

Planning for MC/ServiceGuard

Presented by

Gregory King

of

D-I-S RESEARCH, LTD.

Planning for high availability

MC/ServiceGuard architecture

MC/ServiceGuard packages

Planning for high availability

MC/ServiceGuard architecture

MC/ServiceGuard packages

Project management and planning

System preparation

Planning for high availability

MC/ServiceGuard architecture

MC/ServiceGuard packages

Strategies for high availability

Strategies for disaster tolerance

Hardware considerations

Network terms, concepts & strategies

LVM concepts, issues & strategies

Planning for high availability

MC/ServiceGuard architecture

MC/ServiceGuard packages

Package configuration issues

Package script issues

Application issues

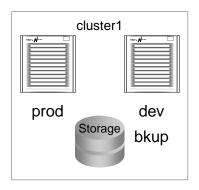
Database issues

Backing up applications managed by packages

Planning for high availability

System configuration for low maintenance

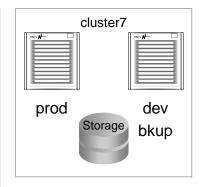
Small clusters: Poor use of resources

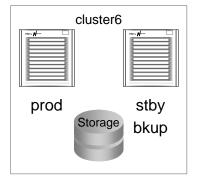


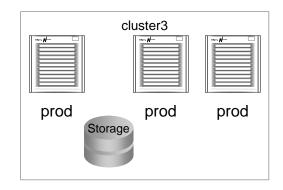
cluster2

prod

- High cost of maintenance
- High cost of training
- High cost of upgrading or patching
- Duplicated storage resources
- Duplicated computing resources
- Consistent change control process almost impossible

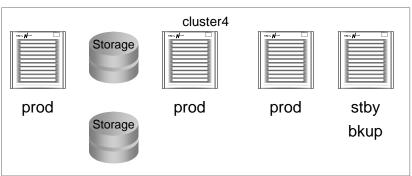


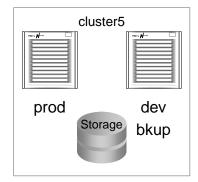




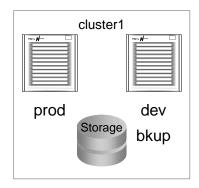
dev

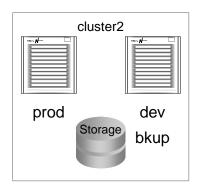
bkup





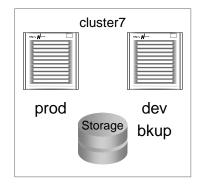
Small clusters: Poor use of resources

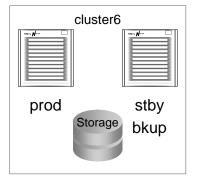


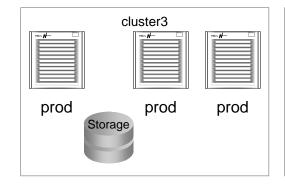


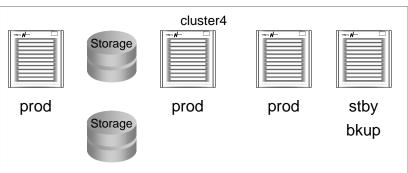
Duplicated resources:

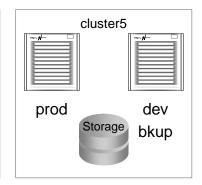
- 3 standby servers sitting idle
- 8 Separate mass storage enclosures
- 5 nodes used for backup servers
- 4 nodes used for development (mixed with production systems)



