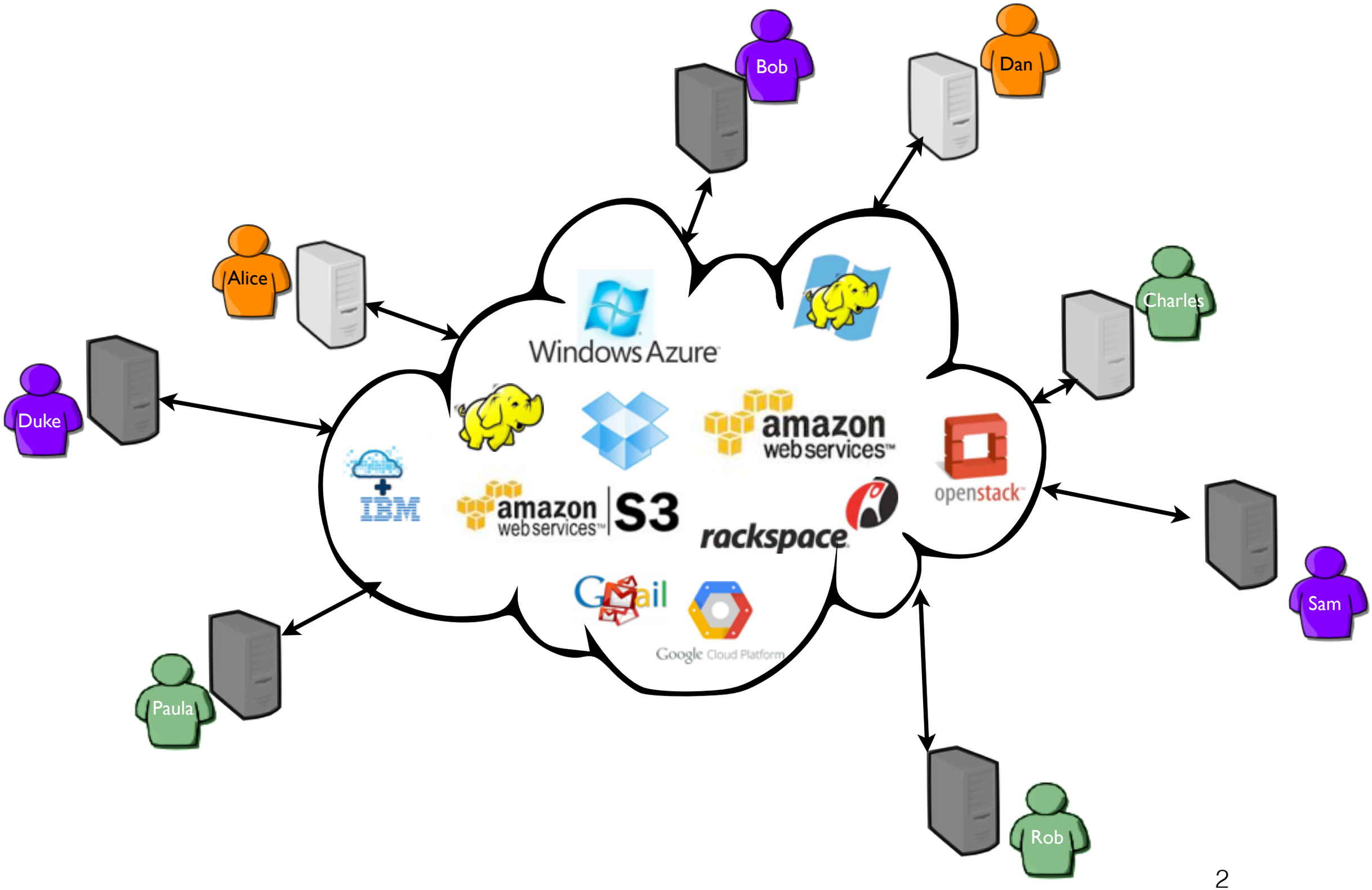
Beyond the Clouds, The Discovery Initiative

Revising OpenStack to operate/use Fog/Edge Computing Infrastructures

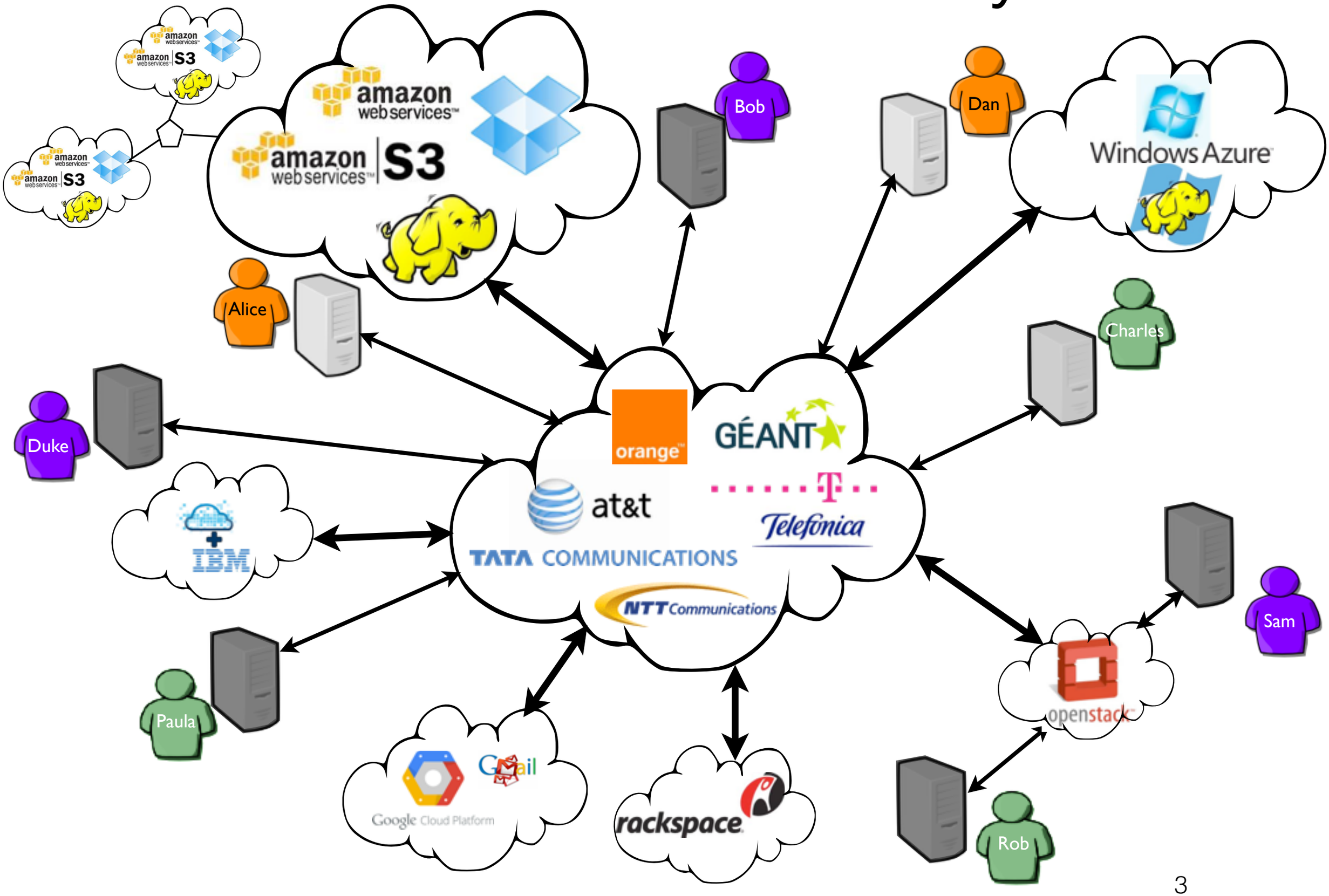
Adrien Lebre

CargoDay - Oct 2016

The Cloud From End-users



The Cloud in Reality



The Trend since 2013: Large off shore DCs

- To cope with the increasing CC demand while handling energy concerns but…

credits: datacentertalk.com - Microsoft DC, Quincy, WA state

The Trend since 2013: Large off shore DCs

- To cope with the increasing CC demand while handling energy concerns but…

credits: google map - Quincy

The Trend since 2013: Large off shore DCs

- To cope with the increasing CC demand while handling energy concerns but…

credits: coloandcloud.com

The Trend since 2013: Large off shore DCs

Jurisdiction concerns

Reliability

CC distance (network overheads)

2012 - 2013
Major brakes for the adoption of the CC model

Jurisdiction concerns
Reliability

CC distance (network overheads)

Big Data - Internet of Things/Everything / Industrial Internet

2016 - xxxx

Localization is a key element to deliver *efficient* as well as *sustainable* **Utility Computing** solutions

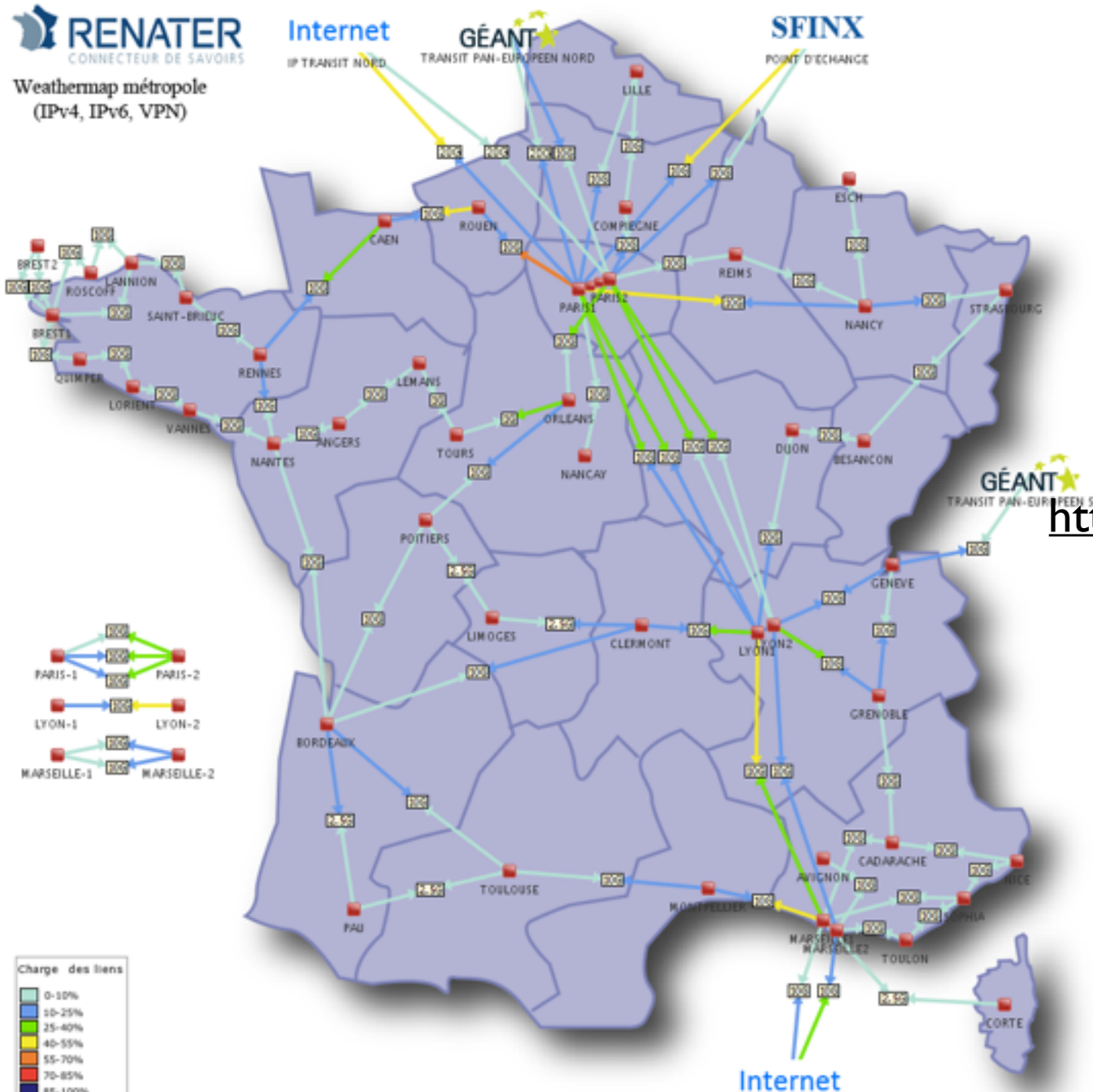
A simple Idea

Bring Clouds back to the cloud

Beyond the Clouds, the DISCOVERY Initiative

- Locality-based UC infrastructures / Fog / Edge

  A promising way to deliver highly efficient and sustainable UC services is to provide UC platforms as close as possible to the end-users.

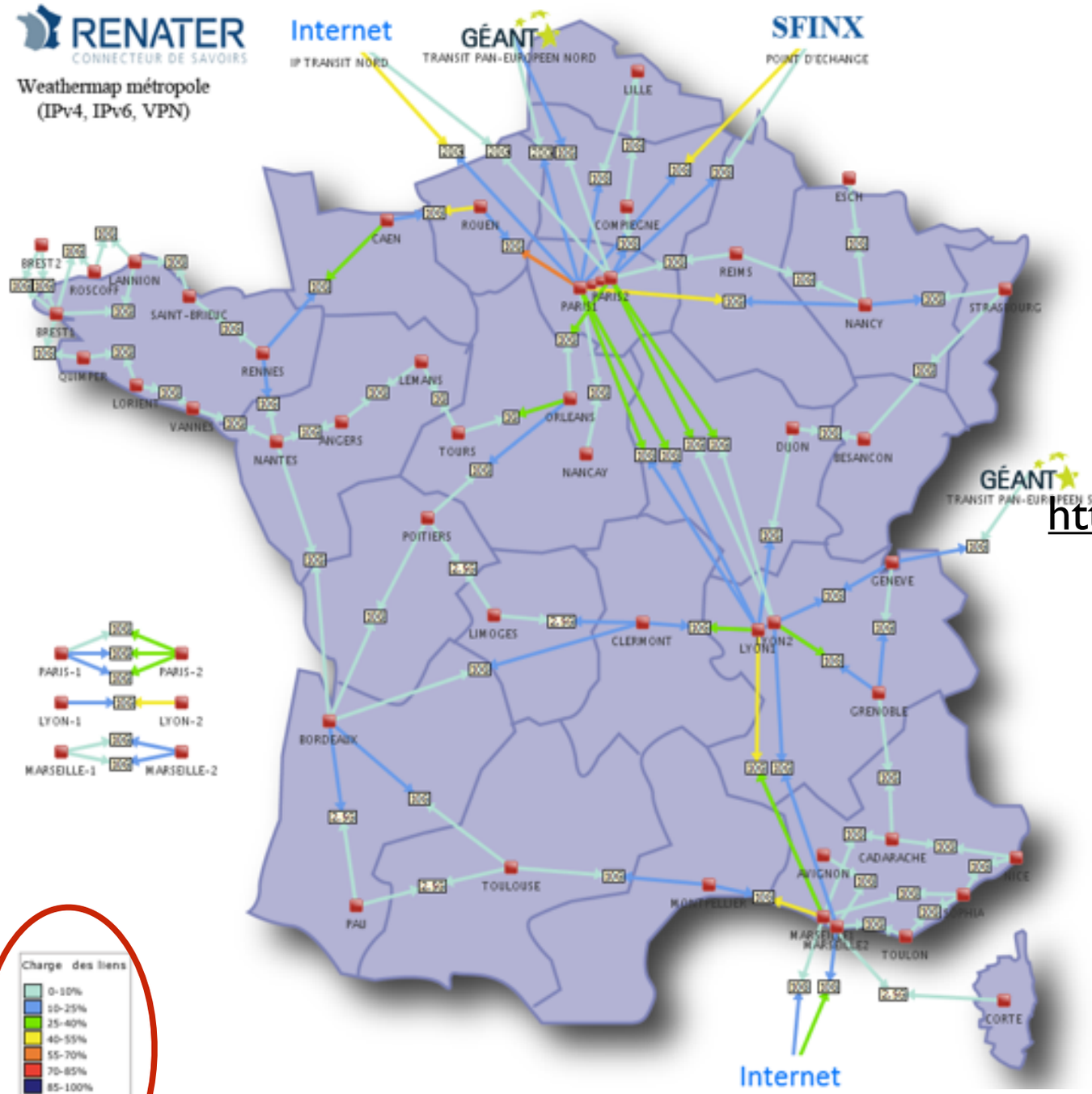


http://www.renater.fr/raccourci?lang=fr

Beyond the Clouds, the DISCOVERY Initiative

- Locality-based UC infrastructures / Fog / Edge

  A promising way to deliver highly efficient and sustainable UC services is to provide UC platforms as close as possible to the end-users.

