Multidomain MDM

MDM Environment scaling through HA and DR setup

Adhish Mahajan, Product Specialist- MDM

Informatica™

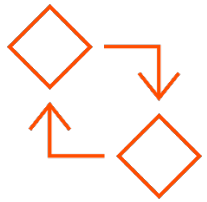
Housekeeping Tips

- Today's Webinar is scheduled for 1 hour
- The session will include a webcast and then your questions will be answered live at the end of the presentation
- All dial-in participants will be muted to enable the speakers to present without interruption
- Questions can be submitted to "All Panelists" via the Q&A option and we will respond at the end of the presentation
- The webinar is being recorded and will be available to view on our INFASupport YouTube channel and Success Portal. The link will be emailed as well.
- Please take time to complete the post-webinar survey and provide your feedback and suggestions for upcoming topics.

Feature Rich Success Portal

Bootstrap trial and POC Customers

Enriched Customer Onboarding experience



Product Learning Paths and Weekly Expert Sessions

Informatica Concierge with Chatbot integrations

Tailored training and content recommendations

Informatica

More Information

Success Portal

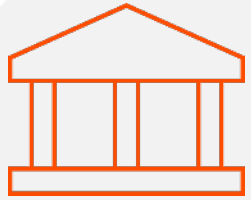
https://success.informatica.com

Communities & Support

https://network.informatica.com

Documentation

https://docs.informatica.com



University

https://www.informatica.com/in/services-and-training/informatica-university.html

Informatica

Safe Harbor

The information being provided today is for informational purposes only. The development, release, and timing of any Informatica product or functionality described today remain at the sole discretion of Informatica and should not be relied upon in making a purchasing decision.

Statements made today are based on currently available information, which is subject to change. Such statements should not be relied upon as a representation, warranty or commitment to deliver specific products or functionality in the future.

Agenda

- Overview
- MDM High Availability Solutions
- ActiveVOS and Process Server
- Infinispan clustering for High Availability (HA)
- Clustering – Elastic Search
- HA requirements for SIF, E360, REST
- JMS and EJB Clustering
- Introduction to Disaster recovery
- Best Practices

Overview – MDM High Availability(HA)

- HA provides maximum flexibility to access the MDM services without much interruption.
- It’s setup helps preventing unplanned downtime. It helps to recover from failures in application components and system wide failures.
- It helps in meeting business load requirements through environment scaling and multi node setup.

MDM High Availability Solutions

1. Non Distributed HA
2. Distributed HA
3. Zero Down time

Non-Distributed HA - Architecture

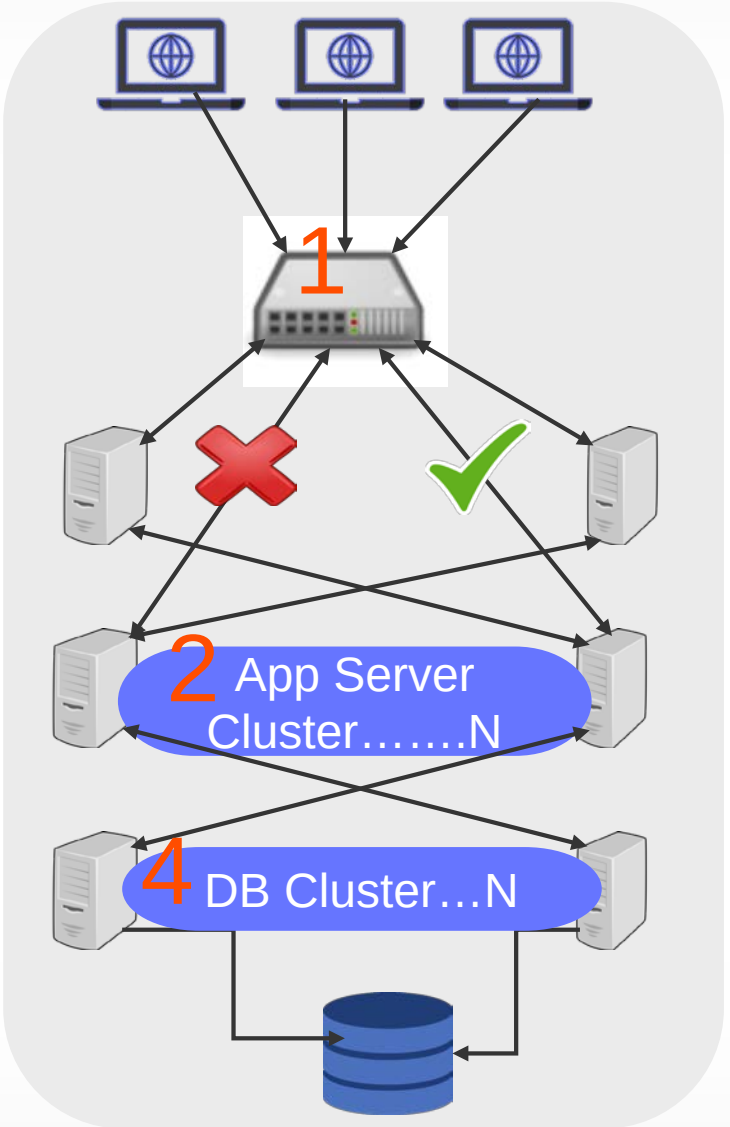
1. Load Balancer
2. Multi node setup
3. Support Continuous Ops when Node failure
4. DB cluster for HA

