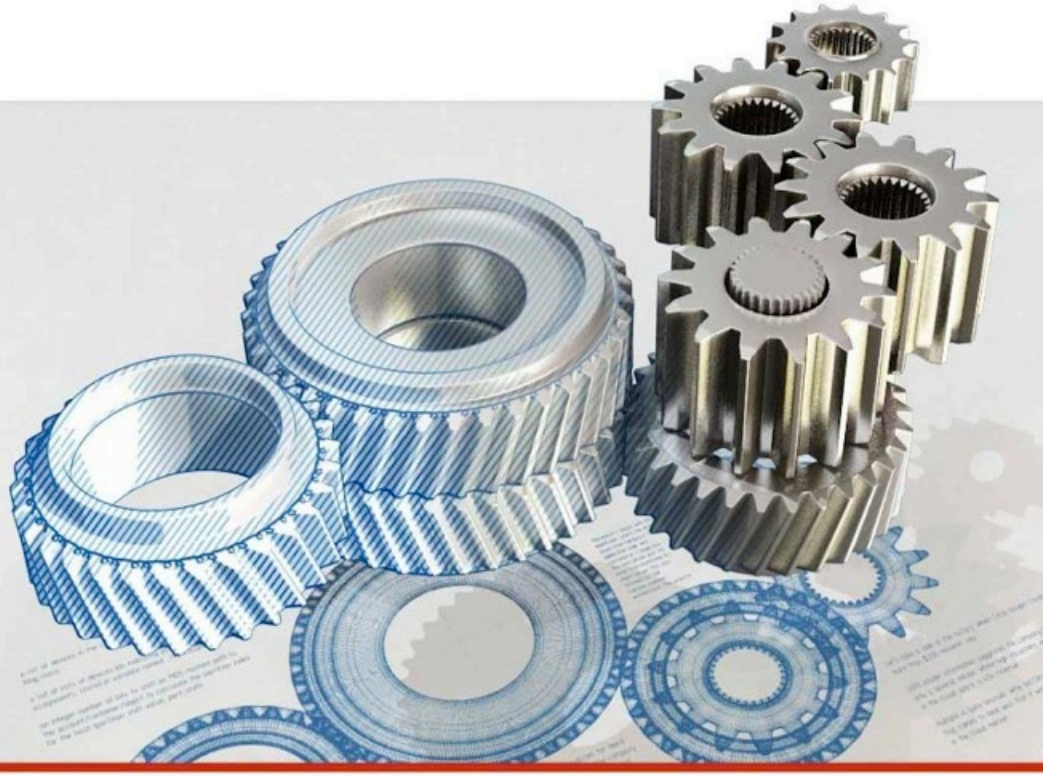
MIRANTIS

Building **Storage as a Service** with OpenStack



BUILT FOR openstack™ CLOUD SOFTWARE

Greg Elkinbard
Senior Technical Director

About the Presenter

BUILT FOR openstack™ CLOUD SOFTWARE

Greg Elkinbard

- Senior Technical Director at Mirantis
- Builds on demand IaaS and PaaS layers for public/private clouds
- Over 2 decades of experience in platform and cloud engineering
- Focus on breakthrough technology for cloud and HPC

