Introduction to IEEE 802.1Qca

Path Control and Reservation

Authors: **Hesham ElBakoury**

Version: 1.0

HUAWEI TECHNOLOGIES CO., LTD.

HUAWEI

GOAL

Present the scope main concepts of the new IEEE 802.1Qca standard which is being developed by IEEE 802.1 to decide whether IEEE 1904.2 is interested to pursue using this new standard for its discovery protocol given its added complexity.

Agenda

- **What is the new IEEE 802.1Qca standard ?**
- **Introduction to IS-IS**
- **Overview of Shortest Path Bridging (SPB) – IEEE 802.1aq**
  - Spanning Tree vs SPB
- **Overview of Path Computation Element**
- **IEEE 802.1Qca Explicit Trees**
  - New IS-IS Topology sub-TLV of LSP.

What is IEEE 802.1Qca ?

- It is a new standard that will provide explicit path control, bandwidth and stream reservation, redundancy (protection or restoration) for data flows.
- The new standard will extend the applications of IS-IS to control bridged networks beyond the capabilities of Shortest Path Bridging (IEEE 802.1aq) by using IS-IS to carry control information for time synchronization and scheduling.
- The new control protocol will provide explicit forwarding path control thus enabling the use of non-shortest path
- IEEE 802.1Qca uses Path Control Element (PCE) that is defined by IETF PCE WG.
- Path Control and Reservation is specified in Clause 45 of IEEE 802.1Q.
- IEEE 802.1Qca will be an amendment to the ongoing Revision of 802.1Q

Status of IEEE 802.1Qca

- **Draft D1.1 was published on Sept 26th 2014.**
- **We are in WG ballot on D1.1 until Oct 26th, unofficial results can be seen here:**
- **http://standards.ericsson.com/ethernet/web_tally/**
- **The plan is to go for a recirc after the next meeting and then sponsor ballot by March.**
  - It could then go to RevCom in July 2015 and get published in late 2015.

1 IS-IS

Why IS-IS ?

- IS-IS : Intermediate System to Intermediate System routing exchange protocol (the ISO protocol for routing within a single routing domain).
- IS-ISis a Link State Protocol similar to the Open Shortest Path First (OSPF)
- Embraced by the large tier1 ISPs.
- Proven to be a very stable and scalable, with very fast convergence.
- Encodes the packet(s) in TLV format.
- Flexible protocol in terms of tuning and easily extensible to new features (MPLS-TE etc).
- It runs directly over Layer 2. (next to IP).

IS-IS MAC Header and Addresses

AllL1ISs	01-80-C2-00-00-14
AllL2ISs	01-80-C2-00-00-15
AllIntermediateSystems	09-00-2B-00-00-05
AllEndSystems	09-00-2B-00-00-04

LSP Generation: What triggers a new LSP

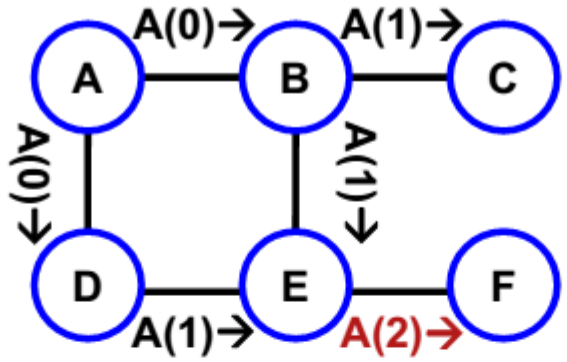
- **When something changes …**
  - adjacency came up or went down
  - interface up/down (connected IP prefix !)
  - redistributed IP routes change
  - inter-area IP routes change
  - an interface is assigned a new metric
  - most other configuration changes
  - periodic refresh

LSP Generation: New LSP

- Create new LSP, install in your own LSPDB and mark it for flooding
- Send the new LSP to all neighbors
- Neighbors flood the LSP further.

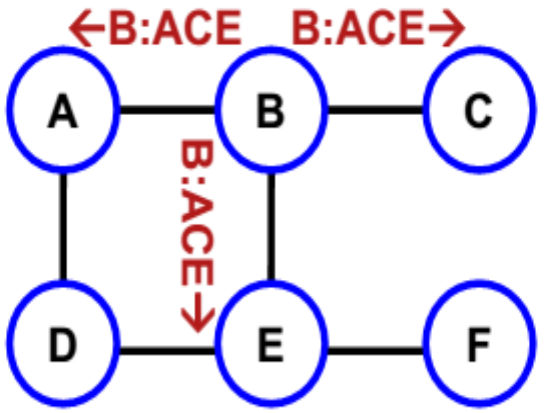
2 Shortest Path Bridging (SPB)

Spanning Tree

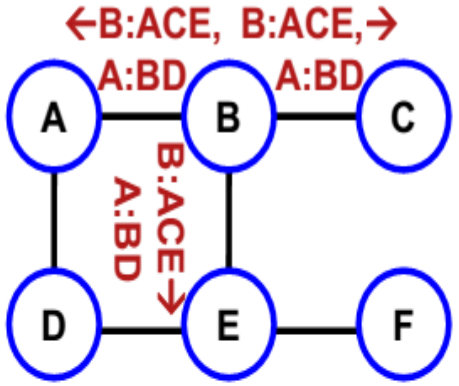


- Let us suppose **A** is the root of the spanning tree.
- **A** advertises “I am the root.”
- **B** and **D** advertise “I am one hop from the root.”
- **E** advertises “I am two hops from the root.”
- **F knows nothing of the path** to **A**; only the distance.

Shortest Path Bridging (1)



- **A** advertises, "I am **A**. I connect to **B** and **D**."
- **B** advertises, "I am **B**. I connect to **A**, **C** and **E**."

