FENIX - Federated engine for information exchange

Federated data & computing infrastructure

Giuseppe Fiameni (CINECA) et al

g.Fiameni@cineca.it

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

cea

CINECA

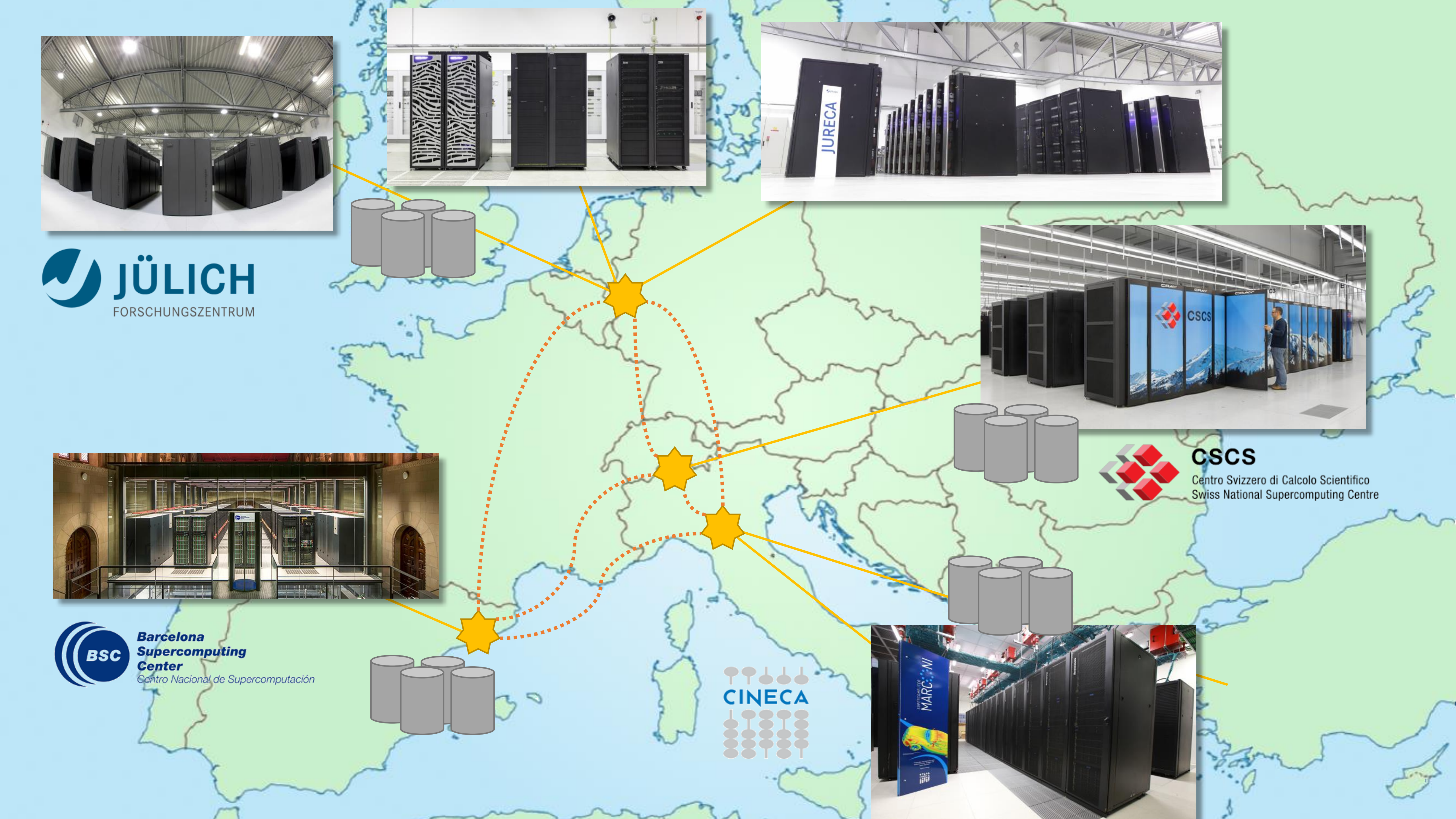
CSCS
Centro Svizzero di Calcolo Scientifico
Swiss National Supercomputing Centre

