

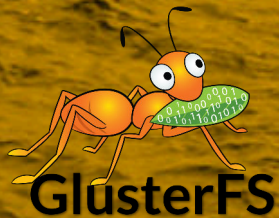


OPEN STORAGE IN THE ENTERPRISE

with GlusterFS and Ceph

Dustin L. Black, RHCA
Principal Technical Account Manager
Red Hat Strategic Customer Engagement

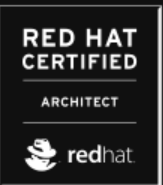
2014-12-11





Dustin L. Black, RHCA
Principal Technical Account Manager
Red Hat, Inc

dustin@redhat.com
@dustinblack



Wouldn't you like to have...

**a single named support contact
who know's your business,
your technology,
and your needs?**

A trusted advisor and technical expert
to analyze your configuration,
advise on your architecture,
and collaborate on your strategy?

An advocate and liaison
connecting you with engineers and
maintainers,
within **Red Hat** and upstream,
ensuring your priorities are also theirs?

**A partner who lives and
breathes open source and
transparency?**

RED HAT Technical Account Management

Premium named-resource proactive support
from your leading experts in open solutions

Contact your sales team or visit redhat.com

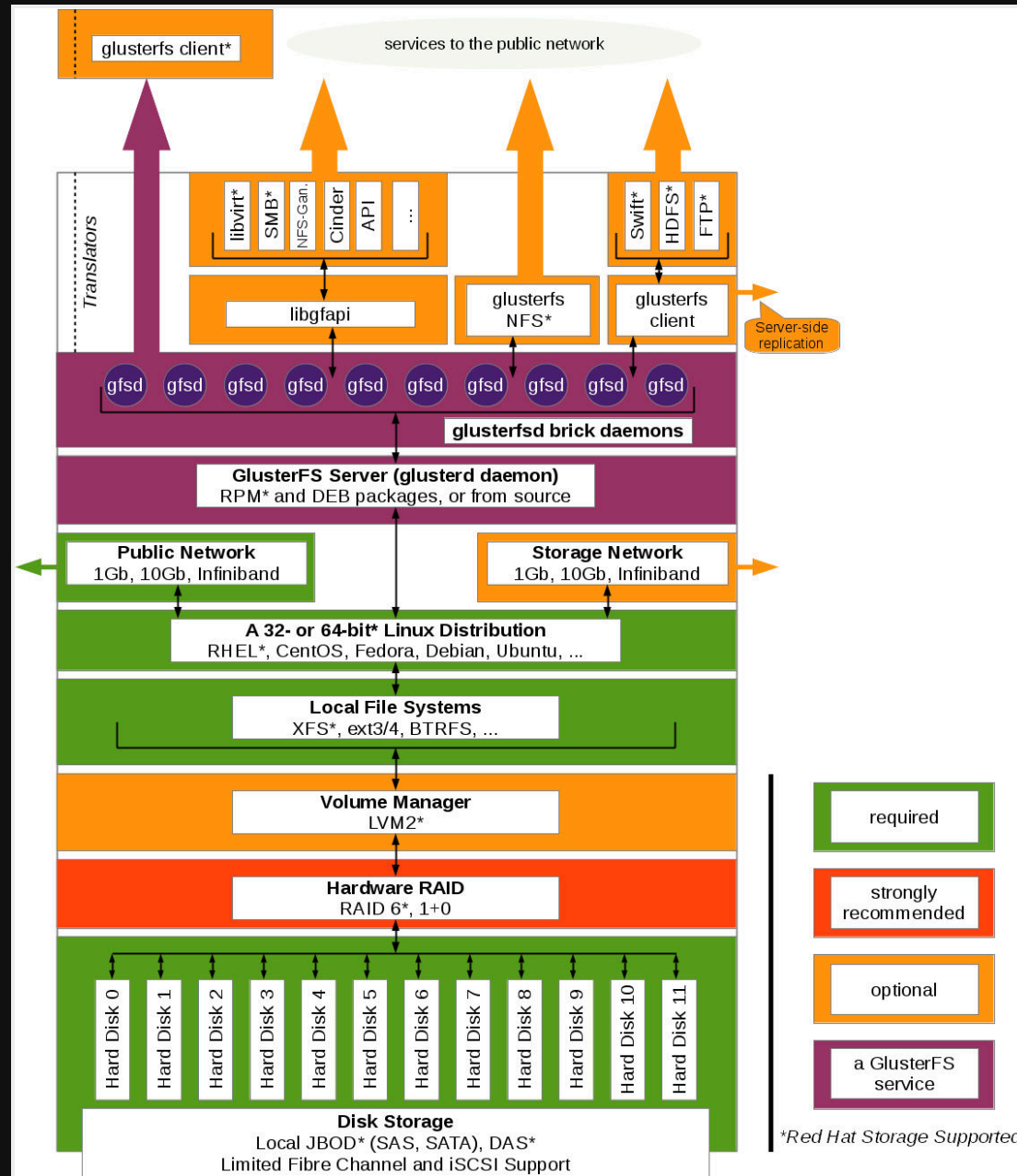
Supporting success. Exceeding expectations.

