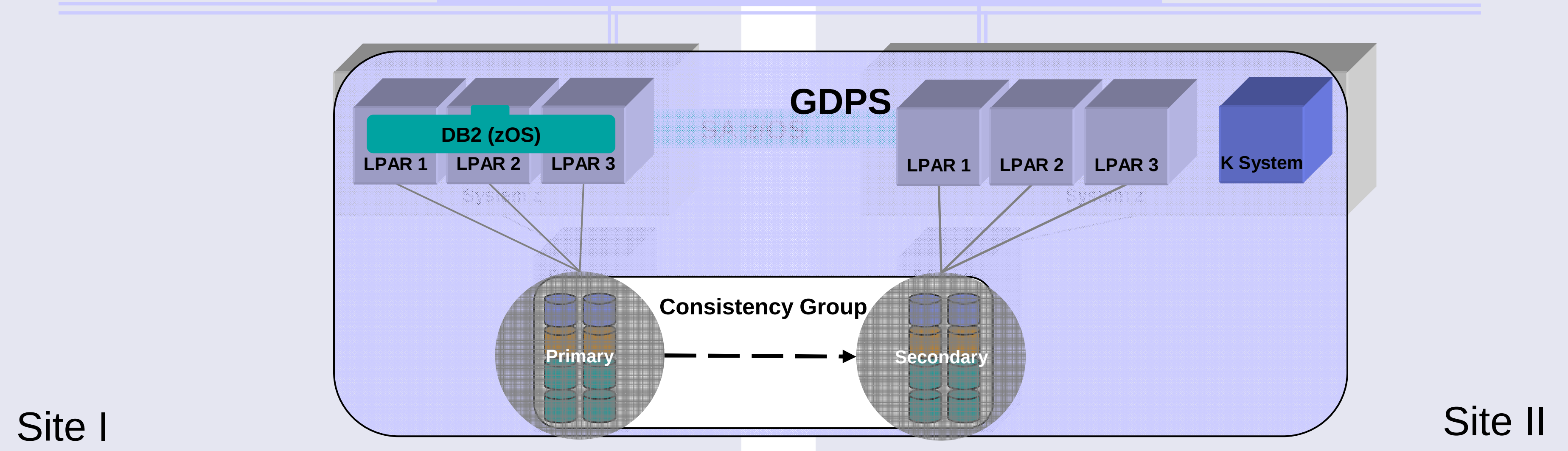
System Automation and High Availability with z/VM and Linux for System z

Overview

- xDR – GDPS/PPRC Multiplatform Resiliency for zSeries
  - GDPS
  - Mirroring Techniques
  - Re-IPL, Site-Takeover, HyperSwap
  - xDR for z/VM
  - xDR for Linux on System z running native in LPAR
- DCM/DDR – Distributed Cluster Management - Distributed Disaster Recovery
  - SA Integration into GDPS: DCM → DDR
  - Supported Scenarios
  - Supported Setups

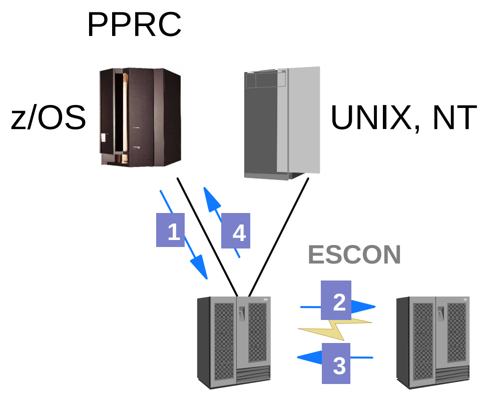
Geographically Dispersed Parallel Sysplex

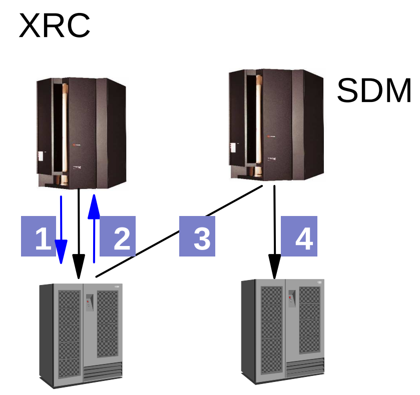
- GDPS is the DR solution for the mainframe
  - SA z/OS is a central part of GDPS
- GDPS features
  - Disk error detection
  - Heartbeat for sanity checks
  - Re-IPL in place
  - Coordinated Site Takeover
  - Coordinated HyperSwap



Mirroring via PPRC / XRC

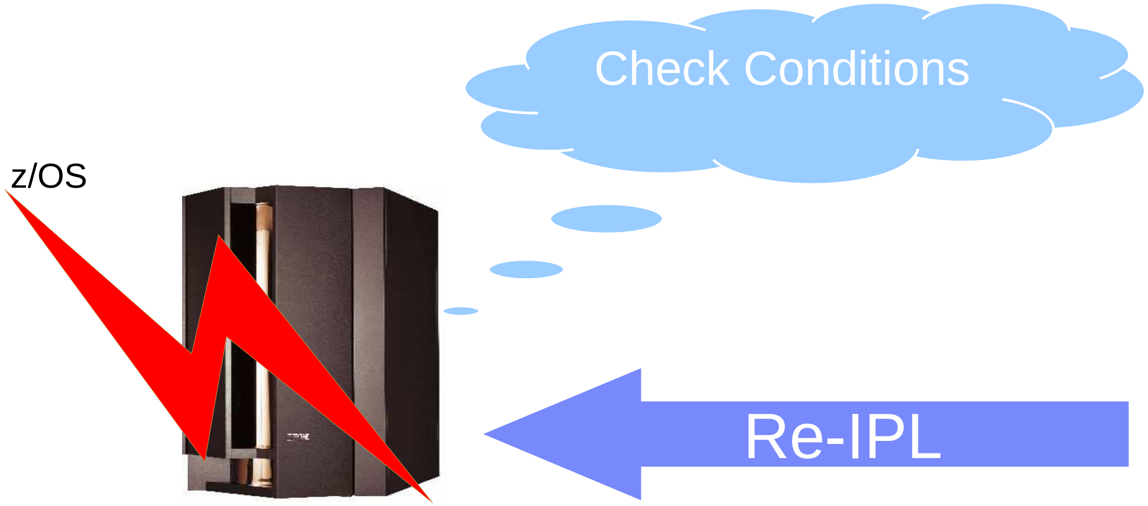
- Metro-Mirror – PPRC – Peer to Peer Remote Copy
  - Synchronous mirroring
  - Multisite Sysplex (fiber distance between sites 100 km - max)
  - No or limited data loss in unplanned failover - user policy
  - Planned and Unplanned reconfiguration support
- XRC – Extended Remote Copy
  - Asynchronous mirroring
  - Supports any distance
  - Production systems in Site 1
  - Limited data loss to be expected in unplanned failover
  - GDPS initiates restart of production systems in Site 2
- Global-Mirror
  - Asynchronous mirroring – Storage based