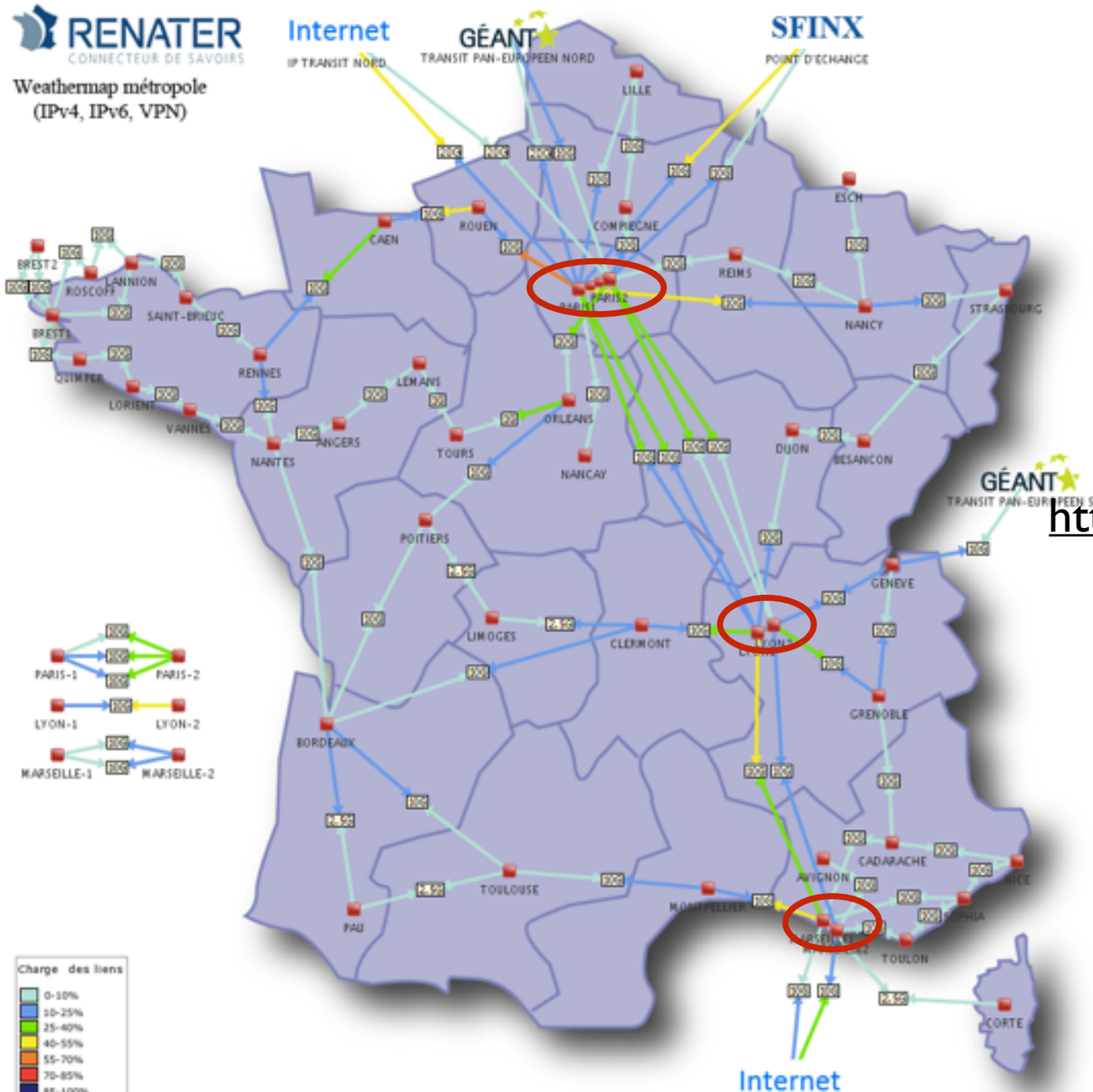


http://www.renater.fr/raccourci?lang=fr

Beyond the Clouds, the DISCOVERY Initiative

- Locality-based UC infrastructures / Fog / Edge

  A promising way to deliver highly efficient and sustainable UC services is to provide UC platforms as close as possible to the end-users.

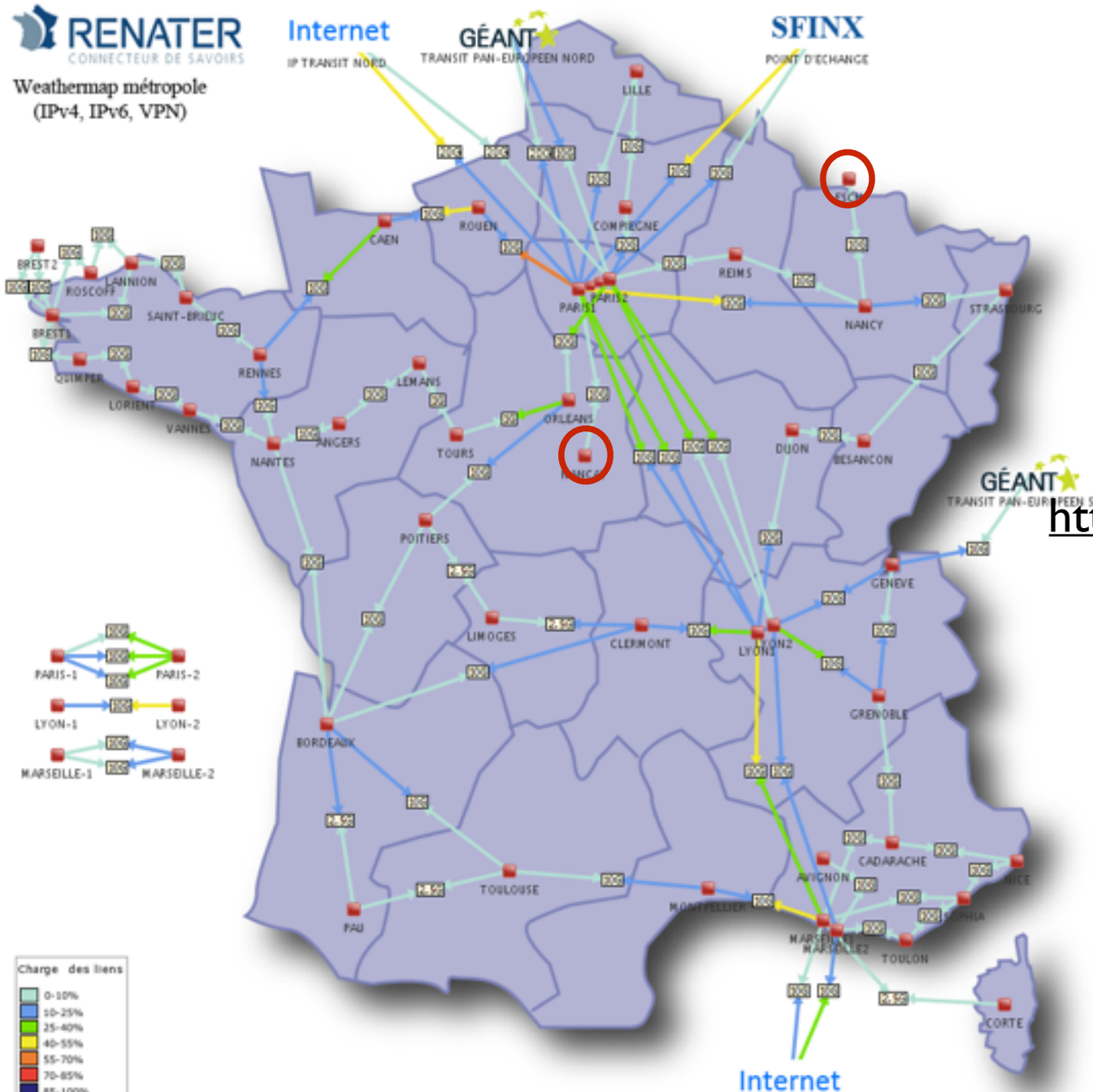


http://www.renater.fr/raccourci?lang=fr

Beyond the Clouds, the DISCOVERY Initiative

- Locality-based UC infrastructures / Fog / Edge

  A promising way to deliver highly efficient and sustainable UC services is to provide UC platforms as close as possible to the end-users.

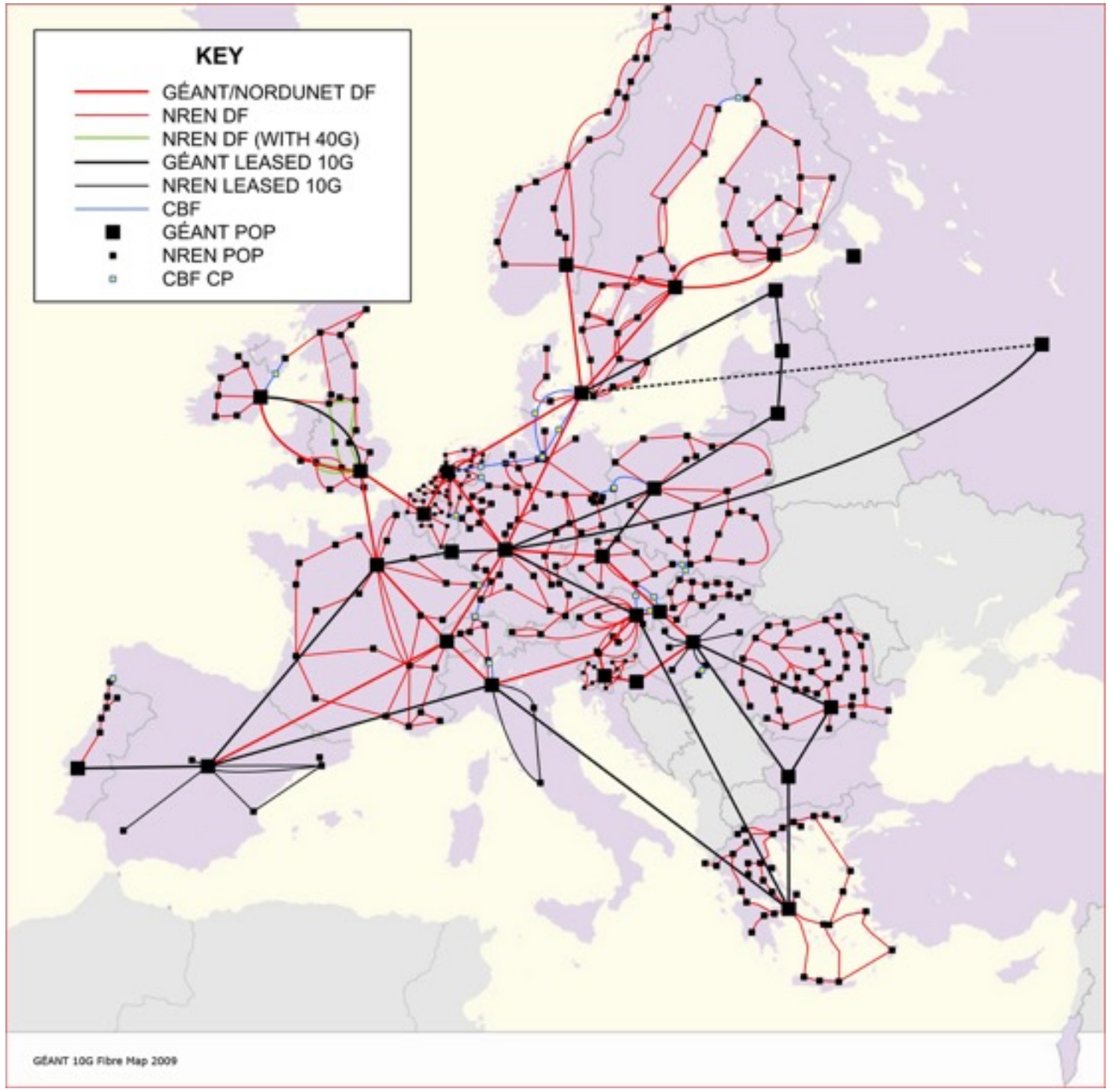


http://www.renater.fr/raccourci?lang=fr

Beyond the Clouds, the DISCOVERY Initiative

- Locality-based UC infrastructures / Fog / Edge

  A promising way to deliver highly efficient and sustainable UC services is to provide UC platforms as close as possible to the end-users.

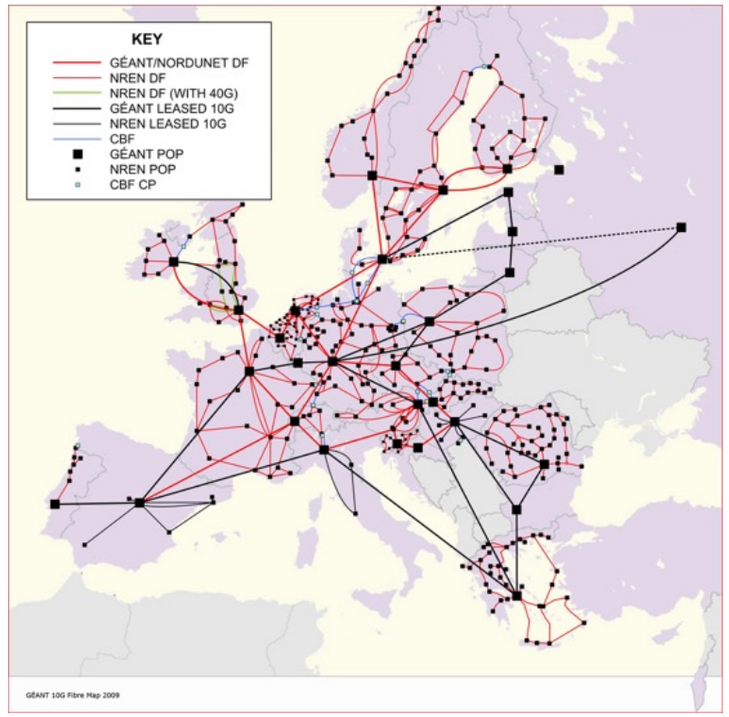


Fog/Edge Computing Infrastructures

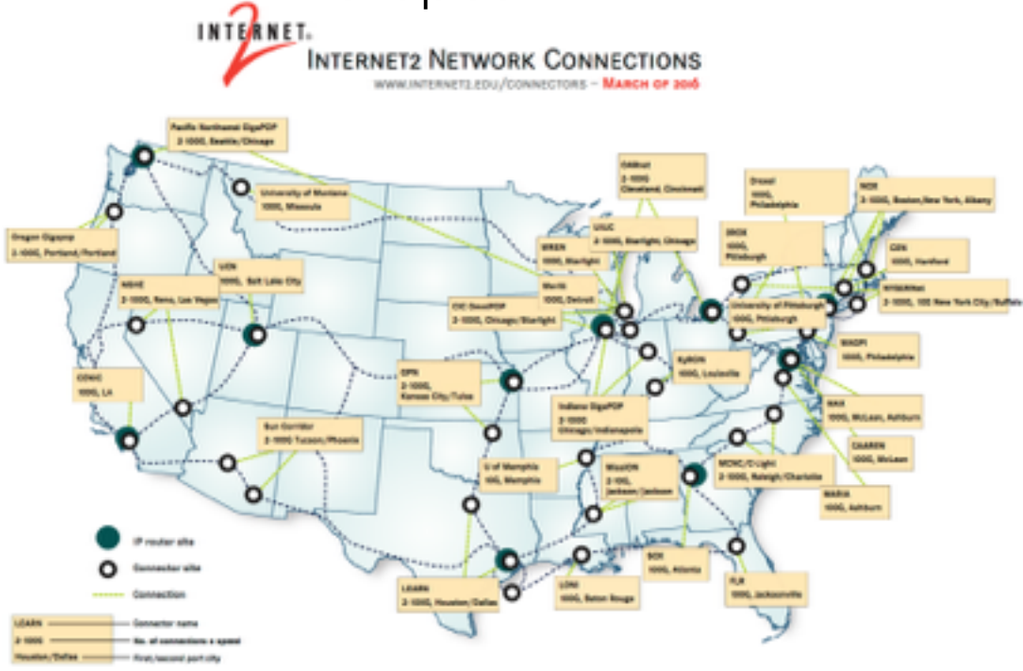
- Leverage network backbones

  Extend any point of presence of network backbones (aka PoP) with servers (from network hubs up to major DSLAMs that are operated by telecom companies, network institutions…).

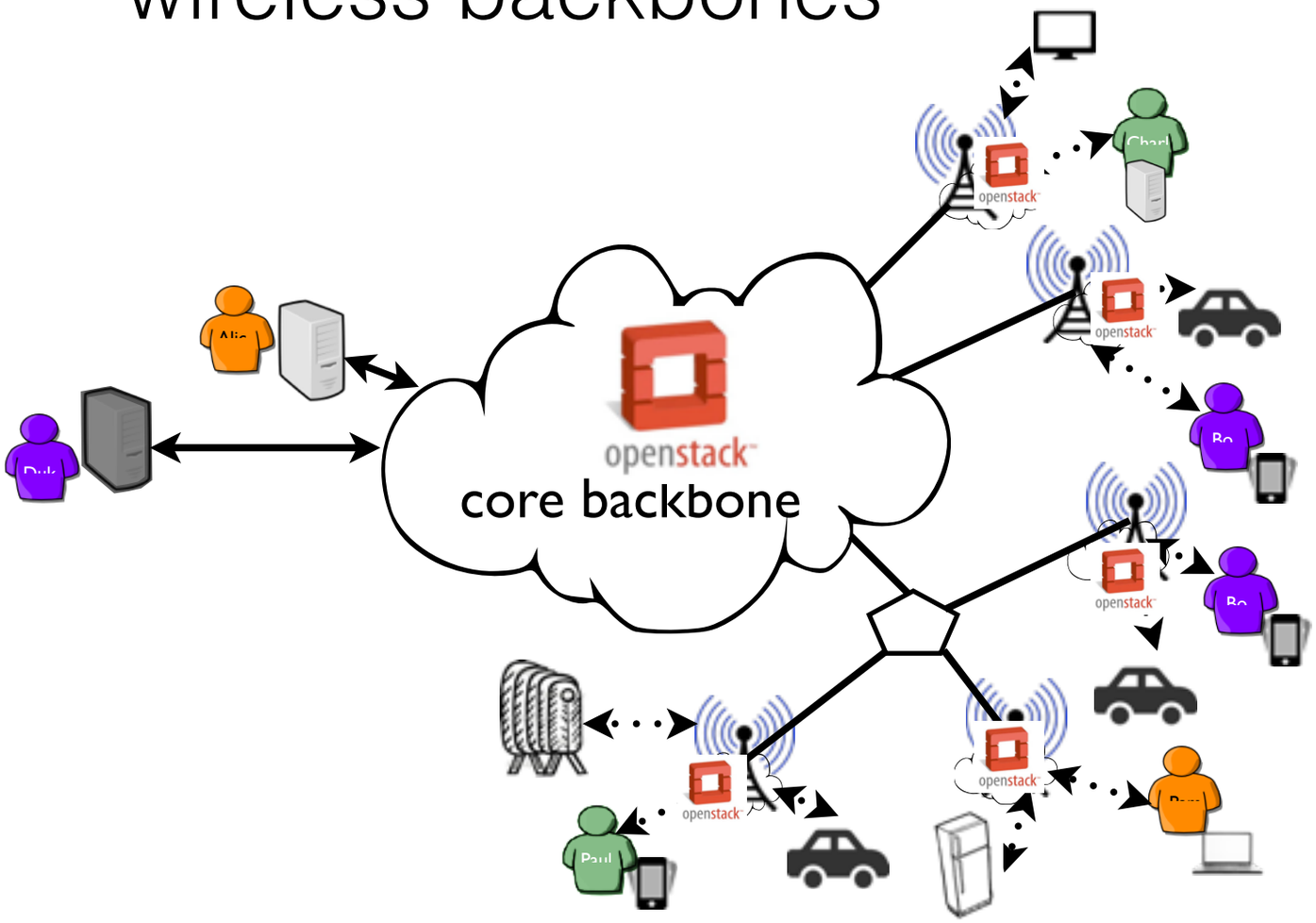
- Extend to the edge by including wireless backbones



