1.1 What is Grid?

1.2 SAS Grid Architecture

1.3 How to run SAS code on the Grid

1.4 Grid Option Sets

1.5 Monitoring
What is SAS Grid
WHAT IS GRID?

- A SAS grid computing environment is one in which SAS computing tasks are distributed among multiple computers on a network, all under the control of SAS Grid Manager.
**SAS® GRID MANAGER**

**KEY CAPABILITIES**

**Workload Management**
- Provides job, host & user management
- Prioritize & schedule jobs using rules-based queues
- Identify, allocate and manage resources

**High Availability**
- Detect grid failure and recovers automatically
- Automatic restart of jobs from last successful checkpoint.

**Performance**
- Increase throughput of all SAS jobs
- Jobs are divided into subtasks for parallel execution
- Integrate and analyze large volumes of data

**Grid-Enabled SAS Products and Solutions**
WHAT IS GRID?  HIGH-LEVEL OVERVIEW

A  B  C  D

SAS Grid Manager
Control Server

SAS Grid Manager
Nodes
## MULTIPLE WAYS TO LEVERAGE GRID

<table>
<thead>
<tr>
<th>Products</th>
<th>Workload Management</th>
<th>Enterprise Scheduling</th>
<th>Parallelized Processing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Any SAS programs</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>SAS Data Integration</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>SAS Web Report Studio</td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>SAS Marketing Automation</td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>SAS Marketing Optimization</td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>SAS Enterprise Miner</td>
<td>✓</td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>SAS Enterprise Guide</td>
<td>✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SAS Add-in for Microsoft Office</td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>SAS Workspace Server</td>
<td>✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SAS Stored Process Server</td>
<td>✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SAS OLAP Server</td>
<td>✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SAS Stored Processes</td>
<td>✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SAS Risk Dimensions</td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>JMP Genomics</td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>SAS Demand Forecasting for Retail</td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>SAS for Customer Experience Analytics</td>
<td></td>
<td></td>
<td>✓</td>
</tr>
</tbody>
</table>
GRID ARCHITECTURE

ARCHITECTURE – CONCEPTUAL VIEW

Grid Node 1 (Grid Control Server)

Grid Node 2

Grid Node 3 (Metadata Server)

Shared/Clustered File System
GRID ARCHITECTURE

ARCHITECTURE – CONCEPTUAL VIEW

Grid Node 1
(Grid Control Server)

Grid Node 2

Grid Node 3
(Metadata Server)

Grid Node 4

Shared/Clustered File System

Grid Node 5
GRID ARCHITECTURE

GRID ARCHITECTURE – CONCEPTUAL VIEW

Grid Node 1
(Grid Control Server)

Grid Node 2

Grid Node 3
(Metadata Server)

Grid Node 4
(Maintenance)

Shared/Clustered File System

Grid Node 5
**SAS® GRID MANAGER**

**WORKLOAD MANAGEMENT**

**Queuing for load balancing**

- **Normal Queue**
- **High Priority Queue**
- **Overnight Queue**
- **Testing Queue**
- **Development Queue**

**Job slots to match capability**

- 10 concurrent jobs
- 10 concurrent jobs
- 3 concurrent jobs

**Utilization Thresholds**

- CPU < 80%
- Mem < 90%
- CPU < 75%
- I/O < 75%
- CPU < 50%
- Mem < 75%
- I/O < 80%

**Advanced job scheduling**

- When file ABCD becomes available, launch the reconciliation process
- Execute the process on the first of each month starting at 10PM
- When process XXXX complete successfully, launch process YYYYY
How to run SAS code on the Grid

