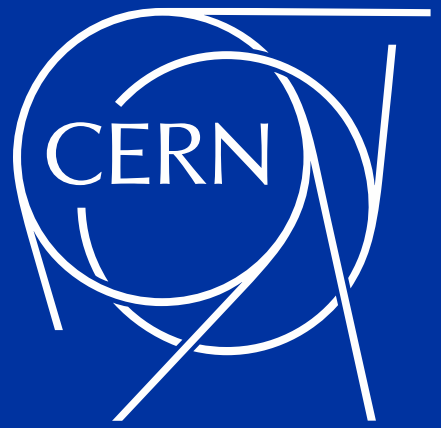


Technology Infrastructures Function

João Fernandes
IT

IT-FTI Group Meeting – 10.12.2024

Outline

Context

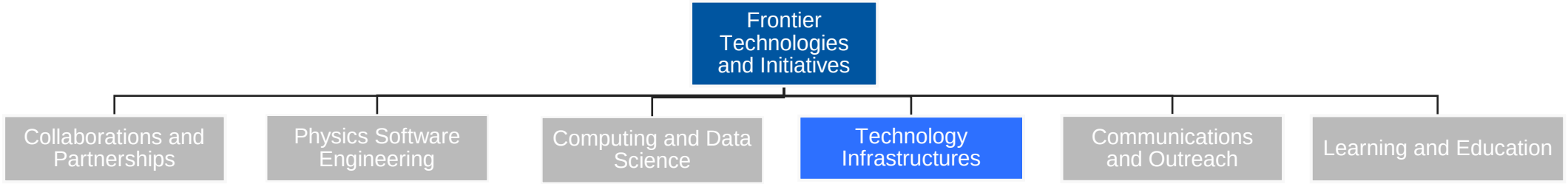
Strategy for External Computing Resources

Cloud Resources Operating Model

Q&A

Technology Infrastructures: Functional overview

The Technology Infrastructures function provides general coordination of access to resources and capabilities from external providers to FTI activities. It also develops expertise in efficient and sustainable management and use of heterogenous computing services such as public clouds, quantum services and HPC, keeping close integration with the IT Architecture function and a direct relationship with IT Capacity Planning and the Technical Coordination teams.



KEY ROLES & RESPONSIBILITIES

- Manage agreements between the FTI Group and external resource providers to support the technical needs of IT Dept. flagship projects
- Develop data analytics tooling to monitor the technology resources usage including costing, manage risks, and contribute to the Department planning and sustainability efforts
- Exploit strategic partnerships in heterogeneous computing services to co-design use cases for future technologies and capabilities
- Establish a direct link to IT Capacity Planning and IT Architecture functions and contribute to the transition of the experience of the FTI projects into Technical Delivery

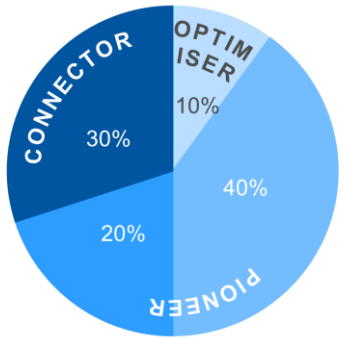
MEASURES OF SUCCESS

- Successful provisioning of resources to FTI projects and initiatives
- Improved understanding of cost management of new architectures in the IT services lifecycle
- Operational access model for quantum services in the CERN Quantum Hub
- Consolidated synergies between the activities of FTI , ARB, Capacity Planning and TC
- Better understanding of areas for future strategic collaborations with external providers

KEY CAPABILITIES

Strategic Planning	Resource Usage
Sustainability	Co-Development

ROLES FRAMEWORK



GOVERNANCE STRUCTURES

Weekly Report meetings
Project meetings with the QTI and NGT
ARB meetings
Capacity Planning meetings
Project meetings with industrial partners

PAIN POINTS ADDRESSED

- Standardisation of metrics for seamless access across diverse computing resources
- Provision of advanced analytics capabilities with increased cost-effectiveness
- Tangible impact of innovation in service operations

