



SAS® Grid Computing

A flexible, centrally managed computing environment for workload balancing, high availability and faster processing

What is SAS® Grid Computing?

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

Why is SAS® Grid Computing important?

With SAS Grid Computing, it is easier and cost-effective to accommodate compute-intensive applications and growing numbers of users resources across available hardware resources and ensure continuous high availability for business analytics applications. You can create a managed, shared environment to process large volumes of SAS programs in an efficient manner.

For whom is SAS® Grid Computing 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 time for data integration, reporting and analytic jobs, and achieve faster results.

Meeting the computing demands of today's business environment with fixed and limited IT budgets is a challenge for every enterprise. Buying the latest and greatest servers (scaling up) to meet peak-demand computing loads is one way to meet those needs, but it can be costly and inefficient. With increasingly complex analytics, growing data volumes and tightening batch windows, businesses are also facing an even greater need for reducing planned and unplanned downtime, and guaranteeing uptime and continuity of services (commonly referred to as high availability).

As data integration, analytics and reporting capabilities grow in strategic importance and encompass increasing numbers of users and large quantities of data, the ability to cost-effectively scale a business analytics system to gain operational flexibility, improve performance and meet peak demands using grid computing becomes a competitive advantage.

SAS Grid Computing enables organizations to create a managed, shared environment to process large volumes of data and analytic programs more efficiently. It provides critical capabilities that are necessary for today's business analytics environments, including workload balancing, job prioritization, high availability and built-in failover, parallel processing and resource assignment and monitoring.

Key Benefits

- **Gain flexibility with dynamic workload balancing to meet changing business demands.** SAS Grid Manager provides a central point for administering policies, programs, queues and job prioritization to achieve business goals across multiple types of users and applications under a given set of constraints. IT can gain flexibility and meet service levels by easily reassigning computing resources to meet peak workloads or changing business demands.
- **Create a highly available SAS computing environment.** The presence of multiple servers in a grid environment enables jobs to run on the best available resource and if a server fails, its jobs can be seamlessly transitioned to another server – providing a highly-available business analytics environment. High availability also enables IT staff to perform maintenance on specific servers without interrupting analytics jobs, as well as introduce additional computing resources without disruption to the business.
- **Accelerate time to results using existing IT infrastructure.** Multiprocessing capabilities allow individual jobs to be divided into sub-tasks that are run in parallel on the best available hardware resource in a grid environment. SAS programs best suited for parallel processing are those with large data sets and long run times, and those with replicate runs of independent tasks running against large data sets. Faster processing of data integration, reporting and analytical jobs will accelerate decision making across the enterprise.



- **Incrementally grow computing resources in a cost-effective manner.** SAS Grid Manager allows organizations to fully utilize all available computing resources now and cost-effectively scale out as needed, adding capacity in single-processing units for incremental IT spending. Adding low-cost commodity hardware resources incrementally eliminates the need to size today's environment for what you expect your needs to be a few years from now.

Product Overview

SAS Grid Computing, a patented technology, uses industry-leading grid computing middleware from Platform Computing to maximize availability of your business analytics environment, balance user and application workloads among available computing resources, and accelerate time to results for competitive advantage. It also gives IT the ability to incrementally add computing resources, eliminating the need to size today's environment for what you expect your needs to be a few years from now, and leverages lower-cost commodity hardware.

Dynamic, resource-based workload balancing

SAS Grid Computing delivers enterprise-class dynamic workload balancing for multiple users and applications by incorporating proven, industry-leading grid computing middleware. 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 can enforce policies to determine job prioritization based on needs of different users and deliver better service levels to business units. It provides flexibility for IT staff to easily identify and allocate computing resources to match changing demands.

