

UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE PATENT TRIAL AND APPEAL BOARD

ZOHO CORPORATION and
ZOHO CORPORATION PVT., LTD.
Petitioner,

v.

MEETRIX IP, LLC,
Patent Owner.

Case IPR2022-00377
U.S. Patent No. 9,253,332

PETITION FOR *INTER PARTES* REVIEW OF
U.S. PATENT NO. 9,253,332

TABLE OF CONTENTS

I.	INTRODUCTION.....	1
II.	MANDATORY NOTICES UNDER 37 C.F.R § 42.8(a)(1) AND (B)	2
	A. Real Party-In-Interest Under 37 C.F.R. § 42.8(b)(1).....	2
	B. Lead and Back-Up Counsel and Service Information Under 37 C.F.R. § 42.8(b)(3) and (b)(4)	2
	C. Related matters Per 37 C.F.R. § 42.8(b)(2)	3
III.	PAYMENT OF FEES – 37 C.F.R. § 42.103	3
IV.	GROUND FOR STANDING UNDER 37 C.F.R. § 42.104(a)	3
V.	CHALLENGE UNDER 37 C.F.R. § 42.104(B) AND RELIEF REQUESTED.....	3
VI.	OVERVIEW OF THE '332 PATENT	5
	A. The '332 Patent Specification.....	5
	B. The Challenged Claims.....	8
	C. Level of Ordinary Skill in the Art.....	9
	D. Claim Construction	9
	E. The '332 Patent Priority Date	10
VII.	THE BOARD'S DISCRETION UNDER 35 U.S.C. § 314(A) SHOULD NOT PRECLUDE INSTITUTION.....	10
VIII.	IDENTIFICATION OF WHERE EACH ELEMENT OF EACH CLAIM IS FOUND IN THE PRIOR ART	14
	A. Ground 1: Claims 1-2, 4-5 and 7 Are Obvious Over Knappe in View of Elliott and the VPN Textbook	14
	B. Ground 2: Claims 3, 6, and 8-11 Are Obvious Over Knappe in View Elliott, the VPN Textbook, and Drell.....	50
	C. Ground 3: Claim 12 is Obvious Over Knappe in Combination with Elliott, VPN Textbook, Drell, and Hoke	77
IX.	CONCLUSION	81

TABLE OF AUTHORITIES

Cases

<i>In re Etter</i> , 756 F.2d 852 (Fed. Cir. 1985) (en banc)	25
<i>In re Mouttet</i> , 686 F.3d 1322 (Fed. Cir. 2012)	25, 27, 56, 79
<i>Intel Corp. v. VLSI Tech. LLC</i> , IPR2020-00158, Paper 16 (May 20, 2020).....	12
<i>Phillips v. AWH Corp.</i> , 415 F.3d 1303 (Fed. Cir. 2005) (en banc)	9, 10
<i>Sand Revolution II, LLC v. Continental Intermodal Group-Trucking LLC</i> , IPR2019-01393, Paper 24, 7 (June 16, 2020).....	12, 13
<i>Sotera Wireless, Inc. v. Masimo Corp.</i> , IPR2020-01019, Paper 12 (PTAB Dec 1, 2020).....	13
<i>Vivid Techs., Inc. v. Am. Sci. & Eng’g, Inc.</i> , 200 F.2d 795 (Fed. Cir. 1999)	10
<i>VMware, Inc. v. Intellectual Ventures I LLC</i> , IPR2020-00470, Paper 13 at 20 (PTAB Aug. 18, 2020).....	13

Statutes

35 U.S.C. §§ 311, 312, 314(a), ...	1
--------------------------------------	---

Rules

37 C.F.R. § 41.10(b)	2
37 C.F.R. § 42.104(a)	3
37 C.F.R. § 42.15(a)	3
37 C.F.R. § 42.6(c)	3

Regulations

37 C.F.R. § 42.100.....	1
83 Fed. Reg. 51341 (October 11, 2018)	9

TABLE OF EXHIBITS

Ex.	Description
1001	U.S. Patent No. 9,253,332 (“the ’332 patent”)
1002	Patent Assignments for the ’332 patent
1003	Declaration of Dr. Henry H. Houh
1004	<i>Curriculum Vitae</i> for Dr. Henry H. Houh
1005	File History of Application No. 13/674,227 (the ’227 application)
1006	U.S. Patent No. 7,180,997 (“Knappe”)
1007	U.S. Patent No. 6,690,654 (“Elliott”)
1008	U.S. Patent No. 7,089,285 (“Drell”)
1009	U.S. Patent No. 6,701,437 (“Hoke”)
1010	Guichard <i>et al.</i> , “MPLS and VPN Architectures, A practical guide to understanding, designing and deploying MPLS and MPLS-enabled VPNs,” Cisco Press, 2001 (“VPN Textbook”)
1011	Declaration of Ingrid Hsieh-Yee, Ph.D., under 37 C.F.R. § 1.68
1012	U.S. Patent No. 3,060,265
1013	U.S. Patent No. 4,139,731
1014	U.S. Patent No. 3,601,530
1015	U.S. Patent No. 4,796,293
1016	A Primer on the T.120 Series Standard
1017	T.120 Standard (1996)
1018	U.S. Patent No. 6,237,025
1019	U.S. Patent No. 6,782,413

Ex.	Description
1020	U.S. Patent No. 6,980,526
1021	Caputo, “Cisco Packetized Voice & Data Integration,” McGraw-Hill, 2000
1022	De Moraes <i>et al.</i> , “The Internet Multicast from ITS: How it was Done and Implications for the Future,” IEEE Communications Magazine, January 1995
1023	U.S. Patent Publication No. 2004/0261116 A1
1024	U.S. Patent No. 6,850,496
1025	U.S. Patent No. 6,553,515
1026	U.S. Patent No. 6,633,848
1027	Computer Desktop Encyclopedia
1046	Declaration of C.J. Alice Chuang

I. INTRODUCTION

Zoho Corporation and Zoho Corporation Pvt., Ltd. (“Petitioner”) respectfully requests *inter partes* review (“IPR”) under 35 U.S.C. §§ 311, 312, 314(a), and 37 C.F.R. § 42.100, for claims 1-12 (“Challenged Claims”) of U.S. Patent No. 9,253,332 (“’332 patent,” Ex_1001). Meetrix IP, LLC (“Patent Owner”) is the owner of record. Ex_1002 (conveying the ’332 patent family to Meetrix).

The ’332 patent is directed to conferencing technology, including “audio and video telecommunications for collaboration over hybrid networks,” where audio data from multiple participants are mixed together and transmitted to conference participants. Ex_1001, Abstract, 1:18-19. The hybrid networks include a Public Switched Telephone Network (PSTN) and the Internet. *Id.*, 4:39-53, FIG. 3. As described below and in the Declaration of Dr. Henry H. Houh (Ex_1003), these concepts were well known by the claimed priority date for the ’332 patent. The primary reference (“Knappe,” Ex_1006) describes a multiparty conferencing system, and teaches receiving and mixing audio data from multiple participants, including a moderator and PSTN telephones. Ex_1006, 1:8-27, 3:41-49, 4:18-28, 6:62-67, 8:52-57. Other claimed concepts, including using virtual private network (VPN) tunneling and voice over IP (VoIP) for conferencing communications, were also well known in the prior art. The combination of these prior art teachings

would have been obvious to one of ordinary skill in the art, and thus the claims of the '332 patent are invalid.

II. MANDATORY NOTICES UNDER 37 C.F.R § 42.8(A)(1) AND (B)

A. Real Party-In-Interest Under 37 C.F.R. § 42.8(b)(1)

The real parties-in-interest are Zoho Corporation and Zoho Corporation Pvt., Ltd.

B. Lead and Back-Up Counsel and Service Information Under 37 C.F.R. § 42.8(b)(3) and (b)(4)

Lead Counsel:

Hector J. Ribera (Reg. No. 54,397)
Marton Ribera Schumann & Chang LLP
548 Market St., Suite 36117
San Francisco, CA 94104
Email: hector@martonribera.com
Tel: (415) 360-2512

Back-up Counsel:

C.J. Alice Chuang (Reg. No. 65,914)
Marton Ribera Schumann & Chang LLP
548 Market St., Suite 36117
San Francisco, CA 94104
Email: cjalice@martonribera.com
Tel: (415) 360-2516

Petitioner consents to electronic service by delivering the documents to the email addresses of primary and backup counsel. A Power of Attorney is being filed concurrently under 37 C.F.R. § 41.10(b).

C. Related matters Per 37 C.F.R. § 42.8(b)(2)

To the best of Petitioner's knowledge, the '332 patent is involved in the following pending district court cases:

Case Name	Case Number	Court	Date Filed
<i>Meetrix IP, LLC v. Zoho Corporation</i>	6:21-cv-01288	W.D. Tex.	12/10/21
<i>Meetrix IP, LLC v. Verizon Comms., Inc.</i>	6:21-cv-01289	W.D. Tex.	12/10/21

III. PAYMENT OF FEES – 37 C.F.R. § 42.103

Petitioner pays the fee for this petition as required by 37 C.F.R. § 42.15(a) contemporaneously with the filing of this petition.

IV. GROUND FOR STANDING UNDER 37 C.F.R. § 42.104(A)

Petitioner certifies under 37 C.F.R. § 42.104(a) that the '332 patent is available for IPR and that Petitioner is not barred or estopped from requesting an IPR challenging the validity of the challenged claims on the identified grounds. Petitioner has not filed a civil action challenging the validity of any claim of the '332 patent.

V. CHALLENGE UNDER 37 C.F.R. § 42.104(B) AND RELIEF REQUESTED

Petitioner requests IPR and cancellation of claims 1-12 of the '332 patent under pre-AIA 35 U.S.C. § 103. Pursuant to 37 C.F.R. § 42.6(c), copies of the Exhibits are filed herewith.

Under 37 C.F.R. §§ 42.104(b)(1) and (2), these grounds are offered as reasons for cancelling claims 1-12 of the '332 patent:

Ground	Claims	Statutory Basis	Reference(s)
1	1-2, 4-5 and 7	§ 103	Knappe, Elliott, and VPN Textbook
2	3, 6, and 8-11	§ 103	Knappe, Elliott, VPN Textbook, and Drell
3	12	§ 103	Knappe, Elliott, VPN Textbook, Drell, and Hoke

U.S. Patent No. 7,180,997 to Knappe (Ex_1006) was filed on September 6, 2002.

U.S. Patent No. 6,690,654 to Elliott, *et al.* (“Elliott” – Ex_1007) was filed on October 1, 1998.

U.S. Patent No. 7,089,285 to Drell (“Drell” – Ex_1008) was filed on October 5, 2000.

U.S. Patent No. 6,701,437 to Hoke (“Hoke” – Ex_1009) was filed on November 9, 1998.

“MPLS and VPN Architectures,” (“VPN Textbook” – Ex_1010) by Guichard *et al.*, was published in 2001. *See* Declaration of Ingrid Hsieh-Yee, (“Yee Decl.” - Ex_1011), ¶¶16, 23, 95-111.

Knappe, Elliott, Hoke, and Drell are prior art under 35 U.S.C. § 102(e). The VPN Textbook is prior art under 35 U.S.C. § 102(b).

VI. OVERVIEW OF THE '332 PATENT

The '332 patent was filed on Nov. 12, 2012, as a continuation of a string of applications, claiming priority to a provisional application filed March 10, 2003.

Ex_1001; Ex_1003 ¶48.

A. The '332 Patent Specification

The '332 patent is directed to a “system and method for supporting a multi-participant voice conference call using PSTN and Internet networks.” Ex_1001, Abstract. The method includes receiving voice data from a PSTN client, a moderator, and at least one remote client connected to the Internet. *Id.*

The specification describes using a hybrid network that connects several Internet-based audio video conference participants (including a moderator) and one or more PSTN participants. Ex_1001, FIG. 3, 4:39-5:18; Ex_1003 ¶50. The hybrid network is depicted in Figure 3 of the patent:

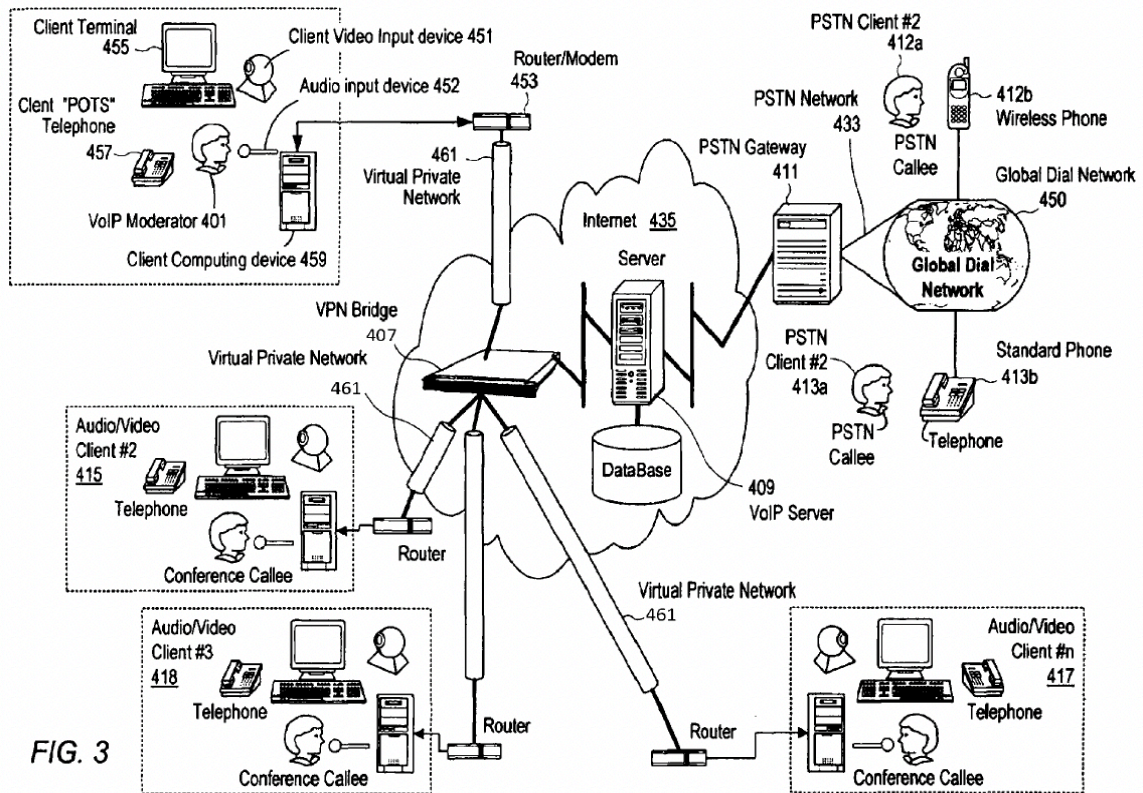


FIG. 3

Ex_1001, FIG. 3; Ex_1003 ¶51.

Specifically, Figure 3 shows “the Internet-based clients 401, 415, 418, and 417 are connected through routers or modems 453 preferably in a virtual private network configuration 461” with “the VPN bridge 407 located in the center of a “star topology,” with connections to the VPN bridge serving as a “VPN tunnel.”

Ex_1001, 5:7-32, 6:20-42, FIGs. 3-5; Ex_1003 ¶52.

The PSTN network includes “telephony clients using standard wired 413 or wireless telephone 412 systems,” with the PSTN clients connected to other clients via PSTN gateway 411. Ex_1001, 4:67-5:1; Ex_1003 ¶53. A connection from the VPN bridge 407 to the a “voice over IP server 409 is used to transfer conference

audio from any participant on the IP network to any participant in the PSTN.”

Ex_1001, 5:10-14. “The voice over IP server 409 is responsible for transcoding audio information from the virtual private network 461 to and from the PSTN gateway 411, thus bridging the PSTN and VPN together. *Id.*, 5:15-18. The ’332 patent discloses that “voice over IP server 409 may use standard H.323 or SIP network protocol to establish communications” between PSTN and Internet clients. *Id.*, 7:47-50; Ex_1003 ¶54.

The specification also discloses performing encoding and decoding of audio data using codecs of the H.323 standard, as well as transmission of collaboration data per the T.120 standard. Ex_1001, 2:55-61, 3:10-12, FIG. 29; Ex_1003 ¶ 55. The specification also describes mixing audio received from the PSTN with audio from the moderator using audio encoder 520a local to the moderator and then delivering “both the PSTN client’s audio and the local moderator client’s audio to remote audio video clients 418.” Ex_1001, 8:60-67; Ex_1003 ¶55.

The patent further discusses the process of combining audio from the moderator with audio from remote clients. Ex_1001, 9:17-20; Ex_1003 ¶56. The mixed audio data is then “placed in virtual private network tunnel for transport to the voice over IP server 409 and onto the gateway for audio presentation to the PSTN[.]” Ex_1001, 9:20-23. The specification also explains that “local moderator client 401 receives audio from PSTN client 412 . . . and mixes audio previously

decoded by the voice over IP decoder 524 with the audio from the remote client's 418 for presentation to the local speaker 454.” *Id.*, 9:4-8; Ex_1003 ¶57.

B. The Challenged Claims

Petitioner requests IPR of claims 1-12. The '332 patent has two independent claims. Independent claim 1 is directed to a method for supporting a multi-participant audio/video conference call, the method comprising:

- receiving first audio data from a PSTN client;
- receiving second audio data from a moderator;
- receiving third audio data, video data, and collaboration data from at least one remote client through a first VPN;
- mixing the first audio data from the PSTN client with the second audio data from the moderator into a first mixed audio data;
- transmitting the first mixed audio data to the remote client through the first VPN tunnel;
- mixing the second audio data from the moderator with the third audio data from the remote client into a second mixed audio data; and
- transmitting the second mixed audio data to the PSTN client.

Ex_1001, 9:52-10:5; Ex_1003 ¶58. Claim 2 further adds the mixing of “the first audio data from the PSTN client with the third audio data from the remote client into a third mixed audio data.” Ex_1001, 10:6-10; Ex_1003 ¶59. Claims 3-7 add further known limitations, such as using IP packets and compressing audio data, to the method steps in claim 1, including “VoIP encoding the second mixed audio data before transmitting the second mixed audio data to the PSTN client.”

Ex_1001, 10:11-29; Ex_1003 ¶60. Independent claim 8 is a system claim requiring mixers, transport outputs, and a VoIP encoder. Ex_1001, 10:30-46; Ex_1003 ¶61. Claims 9-12 add an additional mixer, a speaker, an encoder and decoder, and an audio decompressor to the claimed system. Ex_1001, 10:47-63; Ex_1003 ¶62.

C. Level of Ordinary Skill in the Art

The '332 patent generally relates audio/video conferencing systems. A Person of Ordinary Skill in the Art (“POSITA”) in March 2003 would have had working knowledge of the available audio and video conferencing art. Ex_1003 ¶27. A POSITA would have had a bachelor’s degree in computer science, computer engineering, or an equivalent, and three or more years of professional experience relating to conferencing systems in packet-based networks, or without said professional experience, further education relating to conferencing systems in packet-based networks. *Id.* ¶¶24-28.

D. Claim Construction

During IPR claims are construed according to the standard set forth in *Phillips v. AWH Corp.*, 415 F.3d 1303 (Fed. Cir. 2005) (*en banc*). *See* 83 Fed. Reg. 51341 (October 11, 2018). Petitioner believes that, for purposes of this proceeding, no claim term requires express construction for the Board to evaluate

the patentability of the claims.¹ *Vivid Techs., Inc. v. Am. Sci. & Eng'g, Inc.*, 200 F.2d 795, 803 (Fed. Cir. 1999). Accordingly, this Petition analyzes the claims consistent with their ordinary and customary meaning as would be understood by a POSITA in light of the specification. *Phillips*, 415 F.3d at 1314-17; Ex_1003 ¶31.

