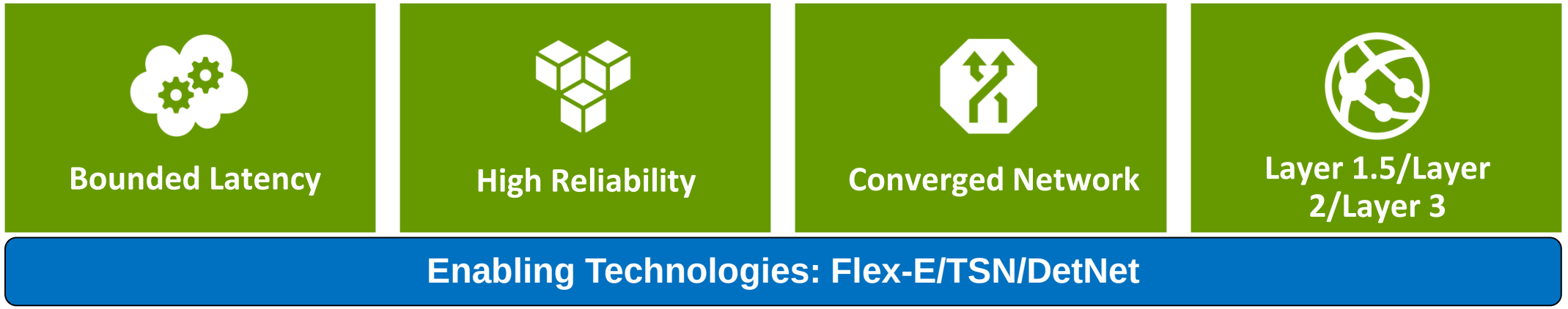
Advances in Deterministic Networking

Andrew G. Malis
Distinguished Engineer
Huawei Technologies

MPLS+SDN+NFVWORLD PARIS 2019

Deterministic Networking

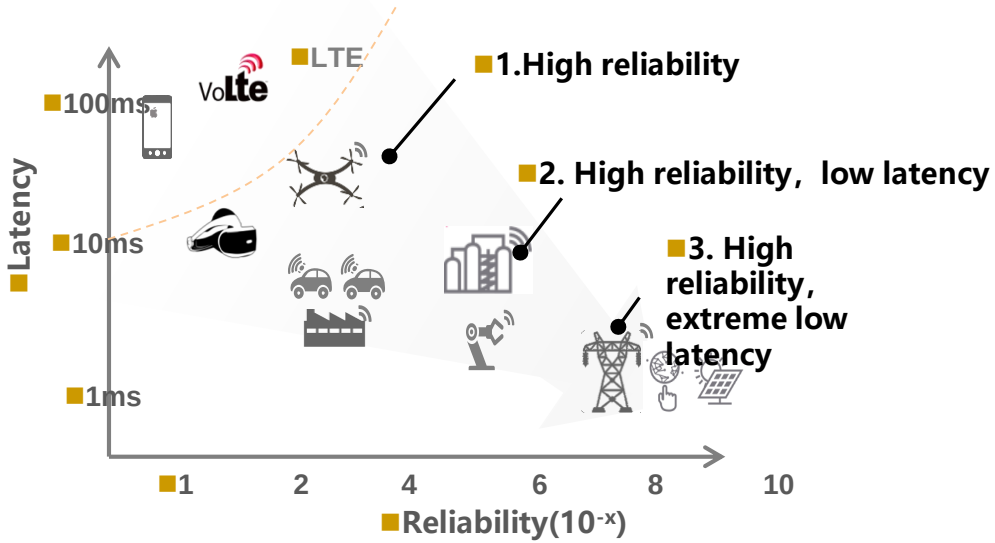
- **Deterministic Networking has the following features :**
  - ✓ Can provide high-quality service delivery with no congestion;
  - ✓ Can guarantee bounded latency, low jitter and high reliability;
  - ✓ Is supplied by a network that is primarily an IP/Ethernet network;



- **TSN**: Time Sensitive Networking (IEEE) , designed for layer 2; http://www.ieee802.org/1/pages/tsn.html
- **DetNet**: Deterministic Networking (IETF), designed for layer 3; https://tools.ietf.org/wg/detnet/
- **Flex-E**: Flexible Ethernet (OIF), designed for layer 1.5; https://www.oiforum.com/technical-work/hot-topics/flex-ethernet-flexe-2/

Use Cases for Deterministic Networking

- Deterministic Networking has the following features :
  - 5G Mobile
    - ✓ Network Slicing
    - ✓ Ultra-Reliable Low Latency Communications (URLLC)
    - ✓ Fronthaul and backhaul networks
  - Remote healthcare
  - Autonomous vehicles
  - Factory automation
  - Smart grid
  - Professional audio and video
  - Mining industries
  - Network-based augmented and virtual reality