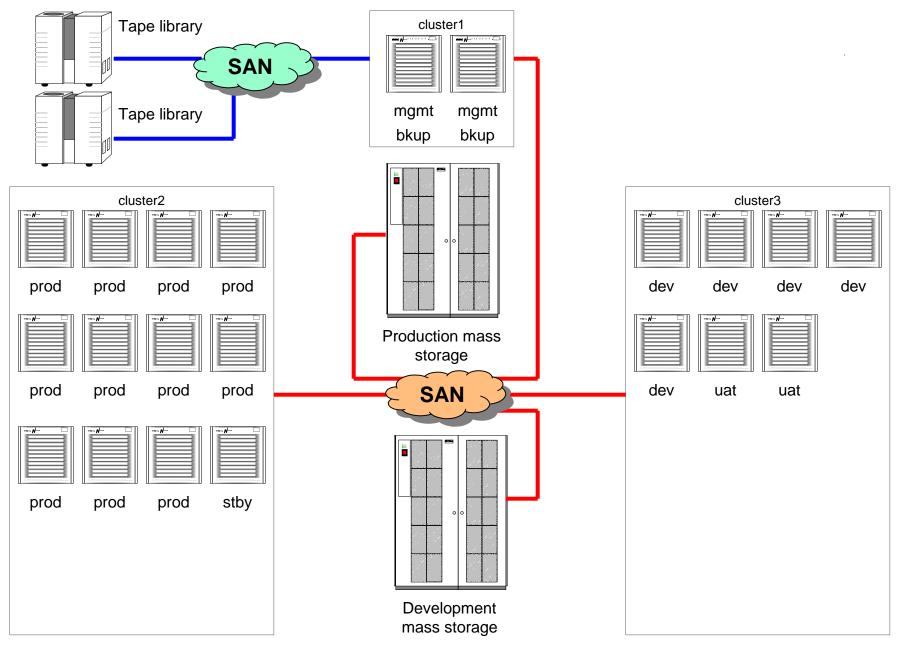








Large clusters: Scalable solutions

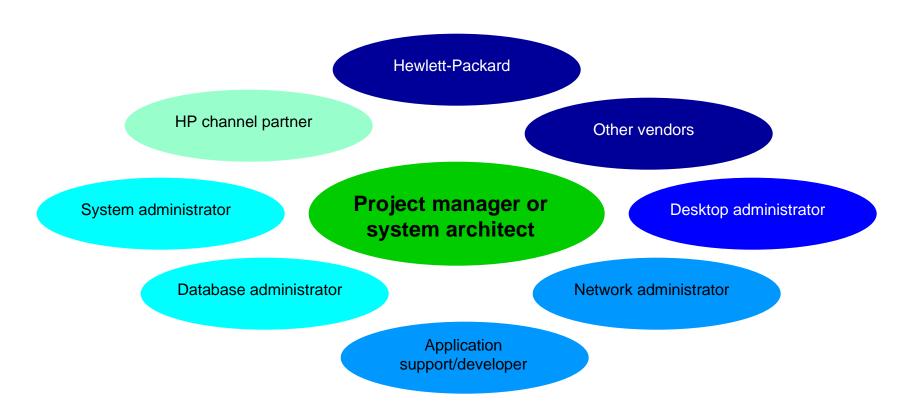


Planning: Converting an existing system

- All machines should be configured consistently or as close as possible to the same configuration for:
 - Kernel tunable parameters
 - LVM standards
 - Password files
 - Patch updates
- Only one to two releases of the OS should be installed on all production or development systems.
- No errors should be reported during boot time. Systems should be very clean.
- All systems should be documented in detail.

The implementation team

The implementation team: roles



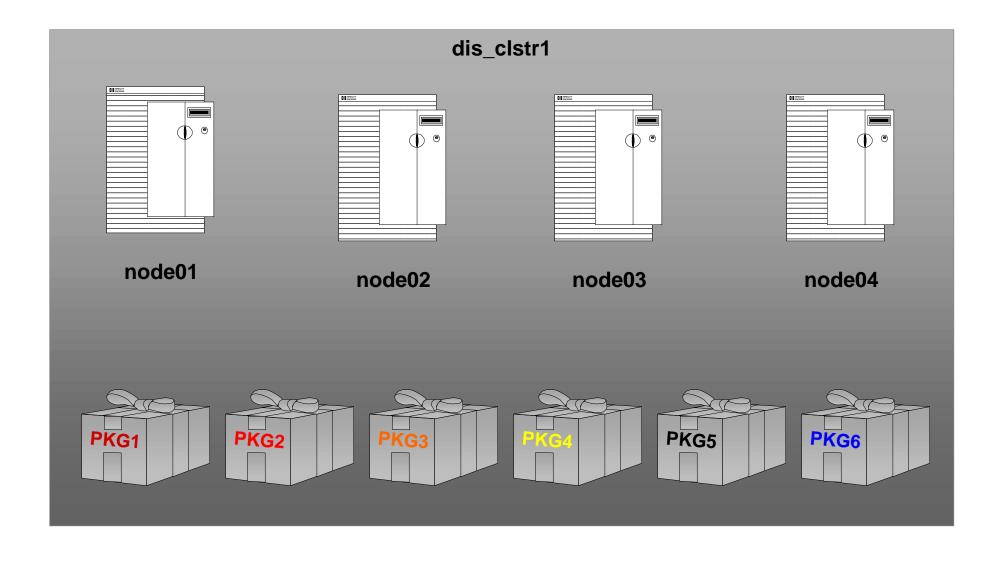
MC/ServiceGuard architecture

Basic terms & concepts

MC/ServiceGuard: Not high availability by itself

Product or component	Protection from
Good system architecture	Unplanned downtime due to OS, database or applications, quick recovery
Change system documentation	Quick recovery from failures & easier planning for change control
Change control processes	Unplanned downtime due to OS, database or applications
Good system administration habits	Unplanned downtime due to OS, database or applications, quick recovery
Firmware	Memory and CPU failure
MC/ServiceGuard: cluster manager	SPU & LAN device failure. Prevents VG from being used by multiple systems
MC/ServiceGuard: package manager	Automated start, stop and monitoring of applications and databases
MC/ServiceGuard: cluster manager	LAN device failure
MC/ServiceGuard: Ivm manager	Volume groups being active on multiple nodes simultaneously
Mirror/UX	PV failures
LVM	Disk controller failures
JFS	Allows quick, accurate recovery of file systems after an SPU failure
Advanced JFS	Dynamic resizing of file systems while in production, de-fragmentation of LV
Event Monitoring System (EMS)	Memory, CPU utilization, I/O
Process Resource Manager (PRM)	Limits or ensures CPU for certain applications

