

Checklists for Upgrading to
Pentaho 8.1

Pentaho Upgrade Checklists

These checklists are designed to give you a reference point for each step of the Upgrade. Included are lists for general steps, details about your server environment, data integration specifics, and Pentaho tools and content.

| Software | Major Version(s) |
|----------|------------------------------|
| Pentaho | 6.0, 6.1, 7.0, 7.1, 8.0, 8.1 |

- [General Upgrade Checklist](#)
- [General Network Environment](#)
- [Pentaho Server \(or BA Server\)](#)
- [Data Integration Environment](#)
- [Business Analytics Content](#)

Table 1: General Update Checklist

| Question | Notes |
|---|-------|
| What is the currently installed Pentaho software and version? | |
| Did you determine your incremental upgrade path (if any)? | |
| What is the installation method used? (Archive or Manual) | |
| Did you back up your data and configurations? | |
| Did you install the new Pentaho version? | |
| Did you restore your backups and custom configurations? | |
| Did you update the Pentaho Design Tools? | |
| Did you start and test your Pentaho server? | |
| Did you contact your CSM and support, before upgrading, to let them know that you have custom configurations? | |

Table 2: General Network Environment

| Question | Notes |
|---|-------|
| What is the minimum and maximum memory allocated to the Pentaho JVMs? | |
| What are the configured environments (Dev, Test, Prod, other)? | |
| What is the server's operating system and version? | |
| What are the server's processor specifications and cores? | |
| What is the available memory to the operating system? | |
| What is/are the physical location and proximity of the server(s) (same country, data center, rack, other)? | |
| What is the username and password of the OS user who runs the Pentaho process? | |
| What plugins have been installed on or removed from the server? | |
| Is a proxy or load balancer used in front of the Pentaho server(s)? | |
| Have other configuration files been modified from their defaults of a fresh installation? If so, are these files backed up? | |
| Have any language localizations been made? | |
| Is Pentaho integrated with other web applications? If so, what URLs are used for the integration? | |
| Has the Pentaho UI been customized or themed? | |
| Is Pentaho clustering being used? | |
| Are DI-only servers set up within the cluster? | |

Table 3: Pentaho Server (or BA Server)

| Question | Notes |
|---|-------|
| What scripts are used to start/run Pentaho (ctlscript.sh, start-pentaho.sh, ctl.sh...)? | |
| How are users authenticated (Pentaho, LDAP, JDBC...)? | |
| How are users authorized for content (Pentaho, JDBC...)? | |
| Are there security extensions in place (SSO, DSP, SQL Generator...)? If so, is the source code backed up? | |
| Have you implemented logging and monitoring? | |
| Have any JNDI datasources been added to context.xml or server.xml? | |
| How many metadata models have been created and used? Are the metadata models secured by role? | |
| How many Mondrian schemas and cubes have been created and used? Are the schemas secured by role? | |
| How many users run reports? Where are those users located geographically? | |
| Have you exported your BA repository? | |
| Have you exported your usage data? | |

Table 4: Business Analytics Content

| Question | Notes |
|---|-------|
| What types of content are being used and how many of each (xanalyzer, xdash, prpt, xaction, waqr, jpivot...)? | |
| Which repository is used to store content (MySQL, PostgreSQL, Oracle, SQL Server)? | |
| Are there any reports that are executed on a schedule? If so, which reports on what schedule? | |
| Have ACLs been applied to specific reports? Or just defaults? | |
| If using PRD or xactions, what are the datasource types (SQL, MDX, PDI...)? | |
| If using PRD- have drill links been defined and are they using manual linking or Pentaho file linking? | |

Table 5: Data Integration

| Question | Notes |
|--|-------|
| What scripts are used to start/run Pan/Kitchen/Carte/DI-only server? | |
| Which repository is used to store transformations and jobs? (files, file repository, database repository, enterprise repository) | |
| What is the frequency of job execution? | |
| What is the number of jobs executing each hour/day/week? | |
| What is the number of transformations executing in those jobs? | |
| Are the jobs running on one server or several servers? | |
| How are jobs scheduled? | |
| What security is used for developers and operators? | |
| Have you exported your DI repository ? | |