

open



USE



IMPROVE



EVANGELIZE

Availability of PostgreSQL in the Datacenter

- Detlef Ulherr
- Sun Microsystems
-

開
放
的
열린
مفتوح
libre
मुक्त
ಮುಕ್ತ
livre
libero
ముక్త
开放的
açık
open
nyílt
:::
πικρ
オープン
livre
ανοικτό
offen
otevřený
öppen
открытый
வெளிப்படை



Outline

- Increasing the availability of PostgreSQL
- Why is High Availability Important?
- What is “Open High Availability Cluster”?
- How Solaris Cluster Provides High Availability
- Cluster Agents
- PostgreSQL and Open HA Cluster
- Cluster in Action



Increasing the Availability of PostgreSQL



PostgreSQL availability with Solaris

- Use the features in Solaris 10 / Nevada
- Use the shared disk approach (Sun Cluster / Open HA Cluster), shared nothing available as well
- Configure a warm standby with `pg_standby`
- Use master slave replication with Slony-I
- Combine the replication with shared disk
- If you need disaster recovery, use Sun Cluster Geo / Open Ha Geo



Solaris Features for more PostgreSQL Availability



Solaris 10 availability features

- Service Management Facility (SMF)
 - Offers process monitoring
- IP Multipathing (IPMP)
 - Protects from network adapter failures
- Isolation with Containers
 - Separate multiple PostgreSQL installations



Why Is High Availability Important?



Why is HA Important?

- Downtime is costly
- Failures are inevitable
 - Hardware, Software, Human Error, Disaster
 - Automated recovery the goal
- With a single physical system, single points of failure are catastrophic
 - Network card dies, CPU misbehaves, Disk drive crashes ...

HA Clusters and Solaris automate the recovery process from inevitable failures, minimizing downtime and cost



What is Open High Availability Cluster?



Open High Availability Cluster

- The open-source code base for Solaris™ Cluster
 - Based on Solaris Cluster 3.2
- Hosted by the HA Clusters community group on opensolaris.org
 - <http://www.opensolaris.org/os/community/ha-clusters/ohac/>
- Code available under the CDDL

opensolaris™



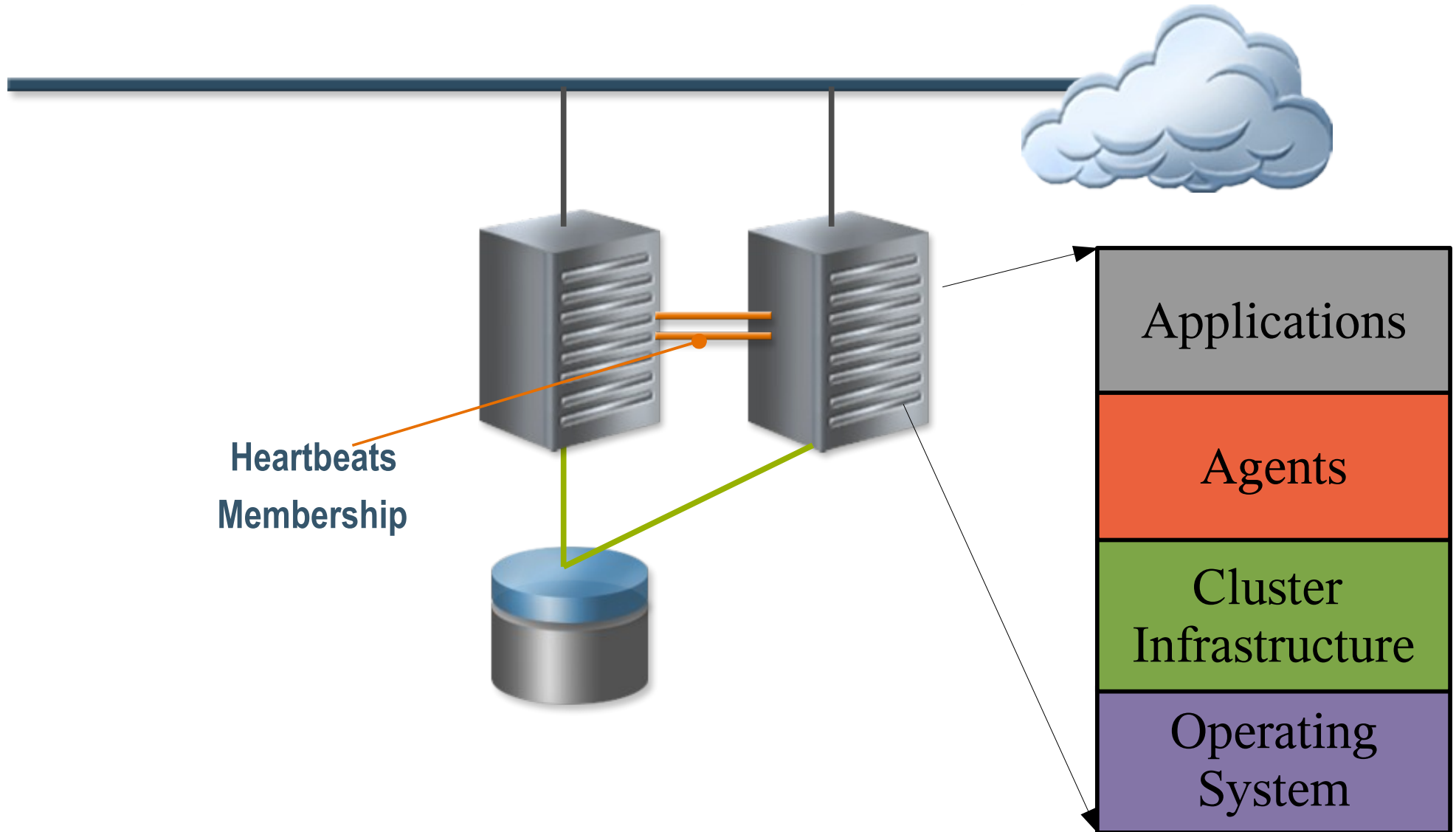
How Solaris Cluster Provides High Availability