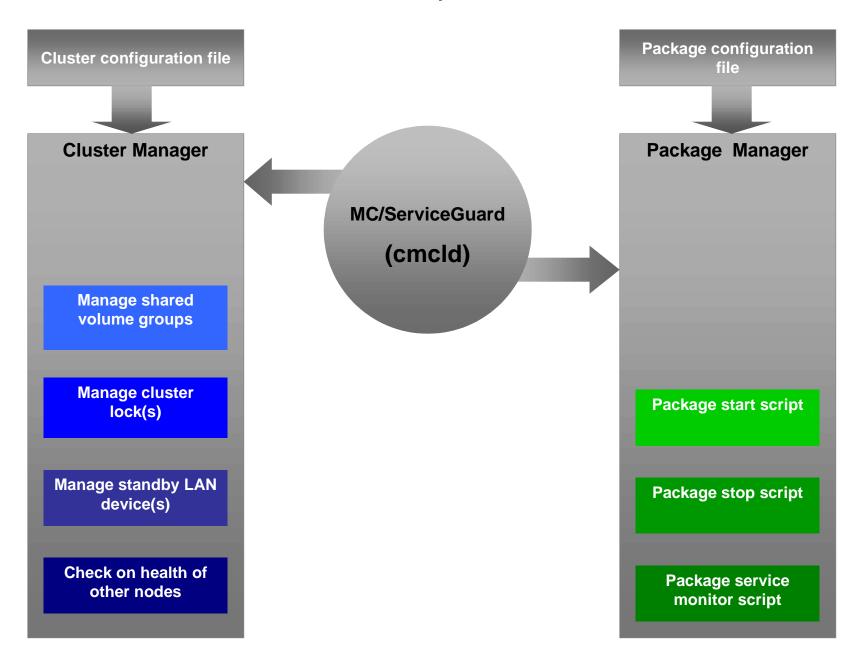
MC/ServiceGuard: Cluster



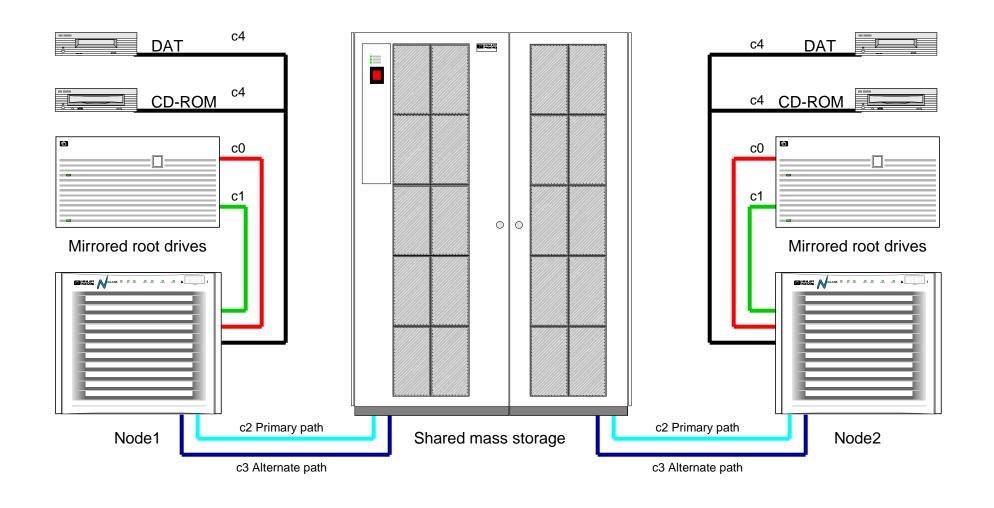
MC/ServiceGuard: Package

Components of a package	Example
Package Name	pkg1
Managed resources IP address(es) Volume groups Logical volumes Managed apps start the application stop the application monitor the application	10.10.10.4 10.10.11.4 VG01 /dev/VG01/Ivol1 /dir1 VG02 /dev/VG02/Ivol1 /dir2 Oracle data for SID1 (resides on VG02) Oracle data for SID2 (resides on VG02) Oracle TNS listener (for SID1) Oracle TNS listener (for SID2) Sybase server 1 app1 (resides on VG00) app2 (resides on VG00)
	app1 (resides on VG00)

