

# 5G Edge Services Use Cases for CATS WG

draft-dunbar-cats-edge-service-metrics-01

Linda Dunbar: [ldunbar@futurewei.com](mailto:ldunbar@futurewei.com)

Kausik Majumdar: [kmajumdar@microsoft.com](mailto:kmajumdar@microsoft.com)

Gyan Mishra: [gyan.s.mishra@verizon.com](mailto:gyan.s.mishra@verizon.com)

Haibo Wang: [Rainsword.wang@huawei.com](mailto:Rainsword.wang@huawei.com)

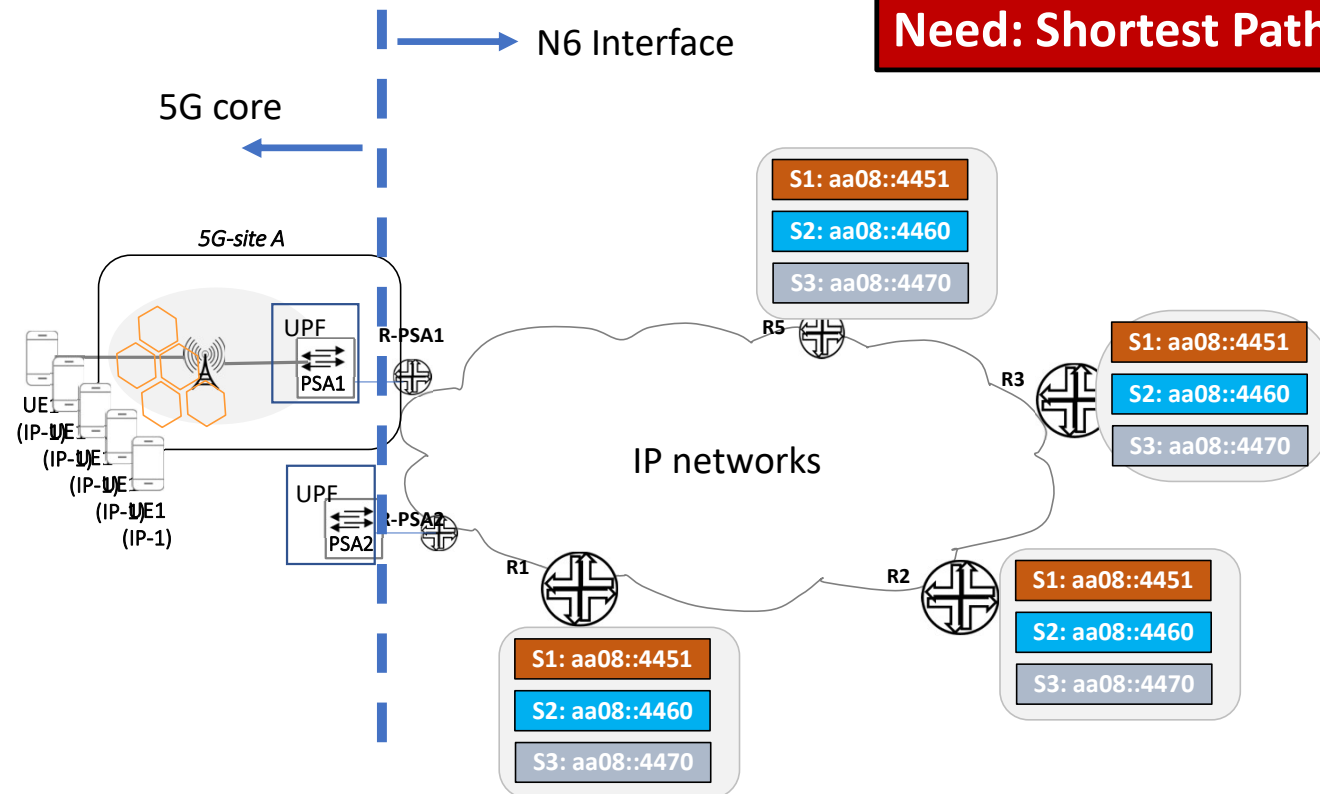
Haoyu Song: [Haoyu.song@futurewei.com](mailto:Haoyu.song@futurewei.com)

