UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE PATENT TRIAL AND APPEAL BOARD

GOOGLE LLC and YOUTUBE, LLC

Petitioner

v.

VOXER, INC. and VOXER IP LLC

(record) Patent Owner

IPR2026-00421
Patent No. 11,146,516

DECLARATION OF DR. HENRY HOUH

TABLE OF CONTENTS

TABLE OF CONTENTS 2

I. ENGAGEMENT AND COMPENSATION 4

II. SUMMARY OF OPINIONS 6

III. QUALIFICATIONS 6

IV. MY UNDERSTANDING OF THE RELEVANT LAW 13

A. Claim Construction 13

B. Anticipation 14

C. Obviousness 15

V. TECHNICAL INTRODUCTION 19

A. The ’516 Patent Disclosure 19

B. Overview of Prosecution History 21

VI. RELEVANT TIMEFRAME FOR DETERMINING OBVIOUSNESS 28

VII. LEVEL OF ORDINARY SKILL IN THE ART 28

VIII. CLAIM CONSTRUCTION 29

IX. PRIORITY DATE FOR THE CHALLENGED CLAIMS 30

A. The E2E Limitation 30

B. The Priority Date of the ’516 Patent is no Earlier than August 8, 2016. 32

C. The Priority Applications Filed Prior to August 8, 2016 Lack Support for the E2E Limitation 33

X. DETAILED EXPLANATION OF THE REASONS FOR UNPATENTABILITY 36

Ground 1. Claims 1-18 are Obvious over Katis and Abbazia 36

A. Overview of the Combination 36
1. Overview of Abbazia 37
2. The Combination 39
B. Rationale (Motivation) Supporting Obviousness 41
C. Reasonable Expectation of Success 46
D. Analogous Art 46
E. Graham Factors 47
F. Claim Mapping 47
XI. SECONDARY CONSIDERATIONS 70
XII. OATH 70

1. I, Henry Houh, hereby declare as follows:

I. ENGAGEMENT AND COMPENSATION

2. I have been retained by Google LLC ("Petitioner") as an expert in the *inter partes* review proceeding described above. I have been asked to provide my technical review, analysis, insights, and opinions regarding the references that Petitioner's counsel have provided to me that form the basis for the grounds of unpatentability set forth in the Petition for *inter partes* review of U.S. Patent No. 11,146,516 ("the '516 patent"). For this service, my billing rate is my standard hourly consulting fee and I am reimbursed for actual expenses. My compensation in no way depends on the outcome of this matter.

3. This declaration sets forth my analyses and opinions based on the materials I have considered thus far and the bases for my opinions. My opinions are also based on my education and experience. I understand that this declaration will be used in the above mentioned inter partes review. In forming my opinions, I have considered the following Exhibits:

Exhibit No.	Description
1001	U.S. Patent No. 11,146,516 ("**'516 patent**").
1004	U.S. Patent Appl. Pub. No. US 2009/0003558 A1 ("**Katis**").
1005	U.S. Pat. No. 7,756,033 ("**Abbazia**").
1006	App. Ser. No. 14/799,528 (Filed July 14, 2015) ("'**528 Specification**").

1007	File History of U.S. Application No. 17/062,060.
1008	File History of U.S. Application No. 12/037,256.
1009	File History of U.S. Application No. 15/584,224.
1010	Provisional Application No. 60/937,552.
1011	Provisional Application No. 60/999,619.
1012	List of Applications in Voxer Patent Family.
1013	Proof of Service, D.I. 6, *Voxer, Inc. v. Google, LLC, et al.*, Civ. Action No. 1:25-cv-00873 (D. Del.).
1014	U.S. Pat. No. 11,777,883.
1015	File History of U.S. Application No. 14/799,528.
1016	Redline of ’528 Application as Compared to ’516 Patent.
1017	App. Ser. No. 10/836,284 (Filed May 3, 2004).
1018	App. Ser. No. 12/028,400 (Filed Feb. 8, 2008).
1019	App. Ser. No. 12/037,256 (Filed Feb. 28, 2008).
1020	App. Ser. No. 13/555,034 (Filed July 20, 2012).
1021	App. Ser. No. 14/198,043 (Filed Mar. 5, 2014).
1022	S. Deering, “RFC1112: Host Extensions for IP Multicasting,” *Internet Engineering Task Force (IETF)* (August 1989).
1023	P. Lewis, “Peering Out a ‘Real Time’ Window,” *New York Times* (Feb. 8, 1995).
1024	“IP Multicast Explained,” *Metaswitch Networks* (copyright 2004-2009).
1025	*Facebook, Inc. v. Voxer IP LLC*, IPR2021-00113, Paper 13 (Patent Owner’s Preliminary Response).

4. I note that I rely on a several publications and non-patent literature materials in this declaration. In my opinion, these documents were published on the respective dates noted in each of the publications. In my opinion, the journals and other venues in which these materials were published are reputable and these materials are the type of materials that experts in the industries and fields of study relating to the methods of the challenged patent would reasonably rely on.

II. SUMMARY OF OPINIONS

5. It is my opinion that:

- Despite claiming the benefit of an earlier priority date, the challenged claims of the '516 patent are not entitled to a priority date before August 8, 2016.
- **Ground 1:** Claims 1–18 are obvious over Katis and Abbazia

III. QUALIFICATIONS

6. I hold a Ph.D. in electrical engineering and computer science from the Massachusetts Institute of Technology ("MIT"), awarded in 1998. I also hold a Master's Degree in electrical engineering and computer science from MIT, awarded in 1991, and two Bachelor's degrees (the first in electrical engineering and computer science and the second in physics) also from MIT, awarded in 1989 and 1990, respectively.

7. I have approximately 30 years of work experience related to topics in software development, networking, Internet communications, Web development, mobile device applications, and Internet advertising. I am currently a technical consultant specializing in Social Networking, Web 2.0, Web Site Development, Data Networking, Optical Networking, Telecommunications, Media Streaming and Voice over IP.

8. In graduate school, I joined the Advanced Network Architecture ("ANA") group at the Laboratory for Computer Science ("LCS") at MIT, then headed by Prof. Tennenhouse and senior research scientist David Clark. Dr. Clark is generally considered to be one of the fathers of the Internet Protocol, having acted as chief protocol architect for the development of the Internet. Later, when Prof. Tennenhouse formed his own research group, the Telemedia Network Systems ("TNS") group, I joined the TNS group as a research assistant.

9. As part of my doctoral research at MIT from 1991-1998, I worked as a research assistant in the Telemedia Network Systems (TNS) group at the Laboratory for Computer Science. The TNS group built a high speed gigabit network and applications which ran over the network, such as remote video capture, processing and display on computer terminals. I set up the group's web server around late 1993, which at the time was one of the first several hundred web servers in existence. The graduate students, including myself, were responsible for designing, installing, and

maintaining our group's network and computer infrastructure, including configuring any network equipment used in the group, including bridges and routers. I also performed the same role later in my own startup businesses.

10. The TNS group built a high-speed gigabit network and applications which ran over the network, such as remote video capture, processing, and display on computer terminals. I designed the core networking hardware and software, including the high-speed data links and the device drivers for the network interface cards.

11. TNS's research activities focused on high-speed networking and the types of applications that require increased network bandwidth and computing power. One of the applications I investigated was live streaming video, and real-time processing of such video. My work was focused on the transport, switching, and routing of data (including video and audio streams), and the integration of TNS's network into the computing environment, including the protocols, IP/ATM integration, operating system software and device drivers. TNS created a computing environment where it was possible to stream audio/video live, and also created software processing modules to further process video to perform advanced functions such as green screening, scene change detection, motion detection, compression, video blending and video overlay, among many other features – all in real-time while displaying the processed video live. We also created multiparty video conferencing

applications, and installed a multicast router as part of the MBONE, and used the MBONE to participate in meetings held by the Internet Engineering Task Force ("IETF"), the body that develops standards for the Internet. My Ph.D. thesis work focused on studying the routing resources required to route streams of video through the network, proposing several alternative labeling strategies at ingress points of a network to speed up network routing, especially for popular applications such as web traffic and longer lived streams such as audio and video. These labeling strategies I analyzed for my thesis is the same type of strategy used in MPLS today.

12. During the early part of my graduate studies, a time when there were perhaps a hundred or so web servers in existence, I set up a web server on one of our lab computers, and created content rich web pages for my research group and my research activities. Eventually, I, together with others I was working with, added live video demonstrations to TNS's web site. TNS's web site was one of the first several hundred web servers to exist, and the first to offer live video demonstrations initiated from the web site to a user's remote computer. I helped to build the web pages that initiated these video sessions via a web interface. I started and maintained the "Commercial Sites Index," the first listing of all the companies on the web. I later sold this list to OpenMarket, then a leader in web servers and eCommerce servers. I also started a company to build web sites and sell targeted online advertisements.

13. Former Vice-President Al Gore visited our group in 1996 and received a demonstration of – and remotely drove – a radio-controlled toy car with a wireless video camera mounted on it; the video was encoded by TNS-designed hardware, streamed over the TNS-designed network and displayed using TNS-designed software.

14. I authored or co-authored at least twelve papers and conference presentations on my group's research at TNS. I also co-edited TNS's final report on its gigabit networking research effort. Some of my papers about web site development were presented at the first World-Wide Web conference held in 1994 at CERN in Geneva, Switzerland. CERN is the birthplace of the Web, and the conference was chaired by Tim Berners-Lee, the inventor of the World-Wide Web. When Tim Berners-Lee formed the World Wide Web Consortium at the MIT LCS in order to manage the progression of web protocols and standards, I assisted him in obtaining the consortium's domain names and setting up the consortium's web server.

15. From 1997 to 1999, I was a Senior Scientist and Engineer at NBX Corporation, a start-up that made business telephone systems for streaming packetized audio over data networks instead of using traditional telephone lines. NBX was later acquired by 3Com Corporation, and the phone system I worked on at NBX was widely used by numerous businesses. As part of my work at NBX, I

designed the core audio reconstruction algorithms for the telephones, as well as the packet transmission algorithms. Our system also utilized multicast to deliver music-on-hold content to all the phones as well as for audio paging, which could activate the phones with a live audio message. I also designed and validated the core packet transport protocol used by the phone system. I also designed the method of utilizing network priority methods in our phone system. I also configured routers in our development and testing infrastructure, and set up tunnels through the Internet that we used to connect our network-based phones at trade shows around the country to our headquarters network.

16. From 2003-2004, I also was the architect for a next-generation web testing platform developed by Empirix, known as e-Test Suite. e-Test Suite is now owned by Oracle Corporation. e-Test provided functional and load testing for web sites. e-Test emulated a user's interaction with a web site and provided web developers with a method of creating various scripts and providing both functional testing (*i.e.*, did the web site provide the correct response) and load testing (*i.e.*, could the web site handle 5000 users simultaneously). As part of this project, I co-wrote a javascript interpreter for e-Test so that it could understand the javascript code present on advanced web sites.

