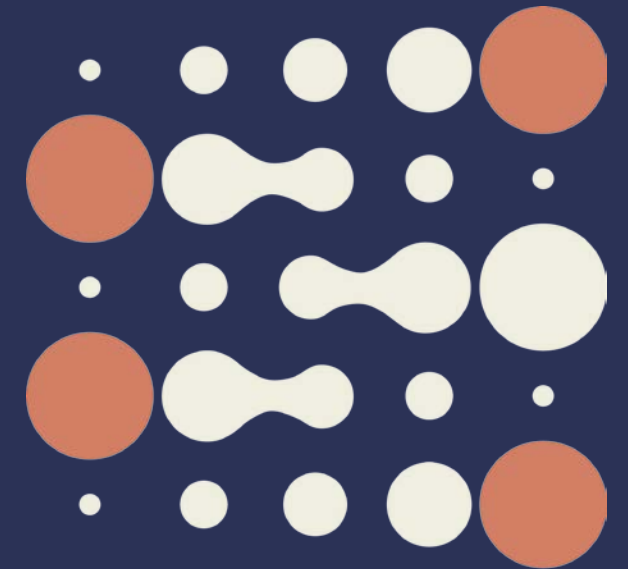


IARC Scientific Council SC/61

Item 9: Request for support from the governing council special fund: computing infrastructure for the IARC Scientific IT platform

Nicolas Tardy, infrastructure manager (ITS)

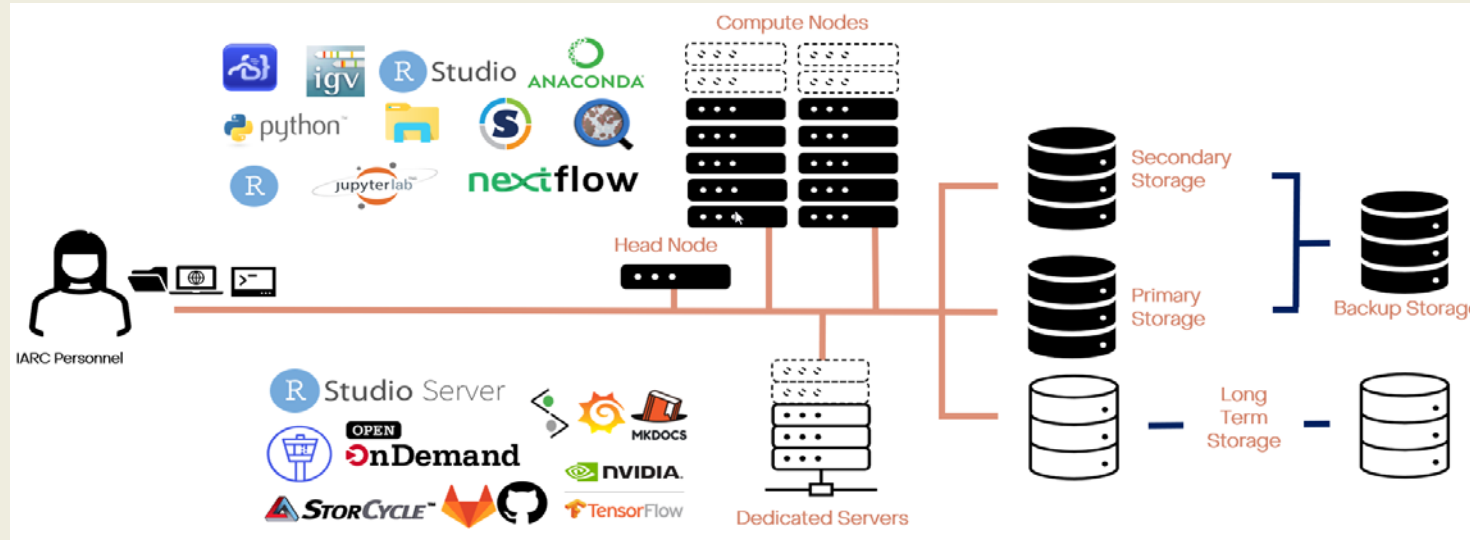
International Agency
for Research on Cancer



The IARC Scientific IT Platform



Scientific IT



Purpose:

- Centralized infrastructure, cost-effective, for storing and analyzing scientific data.
- Ensures secure, high-performance computing resources aligned with worldwide data protection standards.
- A key resource for collaborative, data-centric cancer research and open-science initiatives.

Strategic Importance:

- Enables efficient handling of complex datasets for cancer research.
- Promotes interdisciplinary collaboration and open-science initiatives.
- Supports the IARC Medium-Term Strategy (MTS).



OnDemand provides an integrated, single access point for all of your HPC resources.

