SAP Data Services Failover and Load Balancing

Failover Vs. Load Balancing

- A failover cluster (high availability) is a group of servers that work together to maintain high availability of applications and services. If one of the servers, or nodes, fails, another node in the cluster can take over its workload without any downtime (this process is known as failover).
- Load balancing (scaling) improves the distribution of workloads across multiple computing resources and helps to avoid overload of any single resource.

Failover (High Availability)

You can provide failover for Data Services batch jobs only (no real-time jobs) through using a Windows/AIX cluster. Only the DS Job Server is supported on a Windows Cluster.

You need at minimum 4 machines to achieve high availability. (2 for IPS and 2 for DS)
You cannot install IPS or BI on a Windows cluster as that is not supported.
You also need to have a Windows cluster set up before installing DS.
On the Web server side (Web Tier), we do not specifically provide any clustering support, but it might be achievable through 3rd party software. This setup however will not provide for load-balancing. Server group is configured on DS Management Console which is not supported with Windows Cluster. So you cannot have both high availability and load-balancing at the same time.

Almost Failover (High Availability) and Load Balancing

You need to have 2 machines with both Information Platform Services (or Business Intelligence Platform) and Data Services installed and then configure them in the following way:

1. Point both SIA's to the same CMS database.
2. Give your CMS cluster a name (in the CCM go to Properties of your SIA Node).
3. Set the cluster name rather than individual server name in the DS MC in the BOE Server section (Under Management).
4. You need to point both your 2 Input FRSs and 2 Output FRSs to point to the same highly available file system share.
5. You need to set your SIA's to run under an account that has access to the remote highly available share.
6. Have both DS job servers pointing to same DS repository.
7. Use a Job Server Group to which both of your respective DS job servers belong.
8. Ensure that on each SIA node you have both AJS and EIM APS installed
9. Use the BOE Scheduler and not Data Services scheduler if you want your job schedules to work if one of the servers is down.

Note: For batch jobs that are using the RFC Client (loading of BW and extraction from BW via Open Hub) you have to be on Data Services 4.2 SP1 for this to work since the RFC Service that's running on EIM APS will not work correctly if you have two EIM APSs exposing this service and belonging to a clustered CMS environment.

Note: Clustering of BI with IPS is not supported. You can cluster one BI with another BI, and one IPS with another IPS, however not mix them.

Load Balancing - Scaling

Add additional identical machines to your landscape and install Data Services Job Server.
Point each of your DS Job Servers to the same DS repository which has to be set as the default repository
Create a Job Server Group
For every additional instance of Data Services Job Server you need an additional license

Load Balancing Through Job Server Group
Load balancing is achieved through the logical concept Job Server Group.

A server group automatically measures resource availability on each Job Server in the group and distributes scheduled batch jobs to the Job Server with the lightest load at runtime.

Requirements:

All the Job Servers in an individual server group must be associated with the same repository, which must be defined as a default repository. The Job Servers in the server group must also have:

- Identical SAP Data Services versions
- Identical database server versions
- Identical locale

Each computer can only contribute one Job Server to a server group.

**Job Server Overhead**

Compared to normal Job Servers, Job Servers in a server group each:

- Collect a list of other Job Servers in their server group
- Collect system load statistics every 60 seconds:
  - Number of CPUs (on startup only)
  - Average CPU load
  - Available virtual memory
- Service requests for system load statistics
- Accept server group execution requests

**Load Balance Index**

All Job Servers in a server group collect system load statistics and convert them into a load balance index value for each Job Server. The Job Server with the lowest index value is selected to execute the current job. The software polls all Job Server computers every 60 seconds to refresh the load balance index.

**Installing Second Job Server**
Configuring Second Job Server
DS Repository

Creating Job Server Group
### Server Groups

#### Add a new server group

<table>
<thead>
<tr>
<th>Server Group Name</th>
<th>Repository</th>
</tr>
</thead>
<tbody>
<tr>
<td>sg_DS_LOCAL</td>
<td>DS_LOCAL</td>
</tr>
</tbody>
</table>

Select the job servers you want this server group to contain and click Apply.

<table>
<thead>
<tr>
<th>Select</th>
<th>Name</th>
<th>Job Server</th>
</tr>
</thead>
<tbody>
<tr>
<td>✓</td>
<td>JobServer_1</td>
<td>EBSYS5.0500</td>
</tr>
<tr>
<td>✓</td>
<td>JobServer_2</td>
<td>EBSYS2.3500</td>
</tr>
</tbody>
</table>

#### Status of a server group and the job servers it contains.

<table>
<thead>
<tr>
<th>Server Group Name</th>
<th>Job Server</th>
</tr>
</thead>
<tbody>
<tr>
<td>sg_DS_LOCAL</td>
<td>JobServer_2</td>
</tr>
<tr>
<td></td>
<td>JobServer_1</td>
</tr>
</tbody>
</table>
Ensure Jobs Can Run On Any Job Server

Use the same directory structure across your host computers for source and target file operations
Use relative paths for file names
Use the same source and target connection strings to your databases for all Job Server hosts
Bulk load directories

Additional Info

1603393 - Installation Scenarios - Data Services 4.x
Check KBA: 1799669 - How to cluster Information Platform Services 4.x with Business Intelligence Platform 4.x
Please reference the Data Services Performance Guide to understand the different methods of distributed execution for DS batch jobs.
Jobs will still have to be re-run and recovery built into them. Please see Designer Guide.
The above configuration does not cover load balancing or scaling of Real-Time jobs. See KBA 1592957 for more info.