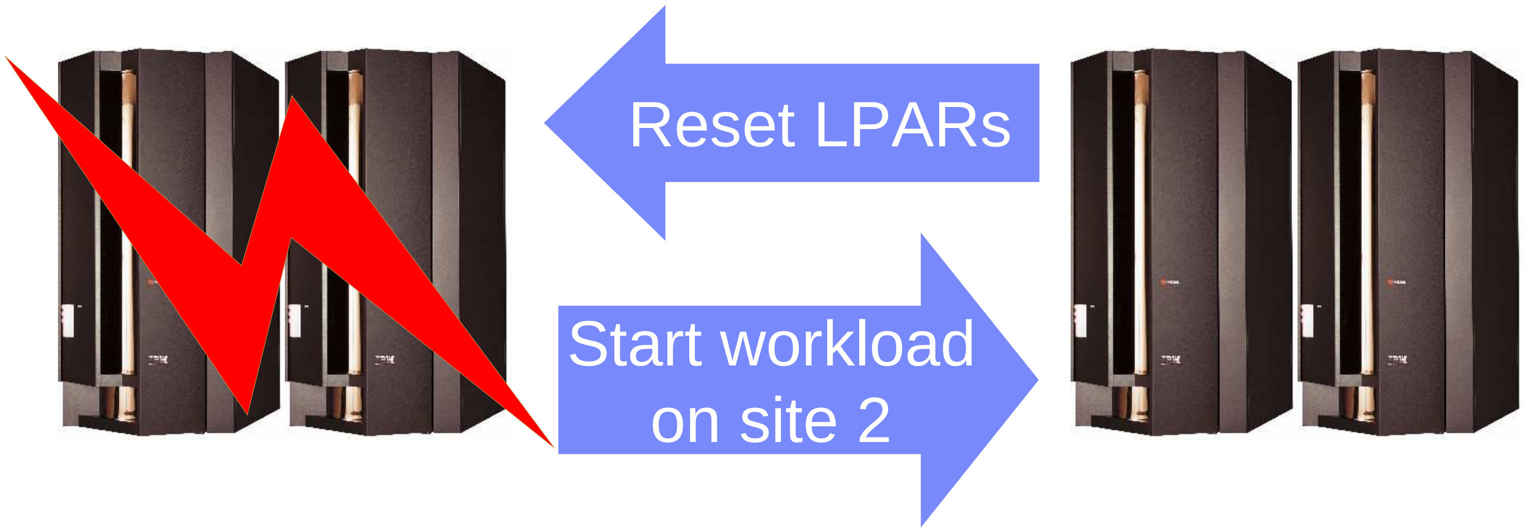
Re-IPL in Place

- when GDPS detects that a z/OS system is no longer active, it verifies whether the policy definition indicates that Auto IPL has been enabled, that the threshold of the number of IPLs in the predefined time window has not been exceeded, and that no planned action is active. If these conditions are met, GDPS automatically re-IPLs the system in place, brings it back into the Parallel Sysplex, and restarts the workload.



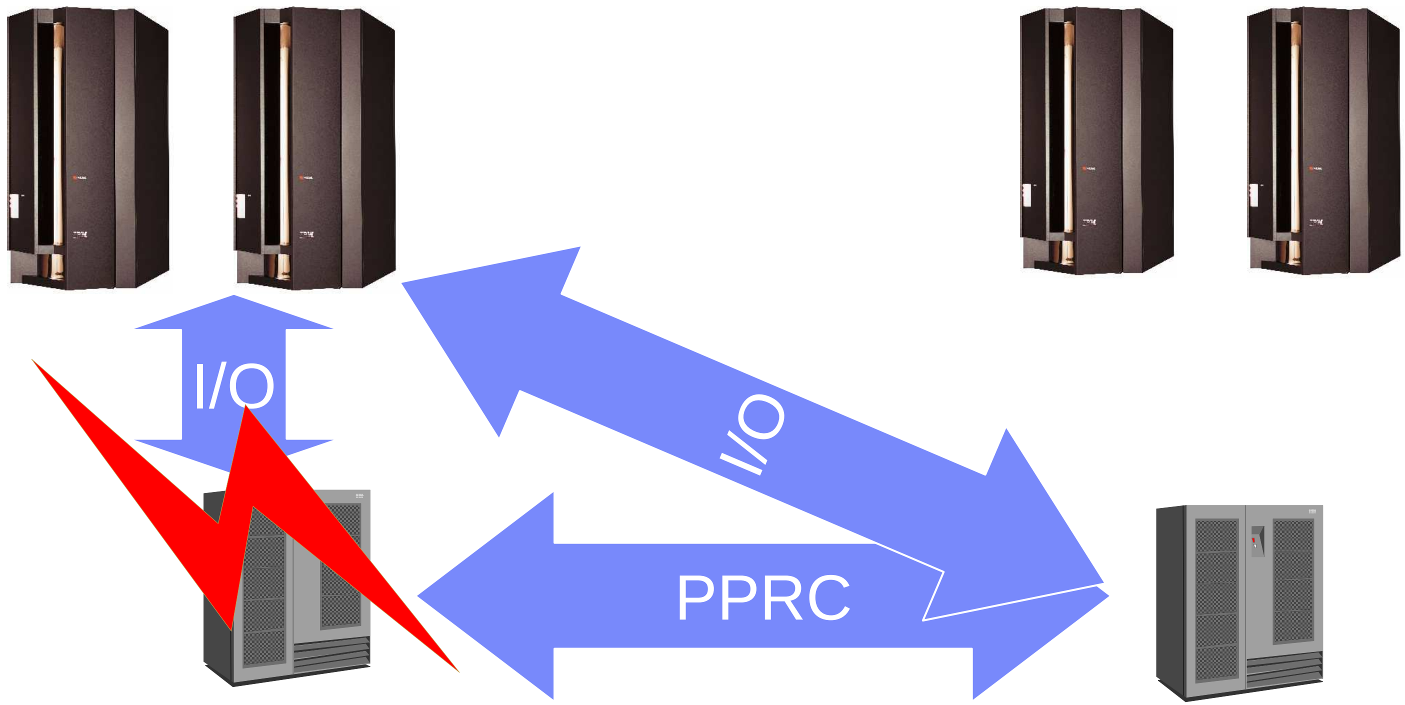
Site takeover

- **when GDPS detects a complete site failure, it can offer a takeover of the workload on site 2.**



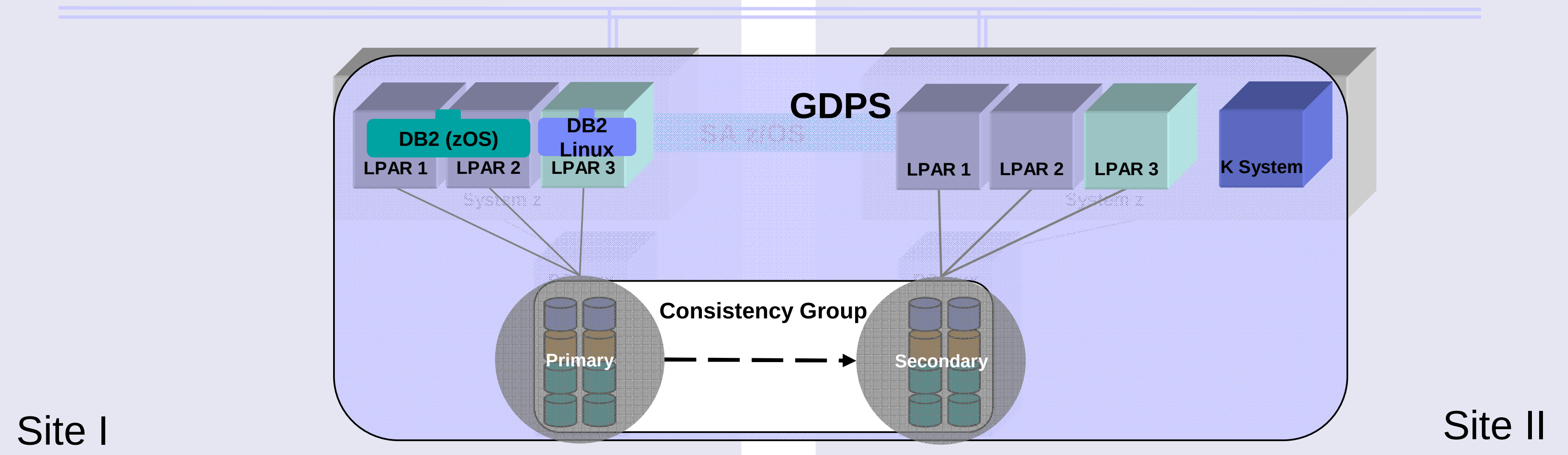
HyperSwap

- when GDPS detects an disk outage, the logical links to the primary disks can be switched to the secondary disks
- Can swap large number of devices very fast
- Includes all volumes attached to the system
- Non-disruptive – applications keep using same device addresses
- Planned (Maintenance) and Unplanned (Continuous Availability) Scenario