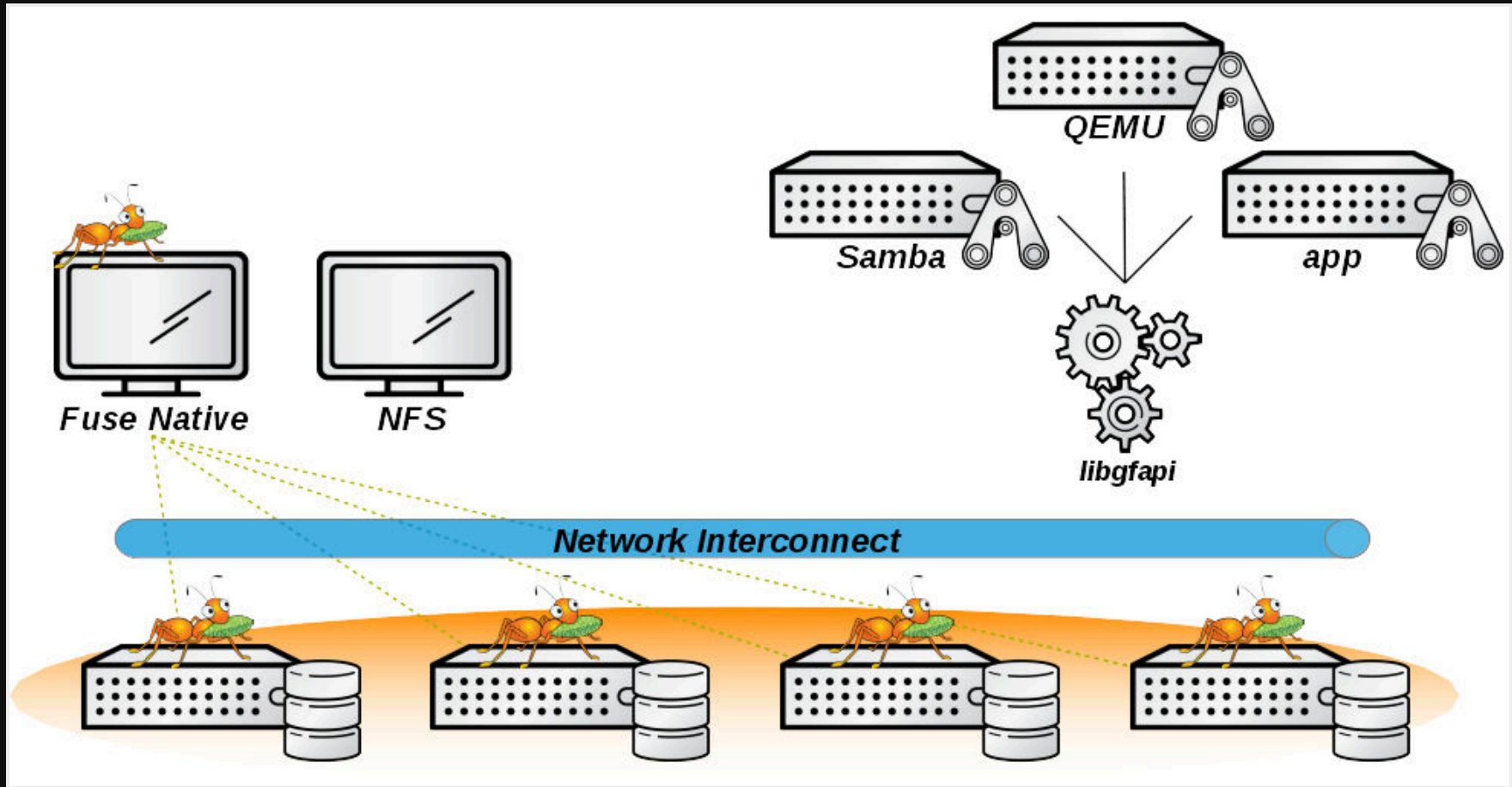
Let's Talk Distributed Storage

- | Decentralize and Limit Failure Points
- | Scale with Commodity Hardware and Familiar Operating Environments
- | Reduce Dependence on Specialized Technologies and Skills

GlusterFS

- | Clustered Scale-out General Purpose Storage Platform
- | Fundamentally File-Based & POSIX End-to-End
 - | Familiar Filesystems Underneath (EXT4, XFS, BTRFS)
 - | Familiar Client Access (NFS, Samba, Fuse)
- | No Metadata Server
- | Standards-Based – Clients, Applications, Networks
- | Modular Architecture for Scale and Functionality



