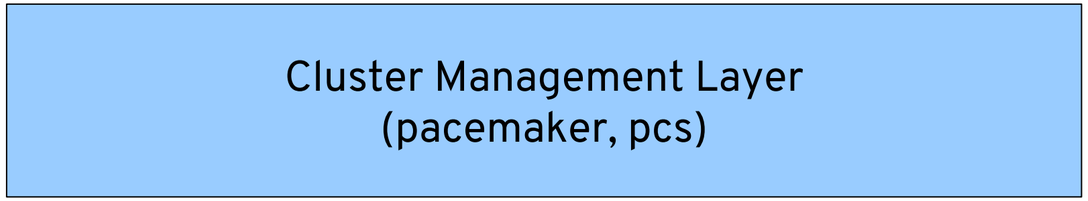
Red Hat HA add-on for SAP

Markus Koch
Partner Enablement Manager EMEA

Agenda

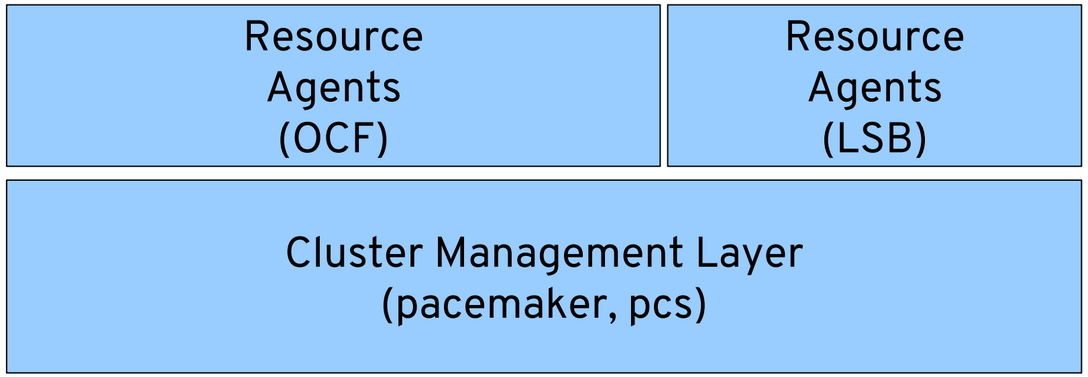
- Pacemaker Overview
- Enhancements
  - Resource Alerts
  - Integration
    - New resource agents
    - New fence agents
  - Multi-site (Tech Preview)
  - Stretch (Tech Preview)
  - Core / UI
- Supported SAP HANA configurations

Basic HA Cluster Architecture



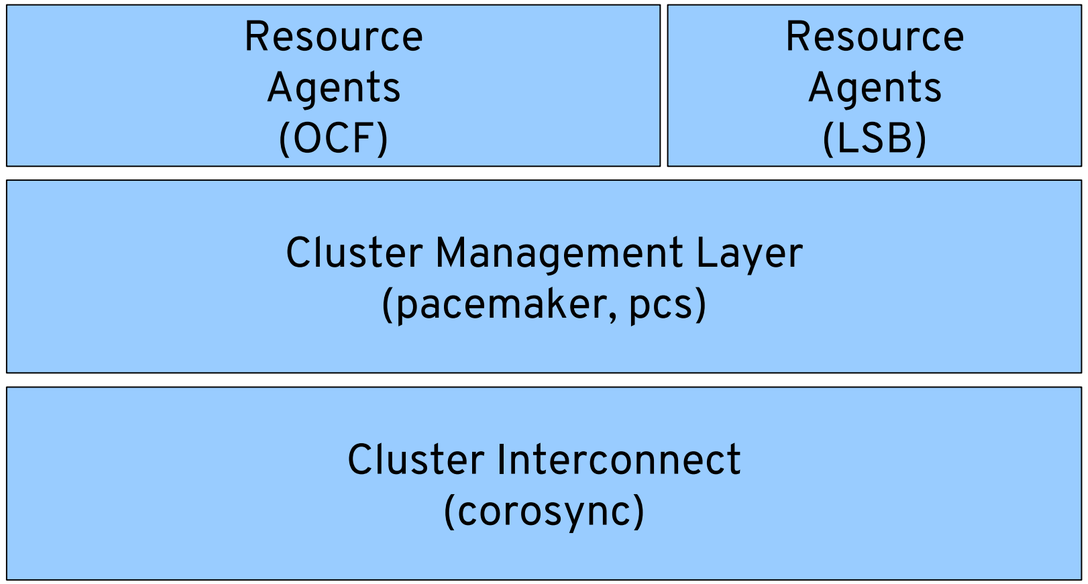
Pacemaker provides all packages to configure and manage a high availabilty cluster via CLI or GUI