E. The '332 Patent Priority Date

The '332 patent issued from U.S. Patent Application No. 13/674,227 (the '227 application), filed November 12, 2012. Ex_1001, 1. The '227 application is a continuation of a series of applications that claim the benefit of U.S. Provisional Patent Application No. 60/453,307, filed March 10, 2003. *Id.* For this Petition, Petitioner relies on March 10, 2003, as the priority date for the Challenged Claims.

VII. THE BOARD'S DISCRETION UNDER 35 U.S.C. § 314(A) SHOULD NOT PRECLUDE INSTITUTION

On June 21, 2022, the Director of the USPTO issued a memorandum setting forth interim procedures for determining institution of IPRs in view of the factors set forth in *Apple Inc. v. Fintiv, Inc.* See https://www.uspto.gov/sites/default/files/documents/interim_proc_discretionary_denials_aia_parallel_district_court_litigation_memo_20220621_.pdf

¹ Petitioner reserves its rights to: (1) respond to assertions by Patent Owner that any claim term requires construction for the purposes of this IPR proceeding; and (2) seek construction of any claim term in other forums as appropriate.

(“Memorandum”). The Memorandum clarified (1) that the PTAB will not rely on the *Fintiv* factors to discretionarily deny institution where a petition presents compelling evidence of patentability, and (2) consistent with *Sotera Wireless*, the PTAB will not discretionarily deny institution where a petitioner presents a stipulation not to pursue in a parallel proceeding the same grounds or any grounds that could have been raised before the PTAB. In light of both these factors, the Board should not exercise its discretion to deny institution here.

First, this Petition presents compelling evidence of unpatentability. As set forth below in this Petition and in the accompany declaration of Dr. Houh (Ex_1003), the subject matter of the ’332 patent relates to well-known concepts of multi-party conferencing between participants in hybrid Internet and PSTN networks. By the ’332 patent’s claimed priority date, there were many available multi-party conferencing systems and a POSITA was well aware of how to use known communications technologies such as audio mixing, audio codecs, VPN tunneling, and IP packets to communicate conferencing data across the hybrid networks. *See e.g.* Ex_1003 ¶¶63-112. Thus, this is not a case in which Petitioner barely meets the IPR standards with obscure technical references. Here, there is compelling evidence that the claims of the ’332 patent merely recite well-established multi-party conferencing technology.

Second, upon the filing of this Petition, Petitioner intends to file in the district court litigation a request for pre-institution stay of proceedings with a stipulation that, if the IPR is instituted, Petition will not rely on any grounds that is raised or could have been in this Petition.

Finally, the remaining *Fintiv* factors weigh against denial:

Fintiv factor 1, likelihood of a stay. As noted, Petition is seeking a pre-institution stay and if needed, will renew the request for a stay upon institution. Even if Patent Owner suggests that district court litigation may not be stayed, the Board “will not attempt to predict how the district court in the related district court litigation will proceed because the court may determine whether or not to stay any individual case, including the related one, based on a variety of circumstances and facts beyond [its] control and to which the Board is not privy.” *Sand Revolution II, LLC v. Continental Intermodal Group-Trucking LLC*, IPR2019-01393, Paper 24, 7 (June 16, 2020) (Informative); *see Intel Corp. v. VLSI Tech. LLC*, IPR2020-00158, Paper 16, *7 (May 20, 2020).

Fintiv factor 2, trial date proximity to Final Written Decision. No trial date has been set in the district court litigation. The district court litigation is still in the early stages, having been transferred between divisions in the W.D. Texas and stayed for several months pending re-assignment. Given the parties have yet to initiate discovery or conduct claim construction, it is likely a Final

Written Decision would issue far in advance of any potential trial date in the district court litigation, particularly as Petitioner will be seeking a stay.

Fintiv factors 3 and 4, investment in the parallel proceedings and overlap in issues.

As of the filing this Petition, the district court remains in the early stages. Although the case was originally filed in December 2021, early procedural motions and an intra-district transfer and stay of the case means that there has been no fact discovery, no substantive motions, and no claim construction proceedings. In short, no substantive action has been taken in the district court and very little court and party investment has been made in that proceeding. *Sand Revolution*, IPR2019-01393, Paper 24 at 10-11 (two-page *Markman* Order did not demonstrate high level of investment of time and resources).

Critically, the invalidity grounds, arguments, and evidence presented in this Petition will be different than those presented in the district court. As noted above, Petitioner will stipulate in the Litigation that if the Board institutes IPR, it will not pursue invalidity on the same grounds at issue in this Petition or any grounds that could have been raised in this Petition. This mitigates concerns of duplicated efforts and conflicting decisions. *See Sotera Wireless, Inc. v. Masimo Corp.*, IPR2020-01019, Paper 12 (PTAB Dec 1, 2020); *see also VMware, Inc. v. Intellectual Ventures I LLC*, IPR2020-00470, Paper 13 at 20 (PTAB Aug. 18, 2020) (finding similar stipulation “mitigates any concerns of duplicative efforts

between the district court and the Board, as well as concerns of potentially conflicting decisions”). As such, these factors weigh against discretionary denial.

VIII. IDENTIFICATION OF WHERE EACH ELEMENT OF EACH CLAIM IS FOUND IN THE PRIOR ART

Petitioner respectfully requests cancellation of claims 1-12 of the '332 patent.

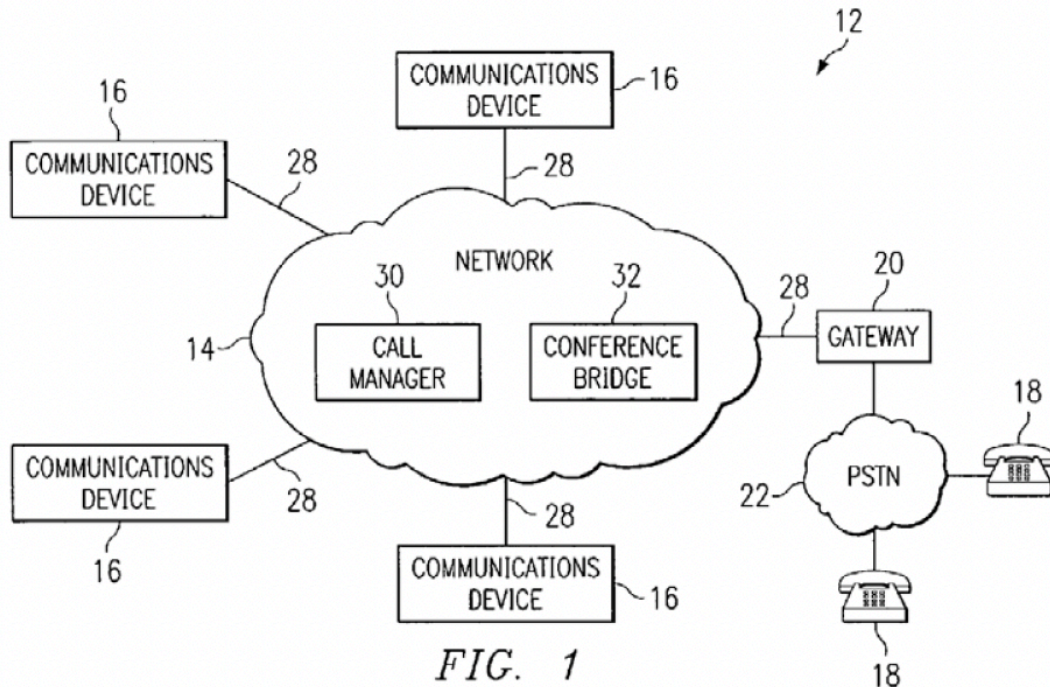
A. Ground 1: Claims 1-2, 4-5 and 7 Are Obvious Over Knappe in View of Elliott and the VPN Textbook

1. Summary of Prior Art

a) Summary of Knappe

Knappe discloses a multi-participant conferencing system using PSTN and Internet networks. Ex_1003 ¶118. Specifically, Knappe discloses a “a multi-party communication session” with “a plurality of participant voice streams from a plurality of respective conference participants.” Ex_1006, Abstract, 1:8-11, 8:46-10:50, FIGS. 1-8; Ex_1003 ¶118.

As shown in Figure 1 below, the multi-participant conferencing system in Knappe connects multiple network participants (communication devices 16) and PSTN telephone participants (PSTN telephones 18):



Ex_1006, FIG. 1; *see also, id.*, 3:41-4:28; Ex_1003 ¶119. Knappe explains that in one embodiment “network 14 is the Internet . . . or other suitable packet-switched network” and “transmits information in Internet (IP) packets” with telephony voice information “transmitted in the Voice over IP (VoIP) format.” Ex_1006, 3:50-56; Ex_1003 ¶120. Specifically, Knappe explains:

codec (coder/decoder) converts audio, video or other suitable signals generated by users from analog signals into digital form. The digital encoded data is encapsulated into IP or other suitable packets for transmission over the network 14. IP packets received from the network 14 are converted back into analog signals and played to the user.

Ex_1006, 4:5-14; Ex_1003 ¶121. Knappe further explains that “at least one participant may be given the priority of moderator for a particular conference call.”

Ex_1006, 8:52-56; Ex_1003 ¶121.

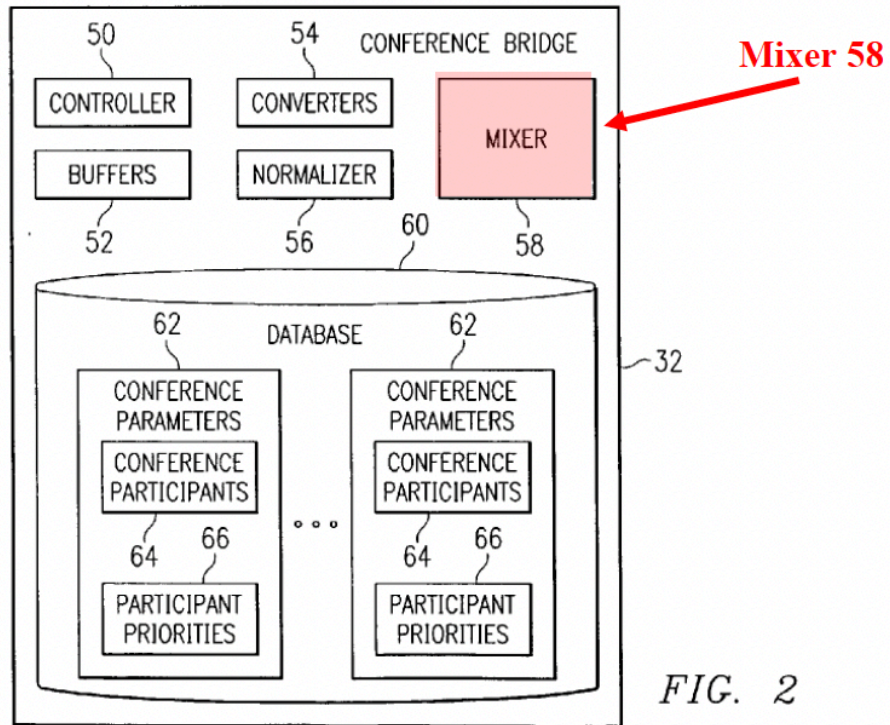
The standard analog telephones 18 participating in the conferencing system “communicate standard telephony signals through PSTN 22 to the gateway 20.”

Ex_1001, 4:19-21; Ex_1003 ¶122. Knappe explains that:

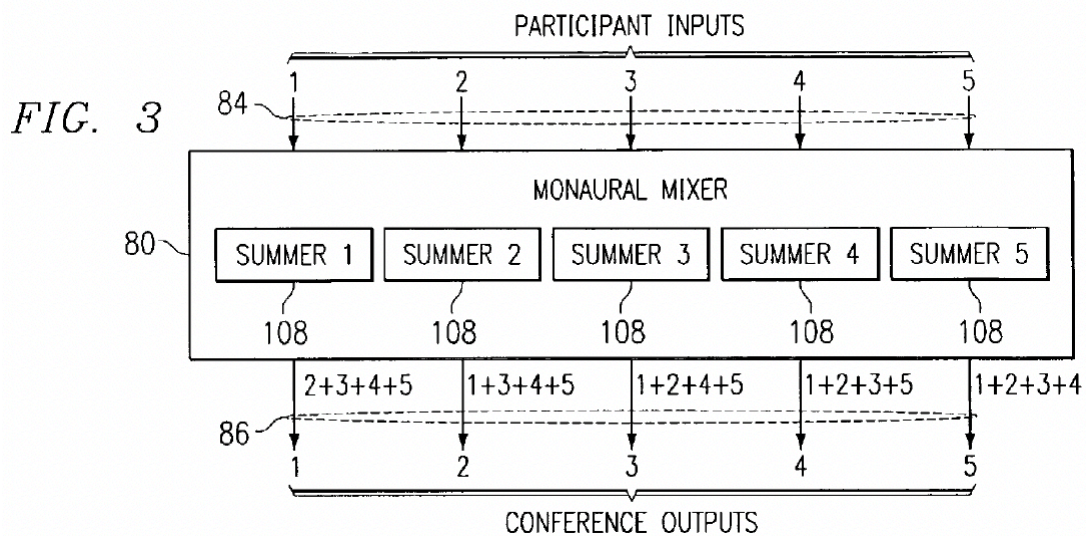
[a]t the gateway 20, the signals are converted to IP packets in the VoIP format. Similarly, VoIP packets received from the network 14 are converted into standard telephony signals for delivery to the destination telephone 18 through PSTN 22. The gateway 20 also translates between the network call control system and the Signaling System 7 (SS7) protocol and/or other signaling protocols used in PSTN.

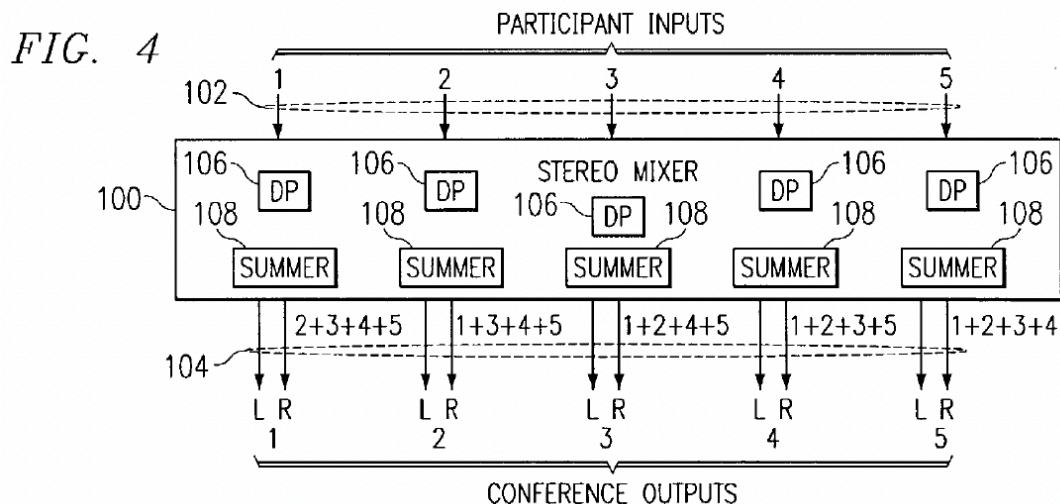
Ex_1006, 9:21-28; Ex_1003 ¶122.

In Knappe, the “participants are connected and stream media through the conference bridge 32.” Ex_1006, 5:32-33; Ex_1003 ¶123. As illustrated in Figure 2 below, Knappe’s conference bridge 32 employs mixer 58, which “includes a plurality of summers or other suitable signal processing resources each operable to sum, add or otherwise combine a plurality of input streams into conference output streams for participants to a conference call.” Ex_1006, 6:33-37.



Id., FIG. 2 (annotated); Ex_1003 ¶123. Figures 3-4 in Knappe show examples of mixers that “combine[] the streams in summers 82 to generate conference output streams 86 for each participant to a conference call.” Ex_1006, 6:60-62 (monaural mixer), 7:36-39 (stereo mixer).





Id., FIGS. 3-4; Ex_1003 ¶124.

The combined data is then transmitted to each of the other participants:

“[d]uring normal operation, each participant receives the audio input of each other participant.” Ex_1006, 7:1-15; Ex_1003 ¶125. The communications devices 16 also “comprise real-time applications that play traffic as it is received.” Ex_1006, 4:5-9; Ex_1003 ¶121.

b) Summary of Elliott

Elliott also discloses a conferencing system that allows for multi-media collaboration between multiple remote participants, including Internet-based and PSTN clients. Ex_1007, Abstract & FIG. 1; Ex_1003 ¶127. Elliott teaches “collaborative Web browsing, audio conferencing, video conferencing, and application sharing.” Ex_1007, 1:40-44; Ex_1003 ¶128. Elliott further teaches the use of “a virtual private network (VPN) ... which uses IP or IP-tunneling” for the

transmission of “voice, fax, data, video” and “[o]ther types of collaboration ... data” during the conference. Ex_1007, 2:66-3:1, 4:29-34, 6:53-54; Ex_1003 ¶129.

Figure 1 of Elliott shows a “session server 12 [that] is a conference server that follows the ITU T.120 standard” which “contains a series of communication and application protocols and services that provide support for real-time, multi-point data communications.” Ex_1007, FIG. 1, 5:49-54.

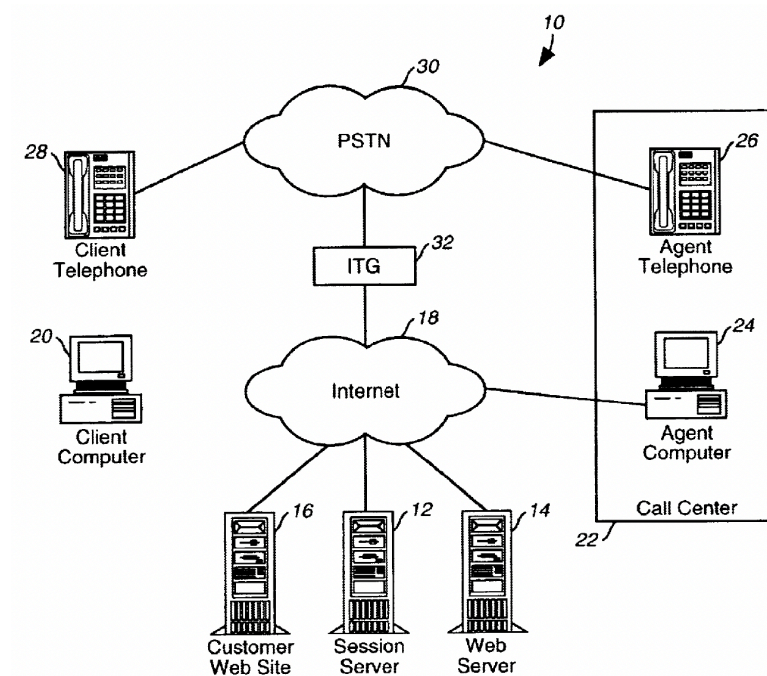


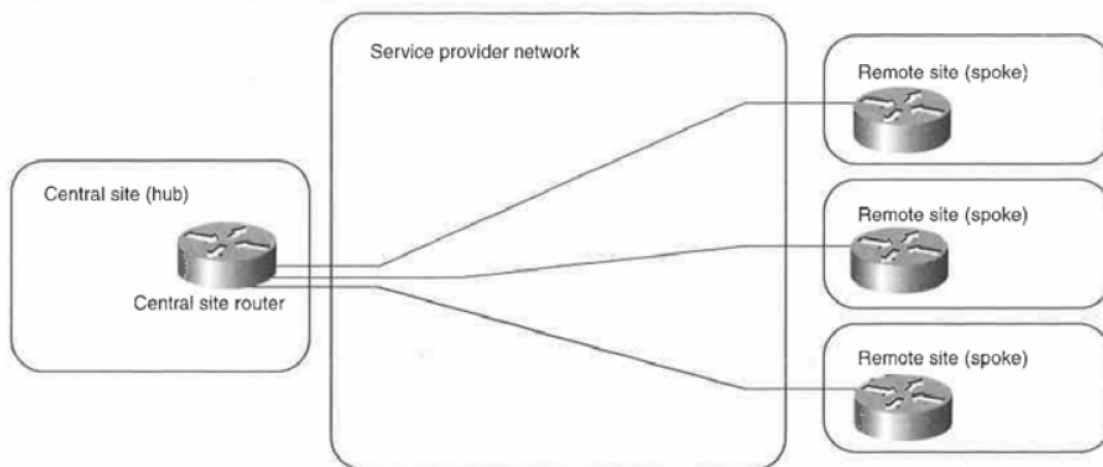
FIG. 1

Id., FIG. 1; Ex_1003 ¶130. “[N]etwork 18 can be either a public or a carrier supported virtual private network (VPN) or local area network (LAN) which uses IP or IP-tunneling. Ex_1007, 4:28-32; Ex_1003 ¶131.

c) **Summary of VPN Textbook**

The VPN Textbook describes well-known topologies for implementing VPNs. Ex_1003 ¶133. The VPN Textbook describes that “[t]he most commonly encountered topology is a hub-and-spoke topology, where a number of remote offices (spokes) are connected to a central site (hub).” Ex_1010 at 129; Ex_1003 ¶133. In the art, the “hub-and-spoke” topology, shown in Figure 7-10 of the VPN Textbook, also is referred to as a “star” topology. Ex_1027 at 438; *see also* Ex_1011 ¶¶16, 23, 79-94; Ex_1003 ¶¶133-134.

