Beyond the cloud !
a "small" overview of cloud challenges

Adrien Lebre / Ascola Project Team
Cumulo NumBio - June 3rd, 2015

Looking back…

- xxx Computing
  Meta / Cluster / Grid / Desktop / "Hive" / Cloud / Sky …

  ⇒ xxx as Utility Computing

- A common objective: provide computing resources (both hardware and software) in a flexible, transparent, secure, reliable, ... way

- Challenges

  Software/Hardware heterogeneity

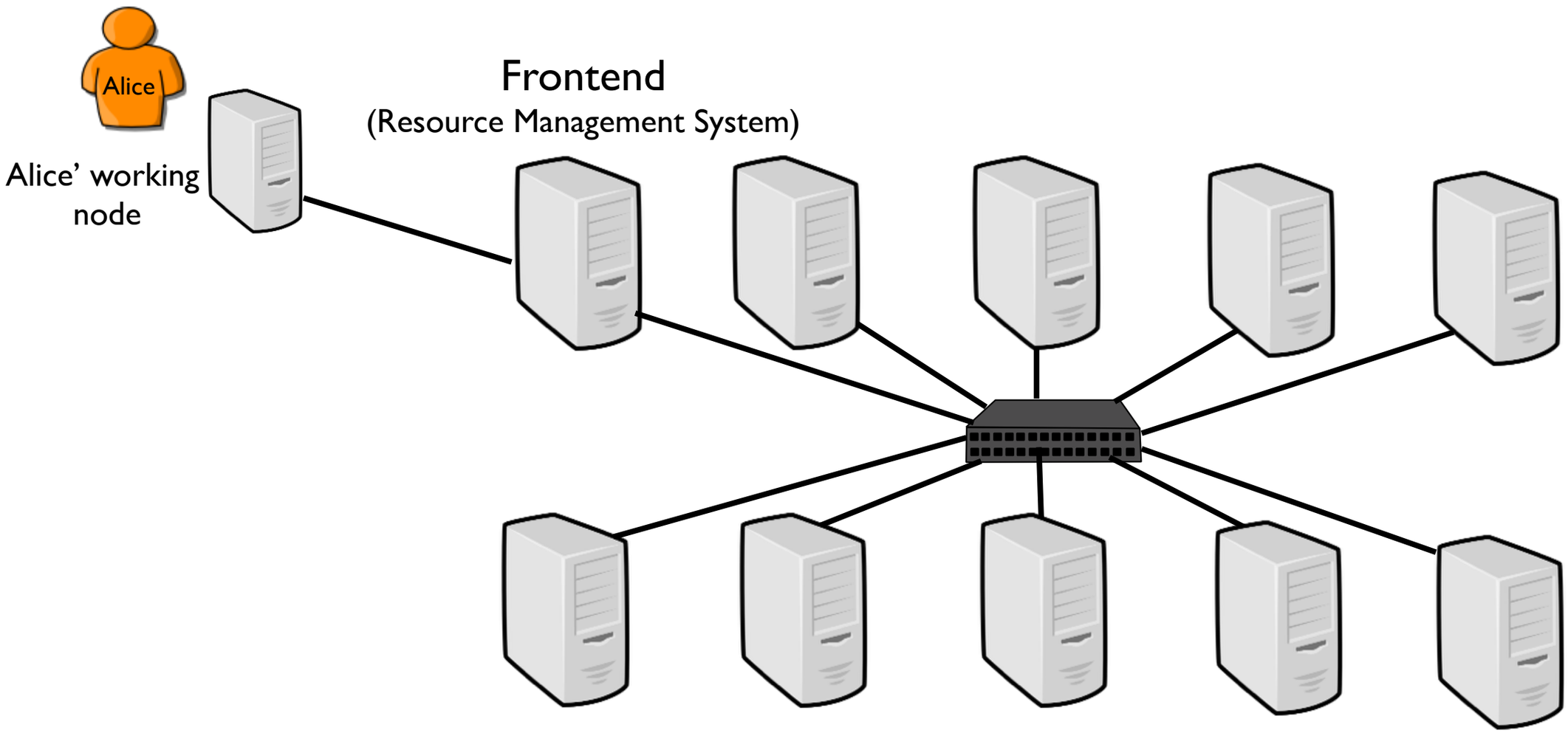
  Security (Isolation between applications, ...)

  Reliability / Resiliency

  Data Sharing ...

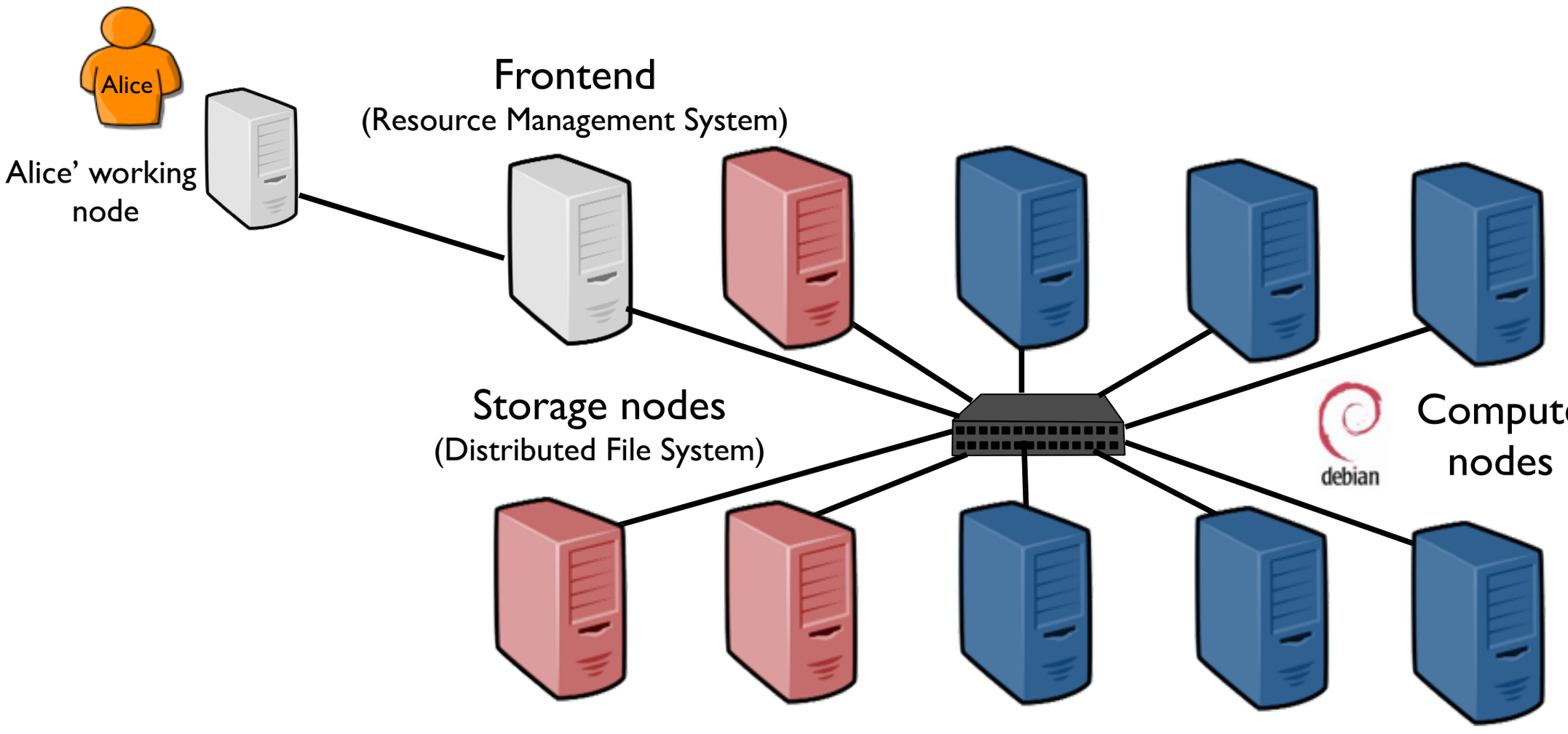
Looking back…

- Network of Workstations 1990 / 20**xx**



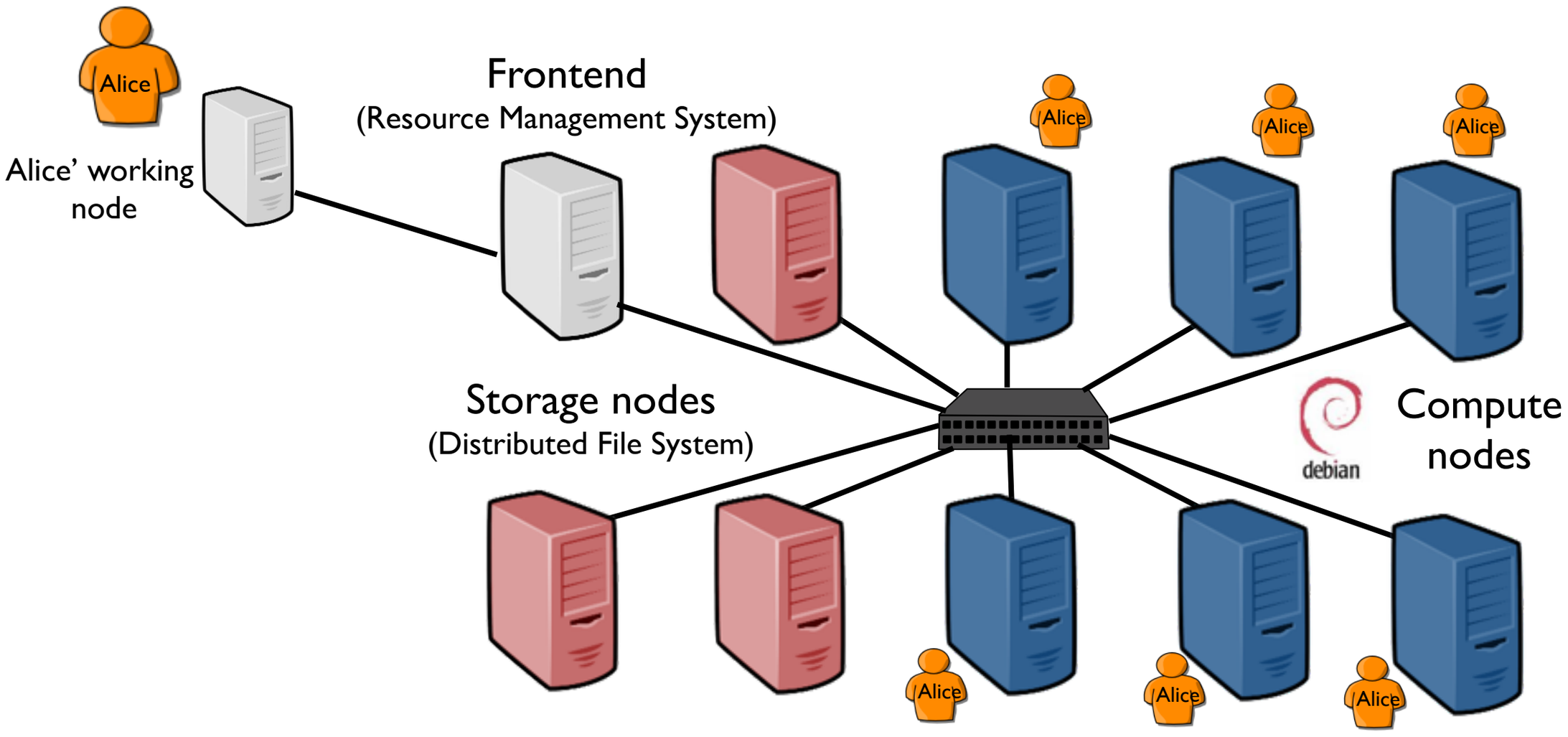
Looking back…

- Network of Workstations 1990 / 20**xx**



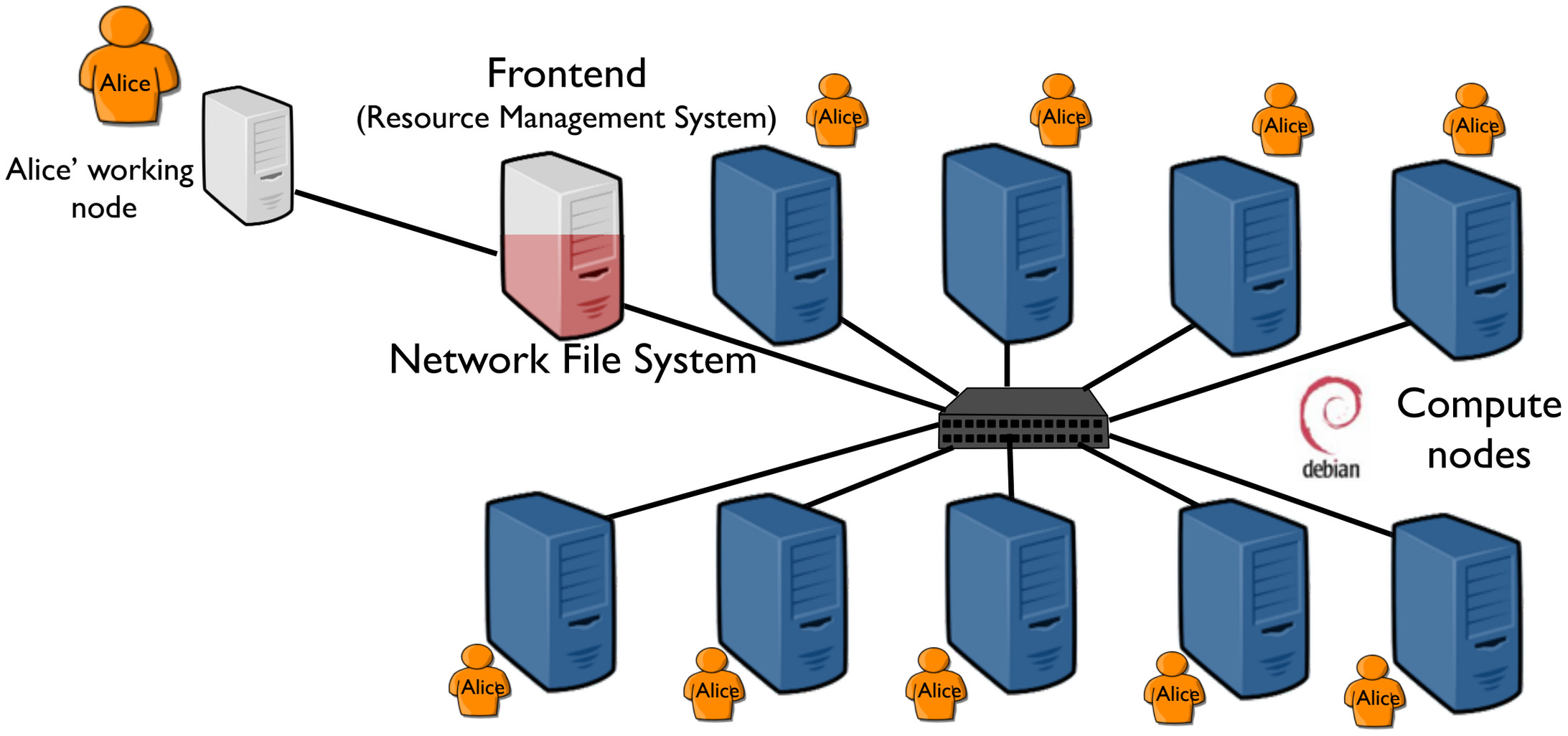
Looking back…

- Network of Workstations 1990 / 20**xx**



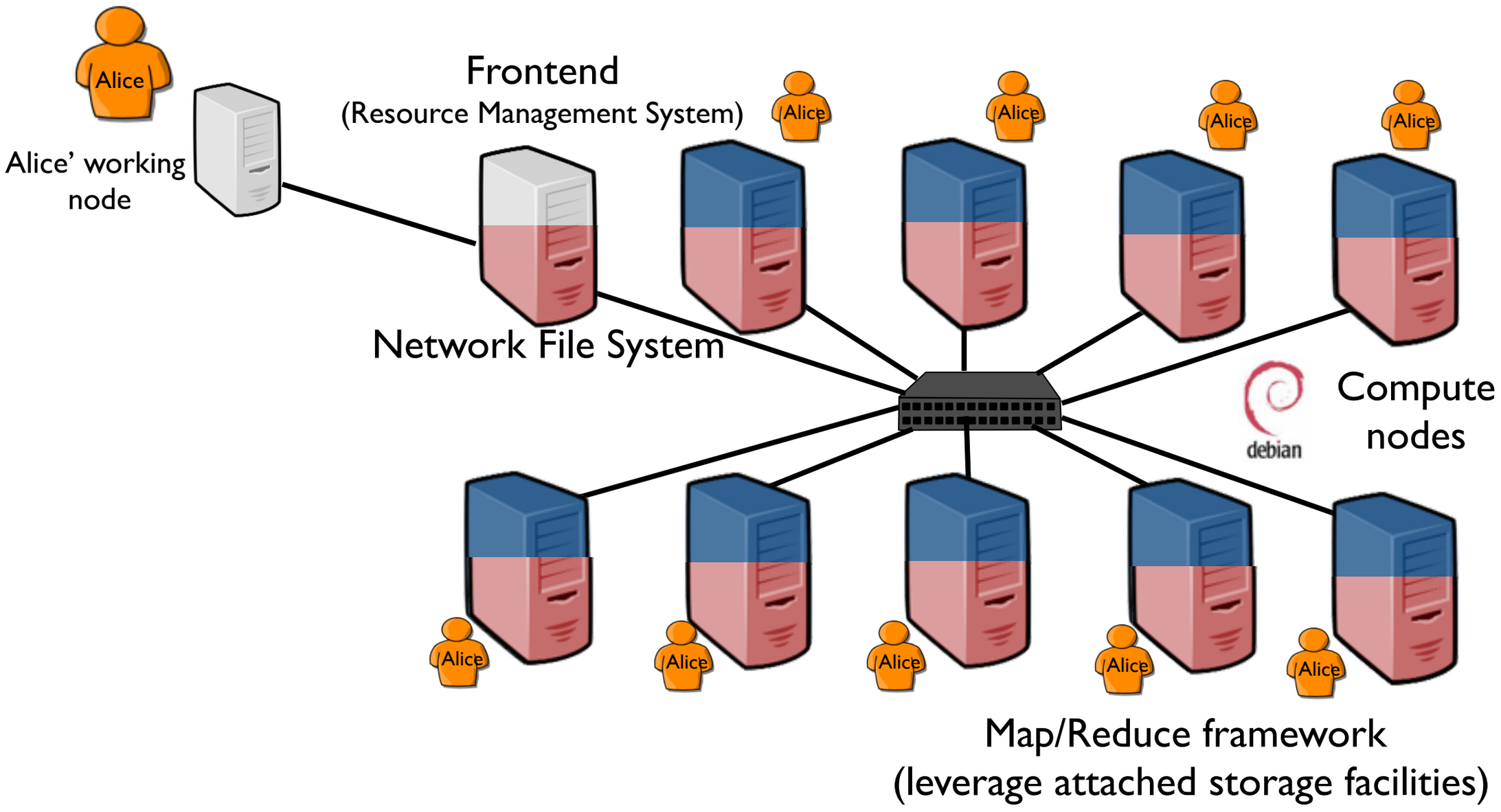
Looking back…

- Network of Workstations 1990 / 20**xx**



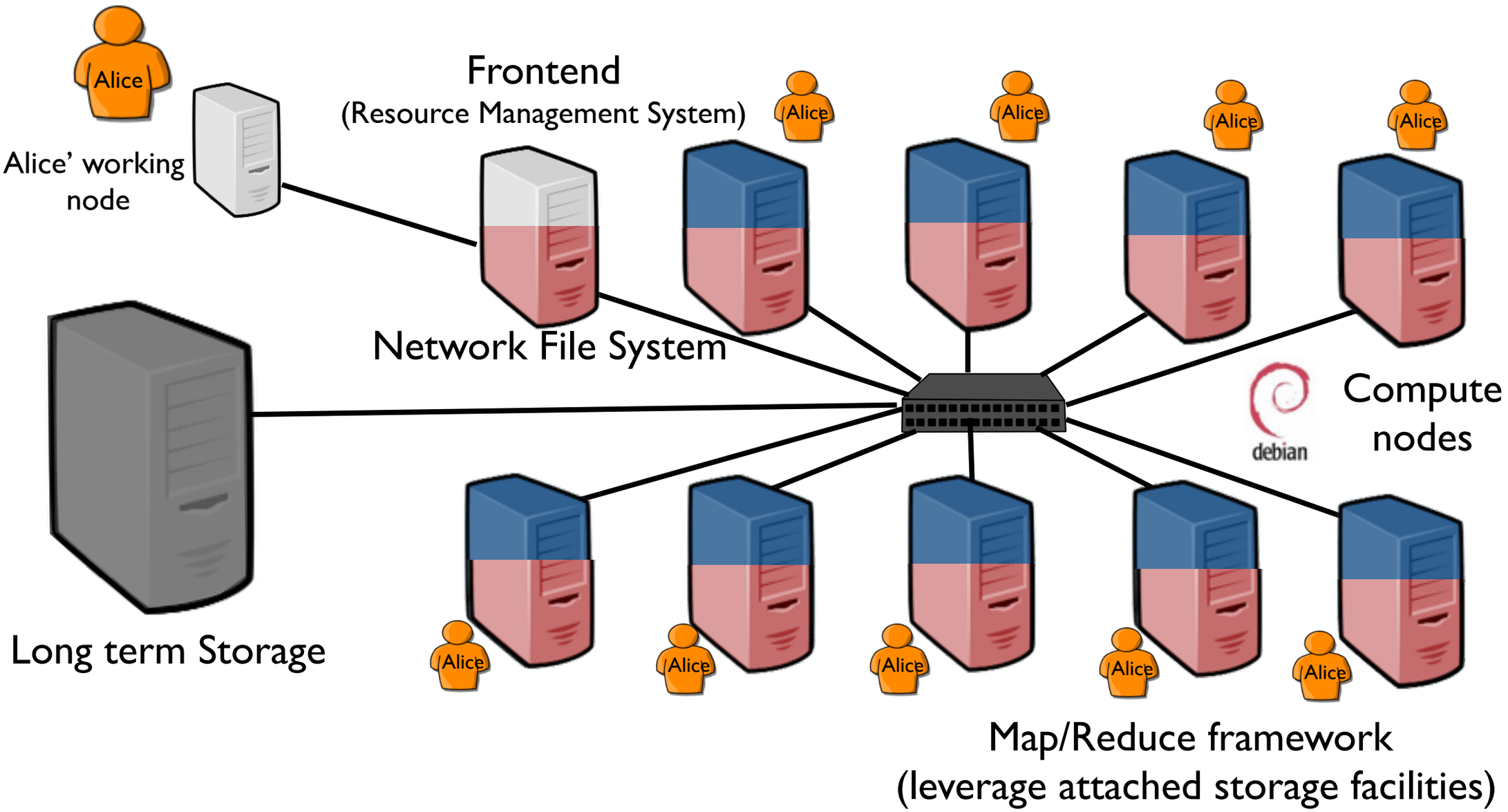
Looking back…

- Network of Workstations 1990 / 20**xx**



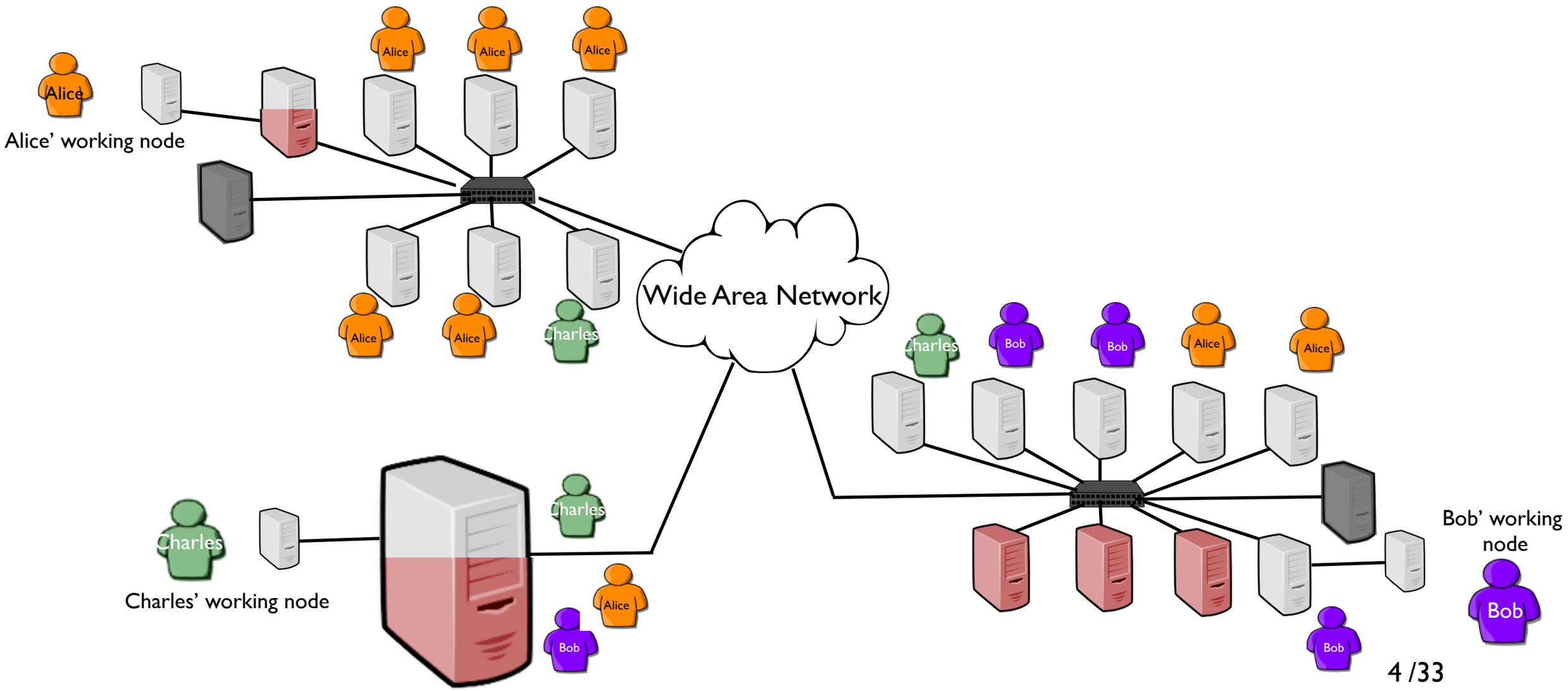
Looking back…

- Network of Workstations 1990 / 20**xx**



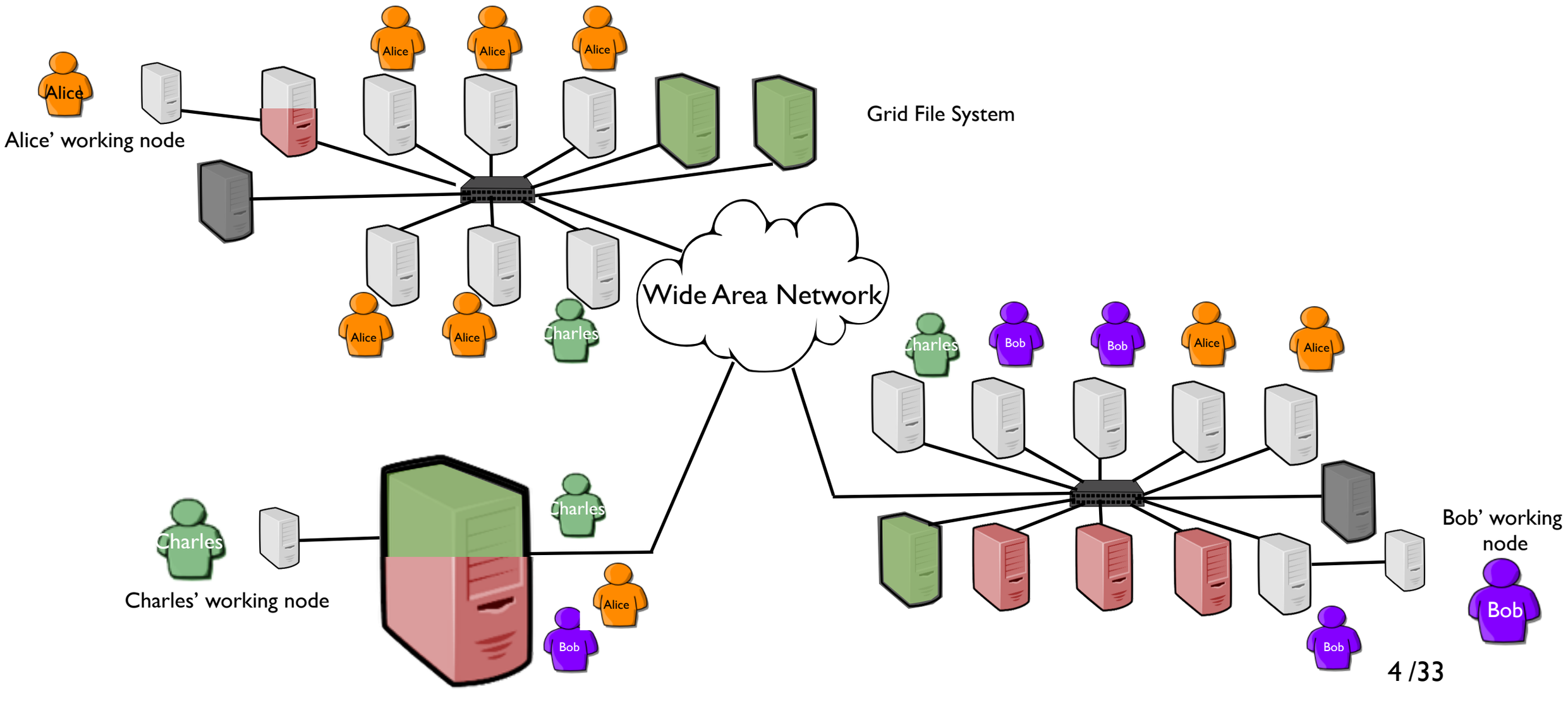
Looking back…

- Network of Workstations 1990 / 20**xx**
- The Grid 1997 / 201**x**



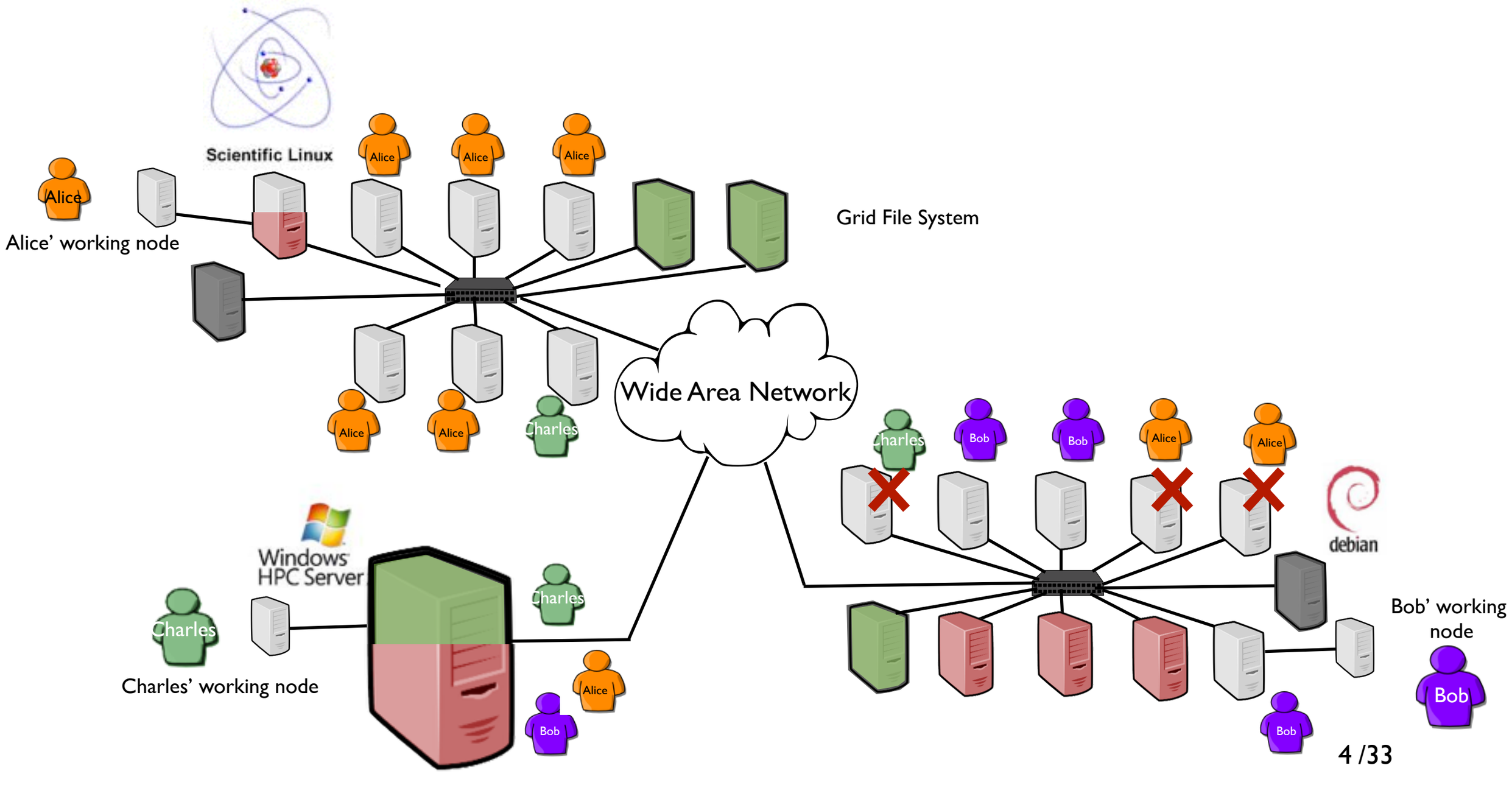
Looking back…

- Network of Workstations 1990 / 20**xx**
- The Grid 1997 / 201**x**



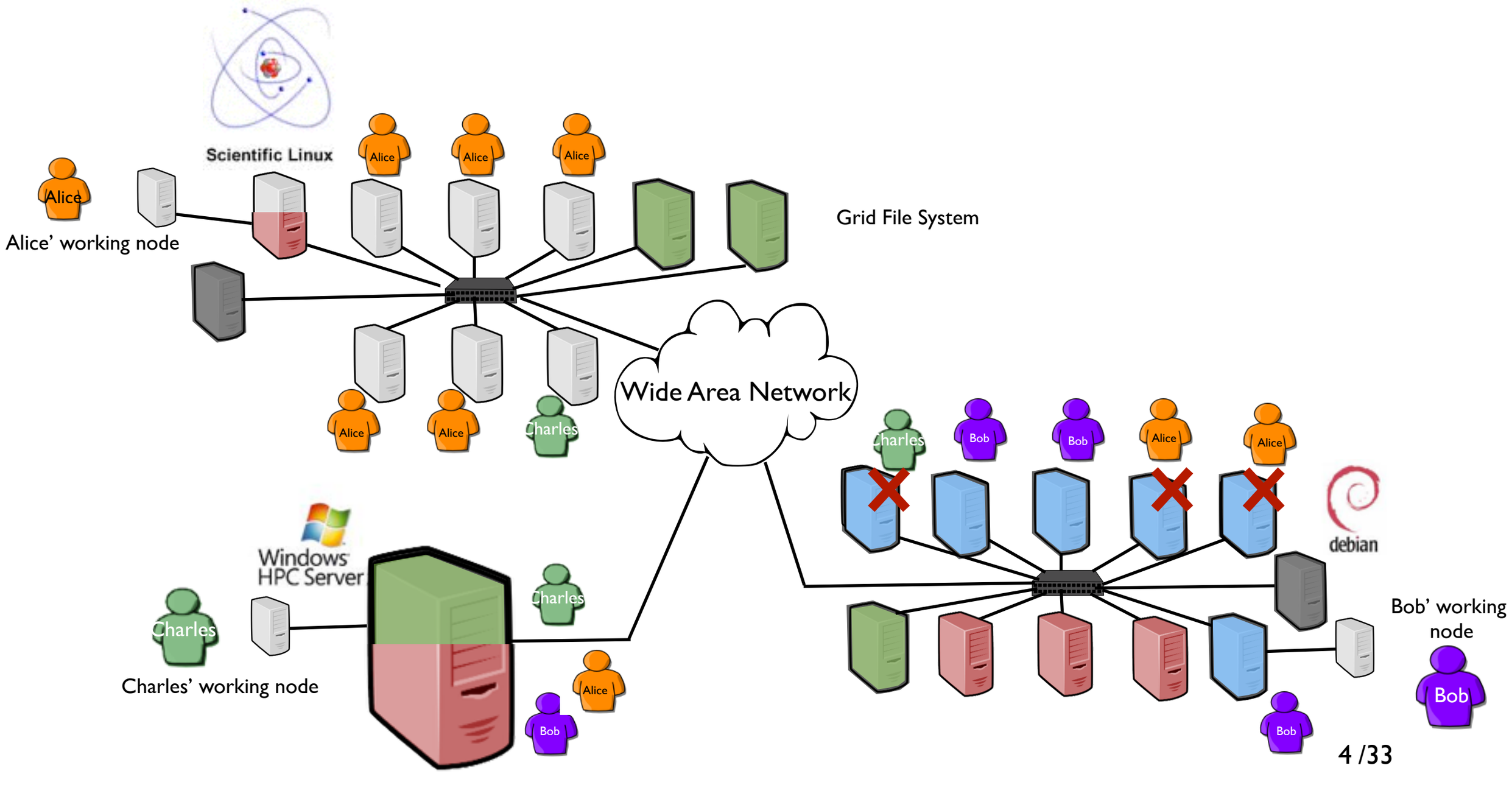
Looking back…

- Network of Workstations 1990 / 20**xx**
- The Grid 1997 / 201**x**



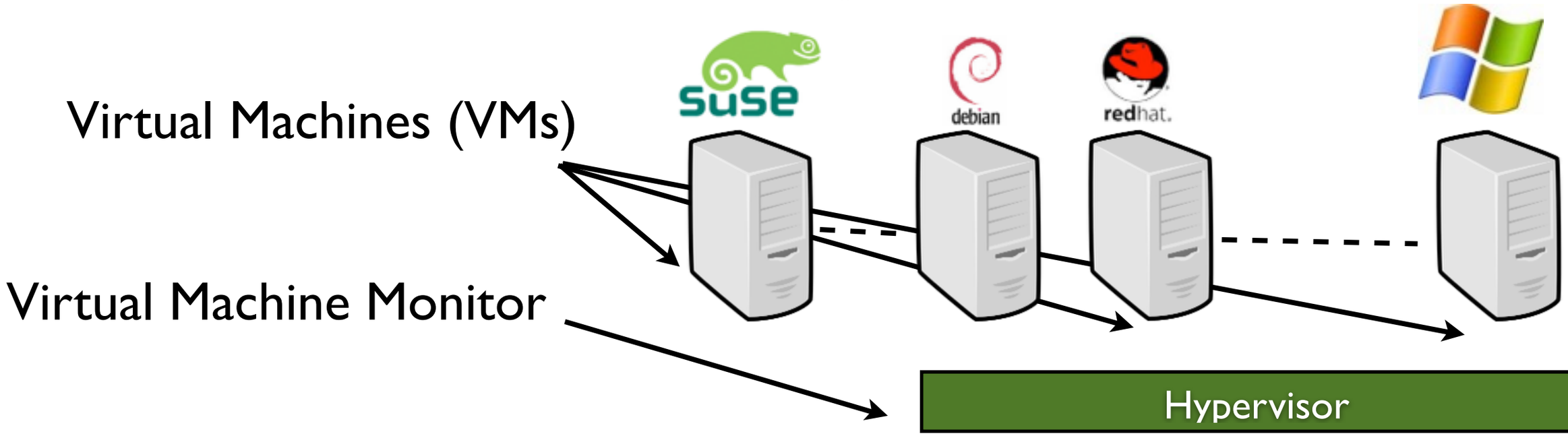
Looking back…

- Network of Workstations 1990 / 20**xx**
- The Grid 1997 / 201**x**



Looking back…

- System virtualization: One to multiple OSes on a physical node thanks to a hypervisor (an operating system of OSes)

