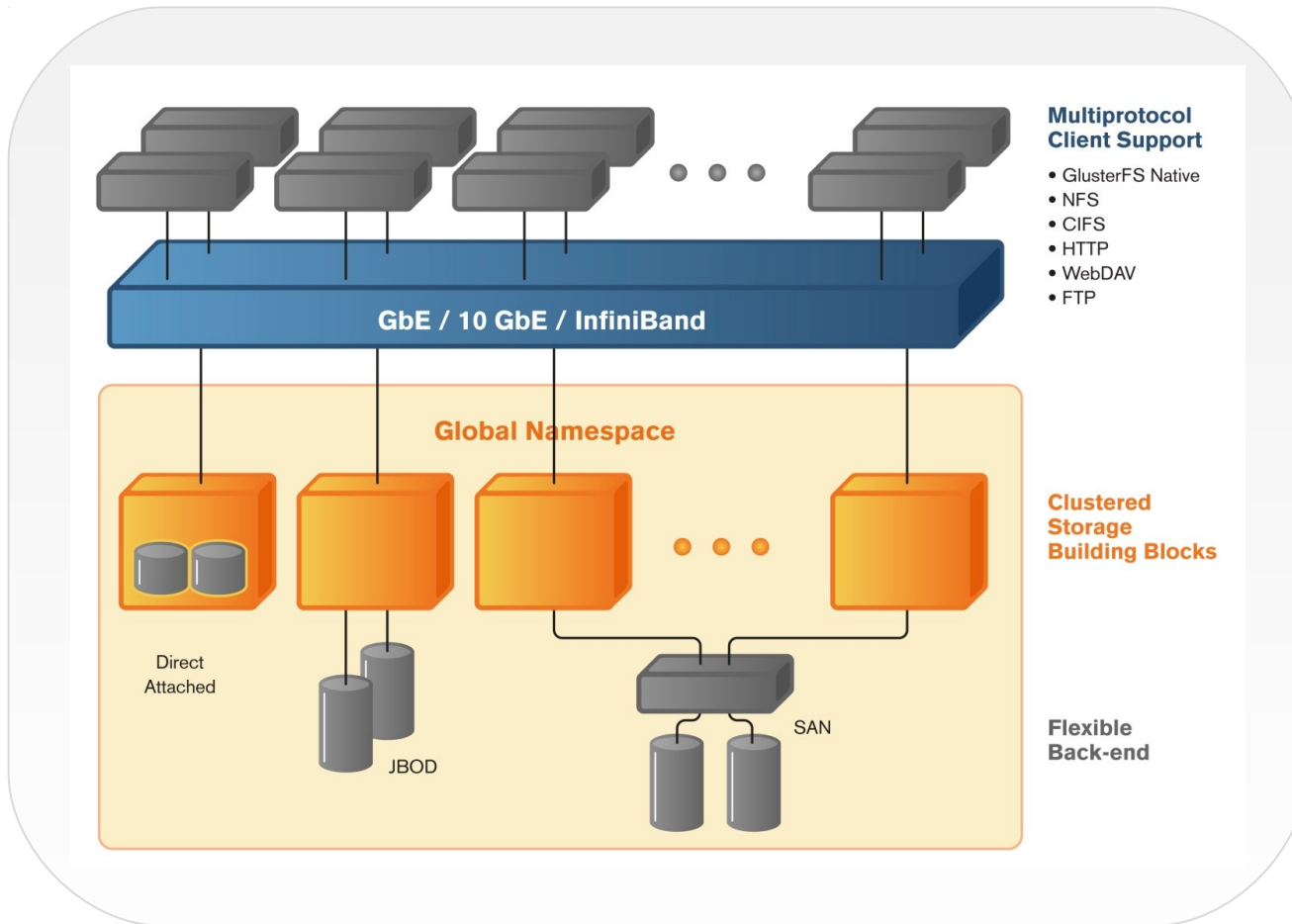# Integrating GlusterFS, qemu and oVirt

