

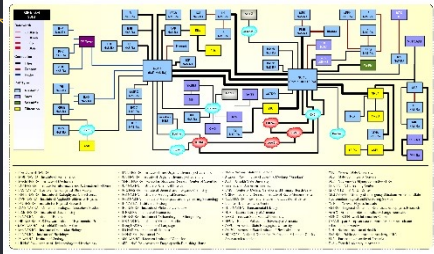
HIGH-PERFORMANCE
ARTIFICIAL INTELLIGENCE
ACTIVITIES AT ASNET-AM:
BRIDGING ADVANCED
COMPUTING AND AI

Hrachya Astsatryan

SIG AI, Poznan, Poland, 11 December 2024

Vision: National Infrastructure for Research and Technology

High-speed network



Software

- Libraries
- Packages
- Middlewares
- Queue systems
- Ticketing systems

Algorithms

- Distributed & parallel
- AI and machine learning
- Data analytics

Computing resources

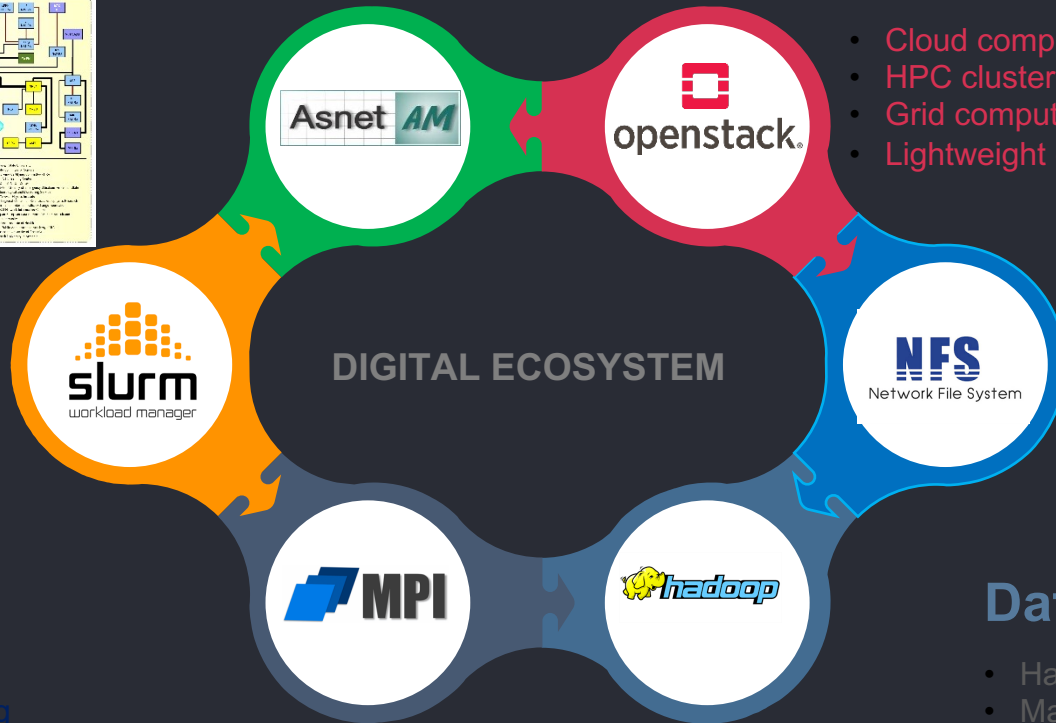
- Cloud computing
- HPC cluster computing
- Grid computing
- Lightweight computing

Storage resources

- Cold
- NAS
- Hybrid
- All-flash

Data processing

- Hadoop/Spark
- MapReduce
- Archiving



AI for HPC

AI to understand and optimize HPC systems.

AI techniques to tackle societal and scientific challenges of relevance

HPC for AI

HPC as a resource for AI to accelerate advanced AI algorithms

AI for HPC: Resource Management

Cooperative memory management system

- **Scientific objective:** to develop an accurate and non-intrusive working set estimation method and remote memory sharing system.
- **Collaboration:** University of Toulouse, France
- **Publication:** IEEE Transactions on Cloud Computing (10.1109/TCC.2020.3018089)

Energy efficiency

- **Scientific objective:** methods to improve energy efficiency in HPC systems (DVFS, power modeling, power cap)
- **Collaboration:** PSNC, Poland
- **Publication:** Scalable computing (10.12694/scpe.v18i2.1286)

Performance optimization of serverless and stateless computing systems

- **Scientific objective:** to utilize an abstraction power of container technologies combined with the scheduling and resource management capabilities of Kubernetes
- **Publication:** Cybernetics and Information Technologies (10.2478/cait-2022-0029)

Optimization of parallel programming models

- **Scientific objective:** performance modeling, benchmarking and profiling of heterogeneous infrastructures
- **Publication:** Concurrency and Computation: Practice and Experience (10.1002/cpe.4073)

