RIMA Cluster Manager 9.2

Edge Manual

Revision: f502c1099

Date: Wed May 28 2025

©2024 NVIDIA Corporation & affiliates. All Rights Reserved. This manual or parts thereof may not be reproduced in any form unless permitted by contract or by written permission of NVIDIA Corporation.

Trademarks

Linux is a registered trademark of Linus Torvalds. PathScale is a registered trademark of Cray, Inc. Red Hat and all Red Hat-based trademarks are trademarks or registered trademarks of Red Hat, Inc. SUSE is a registered trademark of SUSE LLC. NVIDIA, CUDA, GPUDirect, HPC SDK, NVIDIA DGX, NVIDIA Nsight, and NVLink are registered trademarks of NVIDIA Corporation. FLEXlm is a registered trademark of Flexera Software, Inc. PBS Professional, and Green Provisioning are trademarks of Altair Engineering, Inc. All other trademarks are the property of their respective owners.

Rights and Restrictions

All statements, specifications, recommendations, and technical information contained herein are current or planned as of the date of publication of this document. They are reliable as of the time of this writing and are presented without warranty of any kind, expressed or implied. NVIDIA Corporation shall not be liable for technical or editorial errors or omissions which may occur in this document. NVIDIA Corporation shall not be liable for any damages resulting from the use of this document.

Limitation of Liability and Damages Pertaining to NVIDIA Corporation

The NVIDIA Bright Cluster Manager product principally consists of free software that is licensed by the Linux authors free of charge. NVIDIA Corporation shall have no liability nor will NVIDIA Corporation provide any warranty for the NVIDIA Bright Cluster Manager to the extent that is permitted by law. Unless confirmed in writing, the Linux authors and/or third parties provide the program as is without any warranty, either expressed or implied, including, but not limited to, marketability or suitability for a specific purpose. The user of the NVIDIA Bright Cluster Manager product shall accept the full risk for the quality or performance of the product. Should the product malfunction, the costs for repair, service, or correction will be borne by the user of the NVIDIA Bright Cluster Manager product. No copyright owner or third party who has modified or distributed the program as permitted in this license shall be held liable for damages, including general or specific damages, damages caused by side effects or consequential damages, resulting from the use of the program or the un-usability of the program (including, but not limited to, loss of data, incorrect processing of data, losses that must be borne by you or others, or the inability of the program to work together with any other program), even if a copyright owner or third party had been advised about the possibility of such damages unless such copyright owner or third party has signed a writing to the contrary.

Table of Contents

Table of Contents i
0.1 About This Manual iii
0.2 About The Manuals In General iii
0.3 Getting Administrator-Level Support iii
0.4 Getting Professional Services iv

1 Introduction **1**
1.1 Cloud Computing Vs Edge Computing 1
1.2 High Speed Monitoring And Local Processing 1

2 Bright Edge **3**
2.1 Bright Edge 3
2.1.1 Defining The Edge Site 4
2.1.2 Adding Nodes To Pre-existing Edge Sites With `cmsh` 10
2.1.3 Viewing Edge Sites Using `cmsh` 10
2.1.4 Viewing Edge Sites Using Bright View 11
2.1.5 Create Edge ISO 11
2.1.6 Edge ISO Node Installer 13
2.1.7 Edge Directors 16
2.1.8 Edge Nodes 16

3 Installing Slurm To An Edge Site **17**
3.1 Preparation 17
3.2 Installation 17

Preface

Welcome to the *Edge Manual* for RIMA Cluster Manager 9.2.

0.1 About This Manual

This manual is aimed at helping cluster administrators install, understand, configure, and manage the edge computing capabilities of NVIDIA Bright Cluster Manager. The administrator is expected to be reasonably familiar with the *Administrator Manual*.

0.2 About The Manuals In General

Regularly updated versions of the RIMA Cluster Manager 9.2 manuals are available on updated clusters by default at `/cm/shared/docs/cm`. The latest updates are always online at `http://support.brightcomputing.com/manuals`.

- The *Installation Manual* describes installation procedures for the basic cluster.
- The *Administrator Manual* describes the general management of the cluster.
- The *User Manual* describes the user environment and how to submit jobs for the end user.
- The *Cloudbursting Manual* describes how to deploy the cloud capabilities of the cluster.
- The *Developer Manual* has useful information for developers who would like to program with the cluster manager.
- The *Machine Learning Manual* describes how to install and configure machine learning capabilities with the cluster manager.
- The *Containerization Manual* describes how to manage containers with the cluster manager.

If the manuals are downloaded and kept in one local directory, then in most pdf viewers, clicking on a cross-reference in one manual that refers to a section in another manual opens and displays that section in the second manual. Navigating back and forth between documents is usually possible with keystrokes or mouse clicks.

For example: `<Alt>-<Backarrow>` in Acrobat Reader, or clicking on the bottom leftmost navigation button of xpdf, both navigate back to the previous document.

The manuals constantly evolve to keep up with the development of the cluster manager environment and the addition of new hardware and/or applications. The manuals also regularly incorporate customer feedback. Administrator and user input is greatly valued at Bright Computing. So any comments, suggestions or corrections will be very gratefully accepted at `manuals@brightcomputing.com`.

There is also a feedback form available via Bright View, via the menu icon, ☰, following the click-path:

☰→`Help`→`Feedback`

0.3 Getting Administrator-Level Support

If the reseller from whom the cluster manager was bought offers direct support, then the reseller should be contacted.

Otherwise the primary means of support is via the website `https://support.brightcomputing.com`. This allows the administrator to submit a support request via a web form, and opens up a trouble ticket. It is a good idea to try to use a clear subject header, since that is used as part of a reference tag as the ticket progresses. Also helpful is a good description of the issue. The followup communication for this ticket goes via standard e-mail. Section 9.2 of the *Administrator Manual* has more details on working with support.

0.4 Getting Professional Services

Bright Computing normally differentiates between

- professional services (customer asks Bright Computing to do something or asks Bright Computing to provide some service), and
- support (customer has a question or problem that requires an answer or resolution).

Professional services can be provided after consulting with the reseller, or the Bright account manager.

1

Introduction

1.1 Cloud Computing Vs Edge Computing

Cloud computing is traditionally about the concept of end users using resources that are located in a cloud elsewhere. The cloud is the central coordinator, and end users use the resources that are in the cloud rather than using their local resources.