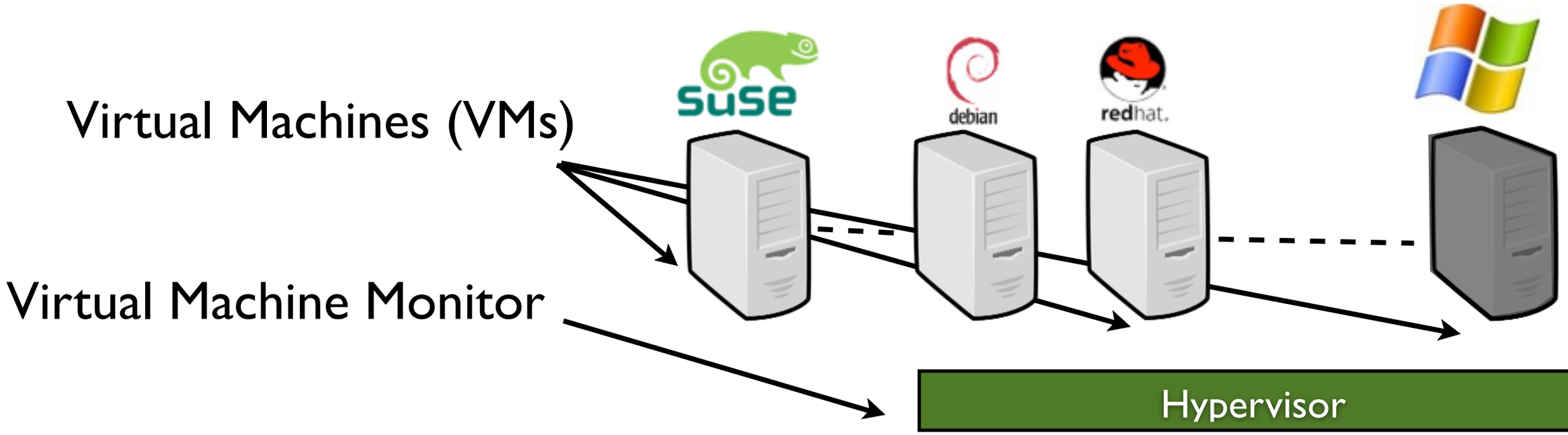


"A ***virtual machine*** (VM) provides a faithful implementation of a physical processor's hardware running in a protected and isolated environment.
Virtual machines are created by a software layer called the ***virtual machine monitor*** (VMM) that runs as a privileged task on a physical processor."

Physical Machine (PM)

Looking back…

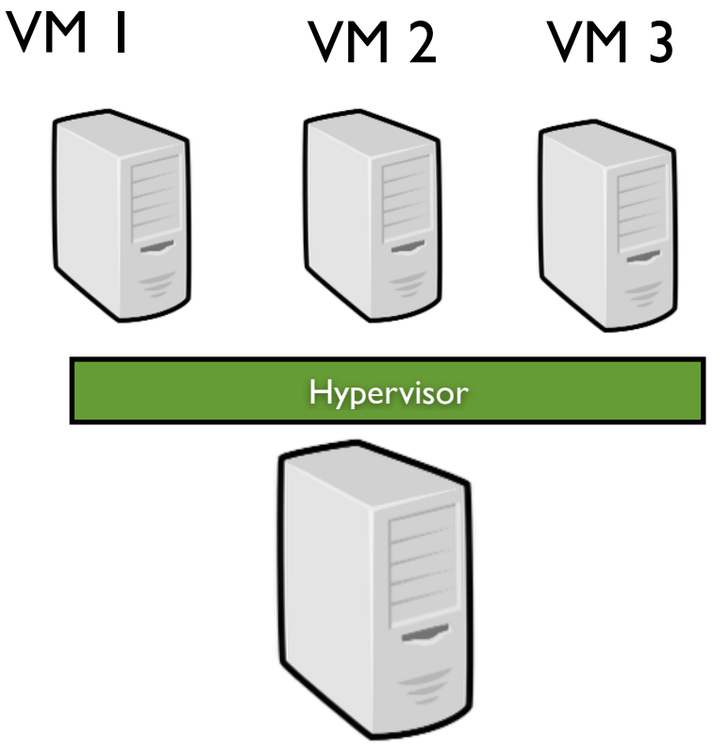
- System virtualization: One to multiple OSes on a physical node thanks to a hypervisor (an operating system of OSes)



"A ***virtual machine*** (VM) provides a faithful implementation of a physical processor's hardware running in a protected and isolated environment.
Virtual machines are created by a software layer called the ***virtual machine monitor*** (VMM) that runs as a privileged task on a physical processor."

Looking back…

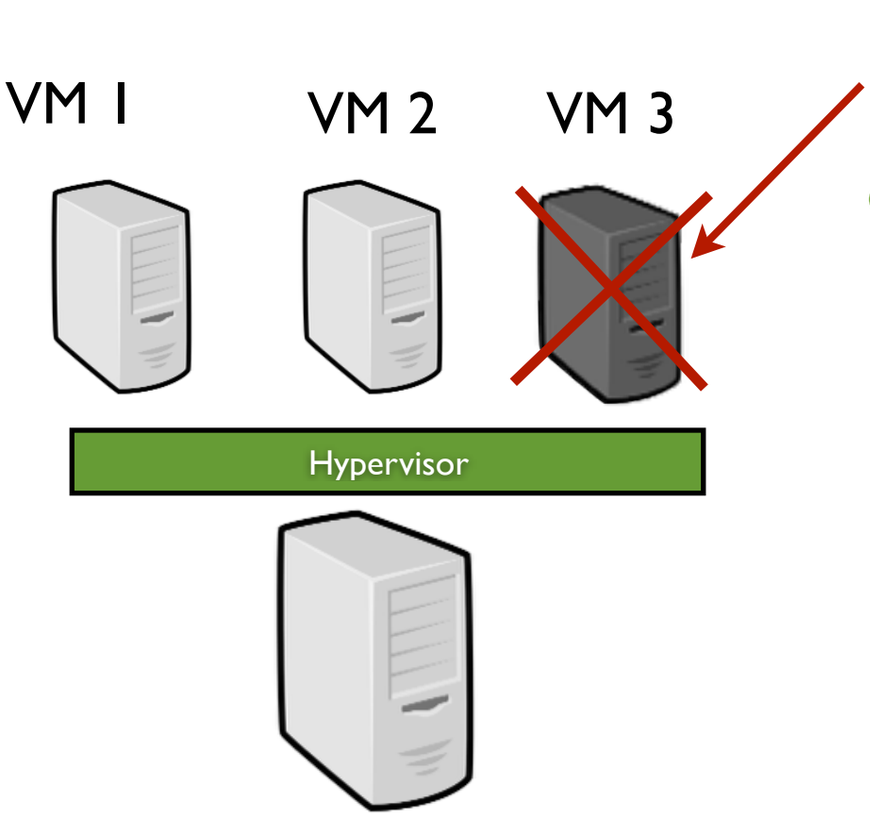
- System virtualization: a great sandbox



- Isolation (“security” between each VM)

Looking back…

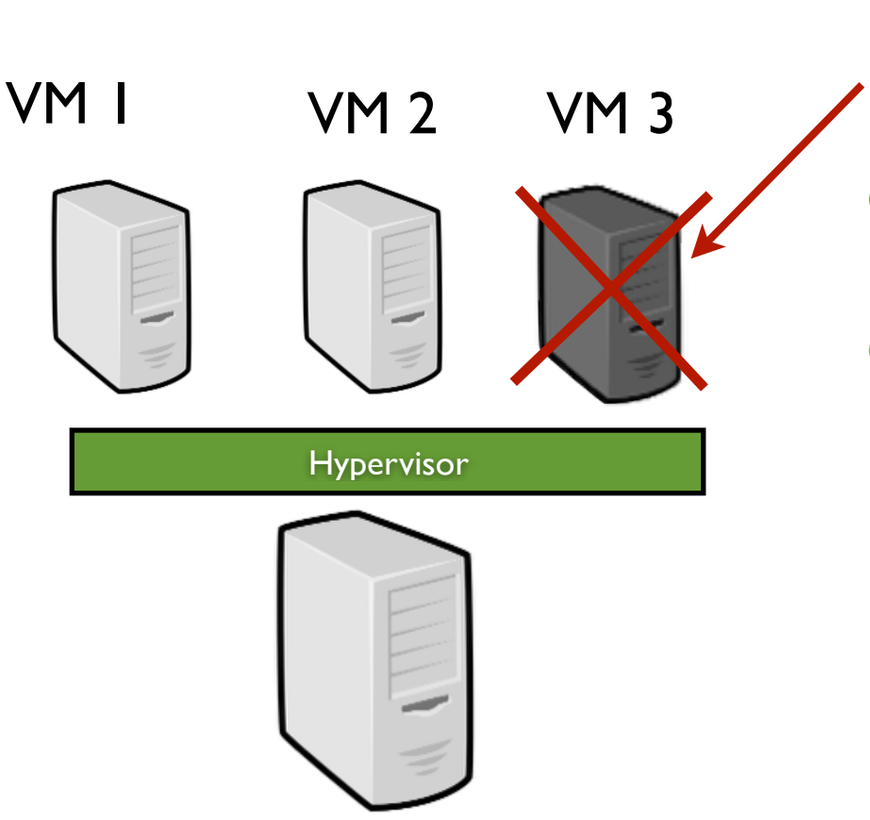
- System virtualization: a great sandbox



- Isolation ("security" between each VM)

Looking back…

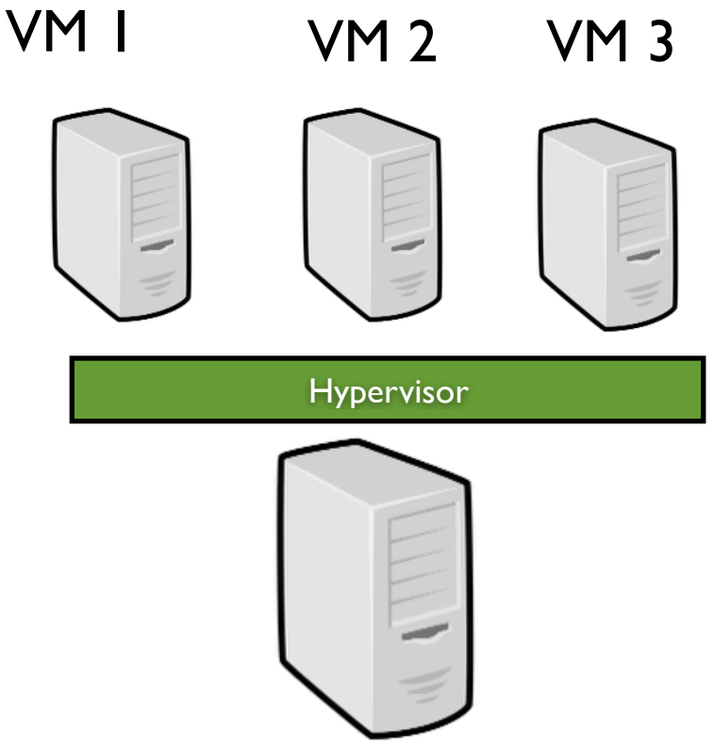
- System virtualization: a great sandbox



- Isolation ("security" between each VM)
- Snapshotting (a VM can be easily resume from its latest consistent state)

Looking back…

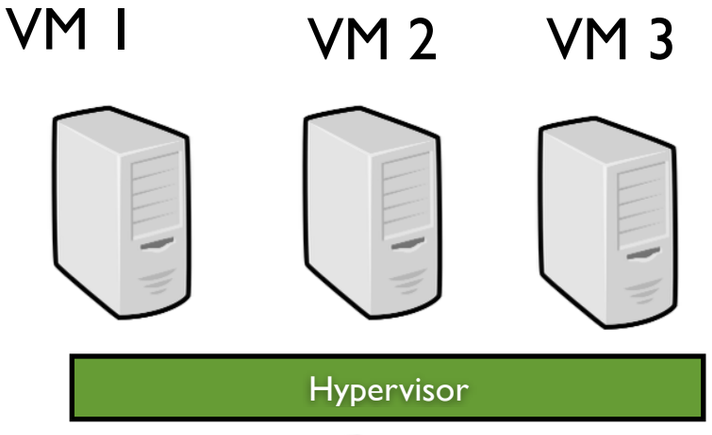
- System virtualization: a great sandbox



- Isolation (“security” between each VM)
- Snapshotting (a VM can be easily resume from its latest consistent state)

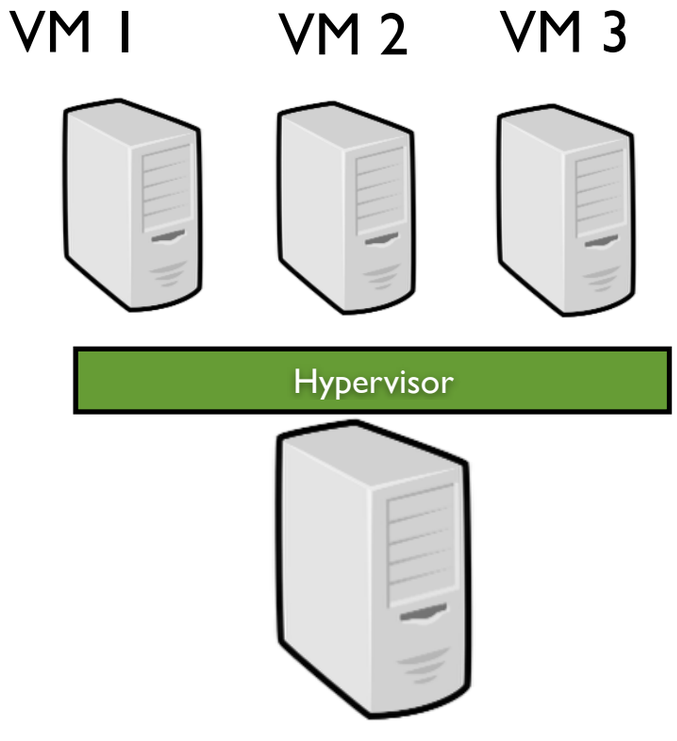
Looking back…

- System virtualization: a great sandbox



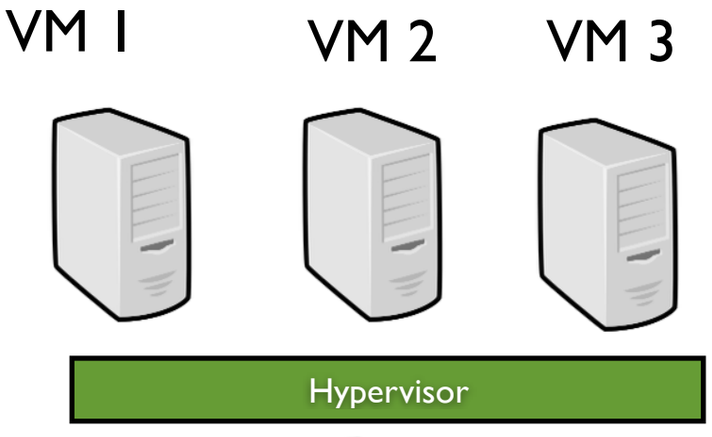
- Isolation (“security” between each VM)
- Snapshotting (a VM can be easily resume from its latest consistent state)

- Suspend/Resume

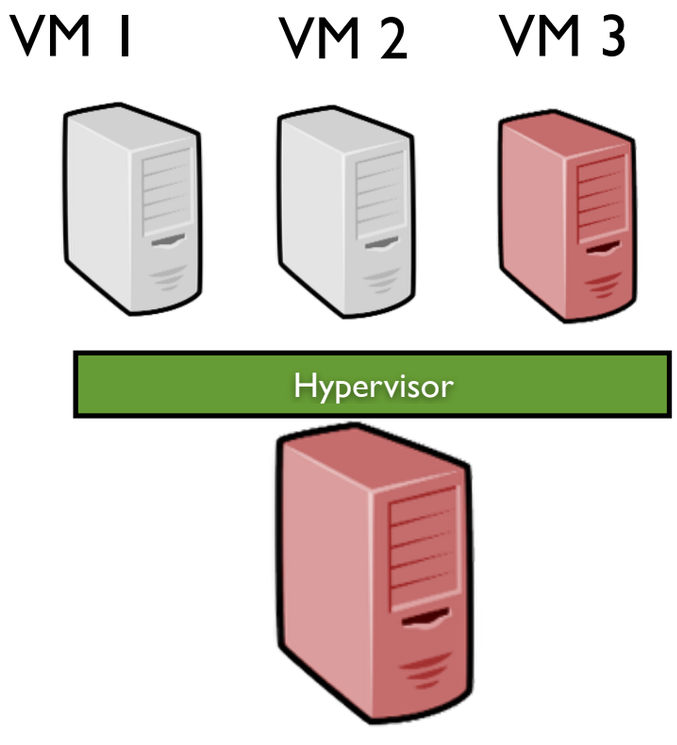


Looking back…

- System virtualization: a great sandbox

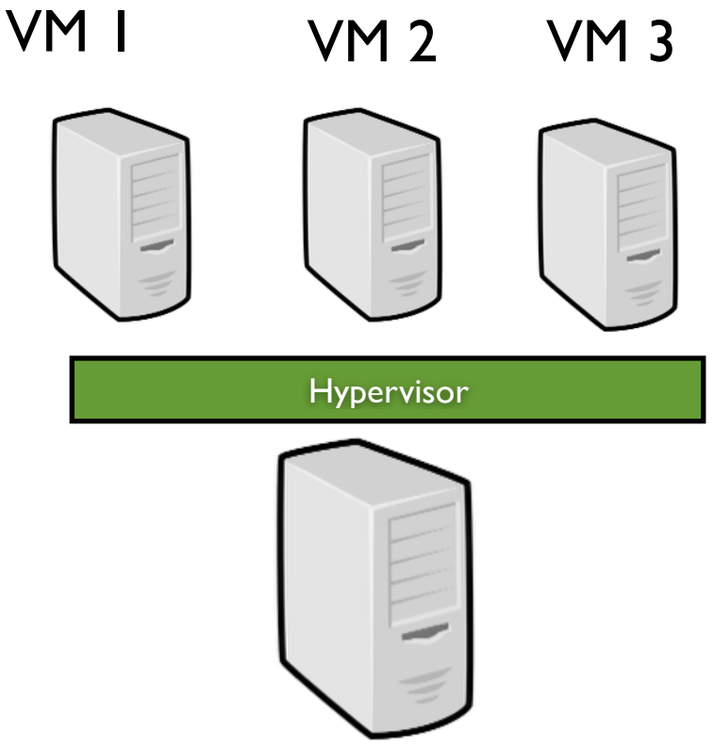


- Isolation (“security” between each VM)
- Snapshotting (a VM can be easily resume from its latest consistent state)
- Suspend/Resume

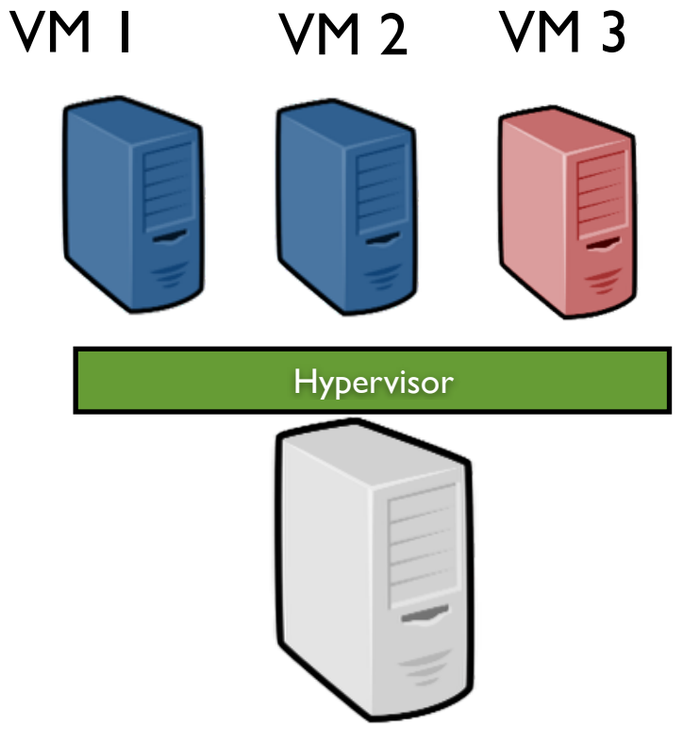


Looking back…

- System virtualization: a great sandbox

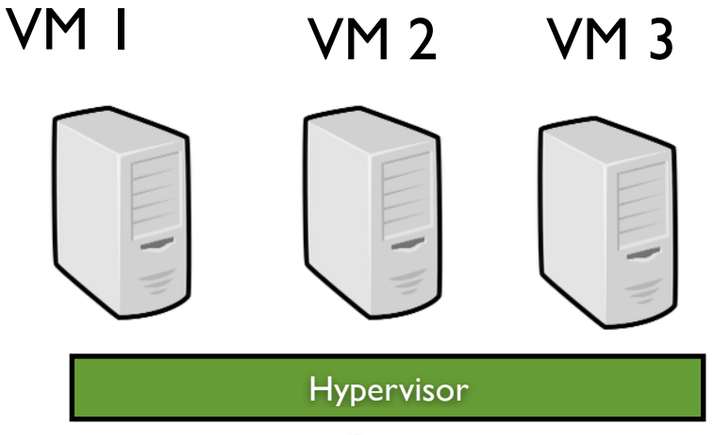


- Isolation (“security” between each VM)
- Snapshotting (a VM can be easily resume from its latest consistent state)
- Suspend/Resume

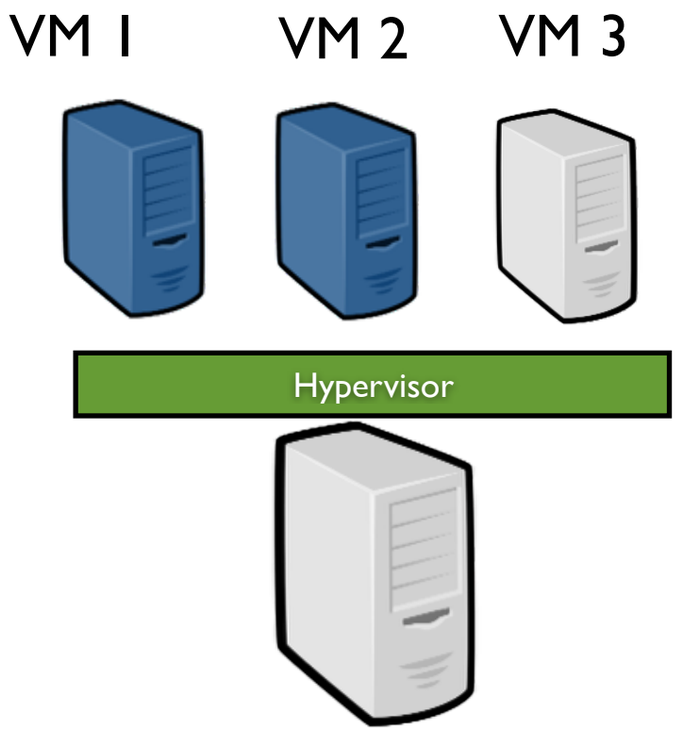


Looking back…

- System virtualization: a great sandbox

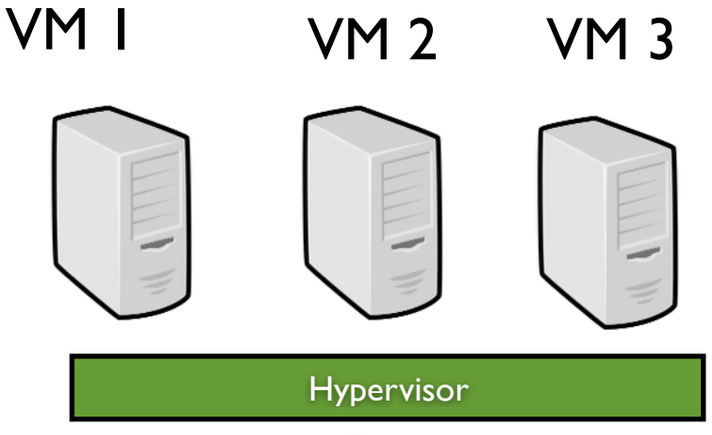


- Isolation ("security" between each VM)
- Snapshotting (a VM can be easily resume from its latest consistent state)
- Suspend/Resume



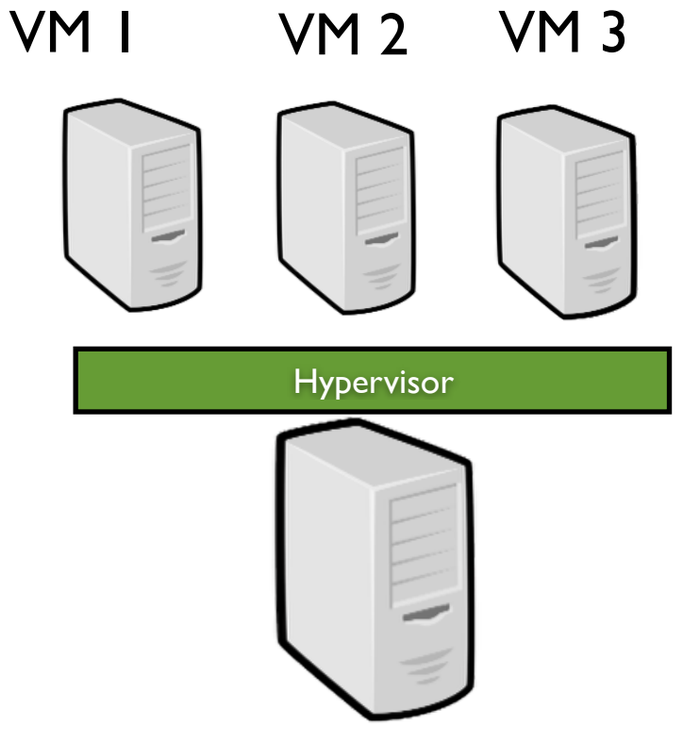
Looking back…

- System virtualization: a great sandbox



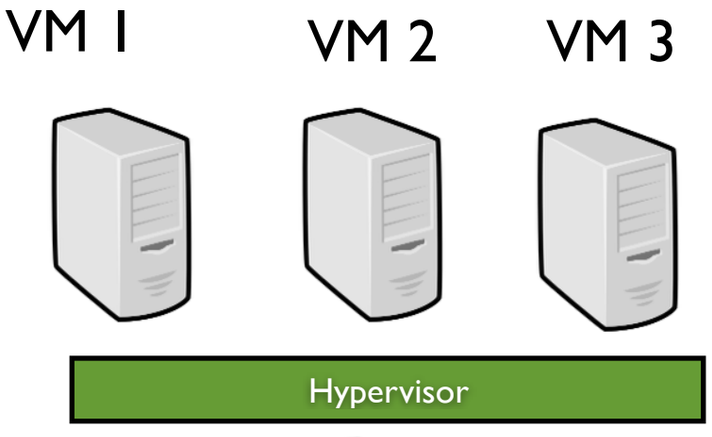
- Isolation (“security” between each VM)
- Snapshotting (a VM can be easily resume from its latest consistent state)

- Suspend/Resume



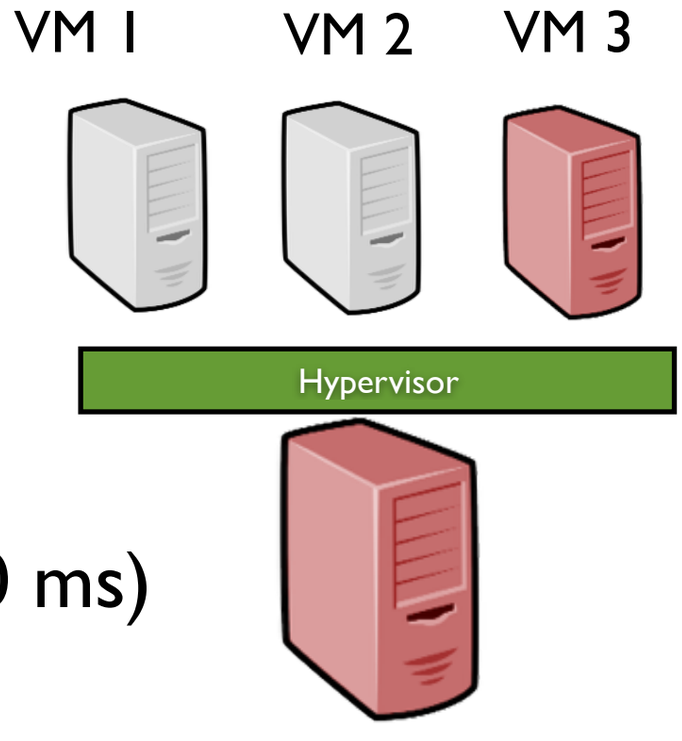
Looking back…

- System virtualization: a great sandbox



- Isolation ("security" between each VM)
- Snapshotting (a VM can be easily resume from its latest consistent state)

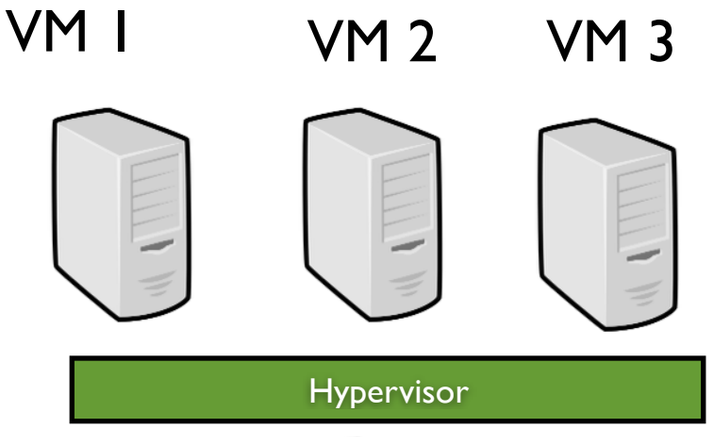
- Suspend/Resume



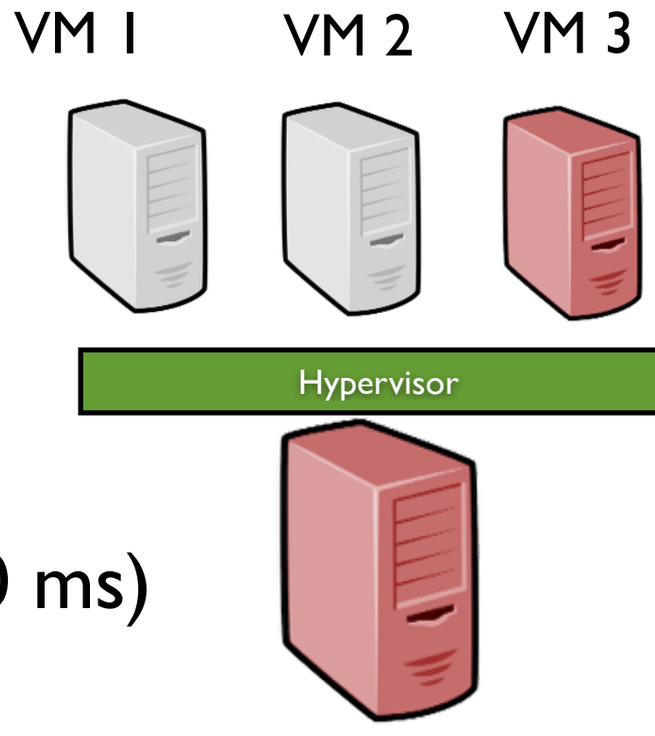
- Live migration
(negligible downtime ~ 60 ms)
Post/Pre Copy

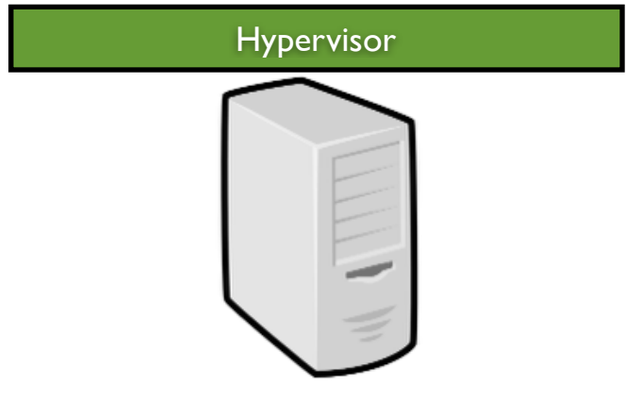
Looking back…

- System virtualization: a great sandbox



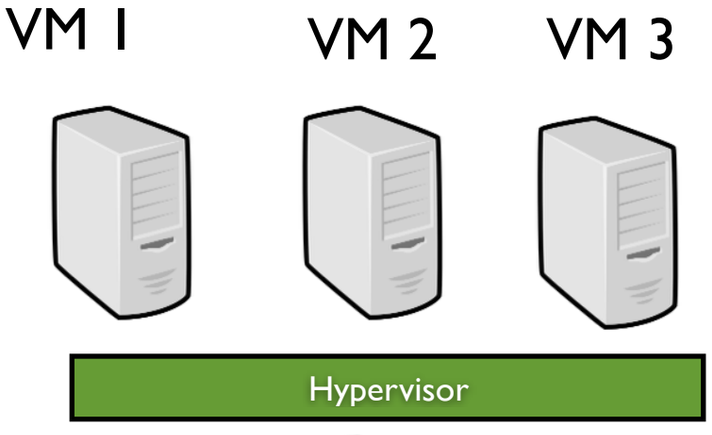
- Isolation ("security" between each VM)
- Snapshotting (a VM can be easily resume from its latest consistent state)
- Suspend/Resume
- Live migration
(negligible downtime ~ 60 ms)
Post/Pre Copy



