Centre de Calcul de l'Institut National de Physique Nucléaire et de Physique des Particules

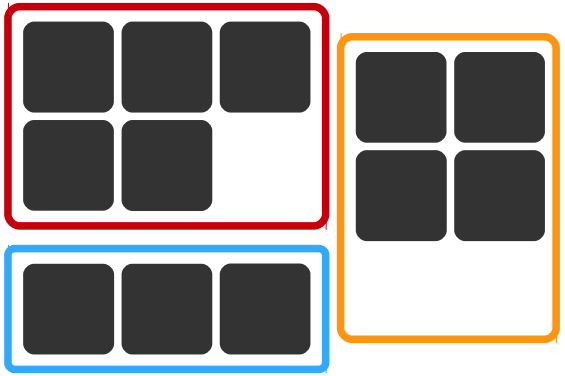
IN2P3-CC cloud computing (IAAS) status

Réseau SARI - Feb '17

Cloud : the big picture

Remote datacenter resource pools
- storage
- CPU
- memory
- networks

Cloudified IT services



IAAS : I want servers, networks and blo devices

PAAS : I want MySQL clusters, email servers

SAAS : I want an online EDMS, ERP service

Cloud user
« The Devops »

IT services
end user

Use cases : R&D, HA and computing clouds

R&D : academic labs and experiments desiring to develop and evaluate new solutions.

→ average availability, mainly private networking, average performance

Public cloud : offload IT services on a public IAAS cloud

Main caracteristics :

- GPFS backend : horizontal capacity and IOPS scaling
- LACP rr 2x10Gbps NICs
- Live migration

→ highest availability, public networking, requested performance

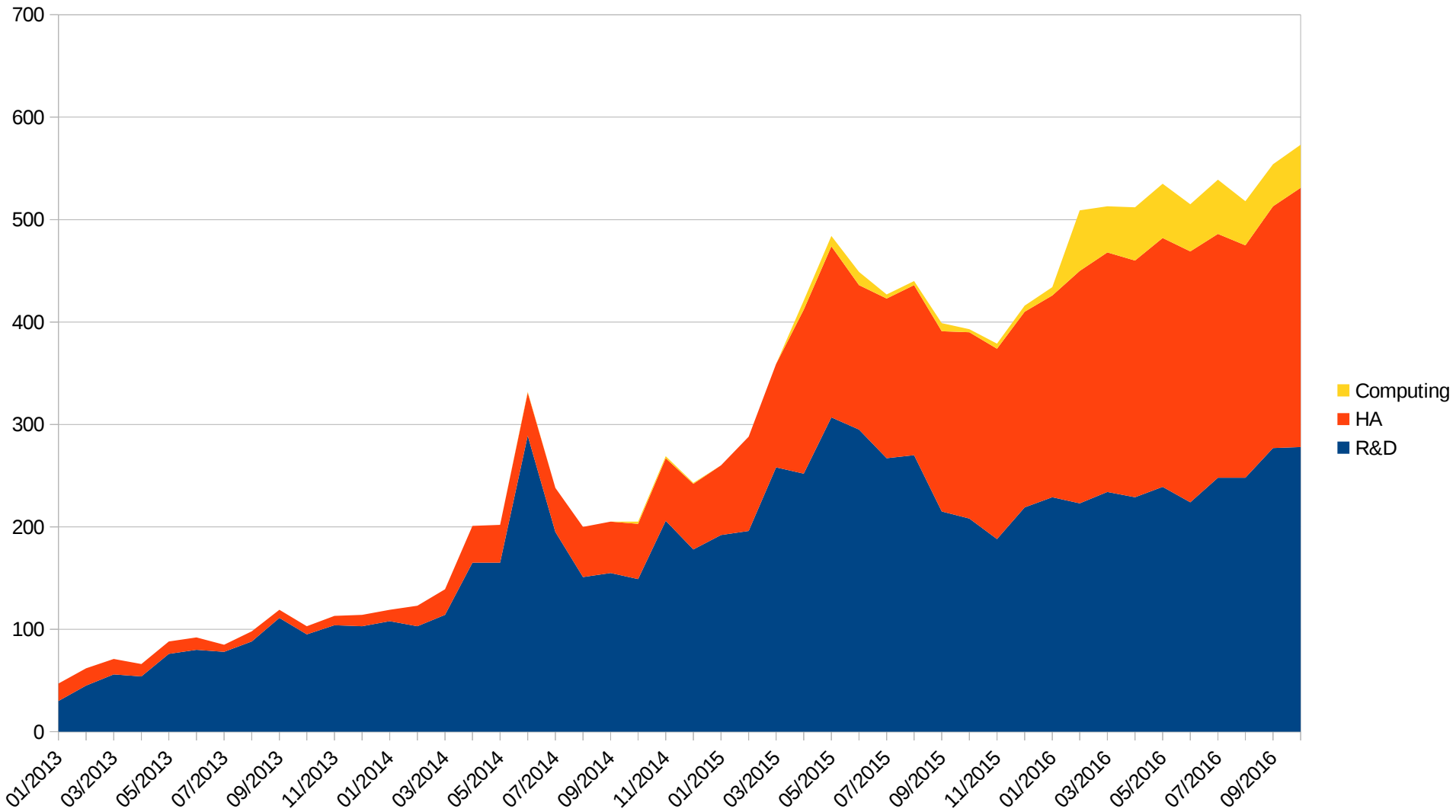
Computing : designed to handle HTC workloads

Main caracteristics :

- ease software deployment and allows specific environment
- specific job management & computing model implementation

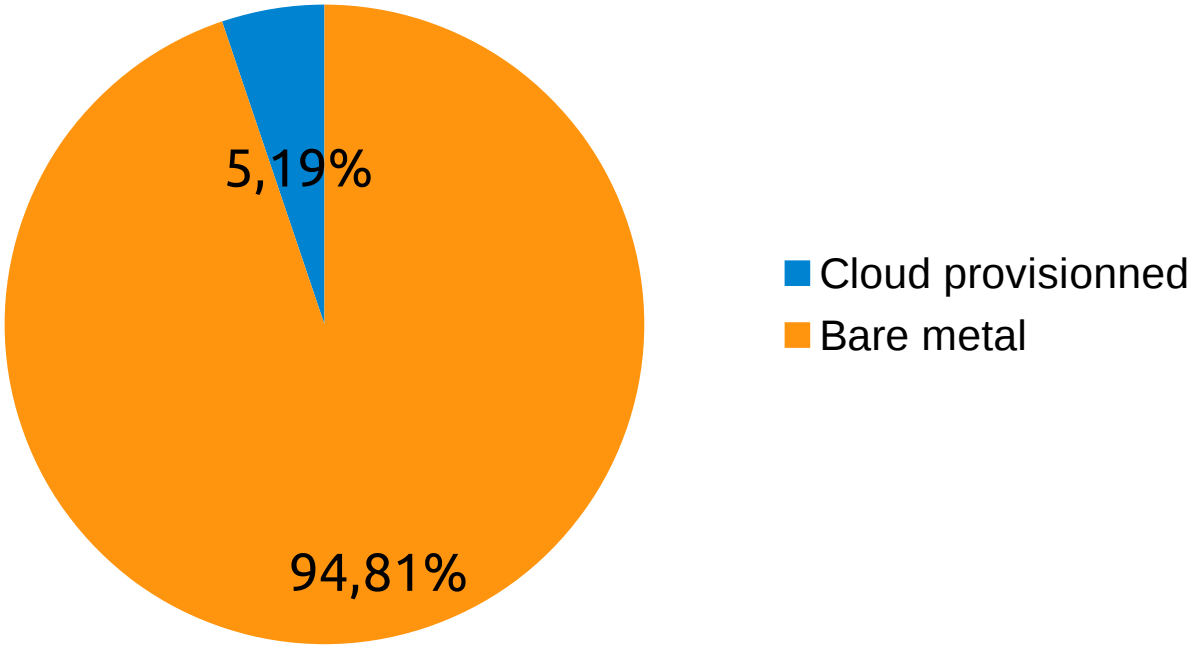
→ low availability, private networking with outgoing internet connectivity, huge CPU cycle accesses (high latency & throughput)

