Getting Started with Visual Analytics Administration and Deployment

Meera Venkataramani, Gary Mehler
A SINGLE SOLUTION

EXPLORE

DISCOVER

DESIGN

DELIVER

REPORT

VISUAL ANALYTICS

WHY?

SAS® and Big Data

SAS Global Forum 2013
CENTRAL ENTRY POINT

INTEGRATION

ROLE-BASED VIEWS

PREPARE
- Manage data
- Load and join data
- Create calculated columns

EXPLORE
- Perform ad-hoc data exploration
- Insights generated through analytic visualizations

DESIGN
- Create dashboard style reports for web or mobile

DELIVER
- SAS® Mobile BI - native tablet applications delivering interactive reports
  - Web and PDF

IN-MEMORY ANALYTICS ENGINE

DELIVERS A SINGLE SOLUTION FOR FASTER, SMARTER DECISIONS
FOR BUSINESS

- Easy-to-use analytics create a starting point for building and analytics driven culture
- Visual exploration of data, combined with analytics, fuel insight discovery for competitive advantage
- Rapid delivery of insights to colleagues via the Web and mobile promote collaboration and fact based decisions

FOR IT

- Control governance and empower colleagues with easy access to all data
- Leverage a highly scalable environment for cost effective growth when the business needs it
BASIC ADMINISTRATION

- Understand LASR data serving
- Understanding hardware
- Load and manage data
- Manage security access
- Monitor system
- Control mobile access
LASR SERVER – HOW IT WORKS

- Big data is processed fast because it is stored in memory
- Big data can be loaded quickly if stored in parallel, distributed storage
  - Hadoop HDFS, Teradata, Greenplum, others
- It can also be loaded from other data sources you have
- We load data into server sessions that hold data and perform analytics
- Server sessions can be limited in total size if needed

Visual Analytics Administrator

- LASR processes running
- Tables available in each process
MANAGING LASR SERVER INSTANCES
KNOWING WHAT DATA IS BEING SERVED UP
### SEEING WHAT IS IN DISTRIBUTED STORAGE

A screenshot of a SAS® interface is shown, displaying a list of directories and files within a distributed storage environment. The interface includes options for managing LASR Tables, HDFS, Resource Monitor, Process Monitor, and Mobile Devices. Each directory and file is listed with properties such as Name, Size, Date Modified, Property, and Value.
BASIC ADMINISTRATION

- Understand LASR data serving
- Understanding hardware
- Load and manage data
- Manage security access
- Monitor system
- Control mobile access
UNDERSTANDING HARDWARE

- LASR Server can run in two modes
  - MPP – across a collection of systems
    » Data and processing of that data is distributed across all systems
    » Operates on Linux systems
  - SMP – single system with plentiful memory and CPU cores
    » Data and processing occurs on a single system
    » Operates on a Linux system
    » Coming in Summer: can operate on a Windows-based server

- Visual Analytics Administrator shows resource use across hardware
DISTRIBUTED DEPLOYMENT (FOR COMMODITY HARDWARE)

SAS® VISUAL ANALYTICS

DESKTOP CLIENTS
- SAS® Management Console

WEB-BASED CLIENTS
- Hub
- Explorer
- Designer
- Viewer
- Data Builder
- Administrator

MOBILE CLIENTS
- iPad
- Android

BLADE ENVIRONMENT

METADATA SERVER

IN-MEMORY STORE

MID-TIER

WORKSPACE SERVER

SAS® LASR ANALYTIC SERVER

Hadoop HDFS

Hadoop RDBMS Nonrelational Click Stream PC Files

SAS® and Big Data
SAS Global Forum 2013
NON-DISTRIBUTED DEPLOYMENT (FOR COMMODITY HARDWARE)
INTRODUCTION TO SAS VISUAL ANALYTICS

Visual Analytics 6.1 comes in 2 versions

- **Distributed**
  - Multiple Machines
  - SAS High Performance Analytics Environment
  - Co-located Data Provider

- **Non-Distributed**
  - Single machine
  - No SAS High-Performance Analytics Environment
  - No co-located data provider
The Distributed version of SAS VA 6.1 comes with one of 3 co-located data Providers

- Distributed
  - Multiple Machines
  - SAS High Performance Analytics Environment
  - Co-located Data Provider
The Distributed version of SAS VA 6.1 comes with one of 3 co-located data Providers

- The 3 choices of Co-Located Data providers are
  - Teradata
  - GreenPlum
  - Hadoop Distributed File System (aka HDFS)
To put it another way:

- **Software**: SAS Visual Analytics 6.1
  - **Mode**: Non Distributed, Distributed
    - **Co-Located Data Provider**: GreenPlum, Teradata, HDFS
HIGH LEVEL VIEW