xDR: SA MP's Integration with GDPS Metro Mirror

- GDPS/Metro Mirror Multi Platform Resiliency for System z (xDR)
  - xDR extends GDPS to not only support z/OS but also Linux on System z (both on z/VM and on LPARs)
  - SA MP is the foundation of xDR
  - xDR provides for Linux on System z:
    - Disk error detection, Heartbeat for sanity checks, Re-IPL in place, coordinated site takeover, coordinated Hyperswap

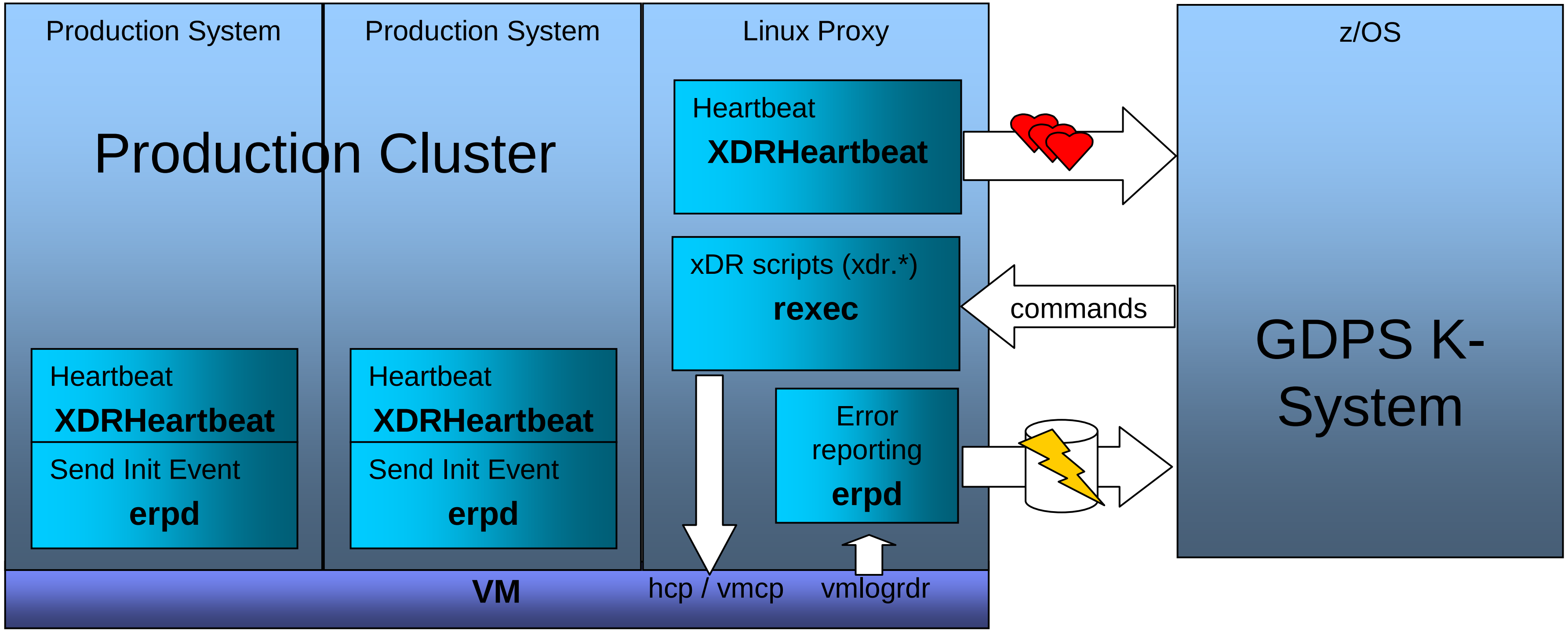


xDR – Cross Platform Disaster Recovery

- xDR extends GDPS to support not only z/OS but also Linux on zSeries
  - Disk error detection
  - Heartbeat for sanity checks
  - Provide Logical Layer to access disks
  - Command Interface for Re-IPL, HyperSwap, Stop Systems, etc.
  - Coordinated Disaster Recovery across OS Platforms
- xDR is not a Product
  - In order to implement xDR, the correct versions of the following products are needed:
    - GDPS
    - System Automation for Multiplatforms
    - Linux
    - VM, if Linux is running on z/VM
- Offical Name: „GDPS/PPRC Multiplatform Resiliency for zSeries“

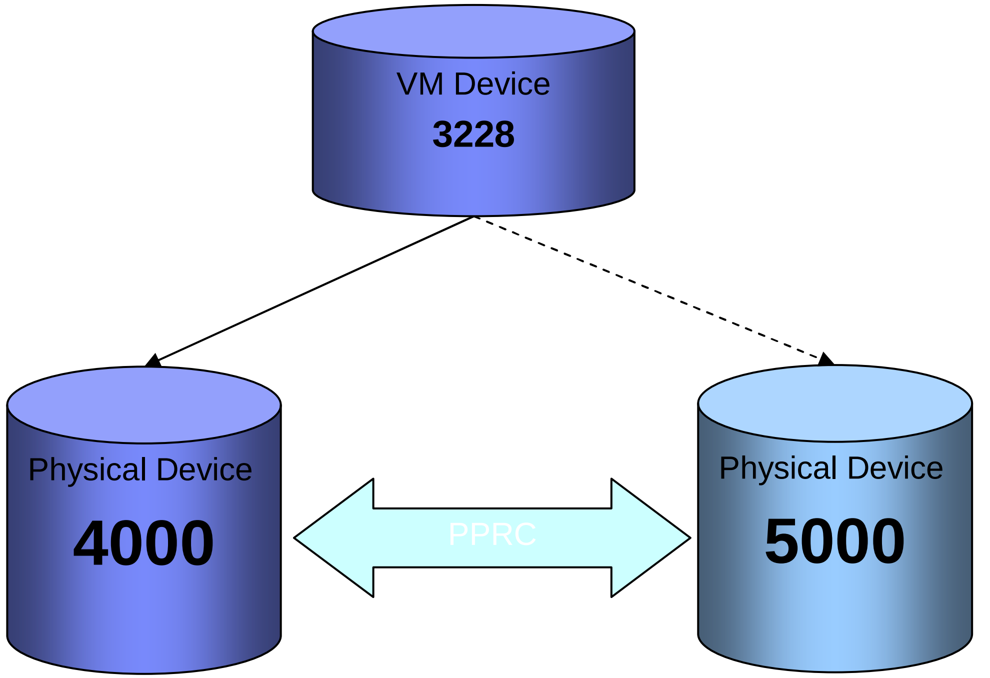
xDR on z/VM

- Proxy
  - One linux system is configured as Proxy for GDPS which has special configuration (Memory locked, Access rights to VM, One-Node-Cluster)
  - Heartbeat for sanity check
  - erpd sends system information and reports disk errors to GDPS
  - CLI via rexec
- Production Nodes
  - Heartbeat for sanity check
  - erpd sends system information
  - CLI via rexec
- The command interface to VM CP is hcp (for SLES8,SLES9) or vmcp (for SLES10)
- The interface to retrieve disk errors from VM is vmlogrdr (linux device)