Overall usage



Projects : 50
Users : 150
Hosts : 80
Aggregates : 16

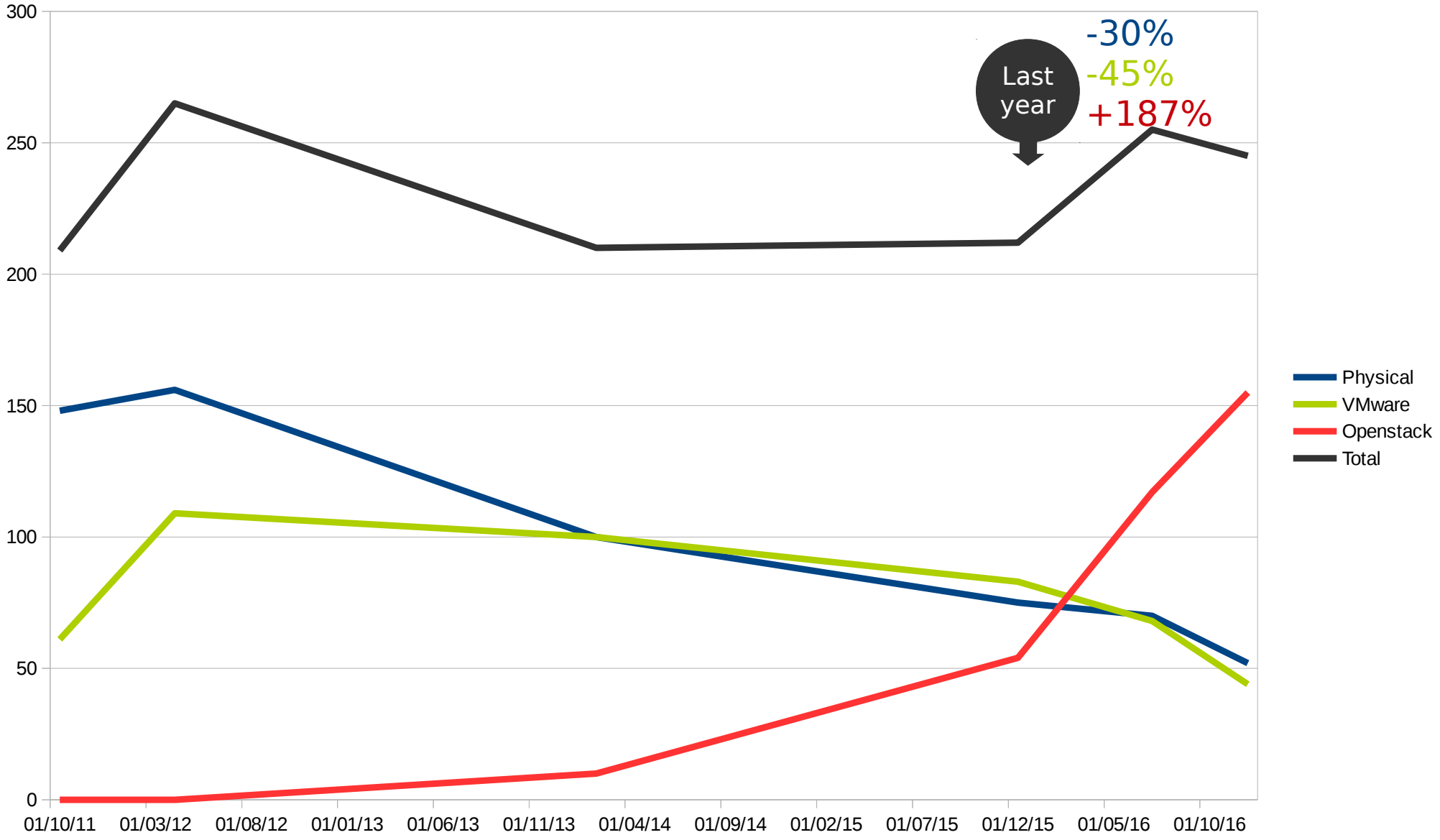
Computing usage



Computing resources :

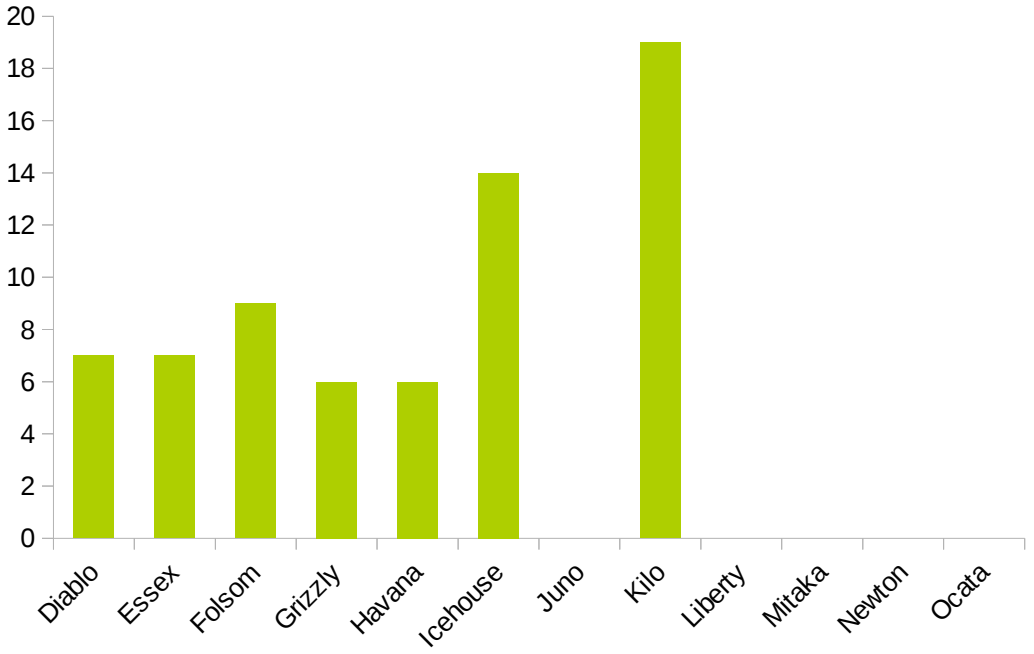
- ~1k HT cores in regard with 25k on HTC farm

Core services deployment (HA private cloud)



Some implementation details

Openstack deployments



Operational components :

Keystone
Glance
Nova
Neutron
Horizon
Cinder
Ceilometer
Swift

In deployment/evaluation :

Heat
Magnum
Rally
Manila

Hardware resources (as of feb ‘17)

Deployment :

- CentOS 7
- RDO packages
- Puppet configuration

Compute clusters (may '16) :

	Cores	RAM	Storage
HA	544	3.6 TB	36 TB
R&D	288	1.2 TB	24 TB
Computing	1216	5.2 TB	40 TB

→ 2048 cores in total, 10TB RAM, 100 TB storage

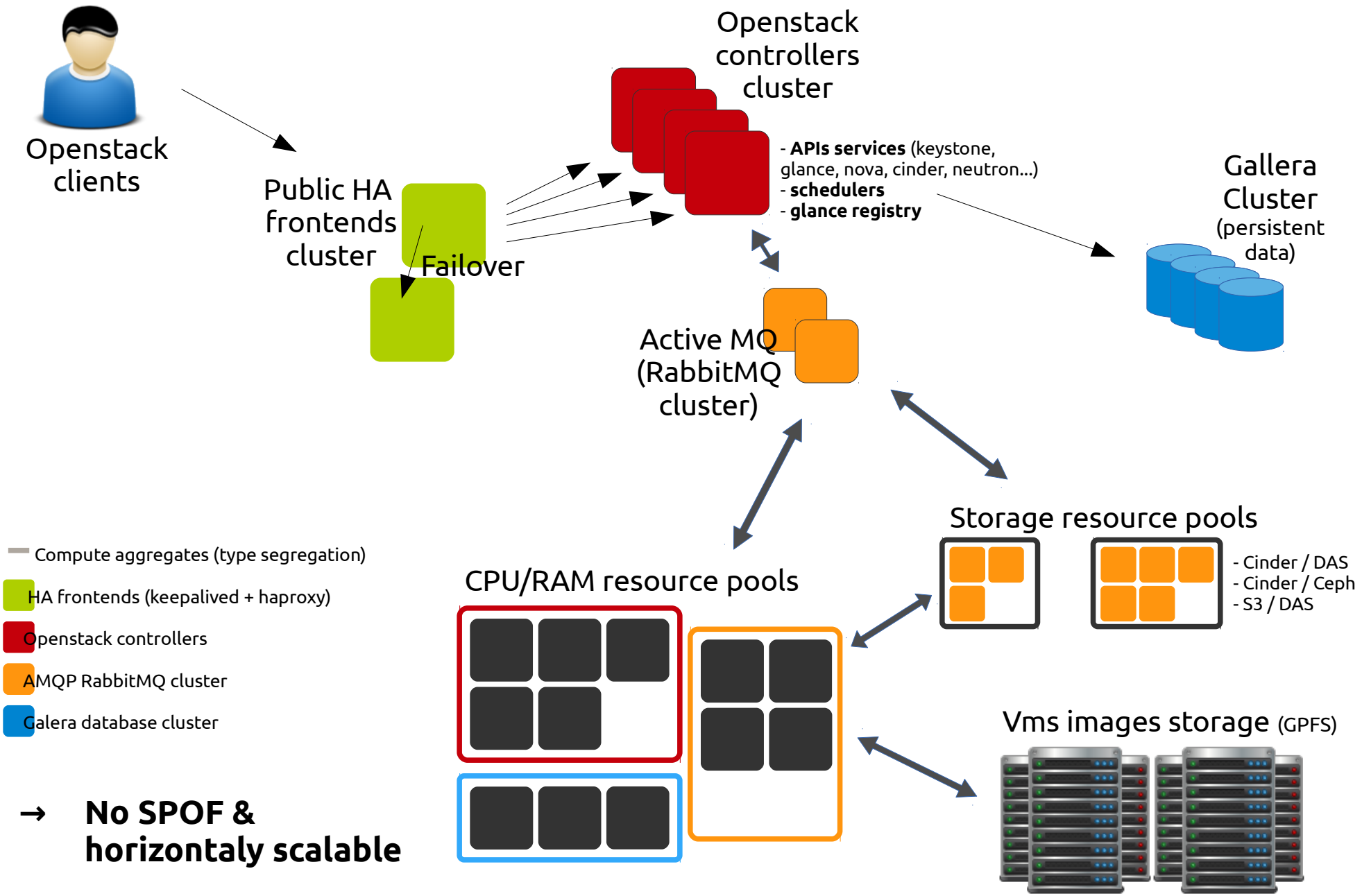
S3 Storage (Swift):