Figure 7-10 *Hub-and-spoke Topology*



Ex_1010 at 130; Ex_1003 ¶134. The VPN Textbook describes this hub-and-spoke (aka star) topology as being the “most commonly encountered topology,” chosen often “for cost or complexity reasons.” Ex_1010 at 129; Ex_1003 ¶ 135.

2. Motivation to Combine Prior Art

a) Reasons to Combine Knappe and Elliott

A POSITA would have been motivated to combine the teachings of Knappe and Elliott to produce the obvious, beneficial, and predictable results of using different types of collaboration data and VPN tunneling within Knappe's system.

Ex_1003 ¶136. A POSITA, when considering Knappe, also would have considered Elliott since they are analogous, both pertaining to the field of multiparty conferencing. *Id.* ¶137; *see e.g.*, Ex_1006, Abstract, 3:26-29; Ex_1007, Abstract. Also, both Knappe and Elliott support computer and PSTN telephone conference participants. Ex_1003 ¶137; *see, e.g.*, Ex_1006, 3:41-49, FIG. 1; Ex_1007, 5:35-39, 6:1-11, 8:16-39, FIG. 1.

A predictable and beneficial result of combining Knappe and Elliott includes the sharing of collaboration data in a multiparty conference. Ex_1003, ¶¶138-144. As discussed above, Knappe teaches a communications system capable of “transmitting audio, video, voice, data and other suitable types of...traffic between source and destination endpoints.” Ex_1006, 3:26-29; Ex_1003 ¶138. To a POSITA, transmitting “data” or other “traffic” would suggest the transmission of various types of data during the conference call to facilitate collaboration. Ex_1003 ¶138. Elliott expressly teaches “collaborative Web browsing, audio conferencing, video conferencing, and application sharing.” Ex_1007, 1:43-44;

Ex_1003 ¶139. Elliott discloses connecting calls “over the public switched telephone network (PSTN) 30” using “session server 12 ... that follows the ITU T.120 standard” capable of “support[ing] real-time, multi-point data communications,” the server “able to set up to conduct video conferences, audio conferences, and data conferences with application sharing,” as well as “[c]ollaborative Web browsing.” Ex_1007, 5:35-61; Ex_1003 ¶139. Exchanging collaboration data was standardized (e.g., under the T.120 standard) and VPN tunneling was conventional by the time of Knappe and Elliott and the earliest claimed priority of the ’332 patent. Ex_1003, ¶140; *see also* Ex_1019, 4:23-27, 6:36-41; Ex_1007, 5:49-56.

It would have been obvious to a POSITA, in view of Elliott’s teachings, to implement Knappe’s transmission of various types of data as collaboration data, because this type of collaboration data would allow a greater variety of information among the participants, improving the experience and enhancing the effectiveness of collaboration in a variety of contexts. Ex_1003, ¶141. Indeed, other patents and printed publications confirm that it was known at the time that additional collaboration data improves the conferencing experience:

collaboration often involves the sharing of visual information—e.g., printed material such as articles, drawings, photographs, charts and graphs, as well as video-tapes and computer-based animations, visualizations and other displays – in such a way that the participants can collectively and interactively examine, discuss, annotate and revise the information. This combination of spoken words, gestures,

visual cues and interactive data sharing **significantly enhances the effectiveness of collaboration in a variety of contexts**[.]

Ex_1018, 1:29-39 (emphasis added); Ex_1003 ¶142. Sharing collaborative information such as written or graphic material during a video conference was known in the conferencing art for over 30 years prior to the '332 patent. Ex_1003 ¶143. Accordingly, a POSITA would have recognized that sharing collaboration data would be beneficial to Knappe's participants. Ex_1003 ¶144. Such a combination would have simply involved combining prior art element according to known elements to yield predictable results, namely allowing participants in Knappe's system to share various types of data during a conference for collaboration purposes. *Id.* The combination of Knappe and Elliott is merely the ordinary use of common techniques to improve the similar conference systems of Knappe and Elliott. *Id.*

Another predictable and beneficial result of combining Knappe and Elliott is the use of a VPN to protect communications between conference participants. Ex_1003 ¶¶145-153. It was known that data transmitted over public IP networks is vulnerable to eavesdropping and other problems. Ex_1003 ¶146; Ex_1010 at 119 (stating that "more and more business-to-business traffic takes place over the Internet," which is "riddled with security issues"). Using a VPN was a known solution for securing communications over public networks. Ex_1003 ¶147; Ex_1010 at 115 ("A Virtual Private Network (VPN) ... [provides] connectivity

amongst multiple sites ... with the same access or security policies as a private network.”). Thus, it would have been obvious to a POSITA to use VPNs with Knappe’s multiparty conferencing to secure communications over the Internet. Ex_1003 ¶148.

The obviousness of adding a VPN feature to Knappe’s conferencing system is further demonstrated by the work of others making similar feature combinations as evidenced by other prior art patents. Ex_1003 ¶149; *see, e.g.*, Ex_1020, 2:65-3:3, 5:52-55 (disclosing “need for a videoconferencing system ... for delivering secure ... services over an IP network” by using a “VPN ... over public networks ... which is encrypted to make it secure from outside access and interference”); Ex_1019, 2:66-3:1, 4:23-27, 6:36-41 (describing a “conference bridge application” that transmits data “over public networks, using technologies such as ... [a] Virtual Private Network (VPN”). Therefore, a POSITA would have found it obvious to combine Knappe’s conferencing system with the VPN technology described in Elliott and the VPN Textbook (discussed below). Ex_1003 ¶¶150-153, 155-162; Ex_1023 ¶¶[1412], [1452]-[1453] (describing “VPN connectivity,” including “hub-and-spoke or mixed topologies” in the context of “services [that] include Internet Access, ... chat and Video-conferencing etc.”).

Specifically, Elliott teaches using a “virtual private network (VPN) ... which uses IP or IP-tunneling” to transmit “voice, fax, data, video” and “[o]ther types of

collaboration ... data.” Ex_1007, 2:66-3:1, 4:29-34, 6:53-54; Ex_1003 ¶151. In view of Elliott, it would have been obvious to a POSITA to use VPN tunneling when communicating audio, video, and collaboration data across Knappe’s Internet network 14 to make the transmitted data secure without the need to lease private data lines and to otherwise reduce costs for implanting a private network securely over the Internet. Ex_1003 ¶151; *see* Ex_1020, 5:52-55 (“VPN refers to a network that is carried over public networks, but which is encrypted to make it secure from outside access and interference.”); Ex_1026, 44:10-12 (“[O]ffices [] create virtual private networks (VPNs) to save money on remote access and interoffice connectivity.”). Indeed, the combination of Knappe and Elliott is merely the ordinary use of the common technique of VPN tunneling of conferencing data, as disclosed in Elliott, in the similar conferencing system of Knappe to yield predictable results, including securing transmitted data and reducing costs associated with leasing private data lines. Ex_1003 ¶152. Petitioner’s proposed combination permits, but does not require, physical incorporation of elements from Elliott and Knappe. *See In re Mouttet*, 686 F.3d 1322, 1332 (Fed. Cir. 2012); *In re Etter*, 756 F.2d 852, 859 (Fed. Cir. 1985) (en banc) (noting that the criterion for obviousness is not whether the references can be combined physically, but whether the claimed invention is rendered obvious by the teachings of the prior art as a whole); Ex_1003 ¶153.

b) Reasons to Combine Knappe and VPN Textbook

A POSITA would have been motivated to combine the teachings of Knappe and the VPN Textbook to produce the obvious, beneficial, and predictable results of providing hub-and-spoke VPN topology when implementing VPN tunneling in the Knappe and Elliott combination, and to provide a cost effective and streamlined VPN implementation. Ex_1003 ¶154. A POSITA would have been familiar with conventional VPN tunneling implementations well documented in readily available textbooks, such as the VPN Textbook. Ex_1003 ¶¶133-135, 155; *see also id.* ¶¶96-97.

A POSITA looking to implement Knappe's system with VPN tunneling would have looked to the VPN Textbook to consider various topologies, including the "most commonly encountered topology is a hub-and-spoke topology, where a number of remote offices (spokes) are connected to a central site (hub)." Ex_1010 at 129; Ex_1027 at 438; Ex_1003 ¶¶156-157. Combining Knappe with the VPN Textbook would merely be a combination of prior art elements (e.g., VPN topology teachings of the VPN Textbook and Knappe's system) according to known methods to yield predictable results—securing traffic exchanged in Knappe's conferencing system. Ex_1003 ¶157. A POSITA would have been able to predict the result of combining a hub-and-spoke VPN topology, as described in the VPN Textbook, in Knappe's system, for example, as modified by Elliott (as

described above), to communicate data between conferencing participants (e.g., remote clients) over secure VPN tunnels. *Id.* ¶158.

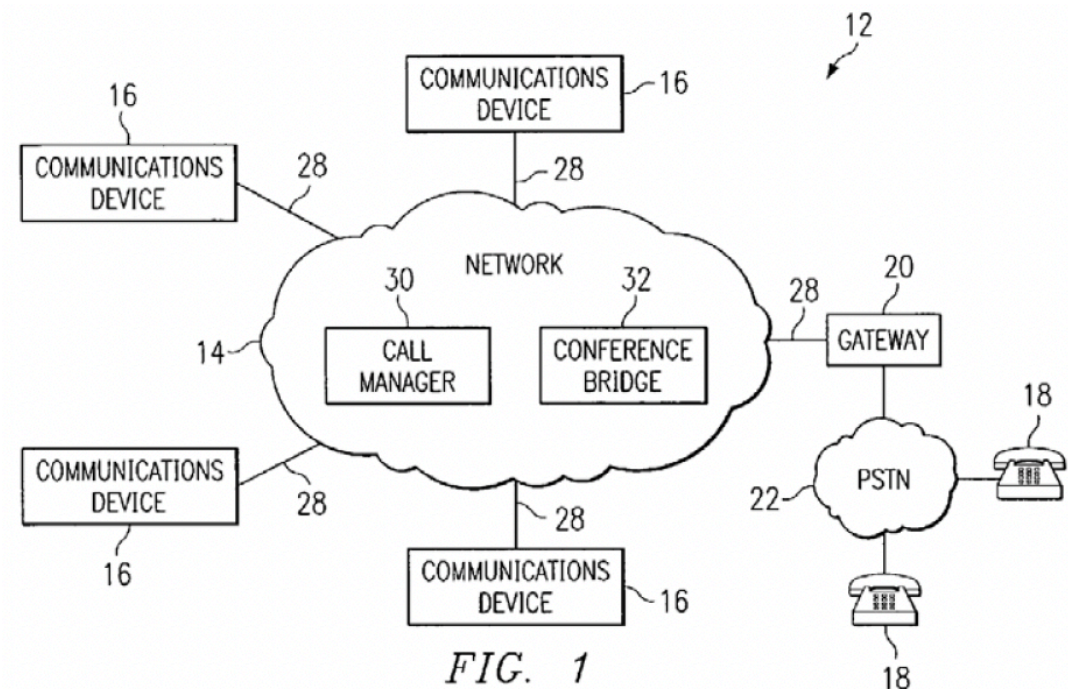
Furthermore, a POSITA would have been motivated to combine the teachings of Knappe, as modified by Elliott, with the teachings of the VPN Textbook for cost and complexity savings. Ex_1010 at 129 (“there are many examples where the customer could benefit from a different topology but has nonetheless chosen the hub-and-spoke topology for cost or complexity reasons”); Ex_1003 ¶159 (implementing VPN tunneling in Knappe’s conferencing system, per Elliott’s teaching, would beneficially constrain cost and limit complexity). A POSITA would have had a reasonable expectation of success in combining Knappe with the VPN Textbook’s hub-and-spoke topology because the topology was conventional and known in the conferencing context. Ex_1003 ¶¶160-162; *see* Ex_1019, 4:23-27, 6:36-41 (describing a “conference bridge application” that transmits data “over public networks, using technologies such as ... Virtual Private Network (VPN)”); Ex_1023, ¶¶[1412], [1451]-[1453] (disclosing VPN connectivity, including hub-and-spoke or mixed topologies, for services including video-conferencing). Petitioner’s proposed combination permits, but does not require, physical incorporation of elements from the VPN Textbook into Knappe’s system. *See In re Mouttet*, 686 F.3d at 1322; Ex_1003 ¶162.

3. Claim 1

[1.P] A method for supporting a multi-participant audio/video conference call, the method comprising:

Knappe discloses “a method and system for improving ... a multiparty communication session ... [that] allow[s] multiple people to participate in ... a ‘conference call.’” Ex_1006, 1:8-27, 8:46-10:50, FIG. 8; Ex_1003 ¶164.

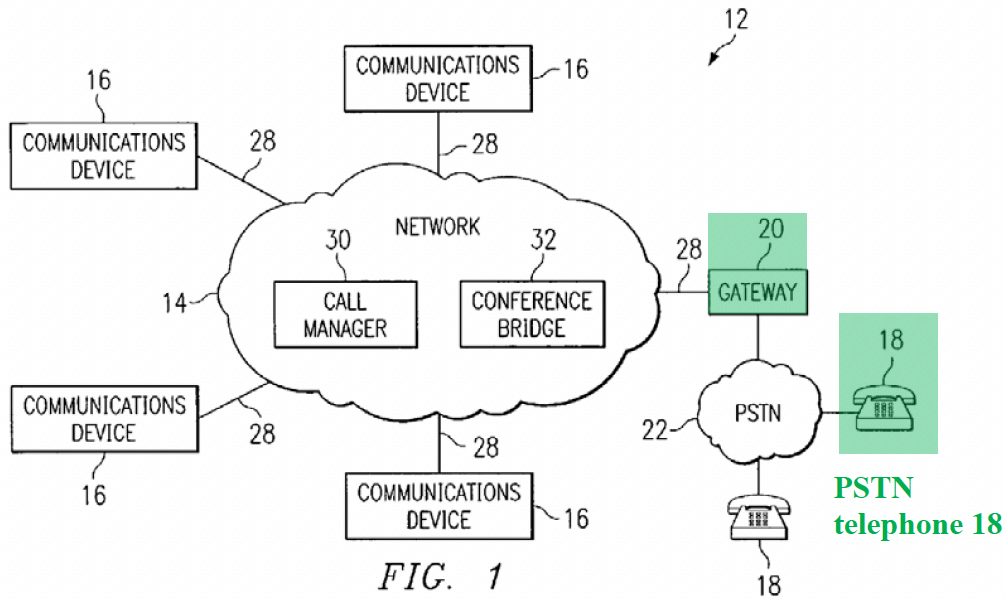
Knappe’s system supports “transmitting audio, video, voice, data ... between source and destination endpoints ... to conduct multiple party telephone conference communication sessions.” Ex_1006, 3:24-31, 4:58-63; Ex_1003 ¶¶118-120, 122-125, 164. As shown in Figure 1 of Knappe, Knappe’s system can support multiple participants in an audio/video conference call:



Ex_1006, FIG. 1. Specifically, Knappe explains that “conference bridge 32 provides conference call and other suitable audio, video, and/or real-time multi-party communication sessions between communication devices 16.” *Id.*, 4:58-5:1; Ex_1003 ¶165. Thus, a POSITA would understand Knappe to support a multi-participant audio/video conference call. Ex_1003 ¶¶118-120, 163-165.

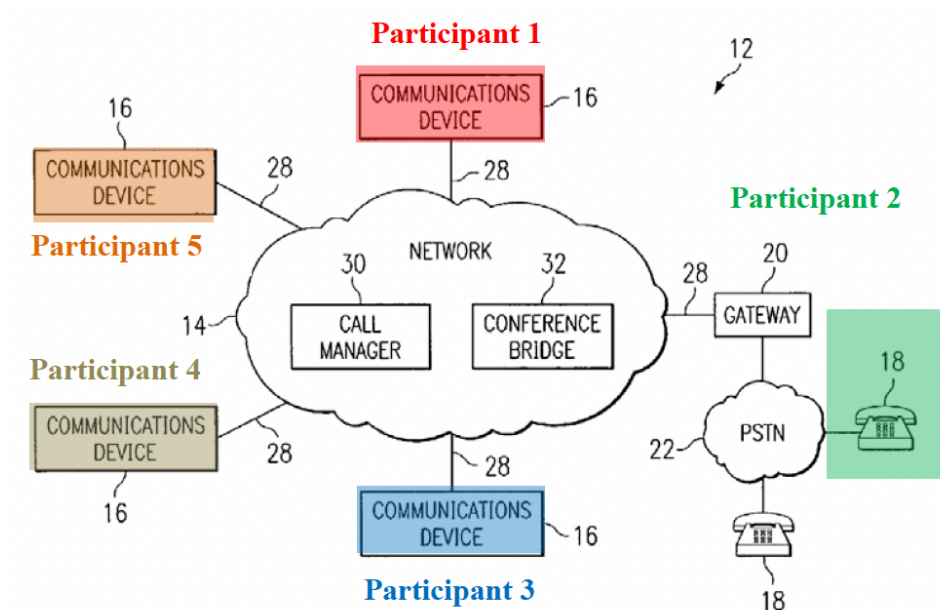
[1.1] receiving first audio data from a Public Switched Telephone Network (PSTN) client;

Knappe teaches “[c]ommunication networks” with one or more telephones 18 and a gateway 20 connected to a PSTN network 22, for “transporting electrical representations of audible sounds.” Ex_1006, 1:16-17, FIG. 1; Ex_1003 ¶167. According to Knappe, “telephones 18 communicate standard telephony signals through PSTN 22 to the gateway 20.” Ex_1006, 4:18-20. The signals from PSTN telephone 18 are “converted to IP packets in the VoIP format” by gateway 20, the VoIP packets “encapsulated in real-time transport protocol (RTP) packets ... or other protocols and formats” for transmission over a network. *Id.*, 3:56-61, 4:20-22, 5:23-33. Thus, a POSITA reviewing Knappe would have understood that audio signals from PSTN telephone 18 are sent to gateway 20, which communicates corresponding audio data to conference bridge 32, which facilitates the multiparty communications, as shown in the annotated version of the ’332 patent’s Figure 1, below.



