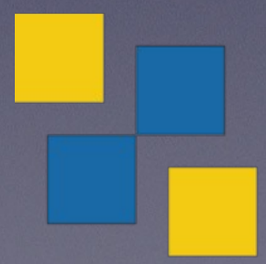
Working with BIANCA as an external collaborator

Linus Nilsson, Marcus Lundberg

UPPMAX, Uppsala University, Sweden

uppmax.uu.se

@UPPMAX



SNIC

UPPMAX

HPC-center at Uppsala University, Sweden

UPPMAX specializes in large-scale analysis of sensitive personal data for **academic** use

- Many users from the life-sciences, some from humanities, a few from the county councils and agencies
- Develop and maintain high-performance computing platforms, services, advanced user support and education

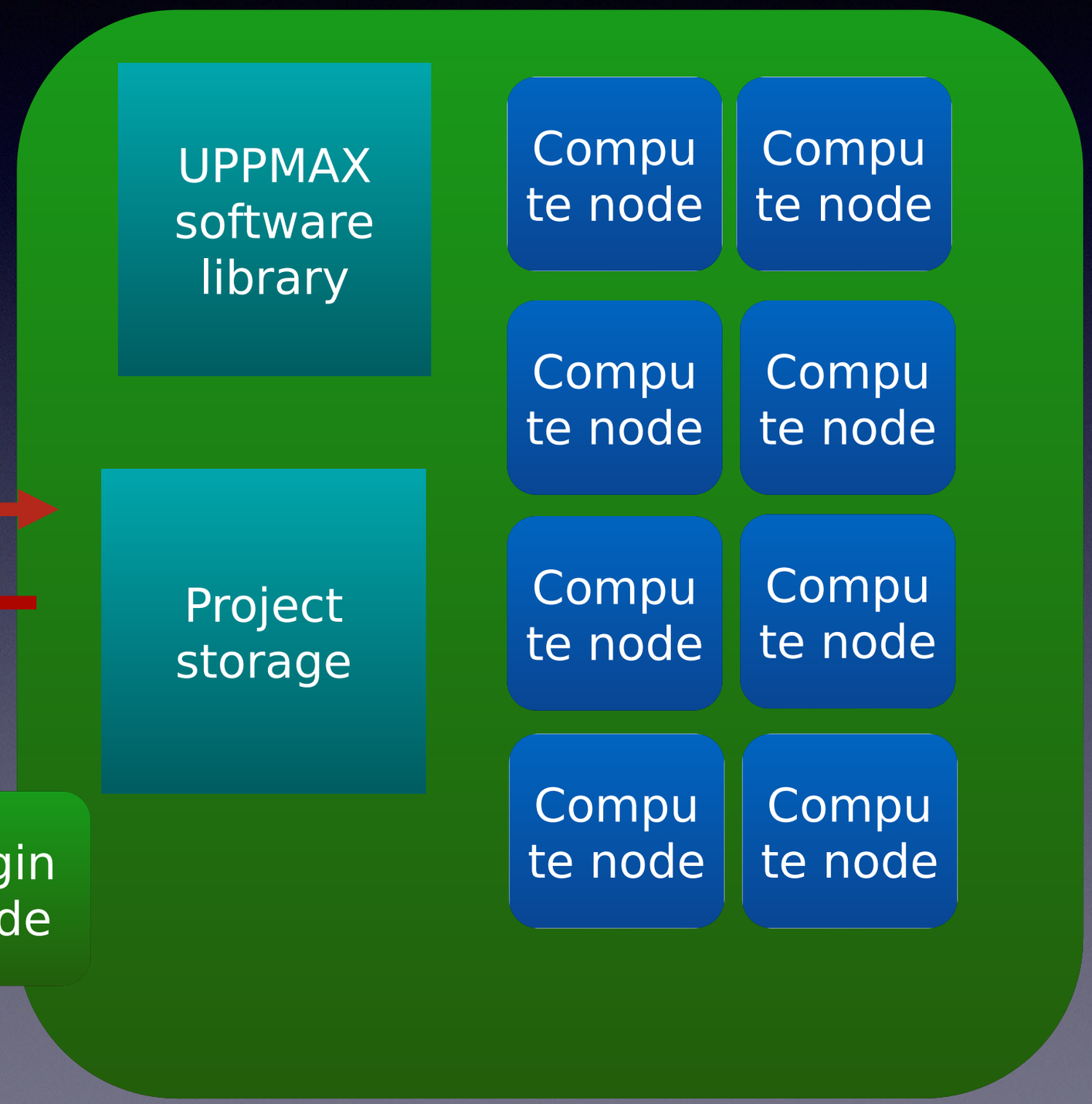
SNIC

Swedish National Infrastructure of Computing

- Funded by the Swedish Research Council and 10 participating universities
- UPPMAX is one of 6 SNIC compute centres
- SNIC a central directory for project and users (“SUPR”)
- SNIC is succeeded by a new organization in 2023

Bianca is a platform for secure computing

- Data transfers logged
- Data **transfers** encrypted
- No internet access
- Full software stack
- Slurm-cluster



