RESILIENT/ADVANCED METHODS TO DO TIMESYNC USING 1588(PTP) BY INTERWORKING WITH OTHER PROTOCOLS.

WSTS 2018, San Jose, 2018

M Ramana Reddy(mrreddy@juniper.net), Juniper Networks

Prabhu Raj VK(prabraj@juniper.net), Juniper Networks

Agenda

1. Introduction to Aggregate Ethernet(LAG) Networks
2. Synchronization Overview and Precision Timing Protocol
3. Current industry challenges with time sync using PTP over LAG(AE) Networks
4. Proposed Solution
5. Summary

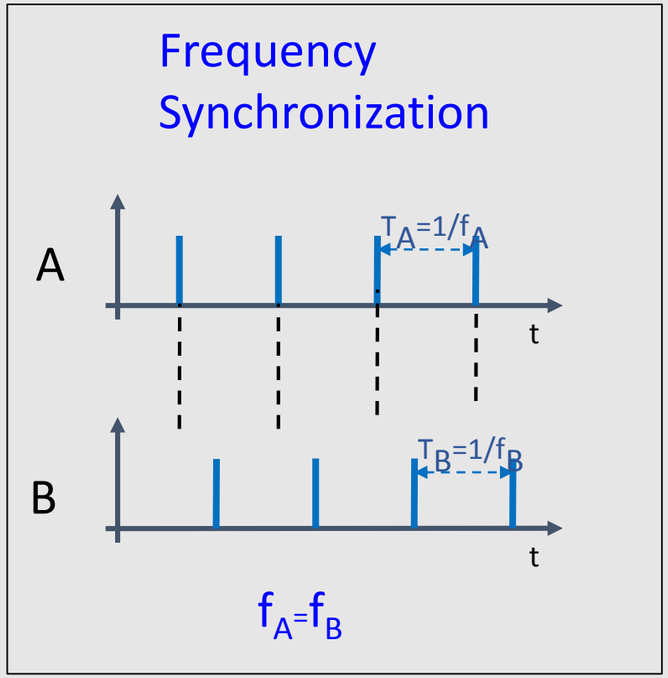
1 Introduction to Aggregate Ethernet(LAG) Networks

Introduction to AggregateEthernet(AE)/Link Aggregation Group(LAG)

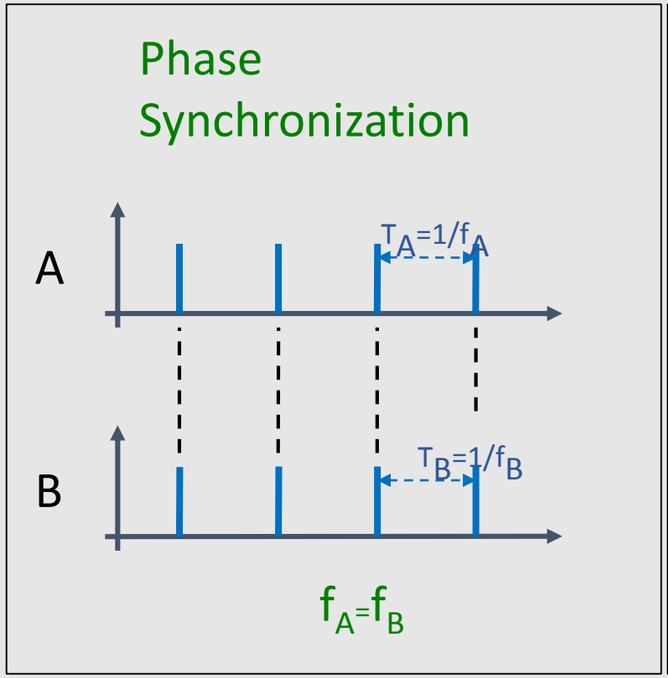
- IEEE 802.3ad link aggregation enables to group Ethernet interfaces to form a single, aggregated Ethernet interface, also known as a *link aggregation group (LAG)* or *bundle*.
- The aggregated Ethernet interfaces that participate in a LAG are called member links. Because a LAG is composed of multiple member links, even if one member link fails, the LAG continues to carry traffic over the remaining links
- Advantages:
  - Increased throughout beyond what a single connection could sustain.
  - Provides connection reliability/redundancy.
  - Can do load balancing among the various links in the Link Aggregation Group. So, the traffic is split evenly between the various links which ensures its effective utilization.