MIRANTIS

Agenda

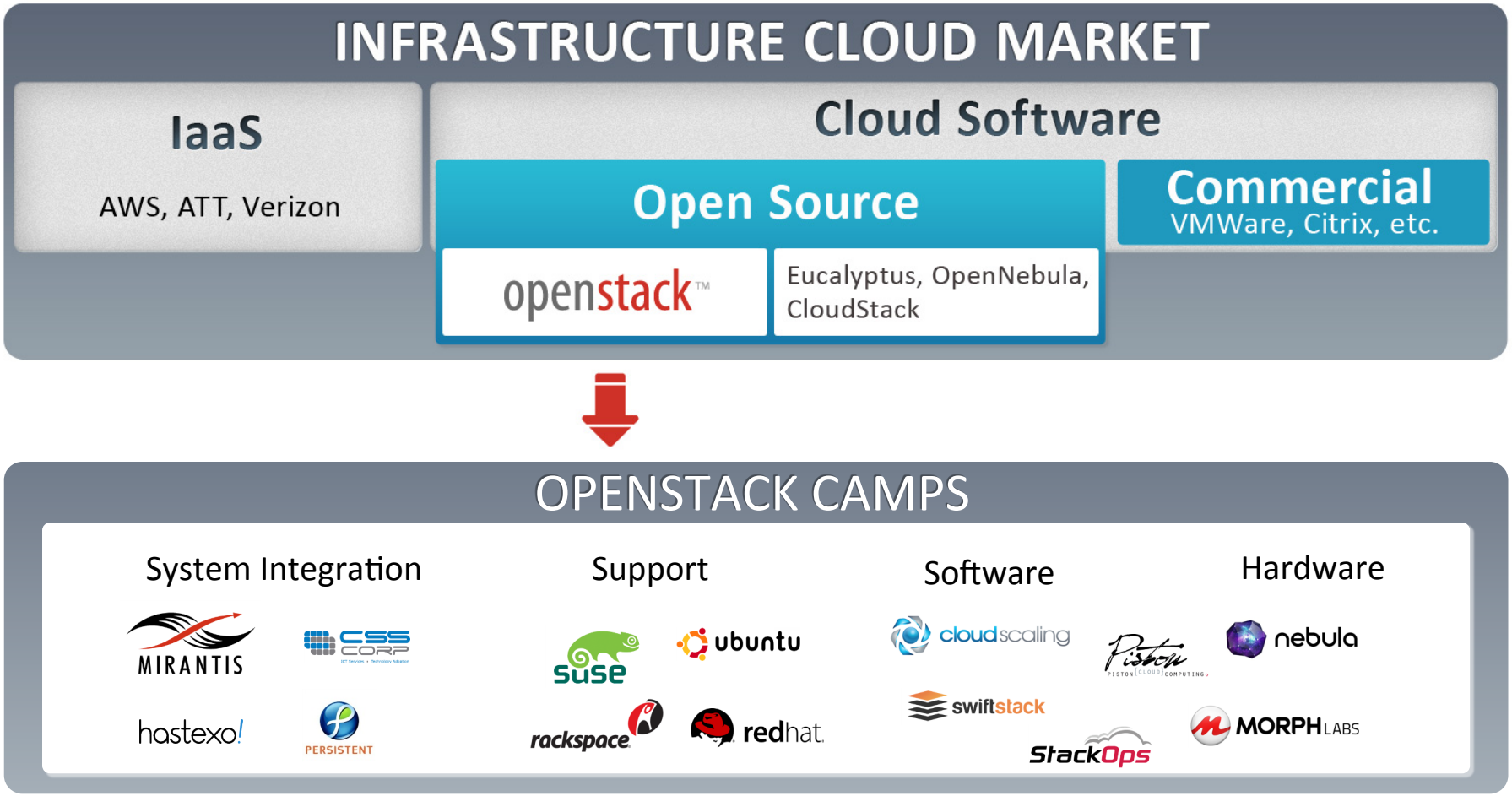
- OpenStack: quick background
- Understanding cloud storage capabilities
- Cloud & legacy storage platforms
- Key open source storage technologies
- Architecting for Opportunities Storage as a Service in OpenStack Cloud
- Q&A – and where to download today's slides

BACKGROUND

About OpenStack

MIRANTIS

Ecosystem


INFRASTRUCTURE CLOUD MARKET
IaaS
AWS, ATT, Verizon
Cloud Software
Open Source
openstack
Eucalyptus, OpenNebula, CloudStack
Commercial
VMWare, Citrix, etc.
OPENSTACK CAMPS
System Integration
MIRANTIS
CSS CORP
hastexo!
PERSISTENT
Support
suse
ubuntu
rackspace
redhat.
Software
cloudscaling
Piston
swiftstack
StackOps
Hardware
nebula
MORPH LABS

What is OpenStack?