17. I worked at BBN Technologies from 2004 to 2009. BBN Technologies, formerly known as Bolt, Beranek and Newman, was a pioneer of the Internet and a

long-established leader in speech recognition and natural language processing. BBN received the first contract to build components of a packet switched network as part of the ARPANET, a U.S. Department of Defense effort funded through the Advanced Projects Research Agency ("ARPA"). While this effort preceded my time at BBN by many years, BBN continued to maintain a strong presence in network research and development, including in wireless networking, device discovery, and protocols for a "body area network."

18. In 2006, as part of my role at BBN Technologies, I helped found PodZinger Inc., now known as RAMP Inc. PodZinger utilized BBN's speech recognition algorithms to search through the spoken words inside audio and video. After managing the creation of the initial prototype system, PodZinger built out a full web-based streaming audio and video search solution when I was the Vice President of Operations and Technology there. The web site had a full load-balanced back end that provided search and indexing solutions for our hosted multimedia (audio and video) content. Our PodZinger system processed videos in order to perform automatic speech recognition on the spoken contents of the video, after which we produced a word transcript we used in a search index for all our processed videos. All of the words in the transcript were individually tagged to frame locations within the video, so that selecting a word in the transcript caused the video to jump to the location within the video at which the word was spoken.

19. I also worked on commercializing a service offered by BBN to monitor call centers and provide analytical data on each call. The business became known as AVOKE Analytics and provided call center analytics for large organizations such as Dell and the Centers for Medicare and Medicaid Services ("CMMS").

20. I am familiar with patents and have numerous provisional and non-provisional patent applications on which I am an inventor. These patent applications deal with various areas of technology including, but not limited to: utilizing network processors, network and packet delivery systems, tagging media content, dynamic presentation of advertising and other content using enhanced metadata, methods for web site testing, and organizing and managing files that are transmitted over a network.

21. I have worked on numerous legal cases that involved video streaming, including those that involve commercial television content providers that use multicast to deliver video content to numerous subscriber end points efficiently.

22. My qualifications and publications are set forth more fully in my curriculum vitae, attached as Ex-1003.

IV. MY UNDERSTANDING OF THE RELEVANT LAW

A. Claim Construction

23. I understand that a claim term in an *inter partes* review is to be interpreted according to standards applied by district courts. First, the language of

the claims themselves is of primary importance in the effort to determine precisely what it is that is patented. The terms used in a claim are generally given the ordinary and customary meaning that the terms would have to a person of ordinary skill in the art in question at the time of the alleged invention, unless the term is expressly defined in the patent. The person of ordinary skill in the art reads the claim in the context of the entire patent, including the specification. Next to the language of the claims, the specification is the single best source for interpreting the claim terms. The claims may also be interpreted using the record of correspondence between the patent applicant and the Patent Office, and also with reference to other sources of evidence that can help to define the meaning of terms to a person of ordinary skill in the relevant time frame. I define the relevant time frame in §VI, below.

B. Anticipation

24. I understand that a claim in an issued patent can be invalid if it is anticipated. In this case, "anticipation" means that there is a single prior art reference that discloses every element of the claim, arranged in the way required by the claim.

25. I understand that an anticipating prior art reference must disclose each of the claim elements expressly or inherently. I understand that "inherent" disclosure means that the claim element, although not expressly described by the prior art reference, must necessarily be present based on the disclosure. I understand

that a mere probability that the element is present is not sufficient to qualify as "inherent disclosure".

C. Obviousness

26. I understand that a claim in an issued patent can be invalid if it is obvious. Unlike anticipation, obviousness does not require that every element of the claim be in a single prior art reference. Instead, it is possible for claim elements to be described in different prior art references, so long as there is motivation or sufficient reasoning to combine the references.

27. I understand that a claim is invalid for obviousness if the differences between the claimed subject matter and the prior art are such that the subject matter as a whole would have been obvious before the effective filing date of the claimed invention to a person having ordinary skill in the art to which said subject matter pertains.

28. I understand, therefore, that when evaluating obviousness, one must consider obviousness of the claim "as a whole". This consideration must be from the perspective of the person of ordinary skill in the relevant art, and that such perspective must be considered as of the effective filing date of the claimed invention.

29. The level of ordinary skill in the art is discussed in §VII, below.

30. The relevant time frame for determining obviousness is discussed in §VI, below.

31. I understand that in considering the obviousness of a claim, one must consider four things. These include the scope and content of the prior art, the level of ordinary skill in the art at the relevant time, the differences between the prior art and the claim, and any "secondary considerations". I understand that these four things were specified by the U.S. Supreme Court in a case called *Graham v. John Deere*, and are therefore called the "*Graham* factors".

32. I understand that "secondary considerations" include real-world evidence that can tend to make a conclusion of obviousness either more probable or less probable. For example, the commercial success of a product embodying a claim of the patent could provide evidence tending to show that the claimed invention is not obvious. In order to understand the strength of the evidence, one would want to know whether the commercial success is traceable to a certain aspect of the claim not disclosed in a single prior art reference (i.e., whether there is a causal "nexus" to the claim language). One would also want to know how the market reacted to disclosure of the invention, and whether commercial success might be traceable to things other than innovation, for example the market power of the seller, an advertising campaign, or the existence of a complex system having many features beyond the claims that might be desirable to a consumer. One would also want to

know how the product compared to similar products not embodying the claim. I understand that commercial success evidence should be reasonably commensurate with the scope of the claim, but that it is not necessary for a commercial product to embody the full scope of the claim.

33. Other kinds of secondary considerations are possible. For example, evidence that the relevant field had a long-established, unsolved problem or need that was later provided by the claimed invention could be indicative of non-obviousness. Evidence that others had tried, but failed to make an aspect of the claim might indicate that the art lacked the requisite skill to do so. Evidence of copying of the patent owner's products before the patent was published might also indicate that its approach to solving a particular problem was not obvious. Evidence that the art recognized the value of products embodying a claim, for example, by praising the named inventors' work, might tend to show that the claim was non-obvious.

34. I further understand that prior art references can be combined where there is an express or implied rationale to do so. Such a rationale might include an expected advantage to be obtained, or might be implied under the circumstances. For example, a claim is likely obvious if design needs or market pressures existing in the prior art make it natural for one or more known components to be combined, where each component continues to function in the expected manner when combined (i.e.,

when there are no unpredictable results). A claim is also likely invalid where it is the combination of a known base system with a known technique that can be applied to the base system without an unpredictable result. In these cases, the combination must be within the capabilities of a person of ordinary skill in the art.

35. I understand that when considering obviousness, one must not refer to teachings in the specification of the patent itself. One can, however, refer to portions of the specification admitted to be prior art, including the "Background" section. Furthermore, a lack of discussion in the patent specification concerning how to implement a disclosed technique can support an inference that the ability to implement the technique was within the ordinary skill in the prior art.

D. Written Description and Priority

36. I understand that a patent claim needs to be described in the application for a patent that was filed with the United States Patent and Trademark Office. I understand to "describe" a claim, the patent application needs to provide sufficient written description to reasonably convey to those skilled in the art that the inventor had 'possession' of the claimed subject matter as of the filing date. This involves an objective inquiry into the four corners of the specification from the perspective of a person of ordinary skill in the art. Based on that inquiry, the specification must describe an invention understandable to that skilled artisan and show that the inventor actually invented the invention claimed. I further understand that while a

patent specification can support a claim even if it does not need to use exactly the same words as the claim, this does not extend into the realm of "obviousness". That is, a patent specification that *merely* renders a claim obvious does not provide written description support for that claim.

37. I understand that for a negative claim limitation, the written description requirement cannot be met through simple disregard of the presence or absence of a limitation. I understand that the written description requirement *may* be satisfied if a PHOSITA would understand the specification to: (1) describe a reason to exclude the relevant element, e.g. by expressly listing the disadvantages of using that element or distinguishing among the element and alternatives to it, or (2) inherently disclose the negative limitation such that it would always be understood by skilled artisans as being necessarily excluded.

38. I understand that a patent application is entitled to the benefit of the filing date of an earlier filed application to which it claims priority only if the disclosure of the earlier application provides support for the claims of the later application. I understand that originally-filed claims may be relied upon for support.

V. TECHNICAL INTRODUCTION

A. The '516 Patent Disclosure

39. I have reviewed the '516 patent. The '516 patent describes a "telecommunication and multimedia management method and apparatus" that

allows users to engage in conversations by exchanging messages that include voice and/or other media. EX1001, 1:38-45, 4:21-33. Specifically, "[u]sers are empowered to conduct communications in either: (i) a near-synchronous or 'live' conversation, providing a user experience similar to a standard full duplex phone call; or (ii) in a series of back and forth time-delayed transmissions (i.e., time-shifted mode)." *E.g.*, EX1001, 4:34-38. "[U]sers engaged in a conversation can seamlessly transition from the live mode to the time-shifted mode and back again." *E.g.*, EX1001, 4:38-40. According to the '516 patent, "[t]his [is a] new form of communication" called "Voxing." EX1001, 4:48-50.

40. The '516 patent provides details of a Voxing system. Of note, a "Device" is the "physical device that runs the Client application," where "Client" refers to "the user application in the communication system, which includes a user interface, persistent data storage, and 'Voxing' functionality." EX1001, 7:65, 7:54-56. "Devices 13 communicate with one another over a communication services network 14, including one or more Servers 16." EX1001, 11:63-67. "Servers are responsible for [A] routing Messages sent back and forth between Users over the network and [B] the persistent storage and archiving of Media payloads." EX1001, 8:6-11; *see also id.* at 22:53-58. Client 12 includes store and stream module 24, which "persistently stores all Media … originating using a Device 13 or received over the network 18 from other Device 13 [sic] and/or users." EX1001, 18:30-33.

Application 78 running on server 16 includes store and stream module 84, which also stores received messages and transmits the message to the next "hop" along the path to the intended recipient. EX1001, 21:60-22:4, 22:52-62. By virtue of these store and stream modules, a user can leave a message for another party even when the sender and/or recipient is unavailable: "the Message [will be] queued for transmission as soon as network connectivity becomes available, resulting in a time-shifted delivery." *E.g.*, EX1001, 18:33-42.

