Architectures and Tools for Large-Scale Workflows
Exa-AToW

F. Bodin

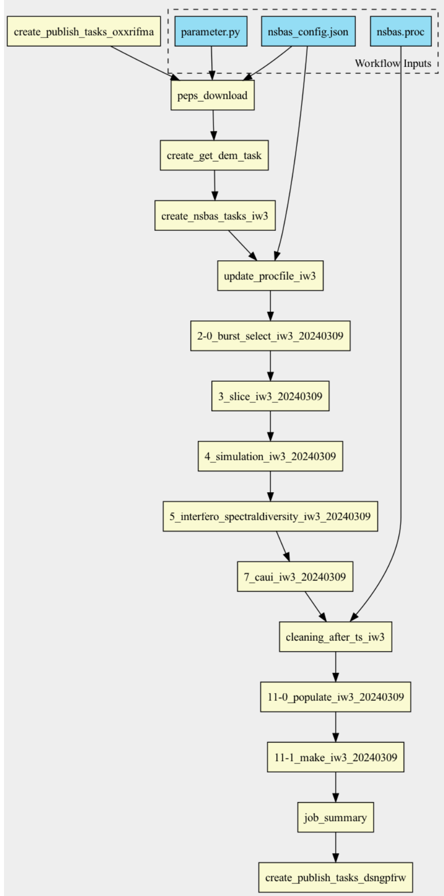
https://numpex.org/

Partners & Roles

- University of Rennes
  - PI: François Bodin / Mark Asch
  - Role: Architecture
- Inria DataMove
  - PI: Olivier Richard
  - Role: Orchestration
- CEA DRF
  - PI: Thierry Deutch
  - Role: Workflow specification
- CNRS IDRIS
  - PI: Guillaume Harry
  - Role: Cybersecurity
- CNRS PYTHEAS
  - PI: Didier Mallarino
  - Role: Sustainability models

Co-coordination:

- François Bodin
- Thierry Deutsch
- Mark Asch



Technology integration & interoperability project

Introduction

- The “Architectures and Tools for Large-Scale Workflows” project considers that
  - The Exascale machine is part of a “datasphere”
  - We have a continuum of distributed resources (*transcontinuum [ETP4HPC]*) which blur the frontier of storage, compute, and network resources
- We aims at developing support for large-scale workflows
  - Mix of data and compute tasks
  - Data-driven Service Oriented Architecture (FAIR)
  - Integration of technologies with the corresponding governance
- With a global optimization approach, considering performance, energy and sustainability

Exa-AToW proposes a metadata centric approach

The Datasphere

- Data sources are multiple and heterogeneous
  - Large scientific instruments
  - Sensors (IoT)
  - Simulations

- From a data perspective, **the computing machine can be seen as a data cache**
  - Dynamic management of data storage

- The data transfer time (and the associated energy) is a strong constraint
  - Transferring 1 Petabyte at the speed of 10Gb/s -> 9 days at the maximum speed
  - In practice it is 45 days with a lot of support

Issues of Accessibility of Computing, Storage, and Datasets

- Resource allocation is done on a large scale over several geographically distributed infrastructures
  - The locality of the datasets must be taken into account
  - Distribution of the computation/HPDA/IA parts in a holistic and efficient way

- Operating conditions must be in line with practices and capabilities
  - Data input and output flows are dimensioning factors
  - Cybersecurity and intellectual property protection is a must
  - Integration with user interfaces (e.g. Jupyter notebook)

- Increasing volatility of energy costs
  - Combined with variable environmental impacts (coal/nuclear)

Sustainability ≠ Energy Efficiency

- Moving from marginal cost
  - Operations and applications
  - Pflops/watt
  - Minimum PUE, reuse of generated heat

