Research in an Open Cloud Exchange

MASSACHUSETTS TECHNOLOGY COLLABORATIVE

BOSTON UNIVERSITY

VE RI TAS

MGHPCC MASSACHUSETTS GREEN HIGH PERFORMANCE COMPUTING CENTER

MIT

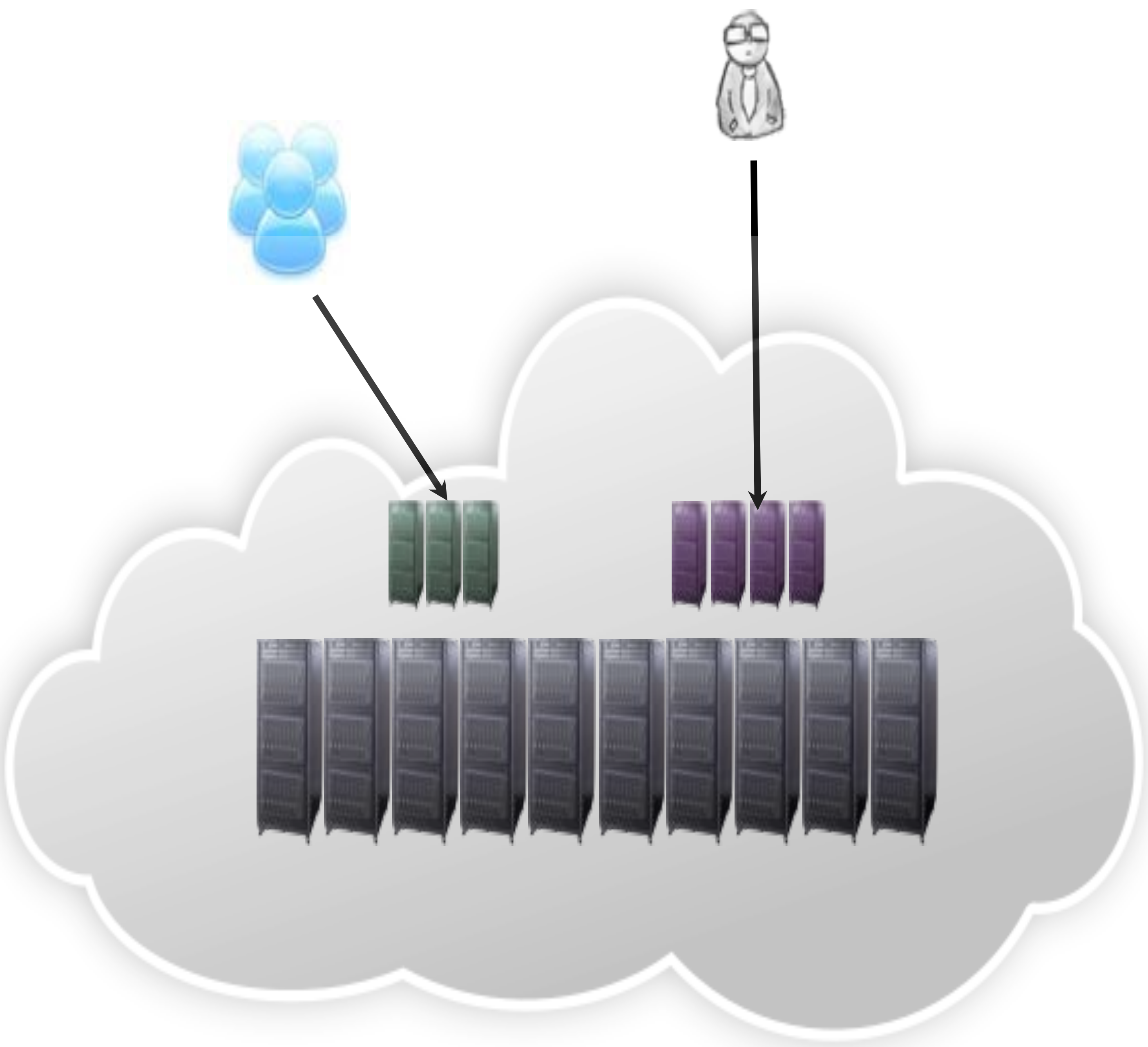
Northeastern

UMASS

CLOUD COMPUTING IS HAVING A DRAMATIC IMPACT

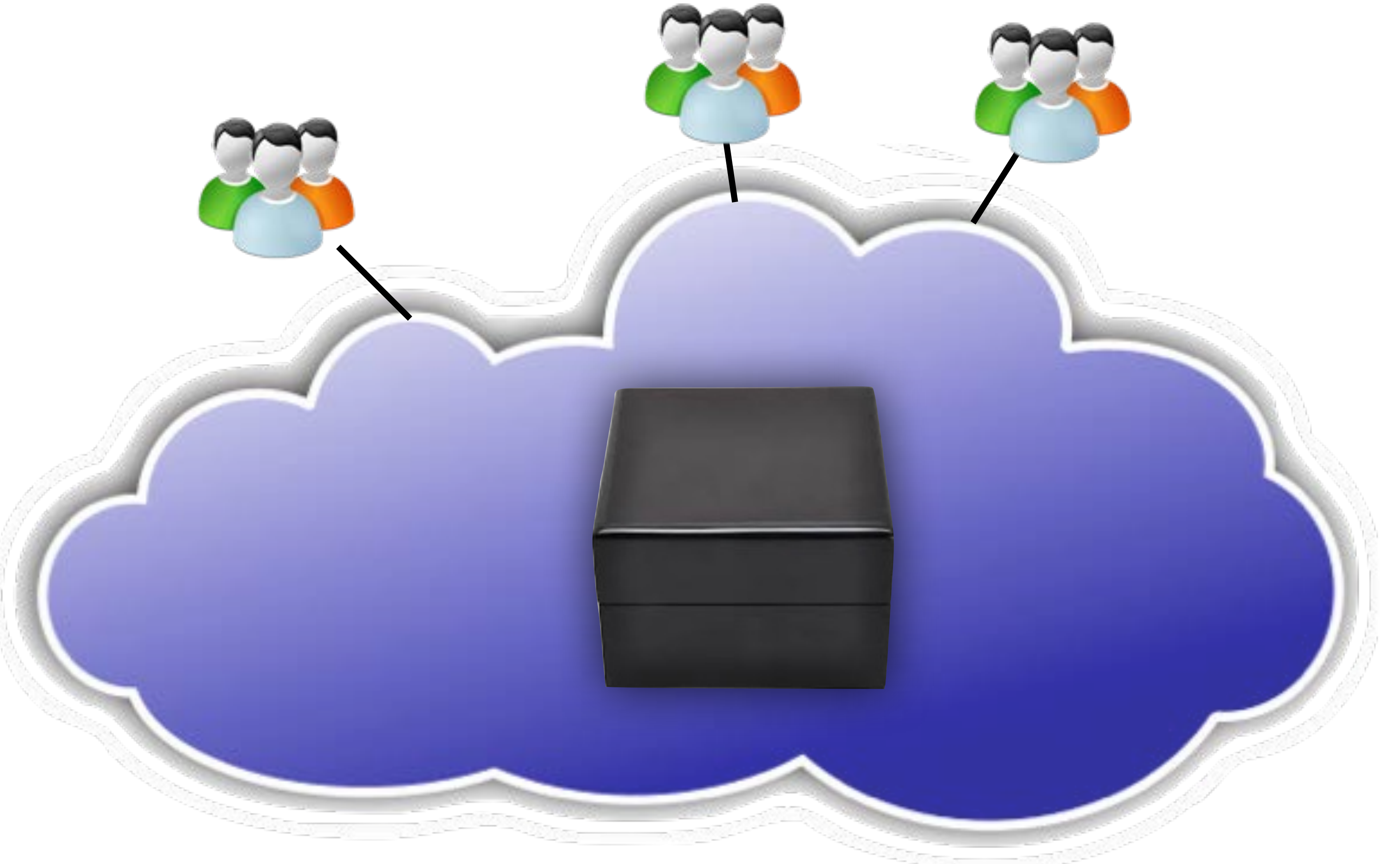
- On-demand access
- Economies of scale

All compute/storage will move to the cloud?



Today's IaaS clouds

- One company responsible for implementing and operating the cloud
- Typically highly secretive about operational practices
- Exposes limited information to enable optimizations



What's the problem

- Lots of innovation above the IaaS level… but
  - consider EnterpriseDB, or Akamai
- Lots of different providers… but
  - bandwidth between providers limited
  - offerings incompatible; switching a problem
  - price challenges to moving
- No visibility/auditing internal processes
- Price is terrible for computers run 24x7x365

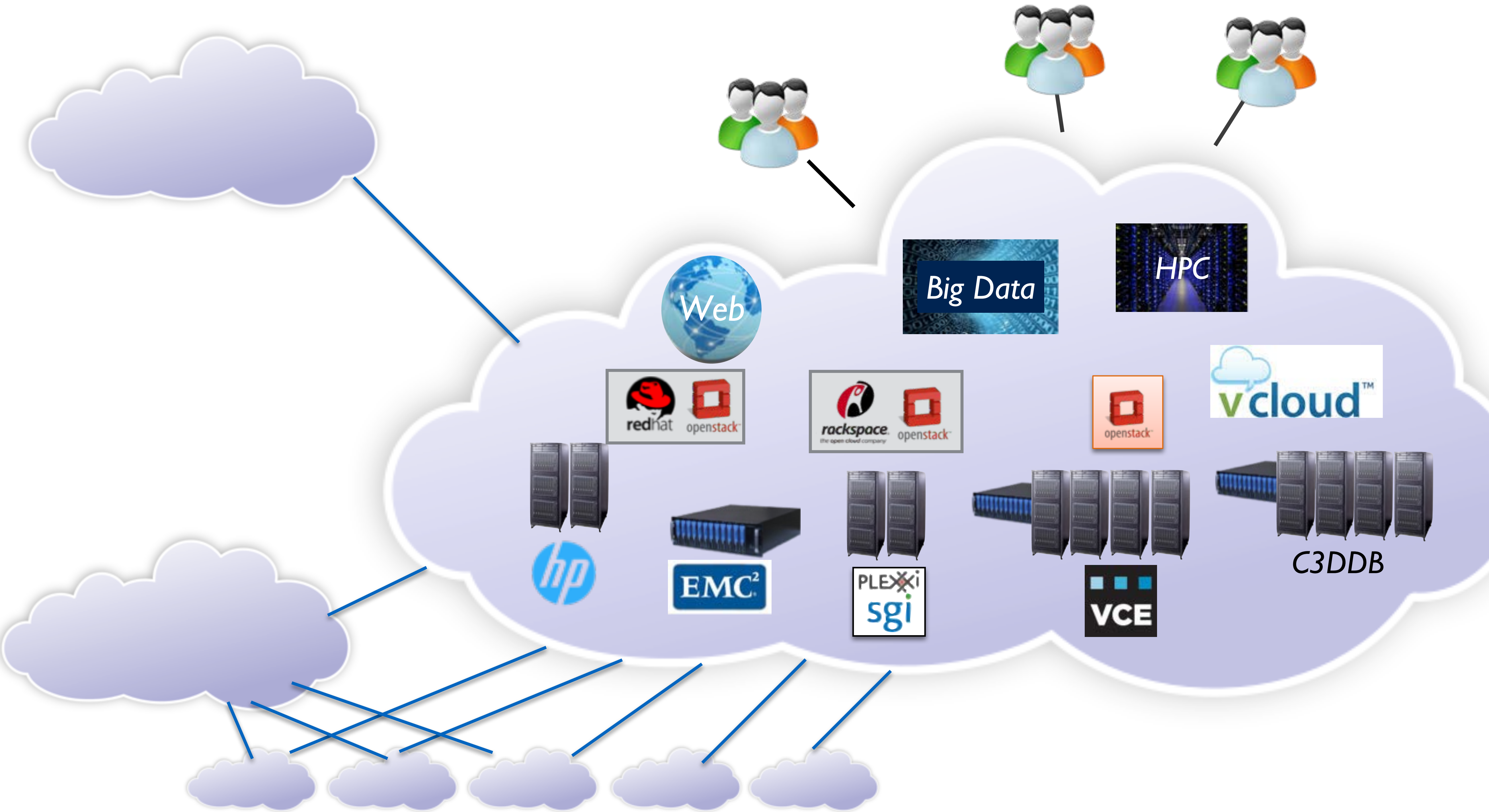
More challenges

- Provider incentive not aligned with efficient marketp
  - stickin
  - advan
  - homo

We are in the equivalent of the pre-Internet world, where AOL and CompuServe dominated on-line access

- Hard for large provider to efficiently support niche markets, radically different economic models…
- Niche providers probably can't support rich ecosystem

Is a different model possible?
An "Open Cloud eXchange (OCX)"



BIG BOX STORE

SHOPPING MALL

CATHEDRAL

BAZAAR

Why is this important

- Anyone can add a new service and compete in a level playing field
- History tells us the opening up to rich community/marketplace competition results in innovation/efficiency:
  - “The Cathedral and the Bazaar” by Eric Steven Raymond
  - “The Master Switch: The Rise and Fall of Information Empires” by Tim Wu
- This could fundamentally change systems research:
  - access to real data
  - access to real users
  - access to scale

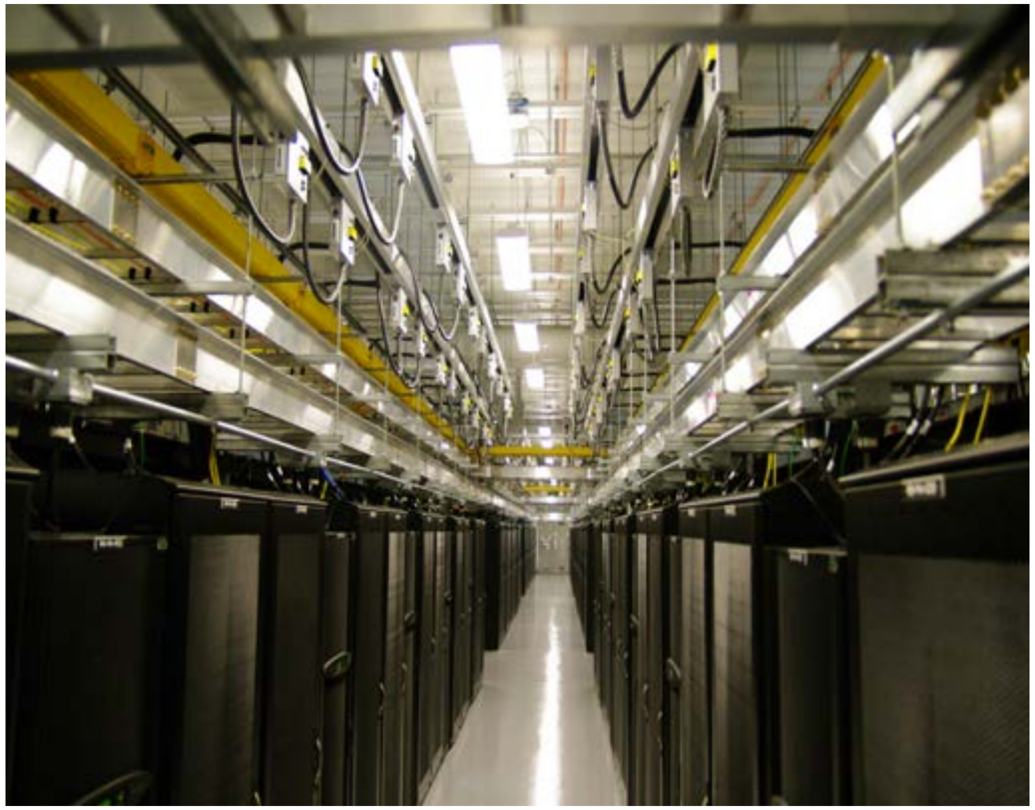
Without that…solving the spherical horse problem…

This isn’t crazy… really

- Current clouds are incredibly expensive…
- Much of industry locked out of current clouds
- lots of great open source software
- lots of great niche markets; markets important to us…
- lots of users concerned by vendor lock in…
- this doesn’t need to be AWS scale to be worth it
  - “Past a certain scale; little advantage to economy of scale” — John Goodhue