- An open source community focused on building a software platform for building private and public cloud
- Core projects include:
  - Nova - compute
  - Swift - object store
  - Glance - image management
  - Keystone - auth
  - Horizon - dashboard
  - Quantum - network orchestration
  - Cinder - block storage
- Open sourced under Apache 2.0 license
- 6000 members in 87 countries

OpenStack capabilities

A brief overview of key problems OpenStack was designed to solve

VMs on demand

- provisioning
- snapshotting

Multi-tenancy

- quotas for different users
- user can be associated with multiple tenants

Volumes

Object storage for VM images and arbitrary files

MIRANTIS

CLOUD STORAGE

Understanding Capabilities

MIRANTIS

OpenStack Storage Architecture

- 3 storage tiers to fill different storage needs
- Ephemeral storage tier
- Persistent storage tier
- Object storage tier

OpenStack Ephemeral Storage

- Tied to lifecycle of individual VM
- Guest Boot Volume
- Typically all that is needed for small guests

- Types
  - Compute server local drives (common option)
    - *Low Latency*
    - *Low Cost*
    - *Longer Live Migration*
    - *No guest H/A*
  - Shared storage NAS or SAN
    - *Higher Cost, More features*
    - *Thin provisioning*
    - *Faster Migration*
    - *Standby H/A*
    - *Not managed by OpenStack*

