In-band Network Telemetry update

Mauro Campanella (GARR), - mauro.campanella@garr.it
Tomas Martinek (CESNET), Damian Parniewicz (PSNC),
Federico Pederzolli and Damu Ding (FBK),

22nd STF Meeting, 25 Feb 2021, virtual

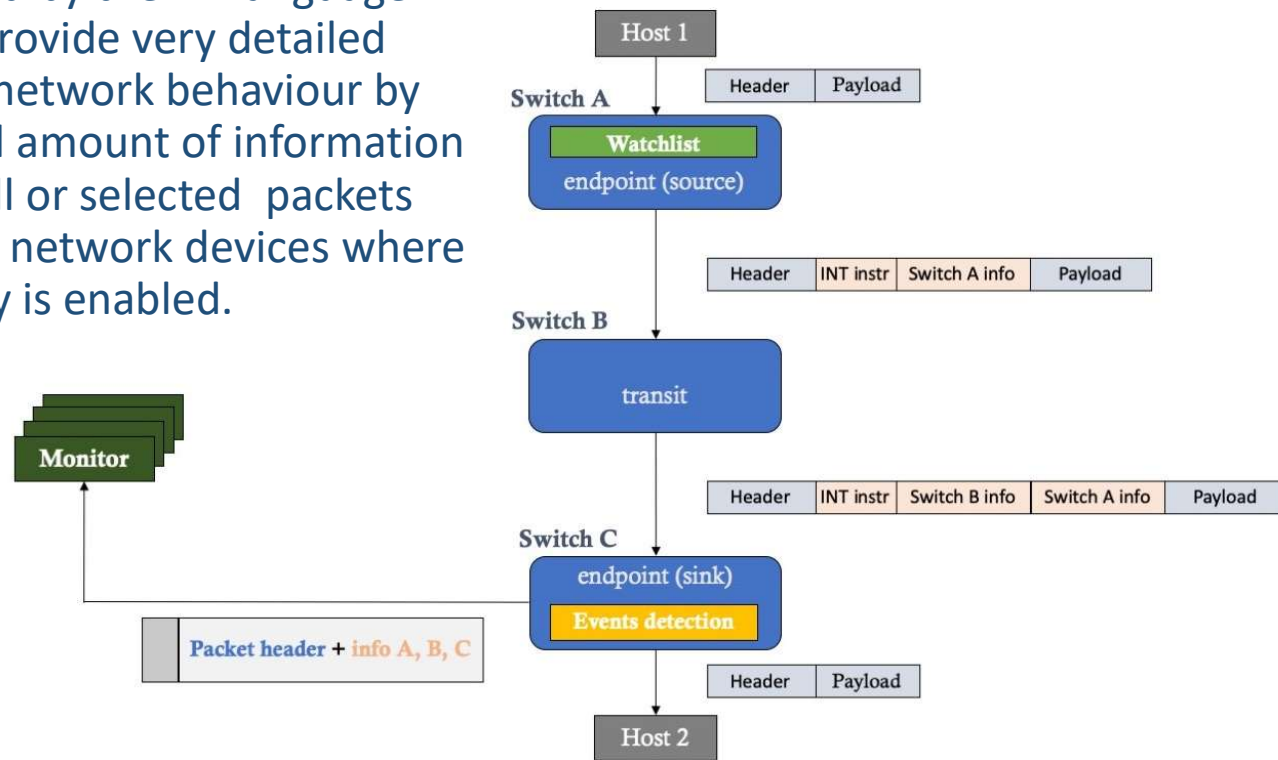Public

www.geant.org

Why INT (and telemetry)?

