Phase 1:
Partner Sales Engineer Training
August 3, 2011

Jacob Shucart – Sales Engineer
jacob@gluster.com – 408.598-4306
August 2011

Agenda

- 2 Goals of Training
- Gluster is simple and powerful
- Our approach and advantages
- Talking to our prospects
- Technical overview
- Sizing and architecture
- Successful POCs
- Q&A

Goals for Today

1. How from a technical perspective, you can show how Gluster can solve customers storage needs
2. How can you get someone started on POC/Trial of Gluster

The Gluster Approach

- **A radically better way to do storage**
  - Scalable, virtualized, commoditized, and centrally managed pool…
  - The same way your customers do computing today
- **Software only**
- **Commodity hardware, public & private clouds**
- **Scale out – capacity, performance, availability**
- **Global Namespace**
- **Standard NAS→no application rewrite required**
- **Purpose built for cloud and virtual environments**
- **Over 300 deployments, multi PB, 200+ nodes**

Gluster Architecture Advantages

- Intelligence in the SW
- Leverage commodity HW
- Scale-out elastically
- Replication for reliability
- Software enables virtualization

- **Software only**
- **Complete storage operating system stack**
- **No metadata server**
  - Fully distributed architecture, no bottleneck
  - Gluster Elastic Hash
- **High performance global namespace**
  - Scale out with linear performance
  - Hundreds of petabytes
  - 1 GbE,10GbE, RDMA
- **High availability**
  - Replication to survive hardware failure
  - Self-healing
  - Data stored in NFS-like native format
- **Stackable userspace design**
  - No kernel dependencies, simple install
  - Match specific workload profiles
  - Early maturity and rich functionality

Customer Benefits from Gluster

Flexibility

- Hardware choice
- Scale both performance and capacity linearly
- Add or remove resources as needed
- Grow, shrink, migrate volumes with no downtime (equiv. to Vmotion for storage)

Ease of use

- Centrally managed
- All storage appears as a centralized pool
- Standard h/w, standard networks, standard protocols
- Deploy in minutes
- Distributed console manager

Performance

- Performance scales linearly
- By scaling load, get high performance from low-cost drives
- Works well across wide variety of workloads
- Supports demanding applications, including those requiring millions of small files, multi-TB files
- Exceptional price to performance

Economics

- Commodity hardware
- Re-purpose existing hardware
- Scale performance, capacity, & redundancy as needed, with usage

Reliability

- *N*-way replication
- Self heal
- Used in production at 100's of enterprises worldwide

Purpose built for virtual environments

- Data managed with Gluster remains POSIX compliant
- No changes required for your applications
- Virtual machine images are just files
- SAN/NAS hybrid available to provide optimal storage for both application data & VMIs

Purpose built for the cloud

- No changes to your application
- Highly available
- Multi-tenant
- Usage-based pricing