JÜLICH
FORSCHUNGSZENTRUM

Co-funded by the European Union

The Human Brain Project

- **Research Communities: The Human Brain Project Goals of the Human Brain Project (HBP)**
  - Enable research aiming for understanding of the human brain
  - Transfer neuroscience knowledge for development of future technologies
- **FET Flagship project funded by EC**
  - Future & Emerging Technologies projects (co-)funded by European Commission
  - Science-driven, seeded from FET, extending beyond ICT
  - Ambitious, unifying goal, large-scale
- **Current HBP status**
  - 114 participants in Specific Grant Agreement 1 (SGA1)
  - SGA1 runs from 2016-18 with an overall budget of about € 110M


JÜLICH
FORSCHUNGSZENTRUM
JURECA
CSCS
CSCS
Centro Svizzero di Calcolo Scientifico
Swiss National Supercomputing Centre
BSC
Barcelona
Supercomputing
Center
Centro Nacional de Supercomputación
CINECA

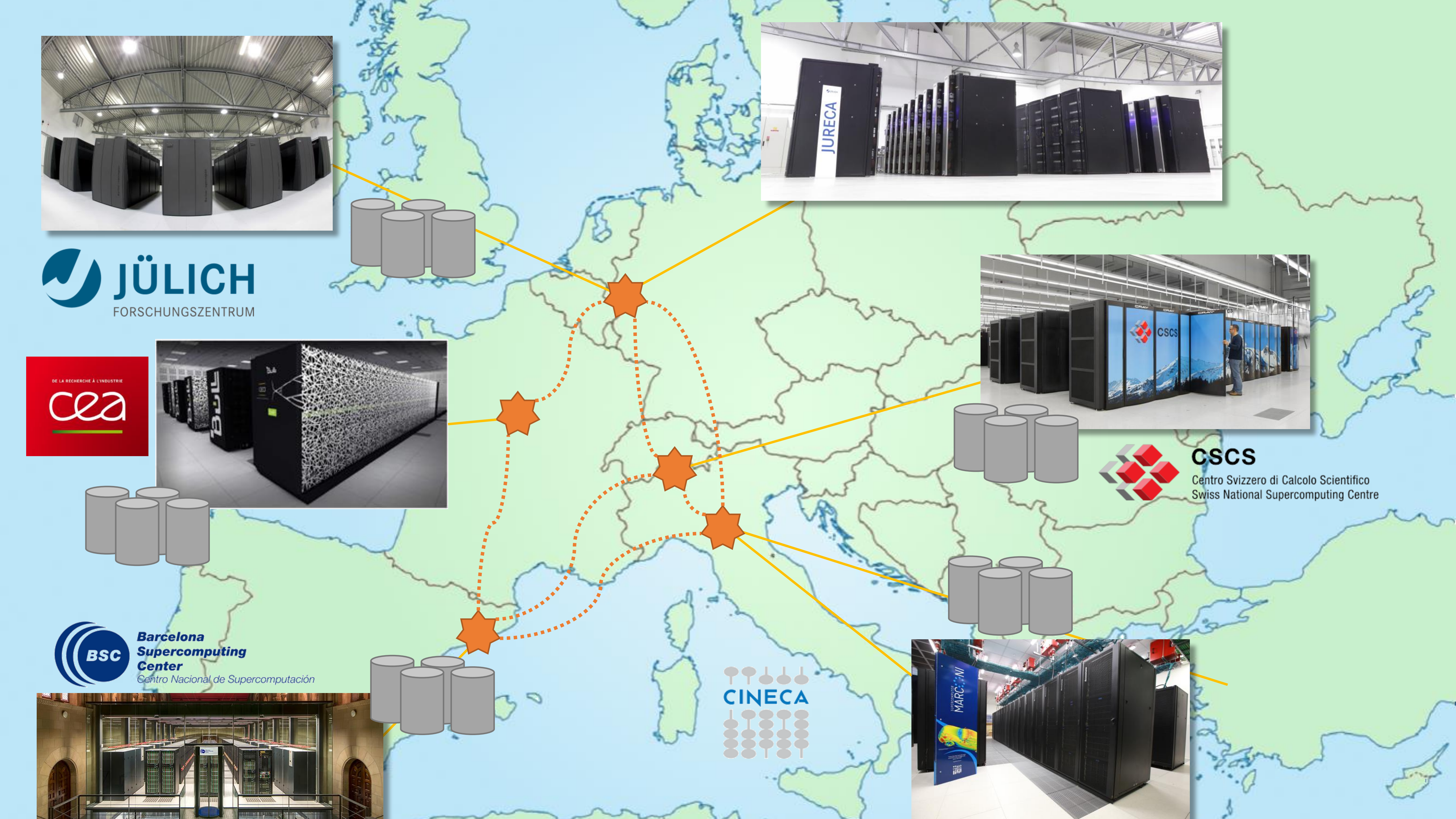
High Performance Analytics & Computing Platform

