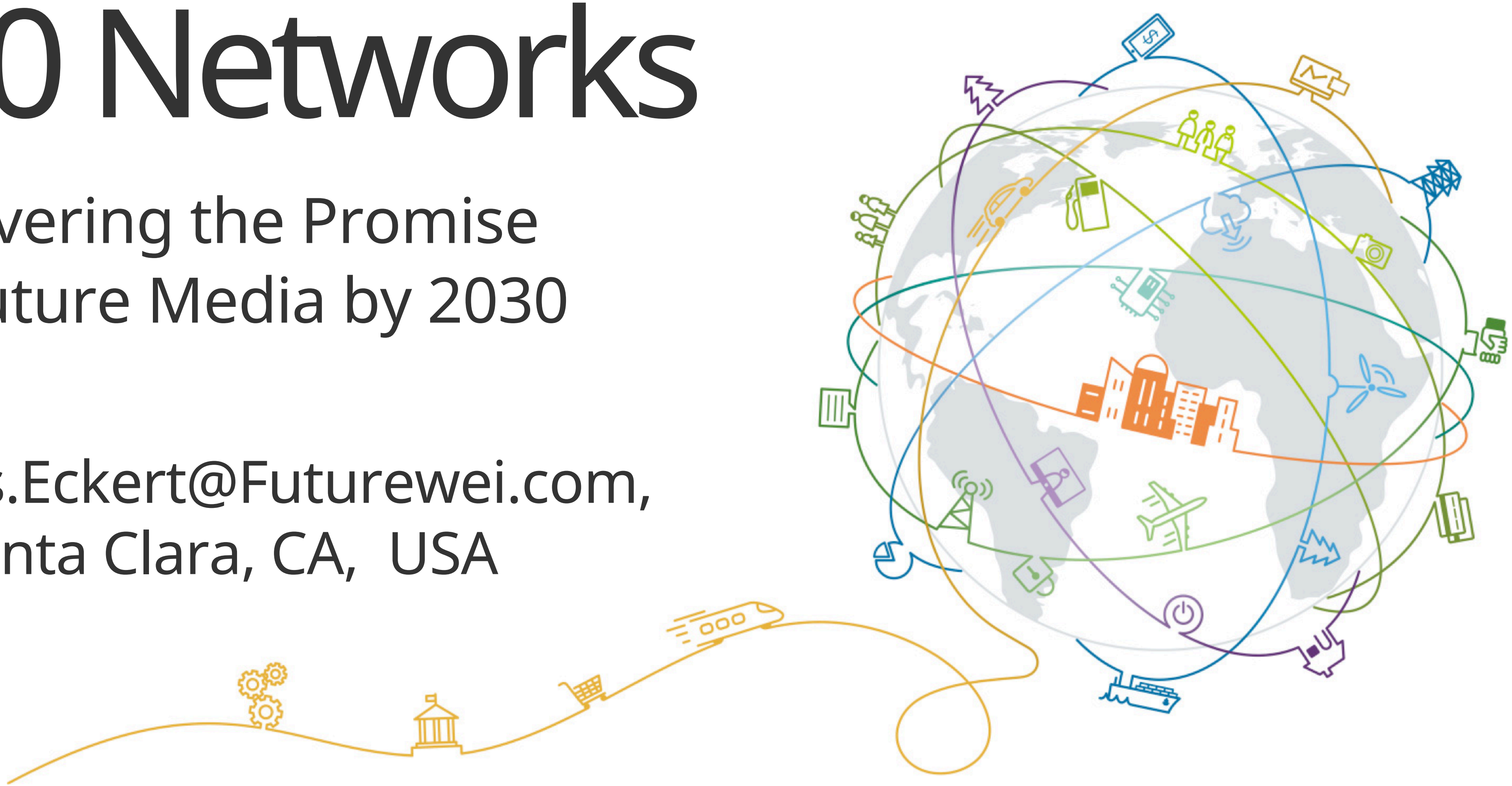
Network Technologies for 2030 Networks

Delivering the Promise
of Future Media by 2030

Toerless.Eckert@Futurewei.com,
Santa Clara, CA, USA



IEEE International Conference on Cloud Networking
6. November 2019, Coimbra, Portugal

Technologies for 2030 Networks

WHY

*more users, (IOT) devices, quality, critical/**future applications**, …*

WHERE

Metropolitan network infrastructures (fixed/mobile)
(for this slide deck)

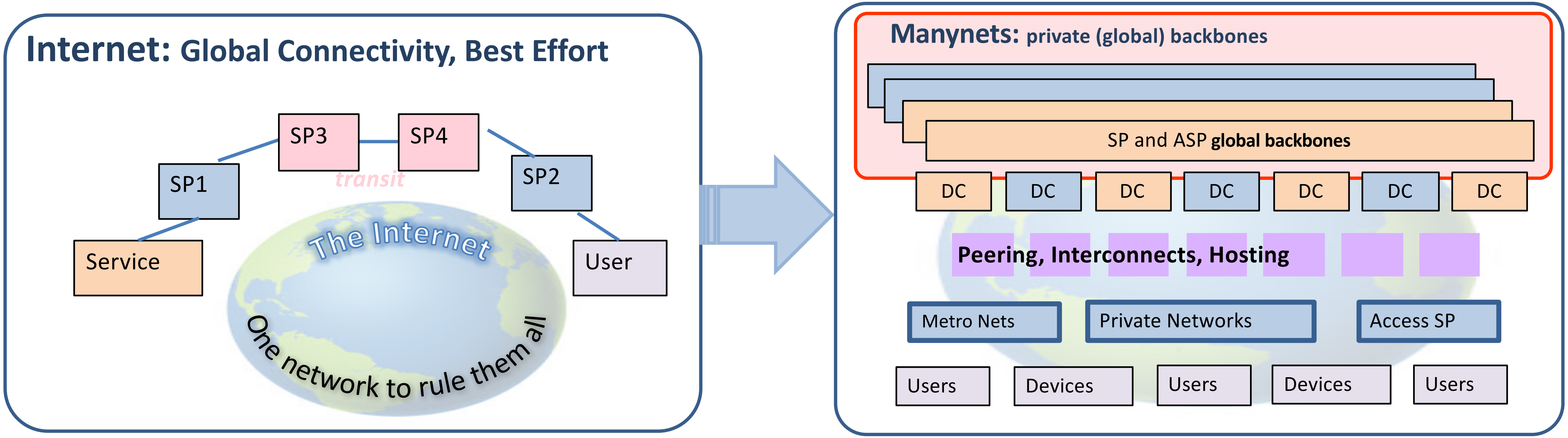
WHAT

New network protocol/services technologies
NewIP, High-Precision, Qualitative Comms, Holographic, …

….

GREAT SUCCESS ! ??

WHERE: from Internet to ManyNets



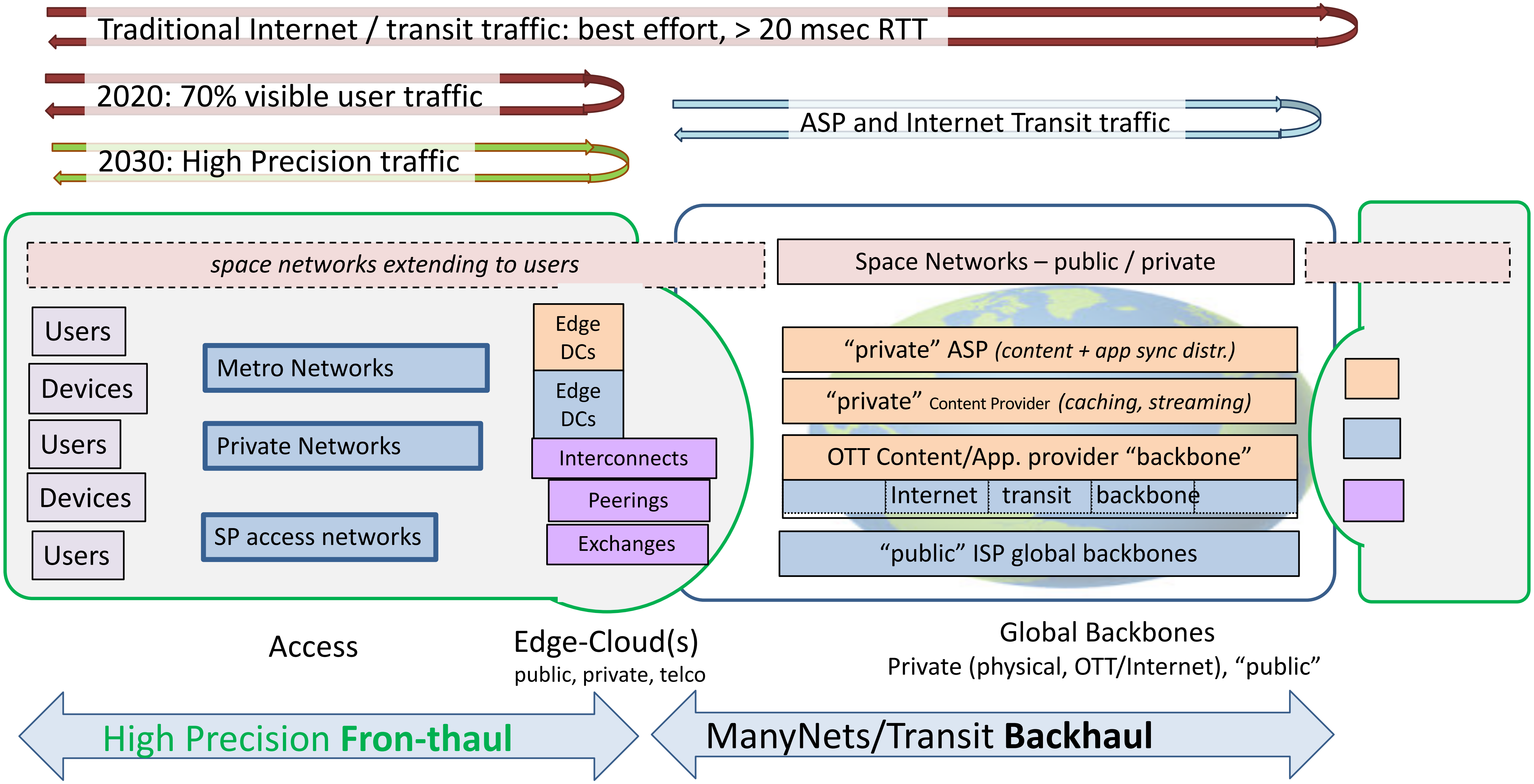
TCP/IP was/still-is designed for multi-transit Internet (1970th)