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**
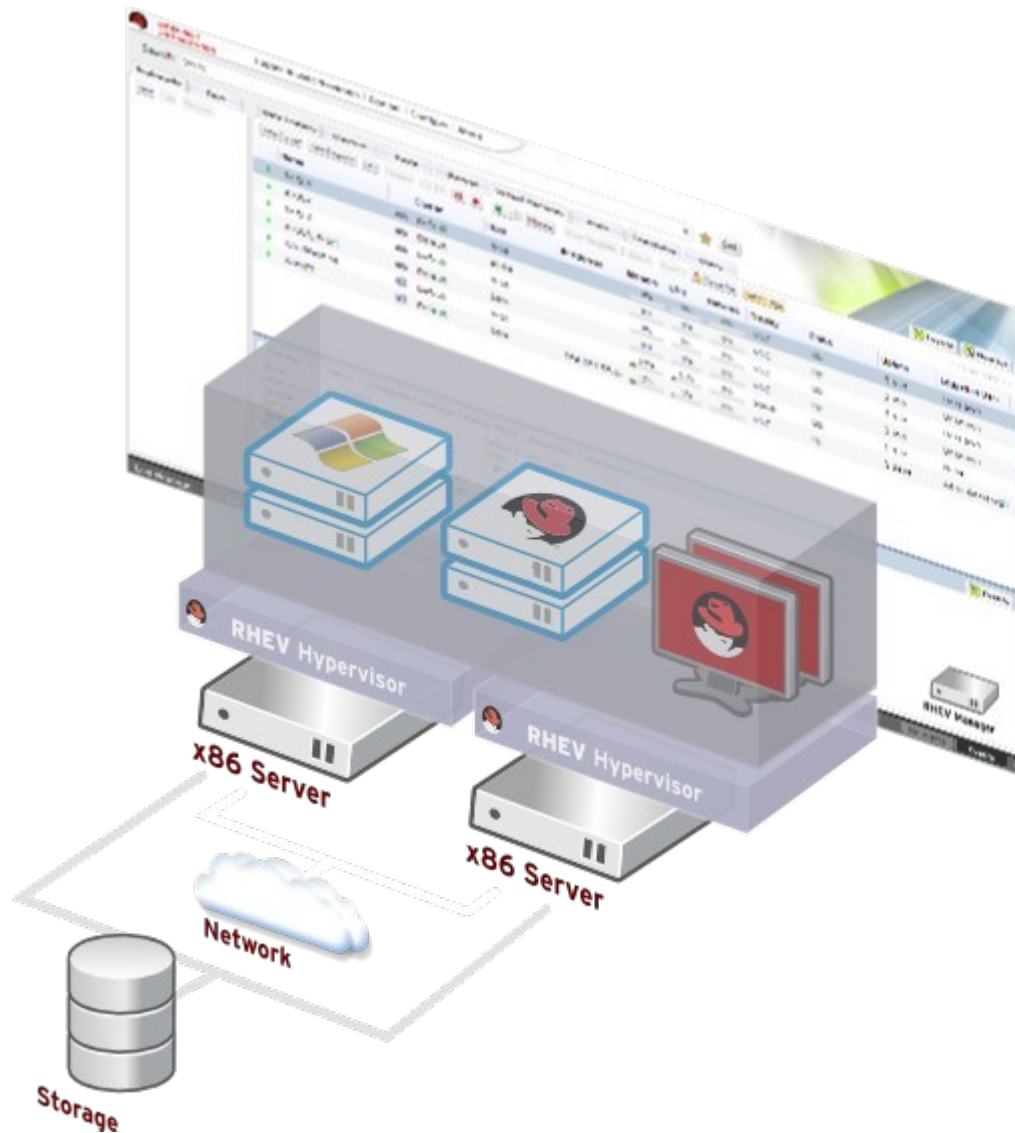- **Linear performance scaling**

