

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

**PETITION FOR INTER PARTES REVIEW
UNDER 35 U.S.C. §§ 311-319 AND 37 C.F.R. § 42.100 ET SEQ.**

TABLE OF CONTENTS

I.	MANDATORY NOTICES	1
A.	Real Parties-in-Interest	1
B.	Notice of Related Matters	1
C.	Notice of Lead and Backup Counsel.....	3
D.	Service Information.....	3
II.	CERTIFICATION OF GROUNDS FOR STANDING	3
III.	GROUNDS OF UNPATENTABILITY.....	4
IV.	INTRODUCTION	4
A.	Summary of the '516 Patent.....	4
B.	Overview of Prosecution History	6
V.	LEVEL OF ORDINARY SKILL IN THE ART	12
VI.	CLAIM CONSTRUCTION	12
VII.	PRIORITY DATE FOR THE CHALLENGED CLAIMS	13
A.	The E2E Limitation	14
B.	The Priority Date of the '516 Patent is no Earlier than August 8, 2016.	15
1.	Legal Standards.....	16
2.	The Priority Applications Filed Prior to August 8, 2016 Lack Support for the E2E Limitation	17
VIII.	DETAILED EXPLANATION OF THE REASONS FOR UNPATENTABILITY	20
Ground 1.	Claims 1-18 are Obvious over Katis and Abbazia.....	20

A.	Overview of the Combination	20
1.	Overview of Abbazia	21
2.	The Combination.....	23
B.	Rationale (Motivation) Supporting Obviousness.....	26
C.	Reasonable Expectation of Success	29
D.	Analogous Art	30
E.	Claim Mapping.....	30
	CLAIM 1	30
	CLAIM 2	39
	CLAIM 3	39
	CLAIM 4	39
	CLAIM 5	40
	CLAIM 6	41
	CLAIM 7	42
	CLAIM 8	42
	CLAIM 9	43
	CLAIM 10	44
	CLAIM 11	45
	CLAIM 12	47
	CLAIM 13	48
	CLAIM 14	48
	CLAIM 15	49
	CLAIM 16	50

CLAIM 17 50

CLAIM 18 52

IX. SECONDARY CONSIDERATIONS 53

X. CONCLUSION.....54

TABLE OF EXHIBITS

Exhibit No.	Description
1001	U.S. Patent No. 11,146,516 (“ ’516 patent ”).
1002	Declaration of Dr. Henry Houh.
1003	C.V. of Dr. Henry Houh.
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, <i>Voxer, Inc. v. Google, LLC, et al.</i> , 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,” <i>Internet Engineering Task Force (IETF)</i> (August 1989).
1023	P. Lewis, “Peering Out a ‘Real Time’ Window,” <i>New York Times</i> (Feb. 8, 1995).
1024	“IP Multicast Explained,” <i>Metaswitch Networks</i> (copyright 2004-2009).
1025	<i>Facebook, Inc. v. Voxer IP LLC</i> , IPR2021-00113, Paper 13 (Patent Owner’s Preliminary Response).

Petitioner respectfully requests *inter partes* review of claims 1-18 of U.S. Patent No. 11,146,516 (“the ’516 patent”).

I. MANDATORY NOTICES

A. Real Parties-in-Interest

The real-parties-in-interest for this petition are Google LLC¹ and YouTube, LLC (collectively, “Petitioner”). YouTube, LLC is a subsidiary of Google LLC.

B. Notice of Related Matters

Petitioner is aware of the following matters that may affect, or be affected by, a decision in this proceeding.

The ’516 patent has been asserted in two litigations, both of which are pending: *Voxer, Inc. v. Google, LLC, et al.*, Civ. Action No. 1:25-cv-00873 (D. Del.) (“the Google Litigation”), filed July 14, 2025, and *Voxer, Inc. v. Amazon.com, Inc.*, Civ. Action No. 1:25-cv-00872 (D. Del.) (“the Amazon Litigation”), filed July 14, 2025. All defendants have filed motions to transfer venue and to dismiss under FRCP 12(b)(6). The Court has not yet ruled on these motions, and no parties have served infringement or invalidity contentions; however, fact discovery is underway. The Markman hearing is scheduled for April 8, 2027, a jury trial in the Amazon Litigation is scheduled for November 27, 2028, and a jury trial in the Google

¹ Google LLC is a subsidiary of XXVI Holdings Inc., which is a subsidiary of Alphabet Inc. XXVI Holdings Inc. and Alphabet Inc. are not real parties-in-interest to this proceeding.

Litigation is scheduled for February 5, 2029.

In addition, Petitioner is concurrently requesting *inter partes* review of two other patents that are related to the '516 patent: USP 11,658,929 (in IPR2026-00422) and USP 11,777,883 (in IPR2026-00423). Both of these patents are also at issue in the Google and Amazon Litigations.

To Petitioner's knowledge, none of the three challenged patents (i.e., the '929, '516, and '883 patents) have been at issue in any other proceedings.