DetNet WG in IETF

- Designed to extend current IP and MPLS networking technology to support Deterministic Networking
  - Derived from the concepts as IEEE 802.1 Time Sensitive Networking for Ethernet
    - A converged network in support of both critical service (deterministic) and best-effort service
    - Bounded packet latency and jitter (packet delay variation)
    - Extremely low packet loss (duplicate transmission option) and in-order packet delivery
  - Needed at layer 3 because many use cases cannot (only) use a layer 2 network
    - Scaling, distance, infrastructure cost
    - Designed to interwork with IEEE 802.1 TSN
  - Addresses layer 3 technology gaps (e.g., data plane, control plane, OAM, and security aspects)
    - Reuse as much as possible existing IETF forwarding functionality (IP, MPLS, QoS)

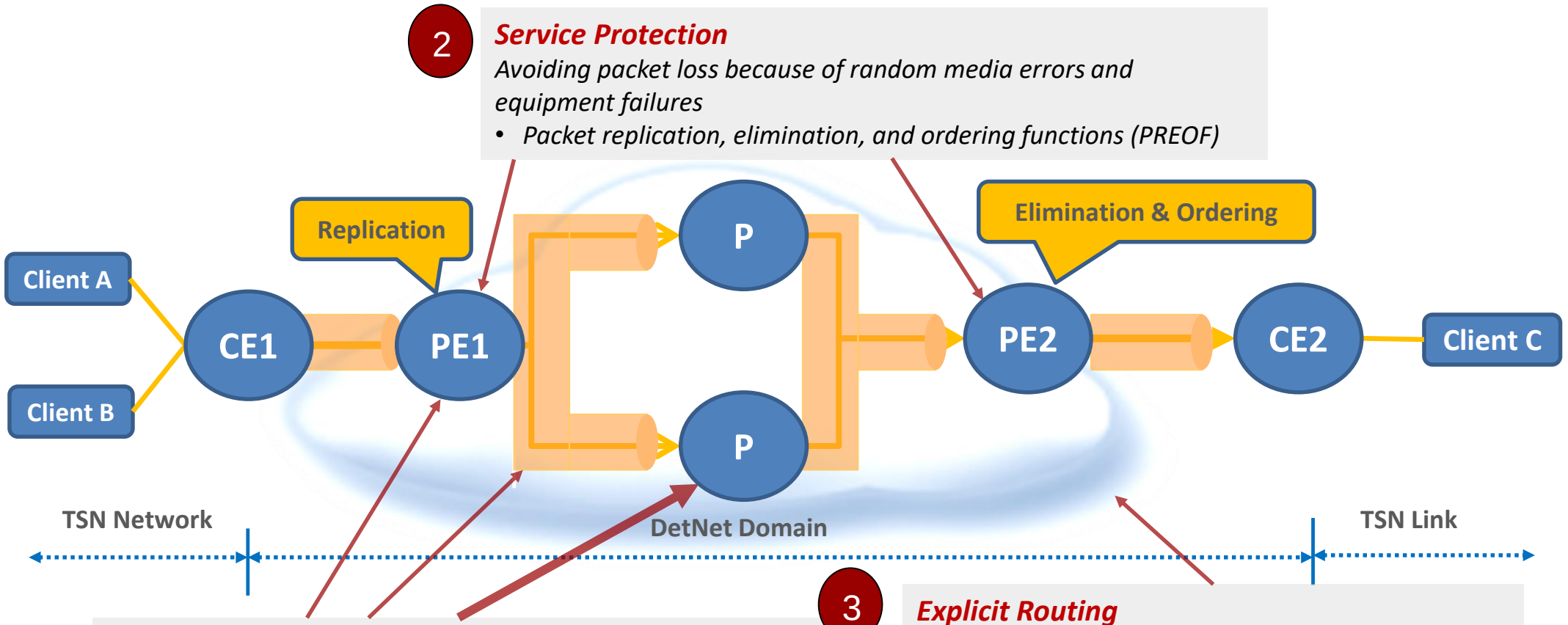
DetNet Data Plane Encapsulation

- Current focus of IETF work is on the DetNet data plane
- Currently two data plane drafts in progress
  - draft-geng-detnet-dp-sol-srv6
    - congestion protection (resource reservation and queuing management for low loss and latency),
    - service protection via packet replication, duplicate elimination, and ordering functions (PREOF) for zero loss
    - explicit routing for further latency management, fine-grained flows, etc.
  - draft-ietf-detnet-dp-sol-mpls
    - MPLS-based data plane that fully supports DetNet functionality:
    - congestion protection/service protection/explicit routing
  - draft-ietf-detnet-dp-sol-ip
    - Simplified IP-based data plane for partial solution that is easier to implement
    - Only provides congestion protection and explicit routing functions based on IP 6-tuple* for flow identification

