SAS® FORUM
UNITED KINGDOM 2015

Approaching a Platform Migration

Approaches to SAS migration and Platform LSF considerations for SAS/Grid

Topics

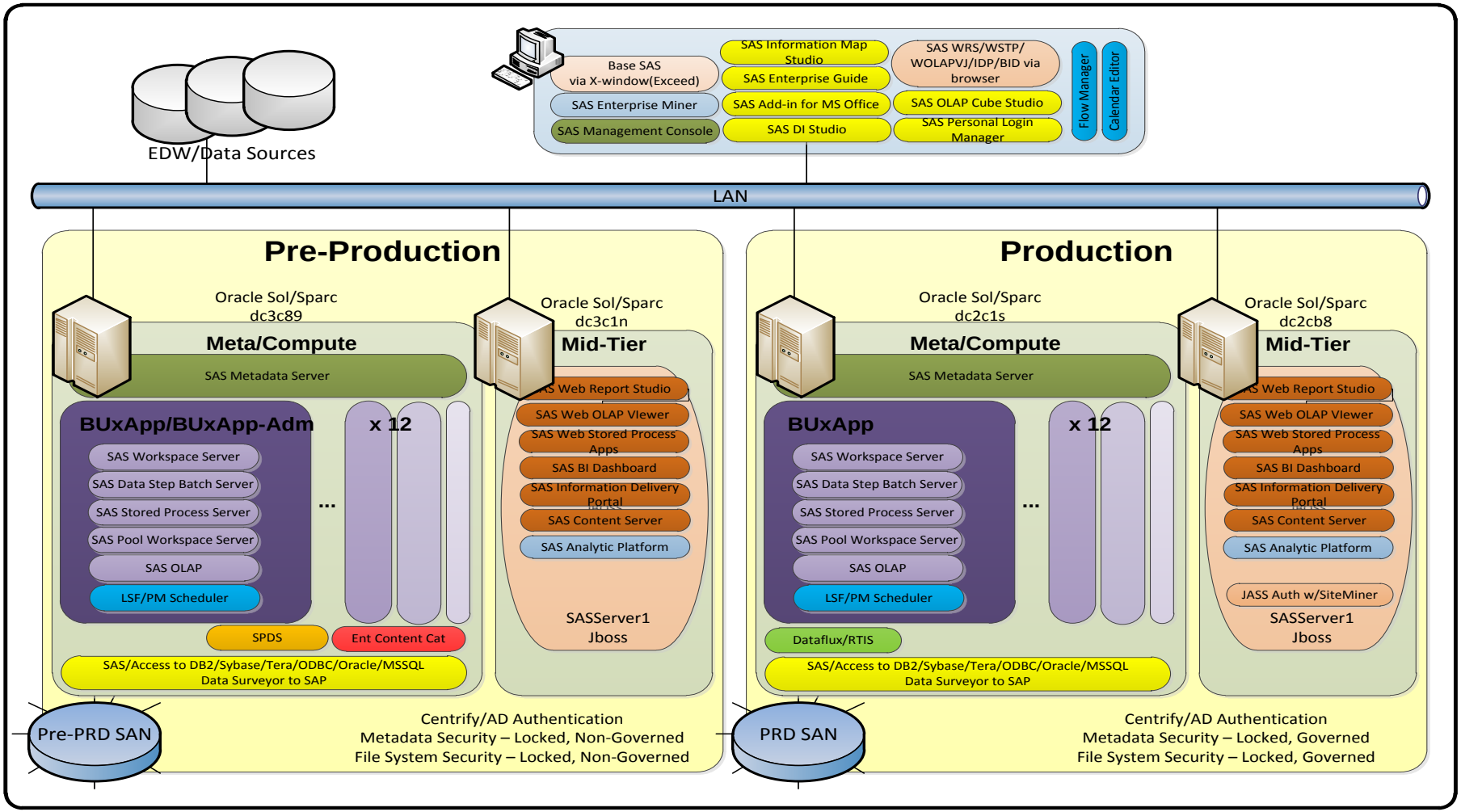
- Scope
- Architecture
- Migration
- Questions
- Grid considerations
- Questions

Bell Canada SAS migration

- 24 months
- 16 Business units
- 50 Developers
- 200 SAS analysts
- 600 Enterprise Guide users
- 4000 Job Flows
- 4000 Web service users
- 96 cpu and 44 Tb disk

