US009270572B2

(12) **United States Patent**
Koganti et al.

(10) **Patent No.: US 9,270,572 B2**
(45) **Date of Patent: Feb. 23, 2016**

(54) **LAYER-3 SUPPORT IN TRILL NETWORKS**

(75) Inventors: **Phanidhar Koganti**, Sunnyvale, CA (US); **Anoop Ghanwani**, Rocklin, CA (US); **Suresh Vobbilisetty**, San Jose, CA (US); **Rajiv Krishnamurthy**, San Jose, CA (US); **Nagarajan Venkatesan**, San Jose, CA (US); **Shunjia Yu**, San Jose, CA (US)

(73) Assignee: **BROCADE COMMUNICATIONS SYSTEMS INC.**, San Jose, CA (US)

(*) Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 0 days.

(21) Appl. No.: **13/312,903**

(22) Filed: **Dec. 6, 2011**

(65) **Prior Publication Data**

US 2012/0281700 A1 Nov. 8, 2012

Related U.S. Application Data

(60) Provisional application No. 61/481,643, filed on May 2, 2011, provisional application No. 61/503,265, filed on Jun. 30, 2011.

(51) **Int. Cl.**
H04L 12/18 (2006.01)
H04L 12/751 (2013.01)

(52) **U.S. Cl.**
CPC ***H04L 45/02*** (2013.01)

(58) **Field of Classification Search**
CPC H04L 43/50; H04L 43/0852; H04L 43/08; H04B 17/003
USPC 370/392, 351, 352
See application file for complete search history.

(56) **References Cited**

U.S. PATENT DOCUMENTS

5,390,173	A	2/1995	Spinney
5,802,278	A	9/1998	Isfeld
5,878,232	A	3/1999	Marimuthu
5,959,968	A	9/1999	Chin

(Continued)

FOREIGN PATENT DOCUMENTS

CN	102801599 A	11/2012
EP	0579567	5/1993

(Continued)

OTHER PUBLICATIONS

U.S. Appl. No. 12/312,903 Office Action dated Jun. 13, 2013.

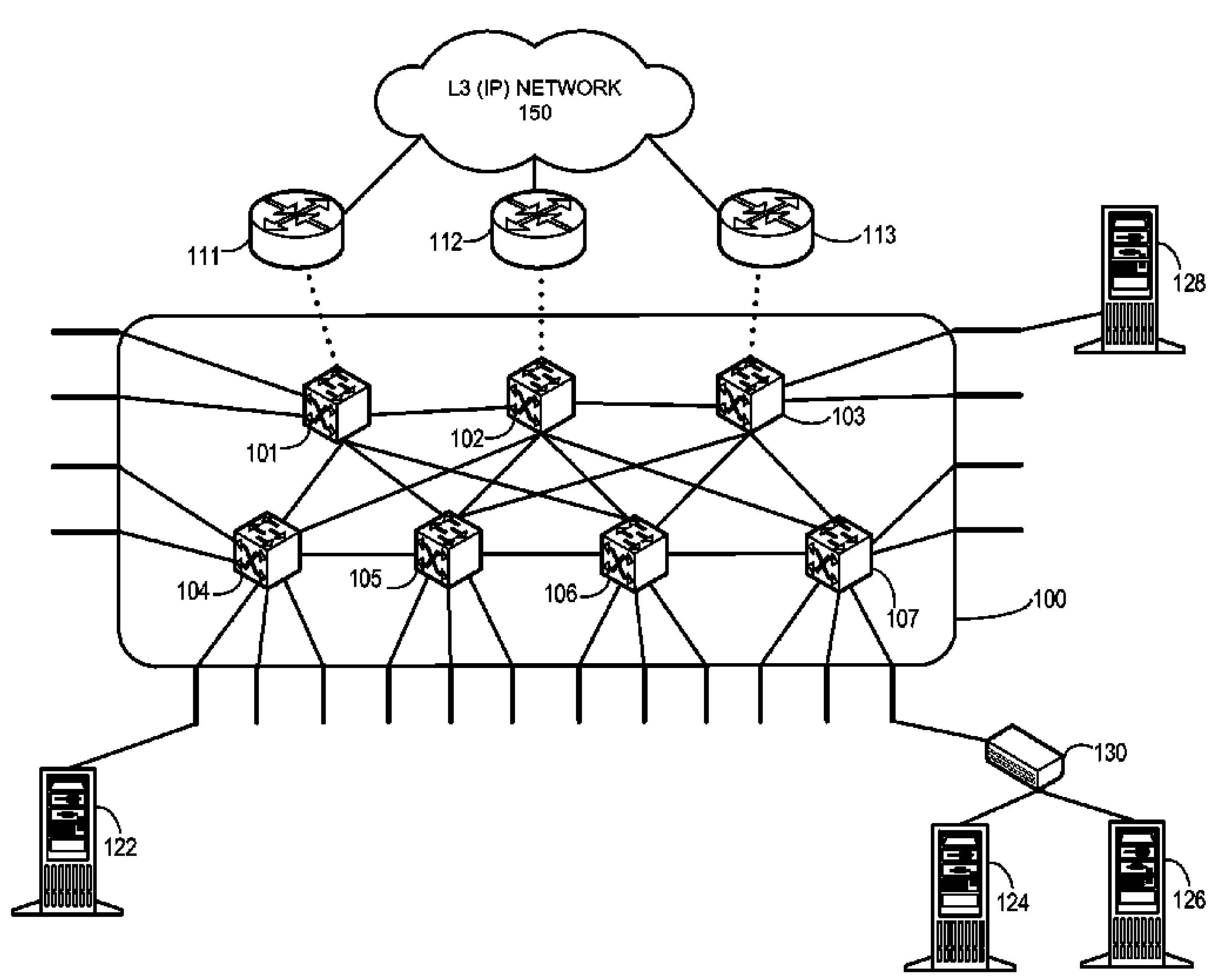
(Continued)

Primary Examiner — Andrew Lai
Assistant Examiner — Zhiren Qin
(74) *Attorney, Agent, or Firm* — Shun Yao; Park, Vaughan, Fleming & Dowler LLP

(57) **ABSTRACT**

One embodiment of the present invention provides a switch. The switch includes an IP header processor and a forwarding mechanism. The IP header processor identifies a destination IP address in a packet encapsulated with an inner Ethernet header, a TRILL header, and an outer Ethernet header. The forwarding mechanism determines an output port and constructs a new header for the packet based on the destination IP address. The switch also includes a packet processor which determines whether (1) an inner destination media access control (MAC) address corresponds to a local MAC address assigned to the switch; (2) a destination RBridge identifier corresponds to a local RBridge identifier assigned to the switch; and (3) an outer destination MAC address corresponds to the local MAC address.

20 Claims, 13 Drawing Sheets



(56) **References Cited**

U.S. PATENT DOCUMENTS

5,973,278 A 10/1999 Wehrill, III
5,983,278 A 11/1999 Chong
6,041,042 A 3/2000 Bussiere
6,064,671 A * 5/2000 Killian 370/389
6,085,238 A 7/2000 Yuasa
6,104,696 A 8/2000 Kadambi
6,185,214 B1 2/2001 Schwartz
6,185,241 B1 2/2001 Sun
6,438,106 B1 8/2002 Pillar
6,498,781 B1 12/2002 Bass
6,542,266 B1 4/2003 Phillips
6,633,761 B1 10/2003 Singhal
6,771,610 B1 8/2004 Seaman
6,873,602 B1 3/2005 Ambe
6,937,576 B1 8/2005 DiBenedetto
6,956,824 B2 10/2005 Mark
6,957,269 B2 10/2005 Williams
6,975,581 B1 12/2005 Medina
6,975,864 B2 12/2005 Singhal
7,016,352 B1 3/2006 Chow
7,061,877 B1 * 6/2006 Gummalla H04B 1/525 370/278
7,173,934 B2 2/2007 Lapuh
7,197,308 B2 3/2007 Singhal
7,206,288 B2 4/2007 Cometto
7,310,664 B1 12/2007 Merchant
7,313,637 B2 12/2007 Tanaka
7,315,545 B1 1/2008 Chowdhury et al.
7,316,031 B2 1/2008 Griffith
7,330,897 B2 2/2008 Baldwin
7,380,025 B1 5/2008 Riggins
7,397,794 B1 7/2008 Lacroute
7,430,164 B2 9/2008 Bare
7,453,888 B2 11/2008 Zabihi
7,477,894 B1 1/2009 Sinha
7,480,258 B1 1/2009 Shuen
7,508,757 B2 3/2009 Ge
7,558,195 B1 7/2009 Kuo
7,558,273 B1 7/2009 Grosser, Jr.
7,571,447 B2 8/2009 Ally
7,599,901 B2 10/2009 Mital
7,688,736 B1 3/2010 Walsh
7,688,960 B1 3/2010 Aubuchon
7,690,040 B2 3/2010 Frattura
7,706,255 B1 4/2010 Kondrat et al.
7,716,370 B1 5/2010 Devarapalli
7,720,076 B2 5/2010 Dobbins
7,729,296 B1 6/2010 Choudhary
7,787,480 B1 8/2010 Mehta
7,792,920 B2 9/2010 Istvan
7,796,593 B1 9/2010 Ghosh
7,808,992 B2 10/2010 Homchaudhuri
7,836,332 B2 11/2010 Hara
7,843,906 B1 11/2010 Chidambaram et al.
7,843,907 B1 11/2010 Abou-Emara
7,860,097 B1 12/2010 Lovett
7,898,959 B1 3/2011 Arad
7,924,837 B1 * 4/2011 Shabtay H04L 12/1886 370/235
7,937,756 B2 5/2011 Kay
7,945,941 B2 5/2011 Sinha
7,949,638 B1 5/2011 Goodson
7,957,386 B1 6/2011 Aggarwal
8,018,938 B1 9/2011 Fromm
8,027,354 B1 9/2011 Portolani
8,054,832 B1 11/2011 Shukia
8,068,442 B1 11/2011 Kompella
8,078,704 B2 12/2011 Lee
8,102,781 B2 1/2012 Smith
8,102,791 B2 1/2012 Tang
8,116,307 B1 2/2012 Thesayi
8,125,928 B2 2/2012 Mehta
8,134,922 B2 3/2012 Elangovan
8,155,150 B1 4/2012 Chung
8,160,063 B2 4/2012 Maltz
8,160,080 B1 4/2012 Arad
8,170,038 B2 5/2012 Belanger
8,194,674 B1 6/2012 Pagel
8,195,774 B2 6/2012 Lambeth
8,204,061 B1 6/2012 Sane
8,213,313 B1 7/2012 Doiron
8,213,336 B2 7/2012 Smith
8,230,069 B2 7/2012 Korupolu
8,239,960 B2 8/2012 Frattura
8,249,069 B2 8/2012 Raman
8,270,401 B1 9/2012 Barnes
8,295,291 B1 10/2012 Ramanathan
8,295,921 B2 10/2012 Wang
8,301,686 B1 10/2012 Appajodu
8,339,994 B2 12/2012 Gnanasekaran
8,351,352 B1 1/2013 Eastlake, III
8,369,335 B2 2/2013 Jha
8,369,347 B2 2/2013 Xiong
8,392,496 B2 3/2013 Linden
8,462,774 B2 6/2013 Page
8,465,774 B2 6/2013 Page
8,467,375 B2 6/2013 Blair
8,520,595 B2 8/2013 Yadav
8,599,850 B2 12/2013 Jha
8,599,864 B2 12/2013 Chung
8,615,008 B2 12/2013 Natarajan
8,706,905 B1 4/2014 McGlaughlin
8,724,456 B1 5/2014 Hong
8,806,031 B1 8/2014 Kondur
8,826,385 B2 9/2014 Congdon
8,937,865 B1 1/2015 Kumar
2001/0005527 A1 6/2001 Vaeth
2001/0055274 A1 12/2001 Hegge
2002/0019904 A1 2/2002 Katz
2002/0021701 A1 2/2002 Lavian
2002/0039350 A1 4/2002 Wang
2002/0054593 A1 * 5/2002 Morohashi H04W 92/20 370/386
2002/0091795 A1 7/2002 Yip
2003/0041085 A1 2/2003 Sato
2003/0123393 A1 7/2003 Feuerstraeter
2003/0174706 A1 9/2003 Shankar
2003/0189905 A1 10/2003 Lee
2003/0216143 A1 11/2003 Roese
2004/0001433 A1 1/2004 Gram
2004/0003094 A1 1/2004 See
2004/0010600 A1 1/2004 Baldwin
2004/0049699 A1 3/2004 Griffith
2004/0057430 A1 3/2004 Paavolainen
2004/0117508 A1 6/2004 Shimizu
2004/0120326 A1 6/2004 Yoon
2004/0156313 A1 8/2004 Hofmeister et al.
2004/0165595 A1 8/2004 Holmgren
2004/0165596 A1 8/2004 Garcia
2004/0213232 A1 10/2004 Regan
2005/0007951 A1 1/2005 Lapuh
2005/0044199 A1 2/2005 Shiga
2005/0074001 A1 4/2005 Mattes
2005/0094568 A1 5/2005 Judd
2005/0094630 A1 5/2005 Valdevit
2005/0122979 A1 6/2005 Gross
2005/0157645 A1 7/2005 Rabie et al.
2005/0157751 A1 7/2005 Rabie
2005/0169188 A1 8/2005 Cometto
2005/0195813 A1 9/2005 Ambe
2005/0207423 A1 9/2005 Herbst
2005/0213561 A1 9/2005 Yao
2005/0220096 A1 10/2005 Friskney
2005/0265356 A1 12/2005 Kawarai
2005/0278565 A1 12/2005 Frattura
2006/0007869 A1 1/2006 Hirota
2006/0018302 A1 1/2006 Ivaldi
2006/0023707 A1 2/2006 Makishima et al.
2006/0034292 A1 2/2006 Wakayama
2006/0059163 A1 3/2006 Frattura
2006/0062187 A1 3/2006 Rune
2006/0072550 A1 4/2006 Davis
2006/0083254 A1 4/2006 Ge
2006/0098589 A1 5/2006 Kreeger

(56) **References Cited**

U.S. PATENT DOCUMENTS

2006/0140130	A1	6/2006	Kalkunte	
2006/0168109	A1	7/2006	Warmenhoven	
2006/0184937	A1	8/2006	Abels	
2006/0221960	A1	10/2006	Borgione	
2006/0235995	A1	10/2006	Bhatia	
2006/0242311	A1	10/2006	Mai	
2006/0245439	A1	11/2006	Sajassi	
2006/0251067	A1	11/2006	Desanti	
2006/0256767	A1*	11/2006	Suzuki	H04L 45/586 370/351
2006/0265515	A1	11/2006	Shiga	
2006/0285499	A1	12/2006	Tzeng	
2006/0291388	A1	12/2006	Amdahl	
2007/0036178	A1	2/2007	Hares	
2007/0083625	A1	4/2007	Chamdani	
2007/0086362	A1	4/2007	Kato	
2007/0094464	A1	4/2007	Sharma	
2007/0097968	A1	5/2007	Du	
2007/0098006	A1	5/2007	Parry	
2007/0116224	A1	5/2007	Burke	
2007/0116422	A1	5/2007	Burke	
2007/0156659	A1	7/2007	Lim	
2007/0177525	A1	8/2007	Wijnands	
2007/0177597	A1	8/2007	Ju	
2007/0183313	A1	8/2007	Narayanan	
2007/0211712	A1	9/2007	Fitch	
2007/0258449	A1	11/2007	Bennett	
2007/0274234	A1	11/2007	Kubota	
2007/0289017	A1	12/2007	Copeland, III	
2008/0052487	A1	2/2008	Akahane	
2008/0065760	A1	3/2008	Damm	
2008/0080517	A1	4/2008	Roy	
2008/0095160	A1	4/2008	Yadav	
2008/0101386	A1	5/2008	Gray	
2008/0112400	A1	5/2008	Dunbar et al.	
2008/0133760	A1	6/2008	Berkvens et al.	
2008/0159277	A1	7/2008	Vobbilisetty	
2008/0172492	A1	7/2008	Raghunath	
2008/0181196	A1	7/2008	Regan	
2008/0181243	A1	7/2008	Vobbilisetty	
2008/0186981	A1	8/2008	Seto	
2008/0205377	A1	8/2008	Chao	
2008/0219172	A1	9/2008	Mohan	
2008/0225852	A1	9/2008	Raszuk	
2008/0225853	A1	9/2008	Melman	
2008/0228897	A1	9/2008	Ko	
2008/0240129	A1	10/2008	Elmeleegy	
2008/0267179	A1	10/2008	Lavinge	
2008/0285458	A1	11/2008	Lysne	
2008/0285555	A1	11/2008	Ogasahara	
2008/0298248	A1	12/2008	Roeck	
2008/0304498	A1	12/2008	Jorgensen	
2008/0310342	A1	12/2008	Kruys	
2009/0022069	A1	1/2009	Khan	
2009/0037607	A1	2/2009	Farinacci	
2009/0042270	A1	2/2009	Dolly	
2009/0044270	A1	2/2009	Shelly	
2009/0067422	A1	3/2009	Poppe	
2009/0067442	A1	3/2009	Killian	
2009/0079560	A1	3/2009	Fries	
2009/0080345	A1	3/2009	Gray	
2009/0083445	A1	3/2009	Ganga	
2009/0092042	A1	4/2009	Yuhura	
2009/0092043	A1*	4/2009	Lapuh	H04L 49/70 370/228
2009/0106405	A1	4/2009	Mazarick	
2009/0116381	A1	5/2009	Kanda	
2009/0129384	A1	5/2009	Regan	
2009/0138577	A1	5/2009	Casado	
2009/0138752	A1	5/2009	Graham	
2009/0161584	A1	6/2009	Guan	
2009/0161670	A1	6/2009	Shepherd	
2009/0168647	A1	7/2009	Holness	
2009/0199177	A1	8/2009	Edwards	
2009/0204965	A1	8/2009	Tanaka	

