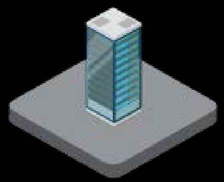
Hewlett Packard Enterprise

What comes after Exascale?

Paolo Faraboschi,
VP and HPE Fellow
Hewlett Packard Labs

February 2023

SYSTEMS OF SYSTEMS

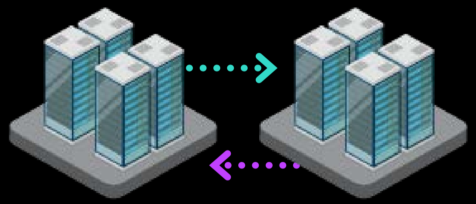


Today

Exascale Supercomputer

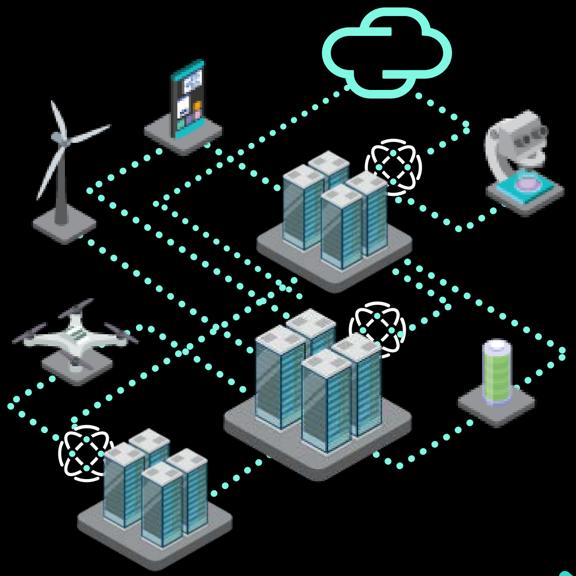
~10x Exascale

Productivity and agility for HPC and AI applications



Federated workflows

For modeling, simulation, data analytics, and artificial intelligence



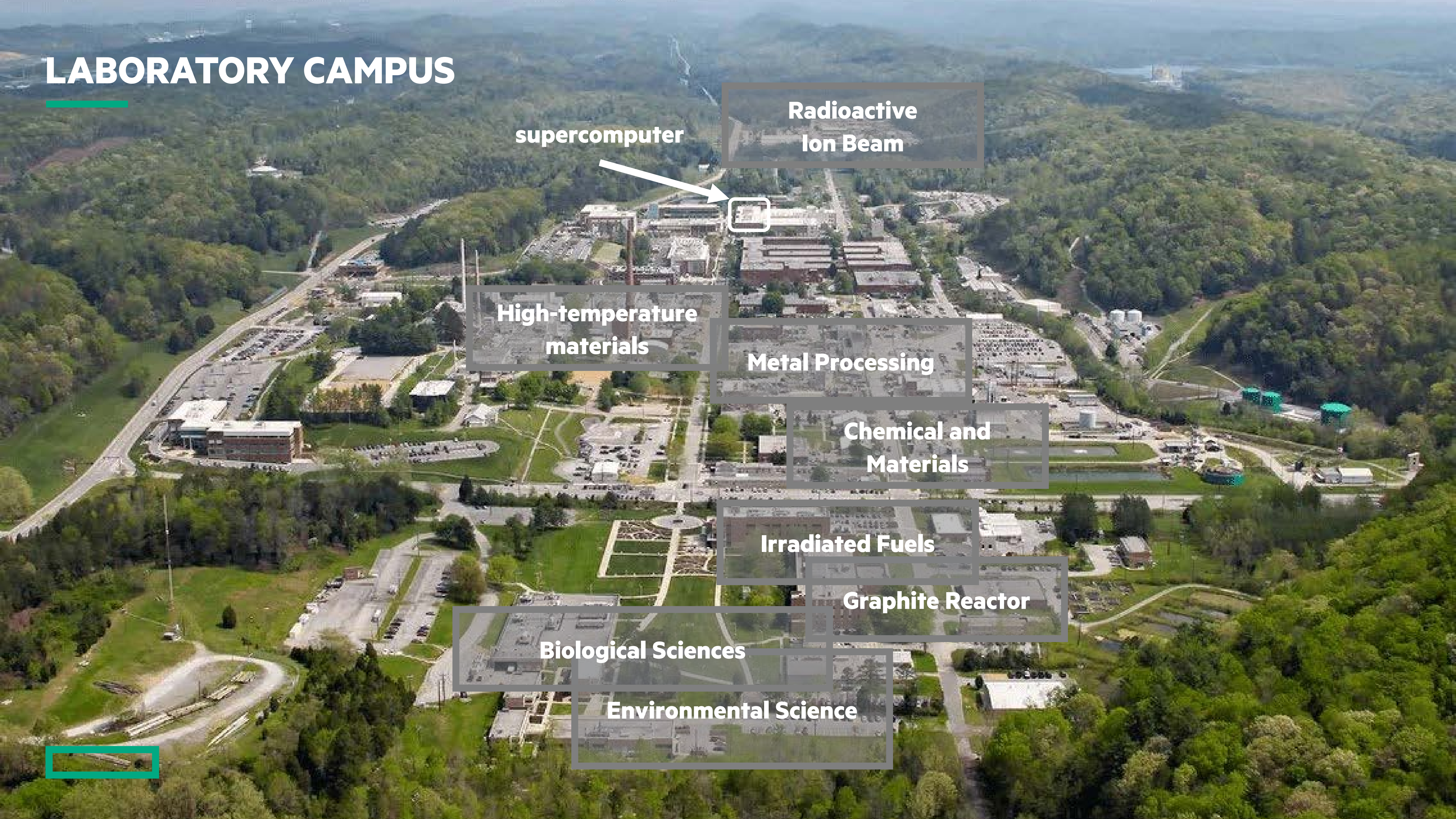
Systems-of-systems

Integrate, automate and optimize workflows that combine theory, simulation, experiments and observations from scientific instruments

