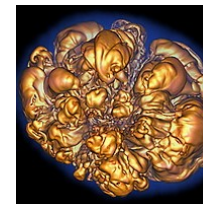
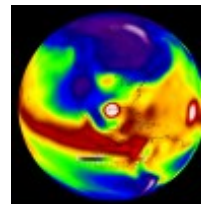
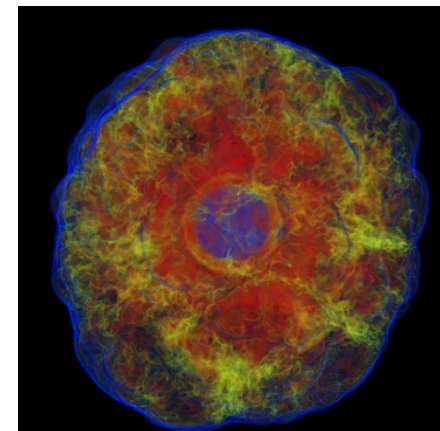
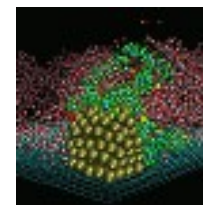
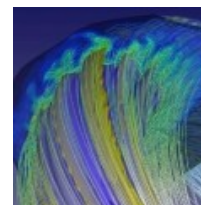
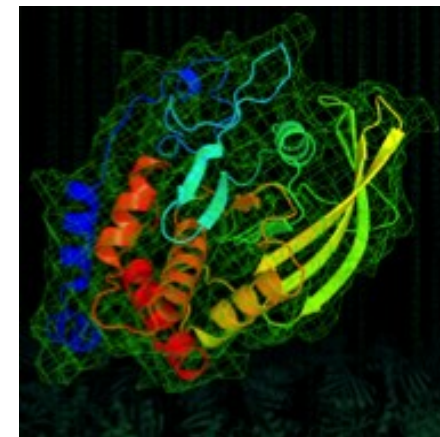
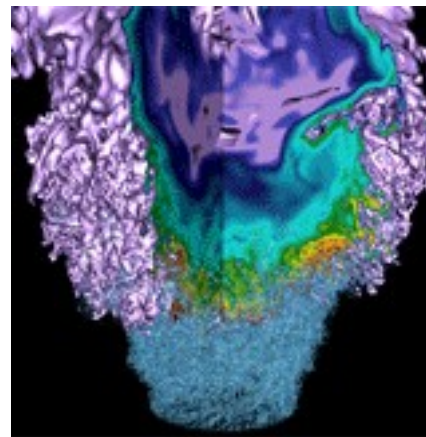


# A slice of the NERSC data collect system



Thomas Davis & Cary  
Whitney

Title

4/21/16

- **NERSC is moving between data centers.**
  - Old data center is a mis-mash of systems.
    - 3 separate building control systems
      - One is Windows 98 based
      - Two are embedded Windows
        - This one is browser based, except..
        - You load an active-x object to check java version
          - And if you do not have the correct version, downloads it from a 3<sup>rd</sup> party website..
        - Then you load java to run a java-ws program..
        - Which then downloads and runs a Windows DLL to provide window decoration.
  - New data center has no chillers.

- **Introduce several new ideas..**

- Clean slate – no legacy to clean up or rebuild.
- Security is paramount.
  - Most BAS/metering protocols are insecure.
- Ability to Fail fast
- Modularity
  - We demand the ability to replace anything, at any time, for any reason.
- Use services that are designed for HA
  - We want failover, maintainability, and scalability
- Zero downtime.
  - We are the first up, and the last down



- **A new way of doing things.**

- System must be first one up, and the last one down.
- System has many sensors, collectors and networking.
  - 1600+ temperature sensors
  - 800+ power meters
    - Substations master breakers, Substation breakers, panel breakers
  - 600+ power strips (Raritan, 50/60 amp is the preferred vendor/model)
    - With breaker level monitoring – 6 breakers per strip.
  - 200+ bacnet devices
    - Many with logical sub-devices.
  - 2k+ ethernet ports
  - 24+ network switches
  - 4 Seismometers
- The data roach hotel.
  - Data checks in, but never checks out..
    - HPSS is your friend.

- No plugins.
  - No java, no active-x, no flash
- HTML5 support.
- Desired to work on Cell phones, tablets, Linux, OSX, Windows.
- Stream processing
  - Ability to calculate a PUE in realtime
  - Fault detection
  - Security incidents
  - Building Control
  - Filtering
- Instrument the data collectors themselves.