The Massachusetts Open Cloud

MGHPCC



15 MW, 90,000 square feet + can grow

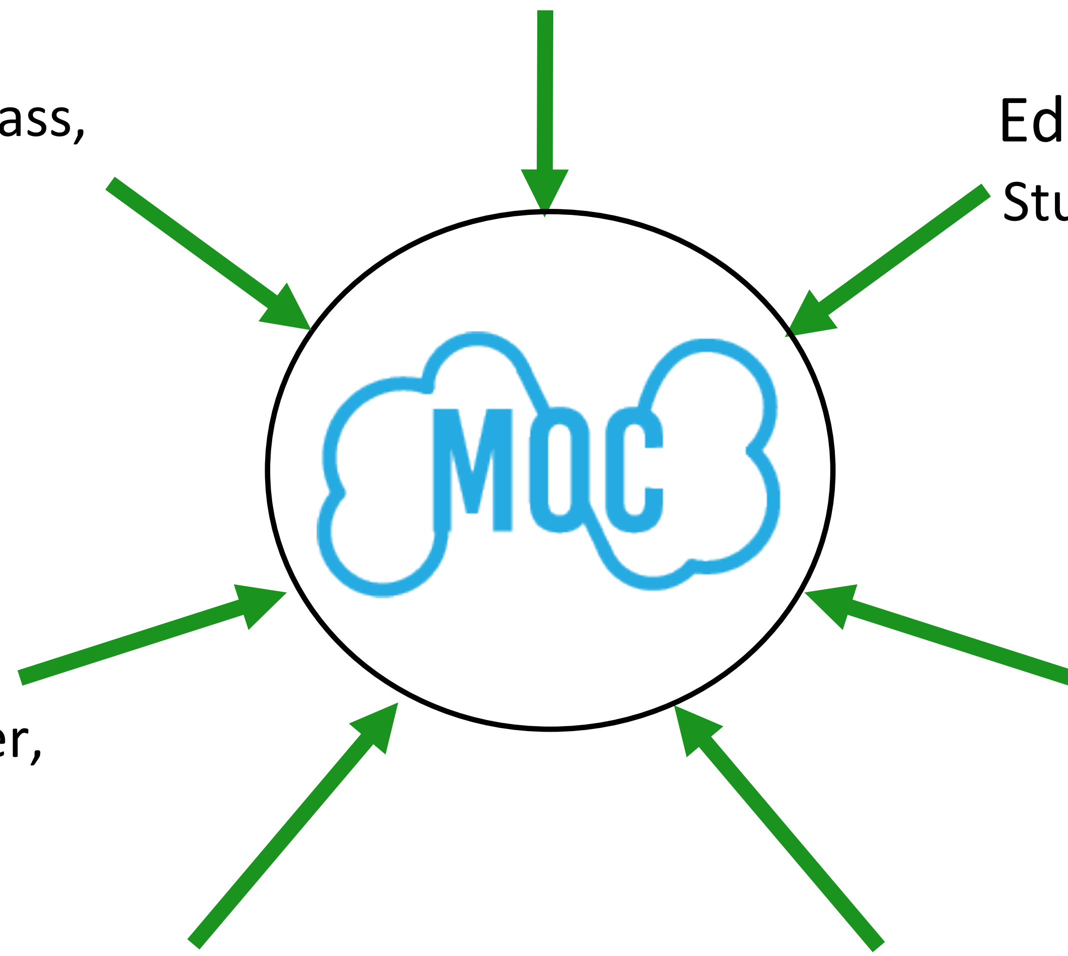
THE MASSACHUSETTS COLLABORATORS

MOC Ecosystem

Users/applications
BigData, HPC, Life Sciences, …

Data
BU, HU, NU, MIT, UMass, Foundations, Govt…

Education and Workforce
Students, industry



Cloud Technology
Operating Systems, Power, Security, Marketplace…

Partners
Brocade, CISCO, Intel, Lenovo, Red Hat, Two Sigma, USAF, Dell, Fujitsu, Mellanox, Cambridge Computer…

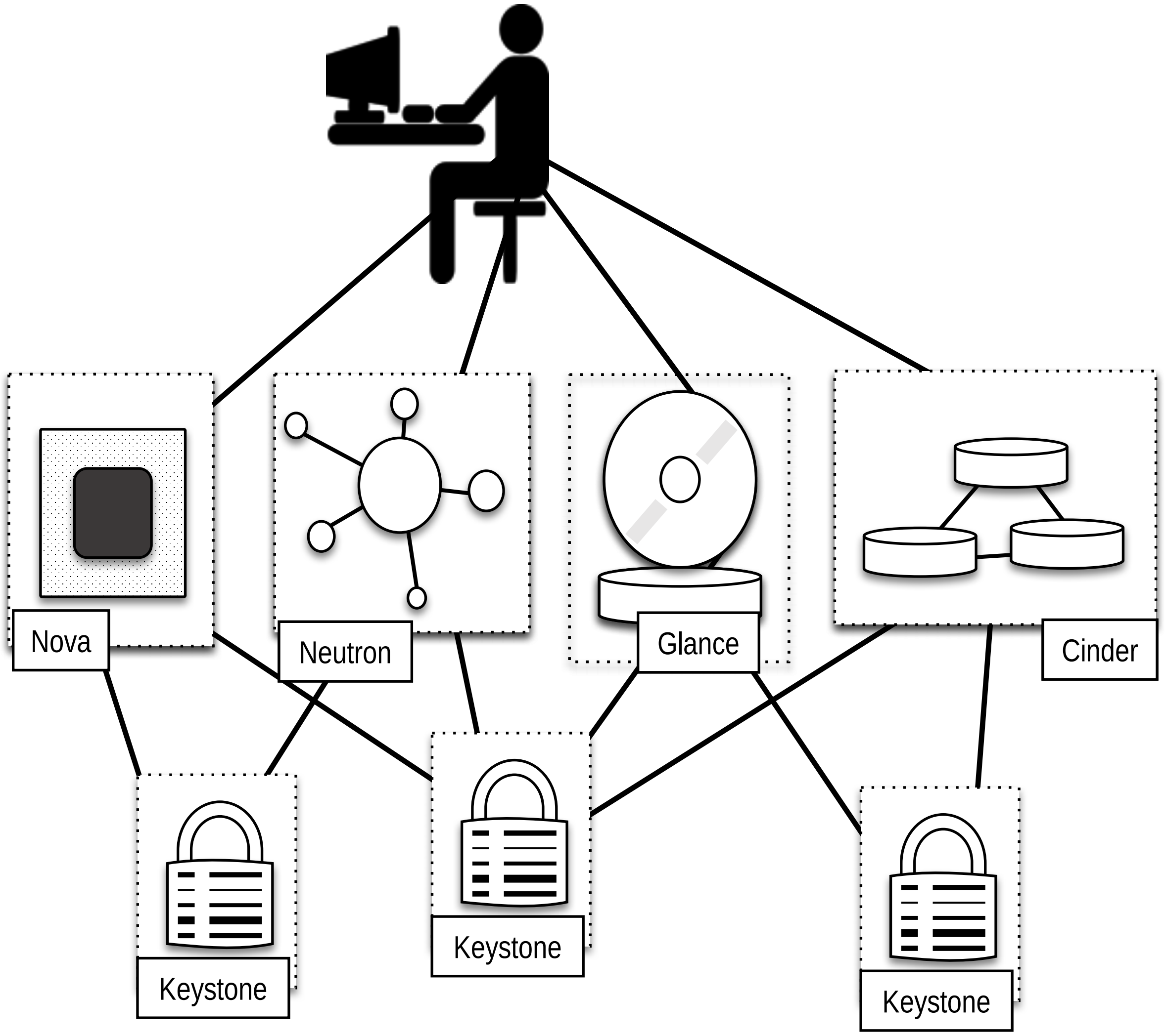
Core Team & Students
OCX model, HIL, Billing, Intermediaries…

University Research IT Partners
BU, HU, NU, UMass, MIT, MGHPCC

HOW DO WE START?

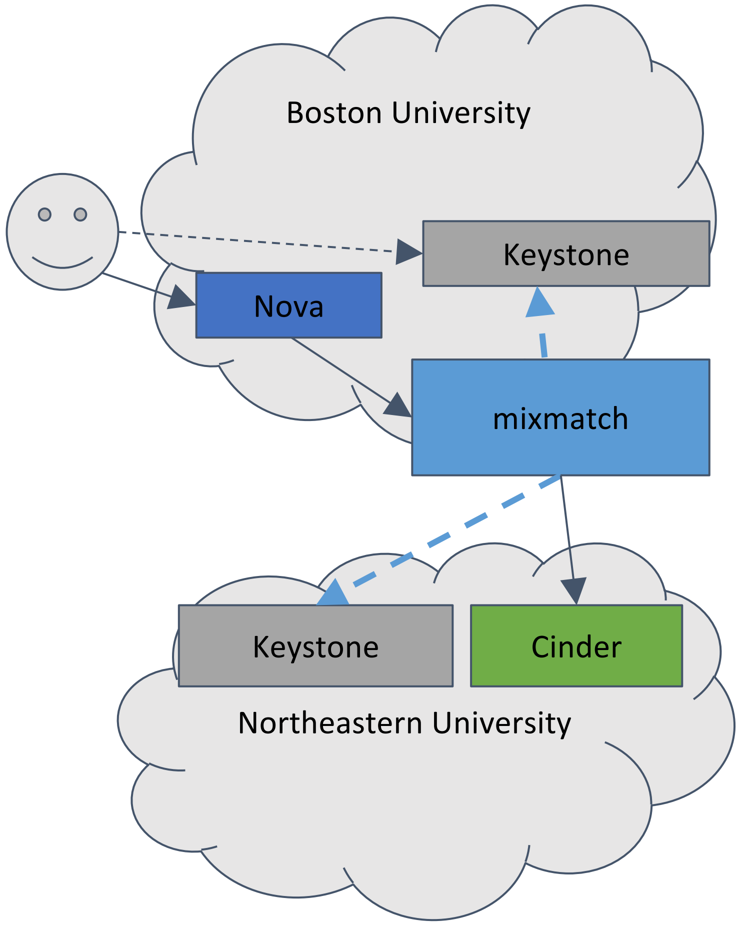
OPENSTACK FOR AN OCX

- OpenStack is a natural starting point
- Mix & Match federation



Mix and Match (Resource Federation)

- Solution
  - Proxy between OpenStack services
- Status of the project
  - Hosted upstream by the OpenStack infrastructure
    - https://github.com/openstack/mixmatch
  - Production deployment planned for Q1 2017
- Team:
  - Core Team: Kristi Nikolla, Eric Juma, Jeremy Freudberg
  - Contributors: Adam Young (Red Hat), George Silvis, Wjdan Alharthi, Minying Lu, Kyle Liberti
- More information:
  - https://info.massopencloud.org/blog/mixmatch-federation/



It's real…

- Available now: Production OpenStack services…
    - Small scale, but growing (couple of hundred servers, 550 TB storage), 200+ users
    - VMs, on-demand Big Data (Hadoop, SPARK...),

- What's coming:
  - Simple GUI for end users
  - OpenShift – Red Hat
  - Federation across universities
  - Rapid/secure Hardware as a Service
  - 20+ PB DataLake
  - Cloud Dataverse

- Platform for enormous range of research projects across BU, NEU, MIT & Harvard

Research challenges

- Marketplace mechanisms
- Hosting Datasets
- Multi-provider cloudlet
- Software defined storage
- HPC on the Cloud
- Secure Hardware Multiplexing

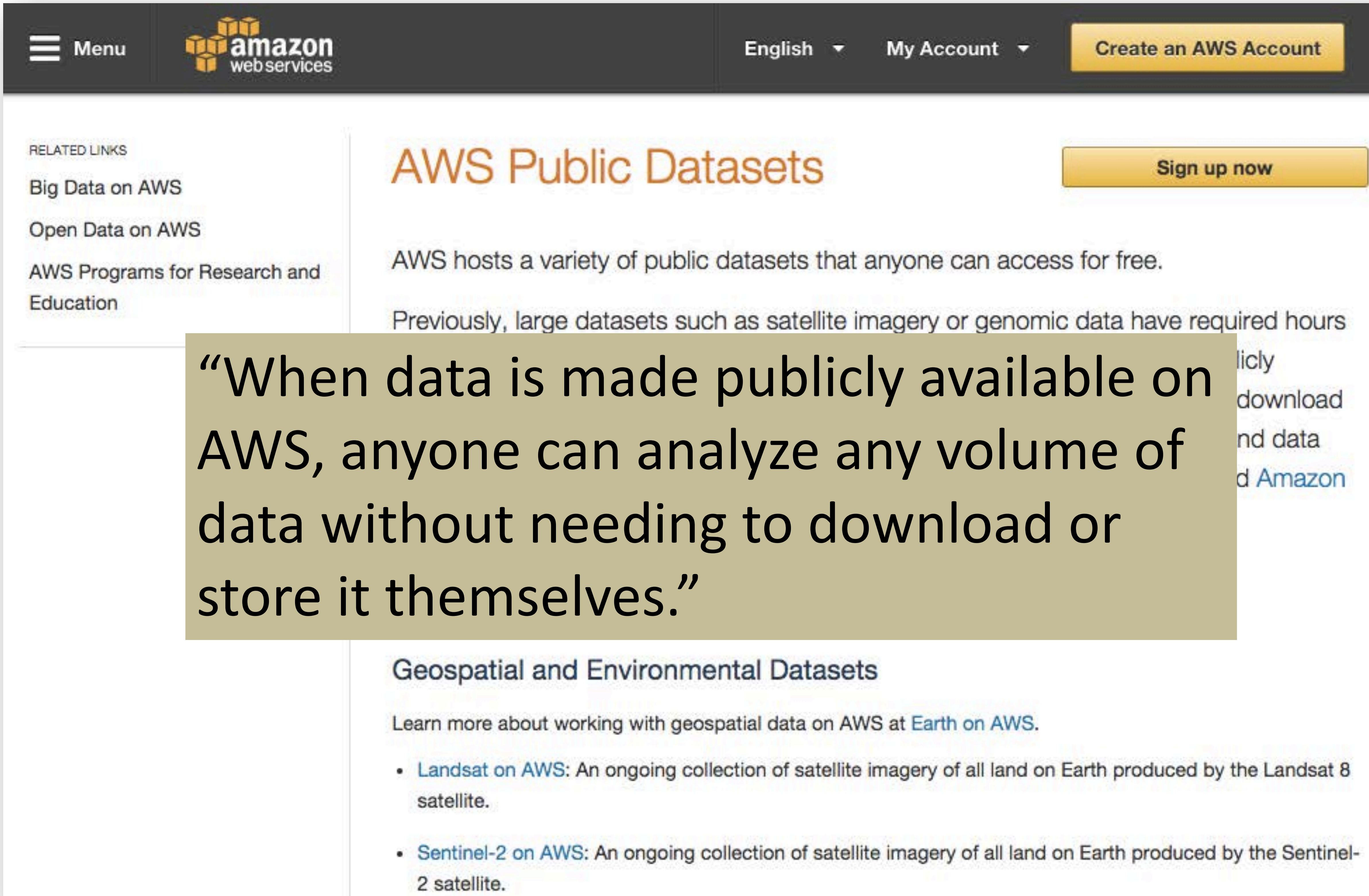
Research challenges

- Marketplace mechanisms
- Hosting Datasets
- Multi-provider cloudlet
- Software defined storage
- HPC on the Cloud
- Secure Hardware Multiplexing

Research challenges

- Marketplace mechanisms
- Hosting Datasets, Mercè Crosas Harvard
- Multi-provider cloudlet
- Software defined storage
- HPC on the Cloud
- Secure Hardware Multiplexing

AWS Public Datasets



"When data is made publicly available on AWS, anyone can analyze any volume of data without needing to download or store it themselves."