Old architecture

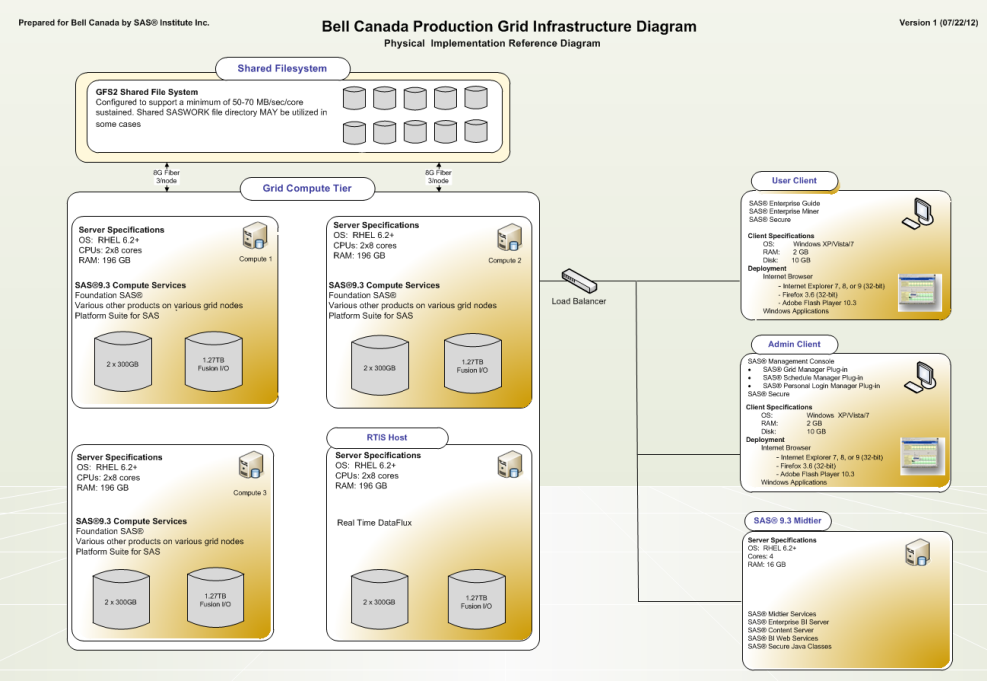
- Two maxed-out Solaris M5000
- Compute and Metadata on both
- Solaris MidTier
- SAS v92
- Platform LSF and Process Manager (not Grid)



New architecture

- Red Hat Enterprise Linux
- Dev, QAT and Prod hardware segregation
- SASv9.4 with SAS/Grid
- Platform LSF 9.1 and Process Manager with Group admin capability

Simpler version



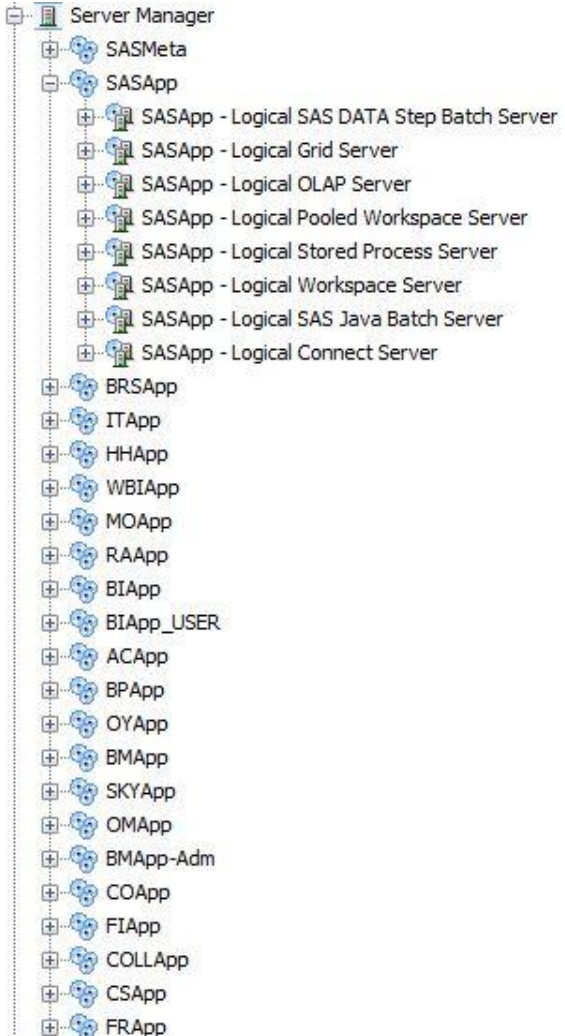
Migration approaches

- Slow re-engineering
- Outsource
- Self service
- Big Bang

‘Instance’ as unit of migration

Instance

- Set of logical servers
- Unix mount point
- Metadata group
- AD group



Slow re-engineering

- Replicate Prod to separate hardware (QAT)
- Multiple instances
- 6 months transition from old Prod to QAT
- Extensive changes and testing
- 2 month cutover from QAT to new Prod

