

Scaling Management

Ericsson Service-Aware Policy Controller

Description

Copyright

© Ericsson AB 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

1	Scaling Management Introduction	1
2	Scaling Management Functions and Concepts	2
2.1	Types of Operation	2
3	Managed Object Model	4





1 Scaling Management Introduction

This document provides an overview of the management model and concepts associated with the Scaling Management managed area.

A managed area is represented by a group of Managed Object Classes (MOCs) within the Managed Object Model (MOM).



2 Scaling Management Functions and Concepts

Scaling Management provides a management interface to configure the following on the Managed Element (ME):

- Scale-out, that is, adding one or more Virtual Machines (VMs) when a cluster requires more processing capacity.
- Graceful scale-in, that is, removing one or more VMs when a cluster needs less processing capacity. Released resources can be allocated again when more processing capacity is required.

A scale-out operation is performed by adding one or more VMs to the Virtualized Network Function (VNF) cluster. Preboot Execution Environment (PXE) boot is used to distribute software to new VMs. For graceful scale-in, the VNF cluster reallocates the resources from VMs to be scaled-in and moves to other VMs to prevent data loss. Performance counters can be used as input to decide which scaling operation is to be performed.

A VM can have one of the following two roles:

- System

When a VM works in role `System`, it cannot be scaled in or scaled out, as it has a system-defined size and a system-defined role. An example is that two VMs are allocated to the Operation and Maintenance (O&M) role within the VNF, which is named SC-1 and SC-2. These VMs cannot be scaled in or scaled out.

- Default-Role

When a VM works in `Default-Role`, it does not have a specialized role like, for example, SC-1 or SC-2. This means that VMs with `Default-Role` can be scaled in or scaled out, depending on the capacity needs of the VNF.

2.1 Types of Operation

Scaling Management supports the following operations:

- Scale-out

The scale-out operation adds a node to the cluster. Further details on how to perform this operation depend on the deployment type:

- Deployment in a Cloud Data Center. See the procedure in [Configure Scale-Out](#).



- Deployment in specific HW. See procedure in [SAPC PNF Scale Out](#).
- Graceful scale-in

The graceful scale-in operation, intended for Cloud scenarios, removes a node from the cluster. The procedure in [Configure Graceful Scale-In](#) provides further details on how to perform this operation.



3 Managed Object Model

The Scaling Management managed area is represented in the Managed Object Model (MOM) as follows:

```
ManagedElement
+-SystemFunction
+-SysM
+-CrM
+-ComputeResourceRole
+-Role
```

For general information about the MOM, MOCs, Managed Objects (MOs), cardinality, and related concepts, refer to [Managed Object Model User Guide](#).

The Scaling Management MOCs are described in [Table 1](#).

Table 1 Scaling Management Managed Object Class Descriptions

Managed Object Class	Description
CrM	The root of the Scaling Management model.
ComputeResourceRole	Describes the service (defined by the role) provided by the compute resource.
Role	Describes a role that determines the type of service provided by a compute resource.