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

Cloud du CCIN2P3 pour l' ATLAS VO

Vamvakopoulos Emmanouil

« ***Rencontres LCG-France*** »

IRFU Saclay

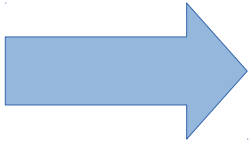
1-2 December 2014

VO needs and objective

Atlas would like to use oportunistic Cloud Resources

- Case for Simulation (Single and MultiCore jobs)

Deploy Virtual Machines as «Worker Nodes» on a Cloud IAAS (infrastructure as a Service) and Interfaced those VMs with the Grid System

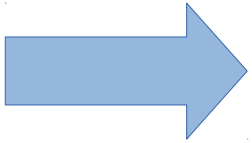


VO needs and objective

Atlas would like to use oportunistic Cloud Resources

- Case for Simulation (Single and MultiCore jobs)

Deploy Virtual Machines as «Worker Nodes» on a Cloud IAAS (infrastructure as a Service) and Interfaced those VMs with the Grid System

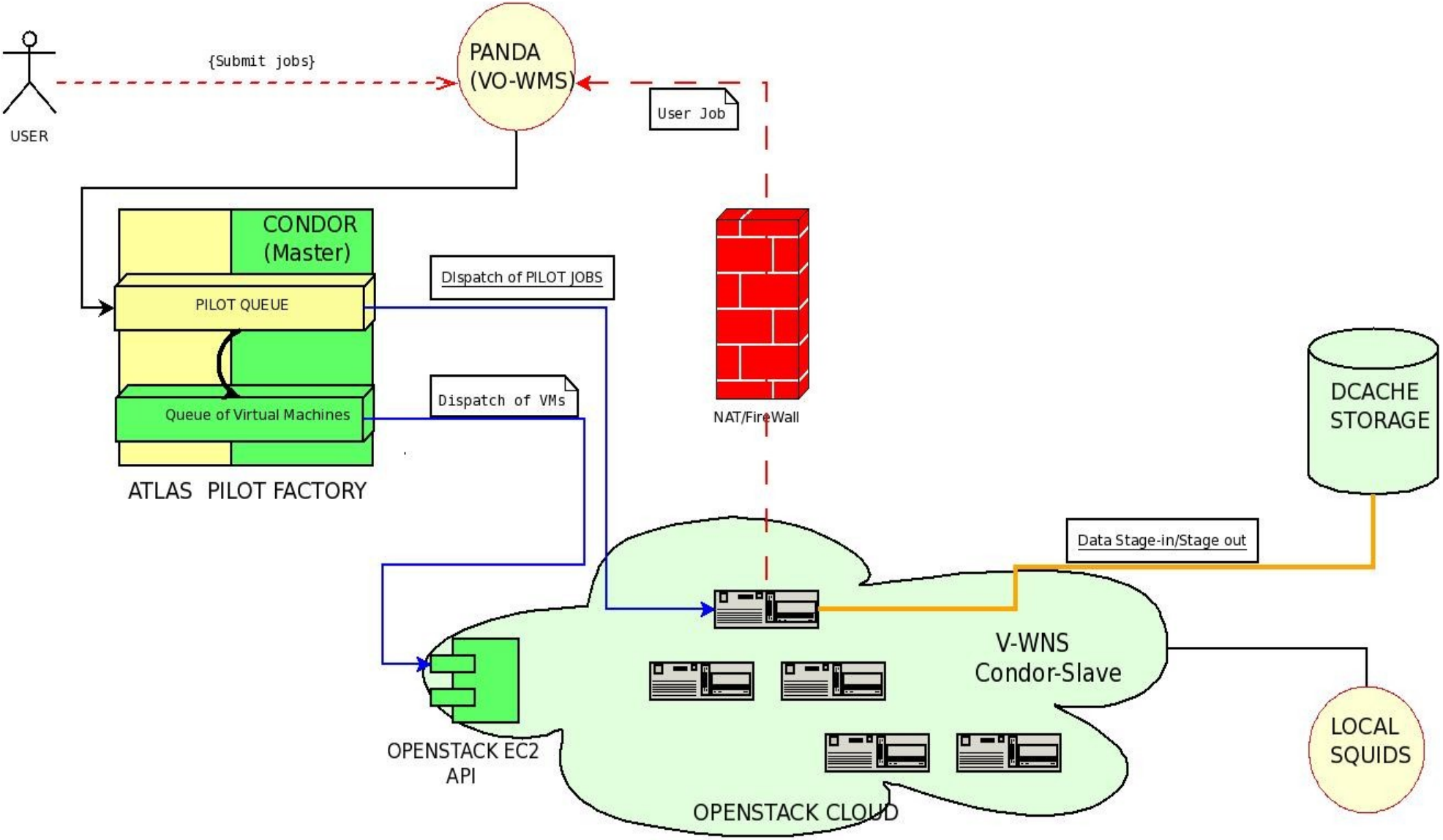


AtlasCloud news and Updates

AtlasCloudOperations Group

- ***atlas-adc-cloudcomputing@cern.ch***
- **Deploy the available, established solutions in order to Bridge the Grid Computing to Cloud Computing**
- **Ensure Cloud Resource smooth operation**