Since ca. 2010:

Scaling global services requires Edge Data-Center (Cloud)

Raise of global ASP, peerings, cloud services, Death of transit

WHERE: 2030 Internet Front-haul / Backhaul



Network 2030 Internet Front-haul

Opportunity for better services

Transit has no business relationship with neither user nor service

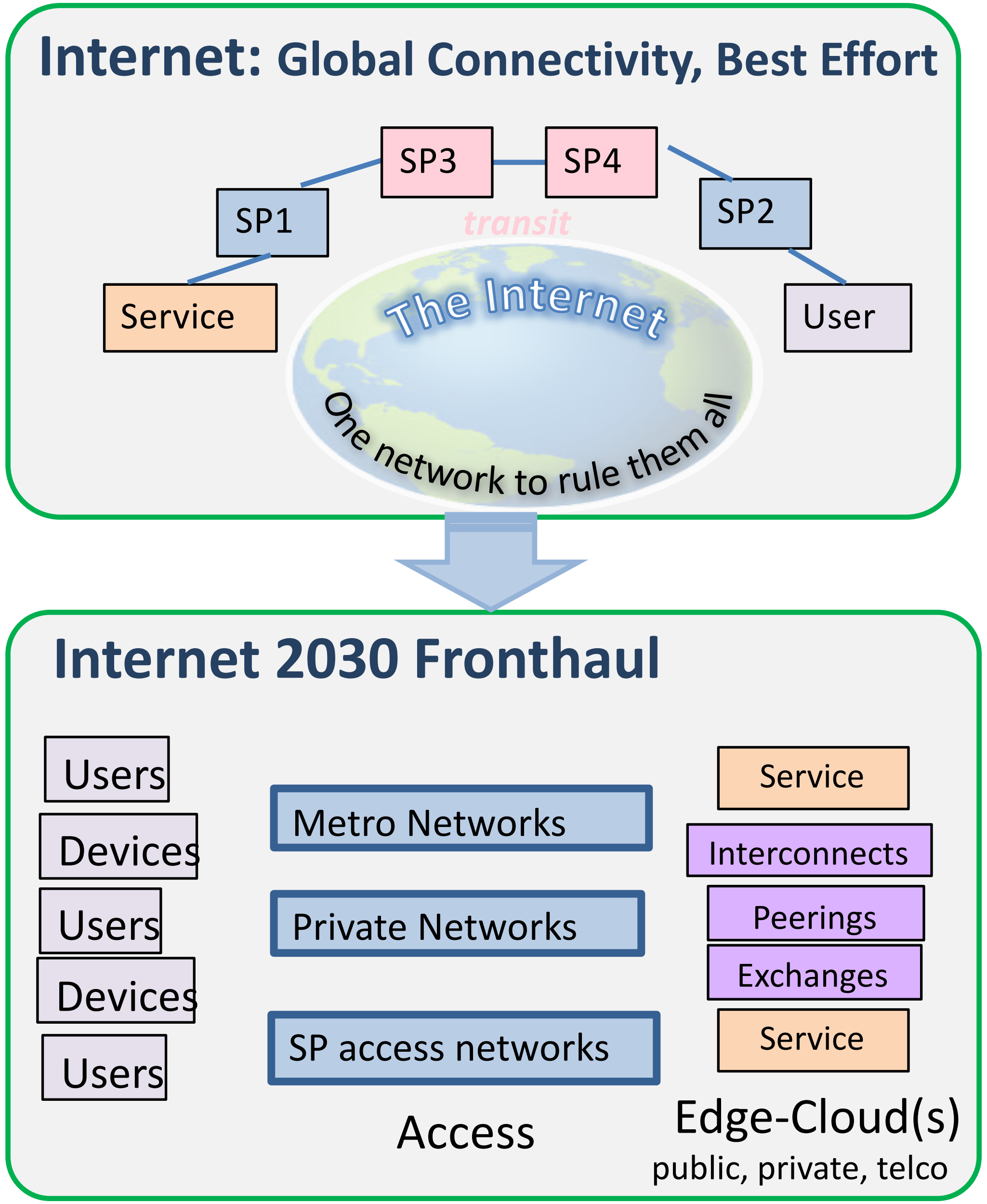
- Key reason for current Internet to be limited to "Best Effort" (BE) services

Fronthaul can be architected without transit.

- Key enabler for better than BE services in 2030 networks.
- User/Service <-> User/Service: 2 networks + interconnect.

Geographical limits (e.g.: 100 mile diameter) of Front-haul permits latency critical services.

- Limited physical latency (low msec).



Technologies for 2030 Networks

Refined 2030 Network Solutions Dependencies

WHY

WHERE

Forgot Something

WHO

Develops 2030 Network services
Operates 2030 Network services

HOW

What technologies do we need to enable **WHO** *?*

WHAT

…

GREAT SUCCESS !!???

WHO: Enable Innovation & Opportunity

Cloud drives Innovation & Opportunity

Software-ization is enabler for new Service Developers and Services Operator

DC-networks: becoming softwareized (P4,..)

Telco-Cloud => Software

5G 'Network', VPN, Security, Content services

MVNO, SD-WAN, ..

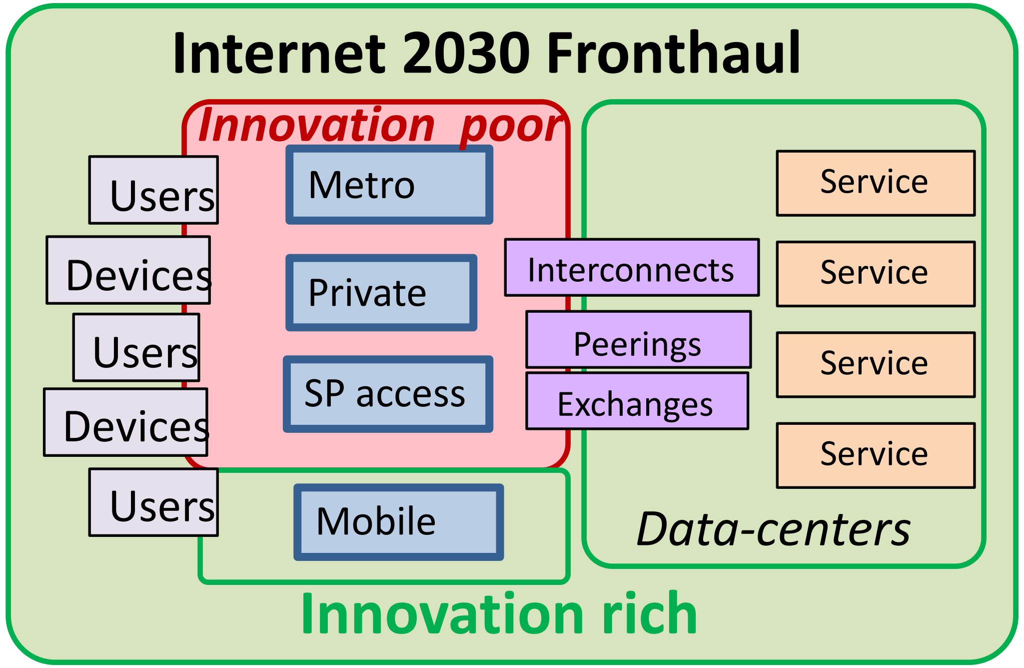
But Software is not 100Gbps…Tbps

Metro-Wireline Networks ?

Today: Speeds, Feed, Commoditization.

Key Technology gap:

Software-izable Tbps Internet 2030 access networks



HOW: Software-ized Metro-Network

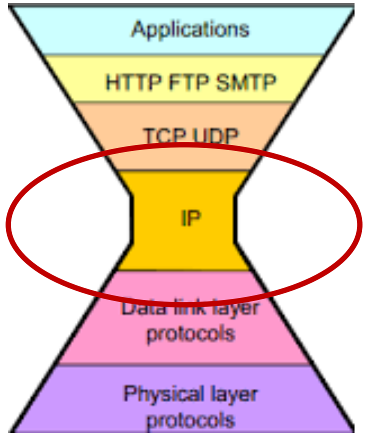
How to deploy & sell new network forwarding plane ?

Slices do NOT allow to program new forwarding planes

Overlays - VMs / Containers in Data-Center

No "bare-metal" option (hop-by-hop) on a metro-network (*)

Example (from last decade): Inability of ICN to get deployed



Innovation Bottleneck

Overlay enough for many network services, but not all:

Speed/Performance, Data-replication,

Latency sensitive: High-Precision, Deterministic, Coordinated, Qualitative, Lossless

Software-ization of Tbps forwarding plane: challenges for a decade

Flexible programming languages (P4 too limited)

Chip cost / power consumption

Virtualization: context overhead VM->Container->Lambda model

Programmable QoS (packet scheduling)

Resource Management (absolute vs. relative)

- ✓ **Innovation Above, Below, and Alongside**
- ✓ **But, Limited Innovation "Inside" the 'Net**
- ✓ **Inside of the network needds to change**
- ✓ **We Are Desperate to Innovate Inside**

Jennifer Rexford,
ACM Sigcomm 2018 Keynote Speech

(*) Commercial offerings, not talking about research networks from NSF.

