Appendix - UFM SLURM Integration

Table of contents

Simple Linux Utility for Resource Management (SLURM) is a job scheduler for Linux and Unix-like kernels.

By integrating SLURM with UFM, you can:

- Assign partition keys (PKeys) to SLRUM nodes that are assigned for specific SLURM jobs.
- Create SHARP reservations based on SLURM nodes assigned for specific SLURM jobs.

Prerequisites

- UFM 6.9.0 (or newer)
- Python 3.0 on SLURM controller
- UFM-SLURM integration files (provided independently)

Automatic Installation

A script is provided to install the UFM-SLURM integration automatically.

1. Using the SLURM controller, extract the UFM-SLURM integration tar file:

   ```
   tar -xf ufm_slurm_integration.tar.gz
   ```

2. Run the installation script using root privileges.

   ```
   sudo ./install.sh
   ```

Manual Installation

To install the UFM-SLURM integration manually:

1. Extract the UFM-SLURM integration tar file:

```
tar -xf ufm_slurm_integration.tar.gz
```

2. Copy the UFM-SLURM integration files to the SLURM controller folder.
3. Change the permissions of the UFM-SLURM integration files to 755.
4. Modify the SLURM configuration file on the SLURM controller, /etc/slurm/slurm.conf, and add/modify the following two parameters:

```
PrologSlurmctld=/etc/slurm/ufm-prolog.sh
EpilogSlurmctld=/etc/slurm/ufm-epilog.sh
```

UFM SLURM Config File

The integration process uses a configuration file located at /etc/slurm/ufm_slurm.conf. This file is used to configure settings and attributes for UFM-SLURM integration.

Here are the contents:

Attribute Name	Description	Optionality
ufm_server	IP of UFM server to connect to	Mandatory
auth_type	Should be token_auth, or basic_auth If you select basic_auth, you need to set ufm_server_user and ufm_server_pass If you select token_auth, you need to set token_auth	Mandatory
ufm_server_user	Username of UFM server used to connect to UFM, if you set auth_type=basic_auth	Mandatory, depends on the auth_type
ufm_server_pass	UFM server user password	Mandatory, depends on the auth_type
token	Generated token when you set uth_typea to token_auth	Mandatory, depends

Attribute Name	Description	Optionality
		on the auth_type
pkey_allocation	By setting pkey_allocation to true, UFM SLURM Integration will use static Pkey assignment to create new Pkey, otherwise it will use the default management Pkey 0x7fff	Mandatory, default is True.
pkey	Hexadecimal string between "0x0001"-"0x7ffe" exclusive	Optional, default is "0x7fff" (This is the default management pkey)
ip_over_ib	PKey is a member in a multicast group that uses IP over InfiniBand	Hidden param, default is True
index0	If true, the API will store the PKey at index 0 of the PKey table of the GUID	Hidden param, default is False
sharp_allocation	By setting sharp_allocation to true, UFM SLURM Integration will create new SHARP allocation with all SLURM job IDs allocated to hosts	Mandatory, default is False
partially_alloc	By setting this to false, UFM will fail the SHARP allocation request if at least one node does not exist in the fabric	Optional, default is False
app_resources_limit	Application resources limitation	Hidden param, default is -1
log_file_name	Name of integration logging file	Optional

Configuring UFM for NVIDIA SHARP Allocation

To configure UFM for NVIDIA SHARP allocation/deallocation you must set sharp_enabled and enable_sharp_allocation to true in gv.cfg file.

Generate token_auth

If you set auth_type=token_auth in UFM SLURM's config file, you must generate a new token by logging into the UFM server and running the following curl command:

```
curl -H "X-Remote-User:admin" -XPOST http://127.0.0.1:8000/app/tokens
```

Then you must copy the generated token and paste it into the config file beside the token_auth parameter.

Prolog and Epilog

After submitting jobs on SLURM, there are two scripts that are automatically executed:

- ufm-prolog.sh – the prolog script is executed when a job is submitted and before running the job itself. It creates the partition key (pkey) assignment and/or NVIDIA SHARP reservation and assigns the SLURM job hosts for them.
- ufm-epilog.sh – the epilog script is executed when a job is complete. It removes the partition key (PKey) assignment and/or NVIDIA SHARP reservation and free the associated SLURM job hosts.

Integration Files

The integration use scripts and configuration files to work, which should be copied to SLURM controller /etc"/slurm. Here is a list of these files:

File Name	Description
ufm-prolog.sh	Bash file which executes jobs related to UFM after the SLURM job is completed
ufm-epilog.sh	Bash file which executes jobs related to UFM before the SLURM job is executed
ufm_slurm.conf	UFM-SLURM integration configuration file
ufm_slurm_prolog.py	Python script file which creates the partition key (pkey) assignment and/or SHARP reservation when the prolog bash script is running

File Name	Description
ufm_slurm_epilog.py	Python script file which removes partition key (pkey) assignment and/or SHARP reservation based on the SLURM job hosts.
ufm_slurm_utils.py	Utility Python file containing functions and utilities used by the integration process

Running UFM-SLURM Integration

Using the SLURM controller, execute the following commands to run your batch job:

```
$ sbatch -N4 slurm_demo.sh
Submitted batch job 1
```

> **Note**
>
> N4 is the number of compute nodes used to run the jobs.
> slurm_demo.sh is the job batch file to be run.

The output and result are stored on the working directory slurm-{id}.out where {id} is the ID of the submitted job.

In the above example, after executing sbatch command, you can see that the submitted job ID is 1. Therefore, the output file would be stored in slurm-1.out.

Execute the following command to see the output:

```
$cat slurm-1.out
```

On the UFM side, a partition key (PKey) is created in case the pkey_allocation parameter is set to true in the configuration file, and the user provided the PKey name including the SLURM job IDs allocated to the hosts. Otherwise it will use the default management PKey.

In addition, the UFM-SLURM will create SHARM AM reservation in case the sharp_allocation parameter is set to true in the ufm_slurm.conf file.

After the SLURM job is completed, the UFM removes the job-related partition key (PKey) assignment and SHARP reservation, if they were created.

From the moment a job is submitted by the SLURM server until its completion, a log file named /tmp/ufm_slurm.log logs all of the actions and errors that occurred during the execution.

This log file can be changed by modifying the log_file_name parameter in /etc/slurm /ufm_slurm.conf.