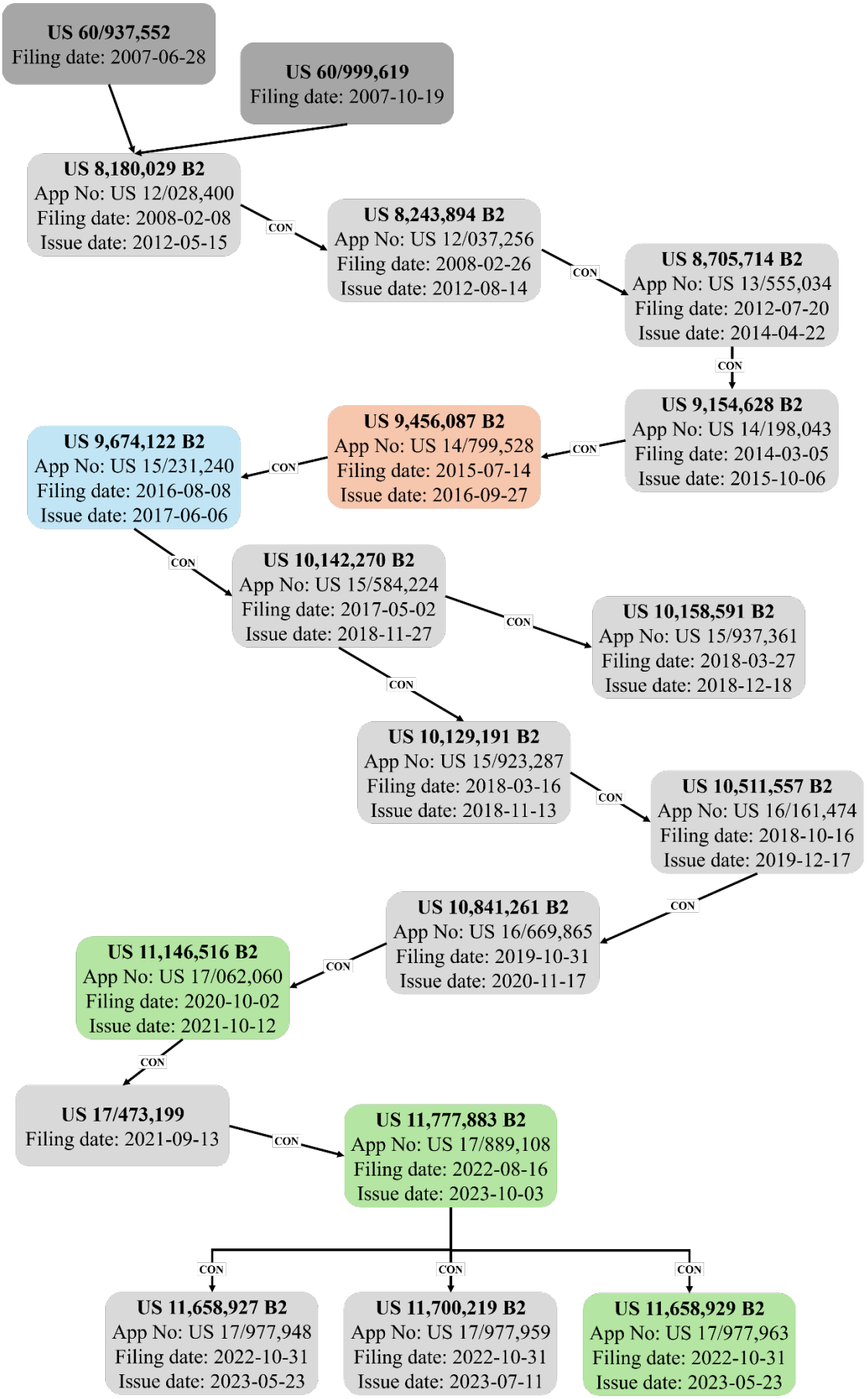
41. In operation, Client 12_1 creates media, stores the media payloads, and "transmits the packets to the receiving Client 12_2 over the network 18." EX1001, 27:35-42. End-to-end encryption is provided between clients 12. EX1001, 12:35-47. "Each Server 16 along the path between the sending Client 12_1 and the receiving Client 12_2 stores the transmitted payloads … and transmits the Vox packets to the next hop." EX1001, 27:42-45. After receiving the message, Client 12_2 stores the payloads, renders the payload information, and sends an acknowledgement to the sender. EX1001, 27:45-54, 21:25-28.

B. Overview of Prosecution History

42. I understand that the '516 patent is a relatively recent patent in an extensive family which claims ultimate priority to two provisional applications filed on June 28, 2007 and October 19, 2007. EX1001, code (63). A graphic illustrating

the immediate patent family to which the ’516 patent belongs is provided below with the ’516 patent highlighted in green.[1]

[1] The other two patents for which Petitioner is concurrently requesting *inter partes* review are also highlighted in green.


US 60/937,552
Filing date: 2007-06-28
US 60/999,619
Filing date: 2007-10-19
US 8,180,029 B2
App No: US 12/028,400
Filing date: 2008-02-08
Issue date: 2012-05-15
CON
US 8,243,894 B2
App No: US 12/037,256
Filing date: 2008-02-26
Issue date: 2012-08-14
CON
US 8,705,714 B2
App No: US 13/555,034
Filing date: 2012-07-20
Issue date: 2014-04-22
CON
US 9,154,628 B2
App No: US 14/198,043
Filing date: 2014-03-05
Issue date: 2015-10-06
CON
US 9,456,087 B2
App No: US 14/799,528
Filing date: 2015-07-14
Issue date: 2016-09-27
CON
US 9,674,122 B2
App No: US 15/231,240
Filing date: 2016-08-08
Issue date: 2017-06-06
CON
US 10,142,270 B2
App No: US 15/584,224
Filing date: 2017-05-02
Issue date: 2018-11-27
CON
US 10,158,591 B2
App No: US 15/937,361
Filing date: 2018-03-27
Issue date: 2018-12-18
CON
US 10,129,191 B2
App No: US 15/923,287
Filing date: 2018-03-16
Issue date: 2018-11-13
CON
US 10,511,557 B2
App No: US 16/161,474
Filing date: 2018-10-16
Issue date: 2019-12-17
CON
US 10,841,261 B2
App No: US 16/669,865
Filing date: 2019-10-31
Issue date: 2020-11-17
CON
US 11,146,516 B2
App No: US 17/062,060
Filing date: 2020-10-02
Issue date: 2021-10-12
CON
US 17/473,199
Filing date: 2021-09-13
CON
US 11,777,883 B2
App No: US 17/889,108
Filing date: 2022-08-16
Issue date: 2023-10-03
CON
CON
CON
US 11,658,927 B2
App No: US 17/977,948
Filing date: 2022-10-31
Issue date: 2023-05-23
US 11,700,219 B2
App No: US 17/977,959
Filing date: 2022-10-31
Issue date: 2023-07-11
US 11,658,929 B2
App No: US 17/977,963
Filing date: 2022-10-31
Issue date: 2023-05-23

43. I am informed that the complete family, however, includes approximately 101 U.S. non-provisional applications. EX1012.

44. I understand that over time, the claims Voxer sought to obtain increasingly diverged from what it described in its applications. For example, in the '256 application, filed in 2008, I understand that Voxer originally described its invention as:

> "code, a method and communication device configured to progressively store streaming media created using the communication device, as the streaming media is created, so that a copy of the streaming media is stored on the communication device. The media is also progressively transmitted, during a transmission session, as the streaming media is created and stored, using a transmission protocol that (i) supports near-real-time communication so that the streaming media is delivered in near real-time to a second communication device associated with a recipient and (ii) guarantees delivery of the streaming media to the second communication device."

EX1008, p.620.

45. I further understand that in contrast, by the later '224 application, filed in 2017, Voxer distinguished its claims from the Enete reference, arguing that "[w]ith Enete, all video communication is synchronous, meaning an end-to-end network connection or session between the parties must first be established before

any video communication may begin." EX1009, p.224. Instead, according to Voxer, "[t]he claims of the present invention are directed to video communication where (1) a recipient can review portions of the video **while the video is transmitted by the sender** and (2) **without first establishing an end-to-end connection** between sending and receiving devices." EX1009, p.225 (emphasis in original). Voxer thus went from describing its alleged invention as streaming media "during a transmission session" to distinguishing it from a system allegedly employing an "end-to-end network connection or session" and presenting claims which affirmatively require streaming media "without first establishing an end-to-end connection between the sending and receiving devices."

46. As another example, I understand that the '552 provisional application providing the foundation for all of Voxer's later applications explicitly stated that Voxer had not yet developed a system which could transmit video: "The only supported Media type for this prototype will be voice. Future versions will support text, video, position, and raw data Media." EX1010, p.14. In my opinion, the specification of the '516 patent itself is also clear on this point: "Although as of the filing of the present application, most conversations are voice only, it is intended that the communication system and method described herein broadly includes conversations including multiple media types, such as voice and video for example." EX1001, 5:34-38. Voxer nonetheless expanded its claims over time to focus on

"video media" as required by all claims of the '516 patent.

47. I further understand that, for example, in the '256 application, Voxer originally presented 22 claims, then added 21 more, then canceled those 43 claims, and added 63 new claims, (EX1008, p.75, p.166, p.344), then after a Non-Final Office Action, canceled all 106 previously-presented claims, and added 81 new claims (*id.*, p.608, p.620). In addition, I understand that Voxer not only pursued a voluminous number of patents over the years (*see* EX1012), but it would present large numbers of claims. This can be seen in the related '883 patent, concurrently challenged by Petitioner, which contains 65 claims. EX1014. I further understand that Voxer would also take actions such as filing applications like the '528 application, with one broad claim, (EX1015, p.124), receive a Non-Final Rejection, and then amend the application to include numerous new claims. *Id.*, pp.184-188.

48. I am informed that during the prosecution of the '060 application leading to the '516 patent filed in 2020, the Office rejected the pending claims for double patenting over Voxer's other patents. EX1007, pp.165-167. The Office also rejected the claims in view of the Mangal[2] principal prior art as well as the Teng,

[2] US 2003/0148785A1 ("Mangal"). During prosecution, the Examiner also cited the US 2003/0208543A1 ("Enete"), US 2008/0062246Al ("Woodworth"), and US 2007/0207782A l) ("Tran") references as pertinent. *Id.* at pp.247-249, 175-178.

Kalaboukis, Yu, Honda, and Mattila secondary references. *Id.* at pp.167-179.

49. I understand that in response, Voxer distinguished the Mangal reference, arguing, among other things that:

> "[w]hile the applicant acknowledges that Mangal contemplates the sending of video media during a PTT transmission, there is absolutely no teaching that such video media is (a) associated with a messaging structure or (b) is routed over the network of Mangal using messaging-like protocols. On the contrary, all video media that is transmitted in the PTT system of Mangal is done so over a predefined, end-to-end, virtual, synchronous connection for a given communication group."

Id. at pp.232, 224-234. I understand that Voxer explicitly argued that Mangal did not teach "the video service infrastructure operates so that a recipient can render at least a portion of the video media of the video message . . . . **without having to first establish an end-to-end connection over the network between the sending device and the recipient device before the sending device begins to transmit the video media of the video message**." *Id.* at p.233.

50. I understand that in response, the Examiner closed prosecution on the merits, and Voxer submitted terminal disclaimers. *Id.* at pp.245-47. Reasons for allowance were not set forth on the record. *Id.* at p.269

VI. RELEVANT TIMEFRAME FOR DETERMINING OBVIOUSNESS

51. I understand that obviousness must be evaluated at the time before the effective filing date of the claimed invention. From the first page of the '516 patent, I understand that the first patent applications provisional applications) were filed on June 28, 2007 and October 19, 2007. EX1001, codes (60), (63)). However, as I explain below in §IX of this declaration, in my opinion, the earliest possible priority date for the challenged claims is August 8, 2016 (i.e., filing of App. No. 15/231,240) because each claim recites a negative limitation ("the E2E Limitation") that is not found in any of the earlier-filed applications. I will assume for the purpose of this declaration that August 8, 2016, is the effective filing date of the claimed invention, with the understanding that my testimony will also be applicable to some time prior to that date. I may refer to this time period as the "relevant time frame," and my testimony concerning obviousness is directed to this timeframe, even if I occasionally do not explicitly use a past tense.

