Problems in the Forwarding Layer

draft-bryant-arch-fwd-layer-ps-00

Stewart Bryant, Uma Chunduri, Toerless Eckert and Alex Clemm

Content

- Problem statement draft overview
- New Application -> New Needs, Services & Capabilities
- Emerging Deployment Models
- Forwarding Layer Shortcomings
- Next Steps

Content

- Problem statement draft overview
- New Application -> New Needs, Services & Capabilities
- Emerging Deployment Models
- Forwarding Layer Shortcomings
- Next Steps

What this Problem Statement does

- Identifies an emerging set of new network and transport requirements that exceed beyond the ability of best effort protocols.
- Makes a case for evolving the forwarding layer to address the new E2E needs to support 5G, B5G, 6G, Network 2030 etc.
- Elaborates on problems/limitations in the forwarding paradigms based on IP and its related protocols.
- Provides the first step in developing a common description and appreciation of the problems in the context of IP as the core of forwarding layer.

Table of Contents

Content

- Problem statement draft overview
- New Application -> New Needs, Services & Capabilities
- Emerging Deployment Models
- Forwarding Layer Shortcomings
- Next Steps

New Applications
drive new Needs

- Beyond 5G will require URLLC beyond the radio edge and include backhaul and core.
- Increasing number of mission-critical applications (tele-operation of equipment, tele-driving, tele-medicine, tactile) are not tolerant of loss and cannot afford retransmissions due to latency constraints
- AR/VR and holographic networking applications require high traffic volumes with predictable latency guarantees, not just optimization.
- Space-terrestrial networks with mega-clusters to ensure connectivity in airspace, over oceans, and beyond exceed the mobility capabilities of existing technologies
- Resource-constrained IoT devices are hampered by the constraints of IPv6, such as address lengths

New Needs (1)

- 5G/B5G Services – end-to-end support
  - URLLC and mMTC are addressed in the radio, but not yet in the backhaul.
  - The fixed network packet layer needs be developed to keep up with the new radios.
- Latency precision category of service:
  - Industrial Internet, industry automation, V2X, Tactile internet etc.
- Volumetric Data transmission
  - Immersive AR/VR media transmission - Predictable throughput (over bandwidth) is necessary for Motion to Photon latency requirements E2E 20 ms for realistic world visualizations.
  - Media evolution from AR/VR to Holograms pushes capacity to 10s of gigabits.

New Needs (2).

- Evolution in Reachability Paradigms
  - Integration of multiple space-broadband constellations with terrestrial networks. To access satellite-networks, geo-coordinate awareness for reachability.
  - Edge compute service data and attributes are location specific. Network layer location awareness for service delivery.
  - ManyNets (Network of many largescale non-transit networks: Sigfox, public clouds), require interworking of disparate network types.
- IoT is stressed by the constraints of IPv6.

New Services

- High-Precision Communications Services
  - Precise service-level objectives (SLOs) for e2e latency, loss
  - In-time (not to exceed) & on-time (not to exceed **and** no sooner than permitted)
  - Coordinated services (matching & interdependent service levels across flows)
- Holographic and Tactile Services
- Qualitative Communications Services
  - Selective suppression of retransmission and dropping of payload portions

New Capabilities

- Manageability:
  - High-precision service assurance
  - Not best-effort service assurance as an afterthought
- High Programmability and Agile Lifecycle:
  - Rapid adaptation to new contexts, applications
- Security and Trustworthiness:
  - Authenticate and authorize traffic without impacting SLOs
- Resilience:
  - Robust to perturbations while avoiding loss or retransmits to meet mission-critical needs
- Privacy-sensitive:
  - Balance need for user privacy with legitimate operational needs.