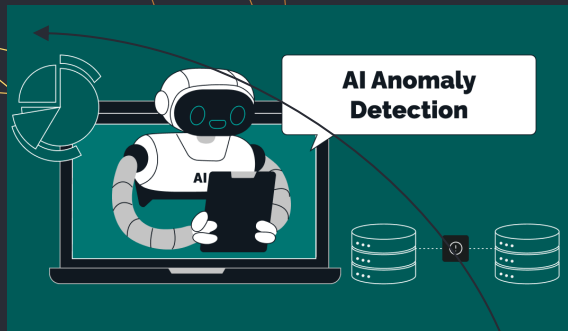
AI for HPC: Memory Pool

Remote direct
memory access

RDMA-KV:
Enhancing
performance of
key-value store in
disaggregated
memory systems

In-Memory
Processing and
Elastic Memory
Management with
RDMA for HPC:
Navigating Irregular
Workloads in HPC
and Serverless
Architectures

AI for HPC: Networking



Leveraging ML in SDN and NFV for advanced traffic management, anomaly detection, and resource optimization

Data

Logs Data

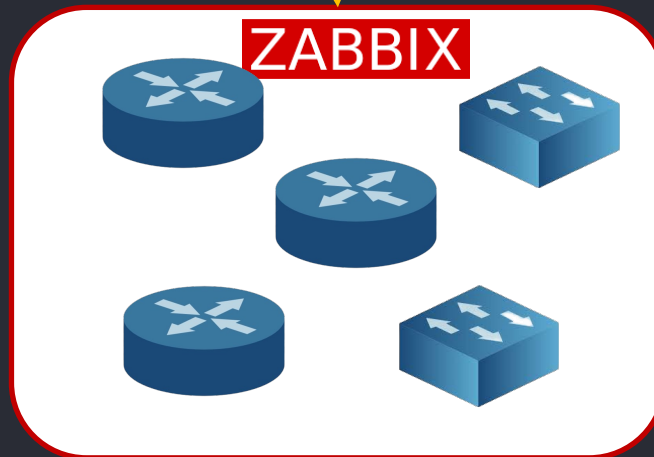
Logs

Logs
Data

Logs

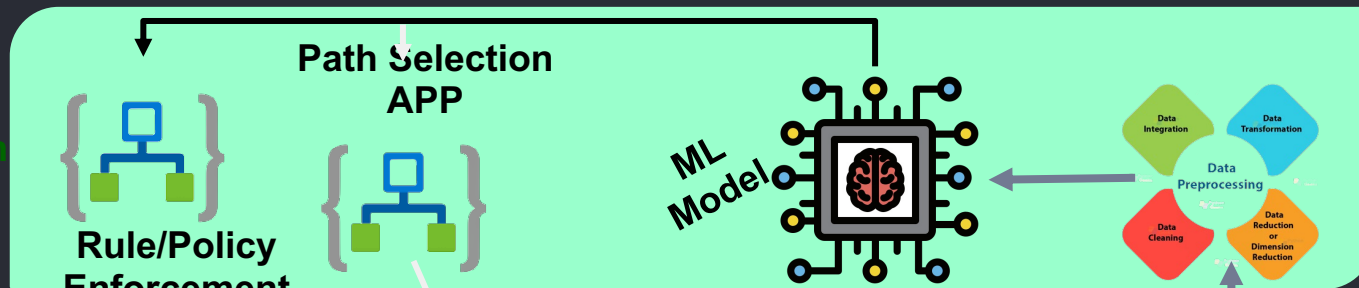
Data

Data

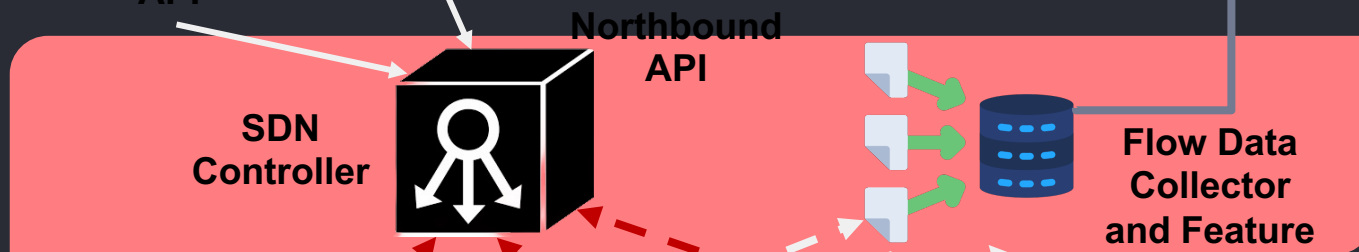


AI for HPC: Networking

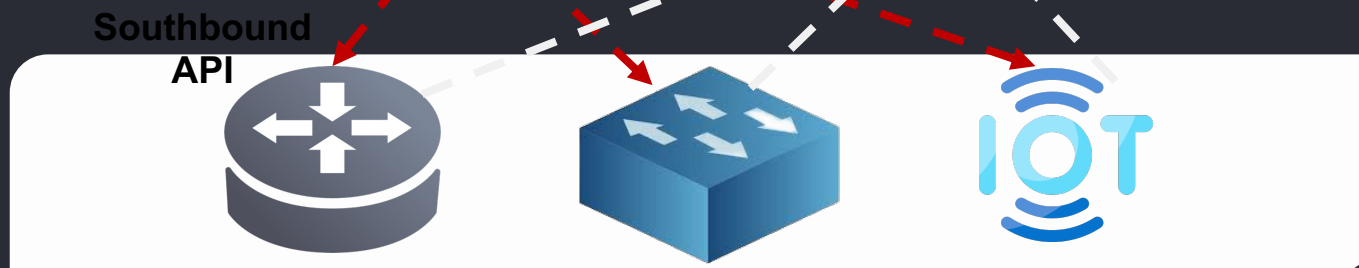
Application
Plane



Control
Plane



Data Plane



HPC for AI: Scientific Applications

Physics

Quantum physics

ML methods and scalable algorithms to retrieve the values of the physical parameters of experimentally studied atomic systems