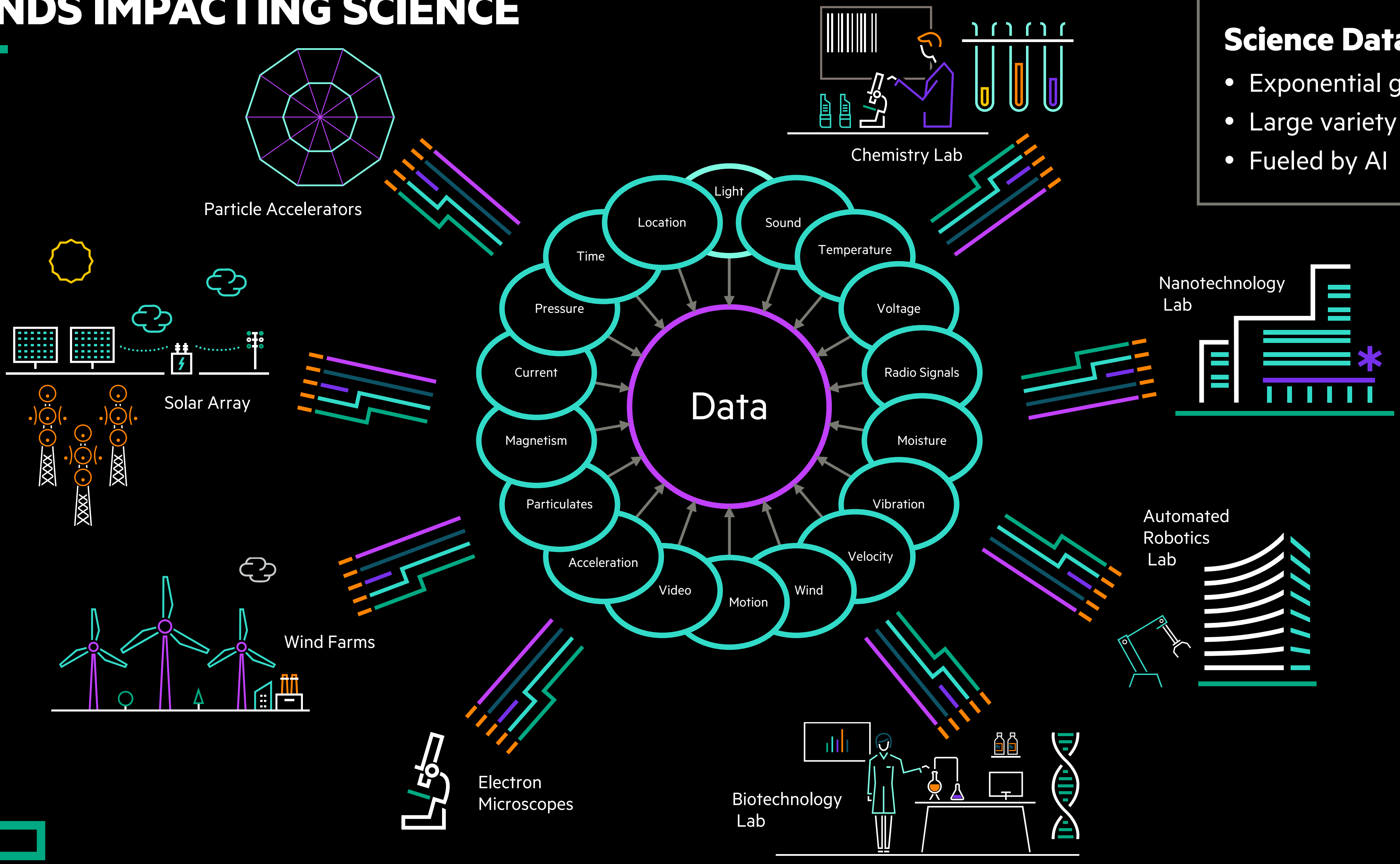
World's fastest
Supercomputers

World's fastest
Workflows

LABORATORY CAMPUS



TRENDS IMPACTING SCIENCE

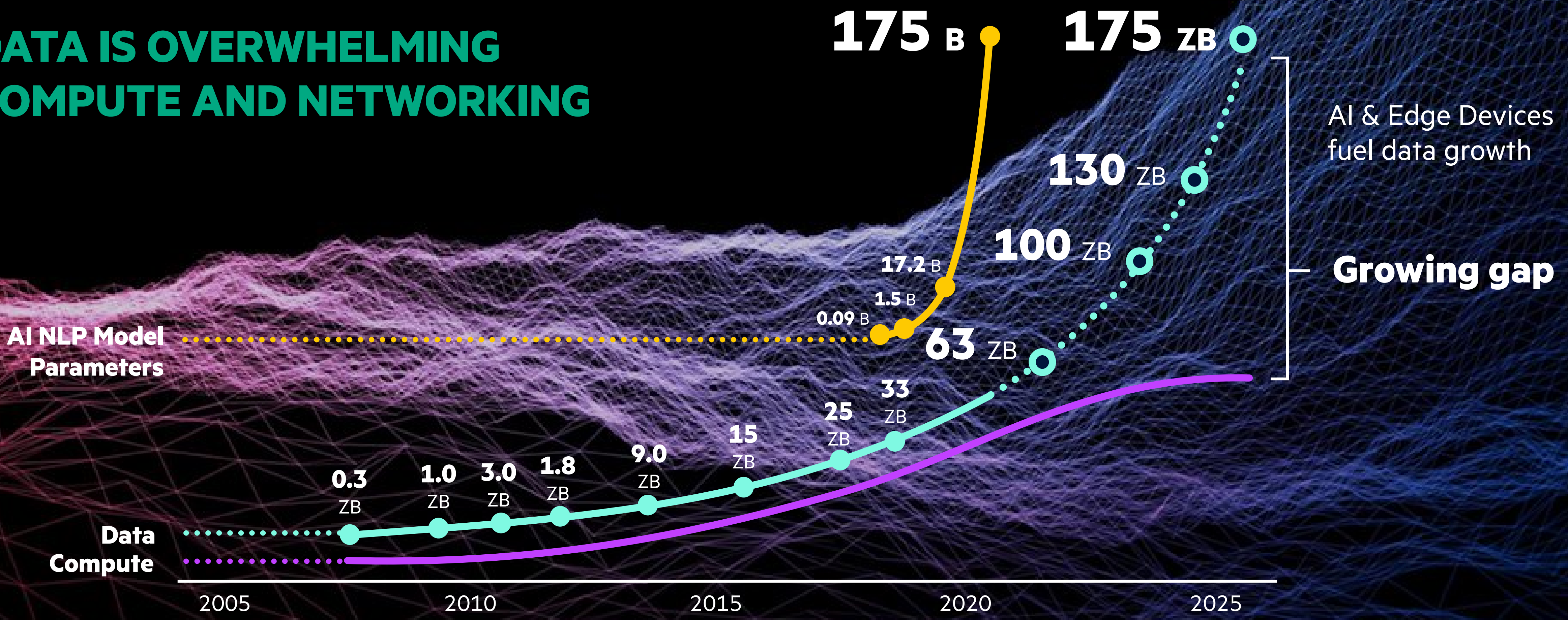


Science Data

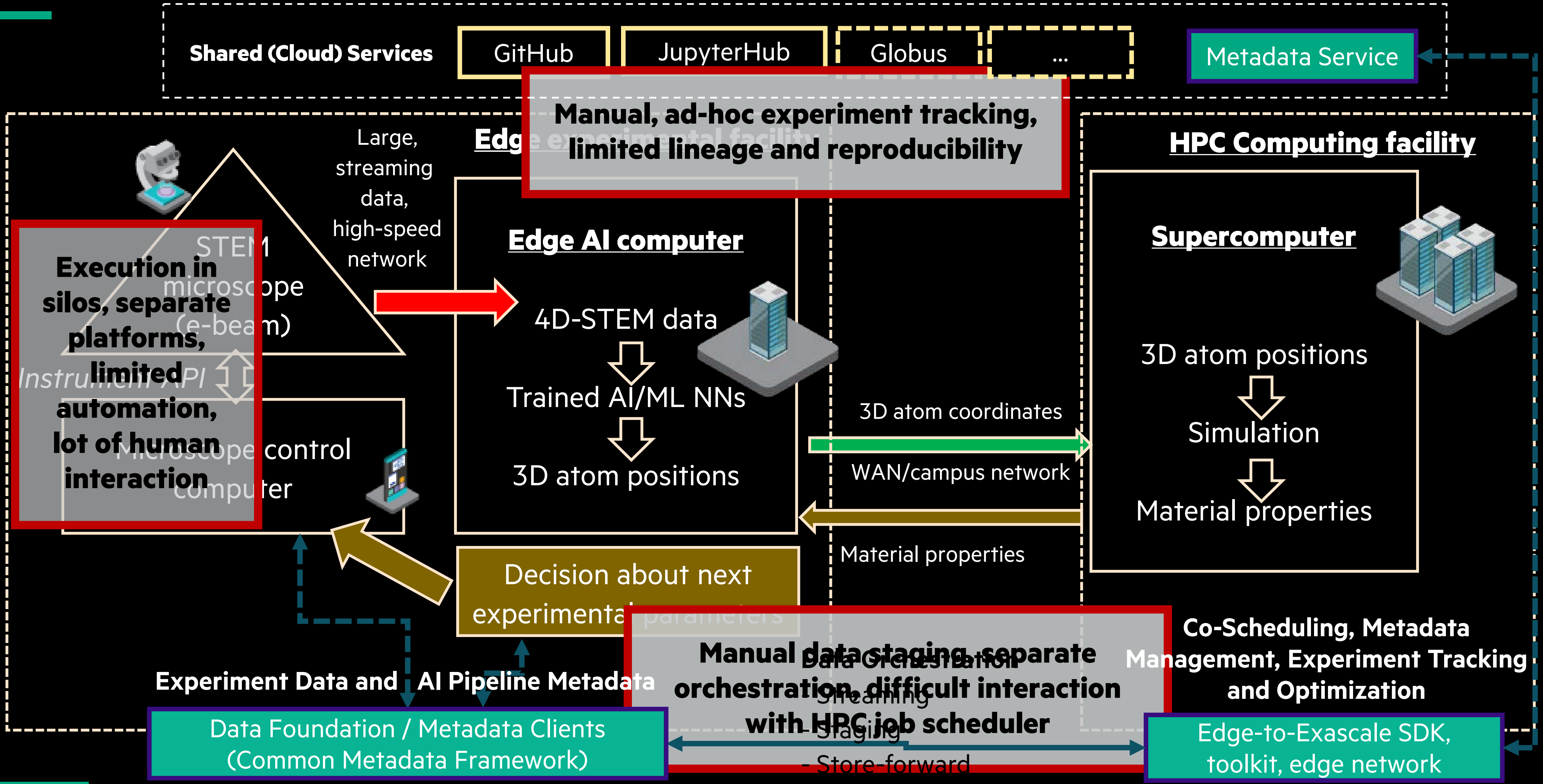
- Exponential growth
- Large variety
- Fueled by AI