- Telemetry is the next step in monitoring , starting with streaming telemetry

- INT adds to monitoring granularity (choice of flows / packet / protocols/…) to monitor and programmability in the data plane, also user-define

- Gather experience with network "Big Data" handling

- Evaluate its usefulness to better monitor and control our networks (creating new knowledge)

From participants to the Telemetry and Big Data Workshop: monitoring procols used today
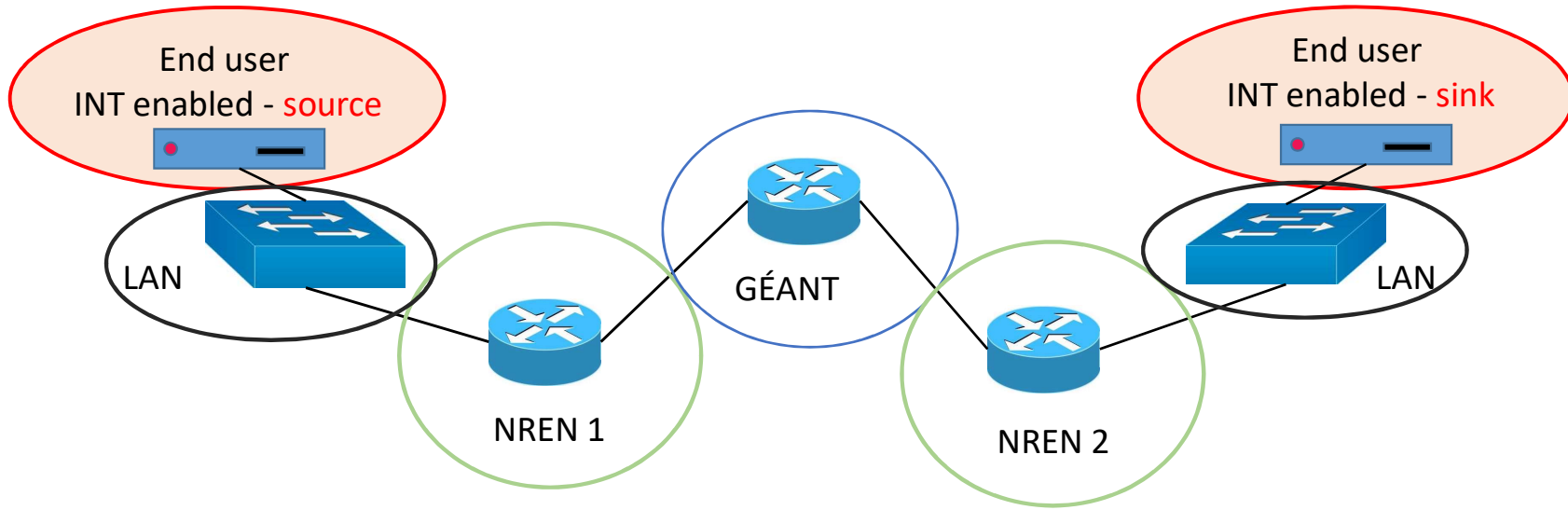

2

GÉANT
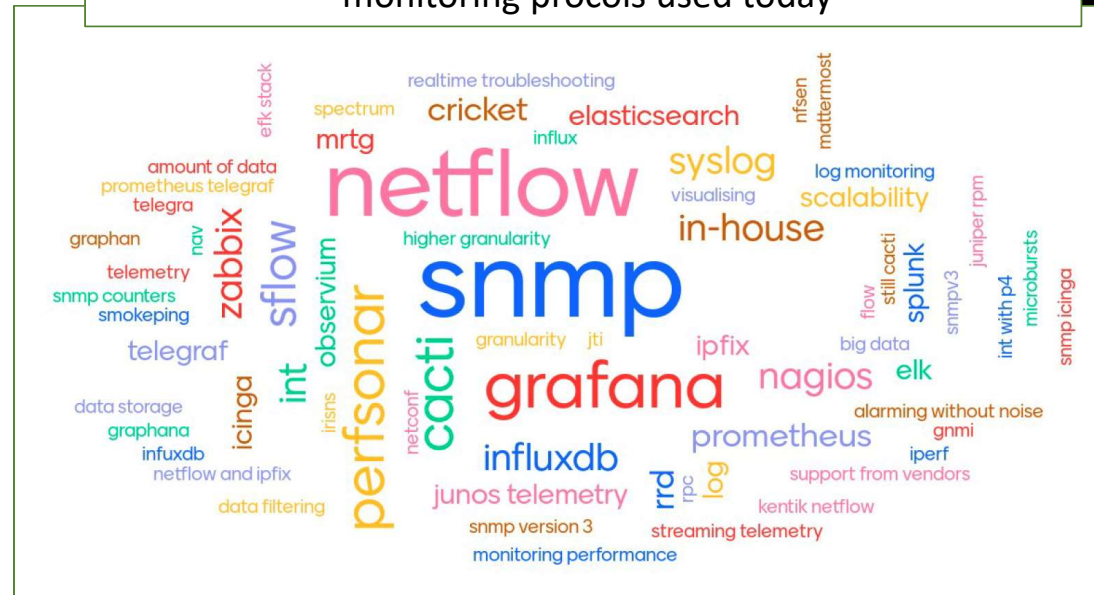
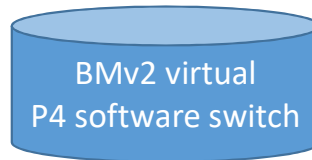# In-Band network Telemetry (INT) summary