- 24TB DAS

Ceph Storage (Cinder):

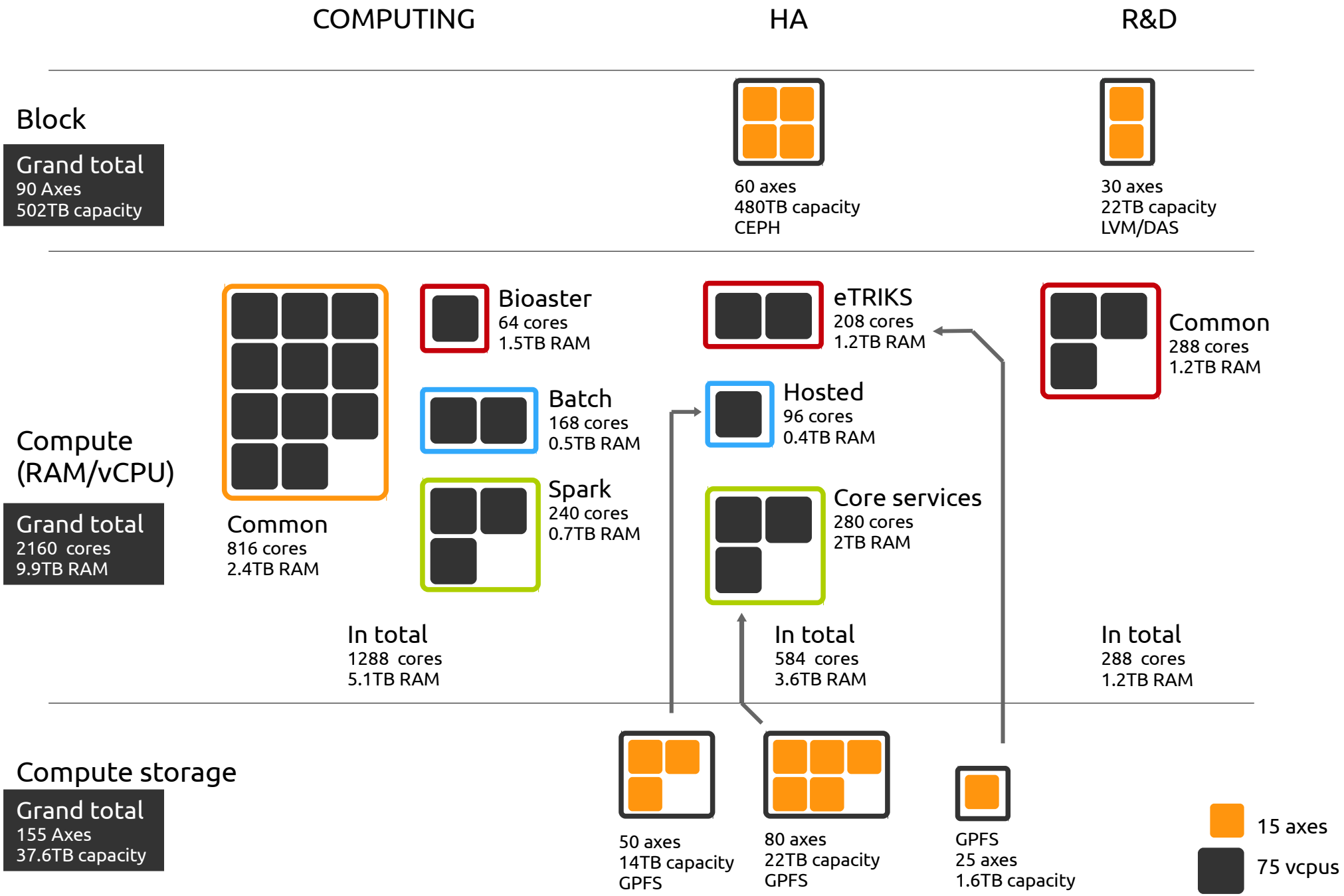
- 480TB DAS

Service architecture



→ **No SPOF & horizontaly scalable**

Resource pools overview (as of march '17)



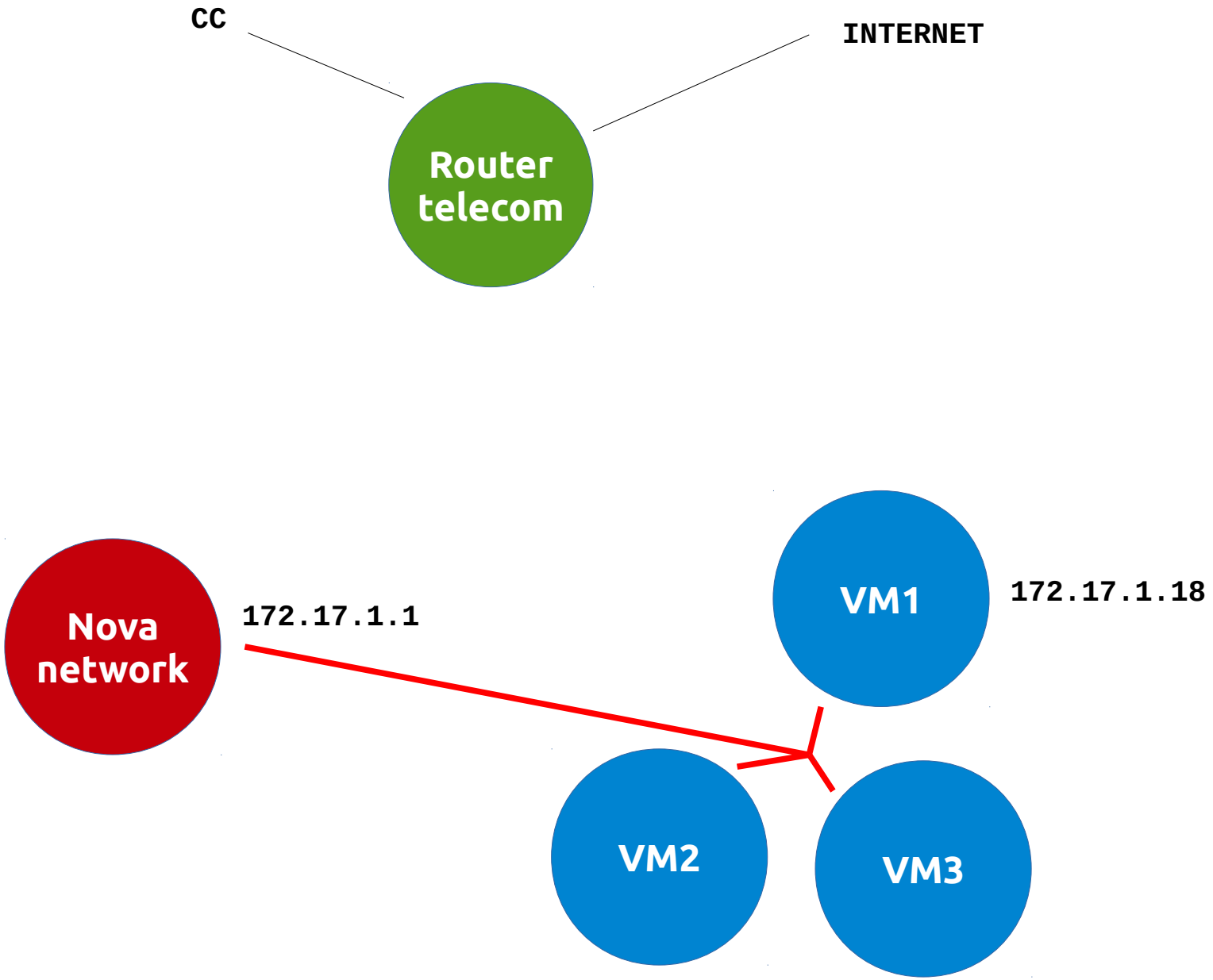
Neutron migration

From legacy networking to Neutron, what's the challenge ?