MIRANTIS

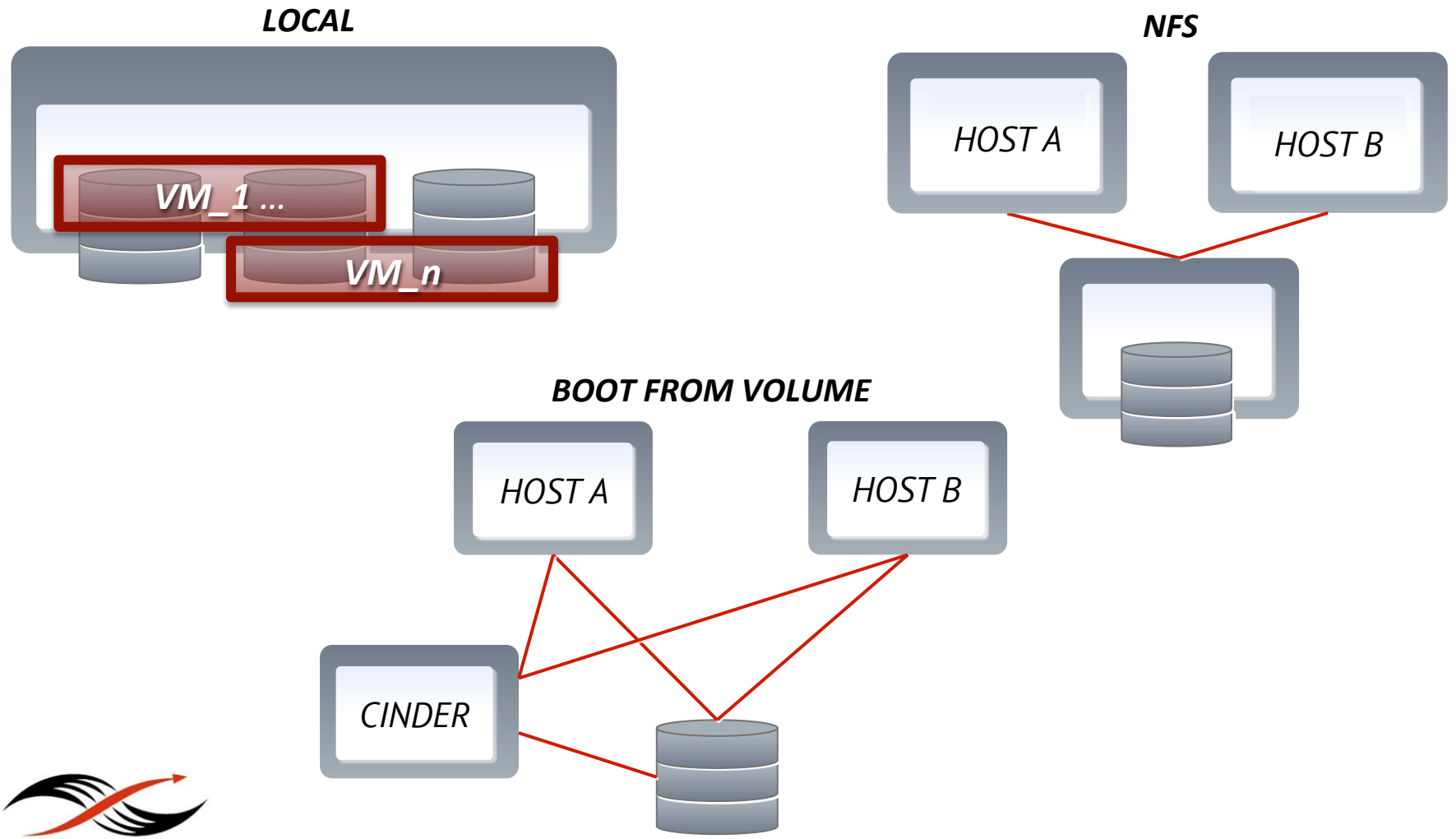
OpenStack Ephemeral Storage Folsom

- Optional Secondary Drive
  - Unformatted Volume
  - Removed when VM is deleted
- Boot from Volume (Cinder)
  - Experimental
  - Manual steps required
  - Better support should come in Grizzly*
  - Persistent or ephemeral lifecycle
  - Shared storage managed by OpenStack

*Grizzly is the next release of OpenStack, due out in Q2 2013

MIRANTIS

Architecting Ephemeral Storage



OpenStack Persistent Storage

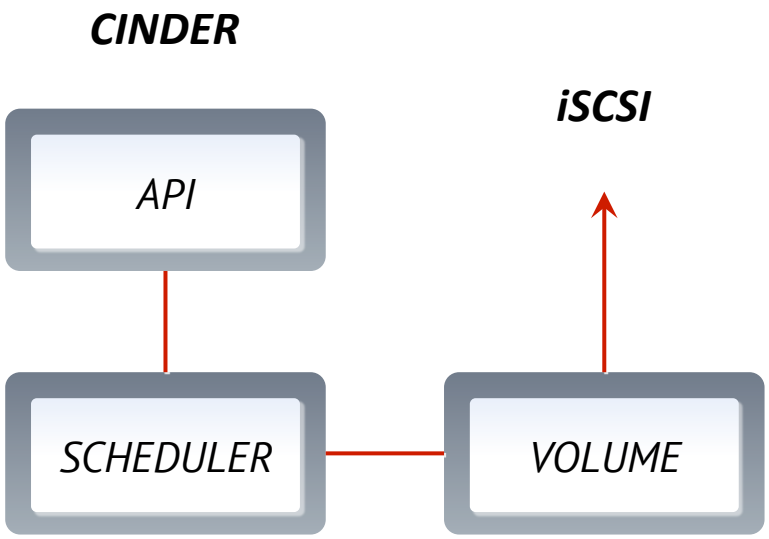
- Projects
  - Essex – Nova-Volumes
  - Folsom – Cinder
- Block Storage Service
  - Create iSCSI volumes on demand
  - Expandable: plug-in storage drivers
  - Snapshots and basic security ACLs
- Basic Service: Linux box with LVM and iSCSI target package
- Advanced Services to complement Cinder: Nexenta, NetApp, Ceph
  - Thin provisioning
  - Native Snapshot support
  - Built in H/A

OpenStack Persistent Storage: Grizzly

High Level Roadmap includes

- Metering
- Quotas
- Cloning
- Backup
- Volume Scheduling (QOS)

Architecting Persistent Storage



OpenStack Object Storage (project "Swift")

