Elettra Sincrotrone Trieste

Computing and Storage Architecture
At
Elettra-Sincrotrone Trieste

What will we talk about

- Introduction – Who, where and what?

- Infrastructure Overview

- Scientific Data Acquisition

- HPC Processing

Who, where and what?
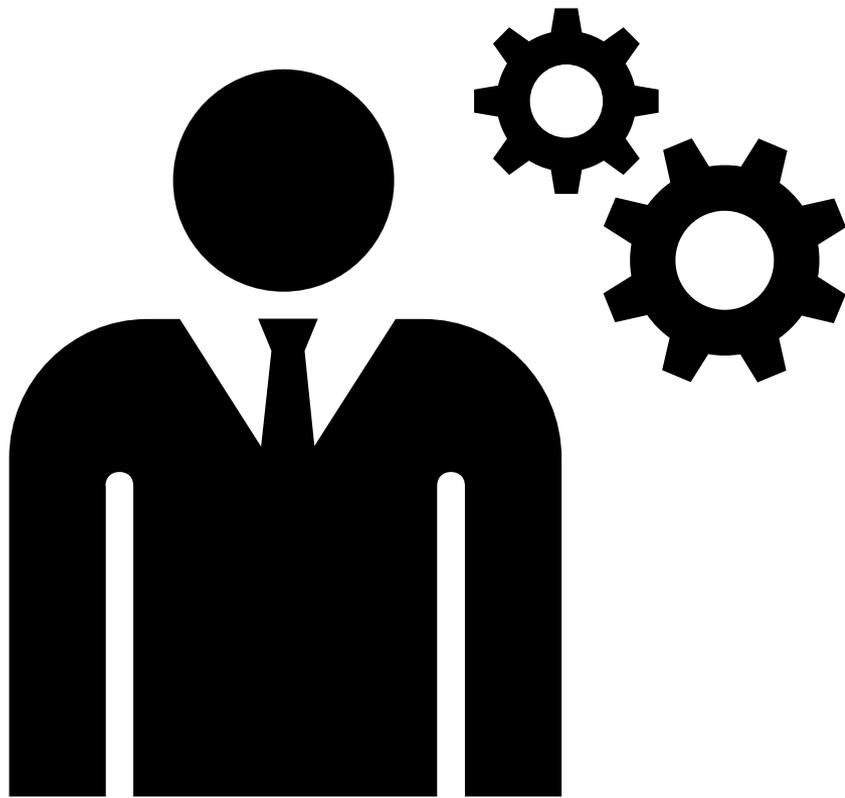
What is Elettra?

- Multidisciplinary research center located in Trieste, Italy

- Running 2 particle accelerators producing synchrotron light

- 30+ beam-lines/laboratories

- 350 people + over 1000 visiting scientists every year

Who am I?

About myself

- Linux system administrator
- Part of the ICT System and Service group in the IT department since 2017
- Working with Virtualization and Storage infrastructure, HPC and much more

Infrastructure Overview

Build on Opensource Software

- Around 99% of infrastructure and 90% of data acquisition runs on opensource software.

- Key elements:
  - Proxmox for Virtualization
  - Ceph for Storage
  - Slurm and Kubernetes for HPC

Elettra – IT Numbers

- 4 Datacenter.

- 150+ physical servers, ~1000 VM.

- 10 PB raw storage.

- Core Network at 10 Gb/s.
(updating to 100 Gb/s)

Elettra - Clusters

- 2 Hyper-converged clusters (Proxmox+Ceph) for machine controls.

- 1 Hyper-converged cluster (Proxmox+Ceph) for support services.

- 1 cluster (Proxmox and Ceph) scientific data handling.

- 1 cluster (Proxmox and Ceph) for backup.

- 1 cluster (Ceph) for S3 storage.

Scientific Data Acquisition

Scientific Data Acquisition

- Different beam-lines, different storage requirments, common acquisition pipeline.

- Pipeline managed through in-house build application (VUO).

- Data retention is 10 years.

Data Acquisition Pipeline

1. Acquired experiment data is saved on the **scratch storage**.

2. Once the acquisition is done, data is transfered to the **online storage** for later consumption.

3. Automatically and on request data is copied (or moved) to the **offline storage** (LTO tapes).

GLOSSARY

- **Scratch**
  Storage space used during acquisition

- **Online**
  Storage space used after acquisition

- **Offline**
  Space (tape) for long-term storage.

