IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE PATENT TRIAL AND APPEAL BOARD

CISCO SYSTEMS, INC.
Petitioner,

v.

VIDEO SOLUTIONS PTE., LTD.
Patent Owner.

U.S. Patent No. 8,446,823

Inter Partes Review No.: IPR2024-00922

DECLARATION OF INGRID HSIEH-YEE, PH.D.

Petitioner's Exhibit 1024
Cisco Systems, Inc. v. Video Solutions Pte. Ltd

Table of Contents

I. Introduction 4
A. Qualifications and Professional Experience 4
B. MARC, OCLC and Cataloging Practices 5
C. Scope of This Declaration 12
D. Person of Ordinary Skill in the Art 14
E. Evidence Considered in Forming My Opinions 15
II. Authenticity and Public Availability of *Duffield* (Exhibit 0001) 20
A. Authentication 20
B. Public Availability 23
1. Copyright Registration Record 23
2. Library of Congress Records 24
3. Linda Hall Library Records 27
4. Princeton University Library Records 29
5. Usage Record 31
6. IEEE Metadata Record 31
C. Summary of My Opinion on *Duffield* 32
III. Authenticity and Public Availability of *Cai* (Exhibit 0002) 34
A. Authentication 34
B. Public Availability 35
1. ACM Metadata Records 35
2. Library Records 37
3. Internet Archive Records 41
4. Usage Record 44
C. Summary of My Opinion on *Cai* 45
IV. Authenticity and Public Availability of *Welzl* (Exhibit 0003) 46
A. Authentication 46
B. Public Availability 49
1. Library of Congress Records 49
2. Texas A&M University Records 52
3. Internet Archive Records 54
4. Usage Record 57
C. Summary of My Opinion on *Welzl* 57

V. Conclusion.....59

I. Introduction

I, Ingrid Hsieh-Yee, Ph.D., do hereby declare as follows:

1. I have personal knowledge of all of the matters about which I testify in this declaration.

2. I have been retained as an independent expert witness on behalf of Cisco Systems, Inc. ("CISCO") for an *inter partes* review ("IPR") proceeding before the United States Patent and Trademark Office ("PTO") regarding the authenticity and public availability of certain prior art references in relation to U.S. Patent No. 8,446,823.

3. I am being compensated for my work in this matter at my customary hourly rate. I am also being reimbursed for any reasonable expenses associated with my work and testimony in this investigation. My compensation is not contingent on the results of my study, the substance of my opinions, or the outcome of this matter.

A. Qualifications and Professional Experience

4. My complete qualifications and professional experience are described in my academic curriculum vitae (**Appendix A**). The following is a brief summary of my relevant qualifications and professional experience.

5. I was a professor in the Department of Library and Information Science at the Catholic University of America for 32 years and was conferred the title of Professor Emerita in December 2022. I have experience working in an academic

library, a medical library, and a legislative library. I hold a Ph.D. and a Masters in Library and Information Studies from the University of Wisconsin-Madison.

6. I am an expert on library cataloging and classification and have published two editions on this subject, *Organizing Audiovisual and Electronic Resources for Access*: *A Cataloging Guide* (2000 and 2006). I taught a variety of courses, including Cataloging and Classification, Advanced Cataloging and Classification, Organization of Internet Resources, Organization of Information, Metadata, Digital Content Creation and Management, Internet Searches and Web Design, Information Literacy Instruction, Advanced Information Retrieval and Analysis Strategies, and The Information Professions in Society. I am familiar with metadata schema design and implementation. In my teaching I covered the design and implementation of metadata in databases, search engines, digital repositories, digital libraries, and digital archives. I also covered how information organization affects the discovery and access to digital resources on the Internet. My research interests cover cataloging and classification, information organization, metadata, information retrieval, information architecture, digital collections, scholarly communication, social media, user interaction with information systems, and others.

B. MARC, OCLC and Cataloging Practices

7. I am very familiar with a library cataloging encoding standard known as the "Machine-Readable Cataloging" standard, also known as "MARC," which

became the national standard for sharing bibliographic data in the United States by 1971 and the international standard by 1973. MARC is the primary communications protocol for the transfer and storage of bibliographic metadata in libraries. Experts in my field reasonably rely upon MARC records when forming their opinions.

8. A MARC record consists of several fields, each of which contains specific data about the work. Each field is identified by a standardized, unique, three-digit code corresponding to the type of data that follows. **Appendix B** is a true and correct copy of Parts VII to X of "Understanding MARC Bibliographic: Machine-Readable Cataloging" (http://www.loc.gov/marc/umb/um07to10.html) from the Library of Congress that explains commonly used MARC fields. For example, the personal author of a work is recorded in Field 100, the title is recorded in Field 245, publisher information is recorded in Field 260, the physical volume and characteristics of a publication are recorded in Field 300, and topical subjects are recorded in Fields 650.

9. I am very familiar with library cataloging and processing practices and the relationships between OCLC (Online Computer Library Center) and library catalogs, and how libraries create and share their records to help the public discover and access information.

10. The Online Computer Library Center (OCLC) is the largest bibliographic network of the world, and its WorldCat database is "the world's most

comprehensive database of information about library collections" (source: https://www.oclc.org/en/worldcat.html), containing more than 559 million records from thousands of member institutions (many of which are libraries of some type) in more than 100 countries. OCLC was founded in 1967 to promote and support library cooperation. According to the "Third Article, Amended Articles of Incorporation of OCLC Online Computer Library Center, Inc.," OCLC was created "to establish, maintain and operate a computerized library network and to promote the evolution of library use, of libraries themselves, and of librarianship, and to provide processes and products for the benefit of library users and libraries, including such objectives as increasing availability of library resources to individual library patrons and reducing the rate of rise of library per-unit costs, all for the fundamental public purpose of furthering ease of access to and use of the ever-expanding body of worldwide scientific, literary and educational knowledge and information" (source: https://www.oclc.org/content/dam/oclc/membership/articles-of-incorporation.pdf).

11. OCLC members can contribute original cataloging records in MARC to the system or derive cataloging records from existing records, an activity referred to as "copy cataloging." When an OCLC participating institution acquires a work, it can create an original MARC record for the work in OCLC's Connexion system (a system for catalogers to create and share MARC records), and the system will

automatically generate a code for the date of record creation in the yymmdd format, and the creating library's OCLC symbol is recorded in subfield "a" of the 040 field. Once the MARC record is in Connexion, it becomes available to other OCLC members for adoption to their local online catalogs (*i.e.*, copy cataloging).

12. After a MARC record is created in Connexion, it also becomes searchable and viewable on WorldCat, which is a free web portal for users to explore more than 10,000 libraries worldwide. Records in WorldCat, however, are not presented in MARC fields. Instead, the data elements are labeled to help users interpret the records. Thus, the information stored in MARC records in Connexion is available to the interested public through the user-friendly WorldCat web portal.

13. Libraries create MARC records for works they acquire, including books, serials, motion pictures, and publications in other formats. MARC records are representations of materials held by libraries and are critical for the discovery and access to those materials. Monograph cataloging is fairly common in libraries, and most libraries make a newly cataloged monograph available to the public soon after the cataloging work is completed, usually within a week.

14. The cataloging of serials and the serial check-in process are discussed here to show how libraries usually provide access to newly received serial issues. According to the glossary of the *RDA: Resource Description and Access* cataloging standard, a serial is "a mode of issuance of a manifestation issued in successive parts,

usually bearing numbering, that has no predetermined conclusion. A serial includes a periodical, monographic series, newspaper, etc." Because the publisher of a serial makes new issues of the serial available successively, a customary cataloging practice is to create one bibliographic record for the serial, and the serial record (encoded in MARC) typically provides information on the beginning date and frequency of the serial, not the dates of individual issues. In other words, libraries typically do not create MARC records for individual issues of a serial. Instead, they rely on a serial check-in system to track the receipt of new issues. A common check-in practice is to date stamp a new issue when it arrives. This practice has become automated since the late 1990s, and libraries now vary in how they share the receipt date of a new serial issue with the public. Some libraries use a date stamp, some affix a label to indicate the receipt date, some pencil in the receipt date, and some do not provide the information to the public.

15. The serial check-in process usually takes less than an hour, and one of the steps involves placing a date stamp on the new issue to document the date the issue is checked in. After that, the holdings information of the serial is updated in the library's catalog so that users know which issues are available for request or access. After serial check-in is completed, the new issue is placed on the shelf with the previous issues of the serial. Libraries with a public periodical room typically place new issues in the periodical room for easy user access. Because information

presented in serials often reflects latest discovery, a general practice of libraries is to make new issues of serials available for user access soon after they are checked in, usually within a week.

16. Conference proceedings can be cataloged as monographs or serials. The monograph treatment of conference proceedings means a MARC record is created for a specific conference proceedings, making it easier for user to discover and access the work. The serial treatment of conference proceedings means a newly received conference proceedings is treated as a part of a serial and no MARC record is created for it. The newly received work is put through the serial check-in process, date stamped, and promptly processed for public access. Libraries usually choose the serial treatment of conference proceedings if the proceedings are from conferences with the same name and held on a regular basis (such as annually). The advantage of the serial treatment is that it saves processing time and the newly received work becomes quickly available to the public.

17. I am personally familiar with many online catalogs, databases, and search engines. In preparing for this declaration, I used the following authoritative information systems to search for records:

- ACM Digital Library (https://dl.acm.org/)
- Google Scholar (https://scholar.google.com/)

- IEEE Xplore Digital Library (https://ieeexplore.ieee.org/Xplore/home.jsp)
- Internet Archive (https://archive.org/)
- OCLC Connexion (https://www.oclc.org/en/connexion.html)
- OCLC FirstSearch (https://firstsearch.oclc.org/)
- Online catalog of the Library of Congress (https://catalog.loc.gov/)
- Online catalog of Linda Hall Library (https://catalog.lindahall.org/discovery/search?vid=01LINDAHALL_INST:LHL)
- Online catalog of Princeton University Library (https://catalog.princeton.edu/)
- Online catalog of the Catholic University of America (https://wrlc-cu.primo.exlibrisgroup.com/discovery/search?vid=01WRLC_CAA:01WRLC_CAA&lang=en)
- Online catalog of the library of the Hong Kong University of Science and Technology (https://library.hkust.edu.hk/)
- Online catalog of the library of Texas A&M University (https://catalog.library.tamu.edu/)

- Public catalog of the US Copyright Office (https://cocatalog.loc.gov/cgi-bin/Pwebrecon.cgi?DB=local&PAGE=First)
- WorldCat (https://search.worldcat.org/)

C. Scope of This Declaration

18. I have been asked to offer an opinion on the authenticity and public availability date of the following documents:

(1) Duffield, N. G., Lakshman, T. V., & Stiliadis, D. (1998, March). On adaptive bandwidth sharing with rate guarantees. In *Proceedings. IEEE INFOCOM'98, the Conference on Computer Communications, Seventeenth Annual Joint Conference of the IEEE Computer and Communications Societies, Gateway to the 21st* Century, Vol. 3, pp. 1122-1130, **Ex. 0001** ("*Duffield*");

(2) Cai, Z., Eisenhauer, G., He, Q., Kumar, V., Schwan, K., & Wolf, M. (2004, October). IQ-services: network-aware middleware for interactive large-data applications. In *Proceedings of the 2nd Workshop on Middleware for Grid Computing*, pp. 11-16, **Ex. 0002** ("*Cai*");

(3) Welzl, M. (2005). *Network Congestion Control: Managing Internet Traffic*. John Wiley & Sons, **Ex. 0003** ("*Welzl*").

19. I am rendering my expert opinion on the authenticity of the documents referenced herein and on whether they were published. I will provide my opinion on when and how these documents were publicly disseminated or otherwise made publicly available to the extent that persons interested and ordinarily skilled in the subject matter or art, exercising reasonable diligence, could have located and retrieved these documents.

20. I am informed by counsel that an item is considered authentic if there is sufficient evidence to support a finding that the item is what it is claimed to be. I am also informed that authenticity can be established based on the contents of the documents themselves, such as the appearance, content, substance, internal patterns, protocols followed or other distinctive characteristics of the item.

21. I am informed by counsel that a given reference is "published" if it is publicly accessible upon a satisfactory showing that such a document has been disseminated or otherwise made publicly available to the extent that persons interested and ordinarily skilled in the subject matter or art could locate it exercising reasonable diligence and obtain the document. I have also been informed by counsel that materials available in a library constitute "printed publications" if they are cataloged and indexed according to general library practices and protocols that make the references available and accessible to members of the interested public.

22. I have been informed by counsel and understand that the patent application that eventually issued as the ’823 patent was filed on June 23, 2009, in the United States, and the earliest priority date of the ’823 patent is June 23, 2009. I have assessed public accessibility of the documents referenced below as of this claimed priority date.

D. Person of Ordinary Skill in the Art

23. I am informed by counsel and understand that the subject matter of this IPR proceeding relates to streaming video data over packet networks.

24. I have been informed by counsel and understand that a “person of ordinary skill in the art” (POSITA) is a hypothetical person who is presumed to be familiar with the relevant field and aware of all pertinent prior art at the time of the invention. This hypothetical person is also a person of ordinary creativity, capable of understanding the scientific principles and literature applicable to the pertinent field.

25. I am informed by counsel and understand that a POSITA in the field of the ’823 patent, as of its June 23, 2009 priority date, would have been someone knowledgeable and familiar with the video codec arts, and with the real-time multimedia transport arts, that are pertinent to the ’823 patent. I have also been informed by counsel and understand that this POSITA would have been familiar with the field of video codecs, including the encoding and transmission of video

data, including MPEG and ITU standards that existed as of the late 2000s. I have further been informed by counsel and understand that this POSITA would also have been familiar with the field of real-time multimedia transport, including the protocols for delivering voice and video data over packet networks and associated control protocols, including IETF and ITU standardized protocols that existed as of the late 2000s. Additionally, I have been informed by counsel and understand that this POSITA would have gained knowledge of these concepts through a mixture of training and work experience, such as by having a bachelor's degree in electrical engineering, computer science, physics, or an equivalent degree or equivalent training, and approximately two years of experience working in the field of video codecs and/or in the field of real-time multimedia transport (where lack of professional experience can be remedied by additional education, and vice versa).

26. It is my opinion that such a person would have been actively engaged in learning about the field, possibly through formal instruction and bibliographic resources. By at least June 2009, such a person would have had access to a vast array of print and electronic resources, including at least those resources referenced below.

E. Evidence Considered in Forming My Opinions

27. In the preparation of this declaration, I have reviewed the documents referenced below and any other documents I reference herein, and each of these is a

type of material that experts in my field would reasonably rely upon when forming their opinions:

(1) The documents referenced above in Section I.C;

(2) Library copy of *Duffield*, obtained from Linda Hall Library, **Appendix 0001A**;

(3) Copyright registration record for IEEE Infocom 1998 proceedings, available at https://cocatalog.loc.gov/cgi-bin/Pwebrecon.cgi?v1=3&ti=1,3&Search%5FArg=IEEE%20INFOCOM&Search%5FCode=TALL&CNT=25&PID=D-GZ-oBntCFL5ChNejCgqKit8olTs&SEQ=20240501145351&SID=1 from the public catalog of the US Copyright Office, **Appendix 0001B**;

(4) Bibliographic and MARC serial records for *Proceedings IEEE INFOCOM*, available at https://lccn.loc.gov/86640672 from the online catalog of the Library of Congress, **Appendix 0001C**;

(5) Bibliographic and MARC records for *Proceedings IEEE INFOCOM'98*, available at https://catalog.lindahall.org/permalink/01LINDAHALL_INST/19lda7s/alma992918883405961 from the online catalog of Linda Hall

Library, **Appendix 0001D**;

(6) Bibliographic and MARC records for *IEEE INFOCOM '98*, available at https://catalog.princeton.edu/catalog/9917891843506421 and https://catalog.princeton.edu/catalog/9917891843506421/staff_view from the online catalog of Princeton University Library, **Appendix 0001E**;

(7) Citations to *Duffield*, obtained from Google Scholar, **Appendix 0001F**;

(8) IEEE metadata record for *Duffield*, available at a https://ieeexplore.ieee.org/abstract/document/662924 from the IEEE Xplore Digital Library, **Appendix 0001G**;

(9) ACM metadata record for *Cai*, available at https://dl.acm.org/doi/abs/10.1145/1028493.1028495, and the contents of *MGC '04 Proceedings* at https://dl.acm.org/doi/proceedings/10.1145/1028493 from the ACM Digital Library, **Appendix 0002A**;

(10) OCLC FirstSearch record and Connexion record for the *MGC '04 Proceedings* (subscriptions required for access) and current

bibliographic and MARC records for this work from the Hong Kong University of Science and Technology Library, available at https://lbdiscover.hkust.edu.hk/marc/991012755654003412, **Appendix 0002B**;

(11) Library record for *Cai*, available at https://wrlc-cu.primo.exlibrisgroup.com/permalink/01WRLC_CAA/1lsipa9/cdi_proquest_miscellaneous_31500804, from the online catalog of the Catholic University of America; website of the ACM *International Conference Proceedings Series* (AICPS) at http://www.acm.org/pubs/icp_series.html; a copy of the series at the ACM Digital Library, archived on May 30, 2002, by Internet Archive and available at https://web.archive.org/web/20020530112256/http://www.acm.org/pubs/icp_series.html; and a copy of the "Browse the ACM Proceedings" webpage of the ACM Digital Library archived on April 30, 2006, by Internet Archive and available at https://web.archive.org/web/20060430033751/http://portal.acm.org/browse_dl.cfm?linked=1&part=series&coll=portal&dl=ACM&CFID=74332594&CFTOKEN=809737, **Appendix 0002C**;

(12) Publisher webpage for *Welzl* available at

https://onlinelibrary.wiley.com/doi/book/10.1002/047002531X, and the front matter of *Welzl* at https://onlinelibrary.wiley.com/doi/epdf/10.1002/047002531X.fmatter, **Appendix 0003A**;

(13) Bibliographic and MARC records for *Welzl*, available at https://lccn.loc.gov/2005015429 from the online catalog of the Library of Congress, **Appendix 0003B**;

(14) Bibliographic and MARC records for *Welzl*, available at https://catalog.library.tamu.edu/Record/in00002093568 from the online catalog of the Texas A&M University Library, **Appendix 0003C**;

(15) Wiley webpage for *Welzl*, archived on November 15, 2005, by Internet Archive, available at https://web.archive.org/web/20051115025024/http://www.wiley.com/WileyCDA/WileyTitle/productCd-047002528X.html, **Appendix 0003D**;

(16) Companion website to *Welzl*, archived on December 24, 2005, by Internet Archive, available at https://web.archive.org/web/20051224175649/http://www.welzl.at/

congestion/index.html, **Appendix 0003E**;

(17) Citations to *Welzl*, obtained from Google Scholar, **Appendix 0003F**.

II. Authenticity and Public Availability of *Duffield* (Exhibit 0001)

A. Authentication

28. **Ex. 0001** is a true and correct copy of **"**On adaptive bandwidth sharing with rate guarantees," ("*Duffield*"), by Duffield et al., in *Proceedings, IEEE INFOCOM '98,* vol. 3, pp. 1122-1130, that I obtained from the Library of Congress. When I began researching *Duffield*, I searched WorldCat for records on the conference proceedings containing *Duffield* and the search results informed me that the Library of Congress held these conference proceedings. After verifying the holdings information at the online catalog of the Library of Congress, I obtained a scan copy of *Duffield* from the Library of Congress through a document vendor. This library copy is presented as **Ex. 0001** in this declaration.

29. **Ex. 0001** is a true and correct copy of *Duffield*. Page 1 shows the covers and spine of a work identified as "Proceedings of the IEEE INFOCOM" that has a call number of "TK 5105.5 .I32a Vol 3 1998." Page 2 is the front cover that shows "IEEE INFOCOM '98, The Conference on Computer Communications" is the name of the conference, which is the "Seventeenth Annual Joint Conference of the IEEE Computer and Communications Societies" that has a conference title of "Gateway

to the 21st Century." The cover also shows the conference was sponsored by the IEEE Computer Society and the IEEE Communications Society, and this volume is "Proceedings Volume 3." Page 3 is the title page that shows the same conference information, conference title, and sponsors as the cover. It also shows the conference took place from March 29 to April 2, 1998 in San Francisco, CA, USA. The title page carries a date stamp of "Library of Congress APR 10 1998 Copyright Office." Page 4 is the copyright page that shows "Proceedings IEEE INFOCOM'98 The Conference on Computer Communications" as the title, and this volume has a "1998" copyright date with the Institute of Electrical and Electronics Engineers, Inc. as the copyright holder. It also shows the IEEE catalog numbers for two formats of the conference proceedings, the ISBNs (International Standard Book Numbers) for four formats of the conference proceedings, and the ISSN (International Standard Serial Number) of the IEEE INFOCOM serial. This page also shows a hand-written call number of "TK5105.5 .I32a Vol. 3, 1998." Pages 5 to 20 are the table of contents of the conference proceedings, and page 16 (internal page xxiv) shows *Duffield* begins on page 1122.

30. Pages 21 to 30 are the full text of *Duffield*, and page 21 (internal page 1122) shows the title, authors, keywords ("fair queueing, scheduling, jitter, delay bounds, real-time traffic, traffic management"), and abstract of this article. The footer shows *Duffield* appears in a work whose ISBN is "0-7803-4383-2" and has a

"1998" copyright date with IEEE is the copyright holder. **Ex. 0001** shows *Duffield* has a total of eight pages, including one table, five figures, nine equations, 24 references and an appendix on "Worst-Case Fairness of FQ-SDS" that contains Equations 10 to 35.

31. My research on WorldCat also identified Linda Hall Library as a holding library of *Proceedings, IEEE INFOCOM '98* that contains *Duffield*, so I obtained a copy of this article from the Linda Hall Library to authenticate *Duffield*. The Linda Hall Library copy is presented as **Appendix 0001A** in this declaration.

32. Pages 1 to 9 of **Appendix 0001A** are the full text of *Duffield* obtained from Linda Hall Library. Page 10 is the cover, page 11 is the title page, and page 12 is the copyright page that shows a hand-written call number of "TK5105.5 .I18 1998 v. 3." Page 13 is the "Message from the General Chair" page and the upper left corner shows an annotation of "5-12-98" that indicates this work was added to the collections of the Linda Hall Library on May 12, 1998. The knowledge of this annotation is based on my personal experience with the Linda Hall Library's collections and their cataloging records. Pages 14 to 29 are the table of contents of the conference proceedings.

33. I have closely compared the Linda Hall Library (**Appendix 0001A**) with the Library of Congress copy (**Ex. 0001**) and found them to contain the same full text of *Duffield*. I have also compared the cover, title page, copyright page, and

table of contents of both library copies and found them to have the same content. One difference is that the Library of Congress copy carries a date stamp of "APR 10 1998" while the Linda Hall Library copy shows a "5-12-98" add date. Another difference is their call numbers are slightly different in that Library of Congress uses "I32" as the Cutter number, and Linda Hall Library uses "I18" as the Cutter number. Cutter numbers are typically based on the author, and can vary by collections, so this difference does not raise concerns about the contents of these copies.

34. The Library of Congress copy of *Duffield* (**Ex. 0001**) is in a condition that creates no suspicion about its authenticity. Specifically, this copy is not missing any intermediate pages, the text on each page appears to flow seamlessly from one page to the next, and there are no visible alterations to the document. In addition, the full text of *Duffield* is the same as that in the Linda Hall Library copy, and the front matter of the conference proceedings is the same as that in the Linda Hall Library copy. Moreover, this copy was found in the custody of a national library. I therefore see no reason to question the authenticity of this copy of *Duffield* (**Ex. 0001**).

B. Public Availability

1. Copyright Registration Record

35. **Appendix 0001B** is a true and correct copy of the copyright registration record for *IEEE INFOCOM '98* that I obtained from the public catalog of the US Copyright Office. I personally located, identified, and obtained this record. It is the

type of record experts in my field would reasonably rely upon when forming their opinions.

36. The "Type of Work" field shows this record is a copyright registration record for a "Serial" and the "Issues Registered" field shows the record is about "no. 17, 1998" of the IEEE INFOCOM proceedings that consists of three volumes and was published on "1998-03-29" (*i.e.*, March 29, 1998) and registered with the US Copyright Office on "1998-04-10" (*i.e.*, April 10, 1998.). **Appendix 0001B** informs my opinion that *IEEE INFOCOMM 1998 Proceedings* was published on March 29, 1998, and registered with the Copyright Office on April 10, 1998.

2. Library of Congress Records

37. **Appendix 0001C** is a true and correct copy of the bibliographic and MARC serial records for IEEE INFOCOMM obtained from the online catalog of the Library of Congress. I personally located, identified, and obtained these records. They are the type of record experts in my field would reasonably rely upon when forming their opinions.

38. Pages 1 to 4 are the bibliographic record that shows this work is treated as a "periodical or newspaper" by the Library of Congress. The bibliographic record shows "Proceedings" as the main title, "IEEE INFOCOM" as the meeting name, and the publication has been published by IEEE Computer Society Press since 1982. The "Item Availability" area shows older issues held by the Library of Congress include

proceedings from 1983 to 2001, meaning the 1998 IEEE INFOCOM conference proceedings is held by Library of Congress.

39. Pages 5 to 8 are the MARC serial record. The first six digit of Field 008 show the record was created on "820819" (*i.e.*, August 19, 1982), and the "c19829999" code following the record creation date indicates this publication is a continuing resource that began in 1982 and has no end date of publication yet. Subfield "a" of Field 040 shows "MCA" created the original record and subfield "d" shows "DLC" (*i.e.*, Library of Congress) modified the MCA record to create their copy cataloging record later. Field 022 shows "0743-166X" as the ISSN of this serial. Field 245 shows the title, and Fields 246 show four variant titles are included to help users discover this serial. Field 362 shows the serial began in 1982 and Field 310 shows the publication frequency is "Annual." Field 500 shows some volumes have distinctive titles. Another Field 550 shows the sponsors. Field 710s and Field 711 show the two sponsors and the conference name are included in the MARC record to help users discover this serial.

40. Subjects are represented in Field 050 by a Library of Congress Classification (LCC) number of "TK5105.5" that represents "General works" on "Computer networks" and in Field 082 by a Dewey Decimal Classification number of "621" that represents "Applied physics." Subjects are also represented in Fields 650 by three Library of Congress subject headings with "Computer networks," "Data

transmission systems," and "Telecommunication" as the main topics, each followed by "Congresses" as a form subdivision to indicate the main topics are treated at a conference. These subject representations support users to conduct subject searches to discover this serial.

41. This MARC serial record has made IEEE INFOCOM proceedings discoverable at the Library of Congress. Since the library's holdings began with the 1983 proceedings, it is likely the serial would have become discoverable at this library by 1983. **Ex. 0001** shows *IEEE INFOCOM '98 Proceedings* has a date stamp of April 10, 1998, it is therefore my opinion that this serial would definitely have been discoverable no later than April 1998. Users would have been able to search for this serial by the ISSN, conference meeting name, and sponsors. They would also have been able to conduct subject searches by the classification numbers and subject headings included in this record to discover this serial.

42. As discussed earlier, a customary library practice in organizing serials is to put newly received issues through the serial check-in process, date stamp the issue, and process it promptly for public access, usually within a week after serial check-in is completed. The date stamp on the Library of Congress copy of *Duffield* (**Ex. 0001**) shows "APR 10 1998," so it is my opinion that the conference proceedings containing *Duffield* would have become publicly accessible by April 17, 1998, at the Library of Congress.

3. Linda Hall Library Records

43. My research on the IEEE INFOCOM proceedings also identified Linda Hall Library as a holding library. **Appendix 0001D** is a true and correct copy of the bibliographic and MARC records for this conference proceedings. I personally located, identified, and obtained the records from the online catalog of the Linda Hall Library. They are the type of record experts in my field would reasonably rely upon when forming their opinions.

44. Pages 1 to 4 are the bibliographic record that shows Linda Hall Library cataloged "Proceedings : IEEE INFOCOM'98" as a book, and its call number is "TK5105.5 .I18 1998." Details of the record include the title, three variant titles, creators, subjects, ISBNs, publisher, and IEEE catalog numbers.

45. Page 5 is the MARC record, and the first six digits of Field 008 show the record was created on "980424" (*i.e.*, April 24, 1998). Subfield "a" of Field 040 shows "NSA" created the original record, and subfield "d" shows "LHL" (*i.e.*, Linda Hall Library) modified the record later to create their copy cataloging record. Field 111 shows the conference name is used as the author of this conference proceedings, Field 245 shows the title and conference information, and Fields 246 show three variant titles are included to help users discover this work. Fields 020 show the ISBNs, and Fields 500 show IEEE catalog numbers for this publication. Fields 710 show the two sponsors are included in the MARC record to help users discover this

conference proceedings.

46. Subjects are represented in Field 090 by a Library of Congress Classification (LCC) number of "TK5105.5" that represents "General works" on "Computer networks" and in Fields 650 by two Library of Congress subject headings with "Computer networks" and "Data transmission systems" as the main topics, each followed by "Congresses" as a form subdivision to indicate the main topics are treated at a conference. These subject representations support users to conduct subject searches to discover this book of conference proceedings.

47. The Linda Hall Library MARC record (**Appendix 0001D**) shows that it was based on an NSA MARC record created on April 24, 1998, and the date annotation on the Linda Hall Library copy of *Duffield* (**Appendix 0001A**) shows *Proceedings, IEEE INFOCOM '98* was added to the Linda Hall Library collections on May 12, 1998. These dates inform my opinion that the cataloging of this work would have been completed by May 12, 1998, making the work discoverable at Linda Hall Library from that date on. Users would have been able to search for this work by ISBNs, conference name, title, variant titles, sponsors, and subject. A customary library practice is to process newly catalog monographs promptly for public access, usually within one week after the cataloging work is completed. So the physical copy of *Proceedings, IEEE INFOCOM '98*, after processing, would have been publicly accessible at this library by May 19, 1998.

4. Princeton University Library Records

48. My research on the 1998 *IEEE INFOCOM* proceedings also found that Princeton University Library cataloged this work as a book and created an original record for it, so I obtained their records as further evidence of the public availability of this work in libraries. **Appendix 0001E** is a true and correct copy of the bibliographic and MARC records from the online catalog of Princeton University Library. I personally located, identified, and obtained these records. They are the type of record experts in my field would reasonably rely upon when forming their opinions.

49. Pages 1 and 2 are the bibliographic record that shows the title, author, format, language, publisher, description, call number, subjects, related organizations, variant titles, ISBNs, and the availability of an electronic version of this work.

50. Pages 3 and 4 are the MARC record, and the first six digits of Field 008 show this record was created on "980529" (*i.e.*, May 29, 1998). Subfield "a" of Field 040 shows "PUL" created the MARC record. According to the *Directory of OCLC Members* (https://www.oclc.org/en/contacts/libraries.html), "PUL" is the OCLC library symbol for Princeton University. Data from Fields 008 and 040 inform my opinion that Princeton University Library treated the conference proceedings containing *Duffield* as a book and created a MARC record for it on May 29, 1998,

making this work discoverable in their online catalog from that date on.

51. Linda Hall Library and Princeton University Library cataloged *IEEE INFOCOM '98 Proceedings* as a book, and their records show that they represent the same conference proceedings containing *Duffield*. For this reason, these two records have many data elements in common, including the eight ISBNs (Fields 020), the conference name as the author (Field 111), a variant title starting with "Seventeenth Annual Joint Conference," (Field 246), a note on bibliographical references and indexes (Field 504), two Library of Congress subject headings (Fields 650), and the sponsors (Fields 710). An interesting difference is the Princeton University MARC record records the title as "IEEE INFOCOM '98" while the Linda Hall Library MARC record records the title as "Proceedings : IEEE INFOCOM'98." Both records use the same Library of Congress classification number, but record them in Field 090 and Field 050, respectively. In addition, the Princeton University MARC record includes five more Library of Congress subject headings (Fields 650) and provides more subject search possibilities for users than the Linda Hall Library MARC record.

52. The MARC record of Princeton University Library (**Appendix 0001E**) was completed on May 29, 1998, making *IEEE INFOCOM '98* discoverable in their online catalog from that date on. Users would have been able to search for this work by ISBNs, conference name, title, variant titles, sponsors, Library of Congress

Classification number, and Library of Congress subject headings. A customary library practice is to process newly catalog monographs promptly for public access, usually within one week after the cataloging work is completed. So the physical copy of *IEEE INFOCOM '98*, after processing, would have been publicly accessible at this library by June 5, 1998.

5. Usage Record

53. Actual usage of a publication is reflected by other publications that make reference to it. I used Google Scholar to review citations to *Duffield* and compile a list of early citations to demonstrate usage. **Appendix 0001F** presents six citations to *Duffield* published in 1999, and the earliest citation appeared in April 1999.

6. IEEE Metadata Record

54. My research on *Duffield* discovered an IEEE metadata record for this article at https://ieeexplore.ieee.org/abstract/document/662924 from the IEEE Xplore Digital Library. **Appendix 0001G** is a true and correct copy of this IEEE record. I personally located, identified, and obtained this record. It is the type of record experts in my field would reasonably rely upon when forming their opinion.

55. Page 1 of **Appendix 0001G** shows the title, authors, and abstract of *Duffield*, and indicates the article is published in the proceedings of IEEE INFOCOM '98, which took place from "29 March 1998" to "02 April 1998." The

"Date Added to IEEE Xplore" field shows the metadata record was added to the IEEE Xplore Digital Library on "06 August 2002," making this article discoverable in this system from that date on. The metadata record includes the ISBN of IEEE INFOCOM '98 (as a book) and the ISSN of the IEEE INFOCOM serial, and these data elements enable users to identify libraries that hold the conference proceedings containing *Duffield*. Since users can search IEEE Xplore freely on the Internet, the record add date means *Duffield* became discoverable on the Internet by August 6, 2002. Users at institutions that subscribed to IEEE Xplore would have been able to access the full text of *Duffield* freely, and users whose institutions did not subscribe to IEEE Xplore would have been able to use the ISBN and ISSN to identify library copies of *Proceedings, IEEE INFOCOM '98* that contains *Duffield*.

C. Summary of My Opinion on *Duffield*

56. The Library of Congress copy of *Duffield* (**Ex. 0001**) is authentic because the copy contains no signs to raise suspicion about its authenticity, and the full text of *Duffield* is the same as the copy obtained from Linda Hall Library (**Appendix 0001A**). The copyright registration record for *IEEE INFOCOM 1998 Proceedings* shows this work was published on March 29, 1998 (**Appendix 0001B**). The date stamp on **Ex. 0001** shows the Library of Congress treated *Proceedings of IEEE INFOCOM '98* as part of a serial, and the 1998 issue was received by the Library of Congress on April 10, 1998, meaning the copy, after serial check in,

would have become publicly accessible by April 17, 1998. Linda Hall Library cataloged *Proceedings, IEEE INFOCOM '98* as a book, and their copy of *Duffield* (**Appendix 0001A**) shows this work was added to their collections on May 12, 1998, meaning the cataloging work was completed by that date, making this work discoverable at the library from that date on. Princeton University Library created their own original MARC record for this work on May 29, 1998, making this work discoverable at their library from that date on. MARC records from the Library of Congress, Linda Hall Library and Princeton University Library (**Appendices 0001C**, **0001D**, **0001E**) show that by the end of May 1998 users would have been able to search for the conference proceedings containing *Duffield* by ISBNs, conference name, variant titles, sponsors, Library of Congress classification number, and Library of Congress subject headings. The physical copies, after processing, would have become publicly accessible as early as April 17, 1998 at the Library of Congress, and no later than June 5, 1998 at Princeton University Library. Citation history shows the earliest citation to *Duffield* appeared in April 1999. *Duffield* was also indexed by IEEE and the metadata record was added to the IEEE Xplore Digital Library on August 6, 2002, making this article discoverable on the Internet and publicly accessible no later than that date.

III. Authenticity and Public Availability of Cai (Exhibit 0002)

A. Authentication

57. **Ex. 0002** is a true and correct copy of "IQ-services: network-aware middleware for interactive large-data applications," ("*Cai*"), by Cai et al., in *Proceedings of the 2nd Workshop on Middleware for Grid Computing*, pp. 11-16, that I obtained from the ACM Digital Library. When I began my research on *Cai*, I searched Google Scholar for records, and found a record for this article that led me to the metadata record for *Cai* at https://dl.acm.org/doi/abs/10.1145/1028493.1028495 in the ACM Digital Library. I then followed the PDF link to obtain a copy of *Cai*. This ACM copy is presented as **Ex. 0002** in this declaration.

58. Page 1 of **Ex. 0002** shows the title, authors, abstract, ACM categories and subject descriptors, general terms, and keywords of this article. The bottom of this page shows *Cai* appears in the proceedings of the *2nd Workshop on Middleware for Grid Computing* that took place in Toronto, Canada. The ISBN (International Standard Book Number) of the proceedings is "1-58113-950-0" and *Cai* has a "2004" copyright with the ACM as the copyright holder. **Ex. 0002** shows *Cai* has a total of six pages, including eight figures (some color), and 21 references. The header of odd-number pages shows a heading of "Middleware 2004 Companion" and the

header of even-number pages shows a heading of "Middleware for Grid Computing."

59. This ACM copy of *Cai* (**Ex. 0002**) is in a condition that creates no suspicion about its authenticity. Specifically, this copy is not missing any intermediate pages, the text on each page appears to flow seamlessly from one page to the next, and there are no visible alterations to the document. Moreover, this copy was found in the custody of the publisher, a place where, if authentic, it would likely be found. I therefore see no reason to question the authenticity of this copy of *Cai* (**Ex. 0002**).

B. Public Availability

1. ACM Metadata Records

60. **Appendix 0002A** is a true and correct copy of the ACM metadata record for *Cai* obtained from https://dl.acm.org/doi/abs/10.1145/1028493.1028495. I personally located, identified, and obtained this record. This is the type of record experts in my field would rely upon when forming their opinions.

61. Pages 1 to 7 are the metadata record that shows the title, authors, and the digital object identifier of *Cai*. It also shows *Cai* appears in *MGC '04: Proceedings of the 2nd Workshop on Middleware for Grid Computing* in October 2004, pages 11 to 16, and the article publication date is "18 October 2004." Page 7

shows this metadata record has a 2024 copyright date with ACM, Inc. as the copyright holder.

62. Pages 8 to 25 are the ACM metadata record (https://dl.acm.org/doi/proceedings/10.1145/1028493) for *MGC'04: Proceedings of the 2nd Workshop on Middleware for Grid Computing* that shows the conference chairs, publisher, conference date and location, ISBN of the proceedings, and its publication date "18 October 2004." The record includes an abstract and lists articles in the proceedings. Pages 13 and 14 show an entry for *Cai* that includes a link to this article. Page 25 shows this metadata record has a 2024 copyright date with ACM, Inc. as the copyright holder.

63. The ACM Digital Library was launched in October 1997, and the general public can search the ACM Digital Library for free on the Internet. The metadata records for *Cai* and the proceedings of *MGC '04* show that these records have been discoverable at the ACM Digital Library since 2004. It is therefore my opinion that users would have been able to discover these records and use them to access *Cai* after October 2004 when the article and the proceedings were published. Evidence of user access to the ACM Digital Library and the *MGC '04 Proceedings* will be presented as **Appendix 0002C** and discussed later in the declaration.

2. Library Records

64. The WorldCat database on OCLC has more than 559 million records for publications in various formats. OCLC provides access to this database in three ways. The general public can use WorldCat, a free web portal, to explore the collections in more than 10,000 libraries worldwide. Librarians can use OCLC Connexion to create and share library records for cataloging, collection management, and interlibrary loan services. Researchers and reference librarians can use OCLC FirstSearch to explore many proprietary databases, including the WorldCat database. The difference between WorldCat the free web portal and the WorldCat database on OCLC FirstSearch is that the latter provides more technical details than the former and requires a subscription for access. For example, the "Date of Entry" field on an OCLC FirstSearch record shows the date when the record was entered into the WorldCat database, indicating the date when the record became discoverable on WorldCat the free web portal, on OCLC Connexion, and on OCLC FirstSearch.

65. To determine when the MGC '04 conference proceedings became available at libraries, I searched the WorldCat database on the OCLC FirstSearch system for records. I also searched OCLC Connexion and the online catalog of the library of the Hong Kong University of Science and Technology for records.

66. **Appendix 0002B** is a true and correct copy of three library records for the *2nd Workshop on Middleware for Grid Computing*. Pages 1 and 2 are the OCLC FirstSearch record, page 3 is the OCLC Connexion record, and pages 5 to 7 are the bibliographic and MARC records obtained from the online catalog of the Hong Kong University of Science and Technology Library. I personally located, identified and obtained these records. These are the type of record experts in my field would reasonably rely upon when forming their opinions.

67. Page 1 of the FirstSearch record in **Appendix 0002B** shows an access link that leads to the ACM metadata record for *MGC '04: Proceedings of the 2nd Workshop on Middleware for Grid Computing* and indicates 253 libraries hold this work. Details of the record include the title, authors (conference chairs), corporate author ("ACM Digital Library"), conference author (the conference name as the author of this proceedings), publisher, year of publication, physical description, language, ISBNs, subject descriptors, genre, notes, Library of Congress call numbers, and variant titles. The "Series" field shows this conference proceedings is part of the *ACM International Conference Proceeding Series*. The "date of Entry" field shows the record was entered into the WorldCat database on "20050321" (*i.e.*, March 21, 2005) and its OCLC number is 58539596. Page 3 is the same library record in OCLC Connexion, the technical system for catalogers to share cataloging records. The Connexion record shows an OCLC number of 58539596, indicates 253

libraries hold this work, and the "Entered" field shows the record was entered into WorldCat on "20050321," confirming this Connexion record is the same as the OCLC FirstSearch record. Subfield "a" of Field 040 shows HNK created the original MARC record. According to the *Directory of OCLC Members* (https://www.oclc.org/en/contacts/libraries.html), HNK is the OCLC library symbol for the library of Hong Kong University of Science and Technology. The FirstSearch record and the Connexion record inform me that the library of the Hong Kong University of Science and Technology created a MARC record for *2nd Workshop on Middleware for Grid Computing* on March 21, 2005, and entered it into the WorldCat database, making this work discoverable in Connexion for catalogers and librarians, in OCLC FirstSearch (for subscribers of OCLC FirstSearch), and in WorldCat, the free web portal, for the general public. Since WorldCat the free web portal was launched in January 1998, it is my opinion that after the MARC record was entered into the WorldCat database on March 21, 2005, the general public would have been able to discover the record for *2nd Workshop on Middleware for Grid Computing* and use it to access the *Cai* article. Users would have been able to search for this work by the title, conference name, ISBNs, and Library of Congress subject headings (see the WorldCat record at https://search.worldcat.org/title/58539596).

68. The original MARC record created on March 21, 2005, by the library of the Hong Kong University of Science and Technology was used by more than

250 libraries for copy cataloging later. In March 2013 a new cataloging standard, Resource Description and Access (RDA), was implemented by the Library of Congress, the British Library, and many major libraries (see https://www.librarianshipstudies.com/2017/07/resource-description-and-access-rda.html), and many OCLC records were subsequently upgraded or replaced. The current bibliographic and MARC records for the *2nd Workshop on Middleware for Grid Computing* at the library of the Hong Kong University of Science and Technology show the March 2005 record was replaced by a MARC record created in 2016.

69. As evidence of this record replacement, I have included the 2016 records in **Appendix 0002B**. The bibliographic record on page 5 show the title, publisher, publication year, physical description, corporate and personal authors, other titles, ISBN and language of this work. These details are the same as the details in the OCLC FirstSearch record created and entered into the WorldCat database on March 21, 2005. The "Other Title" field shows "MGC '04 Proceedings of the 2nd Workshop on Middleware for Grid Computing" and "Proceedings of the 2nd Workshop on Middleware for Grid Computing" are included in the record to help users discover *2nd Workshop on Middleware for Grid Computing*. The MARC record (page 7) shows the record was created on "160829" (*i.e.*, August 29, 2016), and the "s2004" code following the record creation date shows this publication has

a single publication year of 2004. Subfield "a" of Field 040 shows PQKB (ProQuest KB) created the new record. I have reviewed the records of several libraries listed as holding libraries of this conference proceedings in the OCLC FirstSearch system and found that in 2024 these libraries also display the 2016 record for *2nd Workshop on Middleware for Grid Computing*. The replacement of the March 2005 MARC record by a 2016 MARC record does not change the fact that in March 2005 a MARC record for *2nd Workshop on Middleware for Grid Computing* was created and entered into the WorldCat database, making the record discoverable from that date on to users of OCLC FirstSearch, OCLC Connexion, and WorldCat, the free web portal.

3. Internet Archive Records

70. **Appendix 0002B** shows that *Cai* appears in *2nd Workshop on Middleware for Grid Computing*, which is also known as *MGC '04 Proceedings of the 2nd Workshop on Middleware for Grid Computing*, and *Proceedings of the 2nd Workshop on Middleware for Grid Computing*. The OCLC FirstSearch record and the Connexion record also show this publication is part of the *ACM International Conference Proceeding Series*. I have obtained a library record for *Cai* from the online catalog of the Catholic University of America, a webpage of the *ACM International Conference Proceedings Series*, and Internet Archive records to show that by April 30, 2006, interested users would have been able to discover *MGC '04*

Proceedings containing *Cai* on the ACM Digital Library. I personally located, identified, and obtained these records. They are the type of record experts in my field would reasonably rely upon when forming their opinions. These records are presented as **Appendix 0002C** in this declaration.

71. Pages 1 to 3 of **Appendix 0002C** are the library record for *Cai* that shows this article is available online in the *Proceedings of the 2nd Workshop on Middleware for Grid Computing*, which is vol. 76 of the *ACM International Conference Proceeding Series*, and the full text is available at the "ACM Digital Library Complete" system. Pages 4 to 12 are the webpage of *International Conference Proceeding Series* (ICPS) of the ACM that I obtained from http://www.acm.org/pubs/icp_series.html. Page 4 explains that this series was established in 2002 to provide "a mechanism to publish the contents of conferences, technical symposia and workshops in the ACM Digital Library" to increase "their visibility among the international computing community." This webpage lists proceedings in ICPS and page 9 of **Appendix 0002C** shows "MGC '04: Proceedings of the 2nd Workshop on Middleware for Grid Computing" is part of the ICPS of the ACM.

72. Page 14 is a screen capture of a Wayback Machine calendar that shows the webpage of "The ACM International Conference Proceedings Series" (http://www.acm.org/pubs/icp_series.html) was archived 20 times between May 30,

2002 and July 9, 2008, and this particular copy was archived on May 30, 2002, and available at https://web.archive.org/web/20020530112256/http://www.acm.org/pubs/icp_series.html. Pages 15 and 16 are the PDF of this archived webpage that explains the goal of this series "is to enable conferences and workshops to cost effectively produce print proceedings for their attendees, while also providing maximum dissemination of the material through electronic channels, specifically, the ACM Digital Library." Bullets 4 and 5 further explain that "ACM does not intend to carry any inventory or to sell hard copies" and that "ACM will put the full-text of the proceedings papers into the ACM Digital Library, under the heading of the 'ACM International Conference Proceedings'". The goal of the series is to ensure "the proceedings will be available to the widest possible audience of computing professionals."

73. Page 18 is a screen capture of a Wayback Machine that shows the "Browse the ACM Proceedings" webpage of the ACM Digital Library in the ACM Portal was archived 49 times between October 8, 2001 and April 22, 2021, and this particular copy was archived on April 30, 2006 and available at https://web.archive.org/web/20060430033751/http://portal.acm.org/browse_dl.cfm?linked=1&part=series&coll=portal&dl=ACM&CFID=74332594&CFTOKEN=809737. Pages 19 to 23 are the webpage saved in PDF. This archived webpage shows that "AICPS: ACM International Conference Proceeding Series" was available to

users by April 30, 2006.

74. Records in **Appendix 0002C** inform my opinion that *Cai* appears in the *Proceedings of the 2nd Workshop on Middleware for Grid Computing*, which is part of the *ACM International Conference Proceeding Series* (AICPS), and that AICPS was established in 2002 to increase the visibility, dissemination, and online availability (through the ACM Digital Library) of the contents of international conferences, technical symposia and workshops. Internet Archive records show AICPS was discoverable as early as May 30, 2002, and no later than April 30, 2006. It is therefore my conclusion that interested users would have been able to access *Cai* in the ACM Digital Library by April 30, 2006. By at least April 30, 2006, users at institutions with a subscription to the ACM Digital Library would have been able to discover and access *Cai* freely. Since the ACM Digital Library is available to the general public for searching, users without an institutional subscription would still have been able to discover the ACM metadata record for *Cai*, then use it to purchase a copy of this article from ACM.

4. Usage Record

75. Actual usage of a publication is reflected by other publications that make reference to it. Citation history on Google Scholar shows *Cai* has been cited 15 times, including eight times between 2005 and 2008. Excluding citations to *Cai* by the co-authors of *Cai*, I have found that the earliest citation to *Cai* is a dissertation

completed in November 2005: Hay, B. (2005, November). Idact transformation manager (Doctoral dissertation, Montana State University). URL = https://www.cs.montana.edu/techreports/2005/Hay.pdf.

C. Summary of My Opinion on *Cai*

76. *Cai* is an online article available at the ACM Digital Library. I have found the ACM copy of *Cai* (**Ex. 0002**) authentic because it is provided by the publisher and the article itself contains no signs that raise suspicion about its authenticity. The ACM metadata record for *Cai* and *MGC '04 Proceedings of the 2nd Workshop on Middleware for Grid Computing* show that these works were published in October 2004 (**Appendix 0002A**), and users would have been able to discover them by authors, title, and subject terms and categories. Library records from OCLC FirstSearch, OCLC Connexion, and the Hong Kong University of Science and Technology (**Appendix 0002B**) show the conference proceedings containing *Cai* was cataloged and the MARC record was entered into the WorldCat database on March 21, 2005, making the proceedings (and *Cai* contained therein) discoverable and accessible from that date on at three OCLC systems. Because of changes in cataloging standard in March 2013, the 2005 MARC record was later replaced by a record created in 2016. The replacement, however, does not change the fact that the conference proceedings became discoverable in March 2005 to users of OCLC FirstSearch, OCLC Connexion, and the free web portal, WorldCat.org.

Library record of *Cai*, webpage of the *ACM International Conference Proceedings Series* (AICPS), and Internet Archive records of AICPS show that *Cai* appears in a conference proceedings that is part of AICPS, and by April 30, 2006, users would have been able to discover *Cai* through AICPS (**Appendix 0002C**). Users at institutions with a subscription to the ACM Digital Library would have been able to access *Cai* for free, and users whose institutions did not have such a subscription would have been able to use the metadata record for *Cai* to purchase the article from the ACM Digital Library. Citation records show the earliest citation to *Cai* appeared in November 2005, further supporting my opinion that *Cai* would have been publicly accessible at the ACM Digital Library as early as October 2004, and no later than April 30, 2006.

IV. Authenticity and Public Availability of *Welzl* (Exhibit 0003)

A. Authentication

77. **Ex. 0003** is a true and correct copy of portions of *Network Congestion Control: Managing Internet* Traffic, *("Welzl"),* by Michael Welzl, published by John Wiley & Sons in 2005, that I obtained from the Library of Congress. When I began researching this work, I searched WorldCat for records and a record at https://search.worldcat.org/title/60512784 identified the Library of Congress as one of the holding libraries. After verifying the holdings information in the online catalog of the Library of Congress, I requested a scan copy of *Welzl* from the Library

of Congress through a document vendor. This library copy is presented as **Ex. 0003** in this declaration.

78. **Ex. 0003** is a true and correct copy of a portion of *Welzl* obtained from the Library of Congress. Page 1 shows the front cover, spine, and back cover of this book. The front cover shows the title, subtitle, author, publisher, and series to which this work belongs. It also shows a call number label of "TK 5105.875.I57 W454 2005 Copy 1." The back cover shows "0-470-02528-X" as the book's ISBN (International Standard Book Number" and indicates the companion website to this book is at www.welzl.at/congestion. Page 3 is the title page that shows the same title and author information as the front cover, and the full name of the publisher. Page 4 is the copyright page that shows a "2005" copyright date and shows a block of Library of Congress Cataloging-in-Publication (CIP) Data that shows "2005015429" as the Library of Congress control number of this record. This page also shows that a British Library Cataloguing in Publication Data is available. Both CIP records show the 13-digit and 10-digit ISBNs of *Welzl*. Pages 5 to 8 are the table of contents that shows *Welzl* includes six chapters, three appendices, a list of abbreviation, a bibliography and an index. Page 9 (internal page xviii) is the last page of the front matter and it shows information on a list of figures. Pages 10 to 14 are internal pages 15, 37, 128, 164, and 165. Page 15 (internal page 263) is the last page of the index.

79. My research on *Welzl* discovered the publisher's webpage for this book at https://onlinelibrary.wiley.com/doi/book/10.1002/047002531X from the Wiley Online Library that offers free access to its front matter. The publisher webpage and the front matter (https://onlinelibrary.wiley.com/doi/epdf/10.1002/047002531X.fmatter) are presented as **Appendix 0003A** in this declaration. Pages 1 to 3 are the publisher's webpage that shows *Welzl* was first published on "28 July 2005" and has a "2005" copyright date with John Wiley & Sons, Ltd. as the copyright holder. The webpage shows the print ISBN, online ISBN, and DOI (digital object identifier) of *Welzl* and also shows users can freely access the front matter, appendices, bibliography, and index of this book. Pages 5 to 17 are the front matter.

80. I have closely compared the front matter from the publisher (**Appendix 0003A**) with the Library of Congress copy (**Ex. 0003**), and found that the front matter pages in the library copy are the same as their corresponding pages in the publisher copy.

81. My research on *Welzl* also found a scan copy of this book on the Internet at http://pws.npru.ac.th/sartthong/data/files/Traffic__Wiley_Series_on_Communications_Networking.pdf. I have closely compared the Library of Congress copy (**Ex. 0003**) with this online copy and found all the pages in **Ex. 0003** to be the same as

their corresponding pages in the online copy.

82. **Ex. 0003** is in a condition that creates no suspicion about its authenticity. Specifically, the pages in **Ex. 0003** are the same as their corresponding pages in the front matter of *Welzl* offered by the publisher and the same as their corresponding pages in the online copy of this book. The text on each page appears to flow seamlessly, and there are no visible alterations to the document. In addition, **Ex. 0003** is in the custody of a national library. It is therefore my opinion that **Ex. 0003** is from a published book that is authentic and publicly available.

B. Public Availability

1. Library of Congress Records

83. **Appendix 0003B** is a true and correct copy of the bibliographic and MARC records for *Welzl* that I obtained from the online catalog of the Library of Congress. I personally located, identified, and obtained these records. They are the type of record experts in my field would reasonably rely upon when forming their opinions.

84. Pages 1 and 2 of the Library of Congress records for *Welzl* (https://lccn.loc.gov/2005015429) show the author, title, publisher, and copyright date of this work. The physical description field shows the book's front matter ends on page xviii, the last numbered page of the book is page 263, and the book contains illustrations. The records also show the book has two ISBNs, and its LCCN (Library

of Congress Control Number) is "2005015429." These bibliographic details match the information presented in **Ex. 0003**, confirming this record represents *Welzl*. The "Item Available" area shows the book's call number is "TK5105.875.I57 W545 2005" and the copy can be requested at the "Jefferson or Adams Building Reading Rooms." The status of this copy is "Not Charged" (meaning available for user request).

85. Page 3 and 4 are the MARC record that shows the date of record creation and revision. The first six digits of Field 008 show the MARC record was created on "050527" (*i.e.*, May 27, 2005). Field 955 records local cataloging activities, and the first date of "2005-05-27" indicates the MARC record was created on that date; a later note showing "2005-10-25 1 copy rec'd., to CIP ver." indicates a physical copy was received and sent to the Cataloging in Publication Program for record verification on October 25, 2005; and the final note of "2005-10-27 z-processor to BCCD" indicates the physical copy was sent to the Binding and Collection Care Division on October 27, 2005. Subfield "a" of Field 040 shows "DLC" created the original record. According to the *Directory of OCLC Members*, (https://www.oclc.org/en/contacts/libraries.html), "DLC" is the OCLC library symbol for the Library of Congress. Fields 008, 955 and 040 inform my opinion that the Library of Congress created a Cataloging-in-Publication MARC record for *Welzl* on May 27, 2005, before the book was published. This CIP record made *Welzl*

discoverable in the online catalog of the Library of Congress from that date on. After the book was published, a copy was sent to the CIP Division for record verification on October 25, 2005. After that step was completed, the physical copy was sent to the processing unit on October 27, 2005.

86. Field 010 of the MARC record (**Appendix 0003B**) shows the Library of Congress Control Number assigned to *Welzl* and Fields 020 shows the book's ISBNs. Field 100 records the author's name, Field 245 shows the title, subtitle and author statement, Field 260 shows the publisher's offices in England and New Jersey, and a copyright date ("c2005") for the work, Field 300 shows the physical description, and Field 504 shows a note about references and index.

87. Subjects are represented in Field 050 by a Library of Congress Classification (LCC) number of "TK5105.875.I57" that represents "Internet" and in Field 082 by a Dewey Decimal Classification (DDC) number of "004.67/8" that represents "Wide-area networks, including Internet; World Wide Web." Subjects are also represented in Fields 650 by two Library of Congress subject headings "Internet" and "Telecommunication |x Traffic |x Management" ("|x" indicates a topical subdivision). These subject representations support users to conduct subject searches to discover this book.

88. This MARC record (**Appendix 0003B**) shows that *Welzl* was published with a 2005 copyright date and became discoverable on May 27, 2005, at the Library

of Congress when the Cataloging-in-Publication record was created. Users would have been able to conduct subject searches by the LCC number, DDC number and the Library of Congress subject headings in the MARC record to discover this book. They would also have been able to search for this book by the LCCN, ISBNs, author and title.

89. Field 955 of the MARC record (**Appendix 0003B**) shows the physical copy was sent to processing on October 27, 2005. A customary library practice is to process a newly cataloged book promptly to make it publicly accessible. In most academic libraries, and the processing of such a book takes no more than a week after the cataloging work is completed. Based on my experience with the Library of Congress collection, it can take one to three months for a newly cataloged book to become publicly accessible. The Library of Congress MARC record informs my opinion that *Welzl* became discoverable in the online catalog of the Library of Congress on May 27, 2005, when the CIP record was created. The physical copy was sent to processing on October 27, 2005, so the physical copy, a processing, would have become publicly accessible between November 27, 2005, and January 27, 2006.

2. Texas A&M University Records

90. As further evidence of the public availability of *Welzl*, I obtained records from the Texas A&M University Library. **Appendix 0003C** is a true and

correct copy of the bibliographic and MARC records for *Welzl* that I obtained from the online catalog of this university. I personally located, identified, and obtained these records. They are the type of record experts in my field would reasonably rely upon when forming their opinions.

91. The first six digits of Field 008 shows the MARC record was created on "050527" (*i.e.*, May 27, 2005), and subfield "a" of Field 040 shows the record was first created by "DLC" (*i.e.*, Library of Congress). Subfield "d" of Field 040 shows "TXA" modified the original DLC later. Field 948 shows the Texas A&M University Library cataloged *Welzl* on "2005/11/4" (*i.e.*, November 4, 2005). Data from Fields 008, 040, and 948 inform my opinion that the Texas A&M University Library used the Library of Congress MARC record to create a copy cataloging record for *Welzl* and completed their record on November 4, 2005, making *Welzl* discoverable in their online catalog from that date on. A customary library practice is to process a newly cataloged book promptly for public access, usually within a week after the cataloging work is completed. It is therefore my opinion that the physical copy of *Welzl*, after processing, would have become publicly available at this library by November 11, 2005.

92. Since the MARC record of the Texas A&M University Library (**Appendix 0003C**) is based on the Library of Congress MARC record (**Appendix 0003B**), these two records have many data in common, including the Library of

Congress control number in Field 010, the 10-digit ISBN in Field 020, author in Field 100, title in Field 245, publisher information in Field 260 and Field 264, physical description in Field 300, a note on bibliographical references and index in Field 504, classification numbers in Field 050 and Field 082, and subject headings in Fields 650. The MARC record of the Texas A&M University Library has made *Welzl* discoverable since November 4, 2005, when the copy cataloging work was completed. Users would have been able to conduct subject searches by the LCC number, DDC number and the Library of Congress subject headings in the MARC record to discover this book. They would also have been able to search for this book by the LCCN, ISBN, author and title.

3. Internet Archive Records

93. To determine when *Welzl* became available for sale, I searched the Internet Archive for publisher records for this book. **Appendix 0003D** is a true and correct copy of the publisher's webpage for *Welzl*. I personally located, identified, and obtained this record. It is the type of record experts in my field would reasonably rely upon when forming their opinions.

94. Page 1 is a screen capture of a Wayback Machine calendar that shows the publisher webpage for *Welzl* (http://www.wiley.com/WileyCDA/WileyTitle/productCd-047002528X.html) was archived 26 times between "27 Apr 2005" and "13 Dec 2017" and this particular

copy was archived on "NOV 15 2005" and is available at https://web.archive.org/web/20051115025024/http://www.wiley.com/WileyCDA/WileyTitle/productCd-047002528X.html. This screen capture shows the title, author and ISBN of *Welzl*, and indicates it was published in "September 2005" and costs "US $120.00" Pages 2 and 3 are the PDF of the archived webpage that shows a description of *Welzl*, a link to the table of contents, and a link to author information. The book description is identical to the information on the back cover of the Library of Congress copy of *Welzl* (**Ex. 0003**).

95. Internet Archive is a digital library that archives scholarly materials, cultural artifacts, and other web resources on the Internet and provides the public with free access to its collections. I have used the Internet Archive in the past and am familiar with their practice of assigning an extended URL to an archived resource in the format of http://www.archive.org/web/[Year in yyyy][Month in mm][Day in dd][Time code in hh:mm:ss]/[Archived URL] to show the point of time the archived file was captured. Because Internet Archive usually does not archive web resources the moment they appear on the Internet, and a web resource must be available first before it can be archived, the archive date of a file captured by the Internet Archive means that the archived file was available before, and definitely no later than, the archive date. **Appendix 0003D** shows the publisher webpage for *Welzl* was archived on "20051115025024" (*i.e.*, November 15, 2005, at 2 a.m., 50 minutes, 24 seconds).

This archive date means the webpage would have been publicly discoverable on the Internet no later than November 15, 2005, and users would have been able to follow the links on the webpage to purchase a copy of *Welzl*.

96. **Ex. 0003** shows the book has a companion website, so I searched Internet Archive for records of this website to determine when it became publicly accessible. **Appendix 0003E** is a true and correct copy of the companion website that I obtained from Internet Archive. I personally located, identified, and obtained this record. This is the type of record experts in my field would reasonably rely upon when forming their opinions.