As part of the HBP, we build and operate a supercomputing, data and visualization infrastructure that enables scientists to

- Run large-scale, data intensive, interactive brain simulations up to the size of a full human brain
- Manage the large amounts of data used and produced in the Human Brain Project
- Manage complex workflows comprising concurrent simulation, data analysis and visualization workloads

The role of FENIX

- **Deliver a multi-purpose infrastructure offering scalable compute and data services in a federated manner**
- **Support new communities**
  - Neuroscience (remains a main driver to steer the design of the infrastructure)
  - Materials science
  - Genomics
  - Physical science experiments
  - Others communities with similar requirements
- **Supported by national funds and EC through the ICEI Project (Interactive Computing E-Infrastructure)**

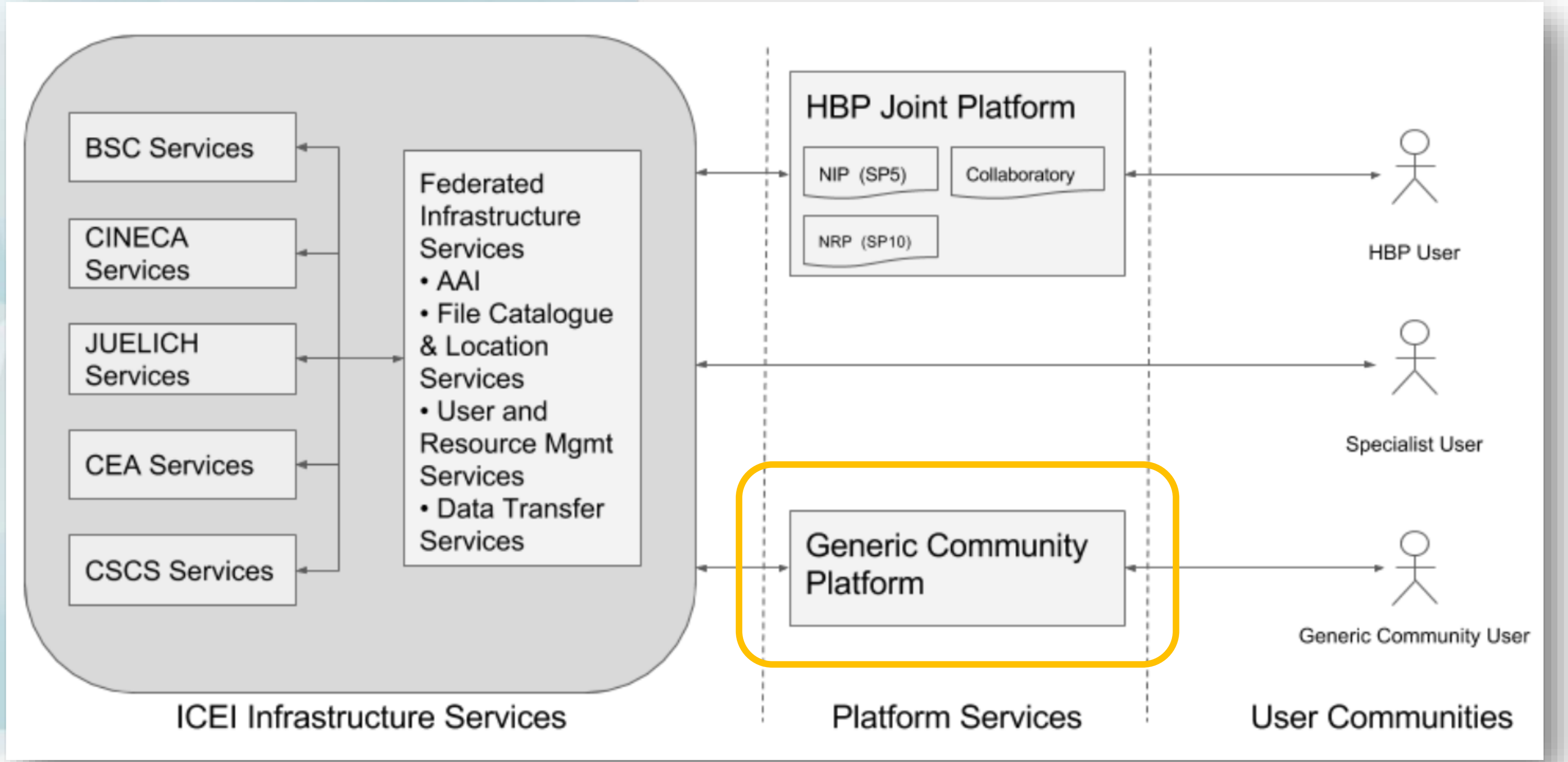

JURECA
JÜLICH
FORSCHUNGSZENTRUM
DE LA RECHERCHE À L'INDUSTRIE
cea
CSCS
Centro Svizzero di Calcolo Scientifico
Swiss National Supercomputing Centre
BSC
Barcelona
Supercomputing
Center
Centro Nacional de Supercomputación
CINECA

Rationale behind FENIX

- **Variety of data sources**
  - Distributed data sources
  - Heterogeneous characteristics
- **HPC systems as source and sink of data**
  - Scalable model simulations creating data
  - Data processing using advanced data analytics methods
- **Aim for data curation, comparative data analysis and for building-up knowledge graphs**

Need for infrastructure to facilitate data sharing and high-performance data processing.

