

# FermiCloud: A private cloud to support Fermilab Scientific Users

S.Timm, K. Chadwick, D. Yocum, G. Garzoglio, H. Kim, P. Mhashilkar, T. Levshina

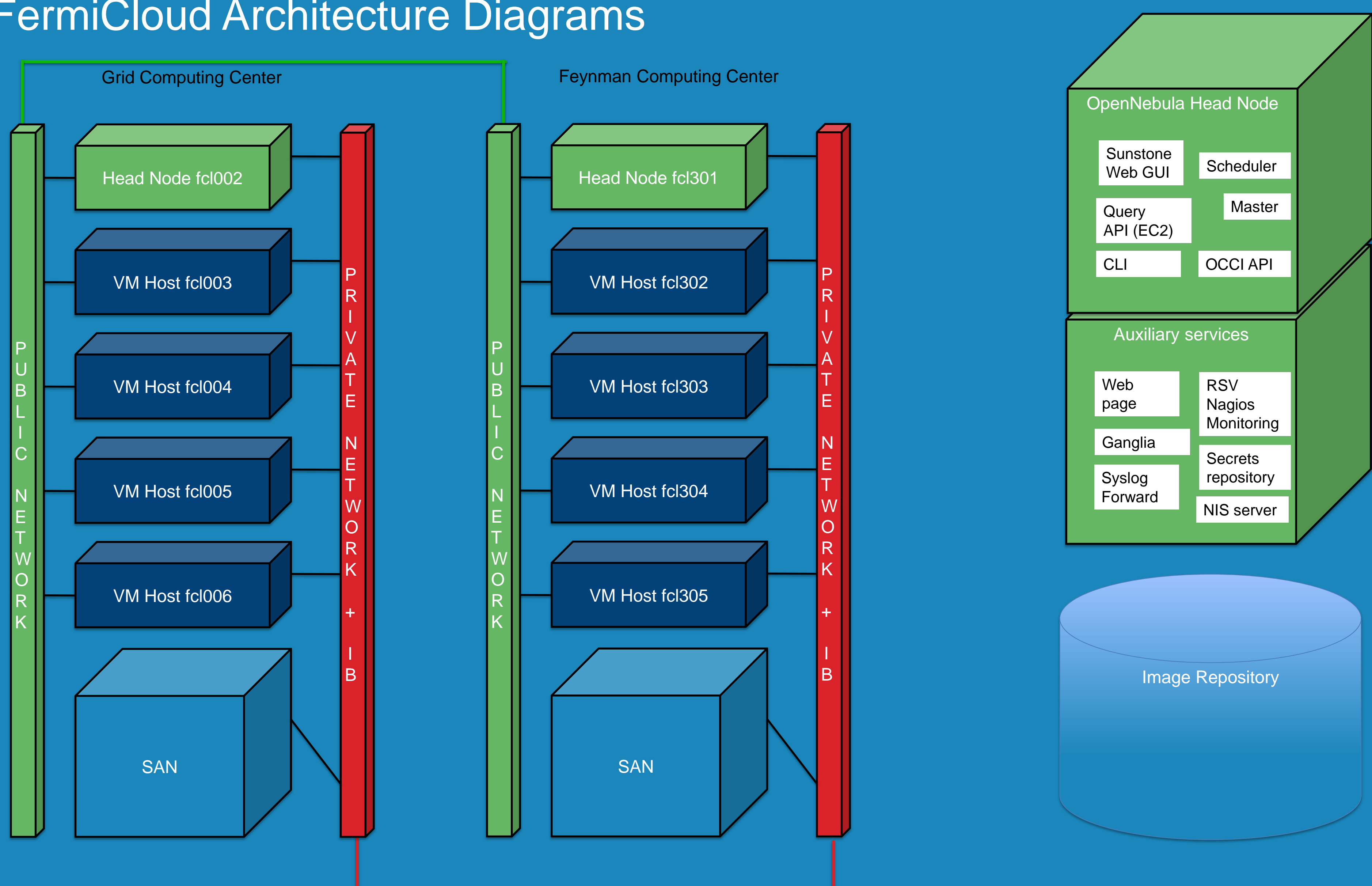
## What is FermiCloud?

- Infrastructure-as-a-service private cloud for Fermilab Scientific Program.
- Integrated into Fermilab site security structure.
- Virtual machines have full access to existing Fermilab network and mass storage devices.
- Scientific stakeholders get on-demand access to virtual machines without system administrator intervention.
- Virtual machines created by users and destroyed or suspended when no longer needed.
- Testbed for developers and integrators to evaluate new grid and storage applications on behalf of scientific stakeholders.
- Ongoing project to build and expand the facility:
  - Technology evaluation, requirements, deployment.
  - Scalability, monitoring, performance improvement.
  - High availability and reliability

## FermiCloud Operations

- Stock virtual machine images are provided for new users.
- Active virtual machines get security patches from site patching services.
- Dormant virtual machines get woken up periodically to get their patches.
- New virtual machines scanned by site anti-virus and vulnerability scanners, don't get network access until they pass.
- Three levels of service:
  - 24 by 7 high availability, can have fixed IP number,
  - 9 by 5 development/integration, use one of a pool of fixed IP's,
  - Opportunistic—Can be pre-empted if idle or if higher-priority users need cloud.