INT was specified by the P4 language community to provide very detailed information on network behaviour by inserting a small amount of information directly inside all or selected packets passing through network devices where INT functionality is enabled.

www.geant.org

End user uses INT on its traffic to measures its IPDV, Loss,...

End user
INT enabled - source

LAN

NREN 1

GÉANT

NREN 2

End user
INT enabled - sink

LAN

Edgecore Wedge100BF-32X   Arista 7170-32c
Tofino (Barefoot/INTEL)

BMv2 virtual
P4 software switch

FPGA

4

GÉANT

Network uses INT to debug network behaviour

End user

End user

LAN

GÉANT

LAN

INT enabled sources

NREN 1

NREN 2

NOC – INT sink

Examples :

Use as passive monitoring tool

- Monitoring flows from a specific user group

- Service performance

- Routing checks

Or active: IPDV and delay measurement (active)

5

www.geant.org

GÉANT
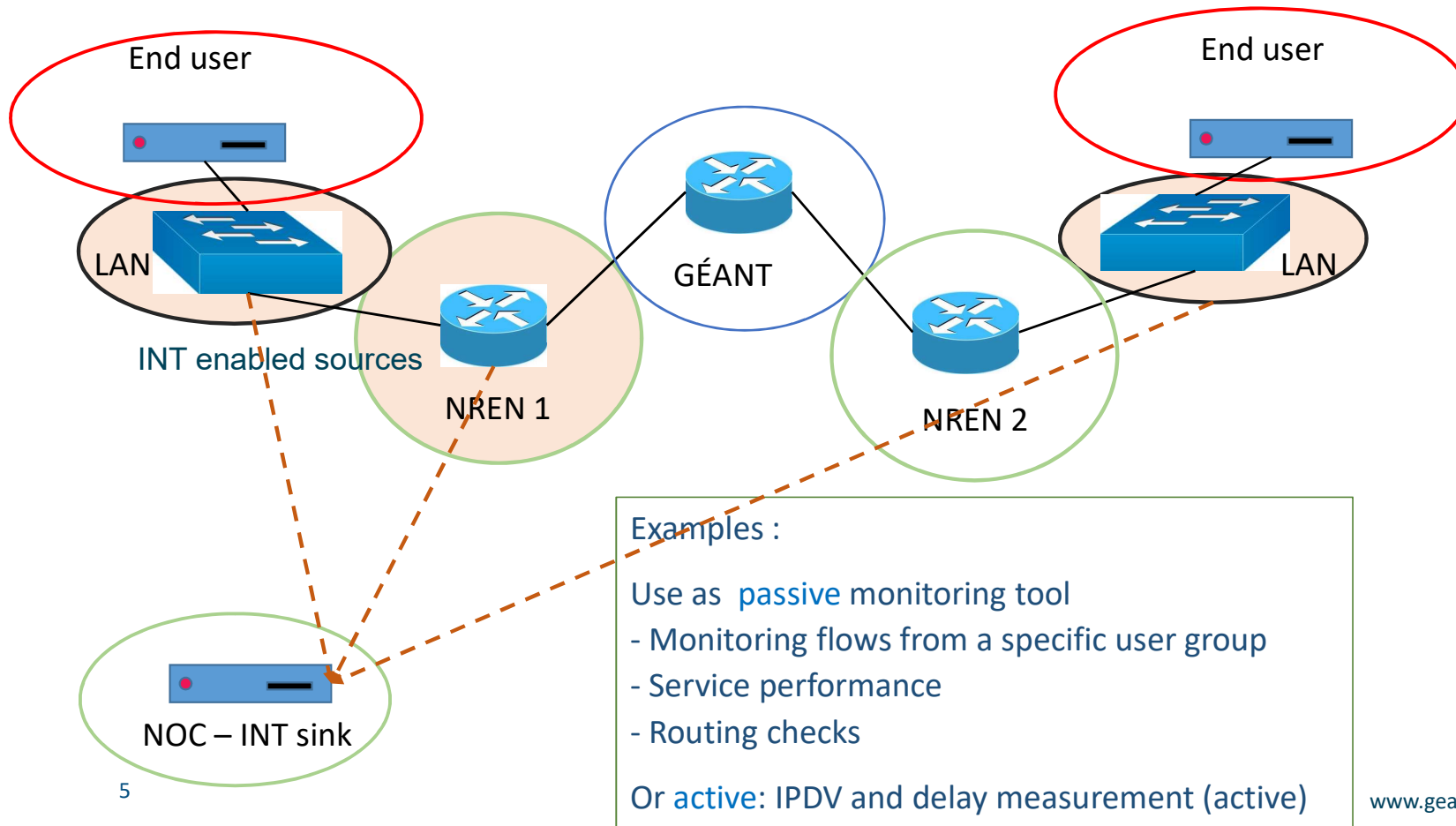
INT : testbed on our networks – "end-user mode"
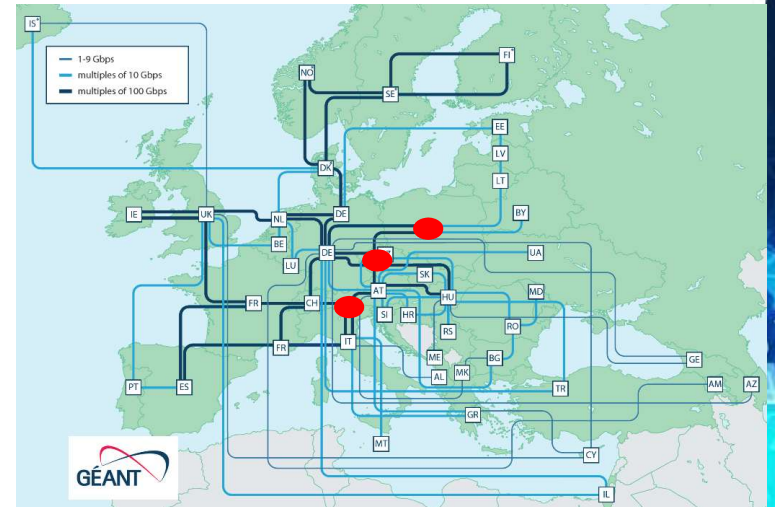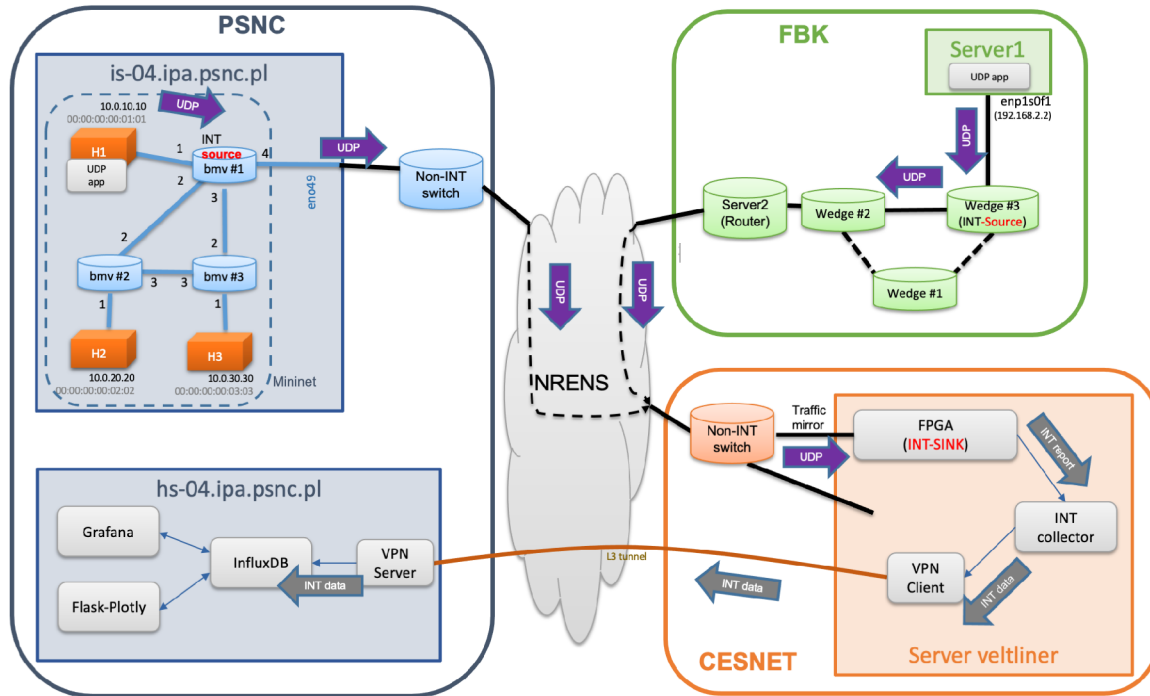
- 3 switch types
- UDP packets flow on NRENs networks
- Collected INT data in CESNET is sent back to PSNC for collection and presentation.