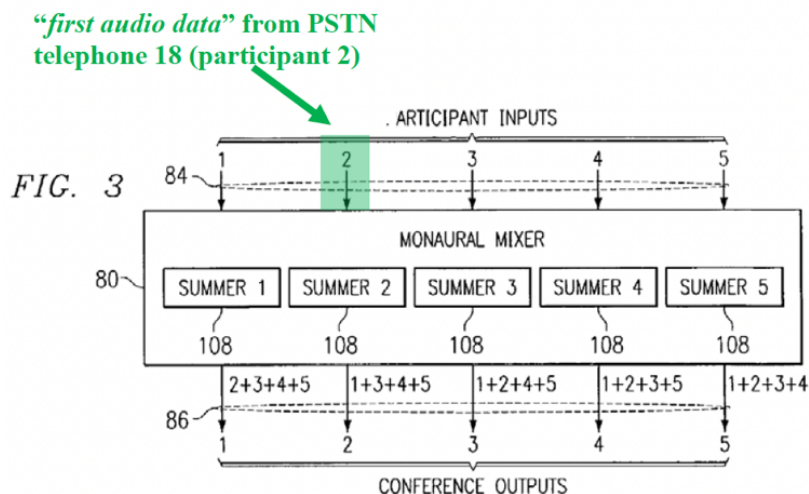
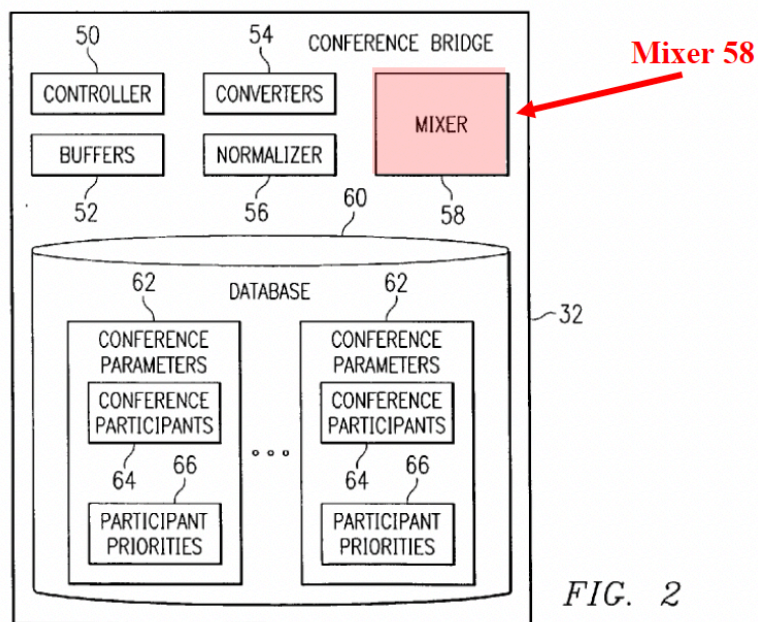
Id., FIG. 1(annotated); Ex_1003 ¶¶168-169.

Thus, a POSITA understands that in Knappe, audio data from PSTN clients, such as telephone(s) 18, are communicated over link 28 and network 14 to conference bridge 32. Ex_1003 ¶¶169-170. As an example, the below Figure 1 of Knappe has been annotated with labels corresponding to participants 1-5 discussed in connection with Figure 3 of Knappe:



Ex_1003 ¶171. Knappe explains that “conference bridge 32” includes buffers 52, including “input buffers [that] receive and buffer packets of input audio streams from participants for processing by conference bridge 32.” Ex_1006, 6:3-6. Thus, Knappe discloses receiving audio data from a PSTN client. Ex_1003 ¶ 171.

Furthermore, as shown in annotated versions of Figures 2-3 of Knappe below corresponding to the annotated version of Figure 1 above, conference bridge 32 includes at least mixer 58, which may correspond to monaural mixer 80 that “receives participant input streams 84 and combines the streams in summers 82 to generate conference output streams 86 for each participant to a conference call.” Ex_1006, 6:59-7:11, FIGS. 2-3.



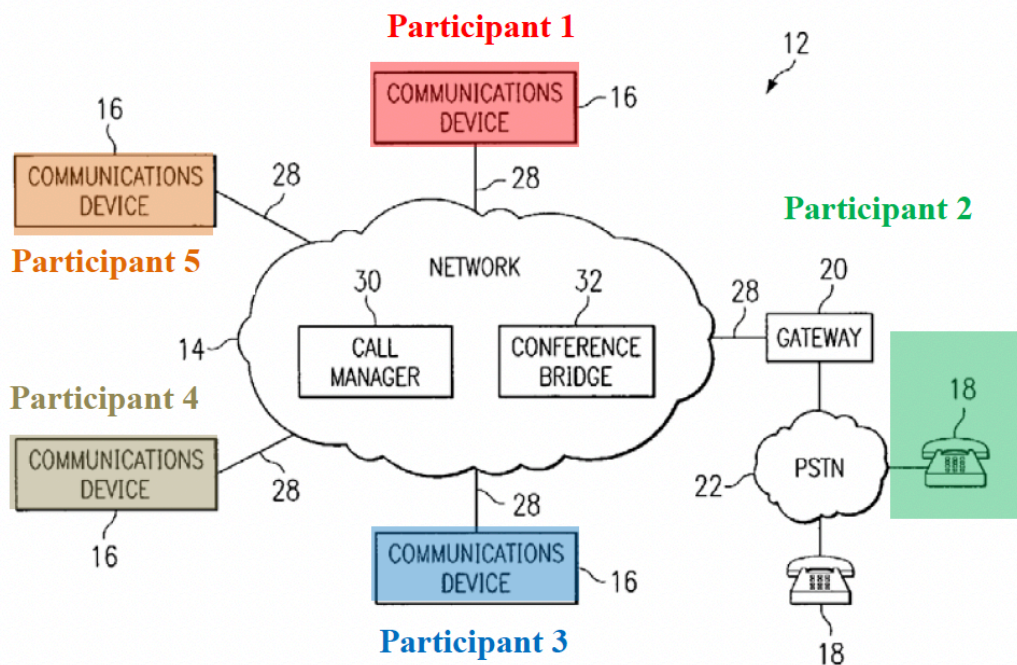
Id., FIGS. 2-3 (annotated); Ex_1003 ¶172. Thus, Knappe discloses conference bridge 32, which receives audio data from a PSTN telephone 18 through gateway 20. Ex_1003 ¶¶166-173.

[1.2] receiving second audio data from a moderator;

Knappe teaches “communication devices 16 comprise IP or other digital telephones, personal and other suitable computers or computing devices ... capable

of communicating real-time audio, video and/or other information over the network 14.” Ex_1006, 3:62-4:1. Knappe further teaches that “at least one participant may be given the priority of moderator for a particular conference call.” *Id.*, 8:52-56, 7:12-15; Ex_1003, ¶175. Thus, a POSITA reading Knappe would understand that one of Knappe’s communications devices 16 may be designated a moderator. Ex_1003 ¶175.

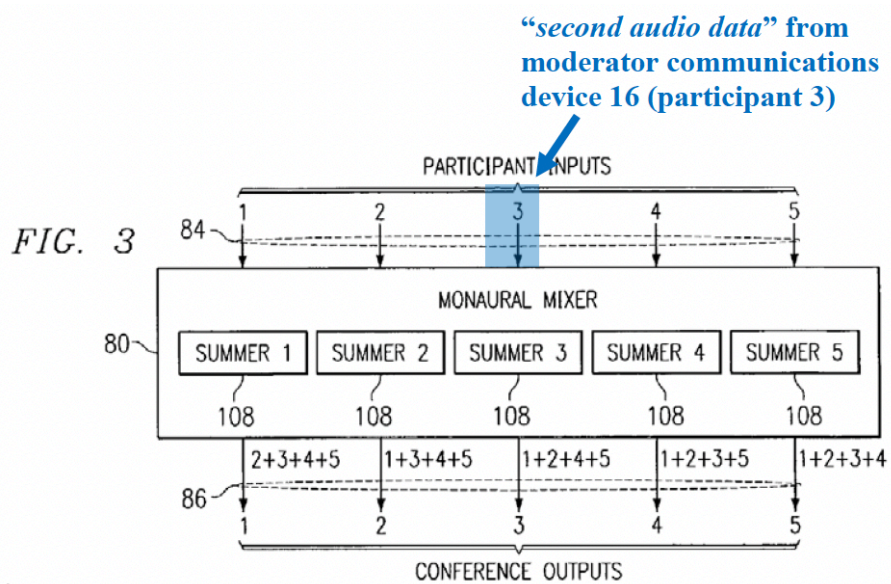
The below annotated Figure 1 from Knappe depicts an example where the moderator is participant 3 in the Knappe system:



Ex_1006, FIG. 1 (annotated); Ex_1003 ¶176.

And as noted in [1.1], Knappe teaches that the “input buffers receive and buffer packets of input audio streams from participants for processing by conference bridge 32,” which would include audio streams from participant 3 (the

moderator). Ex_1006, 6:3-6; Ex_1003 ¶177. Thus, Knappe discloses receiving audio data from a moderator. Ex_1003 ¶¶176-179. Additionally, Knappe teaches that mixer 80 “receives participant input streams 84,” which may include second audio data from a moderator, “and combines the streams in summers 82 to generate conference output streams 86 for each participant to a conference call.” Ex_1006, 6:59-7:11, FIGS. 2-3.



Id., FIG. 3 (annotated); Ex_1003 ¶178. Thus, Knappe discloses that conference bridge 32 receives audio data from a moderator. Ex_1003 ¶¶174-179.

[1.3.1] receiving third audio data, video data, and collaboration data from at least one remote client;

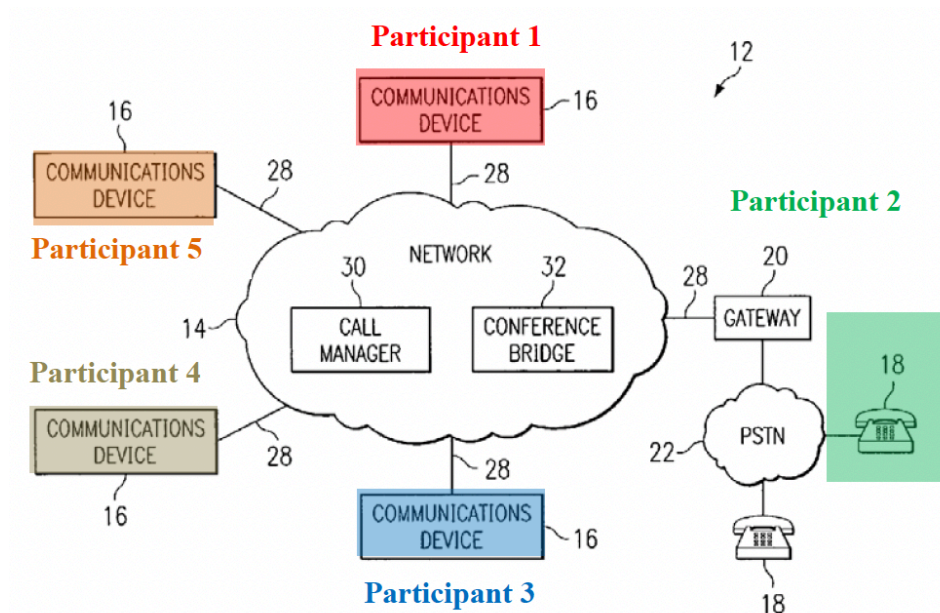
Knappe teaches receiving “third audio data, video data, and collaboration data from at least one remote client.” Specifically, Knappe teaches “communication devices 16 comprise IP or other digital telephones, personal and

other suitable computers or computing devices ... capable of communicating real-time audio, video and/or other information over the network 14” and “transmitting audio, video, voice, data and other suitable types of real-time ... traffic.”

Ex_1006, 3:62-4:1, 3:24-29; Ex_1003 ¶181. Knappe also teaches communication between remote clients over a network and PSTN where “each participant in the conference call may share information with all of the other participants,” contrasting the disclosed system with “a meeting with all persons in the same physical location” and noting the benefit of “minimizing travel expenses.”

Ex_1006, 1:29-47; Ex_1003 ¶182.

A POSITA would have understood that the participants and their devices were physically remote from one another. Ex_1003 ¶183; *see also* ¶119-120, 123. As shown in this annotated Figure 1, Knappe discloses at least a third remote client connected to the network over link(s) 28 exchanging “audio, video, and/or ... other suitable information”:



Ex_1006, FIG. 1 (annotated), 4:58-5:1; Ex_1003 ¶183. As discussed above, Knappe teaches that these conference “participants are connected and stream media through the conference bridge 32.” Ex_1006, 5:31-33; *see also id.*, 4:65-66 (“conference bridge 32 receives media from participating devices 16”), 6:59-67 (“... each participant is assigned a summer 82 that receives audio input streams from each other participant to the conference call”); Ex_1003 ¶184. In Knappe, streamed media include “audio, video, voice, data and other suitable types of real-time ... traffic.” Ex_1006, 3:24-29. A POSITA would have understood “data and other suitable types of real-time ... traffic” to include data to be used for collaboration, including well-known collaboration data such as T.120 standard data. *Id.*, 3:27-28; Ex_1003 ¶185.

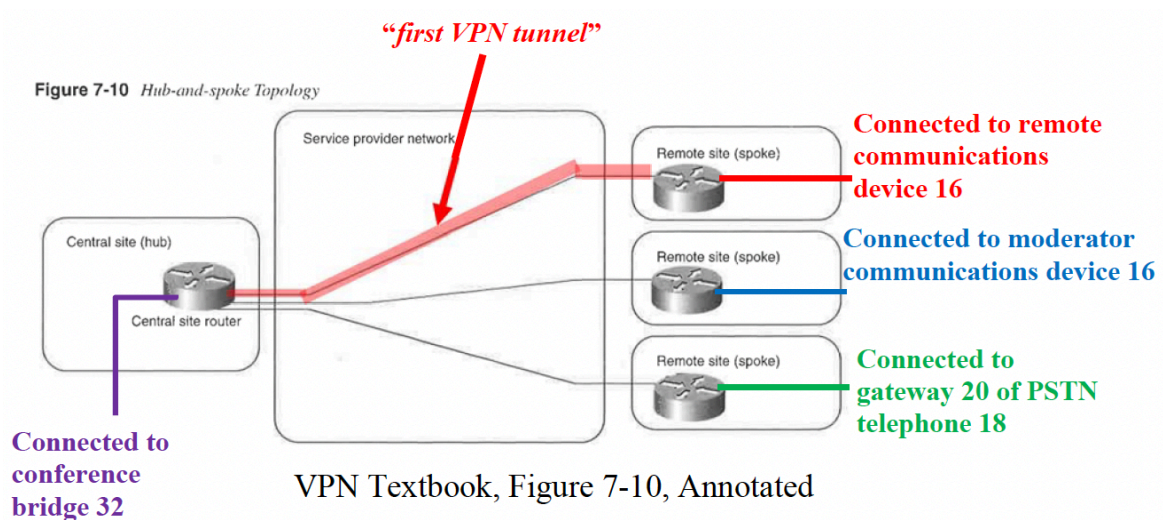
Elliott further teaches “collaboration services [between] two or more parties ... can be readily provided over the Internet and include collaborative Web browsing, audio conferencing, video conferencing, and application sharing,” along with “data conferencing.” Ex_1007, 1:40-44, 5:49-61, 6:46-54; Ex_1003 ¶186. A POSITA would have understood that Knappe’s conferencing system could be implemented with Elliott’s teachings (e.g., application sharing data, web browsing data, and other collaboration data). Ex_1003 ¶¶ 136-144, 186-187. Therefore, a POSITA would have understood the prior art to render obvious this limitation. Ex_1003 ¶188.

[1.3.2] through a first Virtual Private Network (VPN) tunnel

Elliott teaches that such “collaborative Web browsing, audio conferencing, video conferencing, and application sharing” is made over Internet network 18 and can be a “virtual private network (VPN) . . . which uses IP or IP-tunneling.” Ex_1007, 1:43-44, 4:6-12, 4:31-34; Ex_1003 ¶189. A POSITA would have been motivated to combine Elliott’s VPN tunneling for communicating conferencing audio, video, and collaboration data between Knappe’s remote communication devices 16 and conference bridge 32 to provide efficient and secure communications. Ex_1003 ¶¶145-153, 189-190.

Further, the VPN Textbook teaches “a hub-and-spoke topology” where a number of remote sites are connected to a central site (hub) through a VPN tunnel.

Ex_1010 at 129; Ex_1003 ¶191. Accordingly, it would have been obvious that when implementing this topology, Knappe's remote communications device 16, moderator communications device 16, and gateway 20 connected to the PSTN telephone 18 would be connected to different remote site (spoke) routers, while the conference bridge 32 would be connected to the central site (hub) router, as illustrated below:

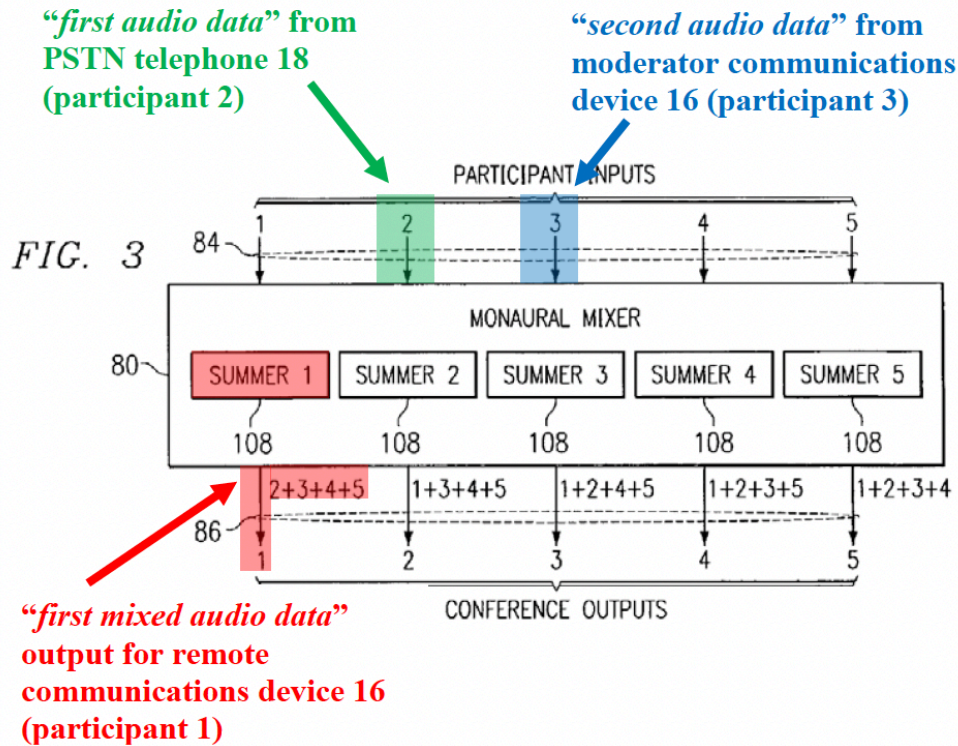


Ex_1010, Fig. 7-10 (annotated); Ex_1003 ¶191-193. As shown, the remote site route (spoke) of the remote communications device 16 and the central router (hub) of the conference bridge 32 are connected through a VPN tunnel. When the conference bridge 32 receives data from or transmits data to the remote communications device 16, the data would be received/sent through the same VPN tunnel, which corresponds to “a first VPN tunnel.” Ex_1003 ¶¶193-194; see also

id. ¶¶154-162. Thus, a POSITA would have understood the prior art to render obvious this limitation. *Id.* ¶¶189-194.