As computing power has become cheaper over time, and resource use has grown, it has in many cases become more financially attractive to shift the emphasis of resource coordination, from the center of the cloud (core of the cloud), over to the local resources which are at the *edge*. These local devices are then called *edge nodes*.

A strong case for edge computing is when the following resource requirements are easier to provide locally via local devices, than via central processing in the cloud:

- low latency
- high bandwidth consumption
- high CPU cycle consumption

For example, a self-driving car requires a low latency, high bandwidth, and high CPU cycle consumption in order to ensure a speedy and safe response to traffic requirements. Attempting to run a self-driving car via central processing in the cloud would be impractically slow or prohibitively dangerous.

Generally, edge computing is regarded as a way to have a geographically spread-out cluster make more local use of its computing resources. A geographically spread-out cluster typically already has plenty of CPU cycles, bandwidth, and low latency at its regular nodes. So, for such a geographically spread-out cluster, making more local use of its computing resources tends to mean granting extra autonomy to the edge computing devices, and making them more independent of the head node.

To achieve this greater autonomy, NVIDIA Bright Cluster Manager uses an *edge director*. This is somewhat similar to the cloud director, but is required to be geographically close to the edge nodes, and is also optimized for edge requirements.

1.2 High Speed Monitoring And Local Processing

The importance of being local and autonomous is often due to the environment that the regular nodes are in. The environment is typically under high speed monitoring by many sensors linked to the regular nodes. The data values obtained by the sensors are processed very quickly by the nodes. Such high speed processing of the monitoring data values can typically only reasonably be achieved by the nodes managing the processing locally as much as possible, rather than having the nodes managed by a head node a large distance away.

2

Bright Edge

2.1 Bright Edge

The Bright Edge feature of NVIDIA Bright Cluster Manager allows a single cluster to span many geographical locations ("one cluster, multiple locations"). Typical use cases are:

- HPC: organizations that have compute resources located in different cities or countries
- IoT: companies that have "edge" locations with the required compute resources at each location

The cluster manager can be used to deploy and manage resources at edge locations from the central head node.

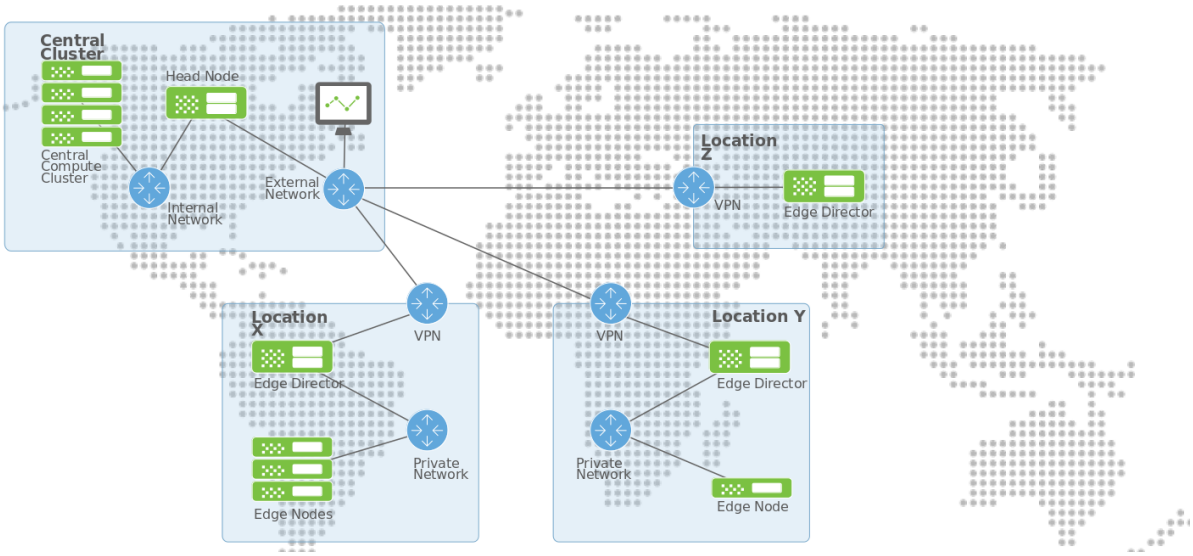

Figure 2.1: Bright Edge: The Big Picture

Bright Edge sites comprise an edge director and edge nodes.

- The edge director must be reachable from the central head node. The edge director forwards requests from the edge nodes to the central head node when required.
- The edge nodes are similar to regular nodes, and are provisioned by them PXE booting off the edge director. Unlike with regular nodes, no direct connection is required between the central head node and the edge nodes.

Items to check before creating edge sites:

- The cluster manager license must allow edge site creation
- The to-be-provisioned edge director must have an IP address that can be reached from the central head node
- Conversely, the central head node must have an IP address that can be reached by the edge director

Creating and deploying edge sites involve the following steps:

- Create the edge site using `cm-edge-setup`
- Create an edge ISO for provisioning the edge director
- Provision the edge director using the edge ISO
- Provision the edge nodes off the edge director

The following sections explain each of the preceding steps in further detail:

2.1.1 Defining The Edge Site

Edge sites are defined in the cluster manager using the TUI-based `cm-edge-setup`. This section goes through a `cm-edge-setup` session on the central head node that creates an edge site definition.

Running `cm-edge-setup` In Interactive Mode

Running `cm-edge-setup` without any options brings up the main edge setup screen (figure 2.2):

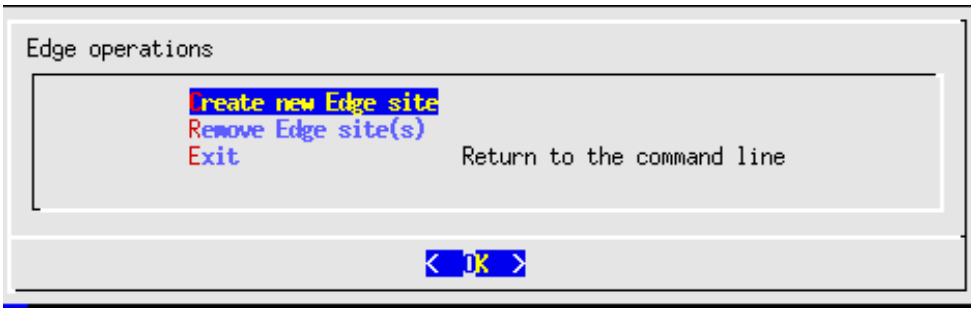

Figure 2.2: Edge setup main screen

A new edge site can be created by entering a series of parameters (figure 2.3):

