**How to Deploy an IBM OpenPOWER Cluster  
for IBM Cloud Private**

Version 1.0

# Introduction

This document, along with referenced links, describes a comprehensive set of instructions, rules, and automation tools for building an IBM® OpenPOWER-based cluster that is created for IBM Cloud Private(ICP).

# High-level deployment steps

**Note:** Each step is described in more detail in the sections that follow.

|  |  |
| --- | --- |
| 1 | [Acquire the hardware](#_Step_1:_Acquire). |
| 2 | [Choose the basic configuration parameters](#_Step_2:_Choose). |
| 3 | [Prepare the deployer node](#_Step_4:_Configure). |
| 4 | [Rack and cable the hardware](#_Step_3:_Rack). |
| 5 | [Configure the cluster using the Cluster POWER-Up tool](#_STEP_5:_CONFIGURE). |
| 6 | [Post-deployment](#_STEP_6:_COMPLETE). |
| Appendix | [Application Software](#_STEP_8:_INSTALL). |

# Step 1: Acquire the hardware

Go to the following link to view the *ICP Design Proposal*, which shows the required hardware. IN BOX FOLDER FOR APPROVAL

[https://github.com/open-power-ref-design/accelerated-db/blob/master/docs/Accelerated%20Database%20Deployment%20Design%20Proposal.pdf](https://github.com/open-power-ref-design/accelerated-db/blob/master/docs/Accelerated%20Database%20Deployment%20Design%20Proposal.pdf%20)

Go to the following link to obtain the bill of materials, which lists the required parts. NEED TO GET BOM SPREADSHEET TEMPLATE

[https://github.com/open-power-ref-design/accelerated-db/blob/master/docs/Accelerated%20Database%20Deployment%20BOM.pdf](https://github.com/open-power-ref-design/accelerated-db/blob/master/docs/Accelerated%20Database%20Deployment%20BOM.pdf%20)

If you do not already have the needed parts, go to the following link to contact an IBM representative for ordering and purchasing assistance.

<https://www-01.ibm.com/marketing/iwm/dre/signup?source=MAIL-power&disableCookie=Yes>

# Step 2: Choose the basic configuration parameters

To facilitate faster automated configuration of the overall solution, collect the parameters in *Table 1* before starting. This data is edited into a *config.yml* file, which is used to automatically configure and deploy the entire solution.

Table 1. Configuration parameters

|  |  |  |
| --- | --- | --- |
| Parameter | Description | Example |
| Domain name |  | Ibm.com |
| Upstream DNS servers | While a domain name system (DNS) server is configured within the cluster, upstream DNS servers must be defined because the names cannot otherwise be resolved. | \*4.4.4.4, 8.8.8.8 as default public upstream DNS servers |
| Deployment node host name | The name of the deployment node. | depnode |
| Management network IP address | Management for the cluster takes place on its own internal network. | 192.168.3.3.24 |
| Data network IP address | Labeled *interconnect* in the config.yml file in the example below. | 10.0.0.1/24 |
| Management switch IP address | Labeled *ipaddr-mgmt-switch* in the config.yml file in the example below. | 192.168.3.5 |
| Data switch IP addresses | Labeled *ipaddr-data-switch* in the config.yml file in example below. | 1.2.3.178 |
| Default login data | Both IDs and passwords. | BMC network, OS Mgmt network |
| Data node hostnames and IPs addresses | Each node in the cluster needs a host name and an IP address for each of the management and data networks. | |  |  |  | | --- | --- | --- | | **Name** | **Management IP** | **Data IP** | | Min-1 | 192.168.3.102 | 10.0.0.2 | | Min-2 | 192.168.3.104 | 10.0.0.4 | | Min-3 | 192.168.3.106 | 10.0.0.6 | | Min-4 | 192.168.3.108 | 10.0.0.8 | |

Go to the following link to see more options in the *config.yml* file.

<https://github.com/open-power-ref-design/ICP-install/blob/master/yamls/config.yaml>

# Step 3: Prepare the deployer node

The deployer node is used to obtain the latest software and deployment tools from GitHub and populate the cluster. The deployer node can be established as a temporary or permanent server. It can be set up as an IBM POWER8® LC or x86 server with the following minimum characteristics:

* Two cores and 32 GB RAM
* Three network-interface connections: 1 GbE Intelligent platform management interface (IPMI), 1 GbE (Mgmt), and 10 GbE (high-speed) .
* Ubunutu 16.04 LTS must be installed before beginning with deployment.