Technologies for 2030 Networks

WHY
WHERE

WHO
HOW

WHAT

Current Internet/private network experience "challenged"

Big Packet Protocol
Coordinated Communications
High-Precision Communications
Qualitative Communications
Lossless Networking
Preferred Path Routing

GREAT SUCCESS !!!!?

BPP – ~~Big~~ Better Packet Protocol

Architecting the next-generation forwarding plane

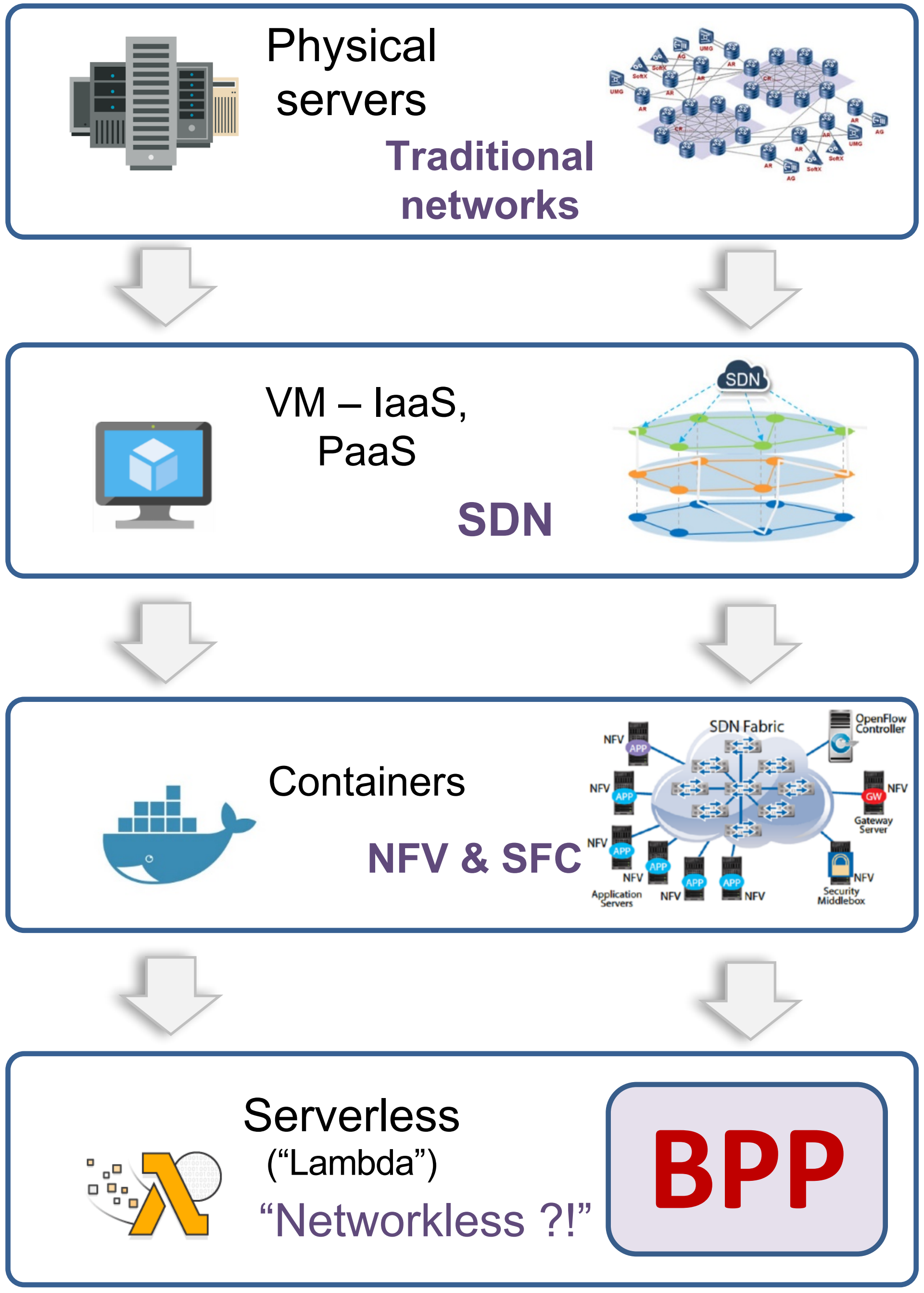
Network evolution so far:
- Primarily evolution of management/(control) plane

Very complex , incrementally grown forwarding planes:
- Duplicating efforts, limited in flexibility, Little/not App-"programmable"
- Various bad limitations. E.g.:

  IPv4 / IPv6, per-feature extension header, parameters per-hop SRv6
  fixed size addresses, minimal changeable by network, …
  MPLS, fixed size addresses, hacked bottom of stack parameters,…
  Ethernet, often same or better QoS features than IP/IP, but all different.

Programmable networks / slices / 2030 services (high precision,..)
- *Very challenging (if not impossible) to build well within this maze*

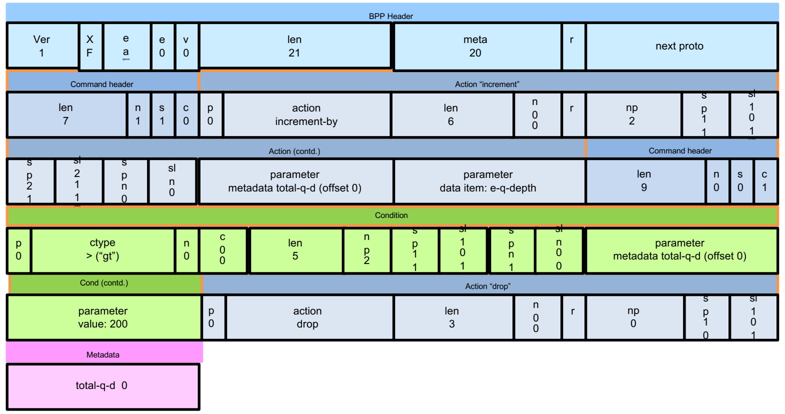
Can we design / frame a better, unified forwarding plane ?
- **Unify / superceed existing solutions, extensible**
- **(SLO) "Service Level Objective" based – "Do What i Mean Networking"**
- *Without all the boring bits of history*



Application/DC vs. Network evolution

BPP – Big Packet Protocol

Key architectural aspect for flexibility



Actions, Parameters and Metadata

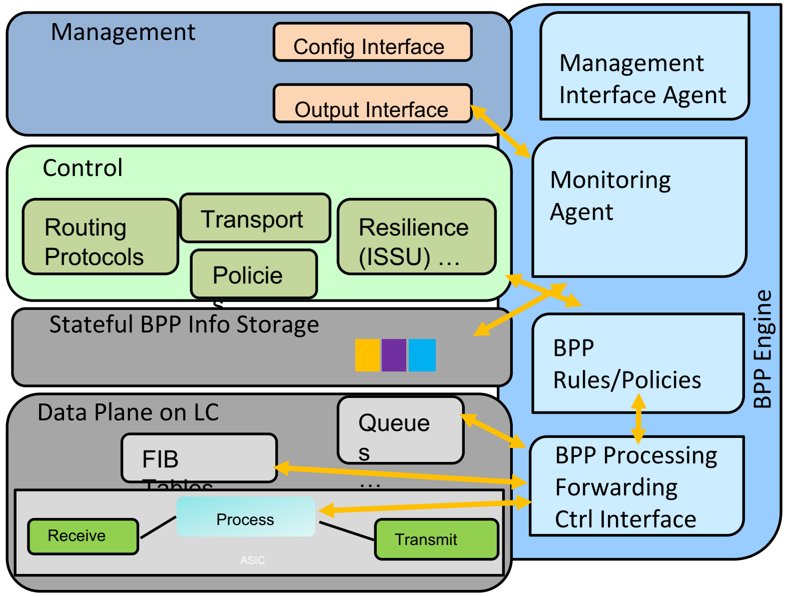
Actions (Commands) as Unit of forwarding plane extensibility
Per action parameters, but cross-action shared metadata
(to minimize headers)
Explicit indication of sequentialization / parallelization of actions

Flowlets: Integrated state-full and state-less Actions

Per flow actions minimize per-packet header
Per-packet actions allow scale to arbitrary number of flows in network.
Not all actions can be performed per packet, e.g.: limit rate to X Mbps

Efficient stacking or merging of headers

Generalization of MPLS/SR concept for efficient layering / tunneling
Look-ahead across multiple headers ?



Support for different address lengths

128 bit too long/inefficient for most private networks , IoT, replace existing solutions
Many open design issues – translation / integration, multiple instances, assignment authorities, …

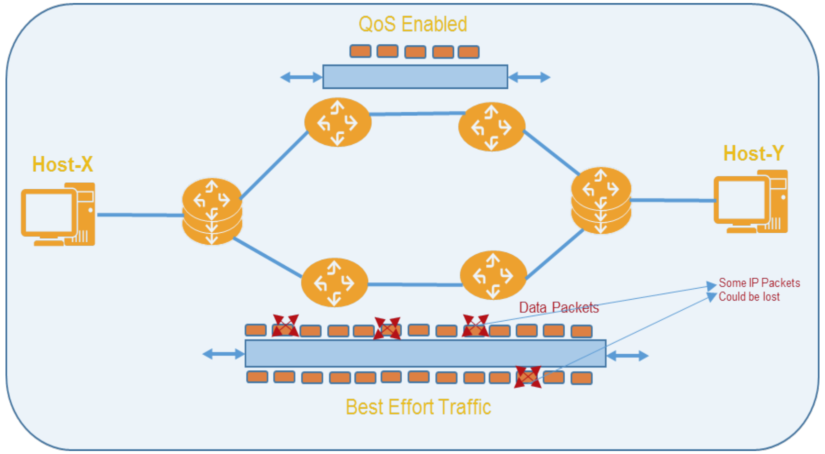
The Internet: ~~Best~~ Lousy Effort (throughput)

What is Internet (throughput) fairness ?
~~"all flows are created equal" ???~~
E.g.: competing 4K TV and Smartphone
Paying more will not get you more throughput during congestion

How to achieve fairness ?
Path-RTT, transport-protocol-aggressiveness

Guarantees ?
Absolute reservation == scale issues
Relative throughput easier !!