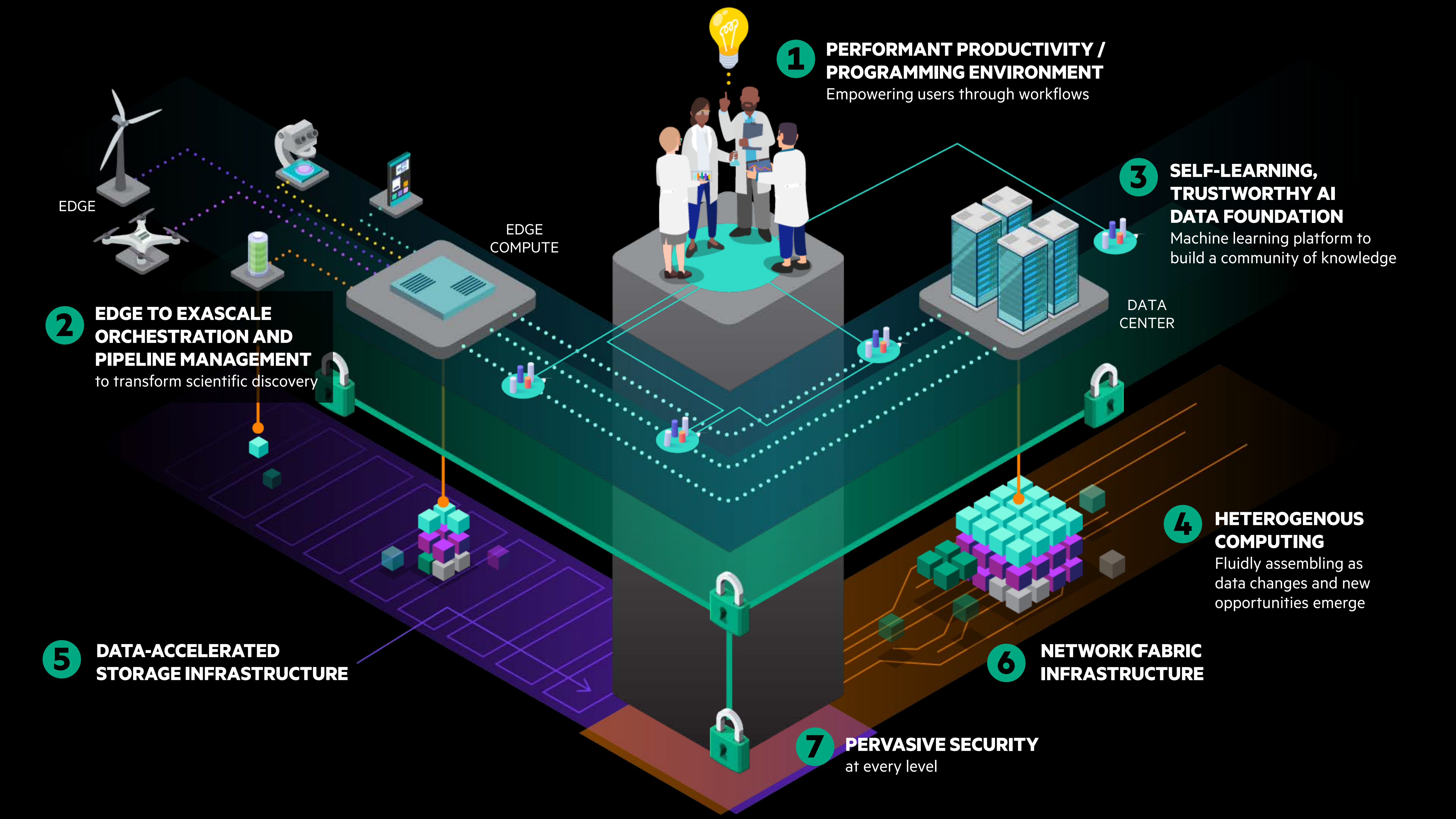
THE DATA AND AI CHALLENGE

DATA IS OVERWHELMING COMPUTE AND NETWORKING



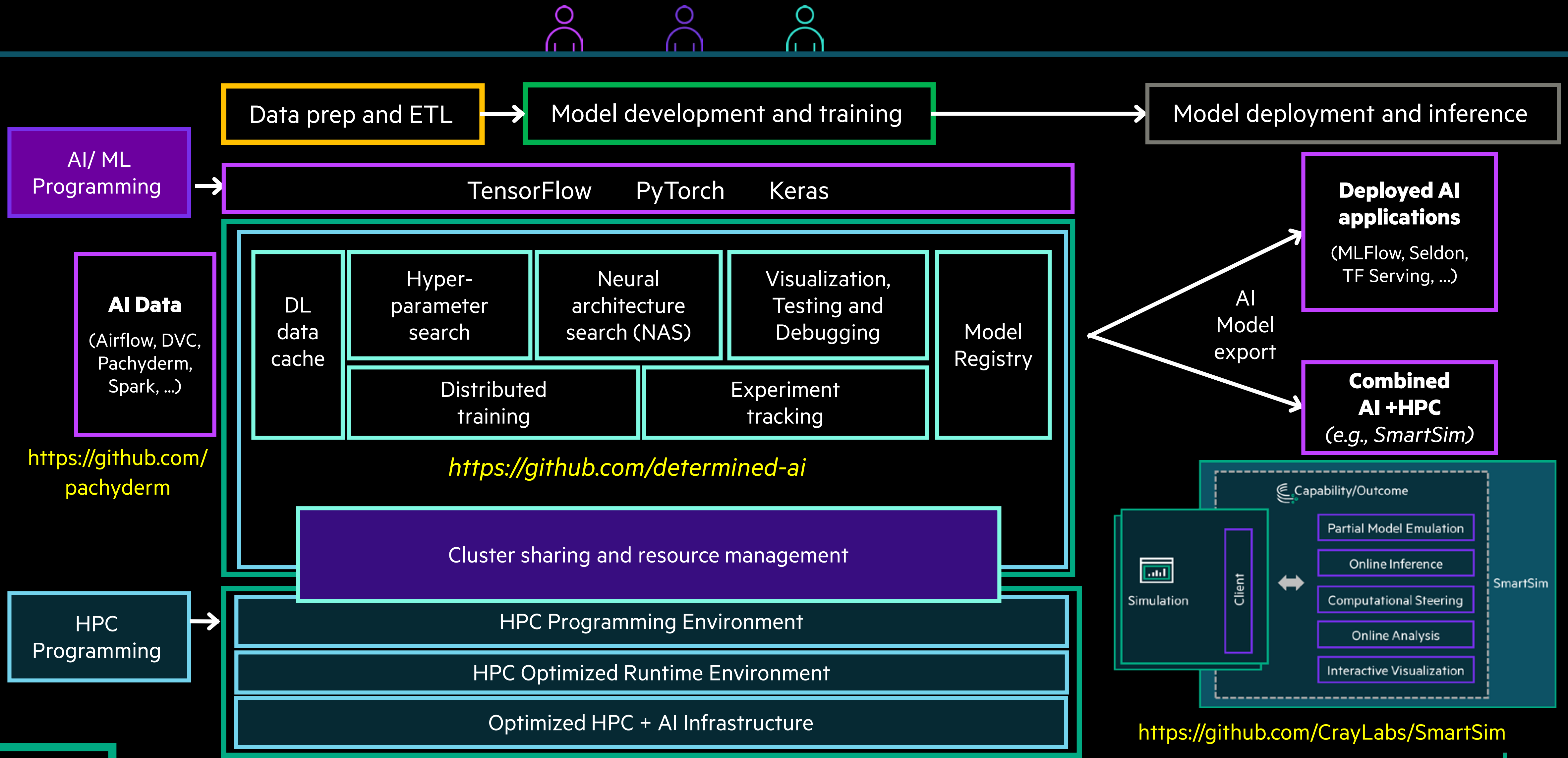
A MOTIVATING EXAMPLE: EDGE-TO-HPC SCIENCE WORKFLOW





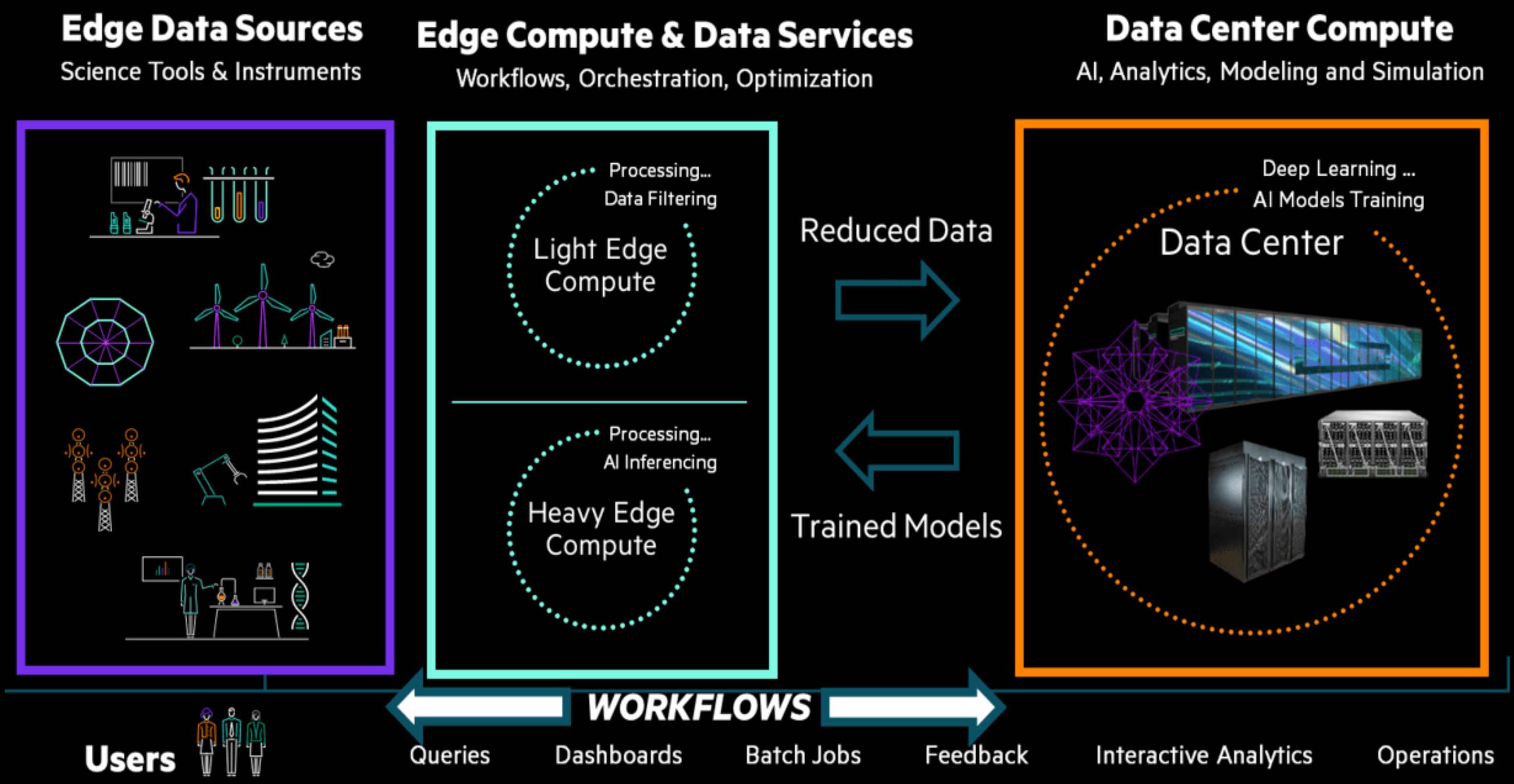
1 PERFORMANT PRODUCTIVITY / PROGRAMMING ENVIRONMENT
Empowering users through workflows
EDGE
EDGE COMPUTE
3 SELF-LEARNING, TRUSTWORTHY AI DATA FOUNDATION
Machine learning platform to build a community of knowledge
DATA CENTER
2 EDGE TO EXASCALE ORCHESTRATION AND PIPELINE MANAGEMENT
to transform scientific discovery
4 HETEROGENOUS COMPUTING
Fluidly assembling as data changes and new opportunities emerge
5 DATA-ACCELERATED STORAGE INFRASTRUCTURE
6 NETWORK FABRIC INFRASTRUCTURE
7 PERVASIVE SECURITY
at every level

(1) PERFORMANCE PRODUCTIVITY & PROGRAMMING ENVIRONMENT



(2) EDGE-TO-EXASCALE ORCHESTRATION

- **Edge-to-Exascale SDK**
  - Helps to manage the lifecycle of edge AI + HPC workflows
  - Coordinates data movement, data transfer nodes, models and job scheduling
- **Edge-friendly job scheduling**
  - Intelligent techniques and tools for orchestrating workflows and refactoring experiments
- **Instrument stack**
  - Software stack based on open platforms



(3) SELF LEARNING, TRUSTWORTHY, AI DATA FOUNDATION

- **Separate layer**
  - Common Metadata Framework (API, client, and server)
  - Agnostic of storage and ML platforms
  - Links data, pipelines, models and outcomes
  - Tracks and learns from workflow history
- **Open, community-based**
  - Built on open-source foundation (Git, DVC, MLMD…)
  - Designed for scale, and exponential growth in scientific data and artifacts
  - Makes data, metadata, models and experiments shareable across teams and communities
- Speed up data-driven scientific discovery

https://github.com/HewlettPackard/cmf

Machine Learning Platform

AI Pipelines
Data Ingestion, Cleanup, Model Engineering, Validation, Serving

Frameworks and MLOps Platforms
Spark, Sklearn, TensorFLow / Pytorch, Mlflow, Kubeflow, etc.

