Benefits of Grid Computing for SAS Applications

SAS Forum International

David Smith: dsmith@platform.com
Alan Wong: awong@platform.com
Thierry Ghenassia: tghenassia@platform.com

June 23, 2005
What We’ll Cover

Introduction to Platform Computing

Grid Computing and Business Intelligence
- What is Grid Computing?
- What problems does Grid solve?
- The Case for Grid Enabling Business Intelligence Applications
- Business Intelligence Trends
- Business Intelligence Time vs. Value

Leveraging Grid Computing in BI Applications – A Sample

Platform and SAS Partnership
- Platform and SAS Integrations
- End-to-End SAS Application Environment (Standard and Grid-Enabled)
- Typical Enterprise BI Application Deployment (Standard and Grid-Enabled)
- What are the benefits of the Platform / SAS grid offering?

Grid and Business Intelligence: The Future
Introduction to Platform Computing

Platform – largest and most established grid software vendor

Founded in 1992 and privately held with $60M in revenues and 400+ employees worldwide

1,600+ customers worldwide in electronics, financial services, manufacturing, and government markets

Customers include: AMD, Bristol-Myers Squibb, Daimler-Chrylser, Deutsche Bank, JP Morgan Chase, Pfizer, Texas Instruments, and the US Department of Defense

Partners include: SAS, Dell, HP, IBM, Intel, NEC, Apple and SGI
Platform’s definition of Grid Computing

Characteristics of a computing grid

- Delegates and manages work across multiple hosts
- Supports multiple, vendor-independent applications
- Supports multiple users
- Supports a heterogeneous, commodity based hardware infrastructure
- Automatically allocates and provisions resources based on application requirements and virtualizes these resources from the user
- Automatically monitors, efficiently load balances, and provides reliable execution of workload
What problems does Grid solve?

- Scalability and performance
- High availability and reliability
- Lower TCO
- Compute resource rightsizing
- Enabling formerly infeasible applications

“Doing more, with less”
Today’s Data Center

Silos of Dedicated Server HW
- Mix of Unix, Mainframe, and Windows depending on application type
- No (limited) Server Sharing between applications

Provisioned for individual application Peak Demand
- Usually one Application per “Box” or cluster
- Increased application demand – add more application specific HW

Scale Up Approach
- Large SMP for application and database scalability

Systems Management
- Complex management and provisioning infrastructure

Limited Data Center View of Infrastructure Utilization
- Application or system specific data – difficult to get the big picture view.
Ideal Future State – Grid for Business Applications

Dynamic Priority Based Scheduling
- Standardized x86 based HW and OS (Linux/Windows)
- Transparent resource sharing between applications

Provisioned for Overall Peak Demand
- Applications share the overall resource pool
- Increased application demand – add more generic infrastructure for overall system usage

Scale Out Approach
- Commodity HW with building block approach

Systems Management
- Simplified tools for provisioning, monitoring and management

Data Center Wide View of Infrastructure Utilization
- Enable billing, chargeback and accounting
Analysts, BI industry experts, and the IT community all believe that scalability is crucial for BI …

“More than half of all BI projects are either never completed or fail” for reasons including high cost of ownership, lack of measurable benefits, and a lack of scalability – ADTMag.com

“User scalability of business intelligence (BI) tools and solutions has become paramount.”, Howard J. Dresner, Gartner Note IGG-04212004-03, Apr 2004
BI requires performance, scalability, flexibility, and reliability, each of which are cornerstones of grid technology.

Many companies run multiple BI applications on dedicated computing infrastructure; as such, dedicated resources are sized for peak capacity, resulting in inefficient overall use of resources.

“Most (BI) vendors are still struggling with the scalability issue.” – DMReview.com March 2004
Business Intelligence Trends

Demand for BI applications is increasing steadily due to:

- Unprecedented data growth
  Annual corporate data growth of 50-100% is the norm (SOx Compliance Journal)
- Real time or right time BI
  Integrating BI data into operational applications can be a competitive differentiator
- BI for the masses
  BI now reaches all parts of an organization, from boardrooms to mailrooms
- Search for a ‘Single Version of the Truth’
  Strategic errors may result from business decisions not made on the same data warehouse; ETL processes have become mission critical
- Consolidation and Standardization
  Move to strategic single-sourcing of BI Vendors to maximize efficiencies

