



SAS® Grid Manager

Workload balancing. High availability. Faster processing.
All in a flexible, centrally managed grid computing environment.

What is SAS® Grid Manager?

SAS Grid Manager uses a patented technology to deliver enterprise-class capabilities that enable many SAS solutions to automatically use a centrally managed grid computing infrastructure to provide workload balancing, high availability and parallel processing for business analytics jobs and processes.

Why is SAS® Grid Manager important?

SAS Grid Manager makes it easy to accommodate compute-intensive applications and a growing number of users cost-effectively across your available hardware resources – while ensuring continuously high availability for your business analytics applications. With SAS Grid Manager, you can create a managed, shared environment for efficiently processing large volumes of SAS programs.

For whom is SAS® Grid Manager designed?

It is designed for CIOs, IT managers, data center managers and grid computing architects seeking to manage SAS programs in a shared environment. It also can be used by statisticians, business analysts and application developers to reduce processing times – and achieve faster results – for data integration, reporting and analytic jobs.

IT budgets in most organizations these days are typically limited, which makes meeting the computing demands of today's business environment a constant challenge. Buying the latest and greatest servers (i.e., scaling up) to meet peak-demand computing loads is one solution – but that can be both costly and inefficient.

As organizations' use of business analytics grows, so does the need for a flexible IT infrastructure that can scale cost-effectively while meeting peak demands and managing growing, increasingly diverse user workloads. SAS Grid Manager is the answer.

SAS Grid Manager enables organizations to create a managed, shared grid computing environment for processing large volumes of data and analytic programs. The solution provides critical capabilities for meeting an organization's business analytics needs, including:

- Workload balancing.
- Job prioritization.
- High availability.
- Parallel processing.
- Resource assignment and monitoring.

Key Benefits

- **Meet changing business demands with dynamic workload balancing.** SAS Grid Manager gives IT greater flexibility to meet service level commitments by easily reassigning computing resources to meet peak workloads or changing business demands. The solution provides a central point of control for administering policies, programs, queues and job prioritization across multiple types of users and applications to achieve business goals under a given set of constraints.

- **Create a SAS computing environment that has high availability.**

Having multiple servers in a grid computing environment enables jobs to run on the best available resource. If a server fails, its jobs can be transitioned seamlessly to another server – providing high availability. In addition, IT staff can perform maintenance on specific servers without interrupting analytics jobs, as well as introduce additional computing resources without disrupting the business.

- **Get faster results from your existing IT infrastructure.**

Multiprocessing capabilities let you divide individual jobs into subtasks that are run in parallel on the best available hardware resource. The SAS programs best-suited for parallel processing are those with large data sets and long run times, as well as those with replicate runs of independent tasks running against large data sets. Processing data integration, reporting and analytical jobs faster accelerates decision making across the enterprise.

- **Grow computing resources incrementally and cost-effectively.**

SAS Grid Manager lets you fully utilize all available computing resources *now*, and cost-effectively scale out as needed, adding capacity in single-processing units to keep IT spending in check. Because you can add low-cost commodity hardware resources incrementally, there's no need to size today's environment for what you anticipate your needs will be a few years from now.



Product Overview

SAS Grid Manager's patented technology uses industry-leading grid computing middleware from Platform Computing to get maximum availability from your business analytics environment. The solution gives you a competitive advantage by enabling you to balance user and application workloads among available computing resources, so you get results much more quickly. IT can add computing resources – in the form of lower-cost commodity hardware – incrementally, eliminating the need to size today's environment for tomorrow's demands.

Dynamic, Resource-Based Workload Balancing

SAS Grid Manager uses proven, industry-leading grid computing middleware to deliver enterprise-class dynamic workload balancing for multiple users and applications. It automates the management and optimization of SAS grids and provides resilience by efficiently distributing the processing of SAS programs across multiple CPUs. Centralized administration lets you enforce policies to determine job

prioritization based on different users' needs, as well as deliver better service levels to business units. And it gives IT staff the flexibility to easily identify and allocate computing resources to match changing demands.

Shared Environment with High Availability

With SAS Grid Manager, there's no need to purchase a separate third-party tool for high availability, because the solution includes several high-availability capabilities for all services critical to a SAS environment. Failover to another node within the grid eliminates the need for a hot standby. And you can perform rolling maintenance with no interruption to users. Guaranteed execution of critical SAS jobs ensures that they are completed correctly; if they are not, you can take corrective action with checkpoint and automatic restart.

