



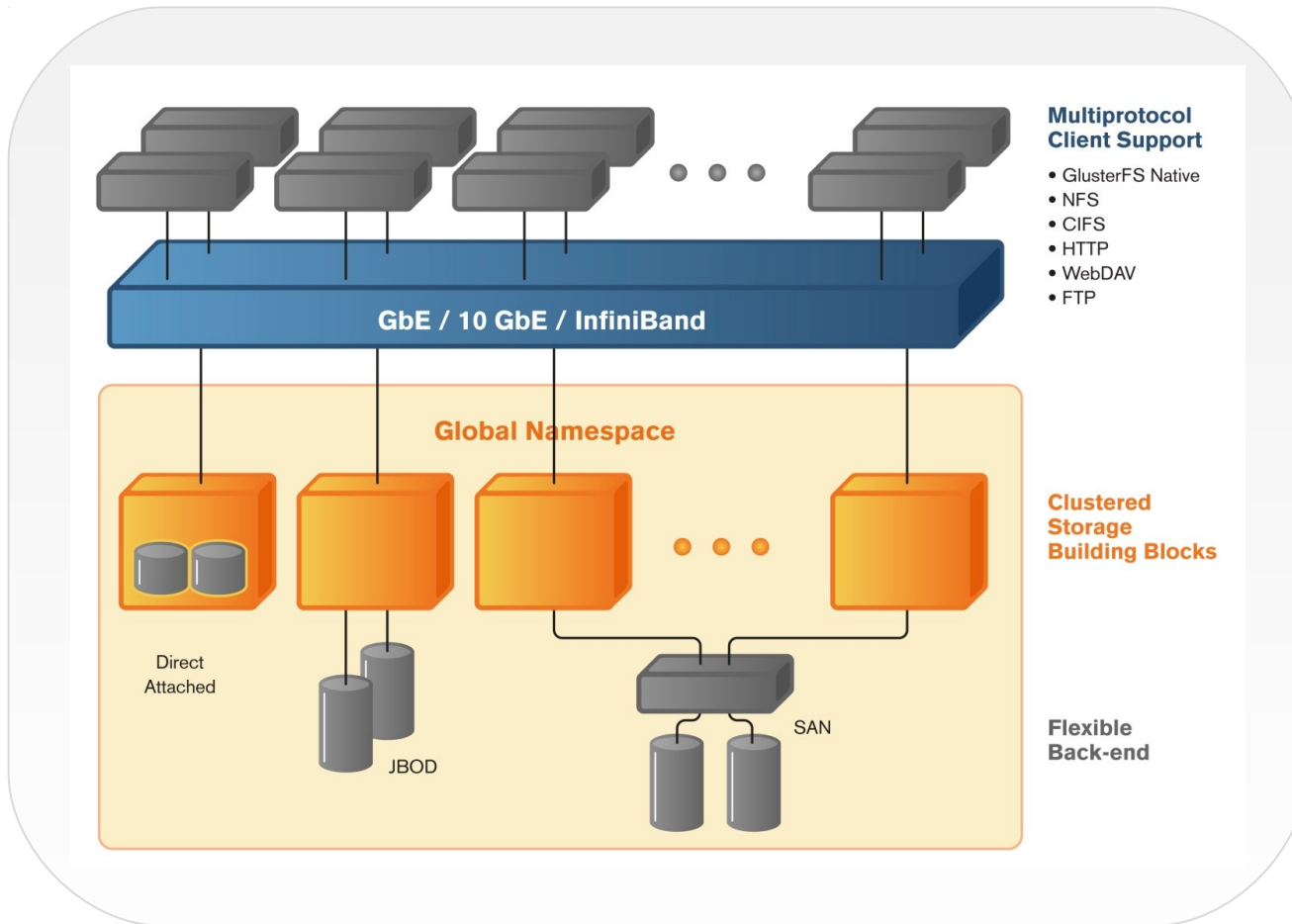
Integrating GlusterFS, qemu and oVirt

Vijay Bellur
Gluster Workshop – LinuxCon EU 2013

Agenda

- **Introduction**
- **GlusterFS as VM Image Store**
- **Integration with QEMU**
- **Integration with oVirt**
- **Q&A**

GlusterFS



- ◆ Global namespace
- ◆ Scale-out clustered storage building blocks
- ◆ Support thousands of clients
- ◆ Access using GlusterFS and NFS protocols
- ◆ Linear performance scaling

