Ubiquitous Edge Platform

Lincoln Bryant
Rob Gardner

US ATLAS Distributed Facilities Meeting, 14 March 2016

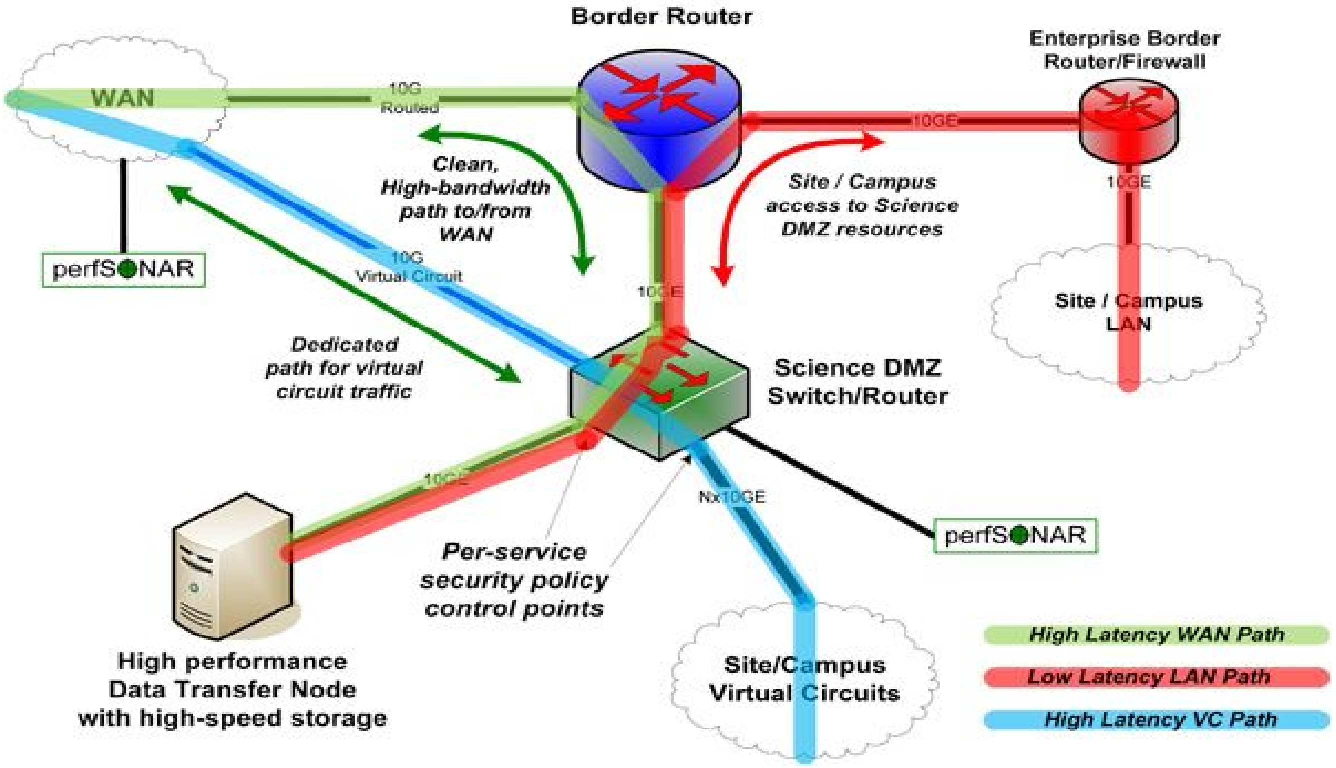
Ubiquitous & Easy “CI Substrate”

- Pioneer a new phase of advanced cyberinfrastructure deployment, allowing sites to flexibly evolve and sustain both on-premise and commercial cloud-based infrastructure
- Hosted services, such as CEs, data caches, squid, etc., could be centrally deployed onto “CI substrates” within a trusted CI zones and remotely operated, upgraded, and optimized for performance
- Extend to shared, opportunistic university clusters and cloud resources

Distributed Virtualized Data Centers

- Reduce IT footprint and ops burden
    - Centralize deployment & ops; reduce local admin cost
- Explore virtualized data center frameworks
    - E.g. container management over bare metal or VMs
- “Blue sky” goal
    - Establish a “trusted pattern” for a “CI substrate” on sites
    - Create distributed virtualized data center(s) overlaying the fabric substrate