97. Page 1 is a screen capture of a Wayback Machine calendar that shows the companion website to *Network Congestion Control: Managing Internet Traffic* (http://www.welzl.at/congestion/index.html) was archived twice between "24 Dec 2005" and "6 Jan 2006" and this particular copy was archived on "DEC 24 2005" and is available at https://web.archive.org/web/20051224175649/http://www.welzl.at/congestion/index.html. Pages 2 to 4 are the PDF of this website, and page 2 instructs users to "visit the Wiley or Amazon pages for details about the book itself." The end of the website on page 4 shows this website was last updated on "25. July 2005."

98. **Appendix 0003E** shows the webpage accompanying the *Welzl* book was first archived on "20051224175649" (*i.e.*, December 24, 2005, at 5 p.m., 56

minutes, 49 seconds), and the website was last updated on July 25, 2005. These dates indicate that the website would have been discoverable by July 25, 2005, when the website was last updated, and definitely no later than December 24, 2005, when the website was first archived by Internet Archive. It is therefore my opinion that this companion website to *Welzl* that contains information for users to visit the publisher's website and Amazon to learn more about *Welzl* was definitely publicly accessible on the Internet no later than December 24, 2005.

4. Usage Record

99. Actual usage of a publication is reflected by other publications that make reference to it. I used Google Scholar to review citations to *Welzl* and compile a list of early citations to demonstrate usage. **Appendix 0003F** presents six citations to *Welzl* published in 2006 and 2007, and the earliest citation appeared in December 2006.

C. Summary of My Opinion on *Welzl*

100. My analysis of the copy of *Welzl* from the Library of Congress (**Ex. 0003**), the front matter from the publisher (**Appendix 0003A**), and an online copy of *Welzl* informs my opinion that the content of **Ex. 0003** is authentic and is from a published book that is publicly accessible. The Library of Congress MARC record (**Appendix 0003B**) shows a CIP record was created on May 27, 2005, the CIP record was verified in October 2005, and the physical copy was sent to processing on

October 27, 2005. It is therefore my opinion that *Welzl* was discoverable at the online catalog of the Library of Congress from May 27, 2005 on, and the physical copy, after processing, would have become publicly accessible by November 3, 2005. The MARC record from the Texas A&M University Library (**Appendix 0003C**) shows the library completed on November 4, 2005, their copy cataloging record that was based on the Library of Congress MARC record. **Appendix 0003C** made *Welzl* discoverable in their online catalog from November 4, 2005 on, and the physical copy, after processing, would have become publicly accessible by November 11, 2005. The publisher's webpage for *Welzl* archived on November 15, 2005 (**Appendix 0003D**), made *Welzl* discoverable on the Internet by that date. The companion website to *Welzl* archived on December 24, 2005 (**Appendix 0003E**) shows users would have been directed to the publisher's website for *Welzl* and Amazon to know the details of *Welzl* and purchase the book if interested. Based on my review and analysis of these documents and records, it is my opinion that *Welzl* was discoverable on the Internet by November 15, 2005, and the book was cataloged and became publicly accessible by November 11, 2005, at the Texas A&M University Library, and also became publicly accessible by January 27, 2006, at the Library of Congress. Citation records show early usage of *Welzl* took place in 2006 and 2007.

V. Conclusion

101. My findings and opinions set forth in this Declaration are based on my work and examination to date. I may continue my examinations in view of additional documents or other factual evidence over the course of this proceeding that may necessitate supplementing and/or refining my opinions. I reserve the right to add to, alter, or delete my opinions and my Declaration upon discovery of additional information. I reserve the right to make such changes as necessary.

102. In signing this Declaration, I understand that the Declaration will be filed as evidence in a contested case before the Patent Trial and Appeal Board of the United States Patent and Trademark Office. I acknowledge that I may be subject to cross-examination in this case and that cross-examination will take place within the United States. If cross-examination is required of me, I will appear for cross-examination within the United States during the time allotted for cross-examination.

103. I hereby declare that all statements made herein on my own knowledge are true and that all statements made on information and belief are believed to be true, and further, that these statements were made with the knowledge that willful false statements and the like so made are made under penalty of perjury.

Date: 5-20-2024 Executed: Ingrid Hsieh-Yee

Ingrid Hsieh-Yee, Ph.D.

EXHIBIT 0001

PROCEEDINGS
OF THE
IEEE INFOCOM

TK 5105
.5
.I32a
Vol 3
1998

IEEE INFOCOM '98

The Conference on Computer Communications

Proceedings Volume 3

Seventeenth Annual Joint Conference of the IEEE Computer and Communications Societies

Gateway to the 21st Century

Sponsored by
IEEE Computer Society
IEEE Communications Society

Tutorials
March 29 to 30, 1998

Technical Sessions
March 31 to April 2, 1998

Conference Committee
General Chair: Roch Guerin, IBM T.J. Watson Research Center
Technical Program Chair: Ian F. Akyildiz, Georgia Institute of Technology

Hotel Nikko
San Francisco, CA, USA

IEEE

IEEE COMMUNICATIONS SOCIETY

Proceedings

IEEE INFOCOM '98

The Conference on Computer Communications

Volume 3

LIBRARY OF CONGRESS
APR 10 1998
COPYRIGHT OFFICE
R

Seventeenth Annual Joint Conference of the IEEE Computer and Communications Societies

Gateway to the 21st Century

---

29 March - 2 April 1998
Hotel Nikko, San Francisco, CA, USA

INFOCOM'98 gratefully acknowledges the generous travel support provided by the National Science Foundation.

Sponsored by
IEEE Computer Society
IEEE Communications Society

IEEE

IEEE COMMUNICATIONS SOCIETY

Proceedings IEEE INFOCOM'98 The Conference on Computer Communications

Abstracting is permitted with credit to the source. Libraries are permitted to photocopy beyond the limit of U.S. copyright law, for private use of patrons, those articles in this volume that carry a code at the bottom of the first page, provided the per-copy fee indicated in the code is paid through the Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923. For other copying, reprint, or republication permission, write to the IEEE Copyrights Manager, IEEE Operations Center, 445 Hoes Lane, Piscataway, NJ 08855-1331. All rights reserved. Copyright © 1998 by The Institute of Electrical and Electronics Engineers, Inc.

TK5105
.5
.I32a
Vol. 3, 1998

IEEE Catalog Number: 98CH36169 (softbound)
98CB36169 (casebound)

ISBN Softbound: 0-7803-4383-2
Casebound: 0-7803-4384-0
Microfiche: 0-7803-4385-9
CD-Rom: 0-7803-4386-7

ISSN 0743-166X

Additional copies of this publication are available from

IEEE Operations Center
P. O. Box 1331
445 Hoes Lane
Piscataway, NJ 08855-1331 USA

1-800-678-IEEE
1-908-981-1393
1-908-981-9667 (FAX)
833-233 (Telex)
email: customer.services@ieee.org

Table of Contents

Volume 1 Tuesday

Session 1A: Quality of Service Routing
Session Chair: Martha Steenstrup, BBN/GTE

QoS Routing in Networks with Uncertain Parameters 3
D. H. Lorenz and A. Orda

QoS Routing Via Multiple Paths Using Bandwidth Reservation 11
N. S. V. Rao and S. G. Batsell

Bounds on End-to-End Performance via Greedy, Multi-Path Routing in Integrated Services Networks 19
S. Mithal

Routing with End to End QoS Guarantees in Broadband Networks 27
A. Orda

Session 1B: Channel Allocation in Wireless Networks
Session Chair: Tony Ephremides, University of Maryland

Effective Bandwidths in Wireless Networks with Multiuser Receivers 35
D. Tse and S. Hanly

Teletraffic Issues Related to Channel Allocation in Digital Mobile Cellular Networks 43
P. L. Hiew and M. Zukerman

Performance Bounds for Dynamic Channel Assignment Schemes Operating under Varying Re-Use Constraints 51
P. Whiting and S. Borst

Channel Assignment Algorithms Satisfying Co-Channel and Adjacent Channel Reuse Constraints in Cellular Mobile Networks 59
S. Sarkar and K. N. Sivarajan

Session 1C: WDM Networks
Session Chair: Bo Li, Hong Kong University of Science and Technology

Cost Effective Traffic Grooming in WDM Rings 69
O. Gerstel, R. Ramaswami, and G. Sasaki

Dynamic Load Balancing in Broadcast WDM Networks with Tuning Latencies 78
I. Balldine and G. N. Rouskas

Unscheduled Multicasts in WDM Broadcast-and-Select Networks 86
E. Modiano

Wavelength Assignment in a WDM Ring: To Minimize Cost of Embedded SONET Rings 94
O. Gerstel, P. Lin, and G. Sasaki

Session 1D: Switching I
Session Chair: Mark Karol, Bell Labs, Lucent Technologies

Minimum Distance Routing in the Bidirectional Shufflenet 102
M. Gerla, P. Palnati, E. Leonardi, and F. Neri

The Multi-Dimensional Shuffle-Exchange Network: A Novel Topology for Regular Network Architectures 110
P. P. To and T. T. Lee

Strictly and Wide-Sense Nonblocking Photonic Switching Under Crosstalk Constraint 118
M. Vaez and C.-T. Lea

Strict-Sense Non-Blocking Conditions for Shuffle/Exchange Networks with Vertical Replication 126
I. Busi and A. Pattavina

Session 2A: Wireless LANs
Session Chair: Vaduvur Bharghavan, University of Illinois Urbana-Champaign

Stability and Performance Analysis of HIPERLAN 134
G. Anastasi, L. Lenzini, and E. Mingozzi

IEEE 802.11 Wireless LAN: Capacity Analysis and Protocol Enhancement 142
F. Calì, M. Conti, and E. Gregori

A Comparison of MAC Protocols for Wireless Local Networks Based on Battery Power Consumption 150
J.-C. Chen, K. M. Sivalingam, P. Agrawal, and S. Kishore

A Near-Optimum Channel Access Protocol Based on Incremental Collision Resolution and Distributed Transmission Queues 158
R. Garcés and J.J. Garcia-Luna-Aceves

Session 2B: Rate Based Congestion Control
Session Chair: Ramesh Nagarajan, Bell Labs, Lucent Technologies

Robust Rate Control for ABR Sources 166
E. Altman, T. Basar, and R. Srikant

LAPLUS: An Efficient, Effective and Stable Switch Algorithm for Flow Control of the Available Bit Rate ATM Service 174
S. Prasad, K. Kiasaleh, and P. Balsara

A Control-Theoretic ABR Explicit Rate Algorithm for ATM Switches with Per-VC Queueing 183
L. Benmohamed and Y. T. Wang

Zero Queueing Flow Control and Applications 192
H. T. Kung and S. Y. Wang

Session 2C: Video Modeling and Coding
Session Chair: Jorg Liebeherr, Polytechnic University, Brooklyn

Modeling Video Traffic in the Wavelet Domain 201
S. Ma and C. Ji

Threshold Autoregressive Models for VBR MPEG Video Traces 209
B. S. Jang and C. Thomson

An Accurate and Treatable Markov Model of MPEG-Video Traffic 217
A. Lombardo, G. Morabito, and G. Schembra

Robust H.263 Video Coding for Transmission over the Internet 225
M. H. Willebeek-LeMair, Z.-Y. Shae, and Y.-C. Chang

Session 2D: End-to-End Protocol Performance
Session Chair: Kang Shin, University of Michigan, Ann Arbor

Timer Reconsideration for Enhanced RTP Scalability 233
J. Rosenberg and H. Schulzrinne

Explicit Window Adaptation: A Method to Enhance TCP Performance 242
L. Kalampoukas, A. Varma, and K.K. Ramakrishnan

TCP Behavior of a Busy Internet Server: Analysis and Improvements 252
H. Balakrishnan, V. N. Padmanabhan, S. Seshan, M. Stemm, and R. H. Katz

TCP Fast Recovery Strategies: Analysis and Improvements 263
D. Lin and H.T. Kung

Session 3A: QoS Based Scheduling
Session Chair: James Roberts, France Telecom

Implementing Fair Queueing in ATM Switches: The Discrete-Rate Approach 272
F. M. Chiussi and A. Francini

Implementing Distributed Packet Fair Queueing in a Scalable Switch Architecture 282
D. C. Stephens and H. Zhang

Decoupled Generalized Processor Sharing: A Fair Queueing Principle for Adaptive Multimedia Applications 291
F. Toutain

Design Considerations for Supporting TCP with Per-Flow Queueing 299
B. Suter, T.V. Lakshman, D. Stiliadis, and A. Choudhury

Session 3B: Wireless Multiple Access
Session Chair: Jim Freebersyser, Army Research Office

Remote-Queueing Multiple Access (RQMA): Providing Quality of Service for Wireless Communication 307
N. R. Figueira and J. Pasquale

D^2MA: A Distributed Access Protocol for Wireless ATM Networks 315
M. Listanti, F. Mascitelli, and A. Mobilia

A Five-Phase Reservation Protocol (FPRP) for Mobile Ad Hoc Networks 322
C. Zhu and M. S. Corson

Judicious Use of Redundant Transmissions Multi-Channel ALOHA Networks with Deadlines 332
Y. Birk and Y. Keren

Session 3C: Performance Analysis I
Session Chair: Udo Krieger, Deutsche Telekom

Cell Loss Asymptotics in Buffer Fed with a Large Number of Independent Stationary Sources 339
N. Likhanov and R. R. Mazumdar

Packet Multiplexers with Adversarial Regulated Traffic 347
S. Rajagopal, M. Reisslein, and K. W. Ross

Computing Queue-Length Distributions for Power-Law Queues 356
M. Roughan, D. Veitch, and M. Rumsewicz

New Bounds and Approximations Using Extreme Value Theory for the Queue Length Distribution in High-Speed Networks 364
J. Choe and N. B. Shroff

Session 3D: IP over ATM Networks
Session Chair: Josep Sole-Pareta, University Politecnico de Cataluna

Performance Issues in VC-Merge Capable Switches for IP over ATM Networks 372
I. Widjaja and A. I. Elwalid

Adaptive Resource Management for Flow-Based IP/ATM Hybrid Switching Systems 381
H. Che, S.-Q. Li, and A. Lin

Flow-Aware Gateway Support for IP-over-ATM 390
K. Lee and A. Fisher

A Simulation Study of Packet Forwarding Methods over ATM: SBR Evaluation..... 401
F. Hoymany and D. Mossé

Panel: Evolution of Internet and Telecommunications
Organizer: Allison Mankin, USC/ISI-East

Panelists from ISP, telco, and Internet development communities discuss the evolution and growth of the communications fabric:

Services (data, phone, etc.)
Reliability
Growth
Who will be the providers?
What technologies will effect the most changes?

Volume 2 Wednesday

Session 4A: Network Security, Management and Pricing
Session Chair: Irene Katzela, University of Toronto

Compact and Unforgeable Key Establishment over an ATM Network..... 411
Y. Zheng and H. Imai

Designing a Distributed Authorization Service..... 419
T. Y. C. Woo and S. S. Lam

CMIS/P++: Extensions to CMIS/P for Increased Expressiveness and Efficiency in the Manipulation of Management Information 430
G. Pavlov, A. Liotta, P. Abbi, and S. Ceri

Incentive-Compatible Pricing Strategies in Noncooperative Networks..... 439
Y. A. Korilis, T. A. Varvarigou, and S. R. Ahuja

Session 4B: Protocol Verification and Testing
Session Chair: Hasan Ural, University of Ottawa

An Extension to Concurrent TTCN..... 447
M. Xu and J. Wu

Promela++: A Language for Constructing Correct and Efficient Protocols 455
A. Basu, G. Morrisett, and T. von Eicken

User-Level Protocol Servers with Kernel-Level Performance 463
J. C. Brustoloni and P. Steenkiste

Towards Formal Semantics for QoS Support 472
J.-P. Richter and H. de Meer

Session 4C: Internet
Session Chair: Jon Crowcroft, University College of London

Adding Voice to a Distributed Game on the Internet 480
J. Bolot and S. F. Parisis

Internet Telephony Gateway Location 488
J. Rosenberg and H. Schulzrinne

Measurement and Analysis of Long-Range Dependent Behavior of Internet Packet Delay 497
M. S. Borella and G. B. Brewster

Simulation of FEC-Based Error Control for Packet Audio on the Internet 505
M. Podolsky, C. Romer, and S. McCanne

Session 4D: Performance Analysis II
Session Chair: Tulin Atmaca, Int-Evry

Performance Analysis of Feedback Controlled Data Packet Transmission over High-Speed Networks 516
Y. Kim and S. Li

The Cost of Qualtity in Networks of Aggregate Traffic 525
N. G. Duffield and S. H. Low

Linear Complexity Algorithms for Maximum Throughput in Radio Networks and Input Queued Switches 533
L. Tassiulas

Transport Layer Adaptable Rate Control (TARC): Analysis of a Transport Layer Flow Control Mechanism for High Speed Networks 540
R. Y. C. Brockett and I. Rubin

Session 5A: Adaptive Wireless Protocols
Session Chair: Aubrey Bush, NSF

Adaptive Use of Error-Correcting Codes for Real-Time Communication in Wireless Networks 548
M. Elaoud and P. Ramanathan

Adaptive Resource Management for DS-CDMA Networks Subject to Energy Constraints 556
S.-J. Oh and K. M. Wasserman

Adaptive Frame Length Control for Improving Wireless Link Throughput, Range and Energy Efficiency 564
P. Lettieri and M. B. Srivastava

Controlled Multimedia Wireless Link Sharing via Enhanced Class-Based Queueing with Channel-State-Dependent Packet Scheduling 572
C. Fragouli, V. Sivaraman, and M. Srivastava

Session 5B: Active Networks
Session Chair: Ken Calvert, Georgia Tech

Active Reliable Multicast 581
L.-W. Lehman, S. Garland, and D. L. Tennenhouse

Improving the Performance of Distributed Applications Using Active Networks 590
U. Legedza, D. Wetherall, and J. Guttag

Self-Organizing Wide-Area Network Caches 600
S. Bhattacharjee, K. Calvert, and E. W. Zegura

DAN: Distributed Code Caching for Active Networks 609
D. Decasper and B. Plattner

Session 5C: QoS and Resource Allocation in Integrated Services Networks
Session Chair: Geoffrey Xie, Naval Postgraduate School

Bandwidth Allocation for Guaranteed versus Best Effort Service Categories 617
E. Altman, A. Orda, and N. Shimkin

SCED+: Efficient Management of Quality of Service Guarantees 625
R. L. Cruz

Enforceable Quality of Service Guarantees for Bursty Traffic Streams 635
E. W. Knightly

On Statistical Multiplexing, Traffic Mixes, and VP Management 643
C.-F. Su and G. de Veciana

Session 5D: Video Communication
Session Chair: Giovanni Pacifici, IBM T.J. Watson Research Center

Scheduling for Large-Scale On-Demand Data Broadcasting 651
D. Aksoy and M. Franklin

A Network-Conscious Approach to End-to-End Video Delivery over Wide Area Networks Using Proxy Servers 660
Y. Wang, Z.-L. Zhang, D. H. C. Du, and D. Su

Some Proposals to Improve Error Resilience in the MPEG-2 Video Transmission over ATM Networks 668
P. Cuenca, A. Garrido, F. Quiles, and L. Orozco-Barbosa

A General Optimal Video Smoothing Algorithm 676
Z. Jiang and L. Kleinrock

Session 6A: Scalable Routing and Multi-Casting
Session Chair: Bulent Yener, New Jersey Institute of Technology

Hierarchical Source Routing through Clouds 685
M. Montgomery and G. de Veciana

Distributed Top-Down Hierarchy Construction 693
D. G. Thaler and C. V. Ravishanka

Hierarchical Routing Using Link Vectors 702
J. Behrens and J.J. Garcia-Luna-Aceves

Forwarding State Reduction for Sparse Mode Multicast Communication 711
J. Tian and G. Neufeld

Session 6B: Mobility Management
Session Chair: Christopher Rose, Rutgers University

Minimizing the Wireless Cost of Tracking Mobile Users: An Adaptive Threshold Scheme 720
Z. Naor and H. Levy

Efficient PCS Call Setup Protocols 728
Y. Cui, D. Lam, J. Widom, and D. C. Cox

Update and Search Algorithms for Wireless Two-Way Messaging: Design and Performance 737
T. Woo, T. F. La Porta, J. Golestani, and N. Agarwal

Architectures and Handoff Schemes for CATV-Based Personal Communications Network 748
N.-F. Huang, C.-A. Su, and H.-C. Chao

Session 6C: End Systems
Session Chair: Simon Lam, University of Texas - Austin

The END: A Network Adapter Design Tool 756
A. Indiresan, A. Mehra, and K. G. Shin

Implementing Protocols in Java: The Price of Portability 765
B. Krupczak, M. Ammar, and K. Calvert

A Comparative Study of High Speed Networks 774
J. Hall, R. Sabatino, S. Crosby, I. Leslie, and R. Black

A Novel Server Selection Technique for Improving the Response Time of a Replicated Service 783
Z. Fei, S. Bhattacharjee, E. W. Zegura, and M. Ammar

Session 6D: Panel - Broadband Wireless Networks
Organizer: Suresh Singh, University of South Carolina

Discuss services to be provided, possible architectures and support for Mobility and QoS in broadband wireless networks. Additional issues include: extending connectivity to users who roam between heterogenous wireless networks (wireless ATM to cellular to satellite); can QoS guarantees be translated (adapted) to changing environments.

Session 7A: Switching II
Session Chair: Murat Azizoglu, University of Washington

A Practical Scheduling Algorithm to Achieve 100% Throughput in Input-Queued Switches 792
A. Mekkittikul and N. McKeown

A Multicast Single-Queue Switch with a Novel Copy Mechanism 800
M. Hashemi and A. Leon-Garcia

Doubling Memory Bandwidth for Network Buffers 808
Y. Joo and N. McKeown

Multiconfiguration Multihop Protocols (MMPs): A New Class of Protocols for Packet-Switched WDM Optical Networks 816
J. P. Jue and B. Mukherjee

Session 7B: ATM Networks
Session Chair: Ioanis Nikolaidis, University of Alberta

OPENET: An Open and Efficient Control Platform for ATM Networks 824
I. Cidon, T. Hsiao, A. Khamisy, A. Parekh, R. Rom, and M. Sidi

UNITE - An Architecture for Lightweight Signalling in ATM Networks 832
G. Hjalmtysson and K.K. Ramakrishnan

Intelligent Voice Smoother for VBR Voice over ATM Networks 841
P. L. Tien and M. C. Yuang

An Adaptive Connection Admission Control Policy for VBR+ Service Class 849
S. Mukherjee, D. Reininger, and B. Sengupta

Session 7C: Resource Allocation and Network Protocols
Session Chair: Joe Touch, USC/ISI

Equilibrium Allocation of Variable Resources for Elastic Traffics 858
S. H. Low

Schemes for Scheduling of Control Messages by Hierarchical Protocols 865
E. Bortnikov and R. Cohen

Admission Control for Booking Ahead Shared Resources 873
A. Greenberg and D. Wischik

Exploring the Performance Impact of QoS Support in TCP/IP Protocol Stacks 883
R. Engel, D. Kandlur, A. Mehra, and D. Saha

Session 7D: Optical Networks
Session Chair: Achille Pattavina, Politecnico Milano

Limited Wavelength Translation in All-Optical WDM Mesh Networks 893
V. Sharma and E. A. Varvarigos

On the Optimal Placement of Wavelength Converters in Wavelength-Routed Networks 902
S. Subramaniam, M. Azizoglu, and A. K. Somani

Provisioning Algorithms for WDM Optical Networks 910
M. Alanyali and E. Ayanoglu

Design of Logical Topologies: A Linear Formulation for Wavelength Routed Optical Networks with No Wavelength Changers 919
R. M. Krishnaswamy and K. N. Sivarajan

Volume 3 Thursday

Session 8A: ATM Switching
Session Chair: Chin-Tau Lea, Hong Kong University of Science and Technology

Performance of ATM Switches with Age Priority Packet Discarding under the ON-OFF Source Model 931
H.-B. Chiou and Z. Tsai

A Delay and Loss Versatile Scheduling Discipline in ATM Switches 939
J. M. Hah and M. C. Yuang

Impact Analysis of Packet-Level Scheduling on an ATM Shared-Memory Switch 947
C. Fulton, S. Li, and A. Lin

Dynamic Flow Switching, A New Communication Service for ATM Networks 955
Q. Bian, K. Shiomoto, and J. Turner

Session 8B: Reliable Multi-Casting
Session Chair: C.S. Raghavendra, The Aerospace Corporation

Optimal Multicast Feedback 964
J. Nonnenmacher and E. W. Biersack

How Bad is Reliable Multicast without Local Recovery? 972
J. Nonnenmacher, M. Lacher, M. Jung, E. Biersack, and G. Carle

Optimal Packing of Group Multi-Casting 980
S. Chen, O. Gunluk, and B. Yener

A Comparison of Server-Based and Receiver-Based Local Recovery Approaches for Scalable Reliable Multicast 988
S. K. Kasera, J. Kurose, and D. Towsley

Session 8C: Congestion Control
Session Chair: Anna Charny, Cabletron and MIT

TCP-Like Congestion Control for Layered Multicast Data Transfer 996
L. Vicisano, L. Rizzo, and J. Crowcroft

Feedback Consolidation Algorithms for ABR Point-to-Multipoint Connections in ATM Networks 1004
S. Fahmy, R. Jain, R. Goyal, B. Vandalore, S. Kalyanaraman, S. Kota, and P. Samudra

A Load Profiling Approach to Routing Guaranteed Bandwidth Flows 1014
I. Matta and A. Bestavros

Multi-Service Connection Admission Control Using Modular Neural Networks 1022
C.-K. Tham and W.-S. Soh

Session 8D: Wireless Networks
Session Chair: Andrew Campbell, Columbia University

Software Radios for Wireless Networking 1030
V. Bose, A. B. Shah, and M. Ismert

MSOCKS: An Architecture for Transport Layer Mobility 1037
D. A. Maltz and P. Bhagwat

Mobile-End Transport Protocol: An Alternative to TCP/IP Over Wireless Links 1046
K.-Y. Wang and S. K. Tripathi

Demand-based Radio Network Planning of Cellular Mobile Communication Systems 1054
K. Tutschku

Session 9A: Video II
Session Chair: T. V. Lakshman, Bell Labs, Lucent Technologies

Layered Video Multicast with Retransmissions (LVMR): Evaluation of Hierarchical Rate Control 1062
X. Li, S. Paul, and M. Ammar

Adaptive Multicast of Multi-Layered Video: Rate-Based and Credit-Based Approaches 1073
B. Vickers, C. Albuquerque, and T. Suda

End-to-End Delay of Videoconferencing over Packet Switched Networks 1084
M. Baldi and Y. Ofek

SAVE: An Algorithm for Smoothed Adaptive Video over Explicit Rate Networks 1093
N.G. Duffield, K.K. Ramakrishnan, and A. R. Reibman

Session 9B: Scheduling Support for Adaptive Bandwidth Sharing
Session Chair: Edward Knightly, Rice University

Packet Fair Queueing Algorithms for Wireless Networks with Location-Dependent Errors 1103
T. S. E. Ng, I. Stoica, and H. Zhang

Minimum Delay Self-Clocked Fair Queueing for Packet-Switched Networks 1112
F. Chiussi and A. Francini

On Adaptive Bandwidth Sharing with Rate Guarantees 1122
N. G. Duffield, T. V. Lakshman, and D. Stiliadis

Optimal Bandwidth/Delay Tradeoff for Feasible- Region-Based Scalable Multimedia Scheduling 1131
W. Zhao, T. Seth, M. Kim, and M. Willebeek-LeMair

Session 9C: Wireless Protocols
Session Chair: C.K. Toh, Hughes Systems

Handover Re-routing Schemes for Connection Oriented Services in Mobile ATM Networks 1139
B. A. J. Banh, G. J. Anido, and E. Dutkiewicz

User Agent Migration Policies in Multimedia Wireless Networks 1147
R. Ramjee, T. La Porta, J. Kurose, and D. Towsley

Data Transfer Scheme for Wireless Real-Time Communications 1156
C.-S. Wu and G.-K. Ma

A Generalized Processor Sharing Approach to Time Scheduling in Hybrid CDMA/TDMA 1164
M. A. Arad and A. Leon-Garcia

Session 9D: Reliable Multicast Protocols
Session Chair: J.J. Garcia-Luna Aceves, University of California – Santa Cruz

Efficient Rate-Controlled Bulk Data Transfer Using Multiple Multicast Groups 1172
S. Bhattacharyya, J. F. Kurose, D. Towsley, and R. Nagarajan

An End-to-End Reliable Multicast Protocol Using Polling for Scaleability 1180
M. P. Barcellos and P. D. Ezhilchelvan

An Error Control Scheme for Large-Scale Multicast Applications 1188
C. Papadopoulos, G. Parulkar, and G. Varghese

A Scalable Control Topology for Multicast Communications 1197
J. Liebeherr and B. S. Sethi

Panel - Active Networks - Hype or Next Big Thing?
Organizer: Tom Chen, GTE Labs

Active networks promise a new level of flexibility to essentially program the network for customized services or computations. This concept is quite different from the traditional view of networks, and it naturally raises controversial issues such as security and code mobility, which are the subjects of numerous current research activities.

This panel brings together a few of the prominent researchers in this new research area to present their views of what active networks will be or should (not) be.

Session 10A: Web
Session Chair: Darleen Fisher, NSF

Reducing Overhead in Flow-Switched Networks: An Empirical Study of Web Traffic 1205
A. Feldmann, J. Rexford, and R. Caceres

Scalable Delivery of Web Pages Using Cyclic Best-Effort Multicast 1214
K. C. Almeroth, M. Ammar, and Z. Fei

Techniques for Developing and Measuring High Performance Web Servers over High Speed ATM Networks 1222
J. C. Hu, S. Mungee, and D. C. Schmidt

The Network Effects of Prefetching 1232
M. Crovella and P. Barford

Session 10B: Potpourri
Session Chair: Jennifer Hou, Ohio State University

Routing Lookups in Hardware at Memory Access Speeds 1241
P. Gupta, S. Lin, and N. McKeown

IP Lookups Using Multiway and Multicolumn Search 1248
B. Lampson, V. Srinivasan, and G. Varghese

Informed-Source Coding-on-Demand (ISCOD) over Broadcast Channels 1257
Y. Birk and T. Kol

A CSMA/CD Compatible MAC for Real-Time Transmissions Based on Varying Collision Intervals 1265
O. Sharon and M. Spratt

Session 10C: Analysis Techniques for Wireless Networks
Session Chair: John Daigle, University of Mississippi

On Wireless Spectrum Estimation and Generalized Graph Coloring 1273
K. Kumaran and S. Khanna

Upper and Lower Bounds of a Class of Channel Assignment Problems in Cellular Networks 1284
A. Sen, T. Roxborough, and S. Medidi

Modeling Fast Fading Channel Dynamics for Packet Data Performance Analysis 1292
Y. Y. Kim and S. Li

Stability of a Type-II Hybrid ARQ Protocol for DS-SSMA Packet Radio Systems 1301
Q. Zhang, T. F. Wong, and J. S. Lehnert,

Session 10D: Performance Models of ATM Networks
Session Chair: Moshe Sidi, Technion, Haifa, Israel

Sizing Exit Buffers in ATM Networks under CBR Traffic 1309
H. Levy, T. Mendelson, M. Sidi, and J. Keren-Zvi

On the Impact of Long-Range-Dependent Traffic in Dimensioning ATM Network Buffer 1317
G. C. Lin and T. Suda

Per-Stream Jitter Analysis in CBR ATM Multiplexors 1325
A. Privalov and K. Sohraby

On the Capabilities of On-Off Models to Capture Arbitrary ATM Sources 1333
S. Galmes and R. Puigjaner

Session 11A: ABR and Fairness
Session Chair: Alex Kolarov, NEC America

Performance of ERICA and QFC for Transporting Bursty TCP Sources with Bursty Interfering Traffic 1341
D. P. Hong and T. Suda

Converging to Approximated Max-Min Flow Fairness in Logarithmic Time 1350
B. Awerbuch and Y. Shavitt

A Stochastic Approximation Approach for Max-Min Fair Adaptive Rate Control of ABR Sessions with MCRs 1358
S. P. Abraham and A. Kumar

A Generalized Max-Min Rate Allocation Policy and Its Distributed Implementation Using ABR Flow Control Mechanism 1366
Y. T. Hou, H. H.-Y. Tzeng, and S. S. Panwar

Session 11B: IEEE 802.14 and IEEE 802.12
Session Chair: Patrick Dowd, University of Maryland and National Security Agency

Deterministic Service Guarantees in 802.12 Networks, Part II: the Cascaded Network Case 1376
P. Kim

Evaluation of Priority and Scheduling Schemes for an IEEE 802.14 MAC Protocol Loaded by Real Traffic 1384
M. V. Ivanovich and M. Zukerman

Adaptive Control Mechanism for Cable Modems MAC Protocols 1392
D. Sala, J. O. Limb, and S. U. Khaunte

A Priority Scheme for the IEEE 802.14 MAC Protocol for Hybrid Fiber-Coax Networks 1400
M. D. Corner, N. Golmie, J. Liebeherr, and D. H. Su

Session 11C: Multicast and Multi-Path Routing Algorithms
Session Chair: Anindo Banerjea, University of Toronto

Loop-Free Multipath Routing Using Generalized Diffusing Computations 1408
W. T. Zaumen and J.J. Garcia-Luna-Aceves

An Efficient Multipath Forwarding Method 1418
J. Chen, P. Druschel, and D. Subramanian

VTDM - A Dynamic Multicast Routing Algorithm 1426
H.-C. Lin and S.-C. Lai

An Efficient Multicast Routing Algorithm for Delay-Sensitive Applications with Dynamic Membership 1433
S.-P. Hong, H. Lee, and B. H. Park

Session 11D: Performance Analysis III
Session Chair: Ness Shroff, Purdue University

A Source Model for VBR Video Traffic Based on M/G/∞ Input Processes 1441
M. Krunz and A. Makowski

Performance and Fluid Simulations of a Novel Shared Buffer Management System 1449
K. Kumaran and D. Mitra

Long-Tailed Loss Rates in a Single Server Queue 1462
P. R. Jelenkovic

Performance Analysis of a Hybrid Priority Control Scheme for Input and Output Queueing ATM Switches 1470
J. Y. Lee and Y. H. Kim

On Adaptive Bandwidth Sharing with Rate Guarantees

N.G. Duffield† T. V. Lakshman* D. Stiliadis*

† AT&T Laboratories
Rm A175, 180 Park Avenue
Florham Park,
NJ 07932-0971 USA

* Bell Labs
Lucent Technologies
101 Crawfords Corner Road
Holmdel, NJ 07733, USA

Keywords—**fair queueing, scheduling, jitter, delay bounds, real-time traffic, traffic management**

Abstract

The objective of recent research in fair queueing schemes has been to efficiently emulate a fluid-flow generalized (weighted) processor sharing (GPS) system, as closely as possible. A primary motivation for the use of fair-queueing has been its use as a means of providing bandwidth guarantees and as a consequence end-to-end delay bounds for traffic with bounded burstiness. The rate guarantees translate to scheduling weights which are set when admission control is done. A consequence of fair queueing systems closely emulating GPS is that when one or more connections are not backlogged, any "excess" bandwidth is distributed to backlogged connections in proportion to their weights. However, weights are set based on the long-term requirements of traffic flows and not in any state-dependent manner that reflects instantaneous needs. In this paper, we question the notion that queueing systems should closely emulate a GPS system. Instead of emulating GPS, we propose three modified scheduling schemes which preserve the rate guarantees of fair queueing (and hence preserve deterministic delay bounds) but adaptively redistribute the excess bandwidth such that either losses are reduced or delays equalized. We compare the performance of the proposed schemes to that of fair queueing using different traffic sources such as voice and video, as well as sources which have aggregate long-range dependent behavior. We find that the proposed schemes, in comparison to packet GPS (PGPS), reduce packet losses and curtail the tails of delay distributions for real-time traffic and hence permit the use of significantly smaller playout buffers for the same network load.

I. Introduction

Much research attention has recently been focused on the development of fair-queueing systems [3], [10], [19], [20], [21], [24] that, within the non-preemption constraints of a packet system, closely emulate fluid flow generalized processor sharing systems (GPS) and operate at high speeds. Demers, Keshav, and Shenker proposed a fair queueing scheme that emulated GPS by using a simulated fluid flow GPS system as a reference and basing packet scheduling decisions on the order of departures in the simulated GPS system. The computational burden of simulating GPS was reduced in the self-clocked fair queueing scheme proposed by Golestani [10] which showed how fair queueing could emulate GPS without simulating GPS for reference. Subsequent work has led to both improved emulation bounds, further reductions in computational complexity and efficient methods for implementation [11], [20], [21], [24]. However, there is no detailed examination of whether the exclusive emulation of GPS is always appropriate for all networking applications.

In this paper, we question the notion that queueing systems must emulate GPS closely. A primary motivation for the expected use of fair queueing in routers and switches is its ability to guarantee worst case end-to-end delay bounds for leaky bucket controlled sources [17]. However, to guarantee worst case delay bounds the scheduler merely needs to isolate flows so that each flow receives its guaranteed share of the link bandwidth [1], [19]. The guarantees do not depend on the property of GPS systems that the unused shares of non-backlogged flows be redistributed to backlogged flows in proportion to their guaranteed fractional share of link bandwidth (or equivalently scheduling weights). A fair queueing system need act as one only when all connections are backlogged. When excess bandwidth (i.e. bandwidth beyond what has already been guaranteed to backlogged connections) is available, the fair queueing system, instead of emulating GPS that distributes excess bandwidth according to long term needs, should redistribute this excess bandwidth in a manner which reflects the current or instantaneous needs of backlogged flows (i.e. a flow state dependent redistribution). The redistribution policy can be picked to optimize measures such as packet loss probabilities or to control the tails of delay distributions.

It has been shown in [1], [23] that even when a FIFO scheduler is used and all sources are leaky-bucket shaped, the mean and 99th percentile of delays are much lower than the worst-case bounds guaranteed by a GPS server. However, there are service disciplines that can distribute the bandwidth in a more elegant manner. Panwar et.al. [16] showed, that the shortest time to extinction policy is actually optimal for scheduling customers with deadlines. Furthermore, a policy that simply serves the longest queue has been shown to require less buffer space to prevent buffer losses than FIFO or other per-flow disciplines [6], [13]. Similarly, policies that serve the queue that is more likely to violate a bound are proved to provide optimal buffer utilization [9]. However, such disciplines by themselves cannot provide any bandwidth guarantees. Actually, low bandwidth constant bit rate connections are likely to starve for long intervals of time, until their queue becomes at least as large as the queues of heavily bursty, high bandwidth connections. Similarly, users that can accept a significant loss rate, using some form of Forward Error Correction, can increase their throughput by simply sending more packets. There is no method to control the bandwidth of individual users.

We propose three modified fair queueing schemes that adaptively redistribute excess bandwidth while, like fair queueing, guaranteeing each flow its specified share of the link bandwidth. This preserves fair queueing's ability to provide worst case end to end delay bounds and the schemes work like fair queueing when there is no excess bandwidth. In addition, the schemes provide worst-case fairness, bounding the time that a connection may not see any service. The excess bandwidth, however, is redistributed as follows:

1. Longest delay first (LDF) that serves the flow with current longest delay.
2. Least time to overflow (LTO) that serves the flow with minimum difference between maximum allowed delay and current delay.
3. Least time to overflow with leaky buckets (LTO_LB) that serves the flow which would cause buffer overflow first if worst case arrivals happen.

We compare the performance of these policies to PGPS using both trace driven simulations as well as simulations with traffic models for various types of sources.

The LDF policy uses excess bandwidth to reduce the variance of the delay distribution. This has the benefit of reducing the playout buffer for voice and video sources. Simulations with video traces and with voice traffic show that indeed this policy performs better than PGPS without any sacrifice of worst case guarantees. Since the deviation from the maximum allowed delay is not taken into account, flows with small delay bounds (like voice) get almost no excess bandwidth in the presence of flows with large delay bounds. Long term congestion or small errors in assigning weights can result in these flows experiencing losses much more than flows with large delay bounds. The LTO policy tries to minimize packet losses by assigning excess bandwidth under the assumption that the flow which is likely to overflow the quickest has the most instantaneous bandwidth need. In doing so, it takes into account the current deviation of each flow from its maximum allowed delay. Simulations with a mix of CBR, voice, and video sources with very different delay bounds show that this policy reduces losses for all classes as well as reduces the variance of delay for each class. The LTO_LB policy goes one step further by determining the connection that is most likely to overflow its buffer under the assumption that arrivals are leaky-bucket shaped.

A scheduling and buffer sharing scheme that tries to guarantee specific delays while ensuring loss-free behavior is presented in [7]. Our proposal is fundamentally different from the scheme presented in [7]. We assume that our system does not have enough buffers to guarantee zero losses. So our goal is not to define a system that will guarantee some specific delays

0-7803-4383-2/98/$10.00 © 1998 IEEE.

This material may be protected by Copyright law (Title 17 U.S. Code)

and at the same time zero losses. Instead, we investigate the effect of state-dependent scheduling mechanisms in minimizing the tails of the delay distributions while providing isolation, and also investigate the effect of such mechanisms on the number of buffers that need to be reserved if limited losses are acceptable.

II. Architecture

Fair Queueing systems attempt to offer the same *normalized service* to any two connections that are continuously backlogged during an interval of time $(t_1, t_2]$, where by normalized service we mean the ratio of the offered service and the allocated rate. The basic feature of such a system is that it treats connections in the same way, irrespective of their burstiness characteristics. Our claim is that end-users are not necessarily interested in a fair distribution of free bandwidth. Instead, for many applications, end-users are merely interested in minimizing their packet losses or seeing low tails in their delay distributions. Of course, the queueing scheme must be able to provide some guarantees regarding worst-case service offered to a connection when the system is fully loaded. Isolation and protection requirements such as these were defined in [1] as necessary to protect the system from misbehaving users that will try to monopolize resources.

A. Preliminaries

The definition of our state-dependent queueing system uses the methodology presented in [20] for designing fair queueing systems. The method is based on the class of servers called Rate Proportional Servers (RPS). The key concept is a virtual time function $v_i(t)$ or *system potential* associated with each connection in the system that represents the total normalized service offered to a connection during its backlogged periods. The virtual time of a connection is defined as a non-decreasing function of time during a system-busy period. When connection i is backlogged, its virtual time increases exactly by the normalized service it receives. That is, for any interval of time $(\tau, t]$ that connection i is continuously backlogged,

$$v_i(t) - v_i(\tau) = \frac{W_i(\tau, t)}{g_i},$$

where, $W_i(\tau, t)$ denotes the amount of service received by session i during the interval $(\tau, t]$ and g_i is the scheduling weight allocated to connection i. When the connection is idle, however, the virtual time of the connection is updated through the *system virtual time*. The system virtual time is a non-decreasing function of time that keeps track of the progress of the total work done by the scheduler. When an idle session i becomes backlogged at time t, its virtual time $v_i(t)$ is set to $v(t)$ to account for the service it missed. Schedulers use different functions to maintain the system virtual time, giving rise to widely different delay and fairness-behaviors. Exactly this flexibility of the rate-proportional servers allows a simple implementation of the system without the requirement of simulating in parallel a fluid system.

At any instant t, the scheduler services only the subset of connections with the minimum virtual time, and each connection in this subset receives service in proportion to its reserved rate g_i. Thus, the scheduler can be seen to increase the virtual times of the connections in this subset at the same rate. A packet server approximates this behavior by sorting the packets in the order of their finishing virtual times under the fluid server, and transmitting them in increasing order of the finishing virtual times. The finishing virtual time of the kth packet of a session i, arriving at the scheduler at time t, is computed as

$$F_i^k = \min(F_i^{k-1}, v(t)) + \frac{L_i^k}{g_i}, \tag{1}$$

where L_i^k is the length of the packet and g_i the allocated rate of session i. That is, the finishing virtual time is computed by adding the service time of the packet at the allocated rate to its starting virtual time. The latter, in turn, is the minimum of the finishing virtual time of the previous packet received from session i, and the current system virtual time.

At the time that a connection becomes backlogged, its virtual time is updated based on the system virtual time function that keeps track of the progress of the total work done by the scheduler. When an idle session i becomes backlogged at time t, its virtual time $v_i(t)$ is set as

$$v_i(t) = \max(v_i(t-), v(t)),$$

to account for the service it missed. Rate-Proportional Servers may use a wide range of functions to maintain the system virtual time, but the function

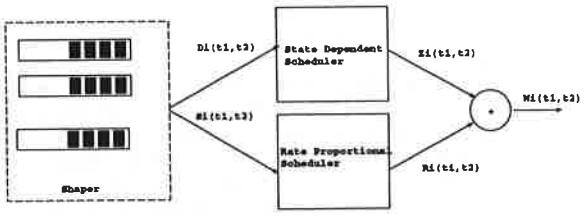


Fig. 1. Scheduling System

chosen must satisfy two fundamental properties: First, during any interval $(t_1, t_2]$ within a system-busy period, the system virtual time function must be increased with a rate of at least one, that is,

$$v(t_2) - v(t_1) \geq t_2 - t_1. \tag{2}$$

Second, the system virtual time function must never exceed the virtual time of any backlogged connection. These two conditions are sufficient to achieve a delay bound equal to that of PGPS.

B. State Dependent Fair Queueing

An illustration of the logical structure of our system is presented in Figure 1. Initially, all arriving packets enter per-connection queues in a shaper device. A packet is considered as having arrived to the system only when the last bit of the packet arrives to the system. Similarly, packet service is considered to be completed only when the last bit of the packet has been served. The shaper releases packets to the scheduler with a rate exactly equal to the allocated. We denote by $S_i(\tau, t)$ the service offered by the shaper to connection i during an interval of time $(\tau, t]$. Packets are transferred from the shaper to the the scheduler with infinite capacity. Let $R_i(\tau, t)$ denote the service offered by the RPS scheduler.

Packets that have not become eligible for service remain in the corresponding connection queue in the shaper, while all the eligible packets wait for service in the RPS scheduler queue. Service is always provided from the RPS scheduler queues, as long as packets are available there. When all scheduler queues are empty, a *state dependent scheduler* (SDS) is invoked and selects a packet from the shaper queues for transmission. The service offered to a connection does not change the state of the shaper. Let us denote with $D_i(t_1, t_2)$ and $Z_i(t_1, t_2)$ the amount of traffic that is forwarded and serviced by the SDS scheduler respectively. If the state dependent scheduler is work-conserving, then the system is work conserving as well. Note also, that for every time t, $D_i(t_1, t_2) = Z_i(t_1, t_2)$. We will refer to these systems as simply FQ-SDS.

An argument that can be made against the above adaptive policies, is that a greedy user will always receive more bandwidth. In such a case, users may request less bandwidth and try to receive more by sending more bursty traffic. Limiting such behavior can be explicity done in many ways. The simplest is policing. In most networks that provide some type of guarantees, it is reasonable to assume that there is some kind of policing. So misbehaving users will be easily isolated. If policing is not a viable option, the adaptive policies can themselves be tailored to counter greedy behavior. Instead of picking the user with maximum queue length or the minimum time to overflow, we can, for example, pick the second such candidate. In this case, users that consistently overflow the buffers cannot depend on receiving any excess bandwidth; it will not be in their best interests to be excessively bursty.

C. Redistributing Excess Bandwidth

This section describes the schemes that we use for the state dependent component of our schedulers. Note that this component is used only to redistribute the excess bandwidth and does not affect the guaranteed rates. The basic idea is that the scheduler, in picking connections to serve, should use information about the state of connections so as to minimize packet losses, reduce average delays for some traffic classes, or shape delay distributions.

Simple Longest Queue First has been shown to minimize the overall packet losses in a system with finite buffers [16]. Here, the scheduler always picks a packet from the longest queue for transmission. However, if the arrival rates and guaranteed shares are such that the expected delays are are not the same for all connections then the longest queue first scheme will always favor high delay (and generally high bandwidth connections). Also, if buffer allocations are not the same, serving the longest queue first may

lead to buffer overflow for connections with a low buffer or low relative bandwidth guarantee.

To account for these differences in buffer sizes and bandwidth requirements, we can simply weight the queue size by the allocated bandwidth to make the scheduler a Longest Delay First scheduler. The scheduler then uses the excess bandwidth to try and equalize the delays of all connections in the system and minimize the packet losses if buffer sizes have been allocated proportional to bandwidth requirements. We will refer to this algorithm as *FQ-LDF*.

Note that this still does not take into account the maximum allowed delay for each connection. An enhancement to FQ-LDF is to explicitly take into account the different delay requirements that applications may have. We would like to be able to allow low maximum delays to voice sources, somewhat higher maximum delays to video sources, and much higher maximum delays for data sources. Also, constant bit rate sources may require extremely low delays, even when their bandwidth allocation is very low. To account for these differences, a maximum delay, that is usually a function of the buffer allocated to a connection, is associated with each connection. The scheduler serves connections that are most likely to exceed their allocated maximum delay. Let D_i denote the maximum delay allocated to connection i. Let us also denote by $d_i = Q_i(t)/g_i$ the delay that the last packet of the queue associated with that connection at time t may see. The connection that should be served is the one that is more likely to overflow and exceed its delay bound. Thus, connections should be served in increasing order of the $(D_i - d_i)$ values.

The above schemes have not taken into account other characteristics of the source that may be known during admission control. Characteristics could be the maximum burst that can be transmitted by a connection or its peak arrival rate. For example, if sources are leaky bucket shaped, as is the case for sources in the Controlled Load traffic class, we could use this information to predict future arrivals from sources and calculate more precisely the connections that are likely to overflow or see delays higher than the allocated. In systems, where a dual, or multiple leaky bucket is used to shape the input traffic, a similar approach can be used. In fact, we do not require that a connection is shaped by a leaky bucket. All we need is that there is a function specified that provides information as to the maximum burst that can arrive for a connection, as well as the worst-case time that this burst may arrive.

The main idea consists of tracking the state of the connection based on its packet arrivals. This tracking is similar to the policing function that is used to drop packets from connections when they exceed their traffic specification. Based on the state of the connection we can estimate the worst-case arrivals for a given connection. In order to minimize the effect of worst-case arrivals for connections, we can serve connections in a way that will minimize the probability that they will be penalized for such an arrival pattern. Note, that since the system is only distributing the instantaneous available bandwidth in a state dependent manner and since connections are unlikely to send their worst case traffic at the same time, such an approach may lead to significant improvements in maximum delays or packet losses.

Let $B_i(\tau, t)$ be the function that describes the maximum amount of traffic that can arrive for connection i during the interval $(\tau, t]$ and let $A_i(\tau, t)$ denote the actual arrivals from session i during the same period. That is, for every interval $(\tau, t]$:

$$A_i(\tau, t) \leq B_i(\tau, t) \tag{3}$$

Then at time t, a maximum burst equal to

$$\Delta B_i \leq B_i(\tau, t) - A_i(\tau, t) \tag{4}$$

may arrive with the peak rate. If the queue size of connection i is equal to $Q_i(t)$ and the maximum buffers allocated to that connection are β_i, then a connection may overflow if

$$\beta_i - Q_i(t) \geq \Delta B \tag{5}$$

Among the connections that may overflow, we will select to transmit a packet from the connection that may overflow sooner, if the maximum burst arrives with the peak rate allocated to that connection. Or,

$$j = min_{k:\beta_k - Q_k(t) \geq \Delta B_k} \left(\frac{\Delta B_k}{P_k}\right) \tag{6}$$

where P_i is the peak rate allocated to connection i.

III. Analysis

A. Preliminaries

Definition 1: A **system busy period** is a maximal interval of time during which the server is never idle.

During a system busy period the server is always transmitting packets.

Definition 2: A **backlogged period for session** i is any period of time during which packets belonging to that session are continuously queued in the system.

Let $Q_i(t)$ represent the amount of session i traffic queued in the server at time t, that is,

$$Q_i(t) = A_i(0, t) - W_i(0, t).$$

A connection is backlogged at time t if $Q_i(t) > 0$.

Definition 3: A **session i busy period** is a maximal interval of time $(\tau_1, \tau_2]$ such that for any time $t \in (\tau_1, \tau_2]$, packets of connection i arrive with rate greater than or equal to g_i, or,

$$A_i(\tau_1, t) \geq g_i(t - \tau_1).$$

A session busy period is the maximal interval of time during which if the session were serviced with exactly the guaranteed rate, it would remain continuously backlogged. It is important to realize the basic distinction between a session backlogged period and a session busy period. The latter is defined only in terms of the arrival function and the allocated rate. Thus, the busy period serves as an invariant for evaluating the worst-case behavior of different scheduling algorithms under the same arrival pattern.

B. Worst-Case Performance

In [19] it was shown that any RPS scheduler has the same worst-case performance as a Weighted Fair Queueing scheduler, and it can thus guarantee the same worst-case delay bounds. However, the addition of the shaping mechanism and the state dependent mechanism affects the arrival pattern at the scheduler. The intuition is that the SDS scheduler is only invoked when there is a free bandwidth available. Worst-case performance usually assumes a fully utilized system. We will now show that the worst-case service offered by the system as described above is not affected by the method by which free bandwidth is distributed.

Our main goal is to provide a lower bound for the service offered by the system to a session i during a session busy period. Note, that according to our definition a session busy period is defined as the maximal interval of time during which the traffic arriving for that connection is at least equal to the allocated bandwidth. Notice, however, that the traffic that arrives to the RPS scheduler can be lower than the traffic arriving to the system. The reason is, that the connection receives excess bandwidth from the SDS server. That means, that the busy periods that the RPS scheduler sees for a connection i, may be different than the connection busy periods as seen by the system. It is easy to verify, however, that a busy period in the RPS system can continue after the end of a busy period as seen by the whole system. We will first consider the fluid modeled RPS system. Let us denote with $W_i^F(\tau, t)$ the service offered by the fluid RPS server and with $W_i^P(\tau, t)$ the service offered by the corresponding packet-by-packet server.

Lemma 1: Let τ be the beginning of a busy period for connection i. The worst case offered service to connection i for any time t during the same busy period is bounded by

$$W_i^F(\tau, t) \geq g_i(t - \tau)$$

The proof is in the Appendix. The basic idea is that since the shaper is offering service at least equal to that reserved for connection i, the scheduler will always have enough packets to service connection i with a rate equal to the reserved.

Note, that we have made no assumption for the SDS server. We can thus assume that it is a packet system. In addition, the packet and fluid RPS systems see exactly the same arrivals, and they are both work conserving systems. Thus, in both systems, the SDS server is serving packets during the same intervals of time. Thus, if the packet server offers different service than the fluid server, this difference is due to the discrepancy between the packet-by-packet and fluid RPS server. But we know from [19] that this difference is bounded by $L_i/g_i + L_{max}/r$ where L_i is the maximum packet size for connection i, L_{max} is the maximum packet size of any connection in the server and r is the link capacity. We can thus write:

Corollary 1: The worst case offered service to the packets of the j-th busy period of connection i that started at time τ is

$$W_{i,j}^P(\tau, t) \geq g_i(t - \tau - \frac{L_i}{g_i} - \frac{L_{max}}{r}).$$

This Corollary is a sufficient condition to prove that the system offers the same worst-case end-to-end delays as PGPS under the rate-proportional assignments [19].

C. Worst-Case Fairness

Worst-case fairness was initially defined by Parekh in [17] and later expanded by Bennett and Zhang [24]. The measure is based on the worst-case delay for clearing the backlog of a session's queue. According to this, for every session i that is continuously backlogged during the interval $(t_1, t_2]$,

$$W_i(t_1, t_2) \geq (t_2 - t_1)g_i - C_i, \tag{7}$$

where $W_i(t_1, t_2)$ represents the service offered to session i during the interval $(t_1, t_2]$, and C_i is a constant. We will call the smallest value of C_i satisfying this inequality as the *worst-case fairness index* (WFI). Bennett and Zhang [24] defined a fairness parameter based on the smallest value of C_i satisfying the inequality, and normalizing it to the allocated rate of the session.

The worst-case fairness index measures the deviation between the service received by a session in the packet-level scheduler and the service received by the same session in a corresponding fluid server, where the session is serviced at a rate of g_i at each instant. Minimizing this parameter improves the traffic mix at the output of the scheduler [24]. However, it should be noted that this parameter does not specify how the free bandwidth left over by idle sessions is distributed across the active ones. Specifically, even for the algorithms that distribute the free bandwidth in a state dependent manner, we can show that they offer the same worst-case fairness as Worst-Case Weighted Fair Queueing [24]. The proof is in the Appendix.

IV. Performance Studies

We simulated the system shown in Figure 1 using a variety of sources to generate traffic and using the different schedulers (FQ-LDF, FQ-LTO, and PGPS) that were described in the previous section. In the simulated system, each source is connected to a leaky bucket shaper. The shaper is connected via an access line to the multiplexing system. For each connection, the scheduler in the multiplexer maintains a separate buffer whose drain time at the guaranteed rate is equal to the connection's maximum allowed delay D_i. The sources used are:

1. Two-state exponential On-Off sources: We used these to emulate compressed voice sources and so used small maximum allowed delays for these sources. The mean load generated per connection is 1/32 of the link capacity.
2. Video teleconference sources: We used the DAR(1) model which is a Markov chain determined by three parameters: the mean, variance, and one frame-lag correlation ρ. The transition matrix is computed as:

$$P = \rho I + (1 - \rho)Q \tag{8}$$

   where I is the identity matrix, and each row of Q consists of the negative binomial (or gamma) probabilities $(f_0, ..., f_K, F_K)$ where $F_K = \sum_{k>K} f_k$ and K is the peak rate. The DAR(1) model matches the autocorrelation of the data over approximately hundred frame lags and has been shown to be a model of video teleconferences accurate enough for traffic studies [5]. We used video sources with a mean rate of 1.5 Mbps, a peak-to-mean ratio of 5, and $\rho = .98$.
3. CBR sources.
4. Data sources. These were modeled as On-Off sources where the on-time is Pareto distributed, and the off-time is exponential. The density of the Pareto distribution is given by

$$f(x) = a^{k-1}(k-1)/x^k, x >= a, a = (k-2)/(k-1)\mu \tag{9}$$

   We set $k = 5/2$, resulting in infinite variance of the on-times and long-range dependence in the aggregate traffic [12], [22]. The mean on-time was set to 6 packet transmission times and the mean off-time to 200.

The leaky bucket shapers were turned off by setting their token rates equal to the access line rate.

We first compared the performance of FQ-LDF with PGPS (which is simulated exactly) using only one type of source. Table I shows packet loss rates for different small buffer sizes from a simulation with 32 voice sources. All connections were served with equal weight and the utilization is about .98. At this high load, PGPS typically has twice the losses of FQ-LDF and redistribution of excess bandwidth to the longest delay queue reduces the aggregate loss rate. We find similar results of lower loss rates for FQ-LDF with video sources as well. With Pareto sources, the loss rates are mostly similar for FQ-LDF and PGPS. This can be accounted for by the heavy-tailed distributions of the on periods. From this property it follows that the most likely manner in which a large burst in the aggregate over all such sources occurs, is that the burst of one source is far larger than the rest. (Such consequences of underlying infinite variance models have been termed the Joseph and Noah effects; see [14], [15]). Most of the other sources have non-backlogged queues and both schedulers give most of the available bandwidth to the one very long queue. Hence, PGPS and FQ-LDF tend to behave almost identically.

Apart from packet losses, we also compared the delay distribution resulting from FQ-LDF to that of PGPS. For the 32 voice source simulation with large enough buffers to avoid packet losses, Figure 2 shows the tail frequencies (the number of times a given delay is exceeded) on a log scale. The plots show that FQ-LDF needs significantly smaller playout buffers than PGPS. We see the same effect for video sources from the boxplots in Figure 3. The utilization is .88 and the buffers are large enough so that there are no losses. The distribution of mean delays experienced by different connections with FQ-LDF and PGPS is shown in the boxplots by showing the median (the line in the middle of the box) of the means, the quartiles (the upper and lower edges of the box enclosing the median), and by showing the range of values outside the quartiles by the lines extending out from the box. The curtailment of the tail of the delay distribution due to the use of FQ-LDF is evident. Thus, for similar loads FQ-LDF has a lower playout buffer requirement than PGPS.

Next, we compare the performance of FQ-LTO with PGPS. FQ-LDF uses the excess bandwidth to equalize delays and it works very well when all sources have the same allowed delays. When different sources have different maximum allowed delays, the FQ-LTO scheduler which takes delays into account is expected to perform better than FQ-LDF and PGPS. We use 32 sources with 8 sources from each of the types described before. CBR sources are guaranteed their mean rate. They generate at most 2 back-to-back packets and have a small buffer allocation of 5 packets. Voice sources are also guaranteed their mean rate and have a maximum buffer allocation of 1200 packets (about 10 ms delay at their guaranteed rates). Video sources are guaranteed two times their mean rates and have a buffer of 12000 packets. Data sources are given a large buffer of 40000 packets. Since they are not delay sensitive, they are guaranteed a rate below their mean rate. Data sources are expected to use some the excess bandwidth from the other classes to make up for their allocation below their means.

Figure 4 shows the box plots for mean delays and 99^{th} percentile delays for the voice (exponential on-off), video, CBR and data classes. The PGPS scheduler gives excess bandwidths to the data class in proportion to the weights. Since the data class has a very large delay allowance, the FQ-LTO scheduler gives priority to the low delay classes in the distribution of excess bandwidth. Consequently, the lower delay classes have very compact delay distributions whereas the very high delay data class (which also has underallocated guaranteed bandwidth) has better tail behavior with PGPS. The FQ-LTO scheduler achieves what one would want in this mixed scenario. Worst case delay bounds are always met even for data (if it is given appropriate guaranteed rates). The lower delay classes have very small tails and hence need drastically lower playout buffers than when using PGPS.

Finally, we study the situation where we have the same mix of sources as before but their guaranteed rates are different even within a source class. We deliberately set D_is to be below what they would be if allocation was strictly based on deterministic bounds, i.e, we assume that some statistical multiplexing will happen. The question then is how well LTO, which tries to prevent buffer overflows and minimize delays, compares with PGPS. The results are shown in Figure 5. We get very low average delays for real-time sources, in comparison to the delays experienced under PGPS, even when their bandwidth allocations are very low. This is because the D_is are set low for real-time sources, and some of the excess bandwidth which would have gone to data sources under the PGPS scheme is now used to further reduce the delays of real time sources. The differences in delays between LTO and PGPS can be very high as is seen by comparing delays for the voice class. Note that if data traffic losses must be prevented, this can be done by giving those connections an appropriate weight.

