Mitglied der Helmholtz-Gemeinschaft

JÜLICH
FORSCHUNGSZENTRUM

Towards the Fenix Infrastructure

D. Pleiter | LSDMA Symposium, Karlsruhe | 29 August 2017

JÜLICH FORSCHUNGSZENTRUM

Fenix Goals

Develop and deploy services that facilitate federation

- Based on European and national resources

Establish infrastructure following the IaaS paradigm for multiple research communities

- Encourage communities to build community specific platforms
- Delegate resource allocation to communities

Architectural features

- Scalable compute resources
- A federated data infrastructure
- Interactive Computing Services providing access to
  - the federated data infrastructure and
  - the scalable compute resources

Disclaimer

The Fenix infra-structure is still in a design and develop-ment phase. Several aspects presented in this talk are to be considered tentative

Fenix Consortium

Currently involved centres

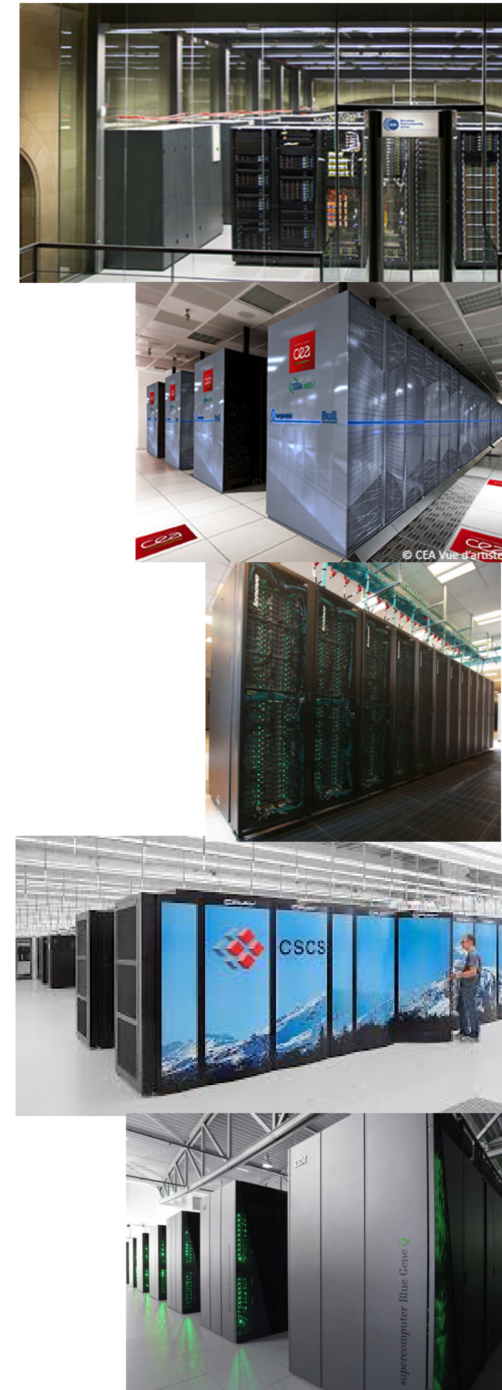
- BSC (ES)
- CEA (FR)
- CINECA (IT)
- CSCS (CH)
- JSC (DE)

Consortium features

- European HPC centres that provide resources within PRACE-2.0
- Strong links to key science drivers

Foreseen extensibility

- Open for more partners and stakeholders



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

Other Research Communities

Materials science

- Data sets from simulations but also experiments
- European community already engaged in enabling data sharing

Genomics

- Explosion of data volumes
- Some groups start to exploit HPC infrastructures

Physical science experiments

- Data from large-scale experiments, e.g. ERIC
- Need for scalable simulations for interpreting experimental results or to process data

Common Features and Requirements

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

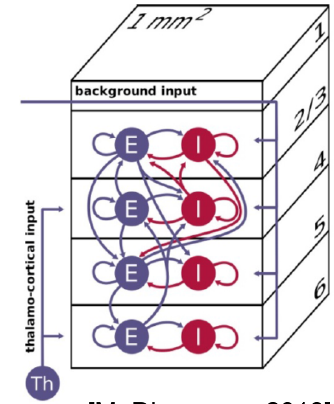
Aim for data curation, comparative data analysis and for building-up knowledge graphs