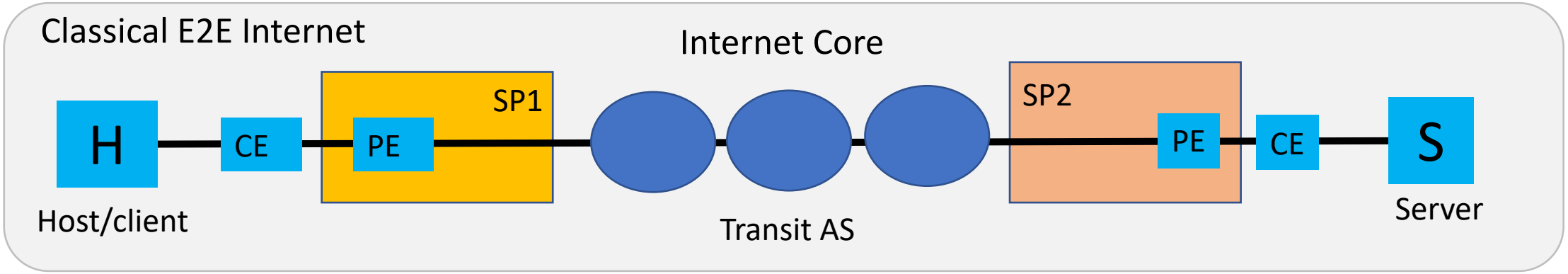
Content

- Problem statement draft overview
- New Application -> New Needs, Services & Capabilities
- Emerging Deployment Models
- Forwarding Layer Shortcomings
- Next Steps

Emerging Deployment Models

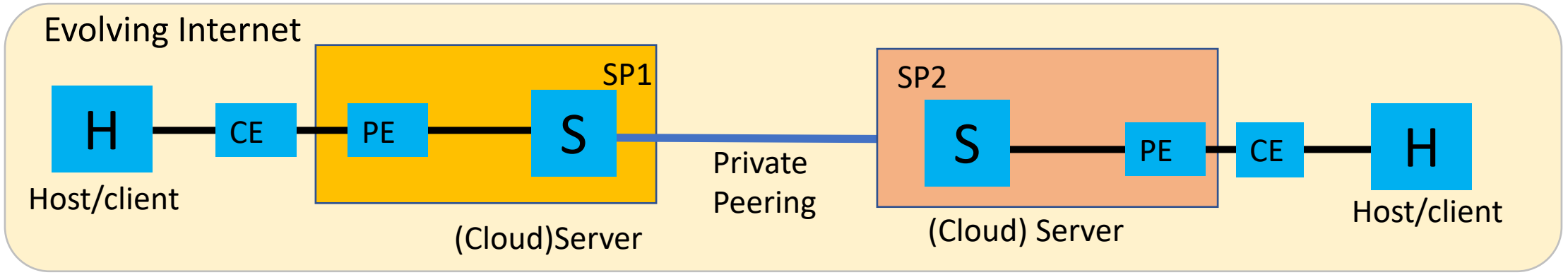
- End to End over the Internet backbone is a minority of traffic
- Commercial deployments do not operate the way they used to.
- Application trajectory to sit in (protected) servers a few hops from the user.
- Applications are becoming self-contained and use their own stack which is tunneled over UDP/IP to the server.

Evolving Network Architecture Facilitates Change



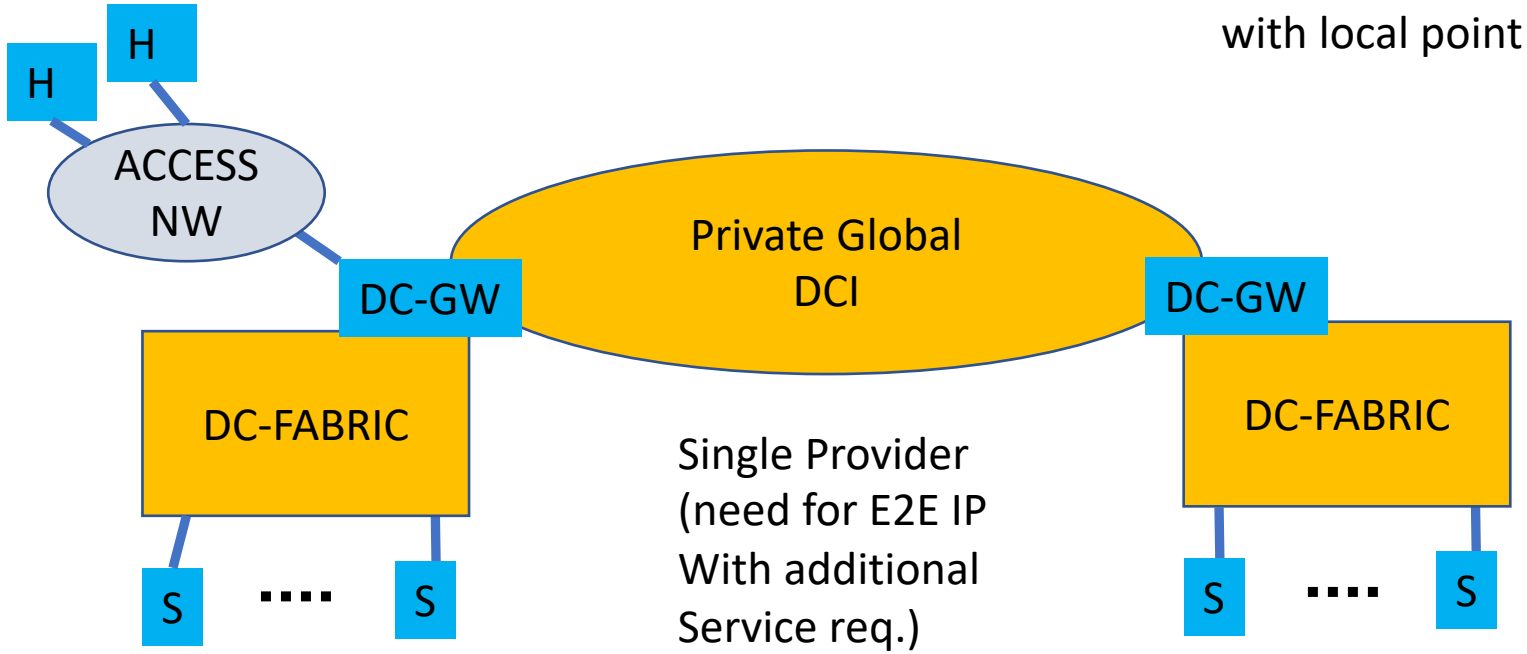
The classical E2E model with connectivity via the internet core is being replaced by a model with edge compute and private interconnections.

