RNP 30 anos
ORGANIZAÇÃO SOCIAL DO MCTI

The Challenge of Delivering Open OSS Data for Research

The RNP contribution to overcome barriers

Alex Moura
alex@rnp.br

Artur Ziviani
artur@lncc.br

Leobino N. Sampaio
leobino@ufba.br

Rafael L. Gomes
rafaellgom@larces.uece.br

NSF Workshop on Overcoming Measurement Barriers to Internet Research (WOMBIR 2021)

Jan 11th 2021

OSS network data requests by research community

CT-Mon surveys with the Brazilian research community (2018 - 2020)

Due to limited funding, where should we invest resources that will deliver the most benefit for research?

A microservices cyberinfrastructure to gather measurement data collected by Operations Support Systems (OSS) that will organize, store, anonymize, convert and deliver datasets from network performance, operations, security, services and applications

OSS MEASUREMENT DATA OF INTEREST BY THE COMMUNITY (2018)

- **Network performance** active **measurements**
- **Flow traces with raw data** truncated at **first 200 bytes**
- **Cyberinfrastructure measurements**: both **baremetal and virtual servers**: CPU, RAM, buffers, cache, I/O, IRQ and network **routers**, **switches, applications** and **services**
- Detailed **network physical topology** (nodes and edges)
- Network equipments **configurations** (routers and switches)
- Services **logs**
- **Circuit utilization** in each PoP
- **Routing tables** (from each router)
- Support **Trouble Tickets**

OSS MEASUREMENT DATA FOR ANALYSIS OF THE IMPACTS OF COVID-19 IN THE NETWORK (2020)

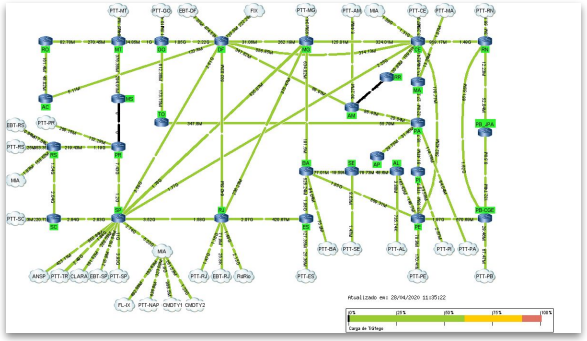
- **Internal** network **traffic utilization**
- **Utilization** traffic **levels at peerings** and exchange points between RNP and commodity SPs
- **Specific** services and applications measurements (e.g. web conferencing system)
- **Network Performance traffic matrix** including RNP's peering partners
- **VPN utilization** measurements from commodity SPs to RNP client institutions
- Specific **traffic matrices** (e.g.: university hospitals, R&D HPC centers)
- **BGP dump updates** received at border routers
- **Periodic routing table dumps** from RNP's border routers (before BGP best-path selection). Allows verification of simultaneous route changes with changes in traffic patterns.
- **Periodic dumps** of **IGP** messages (to follow state of the IGP)