While BI applications continue to become more computationally demanding, BI users are exceedingly concerned about controlling and minimizing costs

- Grid Computing provides technology that will bridge this gap
Business Intelligence Time vs. Value

Framework for Right-Time Business Intelligence
(Courtesy: Bolder Technology)
## Leveraging Grid Computing in BI Applications – A Sample

<table>
<thead>
<tr>
<th>Industry</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Retailing / commercial sales</td>
<td>Permits the use of complex algorithms for highly accurate demographic analysis</td>
</tr>
<tr>
<td>CRM / Call center</td>
<td>Predictive analytics for customer profiling</td>
</tr>
<tr>
<td>Supply Chain/JIT Mfg</td>
<td>Real-time monitoring and analysis for early detection of production issues</td>
</tr>
<tr>
<td>Financial Services</td>
<td>Credit risk analysis for credit card fraud</td>
</tr>
<tr>
<td>Insurance</td>
<td>Complex risk analysis for risk scoring</td>
</tr>
<tr>
<td>Investment/Brokerage</td>
<td>Analysis of and prediction regarding changing market conditions</td>
</tr>
<tr>
<td>Homeland security</td>
<td>Threat identification: Real time identification of high risk candidates</td>
</tr>
<tr>
<td>Life Sciences</td>
<td>Specialized research – dealing with enormous quantities of data and short reporting windows</td>
</tr>
<tr>
<td>Compliance / Regulatory</td>
<td>Sarbanes Oxley, Basel II, etc. are requiring more data be included in risk analyses</td>
</tr>
</tbody>
</table>
Platform and SAS Integrations

Platform recognized early on the benefits that grid could bring to Business Intelligence applications. Platform has had great success with its SAS integration / partnership.

Current Integration: Platform JobScheduler for SAS

- Enabled the automated scheduling and management of SAS jobs in a virtualized, heterogeneous infrastructure environment with enterprise scalability
- Currently integrated with SAS ETL Server, SAS BI Server, and SAS Marketing Automation applications
- Plans are in place for broad deployment across the SAS product line

Upcoming integration: Platform grid enables SAS applications

- Fully automated parallel workload processing of SAS applications across a Platform computing grid.
- Initial launch with ETL Server and Enterprise Miner. Additional applications will follow.
Sample End-to-End SAS Application Environment

ETL
- Packaged
- Legacy
- Other

Analysis
- Data Mart / Data Warehouse
- Cubing (OLAP)
- Mining (Power User) / Modeling (Analyst)

Reporting
- Analytic Reporting
- Enterprise Reporting
- Portal / Groupware
- Business Reporting

People
- Power User (producer)
- Casual User (consumer)
- Business User (dual)

Hardware
- Dedicated H/W
- Dedicated H/W
- Dedicated H/W
Grid Enabled SAS Application Environment

ETL
- Packaged
- Legacy
- Other

Analysis
- Data Mart / Data Warehouse
- Cubing (OLAP)
- Mining (Power User) / Modeling (Analyst)

Reporting
- Analytic Reporting
- Enterprise Reporting
- Portal / Groupware
- Business Reporting

Power User (producer)
Casual User (consumer)
Business User (dual)

SAS/CONNECT

Shared H/W
Typical Enterprise Business Intelligence Application Deployment

ETL

Mainframe Data

CRM Data

Sales Data

Transform

Dedicated H/W

Data Warehouse

Transform

Data Mart

Dedicated H/W

Analytical

Mining / Analysis

Technical Reporting

CRM Data

Sales Data

Data Mining / Analysis

Dedicated H/W

Business Reporting
What are the benefits of the Platform / SAS grid offering?

- Scalability
- Performance
- High Availability
- Accuracy / Consistency
- New Analytic Opportunities
- Fully integrated and supported
- Full automation and management
- Builds on a strong 4 year relationship
- Reduced TCO / Infrastructure Optimization
The majority of the industry’s major Business Intelligence vendors and solution providers are using or planning on deploying grid enabled BI solutions.

75% of companies are implementing, or planning to implement, grid computing architectures and 35% of these deployments are targeted at BI projects. – Evans Data (Sept. 2004)

SAS and Platform Computing will be the first to make Grid Computing a reality for commercially available BI products.
Thank you.