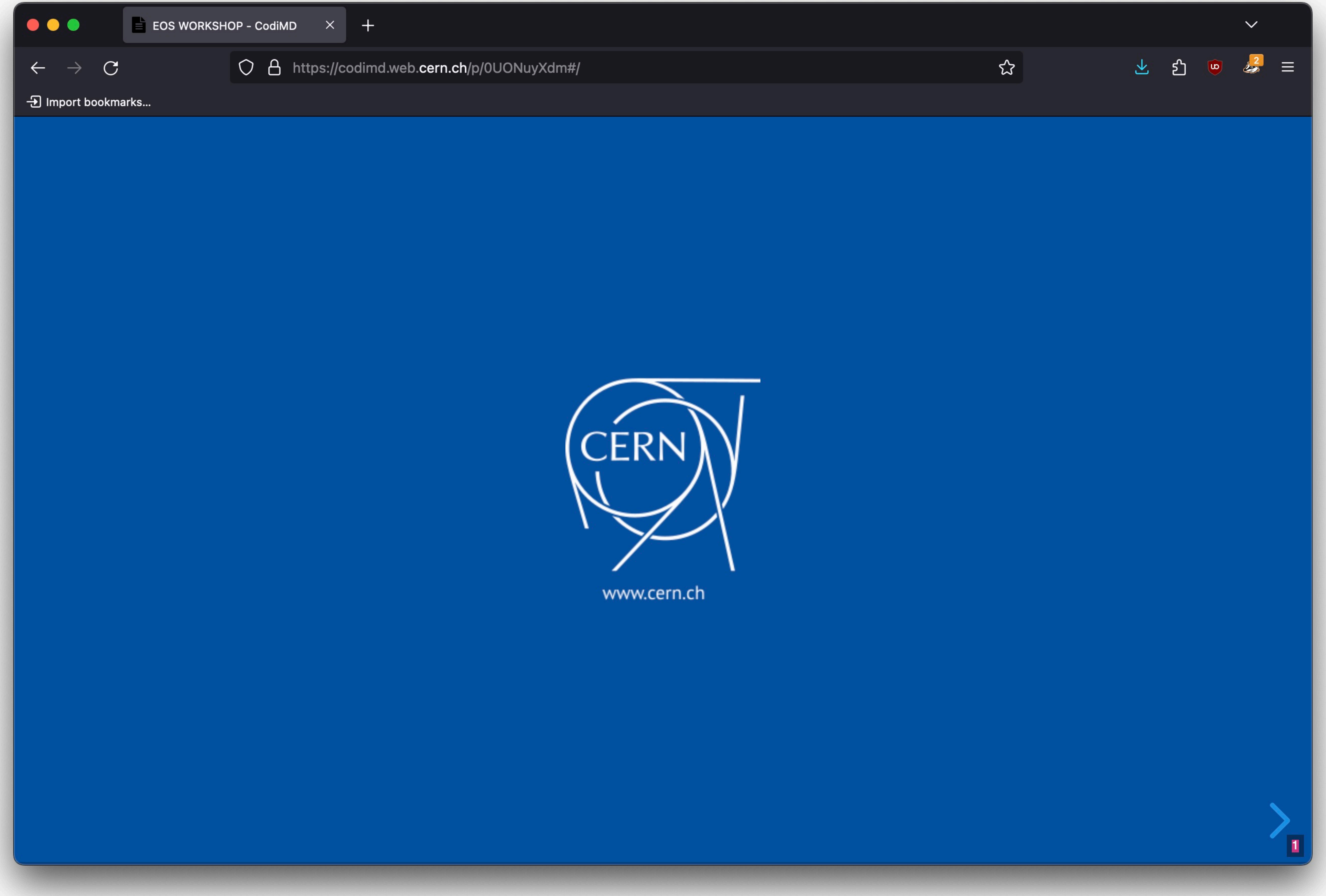
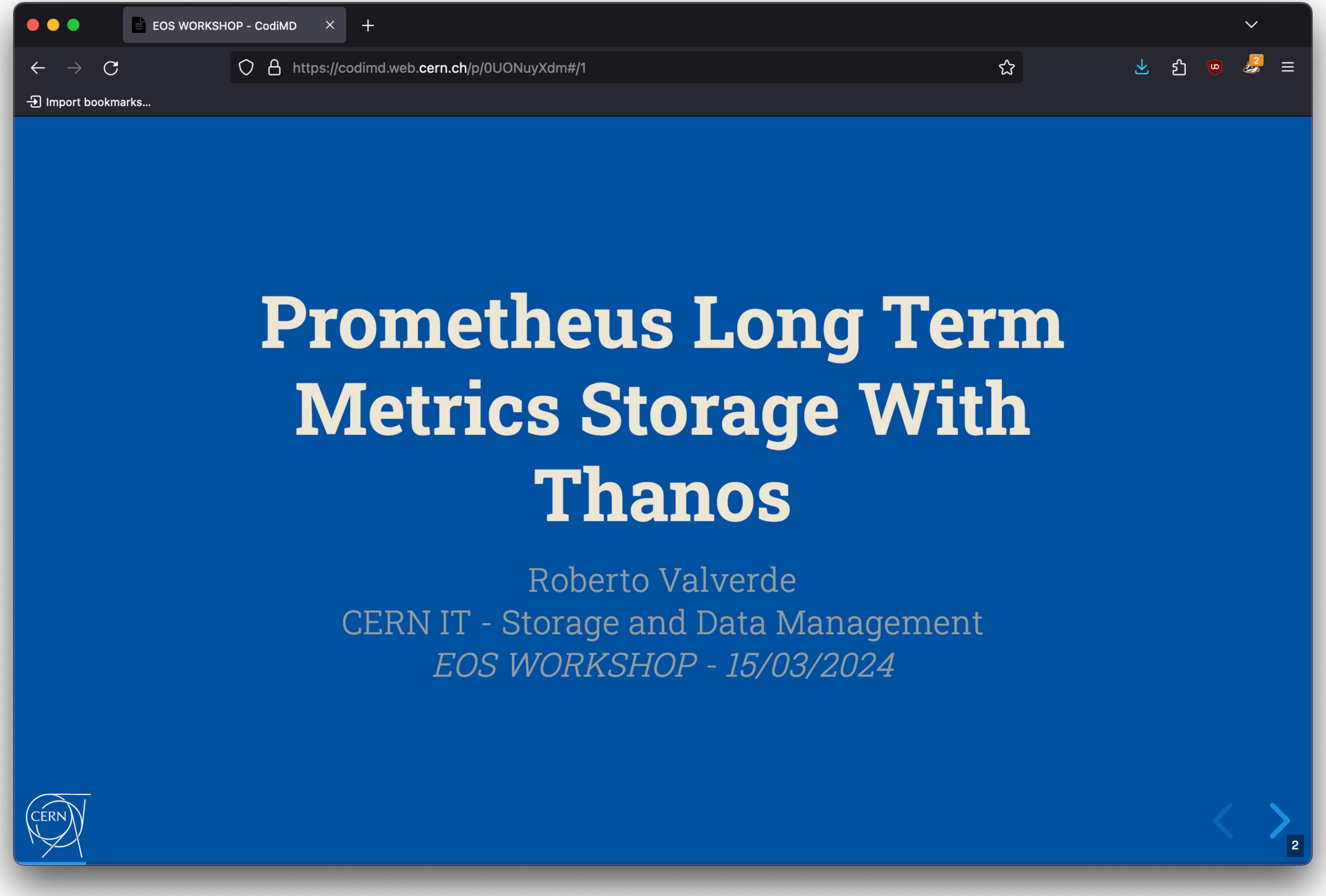

EOS WORKSHOP - CodiMD
https://codimd.web.cern.ch/p/0UONuyXdm#/
Import bookmarks...
CERN
www.cern.ch
1