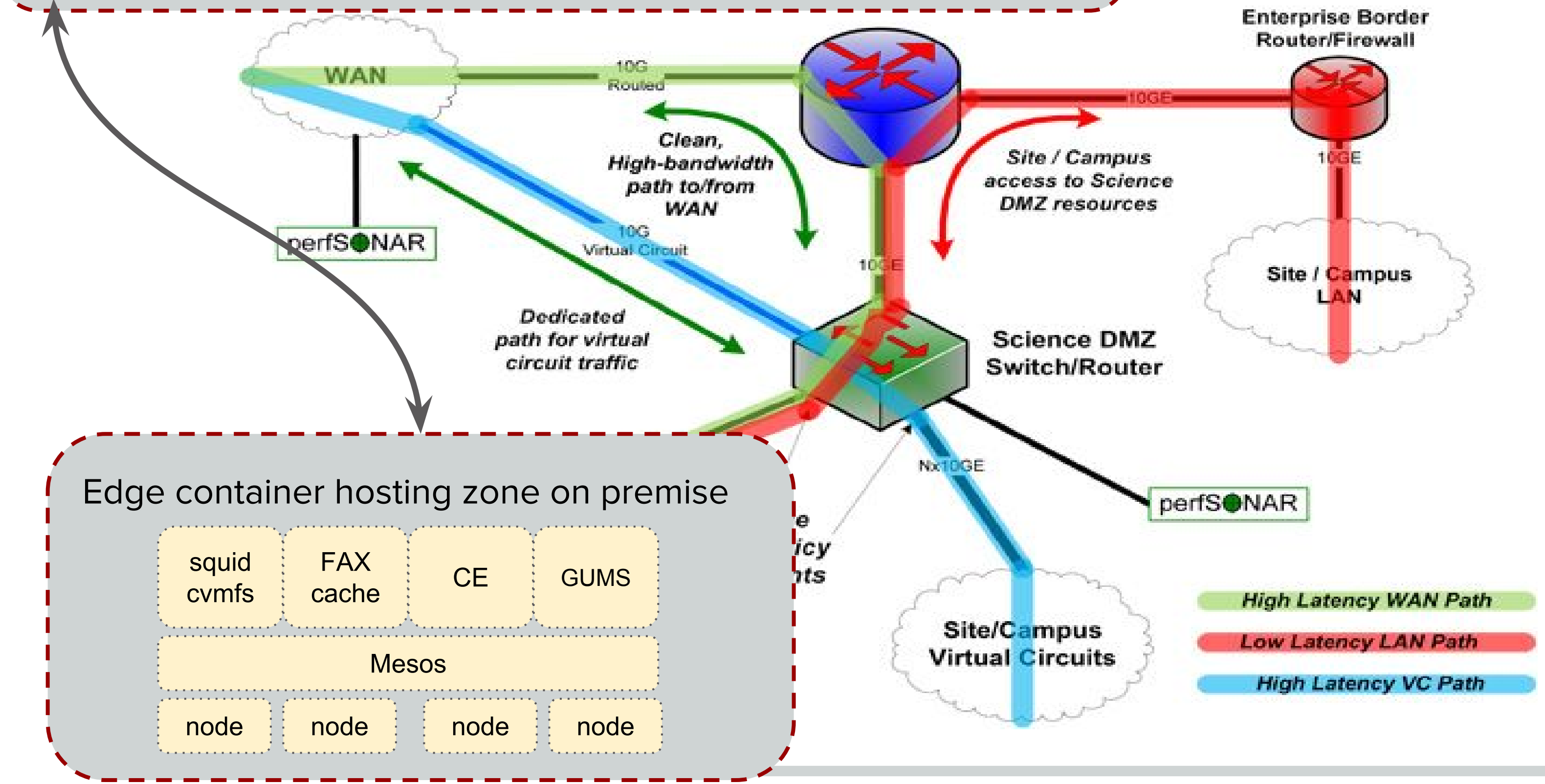
Canonical SciDMZ


Border Router
Enterprise Border Router/Firewall
WAN
10G Routed
10GE
Clean, High-bandwidth path to/from WAN
Site / Campus access to Science DMZ resources
10GE
perfSONAR
10G Virtual Circuit
10GE
Site / Campus LAN
Dedicated path for virtual circuit traffic
Science DMZ Switch/Router
10GE
NxJ0GE
perfSONAR
Per-service security policy control points
High performance Data Transfer Node with high-speed storage
Site/Campus Virtual Circuits
High Latency WAN Path
Low Latency LAN Path
High Latency VC Path