European NREN



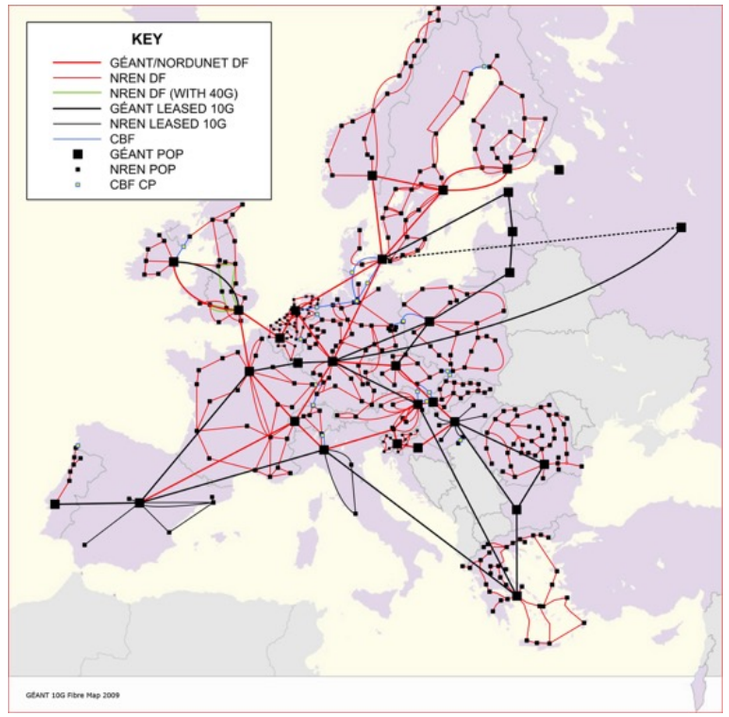
USA NREN



Fog/Edge Computing Infrastructures

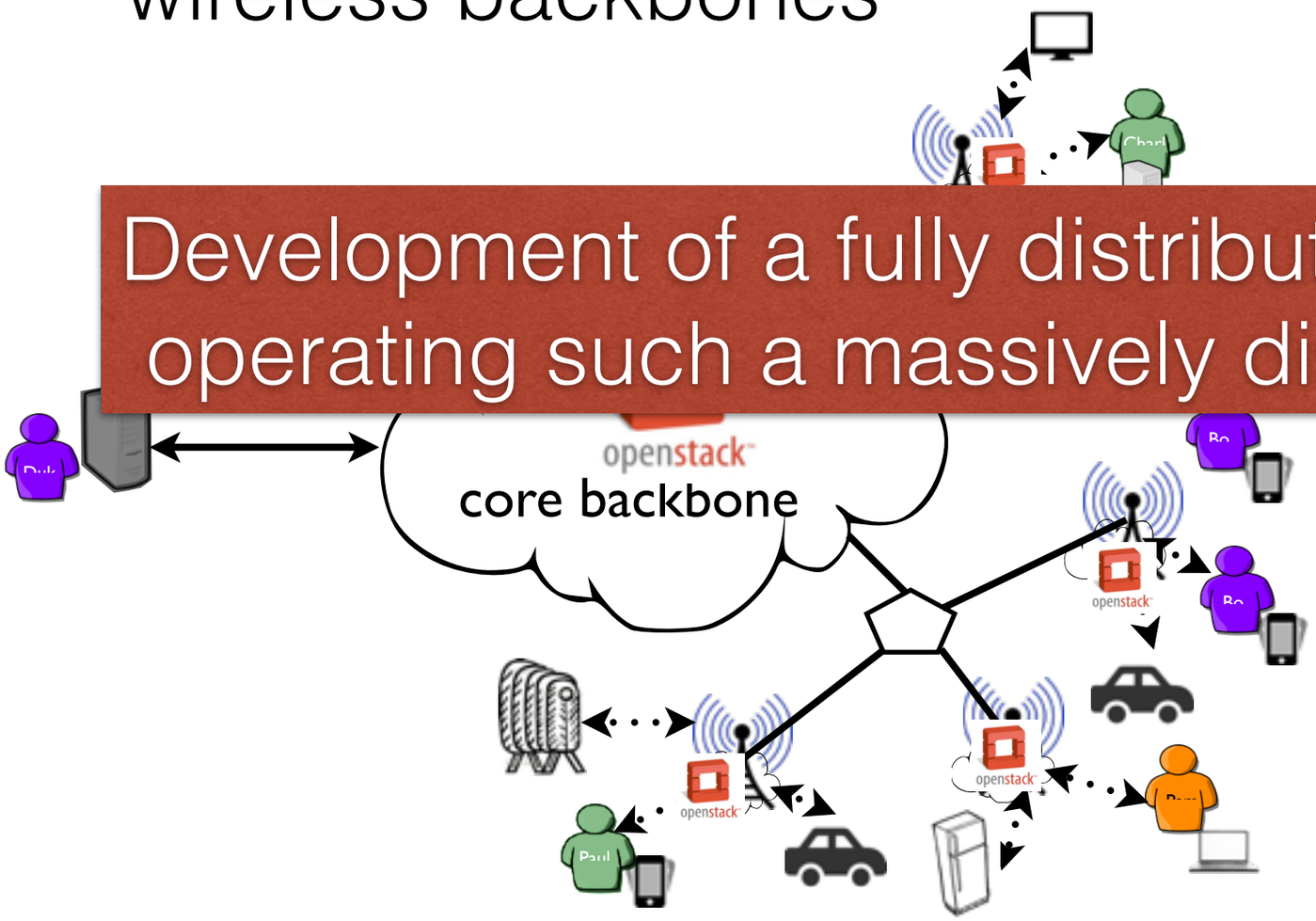
- Leverage network backbones

  Extend any point of presence of network backbones (aka PoP) with servers (from network hubs up to major DSLAMs that are operated by telecom companies, network institutions…).



European NREN

- Extend to the edge by including wireless backbones



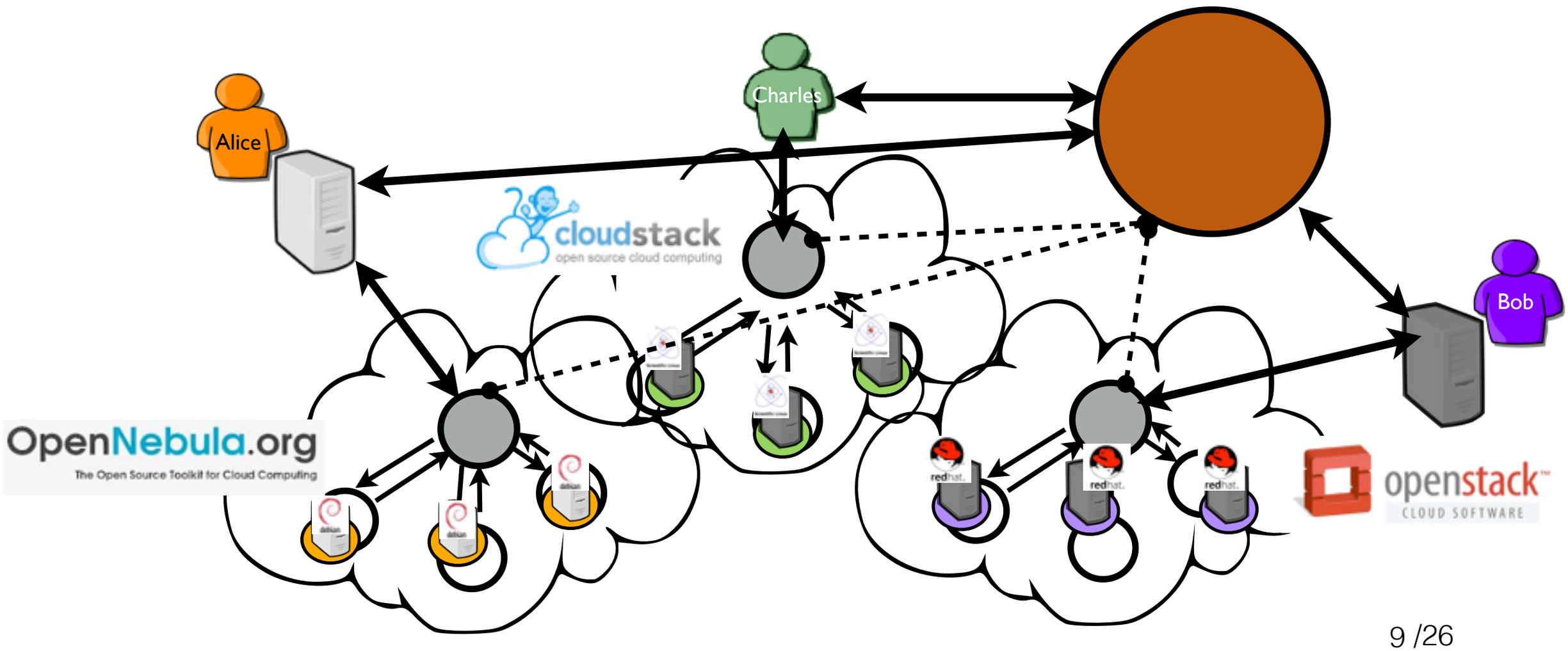
Development of a fully distributed system in charge of operating such a massively distributed infrastructure

USA NREN

What's about Brokering Approaches?

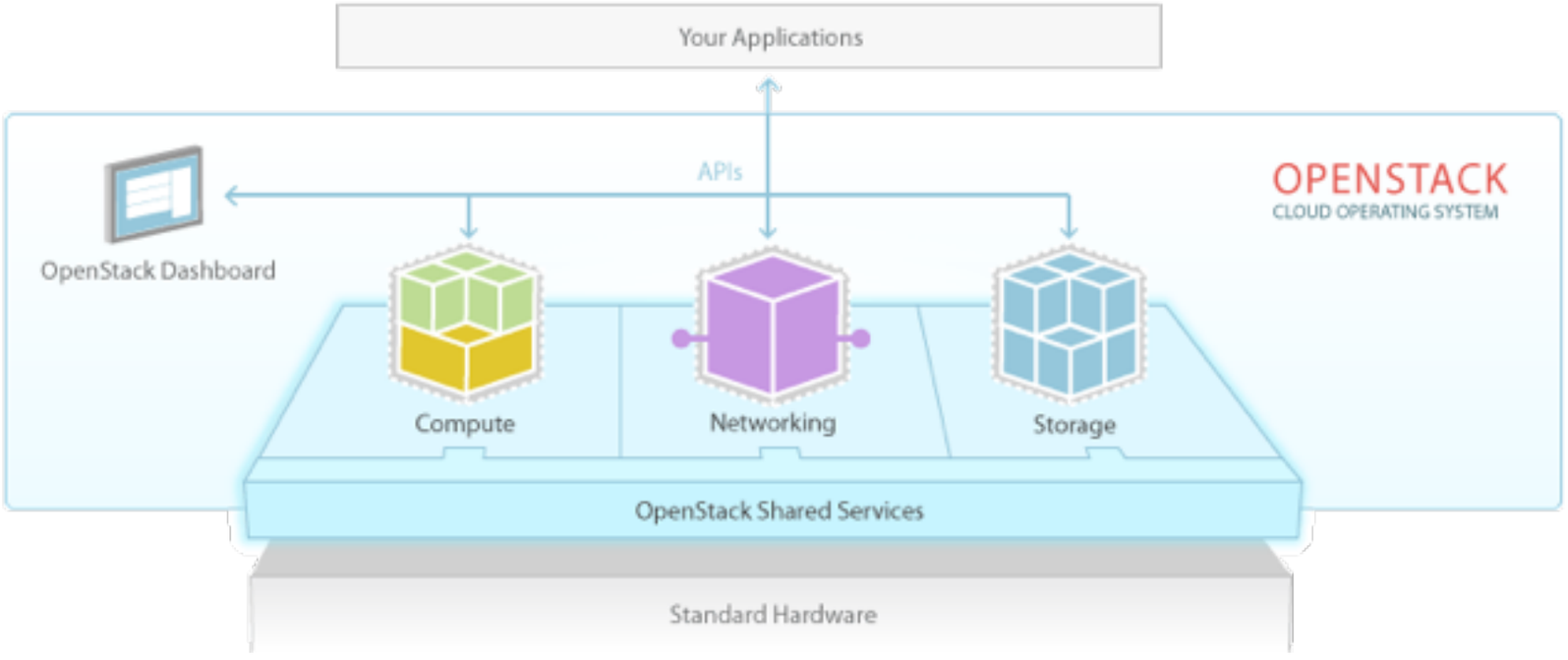
- Sporadic (hybrid computing/cloud bursting) almost ready for production
- While standards are coming (OCCI, ….), current brokers are rather limited

Advanced brokers must reimplement standard IaaS mechanisms while facing the API limitation



Would OpenStack be the solution?

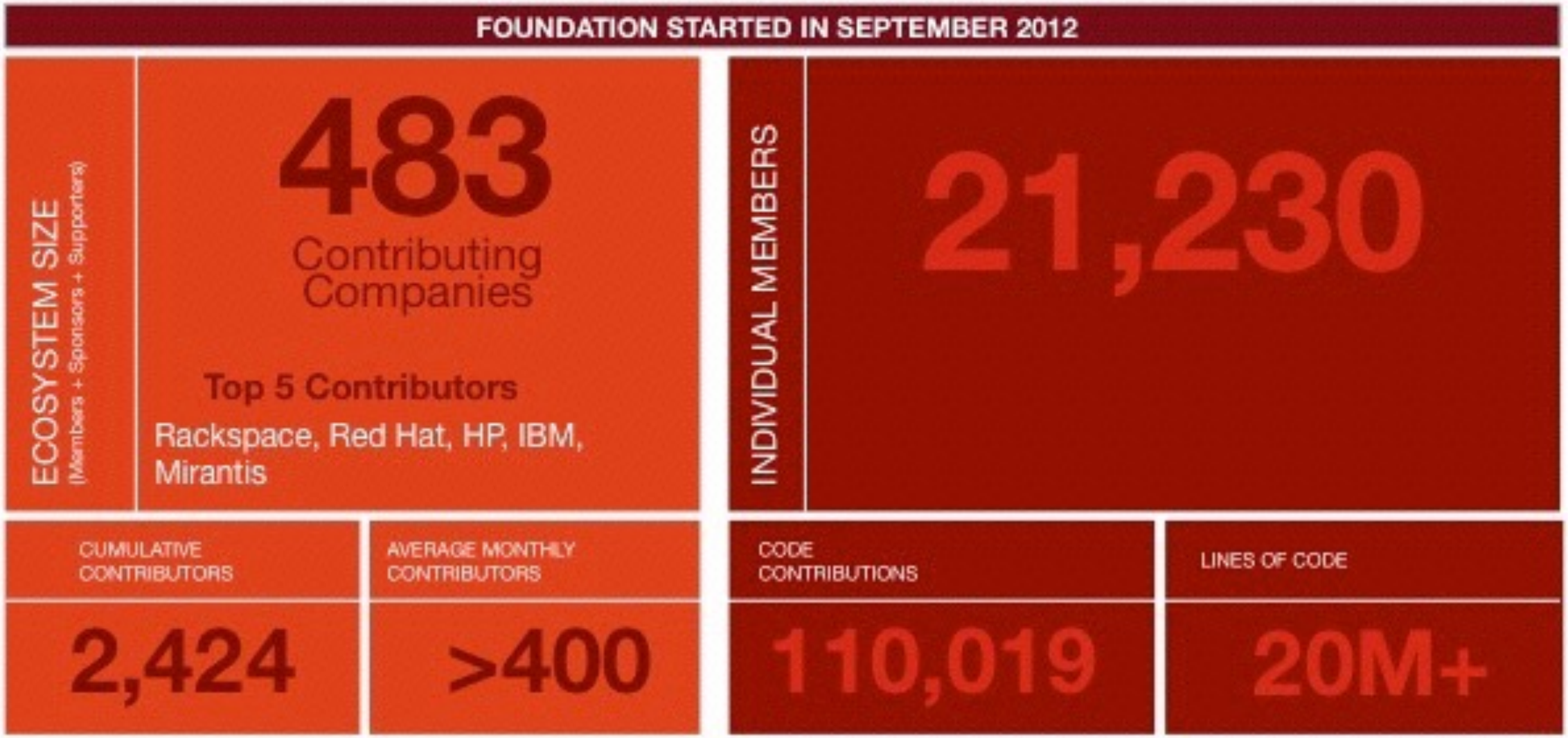
- Do not reinvent the wheel… it is too late



Would OpenStack be the solution?