When talking to prospects about Gluster…

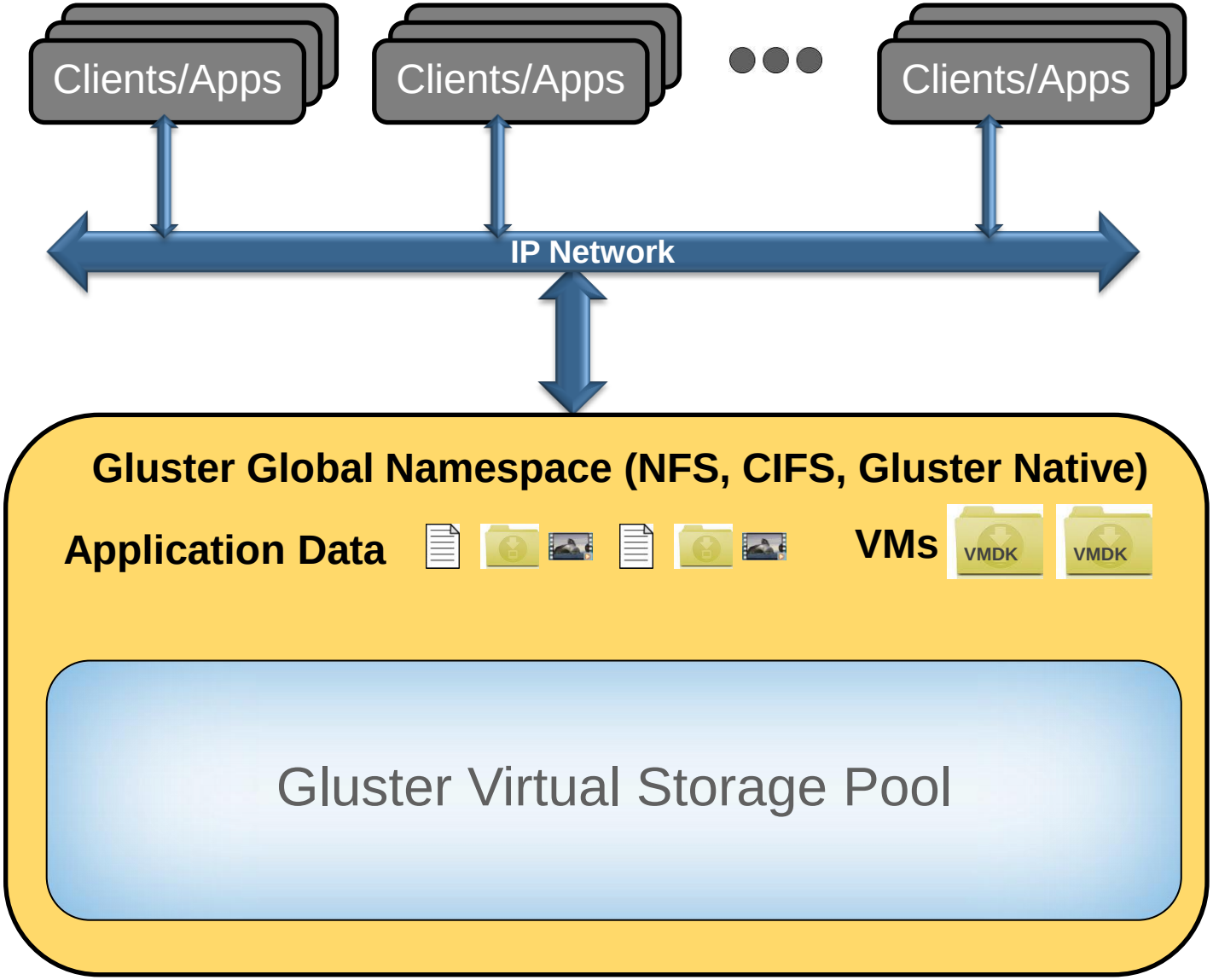
What is Gluster going to solve for prospects?

- Unstructured data grows rapidly (capacity)
- Adding capacity on the fly
- Not all data needs tier 1 costly solutions
- No single point of failure
- Performance needs
- Economics-TCO
- Flexible software approach to storage

What questions come up during calls?

- How does Gluster allow linear scalability?
- What is the difference between the open source version and commercial?
- How does Gluster differ from other distributed file systems?
- What is the ideal configuration for my use case, our recommendations?
- How do you handle high availability?
- Failover?
- RAID?

