UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE PATENT TRIAL AND APPEAL BOARD

CISCO SYSTEMS, INC.,
Petitioner,

v.

VIDEO SOLUTIONS PTE. LTD.
Patent Owner.

IPR2024-00922
U.S. Patent No. 8,446,823

DECLARATION OF DR. IMMANUEL FREEDMAN, UNDER 37 C.F.R. § 1.68 IN SUPPORT OF PETITION FOR *INTER PARTES* REVIEW

Petitioner's Exhibit 1003
Cisco Systems, Inc. v. Video Solutions Pte. Ltd
Page 000001

TABLE OF CONTENTS

Page

I. INTRODUCTION 5

II. MY QUALIFICATIONS AND PROFESSIONAL EXPERIENCE 7

III. RELEVANT LEGAL STANDARDS 13

IV. BACKGROUND 16

A. Overview of the '823 Patent 16

B. Overview of the Prosecution History of the '823 Patent 26

C. Overview of the Challenged Claims of the '823 Patent 28

D. Priority Date of Claims 31

V. CLAIM CONSTRUCTION 32

VI. STATE OF THE ART 33

A. Level of Ordinary Skill in the Art 33

B. State of the Art and the General Knowledge of a POSITA 34

1. Known "Simultaneous Traffic Peaks in Multiple Compressed Video Streams" Problem and Known "Avoiding Coincidence" Solution 38

2. Well-Known "Delay Sensitivity" Problem of Videoconferencing and Well-Known "Dynamic RTD Determination" Solution 44

3. Well-Known "Heterogeneous Packet Network Connection" Problem and Known "Avoiding Increases in the Maximum Determined Delay" Solution 49

VII. IDENTIFICATION OF HOW '823 PATENT CLAIMS 1-8 ARE UNPATENTABLE 59

A. Identification of Challenge 59

Petitioner's Exhibit 1003
Cisco Systems, Inc. v. Video Solutions Pte. Ltd
Page 000002

B. Claims 1-8 are Obvious over Larson in View of Cai, and Further in View of the Knowledge, Skill and Creativity of a POSITA 59

1. Summary of Larson (Ex-1005) 59

2. Summary of Cai (Ex-1009) 65

3. Motivation to Combine the Teachings of Larson with the Teachings of Cai 68

4. Detailed Invalidity Analysis 74

a. Claim 1 74

[1.0.1] A method of managing multiple data streams transported through a common communications resource in a packet network, [1.0.2] wherein data flowing through said resource travels in both directions, and [1.0.3] wherein each stream is subject to data peaks, comprising: 74

i. [1.1] establishing logical two-way connections through the common communications resource between one or more near endpoints and one or more far endpoints; 76

ii. [1.2] determining an undelayed round trip delay for each said logical two-way connection; 78

iii. [1.3] determining the maximum round trip delay among all said logical two-way connections; 82

iv. [1.4] preparing each of said data streams for transmission over respective said logical connections from a said near endpoint to a said far endpoint; 87

Petitioner's Exhibit 1003
Cisco Systems, Inc. v. Video Solutions Pte. Ltd
Page 000003

v. [1.5] determining the location of said data peaks in said data streams; 87

vi. [1.6] transmitting said data streams over said respective logical two-way connections to said one or more far endpoints; and.....89

vii. [1.7] delaying the transmission of one or more of said data streams that are transmitted over a said logical two-way connection having an undelayed round trip delay less than the maximum round trip delay by an amount dependent on the location of said peaks to reduce the degree of coincidence in the data peaks of different streams without increasing the maximum round trip delay for the logical two-way connections......90

b. Claim 2.....94

c. Claim 3.....94

d. Claim 4.....96

e. Claim 5.....97

f. Claims 6 97

g. Claim 7.....97

h. Claim 8.....98

APPENDIX A – MATERIALS RELIED UPON.....101

Petitioner's Exhibit 1003
Cisco Systems, Inc. v. Video Solutions Pte. Ltd
Page 000004

I, Dr. Immanuel Freedman, declare as follows:

I. INTRODUCTION

1. I am over the age of eighteen (18) and otherwise competent to make this Declaration.

2. I have been retained by Cisco Systems, Inc. (Petitioner) as an independent, technical expert consultant in this proceeding before the United States Patent and Trademark Office ("USPTO"). I am being compensated at my rate of $450 per hour for my services. No part of my compensation is dependent on my opinions or on the outcome of this proceeding. I have no financial interest in any of the parties to this proceeding.

3. This Declaration is in support of the petition ("Petition") for *inter partes* review IPR2024-00922 involving U.S. Patent No. 8,446,823 (the "'823 Patent" or "the Patent") (Ex-1001), entitled "Method of Managing the Flow of Time-Sensitive Data Over Packet Networks" and listing Mojtaba Hosseini and Patrick White as the inventors. The '823 Patent issued May 21, 2013, from U.S. Patent Application No. 12/489,977 which was filed on June 23, 2009. Ex-1001, 1.

4. For the purposes of this *inter partes* review proceeding, as I discuss later (in Section IV.D of my Declaration below), I have been instructed to assume that the effective filing date of the claims of the '823 Patent challenged by the

Petitioner's Exhibit 1003
Cisco Systems, Inc. v. Video Solutions Pte. Ltd
Page 000005

Petitioner in this *inter partes* review proceeding is the earliest identified priority date, *i.e.*, the June 23, 2009 filing date of U.S. Patent Application No. 12/489,977.

5. I understand that, according to USPTO records, the '823 Patent is currently assigned to Video Solutions Pte. Ltd. ("Video Solutions" or "Patent Owner"). I assume for purposes of this proceeding that Video Solutions is the patent owner.

6. The '823 Patent is generally directed to the field of packet networks, and in particular to a method of managing the flow of time-sensitive data, such as voice or audio. I am familiar with the technology described in the '823 Patent as of the earliest identified priority date of June 23, 2009.

7. In preparing this Declaration, I have reviewed the '823 Patent (Ex-1001) and its prosecution history (Ex-1002), as well as the patents and documents cited herein, and I have considered these documents in light of the general knowledge in the art as of June 23, 2009. In formulating my opinions, I have relied upon my experience in the relevant art. I have also considered the viewpoint of a person of ordinary skill in the art ("POSITA") in the relevant field, as of June 23, 2009, who I describe below (Section VI.A). For convenience, the materials I considered in arriving at my opinions are listed in Appendix A of this Declaration.

8. I have been asked to provide my technical expertise, analysis, insights, and opinions regarding the '823 Patent and relevant references that form the basis of

Petitioner's Exhibit 1003
Cisco Systems, Inc. v. Video Solutions Pte. Ltd
Page 000006

the grounds of rejection set forth in the accompanying Petition for *inter partes* review of the '823 Patent. As described in detail below, I offer the following opinion in this Declaration:

- A POSITA would have found Claims 1-8 of the '823 Patent to be obvious over U.S. Patent Publication No. 2009/141800 to Larson et al. ("Larson") in view of the article "IQ-Services: Network-Aware Middleware for Interactive Large-Data Applications" ("Cai") and further in view of the knowledge, skill, and creativity of a POSITA.[1]

9. Unless otherwise indicated, I have added the emphasis in quotations to which I have cited (and reproduced) in this Declaration.

II. MY QUALIFICATIONS AND PROFESSIONAL EXPERIENCE

10. My qualifications and professional experience are described in my *Curriculum Vitae*, a copy of which can be found in Exhibit 1004. The following is a brief summary of my relevant qualifications and professional experience.

11. I earned a Bachelor of Science degree in Physics from the University of Durham, England, in 1979. I obtained a Doctorate in Physics from the University of Durham, England in 1986. Between obtaining my undergraduate and

[1] A copy of Larson can be found in Exhibit 1005, and a copy of Cai can be found in Exhibit 1009 (as well as in Ex. 0002 of the Declaration of Ingrid Hsieh-Yee, Ph.D. (Exhibit 1024)).

Petitioner's Exhibit 1003
Cisco Systems, Inc. v. Video Solutions Pte. Ltd
Page 000007

doctoral degree, I developed a microcomputer system for detecting coalmine fires and heatings as a scientist for the National Coal Board and worked as a software engineer for Laser-Scan Ltd. in Cambridge, England.

12. After obtaining my Doctorate, I served as a Research Assistant at University College London from September 1986 to June 1987, where I developed digital image processing algorithms to improve image and stereo-matching quality for a digital terrain modeling system, including software and algorithms for affine transformation, edge filtering, kriging interpolation, and image stereo-matching with sub-pixel acuity. I continued my work with digital image processing as a Research Associate at the University of Maryland, from June 1987 to September 1988. During my time at the University of Maryland, I designed algorithms for filtering, segmenting, clustering, and path planning based on digital images organized by quad-tree data structures.

13. From September 1988 to June 1994, I worked as a Senior Systems Engineer for the Hughes STX Corporation. As part of my work, I developed and applied methods for calibration and determination of spacecraft attitude and the comparison of sky maps from the Cosmic Background Explorer (COBE) mission with sky maps from other missions based on scientific data stored in a spatially-referenced database using a quad-tree data structure. In this role, I led the Systems Engineering and end-to-end development of a novel system for compressing

Petitioner's Exhibit 1003
Cisco Systems, Inc. v. Video Solutions Pte. Ltd
Page 000008

imaging and ancillary data that combined scientific modeling with statistical data compression. I was also charged with designing and developing evaluation tools to ensure user-transparent, system-wide compression of a 380-GB dynamic database at an image quality acceptable to end-user scientists. In public recognition of my work, I received National Aeronautics and Space Administration Group Achievement Awards in 1990 and 1992.

14. After June 1994, I began a six-month stint as a contract Software Engineer for the Federal National Mortgage Association in Washington D.C., for which I developed a graphical user interface to monitor and validate loan servicer input for a Loss Mitigation Project. I then served as an Independent Consultant to Optivision, Inc. for the next six months, where I researched and developed rate control algorithms and software based on the MPEG-2 Test Model 5 for the OPTIVideo^{TM} MPEG-2 video encoder, as well as adaptive quantization algorithms based on the then-JPEG-3 draft standard. In this role, I researched and developed algorithms to improve the quality of gray scale image compression for the medical imaging DICOM Standard by providing a lossless hybrid algorithm encoding image residuals with a diagonal Golomb code based an Enhanced Universal Trellis Coded Quantization algorithm.

15. Between December 1995 and March 1996, I served as a Senior Staff Engineer/Firmware Engineer for General Instrument Inc., Comstream Inc., and

Petitioner's Exhibit 1003
Cisco Systems, Inc. v. Video Solutions Pte. Ltd
Page 000009

Armor Safe Technologies Inc. At Comstream, I worked on integrating an MPEG-2 set top box with OpenTV interactive television middleware programmed in the Microtec C language ported to a Motorola 68340 processor under the pSOS operating system.

16. From January 1996-1997, I was the sole proprietor of Anugraha, where I researched and developed algorithms and processes to compress fine art photography at an image quality acceptable to artists based on the JPEG imaging standard implemented with image pre-processing and adaptive quantization. For the next year or so, I worked as an engineering contractor or consultant for various companies, working primarily on image processing systems and digital interactive television set-top boxes.

17. In October 1998, I began a six-month engagement with Rockwell Collins Inc., where I worked as a Lead Systems Engineer tasked with harmonizing requirements for an MPEG-2 in-flight entertainment system. I then worked for Sun Microsystems Inc. as a Software Engineer until November 1999. During my time at Sun Microsystems Inc., I developed a Distributed Component Object Model (DCOM) software interface between a TV control graphical user interface and a Microsoft broadcast application programming interface (API) with the goal of improving the visual quality of interactive TV displays derived from UDP/IP datagrams synchronized with MPEG-2 audio/video packet data.

Petitioner's Exhibit 1003
Cisco Systems, Inc. v. Video Solutions Pte. Ltd
Page 000010

18. For the next 22 months from January 2000 to October 2002, I worked as the Chief Systems Engineer for Media Logic Systems Ltd. During my time at Media Logic Systems, I designed and developed a live interactive television system (iSeeTV) in which customers communicate with human sales agents in video-enabled call centers. To create this system, I researched and developed tools and encoder systems to improve image quality at prescribed latency and bit rate for distributing live video and audio streams encoded via low latency methods. To perform the above, I was required to understand and implement video codec systems employing the MPEG-2 Simple Profile at Main Level (CATV), MPEG-4 Visual Profile with background sprite coding, and the H.263+ Standard (now known as H.264).

19. Since November 2002, I have been an engineering contractor, and more recently an independent consultant in mathematical modeling, for several companies, such as Cyra Technologies Inc., and Amgen Inc. I also served as a senior research fellow at Merck & Co., Inc., a manager at GlaxoSmithKline Inc., a director at Daiichi Sankyo, Inc., a senior director at Praxis Precision Medicines, and currently serve as a director at Takeda Pharmaceuticals. During this time, I have developed mathematical models and simulations related to various systems, signals, and images. Specifically, I have focused on analyzing, processing, storing, and deriving information from biomedical imaging and other data. Using the

Petitioner's Exhibit 1003
Cisco Systems, Inc. v. Video Solutions Pte. Ltd
Page 000011

information derived from these data, I have created a variety of models related to biology and the effects of drugs on the human body. In recognition of my work, I have received GlaxoSmithKline R&D Recognition Awards in 2012, 2013, and 2016, a Daiichi Sankyo recognition award in 2021, and Takeda recognition awards in 2022 and 2023.

20. In addition to my over thirty years of relevant industry experience, I have authored many publications relating to video encoding and transmission. For example, in 2003, I authored a chapter entitled "Video Compression" for the Internet Encyclopedia. As another example, in 2004, I authored the chapter entitled "Video" for the Berkshire Encyclopedia of Human-Computer Interaction. And, as an additional example, in 2007, I authored a chapter titled "Video Compression" for the Handbook of Computer Networks.

21. I am also a Senior Member of the Institute of Electrical and Electronics Engineers ("IEEE"), serve as the current Philadelphia Chapter Chair of the Communications & Information Theory Societies, and I am a former Chair of the American Association of Pharmaceutical Scientists Pharmaco-Imaging Community. I also served as the 2019 Vice Chair of the IEEE P2673 Intelligence Augmentation for Medical Imaging Standards Working Group. I also have been registered to practice as a patent agent for the United States Patent and Trademark Office since 2002 (Reg. No. 51,704).

Petitioner's Exhibit 1003
Cisco Systems, Inc. v. Video Solutions Pte. Ltd
Page 000012

22. From 2017, I have also volunteered, first as a Research Scholar, and currently as a Voluntary Researcher, with the State University of New York at Buffalo. In this role, I have provided senior authorship and mentorship for doctoral candidates in areas relating to computer modeling and estimation.

III. RELEVANT LEGAL STANDARDS

23. I am not an attorney. In preparing and expressing my opinions and considering the subject matter of the '823 Patent, I am relying on certain basic legal principles that counsel have explained to me. These principles are discussed below.

24. I have been asked to provide my opinions regarding whether the claims of the '823 Patent would have been obvious to a person having ordinary skill in the art (POSITA) at the time of the claimed invention, in light of the prior art. I understand that prior art to the '823 Patent includes patents and printed publications in the relevant art that predate the priority date of the claimed invention recited in the '823 Patent. As I identified above, for purposes of this Declaration, I am applying June 23, 2009 as the earliest identified priority date of the '823 Patent.

25. I have been informed by counsel and I understand that a patent claim is not patentable as obvious (a requirement I understand to be governed by a statute, 35 U.S.C. § 103) if the differences between the patent claim and the prior art are such that the claimed subject matter as a whole would have been obvious at the time the claimed invention was made to a person having ordinary skill in the art to which

Petitioner's Exhibit 1003
Cisco Systems, Inc. v. Video Solutions Pte. Ltd
Page 000013

the subject matter pertains. I have also been informed by counsel and understand that the obviousness analysis takes into account factual inquiries including the level of ordinary skill in the art, the scope and content of the prior art, and the differences between the prior art and the claimed subject matter.

26. I have been informed by counsel and I understand that whether there are any relevant differences between the prior art and the claimed invention is to be analyzed from the view of a person of ordinary skill in the relevant art at the time of the invention. As such, my opinions below as to a person of ordinary skill in the art are as of the time of the invention, even if not expressly stated as such; for example, even if stated in the present tense.

27. In analyzing the relevance of the differences between the claimed invention and the prior art, I have been informed by counsel and I understand that I must consider the impact, if any, of such differences on the obviousness or non-obviousness of the invention as a whole, not merely some portion of it. The person of ordinary skill in the art faced with a problem is able to apply his or her experience and ability to solve the problem and also look to any available prior art to help solve the problem.

28. I have also been informed by counsel and I understand that a precise teaching in the prior art directed to the subject matter of the claimed invention is not needed. I have been informed by counsel and I understand that one may take into

Petitioner's Exhibit 1003
Cisco Systems, Inc. v. Video Solutions Pte. Ltd
Page 000014

account the inferences and creative steps that a person of ordinary skill in the art would have employed in reviewing the prior art at the time of the invention. For example, if the claimed invention combined elements known in the prior art and the combination yielded results that were predictable to a person of ordinary skill in the art at the time of the invention, then this evidence would make it more likely that the claim was obvious. On the other hand, if the combination of known elements yielded unexpected or unpredictable results, or if the prior art teaches away from combining the known elements, then this evidence would make it more likely that the claim that successfully combined those elements was not obvious.

29. I have been informed by counsel and I understand that the U.S. Supreme Court has recognized several rationales for combining the teachings of different references or modifying the teaching of a reference to show obviousness of claimed subject matter. Some of these rationales include the following: (a) combining prior art elements according to known methods to yield predictable results; (b) simple substitution of one known element for another to obtain predictable results; (c) use of a known technique to improve a similar device (method, or product) in the same way; (d) applying a known technique to a known device (method, or product) ready for improvement to yield predictable results; (e) choosing from a finite number of identified, predictable solutions, with a reasonable expectation of success; and (f) some teaching, suggestion, or motivation in the prior

Petitioner's Exhibit 1003
Cisco Systems, Inc. v. Video Solutions Pte. Ltd
Page 000015

art that would have led one of ordinary skill to modify the prior art reference or to combine prior art reference teachings to arrive at the claimed invention.

30. I have been informed by counsel and I understand that objective indicia can be important evidence regarding whether a patent claim is obvious or nonobvious, if it has an appropriate nexus to the claimed invention, *i.e.*, if it is a result of the merits of a claimed invention (rather than the result of design needs or market-pressure advertising or similar activities). Such indicia include: (a) commercial success of products covered by the patent claims; (b) a long-felt need for the invention; (c) failed attempts by others to make the invention; (d) copying of the invention by others in the field; (e) unexpected results achieved by the invention as compared to the closest prior art; (f) praise of the invention by the infringer or others in the field; (g) the taking of licenses under the patent by others; (h) expressions of surprise by experts and those skilled in the art at the making of the invention; (i) the patentee proceeded contrary to the accepted wisdom of the prior art, and, (j) the contemporaneous development of the subject matter claimed by others.