Base SAS
SAS Management Console
SAS Enterprise Guide
SAS Enterprise Miner
SAS Data Integration Studio
Batch Submission
libname bin '/sasdata/sasss1/saslibrary';
title 'K';
proc means data=bin.large(where=(class='K'));
run;

proc means data=bin.large(where=(class='B'));
run;

proc means data=bin.large(where=(class='A'));
run;

title 'LARGE';
proc means data=bin.LARGE;
run;
Re-engineered code to run sections of code in parallel

```sas
options metaserver='precious';
options metaport=8561;
options metauser=sasss1;
options metapass='8791C14621D066064A494262536F57B7';

%let sasappsvr=SASAppGrid;
%let jovar=queue=priority;
%let gerc=%sysfunc(grdsvc_enable(_all_, resource=&sasappsvr; jobopts=jovar));
%put;%put Grid enable return code=&gerc (0 indicates grid was enabled);

%let nnodes=%sysfunc(grdsvc_nnodes(resource=&sasappsvr));
%put;%put Number of grid nodes=&nnodes;

signon grid1;
rsubmit grid1 wait=no;
libname bin '/sasdata/sasss1/saslibrary';
title 'K';
proc means data=bin.large(where=(class='K'));
run;
endrsubmit;
```
Re-engineered code to run sections of code in parallel

```sas
signon grid8;
rsubmit grid8 wait=no;
libname bin '/sasdata/sasss1/saslibrary';
title 'LARGE';
proc freq data=bin.large;
  table class;
run;
endrsubmit;

waitfor grid1 grid2 grid3 grid4 grid5 grid6 grid7 grid8;
rget grid1;
rget grid2;
rget grid3;
rget grid4;
rget grid5;
rget grid6;
rget grid7;
rget grid8;

signoff grid1;
signoff grid2;
signoff grid3;
signoff grid4;
signoff grid5;
signoff grid6;
signoff grid7;
signoff grid8;
```
Re-engineered code to run sections of code in parallel

90% improvement in Run Time

NOTE: SAS Institute Inc., SAS Campus Drive, Cary, NC USA 27513-2414

NOTE: The SAS System used:

real time 9:18.20
cpu time 4:06.20

NOTE: Remote signoff from GRID8 complete.
NOTE: SAS Institute Inc., SAS Campus Drive, Cary, NC USA 27513-2414
NOTE: The SAS System used:

real time 54.70 seconds
cpu time 0.47 seconds
HOW TO RUN SAS CODE ON THE SAS GRID

SAS® GRID MANAGER

SAS® Enterprise Guide
HOW TO RUN SAS CODE ON THE SAS GRID

SAS® GRID MANAGER

SAS® Enterprise Guide
**ENTERPRISE GUIDE**  **ANALYZE LEGACY CODE**

- Analyze legacy code to generate Grid enable code

```plaintext
Proc sort data=ORSTAR.ORSTARBASE out=EGout
   by Product_Category;
run;

Proc means data=EGout.ORCat
   noprint chartype nolabels nway sum nonobs;
   var Total_Retail_Price;
   class Product_Category / order=unformatted ascending;
   output out=EGout.CATEGORYSTATS(label="Summary by Product Category")
      sum= / autoname autolabel ways inherit;
run;
```
• Option to export grid enable code for batch job execution
### Distribution analysis of: MPG_Highway

Variable: MPG_Highway (MPG [Highway])

#### Basic Statistical Measures

<table>
<thead>
<tr>
<th>Location</th>
<th>Variability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>26.84346</td>
</tr>
<tr>
<td>Median</td>
<td>26.00000</td>
</tr>
<tr>
<td>Mode</td>
<td>26.00000</td>
</tr>
<tr>
<td>Std Deviation</td>
<td>5.7412007</td>
</tr>
<tr>
<td>Variance</td>
<td>32.961386</td>
</tr>
<tr>
<td>Range</td>
<td>54</td>
</tr>
<tr>
<td>Interquartile Range</td>
<td>5</td>
</tr>
</tbody>
</table>

#### Basic Confidence Limits Assuming Normality

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>95% Confidence Limits</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>26.8434579</td>
<td>[26.29799951, 27.388916]</td>
</tr>
<tr>
<td>Std Deviation</td>
<td>5.74120072</td>
<td>[5.38062831, 6.1539659]</td>
</tr>
<tr>
<td>Variance</td>
<td>32.9613857</td>
<td>[28.95101601, 37.871296]</td>
</tr>
</tbody>
</table>

#### Tests for Location: Mu=0

<table>
<thead>
<tr>
<th>Test</th>
<th>Statistic</th>
<th>p Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Student's t</td>
<td>96.72920413</td>
<td>0.0000</td>
</tr>
<tr>
<td>Sign</td>
<td>214</td>
<td>0.0000</td>
</tr>
<tr>
<td>Signed Rank</td>
<td>45903</td>
<td>0.0000</td>
</tr>
</tbody>
</table>