[1.4] mixing the first audio data from the PSTN client with the second audio data from the moderator into a first mixed audio data;

As discussed above at [1.1] and [1.2], Knappe teaches receiving “*the first audio data from the PSTN client*” and “*the second audio data from the moderator.*” *Id.* ¶196; *see also id.* ¶¶166-179. Knappe also teaches that mixers 58, 80, and 100, of conference bridge 32 mixes the audio data received from PSTN telephone 18 with audio data received from a moderator’s communications device 16 into “*a first mixed audio data.*” *See id.* ¶196. For example, Knappe explains that “monaural mixer 80 receives participant input streams 84 and combines the streams in summers 82 to generate conference output streams 86 for each participant.” Ex_1006, 6:59-67; Ex_1003 ¶197. As a result, “the conference output stream of participant 1 includes the audio inputs of participants 2-5 ... the conference output stream of participant 2 includes the audio inputs of participants 1 and 3-5,” and so on. Ex_1006, 7:1-11; Ex_1003 ¶197. A POSITA would have understood Knappe to disclose the “*mixing the first audio data from the PSTN client with the second audio data from the moderator into a first mixed audio data,*” as shown in an exemplary implementation in the below annotated Figure 3:



Ex_1006, FIG. 3 (annotated); Ex_1003 ¶198. Therefore, Knappe teaches this limitation. Ex_1003 ¶¶195-199.

[1.5] transmitting the first mixed audio data to the remote client through the first VPN tunnel;

As discussed at [1.3]-[1.4], Knappe teaches generating “the first mixed audio data” (e.g. output streams 86) as well as “remote client[s]” (e.g., remote communications devices 16). *Id.* ¶¶180-199. Knappe also discloses “transmitting audio ... between source and destination endpoints,” including “a conference output stream for delivery to the participant.” Ex_1006, 3:24-29; 6:65-67; Ex_1003 ¶200; *see also* Ex_1006, 5:35-39 (“The conference output stream for a participant includes the media of all other participants, a subset of other

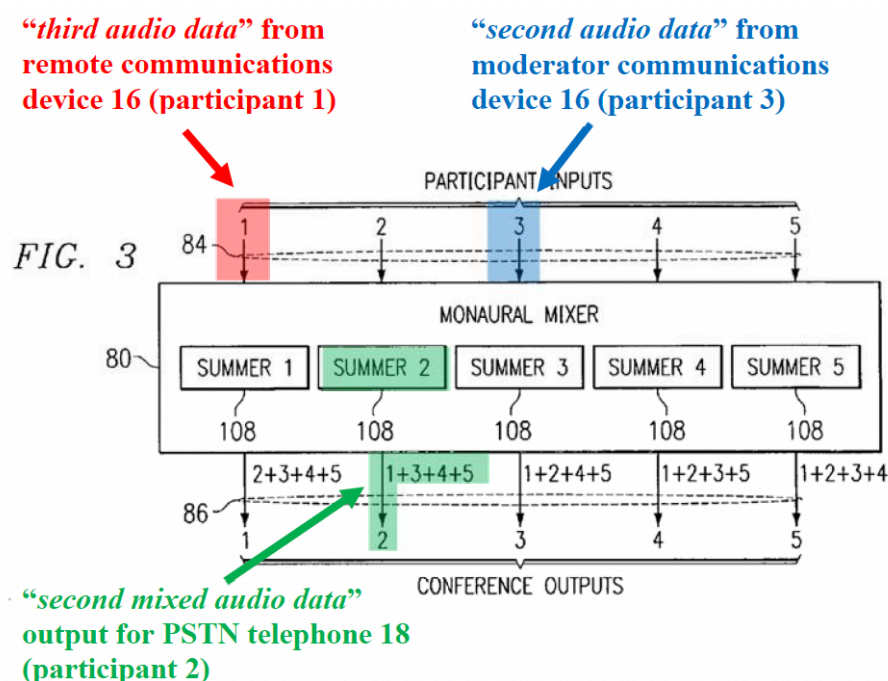
participants or other suitable mix ...”). Thus, a POSITA would have understood Knappe to disclose “*transmitting the first mixed audio data to the remote client.*” Ex_1003 ¶200.

As discussed at [1.3.2], Knappe in combination with Elliott and the VPN Textbook teaches using a VPN tunnel between a central site (e.g., Knappe’s conference bridge 32) and a remote client (e.g., remote communications devices 16). Ex_1003 ¶¶189-194, 201. It would have been obvious to a POSITA to transmit “the first mixed audio data to the remote client” using a VPN tunnel set up between conference bridge 32 and remote communications devices 16. Ex_1003, ¶¶201-202. Thus, the prior art renders obvious this limitation. *Id.* ¶¶200-202.

[1.6] mixing the second audio data from the moderator with the third audio data from the remote client into a second mixed audio data; and

As discussed at [1.2] and [1.3], Knappe teaches receiving “*the second audio data from the moderator,*” and “*the third audio data from the remote client.*” *Id.* ¶¶174-194, 204. Also, as discussed at [1.4], Knappe teaches that mixers 58, 80, and 100, of conference bridge 32, mixes the audio data from PSTN telephone 18 with audio data from a moderator’s communications device 16 into “*a first mixed audio data.*” *See id.* ¶¶195-199, 204. As shown in annotated Figure 3 below, “monaural mixer 80 receives participant input streams 84 and combines the streams in summers 82 to generate conference output streams 86 for each

participant,” which also may include mixing audio data from a moderator’s communications device 16 with another participant’s communications device 16. Ex_1006, 6:59-67; Ex_1003 ¶205. As a result, Knappe’s “conference output stream of participant 2” can comprise “*a second mixed audio data*” resulting from a mix of “the audio inputs of participants 1 and 3-5,” which may include “*the second audio data from the moderator with the third audio data from the remote client*”:



Ex_1006, 7:1-11; Ex_1003 ¶205. Thus, a POSITA would have understood this description in Knappe to disclose this limitation. Ex. 1003 ¶¶203-206.

[1.7] transmitting the second mixed audio data to the PSTN client.

As discussed at [1.5], Knappe discloses “transmitting audio ... between source and destination endpoints,” including “a conference output stream for

delivery to the participant.” Ex_1006, 3:24-29; 6:65-67; Ex_1003 ¶¶200-202, 205; *see also* Ex_1006, 5:35-39 (“The conference output stream for a participant includes the media of all other participants, a subset of other participants or other suitable mix ...”). Further, Knappe teaches the transmitted audio data is received at gateway 20 and converted to standard telephony signals for delivery to destination PSTN telephones 18. Ex_1006, 4:21-25 (“At the gateway 20 ... VoIP packets received from the network 14 are converted into standard telephony signals for delivery to the destination telephone 18 through PSTN 22.”), 7:1-7 (“... each participant receives the audio input of each other participant.”); Ex_1003 ¶209. Thus, a POSITA would have understood Knappe to disclose this limitation. Ex_1003 ¶¶207-210.

4. Claim 2

[2.P] The method of claim 1 further comprising:

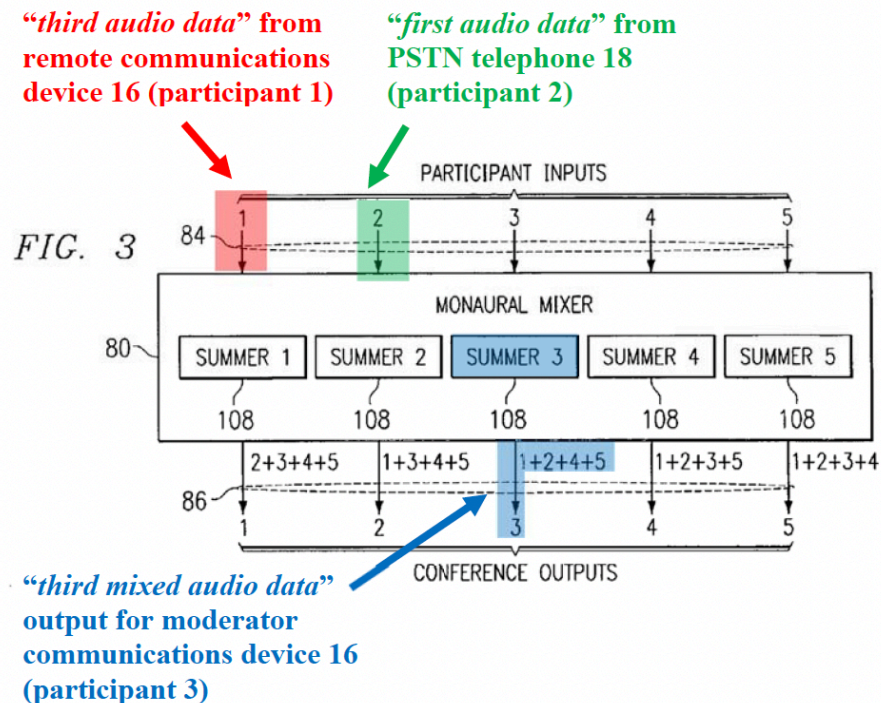
See [1.P]-[1.7] above.

[2.1] mixing the first audio data from the PSTN client with the third audio data from the remote client into a third mixed audio data;

As discussed at [1.1] and [1.3], Knappe teaches receiving “*the first audio data from the PSTN client*” and “*the third audio data from the remote client*” by a mixer (e.g., mixers 58, 80 or 100). *See* Ex_1003, ¶¶ 166-173, 180-188, 213. Also, as discussed at [1.4] and [1.6], Knappe teaches that mixers 58, 80, and 100, of conference bridge 32, can mix the audio data from PSTN telephone 18 with audio

data from a moderator's communications device 16 into *"a first mixed audio data."*
See id. ¶¶194-198, 203-206, 214.

As shown in annotated Figure 3 below, "monaural mixer 80 receives participant input streams 84 and combines the streams in summers 82 to generate conference output streams 86 for each participant," which also may include mixing audio data from PSTN telephone 18 with audio from another participant. Ex_1006, 6:59-67; Ex_1003 ¶¶215-216. As a result, Knappe's "conference output stream of participant 3" can comprise *"a third mixed audio data"* resulting from a mix of "the audio inputs of participants 1-2 and 4-5," which may include *"the first audio data from the PSTN client with the third audio data from the remote client"*:



Ex_1006, FIG. 3, 7:6-7; Ex_1003 ¶216. Thus, a POSITA would have understood this description in Knappe to disclose this limitation. Ex. 1003 ¶¶212-217.

[2.2] playing the third mixed audio data to the moderator.

As discussed at [2.1], Knappe teaches that mixers 58, 80, and 100, of conference bridge 32, can generate “*the third mixed audio data*,” and that “each device 16 receives a conference signal that includes contributions from all other participating devices.” Ex_1006, 4:58-5:1; Ex_1003 ¶¶212-217, 219. Also, as discussed at [1.2], one of Knappe’s communications devices 16 may be designated as a moderator, which can receive the mixed audio data from said mixers.

Ex_1006, 8:52-53 (“... at least one participant may be given the priority of moderator for a particular conference call.”); Ex_1003 ¶¶175-179, 220.

Further, Knappe teaches that “the communication devices 16 comprise real-time applications that play traffic as it is received” with “IP packets received from the network 14 [being] converted back into analog signals and played to the user.” Ex_1006, 4:5-17; Ex_1003 ¶221. Therefore, Knappe discloses communications device 16 receiving “*the third mixed audio data*” and “*playing*” it “*to the moderator*.” Ex_1003 ¶¶218-212.

5. Claim 4

[4.P] The method of claim 1

See [1.P]-[1.7] above.

[4.1] wherein the receiving third audio data from the remote client includes receiving IP packets containing the third audio data corresponding to the remote client and decoding the IP packets containing the third audio data from the remote client.

As discussed at [1.3], Knappe teaches conference bridge 32 “receiving third audio data from the remote client” (e.g., remote communications device 16).

Ex_1003 ¶¶180-185, 223. Knappe also teaches that network 14 uses IP packets for communication. Ex_1006, 3:50-54 (“network 14 is the Internet ... In the Internet embodiment, the network 14 transmits information in Internet Protocol (IP) packets.”); Ex_1003 ¶225.

Knappe further discloses “decoding the IP packets”:

The converters 54 include input and output converters. The input converters receive input packets of a participating from a corresponding buffer and **convert the packet** from the native format of the participant device 16 to a standard format of the conference bridge 32 . . .

Ex_1006, 6:16-21 (*see also id.*, 6:3-15); Ex_1003 ¶226. A POSITA would understand that Knappe’s conversion of the received IP packets from the native format of the participant device to a standard format would include the claimed “decoding.” Indeed, Knappe discloses the use of

[a] codec (coder/decoder) [that] converts audio, video, or other suitable signals generated by users from analog signals into digital format. The digital encoded data is encapsulated into IP or other suitable packets from transmission over the network 14. IP packets received from network 14 are converted back into analog signals and played to the users. It will be understood that the communication

devices 16 may otherwise suitably encode and decode signals transmitted over or received from the network 14.

Ex_1006, 4:4-17; *see also id.*, 4:18-28, 6:3-21; Ex_1003 ¶227. Thus, a POSITA would have understood Knappe to have disclosed this limitation. Ex_1003 ¶¶224-228.

6. Claim 5

[5.P] The method of claim 1

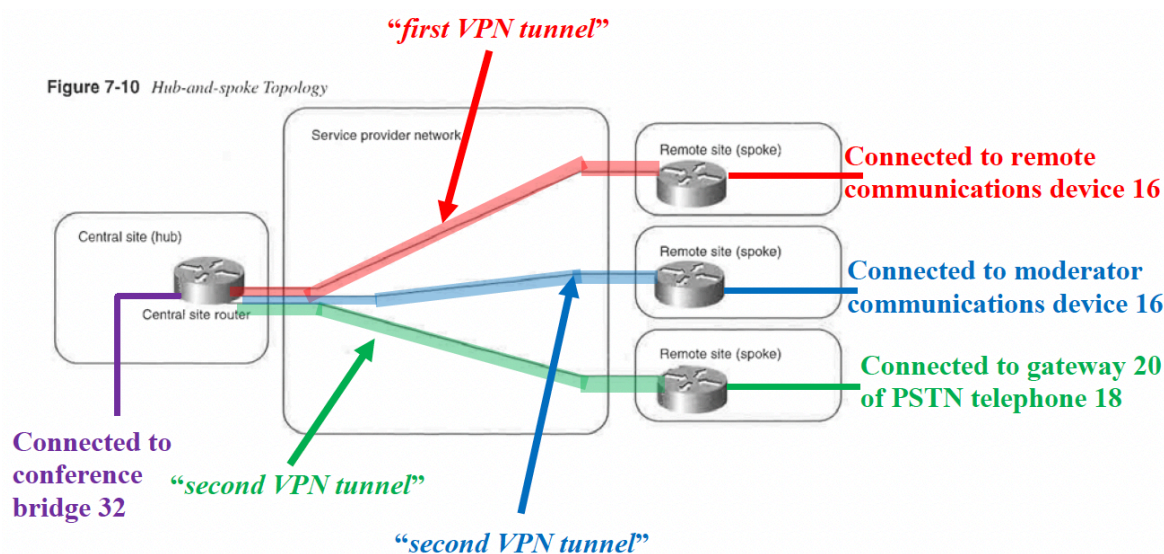
See [1.P]-[1.7] above.

[5.1] wherein the data communicated between the moderator and the PSTN client is transported through a second VPN tunnel.

As a threshold matter, claim 1, from which claim 5 depends, uses the phrases “*first audio data from a Public Switched Telephone Network (PSTN) client,*” “*second audio data from a moderator,*” “*first mixed audio data,*” and “*second mixed audio data,*” but does not recite the phrase “*data communicated between the moderator and the PSTN client.*” *Id.* ¶231. For the purposes of analysis for this IPR only, this Petition will proceed with the assumption that “*the data communicated between the moderator and the PSTN client*” corresponds to one of the data recited in claim 1. *Id.*

As discussed at [1.1]-[1.3], Knappe teaches communicating data between “*the moderator*” (e.g., a communications device 16 assigned to be moderator) and “*the PSTN client*” (e.g., a PSTN telephone 18). Ex_1003 ¶¶166-188, 232. As discussed at [1.5], a POSITA would have been motivated to combine Knappe with

Elliott and the VPN Textbook so that Knappe’s “data” could be “transported through ... a VPN tunnel.” *Id.* ¶¶ 200-202, 233. It would have been obvious to a POSITA to implement “a second VPN tunnel” for transporting “the data communicated between the moderator and the PSTN client” as shown in this annotated Figure 7-10 from the VPN Textbook:



Ex_1010, Fig. 7-10 (annotated); Ex_1003 ¶234. The VPN tunnel between the central (hub) router (e.g., conference bridge 32) and the remote (spoke) routers (e.g., moderator communications device 16 and PSTN telephones 18) may correspond to “a second VPN tunnel.” Ex_1003 ¶235. During a conference, data communications from the moderator communications device 16 may be transmitted to conference bridge 32 through the second VPN tunnel, as labeled in the above figure. Ex_1003 ¶¶234-236. After mixing by conference bridge 32, mixed audio data may be transmitted to the PSTN telephone 18, also through the

second VPN tunnel, as labeled in the above figure. *Id.* ¶236. Thus, the prior art renders obvious this limitation. *Id.* ¶¶230-237.

7. Claim 7

[7] The method of claim 1 further comprising

See [1.P]-[1.7] above.

[7.1] VoIP encoding the second mixed audio data before transmitting the second mixed audio data to the PSTN client.

As discussed at [1.6]-[1.7], Knappe discloses “*transmitting the second mixed audio data to the PSTN client.*” *Id.* ¶¶203-210, 240. Knappe further teaches that such data may be transmitted as VoIP packets. Ex_1006, 3:54-56 (“Telephony voice information is transmitted in the Voice over IP (VoIP) format.”); Ex_1003 ¶241. Knappe explains that during a conference, “VoIP packets received” by gateway 20 may be “converted into standard telephony signals for delivery to destination telephone 18 through PSTN 22.” Ex_1006, 4:17-28; Ex_1003 ¶ 241. Since Knappe communicates using VoIP packets and converts the VoIP packets to standard telephony signals for the PSTN telephone 18, and Knappe discloses “a codec (coder/decoder) converts audio,” a POSITA would understand Knappe’s conference bridge 32 to perform VoIP encoding of the second mixed audio data before transmission to PSTN telephone 18. Ex_1006, 4:9; Ex_1003 ¶242. Thus, Knappe discloses “*VoIP encoding the second mixed audio data before transmitting the second mixed audio data to the PSTN client.*” Ex_1003 ¶¶239-243.

