Planning for MC/ServiceGuard

Presented by

Gregory King

of

D•I•S Research, Ltd.

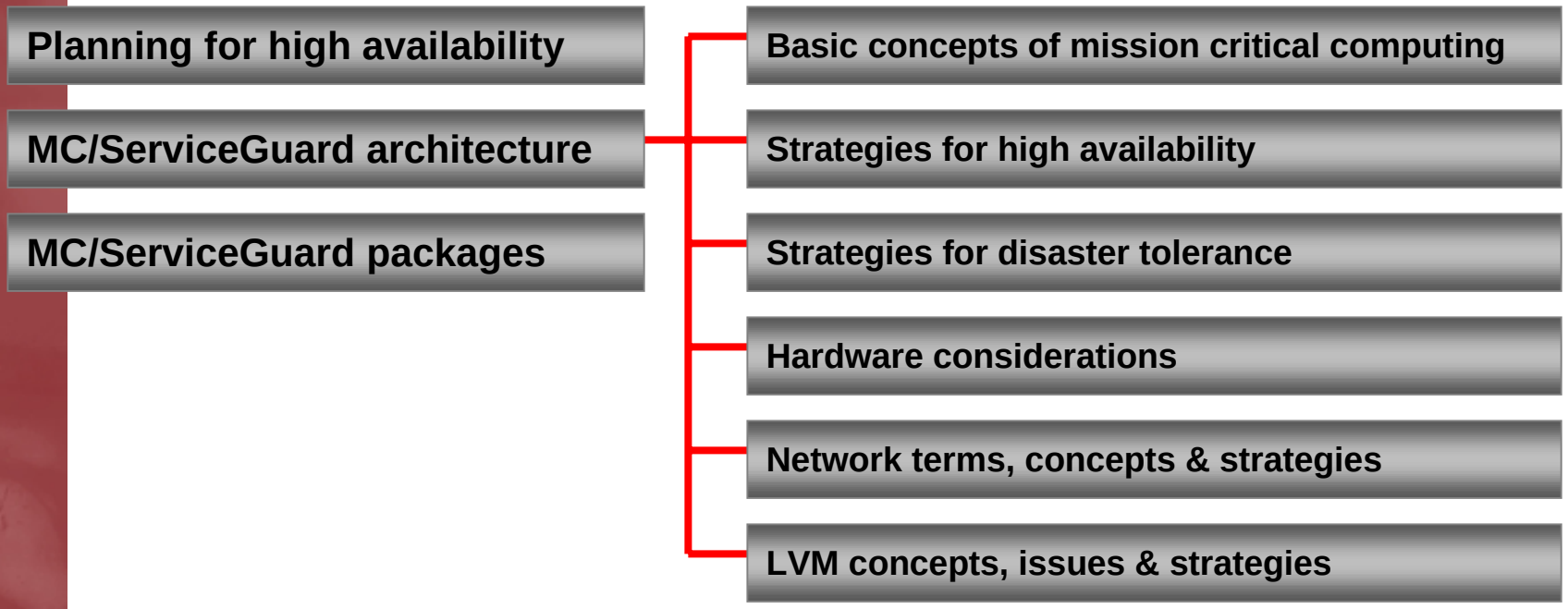
Training Topics

- Planning for high availability
- MC/ServiceGuard architecture
- MC/ServiceGuard packages

Training Topics

- Planning for high availability
- MC/ServiceGuard architecture
- MC/ServiceGuard packages
- System configuration for low maintenance
- The implementation team
- Project management and planning
- System preparation

Training Topics



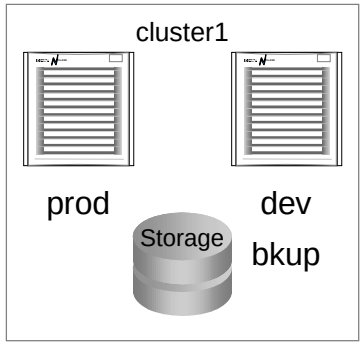
Training Topics

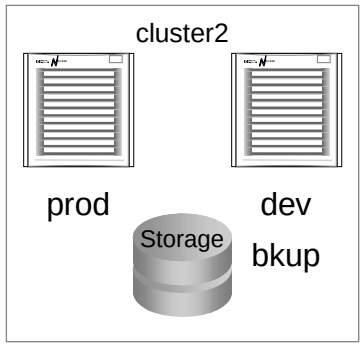
- Planning for high availability
- MC/ServiceGuard architecture
- MC/ServiceGuard packages
  - Basic package concepts
  - Package configuration issues
  - Package script issues
  - Application issues
  - Database issues
  - Backing up applications managed by packages

Planning for high availability

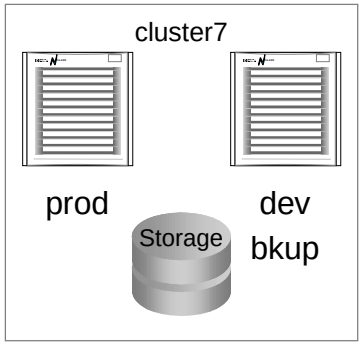
System configuration for low maintenance

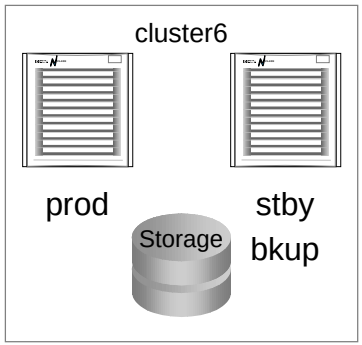
Small clusters: Poor use of resources

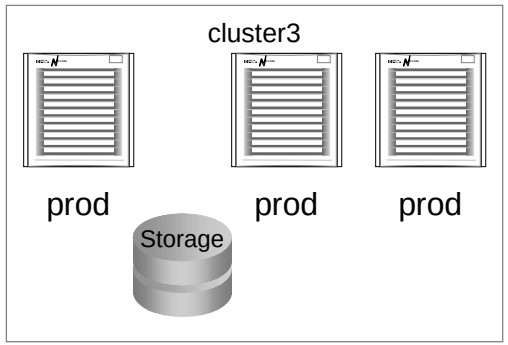


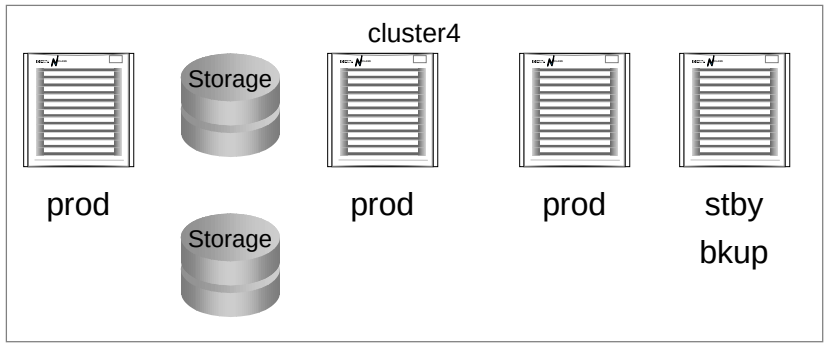


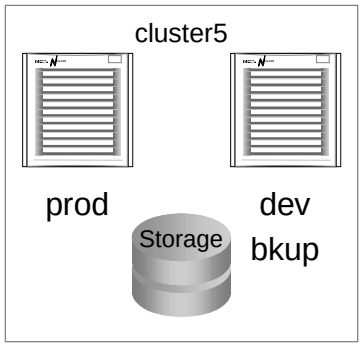
- **High cost of maintenance**
- **High cost of training**
- **High cost of upgrading or patching**
- **Duplicated storage resources**
- **Duplicated computing resources**
- **Consistent change control process almost impossible**



