

32 Tb/s network monitoring tool for the Run 3 LHCb real-time DAQ

Marta Gomis Domènech

CERN Summer Student 2021

Supervised by Rafal Krawczyk and Niko Neufeld



Presentation outline

- About me
- New network architecture of LHCb DAQ for Run 3
- Problem statement
- Tool architecture
- Grafana dashboard demo (UI)
- Project outcomes
- Challenges and conclusions

About me

- From 
- MSc Student in Electrical Engineering 
(Grenoble INP – UGA, France)
 - Specialization: Power Electronics for Smart Grids
 - Interested in data acquisition and data analysis



LHCb
LHCb

ONLINE



Problem statement

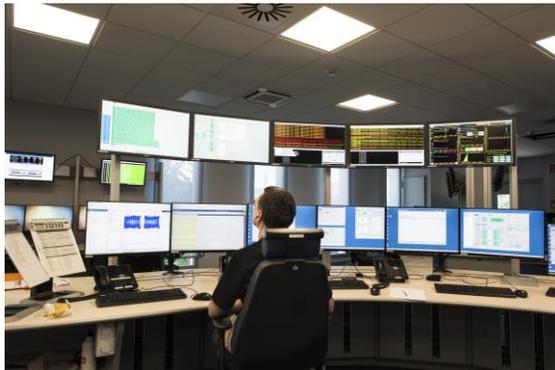
New LHCb Event Builder
network topology for Run 3:



**Switched fabric over
InfiniBand (IB)**



New tools for network
performance monitoring and
network diagnosis



LHCb control room

Final usages:

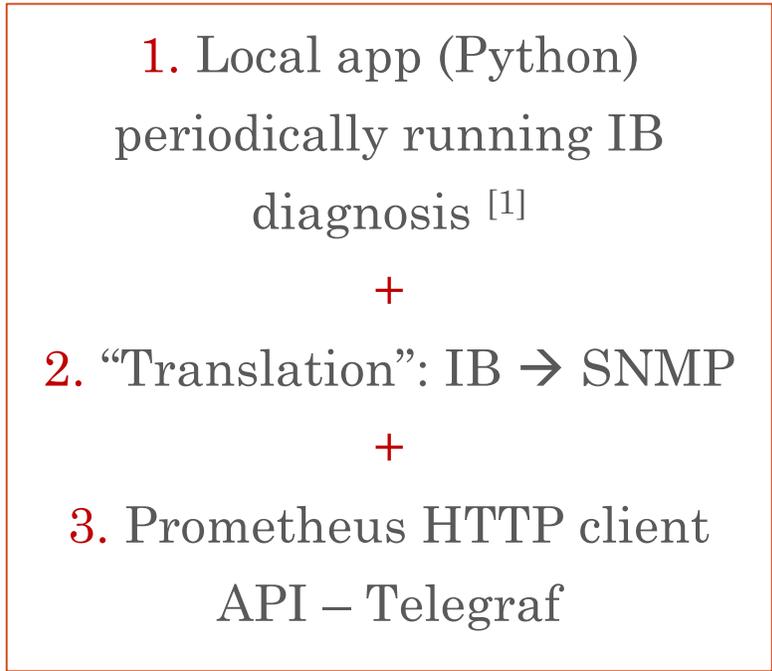
- System installation
- Run time: control room (end users) → network congestion and errors

Tool architecture

Constraints:

- Network switches **unmanaged** →
- Data collection: Prometheus server (sandbox testing environment)
- Data visualization: Grafana web app (UI)
- SNMP Standard compliance

~~SNMP~~ (Internet Standard protocol)



(App based on: [1]
<https://github.com/guilbaults/infiniband-exporter>)



Prometheus



Grafana

1. IB diagnosis

- 1st approach: “black-box”
 - ⚠ Documentation
- **Final approach:** reverse engineering using Linux GDB

Results:

- A. Choice of two IB programs:
 - Counters and links data retrieval (“ibqueryerrors”)
 - Counters reset (“perfquery”)
- B. Modification of IB source code (C language) to fix bugs and unclear methods

2. IB → SNMP

Scrutiny of IB and SNMP documentation

★ RFC Internet Draft [2]

Results:

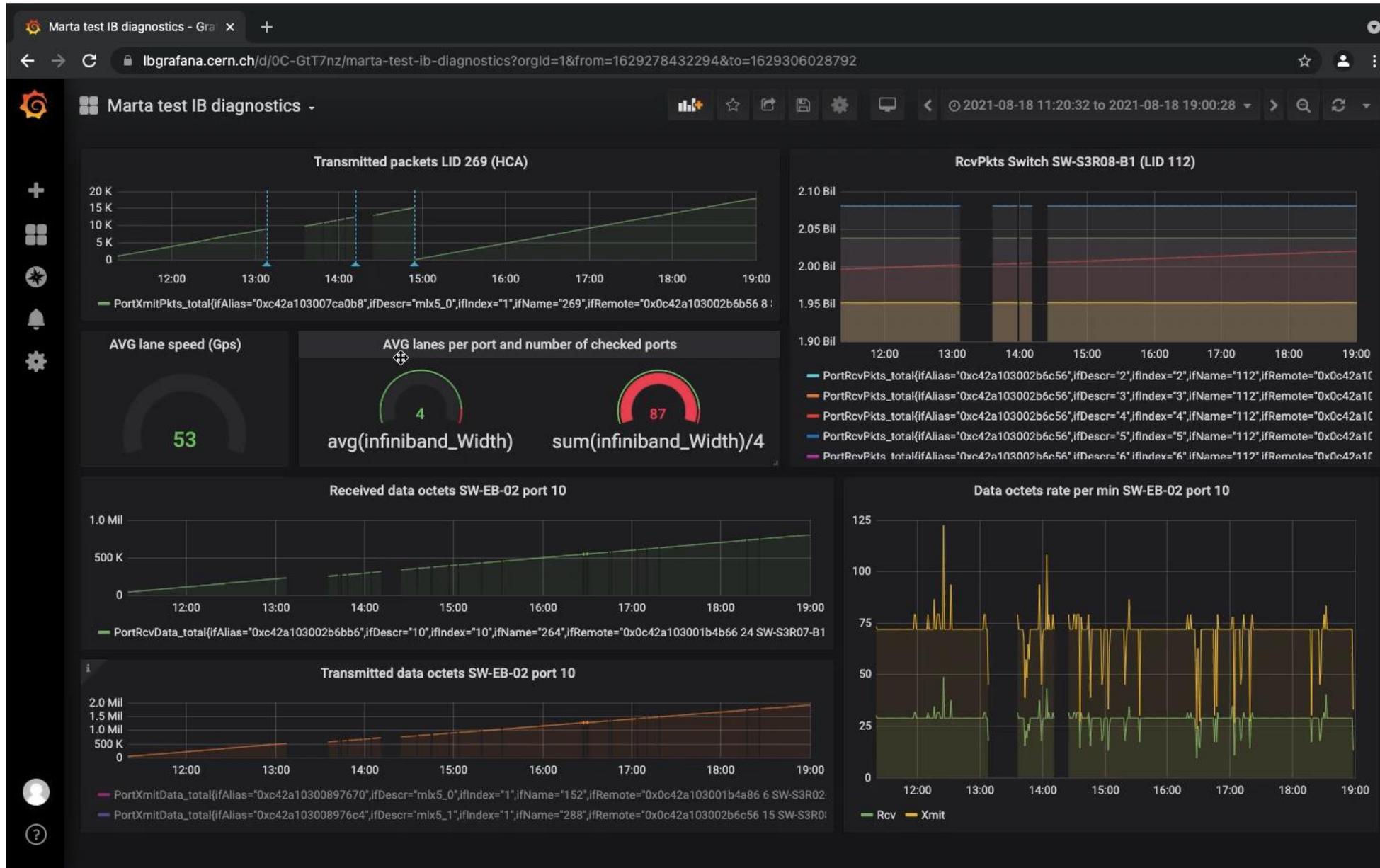
- A. Protocol-compliant mapping table for future reference
- B. Choice of labels and metrics for Prometheus

[2] IP over InfiniBand – Definitions of Managed Objects for Infiniband Interface Types
<https://tools.ietf.org/pdf/draft-ietf-ipoib-ibif-mib-09.pdf>

3. Prometheus API – Telegraf

```
← → ↻ ⚠ Not Secure | tdeb02.lbdaq.cern.ch:9683
# HELP ibIfPortStatInactDiscards_total Total number of
packets discarded due to the port being in the inactive
state.
# TYPE ibIfPortStatInactDiscards_total counter
ibIfPortStatInactDiscards_total{ifAlias="0xc42a103001b4b66",
ifDescr="24",ifIndex="24",ifName="130",ifRemote="0x0c42a1030
02b6bb6 SW-EB-02 264 10"} 4.0
# HELP ifOperStatus Current operational (physical) link
state. 1: LinkUp, 2: Disabled, 3: Polling, 4: Sleep, 5:
PortConfigurationTraining
# TYPE ifOperStatus gauge
ifOperStatus{ifAlias="0xc42a103002b6b5e",ifDescr="Technologi
es",ifIndex="1",ifName="121",ifRemote="0x0c42a103002b6b56
SW-S4R17-B1 111 41"} 1.0
# HELP infiniband_ibqueryerrors_duration_seconds Number of
seconds taken to run ibqueryerrors (both performance and
errors data).
# TYPE infiniband_ibqueryerrors_duration_seconds gauge
infiniband_ibqueryerrors_duration_seconds 0.5073096752166748
```

Grafana dashboard demo



Summary and project outcome

Working diagnosis tool for the Run 3 LHCb real-time DAQ

- Mapping table IB – SNMP: labels and metrics
- Gitlab repository #1: Modifications of IB tools C source code
- Gitlab repository #2: Python exporter of InfiniBand counters to Prometheus, in accordance with SNMP Protocol
- Grafana dashboard: set of panels for data visualization

Challenges and conclusions

Challenges

- Several platforms and tools
- Discussions with different groups in each stage of the project
- Limited time for implementation

Conclusions

- Main objective accomplished: well documented diagnosis tool for the new LHCb Event Builder network for Run 3 

- Better understanding of IB diagnosis tools

Personally:

- Programming competences: new language, experience debugging existing code
- First experience with DAQ, supercomputer networks and communication protocols



Thanks for your attention

Questions