2 Synchronization Overview and Precision Timing Protocol

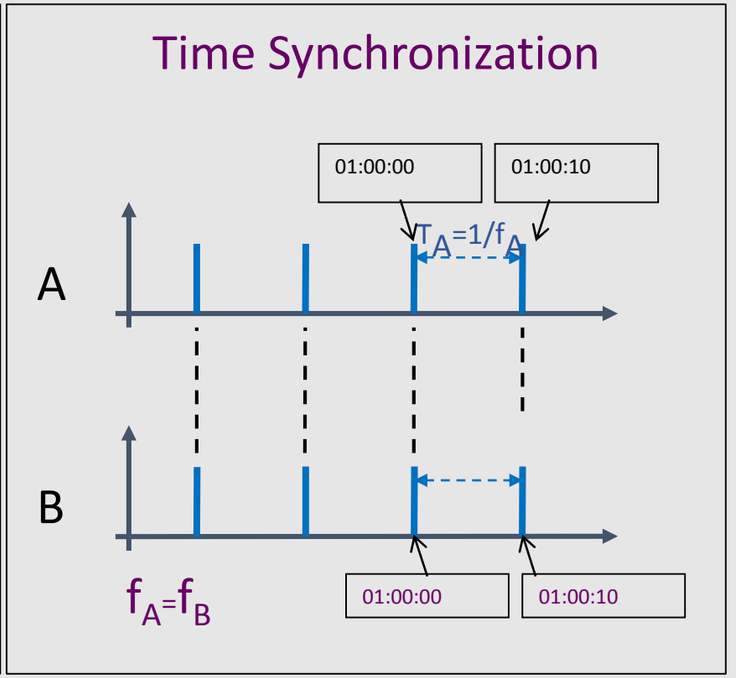
Synchronization Overview



Aligning clocks with respect to *frequency*



Aligning clocks with respect to *phase*



Aligning clocks with respect to *time.* The two clocks must utilize the same epoch. Time synchronization implicitly includes phase and frequency synchronization

IEEE-1588 (Precision Timing Protocol) Overview

Developed by HP/Agilent Laboratories (John Eidson and al.) in 1990s

- Synchronization of test equipment

Industrial Automation needs for high precision applications (e.g., robotics, test & measurements)

- Sub-microsecond synchronization over Controlled Local Area Networks
- Independent of the physical layer
- Transfer frequency and time of the day

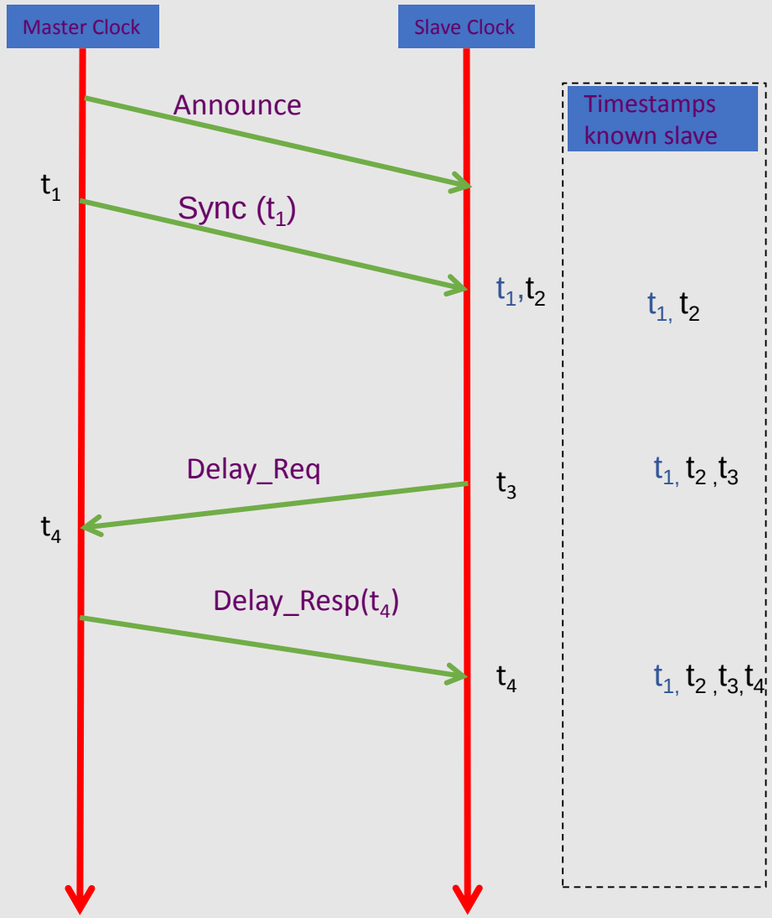
IEEE 1588 PTP Specifications

- 1588 PTP 2002 (also called version 1) - approved in 2002
- 1588 PTP 2008 (also called version 2) - approved in 2008

IEEE 1588 PTP 2008

- Packet based synchronization mechanism
  - UDP/IP layers messaging (multicast and unicast) over Ethernet
- Client/server model
  - Master clock, slave clock (ordinary clock)
  - Intermediary nodes may or may not support IEEE1588 PTP
  - On-path support mechanisms
    - Grandmaster
      - Acts as a master clock of the synchronization chain. May derive time from GNSS (GPS, GLONAS etc)
    - Boundary clock
      - Acts as a slave clock at port that connects to the grandmaster, and as a master to all other ports
      - It isolates the “down stream” clocks from any delays and jitter within the switch/routers
    - Transparent clock
      - It measures residence time of PTP event messages within a node.