Limited by Loss and Capacity:
Higher performance requires lower loss or higher overhead



Internet: Best Effort

- ✓ "Best Effort" Per-Hop Behavior (PHB)
- ✓ ECN / AQM (maybe)
- ✓ "Shortest Path" routing

- ✗ Throughput / Priority Guarantees
- ✗ Latency Guarantee (path/queuing)
- ✗ No-Loss guarantees
- ✗ Jitter

~~Rich~~ QoS ~~services~~ in private IP networks

Archaic

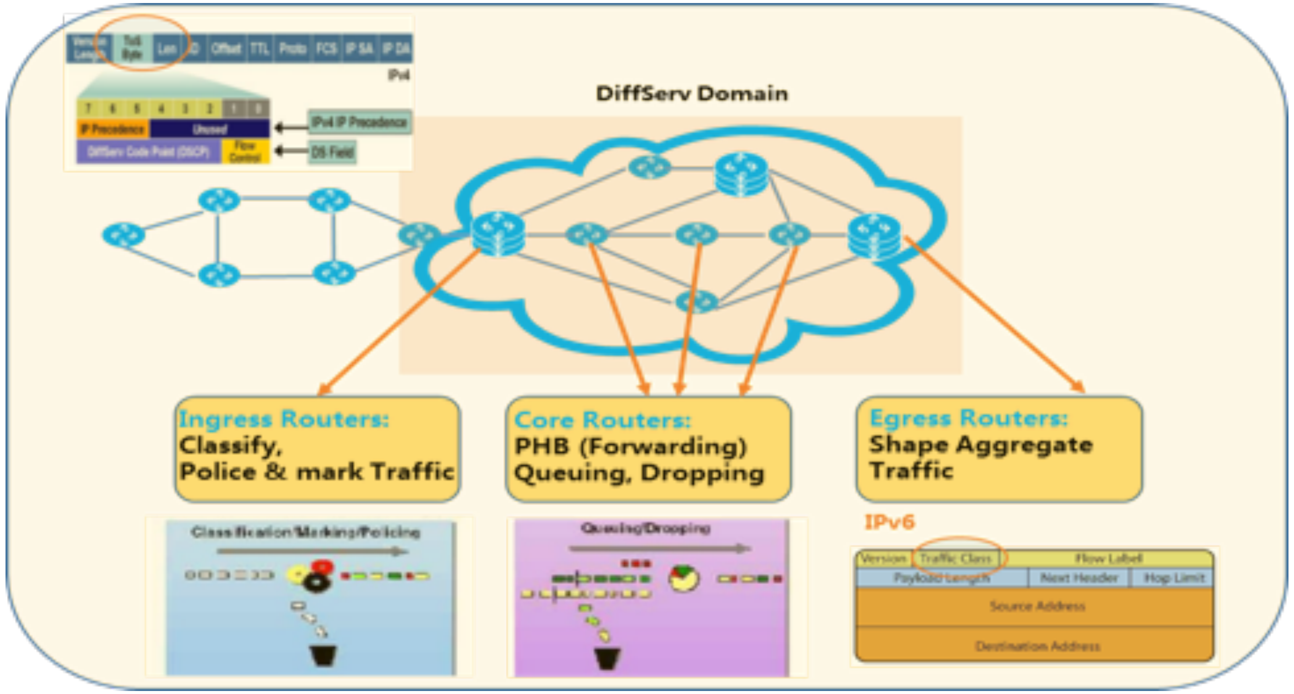
configuration nightmares

DiffServ in L2/L3 VPN / enterprise / DC
- How do you calculate class-weights ?
- Time dependent ?
- Four-level Hierarchical (DiffServ) QoS.

One off SP / content provider hacks.
- ACL for SP video
- Co-location of OTT content servers in SP network

Different IP / MPLS / Ethernet QoS
- Mapping nightmares, service/functionality loss

Per-flow “IntServ” services
- Great for limited private networks, but does not scale to metro networks hop-by-hop



Beyond Internet:

D& ifferentiated Services
Integrated Services

Analyze applications

Understand how network service: (throughput, loss, latency) impact user experience

Not even good MOS for AR/VR today!
MOS = Mean Opinion Score

Initial , intermittent buffering

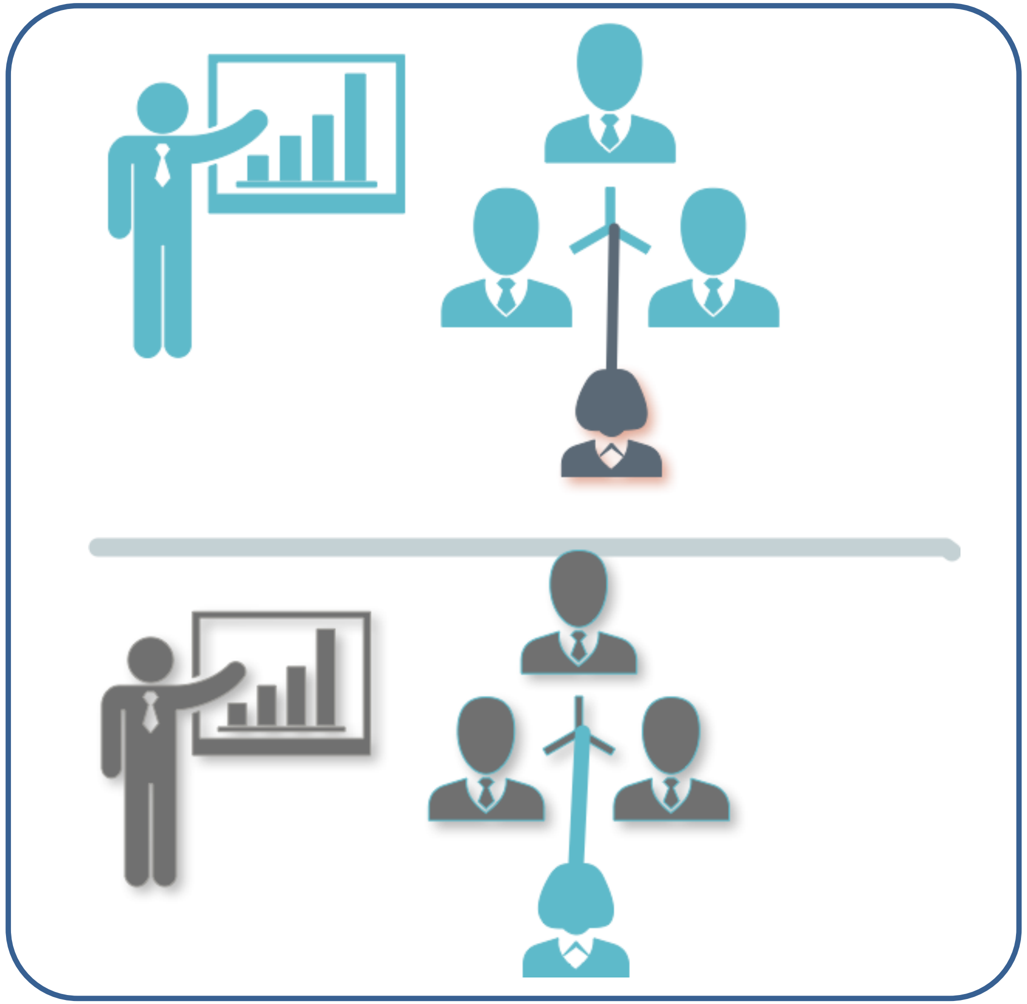
Probability / periodicity of non-congestion loss (NCL)
Random (bit /packet errors) vs. Burst (outage) loss

Ability to conceal NCL (codec dependent)
Ability to correct NCL (FEC, retransmissions)

Normal/Ideal quality (throughput)

Quality Variation by throughput variation
Frequency, fine-grained (concealed)