There were, however, prior proceedings involving the two other patents at issue in the Google and Amazon Litigations: USP 10,142,270 and USP 10,511,557. These two patents are related to the '516 patent and were at issue in proceedings involving Meta Platforms that began before the filing of the '516, '883, and '929 patents. Specifically, the '270 and '557 patents were at issue in *Voxer, Inc. v. Meta Platforms, Inc. f/k/a Facebook, Inc.*, Civ. Action No. 1:20-cv-00655 (W.D. Tex.) ("the Meta Litigation"), filed January 7, 2020. In September 2022, a jury found infringement and awarded damages; however, the Court ultimately vacated the jury verdicts and dismissed the case with prejudice pursuant to the parties' joint motion. In addition, Meta filed IPRs challenging the '270 patent (IPR2021-00115) and the '557 patent (IPR2021-00112 and IPR2021-00113), which were all denied on the merits on May 7, 2021. Also, on May 18, 2023, Meta filed requests for *ex parte* reexamination of the '270 patent (App. No. 90/019,212) and the '557 patent (App.

No. 90/019,211), each of which ultimately resulted in a reexamination certificate confirming the patentability of the challenged claims and adding new claims.

C. Notice of Lead and Backup Counsel

Lead Counsel: Matthew A. Smith (Reg. No. 49,003)
smith@smithbaluch.com
(202) 669-6207

Backup Counsel: Elizabeth Laughton (Reg. No. 70,484)
laughton@smithbaluch.com
(703) 585-8839

Juliet Mitchell Dirba (Reg. No. 53,334)
dirba@smithbaluch.com
(512) 909-4153

D. Service Information

Petitioner consents to electronic service by email on lead and backup counsel at the email addresses listed above.

Post and hand delivery: SMITH BALUCH LLP
77 Newbury Street
3rd Floor
Boston, MA 02116

II. CERTIFICATION OF GROUNDS FOR STANDING

Petitioner certifies that the '516 patent is available for *inter partes* review, and that Petitioner is not barred or estopped from requesting an *inter partes* review on the grounds identified in the petition. This petition is timely filed within one year of the date of service of Patent Owner's complaint. EX1013.

III. GROUNDS OF UNPATENTABILITY

Petitioner requests cancellation of claims 1-18 of the '516 patent based on the following grounds:

Ground	Claim(s) Challenged	35 U.S.C. §	Reference(s)
1	1-18	103	Katis, Abbazia

As explained below (*infra* §VII, beginning on p.13), 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. As a result, each of the relied-upon references is prior art under 35 U.S.C. §102(a)(1)—Katis was published on January 1, 2009, and Abbazia issued on July 13, 2010. 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.

IV. INTRODUCTION

A. Summary of 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; EX1002, ¶39.

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

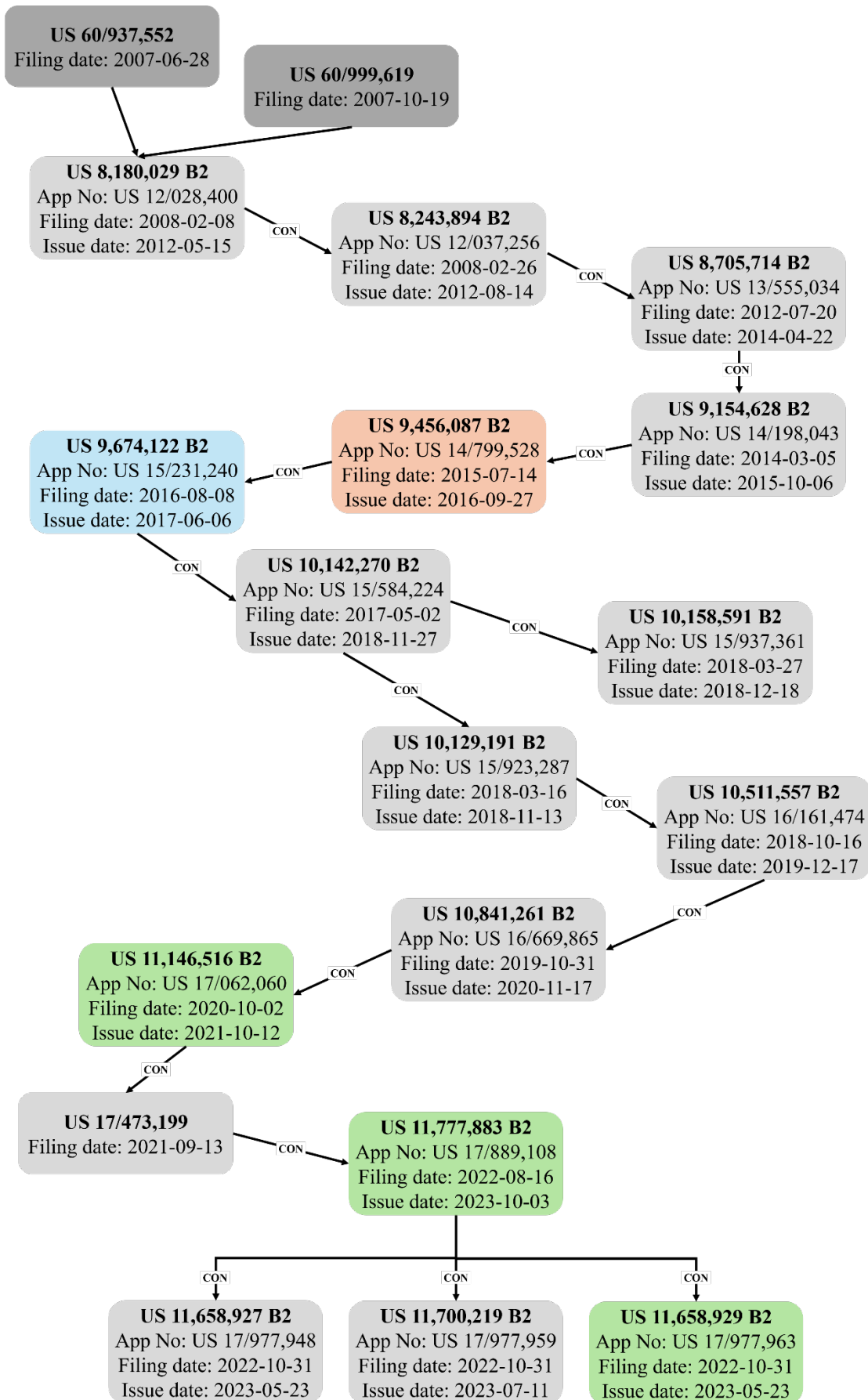
the sender and/or recipient is unavailable: “the Message [will be] queued for transmission as soon as network connectivity becomes available, resulting in a time-shifted delivery.” *E.g.*, EX1001, 18:33-42; EX1002, ¶40.