## FermiCloud Architecture Diagrams



## X.509 Authentication

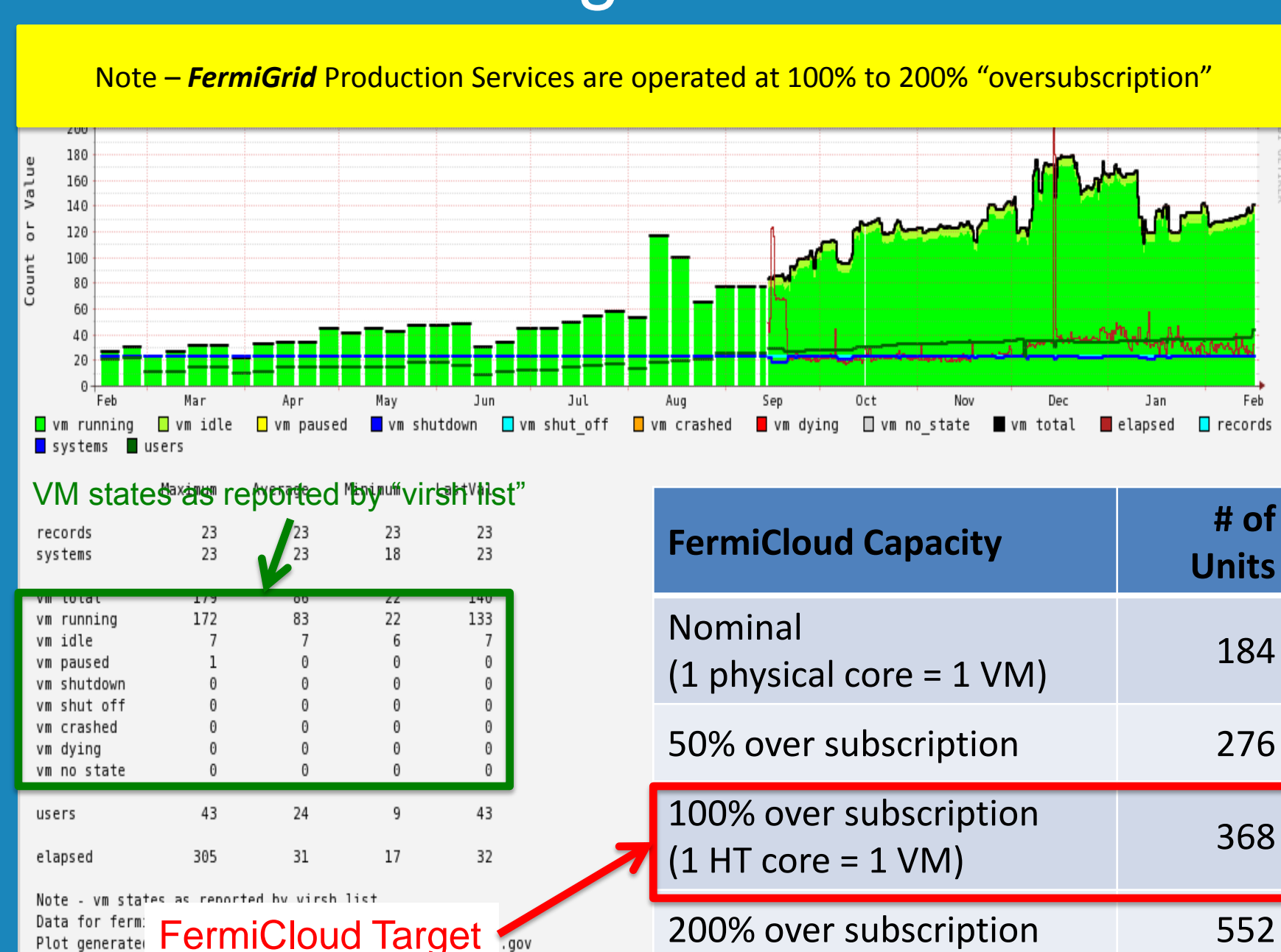
- Use OpenNebula Pluggable authentication feature.
- Wrote X.509 authentication plugin and contributed back to OpenNebula, included in OpenNebula 3.
- X.509 Authentication is integrated into command line tools, EC2 Query API, OCCI API, SunStone management GUI.
- Contributing to standards bodies to make authorization callout to external services, similar to Grid authentication.