Review : Context Diagram



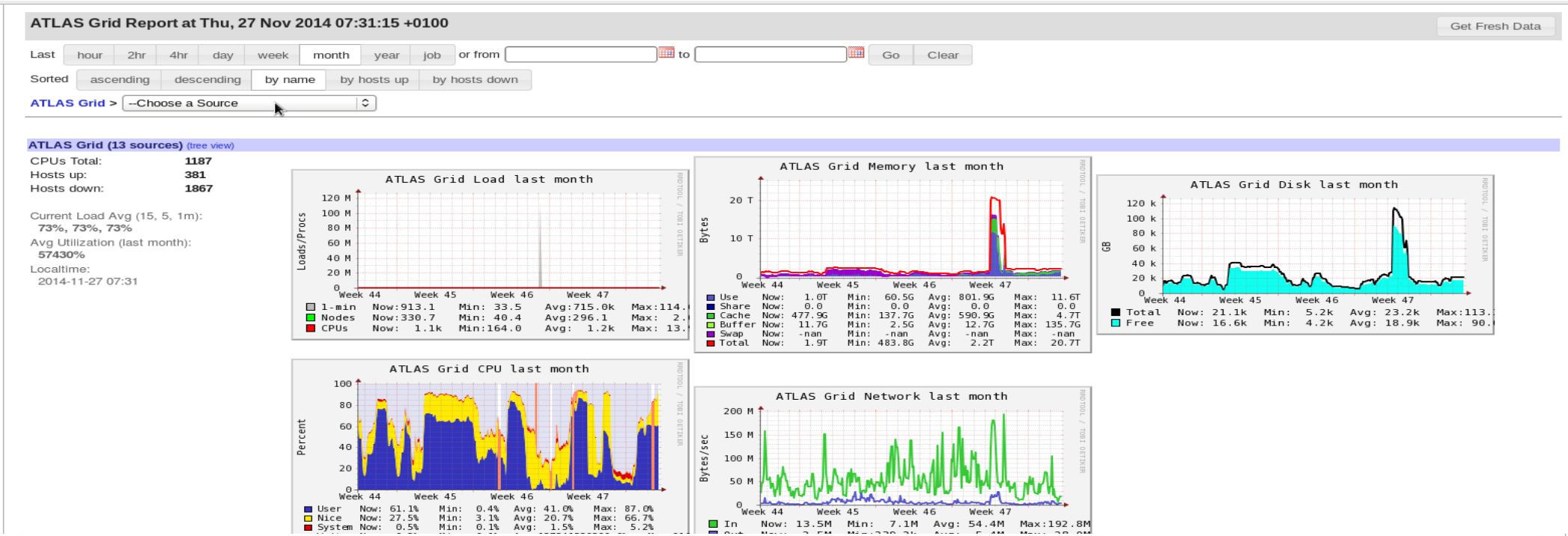
BNL Elastic Cluster

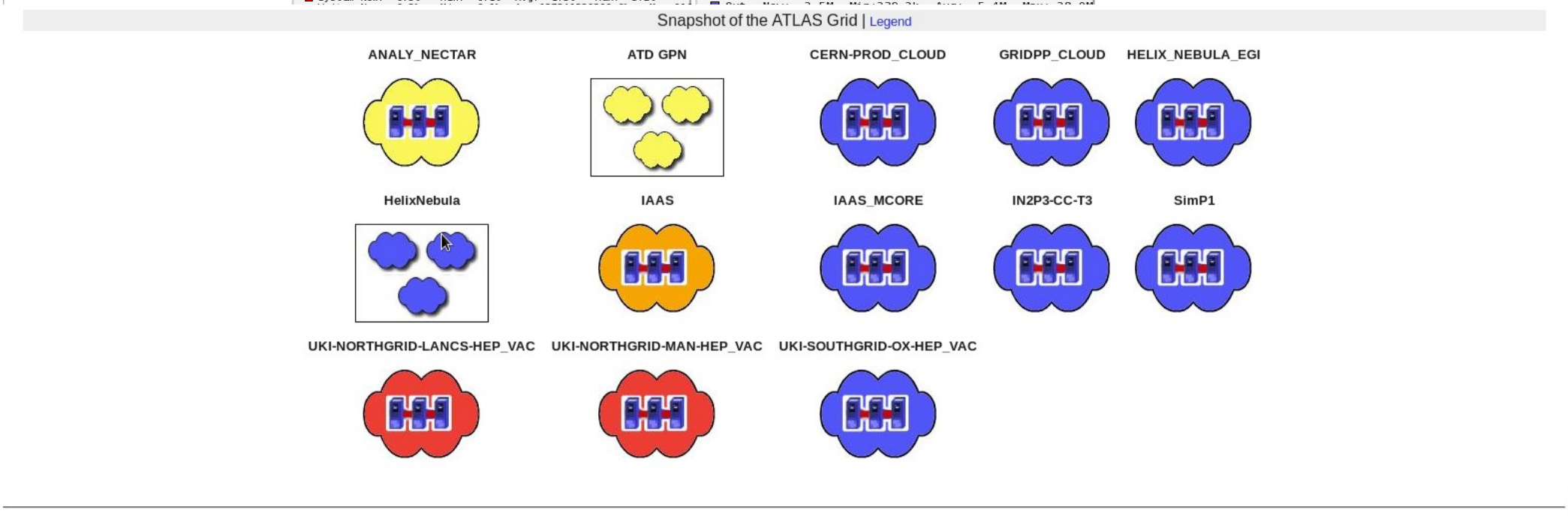
- **https://www.racf.bnl.gov/experiments/usatlas/griddev/AutoPyFactory**
- **BNL/Amazon AWS Pilot**