Buffer size in pkts/conn	Cell Loss Ratio:PGPS	Cell Loss Ratio:FQ-LDF
75	13.18e-3	8.3e-3
100	8.198e-3	4.938e-3
125	5.61e-3	3.61e-3
150	4.245e-3	2.849e-3
250	2.061e-3	1.025e-3
350	9.614e-4	1.054e-4
425	4.876e-4	$< 10^{-6}$
500	2.8478e-4	$< 10^{-6}$

TABLE I. Cell Loss Ratios for 32 Exponential On-Off Sources with Short Buffers

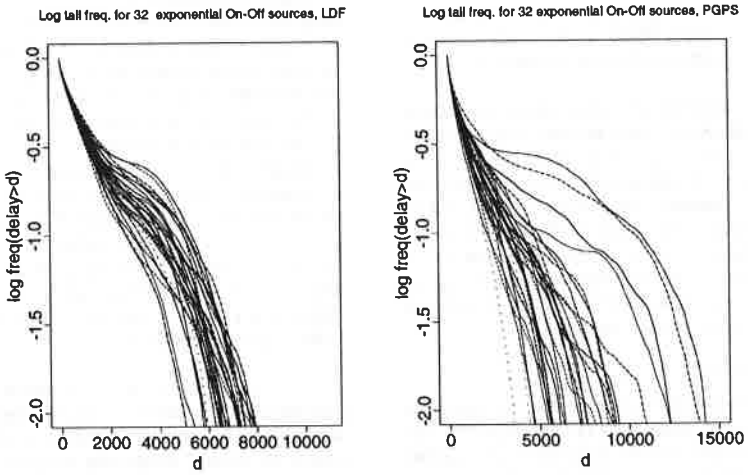


Fig. 2. Log tail frequencies for 32 exponential On-Off sources, according to service discipline (FQ-LDF or PGPS).

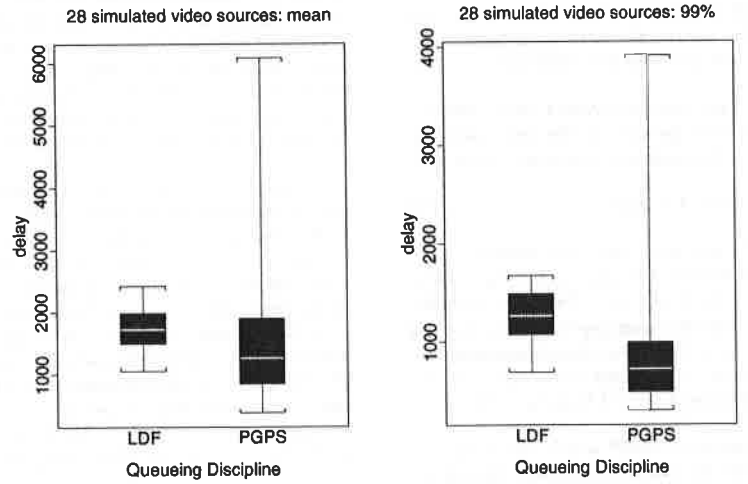


Fig. 3. Boxplots of mean delay (left plot) and 99th percentile delay (right plot) according to excess service discipline (in each plot: FQ-LDF on left, PGPS on right) for 28 simulated video sources. Key: shaded between first and third quartiles; clear bar is median; boundaries at minimum and maximum.

V. Discussion and Concluding Remarks

The main contribution of this paper is the modified fair queueing proposal that redistributes excess bandwidth in a state dependent manner. The drawback of GPS that all fair queueing systems inherit in their close emulation of GPS is that GPS severely restricts state-dependent bandwidth sharing. The only state-dependency in GPS is in the number of backlogged connections. There is no further latitude and sharing is determined by the guaranteed rates which are set based on long term needs of the connections. This restriction on bandwidth sharing is more stringent than that necessary to preserve a key property of fair queueing, the ability to guarantee delay bounds for leaky bucket controlled traffic sources. Consequently, for many applications, there is no need for fair queueing systems to emulate the potentially suboptimal excess bandwidth sharing of GPS.

With this in mind, we propose modified fair queueing schemes that emulate fair queueing only in that they provide rate guarantees and behave like fair queueing when all connections are backlogged. This is sufficient to guarantee worst case delay bounds using our scheme. Furthermore, the

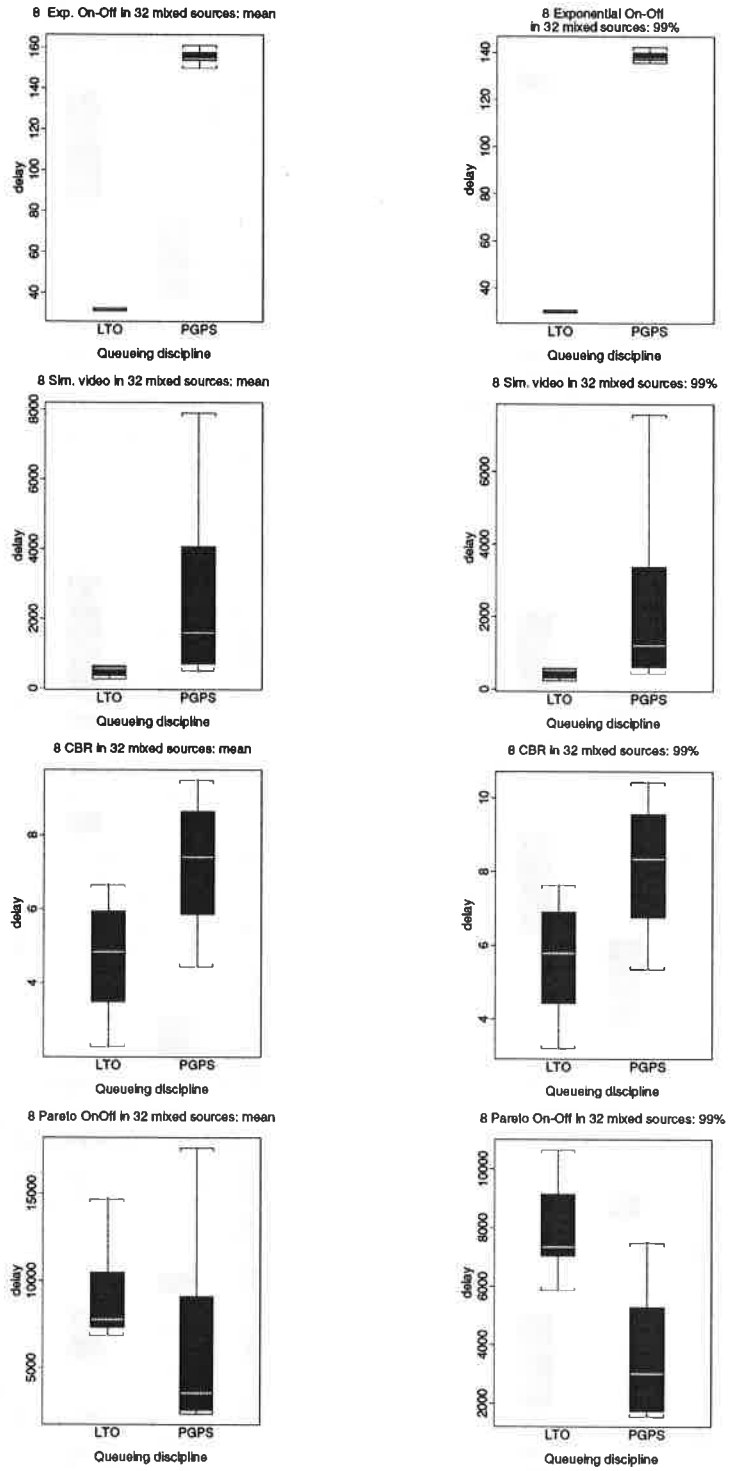


Fig. 4. Boxplots of mean delay (left column) and 99th percentile delay (right column) according to excess service discipline (in each plot: FQ-LTO on left, PGPS on right), displayed by row for each of following 2 classes of sources out of 4 classes being multiplexed: (top to bottom) Exponential On-Off, Simulated Video, CBR, Pareto On-Off. 32 sources total. Key: shaded between first and third quartiles; clear bar is median; boundaries at minimum and maximum.

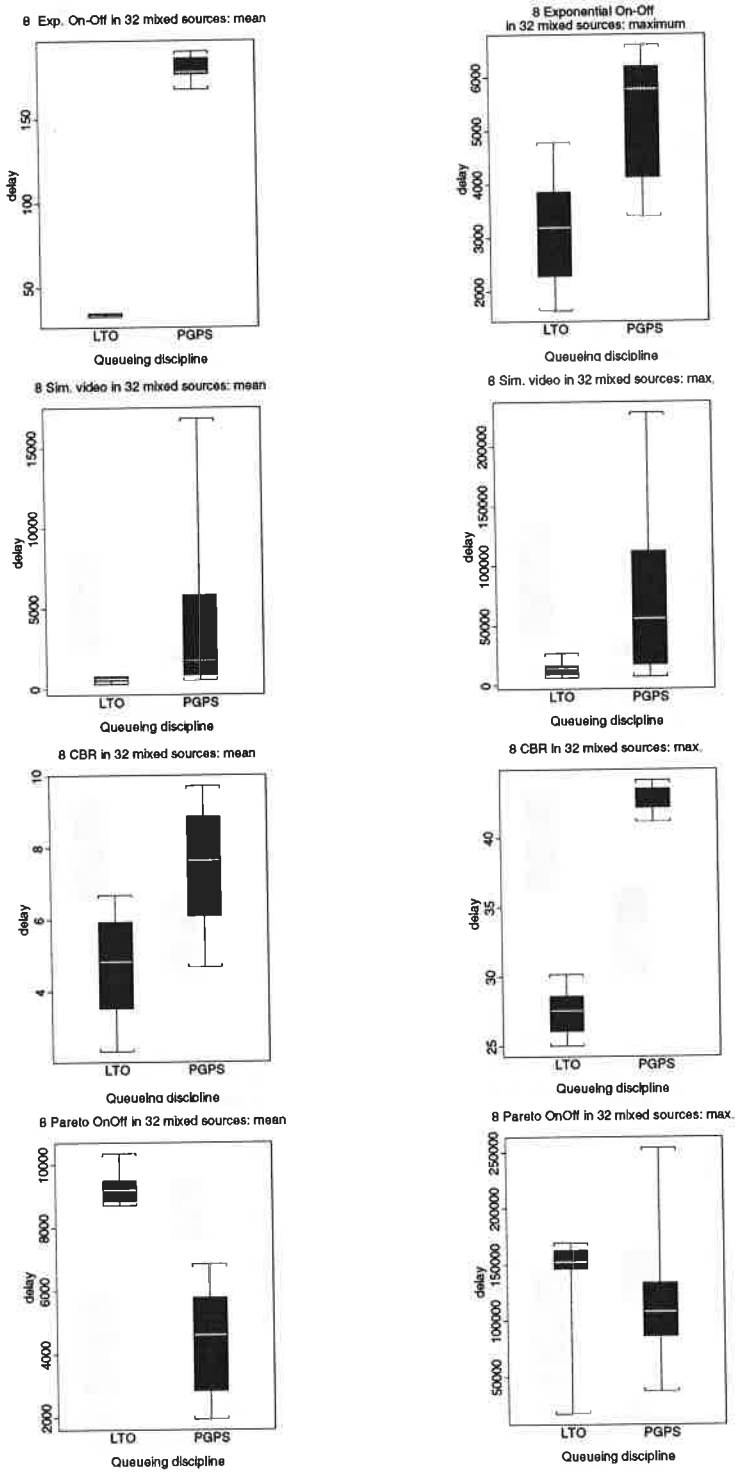


Fig. 5. Boxplots of mean delay (left column) and maximum delay (right column) according to excess service discipline (in each plot: LTO on left, PGPS on right), displayed by row for each of 4 classes of sources: (top to bottom) Exponential On-Off, Simulated Video, CBR, Pareto On-Off. 32 sources total. Key: shaded between first and third quartiles; clear bar is median; boundaries at minimum and maximum.

schemes we propose are worst case fair. Our schemes do not emulate GPS's method for excess bandwidth redistribution. Instead, we proposed three methods (the longest delay first, longest time to overflow, and longest time to overflow taking into account leaky bucket states) for redistributing the excess bandwidth. Simulations show that the FQ-LDF and FQ-LTO policies perform very well in comparison to PGPS. The adaptive bandwidth redistribution reduces packet losses and makes the delay distributions less skewed. The control of delays, in comparison to PGPS, is very significant when very delay sensitive traffic, such as CBR, is mixed with traffic with more laxity in delays. The low delay class has very short tails in comparison to PGPS and hence would need much lower playout buffers. Analytical explanation of our simulation results remains a challenging problem. Unless the fairness requirement of GPS is required for policy reasons, fair queueing systems can improve their performance by adaptive redistribution of excess bandwidth without losing their worst case fairness property or their ability to guarantee worst case delay bounds.

Similar extensions of our scheme can be easily applied to other types of schedulers based either on Earliest Deadline First [8], or on Service Curves [2]. Note, that both of these approaches define how packets must be served in the worst-case so as not to violate some specific delay bounds. However, there is no requirement as to how packets can be served if there is available free bandwidth in order to minimize the tails of the delay distributions. Actually, for both of the above approaches there is no harm if packets are transmitted earlier than expected. Our adaptive bandwidth redistribution techniques only define the method for selecting packets when excess bandwidth is available.

References

[1] D. Clark, S. Shenker, L. Zhang, "Supporting Real-Time Applications in an Integrated Services Packet Network: Architecture and Mechanism", *Proceedings of ACM SIGCOMM 1992*, pp. 14-23.

[2] R. Cruz, "Quality of service guarantees in virtual circuit switched networks," *IEEE Journal on Selected Areas In Communications*, vol. 13, pp. 1048–1056, August 1995.

[3] A. Demers, S. Keshav, S. Shenker, "Analysis and Simulation of a Fair Queueing Algorithm" *Internetworking: Research and Experience*, pp, 3-26, vol. 1, 1990.

[4] N. Duffield, T.V. Lakshman, D. Stiliadis "On Adaptive Bandwidth Sharing with Rate Guarantees," Bell Laboratories Technical Memorandum 113470-971215-07TM.

[5] A. Elwalid, D. Heyman, T. V. Lakshman, D. Mitra, A. Weiss, "Fundamental Bounds and Approximations for ATM Multiplexers with Applications to Video Teleconferencing", *IEEE Journal on Selected Areas in Communications: Special Issue on Fundamental Advances in Networking*, pp. 1004-1016, August 1995.

[6] H. Gail, G. Grover, R. Guerin, S. Hantler, Z. Rosberg, M. Sidi, "Buffer size requirements under longest queue first", *Proceedings IFIP'92*, 1992

[7] L. Georgiadis, R. Guerin and A. Parekh "Optimal multiplexing on a single link: Delay and buffer requirements", in *Processdings of IEEE INFOCOM'94*, pp.524-532.

[8] L. Georgiadis, R. Guerin, V. Peris and K.N Sivarajan "Efficient network QOS provisioning based on per-node traffic shaping," in *Proceedings of IEEE INFOCOM'96*, pp.102-110.

[9] A. Birman, H.R. Gail, S.L. Hantler and Z. Rosberg, "An optimal service policy for buffer systems," *Journal of the Association for Computing Machinery*, pp. 641-57. vol. 42, no. 3, May 1995.

[10] S. J. Golestani, "A Self-Clocked Fair Queueing Scheme for Broadband Applications", *Proceedings of INFOCOM'94*, pp. 636-646., 1994.

[11] P. Goyal, H.M. Vin and H. Chen, "Start-time Fair Queing: A Scheduling Algorithm for Integrated Services Packet Switching Networks", *Proceedings ACM SIGCOMM'96*, pp. 157-169.

[12] K. R. Krishnan, "The Hurst Parameter of Non-Markovian On-Off Traffic Sources", Internal Bellcore Report, 1995.

[13] D. S. Lee, "Weighted Longest Queue First: An Adaptive Scheduling Discipline for ATM Networks", *Proceedings INFOCOM'97*, 1997.

[14] B.B. Mandelbrot and J.W. Van Ness (1968). "Fractional Brownian Motions, Fractional Noises and Applications". *SIAM Review*, **10** 422–437

[15] B.B. Mandelbrot and J.R. Wallis (1968). "Noah, Joseph, and Operational Hydrology". *Water Resour. Res.*, **4** 909–918

[16] S. Panwar, D. Towsley, J. Wolf, "Optimal scheduling policies for a class of queues with customer deadlines to the beginning of services", *Journal of ACM*, vol. 35, no. 4, pp. 832—844, 1988.

[17] A. K. Parekh R. G. Gallager, "A Generalized Processor Sharing Approach to Flow Control in Integrated Services Networks: The Multiple Node Case", *Proceedings of INFOCOM'93*, pp. 521-530, 1993.

[18] J.L. Rexford, A.G. Greenberg, F.G Bonomi, "Hardware Efficient Fair Queueing Architectures for High-Speed Networks", *Proceedings of INFOCOM'96*, pp. 120-128, 1996.

[19] D. Stiliadis, A. Varma, "Latency-Rate Servers: A General Model for Analysis of Traffic Scheduling Algorithms", *Proceedings of INFOCOM'96*, pp. 111-119, 1996.

[20] D. Stiliadis and A. Varma, "Design and analysis of Frame-based Fair Queueing: A New Traffic Scheduling Algorithm for Packet-Switched Networks", *Proceedings of ACM SIGMETRICS '96*, pp. 104-115, May 1996.

[21] S. Suri, G. Varghese and G. Chandramenon, "Leap Forward Virtual Clock: A New Fair Queuing Scheme with Guaranteed Delays and Throughput Fairness" *Proceedings INFOCOM'97*, 1997.

[22] W. Willinger, M.S. Taqqu, R. Sherman, D.V. Wilson, "Self-Similarity Through High-Variability: Statistical Analysis of Ethernet LAN Traffic at the Source Level", *Proceedings of ACM SIGCOMM 1995*.

[23] D. Yates, J. Kurose, D. Towsley, M.G. Hluchyj, "On per-session end-to-end delay distributions and the call admission problem for real-time applications with QoS requirements", *Proceedings of ACM SIGCOMM 1993*, pp.2-12, September 1993.

[24] J.C.R. Bennett, H. Zhang, "WF^2Q: Worst-case Fair Weighted Fair Queueing", *Proceedings of INFOCOM'96*, pp. 120-128, 1996.

Appendix

Worst-Case Fairness of FQ-SDS

Let us denote by t_k, the times that a packet is released by the shaper to the RPS scheduler. In order to be able to calculate an exact bound for the worst-case fairness we need to account for the fluid server that the RPS system is simulating. Let us assume that $W_i^F(\tau, t)$ is the service offered by this fluid to server to connection i during an interval $(\tau, t]$. We will first prove the following lemma:

Lemma 2: Let $W_i^F(\tau, t)$ be the service received by session i in the fluid RPS system during an interval $(\tau, t]$ in which it is continuously backlogged. If at time τ a packet was released from the shaper then,

$$W_i^F(\tau, t) \geq g_i(t - \tau).$$

Proof: We know that since the beginning of the system busy period the shaper releases packets with rate equal to g_i. We also know that the connection virtual time is increasing by the normalized service offered to the connection, when the connection is not idle. Otherwise, it increases based on the system virtual time. Let us denote with $t' < \tau$, the last time that connection virtual time was updated by the use of the system virtual time. Thus, for time t',

$$v(t') = v_i(t') \tag{10}$$

Obviously, this happened after an idle period of connection, where the connection virtual time has remained behind the system virtual time. This update was initiated by a packet arrival. That means that at time t' a packet was released from the shaper. We know that the total arrivals to the RPS scheduler since time t' are bounded by

$$S_i(t', \tau) \leq g_i(\tau - t') \tag{11}$$

We also know that since time t' the connection virtual time has only increased by the normalized service offered to it. Thus,

$$v_i(\tau) - v_i(t') = \frac{W_i^F(t', \tau)}{g_i} \tag{12}$$

$$\leq \frac{S_i(t', \tau)}{g_i} \quad \text{From Eq.11} \tag{13}$$

$$\leq (\tau - t') \tag{14}$$

We also know that the system virtual time is increasing with a rate at least linear to the real time, or

$$v(\tau) \geq v(t') + (\tau - t'). \tag{15}$$

From Eq. (10),(14) and (15) we can conclude that

$$v_i(\tau) \leq v(\tau) \tag{16}$$

Thus when a packet is released from the shaper at time τ, the connection virtual time $v_i(\tau)$ will be updated to a value no larger to that of the system virtual time. During the interval $(\tau, t]$, the connection virtual time is increasing by the normalized service offered to that connection. Thus,

$$v_i(t) - v_i(\tau) = \frac{W_i^F(\tau, t)}{g_i}. \tag{17}$$

Therefore,

$$W_i^F(\tau,t) = g_i(v_i(t) - v_i(\tau)). \quad (18)$$

Since $v_i(t) \geq v(t)$ and $v_i(\tau) \leq v(\tau)$, we can re-write this as

$$W_i^F(\tau,t) \geq g_i(v(t) - v(\tau)). \quad (19)$$

During the interval $(\tau,t]$, the system potential is increasing with a rate at least equal to that of real time. That is,

$$v(t) - v(\tau) \geq t - \tau. \quad (20)$$

From equations (19) and (20), we have $W_i^F(\tau,t) \geq g_i(t-\tau)$. □ We can now evaluate the worst-case fairness index of the packet-based system.

Theorem 1: The queueing delay of a session-i packet arriving at time τ in a packet FQ-SDS system is upper-bounded by

$$d_i^P(\tau) \leq \frac{Q_i^P(\tau)}{g_i} + \frac{L_{max}}{r} + L_i(\frac{1}{g_i} - \frac{1}{r}), \quad (21)$$

where $Q_i^P(\tau)$ is the backlog of session i at time τ, L_i the maximum packet size of session i, and L_{max} the maximum packet size across all sessions.

Proof: Let us denote with t_k the time at which the k-th packet of session i is released from the shaper to the scheduler. Note that this time is identical in both the fluid and packet versions considered. We will prove the theorem for each instant t in the interval $(t_k, t_{k+1}]$. Let, $Q_i^F(t)$ be the backlog in the fluid server at time t and $Q_i^P(t)$ that in the packet-by-packet server. By Lemma 2, the service offered to session i after time t_k is with a rate at least equal to g_i, and without any latency.

Let us assume that the backlog of connection i is cleared at some time t^* in the fluid server. Then we can write

$$\begin{aligned} Q_i^F(t) &= W_i^F(t,t^*) \\ &= W_i^F(t,t_{k+1}) + W_i^F(t_{k+1},t^*) \\ &\geq W_i^F(t,t_{k+1}) + \max(g_i(t^* - t_{k+1}),0) \end{aligned} \quad (22)$$

We assume that if $t^* < t_{k+1}$, then $W_i^F(t_{k+1},t^*) = 0$. We can re-write Eq. (22) as

$$\begin{aligned} t^* &\leq t_{k+1} + \frac{Q_i^F(t) - W_i^F(t,t_{k+1})}{g_i} \\ &\leq (t_{k+1} - t) + t + \frac{Q_i^F(t) - W_i^F(t,t_{k+1})}{g_i}. \end{aligned} \quad (23)$$

The above equation holds for the fluid sever. The time to clear the backlog at time t in the fluid server is thus given by $d_i^F(t) = t^* - t$. The packet sever may lag the fluid server at most by the transmission time of one packet [19]. Therefore, the time $d_i^v(t)$ needed to clear the backlog of connection i in the packet-by-packet server after time t is bounded by

$$\begin{aligned} d_i^v(t) &\leq d_i^F(t) + \frac{L_{max}}{r} \\ &\leq \frac{Q_i^F(t) - W_i^F(t,t_{k+1})}{g_i} + (t_{k+1} - t) + \frac{L_{max}}{r} \end{aligned} \quad (24)$$

To evaluate the above expression, we will consider two separate cases.

Case 1: $Q_i^v(t) \geq Q_i^F(t)$. Let L_i^k denote the size of the kth packet of session i. Then,

$$\begin{aligned} W_i^F(t,t_{k+1}) &= \frac{L_i^k}{g_i} - W_i^F(t_k,t) \\ &\geq \frac{L_i^k}{g_i} - \min(L_i^k, r(t-t_k)). \end{aligned} \quad (25)$$

From Eq. (24) and (25) we can write

$$d_i^v(t) \leq (t_{k+1} - t) + \frac{Q_i^F(t)}{g_i} - \frac{L_i^k}{g_i} + \min(L_i^k, r(t-t_k)) + \frac{L_{max}}{r}. \quad (26)$$

By using the hypothesis that $Q_i^P(t) \geq Q_i^F(t)$, this becomes

$$d_i^P(t) \leq (t_{k+1} - t) + \frac{Q_i^P(t)}{g_i} - \frac{L_i^k}{g_i} + \frac{\min(L_i^k, r(t-t_k))}{g_i} + \frac{L_{max}}{r}. \quad (27)$$

The right-hand side of the above equation is maximized when $L_i^k = r(t - t_k)$. Thus,

$$\begin{aligned} d_i^P(t) &\leq (t_{k+1} - t_k - \frac{L_i^k}{r}) + \frac{Q_i^P(t)}{g_i} - \frac{L_i^k}{g_i} + \frac{L_i^k}{g_i} + \frac{L_{max}}{r} \\ &\leq \frac{Q_i^P(t)}{g_i} + L_i^k(\frac{1}{g_i} - \frac{1}{r}) + \frac{L_{max}}{r} \\ &\leq \frac{Q_i^P(t)}{g_i} + L_i(\frac{1}{g_i} - \frac{1}{r}) + \frac{L_{max}}{r}. \end{aligned} \quad (28)$$

Case 2: $Q_i^F(t) > Q_i^P(t)$. In this case, let $\Delta Q = Q_i^F(t) - Q_i^P(t)$. We can re-write Eq.(24) as

$$d_i^P(t) \leq (t_{k+1} - t) + \frac{Q_i^P(t) + \Delta Q}{g_i} - \frac{W_i^F(t,t_{k+1})}{g_i} + \frac{L_{max}}{r}. \quad (29)$$

At time t, the packet-by-packet server has offered more service to connection i than the fluid server. However, packet $k+1$ has not yet been released by the shaper. Thus, the additional service that fluid server has to offer until time t_{k+1} is equal to the additional service that the packet-by-packet server has offered until time t, plus the service it will offer until time t_{k+1}. That is,

$$W_i^F(t,t_{k+1}) = \Delta Q + W_i^P(t,t_{k+1}). \quad (30)$$

From Eq. (29) and (30),

$$d_i^P(t) \leq (t_{k+1} - t) + \frac{Q_i^P(t)}{g_i} - \frac{W_i^P(t,t_{k+1})}{g_i} + \frac{L_{max}}{r}. \quad (31)$$

Notice that at time t_k, the packet-by-packet server can only be behind the fluid server. Let us assume without loss of generality, that

$$W_i^F(0,t_k) = W_i^P(0,t_k) + \Delta W. \quad (32)$$

For the service offered to connection i by the packet-by-packet server after time t, we can write

$$\begin{aligned} W_i^P(t,t_{k+1}) &= W_i^P(t_k,t_{k+1}) - W_i^P(t_k,t) \\ &\geq L_i^k + \Delta W - W_i^P(t_k,t) \\ &\geq L_i^k + \Delta W - \min(L_i + \Delta W, r(t-t_k)). \end{aligned} \quad (33)$$

Thus, Eq. (31) becomes

$$\begin{aligned} d_i^P(t) &\leq (t_{k+1} - t) + \frac{Q_i^P(t)}{g_i} \\ &\quad -(L_i + \Delta W) + \min(L_i^k + \Delta W, r(t-t_k)) + \frac{L_{max}}{r}. \end{aligned} \quad (34)$$

The right-hand side of the above equation is maximized when $L_i + \Delta W = r(t - t_k)$. Therefore,

$$\begin{aligned} d_i^P(t) &\leq t_{k+1} - t_k - \frac{L_i^k + \Delta W}{r} + \frac{Q_i^P(t)}{g_i} + \frac{L_{max}}{r} \\ &\leq \frac{L_i^k}{g_i} - \frac{L_i^k + \Delta W}{r} + \frac{Q_i^P(t)}{g_i} + \frac{L_{max}}{r} \\ &\leq \frac{Q_i^P(t)}{g_i} + \frac{L_{max}}{r} + L_i^k(\frac{1}{g_i} - \frac{1}{r}) \\ &\leq \frac{Q_i^P(t)}{g_i} + \frac{L_{max}}{r} + L_i(\frac{1}{g_i} - \frac{1}{r}). \end{aligned} \quad (35)$$

This concludes the proof of Theorem 1. Note that the worst-case fairness index of $(L_{max}/r) + L_i/g_i - L_i/r$ given by the theorem is identical to that of WF2Q [24].

ISBN: 0-7803-4383-2

EXHIBIT 0002

Check for updates

IQ-Services: Network-Aware Middleware for Interactive Large-Data Applications

Zhongtang Cai, Greg Eisenhauer, Qi He,
Vibhore Kumar, Karsten Schwan, Matthew Wolf
{ztcai,eisen,qhe,vibhore,schwan,mwolf}@cc.gatech.edu
College of Computing
Georgia Institute of Technology

ABSTRACT

IQ-Services are application-specific, resource-aware code modules executed by data transport middleware. They constitute a 'thin' layer between application components and the underlying computational and communication resources that implements the data manipulations necessary to permit wide-area collaborations to proceed smoothly, despite dynamic resource variations. IQ-Services interact with the application and resource layers via dynamic performance attributes, and end-to-end implementations of such attributes also permit clients to interact with data providers. The joint middleware/resource and provider/consumer interactions implemented with performance attributes may be used to realize effective methods for managing the data flows in the large-data, distributed Grid applications targeted by our research. Experimental results in this paper demonstrate substantial performance improvements attained by coordinating network-level with service-level adaptations of the data being transported and by permitting end users to dynamically deploy and use application-specific services for manipulating data in ways suitable for their current needs.

Categories and Subject Descriptors

C.2.2 [**Computer-Communication Networks**]: Network Protocols—*applications, protocol architecture*; C.2.3 [**Computer-Communication Networks**]: Network Operations—*Network monitoring*; C.2.5 [**Computer-Communication Networks**]: Local and Wide-Area Networks—*high-speed, Internet*; C.3 [**Special-Purpose and Application-Based Systems**]: Real-time and embedded systems

General Terms

design, measurement, performance

Keywords

middleware, network-aware, coordinated adaptation, adaptive transport, wide-area collaboration, large-data transfer, distributed Grid application

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.
2nd Workshop on Middleware for Grid Computing Toronto, Canada
Copyright 2004 ACM 1-58113-950-0 ...$5.00.

1. INTRODUCTION

Distributed applications that 'stress' wide area networks include telepresence, remote collaboration and visualization, remote instrument access and control, and real-time monitoring and surveillance. Problems arise both from their large data volumes and from their interactive nature, the latter requiring data to be transferred with low latency and high predictability. Factors contributing to these problems include the heterogeneity and dynamics of underlying networks, the lack of support for QoS at the network level, and TCP's end-to-end congestion control that results in bursty network traffic coupled with the delivery of unstable QoS over time [19]. In addition, it is difficult to measure available network bandwidth, especially when collaborators or remote users must utilize shared substrates like the Internet.

Our research is developing middleware-based methods of providing to interactive wide area applications the data they need at acceptable levels of quality. The idea of the *IQ-Services resource-aware middleware* is to have middleware execute application-defined services for data filtering, transformation, and scheduling, and to continuously adjust such services' behaviors in accordance with current resource availability and client needs. A key, new aspect of IQ-Services is that they can be installed at runtime, initiated by applications, and used whenever or wherever appropriate. This paper's specific focus is on network behavior, which is captured with dynamic bandwidth measurement [15, 14] techniques. An example is a new IQ-Service, a data downsampling filter installed by a visualization client, which uses cutting planes in order to reduce the amount of data being transmitted to maintain acceptable data rates despite variations in available network bandwidth [13]. As a result and in contrast to adaptive applications or application-specific adaptation frameworks like those developed for multimedia systems, IQ-Services are useful for a wide range of applications and techniques. Other examples of in-stream data filtering explored in our own past work include data reordering, prioritization, and downsampling [12], the use of application-specific [21] or general [20] methods for data compression, and data transformations that implement tradeoffs in the amounts of processing vs. data volumes in the overlay networks middleware

[1]This work is supported in part by the NSF ANIR program and by the DOE MICS High Performance Networks program.

uses for service execution [4].

IQ-Services target high performance applications, and their attainment of high performance is in part because we separate the IQ-Service middleware layer from other functionality required for dynamic adaptation, such as online performance monitoring. Instead, IQ-Services use dynamic *performance attributes* to 'link' the services being adapted with the monitoring and decision-making facilities that drive adaptation processes. In our implementation of IQ-Services with the IQ-ECho middleware [10], for instance, monitoring is performed by the underlying communication protocol. This protocol, termed IQ-RUDP [12], is a reliable UDP(RUDP)-based protocol that implements instrumentation less general than but akin to the instrumented Linux protocol stacks developed in the Web100 project [14]. The composite functionality IQ-Services and IQ-RUDP extends that of MIT's communication congestion manager [3], in that IQ-Services and IQ-RUDP do not limit service adaptations to the application level. Specifically, additional protocol-level adaptations use application-provided performance attributes like desired loss tolerance or even per message attributes like 'priority markups', thereby enabling fine-grain, rapid reactions to changes in network state. Finally, performance attributes have other interesting uses, including (1) to carry information relevant to network communications, (2) to capture system-level information (e.g., provided by the DProc OS-level mechanisms described in [2]), and (3) to indicate client-level service desires, as demonstrated in this paper with a client that specifies currently desired cutting plane positions for high end remote data visualization.

This paper's contributions include (1) experimental validations of our own earlier simulation- and emulation-based results [12], in which we had shown that coordinated application/middleware-, and network-level adaptations can outperform application-only adaptation methods; (2) an overview of the middleware architecture that enables network-aware service adaptation; and (3) Using this architecture to creat and evaluate a network-aware version of GridFTP, termed *IQ-GridFTP*. IQ-GridFTP dynamically adjusts file contents during transfer (via user-supplied data manipulation services) in order to maintain high transfer rates under varying network conditions.

2. DATA-INTENSIVE, DISTRIBUTED APPLICATIONS

There are multiple, ongoing, wide-area collaborations in scientific computing, ranging from high profile efforts like Ligo [11] and DOE's Terascale Supernova Initiative (TSI) [7] to adhoc collaborations between specific sets of researchers. Collaboration tools like Access Grid [1] and grid portal efforts are evidence of the importance of remote collaboration to the broader CS and scientific communities.

Real-Time Collaboration: Our research uses an interactive high performance application, termed SmartPointer [21], to evaluate the utility of IQ-Services. This application permits remote users to collaboratively view and manipulate data produced by a molecular dynamics simulation [21]. Users can specialize data by dynamically deploying suitable middleware-level transformation and filtering services, whose behaviors are adjusted at runtime in accordance with available network resources.

Collaborative environments like these pose problems to developers both because of the heterogeneous platforms and dynamic user behaviors, and because of inherent capabilities expected by end users. One issue is the coordination of data delivery to multiple collaborators, since collaboration activities may be compromised if the information for one collaborator consistently lags that of the others. The resulting soft real-time constraints on data delivery imply a need for consistent frame rates (to insure data freshness), perhaps in preference over always receiving data at the highest levels of resolution. Another issue is fairness across multiple connections, which is particularly important when large data transfers share network links with remote control loops. Sample loops in our application include those that control a remote data filter and those that implement dynamic annotations on clients' displays.

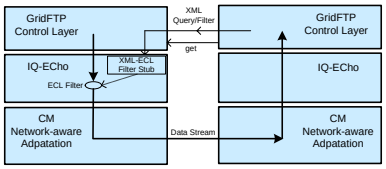


Figure 1: IQ-GridFTP Overview

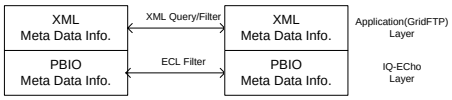


Figure 2: IQ-GridFTP MetaData Mapping

Large-Data Transfers via IQ-GridFTP: IQ-GridFTP extends the Globus GridFTP implementation [9] by realizing its goal of partial file transfers, that is, to transfer both entire files and also certain regions of files. The GridFTP standard represents this functionality with ERET (Extended Retrieve) and ESTO (Extended Store) file access instructions, which include an offset and size parameter to fetch relevant file portions. In the future, we expect file manipulations and transfers to be able to utilize the DFDL (Data Format Description Language) currently under development. This language will provide the high-level descriptions of file contents needed for realizing finer grain content filtering, down to the level of individual file data attributes.

Our intent is to utilize partial file transfers to limit network usage from a server to specific clients, by deploying customized filters at the server-resident data source. Figure 1 demonstrates such functionality by placing a client-specific XML-based, SQL-like query with the server, which then ensures that only relevant file portions and/or file attribute data are transferred. This exploits the fact that data sharing in scientific applications means transferring data on the order of GigaBytes, from one location to another, even when all the data attributes and rows might not be needed or when certain attributes can be combined to reduce the number of bytes being transferred per row [17].

3. IQ-SERVICES: SOFTWARE ARCHITECTURE

IQ-Services are application-specific code modules associated with data transport middleware. Figure 3 shows a prototypical real-time collaboration in which application-level data is distributed to remote collaborators and is manipu-

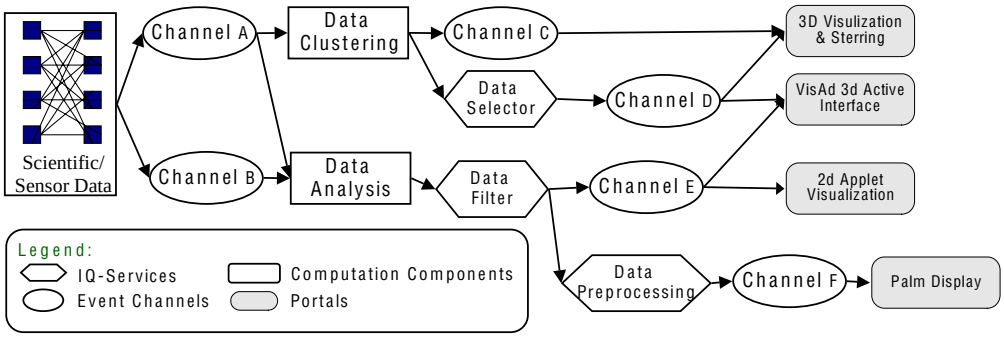


Figure 3: Typical Application Scenario

lated by computational components like data clustering or data analysis [17]. The role of IQ-Services in this scenario is illustrated by the additional data filters and selectors associated as 'handlers' with the publish/subscribe event channels used for data distribution. As evident from the figure, IQ-Services do not replace application-level computations like data analysis or clustering. Instead, they form a 'thin', efficient layer of application-provided functionality that is placed 'into' middleware by applications, the purpose of which is to allow middleware to manipulate data on its path from providers to consumers and to do so in conjunction with information about network behavior provided by communication protocols. In the IQ-ECho publish/subscribe infrastructure used in our implementation, this layer is comprised of dynamically created and deployed 'event handlers'.

IQ-Services may be associated with clients, servers, or intermediate nodes, thereby forming an overlay network. Typical configurations of overlays used in interactive scientific applications are described in [17, 6]. Overlays should be created and changed during an application's execution, due to dynamic changes in end user behavior or when indicated by substantial changes in network or machine resources. IQ-ECho supports both the runtime creation of overlay nodes and the installation of services on those nodes, using dynamic binary code generation techniques [10].

Performance attributes implement IQ-Service/monitoring interactions, which in this paper, are focused on service/protocol information exchanges. Such attributes traverse multiple layers of the protocol stack (1) to provide to the protocol layer certain application-level information like packet priorities or importance to communication protocols, and (2) to provide network-level monitoring information like available bandwidth or current packet loss to middleware-level services. Attributes may also be used to implement end-to-end performance-relevant interactions between data providers and consumers. Examples include a client's use of performance attributes to set parameters in a server's data filter, and a server-side instruction of the middleware handler to upsample the data sent since additional network bandwidth is now available. End-to-end and cross-layer interactions via performance attributes are depicted by the solid, vertical arrow in Figure 4.

Figure 4 also shows how IQ-ECho maps an application-level message submitted to an event channel to a message sent to the underlying protocol stack and communication socket used by a specific source-sink pair: after event submission, the event is mapped to a lower-level facility called

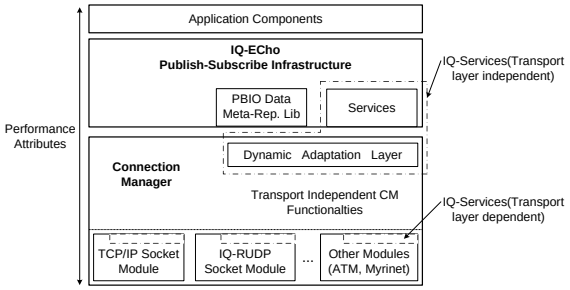


Figure 4: System Architecture Overview

the 'Communication Manager', which then sends the event to one of multiple communication protocols. This permits IQ-ECho to run on top of multiple transport protocols, including the standard TCP protocol and an instrumented version of RUDP developed in our research, termed IQ-RUDP. (see [12] for additional detail on this protocol). In comparison, for TCP, network measurement capabilities are directly associated with communication services, as depicted in Figure 4 by the 'measurement' methods layered between event management and the communication manager. By placing such methods directly 'into' the communication stream, measurements can be performed using the application's native communications, rather than with additional packet trains generated for such purposes.

For instrumented protocols like IQ-RUDP, the Dynamic Adaptation Layer (DAL) shown in Figure 4 layer moderates between protocol- and application-level semantics. The DAL can also implement service-specific performance models. Model outputs exported via performance attributes then 'drive' adaptation methods realized in the DAL, which in turn drive the actual adaptations performed by IQ-Services. Current methods implemented in our work include packet reliability control and coordination across multiple, concurrent transport-level connections. Performance 'models' utilized in our work employ (1) linear bandwidth regression and (2) an end-to-end method for measuring available bandwidth. Finally, the DAL can use additional threads to execute adaptation methods and/or periodically run network measurement tools like Pathload [15] or Netlets [18].

We have developed and deployed a variety of adaptation methods in the DAL. They include (1) coordinated packet Reliability control, which permits applications to define dif-

ferent levels of packet reliability requirements, which are then handled in a coordinated fashion by the combined actions of the IQ-Service/DAL/IQ-RUDP layers, (2) congestion avoidance that uses application-level methods to control the traffic volume imposed on the underlying network, and (3) connection coordination implementing various schemes for imposing traffic on different or multiple network connections used for data transmission. An example of (3) is a collaboration-friendly 'fairness' scheme that improves synchronization among collaborators, by reducing packet rates for flows favored by the network.

4. EXPERIMENTAL EVALUATION

Experimental results demonstrate the ability of the IQ-Services architecture and implementation to dynamically manage traffic behavior in response to network resource availability:

- demonstrations that coordinated network- and middleware-level adaptation can substantially improve the performance of wide area applications;
- sample adaptations performed for interactive high performance applications, focusing on online collaboration via large data sets; and
- comparisons of results attained with different network measurement methods.

Experimentation is performed across wide-area network links and on the NetLab network emulation facility(an Emulab site). Some measurements involve the use of IQ-ECho to transport large-scale interactive data across a 10Gbps link connecting Georgia Tech with machines located at Oak Ridge National Laboratories. The goal is to demonstrate the feasibility of wide-area collaborations like those required in DOE's Supernova Initiative, where multiple remote collaborators seek access to output data produced at rates approximating 1GB/sec for future large-scale simulation scenarios.

4.1 Coordinating the Transfer of Two IQ-RUDP Connections

In this experiment, the DAL layer coordinates message transmissions across two IQ-RUDP connections used by a single application. The experiment emulates two remote clients that receive the same visualization data stream. Since these two clients are engaged in a real-time collaboration, their desire is to receive and display the same data at the same time. The goal of DAL-based coordination in that case is to improve synchronization between the two clients, thereby reducing buffering costs and increasing the 'freshness' of information received by each client.

The experiment uses two connections that share a bottleneck link but have different RTTs and therefore, different throughput. Based on the common TCP throughput model[1], the throughput of a TCP connection will be inversely proportional to its RTT. To better synchronize the two connections, the DAL layer performs rate adjustment based on the RTT information of each connection, made available by the IQ-RUDP layer. Suppose the DAL performs rate control for every G packets sent on the faster connection, and the RTTs of the two flows are RTT_1 and RTT_2, respectively. If $RTT_1 < RTT_2$, the DAL simply adjusts the rate of connection 1 by adding a delay of $RTT_2 - RTT_1$ to its packets. With this adjustment, the connections 1 and 2 then experience similar RTTs. For the results presented in Figure 5, the parameters without rate adjustment are: $RTT_1 = 2.4ms$, $RTT_2 = 6.6ms$, and loss rate $p = 0.6\% - 1.2\%$.

[1] $T = \frac{1.22}{RTT * \sqrt{b*p}}$ [16]

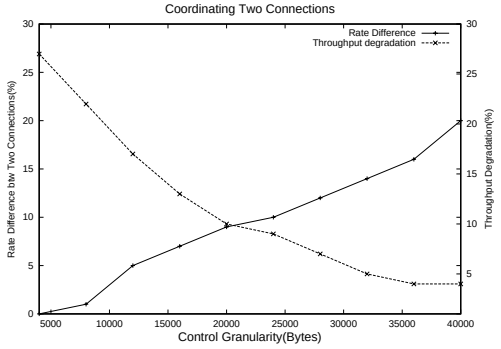


Figure 5: Coordination between Connections

This experiment has two interesting outcomes. First, it demonstrates the viability of application-level, network-aware control over multiple connections. Second, it evaluates 'granularity' issues concerning such control. The purpose of the latter is to better understand potential control limitations due to application constraints like delays in control due to threads scheduling or delays due to differences in the sizes of application-level messages being sent. In Figure 5, the X-value is the size of the data unit G over which a single rate adjustment is effective. The left Y-value is the receiving rate difference between the two receivers. The right Y-value is the throughput degradation of the faster connection. The cross-over point at packet sizes of about 20KBytes represents a reasonable point, where the faster connection's throughput is degraded by about 10% and both connections experience a rate difference of no more than 10%. It is also evident that finer grain control achieves better synchronization.

There are multiple real-life scenarios in which this 'coordination' solution is of interest. One is the co-existence of control and data connections in distributed applications, where a single, large data flow should not be able to unduly delay control flows concerning such data. Another is the co-existence of related flows, such as video and visualization data streams in real-time collaboration.

4.2 Adaptive Downsampling in Congested Networks

In real-time collaboration, it may not be viable to further reduce the rate at which data is provided to end users, since that may violate timeliness guarantee or may cause different end users to get 'out of sync' with respect to the data they are jointly viewing. A solution is to deploy application-level data filters as IQ-Services to downsample the actual data being sent prior to submitting it to the network transport. IQ-Services can regulate the traffic imposed on the underlying network by 'pacing' application-level messages to effectively reduce congestion and maintain better message delivery rates. Furthermore, clients can deploy exactly the filters they wish when and if they need them, and filters can be controlled to deliver the best quality data permitted by current network conditions. This is in stark contrast to earlier work on multimedia data services that offered only a few built-in, rigid data downsampling methods [8].

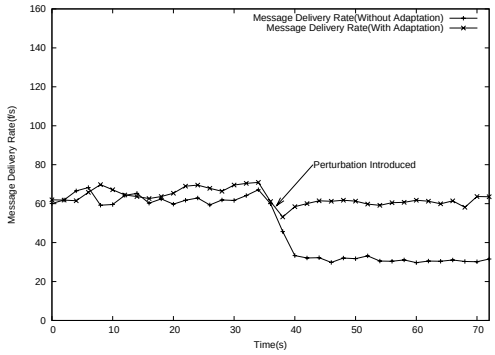


Figure 6: Adaptive Downsampling(ORNL link)

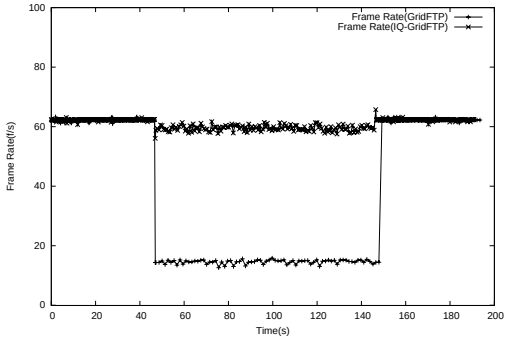


Figure 7: IQ-GridFTP Adaptation(60Mbps Cross Traffic)

Experimental results(Figure 6) demonstrate the necessity and effectiveness of IQ-Services-level adaptation through dynamic data downsampling. Here, 250Mbps cross-traffic is injected as a controlled perturbation into the link from **isleroyale** at Georgia Tech to **cruise** at ORNL, with the bottleneck at Georgia Tech's edge router. When permitting the client to characterize the priority of data, the client installs an IQ-Service data filter at the server side when congestion occurs, and henceforth receives only 'essential' (i.e., as defined by the deployed filter) data at satisfactory speeds. The specific data downsampler used in these experiments removes data relating to visual objects that are not in the user's immediate field of view. That is, the client transfers the current position and viewpoint of the user to the filter (i.e., using attributes), at the server side these values are used to determine what data set the user is currently watching, and that information is then used to transfer appropriately downsampled data to the client. The result is a consequent reduction in the network bandwidth used for data transmission, thereby speeding up data transmission.

This experiment demonstrates the utility of network-initiated data downsampling for maintaining high data rates for limited network bandwidths. By giving end users the ability to define their own IQ-Services for implementing data downsampling, similar methods can be applied to other applications, as shown by our past work on real-time and multimedia applications, with services that implement general rather than application-specific compression methods [20], and by the next sections results attained with GridFTP.

4.3 Experimental Results with IQ-GridFTP

Our implementations of IQ-GridFTP replaces the transport layer of GridFTP with IQ-ECho, thereby making it

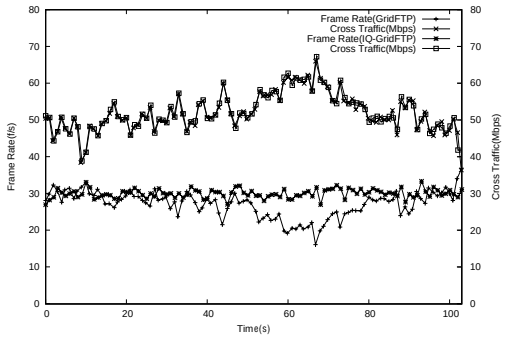


Figure 8: IQ-GridFTP Adaptation(Dynamic Cross Traffic)

easy for end users to deploy desired IQ-Services 'into' FTP transfers. Such filters may be deployed statically, or via dynamic linking methods, or written with E-Code, a highly portable subset of C, which is dynamically executed at the source. XML is used to represent the structure of data contained in the files transported by IQ-GridFTP. Specifically, whenever a client needs to do a partial retrieve on a structured file, say weather.dat, a format description file, say weather.dat.xml (we assume that such a file exists for large structured files), is first fetched from the source. Next, the requesting client creates an IQ-ECho structured data channel using the information from the xml file, and then generates E-Code for the filter using the information supplied. Using IQ-ECho's facilities for dynamic deployment, the IQ-Service that realizes this filter can dynamically: (1) select the specific file attributes that need to be transferred, (2) select the rows required, and (3) perform data reduction operation like averaging, etc., at the data source, thereby reducing the data volumes transmitted.

We compare the performance of the IQ-GridFTP with basic GridFTP in several ways. First, experiments reported in [5], compare the two implementations' effective throughput with and without cross traffic; they demonstrate that IQ-GridFTP achieves throughput similar to that of GridFTP. Second and more interestingly, Figure 7 demonstrates the advantage of utilizing IQ-Services in conjunction with GridFTP. Here, a large number of data files are transferred from an FTP server to a client. In IQ-GridFTP, when the client finds the frame rate to be lower than desired, it creates a data downsampling filter (an image resolution reduction filter when data is being visualized) and deploys it on the server, then uses the filter to reduce data volume, when necessary.

Figure 8 compares the performance of GridFTP and IQ-GridFTP with dynamic cross-traffic derived from an actual trace[2]. Cross-traffic is injected via UDP, thereby reasonably emulating actual network behavior. The data the server transfers to the client is climate data (the size of each record/frame is 172.8K bytes). The client sends a filter to the server to specify whether it wants IQ-GridFTP to adjust data precision when congestion is noticed and if so, what percentage of the data will be adjusted. While the frame rate targeted by the client is 30f/s, the acheived frame rate(averaged over 5 experiments) is improved from 27.34f/s to 30.03f/s, where the normalized standard deviation of frame rate is reduced from 0.11 to 0.04.

[1]Trace file BWY-1063337799-1 from NLANR trace repository(pma.nlanr.net), collected at BWY site on Sept. 12th, 2003.

Additional capabilities of IQ-GridFTP now under development include per connection adaptations like dynamic window size adjustments, as well as fairness improvements or stream synchronization when multiple connections are used for single, large file transfers. The latter is important for storage systems (e.g., DPSS, HPSS) that utilize parallel data transfers and data striping across multiple servers to improve performance.

5. CONCLUSIONS AND FUTURE WORK

The software architecture of IQ-Services shown in Figure 4 offers developers the ability to insert application-specific, lightweight services into data exchange middleware. In this paper, such 'IQ-Services' dynamically adjust the data sent from a distributed scientific collaboration's data providers to its data consumers [11, 7], to guarantee stable rates of data delivery despite runtime variations in available network resources. IQ-Services are also used to implement adaptive file transfers via an IQ-version of GridFTP. We have also used IQ-Services to create resource-aware communication services that apply general compression methods to data being exchanged across wide area networks [20], and we are currently using them to create adaptive remote graphical displays, for the high end 3D visual depictions required by applications like molecular dynamics. Finally, the IQ-Services software architecture has been shown useful for other communication paradigms, such as the M-by-N data exchanges used in remote storage, monitoring, or visualization systems.

The performance improvements attained by use of IQ-Services can be substantial, including up to 25% improvements in message delivery rates when information sources 'pace' the data offered in conjunction with available network bandwidth, and almost threefold improvements in message rates when a client-specific data downsampling service is used to control the amounts of data sent from data server to client.

Future work will compare service-level adaptations that utilize different network-level techniques for assessing current network bandwidth [18, 15]. In addition, it will utilize overlay networks to combine the lightweight data filtering and downsampling methods used in this paper with heavier-weight methods for data transformation and personalization executed by additional machines interposed into the path between data providers and consumers [4, 17], and/or to utilize alternative or concurrent network and machine paths from data providers to consumers.

Acknowledgments

We acknowledge the help of Constantinos Dovrolis and Nagi Rao in defining the network testbeds used in this paper's experimentation. Nagi Rao also made available the Internet-connected machine used for wide-area measurements, and his network measurement methods have been integrated into the current version of IQ-ECho. Neil Bright and the Utah Emulab team spent many hours to set up Georgia Tech's NetLab facility.

6. REFERENCES

[1] AG. Access Grid. http://www-fp.mcs.anl.gov/fl/accessgrid.

[2] S. Agarwala, C. Poellabauer, J. Kong, K. Schwan, and M. Wolf. Resource-Aware Stream Management with the Customizable dproc Distributed Monitoring Mechanisms. In *"Proceedings of High Performance Distributed Computing"*, 2003.

[3] H. Balakrishnan, H. Rahul, and S. Seshan. An Integrated Congestion Management Architecture for Internet Hosts. In *Proceedings of ACM SIGCOMM*, 1999.

[4] F. Bustamante, G. Eisenhauer, P. Widener, K. Schwan, and C. Pu. Active Streams: An Approach to Adaptive Distributed Systems. In *Proceedings of the 8th Workshop in Operating Systems (HotOS-VIII)*, Elmau/Oberbayern, Germany, May 2001.

[5] Z. Cai, G. Eisenhauer, Q. He, V. Kumar, K. Schwan, and M. Wolf. IQ-Services: Network-Aware Middleware for Interactive Large-Data Applications. Technical report, Georgia Tech, Feb. 2004.

[6] P. Chandra, A. Fisher, C. Kosak, and P. Steenkiste. Network Support for Application-Oriented Quality of Service. In *Proceedings of IEEE/IFIP International Workshop on Quality of Service*, May 1998.

[7] DOE-TSI. TeraScale Supernova Initiative. http://www.phy.ornl.gov/tsi.

[8] A. Fox, S. D. Gribble, Y. Chawathe, E. A. Brewer, and P. Gauthier. Cluster-based scalable network services. In *Proceedings of SOSP*, 1997.

[9] Globus. GridFTP. http://www-fp.globus.org/datagrid/gridftp.html.

[10] F. B. Greg Eisenhauer and K. Schwan. Event Services for High Performance Computing. In *Proceedings of High Performance Distributed Computing*, 2000.

[11] GriPhyN. The Grid Physics Network. http://www.griphyn.org.

[12] Q. He and K. Schwan. IQ-RUDP: Coordinating Application Adaptation with Network Transport. In *Proceedings of High Performance Distributed Computing*, July 2002.

[13] C. Isert and K. Schwan. ACDS: Adapting Computational Data Streams for High Performance. In *Proceedings of IPDPS*, May 2000.

[14] M. Mathis. Web100 and the End-to-End Problem. http://www.web100.org/docs/jtech/.

[15] M.Jain and C. Dovrolis. End-to-End Available Bandwidth: Measurement Methodology, Dynamics, and Relation with TCP Throughput. *IEEE/ACM Transactions in Networking*, Aug., 2003.

[16] J. Padhye, V. Firoiu, D. Towsley, and J. Kurose. Modeling TCP Throughput: A Simple Model and its Empirical Validation. In *Proceedings of ACM SIGCOMM*, 1998.

[17] B. Plale, G. Eisenhauer, K. Schwan, J. Heiner, V. Martin, and J. Vetter. From Interactive Applications to Distributed Laboratories. *IEEE Concurrency*, 6(3), 1998.

[18] N. S. V. Rao, S. Radhakrishnan, and B. Y. Cheol. NetLets: Measurement-based Routing for End-to-End Performance over the Internet. In *Proceedings of International Conference on Networking*, 2001.

[19] P. Tinnakornsrisuphap, W. Feng, and I. Philp. On the Burstiness of the TCP Congestion-Control Mechanism in a Distributed Computing System. In *Proceedings of International Conference on Distributed Computing Systems (ICDCS)*, 2000.

[20] Y. Wiseman and K. Schwan. Efficient End to End Data Exchange Using Configurable Compression. In *Proceedings of ICDCS*, Mar. 2004.

[21] M. Wolf, Z. Cai, W. Huang, and K. Schwan. Smart Pointers: Personalized Scientific Data Portals in Your Hand. In *Proceedings of IEEE/ACM Supercomputing Conference*, Nov. 2002.

EXHIBIT 0003

LIBRARY OF CONGRESS
0 016 072 258 7

Network **Congestion** Control

managing internet traffic

MICHAEL WELZL
University of Innsbruck, AUSTRIA

As the Internet becomes increasingly heterogeneous, the issue of congestion control becomes ever more important.

In order to maintain good network performance, mechanisms must be provided to prevent the network from being congested for any significant period of time. Michael Welzl describes the background and concepts of Internet congestion control in an easily comprehensible format. Throughout the book, not just the how, but also the why of complex technologies including the Transmission Control Protocol (TCP) and Active Queue Management are explained. The text also gives an overview of the state-of-the-art in congestion control research and an insight into the future.

Network Congestion Control:

> Presents comprehensive, easy-to-read documentation on the advanced topic of congestion control without heavy maths.
> Aims to give a thorough understanding of the evolution of Internet congestion control: how TCP works, why it works the way it does, and why some congestion control concepts failed for the Internet.
> Explains the Chiu/Jain vector diagrams and introduces a new method of using these diagrams for analysis, teaching & design.
> Elaborates on how the theory of congestion control impacts on the practicalities of service delivery.
> Includes an appendix with examples/problems to assist learning.
> Provides an accompanying website with Java tools for teaching congestion control, as well as examples, links to code and projects/bibliography.

This invaluable text will provide academics and researchers in computer science, electrical engineering and communications networking, as well as students on advanced networking and Internet courses, with a thorough understanding of the current state and future evolution of Internet congestion control. Network administrators and Internet service and applications providers will also find *Network Congestion Control* a comprehensive, accessible self-teach tool.

Companion Website

www.welzl.at/congestion

WILEY
wiley.com

ISBN 0-470-02528-X
9 780470 025284

WELZL

Network **Congestion** Control

managing internet traffic

WILEY

TK 5105
.875
.I57
W454
2005
Copy 1

WILEY

Network **Congestion** Control

managing internet traffic

MICHAEL WELZL



Wiley Series in Communications Networking & Distributed Systems


THE LIBRARY OF CONGRESS
1800

Network Congestion Control

Managing Internet Traffic

Michael Welzl

Leopold Franzens
University of Innsbruck



John Wiley & Sons, Ltd

Copyright © 2005

John Wiley & Sons Ltd, The Atrium, Southern Gate, Chichester,
West Sussex PO19 8SQ, England

Telephone (+44) 1243 779777

Email (for orders and customer service enquiries): cs-books@wiley.co.uk
Visit our Home Page on www.wiley.com

All Rights Reserved. No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means, electronic, mechanical, photocopying, recording, scanning or otherwise, except under the terms of the Copyright, Designs and Patents Act 1988 or under the terms of a licence issued by the Copyright Licensing Agency Ltd, 90 Tottenham Court Road, London W1T 4LP, UK, without the permission in writing of the Publisher. Requests to the Publisher should be addressed to the Permissions Department, John Wiley & Sons Ltd, The Atrium, Southern Gate, Chichester, West Sussex PO19 8SQ, England, or emailed to permreq@wiley.co.uk, or faxed to (+44) 1243 770620.

This publication is designed to provide accurate and authoritative information in regard to the subject matter covered. It is sold on the understanding that the Publisher is not engaged in rendering professional services. If professional advice or other expert assistance is required, the services of a competent professional should be sought.

Other Wiley Editorial Offices

John Wiley & Sons Inc., 111 River Street, Hoboken, NJ 07030, USA

Jossey-Bass, 989 Market Street, San Francisco, CA 94103-1741, USA

Wiley-VCH Verlag GmbH, Boschstr. 12, D-69469 Weinheim, Germany

John Wiley & Sons Australia Ltd, 42 McDougall Street, Milton, Queensland 4064, Australia

John Wiley & Sons (Asia) Pte Ltd, 2 Clementi Loop #02-01, Jin Xing Distripark, Singapore 129809

John Wiley & Sons Canada Ltd, 22 Worcester Road, Etobicoke, Ontario, Canada M9W 1L1

Wiley also publishes its books in a variety of electronic formats. Some content that appears in print may not be available in electronic books.