In operation, Client 12₁ creates media, stores the media payloads, and “transmits the packets to the receiving Client 12₂ 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 12₁ and the receiving Client 12₂ stores the transmitted payloads ... and transmits the Vox packets to the next hop.” EX1001, 27:42-45. After receiving the message, Client 12₂ stores the payloads, renders the payload information, and sends an acknowledgement to the sender. EX1001, 27:45-54, 21:25-28; EX1002, ¶41.

B. Overview of Prosecution History

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.²

² The other two patents for which Petitioner is concurrently requesting *inter partes* review are also highlighted in green.



The complete family, however, includes approximately 101 U.S. non-provisional applications. EX1012.

Over time, the claims Voxer sought to obtain increasingly diverged from what it described in its applications. EX1002, ¶44. For example, in the '256 application, filed in 2008, 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; EX1002, ¶44.

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; EX1002, ¶45. 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). EX1002, ¶45. 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.” EX1002, ¶45.

As another example, 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; EX1002, ¶46. 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 EX1002, ¶46. Despite never successfully developing a system which could in fact transmit video in the manner required by the claims, Voxer expanded its

claims over time to focus on “video media” as required by all claims of the ’516 patent. EX1002, ¶46.

Voxer’s prosecution strategy has been a long war of attrition. 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, Voxer not only pursued a voluminous number of patents over the years (*see* EX1012), but it would regularly present large numbers of claims. This can be seen in the related ’883 patent, concurrently challenged by Petitioner, which contains 65 claims. EX1014. 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; EX1002, ¶47.

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 Mangal³ principal prior art as well as the Teng, Kalaboukis, Yu, Honda, and Mattila secondary references. *Id.* at pp.167-179.

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. 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**

³ US 2003/0148785A1 (“Mangal”). During prosecution, the Examiner also cited the US 2003/0208543A1 (“Enete”), US 2008/0062246A1 (“Woodworth”), and US 2007/0207782A1 (“Tran”) references as pertinent. *Id.* at pp.247-249, 175-178.

the video message.” *Id.* at p.233.

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; EX1002, ¶¶48-50.

V. LEVEL OF ORDINARY SKILL IN THE ART

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. EX1002, ¶52. A higher level of education could substitute for less industry experience, and more industry experience could substitute for the specific level of education. EX1002, ¶52.

VI. CLAIM CONSTRUCTION

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”); EX1002, ¶54. Although it is not clear which (if any) of

those terms will need to be construed,⁴ Petitioner will not take a position in the Litigation that is inconsistent with any of the understandings reflected herein. Also, Petitioner will promptly notify the Board of any claim construction it proposes to the District Court to permit verification of this consistency.

Moreover, it is unlikely that the Board will need to address claim construction in this proceeding. *See Realtime Data, LLC v. Iancu*, 912 F.3d 1368, 1375 (Fed. Cir. 2019) (PTAB need only construe terms to the extent needed to resolve disputes). Because 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 (EX1002, ¶55), the parties are less likely to dispute claim constructions.