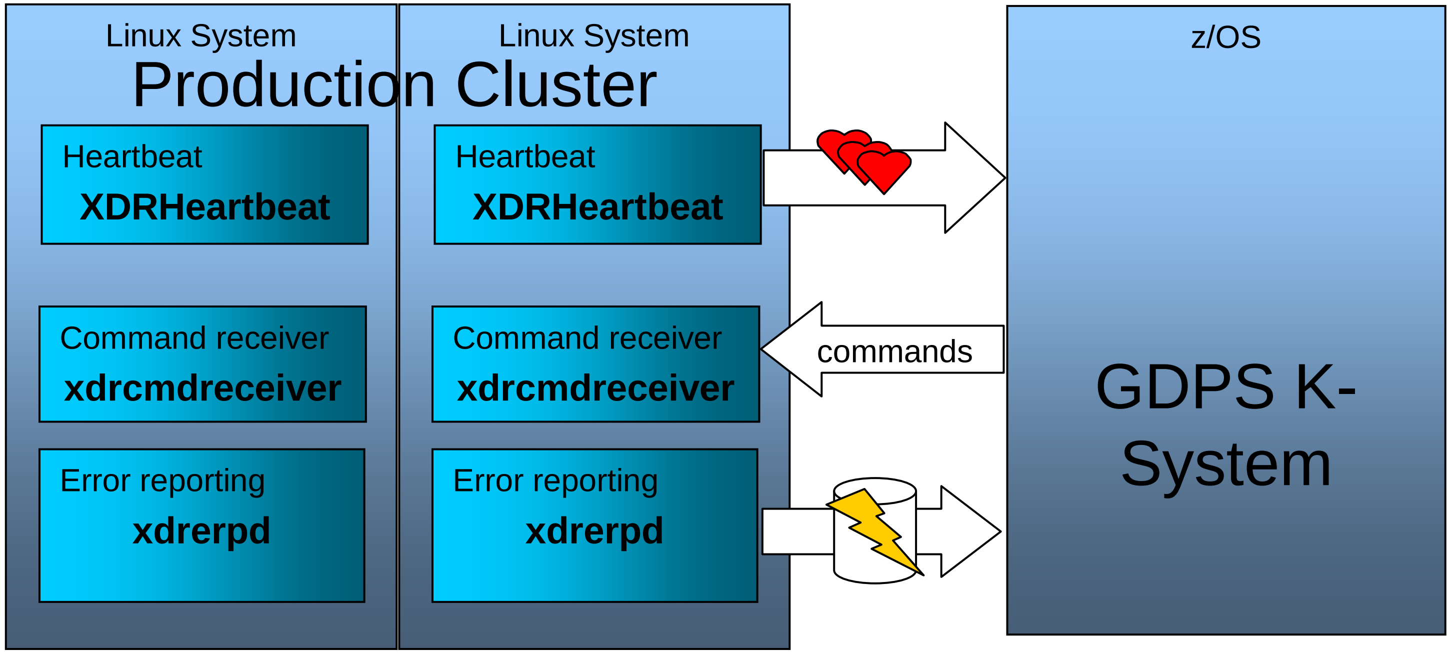
HyperSwap Implementation in z/VM

- Standard VM mechanisms are used for logical mapping
- VM provides a CP interface that allows to configure PPRCed devices and do the HyperSwap



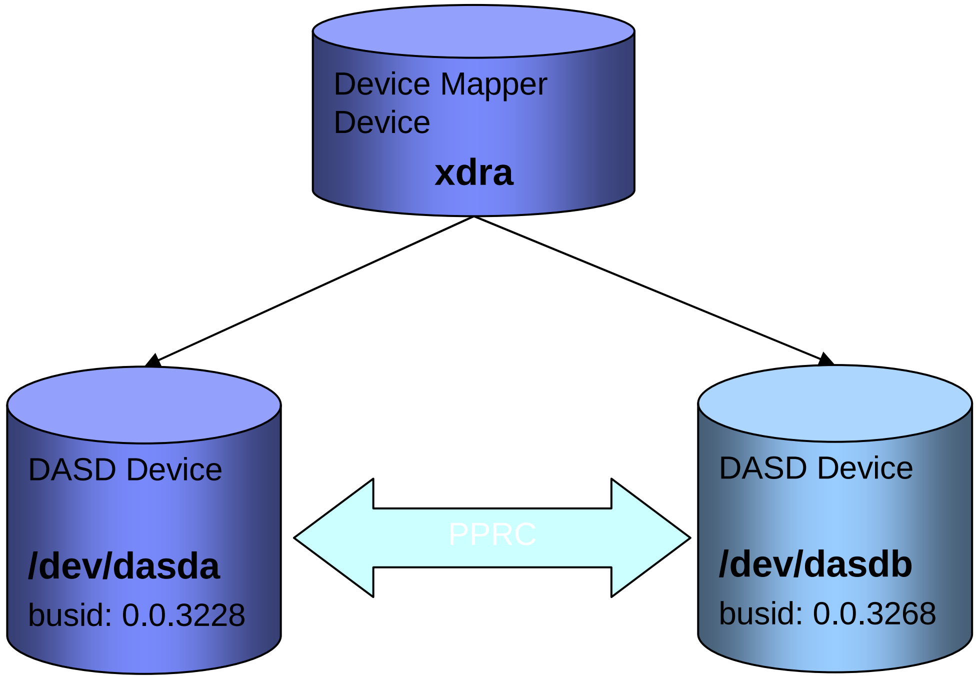
xDR on Linux running native in LPAR

- There is no proxy.
- Heartbeat is used for sanity check
- Each system provides a CLI for GDPS using the Command receiver
- Each system reports disk errors using the error reporting daemon
- The interface to retrieve DASD disk errors is the character device /dev/dasd_eer which is provided by the DASD driver
- The interface to retrieve SCSI disk errors is a NETLINK interface to the device mapper



HyperSwap Implementation in Native Mode

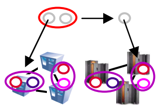
- Linux device mapper is used for logical mapping (target is multipath)
- Multipath tools are used to set up the device pairs during IPL automatically
- Explicit control commands to the DASD devices and the device mapper are used to do the HyperSwap



Enterprise-wide Automation with

IBM Tivoli System Automation Application Manager 3.1 (Distributed Disaster Recovery)

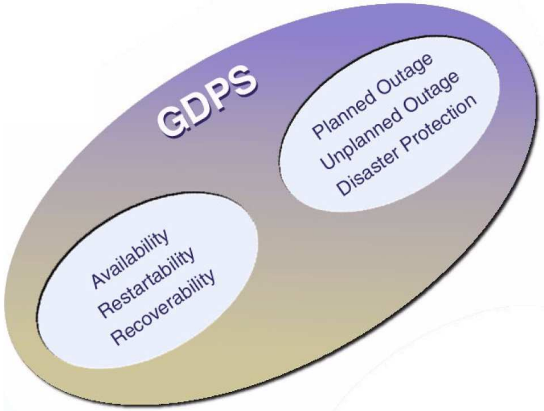
SA Application Manager ⇔ GDPS integration