Basic HA Cluster Architecture



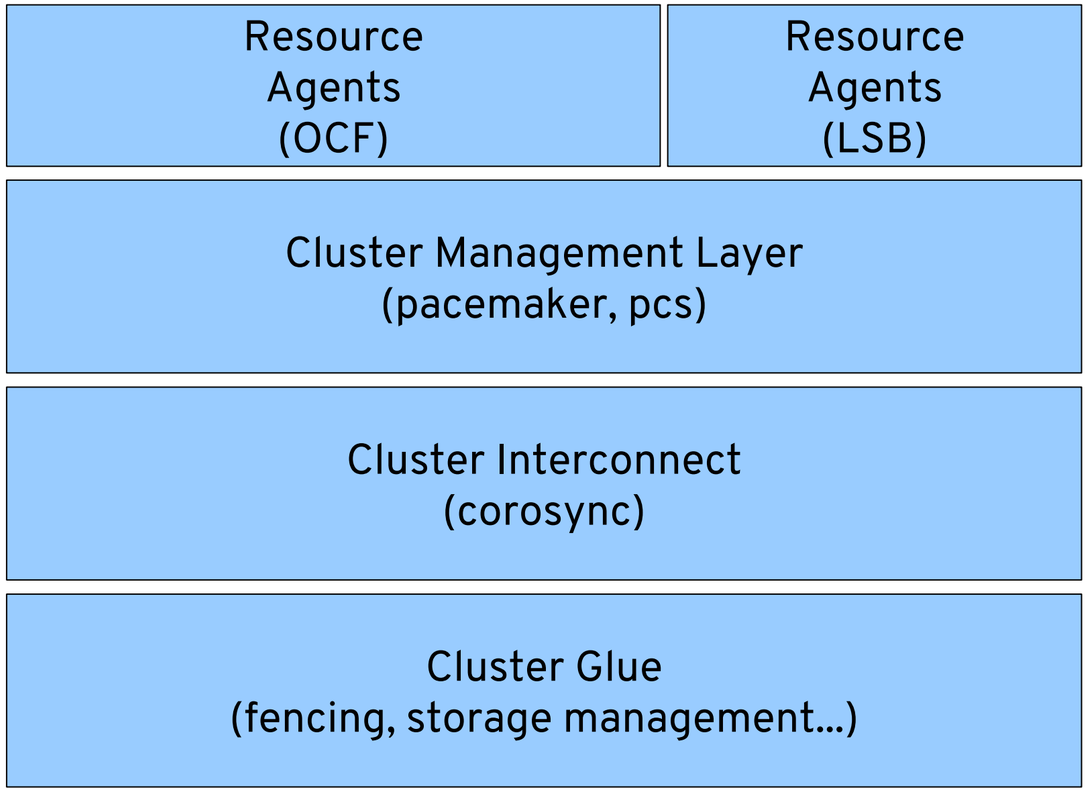
Resource Agents are packages that integrate applications.
They understand applications and its dependencies, so
That they can start, stop and monitor applications

Basic HA Cluster Architecture



Clusters need a heartbeat for internal communications, to check health of nodes, find quorum and other required management communication

Basic HA Cluster Architecture



Fence agents are used to shut-off failed or unresponsive cluster nodes to avoid data corruption