But, AWS public datasets miss key aspects needed in data repositories

- Incentives to share data
- Citation to each version of the data
- Metadata for Discoverability
- Tiered access to non-public data
- Commitment to data archival & preservation

Today's repositories incentivize data sharing by giving credit to data authors through formal citation

Bibliography

Ansolabehere, Stephen; Ban, Pamela; Snyder, James M., Jr., 2017, "State Legislative Historical Elections", doi:10.7910/DVN/LEMNXZ, Harvard Dataverse, V1, UNF:6:8UQYfDIsmIl/tgD+Hrv/8Q==

King, Gary; Pan, Jennifer; Roberts, Molley, 2013, "Replication data for: How Censorship in China Allows Government Criticism but Silences Collective Expression", doi:10.7910/DVN1/22691, Harvard Dataverse, V4

Stephen Ansolabehere; Jonathan Rodden, 2011, "Colorado Data Files for State Legislative Elections", hdl:1902.1/15385, Harvard Dataverse, V2, UNF:5:jNUA7tB3bFeMcC2oGBvdHw==

Sweeney L, **Crosas M**, Bar-Sinai M. 2015, "Sharing Sensitive Data with Confidence: the DataTags System" Journal of Technology Science

A, Whitney J. 2015, "Open Journal Systems and Upgrade Data Publication for Reusable Research"

RR, Duerr R, Haak LL, Haendel M, Herman I, Hodson S, rger A, Proell S, Rauber A, Sacchi S, Smith A, Taylor M, e accessibility of cited data in scholarly publications" PeerJ Computer Science, 1:e1 https://dx.doi.org/10.7717/peerj-cs.1

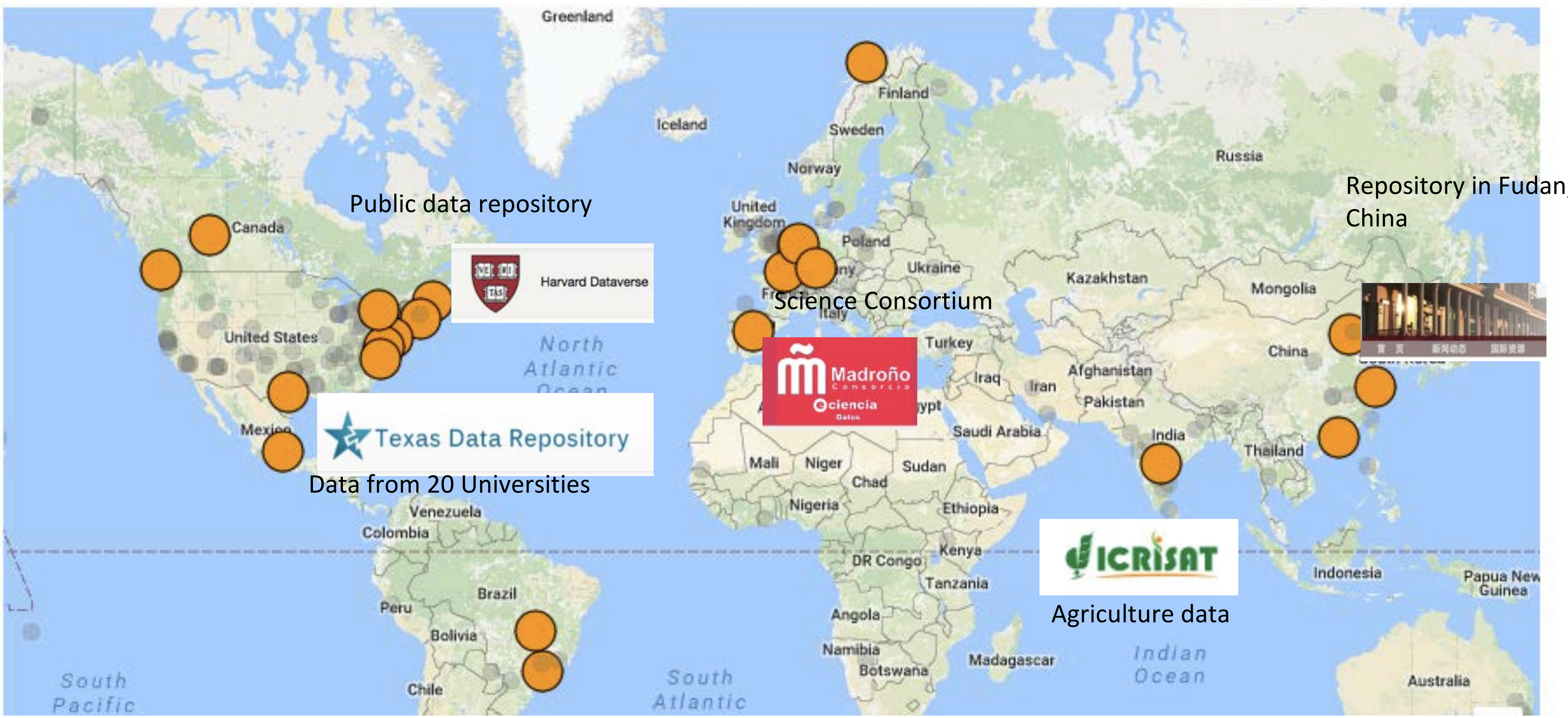
Goodman, A., Pepe, A., Blocker, A.W., Borgman, C.L., Cranmer,K., **Crosas, M.,** Di Stefano, R., Gil, Y., Groth, P., Hedstrom, M., Hogg, D.W., Kashyap,V., Mahabal, A., Siemiginowska, A., Slavkovic, A., 2014. 10 Simple Rules for the Care and Feeding of Scientific Data, PLoS Comput Biol, doi:10.1371/journal.pcbi.1003542

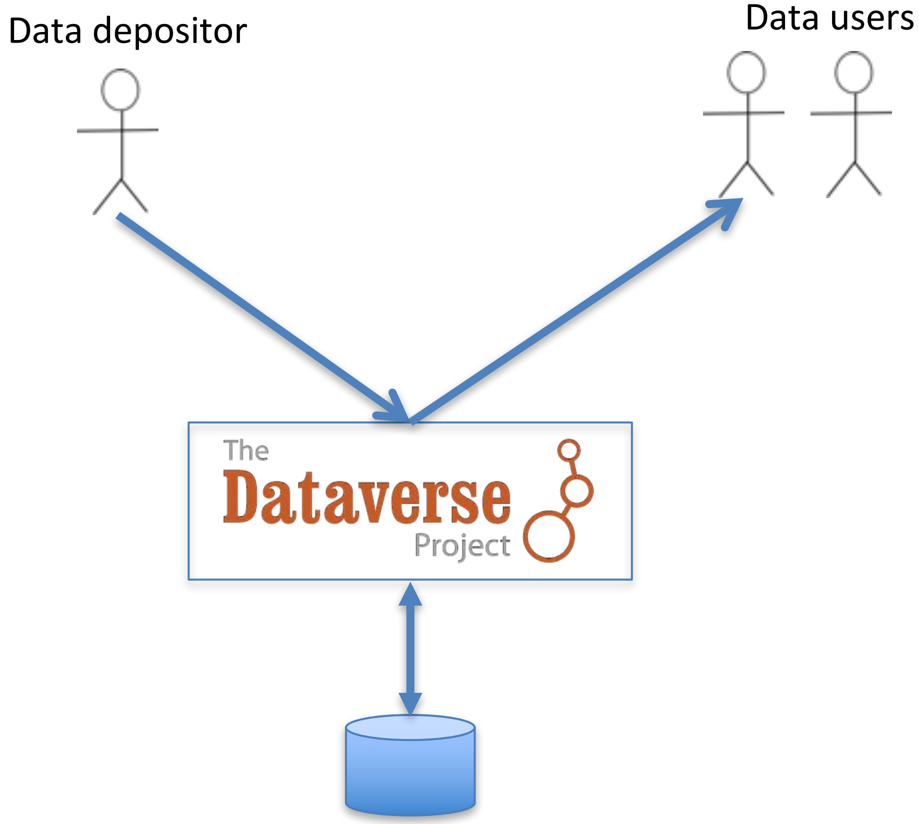
Pepe, A., Goodman, A., Muench, A., **Crosas, M.,** Erdmann, C., 2014. How Do Astronomers Share Data? Reliability and Persistence of Datasets Linked in AAS Publications and a Qualitative Study of Data Practices among US Astronomers. PLoS ONE, DOI: 10.1371/journal.pone.0104798

Persistent citations to datasets published in data repositories

The Dataverse open-source platform enables building any type of data repository

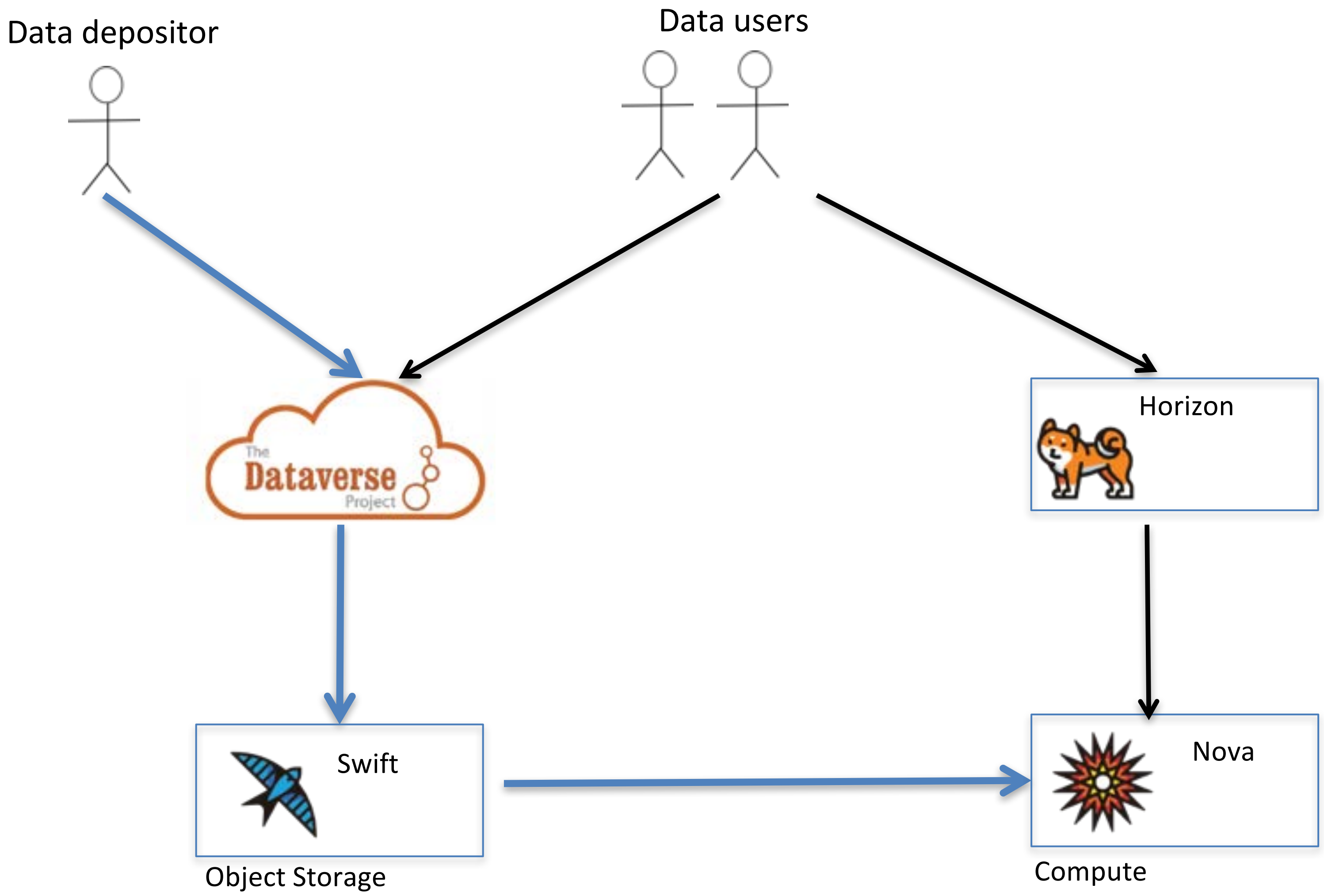
22 Installations　2,133 Dataverses　48,690 Datasets　2,400,322 Downloads