Congestion Loss (*YUCK*), Self-Friendly-ness



Use-case: Digital Actor Single point cloud holographic object in a real scene

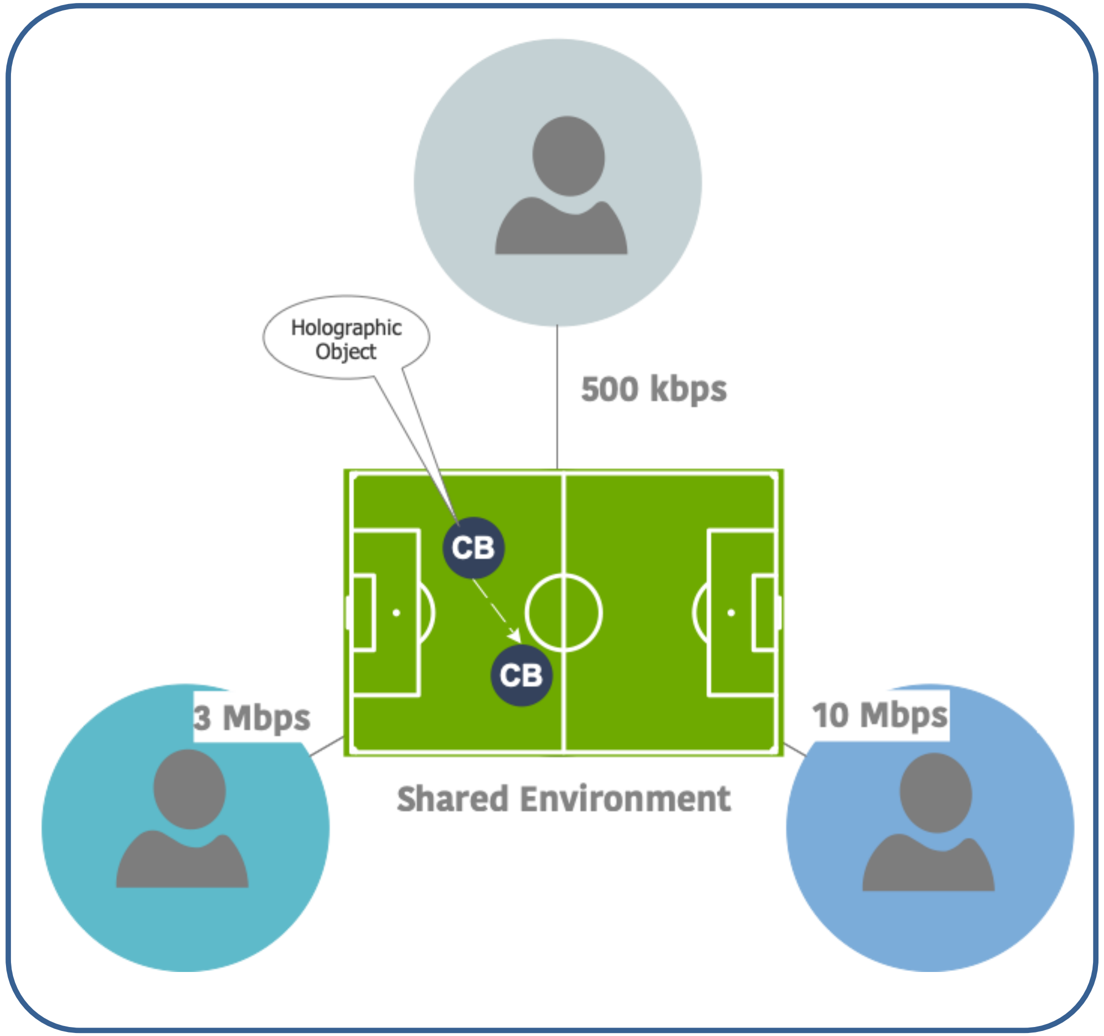
Coordinated Communications

Synchronized latency experience
"Why can player 3 shoot /kill faster?"
(lowest RTT to server)
In-network latency fairness
(vs. clients faking RTT)

Aggregate Network Resource Management
"Get N Gbps for 2 hours, split across parties as you like"

Scale optimizations with guarantees
Multicast: in-network replication to parties
Incast: Many to "virtual/cluster" one

Coordination may need to happen on many hope, or (depending on coordination function) on the edge of the network.



Use case: Synchronized Multi-Part Remote Collaboration

High-Precision Communications

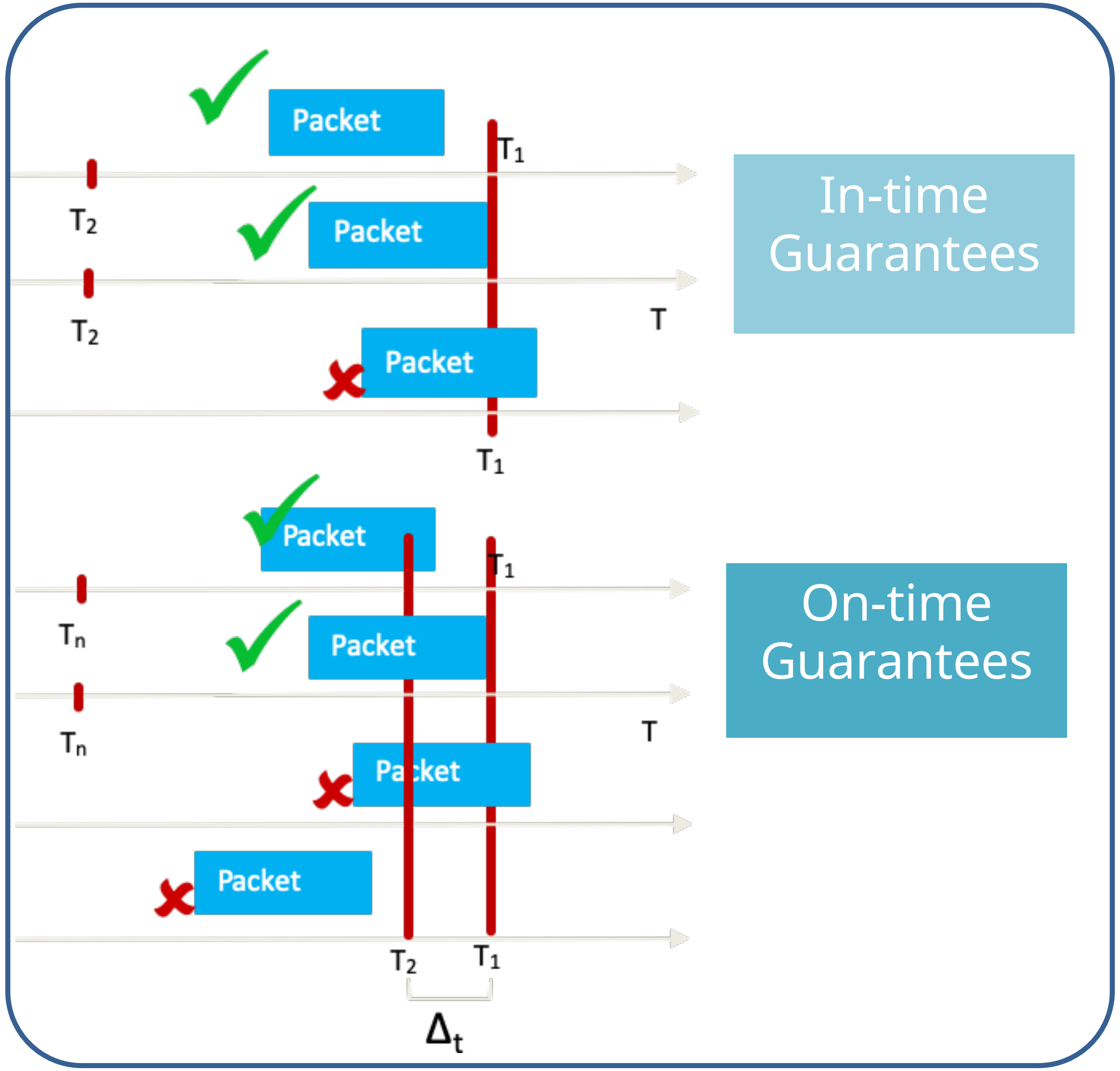
SLO (service level objective) based latency management

Let applications manage network latency (queuing, buffering)
End-to-end min..max latency control

Stateless == per-packet
Avoid scale limits of network flow-awareness

With/without (bandwidth) reservations
E.g: dynamic adjust throughput, but keep low-latency

On-time: Delay early packets
Reduces application side buffering / jitter



Use case: Latency sensitive streaming, control-loops

High-Precision Communications

Example: Flows with different path length/hops

Typical metro network "core" ring part with aggregation

Fair-latency for Flow 1 / Flow 2 ?

Flow 1 disadvantaged by physics (path latency, hops)
But can compensate with high precision communications:

Edge-coordination function:
Assign same minimum end-to-end latency to both flows

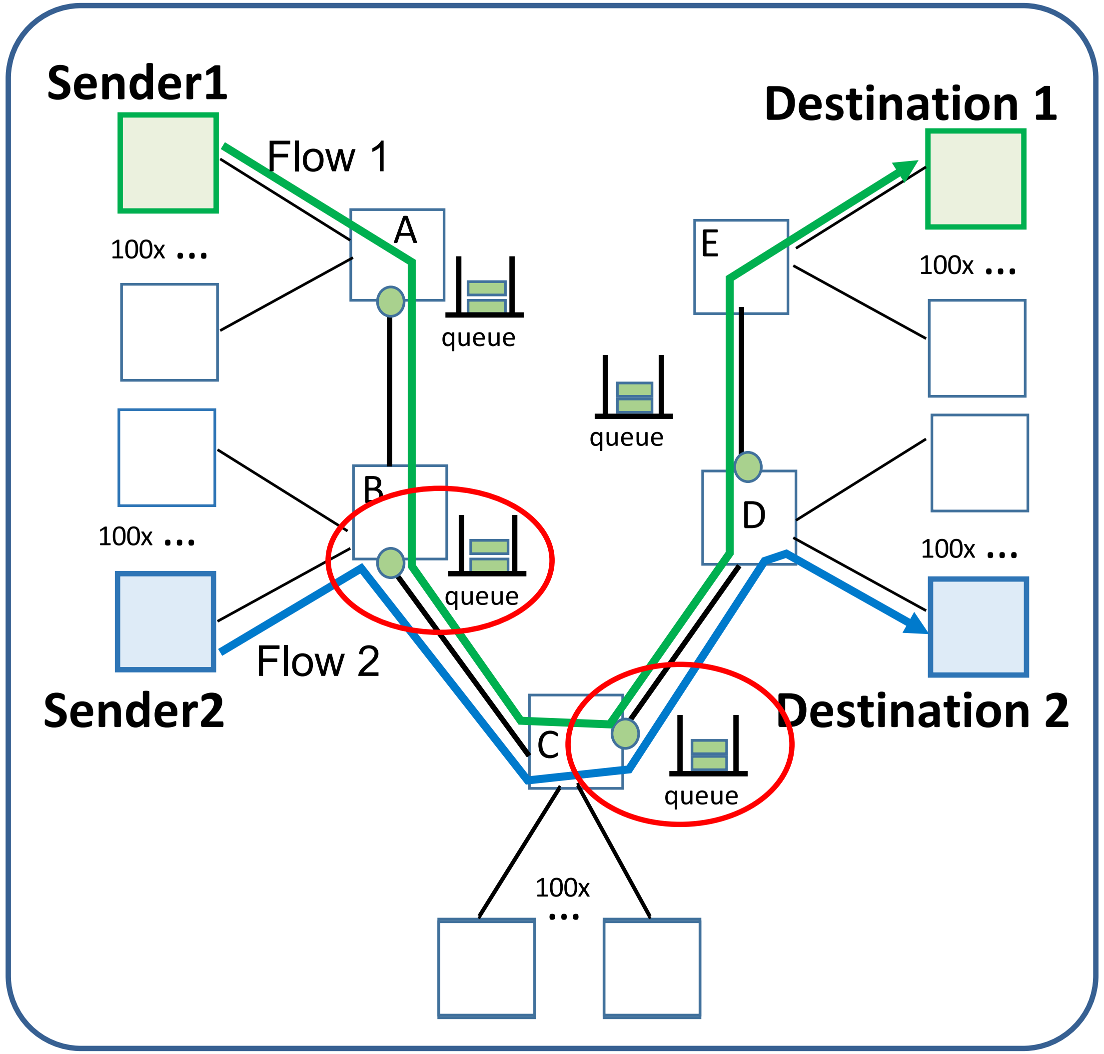
Somewhat larger than flow 1 physical propagation latency.

