

PowerHA SystemMirror introduction and certification preparation (a203872)

—
Antony (Red) Steel
Belisama Pte. Ltd.



IBM Champion

noun \ 'cham-pi-ən \

They're experts. They're leaders.

IBM Champion

IBM

```
#include <std_disclaimer.h>
```

These notes have been prepared by an Australian, so beware of unusual spelling and pronunciation. All comments regarding futures are probably nothing more than the imagination of the speaker and are IBM Confidential till after GA.

Thanks to:

Steven Finnes
Ravi Shankar
Paul Moyer
Dino and the Redbook team



altran

IBM

- With the recent release of PowerHA 7.2.3 and the publication of the redbook, this is a good time to give an update on PowerHA new features. The session will also introduce AIX Administrators to the operation of PowerHA and examples of how PowerHA assists with the management of your virtual environments - beyond maintaining their availability.
- A new redpaper is due out soon - Basic Requirements for HA and DR (on-prem and Cloud), which will review the IBM HA and DR portfolio.
- New certification: IBM PowerHA SystemMirror V7.2.5 AIX Administrator Specialty



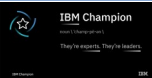
IBM PowerHA SystemMirror V7.2.3 for IBM AIX and V7.2.2 for Linux

Dino Quintero
Jose Martin Abeleira
Adriano Almeida
Bernhard Buehler
Primitivo Cervantes
Stuart Cunliffe
Jes Kiran
Byron Martinez Martinez
Antony Steel
Oscar Humberto Torres
Stefan Velica

 Cloud Power Systems

- Introduction to HA/DR concepts
- Review of the IBM portfolio and where PowerHA sits
- Introduction to PowerHA design and concepts
- Review/reminder of previous PowerHA release content
 - Spit/merge policy standardisation
 - Huge savings with Enterprise Pools and shared processor pool resizing
 - The CBU for Enterprise Systems
- The PowerHA what is new
- PowerHA Certification
- Resource materials

Introduction to High Availability

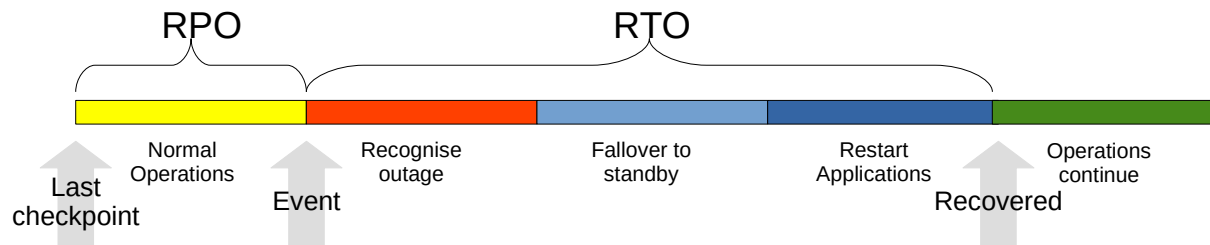


Introduction to High Availability

- High availability is:
 - The reduction to close to zero for downtime (not fault tolerance)
 - A solution to address planned or unplanned down time
 - A solution that is fault resistant, not fault tolerant
 - A solution that should eliminate single points of failure (SPOF)
- PowerHA is not the answer if
 - Cannot afford any downtime – life critical systems
 - Need a fault tolerant solution
 - Environment is not secure
 - Many users with root access
 - Then environment is not stable
 - Change management is not respected
 - No trained administrators
 - Procedures are not well documented
 - Environment is prone to user fiddle factor
 - Applications cannot be controlled
 - Scripts cannot be used to start/stop and recover applications from a failure



- Application managed
 - Concurrent Access - Active/Active across Servers / Sites – Oracle, Db2
 - Update / Log Shipping – Active/Passive – databases, for example SAP/ S4 HANA
- Manage Application
 - PowerHA SystemMirror Standard (HA) and Enterprise (DR)
 - Manage and monitor resources to restart application on working system
- LPAR (VM) managed
 - Simplified Remote Restart (SRR), Virtual Machine Recovery Manager (VMRM) HA and DR
 - Restart OS on working Frame with access to local (HA) or replicated (DR) data
- Typically range of SLA's across different applications, based on:
 - Recovery Time Objective (RTO) – the time before clients can access the application again
 - Recovery Point Objective (RPO) – last transaction saved



- Scale out
 - Many occurrences of the same application to provide redundancy (and throughput)
 - Containers
- Monolithic applications
 - Data replication
 - Mirroring within the data centre
 - Mirroring between the data centre
 - Stretched SAN
 - IP – Geographic Logical Volume Manager (AIX)
 - Replication between data centres
 - Storage based
 - IBM Metro Mirroring / Global Mirroring
 - EMC SRDF
 - Hitachi TrueCopy
 - IBM Hyperswap
 - IP Replication
 - Application Log shipping
 - GLVM

Setting realistic expectations

- What is considered an outage in your environment?
 - Unexpected downtime
 - Maintenance Tasks

Availability

90% (1-nine)

99% (2-nines)

99.9% (3-nines)

99.99% (4-nines)

99.999% (5-nines)

99.9999% (6-nines)

Downtime

36.5 days/year

3.65 days/year

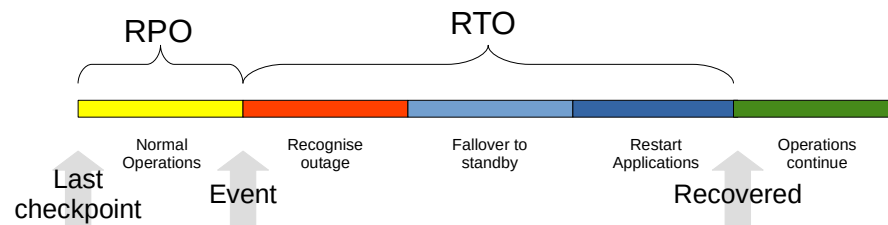
8.76 hours/year

52 minutes/year

5 minutes/year

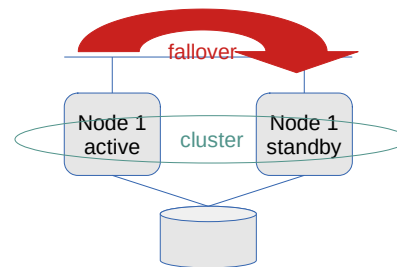
31 seconds/year

- What are the desired:
 - RTO – Recovery Time Objective
 - RPO – Recovery Point Objective

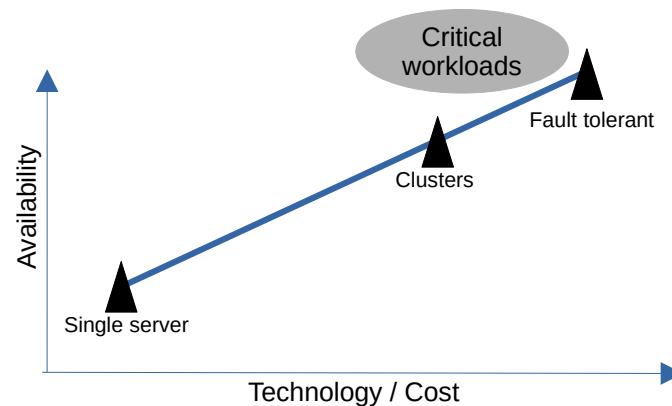


- Environment
 - Well managed and change control
 - Scripts can start / stop / monitor applications

- Cluster based High Availability
 - Business continuity by workload fail-overs
 - Provides for redundant operating environment
 - Automate bringing workloads online
 - Environment specific outage actions
 - Planned and unplanned outage management
- Cluster Solutions
 - Two or more Servers in the cluster
 - Cluster management console and software
 - Communication to achieve health and synchronisation
 - Active-Passive, Active-Active and concurrent models
- Disaster Recovery (DR) Management



- Redundancy
- Planned activities
- Unplanned events
- S/W&H/W updates

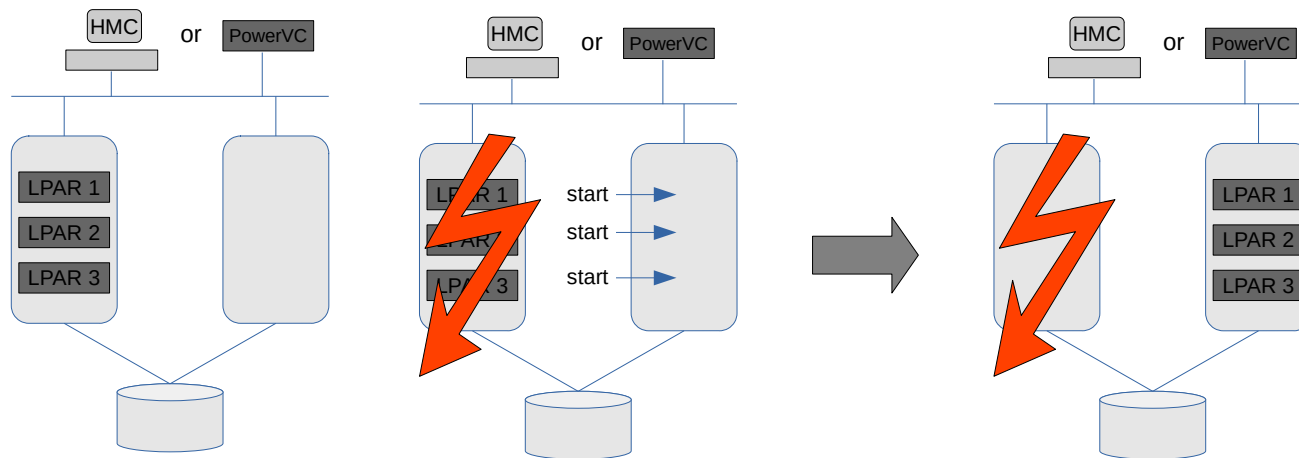


IBM HA/DR “Portfolio”

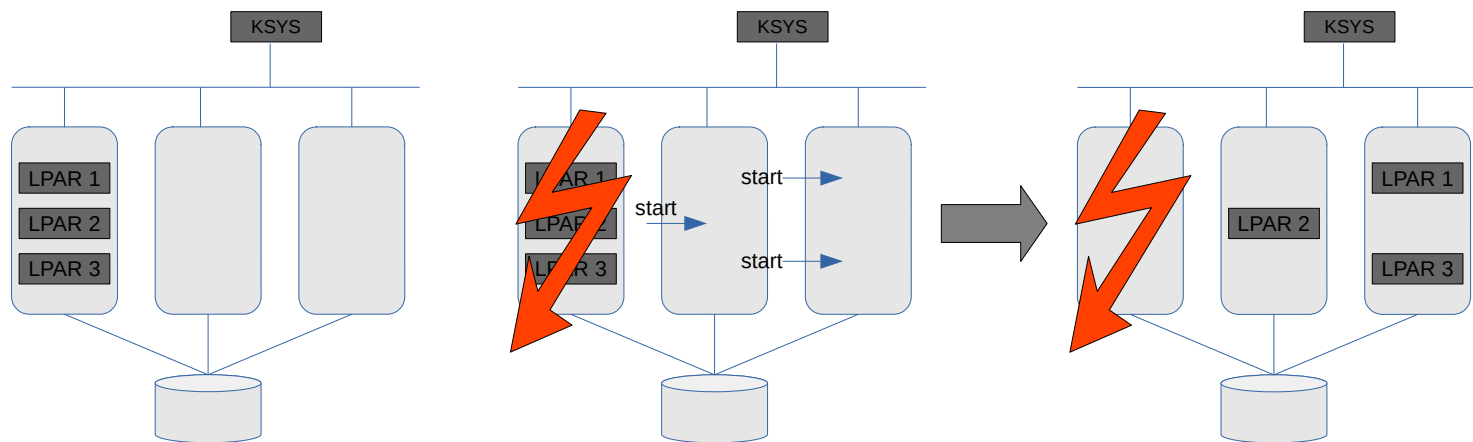


- Options for external management
 - Simplified Remote Restart (SRR)
 - PowerVC Managed SRR
 - VM Recovery Manager HA
 - Monitors host, if host fails, restarts VM on another host
 - Monitors applications / OS (registered hosts with agent installed)
 - For planned outage can use LPM to vacate host, moving VMs to other hosts in group
 - VM Recovery Manager DR
 - Using *ghostdev* and scripts for network changes at DR site