Pinned Apps A featured subset of all available apps

System Installed App

- JupyterLab
- RStudio
- QIPath
- IGV
- Fiji
- NextFlow Tower
- Visual Studio Code
- Mega

powered by OPEN OnDemand

```

# R code for UMAP analysis
# ... (code omitted for brevity) ...
ggplot(MAP_UM, aes(x = UMAP1, y = UMAP2, color = Molecular_clusters)) +
  geom_point(size=1, alpha=0.8) +
  scale_colour_manual(values=distinctive_ols) +
  labs(x = "UMAP dimension 1", y = "UMAP dimension 2", color = "") + theme_bw() + theme(legend.position = "bottom")
  
```

```

Host: osiris.iarc.iarc.fr
Last login: Mon Sep  6 16:41:56 2021 from 10.99.1.26

Welcome tardyn on OSIRIS Cluster, Iarc's HPC Infrastructure

Checking Users Environment

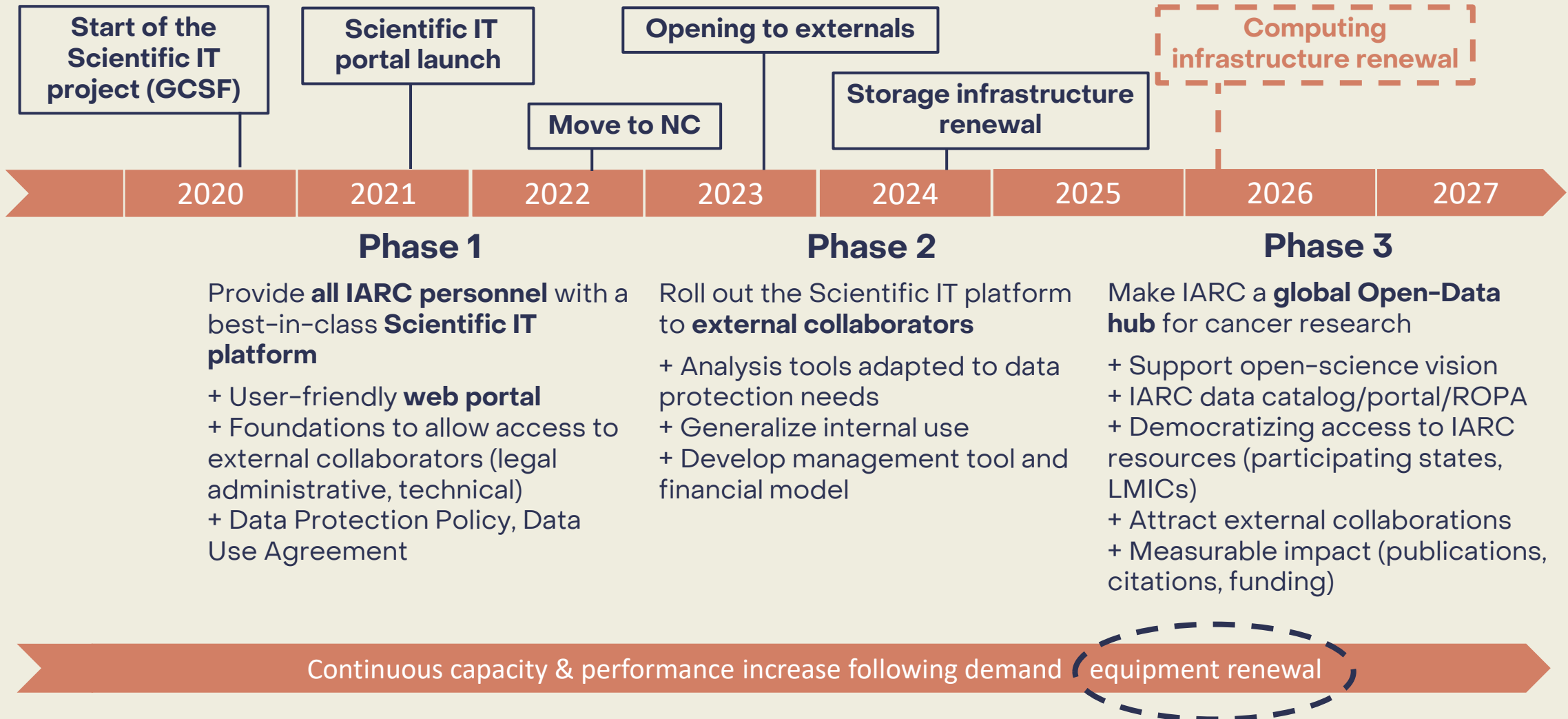
CHECK OUT Osiris documentation on the Scientific Computing Wiki
http://collab.iarc.fr/communities/ScientificComputing/

TIPS: About SCHEDULER: SLURM 20.02.3
visit http://slurm.schedmd.com
Specify the resources you need when submitting a job

TIPS: About DATA
/home : store your personal working file. Space limited. Incremental Backup
/data : store project structure. Check Guideline for more details about backup
/data/references : common references centrally managed
/common : common databases centrally managed
/file : store project file. No backup

CONDA
It is appropriate to install package
conda in /home/tardyn/miniconda3
with conda might be in /app or ask it-services@iarc.fr
  
```

History and future plans





Current Source of Concerns

End of Life for Current Servers

- Servers are eight years old and reaching the end of their lifecycle.
- No longer covered by warranty, increasing the risk of unavailability due to hardware failures.

Capacity to Meet Growing Needs

- Rising computational demands from research activities.
- Current infrastructure struggles to scale with growth.

Upcoming Challenge

- Budget and funding approvals
- Ensure scalability to meet future research demands
- Procure and integrate new High-Performance Computing while minimizing disruption during the transition phase

Funding request and proposal

Request: €250 000 from the Governing Council Special Fund.

Proposed Investment:

- New computing servers totaling 2000 compute cores, 15 TB RAM, and 25 GBps network.
- Including some high-performance graphics processing units (GPU) for artificial intelligence research.
- Supporting infrastructure: network, power supply, and maintenance (five years).

Sustainability Plan:

- Ongoing development of a cost-recovery system for long-term operational support.

Why it matters:

- Ensures continuity and scalability of a critical research platform.
- Aligns with IARC's goals for open science, collaboration, and high-quality research.