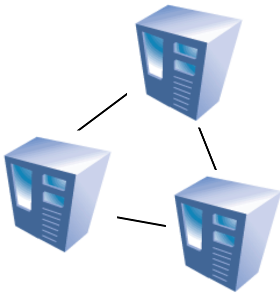


IBM Tivoli System Automation Application Manager

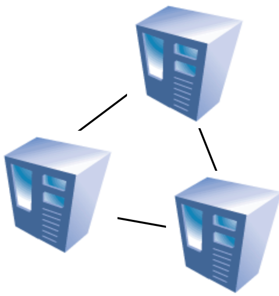
Tivoli software

Distributed Disaster Recovery

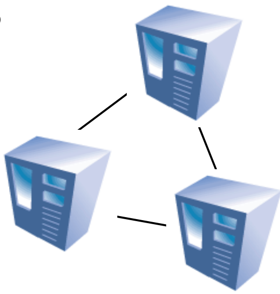




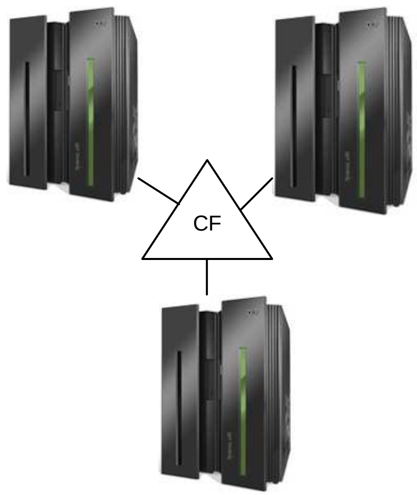
AIX HACMP



Windows MSCS



Linux SA MP



z/OS Sysplex

Distributed Disaster Recovery - Introduction

- SA AM / SA MP ➔ Manages Applications
  - „Resources“ are kept high available (without knowledge of sites)
  - End-to-End scope, cross cluster dependencies, resource grouping to arbitrary abstraction level

- GDPS ➔ Manages Systems
  - IT Infrastructure is dispersed across sites
  - System z scope for servers
  - System z and open systems scope for data replication

- System Automation Manager for Distributed Disaster Recovery
  - Integration with GDPS-DCM
  - Coordinated automation tasks (Site Maintenance, Site Failover etc.)
    - Single point of control for site switching
  - Alerting of serious outages (Cluster faults, Application faults)
    - Could be the beginning of a rolling disaster detected first on open systems, can lead to GDPS takeover prompt
  - First release will support metro mirror distance, global mirror distance will be later

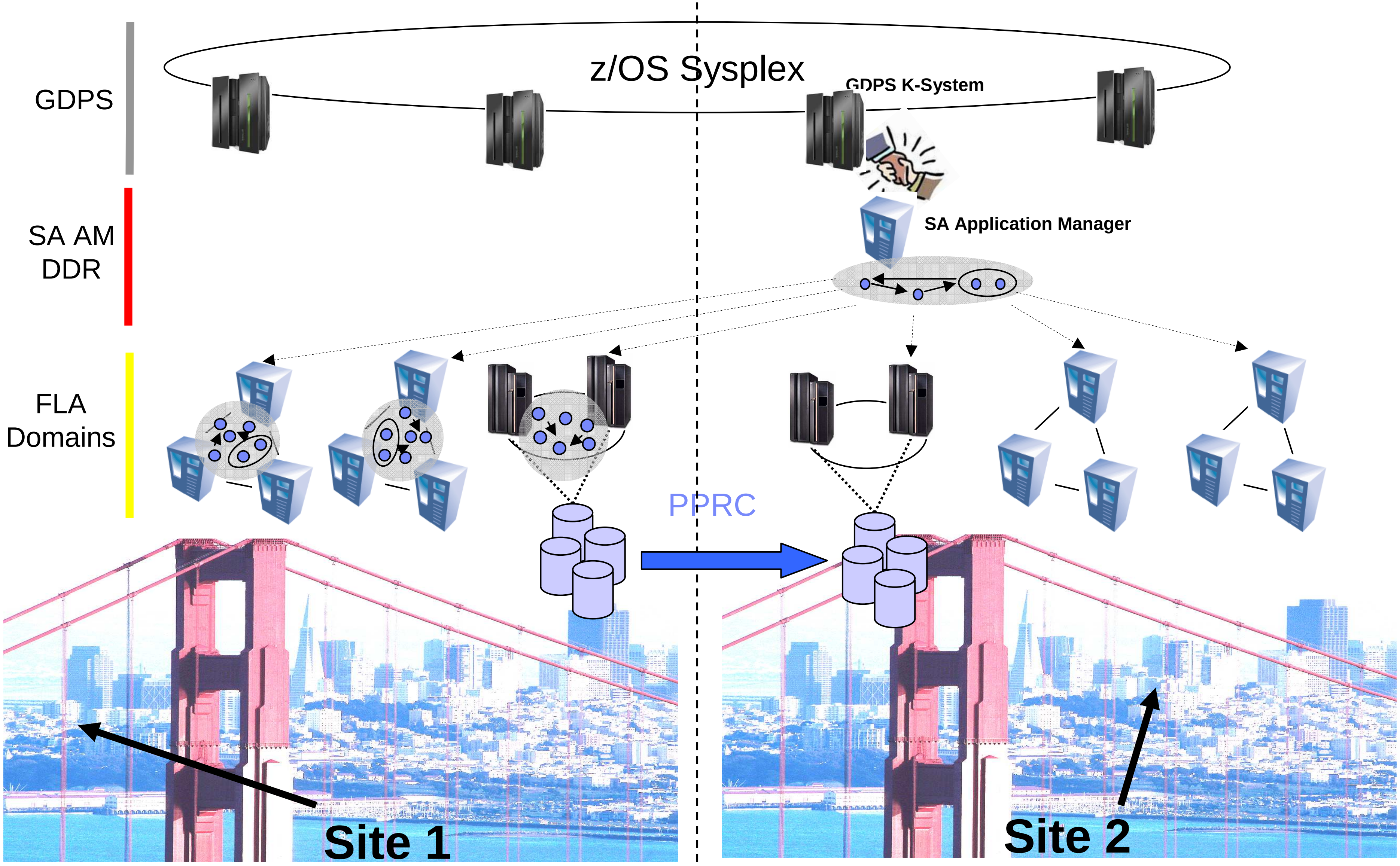
Distributed Disaster Recovery as Feature

- The Distributed Disaster Recovery functionality of the SA Application Manager has to be enabled with help of an additional priced feature license