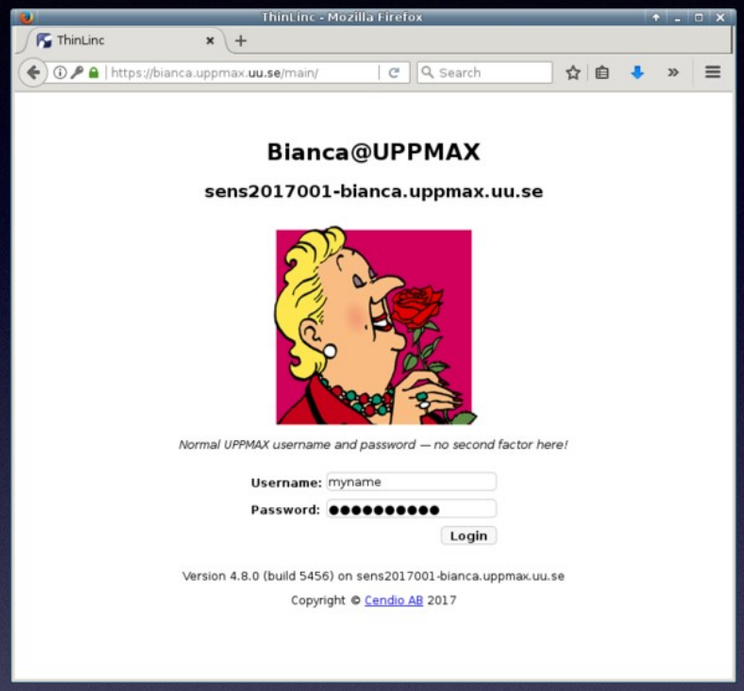
SSH login

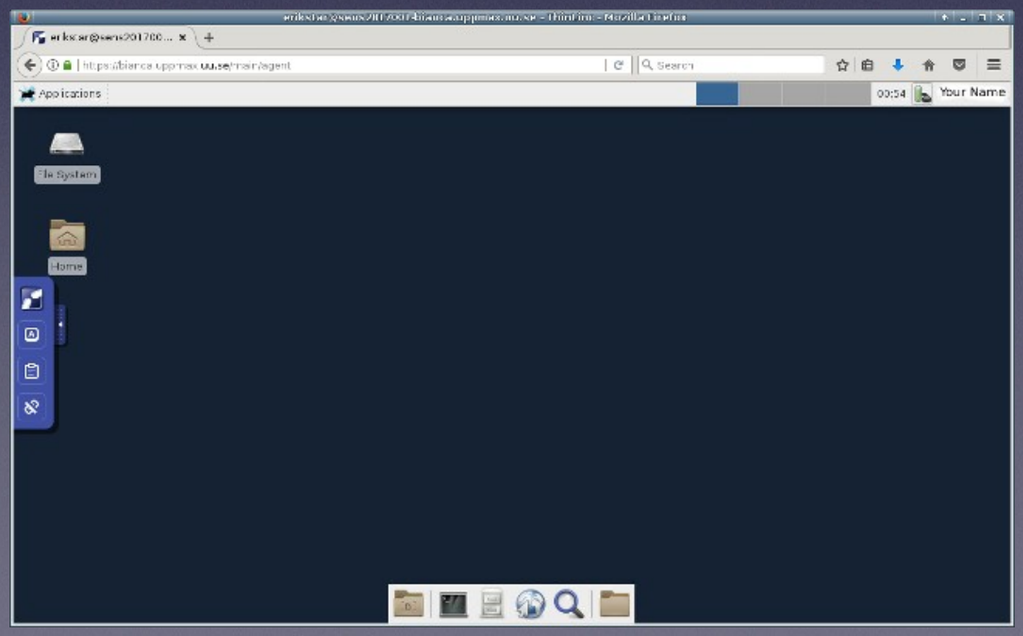
```
[root@sens2017001-bianca ~]# su -l linusn
Last login: Tue Oct  4 08:58:05 CEST 2022 on pts/0
 _   _ ____  ____  __  __      _     __  __
| | | |  _ \|  _ \|  \/  |    / \    \ \/ /   | System:    sens2017001-bianca
| | | | |_) | |_) | |\/| |   / _ \    \  /    | User:      linusn
| |_| |  __/|  __/| |  | |  / ___ \   /  \    |
 \___/|_|   |_|   |_|  |_| /_/   \_\/_/\_\   |

################################################################################

        User Guides: http://www.uppmax.uu.se/support/user-guides
        FAQ: http://www.uppmax.uu.se/support/faq

        Write to support@uppmax.uu.se, if you have questions or comments.


[linusn@sens2017001-bianca ~]$
```