## • Metrics 2.0 implementation

- What the heck is metrics 2.0?
- A different way of representing data
- Attaching meta information to the metric.
- See <http://metrics20.org>

### Traditional systems

```
collectd.df.srv-node-dfs10.df-complex.used
```

```
diskspace._srv_node_dfs10.byte_used  
{  
  server: dfs1  
}
```

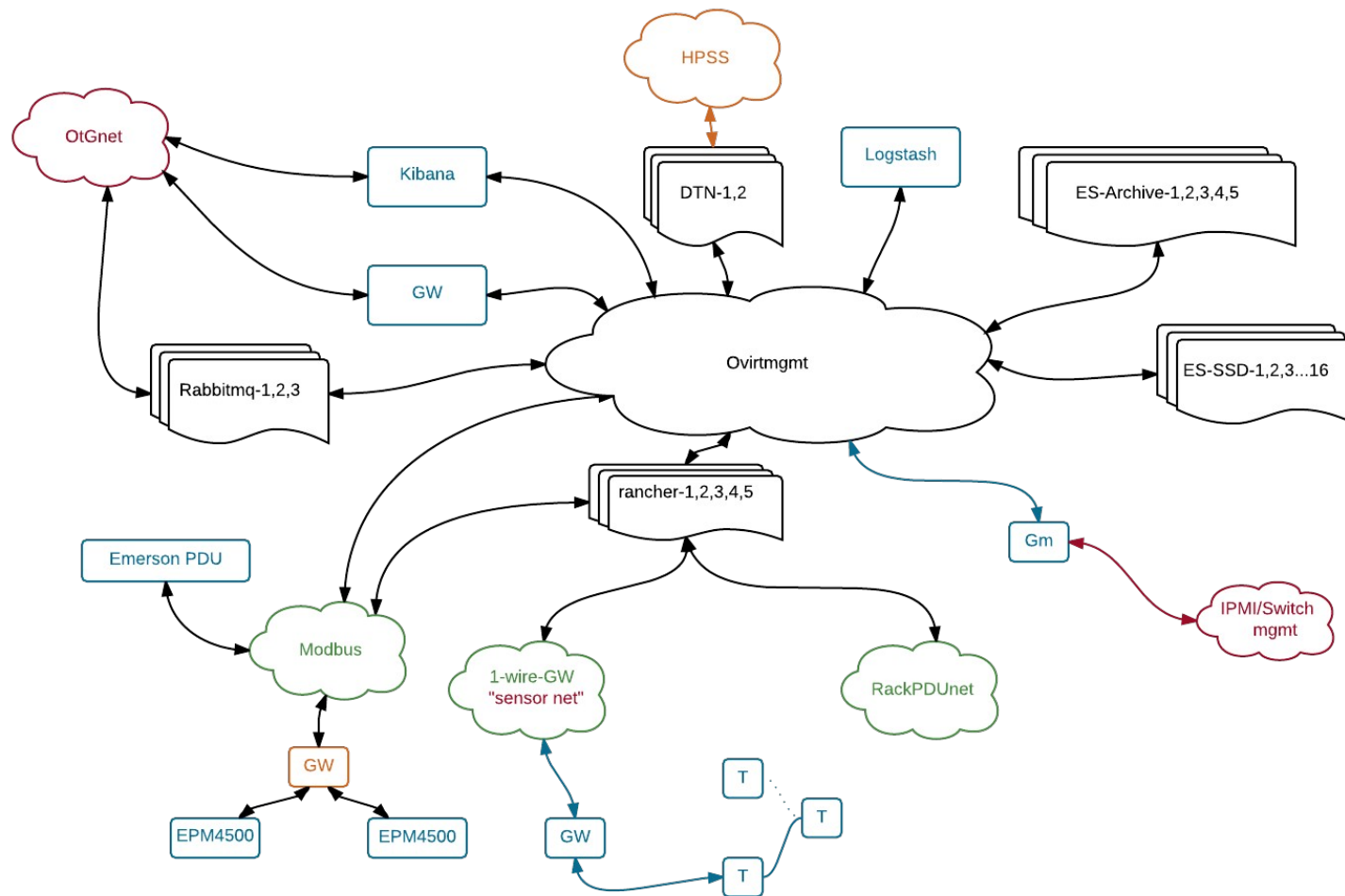
### Metrics 2.0

```
{  
  server: dfs1  
  what: diskspace  
  mountpoint: srv/node/dfs10  
  unit: B  
  type: used  
  metric_type: gauge  
}  
meta: {  
  agent: diamond,  
  processed_by: statsd2  
}
```