AtlasCloud Operation Recommendations

- **Cloud operations**
  - **Use of CernVM (3.X)**
  - **Standard Cloud-init contexualization templates**
  - **Publish to Federated Ganglia Monitor Tool**
  - **(http://agm.cern.ch)**
- **Central AutoListener: «CloudScheduler»**
  - **Not 100% madantory**

Atlas Ganglia Monitoring





http://agm.cern.ch/

WLCG Cloud Usage Resource dashboard

Mixted Sources (VO-WMS/EGI Apel/Ganglia monitoing)

WLCG Clould, Usage

Cloud Resources

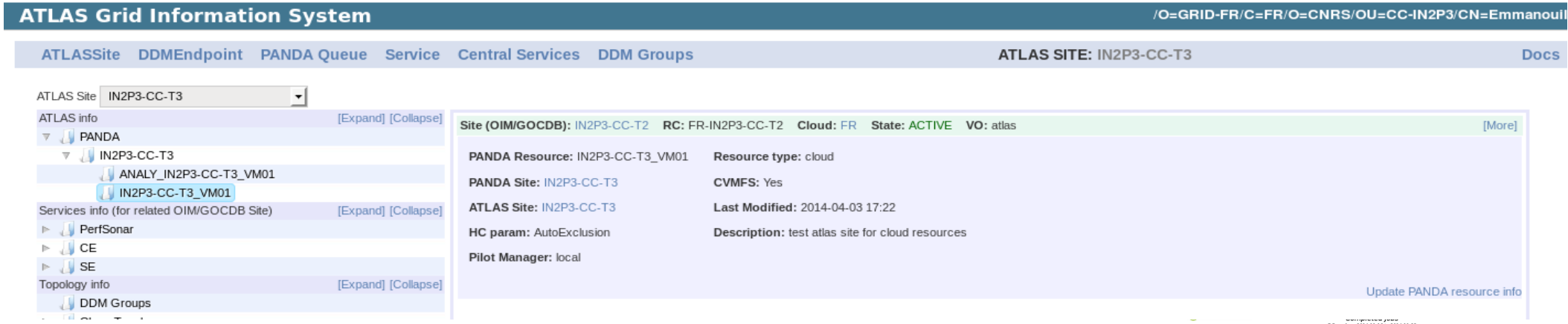
VO: ATLAS

Search:

Country	Resource	Discovered Instances	Monitored Instances	Cores	Jobs
UK	UKI-NORTHGRID-MAN-HEP_VAC	247	246	246	246
Switzerland	CERN-P1	66	N/A	N/A	527
Canada	IAAS	45	4	25	274
Canada	IAAS_MCORE	27	0	0	26
Australia	Australia-NECTAR	13	16	64	43
UK	UKI-NORTHGRID-LANCS-HEP_VAC	8	8	8	7
France	IN2P3-CC-T3_VM02	2	0	0	4
UK	RAL-LCG2_VAC	2	0	0	0
Switzerland	CERN-P1_preprod_MCORE	1	0	0	0
France	IN2P3-CC-T3_VM01	1	1	2	4
Canada	ANALY_IAAS	1	0	0	3
Switzerland	CERN-P1_preprod	1	0	0	4
Australia	ANALY_NECTAR	1	16	64	1
Switzerland	IISAS-FEDCLOUD	0	0	0	0
Switzerland	CESNET-METACLOUD	0	0	0	0
United Kingdom	GRIDPP_MCORE	0	0	0	0
Switzerland	CERN-P1_MCORE_HI	0	0	0	0
UK	UKI-NORTHGRID-LANCS-HEP_CLOUD	0	0	0	0
Australia	ANALY_NECTAR_TEST	0	0	0	0
Switzerland	HELIX_NEBULA_EGI	0	0	0	0
Czech Republic	praguelcg2_RUCIOTEST	0	0	0	0