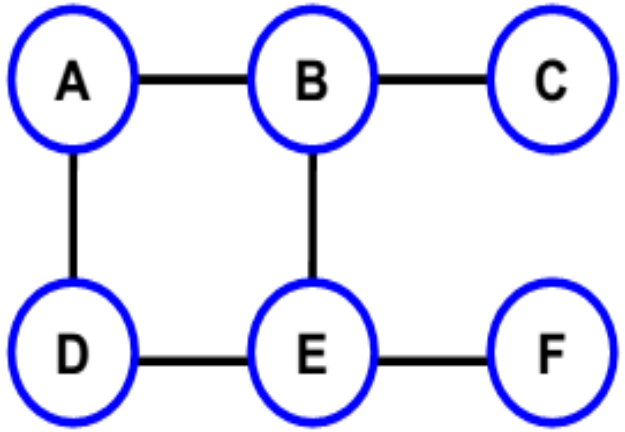


- Furthermore, as soon as B hears A's advertisement, it relays A's information to its neighbors.
- That is, **B** says, "I am **B**. I connect to **A**, **C**, and **E**. **A** says that it connects to **B** and **D**."

✓IS-IS is used to advertise Topology Information.
✓RFC 6329 describes IS-IS Extensions Supporting IEEE 802.1aq Shortest Path Bridging

Shortest Path Bridging (2)

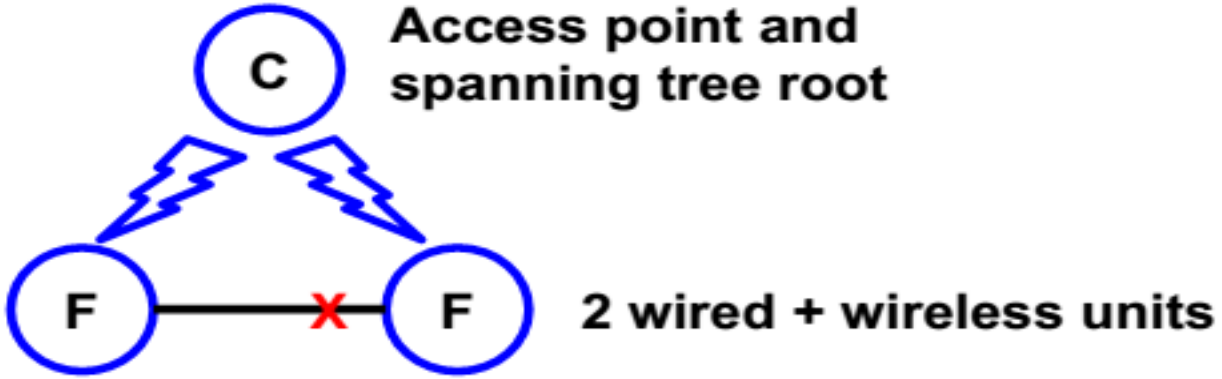
A:BD
B:ACE
C:B
D:AE
E:BDF
F:E



- Eventually, every switch in the network has the state of every other switch, and advertises all of that information to its neighbors.

Spanning Tree vs Shortest Path Bridging

- **Spanning Tree** is a **simpler** protocol and has smaller computational and memory requirements. But, it can route traffic along **inefficient paths**.
- **Shortest Path Bridging** is **more complex** and requires more memory and CPU cycles. But, it routes all traffic along the **shortest path**.



3 Path Computation Element (PCE)

Path Computation Problems

- **There are lots of path computation problems in networking**
  - Constraint-based SPF.
  - Multi-path Problems (Classical Fish Problem in Routing)
  - Point to Multipoint problems
- **More sophisticated algorithms may be needed to solve these problems.**
  - k-shortest paths
  - Linear programming
- **Many problems take considerable computation resources.**
- **Using service planning many problems can be solved off-line**
- **Increasing demand for on-line or rapid response**
  - Network nodes (routers and switches) do not have enough resources (CPU, and memory)
- **All of these problems can be addressed by using dedicated path computation servers**

PCE IETF Working Group

- **History**
  - https://datatracker.ietf.org/wg/pce/history/
  - Group was proposed on 2004-03-01 and started on 2005-01-06 (8 years, old group).
- **Charter**
  - Define PCE Architecture.
  - The PCE Working Group is chartered to specify the required protocols so as to enable a Path Computation Element (PCE)-based architecture for the computation of paths for MPLS and GMPLS Point to Point and Point to Multi-point Traffic Engineered LSPs
  - The WG determines requirements for extensions to existing routing and signaling protocols in support of the PCE architecture and the signaling of inter-domain paths (e.g. RSVP-TE and its GMPLS variations).
    - **Any necessary extensions will be produced in collaboration with the Working Groups responsible for the protocols**
  - The WG defines the required PCEP extensions for **Wavelength Switched Optical Networks** (WSON) while keeping consistency with GMPLS architecture specified in CCAMP WG.

What is PCE (Path Computation Element) ?

- ***PCE: Path Computation Element.*** **An entity that is capable of computing a path through a network based on a representation of the network's topology (obtained by undefined means external to the PCE). A PCE is a higher layer entity in a bridge or an end station.**
  - **It is defined in RFC 4655**
  - PCE can be in a dedicated computation server or in a central provisioning server (NMS)
  - It can be embedded in a router/bridge or every router/bridge
    - The PCE function can be distributed into the routers/bridges.
    - There may be multiple PCEs with different capabilities in different parts of the network