- 50 different VLANs served (1 ↔ 1 projects)
- 6 different modes
  - Private isolated
  - Private internally routed
  - Private SNATED to the outside world
  - Floating IPs (SNAT/DNAT)
  - Public
  - Public distributed into existing networks
- Virtualized (non cloudish) production services implemented
- No supported migration process : you just rely on what's been done by others
- Linux bridges to OVS
- Static network provisionning to SDN
- HA/DVR

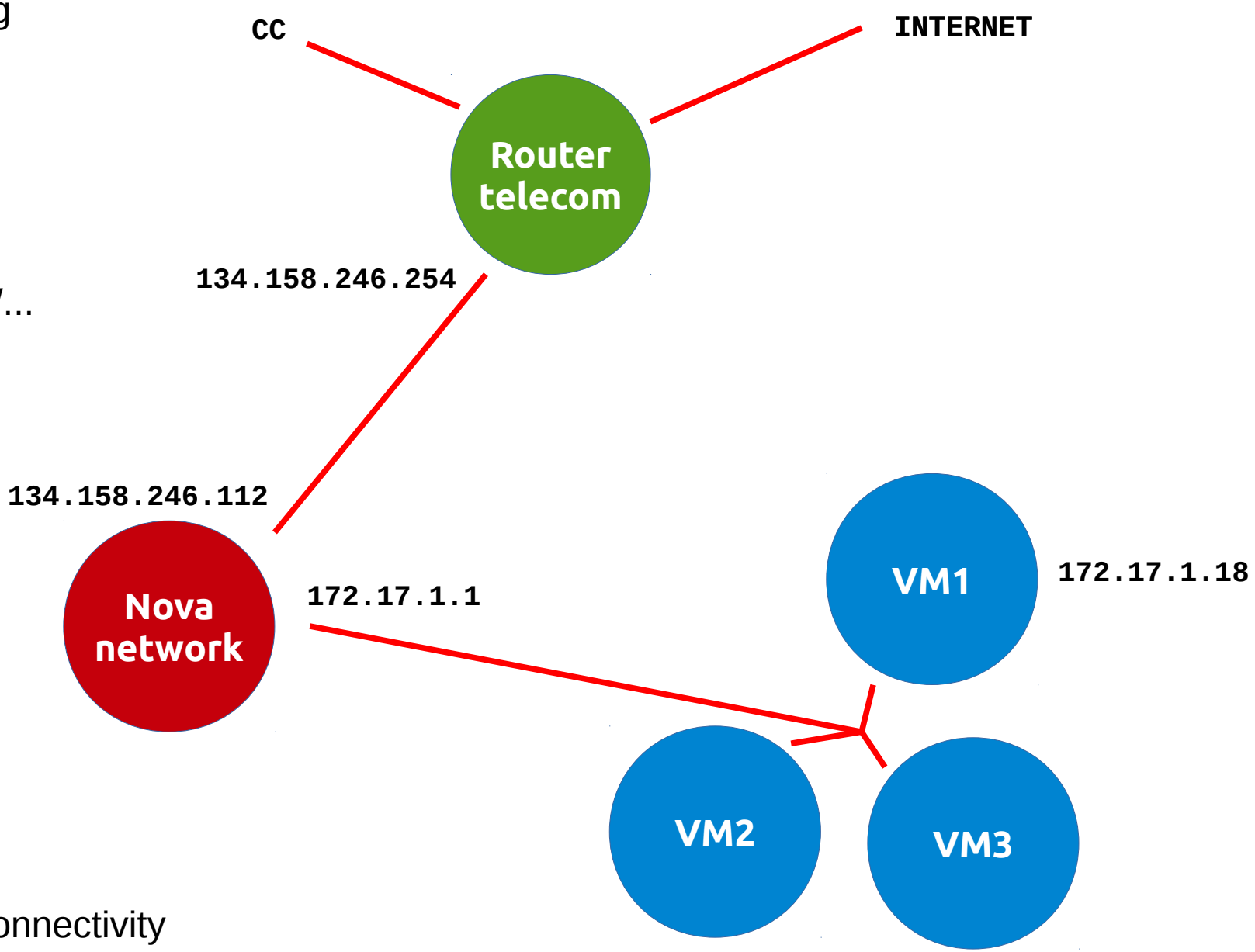
Network type 1
- private subnet
- not routed

Use case :
Likely noone



Network type 1 bis
- private addressing
- not routed
- floating ip

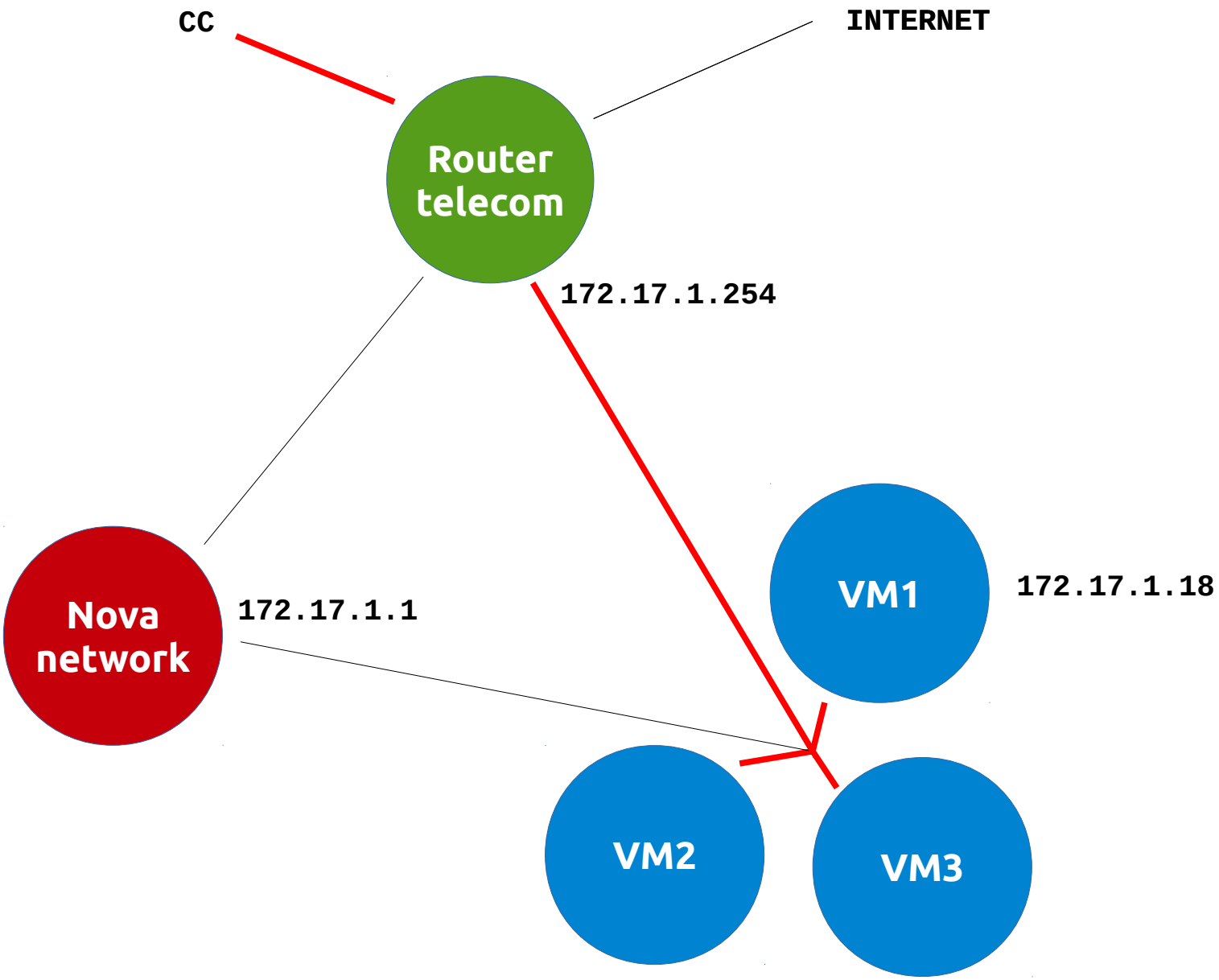
Use case :
- webimatics
- LSST/Euclid
- ecole.info/jenkins/...



