OPENSTACK CLOUD AT YAHOO SCALE: HOW TO AVOID DISASTER

Anders Beitnes
Yahoo Paranoids
Yahoo Scale

- 800 million + monthly users
- 10x traffic spikes
- News, finance, millions/day
- Mail 110+ million daily users, 170+ petabyte data, 30 billion emails sent daily.
- Mobile photo uploads +50% quarter over quarter
- We’re hiring many developers, to experiment, build, and deliver.
OpenStack Reach
Slice and Dice

- Multiple Private Clouds
- The production clusters are in separate security domains.
- We control what enters and leaves the VMs.
- Employee access only.
- We own the VMs and the data they contain, and can inspect it.
- We manage more than VMs
Don’t Panic!
Loneliness can be a good thing

- Who should the Infrastructure nodes talk to?
- Hypervisors should listen to their parents.
- Who should VMs talk to?
Who should talk to the Corporate Infrastructure nodes?

while true; do ssh ihatetowait.corp.ne1.yahoo.com; sleep 5; done
Who should talk to the Production Infrastructure nodes?

Production Security Domain
- Automated Configuration

Corporate Security Domain
- Service Engineers

Infrastructure Security Domain
- 3 x API
- 2 x Rabbit MQ
- 2 x MySQL
Who should the Infrastructure nodes talk to?

Infrastructure Nodes

- nova-compute hypervisor
- DNS IP address assignment
- Operations database
- ETC.
Who should the Hypervisors talk to?

Infrastructure Security Domain

Hypervisor Security Domain
No VM Migration?
Who should the VMs talk to?
Who should talk to the VMs?

- User
- Spice-proxy
- Nova-consoleauth
- Hypervisor

Token, Cookie Credentials
Validate Token, Cookie Credentials
If OK, Connect to SPICE server for VM

Messages for SPICE console
Proxied Messages for SPICE console

Corporate Security Domain
Infrastructure Security domain
Hypervisor Security Domain
One Infrastructure node cluster to rule them all
What about those responsible for the Infrastructure nodes?
Keystone
Keystone policy file

{
    "admin_required": [["role:admin"], ["is_admin:1"]],
    "gsd_admin": [["rule:admin_required"], ["role:GSDAdmin"]],
    "default": [["rule:admin_required"]],

    "get_role": [["rule:gsd_admin"]],
    "get_roles": [["rule:gsd_admin"]],
    "get_role_refs": [["rule:gsd_admin"]],
    "get_all_projects": [["rule:gsd_admin"]],
    "get_project": [["rule:gsd_admin"]],
    "get_project_users": [["rule:gsd_admin"]],
    "get_user": [["rule:gsd_admin"]],
    "get_users": [["rule:gsd_admin"]],
    "get_user_roles": [["rule:gsd_admin"]]
}
Impostors and Spies

How do we determine which physical hosts to trust?

What are the issues with VM snapshot integrity?

How to protect other VMs on the same physical host?

How do we protect our storage systems?
Someone to watch over me
Next generation networks

- What are the issues with Networking at scale?
- L2 networks provide VLAN isolation, but have an associated cost problem
- L3 networks solve the cost issue, but pose a security issue.
L2 Network

Internet

BGP

WAN Router1 — WAN Router2

OSPF

Core sw1 — Core sw2

STP

app1 — app2

rack sw

Backplane
L2 network + Yahoo Scale

Security zone 1

Backplane A
Backplane D
Backplane G
Backplane H

Security zone 2

Backplane C

Security zone 3

Fabric Layer
4-16 wide, build to scale

CLR1
CLR2

Core sw1
Core sw2

Fabric Layer
4 wide

Fabric Layer
8 wide
Next generation networks at Yahoo Scale

Virtual Chassis 1

Virtual Chassis 2

Virtual Chassis 3

Virtual Chassis 4

IBGP

Rack Switch 1

Rack Switch 1

Rack Switch 512

Security zone 1
Thanks!
Credits and attributions

- [https://www.flickr.com/photos/andrewgustar/7280161708](https://www.flickr.com/photos/andrewgustar/7280161708)
  - [https://creativecommons.org/licenses/by-nd/2.0/](https://creativecommons.org/licenses/by-nd/2.0/)
- [https://www.flickr.com/photos/arbron/8278577848/](https://www.flickr.com/photos/arbron/8278577848/)
  - [https://creativecommons.org/licenses/by-nd/2.0/](https://creativecommons.org/licenses/by-nd/2.0/)
- [https://www.flickr.com/photos/clearlyambiguous/131620637](https://www.flickr.com/photos/clearlyambiguous/131620637)
  - [https://creativecommons.org/licenses/by-nd/2.0/](https://creativecommons.org/licenses/by-nd/2.0/)
- [https://www.flickr.com/photos/87913776@N00/460181144](https://www.flickr.com/photos/87913776@N00/460181144)
  - [https://creativecommons.org/licenses/by-nd/2.0/](https://creativecommons.org/licenses/by-nd/2.0/)
- [https://www.flickr.com/photos/macgodbrad/4869419570](https://www.flickr.com/photos/macgodbrad/4869419570)
  - [https://creativecommons.org/licenses/by-sa/2.0/](https://creativecommons.org/licenses/by-sa/2.0/)
- [https://www.flickr.com/photos/sklathill/2255718951/](https://www.flickr.com/photos/sklathill/2255718951/)
  - [https://creativecommons.org/licenses/by-sa/2.0/](https://creativecommons.org/licenses/by-sa/2.0/)