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

Selected HBP Use Cases

GUI based interaction with extreme scale network models

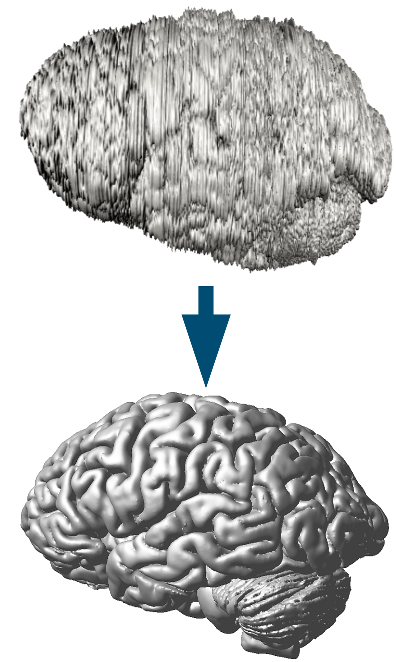
- Various simulators supporting different models
- Need for interactive visualisation of network generation and simulation



[M. Diesmann, 2013]

Enrichment of the human brain atlas with qualitative and quantitative datasets

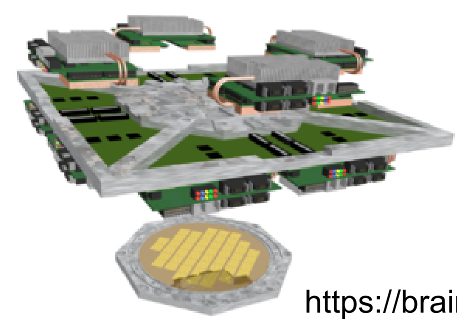
- Spatial and semantic registration of diverse datasets to the human brain atlas



[K. Amunts et al.]

Validation of neuromorphic results

- Analysis of the similarities and differences of results obtained through simulation on HPC and from neuromorphic systems

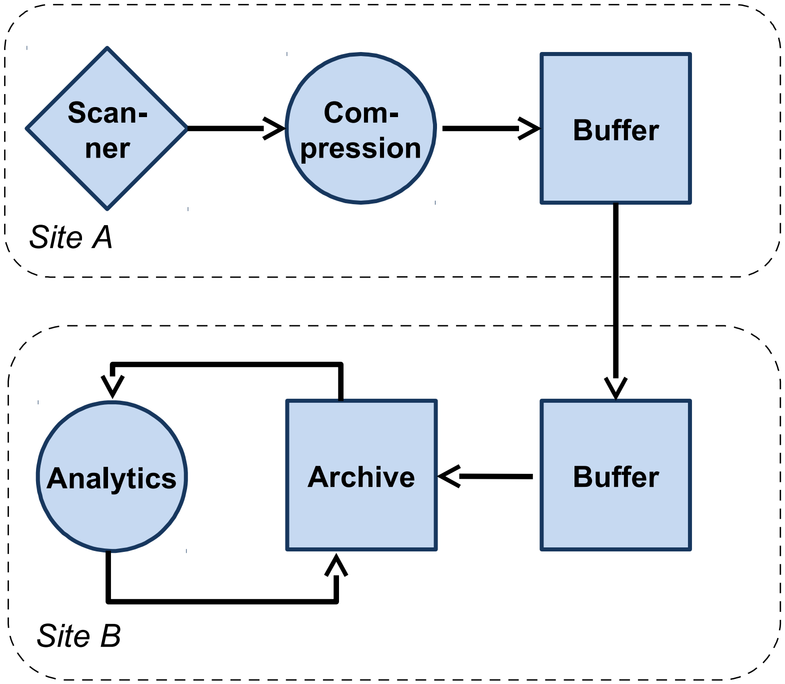


https://brainscales.kip.uni-heidelberg.de/

Use Case Analysis

Analysis of workflow based on abstract infrastructure model

- Data ingest
- Data repository
- Processing station
- Data transport



Use case/workload specific annotation of components

- Data transport
  - Maximum/average required bandwidth
  - Interface requirements
- Data repository
  - Maximum capacity requirements
  - Access control requirements
- Processing station
  - Data processing hardware architecture requirements
  - Required software stacks

Design Principles

- Bring data in **close proximity to the data processing resources**
  - Take advantage of high bandwidth active data repositories and large capacity archive data repositories
- Improve **data resilience**, **data availability** and **data access performance** through federation (and data replication if necessary)
- Data centres involved are independent and provide this **infrastructure as a service** to different research communities
- Make Fenix being perceived by user as **single infrastructure**
  - Single sign-on
  - Data location services
  - Unified transfer mechanisms

Design Principles (cont.)