PTPv2 protocol basics



PTP Accuracy depends:

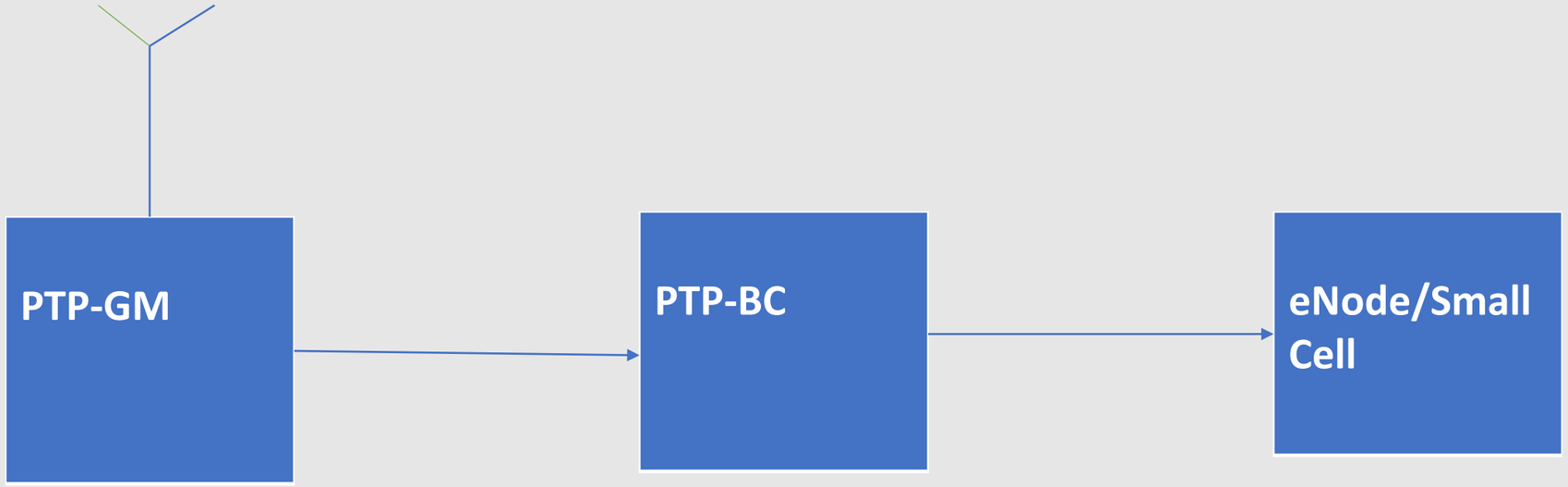
- Packet Delay Variation
- Data path asymmetry
- Time-stamping accuracy
- Oscillator stability

Offset =[(t2 - t1) - (t4 - t3)]÷ 2

3 Current industry challenges with time sync using PTP over LAG(AE) Networks

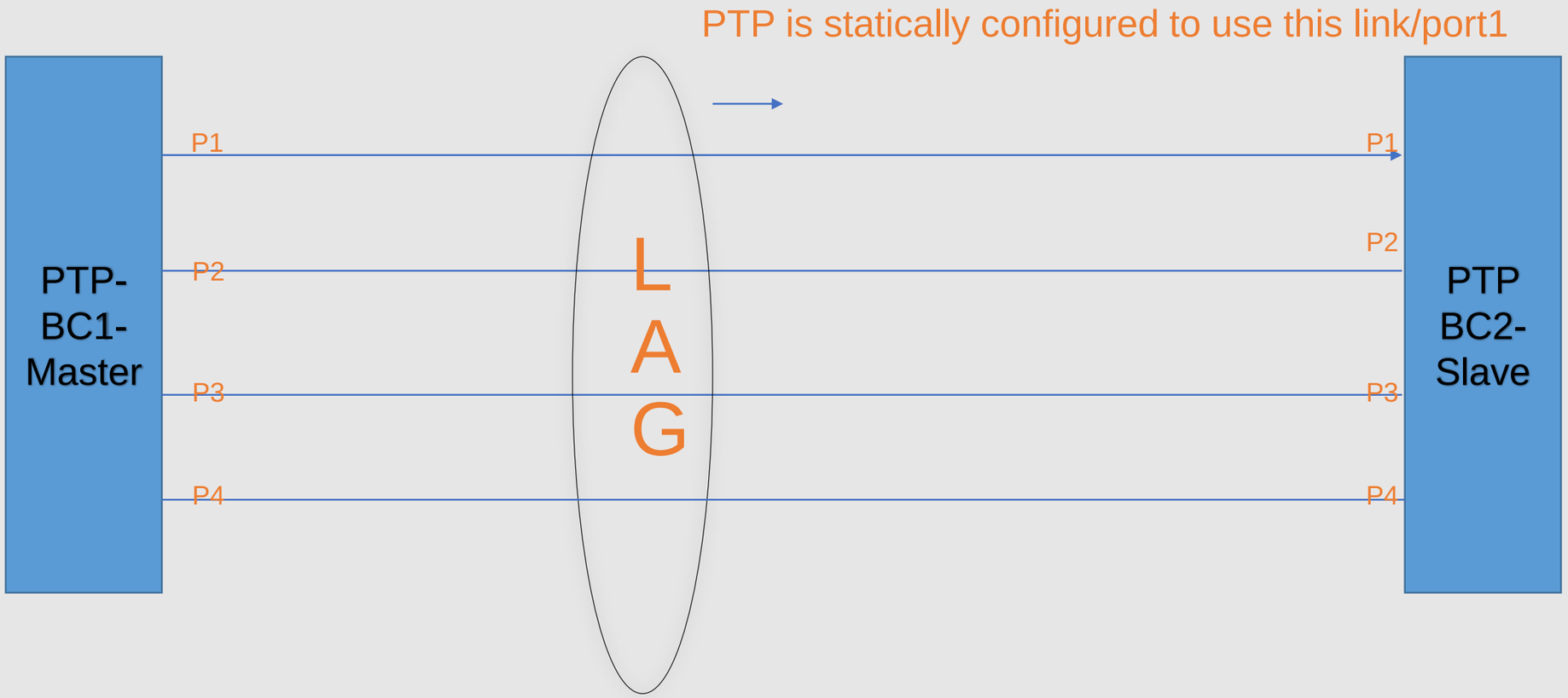
PTP over LAG networks challenges cont..