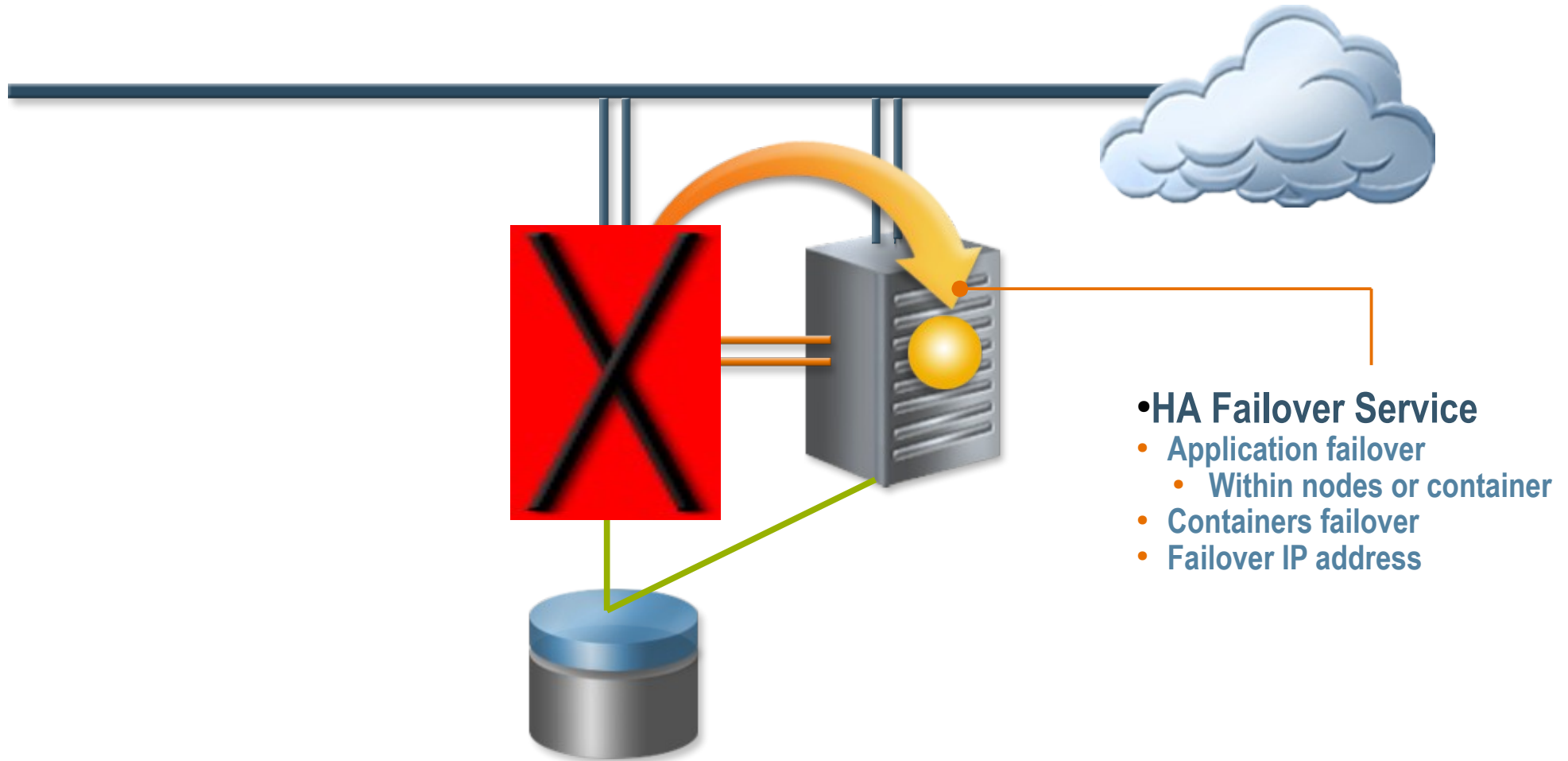
Platform for High Availability

- Tolerates Single Points of Failure (and some double failures)
- Hardware redundancy with off-the-shelf hardware
- Robust software HA infrastructure
 - Monitors the health of cluster infrastructure (hosts, shared storage, and network)
 - Orchestrates recovery of applications and cluster infrastructure (shared storage and IP networking)
 - Integrated tightly with Solaris Operating System
 - Robust membership including quorum to prevent partitions
 - Disk Fencing ensures Data Integrity in spite of failures