Grid-Enabled SAS®

SAS Data Integration Studio (a component of SAS Enterprise Data Integration Server and SAS Data Integration Server) and SAS® Enterprise Miner™ are automatically tailored for parallel processing in a grid computing environment. To achieve maximum processing

efficiency with minimum user intervention, these programs detect the grid environment at the time of execution. The grid-enabled logic that is produced can be saved as SAS Stored Processes for use by other SAS reporting clients to generate results for more users as cost-effectively as possible.

You can set up other SAS solutions, including SAS® Enterprise Guide® and SAS® Risk Dimensions®, to automatically submit SAS jobs to a grid of shared computing resources. All SAS programs can take advantage of a grid computing environment with the addition of programming syntax and a structure that allows the submission of entire programs to the grid or the parallel execution of program steps (subtasks). The grid option of the SAS Code Analyzer in Base SAS software automates this process by generating a new version of a SAS program that has been restructured to run in a distributed environment.

You can schedule your production SAS workflows to run across grid resources using the schedule manager plug-in within the SAS Management Console.

The screenshot displays the SAS Grid Manager Host Dashboard. At the top, there are navigation tabs for 'config', 'graphs', 'alerts', 'grid', 'syslogs', and 'HA'. The main area is titled 'Cluster Filter Options' and includes filters for Cluster, Group, Cacti, Thresholds, Type, Model, Load Stat, Batch Stat, and Refresh. Below this is the 'Host Status' section, which shows a grid of host icons representing different clusters: 'build', 'lsf7damo', and 'lsf7u4'. A pop-up window titled 'HOST: BLD002 GRID: BUILD' is open, displaying detailed information for a specific host, including 'Available Links', 'Running Jobs', 'Host Graphs', 'Remote to Host', 'Load/Batch Status', 'Load Info', and 'Batch Info'. The 'Host Status Legend' at the bottom identifies various host states: Unavailable, Busy/Closed, Idle/Closed, Low Resources, Busy, Idle w/ Jobs, Idle, Starved, and Admin Down.

Platform RTM for SAS allows you to monitor and manage the operation of your SAS grid on a day-to-day basis.

These SAS workflows can be created using SAS Web Report Studio or SAS Data Integration Studio, or they can be ad hoc SAS programs.

Automated, Web-Based Management and Monitoring

SAS Grid Manager includes Platform RTM for SAS, a Web-based tool for monitoring and managing resources, users and jobs running in the grid. Platform RTM also serves as an interface for configuring and managing high-availability services, as well as defining alerts to be sent when thresholds are exceeded.

Scalability for Running More Complex Analytics Faster and Taking Advantage of All Computing Resources

Because SAS has made grid computing an automatic capability within multiple applications, processing times are greatly reduced. As a result, you can integrate, cleanse and analyze larger volumes of data more quickly.

You can schedule a wide variety of SAS jobs across grid environments for optimal resource utilization and faster processing. You can divide individual SAS jobs into subtasks that are then executed in parallel to accelerate processing and increase workload throughput. In today's international organizations, nightly batch-processing windows no longer exist. As a result, data is available 24/7 and can be quickly loaded and analyzed.

Key Features

A Managed, Shared Environment

- Improves efficiency of program distribution and CPU utilization through dynamic, resource-based load balancing.
- Makes computing resources available to multiple users and multiple applications for running larger or more complex analysis.
- Provides job, queue, host and user management across your enterprise.
- Enables job prioritization by rules-based job queues to govern the use of computing resources.
- Provides automatic identification, allocation, management and optimization of computing resources and program flows.
- Allows administrator to easily create a set of metadata-defined grid and SAS options to be applied automatically to workloads submitted to the grid based on the user's identity and the application being used (SAS Enterprise Guide, SAS Enterprise Miner, etc.) to access the grid.
- Simplifies administration of SAS environment through centralized policies.

High Availability

- Provides high-availability capabilities for critical SAS services such as the SAS Metadata Server.
- Uses a grid node as a hot-standby machine for failover.
- Enables nondisruptive, rolling maintenance.
- Detects hardware and software failures in the grid and recovers appropriately.
- Ensures that SAS jobs will be completed optimally.
- Restarts jobs automatically from the last successful checkpoint when used with the SAS checkpoint/restart feature.