DPP- INT distributed test bed in "user mode" IPv4 public addresses, following standard routing on production networks.

6

www.geant.org

The choices and technologies:

Choose INT data and Header position — INT headers contain source and destination timestamp, sequence number, placed between UDP and payload

Coding using P4 — HW programming, limited cycles and memory, lack of complex arithmetic operations, lower flexibility in the use of registers

Collector for INT data — Tuned InFluxDB (up to 260K INT data/s multiple threads)
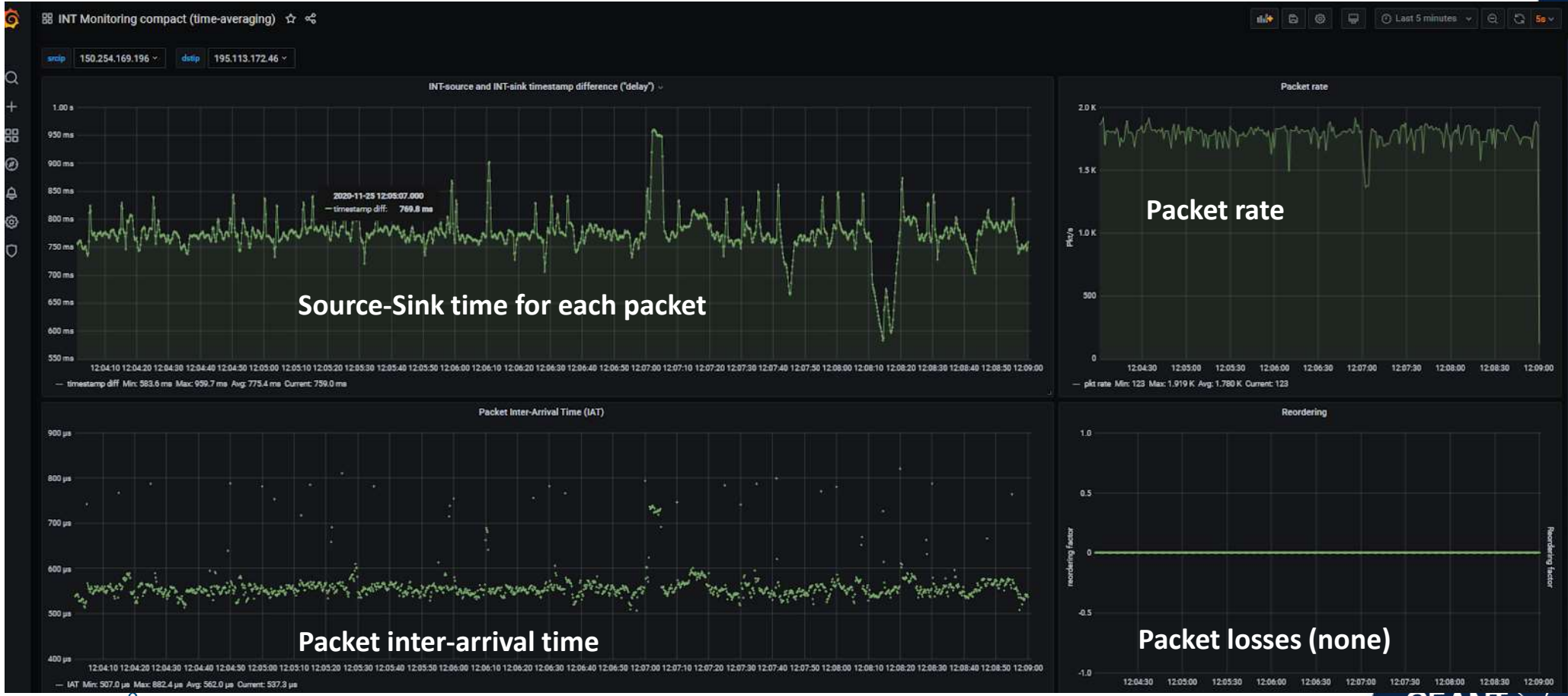
Present — Grafana, Plotly , added IPDV computation in data plane