Outsource

- Like-for-like replication
- Fixed price over 4-6 weeks
- Manual keyboard entry
- Offshore providers following standard template

Self service

- Full access to separate instance on Dev
- Limited access to separate instance on Prod
- 3 months gradual cutover
- Careful user management

Big bang

- Full access to separate instance on Dev
- 6 months functional testing in Dev
- Two months test load in Prod
- Cut-over on a long weekend
- No back out after first 48 hours

Management

- Weekly Governance Committee – sponsor, business primes
- Daily activity call – project manager, tech leads
- Adhoc strategy planning – tech leads, architects

Questions?

- Instances – Logical servers, separate file system, Groups
- Slow re-engineering – migration tool failures, manual verification using DIS, DeployedJobs and JobFlows
- Outsourcing – working within resource limitations, accommodating environment differences
- Self-service – co-ordinating stakeholders
- Big bang – risk vs benefit

Migration to SAS/Grid

- Load balancing
- Governance
- Production hardening
- Delegation and segregation
- Monitoring
- Single node services

Load balancing

- Protecting critical services
- Application specifics
- LSF queues
- Slots
- RTM web interface

Governance

- Allocation of shared resources
- Conflicts and contention
- Delegation of administrative authority
- Dynamic control
- Operational rules

Production hardening

- Change management
- Configuration
- Service guarantees
- Security
- Failover and recovery

Delegation and segregation

- Metadata Roles
- Metadata ACTs
- Internal accounts
- Unix ACLs
- Active Directory Groups + Centrify

Monitoring

- RTM – User and admin access
- LSF commands – bjobs, bstatus, bhist, jhist, lsload
- Unix monitoring tools – top, nmon, Vantage
- SAS MC Schedule Manager
- Platform Flow Manager

Single node services

- SAS/Share and SAS/IntrNet
- Outgoing jdbc
- Incoming sftp
- Xcmd limitations
- DataFlux
- Hadoop

Questions?

- Load balancing
- Governance
- Production hardening
- Delegation and segregation
- Monitoring
- Single node services

Protecting critical services

- Metadata server(s)
- DataFlux
- Grid controller(s)
- SAS/Share
- SAS/IntrNet

Application specifics

- Enterprise Guide – report consumers, analysts, developers
- DI Studio – Jobs, Flows
- Process Manager – Flows, calendars