Suitable for submitting and managing Slurm jobs

Graphical log-in supported

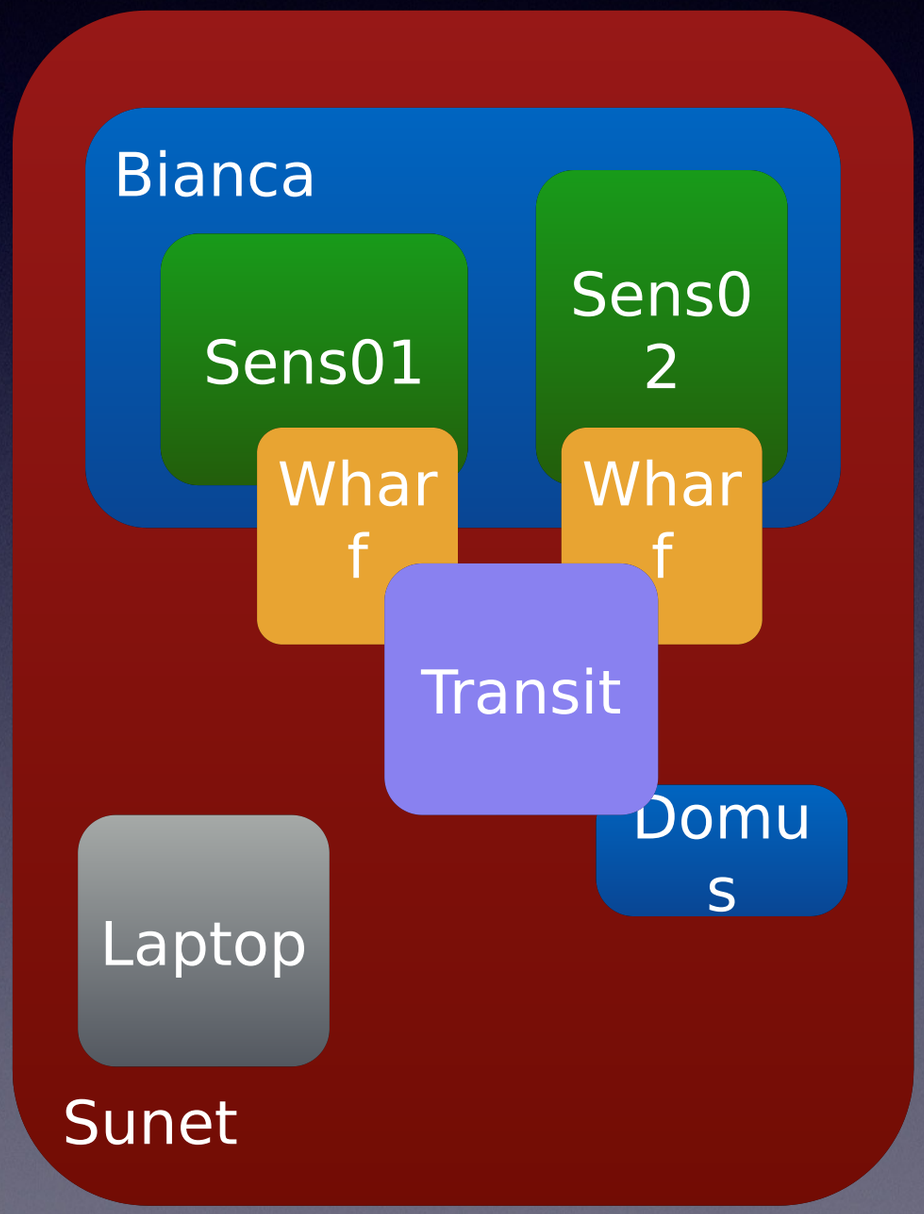
- Remote Desktop via ThinLinc in a browser
- Great for e.g. RStudio, looking at images and results directly





Transferring data

- sftp is the only supported transfer protocol (Filezilla, lftp, etc.)
- A transit server exist that offers more flexibility
  - Useful for:
    - Copying data between SENS projects
    - Downloading data from internet to a SENS project



By the numbers

 200 thin (128GB) nodes 80 fat (256GB+) nodes 4 GPU (2xA100) nodes	 296 projects 212 PI’s from 9 institutions ~700 users
 ~1000 software modules and reference datasets	 ~4.8 PB of data 6.5 PB of storage (soon 11PB)

Bianca is a part of SNIC SENS

- Purpose: To set up and maintain an e-Infrastructure for:
  - Handling sensitive personal data at the National Genomics Infrastructure, NGI
  - High-performance and data-intensive analysis of sensitive personal data
  - Intended user base: life-science users with some HPC experience (i.e. shared Slurm cluster)

Collaborative effort

- Project partners are SNIC, **UPPMAX**, PDC, NGI, and Science for Life Laboratory.

Vetenskapsrådet

UPPSALA UNIVERSITET

Knut och Alice Wallenbergs Stiftelse

SciLifeLab

UPPMAX SNIC-SENS systems

- Producer systems (used by National Genomics Infrastr.)
  - **Miarka (succ. Irma)**: production cluster
  - **Vulpes (succ. Lupus)**: storage for Miarka
  - **Grus**: data delivery service
- Consumer systems
  - **Bianca**: secure compute cluster
  - **Castor**: storage for Bianca
  - **Cygnus**: new storage for Bianca

Allocations

- Projects created from proposals sent and reviewed by SNIC staff at SUPR (supr.snic.se)
- PI must be employed or affiliated with Swedish academia
- Project members and collaborators
  - Create account at SUPR & accept user agreement
  - In practice must connect from SUNET
- Agreements for sensitive data

Costs?