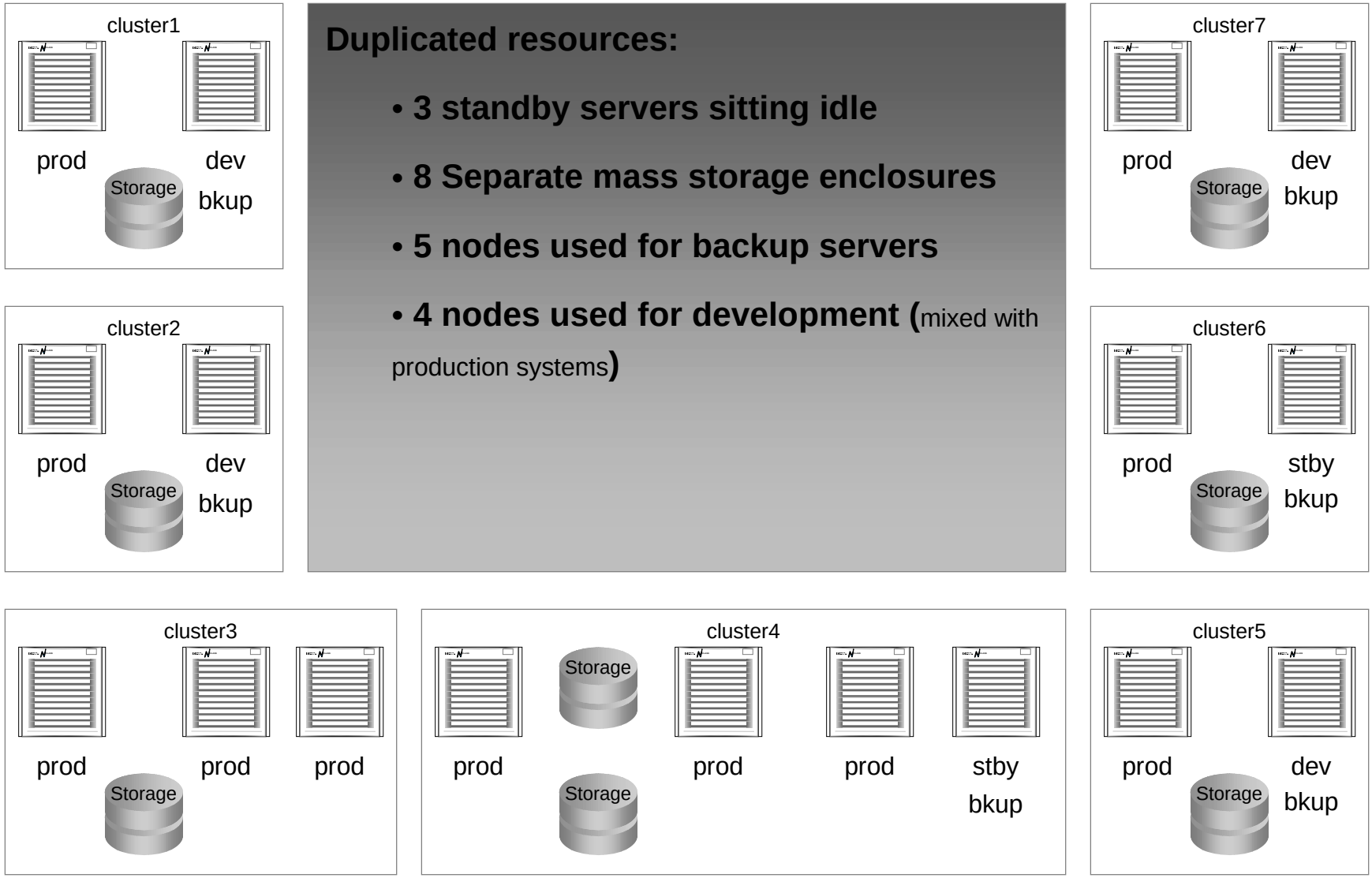




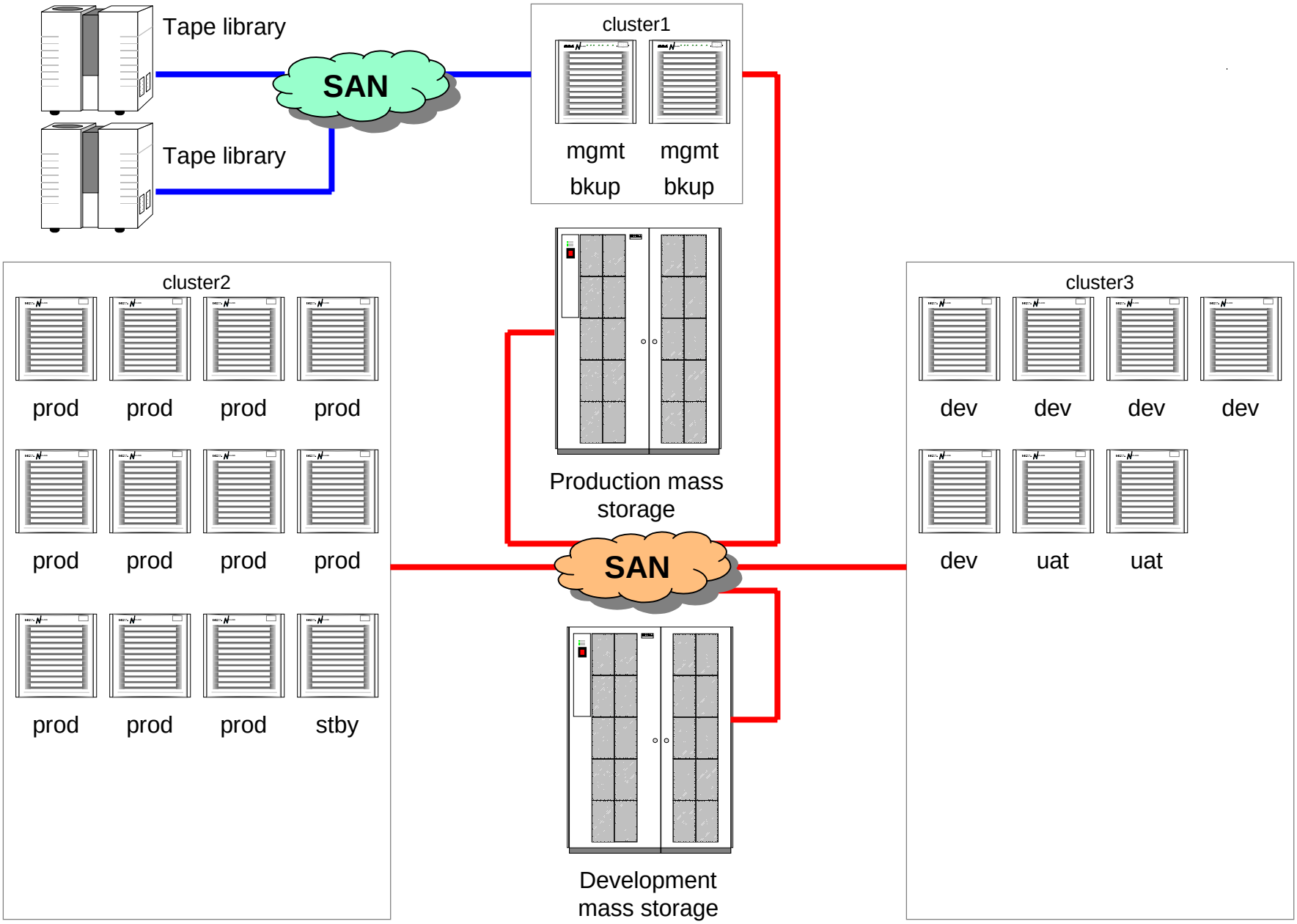




Small clusters: Poor use of resources



Large clusters: Scalable solutions

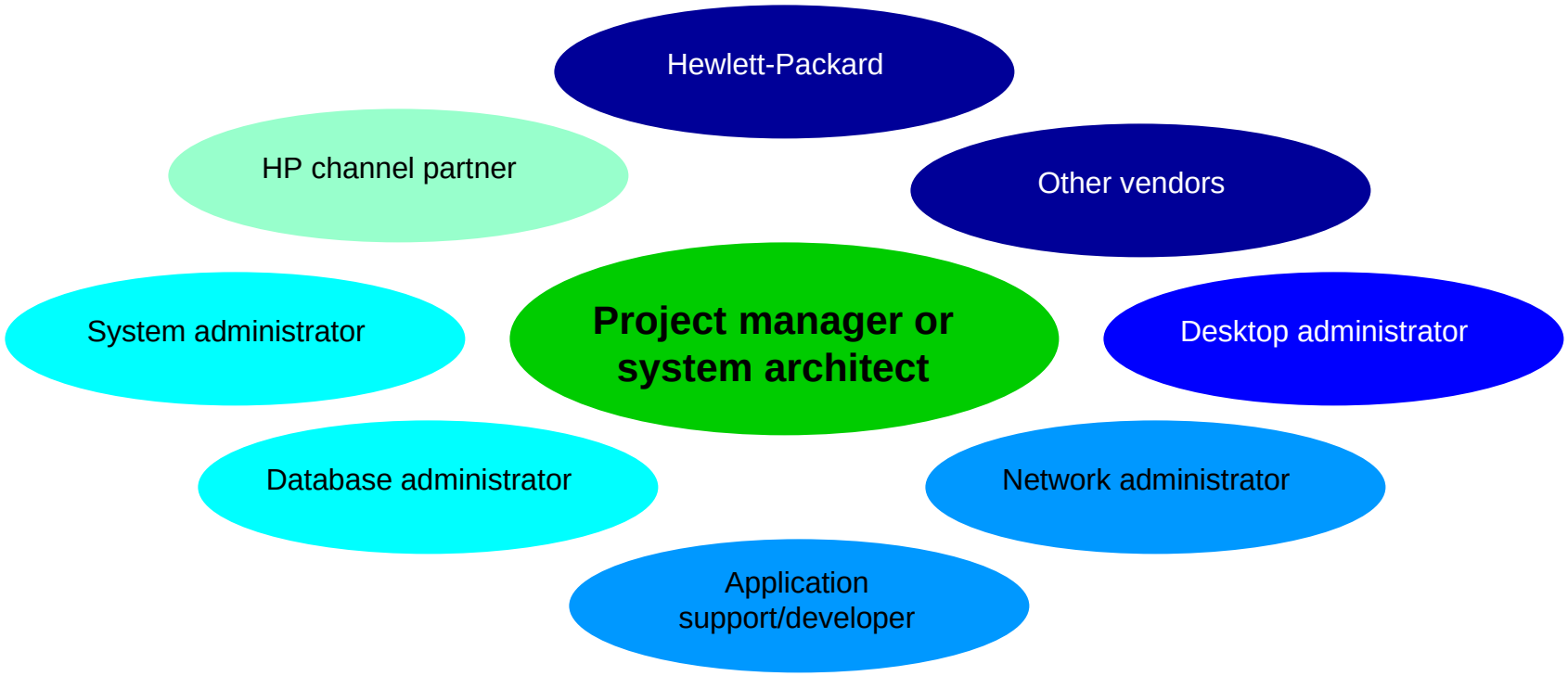


Planning: Converting an existing system

- 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



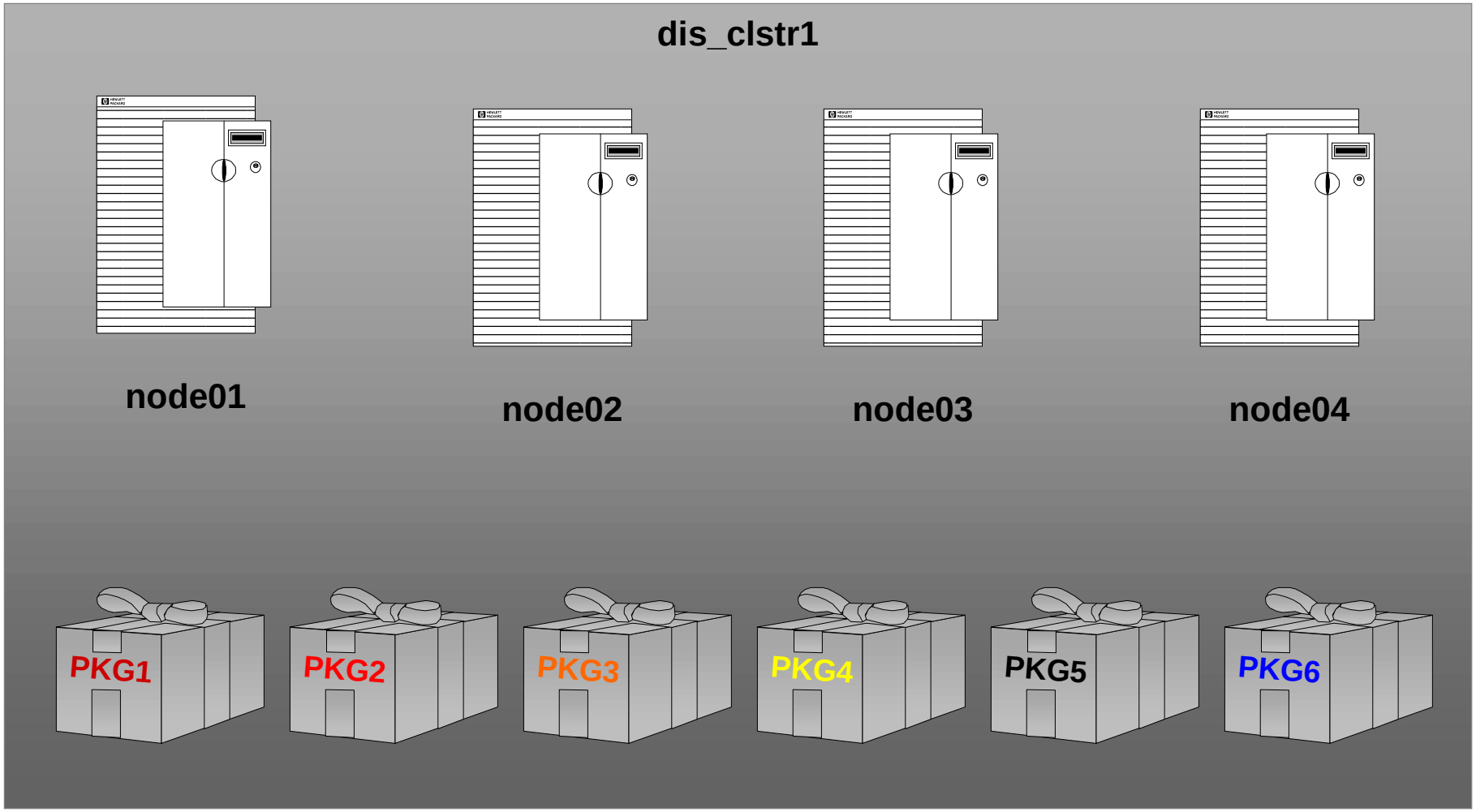
MC/ServiceGuard architecture

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: lvm 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