## Virtualization and MPI

Configuration	#Host Systems	#VM/host	#CPU	Total Physical CPU	HPL Benchmark (Gflops)
Bare Metal without pinning	2	--	8	16	13.9
Bare Metal with pinning (Note 2)	2	--	8	16	24.5
VM no pinning (Notes 2,3)	2	8	1 vCPU	16	8.2
VM with pinning (Notes 2,3)	2	8	1 vCPU	16	17.5
VM+SRIOV with pinning (Notes 2,4)	2	7	2 vCPU	14	23.6

Notes: (1) Work performed by Dr. Hyunwoo Kim of KISTI in collaboration with Dr. Steven Timm of Fermilab. (2) Process/Virtual Machine "pinned" to CPU and associated NUMA memory via use of numactl. (3) Software Bridged Virtual Network using IP over IB (seen by Virtual Machine as a virtual Ethernet). (4) SRIOV driver presents native InfiniBand to virtual machine(s). 2<sup>nd</sup> virtual CPU is required to start SRIOV, but is only a virtual CPU, not an actual physical CPU.

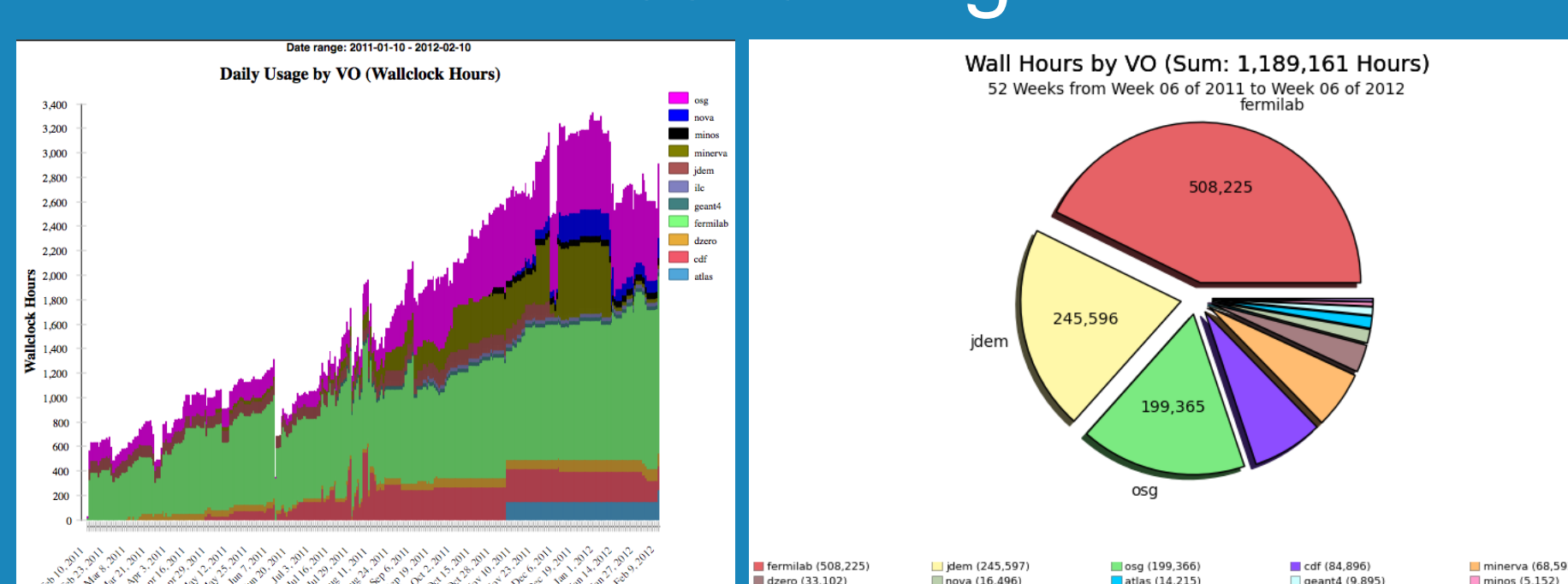
## Monitoring and Metrics



## Grid Cluster On Demand

- Define policy-based expressions for "Idle"
- Detect Idle virtual machines
- Suspend idle virtual machines
- Use vCluster package:
  - Look ahead at batch queue
  - Submit correct virtual machine to FermiCloud
  - Submit to Amazon EC2 if extra capacity needed
  - vCluster a collaboration between Fermilab and KISTI

## Accounting



## High Availability

- Machines in two different buildings
- Mirrored SAN between buildings
- Global shared file system between all nodes
- Copies of all VM's available in both buildings
- Network routable from each building
- Pre-emptive live migration for scheduled outage
- Restart of VM's after unscheduled building failure

