Deterministic Networking

BOF Status

September 19, 2014

Who, What, Why, and How

What is Deterministic Networking?

Same as normal networking, but with the following features for **critical data streams**:

1. **Time synchronization** for network nodes and hosts to better than 1 µs.
2. Software for **resource reservation** for critical data streams (buffers and schedulers in network nodes and bandwidth on links), via configuration, management, and/or protocol action.
3. Software and hardware to ensure **extraordinarily low packet loss ratios**, starting at 10^{-6} and extending to 10^{-10} or better, and as a consequence, a **guaranteed end-to-end latency** for a reserved flow.
4. **Convergence** of critical data streams and other QoS features (including ordinary best-effort) on a single network, even when critical data streams are 75% of the bandwidth.

Who needs Deterministic Networking?

- Two classes of bleeding-edge customers, Industrial and Audio/Video. Both have moved into the digital world, and some are using packets, but now they all realize they must move to Ethernet, and most will move to the Internet Protocols.

1. **Industrial:** process control, machine control, and vehicles.
   - At Layer 2, this is IEEE 802.1 **Time-Sensitive Networking (TSN)**.
   - Data rate per stream very low, but can be large numbers of streams.
   - Latency critical to meeting control loop frequency requirements.
2. **Audio/video:** streams in live production studios.
   - At Layer 2, this is IEEE 802.1 **Audio Video Bridging (AVB)**.
   - Not so many flows, but one flow is 3 Gb/s now, 12 Gb/s tomorrow.
   - Latency and jitter are important, as buffers are scarce at these speeds.

- (You won't find any more market justification in this deck.)

Why such a low packet loss ratio?

Back-of-the-envelope calculations:

1. **Industrial:**
   - Automotive factory floor: 1000 networks • 1000 packets/s/network • 100,000 s/day = 10^{11} packets/day.
   - Machine fails safe when 2 consecutive packets are lost.
   - At a random loss ratio of **10^{-5}**, 10^{-10} is chance of 2 consecutive losses.
   - 10^{11} packets/day • 10^{-10} 2-loss ratio = **10 production line halts/day**.
   - In extreme cases, lost packets can damage equipment or kill people.
2. **Audio video production:** (not distribution)
   - 10^{10} b/s • 10 processing steps • 1000 s/show = 10^{14} bits = 10^{10} packets.
   - Waiting for ACKs and retries = too many buffers, too much latency.
   - Lost packets result in a **flawed master recording**, which is the user's end product.

How such a low packet loss ratio?

1. **Zero congestion loss**.
   - This requires reserving resources along the path. (Think, “IntServ” and “RSVP”) You cannot guarantee anything if you cannot say, “No.”
   - This requires hardware in the form of buffers, shapers, and schedulers. Overprovisioning not useful: its packet loss curve has a tail.
   - Circuits only scale by aggregation in to larger circuits. (MPLS? Others?)
   - 0 congestion loss goes hand-in-hand with finite guaranteed latency.
2. **Seamless redundancy**.
   - 1+1 redundancy: Serialize packets, send on 2 (or more) fixed paths, then combine and delete extras. Paths are seldom automatically rerouted.
   - 0 congestion loss means packet loss is failed equipment or cosmic rays.
   - Zero congestion loss satisfies some customers without seamless redundancy. The reverse is not true in a converged network—if there is congestion on one path, congestion is likely on the other path, as well.