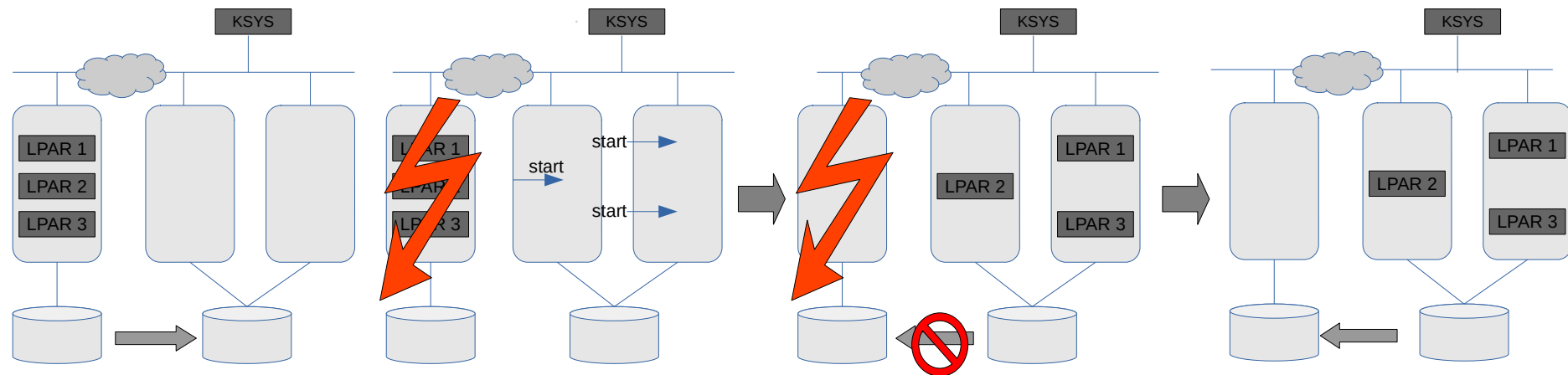
- The simplest option
- Configured via the HMC, NovaLink, or PowerVC
 - Enable an LPAR to be restarted on another Frame if the current Frame fails
 - If managed by PowerVC, the placement of the restarted LPARs can be controlled, otherwise it is a manual process.
- SRR is operating system agnostic and only recognises Frame failures.
- $RTO = \text{Operator intervention} + \text{LPAR IPL} + \text{Application start} (\$)$



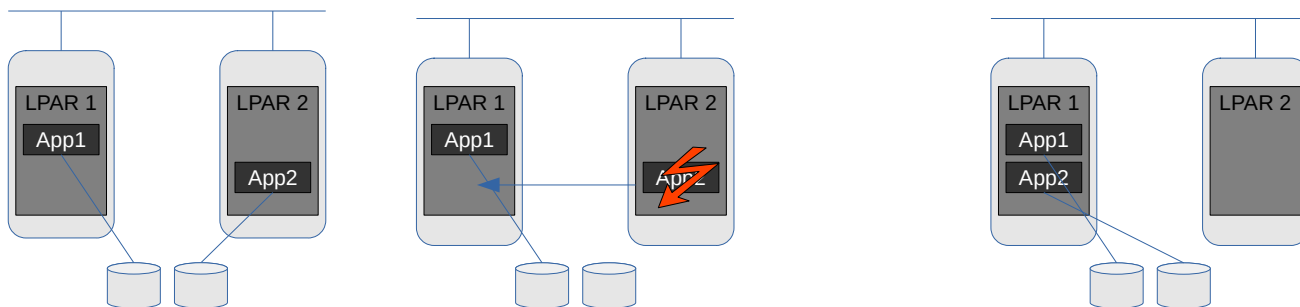
- VMRM HA is designed to automate and manage Simplified Remote Restart within the Data Centre.
- Operating System agnostic and can be scripted to go beyond just responding to Frame failures, using AIX or Linux agents to perform further monitoring.
- Distribution or placement of the virtual machines across the infrastructure can also be controlled.
- $RTO = LPAR\ IPL + Application\ start (\$\$)$



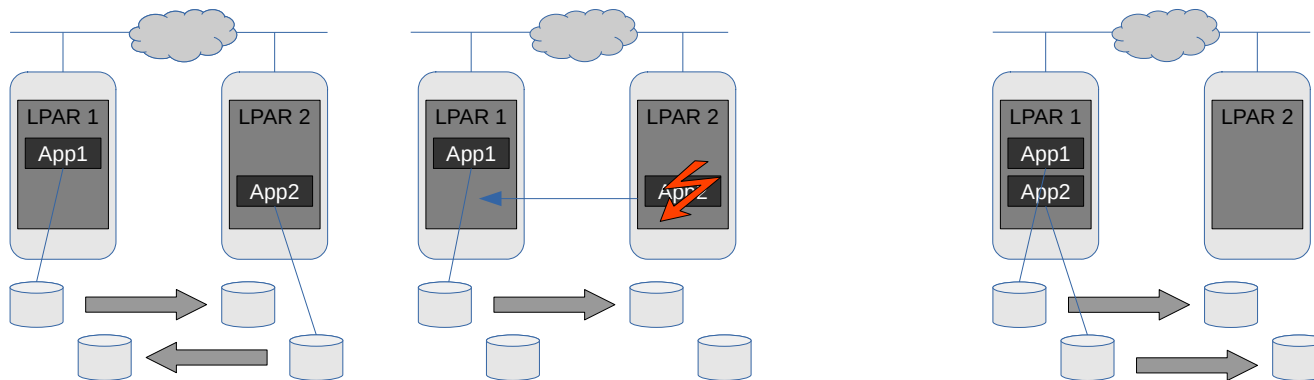
- Similar to VM Recovery Manager, but extends beyond the Data Centre.
- VMRM DR relies on replication to be configured at the storage layer, which it then controls.
- It supports a wide range of IBM and non-IBM storage solutions.
- There is an additional cool feature that allows you to easily run your DR testing by starting the LPAR in DR using cloned LUNs and thus not impacting production availability.
- $RTO = [Site\ Move] + LPAR\ IPL + Application\ start (\$\$)$



- Design principle based on logic of moving resources around the infrastructure to mask failures and allow applications to be restarted and made available for client access.
- PowerHA is also a useful tool for moving workloads around for maintenance and now includes an impressive array of management, monitoring and testing tools.
- RTO = Application start (\$\$\$)

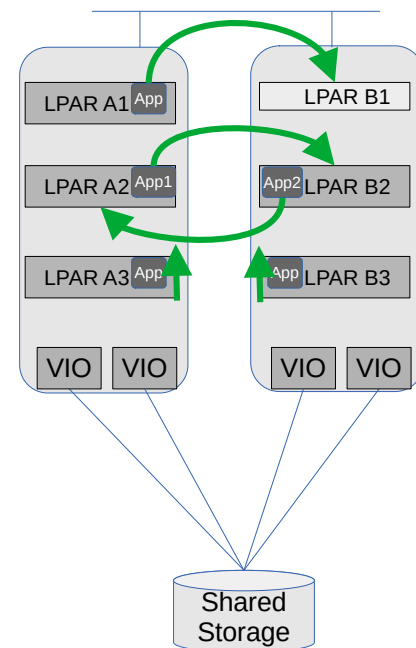


- Extending PowerHA by introducing sites (Data Centres), while controlling the replication of data and coordinating the behaviour of the clusters in each DC.
- In most of the clusters I have worked on, replication is managed by the storage subsystem, with just a few using Geographic Logical Volume Manager (GLVM).
- Has the same management, monitoring and testing tools as PowerHA
- RTO = Application start (\$\$\$\$)

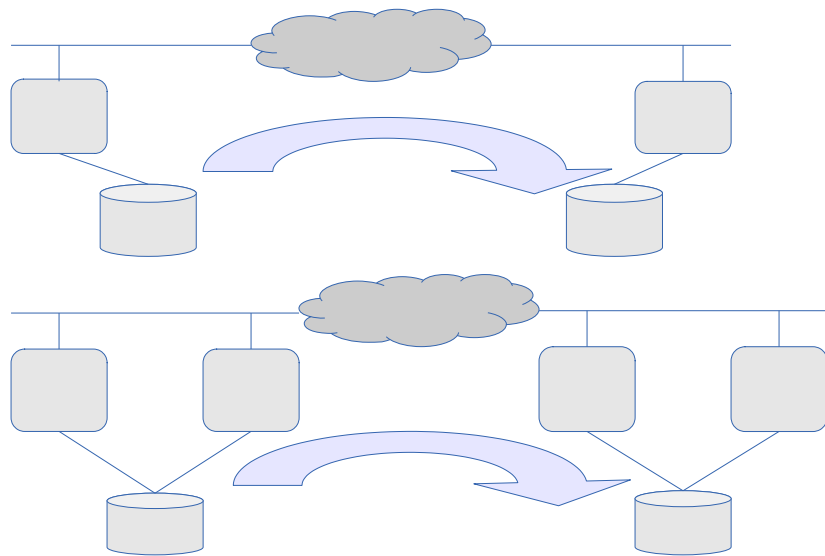


Technology	Storage Requirements	Automated Fallover	Source Server Status	Source VIO Server Status	LPAR / Application Outage
Live Partition Mobility	Shared	No	Active	Active	No (if VM is active)
Simplified Remote Restart	Shared	No	Inactive	Inactive	LPAR Yes
VMRM HA	Shared	Yes	Active or Inactive	Active or Inactive	Only if a server/LPAR outage occurs
VMRM DR	Remote Copy	No	Active or Inactive	Active or Inactive	LPAR Yes
PowerHA Std	Shared	Yes	Active or Inactive	Active or Inactive	Application Yes
PowerHA EE	Remote Copy	Yes	Active or Inactive	Active or Inactive	Application Yes

- One site – PowerHA SystemMirror
 - Dual servers, shared storage
 - Site only single point of failure
 - Cross-site LVM exception.
 - PowerHA Standard across 2 sites using LVM mirroring for data replication
 - VM Recovery Manager HA
- Supported topology configurations
 - Active | Standby
 - Active | Active (Independent Workloads)
 - Active | Active (Concurrent or online on all available nodes)
- Supported Resource Configurations:
 - Dedicated resources
 - Virtualised (NPIV, VSCSI, SSP)
 - Live Partition Mobility awareness
 - AIX 7.2 Live Update awareness



Clusters share storage, (if storage supported by AIX)



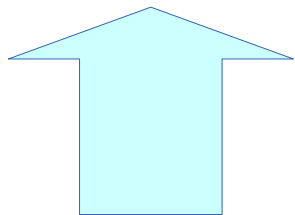
- Disaster Recovery

- Replication
 - GLVM
 - Storage / Database
 - VM Recovery Manager DR
- PowerHA SystemMirror Enterprise Ed.
 - PowerHA managing application and storage replication
 - GLVM
 - SVC; Storewise; MetroMirror; GlobalMirror
 - EMC SRDF / Hitachi TrueCopy/HUR

>> Planning and preparation

PowerHA Availability is not just for failures

- Planned
 - Maintenance
 - Upgrades
 - Testing
 - Development
- Unplanned
 - User Error
 - Application Failure
 - Component Failure
 - Operating System Failure
 - Environmental Disasters



Becoming a more important area -
PowerHA as an administration tool

LPM is an alternative for, ... but not for
(or software upgrades etc)

PowerHA will help to mask or eliminate
.... PowerHA is not fault tolerant (0 downtime)

Introduction to PowerHA SystemMirror for AIX

- PowerHA SystemMirror for AIX Standard Edition
- Cluster management for the data centre
 - Monitors, detects and reacts to events
 - Establishes a heartbeat between the systems
 - Enables automatic switch-over
 - Closely integrated into Power and AIX
 - LVM mirroring; capacity upgrade, live kernel patching ...
 - Multi-cluster federated security, user management, encryption
- IBM shared storage clustering
 - Can enable near-continuous application service
 - Minimise impact of planned & unplanned outages
 - Ease of use for HA operations
- Smart Assists – application agents
 - Out of the box deployment for SAP and other popular applications
- Mature Product
 - 26 Major releases (averaging one a year)
 - Over 18,000 customers worldwide

87% of respondents consider 99.99% (52.56 minutes) of unplanned per server/per annum downtime as the minimum acceptable level of reliability for mission critical servers and applications.