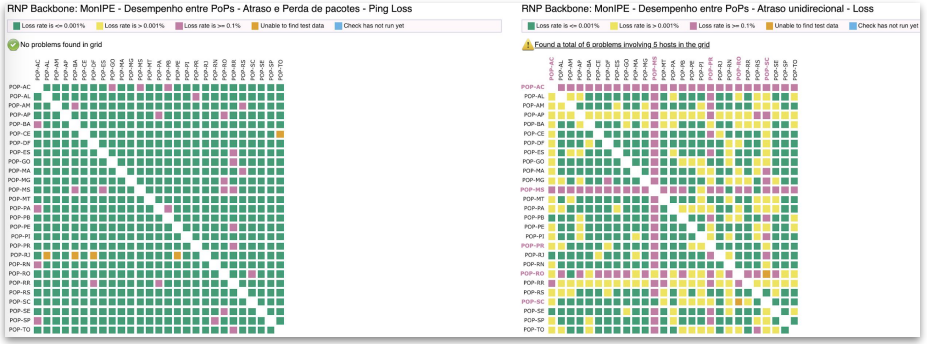
RNP measurements today

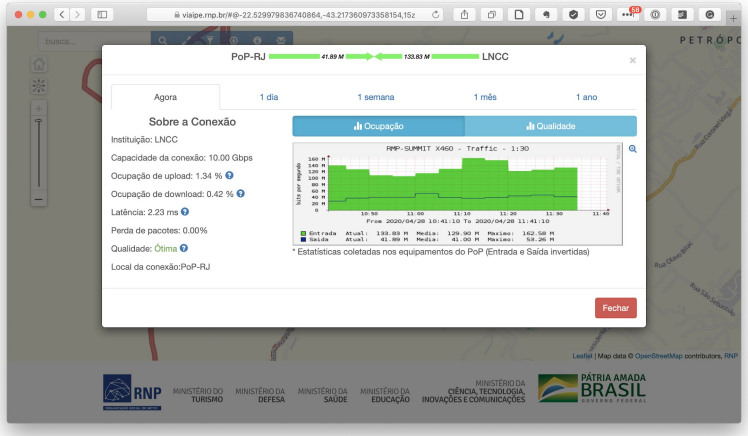
RNP OSS Network Data

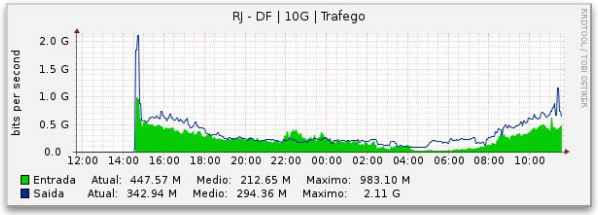
RNP's CT-Mon verified that we have many measurement teams and tools, but also there are integration gaps in internal processes, resources to facilitate data collection, storing, analysis, correlation for research.

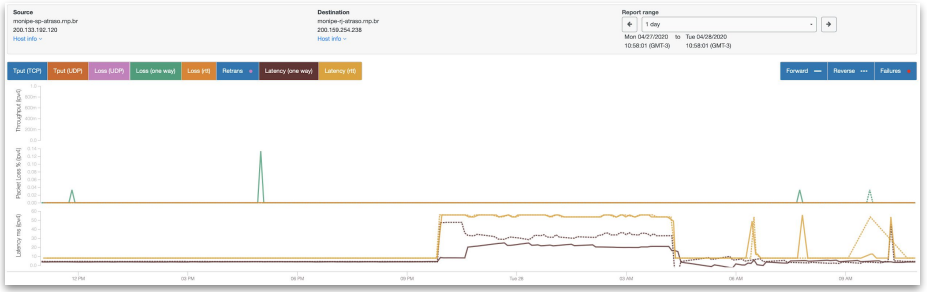
- Tools like perfSONAR and Network Weathermap and other visualizations tools are available, but no easy access to correlations or raw data export
- Not useful scenario to make useful contributions for new research