B. Ground 2: Claims 3, 6, and 8-11 Are Obvious Over Knappe in View Elliott, the VPN Textbook, and Drell

1. Summary of Drell

Drell describes multi-point (MP) conferencing among a plurality of endpoints:

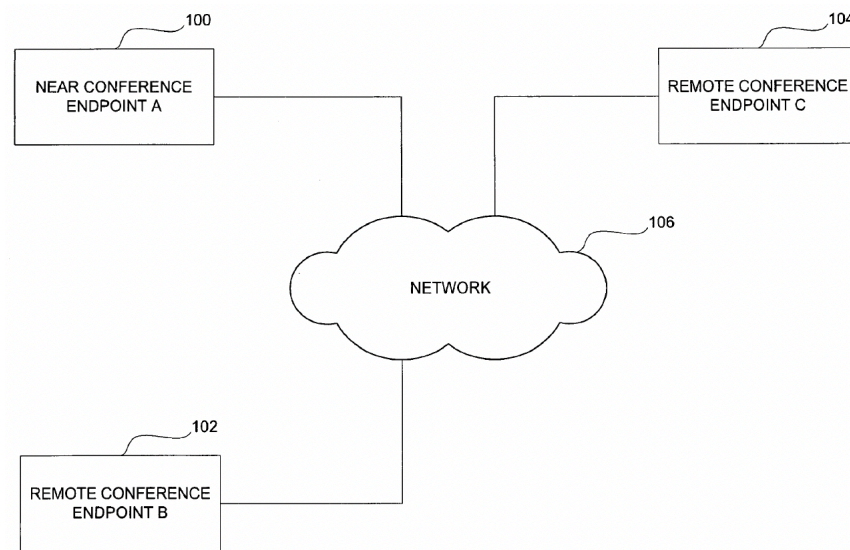


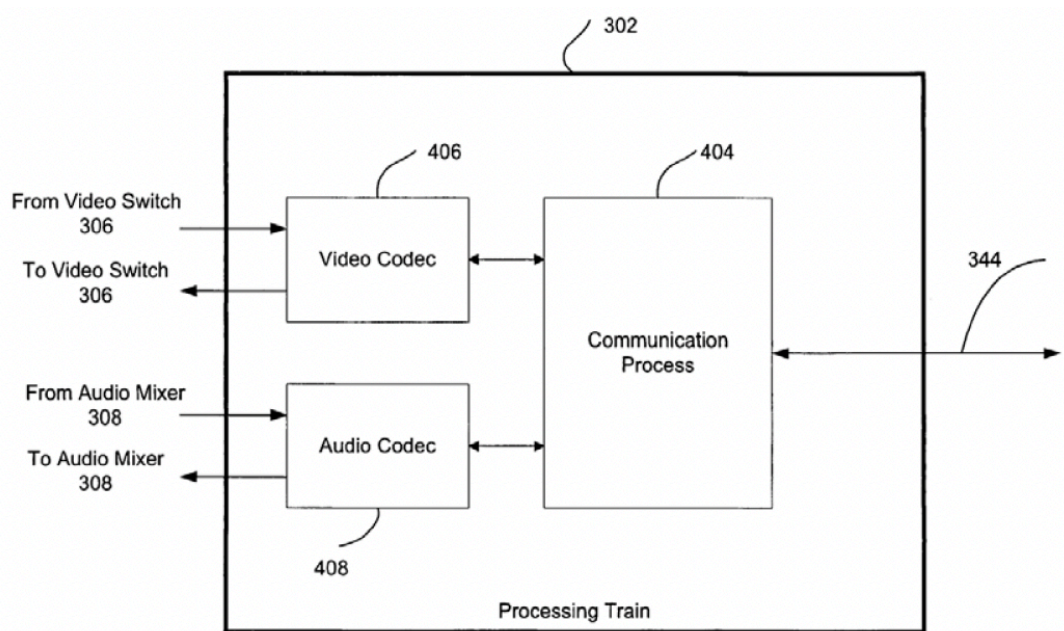
FIG. 1

Ex_1008, FIG. 1; Ex_1003 ¶ 245. Drell explains that “network 106” transmits “audio and video data between and among near conference endpoint 100 and remote conference endpoints 102 and 104.” Ex_1008, 3:20-23; Ex_1003 ¶246.

“Although a single network 106 is shown,” Drell “contemplates the use of two or more networks (for example, the PSTN and the Internet) to connect conference endpoints utilizing different communication protocols.” Ex_1008, 3:33-38; Ex_1003 ¶246.

Drell explains that centrally located conference bridges can have significant cost disadvantages and thus teaches use of “MP conferencing application 234” that is “software-based” and located at a “near conference point 100.” Ex_1008, 1:5-2:8, 4:6-20; Ex_1003 ¶247. Drell explains that “MP conferencing application 234 is configured to instantiate a processing train for each remote conference endpoint” where the “processed audio and video data streams are combined with each other and with locally generated audio and video streams.” Ex_1008, 4:21-29. The “audio and video streams are thereafter distributed to remote conferencing endpoints 102 and 104.” *Id.*, 4:29-31; Ex_1003 ¶248.

Drell further teaches that MP conferencing application 234 uses “a set of codecs” and may “comprise an H.320 process (ISDN-based) or H.323 (packet-based) process.” *Id.*, 2:17-33, FIG. 4; Ex_1003 ¶249.



Ex_1008, FIG. 4; Ex_1003 ¶249. “Audio codec 408 receives the audio data stream ... and applies ... decoding in accordance with a standard ... or proprietary audio compression algorithm” and “video codec 406 receives the video data stream and applies ... decoding in accordance with a standard ... or proprietary video compression algorithm.” Ex_1008, 6:18-27; Ex_1003 ¶250.

Drell further discloses its MP conferencing application 234 includes audio mixing module 308 and a “processing train 302 [that] may include a data codec for coding and encoding still images and the like received from or transmitted to remote conference endpoints 102 and 104.” Ex_1008, 6:53-56; Ex_1003 ¶251.

“In the transmission mode, processing train 302 is configured to receive the audio and video data streams output by audio mixing module 308 and video switching/continuous presence module 306,” along with conferencing data such as the aforementioned “still images,” and “encode[] and combine[] to form a mixed encoded audio/video data stream.” Ex_1008, 5:47-54, 6:55; FIG. 3. Annotated Figure 3 of Drell, below, highlights the audio mixing module 308 and the processing train 302, which a POSITA would have understood to be able to encode and mix “the audio and video data streams” with collaboration data, such as “still images,” to generate mixed data:

MP conferencing application 234 includes audio mixing module 308 and processing train 302

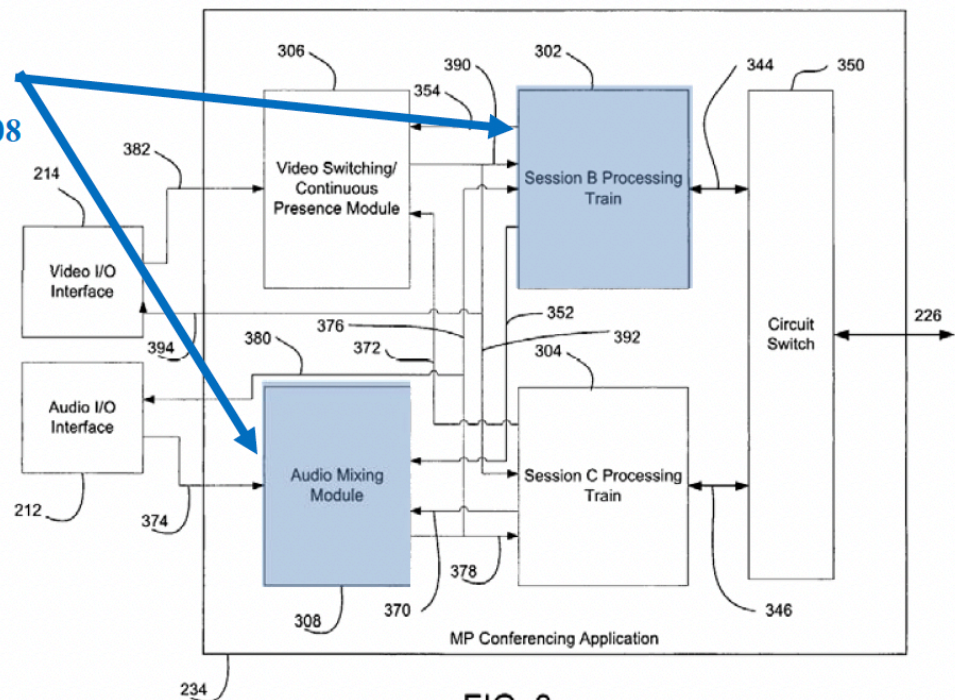


FIG. 3

Ex_1008, FIG. 3 (annotated), 5:47-54, 6:55; Ex_1003 ¶252.

Finally, Drell teaches “[a] video monitor 208 and loudspeaker 210 present images and speech of the remote images and speech.” Ex_1008, 3:45-48; FIG. 2; Ex_1003 ¶253

2. Reasons to Combine Knappe and Drell

A POSITA would have been motivated to combine the teachings of Knappe and Drell to produce the obvious, beneficial, and predictable results of having a user act as moderator as in the Knappe system and implementing various aspects of Knappe’s conferencing system as software and with an audio codec using a standard or proprietary compression algorithm to reduce the amount of data being transmitted. Ex_1003 ¶254. Knappe and Drell are analogous prior art both

pertaining to the same field of multiparty conferencing between Internet-based and PSTN users. *Id.* ¶255.

A POSITA understood that it was routine and would have recognized it is desirable to have a person (e.g. a leader or a moderator) to control the conference. *Id.* ¶255.

Indeed, Knappe teaches that at least “one participant may be given the priority of moderator for a particular conference call” to “control the flow and/or organization of the conference.” Ex. 1006, 8:52-56; 1:53-55; Ex_1003 ¶255. This would be desirable because frequently in business conferencing, at the time and even now, one person is in charge of the conference. Ex_1003 ¶255. Consistent with this understanding, Knappe teaches that at least “one participant may be given the priority of moderator for a particular conference call” to “control the flow and/or organization of the conference.” Ex_1006, 8:52-56; 1:53-55.

Knappe also teaches that its “call manager 30 and conference bridge 32” may “have their have their functionality distributed across and/or at the periphery of the network 14.” *Id.*, 4:31-33. Drell further teaches a “MP conferencing application” that performs functions similar to Knappe’s call manager 30 and conference bridge 32 implemented as software at a conference endpoint. Ex_1008, 1:65-2:6, 6:63-66 (“The present invention advantageously utilizes software-based processing of video and audio data streams to implement a multi-point

conferencing capability in a conference endpoint.”). Thus, it would have been obvious to a POSITA that the moderator communications device 16 of Knappe could implement the call manager and conference bridge 32 functionality (like Drell’s MP conferencing application) as software at the moderator. Ex_1003 ¶256.

A POSITA also would have been motivated to implement Drell’s “[a]udio codec” capable of encoding/decoding “in accordance with a standard ... or proprietary audio compression algorithm” and “video codec 406” capable of encoding/decoding “in accordance with a standard ... or proprietary video compression algorithm” with Knappe’s “communication devices 16 compris[ing] ... [a] codec (coder/decoder).” Ex_1008, 6:18-45; Ex_1006, 4:5-11; Ex_1003 ¶258. This would have the benefit of reducing the amount of data transmitted over Knappe’s network 14. Ex_1003 ¶258. The beneficial results would have been predictable and a POSITA would have had a reasonable expectation of success given Knappe’s mention of a codec and Drell’s discussion of codecs for encoding, decoding, and compressing in the audio and video conferencing context. *Id.* ¶¶259-260.

The combination of Knappe and Drell is merely the ordinary use of a common technique (implementing conference bridging and moderator functionality in software) in the system of Knappe where a participant is designated as moderator to yield predictable results of initiating, organizing,

controlling the conference through software at a conference endpoint and reducing the cost of having a separate conference bridge. *Id.* ¶257.

Petitioner’s proposed combination permits, but does not require, physical incorporation of elements from Drell into Knappe’s system. *See In re Mouttet*, 686 F.3d at 1322; Ex_1003 ¶261.

3. Claim 3

[3.P] The method of claim 1

See [1.P]-[1.7] above.

[3.1] wherein the receiving first audio data from the PSTN client includes receiving IP packets and decoding the IP packets containing the first audio data from the PSTN client at a moderator computer system.

As discussed at [1.1], Knappe teaches that conference bridge 32 receives audio data from the PSTN telephone 18 through gateway 20, disclosing “*receiving first audio data from the PSTN client.*” Ex_1003 ¶¶166-173, 264. Further, as discussed at [1.2], Knappe teaches a moderator communications device 16, which may be implemented as a computer system. Ex_1003 ¶¶175-179, 264.

Knappe further teaches that the audio data from the PSTN telephone is transmitted a VoIP packets. Ex_1006, 3:54-58 (Telephony voice information is **transmitted in the Voice over IP (VoIP) format**”), 4:21-22 (“At the gateway 20, the signals are **converted to IP packets in the VoIP format**”); Ex_1003 ¶265.

Knappe also explains that its conference bridge receives and converts the packets

from the PSTN. Ex_1006, 6:16-21 (“The input **converters receive input packets** . . . and **convert the packet** from the native format of the participant device 16 to a **standard format of the conference bridge 32**”); Ex_1003 ¶266. A POSITA would have understood that as part of converting the received input packets, Knappe’s converters 54 decode the IP packets since in Knappe “converting” of audio is performed with a codec (coder/decoder). Ex_1006, 4:9 (“[A] codec (coder/decoder) converts audio”); Ex_1003 ¶267. Therefore, Knappe discloses *“receiving IP packets and decoding the IP packets containing the first audio data from the PSTN client.”* Id. ¶268.

Knappe further recognizes that the implementation of the conference bridge 32 is not limited to a central facility. *See e.g.* Ex_1006, 4:31-33 (“The call manager 30 and the conference bridge 32 may be located in a central facility or have their functionality distributed across and/or at the periphery of the network 14.”); Ex_1003 ¶269. Drell supplements this idea and teaches that a conference bridge like that in Knappe could be implemented in software at a conference endpoint. Ex_1008, 3:39-41, 4:11-13 (describing near conference point 100 as having “a call manager application 232, and the MP conference application 234”), FIG. 2; Ex_1003 ¶¶269-272. Thus, it would have been obvious to a POSITA to distribute Knappe’s conference bridge 32 functionality as software located at a conference endpoint because doing so would reduce the cost of facilitating

conference communications. Ex_1003¶ 272. Further, it would have been obvious that in the combination, Knappe's communications devices 16, including the device designated as the moderator (as discussed at [1.2]), would include a conference bridge implemented as software similar to Drell's conference endpoint. *Id.*

Drell also teaches that conferencing application 234 (which could be implemented in Knappe's moderator computer) receives packets containing audio data from remote endpoints, including PSTN telephones. Ex_1008, 4:21-31 ("The processing trains [of MP conference application 234] process **audio and video streams received** from remote conference endpoints 102 and 104"); Ex_1003 ¶273. Drell's conferencing application then decodes the received packets. Ex_1008, 6:18-27 ("[a]udio codec 408 [that] receives the audio data stream ... and applies ... decoding in accordance with a standard ... or proprietary audio compression algorithm"), 2:26-33 (describing use of an "H.323 (packet-based) process" where "audio and video signals are thereafter respectively delivered to audio and video codecs for decoding"); Ex_1003 ¶274. A POSITA would have understood that Drell's disclosure of decoding received packetized audio corresponds to converting received audio as Knappe discloses. Ex_1003 ¶274.

In view of Drell's teaching, it would have been obvious to a POSITA for Knappe's conference bridge 32 (implemented as software in the moderator

communications device 16) to decode received IP packet data, because it would allow for mixing of the decoded audio data for each participant (as discussed at [1.4], [1.6], and [2.1]) and for playing mixed audio data to the moderator (as discussed at [2.2]). Ex_1003 ¶275. During operation, the conference bridge 32, implemented as software within the moderator communications device 16, would receive IP packets and convert the audio signals from the remote participants, including audio from the PSTN telephone 18. *Id.* ¶276. Thus, the prior art teaches that receiving audio data from the PSTN telephone 18 includes receiving IP packets and converting the IP packets containing the audio data from the PSTN telephone 18 at the moderator communications device 16, which renders obvious this limitation. *Id.* ¶277.

4. Claim 6

[6.P] The method of claim 1 further comprising

See [1.P]-[1.7].

[6.1] encoding the first mixed audio data into a first compressed audio data before transmitting the first mixed audio data to the remote client.

As discussed at [1.4]-[1.5], Knappe teaches mixing “*the first mixed audio data*” and “*transmitting the first mixed audio data to the remote client.*” Ex_1003 ¶¶194-202, 280. Drell teaches using a compression algorithm to compress mixed audio data before transmission to a remote endpoint. Ex_1008, 6:33-45 (“In the transit mode ... **[a]udio codec 408 encodes the audio data stream output by**

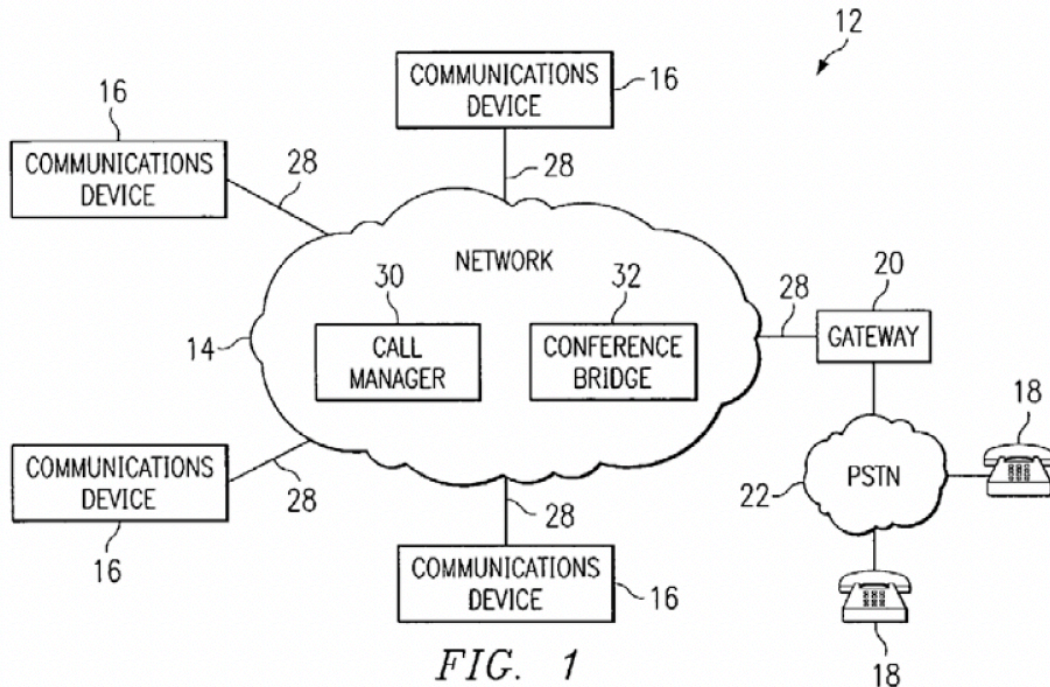
audio mixing module 308 ... using a standard or proprietary audio **compression algorithm** ... and delivers the encoded audio data ...”); Ex_1003 ¶281; *see also id.* ¶282. Since Knappe also discloses “a codec (coder/decoder) converts audio,” (*see* [3.1]) it would have been obvious to a POSITA to use in Knappe a codec that compresses the mixed audio data before transmission to a remote communications device 16 because this would result in the known benefit of reducing the amount of data transmitted over the network. *Id.* ¶283.

Thus, the prior art renders obvious this limitation. *Id.* ¶¶279-284.