Support of Multi-Site Architectures

https://access.redhat.com/articles/27136

- Multi-Site Disaster Recovery Clusters
  - Independent clusters with identical configuration
  - Shared cluster storage is replicated
  - Manual failover in case of disaster
- Stretched Clusters
  - Need to survive failure of one-site or split-brain
  - Surviving site needs to get quorum
  - LAN-like latency (<=2ms RTT)
  - No GFS, clvm, cmirror support

Resource Alerts

- Allows notifications to be sent for any type of pacemaker events
- Sample alerts are provided (snmp, smtp, file)
- Sample files can be used as-is or customized as necessary for each customer environment
- **For Example**
  - When a node in a cluster fails we can configure pacemaker to immediately send out an email to an admin
    - `pcs alert create & pcs alert recipient`

QDevice (Stretch Clusters)

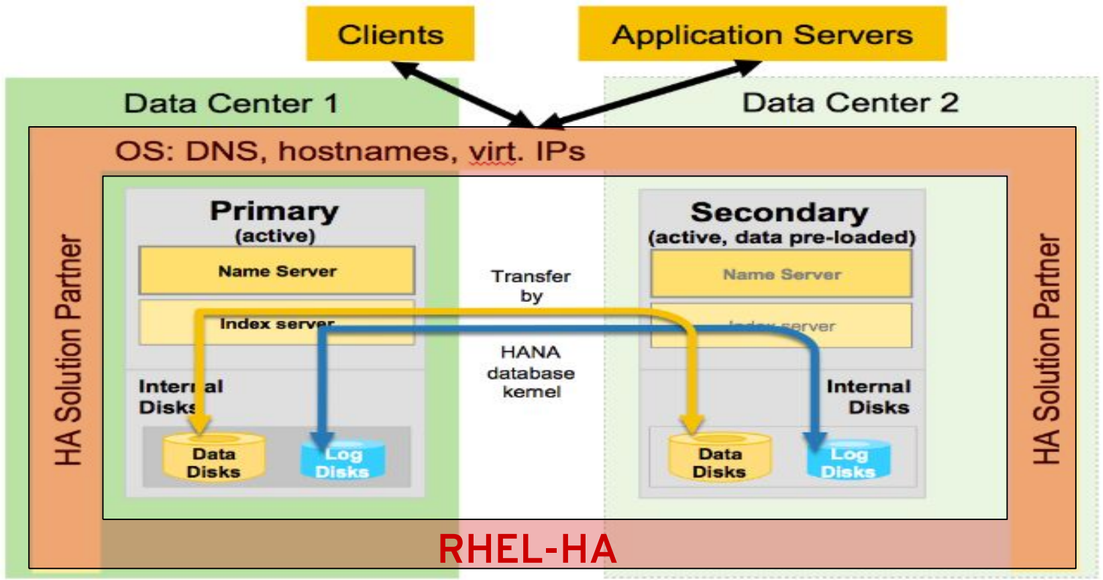
- Tech Preview since 7.3
- Allows cluster to be split in two separate sites
- Requires low-latency connection between sites (<2ms)
- Requires a third site to be the tie breaker
- Configuration through pcs (similar to other HA configuration)
- Can also be used in a 2-node cluster to act as a tiebreaker

Booth (Multi-site Clusters)

- Tech Preview since 7.3
- Allows two clusters in separate sites to coordinate resources
- Allows for higher latency connections
- Requires a third site (arbitrator)
- Configuration through pcs (similar to other cluster configuration)

SAP HANA System Replication

- SAP HANA replicates all data to a secondary SAP HANA system (standard SAP HANA feature).
- Data is constantly pre-loaded on the secondary system to minimize recovery time objective (RTO)
- RHEL-HA support all HANA releases from HANA1.0 SPS08, Scale-Up
- Scale-Out support in progress