IV. BACKGROUND

A. Overview of the ’823 Patent

31. The ’823 Patent relates “to the field of packet networks, and in particular to a method of managing the flow of time-sensitive data, such as voice or

Petitioner’s Exhibit 1003
Cisco Systems, Inc. v. Video Solutions Pte. Ltd
Page 000016

audio", and describes such method as applicable to videoconferencing. Ex-1001, 1:7-52.

32. The '823 Patent acknowledges that it was "common to establish video conferencing sessions over IP networks" such as "LANs, WANs, or virtual networks established over the Internet." Ex-1001, 1:13-22, 2:61-3:6. The '823 Patent also acknowledges that, in "typical" videoconferencing sessions, a "virtual connection is established between a pair of video endpoints," and, in the case of more than two video endpoints, "between each source and each destination endpoint", which "stream video and audio data to each other over other virtual connections (e.g. using RTP)." Ex-1001, 1:13-22, 2:61-3:6, FIG. 1.

33. I have reproduced Figure 1 of the '823 Patent below, which the patent describes as "a schematic diagram of a typical video conferencing system", and annotated the figure to show virtual/logical[2] connections (colored purple) between each pair of endpoints (colored blue (for endpoints at the same location 1) and colored green (for endpoints at different locations 2, 3, 4)):

[2] I note that, in the patent and the file history, the Applicant used "logical connection" and "virtual connection" interchangeably. Ex-1001, 4:42-44, 1:17-22; Ex-1002, 25.

Petitioner's Exhibit 1003
Cisco Systems, Inc. v. Video Solutions Pte. Ltd
Page 000017

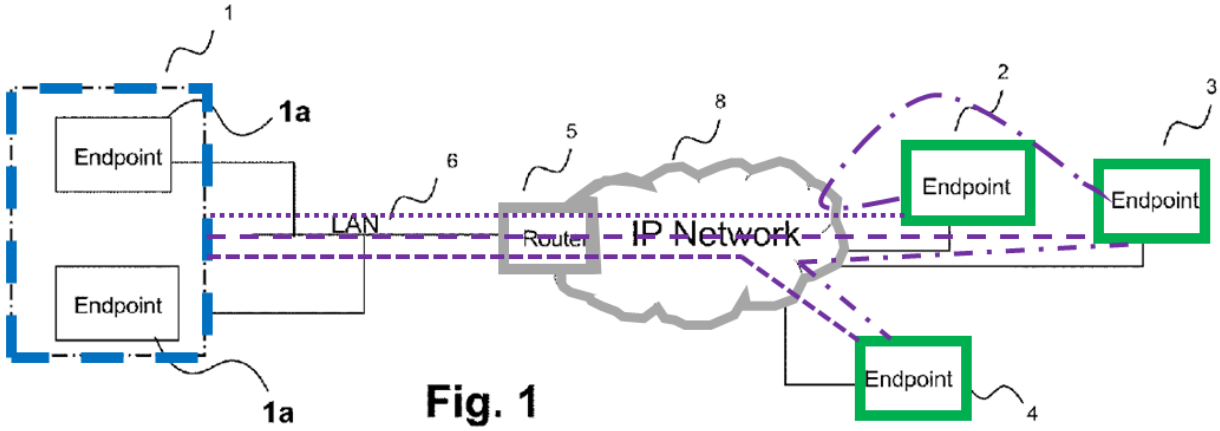


Fig. 1

Ex-1001, FIG. 1; *see also* Ex-1001, FIG. 2 (4:4-8 (patent acknowledging that, in FIG. 2, "[w]ith the exception of the DeSync Control block 32 and delay block 52, *the function of the remaining blocks is known in the art*")), 4:9-20, 4:24-58.

34. The '823 Patent also acknowledges that the video data streamed over each of these bi-directional, logical connections (which I annotate purple above) was in compressed form and, in FIG. 3 (which I also reproduce and annotate below), the patent illustrates a "typical stylized IP Video Signal 44" that includes periodic, larger data peaks called "I-frames" 44a (which contain "the entire picture" (which I annotate in red below)) among smaller "P-frames" 44b (which "only contain information about the changes relative to a previous frame"):

Petitioner's Exhibit 1003
Cisco Systems, Inc. v. Video Solutions Pte. Ltd
Page 000018

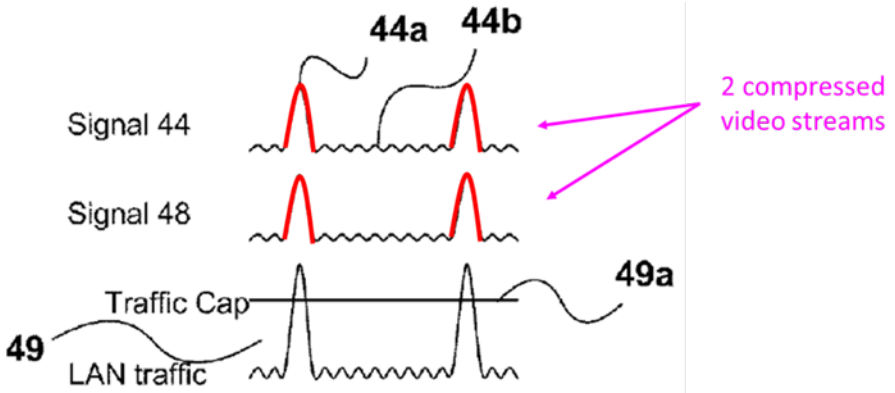


Fig. 3

Ex-1001, 1:20-36, 3:7-20, FIG. 3 (which I reproduced above).

35. The '823 Patent explains that one problem in multi-party videoconference systems is that simultaneous traffic peaks (I-frames) (44a, 48a) from multiple, substantially similar compressed video streams (44, 48)[3] may be communicated across the same network and that instantaneously exceed the

[3] The '823 Patent explains "[t]hese [traffic] peaks … will always coincide if the same source streams to multiple destinations", and that, even if "there is no particular temporal relationship between traffic peaks … coming from different [video] sources", "[f]rom time to time drifting I-frame peaks will coincide" from these "different sources." Ex-1001, 3:15-39; *see also* Section VI.G, Section VI.G.1 of my Declaration (in which I explain and identify exemplary evidence that it was known that different video sources may transmit substantially similar video streams with this "temporal relationship between traffic peaks" if, for example, they use a common group of pictures (GOP) size and sequence).

Petitioner's Exhibit 1003
Cisco Systems, Inc. v. Video Solutions Pte. Ltd
Page 000019

capacity (*e.g.*, 49a) of a common network resource (*e.g.*, router 5 of LAN 6 (FIG. 1)), resulting in congestion and packet losses. Ex-1001, 1:37-45, 3:15-39, 4:49-58. In other words, the sum of the I-frame peaks of signals 44 and 48 (shown in Figure 3 above) exceed the capacity of the common network resource (*e.g.*, of router 5 shown above) shown as traffic cap 49a. *Id*. In my Declaration, I may refer to this problem as the "simultaneous traffic peaks in multiple compressed video streams" problem.

36. The '823 Patent asserts that the solution to this multi-party videoconference problem was to intentionally delay the transmission of one or more of the substantially similar video streams in order to avoid coincidence in their compressed video data peaks (as illustrated in FIG. 4 (which I reproduce and annotate below)):

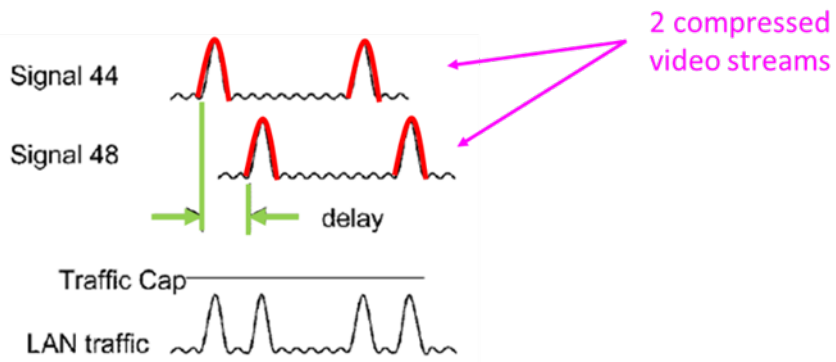


Fig. 4

Petitioner's Exhibit 1003
Cisco Systems, Inc. v. Video Solutions Pte. Ltd
Page 000020

Ex-1001, 5:18-26. The '823 Patent distinguishes its solution of "implementing a **controlled** delay" (which I annotated green above) from prior art solutions that "**indiscriminately** delay[ed] arbitrary packets" to "smooth traffic" and which "further deteriorate[d] the user experience." Ex-1001, 5:18-26, 4:1-3, 1:46-48. In my Declaration, I may refer to this solution to this problem as the "avoiding coincidence" solution.

37. The '823 Patent recognizes that, in its solution, which intentionally applies a controlled delay to one or more video streams of a videoconference and to avoid coincidence in the traffic peaks (I-frames) of such streams, it is important for the system to know **when** these simultaneous traffic peaks **will** occur among the streams. Accordingly, the '823 Patent assumes that, at the time **before** it introduces the controlled delay (*i.e.*, before it performs the "**de**sync function"), the video data from the corresponding source(s) of the video streams are "synchronized" in the sense that the streams **will** have coincident peaks (I-frames) at a given time. Ex-1001, 3:15-25, 4:49-58, 4:66-5:3, 5:12-13, 5:63-65, 6:5-7, FIGS. 3-5.

38. Specifically, the '823 Patent discloses that the system: (i) receives, "from each controlled video source", "the time the last I-frame was transmitted" by the respective source, and where each such time signal is "relative to a common arbitrary real time clock" (Ex-1001, 4:63-65, 5:63-65); (ii) assumes "that I values are roughly periodic" in each of the video streams (Ex-1001, 6:5-9, FIGS. 3-5); and

Petitioner's Exhibit 1003
Cisco Systems, Inc. v. Video Solutions Pte. Ltd
Page 000021

(iii) assumes that the "signal traffic" from each video source is "similar" (Ex-1001, 3:15-20, 3:28-39, 4:49-58, 5:18-20 (FIGS. 3-4), 5:27-30 (FIG. 5), 6:5-9, FIGS. 3-5).

39. In my opinion, a POSITA reading the '823 Patent would understand that, absent the simplistic case of a common video source[4], if the patent's system did not know the timing of **when** these simultaneous traffic peaks **will** occur among the different streams (requiring "synchronization" before "desync" among the different streams), then the system would not know **if** nor **when** the "simultaneous traffic peaks in multiple compressed video streams" problem will occur and, accordingly, would not know **if** nor **when** to apply a controlled delay to one or more of the streams.

40. As I discuss in more detail in Sections VI.B.1 and VII.B.1 of my Declaration (below), in my opinion, the same "avoiding coincidence" solution to the same "simultaneous traffic peaks in multiple compressed video streams" problem was known in the prior art.

[4] I note that, in the Challenged Claims, the "one **or more** near endpoints" (from which the "data streams" are transmitted) recited in Claim 1 is further limited to "**a common** source endpoint" in dependent Claim 7.

Petitioner's Exhibit 1003
Cisco Systems, Inc. v. Video Solutions Pte. Ltd
Page 000022

41. The ’823 Patent also explains that “round-trip delay (RTD)[5] is a very important parameter in a videoconference system”, because as “a two-way communication, like a telephone call”, the communications in each end-to-end connection of the videoconference are delay sensitive in that voice communications are seemingly natural “when the RTD is kept under approximately 150ms” and become “increasingly difficult” with RTD between 150 and 600 ms. Ex-1001, 3:46-4:3, 1:46-52. In my Declaration, I may refer to this problem as the “delay sensitivity” problem.

42. As I discuss in more detail in Section VI.B.2 of my Declaration (below), this same delay sensitivity problem was textbook knowledge for videoconferencing applications implemented over packet networks. And, as I also discuss in more detail in Sections VI.B.2 and VII.B.2-VII.B.3 of my Declaration (below), the identical solution as the ’823 Patent to this problem – techniques to dynamically determine RTD of each connection throughout a multi-party communication session

[5] A POSITA understood the ’823 Patent’s use of the term “Round Trip **Delay** (RT**D**)” refers to the standardized, well-understood “Round Trip **Time** (RT**T**)” value, including because the patent discloses that RTD is “derived from Real-Time Control Protocol (RTCP).” Ex-1001, 1:66-2:1, 4:42-48, 4:59-62; *see also* Section VI.G.2 of my Declaration below.

Petitioner’s Exhibit 1003
Cisco Systems, Inc. v. Video Solutions Pte. Ltd
Page 000023

(Ex-1001, 1:66-2:1, 4:42-48, 4:59-62) – were standardized[6], and published in textbooks and other literature, including to address the delay sensitivities of interactive applications like videoconferencing. *See* Sections VI.B.2, VII.B.2-VII.B.3 of my Declaration (below).

43. The '823 Patent identifies that another problem in multi-party videoconference systems implemented over packet networks was that the RTD experienced by each connection over a packet network may change over time, which may result in the connection with the greatest delay at one moment being replaced by a different connection at another moment. Ex-1001, 5:35-39. In my Declaration, I may refer to this problem as the "heterogeneous packet network connection" problem.

[6] I note that the identical inventors named on the face of the '823 Patent – Mojtaba Hosseini and Patrick White – are also the inventors named on the face of U.S. Pat. No. 8,446,452 (Ex-1010). I also note that U.S. Pat. No. 8,446,452 (Ex-1010) expressly acknowledges that the particular technique disclosed in the '823 Patent to solve this delay sensitivity problem (Ex-1001, 4:42-48, 4:59-62) simply involved each receiving endpoint periodically transmitting standardized information usable to calculate RTD of each connection in a standardized way. Ex-1010, 2:36-40, 4:30-40, 5:4-32, 5:44-50, FIG. 1 (*citing* Ex-1011, Ex-1012); *see also* Section VI.G.2 of my Declaration (below).

Petitioner's Exhibit 1003
Cisco Systems, Inc. v. Video Solutions Pte. Ltd
Page 000024

44. And, the ’823 Patent explains that, “other factors being equal”, its solution to this problem – dynamically determining a maximum among the determined RTDs for the various connections, and choosing to delay transmission of the stream(s) in the connection(s) that have less than the maximum RTD (Ex-1001, 4:66-5:33, 5:66-6:4, 6:22-24, FIGS. 4, 5) – will “smooth net traffic” and “ensure that no party in a multiparty conference will experience RTD greater than would be experienced by parties communicating via the path of greatest point to point delay when all other connections are disconnected” (Ex-1001, 1:66-2:7, 5:66-6:9, 6:22-24). In other words, this solution to the “heterogeneous packet network connection” problem has the benefit of reducing the delay variance across the various videoconference connections while also homogenizing the different participant experiences (by homogenizing the respective delays).

45. As I discuss in more detail in Sections VI.B.3 and VII.B.2-VII.B.3 of my Declaration (below), the same “heterogeneous packet network connection” problem was textbook knowledge for videoconferencing applications implemented over packet networks, and the same solution as the ’823 Patent to this third problem, along with similar (and additional) benefits, was known to POSITAs as of June 2009. *See* Sections VI.B.3, VII.B.2-VII.B.3 of my Declaration (below).

46. Thus, in my opinion and as I discuss in more detail below, the ’823 Patent describes and claims no more than the combination of (i) a known solution to

Petitioner’s Exhibit 1003
Cisco Systems, Inc. v. Video Solutions Pte. Ltd
Page 000025

a known problem occurring when multiple, substantially similar compressed video streams are transmitted over a packet network and through a common communication resource (*e.g.*, a router, network path, etc.), with (ii) standardized and known solutions to well-known problems occurring when implementing an interactive application (*e.g.*, videoconferencing) over a packet network, to achieve known and predictable results. *See* Sections VI.B.1-VI.B.3, VII.B.1-VII.B.4 of my Declaration (below).

B. Overview of the Prosecution History of the '823 Patent

47. The '823 Patent issued from U.S. Patent Application No. 12/489,977 which was filed on June 23, 2009. Ex-1001, p. 1. During prosecution of this application, the claims were rejected as being anticipated by U.S. Pat. No. 5,805,983 to Naidu et al. ("Naidu"). Ex-1002, 32-39. In response, the Applicant argued that Naidu's teaching of equalizing the different transmission delays that respectively existed in "separate RF links between a base station and individual antenna units" because each antenna unit was "at [a] different distance[] from the base station", was not directed to solving any of the identified problems encountered when streaming compressed video data over packet networks. Ex-1002, 25-27.

48. The Applicant then amended the claims "to place the invention more squarely within its intended context of packet networks and logical (virtual

Petitioner's Exhibit 1003
Cisco Systems, Inc. v. Video Solutions Pte. Ltd
Page 000026

connections) passing through a common resource, such as a router", and to recite that only "the data streams for which their respective logical connections have a round trip time less than the maxim[um] [are] delayed so that the peaks no longer coincide and the maximum round trip delay is not increased." Ex-1002, 25-27. This statements by the Applicant are illustrated below in the amendments to claim 1 (to which I have added bracketed notations, and preserved the strikethroughs (deletions) and underlining (additions) made by the Applicant):

> 1. *[1.0.1]* A method of managing multiple data streams transported ~~over~~ through a common communications resource in a packet network, *[1.0.2]* wherein data flowing through said resource travels~~s~~ in both directions, and *[1.0.3]* wherein each stream is subject to data peaks, comprising:
> *[1.1]* establishing logical two-way connections through the common communications resource between one or more near endpoints and one or more far endpoints,
> *[1.2]* determining an undelayed round trip delay for each said ~~each data stream~~logical two-way connection~~;~~,
> *[1.3]* determining the maximum round trip delay among all said logical two-way connections,~~and~~
> *[1.4]* preparing each of said data streams for transmission over respective said logical connections from a said near endpoint to a said far endpoint,
> *[1.5]* determining the location of said data peaks in said data streams,
> *[1.6]* ~~delaying transmission of data peaks in~~transmitting said data

Petitioner's Exhibit 1003
Cisco Systems, Inc. v. Video Solutions Pte. Ltd
Page 000027

streams over said respective logical two-way connections to said one or more far endpoints, and

[1.7] delaying the transmission of one or more of said data streams that are transmitted over a said logical two-way connection having an undelayed round trip delay less than the maximum round trip delay by an amount dependent on the location of said peaks to ~~to at least~~ reduce the degree of coincidence in the data peaks of different streams without increasing the maximum round trip delay for the ~~data streams~~ logical two-way connections.

Ex-1002, 20-21.

49. The Examiner allowed the claims based on these amendments, and the '823 Patent issued on May 21, 2013. Ex-1002, 2, 8-12.