5. Claim 8

[8.P] A system for supporting a multi-participant audio/video conference call comprising:

As discussed at [1.P], Knappe discloses “a method and system for improving ... a multiparty communication session ... [that] allow[s] multiple people to participate in ... a ‘conference call.’” Ex_1006, 1:8-27, 8:46-10:50, FIG. 8; Ex_1003 ¶¶163-165, 286. As shown in Figure 1, Knappe’s system can support multiple participants in an audio/video conference call:

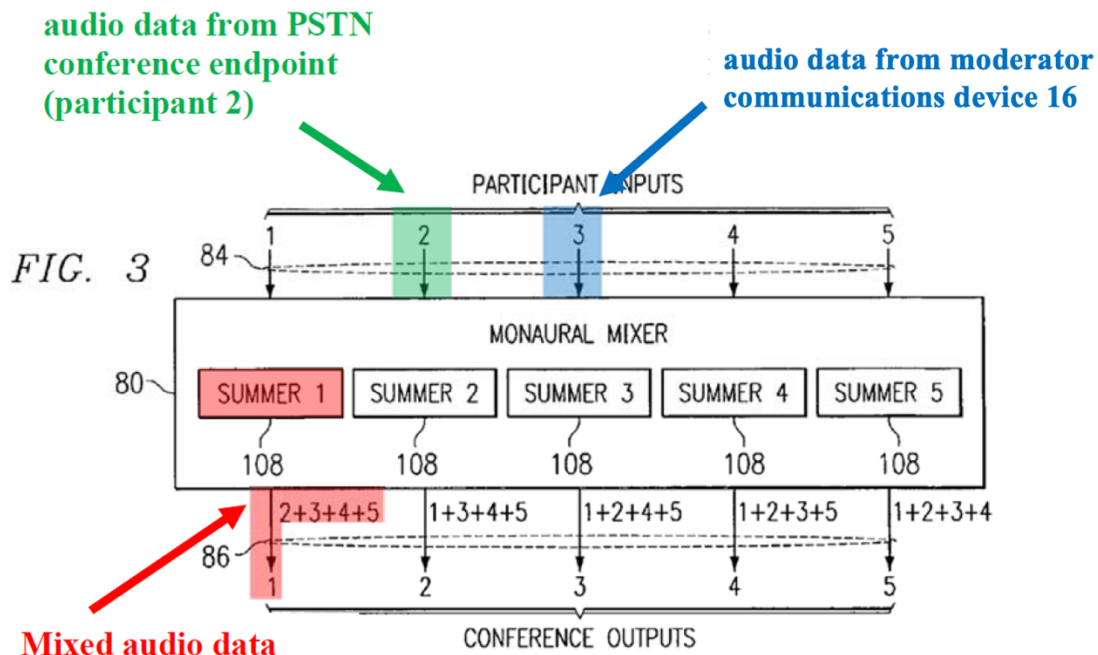


Ex_1006, FIG. 1. Specifically, Knappe explains that “conference bridge 32 provides conference call and other suitable audio, video, and/or real-time multi-party communication sessions between communication devices 16.” *Id.*, 4:58-61. Thus, a POSITA would understand Knappe to support a multi-participate audio/video conference call. Ex_1003 ¶287. Thereby, Knappe teaches the preamble of claim 8. *Id.* ¶¶163-165, 285-288.

[8.1.1] a first mixer configured to mix audio data from a Public Switched Telephone Network (PSTN) client with audio data from a moderator;

As discussed at [1.1], Knappe discloses receiving audio data from a PSTN client. *Id.* ¶¶166-173, 290. As discussed at [1.2], Knappe discloses receiving audio from participant communication devices 16, one of which can be designated

as a moderator. *Id.* ¶¶174-179, 290. As discussed at [1.4], Knappe discloses “mixing the first audio data from the PSTN client with the second audio data from the moderator into a first mixed audio data.” *Id.* ¶¶195-199, 291. Specifically, as discussed in [1.4] and shown in annotated Figure 3 of Knappe below, mixers 58, 80, and 100 of conference bridge 32 mixes audio data from PSTN telephone 18 with audio data from the moderator communication device 16:



Ex_1006, FIG. 3 (annotated); Ex_1003 ¶¶196-198, 291.

In addition, as discussed at [3.1], Drell teaches a conferencing system using MP conferencing application 234 which includes an audio mixing module 308, which can be implemented with Knappe’s system:

MP conferencing application 234 includes audio mixing module 308

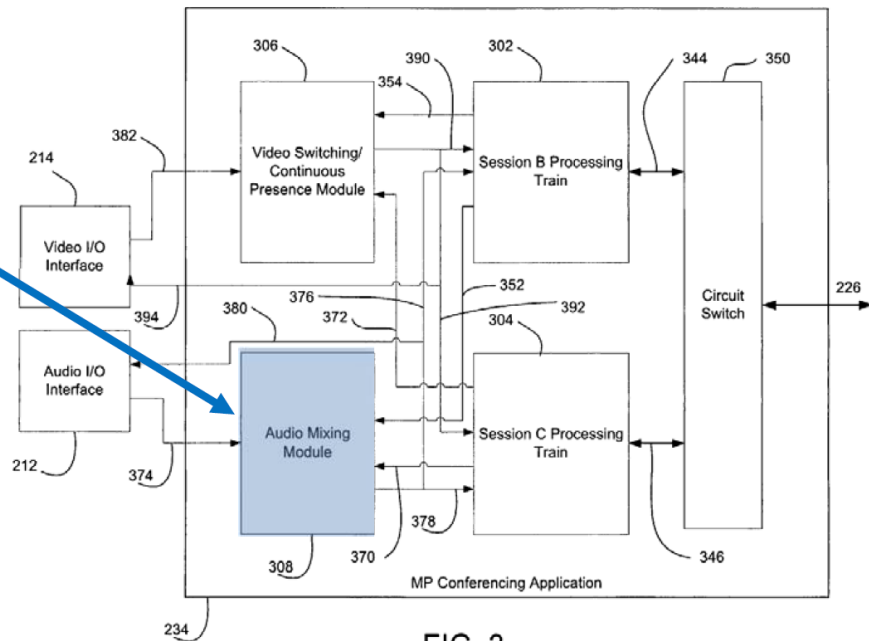


FIG. 3

Ex_1008, FIG. 3 (annotated); Ex_1003 ¶292. The “[a]udio mixing module 308 is configured to combine audio data received from remote conference endpoints . . . with locally generated audio data . . . [where] audio mixing module 308 may simply mix the received audio input data streams.” Ex_1008, 5:8-23; Ex_1003 ¶292. A POSITA would understand that the audio mixing module 308 mixes the locally generated audio (which as discussed at [3.1] can be Knappe’s moderator) with audio data received from a remote PSTN conference endpoint. Ex_1003 ¶292. Thus, the prior art renders obvious this limitation. *Id.* ¶¶289-293.

[8.1.2] and collaboration data from the moderator into a mixed data.

As discussed at [1.3.1], a POSITA would understand that the “audio, video, voice, data and other suitable types of real-time . . . traffic” communicated between conference participants in Knappe would include data to be used for collaboration,

including well-known collaboration data such as T.120 standard data. Ex_1003 ¶295. As discussed at [1.3.1], Elliott teaches that collaboration services, such as collaborative audio and video conferencing, can be provided over the Internet, such that a POSITA would understand that Knappe's conferencing system could include sharing of collaboration data. *Id.*

Drell also teaches that the conferencing application 234 has a processing train 302 that utilizes a data codec to encode "still images" for transmission to conference endpoints. Ex_1008, 6:53-56; Ex_1003 ¶296. A POSITA would understand Drell's "still images" to be "*collaboration data*," as claimed. Ex_1003 ¶¶296-297. It would have been obvious to a POSITA for Knappe's moderator to also be the source of still images, since this would allow for greater variety of collaboration among the participants. *Id.* Thus, a POSITA would have understood Knappe, individually and in combination with Elliott and Drell, to teach receiving "*collaboration data*" from a participant designated as the moderator. *Id.* ¶297.

Drell further explains the that encoded data is multiplexed and output as a single stream by the communication process. Ex_1008 at 2:45-48 ("the audio and video signals are encoded by the audio/video codecs and multiplexed into a **single data stream**"); Ex_1003 ¶298. Since the communication process outputs a "single stream," it would have been obvious to a POSITA that Drell's collaboration data (such as the encoded still images) are combined with the mixed audio data (from

the audio mixing module, discussed at [8.1.1]) for transmission to remote conference endpoints. Ex_1003 ¶299. A POSITA would have understood the audio mixing module 308 and the processing train 302 of Drell, shown below, to encode and mix “the audio and video data streams” with collaboration data, such as “still images,” to generate mixed data:

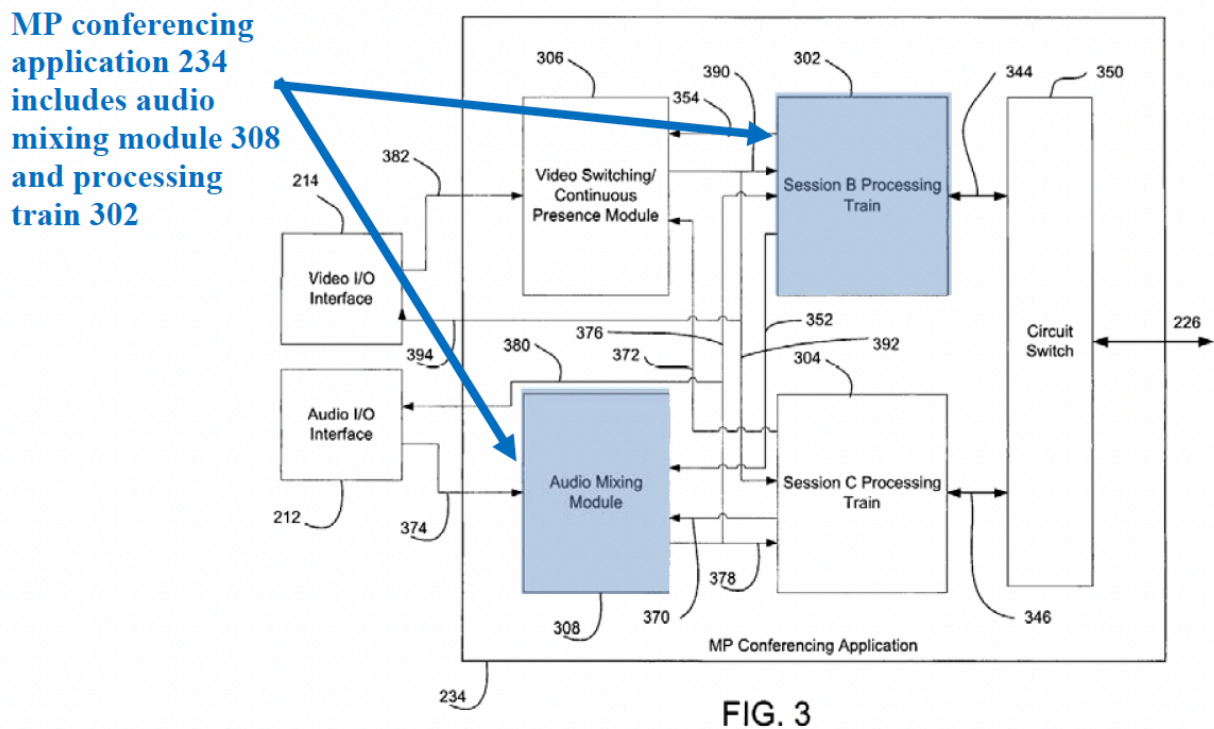


FIG. 3

Ex_1008, FIG. 3 (annotated); Ex_1003 ¶299.

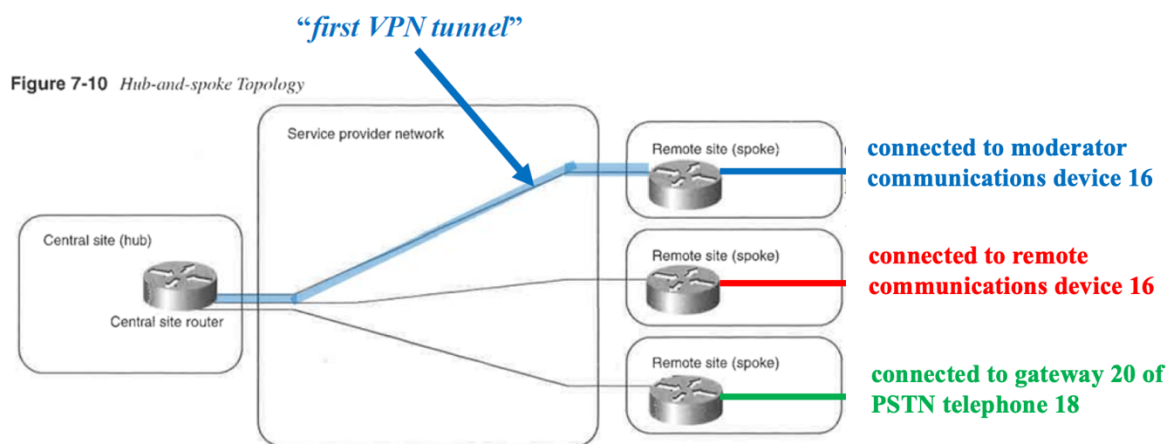
Thus, the prior art teaches using a processing train that further mixes the mixed audio data with collaboration data (e.g., still image data, application data, web browsing data, and other collaboration data per the T.120 standard) from the moderator into a mixed data, which renders obvious this limitation. Ex_1003 ¶¶294-300.

[8.2] a transport output configured to transmit the first mixed data to at least one remote client through a first Virtual Private Network (VPN) tunnel;

As discussed at [8.1.1]-[8.1.2], Knappe, in view of Elliott and Drell, teaches generating mixed data. *Id.* ¶¶289-300, 302. The phrase “*first mixed data*,” however, is not previously recited in the claim. *Id.* ¶303. For the purposes of analysis for this IPR only, this Petition will proceed with the assumption that “*first mixed data*” corresponds to the “*mixed data*” discussed in [8.1.1]-[8.1.2], above. *Id.*

As discussed at [1.5], Knappe also discloses “transmitting audio ... between source and destination endpoints,” including “a conference output stream for delivery to the participant.” Ex_1006, 3:24-29; 6:65-67; Ex_1003 ¶¶200-202, 304. Drell also discloses that its “processing trains ... process the combined signals according to a transmit mode for distribution to the remote endpoints over the network.” Therefore, Drell’s processing trains are able to transmit “*the first mixed data*” generated in [8.1] by the combination of Knappe, Drell, and Elliott. Ex_1008, Abstract, 5:48-54, 6:53-56; Ex_1003 ¶305. Thus, a POSITA would understand that when the collaboration data is mixed with the mixed audio data, as discussed at [8.1.1]-[8.1.2], the mixed data could be transmitted via a transport output as taught in Knappe and Drell. Ex_1003 ¶306.

As discussed at [1.3.2] and [1.5], Knappe can be implemented to transmit data (e.g., from conference bridge 32 to any of communications devices 16) through a VPN tunnel when combined with Elliott and the VPN Textbook. Ex_1003 ¶¶189-194, 200-202, 307. As noted at [1.3.2] when implementing VPN tunneling in Knappe, it would have been obvious to a POSITA to connect each of Knappe's devices located at remote sites to a separate remote router (spoke) that is connected to a router (hub) because this common topology provides a less complex, cost effective implementation. Accordingly, it would have been obvious to a POSITA that Knappe's remote communications device 16, moderator communications device 16, and gateway 20 connected to the PSTN telephone 18 would be connected to different remote site (spoke) routers, with a separate central site (hub) router, as illustrated below in annotated Figure 7-10:



Ex. 1010, Fig. 7-10 (annotated); Ex_1003 ¶308.

As shown, the remote site router (spoke) of the moderator communications device 16 (which as discussed at [3.1] can implement the conference bridge mixing functionality as software per Drell) can transmit data to the remote communications device 16 through a central hub using a “first VPN tunnel.” Thus, a POSITA would have understood Knappe (in combination with the teachings in Drell) in view of Elliott and/or the VPN Textbook to also teach the further limitation of transmitting the first mixed data “*through a first Virtual Private Network (VPN) tunnel.*” Ex_1003 ¶¶301-309.

[8.3] a second mixer configured to mix audio data from the moderator with audio data from the remote client connected to the Internet into a first mixed audio data; and

As discussed at [1.3.1], Knappe discloses receiving audio data from a remote client. As discussed at [1.6], Knappe discloses “*mixing the second audio data from the moderator with the third audio data from the remote client into a ... mixed audio data.*” Ex_1003 ¶¶203-206, 311. As discussed at [1.6], this mixing of audio data may be accomplished by one or more of Knappe’s mixers 58, 80, and 100, thereby rendering this element obvious. *Id.*

[8.4] a VoIP encoder to encode the first mixed audio data into a compressed audio data;

As discussed at [7.1], Knappe discloses “*VoIP encoding the second mixed audio data,*” which is audio data from the moderator mixed with audio data from the remote client. Ex_1003 ¶¶ 239-243, 313. As noted, a POSITA would have

understood Knappe’s disclosure of “[t]elephony voice information [being] transmitted in the Voice over IP (VoIP) format,” the “VoIP packets [being] encapsulated in real-time transport protocol (RTP) packets ... or other suitable packets for transmission over the network 14” to indicate “*a VoIP encoder*” encoding said VoIP packets. *Id.* ¶313.

As discussed at [6.1], Drell teaches using a compression algorithm to compress mixed audio data before transmission to a remote endpoint. Ex_1008, 6:18-45 (disclosing “[a]udio codec” capable of encoding/decoding “in accordance with a standard ... or proprietary audio compression algorithm”); Ex_1003 ¶¶279-284, 314. A POSITA would have recognized the benefit of compressing the mixed audio data generated using Knappe’s mixers (e.g., mixers 58, 80, and 100) for bandwidth savings in transmission. Ex_1003, ¶315. Therefore, a POSITA would have understood the combination of Knappe and Drell to disclose “*a VoIP encoder to encode the first mixed audio data into a compressed audio data.*” *Id.* ¶¶312-316.

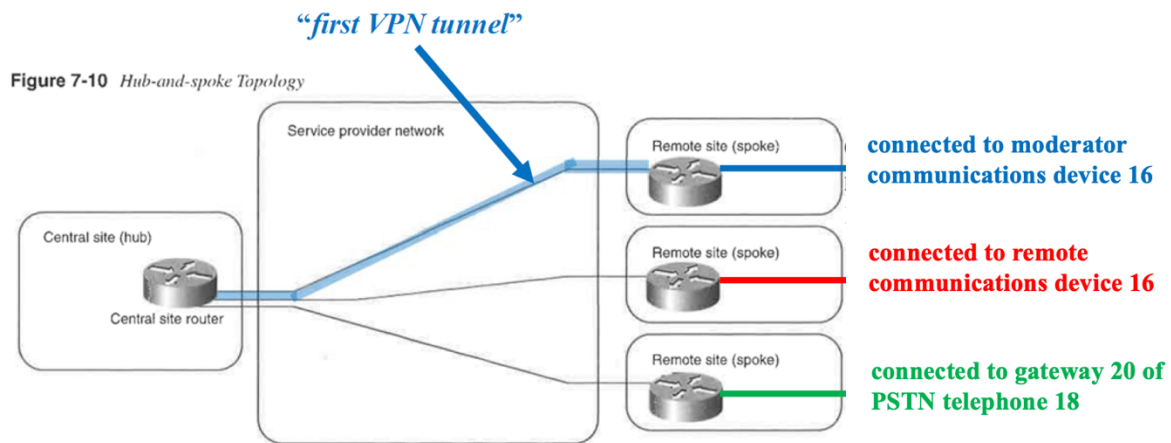
[8.5] the transport output further configured to transmit the compressed audio data to the PSTN client through the first VPN tunnel.

As discussed at [7.1] and [8.4], Knappe and Drell in combination disclose a VoIP encoder that encodes the first mixed audio data (audio data from the moderator mixed with audio data from the remote client) into “*the compressed audio data.*” *Id.* ¶318. Specifically, Drell teaches that “[a]udio code 408 encodes

the audio data stream output by audio mixing module 308 . . . using a standard or proprietary compression algorithm . . . and delivers the encoded audio data to communication process 404” for transmission. Ex_1008, 6:39-45; *see also id.*, 6:46-62, FIG. 5 (describing process for transmission of data following communication process 404); Ex_1003 ¶319.

As discussed at [1.7], Knappe in view of Elliott and the VPN Textbook discloses transmitting the mixed audio data (the audio data from the moderator mixed with audio data from the remote client) to the PSTN. Ex_1003 ¶318. Further, as discussed at [1.3.2] and [8.2], it would have been obvious to a POSITA to implement VPN tunneling when transmitting data in Knappe. *Id.* ¶320. Specifically, the remote site router (spoke) of the moderator communications device 16, which as discussed at [3.1] can implement the conference bridge mixing functionality as software per Drell, can transmit data to the PSTN client through a

central hub using the “first VPN tunnel”:



Ex_1010, Fig. 7-10 (annotated); Ex_1003 ¶320. Thus, a POSITA would have understood the prior art renders obvious this limitation. Ex_1003 ¶¶317-321.