Also reported issues that cause downtime:

Human Error	59%
Security	56%
Software flaws	38%

ITIC's 2020 Reliability poll

- PowerHA SystemMirror for AIX Enterprise Edition
 - Cluster management for the Enterprise
 - Multi-site cluster management
 - Includes the Standard Edition function
 - Low cost mirroring option support

PowerHA SystemMirror editions

- Standard Edition
 - High Availability within a Data Centre
 - Health management of the cluster
 - Service availability against failures in any component of the stack

	Standard	Enterprise
Centralised Management C-SPOC	✓	✓
Cluster resource management	✓	✓
Shared Storage management	✓	✓
Cluster verification framework	✓	✓
Integrated disk heartbeat	✓	✓
SMIT management interfaces	✓	✓
AIX event/error management	✓	✓
Integrated heartbeat	✓	✓
PowerHA DLPAR HA management	✓	✓
Smart Assists	✓	✓
Cross site LVM	✓	✓

- Enterprise Edition
 - Adds long distance failover for Disaster Recovery
 - Low cost host based mirroring support
 - Extensive support for storage array replication
 - IBM DS8K, SVC, XIV, EMC, Hitachi, HP

	Enterprise Edition
Multi Site HA Management	✓
PowerHA GLVM async mode	✓
IBM Metro Mirror support	✓
IBM Global Mirror support DS8700	✓
EMC SRDF sync/async	✓
Hitachi Truecopy	✓
Stretched or linked clusters	✓
DS8000 Hyper Swap	✓



PowerHA SystemMirror editions

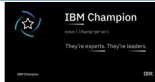
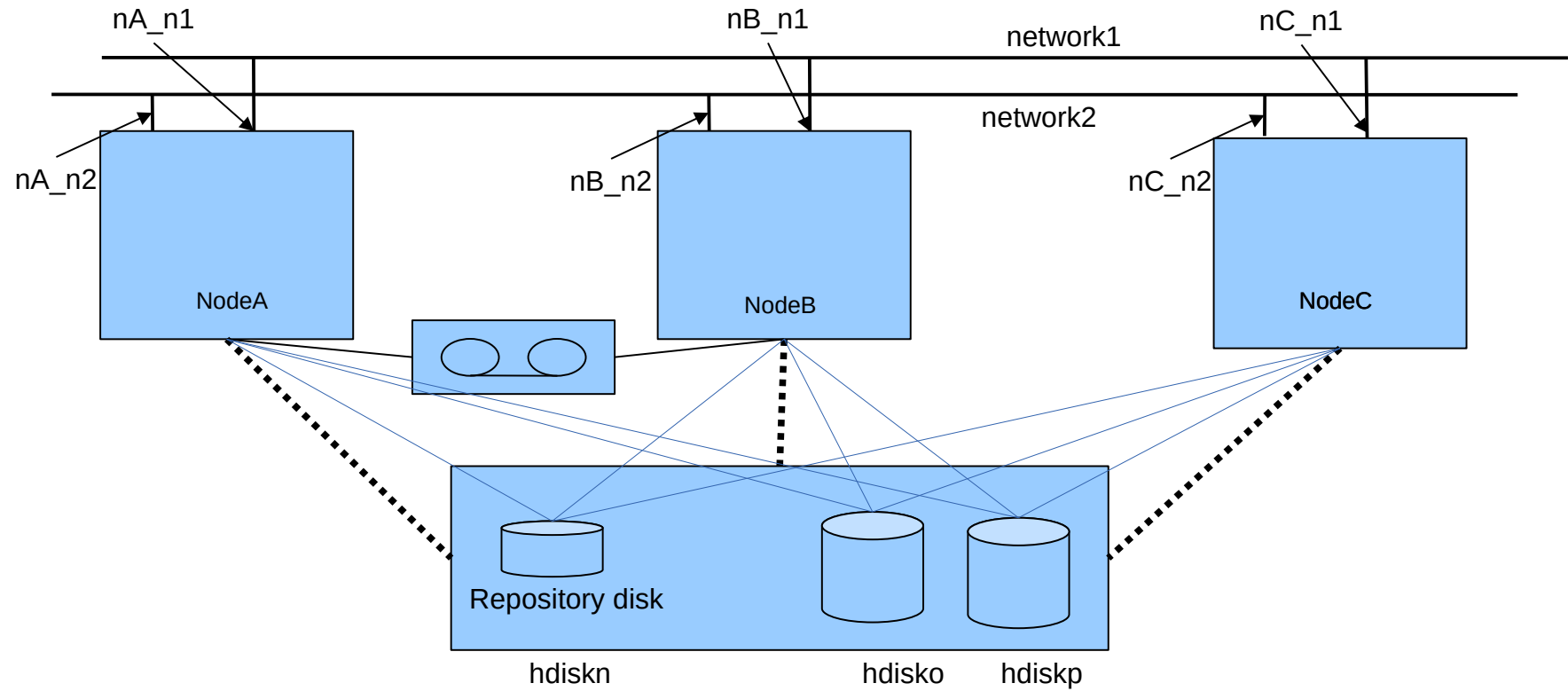
- Standard Edition
 - Supports up to 16 nodes
 - Supports Manual or Smart Assist based Deployments
 - Traditionally shares same common storage enclosure
- Supports 2 Site configurations:
 - No Copy Services Integration
 - No IP Replication Integration
 - Supports Site Specific IPs
 - Can be used with SVC Stretched Clusters
 - Used with Cross Site LVM configurations
 - Supports Split | Merge Policies when configured as a Linked Cluster
- Enterprise Edition
 - Supports up to 16 nodes
 - Supports Stretched or Linked clusters
 - Application Smart Assistants also included for local portion of fallover configuration
 - Provides local & extended cluster remote replication functions
- Can be configured to provide local clustering capabilities at first site and automated fallover to remote site
 - Automates storage level replication
 - Automates IP Replication (GLVM)
 - Integrates with DS8800 Hyperswap
 - Supports up to 2 Sites
 - Supports Split | Merge Policies
 - Higher Price per core

- Designing High Availability
 - A spare should be available for every single hardware and software component that is required to keep application running
 - No 'Single Point of Failure'
 - Whilst a generally accepted principle, not always adhered to, sometimes cut to reduce cost
 - Effects of the failure of a single component not always thought through, for example a single adapter networks with no disk heartbeat network
 - Eliminate single points of failure by:
 - Node
 - Using multiple nodes
 - Power source
 - Using multiple circuits or uninterruptible power supplies
 - Virtualisation layer
 - Two VIO Servers
 - multiple network and storage adapters in each VIO Server

Eliminate single points of failure (SPOF) by:

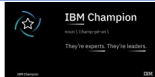
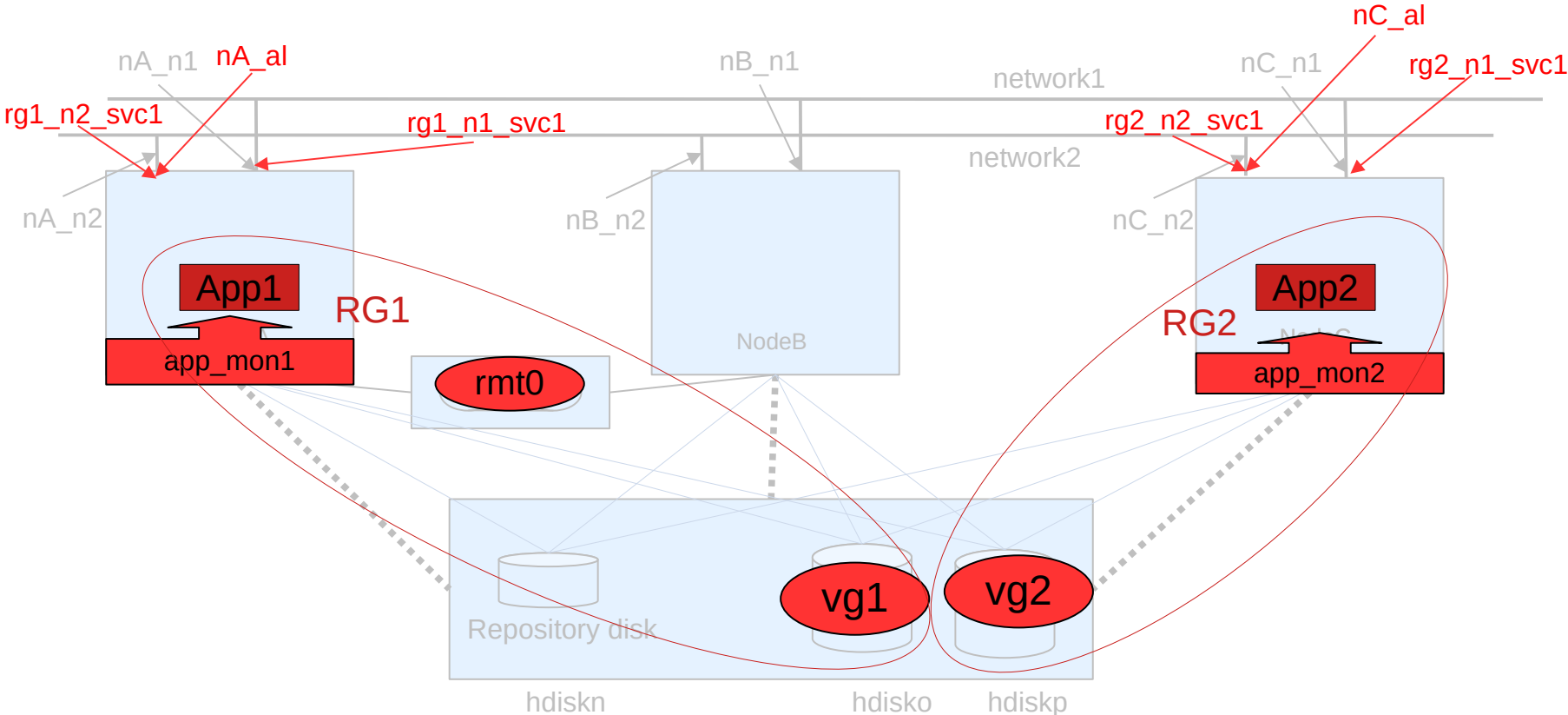
- Network adapter
 - Using redundant network adapters and bonding (ether channel etc)
 - Network Using multiple networks to connect nodes / clients
- TCP/IP subsystem
 - Using non-IP networks to connect nodes (heartbeat)
- Disk adapter
 - Using redundant disk adapter or multi-path hardware
 - hdisks using multiple physical disks with mirroring or raid
- Application
 - Adding node for takeover; configuring application monitor(s)
- Site
 - Adding an additional site