Shared environment with high availability

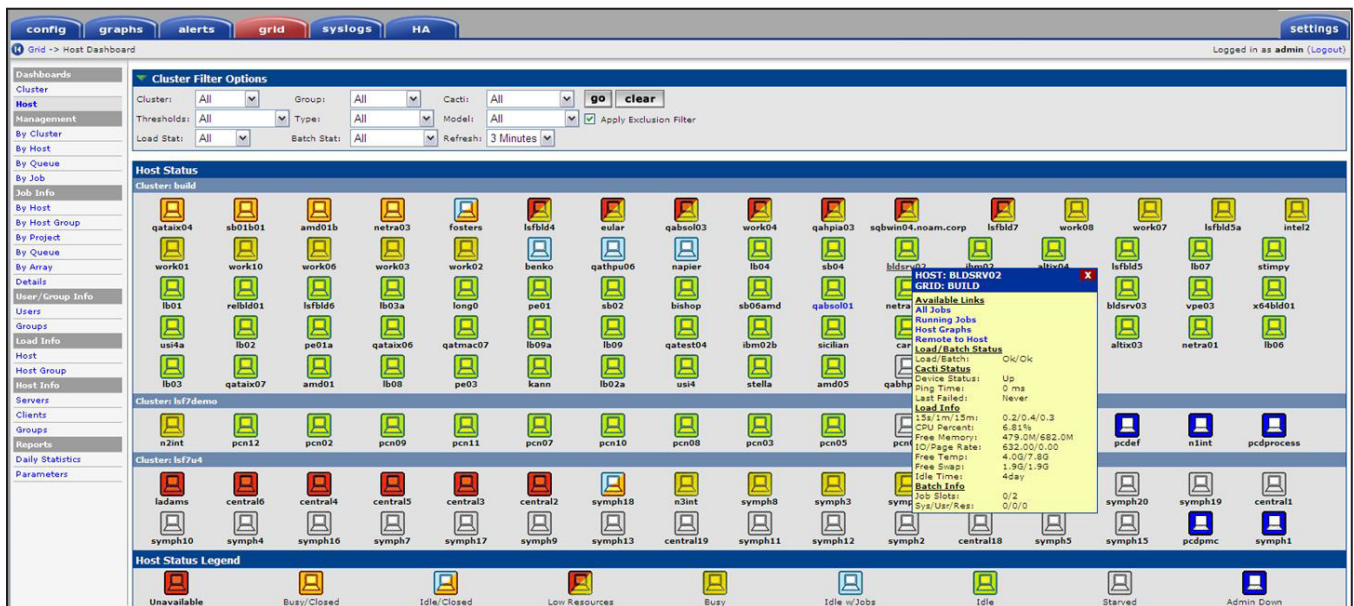
SAS Grid Manager offers several high-availability capabilities for all services that are critical to the SAS environment. Failover of all critical services to another node within the grid eliminates the need for a hot-standby and the need to purchase a separate third-party tool for high availability. SAS Grid Manager allows you to perform rolling maintenance with no interruption to users. You get guaranteed execution of critical SAS

jobs to ensure they are completed in the correct manner; if 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.

Other SAS solutions, including SAS® Enterprise Guide® and SAS® Risk Dimensions®, can be set up 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 structure that allows entire programs to be submitted to the grid or program



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

steps (subtasks) to be executed in parallel. 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.

Your production SAS workflows can be scheduled to run across the grid resources using the schedule manager plug-in within the SAS Management Console. These SAS workflows can be created using SAS Web Report Studio or SAS Data Integration Studio, or they can be ad hoc SAS programs.

Web-based, automated management and monitoring

SAS Grid Manager includes Platform RTM for SAS, a Web-based tool for monitoring and management of resources, users and jobs running in the grid. Platform RTM also serves as an interface to configure and manage high-availability services, and to define alerts to be sent when thresholds have been exceeded.

Scale out to run more complex analytics faster and take advantage of all computing resources

By making grid computing an automatic capability within multiple applications, SAS helps organizations reduce processing times so larger volumes of data can be integrated, cleansed and analyzed more quickly. SAS Grid Computing can schedule a wide variety of SAS jobs across grid environments for optimal resource utilization and faster processing. Individual SAS jobs can be divided into subtasks that are then executed in parallel to accelerate processing and increase throughput of workload. 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

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.