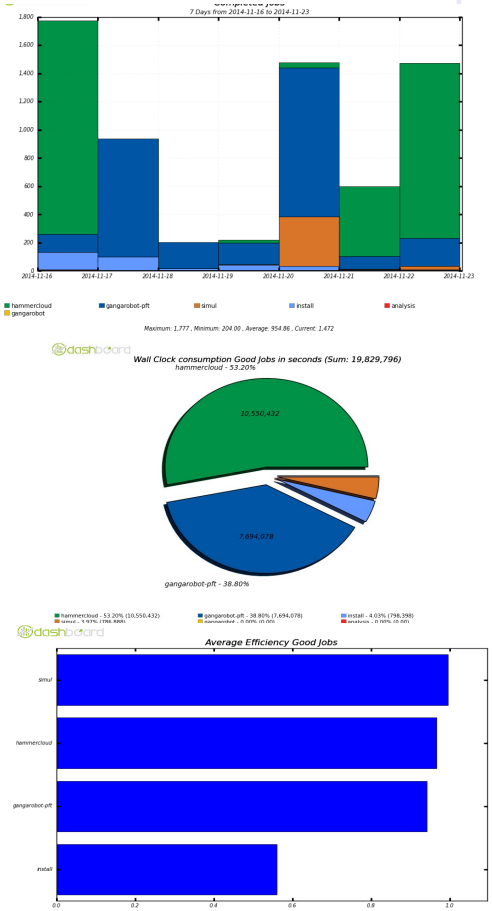
http://cloud-acc-dev.cern.ch

Cloud Resources Intergation at IN2P3-CC



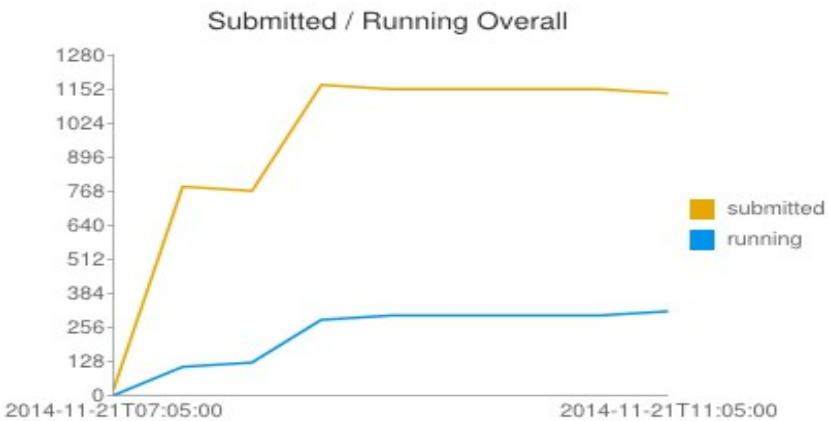
- **IN2P3-CC-T3 --> ATLAS SITE (Agis)**
  - DDM-ENDPOINT-LESS SITE (use of the T1's storage (dCache)
  - **Panda queue IN2P3-CC-T3-VM01 (on-line)**
  - **ANALYSIS_IN2P3-CC-T3-VM01 (off-line)**