- Security, flexibility, **extensibility**, scalability, use of open standards
- Take into account requirements of different sub-communities with respect to the **interfaces** used for data access and for the execution of their **workflows**
- The design is **driven by specific use cases** to ensure that the infrastructure is productively usable for the target science communities soon after deployment

Key Challenges

Common AAI infrastructure

- Federated user identities
- Single sign-on

Federation of storage resources

- Scalable vs. federated access

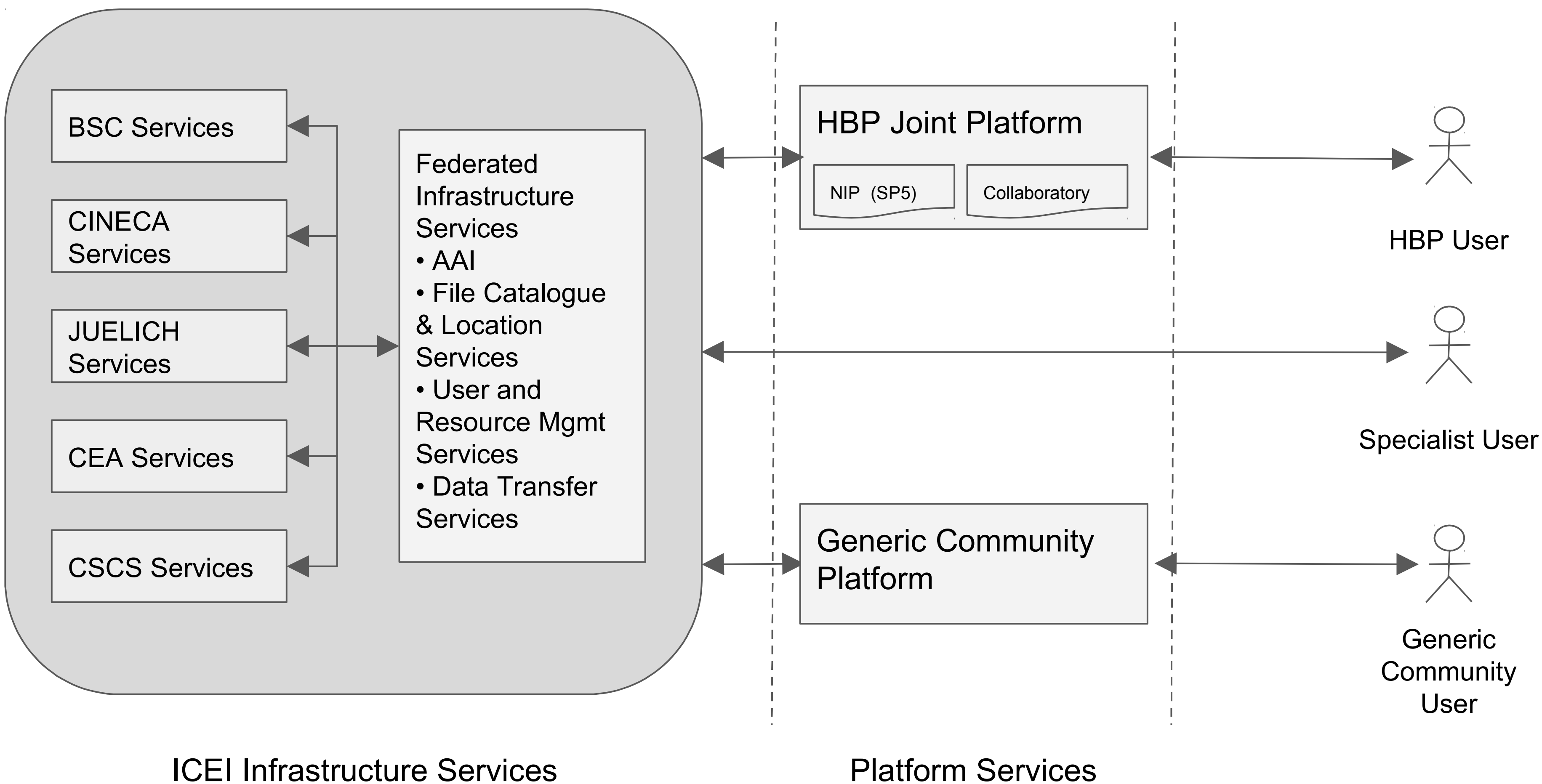
Integration of interactive computing resources

- New type of resource

Management of resource allocation

- Different resource classes
- Delegation of resource allocation to research communities

Architecture Overview



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

Support for interactive user frameworks

- Jupyter notebook
- R
- Matlab/Octave

Scale-out Computing Resources

Highly scalable applications among use casest

- Brain simulators
- DFT-based applications

Focus on increased memory footprint

- Integration of dense memory technologies
  - Initial work through R&D services procured within a Pre-Commercial Procurement of the HBP