Solaris Cluster Stack

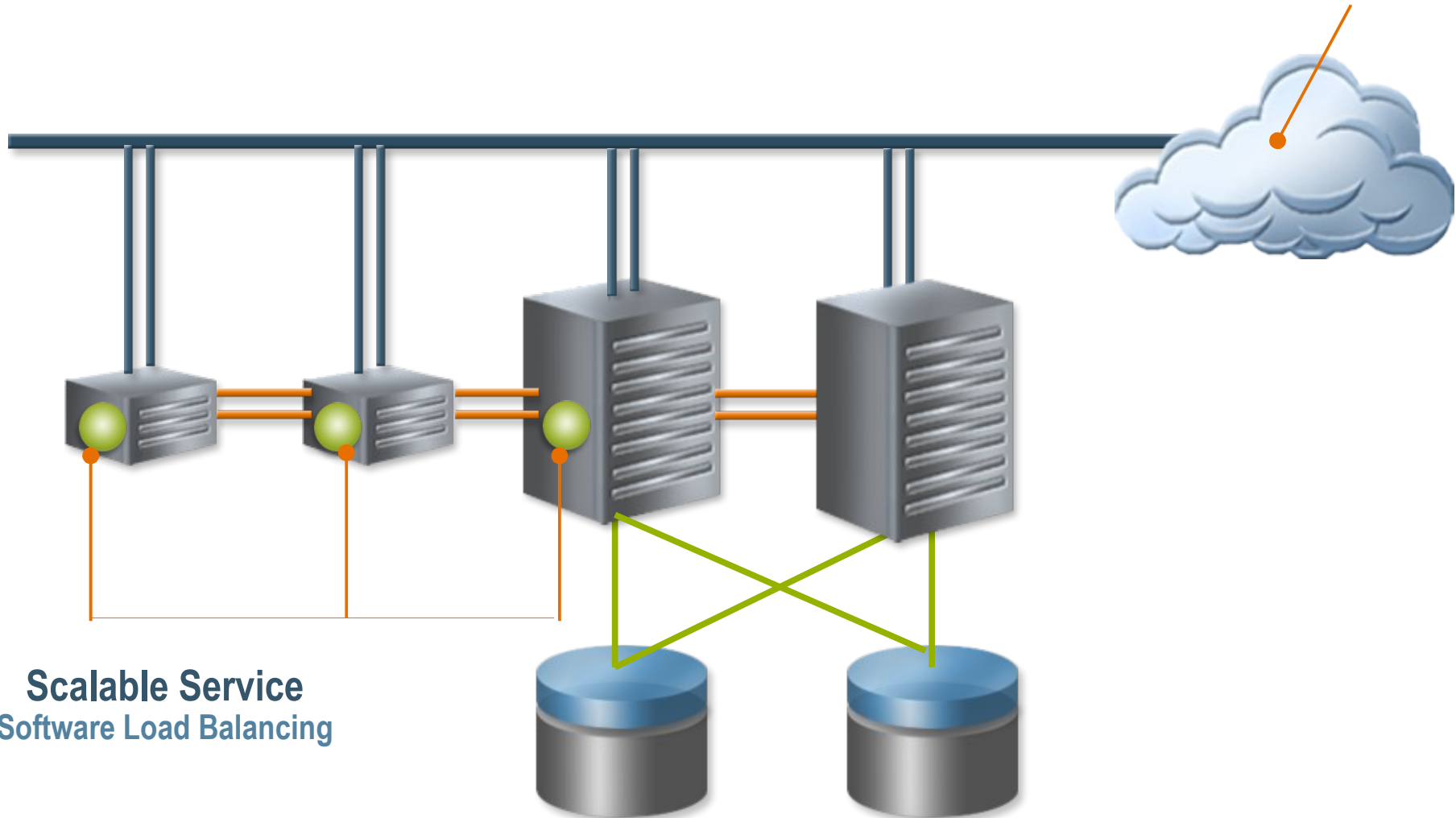


Failover Service



Scalable Service

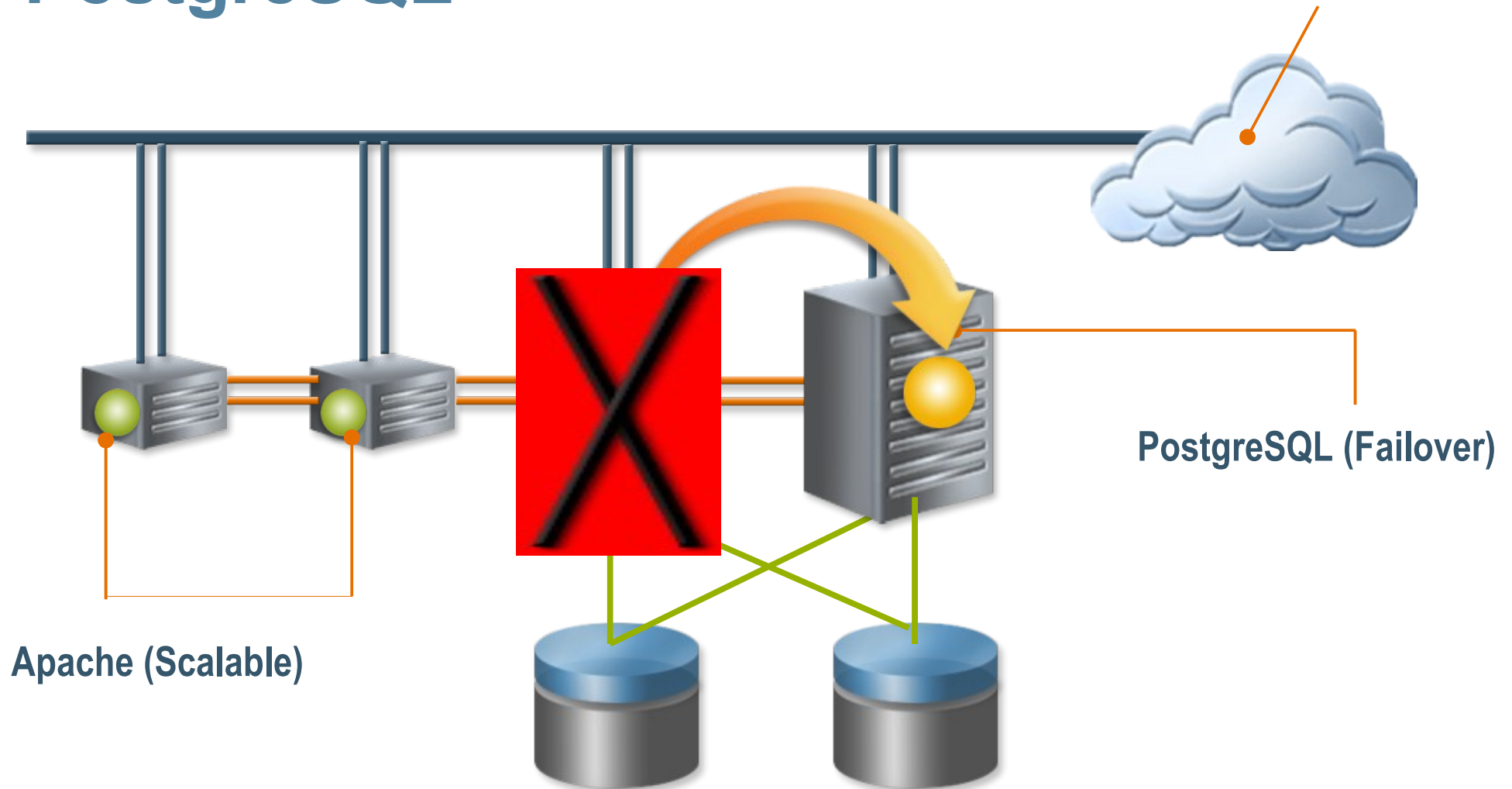
Global Network Service
Provides Global IP address
with failure protection