OPENSTACK COMMUNITY: BROAD SUPPORT AND CONTRIBUTION

openstack
CLOUD SOFTWARE



rackspace
the #1 managed cloud company

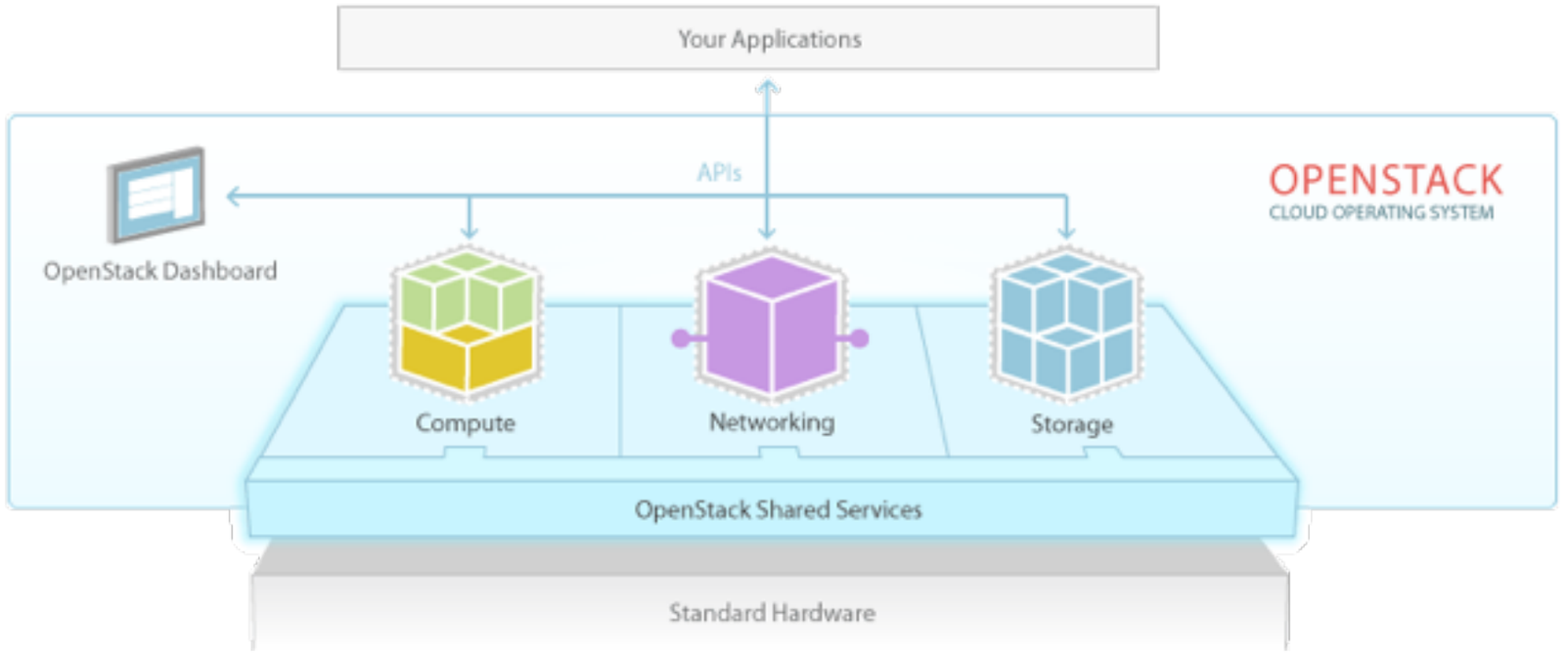
www.rackspace.com 6

2Millions of LOCs just for core-services

Would OpenStack be the solution?

- Do not reinvent the wheel… it is too late
- Few proposals to federate/operate distinct OpenStack DCs

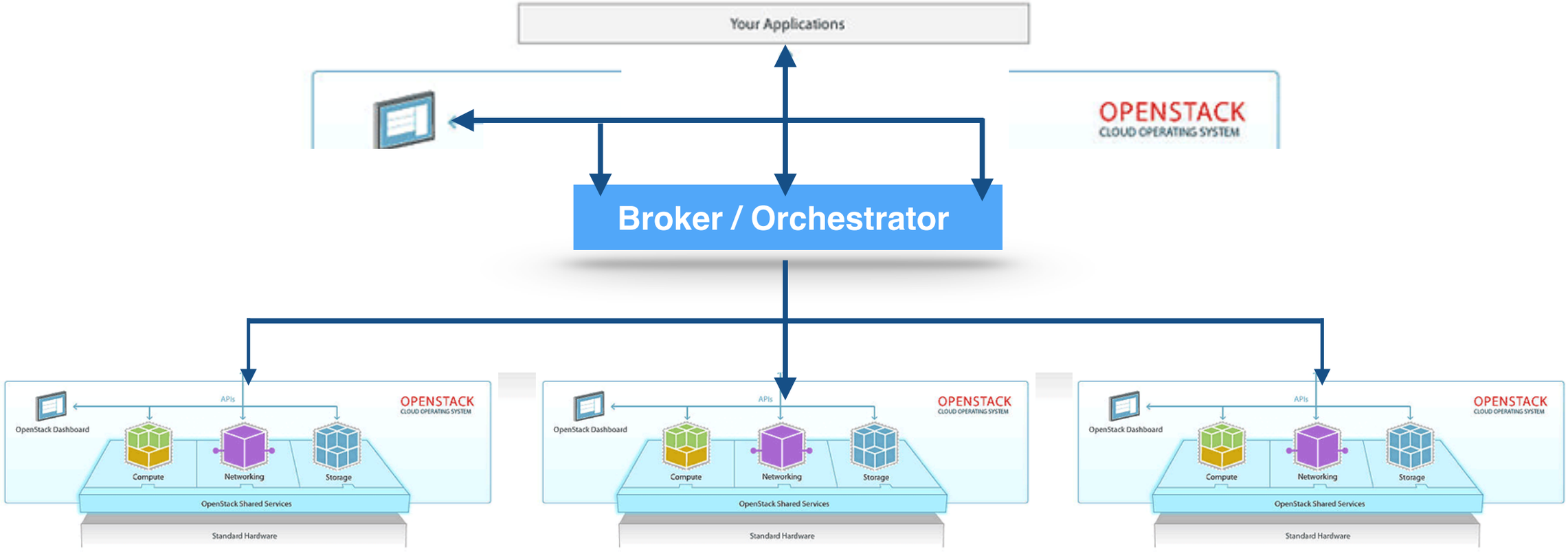
Top/Down: add a substrate to pilot independent OpenStack instances



Would OpenStack be the solution?

- Do not reinvent the wheel… it is too late
- Few proposals to federate/operate distinct OpenStack DCs

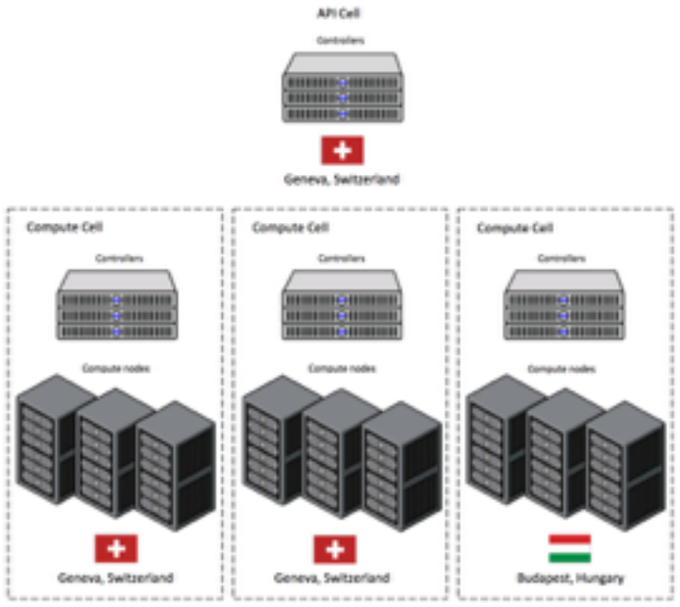
Top/Down: add a substrate to pilot independent OpenStack instances



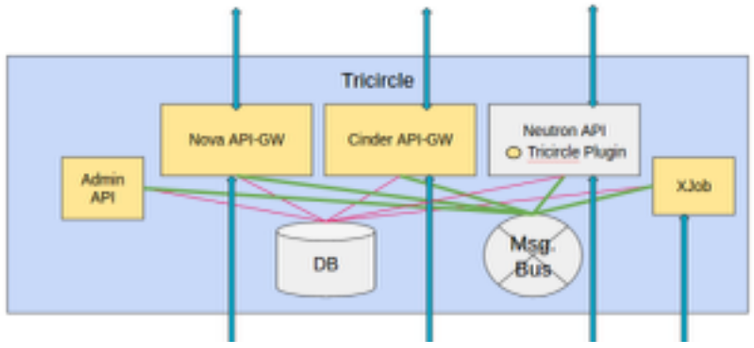
Would OpenStack be the solution?

- Do not reinvent the wheel… it is too late
- Few proposals to federate/operate distinct OpenStack DCs

Top/Down: add a substrate to pilot independent OpenStack instances



Cell v2 (Nova)

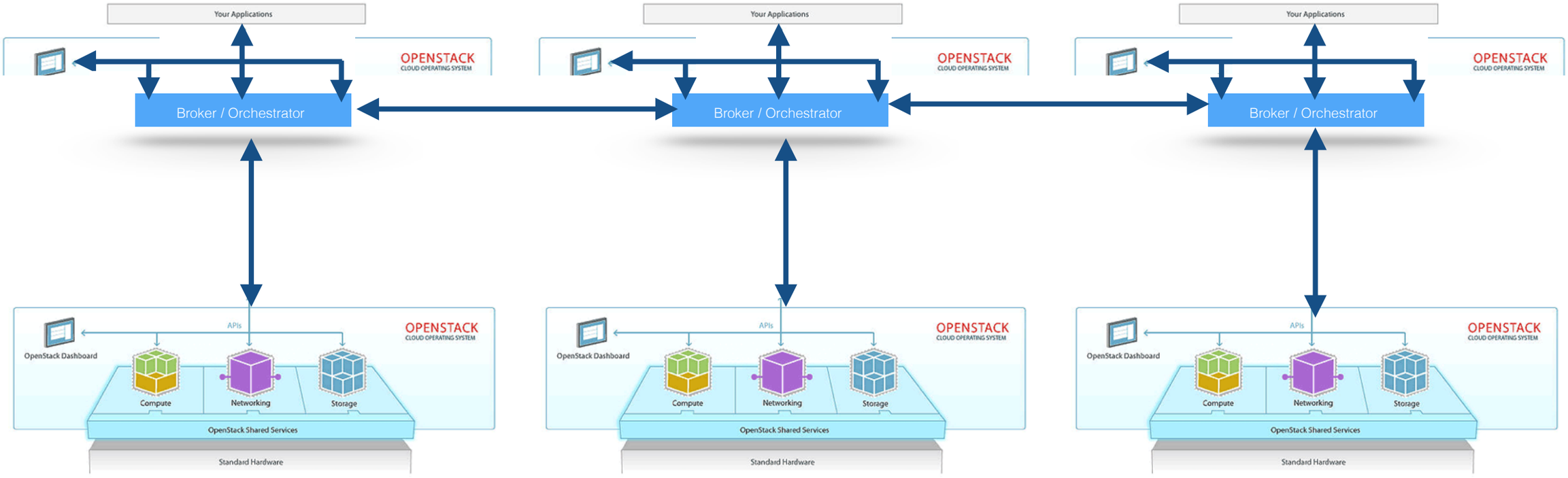


Tricircle (previously cascading OpenStack)

Would OpenStack be the solution?

- Do not reinvent the wheel… it is too late
- Few proposals to federate/operate distinct OpenStack DCs

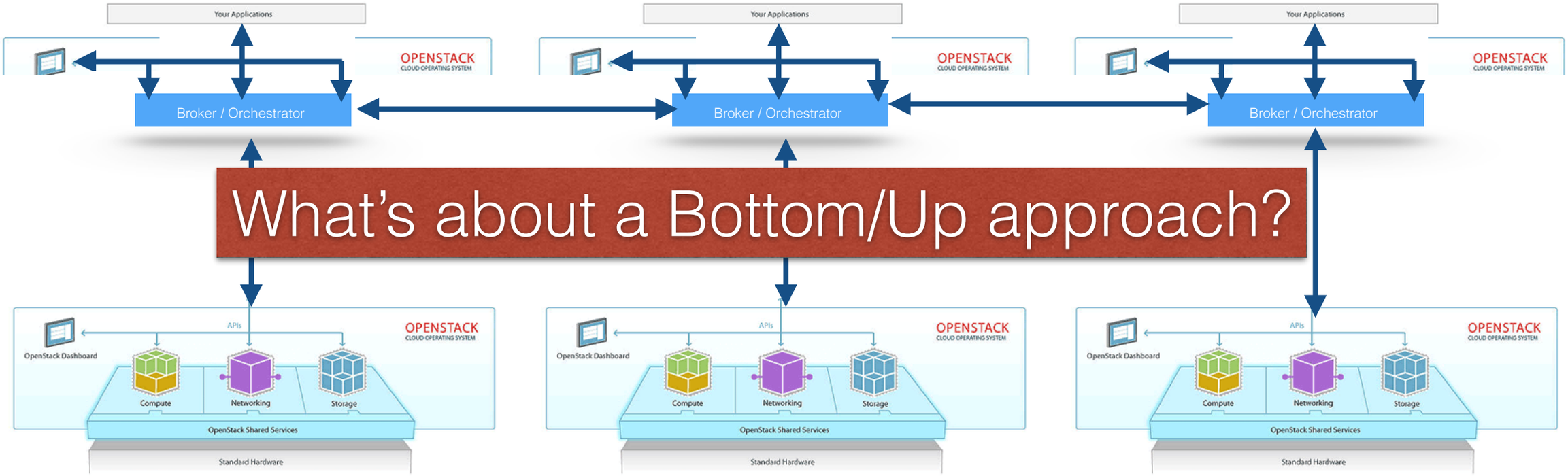
Top/Down: add a substrate to pilot independent OpenStack instances



Would OpenStack be the solution?

- Do not reinvent the wheel… it is too late
- Few proposals to federate/operate distinct OpenStack DCs

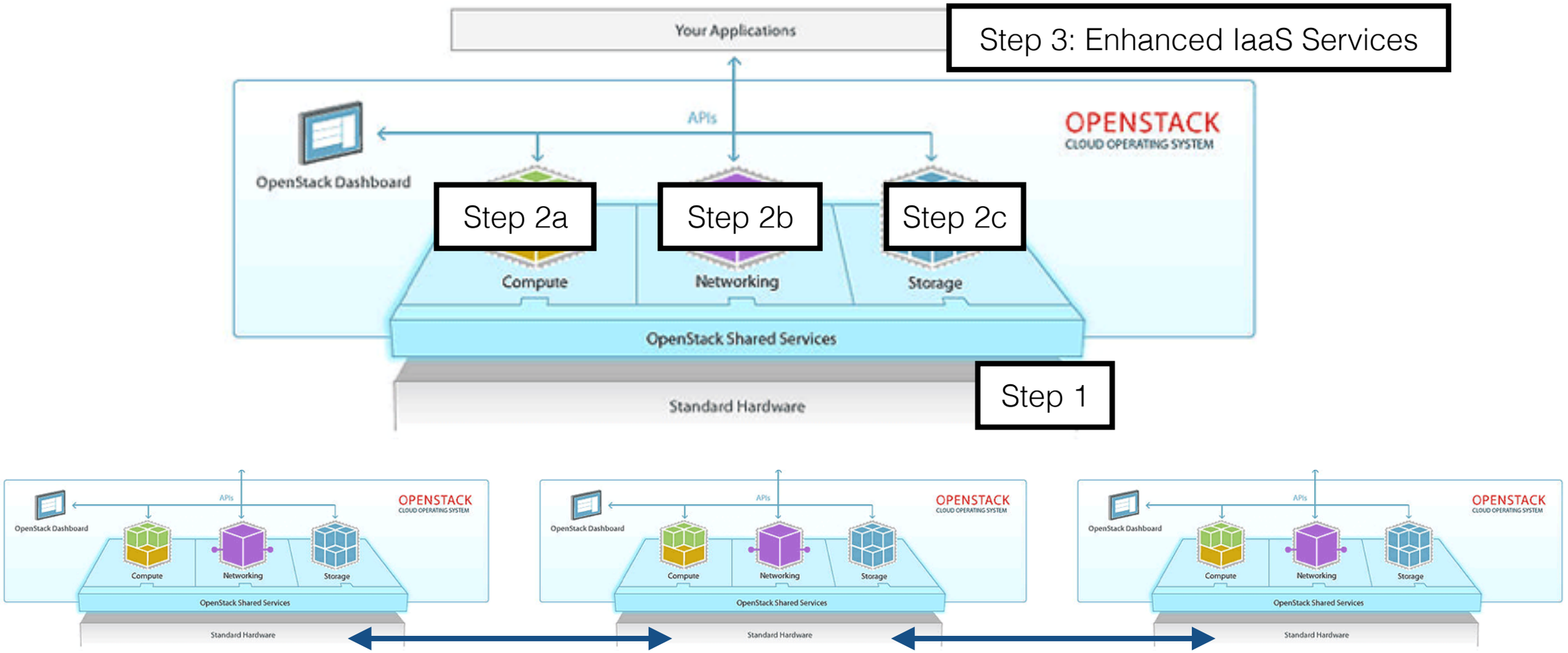
Top/Down: add a substrate to pilot independent OpenStack instances



Would OpenStack be the solution?

- Do not reinvent the wheel… it is too late
- Few proposals to federate/operate distinct OpenStack DCs

Bottom/Up - investigate whether/how OpenStack core services can be cooperative by default using Self* and P2P mechanisms



Would OpenStack be the solution?

- Do not reinvent the wheel… it is too late
- Few proposals to federate/operate distinct OpenStack DCs

Bottom/Up - investigate whether/how OpenStack core services can be cooperative by default

openstack CLOUD SOFTWARE

Technical considerations

SIDEBAR ◂ PREV | UP | NEXT ▸

CONTENTS SEARCH

OPENSTACK MANUALS ▸ OPENSTACK ARCHITECTURE DESIGN GUIDE - CURRENT

Technical considerations

Repurposing an existing OpenStack environment to be massively scalable is a formidable task. When building a massively scalable environment from the ground up, ensure you build the initial deployment with the same principles and choices that apply as the environment grows. For example, a good approach is to deploy the first site as a multi-site environment. This enables you to use the same deployment and segregation methods as the environment grows to separate locations across dedicated links or wide area networks. In a hyperscale cloud, scale trumps redundancy. Modify applications with this in mind, relying on the scale and homogeneity of the environment to provide reliability rather than redundant infrastructure provided by non-commodity hardware solutions.