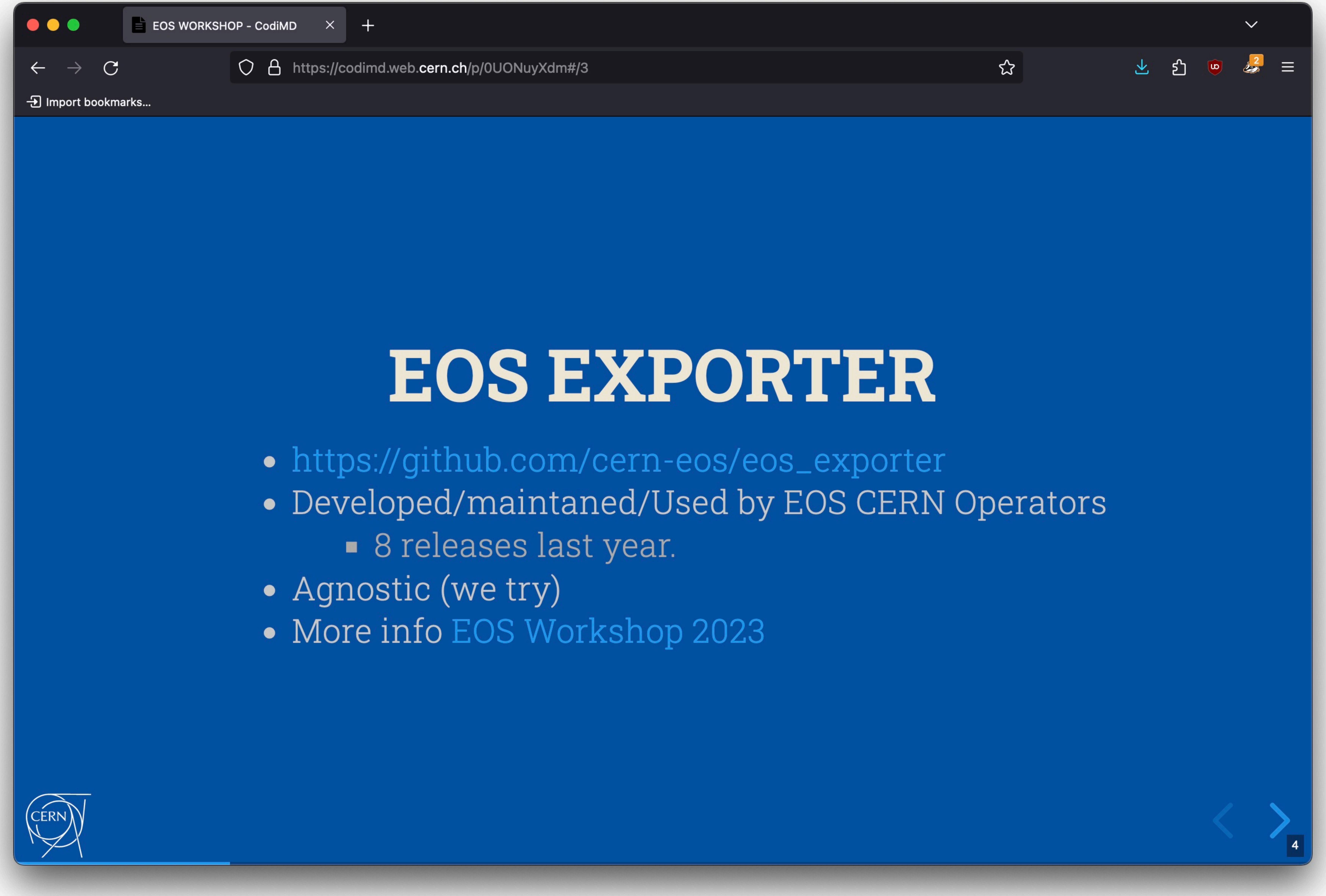
Prometheus Long Term Metrics Storage With Thanos

Roberto Valverde
CERN IT - Storage and Data Management

EOS WORKSHOP - 15/03/2024

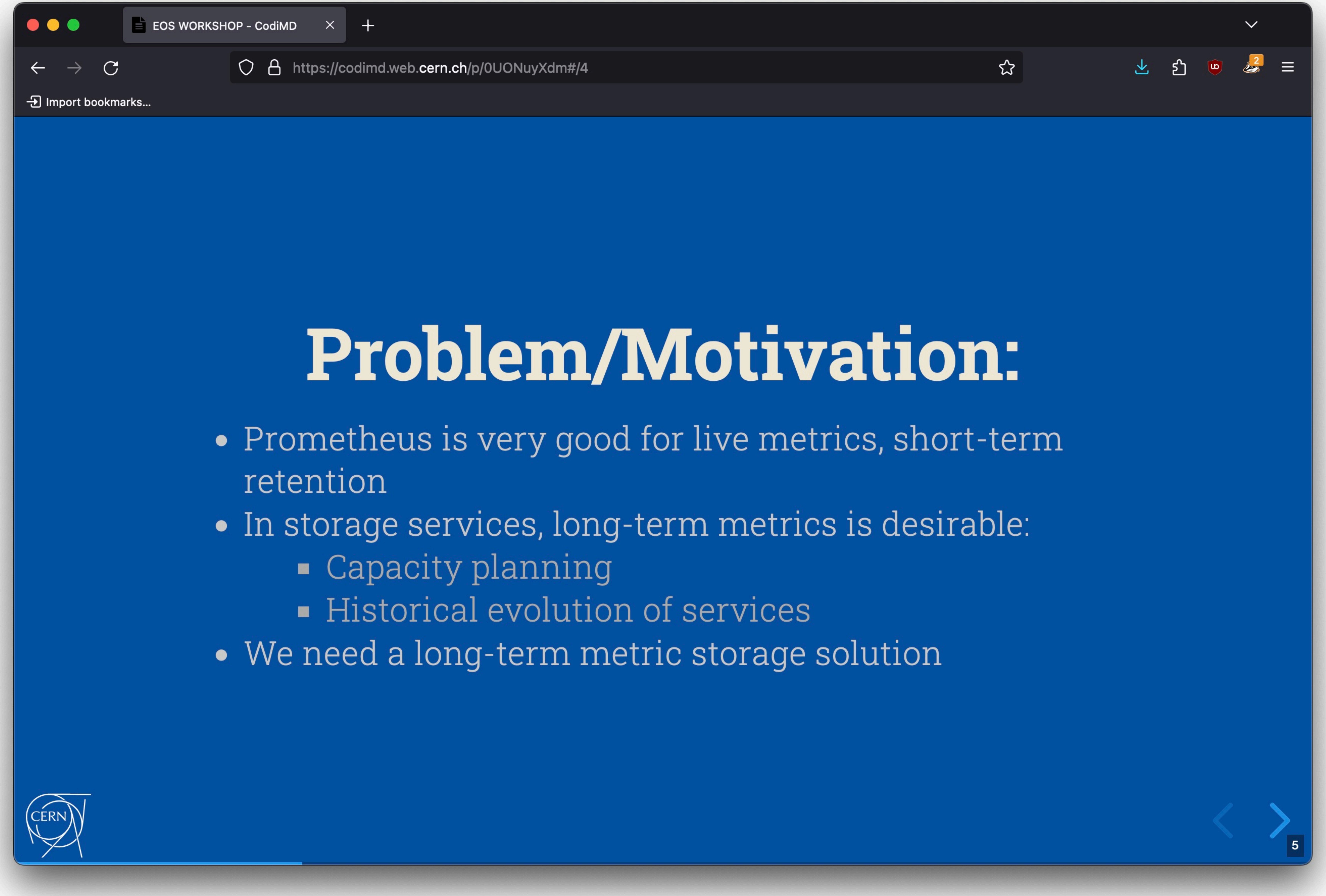
Prometheus

- In Storage and Data Management group we rely on prometheus for many services:
  - Ceph, EOS, cback...
- Pull-based, time-series data model
- PromQL query language: `count by(cluster, eos_version)(eos_node_info)`
- Central server that `scrapes` targets, called `exporters`
  - There is an exporter for that! exporters
  - And there is also an EOS exporter!