A Standard Gluster Deployment



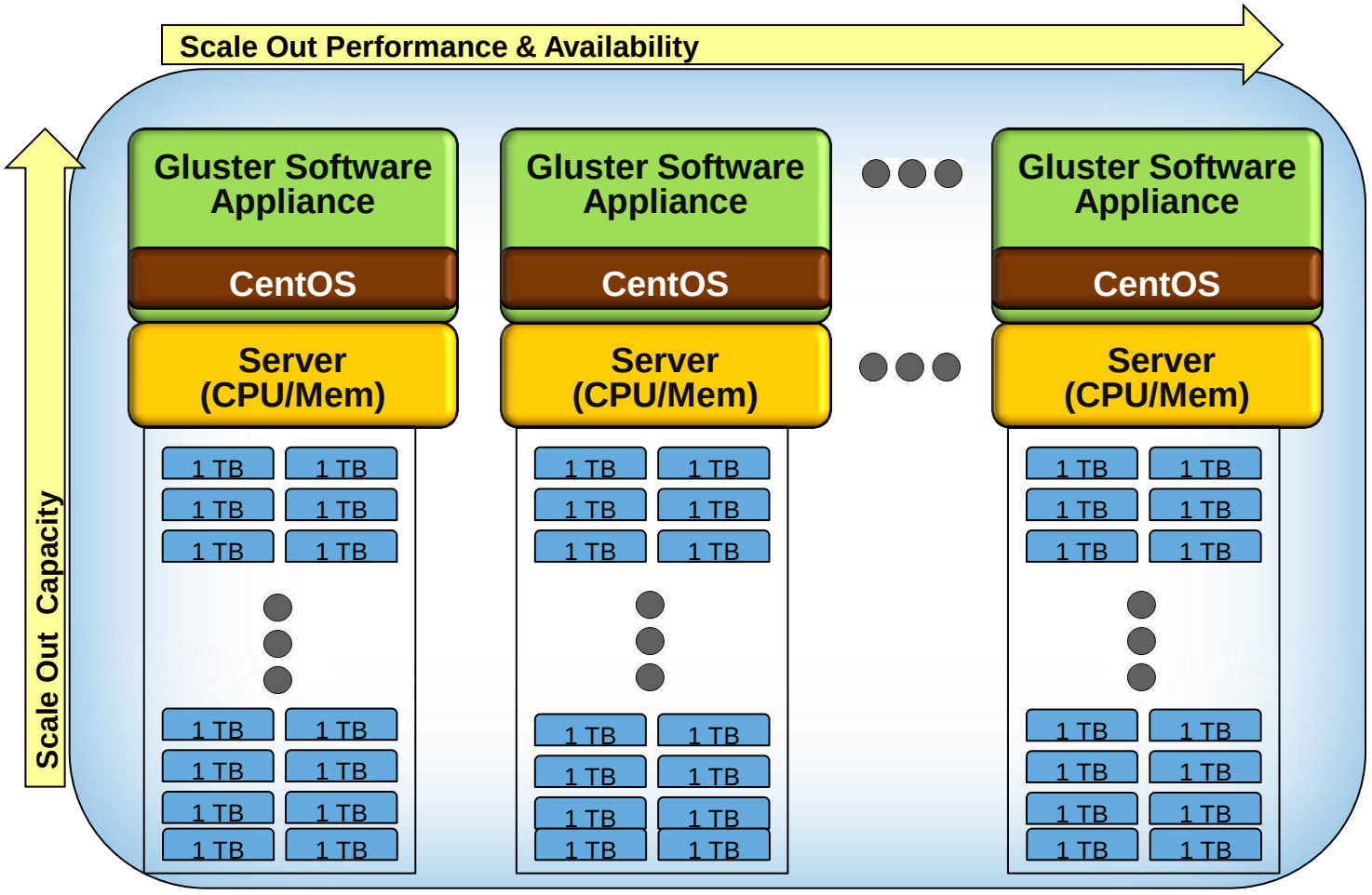
Standard clients running standard apps

Over any standard IP network

Access application data, as files & folders, in a global namespace, using a variety of standard protocols

Stored in a commoditized, virtualized, scale-out, centrally managed pool DAS, SAN, NAS, cloud

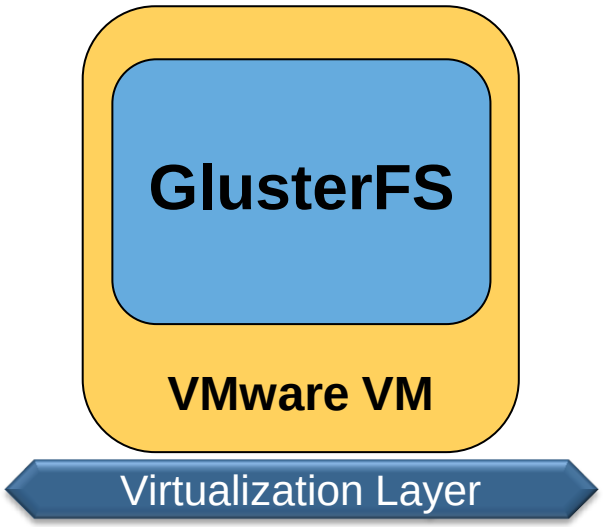
Anatomy of a storage pool: On Premise



Gluster Storage Software Appliances

- Deployed on any VMware certified Hardware
- With any VMware certified underlying storage: DAS, JBOD, SAN
- Aggregate CPU, memory, network, capacity
- Scale out capacity and performance as needed
- Replicate n-way as needed for availability

Virtual Storage Appliance



Simplified configuration

- Works out-of-the-box on virtually any VM Ware certified hardware
- The solution is properly configured from the outset
- Reduces issues associated with improper firmware, inconsistent network cards, etc.
- Agility with Control - increase the speed of deploying storage resources

Simplified management

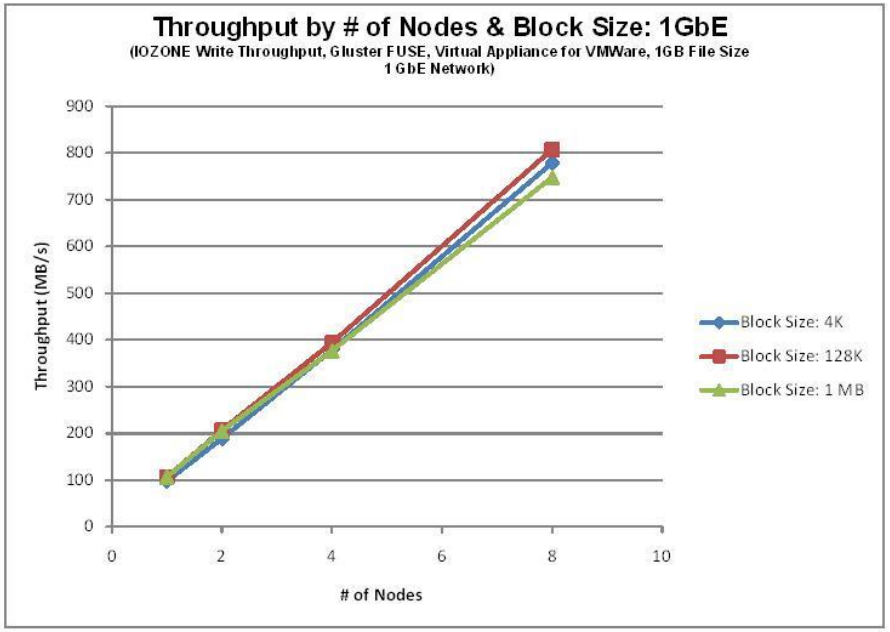
- Storage-on-demand…deploying Gluster on multiple nodes takes minutes
- Deployment, provisioning and upgrades vastly simplified—or automated