Library of Congress Cataloging-in-Publication Data

Welzl, Michael, 1973–
Network congestion control : managing Internet traffic / Michael Welzl.
p. cm.
Includes bibliographical references and index.
ISBN-13: 978-0-470-02528-4 (cloth : alk. paper)
ISBN-10: 0-470-02528-X (cloth : alk. paper)
1. Internet. 2. Telecommunication–Traffic–Management. I. Title.
TK5105.875.I57W454 2005
004.67′8 – dc22

2005015429

British Library Cataloguing in Publication Data

A catalogue record for this book is available from the British Library

ISBN-13 978-0-470-02528-4
ISBN-10 0-470-02528-X

Typeset in 10/12pt Times by Laserwords Private Limited, Chennai, India
Printed and bound in Great Britain by Antony Rowe Ltd, Chippenham, Wiltshire
This book is printed on acid-free paper responsibly manufactured from sustainable forestry in which at least two trees are planted for each one used for paper production.

Contents

Foreword xi

Preface xiii

List of Tables xv

List of Figures xvii

1 Introduction 1
- 1.1 Who should read this book? . . . 1
- 1.2 Contents . . . 2
- 1.3 Structure . . . 4
  - 1.3.1 Reader's guide . . . 5

2 Congestion control principles 7
- 2.1 What is congestion? . . . 7
  - 2.1.1 Overprovisioning or control? . . . 8
- 2.2 Congestion collapse . . . 10
- 2.3 Controlling congestion: design considerations . . . 13
  - 2.3.1 Closed-loop versus open-loop control . . . 13
  - 2.3.2 Congestion control and flow control . . . 14
- 2.4 Implicit feedback . . . 14
- 2.5 Source behaviour with binary feedback . . . 16
  - 2.5.1 MIMD, AIAD, AIMD and MIAD . . . 16
- 2.6 Stability . . . 19
  - 2.6.1 Control theoretic modelling . . . 19
  - 2.6.2 Heterogeneous RTTs . . . 20
  - 2.6.3 The conservation of packets principle . . . 21
- 2.7 Rate-based versus window-based control . . . 21
- 2.8 RTT estimation . . . 23
- 2.9 Traffic phase effects . . . 24
  - 2.9.1 Phase effects in daily life . . . 26
- 2.10 Queue management . . . 26
  - 2.10.1 Choosing the right queue length . . . 27
  - 2.10.2 Active queue management . . . 27
- 2.11 Scalability . . . 28
  - 2.11.1 The end-to-end argument . . . 28

2.11.2 Other scalability hazards . . . 29
2.12 Explicit feedback . . . 31
2.12.1 Explicit congestion notification . . . 32
2.12.2 Precise feedback . . . 32
2.13 Special environments . . . 36
2.14 Congestion control and OSI layers . . . 38
2.14.1 Circuits as a hindrance . . . 39
2.15 Multicast congestion control . . . 40
2.15.1 Problems . . . 42
2.15.2 Sender- and receiver-based schemes . . . 43
2.16 Incentive issues . . . 44
2.16.1 Tragedy of the commons . . . 44
2.16.2 Game theory . . . 44
2.16.3 Congestion pricing . . . 45
2.17 Fairness . . . 47
2.17.1 Max–min fairness . . . 48
2.17.2 Utility functions . . . 49
2.17.3 Proportional fairness . . . 51
2.17.4 TCP friendliness . . . 51
2.18 Conclusion . . . 52

3 Present technology **55**
3.1 Introducing TCP . . . 56
3.1.1 Basic functions . . . 57
3.1.2 Connection handling . . . 59
3.1.3 Flow control: the sliding window . . . 60
3.1.4 Reliability: timeouts and retransmission . . . 61
3.2 TCP window management . . . 62
3.2.1 Silly window syndrome . . . 62
3.2.2 SWS avoidance . . . 62
3.2.3 Delayed ACKs . . . 64
3.2.4 The Nagle algorithm . . . 64
3.3 TCP RTO calculation . . . 65
3.3.1 Ignoring ACKs from retransmissions . . . 66
3.3.2 Not ignoring ACKs from retransmissions . . . 66
3.3.3 Updating RTO calculation . . . 67
3.4 TCP congestion control and reliability . . . 69
3.4.1 Slow start and congestion avoidance . . . 69
3.4.2 Combining the algorithms . . . 71
3.4.3 Design rationales and deployment considerations . . . 73
3.4.4 Interactions with other window-management algorithms . . . 74
3.4.5 Fast retransmit and fast recovery . . . 75
3.4.6 Multiple losses from a single window . . . 77
3.4.7 NewReno . . . 79
3.4.8 Selective Acknowledgements (SACK) . . . 81
3.4.9 Explicit Congestion Notification (ECN) . . . 84

3.5 Concluding remarks about TCP . . . 88
3.6 The Stream Control Transmission Protocol (SCTP) . . . 91
3.7 Random Early Detection (RED) . . . 93
3.8 The ATM 'Available Bit Rate' service . . . 96
3.8.1 Explicit rate calculation . . . 98
3.8.2 TCP over ATM . . . 100

4 Experimental enhancements 103
4.1 Ensuring appropriate TCP behaviour . . . 104
4.1.1 Appropriate byte counting . . . 104
4.1.2 Limited slow start . . . 106
4.1.3 Congestion window validation . . . 107
4.1.4 Robust ECN signalling . . . 108
4.1.5 Spurious timeouts . . . 109
4.1.6 Reordering . . . 113
4.1.7 Corruption . . . 115
4.2 Maintaining congestion state . . . 119
4.2.1 TCP Control Block Interdependence . . . 119
4.2.2 The Congestion Manager . . . 119
4.2.3 MulTCP . . . 121
4.3 Transparent TCP improvements . . . 123
4.3.1 Performance Enhancing Proxies (PEPs) . . . 123
4.3.2 Pacing . . . 126
4.3.3 Tuning parameters on the fly . . . 128
4.4 Enhancing active queue management . . . 129
4.4.1 Adaptive RED . . . 130
4.4.2 Dynamic-RED (DRED) . . . 131
4.4.3 Stabilized RED (SRED) . . . 132
4.4.4 BLUE . . . 133
4.4.5 Adaptive Virtual Queue (AVQ) . . . 133
4.4.6 RED with Preferential Dropping (RED-PD) . . . 134
4.4.7 Flow Random Early Drop (FRED) . . . 135
4.4.8 CHOKe . . . 135
4.4.9 Random Early Marking (REM) . . . 136
4.4.10 Concluding remarks about AQM . . . 137
4.5 Congestion control for multimedia applications . . . 139
4.5.1 TCP-friendly congestion control mechanisms . . . 143
4.5.2 The Datagram Congestion Control Protocol (DCCP) . . . 149
4.5.3 Multicast congestion control . . . 155
4.6 Better-than-TCP congestion control . . . 160
4.6.1 Changing the response function . . . 161
4.6.2 Delay as a congestion measure . . . 164
4.6.3 Packet pair . . . 167
4.6.4 Explicit feedback . . . 169
4.6.5 Concluding remarks about better-than-TCP protocols . . . 175
4.7 Congestion control in special environments . . . 176

5 Internet traffic management – the ISP perspective **181**
5.1 The nature of Internet traffic . . . 182
5.2 Traffic engineering . . . 184
5.2.1 A simple example . . . 185
5.2.2 Multi-Protocol Label Switching (MPLS) . . . 186
5.3 Quality of Service (QoS) . . . 187
5.3.1 QoS building blocks . . . 188
5.3.2 IntServ . . . 190
5.3.3 RSVP . . . 191
5.3.4 DiffServ . . . 191
5.3.5 IntServ over DiffServ . . . 192
5.4 Putting it all together . . . 194

6 The future of Internet congestion control **199**
6.1 Small deltas or big ideas? . . . 200
6.1.1 TCP-friendliness considerations . . . 201
6.1.2 A more aggressive framework . . . 203
6.2 Incentive issues . . . 205
6.2.1 The congestion response of UDP-based applications . . . 206
6.2.2 Will VoIP cause congestion collapse? . . . 209
6.2.3 DCCP deployment considerations . . . 211
6.2.4 Congestion control and QoS . . . 212
6.3 Tailor-made congestion control . . . 214
6.3.1 The Adaptation Layer . . . 214
6.3.2 Implications . . . 215

A Teaching congestion control with tools **219**
A.1 CAVT . . . 220
A.1.1 Writing scripts . . . 223
A.1.2 Teaching with CAVT . . . 224
A.1.3 Internals . . . 225
A.2 *ns* . . . 227
A.2.1 Using *ns* for teaching: the problem . . . 227
A.2.2 Using *ns* for teaching: the solution . . . 228
A.2.3 NSBM . . . 229
A.2.4 Example exercises . . . 233

B Related IETF work **235**
B.1 Overview . . . 235
B.2 Working groups . . . 236
B.3 Finding relevant documents . . . 238

C List of abbreviations **239**

Bibliography **243**

Index **259**

4.5 Matching a fluctuating application stream onto a fluctuating congestion control mechanism . . . 141
4.6 Matching a constant application stream onto a constant congestion control mechanism . . . 142
4.7 Matching an adaptive application stream onto a fluctuating congestion control mechanism . . . 143
4.8 The DCCP generic header (both variants) . . . 153
4.9 TCP congestion avoidance with different link capacities . . . 160
4.10 Packet pair . . . 168
4.11 CADPC/PTP in action. . . . 173

5.1 A traffic engineering problem . . . 185
5.2 A generic QoS router . . . 188
5.3 Leaky bucket and token bucket . . . 189
5.4 IntServ over DiffServ . . . 193

6.1 A vector diagram of TCP Reno. . . . 203
6.2 The test bed that was used for our measurements . . . 206
6.3 Sender rate and throughput of streaming media tools . . . 208
6.4 Average packet length of interactive multiplayer games . . . 209
6.5 The Adaptation Layer . . . 215
6.6 Four different ways to realize congestion control . . . 216

A.1 Screenshot of CAVT showing an AIMD trajectory . . . 221
A.2 Screenshot of the CAVT time line window (user 1 = AIAD, user 2 = MIMD, equal RTTs) . . . 222
A.3 Some CADPC trajectories . . . 224
A.4 Class structure of CAVT . . . 226
A.5 A typical *ns* usage scenario. . . . 228
A.6 A screenshot of NSBM . . . 230
A.7 A screenshot of *nam* . . . 231
A.8 A screenshot of *xgraph* . . . 232
A.9 The congestion collapse scenario with 1 Mbps from source 1 to router 2 . . 233

whether control and/or measurement actions from within the network should be included. Since it seems obvious that adding these functions will complicate things significantly, we postpone such considerations and start with the simpler case of *implicit* feedback, that is, measurements that are taken at the receiver and can be used to deduce what happens within the network.

In order to determine what such feedback can look like, we must ask the question, What can happen to a packet as it travels from source to destination? From an end-node perspective, there are basically three possibilities:

1. It can be delayed.
2. It can be dropped.
3. It can be changed.

Delay can have several reasons: distance (sending a signal to a satellite and back again takes longer than sending it across an undersea cable), queuing, processing in the involved nodes, or retransmissions at the link layer. Similarly, packets can be dropped because a queue length is exceeded, a user is not admitted, equipment malfunctions, or link noise causes a checksum of relevance to intermediate systems to fail. Changing a packet could mean altering its header or its content (payload). If the content changed but the service provided by the end-to-end protocol includes assurance of data integrity, the data carried by the packet become useless, and the conclusion to be made is that some link technology in between introduced errors (and no intermediate node dropped the packet due to a checksum failure). Such errors usually stem from link noise, but they may also be caused by malicious users or broken equipment. If the header changed, we have some form of explicit communication between end nodes and inner network nodes – but at this point, we just decided to ignore such behaviour for the sake of simplicity. We do not regard the inevitable function of placing packets in a queue and dropping them if it overflows as such active participation in a congestion control scheme.

The good news is that the word 'queue' was mentioned twice at the beginning of the last paragraph – at least the factors 'delay' and 'packet dropped' can indicate congestion. The bad news is that each of the three things that can happen to a packet can have quite a variety of reasons, depending on the specific usage scenario. Relying on these factors therefore means that implicit assumptions are made about the network (e.g. assuming that increased delay always indicates queue growth could mean that it is assumed that a series of packets will be routed along the same path). They should be used with care.

Note that we do not have to restrict our observations to a single packet only: there are quite a number of possibilities to deduce network properties from end-to-end performance measurements of series of packets. The so-called *packet pair* approach is a prominent example (Keshav 1991a). With this method, two packets are sent back-to-back: a large packet immediately followed by a small packet. Since it is reasonable to assume that there is a high chance for these packets to be serviced one after another at the bottleneck, the interspacing of these packets can be used to derive the capacity of the bottleneck link. While this method clearly makes several assumptions about the behaviour of routers along the path, it yields a metric that could be valuable for a congestion control mechanism (Keshav 1991b). For the sake of simplicity, we do not discuss such schemes further at this point and reserve additional observations for later (Section 4.6.3).

This material may be protected by Copyright law (Title 17 U.S. Code)

solutions for congestion control. While it is well known that Internet access now spans a broad range of devices and environments (from the cell phone in my pocket to the satellite way above me – in fact, there is now even research on the 'Interplanetary Internet'), we have totally neglected this aspect of network communications up to now. However, it does play a significant role for congestion control.

Briefly put, the influence of heterogeneous environments can falsify the interpretation of network measurements. As already mentioned in Section 2.4, only three things can happen to a single packet as it traverses the network: (i) it can be delayed, (ii) it can be dropped, and (iii) it can be changed. At least two of these effects lend themselves to interpretations by an end-to-end receiver for the sake of providing the sender with feedback about the state of congestion in the network: delay, which can be a result of growing queues, and packet drops, which can be a result of an overflowing queue. We have also seen that, whether increasing delay is interpreted as a sign of congestion or not, measuring the RTT is useful for several things. Here is a brief (non comprehensive) list of different environments with their properties that have an influence on end-to-end delay and loss:

Wireless links: In wireless usage scenarios, bit errors are frequent; they can, for instance, occur when a user passes by a wall or enters an elevator. At the wireless receiver, receiving noisy data means that a checksum will fail. Depending on the specific technology that is used, this can mean two things:

1. The packet is dropped (because its contents are now useless).
2. The packet is delayed, because the link layer retransmits it if an error occurs. Even then, if multiple retransmissions fail, the packet can eventually be dropped.

Typically, bit errors are not statistically independent, that is, they are clustered in time (Karn et al. 2004).

Satellite links: Satellites combine a number of potentially negative factors:

- Connections are noisy, which means that they share the negative properties of wireless links described above. This problem may be compensated for with *Forward Error Correction* (FEC), but this can add delay.
- They show a long delay, which is always a problem because of its impact on the precision of RTT measurements: the larger the delay, the larger its fluctuations. In the case of *GEOs* – satellites in geostationary orbit – the delay is between 239.6 and 279.0 ms (Allman et al. 1999a). While 'low earth orbit' satellites (*LEOs*) seem to alleviate this problem via their significantly reduced one-way propagation delay of about 20 to 25 ms, it may be necessary to route data *between* them ('LEO handoff'), which can cause packet loss and propagation delay fluctuations (Hassan and Jain 2004).
- Satellites are typical examples of so-called long fat pipes – links with a large bandwidth $\times$ delay product. As we have discussed earlier, this value dictates the amount of data a protocol should have 'in flight' (transmitted, but not yet acknowledged) in order to fully utilize a link (Allman et al. 1999a). Such links cause problems for mechanisms like AIMD: additively increasing the rate with a constant factor until the link is saturated can take a very long time (perhaps

hop after the sender; this can be attained by choosing a suitable packet (actually *frame*) type from the underlying link layer technology. For instance, 802.3x defines a PAUSE frame that can be used as a gap packet when its 'pause time' is set to zero.

4.3.3 Tuning parameters on the fly

Another way to control the behaviour of TCP without changing its implementation is to adaptively tune its parameters. One particularly interesting parameter is the buffer size of the sender, which automatically imposes an upper limit on its window. Tuning the maximum window size itself may seem more attractive because its impact is perhaps more immediate, but this parameter is not always available – remember that the goal is to influence TCP without changing its implementation. Since the maximum sender window should ideally equal the bandwidth × RTT product of the connection, the required buffer size varies as widely as the environments that TCP/IP is used in; an example range given in (Hassan and Jain 2004) and (Feng et al. 2003) is a short modem connection with a capacity of 56 kbps and a delay of 5 ms – corresponding with a window size of 36 bytes – versus a long-distance ATM connection with a capacity of 622 Mbps and a delay of 100 ms, which corresponds with a window size of 7.8 MB. Choosing the latter window size for the former scenario wastes over 99% of its allocated memory whereas choosing the former window size for the latter scenario means that up to 99% of the network capacity is wasted.

One simple solution to the problem is to manually tune the buffer size depending on the environment. The problem with this approach is that both the network capacity that is available to an end system and the RTT fluctuate, which is caused by effects such as routing changes or congestion in the network. It is therefore desirable to automatically adapt the buffer size to the given network conditions. While they are not the only methods available (e.g. the memory management technique in Linux kernel version 2.4 also does this), two well-known approaches can be seen as representatives:

1. *Auto-tuning* utilizes TCP header information and the Timestamps option to estimate the bandwidth × RTT product and adapt the buffer size on the fly; this is a sender-side kernel modification where several concurrent TCP connections can share a single buffer (Semke et al. 1998).

2. *Dynamic Right-Sizing (DRS)* is a receiver-side modification that makes the sender change its maximum window by tuning the advertised window (Feng et al. 2003); thus, the *flow control* functionality of TCP is used by this approach. The receiver estimates the current *cwnd* of the sender by monitoring the throughput; the RTT is estimated by measuring the time between sending an ACK and the reception of data that are at least one window beyond the ACKed data. Since this assumes that the sender will transmit new data right away when it receives an ACK, but this may not always be the case (i.e. the sender may not have new data to send), this RTT is interpreted as an upper bound. If the receiver is itself sending back data (remember, TCP connections are bidirectional), the RTT will automatically be estimated by the TCP implementation.

While these techniques do not require changing the TCP code, both auto-tuning and DRS operate at the kernel level. For instance, the advertised receiver window is usually not

function that retains the strengths of BIC yet simplifies its window control and enhances its TCP-friendliness. As the name suggests, the function that was found is cubic; its input parameters are the maximum window size (from normal BIC), a constant scaling factor, a constant multiplicative decrease factor and the time since the last drop event. The fact that this 'real' time dominates the control is what makes it independent of the RTT: for two flows that experience loss at the same time, the result of this function depends on the time since the loss occurred even if they have different RTTs.

The cubic response function enhances the TCP-friendliness of the protocol; at the same time, the fact that the window growth function does not depend on the RTT allows TCP to be more aggressive than CUBIC when the RTT is small. Therefore, CUBIC additionally calculates how fast a TCP sender would open its window under similar circumstances and uses the maximum of this result and the result given by the standard CUBIC function.

TCP Westwood+

TCP Westwood+ (a refined version of 'TCP Westwood', first described in (Grieco and Mascolo 2002)) resembles HSTCP and STCP in that it changes the response function to react not in a fixed manner but in proportion to the current state of the system. It is actually quite a simple change, but there is also a fundamental difference: in addition to the normal inputs of a TCP response function, TCP Westwood+ utilizes the actual rate at which packets arrive at the receiver as an input – this is determined by monitoring the rate of incoming ACKs. Strictly speaking, this protocol therefore uses slightly more implicit feedback than the ones discussed so far. A result of this is that it works well over wireless networks: it will reduce its rate severely only when a significant number of packets are lost, which may either indicate severe congestion or a long series of corruption-based loss events. A single loss event that is due to link noise does not cause much harm – this is different from standard TCP. Also, while this protocol is TCP-friendly according to (Grieco and Mascolo 2004), it does not distinguish between 'low loss' and 'high loss' scenarios.

The response function of TCP Westwood+ is easily explained: its only divergence from standard TCP Reno or NewReno behaviour is the update of *ssthresh* (the starting point for fast recovery) in response to a congestion event. Instead of applying the fixed rate decrease by half, it sets this variable to the product of the estimated bandwidth (as determined from the rate of ACKs) and the minimum RTT – this leads to a more-drastic reduction in a case of severe congestion than in a case of light congestion. The key to the efficiency of TCP Westwood+ is the function that counts and filters the stream of incoming ACKs; this was designed using control theory. While the original version of TCP Westwood was prone to problems from 'ACK compression' (see Section 4.7), this was solved in TCP Westwood+ by means of a slightly modified bandwidth estimation algorithm.

4.6.2 Delay as a congestion measure

When congestion occurs in the network, a queue grows at the bottleneck and eventually, if the queue length cannot compensate, packets will be dropped. It therefore seems reasonable to react upon increasing delay. Note that only changes can be used, as a fixed delay measure does not explain much: two mildly congested queues may yield the same delay as a single severely congested one. Only delay *changes* can be used, that is, measurements must always be (implicitly or explicitly) combined with a previously recorded value. Delay measurements

already control the behaviour of TCP via its RTT estimate, and increasing delay (which might be caused by congestion) will even make it react slower; still, the measurements are only interpreted for the sake of proper self-clocking and not as a real congestion indication.

Using delay in this manner has the potential advantage that a congestion control mechanism can make a more reasonable decision because it has more feedback. At the same time, it is non-intrusive, that is, it does not require any additional packets to be sent, and routers do not have to carry out extra processing. However, this feedback must be used with care, as it is prone to misinterpretations (see Section 2.4) – delay can also be caused by routing or link layer ARQ.

The idea of delay-based congestion control is not a new one; in fact, a mechanism that does this was described as early as 1989 (Jain 1989). In this scheme, the previous and current window and delay values are used to calculate a so-called normalized delay gradient, upon which the decision whether to increase or decrease the rate is based. Since then, several refinements were proposed; we will look at two popular ones.

TCP Vegas

An implicit form of delay-based congestion control can be found in a well-known TCP variant called *TCP Vegas* (Brakmo et al. 1994). This is a sender-side TCP modification that conforms to the normal TCP Reno standard: the specification does not forbid sending *less* than a congestion window, and this is exactly what TCP Vegas does. Like TCP Westwood+, it determines the minimum RTT of all measurements. The expected throughput is then calculated as the current size of *cwnd* divided by the minimum RTT. Additionally, the actual throughput is calculated by sending a packet, recording how many bytes are transferred until reception of the corresponding ACK and dividing this number by the sample RTT; if the sender fully utilizes its window, this should be equal to *cwnd* divided by SRTT. The calculation is carried out every RTT, and the difference between the actual rate and the expected rate governs the subsequent behaviour of the control.

The expected rate minus the actual rate is assumed to be non-negative, and it is compared against two thresholds. If it is underneath the lower threshold, Vegas increases *cwnd* linearly during the next RTT because it assumes that there are not enough data in the network. If it is above the upper threshold (which means that the expected throughput is much lower than the actual sending rate), Vegas decreases *cwnd* linearly. The congestion window remains unchanged when the difference between the expected and the actual rate is between the two thresholds. In this way, it can converge to a stable point of operation – this is a major difference between TCP Vegas and TCP Reno, which always needs to exceed the available capacity in order to detect congestion. Vegas can detect incipient congestion and react early, and this is achieved by monitoring delay: if the sender fully utilizes its congestion window, the only difference between the actual and expected rate calculated by Vegas is given by the minimum RTT versus the most-recent RTT sample.

This mechanism is the core element of TCP Vegas; there are two more features in the protocol: (i) it uses the receipt of certain ACKs as a trigger to check whether the RTO timer expired (this is an enhancement over older implementations with coarse timers), and (ii) it only increases the rate every other RTT in slow start so as to detect congestion during the intermediate RTTs. Despite its many theoretical advantages, Vegas is hardly used in practice nowadays because it is less aggressive than TCP Reno; if it was used in the Internet, it would therefore be 'pushed aside' (Grieco and Mascolo 2004). The value of TCP Vegas lies

limited transmit, 80, 114
MSS, 58
Nagle Algorithm, 65
NewReno, 80–81
Nonce, 87
Padhye equation, 91
partial ACK, 79
PAWS, 85
Peak Hopper, 113
PH-RTO, 113
Pipe algorithm, 83
Reno, 77
Reordering, 113–115
RTO, 61, 65
RTO calculation, 65–69
RTT estimation, 65–69
SACK, 82–85
segment, 57
sliding window, 60
slow start, 69–76
spoofing, 124
spurious timeouts, 109–113
ssthresh, 71
SWS, 62
Tahoe, 73
TCB Interdependence, 119
Timestamps option, 67, 84, 110
Vegas, 165–166
Westwood+, 164
TCP compatible, 52
TCP friendliness, 52, 201–205
Teaching, 219–233
TEAR, 145
TFMCC, 156–157
TFRC, 147–149
Timestamps option, 67, 84, 110
Tools, 219–233
Traffic
  characteristics, 182–183
  engineering, 181, 184–187, 194–198
  management, 181–198
Tragedy of the commons, 44, 75

UDP-Lite, 117
Unicast, 40
Utility function, 49

Vegas, 165–166

WEBRC, 159
Weighted Fair Queuing, 190
Westwood+, 164
Window-based cc, 21, 60
Wireless links, 37
WRED, 137

XCP, 174–175

APPENDIX A

Ingrid Hsieh-Yee

Professor Emerita
Dept. of Library and Information Science
Catholic University of America
Washington, D.C. 20064
E-mail: hsiehyee@cua.edu

Education

Ph.D. 1990. Library and Information Studies, University of Wisconsin-Madison
Minors: Sociology and Psychology

M.A. 1984. Library and Information Studies, University of Wisconsin-Madison.

M.A. 1983. Comparative Literature, University of Wisconsin-Madison.

B.A. 1980. Foreign Languages and Literature, National Taiwan University.

Work Experience

Professor, School/Dept. of Library and Information Science, Catholic University of America, 1990-2022 (Assistant Professor, 1990-1996; Associate Professor, 1997-2004; Professor 2004-2022; Professor Emerita 2022)

Co-Chair, Dept. of Library and Information Science, Catholic University of America, June 2015-August 2016.

Acting Dean, School of Library and Information Science, Catholic University of America, January 2010-June 2012.

Cataloger, Dept. of Legislative Reference Library, Annapolis, Maryland, 1989-1990.

Lecturer, School of Library and Information Studies, University of Wisconsin-Madison, 1988.

Teaching Assistant, School of Library and Information Studies, University of Wisconsin-Madison, 1986-1988.

Cataloger, Health Sciences Library, University of Wisconsin-Madison, 1984-1986.

Areas of Teaching and Research Interests

Information Organization and Access; Metadata; Cataloging & Classification; Information Architecture; Information Retrieval; Digital Collections; Social Media; Scholarly

Communication; Information Behavior; Health Informatics; Human Computer Interaction; Usability Studies

Grants & Honors

Cultural Heritage Information Management Project. IMLS grant. Amount: $498,741. Period: Aug. 2012 to July 2015. Co-PI with Dr. Youngok Choi.

D.C. Health Information Technology (HIT4): Building Capacity & Providing Access in Our Nation's Capital. Dept. of Labor H2B Training Grant. Grant amount: $4,175,500. Grant period: Nov. 2011 to Dec. 2015. Partner with the Metropolitan School of Professional Studies of the Catholic University of America, Children's National Medical Center, D.C. Department of Employment Services, Holy Cross Hospital, Howard University, Center for Urban Progress, Providence Hospital, and Sibley Memorial Hospital.

Capital Health Careers Project. Department of Labor Healthcare Sector and Other High Growth and Emerging Industries Grant. Grant amount: $4,953,999. Grant period: March 2010 – February 2013. Awarded to a group of healthcare organizations and educational institutions in Washington, D.C. Providence Health Foundation of Providence Hospital (Lead institution). Part of the grant supported the development of a Master's degree program in Information Technology with a concentration in Health Information Technology offered by the School of Library and Information Science.

The Washington D.C. School Librarians Project. IMLS grant. Grant amount: $412,660. Grant period: Aug. 2007 – June 2011. The School partnered with the District of Columbia Public Schools (DCPS) and the District of Columbia Library Association to educate and mentor school media specialists for the DCPS system. PI, Jan. 2010 to June 2011.

SIG Member of the Year, American Society for Information Science and Technology (2009).

Most Outstanding Paper of *OCLC Systems & Services* (2001).

ALISE Research Grant (2001).

Most Outstanding Paper of *OCLC Systems & Services* (2000).

Research Grant from ERIC (1999-2000).

Best Research Paper Award; Association for Library and Information Science Education (1998).

Research Grants, Catholic University of America. 1991, 1992, 1993, 1996, 1998, 1999, 2004, 2005, 2006, 2007, 2013-14.

Cooperative Faculty Research Grant, Consortium of Universities in the Washington Metropolitan Area (1993-1994).

Cooperative Research Grant, Council on Library Resources (1993-1994).

Journal of the American Society for Information Science Best Paper Award (1993).

ASIS/ISI Information Science Doctoral Dissertation Scholarship (1989).

HEA Title IIB Fellowship (Dept. of Education) (1989)

Chinese-American Librarians Association Scholarship (1987).

Beta Phi Mu (1985).

Vilas Fellowship, University of Wisconsin-Madison. 1984

Publications

Hsieh-Yee, I. (2021). Can We Trust Social Media? *Internet Reference Services Quarterly*, 1-15. Published online: 28 July, 2021. 10.1080/10875301.2021.1947433.

Bailey, T., & Hsieh-Yee, I. (2020). Combating the Sharing of False Information: History, Framework & Literacy Strategies, *Internet Reference Services Quarterly*, 24 (1-2), 9-30.

Choi, Y., and Hsieh-Yee, I. (2010). Finding Images in an OPAC: Analysis of User Queries, Subject Headings, and Description Notes. *Canadian Journal of Information and Library Science,* 34(3): 271 – 295.

Hsieh-Yee, I. (2008). Educating Cataloging Professionals in a Changing Information Environment. *Journal of Education for Library and Information Science,* 46(2): 93-106.

Vellucci, S. L., Hsieh-Yee, I., and Moen, W.E. (2007). The Metadata Education and Research Information Commons (MERIC): A Collaborative Teaching and Research Initiative. *Education for Information,* 25(3&4): 169-178.

NISO Framework for Guidelines for Building Good Digital Collections. 3rd ed. Baltimore, MD: National Information Standards Organization, 2007. Also available online: http://www.niso.org/framework/framework3.pdf (NISO Working Group members: Priscilla Caplan (chair), Grace Agnew, Murtha Baca, Tony Gill, Carl Fleischhauer, Ingrid Hsieh-Yee, Jill Koelling, and Christie Stephenson.)

Choi, Y., Hsieh-Yee, I., and Kules, B. (2007). Retrieval Effectiveness of TOC and LCSH. *Proceedings of the Joint Conference on Digital Libraries*, pp. 233-234.

Vellucci, S., and Hsieh-Yee, I. (2007). They Didn't Teach Me That in Library School! Building

a Digital Teaching Commons to Enhance Metadata Teaching, Learning and Research. *Proceedings of the National Conference of the Association of College and Research Libraries, Baltimore, MD*, pp. 26-31.

Mitchell, Vanessa, and Ingrid Hsieh-Yee. (2007). Converting Ulrich's Subject Headings to FAST Headings: A Feasibility Study. *Cataloging & Classification Quarterly,* 45(1): 59-85.

Hsieh-Yee, I., Tang, R., and Zhang, S. (2007). User Perceptions of a Federated Search System. *IEEE Technical Committee on Digital Libraries Bulletin,* Summer 3(2) (URL = http://www.ieee-tcdl.org).

Tang, R., Hsieh-Yee, I., and Zhang, S. (2007). User Perceptions of MetaLib Combined Search: An Investigation of How Users Make Sense of Federated Searching." *Internet Reference Services Quarterly,* 12(12): 211-236.

Hsieh-Yee, I., Tang, R., and Zhang, S. (2006). User Perceptions of a Federated Search System. *Proceedings of the Joint Conference on Digital Libraries*, *June 11-15, 2006, Chapel Hill*, p. 338.

Hsieh-Yee, I. (2006). *Organizing Audiovisual and Electronic Resources for Access: A Cataloging Guide*. 2nd ed. Westport, Conn.: Libraries Unlimited.

NISO A Framework of Guidance for Building Good Digital Collections. 2nd ed. Bethesda, MD: National Information Standards Organization, 2004. Framework Advisory Group: Grace Agnew, Liz Bishoff, Priscilla Caplan (Chair), Rebecca Gunther and Ingrid Hsieh-Yee.

Hsieh-Yee, I. (2004). Cataloging and Metadata Education in North American LIS Programs. *Library Resources & Technical Services,* 48(1): 59-68.

Hsieh-Yee, I. (2004). Cataloging and Metadata Education. In Gary E. Gorman (Ed.), *International Yearbook of Library and Information Management 2003: Metadata Applications and Management*, (pp.204-234). London: Facet Publishing.

Yee, P. L., Hsieh-Yee, I., Pierce, G.R., Grome, R., and Schantz, L. (2004). Self-Evaluative Intrusive Thoughts Impede Successful Searching on the Internet. *Computers in Human Behavior,* 20(1): 85-101.

Hsieh-Yee, I. (2003). Cataloging and Metadata Education: A Proposal for Preparing Cataloging Professionals of the 21st Century. A report submitted to the ALCTS-Education Task Force in response to Action Item 5.1 of the *Bibliographic control of Web Resources: A Library of Congress Action Plan*. Approved by the Association for Library Collections and Technical Services. Web version available since April 2003 at http://lcweb.loc.gov/catdir/bibcontrol/CatalogingandMetadataEducation.pdf.

Hsieh-Yee, I. (2002). Cataloging and Metadata Education: Asserting a Central Role in Information Organization. *Cataloging & Classification Quarterly* 34(½): 203-222.

Hsieh-Yee, I., and Smith, M. (2001). The CORC Experience: Survey of Founding Libraries, Part I. *OCLC Systems & Services,* 17: 133-140. (Received "The Most Outstanding Paper of OCLC Systems & Services in 2001" award.)

Hsieh-Yee, I., and Smith, M. (2001). The CORC Experience: Survey of Founding Libraries, Part II, Automated Tools and Usage. *OCLC Systems & Services,* 17: 166-177. (Received "The Most Outstanding Paper of OCLC Systems & Services in 2001" award.)

Hsieh-Yee, I. (2001). ERIC User Services: Changes and Evaluation for the Future. *Government Information Quarterly,* 18: 31-42.

Hsieh-Yee, Ingrid. (2001). Research on Web Search Behavior. *Library and Information Science Research,* 23: 167-185.

Logan, E., and Hsieh-Yee, I. (2001). Library and Information Science Education in the Nineties. *Annual Review of Information Science and Technology,* 35: 425-477.

Hsieh-Yee, I. (Ed.) (2001). *Library and Information Science Research*, 23 (2). A special issue in honor of the retirement of Douglas L. Zweizig.

Hsieh-Yee, I. (2000). *ERIC User Services: Evaluation in a Decentralized Environment*. Washington, D.C.: Dept. of Education.

Hsieh-Yee, Ingrid. (2000). *Organizing Audiovisual and Electronic Resources for Access: A Cataloging Guide*. Littleton, CO: Libraries Unlimited.

Hsieh-Yee, I. (2000). Organizing Internet Resources: Teaching Cataloging Standards and Beyond. *OCLC Systems & Services,* 16: 130-143. (Received "The Most Outstanding Paper of OCLC Systems & Services in 2000" award.)

Hsieh-Yee, I. (1998). The Retrieval Power of Selected Search Engines: How Well Do They Address General Reference Questions and Subject Questions? *Reference Librarian,* 60: 27-47.

Hsieh-Yee, I. (1998). Search Tactics of Web Users in Searching for Texts, Graphics, Known Items and Subjects: A Search Simulation Study. *Reference Librarian,* 60: 61-85. (Received the 1997 Best ALISE Research Paper Award.)

Hsieh-Yee, I. (1997). Access to OCLC and Internet Resources: LIS Educators' Views and Teaching Practices. *RQ,* 36: 569-86.

Hsieh-Yee, I. (1997). Teaching Online and CD-ROM Resources: LIS Educators' Views and Practices. *Journal of Education for Library and Information Science,* 38: 14-34.

Hsieh-Yee, I. (1996). The Cataloging Practices of Special Libraries and Their Relationship with OCLC. *Special Libraries,* 87: 10-20.

Hsieh-Yee, I. (1996). Modifying Cataloging Practice and OCLC Infrastructure for Effective Organization of Internet Resources. In *Proceedings of the OCLC Internet Cataloging Colloquium*. [Online]. Available: http://www.oclc.org/oclc/man/colloq/hsieh.htm

Hsieh-Yee, I. (1996). Student Use of Online Catalogs and Other Information Channels. *College & Research Libraries,* 57: 161-175.

Hsieh-Yee, I. (1995). Ten entries in James S. C. Hu (Ed.), *Encyclopedia of Library & Information Science*, 913, 1028-29, 1036, 1037, 1145-46, 1514, 1575, 1763-64, 2216-27, 2378-79. Taipei, Taiwan: Sino-American Publishing. (Topics include "Advanced Technology/Libraries," "Information Ethics," "Instruction on Cataloging and Classification," "Instruction on Reference Services.")

Hsieh-Yee, I. (1993). Effects of Search Experience and Subject Knowledge on Online Search Behavior: Measuring the Search Tactics of Novice and Experienced Searchers. *Journal of the American Society for Information Science,* 44: 161-174. (Received the 1993 Best JASIS Paper Award.)

Works in Progress

Hsieh-Yee, I. Actionable lessons for debunking false information: A comparative analysis of literacy lessons in communication and library and information science

Hsieh-Yee, I. Representing diverse voices and communities in cataloging and indexing: The evolution of controlled vocabularies

Hsieh-Yee, I. Information Organization and Retrieval in Practice (a monograph scheduled for completion and publication in 2023)

Presentations

Bailey, T. & Hsieh-Yee, I. (February 2020). The Phenomena of Sharing Misinformation and the Need for Information Literacy. Presented at the 2020 Bridging the Spectrum Symposium, Washington, D.C., February 14, 2020.

Hsieh-Yee, I. and Fragan-Fly, J. (May 2018) Trends, Design & Strategies for Digital Scholarship Services. Presented at the 2018 Maryland/Delaware Library Association Conference, Cambridge, MD.

Hsieh-Yee, I. (February, 2018) Research Data Management: What It Takes to Succeed. Presented at the 10th Bridging the Spectrum Symposium, Washington, D.C.

Hsieh-Yee, I. (February, 2017) *Research Data Management: New Competencies and Opportunities for Information Professionals*. Presented at the 9th Bridging the Spectrum Symposium, Washington, D.C.

Hsieh-Yee, I. and Lawton, P. (February, 2017) *Enhancing Catholic Portal Searches with User Terms and LCSH*. Presented at the 9th Bridging the Spectrum Symposium, Washington, D.C.

Hsieh-Yee, I. (2016, October) *Visualizing Data for Information*. Presented at the 2016 Virginia Library Association Conference, Hot Springs, VA.

Hsieh-Yee, I. (2016, August) *Religious Materials Toolbox for Archivists: Solutions to Problems Facing the Profession*. Presented at Archives * Records 2016, Atlanta, GA.

Hsieh-Yee, I. and Lawton, P. (2016, March) *Enhancing Retrieval of Catholic Materials with LCSH Knowledge Structure*. Presented at the 2016 Catholic Library Association Conference, San Diego, CA.

Fagan-Fry, J. and Hsieh-Yee, I. (2016, February) *Approaches to Digital Scholarship at Top Universities around the World: Scholarly Publishing in the Digital Age.* Presented at the 8th Bridging the Spectrum Symposium, Washington, D.C.

Hsieh-Yee and Fagan-Fry, J. (2016, January) *Innovative Services for Digital Scholarship at Top 100 Research Libraries of the World*. Poster presented at the 2016 Annual Conference of the Association for Library and Information Science Education, Boston, Mass.

Hsieh-Yee, I. and Lawton, P. (2015, June). *Crowdsourcing terms for CRRA portal themes.* Poster presented at the third CRRA symposium and annual meeting, Bringing the created toward the Creator: Liturgical art and design since Vatican II. Catholic Theological Union, Chicago, Illinois.

Hsieh-Yee, I. and Lawton, P. (2015, February). *Crowdsourcing terms for thematic exploration in the Catholic Portal.* Poster presented at the 7th Annual Bridging the Spectrum Symposium, Washington, D.C.

Hsieh-Yee, I., James, R., and Fagan-Fry, J. (2015, February). *Support for digital scholarship at top university libraries of the world*. Poster presented at the 7th Annual Bridging the Spectrum Symposium, Washington, D.C.

Hsieh-Yee, I., Zhang, S., Lin, K., and Cherry, S. (2015, February). *Thus said the end users: Summon experience and support for research workflows*. Poster presented at the 7th Annual Bridging the Spectrum Symposium, Washington, D.C.

Yontz, E., Hsieh-Yee, I., & Houston, S. (2015, February). *Healthy Heroes Summer Reading Club: Developing healthy youth at public libraries.* 11th Annual Jean Mills Health Symposium, Greenville, North Carolina.

Yontz, E., Hsieh-Yee, I., and Houston, S. (2015, January). *Healthy youth and libraries: A pilot study.* Association for Library & Information Science Education (ALISE) Annual Conference, Chicago, Illinois.

Hsieh-Yee, I. (2014, May). *Linking CRRA resources to portal themes via authority files*. Presented at the Catholic Research Resources Alliance 2014 Membership Meeting, Marquette, WI.

Hsieh-Yee, I. (2014, April). *Enhancing subject access to CRRA resources*. Presented at the 2014 Catholic Library Association Conference, Pittsburgh, PA.

Hsieh-Yee, I. (2014, January). *Health Information Technology Program: Educational entrepreneurship in action*. Presented at the 2014 annual Conference of the Association for Library and Information Science Education, Philadelphia, PA

Hsieh-Yee, I., Zhang, S., Lin, K., and Cherry, S. (2014, January). *Discovering information through Summon: An analysis of user search strategies and search success*. Paper presented at the 6th Bridging the Spectrum Symposium, Washington, D.C.

Hsieh-Yee, I. (2012, December). *National Digital Stewardship Alliance and SLIS at CUA: An Educational Partnership*. Paper presented at Best Practices Exchange: Acquiring, Preserving, and Providing Access to Government Information in the Digital Era, Annapolis, MD

Choi, Y. and Hsieh-Yee, I. (2010, January). *Finding Images in an OPAC: Analysis of User Queries, Subject Headings, and Description Notes*. Paper presented at 2nd Annual Bridging the Spectrum Symposium, Catholic University of America, Washington, D.C.

Hsieh-Yee, I. and Coogan, J. (2010, January). *Google Scholar vs. Academic Search Premier: What Libraries and Searchers Need to Know*. Paper presented at 2nd Annual Bridging the Spectrum Symposium, Catholic University of America, Washington, D.C.

Hsieh-Yee, I. (2009, November). *Information Science Education: An LIS School's Perspective*. Paper presented at Annual Meeting of the American Society for Information Science and Technology, Vancouver, British Columbia, Canada.

Hsieh-Yee, I., Menard, E., Ya-Ning Chen, A., Shu-Jiun Chen, S., Kalfatovic, M. R., Wisser. K. M. (2009, November). *Information Organization in Libraries, Archives and Museums: Converging Practices and Collaboration Opportunities*. Presented at Annual Meeting of the American Society for Information Science and Technology, Vancouver, British Columbia, Canada. (Organizer and moderator of this panel.)

Hsieh-Yee, I. and Coogan, J. (2009, July). *Catching up to Google Scholar: The Retrieval Power of Academic Search Premier and Google Scholar*. Poster presented at American Library Association Conference, Chicago, Illinois.

Hsieh-Yee, I., with the CUA Scholarly Communications Project Team. (2009, January). *Digital Scholarship@CUA: Developing an Institutional Repository for CUA*. Poster presented at 1st Annual Bridging the Spectrum Symposium, Catholic University of America, Washington, D.C.

Wise, M., Cylke, K., and Hsieh-Yee, I. **(**2009, January). *Digital Talking Books: Meeting the Needs of the Blind and the Handicapped*. Paper presented at the Bridging the Spectrum Symposium, Catholic University of America, Washington, D.C.

Hsieh-Yee, I. (2009, January). *User Expectations of MERIC*. Presented at the Information Organization Competencies for the 21st Century Discussion Session of the 2009 Conference of the Association for Library and Information Science Education, Denver, Colorado.

Choi, Y., and Hsieh-Yee, I. (2008,November). *Subject Access for Images in an OPAC.* Annual Meeting of the American Society for Information Science and Technology, Columbus, Ohio. (Also co-organized a panel on Retrieving and Using Visual Resources: Challenges and Opportunities for Research and Education.)

Hsieh-Yee, I. (2008, June). *Educating Cataloging Professionals in a Changing Information Environment*. National Taiwan University, Taipei, Taiwan.

Vellucci, S. L., Moen, W.E., Hsieh-Yee, I., Marson, B., and Wisser, K. (2008, January) *Building a Metadata Education and Research Community through MERIC (Metadata Education and Research Information Commons): Demo and Stakeholder Input*. A panel presented at the 2008 Conference of the Association for Library and Information Science Education, Philadelphia, Pennsylvania.

Hsieh-Yee, I., Choi,Y. and Kules, B. (2007, October). *Searching for Books and Images in OPAC: Effects of LCSH, TOC and Subject Domains.* A poster presented at the American Society for Information Science and Technology Annual Meeting, Milwaukee, Wisconsin.

Hsieh-Yee, I. and Coogan, J. (2007, August) *Google Scholar vs. Academic Search Premier: A Comparative Analysis.* Presented to the Faculty and Staff of the University of the District

of Columbia.

Hsieh-Yee, I. and Coogan, J. (2007, June). *Google Scholar vs. Academic Search Premier: A Comparative Analysis*. Presented to the Washington Research Library Consortium Community, Catholic University of America, Washington, D.C.

Hsieh-Yee, I., Choi, Y., and Kules, B.. (2007, June). *What Users Need for Subject Access: Table of Contents or Subject Headings?* A poster presented at the 2007 American Library Association Annual Conference, Washington, D.C., June 2007.

Choi, Y., Hsieh-Yee, I., and Kules, B. (2007, June). *Retrieval Effectiveness of TOC and LCSH.* A paper presented at the Joint Conference on Digital Libraries 2007, Vancouver, Canada.

Vellucci, S. L., Hsieh-Yee, I., and Moen, W.E. (2007, May). *If We Build It, Will They Come? Building a Community of Practice for Metadata Stakeholders.* A poster presented at the Rutgers University Research Day, Bridgeton, New Jersey.

Hsieh-Yee, I. (2007, May). *Federated Searching: User Experience & Perceptions.* International Conference on Information Organization & Retrieval, National Taiwan University, Taipei, Taiwan.

Hsieh-Yee, I. (2007, May). *Search Performance of Google Scholar and Academic Search Premier.* International Conference on Information Organization & Retrieval, National Taiwan University, Taipei, Taiwan.

Hsieh-Yee, I. (2007, May) *MERIC: Building a Digital Commons for Metadata Education & Research.* International Conference on Information Organization & Retrieval, National Taiwan University, Taipei, Taiwan.

Hsieh-Yee, I., and Coogan, J. (2007, March/April). *A Comparative Analysis of Google Scholar and Academic Search Premier.* Poster presented at the Association of College & Research Libraries 13th National Conference, Baltimore, Maryland.

Vellucci, S. L. and Hsieh-Yee, I. (2007, March/April) *They Didn't Teach Me That in Library School! Building a Digital Teaching Commons to Enhance Metadata Teaching, Learning and Research*. On-site presentation and Webcast by Elluminate. A contributed paper presented at the Association of College & Research Libraries 13th National Conference, Baltimore, Maryland. The acceptance rate for contributed paper was 20%. This paper was one of 10 conference papers chosen for live webcast during the conference.

Moen, W., Hsieh-Yee, I. and Vellucci, S.L. (2007, January) *A DSpace Foundation for a Teaching & Research Commons: The Metadata Education and Research Information Commons.* A poster session presented at the Open Repositories Conference 2007, San Antonio, Texas.

Tang, R., Hsieh-Yee, I., and Zhang, S. (2006, November) *User Perception of MetaLib Combined Search*. Paper presented at the Annual Meeting of the American Society for Information Science and Technology, Austin, Texas, Nov. 2006.

Hsieh-Yee, I. (2006, November). *Federated Searching: User Perceptions, System Design, and Library Instructions.* Paper presented at the Annual Meeting of the American Society for Information Science and Technology, Austin, Texas. (Panel organizer, moderator, presenter).

Hsieh-Yee, I. (2006, November). *Building a Digital Teaching Commons to Enhance Teaching and Learning: The MERIC Experience and Challenges.* Paper presented at the Annual Meeting of the American Society for Information Science and Technology, Austin, Texas. (Panel organizer, moderator, presenter)

Hsieh-Yee, I. (2006, September). *Search Performance of Google Scholar and Academic Search Premier.* Paper presented at the ERIC Publishers Meeting, Washington, D.C.

Hsieh-Yee, I., Zhang, S., and Rong Tang, R. (2006, June). *User Perceptions of a Federated Search System*. Poster presented at Joint Conference on Digital Libraries, Chapel Hill, North Carolina.

Hsieh-Yee, I. and Zhang, S. (2006, June). *Preparing Users for Federated Search: Implications of a MetaLib User Perceptions Study.* Paper presented at the 2006 Ex Libris User Groups of North America Conference, Knoxville, Tennessee.

Hsieh-Yee, I. (2006, January). *MERIC Organizations and Navigation*. Paper presented at the 2006 ALISE Annual Conference, San Antonio, Texas.

Hsieh-Yee, I. (2006, January). *Metadata and Cataloging Education: Recommended Competencies.* Paper presented at the 2006 ALISE Annual Conference, San Antonio, Texas.

Hsieh-Yee, I. (2005, November). *Digital Library Evaluation: Progress & Next Steps.* Presentation at the Annual Meeting of the American Society for Information Science & Technology, Charlotte, North Carolina.

Hsieh-Yee, I. (2005, August). *Providing Access to Digital Content: Issues for DL Managers.* Presentation at MDK12 Digital Library Steering Committee Meeting, Columbia, Maryland.

Hsieh-Yee, I. (2005, April). *Enhancing Teaching and Learning: The Role of School Library Media Specialists.* Presentation at Meeting of the Baltimore County Public School System School Media Specialists, Baltimore, Maryland.

Hsieh-Yee, I. (2005, January). *Subject Access and Users: Insights & Inspirations from Marcia J. Bates.* Paper presented at the Historical Perspectives SIG, 2005 Conference of the Association for Library and Information Science Education, Boston, Massachusetts.

Hsieh-Yee, I. (2005, January). *Electronic Resource Management: Practice, Employer Expectations, & CE Interests.* Paper presented at Technical Services Education SIG, 2005 Conference of the Association for Library and Information Science Education, Boston, Massachusetts.

Hsieh-Yee, I. (2004, October). *Library Professionals for the Digital Age: Competencies & Preparation.* Paper presented at Bibliographic Access Management Team meeting, Library of Congress, Washington, D.C.

Hsieh-Yee, I. (2004, January). *Cataloging and metadata expertise for the digital era.* Presented at Preparing 21st Century Cataloging and Metadata Professionals: A Workshop for Educators and Trainers, San Diego and sponsored by ALCTS, ALISE, LC, and OCLC.

Hsieh-Yee, I. (2004, January). *Educating catalogers for the digital era.* Paper presented at the Technical Services SIG, 2004 Conference of the Association for Library and Information Science Education, San Diego.

Hsieh-Yee, I. (2003, July). *Cataloging Education for the 21st Century.* A presentation at the Library of Congress, Washington, D.C.

Hsieh-Yee, I. (2002, January) *Metadata Education and Research Priorities: A Delphi Study of Metadata Experts.* Presentation at the 2002 Conference of the Association for Library and Information Science Education, New Orleans.

Hsieh-Yee, I. (2001, November). *A Delphi Study of Metadata: Preliminary Findings.* Poster session at the 2001 Annual Meeting of the American Society for Information Science & Technology, Washington, D.C.

Hsieh-Yee, I. (2001, June). *Resources on Asian American Children: Analysis of Retrieval by Search Engines and WorldCat.* Presentation at the National Conference on Asian Pacific American Librarians, San Francisco.

Hsieh-Yee, I. (2001, January). *Delphi Study on Metadata: Project Design.* Presentation at Research Awards Session, Association for Library & Information Science Education, Washington, D.C.

Hsieh-Yee, I. (2000, May). *Web Search Behavior Research: Progress and Implications.* Presentation at the Symposium on Evaluating Library and Information Science Research, University of Wisconsin-Madison, Madison, Wisconsin.

Hsieh-Yee, I. (2000, March). *ERIC User Services: Evaluation in a Decentralized Environment.* Presentation at the National ERIC Joint Directors/Technical Meeting, Arlington, Virginia.

Hsieh-Yee, I. (2000, January). *Enhancing Learning with Web Technology.* Presentation at Faculty Conversations, Catholic University of America, Washington, D.C.

Hsieh-Yee, I. (2000, January). *From Surrogates to Objects: CUA's Approaches to Organizing Electronic Resources.* Paper presentation at the Annual Conference of the Association for Library and Information Science Education, San Antonio, Texas.

Yee, P., and Hsieh-Yee, I. (1997, November). *Individual Differences in Search Behavior on the WWW.* A poster session presented at the 38th Annual Meeting of the Psychonomic Society, Philadelphia, Pennsylvania.

Hsieh-Yee, I. (1997, April). *Research + Marketing + Preparation = Job!* Presented at the "Workshop on Resume and Interview Techniques," Special Libraries Association, Student Chapter, Catholic University of America, Washington, D.C.

Hsieh-Yee, I. (1997, February). *Creating CyberCatalogers: Education and Training.* Presentation at ALA's Midwinter Meeting, Washington, D.C.

Hsieh-Yee, I. (1997, February). *Search Tactics of Web Users in Searching for Texts, Graphics, Known Items and Subjects: A Search Simulation Study.* Presented at the Conference of the Association for Library and Information Science Education, Washington, D.C.

Hsieh-Yee, Ingrid. "Beginning Your Special Library/Information Center Career." Presented at SLA's "Career Day," Jan. 11, 1997, Catholic University of America.

Hsieh-Yee, I. (1996, September). *The Roles of Library and Information Scientists in Managing Electronic Information.* Presentation at Hamilton College, Clinton, New York.

Hsieh-Yee, I. (1996, May). *The Future of Cataloging as a Profession.* Presented at "The Cataloging Forum, Library of Congress, Washington, D.C.

Hsieh-Yee, I. (1994, October). *The Impact of the Internet on OPACs.* Presented at the Third Workshop on User Interfaces for OPACs, Library of Congress, Washington, D.C.

Reports

Hsieh-Yee, I., with Knowledge Management Competencies and Performance Action Group of the Federal Knowledge Management Initiative. "From Knowledge Management Competencies to Improved Organizational Performance." April 9, 2009.

Hsieh-Yee, I., with Knowledge Practices Action Group of the Federal Knowledge Management Initiative. "KM Practice in Government Agencies: Findings and Recommendations." April 9, 2009.

Hsieh-Yee, I. "Delphi Study on Metadata." 2001. Three quarterly reports submitted to the Association for Library and Information Science Education.

Hsieh-Yee, I. "College Students' Information Channels: Patterns of Use and Possible Factors in Channel Selection." 1995. Submitted to the Catholic University of America.

Hsieh-Yee, I. "The Information-Seeking Patterns of Scholars and Their Use of an Online Information System." 1994. Submitted to the Council on Library Resources.

Book Reviews

Review of *The Measurement and Evaluation of Library Services*, by Sharon L. Baker and F. Wilfrid Lancaster. *Information Processing and Management* 30 (1994): 450-52.

Review of *Subject Access to Films and Videos, by Sheila S. Intner and William E. Studwell; and Cataloging Unpublished Nonprint Materials*, by Verna Urbanski with Bao Chu Chang and Bernard L. Karon. *Information Processing and Management* 30 (1994): 449-50.

Review of *Automated Information Retrieval in Libraries: A Management Handbook*, by Vicki Anders. *Journal of Library and Information Science* 19 (1993): 98-100.

Review of *Full Text Databases*, by Carol Tenopir and Jung Soon Ro. *Information Processing and Management* 28 (1992): 667-68.

Review of *Descriptive Cataloging for the AACR2R And USMARC: A How-to-Do It Workbook*, by Larry Millsap and Terry Ellen Ferl. *Information Processing and Management* 28 (1992): 809-11.

Review of *MARC Manual: Understanding and Using MARC Records*, by Deborah J. Byrne. *Information Processing and Management* 28 (1992): 537-38.

Service

Professional Associations and Societies

- Library of Congress. RDA Training Program for the Profession. Co-authored with Tim Carlton. 2013-2014.
- 2014 Digital Preservation Outreach & Education Survey. Contributed to the design of the survey, 2014.
- National Digital Stewardship Alliance. Outreach Committee. 2011-2014.
- National Digital Stewardship Residency Program. Advisory Group, 2012-2013.
- FEDLINK Health Information Technology Advisory Council, 2011-2015.

- 2012 Joint Conference on Digital Libraries. Program Planning Committee, Pre-Conference Proposals Review Committee, 2012
- Catholic Research Resources Alliance. Five-Year Strategic Plan Task Force, 2011-2012
- Institute of Museum and Library Services. Grant reviewer. 2004, 2005, 2010.

- Association for Library and Information Science Education.
  * ALISE Bodan Wynar Research Paper Award Committee, 2015, 2016, 2017
  * ALISE Eugene Garfield Dissertation Award Competition, Jury, 2013, 2014
  * ALISE Research Grant Competition Committee. Chair, 2012
  * Pratt-Severn Faculty Innovation Award. Chair, 2009, 2010
  * ALISE Doctoral Poster Jury, 2012
  * "Information Organization Competencies for the 21st Century" Discussion session leader. 2009 Conference of the Association for Library and Information Science Education.
  * Assisted Technical Services SIG Convener in organizing a program, ""Building a Metadata Education and Research Community through MERIC (Metadata Education and Research Information Commons): Demo and Stakeholder Input" for the 2008 ALISE conference.
  * Association for Library Collections and Technical Services/Association for Library and Information Science Education (ALCTS/ALISE) Metadata Education and Research Information Center (MERIC) Advisory Board, Co-Chair (with Sherry Vellucci), 2005-2007. Chair, 2008-2009 (leading the effort to build MERIC, a repository and collaborative space for metadata educators, practitioners, and researchers)
  * Technical Services SIG, Convener, 2004-2005. Organized a program on "Electronic Resources Management: Current Practices, Employer Expectations, and Teaching Strategies" for the 2005 conference in Boston, Massachusetts.
  * Technical Services SIG, Convener, 2003-2004. Organized a program on "Organizing Information with Metadata: Desired Competencies and Teaching Innovations" for the 2004 conference.
  * Technical Services SIG, Convener, 1999-2000. Organized a program on "Teaching the Organization of Electronic Resources" for the 2000 conference.
  * Curriculum SIG, Co-convener (with Sibyl Moses), 1996-97. Organized a program on "Government Information Policy" for the 1997 conference.

- American Society for Information Science & Technology.
  * Reviewer, Conference program panel submissions and poster submissions, 2005, 2006, 2007, 2009, 2011, 2012, 2013, 2014, 2015, 2016, 2017
  * Nomination Committee, 2009-2011
  * Information Science Education Special Interest Group. American Society for Information Science and Technology. Chair-Elect, 2007-2008. Chair 2008-2009.
  * Committee on Information Science Education. 1999-2006.
  * Committee on Information Science Education. Organizing Committee for an orientation program for students at ASIS annual meetings, 1999-2001
  * Committee on Information Science Education. Sub-committee on Student Welfare (focusing on issues related to master's education), 1998-2001

* SIG ED. Organizing Committee for the "Seminar on Research and Career Development" for junior researchers. 1995-96 (chair), 1997-2001
* ISI Doctoral Dissertation Proposal Scholarship Jury, 1997; 2001, 2002
* Pratt-Severn Best Student Research Paper Award Jury. Chair. 1997
* 1998 Midyear Meeting (referee of contributed papers), 1997
* Organizer and moderator of the ASIS Doctoral Forum and the Doctoral Research Seminar 1994-1995
* SIG Human Computer Interaction. Chair-Elect, Chair, 1993-1995
* Doctoral Forum Award Jury, 1995
* Best Student Paper Award Jury, 1995

- American Library Association.
  * Committee on Accreditation, External Review Panelist, 2009- (site visiting team 2013-2014; site visiting team 2016-2017)
  * Association for Library Collections and Technical Services Task Force on Competencies and Education for a Career in Cataloging, member, 2008-2009
  * Facilitator for "What They Don't Teach in Library School: Competencies, Education and Employer Expectations for a Career in Cataloging," an Association for Library Collections and Technical Services Preconference, June 22, 2007 in Washington, D.C. Also a local liaison for bringing this program to the Catholic University of America.
  * Facilitator for a discussion on "Effect of Electronic Resources on Technical Services" at ALA's Midwinter Meeting held in Feb. 1997 in Washington, D.C.
  * International Relations Committee, Subcommittee Task Force for IFLA and China, 1994-1997

- Virginia Association of School Librarians. Scholarships and Awards Committee. 2010-2012

- Federal Knowledge Management Initiative, Knowledge Management Practices Action Group. Member. 2009 (leading the effort to build a knowledge management repository)
- Federal Knowledge Management Initiative, Knowledge Management Competencies & Learning Action Group. Member. 2009 (developing an action plan for helping government knowledge workers and government agencies to develop knowledge management competencies)
- National Center for Education Statistics. Technical Review Panel. 2008.
- External evaluator for a case of promotion to full professorship. University of Tennessee. 2008.
- National Information Standards Organization (NISO). Advisory Board, Revision of "IMLS Framework of Guidance for Building Good Digital Collections," 2004, 2007.
- Library of Congress, Bibliographic Control of Web Resources: A Library of Congress Action Plan. Principal Investigator of Action Item 5.1, focusing on cataloging and metadata education for students and new librarians, 2002-2003. (worked with the Association for Library Collections and Technical Services, Education Task Force)

- Chinese American Librarians Association
  * Chinese American Librarians Association Outstanding Library Leadership Award in Memory of Dr. Margaret Chang Fung, Award Committee, 2016-2017
  * Achievement Award Jury, 2000-2001
  * CALA Goal 2000 Task Force, 1997
  * Scholarship Committee, 1995, 1996-1997 (chair)
  * Board of Directors, 1994-1997
  * Publication Committee, 1993-1995
  * International Relations Committee, 1993-1996

- Sailor[SM] Assessment Advisory Group (An impact study of Sailor, Maryland's Public Information Network), 1995

- Editorial boards
  Journal of Library & Information Science. Editorial Board, 2012-
  Chinese American Librarians Association, *Occasional Papers Series*. Editorial Board, 2009-2016.
  Library Quarterly. Editorial Board, 2003-2008
  Bulletin of the Medical Library Association, 1994-97
  Newsletter editor for the Chinese American Librarians Association, 1989-92

- Referee for the following journals
  Information Processing and Management
  Journal of Digital Information
  Journal of Education for Library and Information Science
  Journal of Information Science
  Journal of Library & Information Science
  Journal of Library Metadata
  Journal of the American Society for Information Science & Technology
  Library and Information Science Research
  Library Quarterly

- Expert reviewer, "Digital Library" course, Evaluation module, University of North Carolina, Chapel Hill, 2007-2008.
- Expert reviewer, "Information Organization" course, University of Michigan, Ann Arbor. 2007.

