SAS Analytical Platform - technology direction and some modernization options

Henrik Slettene
Head of Nordic Center of Excellence, Enterprise Analytics
Modernization?

(1990 announcement)

Film → Scanner → Manipulation → Put on CD → Available to play

Source: Kodak’s Hybrid New York Times October 7, 1990
Digitalization

Digital opportunities as competitive edge - and threat

"My business and its IT organization are being engulfed by a torrent of digital opportunities. We cannot respond in a timely fashion, and this threatens the success of the business and the credibility of the IT organization."

<table>
<thead>
<tr>
<th>Region</th>
<th>Agree</th>
<th>Disagree</th>
</tr>
</thead>
<tbody>
<tr>
<td>EMEA</td>
<td>52%</td>
<td>48%</td>
</tr>
<tr>
<td></td>
<td>43%</td>
<td>57%</td>
</tr>
<tr>
<td></td>
<td>53%</td>
<td>47%</td>
</tr>
<tr>
<td></td>
<td>60%</td>
<td>40%</td>
</tr>
<tr>
<td></td>
<td>33%</td>
<td>67%</td>
</tr>
</tbody>
</table>

"The IT organization has the right skills and capabilities in place to meet upcoming challenges."

### SAS technology directions

**What is «modern sas»?**

<table>
<thead>
<tr>
<th>Feature</th>
</tr>
</thead>
<tbody>
<tr>
<td>SAS on Hadoop</td>
</tr>
<tr>
<td>Moving SAS to the data</td>
</tr>
<tr>
<td>In-memory analytics</td>
</tr>
<tr>
<td>Visualization and exploration</td>
</tr>
<tr>
<td>Decision Management</td>
</tr>
<tr>
<td>Data Management including Event Stream Processing</td>
</tr>
<tr>
<td>Improved platform management and deployment</td>
</tr>
</tbody>
</table>
I have SAS… What are possible next steps?

- High Performance Analytics
- Visual Statistics on top of VA
- SAS on SAP HANA
- Decision Management
- Data Management
- SAS on Hadoop
- Analytics Factory
In-Memory Analytics based on LASR Server

In-Memory Analytics Platform
*Shared single copy of data for all three!*

<table>
<thead>
<tr>
<th>Shared Graphical User Interface</th>
<th>SAS® Studio Interface</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>SAS® Visual Analytics</strong>&lt;br&gt;with Approachable Analytics</td>
<td><strong>SAS® In-Memory Statistics for Hadoop</strong></td>
</tr>
<tr>
<td><strong>SAS® Visual Statistics</strong>&lt;br&gt;with Approachable Modelling</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Open to the masses</th>
<th>Reduced number of users</th>
<th>Lower number of users</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>No</strong> statistics skills required</td>
<td><strong>Some</strong> Statistics skills required</td>
<td><strong>Statistics skills</strong> required</td>
<td></td>
</tr>
<tr>
<td><strong>No</strong> programming knowledge required</td>
<td><strong>No</strong> programming knowledge required</td>
<td><strong>Programming knowledge</strong> required</td>
<td></td>
</tr>
<tr>
<td>Focus on <strong>consuming</strong> analytics</td>
<td>Focus on <strong>producing</strong> analytics</td>
<td>Focus on <strong>producing</strong> analytics</td>
<td></td>
</tr>
</tbody>
</table>
SAS and SAP

**Reduced Data Movement**
Short data-to-compute distance

**Massive Performance Boost**
In-memory parallel processing for analytic workloads.

**Simplified IT Landscape**
Analytical and Transactional processing on a single data store.
SAS Decision Manager

SAS® Decision Manager (Project-based Collaboration, Governance)

SAS® Visual Analytics Admin & Reporting
SAS® Model Manager
SAS® Business Rules Manager
SAS® Data Integration Server
SAS® Real-time Decision Manager
SAS® Visual Analytics

Shared Platform Components
Lineage/Metadata
Scheduling
Security
Monitoring
Workflow
SAS Data Management
Unified Data Management architecture

DATA QUALITY SERVICES
- Profile
- Discover
- Verify
- Cleanse
- Standardize
- Match & Consolidate

DATA INTEGRATION SERVICES
- Data Transforms
- Apply Business Rules
- Analytic Transforms

DEPLOYMENT SERVICES
- Orchestrate
- Deploy
- Schedule

ENTERPRISE DATA ACCESS AND DATA FEDERATION

COLLABORATION & GOVERNANCE

BUSINESS & TECHNICAL METADATA
Hadoop and SAS – two scenarios

Hadoop as a Data Platform  
(standalone or as part of a broader ecosystem)

Hadoop as a core component of the next generation of BI and Analytics
Hadoop as a data platform

- Hadoop as a new data store
- Hadoop as an additional input to the EDW
- Hadoop as a foundation for BI & Analytics
- Hadoop as a "data lake" or staging area for the Warehouse
SAS on Hadoop

Today...

**BASE SAS**
- Map Reduce + Pig Scripting + HDFS Commands

**SAS/Access to Hadoop**
- Hive, Hive2 + Direct file access

**SAS/Access to Impala (Cloudera only)**

**SAS Data Integration Studio (Transforms) in Data Management**
- Standard / Advanced:
  - Read/Write HDFS files
  - Submit HiveQL code
  - Execute Map/Reduce code
  - Submit Pig Latin
  - Transfer data to/from Hadoop using Hadoop utilities
  - SQL transforms pushed down with Access to Hadoop engine

**SAS Federation Server**
Virtual and secure access to Hadoop and more traditional sources

**SAS Event Stream Processing Engine**
To bring streaming data from Sensors into Hadoop

Coming very soon...

**SAS Data Loader for Hadoop**
- SAS Data Loader
  - HTML 5 GUI to drive Execution

- SAS Data Quality Accelerator for Hadoop
  - Execute selected DQ routines in Hadoop

- SAS Code Accelerator for Hadoop
  - Execute SAS DS2 code in Hadoop

Point and click data management routines where data stays in Hadoop
Hadoop in the analytics lifecycle

SAS Visual Analytics
SAS Scoring Accelerator for Hadoop
SAS Code Accelerator for Hadoop

Done using the Build Model Tools and other checks

SAS Visualization
SAS Visual Statistics
SAS In-Memory Statistics for Hadoop

Done using either the Data Preparation, Data Exploration or Build Model Tools

SAS High Performance Analytics Offerings supported by relevant clients like SAS Enterprise Miner, SAS/STAT etc.
I have SAS... What are possible next steps?

High Performance Analytics

Visual Statistics on top of VA

SAS on SAP HANA

Decision Management

Data Management

SAS on Hadoop

Analytics Factory

SAS Forum Stockholm 2014