This significantly reduces the problem of rolling out new network services.



Changing Fixed Access Models (1 or 2 Providers)

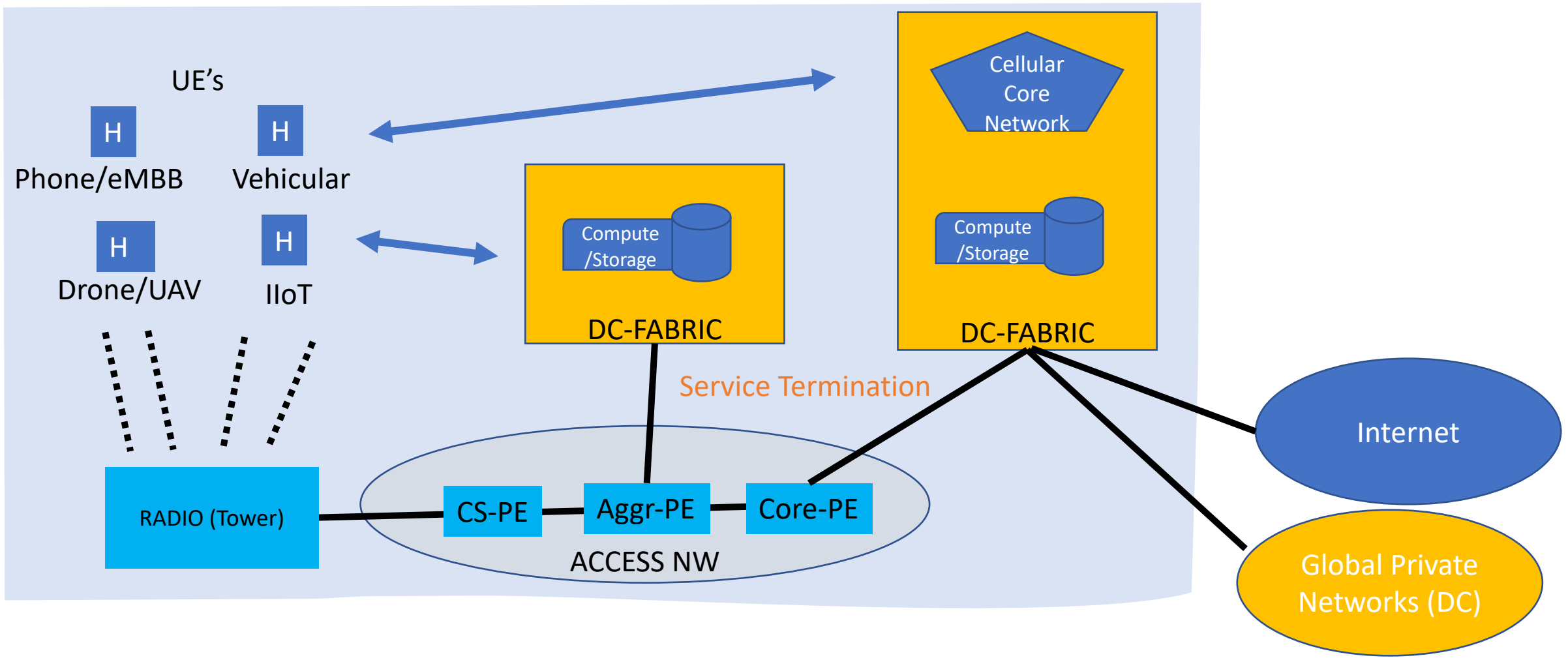
Access network and single global provider with local point of presence.