Seperate view on Historical View or SSB DashBoard

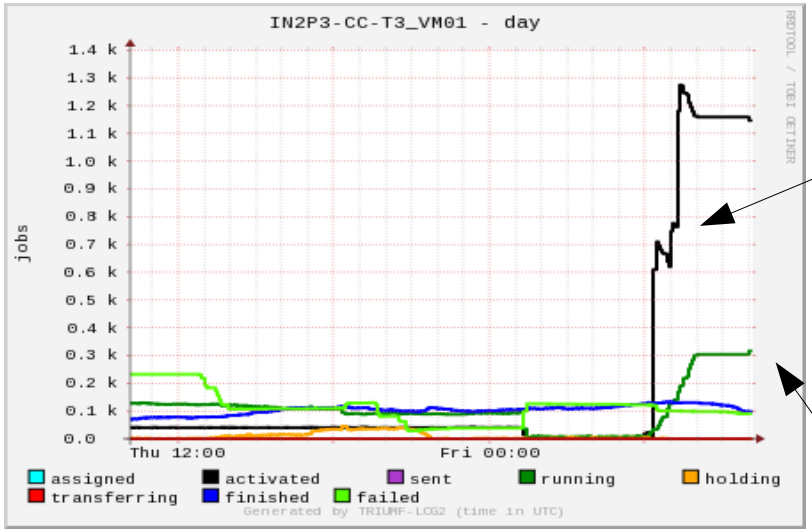


BNL Elastic Cluster Solution: APF + condorEC2

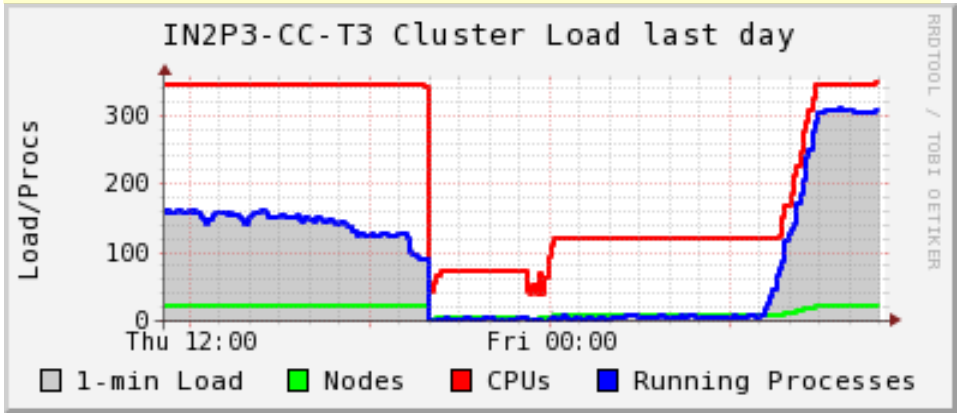
HC test Framework:User



Panda – VO WMS



Condor Cluster



# of Activated Jobs : Ready to dispatch

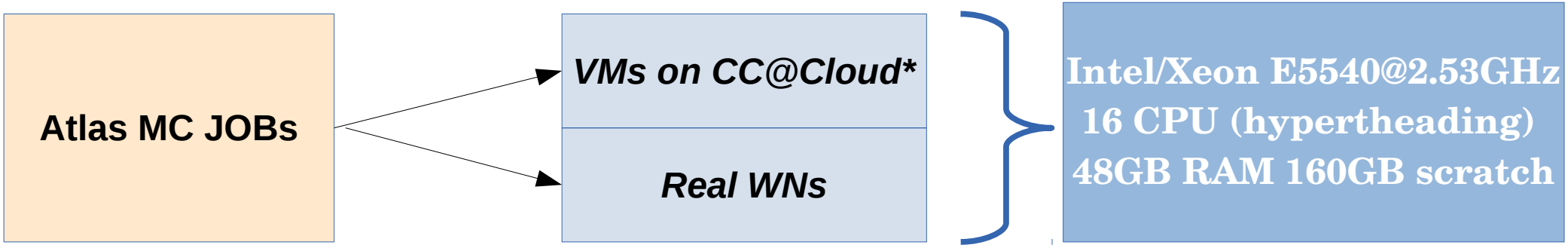
running jobs at Condor Cluster according to the # of Activated jobs

Running Jobs

Submit Pilots | **Submit VMs**

APF+CondorEc2

Benchmark : Atlas MC



- ***CERN-P1 StressTest - mc12 AtlasG4_trf 17.2.2 (512)***
- ***AtlasProduction/17.2.2***
- ***mc12_8TeV.175590.Herwigpp_pMSSM_DStau_MSL_120_M1_000.evgen.EVNT.e1707_tid01212395_00_derHCBM***
- ***Overheads (wallclock) ~ 30 %***
- ****16-VCPU/32GB-RAM/160GB-Scratch***

16 corrurent job
On both partition
For ~7 circles

Next Steps

- Verify BenchMark with latest Hardware
  - Optimization of Memory performance : e.g. Numa topology vs the Size of the Image
- Deploy CernVm 3.x image solution (PaaS)
- Establish a local fairshare policy
- Finilized operation details
- Log and Tracking Activity

Merci

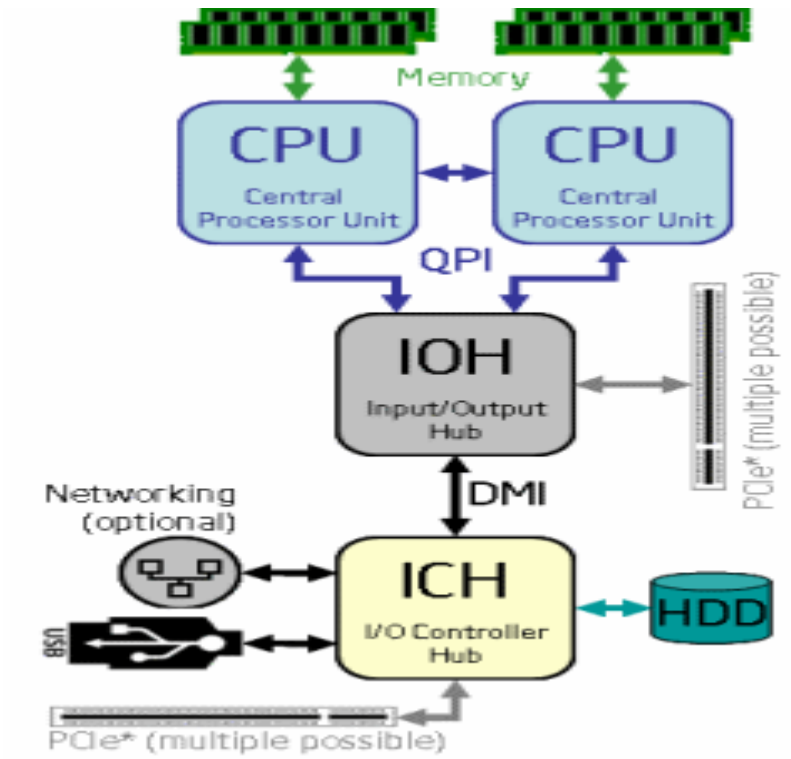
de Vorte Attetion !