If you do not already have Ubuntu, it is available at the following sources:

* Power8-LC servers: <https://www.ubuntu.com/download/server/power8>
* x86 servers: <https://www.ubuntu.com/download/server>

# Step 4: Rack and cable the hardware

The design proposal(make link) prescribes the IBM LC922 server (product ID: 9006-22p) as control plane and worker building block for the Cloud Native configuration (pages 2-3). For the Cognitive Computation configuration, the design proposal prescribes the IBM LC922 server (product ID: 9006-22p) as control plane and AC922 server (product ID: 8335-GTH) as the worker building block to use the unique high-speed NVIDIA® NVLink™ bus between its CPU and GPUs (page 4-5)

Go to the following link for more information about the LC922 specific configuration.

<http://www.redbooks.ibm.com/abstracts/redp5495.html?Open>

Go to the following link for more information about the AC922 specific configuration.

<http://www.redbooks.ibm.com/abstracts/redp5494.html?Open>

This step describes how to cable and rack the servers in the rack so they are networked together correctly. It is not intended to be comprehensive. For example, it is assumed you can cable the servers to the power source as needed.



Figure 1a. Back view of server LC922



Figure 1b. Back view of server AC922

**Note:** While these servers are capable of sharing ports (multi-function ports), automation requires the port to be set up as a baseboard management controller (BMC) for data only.

## **Racking the components**

The racking rules specify where to place the servers and switches and where to connect the cables.

The suggestedracking rules, shown in *Figure 2* on page 6, focus on enabling:

* Rack modularity
* Consistency
* Expandability
* Ease of servicing, repurposing, shipping, and cooling

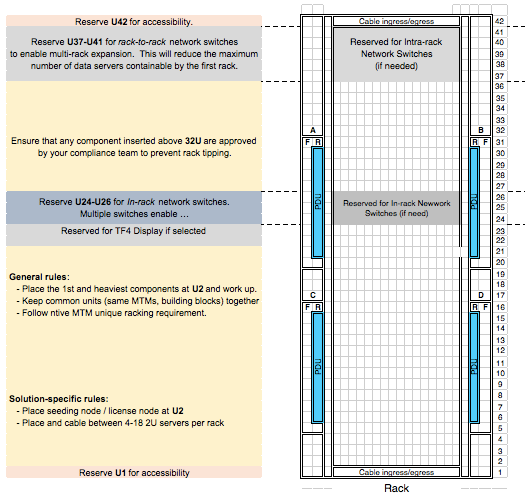


Figure 2. Suggested Racking Rules

## Additional racking rules

The following additional racking rules are for IBM Cloud Private design, which supports in-server storage only. A future version will add support for external storage.

Place the intra-rack (leaf) network switches in slots U24 - U26 as follows:

* Place the 10G/40G/IB data plane switch in slot 26U (parts 8828-E36, 8831-NF2, or 7120-64C)
* Leave slotU25 open. Reserve this for later use of short-depth devices. For more information see the *Bill of Materials* document: NEED SPREADSHEET TEMPLATE

[https://github.com/open-power-ref-design/accelerated-db/blob/master/docs/Accelerated%20Database%20Deployment%20BOM.pdf](https://github.com/open-power-ref-design/accelerated-db/blob/master/docs/Accelerated%20Database%20Deployment%20BOM.pdf%20)

* Place the 1G management plane switch in slot 24U.

The Figure 3 on page 8 shows an example of these additional racking rules.



Figure 3. Minimum Cloud Native cluster configuration



Figure 4. Minimum Cognitive Computing cluster configuration

## **Cable the components together**

To follow the cabling rules, cable the like labels on the servers to the applicable network switches. *Figure 4* shows an example approach for a four-server cluster configuration. The rear server view is shown with labels for five servers named that correspond to LC1, LC2, LC3, AC1, and AC2(from top down).

*\*In a cloud native configuration AC922(noted as AC) will be a LC922(noted as LC).*



Figure 4. Example cabling for a minimum four Cognitive Computing cluster network

*Figure 5* and *Figure 6* show a network-switch port view. The labels on the ports correspond to the server labels above and also to the information in the *config.yml* and *inventory.yml* files.

\*In a cloud native configuration AC922(noted as AC) will be a LC922.



Figure 5. Lenovo G8264CS – 10 GbE data network switch cabling scheme



Figure 6. Lenovo G8052 – 1 GbE management network switch cabling scheme

***\*In a cloud native configuration AC922(noted as AC) will be a LC922.***

# Step 5: Configure the cluster using the Cluster POWER-Up tool

This step covers the power on, initialization, configuration, and installation of a cluster solution. This deployment kit provides an automated method to quickly and more predictably go from assembly to a tuned operational state of the cluster’s infrastructure. This is referred to as *hardware genesis*.