- **Ovirt (RHEL) VM management**
  - VLAN tagging/trunks
    - PXEboot needs untagged packets..
  - Dense servers – 8 nodes in 4U slot.
    - Supermicro Fat-twin based
- **Rancher with Docker containers**
- **Split networks**
  - Have to connect to 10 different networks.
  - 3 public networks
  - 7 private networks

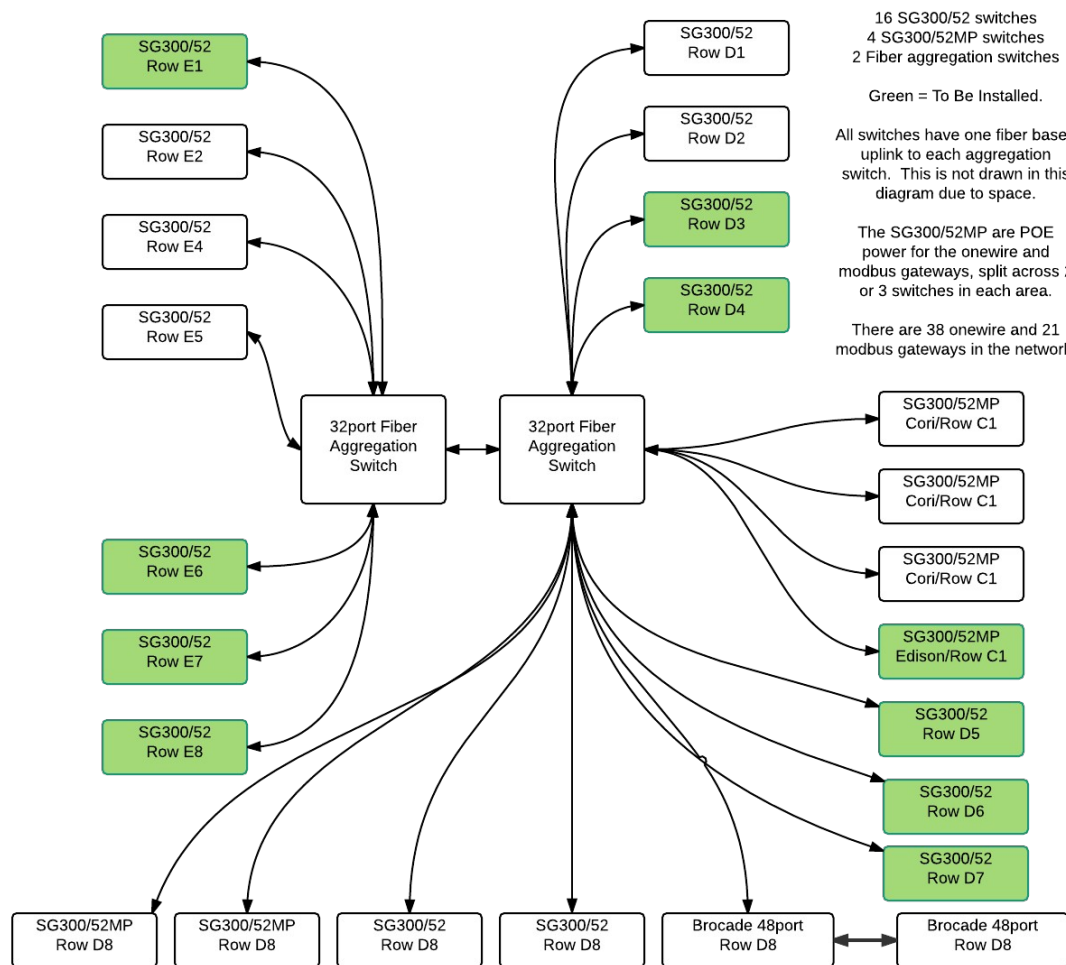


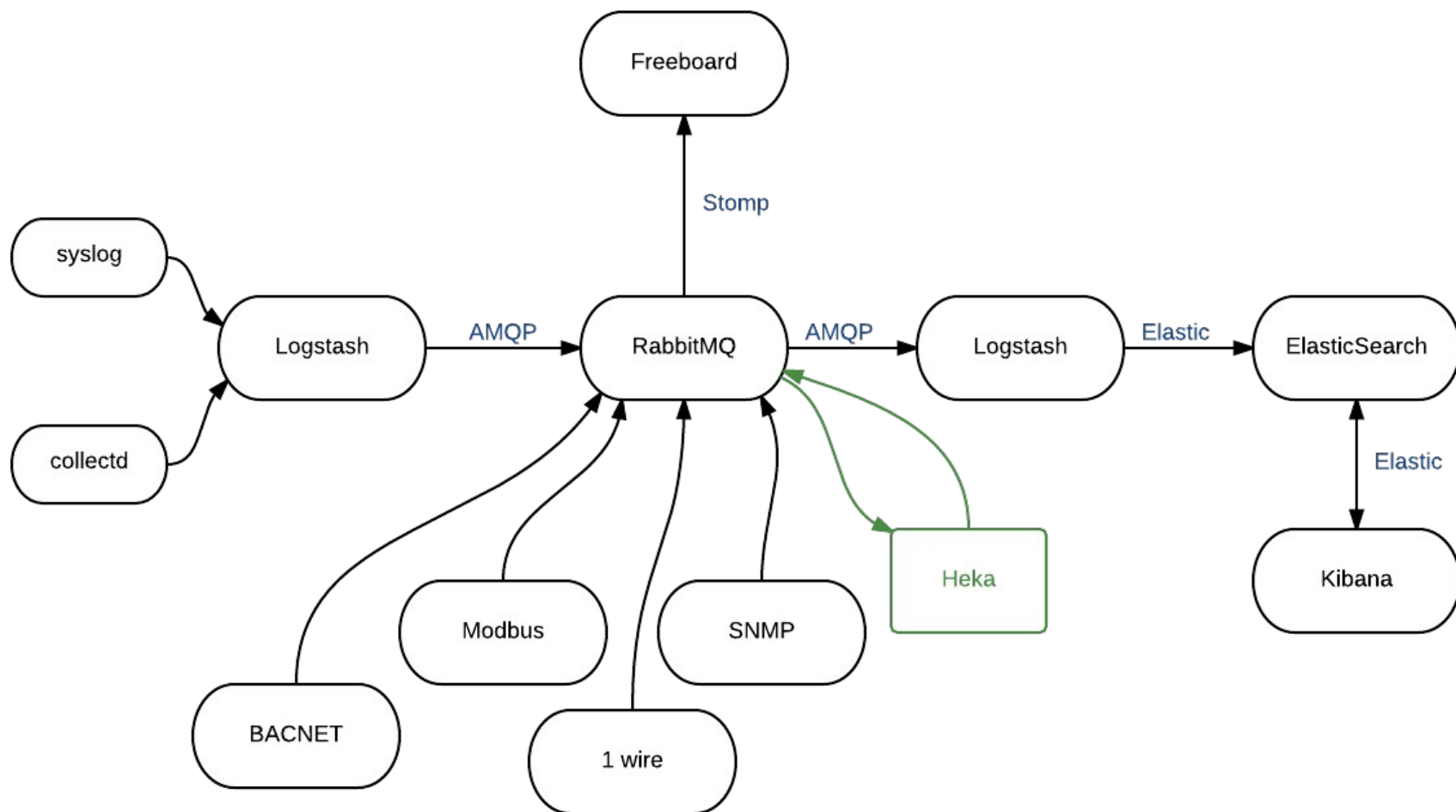






# CRT Switch Network



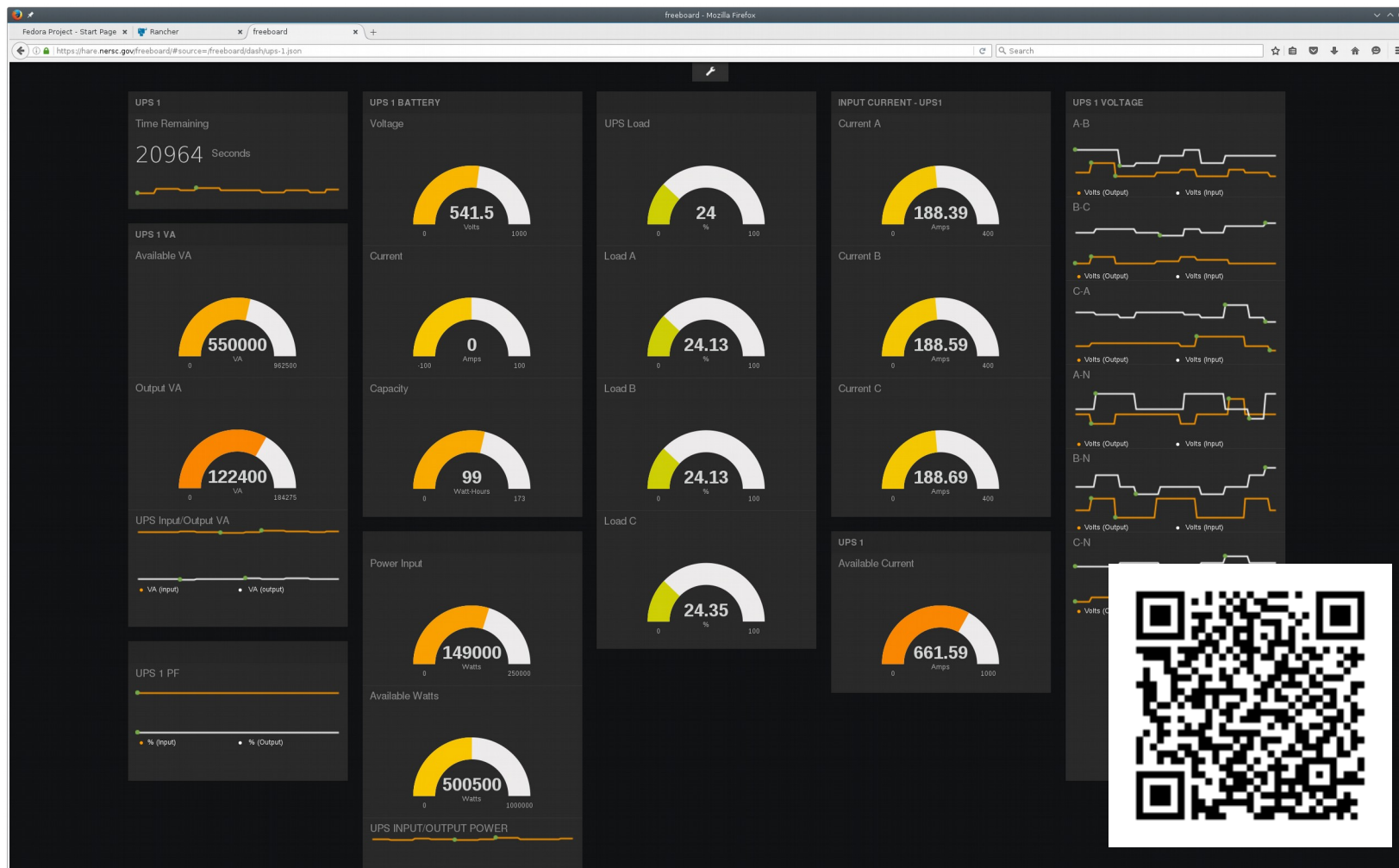


What the logical translates into.

