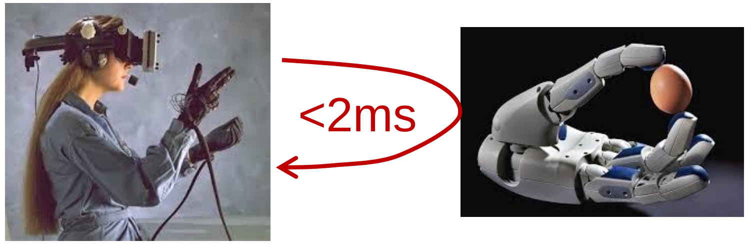
Service Assurance for High-Precision Networks

6th ITU Workshop on Network 2030
Lisbon, Portugal, 14 January 2030

Alexander Clemm alex@futurewei.com

The rise of
high-precision network applications

- Haptic Communications & Tactile Internet
  - Tele-operation of machinery, actuators, tele-surgery
  - End-to-end latency < 5 ms
- Industry 4.0
  - Long-distance industrial control and cyber-physical systems
  - Deterministic, time-guaranteed services
- VR/AR → Holographic-Type-Communication
  - New interaction models, media, training/education, entertainment
  - Effective compression leveraging user interactivity
    requires low latency guarantees even for canned content



Tactile Operator (Client)

Actuators & Sensors (Server)

Image references:
http://www.boonvr.com/data/blog/2017/12/19/history-of-vr/
https://pixels.com/featured/pr2-robot-hand-holding-an-egg-patrick-landmann.html
https://disruptionhub.com/innovation-industry-4-0/
https://themarketfact.com/2019/05/22/global-holographic-display-market-2019/

Characteristics of High-Precision Network Applications

- Stringent SLOs
  - Extremely low latencies coupled with high bandwidth + low loss
    - e.g. Tactile feedback: o(<2ms) round-trip latency
    - Holographic-Type Communications: Gbps→Tbps, o(30 ms) round-trip for user interaction-based optimization schemes
- No graceful degradation
  - Missed SLOs may not merely imply lowered QoE but lead to complete breakdown
    - Examples: loss of illusion of haptic control
  - Dysfunction of the network service as a result
- Mission-Criticality
  - Cannot tolerate occasional breakdowns (even if rare)
    - Remote operation of machinery – compare with 737 Max
  - Guarantees and validation beyond "best effort", "optimization", "priorization"

Image reference: https://www.nbcnews.com/news/us-news/southwest-pilots-union-sues-boeing-saying-it-lied-about-737-n1063561

Technical implications

- Network and protocol engineering:
  engineer networking services to be high-precision by design
  - But: designs can fail, components can fail, unexpected occurrences can happen, engineering assumptions may not always hold