:PTP simple topology with Grandmaster , Boundary Clock and eNodeB/Small Cells interconnected without LAG



PTP over LAG topology

PTP boundary clock devices connected over LAG bundle with primary Link/Port(P1) carrying PTP traffic :



Todays challenge of PTP over LAG networks..

- PTP Switchover over member ports is not seamless :
  - Today with PTP/1588 over LAG with multiple member ports, if one of the bundle member carrying PTP traffic goes down, the traffic does not seamlessly switchover to any of the other available bundle member link/ports ensuring the Link/Port symmetries for the upstream and downstream PTP packets on the same switched Link/Port.
  - As PTP messages are lost due to port carrying PTP traffic is down, Slave will move out-of-phase/frequency locked state and impacting end-nodes needing phase and frequency synchronization.
- Issue of Asymmetricity :
  - Non symmetric path creates asymmetry in propagation delay for the PTP packets exchanged between two nodes can lead to error in phase offset calculation.

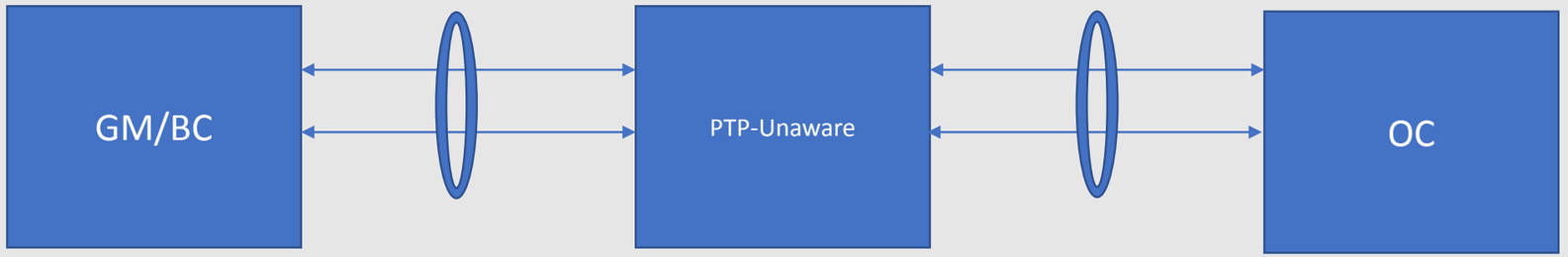
PTP over LAG challenges cont..

- Static configuration of Primary/Secondery on a LAG bundle:
  - One available solution today is to statically configure one of the bundle member as Primary which is carrying the PTP traffic and configure another bundle member Link/Port as secondary to switch over and carry the PTP traffic in the event of primary port/link going down. Basically a static configuration, instructing the PTP stack running on these GM/BC/OC nodes to send and receive packets on the given Link/Port as per configuration.
- But there are two problems associated with the existing solution.

PTP over LAG challenges cont..