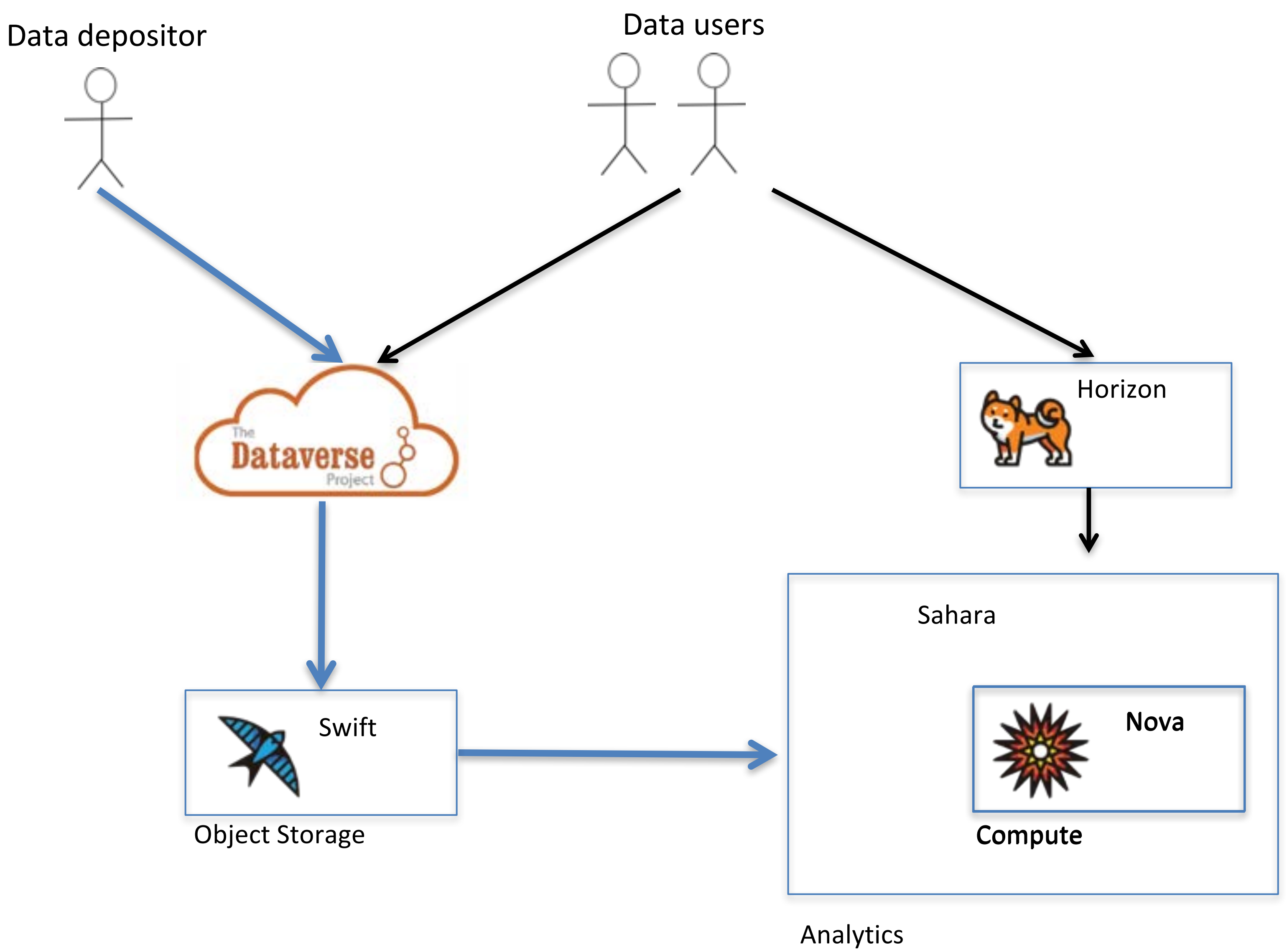


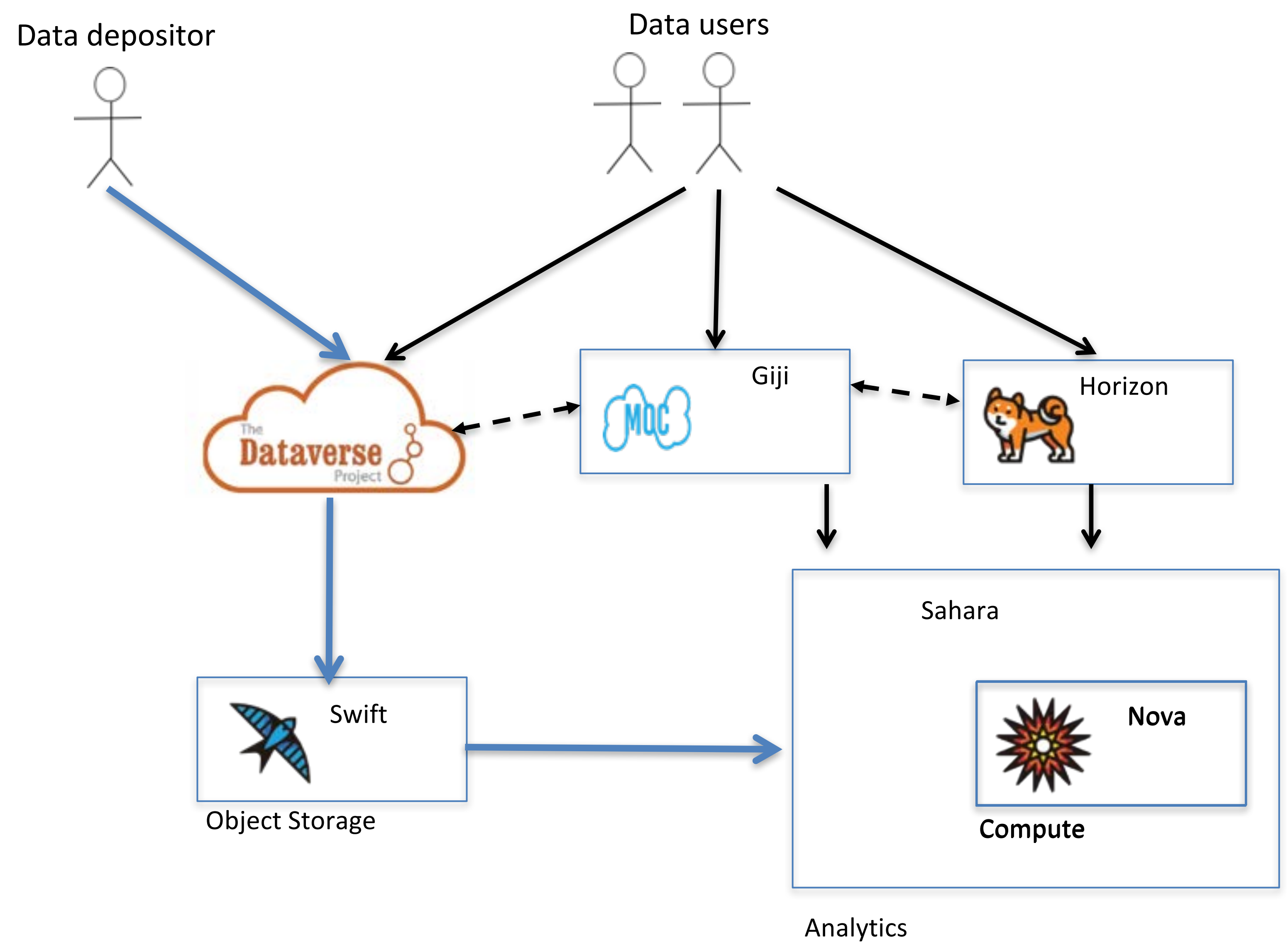


Problems:

- Large datasets
- Lack computational infrastructure


Data depositor
Data users
The Dataverse Project
Horizon
Swift
Nova
Object Storage
Compute


Data depositor
Data users
Dataverse
Project
Horizon
Sahara
Nova
Compute
Swift
Object Storage
Analytics


Data depositor
Data users
Giji
Horizon
Dataverse
Project
Swift
Object Storage
Sahara
Nova
Compute
Analytics

Research challenges

- Marketplace mechanisms
- Hosting Datasets
- Multi-provider cloudlet
- Software defined storage
- HPC on the Cloud
- Secure Hardware Multiplexing

Research challenges

- Marketplace mechanisms
- Hosting Datasets
- Multi-provider cloudlet
- Software defined storage, Peter Desnoyers NU
- HPC on the Cloud
- Secure Hardware Multiplexing

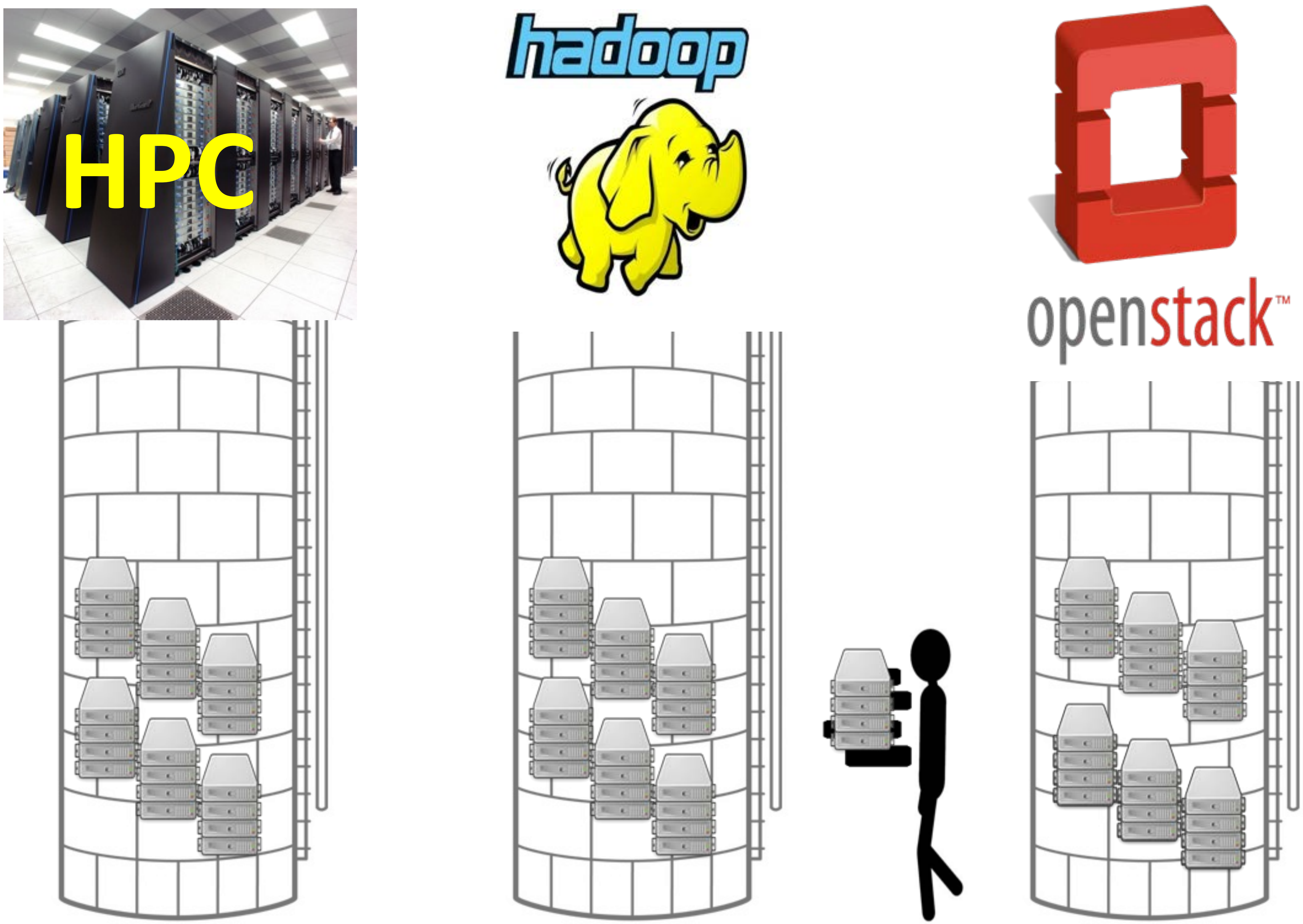
Research challenges

- Marketplace mechanisms
- Hosting Datasets
- Multi-provider cloudlet
- Software defined storage
- HPC on the Cloud: Chris Hill MIT
- Secure Hardware Multiplexing

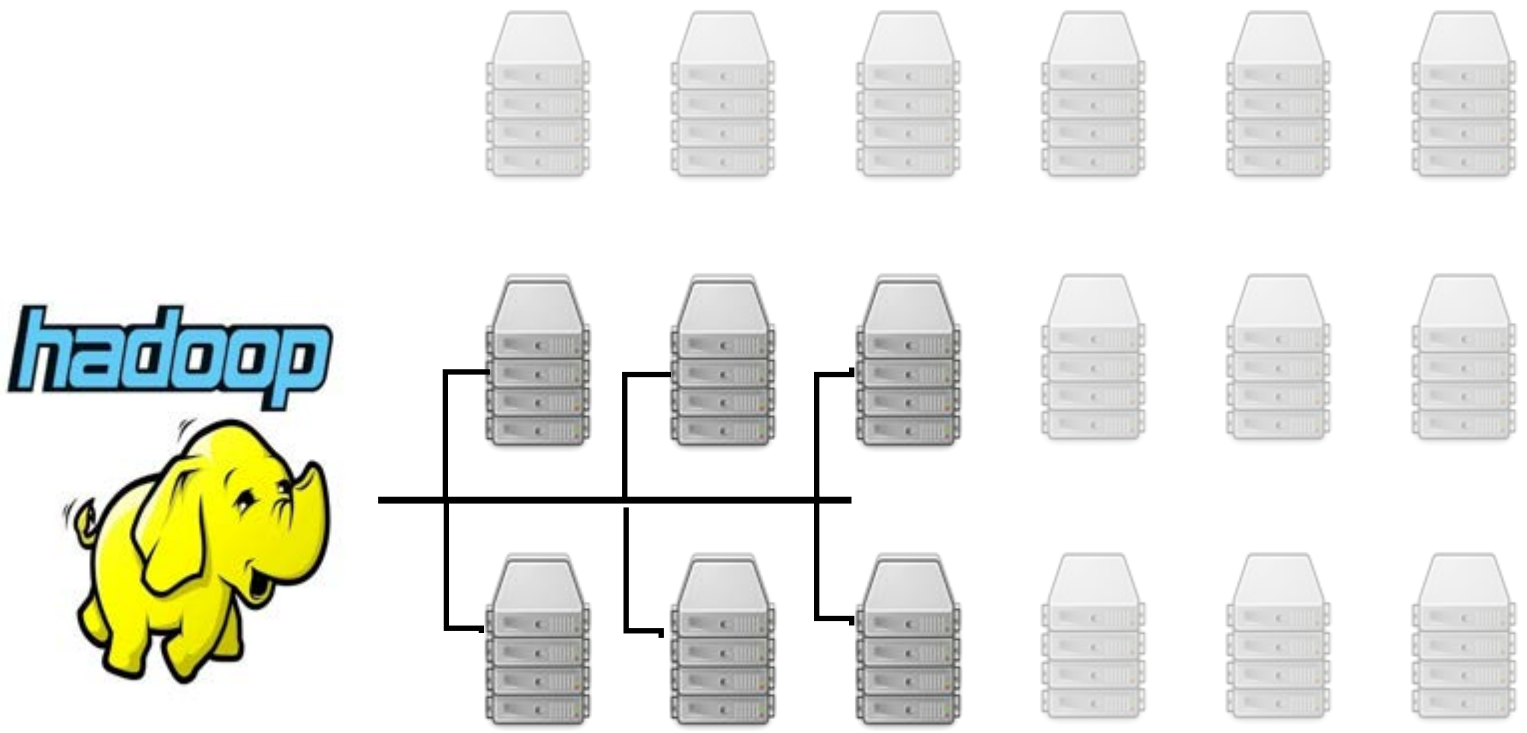
Research challenges

- Marketplace mechanisms
- Hosting Datasets
- Multi-provider cloudlet
- Software defined storage
- HPC on the Cloud
- Secure Hardware Multiplexing: Peter Desnoyers NU, Gene Cooperman NU, Nabil Schear MIT LL, Larry Rudolph & Trammell Hudson Two Sigma, Jason Hennessey BU, …

Datacenter has isolated silos

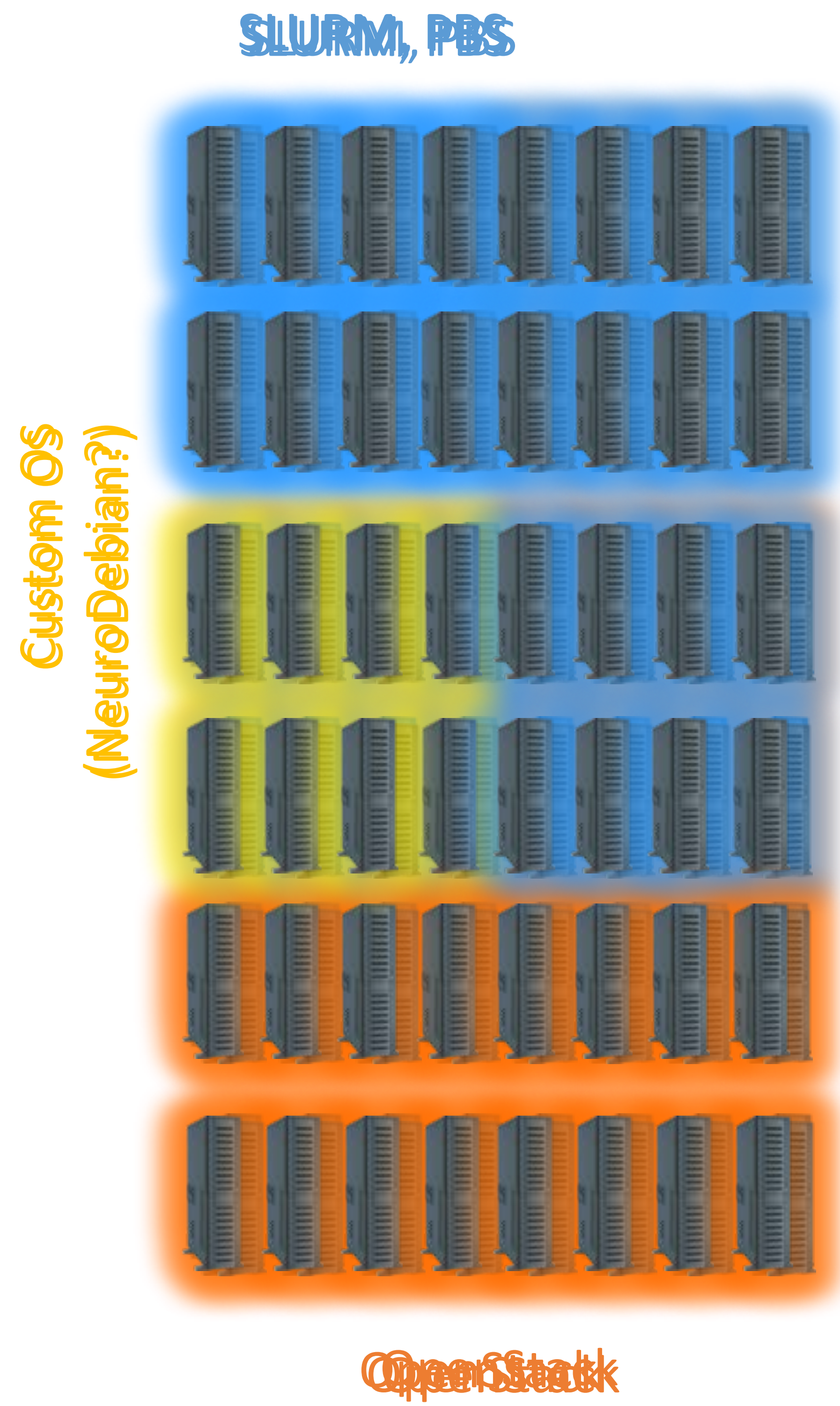


Hardware isolation layer



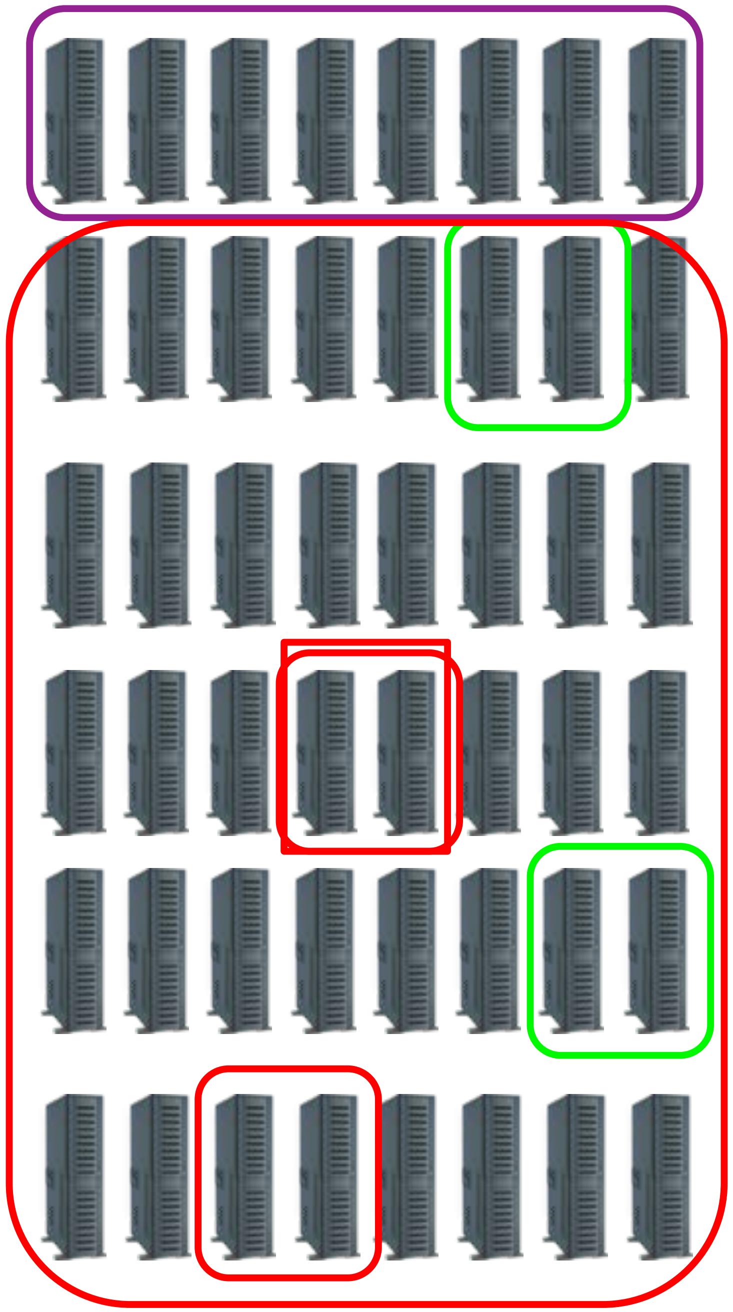
Connect nodes and networks

Hardware Isolation Layer (HIL):
CONVERGING HPC, BIG DATA & CLOUD



What about security?