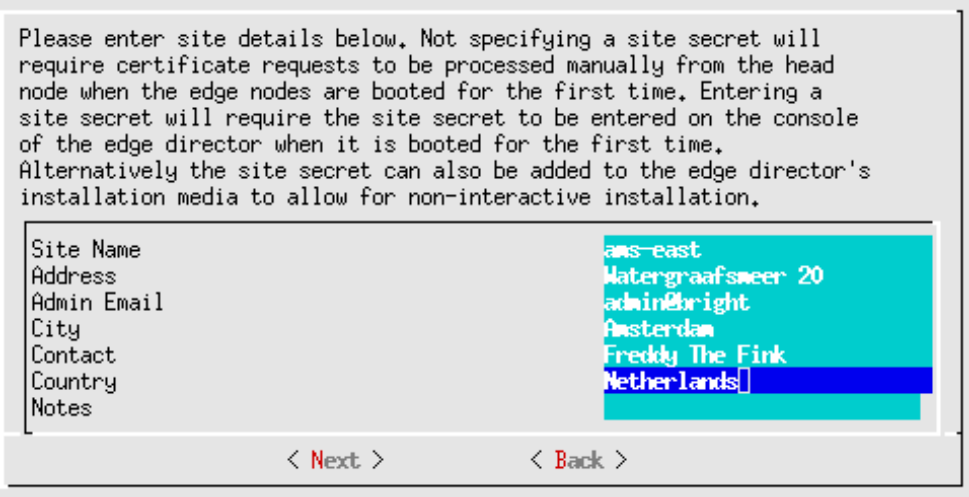

Figure 2.3: Entry of edge site parameters

A secret for the site should be entered (figure 2.4):

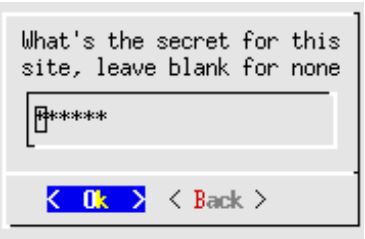

Figure 2.4: Entry of site secret

The site secret entry is reconfirmed by the administrator in a subsequent entry screen.

The next screen after that asks how the external network for the edge director should be set (figure 2.5):

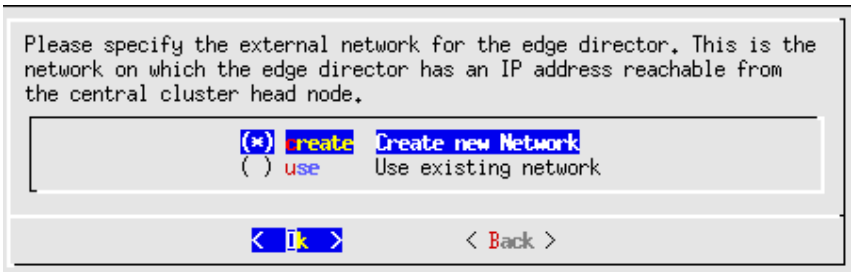

Figure 2.5: Selection of edge external network

- If networks defined as type `EdgeExternal` are found, then these networks are presented for selection (figure 2.6.
- If no networks of type `EdgeExternal` are found, then the only option is to create a new network (figure 2.7).

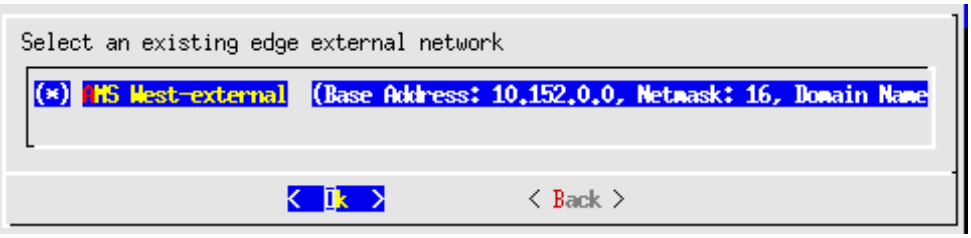

Figure 2.6: Selection of an existing edge external network

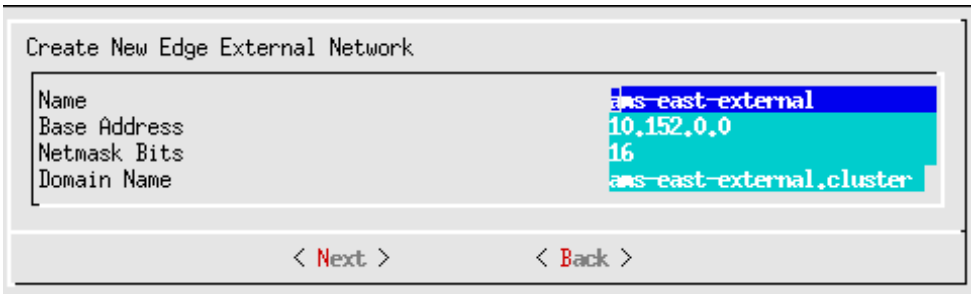

Figure 2.7: Creation of a new edge external network

Similarly to the external network configuration for the edge director, a screen comes up next that asks how the internal network for the edge director (figure 2.8) should be set:

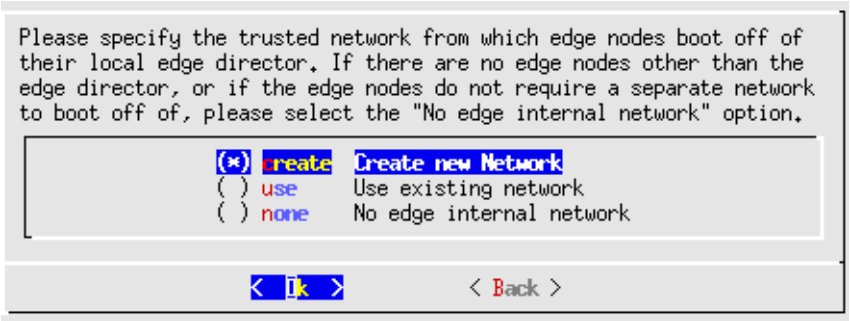

Figure 2.8: Selection of an existing edge internal network

Similarly to the external network configuration, the configuration for trusted internal network for the edge director presents the following similar options:

- If networks defined as type `EdgeInternal` are found, then these networks are presented for selection.
- If no networks of type `EdgeInternal` are found, then the only option is to create a new network (figure 2.9).