Easier transition

- Gluster can be deployed across heterogeneous set of hardware—including legacy equipment
- Much better economics

Virtually no performance hit

Performance Benchmarks



# Nodes	Block Size: 4K	Block Size: 128K	Block Size: 1 MB
1	97.96	107.19	107.08
2	189.08	207.41	205.31
4	378.56	394.4	377.05
8	779.07	807.8	748.23

- **Performance benchmarks show we get close to network saturation with small # of nodes**
- **Significantly better than Isilon/EMC**
- **For most target applications, far exceed requirements at fraction of cost**
- **A customer is achieving 2.8 GB/s replicated throughput**
  - 1 client, 2 servers mirrored. 2.8GB total throughput.
  - (1.3 writes, 1.5 reads) Fusion I/O storage.

Small File Performance per Node (NB: We can scale linearly)

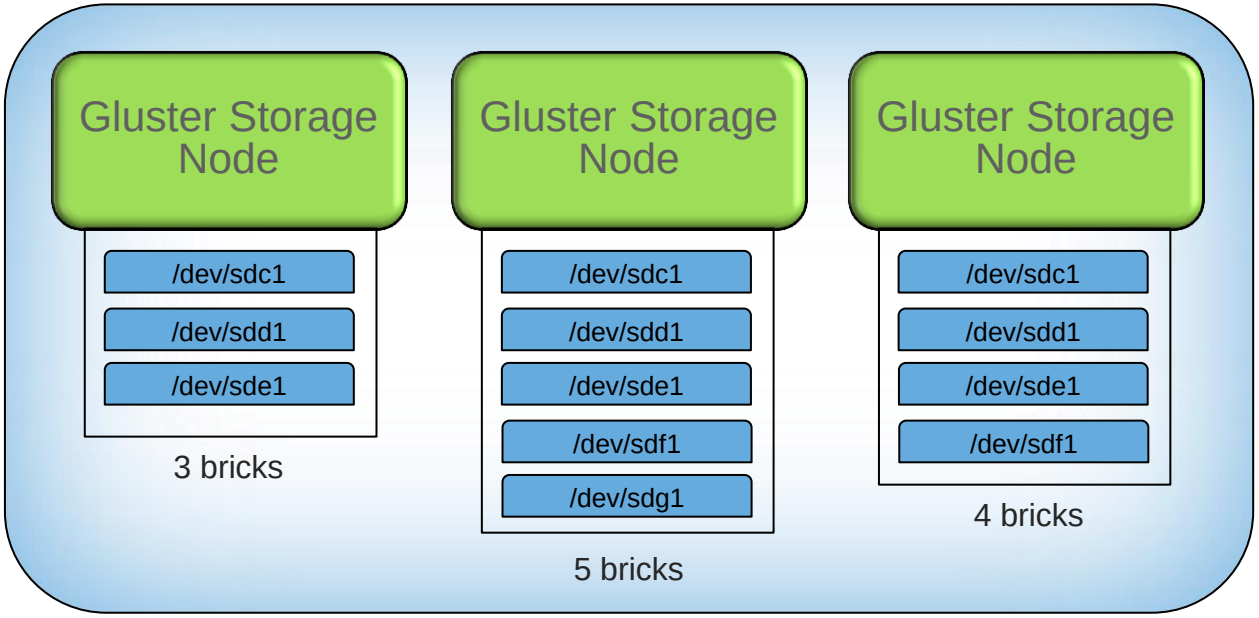
Operations	4K files/sec
Create	319.5
Read	344.8
Write	316.5

1GbE, single node, SATA drives,

Gluster Technical Fundamentals

A Brick

- A brick is the combination of a node and a file system.
- Each brick is limited to 16TB (ext3 limit).
- No limit to the number bricks per node.
- Gluster operates at the brick level, not at the node level.
- Ideally each brick in a cluster should be the same size.



A Gluster cluster with 12 bricks.