Cluster POWER-Up occurs once at the beginning of the cluster solution lifecycle. The open-sourced automation scripts are available and can be reused for maintenance and cluster expansion.

The Cluster POWER-Up tool automatically initializes and configures the hardware by accomplishing the following tasks:

* Reading the *config.yml* files with edited environment-specific changes
* Driving the BMCs to populate the IP addresses to the nodes
* Detecting and populating relevant configuration data to the deployer node
* Deploying the required operating system images to the server nodes
* Configuring the network switches
* Configuring all server management and data nodes (network interfaces, GPU drivers, and so on)

When the Cluster POWER-Up tool completes its process, control of the cluster is transferred to the ICP master node.

All Cluster POWER-Up Cluster tool procedures are built into automation described in *Perform the deployment of* **Cluster POWER-Up**

**Perform the deployment of Cluster POWER-Up** on page 15. Go to the following link for more information about this process in the Cluster POWER-Up deployment README file.

<https://github.com/open-power-ref-design-toolkit/power-up/blob/master/README.rst>

Go to the following link for more information about the procedure overview and deployment automation procedures found in the ICP-install README file.

<https://github.com/open-power-ref-design/ICP-install/blob/master/README.md>

## **Obtain the default configuration file**

The Cluster POWER-Up automation uses a configuration file to specify the target cluster configuration. The deployment tooling uses this YAML text file to specify the IP address locations of the managed switches and the system nodes attached to the switches as well as other useful details for deployment process.

Go to the following link for a copy of the OpenPOWER Cluster for IBM Cloud Private configuration file that deploys and initiates ICP.

<https://github.com/open-power-ref-design/ICP-install/blob/master/yamls/config.yaml>

## **Customize the configuration file for the environment**

The *config.yml* file contains a lot of configuration information. To enable a cluster tailored to specific environment, edit the .yml file with the configuration parameters that were collected in *Step 2*: Choose the basic configuration parameters, replacing the red text with your data. A sample config.yml file has been provided, configuration.yaml. The following excerpt focuses on the lines to edit.

|  |
| --- |
| ~ ~ ~ ~ ~ ~ ~ licensing comment and YAML ~ ~ ~ ~ ~  switches:  mgmt:  - label: mgmt1  class: lenovo  userid: admin *🡸 Type your management switch user id here.*  password: admin *🡸 Type your management switch password here.*  interfaces:  - type: inband  ipaddr: 192.168.3.5 *🡸 Type your management switch IP address here.*  port: 1  data:  - label: data1  class: Lenovo  userid: admin *🡸 Type your data switch user id here.*  password: admin *🡸 Type your data switch password here.*  interfaces:  - type: inband  ipaddr: 1.2.3.178 *🡸 Type your data switch IP address here.*  ~ ~ ~ ~ ~ ~ ~ YAML and comments ~ ~ ~ ~ ~  interfaces:  - label: ipmi  description: pxe interface  iface: eth0  method: dhcp  - label: pxe  description: pxe interface  iface: eth0  method: dhcp  - label: external  description: Organization site or external network  iface: eth1  method: static  address\_list:  - 1.2.3.18-1.2.3.20 *🡸 Type your client IP here().*  - 1.2.3.38 *🡸 Type your client IP here.*  - 1.2.3.76 *🡸 Type your client IP here.*  netmask: 255.255.255.0 *🡸 Type your netmask IP here.*  broadcast: 1.2.3.255 *🡸 Type your broadcast IP here.*  gateway: 1.2.3.1 *🡸 Type your gateway IP here.*  dns\_search: aus.stglabs.ibm.com  dns\_nameservers: 1.2.3.200 *🡸 Type your nameserver IP here.*  - label: interconnect  description: Private 10G Data Network to Interconnect Cluster  iface: eth2  method: static  address\_start: 10.0.0.1  netmask: 255.255.255.0  broadcast: 10.0.0.255  networks:  external:  description: Organization site or external network  addr: 1.2.3.4/24 *🡸 Type your subnet address here.*  broadcast: 1.2.3.255 *🡸 Type your broadcast IP here.*  gateway: 1.2.3.1 *🡸 Type your gateway IP here.*  dns-nameservers: 1.2.3.4 *🡸 Type your nameserver IP here.*  dns-search: aus.stglabs.ibm.com  method: static  eth-port: eth10  interconnect:  description: Private 10G Data Network to Interconnect Cluster  addr: 10.0.0.0/24  broadcast: 10.0.0.255  method: static  eth-port: eth11  ~ ~ ~ ~ ~ ~ ~ bunch of YAML and comments ~ ~ ~ ~ ~  node-templates:  controller1:  hostname: min *🡸 Type your hostname here.*  userid-ipmi: ADMIN *🡸 Type your userid here.*  password-ipmi: admin *🡸 Type your password here.*  cobbler-profile: ubuntu-16.04.1-server-ppc64el  ~ ~ ~ ~ ~ ~ ~ bunch of YAML and comments ~ ~ ~ ~ ~ |

Editable portions of the Config.yml file

Go to the following link for further details on all the config file entries.