Scalability Concerns

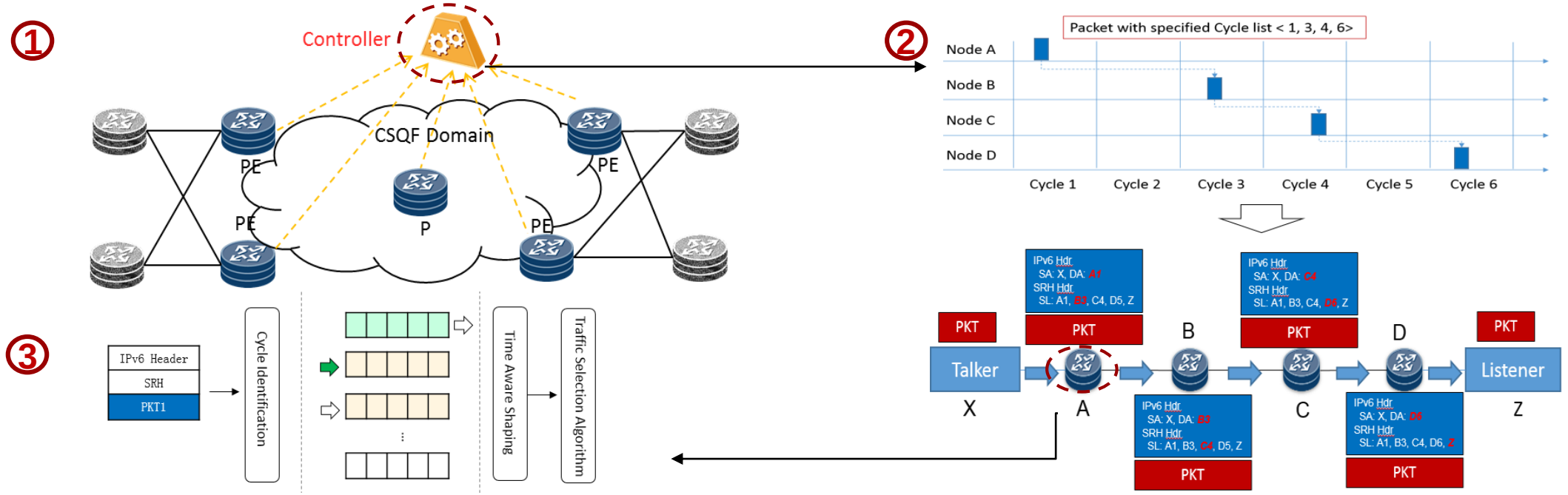
- Scalability is a major issue for the DetNet control plane and data plane (#flows)
- DetNet domains can be expected to support a large number of flows that will need particular queuing disciplines and/or resource allocation, depending on the requirements for each particular flow
  - This could require both a large amount of signaling for dynamic signaling protocols (such as an RSVP-TE session to establish and maintain each flow)
  - Could also require state in each forwarding node on the path of each flow
  - DetNet could have many more flows than classical RSVP-TE deployment
    - Classical RSVP-TE in SP backbone: single full mesh of traffic tunnels between nodes (N2 sessions)
- This generally argues against the use of per-flow:
  - Per-flow state across P nodes
    - Better: aggregated forwarding/queuing state
  - Per-flow signaling (such as RSVP-TE)
    - Better: centrally managed approaches

Data Plane Technologies for Deterministic Networking



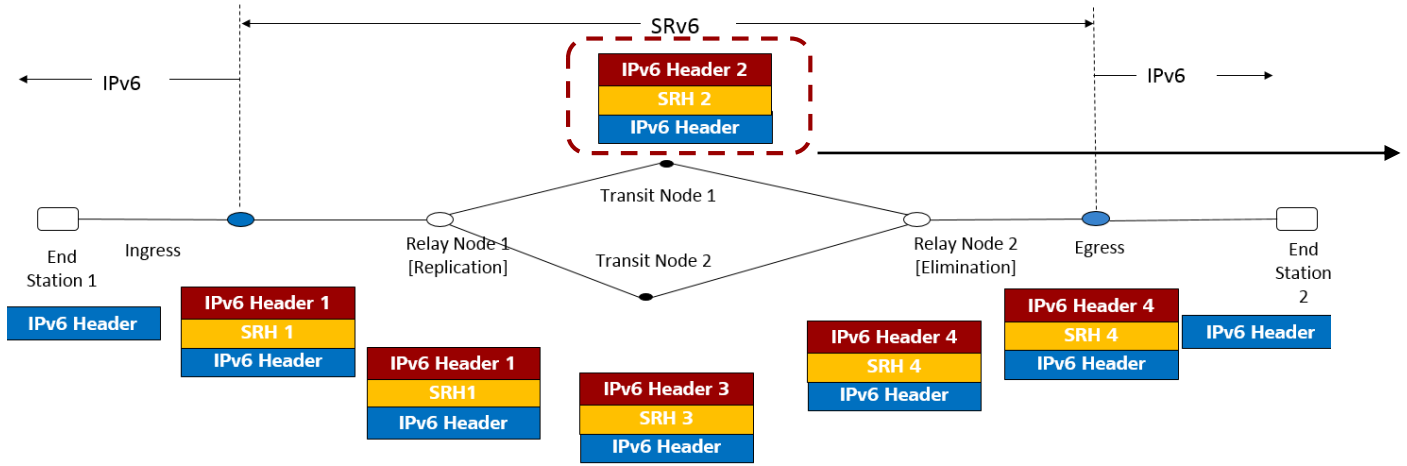
DetNet Data Plane Solution- Congestion Protection