C. Overview of the Challenged Claims of the '823 Patent

50. I reproduce claim 1 of the '823 Patent below (to which I have again added bracketed notations):

1. *[1.0.1]* A method of managing multiple data streams transported through a common communications resource in a packet network, *[1.0.2]* wherein data flowing through said resource travels in both directions, and *[1.0.3]* wherein each stream is subject to data peaks, comprising:

[1.1] establishing logical two-way connections through the common communications resource between one or more near endpoints and one or more far endpoints,

[1.2] determining an undelayed round trip delay for each said logical two-way connection,

Petitioner's Exhibit 1003
Cisco Systems, Inc. v. Video Solutions Pte. Ltd
Page 000028

[1.3] determining the maximum round trip delay among all said logical two-way connections,

[1.4] preparing each of said data streams for transmission over respective said logical connections from a said near endpoint to a said far endpoint,

[1.5] determining the location of said data peaks in said data streams,

[1.6] transmitting said data streams over said respective logical two-way connections to said one or more far endpoints, and

[1.7] delaying the transmission of one or more of said data streams that are transmitted over a said logical two-way connection having an undelayed round trip delay less than the maximum round trip delay by an amount dependent on the location of said peaks to reduce the degree of coincidence in the data peaks of different streams without increasing the maximum round trip delay for the logical two-way connections.

51. As I discussed in Section IV.A of my Declaration (above), many elements of Claim 1 – specifically, elements [1.0.1]-[1.0.3], [1.1], and [1.4]-[1.6] (as I noted with brackets above) – recite features that the '823 Patent itself expressly acknowledges were well-known to a POSITA as of June 2009. *See* Section IV.A (above); Ex-1001, 1:13-36, 2:61-3:20, FIGS. 1, 3, FIG. 2 (4:4-8 (patent acknowledging that, in FIG. 2, "[w]ith the exception of the DeSync Control block 32 and delay block 52, *the function of the remaining blocks is known in the art*")), 4:9-20, 4:24-58.

Petitioner's Exhibit 1003
Cisco Systems, Inc. v. Video Solutions Pte. Ltd
Page 000029

52. Additionally, a POSITA would recognize that there must be at least two "near endpoints" **or** at least two "far endpoints" ([1.1], [1.4], [1.6])[7] in Claim 1, because the claim requires "**multiple** data stream**s** transported through a common communication resource in a packet network" [1.0.1], "coincidence in the data peaks of **different** stream**s**" [1.7], plural "logical two-way connection**s**" established "through the common communication resource" [1.1], "preparing each of **said** data stream**s**" (plural) "for transmission over respective said logical connection**s**" (plural) [1.4], and "transmitting **said** data stream**s**" (plural) over "**said** respective logical two-way connection**s**" (plural) [1.6]. Moreover, a POSITA would also recognize that, without the "delaying" step of element [1.7] of Claim 1, the near endpoint(s) would otherwise have simultaneously transmitted two (or more) substantially similar data streams toward the far endpoint(s) in a manner that the coincident peaks (*e.g.*, I-frames) of the streams would have, at that instant in time, exceeded the capacity of the common communication resource (*e.g.*, network router, network path, etc.).

53. Once the features that the '823 Patent acknowledges to be well-known are stripped away, what remains is a claimed solution that uses the dynamically

[7] This could be, for example, (i) a common video source endpoint and two or more destination endpoints, (ii) two or more video source endpoints and a common destination endpoint, or (iii) two or more video source endpoints and two or more destination endpoints.

Petitioner's Exhibit 1003
Cisco Systems, Inc. v. Video Solutions Pte. Ltd
Page 000030

determined round-trip delay of each connection, and the maximum current delay, to decide which one (or more) of the multiple data streams to intentionally delay.

54. As I discussed in Section IV.A of my Declaration above, and in more detail in Sections VI.B.2 and VI.B.3 of my Declaration below, it was textbook knowledge that videoconferencing systems were delay sensitive, and that the delay in each videoconferencing connection over a packet network varied over time. Accordingly, as I also discussed in these Sections of my Declaration, standardized videoconferencing solutions dynamically determined round-trip time (network transit time) of each connection, and it was known that techniques that did not increase the delay of the video stream in the connection with the longest (maximum) current delay, while intentionally applying a controlled delay to the other stream(s), would reduce the delay variance across the various videoconference, homogenize the different participant experiences (by homogenizing the respective delays), reduce buffering costs, and reduce the playout buffer for the video sources. *See* Section IV.A (above); Sections VI.B.2 and VI.B.3 (below); *see also* Sections VII.B.1-VII.B.4 (below).

D. Priority Date of Claims

55. I have been instructed to assume, for the purpose of this *inter partes* review proceeding, that the priority date of claims 1-8 of the ’823 Patent is the June

Petitioner’s Exhibit 1003
Cisco Systems, Inc. v. Video Solutions Pte. Ltd
Page 000031

23, 2009 filing date of U.S. Patent Application No. 12/489,977 ("Priority Date"). Ex-1001, p. 1.

V. CLAIM CONSTRUCTION

56. It is my understanding that in order to properly evaluate the Challenged Claims of the '823 Patent, the terms of the claims must first be interpreted. It is my understanding that, for the purposes of this *inter partes* review, the claims are to be construed under the so-called *Phillips* standard, under which claim terms are given their ordinary and customary meaning as would be understood by a person of ordinary skill in the art ("POSITA", defined below in Section VI.A) as of the Priority Date (*see* Section IV.D above) in light of the specification and prosecution history, unless the inventor has set forth a special meaning for a term. That is the meaning I have applied in my analysis.

57. I have been informed by counsel that no court or administrative board has construed any of the claims of the '823 Patent. If a board or court construes any term of the '823 Patent, I reserve the right to supplement my opinions in view of such construction. I have also been informed by counsel that the Board does not construe claim terms unless the terms are in controversy, and then only to the extent necessary to resolve the controversy.

58. I have followed these principles in my analysis throughout this Declaration. It is my opinion that no specific construction of any claim term is

Petitioner's Exhibit 1003
Cisco Systems, Inc. v. Video Solutions Pte. Ltd
Page 000032

required because the relied-upon prior art meets each of the claim terms under any proper construction.

VI. STATE OF THE ART

A. Level of Ordinary Skill in the Art

59. I have been informed by counsel and understand that a person of ordinary skill in the relevant art is a hypothetical person who is presumed to be aware of all pertinent prior art, thinks along conventional wisdom in the art, and is a person of ordinary creativity. I have also been informed by counsel and understand that there are multiple factors relevant to determining the level of ordinary skill in the pertinent art, including (1) the levels of education and experience of persons working in the field at the time of the invention; (2) the sophistication of the technology; (3) the types of problems encountered in the field; and (4) the prior art solutions to those problems.

60. In my opinion, a person of ordinary skill in the art ("POSITA") in the field of the '823 Patent, as of its Priority Date (*see* Section IV.D above), would have been someone knowledgeable and familiar with the video codec technology and communications protocols and standards associated with the real-time multimedia transport arts, that are pertinent to the '823 Patent. A POSITA would have had a bachelor's degree in electrical engineering, computer science, physics, computer engineering, or an equivalent degree or equivalent training, and

Petitioner's Exhibit 1003
Cisco Systems, Inc. v. Video Solutions Pte. Ltd
Page 000033

approximately two to three years of experience working with real-time multimedia transport, including videoconferencing systems, video codecs and/or similar technology. Lack of professional experience can be remedied by additional education, and vice versa.

61. A POSITA may work as part of a multi-disciplinary team and draw upon not only his or her own skills, but also take advantage of certain specialized skills of others on the team to solve a given problem.

62. For purposes of this Declaration, in general, and unless otherwise noted, my statements and opinions, such as those regarding my experience and the understanding of a POSITA generally (and specifically related to the references I consulted herein), reflect the knowledge that existed in the field as of the Priority Date of the '823 Patent. Unless otherwise stated, when I provide my understanding and analysis below, it is consistent with the level of a POSITA prior to the Priority Date of the '823 Patent.

B. State of the Art and the General Knowledge of a POSITA

63. In this section, I describe the state of the art, including the general knowledge, skill, and creativity possessed by a POSITA as of June 2009. The prior art references that I discuss in this section, and my discussions of what was known to a POSITA during this timeframe, provide the factual support for my general description of the state of the art as of June 2009, informs the motivations and

Petitioner's Exhibit 1003
Cisco Systems, Inc. v. Video Solutions Pte. Ltd
Page 000034

reasonable expectations of success for the Ground (that I detail in Section VII of my Declaration below).

64. Streaming video data over packet networks (*e.g.*, the Internet), including for interactive (bi-directional) applications such as videoconferencing, was known and widely implemented long before the filing date of the '823 Patent. *See e.g.*, Ex-1006, 5-14; Ex-1005, [0002], [0015]; Ex-1019, 1; Ex-1020, 1. As I discussed in Section IV.A of my Declaration (above), the '823 Patent acknowledges that, as of June 2009, it was "common to establish video conferencing sessions over IP networks" such as "LANs, WANs, or virtual networks established over the Internet." Ex-1001, 1:13-22, 2:61-3:6.

65. Before video data is streamed over a packet network, it is compressed using a standardized codec, such as the Moving Pictures Expert Group (MPEG) codecs. Ex-1005, [0003]-[0004], [0016]-[0017], Ex-1001, 1:23-36, 4:4-8, 4:49-58, 3:7-14, FIG. 3; Ex-1017, [0005]-[0006], [0008]; Ex-1018, 4:19-33; Ex-1006, 11-12. A video stream includes video data that represents a series of frames that make up a video sequence. To encode a video stream, video codecs will designate different frames in different ways. Frames may be designated as reference frames (also referred to as I-frames) or non-reference frames (also referred to as P- frames). MPEG coding, for example, includes "I-frames" which "are coded using information only found in the picture itself" and "P-frames" which are predictive frames that

Petitioner's Exhibit 1003
Cisco Systems, Inc. v. Video Solutions Pte. Ltd
Page 000035

"only contain information about the changes relative to a previous frame" (and thus are "coded with respect to other P-frames or I-frames"). Ex-1005, [0003]; Ex-1001, 1:23-31; Ex-1018, 4:19-23.

66. In this manner, the video data representing the P-frames is much smaller than the data representing the I-frames. Ex-1005, [0004], [0015]; Ex-1001, 1:35-36, 3:7-14, FIG. 3 (which I reproduced in Section IV.A (above) with I-frame 44a and P-frame 44b); Ex-1017, [0008]. As I discussed in Section IV.A of my Declaration (above), the '823 Patent acknowledges that video data being streamed over packet networks was typically in compressed I-frames and P-frames. Ex-1001, 1:13-36, 2:59-3:16, FIGS. 1, 3.

67. An encoder may encode video streams into "groups of pictures" (or "GOPs"), with a GOP referring to a sequence of an I-frame and a number of P-frames, and with the GOP size (generally the number of P-frames between I-frames) being variable, with the understood tradeoff being that a higher frequency of encoding I-frames generally yields a higher picture quality but also increased bandwidth consumption. Ex-1005, [0002], [0004], [0025]-[0028], [0031]-[0032], [0041]-[0044]; FIG. 2A; Ex-1018, 4:26-5:10, 8:15-24, FIGS. 2-3; Ex-1006, 11-12, 150-152; Ex-1001, 1:49-52. Compressed video streams may include various GOPs, or the same GOP, one after the other, and may include varying positions of

Petitioner's Exhibit 1003
Cisco Systems, Inc. v. Video Solutions Pte. Ltd
Page 000036

the I-frame within the same GOP and/or different GOPs. Ex-1005, [0032]; Ex-1023, [0023]-[0024]; FIG. 2; Ex-1018, 5:22-6:7, 6:29-7:12, 8:3-24, FIGS. 4, 6.

68. As of June 2009, it was well-understood that the IETF-standardized Real-time Transport Protocol (RTP) (Ex-1011) was the predominant protocol used for streaming video and audio data over packet networks and between two or more endpoints, including for interactive (bi-directional) multimedia applications such as videoconferencing. Ex-1006, 9-11; Ex-1019, 2; Ex-1005, [0017]-[0018]; Ex-1001, 1:13-22, 4:4-8, 4:42-48. As I discussed in Section IV.A of my Declaration (above), the '823 Patent acknowledges that, when implementing RTP, it was well-understood that logical (rather than physical) two-way connections were established between endpoints at different locations. Ex-1001, 1:13-22, 2:61-3:6, FIG. 1 (logical connections (which I annotate in purple in Section IV.A) between each pair of endpoints (which I annotate in blue and green in Section IV.A), FIG. 2, 4:4-8, 4:9-20, 4:32-58; *see also* Section IV.B (in which I cite to Ex-1002, 25-26); Ex-1006, 67-69 (describing that each participant in an RTP session is identified by a logical (rather than physical) network (IP) address and a logical (rather than physical) port pair); Ex-1006, 21-22; Ex-1006, 55 (identifying ITU-T H.323 adopted RTP); Ex-1022, 6-8 (definitions of "data", "logical channel", "multipoint conference", "network address", "RTP session"); Ex-1022, 19-20, 23-

Petitioner's Exhibit 1003
Cisco Systems, Inc. v. Video Solutions Pte. Ltd
Page 000037

24, 90-94, 107; Ex-1021, 22, 42; Ex-1011 (RTP RFC); Ex-1013 (IP RFC); Ex-1015 (UDP RFC); Ex-1014 (TCP RFC).

69. As I also discussed in Section IV.A of my Declaration (above), the '823 Patent acknowledges that, in typical videoconferencing sessions, all of the endpoints are both senders and receivers, and commonly located endpoints will often stream video data over a common communication resource (*e.g.*, a router, network path, etc.). Ex-1001, 2:61-3:6, FIG. 1 (endpoints 1a, 1b at location 1 (blue) connected to IP network 8 via LAN 6 and router 5 (gray)) (reproduced and annotated in Section IV.A of my Declaration above), FIG. 2, 4:4-11, 4:14-20, 4:24-41); *see also* Ex-1005, [0019]; Ex-1019, 1.

70. Although the '823 Patent described its claimed solution as solving three (3) problems that it identified as occurring within the context of this admittedly well-known framework for streaming video data over packet networks as part of videoconferencing sessions, in my opinion and as I detail below, each of these three problems, the patent's respectively described solution to each problem, and similar (and additional) benefits to those identified by the patent for each such solution, were also known to POSITAs in June 2009.

1. Known "Simultaneous Traffic Peaks in Multiple Compressed Video Streams" Problem and Known "Avoiding Coincidence" Solution

Petitioner's Exhibit 1003
Cisco Systems, Inc. v. Video Solutions Pte. Ltd
Page 000038

71. When multiple, substantially similar video streams[8] are communicated across the same network, a well-known problem was that simultaneous traffic peaks (*e.g.,* simultaneous I-frames in multiple streams) may occur that instantaneously exceed the capacity of a common network resource (*e.g.*, a router, network path, etc.) in such network resulting in congestion and packet losses. Ex-1001, 1:37-45, 3:15-39, FIG. 3. As I identified in Section IV.A of my Declaration (above), I may refer to this well-known problem as the "simultaneous traffic peaks in multiple compressed video streams" problem.

72. In my opinion, the same solution disclosed in the '823 Patent for solving this same problem – intentionally applying a controlled delay to one or more of the video streams to avoid coincidence in their data peaks – was also known to POSITAs. Ex-1001, 5:18-26, FIG. 4. As I identified in Section IV.A of my Declaration (above), I may refer to this known solution to this well-known problem as the "avoiding coincidence" solution.

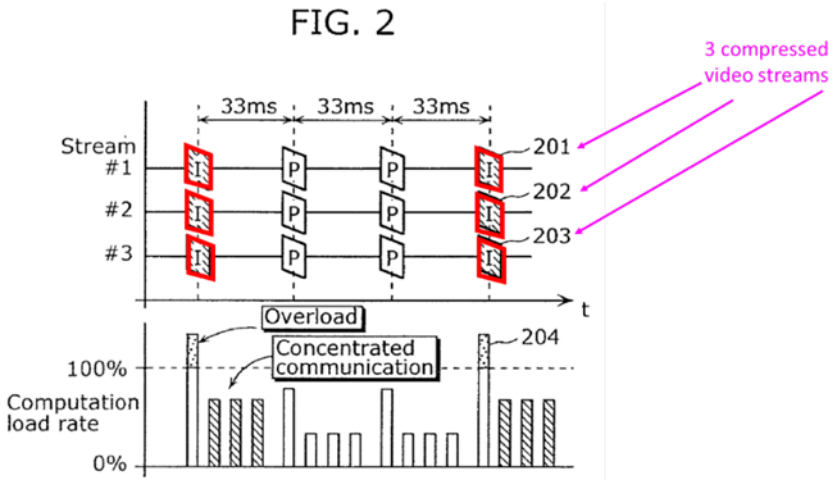
73. For example, Yamauchi (Ex-1017) describes that, when a common

[8] As I discussed in Section IV.A (footnote 3) of my Declaration (above), the '823 Patent explains "[t]hese [traffic] peaks … will always coincide if the same source streams to multiple destinations", and that, even if "there is no particular temporal relationship between traffic peaks … coming from different [video] sources", "[f]rom time to time drifting I-frame peaks will coincide" from these "different sources." Ex-1001, 3:15-39.

Petitioner's Exhibit 1003
Cisco Systems, Inc. v. Video Solutions Pte. Ltd
Page 000039

video source (*e.g.*, a "real-time streaming" source) "generat[es] coded frames of multiple data streams" for "simultaneous[] distribut[ion] to multiple communication terminals", "I-frames 201, 202 and 203 are generated within the same frame period in [the] three different data streams" and, "[t]hus, intra coding for the I-frames are overlapped, coded frames that exceed transmission capacity of the communication apparatus is generated, code amount to be transmitted increases in the period in which an I-frame is generated, causing the overload 204" and "causing packet losses":

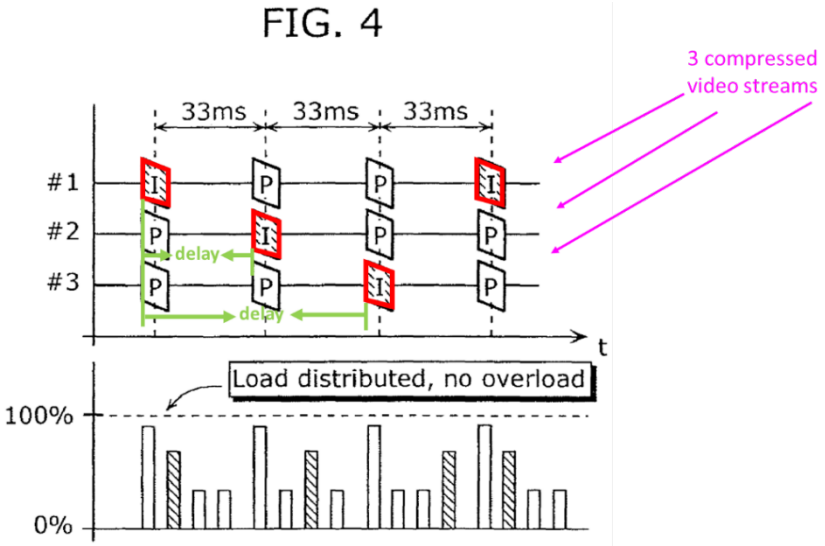
FIG. 2



Ex-1017, Abstract, [0169], [0008], [0009], [0012]-[0013], FIG. 2 (which I annotate above).