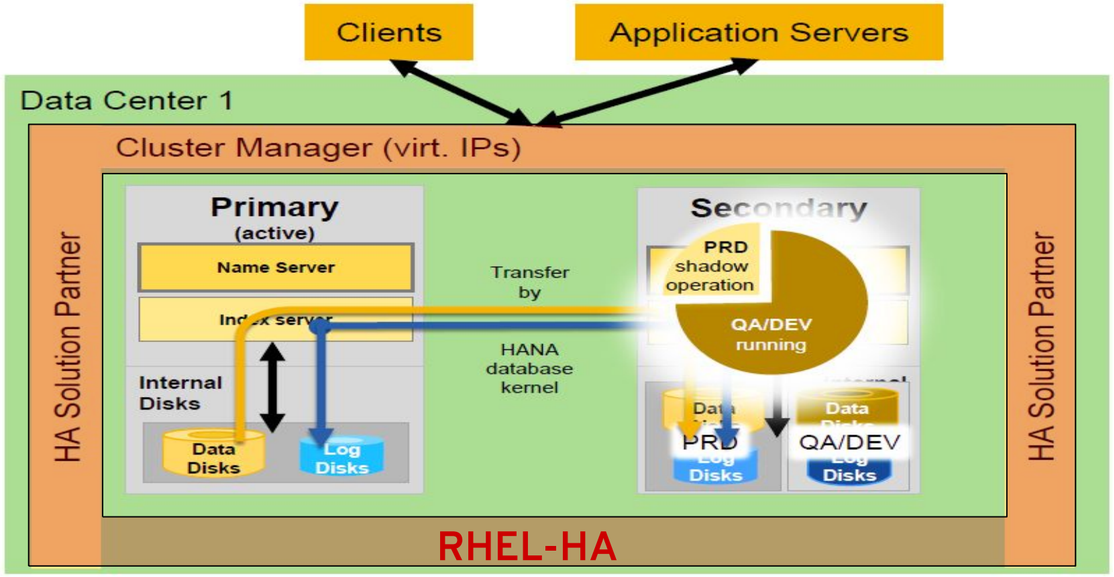


- Support of MCOD, MCD, MCOS
- Additional resource groups and constraints need to be configured
- Support of Active/Active (read enabled) in HANA 2.0

For more details: https://goo.gl/cqFPdb

SAP HANA System Replication Cost Optimized

- Alternative for local high availability
- Allows non-prod systems on secondary, resources are freed for non-prod instances (no/less data preload of production database)
- During take-over the non-prod operation needs to be ended
- Take-Over performance similar to cold start-up of SAP HANA

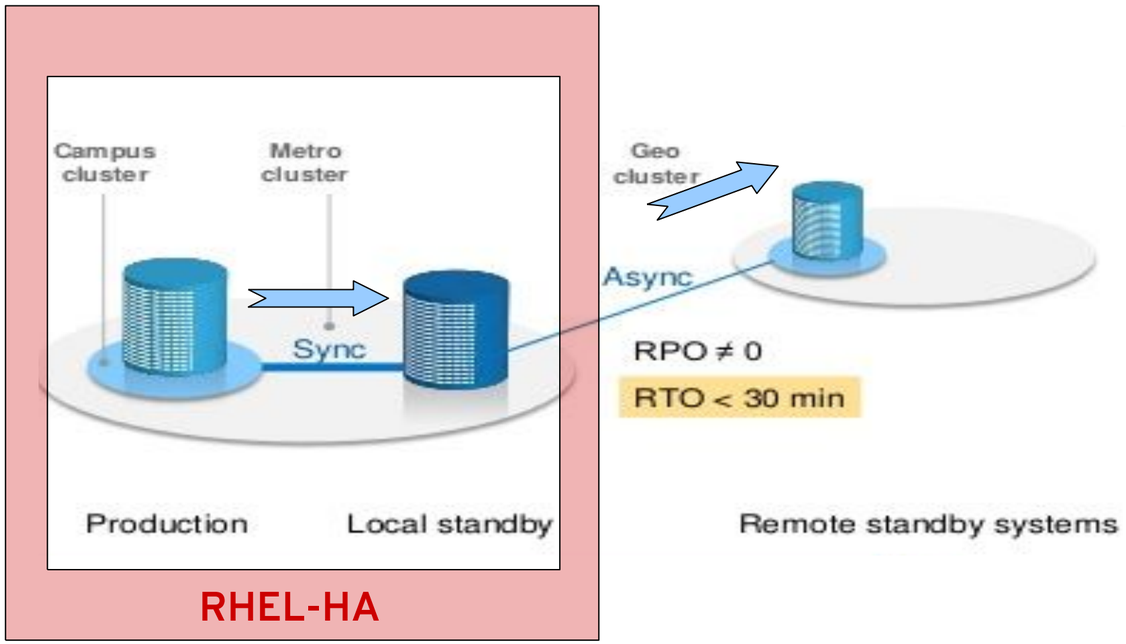


- Setup similar to normal setup
- Additional resource groups and constraints need to be configured