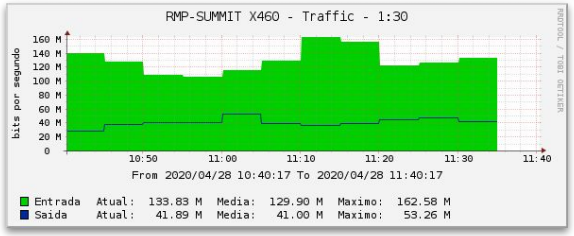












RNP barriers to provide OSS network data for research

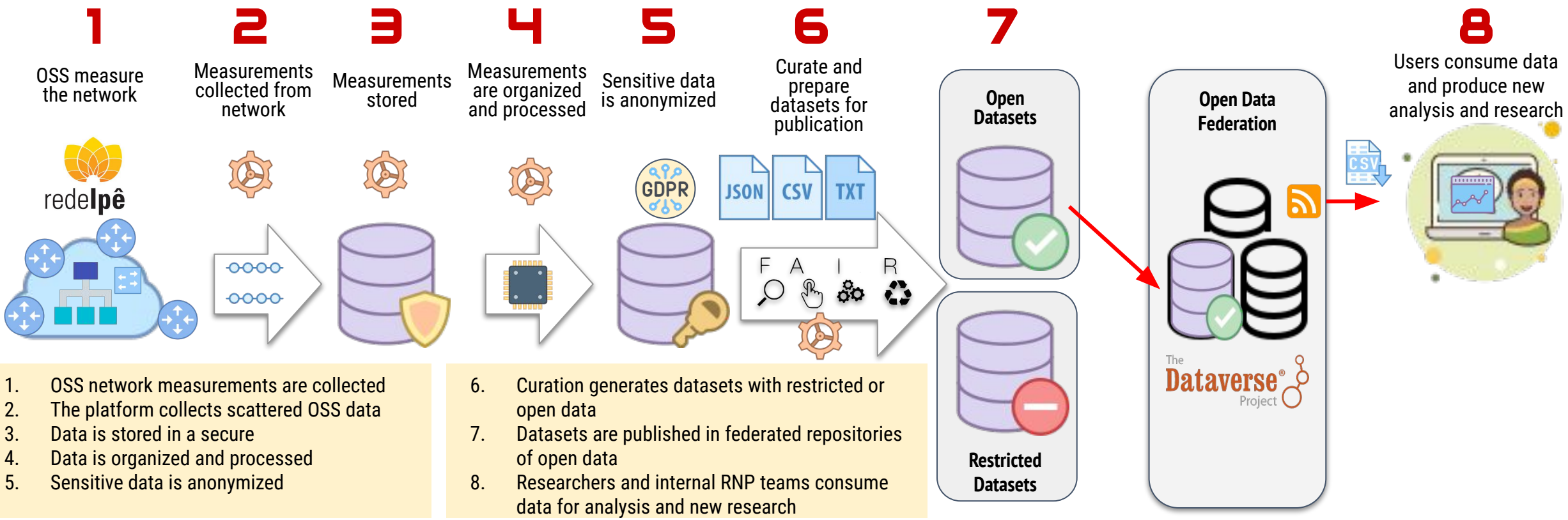
RNP Monitoring Technical Commitee mapped our restrictions and limitations

- Lack of a Data Management Plan, methods and specific processes and procedures for dataset sharing
- More specifically, the current limitations mapped at RNP (Brazilian NREN) are:
  - process with SLA to allow access to datasets of SNMP or flows and equipment configuration
  - proper methods to obfuscate data
  - documentation and metadata for the proper interpretation of collected data
  - capacity to deliver periodic data
  - process to allow access to data from TTS (Trouble Ticketing System)
  - process and tools to obfuscate sensitive information in TTS
  - process and infrastructure to collect and share network topology, telemetry and routing data sets
- Operators that often are not willing to collaborate in sharing raw OSS network data, because of security or privacy concerns

MicroMon

Automated OSS Network Data Sharing for Research

MicroMon: an automated microservices platform to collect, store, organize, anonymize and share OSS network measurement datasets with researchers



1. OSS network measurements are collected
2. The platform collects scattered OSS data
3. Data is stored in a secure
4. Data is organized and processed
5. Sensitive data is anonymized

6. Curation generates datasets with restricted or open data
7. Datasets are published in federated repositories of open data
8. Researchers and internal RNP teams consume data for analysis and new research

Challenges

The CT-Mon mapped the following challenges

- How to balance investments for data acquisition in high speed and large volume network flows (100Gbps and above) and high precision measurements in a data storage? How to finance a long term, accessible data store capable of large volume network data?
- How to define useful data filtering strategies?
- How to ensure data sharing in compliance with local laws and regulations like GDPR?
- How to ensure compliance and anonymization that will not compromise data usefulness?
- How to define an interaction model so that the sharers have an interest (and incentive) to share?
- How to promote open data sharing of some part of network OSS aggregated data that would not compromise operations security to leverage network research?

Thank You!

CT-Mon
Monitoring Technical Commitee
http://wiki.rnp.br/display/ctmon

Aviso Legal
Este documento foi preparado no âmbito de trabalho financiado pela RNP. Nem a RNP nem o governo brasileiro, nem qualquer um de seus funcionários fazem qualquer garantia, expressa ou implícita, ou assumem qualquer responsabilidade legal pela precisão, integridade ou utilidade de qualquer informação, aparelho, produto ou processo divulgado ou representa que seu uso não infringe direitos de propriedade privada. A referência aqui contida a qualquer produto, processo ou serviço comercial específico por nome comercial, marca comercial, fabricante ou de outra forma não constitui necessariamente ou implica seu endosso, recomendação ou favorecimento pela RNP ou pelo governo brasileiro. As visões e opiniões dos autores aqui expressos não necessariamente indicam ou refletem as do governo brasileiro ou da RNP, e não devem ser usadas para fins de publicidade ou endosso de produto.

Disclaimer
This document was prepared as an account of work sponsored by RNP. Neither the Brazilian government nor RNP, nor any of their employees makes any warranty, expressed or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the Brazilian government or RNP. The views and opinions of authors expressed herein do not necessarily state or reflect those of the Brazilian government or RNP, and shall not be used for advertising or product endorsement purposes