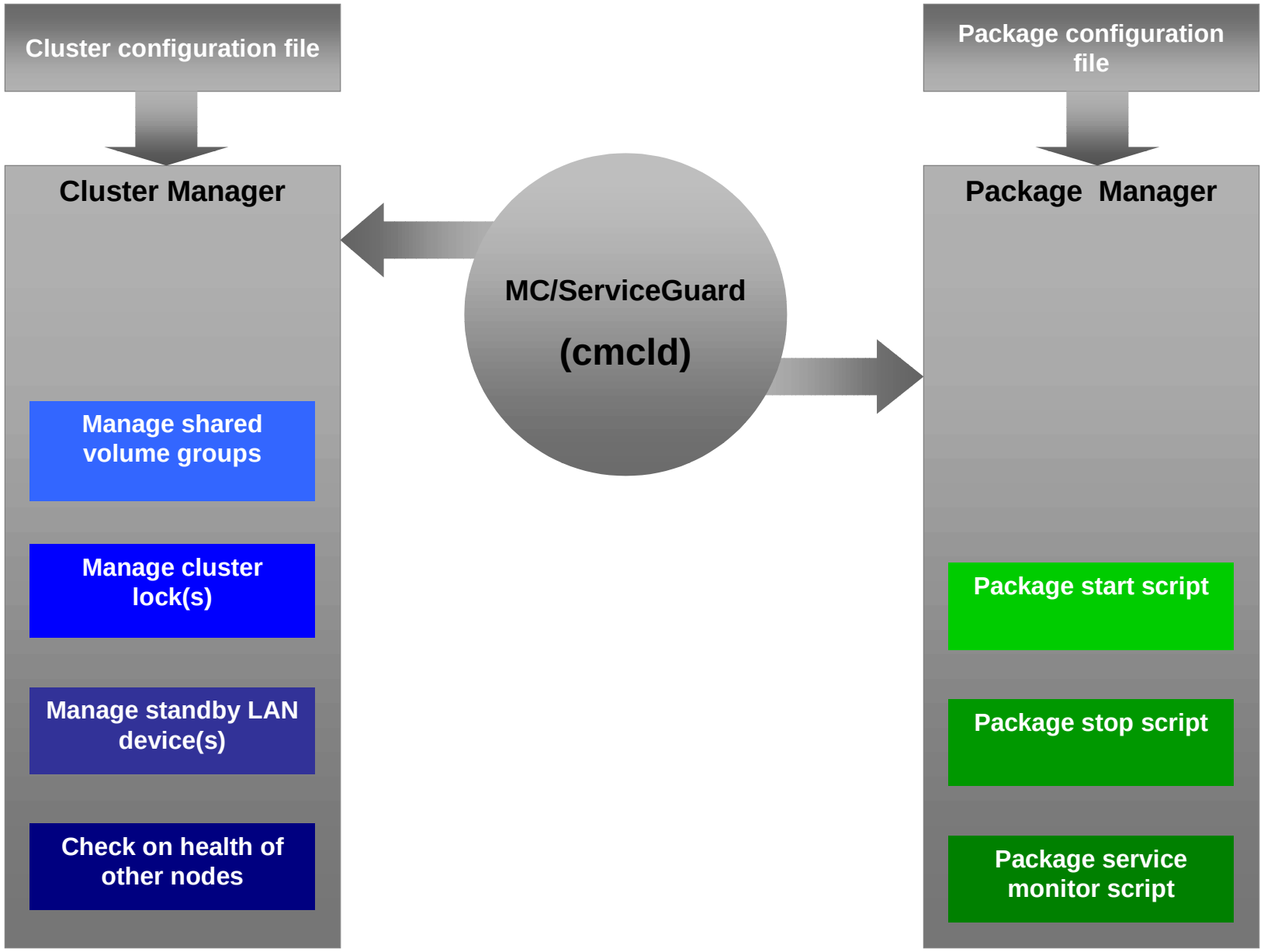
MC/ServiceGuard: Cluster



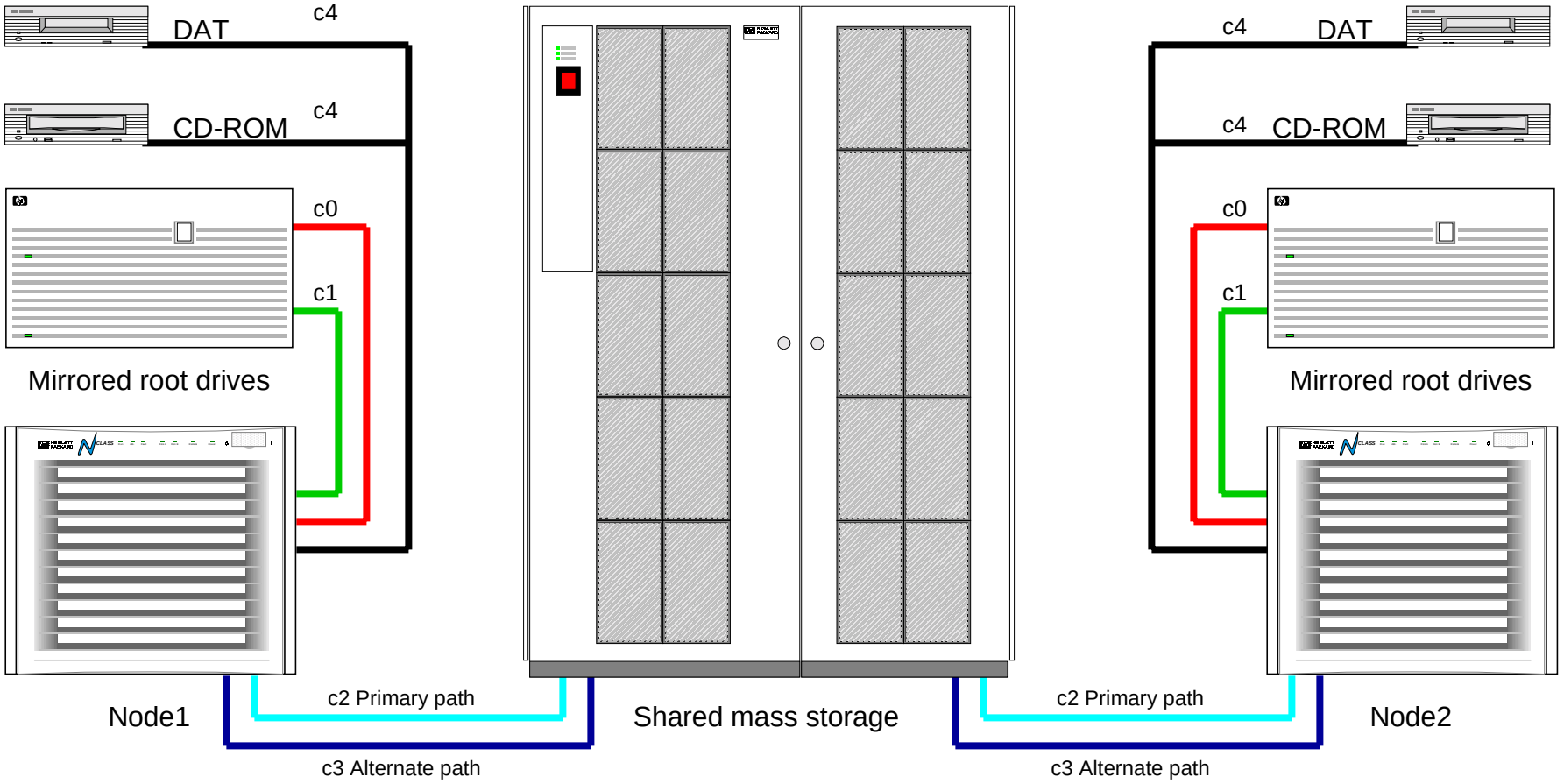
MC/ServiceGuard: Package

Components of a package	Example
Package Name	pkg1
Managed resources IP address(es) Volume groups Logical volumes	10.10.10.4 10.10.11.4 VG01 /dev/VG01/lvol1 /dir1 VG02 /dev/VG02/lvol1 /dir2
Managed apps start the application stop the application monitor the application	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)

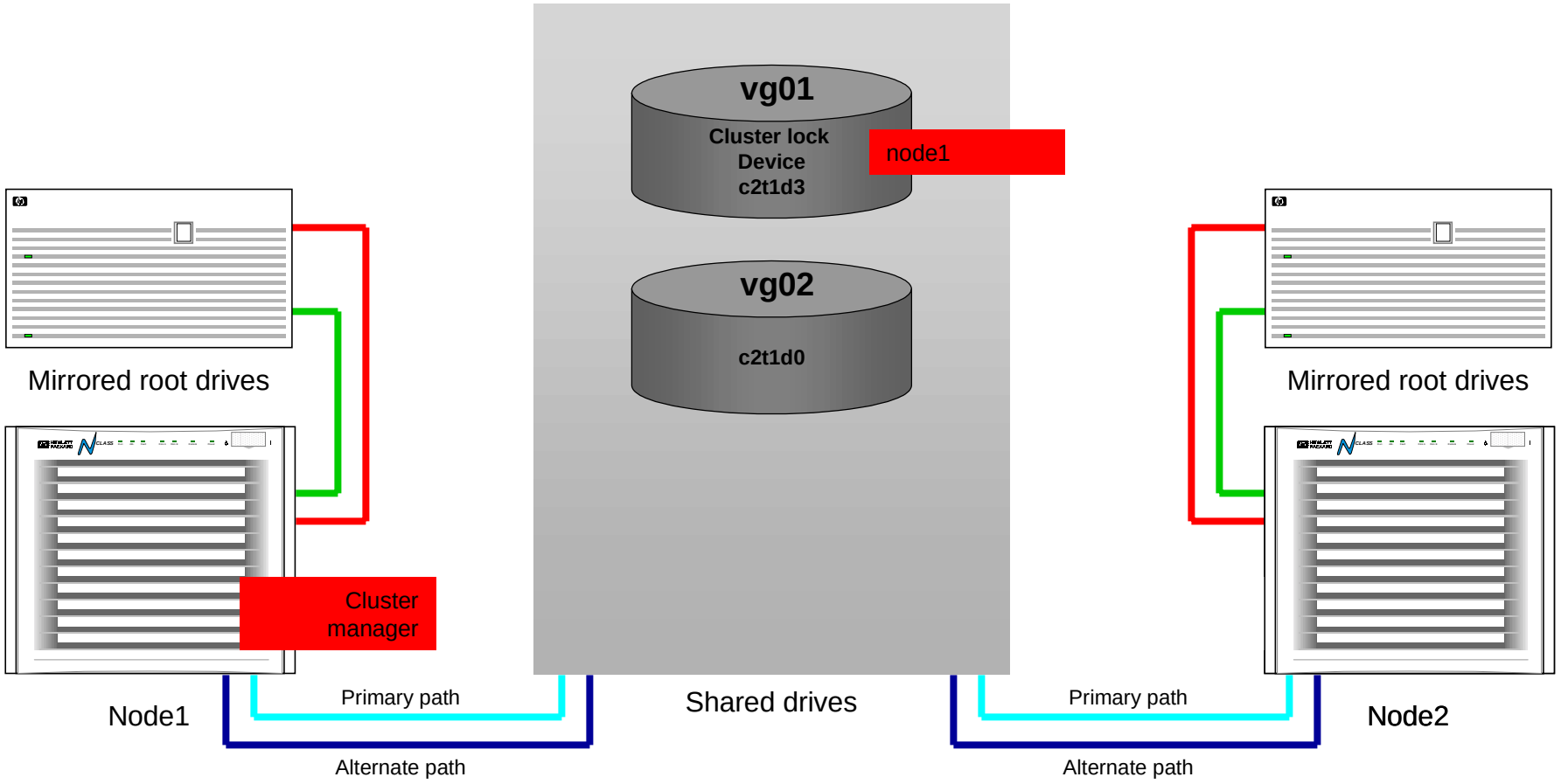
MC/ServiceGuard: Responsibilies of daemon



MC/ServiceGuard: Basic hardware



MC/ServiceGuard: Cluster lock



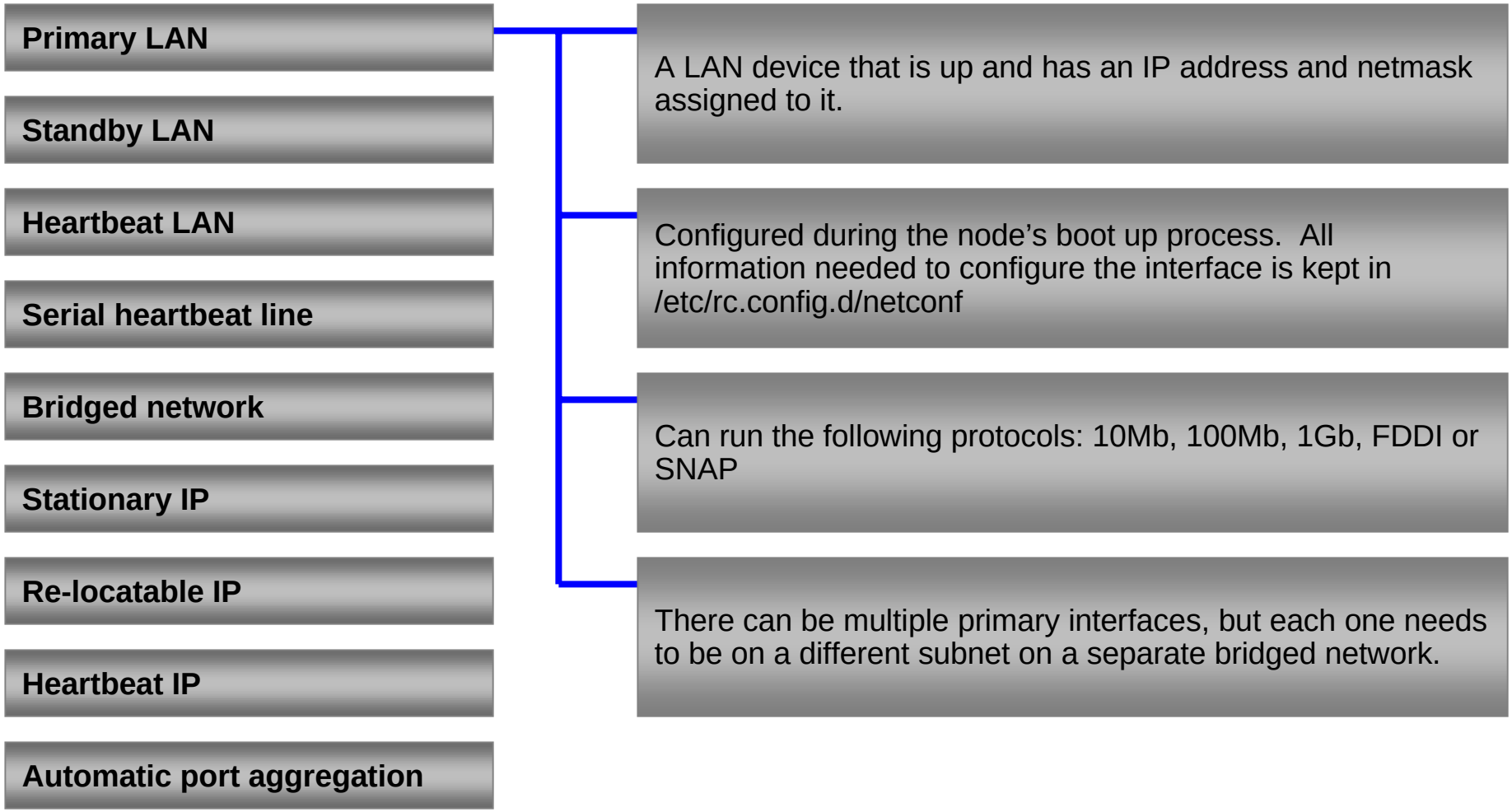
- **Required for two node clusters**
- **Recommended in three node clusters**
- **Can use any PV in any shared volume group**