2009/0213783	A1	8/2009	Moreton	
2009/0222879	A1	9/2009	Kostal	
2009/0232031	A1	9/2009	Vasseur	
2009/0245137	A1	10/2009	Hares	
2009/0245242	A1	10/2009	Carlson	
2009/0246137	A1	10/2009	Hadida	
2009/0252049	A1	10/2009	Ludwig	
2009/0252061	A1	10/2009	Small	
2009/0260083	A1	10/2009	Szeto	
2009/0279558	A1	11/2009	Davis	
2009/0292858	A1	11/2009	Lambeth	
2009/0316721	A1	12/2009	Kanda	
2009/0323708	A1	12/2009	Ihle	
2009/0327392	A1	12/2009	Tripathi	
2009/0327462	A1	12/2009	Adams	
2010/0027420	A1	2/2010	Smith	
2010/0046471	A1	2/2010	Hattori	
2010/0054260	A1	3/2010	Pandey	
2010/0061269	A1	3/2010	Banerjee	
2010/0074175	A1	3/2010	Banks	
2010/0097941	A1	4/2010	Carlson	
2010/0103813	A1	4/2010	Allan	
2010/0103939	A1	4/2010	Carlson	
2010/0131636	A1	5/2010	Suri	
2010/0158024	A1	6/2010	Sajassi	
2010/0165877	A1	7/2010	Shukia	
2010/0165995	A1	7/2010	Mehta	
2010/0168467	A1	7/2010	Johnston	
2010/0169467	A1	7/2010	Shukia	
2010/0169948	A1	7/2010	Budko	
2010/0182920	A1	7/2010	Matsuoka	
2010/0215049	A1	8/2010	Raza	
2010/0220724	A1	9/2010	Rabie	
2010/0226368	A1	9/2010	Mack-Crane	
2010/0226381	A1	9/2010	Mehta	
2010/0246388	A1	9/2010	Gupta	
2010/0257263	A1	10/2010	Casado	
2010/0271960	A1	10/2010	Krygowski	
2010/0272107	A1*	10/2010	Papp	H04L 29/12028 370/392
2010/0281106	A1	11/2010	Ashwood-Smith	
2010/0284414	A1	11/2010	Agarwal	
2010/0284418	A1	11/2010	Gray	
2010/0287262	A1	11/2010	Elzur	
2010/0287548	A1	11/2010	Zhou	
2010/0290473	A1	11/2010	Enduri	
2010/0299527	A1	11/2010	Arunan	
2010/0303071	A1	12/2010	Kotalwar	
2010/0303075	A1	12/2010	Tripathi	
2010/0303083	A1	12/2010	Belanger	
2010/0309820	A1	12/2010	Rajagopalan	
2010/0309912	A1	12/2010	Mehta	
2010/0329110	A1	12/2010	Rose	
2011/0019678	A1	1/2011	Mehta	
2011/0032945	A1	2/2011	Mullooly	
2011/0035489	A1	2/2011	McDaniel	
2011/0035498	A1	2/2011	Shah	
2011/0044339	A1	2/2011	Kotalwar	
2011/0044352	A1	2/2011	Chaitou	
2011/0064086	A1	3/2011	Xiong	
2011/0064089	A1	3/2011	Hidaka	
2011/0072208	A1	3/2011	Gulati	
2011/0085560	A1	4/2011	Chawla	
2011/0085563	A1	4/2011	Kotha	
2011/0110266	A1	5/2011	Li	
2011/0134802	A1	6/2011	Rajagopalan	
2011/0134803	A1	6/2011	Dalvi	
2011/0134925	A1	6/2011	Safrai	
2011/0142053	A1	6/2011	Van Der Merwe	
2011/0142062	A1	6/2011	Wang	
2011/0161494	A1	6/2011	McDysan	
2011/0161695	A1	6/2011	Okita	
2011/0188373	A1	8/2011	Saito	
2011/0194403	A1	8/2011	Sajassi	
2011/0194563	A1	8/2011	Shen	
2011/0228780	A1	9/2011	Ashwood-Smith	
2011/0231570	A1	9/2011	Altekar	
2011/0231574	A1	9/2011	Saunderson	
2011/0235523	A1	9/2011	Jha	

(56) **References Cited**

U.S. PATENT DOCUMENTS

2011/0243133	A9	10/2011	Villait
2011/0243136	A1	10/2011	Raman
2011/0246669	A1	10/2011	Kanada
2011/0255538	A1	10/2011	Srinivasan
2011/0255540	A1	10/2011	Mizrahi
2011/0261828	A1	10/2011	Smith
2011/0268120	A1	11/2011	Vobbilisetty
2011/0268125	A1	11/2011	Vobbilisetty
2011/0273988	A1	11/2011	Tourrilhes
2011/0274114	A1	11/2011	Dhar
2011/0280572	A1	11/2011	Vobbilisetty
2011/0286457	A1*	11/2011	Ee H04L 45/00 370/392
2011/0296052	A1	12/2011	Guo
2011/0299391	A1	12/2011	Vobbilisetty
2011/0299413	A1	12/2011	Chatwani
2011/0299414	A1	12/2011	Yu
2011/0299527	A1	12/2011	Yu
2011/0299528	A1	12/2011	Yu
2011/0299531	A1	12/2011	Yu
2011/0299532	A1	12/2011	Yu
2011/0299533	A1	12/2011	Yu
2011/0299534	A1	12/2011	Koganti
2011/0299535	A1*	12/2011	Vobbilisetty et al. 370/392
2011/0299536	A1	12/2011	Cheng et al.
2011/0317559	A1	12/2011	Kern
2011/0317703	A1	12/2011	Dunbar et al.
2012/0011240	A1	1/2012	Hara
2012/0014261	A1	1/2012	Salam
2012/0014387	A1	1/2012	Dunbar
2012/0020220	A1	1/2012	Sugita
2012/0027017	A1	2/2012	Rai
2012/0033663	A1	2/2012	Guichard
2012/0033665	A1	2/2012	Jacob Da Silva
2012/0033669	A1	2/2012	Mohandas
2012/0075991	A1	3/2012	Sugita
2012/0099567	A1	4/2012	Hart
2012/0099602	A1	4/2012	Nagapudi
2012/0106339	A1*	5/2012	Mishra H04L 43/106 370/235
2012/0131097	A1	5/2012	Baykal
2012/0131289	A1	5/2012	Taguchi
2012/0147740	A1	6/2012	Nakash
2012/0158997	A1	6/2012	Hsu
2012/0163164	A1	6/2012	Terry
2012/0177039	A1	7/2012	Berman
2012/0243359	A1	9/2012	Keesara
2012/0243539	A1	9/2012	Keesara
2012/0275347	A1	11/2012	Banerjee
2012/0294192	A1	11/2012	Masood
2012/0294194	A1	11/2012	Balasubramanian
2012/0320800	A1	12/2012	Kamble
2012/0320926	A1	12/2012	Kamath et al.
2012/0327766	A1	12/2012	Tsai et al.
2012/0327937	A1	12/2012	Melman et al.
2013/0003535	A1	1/2013	Sarwar
2013/0003737	A1	1/2013	Sinicrope
2013/0003738	A1	1/2013	Koganti
2013/0028072	A1	1/2013	Addanki
2013/0034015	A1	2/2013	Jaiswal
2013/0067466	A1	3/2013	Combs
2013/0070762	A1	3/2013	Adams
2013/0114595	A1	5/2013	Mack-Crane et al.
2013/0127848	A1	5/2013	Joshi
2013/0194914	A1	8/2013	Agarwal
2013/0219473	A1	8/2013	Schaefer
2013/0250951	A1	9/2013	Koganti
2013/0259037	A1	10/2013	Natarajan
2013/0272135	A1	10/2013	Leong
2013/0301642	A1	11/2013	Radhakrishnan
2014/0044126	A1	2/2014	Sabhanatarajan
2014/0105034	A1	4/2014	Sun

FOREIGN PATENT DOCUMENTS

EP	1398920	A2	3/2004
EP	1916807	A2	4/2008
EP	2001167	A1	12/2008
WO	2008056838		5/2008
WO	2009042919		4/2009
WO	2010111142	A1	9/2010
WO	2014031781		2/2014

OTHER PUBLICATIONS

U.S. Appl. No. 13/365,808 Office Action dated Jul. 18, 2013.
U.S. Appl. No. 13/365,993 Office Action dated Jul. 23, 2013.
U.S. Appl. No. 13/092,873 Office Action dated Jun. 19, 2013.
U.S. Appl. No. 13/184,526 Office Action dated May 22, 2013.
U.S. Appl. No. 13/184,526 Office Action dated Jan. 28, 2013.
U.S. Appl. No. 13/050,102 Office Action dated May 16, 2013.
U.S. Appl. No. 13/050,102 Office Action dated Oct. 26, 2012.
U.S. Appl. No. 13/044,301 Office Action dated Feb. 22, 2013.
U.S. Appl. No. 13/044,301 Office Action dated Jun. 11, 2013.
U.S. Appl. No. 13/030,688 Office Action dated Apr. 25, 2013.
U.S. Appl. No. 13/030,806 Office Action dated Dec. 3, 2012.
U.S. Appl. No. 13/030,806 Office Action dated Jun. 11, 2013.
U.S. Appl. No. 13/098,360 Office Action dated May 31, 2013.
U.S. Appl. No. 13/092,864 Office Action dated Sep. 19, 2012.
U.S. Appl. No. 12/950,968 Office Action dated Jun. 7, 2012.
U.S. Appl. No. 12/950,968 Office Action dated Jan. 4, 2013.
U.S. Appl. No. 13/092,877 Office Action dated Mar. 4, 2013.
U.S. Appl. No. 12/950,974 Office Action dated Dec. 20, 2012.
U.S. Appl. No. 12/950,974 Office Action dated May 24, 2012.
U.S. Appl. No. 13/092,752 Office Action dated Feb. 5, 2013.
U.S. Appl. No. 13/092,752 Office Action dated Jul. 18, 2013.
U.S. Appl. No. 13/092,701 Office Action dated Jan. 28, 2013.
U.S. Appl. No. 13/092,701 Office Action dated Jul. 3, 2013.
U.S. Appl. No. 13/092,460 Office Action dated Jun. 21, 2013.
U.S. Appl. No. 13/042,259 Office Action dated Mar. 18, 2013.
U.S. Appl. No. 13/042,259 Office Action dated Jul. 31, 2013.
U.S. Appl. No. 13/092,580 Office Action dated Jun. 10, 2013.
U.S. Appl. No. 13/092,724 Office Action dated Jul. 16, 2013.
U.S. Appl. No. 13/092,724 Office Action dated Feb. 5, 2013.
U.S. Appl. No. 13/098,490 Office Action dated Dec. 21, 2012.
U.S. Appl. No. 13/098,490 Office Action dated Jul. 9, 2013.
U.S. Appl. No. 13/087,239 Office Action dated May 22, 2013.
U.S. Appl. No. 13/087,239 Office Action dated Dec. 5, 2012.
U.S. Appl. No. 12/725,249 Office Action dated Apr. 26, 2013.
U.S. Appl. No. 12/725,249 Office Action dated Sep. 12, 2012.
Brocade Unveils "The Effortless Network", http://newsroom.brocade.com/press-releases/brocade-unveils-the-effortless-network--nasdaq-brcd-0859535, 2012.
Foundry FastIron Configuration Guide, Software Release FSX 04.2.00b, Software Release FWS 04.3.00, Software Release FGS 05.0.00a, Sep. 26, 2008.
FastIron and TurboIron 24X Configuration Guide Supporting FSX 05.1.00 for FESX, FWSX, and FSX; FGS 04.3.03 for FGS, FLS and FWS; FGS 05.0.02 for FGS-STK and FLS-STK, FCX 06.0.00 for FCX; and TIX 04.1.00 for TI24X, Feb. 16, 2010.
FastIron Configuration Guide Supporting Ironware Software Release 07.0.00, Dec. 18, 2009.
"The Effortless Network: HyperEdge Technology for the Campus LAN", 2012.
Narten, T. et al. "Problem Statement: Overlays for Network Virtualization", draft-narten-nvo3-overlay-problem-statement-01, Oct. 31, 2011.
Knight, Paul et al., "Layer 2 and 3 Virtual Private Networks: Taxonomy, Technology, and Standardization Efforts", IEEE Communications Magazine, Jun. 2004.
"An Introduction to Brocade VCS Fabric Technology", Brocade white paper, http://community.brocade.com/docs/DOC-2954, Dec. 3, 2012.
Kreeger, L. et al., "Network Virtualization Overlay Control Protocol Requirements", Draft-kreeger-nvo3-overlay-cp-00, Jan. 30, 2012.
Knight, Paul et al., "Network based IP VPN Architecture using Virtual Routers", May 2003.

(56) **References Cited**

OTHER PUBLICATIONS

Louati, Wajdi et al., "Network-based virtual personal overlay networks using programmable virtual routers", IEEE Communications Magazine, Jul. 2005.
U.S. Appl. No. 13/092,877 Office Action dated Sep. 5, 2013.
U.S. Appl. No. 13/044,326 Office Action dated Oct. 2, 2013.
Abawajy J. "An Approach to Support a Single Service Provider Address Image for Wide Area Networks Environment" Centre for Parallel and Distributed Computing, School of Computer Science Carleton University, Ottawa, Ontario, K1S 5B6, Canada.
Office Action for U.S. Appl. No. 13/425,238, filed Mar. 20, 2012, dated Mar. 12, 2015.
Office Action for U.S. Appl. No. 13/786,328, filed Mar. 5, 2013, dated Mar. 13, 2015.
Office Action for U.S. Appl. No. 14/577,785, filed Dec. 19, 2014, dated Apr. 13, 2015.
Mahalingam "VXLAN: A Framework for Overlaying Virtualized Layer 2 Networks over Layer 3 Networks" Oct. 17, 2013 pp. 1-22, Sections 1, 4 and 4.1.
Office action dated Apr. 30, 2015, U.S. Appl. No. 13/351,513, filed Jan. 17, 2012.
Office Action dated Apr. 1, 2015, U.S. Appl. No. 13/656,438, filed Oct. 19, 2012.
Office Action dated May 21, 2015, U.S. Appl. No. 13/288,822, filed Nov. 3, 2011.
Siamak Azodolmolky et al. "Cloud computing networking: Challenges and opportunities for innovations", IEEE Communications Magazine, vol. 51, No. 7, Jul. 1, 2013.
Office Action dated Apr. 1, 2015 U.S. Appl. No. 13/656,438, filed Oct. 19, 2012.
Office action dated Jun. 8, 2015, U.S. Appl. No. 14/178,042, filed Feb. 11, 2014.
Office Action Dated Jun. 10, 2015, U.S. Appl. No. 13/890,150, filed May 8, 2013.
"Switched Virtual Internetworking moved beyond bridges and routers", 8178 Data Communications Sep. 23, 1994, No. 12, New York.
S. Night et al., "Virtual Router Redundancy Protocol", Network Working Group, XP-002135272, Apr. 1998.
Eastlake 3rd., Donald et al., "RBridges: TRILL Header Options", Draft-ietf-trill-rbridge-options-00.txt Dec. 24, 2009.
J. Touch, et al., "Transparent Interconnection of Lots of Links (TRILL): Problem and Applicability Statement", May 2009.
Perlman, Radia et al., "RBridge VLAN Mapping", Draft-ietf-trill-rbridge-vlan-mapping-01.txt Dec. 4, 2009.
Brocade Fabric OS (FOS) 6.2 Virtual Fabrics Feature Frequently Asked Questions.
Perlman, Radia "Challenges and Opportunities in the Design of TRILL: a Routed layer 2 Technology", XP-002649647, 2009.
Nadas, S. et al., "Virtual Router Redundancy Protocol (VRRP) Version 3 for IPv4 and IPv6", Mar. 2010.
Perlman, Radia et al., "RBridges: Base Protocol Specification", draft-ietf-trill-rbridge-protocol-16.txt, Mar. 3, 2010.
Christensen, M. et al., "Considerations for Internet Group Management Protocol (IGMP) and Multicast Listener Discovery (MLD) Snooping Switches", May 2006.
Lapuh, Roger et al., "Split Multi-link Trunking (SMLT)", Oct. 2002.
Lapuh, Roger et al., "Split Multi-link Trunking (SMLT) draft-lapuh-network-smlt-08", 2008.
Office Action for U.S. Appl. No. 13/351,513, filed Jan. 17, 2012, dated Feb. 28, 2014.
Office Action for U.S. Appl. No. 13/598,204, filed Aug. 29, 2012, dated Feb. 20, 2014.
Office Action for U.S. Appl. No. 13/533,843, filed Jun. 26, 2012, dated Oct. 21, 2013.
Office Action for U.S. Appl. No. 13/312,903, filed Dec. 6, 2011, dated Nov. 12, 2013.
Office Action for U.S. Appl. No. 13/092,873, filed Apr. 22, 2011, dated Nov. 29, 2013.
Office Action for U.S. Appl. No. 13/184,526, filed Jul. 16, 2011, dated Dec. 2, 2013.
Office Action for U.S. Appl. No. 13/042,259, filed Mar. 7, 2011, dated Jan. 16, 2014.
Office Action for U.S. Appl. No. 13/092,580, filed Apr. 22, 2011, dated Jan. 10, 2014.
Office Action for U.S. Appl. No. 13/092,877, filed Apr. 22, 2011, dated Jan. 6, 2014.
Zhai F. Hu et al. "RBridge: Pseudo-Nickname; draft-hu-trill-pseudonode-nickname-02.txt", May 15, 2012.
Huang, Nen-Fu et al., "An Effective Spanning Tree Algorithm for a Bridged LAN", Mar. 16, 1992.
Office Action dated Jun. 6, 2014, U.S. Appl. No. 13/669,357, filed Nov. 5, 2012.
Office Action dated Feb. 20, 2014, U.S. Appl. No. 13/598,204, filed Aug. 29, 2012.
Office Action dated May 14, 2014, U.S. Appl. No. 13/533,843, filed Jun. 26, 2012.
Office Action dated May 9, 2014, U.S. Appl. No. 13/484,072, filed May 30, 2012.
Office Action dated Feb. 28, 2014, U.S. Appl. No. 13/351,513, filed Jan. 17, 2012.
Office Action dated Jun. 18, 2014, U.S. Appl. No. 13/440,861, filed Apr. 5, 2012.
Office Action dated Mar. 6, 2014, U.S. Appl. No. 13/425,238, filed Mar. 20, 2012.
Office Action dated Apr. 22, 2014, U.S. Appl. No. 13/030,806, filed Feb. 18, 2011.
Office Action dated Jun. 20, 2014, U.S. Appl. No. 13/092,877, filed Apr. 22, 2011.
Office Action dated Apr. 9, 2014, U.S. Appl. No. 13/092,752, filed Apr. 22, 2011.
Office Action dated Mar. 26, 2014, U.S. Appl. No. 13/092,701, filed Apr. 22, 2011.
Office Action dated Mar. 14, 2014, U.S. Appl. No. 13/092,460, filed Apr. 22, 2011.
Office Action dated Apr. 9, 2014, U.S. Appl. No. 13/092,724, filed Apr. 22, 2011.
Brocade 'An Introduction to Brocade VCS Fabric Technology', Dec. 3, 2012.
Lapuh, Roger et al., 'Split Multi-link Trunking (SMLT) draft-lapuh-network-smlt-08', Jan. 2009.
Office Action for U.S. Appl. No. 13/030,688, filed Feb. 18, 2011, dated Jul. 17, 2014.
Office Action for U.S. Appl. No. 13/365,993, filed Feb. 3, 2012, from Cho, Hong Sol., dated Jul. 23, 2013.
Office Action for U.S. Appl. No. 13/030,806, filed Feb. 18, 2011, dated Dec. 3, 2012.
Office Action for U.S. Appl. No. 13/312,903, filed Dec. 6, 2011, dated Jun. 13, 2013.
Office Action for U.S. Appl. No. 13/087,239, filed Apr. 14, 2011, dated Dec. 5, 2012.
Office action dated Apr. 26, 2012, U.S. Appl. No. 12/725,249, filed Mar. 16, 2010.
Office action dated Sep. 12, 2012, U.S. Appl. No. 12/725,249, filed Mar. 16, 2010.
Office action dated Dec. 21, 2012, U.S. Appl. No. 13/098,490, filed May 2, 2011.
Office action dated Mar. 27, 2014, U.S. Appl. No. 13/098,490, filed May 2, 2011.
Office action dated Jul. 9, 2013, U.S. Appl. No. 13/098,490, filed May 2, 2011.
Office action dated May 22, 2013, U.S. Appl. No. 13/087,239, filed Apr. 14, 2011.
Office action dated Dec. 5, 2012, U.S. Appl. No. 13/087,239, filed Apr. 14, 2011.
Office action dated Feb. 5, 2013, U.S. Appl. No. 13/092,724, filed Apr. 22, 2011.
Office action dated Jan. 10, 2014, U.S. Appl. No. 13/092,580, filed Apr. 22, 2011.
Office action dated Jun. 10, 2013, U.S. Appl. No. 13/092,580, filed Apr. 22, 2011.
Office action dated Jan. 16, 2014, U.S. Appl. No. 13/042,259, filed Mar. 7, 2011.