- RESTful APIs
- Simple but Scalable Service (6+ PB)
- Geo replication
- Architecture
  - DHT with anti-entropy services
  - Proxy Servers
  - Object Servers
- Swift matches most of the S3 functionality and has a S3 API shim as well as native API
- By comparison, AWS EC2's "S3" service also offers
  - Support for regions
  - Large Object Handling
  - Payment transfer bucket API

MIRANTIS

OpenStack Object Storage Hardware

- Basic Hardware: Linux boxes with XFS
  - Low cost hardware for object servers
  - More powerful servers for proxies.
  - Account/Containers need higher IOPS
    - *get a better io subsystem on proxies*
    - *co-locate account and container rings with proxies*
- Alternative Object Stores:
  - Nexenta's enhanced Swift
  - Ceph RadosGW
  - Gluster UFO (Unified File and Object)

STORAGE PLATFORMS

Comparing Cloud and Legacy

MIRANTIS

Key Storage Platform Types

- NAS and Scale Out NAS
- SAN
- Shared filesystems
- Object Storage

NAS

- Large NAS arrays or tightly coupled clusters
  - *Isilon, Ibrix, BlueArc, SONAS, OnTap GX*
  - *Present a global namespace to entire enterprise*
- Centralized and relatively expensive
- Cloud model takes a different approach
  - *Tenants do not need a lot of storage; no global namespace across tenants*
  - *Instead of large storage arrays, offer many smaller ones (give each tenant its own pool of storage)*
- You let the cloud platform manage the provisioning and control of resource allocation

SAN

- Dedicated Storage network
  - Fibre Channel or iSCSI
  - Tighly coupled collection storage switches, storage arrays and storage management software
- Centralized and relatively high-cost
- Cloud model is different
  - Guests do not need a lot of individual storage
  - Instead of large arrays, give each guest its own storage (across a pool of smaller arrays)
- You let the cloud platform manage the provisioning and control of resource allocation

Shared Filesystems

- Often a complement to SAN & replacement for NAS
- Legacy: Veritas, CXFS, GPFS
  - Limited Scalability
  - Relatively high cost
- OpenSource alternatives
  - Lustre, Gluster, Ceph
  - Designed to large scale
  - Much lower costs
  - Higher aggregate performance

Object Storage

- Non-POSIX
- HTTP based APIs (e.g., REST)
- Relaxed consistency
- EMC Atmos was one of the early systems
  - Now all vendors provide a REST interface to legacy storage arrays
- Openstack Swift
  - Higher scalability
  - Lower cost
  - Simplified customization

OPEN SOURCE OPTIONS:

Key Storage Technologies

MIRANTIS

Key Open Source Storage Technologies

- **Gluster**: Distributed network file system
- **Lustre**: Parallel large scale distributed file system
- **Ceph**: High-performance distributed object store
- **Swift**: OpenStack distributed cloud object store

MIRANTIS

Gluster

- Distributed Network File System
- Distributed Metadata Management
- Stackable translators provide different access methods
  - Native
  - NFS
  - Object
  - HDFS
- Good midsize/medium performance, replication based data protection
- Openstack integration into ephemeral storage tier and as a Swift replacement
- Good as guest shared file system for PaaS deployments

Lustre

- Distributed FS for HPC market
- Centralized Metadata management
- Optimized for extreme scalability and read/write performance
- Network equivalent of Raid-0
- Great speed but no network data protection
- Metadata performance is an issue
- Currently no Openstack integration

Ceph

- Distributed object store
- Different access methods
  - Distributed FS
  - Object Store
  - API
  - Block Store
- Good performance, scalability, reliability, sophisticated failure domain isolation
- Well integrated into OpenStack
  - Glance Image Store connector
  - Cinder connector
    - *Persistent storage tier*
    - *Ephemeral storage tier, using boot from volume*
  - Replacement for Swift

Swift

- Distributed Object Storage
- DHT based, lightweight distributed meta data
- Rest API only
- Designed for high scalability and reliability
- Mid level IO performance
- Default object storage for Openstack