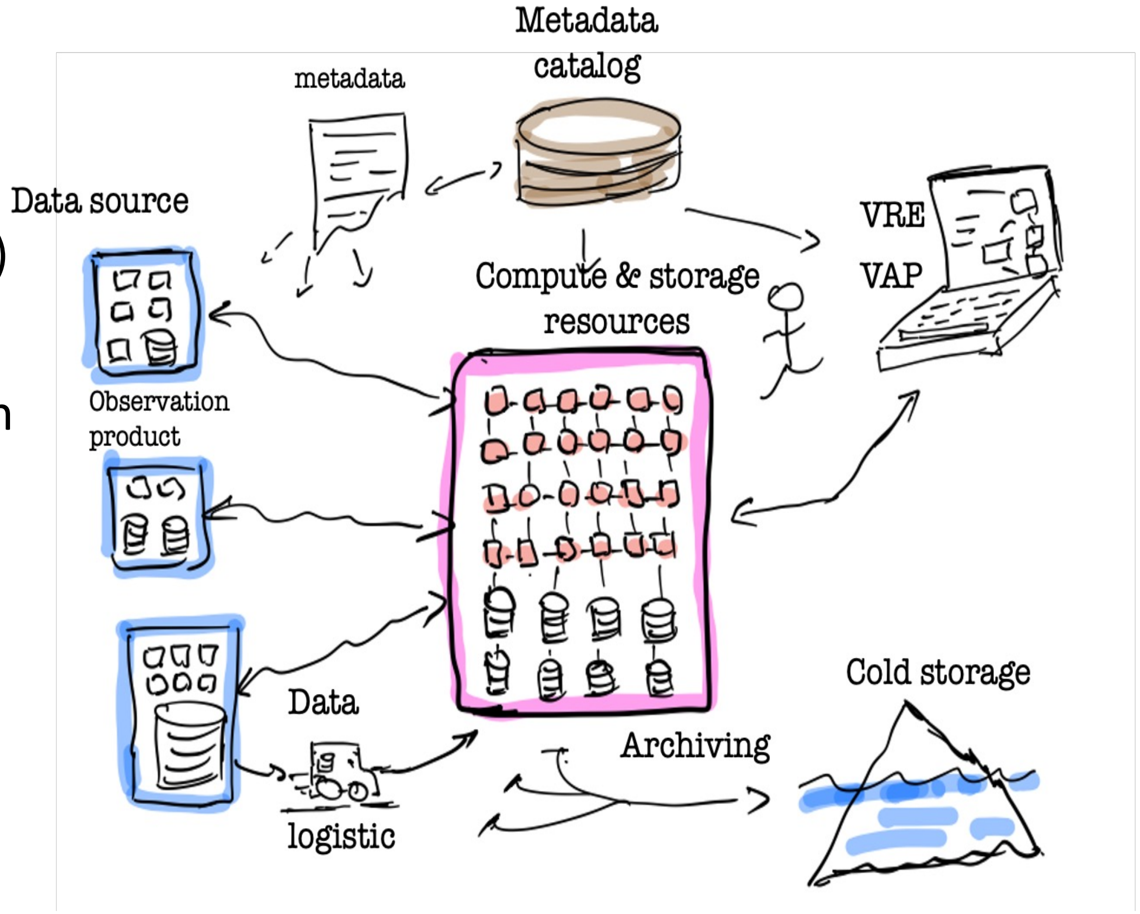
- To full cost
  - Impact of ALL life cycle stages for a comprehensive environmental cost assessment

	Total of CO_2 emission
Manufacturing	~58 %
Operations (5 years in FR)	~42 %

AMD system hosted at CEA (Src: M. Gilliot, Orap Forum 47)

User Considerations

- Use of new environments
  - VRE (Virtual Research Environment)
  - VAP (Virtual Analysis Platform)
- Application workflows composition
- Integration of multiples data sources



The Exa-AToW Challenges

- **D1** How to dynamically allocate distributed resources to accommodate for variations in the power cost

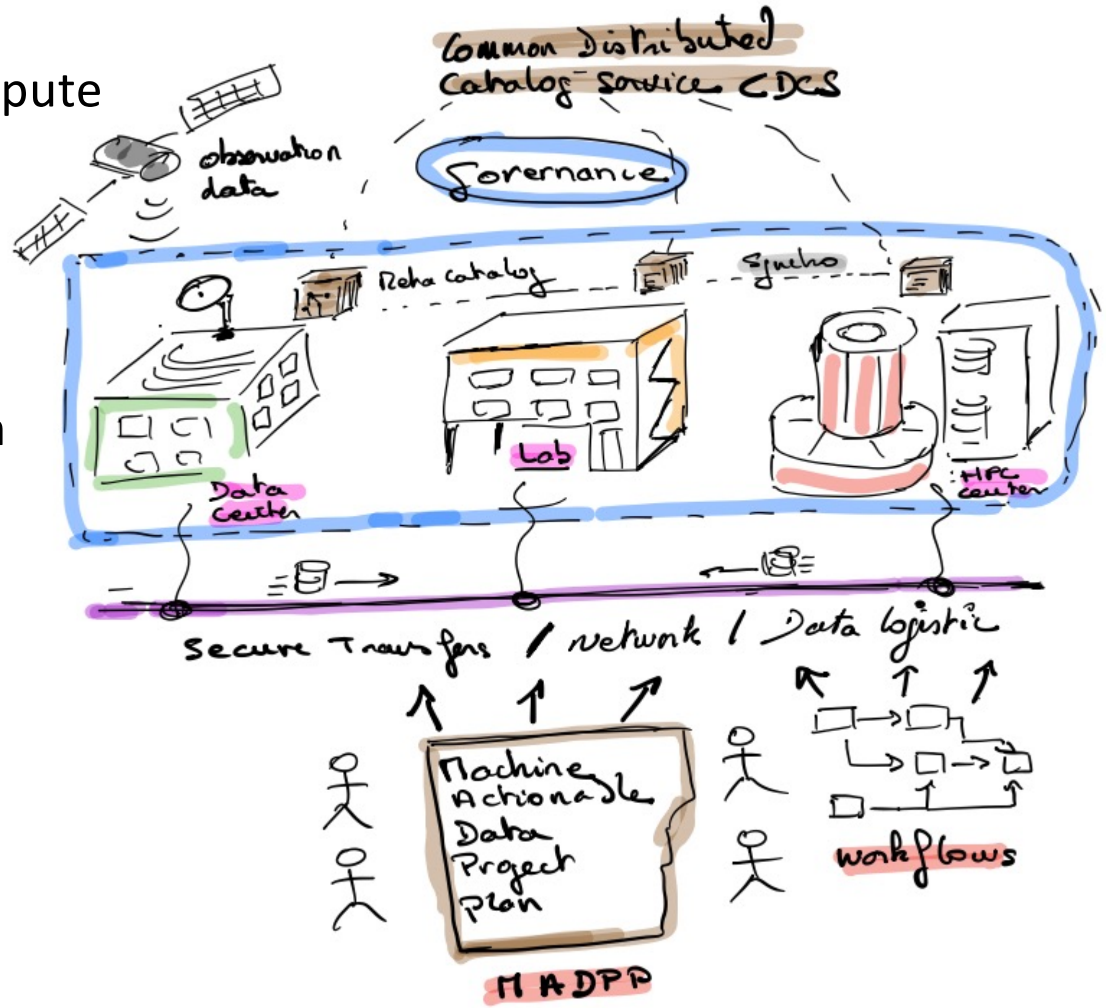
- **D2** How to organize data logistics, data life cycle, data processing, and metadata *standardization*