Secure Cloud Project

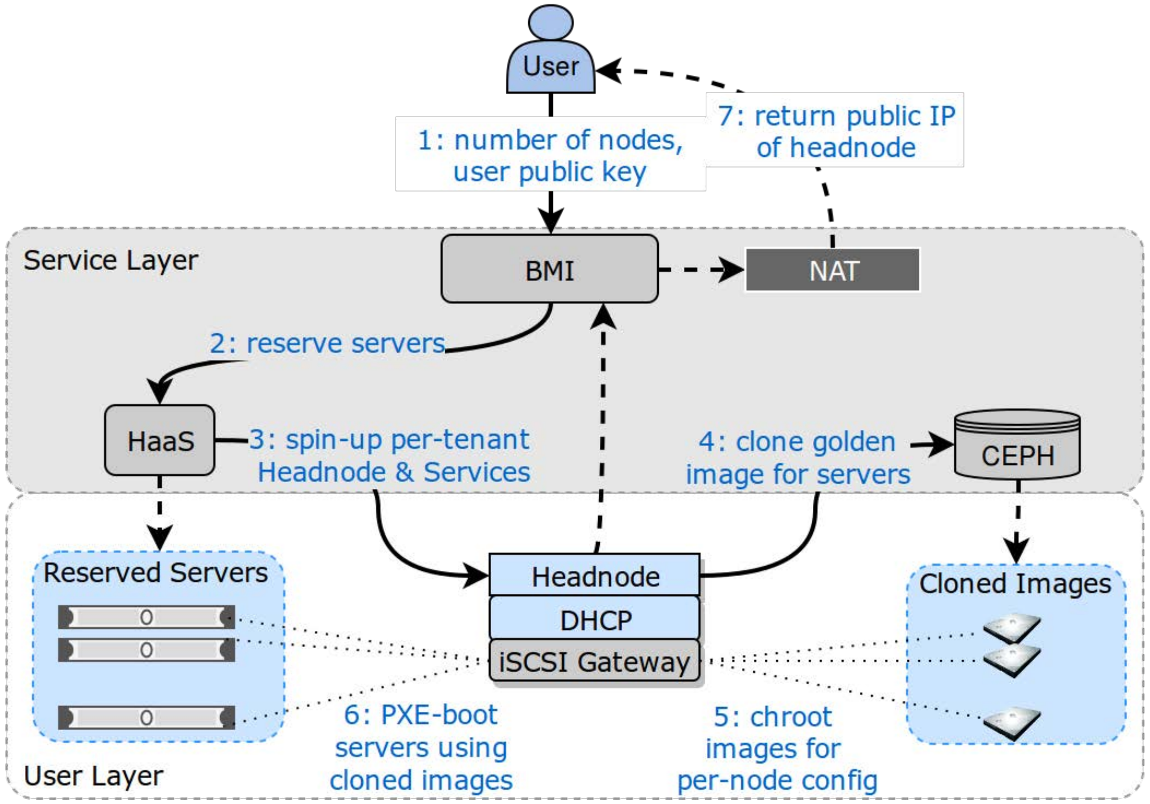


- Shared project with Two Sigma, MIT LL, USAF, Lenovo, Intel
- Integrating attestation infrastructure & secure FW

How fast can we do this?

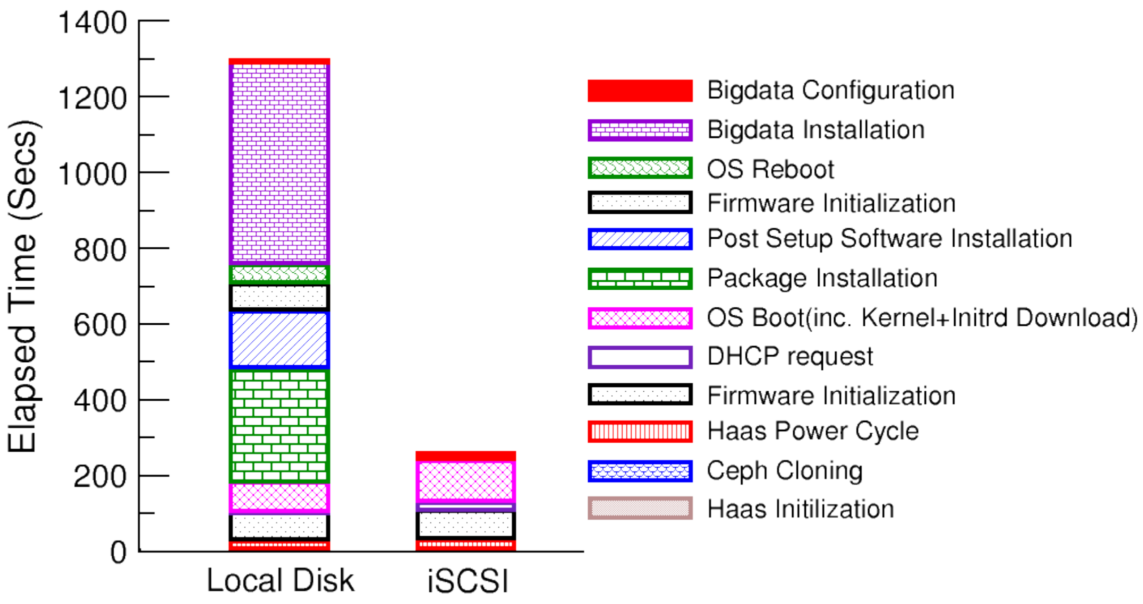
Bare Metal Imaging Service

iSCSI-based



Rapid Bare-Metal Provisioning and Image Management, Ravisantosh Gudimetla and Apoorve Mohan

Able to provision + boot in < 5 min



Turk, A., Gudimetla, R. S., Kaynar, E. U., Hennessey, J., Tikale, S., Desnoyers, P., & Krieger, O. (2016). An Experiment on Bare-Metal BigData Provisioning. In 8th USENIX Workshop on Hot Topics in Cloud Computing (HotCloud 16).

Research challenges

- Can we expose rich information about services while not violating customer privacy
- How can we correlate between the information between the different layers?
- How can we identify source of failures?
- How can we create a Networking Marketplace?

Research challenges

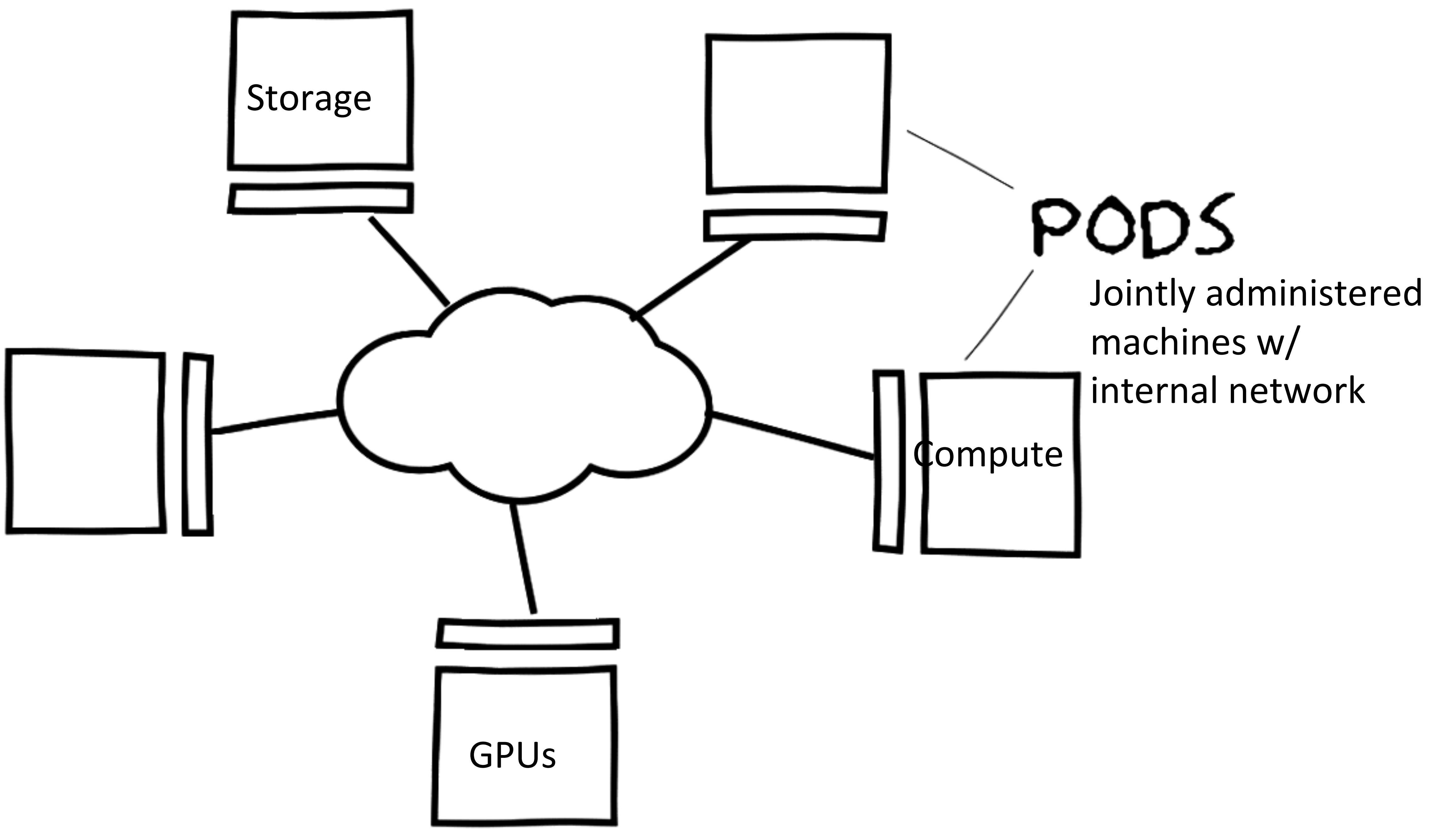
- Can we expose rich information about services while not violating customer privacy
- How can we correlate between the information between the different layers?
- How can we identify source of failures?
- Networking Marketplace: Rodrigo Fonseca Brown

Common view:

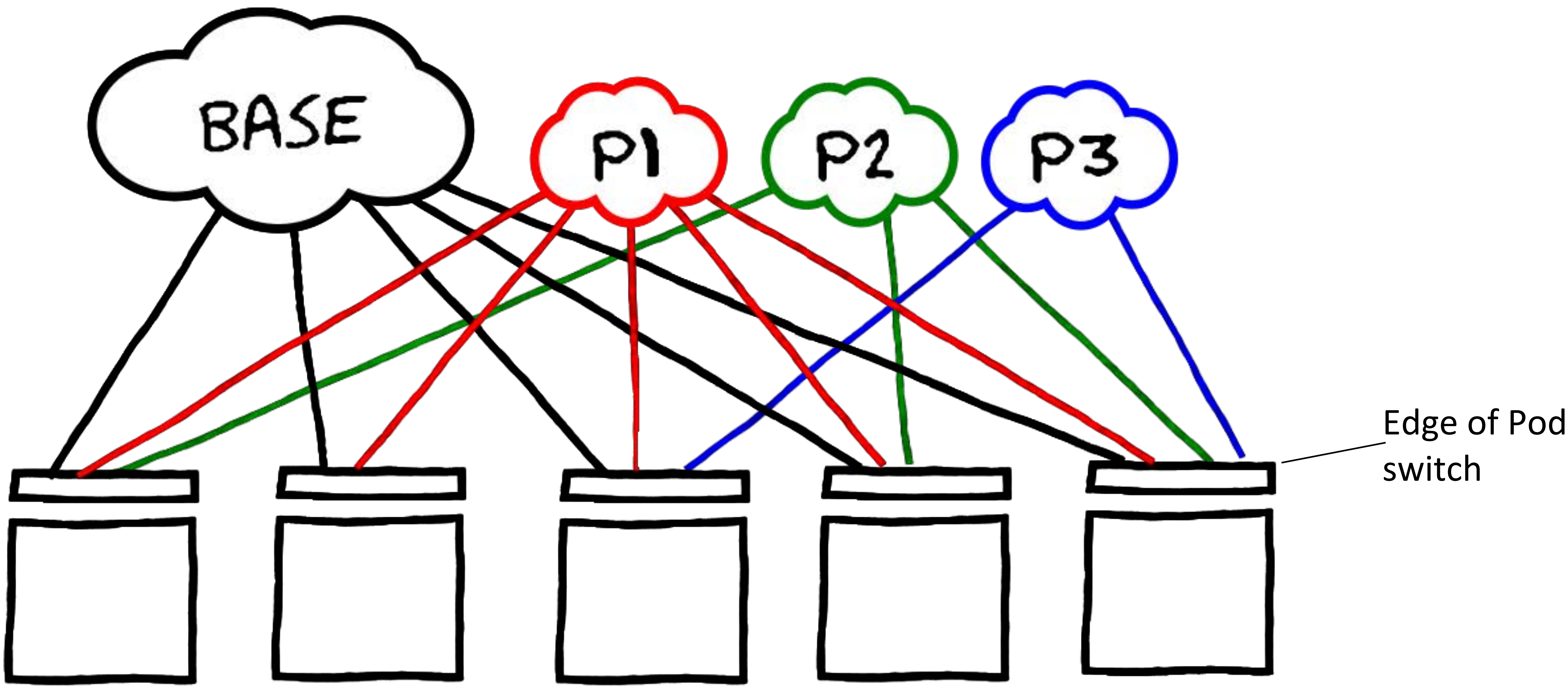
Networking is like air conditioning, or power

Part of the infrastructure, provided by the datacenter

Basic Architecture



Multi-Provider Inter-Pod Network



Research enabled

- New hardware infrastructure; e.g. FPGAs, new processors
- Caching storage from Data Lakes
- Cloud security and composability of security properties; e.g., MACS project
- Smart cities
- Analysis of cloud internal information (logs, metrics) for security, for optimization…
- Highly elastic environments; e.g., 1000 servers for a minute:

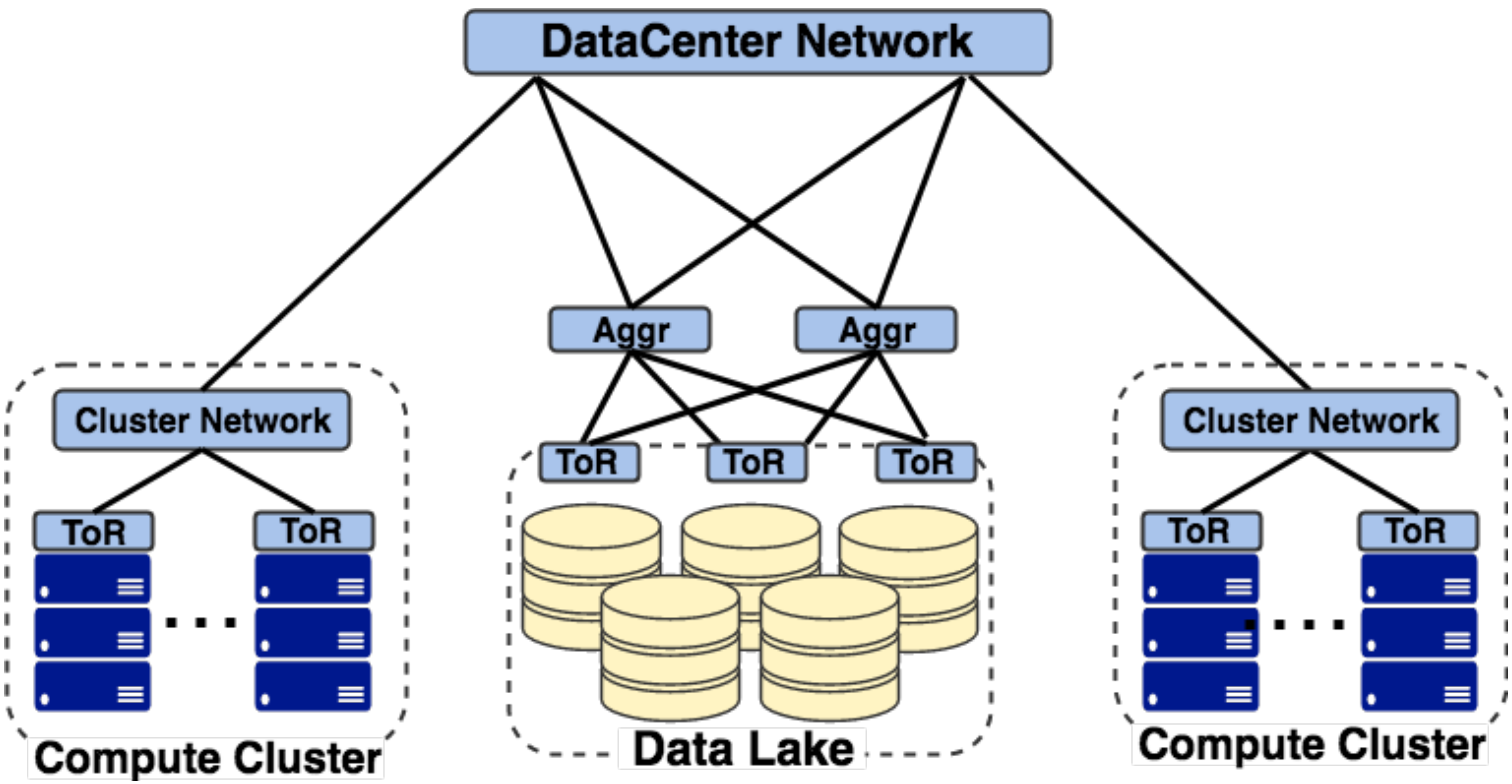
Research enabled

- New hardware infrastructure; e.g. FPGAs, new processors: Martin Herbordt (BU)
- Caching storage from Data Lakes
- Cloud security and composability of security properties; e.g., MACS project
- Smart cities
- Analysis of cloud internal information (logs, metrics) for security, for optimization…
- Highly elastic environments; e.g., 1000 servers for a minute:

Research enabled

- New hardware infrastructure; e.g. FPGAs, new processors
- Caching storage from Data Lakes: Desnoyers NU, Krieger BU
- Cloud security and composability of security properties; e.g., MACS project
- Smart cities
- Analysis of cloud internal information (logs, metrics) for security, for optimization…
- Highly elastic environments; e.g., 1000 servers for a minute:

Data Lake in a typical DC

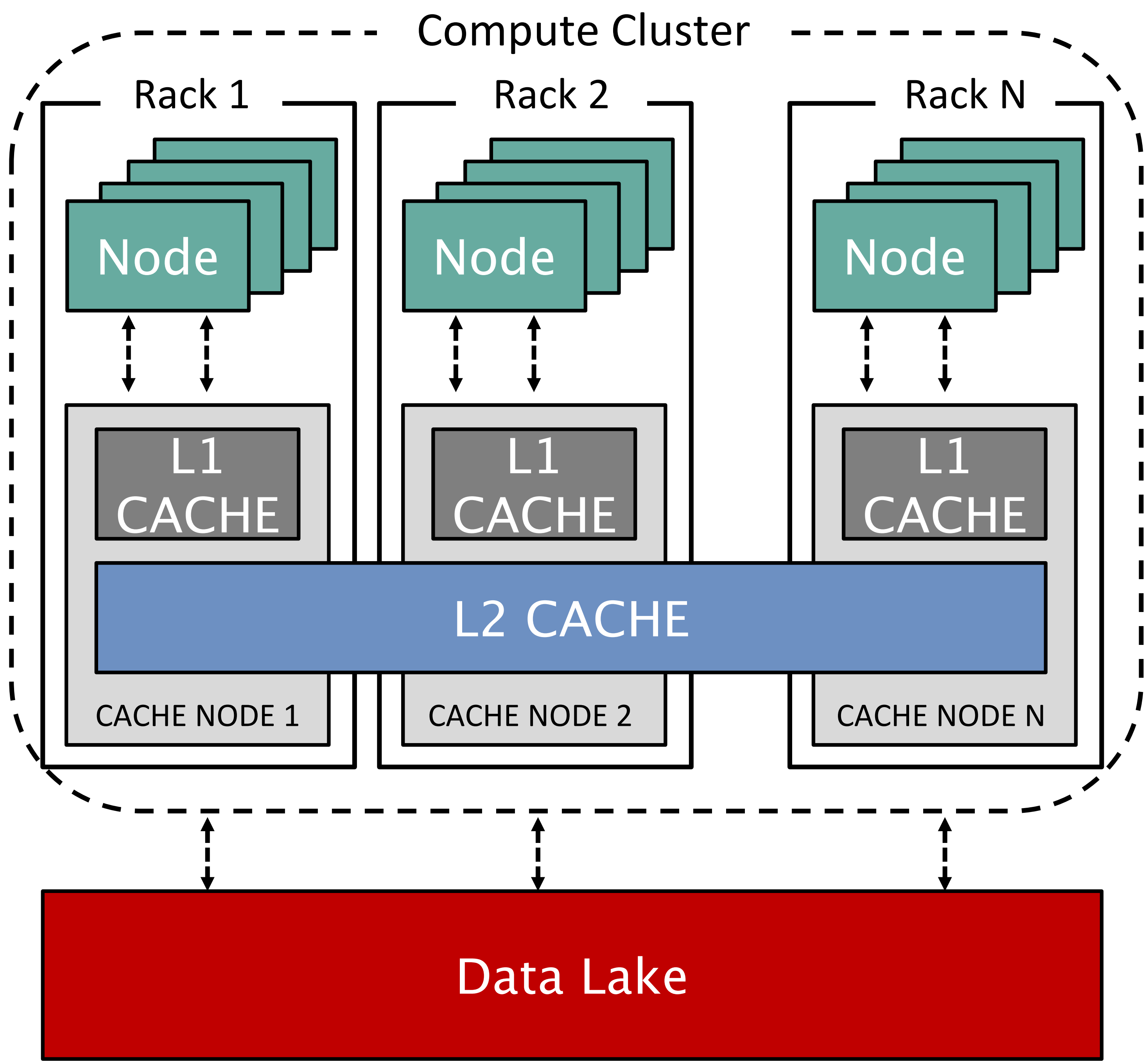


North Eastern Storage Exchange (NESE):
20+PB Harvard, NEU, MIT, BU, UMass

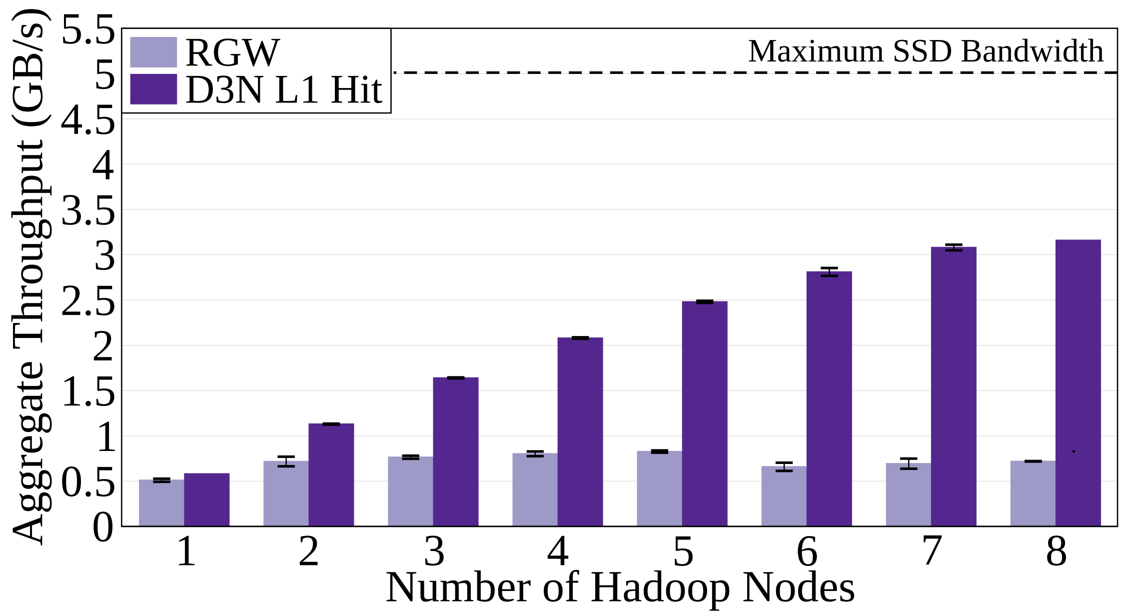
Datacenter scale Data Delivery Network (D3N)

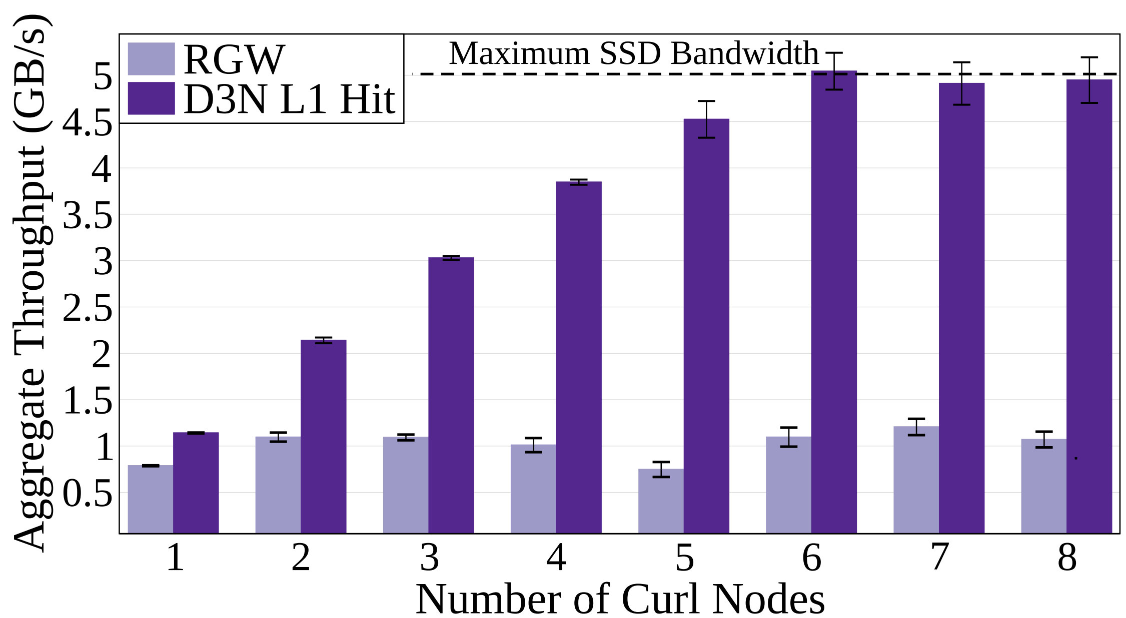
Simple deployment:

- Cache Node per rack
- **L1 :** Rack Local
  - reduce inter rack traffic
- **L2 :** Cluster Local
  - reduce clusters and back-end storage traffic
- Implemented by modifying **CEPH Rados Gateway**



D3N Results





- Exceeds maximum bandwidth Hadoop
- Demonstrates makes sense to share expensive SSDs – faster than local disk
- With extreme benchmark can saturate SSD & 40 Gb NIC
- Will be of enormous value with NESE data lake

Research enabled

- New hardware infrastructure; e.g. FPGAs, new processors
- Caching storage from Data Lakes: Desnoyers NU, Krieger BU
- Cloud security and composability of security properties; e.g., MACS project
- Smart cities
- Analysis of cloud internal information (logs, metrics) for security, for optimization…
- Highly elastic environments; e.g., 1000 servers for a minute:

Modular Approach to Cloud Security

In security, the sum of the parts is often a *hole*.

– Dave Safford, circa 2000

Our goal is to build security systems so that the sum of the parts is a *holistic security guarantee*.