74. To solve this problem (and as I show in the illustrated example below), Yamauchi teaches intentionally applying a controlled delay to its generation of the traffic peaks (I-frames) in two of the compressed video data streams:

Petitioner's Exhibit 1003
Cisco Systems, Inc. v. Video Solutions Pte. Ltd
Page 000040



Ex-1017, FIG. 4 (annotated above), [0019], [0021], [0023], [0033], [0081]-[0086], FIG. 9; *compare with* Ex-1001, FIG. 5, 5:28-32 ("distribute I-frame peaks in signals 48 and 50 between successive peaks of Signal 44" and with P-frames (44b) distributed therebetween in each signal); *see also* Ex-1023, [0023]-[0024], FIG. 2.

75. As another example, Cooper (Ex-1018) describes its "invention" as "relat[ing] to … improved bandwidth utilization during concurrent transmission of a plurality of video signals" and describes known problems in "multi-channel video transmission system[s]" when "the I frames for all of the channels appear at the same time point":

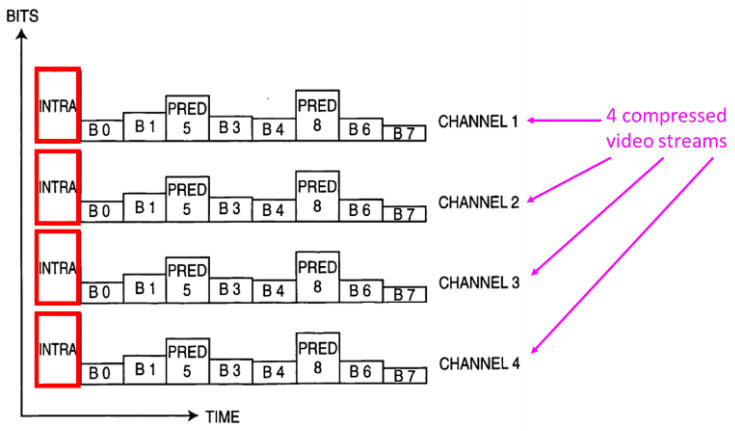


FIG. 3

Ex-1018, 1:6-8, 2:27-3:2, 5:7-21, 6:19-24, FIG. 3 (annotated above).

76. To solve this problem (and as I show in the illustrated example below), Cooper teaches that, "based on use of MPEG coding/compression of the video signals, the Intra-coded (I) frames are temporally displaced by a controlled amount so as to avoid a temporal alignment of I frames among a selected group of channels sharing a common transmission medium":

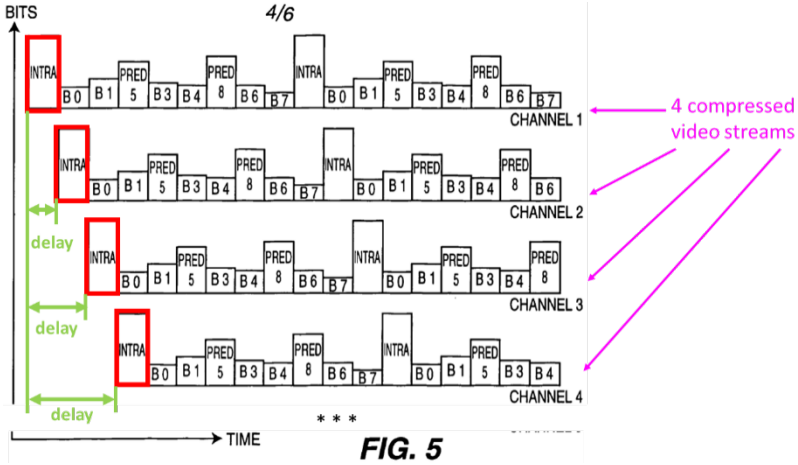


FIG. 5

Ex-1018, 3:5-8, 6:19-24; FIG. 5 (of which I annotate an excerpt (as indicated with *** above)).

Petitioner's Exhibit 1003
Cisco Systems, Inc. v. Video Solutions Pte. Ltd
Page 000042

77. As a further example, Larson (Ex-1005) solves this identical problem – simultaneous traffic peaks (I-frames) in multiple compressed video streams that instantaneously exceed the capacity of a common network resource – in the identical environment as the patent – multi-party videoconferencing applications implemented over packet networks – with the identical solution to this problem disclosed in the '823 Patent – intentionally applying a controlled delay to the transmission of one or more of the streams to avoid coincidence in their data peaks:

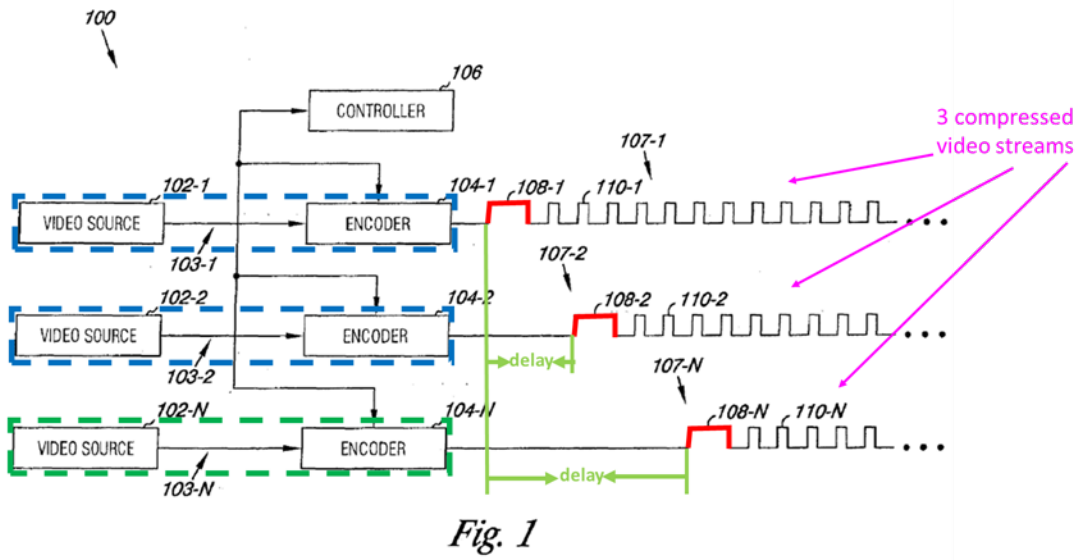


Fig. 1

Ex-1005, [0002], [0015], [0019], [0031], [0033]-[0034], FIG. 1 (annotated above); *see also* Sections VII.B.1, VII.B.4 (elements [1.0.1], [1.0.3], [1.3], [1.5], [1.7]) of my Declaration (below).

78. This state of the art delaying I-frames a controlled amount demonstrates that the '823 Patent's description that prior art systems indiscriminately delayed

Petitioner's Exhibit 1003
Cisco Systems, Inc. v. Video Solutions Pte. Ltd
Page 000043

arbitrary packets is not accurate. *See* ¶36 of Section IV.A of my Declaration (above) (in which I reference Ex-1001, 1:46-48, 4:1-3).

2. Well-Known "Delay Sensitivity" Problem of Videoconferencing and Well-Known "Dynamic RTD Determination" Solution

79. The second problem that the '823 Patent identified it was solving – the delay sensitivity of communications in each end-to-end connection of a videoconference because voice communications are seemingly natural with a round-trip delay (RTD) of less than 150ms and become increasingly difficult with a RTD between 150 and 600 ms – was textbook knowledge to POSITAs in June 2009. Ex-1001, 1:46-52, 3:46-4:3. As I identified in Section IV.A of my Declaration (above), I may refer to this well-known problem as the "delay sensitivity" problem.

80. For example, Perkins, a 2003 textbook[9], (Ex-1006) disclosed that

[9] For example, I have reviewed Exhibit 1029, which I understand to be a copy of the United States Copyright Office's Catalog Record for Colin Perkins, "RTP: Audio and Video for the Internet", Addison-Wesley, 2003 ("Perkins"), available from the Public Catalog Portal at https://cocatalog.loc.gov/cgi-bin/Pwebrecon.cgi?DB=local&PAGE=First, and which identifies that this textbook was published on July 18, 2003. Ex-1029, 1. This information is also consistent with the copyright information in the textbook itself. *See, e.g.*, Ex-1006, p. 003.

Petitioner's Exhibit 1003
Cisco Systems, Inc. v. Video Solutions Pte. Ltd
Page 000044

round-trip delay[10] was an important network parameter in interactive applications (like videoconferencing and Voice over IP calls) because the usability of such applications demanded low latency (an RTD less than approximately 300ms) in sending and receiving data over each connection:

> [R]ound-trip time is important in interactive applications because delay hinders interactivity. Studies have shown that it is difficult to conduct conversations when the total round-trip time exceeds about 300 milliseconds.

Ex-1006, 105-106;

> A system designed for interactive use-for example, video conferencing or telephony-must try to keep the playout delay as small as possible because it cannot afford the latency incurred by the buffering. Studies of human perception point to a limit in round-trip time of about 300 milliseconds as the maximum tolerable for interactive use; this limit implies an end-to-end delay of only 150 milliseconds including network transit time and buffering delay if the system is symmetric.

Ex-1006, 170; *see also* Ex-1006, 33 ("people don't notice [round-trip] delays of less than approximately 300 milliseconds (ms)."); Ex-1021, 34 ("The ITU-T G.114

[10] As I discussed in Section IV.A of my Declaration (footnote 5 above), a POSITA understood the '823 Patent's use of the term "Round Trip **Delay** (RT**D**)" refers to the standardized, well-understood "Round Trip **Time** (RT**T**)" value, including because the patent discloses that RTD is "derived from Real-Time Control Protocol (RTCP)." Ex-1001, 1:66-2:1, 4:42-48, 4:59-62.

Petitioner's Exhibit 1003
Cisco Systems, Inc. v. Video Solutions Pte. Ltd
Page 000045

recommendation limits RTT to 300 ms or less for telephone traffic."), Ex-1021, 26 ("If the delay becomes long (say over 400 or 500 ms), the conversation appears phony … [and] can be quite annoying."); Ex-1009, 11 ("interactive", "distributed applications", including "telepresence" and "remote collaboration and visualization", "stress wide area networks" because they "requir[e] data to be transferred with low latency"); Ex-1019, 1 ("the data of a videoconferencing application are all delay sensitive … media streams").

81. And, in my opinion, the identical solution disclosed in the '823 Patent for solving this identical "delay sensitivity" problem – techniques to dynamically determine RTD throughout a multi-party communication session – was standardized, and published in textbooks and other literature, including to address the delay sensitivities of interactive applications like videoconferencing. Ex-1001, 1:66-2:1, 4:42-48, 4:59-62.

82. For example, the same Perkins textbook (Ex-1006) describes IETF-standardized techniques to use information provided in RTP Control Protocol (RTCP) Receiver Reports (RR) to determine RTD between a sender and each receiver (*i.e.*, for each logical connection of a multi-party videoconferencing session):

> One of the primary uses of RTCP is reception quality reporting, which is accomplished through RTCP receiver report (RR) packets, which are sent by all participants who receive data … The reception quality

Petitioner's Exhibit 1003
Cisco Systems, Inc. v. Video Solutions Pte. Ltd
Page 000046

> feedback in RR packets is useful not only for the sender, but also for other participants and third-party monitoring tools. The feedback provided in RR packets can allow the sender to adapt its transmissions according to the feedback … A sender can use the [Last Sender Report] and [Delay since Last Sender Report] fields to calculate the round-trip time[11] between it and each receiver … The process is shown in Figure 5.4, an example taken from the RTP specification… Note that the calculated value is the network round-trip time, and it excludes any processing at the endpoints … [R]ound-trip time is important in interactive applications because delay hinders interactivity.

Ex-1006, 101, 105-106; *compare with* Ex-1001, 4:42-48, 4:60-62, 5:56-62; *see also* Section VI.B of my Declaration (above) (in which I discuss RTP logical connections); Ex-1006, 285 ("RTCP allows the sender to calculate the round-trip time to each receiver."); Ex-1011, 31, 33-34, FIG. 2 (RTP specification describing same process for a sender to determine RTT between it and each receiver); Ex-1020, 1-2 (describing known techniques to dynamically determine RTT of each connection from "RTCP feedback information" and for use in a "video teleconferencing system in which video data needs to be streamed to multiple recipients over continuously changing networks."); Ex-1019, 2-3 ("As shown in Figure 1 (reproduced below), a

[11] *See* footnote 10 in this Section VI.B.2 of my Declaration (above) and footnote 5 in Section IV.A of my Declaration (above) (identifying "round-trip **delay** (RT**D**)" and "round-trip **time** (RT**T**)" refer to the same value).

Petitioner's Exhibit 1003
Cisco Systems, Inc. v. Video Solutions Pte. Ltd
Page 000047

typical QoS-based Internet videoconferencing architecture can be described as follows: a multipoint videoconferencing system has a *Source* and a group of session participants – *Receivers*. The *Source* is capable of establishing a videoconferencing session by sending compressed live audio and video streams to the Internet via separate RTP multicast channels … While receiving the coming streams, each receiver periodically issues a series of RTCP reports to the network to report its most recent receiving status … In the SessionManager module, the *RTP Monitor* is responsible for tracing the transmission performance and monitoring each session receiver's receiving status by collecting the receiver's individual RTCP reports."); Ex-1007, 164-165 ("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 … Delay measurements already control the behaviour of TCP via its RTT estimate … for the sake of proper self-clocking … Using delay in this manner has the potential advantage that a congestion control algorithm 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

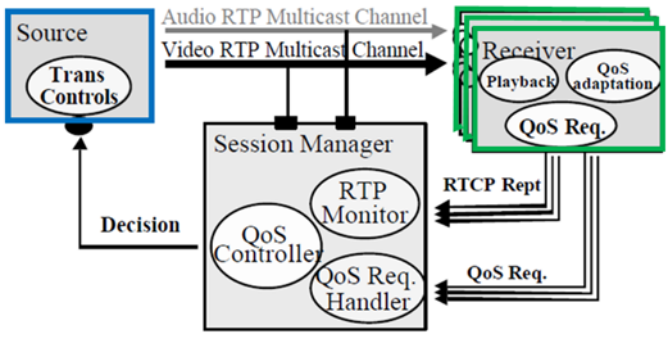


Figure 1. Overall architecture

Petitioner's Exhibit 1003
Cisco Systems, Inc. v. Video Solutions Pte. Ltd
Page 000048

not have to carry out extra processing … 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."); Ex-1009, 14 (§4.1) (disclosing dynamically determining RTT of each connection of a multi-party, interactive application).

3. Well-Known "Heterogeneous Packet Network Connection" Problem and Known "Avoiding Increases in the Maximum Determined Delay" Solution

83. The third problem that the '823 Patent identified it was solving – that the amount of network transit delay (RTD) in packet networks transporting streaming audio and/or video data varied over time, and varied among different end-to-end logical connections – was textbook knowledge to POSITAs in June 2009. Ex-1001, 5:35-39. As I identified in Section IV.A of my Declaration (above), I may refer to this well-known problem as the "heterogeneous packet network connection" problem.

84. For example, the Perkins textbook (Ex-1006) describes this same problem:

> How long does it take packets to transit the network? The answer depends on the route taken … the speed of the links, the number of routers through which a packet must pass, and the amount of queuing delay imposed by each router … In network terms a short path is most often the one with the least hops, even if it covers a longer physical distance … Measures of the delay in itself are not very interesting because they so clearly depend on the location of the source and the

Petitioner's Exhibit 1003
Cisco Systems, Inc. v. Video Solutions Pte. Ltd
Page 000049

> destination. Of more interest is how the network transit time varies from packet to packet … Some sample measurements of the variation in transit time … are presented in Figure 2.10 and Figure 2.11… Note the following … Several step changes in the average transit time, perhaps due to route changes in the network, can be observed … The transit time is not constant; rather it shows significant variation throughout the session. These are all problems that an application, or higher-layer protocol, must be prepared to handle and, if necessary, correct.

Ex-1006, 32-36;

> Sending to a group means that more things can go wrong: Reception quality is no longer affected by a single path through the network, but by the path from the source to each individual receiver. The defining factor in the measurement of loss and delay characteristics for a multicast session is *heterogeneity*.

Ex-1006, 38-39 (original emphasis);

> A multimedia session comprises several media streams, and in RTP each is transported via a separate RTP session. Because… the streams are transported separately across the network, the media will tend to have different playout times.

Ex-1006, 209; *see also* Ex-1006, 105-106 ("[T]he round-trip time … calculated value is the *network* round-trip time, and it excludes any processing at the endpoints.") (original emphasis); Ex-1006, 184 ("When RTP packets flow over a real-world IP network, variation in the interpacket timing is inevitable.")

Petitioner's Exhibit 1003
Cisco Systems, Inc. v. Video Solutions Pte. Ltd
Page 000050

85. Other examples are the Welzl textbook (Ex-1007)[12] and the Black textbook (Ex-1021)[13], which each describe this same "heterogeneous packet network connection" problem:

> [B]oth 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.

Ex-1007, 128;

> What can happen to a packet as it travels from source to destination? From an end-node perspective … It can be delayed … Delay can have

---

[12] I have reviewed the Declaration of Ingrid Hsieh-Yee, Ph.D. (Exhibit 1024) and understand that Exhibit 1007 is a true and correct copy of excerpts of "Network Congestion Control: Managing Internet Traffic", by Michael Welzl, published by John Wiley & Sons in 2005, and that Dr. Hsieh-Yee obtained from the Library of Congress. Ex-1024, p. 46 (¶77); Ex-1024 at Ex. 0003; Ex-1007. I also understand from the the Declaration of Dr. Hsieh-Yee that Welzl (Ex-1007; Ex-1024 at Ex. 0003) was distributed and otherwise made publicly accessible at least as early as May 27, 2005. Ex-1024, pp. 57-58 (¶100).

[13] I have reviewed Exhibit 1030, which I understand to be a copy of the United States Copyright Office's Catalog Record for Uyless Black, "Voice Over IP", 2000, Prentice Hall PTR ("Black"), available from the Public Catalog Portal at https://cocatalog.loc.gov/cgi-bin/Pwebrecon.cgi?DB=local&PAGE=First, and which identifies that this textbook was published on August 4, 1999. Ex-1030, 1. This information is also consistent with the copyright information in the textbook itself. *See, e.g.*, Ex-1021, p. 003.

Petitioner's Exhibit 1003
Cisco Systems, Inc. v. Video Solutions Pte. Ltd
Page 000051

several reasons: distance …, queuing, processing in the involved nodes, or retransmissions in the link layer.

Ex-1007, 15;

> [O]nly three things can happen to a packet as it traverses the network: (i) it can be delayed, (ii) it can be dropped, (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: (i) delay …, and (ii) packet drops… 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.