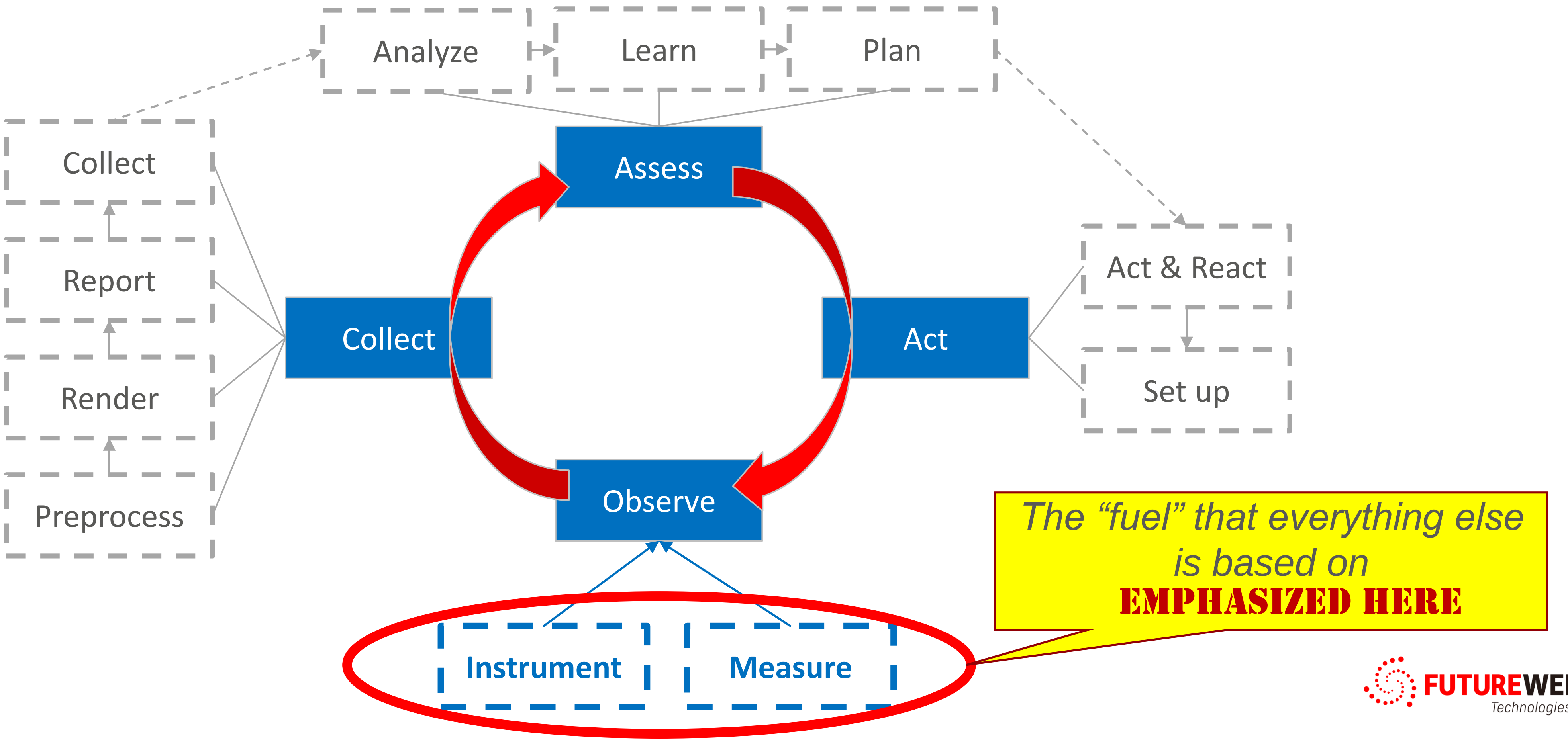
- Network operations and management:
  provide high-precision Service Assurance
  - So: is the current state-of the-art in service assurance sufficient?
    What are the challenges?

“You can’ manage what you can’t measure” (Peter Drucker)
Extensions:
If you can’t measure it, you can’t improve it
If you cannot verify it, you don’t know if you’re getting what you think you’re getting

Image references:
https://efta.org/2018/01/atmia-us-conference-2018-vegas-expect-unexpected/
https://www.theverge.com/2016/10/13/13257470/asteroid-comet-earth-impact-global-warming-period

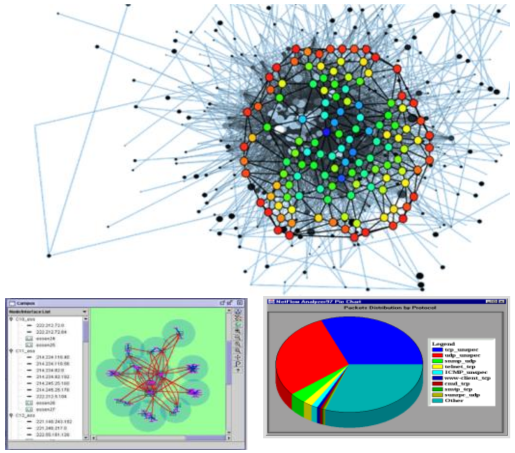
Service Assurance Lifecycle

Service Assurance: the methods, operations, and activities that ensure and verify that services are running smoothly, functioning properly, and meeting their service level objectives



FUTUREWEI Technologies

Flow statistics today



Common techniques: IPFIX & Netflow

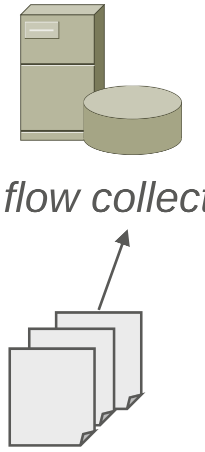
Collect statistics about flows in a flow cache & export

- Flow: packets sharing a common flow key (n-tuple)
- Statistics: counts, num drops, flow term reason, etc, (ca 450 IEs standardized)