SciDMZ with CI Substrate

US ATLAS Facility central ops console:

```
$ slate install osg-squid.3.3 --sites MWT2 ALGT2 SWT2 WT2 NET2
```


WAN
10G Routed
Enterprise Border Router/Firewall
10GE
Clean, High-bandwidth path to/from WAN
Site / Campus access to Science DMZ resources
10GE
perfSONAR
10G Virtual Circuit
10GE
Site / Campus LAN
Dedicated path for virtual circuit traffic
Science DMZ Switch/Router
e
icy
nts
Nx10GE
perfSONAR
Site/Campus Virtual Circuits
High Latency WAN Path
Low Latency LAN Path
High Latency VC Path

Edge container hosting zone on premise

squid cvmfs	FAX cache	CE	GUMS
Mesos			
node	node	node	node

Deploying research software at the edge

Open Science Grid

perfSONAR

Xrootd Cache

globus

Your favorite project here!

Hardware

- Produce reference specification for supportability reasons
  - No more than 2-3 vendor options.
- Cloud providers like Joyent have done a really good job in this space. Something similar to:
  - https://docs.joyent.com/private-cloud/hardware/specs

Operating system

- Many choices to evaluate in this area
- Traditional distributions:
  - EL, Ubuntu, etc
- Upcoming projects building around containers:
  - CoreOS, Boot2Docker, RancherOS, Project Atomic
- Exotic alternatives:
  - SmartOS (Solaris-based, emulating Linux kernel ABI)