- **VUO**
  Virtual Unified Office, core application for managing data acquisition pipeline.
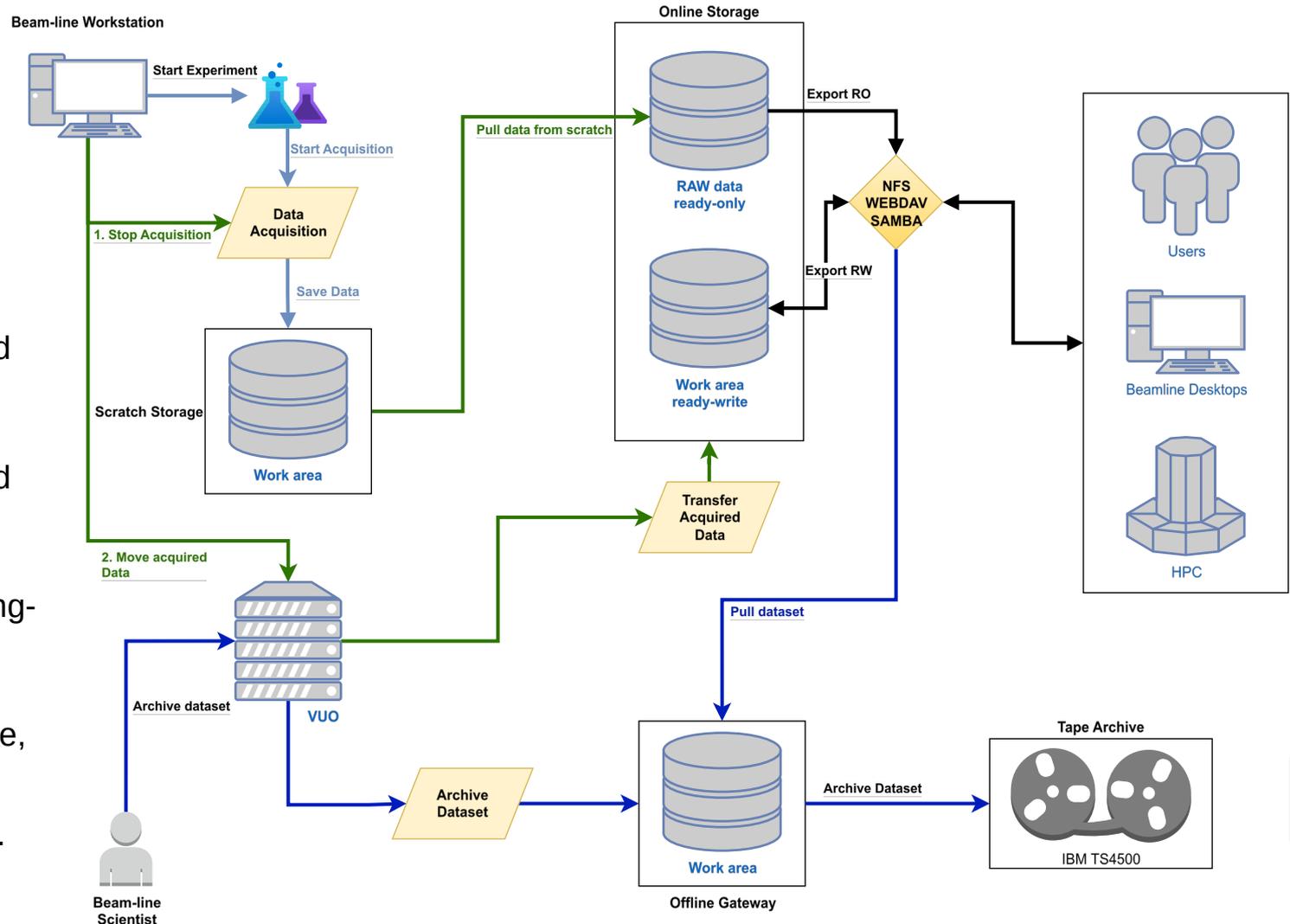
High Perfomance Computing

Elettra
Sincrotrone
Trieste

ISO 9001
OHSAS 18001
BUREAU VERITAS
Certification

Online processing

- Data Processing in-real-time (during acquisition).
- CPU & GPU processing pipeline, in-house development.
    - 3 x Nvidia A100
    - 3 x Nvidia L40S
    - 4 x Nvidia L4
    - 6 x Nvidia T4
- On-premise Kubernetes cluster as computing infrastructure.

Batch processing

- "Offline" (not in real time) processing of data.

- "Kalculus" cluster
  - Slurm batch scheduler.
  - 56 nodes, 3 partitions, 1500+ cpu cores.

QUESTIONS?

Elettra Sincrotrone Trieste

www.elettra.eu