Use for monitoring and security (predominant technique today); continued applicability for high-precision services

SrcIf	SrcIPadd	DstIf	DstIPadd	Protocol	TOS	Pkts	Src Port	Dst Port	...	Bytes /Pkt	Active	Idle
Fa1/0	173.100.21.2	Fa0/0	10.0.227.12	11	80	**1107**	00A2	00A2	...	**1528**	1800	4
Fa1/0	173.100.3.2	Fa0/0	10.0.227.12	6	40	**2491**	15	15	...	**740**	41.5	1
Fa1/0	173.100.6.2	Fa0/0	10.0.227.12	6	40	**2210**	19	19	...	**1040**	24.5	15

flow cache



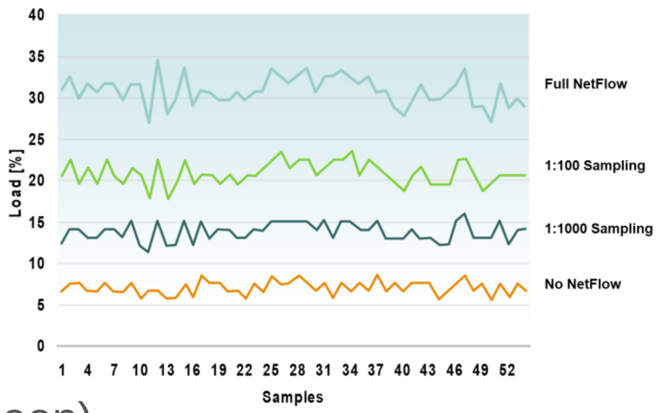
(1) Packet arrives → (2) Flow cache updated → (3) Flow expiration? → (4) aggregate & export

FUTUREWEI Technologies

Image reference: https://thenewstack.io/whats-store-machine-learning-2016/

Challenges

- FE cycles consumption
  - Mitigate by sampling:
    e.g. update cache stats for only 1 in 128 (or 1024) packets
  - Implications:
    - Statistical inaccuracy
    - May miss small flows (may never be seen)
    - May lead to underestimated flow duration (late packets may never be seen)
    - May miss glitches and irregularities of flow behavior
- Flow volume
  - Mitigate by aggregating flow records, longer flow expiration
  - Implications:
    - Coarseness (sub-flows are not distinguished)
    - Staleness of flow data (updates only after minutes, not subseconds)
    - Precludes near-real time control loops on flow statistics
- Functional limitations
  - Static IEs fail to address certain use cases –
    e.g. dynamic flow stats in dependence of certain dynamic conditions
    (e.g. queue occupancy, packet sizes)
  - Not well-suited for real-time control loops



FUTUREWEI Technologies

Is this sufficient for high-precision service assurance?

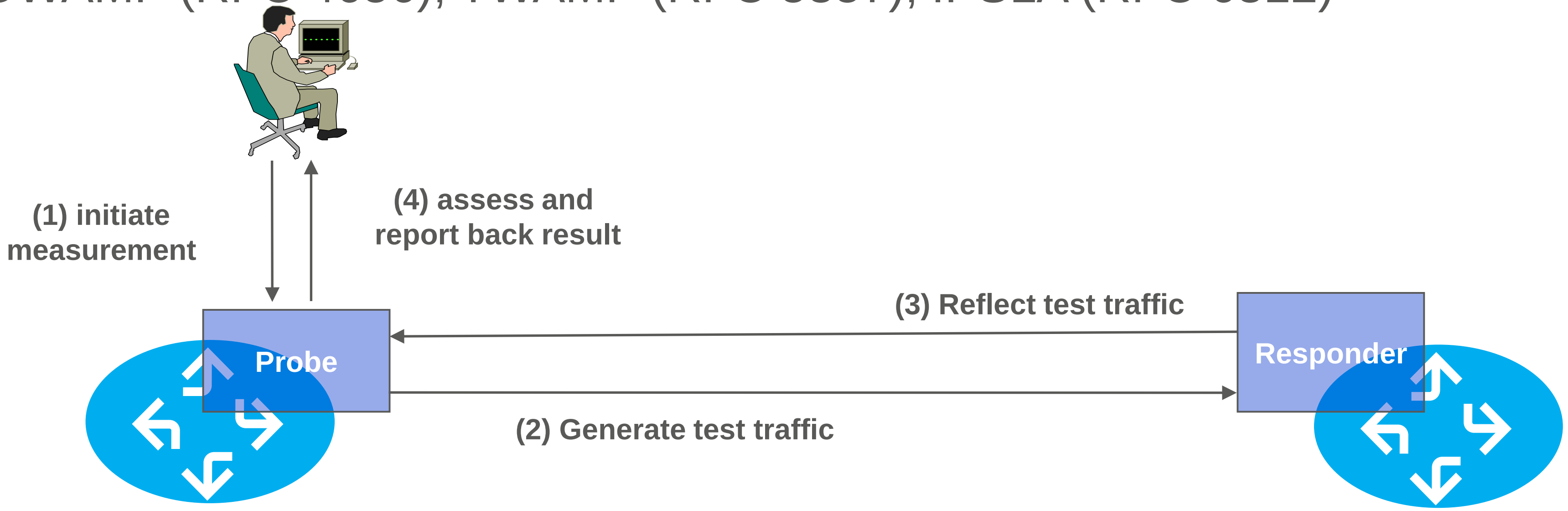
- High-Precision Service Assurance requires accurate flow records
  - Accurate statistics (e.g. of drops, of interarrival rates, etc) needed to assess compliance with SLOs at all times
  - All flows need to be accounted for, including small ones
    - High Precision Services come at a premium – flow records provide one basis for charging
  - Real-time control loops may require shorter export intervals & more records, not less