Self-Learning Data Foundation

Northbound Connectors and Intelligence
Data Selection, AI Pipeline Quality, etc.

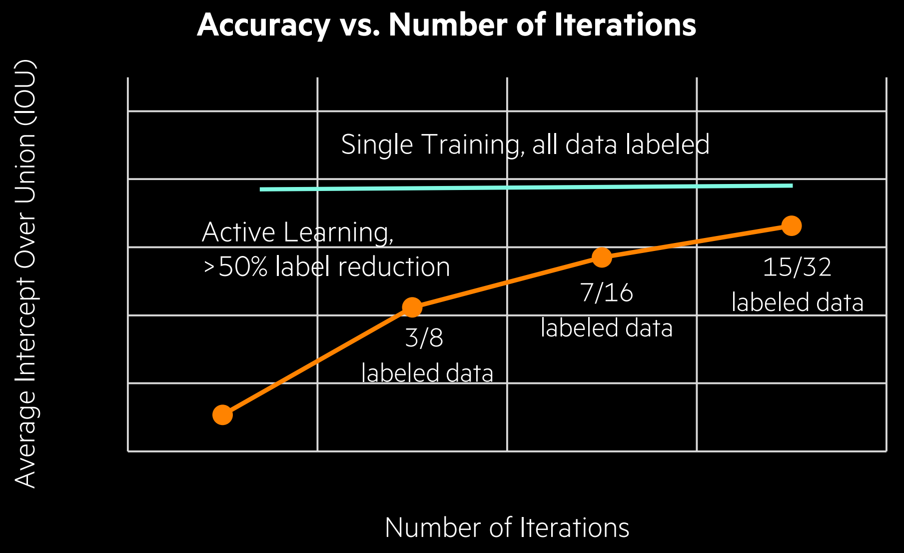
Git for AI Data
Data and Metadata Management, Versioning, Workflow Lineage

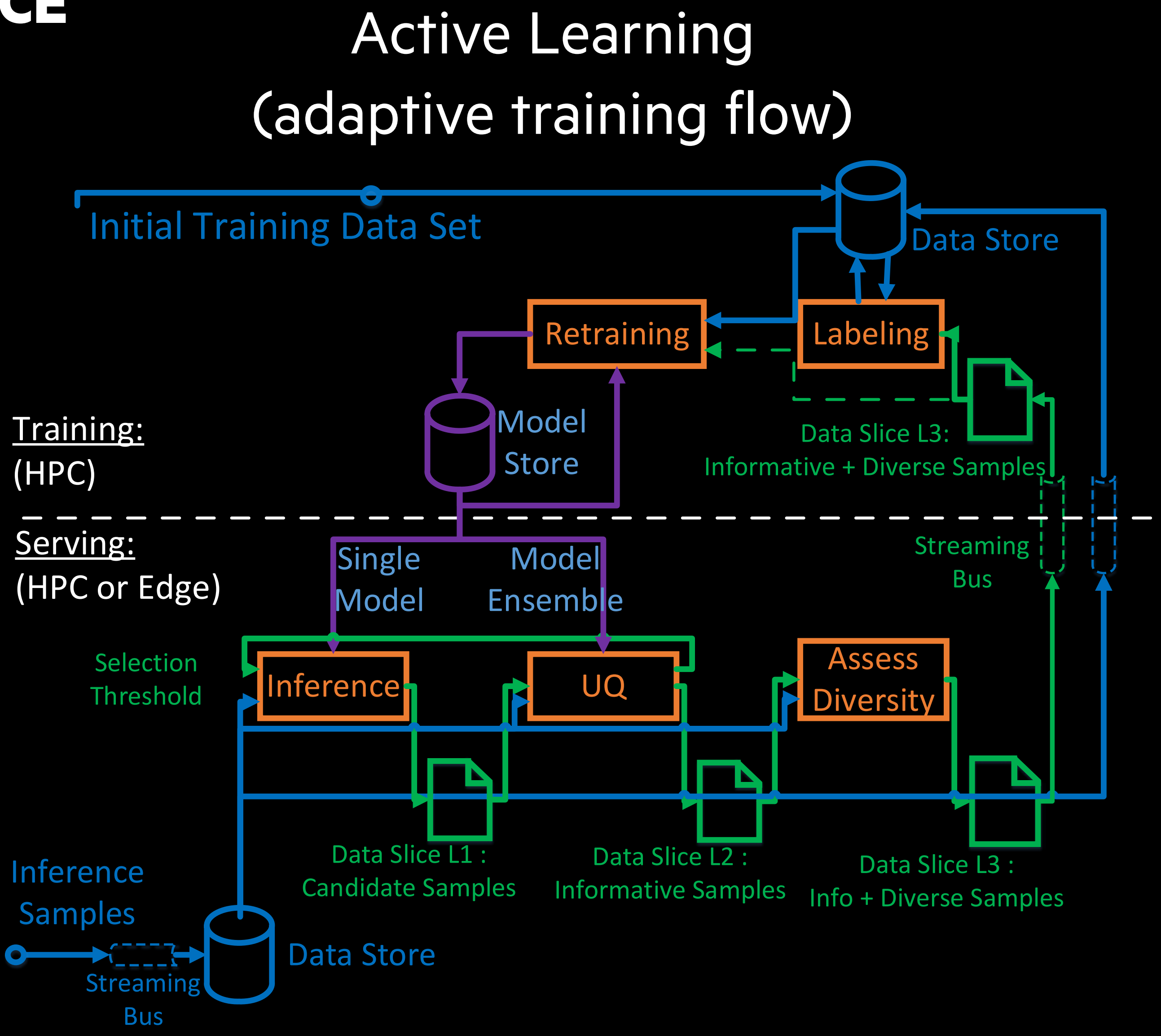
Southbound Connectors and Intelligence
Data Gradation for Tiering & Movement, Data Shaping, etc.

Data Storage Systems
PFS, object stores (HPC, Cloud, Edge)