Scalable Service
Software Load Balancing

Example: Apache and PostgreSQL

Global Network Service
Provides Global IP address
with failure protection

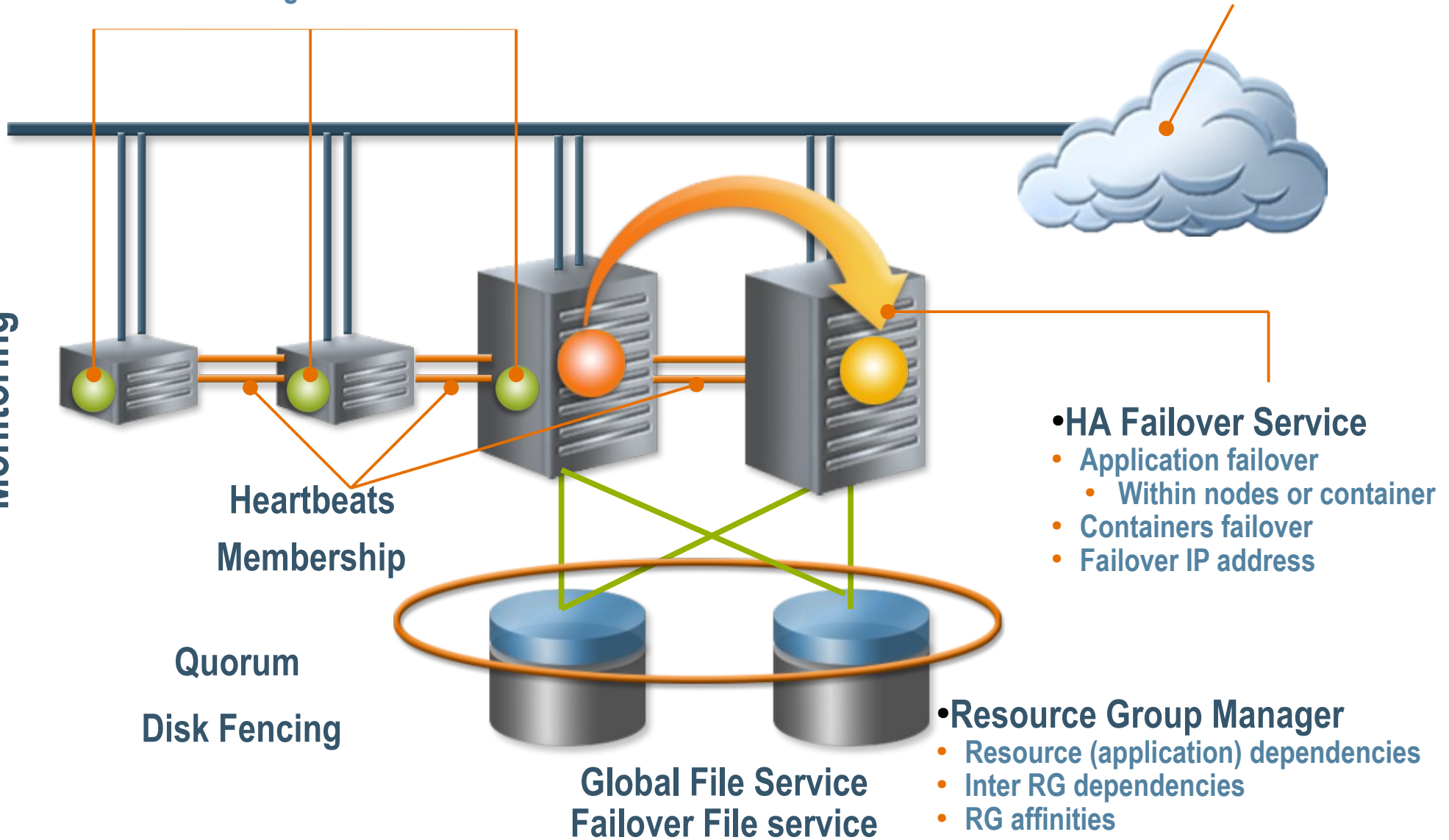


Solaris Cluster Architecture

Scalable Service
Software Load Balancing

Global Network Service
Provides Global IP address
with failure protection

Monitoring



Heartbeats
Membership

Quorum
Disk Fencing

Global File Service
Failover File service

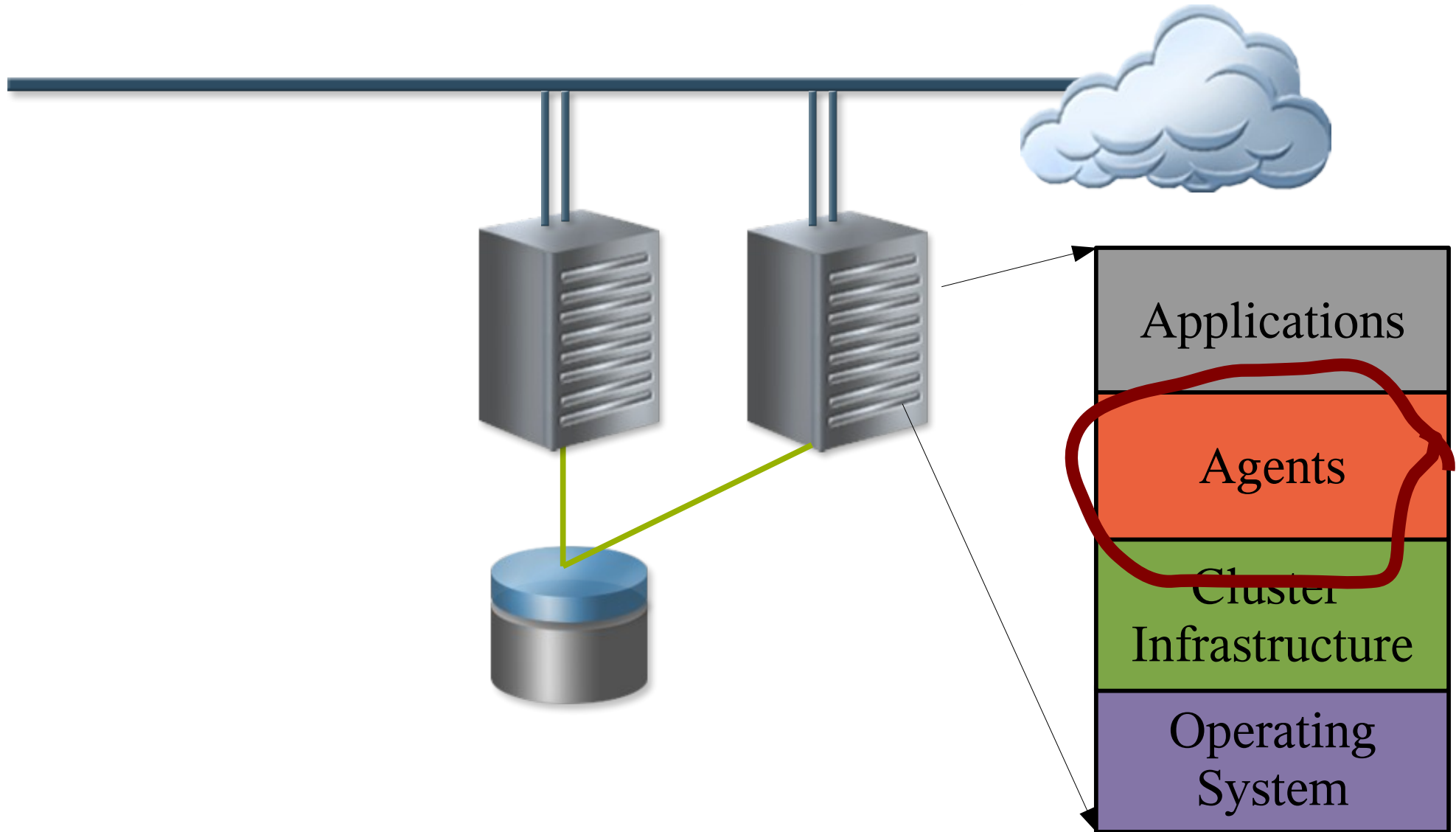
- **HA Failover Service**
 - Application failover
 - Within nodes or container
 - Containers failover
 - Failover IP address

- **Resource Group Manager**
 - Resource (application) dependencies
 - Inter RG dependencies
 - RG affinities



Cluster Agents

Solaris Cluster Stack





Cluster Agents (Data Services)

- Applications run on cluster unmodified (off-the-shelf)
- Cluster Agents are the “glue” layer between applications and cluster infrastructure
 - Custom agent for each application
 - ♦ Interacts with cluster core through APIs
 - ♦ Provides start, stop, and other commands specific to the application to be called by the cluster framework
 - ♦ Provides monitor daemon specific to the application
- Application cannot break into the high availability business on Solaris Cluster without an agent!