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**
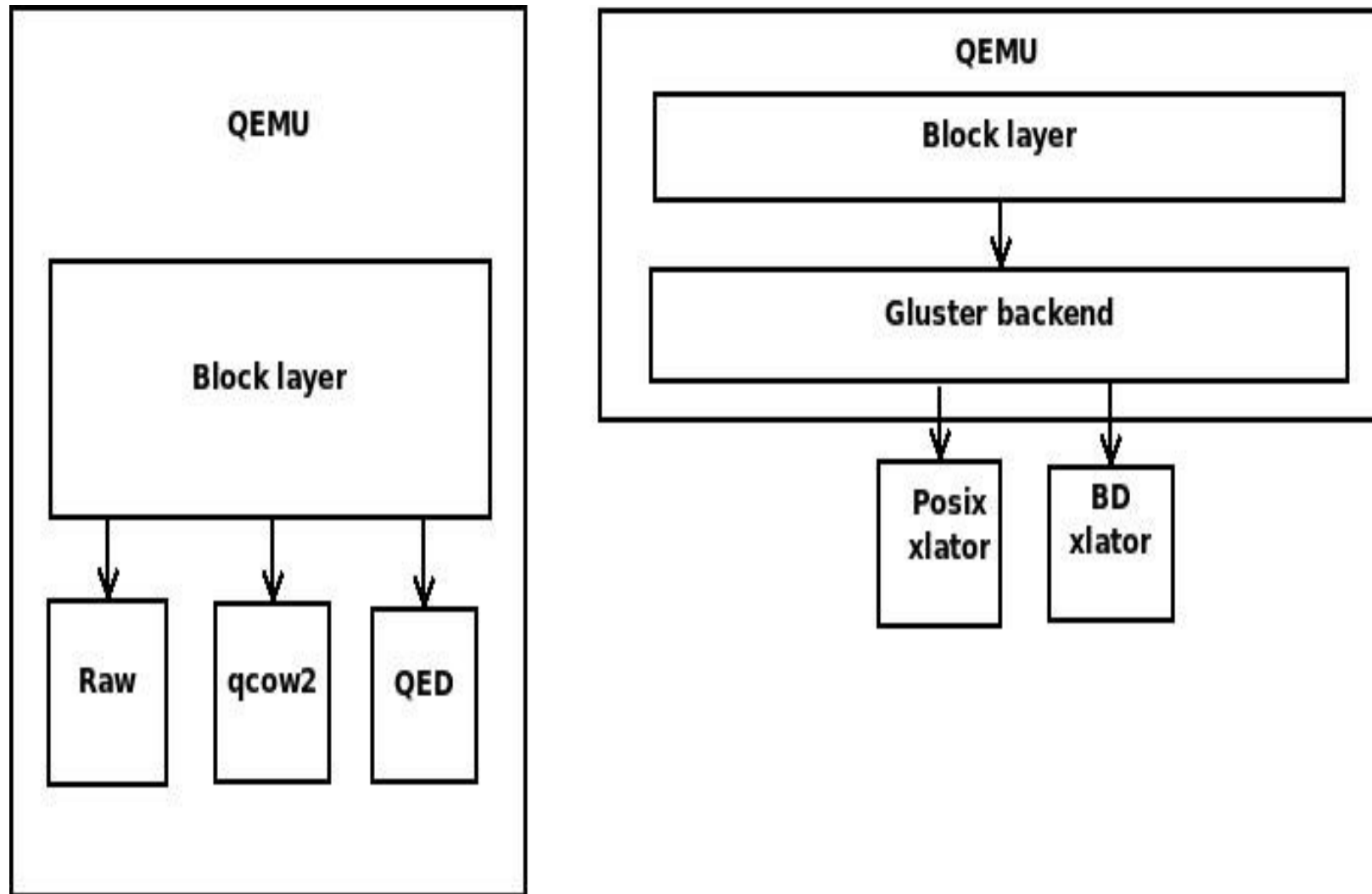
GlusterFS 3.5