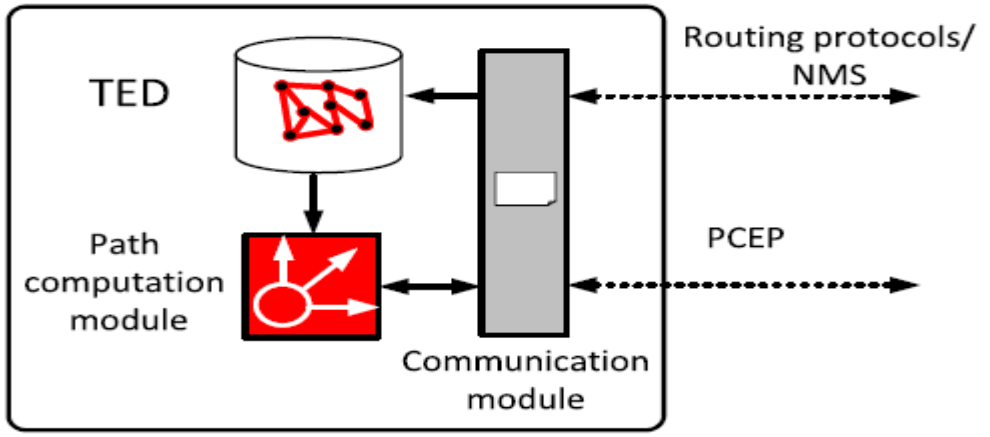
Key Reasons to Deploy PCEs

- **Dynamic and Optimal End-to-End Path Computation.**
  - That is not possible with the current distributed control plane architecture as no node has access to the full network topology.
- **Secure and Private information sharing**
  - The PCE protocol communications between PCE Servers may use a secured channel.
  - The PCE Servers can provide path details in the form of an encrypted key, thereby keeping internal domain topology private.
- **Customizable path computation**
  - PCE Server products provide open software APIs, to allow operators to customize or replace routing algorithms.
- **Improved price/performance ratio**
  - Deploying/upgrading PCE servers is much lower cost than upgrading all NEs in the network
  - Increase life span of the NE
  - PCE servers can be hosted in the cloud in virtual machines reduces CAPEX and OPEX.
- **Simplified operations for upgrading path computation**

PCE Functional Modules

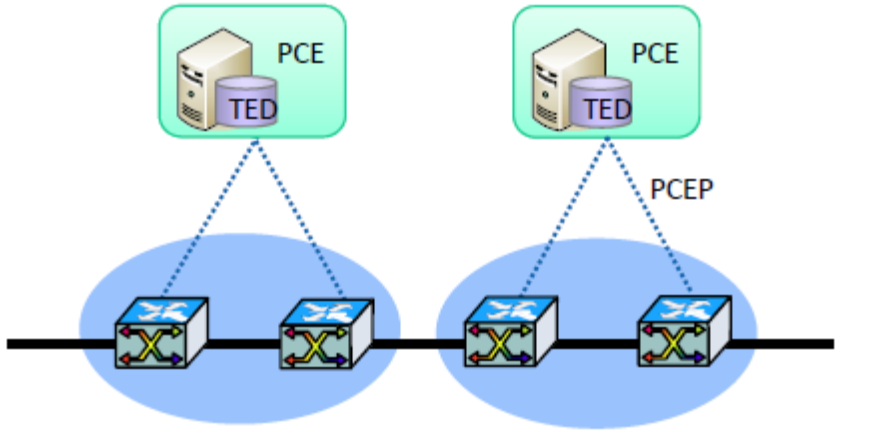
- **Traffic Engineering Data Base (TED)**

  The TED contains the topology and the resource information of the IS-IS domain

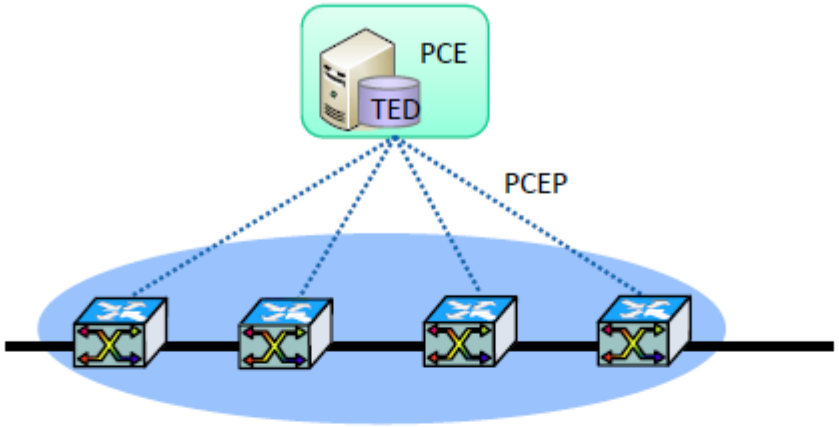


- **Path Computation Module**
  - The algorithm the PCE uses for path computation is not specified by 802.1Qca
- **Communication Module)**
  - PCE Communication Protocol (not defined by IEEE 802.1Qca)
- **A PCE may only include the computation engine.**
  - It may be remote and accessed by PCE Communication Protocol
  - it may be built into some other network component (e.g. Bridge).
- **PCE can be a network component that includes path computation, and the TED.**
  - Can do a number of other functions like configuration and provisioning.

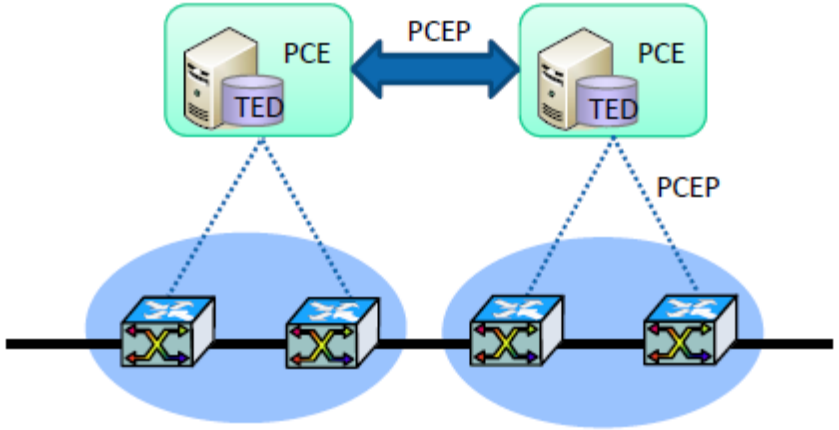
Centralized vs Distributed Path Computation