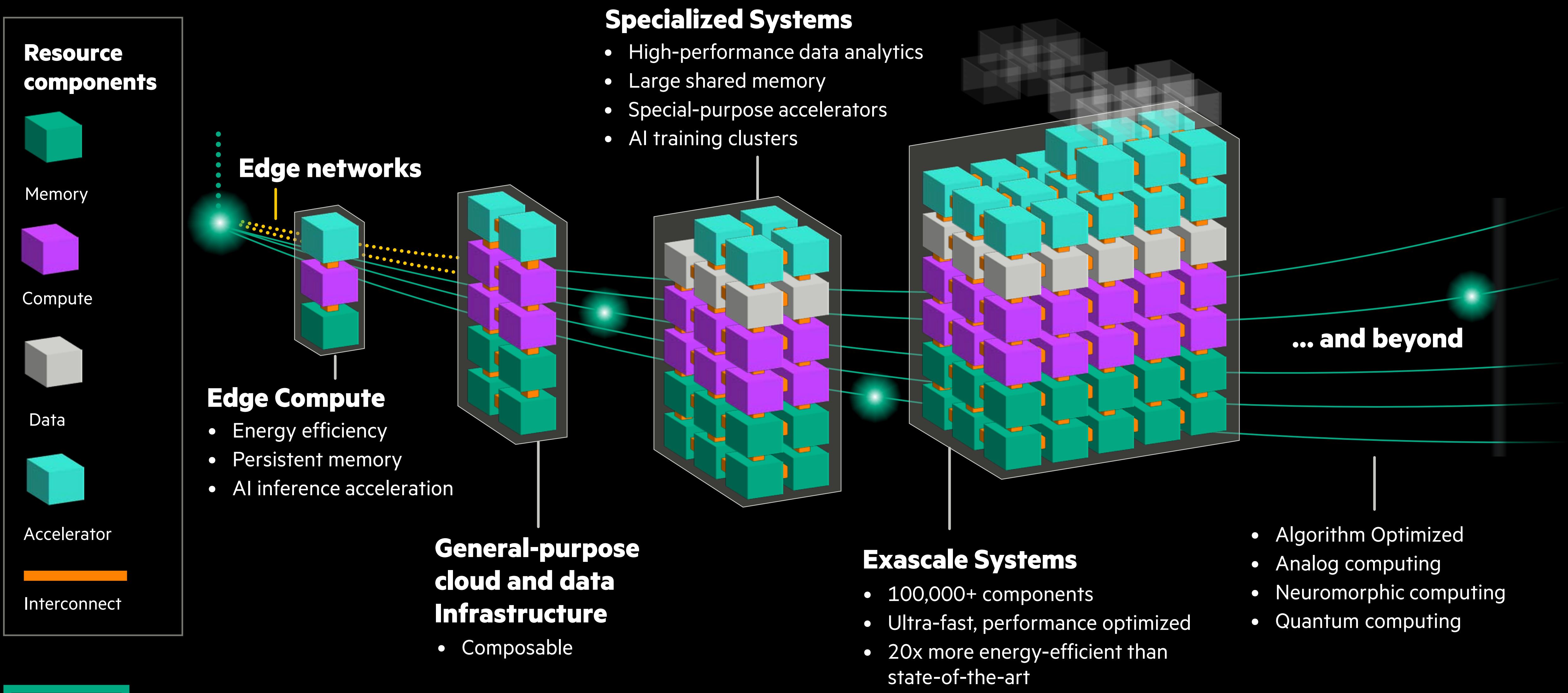
EXAMPLE: NORTHBOUND INTELLIGENCE

- Extract insights from data in complex workflows
- Utilities and infrastructure for lineage and provenance interception of AI/ML workflows
- Meta-learning capabilities: historical correlations help building robust and explainable models
- Identify data of highest importance for quality
- Spans the training and inference flows, closing the loop between monitoring and model update





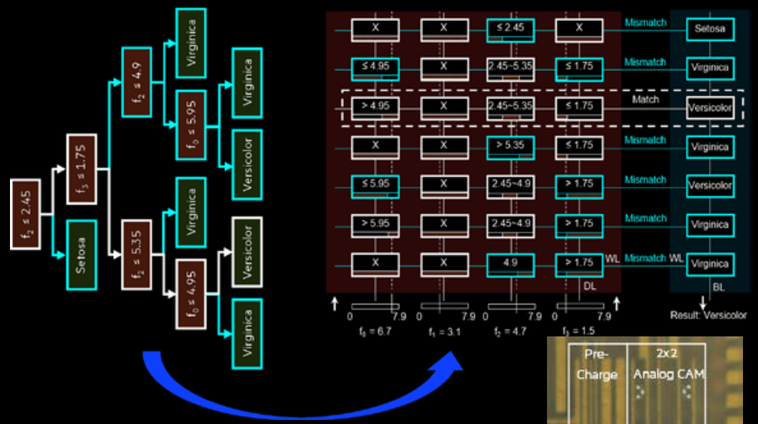
(5) HETEROGENEOUS COMPUTING

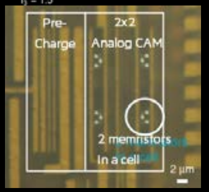


EXAMPLE: ANALOG IN-MEMORY COMPUTING FOR DECISION TREES

X-TIME: eXplainable Tree In-Memory Engine

Analog Computing Element





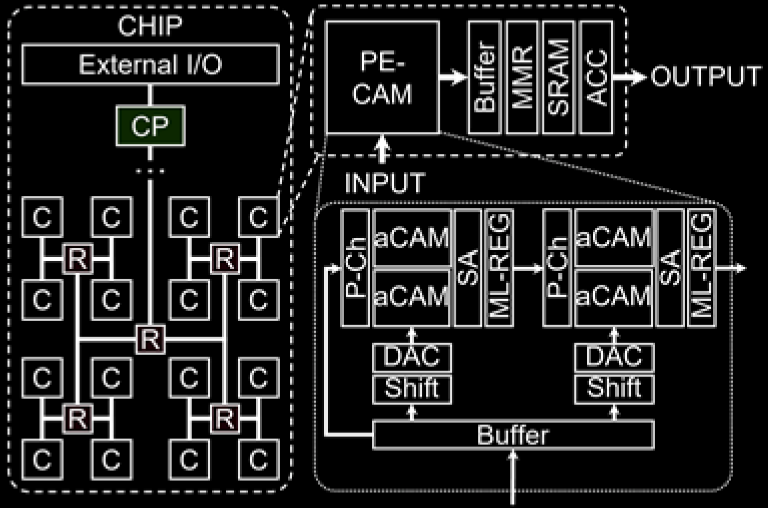
G. Pedretti *et al*,
Nature Comm. 12, 5806 (2021)

68M ACAM elements (4k cores of 256x65) connected with an **H-tree NoC**

On-the-fly parallel reduction avoids overheads for large scale models

20W power budget

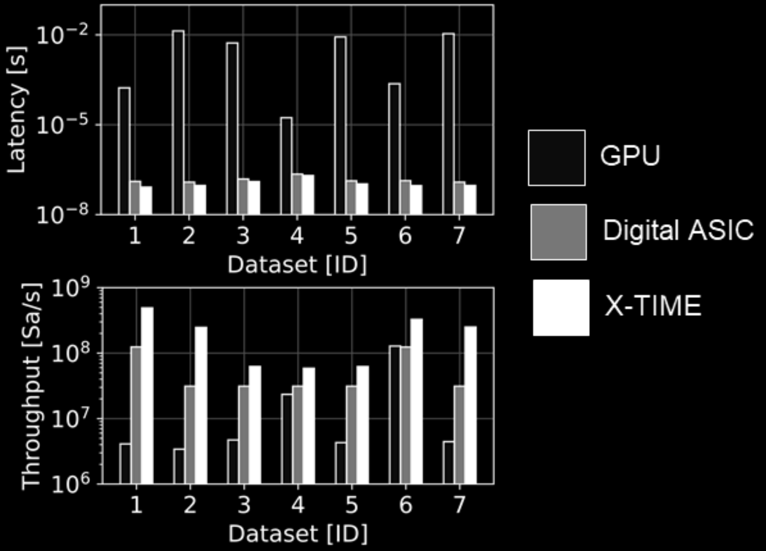
Novel Microarchitecture



Depth and width independent throughput scaling, **massively parallel** node traversal in ACAM