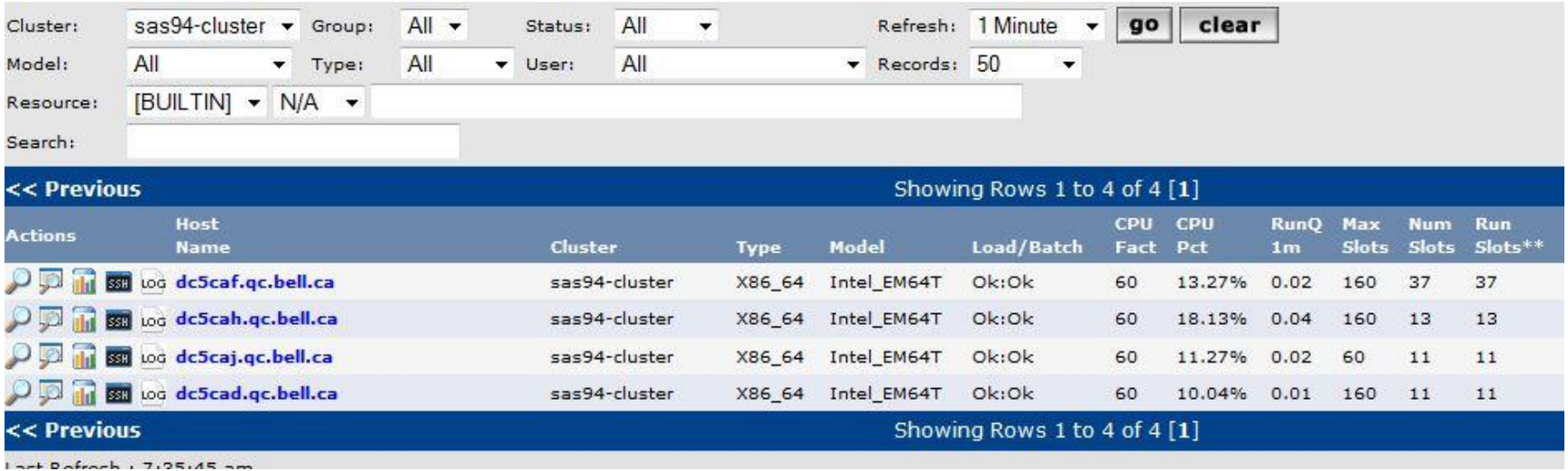
LSF Queues

- Production
- Priority
- Normal
- Express

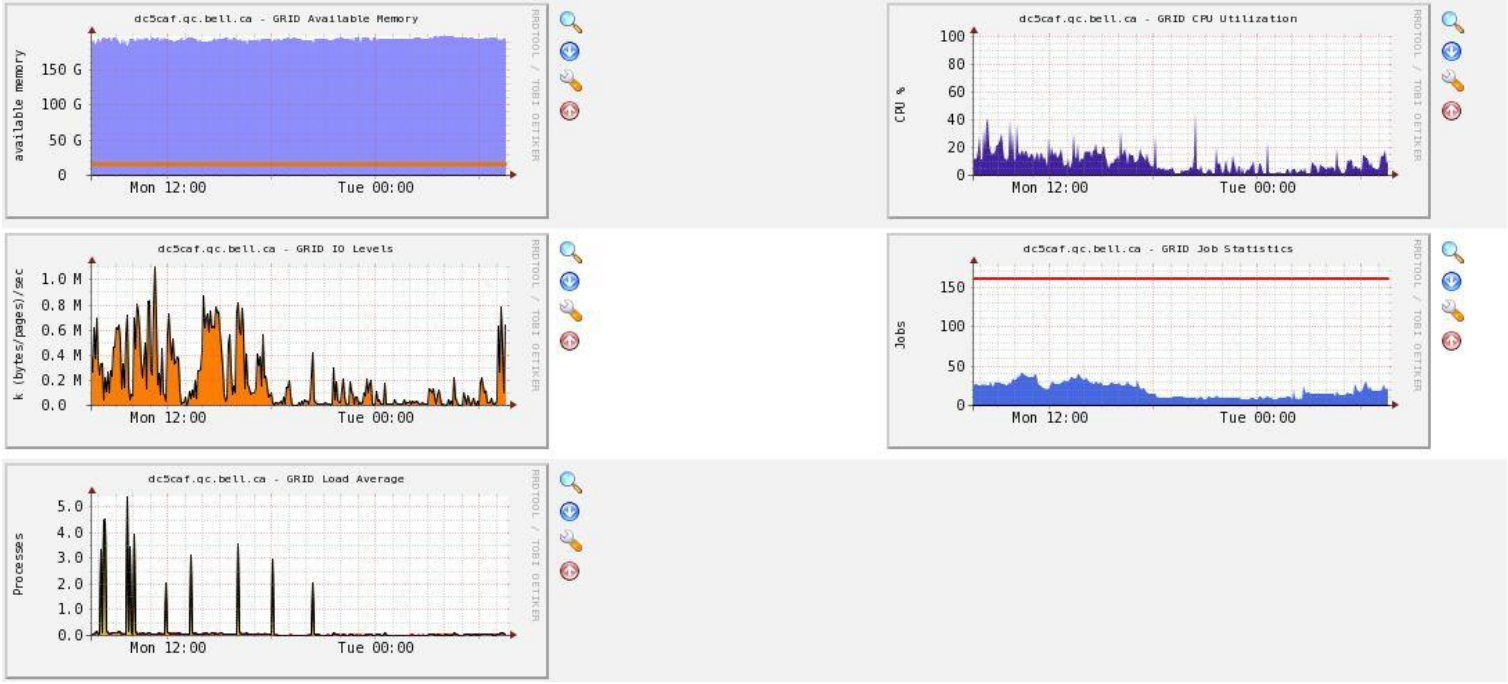
Slots

- 8 per cpu default
- More for ETL
