



Introduction to SAS Grid and some considerations

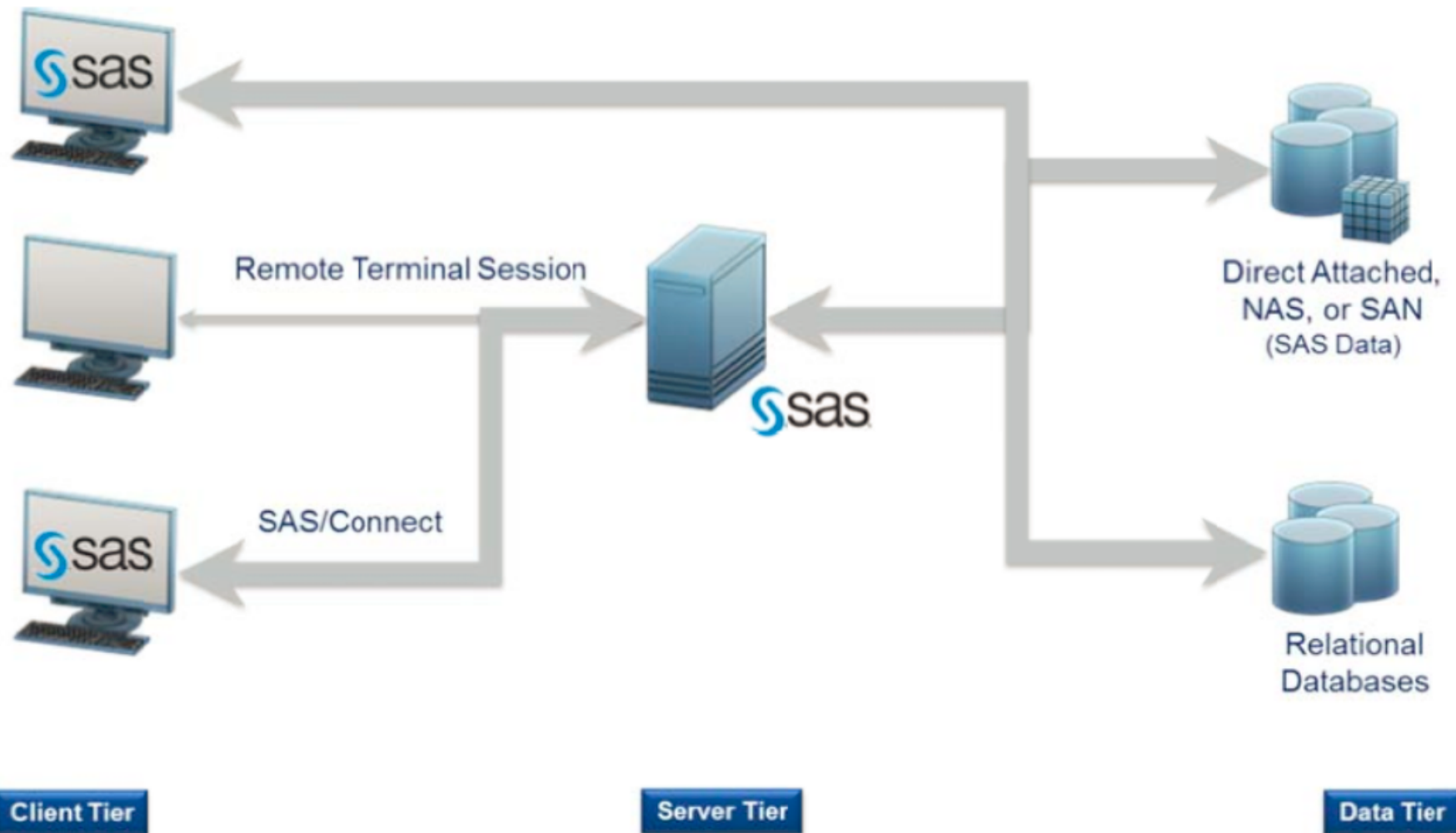
Agenda

- ▶ **1 General Introduction of SAS solutions**
- 2 The benefits of moving to SAS on the Grid**
- 3 Potential problems need to be considered**
- 4 Different ways of Accessing SAS on the Grid**
- 5 Guidance on how to best utilize the power of SAS Grid**

PC SAS

- SAS is installed on user's personal computer directly.
- Easy for users to understand as all things such as data, program, output etc. are on the same machine. No need to connect to a Server.
- CPU speed is slow. The disk space, memory are all very limited.
- Not easy to manipulate big data.