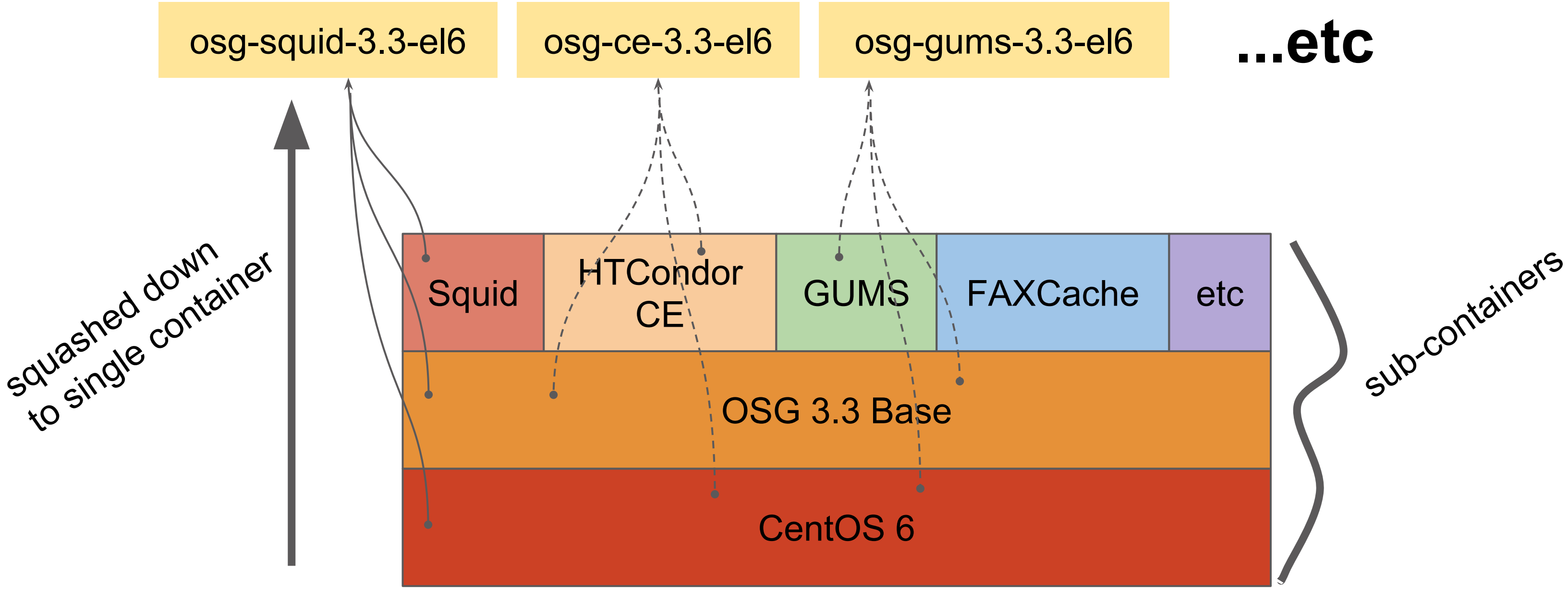
Software

- Microservices-y architecture
  - Follow the Docker model of 1 application per container
- Service discovery and configuration tools
  - Consul, etcd, etc
- Scheduling
  - Kubernetes, Docker Swarm, Fleet, Mesos
  - (HTCondor?)

Software

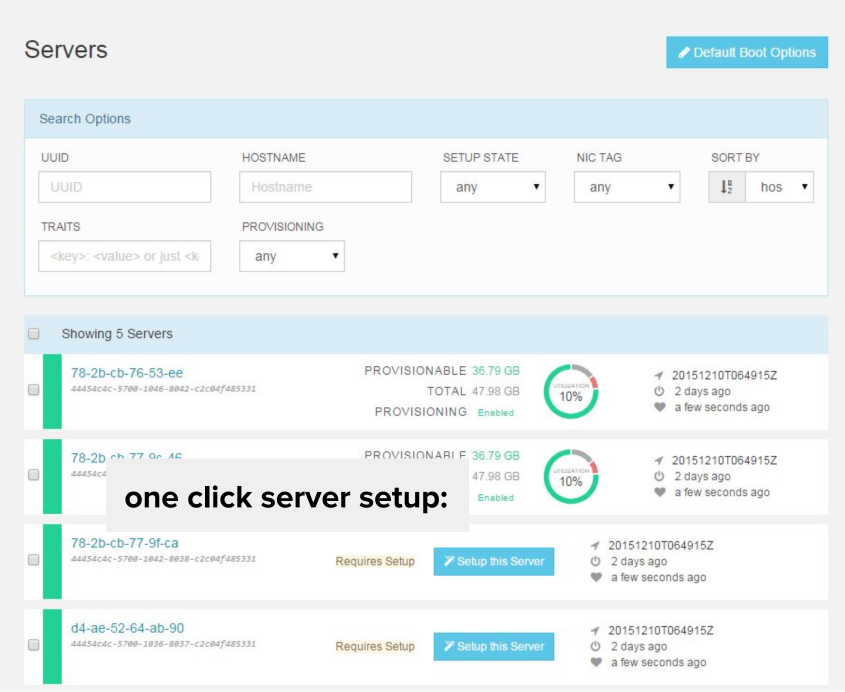
- Dockerized applications created, vetted, maintained by central operations team.
    - Pushed by operators down to subscribed sites
    - Or, pulled by local site admins without interaction with central.
- Built-in monitoring
    - Graphite, ELK, etc

Containerizing Services



Frontier-Squid Containerized

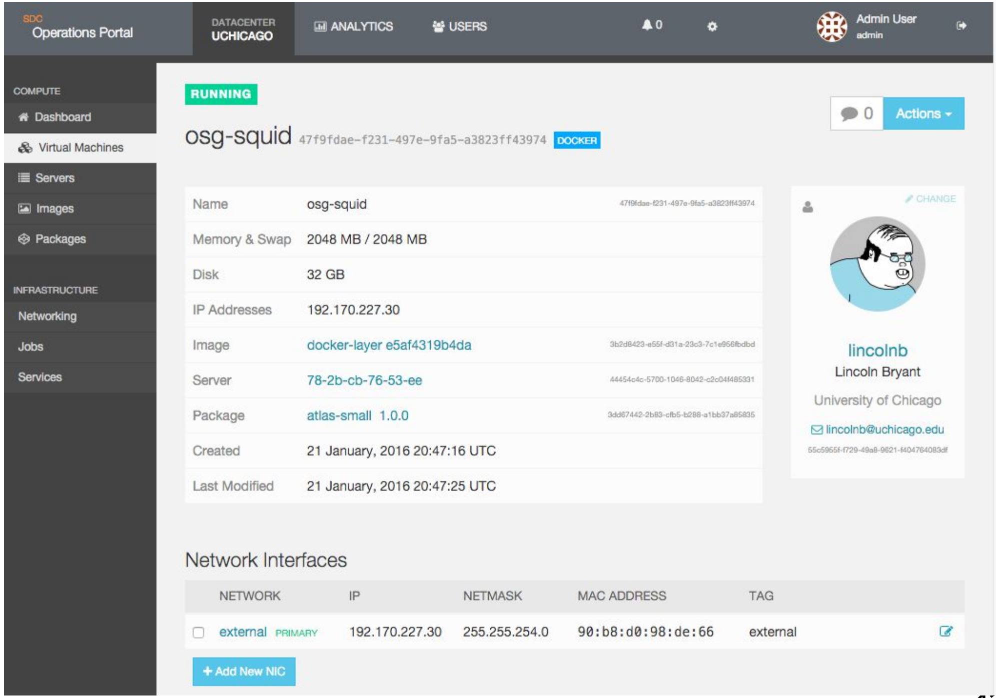
- Deployed in a hybrid cloud @ Midwest Tier 2:



□ Showing 5 Servers

	Server	Capacity	Utilization	Status
□	78-2b-cb-76-53-ee 44454c4c-5700-1046-8042-c2c04f485331	PROVISIONABLE 36.79 GB TOTAL 47.98 GB PROVISIONING Enabled	UTILIZATION 10%	20151210T064915Z 2 days ago a few seconds ago
□	78-2b-cb-77-9c-46 44454c4	PROVISIONABLE 36.79 GB 47.98 GB Enabled	UTILIZATION 10%	20151210T064915Z 2 days ago a few seconds ago
□	78-2b-cb-77-9f-ca 44454c4c-5700-1042-8038-c2c04f485331	Requires Setup	Setup this Server	20151210T064915Z 2 days ago a few seconds ago
□	d4-ae-52-64-ab-90 44454c4c-5700-1036-8037-c2c04f485331	Requires Setup	Setup this Server	20151210T064915Z 2 days ago a few seconds ago

one click server setup:



osg-squid 47f9fdae-f231-497e-9fa5-a3823ff43974 DOCKER

Name	osg-squid	47f9fdae-f231-497e-9fa5-a3823ff43974
Memory & Swap	2048 MB / 2048 MB	
Disk	32 GB	
IP Addresses	192.170.227.30	
Image	docker-layer e5af4319b4da	3b2d8423-e55f-d31a-23c3-7c1e956fbdbd
Server	78-2b-cb-76-53-ee	44454c4c-5700-1046-8042-c2c04f485331
Package	atlas-small 1.0.0	3dd67442-2b83-cfb5-b288-a1bb37a85835
Created	21 January, 2016 20:47:16 UTC	
Last Modified	21 January, 2016 20:47:25 UTC	

lincolnb
Lincoln Bryant
University of Chicago
lincolnb@uchicago.edu
55c5965f-f729-49a8-9621-f404764083df

Network Interfaces

	NETWORK	IP	NETMASK	MAC ADDRESS	TAG
□	external PRIMARY	192.170.227.30	255.255.254.0	90:b8:d0:98:de:66	external

Benefits for ATLAS

- Easily deploy Tier 2 and Tier 3 services
    - PROOF on Demand
    - Remote desktop / NX
    - FAX doors
    - XRootD caches
    - etc

Current pain points

- Many points where human interaction is currently needed
  - Can we automate here?
- Is it possible for me to stand up, then destroy an entire ATLAS site in an automated way?
  - CE, SE, all interactions with AGIS, etc.

Security considerations

- Who has root on the machine?
- Can trusted users allocate resources and start containers remotely without having root?
  - Unprivileged containers are semi-working in newer kernels, but here be dragons..
- Ultimately: What is the correct privilege separation between owner and operator?

Other considerations

- Should there be a VPN / control channel setup such that these nodes are all accessible via the same private IP space?
- Can we use this platform as a testbed for things like SDN?
- What does it look like when we have multiple nodes per site?

Summary

- Platform for edge services on Science DMZs
- Container-based applications, maintained by a central team
- Built-in service discovery, configuration, and monitoring
- Flexible, adaptable to the needs of other projects.

Thank you!
Questions?