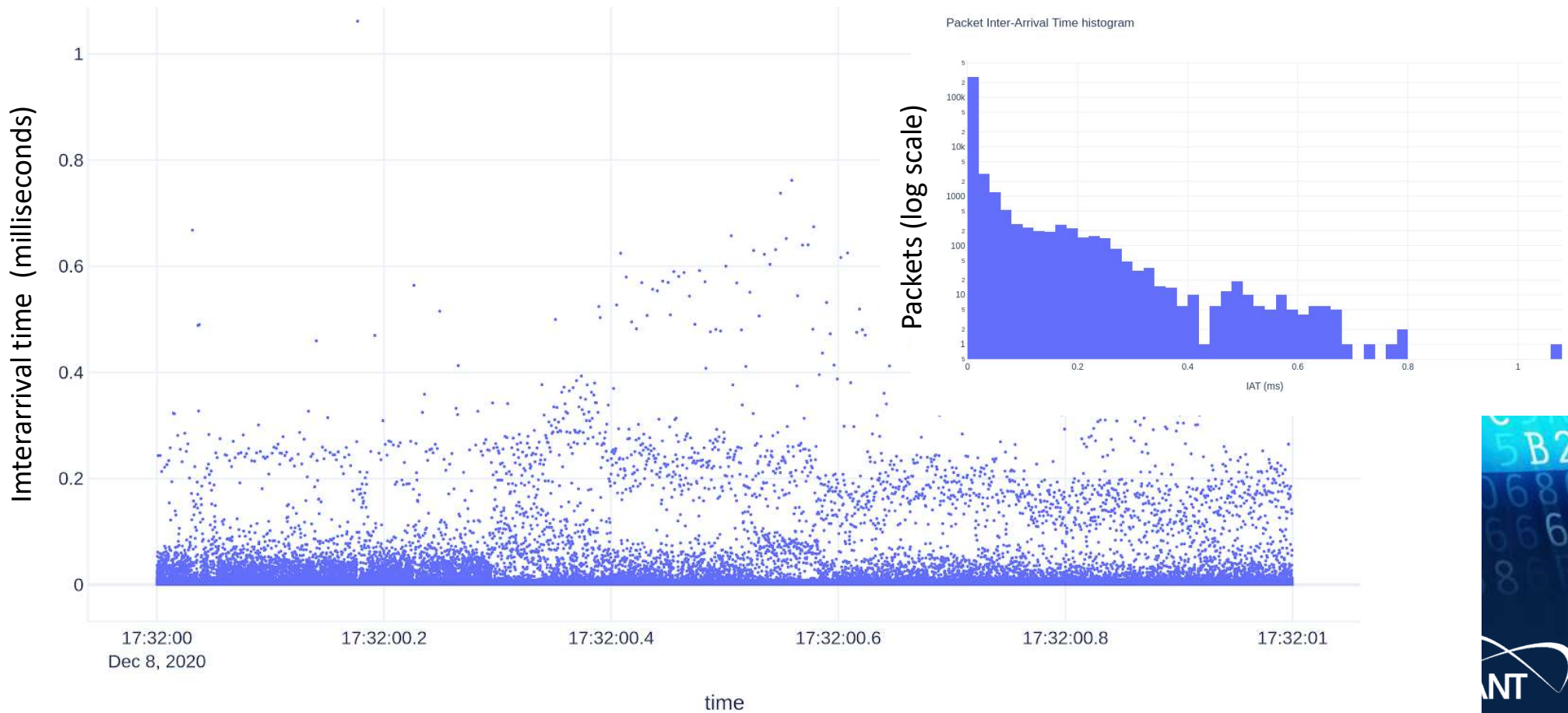**Analyze** — "Knowledge" under development now