Cluster topology

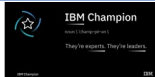
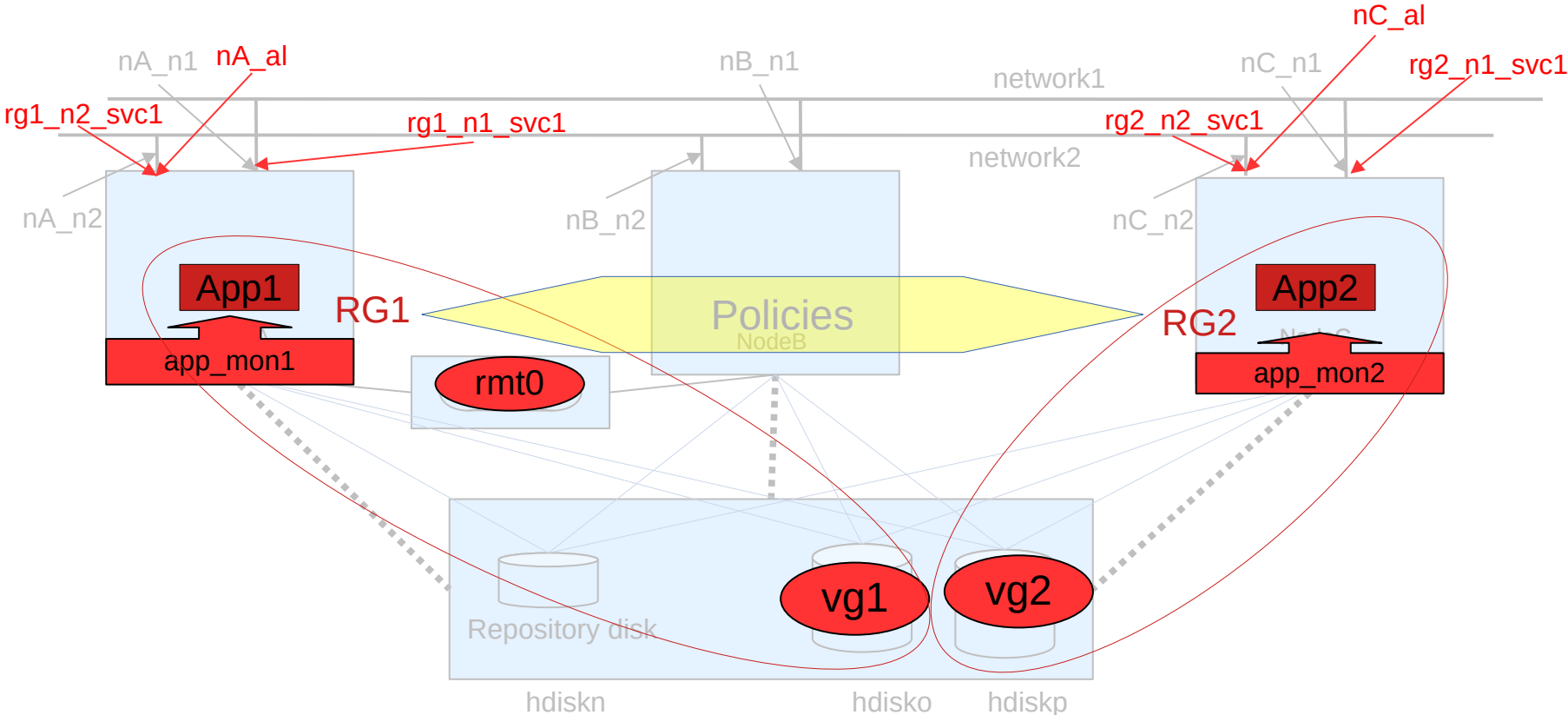


- A collection of all the resources required to run the application. Depending on the application may contain:
 - List of nodes on which to run (“home” node is the first node in the list)
 - IP Address
 - Volume Group / File systems
 - Application start script
 - Application stop script
 - Application monitor script(s)
 - Policies governing
 - Startup
 - Fallover
 - IP Address placement
- Dependencies can be configured to control the behaviour of the resource group(s)

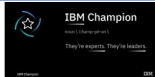
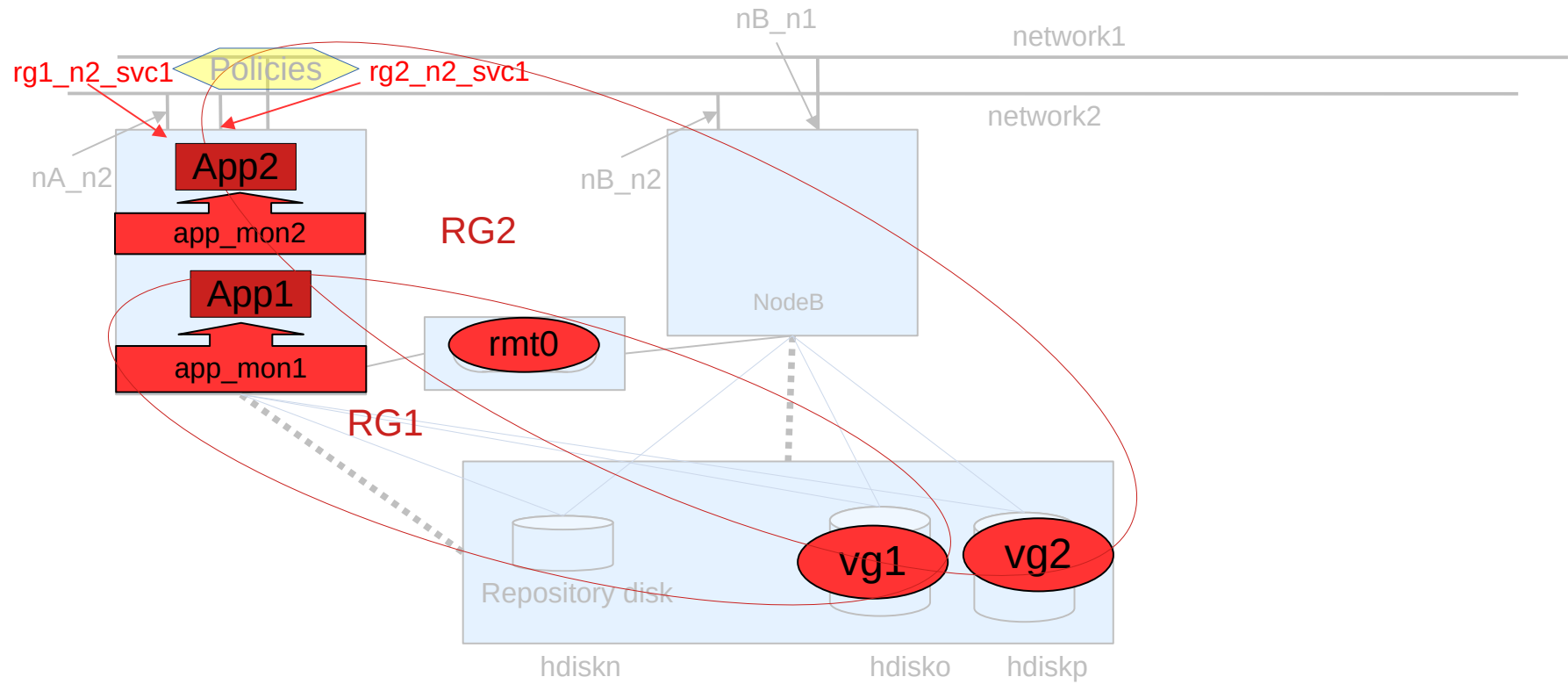
Resource groups



Resource groups



Resource groups

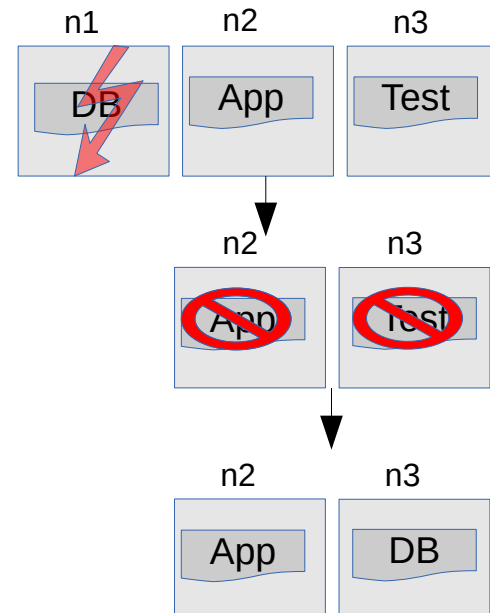


Cluster behaviour

- Resource group policies
 - Startup
 - Online on home node only
 - Online on first available
 - Online on all available
 - Ability to control start up distribution
 - Fallover
 - Fallover to next node in the list
 - Fallover using Dynamic node priority (CPU, Paging space, Disk IO, Adaptive (user defined))
 - Bring offline
 - Fallback
 - Fallback to higher priority node
 - Never fallback
 - Delayed fallback timer
- Resource group dependencies
 - covered in detail below
- IP distribution preferences
 - covered in detail below
- Inter site management policies
 - Online on Both Sites
 - Online on Either Site
 - Prefer Primary Site
 - Ignore

Resource Group dependencies

- Online on same node dependency
 - Resource groups come online on the same node
 - Start after and stop after
- Parent child dependency
 - Child will come online after the parent is stable, will go offline if the parent goes offline. Can have up to 3 levels
- Location (Online on different node) dependency
 - High, intermediate and low
 - High will force intermediate and low to move, intermediate will force low to move
 - Same priority cannot come online on same node
 - Same priority will not cause a movement
- Dynamic Node Priority
 - Processor Utilisation (cl_highest_idle_cpu)
 - Memory Utilisation (cl_highest_free_mem)
 - Disk I/O Utilisation (cl_lowest_disk_busy)
- Dynamic Node Adaptive Fallover – user defined
 - cl_lowest_nonzero_udscript_rc
 - cl_highest_udscript_rc



Online on different node dependency

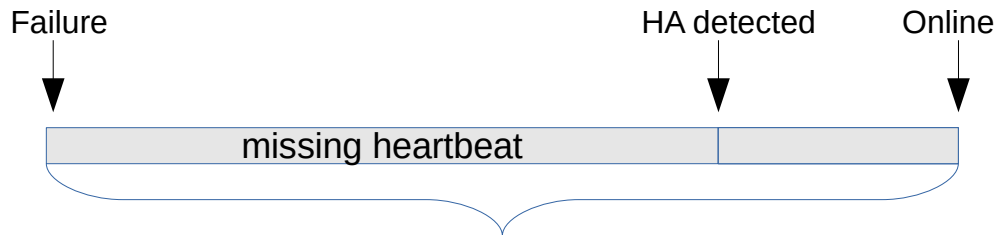
DB: n1,n3,n2	High
App: n2,n3,n1	Intermediate
Test: n3,n2,n1	Low

Parent / Child
DB – parent; App - Child

- Collocation
 - All Service labels will be on the same “adapter”
- Collocation with persistent
 - all service labels will be on the same “adapter” as the persistent IP.
- Collocation with Source
 - all service labels will be on the same “adapter” and the customer can choose the source IP of the outgoing packets
- Anti-collocation
 - all resources of this type will be allocated on the first “adapter” which is not already serving (or serving the least number of) addresses
- Anti-collocation with 1st Source
 - Same as above with the service IP being the source address of all outgoing packets.
- Anti-collocation with Persistent Labels
 - service labels will almost never be on the same “adapter” as the persistent IP, that is, service will occupy a different interface as long as one is available, but if no other is available then they will occupy the same interface.
- Anti-collocation with Persistent Labels and Source
 - Same as above with all outgoing packets having the service IP as the source address.