MIRANTIS

Open Source Platforms & Use Cases

	Consider…				
If you need…	**Gluster**	**Lustre**	**Ceph**	**Cinder**	**Swift**
NAS and Scale Out NAS	✔	✔	✔		
SAN			✔	✔ *consider with Ceph Plugin or other storage systems*	
Shared Filesystems	✔	✔	✔		
Object Storage	✔		✔		✔

STORAGE AS A SERVICE

Opportunities and Architecture
for OpenStack Cloud

MIRANTIS

Concept: Storage as a Service

Multi-tenant service which

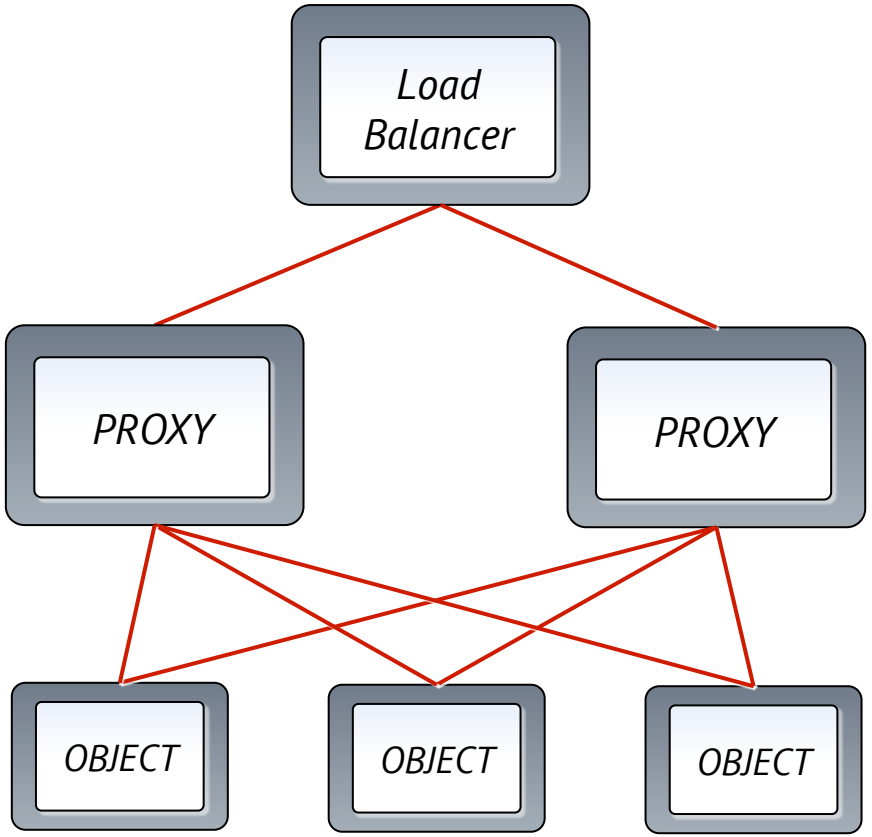
1. Presents storage via a well-defined set of remotely accessible APIs
2. Abstracts the actual storage implementation
3. Can be a foundation for more specialized functions: DR, backup, document sharing, etc.

Using Storage as a Service

- Access Methods
  - Object
  - Block
  - File
- OpenStack a good platform
  - Powerful set of components
  - Enhancements within easy reach

Object Services

This is what Swift was built to do



Access Method: File Services

- Cloud Drive
- Multi Tenant NAS

Use Case: using Swift as a cloud drive

- Swift: backend store
- Commercial or Open Source front end clients
  - Commercial: Gladinet, Cloudberry, ExpanDrive, Webdrive
  - Open Source: Cyberduck, Sparkleshare, Syncany
- Removable Drive or “Windows Explorer-like” UI
- Use Cases
  - Backup to the Cloud
  - Share Pictures, Music, Documents