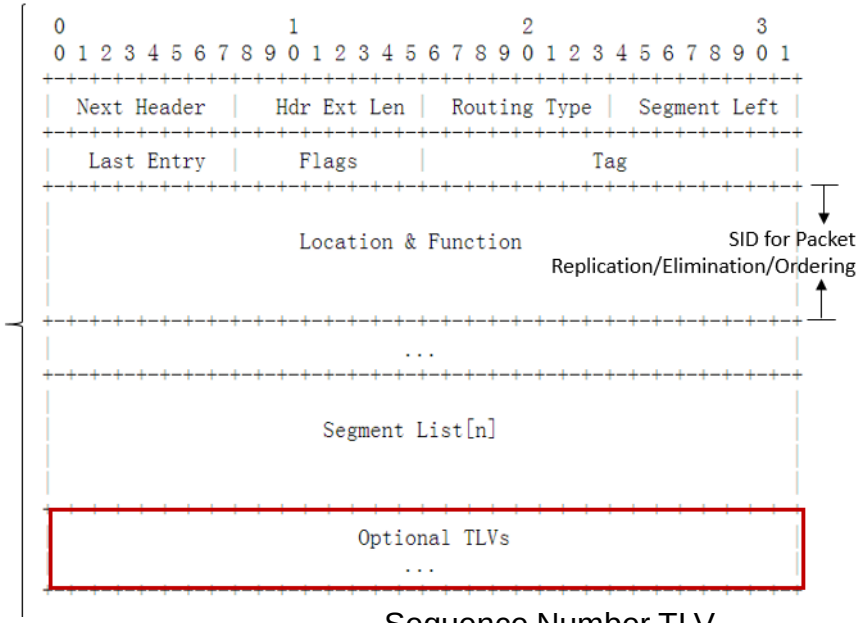
- CSQF： Cycle Specified Queuing and Forwarding:
  - Controller Computation：Collect the topology and latency information of the network, and compute a valid path；
  - Forwarding the packet based on SRH：based on the SRv6 network programming, SID in the SRH can be used to indicate the cycle number of every hop, and the end to end route is also determined；
  - Cyclic Queuing and Scheduling: IEEE 802.1 Qch based queuing management algorithm;



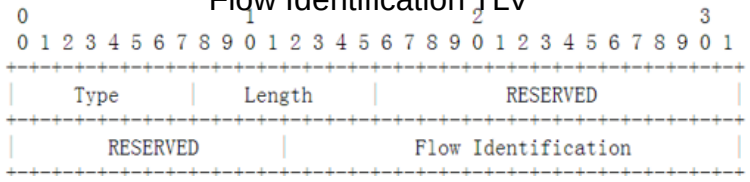
DetNet Data Plane Solution- Seamless Redundancy

- PREOF： Packet Replication、Elimination and Ordering
  - Explicit route to deliver the DetNet flow : steer the packet with the segment List in SRH;
  - Indicating packet processing, including packet replication/elimination/ordering: extend the SID functions;
  - Flow Identification and Sequence Number: define new SRH TLVs;