For more details: https://goo.gl/cqFPdb

SAP HANA Multi Tier System Replication

- Multi Tier System Replication / Replication Chains can make use of cluster
- Tertiary site not managed by cluster
- Replication to tertiary site will be broken in fail-over case
- newer HANA version will support Star Topology, where replication can continue



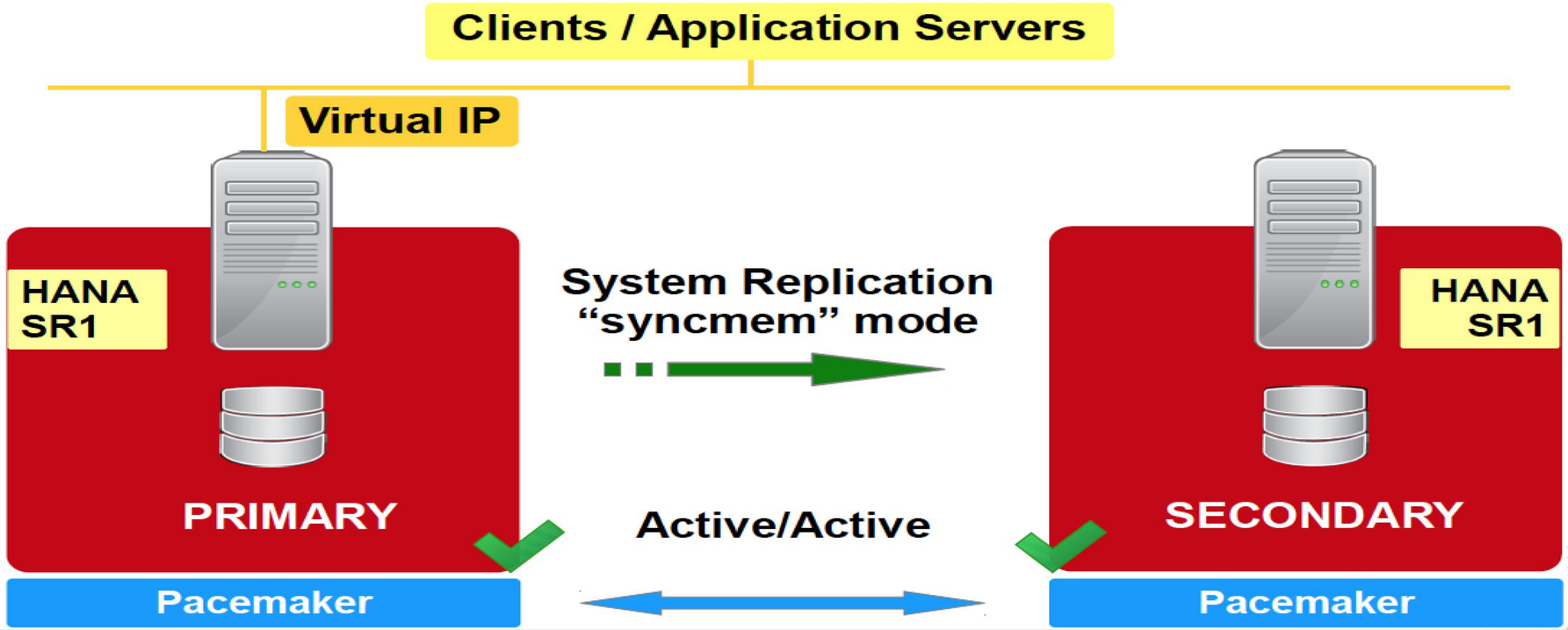
Automated SAP HANA System Replication

Resource Agents

- SAP HANA
  - Manages pre-configured SAP HANA System Replication environment
- SAP HANA Topology
  - Gathers information about the current status of SAP HANA System Replication
- Both are bundled in resource-agents-sap-hana rpm
- Configuration Guide
  - https://access.redhat.com/articles/3004101