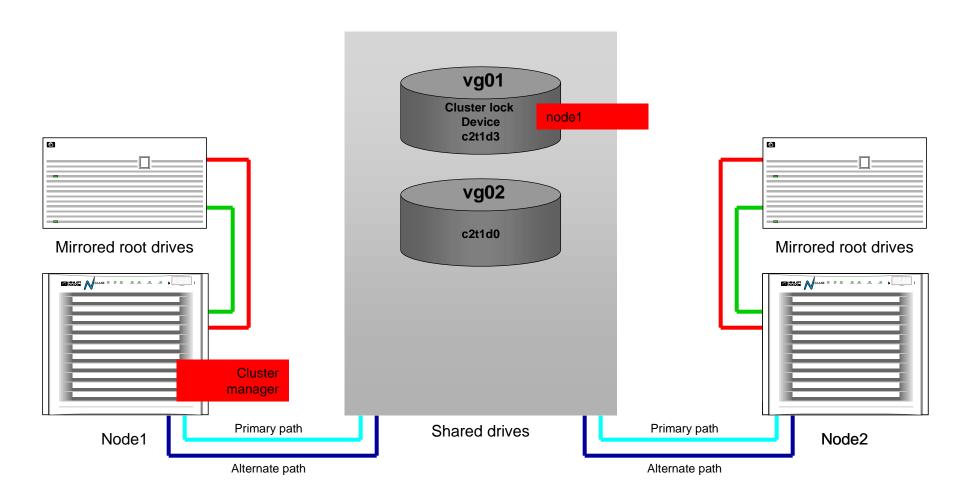
MC/ServiceGuard: Responsibilies of daemon



MC/ServiceGuard: Basic hardware



MC/ServiceGuard: Cluster lock



- •Required for two node clusters
- •Recommended in three node clusters
- •Can use any PV in any shared volume group

Network terms, concepts & strategies

Primary LAN

Standby LAN

Heartbeat LAN

Serial heartbeat line

Bridged network

Stationary IP

Re-locatable IP

Heartbeat IP

Automatic port aggregation

Primary LAN

Standby LAN

Heartbeat LAN

Serial heartbeat line

Bridged network

Stationary IP

Re-locatable IP

Heartbeat IP

Automatic port aggregation

A LAN device that is up and has an IP address and netmask assigned to it.

Configured during the node's boot up process. All information needed to configure the interface is kept in /etc/rc.config.d/netconf

Can run the following protocols: 10Mb, 100Mb, 1Gb, FDDI or SNAP

There can be multiple primary interfaces, but each one needs to be on a different subnet on a separate bridged network.

Primary LAN

Standby LAN

Heartbeat LAN

Serial heartbeat line

Bridged network

Stationary IP

Re-locatable IP

Heartbeat IP

Automatic port aggregation

A LAN device that is up, but has no IP address or netmask assigned to it.

Configured during the node's boot up process. Although no IP is assigned during boot, the interface entry is kept in /etc/rc.config.d/netconf

MC/ServiceGuard LAN manager will move all stationary and re-locatable IP addresses from a failed LAN device to the standby LAN device.

The primary and standby LAN devices should both be of the same type and should both be on the same subnet via a bridged network.

Primary LAN

Standby LAN

Heartbeat LAN

Serial heartbeat line

Bridged network

Stationary IP

Re-locatable IP

Heartbeat IP

Automatic port aggregation

A dedicated LAN used to monitor the health of all nodes in a cluster. This is the method used to detect a failed node.

The heartbeat LAN should be isolated from the normal network infrastructure.

Any network device that is supported by MC/ServiceGuard can be used as a dedicated heartbeat LAN.

Use a single twisted-pair cable to connect two node clusters.

Use a hub to connect three or more nodes.

Primary LAN

Standby LAN

Heartbeat LAN

Serial heartbeat line

Bridged network

Stationary IP

Re-locatable IP

Heartbeat IP

Automatic port aggregation

A standby heartbeat connection used by MC/ServiceGuard in the event the dedicated heartbeat LAN fails due to network saturation.

Only supported in a two node cluster.

Will only function long enough to allow a two node cluster to reform and elect another cluster manager.

Primary LAN

Standby LAN

Heartbeat LAN

Serial heartbeat line

Bridged network

Stationary IP

Re-locatable IP

Heartbeat IP

Automatic port aggregation

A primary interface and it's supporting standby interface are connected to the same subnet via different cables to separate hubs, concentrators, bridges or switches.

Primary LAN

Standby LAN

Heartbeat LAN

Serial heartbeat line

Bridged network

Stationary IP

Re-locatable IP

Heartbeat IP

Automatic port aggregation

An IP address that is permanently associated with a single LAN device.