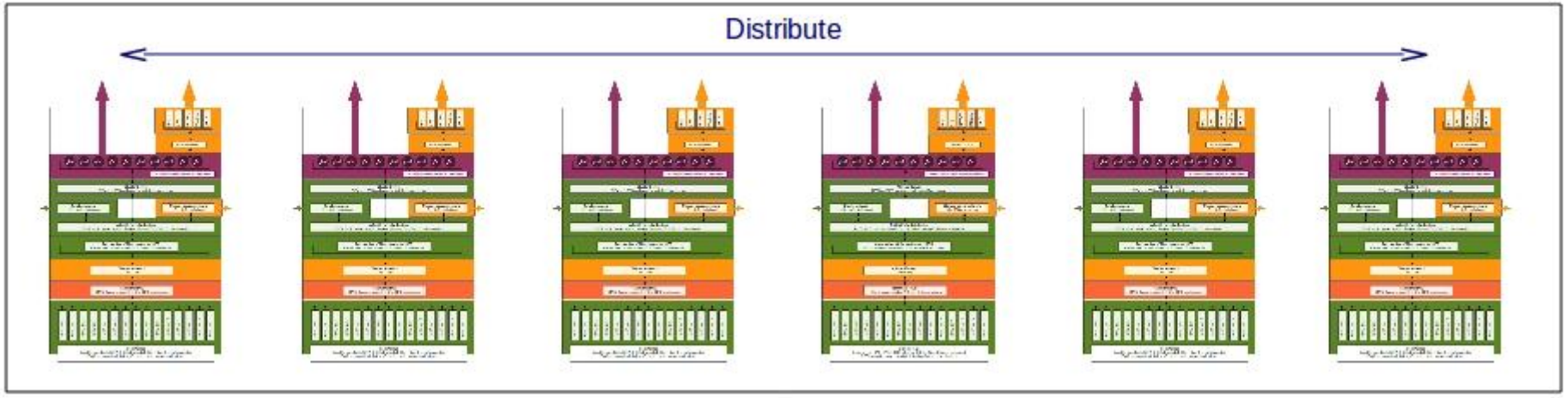
Gluster Technical Fundamentals

- **Elastic Hashing Algorithm**
  - Every folder in a volume is assigned a equal segment of the 32bit number space across bricks.
    - Each folder across 12 bricks, –
      - brick1 = 0- 357913941
      - brick2 = 357913942 - 715827883
      - brick3 = 715827884 – 1073741823
      - …brick12 = 3937053354 – 4294967295
    - Every folder and sub-folder gets the full range, 0 – 4294967295.
  - The EHA hashes the name of the file being read | written.
    - \...\...\...\glusterrules.txt = 815827884
    - We use the Davies-Meyer hash.
  - The file is read | written on the brick that matches the path and filename hash.
  - The avalanche effect of a hashing algorithm prevents hotspots.
    - Regardless of the similarities between filenames/path the hash result is sufficiently different.
    - Additional, redundant steps are taken to prevent hotspots.
  - Adding a brick to a volume updates the EHA graph on each node.
    - Running a volume rebalance after adding a brick physically moves files to match the new EHA graph.
    - Adding a brick (`gluster volume add brick`) without running a rebalance results in new files being written to the new bricks and link files being created on first access for existing files. This can be slow and isn't recommended.

Gluster Technical Fundamentals

Distribute only

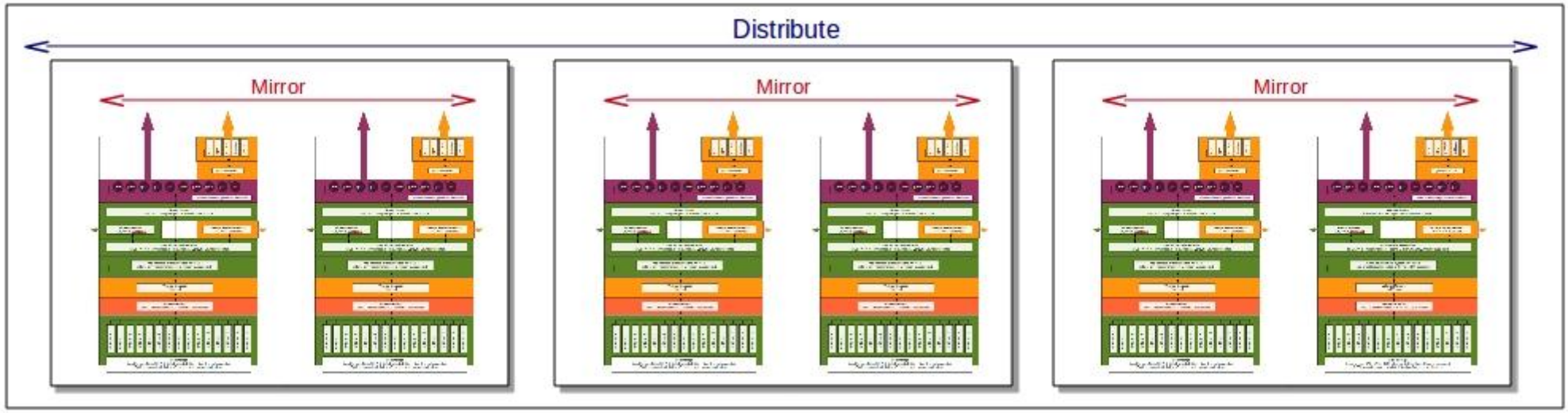
- Non-redundant at the brick level
  - Cuts hardware, software costs in half.
  - Failure of a brick or node results in loss of access to the data on those bricks.
  - Writes destined to the failed brick will fail.
  - Redundant RAID, hardware is strongly recommended.