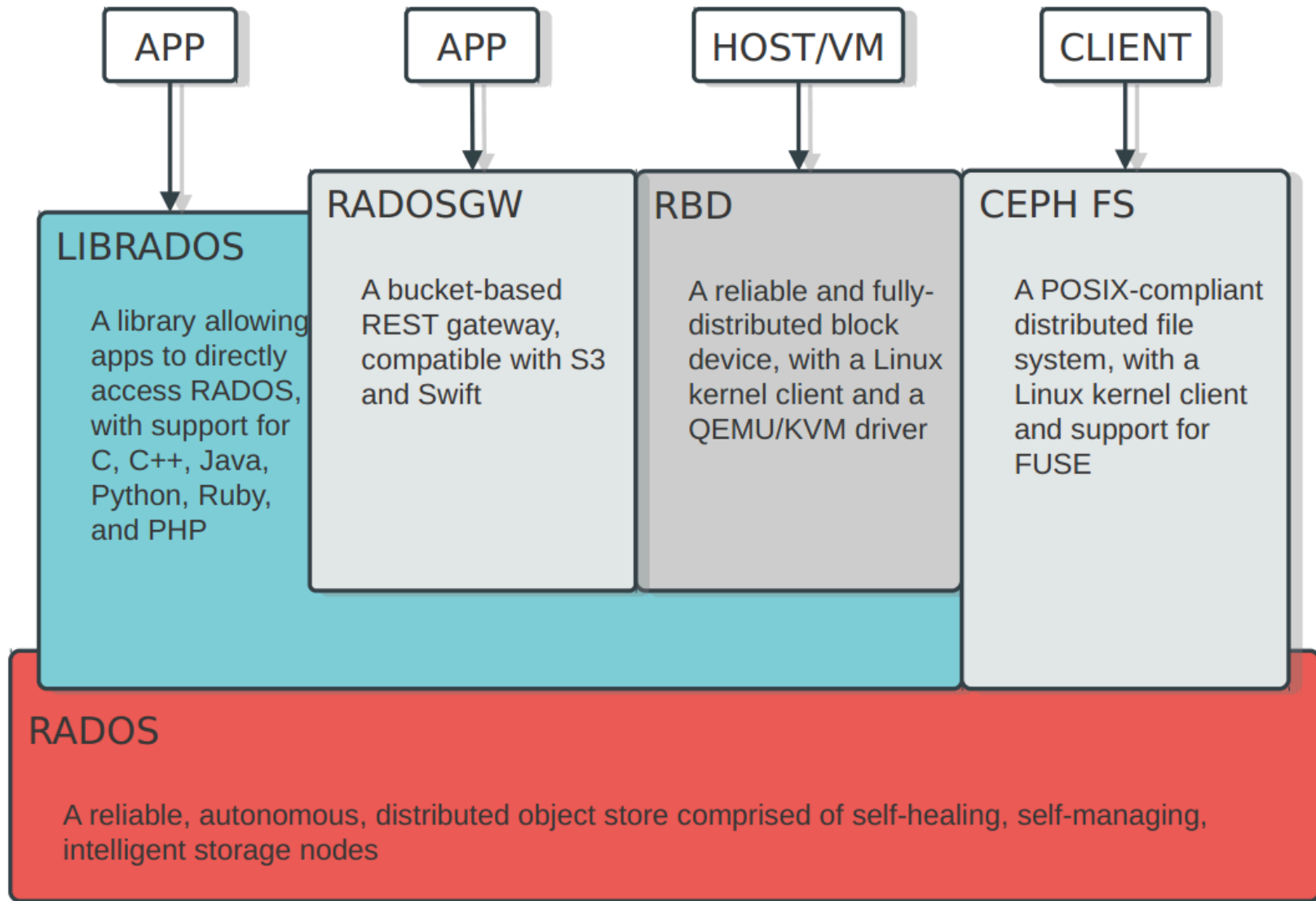


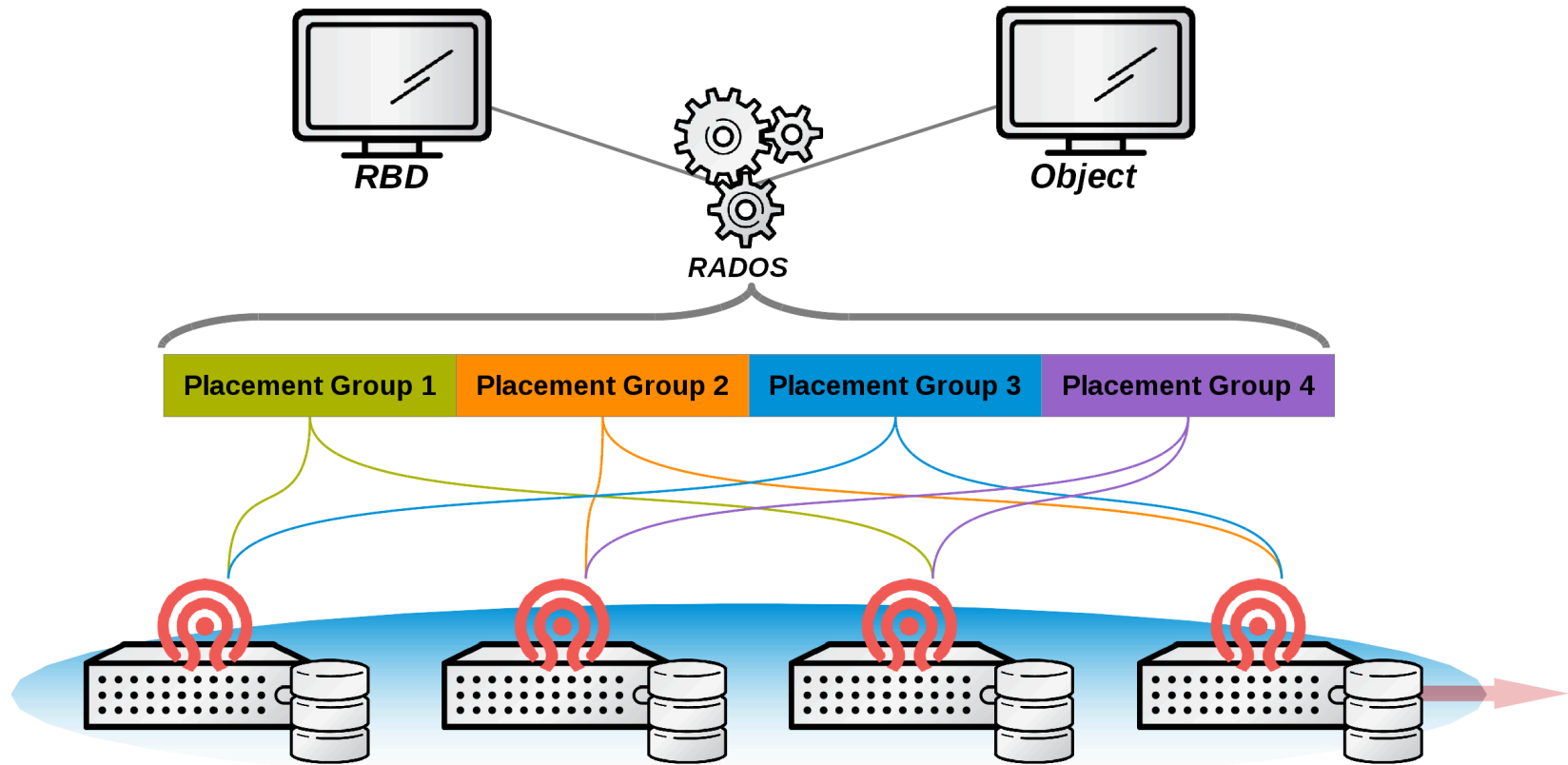
Red Hat Storage Server

- | Enterprise Implementation of GlusterFS
- | Integrated Software Appliance
- | RHEL + XFS + GlusterFS
- | Certified Hardware Compatibility
- | Subscription Model
- | 24x7 Premium Support