VII. PRIORITY DATE FOR THE CHALLENGED CLAIMS

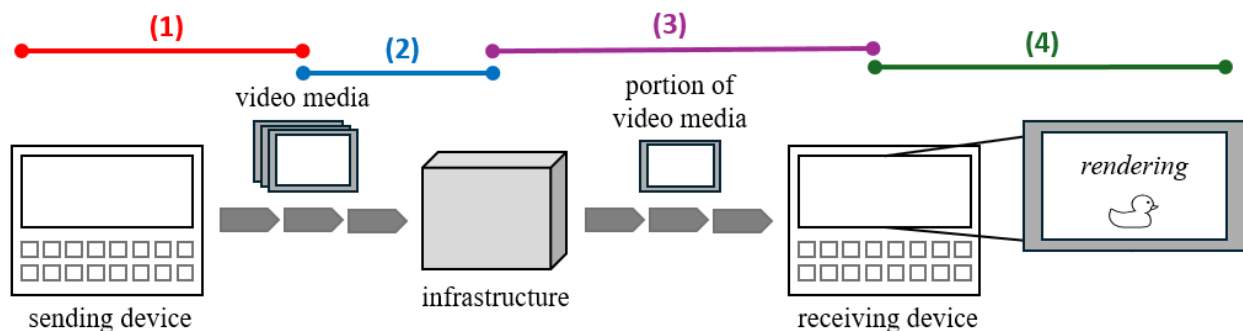
The '516 patent includes a lengthy priority claim (*see* EX1001, codes (60), (63)); however, the earliest possible priority date for the challenged claims is August 8, 2016 (i.e., filing of App. No. 15/231,240) because each claim recites a negative limitation (“the E2E Limitation”) that is not found in any of the earlier-filed applications. EX1002, ¶56.

⁴ The parties have not yet identified the asserted claims, infringement contentions, or invalidity contentions, so it is unknown whether there will be any disputes regarding the proper understanding of the claims.

A. The E2E Limitation

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.**”

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. EX1002, ¶58. These requirements are illustrated in the graphic below.



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. EX1002, ¶59. 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. EX1002, ¶59. 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. EX1002, ¶59. In addition, all of this must occur *without* first establishing an end-to-end connection between the sending and receiving devices. EX1002, ¶59.

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. EX1002, ¶60.

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

All challenged claims of the '516 patent effectively recite the E2E Limitation.

EX1002, ¶61. 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. EX1002, ¶61. As a result, 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.⁵ EX1002, ¶61.

1. Legal Standards

“It is elementary patent law that a patent application is entitled to the benefit of the filing date of an earlier filed application *only if* the disclosure of the earlier application provides support for the claims of the later application, as required by 35 U.S.C. § 112.” *PowerOasis, Inc. v. T-Mobile USA, Inc.*, 522 F.3d 1299, 1306 (Fed. Cir. 2008) (emphasis added). “To satisfy the written description requirement [in § 112,] the disclosure of the prior application must convey with reasonable clarity to those skilled in the art that, as of the filing date sought, the inventor was in possession of the invention.” *Id.* (cleaned up).

A disclosure that merely renders the claims **obvious** does not provide sufficient written description support. *See Ariad Pharm., Inc. v. Eli Lilly & Co.*, 598

⁵ 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.

F.3d 1336, 1352 (Fed. Cir. 2010) (en banc). “[T]he written description [must] actually or inherently disclose the claim element,” and the fact that a PHOSITA “*might* be able to construct [it] ... from the teachings of the disclosure” is insufficient. *PowerOasis*, 522 F.3d at 1306-07; *see D Three Enters., LLC v. SunModo Corp.*, 890 F.3d 1042, 1052 (Fed. Cir. 2018) (“[A]dequate written description does not ask what is permissible, rather, it asks what is disclosed.”).

The legal standard is substantially the same for negative limitations, like the E2E Limitation (“without first establishing”). *See Novartis Pharms. Corp. v. Accord Healthcare, Inc.*, 38 F.4th 1013, 1016-18 (Fed. Cir. 2022). For a negative limitation, “the written description requirement cannot be met through simple disregard of the presence or absence of a limitation.” *Id.* at 1017. But the 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 “distinguish[ing] among the element and alternatives to it” (*id.* at 1016), or (2) “inherently disclos[e] the negative limitation” such that it “would always be understood by skilled artisans as being necessarily excluded” (*id.* at 1017-18).

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

Before it was injected 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. EX1002, ¶62. The application immediately prior to the '240 application was App. Ser. No. 14/799,528, filed July 14, 2015. EX1006. 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); EX1010-1011 (provisional applications); EX1016 (Redline of '528 application as compared to '516 patent); EX1002, ¶62.

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. This limitation is not disclosed in the '528 Specification. EX1002, ¶63. *See Novartis*, 38 F.4th at 1016.