VII. LEVEL OF ORDINARY SKILL IN THE ART

52. I understand that the legal question of obviousness must be assessed from the perspective of a person having ordinary skill in the art ("PHOSITA") at the time of the invention. In my opinion, a person having ordinary skill in the art ("PHOSITA") would have had at least (1) a Bachelor's degree in electrical or computer engineering or a related course of study with at least two years of work

experience in industry in the areas of computer networking, multimedia processing, or a related area of work or (2) a Master's degree in electrical or computer engineering or a related course of study, with at least one year of experience in industry in the areas of computer networking, multimedia processing, or a related area of work. A higher level of education could substitute for less industry experience, and more industry experience could substitute for the specific level of education. My testimony concerning obviousness is offered from this perspective even if I do not explicitly state so.

53. In my opinion, I would have met this standard as of the relevant time August 8, 2016 (i.e., filing of App. No. 15/231,240) and continue to do so today. To the extent I exceed the minimum qualifications of a PHOSITA during the relevant time period, my opinions in this declaration are presented from the perspective of a PHOSITA.

VIII. CLAIM CONSTRUCTION

54. In my opinion, the specification and prosecution history point toward constructions for a large number of terms, some of which are found in the claims. *See, e.g.*, EX1001, 7:46-11:57 ("Glossary"). I have applied an understanding of the claims consistent with the specification and prosecution history in evaluating the claims.

55. I note that the primary reference (Katis) is a publication from the nominal priority chain of the '516 patent and is thus substantially similar to the specification of the '516 patent. I have interpreted the claims consistently in evaluating written description and priority as well as in evaluating the applied prior art.

IX. PRIORITY DATE FOR THE CHALLENGED CLAIMS

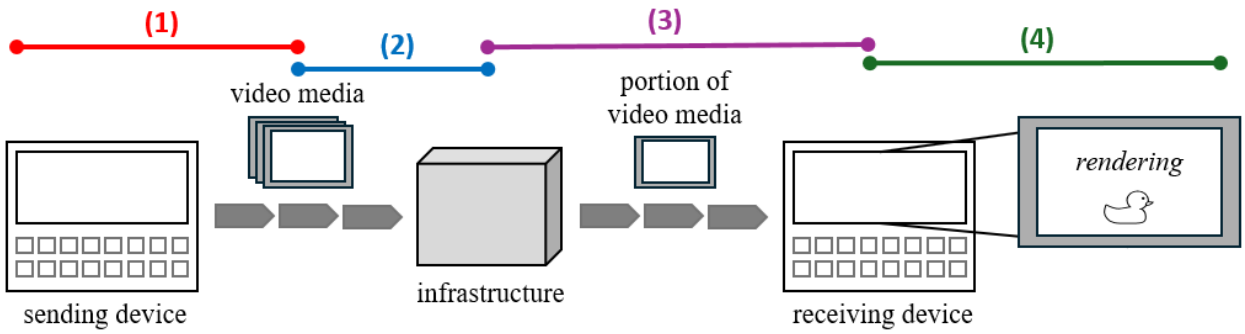
56. As noted above, the '516 patent includes a lengthy priority claim (*see* EX1001, codes (60), (63)); however, in my opinion, the earliest possible priority date for the challenged claims is August 8, 2016 (i.e., filing of App. No. 15/231,240) because each claim recites a negative limitation ("the E2E Limitation") that is not found in any of the earlier-filed applications.

A. The E2E Limitation

57. Claim 1 of the '516 patent recites "receiving … a video message transmitted by a sending device" at an "infrastructure" and "transmitting the video media of the video message from the … infrastructure to a recipient device so that a recipient can render at least a portion of the video media … on the recipient device, … so that the rendering occur[s]: (i) while the video media … is being … transmitted over the network by the sending device; and (ii) without having to first establish an end-to-end connection … between the sending device and the recipient device."

EX1001, 44:35-36, 44:39-52. Together, these limitations are referred to herein as the "**E2E Limitation**."

58. This dense claim language can be summarized with a few architectural requirements and some temporal ones. Architecturally, (1) a sending device transmits video media of a video message; (2) an infrastructure receives that video media; (3) the infrastructure transmits the video media, including portions of it, to a receiving ("recipient") device; and (4) the receiving device is able to render those portions. These requirements are illustrated in the graphic below.



59. As for the temporal requirements, the claim language specifies that the receiving device must be able to render the portions of video media ***while*** the sending device is transmitting the video media—i.e., #4 must occur while #1 is occurring. In addition, the recited "portions of video media" and the "video media" are both "of the video message" (i.e., #4 and #1 involve the same message), and as a result, the sending and receiving devices must concurrently transfer data for the message to/from the infrastructure—i.e., #2 and #3 must also be concurrent. This means that

the sending and receiving devices must each be connected to and in communication with the infrastructure (i.e., the sending and receiving devices must both be online) at the time. In addition, all of this must occur ***without*** first establishing an end-to-end connection between the sending and receiving devices.

60. As a result, the E2E Limitation requires (1) the receiving device to be able to render portions of video media ***while*** the sending device is transmitting the video media (which requires both the sending and receiving devices to be online) and (2) that this occur ***without*** having first established an end-to-end connection between the sending and receiving devices.

B. The Priority Date of the '516 Patent is no Earlier than August 8, 2016.

61. All challenged claims of the '516 patent effectively recite the E2E Limitation. Arguable support for the E2E Limitation, however, was not present in applications in the priority chain of the '516 patent until the filing of U.S. App. Ser. No. 15/231,240, on August 8, 2016. As a result, in my opinion, all claims of the

'516 patent are entitled to a priority date of no earlier than August 8, 2016, and both Katis and Abbazia qualify as prior art.[3]

C. The Priority Applications Filed Prior to August 8, 2016 Lack Support for the E2E Limitation

62. Before it was added into claim language of the application filed August 8, 2016 (the '240 application), the E2E Limitation was not present in the applications that allegedly support the '516 patent's priority chain. The application immediately prior to the '240 application was App. Ser. No. 14/799,528, filed July 14, 2015. EX1006. I understand that this '528 application is representative of prior applications (which do not significantly differ in their disclosure of the E2E Limitation), and will be referred to as "**the '528 Specification**" in the following analysis. EX1006 (the '528 application); EX1018-1021 (prior non-provisional applications); EX1016 (Redline of '528 application as compared to '516 patent). I also reviewed the provisional applications and see no disclosure of the E2E Limitation. EX1010; EX1011.

[3] I understand that even if the claims were entitled to a slightly earlier date, Katis and Abbazia would still both qualify as prior art so long as the priority date is after July 13, 2010.

63. As explained above, the E2E Limitation requires (1) the receiving device to be able to render portions of video media while the sending device is transmitting the video media (which requires the sending and receiving devices to both be online), and (2) that this occur without having first established an end-to-end connection between the sending and receiving devices. In my opinion, this limitation is not disclosed in the '528 Specification.

64. First, the '528 Specification does not expressly describe the E2E Limitation. There is simply no express discussion of the need/ability to avoid or omit an end-to-end connection before exchanging video media between online devices, as required by the E2E Limitation.

65. Second, a PHOSITA would not infer, from the Specification, any reason to avoid, or advantage in avoiding, end-to-end connections before exchanging video media between online devices. In fact, where the '528 Specification discusses interaction between the first and second devices, it implies that an end-to-end connection is useful. For example, the '528 Specification states that client 12 in device 13 includes modules 40 and 44, which facilitate "multi-party ***end-to-end authentication, encryption and security*** … ***between Clients 12***." EX1006, ¶82 (emphasis added). In other words, the '528 Specification teaches that *establishing* an end-to-end connection between online devices provides an *advantage*, i.e., end-to-end encryption. EX1006, ¶82. Also, the '528 Specification elsewhere implies

other advantages to such a connection, such as packet acknowledgement (EX1006, ¶120) and sharing of presence data to optimize network utilization (EX1006, ¶20).

66. In my opinion, Voxer may contend that the '528 Specification also includes *disadvantages* of an end-to-end connection in paragraphs 3, 6, 11, 19, and 39(iv), but such an argument would be misplaced. Each such paragraph is directed to the situation where one of the devices is ***unavailable*** (i.e., the devices are not both online). Indeed, the purpose of these passages is to explain that, when a recipient is unavailable, a sender can still transmit a message. Moreover, the solution described is ***not*** related to an end-to-end connection (or the lack thereof), but rather storing messages in a number of locations so that the recipient can review the messages in a time-shifted mode after the recipient becomes available. *See, e.g.*, EX1006, ¶19, ¶39(iv), ¶110. This is shown, for example, in paragraph 39(iv), which states:

> iv. enabling users to participate in conversations without waiting for a connection to be established with another participant or the network. ***This attribute allows users to begin conversations, participate in conversations, and review previously received time-shifted messages of conversations even when there is no network available, when the network is of poor quality, or other participants are unavailable***.

EX1006, ¶39(iv)(emphasis added).

67. Because the '528 Specification does not describe the E2E Limitation, nor provide disclosure that would allow a PHOSITA to conclude that the E2E

Limitation would be advantageous, in my opinion, it does not provide sufficient support for the E2E Limitation to support a priority claim earlier than the filing of the ’240 Application on August 8, 2016.

X. DETAILED EXPLANATION OF THE REASONS FOR UNPATENTABILITY

Ground 1. Claims 1-18 are Obvious over Katis and Abbazia

68. It is my opinion that claims 1-18 would have been obvious over Katis (EX1004) and Abbazia (EX1005).

69. I understand that both Katis and Abbazia are prior art when the correct priority date of no earlier than August 8, 2016 is used.

70. Katis is a U.S. Patent application publication, published on January 1, 2009, and Abbazia is a U.S. Patent issued on July 13, 2010.

71. I understand that Abbazia was not cited during the prosecution of any application leading to the ’516 patent, and Katis, as an application in the priority chain, was not considered to be prior art.

A. Overview of the Combination

72. In my opinion, Katis in view of Abbazia renders the challenged claims obvious.

73. **Katis** teaches the same system of the claims of the ’516 patent, except that it does not disclose the E2E Limitation, as described above.

74. **Abbazia** discloses the E2E Limitation in the context of a Multicast network for efficient distribution of video messages. I note that a certificate of correction issued August 6, 2013 replaced Abbazia's drawings (EX1005, pp.20-30), but Abbazia's corrected figures 6 and 7 were part of the original application filing on May 3, 2004 (EX1017).