Infrastructure segregation

OpenStack services support massive horizontal scale. Be aware that this is not the case for the entire supporting infrastructure. This is particularly a problem for the database management systems and message queues that OpenStack services use for data storage and remote procedure call communications.

Traditional clustering techniques typically provide high availability and some additional scale for these environments. In the quest for massive scale, however, you must take additional steps to relieve the performance pressure on these components in order to prevent them from negatively impacting the overall performance of the environment. Ensure that all the components are in balance so that if the massively scalable environment fails, all the components are near maximum capacity and a single component is not causing the failure.

Would OpenStack be the solution?

- Do not reinvent the wheel… it is too late
- Few proposals to federate/operate distinct OpenStack DCs

Bottom/Up - investigate whether/how OpenStack core services can be cooperative by default

Technical considerations

SIDEBAR ◂ PREV | UP | NEXT ▸

OPENSTACK MANUALS ▸ OPENSTACK ARCHITECTURE DESIGN GUIDE - CURRENT

Technical considerations

Repurposing an existing OpenStack environment to be massively scalable is a formidable task. When building a massively scalable environment from the ground up, ensure you build the initial deployment with the same principles and choices that apply as the environment grows. For example, a good approach is to deploy the first site as a multi-site environment. This enables you to use the same deployment and segregation methods as the environment grows to separate locations across dedicated links or wide area networks. In a hyperscale cloud, scale trumps redundancy. Modify applications with this in mind, relying on the scale and homogeneity of the environment to provide reliability rather than redundant infrastructure provided by non-commodity hardware solutions.

Infrastructure segregation

OpenStack services support massive horizontal scale. Be aware that this is not the case for the entire supporting infrastructure. This is particularly a problem for the database management systems and message queues that OpenStack services use for data storage and remote procedure call communications.

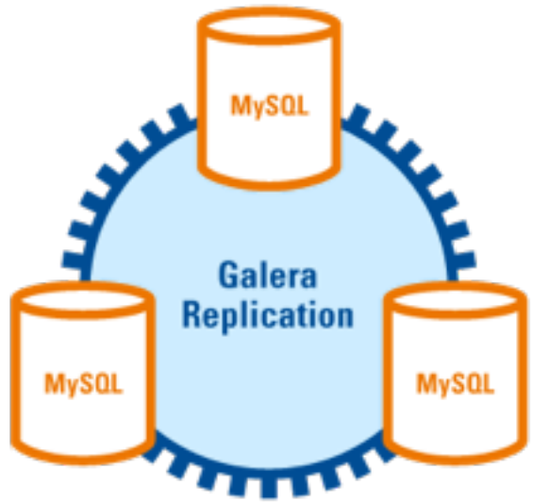
Traditional clustering techniques typically provide high availability and some additional scale for these environments. In the quest for massive scale, however, you must take additional steps to relieve the performance pressure on these components in order to prevent them from negatively impacting the overall performance of the environment. Ensure that all the components are in balance so that if the massively scalable environment fails, all the components are near

Distributing OpenStack Through a Bottom/Up Approach

- Step 1: OpenStack shared services

  A SQL database 

  A messaging queue 



- Active/Active replication

  Production ready but does not scale to our target.

- Key/Value Store systems

  Alternate solutions for storing states over a highly distributed infrastructure

Distributing OpenStack Through a Bottom/Up Approach

- Step 1: OpenStack shared services

  A SQL database

  A messaging queue

- Active/A… …tion

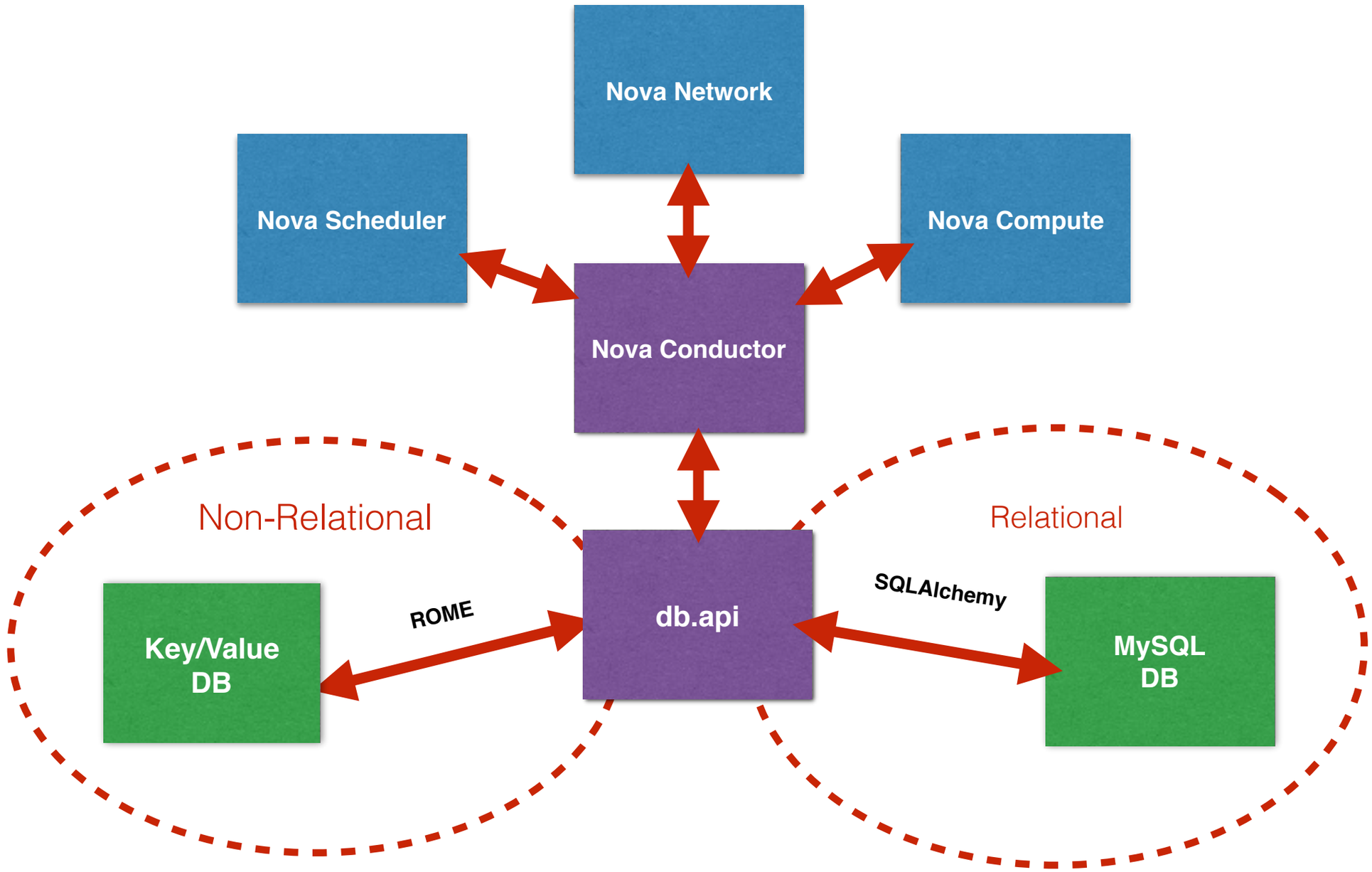
  …y but does not scale to our target.

- …y/Value Store systems

  Alternate solutions for storing states over a highly distributed infrastructure

How can we switch between a SQL solution and a NoSQL system for storing inner states of OpenStack ?

Leveraging a Key/Value Store DB

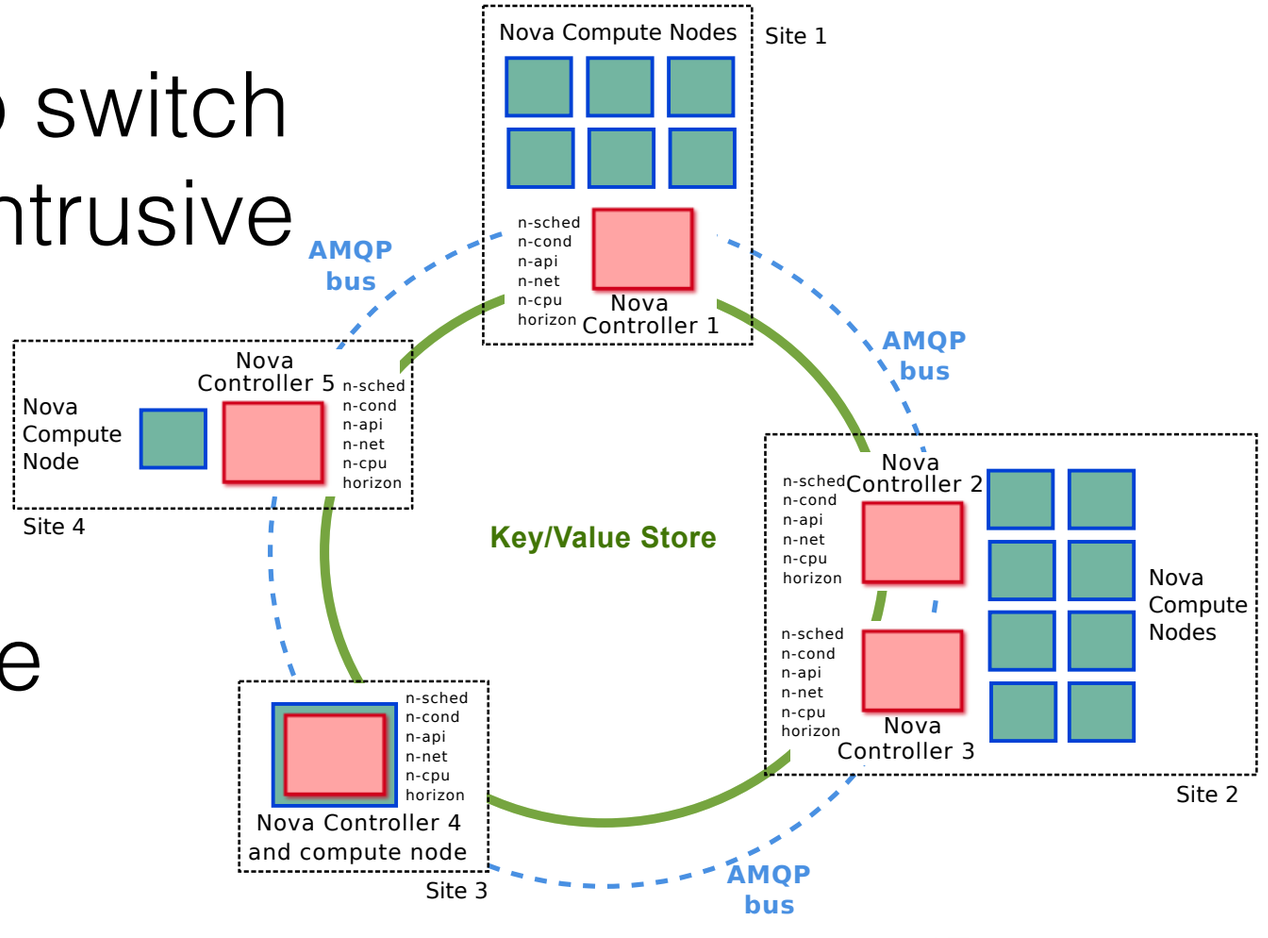


Nova (compute service) - software architecture

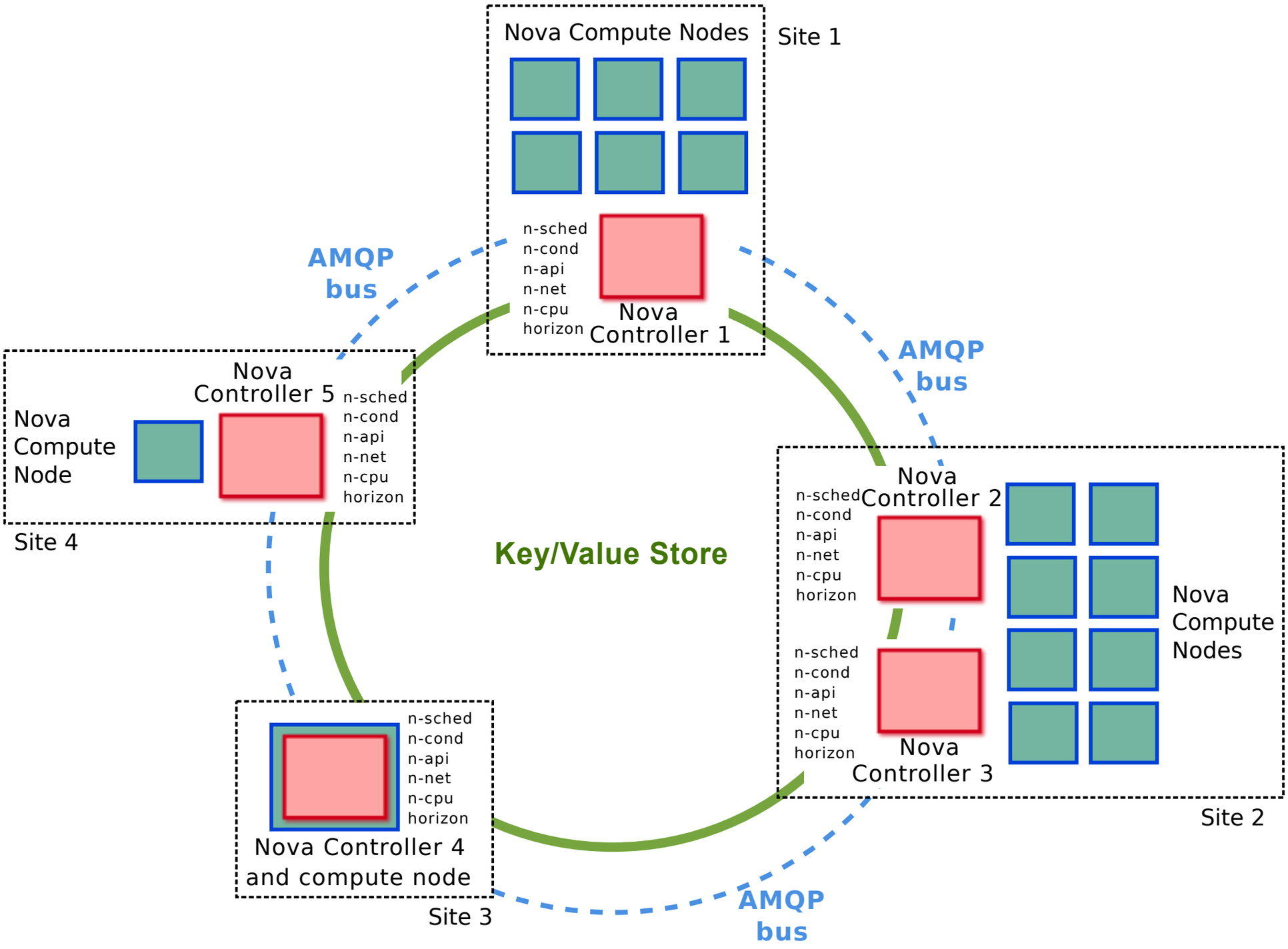
ROME

- Relational Object Mapping Extension for key/value stores
  Jonathan Pastor's Phd
  https://github.com/BeyondTheClouds/rome
- Enables the query of Key/Value Store DB with the same interface as SQLAlchemy

- Enables Nova OpenStack to switch to a KVS without being too intrusive
- The KVS is distributed over (dedicated) nodes
- Nova services connect to the Key/value store cluster

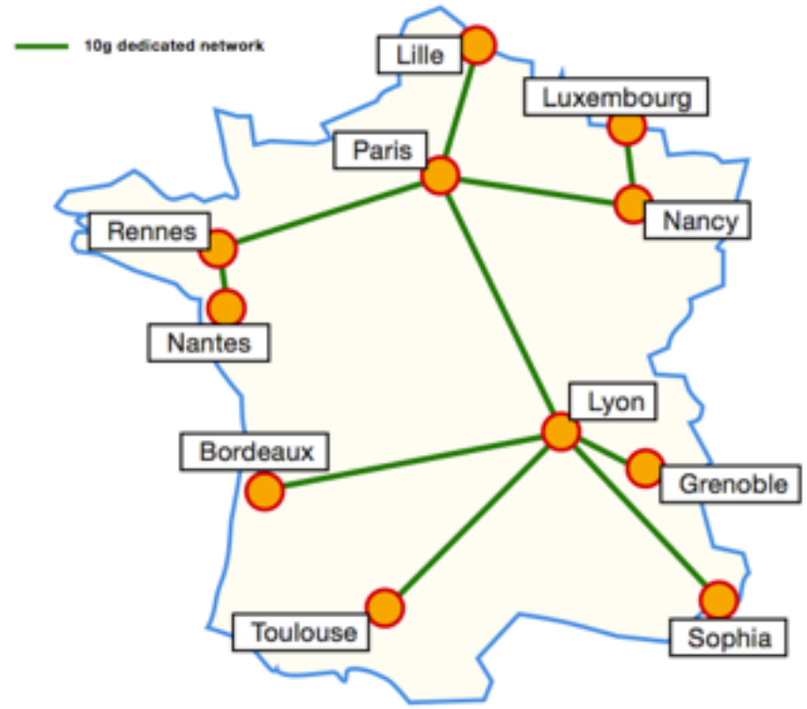


Nova Proof-Of-Concept



Experiments

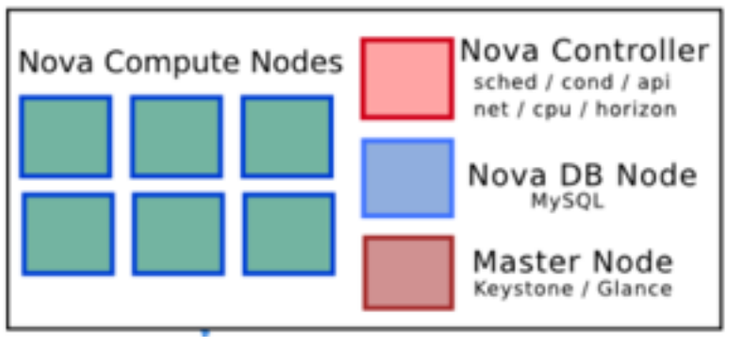
- Experiments have been conducted on Grid'5000
- Mono-site experiments
  ⟹ Evaluate the overhead of using ROME/Redis and the network impact.
- Multi-site experiments
  ⟹ Determine the impact of latency.
  ⟹ Validate compatibility with higher level mechanisms validation