<https://github.com/open-power-ref-design-toolkit/power-up/blob/master/docs/Config-Specification.rst>

## **The inventory file**

The inventory file is a YAML text file that contains the entire inventory of the cluster, captured during the Cluster POWER-Up process. It can be used to feed subsequent automation (management, deployment, and so on). Do not edit this file manually.

Go to the following link for the generic master copy of the latest inventory file.

NEED COPY OF INVENTORY OUT IN REPO(inventory40.yml).

[https://github.com/open-power-ref-design/accelerated-db/blob/mapd\_initial/master\_inventory.yml](https://github.com/open-power-ref-design/accelerated-db/blob/mapd_initial/master_inventory.yml%20)

The file contains the configuration specifics of each network switch and server node. The *Switches* data structure indicates the types of switches (management, spine, or leaf), their IP addresses, and associated log in credentials. The following sample inventory data structure contains the management and leaf switches attributes. The *Nodes* data structure specifies the type of node controller, its network properties, and its system architecture (ppc64 or x86). The following snippet shows the data structure.

|  |
| --- |
| ***nodes:***  ***- label: min***  ***hostname: min-1***  ***rack\_id: null***  ***ipmi:***  ***ipaddrs:***  ***- 192.168.10.21***  ***macs:***  ***- 70:e2:84:14:0a:ac***  ***password: admin***  ***ports:***  ***- 5***  ***switches:***  ***- mgmt1***  ***userid: ADMIN***  ***pxe:***  ***devices:***  ***- eth0***  ***ipaddrs:***  ***- 192.168.20.21***  ***macs:***  ***- 70:e2:84:14:0a:ae***  ***ports:***  ***- 6***  ***rename:***  ***- true***  ***switches:***  ***- mgmt1***  ***data:***  ***devices:***  ***- eth1***  ***- eth2***  ***macs:***  ***- null***  ***- null***  ***ports:***  ***- 1***  ***- 2***  ***rename:***  ***- true***  ***- true***  ***switches:***  ***- data1***  ***- data1***  ***os: &id001***  ***profile: ubuntu-18.04-server-ppc64el***  ***users:***  ***- name: ubuntu***  ***password: $6$Utk.IILMG9.$EepS/sIgD4aA.qYQ3voZL9yI3/5Q4vv.p2s4sSmfCLAJlLAuaEmXDizDaBmJYGqHpobwpU2l4rJW.uUY4WNyv.***  ***groups: sudo***  ***install\_device: /dev/sdj***  ***roles: null***  ***interfaces:***  ***- label: pxe***  ***description: pxe interface***  ***iface: eth0***  ***method: dhcp***  ***- label: external***  ***description: Organization site or external network***  ***iface: eth1***  ***method: static***  ***netmask: 255.255.255.0***  ***broadcast: 9.3.3.255***  ***gateway: 9.3.3.1***  ***dns\_search: aus.stglabs.ibm.com***  ***dns\_nameservers: 9.3.1.200***  ***address: 9.3.3.18***  ***- label: interconnect***  ***description: Private 10G Data Network to Interconnect Cluster***  ***iface: eth2***  ***method: static***  ***netmask: 255.255.255.0***  ***broadcast: 10.0.0.255***  ***address: 10.0.0.1*** |

When Cluster POWER-Up completes, the *inventory40.yml* file is stored on the deployment node in the root ~*/power-up* directory.

## **Perform the deployment of Cluster POWER-Up**

**Perform the deployment of Cluster POWER-Up**

To deploy the Cluster POWER-Up tool, clone the tool repo:

$ git clone https://github.com/open-power-ref-design-toolkit/power-up.git

Create a configuration file that includes this snippet from the [icp-config.yaml](https://github.com/open-power-ref-design/ICP-install/blob/master/icp-config.yaml). Example [config.yaml](https://github.com/open-power-ref-design/ICP-install/blob/master/config.yaml)

After the *config.yaml* is updated, to initiate the cluster deployment:

$ pup deploy myconfig.yml

**NOTE:** You will have to enter your sudo password once the deploy is started and then hit Enter a few times. This process will take over 1 hour to complete.