Backup Slides

Memory performance and Numa Topology

Intel® Xeon® Processor E5540

(8M Cache, 2.53 GHz, 5.86 GT/s Intel® QPI)



« Non-uniform memory access (NUMA) is a computer memory design used in Multiproccesing, where the memory access time depends on the memory location relative to the processor. »

Demostration of Non-uniform memory access Pattern

e.g. numactl --membind 1 --physcpubind 0

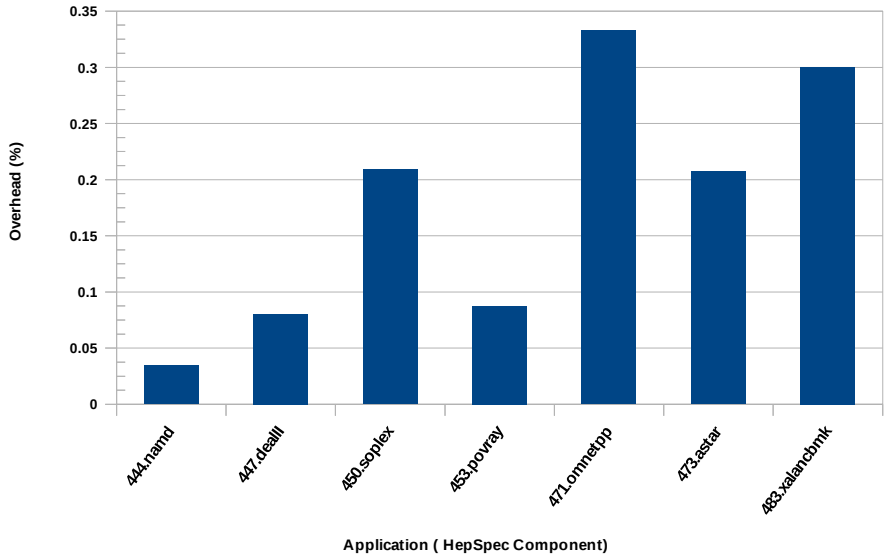
./ramsmp -b 6 -m 32 -p 1 -l 3 (*)

Local --> ~8563.08 MB/s

Remote --> ~ 5855.48 MB/s

http://vmstudy.blogspot.gr/2010/09/kvm-memorycpu-benchmark-with-numa.html
* http://www.alasir.com

Benchmark : HepSpec06



Integer Benchmarks

471.omnetpp C++ Discrete Event Simulation
473.astar C++ Path-finding Algorithms
483.xalancbmk C++ XML Processing

Floating Point Benchmarks

444.namd C++ Biology / Molecular Dynamics
447.dealII C++ Finite Element Analysis
450.soplex C++ Linear Programming, Optimization
453.povray C++ Image Ray-tracing