High availability

- Included with SAS Grid Manager as part of the Platform Suite for SAS.
- Provides high-availability capabilities for critical SAS services such as the SAS Metadata Server.
- Grid node serves as a hot-standby machine for failover.
- Available on all operating systems supported by SAS Grid Manager.
- Detects hardware and software failures in the grid and recovers appropriately.
- Ensures that SAS jobs will be completed optimally.
- When used with the SAS checkpoint/restart feature, jobs can be automatically restarted from the last successful checkpoint.
- Simplifies administration of SAS environment through centralized policies.

Grid-enabled SAS®

- SAS Data Integration Studio and SAS Enterprise Miner are automatically tailored 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.
- SAS Code Analyzer analyzes 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.
- Many SAS solutions and user-written programs can be easily configured for submission to a grid of shared resources.
- Grid-enabled SAS logic can be saved as SAS Stored Processes, which can then be used by all SAS Business Intelligence clients and analytic applications.
- Provides greater resilience for mission-critical applications and high availability for the SAS environment.
- Includes command-line batch submission utility called SASGSUB:
 - Allows you to submit and forget, and reconnect later to retrieve results.
- Enables batch submission to leverage checkpoint and automatically restart jobs.

Real-time monitoring and administration

- Platform RTM for SAS allows you to monitor and manage SAS Grid Computing:
 - Web-based tool enables you to monitor and administer multiple SAS grids.
 - Monitor and administer your grids from any location, as long as you have access to a Web browser.
 - Graphs to track resource usage, users and jobs running on the grid.

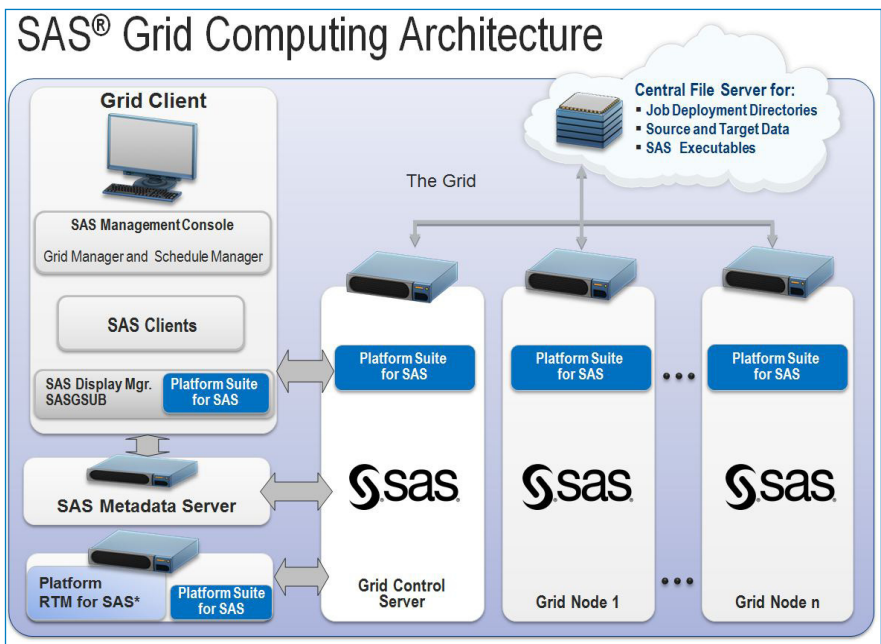
SAS® Grid Computing System Requirements

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

- GUI to modify grid configurations and define alerts when critical thresholds are crossed.
- GUI to configure and manage your critical services for high availability.
- Supports filtering and role definitions to customize 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.
 - Resource and activity information can be filtered, allowing users to focus on specific aspects of grid operations.
 - Provides graphical representation of job distribution.

Flexible infrastructure

- Lets you cost-effectively add computing resources incrementally to meet growing numbers of users, as well as increased business needs.
- Decouples the computing infrastructure from the SAS applications to allow business users to focus on their processes.
- Creates a shared environment to easily and dynamically allocate resources to meet peak needs of different business users over time.
- Eases maintenance of infrastructure by allowing machines to be taken offline without disrupting the business.
- Improves price and performance through the use of commodity hardware.



SAS Grid Computing Architecture