Enables **large scale model inference** without performance loss

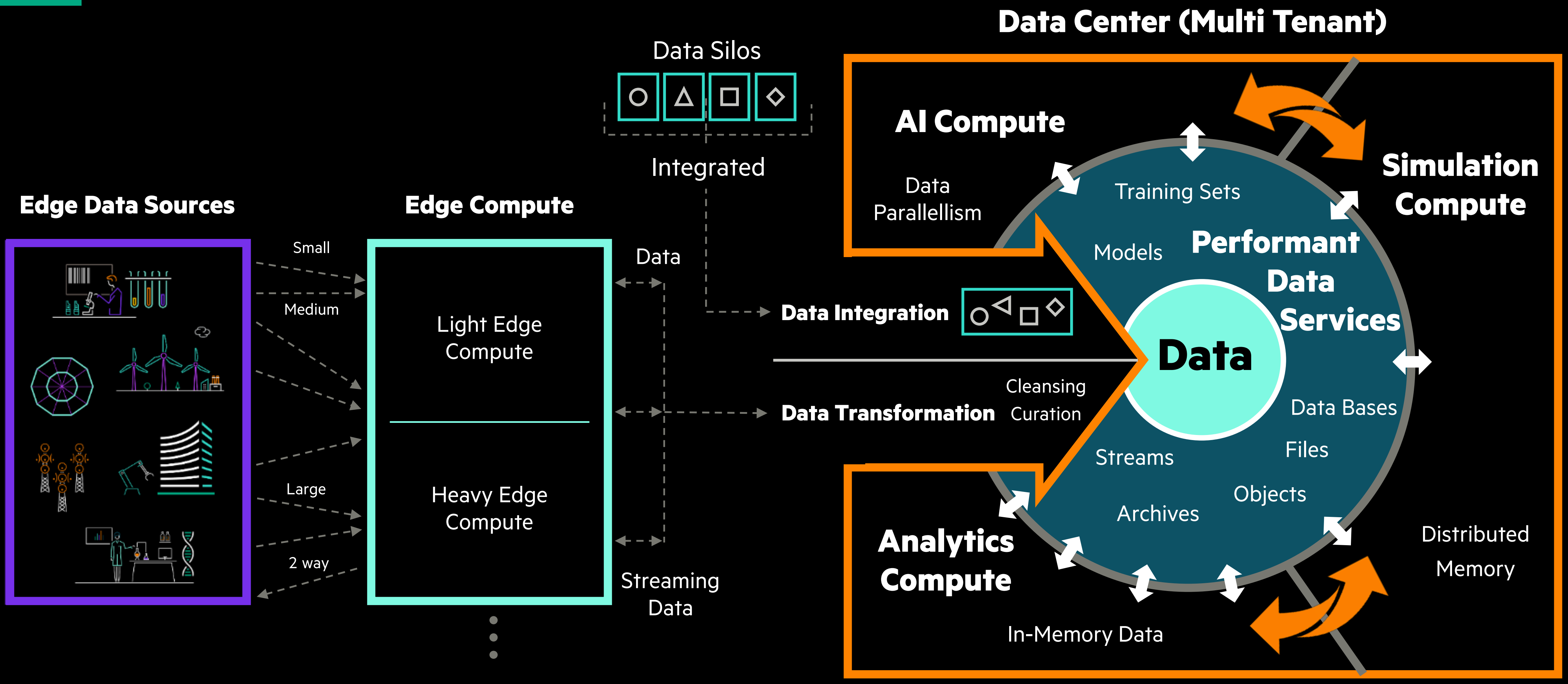
Scalable performance



Up to **9,740x lower latency** compared with GPU (V100) at higher throughput

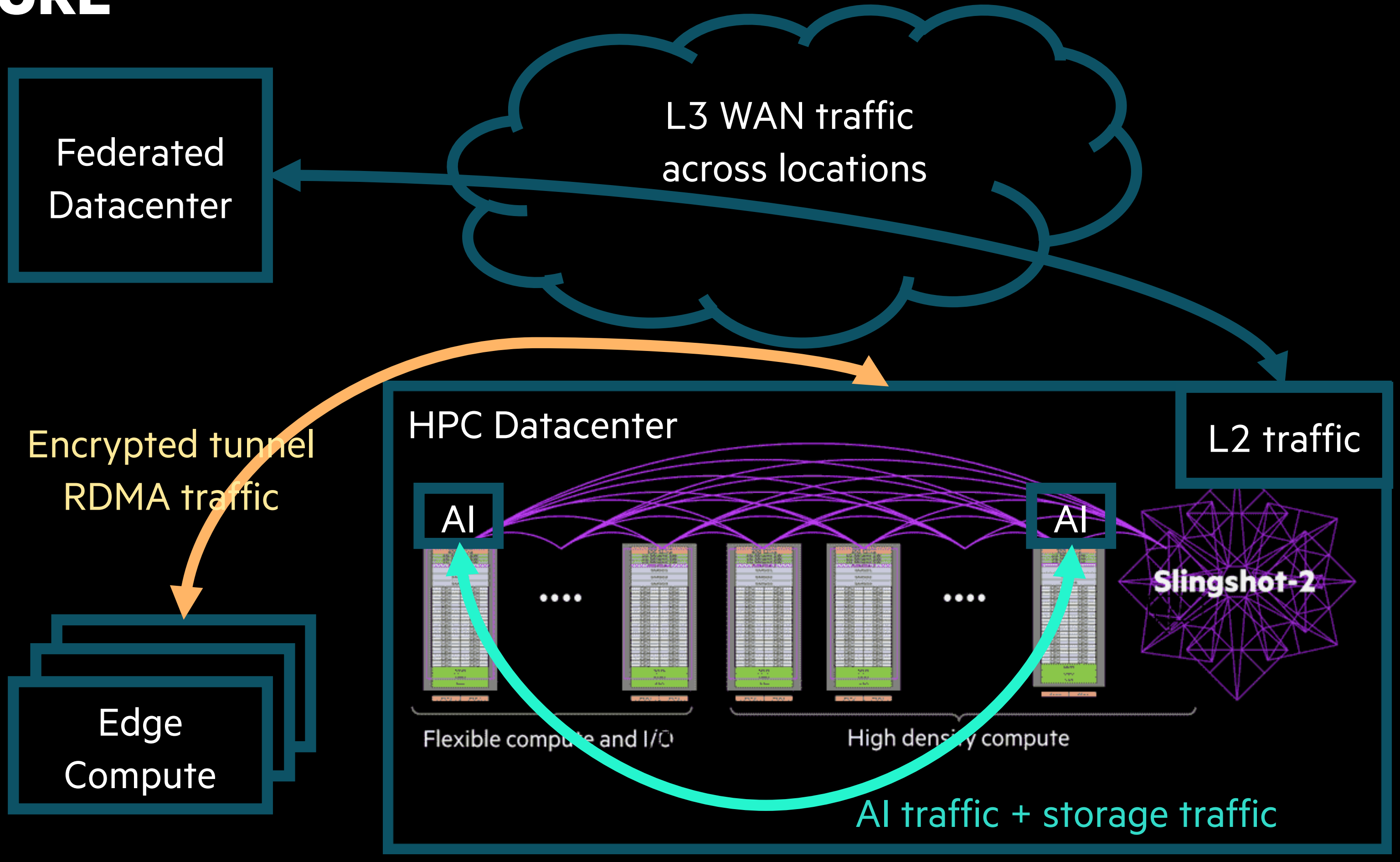
Up to **8x higher throughput** compared to digital ASIC at lower latency

(5) DATA-ACCELERATED STORAGE



(6) NETWORK INFRASTRUCTURE

- **Extending HPC networks to the edge:** compute, storage, devices
- **Accelerators driving injection** (100 Gbps → 200 Gbps)
- **Photonics remains challenging** (copper @ 200 Gbps within rack works!)
- **Blending of supercomputer + cloud:** HPC-aaS, single job with 1,000 instances, QoS, privacy and security
- **Growing interest in HPC functionality** at Ethernet link / transport (RDMA, offloading, isolation, progression, collectives, flow-based congestion, …)