- Less for EG

RTM web interface



RTM web interface



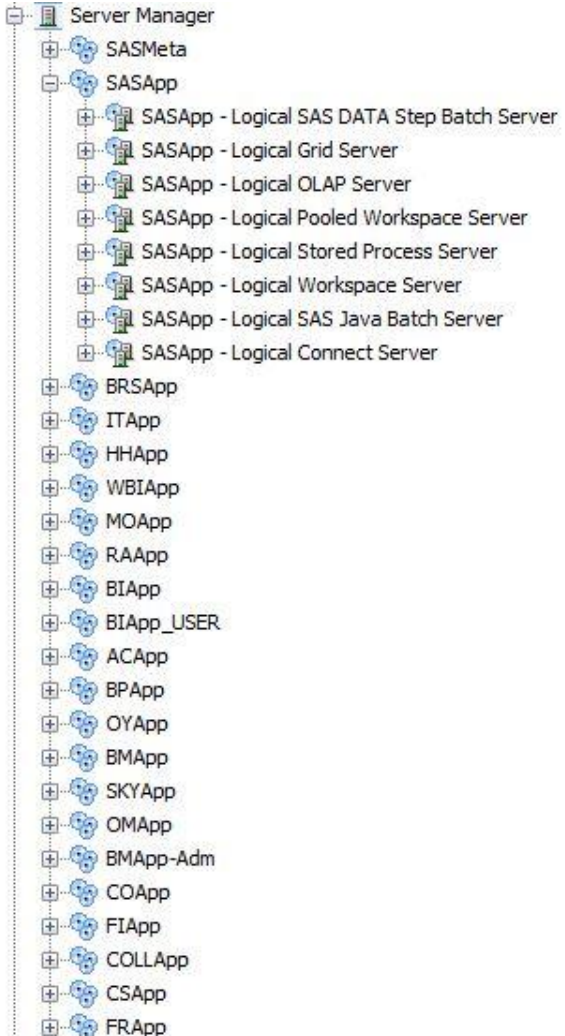
Contact

andrew.farrer@bell.ca

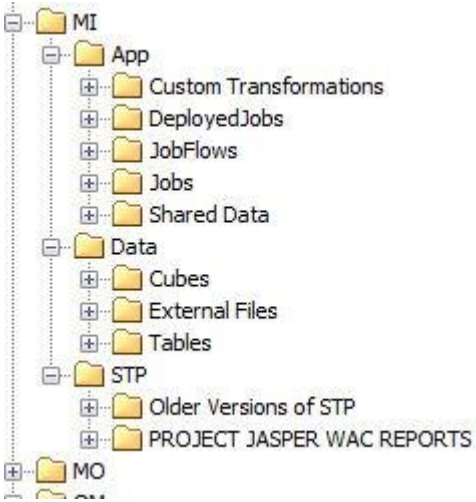