(56) **References Cited**

OTHER PUBLICATIONS

Office action dated Mar. 18, 2013, U.S. Appl. No. 13/042,259, filed Mar. 7, 2011.
Office action dated Jun. 21, 2013, U.S. Appl. No. 13/092,460, filed Apr. 22, 2011.
Office action dated Jan. 28, 2013, U.S. Appl. No. 13/092,701, filed Apr. 22, 2011.
Office action dated Jul. 3, 2013, U.S. Appl. No. 13/092,701, filed Apr. 22, 2011.
Office action dated Jul. 18, 2013, U.S. Appl. No. 13/092,752, filed Apr. 22, 2011.
Office action dated Dec. 20, 2012, U.S. Appl. No. 12/950,974, filed Nov. 19, 2010.
Office action dated May 24, 2012, U.S. Appl. No. 12/950,974, filed Nov. 19, 2010.
Office action dated Jan. 6, 2014, U.S. Appl. No. 13/092,877, filed Apr. 22, 2011.
Office action dated Sep. 5, 2013, U.S. Appl. No. 13/092,877, filed Apr. 22, 2011.
Office action dated Mar. 4, 2013, U.S. Appl. No. 13/092,877, filed Apr. 22, 2011.
Office action dated Jan. 4, 2013, U.S. Appl. No. 12/950,968, filed Nov. 19, 2010.
Office action dated Jun. 7, 2012, U.S. Appl. No. 12/950,968, filed Nov. 19, 2010.
Office action dated Sep. 19, 2012, U.S. Appl. No. 13/092,864, filed Apr. 22, 2011.
Office action dated May 31, 2013, U.S. Appl. No. 13/098,360, filed Apr. 29, 2011.
Office action dated Oct. 2, 2013, U.S. Appl. No. 13/044,326, filed Mar. 9, 2011.
Office action dated Dec. 3, 2012, U.S. Appl. No. 13/030,806, filed Feb. 18, 2011.
Office action dated Jun. 11, 2013, U.S. Appl. No. 13/030,806, filed Feb. 18, 2011.
Office action dated Apr. 25, 2013, U.S. Appl. No. 13/030,688, filed Feb. 18, 2011.
Office action dated Feb. 22, 2013, U.S. Appl. No. 13/044,301, filed Mar. 9, 2011.
Office action dated Jun. 11, 2013, U.S. Appl. No. 13/044,301, filed Mar. 9, 2011.
Office action dated Oct. 26, 2012, U.S. Appl. No. 13/050,102, filed Mar. 17, 2011.
Office action dated May 16, 2013, U.S. Appl. No. 13/050,102, filed Mar. 17, 2011.
Office action dated Aug. 4, 2014, U.S. Appl. No. 13/050,102, filed Mar. 17, 2011.
Office action dated Jan. 28, 2013, U.S. Appl. No. 13/148,526, filed Jul. 16, 2011.
Office action dated Dec. 2, 2013, U.S. Appl. No. 13/184,526, filed Jul. 16, 2011.
Office action dated May 22, 2013, U.S. Appl. No. 13/148,526, filed Jul. 16, 2011.
Office action dated Nov. 29, 2013, U.S. Appl. No. 13/092,873, filed Apr. 22, 2011.
Office action dated Jun. 19, 2013, U.S. Appl. No. 13/092,873, filed Apr. 22, 2011.
Office action dated Jul. 18, 2013, U.S. Appl. No. 13/365,808, filed Feb. 3, 2012.
Office action dated Nov. 12, 2013, U.S. Appl. No. 13/312,903, filed Dec. 6, 2011.
Office action dated Jun. 13, 2013, U.S. Appl. No. 13/312,903, filed Dec. 6, 2011.
Office Action for U.S. Appl. No. 13/092,887, dated Jan. 6, 2014.
Office Action for U.S. Appl. No. 13/098,490, filed May 2, 2011, dated Mar. 27, 2014.
Office Action for U.S. Appl. No. 13/044,326, filed Mar. 9, 2011, dated Jul. 7, 2014.
Office Action for U.S. Appl. No. 13/092,752, filed Apr. 22, 2011, dated Apr. 9, 2014.
Office Action for U.S. Appl. No. 13/092,873, filed Apr. 22, 2011, dated Jul. 25, 2014.
Office Action for U.S. Appl. No. 13/092,877, filed Apr. 22, 2011, dated Jun. 20, 2014.
Office Action for U.S. Appl. No. 13/312,903, filed Dec. 6, 2011, dated Aug. 7, 2014.
Office Action for U.S. Appl. No. 13/351,513, filed Jan. 17, 2012, dated Jul. 24, 2014.
Office Action for U.S. Appl. No. 13/425,238, filed Mar. 20, 2012, dated Mar. 6, 2014.
Office Action for U.S. Appl. No. 13/556,061, filed Jul. 23, 2012, dated Jun. 6, 2014.
Office Action for U.S. Appl. No. 13/742,207 dated Jul. 24, 2014, filed Jan. 15, 2013.
Office Action for U.S. Appl. No. 13/950,974, filed Nov. 19, 2010, from Haile, Awet A., dated Dec. 2, 2012.
Perlman R: 'Challenges and opportunities in the design of TRILL: a routed layer 2 technology', 2009 IEEE GLOBECOM Workshops, Honolulu, HI, USA, Piscataway, NJ, USA, Nov. 30, 2009, pp. 1-6, XP002649647, DOI: 10.1109/GLOBECOM.2009.5360776 ISBN: 1-4244-5626-0 [retrieved on Jul. 19, 2011].
TRILL Working Group Internet-Draft Intended status: Proposed Standard RBridges: Base Protocol Specificaiton Mar. 3, 2010.
Office action dated Aug. 14, 2014, U.S. Appl. No. 13/092,460, filed Apr. 22, 2011.
Office action dated Jul. 7, 2014, for U.S. Appl. No. 13/044,326, filed Mar. 9, 2011.
Office Action dated Dec. 19, 2014, for U.S. Appl. No. 13/044,326, filed Mar. 9, 2011.
Office Action for U.S. Appl. No. 13/092,873, filed Apr. 22, 2011, dated Nov. 7, 2014.
Office Action for U.S. Appl. No. 13/092,877, filed Apr. 22, 2011, dated Nov. 10, 2014.
Office Action for U.S. Appl. No. 13/157,942, filed Jun. 10, 2011.
Mckeown, Nick et al. "OpenFlow: Enabling Innovation in Campus Networks", Mar. 14, 2008, www.openflow.org/documents/openflow-wp-latest.pdf.
Office Action for U.S. Appl. No. 13/044,301, dated Mar. 9, 2011.
Office Action for U.S. Appl. No. 13/184,526, filed Jul. 16, 2011, dated Jan. 5, 2015.
Office Action for U.S. Appl. No. 13/598,204, filed Aug. 29, 2012, dated Jan. 5, 2015.
Office Action for U.S. Appl. No. 13/669,357, filed Nov. 5, 2012, dated Jan. 30, 2015.
Office Action for U.S. Appl. No. 13/851,026, filed Mar. 26, 2013, dated Jan. 30, 2015.
Office Action for U.S. Appl. No. 13/092,460, filed Apr. 22, 2011, dated Mar. 13, 2015.
Office Action for U.S. Appl. No. 13/425,238, dated Mar. 12, 2015.
Office Action for U.S. Appl. No. 13/092,752, filed Apr. 22, 2011, dated Feb. 27, 2015.
Office Action for U.S. Appl. No. 13/042,259, filed Mar. 7, 2011, dated Feb. 23, 2015.
Office Action for U.S. Appl. No. 13/044,301, filed Mar. 9, 2011, dated Jan. 29, 2015.
Office Action for U.S. Appl. No. 13/050,102, filed Mar. 17, 2011, dated Jan. 26, 2015.
Office action dated Oct. 2, 2014, for U.S. Appl. No. 13/092,752, filed Apr. 22, 2011.
Kompella, Ed K. et al., 'Virtual Private LAN Service (VPLS) Using BGP for Auto-Discovery and Signaling' Jan. 2007.
Rosen, E. et al., "BGP/MPLS VPNs", Mar. 1999.
"Switched Virtual Internetworking moves beyond bridges and routers", Sep. 23, 1994, No. 12, New York, US.
Knight, S. et al. "Virtual Router Redundancy Protocol", Apr. 1998, XP-002135272.
Eastlake, Donald et al., "RBridges: TRILL Header Options", Dec. 2009.
Touch, J. et al., "Transparent Interconnection of Lots of Links (TRILL): Problem and Applicability Statement", May 2009.
Perlman, Radia et al., "RBridge VLAN Mapping", Dec. 2009.
"Brocade Fabric OS (FOS) 6.2 Virtual Fabrics Feature Frequently Asked Questions".

(56) **References Cited**

OTHER PUBLICATIONS

Perlman, Radia et al., "RBridges: Base Protocol Specification", Mar. 2010.
Office Action dated Jun. 18, 215, U.S. Appl. No. 13/098,490, filed May 2, 2011.
Office Action dated Jun. 16, 2015, U.S. Appl. No. 13/048,817, filed Mar. 15, 2011.
Touch, J. et al., 'Transparent Interconnection of Lots of Links (TRILL): Problem and Applicability Statement', May 2009, Network Working Group, pp. 1-17.
Zhai F. Hu et al. 'RBridge: Pseudo-Nickname; draft-hu-trill-pseudonode-nickname-02.txt', May 15, 2012.
Office Action dated Jul. 31, 2015, U.S. Appl. No. 13/598,204, filed Aug. 29, 2014.
Office Action dated Jul. 31, 2015, U.S. Appl. No. 14/473,941, filed Aug. 29, 2014.
Office Action dated Jul. 31, 2015, U.S. Appl. No. 14/488,173, filed Sep. 16, 2014.
Office Action dated Aug. 21, 2015, U.S. Appl. No. 13/776,217, filed Feb. 25, 2013.
Office Action dated Aug. 19, 2015, U.S. Appl. No. 14/156,374, filed Jan. 15, 2014.
Office Action dated Sep. 2, 2015, U.S. Appl. No. 14/151,693, filed Jan. 9, 2014.
Office Action dated Sep. 17, 2015, U.S. Appl. No. 14/577,785, filed Dec. 19, 2014.
Office Action dated Sep. 22, 2015 U.S. Appl. No. 13/656,438, filed Oct. 19, 2012.
Office Action dated Nov. 5, 2015, U.S. Appl. No. 14/178,042, filed Feb. 11, 2014.
Office Action dated Oct. 19, 2015, U.S. Appl. No. 14/215,996, filed Mar. 17, 2014.
Office Action dated Sep. 18, 2015, U.S. Appl. No. 13/345,566, filed Jan. 6, 2012.