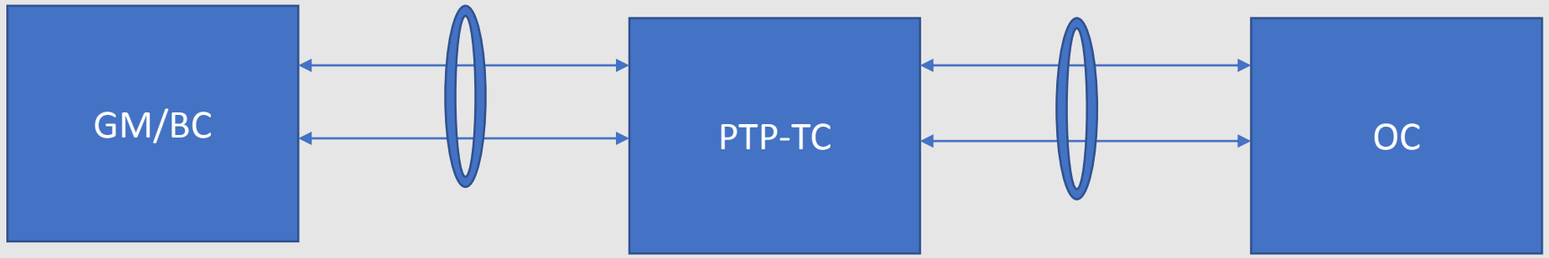
- What if Secondery link goes down:
  - If secondery link/port also goes down PTP connection is lost and PTP traffic cannot switchover to other available Link/Ports despite having the other LAG members connected.
- What if the intermediate nodes are PTP unaware:
  - The static configuration of Primary and secondary Port/Link/Ports will not be possible if the intermediate nodes are Transparent Clock devices/PTP unaware devices as depicted below:

Two PTP BC/OC nodes separated by an intermediate PTP unaware node.



PTP over LAG challenges cont..

- Two PTP BC/OC nodes separated by an intermediate node which can support only PTP TC (Transparent clock) as depicted below :



Proposed Solution:

- Proposal to advertise/learn PTP port priorities over Link Layer Discovery Protocol(LLDP):
  - The proposed solution is to do seamless PTP switchover on the LAG member Link/Port failures without having to bother about the static configuration and dynamically achieve symmetric Link/Port for PTP packet exchange between the two nodes over LAG bundle by learning the PTP port priorities over LLDP.

4 Solution

Proposed Solution: Logical flow and How it works

- Advertise/Learn PTP port priorities:
  - Proposal to add/advertise ptp-port-priorities configuration on any one node and let the other-remote-node learn the PTP port priorities over LLDP. LLDP is widely used Link/Port discovery protocol and supported by many vendors, there-by extending LLDP protocol with an option for PTP port priorities.
- LLDP protocol is extended to carry the ptp port priority on "each member Link/Port" of the LAG.
- How it works:
  - Upon configuring ptp-port-priority on the member-Link/Port, PTP-over-LAG configuration will trigger the LLDP protocol to include the corresponding configured ptp-port-priority.
  - Peer-node upon receiving the LLDP packet with PTP-PORT-PRIORITY TLV , it learns the ptp-port-priority advertised by peer for each of its member-link/port of the LAG.

Proposed Solution: Logical Flow and How it works cont..

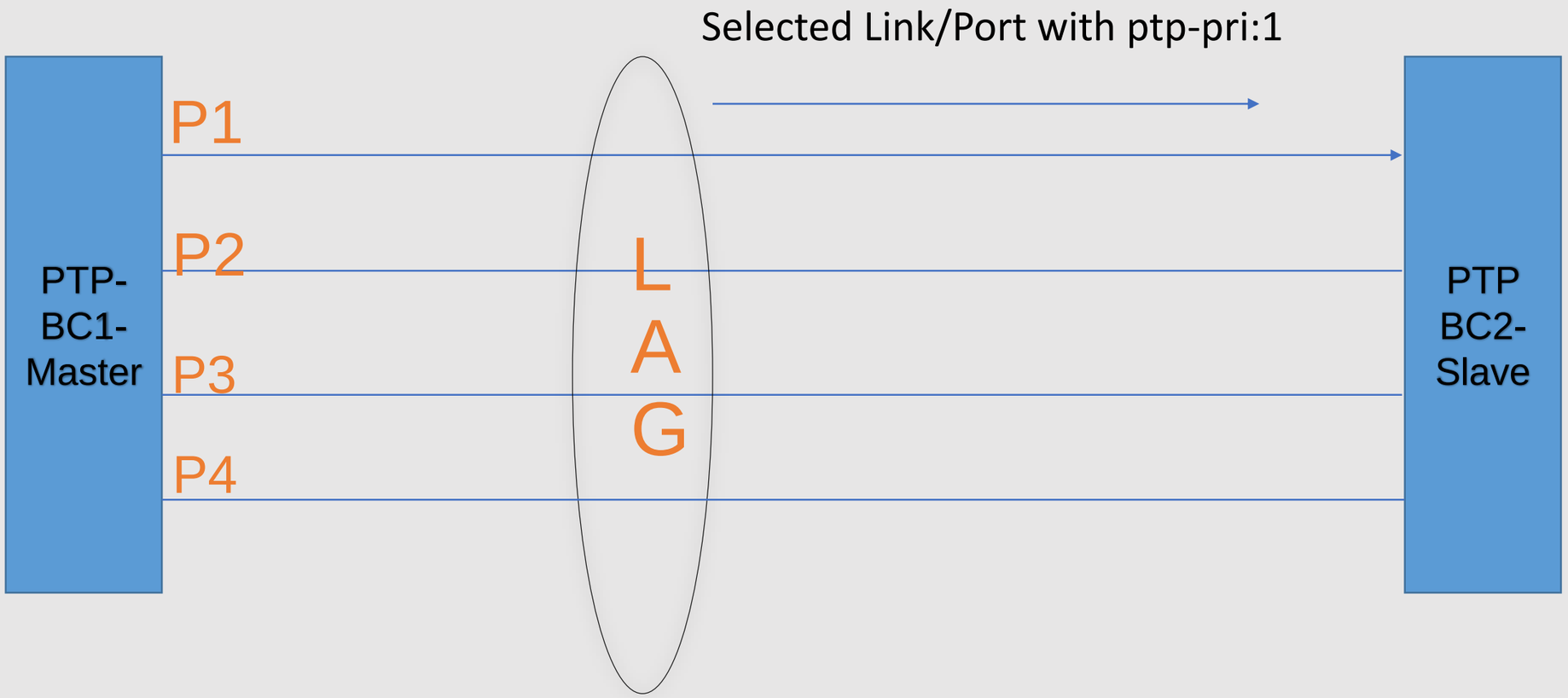
- Learn PTP port priority:
  - Whenever PTP packet is to be sent out of the LAG bundle, the forwarding module shall now choose the member-Link/Port of the LAG with **active-link/port** and with the **highest ptp-port-priority.**
- Simple PTP port priority selection algorithm:
  - A *simple priority selection algorithm can be priority-value-lower (1) implying highest priority* among the LAG members *(reserving 0 and highest value)* .
  - *If there are multiple member-link/ports with same ptp-priority, then choose the 1st active Link/Port.*
- PTP Port priority assignment:
  - The assignment of port-priority in the configuration can be such that the priority assignment order amongst the member links available can be based on the best cost/BW properties. When it is done this way – the switchover happening based on the priority will also fallback to next best available member links.