Catholic University of America

- Academic Senate Committee on Appointments and Promotions, 2005-2008
- Academic Senate Committee on Committees and Rules, 2009-2012, 2019-2022
- Academic Senate representative (for School of Arts & Sciences), 2017-2020

- Academic Senate representative (for School of Library & Information Science) 2003-2012
- Academic Senate Library Committee, Interim Chair (2007), Member, 2008-2012
- Academic Senate Committee on Computing, 1995-2003

- President's Administrative Council, 2010-2012
- Deans' Council, 2010-2012
- Academic Leadership Group, 2010-2012
- Graduate Board, 2010-2012

- School of Arts & Sciences, Committee on Appointments and Promotions, 2015-2019
- School of Arts & Sciences, Academic Council, 2015-2016.
- School of Arts & Sciences, Ordinary Professor Group, 2013-
- Doctoral Dissertation Defense Committee, Chair, Dept. of Psychology, 2016, 2017, 2018, 2019, 2020, 2021, 2022
- Doctoral Dissertation Defense Committee, Chair, Dept. of Education, 2014, 2015, 2017, 2018, 2019
- Doctoral Dissertation Defense Committee, Chair, National Catholic School of Social Services, 2019, 2020, 2021
- Doctoral Dissertation Defense Committee, Chair, School of Nursing, 2006, 2008
- CUA Scholarly Communication Project Team, Member (2007), Chair, 2008-2009
- Dean Search Committee, 1992-1994, 1998-1999, 2002-2003, 2006-2007
- Fulbright Review Panel, 2006
- CUA Service Learning Advisory Board, 2001-2002
- CUA Faculty Conversations on Enhancing Teaching and Learning through Technology, Planning Group, 1999-2001
- CUA Initiative on Technology and Teaching, 1998-2001

Dept. of Library and Information Science

- Admissions Committee, 2007-2009, Chair 2010-2012, Member 2013-2015, Member 2018-2022
- Appointments and Promotions Committee, 1991-2022
- Student advisement, 1990-2022
- 4+1 Bachelor to Master program, Program Advisor, 2014-2022
- LIS Minor, Advisor, 2017-2022
- Curriculum Committee, 2020-21 (Member), 2021-2022 (Chair)
- Scholarship & Awards Committee, 2016-2018, 2020-21
- Comprehensive examination editor, 2020-21
- Symposium, Colloquium, Lecture Series Committee, May 2018-May 2020 (Chair)
- Symposium and Colloquium Committee, fall 2016-May 2018
- Bridging the Spectrum Symposium session moderator, 2018, 2019, 2020, 2021

- Organization of Information Core Course Outcomes Assessment, 2019, 2020
- The Information Professions in Society Core Course Outcomes Assessment, 2019, 2020
- Community Services Librarianship, Course of Study Review and Revision, 2020
- Organization of Information, Course of Study Review and Revision, 2019
- Organization of Information, Competency and Course Offerings Review, 2018-2019
- Accreditation presentation, Chair, June 2015-August 2016
- Interim Co-Chair, June 2015-August 2016.
- Blended/OWL Learning Committee, spring 2016-2018
- Technology Committee, fall 2016-2017
- Comprehensive examination editor, 2016-2017, reader (every year since 1990)
- LIS Advisory Board, 2015-2016 (chair); fall 2016- May 2018 (member)
- Committee on Planning and Assessment, 2015-2016 (chair)
- Senior Faculty Committee, 2014-2016.
- Accreditation Steering Committee, 2014-2016 (Chair, 2015-2016)
- Accreditation Students Standard Committee, co-chair, 2014-2016
- Accreditation Mission, Goals, and Objectives Standard Committee, co-chair, 2014-2016
- Accreditation Curriculum Standard, member 2014-2-16
- Accreditation Administration and Finance Standard, member 2014-2016
- Cultural Heritage Information Management Project (IMLS-funded), Co-PI, 2012-2015
- Cultural Heritage Information Management Forum (scheduled for June 2015), Co-Organizer, 2013-2015
- Health Information Technology Interim Review Committee, 2015 (chair)
- Health Sciences Librarianship Advisory Group, 2015- (chair)
- Comprehensive examination editors, 2013-2014, 2016-2017
- National Digital Stewardship Alliance liaison, 2011-2014
- Advisory Board, Chair 2010-2012
- Academic Honesty Committee, Chair, 2008-2012
- Blended Learning Committee, 2010-2012
- Colloquium Committee, 2010-2012
- Comprehensive Examination Administration, 2010-2012
- Cultural Heritage Information Management Advisory Committee, 2010-2012 (chair), 2013-
- Curriculum Committee, 1991-2003, 2007-2009, Chair 2010-2012, member 2013-
- Curriculum Subcommittee on Comprehensive Examination, Chair 2009-2012
- Health Information Technology Advisory Board, Chair 2010-2012. Member 2013-
- Health Sciences Advisory Committee, 2009, Chair 2010-2012. Member 2013-
- HIT Expert Forum, Chair 2012. Member 2013-2015
- Health Information Technology Student Group Advisor, 2011-2012
- State Council for Higher Education of Virginia, SLIS Representative, 2010-2012
- Symposium Planning Committee, 2010-2012
- Website Management Team, Chair, 2010-2012
- Urban School Librarianship Project (IMLS-Funded), PI, 2007-2011 (chair, 2010-11)

- Failing Grades Committee, 1995-1997 (chair), 2000-2001 (chair), 2004-2005 (chair), 2007 (chair)-2011
- Faculty Search Committee, 1994-1998, 2002-2004, 2006 (chair), Fall 2007-2009, Chair fall 2009-2012
- Recruitment Committee, Chair 2010-2012
- Strategic Planning Committee, Chair 2010-2012
- Technology Committee, 2010-2012
- Accreditation Advisory Committee, 2007-2009
- Accreditation Coordinating Committee, 2007-2009
- Accreditation Steering Committee, 2007-2009
- SLIS Advisory Group, 2007-2009
- Accreditation Curriculum Standard Committee, Co-chair, 2007-2009
- Accreditation Faculty Standard Committee, Co-chair, 2007-2009
- LSC 551 Information Organization Review Team, Co-chair, 2008-2009, 2015-2016.
- Curriculum Subcommittee on Portfolios, 2009
- LSC 555 Information Systems in Libraries and Information Centers Review Team, contributor, 2008-2009
- Redesign of LSC 730 Use and Users of Libraries and Information. 2009-
- Development of a metadata institute that was taught as LSC 715 Organization of Internet Resources in 2008. The institute is being revised and will be offered in 2010 under a new course title.
- Development of lesson plans, assignments, and evaluation rubrics for LSC 606, Cataloging and Classification, for the School's NCATE accreditation. 2008
- Howard and Mathilde Rovelstad Scholarship Committee, Chair, 2004-2007
- Assistant Dean Search Committee, Chair, Fall 2007
- Liaison to the Association for Library Collections and Technical Services to bring its preconference program, Cataloging Education and Employer Expectations, to CUA during the 2007 American Library Association Annual Meeting in Washington, D.C.
- Organizer of the colloquium presentation and reception for Tamar Sadeh of Ex Libris on PRIMO June 2007
- Practicum review and design (work with potential supervisors, such as the American Indian Museum internship description revision) 2006-
- Comprehensive examinations (edits, proctoring, and grading), 1990-
- SLIS Web site redesign: Comments and suggestions. Fall 2007
- Conducted surveys of current students and alumni in preparation for the 2005 re-accreditation, 2004-2005
- Technology Committee, 1992-1999 (chair, 1996-1998), 2002-2003 (member)
- Colloquia Committee 1997-1999, 2002-2003.
- Advisor of the CUA Student Chapter of the American Society for Information Science and Technology, 2002-2003
- Visiting Professor Search Committee, 1999, 2000, 2001
- Leader, Participation in the CORC experiment, 1999-2000
- Advisor of the Special Libraries Association Student Chapter, 1993-1999; the group was recognized for outstanding leadership by SLA in 1999.

- COA planning Committee, Task Force on Electronic Presentation of SLIS Reports (team leader) 1997-1998
- COA Planning Committee, Subcommittee on Technology 1996-1998
- NLM practicum coordinator, 1997-1998
- Computer Literacy Workshops: Assisted with the development and evaluation of the workshops, 1996-1998
- Leader, Participation in the InterCat project, 1995-1997

APPENDIX B

MARC 21 Reference Materials

Part VII: A Summary of Commonly Used MARC 21 Fields
Part VIII: A List of Other Fields Often Seen in MARC Records
Part IX: The Leader
Part X: Field 008 for Books

Part VII:

A Summary of Commonly Used MARC 21 Fields

This is a summary of the MARC 21 tags used most frequently by libraries in entering their own bibliographic records. For full listings of all MARC 21 tags, indicators, and subfield codes, see *MARC 21 Format for Bibliographic Data*.

In the explanations on these pages:

Tags -- The tags (3-digit numbers) are followed by the names of the fields they represent. In this summary, and in the *MARC 21 Format for Bibliographic Data*, if a tag can appear more than once in one bibliographic record, it is labeled repeatable (R). If it can only be used once, it is labeled non-repeatable (NR). For example, a catalog record can have several subjects, so the tags for subject added entries (6XX) are labeled repeatable (R).

Indicators -- The use of indicators is explained in fields where they are used. Indicators are one-digit numbers. Beginning with the 010 field, in every field -- following the tag -- are two character positions, one for Indicator 1 and one for Indicator 2. The indicators are not actually defined in all fields, however. And it is possible that a 2nd indicator will be used, while the 1st indicator remains undefined (or vice versa). When an indicator is undefined, the character position will be represented by the character # (for blank space).

Subfield codes -- All the data in each field (beginning with the 010 field) is divided into subfields, each of which is preceded by a delimiter-subfield code combination. The most common subfield codes used with each tag are shown. Each subfield code is preceded by the character $, signifying a delimiter. The name of the subfield follows the code.

In general, every field MUST have a subfield 'a' (**$a**). One exception that is often seen is in Field 020 (ISBN), when the ISBN information (subfield **$a**) is unavailable but the price (subfield **$c**) is known. Some subfields are repeatable. In this summary, repeatability is noted for only the more common repeatable subfields.

Examples: Examples follow the explanation for each field. For clarity, one space has been placed between the tag and the first indicator, one space has been placed between the second indicator and the first delimiter- subfield code, and one space has been inserted between the delimiter-subfield code and the subfield data.

010 Library of Congress Control Number -- (LCCN)
(NR, or Not Repeatable
Indicators undefined.

Subfield used most often:

$a -- Library of Congress control number

```
Example:    010 ##  $a ###86000988#
```

020 International Standard Book Number -- (ISBN)
(R, or Repeatable)

Indicators undefined.

Subfields used most often:

$a -- International Standard Book Number
$c -- Terms of availability (often a price)
$z -- Cancelled/invalid ISBN (R)

```
Example:    020 ##  $a 0877547637
```

040 Cataloging source -- (NR)

Indicators undefined.

Subfields used most often:

$a -- Original cataloging agency
$c -- Transcribing agency
$d -- Modifying agency (R)

```
Example:    040 ## $a DLC
                   $c DLC
                   $d gwhs
```

100 Main entry -- Personal name -- (primary author)
(NR; there can be only one main entry)

Indicator 1: Type of personal name entry element

0 -- Forename
1 -- Surname (this is the most common form)
3 -- Family name

Indicator 2 undefined.

Indicator 2 became obsolete in 1990. Older records may display 0 or 1

Subfields used most often:

$a -- Personal name
$b -- Numeration
$c -- Titles and other words associated with a name (R)
$q -- Fuller form of name
$d -- Dates associated with a name (generally, year of birth)

```
Example:    100 1# $a Gregory, Ruth W.
```

```
                    $q (Ruth Wilhelme),
                    $d 1910–
```

130 Main entry -- Uniform title -- (NR)

Indicator 1: Nonfiling characters

0-9 -- Number of nonfiling characters present (for initial articles, including spaces)

Indicator 2 undefined.

Indicator 2 became obsolete in 1990. (See 100 above.)

Subfields used most often:

$a -- Uniform title
$p -- Name of part/section of a work (R)
$l -- Language of a work
$s -- Version
$f -- Date of a work

```
Example:    130 0# $a Bible.
                   $p O.T.
                   $p Psalms.
```

240 Uniform title (NR)

Indicator 1: Uniform title printed or displayed

0 -- Not printed or displayed
1 -- Printed or displayed (most common)

Indicator 2: Nonfiling characters

0-9 -- Number of nonfiling characters present (for initial articles, including spaces)

Subfields used most often:

$a -- Uniform title
$l -- Language of a work
$f -- Date of a work

```
Example:    240 10 $a Ile mystérieuse.
                   $l English.
                   $f 1978
```

245 Title Statement (NR)

Indicator 1: Title added entry

(Should the title be indexed as a title added entry?)

0 -- No title added entry

(indicates a title main entry; i.e. no author is given)

1 -- Title added entry

(the proper indicator when an author given in 1XX; the most common situation)

Indicator 2: Nonfiling characters

0-9 -- Number of nonfiling characters present, including spaces; usually set at zero, except when the title begins with an article; e.g., for *The robe*, the second indicator would be set to 4. The letters *T*, *h*, *e*, and the space following them are then ignored in alphabetizing titles. The record will be automatically filed under "*r*" -- for *Robe*.

Subfields used most often:

$a -- Title proper
$h -- Medium (often used for non-book media)
$p -- Name of part/section of a work (R)
$b -- Reminder of title (subtitles, etc.)
$c -- Remainder of title page transcription/Statement of responsibility

```
Example:    245 14 $a The DNA story :
                   $b a documentary history of gene
                      cloning /
                   $c James D. Watson, John Tooze.
```

246 Varying form of title (R)

Indicator 1: Note/title added entry controller
1 -- Note, title added entry
3 -- No note, title added entry

Indicator 2: Type of title
-- No information provided
0 -- Portion of title
1 -- Parallel title
4 -- Cover title
8 -- Spine title

Subfield used most often:
$a -- Title proper

```
Example:    246 3# $a Four corners power review
```

250 Edition statement (NR)

Indicators undefined.

Subfield used most often:
$a -- Edition statement

```
Example:    250 ## $a 6th ed.
```

260 Publication, distribution, etc. (Imprint) (R)

Indicator 1: Sequence of publishing statements
-- No information provided

Indicator 2: Undefined

Subfields used most often:

$a -- Place of publication, distribution, etc. (R)
$b -- Name of publisher, distributor, etc. (R)
$c -- Date of publication, distribution, etc. (R)

```
Example:    260 ## $a New York :
                   $b Chelsea House,
                   $c 1986.
```

300 Physical description (R)

Indicators undefined.
Subfields used most often:

$a -- Extent (number of pages) (R)
$b -- Other physical details (usually illustration information)
$c -- Dimensions (cm.) (R)
$e -- Accompanying material (for example, "teacher's guide" or "manual")

```
Example:    300 ## $a 139 p. :
                   $b ill. ;
                   $c 24 cm.
```

440 Series statement / Added entry--Title

This field was made obsolete in 2008 to simplify the series statement. See 490 and 830.

490 Series statement (No added entry is traced from field) (R)

Indicator 1: Specifies whether series is traced (whether an 8XX tag is also present)

0 -- Series not traced
1 -- Series traced (8XX is in record)

Indicator 2 undefined.
Subfield used most often:

$a -- Series statement (R)
$v -- Volume number (R)

```
Example:    490 1# $a Colonial American craftsmen
```

500 General note (R)

Indicators undefined.
Subfield used most often:

$a -- General note (Used when no specialized note field has been defined for the information. Examples: Notes regarding the index; the source of the title; variations in title; descriptions of the nature, form, or scope of the item.)

```
Example:    500 ##  $a Includes index.
```

504 Bibliography, etc. note (R)

Indicators undefined.

Subfield used most often:

$a -- Bibliography, etc. note

```
Example:    504 ## $a Includes bibliographical
                     references.
```

505 Formatted contents note (R)

Indicator 1: Type of contents note

0 -- Complete contents

1 -- Incomplete contents (used with multivolume set when some volumes are not yet published)

2 -- Partial contents

Indicator 2: Level of content designation

-- Basic

Subfield used most often:

$a -- Formatted contents note

```
Example:    505 0# $a Pride and prejudice -- Emma
                     -- Northanger Abbey.
```

520 Summary, etc. note (R)

Indicator 1: Display constant controller

-- Summary

1 -- Review

2 -- Scope and content

3 -- Abstract

Indicator 2 undefined

Subfields used most often

$a -- Summary, abstract, or annotation

$b -- Expansion of summary note

```
Example:    520 ## $a This basic guide to parliamentary
                     procedure tells how to conduct
                     and participate in a meeting
                     properly.
```

600 Subject added entry -- Personal name (R)

Indicator 1: Type of personal name entry element

0 -- Forename

1 -- Surname (this is the most common form)
3 -- Family name

Indicator 2: Subject heading system/thesaurus (identifies the specific list or file which was used)

0 -- Library of Congress Subject Headings
1 -- LC subject headings for children's literature
2 -- Medical Subject Headings
3 -- National Agricultural Library subject authority file
4 -- Source not specified
5 -- Canadian Subject Headings
6 -- Répertoire de vedettes-matière
7 -- Source specified in subfield $2

(Note regarding Sears subject headings: The MARC 21 format does not provide an assigned indicator for Sears subject headings. Therefore, an indicator of 7 is used, and the MARC defined code "sears" is placed in subfield $2.)

Subfields used most often:

$a -- Personal name (surname and forename)
$b -- Numeration
$c -- Titles and other words associated with a name (R)
$q -- Fuller form of name
$d -- Dates associated with a name (generally, year of birth)
$t -- Title of a work
$v -- Form subdivision (R)
$x -- General subdivision (R)
$y -- Chronological subdivision (R)
$z -- Geographic subdivision (R)
$2 -- Source of heading or term (used with 2nd indicator of 7)

```
Example:    600 10 $a Shakespeare, William,
                   $d 1564-1616
                   $x Comedies
                   $x Stage history.
```

```
Example:    600 10 $a Shakespeare, William,
                   $d 1564-1616
                   $x Knowledge
                   $z Rome
                   $v Congresses.
```

Notice that subfields $v, $x, and $z in the 600 field are repeatable. Subfields $v, $x, $y, and $z do not have to be in alphabetical order. They will be in the order prescribed by the instructions given by the subject heading system.

610 Subject added entry -- Corporate name (R)

*Indicator 1:*Type of corporate name entry element

0 -- Inverted name (not used with AACR2)
1 -- Jurisdiction name
2 -- Name in direct order

Indicator 2: Subject heading system/thesaurus.
See indicator 2 under 600

Subfields used most often:
$a -- Corporate name or jurisdiction name as entry element
$b -- Subordinate unit (R)
$v -- Form subdivision (R)
$x -- General subdivision (R)
$y -- Chronological subdivision (R)
$z -- Geographic subdivision (R)
$2 -- Source of heading or term (used with 2nd indicator of 7)

```
Example:    610 10 $a United States.
                   $b Army Air Forces
                   $v Biography.
```

650 Subject added entry -- Topical term (Most subject headings fit here.) (R)

Indicator 1: Level of subject
-- No information provided

Indicator 2: Subject heading system/thesaurus
(identifies the specific list or file which was used)
0 -- Library of Congress Subject Headings
1 -- LC subject headings for children's literature
2 -- Medical Subject Headings
3 -- National Agricultural Library subject authority file
4 -- Source not specified
5 -- Canadian Subject Headings
6 -- Répertoire de vedettes-matière
7 -- Source specified in subfield $2

Note regarding Sears subject headings: The MARC 21 format does not provide an assigned indicator for Sears subject headings. Therefore, an indicator of 7 is used, and the MARC defined code "sears" is placed in subfield $2.)

Subfields used most often:
$a -- Topical term
$v -- Form subdivision (R)
$x -- General subdivision (R)
$y -- Chronological subdivision (R)
$z -- Geographic subdivision (R)
$2 -- Source of heading or term used with 2nd indicator of 7)

```
Example:    650 #0 $a Theater
                   $z United States
                   $v Biography
                   $v Dictionaries.
```

Notice that subfields $v, $x, and $z in the 650 field are repeatable. Subfields $v, $x, $y, and $z do not have to be in alphabetical order. They will be in the order prescribed by the instructions given by the subject heading system.

651 Subject added entry -- Geographic name (R)

Indicator 1: undefined.

Indicator 2: Subject heading system/thesaurus.

See indicator 2 under 600

Subfields used most often:

$a -- Geographic name
$v -- Form subdivision (R)
$x -- General subdivision (R)
$y -- Chronological subdivision (R)
$z -- Geographic subdivision (R)
$2 -- Source of heading or term (used with 2nd indicator of 7)

```
Example:    651 #0 $a United States
                   $x History
                   $v Chronology.
```

Notice that subfields $v, $x, and $z in the 651 field are repeatable. Subfields $v, $x, $y, and $z do not have to be in alphabetical order. They will be in the order prescribed by the instructions given by the subject heading system.

700 Added entry -- Personal name (R)

Indicator 1: Type of personal name entry element

0 -- Forename
1 -- Surname (this is the most common form)
3 -- Family name

Indicator 2: Type of added entry

-- No information provided (most common; co-authors, editors, etc.)
2 -- Analytical entry (The values for Indicator 2 changed in 1994 with Format Integration, and older records may display additional values. An analytical entry involves an author/title of an item contained in a work.)

Subfields used most often:

$a -- Personal name
$b -- Numeration
$c -- Titles and other words associated with a name (R)
$q -- Fuller form of name
$d -- Dates associated with a name (generally, year of birth)
$e -- Relator term (such as ill.) (R)
$4 -- Relator code (R)

```
Example:    700 1# $a Baldridge, Letitia.
```

710 Added entry -- Corporate name (R)

Indicator 1: Type of corporate name entry element

0 -- Inverted name (not used with AACR2)

1 -- Jurisdiction name

2 -- Name in direct order

Indicator 2: Type of added entry.

See Indicator 2 under 700

-- No information provided

2 -- Analytical entry

Subfields used most often:

$a -- Corporate name or jurisdiction name as entry element

$b -- Subordinate unit (R)

```
Example:    710 2# $a Sunburst Communications (Firm)
```

740 Added entry -- Uncontrolled related/analytical title (R)

Indicator 1: Nonfiling characters

0-9 -- Number of nonfiling characters present (for initial articles, including spaces)

*Indicator 2:*Type of added entry. See Indicator 2 under 700

-- No information provided

2 -- Analytical entry

(This field was redefined in 1994 with Format Integration. Prior to 1994, the field was also used for variant titles, such as a different wording on a spine title. In records created since Format Integration, those variant tiles appear in a 246 field.)

Subfield used most often:

$a -- Title

```
Example:    740 02 $a Uncle Vanya.
```

800 Series added entry -- Personal name (R)

Indicator 1: Type of personal name entry element

0 -- Forename

1 -- Surname

3 -- Family name

Indicator 2 undefined.

Subfields used most often:

$a -- Personal name

$b -- Numeration

$c -- Titles and other words associated with a name (R)

$q -- Fuller form of name

$d -- Dates associated with a name (generally, year of birth)
$t -- Title of a work (the series)
$v -- Volume number

```
Example:    800 1# $a Fisher, Leonard Everett.
                   $t Colonial American craftsmen.
```

830 Series added entry -- Uniform title (R)

Indicator 1 undefined.

Indicator 2: Nonfiling characters

0-9 -- Number of nonfiling characters present (for initial articles, including spaces)

Subfield used most often:

$a -- Uniform title
$v -- Volume number

```
Example:    830 #0 $a Railroads of America (Macmillan)
```

[Back to Top of Page]

Part VIII:

A List of Other Fields Often Seen in MARC Records

001	Control number
003	Control number identifier
005	Date and time of latest transaction
006	Fixed-length data elements -- additional material characteristics
007	Physical description fixed field
008	Fixed length data elements (See Part X)
022	International Standard Serial Number (ISSN)
037	Source of acquisition
041	Language code
043	Geographic area code
050	Library of Congress call number
060	National Library of Medicine call number
082	Dewey Decimal classification number (the one recommended by the Library of Congress; locally-assigned call numbers may appear elsewhere)
110	Main entry -- Corporate name (less frequent under AACR2 rules)
256	Computer file characteristics
263	Projected publication date (indicates a CIP -- Cataloging in Publication -- record)

306 Playing time
508 Creation/production credits note
510 Citation/references note (review sources)
511 Participant or performer note
521 Target audience note (first indicator: 0 = reading grade level, 1 = interest age level, 2 = interest grade level, 3 = special audience characteristics, 4 = motivation interest level)
530 Additional physical form available note
538 System details note
586 Awards note
656 Index term -- Occupation
730 Added entry -- Uniform title
852 Location
856 Electronic location and access
9XX Reserved for local use. (They are used by vendors, systems, or individual libraries to exchange additional data)

[Back to Top of Page]

Part IX:

The Leader

There are 24 positions in the Leader, numbered from 00 to 23. For fuller explanation, see the *MARC 21 Format for Bibliographic Data*.

00-04 Record length (calculated by the computer for each record)
05 Record status
- a = increase in encoding level
- c = corrected or revised
- d = deleted
- n = new
- p = increase in encoding from prepublication (previous CIP)

06 Type of record
- a = language material
- c = printed music
- d = manuscript music
- e = cartographic material
- f = manuscript cartographic material
- g = projected medium
- i = nonmusical sound recording
- j = musical sound recording
- k = 2-dimensional nonprojectable graphic
- m = computer file
- o = kit
- p = mixed materials
- r = 3-dimensional artifact or naturally occurring object
- t = manuscript language material

07 Bibliographic level
- a = monographic component part
- b = serial component part

c = collection
d = subunit
i = integrating resource
m = monograph/item
s = serial

08 Type of control
= no specified type
a = archival

09 Character coding scheme
= MARC-8
a = UCS/Unicode

10 Indicator count (always "2")

11 Subfield code count (always "2")

12-16 Base address of data (calculated by the computer for each record)

17 Encoding level
= full level
1 = full level, material not examined
2 = less-than-full level, material not examined
3 = abbreviated level
4 = core level
5 = partial (preliminary) level
7 = minimal level
8 = prepublication level (CIP)
u = unknown
z = not applicable

18 Descriptive cataloging form
= non-ISBD
a = AACR2
i = ISBD
u = unknown

19 Multipart resource record level
= Not specified or not applicable
a = Set
b = Part with independent title
c = Part with dependent title

20 Length of the length-of-field portion (always "4")

21 Length of the starting-character-position portion (always "5")

22 Length of the implementation-defined portion (always "0")

23 Undefined (always "0")

[Back to Top of Page]

Part X:

Field 008 for Books

Field 008 is used for Fixed Length Data Elements ("Fixed Field Codes"). There are 40 character positions in field 008, numbered from 00-39. Undefined positions must contain either a blank (#) or a fill character (|). Positions 00-17 and 35-39 are defined the same way for all media.

The information shown here for positions 18-34 applies only to books. For explanation of all the positions below and for positions 18-34 for other media, see the *MARC 21 Format for Bibliographic Data*.

Note that field 008 has no indicators or subfield codes.

00-05 Date entered on file (YYMMDD),
where Y=year, M=month, and D=day

06 Type of date/publication status:
- b = no dates given; B.C. date involved
- e = detailed date
- s = single known date/probable date
- m = multiple dates
- r = reprint/reissue date (Date 1) and original date (Date 2)
- n = dates unknown
- q = questionable date
- t = publication date and copyright date
- | = no attempt to code

07-10 Date 1/beginning date of publication

11-14 Date 2/ending date of publication

Date fields contain the year(s) of publication. The type of date(s) in these elements are specified in fixed field element 06: Type of date/publication status. (For further details, see the field 008 description in the *MARC 21 Format for Bibliographic Data*.)

15-17 Place of publication, production, or execution
For example:
- pk# = Pakistan
- cau = California (US)

(For a full list of codes used in these positions, see the *MARC Code List for Countries*.)

18-21 Illustrations (up to 4 codes):
- # = no illustrations
- a = illustrations
- b = maps
- c = portraits
- d = charts
- e = plans
- f = plates
- g = music
- h = facsimiles
- i = coats of arms
- j = genealogical tables
- k = forms
- l = samples
- m = phonodisc, phonowire, etc.
- o = photographs
- p = illuminations
- | = no attempt to code

22 Target audience:
- # = unknown or not specified
- a = preschool

b = primary
c = pre-adolescent
d = adolescent
e = adult
f = specialized
g = general
j = juvenile
| = no attempt to code

23 Form of item:
= none of the following
a = microfilm
b = microfiche
c = microopaque
d = large print
f = braille
r = regular print reproduction
s = electronic
| = no attempt to code

24-27 Nature of contents (up to 4):
= no specified nature of contents
a = abstracts/summaries
b = bibliographies (is one or contains one)
c = catalogs
d = dictionaries
e = encyclopedias
f = handbooks
g = legal articles
i = indexes
j = patent document
k = discographies
l = legislation
m = theses
n = surveys of literature
o = reviews
p = programmed texts
q = filmographies
r = directories
s = statistics
t = technical reports
u = standards/specifications
v = legal cases and notes
w = law reports and digests
z = treaties
| = no attempt to code

28 Government publication:
= not a government publication
i = international intergovernmental
f = federal/national
a = autonomous or semi-autonomous component
s = state, provincial, territorial, dependent, etc.
m = multistate
c = multilocal

l = local
z = other type of government publication
o = government publication -- level undetermined
u = unknown if item is government publication
| = no attempt to code

29 Conference publication:
0 = not a conference publication
1 = conference publication
| = no attempt to code

30 Festschrift:
0 = not a festschrift
1 = festschrift
| = no attempt to code

31 Index:
0 = no index
1 = index present
| = no attempt to code

32 Undefined (since 1990) (Earlier records may contain the values 0 or 1)
= Undefined
| = no attempt to code

33 Literary form:
0 = not fiction (not further specified)
1 = fiction (not further specified)
c = comic strips
d = dramas
e = essays
f = novels
h = humor, satires, etc.
i = letters
j = short stories
m = mixed forms
p = poetry
s = speeches
u = unknown
| = no attempt to code

34 Biography:
= no biographical material
a = autobiography
b = individual biography
c = collective biography
d = contains biographical information
| = no attempt to code

35-37 Language:
A three-letter code. For example: eng fre ger spa rus ita

(For a full list of codes used in these positions, see the *MARC Code List for Languages*.)

38 Modified record:
= not modified
x = missing characters (because of characters unavailable in MARC character set)
s = shortened
d = "dashed-on" information omitted

r = completely romanized/printed cards in script
o = completely romanized/printed cards romanized
| = no attempt to code

39 Cataloging source:
= national bibliographic agency
c = cooperative cataloging program
d = other sources
u = unknown
| = no attempt to code

[Back to Top of Page]

[Back to Table of Contents] -- [Continue to Part 11]



Library of Congress
Library of Congress Help Desk (10/27/2009)

APPENDIX 0001A

On Adaptive Bandwidth Sharing with Rate Guarantees

N.G. Duffield† T. V. Lakshman* D. Stiliadis*

† AT&T Laboratories
Rm A175, 180 Park Avenue
Florham Park,
NJ 07932-0971 USA

* Bell Labs
Lucent Technologies
101 Crawfords Corner Road
Holmdel, NJ 07733, USA

Keywords—**fair queueing, scheduling, jitter, delay bounds, real-time traffic, traffic management**

Abstract

The objective of recent research in fair queueing schemes has been to efficiently emulate a fluid-flow generalized (weighted) processor sharing (GPS) system, as closely as possible. A primary motivation for the use of fair-queueing has been its use as a means of providing bandwidth guarantees and as a consequence end-to-end delay bounds for traffic with bounded burstiness. The rate guarantees translate to scheduling weights which are set when admission control is done. A consequence of fair queueing systems closely emulating GPS is that when one or more connections are not backlogged, any "excess" bandwidth is distributed to backlogged connections in proportion to their weights. However, weights are set based on the long-term requirements of traffic flows and not in any state-dependent manner that reflects instantaneous needs. In this paper, we question the notion that queueing systems should closely emulate a GPS system. Instead of emulating GPS, we propose three modified scheduling schemes which preserve the rate guarantees of fair queueing (and hence preserve deterministic delay bounds) but adaptively redistribute the excess bandwidth such that either losses are reduced or delays equalized. We compare the performance of the proposed schemes to that of fair queueing using different traffic sources such as voice and video, as well as sources which have aggregate long-range dependent behavior. We find that the proposed schemes, in comparison to packet GPS (PGPS), reduce packet losses and curtail the tails of delay distributions for real-time traffic and hence permit the use of significantly smaller playout buffers for the same network load.

I. Introduction

Much research attention has recently been focused on the development of fair-queueing systems [3], [10], [19], [20], [21], [24] that, within the non-preemption constraints of a packet system, closely emulate fluid flow generalized processor sharing systems (GPS) and operate at high speeds. Demers, Keshav, and Shenker proposed a fair queueing scheme that emulated GPS by using a simulated fluid flow GPS system as a reference and basing packet scheduling decisions on the order of departures in the simulated GPS system. The computational burden of simulating GPS was reduced in the self-clocked fair queueing scheme proposed by Golestani [10] which showed how fair queueing could emulate GPS without simulating GPS for reference. Subsequent work has led to both improved emulation bounds, further reductions in computational complexity and efficient methods for implementation [11], [20], [21], [24]. However, there is no detailed examination of whether the exclusive emulation of GPS is always appropriate for all networking applications.

In this paper, we question the notion that queueing systems must emulate GPS closely. A primary motivation for the expected use of fair queueing in routers and switches is its ability to guarantee worst case end-to-end delay bounds for leaky bucket controlled sources [17]. However, to guarantee worst case delay bounds the scheduler merely needs to isolate flows so that each flow receives its guaranteed share of the link bandwidth [1], [19]. The guarantees do not depend on the property of GPS systems that the unused shares of non-backlogged flows be redistributed to backlogged flows in proportion to their guaranteed fractional share of link bandwidth (or equivalently scheduling weights). A fair queueing system need act as one only when all connections are backlogged. When excess bandwidth (i.e. bandwidth beyond what has already been guaranteed to backlogged connections) is available, the fair queueing system, instead of emulating GPS that distributes excess bandwidth according to long term needs, should redistribute this excess bandwidth in a manner which reflects the current or instantaneous needs of backlogged flows (i.e. a flow state dependent redistribution). The redistribution policy can be picked to optimize measures such as packet loss probabilities or to control the tails of delay distributions.

It has been shown in [1], [23] that even when a FIFO scheduler is used and all sources are leaky-bucket shaped, the mean and 99th percentile of delays are much lower than the worst-case bounds guaranteed by a GPS server. However, there are service disciplines that can distribute the bandwidth in a more elegant manner. Panwar et.al. [16] showed, that the shortest time to extinction policy is actually optimal for scheduling customers with deadlines. Furthermore, a policy that simply serves the longest queue has been shown to require less buffer space to prevent buffer losses than FIFO or other per-flow disciplines [6], [13]. Similarly, policies that serve the queue that is more likely to violate a bound are proved to provide optimal buffer utilization [9]. However, such disciplines by themselves cannot provide any bandwidth guarantees. Actually, low bandwidth constant bit rate connections are likely to starve for long intervals of time, until their queue becomes at least as large as the queues of heavily bursty, high bandwidth connections. Similarly, users that can accept a significant loss rate, using some form of Forward Error Correction, can increase their throughput by simply sending more packets. There is no method to control the bandwidth of individual users.

We propose three modified fair queueing schemes that adaptively redistribute excess bandwidth while, like fair queueing, guaranteeing each flow its specified share of the link bandwidth. This preserves fair queueing's ability to provide worst case end to end delay bounds and the schemes work like fair queueing when there is no excess bandwidth. In addition, the schemes provide worst-case fairness, bounding the time that a connection may not see any service. The excess bandwidth, however, is redistributed as follows:

1. Longest delay first (LDF) that serves the flow with current longest delay.
2. Least time to overflow (LTO) that serves the flow with minimum difference between maximum allowed delay and current delay.
3. Least time to overflow with leaky buckets (LTO_LB) that serves the flow which would cause buffer overflow first if worst case arrivals happen.

We compare the performance of these policies to PGPS using both trace driven simulations as well as simulations with traffic models for various types of sources.

The LDF policy uses excess bandwidth to reduce the variance of the delay distribution. This has the benefit of reducing the playout buffer for voice and video sources. Simulations with video traces and with voice traffic show that indeed this policy performs better than PGPS without any sacrifice of worst case guarantees. Since the deviation from the maximum allowed delay is not taken into account, flows with small delay bounds (like voice) get almost no excess bandwidth in the presence of flows with large delay bounds. Long term congestion or small errors in assigning weights can result in these flows experiencing losses much more than flows with large delay bounds. The LTO policy tries to minimize packet losses by assigning excess bandwidth under the assumption that the flow which is likely to overflow the quickest has the most instantaneous bandwidth need. In doing so, it takes into account the current deviation of each flow from its maximum allowed delay. Simulations with a mix of CBR, voice, and video sources with very different delay bounds show that this policy reduces losses for all classes as well as reduces the variance of delay for each class. The LTO_LB policy goes one step further by determining the connection that is most likely to overflow its buffer under the assumption that arrivals are leaky-bucket shaped.

A scheduling and buffer sharing scheme that tries to guarantee specific delays while ensuring loss-free behavior is presented in [7]. Our proposal is fundamentally different from the scheme presented in [7]. We assume that our system does not have enough buffers to guarantee zero losses. So our goal is not to define a system that will guarantee some specific delays

0-7803-4383-2/98/$10.00 © 1998 IEEE.

and at the same time zero losses. Instead, we investigate the effect of state-dependent scheduling mechanisms in minimizing the tails of the delay distributions while providing isolation, and also investigate the effect of such mechanisms on the number of buffers that need to be reserved if limited losses are acceptable.

II. Architecture

Fair Queueing systems attempt to offer the same *normalized service* to any two connections that are continuously backlogged during an interval of time $(t_1, t_2]$, where by normalized service we mean the ratio of the offered service and the allocated rate. The basic feature of such a system is that it treats connections in the same way, irrespective of their burstiness characteristics. Our claim is that end-users are not necessarily interested in a fair distribution of free bandwidth. Instead, for many applications, end-users are merely interested in minimizing their packet losses or seeing low tails in their delay distributions. Of course, the queueing scheme must be able to provide some guarantees regarding worst-case service offered to a connection when the system is fully loaded. Isolation and protection requirements such as these were defined in [1] as necessary to protect the system from misbehaving users that will try to monopolize resources.

A. Preliminaries

The definition of our state-dependent queueing system uses the methodology presented in [20] for designing fair queueing systems. The method is based on the class of servers called Rate Proportional Servers (RPS). The key concept is a virtual time function $v_i(t)$ or *system potential* associated with each connection in the system that represents the total normalized service offered to a connection during its backlogged periods. The virtual time of a connection is defined as a non-decreasing function of time during a system-busy period. When connection i is backlogged, its virtual time increases exactly by the normalized service it receives. That is, for any interval of time $(\tau, t]$ that connection i is continuously backlogged,

$$v_i(t) - v_i(\tau) = \frac{W_i(\tau, t)}{g_i}.$$

where, $W_i(\tau, t)$ denotes the amount of service received by session i during the interval $(\tau, t]$ and g_i is the scheduling weight allocated to connection i. When the connection is idle, however, the virtual time of the connection is updated through the *system virtual time*. The system virtual time is a non-decreasing function of time that keeps track of the progress of the total work done by the scheduler. When an idle session i becomes backlogged at time t, its virtual time $v_i(t)$ is set to $v(t)$ to account for the service it missed. Schedulers use different functions to maintain the system virtual time, giving rise to widely different delay and fairness-behaviors. Exactly this flexibility of the rate-proportional servers allows a simple implementation of the system without the requirement of simulating in parallel a fluid system.

At any instant t, the scheduler services only the subset of connections with the minimum virtual time, and each connection in this subset receives service in proportion to its reserved rate g_i. Thus, the scheduler can be seen to increase the virtual times of the connections in this subset at the same rate. A packet server approximates this behavior by sorting the packets in the order of their finishing virtual times under the fluid server, and transmitting them in increasing order of the finishing virtual times. The finishing virtual time of the kth packet of a session i, arriving at the scheduler at time t, is computed as

$$F_i^k = \min(F_i^{k-1}, v(t)) + \frac{L_i^k}{g_i}, \tag{1}$$

where L_i^k is the length of the packet and g_i the allocated rate of session i. That is, the finishing virtual time is computed by adding the service time of the packet at the allocated rate to its starting virtual time. The latter, in turn, is the minimum of the finishing virtual time of the previous packet received from session i, and the current system virtual time.

At the time that a connection becomes backlogged, its virtual time is updated based on the system virtual time function that keeps track of the progress of the total work done by the scheduler. When an idle session i becomes backlogged at time t, its virtual time $v_i(t)$ is set as

$$v_i(t) = \max(v_i(t-), v(t)),$$

to account for the service it missed. Rate-Proportional Servers may use a wide range of functions to maintain the system virtual time, but the function chosen must satisfy two fundamental properties: First, during any interval $(t_1, t_2]$ within a system-busy period, the system virtual time function must be increased with a rate of at least one, that is,

$$v(t_2) - v(t_1) \geq t_2 - t_1. \tag{2}$$

Second, the system virtual time function must never exceed the virtual time of any backlogged connection. These two conditions are sufficient to achieve a delay bound equal to that of PGPS.

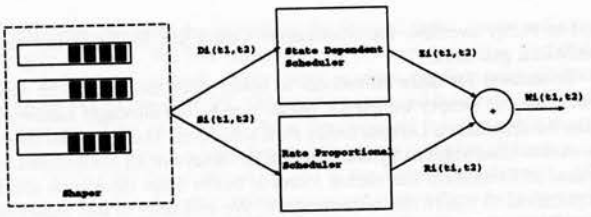


Fig. 1. Scheduling System

B. State Dependent Fair Queueing

An illustration of the logical structure of our system is presented in Figure 1. Initially, all arriving packets enter per-connection queues in a shaper device. A packet is considered as having arrived to the system only when the last bit of the packet arrives to the system. Similarly, packet service is considered to be completed only when the last bit of the packet has been served. The shaper releases packets to the scheduler with a rate exactly equal to the allocated. We denote by $S_i(\tau, t)$ the service offered by the shaper to connection i during an interval of time $(\tau, t]$. Packets are transferred from the shaper to the the scheduler with infinite capacity. Let $R_i(\tau, t)$ denote the service offered by the RPS scheduler.

Packets that have not become eligible for service remain in the corresponding connection queue in the shaper, while all the eligible packets wait for service in the RPS scheduler queue. Service is always provided from the RPS scheduler queues, as long as packets are available there. When all scheduler queues are empty, a *state dependent scheduler* (SDS) is invoked and selects a packet from the shaper queues for transmission. The service offered to a connection does not change the state of the shaper. Let us denote with $D_i(t_1, t_2)$ and $Z_i(t_1, t_2)$ the amount of traffic that is forwarded and serviced by the SDS scheduler respectively. If the state dependent scheduler is work-conserving, then the system is work conserving as well. Note also, that for every time t, $D_i(t_1, t_2) = Z_i(t_1, t_2)$. We will refer to these systems as simply FQ-SDS.

An argument that can be made against the above adaptive policies, is that a greedy user will always receive more bandwidth. In such a case, users may request less bandwidth and try to receive more by sending more bursty traffic. Limiting such behavior can be explicity done in many ways. The simplest is policing. In most networks that provide some type of guarantees, it is reasonable to assume that there is some kind of policing. So misbehaving users will be easily isolated. If policing is not a viable option, the adaptive policies can themselves be tailored to counter greedy behavior. Instead of picking the user with maximum queue length or the minimum time to overflow, we can, for example, pick the second such candidate. In this case, users that consistently overflow the buffers cannot depend on receiving any excess bandwidth; it will not be in their best interests to be excessively bursty.

C. Redistributing Excess Bandwidth

This section describes the schemes that we use for the state dependent component of our schedulers. Note that this component is used only to redistribute the excess bandwidth and does not affect the guaranteed rates. The basic idea is that the scheduler, in picking connections to serve, should use information about the state of connections so as to minimize packet losses, reduce average delays for some traffic classes, or shape delay distributions.

Simple Longest Queue First has been shown to minimize the overall packet losses in a system with finite buffers [16]. Here, the scheduler always picks a packet from the longest queue for transmission. However, if the arrival rates and guaranteed shares are such that the expected delays are are not the same for all connections then the longest queue first scheme will always favor high delay (and generally high bandwidth connections). Also, if buffer allocations are not the same, serving the longest queue first may

lead to buffer overflow for connections with a low buffer or low relative bandwidth guarantee.

To account for these differences in buffer sizes and bandwidth requirements, we can simply weight the queue size by the allocated bandwidth to make the scheduler a Longest Delay First scheduler. The scheduler then uses the excess bandwidth to try and equalize the delays of all connections in the system and minimize the packet losses if buffer sizes have been allocated proportional to bandwidth requirements. We will refer to this algorithm as *FQ-LDF*.

Note that this still does not take into account the maximum allowed delay for each connection. An enhancement to FQ-LDF is to explicitly take into account the different delay requirements that applications may have. We would like to be able to allow low maximum delays to voice sources, somewhat higher maximum delays to video sources, and much higher maximum delays for data sources. Also, constant bit rate sources may require extremely low delays, even when their bandwidth allocation is very low. To account for these differences, a maximum delay, that is usually a function of the buffer allocated to a connection, is associated with each connection. The scheduler serves connections that are most likely to exceed their allocated maximum delay. Let D_i denote the maximum delay allocated to connection i. Let us also denote by $d_i = Q_i(t)/g_i$ the delay that the last packet of the queue associated with that connection at time t may see. The connection that should be served is the one that is more likely to overflow and exceed its delay bound. Thus, connections should be served in increasing order of the $(D_i - d_i)$ values.

The above schemes have not taken into account other characteristics of the source that may be known during admission control. Characteristics could be the maximum burst that can be transmitted by a connection or its peak arrival rate. For example, if sources are leaky bucket shaped, as is the case for sources in the Controlled Load traffic class, we could use this information to predict future arrivals from sources and calculate more precisely the connections that are likely to overflow or see delays higher than the allocated. In systems, where a dual, or multiple leaky bucket is used to shape the input traffic, a similar approach can be used. In fact, we do not require that a connection is shaped by a leaky bucket. All we need is that there is a function specified that provides information as to the maximum burst that can arrive for a connection, as well as the worst-case time that this burst may arrive.

The main idea consists of tracking the state of the connection based on its packet arrivals. This tracking is similar to the policing function that is used to drop packets from connections when they exceed their traffic specification. Based on the state of the connection we can estimate the worst-case arrivals for a given connection. In order to minimize the effect of worst-case arrivals for connections, we can serve connections in a way that will minimize the probability that they will be penalized for such an arrival pattern. Note, that since the system is only distributing the instantaneous available bandwidth in a state dependent manner and since connections are unlikely to send their worst case traffic at the same time, such an approach may lead to significant improvements in maximum delays or packet losses.

Let $B_i(\tau, t)$ be the function that describes the maximum amount of traffic that can arrive for connection i during the interval $(\tau, t]$ and let $A_i(\tau, t)$ denote the actual arrivals from session i during the same period. That is, for every interval $(\tau, t]$:

$$A_i(\tau, t) \leq B_i(\tau, t) \tag{3}$$

Then at time t, a maximum burst equal to

$$\Delta B_i \leq B_i(\tau, t) - A_i(\tau, t) \tag{4}$$

may arrive with the peak rate. If the queue size of connection i is equal to $Q_i(t)$ and the maximum buffers allocated to that connection are β_i, then a connection may overflow if

$$\beta_i - Q_i(t) \geq \Delta B \tag{5}$$

Among the connections that may overflow, we will select to transmit a packet from the connection that may overflow sooner, if the maximum burst arrives with the peak rate allocated to that connection. Or,

$$j = min_{k:\beta_k - Q_k(t) \geq \Delta B_k}(\frac{\Delta B_k}{P_k}) \tag{6}$$

where P_i is the peak rate allocated to connection i.

III. Analysis

A. Preliminaries

Definition 1: A **system busy period** is a maximal interval of time during which the server is never idle.

During a system busy period the server is always transmitting packets.

Definition 2: A **backlogged period for session i** is any period of time during which packets belonging to that session are continuously queued in the system.

Let $Q_i(t)$ represent the amount of session i traffic queued in the server at time t, that is,

$$Q_i(t) = A_i(0, t) - W_i(0, t).$$

A connection is backlogged at time t if $Q_i(t) > 0$.

Definition 3: A **session i busy period** is a maximal interval of time $(\tau_1, \tau_2]$ such that for any time $t \in (\tau_1, \tau_2]$, packets of connection i arrive with rate greater than or equal to g_i, or,

$$A_i(\tau_1, t) \geq g_i(t - \tau_1).$$

A session busy period is the maximal interval of time during which if the session were serviced with exactly the guaranteed rate, it would remain continuously backlogged. It is important to realize the basic distinction between a session backlogged period and a session busy period. The latter is defined only in terms of the arrival function and the allocated rate. Thus, the busy period serves as an invariant for evaluating the worst-case behavior of different scheduling algorithms under the same arrival pattern.

B. Worst-Case Performance

In [19] it was shown that any RPS scheduler has the same worst-case performance as a Weighted Fair Queueing scheduler, and it can thus guarantee the same worst-case delay bounds. However, the addition of the shaping mechanism and the state dependent mechanism affects the arrival pattern at the scheduler. The intuition is that the SDS scheduler is only invoked when there is a free bandwidth available. Worst-case performance usually assumes a fully utilized system. We will now show that the worst-case service offered by the system as described above is not affected by the method by which free bandwidth is distributed.

Our main goal is to provide a lower bound for the service offered by the system to a session i during a session busy period. Note, that according to our definition a session busy period is defined as the maximal interval of time during which the traffic arriving for that connection is at least equal to the allocated bandwidth. Notice, however, that the traffic that arrives to the RPS scheduler can be lower than the traffic arriving to the system. The reason is, that the connection receives excess bandwidth from the SDS server. That means, that the busy periods that the RPS scheduler sees for a connection i, may be different than the connection busy periods as seen by the system. It is easy to verify, however, that a busy period in the RPS system can continue after the end of a busy period as seen by the whole system. We will first consider the fluid modeled RPS system. Let us denote with $W_i^F(\tau, t)$ the service offered by the fluid RPS server and with $W_i^P(\tau, t)$ the service offered by the corresponding packet-by-packet server.

Lemma 1: Let τ be the beginning of a busy period for connection i. The worst case offered service to connection i for any time t during the same busy period is bounded by

$$W_i^F(\tau, t) \geq g_i(t - \tau)$$

The proof is in the Appendix. The basic idea is that since the shaper is offering service at least equal to that reserved for connection i, the scheduler will always have enough packets to service connection i with a rate equal to the reserved.

Note, that we have made no assumption for the SDS server. We can thus assume that it is a packet system. In addition, the packet and fluid RPS systems see exactly the same arrivals, and they are both work conserving systems. Thus, in both systems, the SDS server is serving packets during the same intervals of time. Thus, if the packet server offers different service than the fluid server, this difference is due to the discrepancy between the packet-by-packet and fluid RPS server. But we know from [19] that this difference is bounded by $L_i/g_i + L_{max}/r$ where L_i is the maximum packet size for connection i, L_{max} is the maximum packet size of any connection in the server and r is the link capacity. We can thus write:

Corollary 1: The worst case offered service to the packets of the j-th busy period of connection i that started at time τ is

$$W_{i,j}^P(\tau, t) \geq g_i(t - \tau - \frac{L_i}{g_i} - \frac{L_{max}}{r}).$$

This Corollary is a sufficient condition to prove that the system offers the same worst-case end-to-end delays as PGPS under the rate-proportional assignments [19].

C. Worst-Case Fairness

Worst-case fairness was initially defined by Parekh in [17] and later expanded by Bennett and Zhang [24]. The measure is based on the worst-case delay for clearing the backlog of a session's queue. According to this, for every session i that is continuously backlogged during the interval $(t_1, t_2]$,

$$W_i(t_1, t_2) \geq (t_2 - t_1)g_i - C_i, \quad (7)$$

where $W_i(t_1, t_2)$ represents the service offered to session i during the interval $(t_1, t_2]$, and C_i is a constant. We will call the smallest value of C_i satisfying this inequality as the *worst-case fairness index* (WFI). Bennett and Zhang [24] defined a fairness parameter based on the smallest value of C_i satisfying the inequality, and normalizing it to the allocated rate of the session.

The worst-case fairness index measures the deviation between the service received by a session in the packet-level scheduler and the service received by the same session in a corresponding fluid server, where the session is serviced at a rate of g_i at each instant. Minimizing this parameter improves the traffic mix at the output of the scheduler [24]. However, it should be noted that this parameter does not specify how the free bandwidth left over by idle sessions is distributed across the active ones. Specifically, even for the algorithms that distribute the free bandwidth in a state dependent manner, we can show that they offer the same worst-case fairness as Worst-Case Weighted Fair Queueing [24]. The proof is in the Appendix.

IV. Performance Studies

We simulated the system shown in Figure 1 using a variety of sources to generate traffic and using the different schedulers (FQ-LDF, FQ-LTO, and PGPS) that were described in the previous section. In the simulated system, each source is connected to a leaky bucket shaper. The shaper is connected via an access line to the multiplexing system. For each connection, the scheduler in the multiplexer maintains a separate buffer whose drain time at the guaranteed rate is equal to the connection's maximum allowed delay D_i. The sources used are:

1. Two-state exponential On-Off sources: We used these to emulate compressed voice sources and so used small maximum allowed delays for these sources. The mean load generated per connection is 1/32 of the link capacity.
2. Video teleconference sources: We used the DAR(1) model which is a Markov chain determined by three parameters: the mean, variance, and one frame-lag correlation ρ. The transition matrix is computed as:

$$P = \rho I + (1 - \rho)Q \quad (8)$$

   where I is the identity matrix, and each row of Q consists of the negative binomial (or gamma) probabilities $(f_0, ..., f_K, F_K)$ where $F_K = \sum_{k>K} f_k$ and K is the peak rate. The DAR(1) model matches the autocorrelation of the data over approximately hundred frame lags and has been shown to be a model of video teleconferences accurate enough for traffic studies [5]. We used video sources with a mean rate of 1.5 Mbps, a peak-to-mean ratio of 5, and $\rho = .98$.
3. CBR sources.
4. Data sources. These were modeled as On-Off sources where the on-time is Pareto distributed, and the off-time is exponential. The density of the Pareto distribution is given by

$$f(x) = a^{k-1}(k-1)/x^k, x >= a, a = (k-2)/(k-1)\mu \quad (9)$$

   We set $k = 5/2$, resulting in infinite variance of the on-times and long-range dependence in the aggregate traffic [12], [22]. The mean on-time was set to 6 packet transmission times and the mean off-time to 200.

The leaky bucket shapers were turned off by setting their token rates equal to the access line rate.

We first compared the performance of FQ-LDF with PGPS (which is simulated exactly) using only one type of source. Table I shows packet loss rates for different small buffer sizes from a simulation with 32 voice sources. All connections were served with equal weight and the utilization is about .98. At this high load, PGPS typically has twice the losses of FQ-LDF and redistribution of excess bandwidth to the longest delay queue reduces the aggregate loss rate. We find similar results of lower loss rates for FQ-LDF with video sources as well. With Pareto sources, the loss rates are mostly similar for FQ-LDF and PGPS. This can be accounted for by the heavy-tailed distributions of the on periods. From this property it follows that the most likely manner in which a large burst in the aggregate over all such sources occurs, is that the burst of one source is far larger than the rest. (Such consequences of underlying infinite variance models have been termed the Joseph and Noah effects; see [14], [15]). Most of the other sources have non-backlogged queues and both schedulers give most of the available bandwidth to the one very long queue. Hence, PGPS and FQ-LDF tend to behave almost identically.

Apart from packet losses, we also compared the delay distribution resulting from FQ-LDF to that of PGPS. For the 32 voice source simulation with large enough buffers to avoid packet losses, Figure 2 shows the tail frequencies (the number of times a given delay is exceeded) on a log scale. The plots show that FQ-LDF needs significantly smaller playout buffers than PGPS. We see the same effect for video sources from the boxplots in Figure 3. The utilization is .88 and the buffers are large enough so that there are no losses. The distribution of mean delays experienced by different connections with FQ-LDF and PGPS is shown in the boxplots by showing the median (the line in the middle of the box) of the means, the quartiles (the upper and lower edges of the box enclosing the median), and by showing the range of values outside the quartiles by the lines extending out from the box. The curtailment of the tail of the delay distribution due to the use of FQ-LDF is evident. Thus, for similar loads FQ-LDF has a lower playout buffer requirement than PGPS.

Next, we compare the performance of FQ-LTO with PGPS. FQ-LDF uses the excess bandwidth to equalize delays and it works very well when all sources have the same allowed delays. When different sources have different maximum allowed delays, the FQ-LTO scheduler which takes delays into account is expected to perform better than FQ-LDF and PGPS. We use 32 sources with 8 sources from each of the types described before. CBR sources are guaranteed their mean rate. They generate at most 2 back-to-back packets and have a small buffer allocation of 5 packets. Voice sources are also guaranteed their mean rate and have a maximum buffer allocation of 1200 packets (about 10 ms delay at their guaranteed rates). Video sources are guaranteed two times their mean rates and have a buffer of 12000 packets. Data sources are given a large buffer of 40000 packets. Since they are not delay sensitive, they are guaranteed a rate below their mean rate. Data sources are expected to use some the excess bandwidth from the other classes to make up for their allocation below their means.

Figure 4 shows the box plots for mean delays and 99^{th} percentile delays for the voice (exponential on-off), video, CBR and data classes. The PGPS scheduler gives excess bandwidths to the data class in proportion to the weights. Since the data class has a very large delay allowance, the FQ-LTO scheduler gives priority to the low delay classes in the distribution of excess bandwidth. Consequently, the lower delay classes have very compact delay distributions whereas the very high delay data class (which also has underallocated guaranteed bandwidth) has better tail behavior with PGPS. The FQ-LTO scheduler achieves what one would want in this mixed scenario. Worst case delay bounds are always met even for data (if it is given appropriate guaranteed rates). The lower delay classes have very small tails and hence need drastically lower playout buffers than when using PGPS.

Finally, we study the situation where we have the same mix of sources as before but their guaranteed rates are different even within a source class. We deliberately set D_is to be below what they would be if allocation was strictly based on deterministic bounds, i.e, we assume that some statistical multiplexing will happen. The question then is how well LTO, which tries to prevent buffer overflows and minimize delays, compares with PGPS. The results are shown in Figure 5. We get very low average delays for real-time sources, in comparison to the delays experienced under PGPS, even when their bandwidth allocations are very low. This is because the D_is are set low for real-time sources, and some of the excess bandwidth which would have gone to data sources under the PGPS scheme is now used to further reduce the delays of real time sources. The differences in delays between LTO and PGPS can be very high as is seen by comparing delays for the voice class. Note that if data traffic losses must be prevented, this can be done by giving those connections an appropriate weight.

Buffer size in pkts/conn	Cell Loss Ratio:PGPS	Cell Loss Ratio:FQ-LDF
75	13.18e-3	8.3e-3
100	8.198e-3	4.938e-3
125	5.61e-3	3.61e-3
150	4.245e-3	2.849e-3
250	2.061e-3	1.025e-3
350	9.614e-4	1.054e-4
425	4.876e-4	$< 10^{-6}$
500	2.8478e-4	$< 10^{-6}$

TABLE I. Cell Loss Ratios for 32 Exponential On-Off Sources with Short Buffers

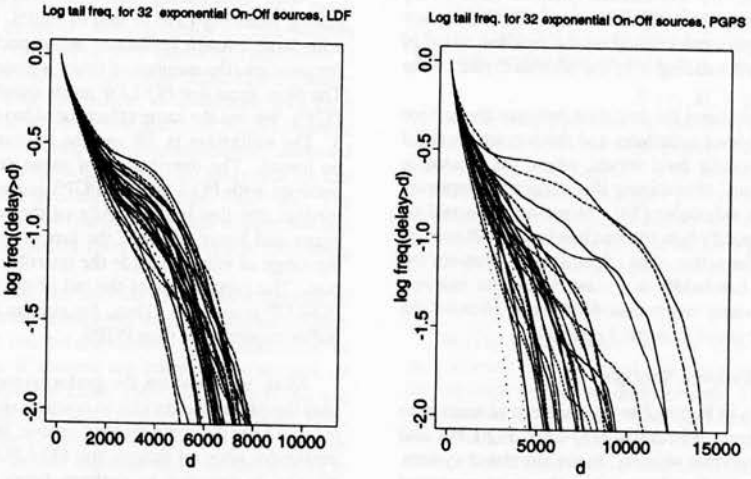


Fig. 2. Log tail frequencies for 32 exponential On-Off sources, according to service discipline (FQ-LDF or PGPS).

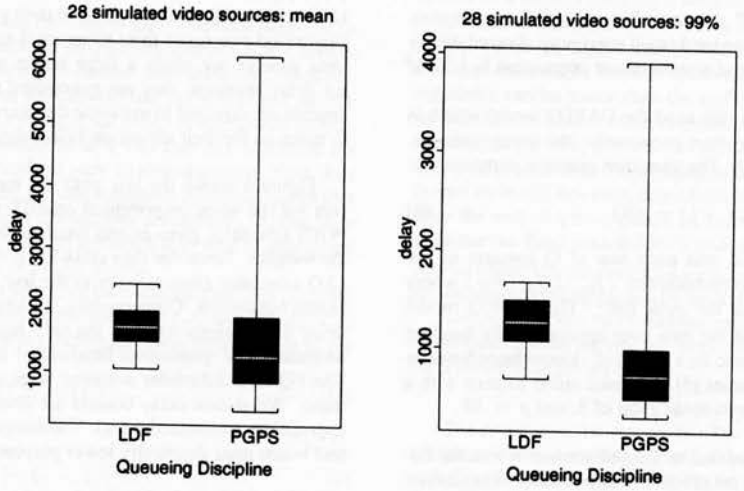


Fig. 3. Boxplots of mean delay (left plot) and 99th percentile delay (right plot) according to excess service discipline (in each plot: FQ-LDF on left, PGPS on right) for 28 simulated video sources. Key: shaded between first and third quartiles; clear bar is median; boundaries at minimum and maximum.