Allow internet/cc connectivity

Network type 2
- private addressing
- routed

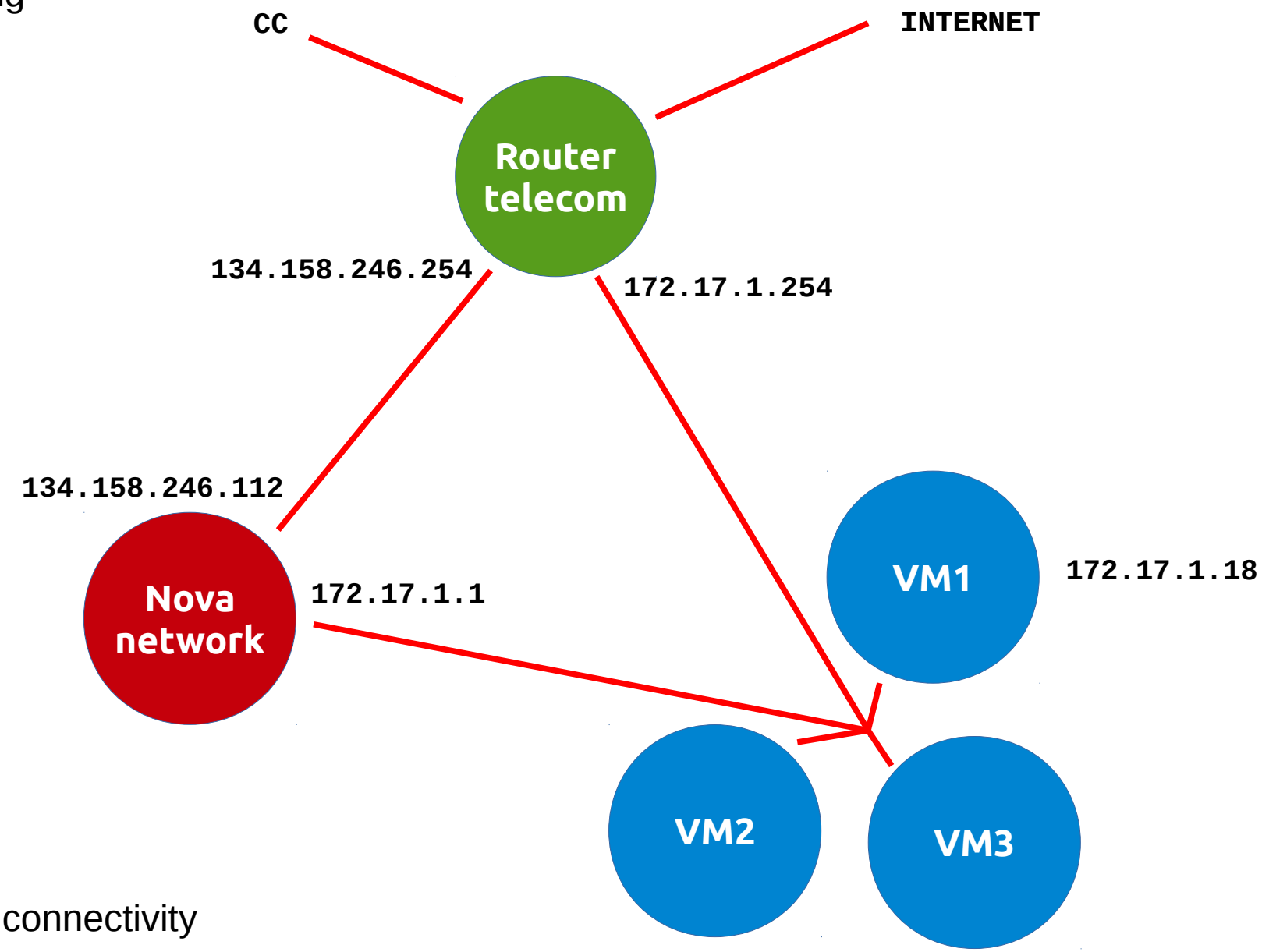
Use case :
- ccin2p3
- htc-atlas



Allow CC connectivity
No NAT for performance
Spare public IPs

Network type 2 bis
- private addressing
- routed
- floating IPs

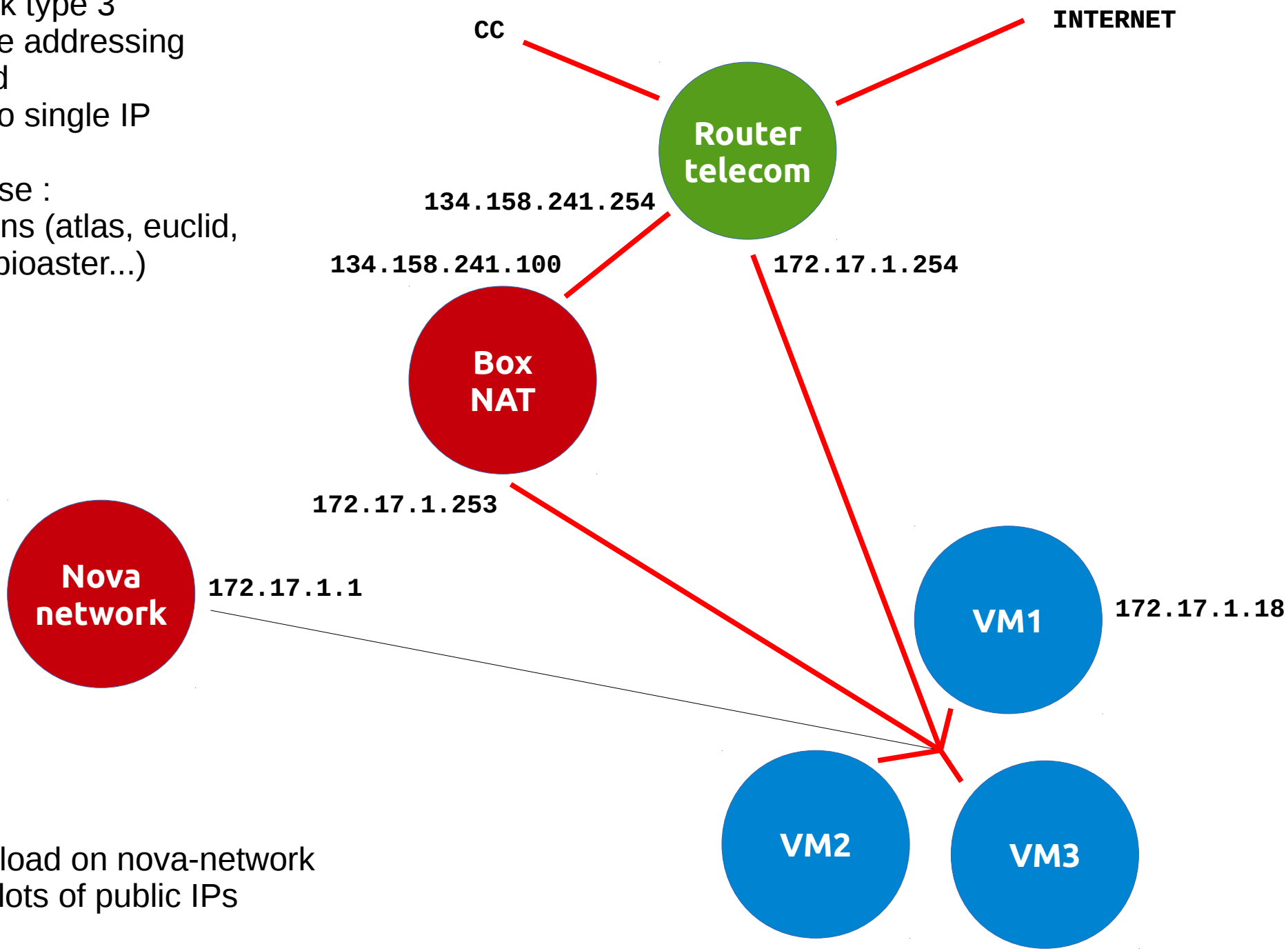
Use case :
- ccin2p3
- htc-atlas



Allow CC/internet connectivity
Spare public IPs

Network type 3
- private addressing
- routed
- NAT to single IP

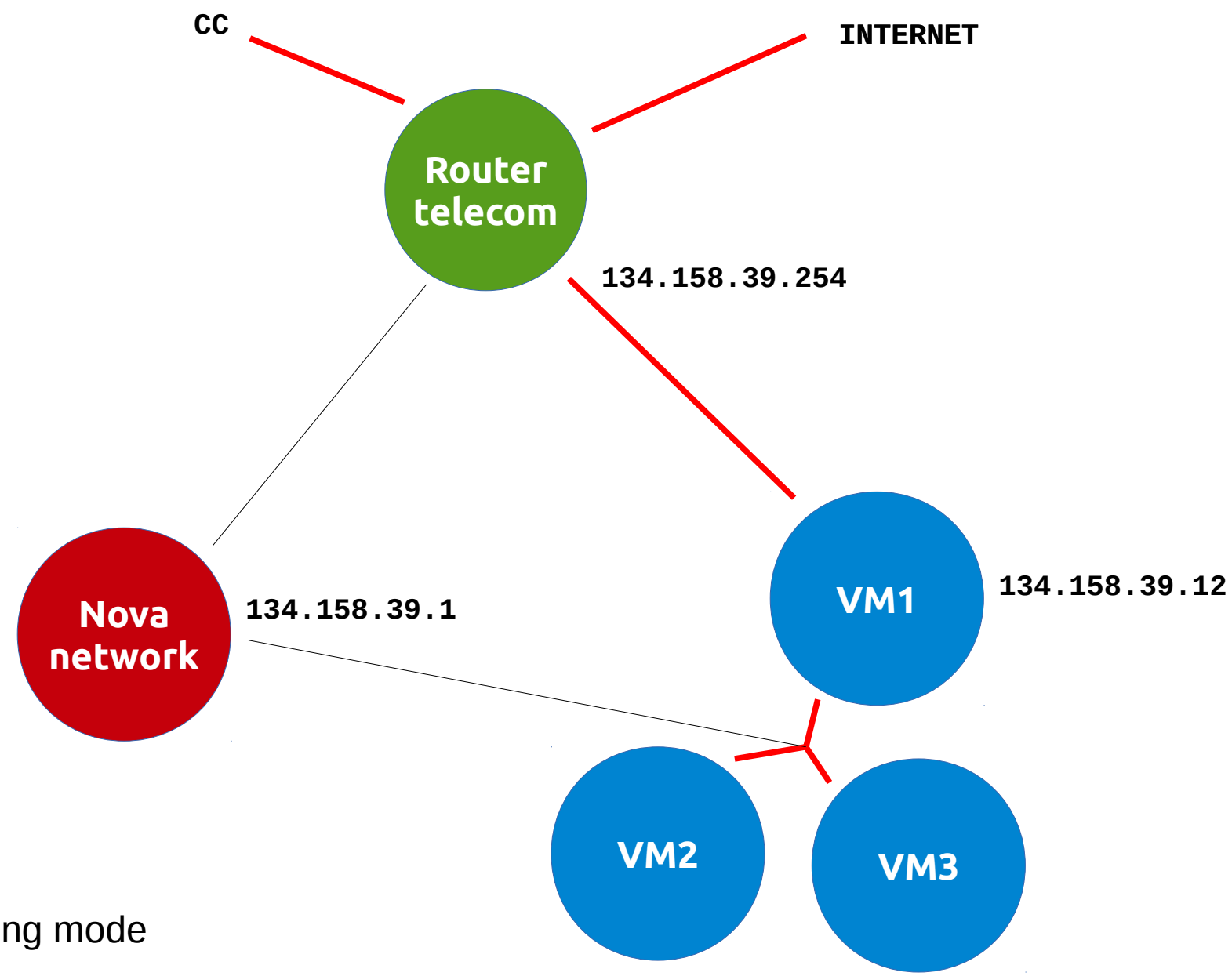
Use case :
- htc Wns (atlas, euclid, LSST, bioaster...)



Lower load on nova-network
Spare lots of public IPs