* cited by examiner

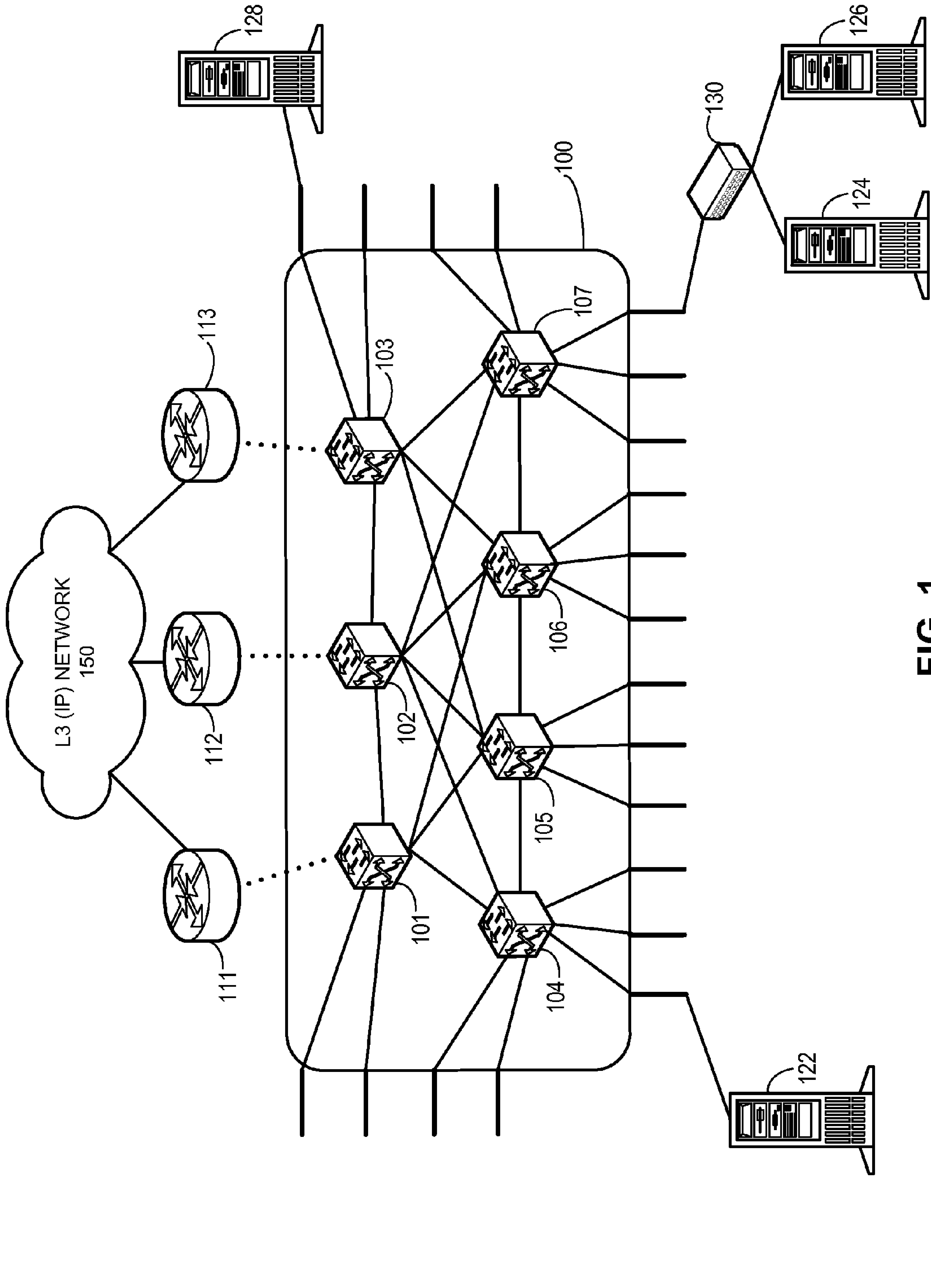

L3 (IP) NETWORK
150
111
112
113
100
101
102
103
104
105
106
107
122
124
126
128
130

FIG. 1

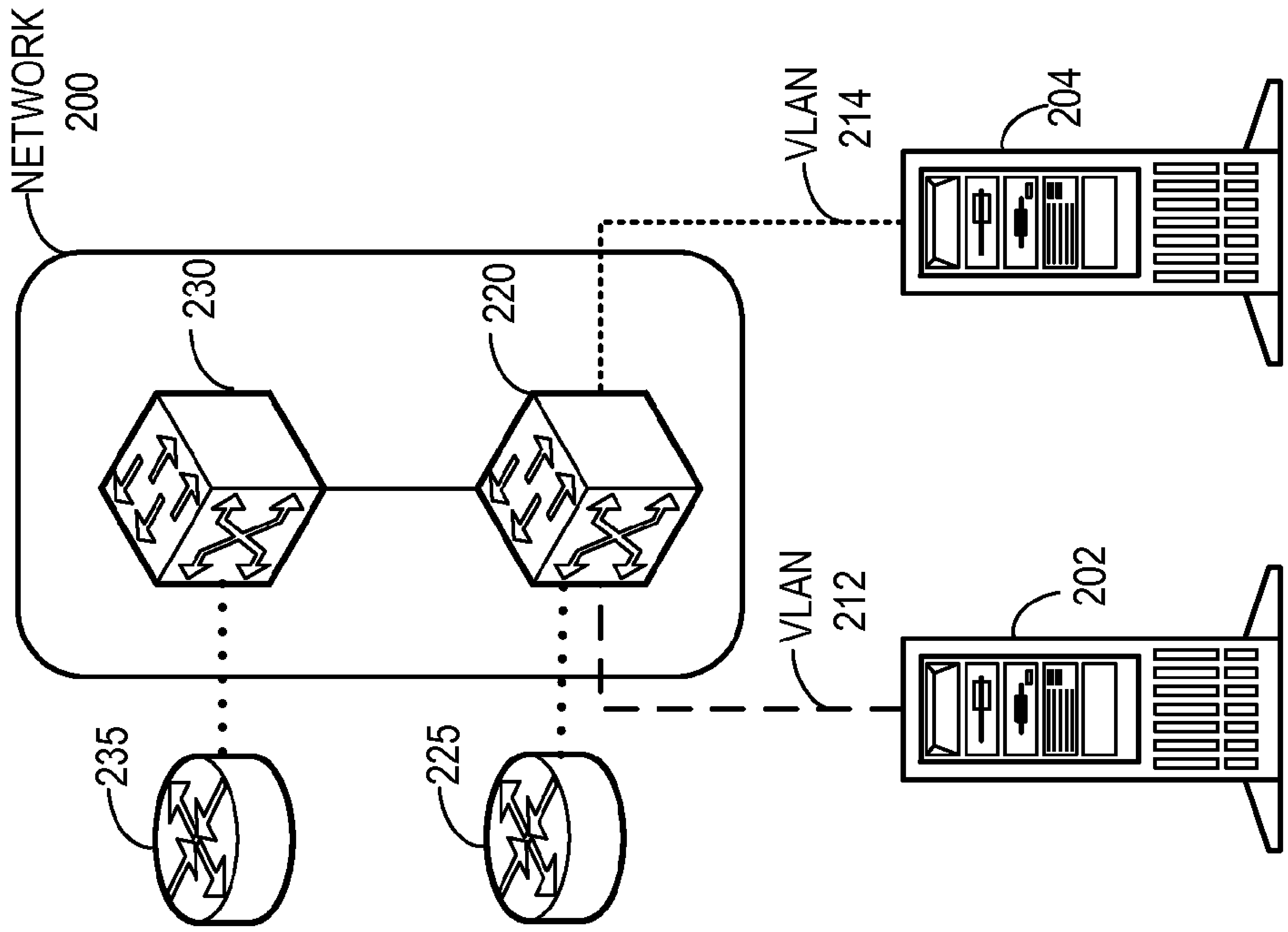

NETWORK 200
230
220
235
225
VLAN 212
VLAN 214
202
204

FIG. 2B

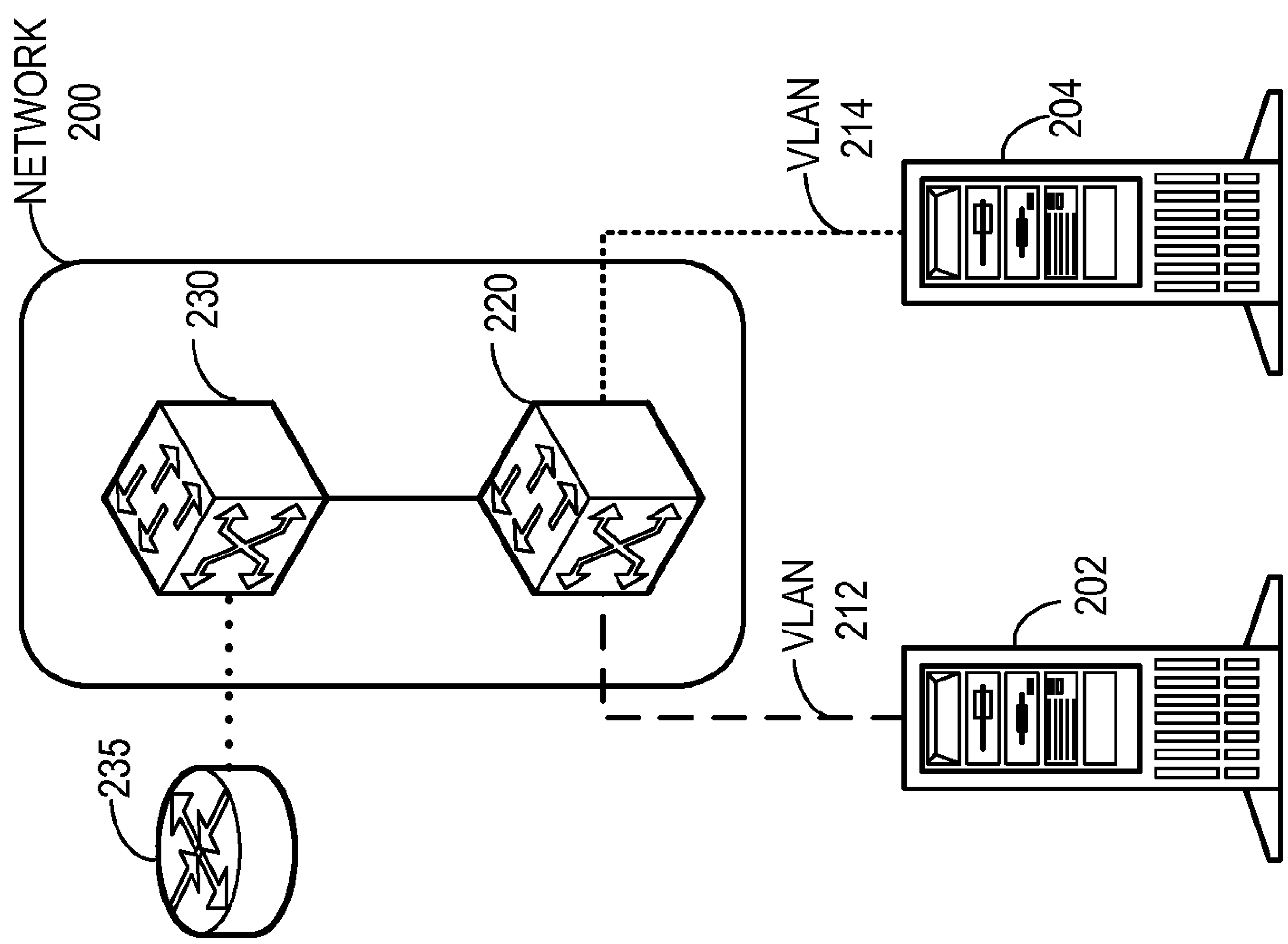

NETWORK 200
230
220
235
VLAN 212
VLAN 214
202
204

FIG. 2A

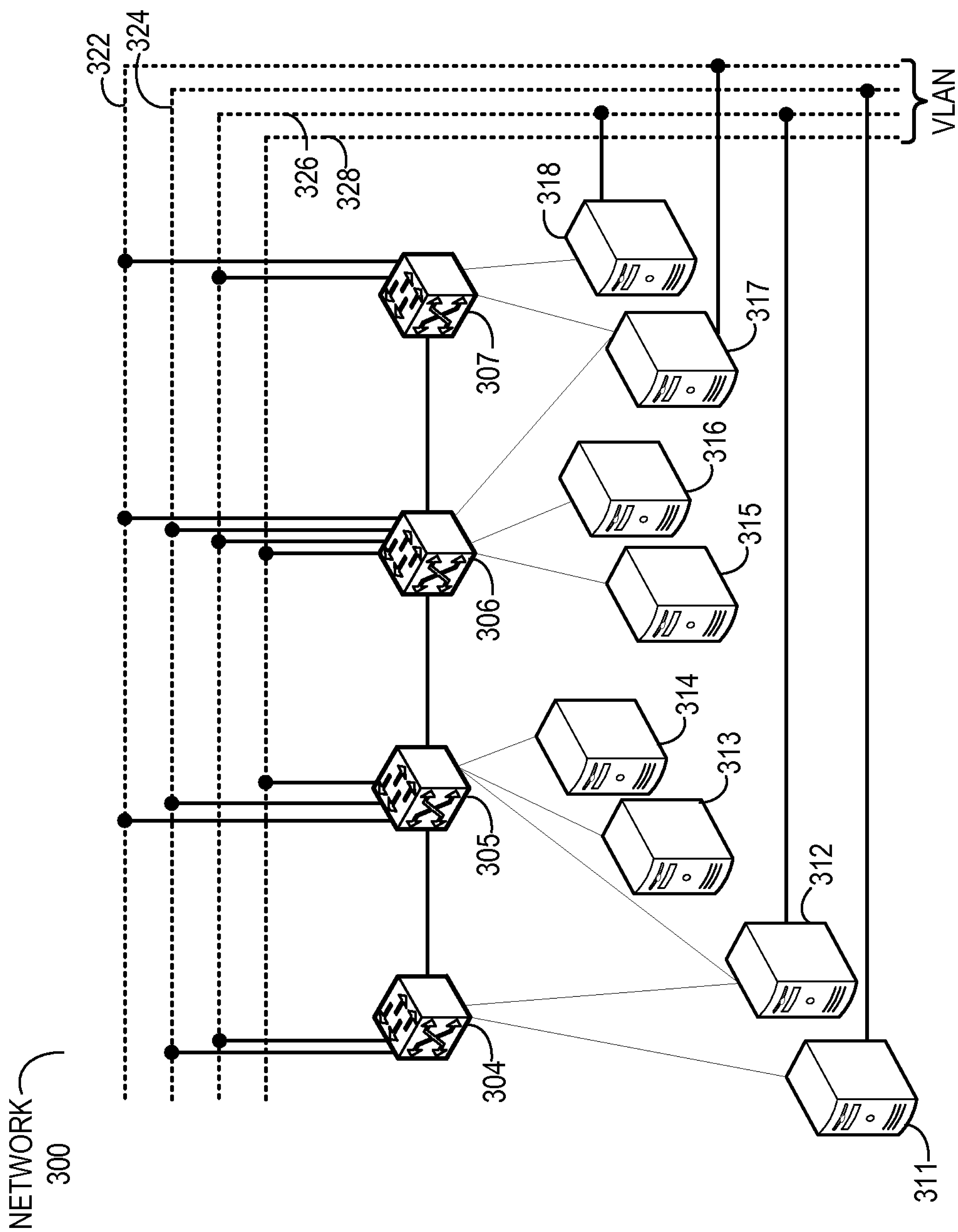

NETWORK 300
304
305
306
307
311
312
313
314
315
316
317
318
322
324
326
328
VLAN

FIG. 3A

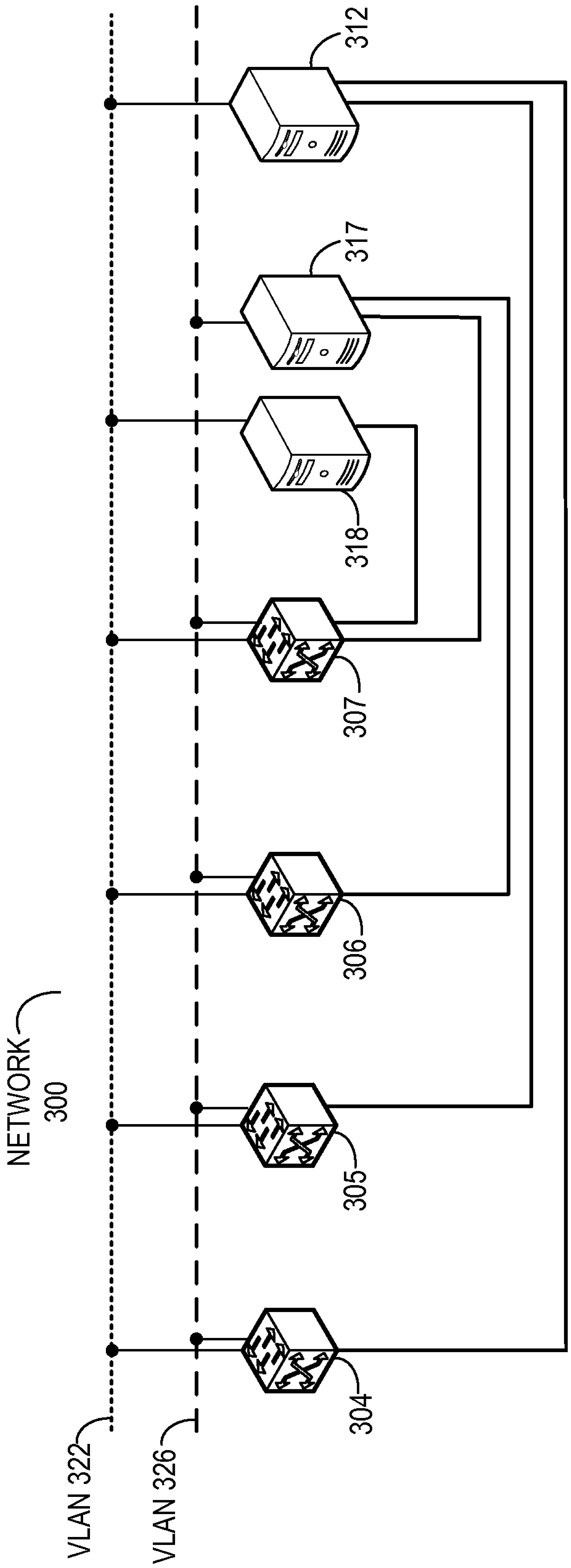


FIG. 3B

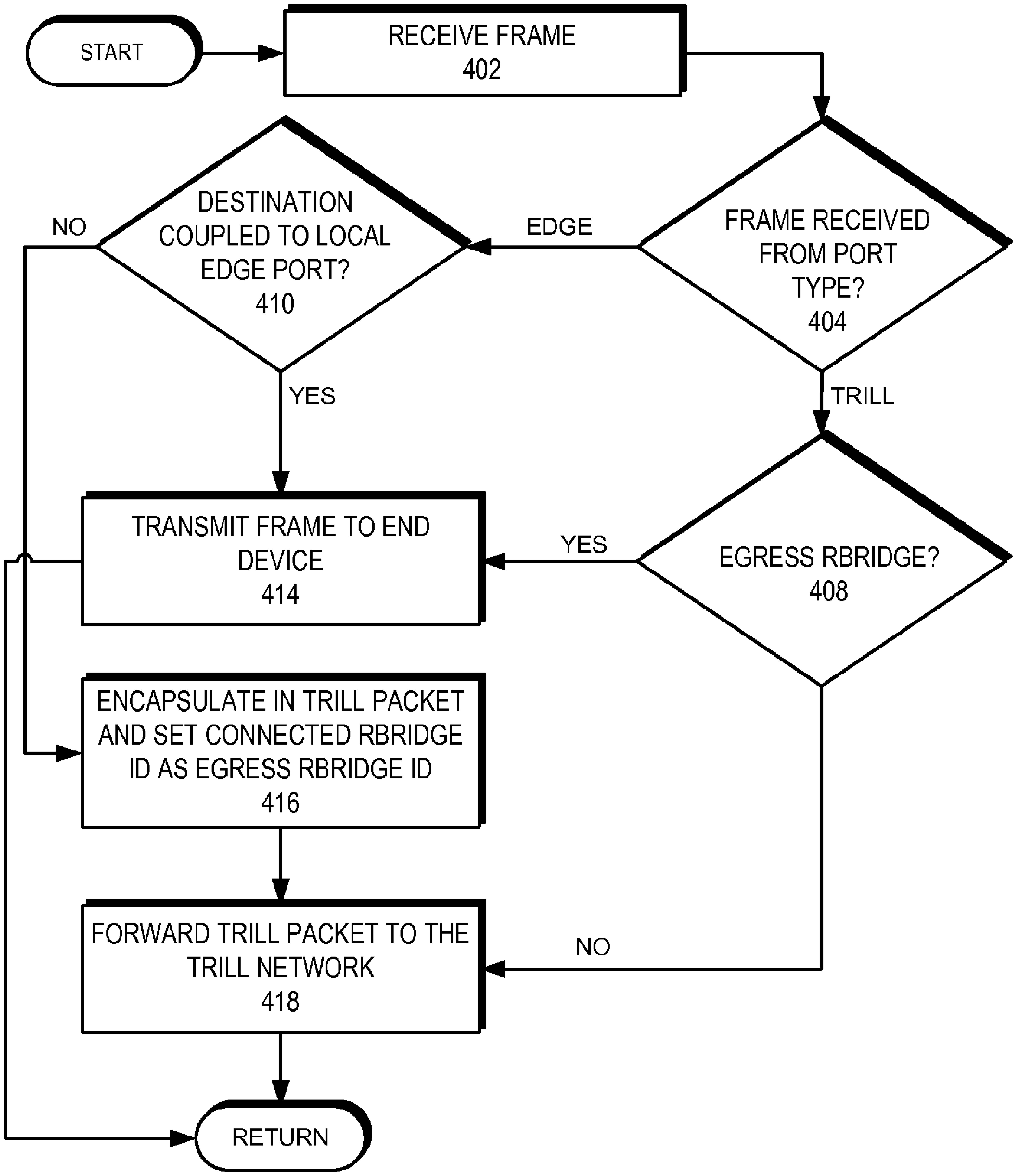

START
RECEIVE FRAME
402
FRAME RECEIVED FROM PORT TYPE?
404
EDGE
DESTINATION COUPLED TO LOCAL EDGE PORT?
410
NO
YES
TRILL
EGRESS RBRIDGE?
408
YES
TRANSMIT FRAME TO END DEVICE
414
ENCAPSULATE IN TRILL PACKET AND SET CONNECTED RBRIDGE ID AS EGRESS RBRIDGE ID
416
NO
FORWARD TRILL PACKET TO THE TRILL NETWORK
418
RETURN

FIG. 4A

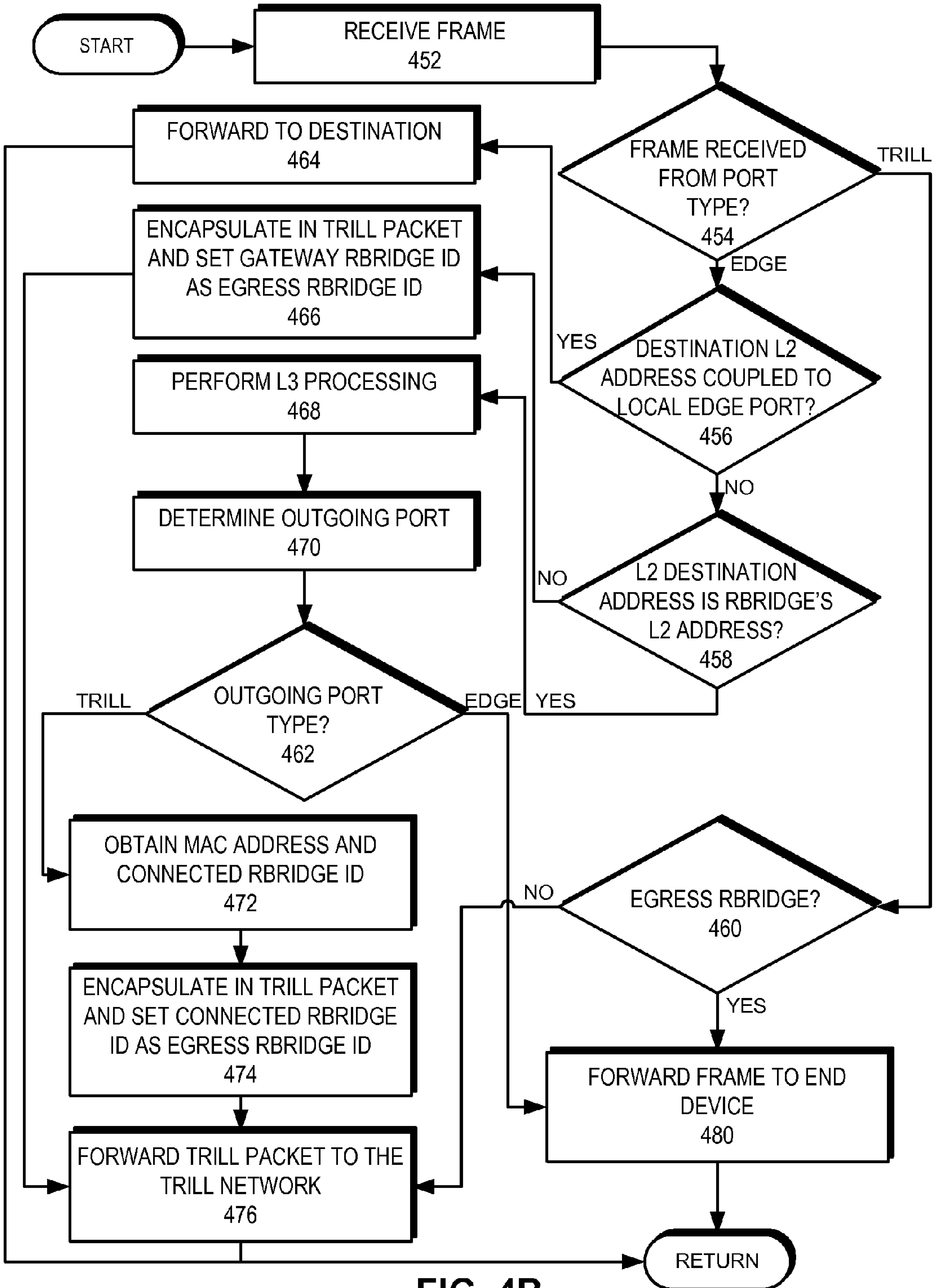

START
RECEIVE FRAME
452
FRAME RECEIVED FROM PORT TYPE?
454
TRILL
EDGE
FORWARD TO DESTINATION
464
ENCAPSULATE IN TRILL PACKET AND SET GATEWAY RBRIDGE ID AS EGRESS RBRIDGE ID
466
YES
DESTINATION L2 ADDRESS COUPLED TO LOCAL EDGE PORT?
456
PERFORM L3 PROCESSING
468
NO
DETERMINE OUTGOING PORT
470
NO
L2 DESTINATION ADDRESS IS RBRIDGE'S L2 ADDRESS?
458
TRILL
OUTGOING PORT TYPE?
462
EDGE
YES
OBTAIN MAC ADDRESS AND CONNECTED RBRIDGE ID
472
NO
EGRESS RBRIDGE?
460
YES
ENCAPSULATE IN TRILL PACKET AND SET CONNECTED RBRIDGE ID AS EGRESS RBRIDGE ID
474
FORWARD FRAME TO END DEVICE
480
FORWARD TRILL PACKET TO THE TRILL NETWORK
476
RETURN