Gluster Technical Fundamentals

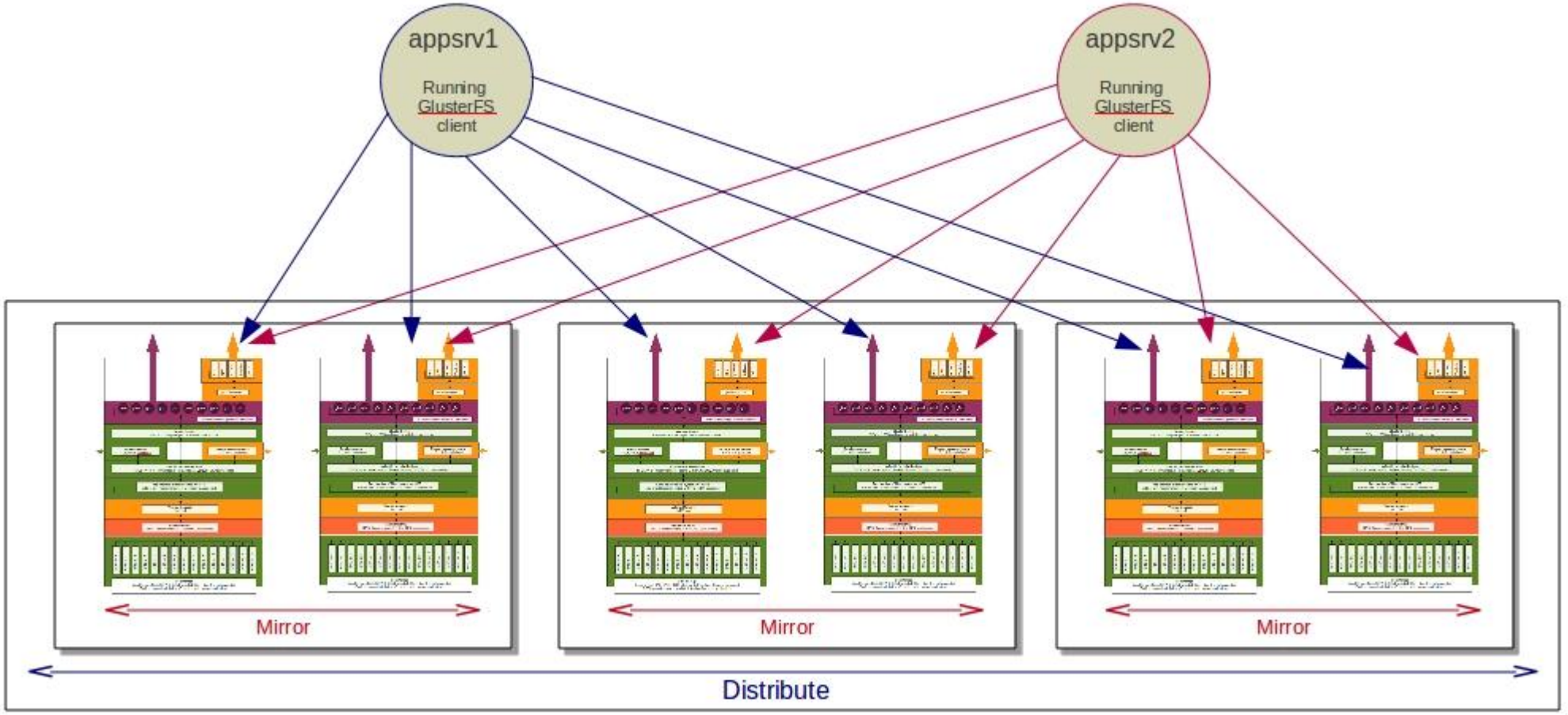
Distribute with Replica

- Redundant at the brick level
  - Failure of a brick or node does not affect I/O.
  - Writes are written simultaneously to each replica.
  - Any number of replicas are supported.
  - Gluster Native, CIFS, and NFS support stateful failover. (Gluster Native only in AWS)
  - Redundant RAID, hardware is strongly recommended.