Network type 4
- public addressing
- routed

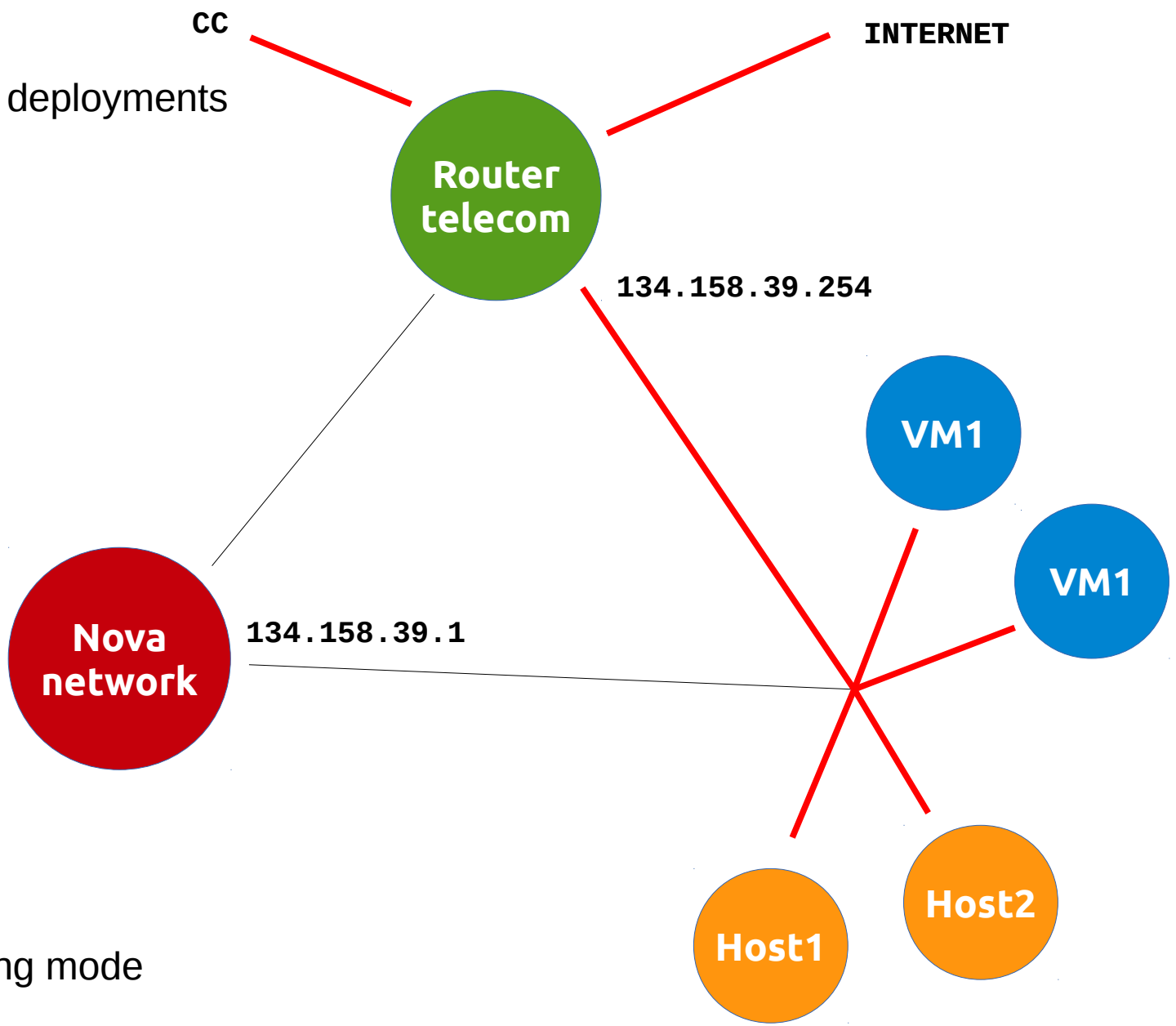
Use case :
- infra-services
- bioaster
- etriks



Legacy networking mode
No NAT (perf)
Public servicing

Network type 5
- public addressing
- routed
- shared with existing deployments

Use case :
- heberge



Legacy networking mode
No NAT (perf)
Public servicing

Monitoring :

Nagios

ELK

Grafana

Nagios®

General
- Home
- Documentation

Current Status
- Tactical Overview
- Map
- Hosts
- Services
- Host Groups
  - Summary
  - Grid
- Service Groups
  - Summary
  - Grid
- Problems
  - Services (Unhandled)
  - Hosts (Unhandled)
  - Network Outages

