

# Single Server System Dimensioning Guide, CEE 6

Cloud Execution Environment

CONFIGURATION MODEL

**Copyright**

© Ericsson AB 2016–2018. All rights reserved. No part of this document may be reproduced in any form without the written permission of the copyright owner.

**Disclaimer**

The contents of this document are subject to revision without notice due to continued progress in methodology, design and manufacturing. Ericsson shall have no liability for any error or damage of any kind resulting from the use of this document.

**Trademark List**

All trademarks mentioned herein are the property of their respective owners. These are shown in the document Trademark Information.



# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	Target Group	2
1.2	System Characteristics	2
<b>2</b>	<b>CEE System</b>	<b>2</b>
<b>3</b>	<b>Hardware Requirements</b>	<b>2</b>
3.1	Server Test Configuration	3
3.2	Network Configuration	4
3.3	CPU Configuration	4
3.3.1	Host OS Cores	5
3.4	RAM Configuration	5
3.4.1	Introduction	6
3.4.2	Configuration	7
3.5	Storage Configuration	8
3.5.1	Local Storage Disk Space	8
3.5.2	Disk Requirements for Atlas	10
3.5.3	Disk Requirements for Nova Snapshots	10
<b>4</b>	<b>Characteristics</b>	<b>10</b>
4.1	General System Limits	11
4.2	Orchestration Interface	11
4.3	Tenant Execution Environment	11
4.3.1	Performance	11
4.3.2	Resiliency	11
4.4	Network	12
4.4.1	Performance	12
4.4.2	Resiliency	12
4.4.3	Tenant Network Limitations	12
4.5	Storage Limitations	12
4.6	In-Service Performance	13
	<b>Reference List</b>	<b>14</b>





# 1 Introduction

This document describes the characteristics of Cloud Execution Environment (CEE) to enable dimensioning and understanding the limitations of CEE. It also describes generic requirements on hardware used for running CEE. The application can have additional requirements.

The document provides general system dimensioning guidelines and does not describe a specific hardware type or model. Dedicated studies to create optimizations for a specific hardware model are outside of the scope of this document.

Storage is measured in gibibyte (GiB), tebibyte (TiB), and mebibyte (MiB) in this document.

1 GiB is equivalent to 1.074 GB.

The following words are used in this document with the meaning specified below:

<b>CSS</b>	The Cloud SDN Switch (CSS) is the virtual switch (vSwitch) component of CEE. It is based on the opensource project openvswitch (OVS) with functional extensions and performance enhancements. For more information, refer to the CEE Architecture Description, Reference [1], and the CSS documentation.
<b>vNIC</b>	A virtual network interface card (vNIC) provides connectivity between CSS and a Virtual Machine (VM). A configuration can provide several vNICs to a VM.
<b>Interface</b>	A network interface. Can be either a physical NIC (PHY) providing CSS with board external connectivity, or a virtual NIC connecting CSS to a VM.
<b>PMD thread</b>	CSS uses a Poll Mode Driver (PMD) technique that continuously polls incoming packets from the NICs, that is, interrupts are not used. To be able to reliably handle all incoming packets, the NIC queues are continuously polled for packets to be handled. This software is executing in one or more threads that are called PMD threads. The execution environment for the PMD threads resides in the Linux user space and is thus isolated from the Linux scheduler to be able to reliably handle a high sustained packet flow without interrupts or delays caused by being scheduled out.
<b>NUMA</b>	Non-Uniform Memory Architecture
<b>Core</b>	A physical core of a processor.



## 1.1 Target Group

Cloud Infrastructure providers and application designers.

## 1.2 System Characteristics

For information on the features of CEE, refer to the [CEE Technical Description](#).

For the characteristics of the used hardware, refer to the product documentation of the hardware used.

# 2 CEE System

This section summarizes the Single Server CEE system.

- The Single Server CEE operates on one Dell server.
- The used switches, data center gateways (DC-GWs), and firewalls are not part of the CEE product, these are the user's responsibility.
- External storage is not available.
- vFuel is used from an installation laptop that is connected to the system for the time of CEE installation, update, and other maintenance activities. When the activity has been completed, vFuel is switched off by using vFuel On Demand Use.