1. Overview of Abbazia

75. Specifically, Abbazia discloses a system such as the one shown in Fig. 1, reproduced here:

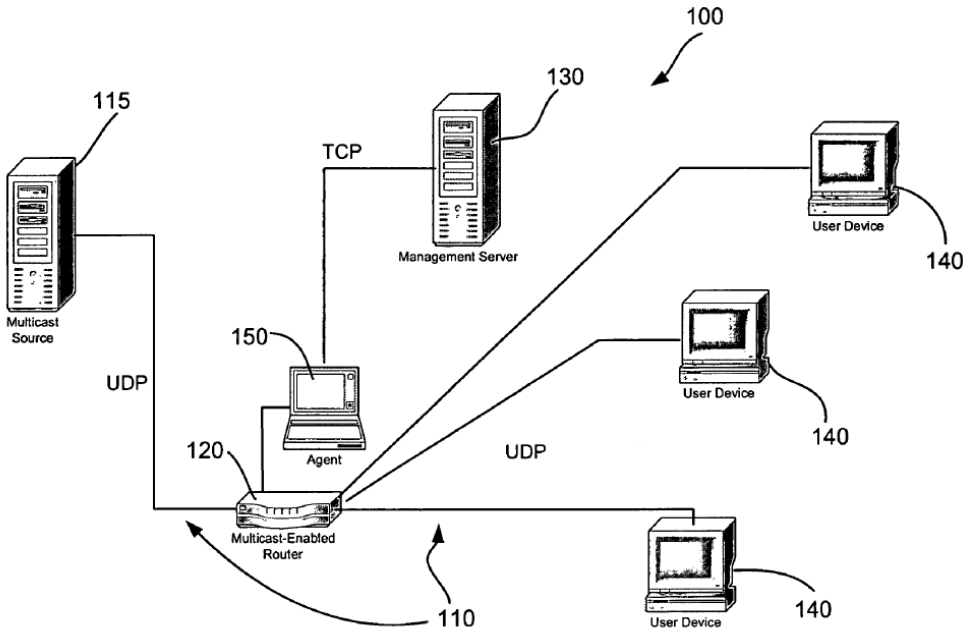


FIG. 1

EX1005, Fig. 1. In Fig. 1, there is a Multicast feed source 115 on the left, which can be any computer or server. EX1005, 3:64-67, 6:7-8. Streaming media (e.g. video) packets from the feed source 115 are sent to a Multicast-enabled router 120, which

in turn sends the packets to user devices 140. EX1005, 12:6-8 (streaming media), 1:17-20 (video), 6:6-13 (packets sent to 120), 4:17-20 (forwarding packets by 120).

76. "Multicast" is a well-known Internet standard for efficient communications, and was well known as of the relevant time. *See* EX1022-1024; *infra* § X.B (discussing the rationale supporting obviousness). In Multicast, a Multicast sender sends a packet to a "Multicast address", which is an IP address reserved for a particular Multicast group. *See, e.g.*, EX1022, p.1 (explaining that, in IP multicasting, packets are delivered to "a set of … hosts identified by a single IP destination address"). When a packet addressed to a Multicast address is received at a Multicast-enabled router, the router forwards the packet to devices on its local network that have opted to receive packets addressed to the Multicast address. *See, e.g.*, EX1022, p.1 (describing function of multicast routers).

77. The Multicast method produces a significant gain in efficiency, because it allows part of the path to be traversed by a single packet. For example, when transmitting a packet from source 115 to devices 140 in Fig. 1 of Abbazia, only one packet (addressed to a Multicast IP address) need be sent from the source 115 to the router 120. In contrast, a non-Multicast (*e.g.* "unicast") transmission mode would send ***three*** packets along this path, each addressed to one of the three user devices 140. Multicast thus reduces traffic by a factor approximately equal to the number of recipients along certain parts of the path to a recipient.

78. To Multicast, Abbazia adds a management server 130. EX1005, Fig. 1, Abstract, 3:64-4:9. The management server listens to Multicast links and collects data, enabling users to monitor statistics and related information concerning the Multicast, and to generate reports in substantially real time. EX1005, 5:13-15; 7:45-8:24.

2. The Combination

79. In my opinion, it would have been obvious to include a Multicast option within the Katis system. Specifically, it would have been obvious to a PHOSITA to allow a Katis server to transmit to a Multicast address for real-time video. In that sense, a Katis server 16 would be a Multicast feed source 115 as taught by Abbazia, and shown below in a modified version of Fig. 1 of Katis (left side) connected to modified Fig. 1 of Abbazia (right side), together indicating a direction of transmission from Katis client 12_1 of device 13_1 (left side) to other users 140 (right side):

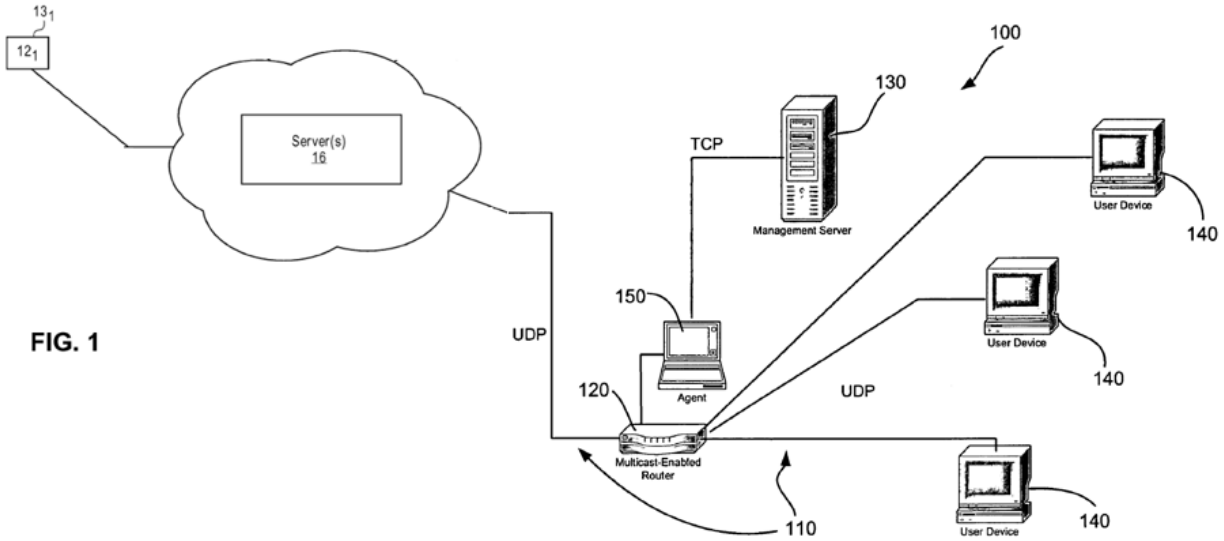


FIG. 1

FIG. 1

In the modified figure, user devices 140 would be user devices of the Voxer system from Katis. The client 12_1 would send real-time video packets to the Voxer server 16, which forwards them to a Multicast address.

80. In the combination, the Katis user would initiate a conversation as usual, but user devices and/or the server would have the option (implemented by simple programming changes at the Katis server 16) to use Multicast. Katis explains that "[a] User creates a Conversation … by inputting a Conversation name [and] a list of Contacts, and an optional scheduled start time," which may be "immediate." EX1004, ¶0257. If Multicast is elected, a Multicast address would be acquired and sent by the server 16 with a notification of an upcoming (or impending) real-time video message to Voxer users, allowing any user in the conversation to receive the real-time video via Multicast and benefit from the improved efficiency. Voxer

packets would be sent to the server(s) 16 as usual, but would be forwarded to the Multicast address for Voxer users in the conversation who elect to receive packets via Multicast (or for whom that election is made). The Multicast address would simply provide an alternate, more efficient, pathway for real-time video, from which any Voxer user in the conversation could elect to receive. Voxer users receiving real-time video from a Multicast address can send reply packets back to the Katis servers directly, with no change. All users would send and receive Vox packets in their usual form, with no change, including the conversation ID, sequence ID, and other data as taught by Katis. *See* EX1004, ¶¶0063, 0065, 0091, 0180-0184, 0191, 0257, 0267. With the combination, a server and/or user device may decide to use Multicast to transmit to a set of recipients in appropriate circumstances (e.g., a large number of recipients, common network infrastructure used for a set of recipients, and/or limited network bandwidth).

81. In my opinion, a PHOSITA would also have found it obvious to use Abbazia's monitoring system to "address the need for real-time management of the multicast feeds from a monitoring perspective, while maintaining low overhead and resource requirements" if desired. EX1005, 2:24-27.

B. Rationale (Motivation) Supporting Obviousness

82. In my opinion, a PHOSITA would have been motivated to add Multicast functionality, in the manner described above, to Katis. Abbazia teaches

that Multicast increases transmission efficiency by reducing the need to send the same packet multiple times to reach multiple endpoints:

> "Methods have been developed for **reducing the overhead** associated with connection based protocols, by using connectionless protocols, such as UDP (user datagram protocol) over an IP network. These connectionless protocols typically rely on a broadcast or '**multicast**' model wherein **a single message is broadcast to a multiple receiving systems without forming a connection** with the individual systems. Essentially, multicasting enables a source to transmit messages to a plurality of hosts which have been collectively identified by a single IP destination address, **rather than to each hosts' unique IP address individually**."

EX1005, 1:35-45 (emphasis added).

83. As I explained above, the E2E Limitation requires that certain steps occur without having first established an end-to-end connection between the sending and receiving devices. *Supra* at §IX.A. I understand that, in an *inter partes* review of a related patent, Voxer took the position that "an end-to-end connection" includes a connection "in which the sending and receiving devices directly communicate with each other." EX1025, p.31 (IPR2021-00113, Patent Owner's Preliminary Response). Abbazia notes that avoiding the need to form such a connection between sender and receiver increases efficiency, stating "[t]he multicast approach eliminates the undesirable overhead associated with forming connections with each system…."

EX1005, 2:4-5, 1:28-34. A PHOSITA would have seen a significant benefit to transmitting packets to a Multicast address to reach multiple intended recipients, in order to achieve the efficiency gain of Multicast techniques. *See, e.g.*, EX1024, p.1 ("IP multicast provides a method of efficient many-to-many communication."). A PHOSITA would have seen this as in accord with Katis' desire to optimize "[t]he utilization of network bandwidth" and "minimizing real-time bandwidth". EX1004, ¶¶0041 (optimizing), 0163 (minimizing), 0143, 0145; *see, e.g.*, EX1024, p.2 ("Sending multiple unicast packets … increas[es] bandwidth requirements and costs"; "[i]nstead, data to multiple destinations can be delivered using multicast." (emphasis omitted)). A PHOSITA likewise would have seen an advantage for transmitting devices, which only have to send one packet to reach N recipients, as opposed to sending N packets. *See, e.g.*, EX1024, p.2 ("Multicast allows the source to send a single copy of data, using a single address for the entire group of recipients."). In the combination, both the client and the server(s) 16 of Katis would only need to transmit 1 packet for every device listening to the Multicast address, and each listening device would receive better network performance because of the reduced traffic Multicast enables. The efficiency gain increases with the number of recipients, and would be especially advantageous where a video message is being broadcast to a large number of users.