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. EX1002, ¶64. *See D Three Enters.*, 890 F.3d at 1052 ("[A]dequate written description does not ask what is permissible, rather, it asks

what is disclosed.”).

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. EX1002, ¶65. In fact, where the ’528 Specification discusses interaction between the first and second devices, it implies that an end-to-end connection is useful. EX1002, ¶65. 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); EX1002, ¶65. 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; EX1002, ¶65. 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). EX1002, ¶65.

Voxer may argue 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. EX1002, ¶66. Each such paragraph is directed to the situation where one of the devices is *unavailable* (i.e., the devices are not both online). EX1002, ¶66. Indeed, the purpose of these passages is to explain that, when a recipient is unavailable, a sender can still transmit a message. EX1002, ¶66.

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; EX1002, ¶66.

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, 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. EX1002, ¶67.

VIII. DETAILED EXPLANATION OF THE REASONS FOR UNPATENTABILITY

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

Claims 1-18 would have been obvious over Katis (EX1004) and Abbazia (EX1005). Both Katis and Abbazia are prior art under AIA 35 U.S.C. §§ 102(a)(1) and/or (a)(2) when the correct priority date of no earlier than August 8, 2016 is used (see §VII beginning on page 13, above). **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.

A. Overview of the Combination

Katis in view of Abbazia renders the challenged claims obvious.

Katis teaches the same system of the claims of the '516 patent, except that it

does not disclose the E2E Limitation, as described above. EX1002, ¶73.

Abbazia discloses the E2E Limitation in the context of a Multicast network for efficient distribution of video messages. 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). EX1002, ¶74.

1. Overview of Abbazia

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

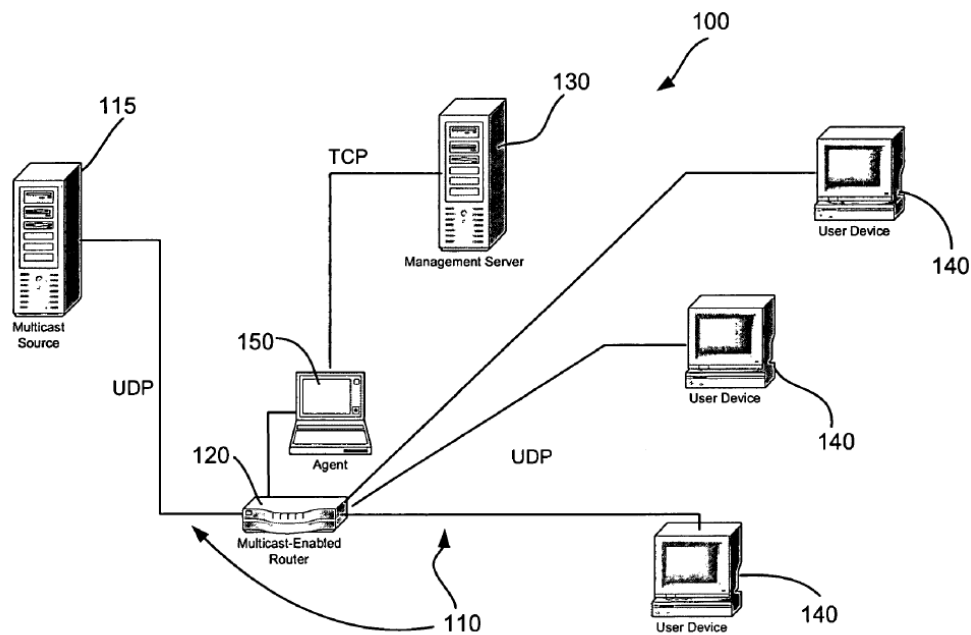


FIG. 1

EX1005, Fig. 1; EX1002, ¶75. 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; EX1002, ¶75.

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); EX1002, ¶75.