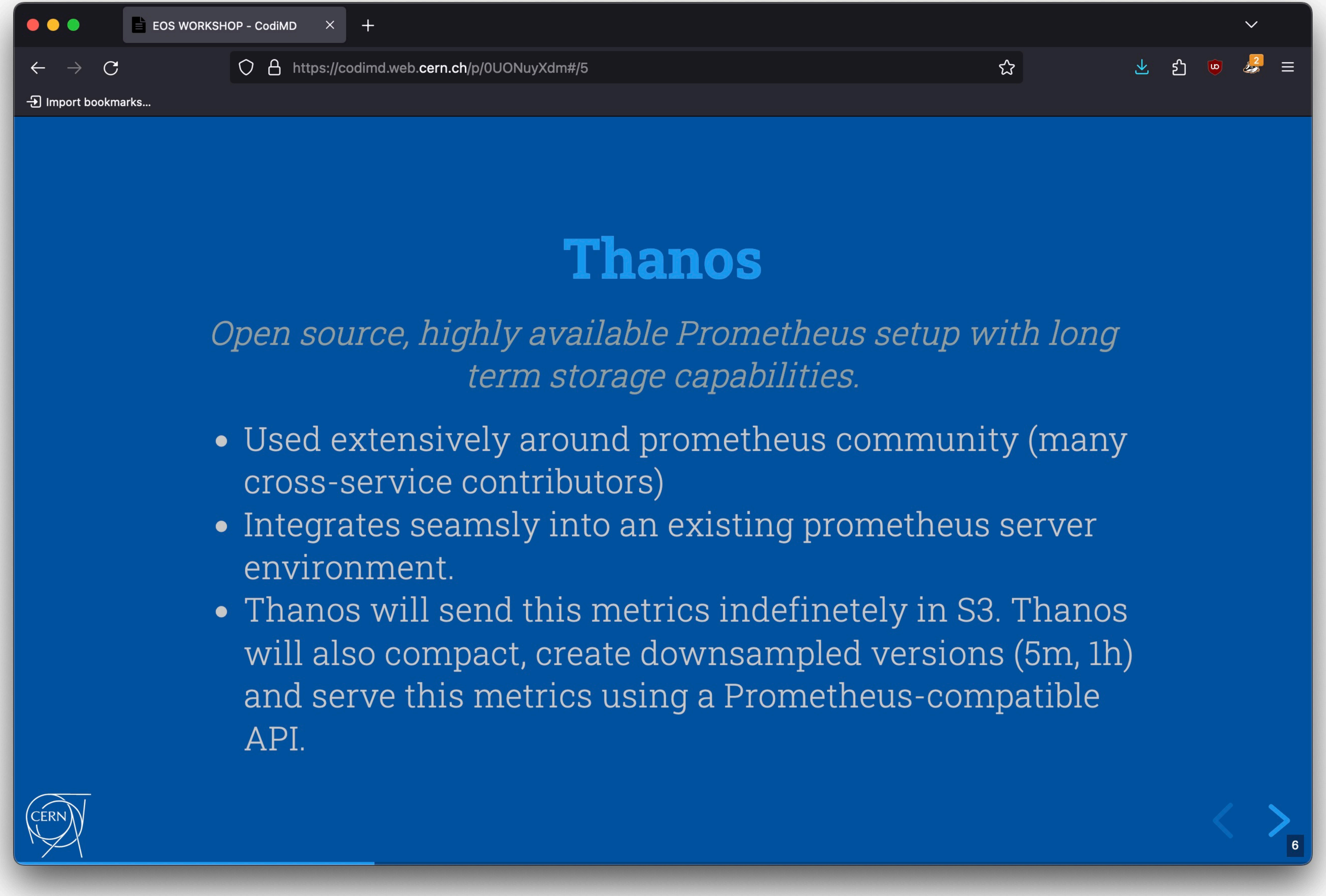
EOS EXPORTER

- https://github.com/cern-eos/eos_exporter
- Developed/maintaned/Used by EOS CERN Operators
    - 8 releases last year.
- Agnostic (we try)
- More info EOS Workshop 2023



Problem/Motivation:

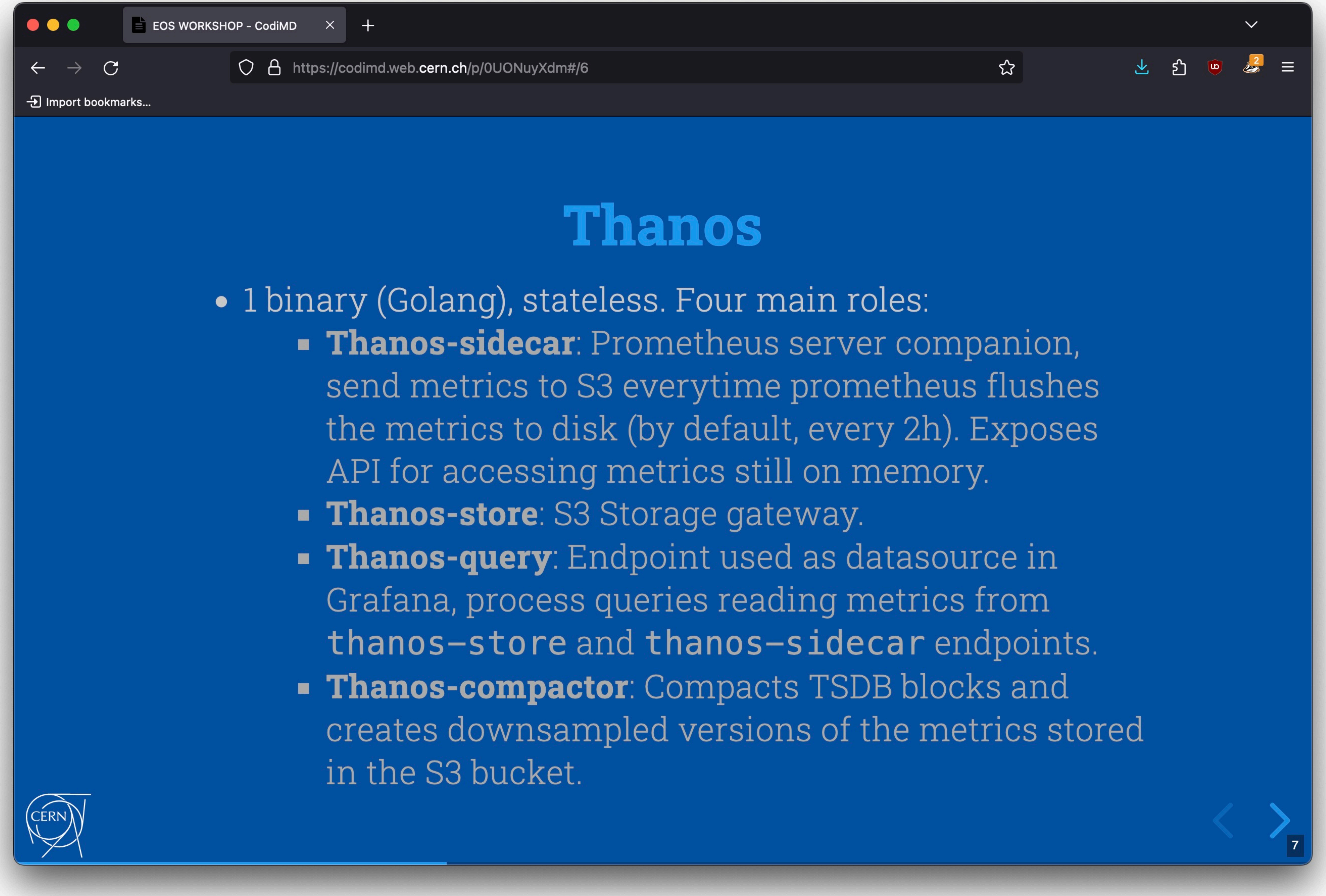
- Prometheus is very good for live metrics, short-term retention
- In storage services, long-term metrics is desirable:
  - Capacity planning
  - Historical evolution of services
- We need a long-term metric storage solution



Thanos

Open source, highly available Prometheus setup with long term storage capabilities.

- Used extensively around prometheus community (many cross-service contributors)
- Integrates seamsly into an existing prometheus server environment.
- Thanos will send this metrics indefinetely in S3. Thanos will also compact, create downsampled versions (5m, 1h) and serve this metrics using a Prometheus-compatible API.



