

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

draft-ietf-ippm-pam

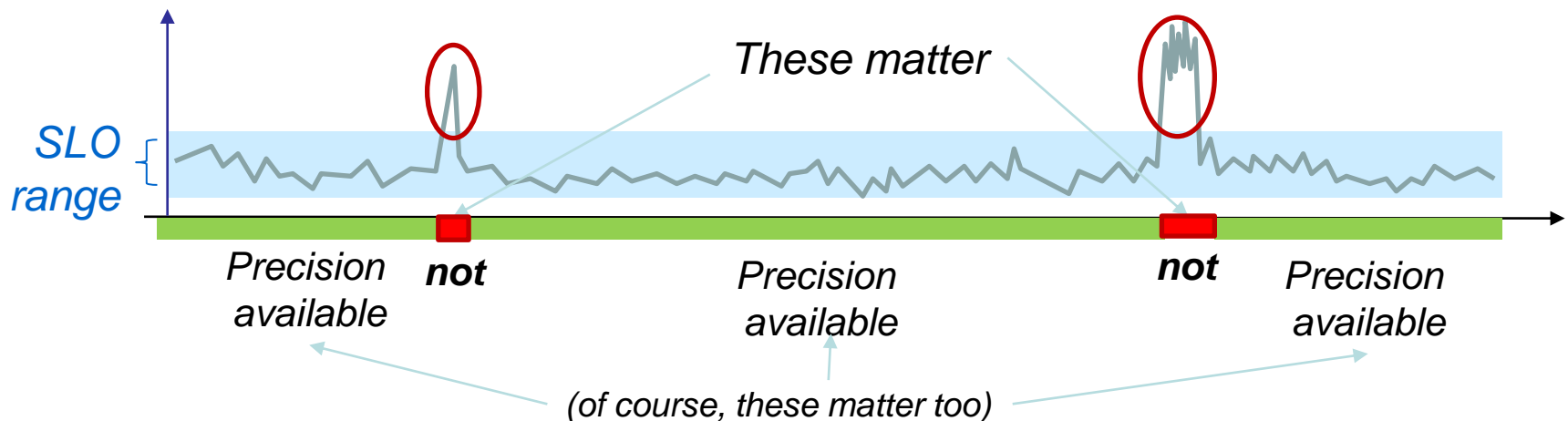
Greg Mirsky
Joel Halpern
Xiao Min
Liuyan Han
Adrian Farrel

Alexander Clemm
John Strassner
Jérôme François
Mohamed Boucadair

IETF-116, March 2023

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

- Addressed comments received in the course of the WG AP
- Introduced Violated Packets Count/Severely Violated Packets Count
- Worked through and resolved discussion Items
- Editorial updates

VPC & SVPC

- Intended to maintain counts of packets for which a performance threshold is violated
- May be used, for example, to distinguish between cases in which violated intervals are caused by isolated violation occurrences or by extensive violations across multiple packets
- Can be compared with the overall amount of traffic to facilitate assessing compliance with statistical SLOs

PAM Policy vs PAM Configuration Settings

- Severity of violations (e.g. presence of threshold levels), availability status (in case of successive violations) may be dependent on operator practices.
- Such setting should not be “hard coded” but customizable/ configurable
- These settings can be deemed to constitute a “PAM policy” earlier, but policy is a big and overloaded term
- Resulting from that conclusion – changed the terminology from “PAM policy” to “PAM configuration settings”

Discussion items

Future work, new documents (outside scope here):

- YANG data model:
 - to control a PAM-capable system
 - to retrieve PAM metrics
- IPFIX Informational Elements
- Statistical SLOs, e.g., histogram and/or bucket
- Additional second-order metrics, e.g., “longest disruption of service time”

Incorporated with clarifications:

- PAM configuration settings
- Considerations to implement an availability state model (itself out of scope)

Next steps

- Welcome comments, questions
- As discussion items have been resolved,
WG LC?

Thank you