Network terms, concepts & strategies

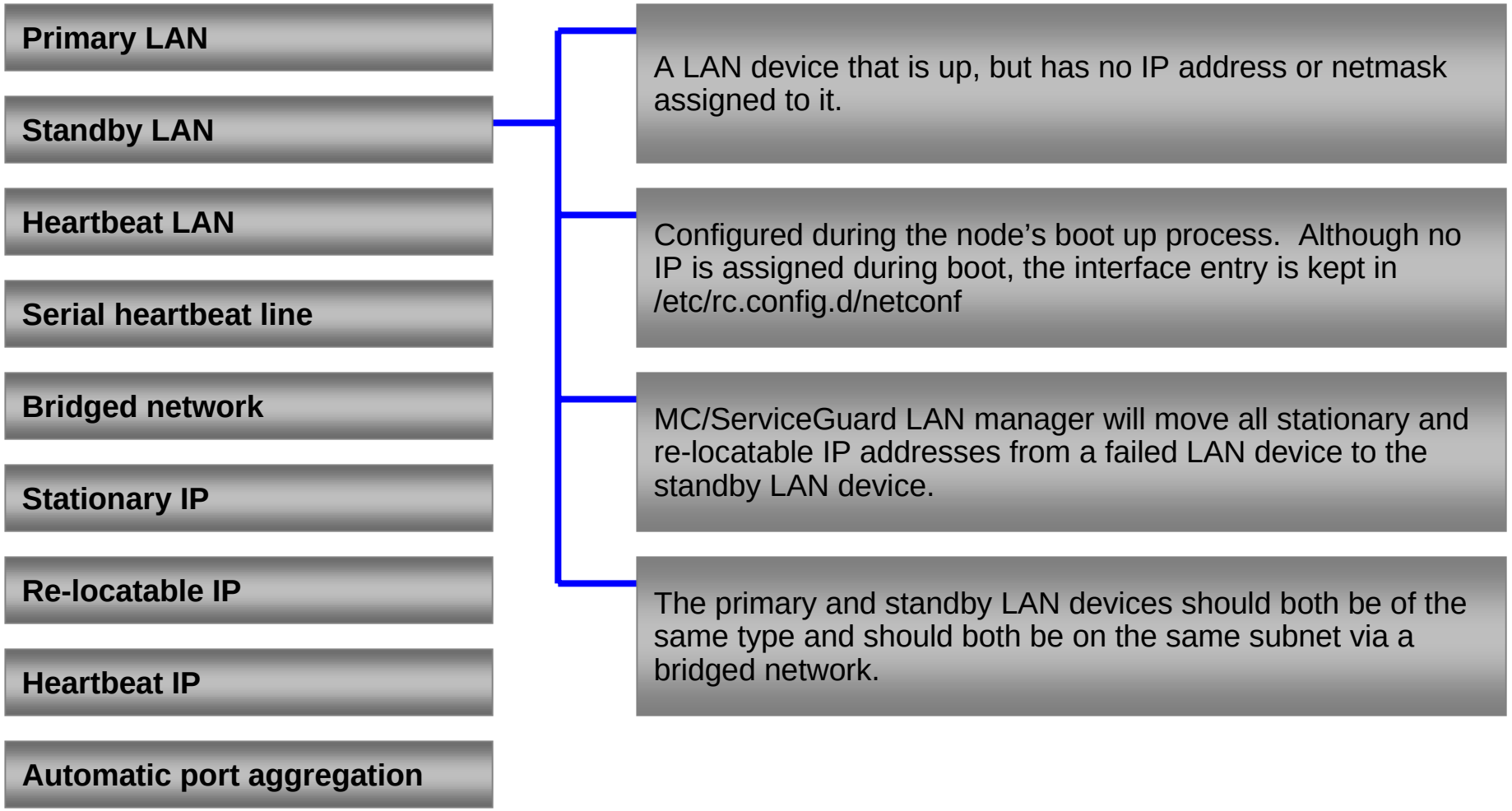
Network Terminology

- **Primary LAN**
- **Standby LAN**
- **Heartbeat LAN**
- **Serial heartbeat line**
- **Bridged network**
- **Stationary IP**
- **Re-locatable IP**
- **Heartbeat IP**
- **Automatic port aggregation**