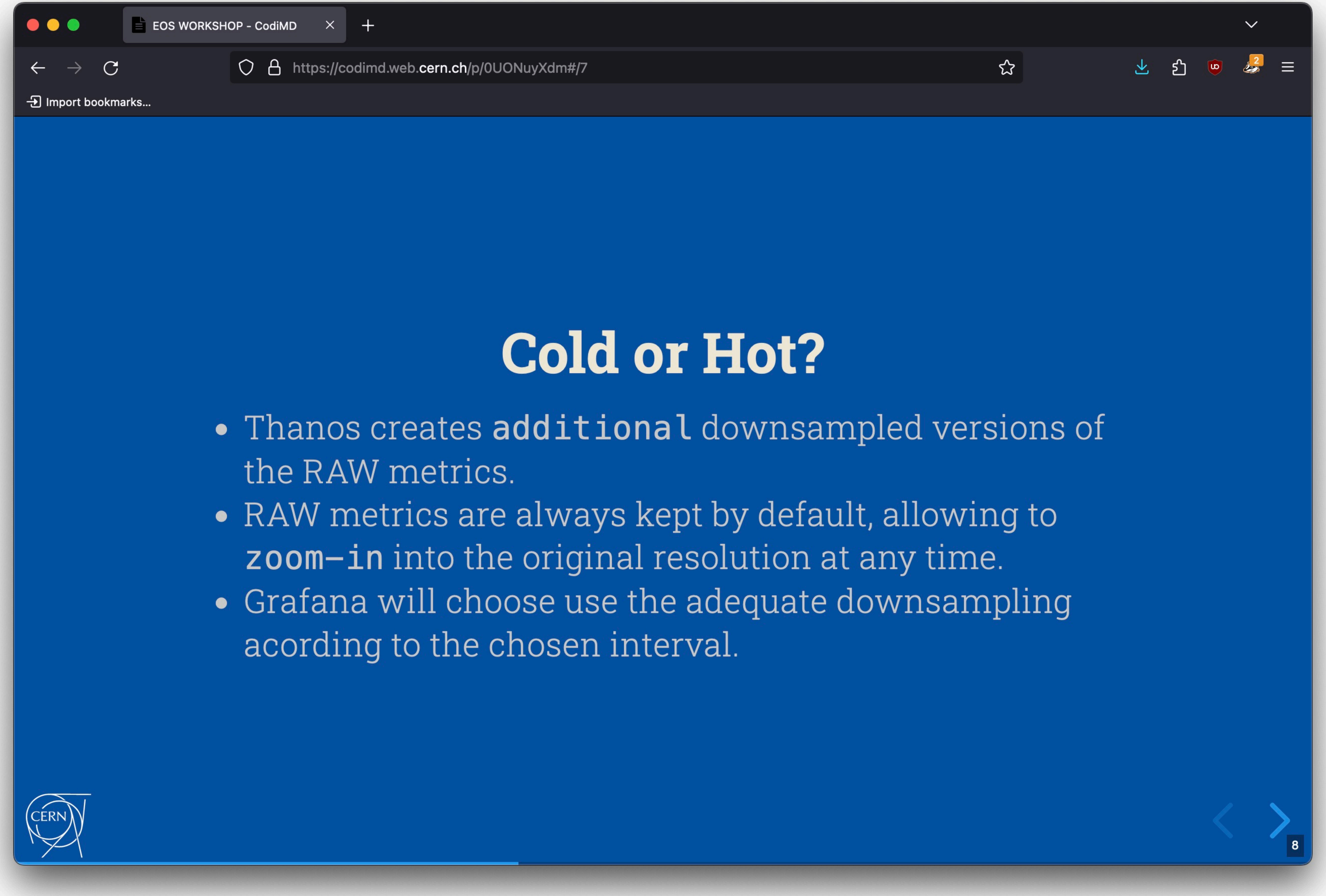
Thanos

- 1 binary (Golang), stateless. Four main roles:
    - **Thanos-sidecar**: Prometheus server companion, send metrics to S3 everytime prometheus flushes the metrics to disk (by default, every 2h). Exposes API for accessing metrics still on memory.
    - **Thanos-store**: S3 Storage gateway.
    - **Thanos-query**: Endpoint used as datasource in Grafana, process queries reading metrics from `thanos-store` and `thanos-sidecar` endpoints.
    - **Thanos-compactor**: Compacts TSDB blocks and creates downsampled versions of the metrics stored in the S3 bucket.

LibreWolf File Edit View History Bookmarks Tools Window Help
Fri 15 Mar 16:52:45
EOS WORKSHOP - CodiMD
https://codimd.web.cern.ch/p/0UONuyXdm#/7
Import bookmarks...
Start call
lient_id: 356566590
thanos4.gif
Screenshot 2024-0...8.19.png
Screenshot 2024-0...11.32.png
Screenshot 2024-0....11.51.png
Screen Recording 2024-0...3.34.mov
Screen Recording 2024-0...6.08.mov
thanos.mov
thanos.gif
thanos2.gif
CERN
8

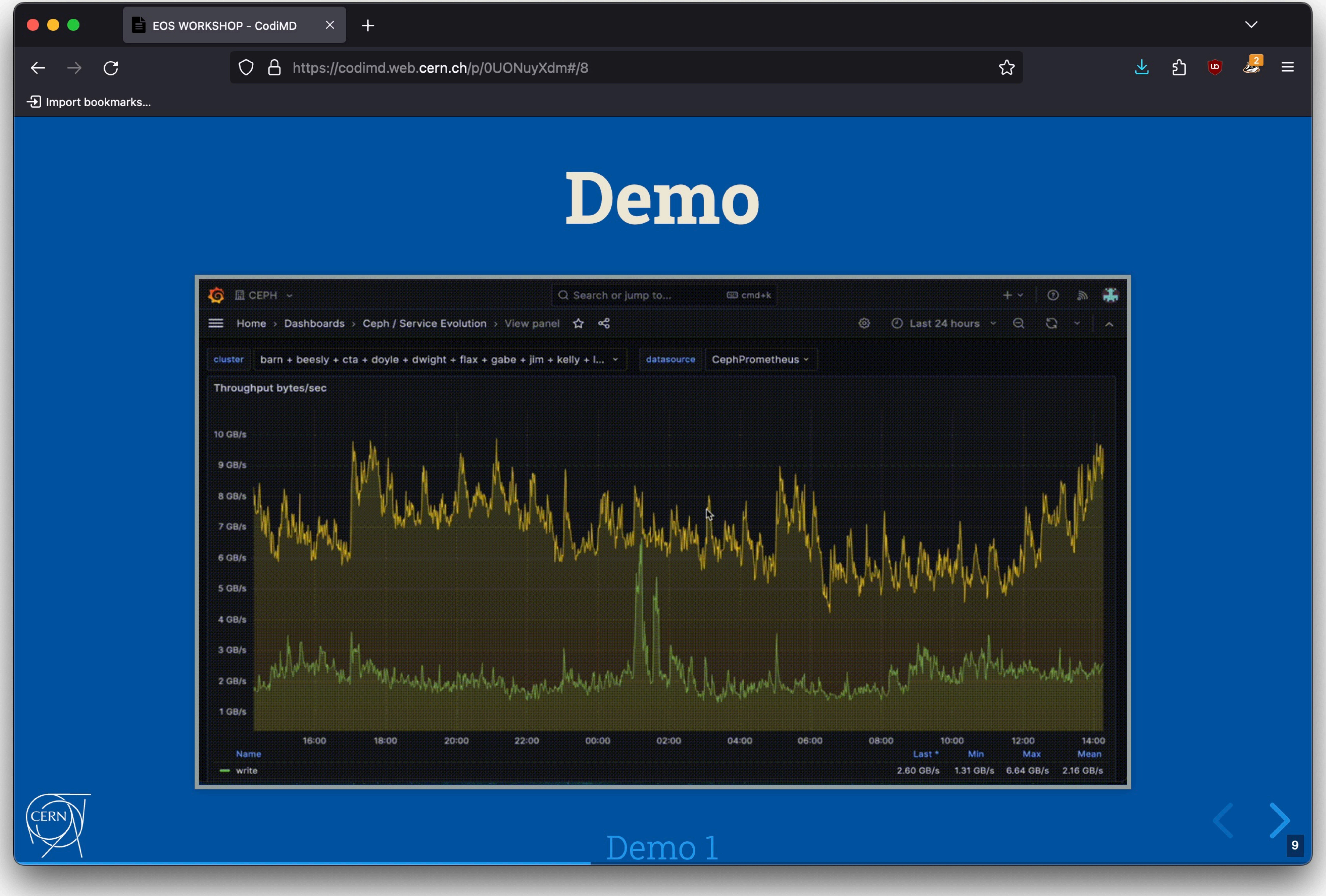
Cold or Hot?

- Thanos creates `additional` downsampled versions of the RAW metrics.
- RAW metrics are always kept by default, allowing to `zoom-in` into the original resolution at any time.
- Grafana will choose use the adequate downsampling acording to the chosen interval.



Cold or Hot?