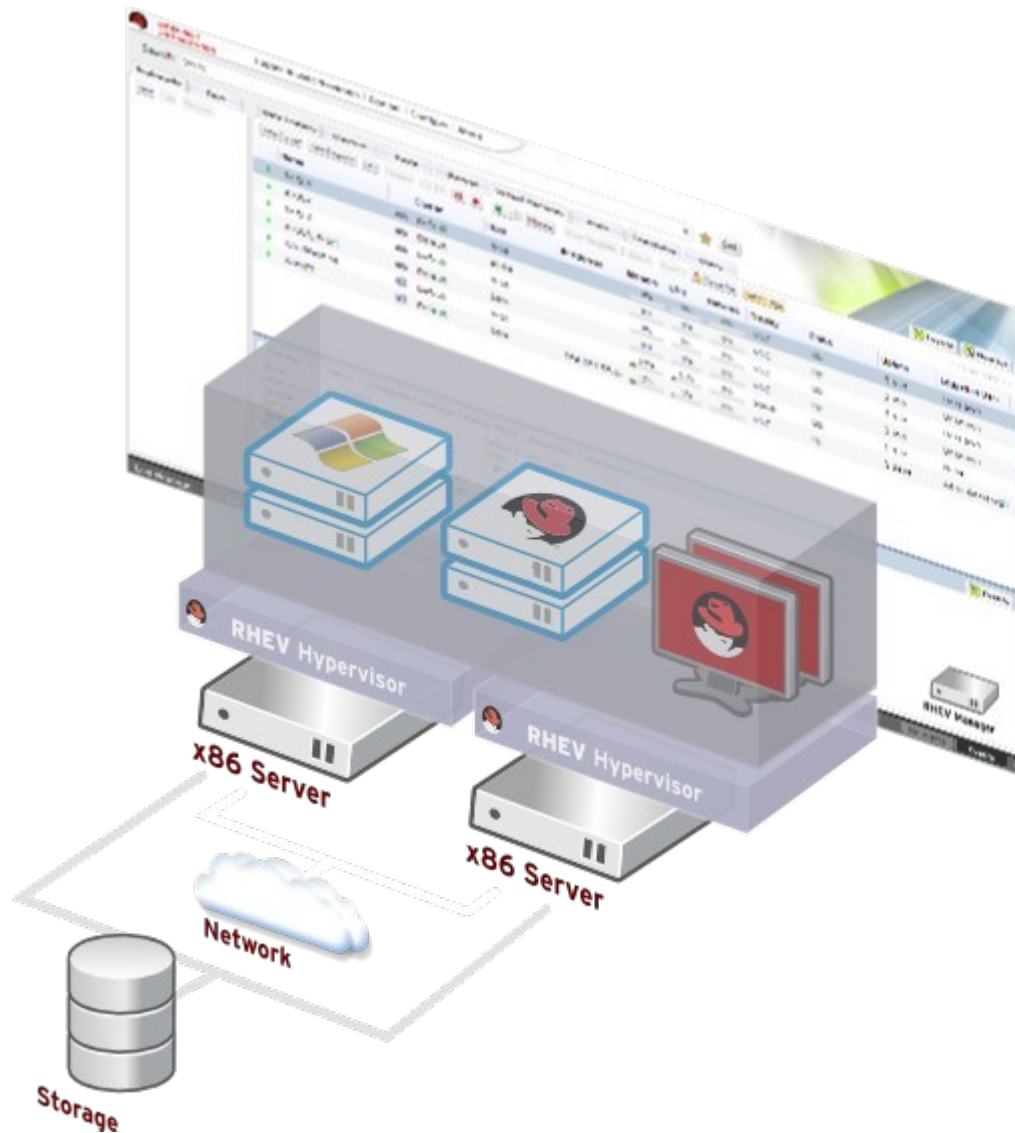
KVM and Qemu



KVM is a full virtualization solution for Linux on x86 hardware containing virtualization extensions.