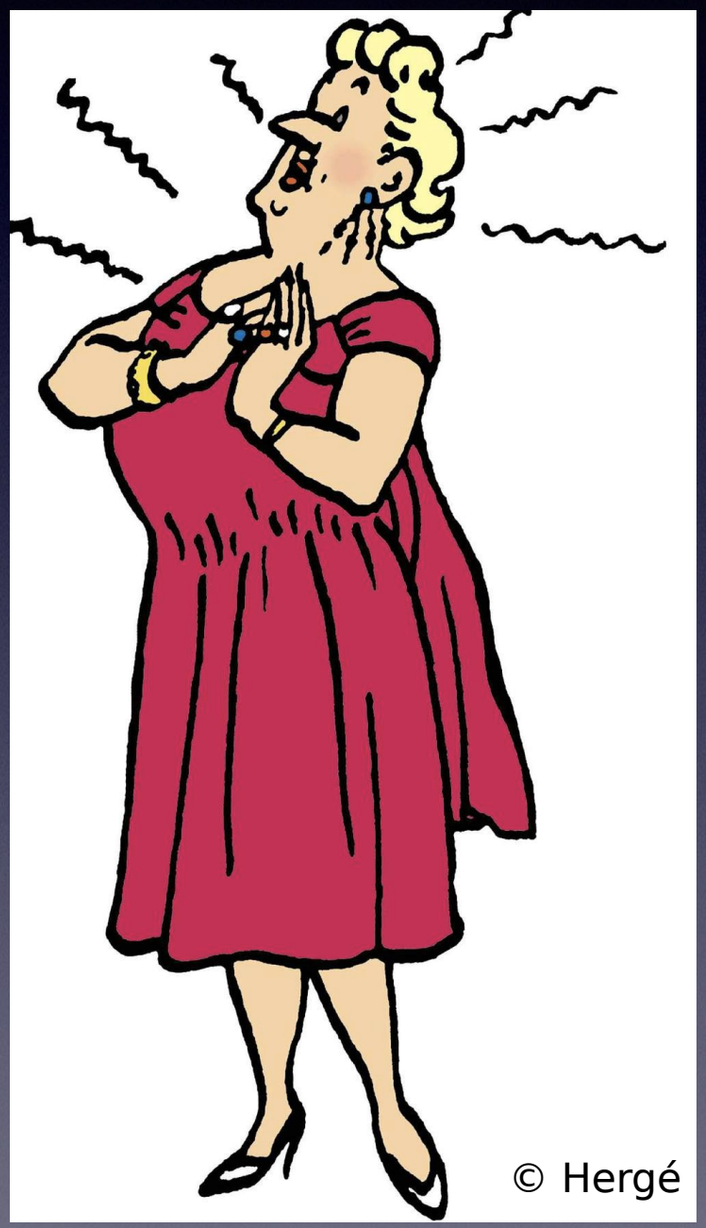
- Swedish researchers & their collaborators ... it is free
- Infrastructure & industry ... needs to pay
- International users... it is possible

SNIC SUPR

Access for foreign users

- Have the Swedish data controller apply for a SNIC-SENS project from supr.snic.se
- GDPR: Sign a personal data processing agreement (UPPMAX can provide help)
- Access to Bianca restricted to SUNET, use a VPN or (if SSH is enough) jump host rackham.uppmax.uu.se

Data agreements



Data Processing Agreements (DPA) on individual PI/project level

Current constraints

- Projects managed via SUPR
- 2FA (TOTP) SUPR-login required for PI
- Bianca login and data transfers requires 2FA authentication
- Direct access to Bianca services only from SUNET
- Currently supported external services are SSH and SFTP

Bianca in 2023-

- SNIC is replaced in 2023
  - Large changes to organization is expected
  - Change in mandate and responsibilities of SNIC-SENS possible outcome
  - At current time of writing, much is unknown

- UPPMAX continues to develop Bianca by hiring new staff in 2022/2023 to increase interoperability between e-services and offering new services

Thank you for listening