Agents Development

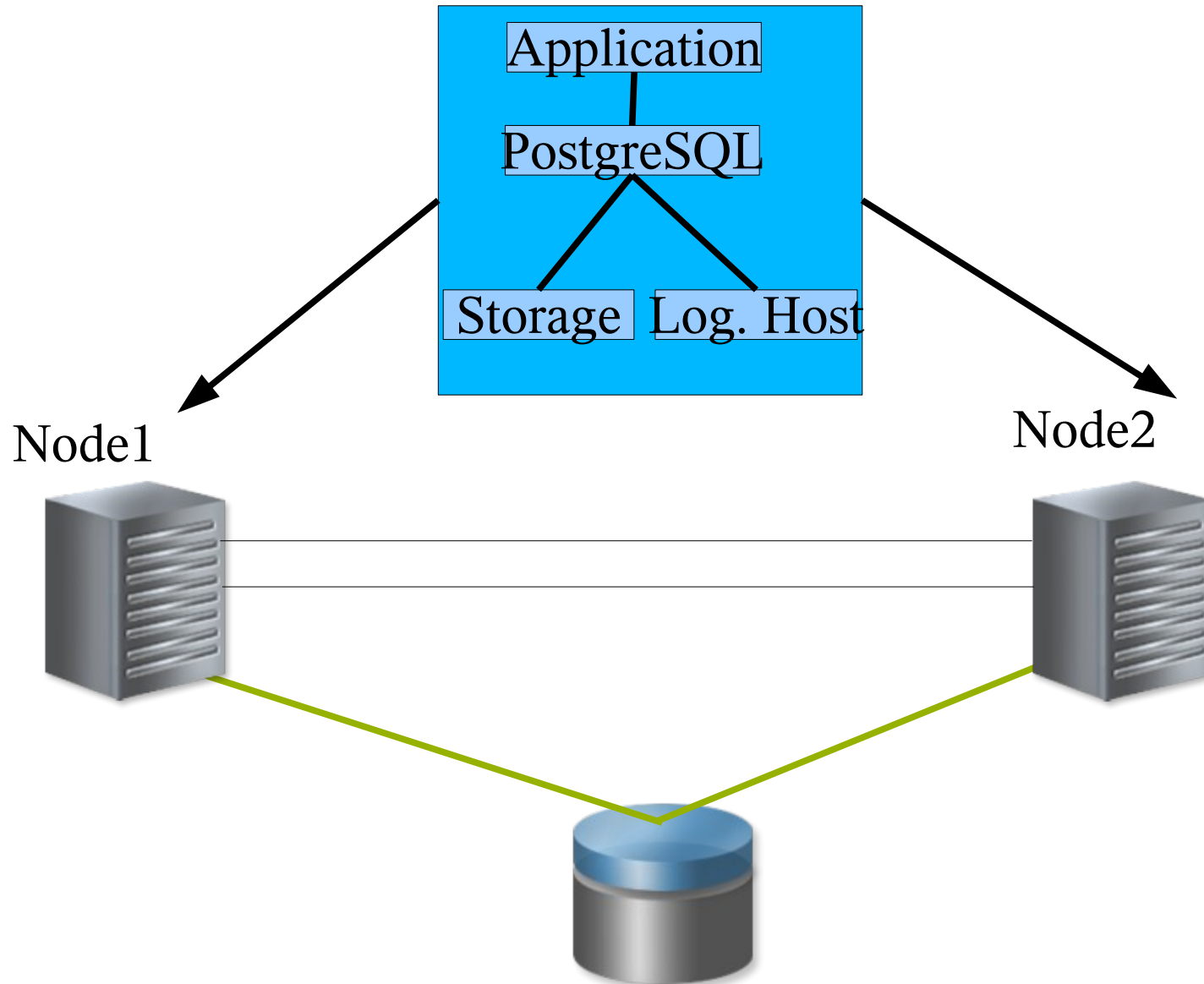
Several choices available

- Sun Cluster Agent Builder
 - Generic Data Service (GDS)
 - Data Service Development Library (DSDL)
 - RGM Application Programming Interface (API)
- Advanced GDS Template available from OHAC

