Installation Paths for CDH with Cloudera Manager

Every phase is required, but there are multiple ways to accomplish every phase, depending on your organization's policies and requirements. Generally, the mechanisms tagged ♞ are easiest for interactive use, and the mechanisms tagged ♣ are best for scripting and automation.

Phase 0: Set up Databases

The CM Server, management services, and some CDH services require databases.

♞ CM Installer

Under the hood, installs "Embedded" PostgreSQL packages and Cloudera Manager packages.

Requires Internet access and sudo access to server machine.

♣ CM "Embedded" Postgresql

```
yum install cloudera-manager-server-db
service cloudera-manager-server-db start
```

Installs a PostgreSQL daemon on port 7432 in private / var/lib/cloudera-scm-server-db directory.

Manually
(MySQL, Oracle, PostgreSQL)

```
CREATE DATABASE …; GRANT …
```

Prepare databases per documentation.

Phase 1: Install CM Server

The Cloudera Manager server is installed on one host.

♣ Packages

```
yum install cloudera-manager-server
    cloudera-manager-daemons
vi /etc/cloudera-scm-server/db.properties
service cloudera-manager-server start
```
Starts CM server on port 7180.

Tarball

```
tar xzvf -C ...
```

Does not require root or use of system packages.

Phase 2: Install CM Agents

The Cloudera Manager Agent is required on every host.

♞ CM Wizard

Asks for SSH creds (password or key) for root or sudo-enabled user.

♣ Packages (on every machine)

```
yum install cloudera-manager-agent
    cloudera-manager-daemons
vi /etc/cloudera-scm-agent/config.ini
service cloudera-manager-agent start
```

Tarball

```
tar xzvf ...
vi config.ini
start agent
```

Phase 3: Install CDH

Installing CDH needs to be done on every host, and can be done either using Parcels (downloaded automatically from Cloudera or manually) or via system packages.

♞ CM Wizard (cont)

Offers choice of CDH version and repo. Installs either packages or parcels.

♣ Parcels

Remote or local repo or manual unpacking. API or UI.

♣ Packages

```
yum install hadoop
zookeeper hue
oozie [et al.]
```

Tarballs (manual)

Unpack per-project CDH tarballs, and configure agent environment. Not suggested. (Prefer parcels.)

Phase 4: Configure CDH

One everything is installed, the cluster can be configured, which includes choosing what processes run on which machines.

♞ CM Wizard (cont)

Choice of services, role assignments, and key configurations. Auto-configures many config options.

♣ API

```
POST /api/<version>/cm/deployment
```

Best for scripting pre-configured deployments.