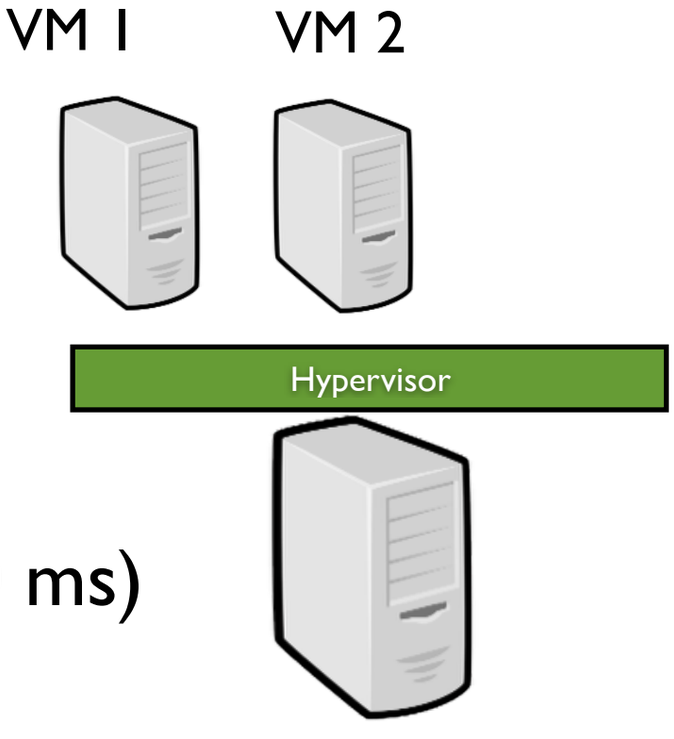


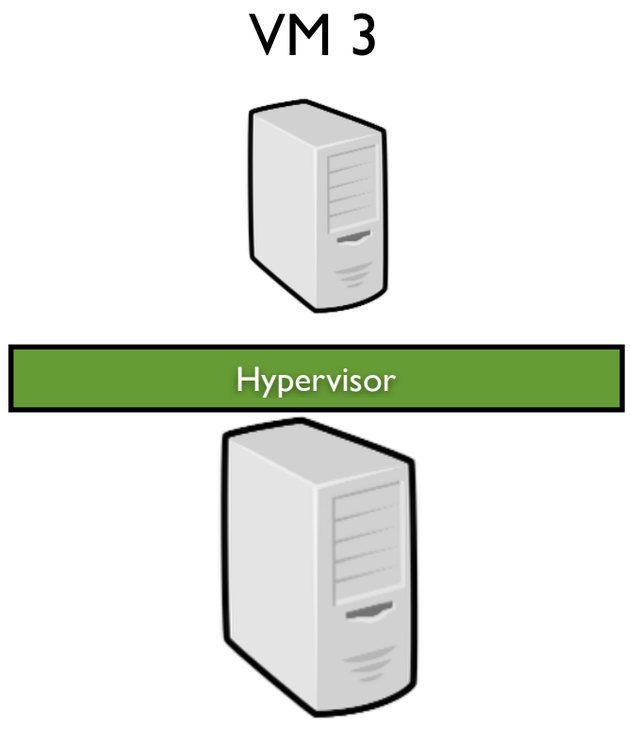
Looking back…

- System virtualization: a great sandbox



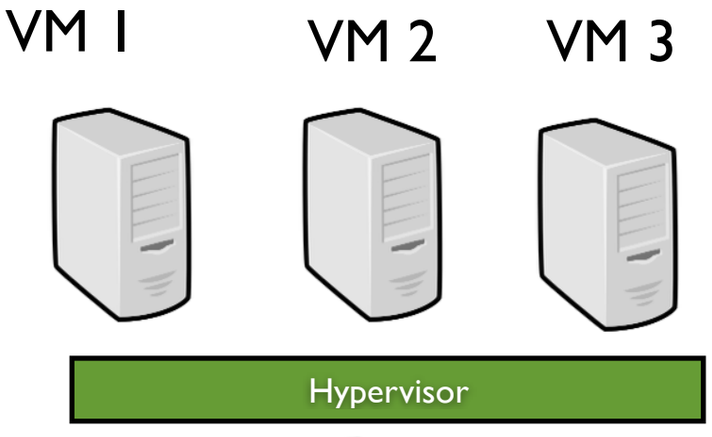
- Isolation ("security" between each VM)
- Snapshotting (a VM can be easily resume from its latest consistent state)
- Suspend/Resume
- Live migration
(negligible downtime ~ 60 ms)
Post/Pre Copy



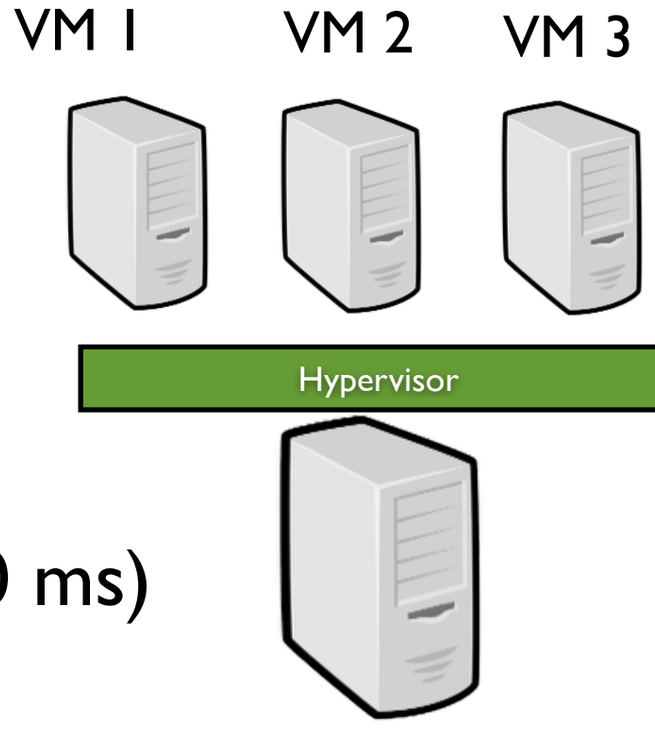


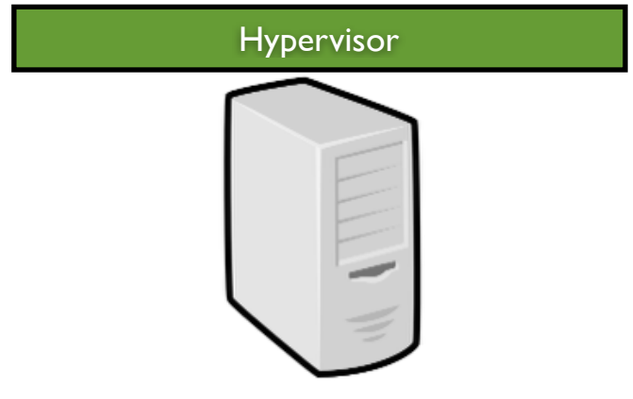
Looking back…

- System virtualization: a great sandbox



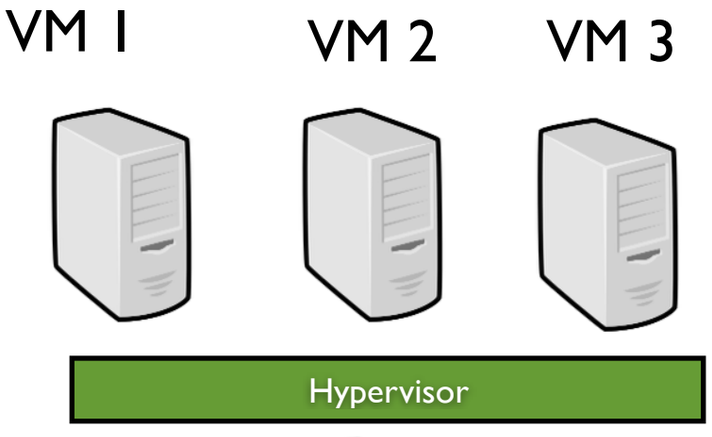
- Isolation ("security" between each VM)
- Snapshotting (a VM can be easily resume from its latest consistent state)
- Suspend/Resume
- Live migration
(negligible downtime ~ 60 ms)
Post/Pre Copy





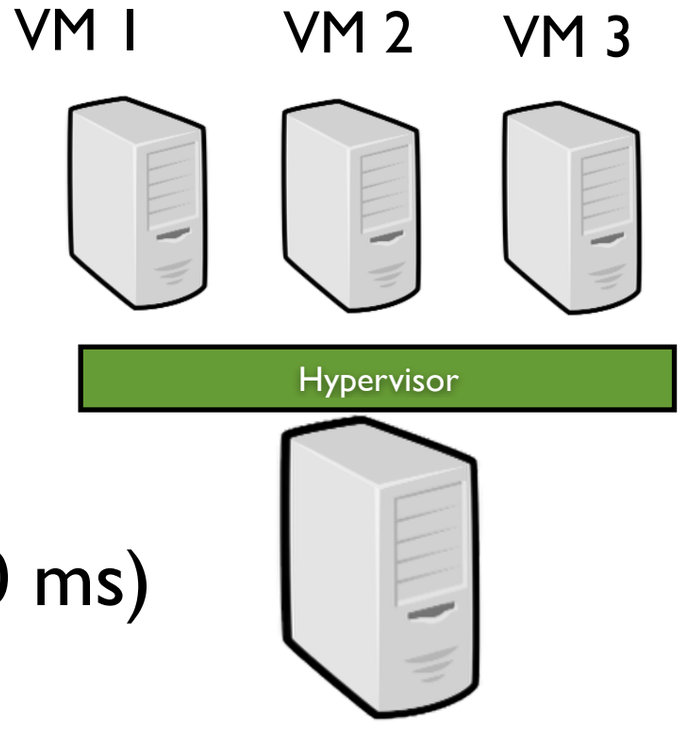
Looking back…

- System virtualization: a great sandbox



- Isolation ("security" between each VM)
- Snapshotting (a VM can be easily resume from its latest consistent state)

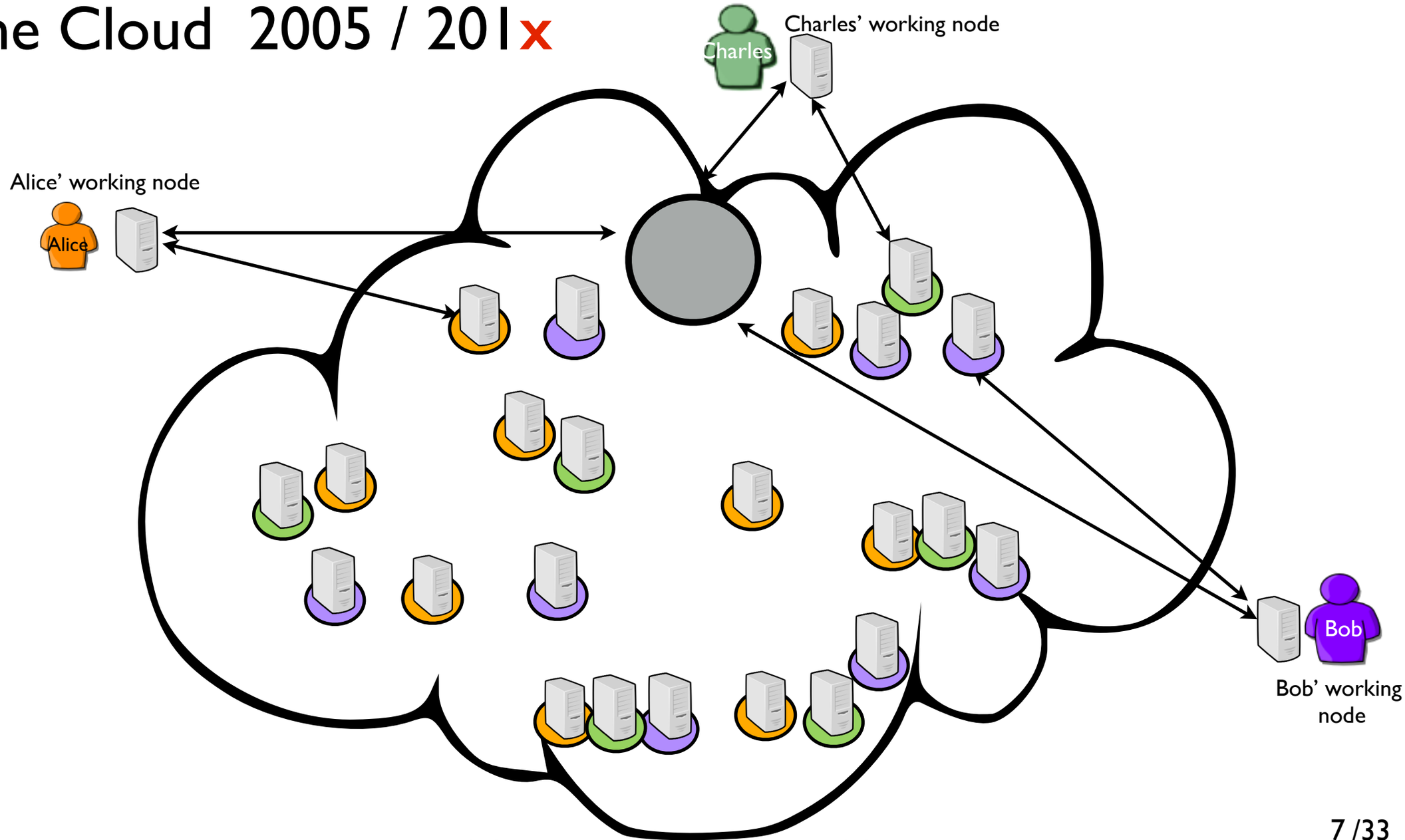
- Suspend/Resume



- Live migration
(negligible downtime ~ 60 ms)
Post/Pre Copy

Looking back …

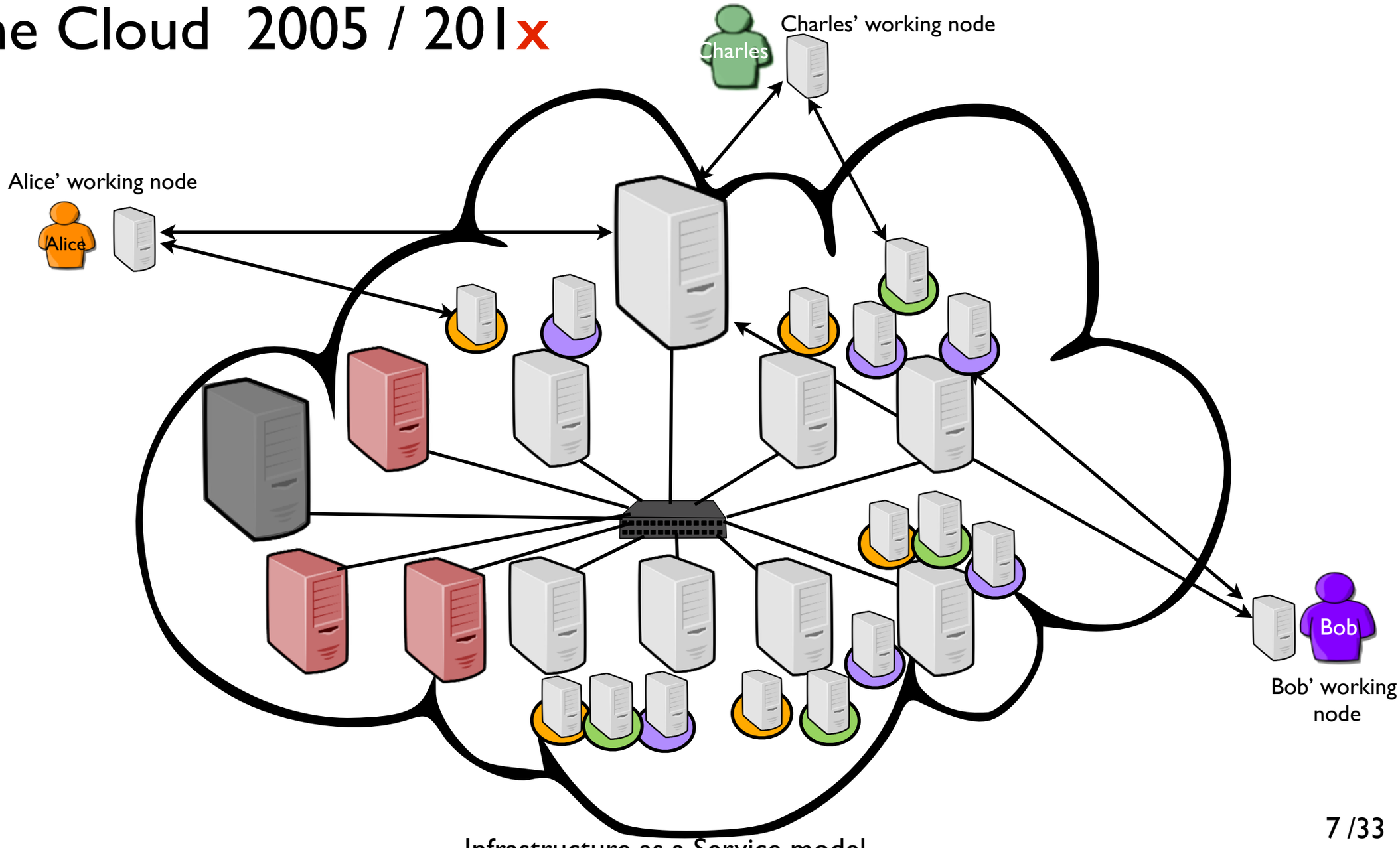
- Network of Workstations 1990 / 20**xx**
- The Grid 1997 / 201**x**
- The Cloud 2005 / 201**x**



Infrastructure as a Service model

Looking back …

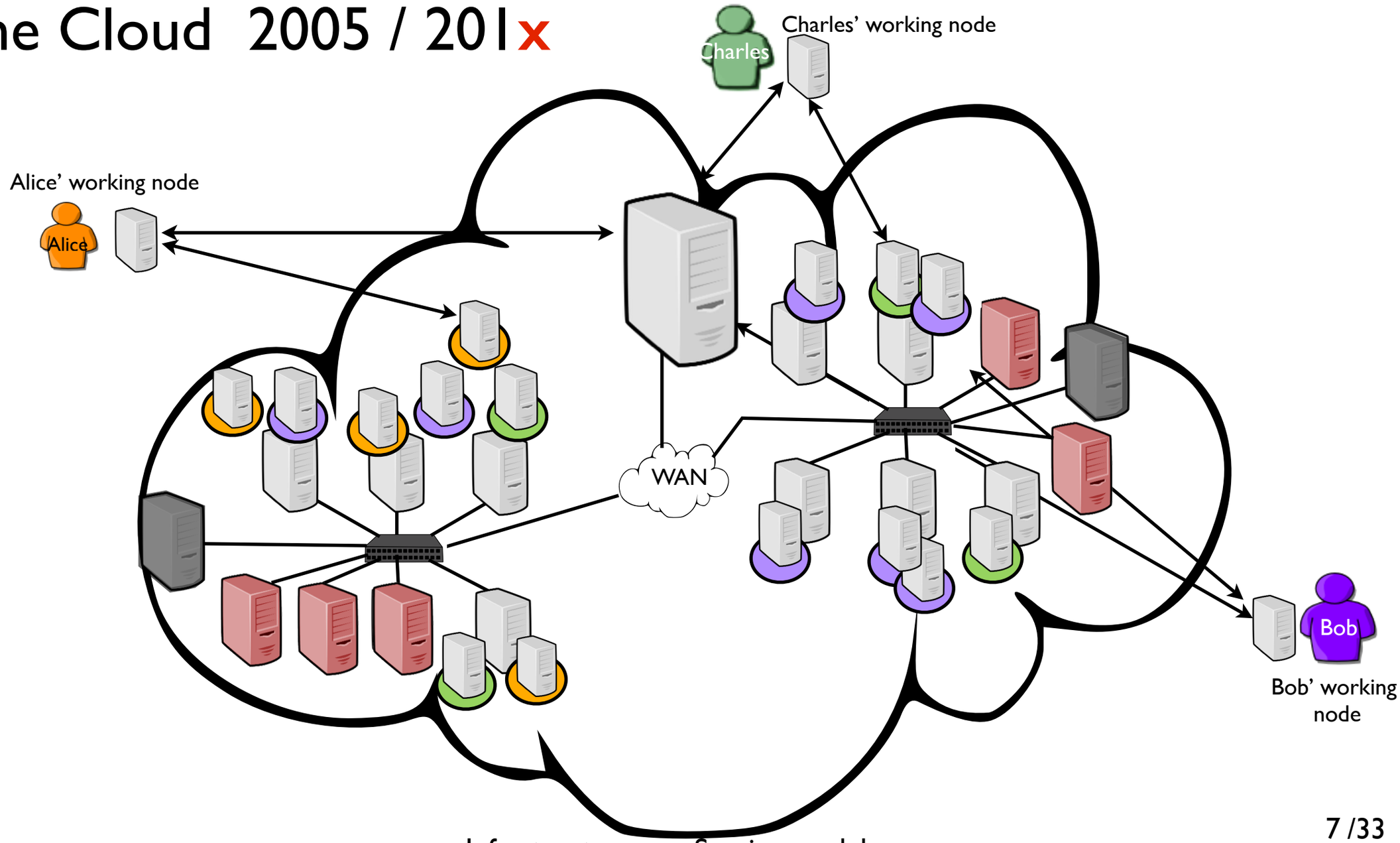
- Network of Workstations 1990 / 20**xx**
- The Grid 1997 / 201**x**
- The Cloud 2005 / 201**x**



Infrastructure as a Service model

Looking back …

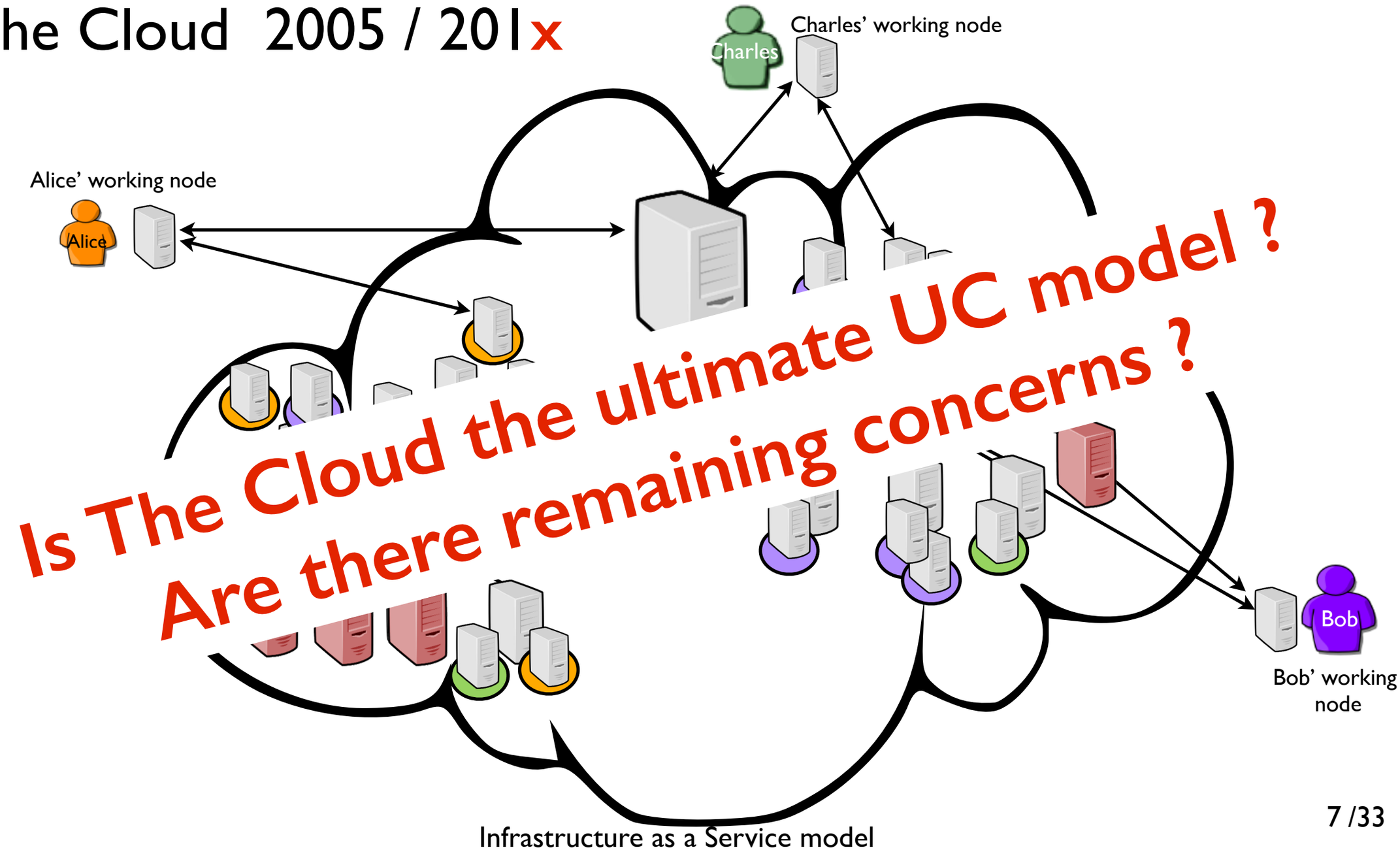
- Network of Workstations 1990 / 20**xx**
- The Grid 1997 / 201**x**
- The Cloud 2005 / 201**x**



Infrastructure as a Service model

Looking back …

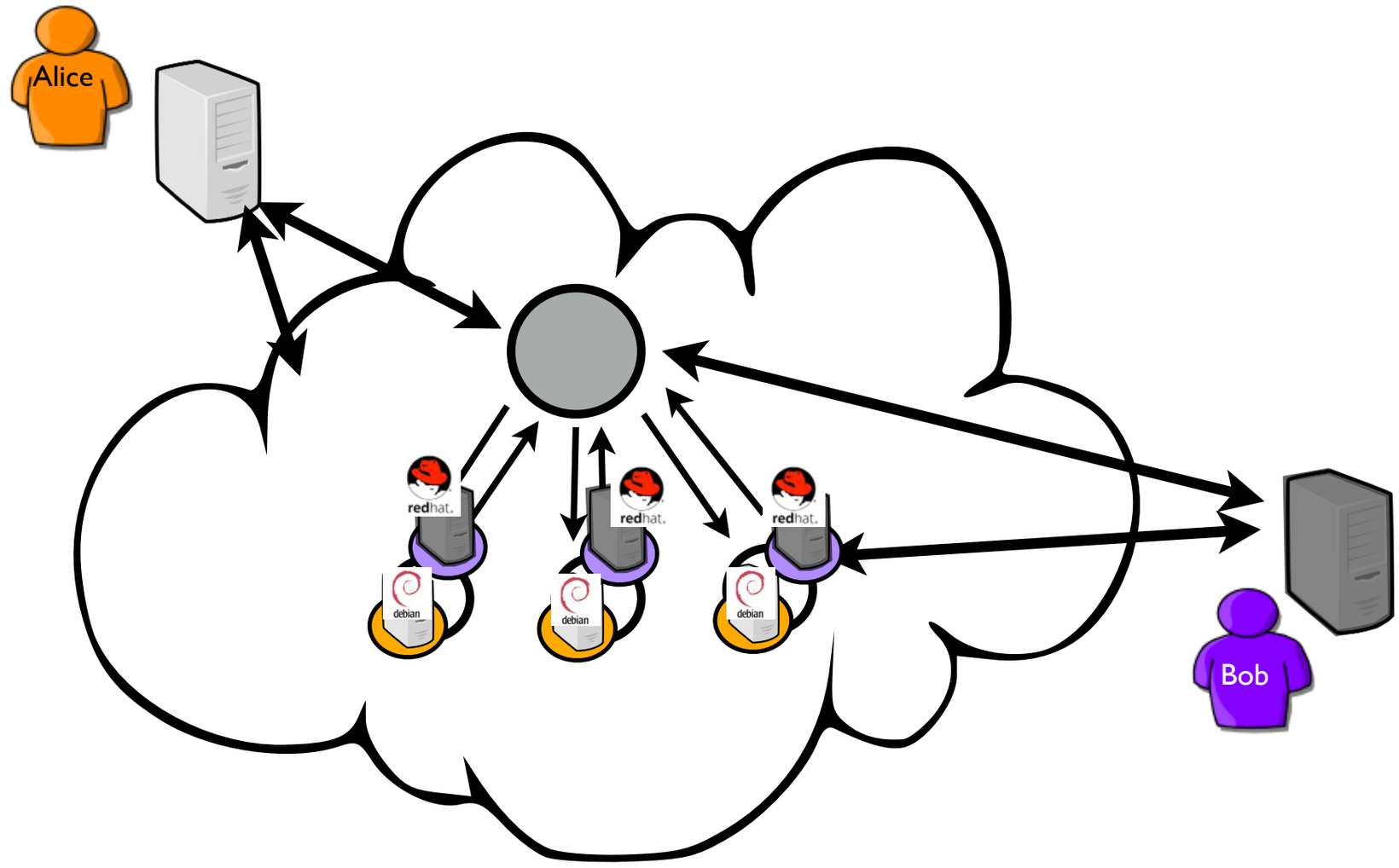
- Network of Workstations 1990 / 20**xx**
- The Grid 1997 / 201**x**
- The Cloud 2005 / 201**x**



Infrastructure as a Service model

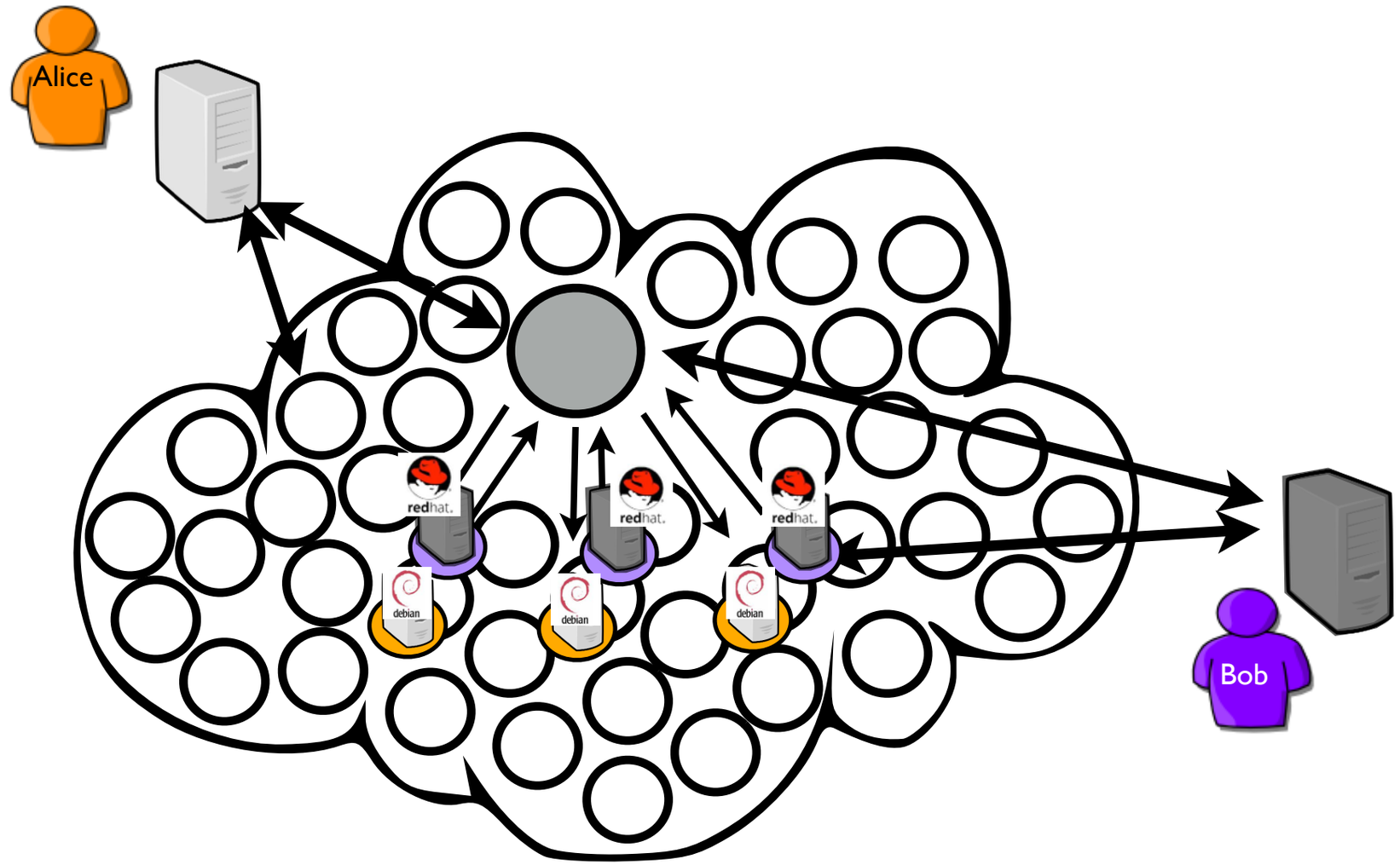
Where we are ?

- IaaS challenges



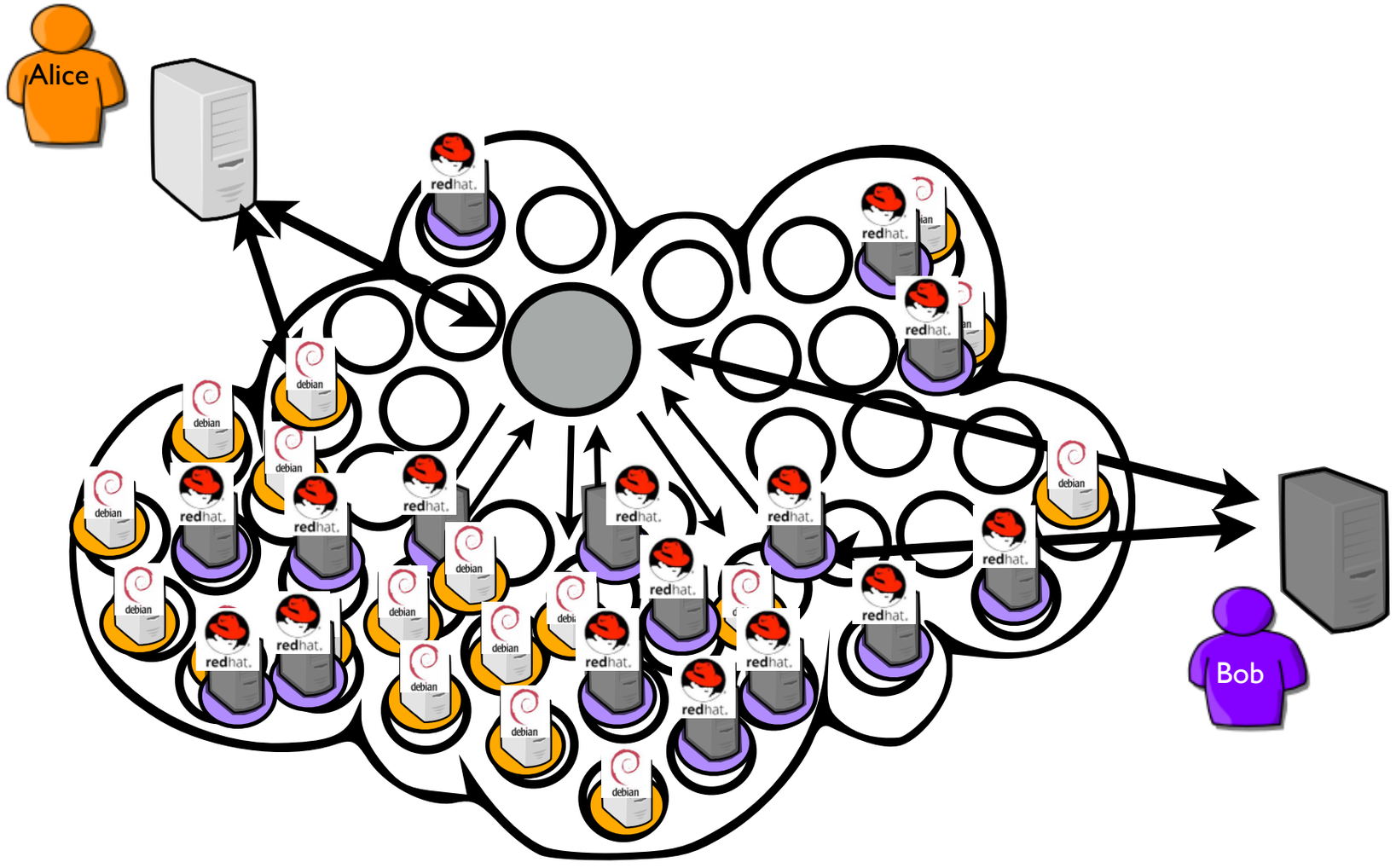
Where we are ?