# 3 Hardware Requirements

CEE is a software product which can run on hardware infrastructures that comply with the generic requirements detailed in this section. Different hardware can imply different CEE system characteristics.

The CEE compute host connected to the customer provided environment is shown in Figure 1.

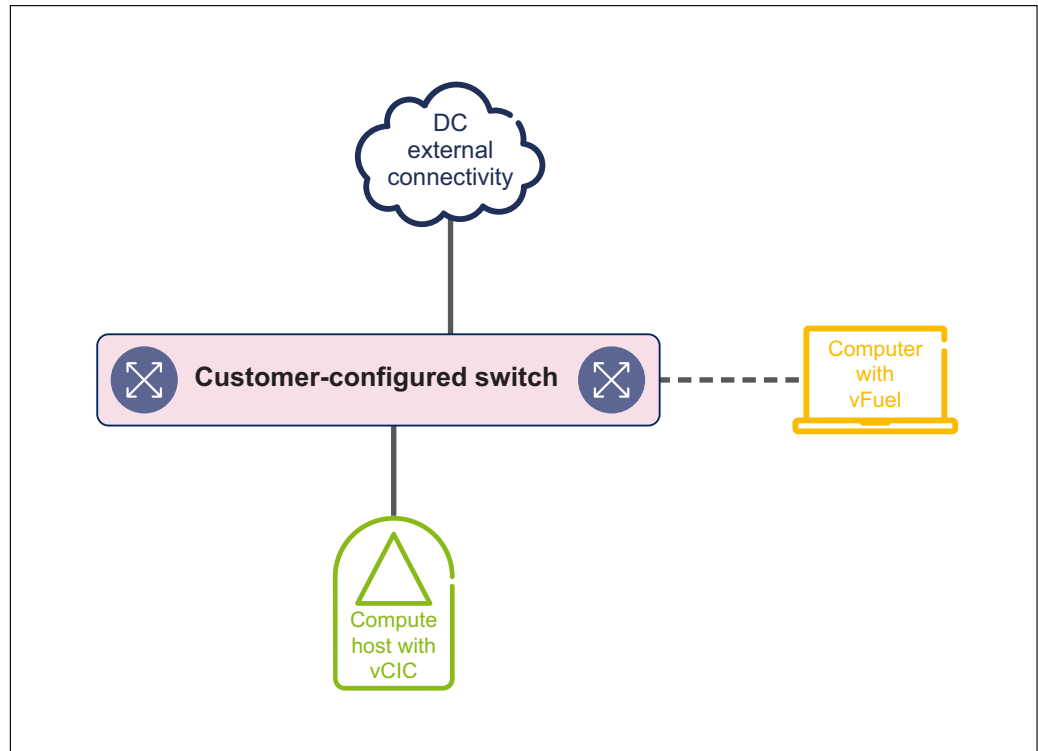


Figure 1 CEE Hardware Environment

### 3.1 Server Test Configuration

This section describes the hardware configuration used to test the single server configuration of CEE.

Table 1 Hardware Configuration Example, Server

Aspect	Requirement
CPU	2x Intel® Xeon® Processor E5-2680 v3 (12 cores per processor, 2 HyperThreads (HT) per core), 48 HTs are available
RAM	128 GiB
NIC	1×10 GE Intel 2P X520 and 1×1 GE unspecified
Onboard Disk	At least 1000 GiB SSD: <ul style="list-style-type: none"> <li>• 2×500 GiB for RAID 0</li> <li>• 2×1000 GiB for RAID 1</li> </ul>
Management Interface	Dell iDRAC



## 3.2 Network Configuration

This section describes hardware configuration for networking.

The physical host network contains at least two interfaces:

- 1×1 GE NIC for the control network
- At least 1×10 GE DPDK compatible NIC for the tenant traffic network

The server management interface must be compatible with Dell iDRAC or HPE iLO.

## 3.3 CPU Configuration

This section provides guidelines on CPU configuration, as the available hardware resources, physical cores, is a crucial parameter. CPUs must be of type Intel Xeon E5 v3 or later, configured with two NUMA nodes.

At CEE installation, the cores are allocated to several owners. Refer to the [Configuration File Guide](#) for more information about the configuration procedure.

