

# Meta Scheduling

User Guide 1.0

November 2024



## Legal Notices

Copyright © 2024. Adaptive Computing Enterprises, Inc. All rights reserved.

This documentation and related software are provided under a license agreement containing restrictions on use and disclosure and are protected by intellectual property laws. Except as expressly permitted in your license agreement or allowed by law, you may not use, copy, reproduce, translate, broadcast, modify, license, transmit, distribute, exhibit, perform, publish, or display any part, in any form, or by any means. Reverse engineering, disassembly, or decompilation of this software, unless required by law for interoperability, is prohibited.

This documentation and related software may provide access to or information about content, products, and services from third-parties. Adaptive Computing is not responsible for and expressly disclaims all warranties of any kind with respect to third-party content, products, and services unless otherwise set forth in an applicable agreement between you and Adaptive Computing. Adaptive Computing will not be responsible for any loss, costs, or damages incurred due to your access to or use of third-party content, products, or services, except as set forth in an applicable agreement between you and Adaptive Computing.

Adaptive Computing, Moab®, Moab HPC Suite, Moab Viewpoint, Moab Wide Area Grid, NODUS Cloud OS™, On-Demand Data Center™, and other Adaptive Computing products are either registered trademarks or trademarks of Adaptive Computing Enterprises, Inc. The Adaptive Computing logo is a trademark of Adaptive Computing Enterprises, Inc. All other company and product names may be trademarks of their respective companies.

The information contained herein is subject to change without notice and is not warranted to be error free. If you find any errors, please report them to us in writing.

Adaptive Computing Enterprises, Inc.  
1100 5th Avenue South, Suite #201  
Naples, FL 34102  
+1 (239) 330-6093  
[www.adaptivecomputing.com](http://www.adaptivecomputing.com)  
[support@adaptivecomputing.com](mailto:support@adaptivecomputing.com)

# Contents

<b>Chapter 1: Meta Scheduling Introduction</b> .....	<b>4</b>
1.1 Requirements .....	4
1.2 Workflow Diagram .....	5
1.3 Architecture .....	6
<b>Chapter 2: Configuring a New Scheduler</b> .....	<b>7</b>
2.1 Using Client Commands Directly on the Moab Server .....	7
2.2 Using the SSH Connection Method .....	8

## Chapter 1: Meta Scheduling Introduction

Moab Workload Manager (a.k.a. Moab) has the capability to interact with any resource manager via an interface called NativeRM. The Meta Scheduling product uses this approach, enabling Moab to query and submit workloads to other schedulers such as SLURM, OpenPBS, and LSF.

Moab gathers information about the nodes and jobs on these schedulers and uses this information to send workloads to them. Automated batch job script converters automatically convert a Torque/Moab batch job script to the target scheduler's format before submitting the job. This makes it very easy for the user as they only have to create a single job script to run workloads on any supported scheduler.

### Contents:

1.1 Requirements .....	4
1.2 Workflow Diagram .....	5
1.3 Architecture .....	6

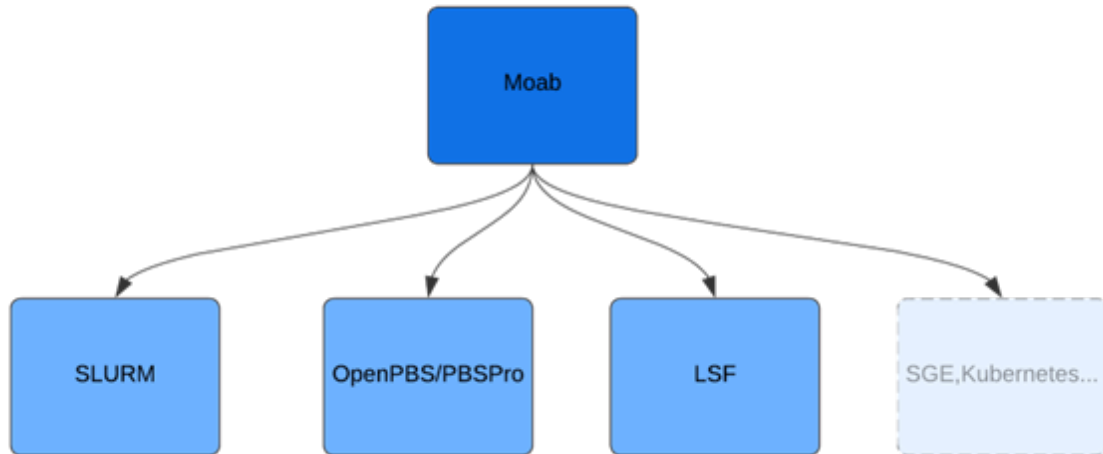
## 1.1 Requirements

The following are required to use Meta Scheduling:

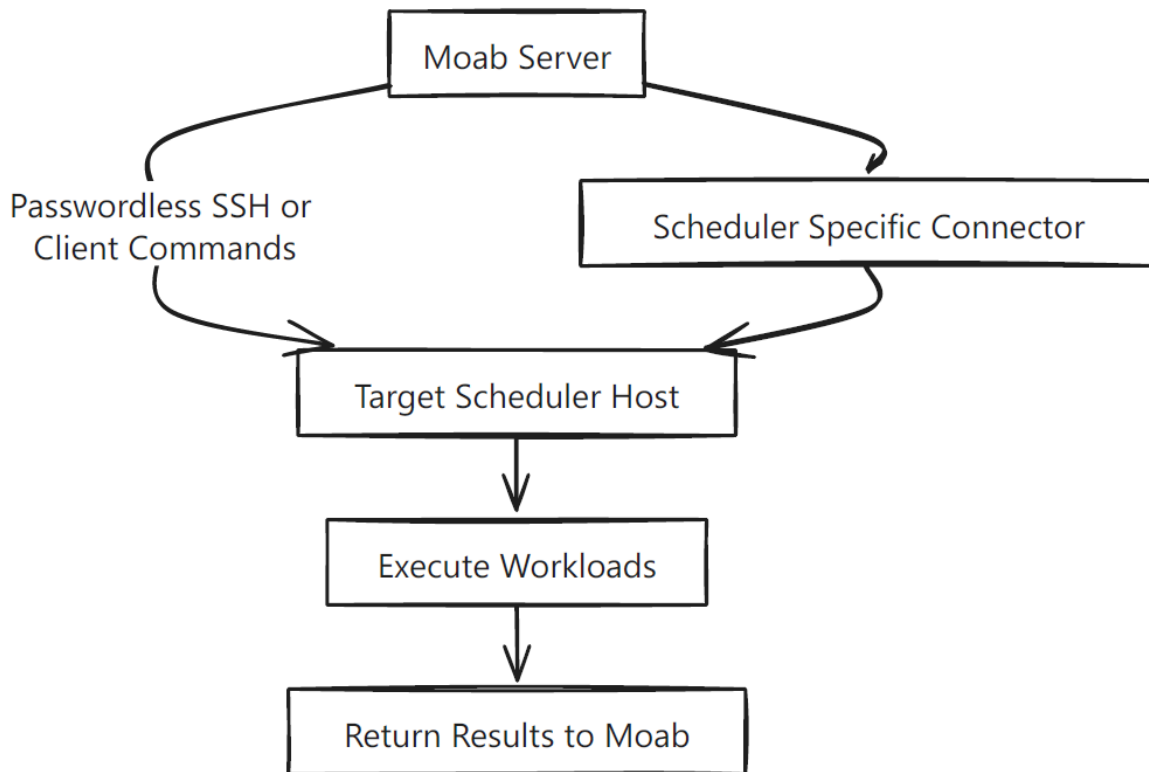
- A Moab Workload Manager installation.
  - A shared file system with the users' home directories on the shared file system.
  - One of the following:
    - Passwordless SSH from the Moab server to the target scheduler host, both as root (or another user that is an administrator on the target scheduler) and as each user submitting workload.
- OR
- Client commands for the target scheduler installed and configured on the Moab server.
  - A scheduler specific connector, requested from [Adaptive Computing](#) as a tarball (as this is a pre-beta product).

## 1.2 Workflow Diagram

Job scripts are converted automatically to the target scheduler format.



# 1.3 Architecture



## Chapter 2: Configuring a New Scheduler

Included in the tarballs for each scheduler are the necessary connectors for workloadquery, clusterquery and torque2<target\_scheduler> conversion programs. Also included are sample configuration snippets, and below are some examples.

### Contents:

2.1 Using Client Commands Directly on the Moab Server .....	7
2.2 Using the SSH Connection Method .....	8

## 2.1 Using Client Commands Directly on the Moab Server

1. In Moab.cfg, add a RMCFG line for the new scheduler, for example for SLURM:

```
RMCFG[slm] TYPE=NATIVE CLUSTERQUERYURL=exec:///opt/moab/tools/slurm-
nativerm/clusterquery WORKLOADQUERYURL=exec:///opt/moab/tools/slurm-
nativerm/workloadquery JOBSUBMITURL=exec:///opt/moab/tools/slurm-nativerm/job-
submit.sh
```

For job submissions, a scheduler specific job script conversion program is included, for example, torque2slurm.