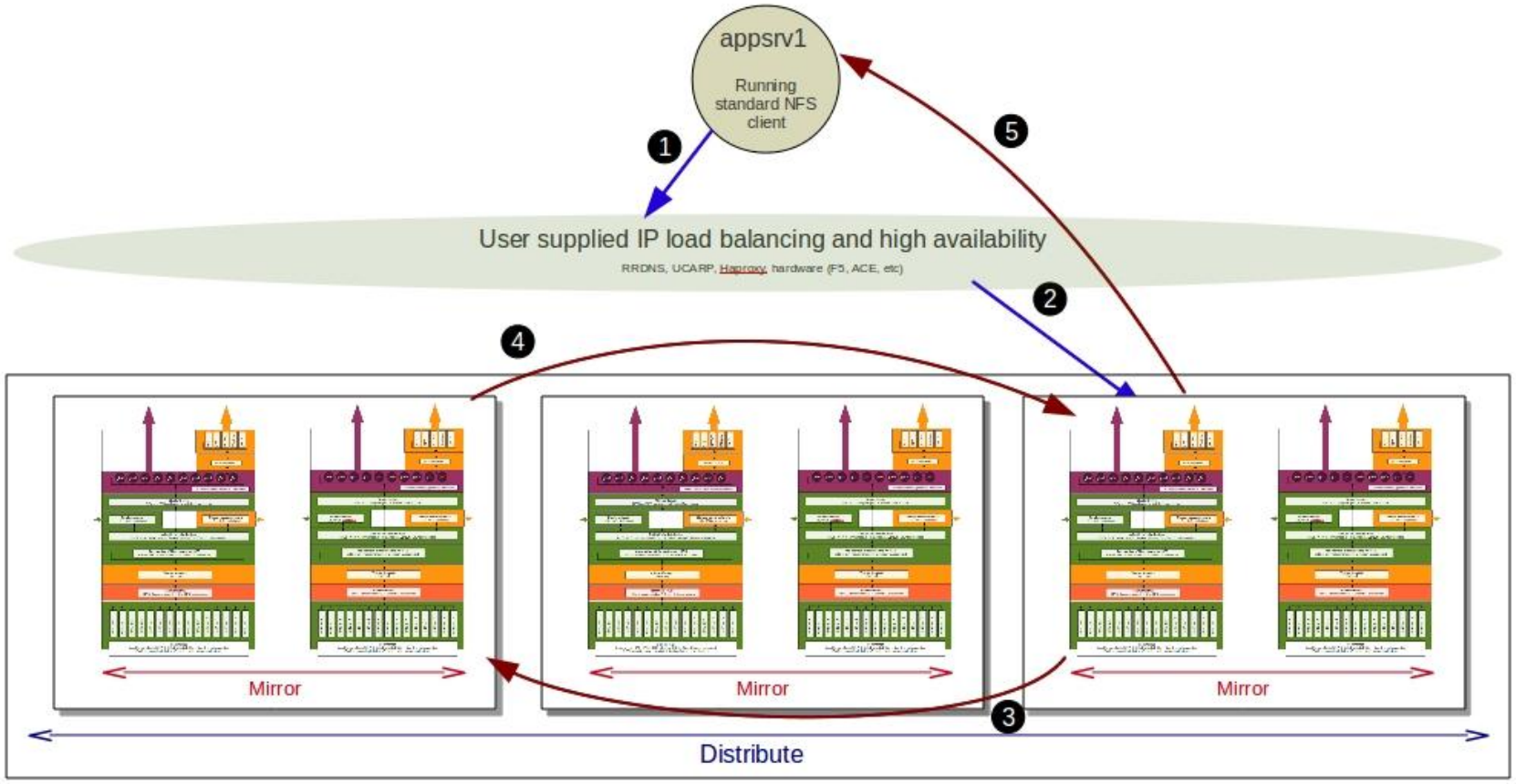
Gluster Technical Fundamentals

GlusterFS Native client data flow



Gluster Technical Fundamentals

NFS, CIFS dataflow



HA for NFS and CIFS

- **Any IP failover tool can work for NFS**
  - Appliance based load balancers with heartbeat such as F5
  - Linux heartbeat, ucarp, CTDB
  - Not all failover works for CIFS as that requires some session handling
- **CTDB is what we use**
  - It is very simple to configure
  - Works for NFS
  - Works for CIFS
  - Is very robust and configurable
- **Round robin DNS for load balancing**
  - You can use any load balancer you want
  - RRDNS is simple to configure and works well
  - Prevents hot spots of activity