Pub: Springer Contemporary Physics
(10.3103/S1068337221040137)

Astrophysics

To process and analyze the exponentially increasing amount of astronomical data using ML

Pub: Data Science
(10.5334/dsj-2024-006)

Earth Sciences

Earth observation and remote sensing

Scalable Data Processing Platform for Earth Observation Data Repositories

Pub: Springer Earth Science Informatics (10.1007/s12145-022-00806-7)

Collaboration: UNIGE, Switzerland

Climate and environment modelling and forecast

Data analytics and service provisioning relying on several prediction digital models and parameterization schemas

Life Sciences

Molecular Dynamics Simulations

To optimize molecular dynamics simulations and service provisioning of molecular dynamics simulations

Pub: Cybernetics and Information Technologies
(10.1515/cait-2017-0056)

EO data processing system independent of the computing infrastructure

Flexible and scalable solutions, taking into account crucial key performance indicators

Global cloud providers

- Like Amazon, Google, or Microsoft, offer user-friendly services along with enormous computational and storage capacities

EO Data repositories

- Sentinel Hub,
- Google Earth Engine,
- WEKEO
- CREODIAS



HPC for AI: Math Applications

Linear Algebra service tradeoff

Research question:
to determine the
combination of
functions (services),
satisfying a user
More: Springer Grid
Computing
(10.1007/978-3-642-
28664-3_22)

Large graph simulations

Research question:
algorithms and methods for
in-memory processing and
storage of large graphs
More: Baltic Journal of
Computing
(10.22364/bjmc.2023.11.2.0
2)

Optimization of mathematical workflows

**Research
question:**
performance-
aware and
energy-aware
mathematical
simulations

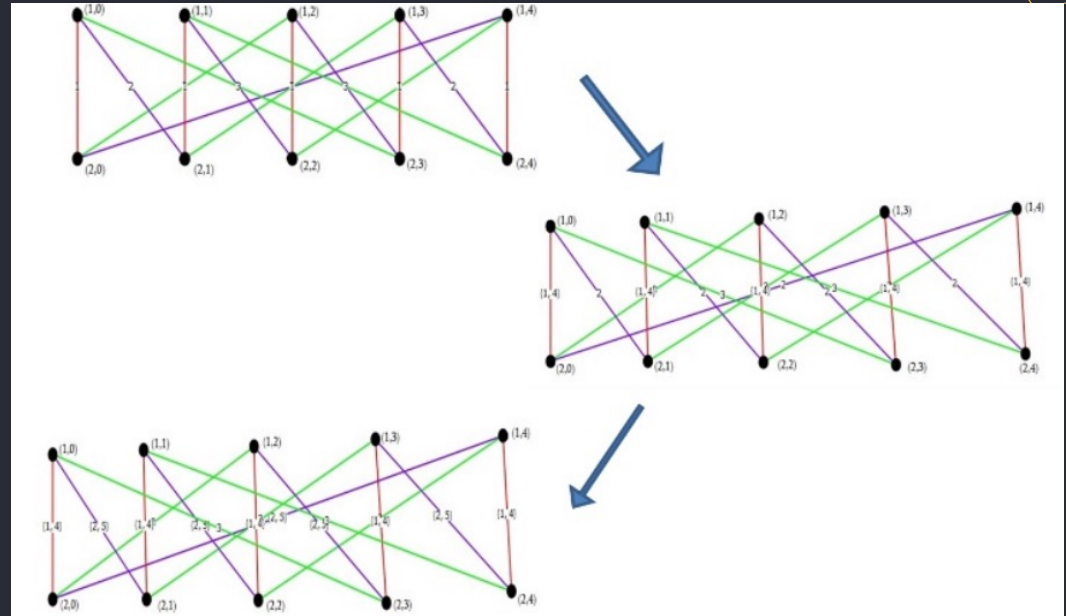
Data quality

**Research
question:** data
quality alerting
models
More:
ADBIS'2019
(10.1007/978-3-
030-30278-8_47)

HPC for AI: Self-organizing Swarms of UAVs

Cloud-based mathematical model designed explicitly for self-organizing UAV swarms

The cloud platform consists of a self-organized UAV-based computing area, a cloud computing environment, and QT service layers



THANK YOU!

DO YOU HAVE ANY QUESTION?



hrach@sci.am



+374 10282030



<https://www.linkedin.com/in/hrachya-astsatryan-97907713/>