PowerCenter High Availability
Informatica PowerCenter High Availability Option provides high availability of all PowerCenter components, seamless failover and recovery of stopped or interrupted work, and simplified set up and management through a web-based administration console.

- PowerCenter HA relies on the underlying IT infrastructure to achieve end-to-end HA. i.e. highly available database, file system, network and hardware servers.

- High availability is not a solution for disaster recovery. You can use high availability features to implement a disaster recovery solution.
PowerCenter High Availability Benefits

- **Resilience.** Highly available systems can tolerate temporary connection failures until a timeout period expires or the failure is resolved. The system tries to reconnect for a specified period of time. If the failure is resolved, there is no interruption in end user activity.

- **Restart and failover.** In highly available systems, when a machine becomes unavailable, processes running on the machine can be restarted on the same machine or on a backup machine. By allowing processes to restart on the same machine or fail over to another machine, the system minimizes or eliminates the downtime due to the failure and maximizes the system operational time.

- **Recovery.** In highly available systems, an interrupted service can complete its operations after it is restarted. A service may be stateful—that is, it records its state of operation in a shared location periodically. When a failure occurs, the system must retrieve the state of the affected service so that it can automatically restart or recover jobs that have terminated abnormally.

Reduce costs and risks associated with data downtime
PowerCenter HA Capabilities

• Service failover from primary to backup services and nodes.
  • Automatically ensures service availability on primary or backup servers should the primary server fail.

• Session and workflow recovery from checkpoints.
  • Automatic or configurable recovery for sessions affected by service failures.

• Resiliency to network and external failures.
  • Automatic re-connect within resilience timeout constraints handles transient network errors and connection failures.

• Centralized web based configuration and administration
  • Ability to create, manage and monitor a high availability configuration of all PowerCenter services through a web-based administration console.
  • Enable Administrators to visually identify single point of failures within the Informatica environment.
PowerCenter High Availability

Automatic Failover Simulation

Integration Service also supports Active-Active mode.
Achieving PowerCenter HA

- **Core Services Availability**
  - At least 2 nodes configured w/ core services for fail-over

- **Application Services Availability**
  - At least 2 nodes configured as primary and backup for services

- **Informatica Services (Tomcat – Service Manager)**
  - Configure service to restart automatically if it terminates unexpectedly

- **External Systems Availability**
  - For Repository, Source/Target/Lookup database to be highly available, use highly available versions of databases (e.g. Oracle RAC, IBM DB2)
  - Use highly available FTP servers and Message Queues.
  - Configure network to be highly available
  - Need shared directory for config, log files, storage (stores state for session and workflow recovery).
  - Shared directory should be on HA file system (e.g. Veritas Cluster File System, IBM GPFS) to remove point of failure
Achieving PowerCenter HA

Key underlying components to achieve a PowerCenter High Availability solution are:

– Highly Available Database
– Highly Available Shared File System
– Redundant Network

General Recommendations:
• Network 1 GB >
• HA CFS w/ Heartbeat and Failover
• Redundant Network

Actual config depend on environment and SLA reqs.
DR with PowerCenter

Incorporating PowerCenter into Disaster Recovery Solutions:

• Primary Data Center should be configured with PowerCenter HA including underlying HA infrastructure.

• Backup Informatica Nodes & Services configured passive (cold standby) mode.

• External Systems are actively replicated across Data Centers by respective vendors.
**DR with PowerCenter**

Incorporating PowerCenter into Disaster Recovery Solutions:

- Primary Data Center should be configured with PowerCenter HA including underlying HA infrastructure.
- Backup Informatica Nodes & Services configured passive (cold standby) mode.
- External Systems are actively replicated across Data Centers by respective vendors.
- Backup Informatica Nodes & Services become active only when Primary Data Center goes down and replication of required data has been completed.
  - Requires scripting / integration with 3rd party operation management tools.
High Availability Scenarios
Detailed Walkthrough
Common Failure Scenarios

• Typical single node setup where most jobs complete within a specific window of time.