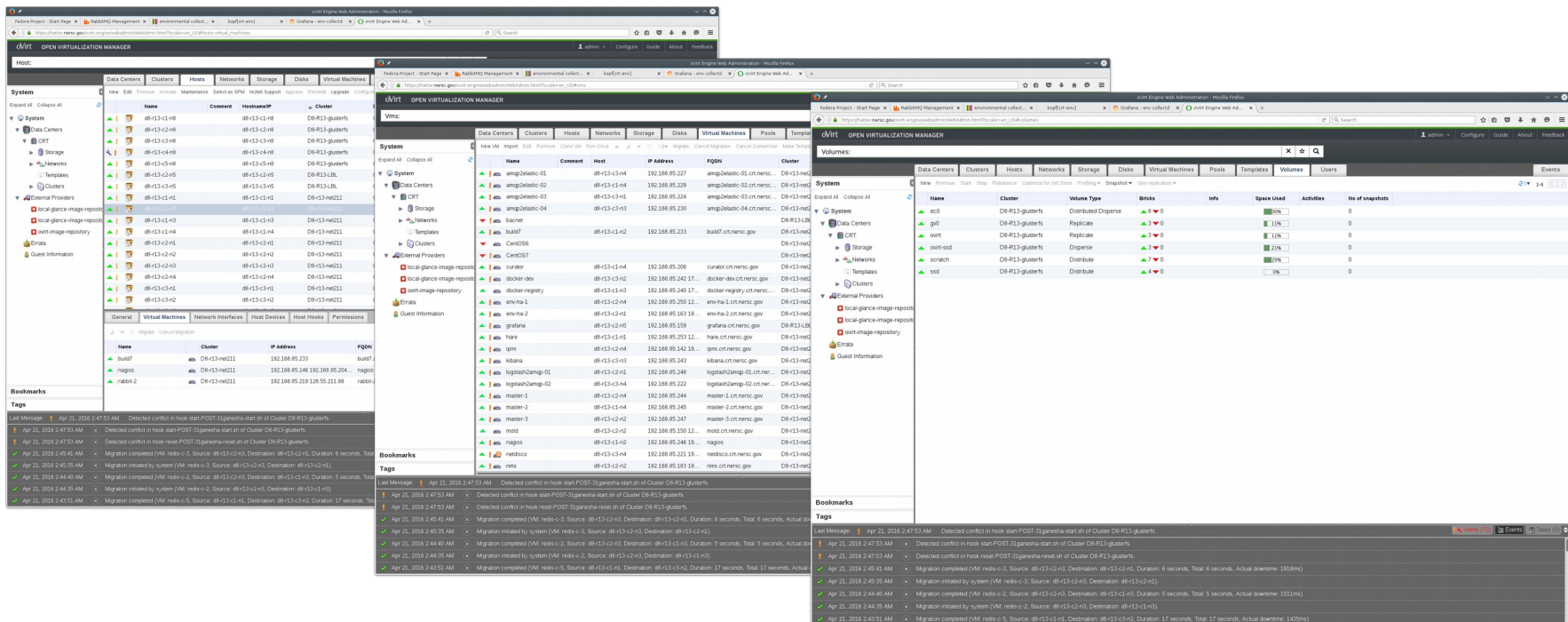


- **15 Elasticsearch (ingest) nodes w/SSD and 64GB ram**
- **5 Elasticsearch (archive) nodes w/8TB disk and 64GB ram**
- **3 dedicated Master nodes (both Consul & Elasticsearch) (vm)**
- **2 internal logstash2amqp nodes (ha) (vm)**
- **1 Kibana/Elasticsearch worker node (no data/not a master) (vm)**
- **4 internal amqp2logstash nodes (vm)**
- **3 rabbitmq-servers in ha/mirror cluster mode (vm)**
- **6 node redis cluster (vm)**
- **8 Rancher/docker nodes**
  - 1 VM for Substation network
  - 2 VM for UPS network
  - 2 real nodes (dual socket, 32 cores, 64GB of ram)
  - 3 floating VM's

# Freeboard and Metrics 2.0 using rabbitmq/Stomp/Websocket



- <http://ovirt.org>
- Redhat's RHEV product, opensource version.
  - We are currently on the 3.6.x version.