acfarrer@gmail.com

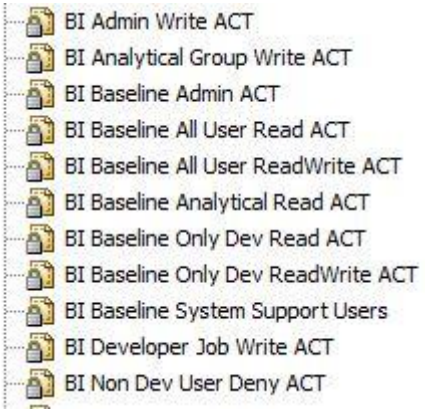
Instance

- Set of logical servers
- Unix mount point
- Metadata group
- AD group



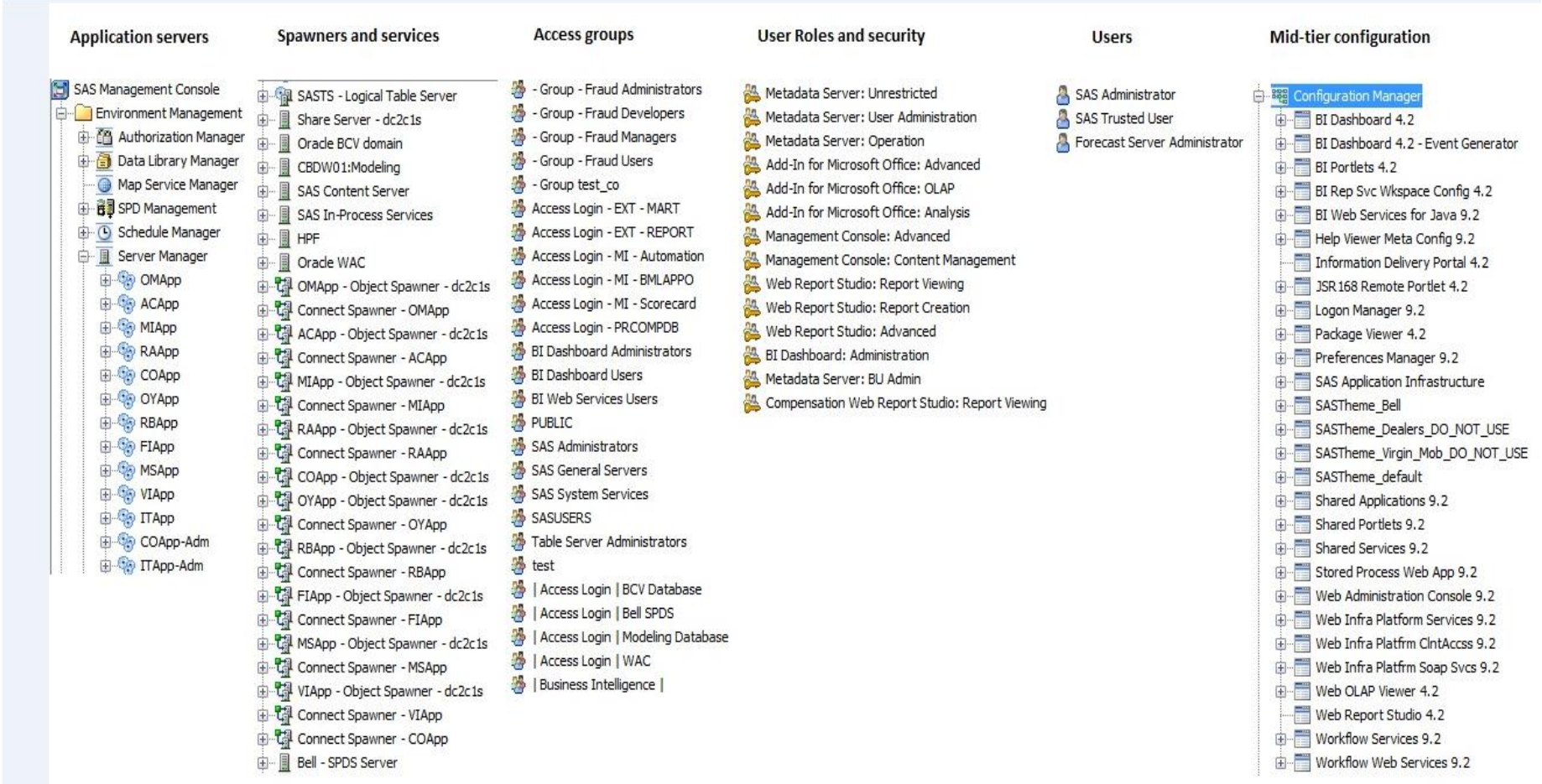