Ceph

- | **Massively scalable, software-defined storage system**
- | **Commodity hardware with no single point of failure**
- | **Self-healing and Self-managing**
 - | Rack and data center aware
 - | Automatic distribution of replicas,
- | **Block, Object, File**
 - | Data stored on common backend filesystems (EXT4, XFS, etc.)
 - | Fundamentally distributed as objects via RADOS
 - | Client access via RBD, RADOS Gateway, and Ceph Filesystem





Inktank Ceph Enterprise

- | Enterprise Implementation of Ceph
- | Combined with management and deployment tools
- | Enterprise-level support with bug escalation and hot patches
- | Bare metal and OpenStack deployments
- | Tested and predefined client and server configurations
- | Support for block, object, and API clients



Use Case:

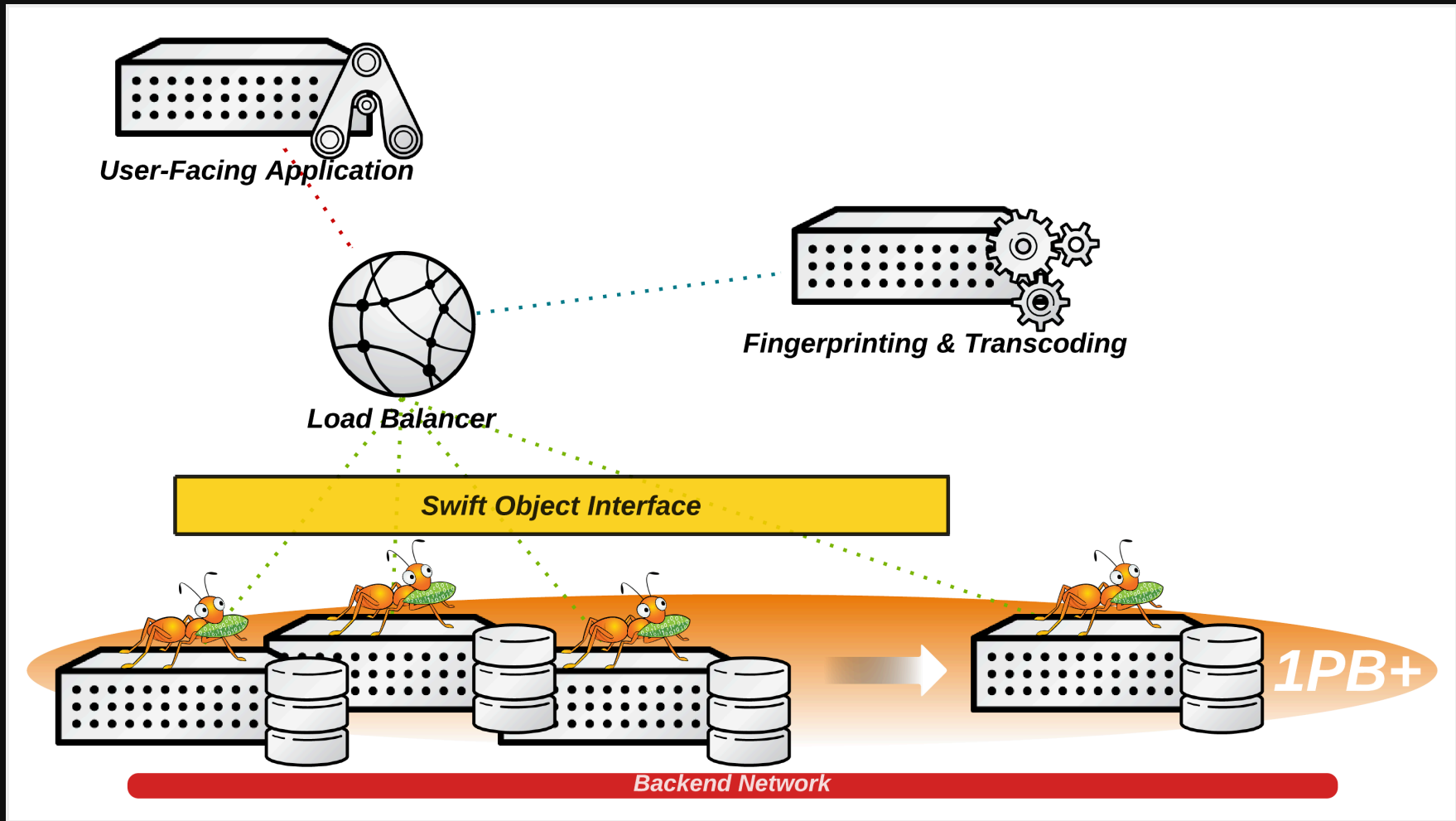
Media Storage via Object Interface

Goals

- | Media file storage for customer-facing app