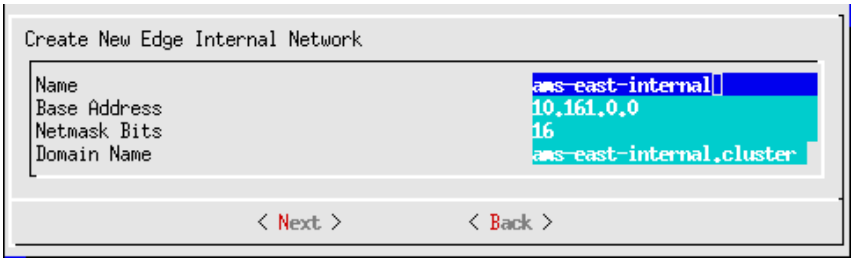

Figure 2.9: Creation of a new edge internal network

The next screen (figure 2.10) allows edge director parameters to be entered:

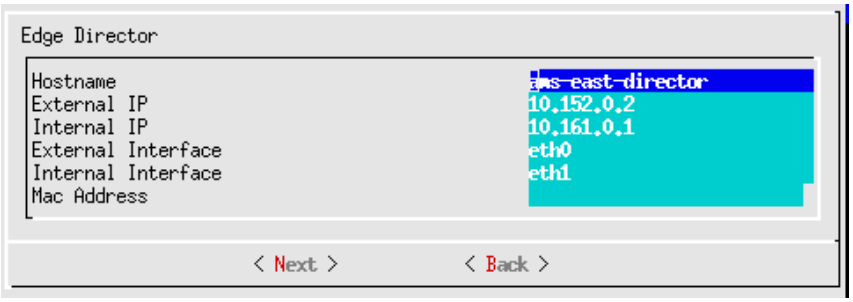

Figure 2.10: Entry of edge director parameters

In the next screen (figure 2.11) Edge director High Availability (HA) can be enabled:

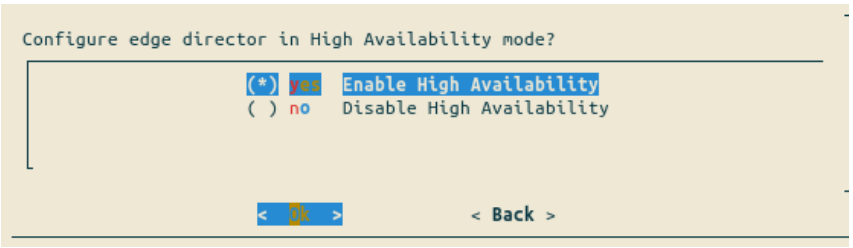

Figure 2.11: Choice of edge director HA

The cluster administrator may want

- `/cm/shared` to be shared from a common network storage to the edge directors
- `/cm/shared` to not be shared, but synchronized between edge directors

How to configure this is described on page 77 of the *Cloudbursting Manual*.

A screen later on (figure 2.12) allows secondary edge director parameters to be set:

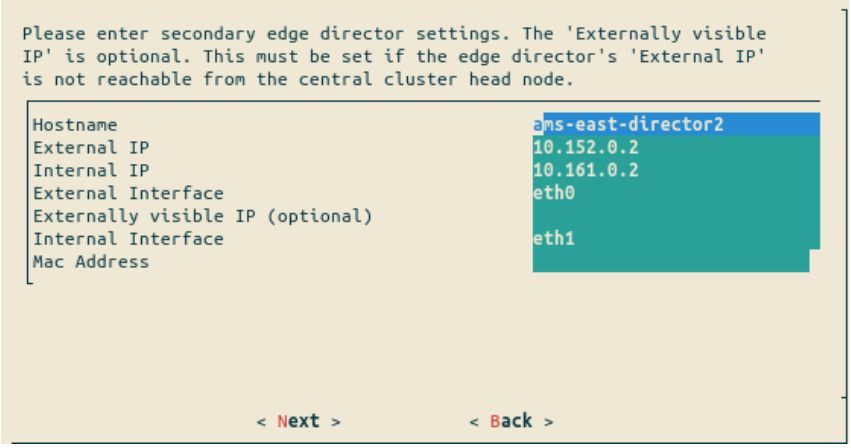

Figure 2.12: Entry of secondary edge director parameters

Shared IP settings for edge directors can be entered in the next screen (figure 2.13):

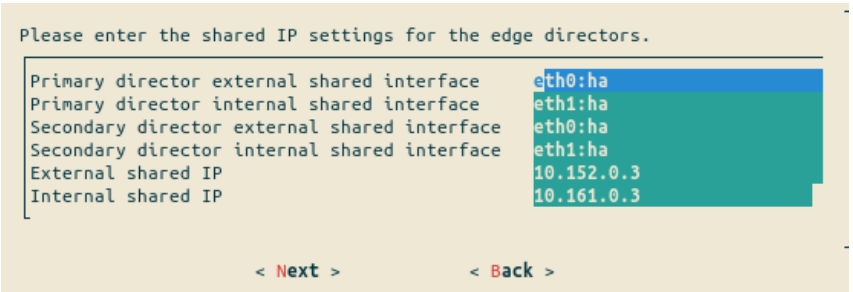

Figure 2.13: Entry of shared IP settings for edge directors

The edge nodes can now be configured. Individual nodes (figure 2.14), or a range of nodes (figure 2.15), can be configured:

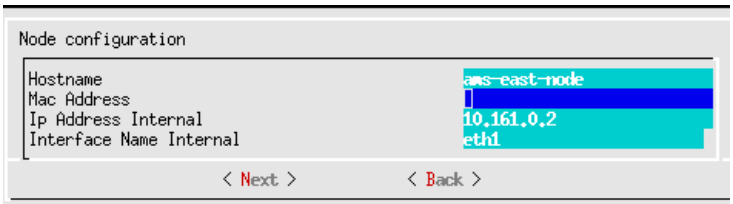

Figure 2.14: Definition of single edge node

Figure 2.15: Define multiple edge nodes

Running cm-edge-setup In Batch Mode

In the preceding section cm-edge-setup was used interactively to define edge sites. It can also be used non-interactively for the same purpose. This is done by saving a site configuration file at the end of the interactive setup. This is a YAML file, and it can be used to re-create the edge sites, or it can be used as a template to create new sites.