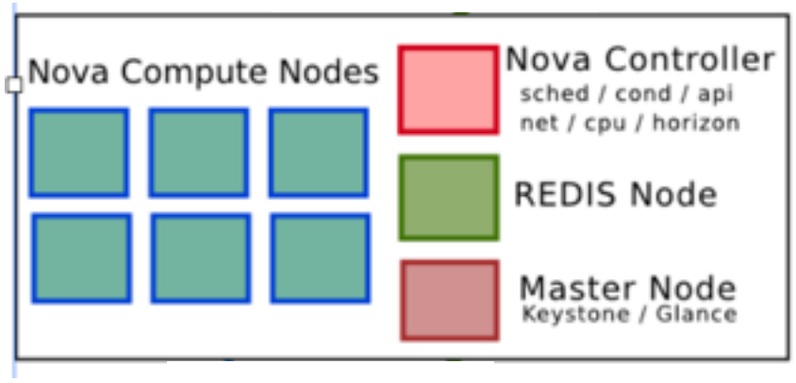
www.grid5000.fr
1500 servers, spread across 10 sites
Full admin rights

Mono-Site Experiments

- Creation of 500 VMs
- Comparison MySQL/SQLAlchemy vs ROME/Redis
  (one dedicated node for the DB server/the REDIS server)



MySQL/SQLAlchemy

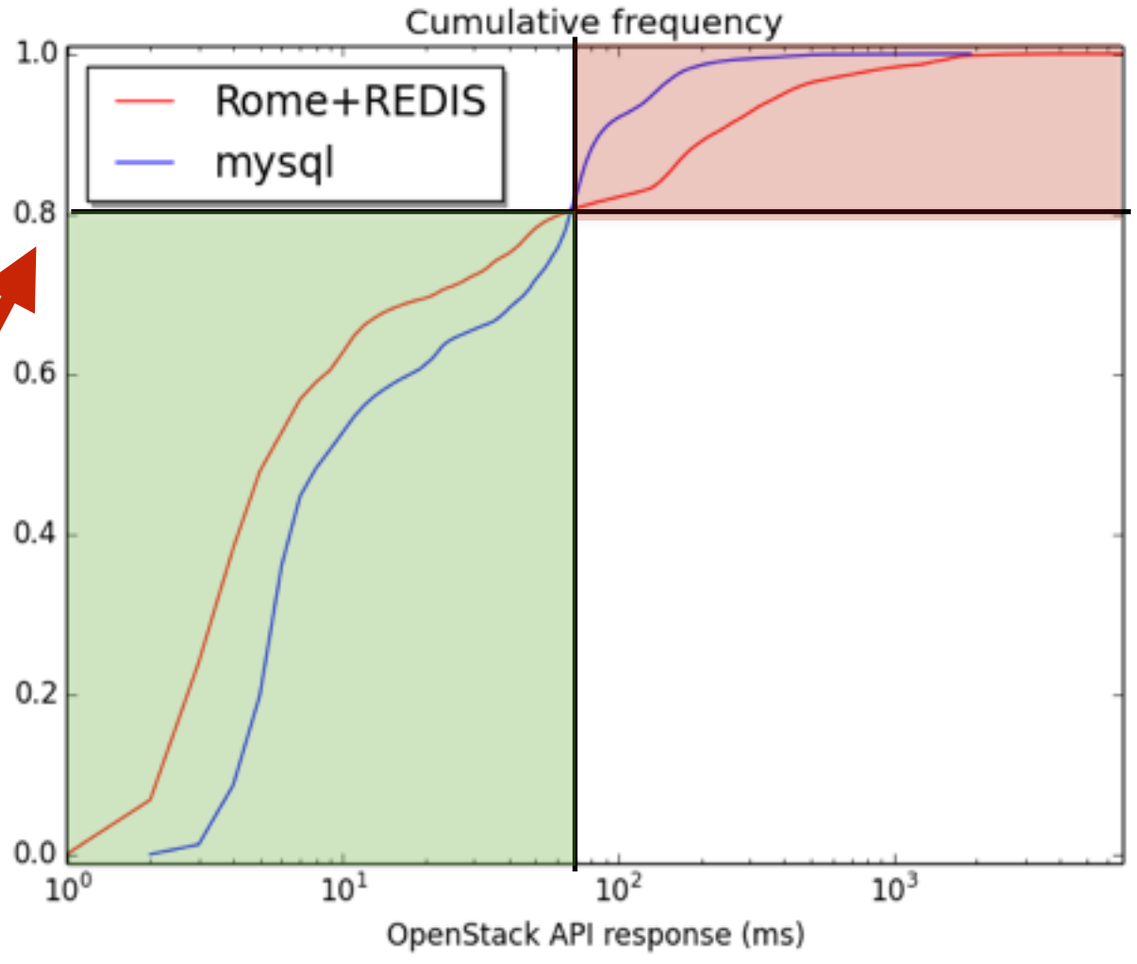


ROME/Redis

Mono-Site Experiments

- Evaluate the overhead of using ROME/Redis
- ROME stores objects in a JSON format: *serialization/deserialization cost*
- ROME reimplements some mechanisms: *join, transaction/session, …*

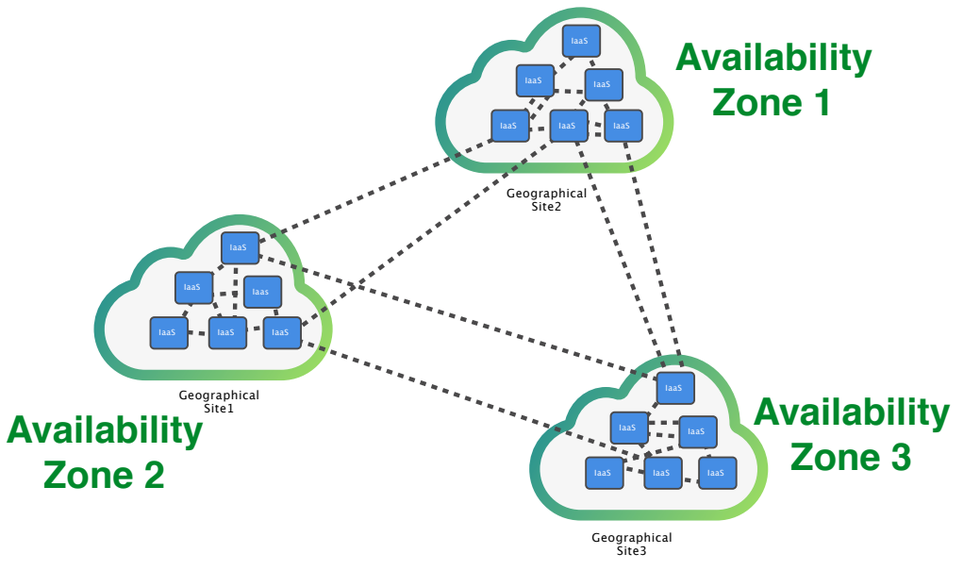
ROME requests are faster for 80% of requests



SQLAlchemy is faster for 20% of requests

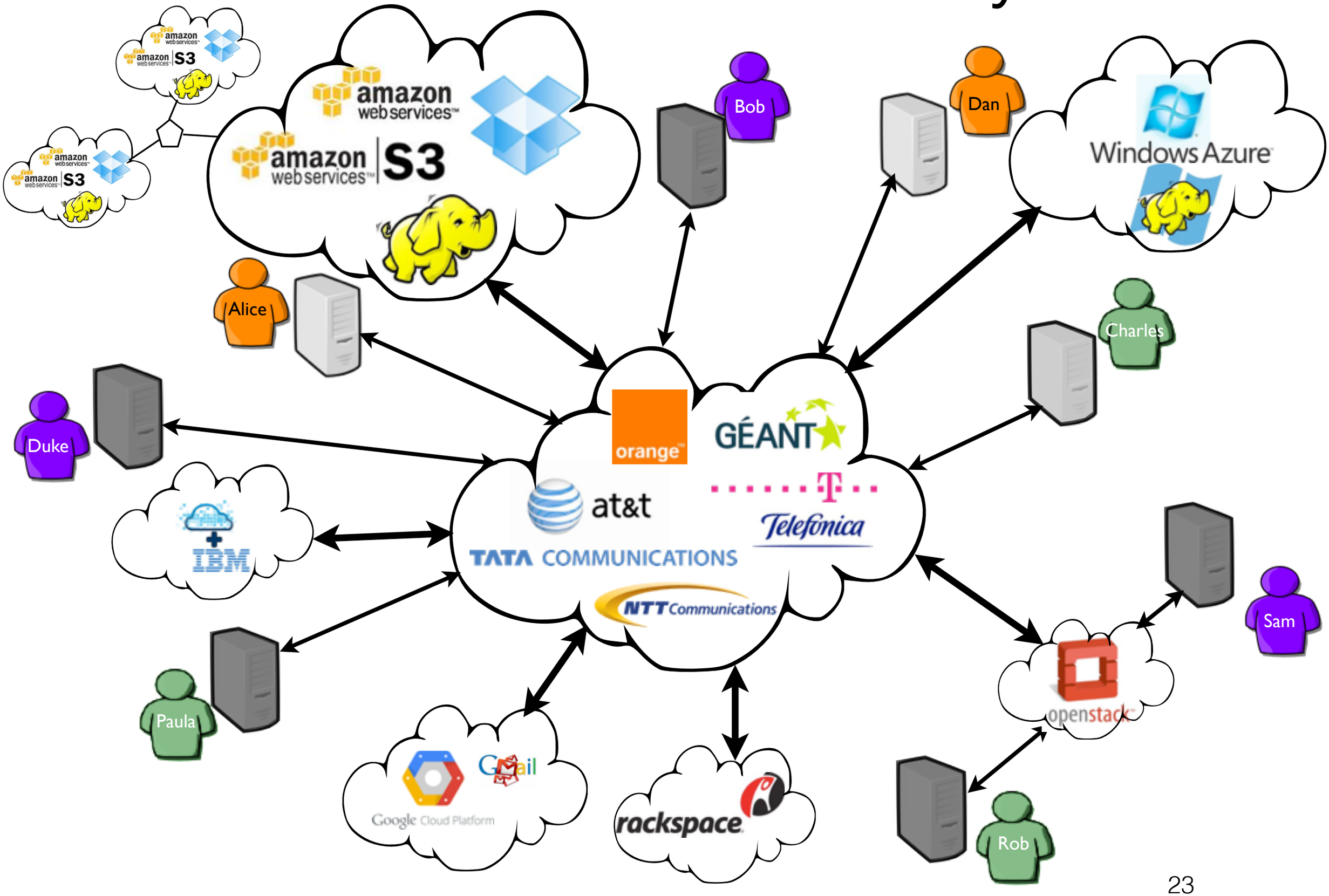
Compatibility with Higher Level Features

- Asses the usage of advanced OpenStack feature:
  host-aggregates / availability zones
- As we targeted a low-level component, ROME is compatible with most of the existing features.
- Performance is not impacted (same order of magnitude)
- VM Repartition is correctly achieved
  (without availability zones the distribution was respectively 26%, 20%, 22%, 32% of the created VMs for a 4 clusters experiments).

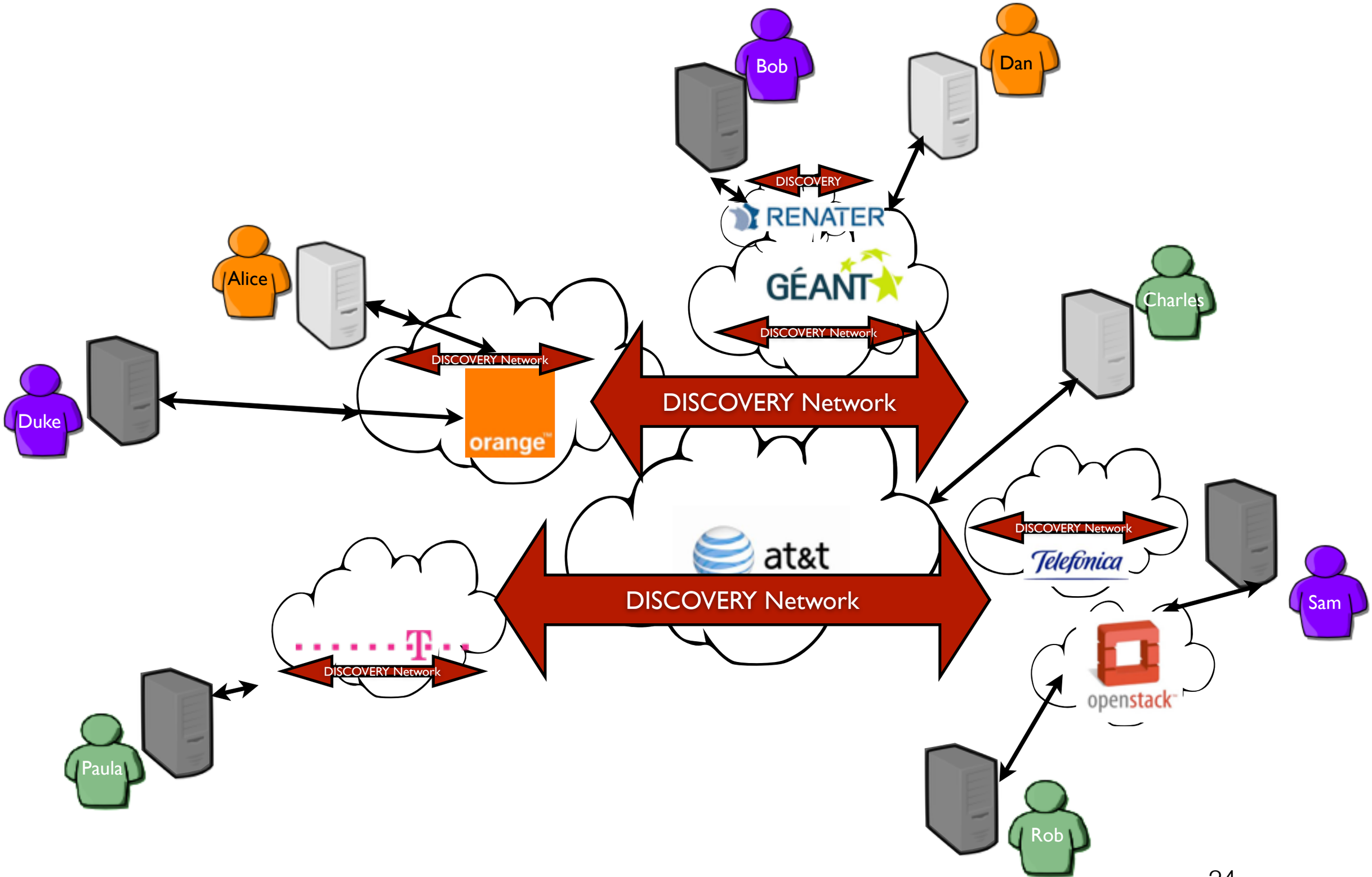


Can we go beyond a research POC ?