**IETF 117 San Francisco**

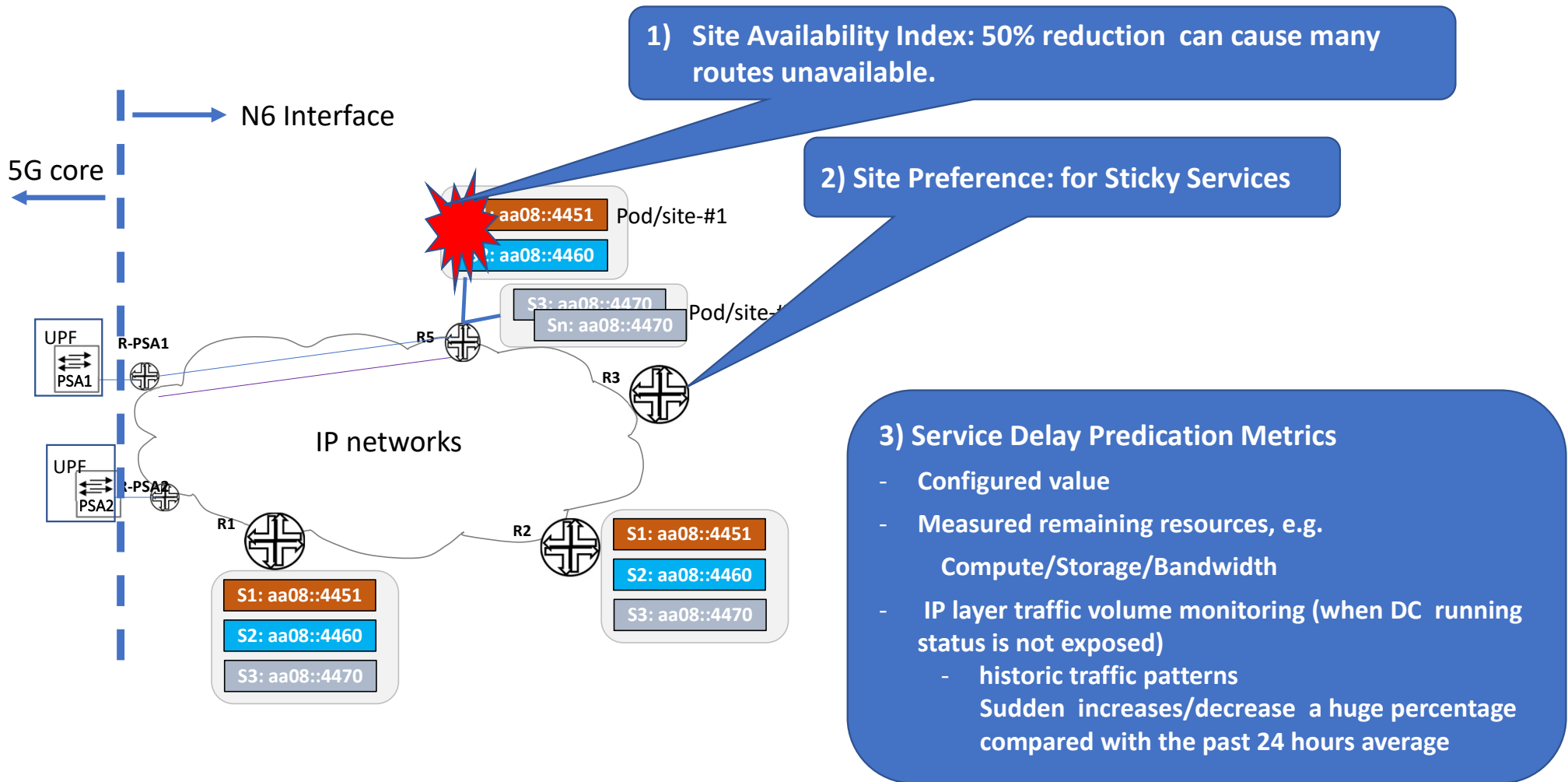
# 5G Edge Computing Background

3GPP TS 23.548 V18.1.1

**Goal: Ultra Low Latency services**  
**But Fewer hops != Lower latency**  
**Need: Shortest Path + Fastest Service processing**



# 5G Edge Service Metadata



1) Site Availability Index: 50% reduction can cause many routes unavailable.