Pros

- Cost effective
- Easy to setup
- Commonly used Architecture

Cons:

- LAN/Network Failures possible
- DR not possible through this solution



Network LB and Failover

Web App

Custom Services

Hub & Process Server

Hub Store

Informatica

Distributed HA - Architecture

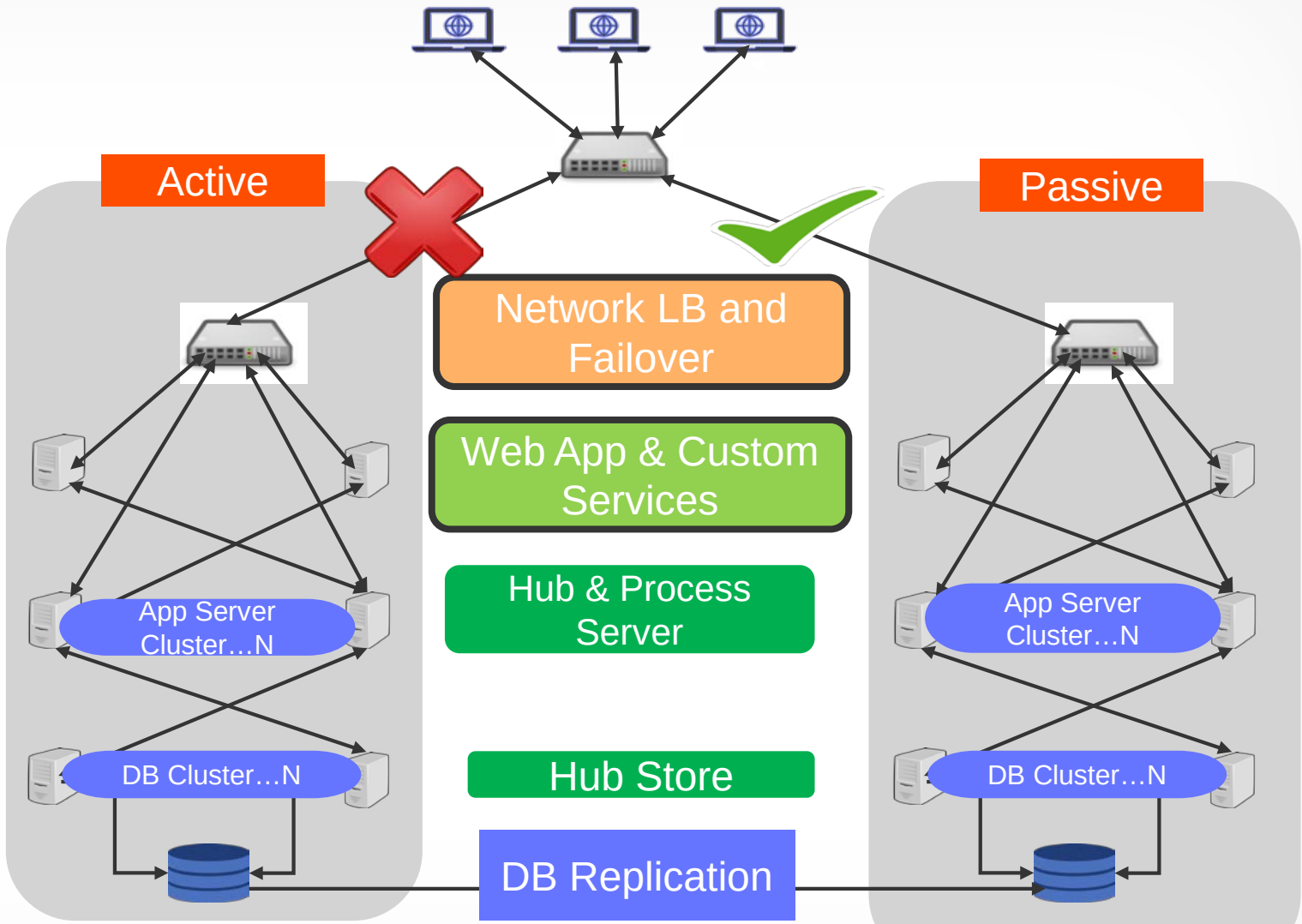
1. DB Replication required
2. Supports HA even during physical failures
3. Active-Passive