Overview of the Fenix Infrastructure

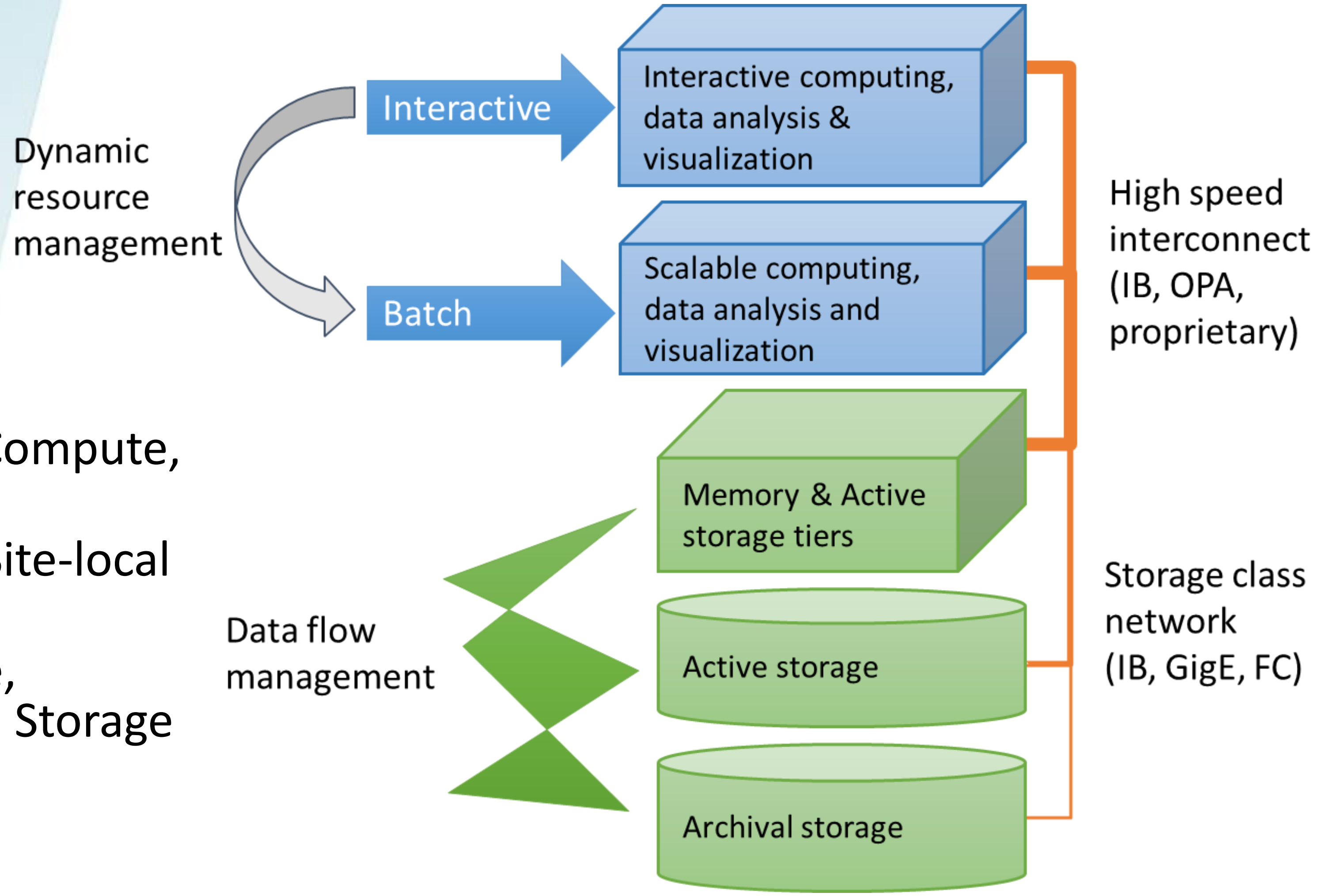


FENIX Services

Specific *service targets*:

- Interactive Computing Services
- Scalable Computing Services
- Federated Data Services

- **Additionally**
  - IaaS environments (SW-defined Compute, Storage and Network)
  - Container Services, DB services, Site-local AAI
  - Scalable and Interactive Compute, Visualisation, Dense memory and Storage tiers
  - Active- and Archival-class Storage



Key challenges

- **Common AAI infrastructure**
  - Federated user identities
  - Single sign-on
- **Federation of storage resources**
  - Scalable vs. federated access
- **Integration of interactive computing resources**
  - New type of resource
- **Management of resource allocation**
  - Different resource classes
  - Delegation of resource allocation to research communities

Key architectural concepts

Interactive Computing Services

- **Interactivity**
  - capability of a system to support distributed computing workloads while permitting
    - Monitoring of applications
    - On-the-fly interruption by the user
- **Architectural requirements**
  - Interactive access
  - Tight integration with scalable compute resources
  - Fast access to data. Improve data movement across multiple storage layers (NVRAM, NVMe, Apache Pass, 3DXPoint, SSD, Disks, Tapes, etc.)
- **Support for interactive user frameworks**
  - Jupyter notebook
  - R
  - Matlab/Octave

Data Store Types

- **Archival Data Repository**
  - Data store optimized for capacity, reliability and availability
  - Used for storing large data products permanently that cannot be easily regenerated
- **Active Data Repository**
  - Data repository localized close to computational or visualization resources
  - Used for storing temporary slave replica of large data objects
- **Upload buffers**
  - Used for keeping temporary copy of large, not easy to reproduce data products, before these are moved to an Archival Data Repository