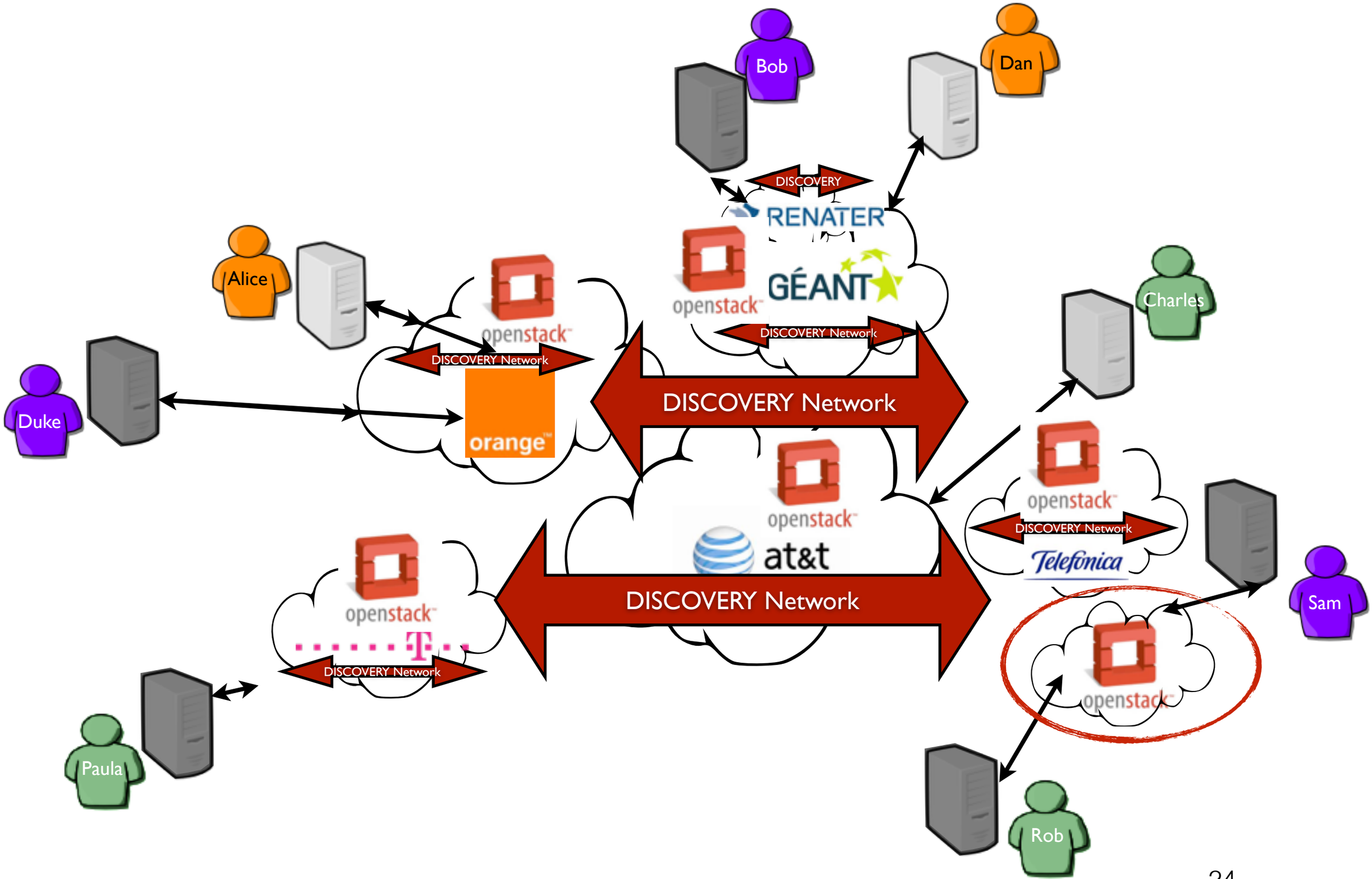
The Cloud in Reality



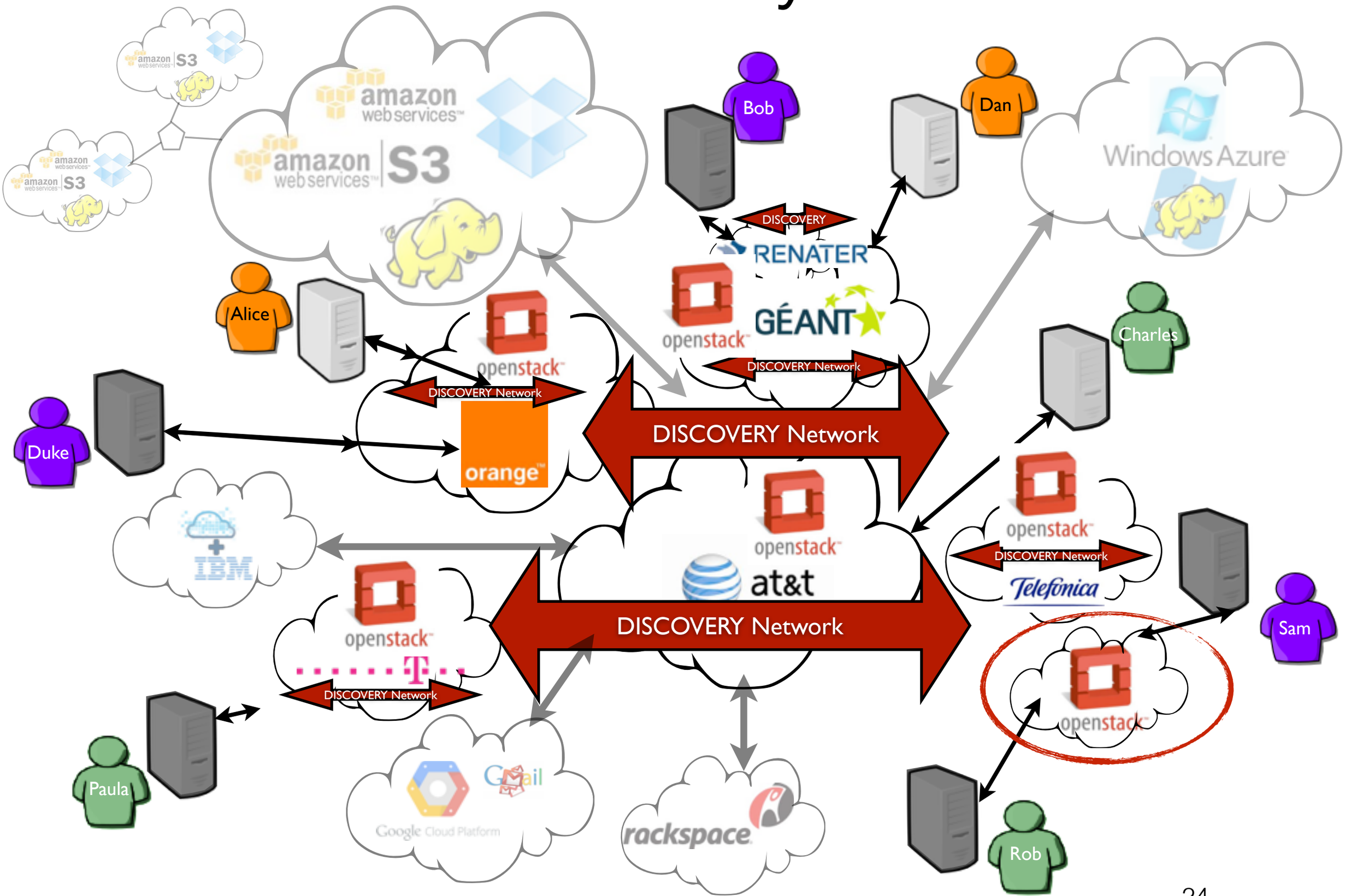
The Discovery Vision



The Discovery Vision



The Discovery Vision



Take Away Message

- Academics and Industrials agree : Fog/Edge Computing is the new trend for delivering Cloud Computing resources
  ⟹ **Fog/Edge computing is coming,**

- Do not reinvent the wheel and take advantage of existing services
  ⟹ **OpenStack to support Massively Distributed Clouds (public/private)**

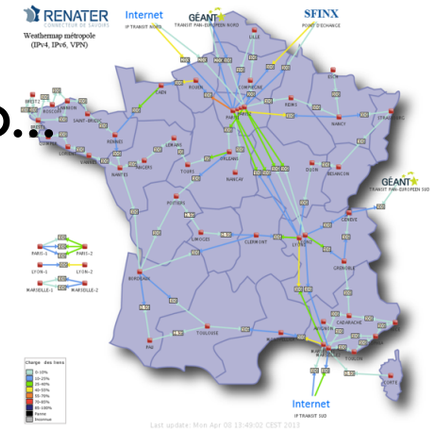
- Several companies/institutes expressed their interests w-r-t the Discovery objectives (Orange, Thales, EU NRENs, ...)
  ⟹ **Creation of a massively distributed clouds WG**
  https://wiki.openstack.org/wiki/Massively_Distributed_Clouds

Changing mentalities/structures takes time !

Beyond Discovery !

- From sustainable data centers to a new source of energy

  A promising way to deliver highly efficient and sustainable UC services is to provide UC platforms as close as possible to the end-users and to...



- Leverage "green" energy (solar, wind turbines...)

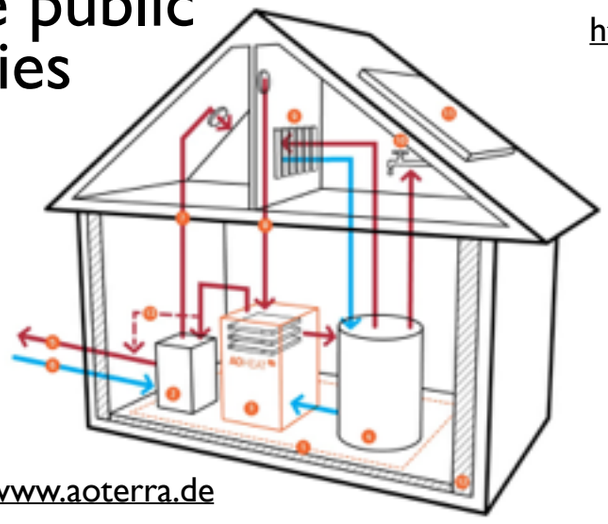
  Transfer the green micro/nano DCs concept to the network PoP

  Take the advantage of the geographical distribution

http://parasol.cs.rutgers.edu

- Leveraging the data furnaces concept

  Deploy UC servers in medium and large institutions and use them as sources of heat inside public buildings such as hospitals or universities



https://www.aoterra.de

Beyond the Cloud, the DISCOVERY Initiative

Localized or ***micro data centers are a fact of life, but by applying a self-contained, scalable and remotely managed solution and process, CIOs can reduce costs, improve agility, and introduce new levels of compliance and service continuity****. Creating micro data centers is something companies have done for years, but often in an ad hoc manner.*

Gartner 2015

Delivering such a system is the objective of Discovery
Thanks - Questions ?

Sagrada Familia microDC (Barcelona, Spain)

Deployment of a new PoP of the Orange French

Additional slides

Just a little bit more…

DISCOVERY - Long term roadmap

- A lot of scientific/technical challenges: making Openstack Fog/Edge computing compliant is not only related to scalability / distribution

Bottom / Up approach

- Shared services [step 1] :
  - Storage backend for Fog/Edge computing (KVS like)
  - Communication layer (scalability, inter-site control,…)
- Compute [step 2a] : Locality at every level (API, scheduler, …)
- Network [step 2b] : Revision of Neutron internals
  (KVS but also SDN functions).
- Storage [step 2c] :
  - S3-like service for Fog/Edge (SWIFT / RADOS under high latency ?)
  - Multi sites VM Image Management (replication/prefetching mechanisms
- Enhanced API [step 3]
  - user authentication, quota management

DISCOVERY - Long term roadmap

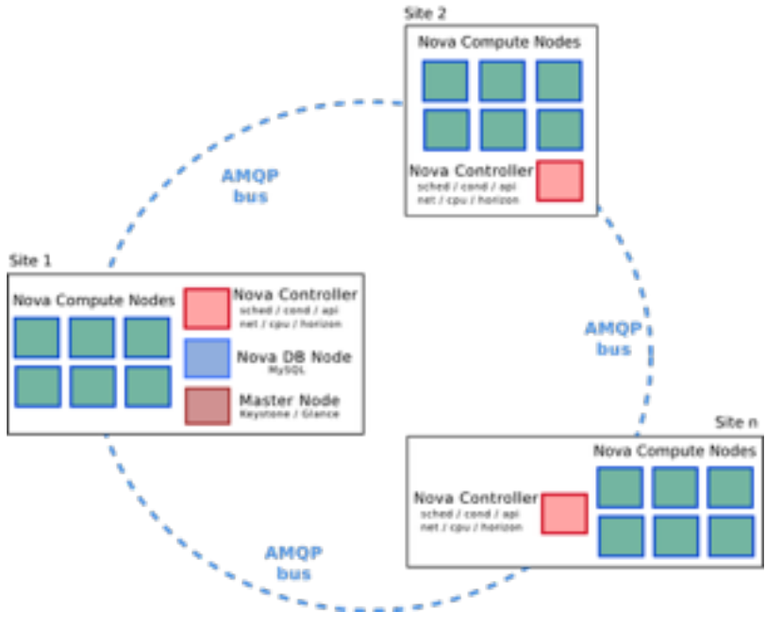
…And Beyond

- Deployment / reconfiguration at each new release/ upgrade throughout the whole infrastructure.

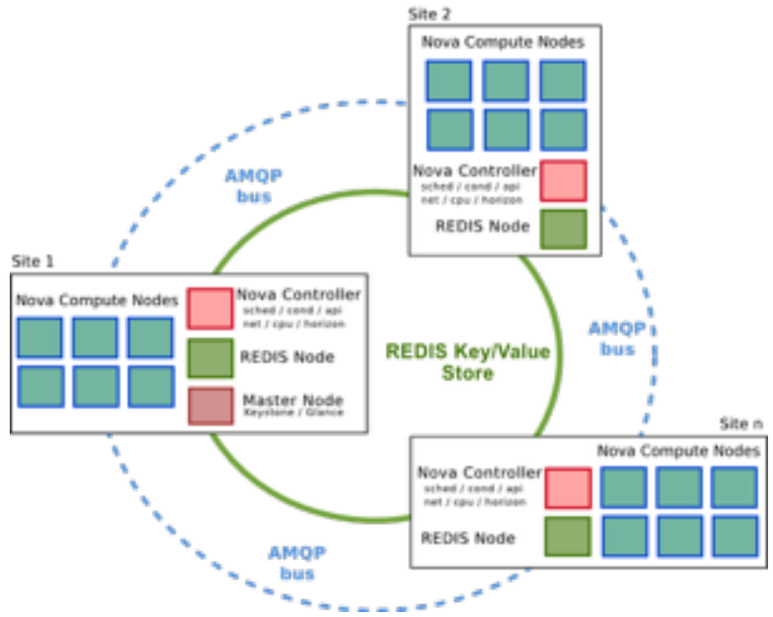
⟹ Makes Openstack vanilla able to support Fog/Edge Cloud use case

Multi-site Experiments

- Creation of 500 VMs, fairly distributed on each controller
- From 2 to 8 sites (emulation of virtual clusters by adding latency thanks to TC)
- Each cluster was containing 1 controller, 6 compute nodes
  (and 1 dedicated node in the case of REDIS).
- MySQL and Redis used in the default configuration
- To fairly compare with MySQL, data replication was not activated in Redis
- Galera experiments have been performed but due to reproducible issues with more than 4 sites, results are not satisfactory enough to be discussed (RR available on demand)



SQLAlchemy+MySQL



ROME+Redis

Multi-Site Experiments

Table 3: Time used to create 500 VMs with a 10ms inter-site latency (in sec.).

Nb of locations	REDIS	MySQL
2 clusters	271	209
4 clusters	263	139
6 clusters	229	123
8 clusters	223	422

Table 4: Time used to create 500 VMs inter-site latency (in sec.).

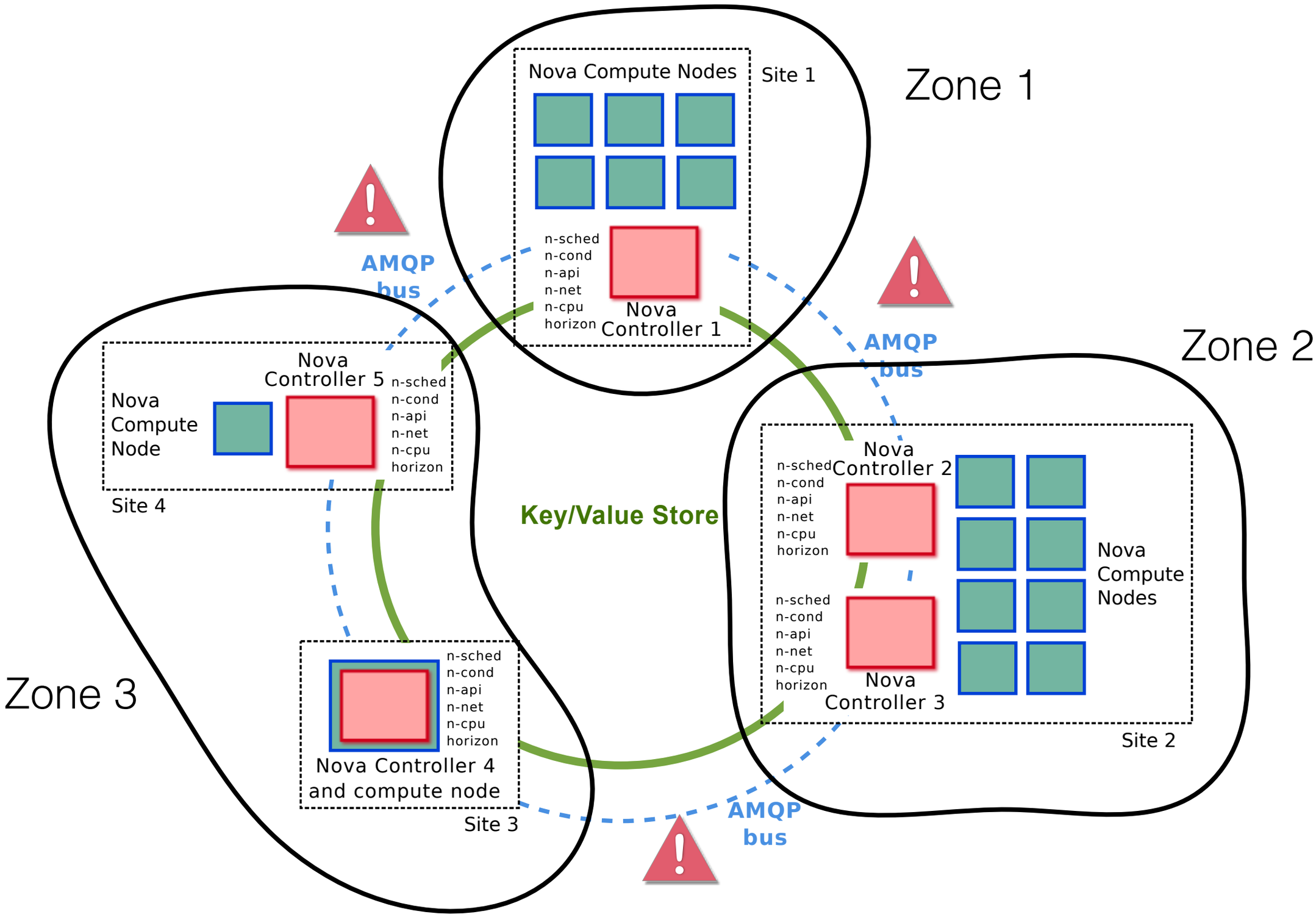
Nb of locations	REDIS	MySQL
2 clusters	723	268
4 clusters	427	203
6 clusters	341	184
8 clusters	302	759

SQL scalability bottleneck

(one SQL server for the whole infrastructure)

Increasing the nb of nodes leads to better reactivity
From 8 clusters, MySQL becomes a bottleneck

But Locality matters !