– Ran Canetti, 2016

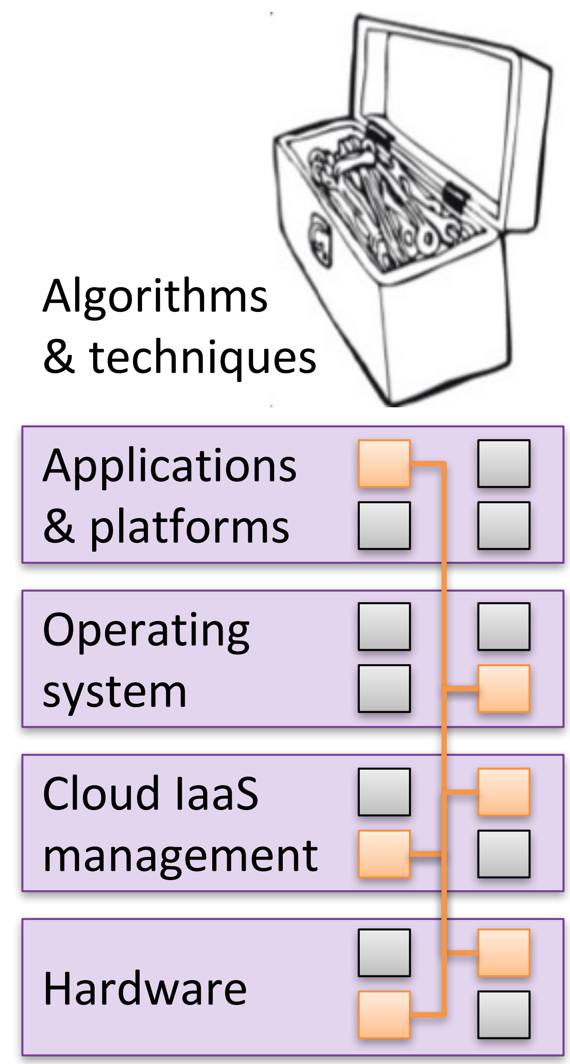
Synergy between MACS and MOC

Types of connections

- *People*: researchers can contribute toward both projects
  - Size of MACS: 13 faculty, 11 postdocs, 25+ graduate students
- *Tech transition*: deploy MACS tech in MOC marketplace
- *Problem creation*: MOC's problems feed MACS research
- *Funding*: joint cloud research has multiplier effect

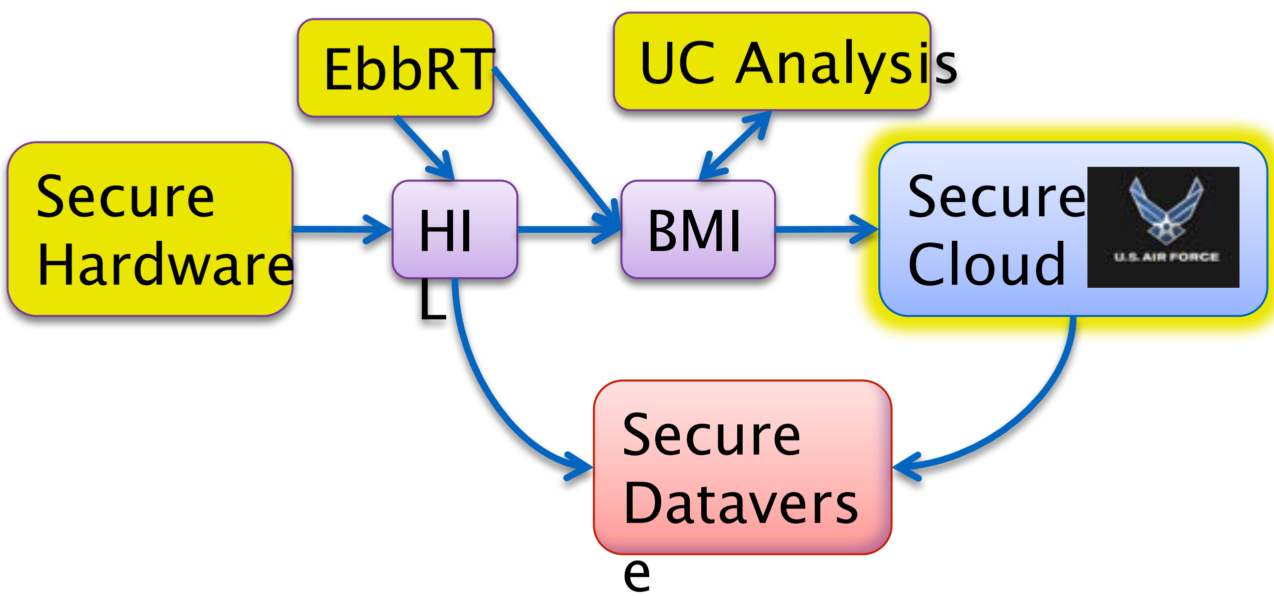
Value that MOC provides to MACS

- *Access*: data, meta-data, scale, problems, and users
- *Unique trust relationships*: federated datacenter

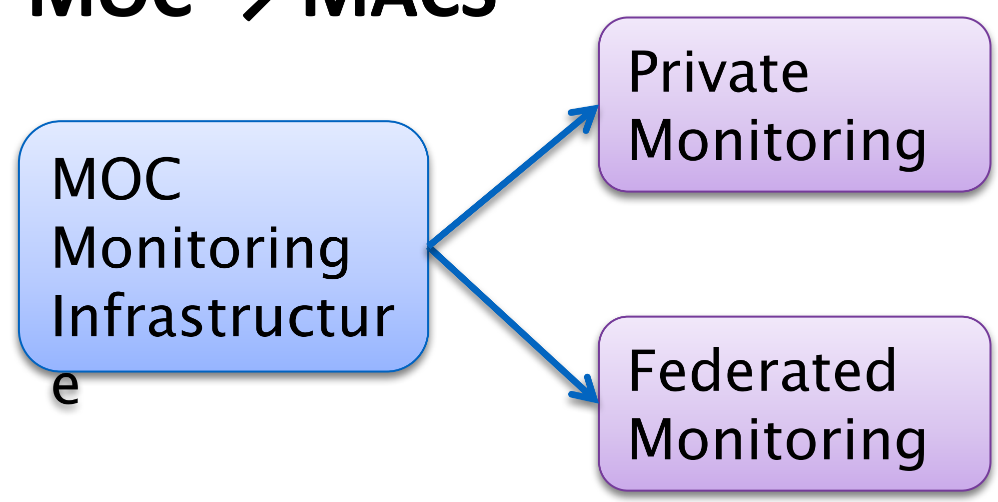


Interplay between MACS and MOC

MACS → MOC



MOC → MACS



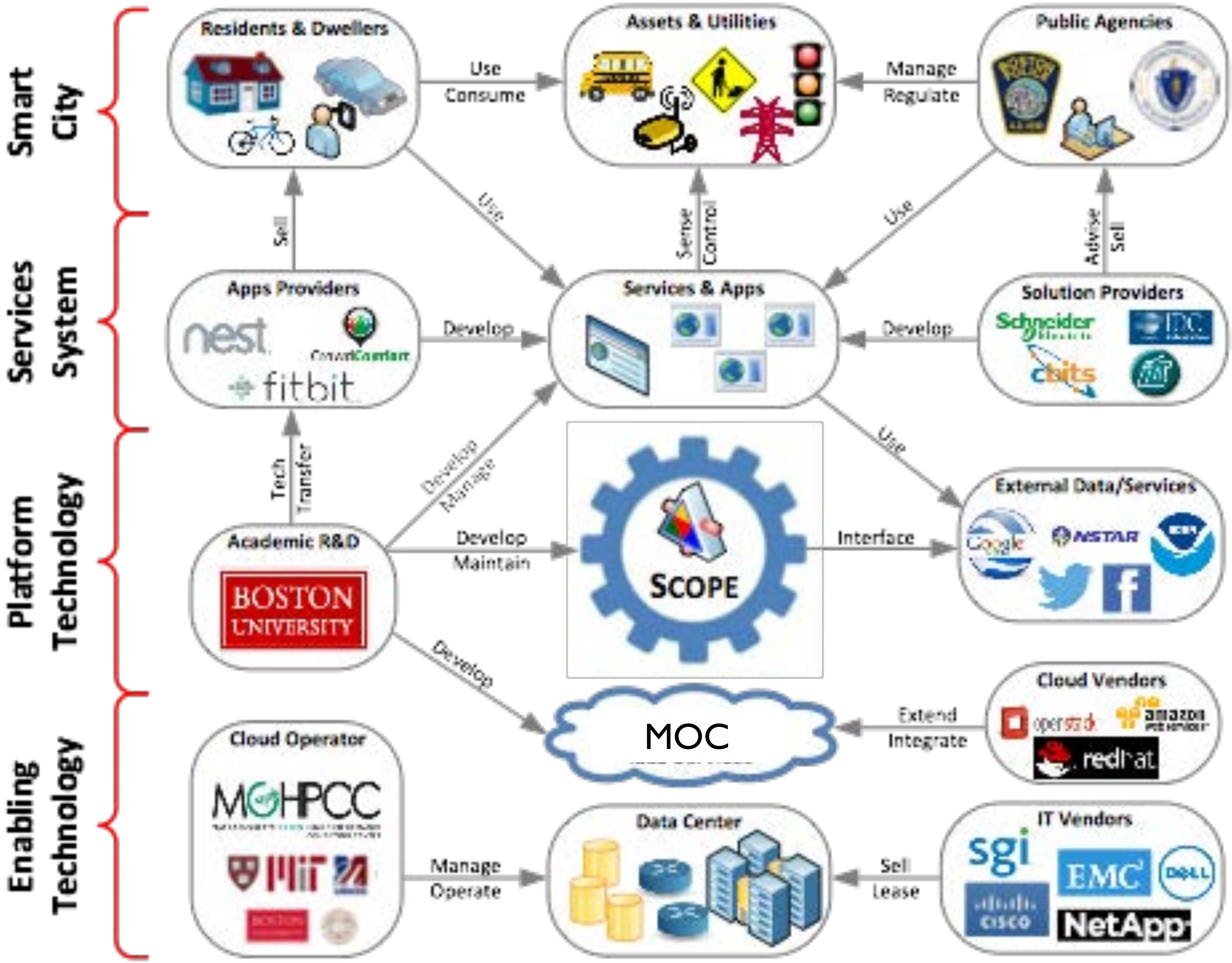
Legend:

- **Yellow** = MACS
- **Blue** = MOC
- **Green** = Joint

Research enabled

- New hardware infrastructure; e.g. FPGAs, new processors
- Caching storage from Data Lakes: Desnoyers NU, Krieger BU
- Cloud security and composability of security properties; e.g., MACS project
- Smart cities: Azer Bestavros BU
- Analysis of cloud internal information (logs, metrics) for security, for optimization…
- Highly elastic environments; e.g., 1000 servers for a minute:

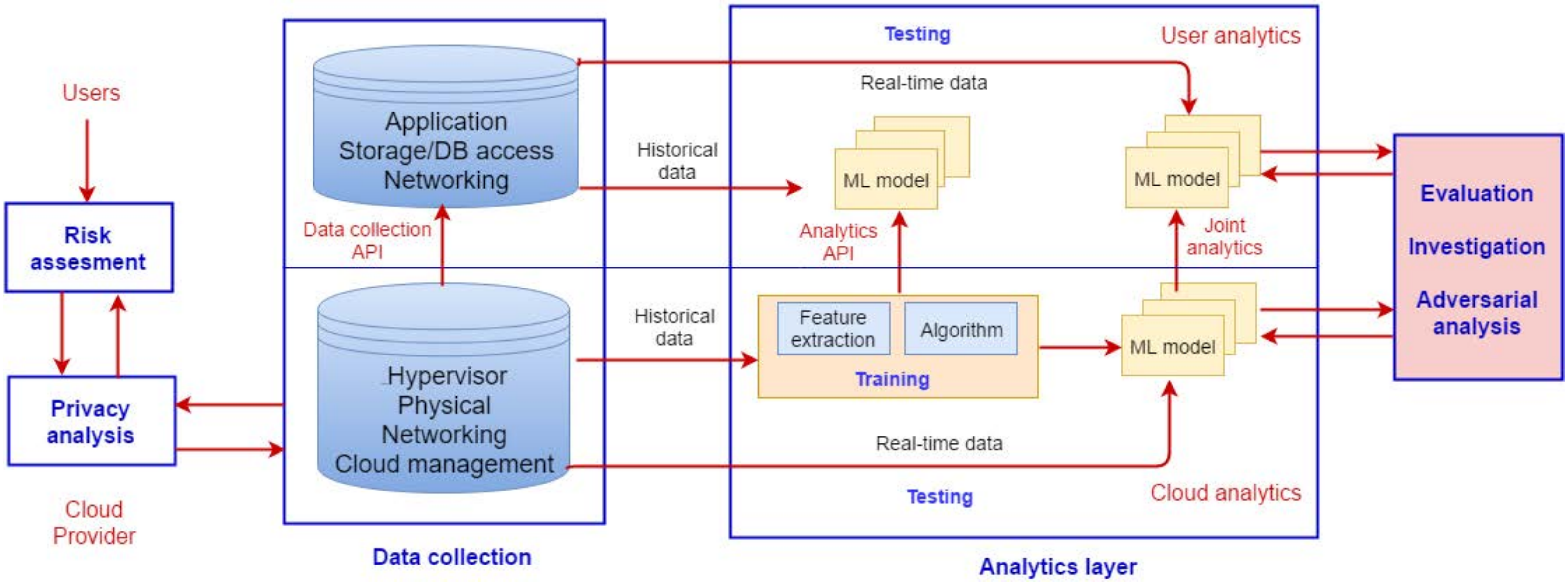
Example: Smart cities



Research enabled

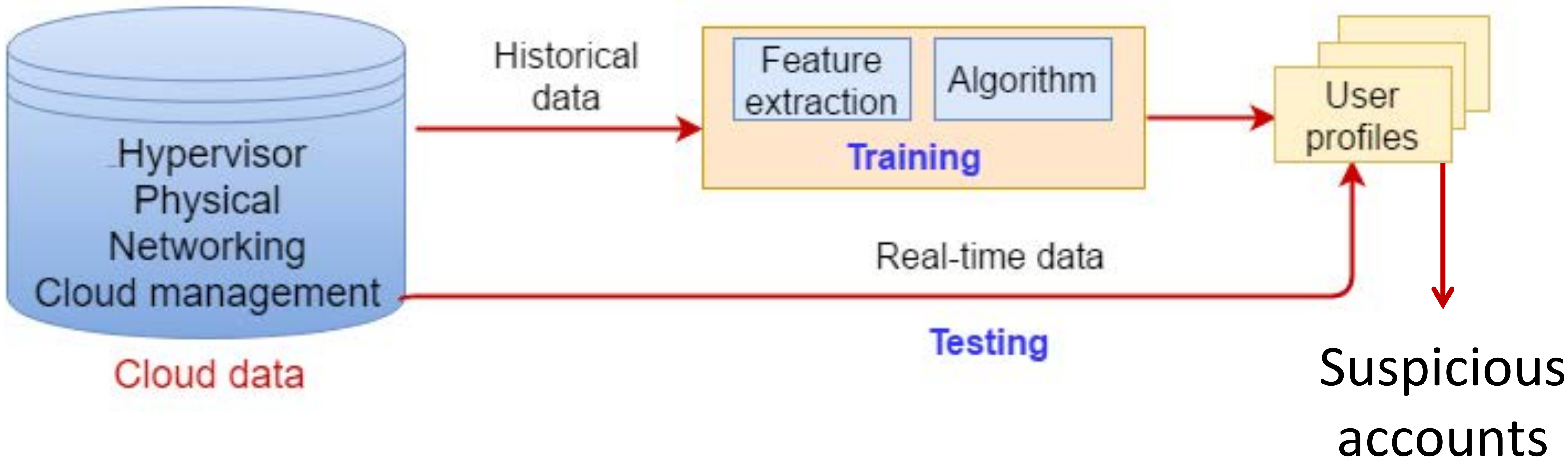
- New hardware infrastructure; e.g. FPGAs, new processors
- Caching storage from Data Lakes: Desnoyers NU, Krieger BU
- Cloud security and composability of security properties; e.g., MACS project
- Smart cities
- Analysis of cloud internal information (logs, metrics) for security, for optimization…: …: Alina Oprea NU
- Highly elastic environments; e.g., 1000 servers for a minute:

Analytics-based defenses



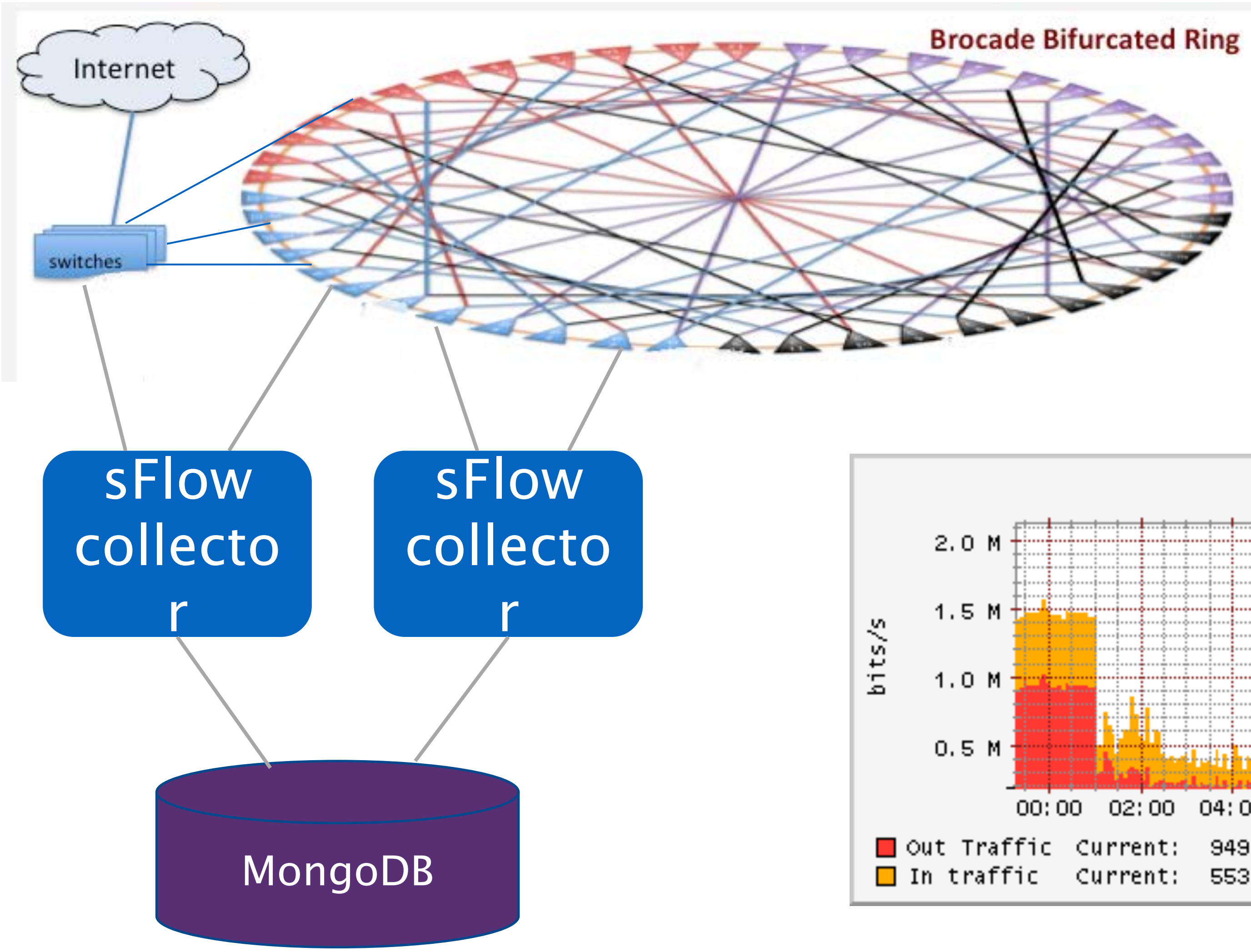
- Goals
  - Correlate data sources from multiple cloud layers
    - Build user, VM and application profiles
  - Machine learning techniques to detect wide range of threats
    - Protection of cloud infrastructure
    - Enable cloud users to protect their resources
  - Provide data collection and analytics APIs to users