Pros

- Handles Network Failures
- Ideal setup for DR

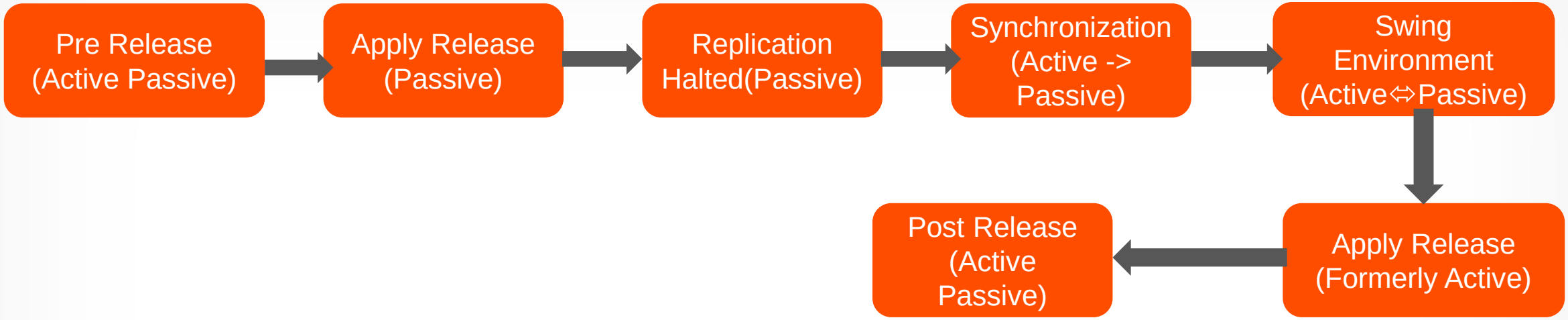
Cons:

- High Cost
- High Maintenance



Informatica

Zero Down time



- Supports continuous Ops during planned downtime
- Active-Passive(Distributed HA) Architecture
- Oracle Golden Gate – Required to manage data replication between environments

Pros :

- Useful for Mission critical apps

Cons :