Metadata objects



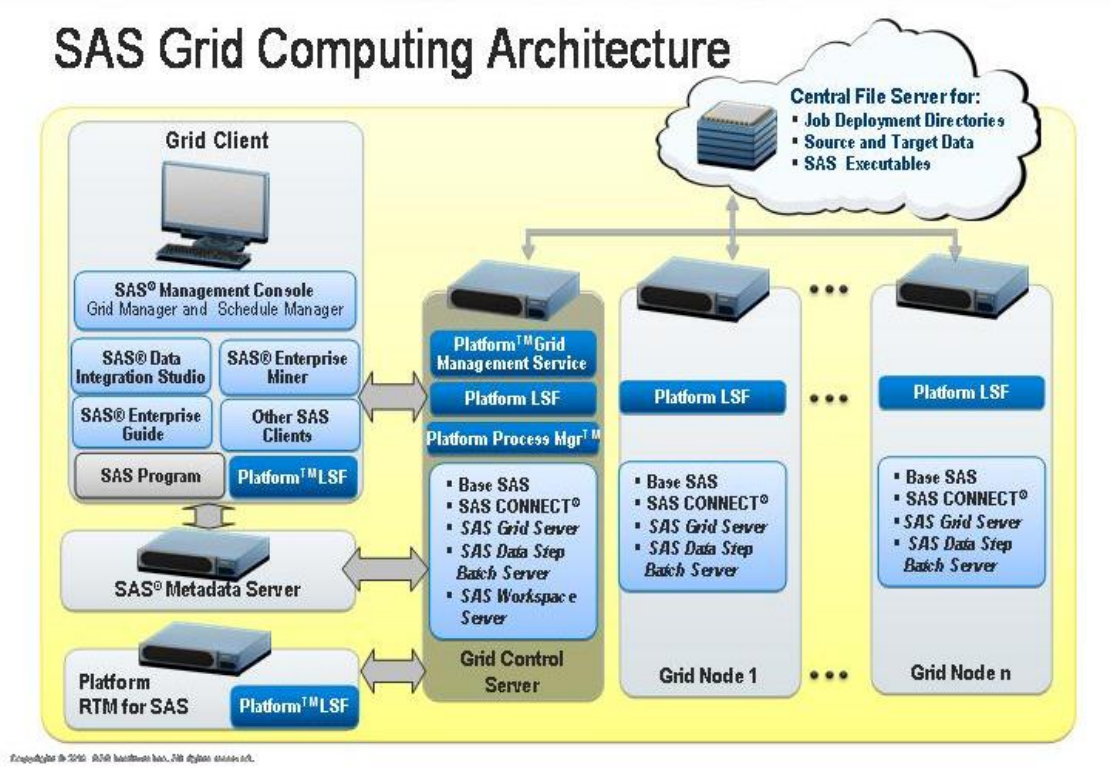


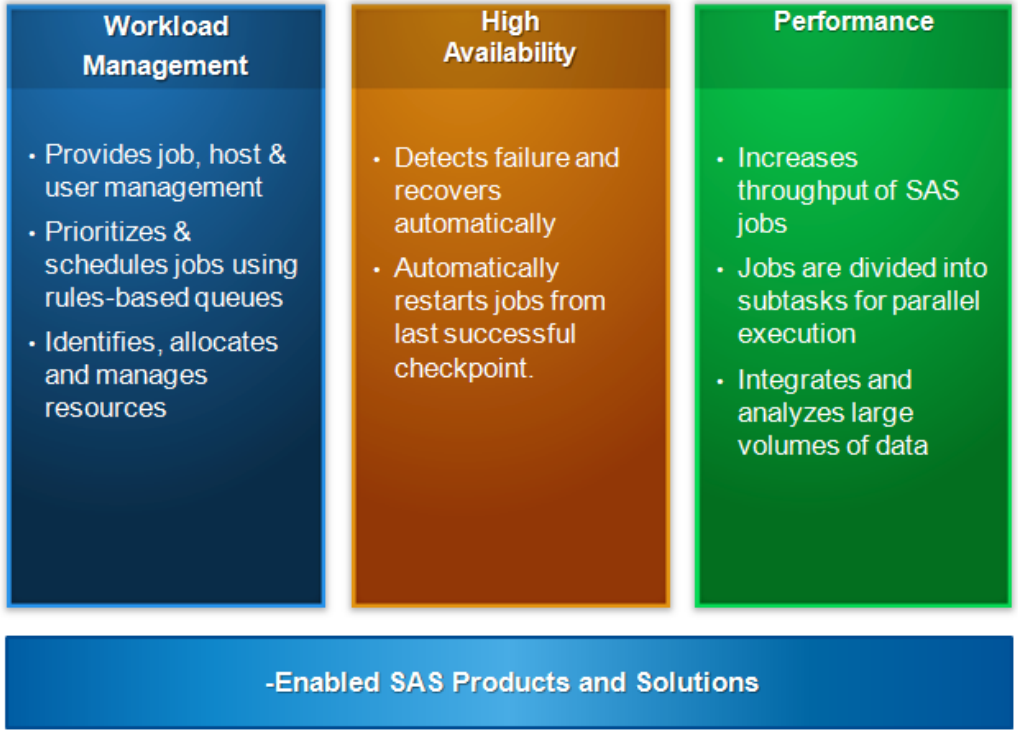
Bell SAS architecture setup - partial overview.jpg - Windows Photo Viewer