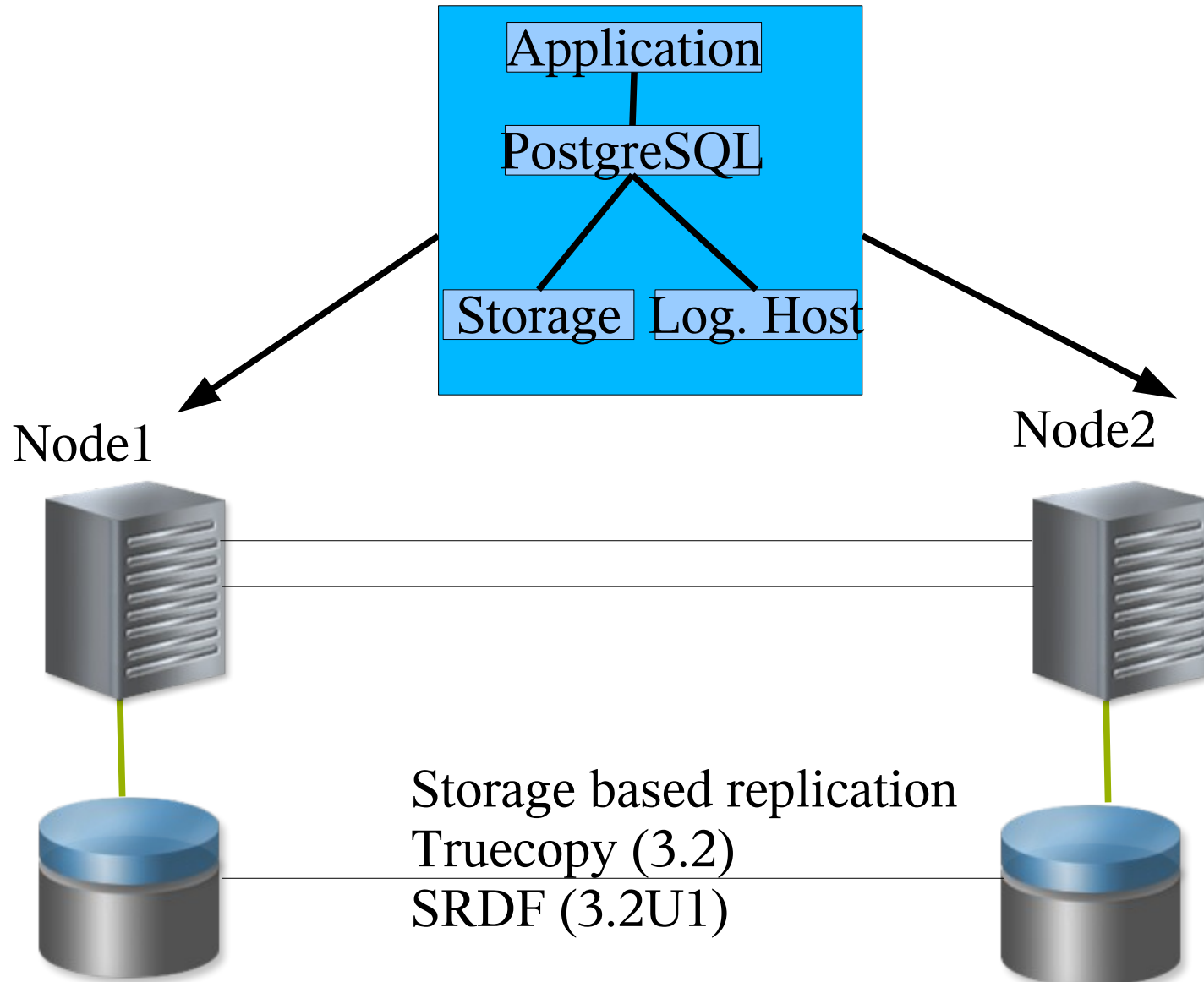


PostgreSQL and Open Ha Cluster

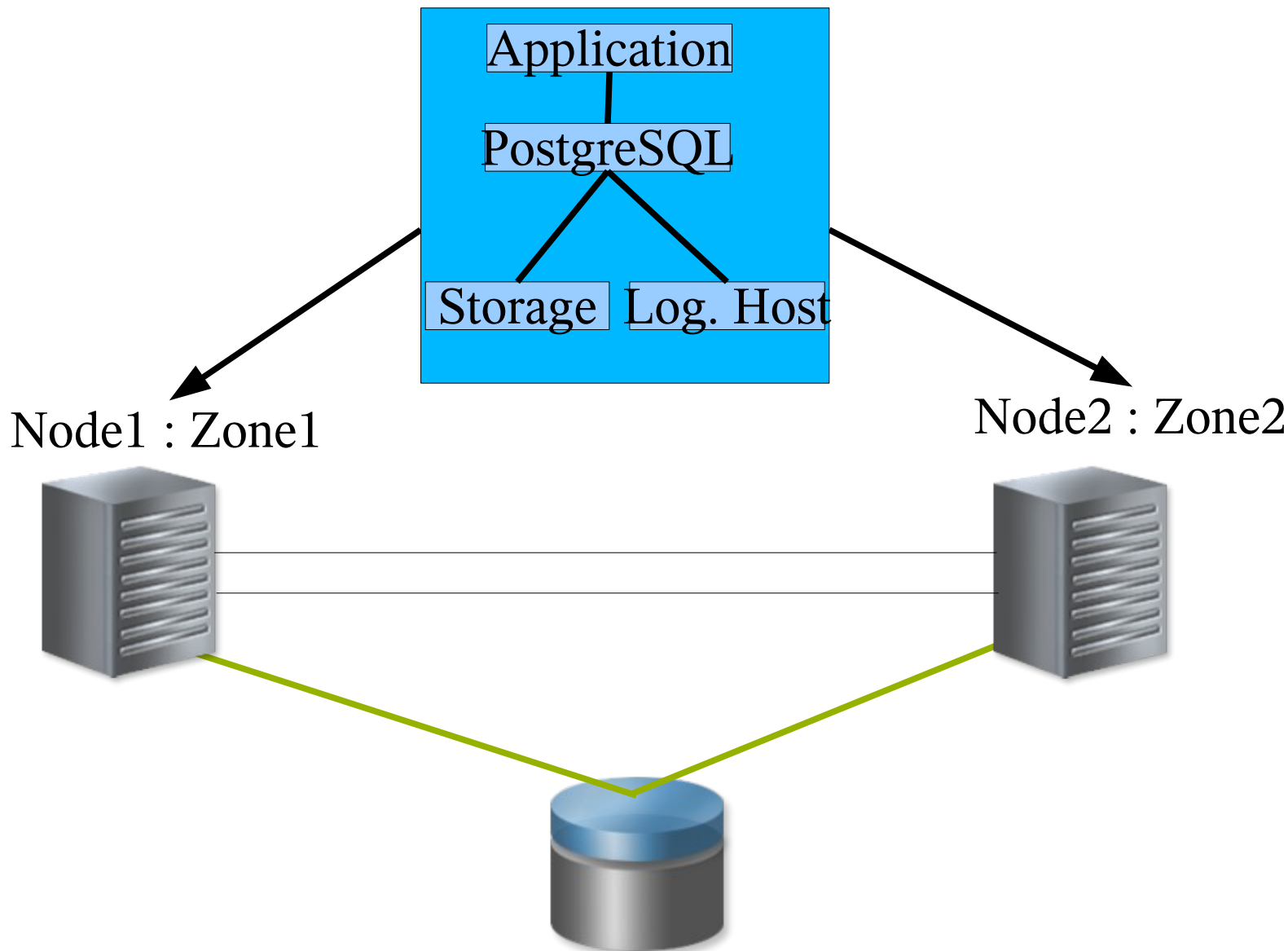
PostgreSQL (Shared Disk) Global Zone



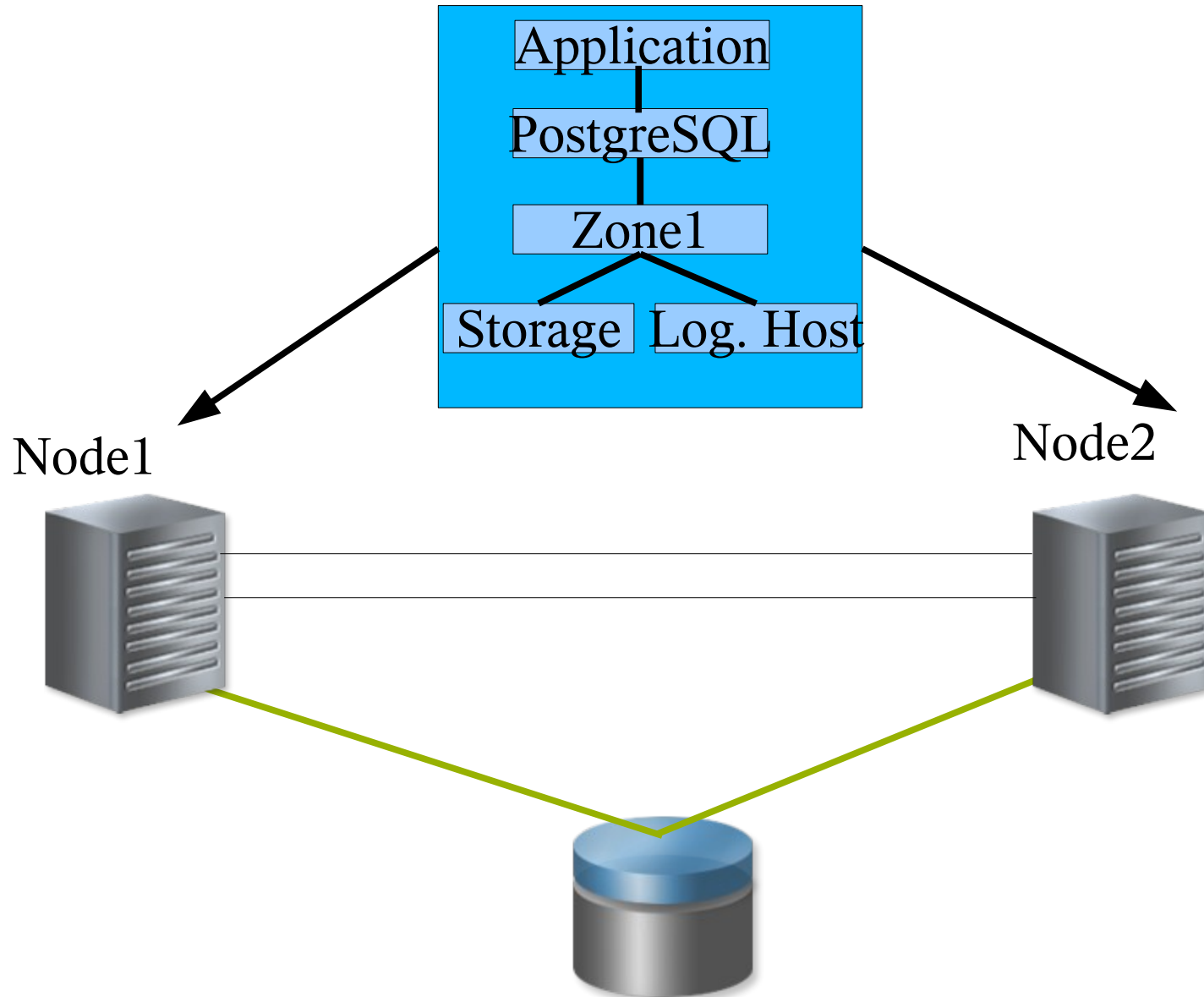
PostgreSQL (Shared Disk) Global Zone



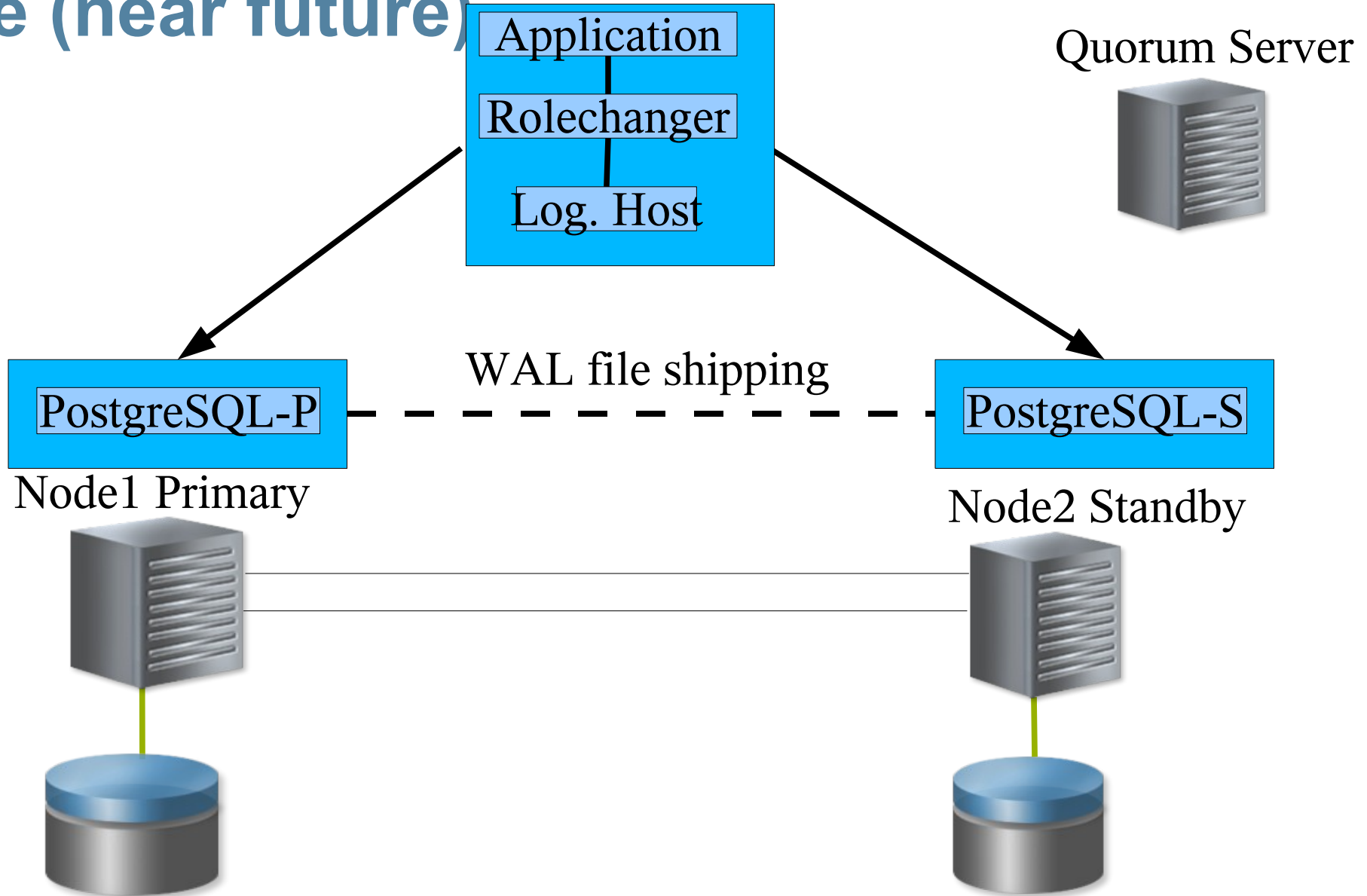
PostgreSQL (Shared Disk) Local Zone



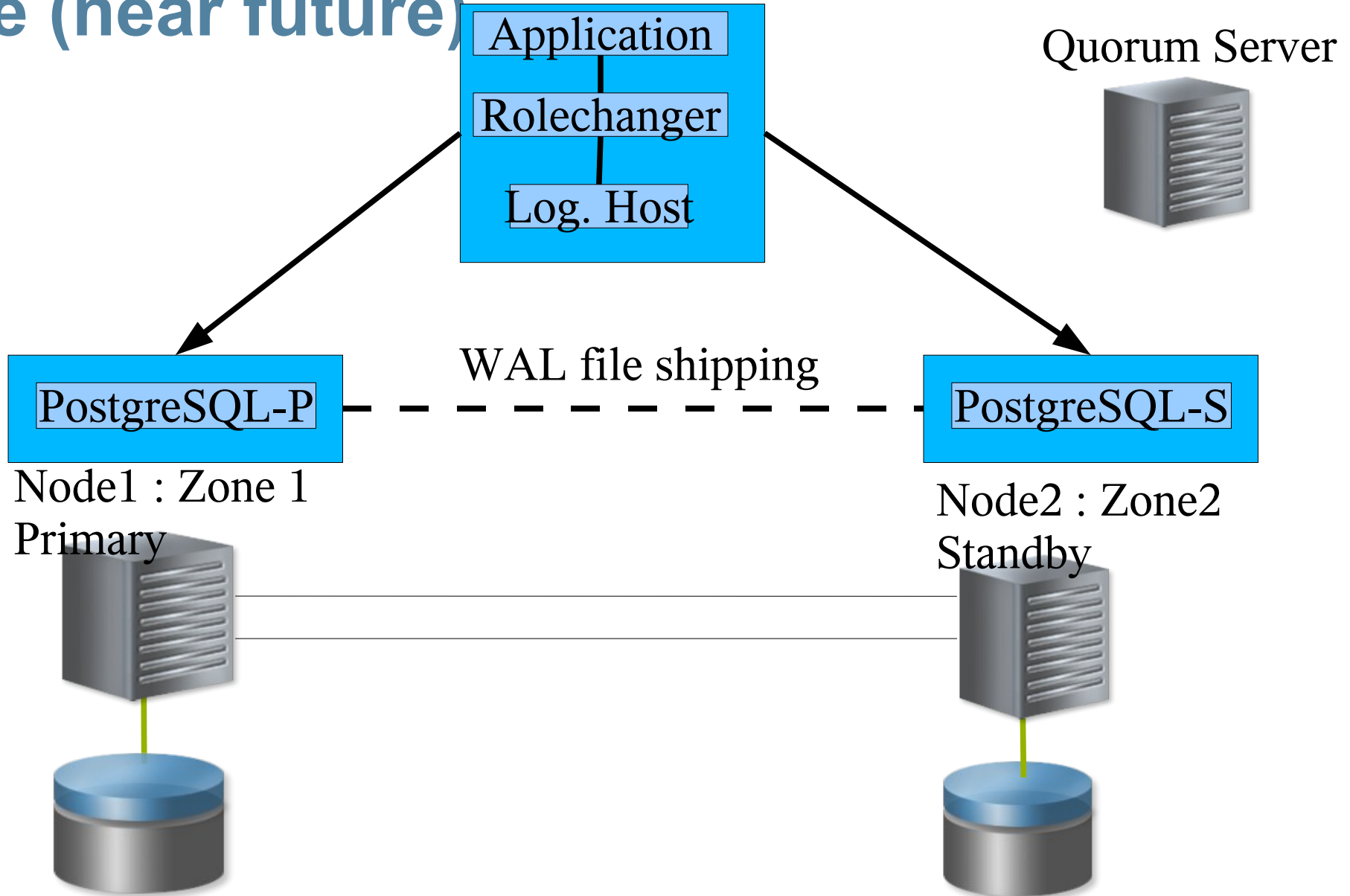