- Very high cost
- High maintenance
- Not available for SQL Server

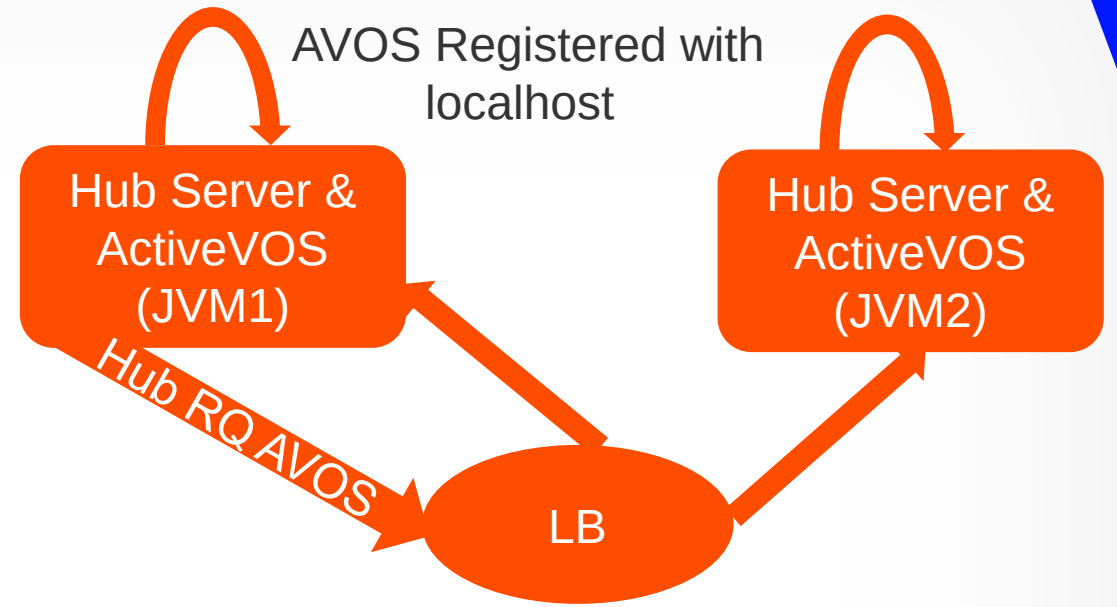
Informatica

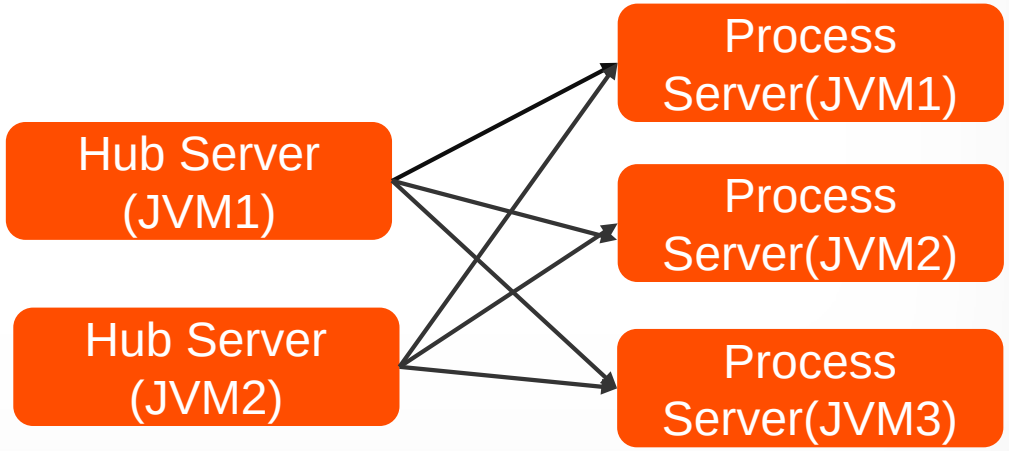
MDM Components

- ActiveVOS and Process Server
- Infinispan
- Elastic search
- SIF/REST
- E360
- JMS & EJB

ActiveVOS & Process Server

- ActiveVOS HA
  - Register workflow with localhost
  - Register workflow with LB(If localhost not allowed)
- Process Server HA
  - Requires registration of all Process Servers(In Hub Console)
  - Hub Server Internally Load Balance requests across Process Servers
  - No need of LB between Hub and Process Server