Source Data

Co-Located Data Provider

In-Memory Analytical Engine

Middle Tier

Load

Action

Action

Action

Browser

Mobile
DATA LOADING – NON-DISTRIBUTED

Source Data

In-Memory Analytical Engine

Mobile

Browser

Middle Tier

Action

Load

Co-located

Data Provider

Load

Data Provider
PACKAGING REALITIES

- VA is not an appliance
  - The customer has to buy the hardware
  - It still does not come in a “ready to go” state.
- You still need someone to install/configure
DEFINITIONS – BLADES

- Blades are very small
DEFINITIONS - CHASSIS

- A **chassis** contains
  - A number of blades
  - Power Supply for the blades
  - Cooling
  - Networking Interfaces
  - Storage Interfaces

Multiple Chassis can be connected together.

- **Animation**
STORAGE

- Applicable to Commodity Hardware
- Teradata and Greenplum have their own pre-defined storage
- VA does not require very fast storage
  - No need for FusionIO/SSD on worker nodes
- VA does not require Shared (clustered) Storage
- Recommended local storage configuration
  - Mirrored drives on SAS Node
  - Mirrored drives on Root Node and Name Node
  - Other Nodes do not need mirroring, but that may make them fragile
    - A disk failure will take out a node
5.1 Order

SAS COMSAT: VA MPP – DISTRIBUTED MODE

Order 09C7ZS (Test, Internal)

Order Actions
- Send SOE
- Mark Complete
- Cancel
- Release
- Reprocess
- View SIDs
- View TLetter
- View SOI
- Link Here
- View Logs
- View Directory
- FSSET Data

Summary | Products | ESD Downloads | Distribution | Processing History

<table>
<thead>
<tr>
<th>Product/Bundle</th>
<th>Expiration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Base SAS</td>
<td></td>
</tr>
<tr>
<td>SAS LASR Analytic Server Distributed Mode</td>
<td>07Dec2012</td>
</tr>
<tr>
<td>SAS High-Performance Server</td>
<td></td>
</tr>
<tr>
<td>SAS LASR Analytic Server</td>
<td></td>
</tr>
<tr>
<td>SAS Threaded Kernel Extensions for Advanced Analytics</td>
<td></td>
</tr>
<tr>
<td>SAS Threaded Kernel Extensions for High-Performance Analytics</td>
<td></td>
</tr>
<tr>
<td>SAS Threaded Kernel Extensions for LASR Analytic Server</td>
<td></td>
</tr>
<tr>
<td>SAS Threaded Kernel Message Passing Interface</td>
<td></td>
</tr>
<tr>
<td>SAS Threaded Kernel Extension for SSH Remote Process Launching</td>
<td></td>
</tr>
<tr>
<td>SAS High-Performance Management Console</td>
<td>21Feb2013</td>
</tr>
<tr>
<td>SAS High-Performance Deployment for Hadoop</td>
<td></td>
</tr>
<tr>
<td>SAS High-Performance Node Installation</td>
<td></td>
</tr>
<tr>
<td>SAS Visual Analytics</td>
<td>21Feb2013</td>
</tr>
</tbody>
</table>
5.1 Order

SAS COMSAT: VA MPP – DISTRIBUTED MODE

Order 09C7ZS (Test, Internal)

<table>
<thead>
<tr>
<th>Order Actions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Send SOE</td>
</tr>
<tr>
<td>Mark Complete</td>
</tr>
<tr>
<td>Cancel</td>
</tr>
<tr>
<td>Release</td>
</tr>
<tr>
<td>Reprocess</td>
</tr>
<tr>
<td>View SIDs</td>
</tr>
<tr>
<td>View TRLetter</td>
</tr>
<tr>
<td>View SOI</td>
</tr>
<tr>
<td>Link Here</td>
</tr>
<tr>
<td>View Logs</td>
</tr>
<tr>
<td>View Directory</td>
</tr>
<tr>
<td>FSSET Data</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Summary</th>
<th>Products</th>
<th>ESD Downloads</th>
<th>Distribution</th>
<th>Processing History</th>
</tr>
</thead>
<tbody>
<tr>
<td>Search Products</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Product/Bundle</th>
<th>Expiration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Base SAS</td>
<td>21Feb2013</td>
</tr>
<tr>
<td>SAS LASR Analytic Server Distributed Mode</td>
<td>07Dec2012</td>
</tr>
<tr>
<td>SAS High-Performance Server</td>
<td></td>
</tr>
<tr>
<td>SAS LASR Analytic Server</td>
<td></td>
</tr>
<tr>
<td>SAS Threaded Kernel Extensions for Advanced Analytics</td>
<td></td>
</tr>
<tr>
<td>SAS Threaded Kernel Extensions for High-Performance Analytics</td>
<td></td>
</tr>
<tr>
<td>SAS Threaded Kernel Extensions for SAS LASR Analytic Server</td>
<td></td>
</tr>
<tr>
<td>SAS Threaded Kernel Message Passing Interface</td>
<td></td>
</tr>
<tr>
<td>SAS Threaded Kernel Extension for SSH Remote Process Launching</td>
<td></td>
</tr>
<tr>
<td>SAS High-Performance Management Console</td>
<td></td>
</tr>
<tr>
<td>SAS High-Performance Deployment for Hadoop</td>
<td></td>
</tr>
<tr>
<td>SAS High-Performance Node Installation</td>
<td></td>
</tr>
</tbody>
</table>
### 5.1 Order