- | Drop-in replacement for legacy object backend
- | 1PB plus 1TB/day growth rate
- | Minimal resistance to increasing scale
- | Multi-protocol capable for future services
- | Fast transactions for fingerprinting and transcoding

Implementation

- | *12 Dell R710 nodes + MD1000/1200 DAS*
 - | Growth of 6 -> 10 -> 12 nodes
- | *~1PB in total after RAID 6*
- | *GlusterFS Swift interface from OpenStack*
- | *Built-in file+object simultaneous access*
- | *Multi-GBit network with segregated backend*





Use Case:

Self-Service Provisioning with Accounting and Chargeback

Goals

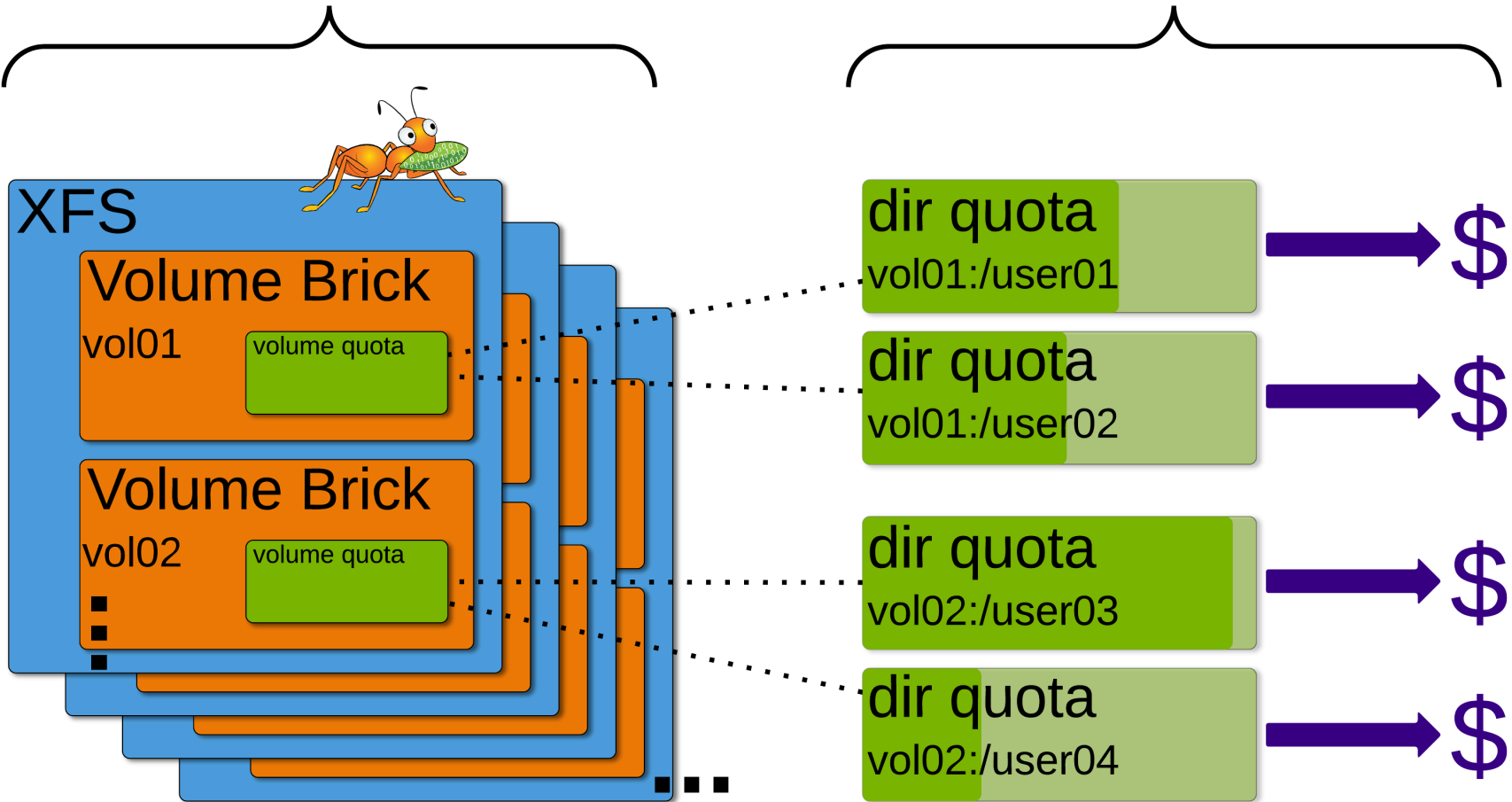
- | Add file storage provisioning to existing self-service virtualization environment
 - | Automate the administrative tasks
- | Multi-tenancy
 - | Subdivide and limit usage by corporate divisions and departments
 - | Allow for over-provisioning
 - | Create a charge-back model
- | Simple and transparent scaling