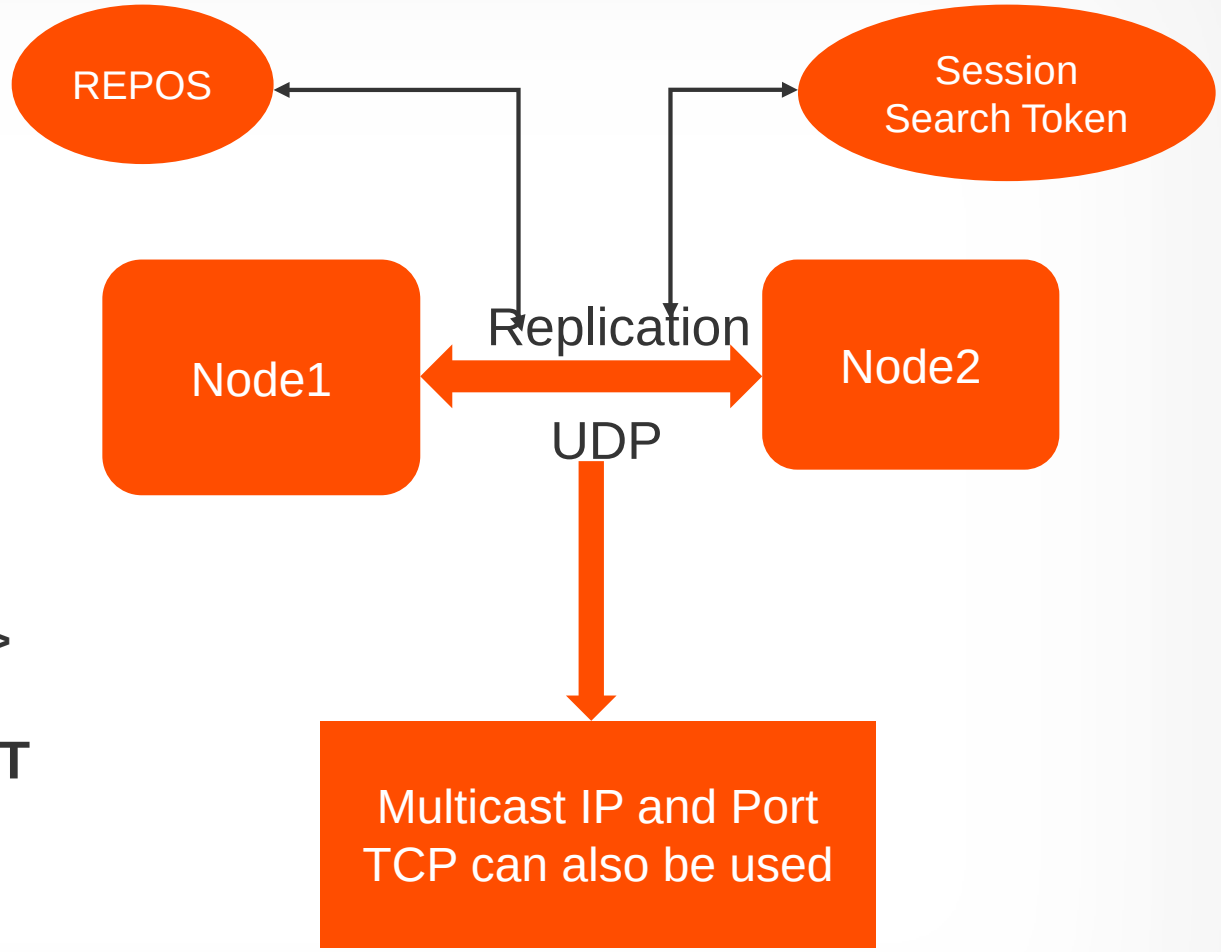
Informatica

Infinispan Clustering

- Replication
  - REPOS CHANGES
  - Runtime Information
    - SESSION MANAGEMENT
    - SEARCH TOKENS
- UDP – Multicast IP and Port.
- Clustername(To identify nodes of same cluster)

  <CacheName>_<DatabaseHost>_<MasterDatabaeName>
- Database Host(cmx_system) ≠ **LOCALHOST**