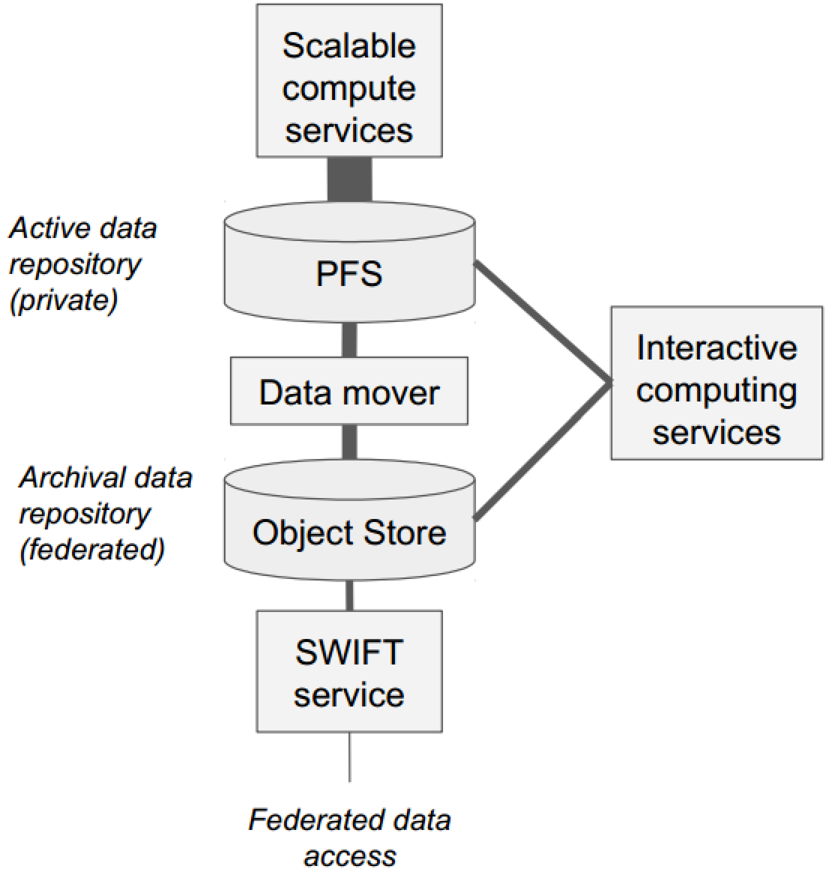
Architectural Concepts: HPC vs. Cloud

- **State-of-the-art: HPC**
  - Highly-scalable parallel file systems
    - Scale to O(10) clients
    - Optimised for parallel read/write streams
  - Interface(s): POSIX
    - Well established interface
    - Wealth of middleware relying on this interface
- **State-of-the-art: Cloud**
  - Solutions for widely distributed storage resources
    - Optimised for flexibility
  - Various interfaces: Amazon S3, OpenStack Swift
    - Typically web-based stateless interfaces
  - Advantages compared to POSIX
    - Suitable for distributed environments (e.g. support for federated IDs)
    - Simple clients
    - Rich mechanisms for access control

Storage Architecture

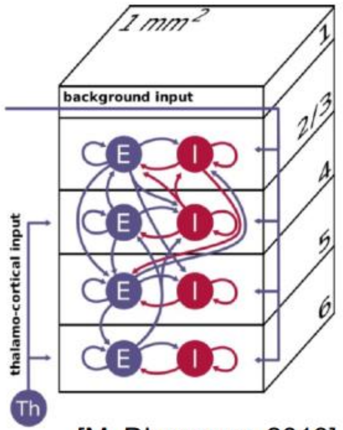
- **Concept**
  - Federate archival data repositories with Cloud interfaces
  - Non-federated active data repositories with POSIX interface accessible from HPC nodes

- **Envisaged implementation: Mandate same technology at all sites**
  - Current candidate: OpenStack SWIFT