Single Provider (need for E2E IP With additional Service req.)

Single “Underlay” provider E2E

for 4G/5G network (Cellular/Access Networks)



Need for E2E IP With stringent underlay needs

Content

- Problem statement draft overview
- New Application -> New Needs, Services & Capabilities
- Emerging Deployment Models
- Forwarding Layer Shortcomings
- Next Steps

Addressing and Packet Formats

Forwarding Layer Shortcomings (1)

- Addresses are:
  - Too long (eg.: IoT, LowPAN, ..)
  - Too short (eg.: content semantics, SR network programming, etc)
  - Always from the same address family
  - Limited semantics
- Packet format / extensibility / flexibility:
  - Only IPv6 going forward(*): RFC8200 very challenging to meet new needs
  - Only few additional types of extension headers can be added
  - But not extensible enough for the network needs of the new application
  - Modifying packet header metadata (beyond ECN, segment-id ?)
  - For example multiple, lightweight per-path network guidance to a packet.

(*) https://www.iab.org/2016/11/07/iab-statement-on-ipv6/

Beyond Best Effort (BBE)

Forwarding Layer Shortcomings (2)

- ‘QoS’ services are limited
  - Only in limited (“controlled”) domains.
  - Not end-to-end and predictable (throughput, latency)
  - Not granular (with DiffServ)
  - Not accurate and guaranteed (even with low-latency-low-loss-services – L4S)
  - Programmability
- DetNet is limited by use of current data planes
  - Scoped for limited to private networks amongst a set of coordinated endpoints.
  - Missing full forwarding plane for IP (e.g. PREOF).
  - Service model for ”inelastic” traffic only (CBR/TSPEC) traffic
- Fairness and differentiated elastic traffic unresolved
  - Internet Fairness == “Do not kill other flows” (not good enough)
  - Component proposal exist: NADA weighted CC, PI^2, Multi-TSPEC, ...
  - But overall solution architecture missing.

User-Network Interface (UNI)

Forwarding Layer Shortcomings (3)

- UNI required to enable high-value applications
  - Application needs to ask more from the network than it can today, and network needs to know more than it does today to help.

- Explicit User/Application <-> signaling missing
  - BBE services (traffic specific Service-Level-Objectives)
  - High Precision Network state feedback (beyond ECN framework)
- Current UNI protocols designed for execution in control plane CPU
  - Problem: Control plane not keeping up with speed of data-plane (scale, latency)
  - Design future UNI protocols for execution in data-plane
  - In data-plane signaling, e.g.: [I-D.han-tsvwg-ip-transport-qos] [I-D.han-tsvwg-enhanced-diffserv]
- Security model missing
  - No trust model: User <-> Network (E.g.: application assumed to only trust DC server app)
  - No comprehensive authentication: User <-> Network
  - Solvable for limited (collaborating) domains – including future (Internet) Access Networks.

Content

- Problem statement draft overview
- New Application -> New Needs, Services & Capabilities
- Emerging Deployment Models
- Forwarding Layer Shortcomings
- Next Steps

Where To Go From Here

- Need converged networks with both BE and BBE
- Need an architectural level discussion
  - About how to address the new requirements.
  - How address (inter)network changes.
- Why not make piecemeal fixes ?
  - X number of services, at-least X or more protocols, that much effort in design, development and standardization

  ➔ Integration then has increased complexity
- Or take a step back to assess the situation
  - Holistically gather and study forwarding requirements

Let’s Talk

- The applications and the network deployments are changing.
- The economics are changing.
- This is an area where the IETF needs to lead.
- Feedback and contributions to this work are welcome.

  Agreements or disagreements?
  What have we missed? Is there anything we have overstated?
- Please join us!

The End