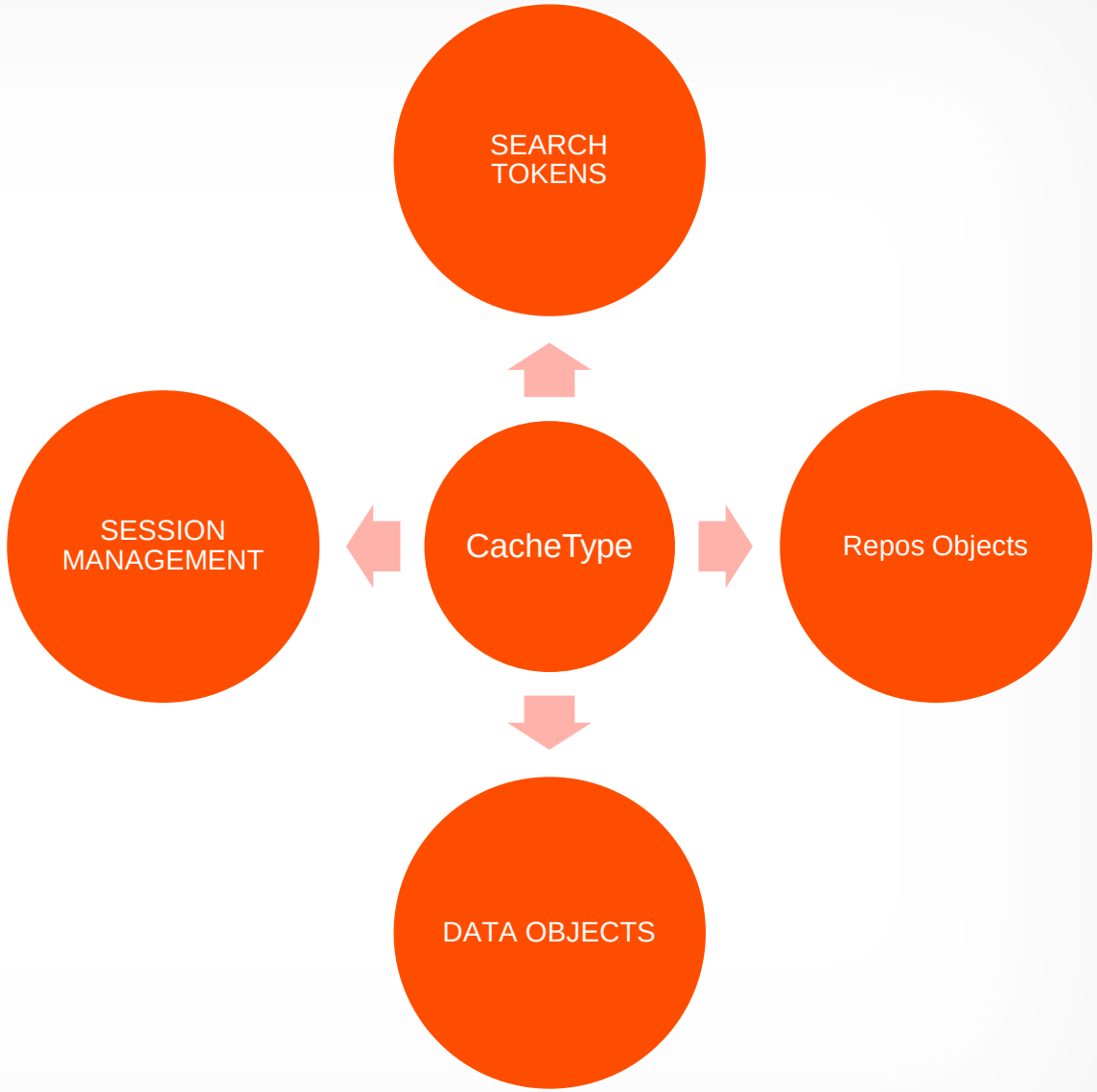
  Should be Database server hostname or IP



Informatica

Infinispan Clustering Contd.