The deploy command starts with the installation of some Linux distribution and finishing by executing the icp-install.sh and installing IBM Cloud Private-CE version 3.1.0. Make sure your configuration file contains the info from [icp-config.yaml](https://github.com/open-power-ref-design/ICP-install/blob/master/yamls/icp-config.yaml) in this repo.

# Step 6: Post-deployment

Currently full deployment automated within the deployment explained in step 5.

# Appendix: Application Software

Further information on IBM Cloud Private including pointers to relevant sites are included.

## **Installing IBM Cloud Private CE**

ICP is installed via automation found in the [scripts](https://github.com/open-power-ref-design/ICP-install/tree/master/scripts) directory in this repo, integrated into the overall deployment automation.

IBM documentation: <https://www.ibm.com/support/knowledgecenter/en/SSBS6K_3.1.0/installing/install_containers_CE.html>

It utilizes 2 files taken from other repos, Docker on POWER and ICP-scripts

<https://github.com/Unicamp-OpenPower/docker_on_power>

<https://github.com/rpsene/icp-scripts>

# References

The following links and documents provide more information related to this document:

LC922:

* [IBM Power Systems LC921 and LC922: Technical Overview and Introduction](http://www.redbooks.ibm.com/abstracts/redp5495.html?Open)
* More IBM Power System® LC922 (9006-22P) reference material located in the [IBM Knowledge Center](https://www.ibm.com/support/knowledgecenter/9006-22P/p9hdx/9006_22p_landing.htm)

AC922

* [IBM Power System AC922 Technical Overview and Introduction](http://www.redbooks.ibm.com/abstracts/redp5494.html?Open)
* More IBM Power System® S822LC (8335-GTB) reference material located in the [IBM Knowledge Center](https://www.ibm.com/support/knowledgecenter/8335-GTH/p9hdx/8335_gth_landing.htm)

IBM Private Cloud

* [IBM Private Cloud documentation](https://www.ibm.com/support/knowledgecenter/en/SSBS6K_3.1.0/kc_welcome_containers.html)

®

#### Copyright and disclaimers

© Copyright International Business Machines Corporation 2017

Printed in the United States of America May 2017

IBM, the IBM logo, and ibm.com are trademarks or registered trademarks of International Business Machines Corp., registered in many jurisdictions worldwide. Other product and service names might be trademarks of IBM or other companies. A current list of IBM trademarks is available on the Web at “Copyright and trademark information” at [www.ibm.com/legal/copytrade.shtml](http://www.ibm.com/legal/copytrade.shtml).

NVLink is a trademark of the NVIDIA Corporation in the United States, other countries, or both.

The OpenPOWER word mark and the OpenPOWER Logo mark, and related marks, are trademarks and service marks licensed by OpenPOWER.

Other company, product, and service names may be trademarks or service marks of others.

All information contained in this document is subject to change without notice. The products described in this document are NOT intended for use in applications such as implantation, life support, or other hazardous uses where malfunction could result in death, bodily injury, or catastrophic property damage. The information contained in this document does not affect or change IBM product specifications or warranties. Nothing in this document shall operate as an express or implied indemnity under the intellectual property rights of IBM or third parties. All information contained in this document was obtained in specific environments, and is presented as an illustration. The results obtained in other operating environments may vary.

This document is intended for the development of technology products compatible with Power Architecture®. You may use this document, for any purpose (commercial or personal) and make modifications and distribute; however, modifications to this document may violate Power Architecture and should be carefully considered. Any distribution of this document or its derivative works shall include this Notice page including but not limited to the IBM warranty disclaimer and IBM liability limitation. No other licenses (including patent licenses), expressed or implied, by estoppel or otherwise, to any intellectual property rights are granted by this document.

THE INFORMATION CONTAINED IN THIS DOCUMENT IS PROVIDED ON AN “AS IS” BASIS. IBM makes no representations or warranties, either express or implied, including but not limited to, warranties of merchantability, fitness for a particular purpose, or non-infringement, or that any practice or implementation of the IBM documentation will not infringe any third party patents, copyrights, trade secrets, or other rights. In no event will IBM be liable for damages arising directly or indirectly from any use of the information contained in this document.

IBM Systems

294 Route 100, Building SOM4

Somers, NY 10589-3216

The IBM home page can be found at [**ibm.com®**](http://www.ibm.com/).

Version 1.1

25 May 2017