- **D3** How to ensure cyber-security at a large-scale on heterogeneous technologies

- **D4** How to deal with the hardware and software heterogeneity in order to run in a repeatable, replicable, and reproducible manner (and thus improving reusability)

Project Components Overview

1. Federation of network, data, and compute resources
2. Metadata Centric Approach
3. Machine Actionable Data/Project Plan (MADPP)
4. Data Logistic
5. Application & Workflow Support
6. Federation Governance



A Collaborative System-of-Systems Approach

- The underlying properties of the problems borrow many concepts of System-of-systems*

  1) Operational Independence of the Components
  2) Managerial Independence of the Components
  3) Geographical distribution of systems
  4) Evolutionary development of SoS
  5) Emergent behaviors of systems-of-systems

- Collaborative systems-of-systems that "*are distinct from directed systems in that the central management organization does not have coercive power to run the system*"*

*1998. M. W. Maier, "Architecting principles for Systems-of-Systems," Systems. Engineering, Vol. 4, No. 1, 1998, pp. 267-84.

Federation of Network, Data, and Compute Resources

- Large scale deployment requires a strong cooperation between all infrastructure stakeholders
  - Coordination for specifying cyber-security requirements in order to allow high interoperability between infrastructures
    - composability of workflows
  - Data logistic, *Data Mesh*
  - Federation level monitoring

- Considering a rich ecosystem
  - ScienceMesh, EGI/WLCG, OSG, EOSC, ELIXIR, …

Metadata Centric Approach

- Metadata in exa-AToW aim at describing properties of
    - Datasets
    - Computations
    - Operational requirements

- It allows the implementation of
    - Access to data sets and use policies
    - Allocation of compute, visualization and storage resources
    - Deploying workflows combining data analysis and numerical models
    - Dealing with provenance and monitoring
    - Allowing data preprocessing, reproducibility and open-science
    - Allow the management of sustainability metrics.

Collaborations linked to use-cases

- Data Terra / Gaia Data
  - NSBAS use-case (https://www.mdpi.com/2072-4292/13/18/3734)
    - Processing Chain Used in the FLATSIM Service
  - Machine Actionable Data Project Plan

- PEPR Diadem
  - Material simulation (https://www.pepr-diadem.fr/ project Diamond)

- SKA
  - Eclat (https://eclat.cnrs.fr/)
    - “centre d'excellence sur le calcul haute performance et l'intelligence artificielle au service de l'instrumentation pour l'astronomie”

Next Steps Focus

1. Experimental platform
   - AAI (for instance Keycloak deployment)
   - Federating resources: Eskemm nodes, TGCC Cluster nodes, Turpan nodes, Slices nodes, …
   - Data logistics simulation
2. Use cases
   - NSBAS CNES-TGCC experiment
   - Diadem
   - SKA
   - …
3. Machine Actionable Data Project Plan (MADPP)
   - Collaboration with Gaia Data / Data Terra
4. AI integration
   - Dynamic workflow management and resources allocation
   - Exa-AToW RAG (LLM) database deployment

Conclusion

- Exa-AToW requires many collaborations
  - Inside Numpex
  - Outside Numpex (data centers, data federations, etc.)

- The nature of exa-AToW is a technology integration project
  - Shares many similarities with collaborative system-of-systems

- Sustainability / energy considerations will be a strong design driver
  - Multidimensional optimization