84. A PHOSITA would have seen additional advantages in using the Multicast system of Abbazia with Katis. As Abbazia notes, "[Multicast] suffers from the inability to guarantee receipt of messages to all systems. For IP networks, multicast is unreliable by design in order to reduce overhead of sending packets to multiple destinations." EX1005, 2:7-10. In other words, as Abbazia notes, a PHOSITA would have understood that a disadvantage of multicast is that the protocol itself does not necessarily ensure that each packet is delivered to the receiver (because recipients do not send acknowledgements to confirm receipt of packets). A PHOSITA, however, would also have known that various techniques could be used to allow the receiver to detect packet loss (e.g., by sequentially numbering the packets), and a PHOSITA would have understood that real-time video playback is highly tolerant of packet loss. Indeed, IP multicast was used for real-time video and other similar applications more than twenty years before the effective filing date of the '516 patent. *See* EX1023. In fact, in a *New York Times* article published in 1995, Mr. Steve Deering (author of RFC1112 (EX1022)) said that the multicast backbone used on the Internet at the time "could become a sort of global video telephone system that allows groups of people -- not just individuals -- to share voice and data and images in 'real time' over the Internet." EX1023, p.1. Furthermore, for real-time video, a Multicast network needs to transmit fewer packets than a unicast network like Katis', and for that reason, will have a lower

error rate due to reduced network congestion. Furthermore, a PHOSITA would generally have found Katis' use of sequence IDs (which are added numbers within packets, one per packet occurring in a sequence among subsequent packets, that allow recipients to observe when a sequence-numbered packet is missing) advantageous with Abbazia's system, because it would increase reliability by allowing devices "to determine if an entire message or some portion of a message is missing". EX1004, Figs. 4C and 4D, ¶¶0115, 0180-0183. This notably tracks Abbazia's suggestion to use sequence numbers. EX1005, 6:65-7:1.

85. A PHOSITA would further have found it obvious to allow a user to encrypt messages to be read, as taught by Katis (EX1004, ¶¶0103, 0102, 0059, 0061, 0170), when including a Multicast address in a conversation, to avoid eavesdropping by third parties. A PHOSITA further would have found it obvious to enable Voxer users to participate in conversations by sending return messages to the Voxer infrastructure including a conversation identifier and sequence ID, based on the teachings of Katis. EX1004, ¶¶0003, 0182, 0019-0020, 0063, 0065-0068, 0038-0053.

86. The combination would have also been obvious because Katis teaches the known problems of developing systems that effectively use network bandwidth (especially for multimedia communications). EX1004, ¶¶0041, 0163, 0143, 0145. As described above, Abbazia teaches a known technique and solution (Multicast in

Abbazia's system) to address this problem, that could have been applied by a PHOSITA with predictable results. In the relevant timeframe, Multicast was part of a known Internet standard (EX1022) and had been successfully used for video broadcasting (EX1023; EX1024).

C. Reasonable Expectation of Success

87. In my opinion, a PHOSITA would have had a reasonable expectation of success in making the proposed combination. The art in the relevant timeframe was predictable. For example, as I explained above, multicast was well-known and a PHOSITA would have known how to implement its use according to the proposed combination. *See, e.g.*, EX1005, 1:35-45; *supra* §X.A (describing the combination), §X.B (addressing the rationale supporting obviousness).

D. Analogous Art

88. In my opinion, Katis and Abbazia are analogous art because they are in the same field as the '516 patent (telecommunications and multimedia communications). EX1001, Title, Abstract, 1:36-4:17; EX1004, Title, Abstract, ¶¶0002-0018; EX1005, Title, Abstract, 1:5-10. Katis addresses the same problems faced by the inventors because it is a publication of an application within the claimed priority chain of the '516 patent. EX1001, Title, Abstract, 1:40-4:23; EX1004, Title, Abstract, ¶¶0002-0018. Abbazia also addresses several problems faced by the inventors including efficiently transmitting and delivering media over a network.

EX1001, 1:47-4:17; EX1004, ¶¶0002-0018, 0041, 0163, 0143, 0145; EX1005, 1:35-45, 2:4-5, 1:28-34.

E. Graham Factors

89. The **level of ordinary skill** in the art is as set forth above in §VII.

90. The **scope and content of the prior art** are discussed throughout the Ground.

91. The differences between the prior art and the claims are discussed throughout the Ground.

92. I am not aware of any **secondary considerations** that would make an inference of non-obviousness more likely, as I discuss in more detail below in paragraph 137.

F. Claim Mapping

93. This section maps the challenged claims to the relevant disclosures of Katis and Abbazia, where the claim text appears in bold-italics, and the relevant mapping follows.

CLAIM 1

"1[Pre]. A method for operating a video message service infrastructure on a network, the method comprising:"

94. In my opinion, the combination renders obvious a **video message service infrastructure**. In the combination, the "**video message service**

infrastructure" includes the Katis system as combined with the system of Abbazia, as shown above in §X.A.2. The **infrastructure** includes Katis' server(s) 16, other components in network 14 as modified (such as multicast-enabled routers), and/or the user devices.

95. Katis teaches that client devices are connected to each other via a **video message service infrastructure** including a plurality of servers on a network. EX1004, ¶¶0101, 0177, 0185-0189, 0061, Fig. 1.

96. Katis shows an infrastructure in Fig. 1, reproduced here:

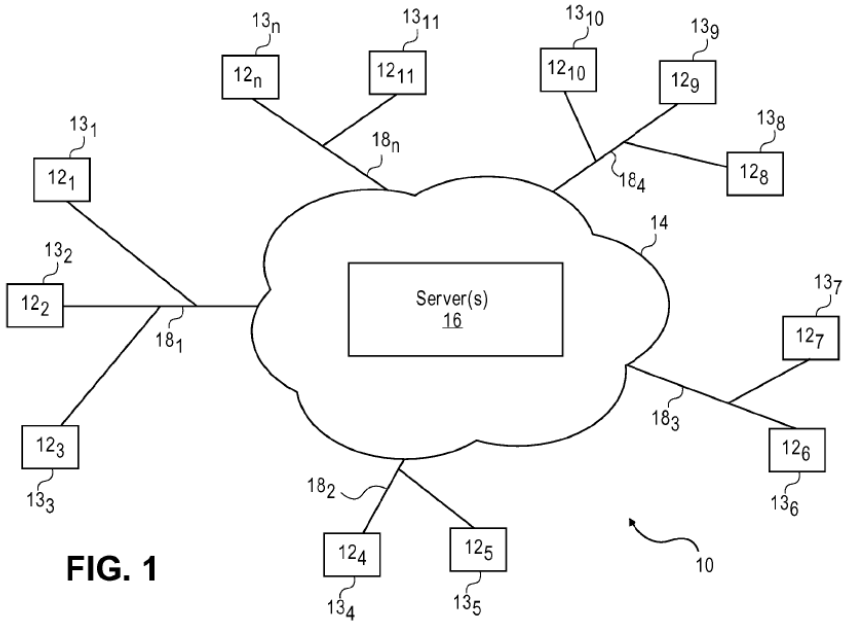


FIG. 1

EX1004, Fig. 1, ¶0022. Fig. 1 shows clients 12 running on devices 13 and in communication over "communication services **network** 14, including one or more **Servers 16**". EX1004, ¶0101 (emphasis added). Katis explains that:

"Clients 12_1 through 12_n communicate with one another and with

> Servers 16 over the networks 18_1 through 18_n and network 14 using individual message units referred to as 'Vox packets', which are described in detail below."

EX1004, ¶0101.

97. The infrastructure allows clients 12 to participate in a **video messaging** conversation thus providing and **operating** a video message **service**. EX1004, Abstract, ¶¶0042-0050, 0177, 0075, 0063-0065, 0003, 0019-0020, 0245, 0060, 0087.

98. The Katis infrastructure participates in the infrastructure of the combination, where the infrastructure of the combination has an alternate Multicast pathway to send real-time video.

"[1a] receiving, at the video message service infrastructure, a video message transmitted by a sending device, video media of the video message transmitted as the video media is created;"

99. In my opinion, the combination renders obvious a **video message service infrastructure receiving a video message transmitted by a sending device**. EX1004, ¶¶0038, 0041, 0063, 0107, 0124, 0132, 0143, 0174, 0194, 0203-0204, 0209, Fig. 1. In the combination, the "infrastructure" includes the Katis system as combined with the system of Abbazia, as shown above in §X.A.2 and discussed in element 1[Pre].

100. Katis teaches that user **sending devices 13 send** messages containing media to servers 16 (part of the **infrastructure**). Katis states:

> "Message: An individual unit of communication from one User to another. Each Message consists of some sort of Media, such as voice or video. Each Message is assigned certain attributes, including: (i) the User sending the message; (ii) the Conversation it belongs to; (iii) an optional or user created Importance Tag; (iv) a time stamp; and (v) the Media payload."

EX1004, ¶0063, *see also* EX1004, ¶¶0038-0040, 0041, 0065, 0132, 0257, 0057, 0059-0060, 0063-0065, 0087.

101. Information establishing a conversation is provided in connection with an initial message from a first user. Katis explains:

> "Create a Conversation 1104 b—A User creates a Conversation through the interface 110 by inputting a Conversation name, a list of Contacts, and an optional scheduled start time. If no start time is designated, it is assumed the start time is immediate. In response, the MCMS application 20 creates a new Conversation in the database 22, associating records for each Participant on the Contacts list."

EX1004, ¶0257.

102. The conversation, in turn, is associated with a list of participants that are designated recipients of the message. EX1004, ¶0132 ("When sending a Message as part of a Conversation that is associated with a number of Contacts, all

of the Contacts receive the Message."); *see also* ¶¶0038, 0041, 0257. As discussed above, it would have been obvious to allow users and/or intermediate servers to elect to use a Multicast address to receive/transmit real-time video messages.

103. Katis also teaches or suggests that **video media of the video message** is **transmitted as the video media is created**. Katis teaches portions of a **video message** can be **streamed** to **another user device** and rendered as that message is still being created, *i.e.*, the message is incomplete. This can be done, for instance, by continuously rendering a message as it is received as Vox packets:

> "The Conversation/Message management service 20f processes the transmission and receipt of Messages in a progressive fashion. **When transmitting, Media may be created while Messages are simultaneously encoded, stored and transmitted**. In other words, **the transmission of Messages may occur simultaneously with the generation of Media by the User** (i.e., while speaking into their Device 13 or generating video). On the receiving side, the receipt, storage, and Rendering of Messages also all occur progressively. **Messages do not need to be completely received before they can be Rendered**. The Rendering of Messages may occur at the same time Messages are being delivered, right up to the MTSD. Further, the service 20 f is also capable of the simultaneous transmission of outgoing Messages and Rendering of incoming Messages. **The progressive nature of the service 20f allows a User to be engaged in a live Conversation while storing and streaming the media of the Conversation for later retrieval** and review as well other

functions described herein."

EX1004, ¶¶0124, 0204-0205, 0042 (emphasis added).