Implementation

- | Dell R510 nodes with local disk
- | ~30TB per node as one XFS filesystem
- | Bricks are subdirectories of the parent filesystem
 - | Volumes are therefore naturally over-provisioned
- | Quotas placed on volumes to limit usage and provide for accounting and charge-back
- | Only 4 gluster commands needed to allocate and limit a new volume; Easily automated

Administration

Automation





Use Case:

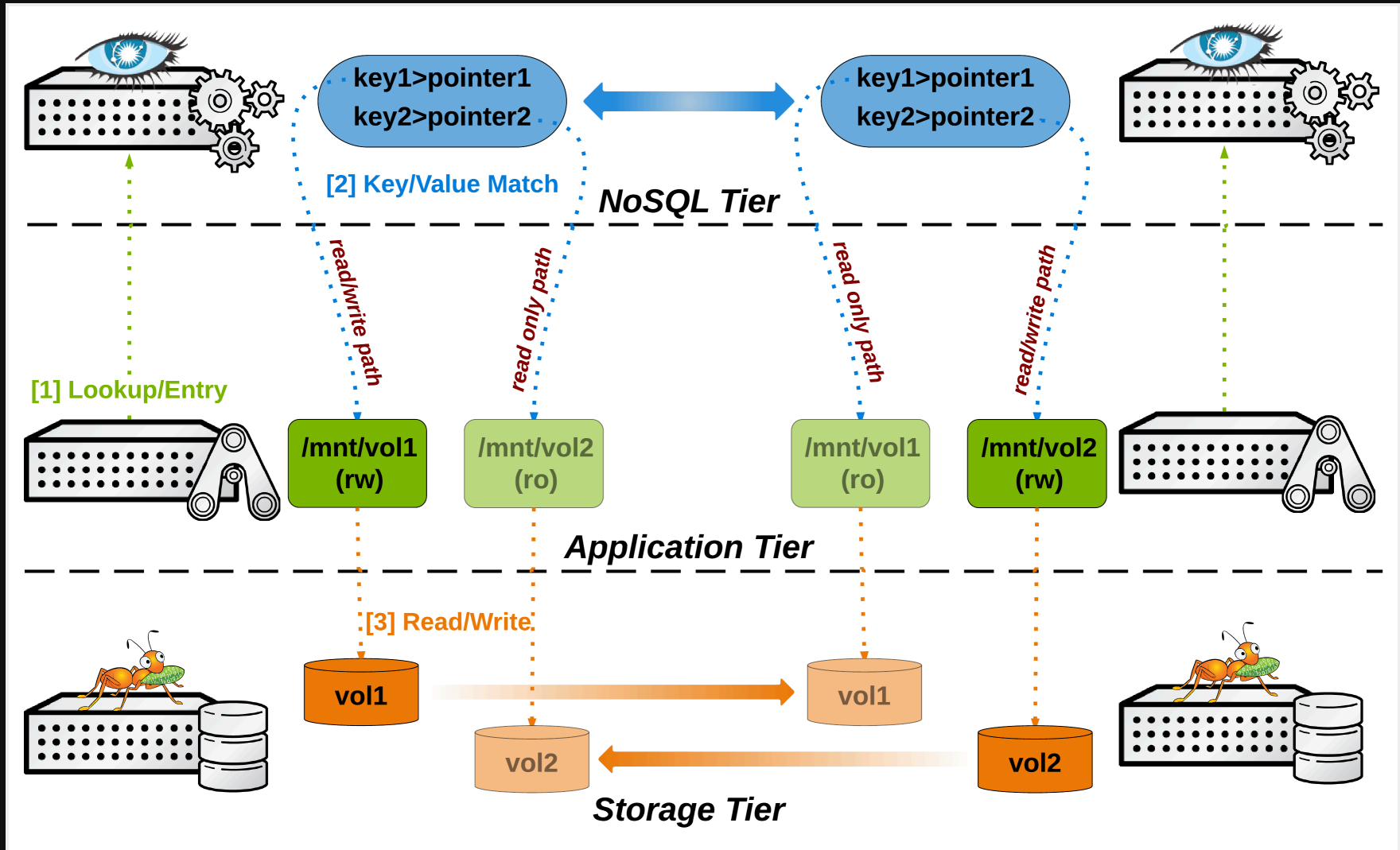
NoSQL Backend with SLA-Bound Geo-Replication