FIG. 4B

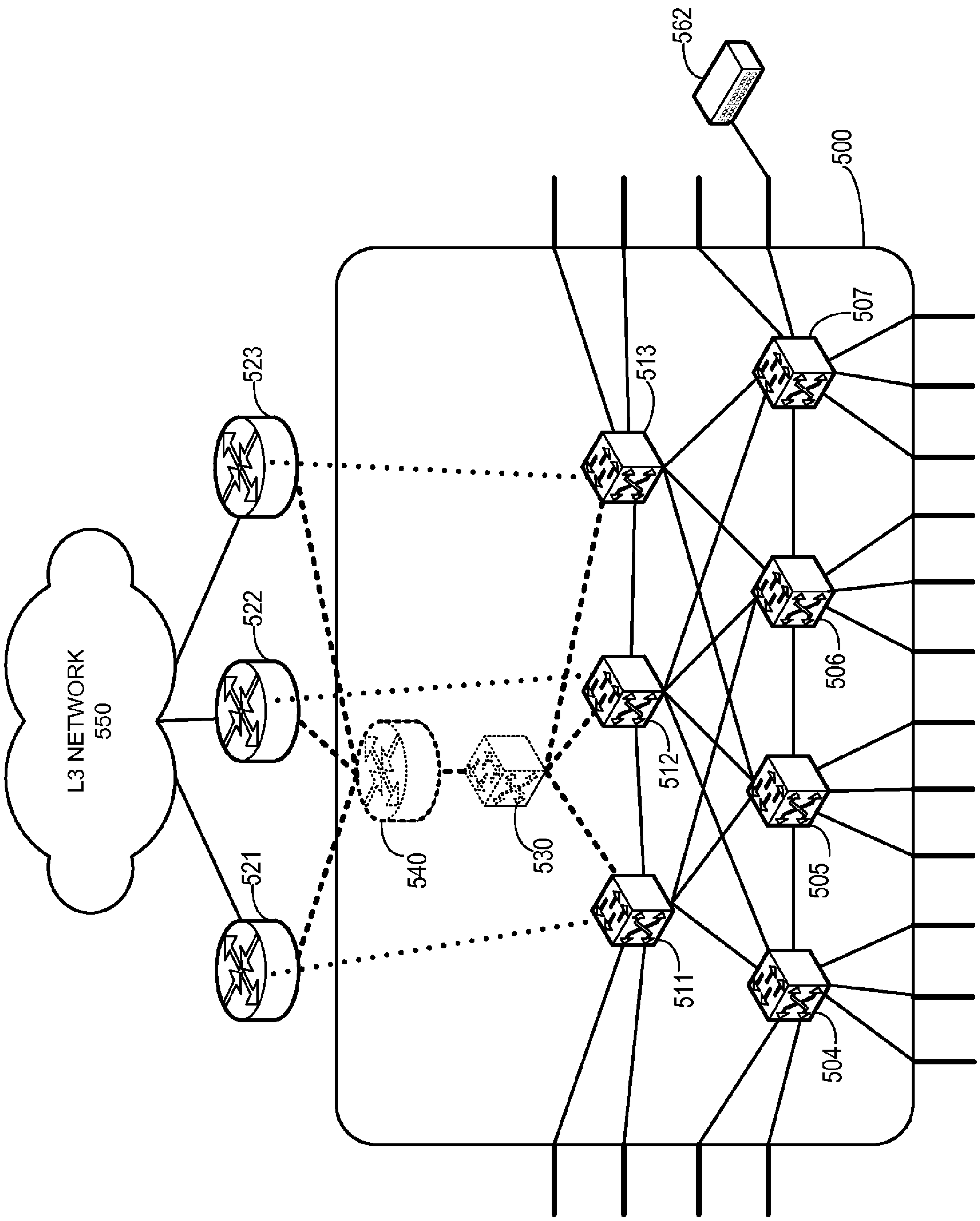

L3 NETWORK
550
521
522
523
540
530
511
512
513
504
505
506
507
500
562

FIG. 5

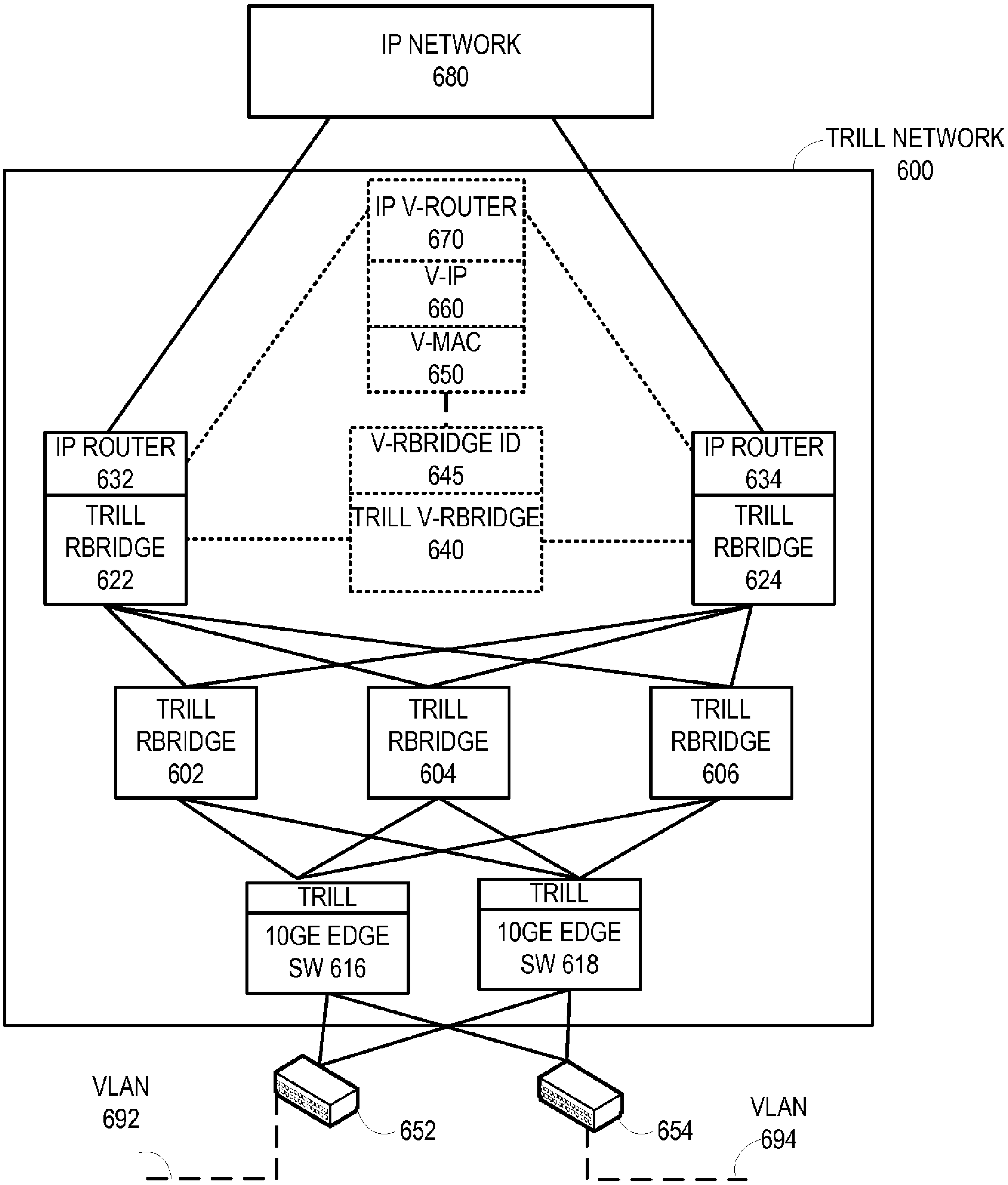

IP NETWORK
680
TRILL NETWORK
600
IP V-ROUTER
670
V-IP
660
V-MAC
650
V-RBRIDGE ID
645
TRILL V-RBRIDGE
640
IP ROUTER
632
TRILL
RBRIDGE
622
IP ROUTER
634
TRILL
RBRIDGE
624
TRILL
RBRIDGE
602
TRILL
RBRIDGE
604
TRILL
RBRIDGE
606
TRILL
10GE EDGE
SW 616
TRILL
10GE EDGE
SW 618
VLAN
692
652
654
VLAN
694

FIG. 6A

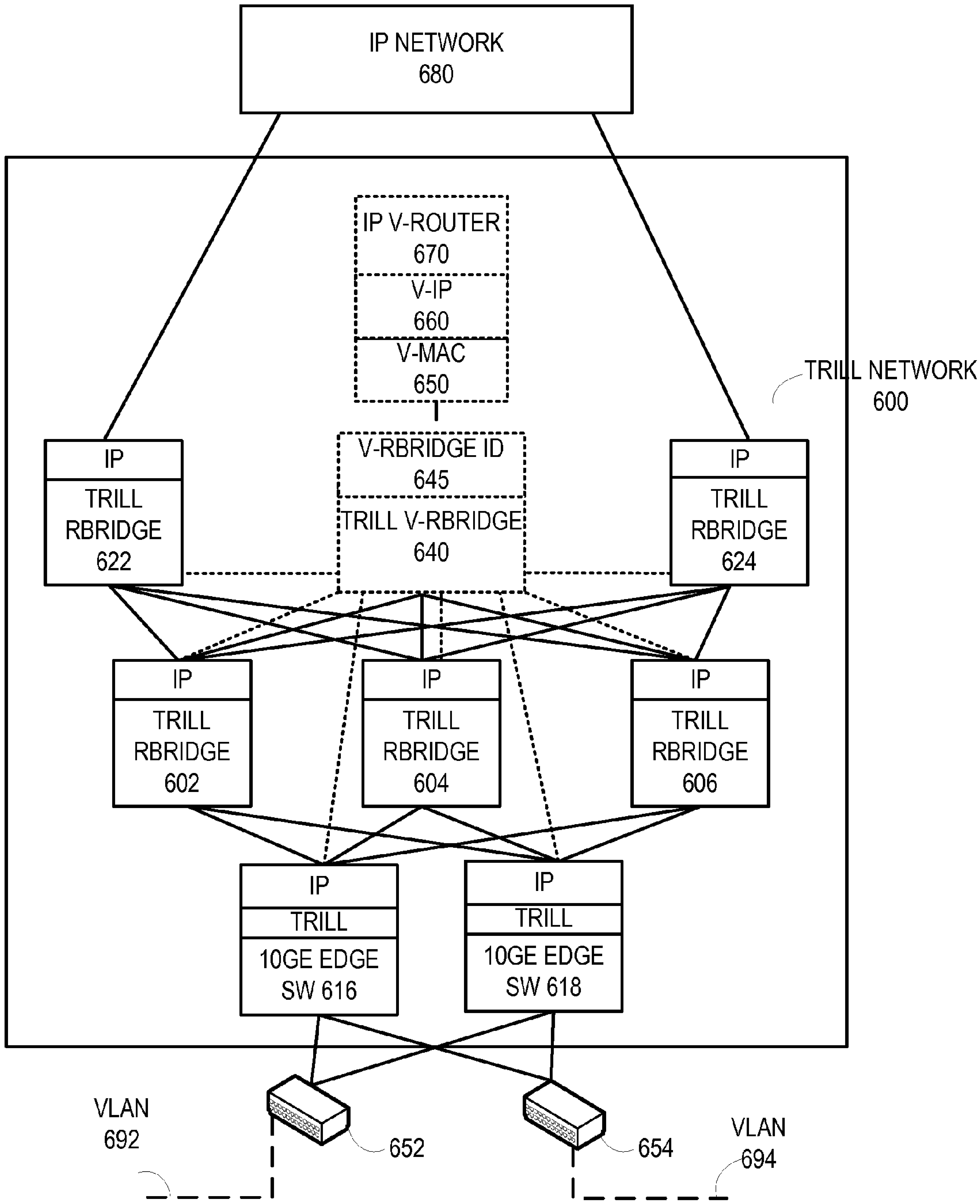

IP NETWORK
680
IP V-ROUTER
670
V-IP
660
V-MAC
650
TRILL NETWORK
600
V-RBRIDGE ID
645
TRILL V-RBRIDGE
640
IP
TRILL
RBRIDGE
622
IP
TRILL
RBRIDGE
624
IP
TRILL
RBRIDGE
602
IP
TRILL
RBRIDGE
604
IP
TRILL
RBRIDGE
606
IP
TRILL
10GE EDGE
SW 616
IP
TRILL
10GE EDGE
SW 618
VLAN
692
652
654
VLAN
694

FIG. 6B

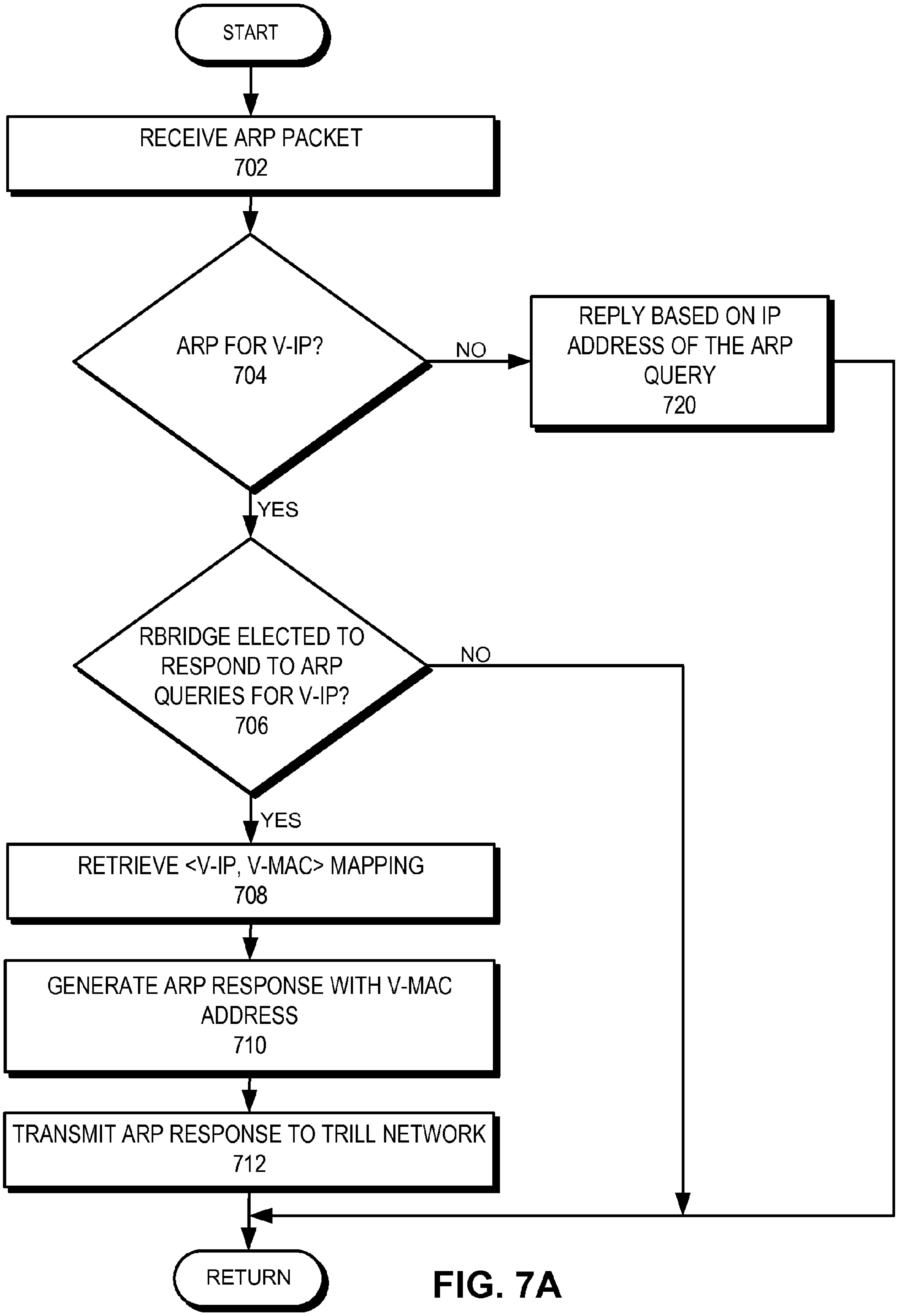

START
RECEIVE ARP PACKET
702
ARP FOR V-IP?
704
NO
REPLY BASED ON IP ADDRESS OF THE ARP QUERY
720
YES
RBRIDGE ELECTED TO RESPOND TO ARP QUERIES FOR V-IP?
706
NO
YES
RETRIEVE <V-IP, V-MAC> MAPPING
708
GENERATE ARP RESPONSE WITH V-MAC ADDRESS
710
TRANSMIT ARP RESPONSE TO TRILL NETWORK
712
RETURN

FIG. 7A

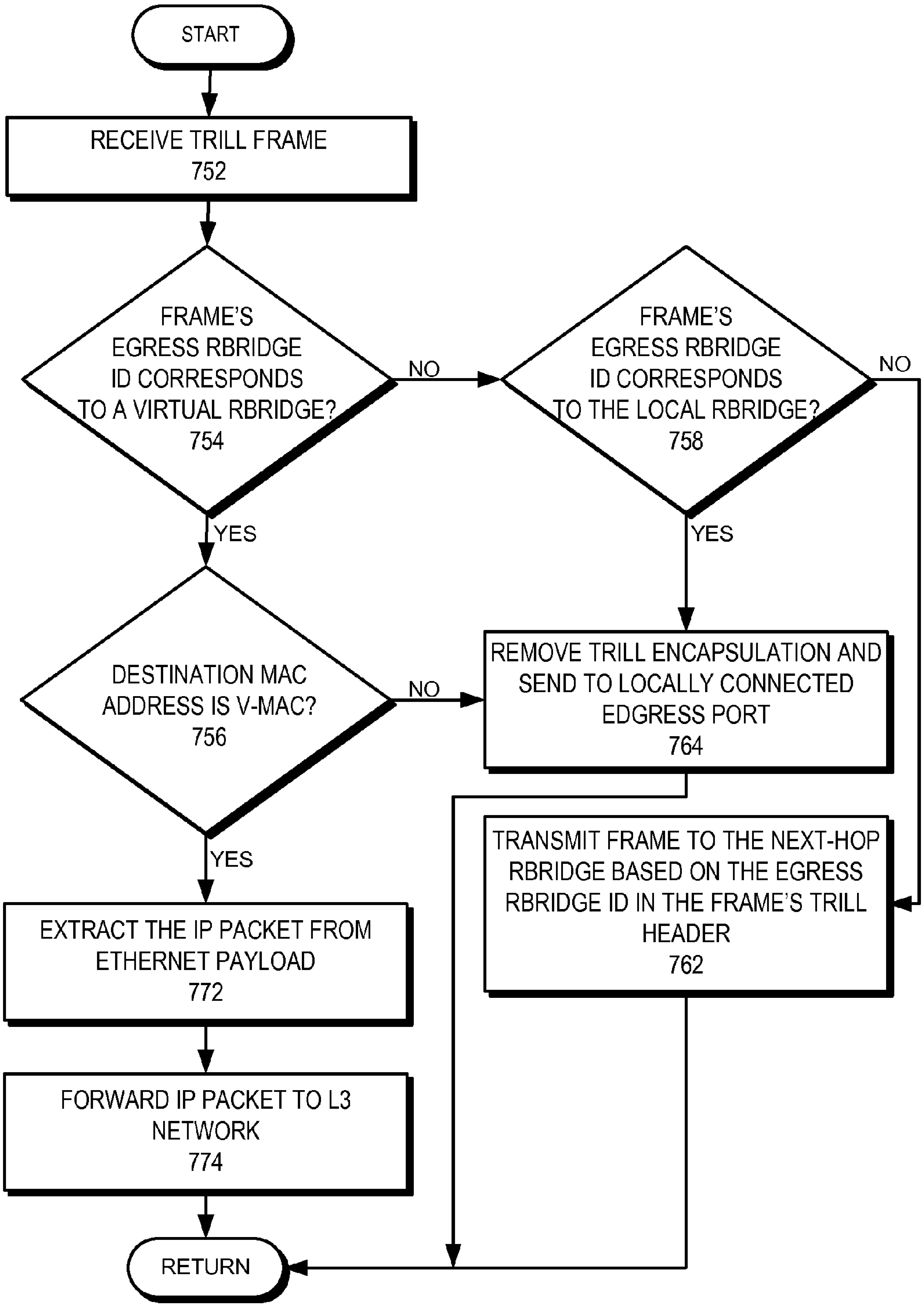
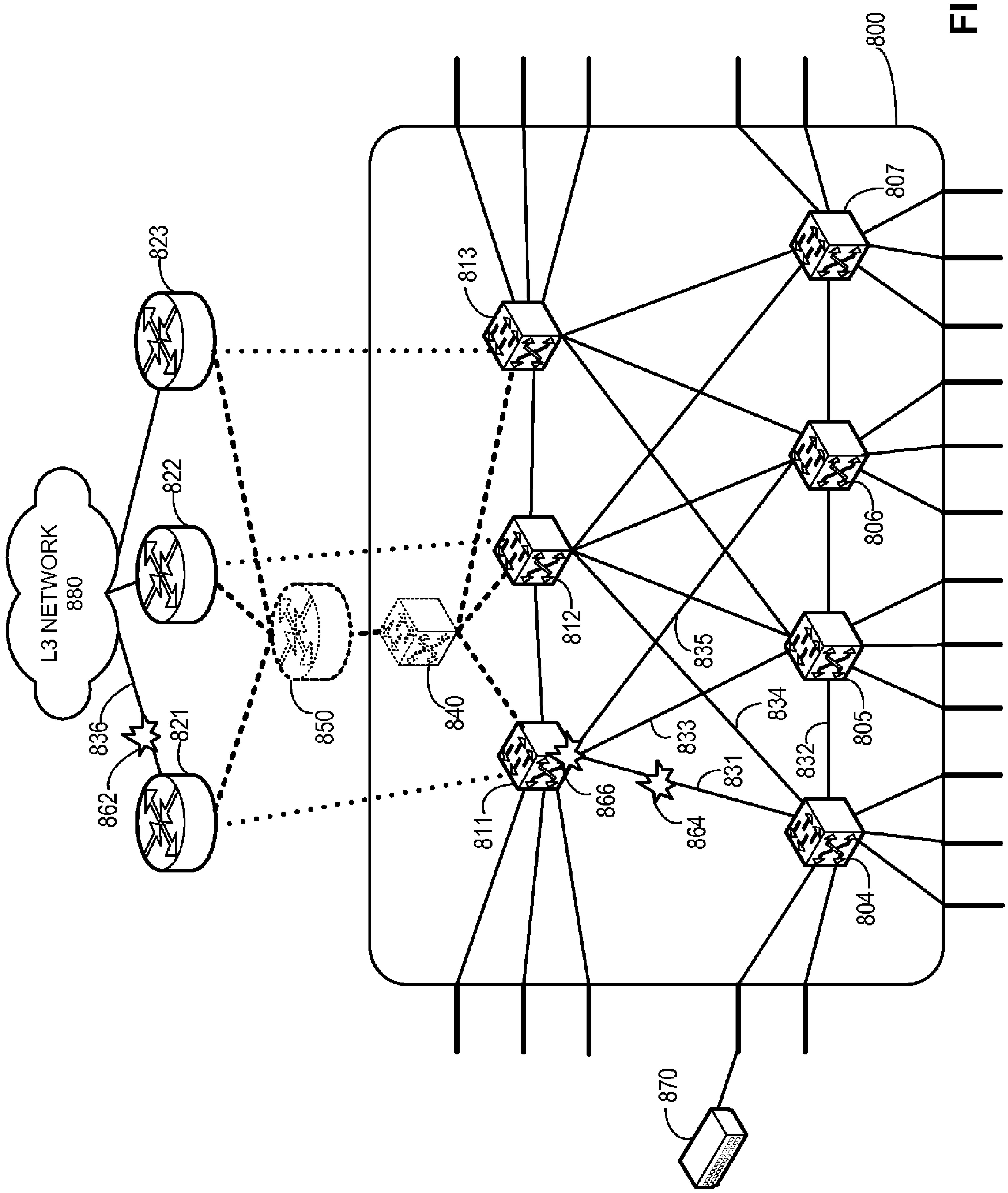