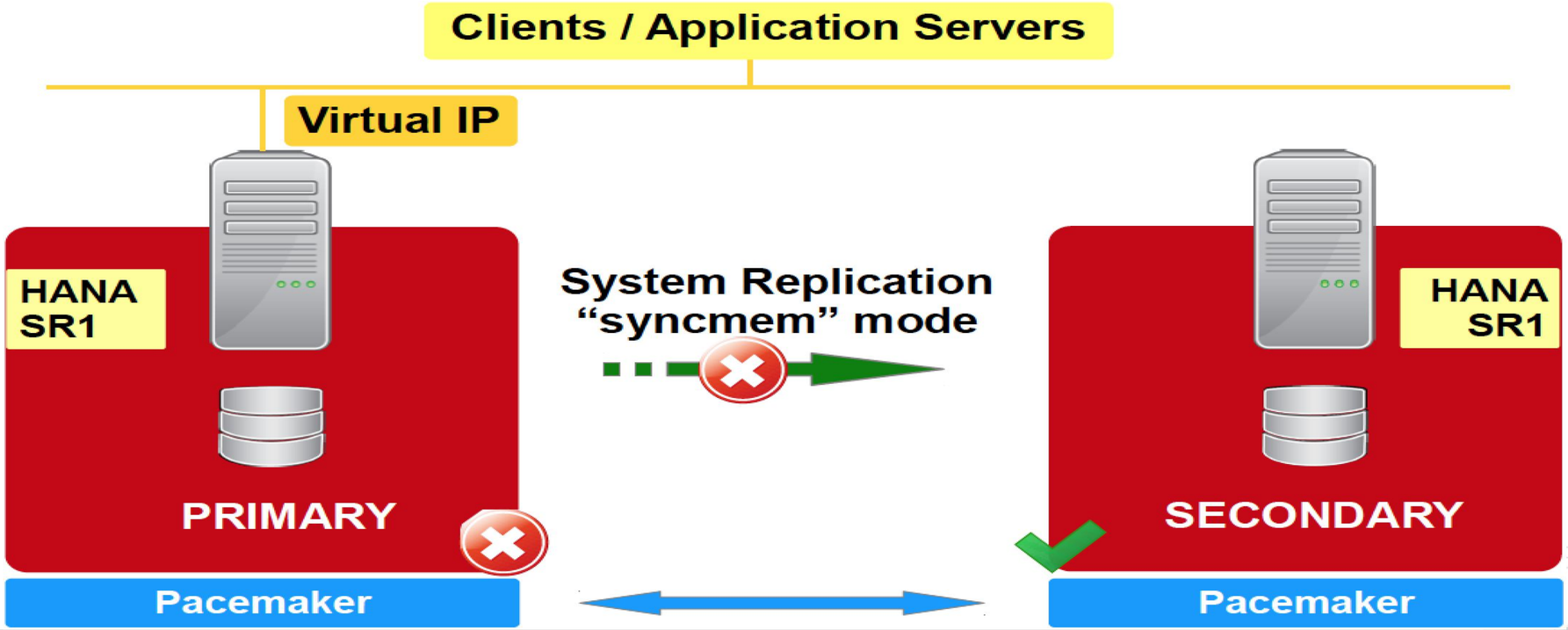
Failover Scenario – System Replication on Pacemaker

- System Replication modes: sync, [syncmem], async
- PREFER_SITE_TAKEOVER = True
- AUTOMATED_REGISTER = False
- No shared storage