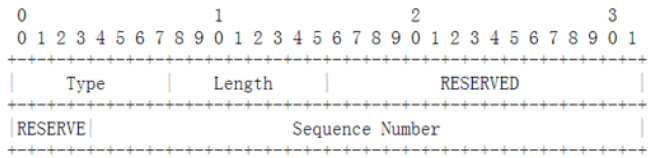




Flow Identification TLV

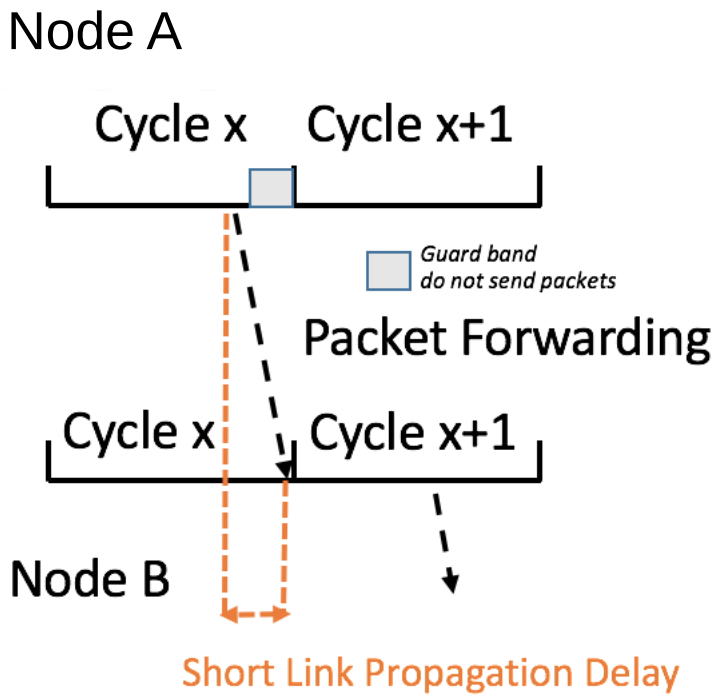


Sequence Number TLV



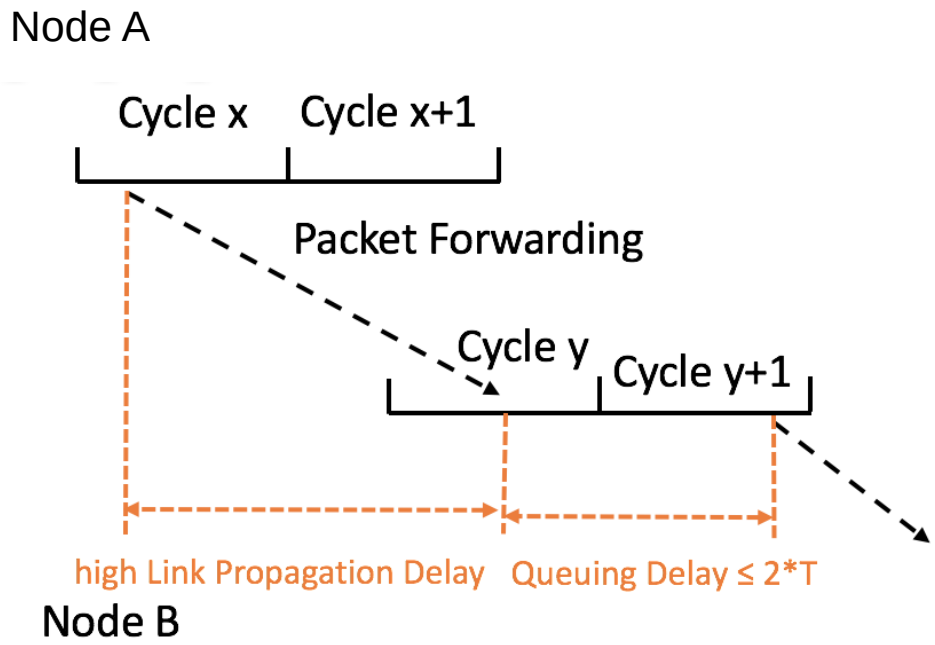
Cyclic Forwarding in TSN

- TSN-CQF: synchronous forwarding scheme without per-flow queuing state on every hop
  - Synchronous packets forwarded across all hops within a single cycle (10..100 µs)
- Challenges
  - Designed for Ethernet campus-sized networks
  - High accuracy time synchronization requirement (ns)
  - Limited physical size due to cycle time
  - The larger the network, the smaller the percentage of traffic that an be synchronous.
  - Example (extreme to make point):
    - 10 µs cycle time: max network size: 2 Km, after < 1 Km only
      < 50% traffic could be synchronous



Large-Scale Network Cyclic Forwarding

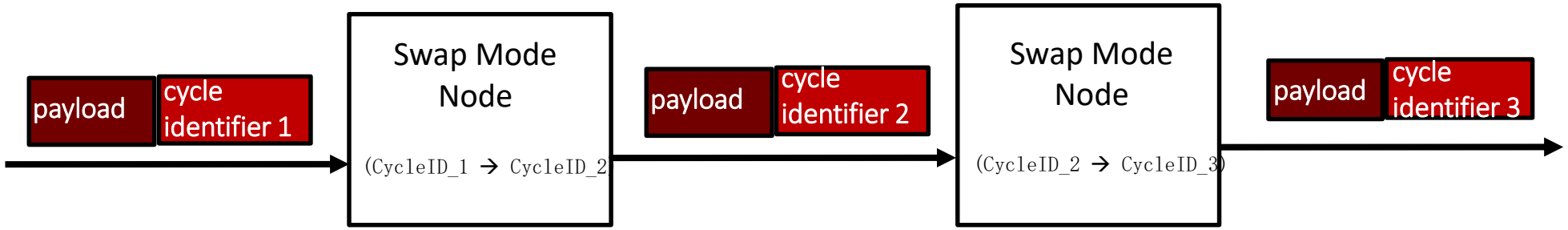
- Carry cycle-identifier in packet
- No synchronous forwarding:
  - Buffer cycles and send after all packets for cycle arrived
- Results
  - Keep the key benefits
    - Easy calculated end-to-end-delay (sum(per-hop-cycle-delay))
    - Tight bounded jitter O(cycle-time) [µs]
  - Eliminated physical scale limitations
    - Can support arbitrary link-propagation delay, hop, end-to-end delay
  - Eliminates need for tight time-synchronization
    - Requires only frequency synchronization in order of cycle time (µs instead of ns)
    - Frequency synchronization has no problems with difficult asymmetric link problem