- IaaS challenges



Where we are ?

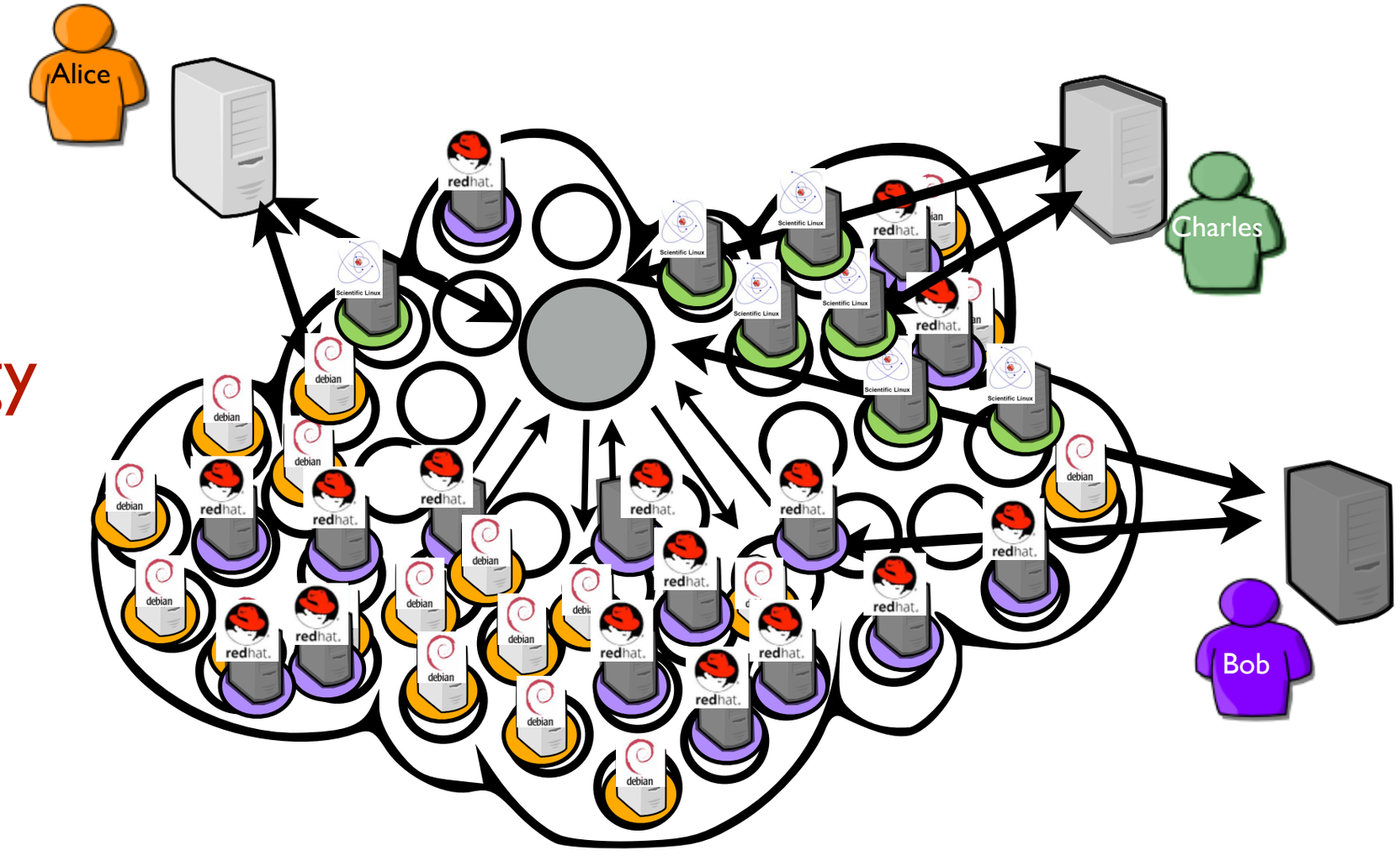
- IaaS challenges



Where we are ?

- IaaS challenges

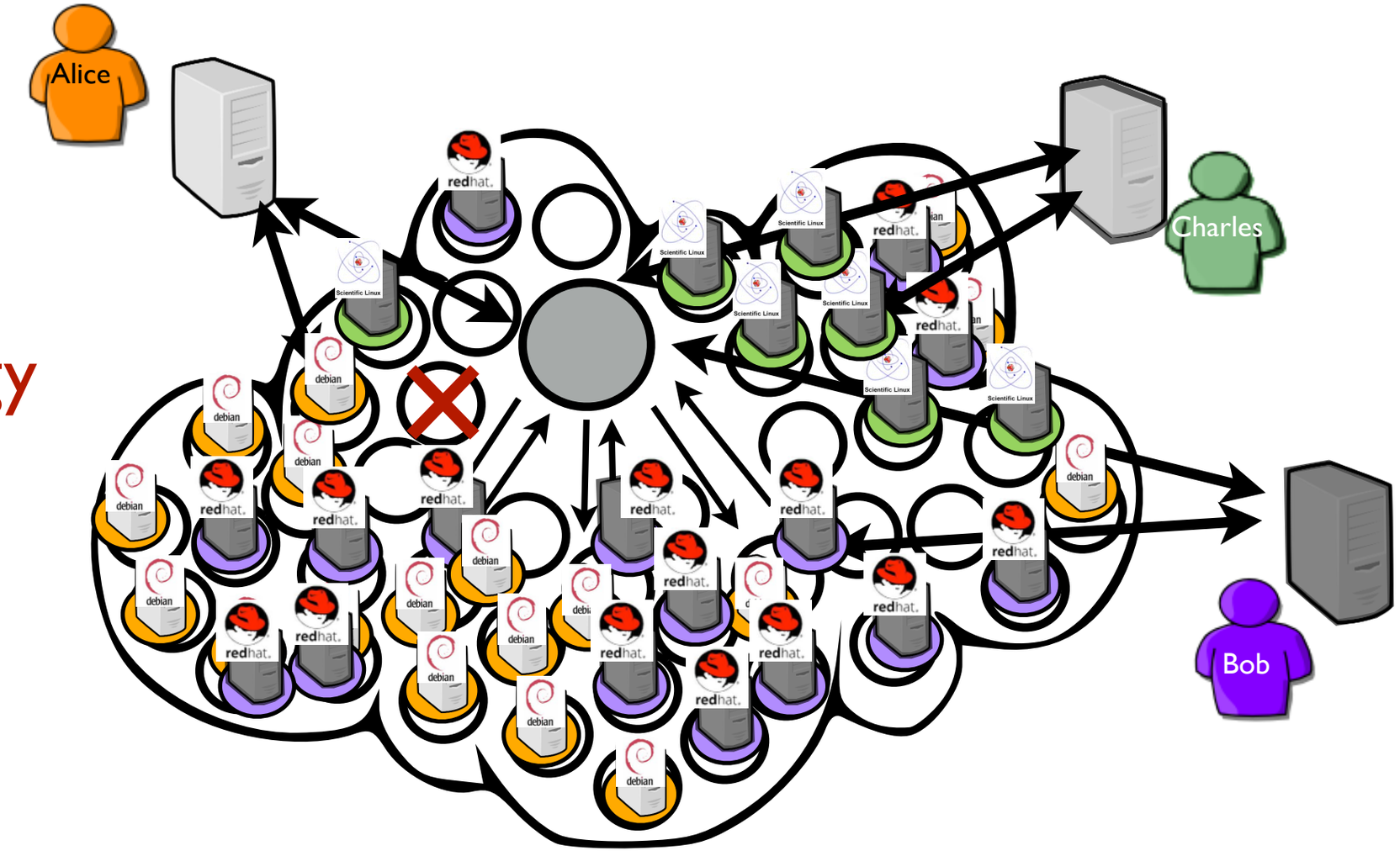
  Scalability / Energy



Where we are ?

- IaaS challenges

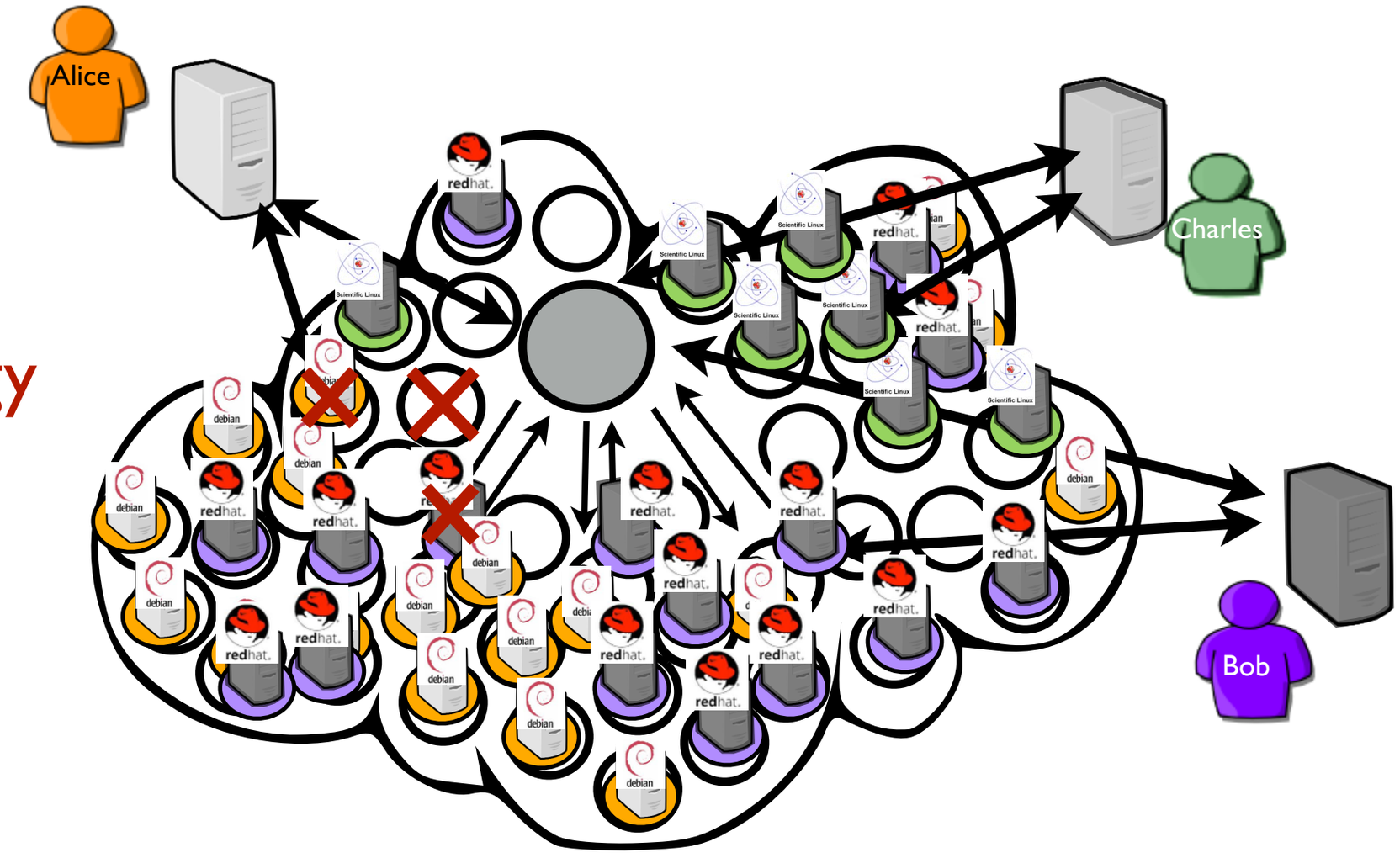
  Scalability / Energy



Where we are ?

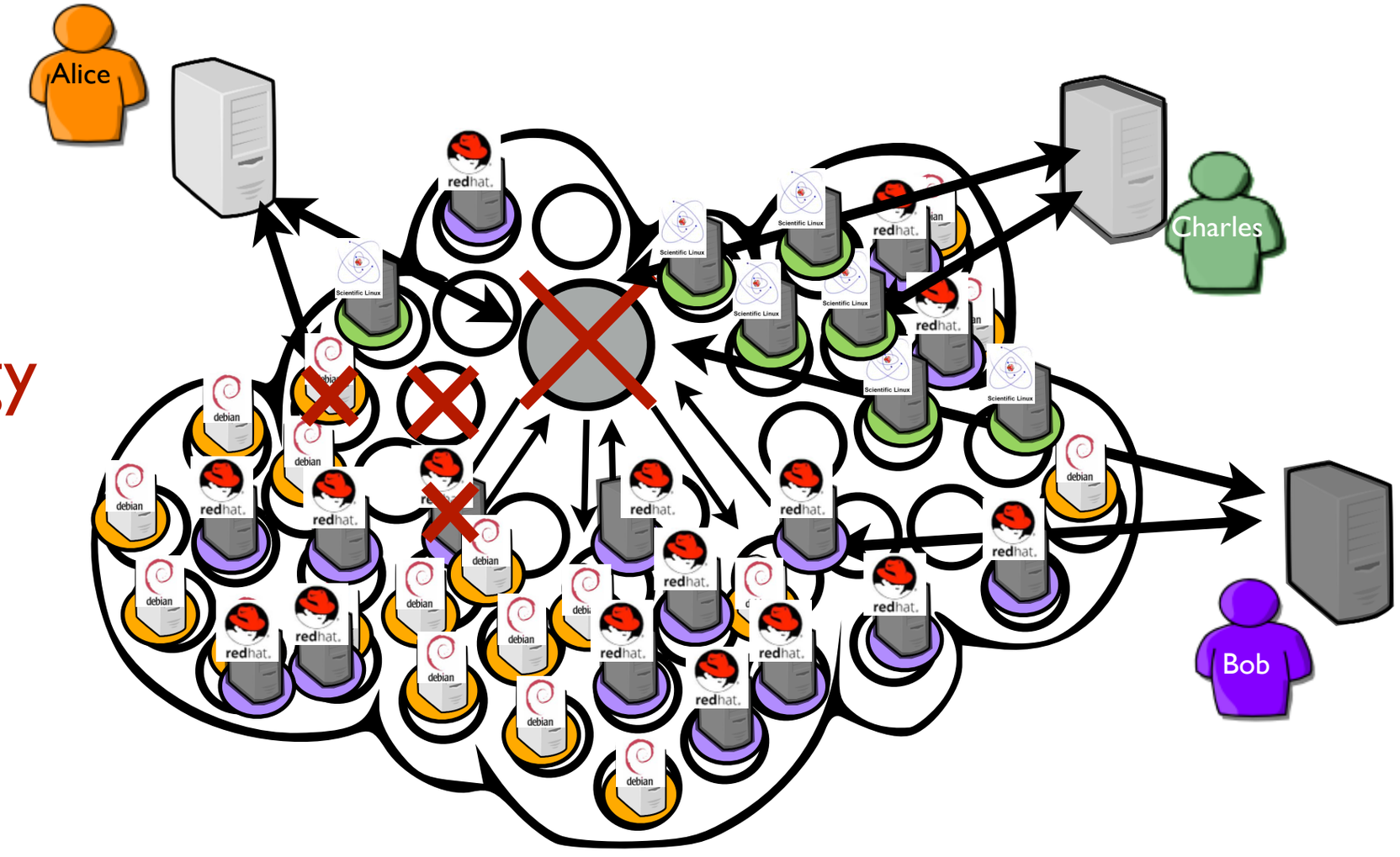
- IaaS challenges

  Scalability / Energy



Where we are ?

- IaaS challenges
  - Scalability / Energy
  - Reliability



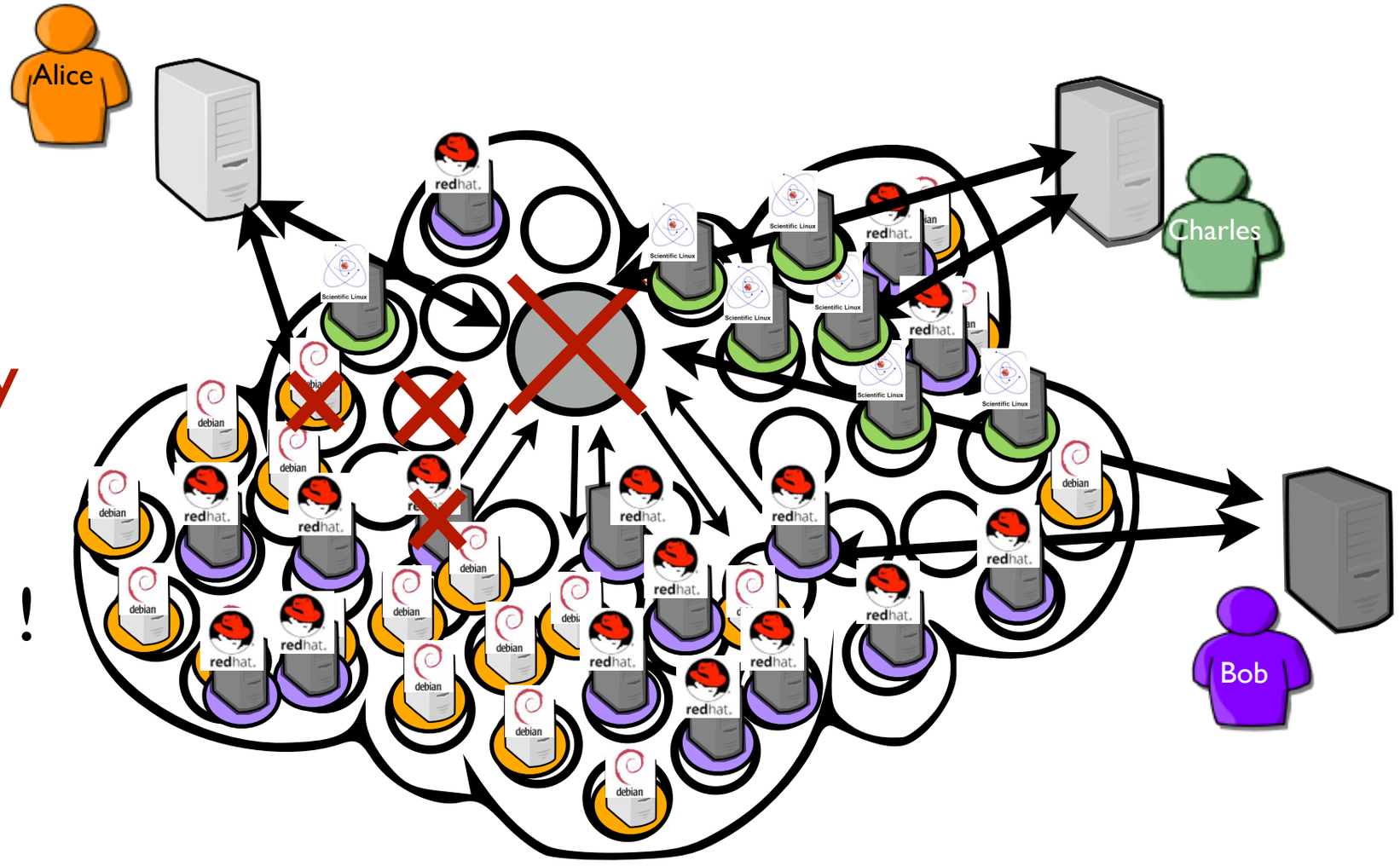
Where we are ?

- IaaS challenges

  Scalability / Energy

  Reliability

  nothing really new !



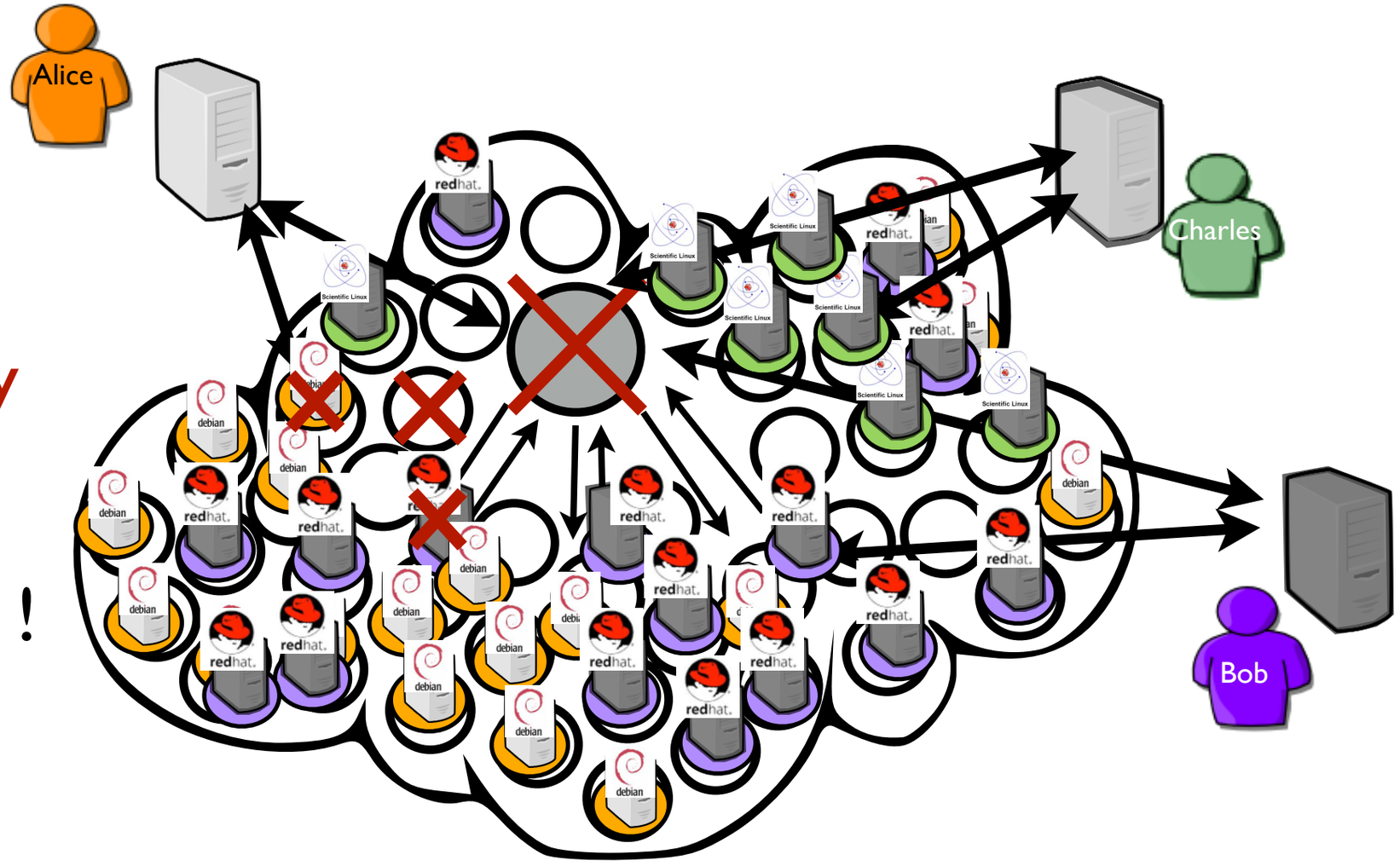
Where we are ?

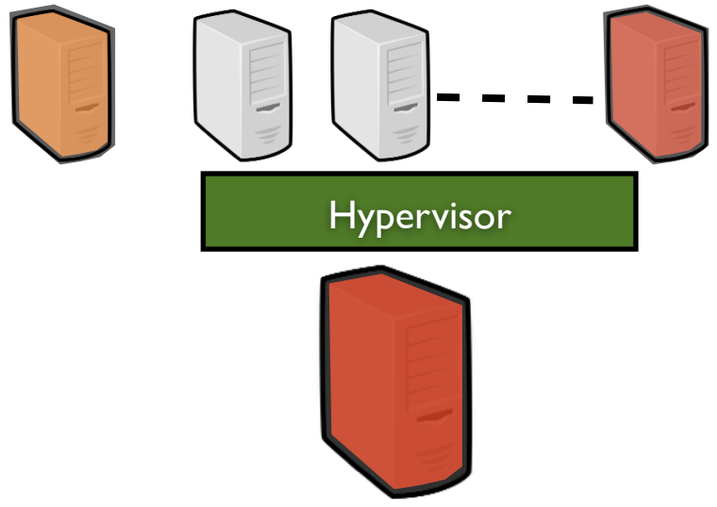
- IaaS challenges

  Scalability / Energy

  Reliability

  nothing really new !





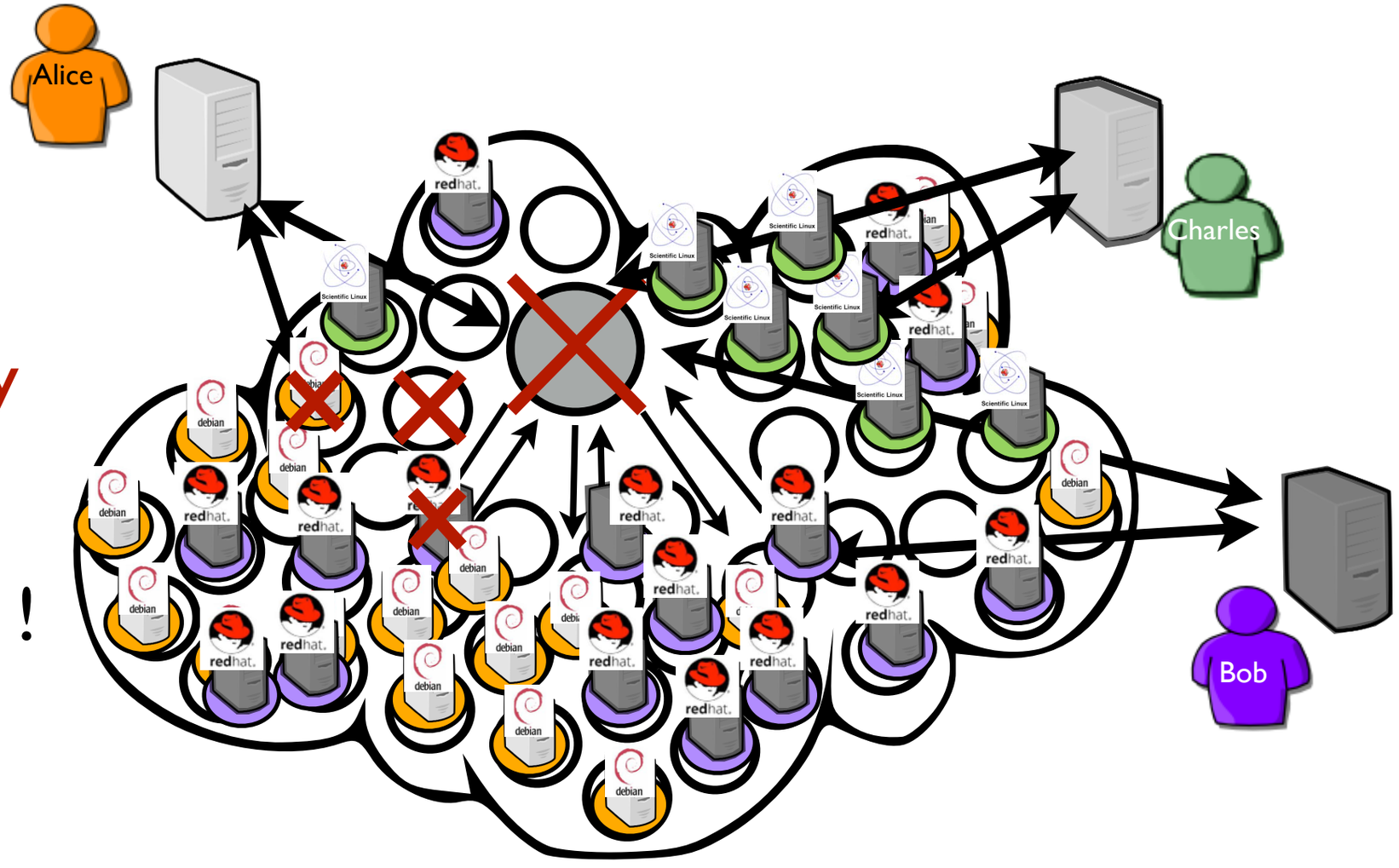
Where we are ?

- IaaS challenges

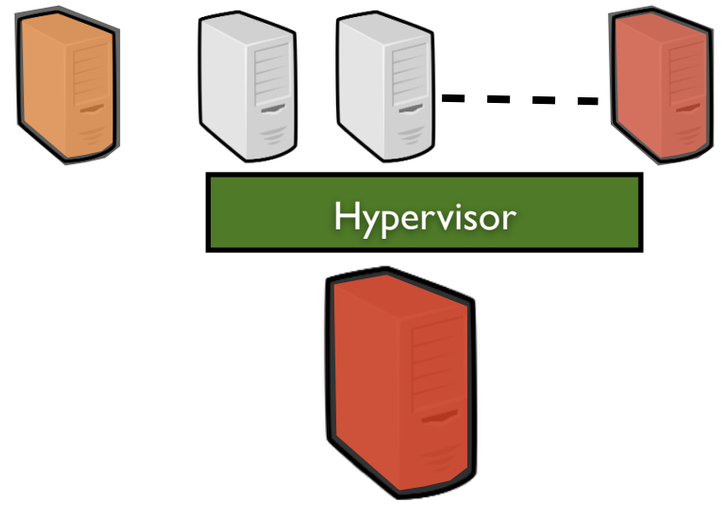
  Scalability / Energy

  Reliability

  nothing really new !



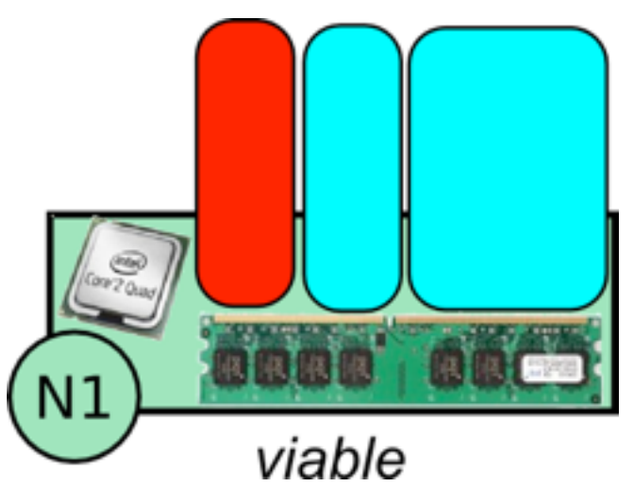
- Virtualize IT impacts performances !
  (difficulty to guarantee performances, SLAs)



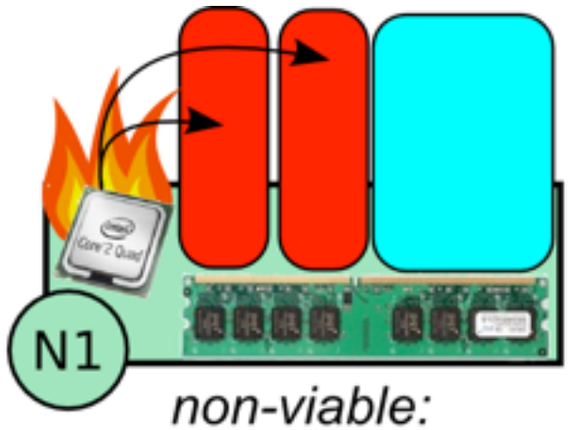
Virtualization and Performance

Viable and Non Viable Configurations

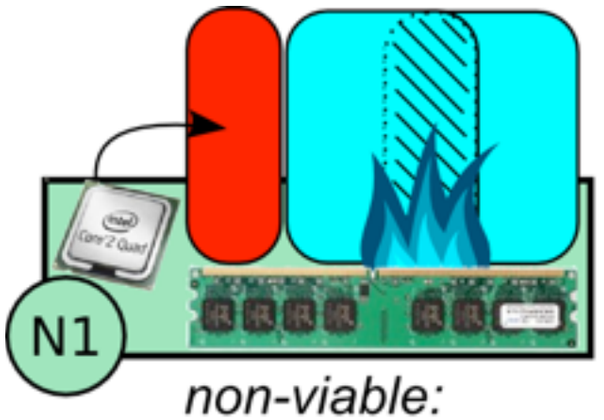
- Fine management of resources (efficiency and energy constraints)
- Find the "right" mapping between needs of VMs and resources provided by PMs



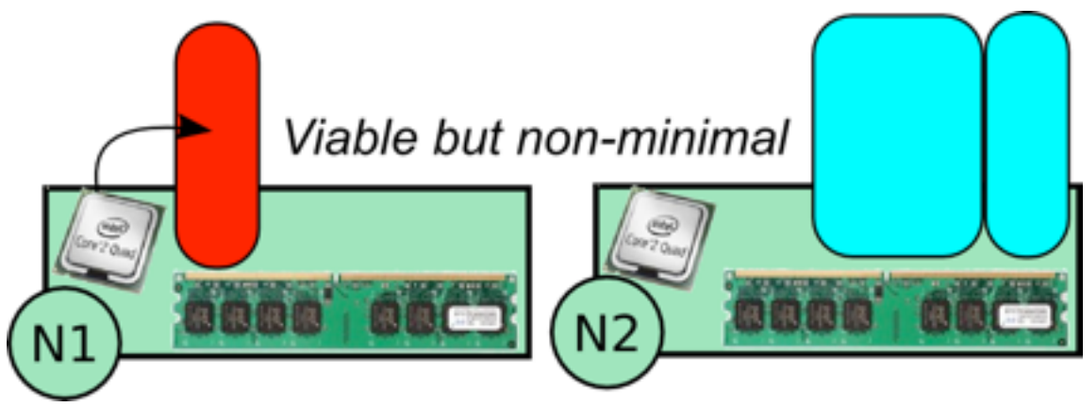
viable

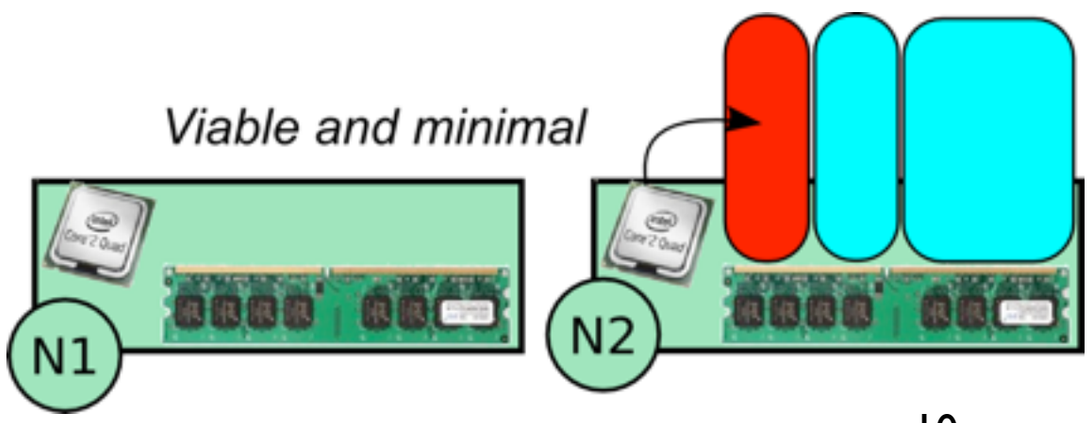


non-viable:
2 active VMs for one CPU



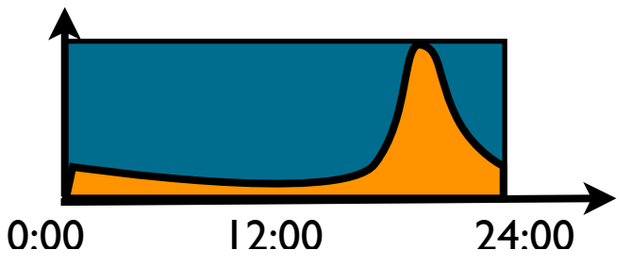
non-viable:
memory overcommitment



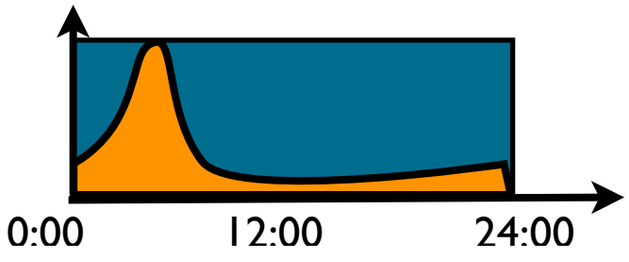


credits: F. Hermenier, OSDI poster session 2008

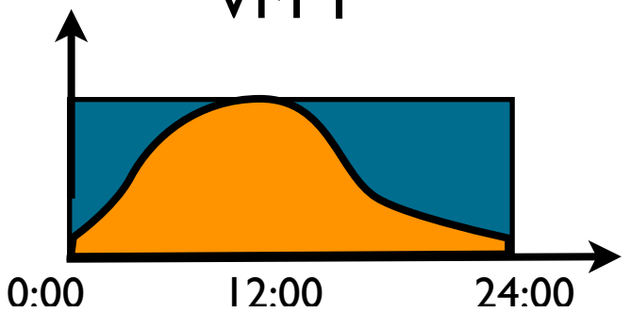
Fluctuations of VM Requirements



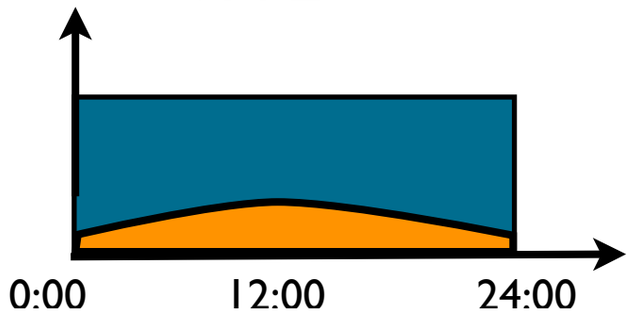
VM 1



VM 2



VM 3



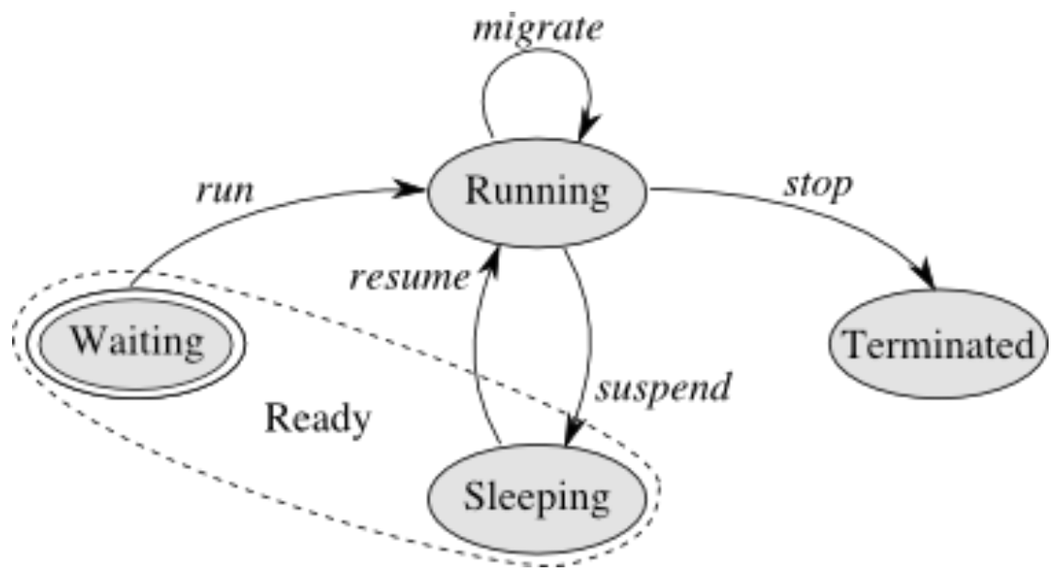
VM 4

- Static placement policies (as delivered by most of the popular Cloud Computing management systems)
  ⇒ Simple but prevent CC providers to maximize the usage of CC resources while guaranteeing VM resource requirements

- Used advanced dynamic placement strategies to relocate VMs according to the scheduler objectives / available resources / waiting queue / ...