2) Site Preference: for Sticky Services

3) Service Delay Predication Metrics

- Configured value
- Measured remaining resources, e.g. Compute/Storage/Bandwidth
- IP layer traffic volume monitoring (when DC running status is not exposed)
  - historic traffic patterns
  - Sudden increases/decrease a huge percentage compared with the past 24 hours average

# Integrate Network Delay with Service Delay Predication

Exemplary algorithm: to compare the cost between Site-i vs. Site-j:

- Service Delay Predication (SerD-i) value,
- Capacity Availability index (CP-i)
- Preference Index (Pref-i), and
- network delay [NetD-i].

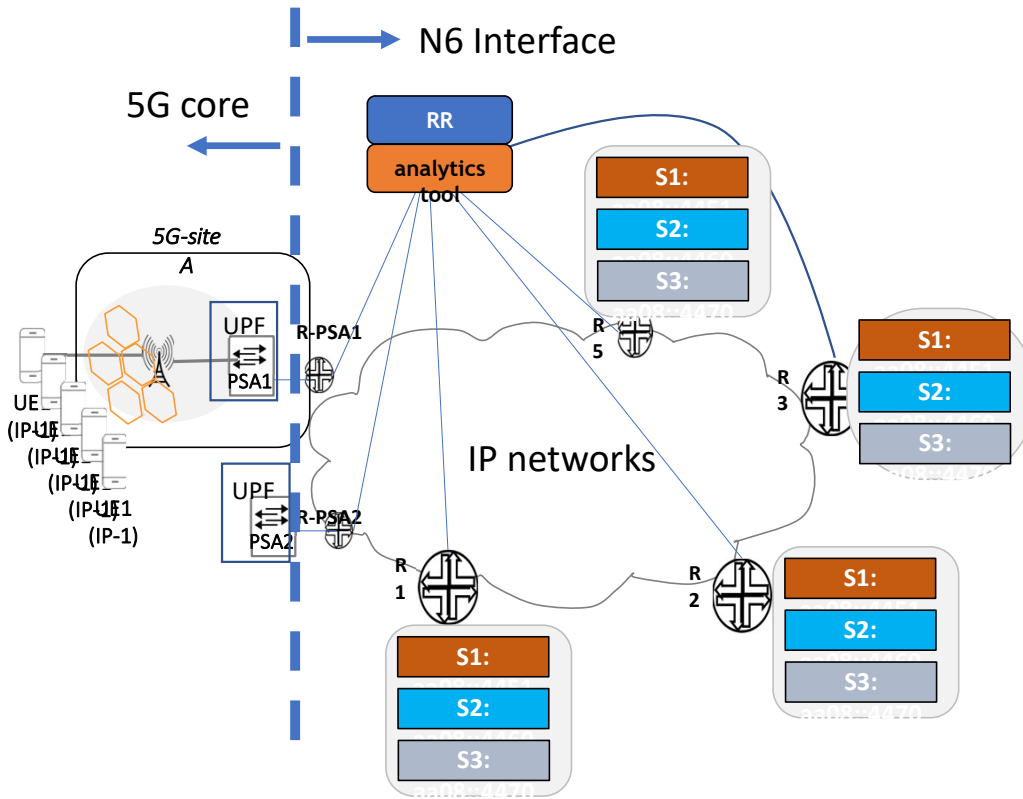
$$\text{Cost-i} = \min\left(w * \left(\frac{\text{SerD-i} * \text{CP-j}}{\text{SerD-j} * \text{CP-i}}\right) + (1-w) * \left(\frac{\text{Pref-j} * \text{NetD-i}}{\text{Pref-i} * \text{NetD-j}}\right)\right)$$

w: Weight, a value between 0 and 1. If smaller than 0.5, Network latency and the site Preference have more influence; otherwise, service delay and capacity have more influence.

When comparing metrics from Site-j with itself, the value from the algorithm is 1.  
Cost-i >1 indicates Site-i costs more than Site-j.

Cost-i <1 indicates Site-i costs less than Site-j.