Proposed Solution: Logical flow cont..

- During a LOS or link/port-defect or fibre-cut on the member-link/port carrying ptp-traffic, the above logic will automatically ensure that the best same Link/Port is selected for ptp-traffic by both the nodes, thereby choosing the best active same Link/Port among the remaining Link/Port members.
- Extensions to LLDP:
  - Proposal to extend LLDP header with new type for ptp priority information propagation.

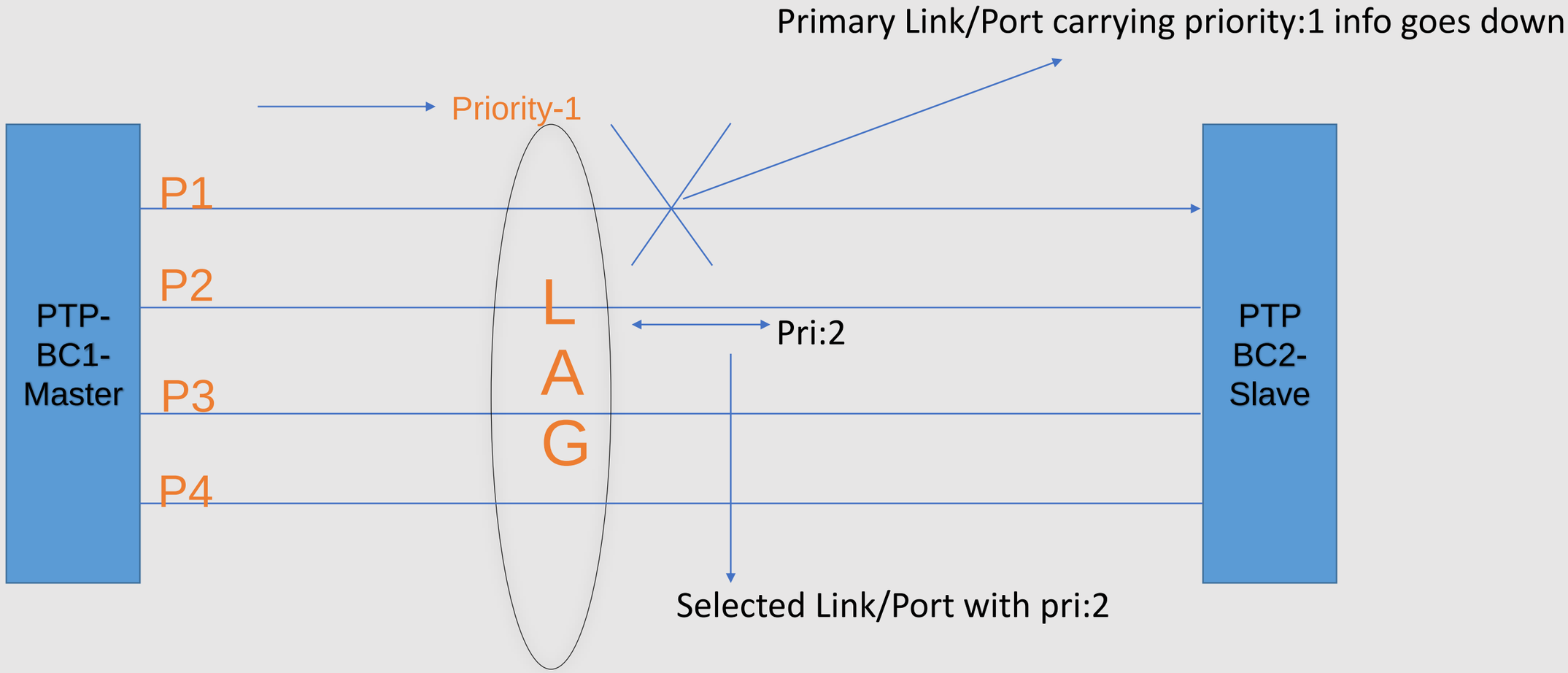
Solution cont..