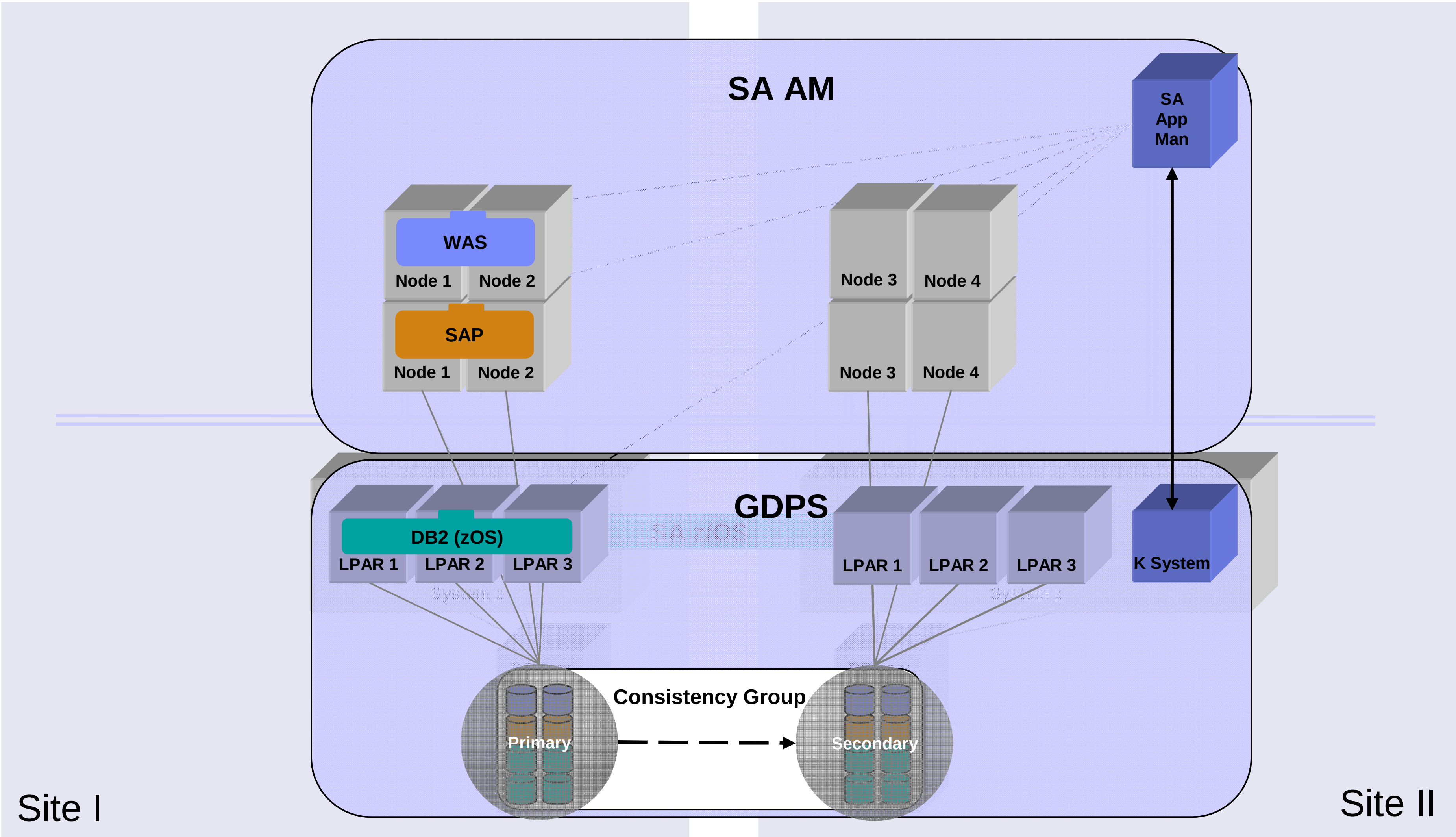
```
IBM Tivoli System Automation end-to-end automation engine
Version: 3.1.0.0.075001, NO_APAR
Features: SA AM for Distributed Disaster Recovery

Usage:
eezdmn [option]
  -START                Starts the automation engine
  -SHUTDOWN | -SHUTD    Stops the automation engine
  -MONITOR  | -M        Displays the current state
  -RECONFIG | -R        Re-configures the automation engine
  -CO                   Starts only the EIF2JMS conversion thread
  -XD ("*" | "<RES_NAME>[,<RES_NAME>]") <DUMPFILE>
                        Dumps (all | specific) resources to a file
  -INSTCERT <fully qualified path to license file>
                        Installs the specified license certificate
```

SA Application Manager - Distributed Disaster Recovery



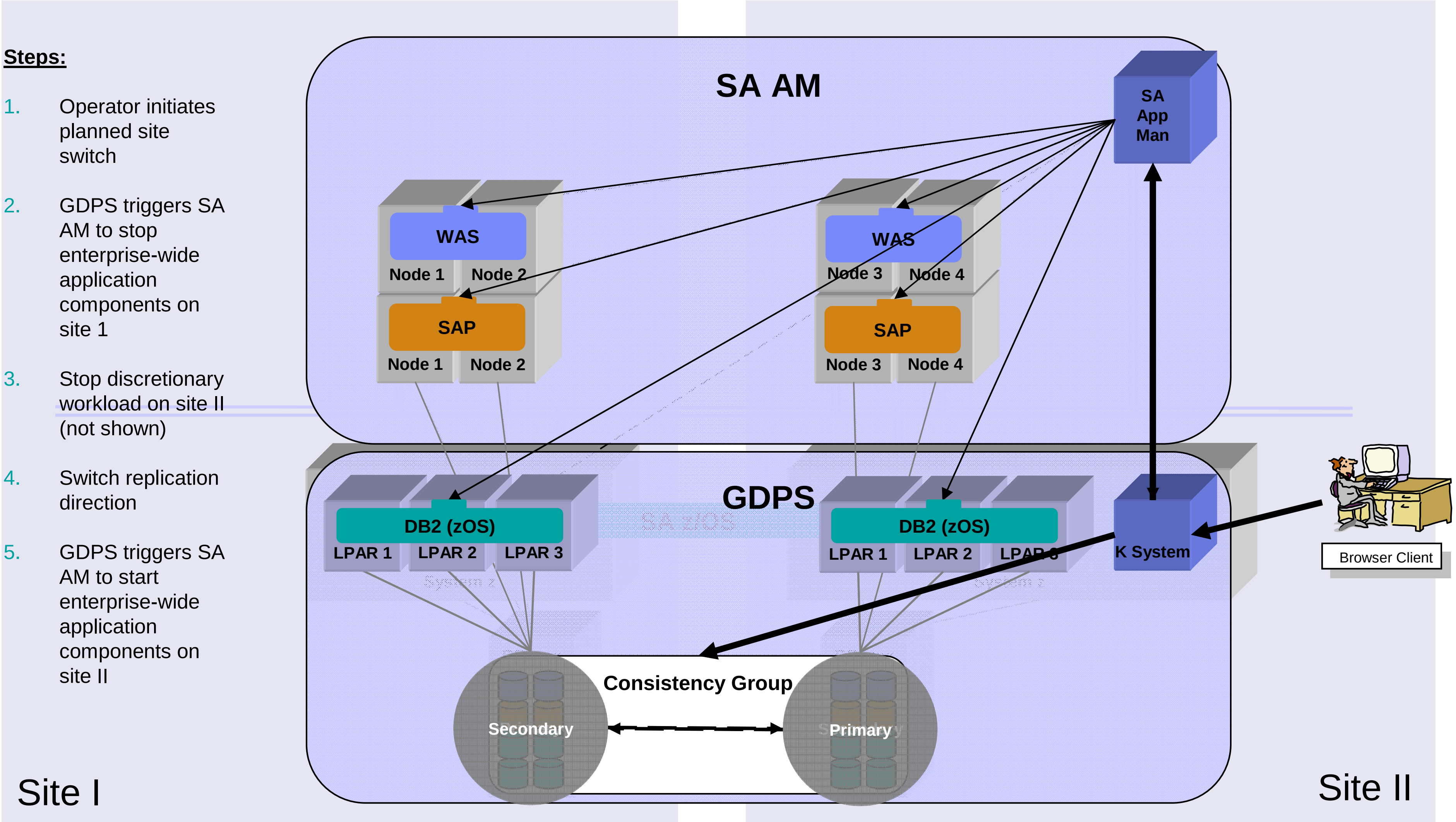
GDPS/SA AM (DDR): An Overview



GDPS/SA AM (DDR) – Planned Site Switch

Steps:

1. Operator initiates planned site switch
2. GDPS triggers SA AM to stop enterprise-wide application components on site 1
3. Stop discretionary workload on site II (not shown)
4. Switch replication direction
5. GDPS triggers SA AM to start enterprise-wide application components on site II

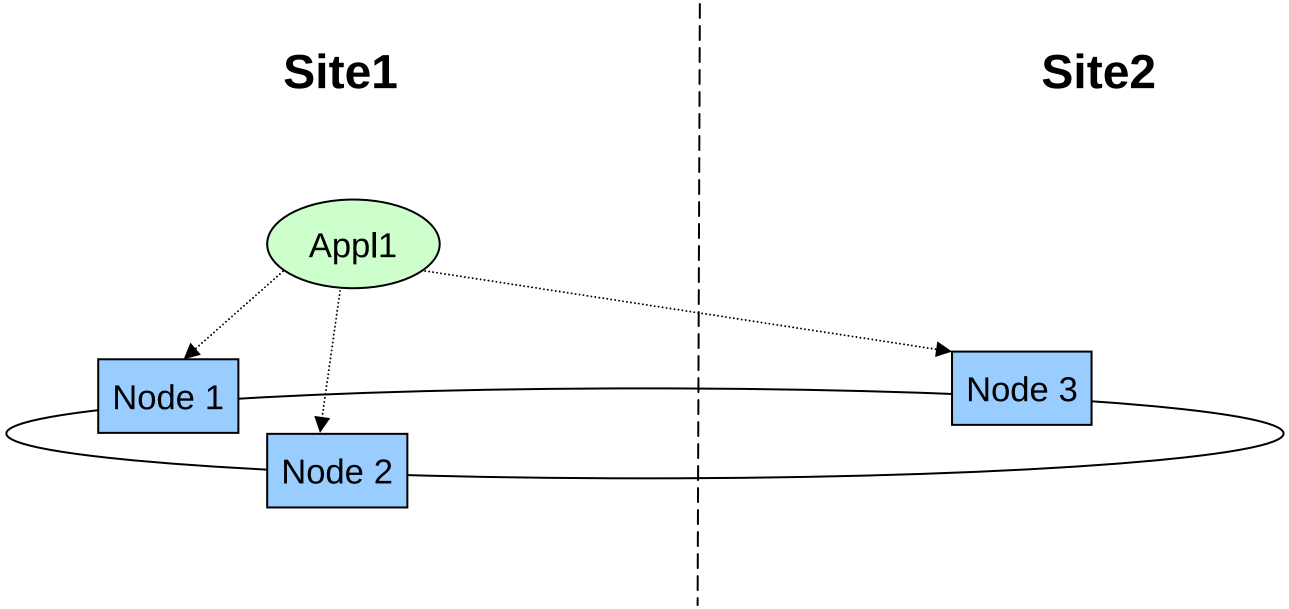


Supported Scenarios

- Coordinated automation tasks
  - System Maintenance
    - Planned Site Switch and Switch Back
    - Maintenance of Backup Site
  - Disk Subsystem Maintenance
    - Planned HyperSwap and Swap Back
  - System or Cluster Failure
    - Unplanned Site Failover and Failback
    - Loss of Backup Site
  - Disk Failure
    - Unplanned HyperSwap and Swap Back
- Alerting of serious outages
  - Cluster Failures
  - Failures of business critical Applications

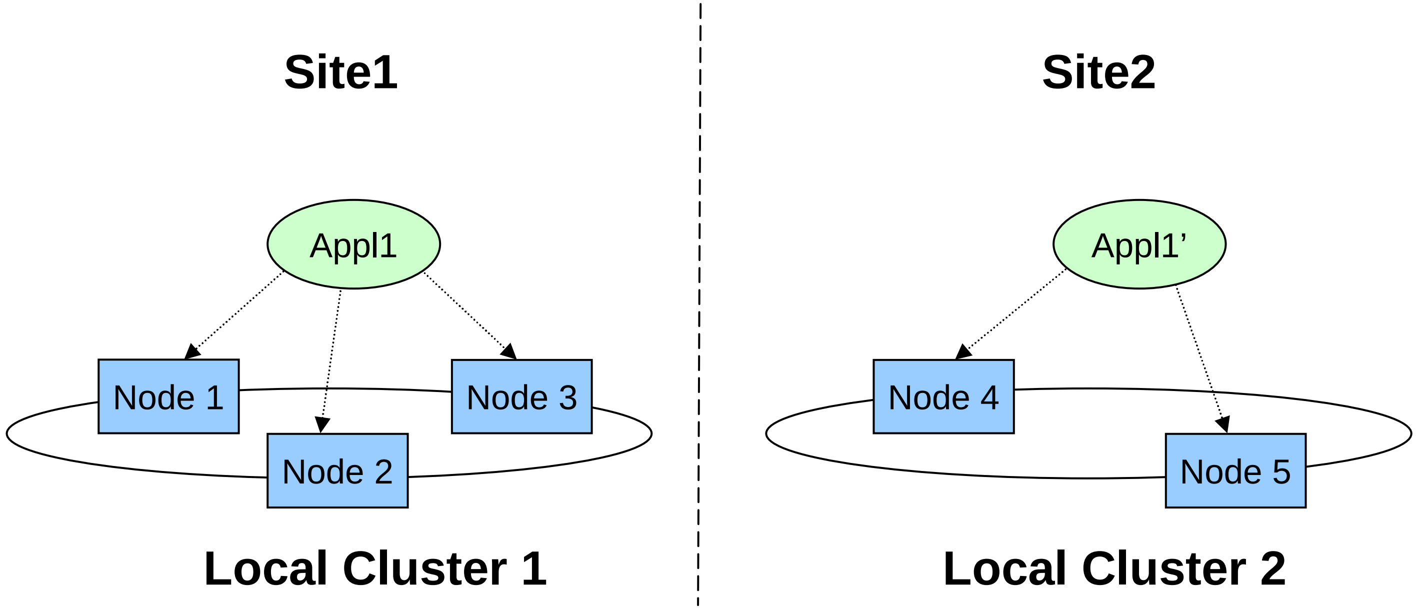
Supported Setups

- Stretched Cluster
  - Active-Active
    - Resources are running concurrently on all nodes at the same time
    - When a site fails, the resources don‘t move to the other site
  - Active-Passive
    - Resources can move to the other site
  - Disk Replication managed by First Level Automation (e.g. HACMP/LVM)