Qemu is the driver for KVM virtualization

oVirt



- Large scale, centralized management for server and desktop virtualization
- Open source alternative to vCenter / vSphere
- Focus on KVM

GlusterFS 3.3

- **Granular data self heal**
- **Pro-active self healing**
- **Online data migration with remove-brick**
- **Support for O_DIRECT in FUSE**

GlusterFS 3.4

- **Improved write-behind behavior**
- **libgfapi with support for DVM**
- **Block Device Translator**
- **Support for volume/cluster management from oVirt**
- **FUSE**
 - **Scatter – Gather patchset**
 - **gid caching**
 - **Configurable fuse queue length**
 - **aio support in posix translator**

GlusterFS 3.5

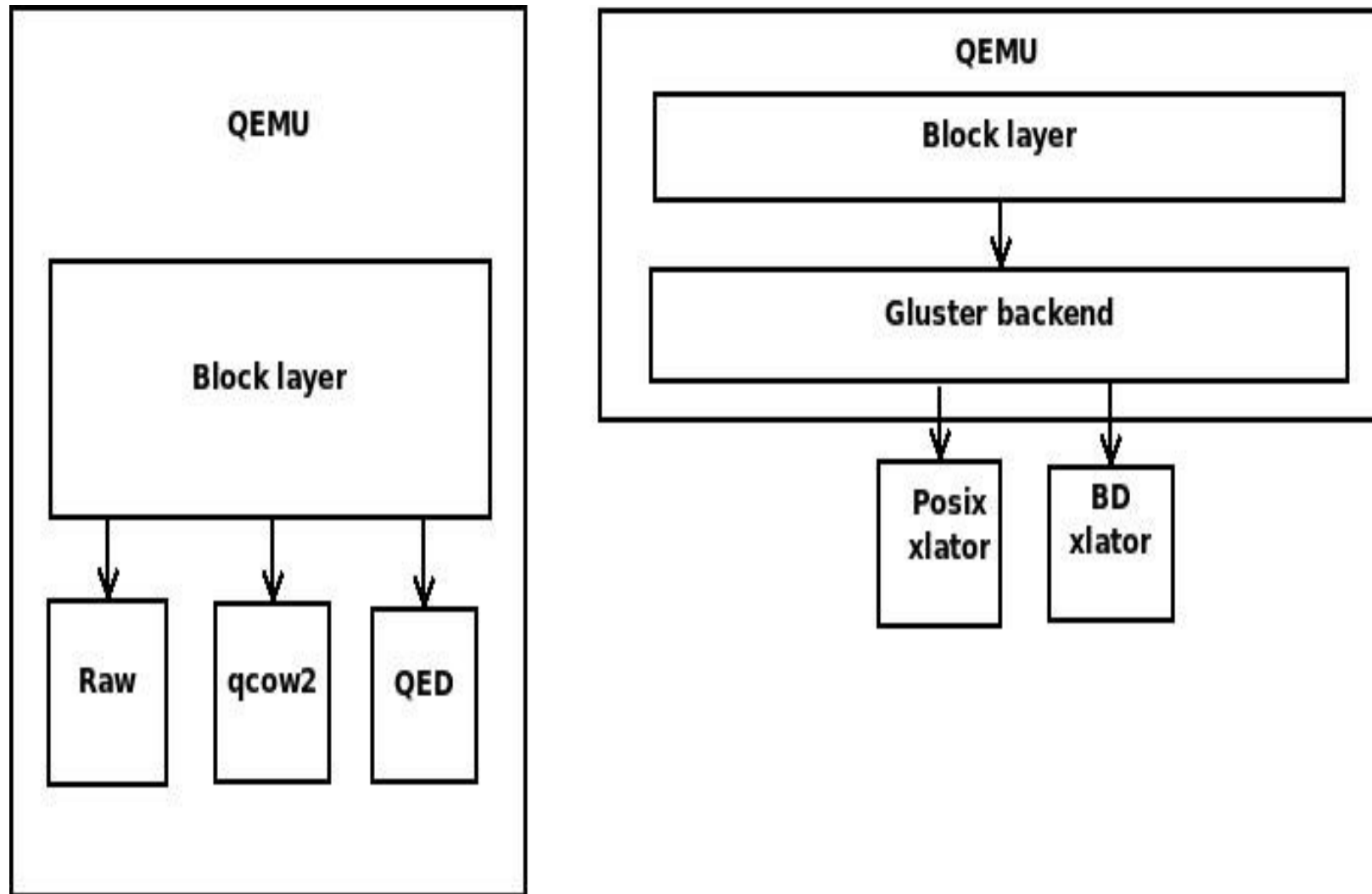
- **fallocate() support**
- **zerofill support in libgfapi**
- **Improved block device translator**
- **File snapshots**