Current state-of-the art cannot provide this, more advances are needed

- Need full coverage beyond best effort: no reliance on sampling, no missed flows or glitches
- Need greater flexibility: custom statistics as demanded by context
- Need greater scale & smaller time scale:
e.g. custom expiration of flow records to enable faster control loops

Measurements today

- Passive and hybrid measurements
  - Passive: observe packets (e.g. packet capture, sniffing) & timestamp observations
  - Hybrid: add markings and other collateral on production traffic
- Active measurements
  - Generate & measure synthetic test traffic using probes and responders to monitor and validate service levels
  - OWAMP (RFC 4656), TWAMP (RFC 5357), IPSLA (RFC 6812)



FUTUREWEI Technologies

Challenges

- Passive measurements & hybrid measurements
  - Need to observe packets raises privacy concerns, encryption issues can be showstopper
  - Hybrid measurements involve stamping and marking – better, may rely on sampling due to performance constraints
- Active measurements:
  - No issues with encryption, eavesdropping, etc, hence preferred, but....
    - Representativeness of production traffic
    - No proof for individual communication instances
    - High CPU (to generate, reflect, receive, analyze)
    - High network bandwidth consumption
  - Mitigate by selective probing, sampling
    - Coverage across time? across communication pairings?
    - i.e. can you measure everywhere, all the time?

Is this sufficient for high-precision service assurance?

High-Precision Service Assurance depends on the ability to measure service levels

- that are delivered for production traffic
- providing coverage for all communication instances
- highly accurate, not relying on statistical sampling
- in ways that do not encroach on privacy, work in the face of encryption
- with acceptable cost for bandwidth, CPU

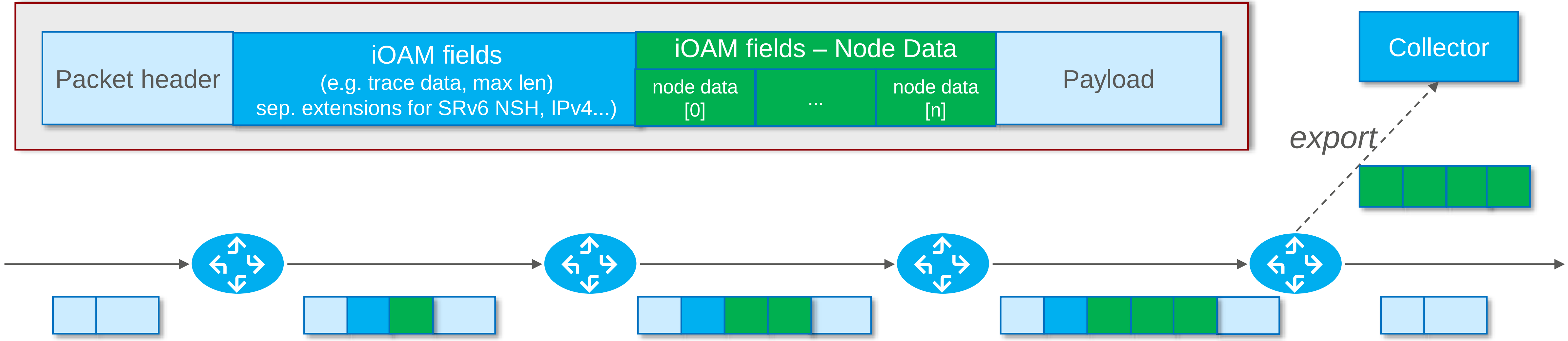
Bonus:

- verifiable & incorruptible – accepable by providers & customers as verdict that SL guarantees are being kept

Current state-of-the art cannot provide this, more advances are needed

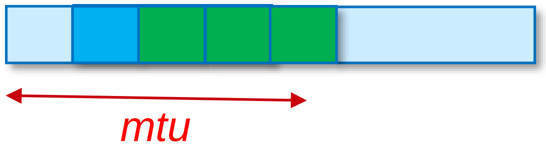
Tracing and IOAM packet telemetry today

- In-situ OAM: assess what happens with a packet while in transit
  - Identify sources of jitter and verify paths
  - Hops add telemetry information to packets that traverse (eg egress queue depth, time stamps)
  - Proof of transit: Update PoT data based on a share of a secret
- Highly relevant for detailed understanding and optimization of service levels (i.e., for high-precision)



Challenges

- Telemetry data size: n data items * path length
  - MTU issues
  - Size variations may cause jitter due to serialization delay
  - Mitigation (1): limit #hops, #data items targeting specific hops (this reduces utility)
  - Mitigation (2): postcard telemetry exported directly from each hop (this is better)
    - Requires off-box correlation/processing/control loops; increases collection complexity



- Telemetry data volume gets large
  - 1 data record per hop, per packet, with possibly multiple data items
  - Intended as troubleshooting tool, not for wholesale SL monitoring and validation – would dwarf volume of production traffic if collected for each flow
  - Mitigation: sampling – at a loss of coverage
- Integration with IP (no extensions)
- Very low-level – no aggregation of data across packets of flows – post-processing required, not well-suited for real-time control loops

FUTUREWEI Technologies

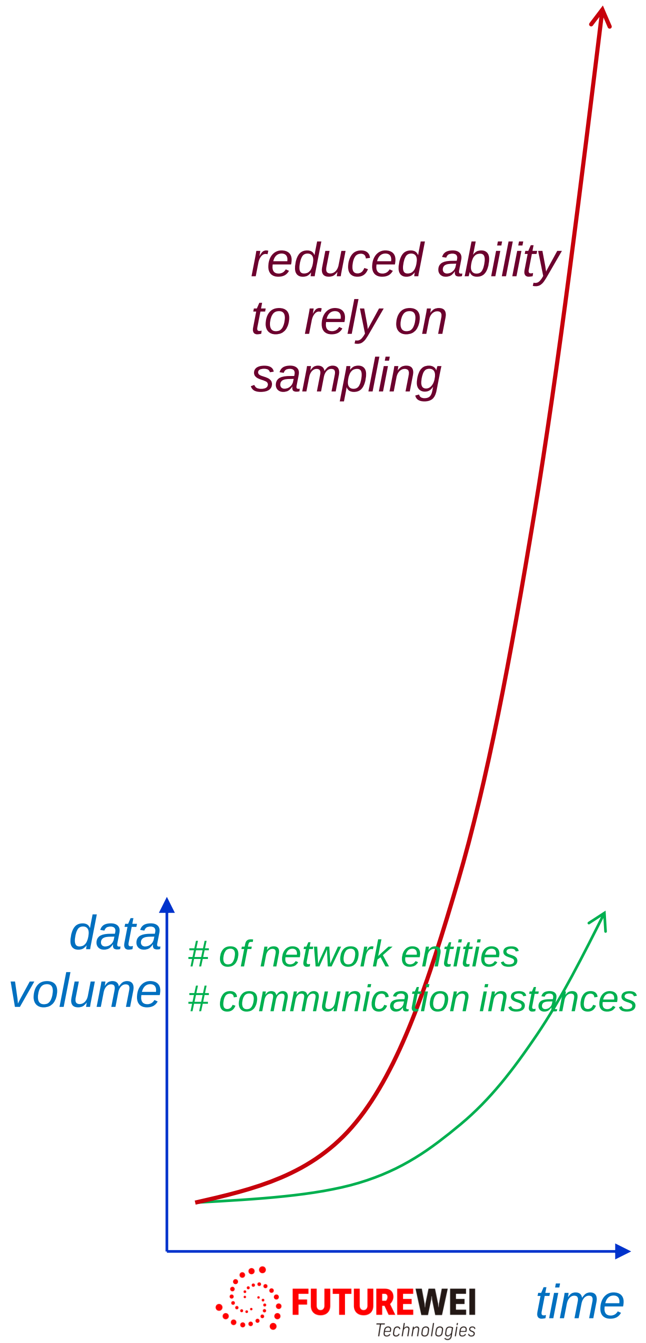
Is this sufficient for high-precision service assurance?