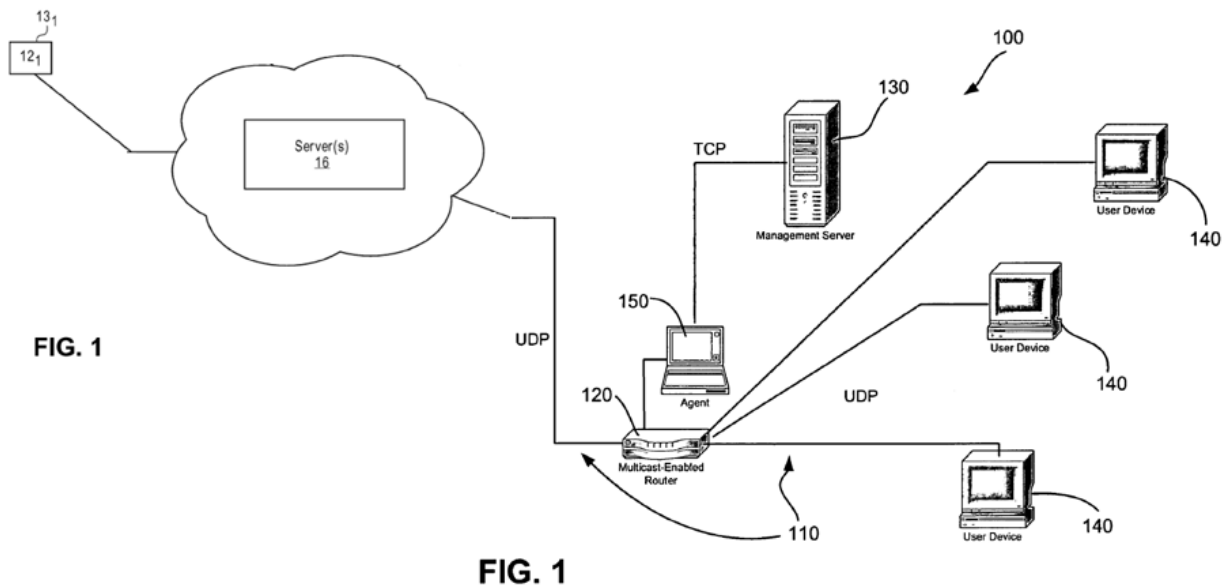
“Multicast” is a well-known Internet standard for efficient communications. EX1002, ¶76 (citing EX1022-1024). In Multicast, a Multicast sender sends a packet to a “Multicast address”, which is an IP address reserved for a particular Multicast group. EX1002, ¶76 (citing EX1022, p.1). When a packet addressed 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. EX1002, ¶76 (citing EX1022, p.1).

The Multicast method produces a significant gain in efficiency, because it allows part of the path to be traversed by a single packet. EX1002, ¶77. 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. EX1002, ¶77.

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; EX1002, ¶78.

2. The Combination

This Ground posits that 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. EX1002, ¶79. 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 12₁ of device 13₁ (left side) to other users 140 (right side):



EX1002, ¶79. In the modified figure, user devices 140 would be user devices of the Voxer system from Katis. EX1002, ¶79. The client 12₁ would send real-time video packets to the Voxer server 16, which forwards them to a Multicast address. EX1002, ¶79.

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. EX1002, ¶80. 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. EX1002, ¶80. 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). EX1002, ¶80. 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. EX1002, ¶80. Voxer users receiving real-time video from a Multicast address can send reply packets back to the Katis servers directly, with no change. EX1002, ¶80. 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. EX1002, ¶80; *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). EX1002, ¶80.

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; EX1002, ¶81.

B. Rationale (Motivation) Supporting Obviousness

A PHOSITA would have been motivated to add Multicast functionality, in the manner described above, to Katis. EX1002, ¶82. 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); EX1002, ¶82.

Abbazia notes that avoiding the need to form 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; EX1002, ¶83. 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. EX1002,

¶83 (citing EX1024, p.1; EX1025). 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; EX1002, ¶83 (citing EX1024, p.2). 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. EX1002, ¶83 (citing EX1024, p.2). 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. EX1002, ¶83. 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. EX1002, ¶83.

A PHOSITA would have seen additional advantages in using the Multicast system of Abbazia with Katis. EX1002, ¶84. 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; EX1002, ¶84. 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). EX1002, ¶84. 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. EX1002, ¶84 (citing EX1023). 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. EX1002, ¶84. 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; EX1002, ¶84. This notably tracks Abbazia's suggestion to use sequence numbers. EX1005, 6:65-7:1; EX1002, ¶84.

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. EX1002, ¶85. 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; EX1002, ¶85.

The combination would also have been obvious under *KSR Int'l Co. v. Teleflex, Inc.*, 550 U.S. 398, 417-21 (2007). Specifically, Katis teaches the known problems of developing systems that effectively use network bandwidth (especially for multimedia communications). EX1004, ¶¶0041, 0163, 0143, 0145; EX1002, ¶86. 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. EX1002, ¶86. In the relevant timeframe, Multicast was part of a known Internet standard and had been successfully used for video broadcasting. EX1002, ¶86 (citing EX1022-1024).

C. Reasonable Expectation of Success

A PHOSITA would have had a reasonable expectation of success in making the proposed combination. EX1002, ¶87. The art in the relevant timeframe was predictable. EX1002, ¶87. For example, 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; EX1002, ¶87; *see also supra* §VIII.A (describing the combination), §VIII.B (rationale supporting obviousness).

D. Analogous Art

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; EX1002, ¶88. 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; EX1002, ¶88. 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; EX1002, ¶88.

E. Claim Mapping

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:”

The combination renders obvious a **video message service infrastructure**. In the combination, the “**video message service infrastructure**” includes the Katis

“Clients 12₁ through 12_n communicate with one another and with Servers 16 over the networks 18₁ through 18_n and network 14 using individual message units referred to as ‘Vox packets’, which are described in detail below.”

EX1004, ¶0101; EX1002, ¶96.

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; EX1002, ¶97.

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. EX1002, ¶98.