Since the system has limited CPU resources, it is necessary to assign the CPUs manually to the different resource owners in order to achieve optimal performance. The cores available for VMs are divided in two pools:

- One pool where VMs are floating and CPUs are over-committed
- And one pool with dedicated CPUs

These pools must be dimensioned according to the application needs.

The number of available CPU IDs depends on the CPU model. For example, the CPU allocation recommended for EPC is shown in Table 2 and Figure 2. This allocation is valid for Dell R630 with Intel Xeon E5-2680 v3 processor.

Table 2 CPU Allocation for Dell R630 with Intel Xeon E5-2680 v3 Processor

CPU Owner	Allocated CPU ID
Tenant VM	6,30,8,32,10,34,7,31,9,33,11,35,13,37,15,39,17,41,19,43,21,45
CSS <sup>(1)</sup>	2,26,3,27
CSS control process <sup>(2)</sup>	1 <sup>(2)</sup>





CPU Owner	Allocated CPU ID
vCIC <sup>(3)</sup>	Uses non-isolated CPUs, see host OS.
Host OS	0,24,1,25,4,28,5,29,12,36,14,38,16,40,18,42,20,44,23,47

(1) In this example, CSS is configured with `normal - perf` mode. Refer to the [Configuration File Guide](#) for more details.

(2) The process does not get a CPU for its exclusive use. A configuration parameter specifies one of the host operating system (OS) CPUs to be used by the CSS control process. The CPU must be selected from a NUMA node to which a physical interface for the tenant network is attached.

(3) The vCIC can be allocated on cores that are shared with the application. In such cases, it must be ensured that the application sharing resources with the vCIC does not exhaust the vCIC resources. Refer to the [Configuration File Guide](#) to configure accordingly.

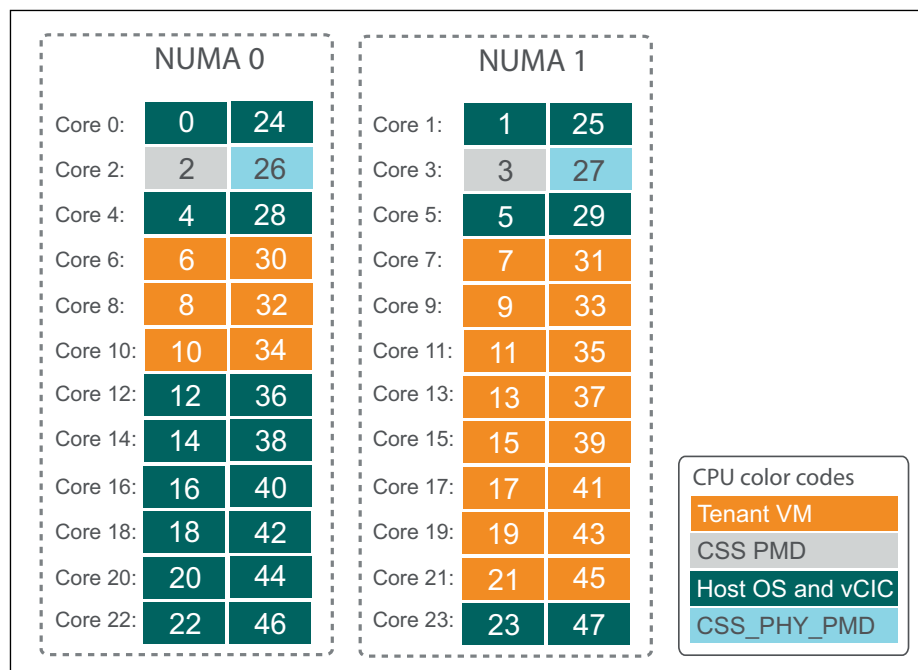


Figure 2 CPU Allocation of the Respective Resource Owner per NUMA Node on Dell R630 with Intel Xeon E5-2680 v3 Processor

### 3.3.1 Host OS Cores

The host OS, monitoring processes and OpenStack agents use the CPUs not reserved for the tenant VMs. In certain single server configurations the vCIC and tenant VMs use non-reserved CPUs, and they share the same CPUs used by the host OS.

## 3.4 RAM Configuration

**Note:** At least 128 GiB of nominal memory is required on the physical server.



This section describes the optimal RAM configuration for the Single Server CEE.

Refer to the [Configuration File Guide](#) for more information about the configuration procedure.