When there is no contention:
This just "delays" flow 2 packets in network to create fairness

Interesting under network contention/congestion

Intelligent latency aware queuing (node B, C) would know from BPP header that flow 2 packets have more time to spare in queuing…
And dequeue flow 1 packets faster

Path/end-to-end latency based automatic queuing priorities.



Use case: Latency sensitive streaming, control-loops

Qualitative Communications

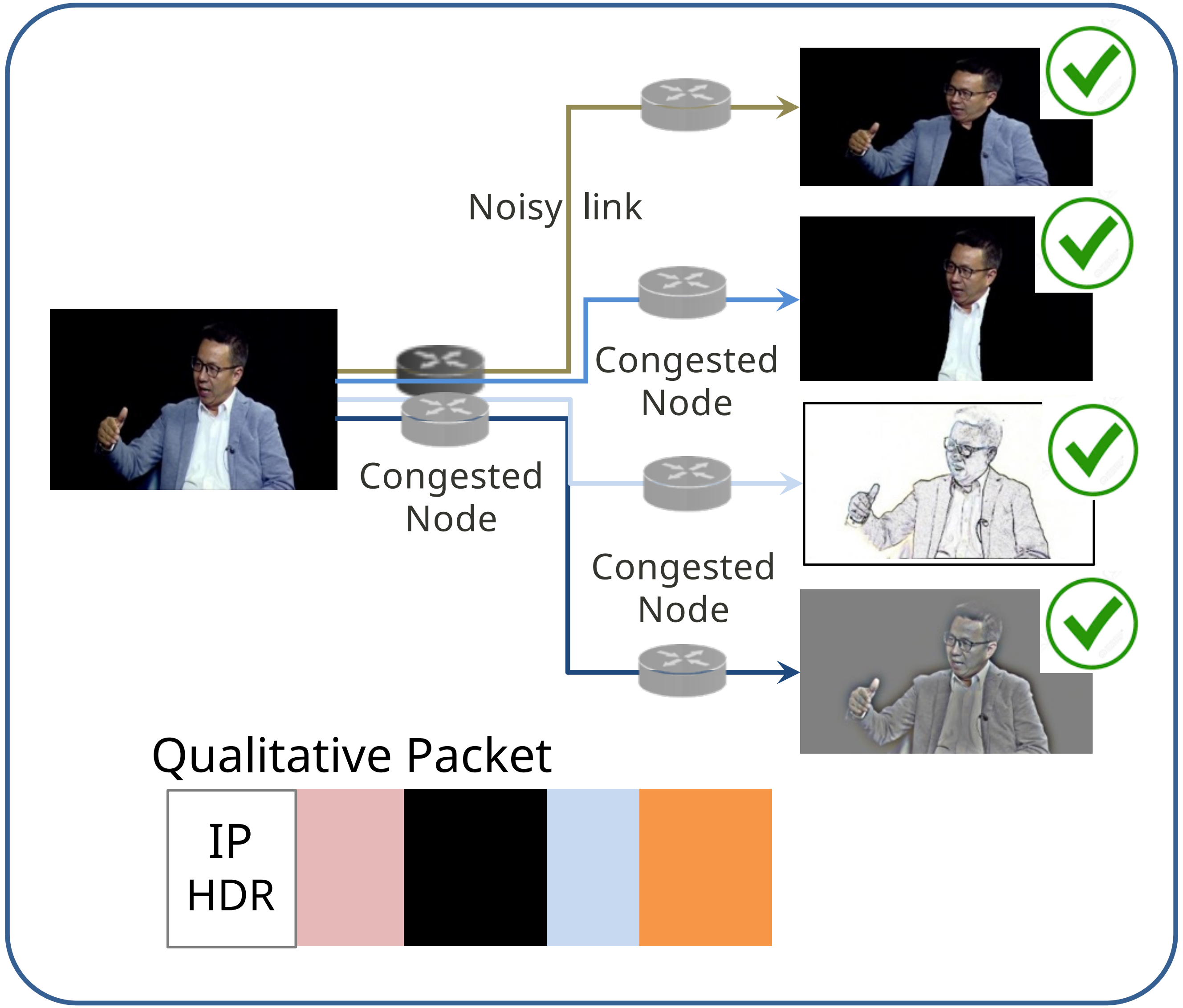
Layered/Object media codecs / transport
e.g.: MPEG SVC, ALC RFC5775

Network awareness of Quality
Mapping to different per-packet network resources
E.g.: congest/(discard) highest-quality level first

Redundancy:
L3 In-network redundancy (FEC) insert/remove

From flows to packets:
When bitrates increase but not packet rates
E.g.: Optical Switched Networks => huge packets

Packet-Chunks, tail-discard / congest / recover



Use case: Scalable Gradual management of quality

2020 State of Traffic Engineering

MPLS/RSVP-TE designed for **Traffic Engineering (TE)**

Traffic flow resource guarantees: bandwidth, latency, reliability
AC with path selection, FRR protection (sub 50ms), per-path/(hop) QoS

Complex, scaling limited for last two decades dominant use-case:

Capacity optimization

Maximize use of non-ECMP SP paths for best effort traffic

Segment Routing (SR-MPLS, SRv6)

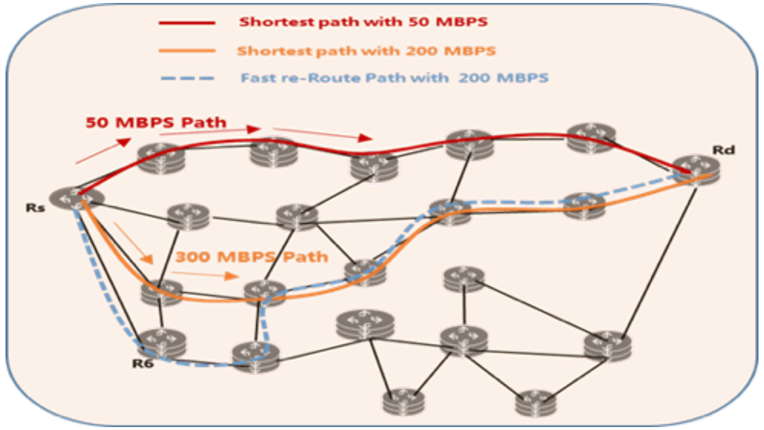
Revives 1981 “loose source-routing”. Ideal for capacity optimization.
Simplifies routing architecture: Only IGP (ISIS, OSPF), no LDP.
Rightfully replaces RSVP-TE for capacity optimization

Source-routing not ideal for real TE problems

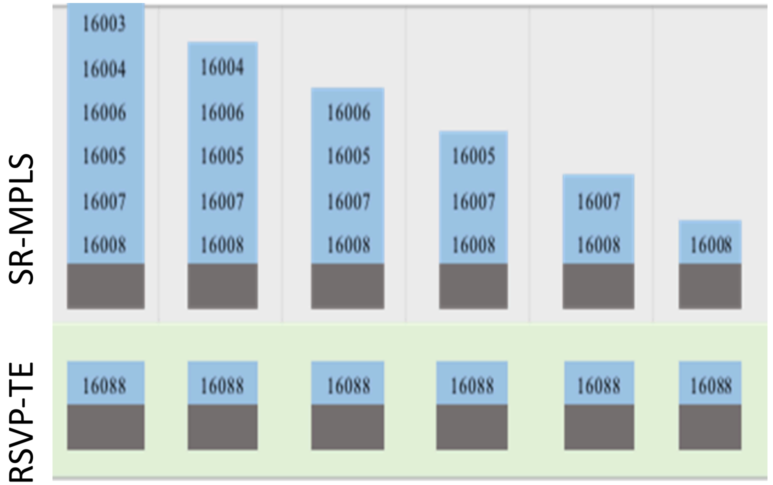
Header tax, undesired re-routing, per-hop-guarantees, monitoring
Difficult/impossible to apply to arbitrary forwarding planes

2030 Networks need to support real TE solutions

With similar optimization as SR did for capacity optimization



Non-equal-cost-path traffic steering



Per-hop header tax (MPLS) SR vs.
Per-flow state based solutions (e.g.: RSVP-TE)

Lossless networking and path steering

Live-live: “near zero loss”, no added latency

Send two copies of same traffic across “disjoint” (no shared failure) paths
Used since 1990th for financial, military, video (SMPTE since 2010th)

All redundant networks have >= 2 disjoint paths.
Value of critical traffic easily >= 2 * normal best effort traffic.

