

Bandwidth Overallocated due to Race Condition

Cloud Execution Environment

OPERATING INSTRUCTIONS

Copyright

© Ericsson AB 2016. 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	Introduction	1
1.1	Alarm Description	1
1.2	Prerequisites	2
2	Procedure	2
2.1	Actions	3



Bandwidth Overallocated due to Race Condition



1 Introduction

This instruction concerns alarm handling.

1.1 Alarm Description

The `Bandwidth Overallocated due to Race Condition` alarm is a primary alarm.

The alarm is issued by the Managed Object (MO) `Node` when the periodic algorithm detects that the bandwidth requirement for the virtual machines (VMs) running on the node exceeds the available bandwidth. For more information, refer to the section on bandwidth based scheduling in *OpenStack Compute API in CEE*.

The severity of the alarm is `MINOR` or `CLEARED`.

The possible alarm causes and fault locations are explained in Table 1.

Table 1 Alarm Causes

Alarm Cause	Description	Fault Reason	Fault Location	Impact
Bandwidth overallocation	The allocated bandwidth exceeds the host capabilities	More VMs were booted on the compute than it was allowed by the bandwidth capabilities of the host	Compute node	The required total bandwidth for the VMs is not available on the compute, which can lead to performance degradation

The following is the consequence for the node if the alarm is not solved:

- The virtual machines running on the affected compute might not have the required network bandwidth, which can lead to a network performance degradation.

The alarm attributes are listed in Table 2.

Table 2 Alarm Attributes

Attribute Name	Attribute Value
Major Type	193
Minor Type	2031718



Attribute Name	Attribute Value
Managed Object Class	Node
Managed Object Instance	Region=<name_of_the_region>, CeeFunction=1, Node=<hostname_of_the_node>
Specific Problem	Bandwidth overallocated due to race condition
Event Type	other (1)
Probable Cause	systemResourcesOverload (207)
Additional Text	;uuid=<HW_UUID_of_failed_server> ⁽¹⁾
Severity	MINOR (5) or CLEARED

(1) The format of this field is expected to change in CEE R6.

Note: The alarm does not specify which VMs are affected.

1.2 Prerequisites

This section provides information on the documents, tools, and conditions that apply to the procedure.

1.2.1 Documents

Not applicable.

1.2.2 Tools

No tools are required.

1.2.3 Conditions

No conditions.

2 Procedure

This section describes the procedure to follow when this alarm is received.



2.1 Actions

Perform the following:

1. Check which VMs are running on the affected compute by issuing the following command on a controller:

```
nova list --host=<affected_compute>
```

2. Check the bandwidth need of the affected VMs on the compute, by issuing the below command on a controller:

```
/etc/zabbix/scripts/bandwidth_allocation_checker.py  
--debug <affected_compute>
```

The printout contains available bandwidth on the node, and the bandwidth used for each VM.

An example of the printout is:

```
===== Checking compute-0-3.domain.tld =====

== Network device: control ==
Getting bw usage for instance name: BWM-2
Bandwidth flavor extraspec not found
Getting bw usage for instance name: BWM-5
Bandwidth flavor extraspec not found
+-----+-----+-----+
| Name      | Total | Used |
+-----+-----+-----+
| in_kbit   | 1000000 | 0 |
| in_kpkt   | 2500 | 0 |
| out_kbit  | 1000000 | 0 |
| out_kpkt  | 2500 | 0 |
+-----+-----+-----+

== Network device: default ==
Getting bw usage for instance name: BWM-2
+-----+-----+-----+-----+-----+-----+-----+-----+
| Name | used_bandwidth_in_kbit | used_bandwidth_in_kpkt | used_bandwidth_out_kbit | used_bandwidth_out_kpkt |
+-----+-----+-----+-----+-----+-----+-----+-----+
| BWM-2 | 40096.0 | 0 | 0 | 0 |
| 24096.0 | 0 | 0 | 0 |
+-----+-----+-----+-----+-----+-----+-----+-----+

Getting bw usage for instance name: BWM-5
+-----+-----+-----+-----+-----+-----+-----+-----+
| Name | used_bandwidth_in_kbit | used_bandwidth_in_kpkt |
```



```
used_bandwidth_out_kbit | used_bandwidth_out_kpkt |
+-----+-----+-----+-----+-----+-----+-----+-----+
| BWM-5 |          40096.0 |          0 |          0 |          |          |          |          |
|          24096.0 |          0 |          0 |          |          |          |          |
+-----+-----+-----+-----+-----+-----+-----+-----+
+-----+-----+-----+-----+
| Name      | Total  | Used  |
+-----+-----+-----+-----+
| in_kbit   | 40000  | 80192 |
| in_kpkt   | 2500   | 0      |
| out_kbit  | 40000  | 48192 |
| out_kpkt  | 2500   | 0      |
+-----+-----+-----+-----+
Overallocation on this compute
1
```

3. Note down the following information from the printout:

- The bandwidth that each VM is using from the bandwidth capacity on the node
- The total bandwidth capacity on the node

4. Determine which VM has at least one Virtual Network Interface Card (vNIC) connected to an SR-IOV Virtual Function (VF) of the host. In this document these tenant VMs are called SR-IOV VMs.

Identify SR-IOV VMs with the below command:

```
root@cic-1:~# neutron port-list -c binding:vnic_type -c
dns_name -c device_id | grep direct | direct | sriovm |
caa7351b-6806-40f3-9d08-cb0107defb57
```

5. Plan how to solve the overallocation issue: select which VMs need to be moved, so that the bandwidth needed for the VMs does not exceed the available bandwidth capacity.

The target host is selected automatically by the system, with regular scheduling during migration.

Note: It is recommended to start the migration with non-SR-IOV VMs, since this has a smaller impact on system traffic.

6. Migrate the non-SR-IOV VMs which do not fit in the available bandwidth on the compute, by issuing the following command on a controller:

```
nova migrate <VM_UUID_to_migrate>
```

Note: Migration of the VMs may cause traffic disturbances.



7. Wait until the VM goes into `VERIFY_RESIZE` state. When this state is reached, confirm the migration:

```
nova resize-confirm <VM_UUID_to_migrate>
```

If migration was successful, the VM goes into `ACTIVE` state.

8. Delete the SR-IOV VMs which do not fit in the available bandwidth, and recreate them using regular booting.

Note: This is required because the migration of SR-IOV VMs is currently not supported.

9. If the alarm is ceased, exit this procedure.

If the alarm remains, collect troubleshooting data as described in the *Data Collection Guideline*. For alarm-specific logs, refer to the table *Data Collection for Alarms and Alerts* in the *Data Collection Guideline*.

10. Contact the next level of maintenance support.

Further actions are outside the scope of this instruction.

11. The job is completed.