IT Cloud Strategy

The IT Department established the need to create a Public Cloud Strategy

Building on the IT department's documented strategy and the recently implemented target operating model, as well as the findings and commitments mentioned above, the following activities will be undertaken:

- An initial set of use-cases will be identified by our project initiation pathways (Business engagement, Innovation and Technical enhancement) that can benefit from deployment on public clouds
- The Architecture and Data Function will propose a framework for the use of public clouds respecting CERN policies and considering potential exit strategies
- Technical Delivery will identify personnel across the different layers of the stack to support the deployment on public clouds of the selected use-cases
- The Resource Management group will develop a means of estimating service costs that can be progressively refined
- The Resource Management group, in collaboration with IPT, will develop a procurement process for public cloud services

These activities will collectively produce a public cloud strategy for the IT department that will be enacted from the beginning of Q4 of 2022.

Published on the 6th May 2022: https://cds.cern.ch/record/2809609

Objectives

- **Complement on-premise capacity with public cloud services (IaaS/PaaS/SaaS)**
- **Prepare the Department to meet the increasing computing needs of CERN**
  - Support future classical computing needs (e.g. NextGen)
  - Equip the QTI2 with access to multiple quantum technologies to ensure its R&D objectives
- **Result in 2024: Final Cloud "Package" with 3 elements:**

  1) Updated Cloud Policy; ☑

  "The usage of external cloud services is overseen by the "Steering Group on Externally Provided Software, Systems, Hardware and Services (STEPS)."

  - Finalised; To be endorsed CERN wide via STEPS

  2) 2 Separate blanket contracts (Cloud and Quantum) with a total of 3+2 vendors

  - Results submitted to the FC 25th of September 2024 and unanimously approved; ☑

  3) Multi-staged and Multi-Governance cloud operating model (2025)

Operating Model Implemented in 3 phases

- **Initiation Phase (now)**
  - Initial deployment of the Resource Operating Model and its Governance
  - Top-level elements of the Governance are all in place, it is possible to make the transition to the Pilot Phase.
    - Estimated timeline: Ready by January 2025

Pilot Phase 2025

- Deployment of 3 initial Pilot Use Cases to build expertise, validate the model and associated processes
- **Minimum Viable Product (MVP):**
  - QTI2 to decide access via sub-licensing to Quantum vendors
    - Access to be only provided to QTI approved R&D projects
  - NextGen: NGT decides access to vendors, timings and amount of resources
    - Technical effort allocated by NGT to operate and develop their own projects
      - **Effort allocation already in the Kubernetes Team and IT-CA**
  - CERN IT Pilot: Heterogenous Architectures Testbed (HAT)
    - Subscription managed by the Capacity Planning function
    - Minimal testing access for checking a particular technology for suitability purposes
    - Limited in budget, resources and access time

Governance - Pilot Phase 2025

Finance & Resource Management

Procurement & Vendor Management

Cloud Enablement Team

- IT Chief Architect
- FinOps Manager
- IT Consulting Lead
- Security
- IT Data Privacy Coordinator

Business Engagement

Training & Upskilling

Heterogeneous Architecture Testbed

Other use cases

Cloud Task Force

FinOps, TD technical experts, Capacity Planning, Resource Management, QTI and NextGen reps./technical experts, Engagement reps.