Configured at boot time using /etc/rc.config.d/netconf

With the exception of standby LAN devices, every LAN device in a node will have a stationary IP address.

Primary LAN

Standby LAN

Heartbeat LAN

Serial heartbeat line

Bridged network

Stationary IP

Re-locatable IP

Heartbeat IP

Automatic port aggregation

An IP address that can move from one LAN device to another. Packages can relocate

Packages bind and unbind IP addresses to one or more LAN devices on a machine. IP addresses on multiple LAN devices must be on separate subnets.

Packages can relocate IP addresses from LAN devices on one node to a LAN device on another node.

ARP cache update preformed every time an IP is bound to, or unbound from a LAN device

Primary LAN

Standby LAN

Heartbeat LAN

Serial heartbeat line

Bridged network

Stationary IP

Re-locatable IP

Heartbeat IP

Automatic port aggregation

A stationary IP address assigned to LAN device for the purpose providing a heartbeat connection between multiple nodes.

Any active LAN device can support IP addresses for both data and heartbeat. This means multiple LAN interfaces will be configured to provide a heartbeat in a cluster.

Primary LAN

Standby LAN

Heartbeat LAN

Serial heartbeat line

Bridged network

Stationary IP

Re-locatable IP

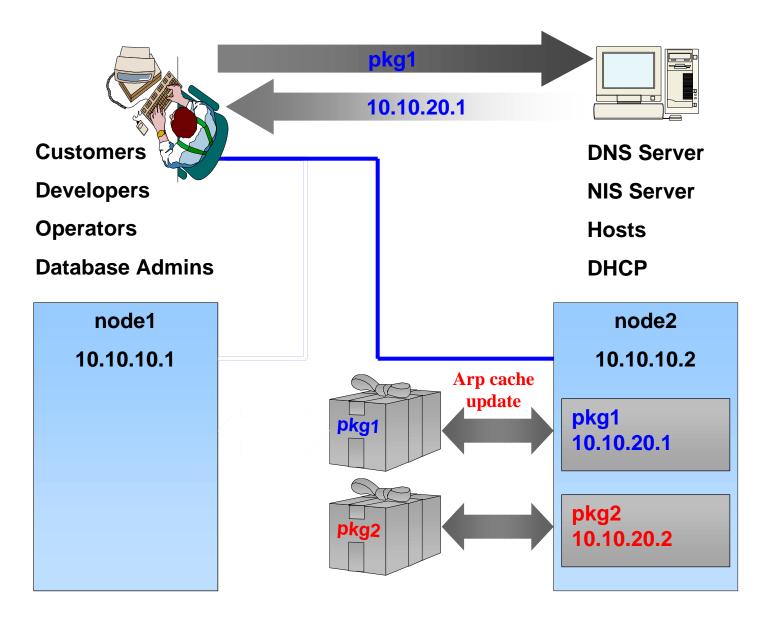
Heartbeat IP

Automatic port aggregation

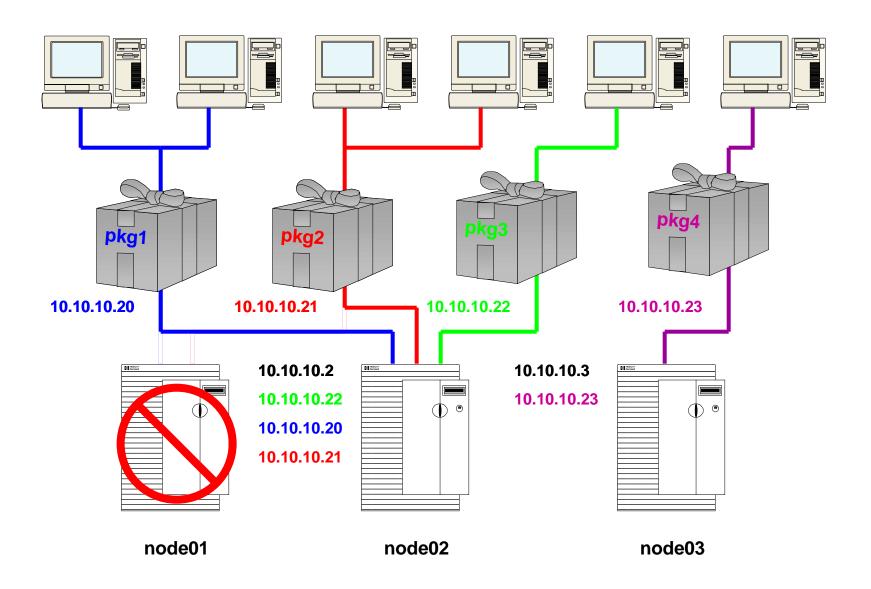
Allows a single IP address to be bound to multiple 100Mb LAN interfaces to provide increased network bandwidth.