The subsections contain the following information:

- Section 3.4.1 on page 6 provides general information about RAM configuration in CEE.
- Section 3.4.2 on page 7 describes the memory sizes for RAM allocation.

### 3.4.1 Introduction

A certain amount of memory is reserved for booting the server, for example, BIOS-reserved memory, page tables, and memory-mapped devices. This reserved memory is called *unmanaged*. The total memory the system can use is called *managed*, and it is the difference between the nominal physical memory and the unmanaged part. The sum of managed and unmanaged parts is relevant for planning purposes, for example, to plan the amount of physical memory to be installed on the physical server. The unmanaged part cannot be changed. The managed memory allocated to the host OS on the compute node is 6 GiB by default. More managed memory can be reserved for the host OS by setting the relevant configuration parameter. The needed amount of memory depends on the values of a number of other parameters as described below.

The memory used for the VMs is allocated to hugepages. This is the memory visible from the inside of the VMs. The 1 GiB hugepages are referred to as **Tenant VM** in the RAM reservation tables of Section 3.4.2 on page 7. In addition to the 1 GiB hugepages, the VMs need memory allocated from the host OS. This memory is used, for example, to emulate devices used by the virtual machine. It is hard to predict the amount of host OS memory used by the emulator since, for example, it depends on the type and number of the used devices. A small VM consumes less than 100 MiB, while it can grow to several hundred MiB in specific cases. About 300 MiB host OS memory would be enough for each VM but we must double it and calculate with 600 MiB as explained below.

In a system using the NUMA architecture, the NUMA location of VMs must be considered. The available memory, that is, the hugepages and the Host OS memory, are evenly distributed between the NUMA nodes. By design, OpenStack Nova allocates VMs on the first NUMA node that fits the VM. Apart from the VMs running on both NUMA nodes, the VMs allocate memory from the NUMA node on which they are running. In a worst case scenario where all VMs are allocated on the same NUMA node, all the memory for the VMs will be allocated from the same NUMA node. In such a scenario most of the memory on the other NUMA node will be unused, and half of the memory on the compute node will be free. To be on the safe side in a dual socket system, the 300 MiB host OS memory per VM must be doubled to cover the case where all VMs are allocated on the same NUMA node.

The processes running on the compute host use 4 KiB memory pages. In addition, CSS also uses 2 MiB hugepages, and QEMU processes of each tenant VM deployed



with hugepages use additional 1 GiB pages. Hugepages for tenant VMs are always reserved symmetrically on each NUMA node, if the required number of hugepages is even. If the required number is odd, one more page is reserved on NUMA node 0.

For details regarding memory allocation in `config.yaml`, refer to the [Configuration File Guide](#).

**Note:** In order to allocate as little memory for the host OS as possible, memory profiling of the host OS for the specific scenario is recommended.

### 3.4.2 Configuration

Table 3 specifies the RAM required for the resource owners.

Table 3 Memory Allocation for System with 128 GiB RAM

RAM Resource Owner	Hugepage size (MiB)	Number of Hugepages (Count)	Total Size (GiB)
CSS	2	1024	2
vCIC	1024	16	16
Unmanaged memory <sup>(1)</sup>			3
Host OS <sup>(2)</sup>			X (Integer) <sup>(2)</sup>
Tenant VM	1024	The same number as the remaining amount of memory in GiB <sup>(3)</sup>	The remaining amount of memory

(1) Refer to the [Multi-Server System Dimensioning Guide](#), CEE 6 to dimension the system in case the nominal physical memory is higher than 128 GiB.

(2) For more information, see Section 3.4.2.1 on page 7.

(3) The remaining amount of memory (GiB) divided by the 1 GiB hugepage size

Each Neutron network created consumes RAM in the vCIC, and it influences the maximum number of virtual tenant networks. See Section 4.4.3 on page 12 for more information.

The default vCIC swap size is 512 MiB. A single server vCIC deployment requires a swap space of 5120 MiB for the vCIC. The swap space can be changed by setting the `vcic_swap_size` optional parameter. For configuration information, refer to the [Configuration File Guide](#).