QUANTUM TECHNOLOGY INITIATIVE 2.0

NexTGen Next Generation Triggers

From 2026: Scaling Phase

- An updated governance model for cloud and associated processes will be a deliverable from the Pilot phase by Dec 2025
- The scaling phase will onboard all the lessons learned and evolution of use cases
- A refined, generalised model to access external heterogeneous resources will be produced
- The Scaling Phase aims to start in January 2026

FTI Tech Infra Function in a Nutshell

- Provision of computing resources not available on-premise
  - Starting by NextGen and QTI
- Provide advanced analytics on resource usage, cost and optimisation
- Generalise a model for heterogenous off-prem resources provisioning
- Establish a direct relationship between FTI, IT Architecture and IT Capacity Planning functions

One More Thing…

Frontier Technology Sailing

- Someone called it once ”Quantum Sailing”, but I think now it can be renamed ”Frontier Tech Sailing” ☺
- Normally happens twice per year: April and October
- We will fill another boat in our next outing in April 2025: the Netherlands
  - If interested, please let me know. ☺

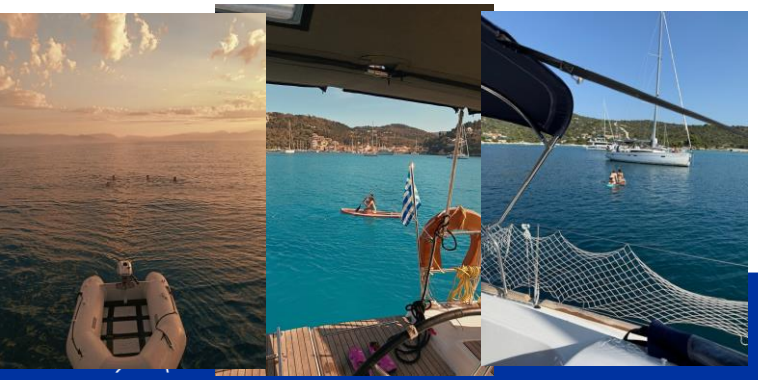
- Below, some pictures from the last outings…

Azores, July 2024

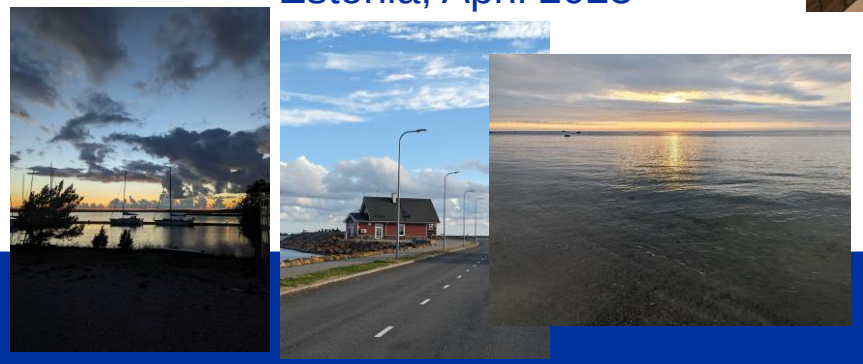
Naples, April 2022

Sicily, October 2023

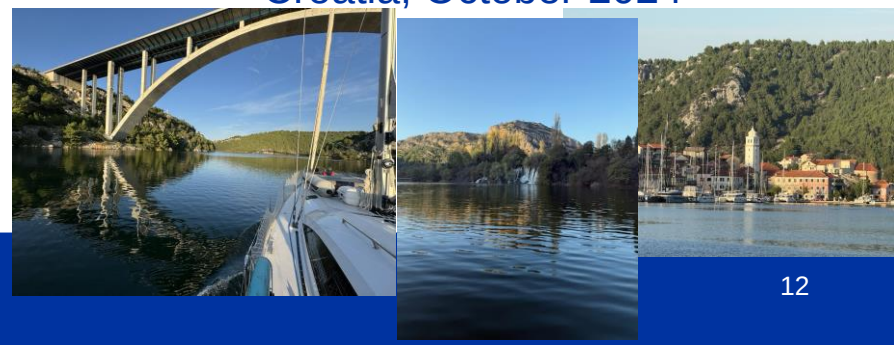
Corfu, October 2022



Estonia, April 2023



Croatia, October 2024



Sailing Netherlands in April

Frontier Technology Sailing Roadmap

- Finland
- Lofoten Islands
- Turkey
- Greece (again) this time Sporades
- Denmark
- Azores again (cross to Flores)
- …

Questions?