Current IEEE 802 Status

IEEE 802 standards now and coming

802.1 Audio Video Bridging is now the Time-Sensitive Networking TG.

- **Time:** A plug-and-play Precision Time Protocol (PTP) profile that allow bridges, routers, or multi-homed end stations to serve as "time relays" in a physical network, regardless of L2/L3 boundaries. (Complete. Enhancements in progress.)
- **Reservation:** A protocol (MSRP) to reserve bandwidth along an L2 path determined by L2 topology protocol, e.g. ISIS. (Complete. Enhancements in progress.)
- **Execution:** Several kinds of resources (shapers, schedulers, etc.) that can be allocated to realize the promises made by the reservation. (See next slide.)
- **Path distribution:** ISIS TLVs to compute and distribute multiple paths through a network. (In progress)
- **Seamless Redundancy:** 1+1 duplication for reliability. (In progress)

IEEE 802 schedulers and shapers

- **AVB Credit-Based Shaper:** Similar to the typical run rate/burst rate shaper, but with really useful mathematical properties. (Done)
  - Only parameter = bandwidth.
  - The impact of any number of shapers = the impact of one shaper with the same total bandwidth.
- **Transmission preemption:** Interrupt (1 level only) transmission of an Ethernet frame with a frame with tight latency requirements, then resume the interrupted frame. (In progress.)
- **Time scheduled:** Every bridge port runs a synchronized, repeating schedule that turns on and off each of the 8 queues with up to nanosecond precision. (In progress.)
- **Synchronized Queuing and Forwarding:** Every flow proceeds in lock-stepped transmission cycles, like arterial blood. (In progress).