“[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;”

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; EX1002, ¶99. In the combination, the “infrastructure” includes the Katis system as combined with the system of Abbazia, as shown above in §VIII.A.2 and discussed in element 1[Pre].

Katis teaches that user **sending devices 13 send** messages containing media to servers 16 (part of the **infrastructure**). EX1002, ¶100. 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; EX1002, ¶100.

Information establishing a conversation is provided in connection with an initial message from a first user. EX1002, ¶101. 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; EX1002, ¶101.

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; EX1002, ¶102. 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. EX1002, ¶102.

Katis also teaches or suggests that **video media of the video message is transmitted as the video media is created**. EX1002, ¶103. 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. EX1002, ¶103. 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); EX1002, ¶103.

“[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”

See analysis for claim element [1a]. EX1002, ¶104. 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. EX1002, ¶104. 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; EX1002, ¶104. In the combination, the video message is transmitted to the recipient device via the Multicast address. EX1002, ¶104.

“[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,”

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; EX1002, ¶105.

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**. EX1002, ¶106.

As explained above in §VIII.A and §VIII.B, it would have been obvious to combine Katis with Multicast functionality, as taught by Abbazia. EX1002, ¶107. 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...”); EX1002, ¶107. 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. EX1002, ¶107. 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**. EX1002, ¶107.

“[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,”

Katis states that media is segmented into indexed packets/payloads such that the **plurality of segments contain[] content of the video media**. EX1002, ¶108. 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...”); EX1002, ¶108.

“[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).”

The **plurality of segments are received at the video message service infrastructure while streamed by the sending device**. EX1002, ¶109. 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).”); EX1002, ¶109.

The packets (**plurality of segments**) are also **stored by the video message service infrastructure**. EX1002, ¶110. 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.”); EX1002, ¶110.

At least some of the segments are streamed to the recipient. EX1002, ¶111. See analysis for claim elements [1a] and [1b]. *See also, e.g.*, EX1004, ¶¶0196-0205; EX1002, ¶111.

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)**. EX1002, ¶112. See analysis for claim element [1b]. EX1004, ¶¶0174, 0019, 0204-0205, 0128, 0124; EX1002, ¶112.

No **end-to-end connection** is **establish[ed]** as discussed under claim element [1c]. EX1002, ¶113.

CLAIM 2

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

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; EX1002, ¶114. 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); EX1002, ¶114.

CLAIM 3

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

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.”). EX1002, ¶115.

CLAIM 4

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

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.”); EX1002, ¶116.

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.”

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; EX1002, ¶117. 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; EX1002, ¶117. 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; EX1002, ¶117.

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. EX1002, ¶118. 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; EX1002, ¶118. Katis states that “current bandwidth conditions” are taken into account. EX1004, ¶0055; EX1002, ¶118.

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.”

This limitation is taught or suggested by Katis. EX1004, ¶¶0041, 0043, 0055, 0143, 0145, 0207, 0218, 0221-0223, 0228, 0236, Fig. 9C; EX1002, ¶119. 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; EX1002, ¶119. Katis states that “current bandwidth conditions” are taken into account. EX1004, ¶0055. See also claim 5 analysis; EX1002, ¶119.

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.”

See claims 5 and 6 analysis. EX1002, ¶120. 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; EX1002, ¶120.

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.”

See claims 5, 6, and 7 analysis. EX1002, ¶121. 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.”); EX1011,⁶ pp. 43-46; EX1002, ¶121; *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

⁶ Katis incorporates by reference U.S. Provisional Application No. 60/999,619 (contained in EX1011). EX1004, ¶0001.

the network.”

See claim elements 1[Pre] and [1a] analysis. EX1002, ¶122. *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.”). EX1002, ¶122.

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 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.”

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; EX1002, ¶123. 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.”); EX1002,

¶123. See also analysis for claim element [1b]. EX1002, ¶123.

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;”

See analysis for claim element [1a]. EX1002, ¶124. Katis teaches allowing the various users to contribute one or more additional messages to a messaging conversation. EX1002, ¶124. 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); EX1002, ¶124. *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; EX1002, ¶124.

“[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;”

See analysis for claim element [1b]. EX1002, ¶125. 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; EX1002, ¶125. 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); EX1002, ¶125.

“[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

See analysis for claim element [1c]. EX1002, ¶126.

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

See analysis for claim element [11b]. EX1002, ¶127. See also EX1004, ¶¶0065, 0080-0081, 0108, 0124, 0128, 0146, 0161, 0183, 0202, 0204-0205, 0299, 0303-0304, 0309-0311, Fig. 13A; EX1002, ¶127. 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; EX1002, ¶127.

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.”

See analysis for claim element [1a]. EX1002, ¶128. Katis teaches or suggests this claim element. EX1002, ¶128. 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); EX1002, ¶128.

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.”

See analysis for claims 1, 10, and 12. EX1002, ¶129. 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; EX1002, ¶129. 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; EX1002, ¶129. See also analysis for claim element [1b]. EX1002, ¶129.