PostgreSQL (Shared Disk) Failover Zone



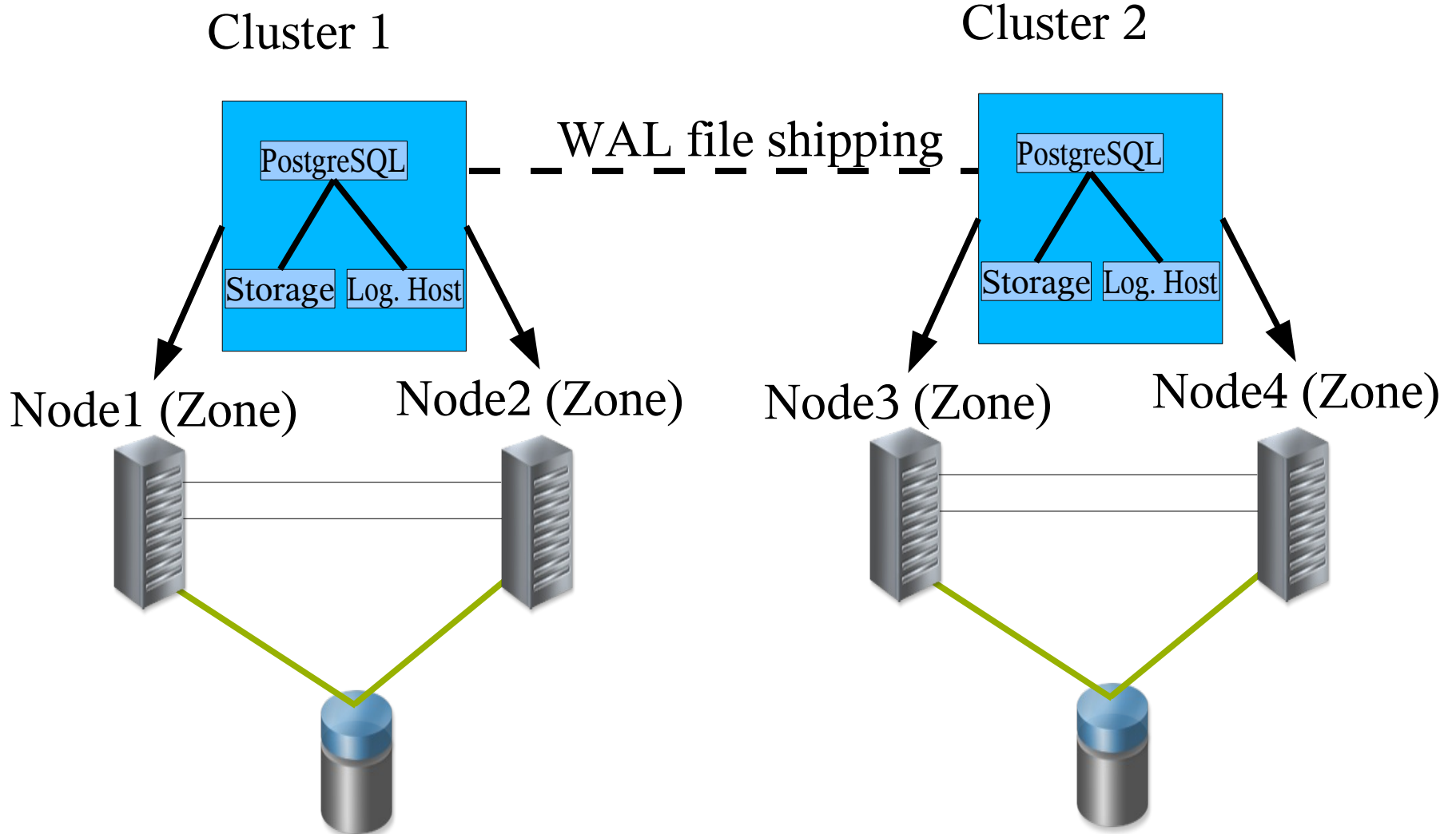
PostgreSQL (Shared Nothing) Global Zone (near future)



PostgreSQL (Shared Nothing) Local Zone (near future)

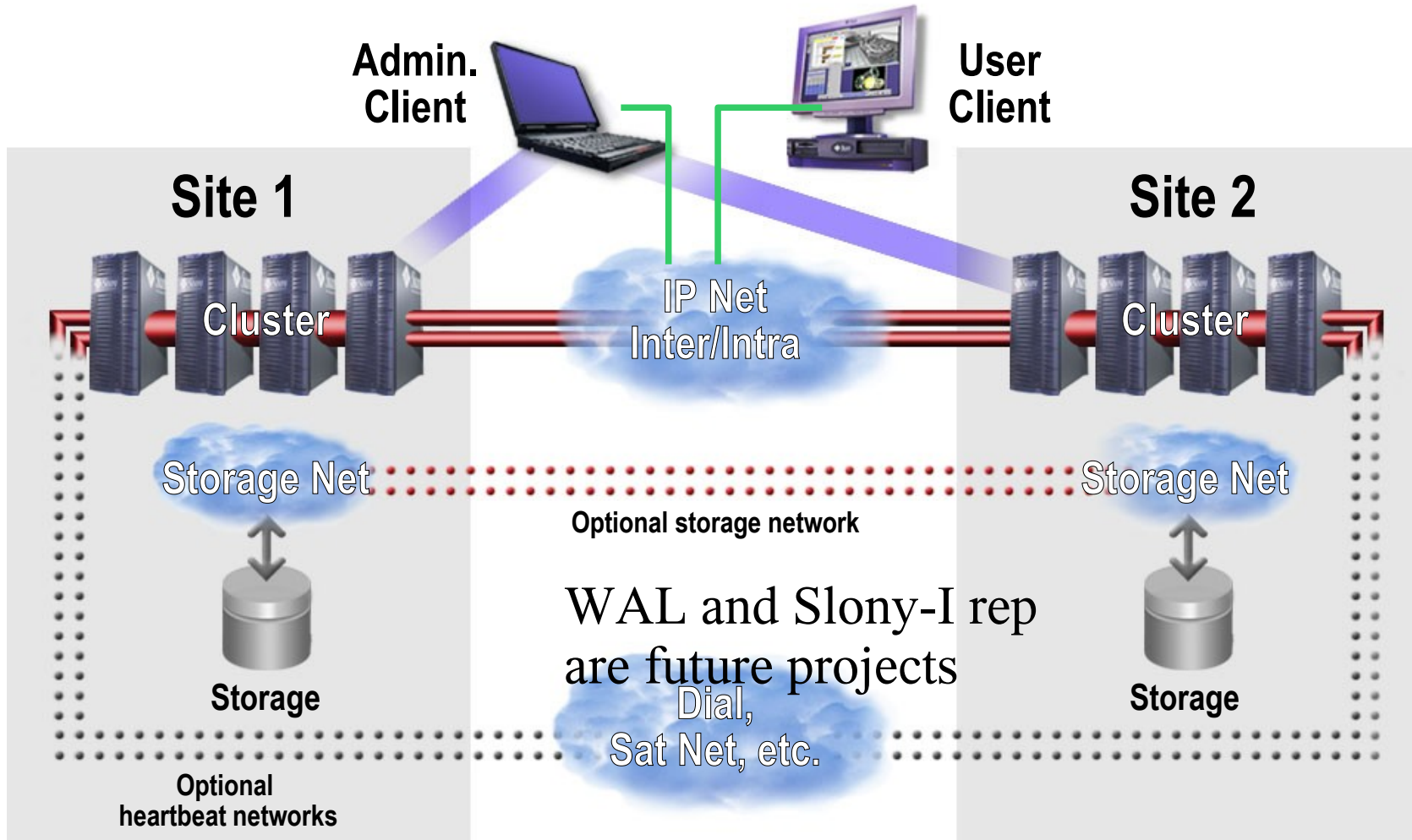


PostgreSQL (Shared Disk) Multiple Clusters (near future)



Geographic Clustering

Multi-Cluster and Multi-Site capability





Cluster in Action



PostgreSQL

- The demo will show a live cluster with PostgreSQL

open



USE



IMPROVE



EVANGELIZE

Thank you!

Detlef Ulherr
detlef.ulherr@sun.com

“open” artwork and icons by chandan:
<http://blogs.sun.com/chandan>

開
放
的
열린
مفتوح
libre
मुक्त
ಮುಕ್ತ
livre
libero
ముక్త
开放的
açık
open
nyílt
ᄒᄒᄒᄒ
πῶς
オープン
livre
ανοικτό
offen
otevřený
öppen
ОТКРЫТЫЙ
வெளிப்படை