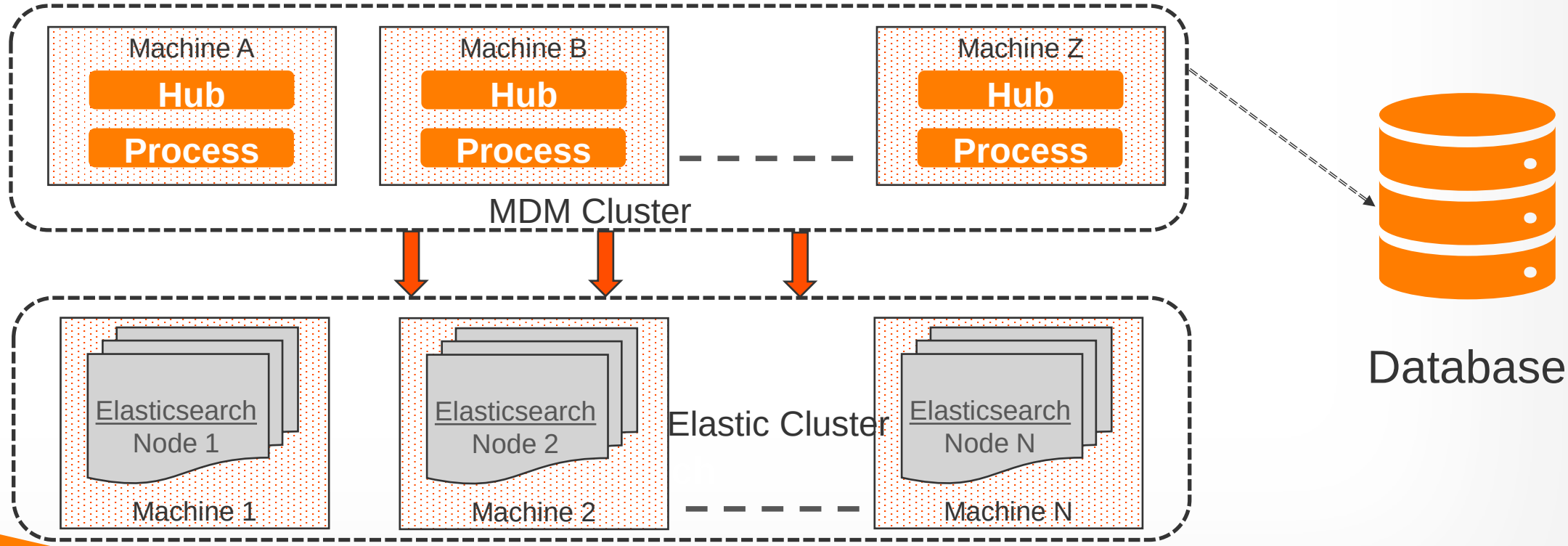
- Replicated cache SYNC mode
- Region Wise Cache(to store different info)
  - REPOS_OBJECTS region – Stores Repos Information like tables, columns etc.
  - DATA_OBJECTS region - Information like SamResources, Hierarchy Type etc.
  - SESSION MANAGEMENT – Stores user session info
  - SEARCH TOKENS – stores search tokens for pagination functionality(SIF/REST)