Generated by SAS ('SASAppGrid', Linux) on August 30, 2012 at 7:05:48 PM
SAS® GRID MANAGER

HOW TO RUN SAS CODE ON THE SAS GRID

SAS® Management Console – SAS® Enterprise Miner Setup

SAS® Grid Manager
SAS® GRID MANAGER

SAS® Data Integration Studio

HOW TO RUN SAS CODE ON THE SAS GRID
SAS® GRID MANAGER

SAS® Data Integration Studio

HOW TO RUN SAS CODE ON THE SAS GRID
SAS® GRID MANAGER

HOW TO RUN SAS CODE ON THE SAS GRID

SAS® Data Integration Studio
• SASGSUB Batch Submit

  sasgsub -gridsubmitpgm MyProgram.sas
             -gridwaitresults
             -gridwaittimeout 3600

  rc=$?
Grid option sets
## GRID OPTIONS IN 9.3 METADATA

**SASAppEM - Grid Server Properties**

<table>
<thead>
<tr>
<th>General</th>
<th>Options</th>
<th>Notes</th>
<th>Extended Attributes</th>
<th>Authorization</th>
</tr>
</thead>
<tbody>
<tr>
<td>Provider:</td>
<td>Platform</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grid Command:</td>
<td>/sas93/config/compute/Lev1/SASApp/GridServer/sasgrid</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Optional**

- **Workload:**
- **Module Name:**
- **Additional Options:** `queue=eMiner; project=Customer Churn; priority=50`

- **Do not require SAS Application Server name as a grid resource.**
SAS 9.3 : MULTIPLE GRID SERVER DEFINITIONS

Server Manager
- SASMeta
- SASApp
- SASAppGrid
  - SASAppGrid - Logical Grid Server
  - SASAppGrid - Grid Server
  - SASAppGrid - Logical Workspace Server
- SASDI
  - SASDI - Logical Grid Server
  - SASDI - Grid Server
  - SASDI - Logical Workspace Server
- SASEG
  - SASEG - Logical Grid Server
  - SASEG - Grid Server
  - SASEG - Logical Workspace Server
- SASEM
  - SASEM - Logical Grid Server
  - SASEM - Grid Server
  - SASEM - Logical Workspace Server

Optional
- Workload:
- Module Name:
- Additional Options: queue=DI

Optional
- Workload:
- Module Name:
- Additional Options: queue=EG

Optional
- Workload:
- Module Name:
- Additional Options: queue=EM
SAS 9.3 : MULTIPLE USERS WITH MULTIPLE GRID CLIENTS
MANAGING USERS AND GRID APPLICATIONS IN 9.4

Using Grid Options Sets
GRID OPTIONS SETS

- Grouping of options commonly used by grid
GRID OPTIONS SETS

• Mapped in a grid server to a grid application for a user/group
• Can be mapped multiple times
• Override default options
MULTIPLE OPTIONS ARE MERGED

Default Values
- SAS Options: \(-\text{memsize}\ 256\text{M} -\text{DBCS}\)
- Required Resources: SASApp
- Grid Options: queue=normal

Options Set Values
- SAS Options: \(-\text{memsize}\ 0\)
- Required Resources: Teradata
- Grid Options: queue=priority

Merged Option Values
- SAS Options: \(-\text{memsize}\ 256\text{M} -\text{DBCS} -\text{memsize}\ 0\)
- Required Resources: SASApp Teradata
- Grid Options: queue=normal queue=priority

Resulting Option Values
- SAS Options: \(-\text{DBCS} -\text{memsize}\ 0\)
- Required Resources: SASApp Teradata
- Grid Options: queue=priority
MONITORING  SAS PROGRAM JOB STATUS ON GRID

• Check job status via RTM or SAS Management Console Grid Manager Plug-in
MONITORING SAS PROGRAM JOB STATUS ON GRID

- Check job status via RTM or SAS Management Console Grid Manager Plug-in
THANK YOU