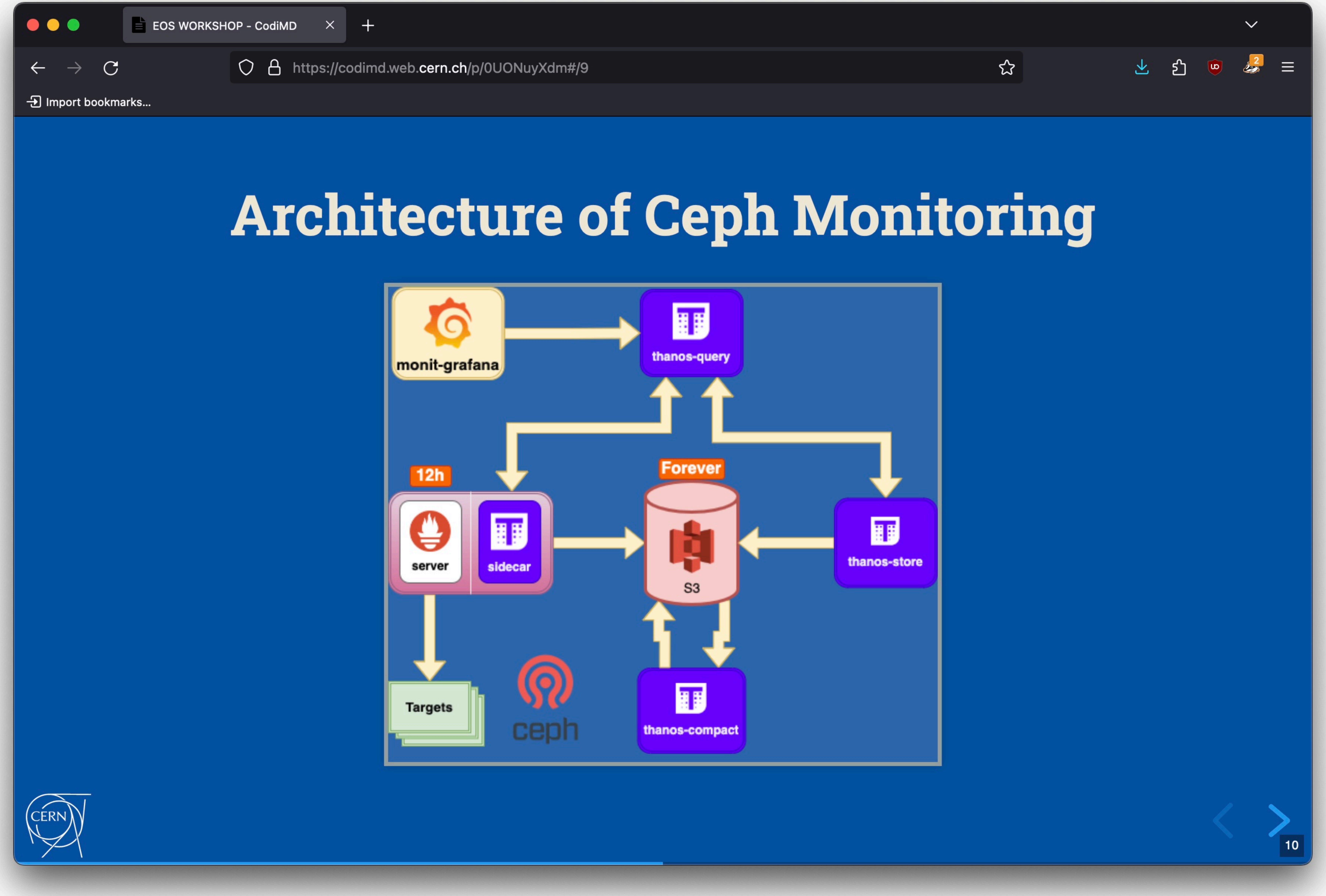
- Thanos creates `additional` downsampled versions of the RAW metrics.
- RAW metrics are always kept by default, allowing to `zoom-in` into the original resolution at any time.
- Grafana will choose use the adequate downsampling acording to the chosen interval.