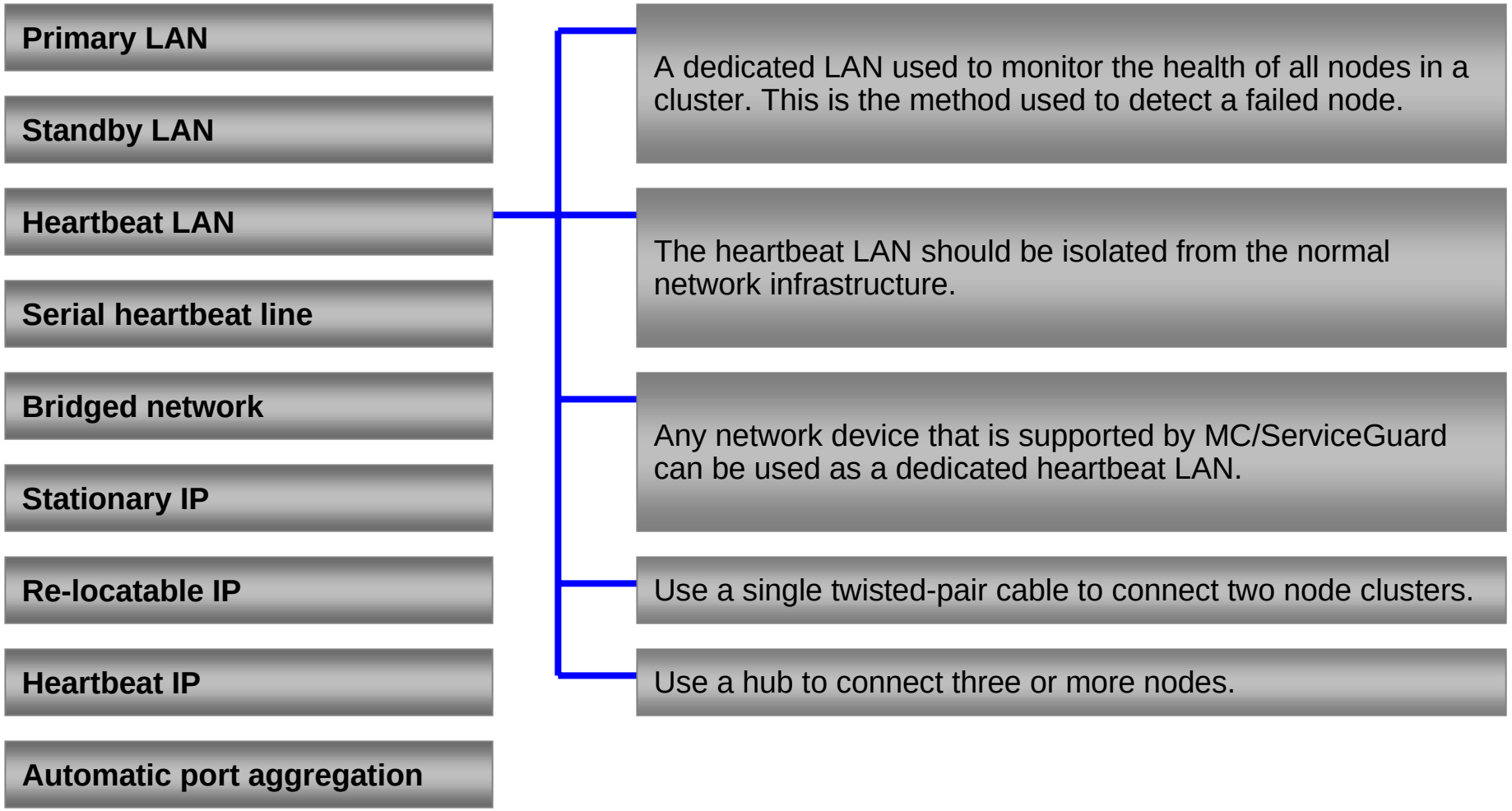
Network Terminology



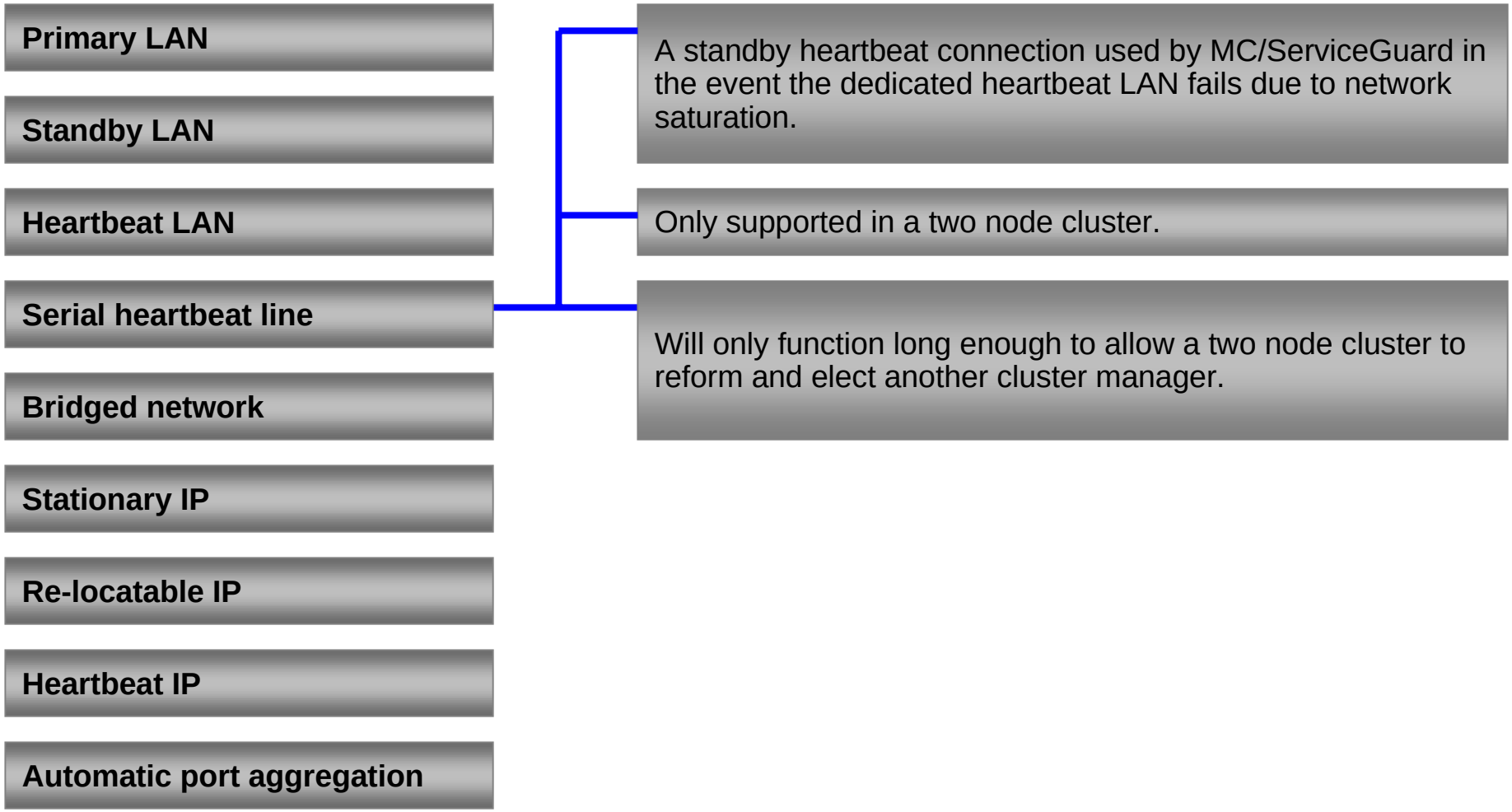
Network Terminology



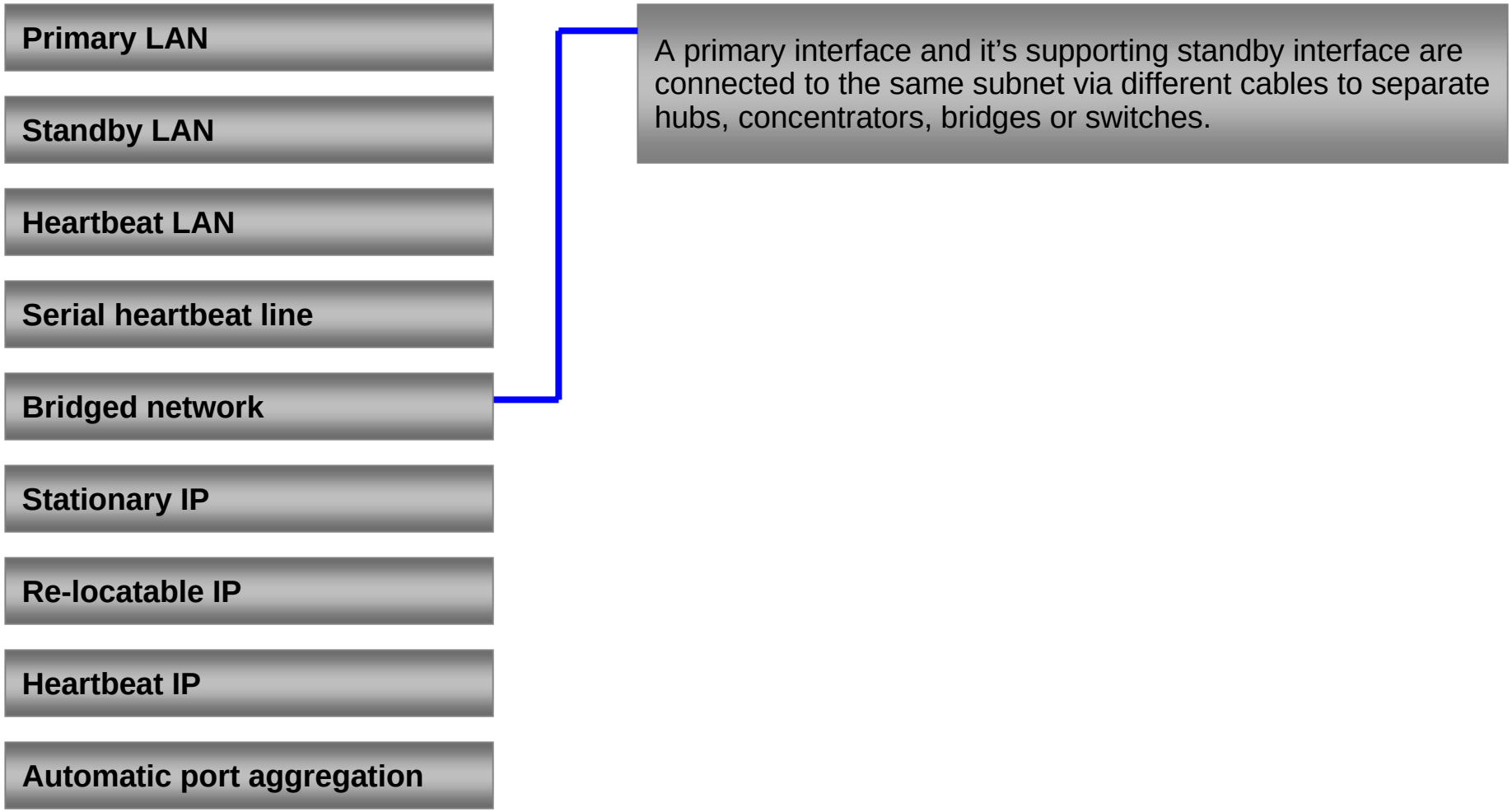
Network Terminology



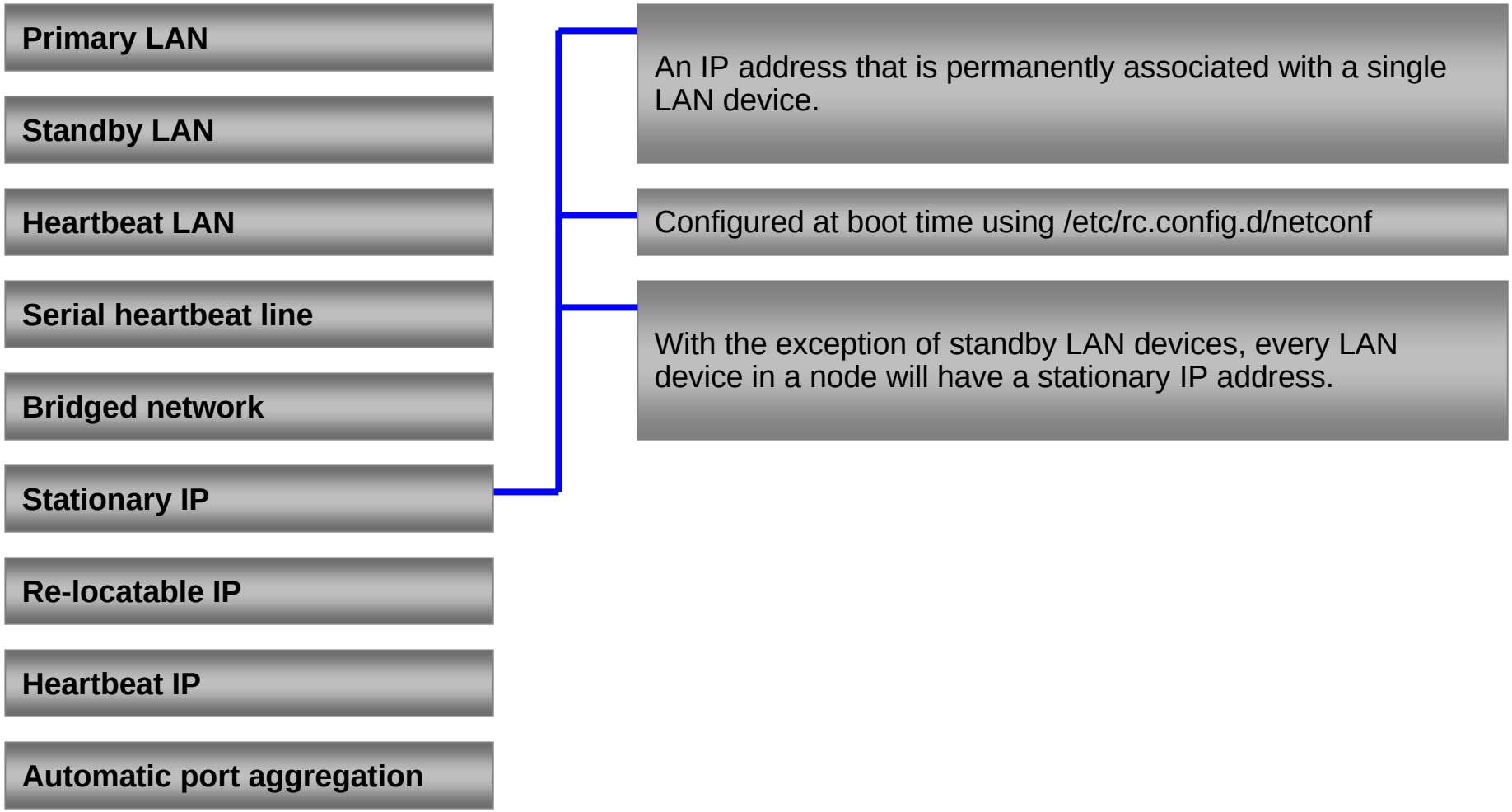
Network Terminology



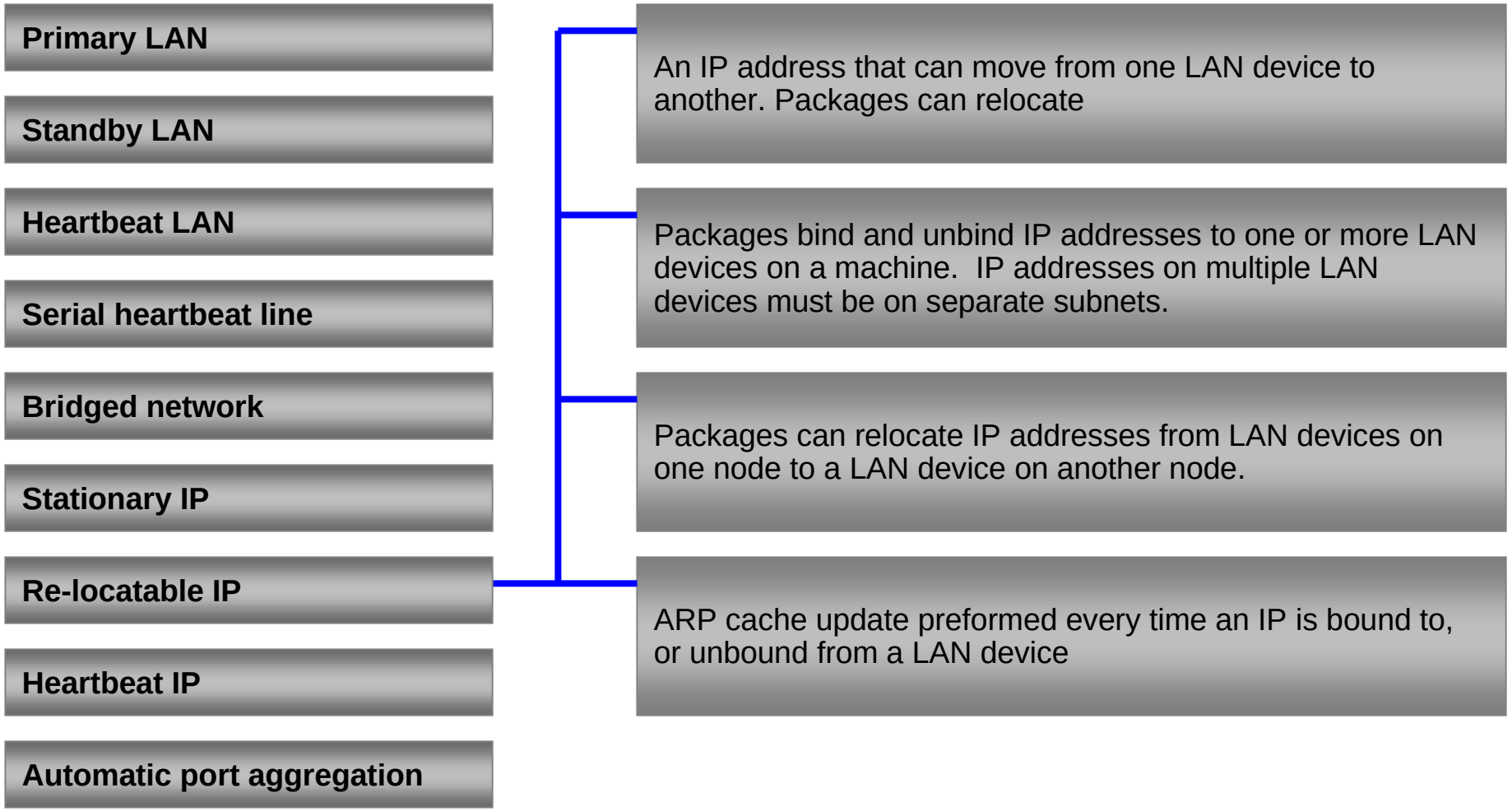
Network Terminology



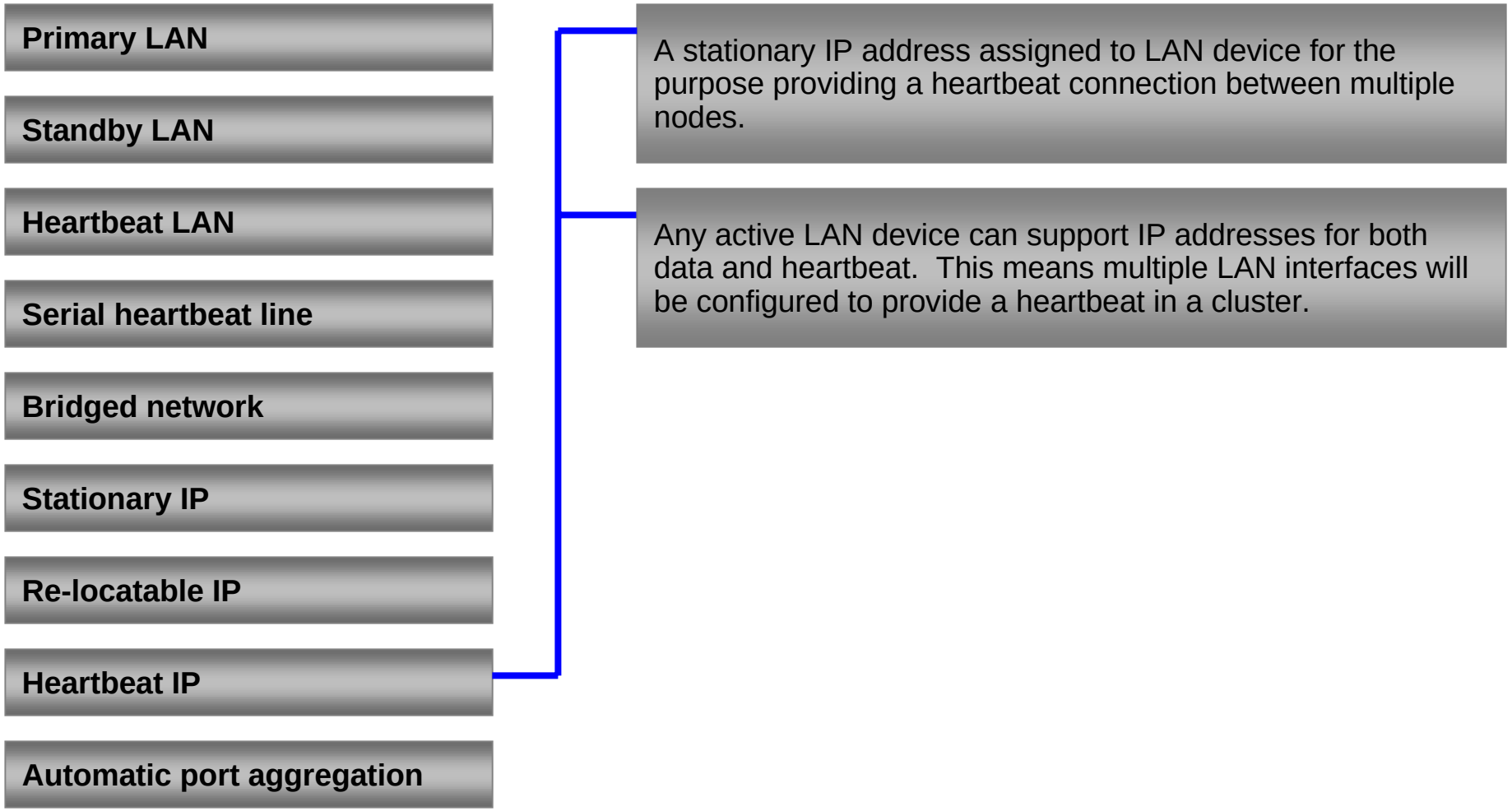
Network Terminology



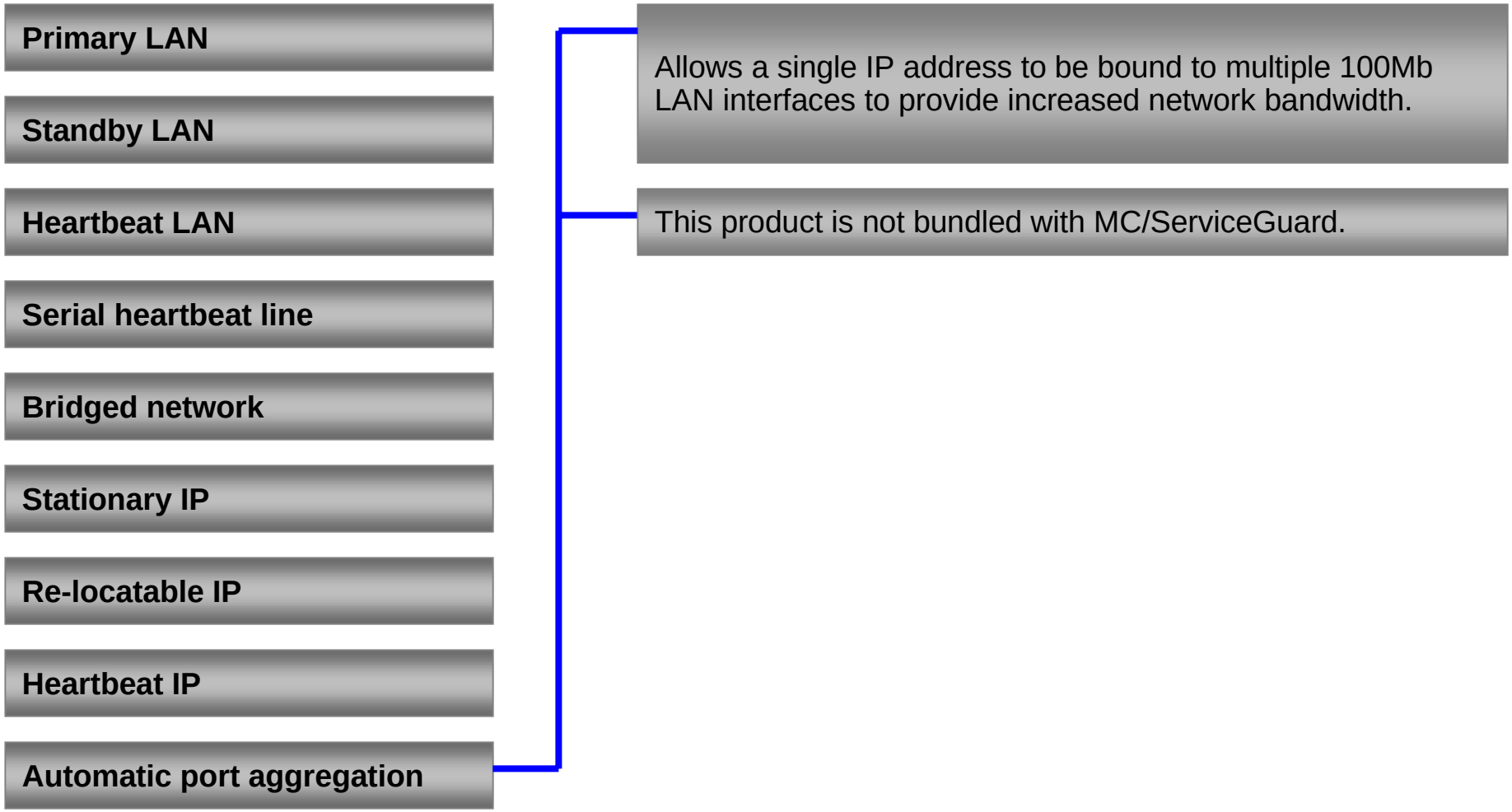
Network Terminology



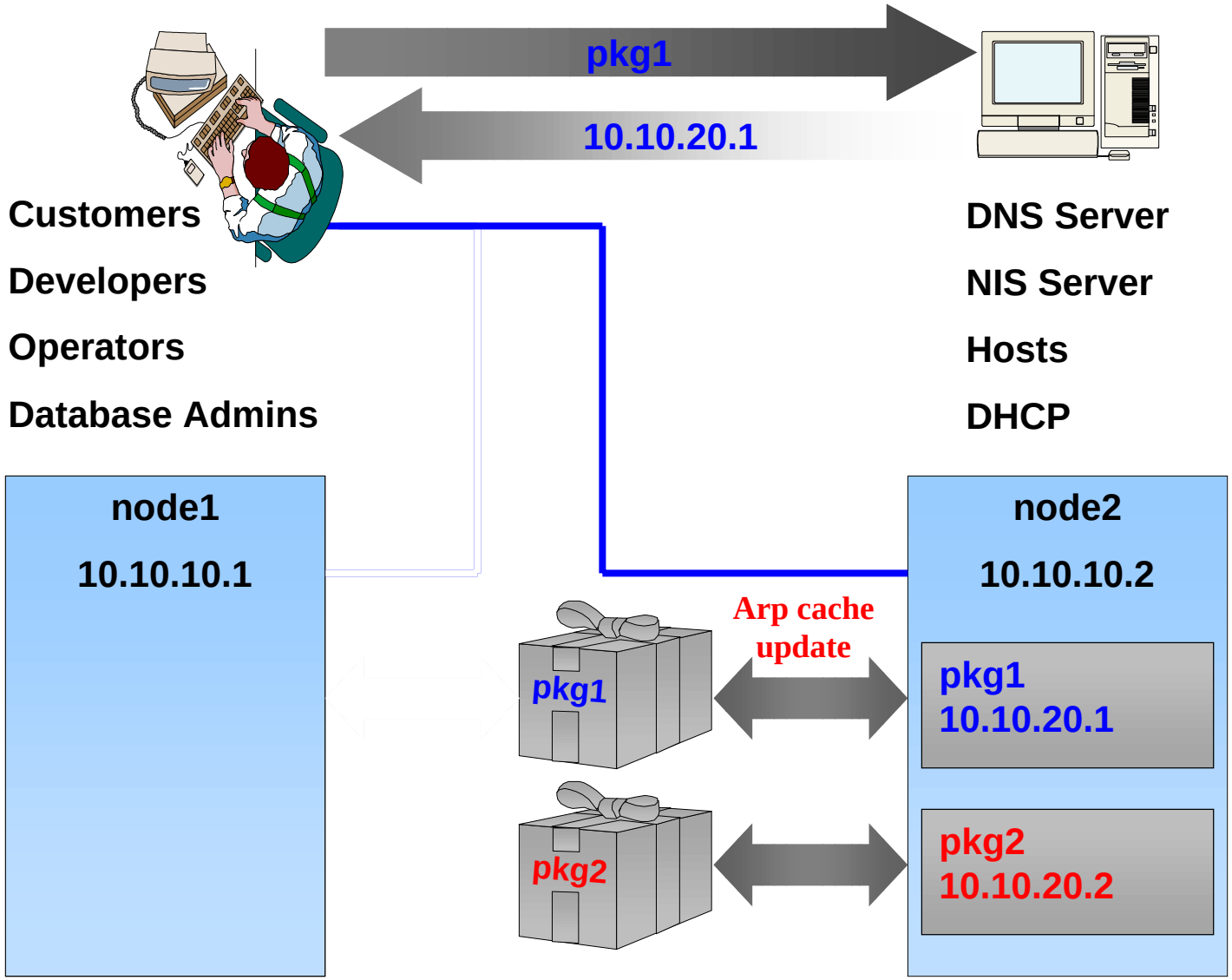
Network Terminology



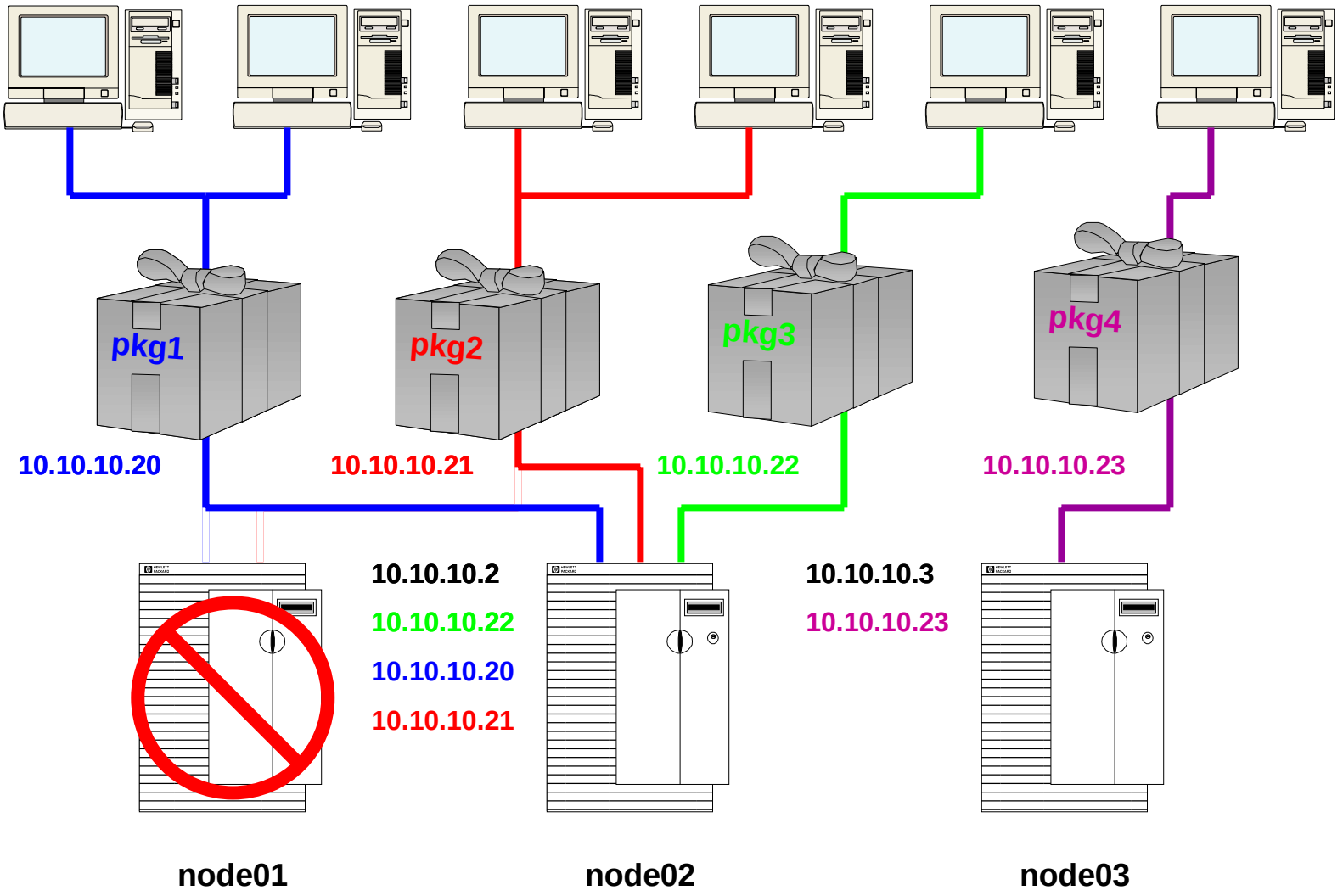
Network Terminology



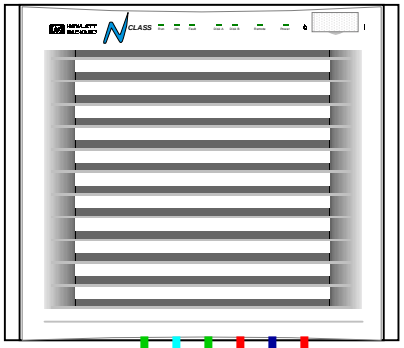