Topology depicting the logical flow of the proposed solution during the normal operation of primary link/port carrying PTP traffic:



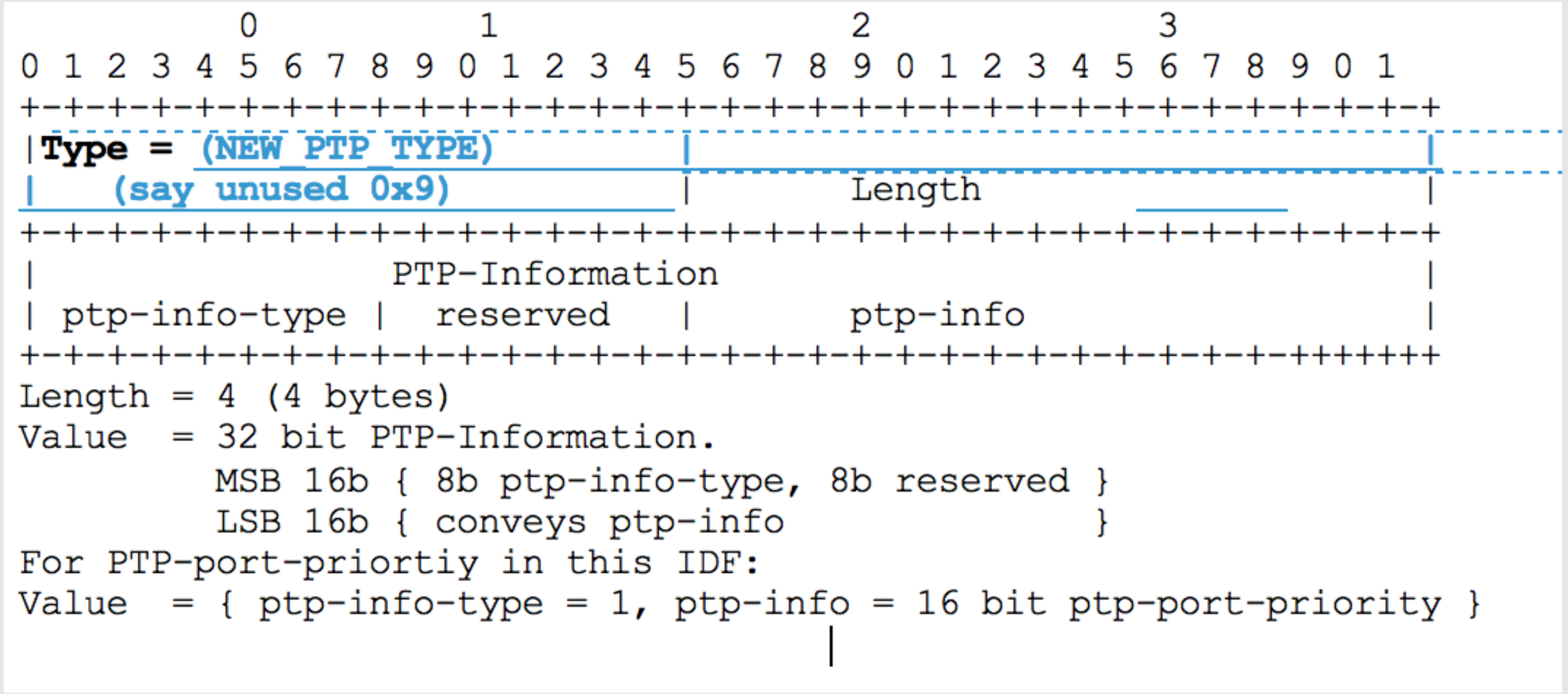
Solution cont..