File Print E-mail Burn Open



SAS Grid Computing Architecture





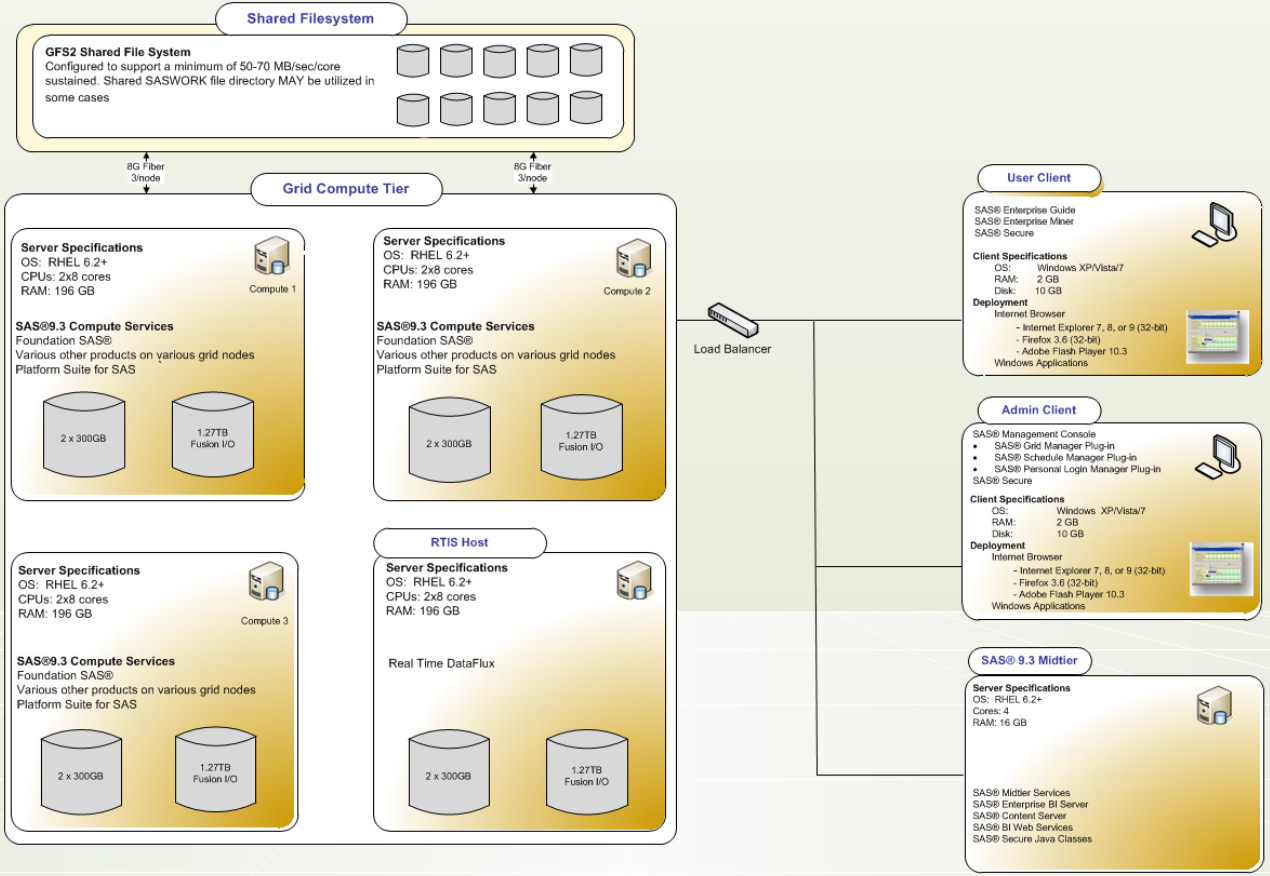
	Grid Environment	Non-Grid Environment
PROS	• Better job/resource manageability • Better High Availability (HA) • Application acceleration, Job parallelization capability • Better utilization of HW • Better flexibility and scalability • Better way to share a single resource pool for two or more environments scenario • Yield better performance overall	• Eliminate one layer of complexity • No impact to the current BU scheduling approach
CONS	• More complex to manage/troubleshoot • Additional training required for SAS Admin and potential SAS users in order to maximize the Grid benefit • Requirement to re-architecture the scheduling approach, share a single scheduler instead.	• No job parallelization • Less job/environment HA • Less hardware utilization

Prepared for Bell Canada by SAS® Institute Inc.

Bell Canada Production Grid Infrastructure Diagram

Physical Implementation Reference Diagram

Version 1 (07/22/12)



Abstract

- In the process of moving 900 users, from SASv9.2 on Solaris to Grid/SASv9.4 on Linux, many lessons were learned. 16 business units had a diverse set of expectations and 4 different strategies were used : Self managed, Outsourced, Re Engineered and Big Bang. The presentation will cover some architecture, the project management structure and each scenario. If time permits, some useful techniques and pitfalls will be discussed.

SAS® FORUM
UNITED KINGDOM 2015

www.SAS.com