Informatica

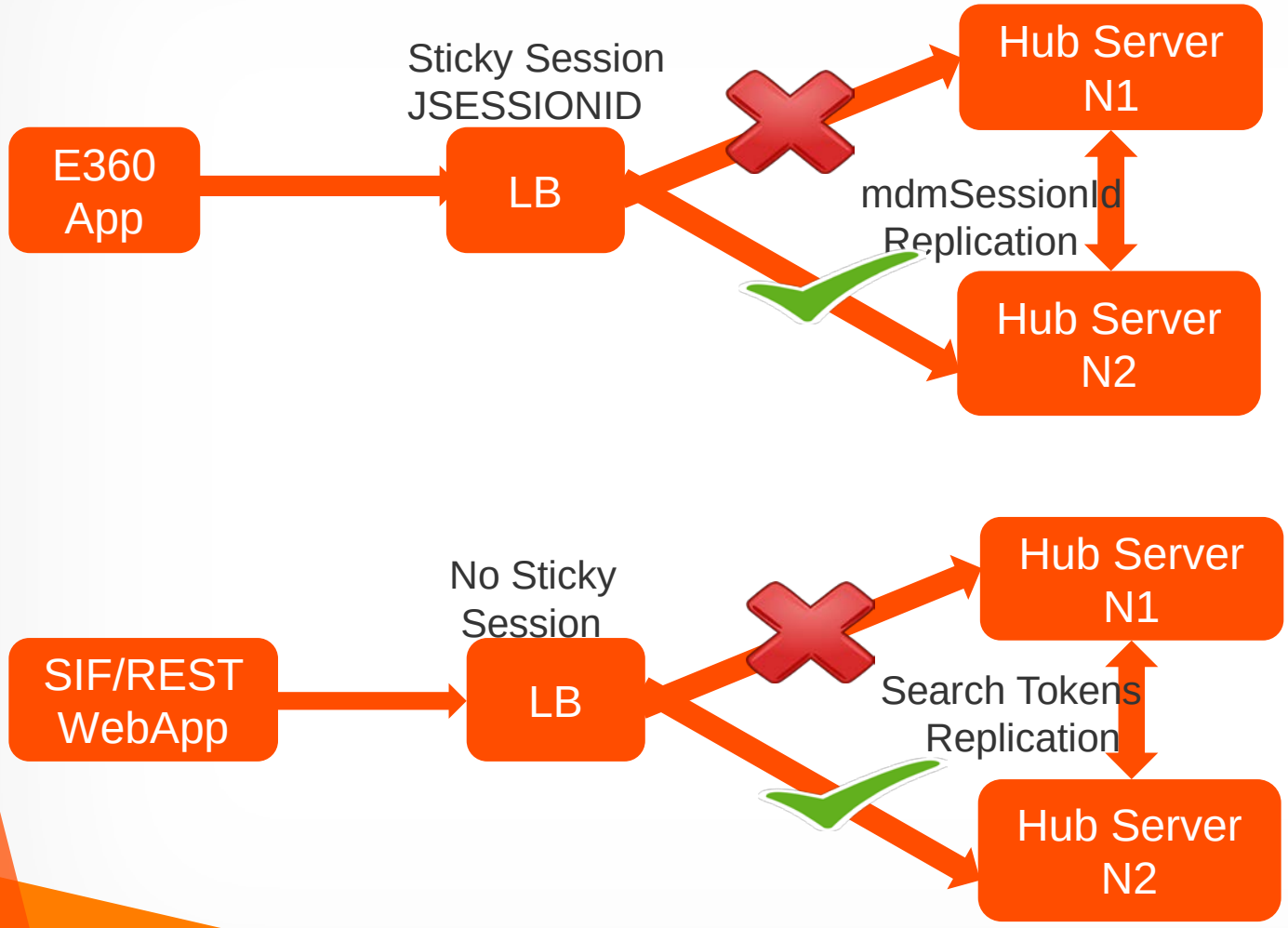
Clustered ES in multi-node MDM

- High Volume Systems with Frequent Loads, Heavy API usage and E360 Searches.
- N should always be an Odd number(To avoid Split Brain)
- Minimum active nodes should be (N/2) + 1 for ES HA
  - For 3 Node ES cluster, 2 should be active at all times
  - Requirement of ES



Informatica

E360, SIF, REST HA requirements



- **E360 Application**
  - Load Balance(LB) - Required for HA
  - One node goes down
    - LB useful for failover
    - Requires user relogin
    - Uncommitted changes needs to redone
  - Enabling Sticky session required
- **SIF, REST based custom application**
  - LB – Required for HA
  - No Sticky session(to avoid all requests going to same node)
  - Application should implement retry logic if node goes down after request sent

Informatica

EJB, JMS Clustering

- **EJB Clustering**
  - HA for EJB calls -> MDM
  - Comma Separated URL(Failover)
  - Configuration at App Server
  - MDM Config, jboss.cluster => true, postinstall/repackage.
- **JMS Clustering**
  - Publish message => Outbound queue
  - HA for Message Consuming App
  - Configuration at App Server

Disaster recovery(DR) - Introduction

- Follows Distributed HA Architecture.
- Passive instance - Geographically distant location.
- Setup to recover - Complete failure in Active instance.

Below configuration required during DR steps :

- CMX_SYSTEM.C_REPOS_DATABASE(Database specific information)
- <ORS_NAME>.C_REPOS_DB_RELEASE(Database specific information)
- Datasource configuration(In App server)
- Cmxserver properties(Hub server installation)
- build.properties(DB URL, Host and port of avos, MDM server)

Disaster recovery(DR) Contd.

- setSiperianEnv.sh(**Server and clustername**)
- Copy Certificates(from **<HubServer>/resources/certificates**)**(Active -> Passive)**
- Copy Certificates(from **<ProcessServer>/resources/certificates**)**(Active -> Passive)**
- Repackage EAR files and redploy.

Best Practices

- **Avoid Split Brain scenarios**
  - ES cluster : Avoid by usage of odd number ES nodes
  - JMS, Infinispan cluster : Avoid by usage of odd number MDM nodes
- **Infinispan** – Use different set of Multicast ports for different Environment to avoid network traffic.
  - For example QA nodes will have different multicast port than Prod nodes
- Recommend to perform the Metadata Changes during downtime and restart the nodes(In Prod).
- SIF / REST calls using paging ensure search tokens are replicated. If the requirement cannot be controlled to a single node.

Thank You

Informatica™