V. Discussion and Concluding Remarks

The main contribution of this paper is the modified fair queueing proposal that redistributes excess bandwidth in a state dependent manner. The drawback of GPS that all fair queueing systems inherit in their close emulation of GPS is that GPS severely restricts state-dependent bandwidth sharing. The only state-dependency in GPS is in the number of backlogged connections. There is no further latitude and sharing is determined by the guaranteed rates which are set based on long term needs of the connections. This restriction on bandwidth sharing is more stringent than that necessary to preserve a key property of fair queueing, the ability to guarantee delay bounds for leaky bucket controlled traffic sources. Consequently, for many applications, there is no need for fair queueing systems to emulate the potentially suboptimal excess bandwidth sharing of GPS.

With this in mind, we propose modified fair queueing schemes that emulate fair queueing only in that they provide rate guarantees and behave like fair queueing when all connections are backlogged. This is sufficient to guarantee worst case delay bounds using our scheme. Furthermore, the

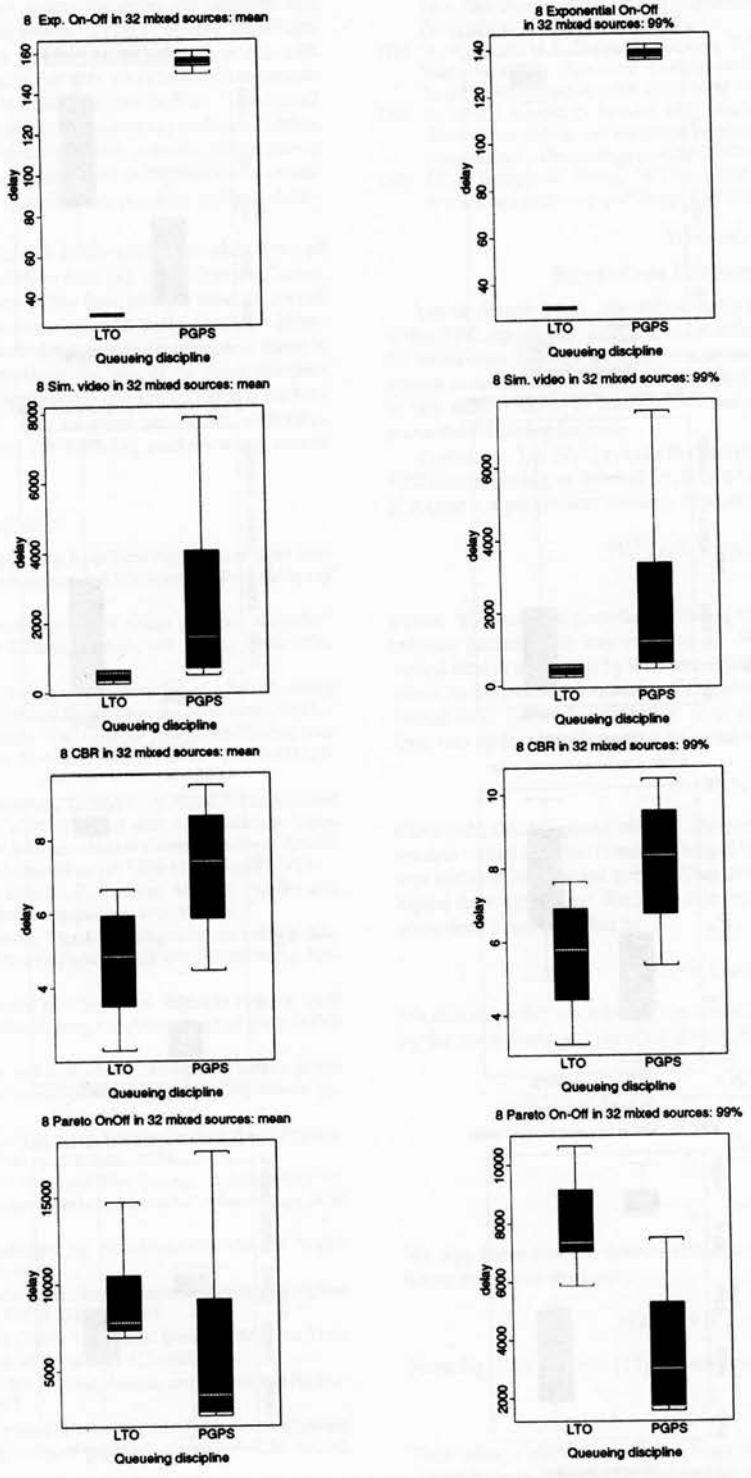


Fig. 4. Boxplots of mean delay (left column) and 99th percentile delay (right column) according to excess service discipline (in each plot: FQ-LTO on left, PGPS on right), displayed by row for each of following 2 classes of sources out of 4 classes being multiplexed: (top to bottom) Exponential On-Off, Simulated Video, CBR, Pareto On-Off. 32 sources total. Key: shaded between first and third quartiles; clear bar is median; boundaries at minimum and maximum.

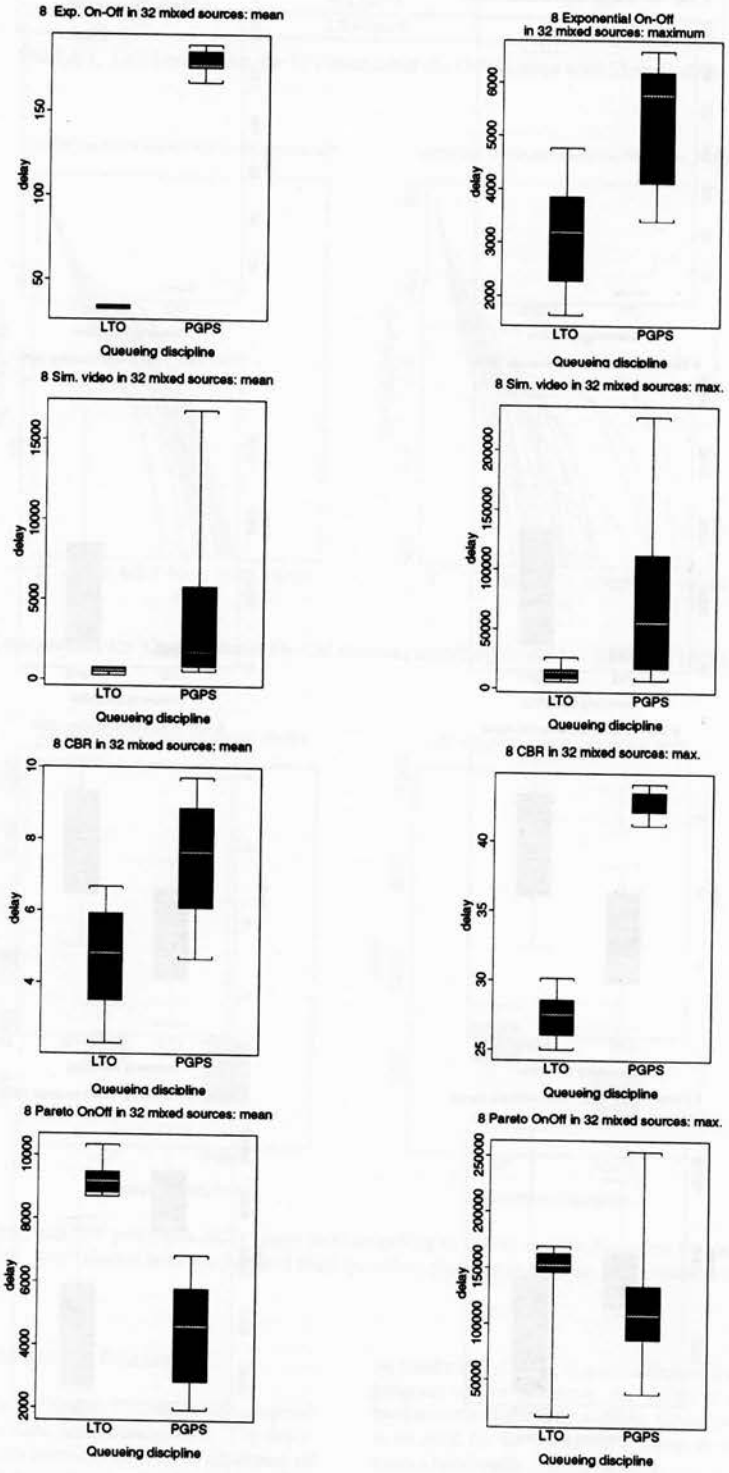


Fig. 5. Boxplots of mean delay (left column) and maximum delay (right column) according to excess service discipline (in each plot: LTO on left, PGPS on right), displayed by row for each of 4 classes of sources: (top to bottom) Exponential On-Off, Simulated Video, CBR, Pareto On-Off. 32 sources total. Key: shaded between first and third quartiles; clear bar is median; boundaries at minimum and maximum.

schemes we propose are worst case fair. Our schemes do not emulate GPS's method for excess bandwidth redistribution. Instead, we proposed three methods (the longest delay first, longest time to overflow, and longest time to overflow taking into account leaky bucket states) for redistributing the excess bandwidth. Simulations show that the FQ-LDF and FQ-LTO policies perform very well in comparison to PGPS. The adaptive bandwidth redistribution reduces packet losses and makes the delay distributions less skewed. The control of delays, in comparison to PGPS, is very significant when very delay sensitive traffic, such as CBR, is mixed with traffic with more laxity in delays. The low delay class has very short tails in comparison to PGPS and hence would need much lower playout buffers. Analytical explanation of our simulation results remains a challenging problem. Unless the fairness requirement of GPS is required for policy reasons, fair queueing systems can improve their performance by adaptive redistribution of excess bandwidth without losing their worst case fairness property or their ability to guarantee worst case delay bounds.

Similar extensions of our scheme can be easily applied to other types of schedulers based either on Earliest Deadline First [8], or on Service Curves [2]. Note, that both of these approaches define how packets must be served in the worst-case so as not to violate some specific delay bounds. However, there is no requirement as to how packets can be served if there is available free bandwidth in order to minimize the tails of the delay distributions. Actually, for both of the above approaches there is no harm if packets are transmitted earlier than expected. Our adaptive bandwidth redistribution techniques only define the method for selecting packets when excess bandwidth is available.

REFERENCES

[1] D. Clark, S. Shenker, L. Zhang, "Supporting Real-Time Applications in an Integrated Services Packet Network: Architecture and Mechanism", *Proceedings of ACM SIGCOMM 1992*, pp. 14-23.

[2] R. Cruz, "Quality of service guarantees in virtual circuit switched networks," *IEEE Journal on Selected Areas In Communications*, vol. 13, pp. 1048–1056, August 1995.

[3] A. Demers, S. Keshav, S. Shenker, "Analysis and Simulation of a Fair Queueing Algorithm" *Internetworking: Research and Experience*, pp, 3-26, vol. 1, 1990.

[4] N. Duffield, T.V. Lakshman, D. Stiliadis "On Adaptive Bandwidth Sharing with Rate Guarantees," Bell Laboratories Technical Memorandum 113470-971215-07TM.

[5] A. Elwalid, D. Heyman, T. V. Lakshman, D. Mitra, A. Weiss, "Fundamental Bounds and Approximations for ATM Multiplexers with Applications to Video Teleconferencing", *IEEE Journal on Selected Areas in Communications: Special Issue on Fundamental Advances in Networking*, pp. 1004-1016, August 1995.

[6] H. Gail, G. Grover, R. Guerin, S. Hantler, Z. Rosberg, M. Sidi, "Buffer size requirements under longest queue first", *Proceedings IFIP'92*, 1992

[7] L. Georgiadis, R. Guerin and A. Parekh "Optimal multiplexing on a single link: Delay and buffer requirements", in *Processdings of IEEE INFOCOM'94*, pp.524-532.

[8] L. Georgiadis, R. Guerin, V. Peris and K.N Sivarajan "Efficient network QOS provisioning based on per-node traffic shaping," in *Proceedings of IEEE INFOCOM'96*, pp.102-110.

[9] A. Birman, H.R. Gail, S.L. Hantler and Z. Rosberg, "An optimal service policy for buffer systems," *Journal of the Association for Computing Machinery*, pp. 641-57. vol. 42, no. 3, May 1995.

[10] S. J. Golestani, "A Self-Clocked Fair Queueing Scheme for Broadband Applications", *Proceedings of INFOCOM'94*, pp. 636-646., 1994.

[11] P. Goyal, H.M. Vin and H. Chen, "Start-time Fair Queing: A Scheduling Algorithm for Integrated Services Packet Switching Networks", *Proceedings ACM SIGCOMM'96*, pp. 157-169.

[12] K. R. Krishnan, "The Hurst Parameter of Non-Markovian On-Off Traffic Sources", Internal Bellcore Report, 1995.

[13] D. S. Lee, "Weighted Longest Queue First: An Adaptive Scheduling Discipline for ATM Networks", *Proceedings INFOCOM'97*, 1997.

[14] B.B. Mandelbrot and J.W. Van Ness (1968). "Fractional Brownian Motions, Fractional Noises and Applications". *SIAM Review*, **10** 422–437

[15] B.B. Mandelbrot and J.R. Wallis (1968). "Noah, Joseph, and Operational Hydrology". *Water Resour. Res.*, **4** 909–918

[16] S. Panwar, D. Towsley, J. Wolf, "Optimal scheduling policies for a class of queues with customer deadlines to the beginning of services", *Journal of ACM*, vol. 35, no. 4, pp. 832—844, 1988.

[17] A. K. Parekh R. G. Gallager, "A Generalized Processor Sharing Approach to Flow Control in Integrated Services Networks: The Multiple Node Case", *Proceedings of INFOCOM'93*, pp. 521-530, 1993.

[18] J.L. Rexford, A.G. Greenberg, F.G Bonomi, "Hardware Efficient Fair Queueing Architectures for High-Speed Networks", *Proceedings of INFOCOM'96*, pp. 120-128, 1996.

[19] D. Stiliadis, A. Varma, "Latency-Rate Servers: A General Model for Analysis of Traffic Scheduling Algorithms", *Proceedings of INFOCOM'96*, pp. 111-119, 1996.

[20] D. Stiliadis and A. Varma, "Design and analysis of Frame-based Fair Queueing: A New Traffic Scheduling Algorithm for Packet-Switched Networks", *Proceedings of ACM SIGMETRICS '96*, pp. 104-115, May 1996.

[21] S. Suri, G. Varghese and G. Chandramenon, "Leap Forward Virtual Clock: A New Fair Queuing Scheme with Guaranteed Delays and Throughput Fairness" *Proceedings INFOCOM'97*, 1997.

[22] W. Willinger, M.S. Taqqu, R. Sherman, D.V. Wilson, "Self-Similarity Through High-Variability: Statistical Analysis of Ethernet LAN Traffic at the Source Level", *Proceedings of ACM SIGCOMM 1995*.

[23] D. Yates, J. Kurose, D. Towsley, M.G. Hluchyj, "On per-session end-to-end delay distributions and the call admission problem for real-time applications with QoS requirements", *Proceedings of ACM SIGCOMM 1993*, pp.2-12, September 1993.

[24] J.C.R. Bennett, H. Zhang, "WF^2Q: Worst-case Fair Weighted Fair Queueing", *Proceedings of INFOCOM'96*, pp. 120-128, 1996.

APPENDIX

WORST-CASE FAIRNESS OF FQ-SDS

Let us denote by t_k, the times that a packet is released by the shaper to the RPS scheduler. In order to be able to calculate an exact bound for the worst-case fairness we need to account for the fluid server that the RPS system is simulating. Let us assume that $W_i^F(\tau,t)$ is the service offered by this fluid to server to connection i during an interval $(\tau,t]$. We will first prove the following lemma:

Lemma 2: Let $W_i^F(\tau,t)$ be the service received by session i in the fluid RPS system during an interval $(\tau,t]$ in which it is continuously backlogged. If at time τ a packet was released from the shaper then,

$$W_i^F(\tau,t) \geq g_i(t-\tau).$$

Proof: We know that since the beginning of the system busy period the shaper releases packets with rate equal to g_i. We also know that the connection virtual time is increasing by the normalized service offered to the connection, when the connection is not idle. Otherwise, it increases based on the system virtual time. Let us denote with $t' < \tau$, the last time that connection virtual time was updated by the use of the system virtual time. Thus, for time t',

$$v(t') = v_i(t') \tag{10}$$

Obviously, this happened after an idle period of connection, where the connection virtual time has remained behind the system virtual time. This update was initiated by a packet arrival. That means that at time t' a packet was released from the shaper. We know that the total arrivals to the RPS scheduler since time t' are bounded by

$$S_i(t',\tau) \leq g_i(\tau - t') \tag{11}$$

We also know that since time t' the connection virtual time has only increased by the normalized service offered to it. Thus,

$$v_i(\tau) - v_i(t') = \frac{W_i^F(t',\tau)}{g_i} \tag{12}$$

$$\leq \frac{S_i(t',\tau)}{g_i} \quad \text{From Eq.11} \tag{13}$$

$$\leq (\tau - t') \tag{14}$$

We also know that the system virtual time is increasing with a rate at least linear to the real time, or

$$v(\tau) \geq v(t') + (\tau - t'). \tag{15}$$

From Eq. (10),(14) and (15) we can conclude that

$$v_i(\tau) \leq v(\tau) \tag{16}$$

Thus when a packet is released from the shaper at time τ, the connection virtual time $v_i(\tau)$ will be updated to a value no larger to that of the system virtual time. During the interval $(\tau,t]$, the connection virtual time is increasing by the normalized service offered to that connection. Thus,

$$v_i(t) - v_i(\tau) = \frac{W_i^F(\tau,t)}{g_i}. \tag{17}$$

Therefore,

$$W_i^F(\tau,t) = g_i(v_i(t) - v_i(\tau)). \quad (18)$$

Since $v_i(t) \geq v(t)$ and $v_i(\tau) \leq v(\tau)$, we can re-write this as

$$W_i^F(\tau,t) \geq g_i(v(t) - v(\tau)). \quad (19)$$

During the interval $(\tau,t]$, the system potential is increasing with a rate at least equal to that of real time. That is,

$$v(t) - v(\tau) \geq t - \tau. \quad (20)$$

From equations (19) and (20), we have $W_i^F(\tau,t) \geq g_i(t-\tau)$. □ We can now evaluate the worst-case fairness index of the packet-based system.

Theorem 1: The queueing delay of a session-i packet arriving at time τ in a packet FQ-SDS system is upper-bounded by

$$d_i^P(\tau) \leq \frac{Q_i^P(\tau)}{g_i} + \frac{L_{max}}{r} + L_i(\frac{1}{g_i} - \frac{1}{r}), \quad (21)$$

where $Q_i^P(\tau)$ is the backlog of session i at time τ, L_i the maximum packet size of session i, and L_{max} the maximum packet size across all sessions.

Proof: Let us denote with t_k the time at which the k-th packet of session i is released from the shaper to the scheduler. Note that this time is identical in both the fluid and packet versions considered. We will prove the theorem for each instant t in the interval $(t_k, t_{k+1}]$. Let, $Q_i^F(t)$ be the backlog in the fluid server at time t and $Q_i^P(t)$ that in the packet-by-packet server. By Lemma 2, the service offered to session i after time t_k is with a rate at least equal to g_i, and without any latency.

Let us assume that the backlog of connection i is cleared at some time t^* in the fluid server. Then we can write

$$\begin{aligned} Q_i^F(t) &= W_i^F(t,t^*) \\ &= W_i^F(t,t_{k+1}) + W_i^F(t_{k+1},t^*) \\ &\geq W_i^F(t,t_{k+1}) + \max(g_i(t^* - t_{k+1}),0) \end{aligned} \quad (22)$$

We assume that if $t^* < t_{k+1}$, then $W_i^F(t_{k+1},t^*) = 0$. We can re-write Eq. (22) as

$$\begin{aligned} t^* &\leq t_{k+1} + \frac{Q_i^F(t) - W_i^F(t,t_{k+1})}{g_i} \\ &\leq (t_{k+1} - t) + t + \frac{Q_i^F(t) - W_i^F(t,t_{k+1})}{g_i}. \end{aligned} \quad (23)$$

The above equation holds for the fluid sever. The time to clear the backlog at time t in the fluid server is thus given by $d_i^F(t) = t^* - t$. The packet sever may lag the fluid server at most by the transmission time of one packet [19]. Therefore, the time $d_i^v(t)$ needed to clear the backlog of connection i in the packet-by-packet server after time t is bounded by

$$\begin{aligned} d_i^v(t) &\leq d_i^F(t) + \frac{L_{max}}{r} \\ &\leq \frac{Q_i^F(t) - W_i^F(t,t_{k+1})}{g_i} + (t_{k+1} - t) + \frac{L_{max}}{r} \end{aligned} \quad (24)$$

To evaluate the above expression, we will consider two separate cases.

Case 1: $Q_i^v(t) \geq Q_i^F(t)$. Let L_i^k denote the size of the kth packet of session i. Then,

$$\begin{aligned} W_i^F(t,t_{k+1}) &= \frac{L_i^k}{g_i} - W_i^F(t_k,t) \\ &\geq \frac{L_i^k}{g_i} - \min(L_i^k, r(t-t_k)). \end{aligned} \quad (25)$$

From Eq. (24) and (25) we can write

$$d_i^v(t) \leq (t_{k+1} - t) + \frac{Q_i^F(t)}{g_i} - \frac{L_i^k}{g_i} + \min(L_i^k, r(t-t_k)) + \frac{L_{max}}{r}. \quad (26)$$

By using the hypothesis that $Q_i^P(t) \geq Q_i^F(t)$, this becomes

$$d_i^P(t) \leq (t_{k+1} - t) + \frac{Q_i^P(t)}{g_i} - \frac{L_i^k}{g_i} + \frac{\min(L_i^k, r(t-t_k))}{g_i} + \frac{L_{max}}{r}. \quad (27)$$

The right-hand side of the above equation is maximized when $L_i^k = r(t - t_k)$. Thus,

$$\begin{aligned} d_i^P(t) &\leq (t_{k+1} - t_k - \frac{L_i^k}{r}) + \frac{Q_i^P(t)}{g_i} - \frac{L_i^k}{g_i} + \frac{L_i^k}{g_i} + \frac{L_{max}}{r} \\ &\leq \frac{Q_i^P(t)}{g_i} + L_i^k(\frac{1}{g_i} - \frac{1}{r}) + \frac{L_{max}}{r} \\ &\leq \frac{Q_i^P(t)}{g_i} + L_i(\frac{1}{g_i} - \frac{1}{r}) + \frac{L_{max}}{r}. \end{aligned} \quad (28)$$

Case 2: $Q_i^F(t) > Q_i^P(t)$. In this case, let $\Delta Q = Q_i^F(t) - Q_i^P(t)$. We can re-write Eq.(24) as

$$d_i^P(t) \leq (t_{k+1} - t) + \frac{Q_i^P(t) + \Delta Q}{g_i} - \frac{W_i^F(t,t_{k+1})}{g_i} + \frac{L_{max}}{r}. \quad (29)$$

At time t, the packet-by-packet server has offered more service to connection i than the fluid server. However, packet $k+1$ has not yet been released by the shaper. Thus, the additional service that fluid server has to offer until time t_{k+1} is equal to the additional service that the packet-by-packet server has offered until time t, plus the service it will offer until time t_{k+1}. That is,

$$W_i^F(t,t_{k+1}) = \Delta Q + W_i^P(t,t_{k+1}). \quad (30)$$

From Eq. (29) and (30),

$$d_i^P(t) \leq (t_{k+1} - t) + \frac{Q_i^P(t)}{g_i} - \frac{W_i^P(t,t_{k+1})}{g_i} + \frac{L_{max}}{r}. \quad (31)$$

Notice that at time t_k, the packet-by-packet server can only be behind the fluid server. Let us assume without loss of generality, that

$$W_i^F(0,t_k) = W_i^P(0,t_k) + \Delta W. \quad (32)$$

For the service offered to connection i by the packet-by-packet server after time t, we can write

$$\begin{aligned} W_i^P(t,t_{k+1}) &= W_i^P(t_k,t_{k+1}) - W_i^P(t_k,t) \\ &\geq L_i^k + \Delta W - W_i^P(t_k,t) \\ &\geq L_i^k + \Delta W - \min(L_i + \Delta W, r(t-t_k)). \end{aligned} \quad (33)$$

Thus, Eq. (31) becomes

$$\begin{aligned} d_i^P(t) &\leq (t_{k+1} - t) + \frac{Q_i^P(t)}{g_i} \\ &\quad -(L_i + \Delta W) + \min(L_i^k + \Delta W, r(t-t_k)) + \frac{L_{max}}{r}. \end{aligned} \quad (34)$$

The right-hand side of the above equation is maximized when $L_i + \Delta W = r(t - t_k)$. Therefore,

$$\begin{aligned} d_i^P(t) &\leq t_{k+1} - t_k - \frac{L_i^k + \Delta W}{r} + \frac{Q_i^P(t)}{g_i} + \frac{L_{max}}{r} \\ &\leq \frac{L_i^k}{g_i} - \frac{L_i^k + \Delta W}{r} + \frac{Q_i^P(t)}{g_i} + \frac{L_{max}}{r} \\ &\leq \frac{Q_i^P(t)}{g_i} + \frac{L_{max}}{r} + L_i^k(\frac{1}{g_i} - \frac{1}{r}) \\ &\leq \frac{Q_i^P(t)}{g_i} + \frac{L_{max}}{r} + L_i(\frac{1}{g_i} - \frac{1}{r}). \end{aligned} \quad (35)$$

This concludes the proof of Theorem 1. Note that the worst-case fairness index of $(L_{max}/r) + L_i/g_i - L_i/r$ given by the theorem is identical to that of WF2Q [24].

IEEE INFOCOM '98

The Conference on Computer Communications

Proceedings Volume 3

Seventeenth Annual Joint Conference of the IEEE Computer and Communications Societies

Gateway to the 21st Century

Sponsored by
IEEE Computer Society
IEEE Communications Society

Tutorials
March 29 to 30, 1998

Technical Sessions
March 31 to April 2, 1998

Conference Committee
General Chair: Roch Guerin, IBM T.J. Watson Research Center
Technical Program Chair: Ian F. Akyildiz, Georgia Institute of Technology

Hotel Nikko
San Francisco, CA, USA

IEEE

IEEE COMMUNICATIONS SOCIETY

Proceedings

IEEE INFOCOM '98

The Conference on Computer Communications

Volume 3

Seventeenth Annual Joint Conference of the IEEE Computer and Communications Societies

Gateway to the 21st Century

---

29 March - 2 April 1998
Hotel Nikko, San Francisco, CA, USA

INFOCOM'98 gratefully acknowledges the generous travel support provided by the National Science Foundation.

Sponsored by
IEEE Computer Society
IEEE Communications Society

IEEE

IEEE COMMUNICATIONS SOCIETY

TK5105.5
.I18
1998
v.3

Proceedings IEEE INFOCOM'98 The Conference on Computer Communications

Abstracting is permitted with credit to the source. Libraries are permitted to photocopy beyond the limit of U.S. copyright law, for private use of patrons, those articles in this volume that carry a code at the bottom of the first page, provided the per-copy fee indicated in the code is paid through the Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923. For other copying, reprint, or republication permission, write to the IEEE Copyrights Manager, IEEE Operations Center, 445 Hoes Lane, Piscataway, NJ 08855-1331. All rights reserved. Copyright © 1998 by The Institute of Electrical and Electronics Engineers, Inc.

IEEE Catalog Number: 98CH36169 (softbound)
98CB36169 (casebound)

ISBN Softbound: 0-7803-4383-2
Casebound: 0-7803-4384-0
Microfiche: 0-7803-4385-9
CD-Rom: 0-7803-4386-7

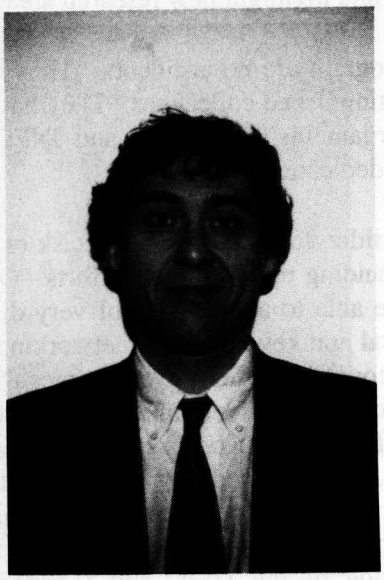
ISSN 0743-166X

Additional copies of this publication are available from

IEEE Operations Center
P. O. Box 1331
445 Hoes Lane
Piscataway, NJ 08855-1331 USA

1-800-678-IEEE
1-908-981-1393
1-908-981-9667 (FAX)
833-233 (Telex)
email: customer.services@ieee.org

Message from the General Chair



Roch Guerin

These are exciting times! The field of computer communications and networking is by all accounts enjoying both unprecedented growth and an incredible rate of technical innovation. Capabilities that only a few years back seemed to belong to the distant future or considered esoteric research topics, are rapidly moving into the reality of commercial offerings. Gigabit switches, be it ATM, Ethernet, Fiber Channel, etc., as well as routers are being announced by established companies and startups alike. High speed networking chipsets capable of sophisticated buffer management and scheduling functions are being developed and offered by numerous vendors. Ubiquitous access and mobility are fast becoming a reality through a combination of new protocols such as mobile IP, DHCP, etc., and infrastructure upgrades such as worldwide satellite networks with broadband versions soon to follow. Wavelength Division Multiplexing systems are being deployed and increasing many folds the capacity of network links. And the list could go on.

This incredible pace of progress and changes is presenting new opportunities and new challenges to the networking research community. The explosion of the capabilities and performance of computer networks is both fostering the creation of new applications and subsuming many of the existing networking technologies. Understanding those emerging applications while migrating existing functions calls for new solutions, techniques, and technologies. Additionally, keeping up with the current pace of innovation requires that we keep on looking forward to develop what will be the next generation technology, protocols, and applications.

The INFOCOM conference is squarely at the junction of all these activities. This is reflected not only in the growing success of the conference, which keeps breaking record after record every year, but also in the quality and diversity of its technical program. Seven outstanding tutorials by renowned experts in their fields, will provide attendees with opportunities to gain an in-depth understanding of the core

LINDA HALL LIBRARY
KANSAS CITY, MISSOURI

Table of Contents

Volume 1 Tuesday

Session 1A: Quality of Service Routing
Session Chair: Martha Steenstrup, BBN/GTE

QoS Routing in Networks with Uncertain Parameters 3
D. H. Lorenz and A. Orda

QoS Routing Via Multiple Paths Using Bandwidth Reservation 11
N. S. V. Rao and S. G. Batsell

Bounds on End-to-End Performance via Greedy, Multi-Path Routing in Integrated Services Networks 19
S. Mithal

Routing with End to End QoS Guarantees in Broadband Networks 27
A. Orda

Session 1B: Channel Allocation in Wireless Networks
Session Chair: Tony Ephremides, University of Maryland

Effective Bandwidths in Wireless Networks with Multiuser Receivers 35
D. Tse and S. Hanly

Teletraffic Issues Related to Channel Allocation in Digital Mobile Cellular Networks 43
P. L. Hiew and M. Zukerman

Performance Bounds for Dynamic Channel Assignment Schemes Operating under Varying Re-Use Constraints 51
P. Whiting and S. Borst

Channel Assignment Algorithms Satisfying Co-Channel and Adjacent Channel Reuse Constraints in Cellular Mobile Networks 59
S. Sarkar and K. N. Sivarajan

Session 1C: WDM Networks
Session Chair: Bo Li, Hong Kong University of Science and Technology

Cost Effective Traffic Grooming in WDM Rings 69
O. Gerstel, R. Ramaswami, and G. Sasaki

Dynamic Load Balancing in Broadcast WDM Networks with Tuning Latencies 78
I. Balldine and G. N. Rouskas

Unscheduled Multicasts in WDM Broadcast-and-Select Networks 86
E. Modiano

Wavelength Assignment in a WDM Ring: To Minimize Cost of Embedded SONET Rings 94
O. Gerstel, P. Lin, and G. Sasaki

Session 1D: Switching I
Session Chair: Mark Karol, Bell Labs, Lucent Technologies

Minimum Distance Routing in the Bidirectional Shufflenet 102
M. Gerla, P. Palnati, E. Leonardi, and F. Neri

The Multi-Dimensional Shuffle-Exchange Network: A Novel Topology for Regular Network Architectures 110
P. P. To and T. T. Lee

Strictly and Wide-Sense Nonblocking Photonic Switching Under Crosstalk Constraint 118
M. Vaez and C.-T. Lea

Strict-Sense Non-Blocking Conditions for Shuffle/Exchange Networks with Vertical Replication 126
I. Busi and A. Pattavina

Session 2A: Wireless LANs
Session Chair: Vaduvur Bharghavan, University of Illinois Urbana-Champaign

Stability and Performance Analysis of HIPERLAN 134
G. Anastasi, L. Lenzini, and E. Mingozzi

IEEE 802.11 Wireless LAN: Capacity Analysis and Protocol Enhancement 142
F. Calì, M. Conti, and E. Gregori

A Comparison of MAC Protocols for Wireless Local Networks Based on Battery Power Consumption 150
J.-C. Chen, K. M. Sivalingam, P. Agrawal, and S. Kishore

A Near-Optimum Channel Access Protocol Based on Incremental Collision Resolution and Distributed Transmission Queues 158
R. Garcés and J.J. Garcia-Luna-Aceves

Session 2B: Rate Based Congestion Control
Session Chair: Ramesh Nagarajan, Bell Labs, Lucent Technologies

Robust Rate Control for ABR Sources 166
E. Altman, T. Basar, and R. Srikant

LAPLUS: An Efficient, Effective and Stable Switch Algorithm for Flow Control of the Available Bit Rate ATM Service 174
S. Prasad, K. Kiasaleh, and P. Balsara

A Control-Theoretic ABR Explicit Rate Algorithm for ATM Switches with Per-VC Queueing 183
L. Benmohamed and Y. T. Wang

Zero Queueing Flow Control and Applications 192
H. T. Kung and S. Y. Wang

Session 2C: Video Modeling and Coding
Session Chair: Jorg Liebeherr, Polytechnic University, Brooklyn

Modeling Video Traffic in the Wavelet Domain 201
S. Ma and C. Ji

Threshold Autoregressive Models for VBR MPEG Video Traces 209
B. S. Jang and C. Thomson

An Accurate and Treatable Markov Model of MPEG-Video Traffic 217
A. Lombardo, G. Morabito, and G. Schembra

Robust H.263 Video Coding for Transmission over the Internet 225
M. H. Willebeek-LeMair, Z.-Y. Shae, and Y.-C. Chang

Session 2D: End-to-End Protocol Performance
Session Chair: Kang Shin, University of Michigan, Ann Arbor

Timer Reconsideration for Enhanced RTP Scalability 233
J. Rosenberg and H. Schulzrinne

Explicit Window Adaptation: A Method to Enhance TCP Performance 242
L. Kalampoukas, A. Varma, and K.K. Ramakrishnan

TCP Behavior of a Busy Internet Server: Analysis and Improvements 252
H. Balakrishnan, V. N. Padmanabhan, S. Seshan, M. Stemm, and R. H. Katz

TCP Fast Recovery Strategies: Analysis and Improvements 263
D. Lin and H.T. Kung

Session 3A: QoS Based Scheduling
Session Chair: James Roberts, France Telecom

Implementing Fair Queueing in ATM Switches: The Discrete-Rate Approach 272
F. M. Chiussi and A. Francini

Implementing Distributed Packet Fair Queueing in a Scalable Switch Architecture 282
D. C. Stephens and H. Zhang

Decoupled Generalized Processor Sharing: A Fair Queueing Principle for Adaptive Multimedia Applications 291
F. Toutain

Design Considerations for Supporting TCP with Per-Flow Queueing 299
B. Suter, T.V. Lakshman, D. Stiliadis, and A. Choudhury

Session 3B: Wireless Multiple Access
Session Chair: Jim Freebersyser, Army Research Office

Remote-Queueing Multiple Access (RQMA): Providing Quality of Service for Wireless Communication 307
N. R. Figueira and J. Pasquale

D^2MA: A Distributed Access Protocol for Wireless ATM Networks 315
M. Listanti, F. Mascitelli, and A. Mobilia

A Five-Phase Reservation Protocol (FPRP) for Mobile Ad Hoc Networks 322
C. Zhu and M. S. Corson

Judicious Use of Redundant Transmissions Multi-Channel ALOHA Networks with Deadlines 332
Y. Birk and Y. Keren

Session 3C: Performance Analysis I
Session Chair: Udo Krieger, Deutsche Telekom

Cell Loss Asymptotics in Buffer Fed with a Large Number of Independent Stationary Sources 339
N. Likhanov and R. R. Mazumdar

Packet Multiplexers with Adversarial Regulated Traffic 347
S. Rajagopal, M. Reisslein, and K. W. Ross

Computing Queue-Length Distributions for Power-Law Queues 356
M. Roughan, D. Veitch, and M. Rumsewicz

New Bounds and Approximations Using Extreme Value Theory for the Queue Length Distribution in High-Speed Networks 364
J. Choe and N. B. Shroff

Session 3D: IP over ATM Networks
Session Chair: Josep Sole-Pareta, University Politecnico de Cataluna

Performance Issues in VC-Merge Capable Switches for IP over ATM Networks 372
I. Widjaja and A. I. Elwalid

Adaptive Resource Management for Flow-Based IP/ATM Hybrid Switching Systems 381
H. Che, S.-Q. Li, and A. Lin

Flow-Aware Gateway Support for IP-over-ATM 390
K. Lee and A. Fisher

A Simulation Study of Packet Forwarding Methods over ATM: SBR Evaluation 401
F. Hoymany and D. Mossé

Panel: Evolution of Internet and Telecommunications
Organizer: Allison Mankin, USC/ISI-East

Panelists from ISP, telco, and Internet development communities discuss the evolution and growth of the communications fabric:
- Services (data, phone, etc.)
- Reliability
- Growth
- Who will be the providers?
- What technologies will effect the most changes?

Volume 2 Wednesday

Session 4A: Network Security, Management and Pricing
Session Chair: Irene Katzela, University of Toronto

Compact and Unforgeable Key Establishment over an ATM Network 411
Y. Zheng and H. Imai

Designing a Distributed Authorization Service 419
T. Y. C. Woo and S. S. Lam

CMIS/P++: Extensions to CMIS/P for Increased Expressiveness and Efficiency in the Manipulation of Management Information 430
G. Pavlov, A. Liotta, P. Abbi, and S. Ceri

Incentive-Compatible Pricing Strategies in Noncooperative Networks 439
Y. A. Korilis, T. A. Varvarigou, and S. R. Ahuja

Session 4B: Protocol Verification and Testing
Session Chair: Hasan Ural, University of Ottawa

An Extension to Concurrent TTCN 447
M. Xu and J. Wu

Promela++: A Language for Constructing Correct and Efficient Protocols 455
A. Basu, G. Morrisett, and T. von Eicken

User-Level Protocol Servers with Kernel-Level Performance 463
J. C. Brustoloni and P. Steenkiste

Towards Formal Semantics for QoS Support 472
J.-P. Richter and H. de Meer

Session 4C: Internet
Session Chair: Jon Crowcroft, University College of London

Adding Voice to a Distributed Game on the Internet 480
J. Bolot and S. F. Parisis

Internet Telephony Gateway Location 488
J. Rosenberg and H. Schulzrinne

Measurement and Analysis of Long-Range Dependent Behavior of Internet Packet Delay 497
M. S. Borella and G. B. Brewster

Simulation of FEC-Based Error Control for Packet Audio on the Internet 505
M. Podolsky, C. Romer, and S. McCanne

Session 4D: Performance Analysis II
Session Chair: Tulin Atmaca, Int-Evry

Performance Analysis of Feedback Controlled Data Packet Transmission over High-Speed Networks 516
Y. Kim and S. Li

The Cost of Qualtity in Networks of Aggregate Traffic 525
N. G. Duffield and S. H. Low

Linear Complexity Algorithms for Maximum Throughput in Radio Networks and Input Queued Switches 533
L. Tassiulas

Transport Layer Adaptable Rate Control (TARC): Analysis of a Transport Layer Flow Control Mechanism for High Speed Networks 540
R. Y. C. Brockett and I. Rubin

Session 5A: Adaptive Wireless Protocols
Session Chair: Aubrey Bush, NSF

Adaptive Use of Error-Correcting Codes for Real-Time Communication in Wireless Networks 548
M. Elaoud and P. Ramanathan

Adaptive Resource Management for DS-CDMA Networks Subject to Energy Constraints 556
S.-J. Oh and K. M. Wasserman

Adaptive Frame Length Control for Improving Wireless Link Throughput, Range and Energy Efficiency 564
P. Lettieri and M. B. Srivastava

Controlled Multimedia Wireless Link Sharing via Enhanced Class-Based Queueing with Channel-State-Dependent Packet Scheduling 572
C. Fragouli, V. Sivaraman, and M. Srivastava

Session 5B: Active Networks
Session Chair: Ken Calvert, Georgia Tech

Active Reliable Multicast 581
L.-W. Lehman, S. Garland, and D. L. Tennenhouse

Improving the Performance of Distributed Applications Using Active Networks 590
U. Legedza, D. Wetherall, and J. Guttag

Self-Organizing Wide-Area Network Caches 600
S. Bhattacharjee, K. Calvert, and E. W. Zegura

DAN: Distributed Code Caching for Active Networks 609
D. Decasper and B. Plattner

Session 5C: QoS and Resource Allocation in Integrated Services Networks
Session Chair: Geoffrey Xie, Naval Postgraduate School

Bandwidth Allocation for Guaranteed versus Best Effort Service Categories 617
E. Altman, A. Orda, and N. Shimkin

SCED+: Efficient Management of Quality of Service Guarantees 625
R. L. Cruz

Enforceable Quality of Service Guarantees for Bursty Traffic Streams 635
E. W. Knightly

On Statistical Multiplexing, Traffic Mixes, and VP Management 643
C.-F. Su and G. de Veciana

Session 5D: Video Communication
Session Chair: Giovanni Pacifici, IBM T.J. Watson Research Center

Scheduling for Large-Scale On-Demand Data Broadcasting 651
D. Aksoy and M. Franklin

A Network-Conscious Approach to End-to-End Video Delivery over Wide Area Networks Using Proxy Servers 660
Y. Wang, Z.-L. Zhang, D. H. C. Du, and D. Su

Some Proposals to Improve Error Resilience in the MPEG-2 Video Transmission over ATM Networks 668
P. Cuenca, A. Garrido, F. Quiles, and L. Orozco-Barbosa

A General Optimal Video Smoothing Algorithm 676
Z. Jiang and L. Kleinrock

Session 6A: Scalable Routing and Multi-Casting
Session Chair: Bulent Yener, New Jersey Institute of Technology

Hierarchical Source Routing through Clouds 685
M. Montgomery and G. de Veciana

Distributed Top-Down Hierarchy Construction 693
D. G. Thaler and C. V. Ravishanka

Hierarchical Routing Using Link Vectors 702
J. Behrens and J.J. Garcia-Luna-Aceves

Forwarding State Reduction for Sparse Mode Multicast Communication 711
J. Tian and G. Neufeld

Session 6B: Mobility Management
Session Chair: Christopher Rose, Rutgers University

Minimizing the Wireless Cost of Tracking Mobile Users: An Adaptive Threshold Scheme 720
Z. Naor and H. Levy

Efficient PCS Call Setup Protocols 728
Y. Cui, D. Lam, J. Widom, and D. C. Cox

Update and Search Algorithms for Wireless Two-Way Messaging: Design and Performance 737
T. Woo, T. F. La Porta, J. Golestani, and N. Agarwal

Architectures and Handoff Schemes for CATV-Based Personal Communications Network 748
N.-F. Huang, C.-A. Su, and H.-C. Chao

Session 6C: End Systems
Session Chair: Simon Lam, University of Texas - Austin

The END: A Network Adapter Design Tool 756
A. Indiresan, A. Mehra, and K. G. Shin

Implementing Protocols in Java: The Price of Portability 765
B. Krupczak, M. Ammar, and K. Calvert

A Comparative Study of High Speed Networks 774
J. Hall, R. Sabatino, S. Crosby, I. Leslie, and R. Black

A Novel Server Selection Technique for Improving the Response Time of a Replicated Service 783
Z. Fei, S. Bhattacharjee, E. W. Zegura, and M. Ammar

Session 6D: Panel - Broadband Wireless Networks
Organizer: Suresh Singh, University of South Carolina

Discuss services to be provided, possible architectures and support for Mobility and QoS in broadband wireless networks. Additional issues include: extending connectivity to users who roam between heterogenous wireless networks (wireless ATM to cellular to satellite); can QoS guarantees be translated (adapted) to changing environments.

Session 7A: Switching II
Session Chair: Murat Azizoglu, University of Washington

A Practical Scheduling Algorithm to Achieve 100% Throughput in Input-Queued Switches 792
A. Mekkittikul and N. McKeown

A Multicast Single-Queue Switch with a Novel Copy Mechanism 800
M. Hashemi and A. Leon-Garcia

Doubling Memory Bandwidth for Network Buffers 808
Y. Joo and N. McKeown

Multiconfiguration Multihop Protocols (MMPs): A New Class of Protocols for Packet-Switched WDM Optical Networks 816
J. P. Jue and B. Mukherjee

Session 7B: ATM Networks
Session Chair: Ioanis Nikolaidis, University of Alberta

OPENET: An Open and Efficient Control Platform for ATM Networks 824
I. Cidon, T. Hsiao, A. Khamisy, A. Parekh, R. Rom, and M. Sidi

UNITE - An Architecture for Lightweight Signalling in ATM Networks 832
G. Hjalmtysson and K.K. Ramakrishnan

Intelligent Voice Smoother for VBR Voice over ATM Networks 841
P. L. Tien and M. C. Yuang

An Adaptive Connection Admission Control Policy for VBR+ Service Class 849
S. Mukherjee, D. Reininger, and B. Sengupta

Session 7C: Resource Allocation and Network Protocols
Session Chair: Joe Touch, USC/ISI

Equilibrium Allocation of Variable Resources for Elastic Traffics 858
S. H. Low

Schemes for Scheduling of Control Messages by Hierarchical Protocols 865
E. Bortnikov and R. Cohen

Admission Control for Booking Ahead Shared Resources 873
A. Greenberg and D. Wischik

Exploring the Performance Impact of QoS Support in TCP/IP Protocol Stacks 883
R. Engel, D. Kandlur, A. Mehra, and D. Saha

Session 7D: Optical Networks
Session Chair: Achille Pattavina, Politecnico Milano

Limited Wavelength Translation in All-Optical WDM Mesh Networks 893
V. Sharma and E. A. Varvarigos

On the Optimal Placement of Wavelength Converters in Wavelength-Routed Networks 902
S. Subramaniam, M. Azizoglu, and A. K. Somani

Provisioning Algorithms for WDM Optical Networks 910
M. Alanyali and E. Ayanoglu

Design of Logical Topologies: A Linear Formulation for Wavelength Routed Optical Networks with No Wavelength Changers 919
R. M. Krishnaswamy and K. N. Sivarajan

Volume 3 Thursday

Session 8A: ATM Switching
Session Chair: Chin-Tau Lea, Hong Kong University of Science and Technology

Performance of ATM Switches with Age Priority Packet Discarding under the ON-OFF Source Model 931
H.-B. Chiou and Z. Tsai

A Delay and Loss Versatile Scheduling Discipline in ATM Switches 939
J. M. Hah and M. C. Yuang

Impact Analysis of Packet-Level Scheduling on an ATM Shared-Memory Switch 947
C. Fulton, S. Li, and A. Lin

Dynamic Flow Switching, A New Communication Service for ATM Networks 955
Q. Bian, K. Shiomoto, and J. Turner

Session 8B: Reliable Multi-Casting
Session Chair: C.S. Raghavendra, The Aerospace Corporation

Optimal Multicast Feedback 964
J. Nonnenmacher and E. W. Biersack

How Bad is Reliable Multicast without Local Recovery? 972
J. Nonnenmacher, M. Lacher, M. Jung, E. Biersack, and G. Carle

Optimal Packing of Group Multi-Casting 980
S. Chen, O. Gunluk, and B. Yener

A Comparison of Server-Based and Receiver-Based Local Recovery Approaches for Scalable Reliable Multicast 988
S. K. Kasera, J. Kurose, and D. Towsley

Session 8C: Congestion Control
Session Chair: Anna Charny, Cabletron and MIT

TCP-Like Congestion Control for Layered Multicast Data Transfer 996
L. Vicisano, L. Rizzo, and J. Crowcroft

Feedback Consolidation Algorithms for ABR Point-to-Multipoint Connections in ATM Networks 1004
S. Fahmy, R. Jain, R. Goyal, B. Vandalore, S. Kalyanaraman, S. Kota, and P. Samudra

A Load Profiling Approach to Routing Guaranteed Bandwidth Flows 1014
I. Matta and A. Bestavros

Multi-Service Connection Admission Control Using Modular Neural Networks 1022
C.-K. Tham and W.-S. Soh

Session 8D: Wireless Networks
Session Chair: Andrew Campbell, Columbia University

Software Radios for Wireless Networking 1030
V. Bose, A. B. Shah, and M. Ismert

MSOCKS: An Architecture for Transport Layer Mobility 1037
D. A. Maltz and P. Bhagwat

Mobile-End Transport Protocol: An Alternative to TCP/IP Over Wireless Links 1046
K.-Y. Wang and S. K. Tripathi

Demand-based Radio Network Planning of Cellular Mobile Communication Systems 1054
K. Tutschku

Session 9A: Video II
Session Chair: T. V. Lakshman, Bell Labs, Lucent Technologies

Layered Video Multicast with Retransmissions (LVMR): Evaluation of Hierarchical Rate Control 1062
X. Li, S. Paul, and M. Ammar

Adaptive Multicast of Multi-Layered Video: Rate-Based and Credit-Based Approaches 1073
B. Vickers, C. Albuquerque, and T. Suda

End-to-End Delay of Videoconferencing over Packet Switched Networks 1084
M. Baldi and Y. Ofek

SAVE: An Algorithm for Smoothed Adaptive Video over Explicit Rate Networks 1093
N.G. Duffield, K.K. Ramakrishnan, and A. R. Reibman

Session 9B: Scheduling Support for Adaptive Bandwidth Sharing
Session Chair: Edward Knightly, Rice University

Packet Fair Queueing Algorithms for Wireless Networks with Location-Dependent Errors 1103
T. S. E. Ng, I. Stoica, and H. Zhang

Minimum Delay Self-Clocked Fair Queueing for Packet-Switched Networks 1112
F. Chiussi and A. Francini

On Adaptive Bandwidth Sharing with Rate Guarantees 1122
N. G. Duffield, T. V. Lakshman, and D. Stiliadis

Optimal Bandwidth/Delay Tradeoff for Feasible- Region-Based Scalable Multimedia Scheduling 1131
W. Zhao, T. Seth, M. Kim, and M. Willebeek-LeMair

Session 9C: Wireless Protocols
Session Chair: C.K. Toh, Hughes Systems

Handover Re-routing Schemes for Connection Oriented Services in Mobile ATM Networks 1139
B. A. J. Banh, G. J. Anido, and E. Dutkiewicz

User Agent Migration Policies in Multimedia Wireless Networks 1147
R. Ramjee, T. La Porta, J. Kurose, and D. Towsley

Data Transfer Scheme for Wireless Real-Time Communications 1156
C.-S. Wu and G.-K. Ma

A Generalized Processor Sharing Approach to Time Scheduling in Hybrid CDMA/TDMA 1164
M. A. Arad and A. Leon-Garcia

Session 9D: Reliable Multicast Protocols
Session Chair: J.J. Garcia-Luna Aceves, University of California – Santa Cruz

Efficient Rate-Controlled Bulk Data Transfer Using Multiple Multicast Groups 1172
S. Bhattacharyya, J. F. Kurose, D. Towsley, and R. Nagarajan

An End-to-End Reliable Multicast Protocol Using Polling for Scaleability 1180
M. P. Barcellos and P. D. Ezhilchelvan

An Error Control Scheme for Large-Scale Multicast Applications 1188
C. Papadopoulos, G. Parulkar, and G. Varghese

A Scalable Control Topology for Multicast Communications 1197
J. Liebeherr and B. S. Sethi

Panel - Active Networks - Hype or Next Big Thing?
Organizer: Tom Chen, GTE Labs

Active networks promise a new level of flexibility to essentially program the network for customized services or computations. This concept is quite different from the traditional view of networks, and it naturally raises controversial issues such as security and code mobility, which are the subjects of numerous current research activities.

This panel brings together a few of the prominent researchers in this new research area to present their views of what active networks will be or should (not) be.

Session 10A: Web
Session Chair: Darleen Fisher, NSF

Reducing Overhead in Flow-Switched Networks: An Empirical Study of Web Traffic 1205
A. Feldmann, J. Rexford, and R. Caceres

Scalable Delivery of Web Pages Using Cyclic Best-Effort Multicast 1214
K. C. Almeroth, M. Ammar, and Z. Fei

Techniques for Developing and Measuring High Performance Web Servers over High Speed ATM Networks 1222
J. C. Hu, S. Mungee, and D. C. Schmidt

The Network Effects of Prefetching 1232
M. Crovella and P. Barford

Session 10B: Potpourri
Session Chair: Jennifer Hou, Ohio State University

Routing Lookups in Hardware at Memory Access Speeds 1241
P. Gupta, S. Lin, and N. McKeown

IP Lookups Using Multiway and Multicolumn Search 1248
B. Lampson, V. Srinivasan, and G. Varghese

Informed-Source Coding-on-Demand (ISCOD) over Broadcast Channels 1257
Y. Birk and T. Kol

A CSMA/CD Compatible MAC for Real-Time Transmissions Based on Varying Collision Intervals 1265
O. Sharon and M. Spratt

Session 10C: Analysis Techniques for Wireless Networks
Session Chair: John Daigle, University of Mississippi

On Wireless Spectrum Estimation and Generalized Graph Coloring 1273
K. Kumaran and S. Khanna

Upper and Lower Bounds of a Class of Channel Assignment Problems in Cellular Networks 1284
A. Sen, T. Roxborough, and S. Medidi

Modeling Fast Fading Channel Dynamics for Packet Data Performance Analysis 1292
Y. Y. Kim and S. Li

Stability of a Type-II Hybrid ARQ Protocol for DS-SSMA Packet Radio Systems 1301
Q. Zhang, T. F. Wong, and J. S. Lehnert,

Session 10D: Performance Models of ATM Networks
Session Chair: Moshe Sidi, Technion, Haifa, Israel

Sizing Exit Buffers in ATM Networks under CBR Traffic 1309
H. Levy, T. Mendelson, M. Sidi, and J. Keren-Zvi

On the Impact of Long-Range-Dependent Traffic in Dimensioning ATM Network Buffer 1317
G. C. Lin and T. Suda

Per-Stream Jitter Analysis in CBR ATM Multiplexors 1325
A. Privalov and K. Sohraby

On the Capabilities of On-Off Models to Capture Arbitrary ATM Sources 1333
S. Galmes and R. Puigjaner

Session 11A: ABR and Fairness
Session Chair: Alex Kolarov, NEC America

Performance of ERICA and QFC for Transporting Bursty TCP Sources with Bursty Interfering Traffic 1341
D. P. Hong and T. Suda

Converging to Approximated Max-Min Flow Fairness in Logarithmic Time 1350
B. Awerbuch and Y. Shavitt

A Stochastic Approximation Approach for Max-Min Fair Adaptive Rate Control of ABR Sessions with MCRs 1358
S. P. Abraham and A. Kumar

A Generalized Max-Min Rate Allocation Policy and Its Distributed Implementation Using ABR Flow Control Mechanism 1366
Y. T. Hou, H. H.-Y. Tzeng, and S. S. Panwar

Session 11B: IEEE 802.14 and IEEE 802.12
Session Chair: Patrick Dowd, University of Maryland and National Security Agency

Deterministic Service Guarantees in 802.12 Networks, Part II: the Cascaded Network Case 1376
P. Kim

Evaluation of Priority and Scheduling Schemes for an IEEE 802.14 MAC Protocol Loaded by Real Traffic 1384
M. V. Ivanovich and M. Zukerman

Adaptive Control Mechanism for Cable Modems MAC Protocols 1392
D. Sala, J. O. Limb, and S. U. Khaunte

A Priority Scheme for the IEEE 802.14 MAC Protocol for Hybrid Fiber-Coax Networks 1400
M. D. Corner, N. Golmie, J. Liebeherr, and D. H. Su

Session 11C: Multicast and Multi-Path Routing Algorithms
Session Chair: Anindo Banerjea, University of Toronto

Loop-Free Multipath Routing Using Generalized Diffusing Computations 1408
W. T. Zaumen and J.J. Garcia-Luna-Aceves

An Efficient Multipath Forwarding Method 1418
J. Chen, P. Druschel, and D. Subramanian

VTDM - A Dynamic Multicast Routing Algorithm 1426
H.-C. Lin and S.-C. Lai

An Efficient Multicast Routing Algorithm for Delay-Sensitive Applications with Dynamic Membership 1433
S.-P. Hong, H. Lee, and B. H. Park

Session 11D: Performance Analysis III
Session Chair: Ness Shroff, Purdue University

A Source Model for VBR Video Traffic Based on M/G/∞ Input Processes 1441
M. Krunz and A. Makowski

Performance and Fluid Simulations of a Novel Shared Buffer Management System 1449
K. Kumaran and D. Mitra

Long-Tailed Loss Rates in a Single Server Queue 1462
P. R. Jelenkovic

Performance Analysis of a Hybrid Priority Control Scheme for Input and Output Queueing ATM Switches 1470
J. Y. Lee and Y. H. Kim

APPENDIX 0001B

Copyright United States Copyright Office

Try the Copyright Public Records System (CPRS) pilot with enhanced search features and filters.

Help | Search | History | Titles | Start Over

Public Catalog

Copyright Catalog (1978 to present) at DC4
Search Request: Left Anchored Title = IEEE Infocom
Search Results: Displaying 3 of 27 entries

previous next

Labeled View

IEEE Infocom . : proceedings.

Type of Work: Serial
Title: IEEE Infocom . : proceedings.
Title Qualifier: ALL HLDGS GO HERE
Serial Publication Year: 1998
Imprint: Silver Spring, MD : IEEE Computer Society Press, 1985-.
Description: print material.
Frequency: Annually.
Publication History: 1985-
Continues: Infocom.
Copyright Claimant: Institute of Electrical and Electronics Engineers, Inc.
Issues Registered: no. 17, 1998. In 3 vols. Created 1998; Pub. 1998-03-29; Reg. 1998-04-10; TX0004757270
Other Title: Infocom
IEEE Infocom ...
Names: Infocom
IEEE
Institute of Electrical and Electronics Engineers, Inc.

previous next

Save, Print and Email (Help Page)

Select Download Format Full Record Format for Print/Save

Enter your email address: Email

Help Search History Titles Start Over

Contact Us | Request Copies | Get a Search Estimate | Frequently Asked Questions (FAQs) about Copyright | Copyright Office Home Page | Library of Congress Home Page

APPENDIX 0001C

LIBRARY | CATALOG

1 of 1

PERIODICAL OR NEWSPAPER

Proceedings

Full Record MARC Tags

Meeting name

IEEE INFOCOM.

Main title

Proceedings / IEEE INFOCOM.

Published/Created

Silver Spring, MD : IEEE Computer Society Press, ©1982-

Request this Item | LC Find It

Links

Links — Electronic copy from HathiTrust http://catalog.hathitrust.org/Record/002370312

More Information

LCCN Permalink — https://lccn.loc.gov/86640672

Publication history — 1982-

Description — volumes : illustrations ; 28 cm

Issues for <'91-> published in three or more vol.

Current frequency	Annual
ISSN	0743-166X
Linking ISSN	0743-166X
LC classification	TK5105.5 .I32a
Variant title	Title on IEEE Computer Society Proceedings list: INFOCOM Additional title on IEEE Computer Society Proceedings list: IEEE Computer and Communications Societies, Annual Joint Conference of the
Other title	Annual Joint Conference of the IEEE Computer and Communications Societies <1988>-2005 IEEE International Conference on Computer Communications 2006-
Serial key title	Proceedings (IEEE INFOCOM)
Abbreviated title	Proceedings (IEEE INFOCOM)
Related names	IEEE Computer Society. IEEE Communications Society. IEEE INFOCOM.
LC Subjects	Computer networks--Congresses. Data transmission systems--Congresses. Telecommunication--Congresses.
Other Subjects	Computer networks. Data transmission systems. Telecommunication.
Form/Genre	Conference papers and proceedings. Bibliographies. Periodicals.
Browse by shelf order	TK5105.5
Notes	Published: Los Alamitos, CA, <'91-> Some volumes have distinctive titles. Sponsored by: the IEEE Computer Society and the IEEE

Communications Society.

Reproduction no./Source	IEEE Service Center, Publications Sales, 445 Hoes Ln., POB 1331, Piscataway, NJ 08855-1331
Additional formats	Online version: IEEE INFOCOM. Proceedings 2641-9874 (DLC) 2005262135 (OCoLC)61137311
LCCN	86640672
Invalid LCCN	sn 84009799
Dewey class no.	621
National bib agency no.	627611-8 009844672
Other system no.	(OCoLC)ocm08704142
Additional Links	https://ieeexplore.ieee.org/xpl/conhome.jsp?punumber=1000359 To access, click on INFOCOM at http://csdl2.computer.org/persagen/ProceedingByFilter.jsp?filter=I
Type of material	Periodical or Newspaper
Content type	text
Media type	unmediated
Carrier type	volume
Source of description	Description based on: 1982; title from title page. Latest issue consulted: 25th (2006).