Dynamic VM Placement Policies

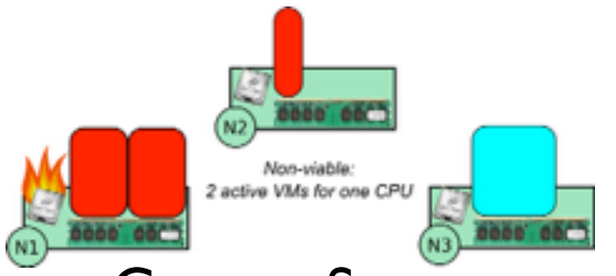
- Generale idea: leverage VM capabilities to manipulate VEs in a similar way of usual processes on a laptop (a VE is a users' working environment, possibly composed of several interconnected VMs)



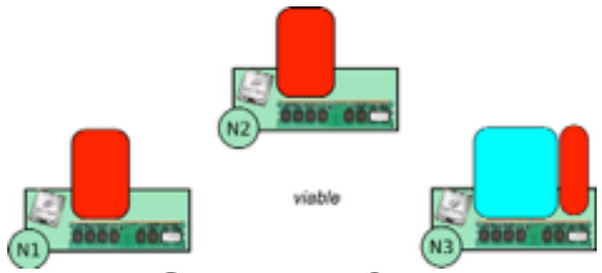
- Each VE is in a particular state

- Perform VE context switches (a set of VM context switches) to reschedule/rebalance the LUC infrastructure

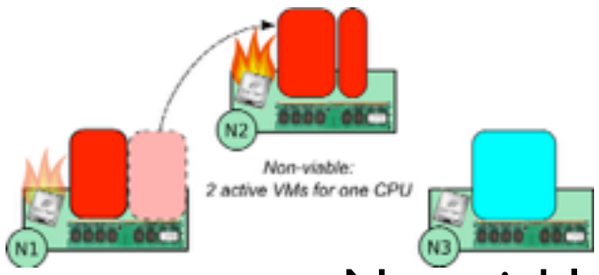
Placement constraints (btrPlace)

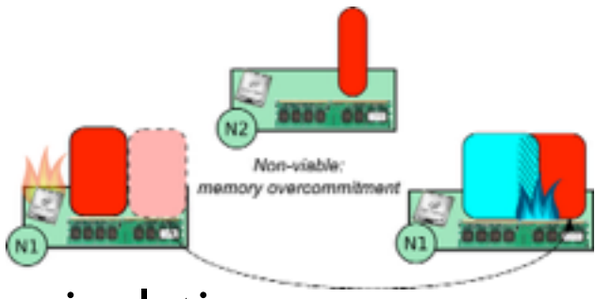


Current Status

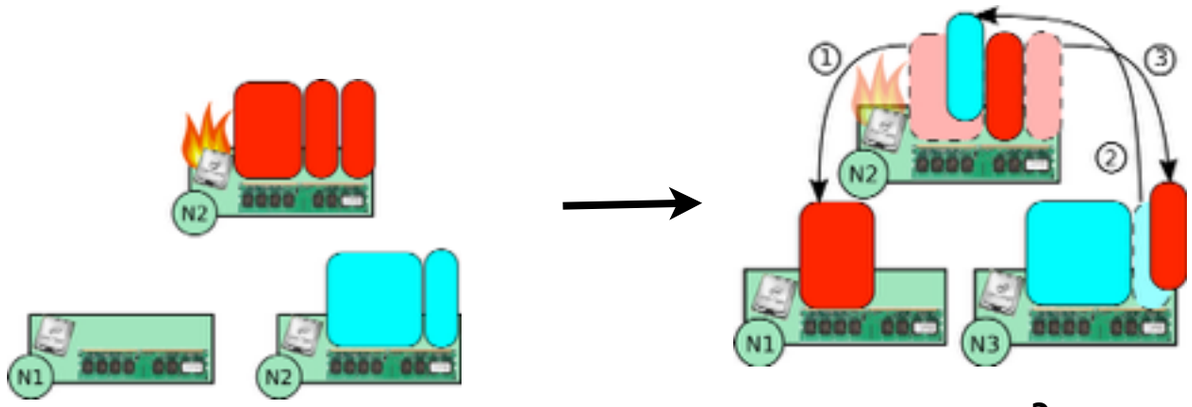


Correct Status

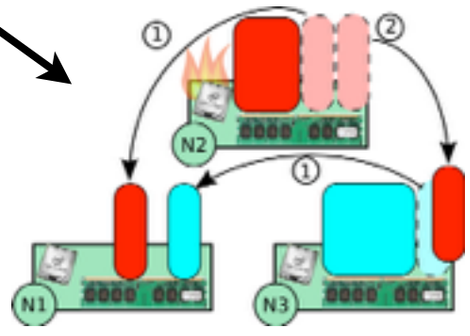




Non-viable manipulations

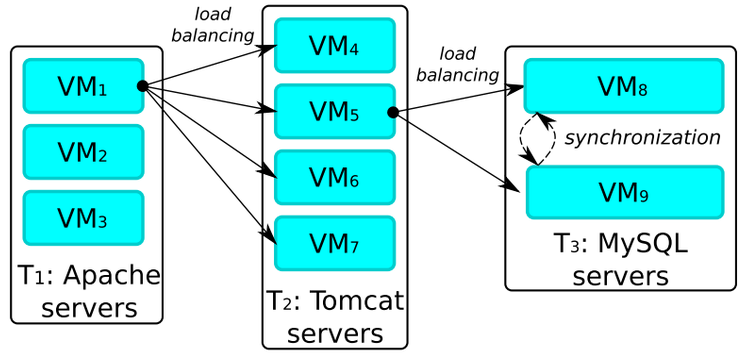


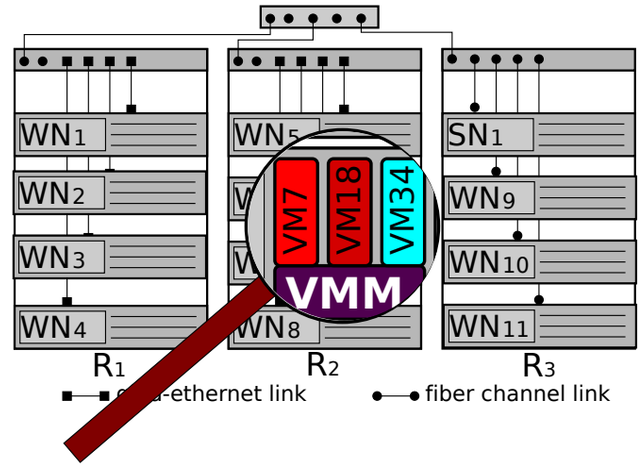
cost: 3



cost: 2

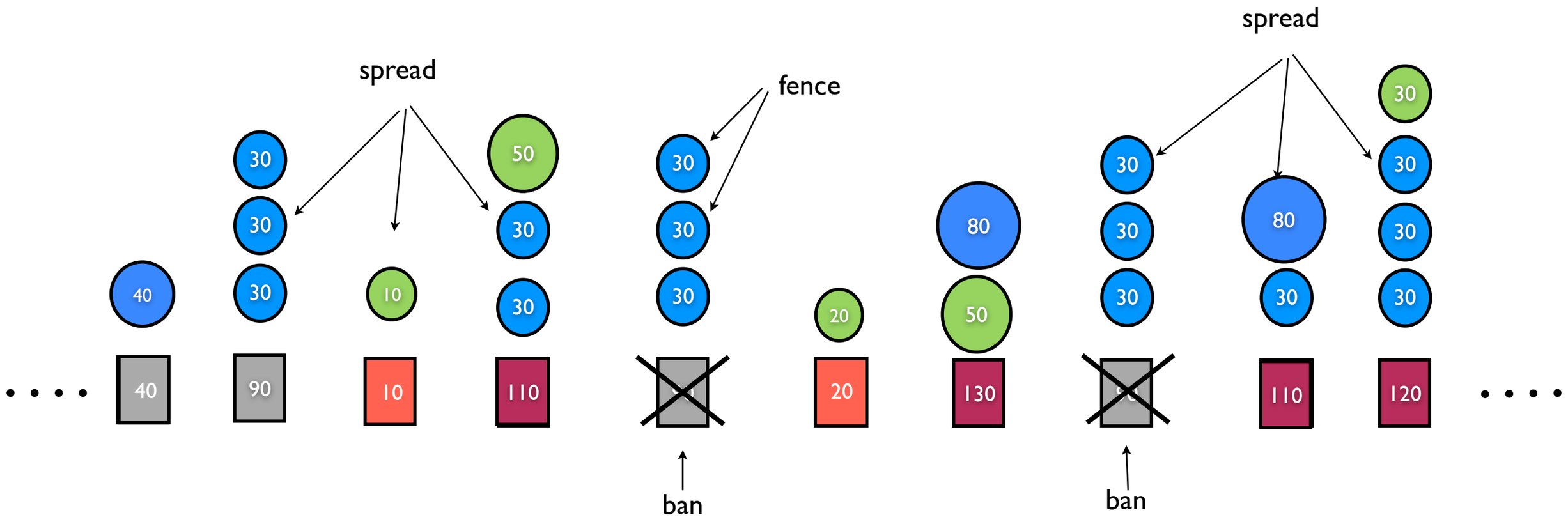
- Find the "right" mapping between needs of VMs, their constraints and resources provided by PMs





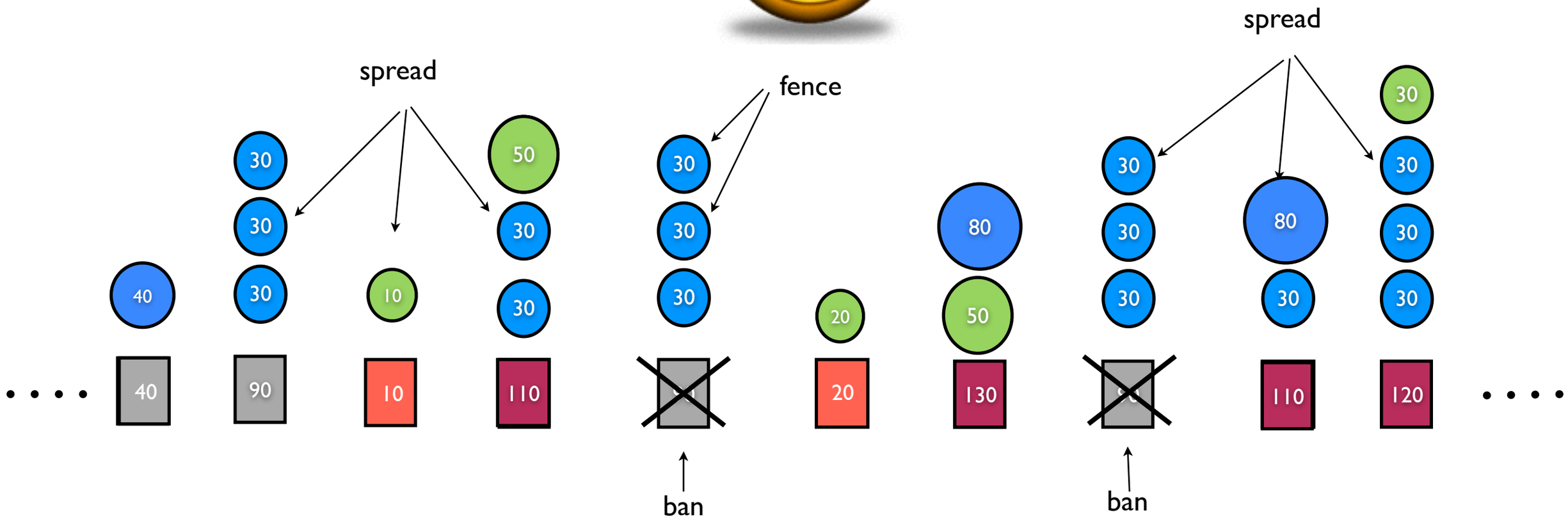
credits: F. Hermenier, Sophia Antipolis University, www.btrplace.org

a Small Example

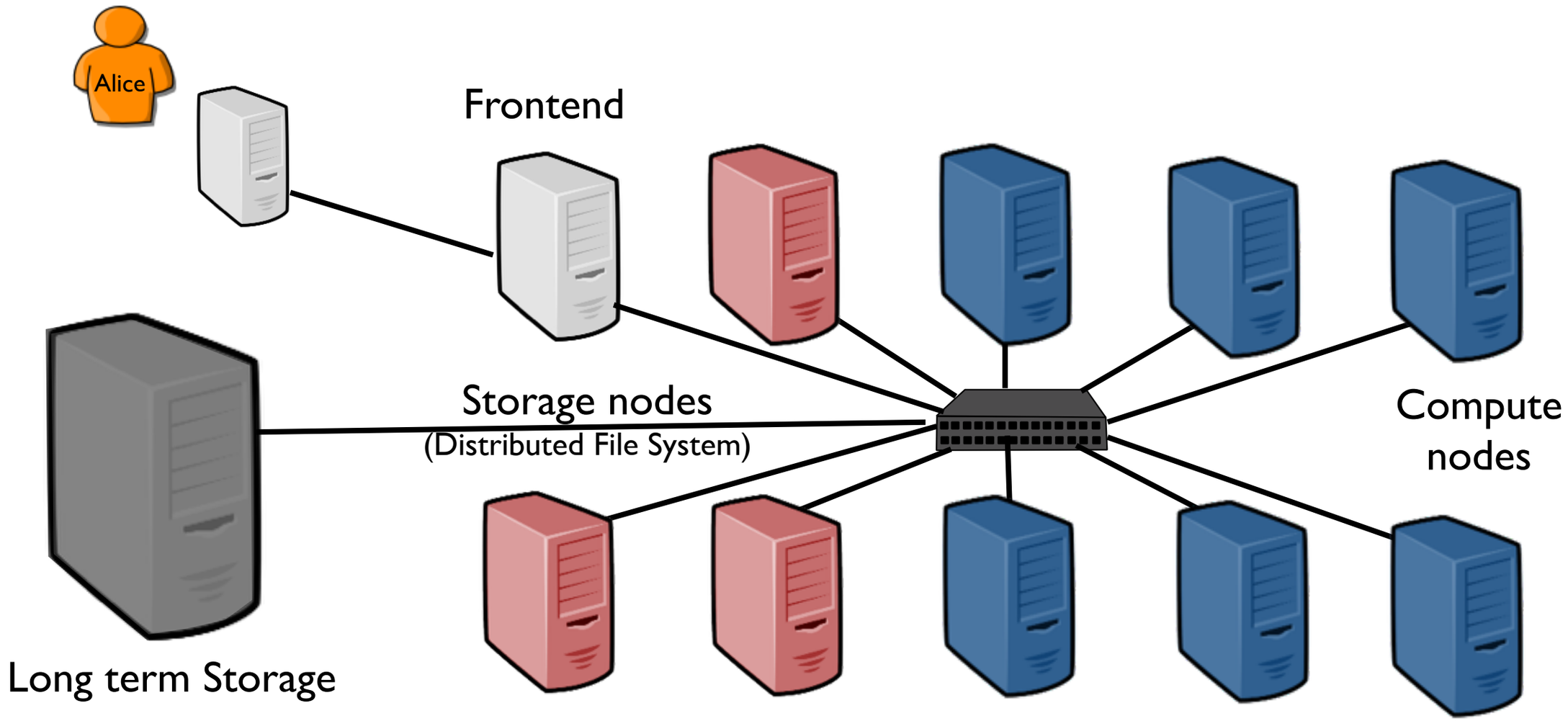


a Small Example

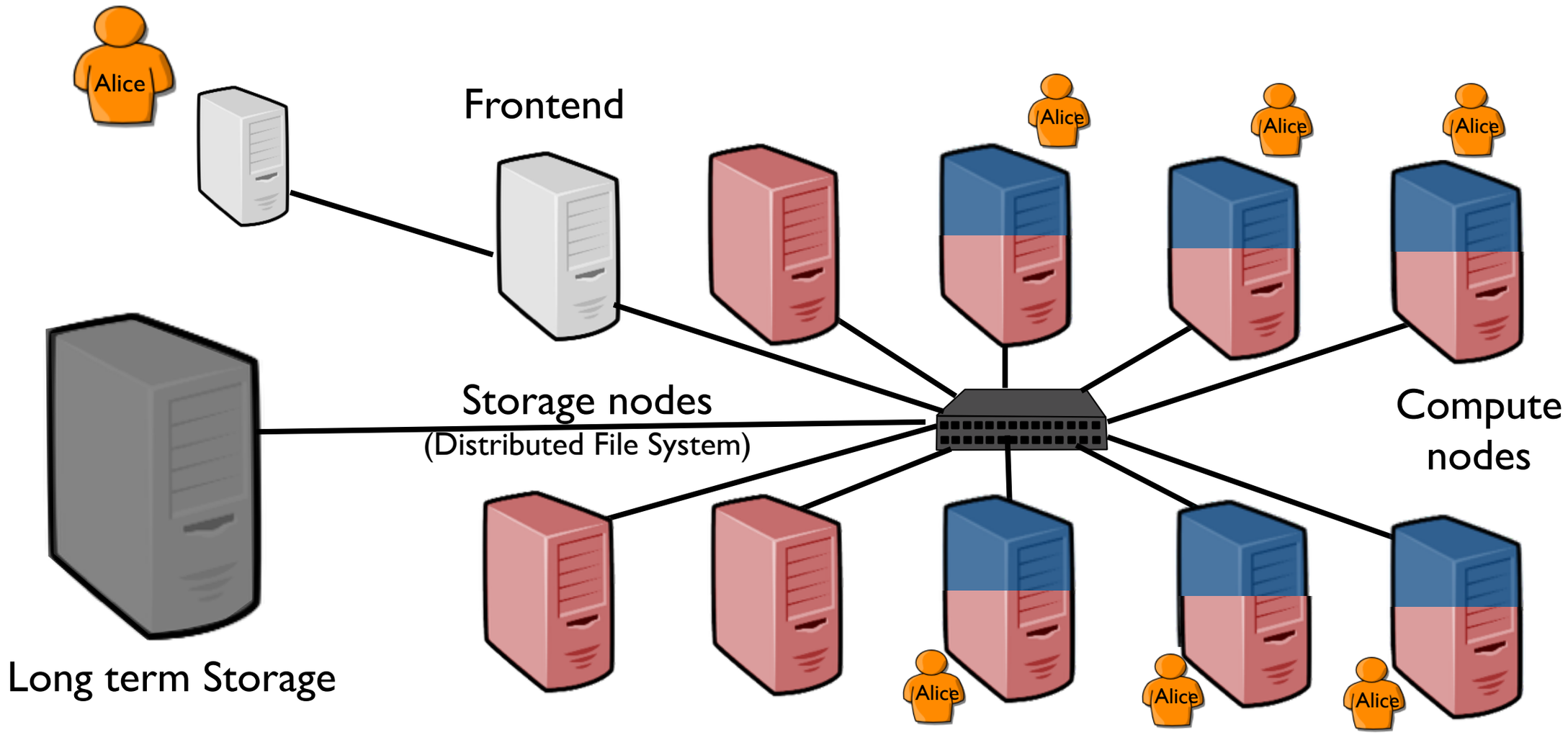
Only CPU is considered in this simple example