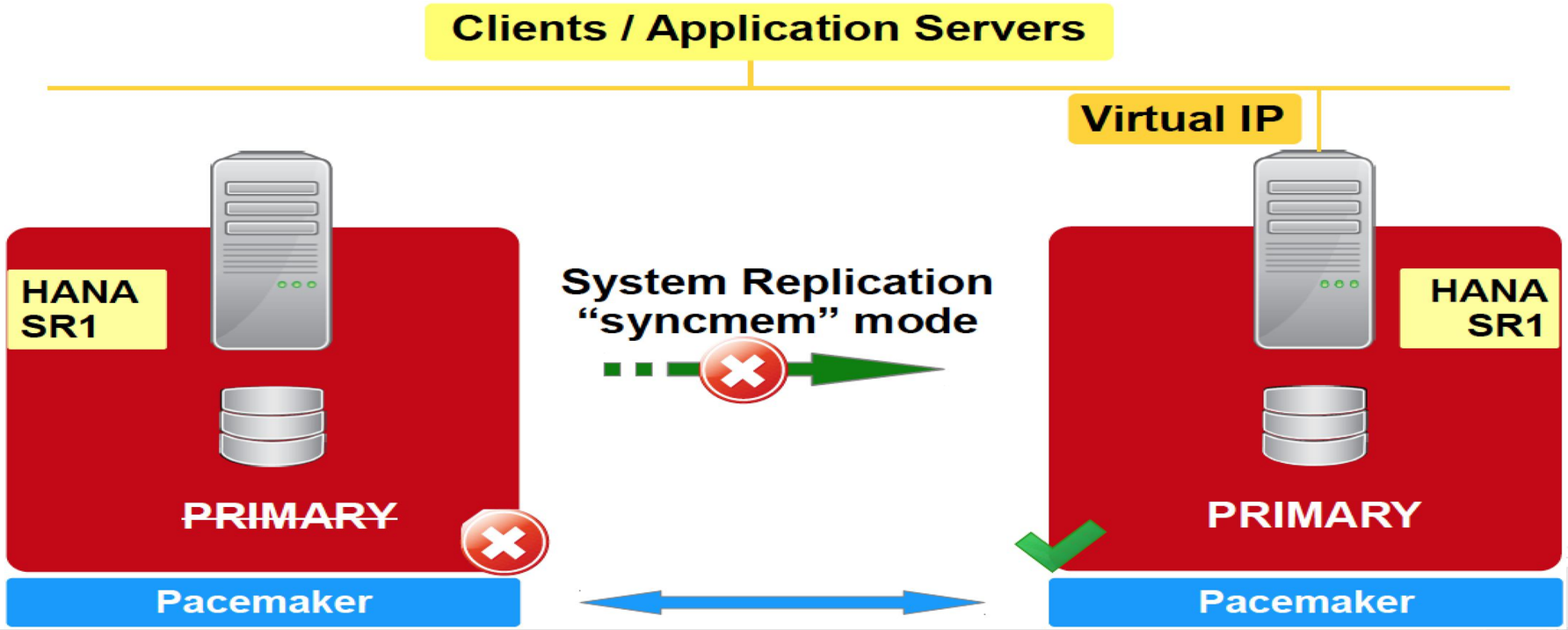
Failover Scenario – Primary Node Down

- Primary node down
- System Replication interrupted
- Pacemaker cluster fence the primary node



