

Precision Availability Metrics for SLO-Governed End-to-end Services

draft-mhmcsfh-ippm-pam

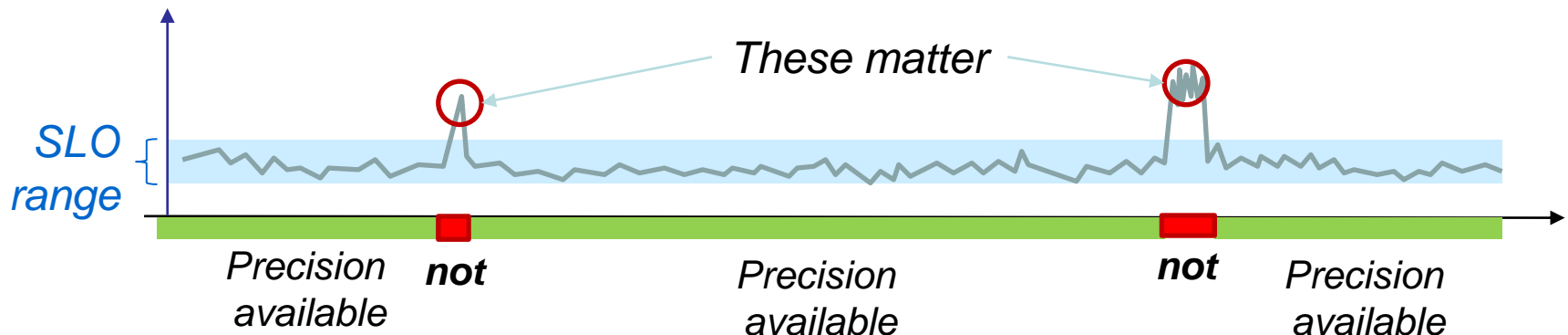
Greg Mirsky
Joel Halpern
Xiao Min
Liuyan Han
Adrian Farrel

Alexander Clemm
John Strassner
Jérôme François
Mohamed Boucadair

IETF-115, November 2022

Precision Availability Metrics

- Observation #1: SLOs are key – you need to count what counts
 - Critical performance metrics reflected a set of SLOs
 - In some use cases, the complete history of each SLO is not needed
 - Capturing violations (and asserting their absence) is often sufficient (and more efficient to retain)
- Observation #2: Analogy between service and system failures
 - Inability to deliver contracted SLOs is a failure
 - Precision Availability is a form of availability



Update

- Welcome Adrian Farrel
- Addressed comments received

SLA \neq Σ SLO

We use definitions of SLA and SLO from draft-ietf-teas-ietf-network-slices. An SLA may include SLOs and Service Level Expectations (SLEs). Unlike SLOs, SLEs can be formally specified but are challenging to measure.

On the Usefulness of Ratio Metrics

Because the transition between service availability/unavailability states is based on the pre-determined number of consecutive intervals, shorter conditions may need to be adequately reflected. That is where ratio provides helpful information, e.g., VIR is the ratio of the sum of VIs and SVIs to the total number of time intervals.

Composite Services and PAM

A composite service might include a set of connectivity constructs (e.g., p2p, p2mp, a2a) (see draft-ietf-teas-ietf-network-slice-nbi-yang for more detail). An SLO might apply to all the constructs, or some constructs are assigned different sets of SLOs. For the purpose of PAM, each connectivity construct that composes the service can be monitored for its own SLO conformance as a sub-service. The composition of PAMs of these sub-services can be viewed as PAM of the composite service.

Discussion items

- Metrics: individual packets that violate SLO(s), e.g., counts of violations related to individual packets may be added in the future?

Future work (beyond this draft)

- YANG data model
- IPFIX Informational Elements
- Support for statistical SLOs, e.g., histogram and/or bucket
- Policies to define violated time unit, configure metrics
- Additional second-order metrics, e.g., “longest disruption of service time”

Next steps

- Welcome comments, questions
- Complete WG adoption poll
- Resolve items noted for further discussion (Metrics)

Thank you