The job-submit.sh script looks similar to this:

```
#!/bin/bash
for i in "$@"
do
    export $i
done
if [ ! -d "$IWD/converted_scripts/" ]
then
    echo "${IWD}/converted_scripts does not exist, let's create it"
    mkdir ${IWD}/converted_scripts
fi
/opt/moab/tools/slurm-nativerm/torque2slurm $EXEC ${IWD}/converted_
scripts/${JNAME}.slurm
chmod +x ${IWD}/converted_scripts/${JNAME}.slurm
sbatch ${IWD}/converted_scripts/${JNAME}.slurm | awk '{print $NF}'
```

2. Restart Moab:

```
# mschedctl -R

RM[slm]          State: Active  Type: NATIVE:AGFULL  ResourceType: COMPUTE
Timeout:         30000.00 ms
```

```
Cluster Query URL:  exec:///opt/moab/tools/slurm-nativerm/cq.sh
Workload Query URL: exec:///opt/moab/tools/slurm-nativerm/wq.sh
Job Submit URL:    exec:///opt/moab/tools/slurm-nativerm/submit.sh
Partition:         slm
Event Management:  (event interface disabled)
RM Performance:    AvgTime=0.49s  MaxTime=3.27s  (114604 samples)
RM Languages:      NATIVE
RM Sub-Languages:  NATIVE
```

### 3. Check the resource manager status:

```
# mdiag -R -v
```

### 4. Submit a job to the target scheduler using the partition name:

```
msub -l nodes=1:ppn=1,partition=slm sleep.sh
```

## 2.2 Using the SSH Connection Method

### 1. In Moab.cfg, add a RMCFG line for the new scheduler, for example for SLURM:

```
RMCFG[slm] TYPE=NATIVE CLUSTERQUERYURL=exec:///opt/moab/tools/slurm-
nativerm/clusterquery.sh WORKLOADQUERYURL=exec:///opt/moab/tools/slurm-
nativerm/workloadquery.sh JOBSUBMITURL=exec:///opt/moab/tools/slurm-nativerm/job-
submit.sh
```

These scripts are simply shell scripts that SSH to the target scheduler host and run the connector programs, such as:

```
clusterquery.sh:
#!/bin/bash
ssh metaslurm /usr/local/sbin/clusterquery
```

```
workloadquery.sh:
#!/bin/bash
ssh metaslurm /usr/local/sbin/workloadquery
```

For job submissions, a scheduler specific job script conversion program is included, for example, `torque2slurm`.

The `job-submit.sh` script looks similar to this:

```
#!/bin/bash
for i in "$@"
do
    export $i
done
if [ ! -d "$IWD/converted_scripts/" ]
then
    echo "${IWD}/converted_scripts does not exist, let's create it"
    mkdir ${IWD}/converted_scripts
fi
/opt/moab/tools/slurm-nativerm/torque2slurm $EXEC ${IWD}/converted_
scripts/${JNAME}.slurm
```



```
chmod +x ${IWD}/converted_scripts/${JNAME}.slurm  
ssh metaslurm "sbatch ${IWD}/converted_scripts/${JNAME}.slurm" | awk '{print  
$NF}'#!/bin/bash
```

## 2. Restart Moab:

```
# mschedctl -R  
  
RM[slm]      State: Active  Type: NATIVE:AGFULL  ResourceType: COMPUTE  
Timeout:    30000.00 ms  
Cluster Query URL:  exec:///opt/moab/tools/slurm-nativerm/cq.sh  
Workload Query URL: exec:///opt/moab/tools/slurm-nativerm/wq.sh  
Job Submit URL:    exec:///opt/moab/tools/slurm-nativerm/submit.sh  
Partition:    slm  
Event Management: (event interface disabled)  
RM Performance:  AvgTime=0.49s  MaxTime=3.27s  (114604 samples)  
RM Languages:   NATIVE  
RM Sub-Languages: NATIVE
```

## 3. Check the resource manager status:

```
# mdiag -R -v
```

## 4. Submit a job to the target scheduler using the partition name:

```
msub -l nodes=1:ppn=1,partition=slm sleep.sh
```