Are you using PowerHA features

- Are you aware of / using
 - Fast failure detection
 - Application monitoring
 - Check application health to allow PowerHA to respond
 - Disk handling changes
 - Enhanced concurrent mode volume groups – now the default. Fast disk takeover
 - No longer required to manage SCSI reserves
 - Cluster Single Point of Control (C-SPOC)
 - Manage disks, volume groups ..
 - Cluster test tool
 - Run default or customised test suite
 - Exploitation of mount guard
 - Nightly Verification of configuration

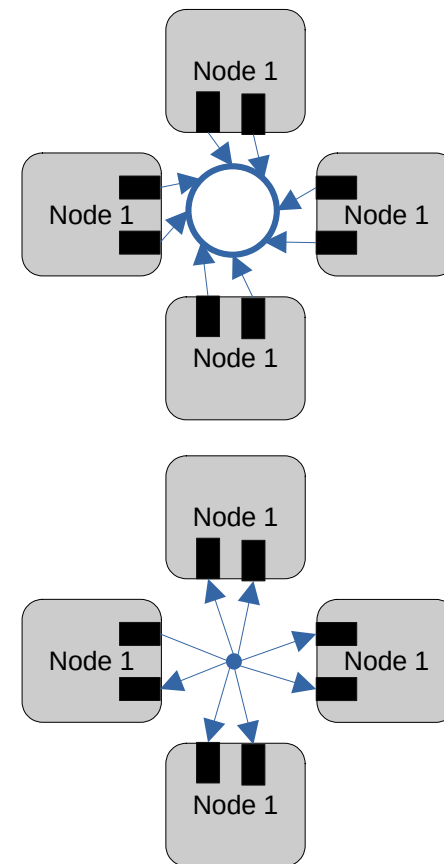


Remember that in the new versions of PowerHA, the developers used feedback from the field/PMRs to stop users making common mistakes

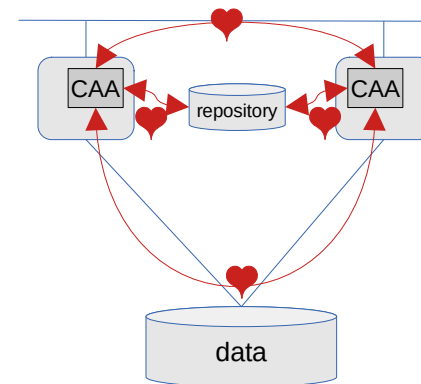
- File collections
 - Sets of files that can be replicated across the cluster at regular intervals or on synchronisation
 - Default sets contain important AIX and PowerHA configuration files
- Smart Assists
 - Installable set of selectable configurations and components (application controllers, monitors) required to make the following applications highly available with PowerHA:
 - Oracle
 - WebSphere
 - WebSphere MQSeries
 - DB2
 - FileNet P8
 - Lotus Domino Server
 - Tivoli Storage Manager
 - Tivoli Directory Server
 - SAP Netweaver
 - SAP MaxDB
 - SAP liveCache Hot Standby
- Cluster command line
 - used by Smart Assists
 - options to build, manage and check clusters
 - clmgr
 - `clmgr <ACTION> <CLASS> [<NAME>] [<ATTRIBUTES...>]`
 - `clmgr -v query resource_group`
- Application availability analysis

- Kernel based
- A set of services/tools embedded in AIX to help manage a cluster of AIX nodes and/or help run cluster software on AIX
 - IBM cluster products (including RSCT, PowerHA, and the VIOS) will use and/or call CAA services/tools.
 - CAA services can assist in the management and monitoring of an arbitrary set of nodes and/or running a third-party cluster.
- CAA does not form a cluster by itself. It is a tool set.
 - There is no notion of quorum. (If 20 nodes of a 21 node cluster are down, CAA still runs on the remaining node).
 - CAA does not eject nodes from a cluster. CAA provides tools to fence a node but never fences a node and will continue to run on a fenced node
- Requires a repository disk (protected at the storage level)
- By default all interfaces monitored
- `snap caa` to collect PD data
- Can secure communications between nodes

- PowerHA 6.1
 - Heartbeat Rings: detailed protocol
 - Leader, Successor, Mayor, Crown prince ...
 - Difficult to add/delete nodes
 - Requires IP aliases management in the subnet
- PowerHA >7.1
 - Multicast based protocol
 - Discover and use as many adapters as possible
 - Use network and SAN as available
 - Adapt to the environment: delay, subnet etc
 - Kernel based cluster message handling
- PowerHA 7.1.3 and above
 - Added unicast option



- Simple setup
- Multiple channels of communication
 - Network
 - SAN
 - Supports virtual fibre
 - LPAR uses sfwcomm VLAN device which uses VLAN 3358 to communicate with VIO Server, Server passes packets out fibre to other nodes *
 - WARNING: Not all current fibre adapters support SAN heartbeat
 - Central repository disk
 - Contains matrix with state of all nodes in cluster
 - 3 independent / redundant modes of communication

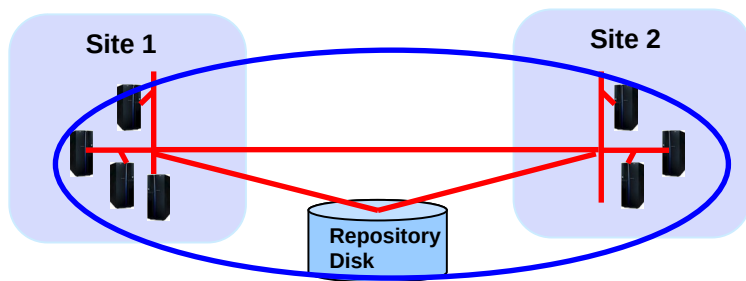


* IBM PowerHA SystemMirror V7.2 for IBM AIX Updates (Redbook)

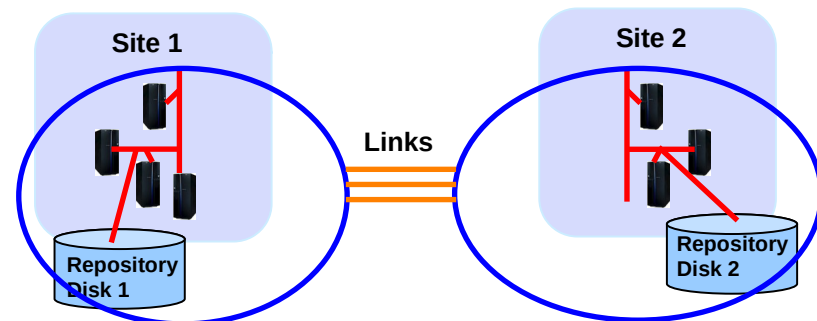
- PowerHA manages applications through an Application Controller
 - Application controllers are scripts, that can be started in one of two modes:
 - Foreground (HA event processing will wait for the completion of the application start script)
 - Background
 - Poorly designed scripts may cause hangs (config_too_long in HA events)
 - With earlier versions of PowerHA, the return codes were not checked, now will cause “EVENT ERROR” if return code is not 0
- Application provisioning
 - Ability to move CPU and memory resources with the application (manage license costs)
 - works with CoD / Enterprise Pools
- PowerHA uses Application Monitors to check on the application health
 - Each Application controller can have 0 or more monitors (at least 1 is strongly recommended)
 - Application monitor can trigger a response
 - Application restart
 - Notification
 - Resource group failover
 - See the monitors in the Smart Assists for examples

- Application monitors (cont)
 - The application monitor is bound to the application controller and there are two types:
 - Process Monitor
 - Checks the process table (entry type and number)
 - Custom Monitor
 - Your own script to check on the application, for example, read for a record in a table
 - Three modes:
 - Startup Monitor
 - Invoked when application started and used to confirm application running
 - Required by other dependencies, for example when to start the child resource group
 - Keeps checking until return code 0
 - Long running Monitors
 - Run at regular intervals to confirm application is still running, else trigger a response
 - Both
 - same monitor for checking startup and long term running

- Standard
 - Cross site LVM is exception, 2 sites with SAN connectivity and LVM mirroring between sites
- Enterprise (DR), introduces storage replication management / hyperswap
 - Stretched
 - Essentially one cluster operating across and aware of 2 sites
 - Linked
 - Two clusters coordinating activity and responses to events

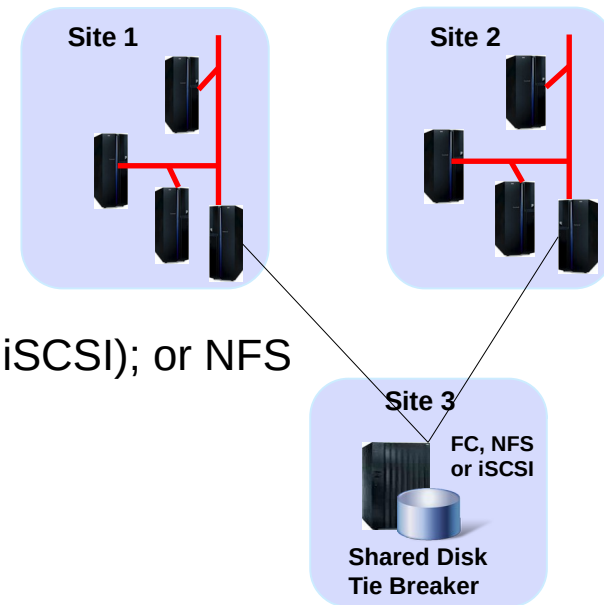


Multi Sites with Stretched Cluster

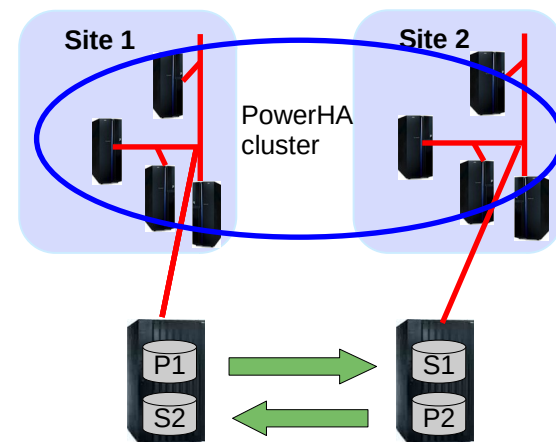


Multi Sites with Linked Clusters

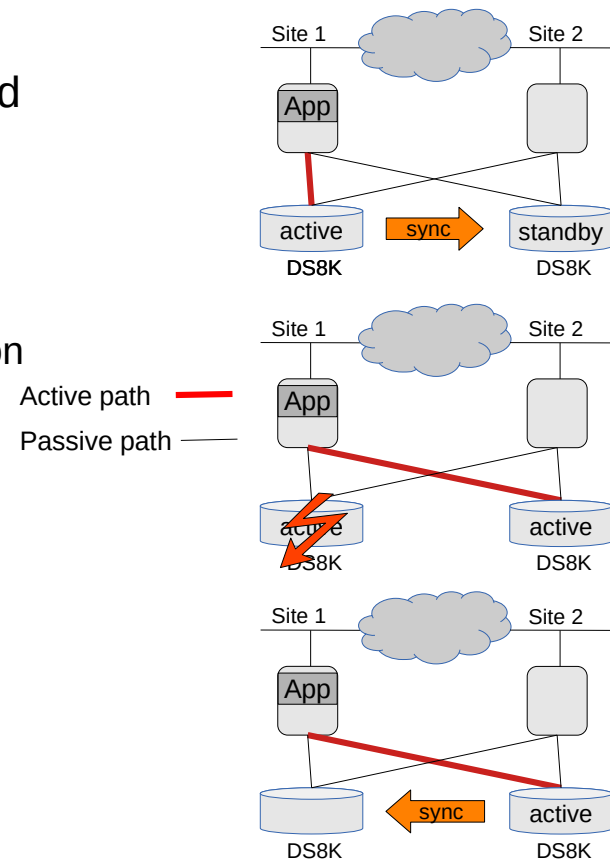
- Site split and merge policies
 - Majority rule
 - Site with > half nodes or site with node with lowest ID wins
 - Manual
 - Administrator controlled
 - Tie breaker
 - Site holding tie breaker wins
- Tie breaker support introduced (third site SCSI reserver (FC; iSCSI); or NFS)
 - SCSI reservation FC/iSCSI disk or NFS file as the tie breaker
 - Losing side is quiesced
 - More suited for Linked Clusters