Quick Search:

Reports
- Availability
- Trends
- Alerts
  - History
  - Summary
  - Histogram
- Notifications
- Event Log

System
- Comments
- Downtime
- Process Info
- Performance Info
- Scheduling Queue
- Configuration

Current Network Status
Last Updated: Tue Apr 4 13:31:15 CEST 2017
Updated every 90 seconds
Nagios® Core™ 3.4.1 - www.nagios.org
Logged in as *Mattieu Puel*

View Service Status Detail For This Service Group
View Status Overview For This Service Group
View Status Summary For This Service Group
View Service Status Grid For All Service Groups

Host Status Totals

Up	Down	Unreachable	Pending
48	0	0	0

All Problems	All Types
0	48

Service Status Totals

Ok	Warning	Unknown	Critical	Pending
94	0	0	0	0

All Problems	All Types
0	94

Status Grid For Service Group 'openstack-servicegroup'

OPENSTACK (openstack-servicegroup)

Host	Services	Actions
cckeystone	Check cckeystone Certificate	
ccoscinder001	Check Cinder volume	
ccoscinder002	Check Cinder volume	
ccoscinder003	Check Cinder volume	
ccoscinder052	Check Cinder volume	
ccoscinder053	Check Cinder volume	
ccoscinder054	Check Cinder volume	
ccoscinder055	Check Cinder volume	
ccoscinder056	Check Cinder volume	
ccosctl01	Check Controller Check Disk Check Rabbitmq	
ccosctl02	Check Controller Check Rabbitmq	
ccosctl06	Check Disk	
ccosdb	Check Mysql connections	
ccosha01	HA Proxy	
ccosha02	HA Proxy	
ccosmonitor	Check Ceilometer Check Cinder Check Core Aggregate Check Glance Check Glance Transfer Check Heberge Aggregate Check Keystone Check Metadata Check Nova Check Nova SSH keypairs Check Spawn VM Check Swift	
ccosndli0001	Check Disk Check compute	
ccosndli0002	Check Disk Check compute	
ccosndli0003	Check Disk Check compute	
ccosndli0004	Check Disk Check compute	
ccosndli0005	Check Disk Check compute	
ccosndli0006	Check Disk Check compute	
ccosndli0007	Check Disk Check compute	
ccosndli0008	Check Disk Check compute	
ccosndli0009	Check Disk Check compute	
ccosndli0010	Check Disk Check compute	
ccosndli0011	Check Disk Check compute	
ccosndli0012	Check Disk Check compute	
ccosndli0013	Check Disk Check compute	
ccosndli0014	Check Disk Check compute	
ccosndli0015	Check Disk Check compute	
ccosndli0016	Check Disk Check compute	
ccosndli0022	Check Disk Check compute	

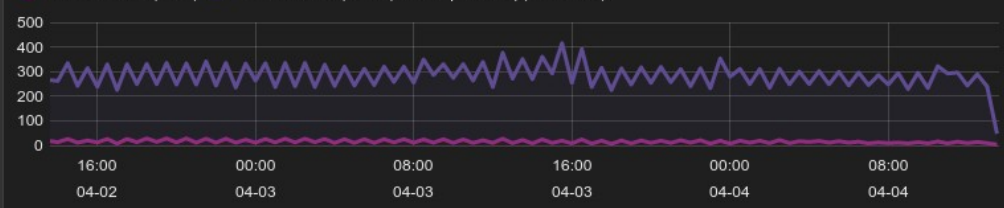
Kibana 3 - Openstack - Google Chrome

Kibana 3 - Openstac ×

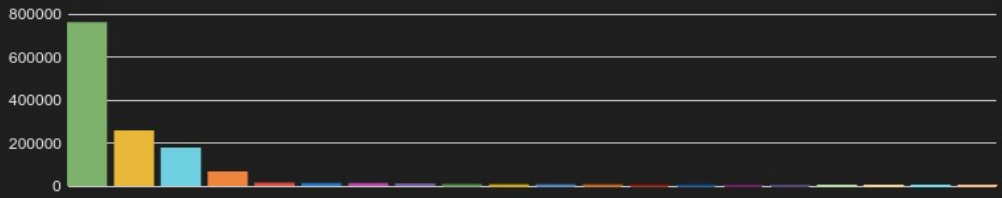
cceventmole.in2p3.fr/#/dashboard/file/Openstack.json

API 200+

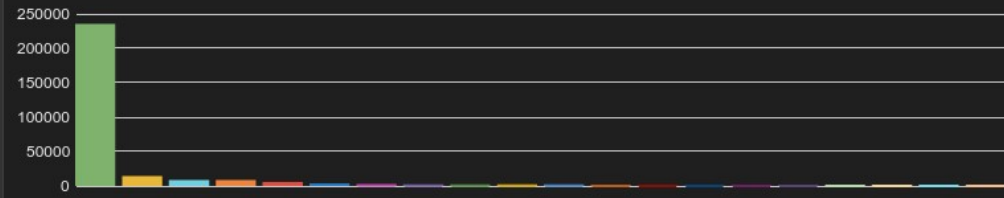
API 400+ AND 500+



USERAGENT

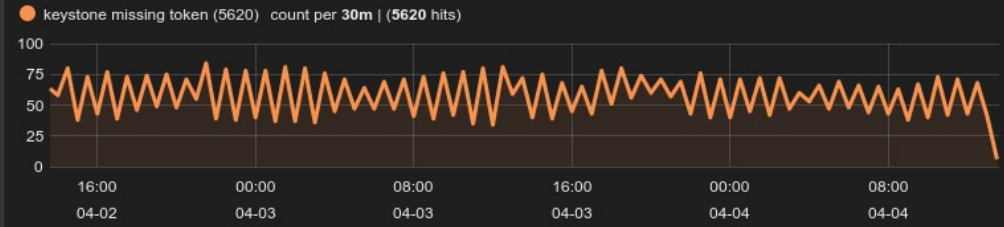


REQUEST BY TENANT

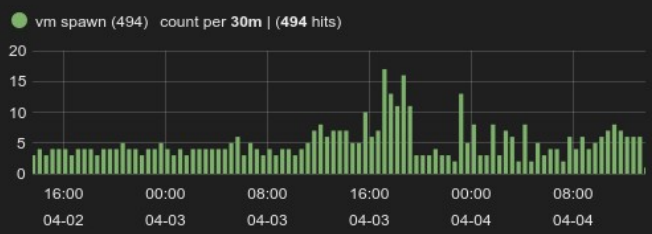


GET TOKEN

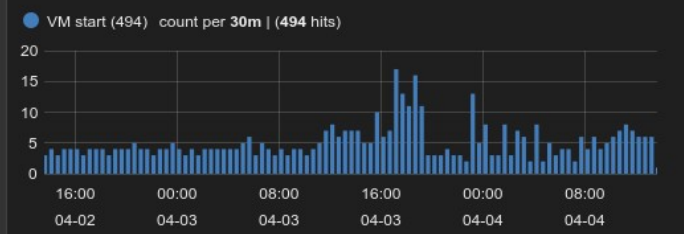
MISSING TOKEN



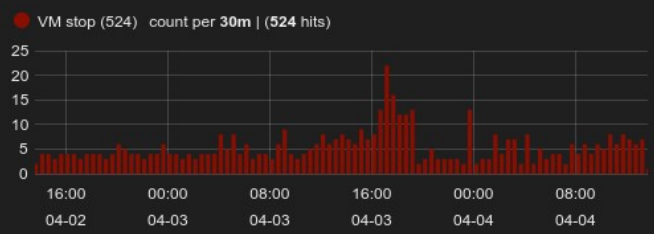
VM CREATE



VM START

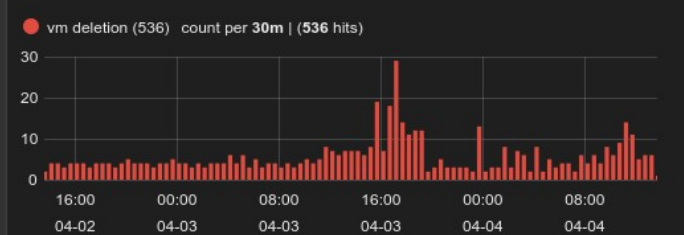


VM STOP

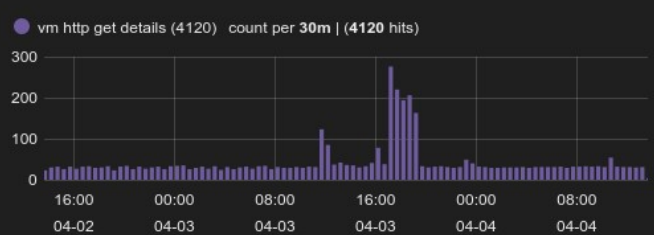


VM REBOOT

VM DELETE



VM GET DETAILS



Grafana - Average by usage/role - Google Chrome

Grafana - Average b

https://ccsampler.in2p3.fr/dashboard/db/average-by-usage-role?var-site_usage=openstack&var-site_role=accounting&var-precision=$__auto_interval&var-environment=production_git