START
RECEIVE TRILL FRAME
752
FRAME'S EGRESS RBRIDGE ID CORRESPONDS TO A VIRTUAL RBRIDGE?
754
NO
FRAME'S EGRESS RBRIDGE ID CORRESPONDS TO THE LOCAL RBRIDGE?
758
NO
YES
YES
DESTINATION MAC ADDRESS IS V-MAC?
756
NO
REMOVE TRILL ENCAPSULATION AND SEND TO LOCALLY CONNECTED EDGRESS PORT
764
TRANSMIT FRAME TO THE NEXT-HOP RBRIDGE BASED ON THE EGRESS RBRIDGE ID IN THE FRAME'S TRILL HEADER
762
YES
EXTRACT THE IP PACKET FROM ETHERNET PAYLOAD
772
FORWARD IP PACKET TO L3 NETWORK
774
RETURN

FIG. 7B



L3 NETWORK
880
821
822
823
836
862
850
840
811
812
813
866
864
831
832
833
834
835
804
805
806
807
800
870

FIG. 8

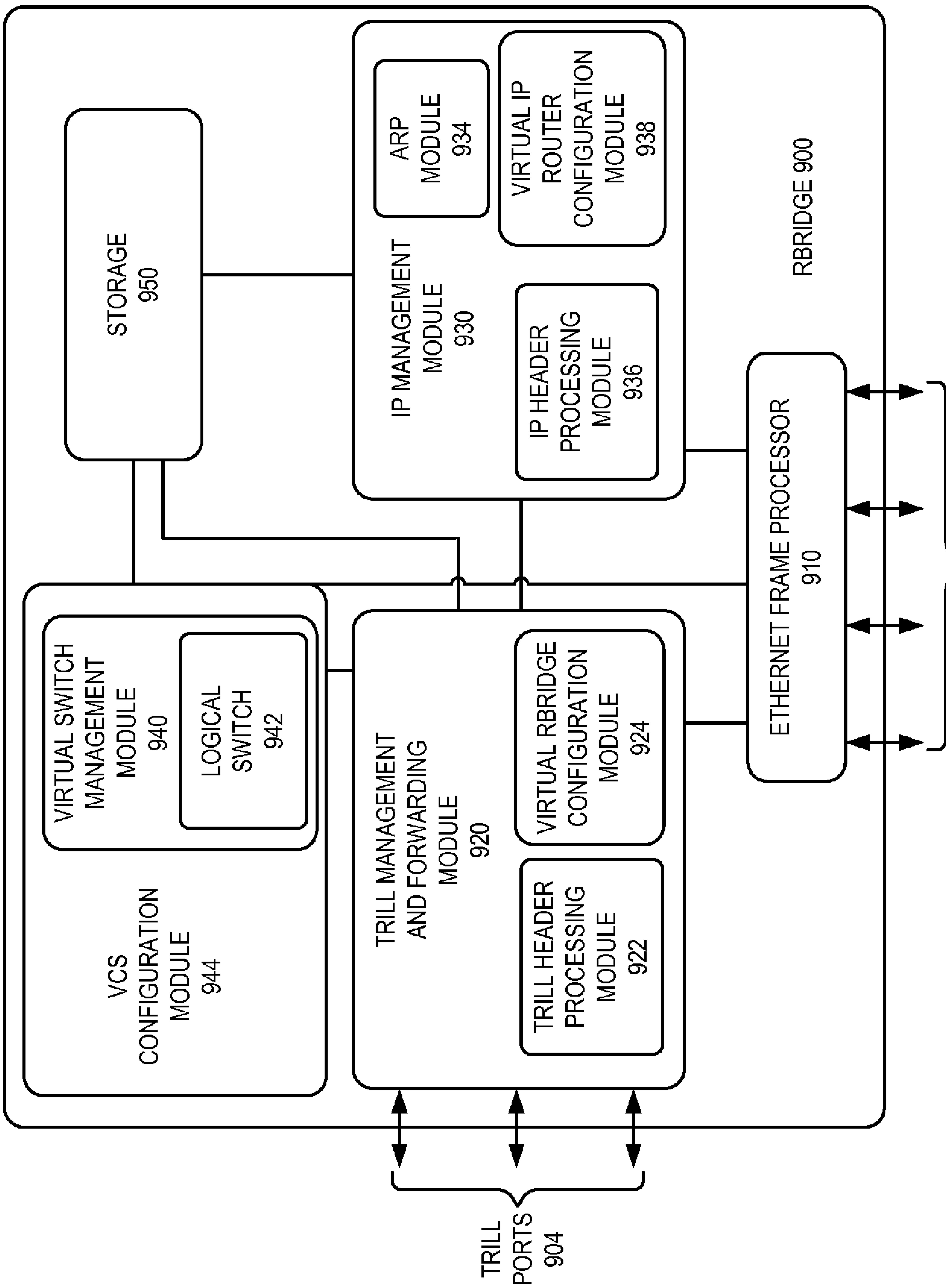

STORAGE 950
IP MANAGEMENT MODULE 930
ARP MODULE 934
VIRTUAL IP ROUTER CONFIGURATION MODULE 938
IP HEADER PROCESSING MODULE 936
RBRIDGE 900
VCS CONFIGURATION MODULE 944
VIRTUAL SWITCH MANAGEMENT MODULE 940
LOGICAL SWITCH 942
TRILL MANAGEMENT AND FORWARDING MODULE 920
VIRTUAL RBRIDGE CONFIGURATION MODULE 924
TRILL HEADER PROCESSING MODULE 922
ETHERNET FRAME PROCESSOR 910
EDGE PORTS 902
TRILL PORTS 904

FIG. 9

LAYER-3 SUPPORT IN TRILL NETWORKS

RELATED APPLICATIONS

This application claims the benefit of U.S. Provisional Application No. 61/481,643, titled "Layer-3 Support in Virtual Cluster Switching," by inventors Phanidhar Koganti, Anoop Ghanwani, Suresh Vobbilisetty, Rajiv Krishnamurthy, Nagarajan Venkatesan, and Shunjia Yu, filed 2 May 2011, and U.S. Provisional Application No. 61/503,265, titled "IP Routing in VCS," by inventors Phanidhar Koganti, Anoop Ghanwani, Suresh Vobbilisetty, Rajiv Krishnamurthy, Nagarajan Venkatesan, and Shunjia Yu, filed 30 Jun. 2011, which are incorporated by reference herein.

The present disclosure is related to U.S. patent application Ser. No. 13/087,239, titled "Virtual Cluster Switching," by inventors Suresh Vobbilisetty and Dilip Chatwani, filed 14 Apr. 2011, and U.S. patent application Ser. No. 12/725,249, titled "Redundant Host Connection in a Routed Network," by inventors Somesh Gupta, Anoop Ghawani, Phanidhar Koganti, and Shunjia Yu, filed 16 Mar. 2010, the disclosures of which are incorporated by reference herein.

BACKGROUND

1. Field

The present disclosure relates to network design. More specifically, the present disclosure relates to a method and system for constructing a scalable switching system that supports layer-3 routing while facilitating automatic configuration.

2. Related Art

The growth of the Internet has brought with it an increasing demand for bandwidth. As a result, equipment vendors race to build larger and faster switches with versatile capabilities, such as layer-3 forwarding, to move more traffic efficiently. However, the size of a switch cannot grow infinitely. It is limited by physical space, power consumption, and design complexity, to name a few factors. Furthermore, switches with higher capability are usually more complex and expensive. More importantly, because an overly large and complex system often does not provide economy of scale, simply increasing the size and capability of a switch may prove economically unviable due to the increased per-port cost.

One way to increase the throughput of a switch system is to use switch stacking. In switch stacking, multiple smaller-scale, identical switches are interconnected in a special pattern to form a larger logical switch. The amount of required manual configuration and topological limitations for switch stacking becomes prohibitively tedious when the stack reaches a certain size, which precludes switch stacking from being a practical option in building a large-scale switching system.

Meanwhile, layer-2 (e.g., Ethernet) switching technologies continue to evolve. More routing-like functionalities, which have traditionally been the characteristics of layer-3 (e.g., Internet Protocol or IP) networks, are migrating into layer-2. Notably, the recent development of the Transparent Interconnection of Lots of Links (TRILL) protocol allows Ethernet switches to function more like routing devices. TRILL overcomes the inherent inefficiency of the conventional spanning tree protocol, which forces layer-2 switches to be coupled in a logical spanning-tree topology to avoid looping. TRILL allows routing bridges (RBridges) to be coupled in an arbitrary topology without the risk of looping by implementing routing functions in switches and including a hop count in the TRILL header.

While TRILL brings many desirable features to layer-2 networks, some issues remain unsolved when layer-3 processing is desired.

SUMMARY

One embodiment of the present invention provides a switch. The switch includes an IP header processor and a forwarding mechanism. The IP header processor identifies a destination IP address in a packet encapsulated with an inner Ethernet header, a TRILL header, and an outer Ethernet header. The forwarding mechanism determines an output port and constructs a new header for the packet based on the destination IP address. The switch also includes a packet processor which determines whether (1) an inner destination media access control (MAC) address corresponds to a local MAC address assigned to the switch; (2) a destination RBridge identifier (RBridge ID) corresponds to a local RBridge identifier assigned to the switch; and (3) an outer destination MAC address corresponds to the local MAC address.

In a variation on this embodiment, the packet processor determines a first virtual local area network (VLAN) tag in the inner Ethernet header, wherein the new header includes a new inner Ethernet header which comprises a second VLAN tag.

In a variation on this embodiment, the switch includes a control mechanism which forms a virtual cluster switch in conjunction with one or more additional switches.

In a variation on this embodiment, the virtual cluster switch is an Ethernet fabric switch functioning as a logical Ethernet switch.

In a variation on this embodiment, the switch includes a switching mechanism switches the packet between VLANs based on the destination IP address.

In a variation on this embodiment, the RBridge identifier is a virtual RBridge identifier and the destination IP address is a virtual IP address assigned to a virtual IP router associated with the virtual RBridge identifier.

In a variation on this embodiment, the virtual IP router is formed by operating the switch in conjunction with at least another physical switch as a single logical router.

BRIEF DESCRIPTION OF THE FIGURES

FIG. **1** illustrates an exemplary TRILL network that includes a plurality of RBridges with IP processing capabilities, in accordance with an embodiment of the present invention.

FIG. **2**A illustrates an exemplary configuration of end devices belonging to different VLANs and coupled to a TRILL network, wherein one RBridge is IP capable, in accordance with an embodiment of the present invention.

FIG. **2**B illustrates an exemplary configuration of end devices belonging to different VLANs and coupled to a TRILL network, wherein all RBridges are IP capable, in accordance with an embodiment of the present invention.

FIG. **3**A illustrates an exemplary TRILL network with multiple VLANs, in accordance with an embodiment of the present invention.

FIG. **3**B illustrates an exemplary TRILL network with multiple VLANs, wherein each RBridge belongs to all VLANs, in accordance with an embodiment of the present invention.

FIG. **4**A presents a flowchart illustrating the process of an RBridge transmitting a frame, in accordance with an embodiment of the present invention.

FIG. 4B presents a flowchart illustrating the process of an IP-capable RBridge transmitting a frame, in accordance with an embodiment of the present invention.

FIG. 5 illustrates an exemplary network where a virtual RBridge and an associated virtual IP router are created based on a plurality of physical gateway RBridges with IP processing capabilities, in accordance with an embodiment of the present invention.

FIG. 6A illustrates an exemplary configuration of how a virtual RBridge and an associated virtual IP router can be logically coupled to a number of gateway RBridges in a TRILL network, in accordance with an embodiment of the present invention.

FIG. 6B illustrates an exemplary configuration of how a virtual RBridge and an associated virtual IP router can be logically coupled to all RBridges in a TRILL network where each RBridge has IP processing capability, in accordance with an embodiment of the present invention.

FIG. 7A presents a flowchart illustrating the process of a gateway RBridge associated with a virtual RBridge responding to an Address Resolution Protocol (ARP) query, in accordance with an embodiment of the present invention.

FIG. 7B presents a flowchart illustrating the process of a gateway RBridge associated with a virtual RBridge forwarding a TRILL frame, in accordance with an embodiment of the present invention.

FIG. 8 illustrates a scenario where one of the RBridges associated with the virtual RBridge experiences a link failure and/or a node failure, in accordance with an embodiment of the present invention.

FIG. 9 illustrates an exemplary architecture of a switch with IP processing capabilities, in accordance with an embodiment of the present invention.

DETAILED DESCRIPTION

The following description is presented to enable any person skilled in the art to make and use the invention, and is provided in the context of a particular application and its requirements. Various modifications to the disclosed embodiments will be readily apparent to those skilled in the art, and the general principles defined herein may be applied to other embodiments and applications without departing from the spirit and scope of the present invention. Thus, the present invention is not limited to the embodiments shown, but is to be accorded the widest scope consistent with the claims.

Overview

In embodiments of the present invention, the problem of providing scalable and flexible layer-3 (e.g., IP) support in a TRILL network is solved by facilitating IP routing in a number of RBridges in the TRILL network. The availability of IP processing within a TRILL network allows cross-layer-2-domain traffic (e.g., traffic across different VLANs) to be forwarded within a TRILL network, which reduces forwarding overhead. Usually, the IP router portion of one of these IP-capable RBridges is assigned as a default gateway router to an end device coupled to a TRILL network. Wherever the end device sends a frame to outside of its local network (e.g., a VLAN), the frame is forwarded to and processed by the IP router portion of the RBridge. This layer-3 processing occurs within the TRILL network. Note that, in a conventional TRILL network, such layer-3 processing has to be done by an IP router residing outside the TRILL network.

In some embodiments, the end-device may be coupled to the TRILL network via an ingress RBridge without IP processing capability. Under such a scenario, the TRILL RBridge portion of an IP-capable RBridge acts as an egress RBridge and the IP router portion of the RBridge can act as the default gateway router. A frame from the end device is received at the ingress RBridge and encapsulated in a TRILL packet, wherein the TRILL packet sets the egress RBridge identifier as the destination RBridge identifier, and the MAC address of the egress RBridge as the inner destination MAC address. The packet is then forwarded though the TRILL network and reaches the egress RBridge, where the outer destination MAC address of the packet is the MAC address of the egress RBridge. The IP router portion of the egress RBridge then processes the IP header in the frame and makes the layer-3 forwarding decision based on the destination IP address of the frame.

In some embodiments, the IP router portion of an IP-capable RBridge may be associated with multiple VLANs associated with the TRILL network. If the destination end device of the frame belongs to one of the associated VLANs, the IP router can obtain the MAC address of the destination end device using ARP requests within that VLAN. The corresponding RBridge of the IP router then sets the RBridge to which the destination end device is coupled as the egress RBridge and forwards the frame to the egress RBridge over the TRILL network.

Although the present disclosure is presented using examples based on the TRILL protocol, embodiments of the present invention are not limited to TRILL networks, or networks defined in a particular Open System Interconnection Reference Model (OSI reference model) layer.

The term "RBridge" refers to routing bridges, which are bridges implementing the TRILL protocol as described in IETF Request for Comments (RFC) "Routing Bridges (RBridges): Base Protocol Specification," available at http://tools.ietf.org/html/rfc6325, which is incorporated by reference herein. Embodiments of the present invention are not limited to applications among RBridges. Other types of switches, routers, and forwarders can also be used.

In this disclosure, the term "edge port" refers to a port which sends/receives data frames in native Ethernet format. The term "TRILL port" refers to a port which sends/receives data frames encapsulated with a TRILL header and outer MAC header.

The term "end device" refers to a network device that is typically not TRILL-capable. "End device" is a relative term with respect to the TRILL network. However, "end device" does not necessarily mean that the network device is an end host. An end device can be a host, a conventional layer-2 switch, or any other type of network device. Additionally, an end device can be coupled to other switches or hosts further away from the TRILL network. In other words, an end device can be an aggregation point for a number of network devices to enter the TRILL network.

The term "IP-capable RBridge" refers to a physical RBridge that can process and route IP packets. An IP-capable RBridge can be coupled to a layer-3 network and can forward IP packets from end devices to the layer-3 network. A number of IP-capable RBridges can form a virtual RBridge and a corresponding virtual IP router, thereby facilitating a virtual gateway router for end devices that supports redundancy and load-balancing. In this disclosure, an RBridge which forms a virtual RBridge and a virtual IP router is also referred to as a "gateway" RBridge. A gateway RBridge responds to ARP requests for the virtual IP address with a virtual MAC address. In various embodiments, any arbitrary number of gateway RBridges can form the virtual RBridge. As gateway RBridges can process both TRILL and IP packets, in this disclosure the term "gateway RBridge" can refer to a physical RBridge in a TRILL network or a physical router in an IP network.