Item Availability

CALL NUMBER	Electronic Resource
Request in	.Electronic Journal
Links available	Electronic copy from HathiTrust http://catalog.hathitrust.org/Record/002370312

CALL NUMBER	TK5105.5 .I32a Set 1
Request in	Jefferson or Adams Building Reading Rooms
Status	Not Charged
Latest receipts	[year] 2009, v. 4 (2009) [year] 2009, v. 3 (2009) [year] 2009, v. 2 (2009) [year] 2009, v. 1 (2009) [year] 2007, v. 5 (2007) [year] 2007, v. 4 (2007) [year] 2007, v. 3 (2007) [year] 2007, v. 2 (2007) [year] 2007, v. 1 (2007) ed. 2005, v. 4 (2005) ed. 2005, v. 3 (2005) ed. 2005, v. 2 (2005) ed. 2005, v. 1 (2005) ed.23, v. 4 (2004) [year] 2004, v. 3 (2004) ed. 2004, v. 2 (2004) ed. 2004, v. 1 (2004) [year] 2003, v. 3 [year] 2003, v. 2 [year] 2003, v. 1 [year] 2002, v. 3 [year] 2002, v. 2 [year] 2002, v. 1
Older receipts	1983-2001

LIBRARY | CATALOG

1 of 1

PERIODICAL OR NEWSPAPER

Proceedings

Full Record **MARC Tags**

000		03928cas a2200769 a 4500
001		11311812
005		20231025065735.0
008		820819c19829999cauar b 1 a0eng c
010	__	|**a** 86640672 |**z** sn 84009799
016	7_	|**a** 627611-8 |**2** DE-600
016	7_	|**a** 009844672 |**2** Uk
022	0_	|**a** 0743-166X |**l** 0743-166X |**2** 1
035	__	|**a** (OCoLC)ocm08704142
037	__	|**b** IEEE Service Center, Publications Sales, 445 Hoes Ln., POB 1331, Piscataway, NJ 08855-1331
040	__	|**a** MCA |**b** eng |**c** MCA |**d** NSD |**d** OCL |**d** NST |**d** AIP |**d** NST |**d** DLC |**d** NST |**d** DLC | **d** NST |**d** WAU |**d** NST |**d** ELM |**d** NST |**d** NSD |**d** MYG |**d** OCLCQ |**d** MYG |**d** CGU | **d** CUI |**d** TJC |**d** OCLCQ |**d** HEBIS |**d** UKMGB |**d** OCLCQ |**d** OCLCF |**d** OCLCO |**d** OCL | **d** OCLCO |**d** I8M |**d** OCLCQ |**d** CSJ |**d** OCLCQ |**d** U3G |**d** NJR |**d** DLC |**d** NBS
042	__	|**a** pcc |**a** nsdp
050	00	|**a** TK5105.5 |**b** .I32a
082	10	|**a** 621 |**2** 12
111	2_	|**a** IEEE INFOCOM.
210	0_	|**a** Proceedings |**b** (IEEE INFOCOM)
222	_0	|**a** Proceedings |**b** (IEEE INFOCOM)
245	10	|**a** Proceedings / |**c** IEEE INFOCOM.
246	1_	|**i** Title on IEEE Computer Society Proceedings list: |**a** INFOCOM
246	1_	|**i** Additional title on IEEE Computer Society Proceedings list: |**a** IEEE Computer and

		Communications Societies, Annual Joint Conference of the
246	13	|**a** Annual Joint Conference of the IEEE Computer and Communications Societies | **f** <1988>-2005
246	13	|**a** IEEE International Conference on Computer Communications |**f** 2006-
260	__	|**a** Silver Spring, MD : |**b** IEEE Computer Society Press, |**c** ©1982-
300	__	|**a** volumes : |**b** illustrations ; |**c** 28 cm
310	__	|**a** Annual
336	__	|**a** text |**b** txt |**2** rdacontent
337	__	|**a** unmediated |**b** n |**2** rdamedia
338	__	|**a** volume |**b** nc |**2** rdacarrier
362	0_	|**a** 1982-
500	__	|**a** Published: Los Alamitos, CA, <'91->
500	__	|**a** Some volumes have distinctive titles.
515	__	|**a** Issues for <'91-> published in three or more vol.
550	__	|**a** Sponsored by: the IEEE Computer Society and the IEEE Communications Society.
588	__	|**a** Description based on: 1982; title from title page.
588	__	|**a** Latest issue consulted: 25th (2006).
590	__	|**a** SERBIB/SERLOC merged record
650	_0	|**a** Computer networks |**v** Congresses.
650	_0	|**a** Data transmission systems |**v** Congresses.
650	_0	|**a** Telecommunication |**v** Congresses.
650	_7	|**a** Computer networks. |**2** fast |**0** (OCoLC)fst00872297
650	_7	|**a** Data transmission systems. |**2** fast |**0** (OCoLC)fst00887993
650	_7	|**a** Telecommunication. |**2** fast |**0** (OCoLC)fst01145830
655	_7	|**a** Conference papers and proceedings. |**2** fast |**0** (OCoLC)fst01423772
655	_7	|**a** Bibliographies. |**2** lcgft |**0** https://id.loc.gov/authorities/genreForms/gf2014026048
655	_7	|**a** Periodicals. |**2** lcgft |**0** https://id.loc.gov/authorities/genreForms/gf2014026139
710	2_	|**a** IEEE Computer Society.
710	2_	|**a** IEEE Communications Society.
711	2_	|**a** IEEE INFOCOM.
776	08	|**i** Online version: |**a** IEEE INFOCOM. |**t** Proceedings |**x** 2641-9874 |**w** (DLC) 2005262135 |**w** (OCoLC)61137311
850	__	|**a** AU |**a** AzTeS |**a** CSt |**a** CU |**a** CU-A |**a** CU-I |**a** CaMWU |**a** CaOWtU |**a** CoFS |**a** CoU |**a** DGW |**a** DLC |**a** DeU |**a** FU |**a** GAT |**a** IaAS |**a** IaU |**a** KMK |**a** KyU |**a** LU |**a** MCM |**a** MiD |**a** MiDW |**a** MiU |**a** MnU |**a** NBuU-SE |**a** NIC |**a** NN |**a** NNC |**a** NRU |**a** NSyU |**a** NcD |**a** NcU |**a** NjR |**a** OCU |**a** OrCS |**a** PPD |**a** RPB-S |**a** TU |**a** TxArU |**a** TxHR |**a** UPB |**a** UU |**a** ViBlbV |**a** ViU |**a** WaU8
859	41	|**u** https://ieeexplore.ieee.org/xpl/conhome.jsp?punumber=1000359

859	41	|**z** To access, click on INFOCOM at |**u** http://csdl2.computer.org/persagen/ProceedingByFilter.jsp?filter=I
890	__	|**a** Proceedings; (IEEE INFOCOM) Silver Spring, MD. |**i** 86-640672
906	__	|**a** 7 |**b** cbc |**c** serials |**d** 3 |**e** ncip |**f** 19 |**g** n-oclcserc
920	__	|**a** Keep 1
984	__	|**a** srvf |**d** 2002-04-24
985	__	|**e** eserial 200406
991	__	|**b** c-GenColl |**h** TK5105.5 |**i** .I32a |**w** SERIALS
991	__	|**b** c-GenColl |**h** TK5105.5 |**i** .I32a |**p** 00003514031 |**t** Copy 1 |**v** * 8888 |**w** CCF
991	__	|**b** c-GenColl |**h** TK5105.5 |**i** .I32a |**p** 00021151531 |**t** Copy 1 |**v** 3 1992 * 8888 |**w** CCF
991	__	|**b** c-GenColl |**h** TK5105.5 |**i** .I32a |**p** 0002115152A |**t** Copy 1 |**v** 1 1992 * 8888 |**w** CCF
991	__	|**b** c-GenColl |**h** TK5105.5 |**i** .I32a |**p** 00021151518 |**t** Copy 1 |**v** 2 1992 * 8888 |**w** CCF
992	__	|**b** B/L |**w** SERLOC

Request this Item | LC Find It

Item Availability

CALL NUMBER	Electronic Resource
Request in	.Electronic Journal
Links available	Electronic copy from HathiTrust http://catalog.hathitrust.org/Record/002370312

CALL NUMBER	TK5105.5 .I32a Set 1
Request in	Jefferson or Adams Building Reading Rooms
Status	Not Charged
Latest receipts	[year] 2009, v. 4 (2009) [year] 2009, v. 3 (2009) [year] 2009, v. 2 (2009) [year] 2009, v. 1 (2009)

[year] 2007, v. 5 (2007)
[year] 2007, v. 4 (2007)
[year] 2007, v. 3 (2007)
[year] 2007, v. 2 (2007)
[year] 2007, v. 1 (2007)
ed. 2005, v. 4 (2005)
ed. 2005, v. 3 (2005)
ed. 2005, v. 2 (2005)
ed. 2005, v. 1 (2005)
ed.23, v. 4 (2004)
[year] 2004, v. 3 (2004)
ed. 2004, v. 2 (2004)
ed. 2004, v. 1 (2004)
[year] 2003, v. 3
[year] 2003, v. 2
[year] 2003, v. 1
[year] 2002, v. 3
[year] 2002, v. 2
[year] 2002, v. 1

Older receipts 1983-2001

APPENDIX 0001D

Top
Send to
Get It
Details
Virtual Browse
Links

CONFERENCE PROCEEDING

Proceedings : IEEE INFOCOM'98 : the conference on computer communications : seventeenth annual joint conference of the IEEE Computer and Communications Societies : gateway to the 21st century, 29 March - 2 April 1998, Hotel Nikko, San Francisco, CA, USA / sponsored by IEEE Computer Society, IEEE Communications Society

IEEE INFOCOM (17th : 1998 : San Francisco, Calif.); IEEE Computer Society; IEEE Communications Society

c1998

Available at Linda Hall Library Books - 2nd Floor (TK5105.5 .I18 1998) >

Send to

Email | QR | Permalink | Citation

EasyBib | Export RIS | Export to Excel

Get It

Back to Locations

LOCATION ITEMS

Linda Hall Library
Available , Books - 2nd Floor ; TK5105.5 .I18 1998
Holdings: v. 1-3

Item in place (0 requests) v.3
Loanable

Item in place (0 requests) v.2
Loanable

Item in place (0 requests) v.1
Loanable

Details

Title	Proceedings : IEEE INFOCOM'98 : the conference on computer communications : seventeenth annual joint conference of the IEEE Computer and Communications Societies : gateway to the 21st century, 29 March - 2 April 1998, Hotel Nikko, San Francisco, CA, USA / sponsored by IEEE Computer Society, IEEE Communications Society
Other title	1998 IEEE INFOCOM Seventeenth Annual Joint Conference of the IEEE Computer and Communications Societies INFOCOM'98
Creator/Contributor	IEEE INFOCOM (17th : 1998 : San Francisco, Calif.) › IEEE Computer Society › IEEE Communications Society ›
Subjects	Computer networks › Data transmission systems ›
Identifier	ISBN : 9780780343832 (softbound) ISBN : 0780343832 (softbound) ISBN : 9780780343849 (casebound) ISBN : 0780343840 (casebound) ISBN : 9780780343856 (microfiche) ISBN : 0780343859 (microfiche) ISBN : 9780780343863 (CD-Rom) ISBN : 0780343867 (CD-Rom) OCLC : (OCoLC)39833184
Publisher	Los Alamitos, Calif. : IEEE Computer Society Press
Creation date	c1998
Format	3 v. : ill. ; 28 cm.
General notes	"IEEE Catalog number: 98CH36169 (softbound)"--T.p. verso. "IEEE Catalog number: 98CB36169 (casebound)"--T.p. verso. Includes bibliographical references and author indexes.

Virtual Browse

‹ eedings :)COM '95, | Proceedings : IEEE INFOCOM'96 : | Proceedings : IEEE INFOCOM'97 : | Proceedings : IEEE INFOCC ›



Links

Display Source Record

```
leader  01766nam a2200421Ia 4500
001     992918883405961
005     20100505173500.0
008     980424s1998    nyua     b    101 0 eng d
019     ##$a39012339
020     ##$a9780780343832 (softbound)
020     ##$a0780343832 (softbound)
020     ##$a9780780343849 (casebound)
020     ##$a0780343840 (casebound)
020     ##$a9780780343856 (microfiche)
020     ##$a0780343859 (microfiche)
020     ##$a9780780343863 (CD-Rom)
020     ##$a0780343867 (CD-Rom)
035     ##$a(OCoLC)39833184
035     ##$9403757
035     ##$a(MoKL)291888-lhalldb
035     ##$a(lhalldb)291888-lhalldb
035     ##$z(OCoLC)39012339
040     ##$aNSA $cNSA $dLHL
049     ##$aLHLA
090     ##$aTK5105.5 $b.I18 1998
111     2#$aIEEE INFOCOM $n(17th : $d1998 : $cSan Francisco, Calif.)
245     10$aProceedings : $bIEEE INFOCOM'98 : the conference on computer communications : seventeenth annual joint conference of the IEE
246     14$a1998 IEEE INFOCOM
246     30$aSeventeenth Annual Joint Conference of the IEEE Computer and Communications Societies
246     30$aINFOCOM'98
260     ##$aLos Alamitos, Calif. : $bIEEE Computer Society Press, $cc1998.
300     ##$a3 v. : $bill. ; $c28 cm.
500     ##$a"IEEE Catalog number: 98CH36169 (softbound)"--T.p. verso.
500     ##$a"IEEE Catalog number: 98CB36169 (casebound)"--T.p. verso.
504     ##$aIncludes bibliographical references and author indexes.
650     #0$aComputer networks $vCongresses
650     #0$aData transmission systems $vCongresses
710     2#$aIEEE Computer Society
710     2#$aIEEE Communications Society
948     ##$aLTI 09/01/2008
948     ##$aLTI 07/14/2010
```

APPENDIX 0001E

Skip to search Skip to main content

IEEE INFOCOM '98 : the conference on computer communications : Seventeenth Annual Joint Conference of the IEEE Computer and Communications Societies : Gateway to the 21st century, 29 March-2 April 1998, Hotel Nikko, San Francisco, CA, USA.

Author

IEEE INFOCOM (17th : 1998 : San Francisco, Calif.)

Format

Book

Language

English

Published/Created

Piscataway, NJ : Available from IEEE Operations Center, ©1998.

Description

3 volumes (xxviii, 1477 pages) : illustrations ; 28 cm

Available Online

Online Content
IEEE Electronic Library (IEL)

Copies in the Library

ReCAP - Remote Storage
TK5105.5 .I48 1998
Available

Location has
Vol. 1-v. 3

Subject(s)

Computer networks—Congresses
Data transmission systems—Congresses
Telecommunication—Congresses
Wireless communication systems—Congresses
Local area networks (Computer networks)—Congresses
Asynchronous transfer mode—Congresses
Internet—Congresses

Related name

IEEE Computer Society
IEEE Communications Society

Notes

"Sponsored by IEEE Computer Society, IEEE Communications Society.

Bibliographic references

Includes bibliographical references and indexes.

Other format(s)

Also available in an electronic version.

Other title(s)

Gateway to the 21st century
Seventeenth Annual Joint Conference of the IEEE Computer and Communications Societies
Conference on Computer Communications

ISBN

0780343832 ((softbound))
9780780343832 ((softbound))
0780343840 ((casebound))
9780780343849 ((casebound))
0780343859 ((microfiche))
9780780343856 ((microfiche))
0780343867 ((CD-ROM))
9780780343863 ((CD-ROM))

OCLC

39833184

Statement on language in description

Princeton University Library aims to describe library materials in a manner that is respectful to the individuals and communities who create, use, and are represented in the collections we manage. Read more...

Ask a Question Suggest a Correction Report Harmful Language

Skip to search Skip to main content

Staff view

LEADER 03316cam a2200757Ma 4500
001 9917891843506421
005 20221219061640.0
008 980529s1998 caua b 101 0 eng d
016 7 a| 006435815 2| Uk
019 a| 39012339
020 a| 0780343832 q| (softbound)
020 a| 9780780343832 q| (softbound)
020 a| 0780343840 q| (casebound)
020 a| 9780780343849 q| (casebound)
020 a| 0780343859 q| (microfiche)
020 a| 9780780343856 q| (microfiche)
020 a| 0780343867 q| (CD-ROM)
020 a| 9780780343863 q| (CD-ROM)
035 a| (CStRLIN)NJPG98-B33686
035 9| CGA2272TS
035 a| (NjP)1789184-princetondb
035 z| (OCoLC)39012339
035 a| (OCoLC)ocm39833184
035 z| (NjP)Voyager1789184
035 a| (OCoLC)39833184 z| (OCoLC)39012339
040 a| PUL b| eng c| PUL d| OCL d| E9X d| OCLCQ d| BAKER d| ZWZ d| OCLCQ d| BDX d| OCLCQ d| OCLCO d| UKMGB d| OCLCF d| OCLCQ d| OCLCO d| LWA d| OCLCO
050 4 a| TK5105.5 b| .I53 1998
082 1 4 a| 621 2| 12
111 2 a| IEEE INFOCOM n| (17th : d| 1998 : c| San Francisco, Calif.)
245 1 0 a| IEEE INFOCOM '98 : b| the conference on computer communications : Seventeenth Annual Joint Conference of the IEEE Computer and Communications Societies : Gateway to the 21st century, 29 March-2 April 1998, Hotel Nikko, San Francisco, CA, USA.
246 3 0 a| Gateway to the 21st century
246 3 0 a| Seventeenth Annual Joint Conference of the IEEE Computer and Communications Societies
246 3 0 a| Conference on Computer Communications
260 a| Piscataway, NJ : b| Available from IEEE Operations Center, c| ©1998.
300 a| 3 volumes (xxviii, 1477 pages) : b| illustrations ; c| 28 cm
504 a| Includes bibliographical references and indexes.
500 a| "Sponsored by IEEE Computer Society, IEEE Communications Society.
530 a| Also available in an electronic version.
650 0 a| Computer networks v| Congresses.

650 0 a| Data transmission systems v| Congresses.
650 0 a| Telecommunication v| Congresses.
650 0 a| Wireless communication systems v| Congresses.
650 0 a| Local area networks (Computer networks) v| Congresses.
650 0 a| Asynchronous transfer mode v| Congresses.
650 0 a| Internet v| Congresses.
650 6 a| Réseaux d'ordinateurs v| Congrès.
650 6 a| Télécommunications v| Congrès.
650 6 a| Transmission sans fil v| Congrès.
650 6 a| Mode de transfert asynchrone v| Congrès.
650 6 a| Internet v| Congrès.
650 6 a| Réseaux locaux (Informatique) v| Congrès.
650 7 a| Asynchronous transfer mode 2| fast
650 7 a| Computer networks 2| fast
650 7 a| Data transmission systems 2| fast
650 7 a| Internet 2| fast
650 7 a| Telecommunication 2| fast
650 7 a| Wireless communication systems 2| fast
650 7 a| Réseaux d'ordinateurs x| Congrès. 2| ram
650 7 a| Mode de transfert asynchrone x| Congrès. 2| ram
650 7 a| Protocoles de réseaux d'ordinateurs x| Congrès. 2| ram
655 7 a| Conference papers and proceedings 2| fast
710 2 a| IEEE Computer Society.
710 2 a| IEEE Communications Society.
911 a| 19980604
912 a| 19990608053239.0
914 a| (OCoLC)ocm39833184 b| OCoLC c| match d| 11172022 e| processed f| 39833184
998 a| 05/29/98 s| 9115 n| NjP w| CINC98B430 d| 05/29/98 c| MWC b| MWC i| 980529 l| NJPG
AVA 0| 9917891843506421 8| 22561327490006421 a| 01PRI_INST b| recap c| Remote Storage d| TK5105.5 .I48 1998 e| available f| 3 g| 0 j| pa k| 0 p| 1 q| ReCAP t| Vol. 1-v. 3 t|
AVE 8| 53880301020006421 e| Available l| online a| 01PRI_INST 0| 9917891843506421

APPENDIX 0001F

Appendix 0001F
Citations to Duffield

Source Document

Duffield, N. G., Lakshman, T. V., & Stiliadis, D. (1998, March). On adaptive bandwidth sharing with rate guarantees. In *Proceedings. IEEE INFOCOM'98, the Conference on Computer Communications. Seventeenth Annual Joint Conference of the IEEE Computer and Communications Societies. Gateway to the 21st Century (Cat. No. 98* (Vol. 3, pp. 1122-1130). IEEE.

Citations: 1999 (in chronological order)

Boykin, M. S., & Znati, T. F. (1999, April). Design and analysis of a quantum-based QoS-aware fair share server for integrated services networks. In *Proceedings 32nd Annual Simulation Symposium* (pp. 34-42). IEEE.

Wang, X., & Meditch, J. S. (1999, May). A predictive bandwidth management scheme and network architecture for real-time VBR traffic. *Microprocessors and Microsystems*, *22*(10), 597-604.

Shakkottai, S., & Srikant, R. (1999, August). Scheduling real-time traffic with deadlines over a wireless channel. In *Proceedings of the 2nd ACM international workshop on Wireless mobile multimedia* (pp. 35-42).

Chaskar, H. M., & Madhow, U. (1999, December). Fair scheduling with tunable latency: A Round Robin approach. In *Seamless Interconnection for Universal Services. Global Telecommunications Conference. GLOBECOM'99.(Cat. No. 99CH37042)* (Vol. 2, pp. 1328-1333). IEEE.

Chiussi, F. M., & Francini, A. (1999, December). Providing QoS guarantees in packet switches. In *Seamless Interconnection for Universal Services. Global Telecommunications Conference. GLOBECOM'99.(Cat. No. 99CH37042)* (Vol. 2, pp. 1582-1590). IEEE.

Lee, D. S. (1999, December). On the optimality and the asymptotic optimality of the smallest weighted available buffer policy. *Advances in Applied Probability*, *31*(4), 1118-1150.

APPENDIX 0001G

IEEE.org IEEE *Xplore* IEEE SA IEEE Spectrum More Sites Subscribe Donate Cart Create Account Personal Sign In

Browse My Settings Help Institutional Sign In

Institutional Sign In

All

ADVANCED SEARCH

On adaptive bandwidth sharing with rate guarantees

Publisher: IEEE Cite This PDF

N.G. Duffield ; T.V. Lakshman ; D. Stiliadis **All Authors** •••

21 Cites in Papers

3 Cites in Patents

58 Full Text Views

Abstract

Authors

References

Citations

Keywords

Metrics

More Like This

Abstract: The objective of research in fair queueing schemes has been to efficiently emulate a fluid-flow generalized (weighted) processor sharing (GPS) system as closely as possib... **View more**

Metadata

Abstract:
The objective of research in fair queueing schemes has been to efficiently emulate a fluid-flow generalized (weighted) processor sharing (GPS) system as closely as possible. A primary motivation for the use of fair queueing has been its use as a means of providing bandwidth guarantees and as a consequence end-to-end delay bounds for traffic with bounded burstiness. The rate guarantees translate to scheduling weights which are set when admission control is done. A consequence of fair queueing systems closely emulating GPS is that when one or more connections are not back-logged, any "excess" bandwidth is distributed to back-logged connections in proportion to their weights. However weights are set based on the long-term requirements of traffic flows and not in any state-dependent manner that reflects instantaneous needs. We question the notion that the queueing system should closely emulate a GPS system. Instead of emulating GPS, we propose three modified scheduling schemes which preserve the rate guarantees of fair queueing (and hence preserve deterministic delay bounds) but adaptively redistribute the excess bandwidth such that either losses are reduced or delays equalized. We compare the performance of the proposed schemes to that of fair queueing using different traffic sources such as voice and video, as well as sources which have aggregate long-range dependent behavior. We find that the proposed schemes, in comparison to packet GPS (PGPS), reduce packet losses and curtail the tails of delay distributions for real-time traffic and hence permit the use of significantly smaller playout buffers for the same network load.

Published in: Proceedings. IEEE INFOCOM '98, the Conference on Computer Communications. Seventeenth Annual Joint Conference of the IEEE Computer and Communications Societies. Gateway to the 21st Century (Cat. No.98

Date of Conference: 29 March 1998 - 02 April 1998

DOI: 10.1109/INFCOM.1998.662924

Date Added to IEEE *Xplore*: 06 August 2002

Publisher: IEEE

Print ISBN:0-7803-4383-2

Conference Location: San Francisco, CA, USA

Print ISSN: 0743-166X

Authors

References



Citations

Keywords

Metrics

More Like This

Fair-rate GPS: a new class for decoupling delay and bandwidth properties
Proceedings ISCC 2002 Seventh International Symposium on Computers and Communications
Published: 2002

An Aggregated Delay|Bandwidth Star Scheme for Admission Control
2007 IEEE 18th International Symposium on Personal, Indoor and Mobile Radio Communications
Published: 2007

Show More

IEEE Personal Account

CHANGE USERNAME/ PASSWORD

Purchase Details

PAYMENT OPTIONS

VIEW PURCHASED DOCUMENTS

Profile Information

COMMUNICATIONS PREFERENCES

PROFESSION AND EDUCATION

TECHNICAL INTERESTS

Need Help?

US & CANADA: +1 800 678 4333

WORLDWIDE: +1 732 981 0060

CONTACT & SUPPORT

Follow

About IEEE *Xplore* | Contact Us | Help | Accessibility | Terms of Use | Nondiscrimination Policy | IEEE Ethics Reporting | Sitemap | IEEE Privacy Policy

A not-for-profit organization, IEEE is the world's largest technical professional organization dedicated to advancing technology for the benefit of humanity.

© Copyright 2024 IEEE - All rights reserved.

IEEE Account

» Change Username/Password

» Update Address

Purchase Details

» Payment Options

» Order History

» View Purchased Documents

Profile Information

» Communications Preferences

» Profession and Education

» Technical Interests

Need Help?

US & Canada: +1 800 678 4333

Worldwide: +1 732 981 0060

» Contact & Support

About IEEE *Xplore* | Contact Us | Help | Accessibility | Terms of Use | Nondiscrimination Policy | Sitemap | Privacy & Opting Out of Cookies

A not-for-profit organization, IEEE is the world's largest technical professional organization dedicated to advancing technology for the benefit of humanity.
© Copyright 2024 IEEE - All rights reserved. Use of this web site signifies your agreement to the terms and conditions.



APPENDIX 0002A

Journals Magazines Proceedings Books SIGs Conferences People

Conference Proceedings Upcoming Events Authors Affiliations Award Winners

Home > Conferences > MIDDLEWARE > Proceedings > MGC '04 > IQ-services: network-aware middleware for interactive large-data applications

ARTICLE

IQ-services: network-aware middleware for interactive large-data applications

Authors: Zhongtang Cai, Greg Eisenhauer, Qi He, Vibhore Kumar, Karsten Schwan, Matthew Wolf Authors Info & Claims

MGC '04: Proceedings of the 2nd workshop on Middleware for grid computing • October 2004 • Pages 11–16

• https://doi.org/10.1145/1028493.1028495

Published: 18 October 2004 Publication History

Check for updates

5 330 Get Access

MGC '04: Proceedi...
IQ-services:...
Pages 11–16

← Previous
Next →

ABSTRACT
References
Cited By
Index Terms
Recommendations

ACM DIGITAL LIBRARY

ABSTRACT

IQ-Services are application-specific, resource-aware code modules executed by data transport middleware. They constitute a 'thin' layer between application components and the underlying computational and communication resources that implements the data manipulations necessary to permit wide-area collaborations to proceed smoothly, despite dynamic resource variations. IQ-Services interact with the application and resource layers via dynamic performance attributes, and end-to-end implementations of such attributes also permit clients to interact with data providers. The joint middleware/resource and provider/consumer interactions implemented with performance attributes may be used to realize effective methods for managing the data flows in the large-data, distributed Gird applications targeted by our research. Experimental results in this paper demonstrate substantial performance improvements attained by coordinating network-level with service-level adaptations of the data being transported and by permitting end users to dynamically deploy and use applications-specific services for manipulating data in ways suitable for their current needs.

References

1. AG. Access Grid. http://www-fp.mcs.anl.gov/fl/accessgrid.

2. S. Agarwala, C. Poellabauer, J. Kong, K. Schwan, and M. Wolf. Resource-Aware Stream Management with the Customizable dproc Distributed Monitoring Mechanisms. In "Proceedings of High Performance Distributed Computing", 2003.

3. H. Balakrishnan, H. Rahul, and S. Seshan. An Integrated Congestion Management Architecture for Internet Hosts. In Proceedings of ACM SIGCOMM, 1999.

Show All References

Cited By

View all

Abbasi H, Wolf M, Eisenhauer G, Klasky S, Schwan K and Zheng F. (2010). DataStager. *Cluster Computing*. **13**:3. (277-290). Online publication date: 1-Sep-2010.

https://doi.org/10.1007/s10586-010-0135-6

Abbasi H, Wolf M, Eisenhauer G, Klasky S, Schwan K and Zheng F. DataStager. Proceedings of the 18th ACM international symposium on High performance distributed computing. (39-48).

https://doi.org/10.1145/1551609.1551618

Kumar V, Cooper B, Cai Z, Eisenhauer G and Schwan K. (2007). Middleware for enterprise scale data stream management using utility-driven self-adaptive information flows. *Cluster Computing*. **10**:4. (443-455). Online publication date: 1-Dec-2007.

https://doi.org/10.1007/s10586-007-0040-9

Show All Cited By

Index Terms

IQ-services: network-aware middleware for interactive large-data applications

Computer systems organization

Embedded and cyber-physical systems

Real-time systems

Networks

Feedback

Network properties

Network structure

Network topology types

Ring networks

Fiber distributed data interface (FDDI)

Network protocols

Transport protocols

Network services

Network monitoring

Network types

Public Internet

Recommendations

IQ-Services: network-aware middleware for interactive large-data applications: Research Articles

Middleware for Grid Computing

IQ-Services are application-specific, resource-aware code modules

Read More

IQ-RUDP: Coordinating Application Adaptation with Network Transport

HPDC '02: Proceedings of the 11th IEEE International Symposium on High Performance Distributed Computing

Our research addresses the efficient transfer of large data across wide-area networks, focusing on applications like remote visualization and real-tim...

Read More

Developing component based adaptive applications in mobile environments

SAC '05: Proceedings of the 2005 ACM symposium on Applied computing

Today, although the system supports for developing distributed applications become mature, they are inadequate for mobile environmen...

Read More

Comments

DL Comment Policy

Comments should be relevant to the contents of this article, (sign in required).

Got it

0 Comments

Share Best Newest Oldest

Nothing in this discussion yet.

View Table Of Contents

Categories

Journals
Magazines
Books
Proceedings
SIGs
Conferences
Collections
People

About

About ACM Digital Library
ACM Digital Library Board
Subscription Information
Author Guidelines
Using ACM Digital Library
All Holdings within the ACM Digital Library
ACM Computing Classification System
Digital Library Accessibility

Join

Join ACM
Join SIGs
Subscribe to Publications
Institutions and Libraries

Connect

Contact
Facebook
X
Linkedin
Feedback
Bug Report

The ACM Digital Library is published by the Association for Computing Machinery.
Copyright © 2024 ACM, Inc.
Terms of Usage | Privacy Policy | Code of Ethics







ACM DIGITAL LIBRARY
Association for Computing Machinery

Journals Magazines Proceedings Books SIGs Conferences People

Conference Proceedings Upcoming Events Authors Affiliations Award Winners

Home > *Conferences* > *MIDDLEWARE* > *Proceedings* > *MGC '04*

MGC '04: Proceedings of the 2nd workshop on Middleware for grid computing



2004 Proceeding

Conference Chairs:


Bruno Schulze,


Radha Nandkumar, + 1

Publisher: Association for Computing Machinery, New York, NY, United States

Conference: Toronto Ontario Canada • October 18 - 22, 2004

ISBN: 978-1-58113-950-1

Published: 18 October 2004

Recommend ACM DL

ALREADY A SUBSCRIBER? SIGN IN

Get Alerts for this Conference

Save to Binder

Export Citation

Bibliometrics

Citation count 224

Downloads (6 weeks) 8

Sections

MGC '04: Proceedings of the 2nd workshop on Middleware for grid computing

2004

← Previous

Next →

Abstract

Proceeding Downloads

Contributors

Recommendations

Acceptance Rates

Comments

Abstract

The Middleware Advanced Workshops program started in 2000 with the 1st Workshop on Reflective Middleware in New York. The surprising success of this first workshop, both in number of submission and number of participants, motivated the continuation of the Workshop program in the following events of the Middleware series. In Heidelberg, Middleware'2001 hosted two well attended workshops on Mobile Computing and Electronic Commerce, always focusing on research within the middleware area. In Rio, Middleware'2003 expanded the program with an open Call for Workshops that resulted in four high-quality workshops: Mobile and Ad-Hoc Computing, Reflective and Adaptive Middleware, Grid Computing, and Model-Driven Approaches.

The preparation for the 2004 edition of the Middleware workshops started in late 2003 with the Call for Workshop Proposals. We received seven good, well planned proposals and, with the help of experts in the Middleware field, four workshops were selected. The selection criteria was based on adherence to the conference theme and objectives, capacity of attracting a good number of researchers and practitioners, whether the proposal was backed by well-known, active researchers in the field, and the overall

formal quality of the proposal.

For the first time, we will host a Doctoral Symposium, equivalent to similar events in other conferences such as OOPSLA. It is organized by Edward Curry from the National University of Ireland and Doug Lea from SUNY Oswego. During the Doctoral Symposium, 8 PhD candidates will present their theses work to a first-class Mentor Committee composed of some of the top Middleware researchers in the world. The goal of the symposium is to expose the students to helpful criticism before their thesis defense. Students beginning their research are also encouraged to attend the workshop as observers to better understand how to structure a research project and learn from the experience of others. The first edition of the Symposium will be highly international with the participation of PhD candidates from Brazil, Germany, Netherlands, Spain, Switzerland, and the United Kingdom.

The Workshop on Middleware for Grid Computing was very successful in Rio and now, organized by Bruno Schulze (LNCC, Brazil) and Radha Nandkumar (NCSA/UIUC, USA), is in its second edition. It received 48 paper submissions from which 15 papers were selected. This workshop will also host a specific poster session with 9 posters. The Workshop on Middleware for Pervasive and Ad-Hoc Computing also

started well in Rio and this year, organized by Paddy Nixon from Strathclyde University (UK) grew both in quantity and quality attracting papers from the lead researchers in the international community.

Last but not least, the Workshop on Adaptive and Reflective Middleware continues very healthy. In its third edition it continues to be a venue for researchers working on the middleware core, finding new ways of structuring it to make it more flexible, adaptable, and efficient. This year, it is being organized by Fabio Costa (UFG, Brazil) and Nanbor Wang (Tech-X Corp., USA).

Proceeding Downloads

PDF Front matter (TOC, Message from the chair, The 2nd International Workshop on Middleware for Grid Computing (MGC), Program committee)

ARTICLE

Introduction to the 2nd International Workshop on Middleware for Grid Computing (MGC)

 Bruno Schulze,  Radha Nandkumar

pp

3 • https://doi.org/10.1145/1028493.1030195

As network performance has outpaced computational power and storage capacity, a new paradigm has evolved to enable the sharing and coordinated use of geographically distributed resources, popularly known as "Grid" computing. The aims are to couple ...

5

ARTICLE

Resource allocation in a middleware for streaming data

Liang Chen, Gagan Agrawal

pp 5–
10 • https://doi.org/10.1145/1028493.1028494

Increasingly, a number of applications rely on, or can potentially benefit from, analysis and monitoring of data streams. To support processing of streaming data in a grid environment, we have been developing a middleware system called GATES (...

16 782

Get Access

ARTICLE

IQ-services: network-aware middleware for interactive large-data applications

Zhongtang Cai, Greg Eisenhauer, Qi He, Vibhore Kumar, Karsten Schwan, + 1

pp 11–

16 • https://doi.org/10.1145/1028493.1028495

IQ-Services are application-specific, resource-aware code modules executed by data transport middleware. They constitute a 'thin' layer between application components and the underlying computational and communication resources that implements the data ...

5 330

Get Access

ARTICLE

GridBox: securing hosts from malicious and greedy applications

 Evgueni Dodonov,

 Joelle Quaini Sousa,

 Hélio Crestana Guardia

pp 17–22 • https://doi.org/10.1145/1028493.1028496

Security is an important concern in providing the infrastructure for the implementation of general purpose computational grids. However, most grid implementations focus their security concerns in correctly authenticating users and hosts and in the ...

12 703

Get Access

ARTICLE

Architecture for resource allocation services supporting interactive remote desktop sessions in utility grids

 Vanish Talwar,  Bikash Agarwalla,

Sujoy Basu, Raj Kumar, Klara Nahrstedt

pp 23–28 • https://doi.org/10.1145/1028493.1028497

Emerging large scale utility computing systems like Grids promise computing and storage to be provided to end users as a utility. System management services deployed in the middleware are a key to enabling this vision. Utility Grids provide a challenge ...

0 806

Get Access

ARTICLE

CoDIMS-G: a data and program integration service for the grid

V. Fontes, B. Schulze, M. Dutra, F. Porto, A. Barbosa

pp 29–34 • https://doi.org/10.1145/1028493.1028498

Grid services provide an important abstract layer on top of heterogeneous components (hardware and software) that take part into a Grid environment. In this scenario, applications, like scientific visualization, require access to data of non-...

1 579

Get Access

ARTICLE

Checkpointing-based rollback recovery for parallel applications on the

InteGrade grid middleware

Raphael Y. de Camargo,
Andrei Goldchleger, Fabio Kon,
Alfredo Goldman

pp 35–40 • https://doi.org/10.1145/1028493.1028499

InteGrade is a grid middleware infrastructure that enables the use of idle computing power from user workstations. One of its goals is to support the execution of long-running parallel applications that present a considerable amount of communication ...

14 631
Get Access

ARTICLE

A tool for the design and evaluation of hybrid scheduling algorithms for computational grids

B. A. Vianna, A. A. Fonseca,
N. T. Moura, L. T. Menezes,
H. A. Mendes, + 3

pp 41–46 • https://doi.org/10.1145/1028493.1028500

One of the objectives of computational grids is to offer applications the collective computational power of distributed but typically shared heterogeneous resources. Unfortunately, efficiently harnessing the performance potential of such systems (i.e. ...

1 1,313
Feedback

Get Access

ARTICLE

Monitoring the connectivity of a grid

Sergio Andreozzi, Augusto Ciuffoletti, Antonia Ghiselli

pp 47–51 • https://doi.org/10.1145/1028493.1028501

Grid computing is a new paradigm that enables the distributed coordination of resources and services which are geographically dispersed, span multiple trust domains and are heterogeneous. Network infrastructure monitoring, while vital for activities ...

5 539

Get Access

ARTICLE

Towards a flexible security framework for peer-to-peer based grid computing

A. Detsch, L. P. Gaspary, M. P. Barcellos, G. G. H. Cavalheiro

pp 52–56 • https://doi.org/10.1145/1028493.1028502

The dynamic, multi-organization nature of large-scale grid computing introduces security issues that must be addressed before grid systems can become widely popular. This paper proposes P2PSLF (Peer-to-Peer Security Layer Framework), a flexible security ...

14 1,809

Get Access

ARTICLE

Grid middleware services for virtual data discovery, composition, and integration

Yong Zhao, Michael Wilde, Ian Foster, Jens Voeckler, Thomas Jordan, + 2

pp 57–62 • https://doi.org/10.1145/1028493.1028503

We describe the services, architecture and application of the GriPhyN Virtual Data System, a suite of components and services that allow users to describe virtual data products in declarative terms, discover definitions and assemble workflows based on ...

26 1,481

Get Access

ARTICLE

Data pipelines: enabling large scale multi-protocol data transfers

Tevfik Kosar, George Kola, Miron Livny

pp 63–68 • https://doi.org/10.1145/1028493.1028504

Collaborating users need to move terabytes of data among their sites, often involving multiple protocols. This process is very fragile and involves considerable human

involvement to deal with failures. In this work, we propose data pipelines, an ...

5 438

Get Access

ARTICLE

Advanced delivery mechanisms in the GRelC project

Giovanni Aloisio, Massimo Cafaro, Sandro Fiore

pp 69–74 • https://doi.org/10.1145/1028493.1028505

Today many Data Grid applications need to manage and process a very large amount of data distributed across multiple grid nodes. Several applications often access large databases (i.e. protein data banks, in the bioinformatics field) without any data ...

7 354

Get Access

ARTICLE

A grid service broker for scheduling distributed data-oriented applications on global grids

Srikumar Venugopal, Rajkumar Buyya, Lyle Winton

pp 75–80 • https://doi.org/10.1145/1028493.1028506

Large communities of researchers distributed around the world are engaged in analyzing huge collections of data generated by

scientific instruments and replicated on distributed resources. In such an environment, scientists need to have the ability to ...

111 1,738

Get Access

ARTICLE

Autonomic WWW server management with distributed resources

Takuya Araki

pp 81–86 • https://doi.org/10.1145/1028493.1028507

If many people access a Web server at one time, the server might not be able to respond within an acceptable time or even provide the service. Therefore, enough servers should be assigned to a service to guarantee quality of service. But reserving a lot ...

1 571

Get Access

ARTICLE

A hierarchical process execution support for grid computing

Fábio R. L. Cicerre,

Edmundo R. M. Madeira,

Luiz E. Buzato

pp 87–92 • https://doi.org/10.1145/1028493.1028508

Grid is an emerging infrastructure used to share resources among virtual organizations

in a seamless manner and to provide breakthrough computing power at low cost. Nowadays there are dozens of academic and commercial products that allow execution of ...

1 979

Get Access

Contributors



Bruno Richard Schulze
Laboratorio Nacional de Computacao Cientifica, Petropolis

Publication Years	1997 - 2019
Publication counts	51
Citation count	146
Available for Download	15
Downloads (cumulative)	6,929
Downloads (12 months)	26
Downloads (6 weeks)	1



Radha Nandkumar
University of Illinois Urbana-Champaign

Publication Years	2004 - 2008
Publication counts	4
Citation count	5
Available for Download	0
Downloads (cumulative)	0
Downloads (12 months)	0
Downloads (6 weeks)	0

Average Downloads per Article	462	Average Downloads per Article	0
Average Citation per Article	3	Average Citation per Article	1



Fabio Kon
University of São Paulo

Publication Years	1997 - 2024
Publication counts	106
Citation count	1,126
Available for Download	36
Downloads (cumulative)	22,835
Downloads (12 months)	2,122
Downloads (6 weeks)	263
Average Downloads per Article	634
Average Citation per Article	11

Recommendations

MPAC '04: Proceedings of the 2nd workshop on Middleware for pervasive and ad-hoc computing

Read More

DSM '05: Proceedings of the 2nd international doctoral symposium on Middleware

Read More

WiSe '03: Proceedings of the 2nd ACM workshop on Wireless security

Read More

Acceptance Rates

Overall Acceptance Rate 14 of 36 submissions, 39%

Year	Submitted	Accepted	Rate
MGC '12	23	9	39%
MGC '11	13	5	38%
Overall	36	14	39%

Comments

DL Comment Policy

Comments should be relevant to the contents of this article, (sign in required).

Got it

0 Comments

Share

Best Newest Oldest

Nothing in this discussion yet.

Machinery. Copyright © 2024 ACM, Inc.
Terms of Usage | Privacy Policy | Code of Ethics







APPENDIX 0002B

FirstSearch University Libraries THE CATHOLIC UNIVERSITY OF AMERICA

WorldCat Detailed Record

| Back to CUA Libraries

- Click on a checkbox to mark a record to be e-mailed or printed in Marked Records.

Home Databases Searching Results Staff View | My Account | Options | Comments | Exit | Hide tips

List of Records Detailed Record Marked Records Saved Records Go to page

Subjects Libraries E-mail Bib Print Export Help

WorldCat results for: nb: 1581139500 and yr: 2004. Record 1 of 8.

Prev 1 Next Mark:

Detailed Record Add/View Comments

2nd Workshop on Middleware for Grid Computing : [October 18-22, 2004] /

Bruno **Schulze**; Radha **Nandkumar**; Fabio **Kon**

2004
English Internet Resource Computer File 92 pages : illustrations.
New York, NY : ACM Press, ; ISBN: 1581139500 9781581139501

GET THIS ITEM

Access: http://portal.acm.org.proxycu.wrlc.org/toc.cfm?id=1028493

Availability: Check the catalogs in your library.

- Libraries worldwide that own item: 253 WRLC

External Resources:

- Find it@CU Find or Request It at CU
- ILL@CU Request it via ILL
- Cite This Item

FIND RELATED

More Like This: Search for versions with same title and author | Advanced options ...

Find Items About: ACM Digital Library. (5)

Title: 2nd Workshop on Middleware for Grid Computing : [October 18-22, 2004] /

Author(s): Schulze, Bruno.
Nandkumar, Radha.
Kon, Fabio.

Corp Author(s): ACM Digital Library.

Conf Author(s): Workshop on Middleware for Grid Computing (2nd : 2004 : Toronto, Ont.)

Publication: New York, NY : ACM Press,

Year: 2004

Description: 92 pages : illustrations.

Language: English

Series: ACM international conference proceeding series; **Variation:** ACM international conference proceeding series.

Standard No: **ISBN:** 1581139500; 9781581139501

Access: http://portal.acm.org.proxycu.wrlc.org/toc.cfm?id=1028493

SUBJECT(S)

Descriptor: Middleware -- Congresses.
Computational grids (Computer systems) -- Congresses.
Middleware -- Congrès.
Grilles informatiques -- Congrès.
Computational grids (Computer systems)
Middleware.

Genre/Form: Conference papers and proceedings.

Note(s): Includes bibliographical references./ **Reproduction:** Electronic reproduction./ New York, N.Y. :/ ACM,/ c2004.

General Info: **Other format available:** Print version:; International Workshop on Middleware for Grid Computing (2nd : 2004).; 2nd Workshop on Middleware for Grid Computing.; New York : ACM, 2004

Class Descriptors: **LC:** QA76.76.M54

Other Titles: Second Workshop on Middleware for Grid Computing; Proceedings of the 2nd Workshop on Middleware for Grid computing.

Responsibility: organizers: Bruno Schulze, Radha Nandkumar ; [general chairs: Fabio Kon].

Material Type: Conference publication (cnp); Document (dct); Internet resource (url); eBook (ebk)

Document Type: Internet Resource; Computer File

Date of Entry: 20050321

Update: 20240127

Accession No: **OCLC:** 58539596

Database: WorldCat

Subjects Libraries E-mail Bib Print Export Help

WorldCat results for: **nb: 1581139500 and yr: 2004**. Record **1** of **8**.

English | Español | Français | عربي | 日本語 | 한국어 | 中文(繁體) | 中文(簡体) |
Options | Comments | Exit

© 1992-2024 OCLC
Terms & Conditions
For information on how your personal data is processed when you use FirstSearch, please see your institution's privacy policy.

WorldCat: 2nd Workshop on Middleware for Grid Computing

OCLC 58539596 No holdings in CUA - 253 other holdings

Books | Rec Stat c | Entered 20050321 | Replaced 20240127100239.6

Type	a	ELvl		Srce	d	Audn		Ctrl		Lang	eng
BLvl	m	Form	o	Conf	1	Biog		MRec		Ctry	nyu
		Cont	b	GPub		LitF	0	Indx	0		
Desc	a	Ills	a	Fest	0	DtSt	s	Dates	2004 ,		

007		c $b r	Functions
040		HNK ‡b eng ‡c HNK ‡d OCL ‡d OCLCQ ‡d OCLCO ‡d OCLCF ‡d COO ‡d OCLCO ‡d OCLCQ ‡d OCL ‡d OCLCO ‡d OCLCQ ‡d OCLCO ‡d OCLCQ ‡d OCLCL	Functions
020		1581139500	Functions
020		9781581139501	Functions
050	4	QA76.76.M54 $b W675 2004	Functions
090		$b	Functions
049		CUAA	Functions
111	2	Workshop on Middleware for Grid Computing $n (2nd : $d 2004 : $c Toronto, Ont.)	Functions
245	1 0	2nd Workshop on Middleware for Grid Computing : $b [October 18-22, 2004] / $c organizers: Bruno Schulze, Radha Nandkumar ; [general chairs:	Functions
246	3	Second Workshop on Middleware for Grid Computing	Functions
260		New York, NY : $b ACM Press, $c 2004.	Functions
300		92 pages : $b illustrations.	Functions
336		text $b txt $2 rdacontent	Functions
337		computer $b c $2 rdamedia	Functions
338		online resource $b cr $2 rdacarrier	Functions
490	1	ACM international conference proceeding series	Functions
504		Includes bibliographical references.	Functions
533		Electronic reproduction. $b New York, N.Y. : $c ACM, $d c2004. $n System requirements: Adobe Acrobat Reader. $n Mode of access: World Wide Web. $n Made available through ACM Digital Library under title: ACM	Functions
650	_0	Middleware ‡v Congresses.	Functions
650	_0	Computational grids (Computer systems) ‡v Congresses.	Functions
650	6	Middleware $0 (CaQQLa)201-0274185 $v Congrès. $0	Functions
650	6	Grilles informatiques $0 (CaQQLa)201-0360430 $v Congrès. $0 (CaQQLa)201-0378219	Functions



BOOK

2nd Workshop on Middleware for Grid Computing

[Place of publication not identified] ACM 2004

Online access

TOP
VIEW IT
DETAILS
SEND TO
LINKS

View it

ACM Digital Library Complete SHOW LICENSE

Details

Title	2nd Workshop on Middleware for Grid Computing
Imprint	[Place of publication not identified] ACM 2004
Physical Description	1 online resource (92 pages)
Notes	Bibliographic Level Mode of Issuance: Monograph English
Subjects	Engineering & Applied Sciences Computer Science
Author	Kon, Fabio Association for Computing Machinery-Digital Library. Nandkumar, Radha Schulze, Bruno
Other Title	MGC '04 Proceedings of the 2nd Workshop on Middleware for Grid Computing Proceedings of the 2nd Workshop on Middleware for Grid Computing
Identifier	ISBN : 1-58113-950-0
Language	English

Send to

REFWORKS EXPORT BIBTEX EXPORT RIS ENDNOTE E-MAIL

CITATION PRINT QR

Links

Display source record: MARC BIBFRAME 2.0

Permanent Link: https://lbdiscover.hkust.edu.hk/bib/991012755654…

```
LDR 01565cam-a2200445zu-4500
001 991012755654003412
005 20210807002030.0
006 m     o  u
007 cr |||||||||||
008 160829s2004    xx      o           eng d
035    |a(CKB)3170000000002356
035    |a(SSID)ssj0001138895
035    |a(PQKBManifestationID)11621802
035    |a(PQKBTitleCode)TC0001138895
035    |a(PQKBWorkID)11182862
035    |a(PQKB)11383441
035    |a(WaSeSS)IndRDA00013892
035    |a(EXLCZ)993170000000002356
040    |aPQKB
041    |aeng
050 14 |aQA76.76.M54
100 1  |aKon, Fabio
110    |aInternational Workshop on Middleware for Grid Computing|eCorporate Author
111 2  |aInternational Workshop on Middleware for Grid Computing|n(2nd :|d2004)
245 10 |a2nd Workshop on Middleware for Grid Computing
246 1  |aMGC '04 Proceedings of the 2nd Workshop on Middleware for Grid Computing
246 1  |aProceedings of the 2nd Workshop on Middleware for Grid Computing
264 31 |a[Place of publication not identified]|bACM|c2004
300    |a1 online resource (92 pages)
336    |atext|btxt
337    |acomputer|bc
338    |aonline resource|bcr
500    |aBibliographic Level Mode of Issuance: Monograph
546    |aEnglish
650  7 |aEngineering & Applied Sciences|2HILCC
650  7 |aComputer Science|2HILCC
710 2  |aAssociation for Computing Machinery-Digital Library.
700    |aNandkumar, Radha|eContributor
700    |aSchulze, Bruno|eContributor
776    |z1-58113-950-0
```

APPENDIX 0002C

CONFERENCE PROCEEDING

IQ-services: network-aware middleware for interactive large-data applications

Cai, Zhongtang ; Eisenhauer, Greg ; He, Qi ; Kumar, Vibhore ; Schwan, Karsten ; Wolf, Matthew

ACM International Conference Proceeding Series; Vol. 76: Proceedings of the 2nd workshop on Middleware for grid computing; 18-22 Oct. 2004, 2004, p.11-16

“ *IQ-Services are application-specific, resource-aware code modules executed by data transport middleware...* “

Available Online

TOP

SEND TO

VIEW ONLINE

DETAILS

LINKS

Send to

SHARE CITATION EMAIL PERMALINK PRINT REFWORKS EASYBIB

EXPORT RIS ENDNOTE EXPORT BIBTEX

https://wrlc-cu.primo.exlibrisgroup.com/permalink/01WRLC_CAA/1lsipa9/cdi_proquest_miscellaneous_31500804

COPY THE PERMALINK TO CLIPBOARD

View Online

Full text availability

ACM Digital Library Complete

Details

Title	IQ-services: network-aware middleware for interactive large-data applications
Creator	Cai, Zhongtang Eisenhauer, Greg He, Qi Kumar, Vibhore Schwan, Karsten Wolf, Matthew
Is Part Of	ACM International Conference Proceeding Series Vol. 76: Proceedings of the 2nd workshop on Middleware for grid computing 18-22 Oct. 2004, 2004, p.11-16
Subject	Middleware
Keywords	adaptive transport ; wide-area collaboration ; network-aware ; coordinated adaptation ; large-data transfer

Keywords	adaptive transport ; wide-area collaboration ; network-aware ; coordinated adaptation ; large-data transfer
Description	IQ-Services are application-specific, resource-aware code modules executed by data transport middleware. They constitute a 'thin' layer between application components and the underlying computational and communication resources that implements the data manipulations necessary to permit wide-area collaborations to proceed smoothly, despite dynamic resource variations. IQ-Services interact with the application and resource layers via dynamic performance attributes, and end-to-end implementations of such attributes also permit clients to interact with data providers. The joint middleware/resource and provider/consumer interactions implemented with performance attributes may be used to realize effective methods for managing the data flows in the large-data, distributed Gird applications targeted by our research. Experimental results in this paper demonstrate substantial performance improvements attained by coordinating network-level with service-level adaptations of the data being transported and by permitting end users to dynamically deploy and use applications-specific services for manipulating data in ways suitable for their current needs.
Publisher	ACM
Language	English
Source	ACM Digital Library
Identifier	ISBN: 1581139500 ISBN: 9781581139501 DOI: 10.1145/1028493.1028495

Links

Search in Amazon ›
Report a problem ›
Search in Google Books ›

Proceedings

International Conference Proceeding Series (ICPS)

Search within ICPS

ICPS Proceedings

The ACM International Conference Proceeding Series (ICPS) provides a mechanism to publish the contents of conferences, technical symposia and workshops in the ACM Digital Library, increasing their visibility among the international computing community.
ACM has published 45,000 research papers in the ACM Digital Library from over 1,350 conferences through the ICPS program since it was established in 2002. All ICPS published articles appear in ... (More)

Search Proceedings by title:

eg. SOUPS

Clear Search

Browse Proceedings alphabetically

A

91 Conferences

- **AAA-IDEA: Advanced Architectures and Algorithms for Internet Delivery and Applications**
- **AADNIC-ABMECR: Africa-Asia Dialogue Network (AADN) International Conference on Advances in Business Management and Electronic Commerce Research**
- **AAIA: Advances in Artificial Intelligence and Applications**
- **ACAI: Advances in Computing and Artificial Intelligence**
- **ACCICI: Advances in Computing, Communications and Informatics**

Proceedings

- ACET: Advances in Computer Entertainment Technology
- ACI: Animal-Computer Interaction
- ACIT: Applied Computing and Information Technology
- ACL2: ACL2 Theorem Prover and its Applications
- ACM TURC: ACM Turing Celebration Conference - China
- ACPIS: Aspects, Components, and Patterns for Infrastructure Software
- ACSAC: Annual Computer Security Applications Conference
- ACSW: Australasian Computer Science Week
- ACWR: Conference on Wireless Technologies for Humantiarian Relief
- ACySE: Agents and CyberSecurity
- AD: Array Databases
- ADAPT: Adaptive Self-tuning Computing Systems
- ADC: Australasian Database Conference
- ADCS: Australasian Document Computing Symposium
- ADMIT: Algorithms, Data Mining, and Information Technology
- ADPUC: Advanced Data Processing in Ubiquitous Computing
- AfriCHI: African Conference for Human Computer Interaction
- AH: Augmented Human
- AHS: Augmented Humans
- AI&DataMining: Integrating AI and Data Mining
- AI2A: Artificial Intelligence, Automation and Algorithms
- AIACT: Artificial Intelligence,Automation and Control Technologies
- AIAM: Artificial Intelligence and Advanced Manufacturing
- AICCC: Artificial Intelligence and Cloud Computing Conference
- AICS: Artificial Intelligence and Computer Science
- AICSconf: Artificial Intelligence and Complex Systems Conference

Proceedings

- AIDR: Artificial Intelligence for Data Discovery and Reuse
- AIEE: Artificial Intelligence in Electronics Engineering
- AIIP: AI for an Intelligent Planet
- AIIPCC: Artificial Intelligence, Information Processing and Cloud Computing
- AIMLR: Artificial Intelligence, Machine Learning and Robotics
- AIMLSystems: AI-ML-Systems
- AIMS: Autonomous and Intelligent Mobile Systems
- AINTEC: Asian Conference on Internet Engineering
- AIPR: Artificial Intelligence and Pattern Recognition
- AIR: Advances In Robotics
- AIRC: Artificial Intelligence, Robotics and Control
- AIRWeb: Adversarial Information Retrieval on the Web
- AISC: Australasian Information Security Conference
- AISS: Advanced Information Science and System
- AISTECS: Advanced Interconnect Solutions and Technologies for Emerging Computing Systems
- AIUCD: AIUCD Annual Conference
- AIVR: Artificial Intelligence and Virtual Reality
- AIW-mtg: Artificial Intelligence - Workshops
- ALICE: Alice Symposium
- AlMoDEP: Algorithms and Models for Distributed Event Processing
- AM: Audio Mostly
- AMECSE: Africa and Middle East Conference on Software Engineering
- AMME: Annual Meeting on Management Engineering
- AND: Analytics for Noisy Unstructured Text Data
- ANDARE: Autotuning and Adaptivity Approaches for Energy Efficient HPC Systems

Proceedings

- AOW: Australasian Ontology Workshop
- APBC: Asia-Pacific Bioinformatics Conference
- APCCM: Asia-Pacific Conference on Conceptual Modeling
- APCHI: Asia Pacific Conference on Computer Human Interaction
- APCHIUX: Asia Pacific Symposium of HCI and UX Design
- APCORISE: Asia Pacific Conference on Research in Industrial and Systems Engineering
- APGV: Applied Perception in Graphics and Visualization
- APIT: Asia Pacific Information Technology Conference
- APNet: Asia-Pacific Workshop on Networking
- APPIS: Applications of Intelligent Systems
- APVis: Asia-Pacific Symposium on Information Visualisation
- ArabWIC: Annual International Conference Research
- ARES: Availability, Reliability and Security
- ARM: Adaptive and Reflective MIddleware
- ARMOR: Approaches to Mobiquitous Resilience
- ARTECH: Digital Arts
- ASBD: Architectures and Systems for Big Data
- ASE BigData: Academy for Science and Engineering BigData
- Asian CHI: Asian CHI Symposium
- AsianPLoP: Asian Conference on Pattern Languages of Programs
- ASSE: Asia Service Sciences and Software Engineering
- ASWEC: Australasian Software Engineering Conference
- ATACCS: Application and Theory of Automation in Command and Control Systems
- AUIC: User Interface
- AUS-CE: Australasian Conference on Computing Education
- AUS-CSC: Australasian Conference on Computer Science

Proceedings

AusDM: Australasian Data Mining

AusGrid: Grid Computing and e-Research

- AusGrid '08: Proceedings of the sixth Australasian workshop on Grid computing and e-research - Volume 82

View Conference →

Autonomics: Autonomic Computing and Communication Systems

AVI: Advanced Visual Interfaces

AWICT: Advanced Wireless Information, Data, and Communication Technologies

B	27 Conferences
C	89 Conferences
D	45 Conferences
E	51 Conferences
F	11 Conferences
G	13 Conferences
H	16 Conferences
I	280 Conferences
J	3 Conferences
K	3 Conferences
L	12 Conferences
M	57 Conferences

MAASC: Middleware and Architectures for Autonomic and Sustainable Computing

MAB: Media Architecture Biennale Conference

Proceedings

MALETS: Machine Learning Technologies in Software Engineering

ManLang: Managed Languages and Runtimes

MANPU: Comics Aanalysis, Processing and Understanding

Mashups: Web APIs and Services Mashups

MB: Methods and Techniques in Behavioral Research

MCHPC: Memory Centric Programming for HPC

MDAC: Massive Data Analytics on the Cloud

MDI: Model-Drive Interoperability

MDS: Middleware Doctoral Symposium

MEDEA: Memory Performance: Dealing with Applications, Systems and Architecture

MedPRAI: Mediterranean Conference on Pattern Recognition and Artificial Intelligence

MEMSYS: Memory Systems

MES: Many-Core Embedded Systems

MexIHC: Mexican HCI Conference

MG: Mardi Gras Conference

MGC: Middleware for Grids, Clouds and e-Science

- MCG '06: Proceedings of the 4th international workshop on Middleware for grid computing
- MGC '05: Proceedings of the 3rd international workshop on Middleware for grid computing
- MGC '04: Proceedings of the 2nd workshop on Middleware for grid computing

View Conference ⟶

MICML: Mathematics, Intelligent Computing and Machine Learning

Middleware: Middleware

MIDI: Multimedia, Interaction, Design and Innovation

MidSens: Middleware Tools, Services and Run-Time Support for Sensor Networks

MindTrek: Mind Trek

MIRAGE: Computer Vision / Computer Graphics Collaboration Techniques and Applications

Proceedings

- MISNC: Multidisciplinary International Social Networks Conference
- mLearn: Mobile and Contextual Learning
- MLIS: Machine Learning for Interactive Systems
- MLMI: Machine Learning and Machine Intelligence
- MLNLP: Machine Learning and Natural Language Processing
- MLPA: Modules and Libraries for Proof Assistants
- MLSDA: Machine Learning for Sensory Data Analysis
- MMUI: Multimodal User Interaction
- MobileHCI: Human Computer Interaction with Mobile Devices and Services
- Mobility: Mobile Technology, Applications, and Systems
- MOBILWARE: Mobile Wireless Middleware, Operating Systems, and Applications
- MobiMedia: Mobile Multimedia Communications
- MOBIQUITOUS: Mobile and Ubiquitous Systems
- MOCO: Movement and Computing
- MOCR-AND: Multilingual OCR and Analytics for Noisy Unstructured Text Data
- MOCR: Multilingual OCR
- MODDM: Model Driven Development for Middleware
- MoDeVVa: Model-Driven Engineering, Verification and Validation
- MODULARITY: Modularity (Formerly Known as AOSD)
- MoMM: Mobile Computing and Multimedia
- MONA: Monitoring, Adaptation and Beyond
- MORSE/VAO: Model-Driven Robot Software Engineering and View-based Software-Engineering
- MORSE: Model-Driven Robot Software Engineering
- MPAC: Middleware for Pervasive and Ad-Hoc Computing
- MPLR: Managed Programming Languages and Runtimes

Proceedings

MSIE: Management Science and Industrial Engineering

MUC: Message Understanding

MUM: Mobile and Ubiquitous Multimedia

MundC: Mensch und Computer

MWSOC: Middleware for Service Oriented Computing

N	14 Conferences	
O	5 Conferences	
P	36 Conferences	
Q	4 Conferences	
R	15 Conferences	
S	71 Conferences	
T	9 Conferences	
U	7 Conferences	
V	16 Conferences	
W	38 Conferences	
X	2 Conferences	

Categories

Journals
Magazines
Books
Proceedings
ences

About

About ACM Digital Library
ACM Digital Library Board
Subscription Information
Author Guidelines
Using ACM Digital Library
All Holdings within the ACM Digital Library

Proceedings

Join

Join ACM
Join SIGs
Subscribe to Publications
Institutions and Libraries

Connect

Contact
Facebook
X
Linkedin
Feedback
Bug Report

The ACM Digital Library is published by the Association for Computing Machinery. Copyright © 2024 ACM, Inc.