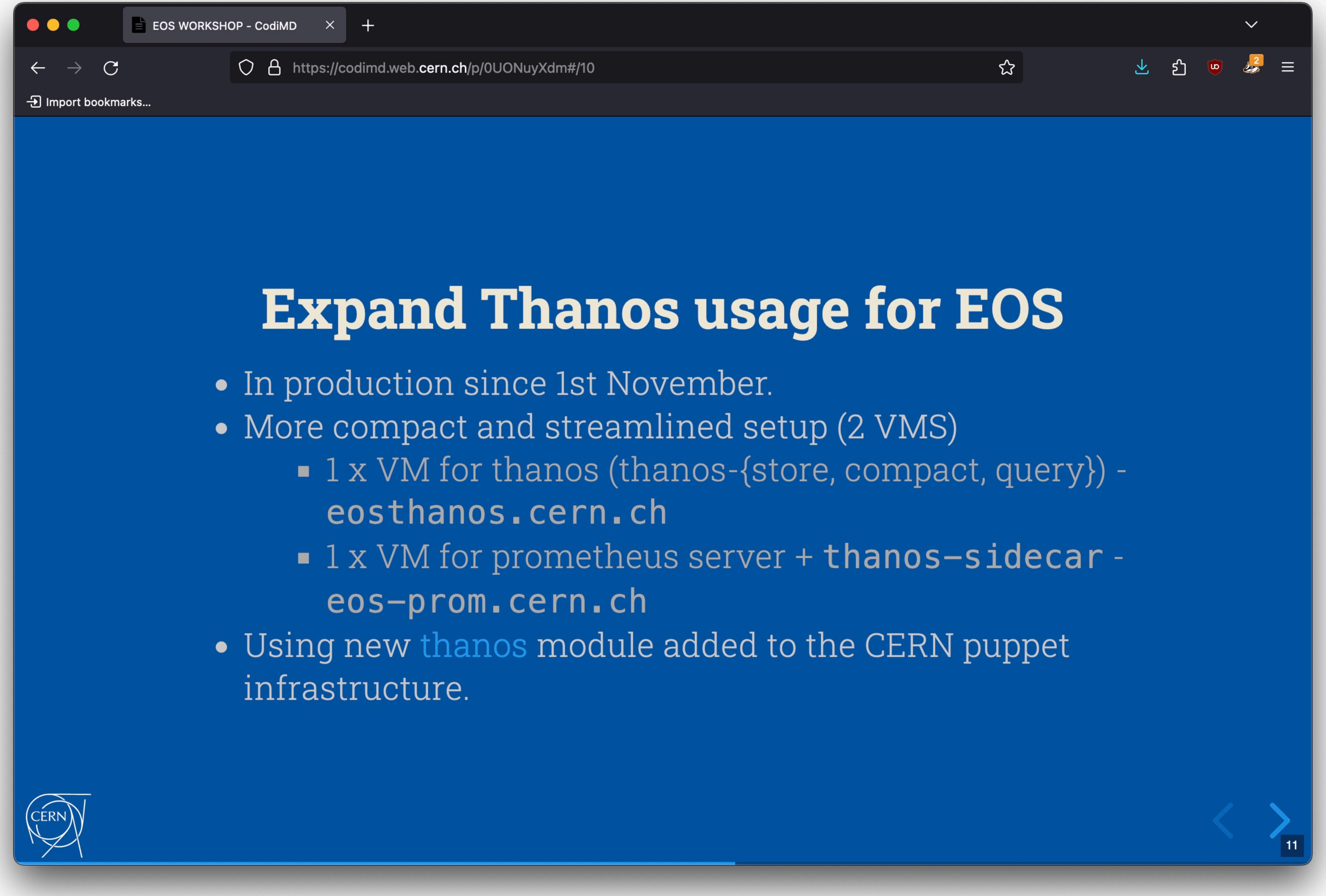
Demo

Name	Last *	Min	Max	Mean
write	2.60 GB/s	1.31 GB/s	6.64 GB/s	2.16 GB/s

Demo 1

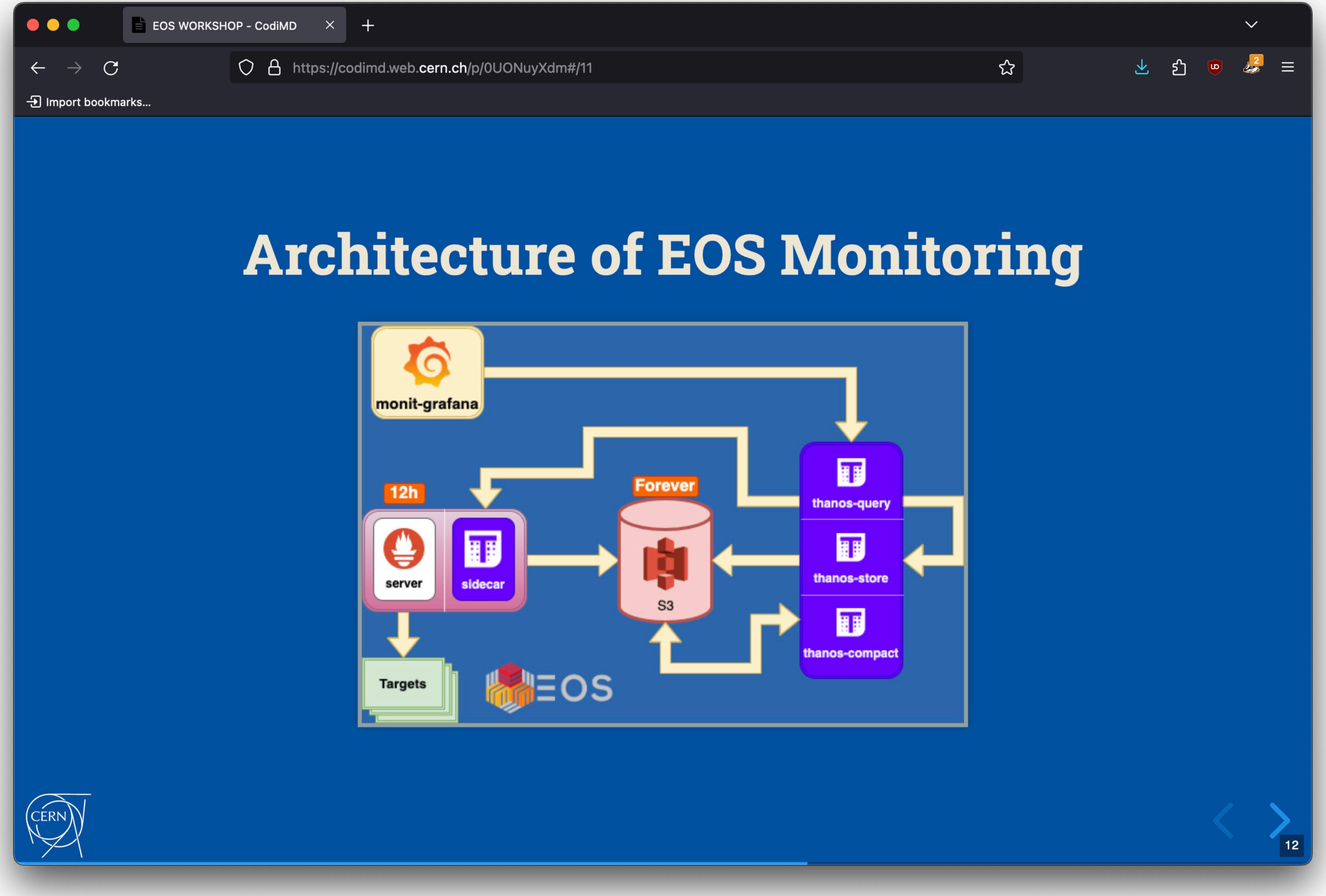


Architecture of Ceph Monitoring

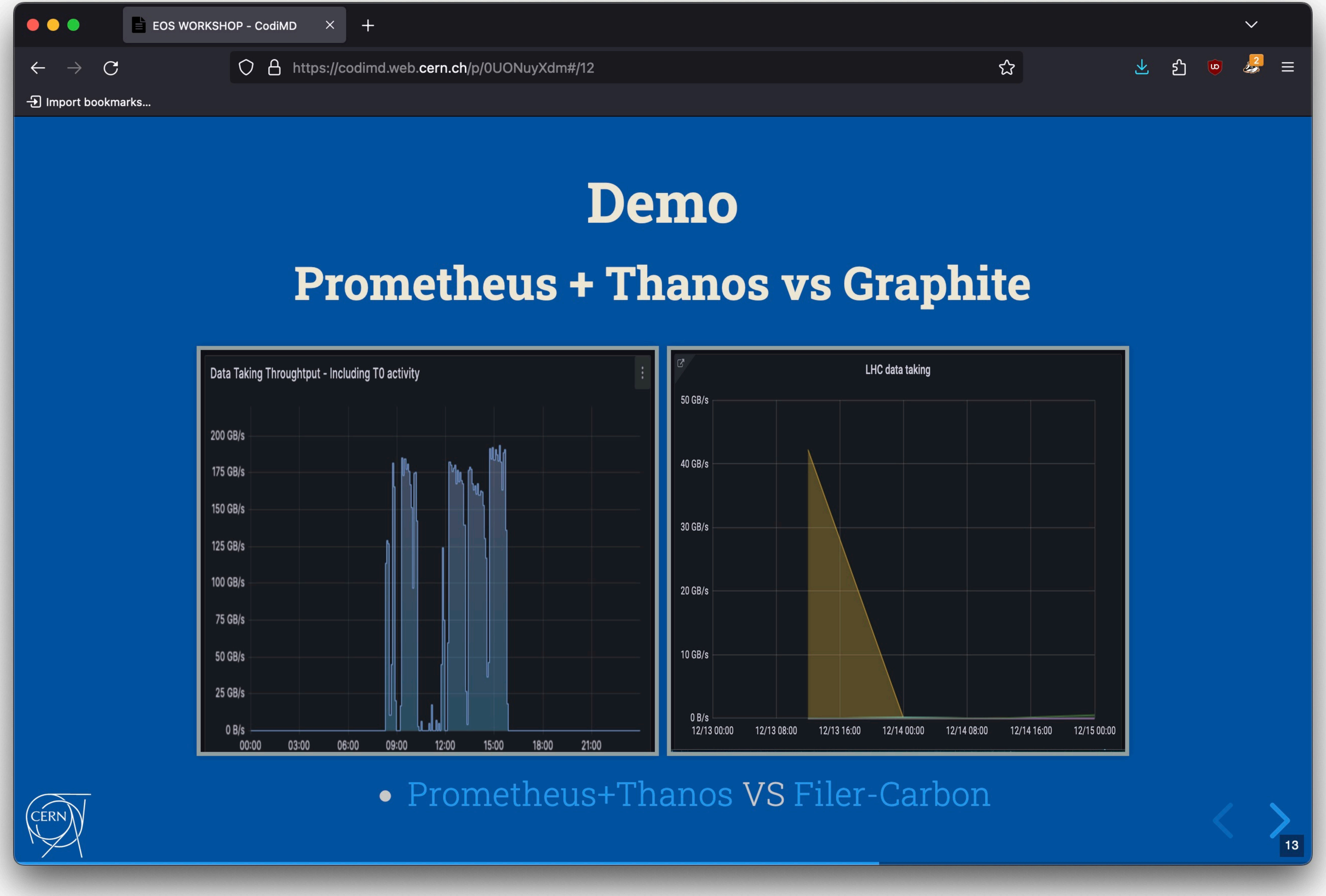


Expand Thanos usage for EOS

- In production since 1st November.
- More compact and streamlined setup (2 VMS)
  - 1 x VM for thanos (thanos-{store, compact, query}) - `eosthanos.cern.ch`
  - 1 x VM for prometheus server + `thanos-sidecar` - `eos-prom.cern.ch`
- Using new thanos module added to the CERN puppet infrastructure.