SAS Unix Server



SAS Unix Server

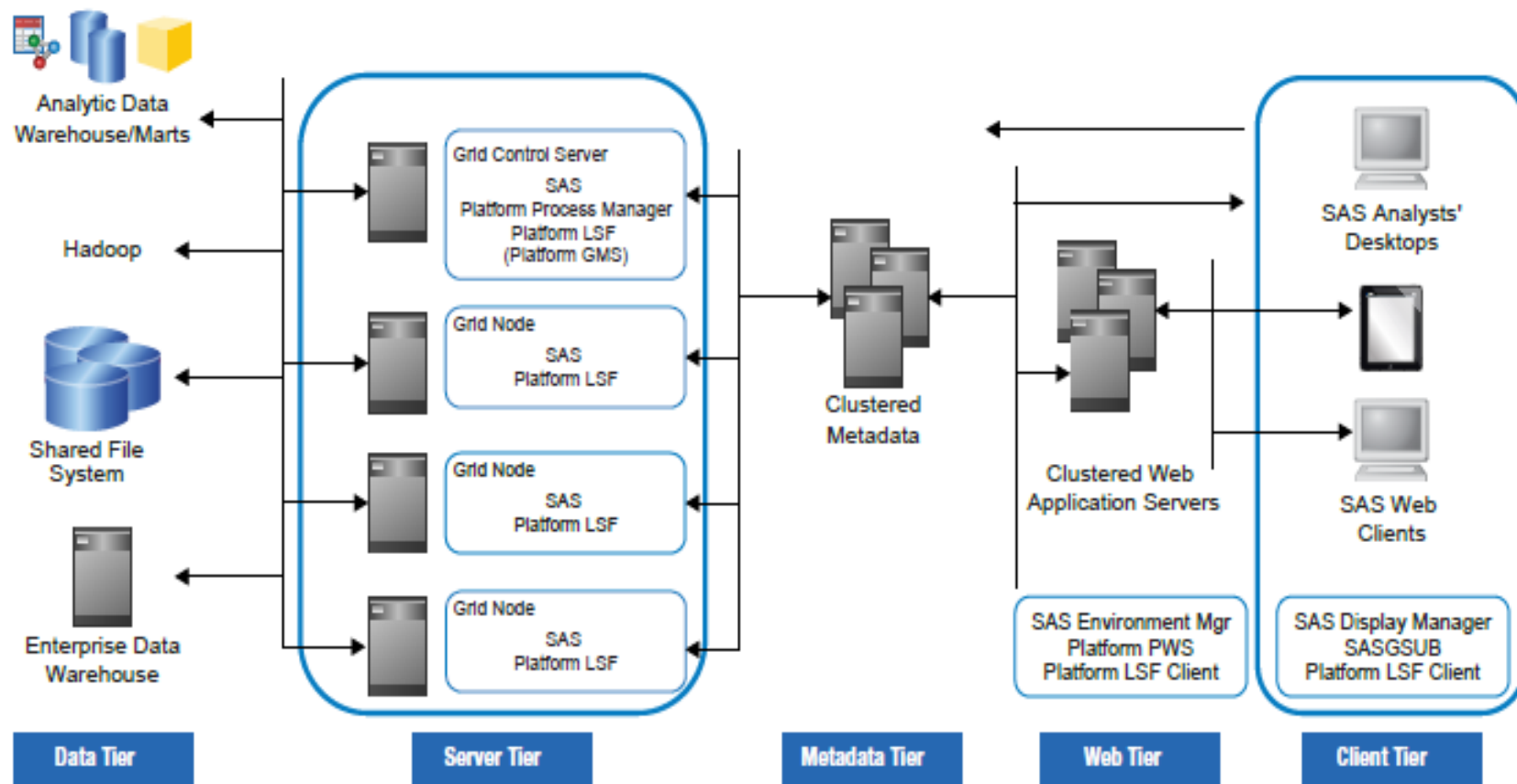
- SAS is installed on a super computer with Unix operating system like AIX, HP Unix etc.
- Often database like Oracle, DB2 etc. are also installed on the same server so that SAS can access data easily.
- Big data is no longer an issue. User can operate on data with hundreds of millions of observations. For relatively smaller data. The execution time is neglectable.