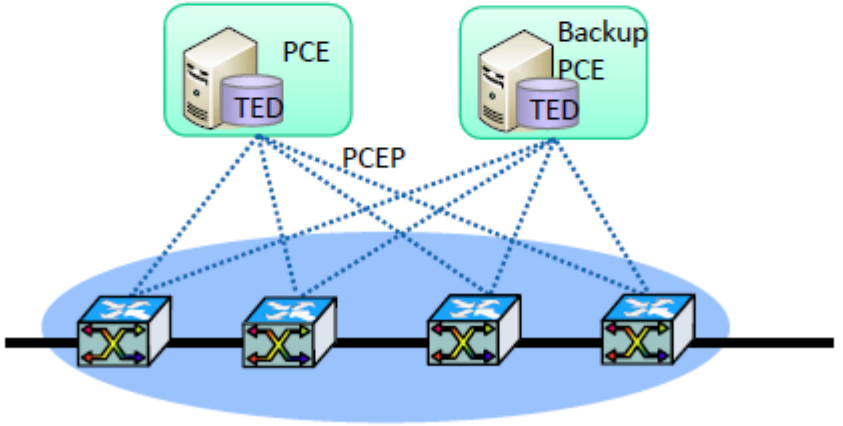
Distributed path computation without collaboration



Centralized Path Computation



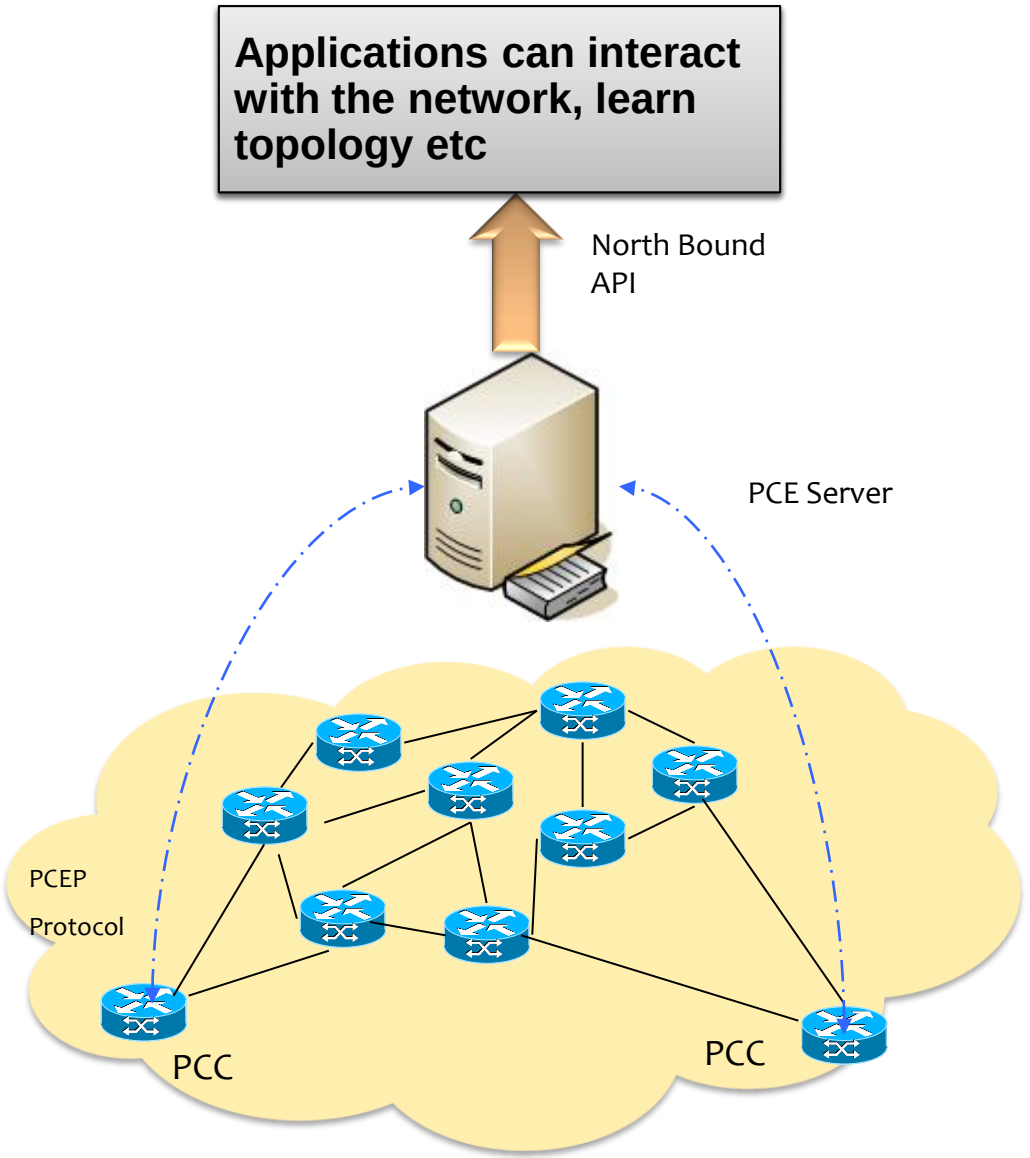
Distributed path computation with collaboration