This product is not bundled with MC/ServiceGuard.

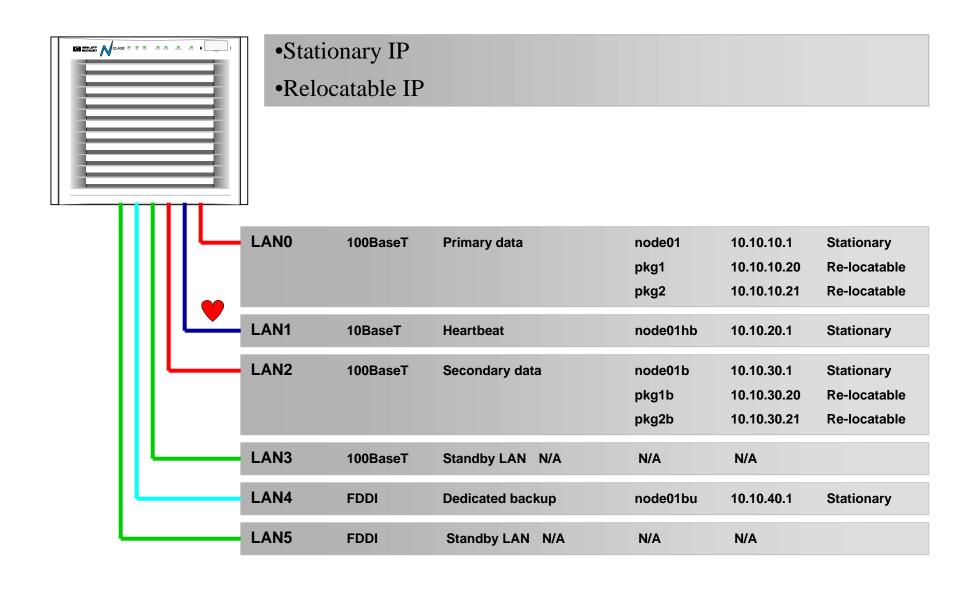
IP Addresses: The key to connectivity



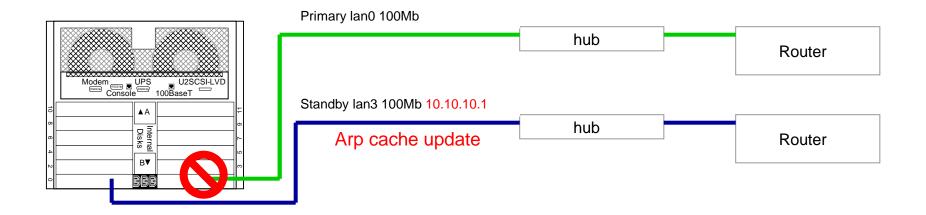
How Packages Keep People Working



Network names, addresses & devices



Standby LAN: LAN Failover

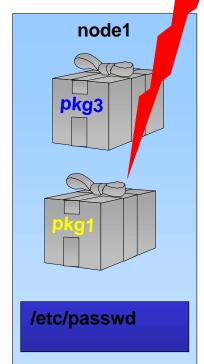


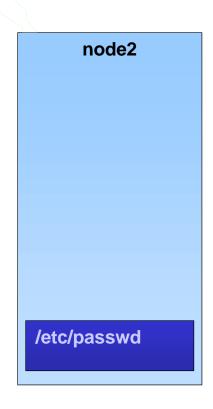
- •Must be same type LAN device as primary
- •Must be on same subnet as the device it is supporting
- Sub-second fail-over
- Both stationary and relocatable IP addresses are moved
- •Arp cache update initiated for any fail-over
- Automatic failback once primary is back on-line

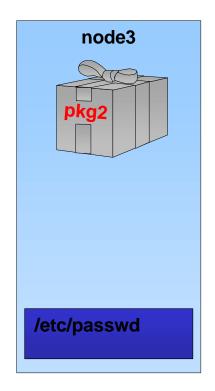
NIS: When to Use NIS



- Use NIS if many users log onto one or more packages
- •Usually manage only passwd & group
- •Ensure the services file is found first using nsswitch.conf