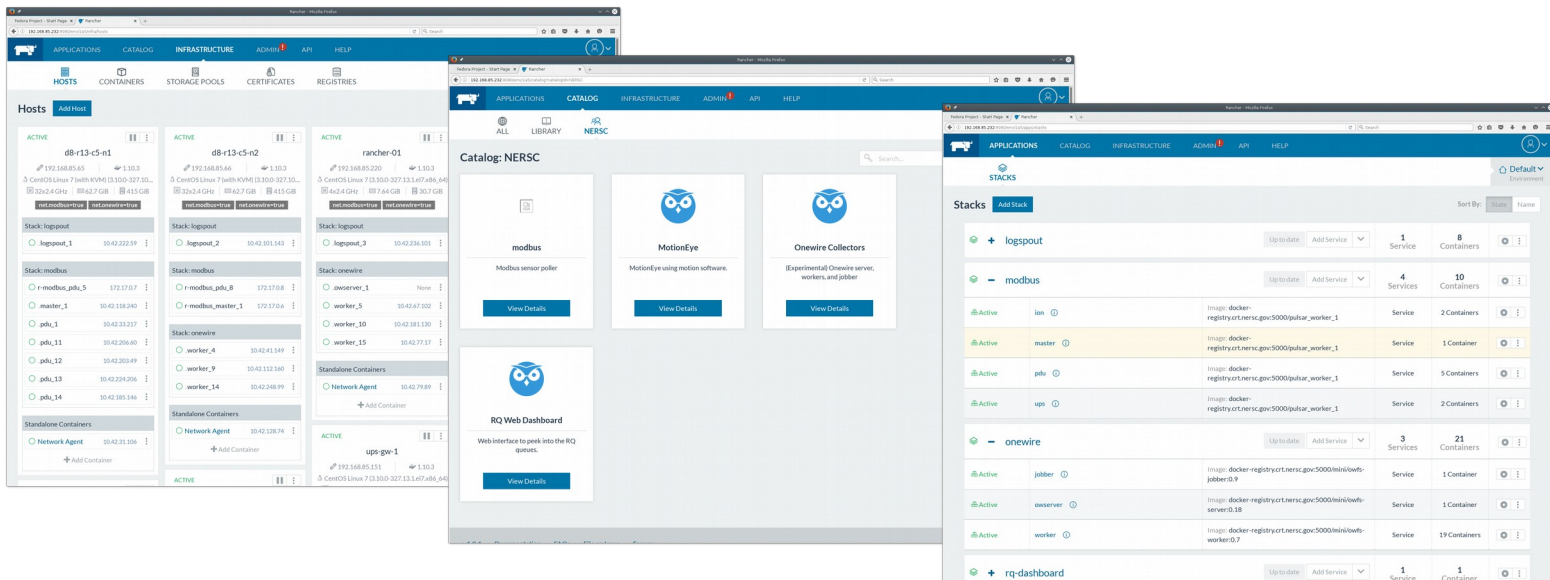


The image displays three overlapping screenshots of the oVirt web management interface. The leftmost screenshot shows the 'System' overview with a table of hosts and clusters. The middle screenshot shows the 'Virtual Machines' tab, listing various VMs like 'amp-beltic-01' through 'amp-beltic-04'. The rightmost screenshot shows the 'Storage' tab, displaying storage pools and their usage. All screenshots show a detailed view of the infrastructure, including host names, IP addresses, and storage configurations.

- KVM/qemu based
- GlusterFS, nfs, iscsi volume management
- Web interface
- Understands VLANs, bridges
- Can talk to Openstack components
- Manages images, templates for you.



- <http://rancher.com>
  - Recently released 1.0.1
    - Oh no.. a x.0 release!
    - Does have some irritating bugs.
    - But it does solve several major problems!