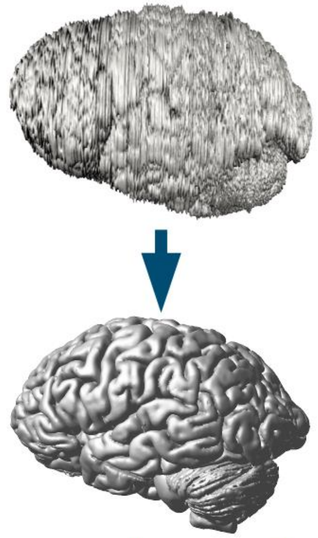


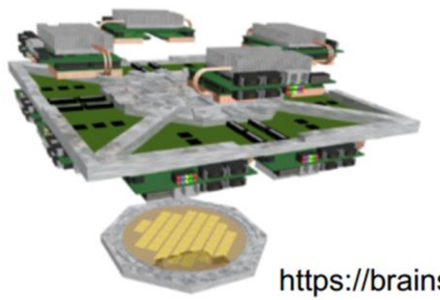
Selected Use Cases

- **GUI based interaction with extreme scale network models**
  - Various simulators supporting different models
  - Need for interactive visualisation of network generation and simulation
- **Enrichment of the human brain atlas with qualitative and quantitative datasets**
  - Spatial and semantic registration of diverse datasets to the human brain atlas
- **Validation of neuromorphic results**
  - Analysis of the similarities and differences of results obtained through simulation on HPC and from neuromorphic systems



[M. Diesmann, 2013]

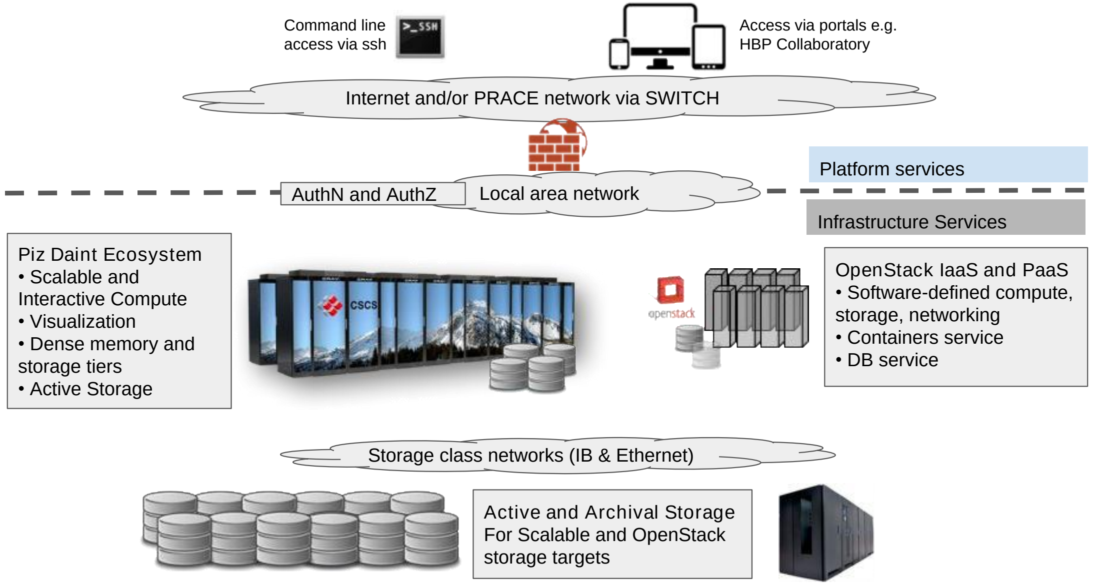

[K. Amunts et al.]


https://brainscales.kip.uni-heidelberg.de/

Scalable Computing Services

Scalable computing services are a key element of the Fenix Infrastructure

- *Piz Daint* at CSCS will form a major part of these services
  - A hybrid multi-core system with 7135 nodes
  - >27 PFlop/s aggregate peak
- The Piz Daint environment offers
  - Scalable and Interactive Computing
  - Visualization
  - Dense memory and storage tiers
  - High-throughput Active Storage
  - **All within one system**



Thank you!

Credits

- **BSC**
  - Javier Bartolome, Sergi Girona and others
- **CEA**
  - Hervé Lozach, Jacques-Charles Lafoucriere, Jean-Philippe Nomine, Gilles Wiber and others
- **CINECA**
  - Carlo Cavazzoni, Giuseppe Fiameni, Roberto Mucci, Debora Testi and others
- **CSCS**
  - Colin McMurtrie, Sadaf Alam, Thomas Schulthess and others
- **Jülich Supercomputing Centre**
  - Anna Lührs, Björn Hagemeier, Boris Orth, Thomas Lippert and others