The term "IP router" refers to the IP-capable portion of an RBridge or a stand-alone IP router. In this disclosure, the terms "IP router" and "router" are used interchangeably.

The term "frame" refers to a group of bits that can be transported together across a network. "Frame" should not be interpreted as limiting embodiments of the present invention to layer-2 networks. "Frame" can be replaced by other terminologies referring to a group of bits, such as "packet," "cell," or "datagram."

The term "RBridge identifier" refers to a group of bits that can be used to identify an RBridge. Note that the TRILL standard uses "RBridge ID" to denote a 48-bit intermediate-system-to-intermediate-system (IS-IS) System ID assigned to an RBridge, and "RBridge nickname" to denote a 16-bit value that serves as an abbreviation for the "RBridge ID." In this disclosure, "RBridge identifier" is used as a generic term and is not limited to any bit format, and can refer to "RBridge ID" or "RBridge nickname" or any other format that can identify an RBridge.

Network Architecture

FIG. **1** illustrates an exemplary TRILL network that includes a plurality of RBridges with IP processing capabilities, in accordance with an embodiment of the present invention. As illustrated in FIG. **1**, a TRILL network **100** includes RBridges, **101**, **102**, **103**, **104**, **105**, **106**, and **107**. RBridges **101**, **102**, and **103** are IP capable and coupled to a layer-3 network **150** as IP routers, **111**, **112**, and **113**, respectively. For example, RBridge **101** and IP router **111** are the same physical device (represented by dotted lines), where its TRILL RBridge portion is denoted by RBridges **101** and its IP router portion is denoted by router **111**. Similarly, RBridge **102** and IP router **112**, and RBridge **103** and IP router **113**, are the same physical devices, respectively.

RBridges in network **100** use edge ports to communicate to end devices and TRILL ports to communicate to other RBridges. For example, RBridge **104** is coupled to end device **122** via an edge port and to RBridges **105**, **101**, and **102** via TRILL ports. An end host coupled to an edge port may be a host machine or an aggregation node. For example, end devices **122**, **124**, **126**, and **128** are host machines, wherein end devices **122** and **128** are directly coupled to network **100**, and end devices **124** and **126** are coupled to network **100** via their aggregation node, a layer-2 bridge **130**.

In FIG. **1**, end device **128** is directly coupled to RBridge **103**. Hence, IP router **113** can act as the default gateway for end device **128**. Consequently, all frames from end device **128** destined to IP network **150** are received at IP router **113** and forwarded to network **150**. On the other hand, RBridge **104** couples end device **122** to network **100** and acts as the ingress RBridge for all frames from end device **122**. One of the IP-capable RBridges (e.g., RBridge **101**) acts as the egress RBridge for frames from end device **122** to network **150**. Under such a scenario, the frame destined to network **150** is encapsulated in a TRILL packet with the RBridge identifier of RBridge **101** as the destination RBridge identifier, and the MAC address of RBridge **101** as the inner destination MAC address. The TRILL packet is then forwarded to RBridge **101**, where the outer destination MAC address of the packet is the MAC address of RBridge **101**. IP router **111** then processes the IP header in the frame and makes the layer-3 forwarding decision based on the destination IP address of the frame.

During operation that does not involve layer-3 processing in RBridges, an end device coupled to the TRILL network may select the default gateway from a layer-3 network and use the corresponding IP address as a default gateway router address. For example, in FIG. **1**, end device **128** selects the default gateway router from IP network **150**. Any frame destined to network **150** from end device **128** is sent to the default gateway. Under such a scenario, if end devices **122** and **128** are on different VLANs, any communication between these end devices will go through network **150**. If end device **128** sends a frame to end device **122**, the frame first goes to the default gateway in network **150**. Consequently, the default gateway processes the IP header in the frame and makes layer-3 forwarding decision toward end device **122**. As a result, routing and bandwidth management will be inefficient and the frame will incur higher latency.

In embodiments of the present invention, as illustrated in FIG. **1**, each frame destined to end device **122** from end device **128**, wherein the end devices are on different VLANs, is received at RBridge **103**. IP router **113** processes the IP header in the frame and makes the forwarding decision toward end device **122** (which involves forwarding the frame on end device **122**'s VLAN through TRILL network **100**). Consequently, RBridge **103** forwards the frame to a corresponding egress RBridge **104** over TRILL network **100**. RBridge **104**, in turn, transmits the frame to end device **122**. Hence, enabling layer-3 support on RBridges in a TRILL network provides higher efficiency in routing and bandwidth management.

In some embodiments, the TRILL network may be a virtual cluster switch (VCS). In a VCS, any number of RBridges in any arbitrary topology may logically operate as a single switch. Any new RBridge may join or leave the VCS in "plug-and-play" mode without any manual configuration.

Note that TRILL is only used as a transport between the switches within network **100**. This is because TRILL can readily accommodate native Ethernet frames. Also, the TRILL standards provide a ready-to-use forwarding mechanism that can be used in any routed network with arbitrary topology. Embodiments of the present invention should not be limited to using only TRILL as the transport. Other protocols (such as multi-protocol label switching (MPLS)), either public or proprietary, can also be used for the transport.

Routine Across VLANs

FIG. **2**A illustrates an exemplary configuration of how end devices belonging to different VLANs and coupled to a TRILL network, wherein one RBridge is IP capable, in accordance with an embodiment of the present invention. In this example, a TRILL network **200** includes TRILL RBridges **220** and **230**. End device **202** is coupled to RBridge **220** over VLAN **212**, and end device **204** is coupled to RBridge **220** over VLAN **214**.

In the example in FIG. **2**A, RBridge **230** is IP capable and IP router **235** is the IP router portion of RBridge **230** (denoted in dotted line). IP router **235** functions as a default gateway router for end devices **202** and **204**. Consequently, although RBridge **220** couples both end devices **202** and **204** to network **200**, any traffic between end devices **202** and **204** will be routed via IP router **235** because end devices **202** and **204** belong to different VLANs. For example, if end device **202** sends a frame to end device **204**, it first assembles an IP packet with end device **204**'s IP address. Based on its local forwarding table, end device **202** realizes that it does not have a direct route to end device **204**, and therefore needs to send the packet to gateway router **235**. Hence, end device **202** encapsulates the IP packet in an Ethernet frame, whose destination MAC address is set to be gateway router **235**'s MAC address. Note that, if end device **202** has no knowledge of IP router **235**'s MAC address, end device **202** can send out an ARP request corresponding to the IP address of router **235**. Router **235** then replies to the ARP request with its MAC address. Subsequently, end device **202** forwards the frame to RBridge **230** via ingress RBridge **220**. IP router **235**, in turn, receives the frame and removes its layer-2 header (including the VLAN tag corresponding to VLAN **212**). IP router **235** then performs a lookup in its IP forwarding table based on the packet's destination IP address, and encapsulates the packet with a new Ethernet header which includes a VLAN tag corresponding VLAN **214**. RBridge **230** then encapsulates the Ethernet frame with a TRILL header and forwards it to end device **204** via egress RBridge **220**.

FIG. **2**B illustrates an exemplary configuration of end devices belonging to different VLANs and coupled to a TRILL network, wherein all RBridges are IP capable, in accordance with an embodiment of the present invention. In this example, a TRILL network **200** includes TRILL RBridges **220** and **230**. End device **202** is coupled to RBridge **220** over VLAN **212**, and end device **204** is coupled to RBridge **220** over VLAN **214**.

In the example in FIG. **2**B, both RBridges **220** and **230** are IP capable and IP routers **225** and **235** are the IP router portion of RBridges **220** and **230**, respectively. Under such a scenario, IP router **225** can be the default gateway router for end devices **202** and **204**. Consequently, any traffic between end devices **202** and **204** can be routed via IP router **225**. For example, if end device **202** sends a frame to end device **204**, it assembles an IP packet with end device **204**'s IP address, encapsulates the IP packet in an Ethernet frame with destination MAC address as router **225**'s MAC address, and forwards the frame to RBridge **225** via ingress RBridge **220**. Note that, if end device **202** has no knowledge of IP router **225**'s MAC address, end device **202** obtains the IP address of router **225** using ARP. IP router **225**, in turn, receives the frame, performs a lookup in its IP forwarding table, encapsulates the packet with a new Ethernet header which includes a VLAN tag corresponding VLAN **214**, and forwards it to end device **204** via egress RBridge **220**. As the cross-layer-2-domain frame does not need to traverse through TRILL network **200**, IP-processing capability at RBridge **220** thereby reduces the bandwidth usage in network **200**.

Distributed Layer-3 Processing

In some embodiments, layer-3 processing capabilities can be distributed to multiple or all TRILL RBridges. In some embodiments, layer-3 processing capabilities associated different the VLANs can be distributed selectively across multiple RBridges. FIG. **3**A illustrates an exemplary TRILL network with multiple VLANs, in accordance with an embodiment of the present invention. In the example in FIG. **3**A, network **300** includes RBridges **304**, **305**, **306**, and **307**. Each of these RBridges is IP capable. RBridge **304** is coupled to end devices **311** and **312**; RBridge **305** is coupled to end devices **312**, **313**, and **314**; RBridge **306** is coupled to end devices **315**, **316**, and **317**; and RBridge **307** is coupled to end devices **317** and **318**. RBridges **305** and **306** belong to VLAN **328**; RBridges **304**, **306**, and **307**, and end devices **312** and **318** belong to VLAN **326**; RBridges **304**, **305**, and **306**, and end device **311** belong to VLAN **324**; and RBridges **305**, **306**, and **307**, and end device **317** belong to VLAN **322**.

In some embodiments, a layer-3 interface on an RBridge corresponding to a VLAN is a Switch Virtual Interface (SVI). For example, RBridge **304** in FIG. **3**A has SVIs for VLANs **324** and **326** (although these SVIs can be on the same physical interface). Consequently, RBridge **304** and end device **318**, and RBridge **304** and end device **311**, are on the same VLAN segment. If end device **311** sends a frame to end device **318**, the destination is outside of VLAN **324**. Consequently, end device **318** sets the destination MAC address of the frame as the MAC address of the SVI on VLAN **324** at RBridge **304**, which is the layer-3 gateway on VLAN **324**. End device **318** then forwards the frame to RBridge **304**. Upon receiving the frame, RBridge **304** recognizes that the frame's destination MAC address is a local MAC address. RBridge **304** then removes the frame's Ethernet header, performs a lookup in its IP forwarding table based on the frame's destination IP address, and encapsulates the frame with a new Ethernet header with a destination MAC address corresponding to end device **318** in VLAN **326**. Finally, RBridge **304** forwards the frame to end device **318** via egress RBridge **307**.

However, when end device **311** sends a frame to end device **317**, RBridge **304** cannot forward the frame to end device **317** because RBridge **304** does not have an SVI on VLAN **322**, to which end device **317** belongs. As a result, upon receiving a frame destined to end device **317** from end device **311**, RBridge encapsulates the frame using a TRILL header with egress RBridge identifier corresponding to RBridge **306** because it has SVIs to all VLANs. RBridge **304** then forwards the frame to RBridge **306**. The frame is routed though the TRILL network and reaches RBridge **306** when the outer destination MAC addresses match the MAC address of RBridge **306**. Upon receiving the frame, RBridge **306** recognizes that the frame's outer destination MAC address is a local MAC address. RBridge **306** then removes the TRILL encapsulations, encapsulates the IP packet with a new Ethernet header with a destination MAC address corresponding to end device **317** in VLAN **322**, and forwards the frame accordingly.

FIG. **3**B illustrates an exemplary TRILL network with multiple VLANs, wherein each RBridge belongs to all VLANs, in accordance with an embodiment of the present invention. In this example, TRILL network **300** includes RBridges **304**, **305**, **306**, and **307**. Each of these RBridges is IP capable. End device **312** is coupled to RBridges **304** and **305**, end device **317** is coupled to RBridges **306** and **307**, and end device **318** is coupled to RBridge **307**. All RBridges in network **300** have SVIs for VLANs **322** and **326**. End devices **312** and **318** belong to VLAN **322**, and end device **317** belongs to VLAN **326**.

In this example, if end device **317** sends a frame to end device **318**, the frame can be routed on layer-3 at RBridge **307** because RBridge **307** has SVIs for VLANs **322** and **326**. As the frame does not travel to any other RBridge in network **300**, it incurs lower latency while saving bandwidth in network **300**. Similarly, if end device **317** sends a frame to end device **312**, the frame can be routed on layer-3 at the IP router portion of either RBridge **306** or **307** as both have SVIs for VLANs **322** and **326**. If all RBridges in the TRILL network have SVIs for all VLANs, inter-VLAN switching is possible at each RBridge.

Frame Processing

FIG. **4**A presents a flowchart illustrating the process of an RBridge transmitting a frame, in accordance with an embodiment of the present invention. During operation, an RBridge receives a frame (operation **402**) and determines the type of port at which the frame was received (operation **404**). If the frame is received at an edge port, then the RBridge checks whether the destination is coupled to a local edge port (operation **410**). If the destination is not coupled to a local edge port, the RBridge encapsulates the frame in a TRILL packet and sets the RBridge identifier of the RBridge to which the end device is coupled as the egress RBridge identifier (operation **416**). The RBridge then forwards the TRILL packet to the TRILL network (operation **418**). Note that the MAC learning process allows an RBridge to learn about the port to which the end device is coupled.

If the frame is received on an edge port and the destination is coupled to a local edge port (operation **410**), then the

RBridge transmits the frame to the destination end device coupled to a local edge port (operation **414**).

If the frame is received from a TRILL port (operation **404**), the RBridge checks whether itself is the egress RBridge of the TRILL packet (operation **408**). If not, then the RBridge forwards the TRILL packet to the TRILL network (operation **418**). Otherwise, the RBridge transmits the frame to the destination end device coupled to a local edge port (operation **414**).

FIG. **4**B presents a flowchart illustrating the process of an IP-capable RBridge transmitting a frame, in accordance with an embodiment of the present invention. The exemplary process in FIG. **4**B is also applicable to embodiments with distributed layer-3 processing, as described in conjunction with FIG. **3**A. During operation, an RBridge receives a frame (operation **452**) and determines the type of port at which the frame is received (operation **454**). If the frame is received at an edge port, then the RBridge inspects the frame to determine whether the end device with the destination MAC address is coupled to a local edge port (operation **456**). If so, the frame is forwarded to the destination via the TRILL network (operation **464**), as described in conjunction with FIG. **4**A.

If the frame's destination MAC address is not coupled to a local edge port, then the RBridge determines whether the frame's destination MAC address is the RBridge's MAC address (operation **458**). If the destination MAC address is not the RBridge's MAC address, then the RBridge encapsulates the frame in a TRILL packet and sets the RBridge identifier of a gateway RBridge as the egress RBridge identifier (operation **466**). The RBridge then forwards the TRILL packet to the TRILL network (operation **476**). On the other hand, if the frame's destination MAC address is the RBridge's MAC address (operation **458**), then the RBridge performs layer-3 processing on the frame (operation **468**) and determines the outgoing port (operation **470**).

The RBridge then determines the type of the outgoing port (operation **462**). If the outgoing port is an edge port, which means the destination end device is coupled locally, the RBridge forwards the frame, which is Ethernet encapsulated with the end device's MAC address as the destination MAC address, to the destination end device (operation **480**). In some embodiments, the end device can be a layer-3 (e.g., IP) router. If the outgoing port is a TRILL port, then the end device is connected to a remote RBridge. Hence, the RBridge obtains the RBridge identifier of the RBridge to which the destination end device is coupled to based on the MAC address of the destination end device (operation **472**). The RBridge then encapsulates the frame in a TRILL packet and sets the obtained RBridge identifier as the egress RBridge identifier (operation **474**). The RBridge then forwards the TRILL packet to the TRILL network (operation **476**).

If the frame is received from a TRILL port (operation **454**), the RBridge checks whether itself is the egress RBridge of the TRILL packet (operation **460**). If not, then the RBridge forwards the TRILL packet to the TRILL network (operation **476**). Otherwise, the RBridge forwards the frame to the destination end device coupled to a local edge port (operation **480**). In some embodiments, the end device can be a layer-3 router, in which case the forwarding includes layer-3 processing on the frame.

Virtual Switch Formation

In some embodiments, a number of TRILL RBridges with IP processing capabilities may act as layer-3 routers for an end device. These RBridges can form a virtual RBridge, which is assigned with a virtual RBridge identifier. Furthermore, these RBridges form a virtual IP router, which is assigned with a virtual IP address and a corresponding virtual MAC address. This virtual IP router operates as a default gateway router, which can provide redundancy and load balancing.

FIG. **5** illustrates an exemplary network where a virtual RBridge and an associated virtual IP router are created based on a plurality of physical gateway RBridges with IP processing capabilities, in accordance with an embodiment of the present invention. As illustrated in FIG. **5**, a TRILL network **500** includes RBridges **504**, **505**, **506**, **507**, **511**, **512**, and **513**. RBridges **511**, **512**, and **513** operate as gateway RBridges and are coupled to a layer-3 network **150** as IP routers **521**, **522**, and **523**, respectively. For example, gateway RBridge **511** and IP router **521** are same physical device (represented by dotted lines), where its TRILL RBridge portion is denoted by gateway RBridge **511** and its IP router portion is denoted by IP router **521**. Similarly, gateway RBridge **512** and IP router **522**, and gateway RBridge **513** and IP router **523** are the same physical devices, respectively.

Gateway RBridges **511**, **512**, and **513** form a virtual RBridge **530** by operating as a single logical RBridge in TRILL network **500**. Similarly, the corresponding IP routers **521**, **522**, and **523** form a virtual IP router **540** by operating as a single logical IP router. An end device **562** coupled to network **500** through RBridge **507** can use virtual IP router **540** as the default gateway router to layer-3 network **550**.

In embodiments of the present invention, as illustrated in FIG. **1**, Virtual RBridge **530** is considered to be logically coupled to gateway RBridges **511**, **512**, and **513**, optionally with zero-cost links represented by dashed lines. Furthermore, gateway RBridges **511**, **512**, and **513** can advertise their respective connectivity (optionally via zero-cost links) to virtual RBridge **530**. As a result, other RBridges in the TRILL network can learn that virtual RBridge **530** is reachable via gateway RBridges **511**, **512**, and **513**, and establish TRILL paths to virtual RBridge **530** using a corresponding virtual RBridge identifier through these gateway RBridges.

All the IP-layer router portions of these gateway RBridges are configured to operate as the layer-3 gateway router (i.e., virtual IP router **540**) for end device **562**. End device **562** uses virtual IP router **540** as the default gateway. Because virtual RBridge **530** is associated with virtual IP router **540**, incoming frames from end device **562** destined to network **550** are marked with virtual RBridge **530**'s identifier as the egress RBridge identifier. Consequently, all frames from end device **562** to network **550** are delivered to one of the gateway RBridges **511**, **512**, and **513**. Hence, load balancing can be achieved among gateway RBridges **511**, **512**, and **513** for frames sent to virtual RBridge **530**.

FIG. **6**A illustrates an exemplary configuration of how a virtual RBridge and an associated virtual IP router can be logically coupled to a number of gateway RBridges in a TRILL network, in accordance with an embodiment of the present invention. In this example, a TRILL network **600** includes a number of TRILL RBridges **602**, **604**, and **606**. Network **600** also includes RBridges **616** and **618**, each with a number of edge ports which can be coupled to external networks. For example, RBridges **616** and **618** are coupled with end devices **652** and **654** via 10GE edge ports. RBridges in network **600** are in communication with each other using TRILL protocol.