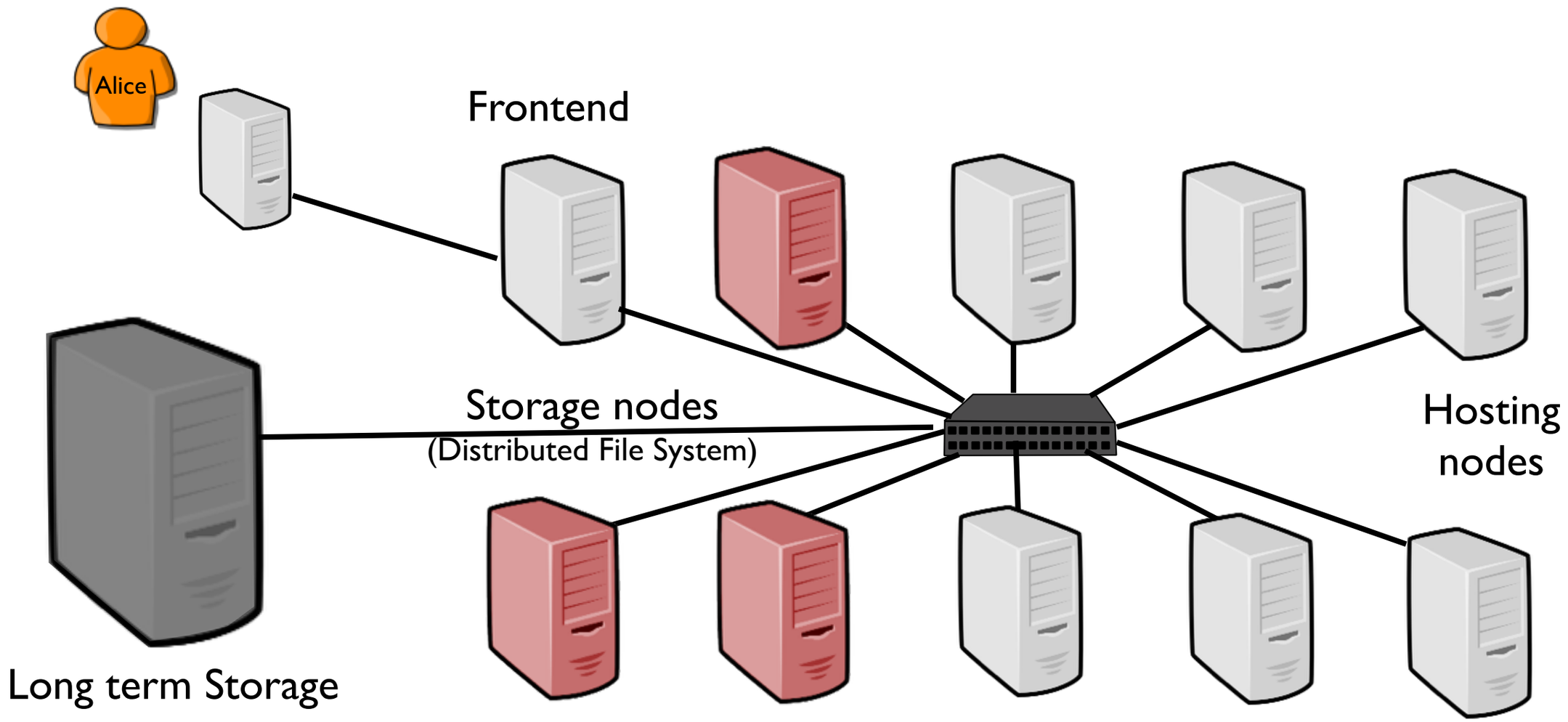
Another with Map/Reduce



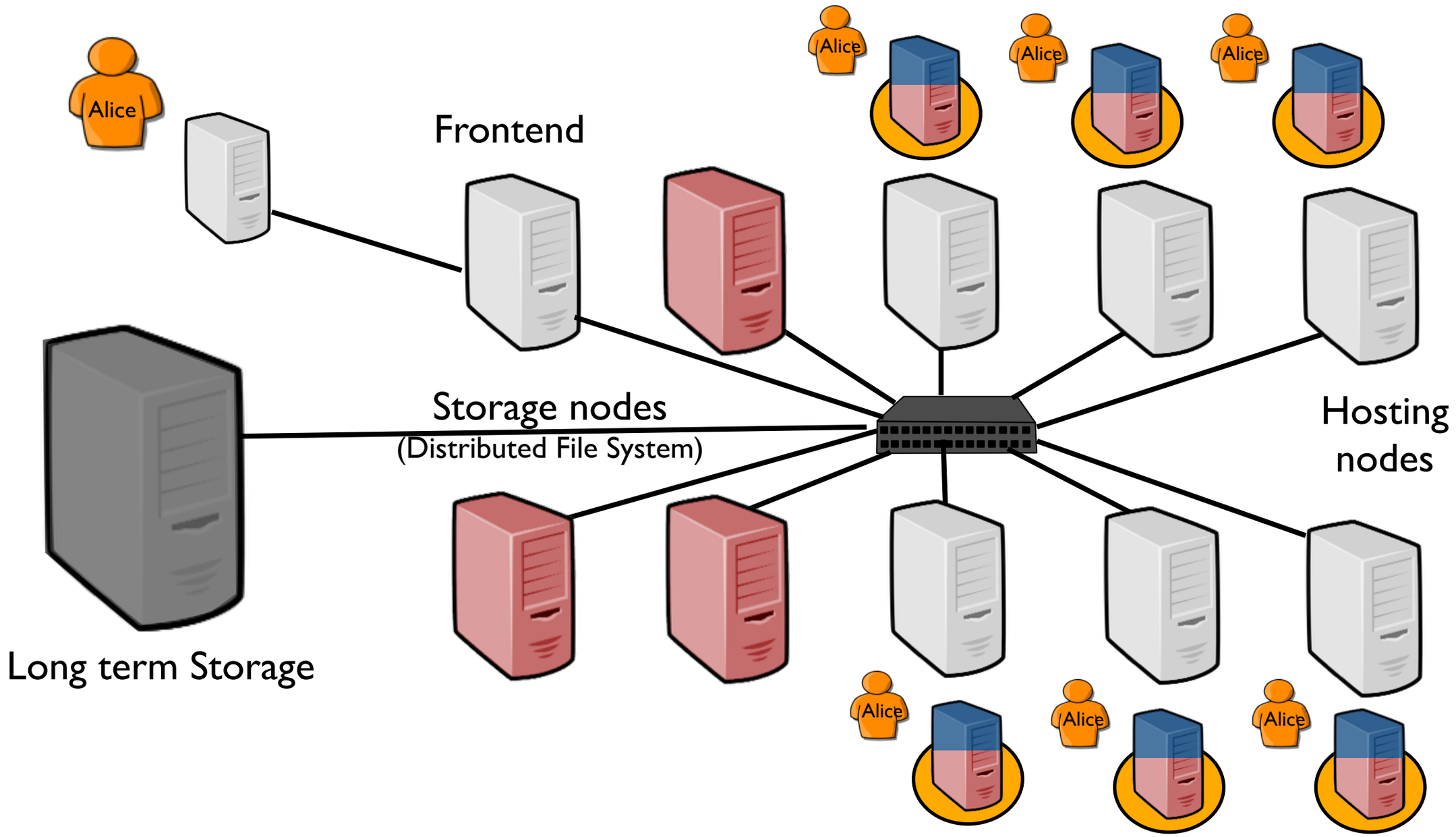
Another with Map/Reduce



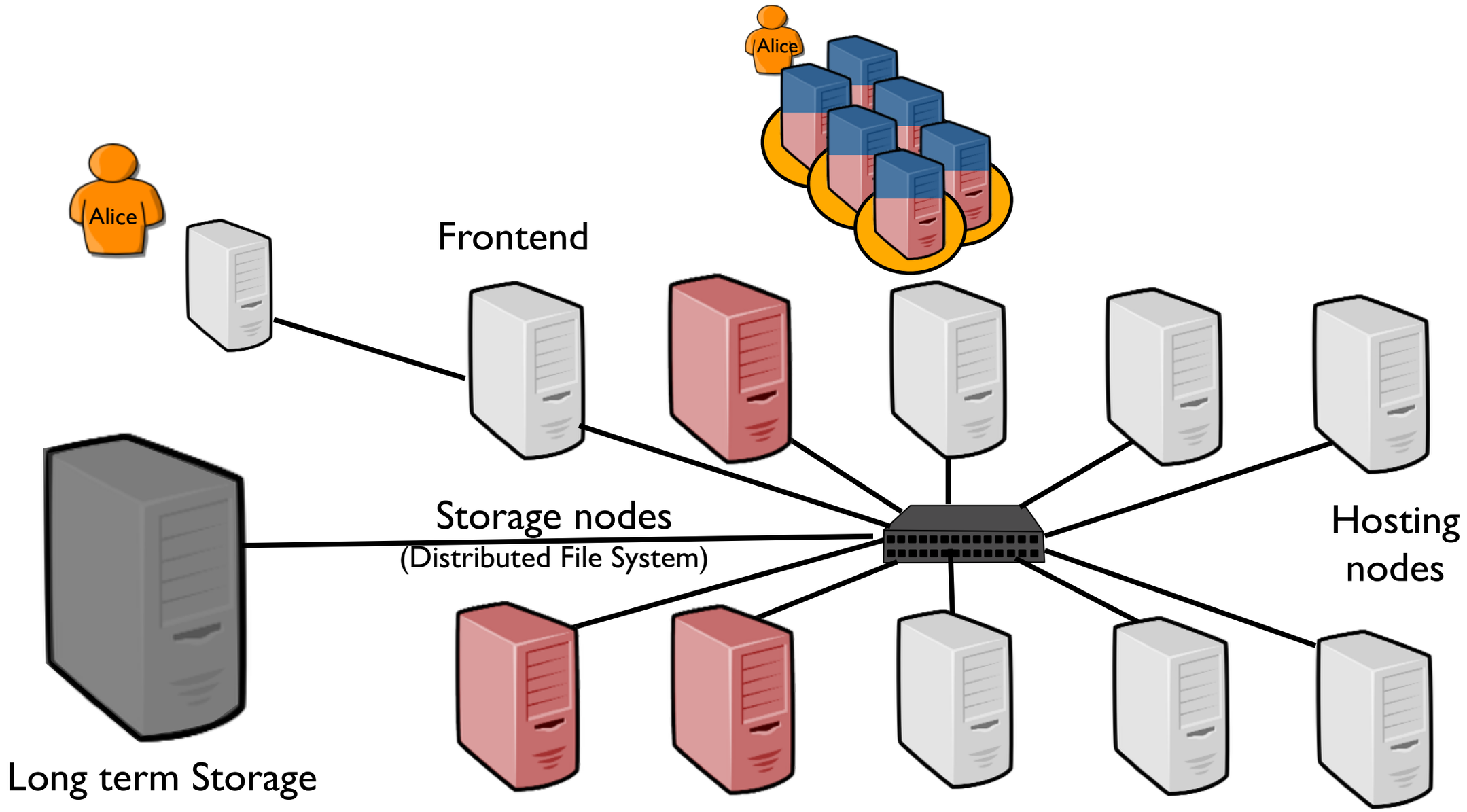
Another with Map/Reduce



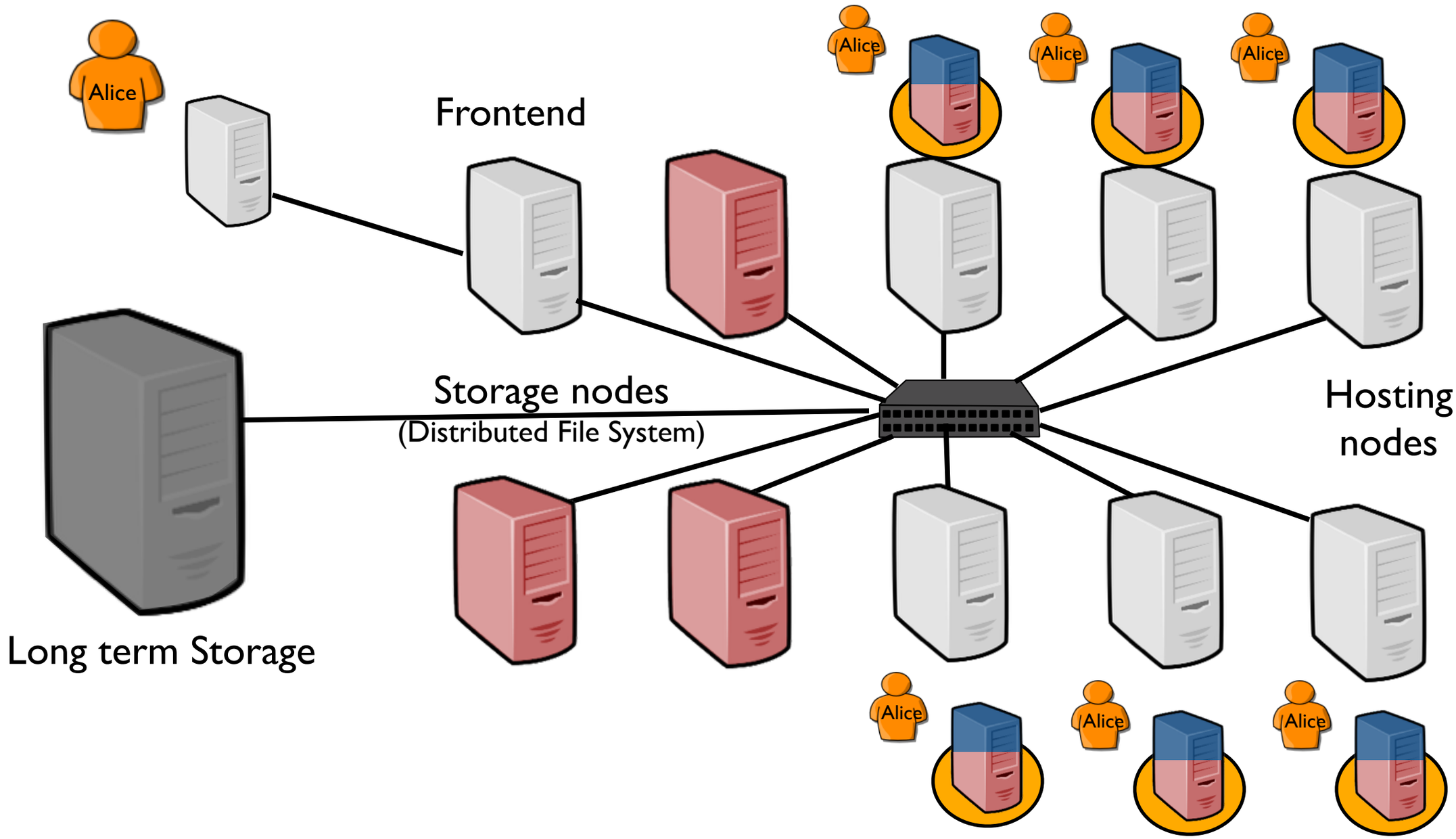
Another with Map/Reduce



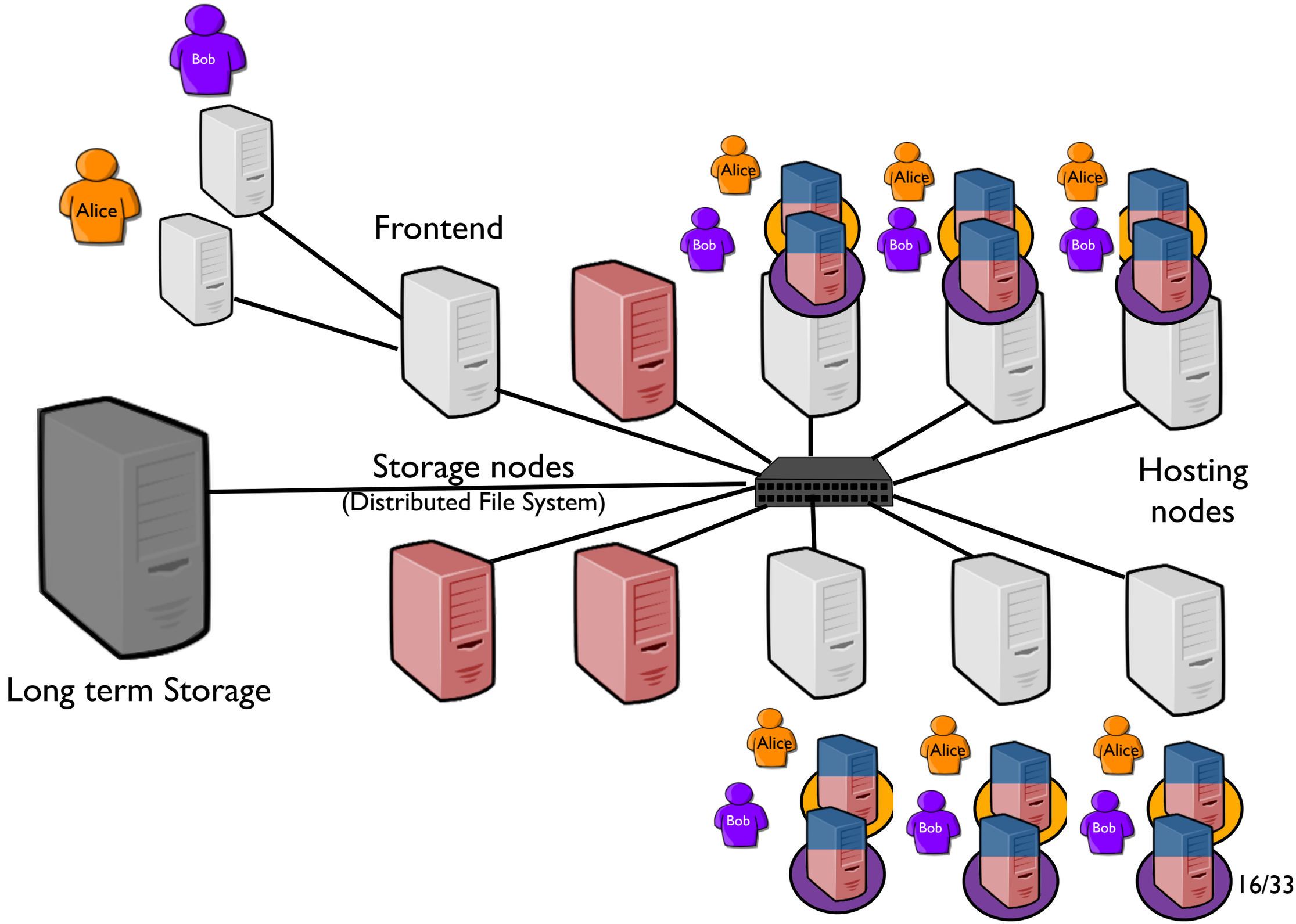
Another with Map/Reduce



Another with Map/Reduce



Another with Map/Reduce



Virtualization and Performance

- Virtualization

Contextualization / portability / "isolation"

Hard to guarantee (reproductible) performances

- Scheduling

  Mainly static ⇒ lead to energy wastes

  Dynamic scheduling strategies
  ⇒ Good achievements but still "food" for researchers

- Other approaches : application / nested virtualisation

VMs make the control of performance harder but clouds are so good….

Anything else ?

The Current Trend: Large off shore DCs

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

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

The Current Trend: Large off shore DCs

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

credits: google map - Quincy

The Current Trend: Large off shore DCs

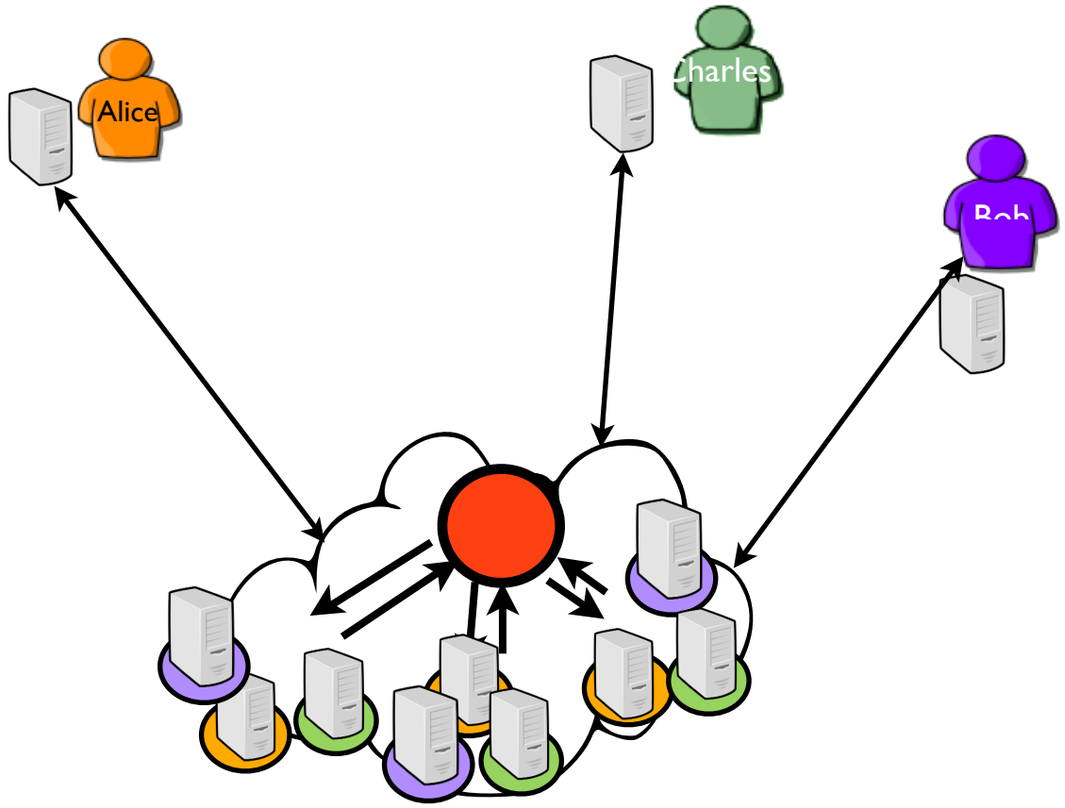
- To cope with the increasing UC demand while handling energy concerns but…

credits: coloandcloud.com

Inherent limitations of current solutions

- Large off shore DCs to cope with the increasing UC demand while handling energy concerns but…

  1. Externalization of private applications/data (jurisdiction concerns, PRISM NSA scandal, Patriot Act)

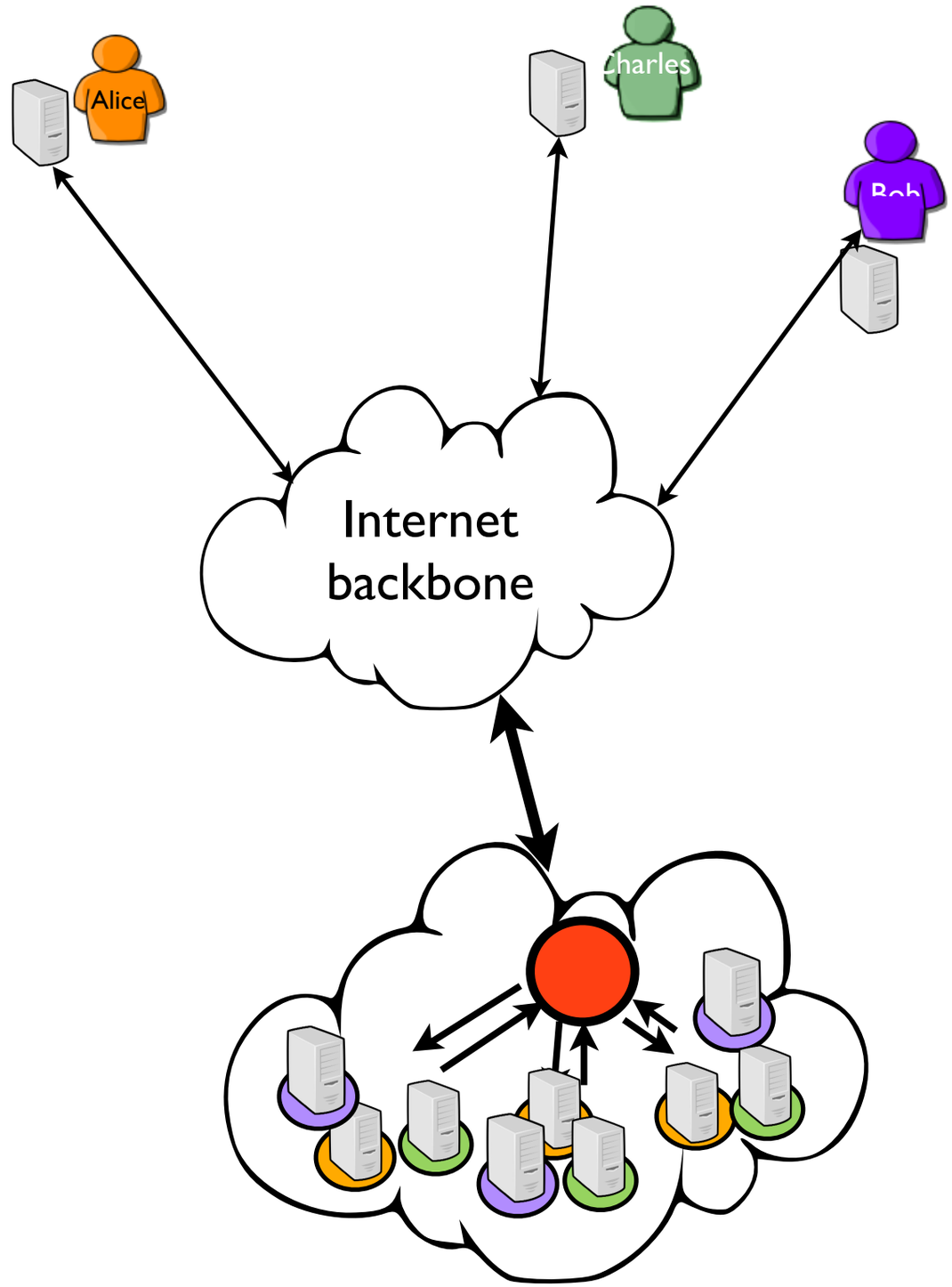


Inherent limitations of current solutions

- Large off shore DCs to cope with the increasing UC demand while handling energy concerns but…

  1. Externalization of private applications/data (jurisdiction concerns, PRISM NSA scandal, Patriot Act)

  2. Overhead implied by the unavoidable use of the Internet to reach distant platforms



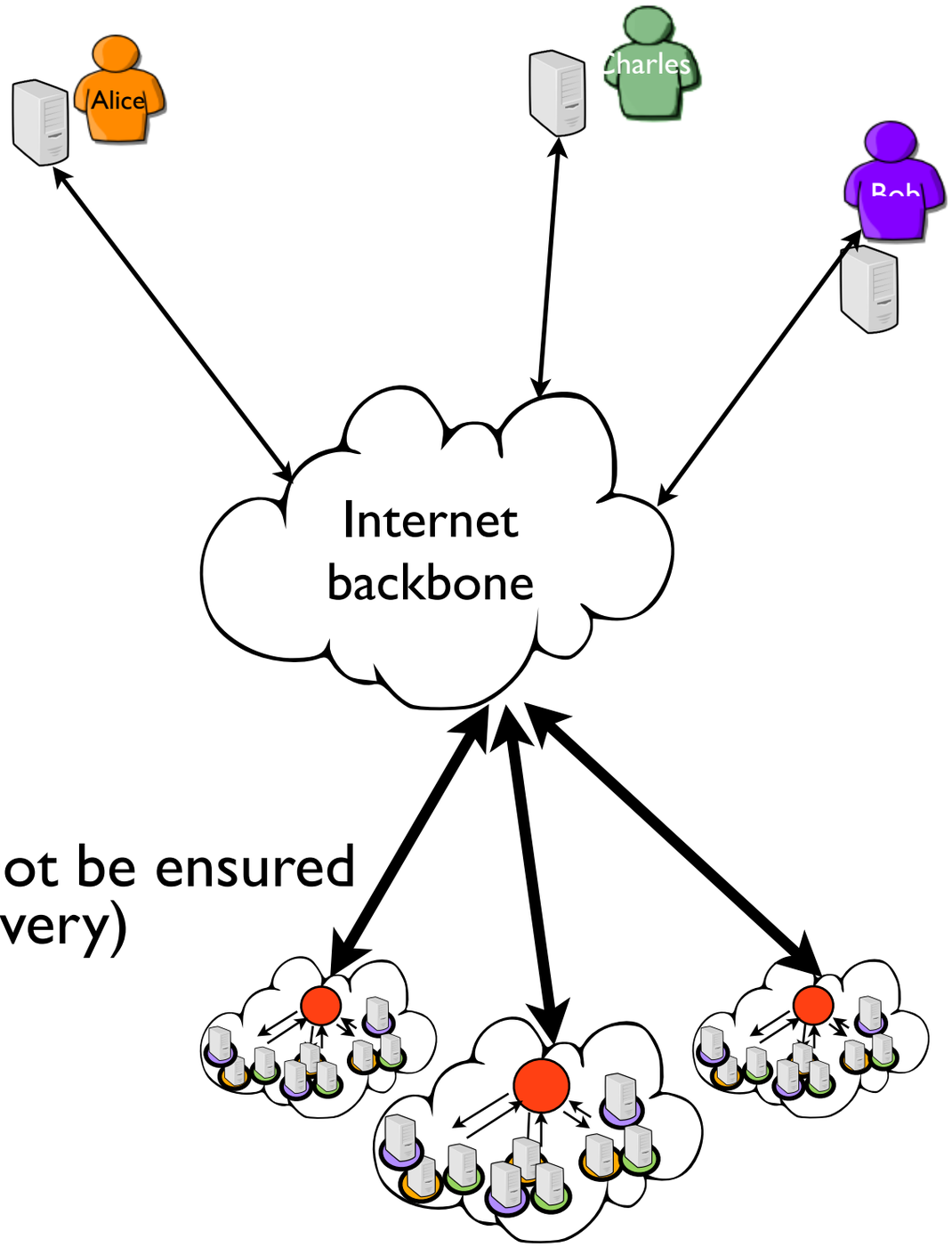
Inherent limitations of current solutions

- Large off shore DCs to cope with the increasing UC demand while handling energy concerns but…

1. Externalization of private applications/data (jurisdiction concerns, PRISM NSA scandal, Patriot Act)

2. Overhead implied by the unavoidable use of the Internet to reach distant platforms

3. The connectivity to the application/data cannot be ensured by centralized dedicated centers (disaster recovery)



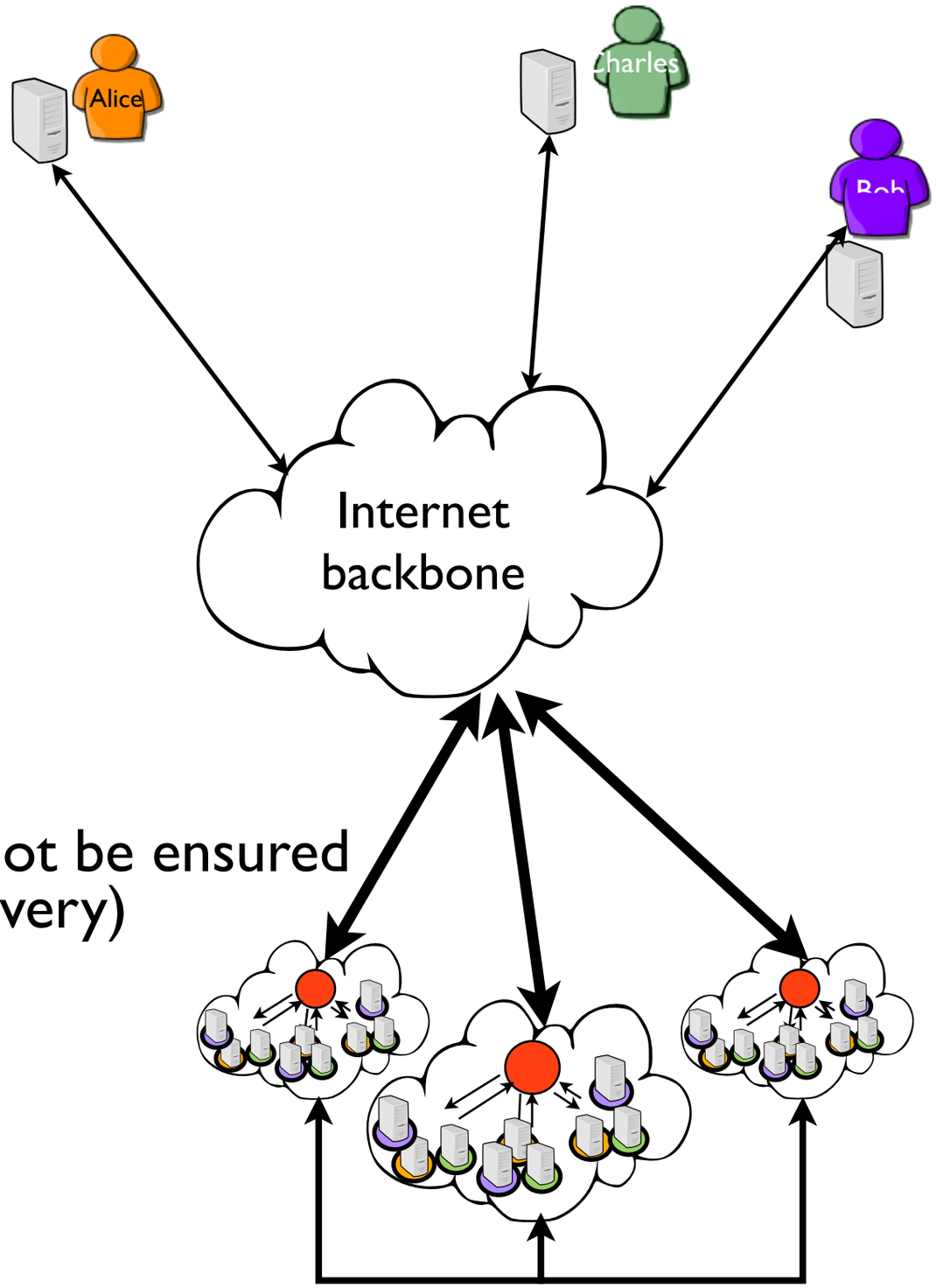
Inherent limitations of current solutions

- Large off shore DCs to cope with the increasing UC demand while handling energy concerns but…

1. Externalization of private applications/data (jurisdiction concerns, PRISM NSA scandal, Patriot Act)

2. Overhead implied by the unavoidable use of the Internet to reach distant platforms

3. The connectivity to the application/data cannot be ensured by centralized dedicated centers (disaster recovery)



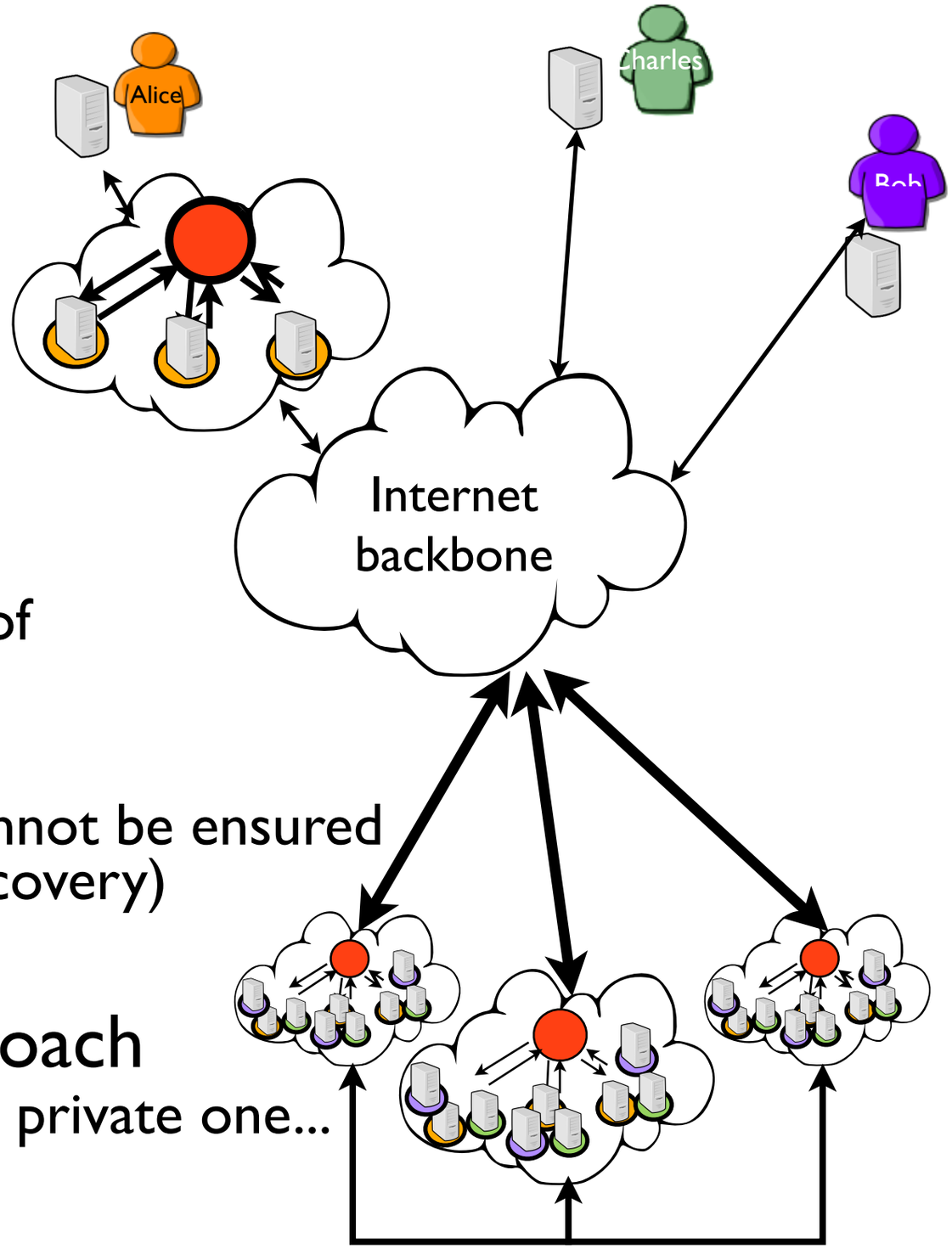
Inherent limitations of current solutions

- Large off shore DCs to cope with the increasing UC demand while handling energy concerns but…

  1. Externalization of private applications/data (jurisdiction concerns, PRISM NSA scandal, Patriot Act)

  2. Overhead implied by the unavoidable use of the Internet to reach distant platforms

  3. The connectivity to the application/data cannot be ensured by centralized dedicated centers (disaster recovery)



- Hybrid platforms: a promising approach

  It depends how you are going to extend the private one...

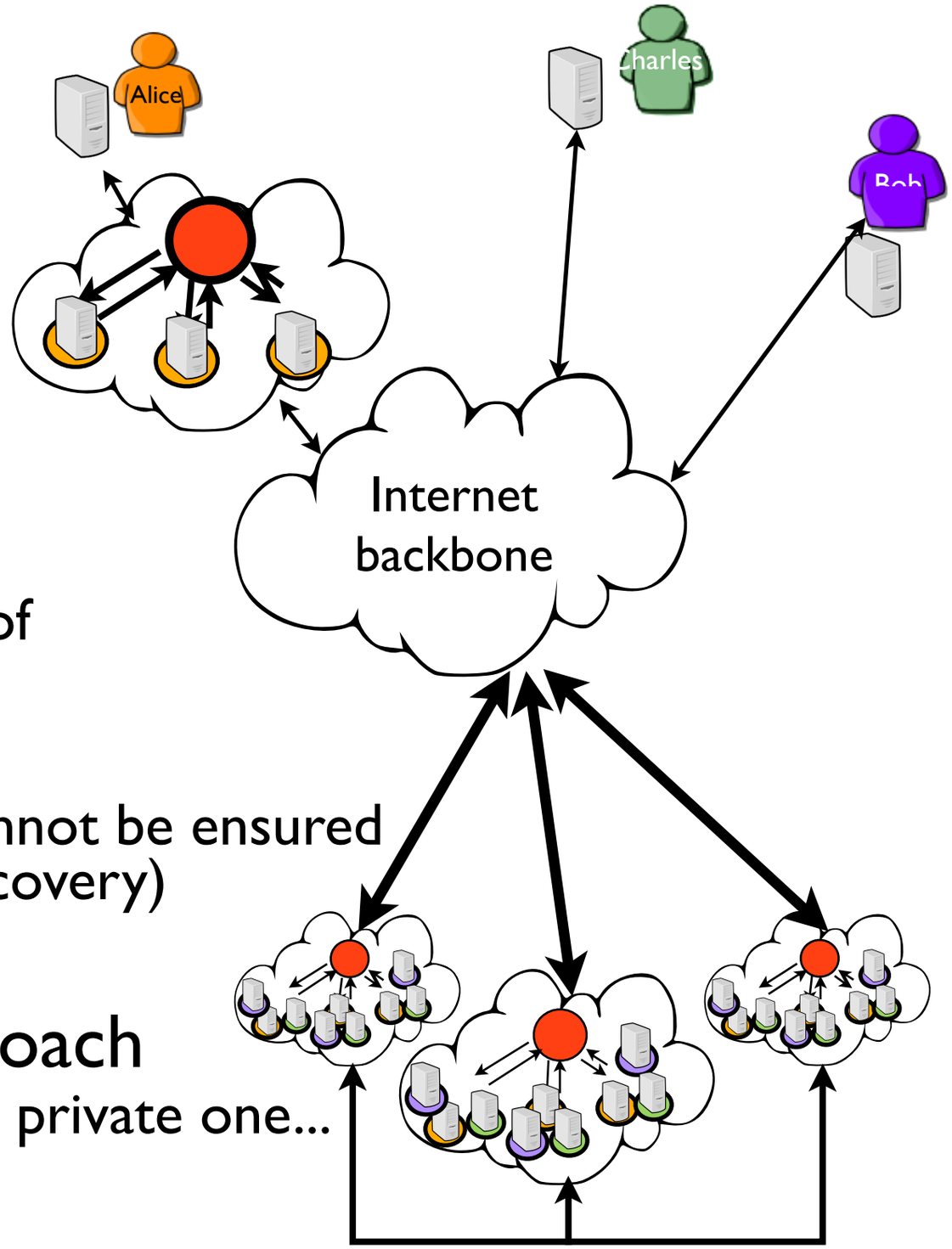
Inherent limitations of current solutions

- Large off shore DCs to cope with the increasing UC demand while handling energy concerns but…

  1. Externalization of private applications/data (jurisdiction concerns, PRISM NSA scandal, Patriot Act)

  2. Overhead implied by the unavoidable use of the Internet to reach distant platforms

  3. The connectivity to the application/data cannot be ensured by centralized dedicated centers (disaster recovery)



- Hybrid platforms: a promising approach
  It depends how you are going to extend the private one...

Is there a way to address these concerns "all in one" ?

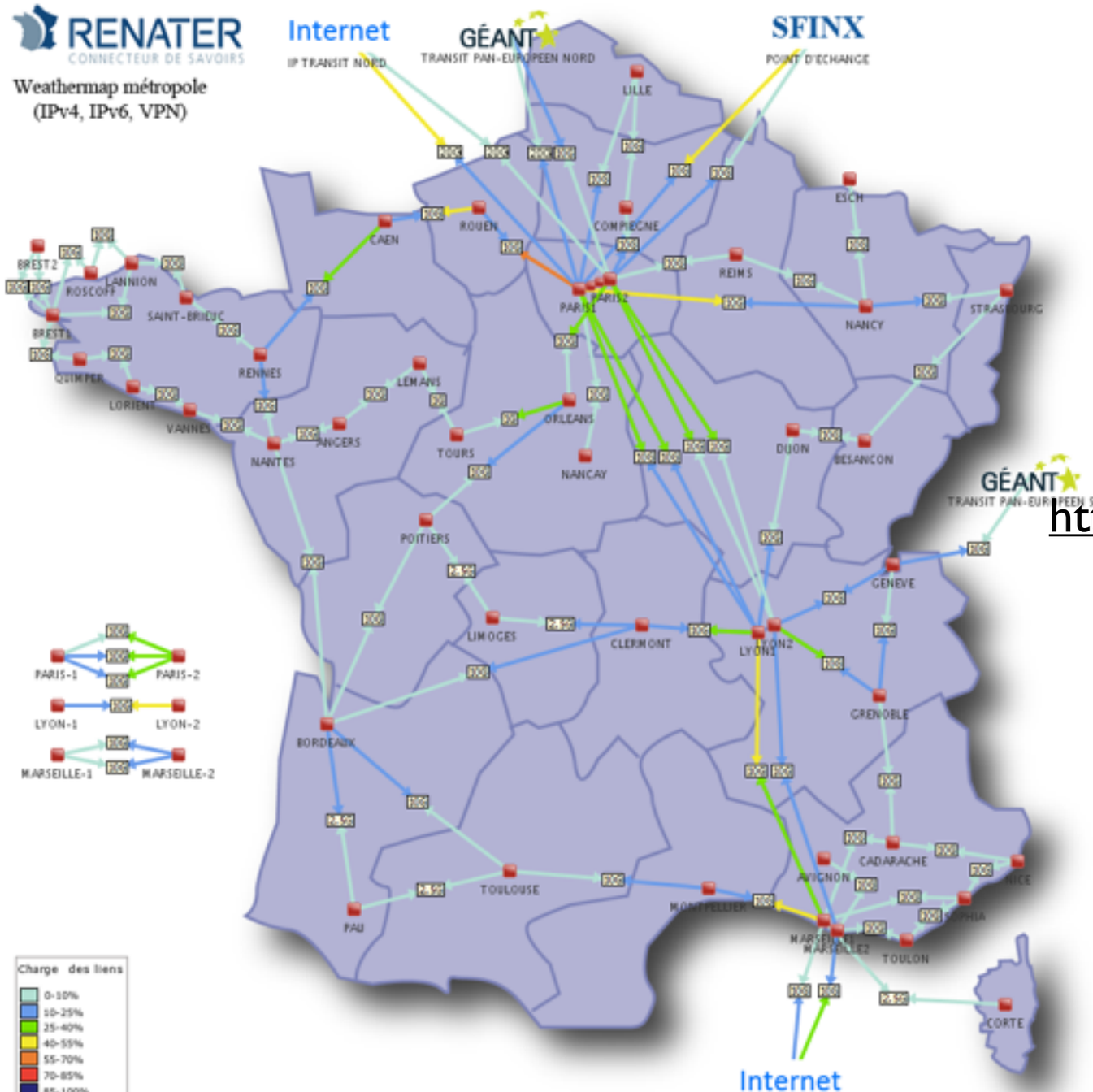
μDC at the edge !
How and where the μDC concept can be deployed ?