Example

```
[root@headnode ~]# cat /root/cm/edge/ams-west.yaml
######################################################################
##  This config file should be used with cm-edge-setup tool
##  Example:
##      cm-edge-setup -c <filename>
##
##  Generated by:
##      cm-edge-setup
##      cluster-tools-8.2-112301_cm8.2_b7ed6dbd8a
##      cmdline: /cm/local/apps/cm-setup/bin/cm-edge-setup
##  Generate on host:
##      smcluster
##  Date of generation:
##      Thu Dec  6 11:05:29 2018
##  MD5 checksum of everything after the closing comment:
##      5ca8aef31f8a047677c220824474e747
##      to compare:  grep -v '^##' <this_file> | md5sum
######################################################################
edge_sites:
- address: Kings
  admin_email: admin@bright
  city: Amsterdam
  contact: admin
  country: Amsterdam
  edge_director:
    category: edge-director
    hostname: ams-west-director
    interface_name_external: eth0
    interface_name_internal: eth1
    ip_address_external: 10.2.125.125
    ip_address_internal: 10.161.255.254
    mac_address: ''
    secondary:
      externally_visible_ip: ''
      hostname: edge-director2
      interface_name_external: eth0
      interface_name_internal: eth1
      ip_address_external: 10.2.125.126
      ip_address_internal: 10.161.255.253
      mac_address: ''
    shared_ip:
      ip_address_external: 10.2.125.127
      ip_address_internal: 10.161.255.252
      primary_interface_name_external: eth0:ha
      primary_interface_name_internal: eth1:ha
      secondary_interface_name_external: eth0:ha
      secondary_interface_name_internal: eth1:ha
    software_image: default-image
```

```
  edge_nodes:
  - category: edge-director
    hostname: ams-west-node001
    interface_name_internal: eth1
    ip_address_internal: 10.161.0.1
    mac_address: ''
  external_network:
    base_address: 10.2.0.0
    domainname: brightcomputing.com
    name: externalnet
    netmask_bits: 16
  internal_network:
    base_address: 10.161.0.0
    domainname: ams-west-internal.cluster
    name: ams-west-internal
    netmask_bits: 16
  notes: ''
  secret: xxxxxx
  site_name: ams-west
meta:
  command_line: /cm/local/apps/cm-setup/bin/cm-edge-setup
  date: Thu Dec  6 11:05:29 2018
  generated_with: Edge
  hostname: smcluster
  package_name: cluster-tools-8.2-112301_cm8.2_b7ed6dbd8a
  package_version: '112301'
```

2.1.2 Adding Nodes To Pre-existing Edge Sites With `cmsh`

Edge nodes can also be added to an existing edge site. This is typically required when no edge nodes were added during `cm-edge-setup`, or if the site is being expanded by adding more nodes. The addition can be done in the usual way, which is to first add the required node object with `cmsh` (section 2.5.3 of the *Administrator Manual*). The nodes are then added to the relevant edge site(s).

Adding nodes to an edge site can be done as follows:

Example

```
[root@smcluster ~]# cmsh
[smcluster]% edgesite
[smcluster->edgesite]% use ams-west
[smcluster->edgesite->ams-west]% append nodes edge-node005 edge-node006
[smcluster->edgesite->ams-west]% commit
[smcluster->edgesite->ams-west]% list
Name (key)   Director            Nodes
------------ ------------------- ------------------------------------------------------------
ams-west     ams-west-director   ams-west-node001,ams-west-director,edge-node005,edge-node006
dell-edge    dell-edge-director  dell-edge-director
```

2.1.3 Viewing Edge Sites Using `cmsh`

Edge sites can be viewed from the `edgesite` mode of `cmsh`

Example

```
[root@smcluster ~]# cmsh
[smcluster]% edgesite
[smcluster->edgesite]% list
Name (key)               Director                 Nodes
```

```
----------------------- ----------------------- ----------------------------------------
ams-west                ams-west-director       ams-west-node001,ams-west-director
dell-edge               dell-edge-director      dell-edge-director
[smcluster->edgesite]%

[smcluster->edgesite]% show ams-west
Parameter                        Value
------------------------------- ----------------------------------------------
Address                         Kings
Administrator e-mail            admin@bright
City                            Amsterdam
Contact                         admin
Country                         Amsterdam
Director                        ams-west-director
Name                            ams-west
Nodes                           ams-west-node001,ams-west-director
Notes
Revision
Secret                          *********
```

2.1.4 Viewing Edge Sites Using Bright View

Edge sites can also be viewed via the clickpath `Datacenter Infrastructure`→`Edge Sites` (figure 2.16). Properties of an edge site can be managed via editing a particular edge site.