Ex-1007, 37;

> Round-trip time (RTT) is a measure of the time it takes to send a packet to a destination node and receive a reply from that node … [I]n the Internet … large variations to RTT do occur. Due to the asynchronous nature of the Internet, RTT is not consistent.

Ex-1021, 34;

> Hop distance is a term used to describe the number of hops between a sender and a receiver. It is a critical aspect in internet telephony because more hops means more delay, and more variable delay. Hop distance must consider round-trip time (RTT), because of the interactive, real-time nature of telephone conversations… [P]acket switches, such as routers, provide variable paths, and variable delay, which is sometimes a low delay and sometimes a lengthy delay.

Ex-1021, 36-37; Ex-1021, 135; Ex-1019, 1 ("Heterogeneity problem … [I]n a real-

Petitioner's Exhibit 1003
Cisco Systems, Inc. v. Video Solutions Pte. Ltd
Page 000052

time videoconferencing session with many participants, some receivers may suffer congestion when a link or links on the path from source to the receiver have limited bandwidth or become congested with other traffic, while others are just fine."); Ex-1020, 1 ("More and more applications and services need to transmit multimedia data over increasingly congested networks. One such application is the video teleconferencing system in which video data needs to be streamed to multiple recipients over continuously changing networks. This type of video teleconferencing system must be able to quickly adapt to changing network status."); Ex-1005, [0074].

86. And, in my opinion, the same solution disclosed in the '823 Patent for solving this third problem – dynamically determining a maximum among the network transit delays (RTDs) for the various connections, and choosing to intentionally delay transmission of the stream(s) in the connection(s) that have less than the determined maximum current delay (Ex-1001, 4:66-5:33, 5:66-6:4, 6:22-24, FIGS. 4, 5) – was also known to POSITAs, as were similar (and additional) benefits from this solution to those that the '823 Patent identifies (Ex-1001, 1:66-2:7 (which I discuss in ¶44 in Section IV.A of my Declaration above)).

Petitioner's Exhibit 1003
Cisco Systems, Inc. v. Video Solutions Pte. Ltd
Page 000053

87. For example, Duffield (Ex-1008)[14] identifies the prevalence of this same third problem ("variance of the delay distribution") in implementing multi-connection, "delay sensitive" applications, like "video teleconferences", over packet networks, and teaches that "Longest Delay First (LDF)" schemes, "that serve[] the flow with the current longest delay", solve this problem as well as "burstiness" problems:

> 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 … 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 … When excess bandwidth (i.e. bandwidth beyond what has already been guaranteed to backlogged connections) is available, the fair queueing system … 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

[14] I have reviewed the Declaration of Ingrid Hsieh-Yee, Ph.D. (Exhibit 1024) and understand that Exhibit 1008 is a true and correct copy of excerpts of "On Adaptive Bandwidth Sharing with Rate Guarantees," by Duffield et al., in *Proceedings, IEEE INFOCOM '98,* vol. 3, pp. 1122-1130, and that Dr. Hsieh-Yee obtained from the Library of Congress. Ex-1024, p. 20 (¶29); Ex-1024 at Ex. 0001; Ex-1008. I also understand from the the Declaration of Dr. Hsieh-Yee that Duffield (Ex-1008; Ex-1024 at Ex. 0001) was distributed and otherwise made publicly accessible at least as early as April 17, 1998. Ex-1024, pp. 32-33 (¶56).

Petitioner's Exhibit 1003
Cisco Systems, Inc. v. Video Solutions Pte. Ltd
Page 000054

excess bandwidth … is redistributed as follows: **1. Longest delay first (LDF) that serves the flow with current longest delay… 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 [Packet Generalized Processor Sharing (PGPS) systems] without any sacrifice of worst case guarantees.

Ex-1008, 1122 (bold emphasis added);

State Dependent Fair Queueing An illustration of the logical structure of our system is presented in Figure 1. When all scheduler queues are empty, a state dependent scheduler (SDS) is invoked … Redistributing Excess Bandwidth This section describes the schemes that we use for the state dependent component of our schedulers … 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 distribution … 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*.

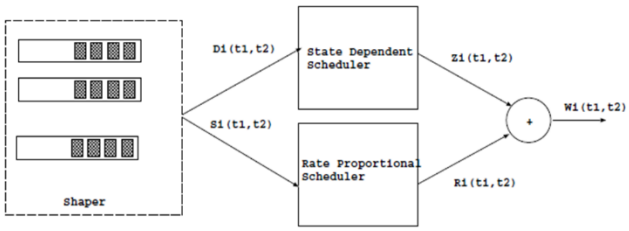


Fig. 1. Scheduling System

Petitioner's Exhibit 1003
Cisco Systems, Inc. v. Video Solutions Pte. Ltd
Page 000055

Ex-1008, 1123-1124 (italicized original emphasis);

> We simulated the system shown in Figure 1 using… Video teleconference sources … We first compared the performance of FQ-LDF with PGPS … **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** … 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 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** ... **FQ-LDF** uses the excess bandwidth to equalize delays and it **works very well when all sources have the same allowed delays**… Simulations show that **the FQ-LDF** ... **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**

Petitioner's Exhibit 1003
Cisco Systems, Inc. v. Video Solutions Pte. Ltd
Page 000056

significant when very delay sensitive traffic ... **is mixed with traffic with more laxity in delays**.

Ex-1008, 1125-1129 (bold emphasis added).

88. As established in the above excerpts, Duffield (Ex-1008) teaches that Longest Delay First techniques – that choose to intentionally delay the stream(s) in connection(s) which are dynamically determined to have less than the longest current delay – solve the same "heterogeneous packet network connection" problem and achieve similar (and additional) benefits to those identified by the '823 Patent, including "reduc[ing] the variance of the delay distribution" for "video teleconference" sources, "reduc[ing] the playout buffer for voice and video sources", and "reduc[ing] the aggregate [packet] loss rate" especially "when all sources have the same allowed delays" and when "very delay sensitive traffic… is mixed with traffic with more laxity in delays." Ex-1008, 1122, 1125-1129; *compare* Ex-1001, 1:66-2:7 (which I discuss in ¶44 in Section IV.A of my Declaration (above)).

89. As another example, Cai (Ex-1009) discloses a "'coordination' solution" to this same third problem ("the heterogeneity and dynamics of underlying [packet] networks") in this "interactive application" context (scenarios in which there is a "coexistence of related flows, such as video and visualization data streams in real-time collaboration"), which involves dynamically determining the RTT of each connection, determining a current maximum RTT (RTT_2 (connection 2)) among the determined RTTs for the various connections, and choosing to

Petitioner's Exhibit 1003
Cisco Systems, Inc. v. Video Solutions Pte. Ltd
Page 000057

intentionally delay transmission of the stream(s) in the connection(s) that have less than the current maximum determined RTT (choosing to intentionally delay transmission of connection 1 with $RTT_1 < RTT_2$). Ex-1009, 14 (§4.1); *see also* Section VII.B.2 of my Declaration (below).

90. Cai also identifies that this "'coordination' solution" achieves similar benefits to those identified by the '823 Patent for the same solution to this same third problem. *Compare* Ex-1009, 14 (benefits of "improve[d] synchronization between the two clients, thereby reducing buffering costs and increasing the 'freshness' of information received by each client") with Ex-1001, 1:66-2:7 ("delay[ing]" the "connection(s) having the least RTD", with "zero delay to the connection having greatest RTD", will "smooth net traffic" and "ensure that no party in a multiparty conference will experience RTD greater than would be experienced by parties communicating via the path of greatest point to point delay when all other connections are disconnected"); *see also* Section IV.A of my Declaration (¶44 above); Section VII.B.2 of my Declaration (below).

91. In other words, in my opinion, a POSITA would recognize that both the '823 Patent (Ex-1001) and Cai (Ex-1009) teach that this solution to the "delay sensitivity" and "heterogeneous packet network connection" problems has the benefit of reducing the delay variance across the various videoconference connections while also homogenizing the different participant experiences (by

Petitioner's Exhibit 1003
Cisco Systems, Inc. v. Video Solutions Pte. Ltd
Page 000058

homogenizing the respective delays).

92. Therefore, in my opinion, dynamically determining a current maximum network transit delay (RTD) among the various logical connections, and choosing to intentionally delay transmission of the stream(s) in the connection(s) that have less than the current maximum determined delay, was a known solution with known and predictable benefits to the well-known heterogeneity problem existing in implementing interactive, multi-party, delay-sensitive applications (like videoconferencing) over packet networks.

VII. IDENTIFICATION OF HOW '823 PATENT CLAIMS 1-8 ARE UNPATENTABLE

A. Identification of Challenge

93. Based on my understanding of the law of obviousness that I discuss in Section III of my Declaration (above), it is my opinion that Claims 1-8 of the '823 Patent are obvious over U.S. Patent Publication No. 2009/0141800 to Larson et al. ("Larson") (Ex-1005) in view of Cai (Ex-1009), and further in view of the knowledge, skill and creativity of a POSITA.

B. Claims 1-8 are Obvious over Larson in View of Cai, and Further in View of the Knowledge, Skill and Creativity of a POSITA

1. Summary of Larson (Ex-1005)

94. Larson, like the '823 Patent, discloses a technique for handling the known problem of simultaneous traffic peaks (I-frames) in multiple video streams that instantaneously exceed the capacity of a common network resource when

Petitioner's Exhibit 1003
Cisco Systems, Inc. v. Video Solutions Pte. Ltd
Page 000059

implementing multi-party videoconferencing applications over packet networks. Larson explains that "[v]ideo conferencing may … use multiple compressed video channels", that "[b]andwidth requirements for transmitting compressed video streams may be affected when one transmission line, having a given bandwidth limit, is used to transmit multiple video streams." Ex-1005, [0015], [0019], [0073]; *see also* Ex-1005, [0069] ("network 420 can be … the Internet, a local area network, a wide area network, a metropolitan area network, an enterprise network...").

95. Larson also explains that, "[f]or instance, a communication network may have three compressed video streams that are transmitted over a shared transmission line having a given bandwidth limit" and, "[w]hen, for instance, two of the three video streams are each transmitting I-frames, their respective bandwidth allocation may simultaneously be fully utilized and insufficient to transmit each I-frame as one video frame", which may result in "the communication network … hav[ing] latency due to insufficient bandwidth allocated to transmit each I-frame as one video frame." Ex-1005, [0019], [0020]; *see also* [0005] (describing that, when "bandwidth devoted to more than one data channel [is] insufficient for transmission of substantially simultaneous I-frames", the result "may be experienced by a participant in a video conference, for instance, as latency in viewing other participants in the video conference).

Petitioner's Exhibit 1003
Cisco Systems, Inc. v. Video Solutions Pte. Ltd
Page 000060

96. In my opinion, this concept is the same "simultaneous traffic peaks (I-frames) in multiple video streams" problem in the same multi-party videoconferencing environment as described in the '823 Patent, and with similar described user experience degradation results. *Compare* with Ex-1001, 3:30-39, FIG. 3.

97. Like the '823 Patent, Larson discloses a technique in which "[i]nitiation of video streams can be staggered such that I-frames do not transmit at the same time." Ex-1005, [0033]. Larson discloses that "[t]he controller 106 can, in one or more embodiments, stagger initiation of video streams from the encoders such that transmission of I-frames in the video streams are non-overlapping." Ex-1005, [0030]. More specifically, as I shown in the example in the figure that I reproduce and annotate below, Larson discloses "video streams from encoders can be staggered such that I-frames do not overlap … [for] example … the video stream 107-1 from encoder 104-1 can start first, followed by video stream 107-2 from encoder 104-2, and then followed by video stream 107-N from encoder 104-N":

Petitioner's Exhibit 1003
Cisco Systems, Inc. v. Video Solutions Pte. Ltd
Page 000061

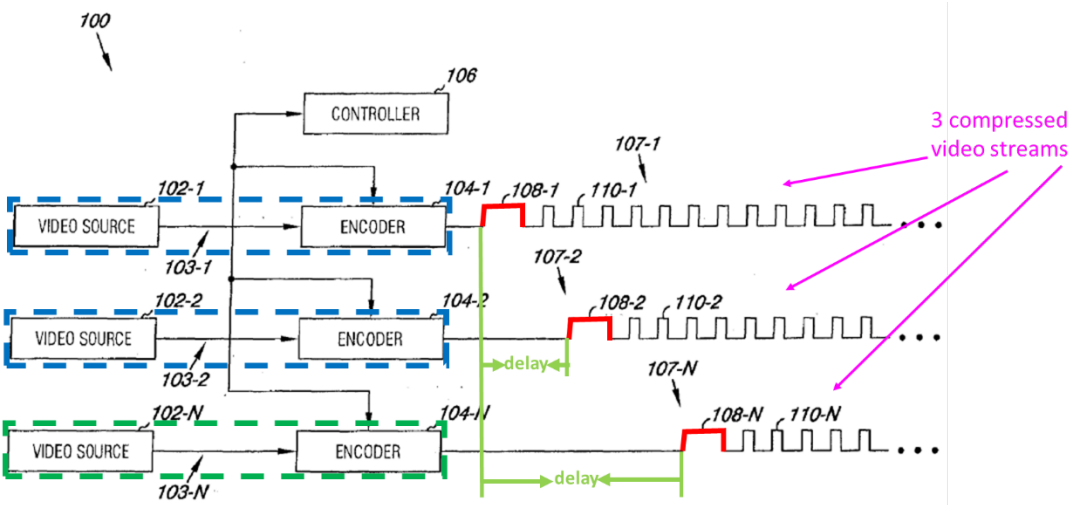


Fig. 1

Ex-1005, [0031], FIG. 1 (in which I have annotated the I-frames with red, each compressed video stream with light purple, and the intentionally applied controlled delay to video streams 107-2 and 107-N with light green).

98. Like the '823 Patent's assumption that, at the time **before** the system introduces the controlled delay (*i.e.*, before it performs the "**de**sync function"), the video data from the corresponding source(s) of the video streams are "synchronized" in the sense that the video streams **will** have coincident peaks (I-frames) at a given time, Larson recognizes the importance of the controller knowing **when** these simultaneous traffic peaks (I-frames) **will** occur among the different streams and, accordingly, Larson discloses that, in order "to transmit **later** I-frames in a non-overlapping fashion", "[o]ne or more methods of maintaining synchronization of… transmitting video streams may be implemented." Ex-1005, [0033], [0038]-[0039]; *see also* Ex-1005, [0031] ("the plurality of encoders can provide coordination

information to the controller 106"), [0032] ("[c]ontinuing coordination of transmission may be facilitated by, for example, encoding each video stream with a common number of pictures in their respective groups of pictures"), [0048], [0050]-[0051], [0060]-[0061]; *see also* Section IV.A of my Declaration (above) (in which I discuss Ex-1001, 3:15-25, 3:28-39, 4:49-58, 4:66-5:3, 5:12-13, 5:18-20, 5:27-30, 5:63-65, 6:5-9, FIGS. 3-5).

99. Larson further discloses that:

> In one or more embodiments, a controller can stagger video streams by, for example, **an interval proportional to the number of pictures per group of pictures divided by the number of video streams**. For example, an embodiment can use a group of pictures size of 18, and three video sources. In such an embodiment, a controller can stagger initiation of video streams 107 **by 18/3=6 pictures**. In this example, the controller can **start** the second video stream **six frames after the first**, and can **start** the third video stream **twelve frames after the first**. In this example, **an I-frame can be transmitted at pictures 1, 6, and 12**. The remaining pictures can be, for example, P-frames.

Ex-1005, [0037]-[0038].

100. In other words, Larson discloses intentionally delaying the transmission of one or more of the video data streams transported through a common communication resource in a packet network to reduce the degree of coincidence in the data peaks of different streams, and where the amount of the intentionally introduced delay is an amount dependent on the determined location of the data

Petitioner's Exhibit 1003
Cisco Systems, Inc. v. Video Solutions Pte. Ltd
Page 000063

peaks in the video data streams, as recited in the Challenged Claims.

101. I have reproduced a similar figure from the '823 Patent below, which shows the same "avoiding coincidence" solution to the same "simultaneous traffic peaks (I-frames) in multiple video streams" problem in the same multi-party videoconferencing environment as Larson:

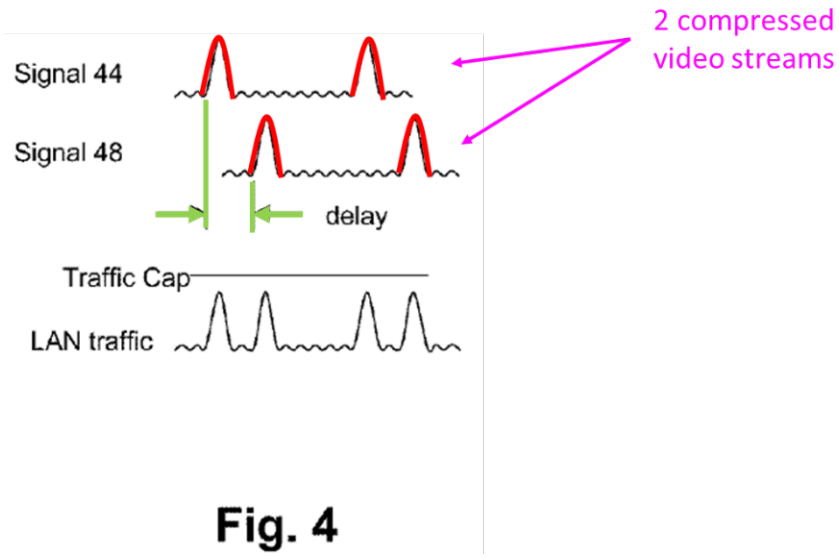


Fig. 4

Ex-1001, 5:18-26, FIG. 4 (which I reproduced and annotated above).

102. Larson does not explicitly state how to identify which compressed video stream(s) to intentionally apply its controlled delay.

103. However, as I explain in detail below, in my opinion, a POSITA would recognize that Larson does not go into detail on which compressed video stream to delay because (i) a POSITA would have the requisite background knowledge on making this selection, and (ii) it was known in the art that introducing additional delay to the video stream of the connection with the maximum (longest) current

Petitioner's Exhibit 1003
Cisco Systems, Inc. v. Video Solutions Pte. Ltd
Page 000064

delay, especially in scenarios where the various videoconferencing connections have widely disparate network transit delays (*e.g.*, participants in widely disparate geographic locations) and with at least one of such delays approaching the well-known usability limit for interactive applications (300ms), would make the performance and user experience of the multi-party, delay-sensitive videoconference worse. *See* my analysis in claim elements [1.3] and [1.7] (below); Section VI.B.2 of my Declaration (Ex-1006, 33, 101-106, 170, 285; Ex-1021, 26, 34; Ex-1009, 11, 14 (§4.1); Ex-1007, 164-165; Ex-1019, 1, 2-3; Ex-1020, 1-2); Section VI.B.3 of my Declaration (Ex-1006, 32-36, 38-39, 105-106, 184, 209; Ex-1007, 15, 37, 128; Ex-1021, 34, 36-37, 135; Ex-1019, 1; Ex-1020, 1; Ex-1008, 1122, 1125-1129; Ex-1009, 14 (§4.1); Ex-1005, [0074]).