- PowerHA SystemMirror Enterprise Edition with storage mirroring
 - IBM and most OEM storage support
 - Synchronous
 - No Data Loss (RPO = 0)
 - Metropolitan distances support (80-150km) as response time/latency impacted by distance - for light to travel 100 km round trip in fibre (200 km) takes around 2 ms
 - Asynchronous
 - Disaster can cause data loss (RPO > 0)
 - Much greater distances supported
- PowerHA
 - Manages automation of workload to mirror relationship management
 - PowerHA supports Metro & Global Mirror
 - Hyperswap support for Metro Mirror
 - Has support for:
 - Consistency Groups provide data integrity for databases
 - Consistency Groups can span multiple disk subsystems
 - Multiple consistency groups supported between sites
 - Adapts to failures and manages the replication as required



- Continuous Availability against Storage failures
- Substitutes storage secondary to take the place of failed primary device
 - Non-disruptive - applications keep running
 - Key value add to HA/DR deployments
- Customer Benefits
 - HyperSwap device configuration transparent to application
 - Unplanned HyperSwap:
 - Continuous availability against storage failures
 - Planned HyperSwap:
 - Storage Maintenance without downtime
 - Storage migration without downtime



Installation



- PowerHA is very flexible
 - Options in topology, storage design, networking, resource group behaviour
 - Many options make it complex, need to understand these so
 - There is no unexpected behaviour
 - Think through all failure scenarios
 - Availability is not compromised
 - optimal availability is achieved
 - PowerHA still needs a reboot after installation
 - Check recommended levels
 - FLRT (<https://www14.software.ibm.com/webapp/set2/flrt/power>)
 - PowerHA reference information (<https://www.ibm.com/docs/en/powerha-aix/7.2?topic=reference-information>)
 - Standard edition filesets
 - CAA packages (part of the AIX install media)
 - bos.cluster.rte
 - bos.ahafs
 - bos.clvm.enh
 - devices.commom.IBM.storfwor

PowerHA installation

- Standard edition filesets (cont)
 - PowerHA packages
 - cluster.license electronic license file
 - cluster.es.server base cluster filesets
 - cluster.doc.en_US.es clinfo and clstat samples and include files and a Web Based Monitor
 - cluster.es.client PowerHA SystemMirror PDF Documentation
 - cluster.es.cspoc cluster client binaries and libraries, plus Web based Smit for PowerHA
 - cluster.es.migcheck C-SPOC and dsh
 - cluster.es.nfs Migration support
 - cluster.es.nfs NFS Server support
 - cluster.msg.en_US.es U.S. English message catalogue
 - cluster.man.en_US.es man pages - U.S. English
 - Optional packages
 - cluster.doc.en_US.assist Smart Assist PDF documentation
 - cluster.hativoli PowerHA SystemMirror Tivoli Server and Client
 - cluster.es.assist Smart Assist filesets
 - cluster.msg.en_US.assist U.S. English Smart Assist messages
 - cluster.es.cfs GPFS support
 - cluster.es.worksheets Online Planning worksheets

PowerHA Enterprise Edition filesets

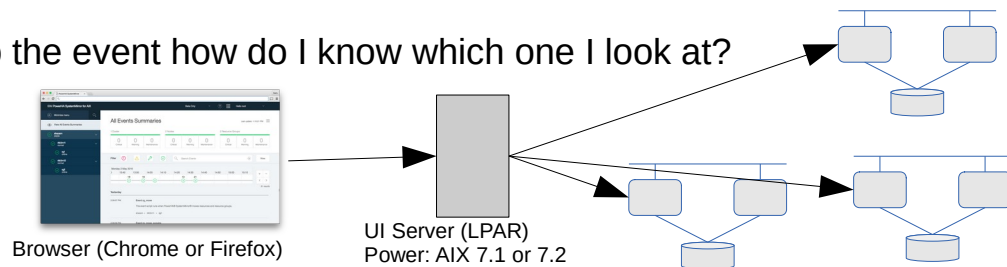
Replication Type	Replication Type File Sets to Install
ESS Direct Management PPRC	cluster.es.pprc.rte cluster.es.pprc.cmds cluster.msg.en_US.pprc
ESS DS6000/DS8000 Metro Mirror DSCLI PPRC	cluster.es.spprc.cmds cluster.es.spprc.rte cluster.es.cgpprc.cmds cluster.es.cgpprc.rte cluster.msg.en_US.svcpprc
San Volume Controller (SVC) & Storwize Family	cluster.es.svcpprc.cmds cluster.es.svcpprc.rte cluster.msg.en_US.svcpprc
XIV, DS8800 in-band and Hyperswap, DS8700/DS8800 Global Mirror	cluster.es.genxd.cmds cluster.es.genxd.rte cluster.msg.en_US.genxd
Geographic Logical Volume Mirroring (GLVM)	cluster.doc.en_US.glvms.pdf cluster.msg.en_US.glvms cluster.xd.glvms glvms.rpv *(file sets in base AIX)
EMC SRDF	cluster.es.sr.cmds cluster.es.sr.rte cluster.msg.en_US.sr
Hitachi TrueCopy / Universal Replicator	cluster.es.tc.cmds cluster.es.tc.rte cluster.msg.en_US.tc

Plus: cluster.xd.license

PowerHA GUI



- Manage one or more clusters
 - creating, deleting, and modifying existing clusters and cluster configuration objects:
 - Resource groups (RGs)
 - Volume groups (VGs); File systems
 - Service IP addresses; Persistent IP addresses
 - Manage snapshots and clone clusters from snapshots
 - Application availability and monitoring metrics
 - Cross cluster verification
 - Health checking – cluster health at a glance
 - I want to know about the health of my environment ... where the resources are, active node or standby node, etc.
 - The kind of events that get posted is overwhelming, you get lost with the important and non-important info.
 - There might be 100 lines relevant to the event how do I know which one I look at?



- In just a few minutes we were able to:
 - determine which resource(s) had a problem

The screenshot displays the IBM PowerHA SystemMirror for AIX monitoring interface. On the left, a sidebar lists resources with their status: demo_cluster (Warning), phac1 (Maintenance, offline), rg1 (Maintenance, offline), rg2 (Maintenance, offline), rg3 (Offline), phac2 (Normal), rg1 (Maintenance, offline), rg2 (Maintenance, offline), rg3 (Online), and cluster2 (Stable). The main area shows 'All Events Summaries' with a dashboard for 2 Clusters, 5 Nodes, and 12 Resource groups. Each has a bar chart showing counts for Critical, Warning, and Maintenance states. Below this is a filter section and a timeline chart for 2016 showing event counts per month. The event log at the bottom shows two 'Event external_resource_state_change_complete' events at 11:30:58 AM and one 'Event RG_OFFLINE_STATE' event at 11:30:54 AM.



- In just a few minutes we were able to:
 - determine which resource(s) had a problem
 - identify the problem

The screenshot displays the IBM PowerHA SystemMirror for AIX monitoring interface. The main dashboard shows the cluster 'demo_cluster' with 2 nodes and 6 resource groups. The status indicators show 1 Critical, 0 Warning, and 0 Maintenance events for the nodes, and 4 Critical, 0 Warning, and 0 Maintenance events for the resource groups. A detailed event view on the right shows the 'Event node_down_complete' at 4:41:10 PM, with a description: 'The event script runs after the node_down event completes successfully. demo_cluster » phac1'. The interface also includes a calendar view for Thursday 18 August 2016, a search bar, and a filter section.



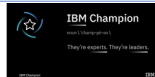
- In just a few minutes we were able to:
 - determine which resource(s) had a problem
 - identify the problem
 - solve the problem

The screenshot shows the IBM PowerHA SystemMirror for AIX management console. The main window displays the 'demo_cluster' logs. The 'Logs' tab is active, showing a list of log files on the left. The main content area is divided into three panels:

- Left Panel (Logs):** A list of log files including 'Events log hacmp.out', 'phac1 Thursday 08/18/2016 11:00:00 PM', 'phac2 Thursday 08/18/2016 11:00:00 PM', 'Cluster log cluster.log', 'Error report', 'Cluster utility log clutils.log', and 'Cluster manager log'.
- Middle Panel (phac1 Cluster verification log (clverify.log)):** Shows the output of a cluster verification command. The log indicates that the check passed successfully.

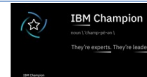

```

4817 BEGIN CHECK
4818
4819 END CHECK
4820 Check: PASSED
4821
4822
4823
4824 Total elapse time: 17 second(s)
4825
4826 Verification has completed normally.
4827
4828
4829 SUMMARY REPORT
4830
4831 Results of: clver -c -V normal -O -t -r -
4832
4833 Node      Errors Logged
4834 -----
4835 All       0
4836 phac1    0
4837 phac2    0
4838
4839 Totals   0
4840
            
```
- Right Panel (phac1 Events log (hacmp.out)):** Shows a sequence of events from the HACMP system. Key events include:
 - 1.87210: HACMP Event Preamb
 - 1.87211: Stop cluster services request with 'Grace'
 - 1.87212: Enqueued rg_move release event for resource
 - 1.87213: Enqueued rg_move release event for resource
 - 1.87214: Enqueued rg_move release event for resource
 - 1.87215: Enqueued rg_move release event for resource
 - 1.87216: Enqueued rg_move release event for resource
 - 1.87217: Enqueued rg_move release event for resource
 - 1.87218: Enqueued rg_move release event for resource
 - 1.87219: Node Down Completion Event has been enque
 - 1.87220: :check_for_site_down[+54] [{ high = high
 - 1.87221: :check_for_site_down[+54] version=1.4
 - 1.87222: :check_for_site_down[+55] :check_for_site
 - 1.87223: :check_for_site_down[+55] :check_for_site
 - 1.87224: :check_for_site_down[+55] :check_for_site

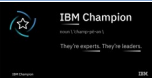


- In just a few minutes we were able to:
 - determine which resource(s) had a problem
 - identify the problem
 - solve the problem
 - and watch as PowerHA health returns.