Centralized Path Computation with backup

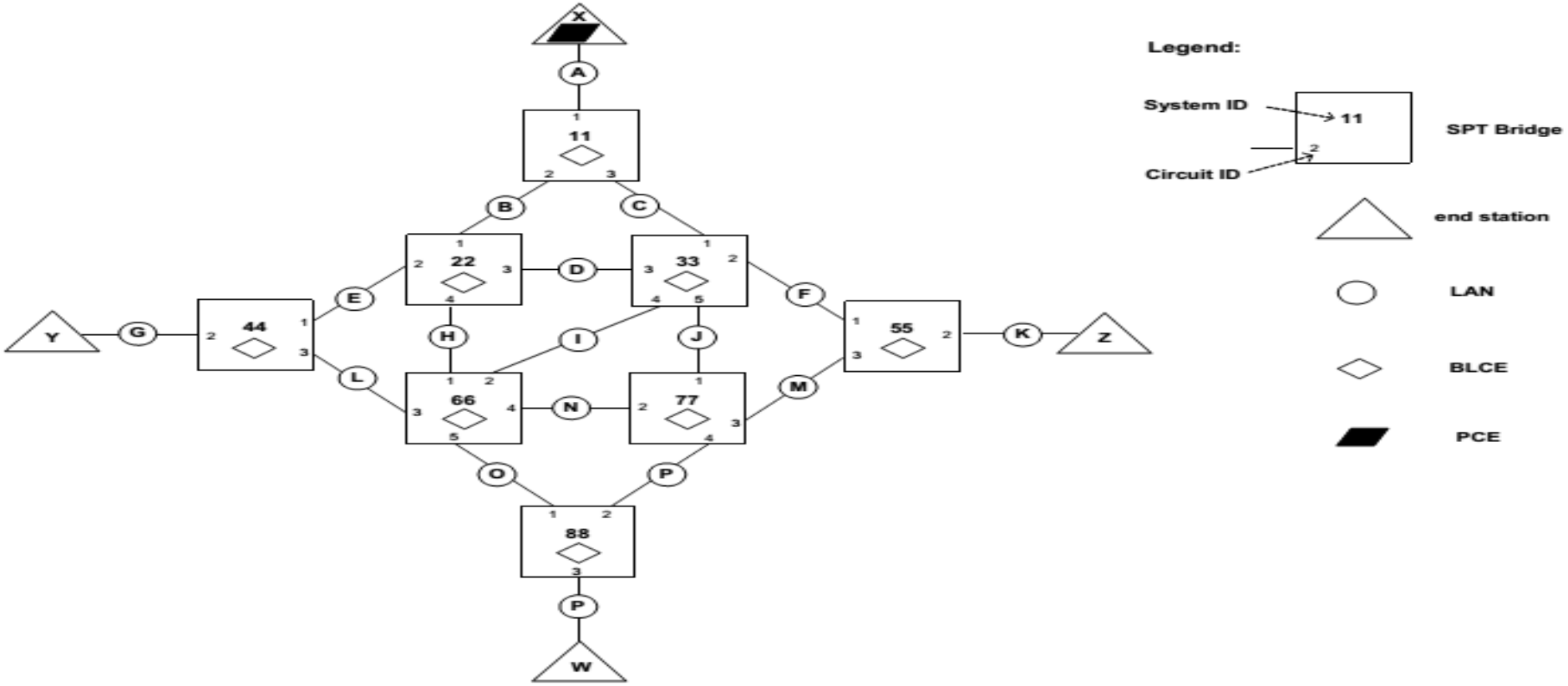
Example of Network Service – Path Computation Element

- **Applications such as VoIP, video and collaboration depend upon strict QoS compliance. Increasing use of these applications, along with the continuing need to minimize costs, adds incentives that seem sure to lead to widespread PCE adoption.**
- **PCE can provide northbound RESTful Open API to provide network information (e.g. topology) to applications and services.**
  - This API can be used by applications to request network path with certain bandwidth and QoS requirements.



4 IEEE 802.1Qca Explicit Trees

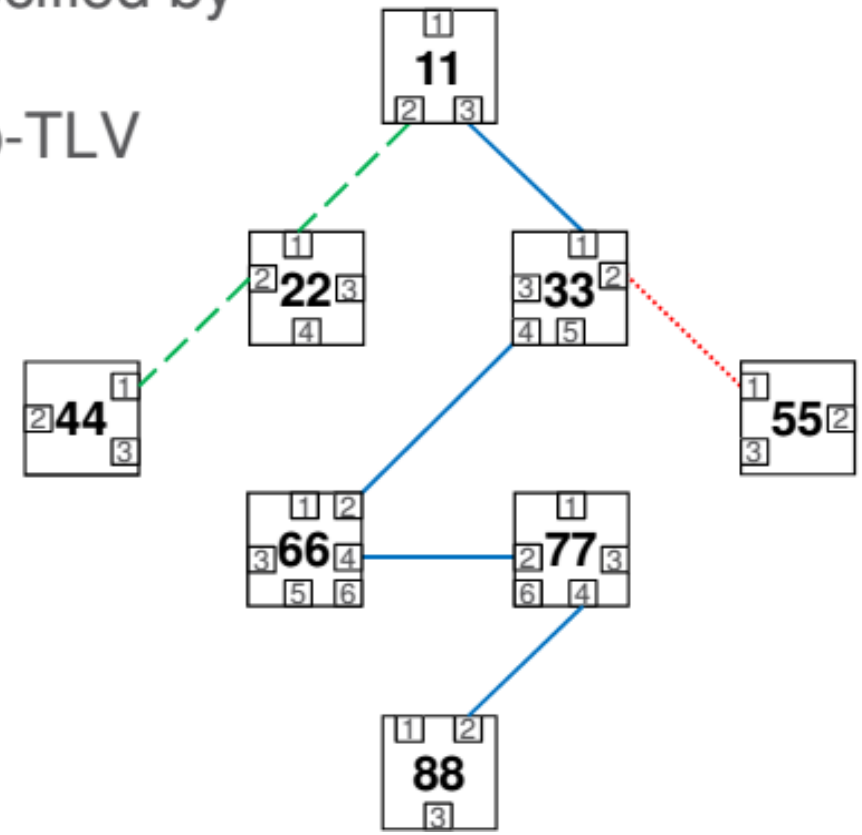
PCE Explicit Trees



- This Figure shows a region controlled by a single PCE residing in end station X connected to SPT Bridge 11. The PCE has IS-IS adjacency established with SPT Bridge 11.
- PCE Computes ET which is installed using IS-IS extensions.
- Stateful PCE can compute ET based on bandwidth requirements.
- Path Congruency algorithms to support bidirectional paths.