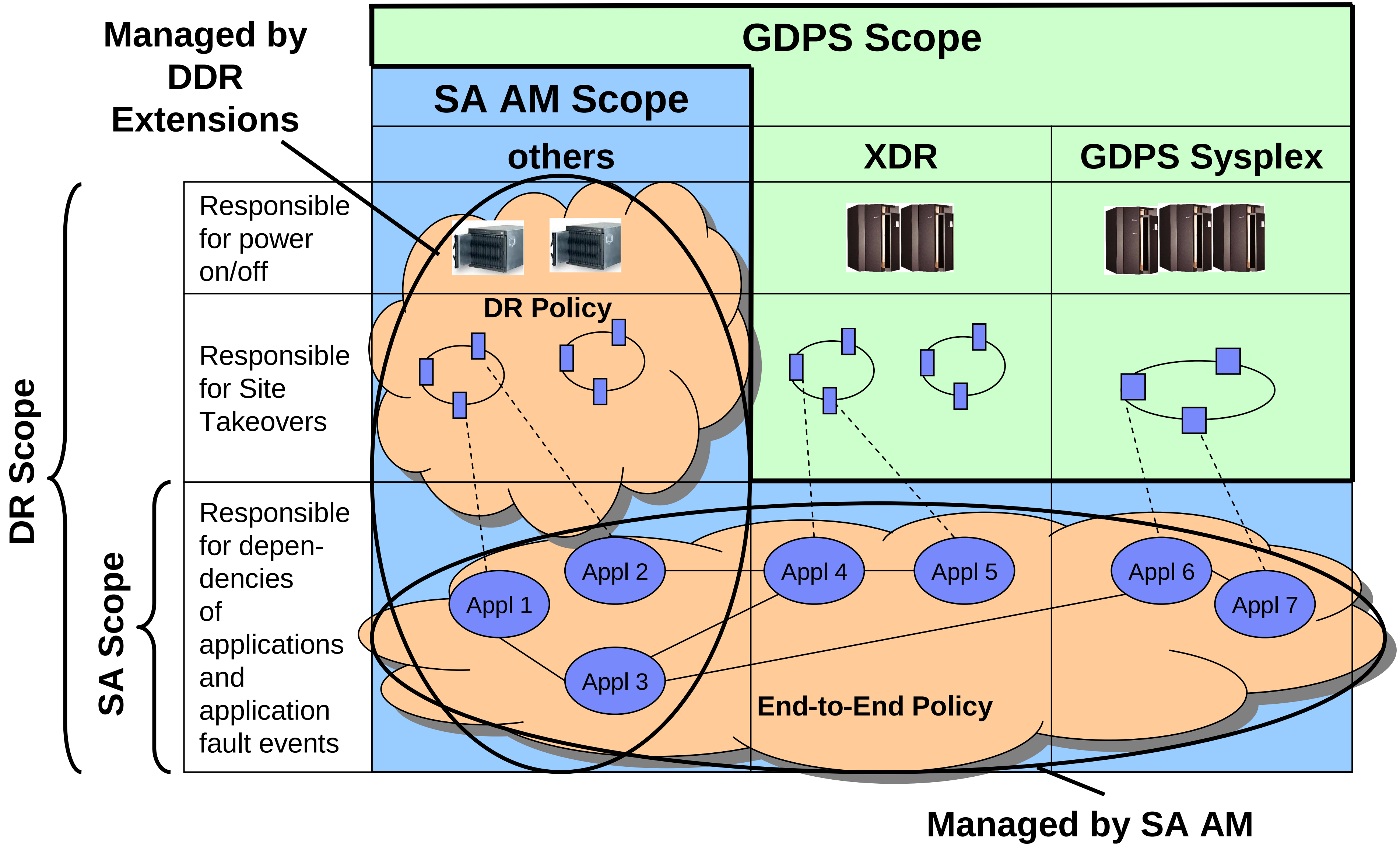


Supported Setups

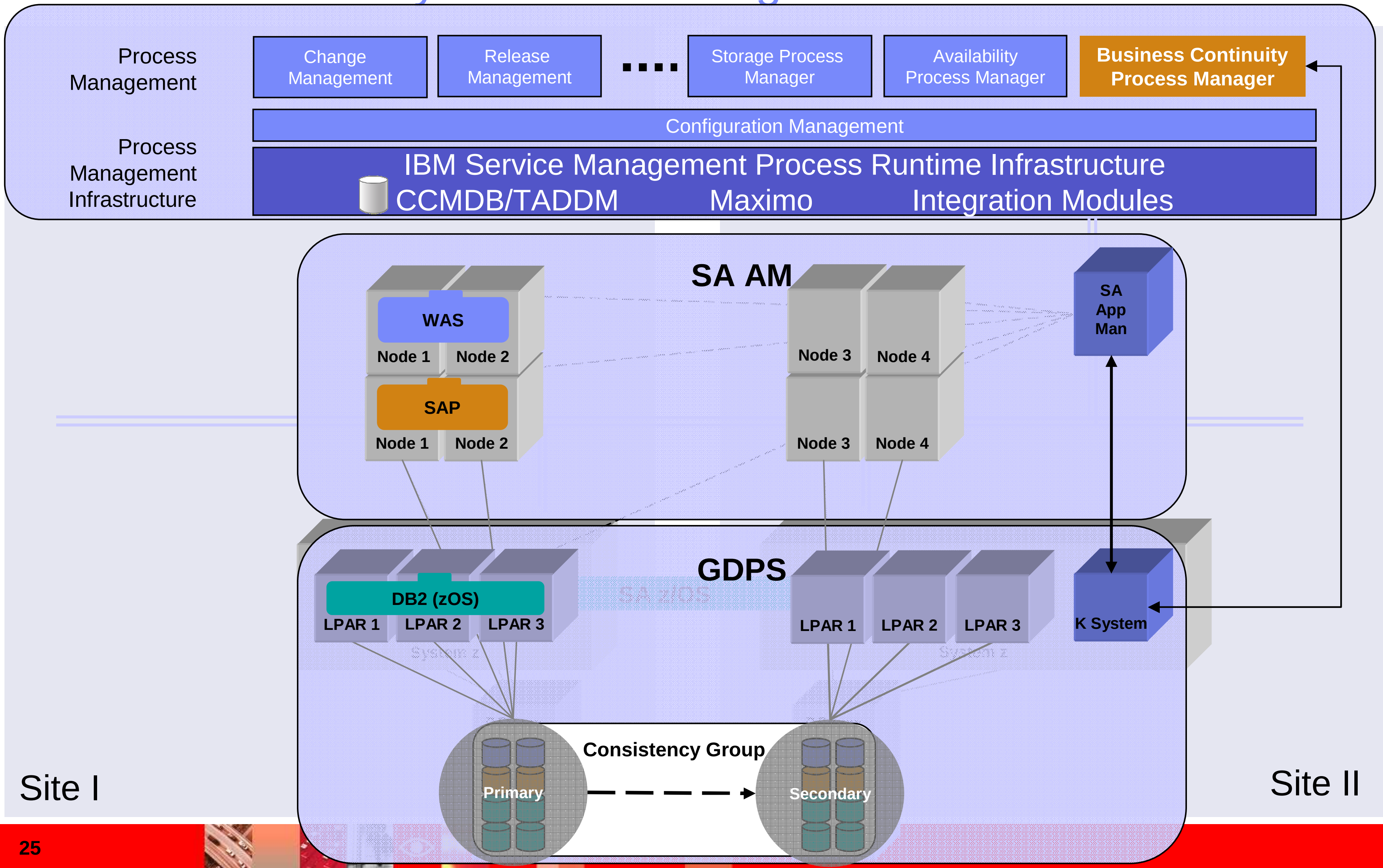
- Non-Stretched Cluster
  - Active-Passive
    - Resources can move to the local cluster on the other site
  - Disk Replication managed by GDPS ➔ PPRC



Responsibilities



Business Continuity Process Manager



Need More Information?

Please contact us:

Arwed Tschoeke
Arwed.tschoeke@de.ibm.com

Joachim Weber
jweber@de.ibm.com