Also included in network **600** are RBridges **622** and **624**, which are layer-3 capable and coupled to an IP network **680**. Gateway RBridges **622** and **624** form virtual RBridge **640** with a virtual RBridge identifier **645**. Physically co-located IP Routers **632** and **634** within gateway RBridges **622** and **624**, respectively, form a virtual IP router **670** which is assigned a virtual IP address **660** and a virtual MAC address **650**. Virtual IP address **660** maps to virtual MAC address **650** for ARP requests directed to virtual IP router **670**. Furthermore, virtual RBridge identifier **645** is associated with virtual MAC address **650**. End devices **652** and **654** can set virtual IP address **660** as their default gateway router address and use ARP to obtain virtual MAC address **650**. End devices **652** and **654** send frames with virtual MAC address **650** as the destination address into network **600**. The frames are encapsulated in TRILL packets and routed toward virtual RBridge **640** using the corresponding virtual RBridge identifier **645**.

In some embodiments, a virtual IP address can be assigned for each VLAN associated with a TRILL network. For example, in FIG. **6**A, end device **652** may belong to VLAN **692**, and end device **654** may belong to VLAN **694**. Different virtual IP addresses may be used for VLANs **692** and **694**, respectively. End devices **652** and **654** then use the virtual IP address associated with VLAN **692** and VLAN **694** as their respective default gateway router addresses. Consequently, end devices **652** and **654** perceive virtual IP router **670** to be in VLAN **692** and VLAN **694**, respectively. For ARP requests for either virtual IP address, the same virtual MAC address **650** is sent in reply. All data frames injected to TRILL network **600** with virtual MAC address **650** as the destination MAC address are routed toward virtual RBridge **640**.

Note that in one embodiment, the virtual MAC address is known to all RBridges in the network **600**. Otherwise, both IP routers **632** and **634** receive a frame forwarded to virtual MAC address **650** and results in packet duplication. Hence, after formation of virtual RBridge **640** and virtual IP router **670**, all RBridges in network **600** are provided with the knowledge about virtual MAC address **650**. That is, virtual MAC address **650** is always "known" to all ingress RBridges in network **600**, and frames destined to virtual MAC address **650** are routed through network **600** using TRILL unicast.

In some embodiments, only one gateway RBridge is elected to reply to ARP requests for the virtual IP address. This election can also be VLAN specific.

FIG. **6**B illustrates an exemplary configuration of how a virtual RBridge and an associated virtual IP router can be logically coupled to all RBridges in a TRILL network where each RBridge has IP processing capability, in accordance with an embodiment of the present invention. In this example, all RBridges in TRILL network **600** have IP processing capabilities. Even though only RBridges **622** and **624** are connected to an IP network, IP processing capacity at all RBridges enables them to route across VLANs, as described in conjunction with FIG. **3**B. For example, any traffic between VLANs **692** and **694** can be switched at RBridges **616** and **618** without requiring the traffic to travel to another RBridge in network **600**.

In some embodiments, all RBridges in network **600** are associated with virtual RBridge **640** and a virtual IP router **670**, and share a virtual RBridge identifier **645**, a virtual IP address **660**, and a virtual MAC address **650**. In some embodiments, all RBridges in network **600** may be connected to IP network **680**.

ARP and Frame Processing in a Virtual Switch

FIG. **7**A presents a flowchart illustrating the process of a gateway RBridge associated with a virtual RBridge responding to an Address Resolution Protocol (ARP) query, in accordance with an embodiment of the present invention. Upon receiving an ARP request packet for an IP address (operation **702**), the gateway RBridge checks whether the ARP request is for a virtual IP address (operation **704**). If not, the gateway RBridge responds based on the IP address in the ARP request (assuming that IP address is the gateway RBridge's physical IP address) (operation **720**). Otherwise, the gateway RBridge checks whether it is elected to respond to an ARP request for the virtual IP address (operation **706**). If not, the ARP request is discarded. Otherwise the gateway RBridge retrieves the virtual MAC address for the virtual IP address (operation **708**) and generates an ARP reply containing the virtual MAC address (operation **710**). The gateway RBridge transmits the ARP reply to the TRILL network (operation **712**). Note that an ARP request is disseminated in the TRILL network using multicast and each IP-capable RBridge, including the one elected to respond to ARP requests for the virtual IP address, receives the query. However, the ARP reply is sent as a unicast transmission in the TRILL network to the end device.

FIG. **7**B presents a flowchart illustrating the process of a gateway RBridge associated with a virtual RBridge forwarding a TRILL frame, in accordance with an embodiment of the present invention. Upon receiving a TRILL frame (operation **752**), the RBridge checks whether the egress RBridge identifier in the TRILL header of the frame corresponds to a virtual RBridge (operation **754**). If the identifier does not correspond to the virtual RBridge, the RBridge inspects whether the egress RBridge identifier in the TRILL header of the frame corresponds to the local RBridge. If not, then the TRILL frame is forwarded to the next-hop RBridge based on the egress RBridge identifier (operation **762**). Otherwise, the RBridge removes the TRILL encapsulation and send the frame to a local egress port (operation **764**). If the RBridge identifier corresponds to the virtual RBridge, the RBridge checks whether the destination MAC address of the Ethernet frame encapsulated in the TRILL frame is the associated virtual MAC address (operation **756**). If so, then the frame is destined to an IP network the gateway RBridge is coupled to. Hence, the IP packet is extracted from the Ethernet payload of the frame (operation **772**). The gateway RBridge checks the IP address of the IP packet and performs layer-3 IP forwarding toward the IP network (operation **774**). On the other hand, if the destination MAC address is not the virtual MAC address, then the virtual RBridge is for multi-homed layer-2 end devices. Accordingly, the RBridge removes the TRILL encapsulation and send the frame to locally connected egress port (operation **764**). Operation of virtual RBridges for multi-homed end devices, such as forwarding multicast frames, is specified in the U.S. Patent Publication No. 2010/0246388, titled "Redundant Host Connection in a Routed Network," the disclosure of which is incorporated herein in its entirety.

Failure Handling

FIG. **8** illustrates a scenario where one of the RBridges associated with the virtual RBridge experiences a link failure and/or a node failure, in accordance with an embodiment of the present invention. In this example, in a TRILL network **800**, RBridges **811**, **812**, and **813** form a virtual RBridge **840**, and their respective IP-router portions denoted as IP routers **821**, **822**, and **823** form a virtual IP router **850**. Also included are four RBridges **804**, **805**, **806**, and **807**. An end device **870** is connected to network **800** using RBridge **804** as the ingress RBridge. Virtual IP router **850** is set as a default gateway router for end device **870**. Hence, all frames destined to network **880** from end device **870** have the virtual MAC address assigned to virtual IP router **850** as the destination MAC address. Note that these frames can be forwarded by gateway RBridges **811**, **812**, and **813** for load balancing. Gateway RBridges **811**, **812**, and **813** also provide redundancy among each other to handle failures.

Suppose that a failure **864** occurs to link **831** adjacent to gateway RBridge **811**. As a result, link **831** is removed from routing decisions in network **800**. All frames from end device **870** are still using the virtual MAC address as the destination address, and thus are still forwarded to any of the gateway RBridges via alternative links (e.g., links **832**, **833**, and **834**).

Suppose that a failure **862** occurs during operation that fails link **836** adjacent to IP router **821**. Consequently, IP router **821** is disconnected from network **880** and is incapable of forwarding frames to network **880**. Under such a scenario, IP router **821** is removed from virtual IP router **850**. As a result, IP router **821** stops operating as a layer-3 gateway router for end device **870**. However, gateway RBridge **811** still remains connected to network **800** and continues to operate as a regular TRILL RBridge. As virtual IP router **850** still operates as a default gateway for end device **870**, IP routers **822** and **823** can continue to operate as layer-3 gateway routers (as virtual IP router **850**) for end device **870**. Hence, all frames from end device **870** to network **880** are then distributed among gateway RBridges **812** and **813**.

In some embodiments, with failure **862**, an elected gateway RBridge stops responding to ARP requests for the virtual IP address and notifies other gateway RBridges. Consequently, the other gateway RBridges then elect among themselves another gateway RBridge to respond to ARP requests.

In some embodiments, with failure **862**, IP router **821** might not immediately remove its membership from virtual IP router **850** and might continue to receive layer-3 traffic from end devices. Under such circumstances, gateway RBridge **811**, the TRILL counterpart of IP router **821**, forwards the layer-3 traffic with TRILL encapsulation to other gateway RBridges (e.g., gateway RBridge **812**) which, in turn, forward the traffic to network **880**. However, if all similar IP routers suffer link failures and lose their connection to network **880**, IP router **821** along with the other gateway RBridges with link failures are removed from virtual IP router **850**. However, all gateway RBridges continue operating as TRILL RBridges.

Suppose that a node failure **866** occurs at gateway RBridge **811** (and essentially IP router **821** as they are the same physical device). As a result, links **831**, **833**, **835**, and **836** fail as well. Consequently, gateway RBridge **811** and IP router **821** are disconnected from both network **800** and network **880**, and are incapable of transmitting to or receiving from either network. Under such a scenario, IP router **821** is removed from virtual IP router **850** and gateway RBridge **811** is removed from virtual RBridge **840**. As a result, IP router **821** stops operating as a layer-3 gateway node. Furthermore, gateway RBridge **811** is disconnected from network **800** and removed from all TRILL routes in network **800**.

With failure **866**, as virtual IP router **850** still operates as a default gateway for end device **870**, routers **822** and **823** continue operating as layer-3 gateway nodes for end device **870**. Hence, all frames from end device **870** to network **880** are distributed between gateway RBridges **812** and **813**. Furthermore, if IP router **821** had been an elected router, it stops responding to ARP requests for the virtual IP address. Other RBridges coupled to the failed gateway RBridge can detect the failure and notify all RBridges, including other active gateway RBridges. Consequently, the active gateway RBridges can elect another gateway RBridge to respond to ARP requests.

Exemplary Switch System

FIG. **9** illustrates an exemplary architecture of a switch with IP processing capabilities, in accordance with an embodiment of the present invention. In this example, an RBridge **900** includes a number of TRILL ports **904**, a TRILL management and forwarding module **920**, an IP management module **930**, an Ethernet frame processor **910**, and a storage **950**. TRILL management and forwarding module **920** further includes a TRILL header processing module **922**. IP management module **930** further includes an ARP module **934** and an IP header processing module **936**.

TRILL ports **904** include inter-switch communication channels for communication with one or more RBridges. This inter-switch communication channel can be implemented via a regular communication port and based on any open or proprietary format. Furthermore, the inter-switch communication between RBridges is not required to be direct port-to-port communication.

During operation, TRILL ports **904** receive TRILL frames from (and transmit frames to) other RBridges. TRILL header processing module **922** processes TRILL header information of the received frames and performs routing on the received frames based on their TRILL headers, as described in conjunction with FIG. **4**B. TRILL management and forwarding module **920** forwards frames in the TRILL network toward other RBridges and frames destined to a layer-3 node toward the IP management module **930**. IP header processing module **936** forwards frames across VLANs.

In some embodiments, RBridge **900** may form a virtual RBridge and a virtual IP address, wherein TRILL management and forwarding module **920** further includes a virtual RBridge configuration module **924**, and IP management module **930** further includes a virtual IP router configuration module **938**. TRILL header processing module **922** generates the TRILL header and outer Ethernet header for ingress frames corresponding to the virtual RBridge. Virtual RBridge configuration module **924** manages the communication with gateway RBridges and handles various inter-switch communications, such as link and node failure notifications. Virtual RBridge configuration module **924** allows a user to configure and assign the identifier for the virtual RBridges, and decides whether a frame has to be promoted to layer-3, as described in conjunction with FIG. **7**B.

Furthermore, virtual IP router configuration module **938** handles various inter-switch communications, such as layer-3 link failure notifications. Virtual IP router configuration module **938** allows a user to configure and assign virtual IP addresses and a virtual MAC address.

ARP module **934** is responsible for ARP request replies, as described in conjunction with FIG. **4**B. ARP module **934** also maintains mappings between a virtual MAC address and a virtual IP address and stores the mappings in Storage **950**. Storage **950** also includes TRILL and IP routing information.

In some embodiments, gateway RBridge **900** may include a number of edge ports **902**, as described in conjunction with FIG. **1**. Edge ports **902** receive frames from (and transmit frames to) end devices. Ethernet frame processor **910** extracts and processes header information from the received frames. Ethernet frame processor **910** forwards the frames to IP management module **930** if there is no other intermediate RBridge between the end device and RBridge **900**.

In some embodiments, gateway RBridge **900** may include a VCS configuration module **944** that includes a virtual switch management module **940** and a logical switch **942** as described in conjunction with FIG. **1**. VCS configuration module **944** maintains a configuration database in storage **950** that maintains the configuration state of every switch within the VCS. Virtual switch management module **940** maintains the state of logical switch **942**, which is used to join other VCS switches. In some embodiments, logical switch **942** can be configured to operate in conjunction with Ethernet frame processor **910** as a logical Ethernet switch.

Note that the above-mentioned modules can be implemented in hardware as well as in software. In one embodiment, these modules can be embodied in computer-executable instructions stored in a memory which is coupled to one or more processors in gateway RBridge **900**. When executed, these instructions cause the processor(s) to perform the aforementioned functions.

In summary, embodiments of the present invention provide a switch, a method and a system for providing layer-3 support in a TRILL network. In one embodiment, the switch includes an IP header processor and a forwarding mechanism. The IP header processor identifies a destination IP address in a packet encapsulated with an inner Ethernet header, a TRILL header, and an outer Ethernet header. The forwarding mechanism determines an output port and constructs a new header for the packet based on the destination IP address. The switch also includes a packet processor which determines whether (1) an inner destination media access control (MAC) address corresponds to a local MAC address assigned to the switch; (2) a destination RBridge identifier corresponds to a local RBridge identifier assigned to the switch; and (3) an outer destination MAC address corresponds to the local MAC address. Such configuration provides a scalable and flexible solution to enable layer-3 processing in the switch.

The methods and processes described herein can be embodied as code and/or data, which can be stored in a computer-readable non-transitory storage medium. When a computer system reads and executes the code and/or data stored on the computer-readable non-transitory storage medium, the computer system performs the methods and processes embodied as data structures and code and stored within the medium.

The methods and processes described herein can be executed by and/or included in hardware modules or apparatus. These modules or apparatus may include, but are not limited to, an application-specific integrated circuit (ASIC) chip, a field-programmable gate array (FPGA), a dedicated or shared processor that executes a particular software module or a piece of code at a particular time, and/or other programmable-logic devices now known or later developed. When the hardware modules or apparatus are activated, they perform the methods and processes included within them.

The foregoing descriptions of embodiments of the present invention have been presented only for purposes of illustration and description. They are not intended to be exhaustive or to limit this disclosure. Accordingly, many modifications and variations will be apparent to practitioners skilled in the art. The scope of the present invention is defined by the appended claims.

What is claimed is:

1. A switch, comprising:

layer-2 processing circuitry configured to determine that:

outer and inner destination media access control (MAC) addresses of a packet correspond to a MAC address assigned to the switch, wherein the packet is encapsulated with an inner Ethernet header, a routable header, and an outer Ethernet header;

encapsulation circuitry configured to determine that:

a destination switch identifier of the routable header corresponds to a switch identifier assigned to the switch, wherein the routable header is placed between the outer and inner Ethernet headers;

Internet Protocol (IP) processing circuitry configured to lookup a destination IP address of a layer-3 header of the packet in a local layer-3 forwarding table in the switch, wherein the layer-3 header is distinct from the routable header, and wherein the destination IP address is a virtual IP address assigned to a virtual IP router, which is formed based on the switch in conjunction with at least another physical switch to operate as a single router; and

forwarding circuitry configured to determine an output port and construct a new header for the packet based on looking up the destination IP address in the local layer-3 forwarding table.

2. The switch of claim **1**, wherein the layer-2 processing circuitry is further configured to determine a first virtual local area network (VLAN) tag in the inner Ethernet header; and

wherein the new header includes a new inner Ethernet header comprising a second VLAN tag.

3. The switch of claim **1**, wherein the switch is a member of a network of interconnected switches, wherein the network of interconnected switches is controlled as a single logical switch.

4. The switch of claim **1**, further comprising switching circuitry configured to switch the packet between VLANs based on the destination IP address.

5. The switch of claim **1**, wherein the destination switch identifier is a virtual switch identifier; and

wherein the virtual IP router is associated with the virtual switch identifier.

6. The switch of claim **1**, wherein the IP processing circuitry is further configured to map the virtual IP address to a virtual media access control (MAC) address.

7. The switch of claim **1**, further comprising Address Resolution Protocol (ARP) circuitry configured to generate an ARP response for an IP address assigned to the switch, wherein the ARP response comprises a MAC address assigned to the switch.

8. A method, comprising:

determining that:

outer and inner destination media access control (MAC) addresses of a packet correspond to a MAC address assigned to a switch, wherein the packet is encapsulated with an inner Ethernet header, a routable header, and an outer Ethernet header; and

a destination switch identifier of the routable header corresponds to a switch identifier assigned to the switch, wherein the routable header is placed between the outer and inner Ethernet headers;

looking up a destination Internet Protocol (IP) address of a layer-3 header of the packet in a local layer-3 forwarding table in the switch, wherein the layer-3 header is distinct from the routable header, and wherein the destination IP address is a virtual IP address assigned to a virtual IP router, which is formed based on the switch in conjunction with at least another physical switch to operate as a single router; and

determining an output port and constructing a new header for the packet based on looking up the destination IP address in the local layer-3 forwarding table.

9. The method of claim **8**, further comprising:

determining a first virtual local area network (VLAN) tag in the inner Ethernet header; and

including in the new header a new inner Ethernet header comprising a second VLAN tag.

10. The method of claim **8**, wherein the switch is a member of a network of interconnected switches wherein the network of interconnected switches is controlled as a single logical switch.

11. The method of claim **8**, further comprising switching the packet between VLANs based on the destination IP address.

12. The method of claim **8**, wherein the destination switch identifier is a virtual switch identifier; and

wherein the virtual IP router is associated with the virtual switch identifier.

13. The method of claim **8**, further comprising mapping the virtual IP address to a virtual media access control (MAC) address.

14. The method of claim **8**, further comprising generating an Address Resolution Protocol (ARP) response for an IP address assigned to the switch, wherein the ARP response comprises a MAC address assigned to the switch.

15. A computing system, comprising:

a processor; and

a non-transitory computer-readable storage medium storing instructions which when executed by the processor causes the processor to perform a method, the method comprising:

determining that:

outer and inner destination media access control (MAC) addresses of a packet correspond to a MAC address assigned to the computing system, wherein the packet is encapsulated with an inner Ethernet header, a routable header, and an outer Ethernet header; and

a destination switch identifier of the routable header corresponds to a switch identifier is assigned to the computing system;

looking up a destination Internet Protocol (IP) address of a layer-3 header of the packet in a local layer-3 forwarding table in the computing system, wherein the layer-3 header is distinct from the routable header and wherein the destination IP address is a virtual IP address assigned to a virtual IP router, which is formed based on the switch in conjunction with at least another physical switch to operate as a single router; and

determining an output port and constructing a new header for the packet based on looking up the destination IP address in the local-layer-3 forwarding table.

16. The computing system of claim **15**, within the method further comprises:

determining a first virtual local area network (VLAN) tag in the inner Ethernet header; and

including in the new header a new inner Ethernet header comprising a second VLAN tag.

17. The computing system of claim **15**, wherein the computing system is a member of a network of interconnected switches, wherein the network of interconnected switches is controlled as a single logical switch.

18. The computing system of claim **15**,

wherein the destination switch identifier is a virtual switch identifier; and

wherein the virtual IP router is associated with the virtual switch identifier.

19. The computing system of claim **15**, wherein the method further comprises generating an Address Resolution Protocol (ARP) response for an IP address assigned to the computing system, wherein the ARP response comprises a MAC address assigned to the computing system.

20. The computing system of claim **15**, whether the method further comprises switching the packet between VLANs based on the destination IP address.

* * * * *