2. Summary of Cai (Ex-1009)

104. Cai, like the '823 Patent, discloses a technique, when implementing "interactive", "distributed applications" "over multiple connections", including "telepresence" and "remote collaboration and visualization" applications, for handling the known problems associated with the "heterogeneity and dynamics of underlying [packet] networks" and that such applications "interactive nature … requir[es] data to be transferred with low latency." Ex-1009, 11, 14 (§4.1); Ex-1009, 12 (§2: "Real-Time Collaboration … environments … pose problems to developers both because of the heterogeneous platforms and dynamic user behaviors, and

Petitioner's Exhibit 1003
Cisco Systems, Inc. v. Video Solutions Pte. Ltd
Page 000065

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."); *see also* Sections VI.B.2 and VI.B.3 of my Declaration (above).

105. Like the '823 Patent, Cai discloses a "'coordination' solution" to these same problems in this same "interactive application" context (scenarios in which there is a "coexistence of related flows, such as video and visualization data streams in real-time collaboration"), and where this solution involves dynamically determining the RTT of each connection, determining a current maximum RTT (RTT_2 (connection 2)) among the determined RTTs for the connections, and choosing to intentionally delay transmission of the stream(s) in the connection(s) that have less than the current maximum determined RTT (choosing to intentionally delay transmission of the stream in connection 1 with $RTT_1 < RTT_2$):

> 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 experiment uses two connections that share a bottleneck link but have different RTTs and therefore, different throughput**… To better synchronize the two connections, the [Dynamic Adaptation Layer (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

Petitioner's Exhibit 1003
Cisco Systems, Inc. v. Video Solutions Pte. Ltd
Page 000066

> 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.

Ex-1009, 14 (§4.1).

106. Cai also identifies that this "'coordination' solution" achieves similar benefits to those identified by the '823 Patent for the same solution to these same problems. *Compare* Ex-1009, 14 (benefits of "improve[d] synchronization between the two clients, thereby reducing buffering costs and increasing the 'freshness' of information received by each client") with Ex-1001, 1:66-2:7 ("delay[ing]" the "connection(s) having the least RTD", with "zero delay to the connection having greatest RTD", will "smooth net traffic" and "ensure that no party in a multiparty conference will experience RTD greater than would be experienced by parties communicating via the path of greatest point to point delay when all other connections are disconnected"); *see also* Section VI.B.3 of my Declaration (above).

107. In other words, in my opinion, a POSITA would recognize that both the '823 Patent (Ex-1001) and Cai (Ex-1009) teach that this solution to the "delay sensitivity" and "heterogeneous packet network connection" problems has the benefit of reducing the delay variance across the various videoconference connections while also homogenizing the different participant experiences (by

Petitioner's Exhibit 1003
Cisco Systems, Inc. v. Video Solutions Pte. Ltd
Page 000067

homogenizing the respective delays).

3. Motivation to Combine the Teachings of Larson with the Teachings of Cai

108. The ’823 Patent (Ex-1001), Larson (Ex-1005) and Cai (Ex-1009) are all directed to implementing multi-party interactive applications (like videoconferencing) over a packet network, where there are multiple logical connections between participants in different locations and where multiple, substantially similar video streams are communicated through a shared communication resource. *See, e.g.*, Ex-1001, 1:13-22, 2:61-3:6, 3:15-20, 4:4-11, 4:14-20, 4:24-41, FIGS. 1-3 (which I discuss, for example, in Section IV.A of my Declaration (above)); Ex-1005, [0015], [0019], [0070], FIGS. 1, 4 (which I discuss, for example, in Section VII.B.1 of my Declaration (above)); Ex-1009, 11, 41 (§4.1) (which I discuss, for example, in Section VII.B.2 of my Declaration (above)).

109. Applying Cai’s teachings of choosing to intentionally delay transmission of the video stream(s) in the interactive application connection(s) that have less than the current maximum determined RTT among the multiple connections, to Larson’s teachings of its videoconferencing system, and in order to identify which compressed video stream(s) to intentionally apply Larson’s controlled delay among the multiple connections, would be applying a known technique to a known method ready for improvement to achieve known and predictable results.

Petitioner’s Exhibit 1003
Cisco Systems, Inc. v. Video Solutions Pte. Ltd
Page 000068

110. Combining the teachings of Larson and the teachings of Cai would be achieved through known methods.

111. These known methods include routine software application programming to dynamically determine the round-trip delay (RTT)[15] of each connection of the interactive application (as taught by Cai) by implementing the standardized RTP protocol, disclosed as being used by Larson's videoconferencing system, in the standardized manner to have Larson's controller dynamically (i) receive RTCP reports for each terminal-to-terminal connection, and (ii) use standardized information in such reports to determine the current RTT of each connection. Ex-1009, 14 (§4.1); Ex-1005, [0017]-[0019], [0031]; Ex-1006, 101-106, 285; Ex-1011, 31, 33-34, FIG. 2; *see also* Sections VI.B.2, VII.B.1 and VII.B.2 of my Declaration (above).

112. These known methods also include routine software application programming to have Larson's controller dynamically determine the current maximum (longest) RTT among the determined RTTs for the various connections (as taught by Cai), and use this determination to choose which, of the multiple compressed video stream(s), to intentionally apply Larson's controlled delay (as also

[15] *See* footnote 10 in Section VI.B.2 of my Declaration (above) and footnote 5 in Section IV.A of my Declaration (above) (identifying "round-trip **delay** (RT**D**)" and "round-trip **time** (RT**T**)" refer to the same value).

Petitioner's Exhibit 1003
Cisco Systems, Inc. v. Video Solutions Pte. Ltd
Page 000069

taught by Cai). Ex-1009, 14 (§4.1); Ex-1005, [0030]-[0031], [0040]; Ex-1008, 1122, 1125-1129 (which I discuss in Section VI.B.3 of my Declaration above); *see also* Sections VII.B.1 and VII.B.2 of my Declaration (above).

113. Cai discloses that dynamically using a maximum current RTT first (longest delay first) transmission scheme is beneficial in a multi-party, interactive application context (like videoconferencing) because it addresses the "delay sensitivity" and "heterogeneous packet network connection" problems of such applications while reducing the delay variance across the various videoconference connections while also homogenizing the different participant experiences (by homogenizing the respective delays). Ex-1009, 11, 14 (§4.1), 12 ("Real-Time Collaboration" section); *see also* Ex-1008, 1122, 1125-1129 (which I discuss in Section VI.B.3 of my Declaration above); Section VII.B.2 of my Declaration (above).

114. In my opinion, a POSITA would have been motivated to combine the teachings of Cai to identify which of Larson's compressed video stream(s) to intentionally delay because doing so would simply involve applying a known technique to a known method ready for improvement to yield known and predictable results.

115. For example, adding Cai's dynamic determination of RTT of each interactive application connection feature would be, as was textbook knowledge,

Petitioner's Exhibit 1003
Cisco Systems, Inc. v. Video Solutions Pte. Ltd
Page 000070

using the standardized and conventional control protocol (RTCP) for Larson's disclosed multimedia transport protocol (RTP) in its standardized and conventional way to determine an important network parameter to monitor during Larson's delay-sensitive videoconferencing sessions. Ex-1009, 11, 14 (§4.1); Ex-1005, [0017]-[0019], [0031]; *see also* Section VI.B.2 of my Declaration (above) (in which I discuss Ex-1006, 33, 101-106, 170, 285; Ex-1011, 31, 33-34, FIG. 2; Ex-1021, 26, 34; Ex-1007, 164-165; Ex-1019, 1, 2-3; Ex-1020, 1-2); Section VII.B.2 of my Declaration (above).

116. And, for example, in my opinion, a POSITA would readily recognize that adding Cai's features of dynamically ranking the determined connection RTTs from longest to shortest, while applying a maximum current RTT first transmission scheme, to identify which of Larson's compressed video stream(s) to intentionally delay would be a suitable design choice among a finite number of network parameters from which this selection decision could be made without adversely impacting the performance and user experience of the multi-party videoconference, especially in scenarios where the various connections have widely disparate determined RTTs (*e.g.*, participants in widely disparate geographic locations) and with an RTT approaching the well-known usability limit for interactive applications (300ms). *Id*.; *see also* Section VI.B.3 of my Declaration (above) (in which I discuss Ex-1006, 32-36, 38-39, 105-106, 184, 209; Ex-1021, 34, 36-37,

Petitioner's Exhibit 1003
Cisco Systems, Inc. v. Video Solutions Pte. Ltd
Page 000071

135; Ex-1007, 15, 37, 128; Ex-1019, 1; Ex-1020, 1; Ex-1008, 1122, 1125-1129; Ex-1005, [0074], Ex-1009, 14 (§4.1)).

117. In my opinion, in view of my analysis in this Section VII.B.3 of my Declaration (and as I discuss further in my claims analysis in Section VII.B.4 of my Declaration (below)):

a. Larson and Cai themselves provided teaching, suggestion or motivation that would have led a POSITA to identify which of Larson's compressed video streams to intentionally delay among Larson's multiple videoconferencing connections, using Cai's features of dynamically determining the maximum current RTT among the videoconference connections, and using this value as the criteria for selecting the connection(s) to intentionally apply a controlled delay (namely to the stream(s) in the connection(s) that currently have determined RTTs less than this value), because these features addressed the "heterogeneity" and "delay sensitivity" problems experienced in the videoconferencing environment in which Larson's avoiding coincidence solution was disclosed as being implemented. *See* my analysis for claim elements [1.3] and [1.7] in Section VII.B.4 of my Declaration (below).

b. The general knowledge in the art was that Cai's use of dynamically determined RTTs in each connection of a multi-party videoconferencing

Petitioner's Exhibit 1003
Cisco Systems, Inc. v. Video Solutions Pte. Ltd
Page 000072

application was an accepted, known, predictable and standardized (for RTP) mechanism to provide additional feedback information for controlling implementation of this delay-sensitive application by using software programming, and the use of this programming was identified as achieving known "homogenizing" benefits in the '823 Patent (the same known benefits disclosed by Cai and other prior art (*e.g.*, EX-1008 (Duffield)). *See* my analysis for claim element [1.2] in Section VII.B.4 of my Declaration (below); *see also* Sections VI.B.2, VII.B.2 of my Declaration (above).

c. The beneficial aspects of implementing Larson's avoiding coincidence solution in Larson's videoconferencing environment based on choosing to intentionally delay the compressed video stream(s) of the videoconferencing connection(s) currently experiencing less than the maximum current RTT among all connections, would be accomplished using a straight-forward exercise of programming per standardized RTP techniques and Cai's known techniques. *See* my analysis for claim elements [1.3] and [1.7] in Section VII.B.4 of my Declaration (below); *see also* Sections VI.B.3, VII.B.2 of my Declaration (above).

d. A POSITA would also be confident the combination would have a reasonable likelihood of success given that such programming was well

Petitioner's Exhibit 1003
Cisco Systems, Inc. v. Video Solutions Pte. Ltd
Page 000073

within the skill level of a POSITA and was disclosed in multiple references as resolving textbook knowledge problems, while achieving the same (and additional) benefits identified by the '823 Patent, in the same multi-party videoconferencing environment. *See* my analysis for claim elements [1.3] and [1.7] in Section VII.B.4 of my Declaration (below); *see also* Sections VI.B.3, VII.B.2 of my Declaration (above).

4. Detailed Invalidity Analysis

a. <u>Claim 1</u>

[1.0.1] A method of managing multiple data streams transported through a common communications resource in a packet network, [1.0.2] wherein data flowing through said resource travels in both directions, and [1.0.3] wherein each stream is subject to data peaks, comprising:

118. Larson discloses a method of managing multiple "compressed video streams" [*managing multiple data streams*] transported through "a shared transmission line having a given bandwidth limit" [*common communications resource*] in "a communication network." Ex-1005, [0015], [0019], [0021]. Larson discloses that such encoded video data would be transmitted as "packets" over such a communication network [*packet network*]. Ex-1005, [0002], [0017]-[0018], [0065], [0069] (disclosing well-known packet networks as examples of the "videoconferencing network in which embodiments of the present disclosure can be

Petitioner's Exhibit 1003
Cisco Systems, Inc. v. Video Solutions Pte. Ltd
Page 000074

implemented"); Ex-1021, 31-32 (identifying these disclosed examples as being packet networks).

119. Additionally, as I identify in Section VI.B of my Declaration (above), the '823 Patent acknowledges that, in typical videoconferencing sessions, commonly located videoconferencing terminals (such as Larson's 418-1 and 418-2) will often stream video data over a common communication resource (*e.g.*, a router). Ex-1005, [0070], [0019], FIG. 4; Ex-1001, 2:61-3:6, FIG. 1, FIG. 2, 4:4-11, 4:14-20, 4:24-41; Ex-1019, 1. Accordingly, in my opinion, a POSITA understood Larson's multiple compressed video streams, sent to/from commonly located terminals, to be transported through a common resource, such as a router [*common communications resource*].

120. In my opinion, a POSITA understood that Larson discloses compressed video data flowing through the common communication resource in both directions as recited in the Challenged Claims. For example, Larson discloses that "[v]ideo conferencing may … use multiple compressed video channels with reasonably predictable bandwidth requirements", which "may be affected when one transmission line, having a given bandwidth limit, is used to transmit" the "compressed video streams" carrying such videoconferencing data, and Larson's preferred embodiments are directed to videoconferencing applications. Ex-1005, [0015], [0019], [0025], [0065]-[0074], FIG. 4. Additionally, Larson discloses

Petitioner's Exhibit 1003
Cisco Systems, Inc. v. Video Solutions Pte. Ltd
Page 000075

videoconferencing embodiments that include "video sources originating from different sites", with some "videoconferencing terminals … in a same location" and others "at different locations", and in which signals from these terminals will "reach each of the other terminals." Ex-1005, [0065], [0070], [0074]; FIG. 4. And, as acknowledged by the '823 Patent, a POSITA understood conventional videoconferences were two-way communications of video data. Ex-1001, 3:48-49, 1:13-22 (discussed in Sections IV.A and VI.B of my Declaration (above)).

121. Accordingly, it is my opinion that a POSITA understood that Larson discloses compressed video data flowing through the common communication resource in both directions as recited in the Challenged Claims.

122. Larson further discloses that each compressed video stream includes "a number of I-frames" "and P-frames" and that, "[d]ue to the difference in compression ratios between I-frames and P-frames, among other factors, transmission of I-frames may occupy more bandwidth during transmission over the communications network" [*each stream is subject to data peaks*]. Ex-1005, [0028], FIG. 1, [0004], [0015], [0025], [0026].

i. [1.1] establishing logical two-way connections through the common communications resource between one or more near endpoints and one or more far endpoints;

123. Larson discloses "[v]ideo conferencing terminals 418-1, 418-2 … can be in a same location and connected to the network 420" [*one or more near*

Petitioner's Exhibit 1003
Cisco Systems, Inc. v. Video Solutions Pte. Ltd
Page 000076

endpoints] (which I annotate blue in FIG. 4 reproduced below), that "a video conference can involve video conferencing terminals from different sites, for example video conference terminals 418-4 and 418-N, as illustrated in FIG. 4" [*one or more far endpoints*] (which I annotate green in FIG. 4 reproduced below), and that these "[v]ideo conferencing terminals can … be interconnected to each other through a network 420, as illustrated in FIG. 4" (which is *a packet network* and includes *the common communications resource*). Ex-1005, [0069], [0070], [0066] ("encoder can be internal to a [video conferencing terminal (VC)]"); *see also* my above discussion for element [1.0.1].

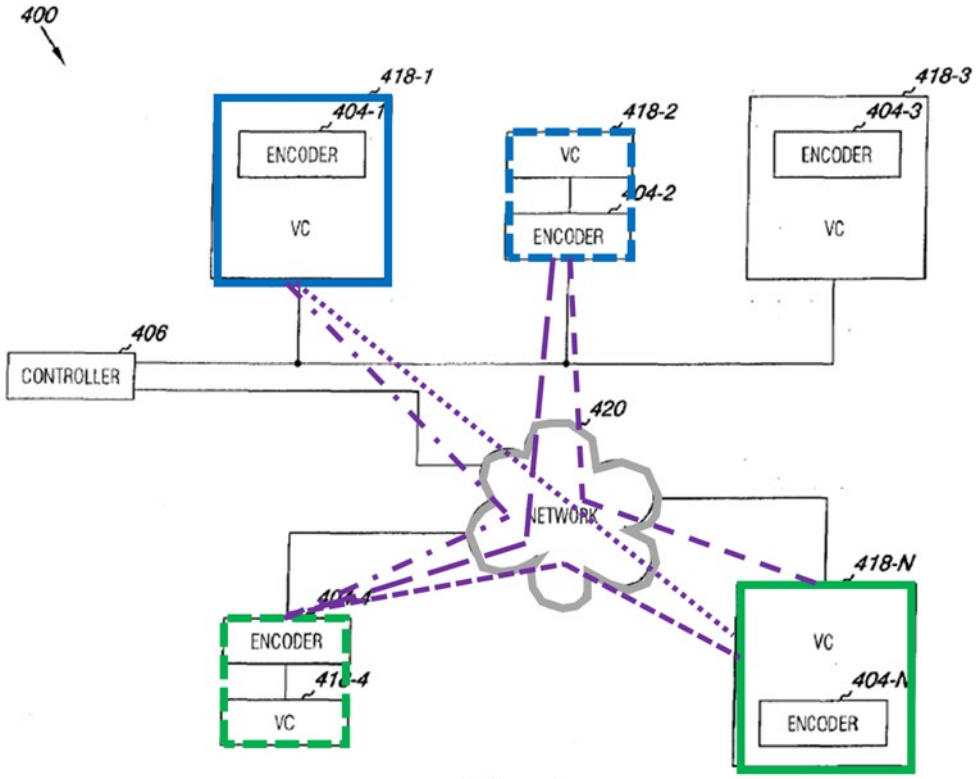


Fig. 4

Petitioner's Exhibit 1003
Cisco Systems, Inc. v. Video Solutions Pte. Ltd
Page 000077

124. Larson also discloses using the standardized "Real-Time Transport Protocol (RTP)" to transport the compressed video data in each stream over the packet network. Ex-1005, [0017]-[0019]. As I discussed in Sections IV.A, IV.B and VI.B of my Declaration (above), the '823 Patent acknowledges that, when implementing RTP, it was well understood that logical (rather than physical) two-way connections were established between endpoints, and that in typical videoconferencing sessions, all of the endpoints are both senders and receivers. Section IV.A (in which I discuss Ex-1001, 1:13-22, 2:61-3:6, FIG. 1, FIG. 2, 4:4-8, 4:9-20, 4:32-58); Section IV.B (in which I discuss Ex-1002, 25-26); Section VI.B (in which I discuss Ex-1006, 67-69, 21-22; Ex-1006, 55; Ex-1022, 6-8, 19-20, 23-24, 90-94, 107).