"[1b] transmitting the video media of the video message from the video message service infrastructure to a recipient device so that a recipient can render at least a portion of the video media of the video message on the recipient device, the video message service infrastructure operating so that the rendering occurring:

(i) while the video media of the video message is being created and transmitted over the network by the sending device; and"

104. See analysis for claim element [1a]. Video media of the video message is **transmitted** to and **rendered** at a recipient device **while the video media of the video message is being created and transmitted**. EX1004, ¶¶0174, 0124, 0019, 0204-0205, 0128. A PHOSITA would have understood that this concurrent transmission from the first (transmitting) device to the second (recipient) device occurs via the video message service infrastructure. *E.g.*, EX1004, ¶¶0197, 0203-0205. In the combination, the video message is transmitted to the recipient device via the Multicast address.

"[1c] (ii) without having to first establish an end-to-end connection over the network between the sending device and the recipient device before the sending device begins to transmit the video media of the video message,"

105. Regarding the claim language that **a sending device begins to transmit the video media of the video message**, *see* immediately above, elements [1a] and [1b]. EX1004, ¶0128.

106. In my opinion, the combination renders obvious doing so **without having to first establish an end-to-end connection over the network between the sending device and any recipient devices respectively.**

107. As explained above in §X.A and §X.B, it would have been obvious to combine Katis with Multicast functionality, as taught by Abbazia. Abbazia teaches that Multicast functionality avoids connections with target devices to gain efficiency. EX1005, 1:35-45, 2:4-5 ("The multicast approach eliminates the undesirable overhead associated with forming connections with each system…."). Specifically, in Multicast, a packet is transmitted to single Multicast IP address, and then network-level devices, such as routers, forward it to its end destination. When the sending device sends to a Multicast address (e.g. when livestreaming a video to a large audience and all users opt to use Multicast), no **end-to-end connection** is formed **with the recipient device before the sending device begins to transmit the video media of the video message**.

"[1d] wherein the video media of the video message is segmented into a plurality of segments, the plurality of segments containing content of the video media,"

108. Katis states that media is segmented into indexed packets/payloads such that the **plurality of segments contain[] content of the video media**. For example, Katis states that "the Media is segmented and placed into the payloads of a plurality of Vox packets 95." EX1004, ¶¶0183, 0019, 0205, 0311 ("The Messages are segmented from a stream of Media….").

"[1e] the plurality of segments are received at the video message service infrastructure while streamed by the sending device, stored by the video message service infrastructure, and at least some of the segments are streamed to the recipient so that the recipient can render at least the portion of the video media of the video message on the recipient device while (i) is occurring and without first establish the end-to-end connection of (ii)."

109. **The plurality of segments are received at the video message service infrastructure while streamed by the sending device**. See analysis for claim elements [1a] and [1d]. EX1004, ¶¶0124, 0203-0205, 0042; *see also* EX1004, ¶¶0183 ("The packets are then continually, stored, transmitted, received, stored and rendered in streams (i.e., streaming media) …."), 0197 (A sending device "converts the Media into indexed media payloads …, creates Vox packets, and transmits the packets to the receiving Client 12 (box 134) over the network 18. Each Server 16 along the path between the sending Client 12 and the receiving Client 12 stores the transmitted payloads in the PIMB 85 and transmits the Vox packets to the next hop (box 133).").

110. The packets (**plurality of segments**) are also **stored by the video message service infrastructure**. See analysis for claim element [1a]. EX1004, ¶¶0124, 0203-0205, 0042, 0020, 0045, 0061, 0065, 0084, 0093, 0100, 0121, 0143, 0145, 0172, 0174, 0183-0184; *see also* EX1004, ¶0203 ("In FIG. 8F, the sequence for a Server 16 to receive Vox packets from the previous hop on the network 18, store, archive and transmit the Vox packets to the next hop is illustrated.").

111. **At least some of the segments are streamed to the recipient**. *See* analysis for claim elements [1a] and [1b]. *See also, e.g.*, EX1004, ¶¶0196-0205.

112. Streaming occurs **so that the recipient can render at least the portion of the video media of the video message on the recipient device while (i) is occurring** (i.e., **the video media of the video message is being created and transmitted over the network by the sending device**). *See* analysis for claim element [1b]. EX1004, ¶¶0174, 0019, 0204-0205, 0128, 0124.

113. No **end-to-end connection** is **establish[ed]** as discussed under claim element [1c].

CLAIM 2

"2. The method of claim 1, further comprising indexing the plurality of segments respectively."

114. This limitation is taught or suggested by Katis. EX1004, ¶¶0019, 0038, 0040, 0041, 0065, 0132, 0063, 0148-0149, 0155, 0181-0183, 0193, 0228, 0246,

0291, 0294, 0293, 0295-0297, 0302, 0309, 0311, Figs. 13A, 8B, 4C, 4D, claim 2, claim 19. For example, Katis states that "hardware 34 receives the raw audio and/or video signals and provides them to the Encoder Receiver 28a, which encodes the signals into **indexed** media payloads." EX1004, ¶0155 (emphasis added). Katis states that "each packet 95 is **indexed**, time-stamped, and given a sequence identifier." EX1004, ¶¶0183, 0311 (emphasis added).

CLAIM 3

"3. The method of claim 1, further comprising timestamping the plurality of segments respectively."

115. See analysis for claim 2 (describing time-stamping); *see* EX1004, claim 19 ("The method of claim 1, wherein the media comprises one or more indexed media payloads each having a time stamp and unique sequence number.").

CLAIM 4

"4. The method of claim 1, further comprising providing a sequence identifier for the plurality of segments respectively."

116. See analysis for claim 2 (describing the use of sequence identifiers); *see* EX1004, claim 19 ("The method of claim 1, wherein the media comprises one or more indexed media payloads each having a time stamp and unique sequence number.").

CLAIM 5

"5. The method of claim 1, further comprising trans-coding [sic] the content contained in the plurality of segments from a first codec used by the sending device into a second codec used by the receiving device."

117. It would have been obvious to perform this function as suggested by Katis. EX1004, ¶¶0119, 0139, 0059, 0085, 0074, 0082, 0139, 0061, 0102, 0221-0222, 0236, Fig. 9C. For example, Katis defines encoding as:

> "Encoding: The process of translating Media either created by a User (such as voice or video) or otherwise originating on a Device (such as GPS or other sensor data), and converting the media into digital data to be processed by a Client."

EX1004, ¶0082. Katis states that "[s]ervers provide . . . transcoding . . . and the optimization of the transmission of data across the network." EX1004, ¶0061. Katis further states that "different Codec or Codecs, each having a different bit rate versus quality tradeoff, may be used." EX1004, ¶0221. Katis states that "the rendering and encoding module 21 is responsible for performing all rendering tasks for the MCMS application 20. These tasks include rendering Media appropriately for the device 13 running application 12." EX1004, ¶¶0128, 0139, 0161.

118. It would have been obvious to also do this in the combination, especially when it is known that the encoding of the sending device for real-time video does not match the encoding of the recipients, or where network conditions

generally prescribe a lower use of bandwidth. For example, Katis states that "[w]hen insufficient underlying network resources are available, the quality of the transmitted Media streams can be degraded." EX1004, ¶0143. Katis states that "current bandwidth conditions" are taken into account. EX1004, ¶0055.

CLAIM 6

"6. The method of claim 1, further comprising degrading a quality of the content contained in the plurality of segments as received by the video message service infrastructure for transmission to the recipient devices when conditions on the network warrant."

119. This limitation is taught or suggested by Katis. EX1004, ¶¶0041, 0043, 0055, 0143, 0145, 0207, 0218, 0221-0223, 0228, 0236, Fig. 9C. For example, Katis states that "[w]hen insufficient underlying network resources are available, the quality of the transmitted Media streams can be degraded." EX1004, ¶0143. Katis states that "current bandwidth conditions" are taken into account. EX1004, ¶0055. See also claim 5 analysis.

CLAIM 7

"7. The method of claim 6, wherein degrading the quality of the content contained in the plurality of segments comprises reducing a bit rate used to represent the content relative a complete bit rate representation of the content as received by the video message infrastructure."

120. See claims 5 and 6 analysis. For example, Katis states: "if the MABR [Maximum Available Bit Rate] is less than the AIBR [Aggregate Ideal Bit Rate],

then a series of procedures are applied . . . [including] adjustments to the payload content and quality . . . to meet current network conditions" (¶0218) and the "bit rate may be progressively lowered" (¶0223). EX1004, ¶¶0207, 0210, 0215, 0218-0223, Fig. 9C.

CLAIM 8

"8. The method of claim 1, further comprising generating two or more bit rate versions used to represent the content contained in the plurality of segments of the video message, each of the two or more bit rate versions representing different degraded quality versions of the video media of the video message."

121. See claims 5, 6, and 7 analysis. For example, it would have been obvious to provide different degraded-quality versions of the message to different recipients, as suggested by Katis. EX1004, ¶¶0055 ("enabling the system to optimize the utilization of network bandwidth by making tradeoffs between timeliness and media quality, using the presence and intentions of the recipient(s) (i.e., to either review the media in real time or in a time-shifted mode), as well as measures of network latency, network degradation, packet loss or damage, and/or current bandwidth conditions."), 0207-0237 (discussion of how recipient devices receive degraded version of video media), 0143 ("The Store and Stream module 24 … optimizes the delivery speed and bandwidth usage across the network 18 to meet a User's immediate performance requirements, while managing network quality and capacity. … When insufficient underlying network resources are available, the

quality of the transmitted Media streams can be degraded."), 0207 ("The PQM 26g relies on a metric called the Maximum Available Bit Rate (MABR), which is a continually computed approximation of actual transmission capacity or bandwidth (i.e., a measure of the capability of the network at a given point in time) between a sending and receiving node pair."); EX1011,[4] pp. 43-46; *see also* EX1004, ¶¶0162-0169, 0017 ("When a recipient is not going to render data immediately, these prior art systems unnecessarily use bandwidth when it is not needed ….").

CLAIM 9

"9. The method of claim 1, further comprising implementing the video message service infrastructure with a plurality of servers on the network."

122. See claim elements 1[Pre] and [1a] analysis. See also EX1004, ¶0174 ("It should be noted that only a single Server 16 is illustrated in FIG. 1 for simplicity. It should be understood that in actual embodiments, multiple servers or a 'server farm' may be used for a network with a large number of Users.").

CLAIM 10

"10. The method of claim 1, further comprising: storing the content of the plurality of segments of the video message in a storage location accessible by the video message service infrastructure; and

[4] I understand that Katis incorporates by reference U.S. Provisional Application No. 60/999,619 (contained in EX1011). EX1004, ¶0001.

providing from the storage location to the recipient device the content of the plurality of segments of the video message for rendering in a time-shifted mode."

123. This limitation is taught or suggested by Katis. EX1004, ¶¶0020, 0040, 0045, 0061, 0065, 0067, 0084, 0093, 0100, 0121, 0127, 0143, 0145, 0172, 0174, 0184, 0204-0205, claim 6. For example, Katis states that "the media payloads generated by a Client 12 are stored in multiple locations" (including sending device 13, intermediate server(s) 16, and receiving device 13) and claim 6 of Katis recites: "rendering on the first communication device the persistently stored received media in a time-shifted mode." EX1004, ¶0184, claim 6; *see also* EX1004, ¶0205 ("It will also be continually rendered as the Media is being received when in the near real-time mode or from storage when in the time-shifted mode."). See also analysis for claim element [1b].