- High-Precision Service Assurance needs visibility into telemetry across a path
  - Optimize high-precision service levels
  - Identify causes for jitter, reconstruct QoS and policy decisions
- Need full coverage to detect "glitches" while keeping bandwidth and CPU tax at an acceptable level
  - No random sampling but schemes that ensure full coverage under "interesting" conditions
- Provide at packet- as well as flow-level
  - Allow to e.g. also capture variations in packet telemetry across totality of flow
- Enable real-time actionable information
  - Enable local control loops with minimal dependence on external systems

Current state-of-the art cannot provide this, more advances are needed

Service Assurance Challenges for High-Precision Networks

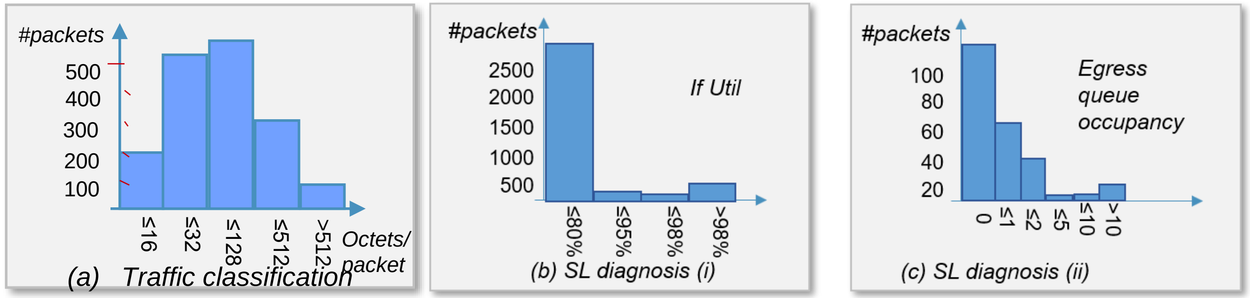
- Accuracy
  - As high-precision service level guarantees are expected, high-precision measurements and instrumentation are required
- Coverage
  - For mission-critical services, ***every*** service instance must be assured and validated
- Sampling and Scale
  - Sampling as a technique to achieve scale will no longer be acceptable
- Real-time control loops at scale
  - Moving beyond validation that occurs after-the-fact will require actionable real-time intelligence for every service instance
- Verifiability and incorruptibility
  - Mission-criticality, guarantees require ability to verify
  - Can measurements, statistics, telemetry be used
    - for charging of high-precision service delivery?
    - for insurance?
    - for proof in a court of law?