The screenshot displays the IBM PowerHA SystemMirror for AIX monitoring interface. The top header shows 'IBM PowerHA SystemMirror for AIX' and 'Hello root'. The main content area is titled 'demo_cluster' and shows a status of 'Stable'. It includes tabs for 'Events summaries', 'Logs', 'General', 'Network', 'Storage', 'Replicated storage', and 'Applications'. A 'Snapshots' section shows '2 Nodes' and '6 Resource groups', each with '0' Critical, '0' Warning, and '0' Maintenance events. A 'Filter' section includes icons for Critical (red exclamation mark), Warning (yellow triangle), Maintenance (wrench), and Stable (green checkmark), along with a search bar and a 'Reset' button. A timeline for 2016 shows event counts: May (352), Jun (50, 24), Jul (122, 22), and Aug (427). A 'Today' section shows an event at 3:55:34 PM with the title 'Event RG_ONLINE_STATE' and the description 'Description is not available'.

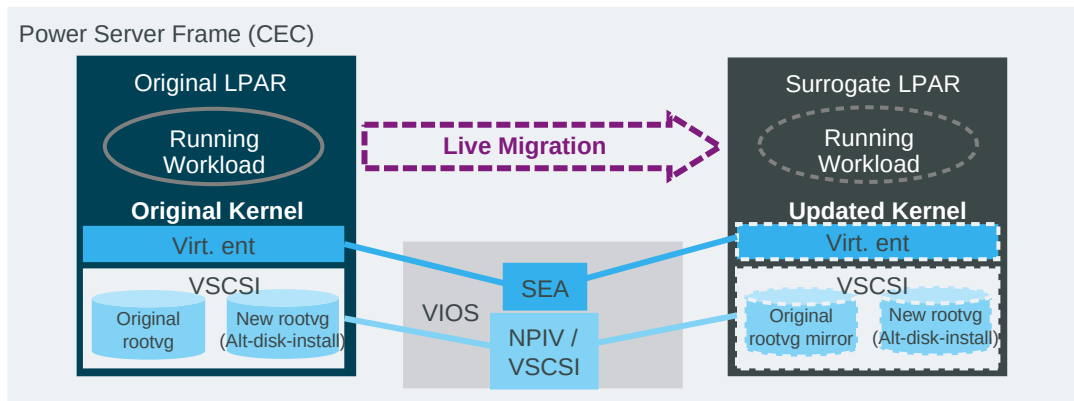


Management features



- C-SPOC vs lazy update
 - When changes to disks, volume groups, file systems are made through C-SPOC, all nodes that share the resource group will be updated
 - If the change is made outside C-SPOC, PowerHA will detect that a change has been made, so will export and re-import the volume group.
- AIX upgrades
 - Support for AIX live update
- PowerHA updates
 - method often depends on starting and final PowerHA versions
 - Rolling updates
 - snapshot conversions
 - easy update tool

- Allows live update of the AIX Kernel and device drivers via i-fix without a reboot
 - No restrictions on the types or number of interim fixes (i-fixes)
 - Uses existing maintenance model and administrative tools (NIM, emgr, geninstall)
 - Introduced in: AIX 7.2 TL0 SP0



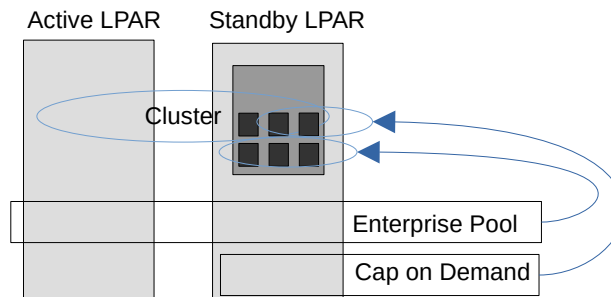
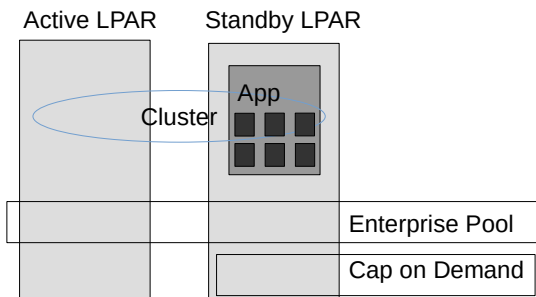
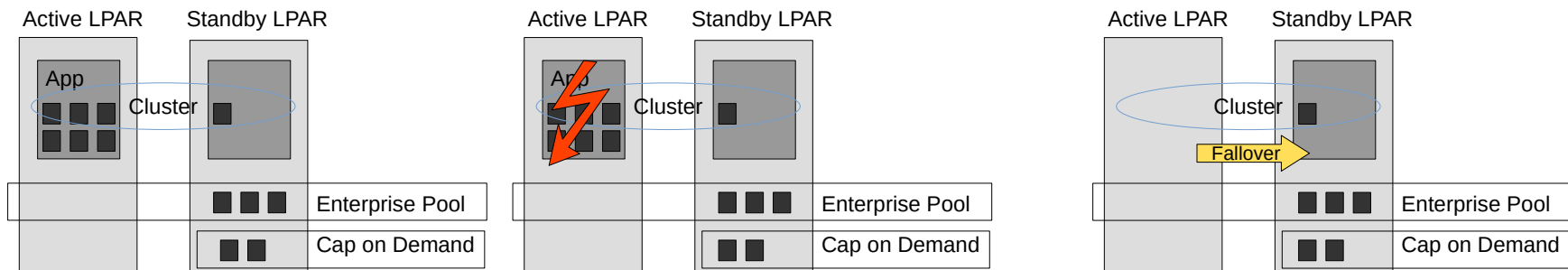
- Additional Software Requirements

- Requires Power firmware 730_66 or later, VIOS 2.2.3.50 or later
- PowerHA 7.2 & RSCT 3.2.1.0 are required for PowerHA support
- PowerSC 1.1.4.0 for environments using PowerSC

- Additional Information

- Deeper dive into AIX Live Update
<https://ibm.biz/BdHFKw>
- Hands on AIX Live Update article
<https://ibm.biz/BdHFTd>

- PowerHA Enterprise Pool and CoD integrated management
 - Significant cost savings in both hardware and software



Log file locations

Log name	Location	size	Description
cluster.log	/var/hacmp/adm		Good high level view, output from PowerHA scripts and daemons
cluster.mmddyyyy	/var/hacmp/adm/history		daily history file
hacmp.out	/var/hacmp/log	rolled .1-.7	Output of all resource group activity with summaries
grpqlsm	/var/ha/log	trimmed	RSCT group services messages
clverify.log, working files	/var/hacmp/clverify	Previous successful and unsuccessful verification attempts.	Previous successful and unsuccessful verification attempts.
clavan.log	/var/hacmp/log	Application availability logs	Details of when each application is started or stopped
clpol	/usr/es/sbing/cluster/etc	Priority override data	Priority override data
clutils.log	/var/hacmp/log	cluster utilities and file propagation	cluster utilities and file propagation
cl_testtool.log	/var/hacmp/log	rolled	Cluster test tool logs and hacmp.out summaries for each step
autoverify.log	/var/hacmp/log	autoverification logs	autoverification logs
discovery data	/usr/es/sbin/cluster/etc	cllip_config	cllip_config
sa.log	/var/hacmp/log	Cluster discovery logs	Smart assist logs
clcomd.log / clcomddiag.log	/var/log/clcomd	Logging for the clcomd daemon	Logging for the clcomd daemon
clinfo.log	/var/hacmp/log	rolled .1-.7	output from event scripts
cspoc.log .long .remote	/var/hacmp/log	Summary of all cluster activity	messages and output from all CSPOC commands
clstrmgr.debug .long	/var/hacmp/log	rolled .1-.7	

Review of updates



PowerHA 7.2.5

- Dec 2018 PowerHA SystemMirror for AIX 7.2.3
- Dec 2019 PowerHA SystemMirror for AIX 7.2.4
- PowerHA SystemMirror for AIX 7.2.5, Announced Oct 6 2020 and GA Dec 10
 - Geographic Logical Volume Manager (GLVM) configuration wizard
 - Configure and orchestrate multiple parallel GLVM instances, particularly for cloud deployments
 - Graphical User Interface (GUI) improvements
 - Support for cross-cluster verification
 - Manage administrative authorities' users through the GUI
 - Cluster deployment and management:
 - Removal of hostname dependency for cluster management
 - Bulk cluster notification enhancement
 - Activity log server hostname support
 - SSH server key implementation
 - Increased application resiliency during manage and unmanage activity
 - Cloud tie-breaker via lock file for on premise or cloud deployments
 - PowerHA currency support for IBM Tivoli Directory Server, IBM Tivoli Storage Manager, Oracle, IBM Db2, and SAP NetWeaver



Highlights from previous releases

- Resource optimised High Availability (ROHA)
 - Support for Enterprise Pool, CoD
 - flexible Capacity management across various Power pools
 - Enterprise Pool Integration
 - Manipulate Shared Processor Pool Sizes
 - New HMC Integration Tunables
- Advanced split handling policies
 - Node halt & disk fencing quarantine policies
 - Disk & NFS tie breaker support
 - Flexibility to avoid the need of a NAS backed device as tie breaker disk
 - Manual Fallover
- Automated repository replacement
 - Define multiple repository disks & auto replacement behaviour with AIX 7.2
- Network & Node Failure detection improvements
 - Fine granular controls to change failure detection times
 - Improved network interface failure detection methods
- Non-Disruptive Upgrade Support for HA
 - Cluster software upgrade without disruption to workload
 - Easy PowerHA Migration without downtime
- rootvg disk failure detection improvements
- AIX Live Update Support & LPM Support Enhancements
 - Adapts to LPM operation
 - Live update for AIX supported, but only one node at a time.

Highlights from previous releases (cont)

- Enhanced AIX and detailed cluster verification checks
 - Validation with a number of new checks including AIX Runtime Expert Settings (optional)
- Graphical Management
 - Cluster administration capabilities
 - Role based UI access control management
 - Multi tenancy support through HA zones
- Log Analyser tool
 - Analyse live or offline cluster FFDC information
 - Easy problem determination
 - Pattern based analysis support
- REST based HMC communication support
 - Setup for active node halt or ROHA policies using REST
- Novalink support (P8 and above)
- PowerHA (CAA) auto restart for VM restart DR
 - PVID based repository disk management
- Easy Update management
 - Local file system or NIM based update
- Quarantine Policies
 - Define Critical resource group
 - Split handling protections
 - HMC Node Halt Policy
 - SCSI-3 Node Fence Policy

Highlights from previous releases (cont)