➢ **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

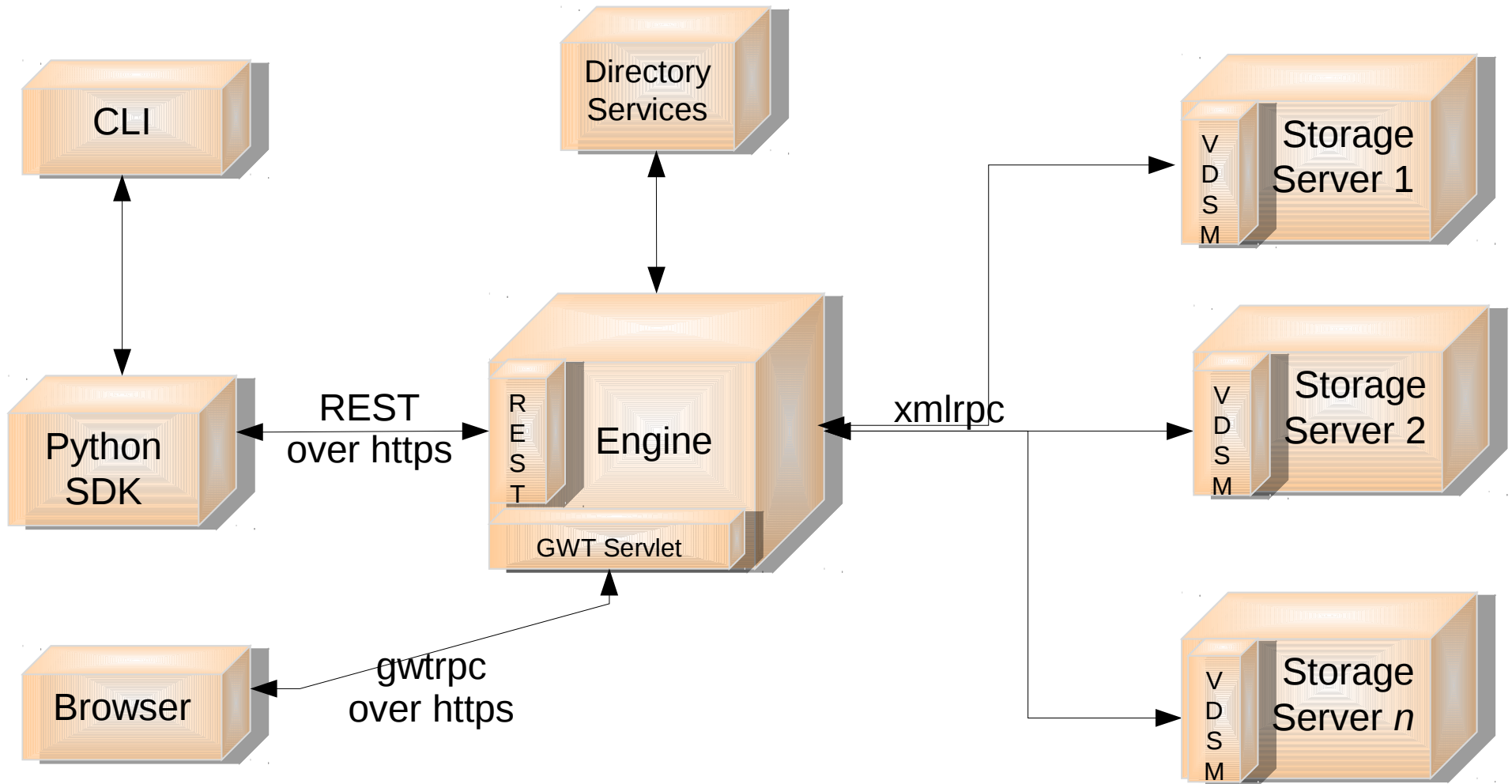➢ libgfapi – possible with qemu 1.3 and GlusterFS 3.4