Average by usage/role

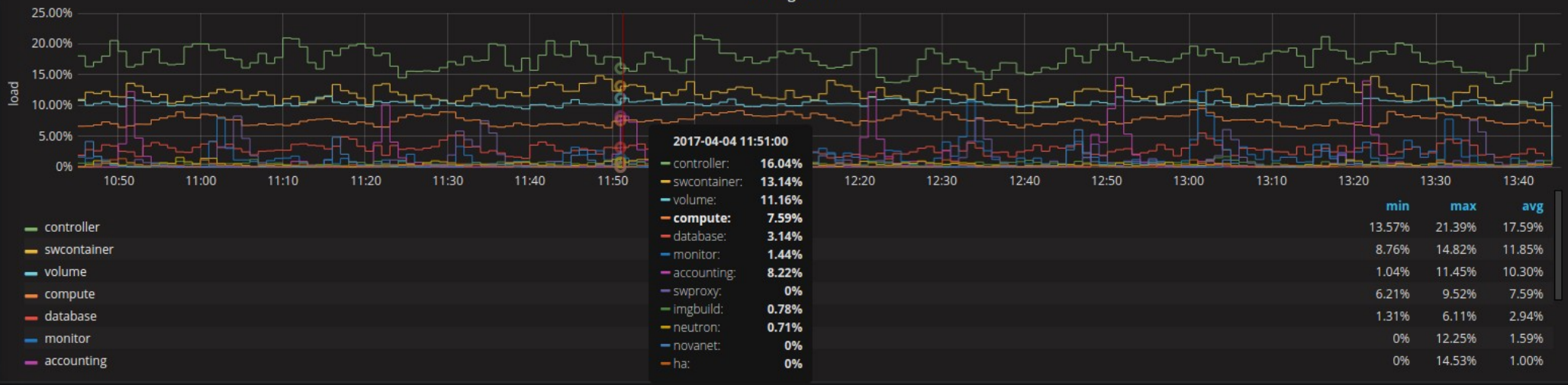
Zoom Out | Last 3 hours

site_usage: openstack | site_role: accounting | precision: auto | environment: production_git

host count

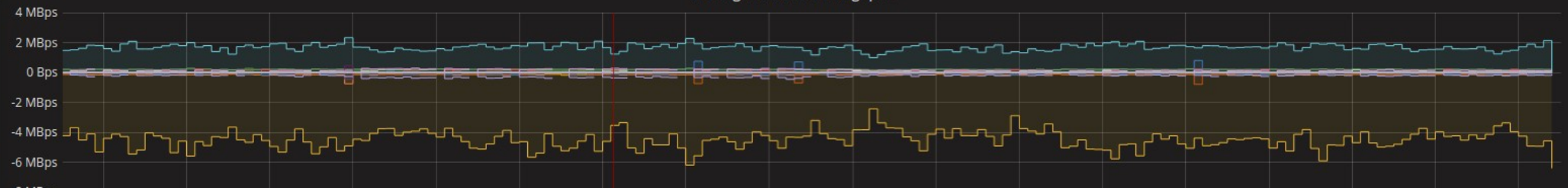
facter.site_role	Unique Count
compute	86
swcontainer	6
volume	11
neutron	2
ha	4
controller	3
imgbuild	1
accounting	2
swproxy	1
novanet	1

average load



	min	max	avg
controller	13.57%	21.39%	17.59%
swcontainer	8.76%	14.82%	11.85%
volume	1.04%	11.45%	10.30%
compute	6.21%	9.52%	7.59%
database	1.31%	6.11%	2.94%
monitor	0%	12.25%	1.59%
accounting	0%	14.53%	1.00%

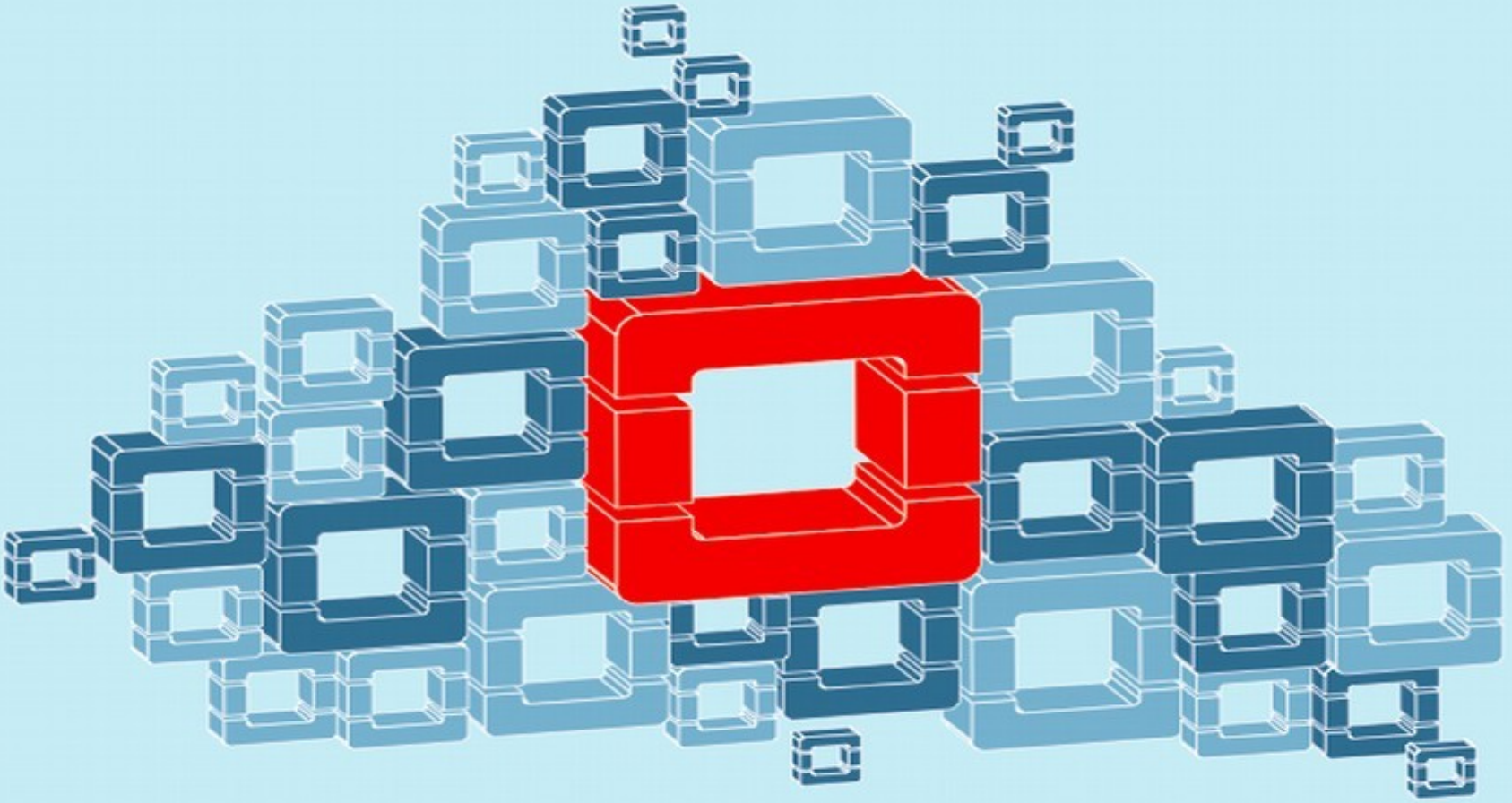
average network throughput



What's next ?

What’s next in the mid-term

- Public cloud platform enhancements (storage & CPU) : targeted usage within 3 years : 300 instances, 1.4TB RAM, 800 vCPUs, 50 TB storage
- FSS with Synergy for computing
- Finalize resources orchestration (Heat)
- Container orchestration (Magnum/Kubernetes)
- Cinder refactoring with Ceph backend
- AAI integration
- Interest in shared FS aaS (Manila)



(thank you)

Questions ?

Divers

- Updates : better with puppet
- Cout ETP → taille du deploiement