Grid-Enabled SAS®

- Automatically tailors SAS Data Integration Studio and SAS Enterprise Miner for parallel processing and job submission in a grid environment.
- Balances the load of many SAS Enterprise Guide users through easy submission to the grid.
- Provides load balancing for all SAS servers to improve throughput and response time of all SAS clients.
- Uses SAS Code Analyzer to analyze job dependencies in SAS programs, and generates grid-ready code:
 - Used by SAS Data Integration Studio and SAS Enterprise Guide to import SAS programs.
- Provides automated session spawning and distributed processing of SAS programs across a set of diverse computing resources.
- Speeds up processing of applicable SAS programs and applications, and provides more efficient computing resource utilization.
- Enables scheduling of production SAS workflows to be executed across grid resources:
 - Provides a process flow diagram to create SAS flows of one or more SAS jobs that can be simple or complex to meet your needs.
 - Uses all of the policies and resources of the grid.
- Enables many SAS solutions and user-written programs to be configured easily for submission to a grid of shared resources.
- Integrates with all SAS Business Intelligence clients and analytic applications by storing grid-enabled code as SAS Stored Processes.
- Provides greater resilience for mission-critical applications and high availability for the SAS environment.
- Includes a command-line batch submission utility called SASGSUB:
 - Allows you to submit and forget, and reconnect later to retrieve results.
 - Enables integration with other standard enterprise schedulers.
- Enables batch submission to leverage checkpoint and automatically restart jobs.
- Applies grid policies to SAS workspace servers when they are launched through the grid.

SAS® Grid Computing System Requirements

To learn more about SAS Grid Computing system requirements, download white papers, view screenshots and see other related material, please visit sas.com/grid.

Key Features (continued)

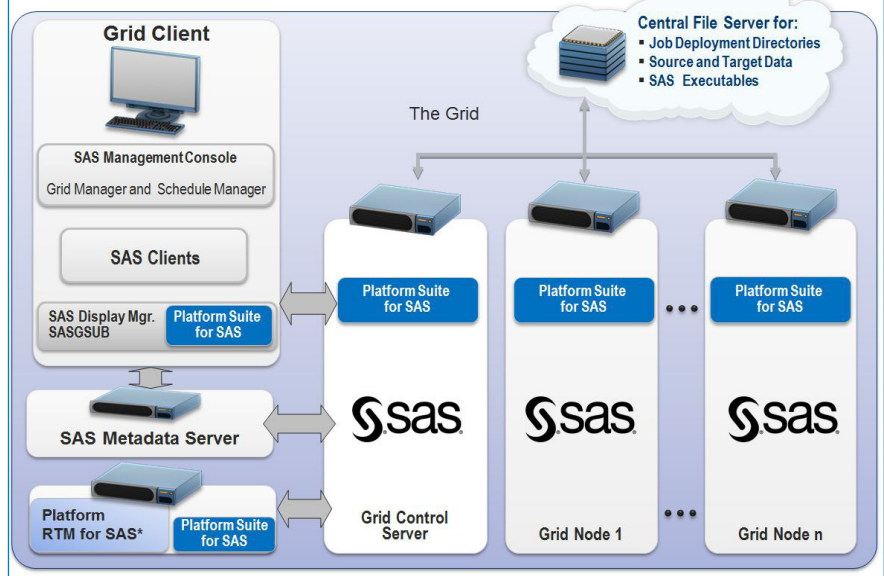
Real-Time Monitoring and Administration

- Platform RTM for SAS allows you to monitor and manage SAS Grid Manager:
 - Provides a Web-based tool for monitoring and administering multiple SAS grids.
 - Lets you monitor and administer your grids from any location, as long as you have access to a Web browser.
 - Includes graphs for tracking resource usage, users and jobs running on the grid.
 - Provides a GUI for modifying grid configurations and defining alerts when critical thresholds are crossed.
 - Provides a GUI for configuring and managing critical services for high availability.
 - Supports filtering and role definitions for customizing displays and activities that can be performed.
- Grid manager plug-in provides monitoring capabilities from SAS Management Console:
 - Provides real-time monitoring of SAS programs and computing resources.
 - Enables filtering of resource and activity information, allowing users to focus on specific aspects of grid operations.
 - Provides graphical representation of job distribution.

Flexible Infrastructure

- Allows you to add computing resources incrementally to cost-effectively accommodate a growing number of users, as well as meet increased business needs.
- Decouples the computing infrastructure from the SAS applications to allow business users to focus on their processes.
- Creates a shared environment for easily and dynamically allocating resources to meet peak needs of different business users over time.

SAS® Grid Computing Architecture



SAS® Grid Computing architecture.