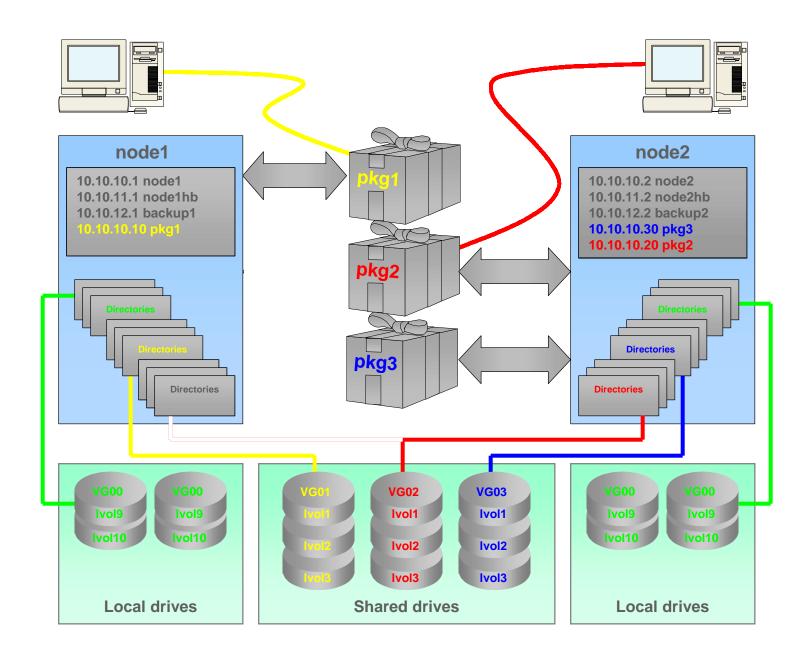




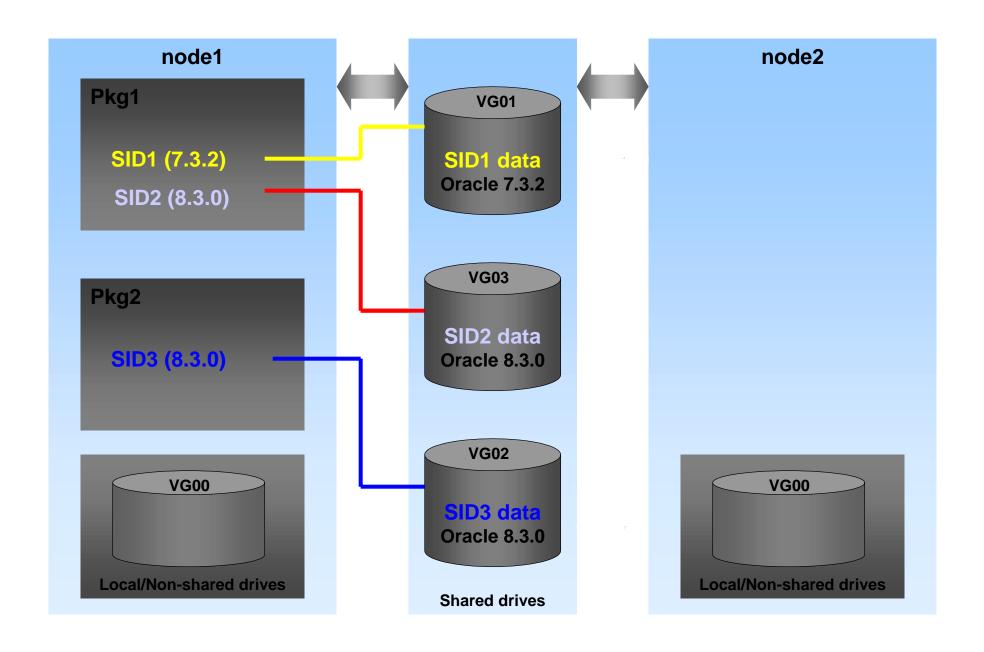


LVM concepts, issues & strategies

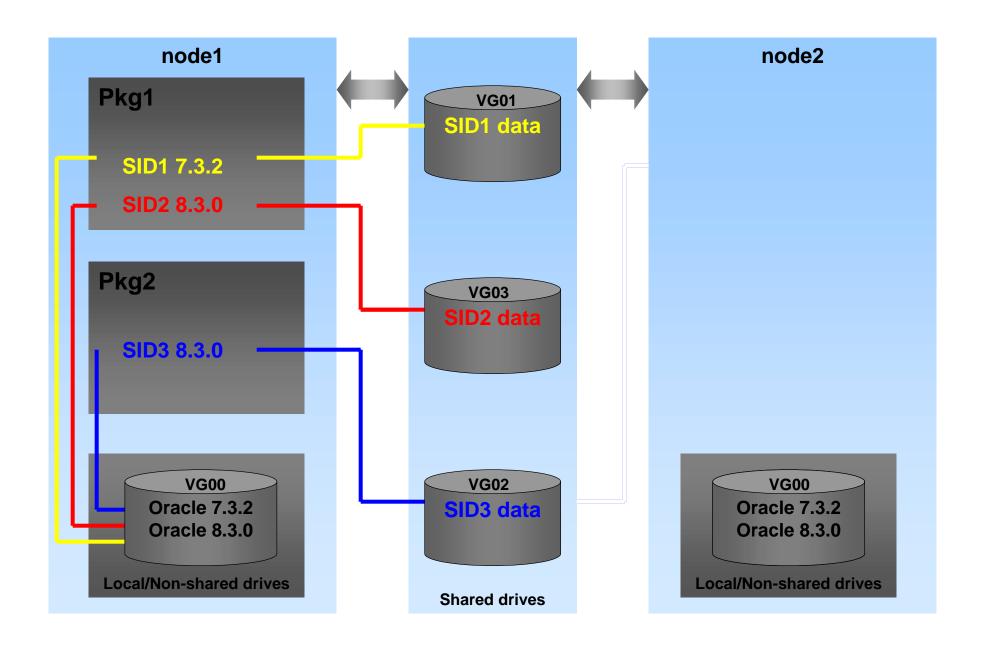
LVM: The mechanics of moving a package



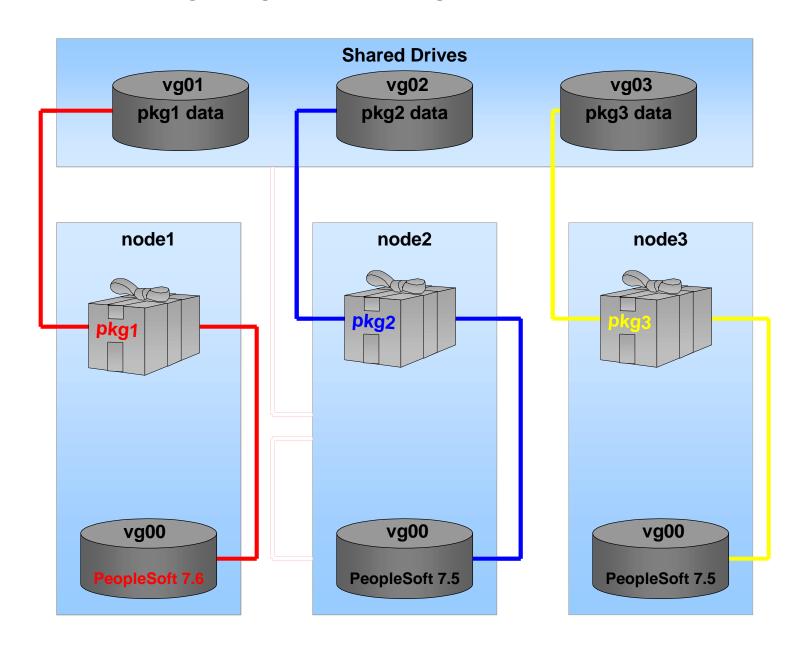