Time synchronization to be tuned to sufficient precision (few microseconds, ns), need also node internal clock stability

GÉANT

5 minutes of the INT monitored flow from PSNC to CESNET

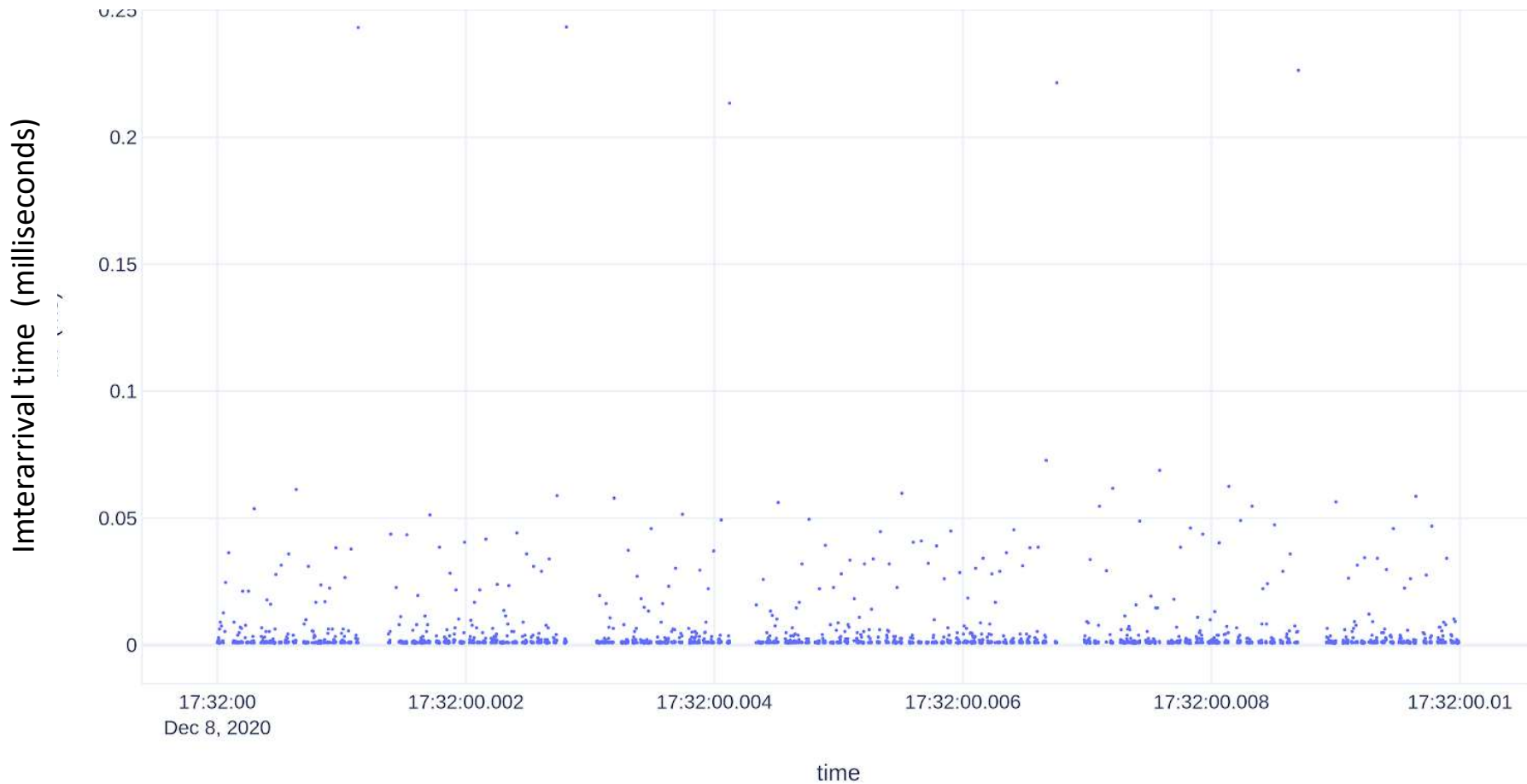# INT example: 1 second of interarrival packet time for a UDP flow of 260 K packets / second (4 microseconds averge) between two NRENs