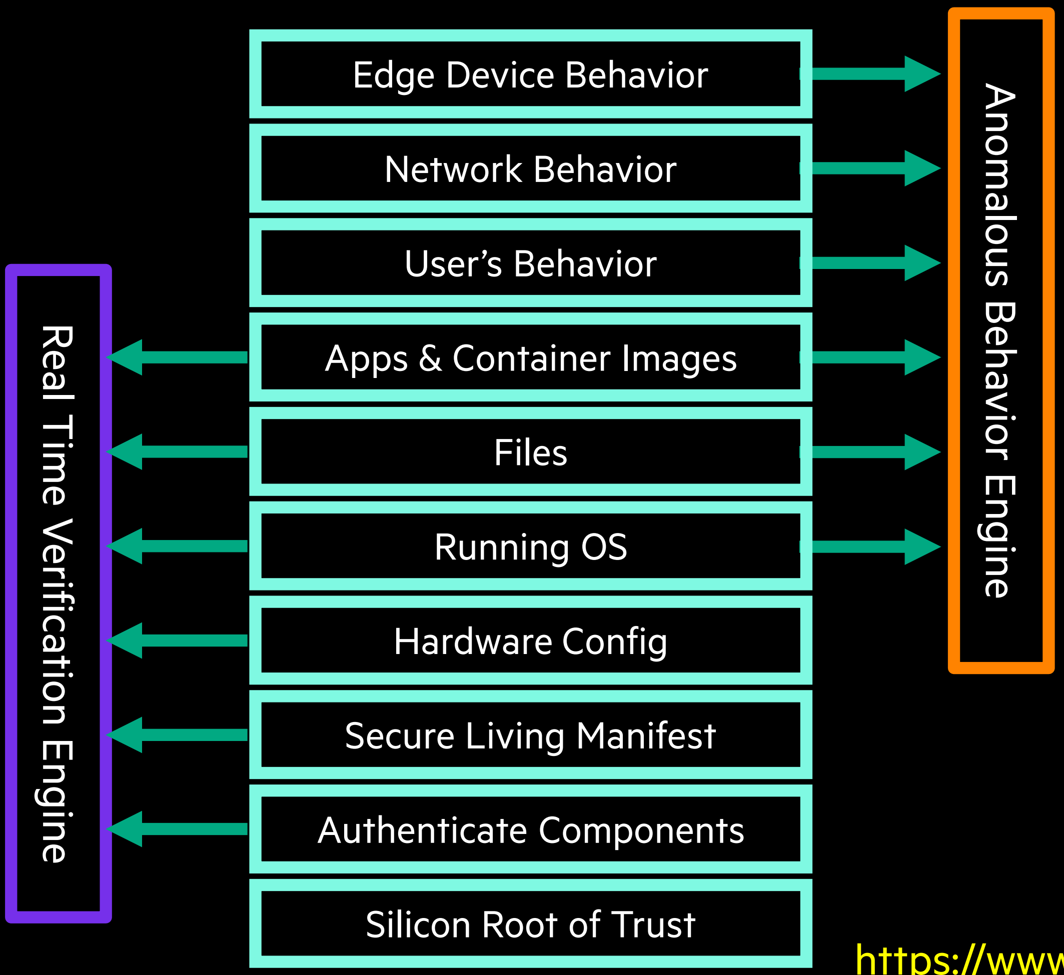
Converged Ethernet for interoperability

Performance at scale

Congestion control + multi-tenancy

(7) PERVASIVE SECURITY

Platform trust + Real time detection



- Advanced Persistent Threats (APT)
  - Breech-to-discovery time: ~100 days *[FireEye 2018]*
- Supply Chain: top security concern of most governments
  - How to secure what is manufactured, delivered, where
- Extend silicon root of trust in real time
  - Detect zero-day threats without signatures
  - Continuous kernel integrity check (in seconds)
  - Verification of pluggable hw components and firmware
- IDevID and Platform Certificates (launched June '21)
  - Factory-issued X.509 server and TCG certificates
  - Manifest of pluggable hardware components
- Deep Supply Chain Attestation
  - Manifest of ***all*** parts, signed, maintained via blockchain

https://www.hpe.com/us/en/security/project-aurora.html

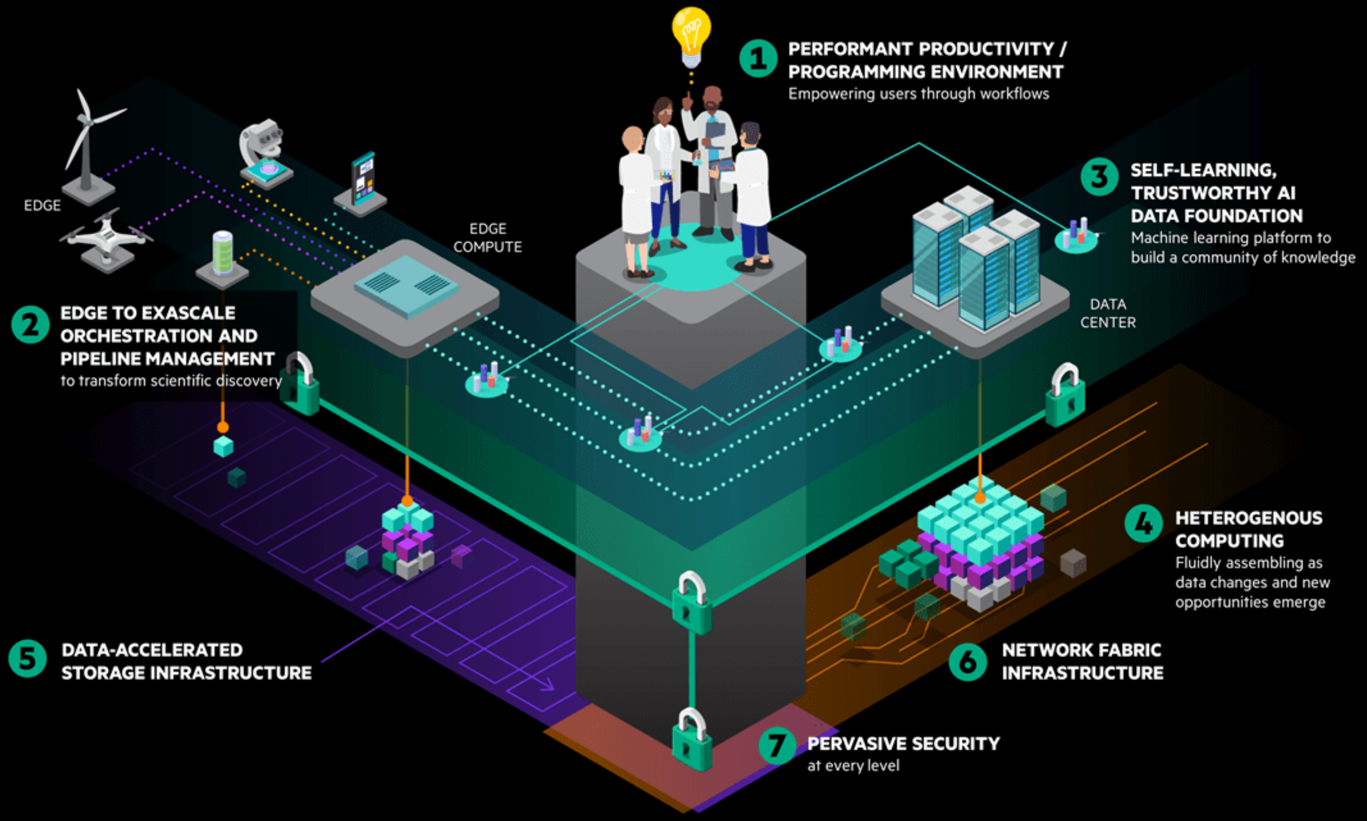
WRAPPING UP

- AI and data are disrupting science
  *"AI will **not** replace scientists,*
  but scientists that use AI
  will replace those who don't"

- Complex science workflows will span from the experimental edge to extreme-scale computing

- Today's Exascale generation may be the last of "monolithic" supercomputers

- The next breakthrough will require a "Systems-of-Systems" view



SYSTEMS OF SYSTEMS