Cyclic Queuing Can Work For Any Forwarding Plane

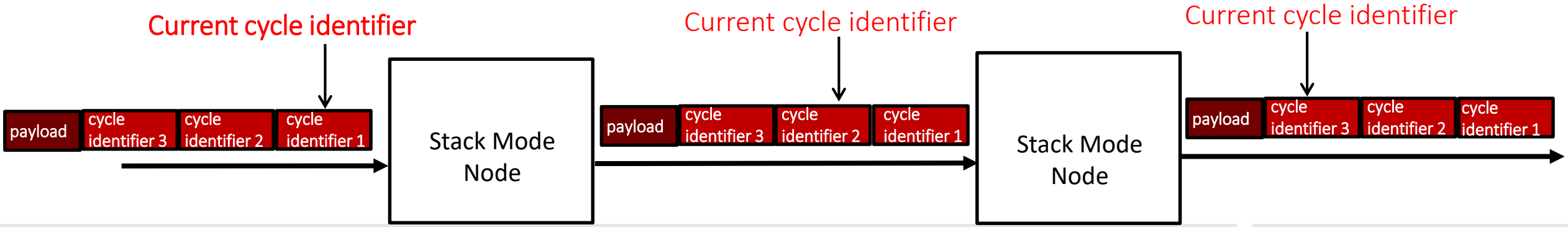
- Swap Mode (similar to MPLS): In-packet Cycle-identifier (applicable to any forwarding plane)

  Each node pre-provisioned with cycle-mapping table (e.g.: from PCECC)



- Stack Mode (for SR-MPLS, SRv6): In-packet: stack of cycle-identifiers, one for each hop

  Each hop maps based on next cycle-identifier



DetNet Control Plane

- The DetNet WG hasn't yet defined any control plane mechanisms or protocols
  - These would be used for:
    - Instantiating DetNet flows in a DetNet domain
      - Queuing buffer reservations, link BW reservations, etc.
    - DetNet Service Label allocation and distribution, when the DetNet MPLS encapsulation is in use
    - Advertising node and link resources, capabilities and adjacencies, both static and dynamic
- The IETF has defined many possible protocols and mechanisms that could be used or extended for DetNet use, there is no real need to invent anything new from scratch
- These generally fit into three general categories:
  - Distributed signaling
  - Centralized control and management
  - Hybrid (combination of the two)

Distributed Signaling Protocols

- The IETF has had a long history of defining distributed signaling protocols for flow establishment and routing and MPLS label distribution
- Such protocols include (but are not limited to):
  - RSVP and RSVP-TE, used for flow and LSP establishment with resource reservations, the latter with the ability to define a constrained and ordered set of transit nodes for the flow, the latter also used by GMPLS for signaling in non-IP transport networks (also working in conjunction with IGPs)
  - IGP extensions for Segment Routing
  - BGP, along with EVPN and other extensions