#### 3.4.2.1 RAM Allocation for Host OS

The variable X corresponding to the minimum size of RAM in GiBs allocated to the Host OS can be calculated using the following formula:



**Note:** The formula described in this section is used to calculate the minimum amount of RAM required for the host OS. Depending on the configuration, more memory can be reserved for the host OS.

$X = 6 + n * 0.6$  rounded up to the next **odd** integer,  
where  $n$  is the maximum number of VMs planned on the compute host.

For example, if the compute host is to host 10 VMs,  $6 + 10 * 0.6 = 12$  is rounded up to the next odd integer, therefore  $X = 13$

## 3.5 Storage Configuration

**Note:** The physical server must have an SSD or HDD with a minimum capacity of 300 GiB.

This section describes the local storage implementations, and disk requirements.

Refer to the [Configuration File Guide](#) for more information about the configuration procedure of CEE.

**Note:** Distributed storage is not available for single server deployments of CEE.

### 3.5.1 Local Storage Disk Space

This section lists requirements on disk space.

Table 4 shows the dimensioning of the various partitions of the tested configuration of the Single Server CEE.

Table 4 Local Storage Disk Space

Use	Size	Partition	Note
Root partition of the vCIC	50 GiB	/	By default, the swap file is located on the root partition, and consumes disk from the allocated space for the root partition.
Logs and core/crash dumps of the vCIC	40 GiB <ul style="list-style-type: none"><li>• 10 GiB for logs</li><li>• 30 GiB for core and crash dumps</li></ul>	/var/log	If the size of this area is increased, it increases the storage area for core and crash dumps. The 10 GiB for logs is a constant value.



Use	Size	Partition	Note
Database for OpenStack and Zabbix (MySQL) on the vCIC	40 GiB	/var/lib/mysql	
Glance repository in Swift on vCIC	40 GiB	/var/lib/glance	<p>The size might need to be adjusted depending on the amount and size of images stored in Glance.</p> <p>It includes temporary storage for SFTP/SCP transfers.</p> <p>In Single Server CEE the images are replicated in the same vCIC. Consequently, allocating 40 GiB for Glance allows a maximum of 20 GiB of images to be stored.</p>
Root partition of the compute host	50 GiB	/	
Logs and core/crash dumps of the compute host	40 GiB <ul style="list-style-type: none"> <li>• 10 GiB for logs</li> <li>• 30 GiB for core and crash dumps</li> </ul>	/var/log	If the size of this area is increased, it increases the storage area for core and crash dumps. The 10 GiB for logs is a constant value.
<b>Sum for the vCIC</b>	170 GiB		The sum of disk space used by the vCIC
<b>Sum for the compute host</b>	90 GiB		The sum of disk space used by the compute host
<b>Sum for the vCIC and the compute host</b>	260 GiB		The sum of disk space used by the vCIC and the compute host

The remaining disk space is used as ephemeral storage for the VMs. The size of the ephemeral storage can be calculated by removing the storage area used by the vCIC and the compute host from the total disk space.



### 3.5.2 Disk Requirements for Atlas

Atlas uses a fixed disk size value of 10GiB and a configurable ephemeral storage size with default value of 120GiB. When images are loaded using Atlas from Images or Catalog Panel (as part of an .ova file), the image is temporarily stored in ephemeral storage in Atlas. To support loading of large images, the recommendation is to use 120 GiB for the Atlas ephemeral storage.

In case the local disk is used as ephemeral storage (no centralized storage or distributed storage), the Atlas VM occupies 130 GiB (10 GiB disk + 120 GiB ephemeral) of the local disk on the compute host where it is running.

To reduce the ephemeral disk allocated to Atlas, the size of the ephemeral disk can be reduced from 120 GiB to a minimum of 10 GiB.

**Note:** 30% of the ephemeral disk in Atlas is used as a temporary storage for images or .ova files. Consequently, the size of the ephemeral disk needs to be adjusted according to the size of .ova files to be loaded. Using a reduced ephemeral disk size of 10 GiB implies that it can be impossible to load .ova files that contain images larger than 3 GiB.

### 3.5.3 Disk Requirements for Nova Snapshots

Nova snapshots are stored in the `/var/lib/glance` partition of the CIC node.