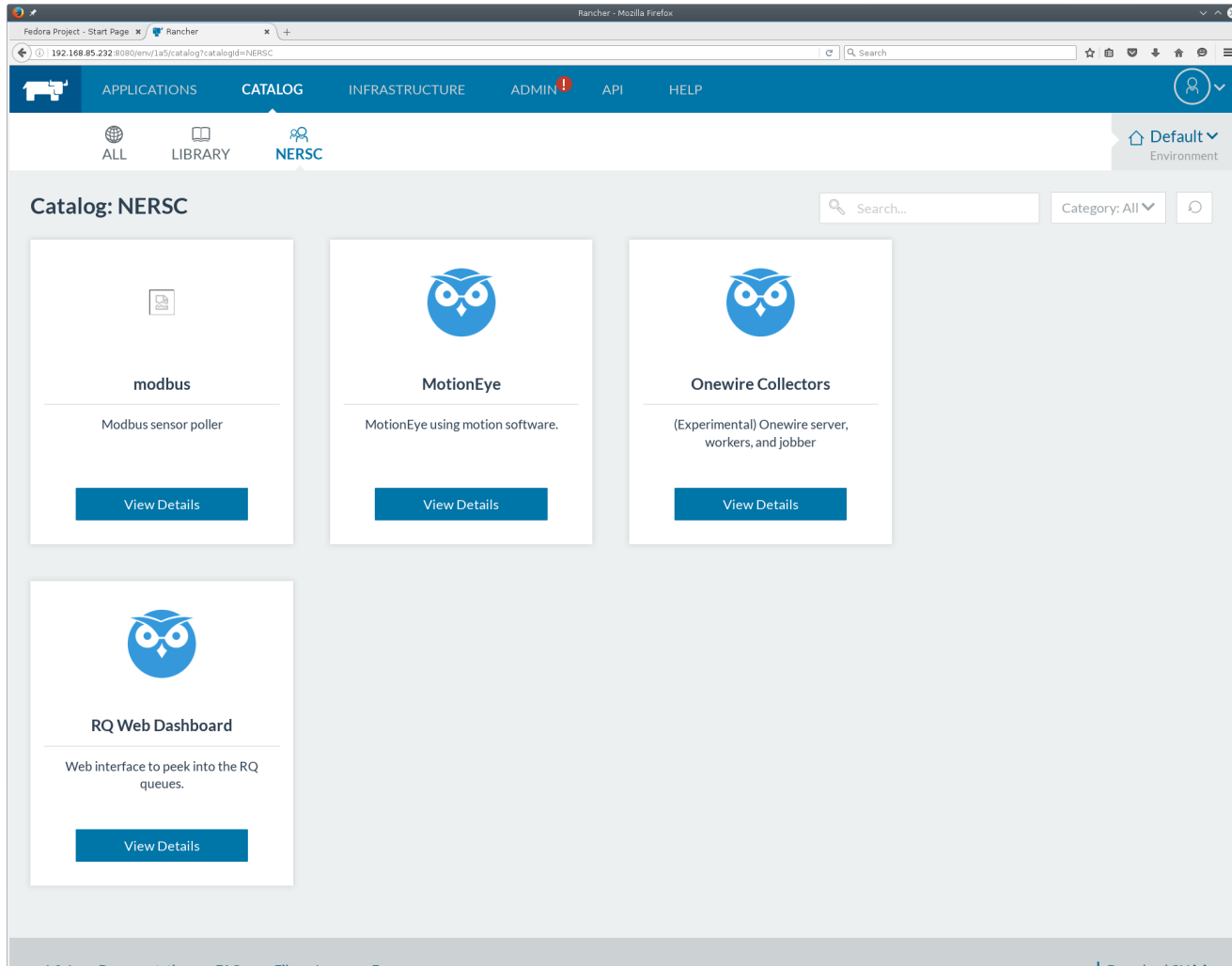




- Supports docker, kubernetes, and Docker Swarm.
- Internal, encrypted network between containers.
- HTML and CLI interfaces.
  - Docker compose file support
- Catalog for applications.
- Running on both VM's and bare metal systems.
- They even have a Docker based OS – RancherOS
  - Like CoreOS/AtomicOS, but geared towards virtual machines.
- Convoy for docker volume management using GlusterFS/NFS.

- **One click deployment of application stack.**
  - **Scales containers, creates internal network, manages volumes**
  - **Rollback, upgrades are simplified.**
  - **Schedules containers on the proper nodes.**
    - **Based on tags.**
    - **We use this for network routing for example.**

# Catalog..



Setup.



Rancher - Mozilla Firefox

Fedora Project - Start Page x RabbitMQ Management x environmental collect... x kopf[crd-env] x Grafana - env collectd x svirt Engine Web Ad... x freeboard x Rancher x +

192.168.85.232:8080/env/1a5/catalog/nerSC:oneWire?catalogId=NERSC

APPLICATIONS CATALOG INFRASTRUCTURE ADMIN API HELP

ALL LIBRARY NERSC

Default Environment

Catalog: Onewire Collectors

### Add Onewire Collectors Stack

(Experimental) Onewire server, workers, and jobber

Catalog: NERSC  
Category: Onewire  
Maintainer: Thomas Davis <tadavis@lbl.gov>

#### Template Version

0.1

Select a version of the template to deploy

#### New Stack

Name: onewire Description: Description

#### Configuration Options

Worker Nodes:\* 1

Number of workers for collection.

☒ Start services after creating

PREVIEW

Launch Cancel

# Production..



Rancher - Mozilla Firefox

Fedora Project - Start Page x RabbitMQ Management x environmental collect... x kopf[ert-erv] x Grafana - env collectd x oVirt Engine Web Ad... x freeboard x Rancher x