Note: *Highly inefficient code like full join two very large datasets can still slow down the whole server if not have it crashed.*

SAS Unix Server

- Once a SAS Unix Server is deployed, the speed of CPU is settled. Also it is not convenient to improve disk space or add more memory.
- System manager need to evaluate carefully when deploying SAS Unix Server to strike a balance between cost and capacity.
- Downtime is needed to maintain the server. During maintenance period users cannot work on the Server.

Grid Architecture Diagram - Platform Suite for SAS Shared



SAS Grid Topology

- Grid Control Server

- Generally Grid Control Server is one of the Grid Nodes but with Platform Web Services (PWS) installed so this computer can receive SAS job from clients and distribute the jobs to all Grid Nodes based on workload of each Node. Grid Control Server itself can also execute SAS programs. However, it is recommended to keep this computer as a process manager only to ensure efficiency.

- Grid Nodes

- Grid Nodes are a cluster of peer computers which can receive SAS job distributed to it and return result to requester once program is completed.

SAS Grid Topology

- Shared file system

- This component is used to store data and shared directories for jobs that run on the Grid.

- Load Sharing Facility (LSF)

- This facility dispatches all jobs submitted to it and returns the status of each job. LSF also manages any resource requirements and performs load balancing across machines in a grid environment. A server version of Platform LSF must be on the Grid Control Server and each Grid Node.

- Grid Clients

- Through Grid Clients users are able to submit SAS programs to SAS on the Grid for execution. Grid Clients can be PC SAS, SAS EG, line interactive mode etc.

Agenda

- 
- 1 General Introduction of SAS solutions
 - 2 The benefits of moving to SAS on the Grid
 - 3 Potential problems need to be considered
 - 4 Different ways of Accessing SAS on the Grid
 - 5 Guidance on how to best utilize the power of SAS Grid

Benefits of SAS Grid

- **Better performance**

- First and foremost, SAS on the Grid provides users with a better performance. In particular, the server is able to break down large program into smaller sub tasks and distribute them to different SAS Grid nodes to run all sub tasks in a parallel mode and finish the whole program faster.

- **Scalable**

- If the performance of server is slow due to data or user increases, system administrator can increase the capacity by adding servers to SAS Grid.

- **More Stable**

- As the SAS on the Grid is actually a cluster of peer computers and programs are equally distributed to these computers with no preference, failing on one node is totally not a problem.

Agenda

- 1 General Introduction of SAS solutions
- 2 The benefits of moving to SAS on the Grid
- ▶ 3 Potential problems need to be considered
- 4 Different ways of Accessing SAS on the Grid
- 5 Guidance on how to best utilize the power of SAS Grid

Potential problems

- network speed is one of the key factor that impact SAS Grid performance
- DDE (Dynamic Data Exchange) is not available on SAS on the Grid.
- X command is different.
- SAS environment will be different so there may be potential issues.

