

ParaStation MOM

Software Product Detailed Description

Product: ParaStation MOM

Date: June 2012

Document number: PSMOM-1.0en

This software product description documents the functionality provided by the [ParaStation MOM](#) as well as the system prerequisites required for installation and operation, licensing scheme and other useful information.

Overview

The [ParaStation MOM](#) as part of the [ParaStationV5](#) cluster suite offers a tight integration between the [ParaStation](#) process management and the PBS batch system. The [PSMOM](#) is a complete replacement of the PBS execution daemon *pbs_mom*. It is implemented as a plug-in for the [ParaStation](#) process management daemon *psid*. Using the [PSMOM](#) plug-in the *psid* is therefore the only daemon needed on the compute nodes for batch system operation. Furthermore the [ParaStation MOM](#) offers a fast and reliable job startup based on the [ParaStation RDP](#) communication layer and offers various unique features.

Features

- Advanced accounting for all processes of serial and parallel jobs.
- A controlled ssh log-in for users to the compute nodes is provided using a PAM module.
- Secure communication between the [ParaStation](#) process management and the PBS batch system.
- Inhibits manipulation of the node list by users and the specification of the resources management system is followed exactly.
- The [ParaStation](#) process management and the PBS resources management share the same state information of a node.
- Resource-limits can be enforced for parallel jobs with a minimum of effort, since all necessary data is available from the [ParaStation](#) accounting layer.
- Advanced cleanup features to identify and kill user processes even started via the "nohup" command as well as user processes which turned themselves into daemons.
- Reliable and highly scalable node-to-node communication using the [ParaStation RDP](#) communication layer.
- Faster node down detection to prevent common job errors.
- Easy definition of global resource limits (/etc/security/limits.conf) using PAM.
- Avoiding an additional daemon on the compute nodes has positive effects regarding operating system jitter.
- Fully integrated in [ParaStation MPI](#) as a lightweight plug-in.
- On-the-fly reconfiguration and information gathering using the *psadmin* command.

Accounting

The [ParaStation MOM](#) is able to account all processes of serial and parallel jobs. When using the [PSMOM](#) the compute processes are started under the control of the [ParaStation](#) process management daemon *psid* and therefore also under the control of the [PSMOM](#). With the combined information from [PSMOM](#) and *psid* the job-script itself and all corresponding compute processes can be identified and correlated. This gives the [PSMOM](#) the ability to account every process of a parallel job and report more accurate accounting data to the batch system.

Software Product Detailed Description

SSH Login for ordinary users

To allow ssh logins to compute nodes for ordinary users, the [ParaStation MOM](#) is using a PAM module which is loaded into the SSH daemon process. Every new ssh connection will be started using two internal steps. The first step is the authentication process which is not influenced by the [PSMOM](#). All usual mechanisms for the ssh authentication process can be used. This includes the ssh password login or the public key authentication. Therefore it is possible to use MPI versions which rely on the process startup via ssh.

If the user has successfully authenticated itself to the SSH daemon the second step is to open a new SSH session. The [ParaStation MOM](#) will perform different checks to verify if the user should be allowed to login into the local compute node. If the last job from a user on a node terminates all leftover SSH sessions will also be terminated.

Scalability

For node to node communication the [ParaStation MOM](#) is using the ParaStation RDP communication layer to overcome scalability issues. Within [PSMOM](#) there are no limitations with respect to the number of processes per application and the number of simultaneously executing processes per compute node. The decisive factors for these are the available system resources such as, for example, the available memory and configured limits, like maximum number of processes per node.

The [ParaStation MOM](#) is tested with jobs running on 3200 nodes in parallel and sizes up to 25,600 processes per job. No limiting factors are foreseeable within [ParaStation MOM](#).

Usability

The [ParaStation MOM](#) can be reconfigured on-the-fly using the *psiadmin* command without the need of a restart. All configuration options are explained directly using the plug-in help function in the *psiadmin* without the need for extern documentation. Furthermore detailed information about running jobs, job statistics, the internal state and open network/SSH connections can be shown.

Prerequisites

To install and operate [ParaStation MOM](#), the following prerequisites must be met:

- A recent version of the ParaStation process management psmgmt is needed
- The TORQUE Resource Manager 3.x and 2.x branches are supported
- All PBS compatible schedulers including Moab and Maui are supported
- [ParaStation MOM](#) is supported on openSuSE, SLES, RHES, Fedora, CentOS

Free disk space: The installation of the [ParaStation MOM](#) package requires approximately 1MB free disk space per node. Administrator privileges are required for installation. The software can be exported from a file sever via NFS or Lustre to all cluster nodes.

Media

The [ParaStation MOM](#) package is available in binary format as pre-compiled RPM package ready for installation on the target system. Please contact support@par-tec.com how to obtain this package.

ParaStation MOM

Software Product Detailed Description

License

ParaStation MOM is not freeware. Details can be found in the ParaStation license agreement on www.par-tec.com.

This product includes software developed by NASA Ames Research Center, Lawrence Livermore National Laboratory, and Veridian Information Solutions, Inc. Visit www.OpenPBS.org for OpenPBS software support, products, and information.

Support

After signing a support contract, support for **PSMOM** is granted for the agreed period. The maximum response time is next business day. Support is available by telephone, email, and/or remote login. On-site support at the installation site is not included.

The support comprises all **ParaStationV5** components as well as the open source software utilized (if applicable). Other open source software tools that have been provided free of charge are only supported if resources are available, a general claim cannot be advanced on the basis of this support agreement.

Scope of delivery

ParaStation MOM comprises the following components:

- Software RPM packages (psmom),
- Support as agreed in the support contract.

Copyright

ParTec, ParaStation and **ParaStationV5** are registered trademarks of ParTec Cluster Competence Center GmbH. All other product and brand names are trademarks or registered trademarks of their respective owners.

The information in this version of the software product detailed description is valid as from the time of publishing. Errors & omissions excluded.

Further information

For further information about **ParaStation MOM** and related products visit <http://www.par-tec.com> or send an email to sales@par-tec.com.