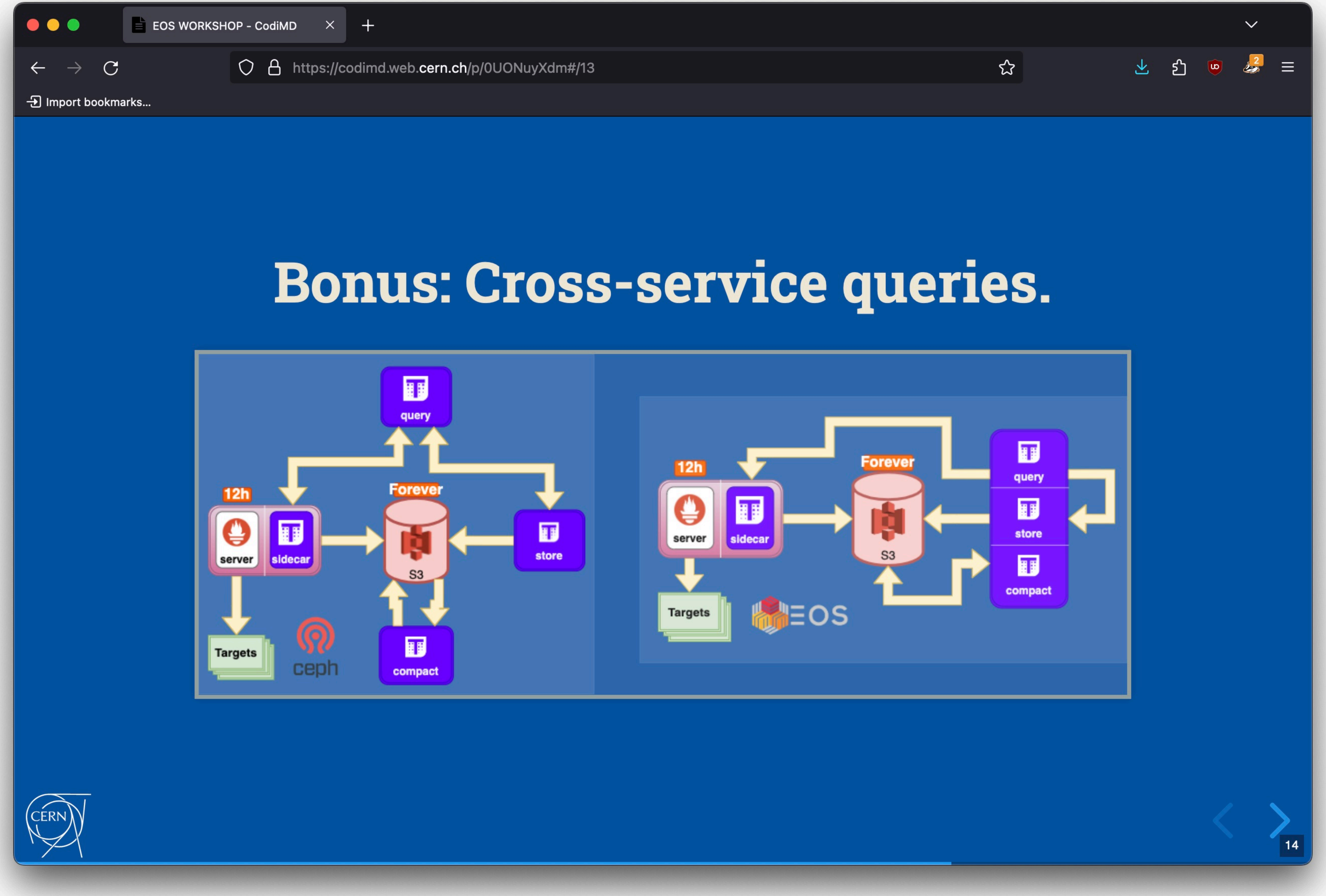
Architecture of EOS Monitoring



Demo

Prometheus + Thanos vs Graphite

- Prometheus+Thanos VS Filer-Carbon

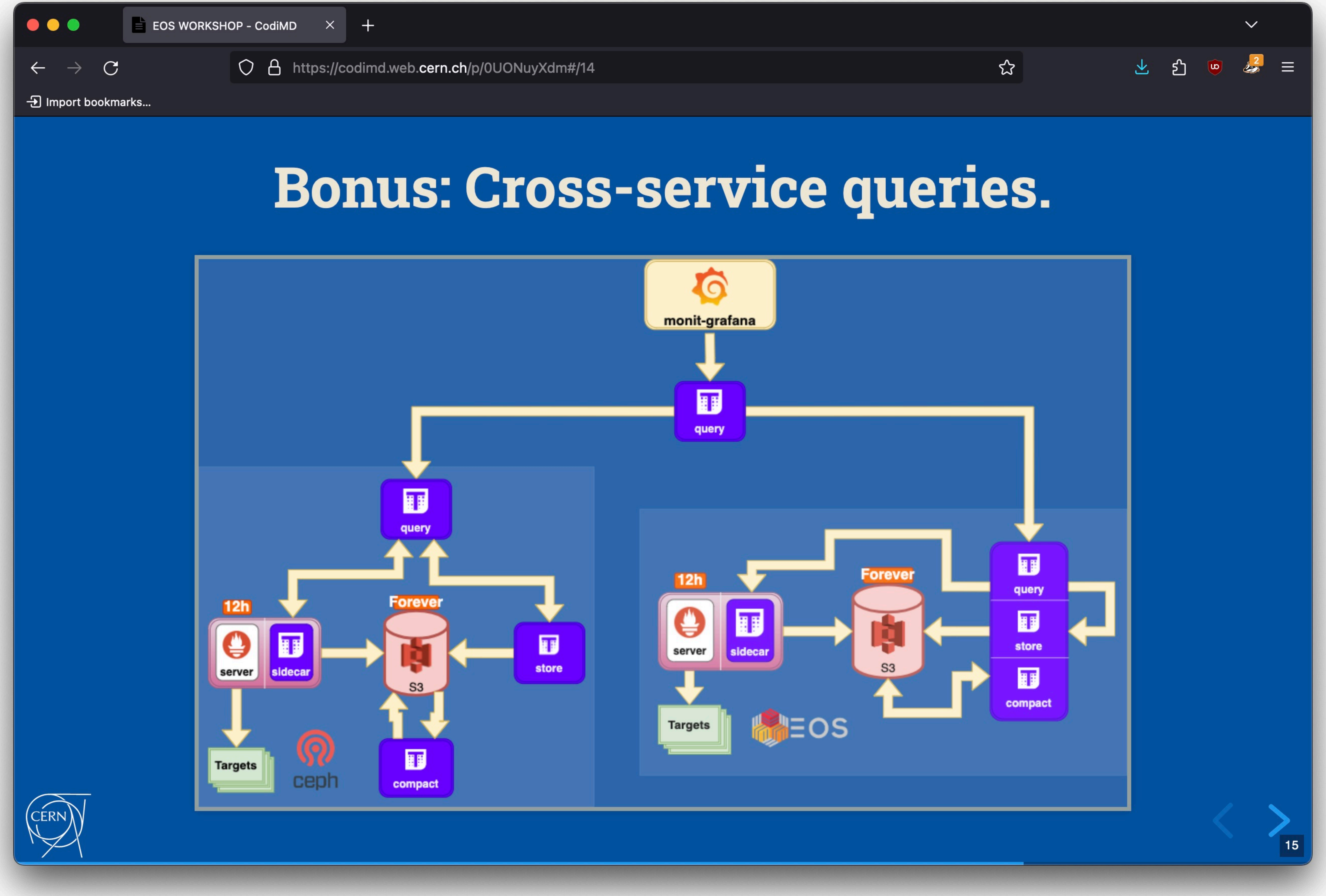


EOS WORKSHOP - CodiMD

https://codimd.web.cern.ch/p/0UONuyXdm#/13

Import bookmarks...

Bonus: Cross-service queries.



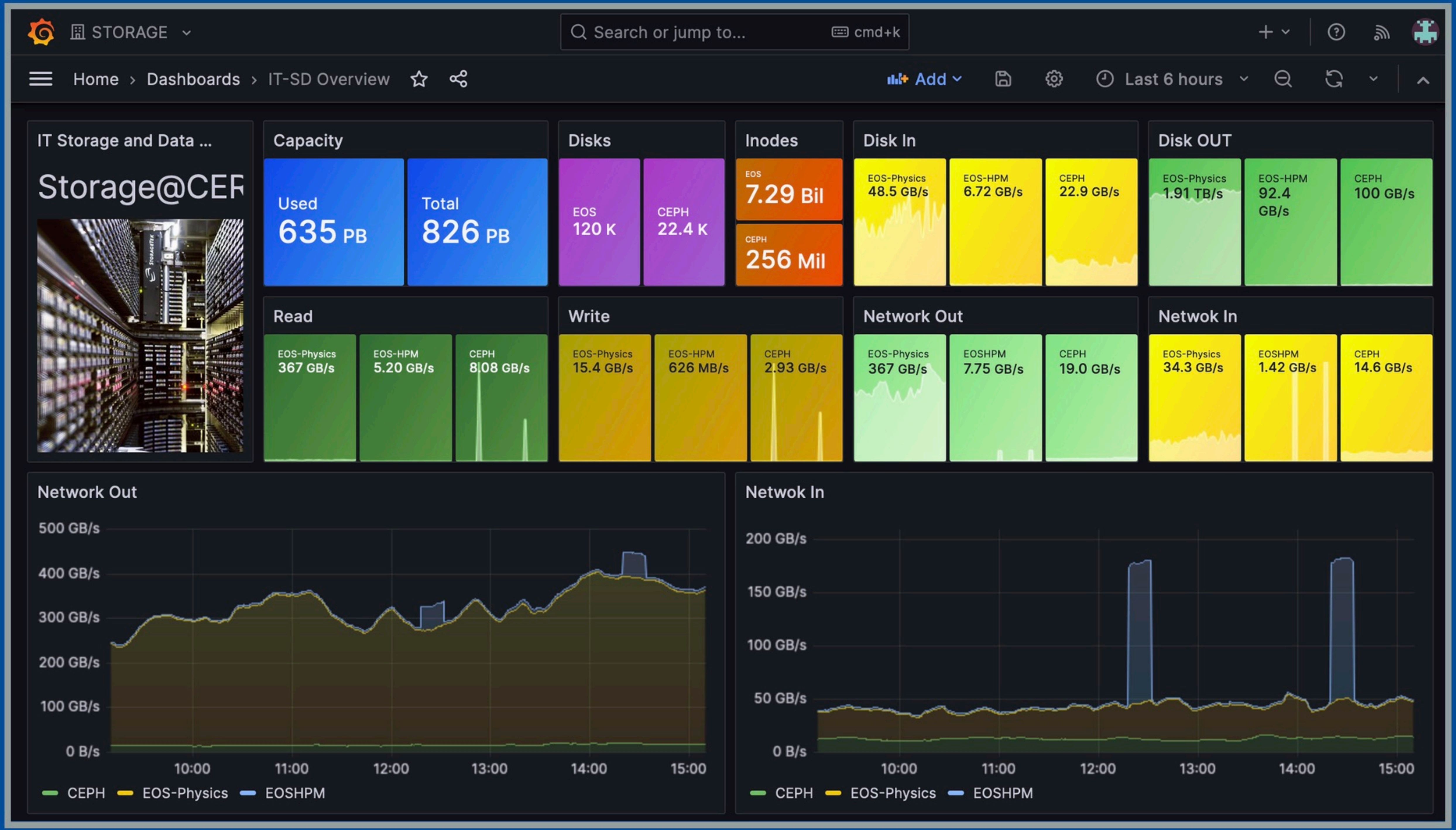
Bonus: Cross-service queries.

EOS WORKSHOP - CodiMD

https://codimd.web.cern.ch/p/0UONuyXdm#/15

Import bookmarks...

Bonus: Cross-service queries.



IT Storage and Data ...