#### SAS ORDER INFORMATION: VA MPP – DISTRIBUTED MODE

<table>
<thead>
<tr>
<th>Your Order Information:</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Order Number:</td>
<td>09C7ZS</td>
</tr>
<tr>
<td>SAS Installation Key:</td>
<td>SJSI-96H5-7ST3</td>
</tr>
<tr>
<td>Deployment Type:</td>
<td>Planned (Deployment plan needed)</td>
</tr>
<tr>
<td>Date:</td>
<td>December 03, 2012</td>
</tr>
<tr>
<td>Space Required for SAS Software Depot:</td>
<td>11.7GB</td>
</tr>
<tr>
<td>Software Status:</td>
<td>Pre-production</td>
</tr>
<tr>
<td>SAS Installation Representative</td>
<td>Temporary SAS User</td>
</tr>
<tr>
<td>E-mail Address</td>
<td></td>
</tr>
<tr>
<td>Tech Support Site Number</td>
<td>70068118</td>
</tr>
</tbody>
</table>

#### Information for Tech Support Site 70068118:

| Site Name:                                         | Pre-Prod LAX VA 6.1 Distributed |
| Tech Support Site Number:                          | 70068118 |
| Contracts Site Number:                             | 545535 |
| Operating System:                                 | Linux® for X64 |
| Product:                                          | SAS 9.3 64-bits TS1M2, Rev. 930_12w50 |

<table>
<thead>
<tr>
<th>Product(s) Ordered</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Base SAS</td>
<td></td>
</tr>
<tr>
<td>SAS LASR Analytic Server Distributed Mode</td>
<td></td>
</tr>
<tr>
<td>SAS Visual Analytics</td>
<td></td>
</tr>
</tbody>
</table>
5.1 Order

SAS COMSAT: VA SMP – NON-DISTRIBUTED MODE
5.1 Order

SAS COMSAT: VA SMP – NON-DISTRIBUTED MODE

Order 09C7ZT (Test, Internal)

Order Actions
- Send SOE
- Mark Complete
- Cancel
- Release
- Reprocess
- View SIDs
- View TLetter
- View SOI
- Link Here
- View Logs
- View Directory
- FSSET Data

Summary

Products

ESD Downloads

Distribution

Processing History

Product/Bundle

Base SAS

SAS LASR Analytic Server Non-Distributed Mode

SAS High-Performance Server
SAS LASR Analytic Server
SAS Threaded Kernel Extensions for Advanced Analytics
SAS Threaded Kernel Extensions for High-Performance Analytics
SAS Threaded Kernel Extensions for LASR Analytic Server
SAS Threaded Kernel Message Passing Interface
SAS Threaded Kernel Extension for SSH Remote Process Launching
SAS Threaded Kernel Extension SMP Components

Expiration

21Feb2013
07Dec2012
### 5.1 Order

**SAS ORDER INFORMATION: VA SMP – NON-DISTRIBUTED MODE**

#### Your Order Information:

<table>
<thead>
<tr>
<th>Order Number:</th>
<th>09C7ZT</th>
</tr>
</thead>
<tbody>
<tr>
<td>SAS Installation Key:</td>
<td>WFHP-1278-Y9LD</td>
</tr>
<tr>
<td>Deployment Type:</td>
<td>Planned (Deployment plan needed)</td>
</tr>
<tr>
<td>Date:</td>
<td>December 03, 2012</td>
</tr>
<tr>
<td>Space Required for SAS Software Depot:</td>
<td>11.4GB</td>
</tr>
<tr>
<td>Software Status:</td>
<td>Pre-production</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Temporary SAS User</th>
<th>E-mail Address</th>
<th>Tech Support Site Number</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>70068118</td>
</tr>
</tbody>
</table>