• Unexpected failures require sessions to be restarted losing precious time.
Common Failure Scenarios

- Typical single node setup where most jobs complete within a specific window of time.
- Unexpected failures require sessions to be restarted losing precious time.
- Transient network failures result in the PowerCenter losing connectivity to sources and targets.
How Does It Work? Network Resiliency

- With HA, session does not immediately fail once connectivity to target is lost.

**Domain**

**Node 1**

- Domain
- Gateway
- Log
- Repository
- Integration#1
- Integration#2

**Node 2**

- Domain
- Gateway
- Log
- Repository
- Integration#1
- Integration#2

**HA File System**

**Shared Directory**

**HA Database**

**Source DB**

**Target DB**

**Repository DB**
How Does It Work? Network Resiliency

• With HA, session does not immediately fail once connectivity to target is lost.

• Session will try to reconnect to the target for a specific amount of time.

With HA, session does not immediately fail once connectivity to target is lost.

Session will try to reconnect to the target for a specific amount of time.
How Does It Work? Network Resiliency

- With HA, session does not immediately fail once connectivity to target is lost.
- Session will try to reconnect to the target for a specific amount of time.
- Once transient network failure is resolved, session resumes processing.

- With HA, session does not immediately fail once connectivity to target is lost.
- Session will try to reconnect to the target for a specific amount of time.
- Once transient network failure is resolved, session resumes processing.
How Does It Work? Failover & Recovery

- Node 1 is running on a machine that encounters an unexpected failure.
How Does It Work? Failover & Recovery

• Node 1 is running on a machine that encounters an unexpected failure.

• Integration Service fails-over to Node 2
How Does It Work? Failover & Recovery

• Node 1 is running on a machine that encounters an unexpected failure.

• Integration Service fails-over to Node 2.

• Workflow and session restart with recovery and continue from last checkpoint.

How Does It Work? Failover & Recovery

- **Domain**
  - **HA File System Shared Directory**
  - **Source DB**
  - **Target DB**
  - **HA Database**

- **Node 2**
  - **Repository**
  - **Integration#1**
  - **Integration#2**

- **Node 1** is running on a machine that encounters an unexpected failure.
Recovery

Architectural recovery
• If the Service Manager and Repository Service recover, but the Integration Service cannot recover the restart is not successful and has little value to a production environment

Procedural recovery
• Recovery strategy set to the workflow/session level which can recovered manually or automatically.
## Comparing Features of Each Option

The following table compares the feature difference between the PowerCenter base product, and combinations of the base product with the high availability, server grid, and session on grid options:

<table>
<thead>
<tr>
<th>Product Options / Feature</th>
<th>Base Product</th>
<th>Base Product + Server Grid + Session on Grid</th>
<th>Base Product + High Availability</th>
<th>Base Product + High Availability + Server Grid + Session on Grid</th>
</tr>
</thead>
<tbody>
<tr>
<td>Internal PowerCenter Resilience</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>Repository Database Resilience</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>Automatic Failover of the Domain to a Gateway Node</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>Automatic Restart of Application Services on the same node</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>Automatic restart of the Integration Service and Repository Service on a Backup Node</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>Run Workflows or Session on a Grid</td>
<td>✔️</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Powercenter File System certification

The following shared file systems are certified by Informatica for use in Integration Service failover and session recovery

**Storage Array Network:**

Veritas Cluster Files System (VxFS)

IBM General Parallel File System (GPFS)

**Network Attached Storage using NFS v3 protocol:**

EMC UxFS hosted on an EMV Celerra NAS appliance

NetApp WAFL hosted on a NetApp NAS appliance

For more information, see the *Statement of Support Regarding File System Support for Informatica PowerCenter High Availability Service Failover and Session Recovery* on my.informatica.com.
Thanks