Use Case: Multi tenant NAS

- Physical hosts or virtual guests
- OpenStack can be extended to manage NAS storage pools; creates or partitions storage, on demand
  - Virtual storage array
    - *Quantum creates network segregation*
    - *Nova-compute provides a pool for Virtual NAS heads*
    - *Local drives or Cinder provide back end storage*
  - Physical storage array
    - *Quantum creates network segregation*
    - *Cinder partitions and manages physical storage arrays*
    - *Secure separation requires internal virtualization support in storage array, as in NetApp Virtual Filer*

Multi Tenant iSCSI SAN

- Cinder already acts as a SAN for the virtual guests
- Can be expanded to SAN services to bare metal nodes (where there is no hypervisor)
  - Use Quantum to partition the network
  - Use OpenStack Cinder to provision the storage and set ACLs
- Physical hosts require broader security than guests
  - Use VSA (Virtual Storage Array) with storage provided by local drives or Cinder

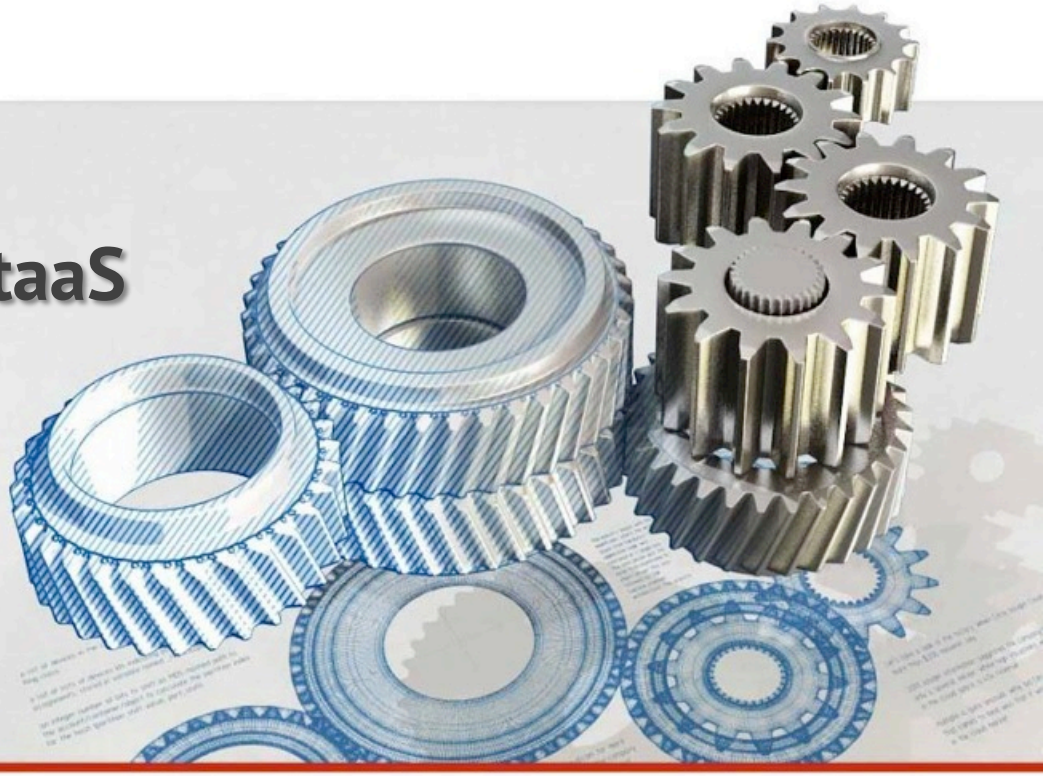
Takeaways

- OpenStack and its 3 storage tiers
  - Ephemeral, Persistent and Object Stores
- Open Source storage projects solve key storage needs at a lower cost
- Storage as a Service can be implemented with OpenStack

MIRANTIS

Download slides at
http://bit.ly/mirantis-StaaS

Q&A



BUILT FOR openstack™ CLOUD SOFTWARE