Locality Based Utility Computing Toward LUC Infrastructures

Beyond the Clouds, the DISCOVERY Initiative

- Locality-based UC infrastructures

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

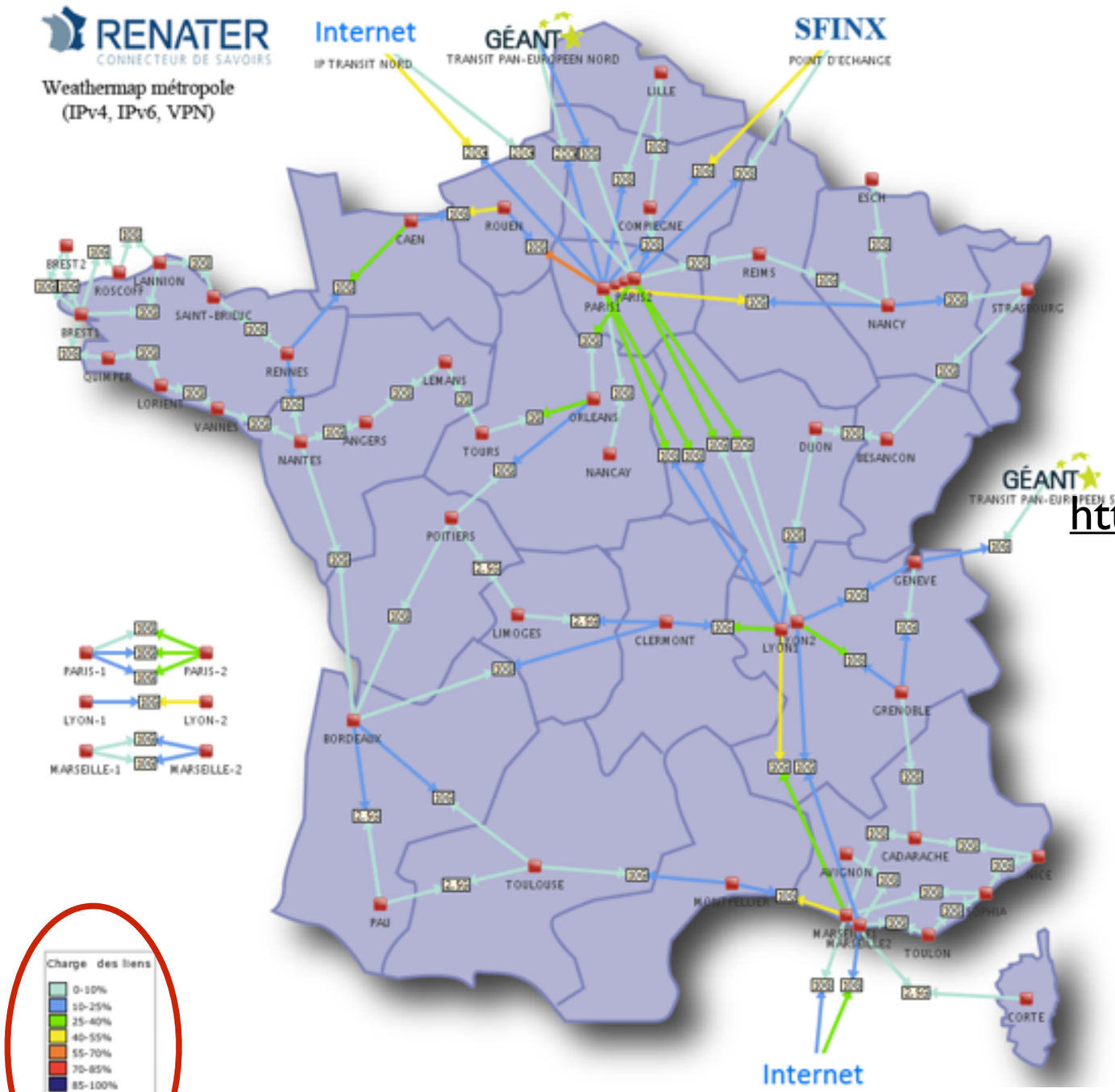


http://www.renater.fr/raccourci?lang=fr

Beyond the Clouds, the DISCOVERY Initiative

- Locality-based UC infrastructures

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

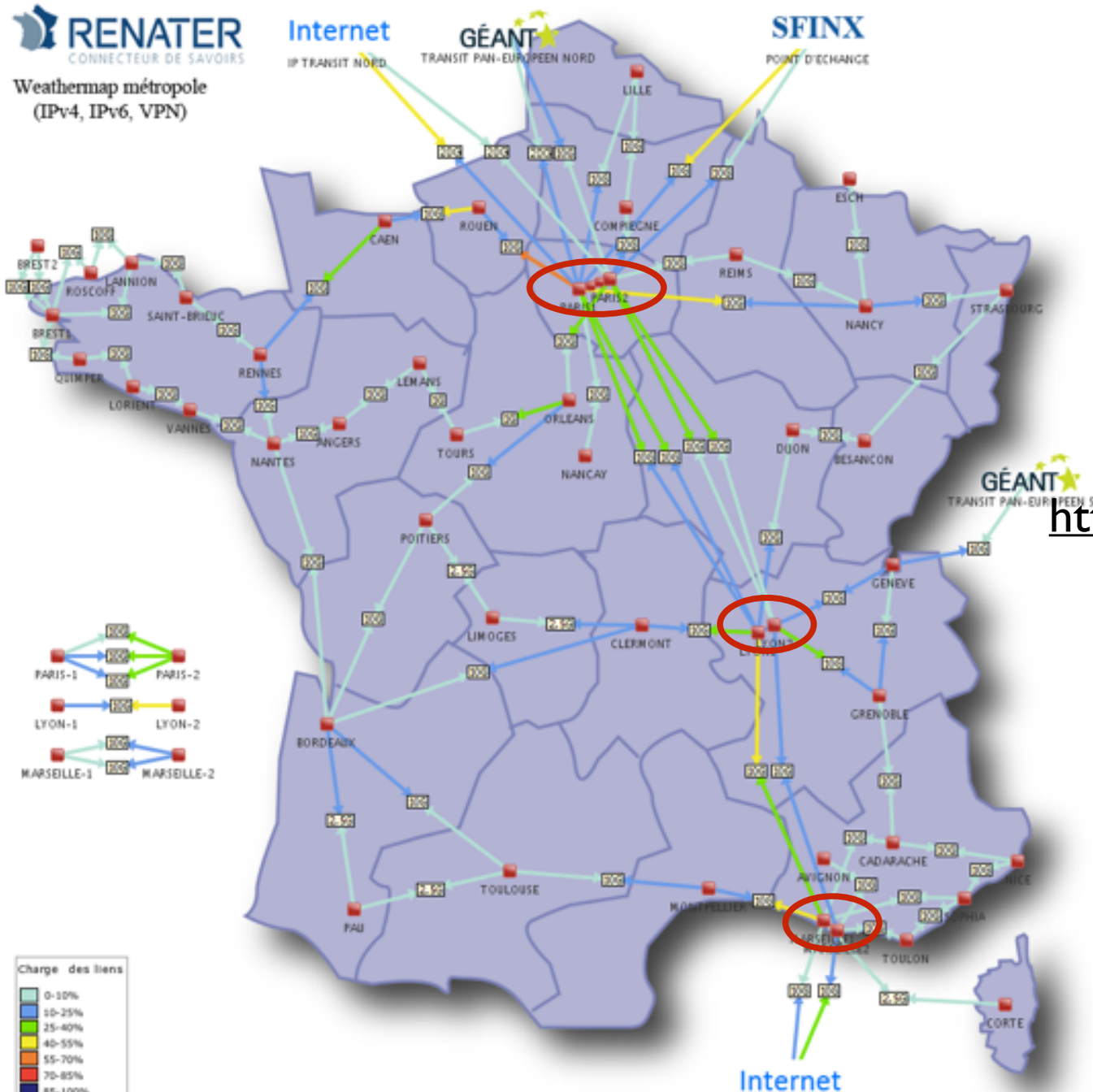


http://www.renater.fr/raccourci?lang=fr

Beyond the Clouds, the DISCOVERY Initiative

- Locality-based UC infrastructures

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

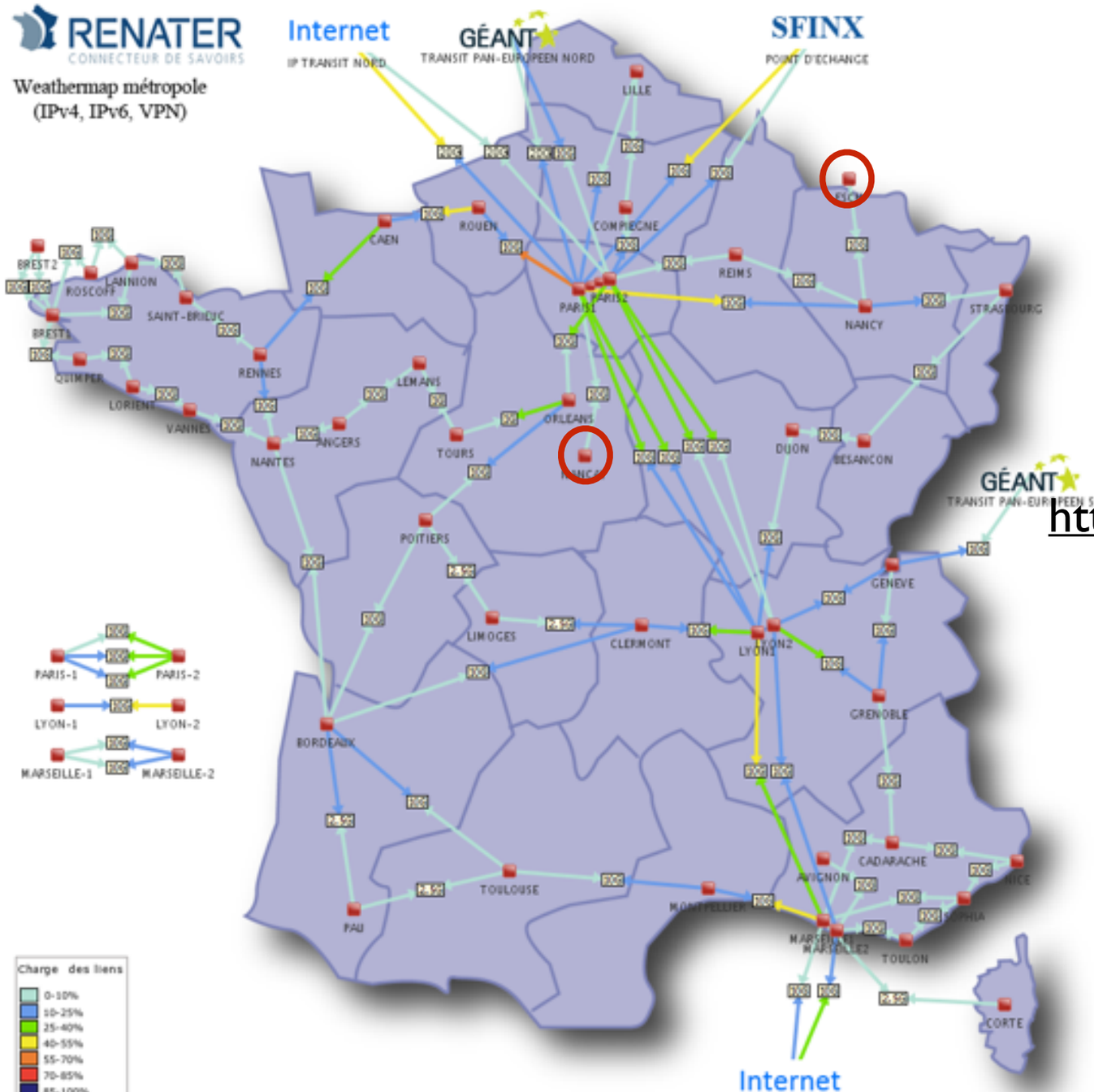


http://www.renater.fr/raccourci?lang=fr

Beyond the Clouds, the DISCOVERY Initiative

- Locality-based UC infrastructures

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

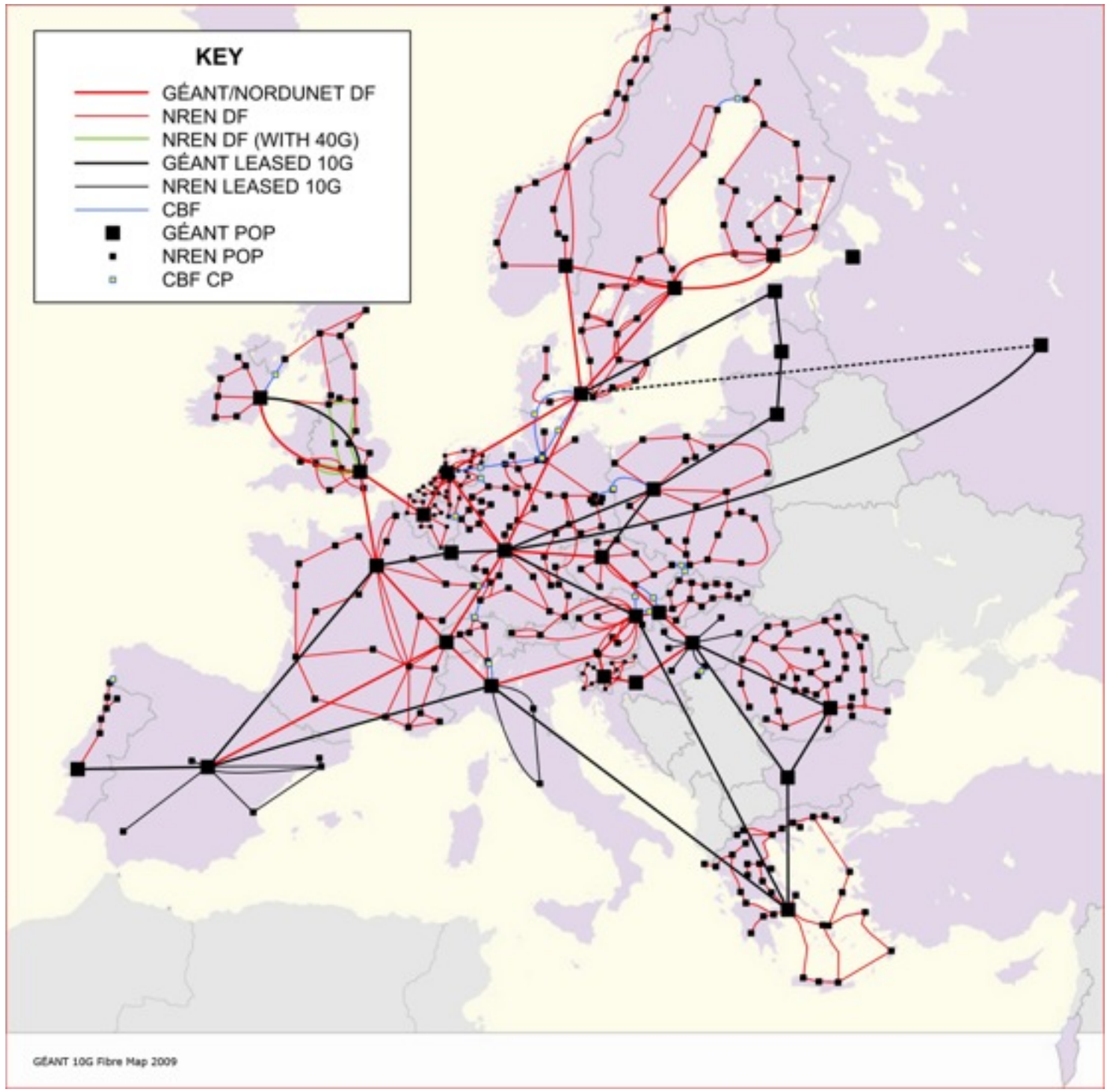


http://www.renater.fr/raccourci?lang=fr

Beyond the Clouds, the DISCOVERY Initiative

- Locality-based UC infrastructures

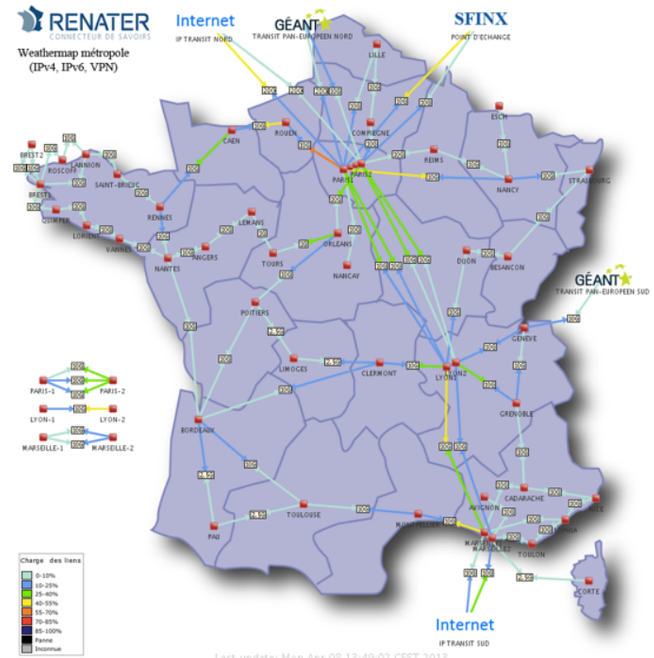
  A promising way to deliver highly efficient and sustainable UC services is to provide UC platforms as close as possible to the end-users.



Beyond the Cloud, the DISCOVERY Initiative

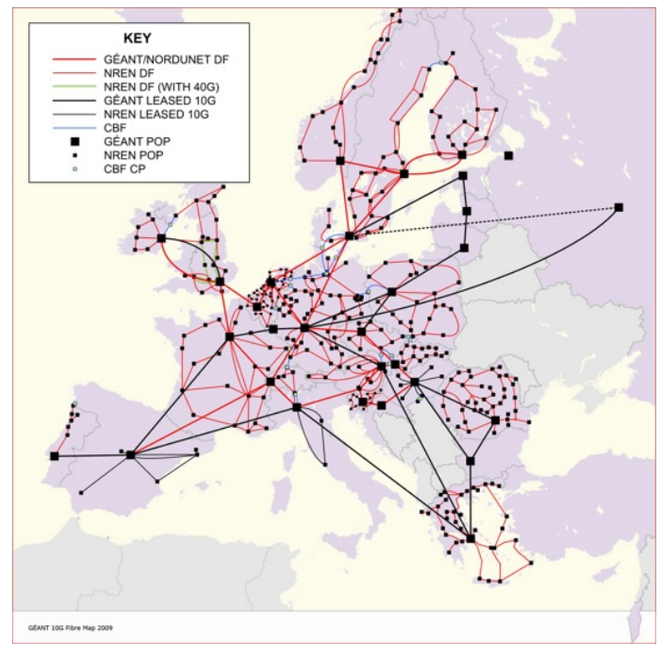
- Locality-based UC infrastructures

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



- Leveraging network backbones

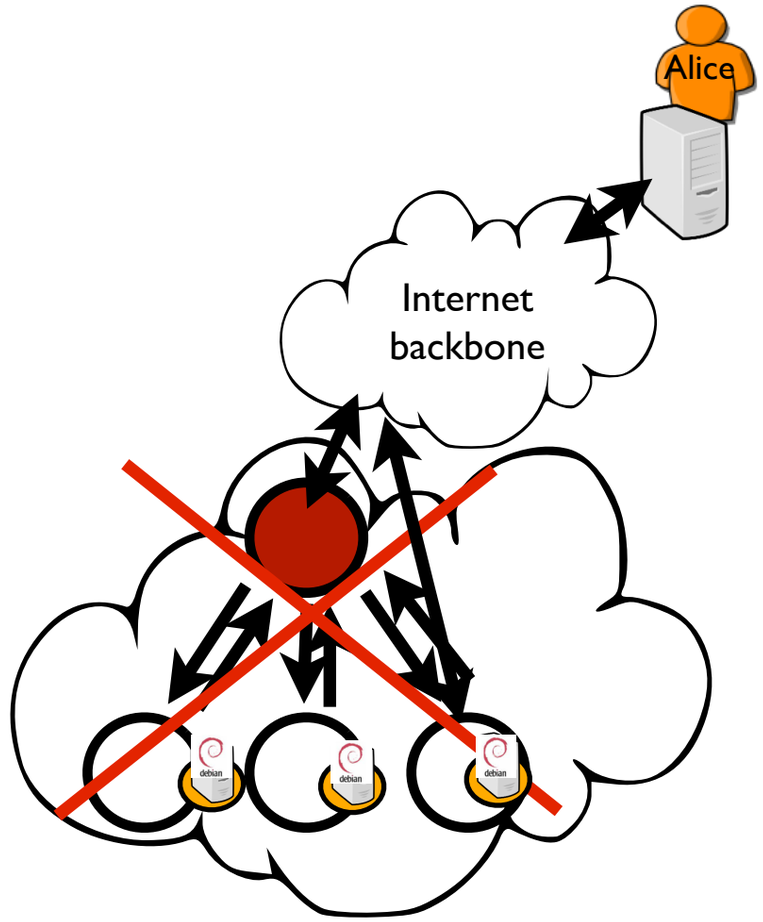
  Extend any point of presence of network backbones with UC servers (from network hubs up to major DSLAMs that are operated by telecom companies and network institutions).



⇒ Operating such widely distributed resources requires the definition of a fully distributed system

The DISCOVERY Proposal

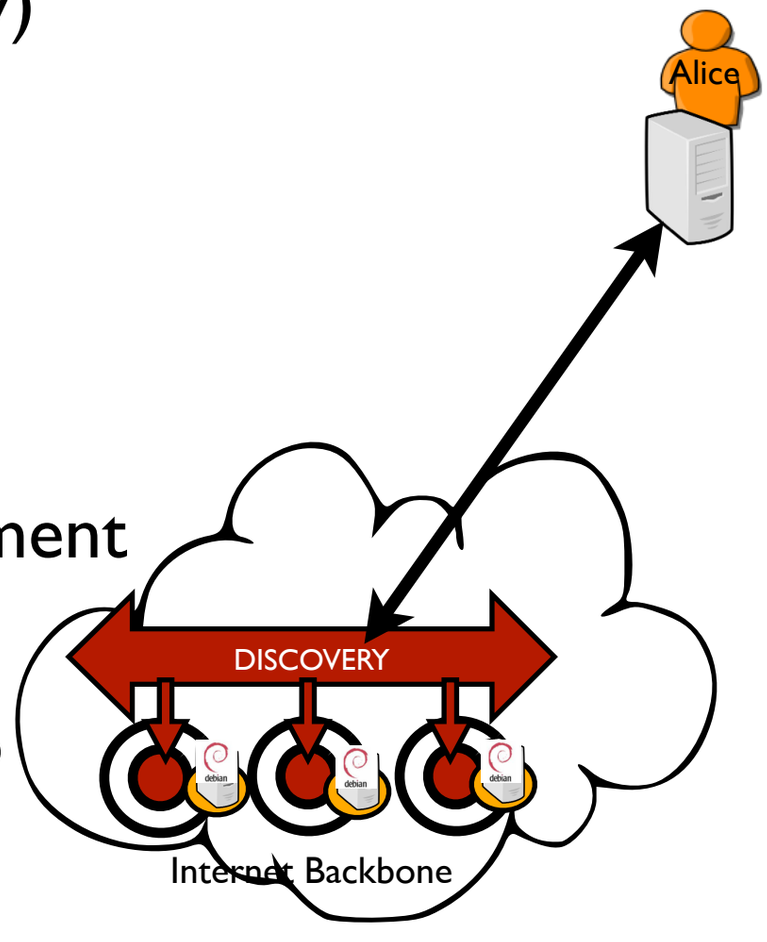
- DIStributed and COoperative framework to manage Virtual EnviRonments autonomously



The DISCOVERY Proposal

- DIStributed and COoperative framework to manage Virtual EnviRonments autonomously
- The LUC OS
  - A fully distributed IaaS system and not a distributed system of IaaS systemS. We want to/must go further than high level cloud APIs (cross-cutting concerns such as energy/security)
  - Leverage P2P algorithms and self-* approaches
- lots of scientific/technical challenges

Cost of the network !? partial view of the system !? Impact on the others VMs !?, management of VM images !? Which software abstractions to make the development easier and more reliable (distributed event programming)? How to take into account locality aspects? …



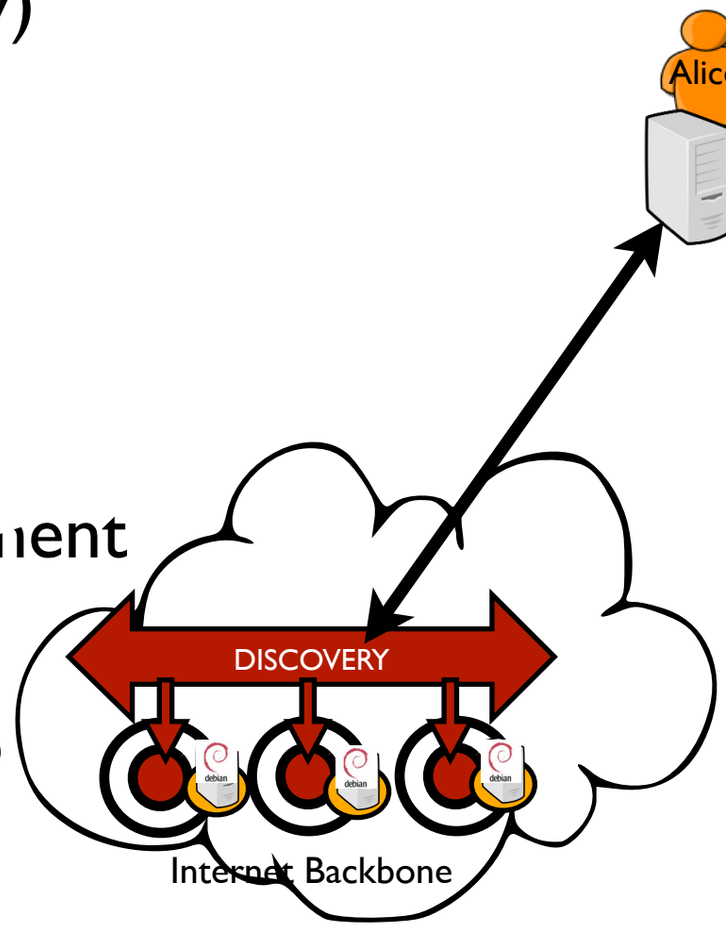
The DISCOVERY Proposal

- DIStributed and COoperative framework to manage Virtual EnviRonments autonomously
- The LUC OS
  - A fully distributed IaaS system and not a distributed system of IaaS systemS. We want to/must go further than high level cloud APIs (cross-cutting concerns such as energy/security)
  - Leverage P2P algorithms and self-*
- lots of scientific/tec[illegible]

Cost [illegible]

?? A distributed version of the EGI Core that directly manipulates resources http://www.egi.eu/infrastructure/cloud/ ??

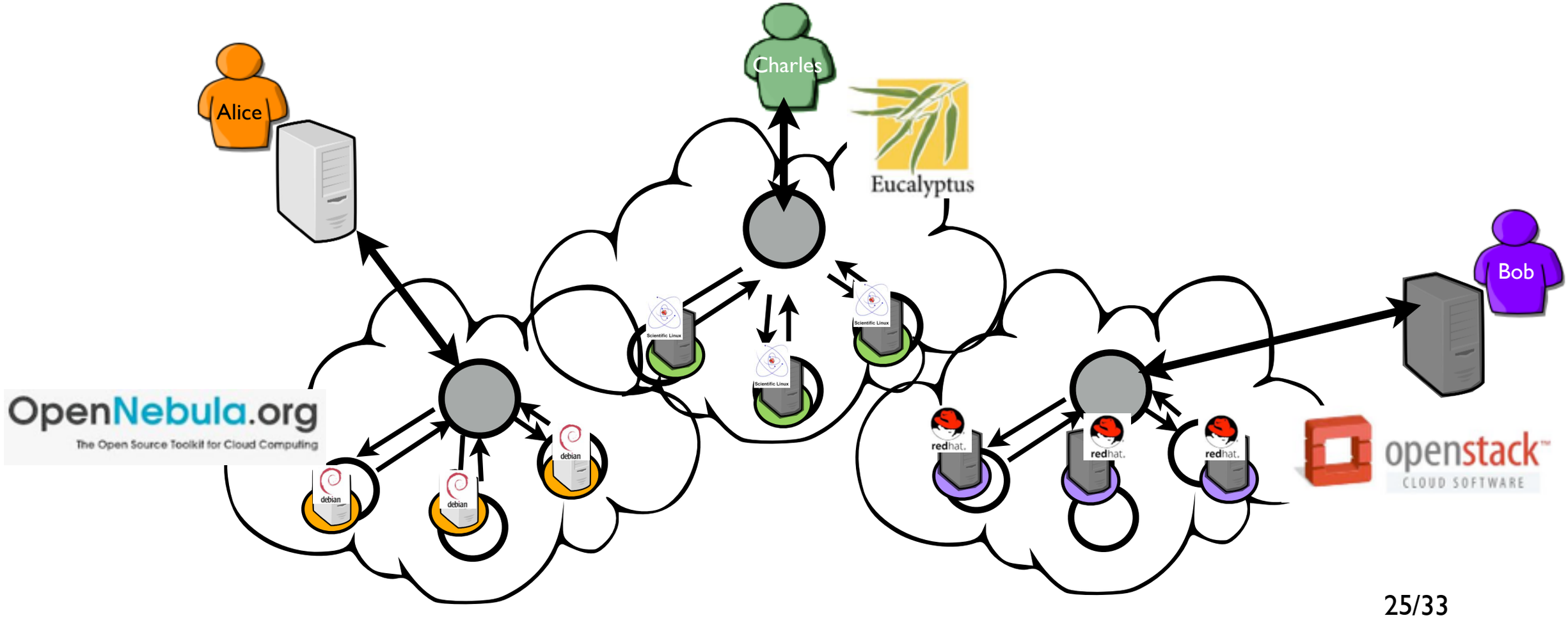
[illegible]ement
[illegible]ns to
[illegible]ore reliable
[illegible]ng)? How to take into
[illegible]ts? …



Why not a broker ?

- “federation of clouds” (sky computing)

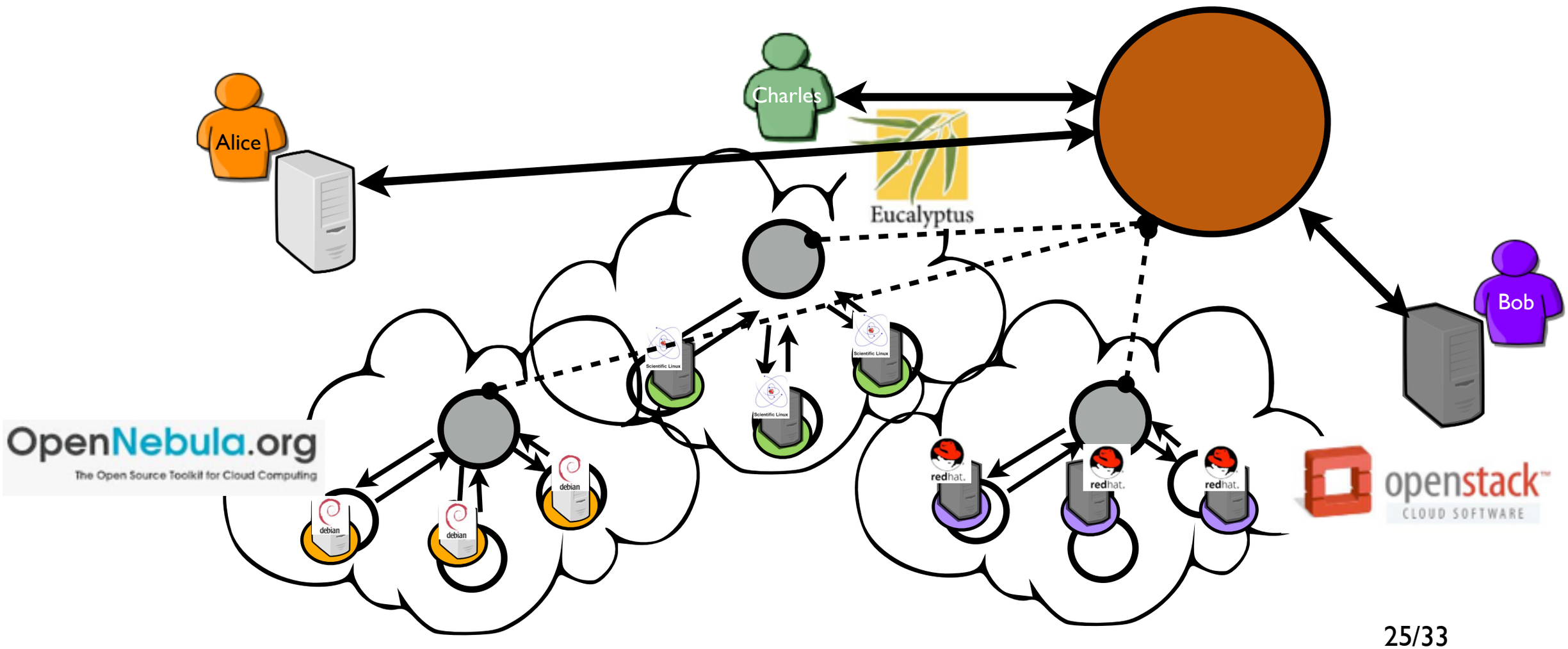
  Sporadic (hybrid computing/cloud bursting) almost ready for production
  While standards are coming (OCCI, OVF, ….), current brokers are rather limited



Why not a broker ?

- "federation of clouds" (sky computing)

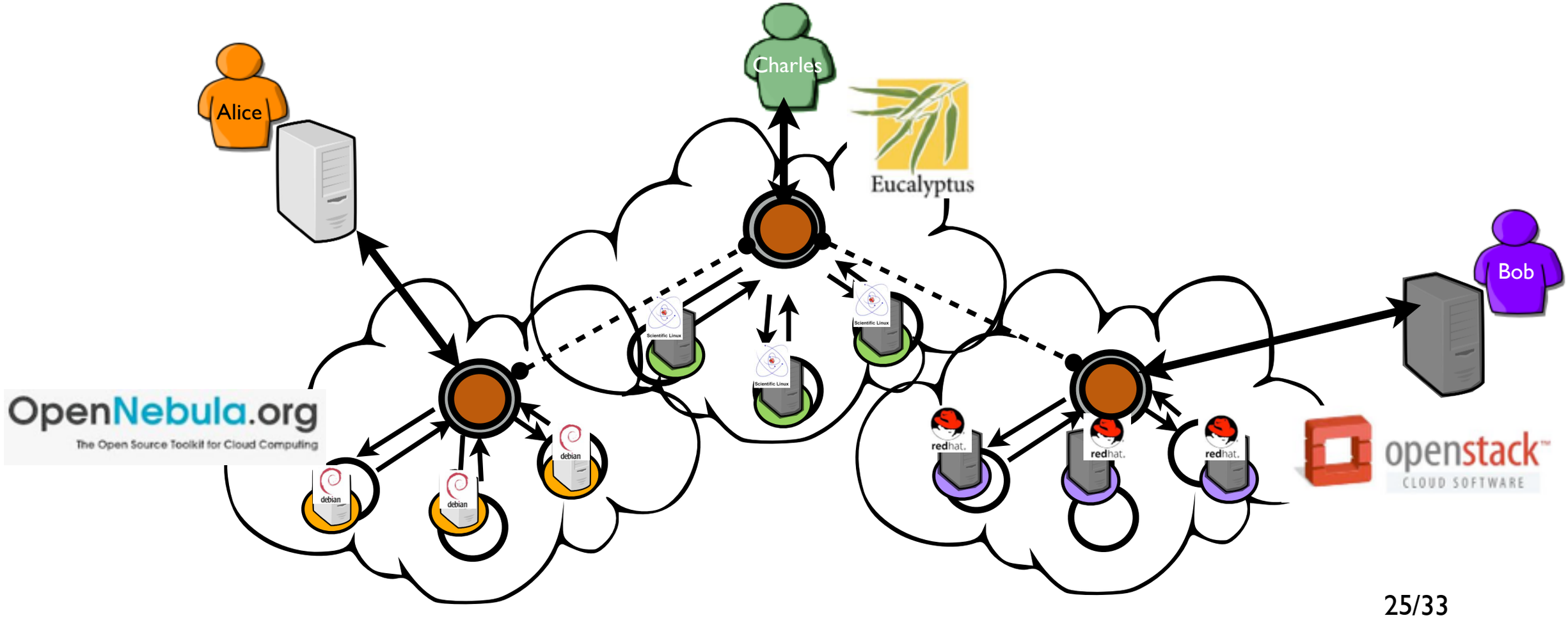
  Sporadic (hybrid computing/cloud bursting) almost ready for production
  While standards are coming (OCCI, OVF, ….), current brokers are rather limited



Why not a broker ?

- “federation of clouds” (sky computing)

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

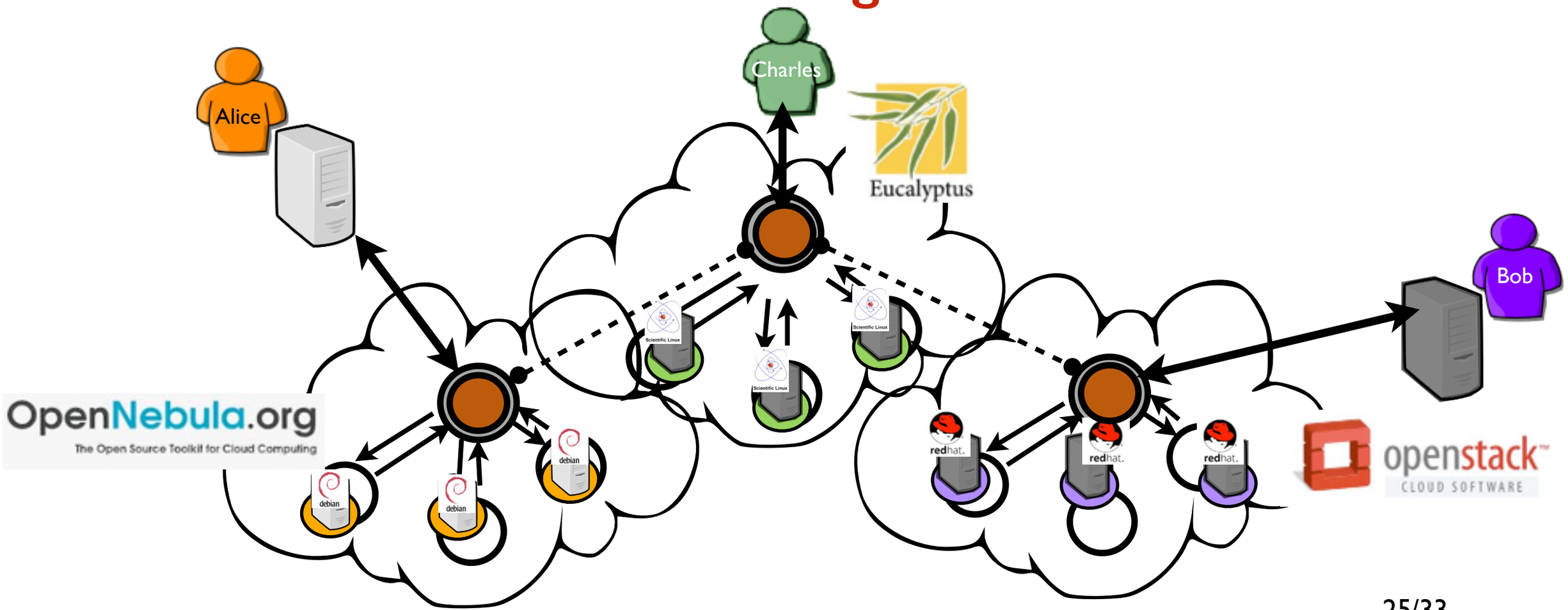


Why not a broker ?