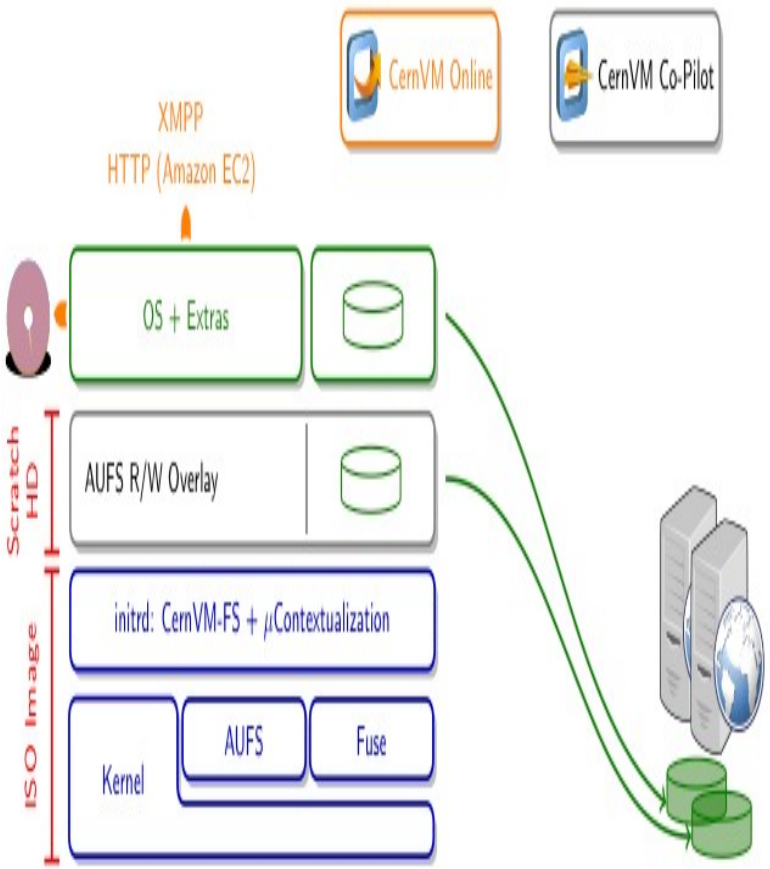
- Total overheads (on HepSpec06 mark) ~17%
- The individual overheads (WallClocks) per HepSpec06 component exhibit large dispersion
- Floatpoint vs integer Application ?

Image Softwate Stack and Size

- BoxGrider
  - SL 6.5
  - HepOS lib
  - CVMFS
  - HTCondor
  - Cloud init / modules
    - CVMFS
    - Condor
    - Ganglia
  - Vo Profile from CVMFS
  - Site Configuration from CVMFS/Agis
- Valid Defaults only for CC

32GB RAM, 16VCPU, 160GB scratch space

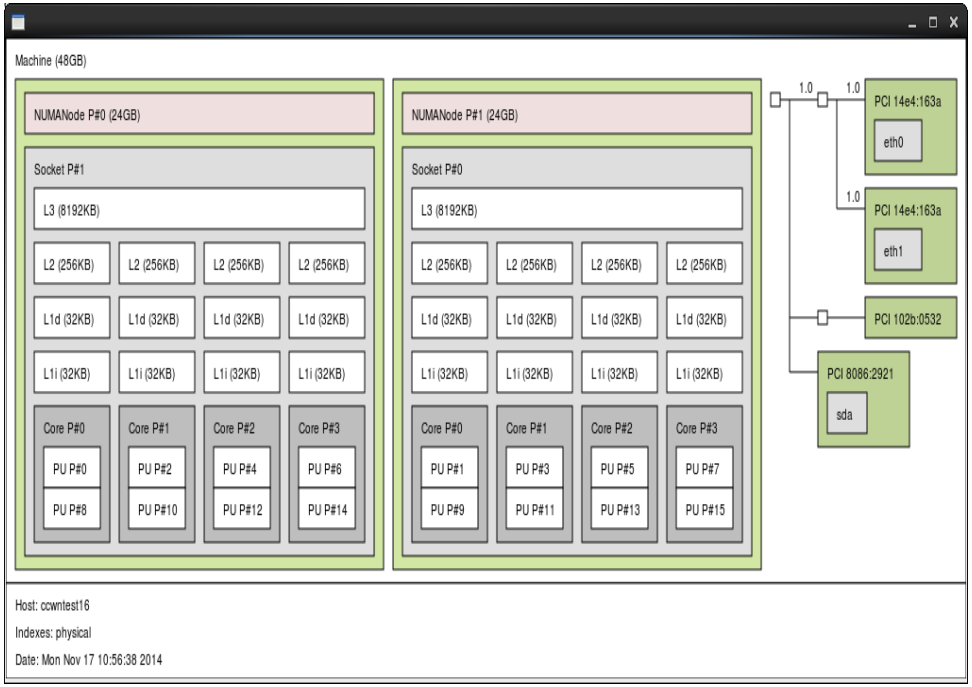
« µCernVM is the heart of the CernVM 3 virtual appliance.
It is based on Scientific Linux 6 combined with a custom, virtualization-friendly Linux kernel. This image is also fully RPM based; you can use yum and rpm to install additional packages ».



