FairShareScheduler installation guide for the OpenStack IceHouse

This document provides the step-by-step procedure on how to install and configure the FairShareScheduler in a OpenStack IceHouse based infrastructure which must be already installed by using the “packstack” tool (“rdo-release-icehouse-3”).

1. Installation of the FairShareScheduler
2. Creation of the database “scheduler_priority_queue”
3. Configuration
4. Changes in the OpenStack's source code
5. Change the OpenStack's default values
6. Restart of the OpenStack services

1. Installation of the FairShareScheduler

```
$ git clone https://github.com/CloudPadovana/openstack-fairshare-scheduler.git

$ python -m compileall -f ./openstack-fairshare-scheduler/src/fairsharescheduler/

$ for svc in api cert compute conductor scheduler; do service openstack-nova-$svc stop; done

$ cp -r openstack-fairshare-scheduler/src/fairsharescheduler/ \
/usr/lib/python2.6/site-packages/nova/scheduler
```

1.1 Dependencies

```
$ yum install mysql-connector-python

$ wget http://www.webwareforpython.org/downloads/DBUtils/DBUtils-1.1.tar.gz

$ tar xzvf DBUtils-1.1.tar.gz

$ cd DBUtils-1.1/

$ python setup.py install
```

2. Creation of the database “scheduler_priority_queue”

The persistent queue of the FairShareScheduler is created by executing the following command:

```
$ mysql -uroot -p < openstack-fairshare-scheduler/script/create_db_fairshare_scheduler.sql
```

To verify the database creation, please check if the “**scheduler_priority_queue**” was created:

```
mysql> show databases;
+-------------------------+
| Database                |
+-------------------------+
| information_schema      |
| glance                  |
| keystone                |
| mysql                   |
| nova                    |
| ovs_neutron             |
| scheduler_priority_queue |
| test                    |
+-------------------------+
8 rows in set (0.00 sec)
```

2.1 Grant permissions to the “nova” user

```
$ mysql -r root -p
mysql> SHOW GRANTS FOR 'nova';
+--------------------------------------------------------------------------------------------------+
| Grants for nova@%                                                                                |
+--------------------------------------------------------------------------------------------------+
| GRANT USAGE ON *.* TO 'nova'@'%' IDENTIFIED BY PASSWORD '*F2E6A6AE17C926152E608D2D0FBFA29864E16' |
| GRANT ALL PRIVILEGES ON `nova`.* TO 'nova'@'%' WITH GRANT OPTION                                 |
+-----------------------------------------------------------------------------------------------------+
2 rows in set (0.00 sec)

mysql> GRANT SELECT ON `keystone`.* TO 'nova'@'%';
mysql> GRANT ALL PRIVILEGES on scheduler_priority_queue.* to 'nova'@'%';
mysql> FLUSH PRIVILEGES;
mysql> SHOW GRANTS FOR 'nova';
+-----------------------------------------------------------------------------------------------------+
| Grants for nova@%                                                                                   |
+-----------------------------------------------------------------------------------------------------+
| GRANT USAGE ON *.* TO 'nova'@'%' IDENTIFIED BY PASSWORD '*F2E6A6AE17C926152E608D2D0FBFA29864E16' |
| GRANT SELECT ON `keystone`.* TO 'nova'@'%'                                                          |
| GRANT ALL PRIVILEGES ON `nova`.* TO 'nova'@'%' WITH GRANT OPTION                                    |
| GRANT ALL PRIVILEGES ON `scheduler_priority_queue`.* TO 'nova'@'%'                                  |
+-----------------------------------------------------------------------------------------------------+
4 rows in set (0.01 sec)
```

3. Configuration

In order to configure the FairShareScheduler, some Nova parameters values must be updated within the [DEFAULT] section of /etc/nova/nova.conf:

```
[DEFAULT]
…

#
# Options defined in nova.scheduler.manager
#
e changes must be
# Default driver to use for the scheduler (string value)
scheduler_driver=nova.scheduler.fairsharescheduler.fairshare_scheduler.FairShareScheduler

# the topic scheduler nodes listen on (string value)
scheduler_topic=scheduler
…

#
# Options defined in nova.openstack.common.notifier.api
#

# Driver or drivers to handle sending notifications (multi valued)
notification_driver=nova.openstack.common.notifier.rpc_notifier
…

#
# Options defined in nova.openstack.common.notifier.rpc_notifier
#

# AMQP topic used for OpenStack notifications (list value)
notification_topics=notifications
notify_on_state_change=vm_and_task_state
```