Centralized Control and Management

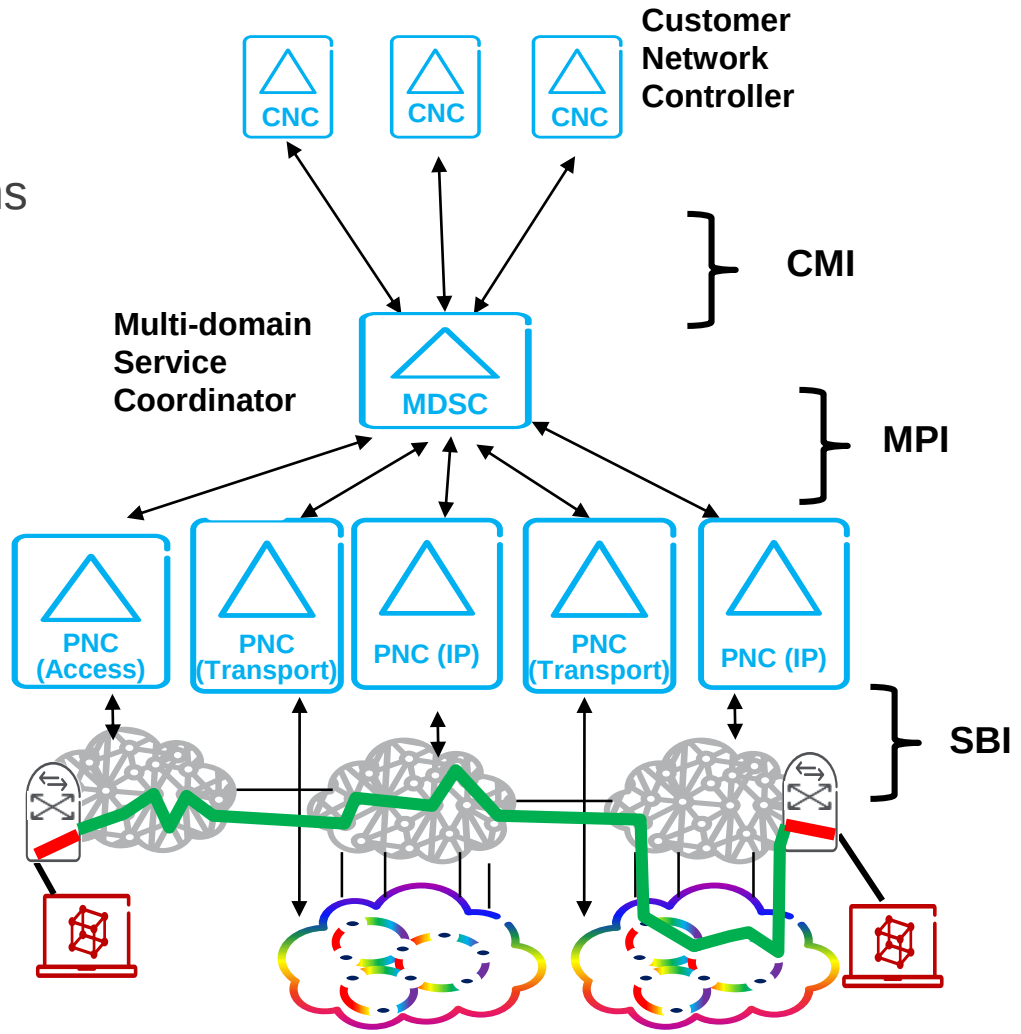
- In modern network architectures, SDN-like principles for centrally managing networks and domains are increasingly popular
- Controller is the sole source of routing and forwarding information
- Some example mechanisms for central management of networks and domains that could be used for DetNet flow establishment and label distribution include:
  - Segment Routing (SR-MPLS and SRv6) data plane in conjunction with centralized flow establishment and complete label stack or v6 address list distribution to Detnet domain entry nodes
  - NETCONF/YANG management of all nodes in a DetNet domain with traditional (non-SR) MPLS and/or IP forwarding plane
  - Centralized PCE with PCEP to all nodes in a DetNet domain with traditional (non-SR) MPLS forwarding plane

Hybrid Architectures

- Of course, a network architecture doesn't necessarily need to be strictly using distributed signaling or centralized management. Hybrid architectures are also possible, such as:
  - A centralized controller/PCE gathers the information from the protocols and signaling active in the network, either by direct polling/reporting from the nodes, and/or snooping on protocol exchanges (IGP, BGP, etc.) It then:
    - Uses PCEP to the edge nodes, where the PCE is used instead of local route computation in an edge node, and PCEP is used to initiate RVSP-TE signaling for a new traffic engineered LSP
    - Uses NETCONF/YANG to collect performance data and resource utilization for nodes and links in the network, and these are then pushed to edge nodes for use in local path computations for RSVP/TE
  - Many other combinations are also possible

ACTN and DetNet Synergy

- ACTN(Abstraction and Control of TE Network) enables mechanisms for central management of networks and domains
  - It could be used for DetNet flow establishment and label distribution
- ACTN works with various data plane technologies, including MPLS, SR-MPLS, SRv6
  - Can support SRv6-based、MPLS-based and IP-based DetNet data planes
- ACTN works with both centralized and hybrid approaches
  - centralized path computation and either centralized path provisioning (PCE Central Controller or NETCONF/YANG)
  - or distributed path provisioning using RSVP-TE