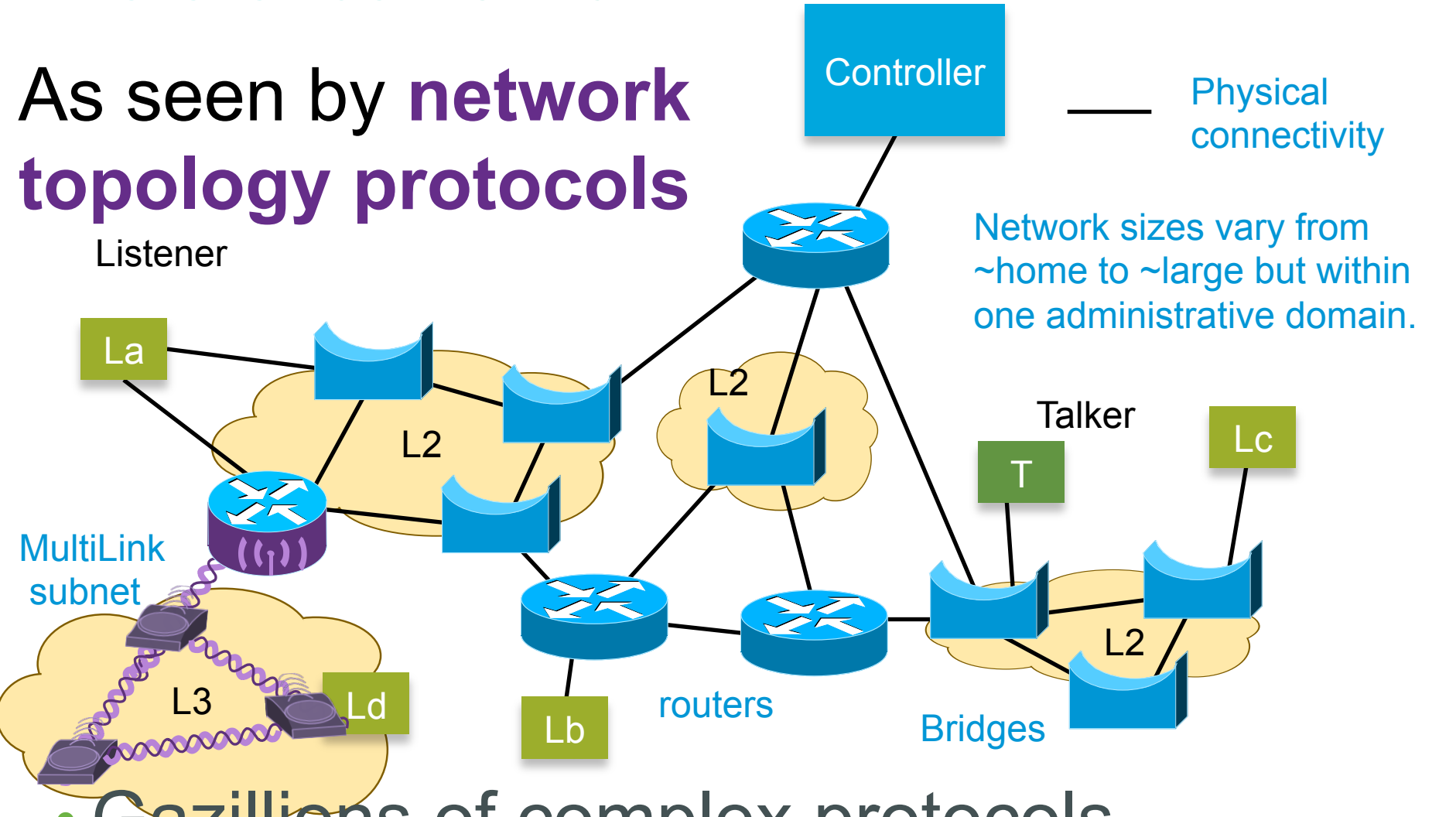
IEEE AVB standards success

- The AVB TG gained some slight traction for the home – every Apple laptop supports AVB – but this is not widely used. Yet.
- The biggest use case today is audio and video studios and their equipment suppliers, e.g. Dolby, Gibson, Harman, Riedel, Extreme, Arista.
- We anticipate shipment of AVB in automobiles in 2015.
- The AVnu Alliance, an industry consortium, was created to promote the AVB standards (as Ethernet Alliance supports IEEE 802.3 standards). Originally oriented towards audio and video in the home and studio, focus now approximately even between AV and industrial/automotive.

Mixed L2/L3 need

Reference network

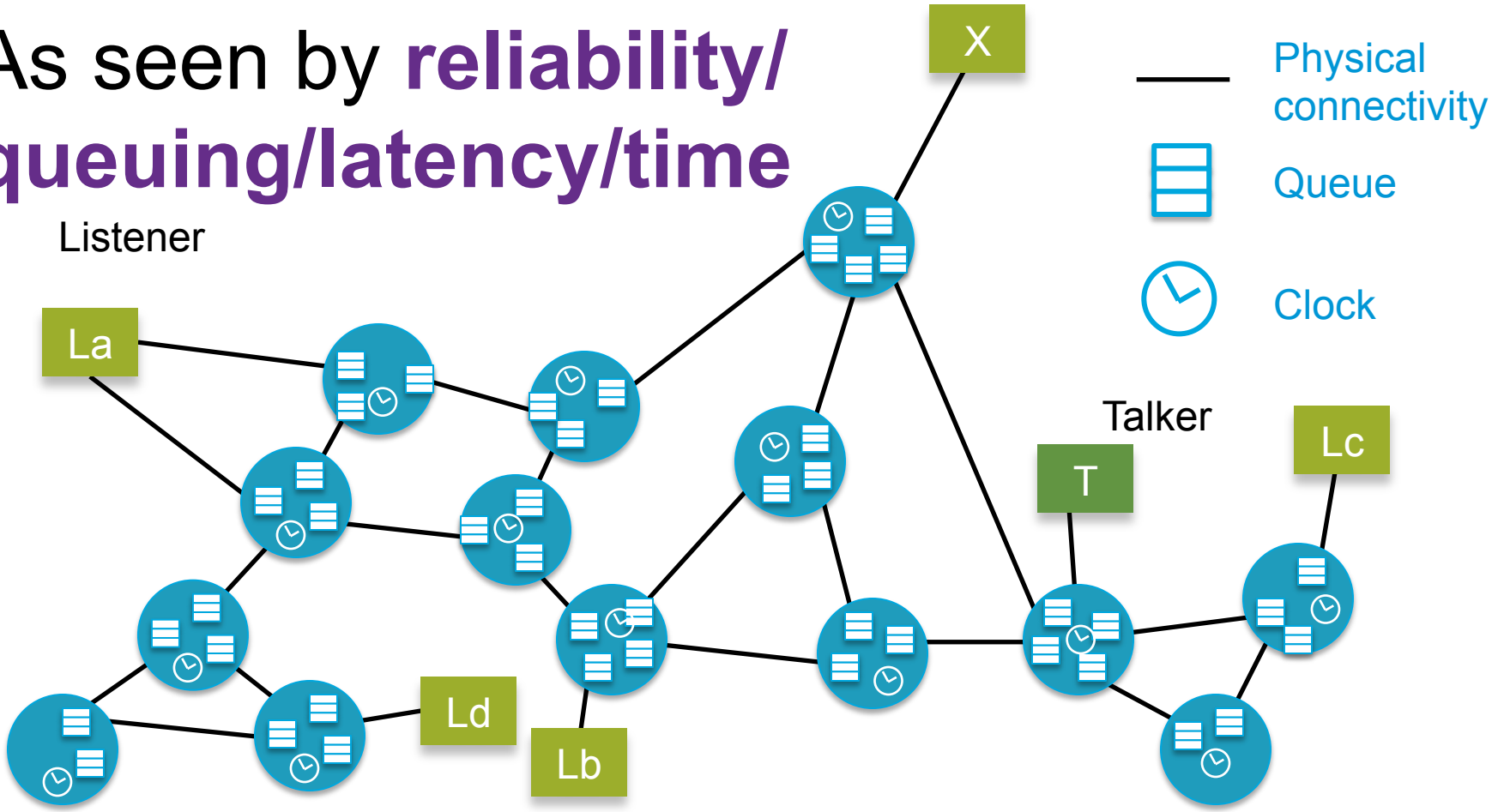
As seen by **network topology protocols**