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

CLI

Directory Services

Python SDK

REST over https

R E S T

Engine

GWT Servlet

xmlrpc

V D S M — Storage Server 1

V D S M — Storage Server 2

V D S M — Storage Server n

Browser

gwtrpc over https

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
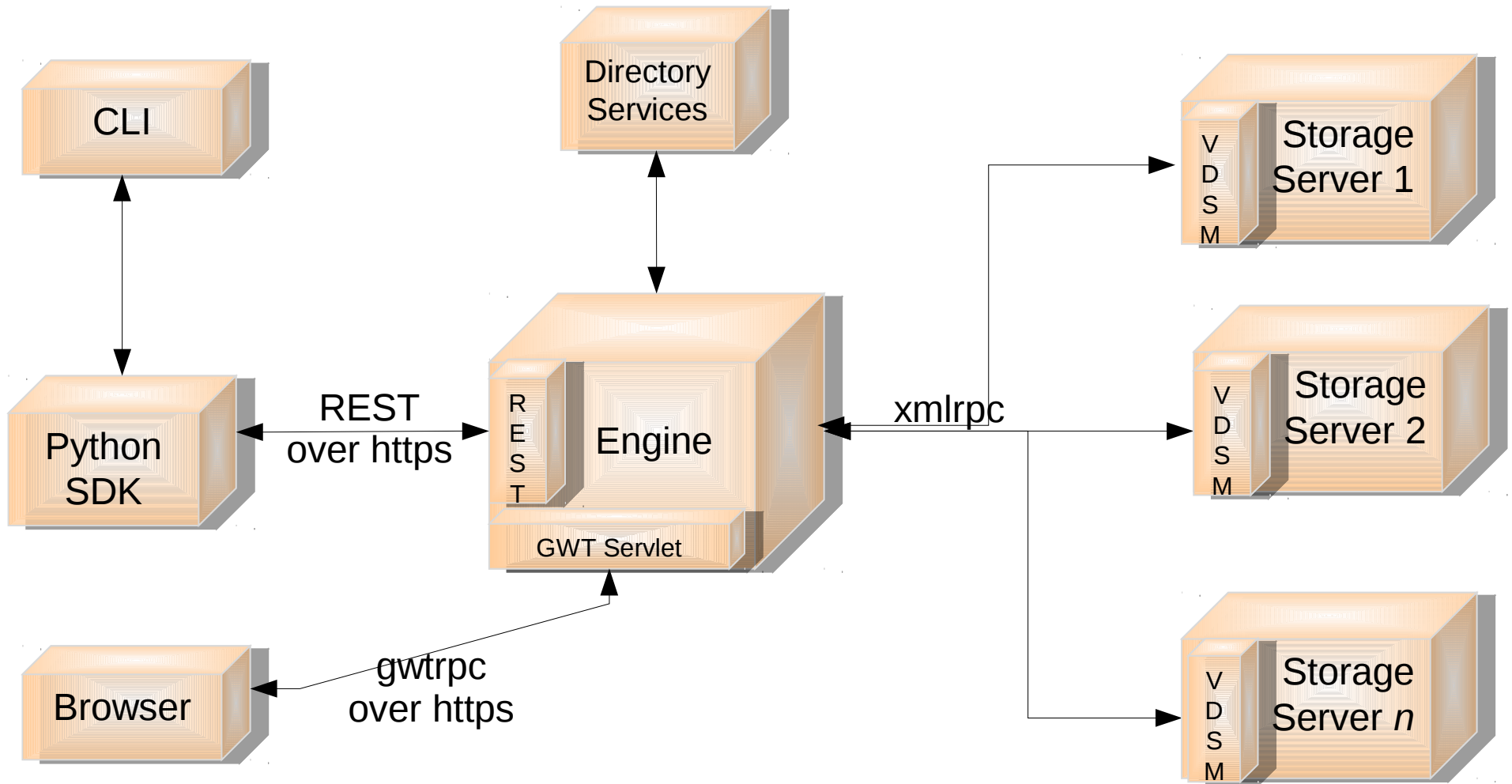
oVirt Gluster Management Features

- ➢ 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

The Future of Storage is *Open* for Business

Events

Gluster as Virt Store

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

Create PosixFS storage domain (VFS type = glusterfs)

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