QEMU & GlusterFS Integration

Integration possible through:

- GlusterFS native client - possible with qemu 1.2 and GlusterFS 3.3
- GlusterFS NFS – possible with qemu 1.2 and GlusterFS 3.3
- libgfapi – possible with qemu 1.3 and GlusterFS 3.4

QEMU & GlusterFS Integration

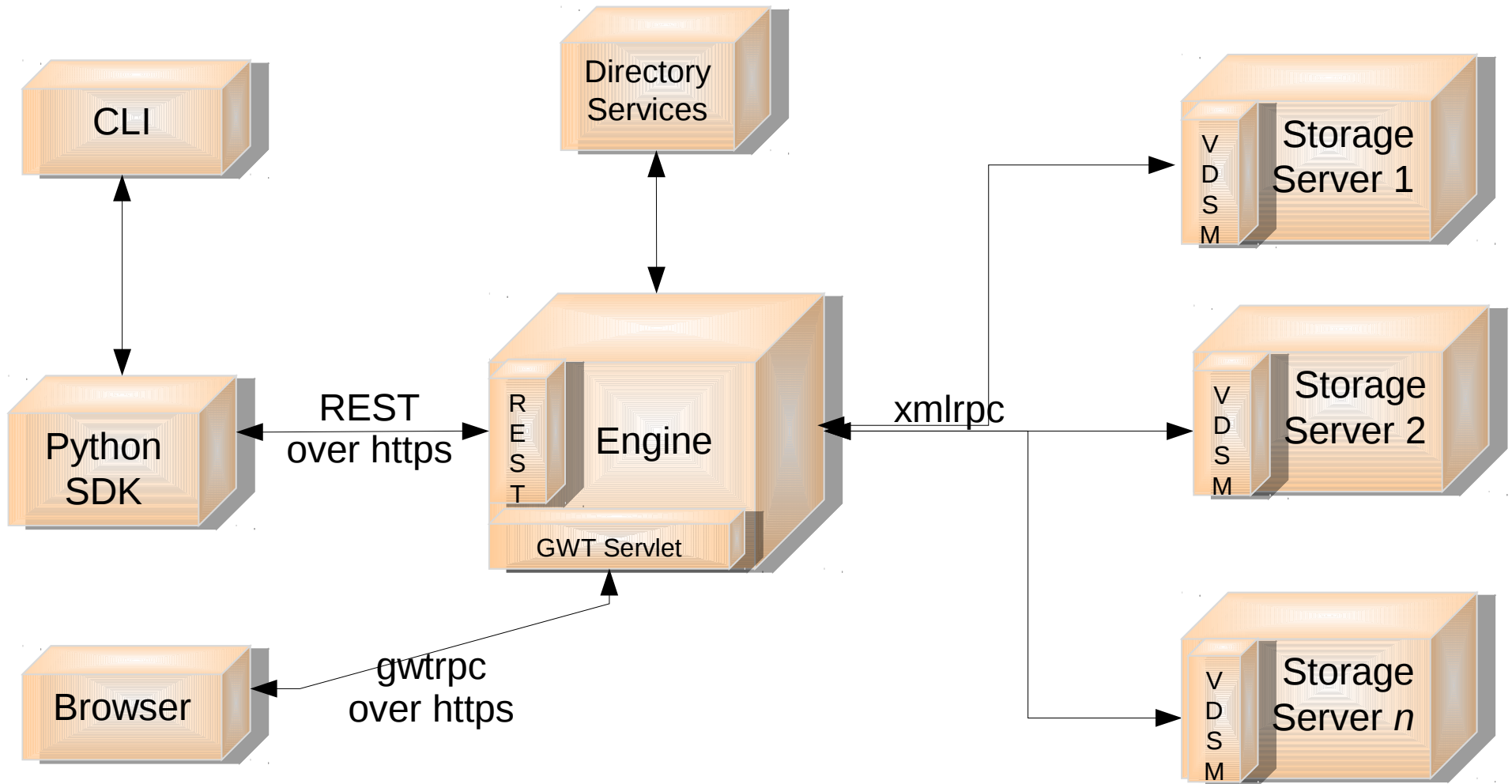


QEMU & GlusterFS Integration

➤ FIO Numbers (Seq read, 4 files with direct io, qemu options: if=virtio, cache=none)