125. Accordingly, in my opinion, a POSITA understood Larson as disclosing logical two-way connections between its video conferencing terminals (which I annotated purple above between each pair of terminals). *Id*.

ii. [1.2] determining an undelayed round trip delay for each said logical two-way connection;

126. As I explained in my analysis of claim element [1.1] (above), in my opinion, a POSITA understood Larson disclosing implementing RTP in its standardized and conventional way, including to transport the compressed video data in each stream over the packet network, and to establish logical two-way connections between each pair of its video conferencing terminals. Ex-1005, [0017]-[0019].

Petitioner's Exhibit 1003
Cisco Systems, Inc. v. Video Solutions Pte. Ltd
Page 000078

127. Larson discloses (i) that the different connections between "[videoconferencing] terminals at different locations" may have different "latency related to transmission lengths", (ii) a feedback loop to the controller 106 in order to send it "coordination information", and (iii) the controller has a microprocessor and memory with programmable instructions. Ex-1005, [0074], [0030]-[0031], [0040]. Larson does not disclose the details in implementing RTP in its videoconferencing system for determining current round-trip delay (RTD)[16] for each logical two-way connection (element [1.2]).

128. However, as I discussed in Section IV.A of my Declaration (above), the '823 Patent (and a different patent naming the same inventors) acknowledges that a POSITA understood that, in typical endpoint-to-endpoint videoconferencing sessions, the respective signals sent over each logical connection "include both the video signal (e.g. carried in Real Time Protocol-RTP) and round trip delay (RTD) information on the RTP flow (e.g. derived from RTP Control Protocol RTCP)." Section IV.A (in which I discuss Ex-1001, 4:42-48, 4:5-9, 4:59-62; Ex-1010, 2:36-40, 4:30-40, 5:4-32, 5:44-50, FIG. 1).

[16] *See* footnote 10 in Section VI.B.2 of my Declaration (above) and footnote 5 in Section IV.A of my Declaration (above) (identifying "round-trip **delay** (RT**D**)" and "round-trip **time** (RT**T**)" refer to the same value).

Petitioner's Exhibit 1003
Cisco Systems, Inc. v. Video Solutions Pte. Ltd
Page 000079

129. In addition, as I discussed in Section VI.B.2 of my Declaration (above), it was textbook knowledge (i) that the delay sensitivity of videoconferencing applications made RTD an important network parameter to measure for each terminal-to-terminal connection, and (ii) to use standardized information provided in RTCP reports to determine the initial (*i.e.*, undelayed) RTD for each logical connection, especially in videoconferencing systems, and throughout a multi-party communication session. Section VI.B.2 (in which I discuss Ex-1006, 33, 101, 105-106, 170, 285; Ex-1011, 31, 33-34, FIG. 2; Ex-1021, 34; Ex-1020, 1-2; Ex-1019, 1-3).

130. Thus, in my opinion, a POSITA understood that Larson's implementation of RTP included determining an undelayed round trip delay for each said logical two-way connection.

131. To the extent that Patent Owner asserts that Larson does not explicitly state that it determines an undelayed round trip delay for each said logical two-way connection, in my opinion, it would have been obvious to a POSITA to modify the teachings of Larson with Cai's teachings of dynamically determining the undelayed RTD of each interactive application connection, and using standardized techniques for doing so with RTP, as I discussed in my "Motivation to Combine the Teachings of Larson with the Teachings of Cai" section (above). *See* Section VII.B.3 of my Declaration (above).

Petitioner's Exhibit 1003
Cisco Systems, Inc. v. Video Solutions Pte. Ltd
Page 000080

132. In my opinion, a POSITA would have been confident that having Larson's videoconferencing terminals (*e.g.*, 418-1, 418-2, 418-4, 418-N) communicate standardized RTCP receiver reports for each logical connection (*e.g.*, 418-1-to/from-418-N, 418-2-to/from-418-4, 418-N-to/from-418-4, etc.), and when each terminal has respectively received RTP packets over a respective connection during a respective reporting interval, provides the readily apparent and predictable benefit of using the standardized and conventional control protocol (RTCP) for Larson's disclosed multimedia transport protocol (RTP) in its standardized and conventional way. EX-1005, [0017]-[0019], [0031]; Section VI.B.2 (in which I discuss Ex-1006, 101-106, 285; Ex-1011, 31, 33-34, FIG. 2; Ex-1019, 2-3; Ex-1020, 1-2; Ex-1007, 164-165; Ex-1009, 14 (§4.1)); *see also* Section VII.B.3 of my Declaration (above). It is also my opinion that a POSITA would have been confident that having Larson's controller receive such standardized RTCP reports (*e.g.*, as part of its "coordination information") to dynamically determine RTDs of each logical connection in the standardized and conventional way would have reasonably achieved the widely understood and predictable benefits of having the controller receive additional feedback information (i) that was well-understood to be an important network parameter to monitor during Larson's delay-sensitive videoconferencing sessions, and (ii) to further inform decision-making while implementing Larson's "avoiding coincidence" solution. *Id*.

Petitioner's Exhibit 1003
Cisco Systems, Inc. v. Video Solutions Pte. Ltd
Page 000081

iii. [1.3] determining the maximum round trip delay among all said logical two-way connections;

133. As I explained in my analysis of claim elements [1.1] and [1.2] (above), in my opinion, a POSITA understood Larson disclosing implementing RTP in its standardized and conventional way, including to transport the compressed video data in each stream over the packet network, and to establish logical two-way connections between each pair of its video conferencing terminals, and it would have been obvious to POSITAs for Larson to use standardized RTCP reports, and to have Larson's controller dynamically determine RTDs of each logical connection (as taught by Cai) from such reports, in the standardized and conventional way.

134. Larson discloses (i) that the different connections between "[videoconferencing] terminals at different locations" (*e.g.*, co-located 418-1 and 418-2, different locations 418-4, 418-N) may have different "latency related to transmission lengths" – *i.e.*, it expressly recognizes the well-known "heterogeneous packet network connection" problem (*see* Section VI.B.3 (above)) –, (ii) a feedback loop to the controller 106 in order to send it "coordination information", and (iii) the controller has a microprocessor and memory with programmable instructions. Ex-1005, [0074], FIG. 4, [0030]-[0031], [0040].

135. As I explained in my summary of Larson in Section VII.B.1 (above), and in my analysis of claim element [1.7] (below), like the '823 Patent, Larson discloses solving the problem of simultaneous traffic peaks (which I annotate in red

Petitioner's Exhibit 1003
Cisco Systems, Inc. v. Video Solutions Pte. Ltd
Page 000082

below) in multiple compressed video streams (which I annotate in light purple below), in a multi-party videoconferencing application (*e.g.*, 418-1, 418-2, 418-4, 418-N) environment implemented over a packet network, with its disclosed solution that intentionally applies a controlled delay (which I annotate in light green below) to the transmission of one or more of the streams (*e.g.*, stream 107-2 from 418-2 to 418-4, stream 107-N from 418-N o 418-4), while not increasing the delay of the other (*e.g.*, stream 107-1 from 418-1 to 418-4), to avoid coincidence in their data peaks:

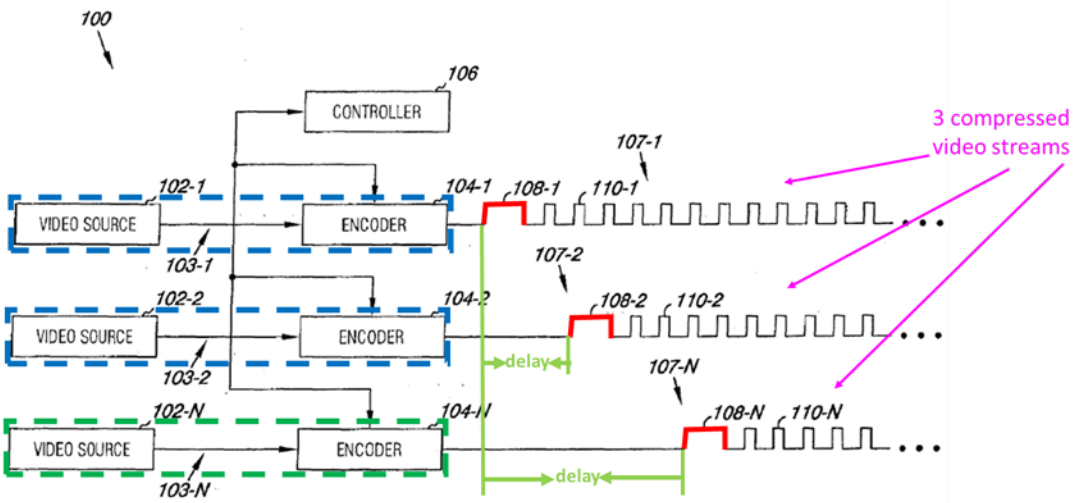


Fig. 1

Ex-1005, [0005], [0019], [0020], [0030]-[0031], [0033]-[0034]); Ex-1005, [0067] ("encoders illustrated in FIG. 4 can function substantially as described with regard to FIG[] 1"), [0068] ("controller 406 can function substantially as described with regard to FIG. 1"), [0070] ("Transmission of video sources can be coordinated between the number of sites as described in the present disclosure."), [0071] ("the

Petitioner's Exhibit 1003
Cisco Systems, Inc. v. Video Solutions Pte. Ltd
Page 000083

controller can coordinate encoding of the video sources such that transmission of an I-frame from, for example, a first video source is completed before a transmission of an I-frame from a second video source is initiated.").

136. Given that Larson discloses that its "avoiding coincidence" solution[17] was to be used to improve the performance and user experience of a multi-party videoconference being implemented over a packet network, in my opinion, a POSITA reading Larson would understand that it would **not** select to intentionally delay transmission of a video stream that would make the performance and user experience of the multi-party videoconference worse. *See also* Sections VII.B.1 and VII.B.3 of my Declaration (above).

137. And, given that it was textbook knowledge that (i) network transit delays (RTDs) in the different terminal-to-terminal connections are different from each other and changing over time in packet networks (as also recognized by Larson), and (ii) the delay sensitivity of videoconferencing applications made RTD an important network parameter to measure for each terminal-to-terminal connection, in my opinion, a POSITA reading Larson would understand that, if additional delay was intentionally added to the video stream (*e.g.*, stream 107-1) in the connection with the maximum (longest) current RTD (*e.g.*, connection between

[17] *See* my note in paragraphs 36 and 72 of my Declaration (above) regarding my use of the short-hand "'avoiding coincidence' solution" in my Declaration.

Petitioner's Exhibit 1003
Cisco Systems, Inc. v. Video Solutions Pte. Ltd
Page 000084

terminals 418-1 and 418-N), then the result would be making the experience even worse for that videoconference 418-N participant (and more likely causing that connection to be dropped), while leaving "unused" bandwidth (available from the connection(s) with less than the current maximum RTD (*e.g.*, connection between terminals 418-2 and 418-4)), both results of which are contrary to expressly stated purposes of Larson's "avoiding coincidence" solution. Ex-1005, [0005], [0018], [0020], [0073]-[0074]; Section VI.B.2 of my Declaration (in which I discuss Ex-1006, 33, 101-106, 170, 285; Ex-1021, 26, 34; Ex-1009, 11; Ex-1019, 1, 2-3; Ex-1007, 164-165; Ex-1020, 1-2); Section VI.B.3 of my Declaration (in which I discuss Ex-1006, 32-36, 38-39, 105-106, 184, 209; Ex-1007, 15, 37, 128; Ex-1021, 34, 36-37, 135; Ex-1019, 1; Ex-1020, 1; Ex-1008, 1122, 1125-1129; Ex-1009, 11, 14 (§4.1)); *see also* Section VII.B.3 of my Declaration (above).

138. In other words, a POSITA reading Larson would understand that, in implementing Larson's "avoiding coincidence" solution, especially in scenarios where the various connections have widely disparate determined RTTs (*e.g.*, participants in widely disparate geographic locations) and with an RTT approaching the well-known usability limit for interactive applications (300ms), you may cause more problems than you are solving if you do not continually "smooth" (homogenize) the delays in the various terminal-to-terminal videoconference connections so that Larson's intentional addition of delay(s) to video stream(s) in

Petitioner's Exhibit 1003
Cisco Systems, Inc. v. Video Solutions Pte. Ltd
Page 000085

one or more, but not all, connections does not cause delays in one connection (*e.g.*, connection between terminals 418-1 and 418-N) to get excessive (especially if the other connection(s) has/have the ability to absorb some additional delay). *Id*.

139. Given also that dynamically determining a current maximum (longest) network transit delay (RTD) among the various logical connections (as taught by Cai), and using this value as the criteria for selecting the connection(s) to intentionally apply a controlled delay (namely to the video stream(s) (*e.g.*, stream 107-2) in the connection(s) that currently have RTDs less than this value (*e.g.*, connection between terminals 418-2 and 418-4)) (as also taught by Cai), was a known solution to the well-known heterogeneity problem existing in implementing multi-party, delay-sensitive videoconferencing applications over packet networks (like Larson's), in my opinion, selecting this known solution – to avoid increasing the determined current maximum RTD (*e.g.*, of the stream 107-1 connection between terminals 418-1 and 418-N) – among a finite number of network parameters from which this selection decision could be made, would be a mere design choice to POSITAs with known and predictable benefits including reducing the delay variance across the various videoconference connections while also homogenizing the different participant experiences (by homogenizing the respective delays), reducing buffering costs, and reducing the playout buffer for the video sources. *See* EX-1009, 11, 12 ("Real-Time Collaboration" section), 14(§4.1)) (which I discuss in Section

Petitioner's Exhibit 1003
Cisco Systems, Inc. v. Video Solutions Pte. Ltd
Page 000086

VII.B.2 of my Declaration); *see also* Sections VI.B.3 and VII.B.3 of my Declaration (in which I discuss Ex-1008, 1122, 1125-1129; Ex-1009, 14 (§4.1) and *compare* with Ex-1001, 1:66-2:7 (which I discuss in ¶44 in Section IV.A of my Declaration above)).

iv. [1.4] preparing each of said data streams for transmission over respective said logical connections from a said near endpoint to a said far endpoint;

140. Larson discloses "two or more uncompressed video streams 103-1, 103-2 …, as illustrated in FIG. 1, having predictable bandwidth requirements, can be compressed into a number of I-frames 108-1, 108-2 … and P-frames 110-1, 110-2 …, respectively" [*preparing each of said data streams for transmission*], which, as I explained in my analysis of elements [1.1] and [1.3] above, Larson discloses as subsequently being a compressed video stream 107-1 sent over the logical connection between, for example, terminal 418-1 [*from a said near endpoint*] and terminal 418-N [*to a said far* endpoint], and a compressed video stream 107-2 sent over the logical connection between, for example, terminal 418-2 [*from a said near endpoint*] and terminal 418-4 [*to a said far* endpoint]. Ex-1005, [0028], [0022]-[0024], [0030]; FIG. 1.

v. [1.5] determining the location of said data peaks in said data streams;

Petitioner's Exhibit 1003
Cisco Systems, Inc. v. Video Solutions Pte. Ltd
Page 000087

141. Larson discloses the controller determines the location of the respective I-frame in each of the compressed video data streams that are being prepared for transmission (*i.e.*, prior to transmission) and in order to determine if they will overlap [*determining the location of said data peaks in said data streams*].

142. Like the '823 Patent's assumption that, at the time **before** the system introduces the controlled delay (*i.e.*, before it performs the "**de**sync function"), the video data from the corresponding source(s) of the video streams are "synchronized" in the sense that the video streams **will** have coincident peaks (I-frames) at a given time, Larson recognizes the importance of the controller knowing **when** these simultaneous traffic peaks (I-frames) **will** occur among the different compressed video streams and, accordingly, Larson discloses that, in order "to transmit **later** I-frames in a non-overlapping fashion", "[o]ne or more methods of maintaining synchronization of … transmitting video streams may be implemented." Ex-1005, [0033], [0038]-[0039]; *see also* Ex-1005, [0031] ("the plurality of encoders can provide **coordination information** to the controller 106"), [0032] ("In one or more embodiments, each encoder can continue to transmit more than one group of pictures such that **after the end of the first group of pictures, a new group of pictures is transmitted, which can begin with another I-frame (not shown). Continuing coordination of transmission** may be facilitated by, for example, encoding each

Petitioner's Exhibit 1003
Cisco Systems, Inc. v. Video Solutions Pte. Ltd
Page 000088

video stream with a common number of pictures in their respective groups of pictures. Thus, I-frames can be generated by initiating encoding and **by reaching the end of a group of pictures (e.g., starting a new group of pictures)**."), [0048], [0050]-[0051], [0060]-[0061]; *see also* Sections IV.A, VII.B.1 of my Declaration above (in which, at ¶¶36, 37, 98 (above), I discuss Ex-1001, 3:15-25, 3:28-39, 4:49-58, 4:66-5:3, 5:12-13, 5:18-20, 5:27-30, 5:63-65, 6:5-9, FIGS. 3-5). In other words, just like the '823 Patent, Larson discloses that the controller uses periodically received "coordination information" (*e.g.*, stream content (*e.g.*, GOP) and timing information) for each of the compressed video data streams to determine the location of the respective I-frame in each of the streams that are being prepared for transmission (*i.e.*, prior to transmission) and in order to determine if they will overlap. *Id*.

vi. [1.6] transmitting said data streams over said respective logical two-way connections to said one or more far endpoints; and

143. Larson discloses that, in implementing its "avoiding coincidence" solution[18], the controller will coordinate the transmission timing of compressed video stream 107-1 over the respective logical connection between, for example, terminals 418-1 and 418-N, and of compressed video stream 107-2 over the

[18] *See* my note in paragraphs 36 and 72 of my Declaration (above) regarding my use of the short-hand "'avoiding coincidence' solution" in my Declaration.

Petitioner's Exhibit 1003
Cisco Systems, Inc. v. Video Solutions Pte. Ltd
Page 000089

respective logical connection between, for example, terminals 418-2 and 418-4 [*said data streams over said respective logical two-way connections to said one or more far endpoint*s]. Ex-1005, [0030] ("controller 106 can … stagger initiation of video streams from the encoders such that transmission of I-frames in the video streams are non-overlapping[, and] … can coordinate transmission such that an I-frame from only one encoder is allowed to transmit at any given time."); [0071] ("the controller can coordinate encoding of the video sources such that transmission of an I-frame from, for example, a first video source is completed before a transmission of an I-frame from a second video source is initiated."); [0031], [0033], [0034]; FIGS. 1, 4; *see also* my analysis of element [1.7] below (for the actual *transmitting*).

vii. [1.7] delaying the transmission of one or more of said data streams that are transmitted over a said logical two-way connection having an undelayed round trip delay less than the maximum round trip delay by an amount dependent on the location of said peaks to reduce the degree of coincidence in the data peaks of different streams without increasing the maximum round trip delay for the logical two-way connections.