192.168.85.232:8080/env/1a5/apps/stacks Search

APPLICATIONS CATALOG INFRASTRUCTURE ADMIN! API HELP

STACKS Default Environment

Stacks Add Stack Sort By: State Name

logspout			Up to date	Add Service	1 Service	8 Containers	
Active	logspout	Image: rancher/logspout-logstashv0.2.0	Service	8 Containers			
modbus			Up to date	Add Service	4 Services	10 Containers	
Active	ion	Image: docker-registry.crt.nersc.gov:5000/pulsar_worker_1	Service	2 Containers			
Active	master	Image: docker-registry.crt.nersc.gov:5000/pulsar_worker_1	Service	1 Container			
Active	pdu	Image: docker-registry.crt.nersc.gov:5000/pulsar_worker_1	Service	5 Containers			
Active	ups	Image: docker-registry.crt.nersc.gov:5000/pulsar_worker_1	Service	2 Containers			
onewire			Up to date	Add Service	3 Services	21 Containers	
Active	jobber	Image: docker-registry.crt.nersc.gov:5000/mini/owfs-jobber.0.9	Service	1 Container			
Active	owserver	Image: docker-registry.crt.nersc.gov:5000/mini/owfs-server.0.18	Service	1 Container			
Active	worker	Image: docker-registry.crt.nersc.gov:5000/mini/owfs-worker.0.7	Service	19 Containers			
rq-dashboard			Up to date	Add Service	1 Service	1 Container	
Active	rq-dashboard	Image: docker-registry.crt.nersc.gov:5000/mini/rq-dashboard:0.1 Ports: 9181	Service	1 Container			

v1.0.1 Documentation FAQs File an Issue Forums Download CLI

- Service discovery/registration with Consul
  - Uses Nagios checks for health and integration in nagios.
  - DNS routing and load balancing
- Used by onewire stack to route onewire api requests to the proper gateway.
- Nagios also checks for service health.
- Also used to provide round-robin DNS name service for internal services.
- Future uses include configuration templating and key value pairs.