IP Addresses: The key to connectivity



How Packages Keep People Working



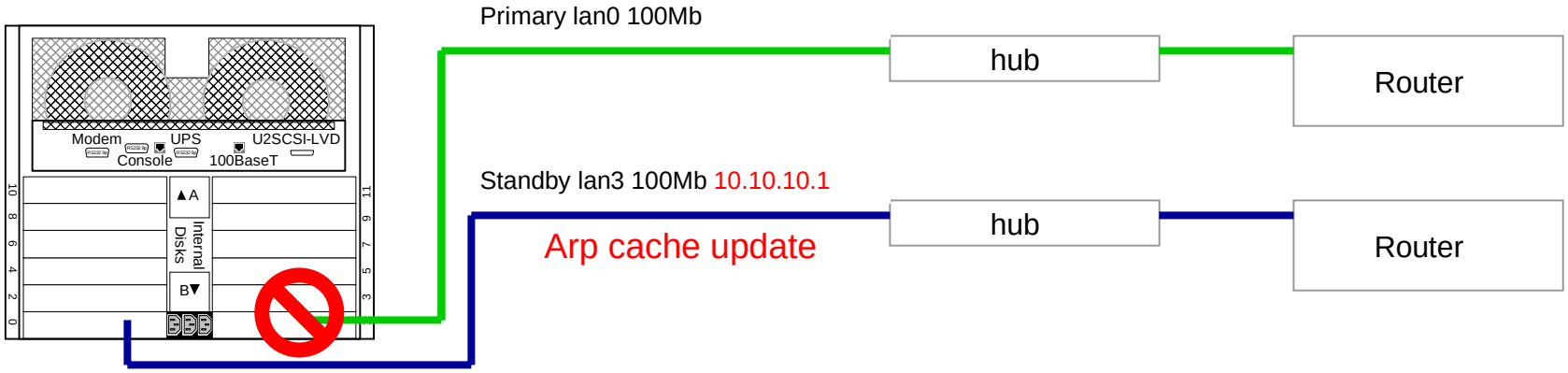
Network names, addresses & devices



- Stationary IP
- Relocatable IP

LAN0	100BaseT	Primary data	node01	10.10.10.1	Stationary
			pkg1	10.10.10.20	Re-locatable
			pkg2	10.10.10.21	Re-locatable
LAN1	10BaseT	Heartbeat	node01hb	10.10.20.1	Stationary
LAN2	100BaseT	Secondary data	node01b	10.10.30.1	Stationary
			pkg1b	10.10.30.20	Re-locatable
			pkg2b	10.10.30.21	Re-locatable
LAN3	100BaseT	Standby LAN N/A	N/A	N/A	
LAN4	FDDI	Dedicated backup	node01bu	10.10.40.1	Stationary
LAN5	FDDI	Standby LAN N/A	N/A	N/A	