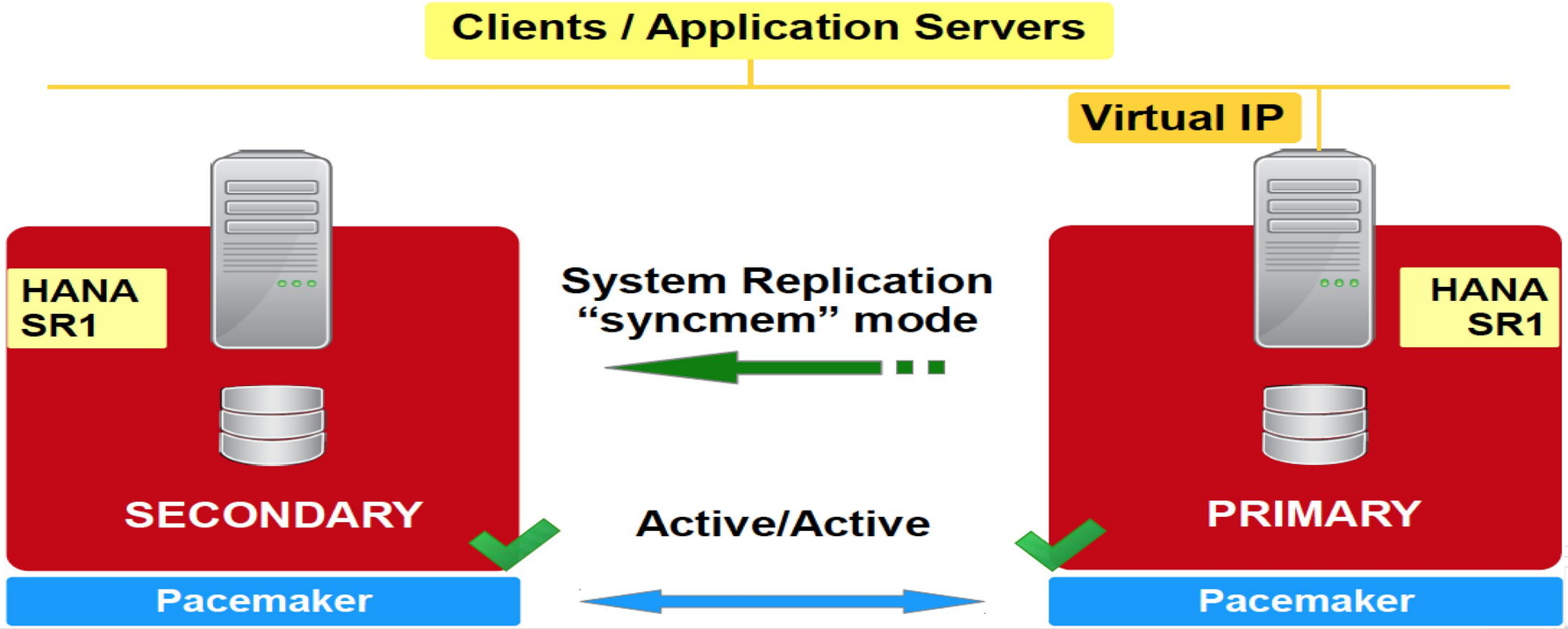
Failover Scenario – Secondary Node Take-Over

- Secondary becomes the new Primary
- Virtual IP binds to the new Primary node
- Previous Primary remains Primary, because “AUTOMATED_REGISTER = False”, and Administrator must decide if the setup failback or register the old Primary as the new secondary before HANA System Replication can start again



Failover Scenario – What if “AUTOMATED_REGISTER = True”

- Wait for “DUPLICATE_PRIMARY_TIMEOUT” timeout
- Former Primary registers as the new Secondary
- System Replication starts, in the opposite direction



The Rules of HA

- Keep it simple
- Keep it simple
- Prepare for failure
- Complexity is the enemy of reliability
- Test your HA setup

Additional Information

- Tutorial: https://goo.gl/Yd7A8n
- Documentation RHEL 7 Clustering: https://goo.gl/HymSD6
- Knowledge Base Index: https://access.redhat.com/articles/47987
  - Components, Concepts & Features: https://access.redhat.com/articles/2662011
  - Recommended Deployment Practises: https://access.redhat.com/articles/40051
  - SAP HANA system replication in pacemaker cluster https://access.redhat.com/articles/3004101
  - SAP Netweaver in pacemaker cluster https://access.redhat.com/articles/3150081
  - Cluster Review: https://access.redhat.com/articles/2359891
  - Introduction to Cluster Tools: https://access.redhat.com/articles/2123451
  - Introduction to Failover: https://access.redhat.com/articles/2123551

redhat.

THANK YOU

plus.google.com/+RedHat

facebook.com/redhatinc

linkedin.com/company/red-hat

twitter.com/RedHatNews

youtube.com/user/RedHatVideos