Future vision: Elastic access to scalable compute resources

- Allow users of Interactive Computing Services to launch scalable applications

Architectural Concepts: 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

Architectual Concepts: HPC vs. Cloud

State-of-the-art: HPC

- Highly-scalable parallel file systems
  - Scale to $O(10^5)$ 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

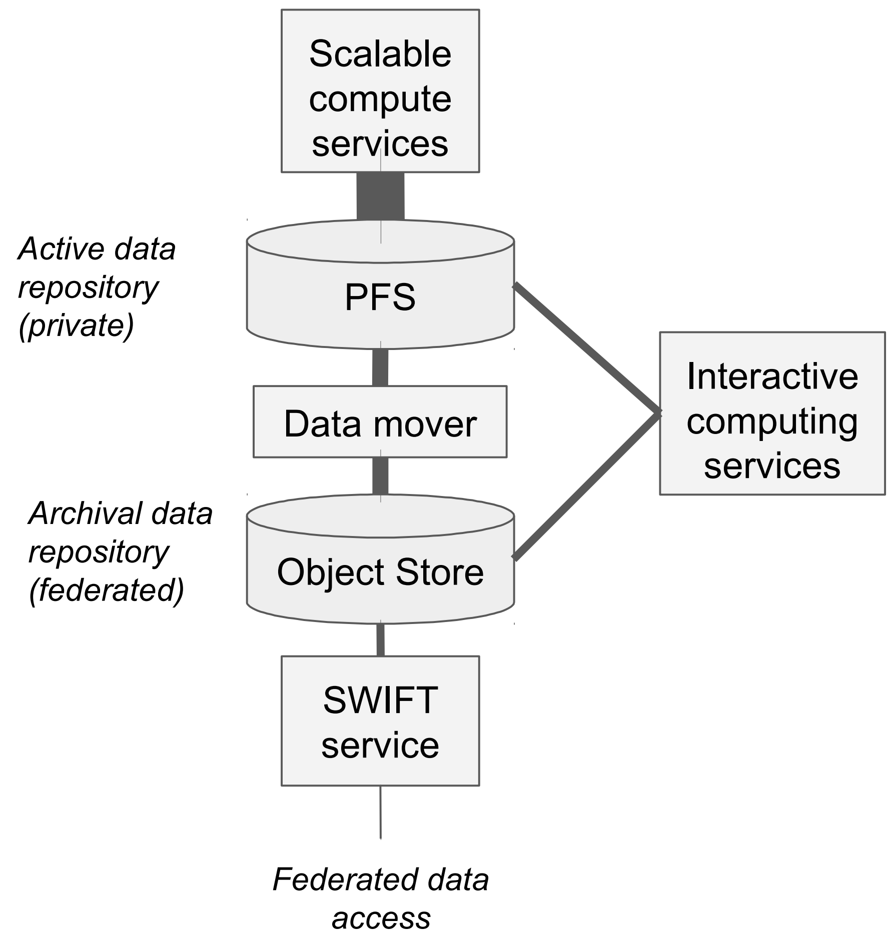
Storage Architecture

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



Fenix Credits

Fenix Credit =
Currency for authorising resource consumption

Different types of resources

- Scalable compute resources (N_{node} × time)
- Interactive computing services (N_{node} × time)
- Active data repositories (capacity × time)
- Archival data repositories (capacity)

Business model

- Resource providers provide Fenix Credits based on TCO analysis
- Fenix Communities allocated resources based on peer-review process

Summary and Outlook

Strong science drivers towards data-oriented, federated HPC infrastructures

- Examples: Brain research, materials science

Many opportunities and challenges

- Federation of services including AuthN
- POSIX vs. Cloud storage technologies
- Integration of interactive computing services
- New models for allocating HPC and data resources to research communities

Fenix

- Group of (currently) 5 European supercomputing centres committing to federate relevant services

Credits

BSC

- Javier Bartolome, Sergi Girona and others

CEA

- Hervé Lozach, Jacques-Charles Lafoucriere, Jean-Philippe Nomine, Gilles Wiber and others

CINECA

- Carlo Cavazzoni, Giovanni Erbacci, Giuseppe Fiameni, Roberto Mucci and others

CSCS

- Colin McMurtrie, Sadaf Alam, Thomas Schulthess and others

Jülich Supercomputing Centre

- Anna Lührs, Björn Hagemeier, Boris Orth, Thomas Lippert and others