Sizing and Architecture

- **Gluster performance relies on hardware**
  - Number of systems depends on performance and capacity
  - There are many ways to meet customer needs
  - 2U & 4U DAS systems and JBODS are great building blocks
- **Capacity-centric environments**
  - 2U & 4U DAS systems with multiple JBODS
  - Lower RAM and CPU requirements
  - Lower network requirements
- **Mixed capacity and performance environments**
  - 2U & 4U DAS systems with 1-2 JBODS max
  - Higher RAM and CPU requirements
  - Low to high network requirements
- **High performance environments**
  - 1U or 2U systems with no JBODS
  - Highest RAM and CPU requirements
  - Fast disks and fast network

Successful POCs when Needed

- **Most users tend to want to evaluate new technology**
  - Evaluations typically only take 5-10 days to run
  - They take longer because they have not committed the time…try and set timelines
- **Gather information about required environment**
  - What does their test environment look like vs. what will be deployed?
  - What are the performance/capacity requirements?
  - IOPS vs. Throughput?
  - What does the workload look like? File sizes, transactions, etc
- **Identify some test cases that work**
  - Is there a smaller set of data that can work for a POC?
  - Does the prospect know? Can we tell them what tests they should run?
  - Do they need hardware for the test?
  - Do they have their own test plan?
- **Validate the test cases and define success criteria**
  - Have you defined the timeline?
  - Are the prospect's goals achievable within that timeline?
  - Will they move forward if their criteria are met?
- **Gluster is here to help with this process-USE US**

Demonstration

Strong Gluster Community

Who Is Using Gluster

- **125,000+ downloads**
  - ~7,000 /month
- **300+ 'known' deployments**
  - 45 countries
  - Over 5 PB of data
- **1,200+ registered users**
  - Mailing lists, IRC chat
- **Community contributions**
  - Diverse testing environments
  - Bugs identification and fixes
  - Code contributions

Gluster Customers

- **Scale-Out NAS**
  - **High Performance**
  - **Archive**
- **Public Cloud**
  - **Retail + AWS**
  - **Service providers**
- **Private Cloud**
- **Virtual storage for VM environments**
- **From single TB to PB-scale**

Pandora Internet Radio

- 1.2 PB of audio served per week
- 13 million files
- Over 50 GB/sec peak traffic

Problem

- Explosive user & title growth
- As many as 12 file formats for each song
- 'Hot' content and long tail

Solution

- Three data centers, each with a six-node GlusterFS cluster
- Replication for high availability
- 250+ TB total capacity

Benefits

- Easily scale capacity
- Centralized management; one administrator to manage day-to-day operations
- No changes to application
- Higher reliability

Brightcove

- Over 1 PB currently in Gluster
- Separate 4 PB project in the works

Problem

- Cloud based online video platform
- Explosive customer & title growth
- Massive video in multiple locations
- Costs rising, esp. with HD formats

Solution

- Complete scale-out based on commodity DAS/JBOD
- Replication for high availability
- 1PB total capacity

Benefits

- Easily scale capacity
- Centralized management; one administrator to manage day-to-day operations
- Higher reliability
- Path to multi-site

Cincinnati Bell Technology Solutions

- **Large scale VM storage**
- **Low cost service delivery for enterprise customer**
- **Drastic reduction in provisioning time**

Problem

- Host a dedicated enterprise cloud solution
- Large scale VMware environment
- Need high availability

Solution

- Gluster for VM storage, NFS to clients
- SAS drives on back-end
- Replication for high availability

Benefits

- Storage provisioning from 6 wks. to 15 min.
- Vendor agnostic storage
- Low cost of service delivery
- Elastic growth

Partners Healthcare

Private Cloud: Centralized Storage as a Service

PARTNERS HEALTHCARE

- **Over 500 TB**
- **9 Sun "Thumper" systems in cluster**

Problem

- Capacity growth from 144TB to 1+PB
- Multiple distributed users/departments
- Multi OS access - Windows, Linux and Unix

Solution

- GlusterFS Cluster
- Solaris/ZFS/x4500 w/ InfiniBand
- Native CIFS/ NFS access

Benefits

- Capacity on demand / pay as you grow
- Centralized management
- Higher reliability
- OPEX decreased by 10X

Commercial Deployments

Customer	Nodes	Capacity	Use Case
Large Systems Integrator	132	unknown	HPC – nuclear research
Healthcare	12	1PB	cloud
CDN	16	1.4PB	CDN – HD Video
Online Music Provider	44	500TB	Music, cloud
Large eTail	12	100TB	Thumbnail images
Security Company	10	250TB	Internal cloud storage, shared files
Semiconductor Manufacturer	12	100TB	Code files for dev team
Telco	12	250TB	VM image storage
University	16	100TB	DNA sequencing
Research institute	16	50TB	Research
Bank	4	50TB	Financial Modeling

- **Send any questions after to: gpartner@gluster.com**

THANK YOU!
Q&A

Scale-Out NAS for the Cloud and Virtual Environments

Thank You

Jacob Shucart | jacob@gluster.com
408-598-4306