- Gazillions of complex protocols

Reference network

As seen by **reliability/ queuing/latency/time**



- Just nodes, queues, clocks, and wires!!

Mixed L2/L3 = IEEE/IETF cooperation

- Both bridges and routers are important parts of these networks. Neither is going away.
- Every box along the path must reserve resources, and participate in the reservation protocols, whether a bridge or a router.
- Reservations from pre-configuration, management, or protocol.
- Hosts = applications can participate in the protocols.
- Hosts and operations managers don't know or care whether network is bridged or routed. One Host UNI, one operator view.
- There are valid use cases for application-driven peer-to-peer control flow models, for centrally controlled models, and for mixed scenarios.

Next steps

Divide and conquer: DetNet WG

- Time synchronization is a separate problem. DetNet needs time synch, but other Working Groups and/or SDOs will provide.
- Low-level (typically hardware) queuing, buffering, shaping, and scheduling mechanisms can be left to IEEE 802.
- What is needed (one opinion!): Protocols for:
  - Determining the physical topology and/or the L2/L3 topology of the combined network.
  - Taking reservation requests, and accepting or refusing them.
  - Creating multiple unicast and multicast paths through the network.
  - Setting parameters in routers and bridges to control queuing, buffering, shaping, and scheduling mechanisms along those paths.
  - IEEE 802 has standards completed or in progress for all of these pieces, but they are only for L2, and they do not include a centrally-controlled model. DetNet WG will determine their relevance to the DetNet problem.

IETF Progress

- Side meeting held at IETF90 in Toronto. 25-30 attendees. All positive comments, good advice – particularly with regard to bringing problems, not solutions, to IETF.
- IETF mailing list started: detnet@ietf.org
- First version of Problem Statement draft uploaded and now being edited.
- Several versions of a Working Group charter proposal circulated on mailer.
- Active discussion on mailer on further documents required for BOF at IETF91 (Honolulu).
- BOF request submitted. 2-hour session.

Short-term decisions

- BOF at IETF91?
- In which Area?