6. Claim 9

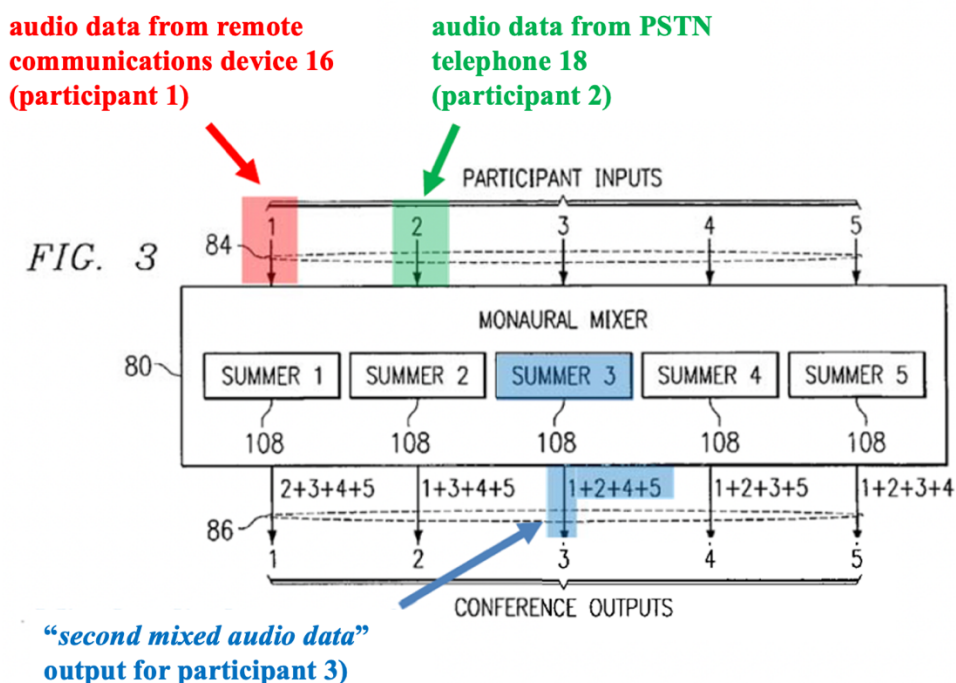
[9.P] The system of claim 8 further comprising:

See [8.P]-[8.5].

[9.1] a third mixer configured to mix the audio data from the PSTN client with the audio data from the remote client into a second mixed audio data; and

As discussed at [2.1], Knappe discloses “mixing the ... audio data from the PSTN client with the ... audio data from the remote client into a ... mixed audio data.” Ex_1003 ¶¶212-217, 324. As discussed at [1.2], [1.4], and [2.1], this mixing of audio data may be accomplished by one or more of Knappe’s mixers 58, 80, and 100. *Id.* ¶¶174-179, 195-199, 212-217, 325. For example, as shown in annotated Figure 3 below, Knappe discloses mixing “audio data from the PSTN

client” (e.g. participant 2) with “*the audio data from the remote client*” (e.g., participant 1), and providing that “*mixed audio data*” to “*the moderator*” (e.g., participant 3):

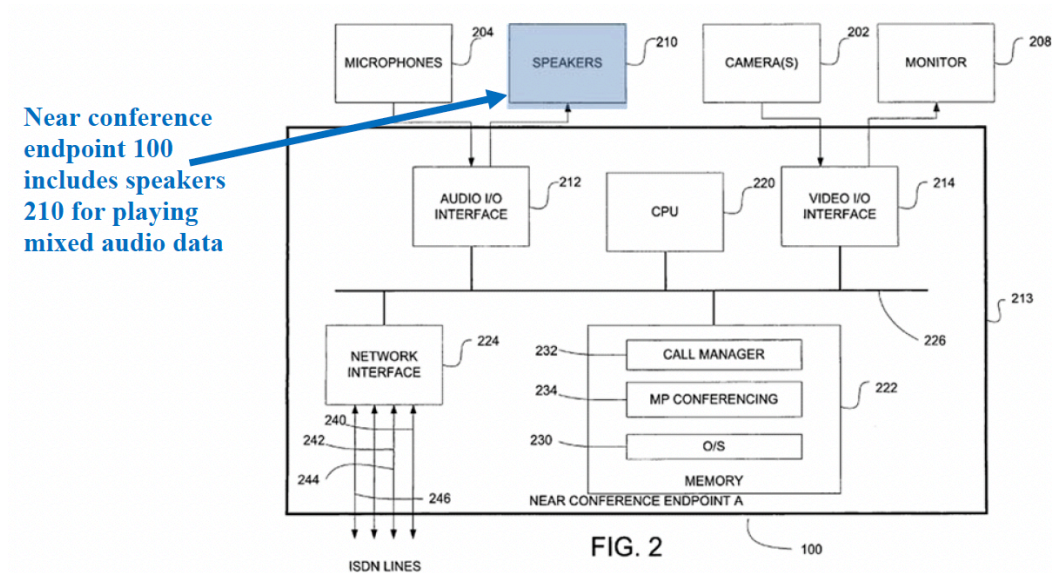


Ex_1006, FIG. 3 (annotated); Ex_1003 ¶326. Thus, Knappe discloses this limitation. Ex_1003 ¶¶323-327.

[9.2] a speaker configured to play the second mixed audio data to the moderator.

As discussed at [9.1], Knappe discloses generating “*the second mixed audio data.*” *Id.* ¶¶323-327, 329. Drell further discloses a “loudspeaker 210.” Ex_1008, 3:45-46, FIGS. 2-3; Ex_1003 ¶330. As discussed at [3.1], Drell’s near conference endpoint A, including its MP conferencing application 234, audio I/O interface

212, and other components, may be implemented with Knappe's conferencing system. Ex_1003 ¶¶269-277, 331. Drell's Figure 2 shows audio I/O interface 212 coupled to loudspeaker 210 and MP conferencing application 234, which has been shown above, as well as in the below Figure 3, to have audio mixing capabilities:



Ex_1008, FIG. 2 (annotated); Ex_1003 ¶332. Drell's Figure 3 shows that same audio I/O interface 212 able to receive mixed audio from audio mixing module 308 in MP conferencing application 234:

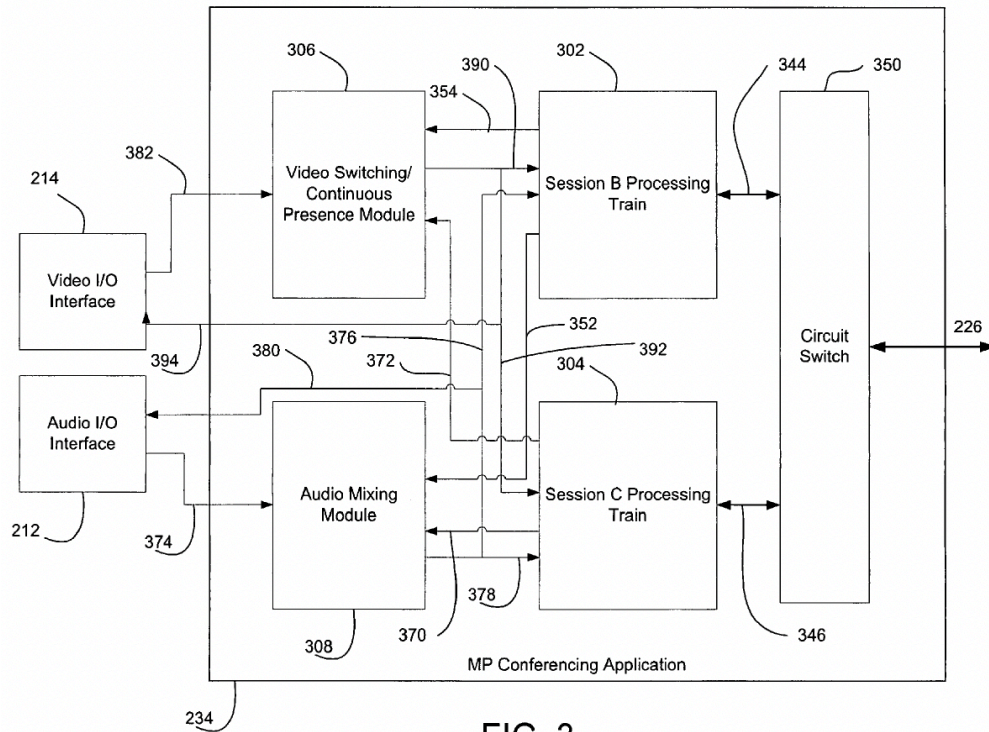


FIG. 3

Ex_1008, FIG. 3; Ex_1003 ¶¶333. Thus, the prior art renders obvious this limitation. Ex_1003 ¶¶328-334.

7. Claim 10

[10.P] The system of claim 8 further comprising

See [8.P]-[8.5].

[10.1] a VoIP decoder configured to receive IP packets containing the audio data from the PSTN client, and further configured to decode the IP packets containing the audio data from the PSTN client.

As discussed at [3.1], Knappe discloses a “codec (coder/decoder)” for “receiving first audio data from the PSTN client include[ing] receiving IP packets and decoding the IP packets containing the first audio data from the PSTN client.” See Ex_1006, 4:5-27, 6:3-21 (describing a “codec (coder/decoder)” to encode data

and “converters” to “receive input packets ... and convert the packet from the native format of the participants device 16 to a standard format of the conference bridge 32”); Ex_1003 ¶¶263-277, 337. Also, as discussed at [7.1], Knappe discloses “*VoIP encoding the second mixed audio data before transmitting the ... mixed audio data.*” Ex_1003 ¶¶239-243, 338. Therefore, the prior art renders obvious this limitation. *Id.* ¶¶336-339.

8. Claim 11

[11.P] The system of claim 8 further comprising

See [8.P]-[8.5].

[11.1] at least one audio decompressor configured to decode IP packets containing the audio data from the remote client.

As discussed at [1.3.1], Knappe discloses receiving audio data from a remote client. *Id.* ¶342. As discussed at [1.6], Knappe discloses “*mixing the second audio data from the moderator with the third audio data from the remote client into a ... mixed audio data.*” *Id.* Knappe also discloses that data, including audio from remote communications device 16, is transmitted as IP packets. *E.g.* Ex_1006, 4:3:53-54 (“In the Internet embodiment, the network 14 transmits information in Internet Protocol (IP) packets”); 4:9-17 (“digital encoded data is encapsulated into IP or other suitable packets for transmission over network 14”); Ex_1003 ¶343.

Knappe also discloses converting audio data into and back from IP packets. Ex_1006, 4:13-14 (“IP packets received from the network 14 are converted back

into analog signals and played to the user”); 6:16-21 (“The input converters receive input packets of a participant . . . and convert the packet from the native format of the participant’s device 16 to a standard format[.]. Conversely, the output converters . . . convert the conference output streams to the native format of participant’s devices.”); Ex_1003 ¶344. And as discussed at [3.1] and [4.1], it would have been obvious to a POSITA that Knappe’s conversion of IP packets would include “*decoding the IP packets containing the first audio data from the PSTN client.*” Ex_1003 ¶345.

As discussed at [3.1], Drell also explains that conferencing application 234 (which can be implemented at Knappe’s moderator communications device 16) uses audio codec 408 to decode packets. Ex_1008, 6:18-27 (“[a]udio codec 408 [that] receives the audio data stream . . . and applies . . . decoding in accordance with a standard . . . or proprietary audio compression algorithm.”); Ex_1003 ¶346. A POSITA would have understood that decoding audio using an audio compression algorithm is performing decompression and thus discloses an “*audio decompressor.*” Ex_1003 ¶¶346-347. Thus, the prior art renders obvious this limitation. *Id.* ¶348.

C. Ground 3: Claim 12 is Obvious Over Knappe in Combination with Elliott, VPN Textbook, Drell, and Hoke

1. Summary of Hoke

Hoke describes “computer systems for processing communications in a virtual private network.” Ex_1009, Abstract; Ex_1003 ¶350. Hoke further describes the compression and decompression of data communicated over VPN. *See, e.g.*, Ex_1009, FIGS. 3-4 & 7, 2:50, 5:40-42 (“Many of the elements may be configured using well-known structures, particularly those designated as relating to various compression or encryption techniques.”). Hoke explains that “[w]hen VPN traffic is sent between source and destination endstations ..., the VPN unit serving the source endstation processes the data packet, encrypts it, compresses it (if necessary), and ... [t]he receiving VPN unit authenticates, decrypts and decompresses the packet (as necessary) before forwarding it toward the destination endstation.” Ex_1009, 8:38-50; Ex_1003 ¶¶350-351.

2. Reasons to combine Knappe and Hoke

A POSITA would have been motivated to combine the teachings of Knappe and Hoke to produce the obvious, beneficial, and predictable results of compressing data before it is transmitted over a VPN tunnel. Ex_1003 ¶352. As discussed above, Knappe teaches a communications system capable of transmitting conferencing data—“audio, video, voice, data and other suitable types of ... traffic”—between source and destination endpoints. Ex_1006, 3:26-29. Also in

the context of conferencing, Elliott teaches the sharing of collaboration data, including “collaborative Web browsing, audio conferencing, video conferencing, and application sharing” and the use of “a virtual private network (VPN) ... which uses IP or IP-tunneling” for the transmission of “voice, fax, data, video” and “[o]ther types of collaboration ... data” during a conference. Ex_1007, 1:43-44 2:66-3:1, 4:29-34, 6:53-54; Ex_1003 ¶¶353-354. As explained above, a POSITA would have been motivated to combine the conferencing system of Knappe to include sharing of collaboration data through a VPN as taught in Elliot and the VPN Textbook. Ex_1003 ¶¶136-162, 355. Thus, in considering Knappe, in view of Elliott and the VPN Textbook, a POSITA would have considered the teachings of Hoke, as well, for teachings relevant to VPN tunneling implementations. *Id.*

In particular, a POSITA would have recognized the desirability of compressing data before it is transmitted over a network, as it was well known in the art that compressing data would reduce bandwidth usage and improve performance. *Id.* ¶356. As Hoke describes the compression and decompression of data communicated over VPN, it would have been obvious to a POSITA to implement the compression techniques in Hoke with an implementation of Knappe with the VPN tunneling of Elliott and the VPN Textbook. *Id.* ¶357; *see* Ex_1009, FIGS. 3-4 & 7, 2:50, 5:40-42. For example, Hoke’s description of a VPN unit that performs both encryption and compression would have suggested to a POSITA to

include a compression feature in the VPN routers described in the VPN Textbook. Ex. 1003 ¶358.

It would have been obvious to a POSITA to compress data transmitted over VPN tunnels implemented in Knappe, in view of Elliott and the VPN Textbook, to minimize bandwidth usage and improve performance, particularly for lower speed links. *Id.* ¶359. A POSITA would expect this application of known compression algorithms with known VPN techniques to successfully yield predictable and beneficial results. *Id.* ¶360; *see also* Ex_1009, 10:16-18 (“Many compression algorithms are known”), 5:41-48 (“Many of the elements may be configured using well-known structures, particularly those designated as relating to various compression or encryption techniques. ... those of ordinary skill in the art will be able to implement the particular methods without undue experimentation.”). Any modifications needed to Knappe’s system to accommodate the teachings of Hoke would have been within the level of ordinary skill in the art. Ex_1003, ¶361. Petitioner’s proposed combination permits, but does not require, physical incorporation of elements from Hoke into Knappe. *See In re Mouttet*, 686 F.3d at 1322; Ex_1003 ¶361.

3. Claim 12

[12.P] The system of claim 8 further comprising

See [8.P]-[8.5].

[12.1] an encoder configured to encode the mixed data into a compressed data before transmitting the compressed data to the remote client.

As discussed at [6.1] and [8.4], Knappe, in combination with Drell, discloses a VoIP encoder configured to encode mixed audio data into a compressed audio data before transmitting the compressed data to the remote client. Ex_1003 ¶¶279-284, 312-316, 364. As discussed at [8.1.1]-[8.1.2], Knappe and Drell in combination also disclose the generation of “*the mixed data.*” *Id.* ¶¶289-300, 365.

To the extent the claim requires applying a single compression process to a mixture of audio and collaboration data, Hoke teaches compressing data that is transmitted over VPN tunnels. *Id.* ¶366. Specifically, Hoke teaches that “VPN units may transmit **VPN traffic in... tunnel... format.**” Ex. 1009, 9:6-10. In tunnel mode, “a data packet is routed to/from the VPN units serving the communicating endstations or remote clients.” *Id.*, 9:10-12. In this context, Hoke teaches that a “VPN unit serving the source endstation processes the data packet, encrypts it, **compresses it** (if necessary)[.]” *Id.* at 8:39-44; Ex_1003 ¶366.

Hoke further teaches that “[m]any compression algorithms are known, but in one embodiment of the invention LZW compression is used.” Ex_1009, 10:16-18. The compression algorithm, as taught by Hoke, corresponds to the claimed “*encoder.*” Ex_1003 ¶367. It would have been obvious to a POSITA to compress the mixed data before it is transmitted over the first VPN tunnel (as discussed at [8.2]), because compressing the mixed data would minimize the utilized bandwidth

and improve the perceived performance of lower-speed links. *Id.* ¶368; *see also id.* ¶¶352-361. In combination, packetized mixed data from Knappe's moderator communications device 16 would be sent to a router (as taught by the VPN Textbook), as discussed at [8.2], and as described in this section would apply a compression algorithm (as taught by Hoke) before transmission over the first VPN tunnel. *Id.* ¶369.

Thus, the prior art teaches that the mixed audio data from the moderator communications device 16 is provided to the VPN router, which compressed the mixed data before transmitting the compressed data over the VPN tunnel to the remote communications device 16, which renders obvious this limitation.

Ex_1003 ¶¶363-370.

IX. CONCLUSION

For the reasons given above, Petitioner has demonstrated a reasonable likelihood of success in establishing the unpatentability of claims 1-12 of the '332 patent and respectfully requests institution of IPR for each of the grounds presented herein and cancellation of the challenged claims.

Respectfully submitted,

Date: December 16, 2022

By: Hector Ribera/

Hector J. Ribera (Reg. No. 54,397)
548 Market St., Suite 36117, San
Francisco, CA 94104
Email: hector@martonribera.com
Tel: (415) 360-2511

Attorneys for Petitioner

CERTIFICATE OF WORD COUNT PURSUANT TO 37 C.F.R. § 42.24(D)

The undersigned hereby certifies that the word count for the above Petition for *inter partes* review is 13,778, which is less than the 14,000 allowed under 37 CFR § 42.24(a)(1)(i).

Date: December 16, 2022

By: /Hector Ribera/

Hector J. Ribera (Reg. No. 54,397)
Marton Ribera Schumann & Chang LLP
548 Market St., Suite 36117, San
Francisco, CA 94104 Email:
hector@martonribera.com
Tel: (415) 360-2511

Attorneys for Petitioner

CERTIFICATE OF SERVICE PURSUANT TO
37 C.F.R. § 42.6 AND 37 C.F.R. § 42.105

Pursuant to 37 C.F.R. § 42.6(e)(4) and 42.105(b), the undersigned certifies that on December 16, 2022, a complete and entire copy of this Petition for *inter partes* review and all supporting exhibits were provided via FedEx First, to the Patent Owner by serving the correspondence address of record as follows:

ATTN: THOMAS MARTIN
AMEDEO FERRARO
WESLEY MEINERDING
MARTIN & FERRARO, LLP
1557 LAKE O'PINES STREET, NE
HARTVILLE, OH 44632

Date: December 16, 2022

By: /Hector Ribera/
Hector J. Ribera (Reg. No. 54,397)
Marton Ribera Schumann & Chang LLP
548 Market St., Suite 36117, San
Francisco, CA 94104
Email: hector@martonribera.com
Tel: (415) 360-2512

Attorneys for Petitioner