Goals

- | Replace legacy database key/blob architecture
- | Divide and conquer
 - | NoSQL layer for key/pointer
 - | Scalable storage layer for blob payload
- | Active/Active sites with 30-minute replication SLA
- | Performance tuned for small-file WORM patterns

Implementation

- | HP DL170e nodes with local disk
- | ~4TB per node
- | Cassandra replicated NoSQL layer for key/pointer
- | GlusterFS parallel geo-replication for data payload site copy exceeding SLA standards
- | Worked with Red Hat Engineering to modify application data patterns for better small-file performance





Use Case:

Storage & Compute Consolidation for Scientific Research

Goals

- | **Scale with storage needs**
 - | Eliminate need to move data between backends
 - | Keep pace with exponential demand
- | **Reduce administrative overhead; Spend more time on the science**
- | **Control and predict costs**
 - | Scale on demand
 - | Simple chargeback model
- | **Efficient resource consumption**

Implementation

- | Dell PowerEdge R720 Servers
- | OpenStack + Ceph
 - | HPC and Storage on the same commodity hardware
 - | Simple scaling, portability, and tracking for chargeback and expansion
- | 400TB virtual storage pool
 - | Ample unified storage on a flexible platform reduces administrative overhead



Use Case:

Multi-Petabyte RESTful Object Store

Goals

- | Object-based storage for thousands of cloud service customers
- | Seamlessly serve large media & backup files as well smaller payloads
- | Quick time-to-market and pain-free scalability
- | Highly cost-efficient with minimal proprietary reliance
- | Standards-based for simplified hybrid cloud deployments

Implementation

- | **Modular server-rack-row "pod" system**
 - | 6x Dell PowerEdge R515 servers per rack
 - | 10x 3TB disks per server; Total 216TB raw per rack
 - | 10x racks per row; Total 2.1PB raw per row
 - | 700TB triple-replicated customer objects
 - | Leaf-Spine mesh network for scale-out without bottleneck
- | **Ceph with RADOS Gateway**
 - | S3 & Swift access via RESTful APIs
 - | Tiered storage pools for metadata, objects, and logs
- | **Optimized Chef recipes for fast modular scaling**

Questions?

people.redhat.com/dblack



Do it!

- | Build a test environment in VMs in just minutes!
- | Get the bits:
 - | **Fedora 21** has GlusterFS and Ceph packages natively
 - | RHSS 3.0 ISO available on the **Red Hat Portal**
 - | Go upstream: **gluster.org / ceph.com**

RED HAT Technical Account Management

Premium named-resource proactive support
from your leading experts in open solutions

Contact your sales team or visit redhat.com

Supporting success. Exceeding expectations.