Revising OpenStack

OPENSTACK COMMUNITY: BROAD SUPPORT AND CONTRIBUTION

openstack
CLOUD SOFTWARE

FOUNDATION STARTED IN SEPTEMBER 2012

ECOSYSTEM SIZE
(Members + Sponsors + Supporters)

483
Contributing Companies

Top 5 Contributors
Rackspace, Red Hat, HP, IBM, Mirantis

INDIVIDUAL MEMBERS

21,230

CUMULATIVE CONTRIBUTORS	AVERAGE MONTHLY CONTRIBUTORS	CODE CONTRIBUTIONS	LINES OF CODE
2,424	>400	110,019	20M+

rackspace
the #1 managed cloud company

www.rackspace.com 6

Revising OpenStack

OPENSTACK COMMUNITY: BROAD SUPPORT AND CONTRIBUTION

openstack
CLOUD SOFTWARE

FOUNDATION STARTED IN SEPTEMBER 2012

ECOSYSTEM SIZE
(Members + Sponsors + Supporters)

483
Contributing Companies

Top 5 Contributors
Rackspace, Red Hat, HP, IBM, Mirantis

INDIVIDUAL MEMBERS

21,230

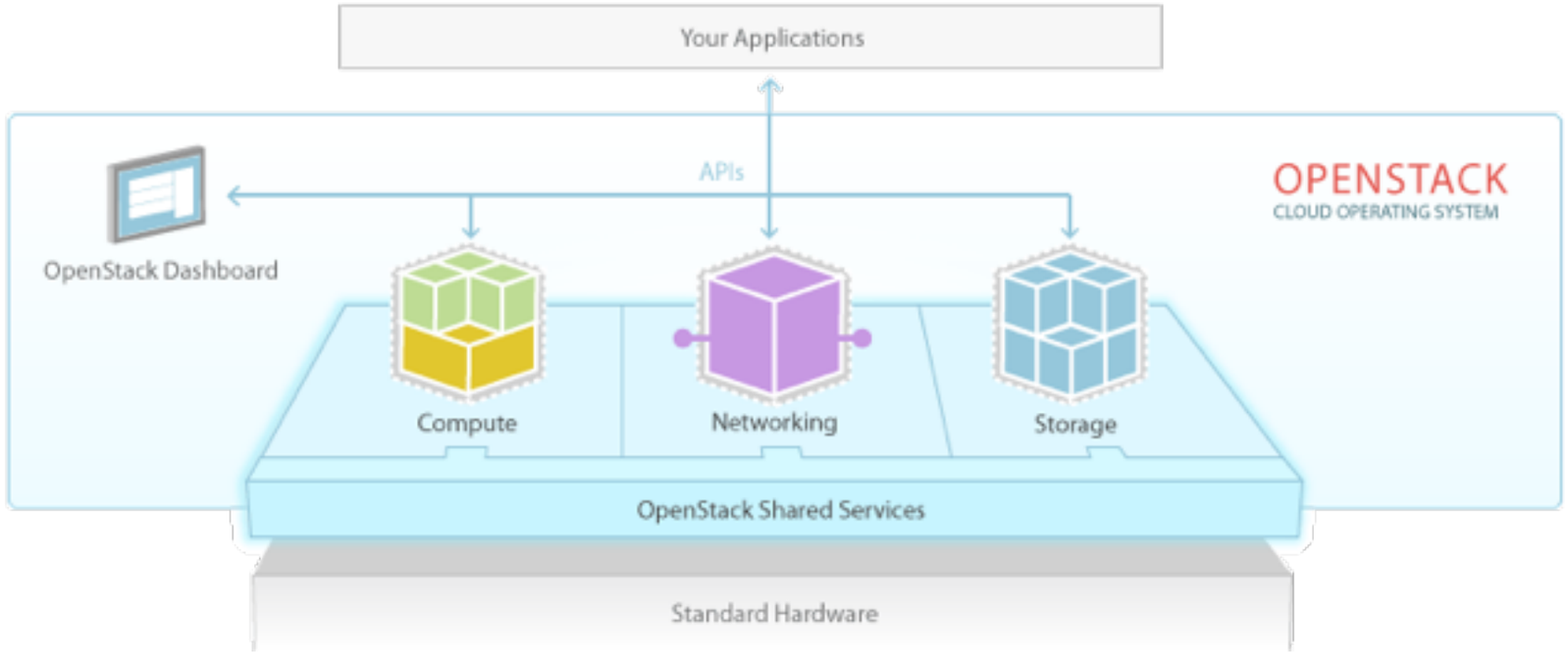
CUMULATIVE CONTRIBUTORS	AVERAGE MONTHLY CONTRIBUTORS	CODE CONTRIBUTIONS	LINES OF CODE
2,424	>400	110,019	20M+

rackspace
the #1 managed cloud company

www.rackspace.com 6

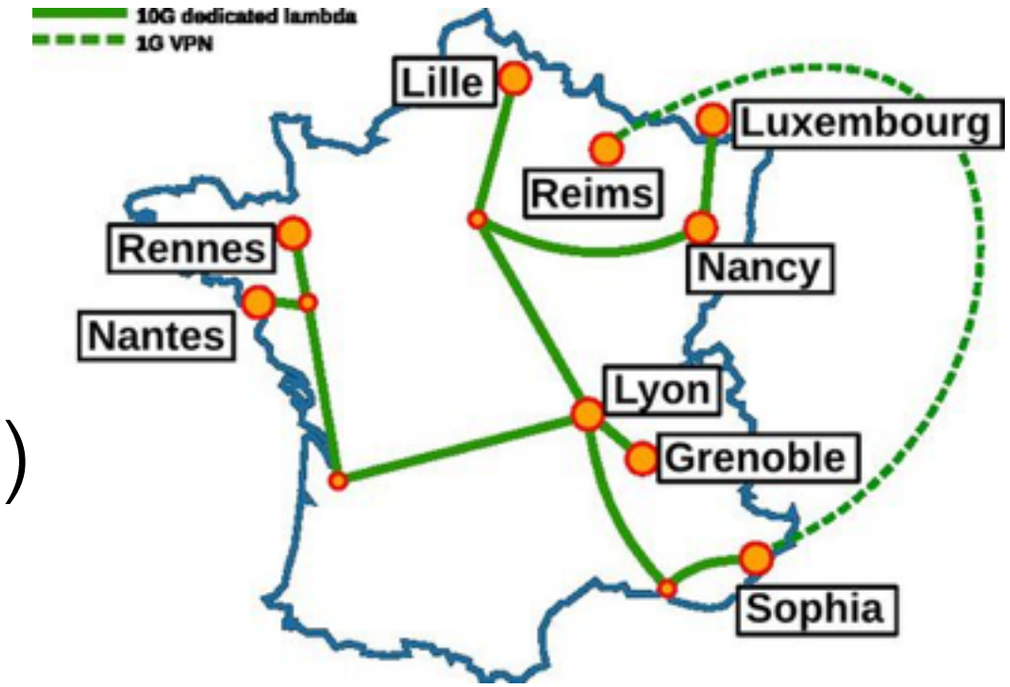
2Millions of LOCs just for core-services

OpenStack Ecosystem



Understanding OpenStack

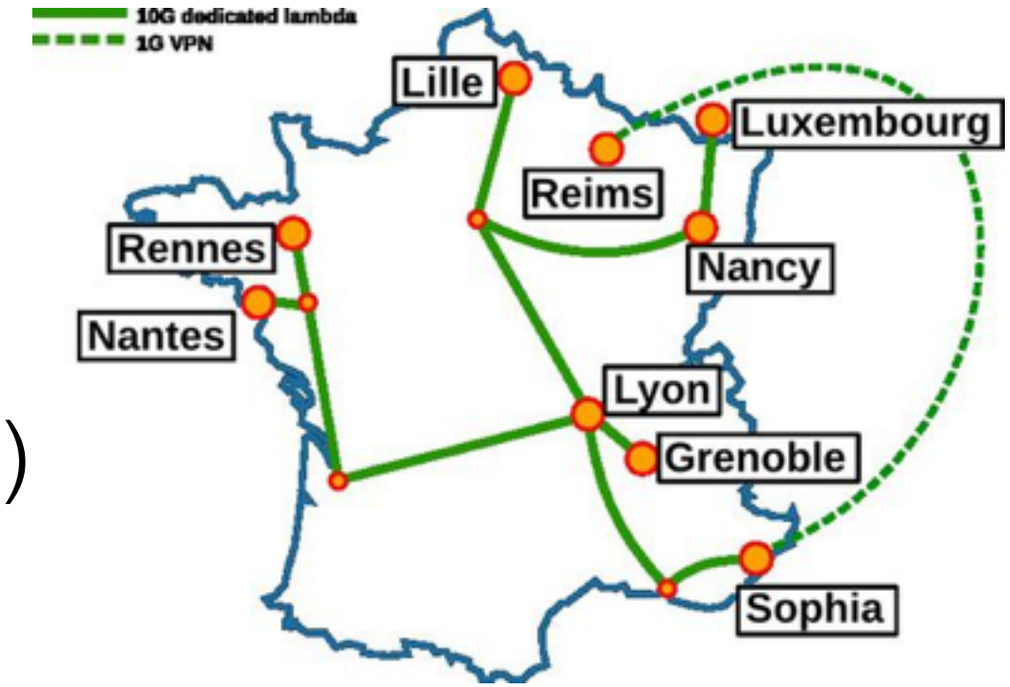
- Docs/white papers
- Performance evaluations (Kolla-G5K: Rally + Grid5000)
- OS profiler



Cross-project profiling library to generate 1 trace per request, that goes through all involved services.

Understanding OpenStack

- Docs/white papers
- Performance evaluations (Kolla-G5K: Rally + Grid5000)
- OS profiler



Cross-project profiling library to generate 1 trace per request, that goes through all involved services.

OS Profiler to understand OpenStack performance

Discovery Task forces

- Today: provided by Orange and Inria (with the support of RENATER)
- In addition to permanent staffs

7 Phds

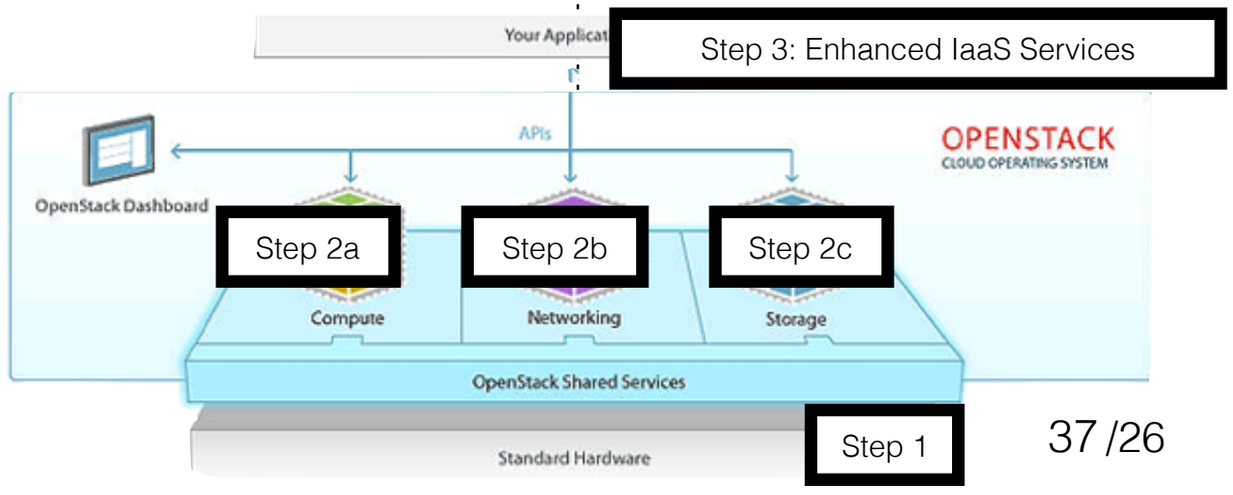
- Locality based Overlay networks - step 1
- Monitoring - step 1
- Security enforcement - step 1
- VM scheduling policies - step 2a
- Distributed SDN capabilities - step 2b
- Image management - step 2c
- Locality from the application elasticity view point - step 3

6 post docs

- Cost benefit analysis and energy opportunity - general
- Identification of Neutron challenges - step 2b
- Deployment/Reconfiguration of OpenStack - general
- VM placement strategies - step 2c
- Data scheduling policies - step 2a/2c
- Use-cases / validations - step 3

3 engineers

- Core developper (soon !)
- Sys Admin
- GUI/command line developper



Discovery Task forces

- Today: provided by Orange and Inria (with the support of RENATER)
- In addition to permanent staffs

7 Phds

- Locality based Overlay networks - step 1
- Monitoring - step 1
- Security enforcement - step 1
- VM scheduling p… step 2a
- D…
- I… …agement - st… 2c
- Locality from the application elasticity view point - step 3

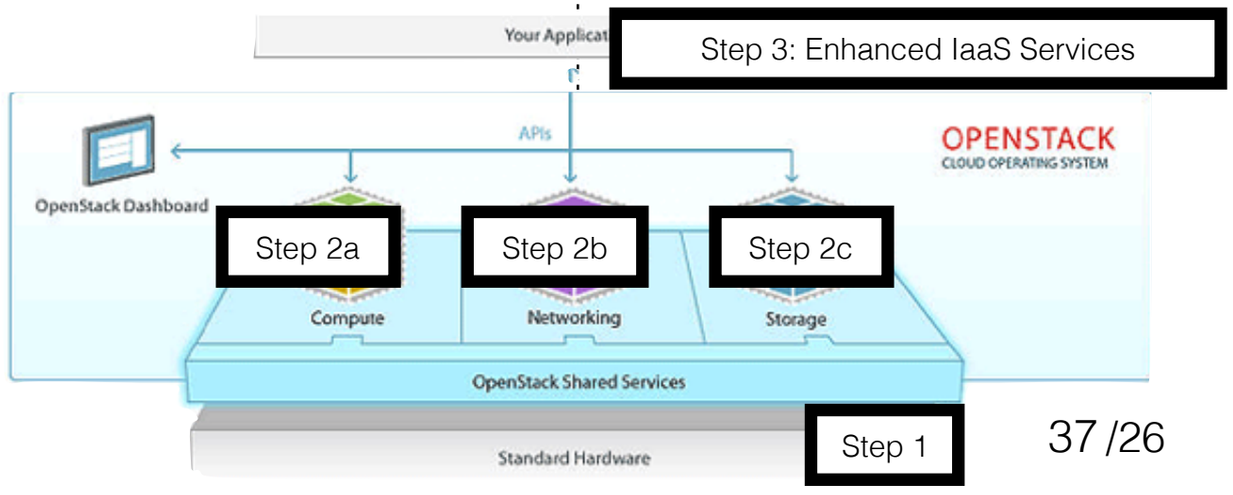
6 post docs

- Cost benefit analysis and ener… opportunity - general
- Identification of … step 2b
- D… …of
- …nt strategies - step 2c
- …a scheduling policies - step 2a/2c
- Use-cases / validations - step 3

…pper …soon !)

- Sys Admin
- GUI/command line developper

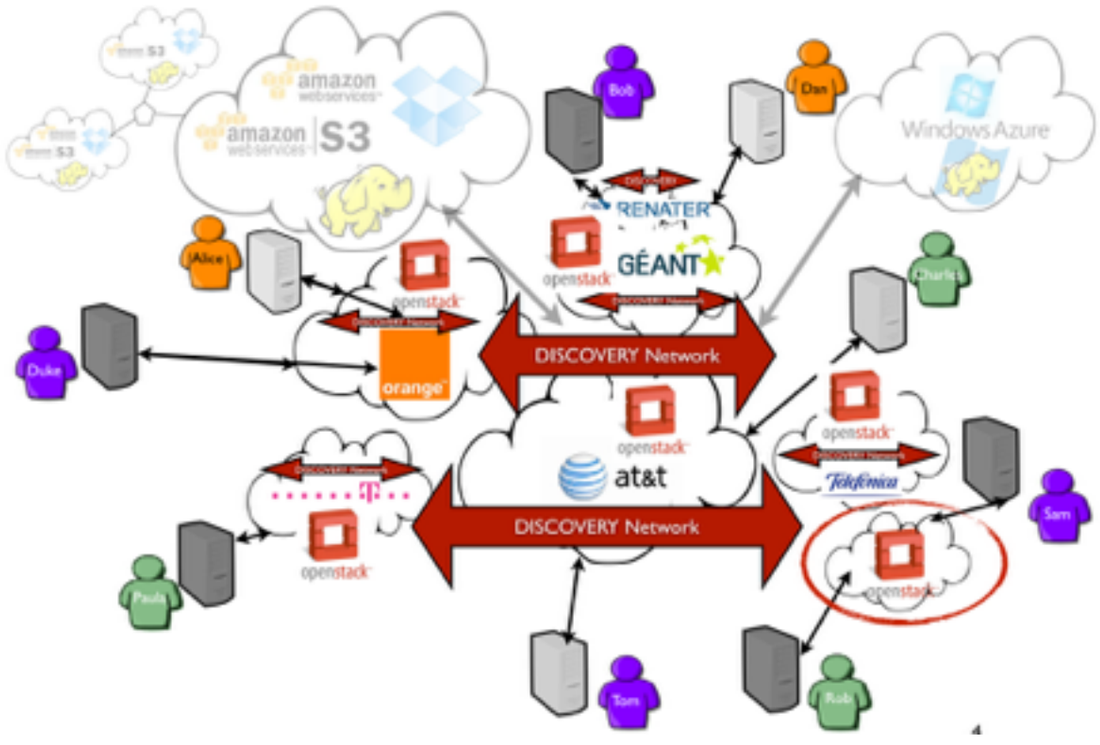
See Open Positions on the Discovery website. or just come and contribute ;)



The DISCOVERY Initiative

- Several researchers, engineers, stakeholders of important EU institutions and SMEs have been taking part to numerous brainstorming sessions (BSC, CRS4, Unine, EPFL, PSNC, Interoute, Orange Labs, Peerialism, TBS Group, XLAB, …)

http://beyondtheclouds.github.io/



RENATER
CONNECTEUR DE SAVOIRS

Inria

orange™

discovery-contact@inria.fr