#### Information for Tech Support Site 70068118:

<table>
<thead>
<tr>
<th>Site Name:</th>
<th>Pre-Prod LAX VA 6.1 SMP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tech Support Site Number:</td>
<td>70068118</td>
</tr>
<tr>
<td>Contracts Site Number:</td>
<td>545535</td>
</tr>
<tr>
<td>Operating System:</td>
<td>Linux® for X64</td>
</tr>
<tr>
<td>Product:</td>
<td>SAS 9.3 64-bits TS1M2, Rev. 930_12w50</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Product(s) Ordered</th>
</tr>
</thead>
<tbody>
<tr>
<td>Base SAS</td>
</tr>
<tr>
<td>SAS LASR Analytic Server Non-Distributed Mode</td>
</tr>
<tr>
<td>SAS Visual Analytics</td>
</tr>
</tbody>
</table>
BASIC ADMINISTRATION

- Understand LASR data serving
- Understanding hardware
- **Load and manage data**
- Manage security access
- Monitor system
- Control mobile access
ACCESSING DATA

- Access new information
  - Excel spreadsheets
  - CSV-exported data
  - SAS datasets on a desktop

- Access existing information
  - SAS data on a network
  - Other data in databases
  - Data available from SAS Information Maps

- **Visual Data Builder**
IMPORTING LOCAL SPREADSHEETS, CSV FILES, ETC
IMPORTING SAS INFORMATION MAPS
JOINING DATA PRIOR TO LOADING INTO LASR MEMORY
SCHEDULING RECURRING TASKS TO KEEP DATA CURRENT
MANIPULATING DATA

- Perform basic joins on data before it is loaded
- Can load directly into LASR or into other SAS library types
- If loading into LASR, can stage data in distributed storage as a pre-step
- Can schedule to re-run on-demand to keep data fresh
BASIC ADMINISTRATION

- Understand LASR data serving
- Understanding hardware
- Load and manage data
- **Manage security access**
- Monitor system
- Control mobile access
MANAGE SECURITY ACCESS

- Identities and access managed through SAS Metadata Server
- New table definitions can be added and kept current
  - Like Register Tables and Update Table Metadata in SAS Management Console
- Folder-level, table-level, and row-level permission can be granted
- LASR Server sessions can be managed and maintained
- Roles and capability enabled
ROLES AND CAPABILITIES CAN LIMIT FUNCTIONAL ACCESS
SECURING DATA SOURCES, REPORTS, EXPLORATIONS
SETTING ROW-LEVEL SECURITY ON DATA SOURCES

SAS® and Big Data
SAS Global Forum 2013

Edit Permission Condition

Enter an expression to specify the rows that can be viewed by user Mehler, Gary J.

- Data Items
  - Character
    - Customer_Age_Group
    - Customer_Country
    - Customer_FirstName
    - Customer_Gender
    - Customer_Group
    - Customer_LastName
    - Customer_Name
    - Customer_Type
  - Numeric
- Operators
  - Search type
    - Numeric (simple)
      - x - y (subtract)
      - x * y (multiply)
      - x / y (divide)
      - x + y (add)
    - Comparison
      - Between (x between y and z)
      - Missing (is missing)
      - Not Missing (is not missing)
- AND
  - (Customer_Group = CatalogSales)
  - (Customer_Country = USA)

{(Customer_Group = CatalogSales) AND (Customer_Country = USA)}
BASIC ADMINISTRATION

- Understand LASR data serving
- Understanding hardware
- Load and manage data
- Manage security access
- **Monitor system**
- Control mobile access
MONITOR SYSTEM

- Observe memory, CPU, and network utilization
- See which LASR server processes are consuming memory or processing
- Determine which tables are available and their status
- Historic and other reports coming from Environment Manager later this year
BROWSING OVERALL HEALTH AND ACTIVITY
BASIC ADMINISTRATION

- Understand LASR data serving
- Understanding hardware
- Load and manage data
- Manage security access
- Monitor system
- Control mobile device access
CONTROL MOBILE DEVICE ACCESS

- Monitor login history
- Can manage via blacklist
  - A user or device gets added to blacklist as needed
  - Helpful if staff leaves organization, or a mobile device is lost
- Can manage via whitelist
  - Individuals or groups are enabled to access via mobile devices
  - Individuals or devices can be removed from access list if needed
BLACKLISTING IS IN USE BY DEFAULT
Big Data

STEPS TO CONQUER COMPLEXITY

TURN CHALLENGE INTO OPPORTUNITY
• 53% leverage half of their valuable data
• 24% believe vast amounts of important data goes untapped

GAIN MAXIMUM VALUE FROM YOUR DATA
• SAS® Analytics + Powerful Visualizations + Sharing
• High Speed Performance + Cost Efficient Scalability

Source: Economist Intelligence Unit 2011 Report, Sponsored by SAS, 2011