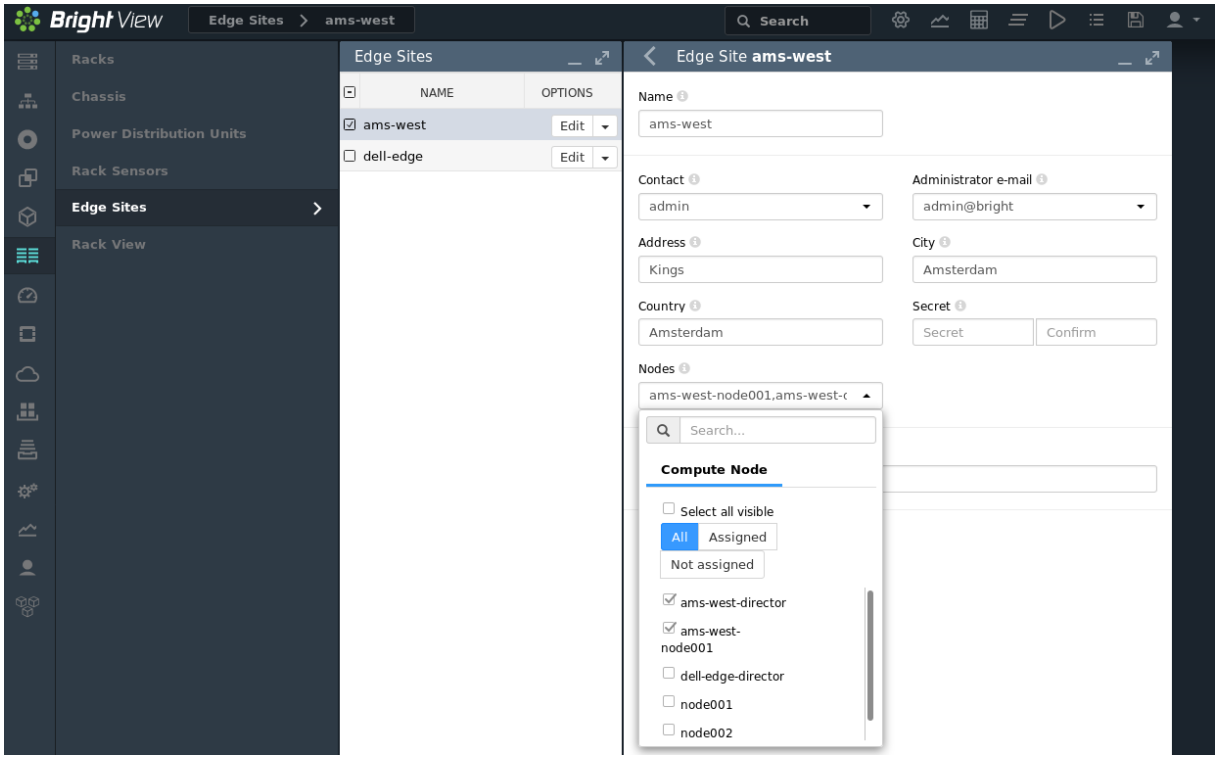

Figure 2.16: Edge sites in Bright View

2.1.5 Create Edge ISO

The next step in the deployment is to create the edge ISO on the head node. Typically, the edge ISO is configured so that the edge director boots from it the first time, and carries out a FULL install using the

ISO for its source of files that will be installed on the edge director. The edge director is also configured to allow a boot from the hard drive.

If booting from the ISO after the first time, and if the partitions on the edge director have not changed, then a SYNC install is carried out against the central head node. If booting after the first time, and if there is no ISO, then the edge director simply boots from its local hard drive, and no files are synced with the central head node

There are two ways to create the edge ISO:

1. The edge ISO can be created with a site-specific auto-generated wrapper script. This is the recommended approach. When an edge site is created, CMDaemon on the head node creates a wrapper script at `/var/spool/cmd/edge/create-<site-name>-iso.sh`. The wrapper script then provides all the site-specific information that needs to be provided for the edge node-installer. The behavior of the wrapper script depends on whether the setup was done with edge director HA enabled.

   - The wrapper script that is generated when edge director HA configuration is enabled, can be used as follows:
     - The following command produces an ISO that can be used for booting the primary edge director:
       ```
       /var/spool/cmd/edge/create-dell-edge-iso.sh --primary
       ```
     - The following command produces an ISO that can be used for booting the secondary edge director:
       ```
       /var/spool/cmd/edge/create-dell-edge-iso.sh --secondary
       ```
     - The following command produces two ISOs, one for the primary and one for the secondary:
       ```
       /var/spool/cmd/edge/create-dell-edge-iso.sh --both
       ```
     - The following command produces an ISO that can be used to boot both the primary and secondary edge directors, but the correct option for the edge director IP must be manually chosen in the node-installer screen:
       ```
       /var/spool/cmd/edge/create-dell-edge-iso.sh
       ```
   - The wrapper script that is generated in the case when no edge director HA configuration is selected, takes no additional arguments. It produces an ISO that is used to boot the single edge director, and is invoked as:
     ```
     /var/spool/cmd/edge/create-dell-edge-iso.sh
     ```

2. Alternatively, the edge site ISO can be created manually by setting the options to `create-edge-iso`:

   Example

   ```
   [root@headnode ~]# create-edge-iso --help
   usage: create-edge-iso [-h] [-v] [-d] [-c] [-f EDGEINTERFACE]
                          [-e EDGEDIRECTORIP] [-m HEADNODEIP] [-g DEFAULTGATEWAY]
                          [-k KERNELIMAGE] [-i IMAGENAME] [-s] [-p PATHTOISOFILE]
                          [-n]

   Create edge iso

   optional arguments:
     -h, --help            show this help message and exit
   ```

```
-v, --verbose         Turn on verbose logging
-d, --debug           Turn on debug mode, iso work directory will not be cleaned up
-c, --includecmshared
                      Include /cm/shared on iso
-f EDGEINTERFACE, --edgeinterface EDGEINTERFACE
                      Name of interface on edge node
-e EDGEDIRECTORIP, --edgedirectorip EDGEDIRECTORIP
                      IP[/Netmask bits] of edge director
                      If Netmask bits is not specified, defaults to /16
-m HEADNODEIP, --headnodeip HEADNODEIP
                      IP[:port] of head node
                      If port is not specified, defaults to :8081
-g DEFAULTGATEWAY, --defaultgateway DEFAULTGATEWAY
                      Gateway for edge director to reach central head node
-k KERNELIMAGE, --kernelimage KERNELIMAGE
                      Name of image whose kernel will be used for booting iso
-i IMAGENAME, --imagename IMAGENAME
                      Name of software image to include on iso
-s, --sitesecret      Prompt user to enter Edge site secret
-p PATHTOISOFILE, --pathtoisofile PATHTOISOFILE
                      Path to iso file name
-n, --donotstoresecret
                      Inform node-installer not to store the secret on the edge director
```

2.1.6 Edge ISO Node Installer

The edge ISO is used to provision the edge director. The node installer displays the following screens when booting from the edge ISO:

Figure 2.17: Edge node-installer ISO boot menu

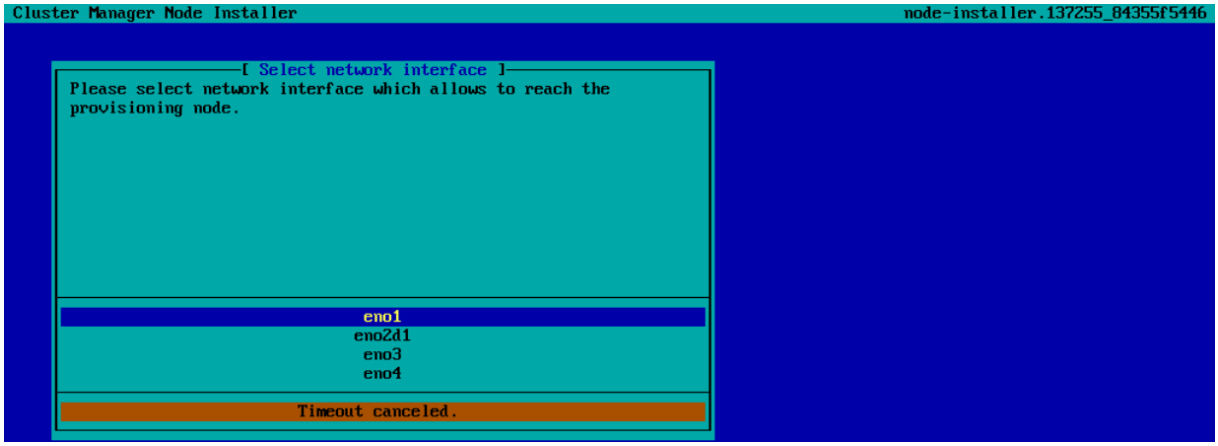

Figure 2.18: Edge node-installer select interface

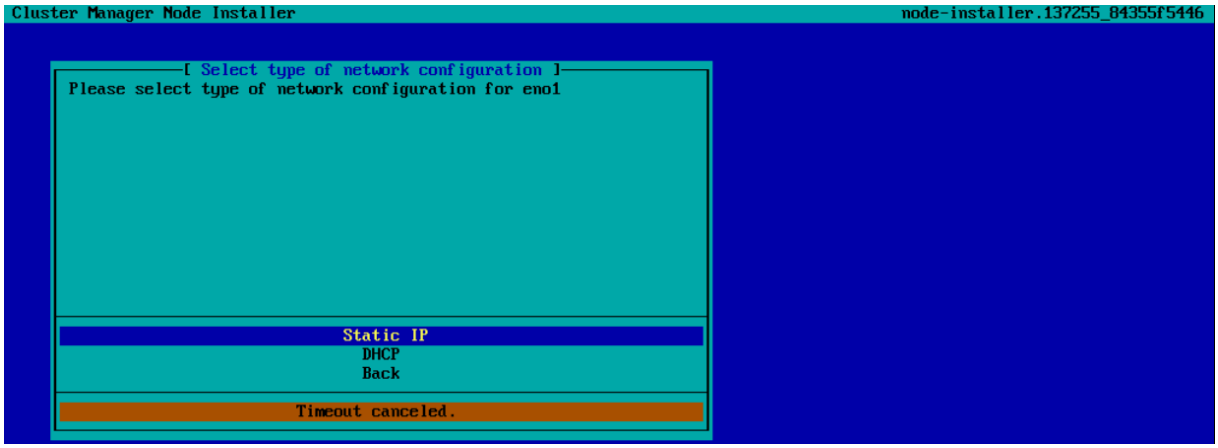

Figure 2.19: Edge director IP Static/DHCP selection

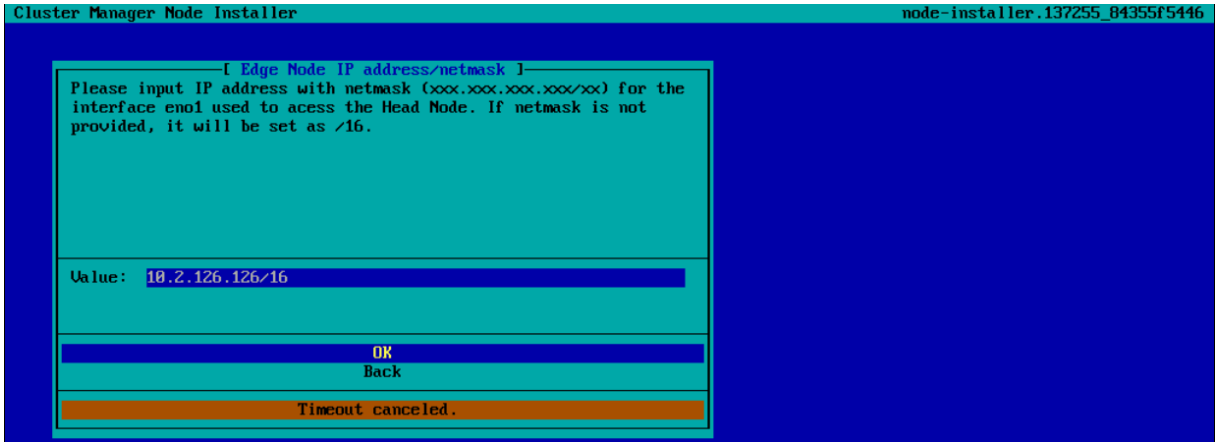

Figure 2.20: Edge director IP address and netmask

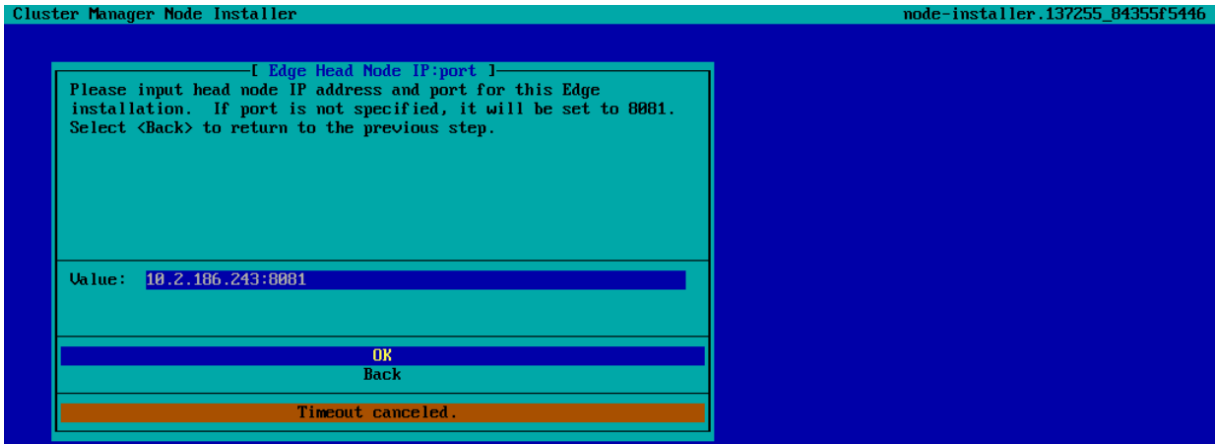

Figure 2.21: Central head node IP address and port

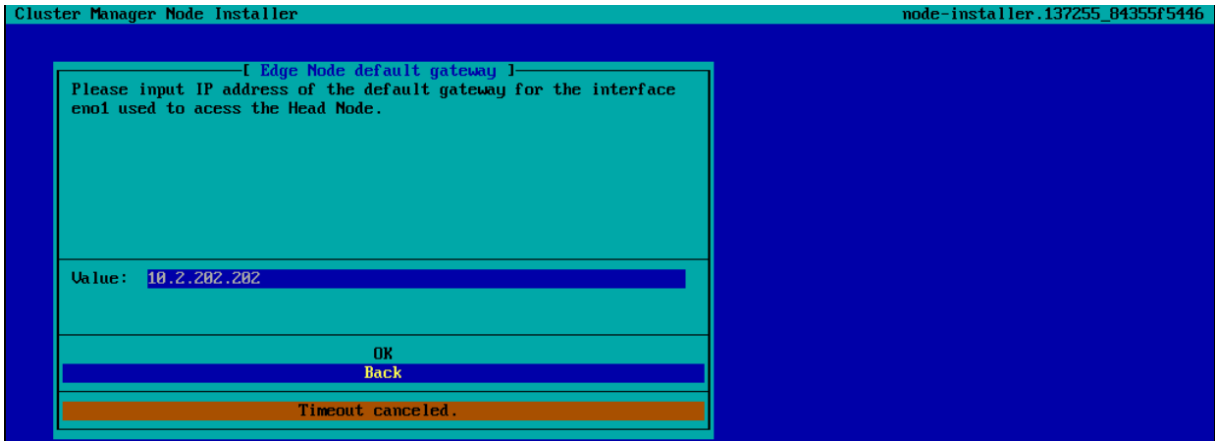

Figure 2.22: Gateway IP address to central head node

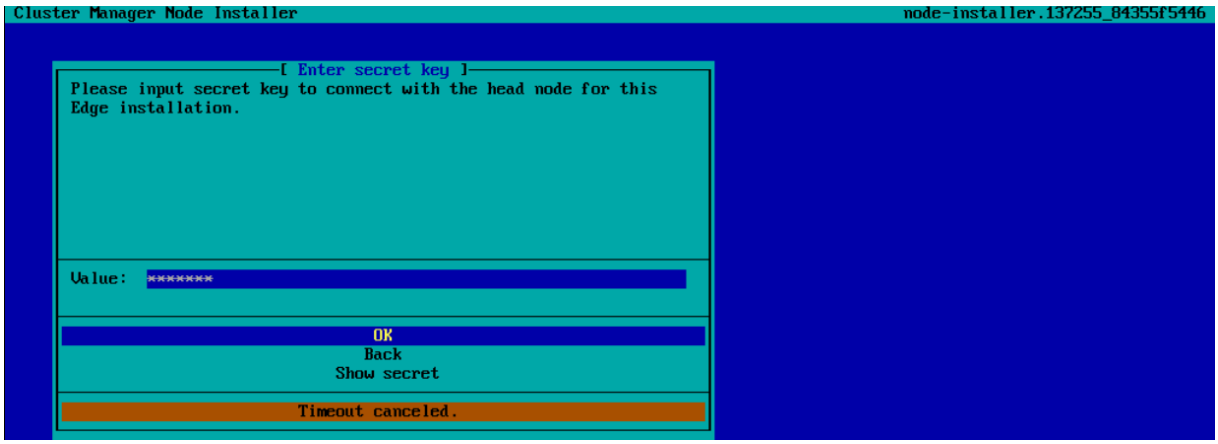

Figure 2.23: Edge site secret

2.1.7 Edge Directors

Edge directors can be provisioned from the head node, but are normally provisioned using the software image on the edge ISO/USB. This means:

- The ISO/USB should have a software image included in it
- The ISO/USB should have `/cm/shared` included in it

If the edge director is booting from the ISO/USB, it means that:

- There is a minimal overhead when only updates, rather than an entire filesystem, are synced from the head node to the edge director
- A `FULL` install of the edge director only takes place during the first installation of that director, or if the director disk partitions have changed.
- If the edge director has already been installed previously, and its disk partitions are unchanged, then a `SYNC` install is carried out, so that local files on the edge director can get updated against the head node

If there is no ISO/USB available to the edge director, then the director simply boots off its local drive, and no SYNC install is followed. An explicit `imageupdate` can however be carried out afterwards when needed, if connectivity is there, to update the software image.

Once the edge director is in the `UP` state, it is responsible for the following local operations:

- Ramdisk creation for the edge nodes
- Power control for itself and the edge nodes
- Device state (`UP`, `DOWN`) check via ICMP ping to the edge nodes
- Monitoring for the edge nodes

2.1.8 Edge Nodes

Edge nodes are expected to PXE boot off the edge internal network. The edge director provisions edge nodes in the same way that the head node provisions regular nodes.

3

Installing Slurm To An Edge Site

On some edge sites there may be a need to run a workload manager.

The Slurm workload manager can be run on an edge site if the cluster is prepared and software installed as in sections 3.1 and 3.2.

3.1 Preparation

1. The edge director must be `UP` according to `cmsh` or Bright View.
2. No MySQL or MariaDB installation must be present on the edge director or it will conflict with the automatic installation of `cm-mariadb`.

3.2 Installation

- The cluster manager script `cm-wlm-setup` (Section 7.3 of the *Administrator Manual*) is then run on the head node.
  - At the `Select installation type` screen, the edge name should be selected, and the edge director should then be specified as only server role node.
- By default there are no user home directories on the edge director or edge nodes. These must therefore be mounted or created, otherwise jobs cannot run on the edge nodes. This is true for all WLMs running on edge nodes.