Storage@CEF

Capacity: Used 635 PB; Total 826 PB

Disks: EOS 120 K; CEPH 22.4 K

Inodes: EOS 7.29 Bil; CEPH 256 Mil

Disk In: EOS-Physics 48.5 GB/s; EOS-HPM 6.72 GB/s; CEPH 22.9 GB/s

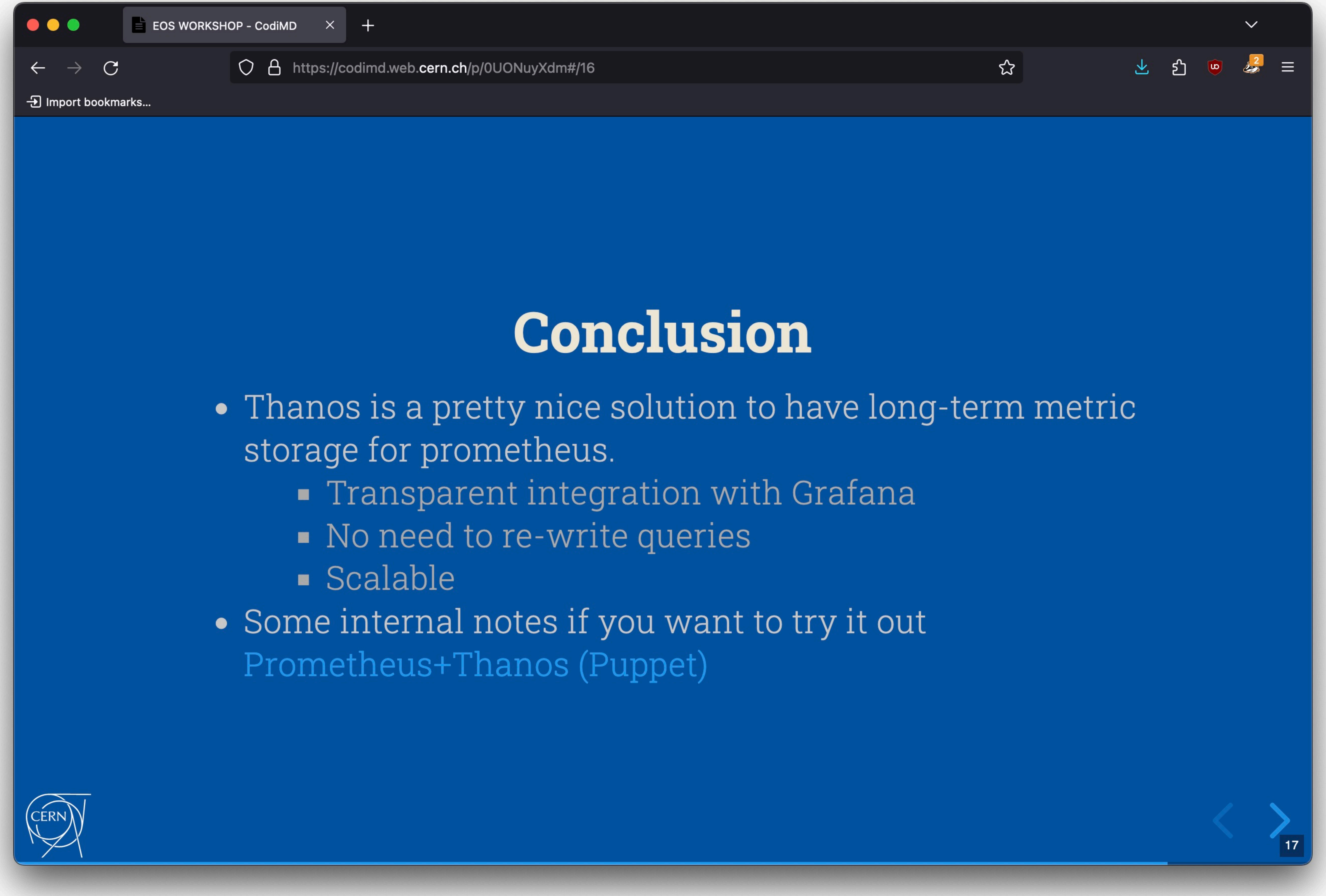
Disk OUT: EOS-Physics 1.91 TB/s; EOS-HPM 92.4 GB/s; CEPH 100 GB/s

Read: EOS-Physics 367 GB/s; EOS-HPM 5.20 GB/s; CEPH 8.08 GB/s

Write: EOS-Physics 15.4 GB/s; EOS-HPM 626 MB/s; CEPH 2.93 GB/s

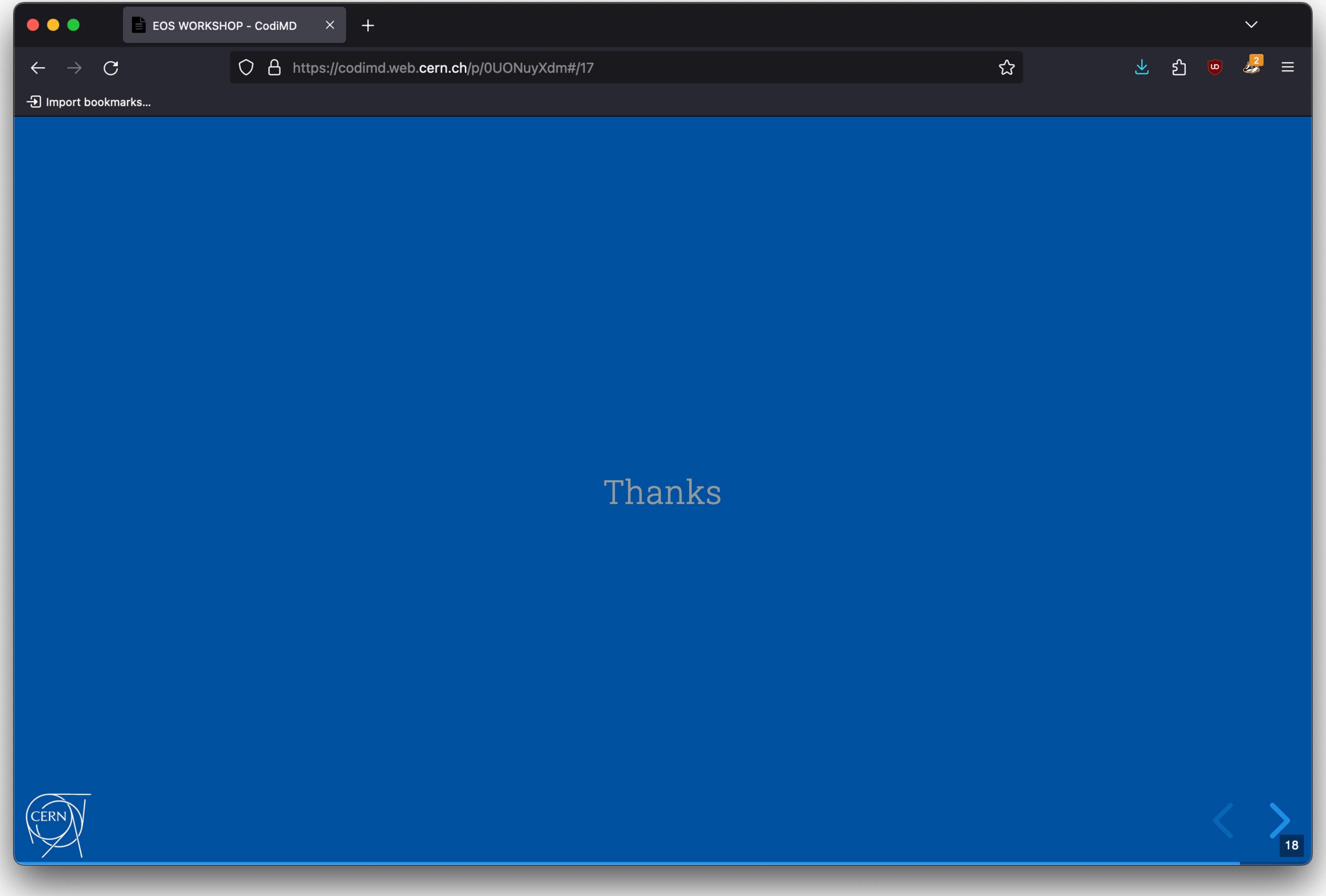
Network Out: EOS-Physics 367 GB/s; EOSHPM 7.75 GB/s; CEPH 19.0 GB/s

Netwok In: EOS-Physics 34.3 GB/s; EOSHPM 1.42 GB/s; CEPH 14.6 GB/s



Conclusion

- Thanos is a pretty nice solution to have long-term metric storage for prometheus.
  - Transparent integration with Grafana
  - No need to re-write queries
  - Scalable
- Some internal notes if you want to try it out Prometheus+Thanos (Puppet)


EOS WORKSHOP - CodiMD
https://codimd.web.cern.ch/p/0UONuyXdm#/17
Import bookmarks...
CERN
18

Thanks