PREOF – Packet Replication Elimination Ordering Function

DetNet term for live-live solution. TSN also uses this.
Only working there in simpler topologies than those needed for 2030

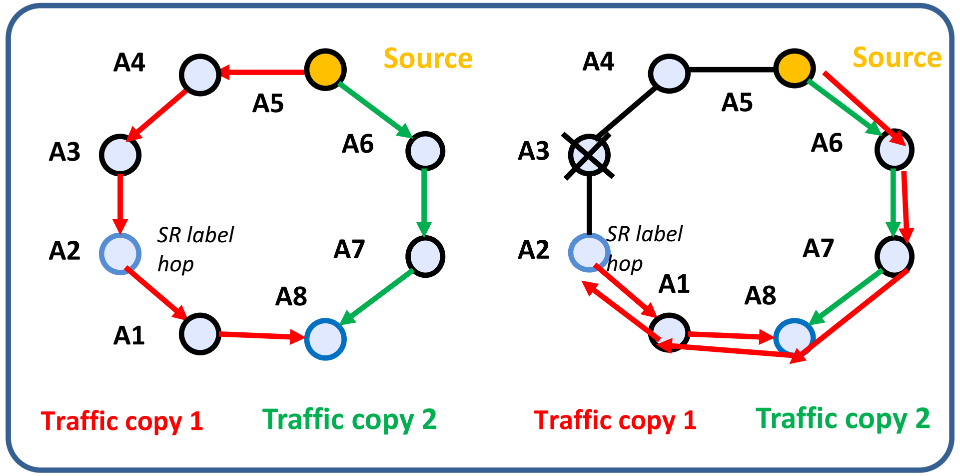
(SR) Loose path-steering issues

Unexpected / unnecessary rerouting: When one path fails, rerouting it undesirable – but unavoidable with loose paths.

Header tax redundant when per-hop state needed anyhow:
Strict latency management (DetNet, TSN) (“shaping state”)
Monitoring / Diagnostic state, MTU

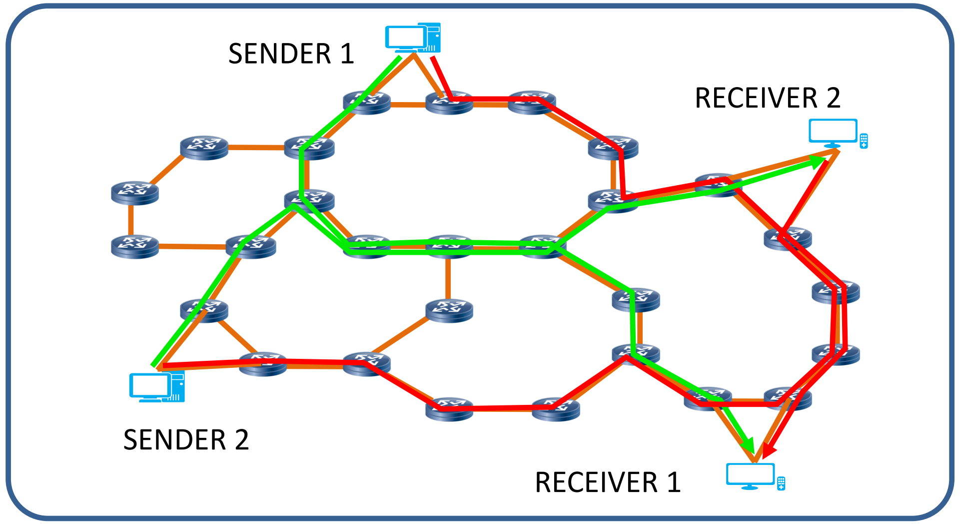
Real networks have long, complex paths

Especially in metro-space (up to 30 hops)
Special solutions for rings insufficient



Live-live transmission
In ring with SR

Rerouting under Failure



Sender sends same data twice, **red** and **green**
link/node failure interrupt only one copy of data

PPR – Preferred Path Routing

Novel Control-plane Signaling Architecture

Use IGP to signal TE information / replace RSVP-TE
PCE / AC (Path Computation Engine / Admission Controller) to network
Centralized PCE/AC or decentralized (edge nodes)

Unified for current and future forwarding planes

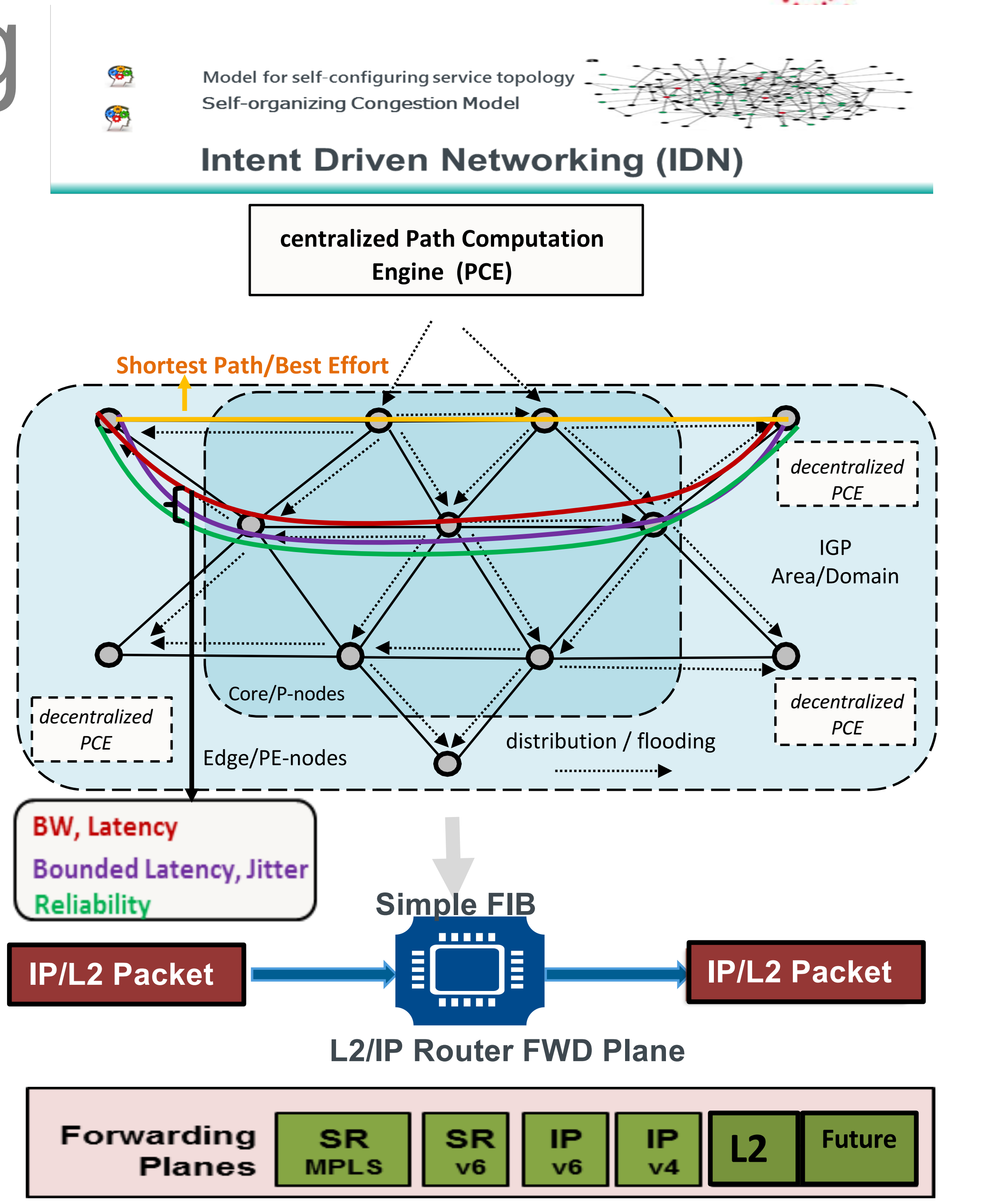
Strict/loose forwarding entries: IPv4, IPv6, SR-MPLS, Ethernet, future
PPR-ID: Identifier for a path == forwarding plane address
/32 for IPv4, /128 for IPv6, /20 for SR-MPLS,...
Per-hop QoS setup as necessary/desired
SW-update-only for existing forwarding planes

Orthogonal / combinable with SR

PPR when hop-by-hop state required.
SR for lightweight capacity maximization.

All the benefits of RSVP-PE without the limits

Efficient signaling via IGP instead of heavy hop-by-hop on-path
RSVP-TE only for MPLS forwarding plane.



Fully backward compatible + extensible, lean path routing information on the Wire

PPR Graphs – scaling state

Point 2 point paths N^2 scaling issue

Worst case: network wide path optimization N source to N dest.

Full mesh NxN paths (RSVP-TE) . N^2/k path across single node

SR: N*l source-route stacks in headers on each source node + N' SID state

k = number of parallel path in network, l = average number of hops in network

PPR can signal "Graphs" from PCE to network

Aggregate per-hop-state whenever possible (determined by PCE)

No change in forwarding plane. Still simple "destination" forwarding entries

PPR-Tree

(Sub)set of senders in network to one destination.

Multiple Trees to same destination for different policies easily possible.

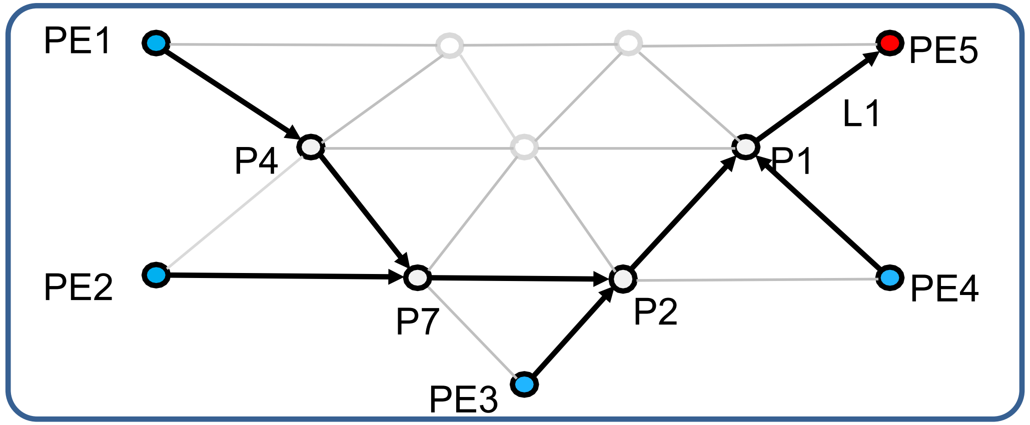
Set of 1 tree for each destination like a "Flex-Topology", but with paths calculated by controller – easier than reverse-engineering SPF metrics.