Agenda

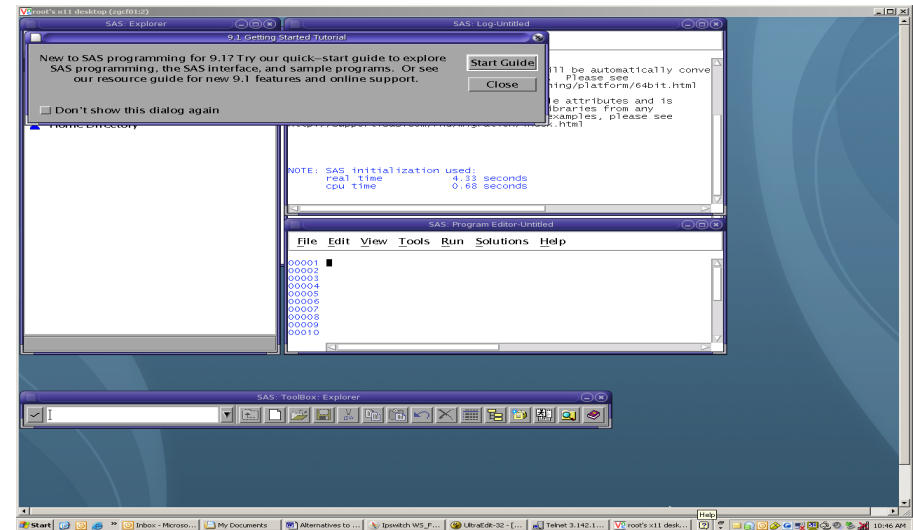
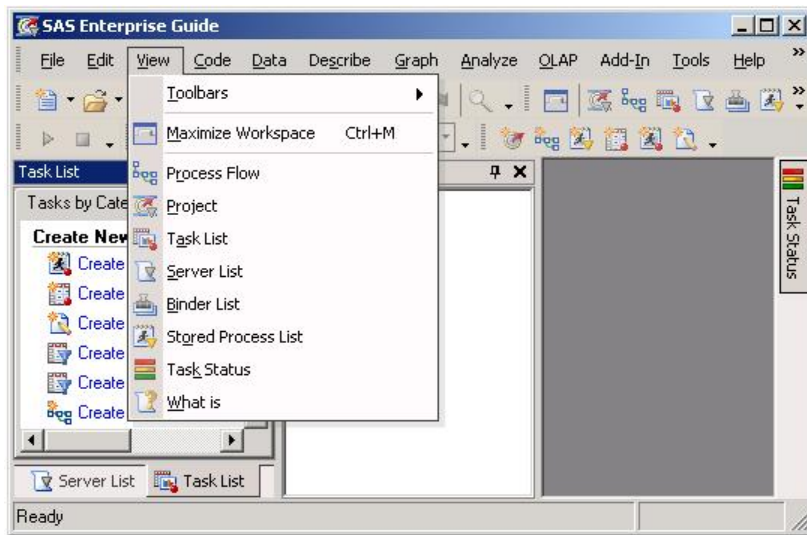
- 1 General Introduction of SAS solutions
- 2 The benefits of moving to SAS on the Grid
- 3 Potential problems need to be considered
- ▶ 4 Different ways of Accessing SAS on the Grid
- 5 Guidance on how to best utilize the power of SAS Grid

Ways of accessing SAS Grid

- SAS sign on: User has SAS installed on their PC. They can work on local SAS directly or submit program to SAS on the Grid via SAS sign on.
- SAS line interactive mode

Ways of accessing SAS Grid

- SAS EG(Enterprise Guide)
- X Window



Agenda

- 1 General Introduction of SAS solutions
- 2 The benefits of moving to SAS on the Grid
- 3 Potential problems need to be considered
- 4 Different ways of Accessing SAS on the Grid
- ▶ 5 Guidance on how to best utilize the power of SAS Grid

Guidance on how to use SAS Grid

- Schedule job is easy to use. Try schedule job if a program need to be regularly executed and take a long time to run.
- When operating via client window, minimize streaming output.
- Batch submit code to SAS on the Grid via line interactive mode when it takes a long time to execute the code.

Any Questions

Stay tuned for updates