There are certain disk requirements for the Nova snapshots to work. Depending on the requirements and frequency on Nova snapshots, the system must be dimensioned with free disk space, according to the following guidelines:

- For a successful Nova snapshot, the amount of free space in disk partition `/var/lib/nova` in the compute host must be **at least** double of the snapshot/VM size, since the snapshot is first extracted locally in the compute node before it is uploaded to the Glance/Swift store.
- The disk space needed in the `/var/lib/nova` partition of the compute disk must have free space **at least** twice the size of VMs root disk. The reason is that during the extraction of the snapshot, first the delta of the VM disk will be extracted, after which the complete disk will be extracted.
- Disk partition `/var/lib/glance` in the CIC node must have free space **at least** twice of the root disk size of the VM, in order to accommodate the snapshot.

## 4 Characteristics

This section describes the system characteristics of CEE.



## 4.1 General System Limits

For the list of system limits, see Table 5.

Table 5 General System Limits

Slogan	Limit
Number of physical servers	One physical server is used.
Number of cores occupied by infrastructure	See Table 2 for information about CPU allocation.

## 4.2 Orchestration Interface

The system limits for orchestration are listed in Table 6.

Table 6 Orchestration Limits

Slogan	Limits
Number of tenants	The maximum number of supported tenants is 50.

## 4.3 Tenant Execution Environment

This section describes the tenant-related limits on the environment.

### 4.3.1 Performance

Performance limits are listed in Table 7.

Table 7 Tenant Execution Performance

Slogan	Limits
Oversubscription	CPU overcommit: supported. Memory overcommit: not supported Disk overcommit: not supported.

### 4.3.2 Resiliency

The Single Server CEE does not provide resiliency for tenant execution, due to the reduced hardware resources.



## 4.4 Network

### 4.4.1 Performance

For networking performance characteristics of CEE, refer to the [Multi-Server System Dimensioning Guide, CEE 6](#).

### 4.4.2 Resiliency

Resiliency is not provided by the Single Server CEE, due to the reduced hardware resources.

### 4.4.3 Tenant Network Limitations

Limitations of the tenant network are listed in Table 8.

Table 8 Tenant Network Limitations

Slogan	Limits
Number of virtual networks	The theoretical aggregated maximum number of virtual tenant networks per CEE region is 4050 for segmentation type vlan. Since each Neutron network created consumes RAM in the vCIC, this theoretical maximum cannot be reached. The default configuration of RAM for vCIC allows 1000 networks. Additional memory is needed if more Neutron Networks are created.
Number of vNICs per guest VM	The maximum number of vNICs per guest VM is 10 (+ 1 Trunk vNIC).
Number of Trunk vNIC attached vLANs	The number of Trunk vNIC attached vLANs is limited to 100.
Number of vNICs per physical server	CSS supports up to 64 vNICs per NUMA node with default RAM allocation of 1 GiB for CSS per NUMA node.
MTU	L2 Packet MTU size is 2140 bytes.

## 4.5 Storage Limitations

This section describes CEE characteristics on storage.

Only local storage is supported on Single Server CEE.

For tenants, ephemeral storage (non-persistent block storage) is supported on local disks of the compute hosts.





There is no support for distributed local storage or for any shared file system in CEE.

Swift uses the Local Storage.

Object storage through Swift is only used for CEE infrastructure.

Management of VM images is supported by the OpenStack image service.

For the Single Server CEE, only “boot from image” is supported.

If data is stored on a local disk, it is erased in case of disk failure or rollback from a failed update, meaning that the VM disappears. The application must be designed accordingly.

## 4.6 In-Service Performance

This section lists the characteristics on in-service performance.

Table 9 In-Service Performance

Slogan	Characteristics
Guest execution retainability	Guest execution is not interrupted at a virtual infrastructure management cluster restart or. At CEE software update the compute nodes are restarted sequentially and this causes VM evacuations or VM restart.
Update availability	When the CEE software update is running, the OpenStack API service is migrated to each vCIC sequentially, therefore the service is unavailable for about one minute during the migration.
Restart availability	The OpenStack API service is not available during the restart of the Virtual Infrastructure Manager (VIM) cluster.



## Reference List

- [1] CEE Architecture Description, 5/155 53-AZE 102 01, available at Ericsson Support
- [2] Limitations and Workarounds for Cloud Execution Environment (CEE) <release>, 5/109 21-AZE 102 01/5-n, where "n" is different for each release