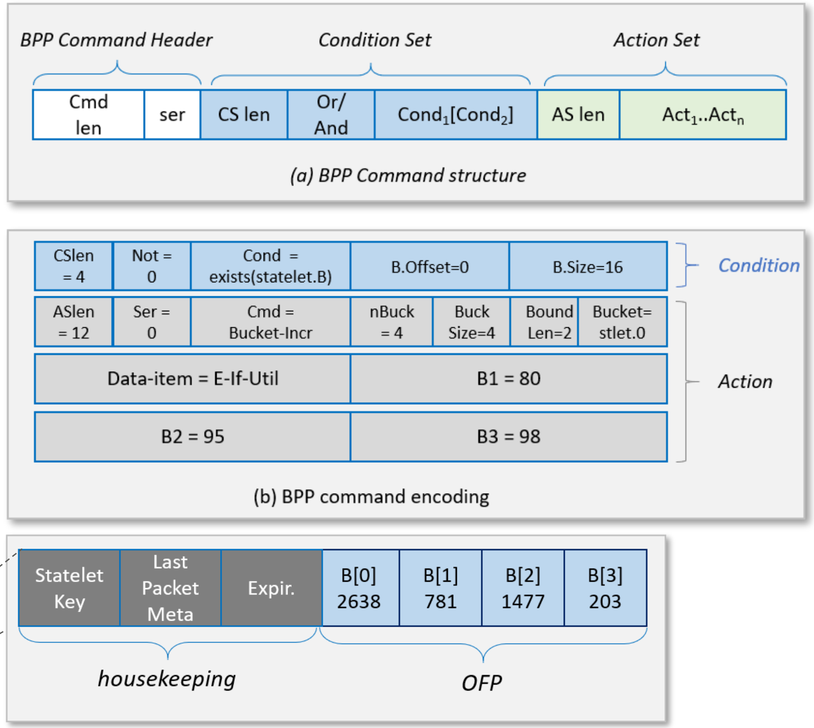


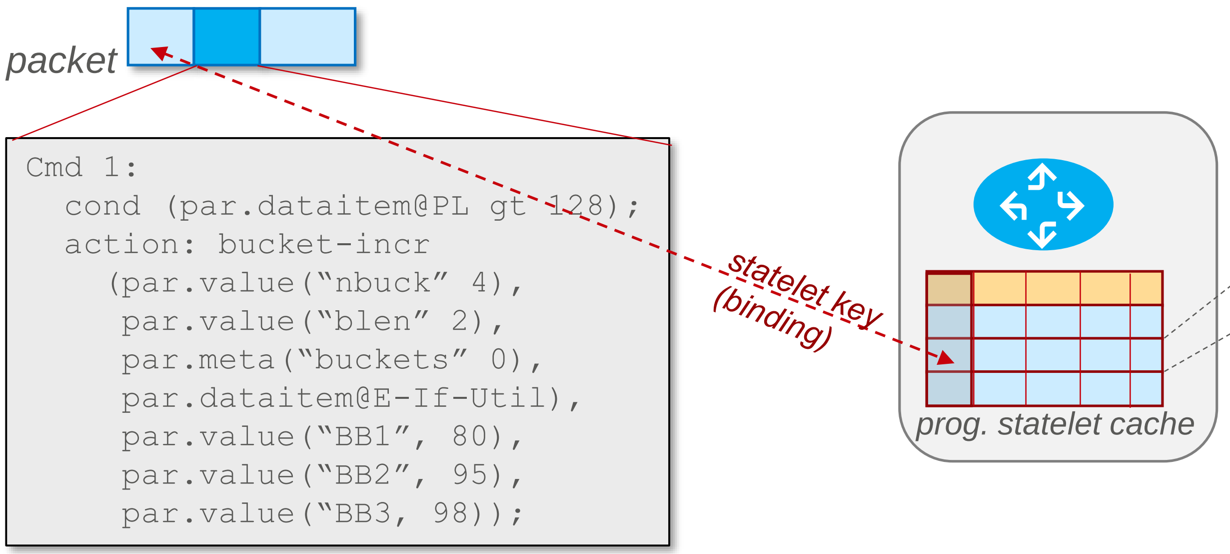
FUTUREWEI Technologies

One solution approach: Operational Flow Profiles (OFPs)

- Packet-programmable statelet cache updated in conjunction with packet forwarding – leverage BPP aka New IP
- Rethink & combine flow records and IOAM: dynamically programmable context-dependent custom stats
- Mitigate scale issues by aggregating data across flows
- Faster time scales through ability to custom-trigger expiration & “micro-flow support”








```
Cmd 1:
  cond (par.dataitem@PL gt 128);
  action: bucket-incr
    (par.value(“nbuck” 4),
     par.value(“blen” 2),
     par.meta(“buckets” 0),
     par.dataitem@E-If-Util),
     par.value(“BB1”, 80),
     par.value(“BB2”, 95),
     par.value(“BB3, 98));
```

A.Clemm, U.Chunduri: “Network-Programmable Operational Flow Profiling”, IEEE Communications Magazine Vol 57 No 7, July 2019.

Conclusions

Today's state-of-the-art for Service Assurance is ill-equipped
to meet the challenge imposed by High-Precision Networks and Services

High-Precision Service Assurance must become part of the design,
not Best-Effort Service Assurance after the fact

More advances are needed
which implies opportunities for research and innovation

Thank You.

FUTUREWEI Technologies