PPR "Forest"

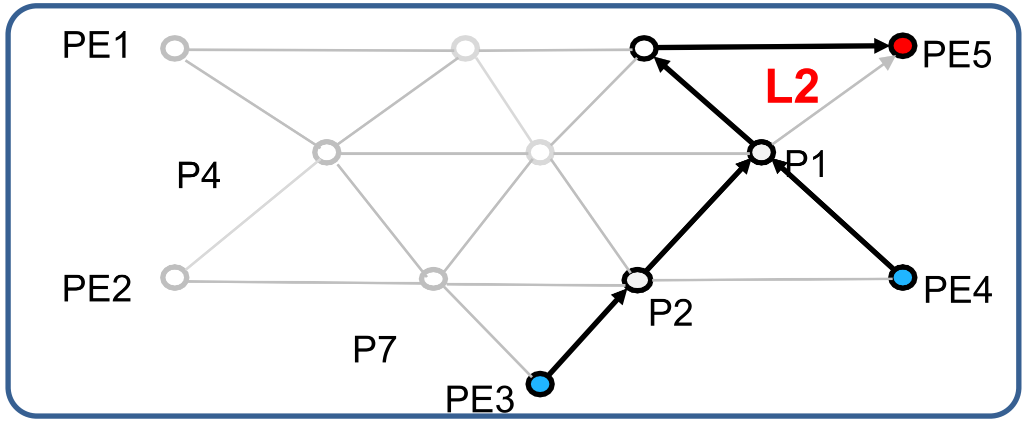
Shared graph for multiple destinations.

Allows for minimum amount of state in network, e.g.: for live-live services

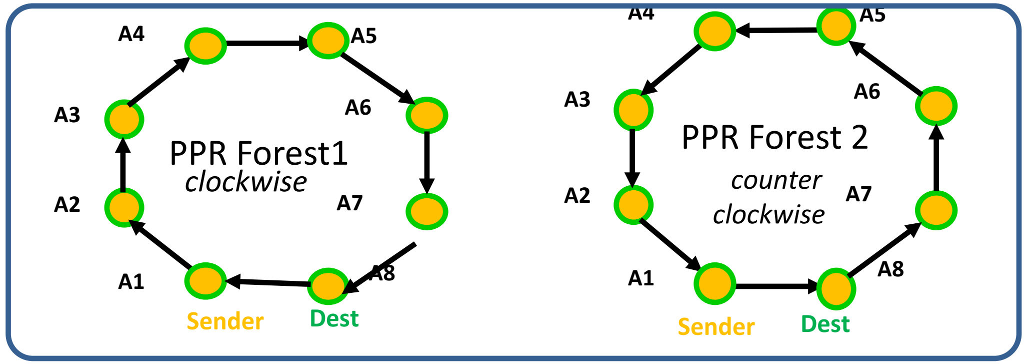
Controller calculated path == better live-live paths than distributed MRT live-live



PPR Tree: all possible sender to destination PE5



Better PPR Tree for some subset of sender



2 PPR Forests representing any path in ring topology

Technologies for 2030 Networks

WHY, WHERE, WHO, HOW, WHAT

GREAT SUCCESS !!!!!

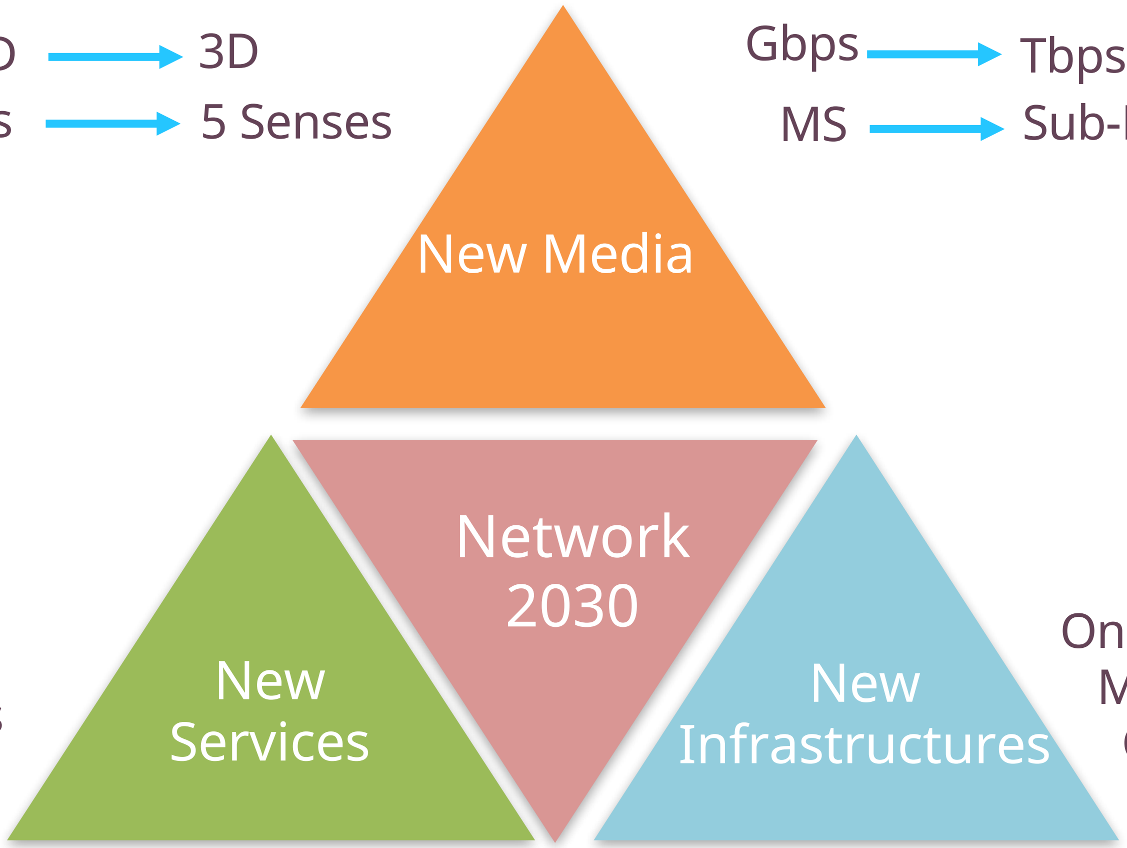
But wait… there is more (to do)

Vision Network 2030

Holographic Society and Industry

2D → 3D
2 Senses → 5 Senses

Gbps → Tbps
MS → Sub-MS Latency



Next wave of Internet Innovations

OneNet → ManyNets
Mbps → Gbps/Tbps Access
Computing in Networks

Lossless Networking
High-Precision Communications
Holographic Type Communications

Converged Terrestrial and Space
Fidelity and Trust
Federated Networks

Publications and Talks

Concepts

- A New Way to Evolve the Internet, A Keynote Speech at IEEE NetSoft 2018, Montreal, Canada, June 2018
- What if we reimagine the Internet?, A Keynote Speech at IEEE ICII 2018, Bellevue, Washington, USA, Oct 2018

Framework and Architecture

- A New Framework and Protocol for Future Networking, ACM Sigcomm 2018 NEAT Workshop, Budapest, August 20, 2018
- New IP: Design for Future Internet with New Service Capabilities Envisioned, IEEE ICC Industry Tutorial, 2019

Market Drivers and Requirements

- Towards a New Internet for the Year 2030 and Beyond, ITU IMT-2020/5G Workshop, Geneva, Switzerland, July 2018
- Network 2030: Market Drivers and Prospects, ITU-T 1st Workshop on Network 2030, New York City, New York, October 2018
- Next Generation Networks: Requirements and Research Directions, ETSI New Internet Forum, the Hague, the Netherlands, October 2018
- The Requirements for the Internet and the Internet Protocol in 2030, ITU-T 3rd Workshop on Network 2030, London, Feb 2019

New Technologies

- Preferred Path Routing – A Next-Generation Routing Framework beyond Segment Routing, IEEE Globecom 2018, December 2018
- Flow-Level QoS Assurance via In-Band Signaling, 27th IEEE WOCC 2018 , 2018
- Using Big Packet Protocol Framework to Support Low Latency based Large Scale Networks, ICNS 2019, Athens, 2019

Use Cases and Verticals

- A Novel Multi-Factored Replacement Algorithm for In-Network Content Caching, EUCNC 2019, Valencia, Spain
- Distributed Mechanism for Computation Offloading Task Routing in Mobile Edge Cloud Network, ICNC 2019, Honolulu, USA
- Enhance Information Derivation by In-Network Semantic Mashup for IoT Applications, EUCNC 2018, Ljubljana, Slovenia
- Latency Guarantee for Multimedia Streaming Service to Moving Subscriber with 5G Slicing, ISNCC 2018, Rome, Italy

References

- Holographic content considerations methods for efficient data transmission and content creation methodologies
- Point Cloud Compression in MPEG MP20 Workshop Hong kong 2017
- Keynote: the near future of immersive experiences: where we are on the journey, what lies ahead, and what it takes to get there.
- Architectures and codecs for real-time light field streaming journal of imaging science and Technology , January 2017
- A Dynamic Compression Technique for Streaming Kinect-Based Point Cloud Data (2017 International Conference on Computing, Networking and Communications (ICNC): Multimedia Computing and Communications)
- Technical White Paper on Mobile Bearer Network Requirements for Mobile Video Services.
- On the Support of Light Field and Holographic Video Display Technology, Light Field Lab, Inc., San Jose, CA. “The road to immersive communication,” Proceedings of the IEEE, vol. 100, Apr. 2012.

Thank You

Comments, Curious, Questions?

Toerless.Eckert@futurewei.com