CLAIM 11

"11[Pre]. The method of claim 1, further comprising:

[11a] receiving, over the network at the video message service infrastructure, a second video message transmitted by a second sending device, second video media of the second video message transmitted as the second video media is created;"

124. See analysis for claim element [1a]. Katis teaches allowing the various users to contribute one or more additional messages to a messaging conversation. Katis states:

> "Depending on preferences and priorities, the recipient(s) might **participate** in the **conversation** in real time, or they might simply be notified that the message is ready for retrieval. In the latter case, the recipient participates in the conversation in a time-shifted mode by reviewing and **replying to the recorded message** at their convenience."

EX1004, ¶0038 (emphasis added); *see also* EX1004, ¶0065 ("Users generally participate in a Conversation using their Devices by either Reviewing Messages in real time or in a time-shifted mode, or creating and sending Messages of a Conversation as desired."); ¶¶0063-64, 0080-0081, 0101, 0124, 0128, 0146, 0161, 0183, 0202, 0299, 0302-0304, 0309-0311, Fig. 13A; 0205, 0208, 0041, 0045, 0065, 0067, 0100, 0135, claim 1, claim 7.

"[11b] transmitting the second video media from the video message service infrastructure over the network to the recipient device so that the recipient can render at least a portion of the second video media on the recipient device, the video message service infrastructure operating so that the rendering occurs: (i) as the second video media of the second video message is created and transmitted over the network by the second sending device;"

125. See analysis for claim element [1b]. Katis teaches that any of the **users** can **transmit one or more additional messages** in a conversation in a similar manner as an original message sent. EX1004, ¶¶0019, 0065, 0069, 0081, 0183. For example, Katis states:

> "When transmitting, Media may be created while Messages are simultaneously encoded, stored and transmitted. In other words, the transmission of Messages may occur simultaneously with the generation of Media by the User . . . .[T]he service 20f is also capable of the simultaneous transmission of outgoing Messages and Rendering of incoming Messages. The progressive nature of the service 20f allows a User to be engaged in a live Conversation while storing and streaming the media of the Conversation for later retrieval and review as well other functions described herein."

EX1004, ¶¶0124, 0311 (emphasis added).

"[11c] (ii) without having to first establish an end-to-end connection over the network between the second sending device and the recipient device before the second sending device begins to transmit the second video media of the second video message; and

126. See analysis for claim element [1c].

[11d] (iii) simultaneously with the video media of the video message."

127. See analysis for claim element [11b]. See also EX1004, ¶¶0065, 0080-0081, 0108, 0124, 0128, 0146, 0161, 0183, 0202, 0204-0205, 0299, 0303-0304, 0309-0311, Fig. 13A. For example, Katis describes:

> "Mixing: The Rendering of one or more Media streams. For example, the Media stream from two Participants of a Conversation may be Mixed when Rendered, creating a User experience similar to

a conversation where multiple people are speaking at the same time."

EX1004, ¶0081.

CLAIM 12

"12. The method of claim 1, further comprising providing an application for execution on the sending device, the application enabling a sender to create and stream the plurality of segments containing content of the video message as the video media is created."

128. See analysis for claim element [1a]. Katis teaches or suggests this claim element. For example, Katis defines a "Client" as: "Client: A Client is the user application in the communication system, which includes . . . 'Voxing functionality.' Users interact with the Client application, and the Client application manages all communications . . . ." EX1004, ¶¶0059, 0060, 0101, 0121, Fig. 14A; *see also* EX1004, ¶¶0238-0244 (describing the user interface).

CLAIM 13

"13. The method of claim 1, further comprising providing an application for execution on the recipient device, the application enabling the recipient to selectively render the plurality of segments containing the content of the video message: at least partially as the content of the plurality of segments of the video message is created, transmitted by the sending device and delivered by the video message service infrastructure; and in a time-shifted mode by rendering the video media of the video message out of storage."

129. See analysis for claims 1, 10, and 12. This limitation is taught or suggested by Katis. EX1004, ¶¶0020, 0045-0046, 0048, 0061, 0065, 0084, 0093,

0100, 0121, 0143, 0145, 0172, 0174, 0184, 0204, claim 6. For example, Katis states that that "the media payloads generated by a Client 12 are stored in multiple locations" and claim 6 of Katis recites: "rendering on the first communication device the persistently stored received media in a time-shifted mode." EX1004, ¶0184, claim 6. See also analysis for claim element [1b].

CLAIM 14

"14. The method of claim 1, wherein the sending device is arranged to initiate transmission of the video message to the video message service infrastructure independently of the video message service infrastructure ascertaining a location of any of one or more recipient devices on the network."

130. In my opinion, this limitation is taught or suggested by the Katis and is implemented in the combination. EX1004, ¶¶0040, 0047, 0122, 0142, 0145, 0174 0191. For example, Katis states that "to initiate a conversation, one can simply begin transmitting to an individual or a group." EX1004, ¶0040. Katis states that "Message transmission is essentially "full-duplex', enabling any party to send a Message at any time, even while another party is also sending a Message, or if the other party is unavailable or otherwise engaged." EX1004, ¶0142. This transmission by a sending device occurs **independently of the video message service infrastructure ascertaining a location of any of one or more recipient devices on the network** both according to Katis and in the combination when using Multicast. Katis teaches that an initiating user can transmit to a server 16, and that

the server 16 stores the message and forwards it, if appropriate. It is a server 16 that must determine the location of recipients (including potentially a Multicast address) on the network, not the sending device, making the ascertainment **independent**. EX1004, ¶¶0174, 0191.

CLAIM 15

"15. The method of claim 1, wherein the sending device is arranged to transmit the video message to the video message service infrastructure regardless of whether any of one or more recipient devices are available or not available on the network."

131. See analysis for claim 14. See also EX1004, ¶0047 (discussing beginning a conversation when "other participants are unavailable").

CLAIM 16

"16. The method of claim 1, further configured to provide a notification of the video message to one or more recipient devices respectively."

132. Katis teaches a **notification** can be provided to each participant when that user's device receives a message. EX1004, ¶¶0107, 0038, 0268, 0137. The combination infrastructure ascertains devices by their IP addresses and provides the notification to the second users. EX1004, ¶¶0191, 0061, 0115; EX1005, 1:55-60, 5:30-34.

CLAIM 17

"17. The method of claim 1, further comprising:

receiving one or more identifiers with the received video message, the one or more identifiers associated with one or more recipients of the video message; resolving the one or more identifiers into one or more IP addresses on the network, each of the one or more IP addresses associated with one or more recipient devices associated with the one or more recipients respectively; and using the one or more IP addresses for routing the plurality of segments over the network when transmitting the video message from the video message service infrastructure to the one or more recipient devices."

133. In my opinion, this limitation is taught or suggested by Katis and obvious in view of the combination. EX1004, ¶¶0063, 0073, 0179, 0185, 0190, 0191, 0132, 0105. For example, Katis states:

> "The address of a target Client 12 needs to be known so that the underlying network delivers the Vox packet 95 to the correct location. With IPv4 networks, the address is typically an IPv4 Address, which is a 32-bit number that uniquely identifies a host within the network. For other networking technologies, the address could be some other type of identifier. IP networks use the Domain Name System (DNS) to resolve human-readable names into IP addresses, and the Address Resolution Protocol (ARP) to resolve IP addresses into physical addresses. Regardless of the underlying networking technology, the system 10 uses one of the above mentioned or other known addressing schemes for delivery of Vox packets 95 to the correct location."

EX1004, ¶0191. When a user initiates a conversation, a list of recipients is associated with conversation identifier and information establishing the conversation is

provided with an initial message from a first user. EX1004, ¶¶0063-0065, 0038, 0040, 0041, 0065, 0132, 0091, 0191, 0257, 0267, 0087, 0163, 0180-0181, 0247; EX1011,[5] p.39 ("2. Payload Data a. Conversation Identifier b. Message Identifier"), p.44 ("Data Elements … 3. Recipient Data b. Identifier (transmission target)"). When the user sends a message, it includes this identifier and the server translates that identifier into recipients, and then ascertains their IP addresses for actual transmission. *See id.* In the combination, when a message is sent to a Multicast address (an IP address reserved for Multicast), the server **resolves** an **identifier** (a conversation identifier) into the Multicast address, which is an **IP address** associated with one or more recipient devices (users in the conversation who elect to receive messages via Multicast). A message (segmented into Vox packets) is sent to the server(s) 16 as usual, but is then forwarded to the Multicast address for Voxer users in the conversation who elect to receive messages via Multicast and then transmitted ultimately to the recipient users. The multicast IP address is then used to route packets of the message at each multicast router in the network to each client device which is part of the multicast group.

[5] Katis incorporates by reference U.S. Provisional Application No. 60/999,619 (contained in EX1011). EX1004, ¶0001.

CLAIM 18

"18[Pre]. The method of claim 1, further comprising:

[18a] configuring the video message service infrastructure to include a plurality of servers located on the network, the plurality of servers collectively configured for performing:"

134. See analysis for claim elements 1[Pre] and [1a] and claim 9.

"[18b] receiving the video message; ascertaining one or more locations on the network for one or more recipient devices, the one or more locations derived from one or more identifiers received with the video message and for identifying the one or more recipients of the video message; and transmitting the video message to the one or more recipient devices so that at least a portion of the video message can be rendered at the same time as the sending device is sending the video message."

135. See analysis for claim element [1a] (**receiving the video message**); claim element [1b] (**transmitting the video message**); and claim 17 (**identifiers received with the video message**).

136. In my opinion, the combination renders obvious **ascertaining one or more locations on the network for one or more recipient devices, the one or more locations derived from one or more identifiers received with the video message and for identifying the one or more recipients of the video message**. The combination infrastructure ascertains recipient devices by their IP addresses. EX1004, ¶¶0191, 0061, 0115; EX1005, 1:55-60, 5:30-34. In the combination, the

server 16 determines that recipients in a conversation are associated with either their own IP addresses or a Multicast IP address. Both a single device's IP address, and a Multicast IP address, are addresses "for" a recipient device. In addition, the multicast routers then route multicast packets to each client device which is part of the multicast group.

XI. SECONDARY CONSIDERATIONS

137. It is my understanding that, in *ex parte* reexaminations of the '557 and '270 patents filed by Meta, Voxer argued that there were objective indicia of non-obviousness. In my opinion, even if any such evidence applies to the challenged claims and is considered, it would not overcome a determination of obviousness for at least the reason that, as explained above, Katis itself is prior art to the '516 patent, and it discloses any alleged merits of the claimed invention.

XII. OATH

138. This declaration and my opinions herein are made to the best of my knowledge and understanding, and based on the material available to me, at the time of signing this declaration. I declare under penalty of perjury under the laws of the United States of America that all statements made of my own knowledge are true and that all statements made on information and belief are believed to be true. I understand that willful false statements and the like are punishable by fine or

imprisonment, or both (18 U.S.C. 1001) and may jeopardize the validity of the application or any patent issuing thereon.

July 13, 2026
Date

Henry H. Houh.
Dr. Henry Houh