Terms of Usage | Privacy Policy | Code of Ethics

http://www.acm.org/pubs/icp_series.html Go
20 captures
30 May 2002 - 9 Jul 2008
APR MAY AUG
2001 30 2002 2003
About this capture

Publications
The First Society in Computing

The ACM International Conference Proceedings Series

ACM is pleased to announce a new concept in Conference Proceedings publishing. The goal of this program is to enable
proceedings for their attendees, while also providing maximum dissemination of the material through electronic channels,

Here is an outline of how the program works:

The Wayback Machine - https://web.archive.org/web/20020530112256/http://www.acm.org:80/pubs/icp_series.html

The ACM International Conference Proceedings Series

ACM is pleased to announce a new concept in Conference Proceedings publishing. The goal of this program is to enable conferences and workshops to cost effectively produce print proceedings for their attendees, while also providing maximum dissemination of the material through electronic channels, specifically, the ACM Digital Library.

Here is an outline of how the program works:

1. Conference Organizers must fill in and submit the form below. The ACM Publications Board will review proposals from conference organizers to determine whether the conference proceedings are suitable for publication by ACM. The Publications Board will work with ACM Staff to solicit support of appropriate ACM Special Interest Groups (SIGs) should overlap of interests occur.
2. Upon acceptance, ACM Staff then works with the conference organizers to understand the requirements for the proceedings. Based on the input of the organizers (# of copies needed, estimation of pages, color, etc.) ACM will use the leverage it has with the vendors to get very reasonable prices for printing and shipping the proceedings to the location of the conference, or wherever the organizers choose.
3. The proceedings volume is identified as being part of the ACM International Proceedings Series, lending ACM's imprimatur to this publication.
4. ACM does not intend to carry any inventory or to sell hard copies; rather
5. ACM will put the full-text of the proceedings papers into the ACM Digital Library, under the heading of the "ACM International Conference Proceedings". There are over 33,000 individual subscribers to the DL, and over 700 library subscriptions. This guarantees that the proceedings will be available to the widest possible audience of computing professionals.
6. The relevant rights expressed in the ACM Rights and Responsibilities Policy (http://www.acm.org/pubs/rights.html) would apply to the proceedings, such as
    - availability at low cost, taking into account the needs of readers in economically emerging countries and in economically undeveloped countries
    - enabling fast access to the electronic version of each article, throughout the world permitting low cost purchasing of individual copies of articles
    - allowing a submission to be posted on home pages and public repositories before and after review
    - allowing metadata information, e.g., bibliographic, abstract, and keywords, for their individual work to be openly available
    - allowing authors the right to reuse their figures in their own subsequent publications for which they have granted ACM copyright,
    - respecting fair use provisions of US copyright law, and
    - allowing libraries to fill interlibrary loan requests, for teaching, research, and other not-for-profit uses.
7. Permission to include the proceedings in appropriate SIG CD publications.

To propose a volume for the ACM International Conference Proceedings Series, please fill out and submit the following information to mandelbaum@acm.org:

A. Acronym and Full Title of Meeting: _____

B. Conference Dates and Location: _____

C. Legal/Financial Sponsor(s): _____

D. Cooperating Organizations: _____

E. This event is a: Conference____; Symposium____; Workshop_____; Other _____

F. It is held: Annually____; Biennially_____; One-time_____; Other _____

G. Expected Attendance #: _____; # of Days _____; # of Sessions _____ ; # of Papers_____
Total Revenue Budget_____ Total Expense Budget_____

H. Is the sponsor a non-profit organization? Yes___; No____

I. Please provide a brief outline of the topics to be covered.
_____
_____
_____

J. URL: _____

K. General Chair:
Name: _____
Address: _____
Affiliation: _____
e-mail: _____ phone #: _____

L. Program Chair:
Name: _____
Affiliation: _____
e-mail: _____ phone #: _____

M. Paper Review Committee Names

N. Which ACM Special Interest Group(s) is/are most relevant: _____
(http://www.acm.org/sigs/guide98.html)

O. If this is a repeated event, how many submissions (S) and acceptances (A) were there for the past two events?
(Event -1) S:____; A:_____
(Event -2) S:____; A:_____

ACM Links

HOME || ABOUT ACM || MEMBERSHIP || PUBLICATIONS || SPECIAL INTEREST GROUPS (SIGs) || EDUCATION || EVENTS & CONFERENCES || AWARDS || CHAPTERS || COMPUTING & PUBLIC POLICY || PRESSROOM

ACM/Publications Last Update: April 25, 2002 by Patrick J. De Blasi
©2001 Association for Computing Machinery

https://web.archive.org/web/20060430033751/http://portal.acm.org/browse_dl.cfm?linked=1&p 150%

http://portal.acm.org/browse_dl.cfm?linked=1&part=series&coll=portal&dl=ACM&CFID=7433 Go

49 captures
8 Oct 2001 - 22 Apr 2021

OCT 2001 | APR 30 2006 | OCT 2011

About this c

Search within the ACM Proceedings:

SEARCH Advanced Search

Browse the ACM Proceedings:

- **ACM Policy:** ACM Policy
- **ACM-SE:** ACM Southeast Regional Conference
- **ACM/CSC-ER:** ACM/CSC-ER
- **AFRIGRAPH:** Computer graphics, virtual reality, visualisation and interaction in Africa
- **AGENTS:** International Conference on Autonomous Agents
- **AICPS:** ACM International Conference Proceeding Series
- **ANCS:** Symposium On Architecture For Networking And Communications Systems
- **ANNA:** Analysis of Neural Net Applications Conference
- **ANSS:** Annual Simulation Symposium

The Wayback Machine - https://web.archive.org/web/20060430033751/http://portal.acm.org:80/...



Subscribe (Full Service) Register (Limited Service, Free) Login

Search: The ACM Digital Library The Guide

SEARCH

THE ACM DIGITAL LIBRARY

Feedback Report a problem Satisfaction survey

Portal → DL Home → Proceedings

Search within the ACM Proceedings:

SEARCH Advanced Search

Browse the ACM Proceedings:

ACM Policy: ACM Policy
ACM-SE: ACM Southeast Regional Conference
ACM/CSC-ER: ACM/CSC-ER
AFRIGRAPH: Computer graphics, virtual reality, visualisation and interaction in Africa
AGENTS: International Conference on Autonomous Agents
AICPS: ACM International Conference Proceeding Series
ANCS: Symposium On Architecture For Networking And Communications Systems
ANNA: Analysis of Neural Net Applications Conference
ANSS: Annual Simulation Symposium
AOSD: Aspect-oriented software development
APL: International Conference on APL
ASE: Automated Software Engineering
ASPDAC: with EDA Technofair Design Automation Conference Asia and South Pacific
ASPLOS: Architectural Support for Programming Languages and Operating Systems
ASSETS: ACM SIGACCESS Conference on Assistive Technologies
AVI: AVI
C&C;: Creativity and Cognition
CASES: International Conference on Compilers, Architecture and Synthesis for Embedded Systems
CAW: Computer Architecture Workshop
CC: Critical Computing
CCS: Conference on Computer and Communications Security
CCSC: Consortium for Computing Sciences in Colleges
CCU: Contemporary Computing in Ukraine
CF: Conference On Computing Frontiers
CFP: Computers, Freedom and Privacy
CGO: Code Generation and Optimization
CHI: Conference on Human Factors in Computing Systems
CIKM: Conference on Information and Knowledge Management
CITC: Conference On Information Technology Education
CODES: International Conference on Hardware Software Codesign
COLT: Annual Workshop on Computational Learning Theory
COMM: Applications, Technologies, Architectures,and Protocols for Computer Communication
CPR: Special Interest Group on Computer Personnel Research Annual Conference
CQL: Symposium on Computers and the Quality of Life
CSC: ACM Annual Computer Science Conference
CSCW: Computer Supported Cooperative Work
CUU: ACM Conference on Universal Usability
CVE: Collaborative Virtual Environments
CoNEXT: International Conference On Emerging Networking Experiments And Technologies
DAC: Annual ACM IEEE Design Automation Conference

DAMON: Damon
DARE: Designing Augmented Reality Environments
DATE: Design, Automation, and Test in Europe
DEBS: Distributed event-based systems
DIAL M: Workshop on Discrete Algothrithms and Methods for MOBILE Computing and Communications
DIM: Workshop On Digital Identity Management
DIS: Symposium on Designing Interactive Systems
DL: International Conference on Digital Libraries
DMKD: Data Mining And Knowledge Discovery
DOLAP: Data Warehousing and OLAP
DPDS: International Symposium on Databases for Parallel and Distributed Systems
DPPI: Designing Pleasurable Products And Interfaces
DRM: ACM Workshop On Digital Rights Management
DSL: DSL
DUX: Designing For User Experiences
DocEng: Document Engineering
Dynamo: Workshop on Dynamic and Adaptive Compilation and Optimization
EC: Electronic Commerce
ECA: Ethics in the Computer Age
ECLIPSE: OOPSLA workshop on eclipse technology eXchange
EDTC: European Design and Test Conference
EMSOFT: International Conference On Embedded Software
ERLANG: Annual ERLANG Workshop
ETRA: Eye Tracking Research & Application
EW: ACM SIGOPS European Workshop
EuroDAC: European Design Automation Conference
FMSE: Workshop on Formal Methods in Security Engineering
FMSP: Formal Methods in Software Practice
FOIS: Formal Ontology in Information Systems
FPCA: Functional Programming Languages and Computer Architecture
FPGA: International Symposium on Field Programmable Gate Arrays
GECCO: Genetic And Evolutionary Computation Conference
GIR: Workshop On Geographic Information Retrieval
GIS: Geographic Information Systems
GLVLSI: Great Lakes Symposium on VLSI
GRAPHICON: GRAPHICON
GRAPHITE: Computer graphics and interactive techniques in Australasia and South East Asia
GROUP: Conference on Supporting Group Work
HMI: History of Medical Informatics
HOPL: History of Programming Languages
HPW: History of Personal Workstations
HSNC: History of Scientific and Numeric Computation
HT: Conference on Hypertext and Hypermedia
HWWS: SIGGRAPH/EUROGRAPHICS Workshop On Graphics Hardware
Haskell: Haskell
Hypercube: Hypercube Concurrent Computers and Applications
ICAIL: International Conference on Artificial Intelligence and Law
ICCAD: International Conference on Computer Aided Design
ICE: International Conference on Information and Computation Economies
ICER: International Computing Education Research Workshop
ICFP: International Conference on Functional Programming
ICIS: International Conference on Information Systems
ICMI: ICMI
ICS: International Conference on Supercomputing
ICSE: International Conference on Software Engineering
ICSOC: International Conference On Service Oriented Computing

IDC: Interaction Design And Children
IEA/AEI: International conference on Industrial and engineering applications of artificial intelligence and expert systems
IHIS: Interoperability Of Heterogeneous Information Systems
IMC: Internet Measurement Conference
IOPADS: Workshop on I/O in Parallel and Distributed Systems
IPSN: Information Processing In Sensor Networks
IQIS: Information Quality in Informational Systems
IRAL: International Workshop on Information Retrieval with Asia Languages
IRP2PN: Information Retrieval In Peer-To-Peer Networks
IRTAW: International Workshop on Real-time Ada Issues
ISCA: International Conference on Computer Architecture
ISLPED: International Symposium on Low Power Electronics and Design
ISMIS: International Symposium on Methodologies for Intelligent Systems
ISMM: International Symposium on Memory Management
ISPD: International Symposium on Physical Design
ISPW: International Software Process Workshop
ISSAC: International Conference on Symbolic and Algebraic Computation
ISSS: International Symposium on Systems Synthesis
ISSTA: International Symposium on Software Testing and Analysis
ISW: Internaional Symposium On Wikis
ITiCSE: Annual Joint Conference Integrating Technology into Computer Science Education
IUI: International Conference on Intelligent User Interfaces
IVME: Interpreters, Virtual Machines And Emulators
IWSSD: International Workshop on Software Specifications & Design
InfoSecCD: Information security curriculum development
JAVA: Java Grande Conference
KCAP: International Conference On Knowledge Capture
KDD: Conference on Knowledge Discovery in Data
LCTES: Language, Compiler and Tool Support for Embedded Systems
LFP: Conference on LISP and Functional Programming
MDM: International Conference On Mobile Data Management
MICRO: International Symposium on Microarchitecture
MM: International Multimedia Conference
MMDB: ACM International Workshop On Multimedia Databases
MSP: Memory System Performance
MSWiM: International Workshop on Modeling Analysis and Simulation of Wireless and Mobile Systems
MVL: Multiple-Valued Logic
Middleware: Middleware Conference
MobiCom: International Conference on Mobile Computing and Networking
MobiDE: International Workshop on Data Engineering for Wireless and Mobile Access
MobiHoc: International Symposium on Mobile Ad Hoc Networking & Computing
MobiSys: International Conference On Mobile Systems, Applications And Services
NOSSDAV: International Workshop on Network and Operating System Support for Digital Audio and Video
NPAR: Non-Photorealistic Animation and Rendering
NPIVM: New Paradigms in Information Visualization and Manipulation
NSPW: New Security Paradigms Workshop
NetGames: Network and System Support for Games
OODS: International Workshop on Object-Oriented Database Systems
OOPSLA: Conference on Object Oriented Programming Systems Languages and Applications
OOPWORK: OOPWORK
PACT: PACT
PADS: Workshop on Parallel and Distributed Simulation
PASCO: International Symposium on Parallel Symbolic Computation
PASTE: Workshop on Program Analysis for Software Tools and Engineering
PDC: Participatory Design

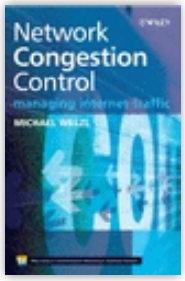
PEPM: ACM/SIGPLAN Workshop Partial Evaluation and Semantics-Based Program Manipulation
PLDI: Conference on Programming Language Design and Implementation
PODC: Annual ACM Symposium on Principles of Distributed Computing
PODS: Symposium on Principles of Database Systems
POMC: ACM Workshop On Principles Of Mobile Computing
POPL: Annual Symposium on Principles of Programming Languages
PPDP: International Conference on Principles and Practice of Declarative Programming
PPEALS: Symposium on Principles and Practice of Parallel Programming
PPoPP: Principles and Practice of Parallel Programming
PRS: Parallel Rendering Symposium
PVG: Parallel and large-data visualization and graphics
RBAC: ACM Workshop on Role Based Access Control
RECOMB: Annual Conference on Research in Computational Molecular Biology
RKRAS: Research In Knowledge Representation For Autonomous Systems
RULE: Workshop On Rule-Based Programming
SAC: Symposium on Applied Computing
SACMAT: Symposium on Access Control Models and Technologies
SASN: Workshop on Security of ad hoc and Sensor Networks
SBCCI: SBCCI
SC: Conference on High Performance Networking and Computing
SCA: Symposium on Computer Animation
SCC: Symposium on Compiler Construction
SCCG: Spring Conference on Computer Graphics
SCG: Annual Symposium on Computational Geometry
SCM: Software Configuration Management Workshop
SCN: Simulation of Computer Networks
SENSYS: Conference On Embedded Networked Sensor Systems
SESPSDE: Software Engineering Symposium on Practical Software Development Environments
SETA: Symposium on Environments and Tools for Ada
SI3D: Symposium on Interactive 3D Graphics
SIGAda: Annual International Conference on Ada
SIGCSE: Technical Symposium on Computer Science Education
SIGDOC: ACM Special Interest Group for Design of Communications
SIGFORTH: SIGFORTH
SIGGRAPH: International Conference on Computer Graphics and Interactive Techniques
SIGIR: Annual ACM Conference on Research and Development in Information Retrieval
SIGMETRICS: Joint International Conference on Measurement and Modeling of Computer Systems
SIGMOD: International Conference on Management of Data
SIGPLAN: SIGPLAN
SIGSMALL/PC: Symposium on Small Systems
SIGSOFT: Foundations of Software Engineering
SIGUCCS: User Services Conference
SLIP: International Workshop on System-Level Interconnect Prediction
SLIPE: Symposium on Language Issues in Programming Environments
SODA: Symposium on Discrete Algorithms
SOSP: ACM Symposium on Operating Systems Principles
SPAA: ACM Symposium on Parallel Algorithms and Architectures
SPDT: Symposium on Parallel and Distributed Tools
SPM: ACM Symposium on Solid and Physical Modeling
SSR: Symposium on Software Reusability
SSS: Workshop On Storage Security And Survivability
STOC: Annual ACM Symposium on Theory of Computing
SV: Software Visualization
SVR: SIGGRAPH Video Review
SWS: Workshop On Secure Web Services
SYMSAC: Symposium on Symbolic and Algebraic Manipulation

TAPIA: Richard Tapia Celebration Of Diversity In Computing
TARK: Theoretical Aspects Of Rationality And Knowledge
TAU: Timing Issues In The Specification And Synthesis Of Digital Systems
TDES: Trends and Direction in Expert Systems
TLDI: Types In Languages Design And Implementation
UIST: Symposium on User Interface Software and Technology
VAST: Virtual reality, archeology, and cultural heritage
VEE: ACM/Usenix International Conference On Virtual Execution Environments
VIS: IEEE Visualization
VLDB: Very Large Data Bases
VRML: Virtual Reality Modeling Language Symposium
VRST: Virtual Reality Software and Technology
VolVis: Symposium on Volume Visualization
WACC: International Conference on Work activities Coordination and Collaboration
WADAS: Washington Ada Symposium
WIDM: Workshop On Web Information And Data Management
WMASH: Wireless Mobile Applications And Services On WLAN Hotspots
WMC: International Workshop on Mobile Commerce
WMI: Wireless Mobile Internet
WORM: Workshop on Rapid Malcode
WOSP: Workshop on Software and Performance
WOSS: Workshop on Self-healing systems
WPDD: Workshop on Parallel & Distributed Debugging
WPES: Workshop On Privacy In The Electronic Society
WS: Wireless Security
WSC: Winter Simulation Conference
WSNA: International Workshop on Wireless Sensor Networks and Applications
WUAUC: Workshop on Universal Accessibility of Ubiquitous Computing
WWW: International World Wide Web Conference
Web3D: 3D technologies for the World Wide Web
WoWMoM: International Workshop on Wireless Mobile Multimedia
XMLSEC: Workshop On XML Security

The ACM Portal is published by the Association for Computing Machinery. Copyright © 2006 ACM, Inc.
Terms of Usage Privacy Policy Code of Ethics Contact Us

APPENDIX 0003A

Adve



Network Congestion Control: Managing Internet Traffic

Author(s): Michael Welzl

First published: 28 July 2005

Print ISBN: 9780470025284 | Online ISBN: 9780470025314 | DOI: 10.1002/047002531X

Copyright © 2005 John Wiley & Sons, Ltd

About this book

As the Internet becomes increasingly heterogeneous, the issue of congestion control becomes ever more important.

In order to maintain good network performance, mechanisms must be provided to prevent the network from being congested for any significant period of time. Michael Welzl describes the background and concepts of ... Show all

Table of Contents

Export Citation(s)

Free Access

Front Matter (Pages: i-xviii)

Summary | PDF | Request permissions

CHAPTER 1

Introduction (Pages: 1-6)

Summary | PDF | Request permissions

CHAPTER 2

Congestion control principles (Pages: 7-53)
Summary | PDF | Request permissions

CHAPTER 3

Present technology (Pages: 55-101)
Summary | PDF | Request permissions

CHAPTER 4

Experimental enhancements (Pages: 103-179)
Summary | PDF | Request permissions

CHAPTER 5

Internet traffic management – the ISP perspective (Pages: 181-198)
Summary | PDF | Request permissions

CHAPTER 6

The future of Internet congestion control (Pages: 199-217)
Summary | PDF | Request permissions

Free Access
Appendix A: Teaching congestion control with tools (Pages: 219-233)
Summary | PDF | Request permissions

Free Access
Appendix B: Related IETF work (Pages: 235-238)
Summary | PDF | Request permissions

Free Access
Appendix C: List of abbreviations (Pages: 239-242)
First Page | PDF | Request permissions

Free Access
Bibliography (Pages: 243-257)

First Page | PDF | Request permissions

Free Access

Index (Pages: 259-263)

First Page | PDF | Request permissions

ABOUT WILEY ONLINE LIBRARY

Privacy Policy
Terms of Use
About Cookies
Manage Cookies
Accessibility
Wiley Research DE&I Statement and Publishing Policies
Developing World Access

HELP & SUPPORT

Contact Us
Training and Support
DMCA & Reporting Piracy

OPPORTUNITIES

Subscription Agents
Advertisers & Corporate Partners

CONNECT WITH WILEY

The Wiley Network
Wiley Press Room

Copyright © 1999-2024 John Wiley & Sons, Inc or related companies. All rights reserved, including rights for text and data mining and training of artificial technologies or similar technologies.

Network Congestion Control

Network Congestion Control: Managing Internet Traffic Michael Welzl
© 2005 John Wiley & Sons, Ltd ISBN: 0-470-02528-X

Network Congestion Control

Managing Internet Traffic

Michael Welzl

Leopold Franzens
University of Innsbruck



John Wiley & Sons, Ltd

10.1002/0470025131X.fmatter, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/0470025131X.fmatter; Wiley Online Library on [12/05/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Copyright © 2005 John Wiley & Sons Ltd, The Atrium, Southern Gate, Chichester, West Sussex PO19 8SQ, England

Telephone (+44) 1243 779777

Email (for orders and customer service enquiries): cs-books@wiley.co.uk
Visit our Home Page on www.wiley.com

All Rights Reserved. No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means, electronic, mechanical, photocopying, recording, scanning or otherwise, except under the terms of the Copyright, Designs and Patents Act 1988 or under the terms of a licence issued by the Copyright Licensing Agency Ltd, 90 Tottenham Court Road, London W1T 4LP, UK, without the permission in writing of the Publisher. Requests to the Publisher should be addressed to the Permissions Department, John Wiley & Sons Ltd, The Atrium, Southern Gate, Chichester, West Sussex PO19 8SQ, England, or emailed to permreq@wiley.co.uk, or faxed to (+44) 1243 770620.

This publication is designed to provide accurate and authoritative information in regard to the subject matter covered. It is sold on the understanding that the Publisher is not engaged in rendering professional services. If professional advice or other expert assistance is required, the services of a competent professional should be sought.

Other Wiley Editorial Offices

John Wiley & Sons Inc., 111 River Street, Hoboken, NJ 07030, USA

Jossey-Bass, 989 Market Street, San Francisco, CA 94103-1741, USA

Wiley-VCH Verlag GmbH, Boschstr. 12, D-69469 Weinheim, Germany

John Wiley & Sons Australia Ltd, 42 McDougall Street, Milton, Queensland 4064, Australia

John Wiley & Sons (Asia) Pte Ltd, 2 Clementi Loop #02-01, Jin Xing Distripark, Singapore 129809

John Wiley & Sons Canada Ltd, 22 Worcester Road, Etobicoke, Ontario, Canada M9W 1L1

Wiley also publishes its books in a variety of electronic formats. Some content that appears in print may not be available in electronic books.

Library of Congress Cataloging-in-Publication Data

Welzl, Michael, 1973–
Network congestion control : managing Internet traffic / Michael Welzl.
p. cm.
Includes bibliographical references and index.
ISBN-13: 978-0-470-02528-4 (cloth : alk. paper)
ISBN-10: 0-470-02528-X (cloth : alk. paper)
1. Internet. 2. Telecommunication–Traffic–Management. I. Title.
TK5105.875.I57W454 2005
004.67′8 – dc22

2005015429

British Library Cataloguing in Publication Data

A catalogue record for this book is available from the British Library

ISBN-13 978-0-470-02528-4
ISBN-10 0-470-02528-X

Typeset in 10/12pt Times by Laserwords Private Limited, Chennai, India
Printed and bound in Great Britain by Antony Rowe Ltd, Chippenham, Wiltshire
This book is printed on acid-free paper responsibly manufactured from sustainable forestry in which at least two trees are planted for each one used for paper production.

10.1002/0470025301X.fmatter, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/0470025301X.fmatter, Wiley Online Library on [12/05/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

All my life, I enjoyed (and am still enjoying) a lot of support from many people – family, friends and colleagues alike, ranging from my grandmother and my girlfriend to my Ph.D. thesis supervisors. I sincerely thank them all for helping me along the way and dedicate this book to every one of them. This is not balderdash, I really mean it!

10.1002/047002531X.fmatter, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/047002531X.fmatter, Wiley Online Library on [12/05/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Contents

Foreword **xi**

Preface **xiii**

List of Tables **xv**

List of Figures **xvii**

1 Introduction **1**
1.1 Who should read this book? . . . 1
1.2 Contents . . . 2
1.3 Structure . . . 4
1.3.1 Reader's guide . . . 5

2 Congestion control principles **7**
2.1 What is congestion? . . . 7
2.1.1 Overprovisioning or control? . . . 8
2.2 Congestion collapse . . . 10
2.3 Controlling congestion: design considerations . . . 13
2.3.1 Closed-loop versus open-loop control . . . 13
2.3.2 Congestion control and flow control . . . 14
2.4 Implicit feedback . . . 14
2.5 Source behaviour with binary feedback . . . 16
2.5.1 MIMD, AIAD, AIMD and MIAD . . . 16
2.6 Stability . . . 19
2.6.1 Control theoretic modelling . . . 19
2.6.2 Heterogeneous RTTs . . . 20
2.6.3 The conservation of packets principle . . . 21
2.7 Rate-based versus window-based control . . . 21
2.8 RTT estimation . . . 23
2.9 Traffic phase effects . . . 24
2.9.1 Phase effects in daily life . . . 26
2.10 Queue management . . . 26
2.10.1 Choosing the right queue length . . . 27
2.10.2 Active queue management . . . 27
2.11 Scalability . . . 28
2.11.1 The end-to-end argument . . . 28

10.1002/047002531X.fmatter, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/047002531X.fmatter, Wiley Online Library on [12/05/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

2.11.2 Other scalability hazards 29
2.12 Explicit feedback 31
2.12.1 Explicit congestion notification 32
2.12.2 Precise feedback 32
2.13 Special environments 36
2.14 Congestion control and OSI layers 38
2.14.1 Circuits as a hindrance 39
2.15 Multicast congestion control 40
2.15.1 Problems 42
2.15.2 Sender- and receiver-based schemes 43
2.16 Incentive issues 44
2.16.1 Tragedy of the commons 44
2.16.2 Game theory 44
2.16.3 Congestion pricing 45
2.17 Fairness 47
2.17.1 Max–min fairness 48
2.17.2 Utility functions 49
2.17.3 Proportional fairness 51
2.17.4 TCP friendliness 51
2.18 Conclusion 52

3 Present technology **55**
3.1 Introducing TCP 56
3.1.1 Basic functions 57
3.1.2 Connection handling 59
3.1.3 Flow control: the sliding window 60
3.1.4 Reliability: timeouts and retransmission 61
3.2 TCP window management 62
3.2.1 Silly window syndrome 62
3.2.2 SWS avoidance 62
3.2.3 Delayed ACKs 64
3.2.4 The Nagle algorithm 64
3.3 TCP RTO calculation 65
3.3.1 Ignoring ACKs from retransmissions 66
3.3.2 Not ignoring ACKs from retransmissions 66
3.3.3 Updating RTO calculation 67
3.4 TCP congestion control and reliability 69
3.4.1 Slow start and congestion avoidance 69
3.4.2 Combining the algorithms 71
3.4.3 Design rationales and deployment considerations 73
3.4.4 Interactions with other window-management algorithms 74
3.4.5 Fast retransmit and fast recovery 75
3.4.6 Multiple losses from a single window 77
3.4.7 NewReno 79
3.4.8 Selective Acknowledgements (SACK) 81
3.4.9 Explicit Congestion Notification (ECN) 84

10.1002/047002531X.fmatter, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/047002531X.fmatter, Wiley Online Library on [12/05/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

3.5 Concluding remarks about TCP . . . 88
3.6 The Stream Control Transmission Protocol (SCTP) . . . 91
3.7 Random Early Detection (RED) . . . 93
3.8 The ATM 'Available Bit Rate' service . . . 96
3.8.1 Explicit rate calculation . . . 98
3.8.2 TCP over ATM . . . 100

4 Experimental enhancements 103
4.1 Ensuring appropriate TCP behaviour . . . 104
4.1.1 Appropriate byte counting . . . 104
4.1.2 Limited slow start . . . 106
4.1.3 Congestion window validation . . . 107
4.1.4 Robust ECN signalling . . . 108
4.1.5 Spurious timeouts . . . 109
4.1.6 Reordering . . . 113
4.1.7 Corruption . . . 115
4.2 Maintaining congestion state . . . 119
4.2.1 TCP Control Block Interdependence . . . 119
4.2.2 The Congestion Manager . . . 119
4.2.3 MulTCP . . . 121
4.3 Transparent TCP improvements . . . 123
4.3.1 Performance Enhancing Proxies (PEPs) . . . 123
4.3.2 Pacing . . . 126
4.3.3 Tuning parameters on the fly . . . 128
4.4 Enhancing active queue management . . . 129
4.4.1 Adaptive RED . . . 130
4.4.2 Dynamic-RED (DRED) . . . 131
4.4.3 Stabilized RED (SRED) . . . 132
4.4.4 BLUE . . . 133
4.4.5 Adaptive Virtual Queue (AVQ) . . . 133
4.4.6 RED with Preferential Dropping (RED-PD) . . . 134
4.4.7 Flow Random Early Drop (FRED) . . . 135
4.4.8 CHOKe . . . 135
4.4.9 Random Early Marking (REM) . . . 136
4.4.10 Concluding remarks about AQM . . . 137
4.5 Congestion control for multimedia applications . . . 139
4.5.1 TCP-friendly congestion control mechanisms . . . 143
4.5.2 The Datagram Congestion Control Protocol (DCCP) . . . 149
4.5.3 Multicast congestion control . . . 155
4.6 Better-than-TCP congestion control . . . 160
4.6.1 Changing the response function . . . 161
4.6.2 Delay as a congestion measure . . . 164
4.6.3 Packet pair . . . 167
4.6.4 Explicit feedback . . . 169
4.6.5 Concluding remarks about better-than-TCP protocols . . . 175
4.7 Congestion control in special environments . . . 176

10.1002/047002531X.fmatter, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/047002531X.fmatter, Wiley Online Library on [12/05/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

5 Internet traffic management – the ISP perspective 181
5.1 The nature of Internet traffic . . . 182
5.2 Traffic engineering . . . 184
5.2.1 A simple example . . . 185
5.2.2 Multi-Protocol Label Switching (MPLS) . . . 186
5.3 Quality of Service (QoS) . . . 187
5.3.1 QoS building blocks . . . 188
5.3.2 IntServ . . . 190
5.3.3 RSVP . . . 191
5.3.4 DiffServ . . . 191
5.3.5 IntServ over DiffServ . . . 192
5.4 Putting it all together . . . 194

6 The future of Internet congestion control 199
6.1 Small deltas or big ideas? . . . 200
6.1.1 TCP-friendliness considerations . . . 201
6.1.2 A more aggressive framework . . . 203
6.2 Incentive issues . . . 205
6.2.1 The congestion response of UDP-based applications . . . 206
6.2.2 Will VoIP cause congestion collapse? . . . 209
6.2.3 DCCP deployment considerations . . . 211
6.2.4 Congestion control and QoS . . . 212
6.3 Tailor-made congestion control . . . 214
6.3.1 The Adaptation Layer . . . 214
6.3.2 Implications . . . 215

A Teaching congestion control with tools 219
A.1 CAVT . . . 220
A.1.1 Writing scripts . . . 223
A.1.2 Teaching with CAVT . . . 224
A.1.3 Internals . . . 225
A.2 *ns* . . . 227
A.2.1 Using *ns* for teaching: the problem . . . 227
A.2.2 Using *ns* for teaching: the solution . . . 228
A.2.3 NSBM . . . 229
A.2.4 Example exercises . . . 233

B Related IETF work 235
B.1 Overview . . . 235
B.2 Working groups . . . 236
B.3 Finding relevant documents . . . 238

C List of abbreviations 239

Bibliography 243

Index 259

10.1002/0470025131X.fmatter, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/0470025131X.fmatter, Wiley Online Library on [12/05/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Foreword

The Internet is surely the second most extensive machine on the planet, after the public switched telephone network (PSTN), and it is rapidly becoming as ubiquitous. In fact, the distinction between the two is fast diminishing as the vision of the unified telecommunication network begins to be realized, and telecommunication operators deploy voice over IP (VoIP) technology. One of the biggest issues involved in the transition from PSTN to VoIP is ensuring that the customer sees (hears!) the best possible Quality of Service at all times. This is a considerable challenge for the network's designers and engineers.

Meanwhile, national governments – and also the European Commission – are implicitly assuming the emergence of the 'Information Society', and even funding research in pursuit of it. Critical applications including health, education, business and government are going to be increasingly dependent on information networks, which will inevitably be based on Internet (and Web) technologies. The penetration of broadband access into homes as well as businesses is rapidly bringing Web (and Internet) into everyone's lives and work.

The Internet was never foreseen as the more commercial network that it has now become: an informal tool for researchers has become a cornerstone of business. It is crucial that the underlying technology of the Internet is understood by those who plan to employ it to support critical applications. These 'enterprise owners', whether they be governments or companies, need to understand the principles of operation of the Internet, and along with these principles, its shortcomings and even its vulnerabilities. It does have potential shortcomings, principally its unproven ability to act as a critical support infrastructure; and it does have vulnerabilities, including its inability to cope with distributed denial-of-service attacks. These are arguably among the most pressing topics for Internet research.

It is particularly important that there are no unwarranted assumptions about the ability of the Internet to support more commercial activities and various critical applications. People involved in managing and operating Internet-based networks, and those who are considering its potential, will be suitably educated by Michael Welzl's book.

Congestion – the overloading of switches or routers with arriving traffic packets – is a consequence of the design of the Internet. Many mechanisms have been proposed to deal with it, though few have been deployed as yet. This book covers the theory and practical considerations of congestion, and gives an in-depth treatment of the subject.

'Network Congestion Control: Managing Internet Traffic' is a welcome addition to the Wiley Series in Communications Networking & Distributed Systems.

David Hutchison
Lancaster University
April 2005

10.1002/0470025131.fmatter, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/0470025131.fmatter; Wiley Online Library on [12/05/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Preface

Some people raised their eyebrows when I told them that I was writing a book on congestion control, and said, 'Is this topic really large enough for a book?' Well, it certainly is, and I am sure that it is needed. For example, there are quite a number of Ph.D. students out there who work in this field–yet, when they start out, they do not have a comprehensive introductory book about the subject. This book is for them and, of course, for anyone else with an interest in this fascinating topic–graduate students, teachers in academia and network administrators alike.

While the original page estimate was only a little lower than the actual outcome, I am now convinced that it would have been possible to write a book of twice this size on congestion control–but this would have meant diverging from the original goals and including things that are already nicely covered in other places. Instead of overloading this book, I therefore choose to recommend two books that were published last year as complementary material: (Hassan and Jain 2004) and (Srikant 2004).

Even if there is only one author, no book is the work of a single person. In my case, there are many people who provided help in one way or another – Anil Agarwal, Simon Bailey, Sven Hessler and Murtaza Yousaf proofread the text; this was sometimes a hectic task, especially towards the end of the process, but they all just kept on providing me with valuable input and constructive criticism. Neither you nor I would be happy with the result if it was not for these people. I would like to point out that I never personally met Anil–we got in touch via a technical discussion in the end-to-end interest mailing list of the IRTF, and he just volunteered to proofread my book. This certainly ranks high in the list of 'nicest things that ever happened to me', and deserves a big thanks.

I would like to thank Craig Partridge for providing me with information regarding the history of congestion control and allowing me to use his description of the 'global congestion collapse' incident. Further thanks go to Martin Zwicknagl for his Zillertaler Bauernkrapfen example, Stefan Hainzer for bringing an interesting article about fairness to my attention, and Stefan Podlipnig for numerous discussions which helped to shape the book into its present form. Two tools are described in Appendix A, where it is stated that they 'were developed at our University'. Actually, they were implemented by the following students under my supervision: Christian Sternagel wrote CAVT, and Wolfgang Gassler, Robert Binna and Thomas Gatterer wrote NSBM. The congestion control behaviour analyses of various applications described in Chapter 6 were carried out by Muhlis Akdag, Thomas Rammer, Roland Wallnöfer, Andreas Radinger and Marcus Fischer under my supervision. I would like to mention Michael Trawöger because he insisted that he be named *here*; he had a 'cool cover idea' that may or may not have made it onto the final book's front page. While the right people to thank for this are normally the members of the Wiley graphics

10.1002/047002531X.fmatter, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/047002531X.fmatter; Wiley Online Library on [12/05/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

department, if the cover of this book is the best that you have ever seen, it surely was his influence.

I would like to thank two people whom I have never been in touch with but the work that they have done had such a major influence on this book: Raj Jain and Grenville Armitage. Like nobody else, Raj Jain is able to explain seemingly complex things in a simple manner; reading his papers assured me that it would indeed be possible to write a book that is easily comprehensible yet covers the whole range of congestion control issues. I had the writing style of his early papers in mind when I wrote certain sections of this book, especially Chapter 2. Grenville Armitage had a somewhat similar influence, as he impressed me with his book 'Quality of Service in IP Networks' (Armitage 2000). It is easily readable and introductory, but it still manages to cover all the important topics. I used this book as some sort of a 'role model'−Grenville Armitage basically did with QoS what I tried to do with congestion control.

I should not forget to thank the people who helped me at the publisher's side of the table−David Hutchison, Birgit Gruber, Joanna Tootill and Julie Ward. These are of course not the only people at John Wiley & Sons who were involved in the production of this book−while I do not know the names of the others, I thank them all! Figures 3.13, 4.11, 6.1 and A.5 were taken from (Welzl 2003) with kind permission of Springer Science and Business Media (Kluwer Academic Publishers at the time the permission was granted).

Finally, I would like to name the people whom I already mentioned in the 'dedication': my Ph.D. thesis supervisors, who really did a lot for me, were Max Mühlhäuser and Jon Crowcroft. My grandmother, Gertrud Welzl, provided me with an immense amount of support throughout my life, and I also enjoy a lot of support from my girlfriend, Petra Ratzenböck; I really strained her patience during the final stages of book writing. As I write this, it strikes me as odd to thank my girlfriend, while most other authors thank their wives−perhaps the time has come to change this situation.

10.1002/047002531X.fmatter, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/047002531X.fmatter; Wiley Online Library on [12/05/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

List of tables

2.1 Possible combinations for using explicit feedback 33

4.1 Active queue management schemes 138
4.2 Differences between CADPC/PTP and XCP 175
4.3 Applicability of mechanisms in this section for special environments 178

A.1 Common abstraction levels for network analysis and design 222

10.1002/047002531X.fmatter, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/047002531X.fmatter; Wiley Online Library on [12/05/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

List of figures

2.1 Congestion collapse scenario 11
2.2 Throughput before (a) and after (b) upgrading the access links 12
2.3 Data flow in node 2 12
2.4 Vector diagrams showing trajectories of AIAD, MIMD, MIAD (a) and AIMD (b) 17
2.5 Rate evolvement with MIAD (a) and AIMD (b) 18
2.6 Simple feedback control loop 20
2.7 AIMD trajectory with $RTT_{\text{customer 0}} = 2 \times RTT_{\text{customer 1}}$ 20
2.8 (a) The full window of six packets is sent. (b) The receiver ACKs 22
2.9 Three CBR flows – separate (a) and interacting (b) 25
2.10 Topology used to simulate burstiness with CBR flows 25
2.11 Choke packets 33
2.12 Explicit rate feedback 35
2.13 Hop-by-hop congestion control 35
2.14 Unicast, broadcast, overlay multicast and multicast 41
2.15 Scenario for illustrating fairness (a); zooming in on resource A (b) 48
2.16 Utility functions of several types of applications 50

3.1 The TCP header 56
3.2 Connection setup (a) and teardown (b) procedure in TCP 59
3.3 The buffer of a TCP sender 60
3.4 Silly window syndrome avoidance 63
3.5 Slow start (a) and congestion avoidance (b) 71
3.6 Evolution of *cwnd* with TCP Tahoe and TCP Reno 72
3.7 A sequence of events leading to Fast Retransmit/Fast Recovery 78
3.8 The TCP SACK option format 81
3.9 How TCP uses ECN 87
3.10 Standards track TCP specifications that influence when a packet is sent . . 88
3.11 The marking function of RED 94
3.12 The marking function of RED in ‘gentle’ mode 96
3.13 Proportional rate adaptation as in CAPC. 100

4.1 A spurious timeout 110
4.2 The congestion manager 120
4.3 Connection splitting 124
4.4 Pacing 127

10.1002/0470025131X.fmatter, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/0470025131X.fmatter, Wiley Online Library on [12/05/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

4.5 Matching a fluctuating application stream onto a fluctuating congestion control mechanism . . . 141
4.6 Matching a constant application stream onto a constant congestion control mechanism . . . 142
4.7 Matching an adaptive application stream onto a fluctuating congestion control mechanism . . . 143
4.8 The DCCP generic header (both variants) . . . 153
4.9 TCP congestion avoidance with different link capacities . . . 160
4.10 Packet pair . . . 168
4.11 CADPC/PTP in action. . . . 173

5.1 A traffic engineering problem . . . 185
5.2 A generic QoS router . . . 188
5.3 Leaky bucket and token bucket . . . 189
5.4 IntServ over DiffServ . . . 193

6.1 A vector diagram of TCP Reno. . . . 203
6.2 The test bed that was used for our measurements . . . 206
6.3 Sender rate and throughput of streaming media tools . . . 208
6.4 Average packet length of interactive multiplayer games . . . 209
6.5 The Adaptation Layer . . . 215
6.6 Four different ways to realize congestion control . . . 216

A.1 Screenshot of CAVT showing an AIMD trajectory . . . 221
A.2 Screenshot of the CAVT time line window (user 1 = AIAD, user 2 = MIMD, equal RTTs) . . . 222
A.3 Some CADPC trajectories . . . 224
A.4 Class structure of CAVT . . . 226
A.5 A typical *ns* usage scenario. . . . 228
A.6 A screenshot of NSBM . . . 230
A.7 A screenshot of *nam* . . . 231
A.8 A screenshot of *xgraph* . . . 232
A.9 The congestion collapse scenario with 1 Mbps from source 1 to router 2 . . . 233

10.1002/047002531X.fmatter, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/047002531X.fmatter, Wiley Online Library on [12/05/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

APPENDIX 0003B

LIBRARY | CATALOG

1 of 1

BOOK

Network congestion control : managing Internet traffic

Full Record MARC Tags

Personal name

Welzl, Michael, 1973-

Main title

Network congestion control : managing Internet traffic / Michael Welzl.

Published/Created

Chichester, West Sussex, England ; Hoboken, NJ : J. Wiley, c2005.

Request this Item | LC Find It

Links

Links

Publisher description http://www.loc.gov/catdir/enhancements/fy0622/2005015429-d.html
Table of contents only http://www.loc.gov/catdir/enhancements/fy0622/2005015429-t.html
Contributor biographical information http://www.loc.gov/catdir/enhancements/fy0653/2005015429-b.html

More Information

LCCN Permalink	https://lccn.loc.gov/2005015429
Description	xviii, 263 p. : ill. ; 25 cm.
ISBN	9780470025284 (cloth : alk. paper) 047002528X (cloth : alk. paper)
LC classification	TK5105.875.I57 W454 2005
LC Subjects	Internet. Telecommunication--Traffic--Management.
Browse by shelf order	TK5105.875.I57
Notes	Includes bibliographical references (p. [243]-257) and index.
LCCN	2005015429
Dewey class no.	004.67/8
National bib agency no.	013205702
Other system no.	(OCoLC)ocm60512784
Type of material	Book

Item Availability

CALL NUMBER	TK5105.875.I57 W454 2005 Copy 1
Request in	Jefferson or Adams Building Reading Rooms
Status	Not Charged

LIBRARY | CATALOG

1 of 1

BOOK

Network congestion control : managing Internet traffic

Full Record **MARC Tags**

000		01495cam a22003254a 4500
001		13985151
005		20070110144108.0
008		050527s2005 enka b 001 0 eng
906	__	|**a** 7 |**b** cbc |**c** orignew |**d** 1 |**e** ecip |**f** 20 |**g** y-gencatlg
925	0_	|**a** acquire |**b** 1 shelf copy |**x** policy default
955	__	|**a** jf17 2005-05-27; |**i** jf17 2005-05-27; |**e** jf17 2005-05-27 to Dewey |**a** aa07 2005-06-02 |**a** ps12 2005-10-25 1 copy rec'd., to CIP ver. |**f** pv15 2005-10-27 z-processor to BCCD
010	__	|**a** 2005015429
016	7_	|**a** 013205702 |**2** Uk
020	__	|**a** 9780470025284 (cloth : alk. paper)
020	__	|**a** 047002528X (cloth : alk. paper)
035	__	|**a** (OCoLC)ocm60512784
040	__	|**a** DLC |**c** DLC |**d** UKM |**d** BWKUK |**d** BAKER |**d** DLC
042	__	|**a** pcc
050	00	|**a** TK5105.875.I57 |**b** W454 2005
082	00	|**a** 004.67/8 |**2** 22
100	1_	|**a** Welzl, Michael, |**d** 1973-
245	10	|**a** Network congestion control : |**b** managing Internet traffic / |**c** Michael Welzl.
260	__	|**a** Chichester, West Sussex, England ; |**a** Hoboken, NJ : |**b** J. Wiley, |**c** c2005.
300	__	|**a** xviii, 263 p. : |**b** ill. ; |**c** 25 cm.
504	__	|**a** Includes bibliographical references (p. [243]-257) and index.

650	_0	|**a** Internet.
650	_0	|**a** Telecommunication |**x** Traffic |**x** Management.
856	42	|**3** Publisher description |**u** http://www.loc.gov/catdir/enhancements/fy0622/2005015429-d.html
856	41	|**3** Table of contents only |**u** http://www.loc.gov/catdir/enhancements/fy0622/2005015429-t.html
856	42	|**3** Contributor biographical information |**u** http://www.loc.gov/catdir/enhancements/fy0653/2005015429-b.html

Request this Item | LC Find It

Item Availability

CALL NUMBER	TK5105.875.I57 W454 2005 Copy 1
Request in	Jefferson or Adams Building Reading Rooms
Status	Not Charged

APPENDIX 0003C



Network congestion control : managing Internet traffic /

Show other versions (1)

Main Author:	Welzl, Michael, 1973-
Format:	Book
Language:	English
Published:	Chichester, West Sussex, England ; Hoboken, NJ : J. Wiley, [2005]
Subjects:	Internet. Telecommunication > Traffic > Management.

Holdings | Description | Other Versions (1) | Similar Items | Staff View

MARC

LEADER	00000pam a22000004a 4500	
001	in00002093568	
005	20151102145132.0	
008	050527s2005 enka b 001 0 eng	
010		**|a** 2005015429
015		**|a** GBA544029 **|2** bnb
016	7	**|a** 013205702 **|2** Uk
019		**|a** 60512784
020		**|a** 047002528X (cloth : alk. paper)
024	3	**|a** 9780470025284

035			**|a** (OCoLC)ocm62179107
040			**|a** DLC **|c** DLC **|d** UKM **|d** BWKUK **|d** BAKER **|d** TXA **|d** UtOrBLW
042			**|a** pcc
049			**|a** TXAM
050	0	0	**|a** TK5105.875.I57 **|b** W454 2005
082	0	0	**|a** 004.67/8 **|2** 22
100	1		**|a** Welzl, Michael, **|d** 1973-
245	1	0	**|a** Network congestion control : **|b** managing Internet traffic / **|c** Michael Welzl.
264		1	**|a** Chichester, West Sussex, England ; **|a** Hoboken, NJ : **|b** J. Wiley, **|c** [2005]
264		4	**|c** ©2005
300			**|a** xviii, 263 pages : **|b** illustrations ; **|c** 25 cm.
336			**|a** text **|b** txt **|2** rdacontent
337			**|a** unmediated **|b** n **|2** rdamedia
338			**|a** volume **|b** nc **|2** rdacarrier

504			**|a** Includes bibliographical references (pages [243]-257) and index.
650		0	**|a** Internet.
650		0	**|a** Telecommunication **|x** Traffic **|** **x** Management.
948			**|a** cataloged **|b** h **|** **c** 2005/11/4 **|d** c **|** **e** rdavis **|f** 12:26:18 pm
994			**|a** C0 **|b** TXA
999			**|a** MARS
999	f	f	**|s** 5bcbf860-ed7d-3663-ba85-10010b83111c **|** **i** 0b499c54-4983-3082-8eae-573a97cfeeda **|** **t** 0
952	f	f	**|p** normal **|a** Texas A&M University **|** **b** College Station **|** **c** Sterling C. Evans Library **|d** Evans: Library Stacks **|t** 0 **|** **e** TK5105.875.I57 W454 2005 **|h** Library of Congress classification **|** **i** unmediated -- volume **|** **m** A14833278582
998	f	f	**|a** TK5105.875.I57 W454 2005 **|t** 0 **|** **l** Evans: Library Stacks

APPENDIX 0003D

https://web.archive.org/web/20051115025024/http://www.wiley.com/WileyCDA/WileyTitle/p 150%

http://www.wiley.com/WileyCDA/WileyTitle/productCd-047002528X.html Go

APR NOV DEC
15
2004 2005 2006

26 captures
27 Apr 2005 - 13 Dec 2017

About this capture

Related Subjects

General Networking

Optical Communications

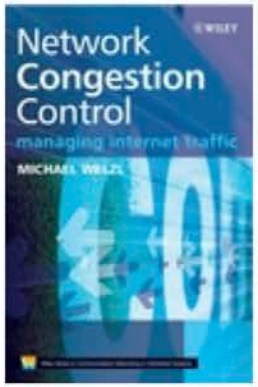
Satellite Communications

Signal Processing

Antennas & Propagation

Audio & Speech Processing and Broadcasting

Networks



Network Congestion Control: Managing Internet Traffic

Michael Welzl
ISBN: 0-470-02528-X
Hardcover
282 pages
September 2005

US $120.00 Add to Cart

This price is valid for United States. Change location to view local pricing and availability.

Description **As the Internet becomes increasingly heterogeneous, the issue of congestion**

The Wayback Machine - https://web.archive.org/web/20051115025024/http://www.wiley.com:80/WileyCDA/WileyTitle/prod…

Related Subjects

General Networking

Optical Communications

Satellite Communications

Signal Processing

Antennas & Propagation

Audio & Speech Processing and Broadcasting

Communication System Security

Join an Engineering Mailing List

Related Titles

Networks

Parlay / OSA: From Standards to Reality (Hardcover)
by Musa Unmehopa, Kumar Vemuri, Andy Bennett

Ad Hoc Wireless Networks: A Communication- Theoretic Perspective (Hardcover)
by Ozan Tonguz, Gianluigi Ferrari

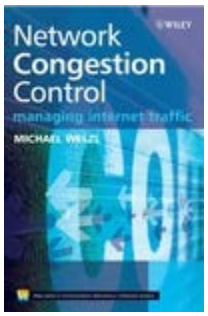
Protocols and Architectures for Wireless Sensor Networks (Hardcover)
by Holger Karl, Andreas Willig

Quality of Service in a Cisco Networking Environment (Hardcover)
by Gilbert Held

Ethernet Networks: Design, Implementation, Operation, Management , 4th Edition (Paperback)
by Gilbert Held

Security for Ubiquitous Computing (Hardcover)
by Frank Stajano

Networks



Network Congestion Control: Managing Internet Traffic

Michael Welzl
ISBN: 0-470-02528-X
Hardcover
282 pages
September 2005

US $120.00 Add to Cart

This price is valid for United States. Change location to view local pricing and availability.

Description

Table of Contents

Author Information

As the Internet becomes increasingly heterogeneous, the issue of congestion control becomes ever more important.

In order to maintain good network performance, mechanisms must be provided to prevent the network from being congested for any significant period of time. Michael Welzl describes the background and concepts of Internet congestion control, in an accessible and easily comprehensible format. Throughout the book, not just the how, but the why of complex technologies including the Transmission Control Protocol (TCP) and Active Queue Management are explained. The text also gives an overview of the state-of-the-art in congestion control research and an insight into the future.

Network Congestion Control:

- Presents comprehensive, easy-to-read documentation on the advanced topic of congestion control without heavy maths.
- Aims to give a thorough understanding of the evolution of Internet congestion control: how TCP works, why it works the way it does, and why some congestion control concepts failed for the Internet.
- Explains the Chiu/Jain vector diagrams and introduces a new method of using these diagrams for analysis, teaching & design.
- Elaborates on how the theory of congestion control impacts on the practicalities of service delivery.
- Includes an appendix with examples/problems to assist learning.
- Provides an accompanying website with Java tools for teaching congestion control, as well as examples, links to code and projects/bibliography.

This invaluable text will provide academics and researchers in computer science, electrical engineering and communications networking, as well as students on advanced networking and Internet courses, with a thorough understanding of the current state and future evolution of Internet congestion control. Network

administrators and Internet service and applications providers will also find *Network Congestion Control* a comprehensive, accessible self-teach tool.

Third-Generation Systems and Intelligent Wireless Networking: Smart Antennas and Adaptive Modulation (Hardcover)
by J. S. Blogh, Lajos Hanzo

Printer-ready version E-mail a friend

Copyright © 2000-2005 by John Wiley & Sons, Inc. or related companies. All rights reserved. Please read our Privacy Policy

APPENDIX 0003E

https://web.archive.org/web/20051224175649/http://www.welzl.at/congestion/index.html 150%

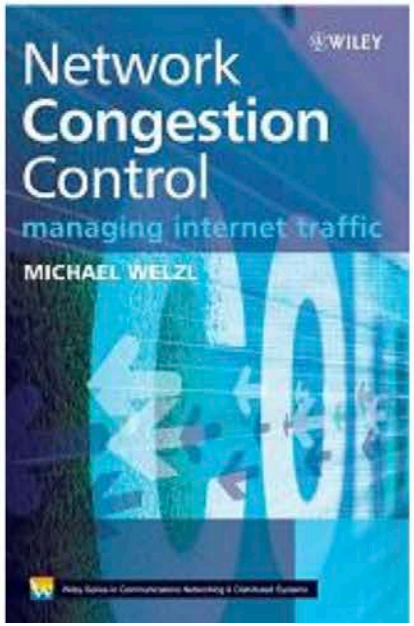
http://www.welzl.at/congestion/index.html Go NOV DEC JAN

2 captures
24 Dec 2005 - 6 Jan 2006

2004 24 2005 2006 About this capture

Network Congestion Control: Managing Internet Traffic

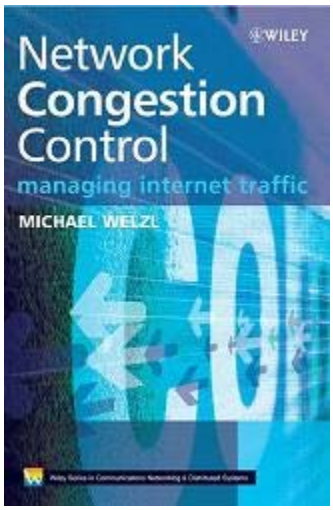
Accompanying website



The Wayback Machine - https://web.archive.org/web/20051224175649/http://www.welzl.at:80/congestion/index.html

Network Congestion Control: Managing Internet Traffic

Accompanying website



Please visit the Wiley or Amazon pages for details about the book itself.

Note: this page contains a lot of links to people because these links tend to be more stable than direct links to pages related to protocols or mechanisms. However, a link to person X doesn't mean that I consider the work done by this person more important than the work done by person Y (who is not mentioned here) - probably, I just liked the web page of person X better :-)

Please drop me a line if you notice a malfunctioning link!

Preface

- In April 2005, some progress was made regarding the issue that is mentioned in the last sentence.

Chapter 2: Congestion control principles

- This chapter starts out with a reference to Keshav's thesis (scroll down a bit), which is an important congestion control reference.
- This simulation script was used to generate figure 2.2.
- Many of the basic principles in this chapter are nicely described in papers by Raj Jain, whose list of publications is a really good starting point for background reading. This is also the place to look for the paper that introduced the AIMD diagrams. Congestion collapse
- Speaking of important papers, this is the origin of the end-to-end argument; if you're a Ph.D. student working on computer networks, you simply must read this paper - it will help you understand why some things are the way they are in today's Internet.
- Phase effects are described here.
- Here are slides from Van Jacobson's SIGCOMM 2001 speech (mentioned in section 2.14.2).
- Some material on the tragedy of the commons; take it with a grain of salt...
- This tutorial by Jean-Yves Le Boudec discusses fairness issues; proportional fairness is a particularly interesting concept that was introduced by Frank Kelly.

Chapter 3: Present technology

- This chapter is mainly based on RFCs, which can be obtained from the RFC-editor (and a million other places). For a quick reference to TCP RFCs, I recommend figure 3.10 on page 88 and the TCP roadmap document, which is still an Internet-draft at the time of writing (albeit an extremely useful one). Of course, you also shouldn't miss the famous congestion avoidance and control paper.
- Sally Floyd made significant contributions to the congestion control mechanisms in today's Internet; she maintains special web pages, e.g., about ECN, SACK, RED... her website, and the ICIR pages in general, are good sources for any such information; for instance, there's the TBIT page which documents the current state of TCP implementations in the Internet.
- This is the squirrel that is mentioned in footnote 13, section 3.4.2 (page 72).
- Some SCTP material is here.
- ATM ABR is thoroughly described in the freely available traffic management specification of the ATM Forum.

Chapter 4: Experimental enhancements

- Many of the enhancements that are described at the beginning of this chapter are specified in experimental RFCs. For these and other TCP enhancements, I would once again like to recommend Sally Floyd's page - the number of literature pointers that she maintains is truly impressive!
- The Congestion Manager (section 4.2.2, page 119).
- These slides by Van Jacobson are mentioned in section 4.4 (page 129).
- Steven Low's group is the place to look for papers on REM, FAST and a lot of maths.
- The DCCP page at ICIR.
- Injong Rhee is the author of TEAR, BIC and CUBIC; he maintains pages for each of these protocols. Similarly, Saverio Mascolo's homepage is probably the best source for information about TCP Westwood. Highspeed TCP and Quick-start are, again, Sally Floyd projects, and CADPC/PTP is my own little contribution to the congestion control community. I maintain a dedicated webpage. Dina Katabi is the main author of XCP.

Chapter 5: Internet traffic management – the ISP perspective

- Section 5.1, "the nature of Internet traffic" - here, again, I recommend Sally Floyd's related page.
- For more information about QoS, I recommend the book "Quality of Service in IP Networks" by Grenville Armitage.

Chapter 6: The future of Internet congestion control

- Some of the things in this chapter are my own research efforts - e.g., "Tailor-made Congestion Control" is a FWF (Austrian Science Fund) supported research project at the University of Innsbruck.
- If you're interested in the most recent congestion control research efforts, it might be a good idea to look at the online technical programs of the ACM SIGCOMM and IEEE INFOCOM conferences, both of which contain freely available papers. ACM CCR is also a good source that is freely available to the public.

Appendix A: Teaching congestion control with tools

- CAVT has a webpage of its own.
- I maintain a small ns webpage, which contains the NSBM tool and the throughput.pl perl script.
- This is not mentioned in the book, and not related to congestion control: if you like these tools and plan to use them in a general networking course, you might also like the irvtool which is nice for teaching routing basics.

Appendix B: Related IETF work

- The Internet Engineering Task Force (IETF) and the Internet Research Task Force (IRTF) with its end-to-end research group. Note that there has been talk about an IRTF congestion control research group, so keep track of the IRTF homepage and see what's happening!
- Some relevant IETF working groups:
  - TCP Maintenance and Minor Extensions (tcpm)
  - Transport Area Working Group (tsvwg)
  - Datagram Congestion Control Protocol (dccp)
  - Audio/Video Transport (avt)
  - Reliable Multicast Transport (rmt)
  - Signaling Transport (sigtran)
  - Robust Header Compression (rohc)

Last update: 25. July 2005
Back to Michael's page

APPENDIX 0003F

Appendix 0003F
Citations to Welzl

Source Document

Welzl, M. (2005). *Network congestion control: managing internet traffic*. John Wiley & Sons.

Citations: 2006 – 2007 (in chronological order)

Rusmin, P. H., Machbub, C., & Harsoyo, A. (2006, December). White: Controlling The Internet Congestion With Integral Controller. In *2006 9th International Conference on Control, Automation, Robotics and Vision* (pp. 1-6). IEEE.

Mohamed, E. D., Petriu, D., & Rolia, J. (2007, May). Load-balancing data traffic among inter-domain links. *IEEE Journal on Selected Areas in Communications*, *25*(5), 1022-1033.

Makki, S., & Wunnava, S. V. (2007, October). Dealing with Challenges in Managing Heterogeneous Networks by Using Mobile Agent Technology. *Int. J. of Computer Science and Network Security*, 316-321.

Alcaraz, S., Juiz, C., Gilly, K., & Puigjaner, R. (2007, November). Promoting web traffic over a DiffServ architecture. In *2007 22nd international symposium on computer and information sciences* (pp. 1-6). IEEE.

Yang, L., Liu, C., Schopf, J. M., & Foster, I. (2007, November). Anomaly detection and diagnosis in grid environments. In *Proceedings of the 2007 ACM/IEEE conference on Supercomputing* (pp. 1-9).

Jang, S. H., Kim, I. K., & Lee, J. S. (2007, December). Node Availability-Based Congestion Control Model Using Fuzzy Logic for Computational Grid. In *Future Generation Communication and Networking (FGCN 2007)* (Vol. 1, pp. 307-312). IEEE.