Topology depicting proposed logic flow of the events during the switchover of ptp traffic from one link to other :



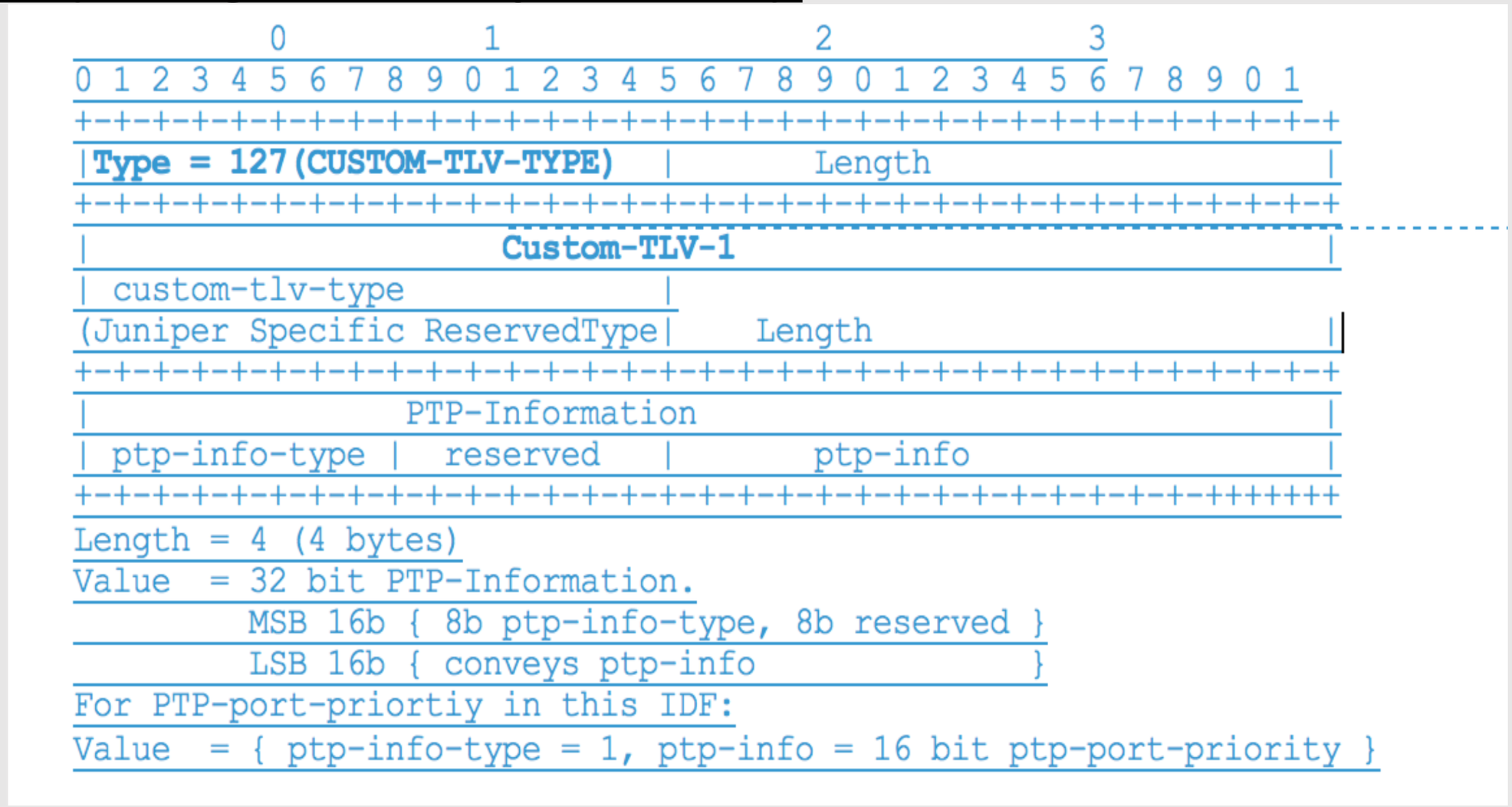
Proposed changes to LLDP header

NEW LLDP TLV for carrying-ptp-port-priority



Proposed changes to LLDP header cont

- **Proposed NEW LLDP TLV for carrying-ptp-port-priority – using Custom-TLV(or Organization-Specifiv-TLV):**



Solution cont..

- The other method to dynamically derive symmetric (same) Link/Port for PTP packet exchange between two nodes and to do graceful switchover of Link/Ports both sides of the LAG-bundle whether PTP aware or PTP TC node at one end of the lag-bundle is to use optional TLVs of LLDP protocol as shown below:

LLDP Ethernet frame structure

Preamble	Destination MAC	Source MAC	Ethertype	Chassis ID TLV	Port ID TLV	Time to live TLV	Optional TLVs	End of LLDPDU TLV	Frame check sequence
	01:80:c2:00:00:0e, or 01:80:c2:00:00:03, or 01:80:c2:00:00:00	Station's address	0x88CC	Type=1	Type=2	Type=3	Zero or more complete TLVs	Type=0, Length=0	

5 Summary

Summary

- The proposed solution claims both models to achieve the required functionality of symmetric PTP packet exchange by dynamically deriving the member Link/Ports of the LAG bundle at both end nodes and to do graceful switchover as against the present day static method of configuring primary and secondery members.
- Proposed method/idea submitted for patent filing with US PTO and others.

Thank you

Acknowledgements:
Kamatchi Gopalakrishnan
Prabhu Raj VK