	Aggregate BW(kB/s)	Min BW(kB/s)	Max BW(kB/s)
Base	63076	15769	17488
FUSE mount	29392	7348	9266
QEMU-GlusterFS native integration	53609	13402	14909

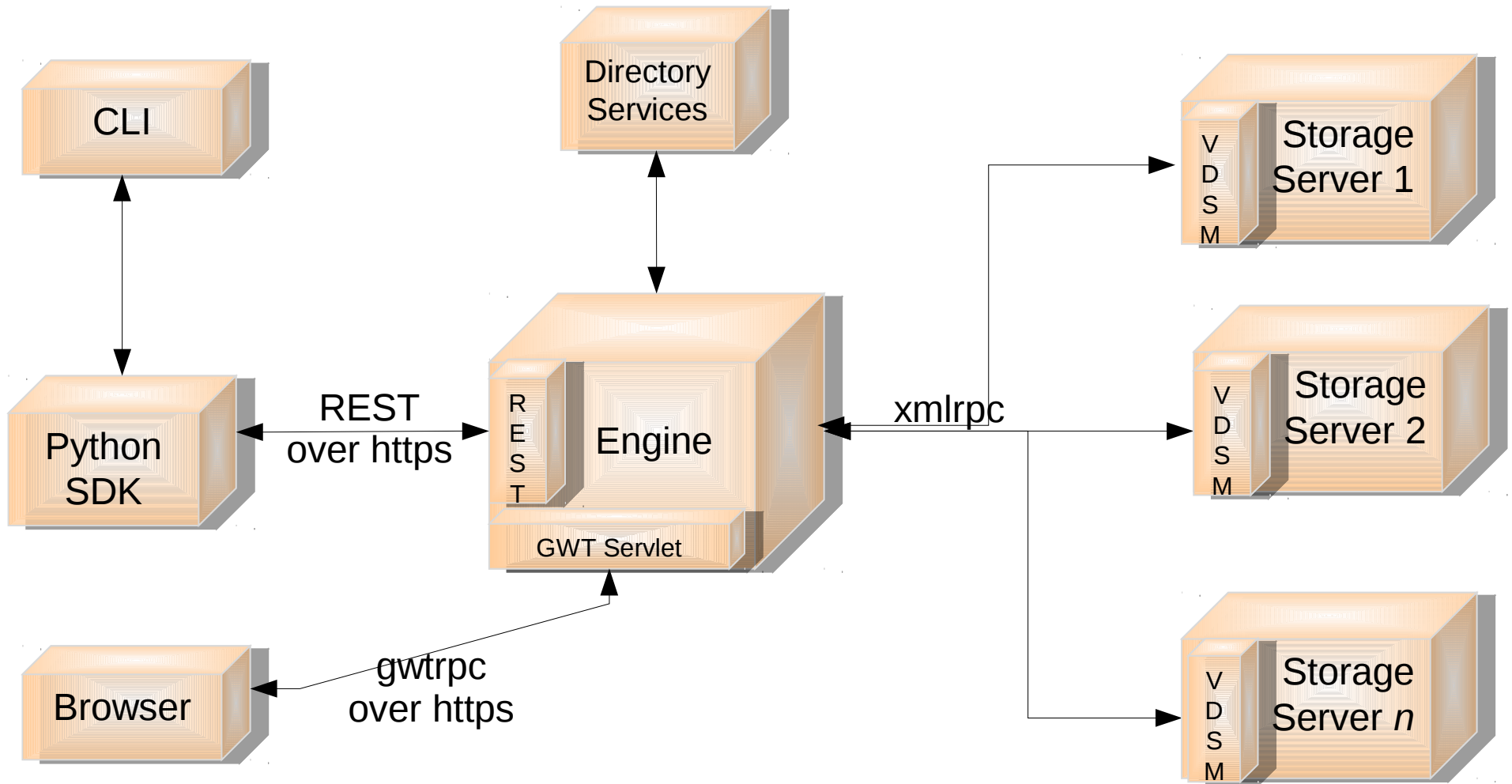
oVirt & GlusterFS - Architecture



oVirt & GlusterFS

- New feature in oVirt 3.1
http://wiki.ovirt.org/wiki/Features/Gluster_Support
- Cluster Configuration Modes
 - Virtualization only
 - Gluster only
 - Virtualization + Gluster (default)
- New entities (Volumes, Bricks, Volume Options)
- VDSM verbs for gluster management

oVirt & GlusterFS - Architecture



oVirt Gluster Management Features

- Cluster Management
 - Create Cluster
 - Add / Remove Storage Servers
 - Delete Cluster
- Volume Management
 - Create Volume
 - Add / Remove bricks
 - Start / Stop / Delete volume

oVirt Gluster Management Features

- Advanced Search
- Multi Level Administration & RBAC
- Audit & Alerts
- ReST API
- Python SDK & CLI