- "federation of clouds" (sky computing)

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

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



The Discovery Initiative Pros/Cons

- Pros

  Locality (jurisdiction concerns, latency-aware apps, minimize network overhead)

  Reliability/redundancy (no critical point/location/center)
  The infrastructure is naturally distributed throughout multiple areas

  Lead time to delivery
  Leverage current PoPs and extend them according to UC demands

  Energy footprint (on-going investigations with RENATER)

  Bring back part of the revenue to NRENs/Telcos

- Cons

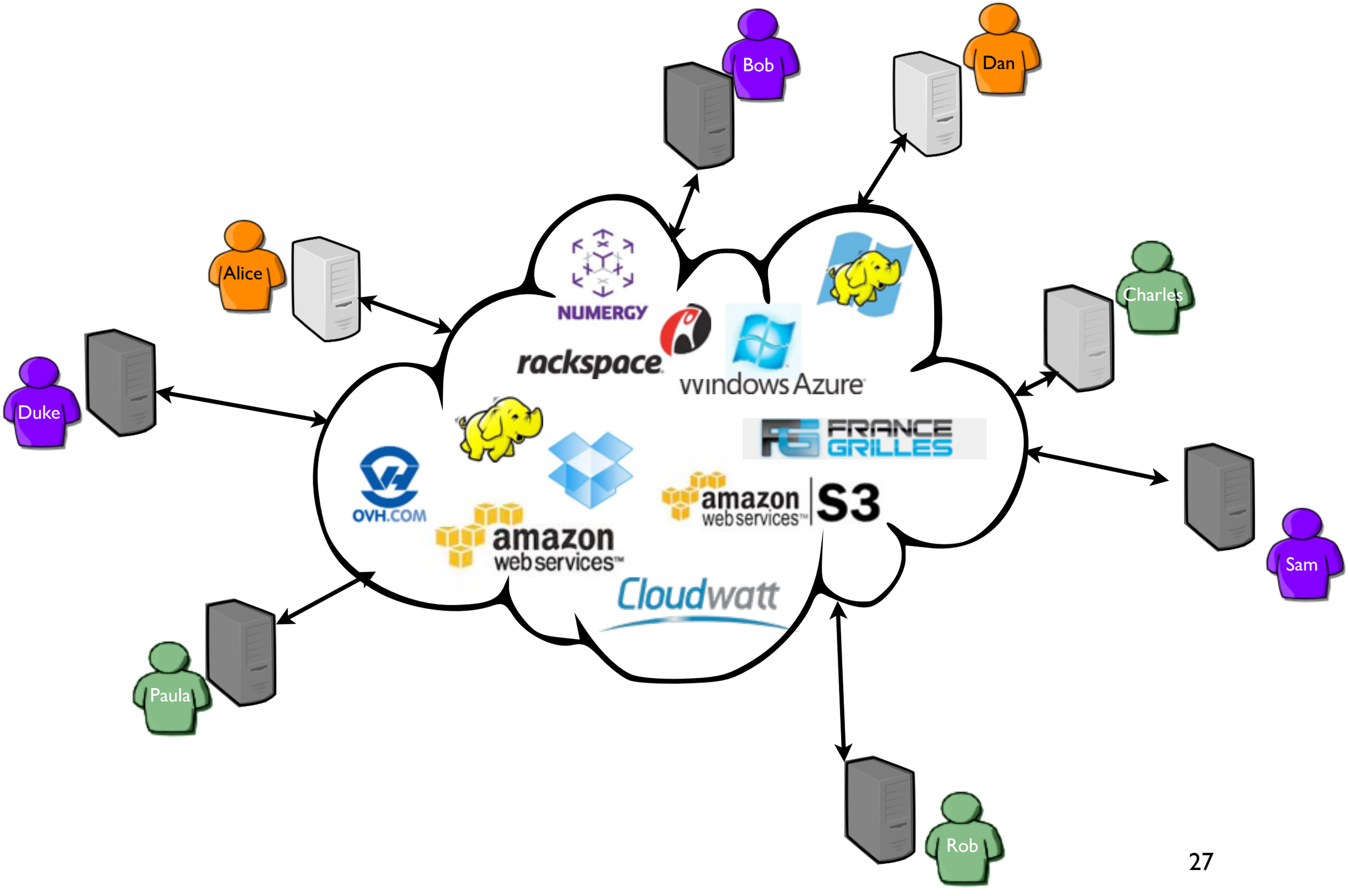
  Security concerns (in terms of who can access to the PoPs)

  Operate a fully IaaS in a unified but distributed manner at WAN level

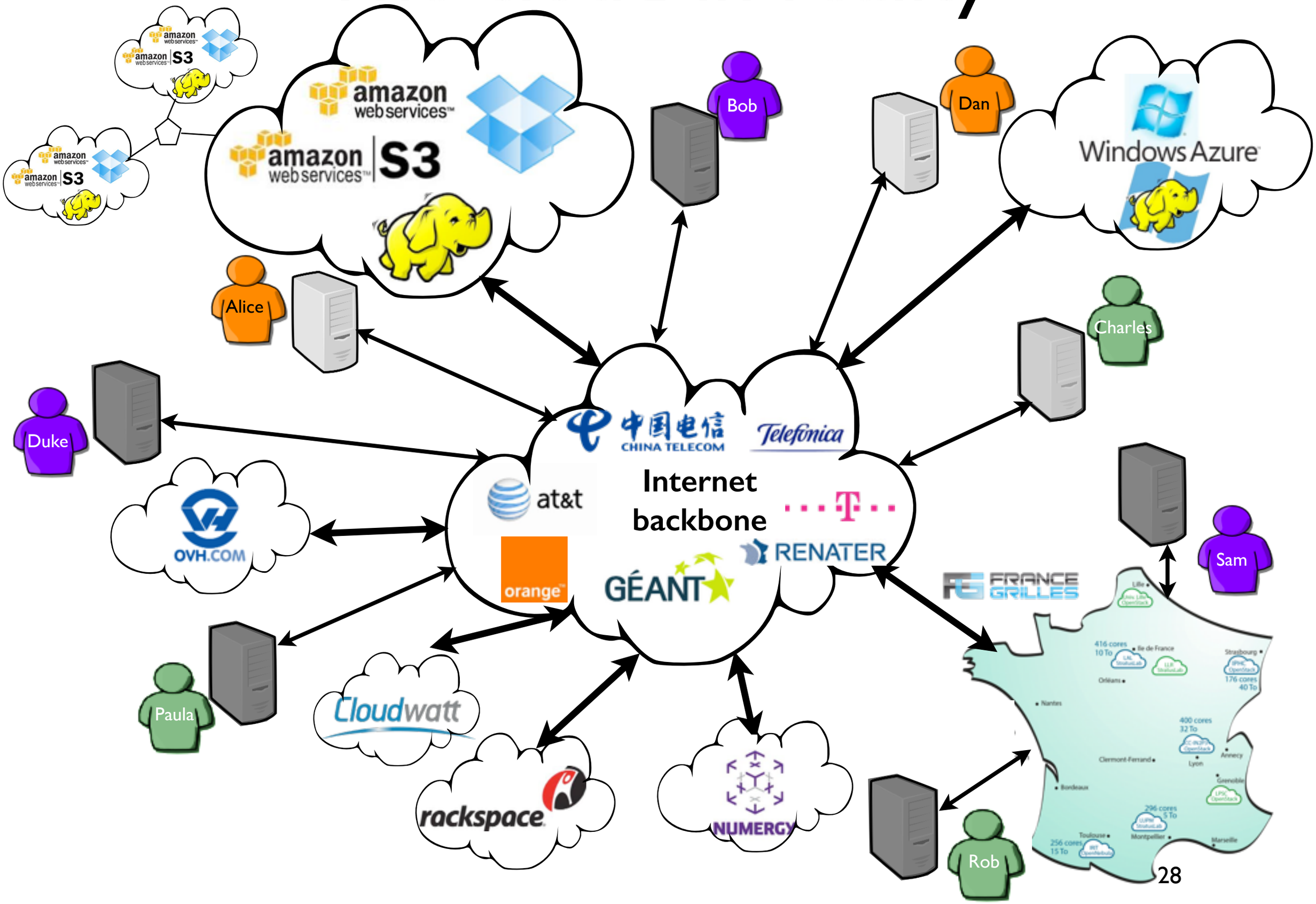
  Not suited for all kinds of applications : Large tightly coupled HPC workloads
  50 nodes/1000 cores, 200 nodes / 4000 cores (5 racks),
  so1000 nodes in one PoP does not look realistic …

  Peering agreement / economic model between network operators

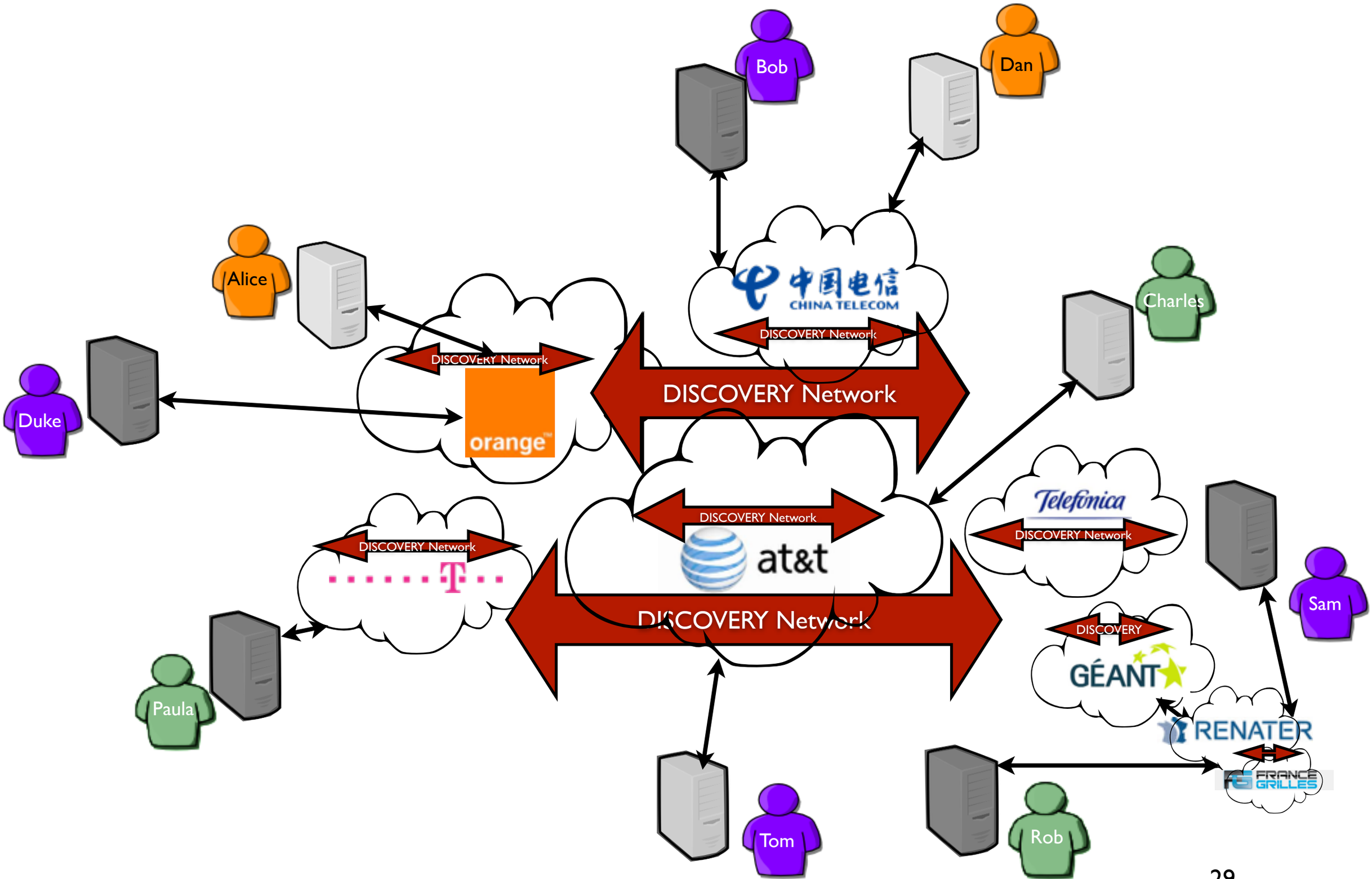
The cloud from end-users



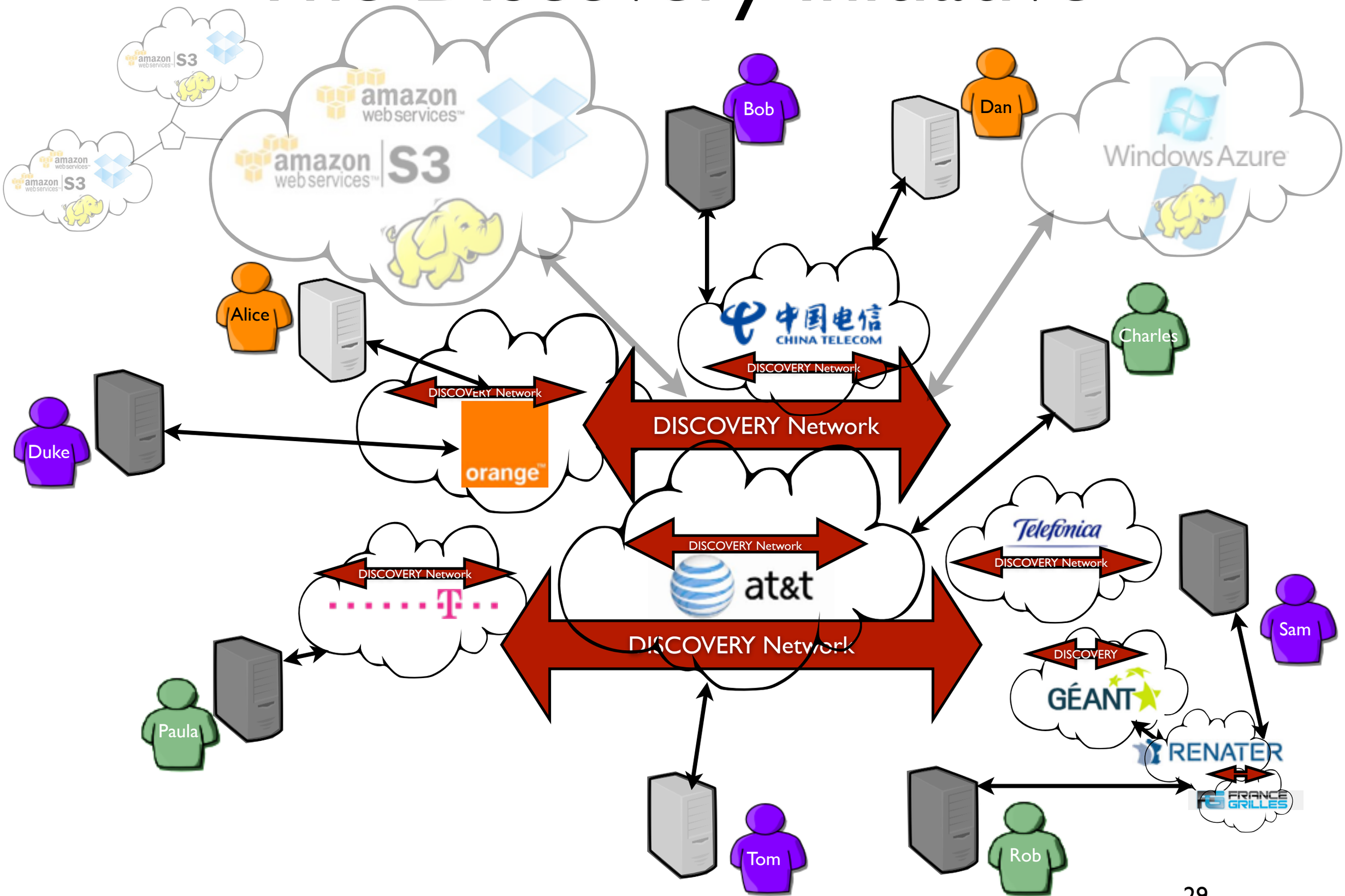
The cloud in reality



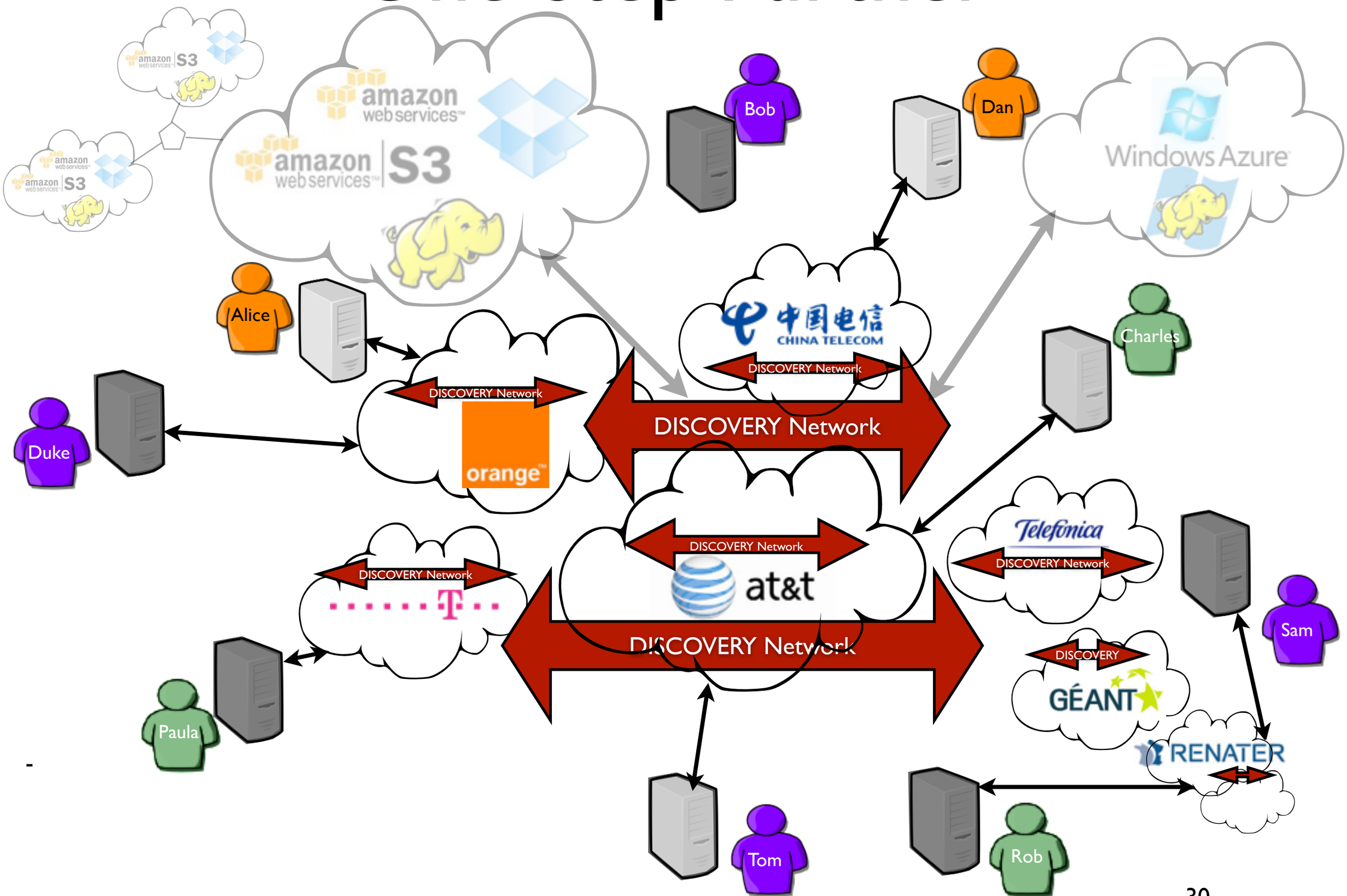
The Discovery Initiative



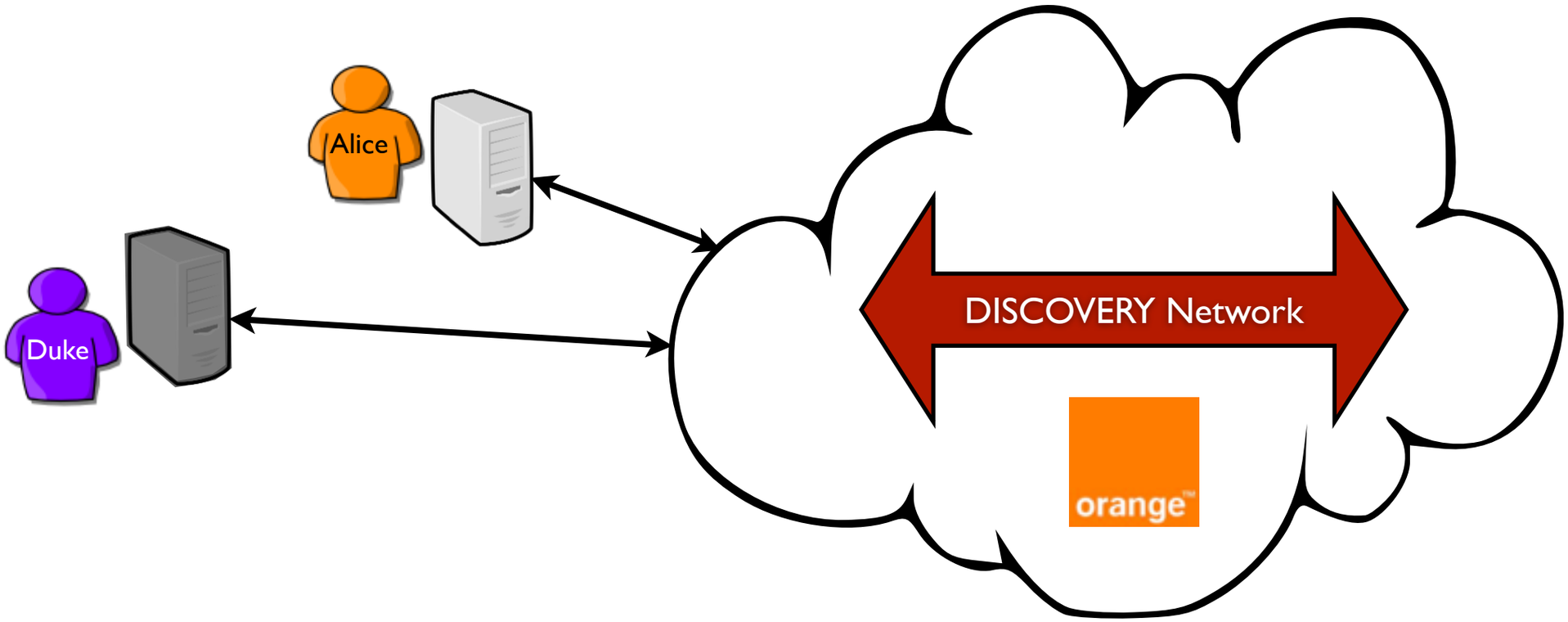
The Discovery Initiative



One Step Further

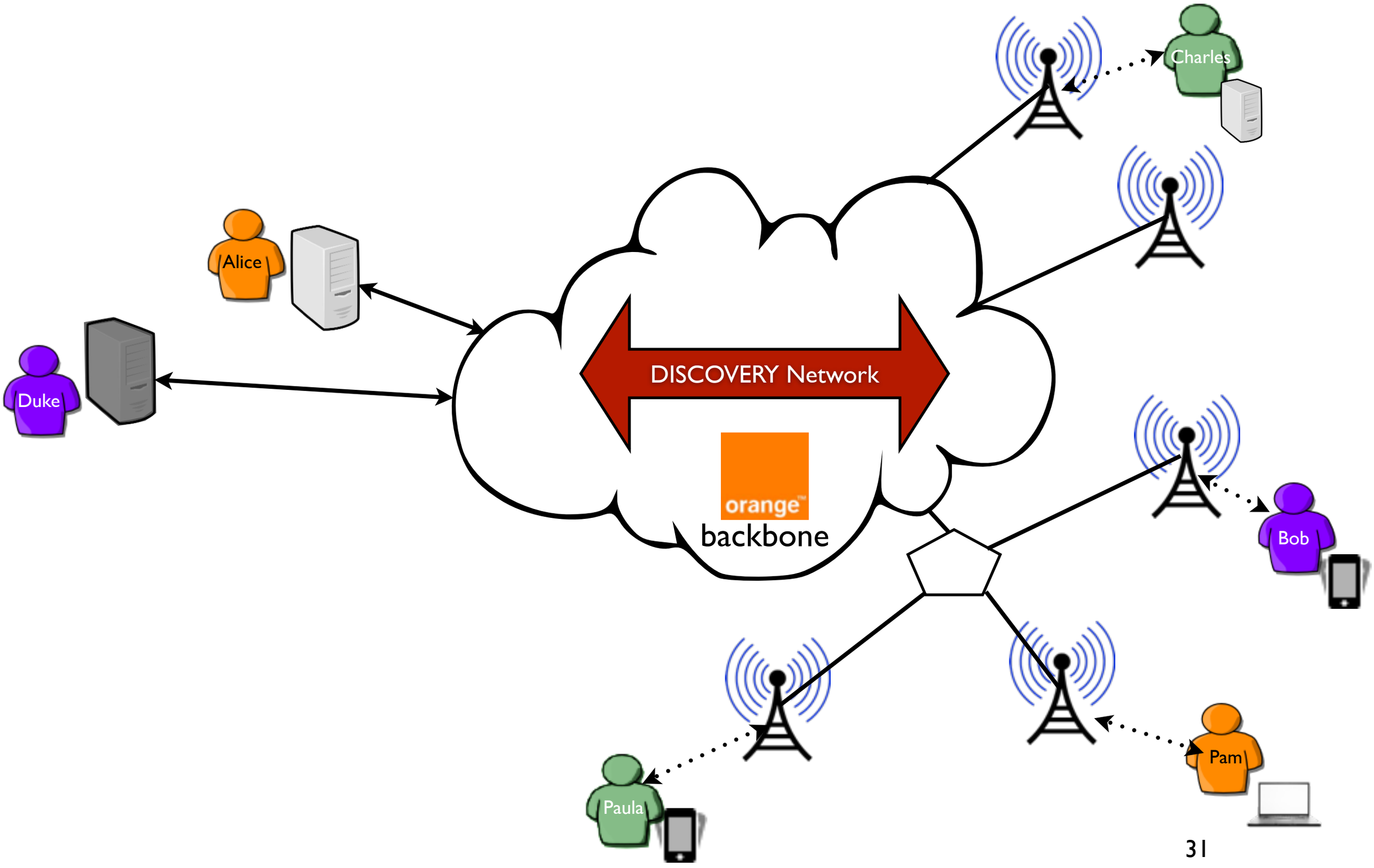


One Step Further

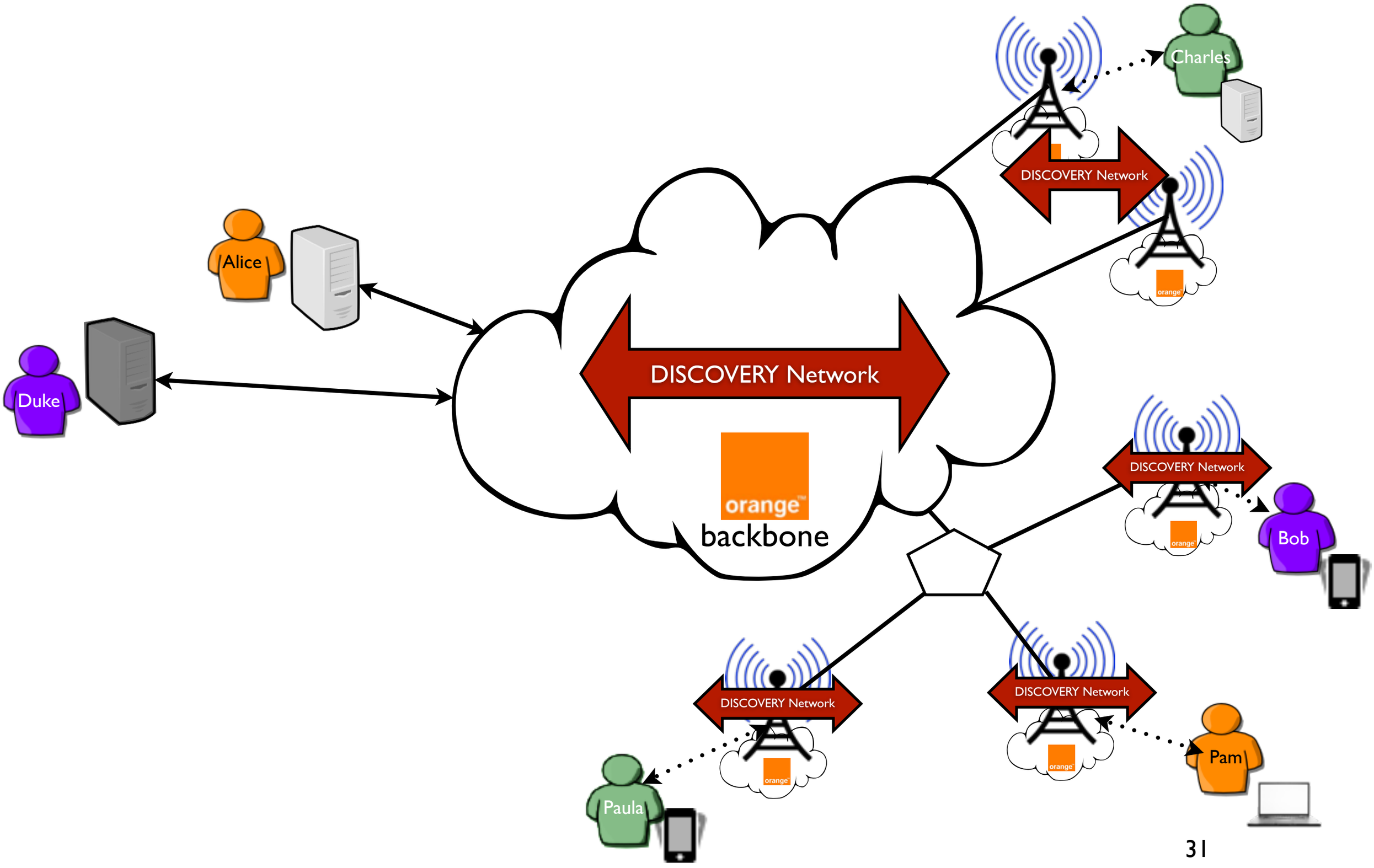


-

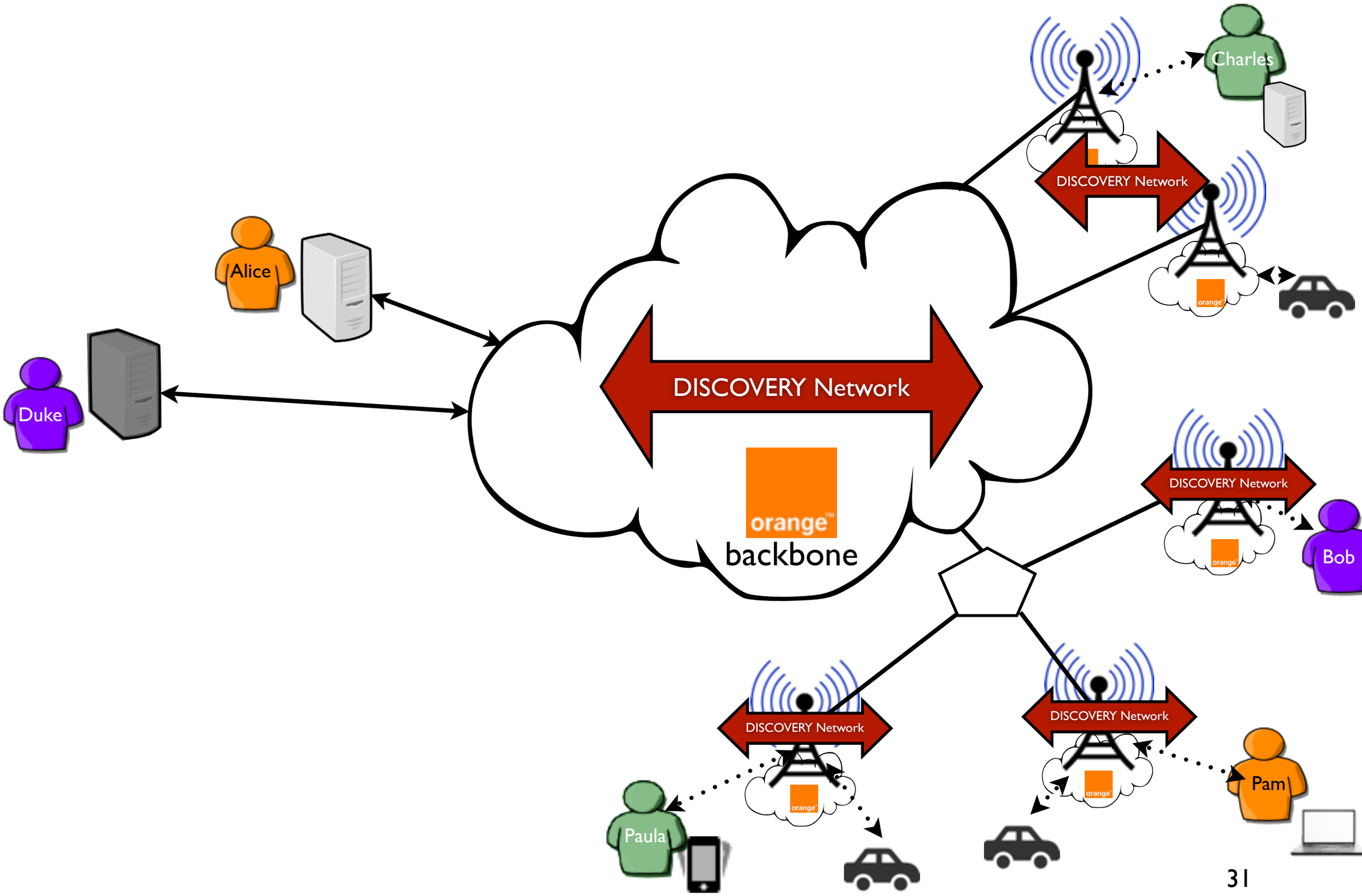
Radio Base Stations



Radio Base Stations



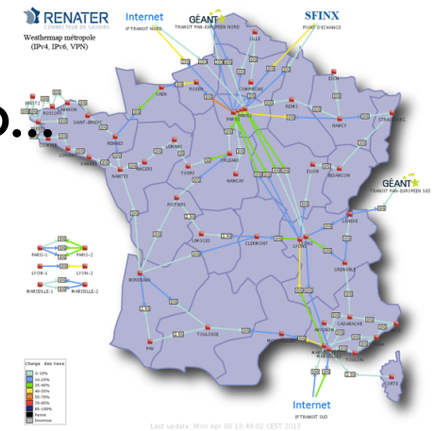
Radio Base Stations



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...)

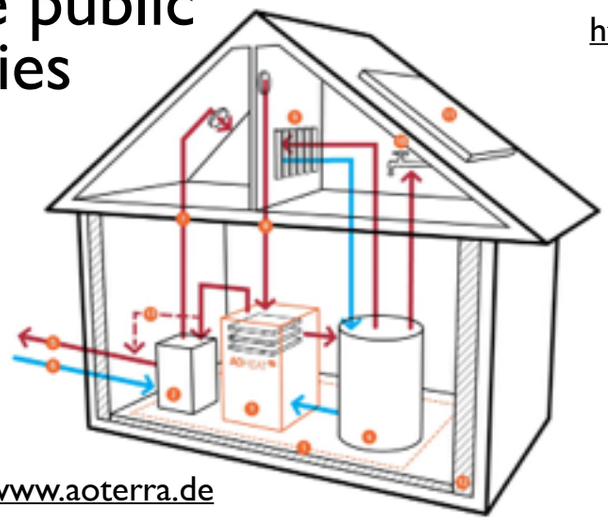
  Transfer the green micro/nano DCs concept to the network PoP

  Take the advantage of the geographical distribution

- 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

http://parasol.cs.rutgers.edu



https://www.aoterra.de

Beyond the Clouds

- **Clouds give the illusion that the infrastructure disappeared but it did not !**

- Thank you / Questions ?
  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/

adrien.lebre@inria.fr