Standby LAN: LAN Failover

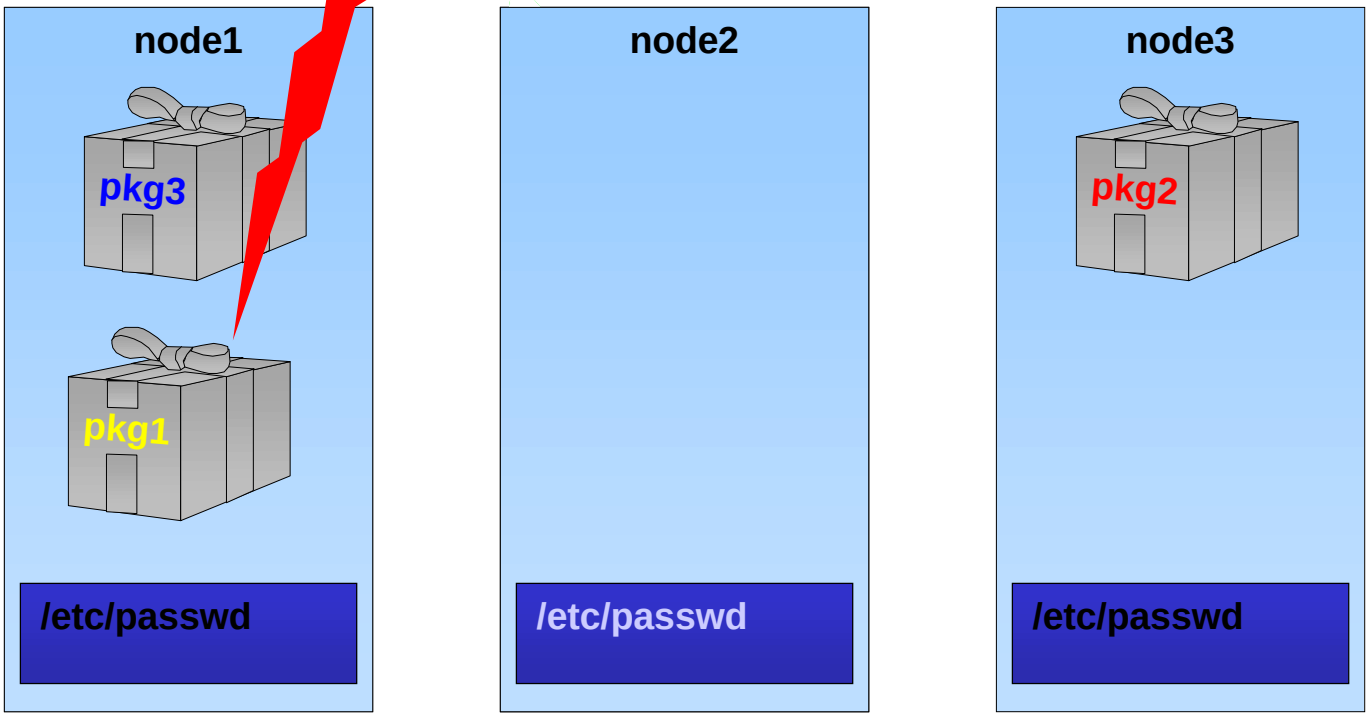


- Must be same type LAN device as primary
- Must be on same subnet as the device it is supporting
- Sub-second fail-over
- Both stationary and relocatable IP addresses are moved
- Arp cache update initiated for any fail-over
- Automatic failback once primary is back on-line