moreover please add the following configuration keys still in the [DEFAULT] section:

```
[DEFAULT]
…

#
# Options for FairShareScheduler
#

# time-window length (day)
period_length=7

# number of time-windows
num_of_periods=3

# update rate fairShare (minute)
rate=5

# definition of weights
age_weight=1000
fair_share_vcpus_weight=10000
```

```
fair_share_memory_weight=7000
# decay_weight="0.5"

# number of workers which process the users requests in parallel (1=single thread)
thread_pool_size=1

# mysql parameters
mysql_user="nova"
mysql_passwd="NOVA_DB_PASS"
mysql_host="localhost"
mysql_scheduler_db="scheduler_priority_queue"
mysql_pool_size=10

# definition of the projects and users shares:
# default value for project share.
default_project_share=10

# the list of projects with the proper assigned share's value
#project_shares={'prjX_name':shareX, … 'prjY_name':shareY}

# the list of users with the proper assigned share's value
#user_shares={'prjN_name':{'usrX_name':shareX, 'usrY_name':shareY}, … }
```

Remark: set up the right values for the **mysql_host** and **mysql_passwd** configuration keys.

Finally take care that the file permissions must be:

```
$-rw-r-----. 1 nova nova 90295 18 set 11:20 /etc/nova/nova.conf
```

3.1 How to define the share's values for projects and users

For simplicity let’s say that we have only two projects, "A" and "B". The site administrator wishes to make a dynamic resource partitioning in such a way to "A" is assigned the 70% of the average total number of VMs that can be created on top of the available compute hosts and therefore "B" has the remaining 30% (please remark that the “share” is a concept different from the "quota" in the cloud's terminology). This partitioning model allows to maximize the resource utilization so that "B" can use, for “short” periods, even all resources belonging to "A" just if not used and vice versa.
To do it the administrator must edit the **project_shares** parameter in “/etc/nova/nova.conf”as follows:

```
project_shares={'A':70, 'B':30}
```

Moreover if a new project “C” is created, by default it is assigned to it the 10% of share which can be changed only by setting up the **project_shares** as above.
Finally, to every user belonging the same project by default are assigned equals shares (1%). For example if the administrator wishes to change the share of same specific users, let's say user_1 belonging to “A” and user_1 and user_3 belonging to “B” having shares 10, 21 and 13 respectively, he has to modify the **user_share** parameter as follows:

```
user_shares={'A':{'user_1':10}, 'B':{'user_1':21, 'user_3':13}}
```

4. Changes in the OpenStack's source code

Edit the manager.py (please, pay attention to respect the indentation of the source code)

```
$ vi /usr/lib/python2.6/site-packages/nova/scheduler/manager.py

class SchedulerManager(manager.Manager):
    """Chooses a host to run instances on."""

    target = messaging.Target(version='2.9')

    def __init__(self, scheduler_driver=None, *args, **kwargs):
        if not scheduler_driver:
            scheduler_driver = CONF.scheduler_driver
        self.driver = importutils.import_object(scheduler_driver)
        self.compute_rpcapi = compute_rpcapi.ComputeAPI()
        super(SchedulerManager, self).__init__(service_name='scheduler', *args, **kwargs)
        self.additional_endpoints.append(_SchedulerManagerV3Proxy(self))
    def cleanup_host(self): <== Add this method
        LOG.info("Destroying the scheduler manager...")
        self.driver.destroy()
```

To compile the code:

```
$ python -m compileall -f /usr/lib/python2.6/site-packages/nova/scheduler
```

5. Change the OpenStack's default values

To give to the scheduler the full resource management control, the default values of **cores**, **instances** and **ram** OpenStack's parameters must be changed as follows:

```
$ nova quota-class-update --cores 10000 default
$ nova quota-class-update --instances 10000 default
$ nova quota-class-update --ram 512000 512000
$ nova quota-defaults
+-----------------------------+-------+
| Quota                       | Limit |
+-----------------------------+-------+
| instances                   | 10000 |
| cores                       | 10000 |
| ram                         | 51200 |
| floating_ips                | 10    |
| fixed_ips                   | -1    |
| metadata_items              | 128   |
| injected_files              | 5     |
| injected_file_content_bytes | 10240 |
| injected_file_path_bytes    | 255   |
| key_pairs                   | 100   |
| security_groups             | 10    |
| security_group_rules        | 20    |
+-----------------------------+-------+
```

6. Restart the OpenStack services

```
$ for svc in api cert compute conductor scheduler; do service openstack-nova-$svc start; done
```

In order to check if the installation and configuration steps were executed successfully, please check the “/var/log/nova/scheduler.log” which should provide several messages logged by the **nova.scheduler.fairsharescheduler** class.