LVM: Binaries on shared drives



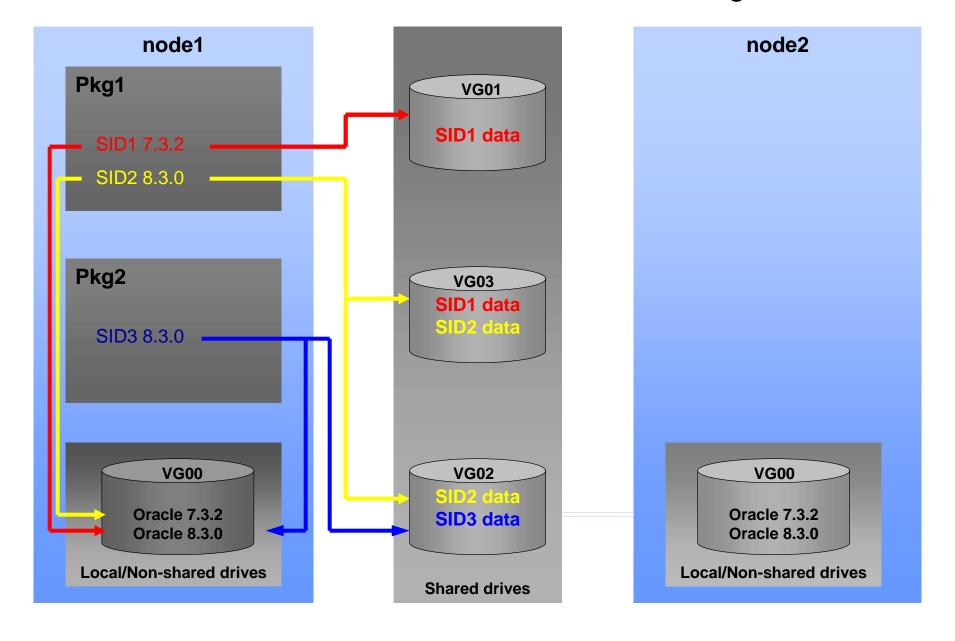
LVM: Binaries on local drives



LVM: Rolling upgrade using binaries on local drives



LVM: Don't Share Data Between Packages



NFS: Using NFS with Packages