NIS: When to Use NIS

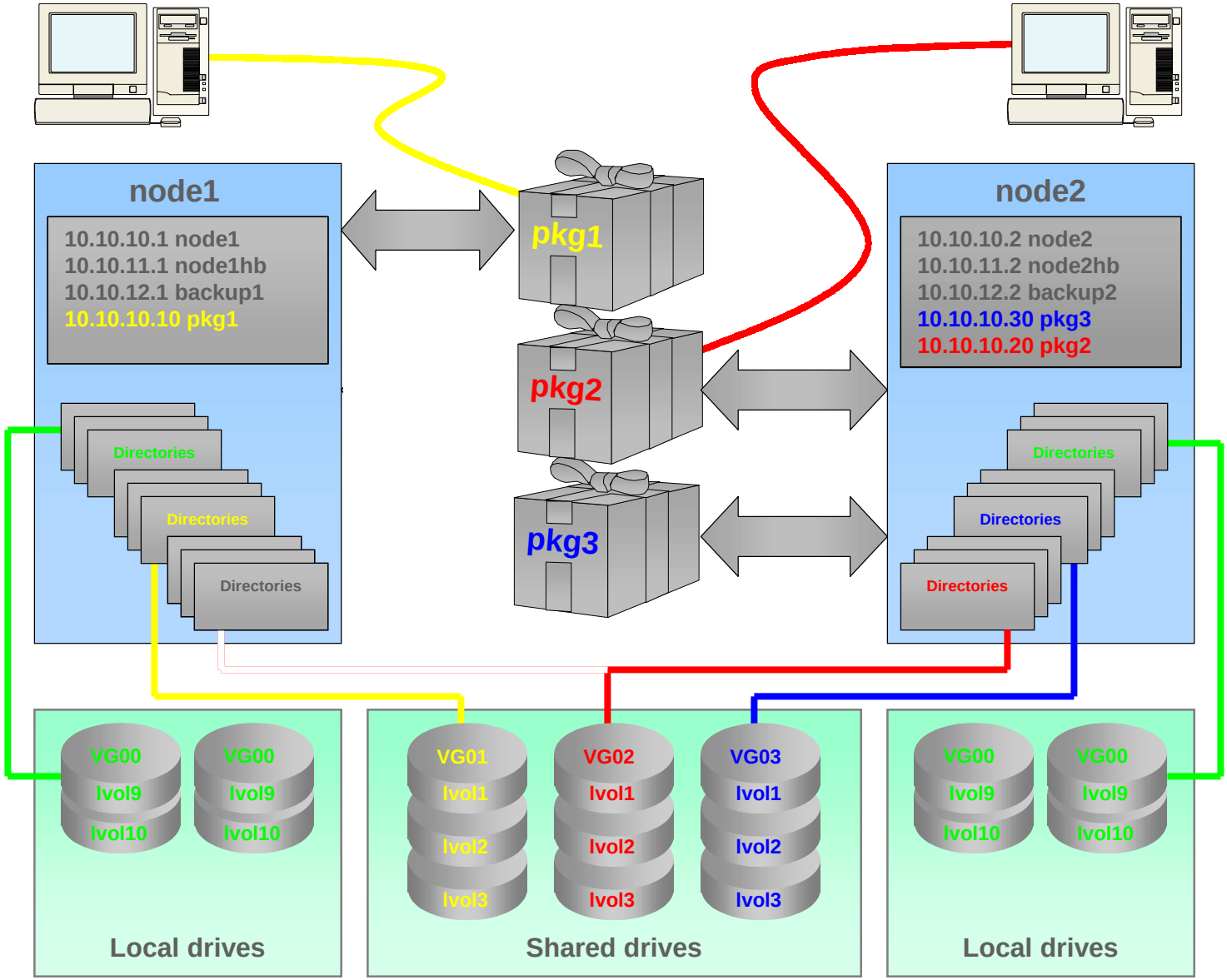
Customer

- **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**

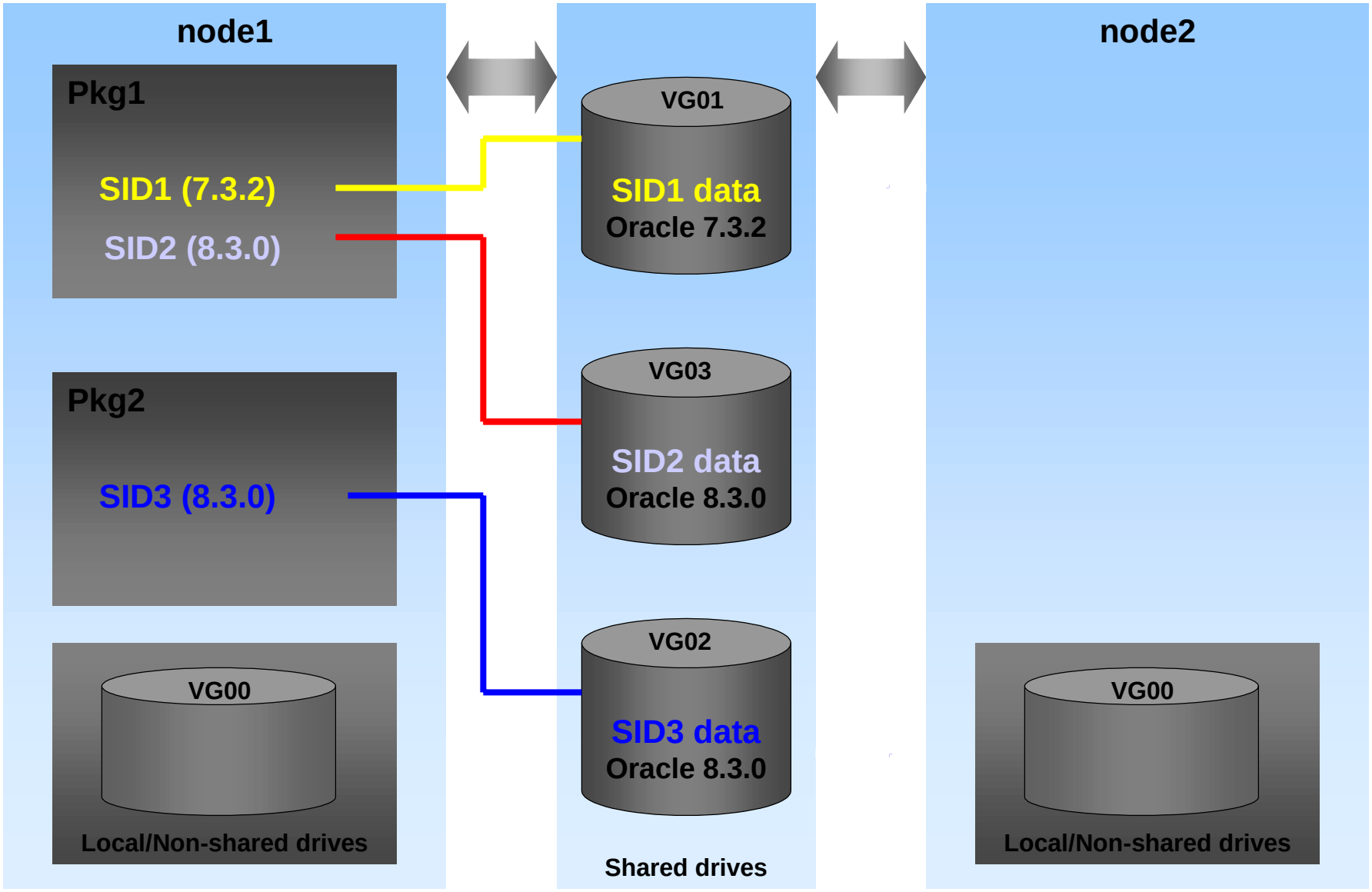


LVM concepts, issues & strategies

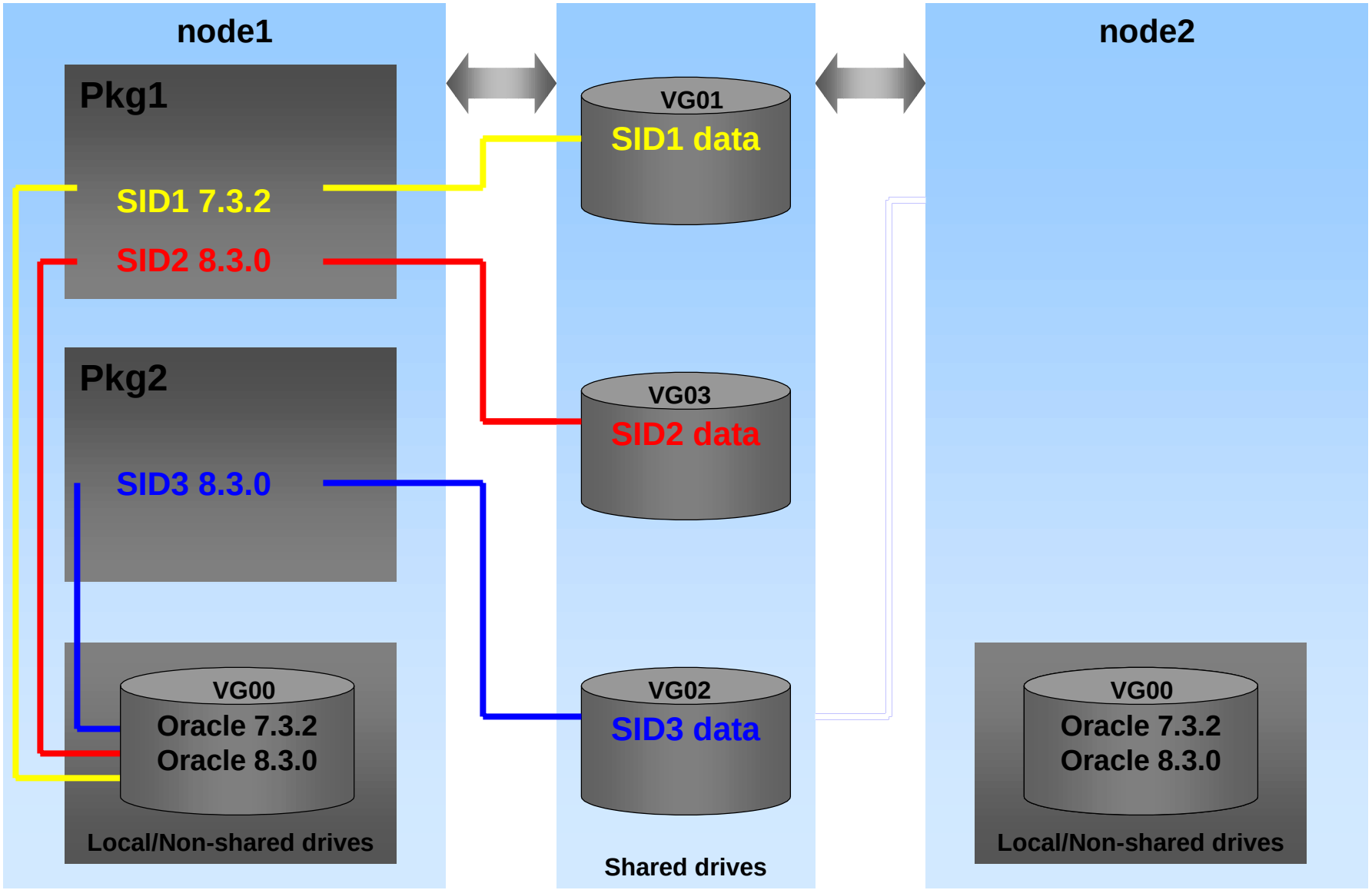
LVM: The mechanics of moving a package



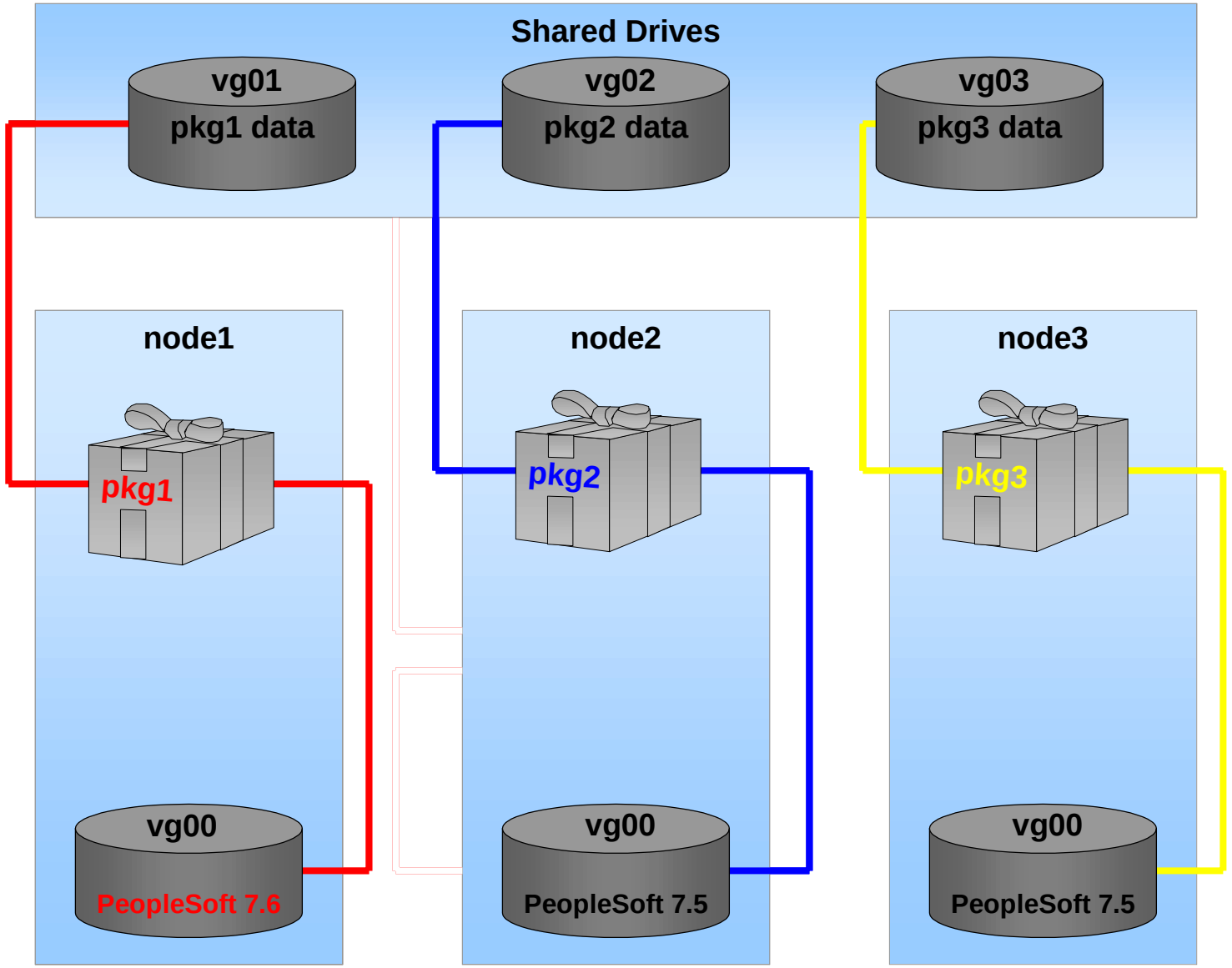
LVM: Binaries on shared drives



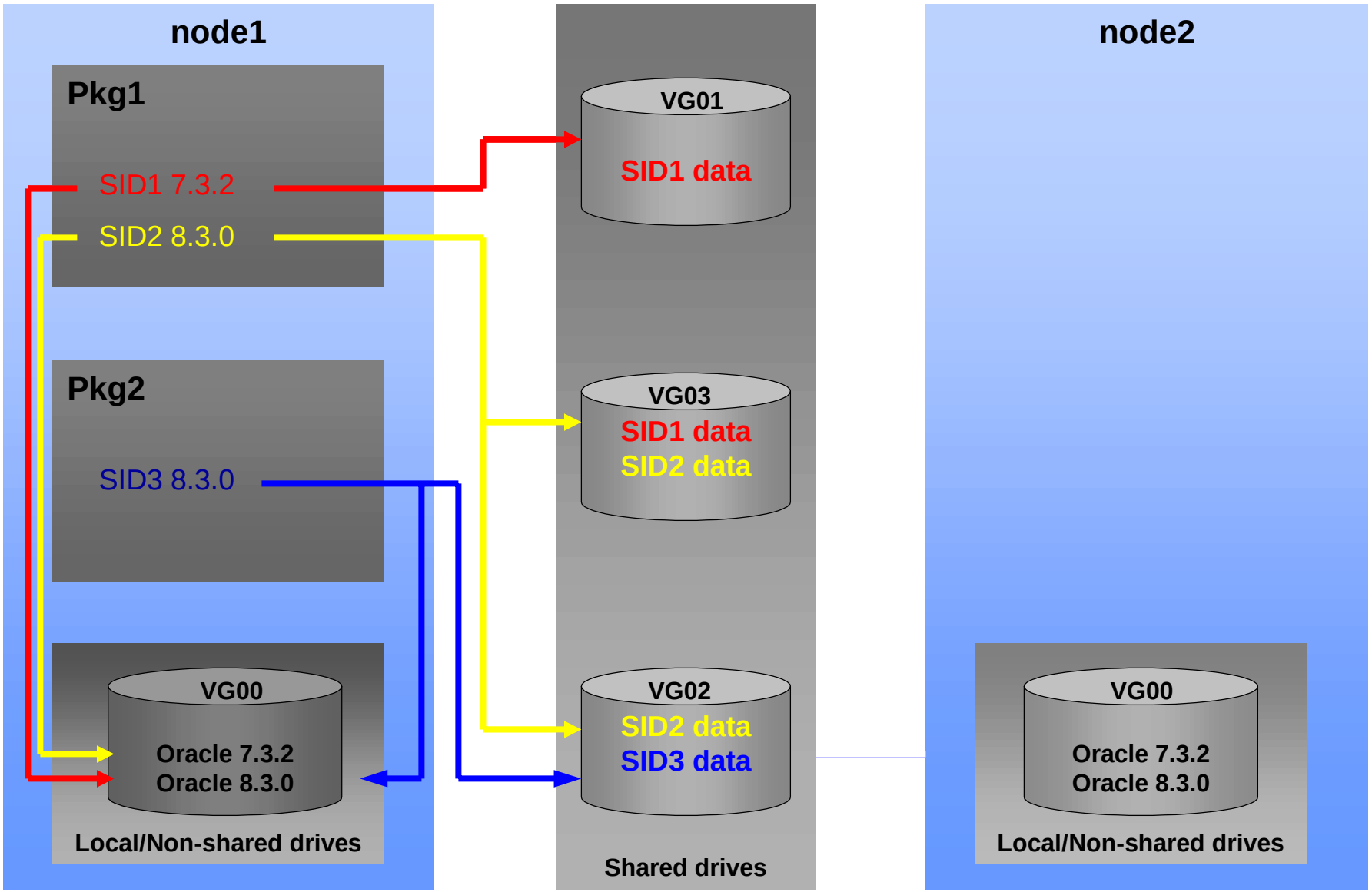
LVM: Binaries on local drives



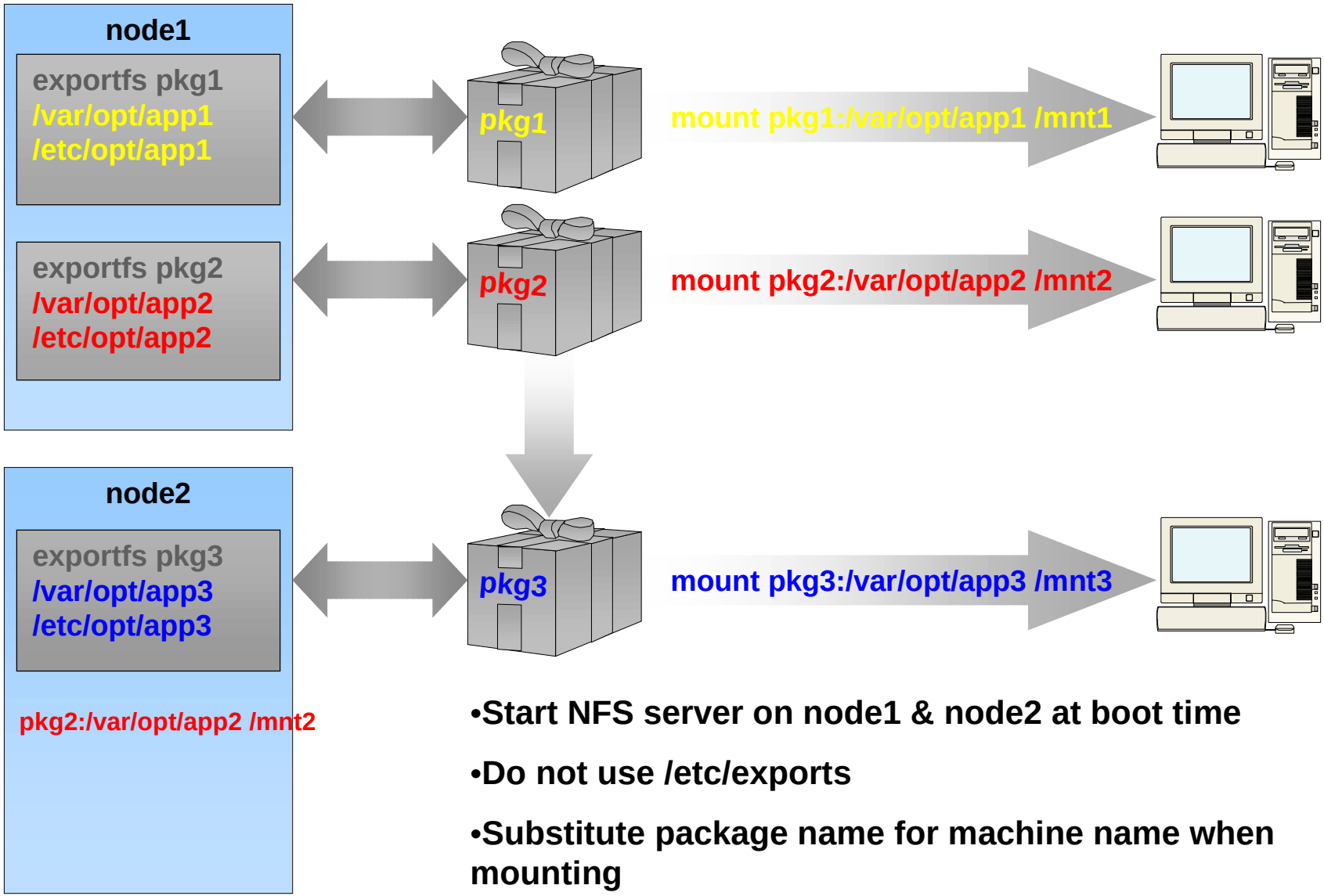
LVM: Rolling upgrade using binaries on local drives



LVM: Don’t Share Data Between Packages



NFS: Using NFS with Packages



- **Start NFS server on node1 & node2 at boot time**
- **Do not use /etc/exports**
- **Substitute package name for machine name when mounting**