- **Small image size ~ 20MB**
- **Include cvmfs, condor, ganglia**
- **Contexualization : Hepix, EC2,Cloud-Init**
- **Strip kernel with latest paravirtualized drivers**
- **Versioned OS on CERVM-FS**
- **Support System update (via UnionFS)**

http://cernvm.cern.ch
J Blomer et al.; 2014 J. Phys.: Conf. Ser. 513 032007 "Micro-CernVM: slashing the cost of building and deploying virtual machines

KVM test + NUMA




```
[root@ccwntest16 vamvakop]# virsh numatune cloudatlas
numa_mode      : strict
numa_nodeset   : 0

[root@ccwntest16 vamvakop]# virsh numatune cloudatlas2
numa_mode      : strict
numa_nodeset   : 1
```

Per-node process memory usage (in MBs)

PID	Node 0	Node 1	Total
17998 (qemu-kvm)	14710.16	0.01	14710.16
18039 (qemu-kvm)	5.43	14744.73	14750.15
Total	14715.58	14744.73	29460.32

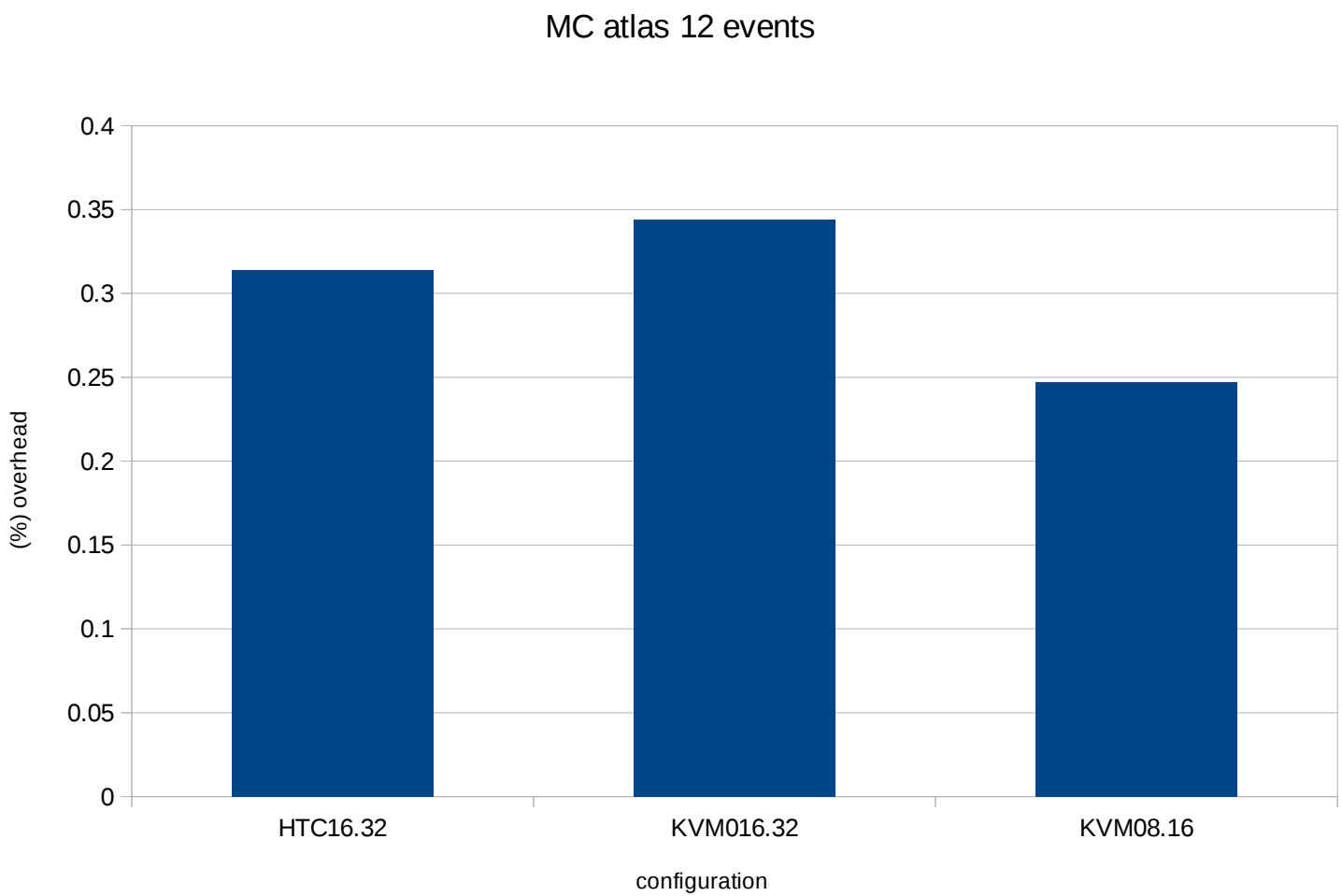
virsh vcpuinfo cloudatlas |grep CPU: |paste - -

CPU:	0	CPU:	0
VCPU:	1	CPU:	2
VCPU:	2	CPU:	4
VCPU:	3	CPU:	6
VCPU:	4	CPU:	8
VCPU:	5	CPU:	10
VCPU:	6	CPU:	12
VCPU:	7	CPU:	14

Cgroup enabled

- Memory
- cpuset

Tests ...


MC atlas 12 events
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
(%) overhead
HTC16.32
KVM016.32
KVM08.16
configuration