ReST API

Events

oVirt Engine Web Administration

127.0.1.1:8080/webadmin/webadmin/WebAdmin.html#events

oVirt Open Virtualization Manager

Logged in user: admin@internal | Configure | Guide | About | Sign Out

Search: Events

Clusters Hosts Volumes Users

Tree

- System
 - Clusters
 - Default
 - Hosts
 - Volumes
 - data
 - Hosts
 - server1
 - server2
 - Volumes
 - videos

Basic View Advanced View

Time	Message	Event ID	User	Host	Cluster	Gluster Volume	Correlation
2012-Oct-15, 23:07:27	Volume Option set on videos.	4002	admin@internal		data	videos	6825fa...
2012-Oct-15, 23:06:46	Gluster Volume music created.	4000	admin@internal		data	music	68b193...
2012-Oct-15, 23:06:46	Volume Option set on music.	4002	admin@internal		data	music	1547f3...
2012-Oct-15, 23:06:44	Volume Option set on music.	4002	admin@internal		data	music	1f3d71...
2012-Oct-15, 23:06:41	Volume Option set on music.	4002	admin@internal		data	music	625d72...
2012-Oct-15, 22:57:08	Gluster Volume videos started.	4004	admin@internal		data	videos	76cd11...
2012-Oct-15, 22:56:33	Gluster Volume videos created.	4000	admin@internal		data	videos	721c23...
2012-Oct-15, 22:56:33	Volume Option set on videos.	4002	admin@internal		data	videos	3c73a2...
2012-Oct-15, 22:56:31	Volume Option set on videos.	4002	admin@internal		data	videos	316769...
2012-Oct-15, 22:56:29	Volume Option set on videos.	4002	admin@internal		data	videos	31533c...
2012-Oct-15, 22:54:49	Available memory of host server2 [608 MB] is under defined threshold [10: 531			server2	data		
2012-Oct-15, 22:54:47	Warning, Low disk space. Host server2 has less than 1000 MB of free spa 23			server2	data		
2012-Oct-15, 22:54:22	Detected new Host server2. Host state was set to Up.	13		server2	data		
2012-Oct-15, 22:54:18	Host server2 was added by admin@internal.	42	admin@internal	server2	data		38a178...
2012-Oct-15, 22:53:14	Host server2 was removed by admin@internal.	44	admin@internal	server2	Default		5eb084...
2012-Oct-15, 22:53:00	Available memory of host server2 [608 MB] is under defined threshold [10: 531			server2	Default		
2012-Oct-15, 22:52:59	Warning, Low disk space. Host server2 has less than 1000 MB of free spa 23			server2	Default		

Bookmarks

Tags

Last Message: 2012-Oct-15, 23:07:27 Volume Option set on videos.

Alerts (0) Events Tasks (0)

Gluster as Virt Store

- Optimize volume for virt store (Sets optimization options on volume)

Name	Volume Type	Number of Bricks	Transport Types	Status
mv10	Distribute	1	TCP	Up

Create PosixFS storage domain (VFS type = glusterfs)

New Domain

Name:

Data Center:

Domain Function / Storage Type: Format:

Use Host:

Path:
Path to device to mount / remote export

VFS Type:

Mount Options:

Upcoming features (oVirt 3.4 and beyond)

- Async Tasks
- UFO configuration

Upcoming features (oVirt 3.2 and beyond)

- Storage Device management
- Reports
- Gluster native storage domain
- Volume Quotas

Contribute

- Git repository

 - `git://gerrit.ovirt.org/ovirt-engine`

 - `git://gerrit.ovirt.org/vdsm`

 - `git://git.gluster.com/glusterfs.git`

- IRC Channels

 - #ovirt on oftc

 - #gluster on freenode

Thank You!

Q & A