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.”

This limitation is taught or suggested by the Katis and is implemented in the combination. EX1004, ¶¶0040, 0047, 0122, 0142, 0145, 0174 0191; EX1002, ¶130. For example, Katis states that “to initiate a conversation, one can simply begin transmitting to an individual or a group.” EX1004, ¶0040; EX1002, ¶130. 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. EX1002, ¶130. 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. EX1002, ¶130. 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; EX1002, ¶130.

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.”

See analysis for claim 14. EX1002, ¶131. See also EX1004, ¶0047

(discussing beginning a conversation when “other participants are unavailable”);
EX1002, ¶131.

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.”

Katis teaches a **notification** can be provided to each participant when that user’s device receives a message. EX1004, ¶¶0107, 0038, 0268, 0137; EX1002, ¶132. 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; EX1002, ¶132.

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.”

This limitation is taught or suggested by Katis and obvious in view of the combination. EX1004, ¶¶0063, 0073, 0179, 0185, 0190, 0191, 0132, 0105; EX1002, ¶133. 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; EX1002, ¶133. 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; EX1011,⁷ p.39 (“2. Payload Data a. Conversation Identifier b. Message Identifier”), p.44 (“Data Elements ... 3. Recipient Data b. Identifier (transmission target)”); EX1002, ¶133. When the user sends a message, it includes this identifier

⁷ Katis incorporates by reference U.S. Provisional Application No. 60/999,619 (contained in EX1011). EX1004, ¶0001.

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). EX1002, ¶133. 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. EX1002, ¶133. 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. EX1002, ¶133.

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:”

See analysis for claim elements 1[Pre] and [1a] and claim 9. EX1002, ¶134.

“[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.”

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**). EX1002, ¶135.

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**. EX1002, ¶136. The combination infrastructure ascertains recipient devices by their IP addresses. EX1004, ¶¶0191, 0061, 0115; EX1005, 1:55-60, 5:30-34; EX1002, ¶136. 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. EX1002, ¶136. Both a single device’s IP address, and a Multicast IP address, are addresses “for” a recipient device. EX1002, ¶136. In addition, the multicast routers then route multicast packets to each client device which is part of the multicast group. EX1002, ¶136.

IX. SECONDARY CONSIDERATIONS

In *ex parte* reexaminations of the ’557 and ’270 patents filed by Meta, an unrelated third-party (*see* §I.B (notice of related matters)), Voxer argued that there

were objective indicia of non-obviousness. The evidence submitted by Voxer in these reexaminations, however, does not undercut Petitioner's showing here. First, there is no nexus between that evidence and the claims of the '516 patent. *See Fox Factory, Inc. v. SRAM, LLC*, 944 F.3d 1366, 1373 (Fed. Cir. 2019) (explaining that "the evidence of secondary considerations must have a 'nexus' to the claims"). Indeed, the *ex parte* reexaminations involved different patents with different claimed subject matter. Also, Petitioner respectfully requests an opportunity to respond to any argument such a nexus exists. *See id.* ("The patentee bears the burden of showing that a nexus exists."). Second, as explained above, Katis is prior art to the '516 patent, and it discloses any alleged merits of the claimed invention, rendering any evidence of secondary considerations ineffective. EX1002, ¶137; *see ClassCo, Inc. v. Apple, Inc.*, 838 F.3d 1214, 1220 (Fed. Cir. 2016) ("A nexus may not exist where, for example, the merits of the claimed invention were 'readily available in the prior art.'"). Finally, the examiner did not rely on Voxer's purported evidence when confirming the patentability of the claims in the *ex parte* reexaminations, and Voxer did not argue secondary considerations during its trial with Meta.

X. CONCLUSION

Petitioner respectfully requests that claims 1-18 of the '516 patent be canceled.

Date: <u>July 13, 2026</u>	<u>/Matthew A. Smith/ (RN 49,003)</u> Matthew A. Smith SMITH BALUCH LLP 700 Pennsylvania Ave SE Ste. 2060 Washington, DC 20003 (202) 669-6207 smith@smithbaluch.com <i>Counsel for Petitioner Google LLC and YouTube, LLC</i>
----------------------------	--

CERTIFICATE OF SERVICE

The undersigned hereby certifies that the foregoing petition for *inter partes* review, together with all exhibits and other documents filed therewith, was served by Priority Mail Express on the date below, on the Patent Owner's counsel of record at the United States Patent & Trademark Office having the following address:

**76743 - BEYER LAW GROUP LLP / VOXER IP
P. O. Box 51887
Palo Alto, CA
UNITED STATES**

Date: July 13, 2026

/Matthew A. Smith/ (RN 49,003)

CERTIFICATE OF WORD COUNT

The undersigned hereby certifies that the foregoing petition for *inter partes* review contains 10,740 words according to the word processing program used to prepare it.

Date: July 13, 2026

/Matthew A. Smith/ (RN 49,003)