Explicit Strict Spanning Tree Topology

- Each hop of a strict explicit tree is exactly specified by its descriptor Topology sub-TLV

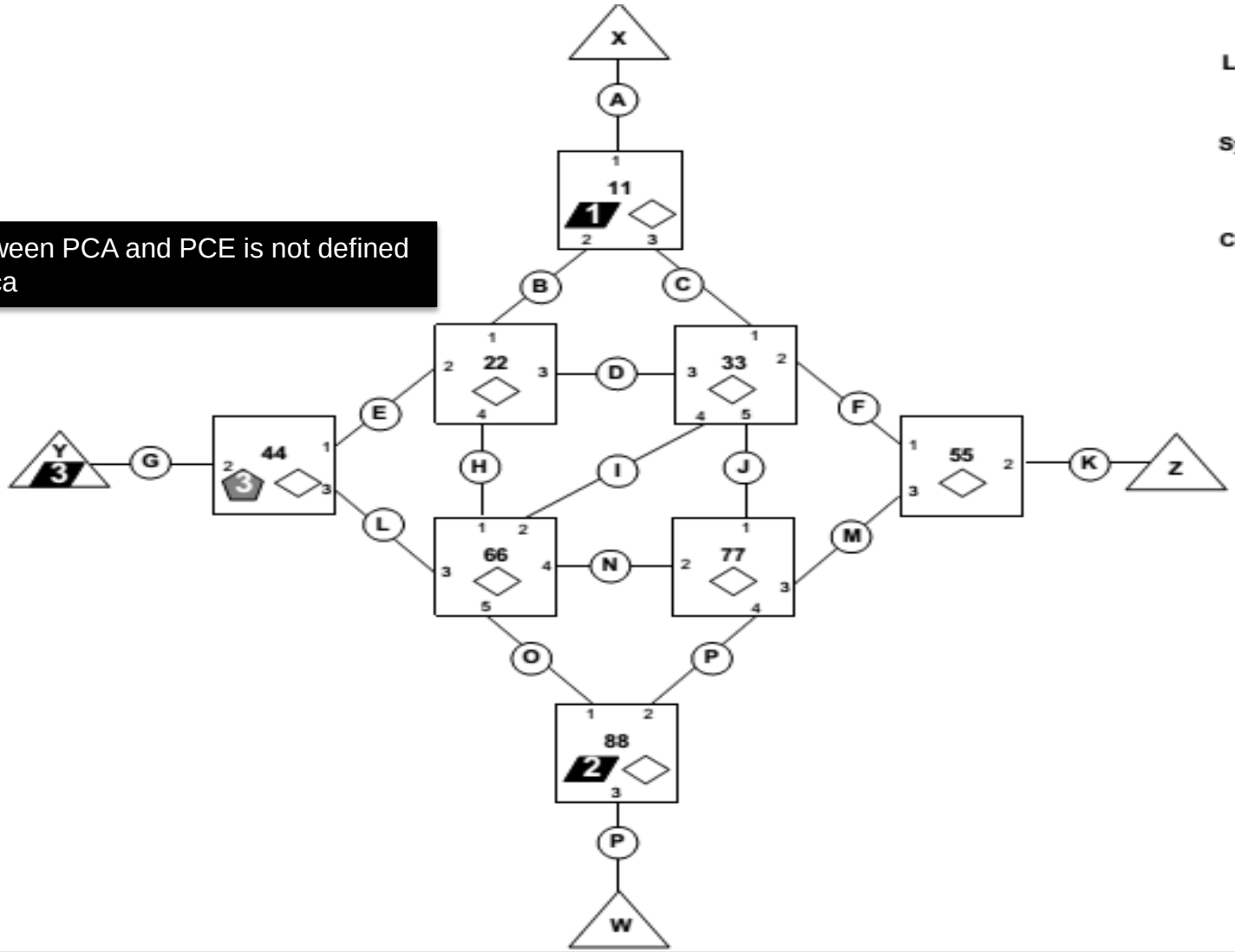


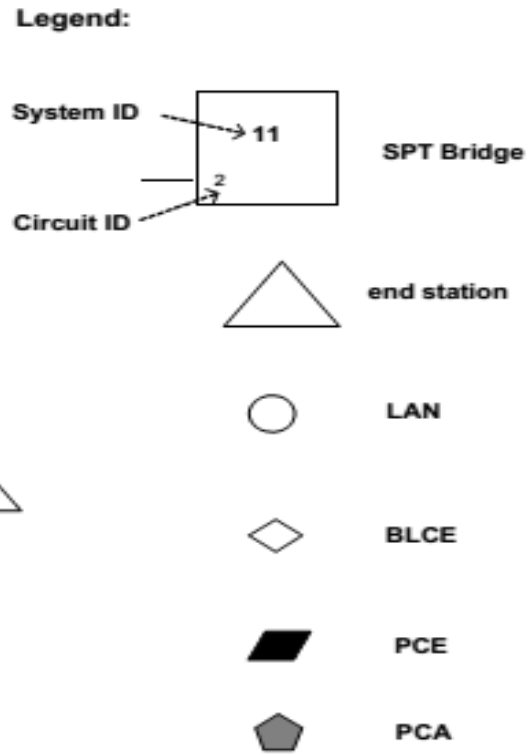
Descriptor

11; *Root, TEP*	Branch 1
33	
66, 4; *Circuit*	
77	
88; *Leaf, TEP*	
11; *Root, TEP*	Branch 2
22	
44; *Leaf, TEP*	
33	Branch 3
55; *Leaf, TEP*	