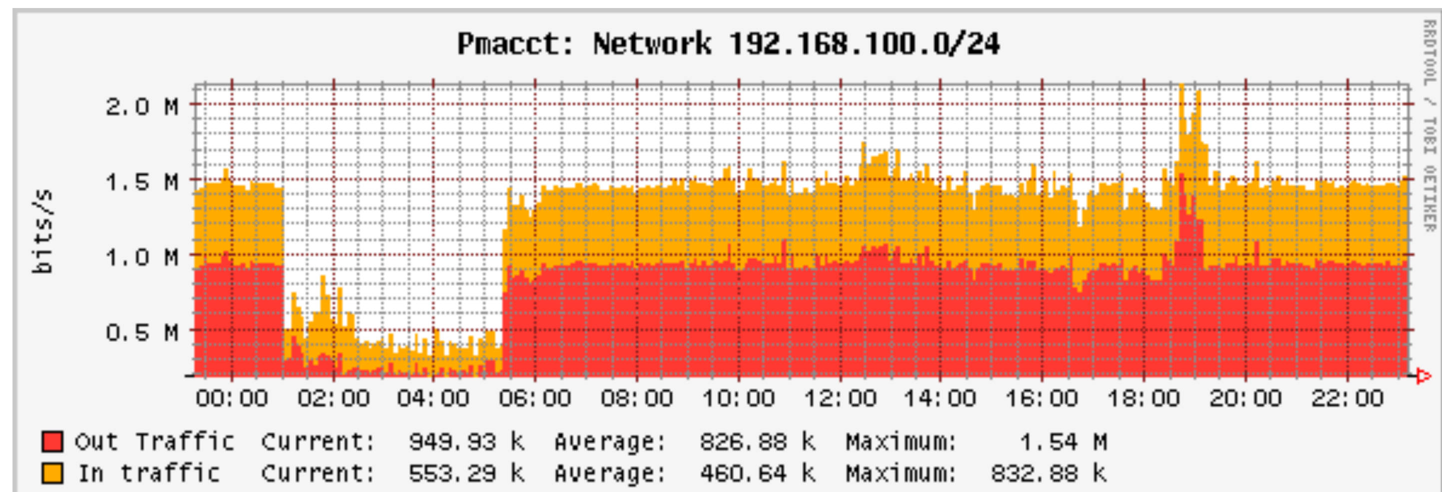
Behavior-based authentication



- Detect credential compromise
  - Developers leak their AWS passwords in GitHub
- Build user profiles based on historical data
  - Login information (IP address, time)
  - VM usage (CPU, memory, disk)
- Anomaly detection
  - Detect unusual activities

Network traffic analysis





Use cases

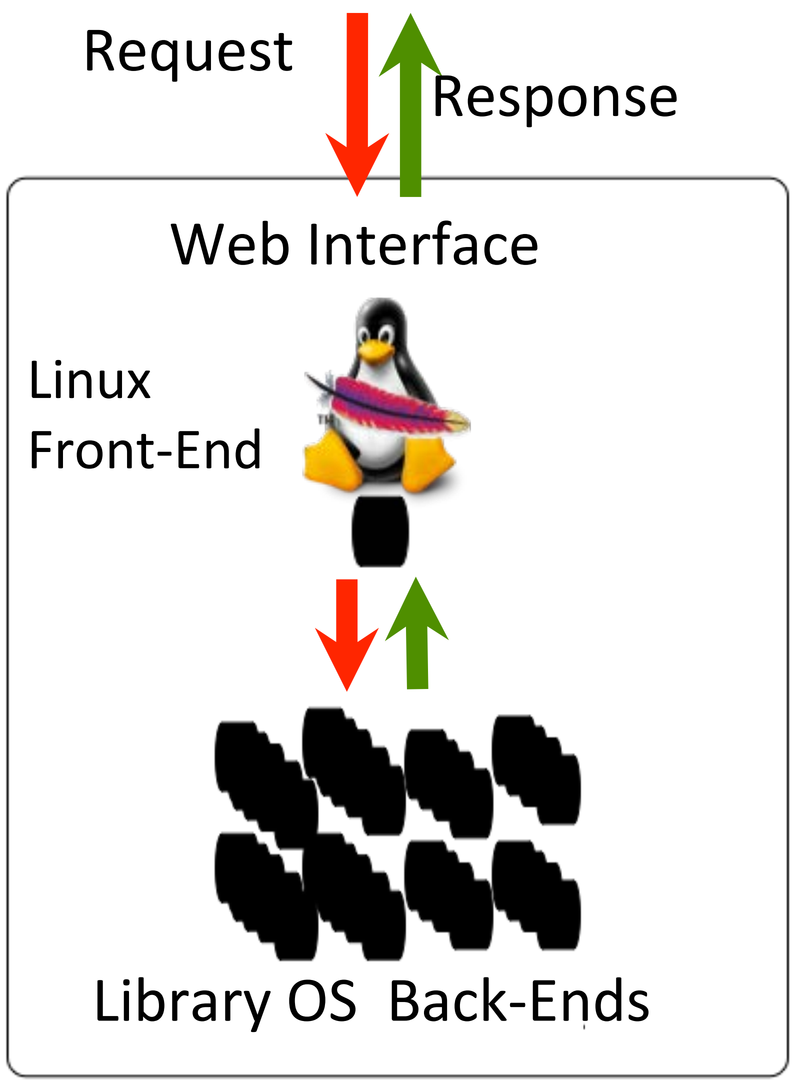
- Detect suspicious communication with external IP addresses
- Detect data exfiltration attempts
- Prevent cloud abuse
  - Malware infection, application exploits , illegal use of cloud

Research enabled

- New hardware infrastructure; e.g. FPGAs, new processors
- Caching storage from Data Lakes: Desnoyers NU, Krieger BU
- Cloud security and composability of security properties; e.g., MACS project
- Smart cities
- Analysis of cloud internal information (logs, metrics) for security, for optimization…: …: Alina Oprea NU
- Highly elastic environments; e.g., 1000 servers for a minute: Jonathan Appavoo BU

Example Supporting Interactive, Bursty HPC Applications: OSDI 2016

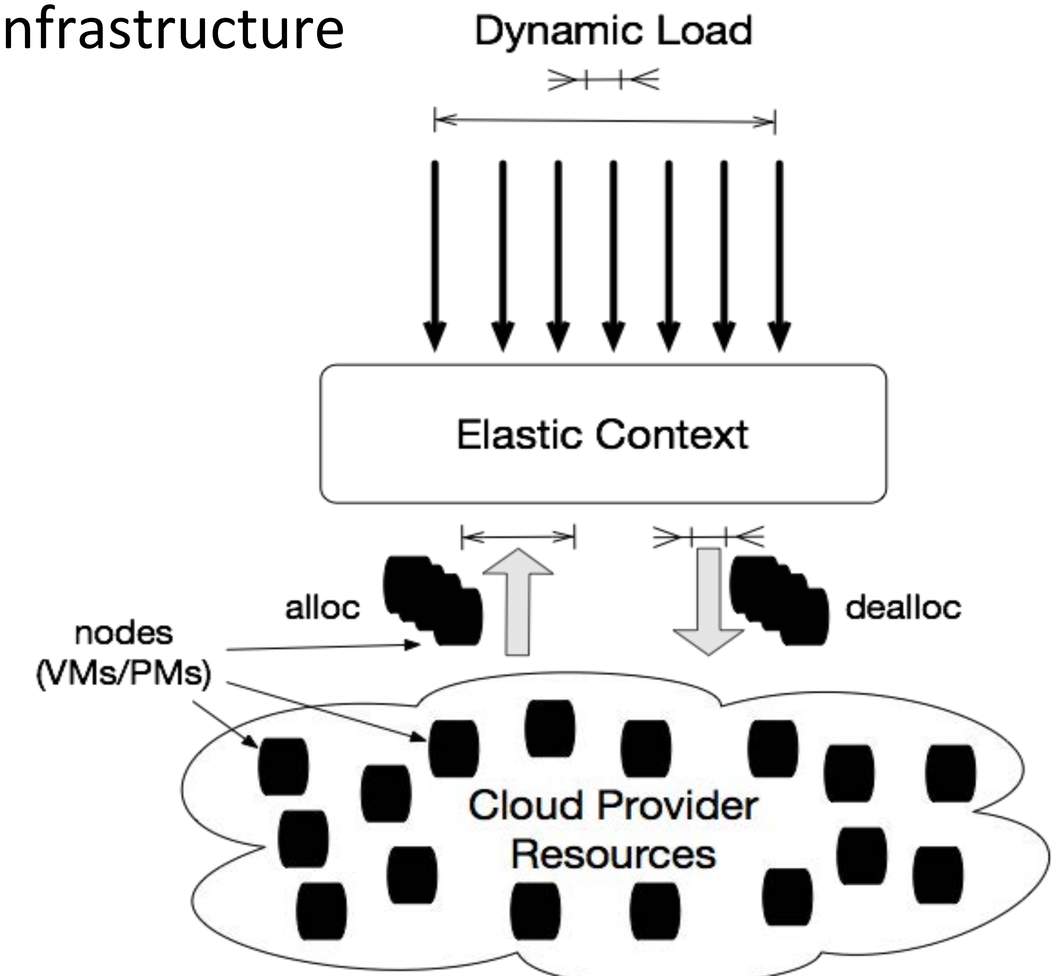
Elastic Software



EbbRT distributed library OS [Appavoo BU]:

- Front-end Linux allocates bare-metal back-end nodes on demand
- Back-end nodes library OS customized to single application needs

Elastic Infrastructure

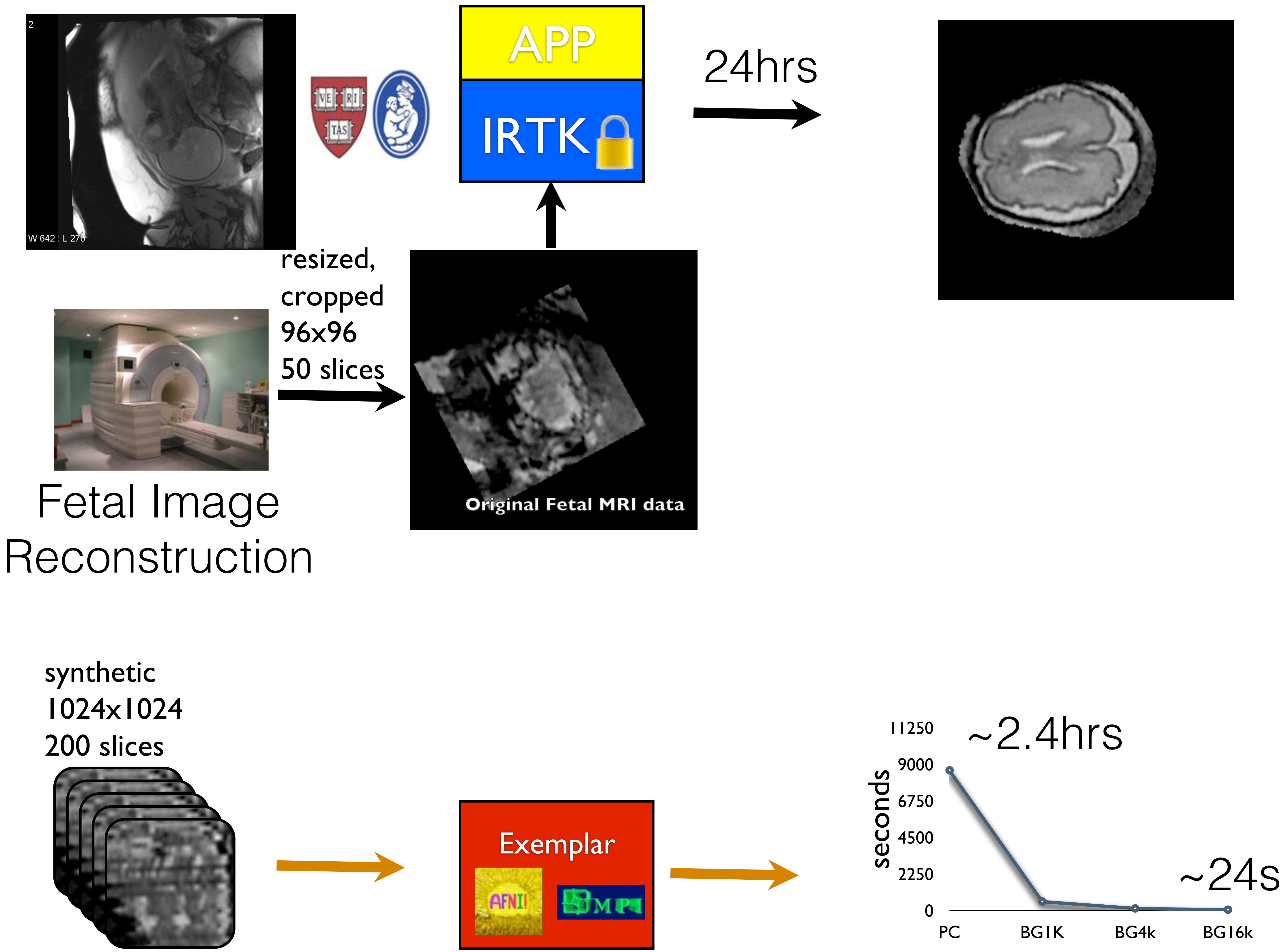


Infrastructure as Elastic Resource Pool

XSP compute service based on Kittyhawk [Appavoo IBM]

- Fast provisioning based on broadcast
- Hardware level based on HaaS
- IaaS level by pre-allocating VMs out of OpenStack

[Appavoo]



W 642 : L 270
APP
IRTK
24hrs
resized,
cropped
96x96
50 slices
Original Fetal MRI data
Fetal Image
Reconstruction
synthetic
1024x1024
200 slices
Exemplar
AFNI!
~2.4hrs
~24s
seconds
11250
9000
6750
4500
2250
0
PC
BG1K
BG4k
BG16k

Red Hat Collaboratory

- Monitoring and Analytics
- OpenShift on the MOC
- Datacenter scale Data Delivery Network (D3N)
- HIL & QUADS
- Accelerator Testbed
- Big Data Analytics and Cloud Dataverse

End-to-end POC: Radiology in the cloud targeting OpenShift with accelerators

Concluding remarks

- MOC a functioning small scale cloud for region today:
  - http://info.massopencloud.org
- Key driver is the OCX Model:
  - Key enablers going on in OpenStack (been a challenge)
  - could become important component of clouds
  - Major research challenge & opportunities
  - Enabling research to co-exists with production:
    - real data, real users, real scale
- Get involved: use it, internships, expose research
- Start replicated the model elsewhere