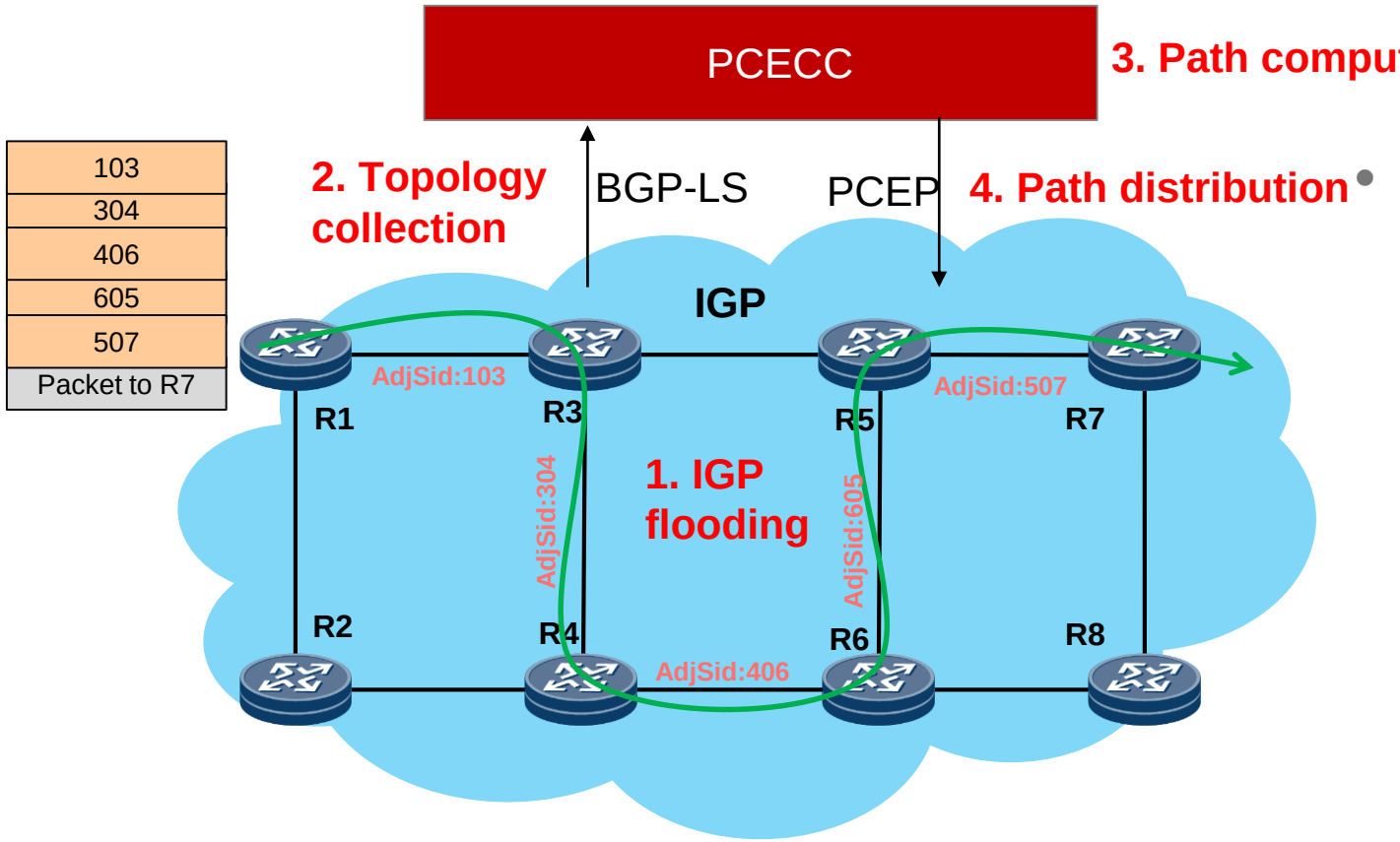
Comparison of Centrally Managed Approaches

- NETCONF/YANG
  - wasn’t designed for high duty cycle environments such as individual flow establishment, thus depending on the implementation, there could be a high cost of usage in terms of node system resources, plus there could be a good deal of state installed in each node to support each flow.
- PCE Central Controller (PCECC) with traditional MPLS LSPs is a better approach,
  - as the signaling was designed for sort of application. However, this could still lead to a good deal of state installed in each node to support each flow.
- PCECC combined with SR-MPLS or SRv6
  - is a much more scalable approach than the others, as it allows the details of packet or flow treatment to be encoded directly in the SIDs on each packet in a flow, which substantially reduces dynamic signaling overhead and reduces the amount of state in network nodes. This approach also allows the integration of DetNet domains with general SR-based backbone networks.

Management Alternatives for MPLS-SR and SRv6

- The IETF is working on a number of alternative signaling mechanisms for MPLS-SR and SRv6
  - Extensions to IGPs and BGP to support distributed signaling
  - BGP-LS and BGP route reflectors can be added for a hybrid solution
  - A promising mostly-centralized approach could be to use a PCECC to push paths represented by SID lists while using BGP-LS to collect network topology and link state information
    - An IGP is used for the usual link state flooding in order to establish adjacencies, but not for DetNet flow path calculations, only for best effort traffic
    - DetNet flow path calculations are centralized in the controller
    - DetNet extension to mechanisms discussed in draft-zhao-pce-pcep-extension-pce-controller-sr

PCE Central Controller (PCECC) SR-based DetNet Path Establishment



Highlights:

- Agility: Centralized path calculations for optimal global resource utilization
- Scalability: Path state is only maintained on ingress nodes, forwarding instructions to each transit node are encoded in the SIDs
- PCECC can have an NBI to receive path requests or orchestration

Conclusion

- DetNet can provide network service with bounded latency, low jitter and almost 0 loss;
- SRv6 is one of the most promising data plane solutions for DetNet, which can support all the data plane functions provided by DetNet, including congestion protection/seamless redundancy/explicit route;
- SR can be combined with ACTN and/or centralized path computation for overall network resource utilization optimization and with EVPN and/or VPN+ signaling for VPN support
- Much of this work extends beyond DetNet to wide-area control planes in general, especially where SR is used

THANK YOU

www.huawei.com