10 milliseseconds of interarrival packet time for a UDP flow of 260 K packets / second (4 microseconds average)

Status

- INT code for this tests is stable using INT Spec 0.5.
  P4 uses v16 in BMv2 and Tofino, P4v14 in FPGA.
  Available on request for now from Github for the three platforms
- Time synchronization under evaluation for improvement and stability in the testbed. To be checked for all types of hardware
- Data collection and presentation tools and configuration optimized
- INT experience detailed report to be published soon

GÉANT

Ongoing developments

- Testing Transit node INT functionality and TCP header INT parser

- Basic INT code implementation over DPDK to provide an INT tool that does not require special hardware, runs in off-the-shelf computing equipment and that can have many Gigabit/s performance (in testing phase)

- Development of a BMv2 INT docker image as a tool to download for easier testing and trials with INT basic use case and P4 programming (timestamps, sequance numbers , including collection and visualization, based on what is already being used in PSNC

- Compliance to INT standard 1.0 and then 2.0

- Improving timestamping

Next steps

- Increase testing topology (up to transatlantic?)
- Establish collaborations to:
    - Identify and develop new use cases
    - Further improving the basic tools
    - Discuss and disseminate the knowledge gathered
    - In-depth data analysis
    - Standardize approach to INT and Data Plane Programming

If interested please drop a mail to mauro.Campanella@garr.it

GÉANT

Summary

- INT (and Data Plane Programming) (using P4) is not business-as-usual, requires specific expertise, however it offers a great tool for monitoring, debugging and providing information to control plane, in real time.

- P4/INT is more and more available in various platforms (switches and linecards)

- INT is a powerful magnifying glass on network behaviour

- Time synchronization between nodes is important

- As a function of the use case, the INT/P4 bases tools may require handling of large amount of "raw" data, to be used for analytics and more. It need further insight, tools and equipment to scale.

GÉANT

References

- The GÉANT First Telemetry and Big Data Workshop
  https://wiki.geant.org/display/PUB/Telemetry+and+Big+Data+Workshop

- INT Tests in NREN networks – Report draft – soon available

- For INT use cases and results see section D, pag. 2 in

  "A Survey on Data Plane Programming with P4: Fundamentals, Advances, and
  Applied Research", Frederik Hauser, Marco Häberle, Daniel Merling, Steffen
  Lindner, Vladimir Gurevich, Florian Zeiger, Reinhard Frank, and Michael Menth
  26 Jan 2021, to be published in" Communications Surveys & Tutorials (COMST)
  journal -- *https://arxiv.org/pdf/2101.10632.pdf*

GÉANT

Thank you

Any questions?

www.geant.org