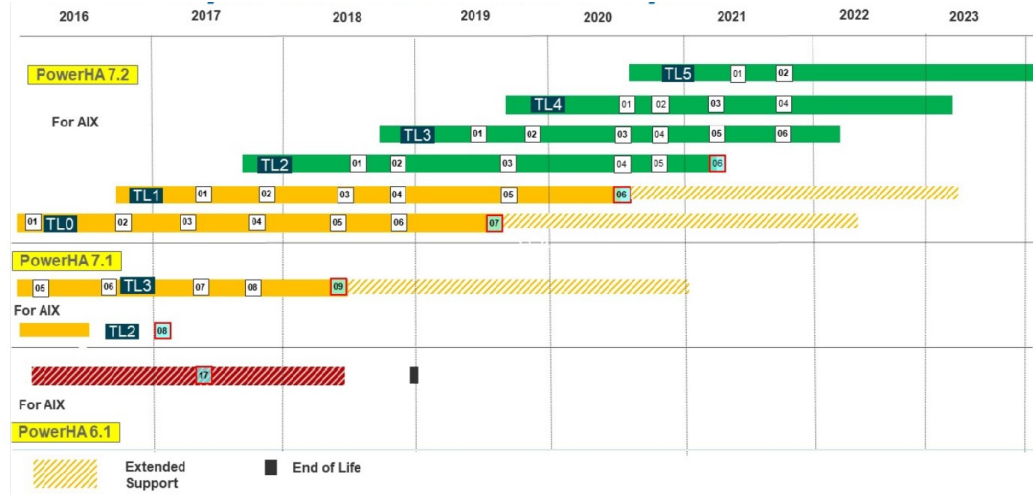
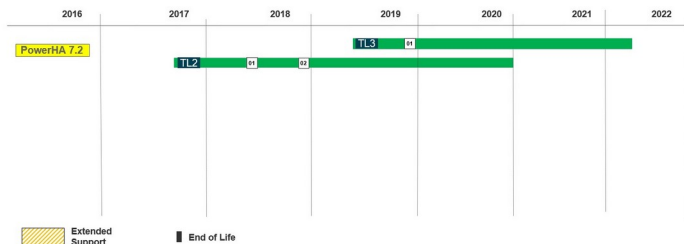
- **Cloud Backup Management**
 - Customers can now use IBM Cloud or Amazon AWS, as a data backup option. Collects the flash copy of the application data at the defined backup schedule, and copies to the IBM Cloud Object Storage S3 or Amazon AWS S3 storage using S3 boto interface.
- **Enhanced Split/Merge Policies**
 - Expanded choices of Split/Merge handling
 - Consistent policies for all types of clusters
 - Standard, Stretched, & Linked Clusters

	Pre-7.2.1 Split-Merge	7.2.1 Split-Merge
Std Cluster	None - Majority	None – majority TB(disk) – TB(disk) TB(NFS) – TB(NFS) Manual - Manual
Stretched cluster	None – Majority TB(disk) – TB(disk) TB(NFS) – TB(NFS)	
Linked cluster	None – Majority TB(disk) – TB(disk) TB(NFS) – TB(NFS) Manual - Manual	

- Miscellaneous tidy
 - Some customers requested use of rsyslog (AIX uses syslog by default)
 - C-SPOC support for LVM preferred read
 - Oracle now has options for shutting down oracle (immediate, abort, normal, transactional)
 - Administration events, so that administrator initiated tasks are logged and handled the same way as PowerHA events.
 - Event serial number to assist tracking through the various PowerHA logs
 - Improvements in recovering from script failure and option to cancel remaining event processing
 - Failures in pre/post/notify event processing can stop the processing of the event
 - Improvements in the log analyser (added options and strings not case sensitive)
 - Network flapping events – define a threshold to be reached before responding

- After 7.2.2 was released, a review conducted:
 - Now that I know where the problem is, I want to be able to fix it.”
 - Health dashboard with solutions
 - The ability to take actions on the managed clusters was asked for, which implied a higher level of security and organisation
 - User management
 - Role-based access control
 - Cluster zones
- Changes introduced
 - 5 customers were involved in the research
 - Design thinking principles incorporated
 - Adds more administrative capabilities
 - Modify existing cluster configuration objects
 - Delete existing cluster configuration objects
 - Installation improved
 - Highly available

- PowerHA release Life cycle strategy
 - Implement technology level release strategy
 - Major releases as necessary
 - Minor release updates (Technology Level 0 to Major release)
 - At least two technology levels per major release
 - Considering
 - Additional 2 year service offering for last TL
- New command
 - `halevel -s`



IBM PowerHA SystemMirror V7.2.5 AIX Administrator Specialty



- A candidate pursuing the PowerHA SystemMirror Administrator Specialty has knowledge and experience deploying PowerHA clusters, both within and across data centres, and a good understanding of various deployment methods. They should have hands-on experience administering the AIX operating system, have a good understanding of SAN and network infrastructures, and have some practical experience in deploying applications that are controlled by the cluster software.
- The candidate is a technical professional who can:
 - Install, deploy, and configure PowerHA clusters on AIX.
 - Identify and troubleshoot issues in a deployment at the various levels of the solution: OS, RSCT, CAA or PowerHA cluster level.
 - Tune or customise the solution to achieve desired behaviour, including operating system, CAA, or PowerHA settings.
- Details of exam
 - Number of questions: 40
 - Number of questions to pass: 26
 - Time allowed: 75 mins.

<https://www.ibm.com/certify/specialty?id=S1000-008>

- Exam breakdown
 - 20% Planning
 - 23% Configuration and deployment
 - 35% Cluster management, administration and upgrade
 - 22% Customisation and troubleshooting
- Sample exam available – has 7 questions
- Preparation
 - Keep calm
 - Read the questions / answers carefully
 - Check the sample exam for benchmark

Professional
Certification
Program**Exam S1000-008 - IBM PowerHA
SystemMirror v7.2.5 for AIX Administrator
Specialty**

1. Which two features are supported for heartbeat communication in V7 clusters?
 - A. TMSCSI
 - B. CSPOC
 - C. SWWCONN
 - D. IP Multicast heartbeating
 - E. IP Unicast heartbeating
2. Which configuration file should be updated with the IP addresses that will be used for initial communication between nodes of each cluster member prior to the initial cluster synchronization?
 - A. /etc/cluster/rhosts
 - B. /etc/resolv.conf
 - C. /etc/netshvc.conf
 - D. /etc/cluster.conf

<https://www.ibm.com/certify/specialty?id=S1000-008>

References



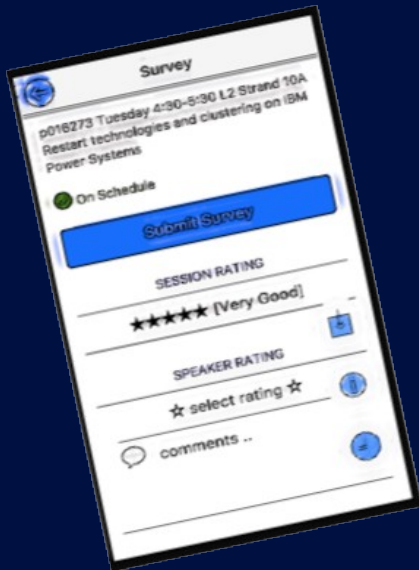
- IBM PowerHA SystemMirror for AIX 7.1.3 Best Practices and Migration Guide
 - <http://www.redbooks.ibm.com/abstracts/sg248234.html?Open>
- IBM PowerHA for AIX community
 - <https://community.ibm.com/community/user/power/communities/community-home?CommunityKey=32520692-50a7-4829-91c5-038ab4df00b1>
- PowerHA SystemMirror for AIX Reference Information
 - <https://www.ibm.com/docs/en/powerha-aix/7.2?topic=reference-information>
- IBM Support – subscribe to notifications
 - <https://www.ibm.com/support/pages/node/718119>
- IBM PowerHA SystemMirror V7.2 for IBM AIX Updates
 - <http://www.redbooks.ibm.com/Redbooks.nsf/RedpieceAbstracts/sg248278.html?Open>
- IBM PowerHA SystemMirror V7.2.3 for IBM AIX SG24-8434-00
 - <http://www.redbooks.ibm.com/abstracts/sg248434.html?Open>
- Implementing IBM VM Recovery Manager for IBM Power Systems
 - <http://www.redbooks.ibm.com/abstracts/sg248426.html?Open>
- Basic Requirements for HA and DR (on-prem and Cloud) (published soon)

- PowerHA Release Notes
 - <https://www.ibm.com/docs/en/powerha-aix/7.2?topic=release-notes>
- PowerHA Comments & Questions:
 - hafeedbk@us.ibm.com
- IBM FLRT
 - <https://www14.software.ibm.com/webapp/set2/flrt/power>

- Jaqui Lynch's good AIX practices and VIO maintenance
 - <https://www.youtube.com/channel/UCYH6OdgB6rV1rPxYt6FWHpw>
- Shawn Bodily's PowerHA implementation, management and upgrade
 - <https://www.youtube.com/channel/UCK-QtfzCBTQRihxQLjfDVHg>
 - Non-disruptive upgrade from 713 to 72
 - <https://www.youtube.com/watch?v=1Kzm7I2mRyE>
 - AIX Live update (PowerHA node)
 - <https://www.youtube.com/watch?v=BJAnpN-6Sno>
 - Automated Repository Replacement
 - <https://www.youtube.com/watch?v=HJZZDCXLwTk>
 - rootvg Failure monitoring
 - <https://www.youtube.com/watch?v=OZcrhVGvkBg>
- Michael Herrera's video on ROHA: Flexible Capacity Management
 - <https://www.youtube.com/watch?v=G-zefev-XYU>

Thank you a203872 - PowerHA introduction and certification preparation

Antony (Red) Steel
antony.steel@belisama.com.sg



**Please don't forget to complete the
session evaluation!**

© 2021 International Business Machines Corporation. No part of this document may be reproduced or transmitted in any form without written permission from IBM.

U.S. Government Users Restricted Rights — use, duplication or disclosure restricted by GSA ADP Schedule Contract with IBM.

Information in these presentations (including information relating to products that have not yet been announced by IBM) has been reviewed for accuracy as of the date of initial publication and could include unintentional technical or typographical errors. IBM shall have no responsibility to update this information. **This document is distributed “as is” without any warranty, either express or implied. In no event, shall IBM be liable for any damage arising from the use of this information, including but not limited to, loss of data, business interruption, loss of profit or loss of opportunity.** IBM products and services are warranted per the terms and conditions of the agreements under which they are provided.

IBM products are manufactured from new parts or new and used parts. In some cases, a product may not be new and may have been previously installed. Regardless, our warranty terms apply.”

Any statements regarding IBM's future direction, intent or product plans are subject to change or withdrawal without notice.

Performance data contained herein was generally obtained in a controlled, isolated environments. Customer examples are presented as illustrations of how those

customers have used IBM products and the results they may have achieved. Actual performance, cost, savings or other results in other operating environments may vary.

References in this document to IBM products, programs, or services does not imply that IBM intends to make such products, programs or services available in all countries in which IBM operates or does business.

Workshops, sessions and associated materials may have been prepared by independent session speakers, and do not necessarily reflect the views of IBM. All materials and discussions are provided for informational purposes only, and are neither intended to, nor shall constitute legal or other guidance or advice to any individual participant or their specific situation.

It is the customer's responsibility to insure its own compliance with legal requirements and to obtain advice of competent legal counsel as to the identification and interpretation of any relevant laws and regulatory requirements that may affect the customer's business and any actions the customer may need to take to comply with such laws. IBM does not provide legal advice or represent or warrant that its services or products will ensure that the customer follows any law.

- Information concerning non-IBM products was obtained from the suppliers of those products, their published announcements or other publicly available sources. IBM has not tested those products about this publication and cannot confirm the accuracy of performance, compatibility or any other claims related to non-IBM products. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products. IBM does not warrant the quality of any third-party products, or the ability of any such third-party products to interoperate with IBM's products. **IBM expressly disclaims all warranties, expressed or implied, including but not limited to, the implied warranties of merchantability and fitness for a purpose.**
- The provision of the information contained herein is not intended to, and does not, grant any right or license under any IBM patents, copyrights, trademarks or other intellectual property right.
- IBM, the IBM logo, ibm.com and [names of other referenced IBM products and services used in the presentation] are trademarks of International Business Machines Corporation, registered in many jurisdictions worldwide. Other product and service names might be trademarks of IBM or other companies. A current list of IBM trademarks is available on the Web at "Copyright and trademark information" at: www.ibm.com/legal/copytrade.shtml





Backup slides