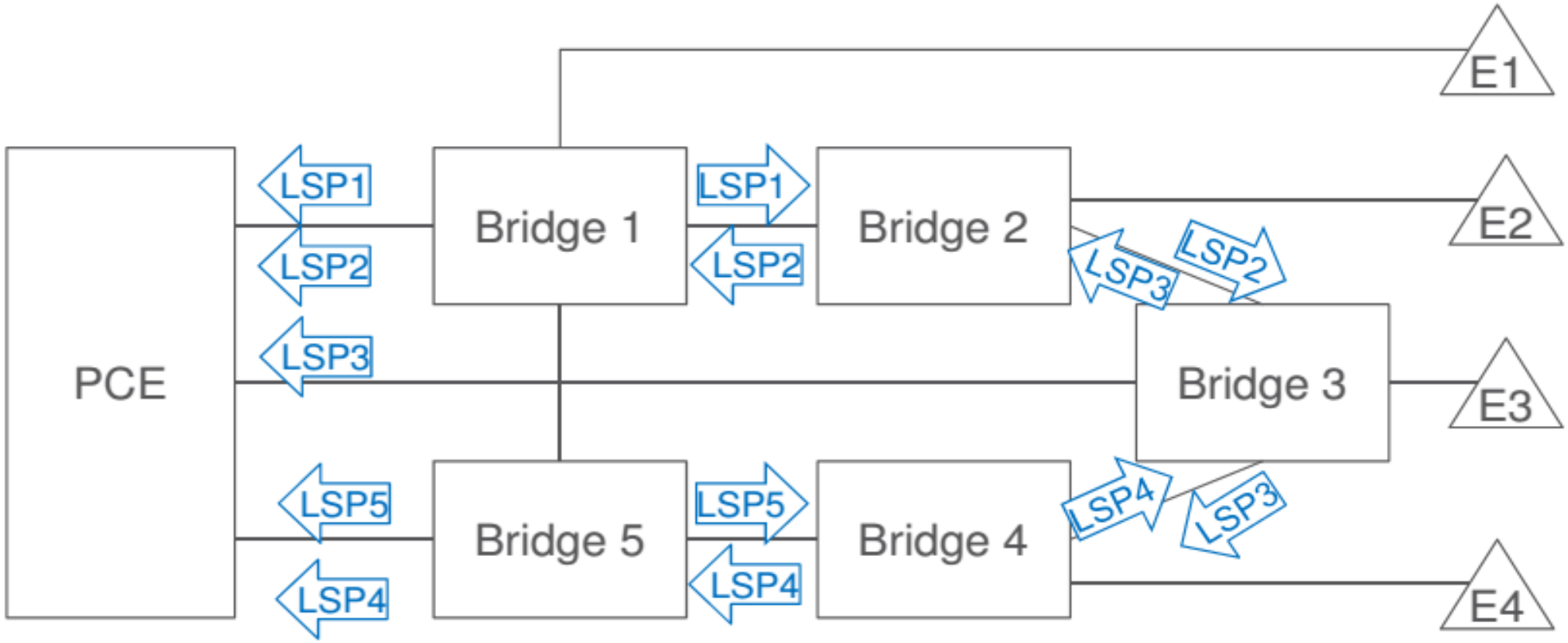
Using Multiple PCEs

The protocol between PCA and PCE is not defined by IEEE 802.1Qca



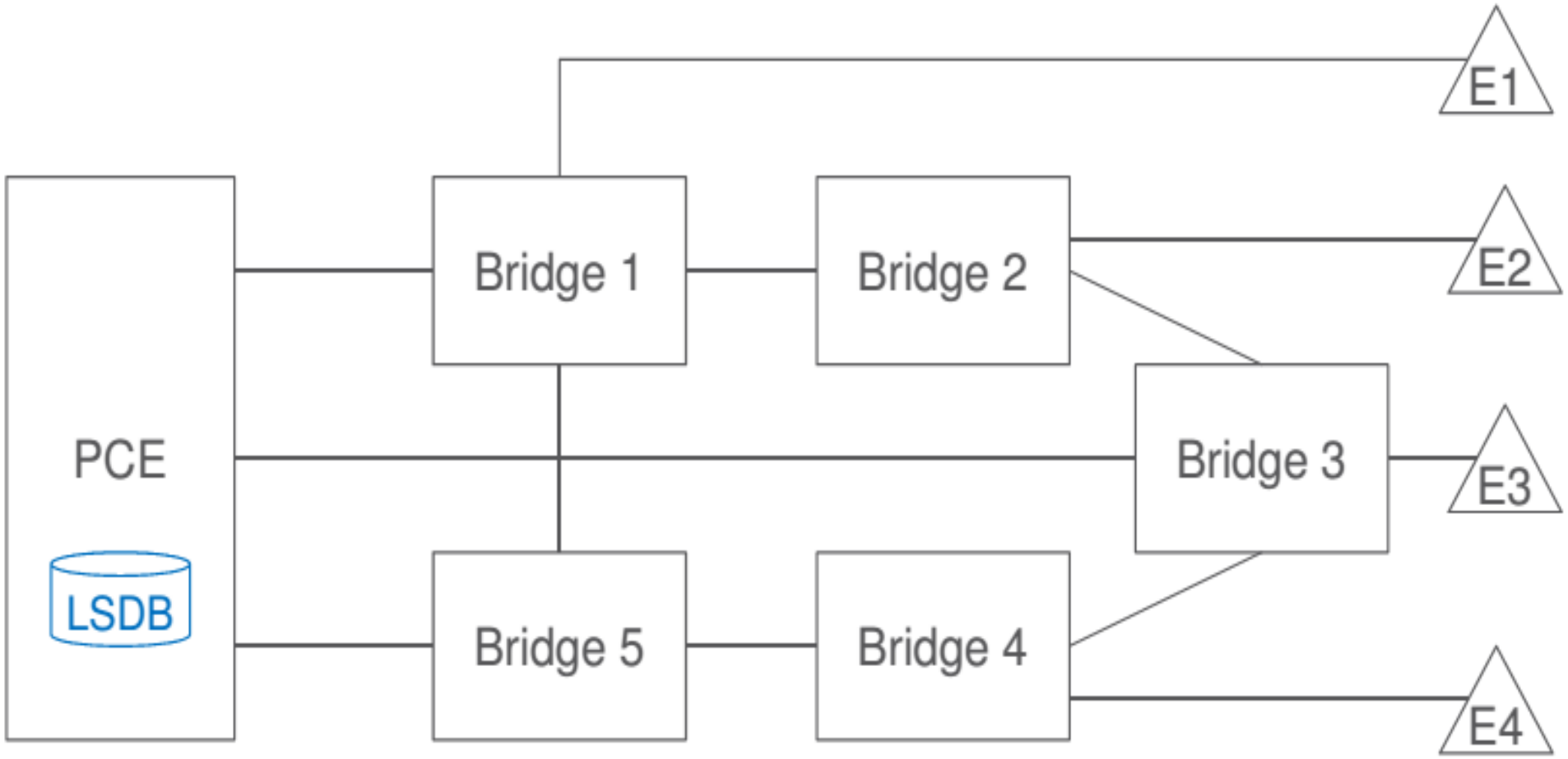


Conveying End Station MAC Addresses usingl LSP Flood

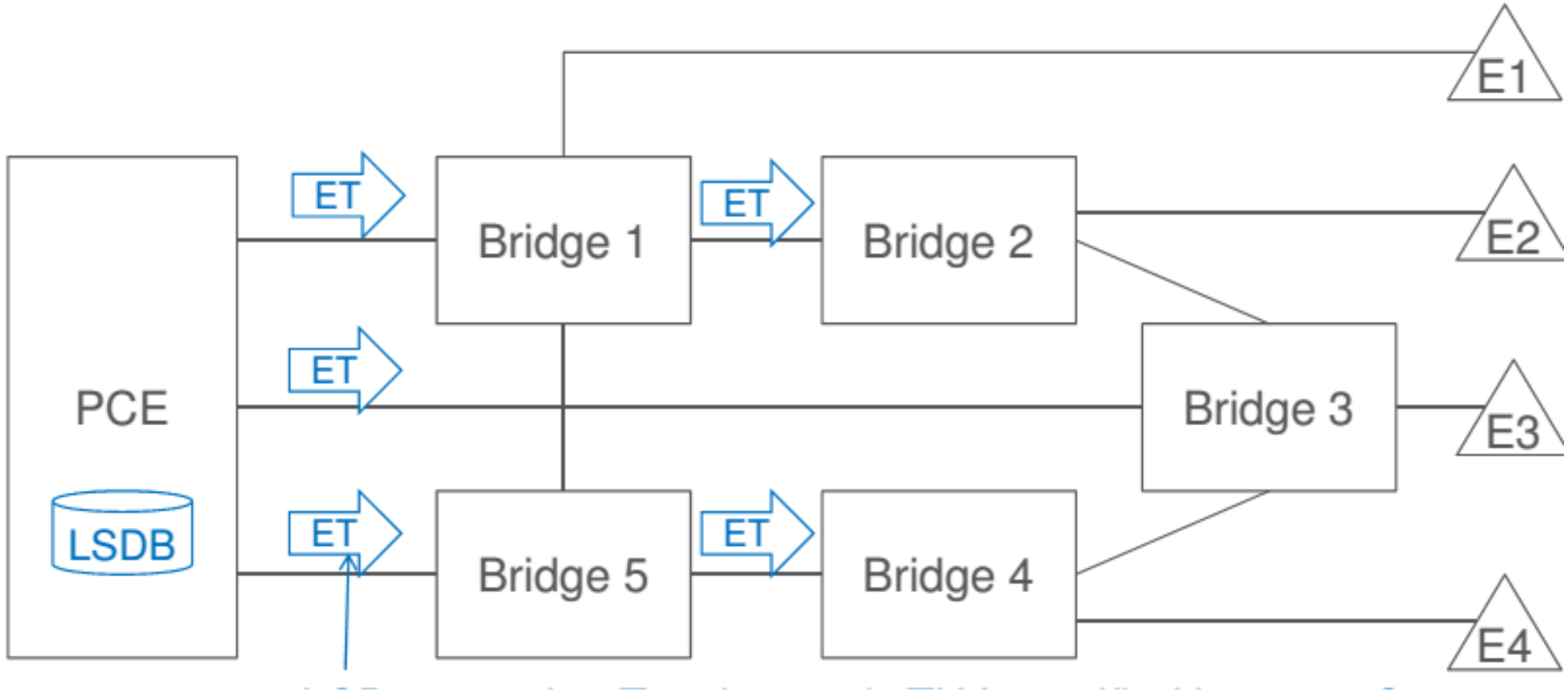


- End Station MAC is conveyed by LSP
- Not all LSPs are shown in the figure

Building LSDB



Flooding Explicit Static Trees Using PCE



The trees associated with the ST ECT Algorithm are static in the sense that no other entity including IS-IS can update them but the tree owner PCE.

Thank you

www.huawei.com