144. Larson discloses first transmitting (undelayed) one of the data streams (*e.g.*, compressed video stream 107-1 that is transmitted over the logical two-way connection between, for example, terminals 418-1 and 418-N), while delaying the transmission of another one (or more) of said data streams (*e.g.*, *delaying the transmission of compressed video stream* 107-2 that is transmitted over the

Petitioner's Exhibit 1003
Cisco Systems, Inc. v. Video Solutions Pte. Ltd
Page 000090

respective logical two-way connection between, for example, terminals 418-2 and 418-4) "such that I-frames do not overlap" in these transmitted data streams [*to reduce the degree of coincidence in the data peaks of different streams*]:

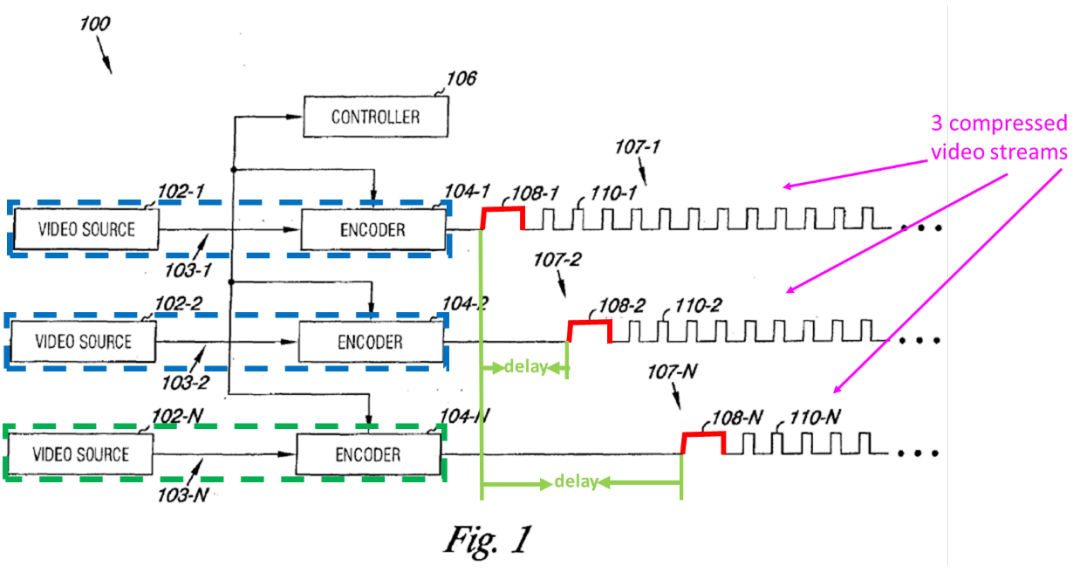


Fig. 1

Ex-1005, [0031] ("video streams from encoders can be staggered such that I-frames do not overlap… [for] example… the video stream 107-1 from encoder 104-1 can start first, **followed by** video stream 107-2 from encoder 104-2…") (emphasis added), Ex-1005, [0033]-[0034].

145. Larson also discloses that the amount of this intentionally added delay can be *an amount dependent on the location of said peaks of different streams* (element [1.7]) and where *the delays are introduced so as to evenly distribute the data peaks in different streams* (claim 6):

> [I]nitiation of transmission of the first and second video streams, for example, can be staggered… a controller can stagger video streams by, for example, **an interval proportional to the number of pictures per**

Petitioner's Exhibit 1003
Cisco Systems, Inc. v. Video Solutions Pte. Ltd
Page 000091

> **group of pictures divided by the number of video streams**. For example, an embodiment can use a group of pictures size of 18, and three video sources. In such an embodiment, a controller can stagger initiation of video streams 107 **by 18/3=6 pictures**. In this example, the controller can **start** the second video stream **six frames after the first**, and can **start** the third video stream **twelve frames after the first**. In this example, **an I-frame can be transmitted at pictures 1, 6, and 12**. The remaining pictures can be, for example, P-frames.

Ex-1005, [0037]-[0038], [0039], FIG. 1.

146. As I explained in my analysis of claim element [1.3] above, in my opinion, a POSITA would recognize that Larson does not go into detail on which compressed video stream(s) **to** delay because (i) POSITAs would have the requisite background knowledge on making this selection, and (ii) it was known in the art that introducing additional delay to the video stream of the connection with the maximum (longest) current RTD, would make the performance and user experience of the multi-party, delay-sensitive videoconference worse, especially in scenarios where the various connections have widely disparate determined RTTs (*e.g.*, participants in widely disparate geographic locations) and with an RTT approaching the well-known usability limit for interactive applications (300ms). *See* Section VII.B.2 (element [1.3]) above; Section VI.B.2 above (in which I discuss Ex-1006, 33, 101-106, 170, 285; Ex-1021, 26, 34; Ex-1009, 11; Ex-1019, 1, 2-3; Ex-1007, 164-165; Ex-1020, 1-2); Section VI.B.3 above (in which I discuss Ex-1006, 32-36, 38-39,

Petitioner's Exhibit 1003
Cisco Systems, Inc. v. Video Solutions Pte. Ltd
Page 000092

105-106, 184, 209; Ex-1007, 15, 37, 128; Ex-1021, 34, 36-37, 135; Ex-1019, 1; Ex-1020, 1; Ex-1008, 1122, 1125-1129; Ex-1009, 11, 14 (§4.1)); *see also* Section VII.B.3 of my Declaration (above).

147. Accordingly, as I also explained in my analysis of claim element [1.3] (above) and in my "Motivation to Combine the Teachings of Larson with the Teachings of Cai" section (above), in my opinion, a POSITA would have been motivated to determine the current maximum RTD among the videoconference connections (as taught by Cai), and to use this value as the criteria for selecting the connection(s) to intentionally apply the controlled delay amount (namely to the video stream(s) in the connection(s) that currently have RTDs less than this value (*e.g.*, stream 107-2 in connection between terminals 418-2 and 418-4)), without increasing the delay of the compressed video stream having current maximum RTD (*e.g.*, stream 107-1 of the connection between terminals 418-1 and 418-N) (as also taught by Cai). *Id*.

148. Further, as I explained in these same Sections of my Declaration (above), applying Cai's teachings would simply involve implementing known and standardized solutions to the "heterogeneity" and "delay sensitivity" problems that were well-known to exist in videoconferencing systems like Larson's to achieve known and predictable benefits, including reducing the delay variance across the various videoconference connections while also homogenizing the different

Petitioner's Exhibit 1003
Cisco Systems, Inc. v. Video Solutions Pte. Ltd
Page 000093

participant experiences (by homogenizing the respective delays) [*without increasing the maximum round trip delay for the logical two-way connections*] with a reasonable expectation of success. *Id*.; *see also* Sections VI.B.3 and VII.B.2 above (in which I discuss Ex-1008, 1122, 1125-1129; Ex-1009, 14 (§4.1) and *compare* with Ex-1001, 1:66-2:7 (which I discuss in ¶44 in Section IV.A of my Declaration above)).

149. Additionally, in my opinion, a POSITA would readily recognize that the current maximum determined RTD was a suitable design choice among a finite number of network parameters from which this selection decision could be made. *Id*.

b. <u>Claim 2</u>

A method as claimed in claim 1, wherein the transmission of said data streams is delayed so as to remove any coincidence in said data peaks of different data streams.

150. Larson discloses this limitation for the reasons that I discussed above regarding claim element [1.7].

c. <u>Claim 3</u>

A method as claimed in claim 1, wherein the delays introduced into the different data streams are ranked in accordance with the determined round trip delays for each logical two-way connection.

Petitioner's Exhibit 1003
Cisco Systems, Inc. v. Video Solutions Pte. Ltd
Page 000094

151. As I explained above in my analysis of claim elements [1.3] and [1.7], Larson in view of Cai renders obvious using the current maximum determined RTD, among the dynamically determined current RTDs of the videoconference connections, as the criteria to select the "one or more of said data streams" **to** delay to avoid coincidence "in the data peaks of different streams" (*e.g.*, in the respective I-frames of the not delayed stream 107-1 (the stream with the current maximum determined RTD) and the delayed stream 107-2 (the stream with a current determined RTD less than the current maximum determined RTD)). *See* Section VII.B.2 (elements [1.3], [1.7]) above.

152. In implementing this known solution for two video streams, in my opinion, a POSITA readily recognized that the current rank order of (1) zero delay to stream 107-1, (2) delay amount to stream 107-2, is in accordance with (1) current maximum RTD for connection 418-1-to/from-418-N, (2) less than current maximum RTD for connection 418-2-to/from-418-4. *See* Section VII.B.2 (element [1.7]) above; Ex-1009, 14 (rank of (1) zero delay, (2) delay of $RTT_2 - RTT_1$, because $RTT_1 < RTT_2$); Ex-1008, 1122 ("Longest delay first (LDF) … policy uses excess bandwidth to reduce the variance of the delay distribution").

153. Additionally, in my opinion, in implementing this known solution for three or more video streams, it would have been obvious to a POSITA to, consistent with achieving the known benefits of this known solution (reducing the delay

Petitioner's Exhibit 1003
Cisco Systems, Inc. v. Video Solutions Pte. Ltd
Page 000095

variance across the videoconference connections), dynamically rank the respective added delay amount from most to least (*e.g.*, delay amounts to streams 107-3, 107-2, 107-1) by connections with most to least current excess bandwidth (shortest to longest current RTD). *Id*.

d. Claim 4

A method as claimed in claim 3, wherein the amount of delay that can be introduced into a data stream is capped at a predetermined value.

154. The "maximum tolerable" limit for videoconferencing use was well-known to be an RTD of around 300ms. *See* Section VI.B.2 above (in which I discuss Ex-1006, 105-106, 170; Ex-1021, 26, 34). In my opinion, for the same reasons that a POSITA would not intentionally apply additional delay to the video stream in the videoconferencing connection with the current maximum RTD – it would make the experience even worse for that participant (and more likely cause that connection to be dropped), while leaving "unused" bandwidth – it would have been obvious to a POSITA implementing Larson's "avoiding coincidence" solution[19] to cap the amount of delay that can be introduced into any data stream such that the current determined RTD plus added delay for a connection never exceeded this 300ms limit.

[19] *See* my note in paragraphs 36 and 72 of my Declaration (above) regarding my use of the short-hand "'avoiding coincidence' solution" in my Declaration.

Petitioner's Exhibit 1003
Cisco Systems, Inc. v. Video Solutions Pte. Ltd
Page 000096

See Section VII.B.2 (elements [1.3], [1.7]) above; Section VI.B.2 above (in which I discuss Ex-1006, 105-106, 170; Ex-1021, 26, 34).

e. Claim 5

A method as claimed in claim 3, wherein the round trip delays are monitored over a period of time, and the rank order is changed as necessary to take into account changing round trip delays.

155. Larson discloses this limitation for the reasons that I discussed above regarding claim elements [1.2] and [1.3], and claim 3.

f. Claims 6

A method as claimed in claim 3, wherein the delays are introduced so as to evenly distribute the data peaks in different streams.

156. Larson discloses this limitation for the reasons that I discussed above regarding claim element [1.7]. Ex-1005, FIG. 1, [0037]-[0038] (for two stream example, discloses I-frames can be transmitted at picture 1 (stream 107-1) and picture 9 (stream 107-2); for three stream example, discloses I-frames can be transmitted at picture 1 (stream 107-1), picture 6 (stream 107-2) and picture 12 (stream 107-2)).

g. Claim 7

A method as claimed in claim 1, wherein the logical connections are established between a common source endpoint and multiple destination endpoints.

157. Larson discloses "[v]ideo conferencing terminals 418-1, 418-2, 418-

Petitioner's Exhibit 1003
Cisco Systems, Inc. v. Video Solutions Pte. Ltd
Page 000097

3… can be in a same location and connected to the network 420 collectively", but provides limited details regarding co-located terminals being part of a common video source. Ex-1005, FIG. 4, [0070].

158. In my opinion, a POSITA would recognize that Larson does not go into detail on implementing its solution for a common video source transmitting the same video data to multiple destination endpoints as respective compressed video streams over different connections because, as the '823 Patent acknowledges, a common video source was well-known to be the simplest case for a controller (like Larson's) determining **when** simultaneous traffic peaks (I-frames) **will** occur among the different compressed video streams (as I-frame "peaks … will **always** coincide if the same source streams to multiple destinations") and, accordingly, such implementation is well within the knowledge, skill and creativity of a POSITA. *See* my analysis of claim element [1.5] above; Ex-1001, 2:46-47, 2:50, 3:6-30, 3:27-28. Accordingly, in my opinion, Larson renders this limitation obvious.

h. <u>Claim 8</u>

A method as claimed in claim 1, wherein said data streams represent video streams and said data peaks correspond to I-frames.

159. Larson discloses this limitation for the reasons that I discussed above regarding claim elements [1.0.1], [1.0.3], [1.5].

Petitioner's Exhibit 1003
Cisco Systems, Inc. v. Video Solutions Pte. Ltd
Page 000098

CONCLUSION AND DECLARATION

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.

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.

I declare that all statements made herein of my knowledge are true, and that all statements made on information and belief are believed to be true, and that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code.

[SIGNATURE ON FOLLOWING PAGE]

Petitioner's Exhibit 1003
Cisco Systems, Inc. v. Video Solutions Pte. Ltd
Page 000099

Dated: 6/15/2024

By: 

Dr. Immanuel Freedman

Petitioner's Exhibit 1003
Cisco Systems, Inc. v. Video Solutions Pte. Ltd
Page 000100

APPENDIX A – MATERIALS RELIED UPON

Ex-1001	U.S. Patent No. 8,446,823 to Hosseini et al. ("'823 Patent"), Filing Date of June 23, 2009
Ex-1002	Prosecution History of U.S. Patent No. 8,446,823.
Ex-1003	Intentionally Left Blank
Ex-1004	My Curriculum Vitae
Ex-1005	U.S. Patent Publication Serial No. 2009/0141800 to Larson ("Larson"), Patent Application Filing Date of September 3, 2008 and Publication Date of June 4, 2009
Ex-1006	Colin Perkins, "RTP: Audio and Video for the Internet", Addison-Wesley, 2003 ("Perkins").
Ex-1007	Michael Welzl, "Network Congestion Control: Managing Internet Traffic", John Wiley & Sons Ltd., August 2005 ("Welzl").
Ex-1008	N. G. Duffield, et al, "On Adaptive Bandwidth Sharing with Rate Guarantees," 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, San Francisco, CA, USA, 1998, pp. 1122-1130, vol.3) ("Duffield").
Ex-1009	Z. Cai et al., "IQ-Services: Network-Aware Middleware for Interactive Large-Data Applications", 2004, Proceedings of the 2nd Workshop on Middleware for Grid Computing (MGC '04), Association for Computing Machinery, New York, NY, USA, pp. 11–16 ("Cai").
Ex-1010	U.S. Patent No. 8,446,452 to Hosseini et al. ("'452 Patent"), Filing Date of November 7, 2008.
Ex-1011	H. Schulzrinne, et al., "RTP: A Transport Protocol for Real-Time Applications," *IETF RFC 3550*, July 2003, archived at Internet Archive Wayback Machine on August 18, 2003 at https://web.archive.org/web/20030818200653/http://www.ietf.org/rfc/rfc3550.txt.

Petitioner's Exhibit 1003
Cisco Systems, Inc. v. Video Solutions Pte. Ltd
Page 000101

Ex-1012	M. Handley, et al., "TCP Friendly Rate Control (TFRC): Protocol Specification," *IETF RFC 3448*, January 2003, archived at Internet Archive Wayback Machine on April 25, 2003 at https://web.archive.org/web/20030425051207/http://www.ietf.org/rfc/rfc3448.txt.
Ex-1013	Intentionally Left Blank.
Ex-1014	"Transmission Control Protocol," *IETF RFC 793*, J. Postel, ed., September 1981, archived at Internet Archive Wayback Machine on February 2, 2007 at https://web.archive.org/web/20070202201546/https://tools.ietf.org/html/rfc793.
Ex-1015	"User Datagram Protocol," *IETF RFC 768*, J. Postel, ed., August 28, 1980, archived at Internet Archive Wayback Machine on January 27, 2007 at https://web.archive.org/web/20070127071930/http://tools.ietf.org/html/rfc768.
Ex-1016	"Internet Protocol," *IETF RFC 791*, J. Postel, ed., September 1981, archived at Internet Archive Wayback Machine on February 4, 2007 at https://web.archive.org/web/20070204151303/https://tools.ietf.org/html/rfc791
Ex-1017	U.S. Patent Publication Serial No. 2008/0320170 to Yamauchi et al. ("Yamauchi"), Patent Application Filing Date of June 24, 2008, Claimed Priority Date of June 25, 2007 and Publication Date of December 25, 2008.
Ex-1018	International Patent Publication Serial No. WO 2004/047444 to Cooper et al. ("Cooper"), Patent Publication Filing Date of October 27, 2003, Claimed Priority Date of November 15, 2002, and Publication Date of June 3, 2004.
Ex-1019	W. Zhu et al., "JQoS: Design and Implementation of a QoS-based Internet Videoconferencing System Using the Java Media Framework", Proceedings of the IEEE Canadian Conference on Electrical and Computer Engineering (Ontario, Canada, May 13--16 2001), July 2001.
Ex-1020	H. Song et al., "Dynamic Flow Control Using Packet Loss Rate and Round Trip Time", Proceedings of the 4th WSEAS International

Petitioner's Exhibit 1003
Cisco Systems, Inc. v. Video Solutions Pte. Ltd
Page 000102

	Conference on Electronics, Control and Signal Processing, 2005.
Ex-1021	Uyless Black, "Voice Over IP", 2000, Prentice Hall PTR ("Black").
Ex-1022	ITU-T Recommendation H.323, "Packet-Based Multimedia Communications Systems", June 2006 ("H.323").
Ex-1023	U.S. Patent Publication Serial No. 2009/0135906 to Comer et al. ("Comer"), Patent Application Filing Date of July 27, 2006, Claimed Priority Date of July 28, 2005 and Publication Date of May 28, 2009.
Ex-1024	Declaration of Ingrid Hsieh-Yee, Ph.D.
Ex-1025	Intentionally Left Blank
Ex-1026	Intentionally Left Blank
Ex-1027	Intentionally Left Blank
Ex-1028	Intentionally Left Blank
Ex-1029	United States Copyright Office Catalog Record for Colin Perkins, RTP: Audio and Video for the Internet", Addison-Wesley, 2003. (Ex-1006)
Ex-1030	United States Copyright Office Catalog Record for Uyless Black, "Voice Over IP", 2000, Prentice Hall PTR (Ex-1021).

Petitioner's Exhibit 1003
Cisco Systems, Inc. v. Video Solutions Pte. Ltd
Page 000103