# Using BGP to distribute Edge Service Metadata



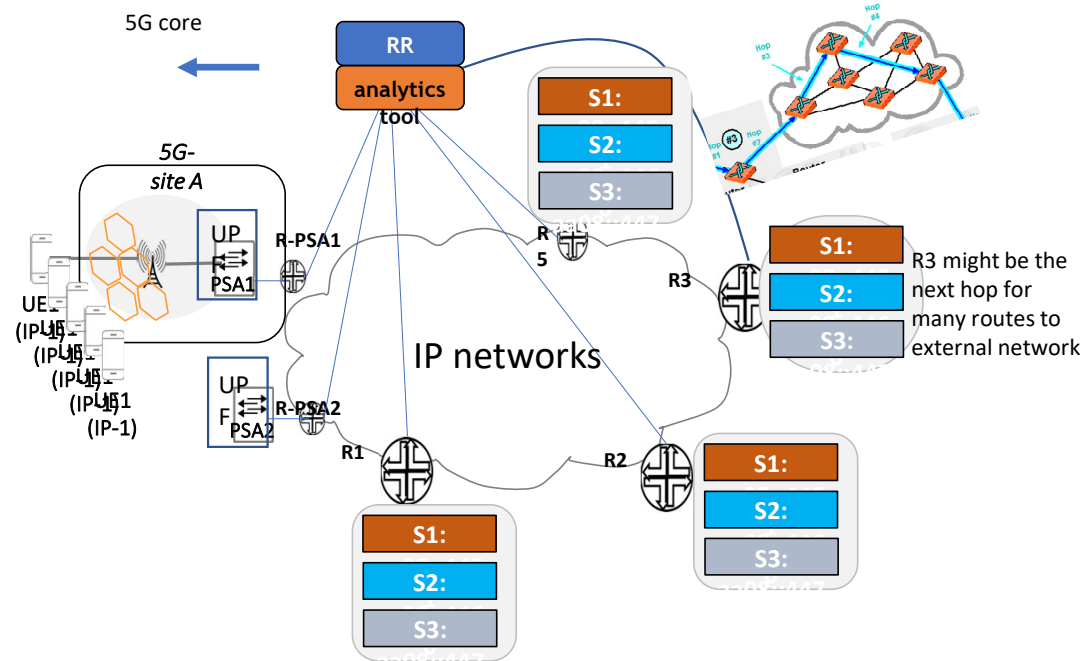
## Scope of Service Metrics Advertisement

- Each ultra-low latency EC service might be requested by a small group of UEs → Egress only needs distribute to a small set of ingress nodes
- BGP RT Constrained Distribution [RFC4684] with RT being the interested Service ID (IP prefixes)

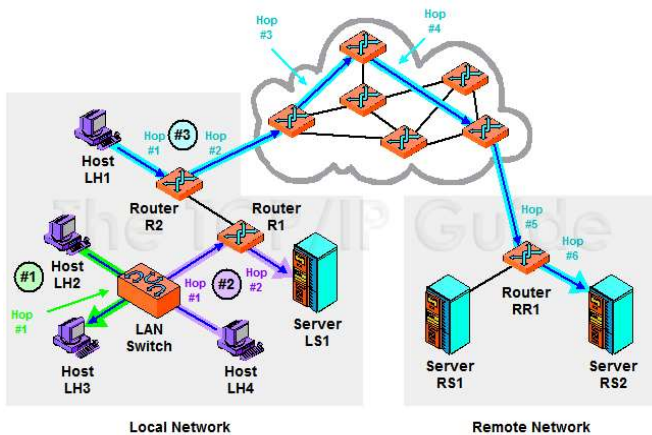
<https://datatracker.ietf.org/doc/draft-ietf-idr-5g-edge-service-metadata/>

# One Egress is the Nexthop for Many Routes

- only a subset of routes with the egress router as Next Hop are considered as Edge Services.
- The Metadata should not impact other routes that have the egress router as their Next Hop
- One egress node can be the next hop for many routes that are NOT considered as the Edge services. Therefore, should not be impacted when the site attached to the egress changes its Availability Index



# Site Availability Index vs. Changing weight to Nexthop



a single changes to the metric or weight to one Next Hop can trigger changes to all the routes that have the node as their Next Hop

## Next Step:

- Solicit feedback