Cloud Based Analytics for SAP

Gary Patterson, Global Lead for Big Data
About Virtustream

A Dell Technologies Business
• 2,300+ employees
• 20+ data centers
• Major operations in 10 countries
• One of the fastest growing Dell EMC companies

Technology Leadership
• IP Leader in cloud enterprise application assurance and security
• Inventor of the µVM™ cloud technology
• Developer of xStream® Cloud Management Platform software
• Operator of leading enterprise-class cloud IaaS
Enterprise Big Data Use Cases

- Consumer
  - Behavior
- Competitive
  - Advantage
- Customer
  - Retention
- Marketing
  - Demand
- Product
  - Development
- Performance
- Supply
  - Scheduling
- Production
- Industry
  - Bio-Informatics
- Operational
  - Efficiency
- Monitoring
- Security
  - Risk
  - Fraud
  - Compliance
- Preventive
  - Maintenance
- Fault
- Insight
- Fault
Complexities of Managing Big Data

- Hardware Platform
  - Setup and Configuration

- Software Platform
  - Setup and Configuration

- Big Data in Production at scale...
  - Tool integration problems
  - Broken nodes
  - Performance tuning
  - Job problems
  - Capacity problems
  - Upgrade Infrastructure
  - Upgrade software
  - Tool integration problems

© 2017 Virtustream, Inc. All Rights Reserved.
Virtustream Big Data Service

Production-ready Cloud Big Data analytics
Instant access to Hadoop and Spark platform-as-a-service for rapid *time to value*

Built for performance, reliability
Rapid scalability to adjust capacity and reliable job completion

Enterprise-class security
Designed from the ground up to run sensitive data

Flexible Use cases
Hadoop, Spark and HANA integrated analytics

Full suite of managed services
Solves the infrastructure, software and people challenges for Big Data

Predictable pricing
Bundled, predictable and all-inclusive pricing for *lower TCO* and *higher ROI*
Production-Ready Big Data Analytics
Instant access to Hadoop and Spark platform-as-a-service

- Deployment in hours, not months
- Portal - Analytics workbench access to built-in Hadoop tools
- Rapid scalability to adjust capacity
- Data analysis from terabytes to multi-petabytes
- Lower TCO to build, operate and adapt Big Data solutions
- Faster *time to value*, higher ROI
Enterprise-Class Requirements

Ease of Administration

Administration
• Access to ticketing system
• Monitoring and alerts
• User account management

Cluster Access
• Compute & Storage dashboard
• Real-time job status
• Analytics workbench with built-in Hadoop tools

Performance, Reliability

Fast and reliable job completion
• Optimized for analyzing large and varied volumes of data
• Infrastructure tuned for speed
• Delivers fast job completion
• Low job failures leads to cost efficiency

Security

Designed for sensitive data
• Robust network authentication
• Data encryption
• Secure, dedicated connectivity options
• Industry compliances
Implement Hadoop strategies for optimizing HANA storage, optimizing compute or acting as a staging area before moving data to HANA

- Archiving old transactions from HANA to Hadoop
- Utilizing VORA to execute federated queries from HANA + Hadoop
- Offloading HANA processing
- Offloading SAP core processing (e.g., CRM, BU data - customer target regeneration, Churn analysis)
- Storing IoT data (e.g., Sensor or Telemetry data) before cleansing, reducing and moving to HANA
- Data Science platform for predictive analytics such as forecasting models, predictive maintenance, etc.

Specialized SAP Solutions
- SAP Predictive Maintenance and Service
- SAP Connected Manufacturing
- SAP Connected Health
- SAP Connected Logistics
- SAP IT Operational Analytics (ITOA)
Virtustream Big Data Service with SAP HANA
Operational insights for enterprise applications

- Operationalizes SAP landscape into Virtustream Big Data Service
- Supports specialized SAP Big Data use cases
- Deep expertise in SAP Business Warehouse (BW) and SAP HANA
- Reduces the time, cost, and effort of integrating multiple Big Data tools
- Reduced cost of running HANA

Enables a holistic SAP BW and HANA Data Management and Optimization strategy that leverages Hadoop
# Virtustream Big Data Service Components

<table>
<thead>
<tr>
<th>Component</th>
<th>Open Source Stack</th>
<th>SAP Big Data Service</th>
<th>SAP HANA</th>
<th>Spark</th>
<th>SAP HANA Vora</th>
</tr>
</thead>
<tbody>
<tr>
<td>Advanced Analytics / ML</td>
<td>Mahout, R, Spark ML</td>
<td>Spark ML, Spark R</td>
<td>PAL</td>
<td>Spark ML</td>
<td>PAL, Spark</td>
</tr>
<tr>
<td>Reporting/Dashboard/Visualization</td>
<td>Zeppelin</td>
<td>Zeppelin</td>
<td>Business Objects, Lumira</td>
<td>Zeppelin</td>
<td>Business Objects, Lumira</td>
</tr>
<tr>
<td>Analytics</td>
<td>Hive, Spark</td>
<td>Hive, Spark</td>
<td>HANA</td>
<td>Storm</td>
<td>Vora/Spark</td>
</tr>
<tr>
<td>SQL / DW</td>
<td>Hive</td>
<td>Hive</td>
<td>BW</td>
<td>Spark SQL Hive</td>
<td>Spark SQL/BW</td>
</tr>
<tr>
<td>No SQL / Time Series</td>
<td>Cassandra / HBASE, Open TSDB</td>
<td>HBASE</td>
<td>HANA</td>
<td>HBASE</td>
<td>HANA</td>
</tr>
<tr>
<td>Data Integration and Processing</td>
<td>Flume/SQOOP/Kafka, MapR, PIG, Spark</td>
<td>Spark Streaming, Flume, Sqoop, Kafka</td>
<td>SAP Data Services</td>
<td>Spark Streaming</td>
<td>Spark Streaming, Sqoop, Data Services</td>
</tr>
<tr>
<td>Cluster Manager / Admin Storage Service</td>
<td>YARN/ HDFS</td>
<td>YARN/ HDFS</td>
<td>AppDirector / HANA</td>
<td>YARN</td>
<td>YARN/ HDFS</td>
</tr>
<tr>
<td>Hypervisor / CMP</td>
<td>Open Stack</td>
<td>Docker Containers</td>
<td>VMware</td>
<td>OpenStack/ xStream</td>
<td>VMware / Docker xStream</td>
</tr>
</tbody>
</table>

© 2017 Virtustream, Inc. All Rights Reserved.
Big Data Managed Services

Managed Services delivers on-demand provisioning of experts to improve ongoing effectiveness of a client’s cloud based information technology assets.

<table>
<thead>
<tr>
<th>Application Management</th>
<th>Performance Tuning, Capacity Planning, Resource Optimization</th>
</tr>
</thead>
<tbody>
<tr>
<td>DB &amp; App Administration</td>
<td>Hadoop, SAP Vora, Administration Upgrades, Patches, Issue Resolution</td>
</tr>
<tr>
<td>VM &amp; OS Administration</td>
<td>Admin, Installation, Upgrades, Patches, Security</td>
</tr>
<tr>
<td>Platform Services</td>
<td>Onboarding, Admin, Provisioning, Backup and Disaster Recovery, 24x7, Helpdesk</td>
</tr>
<tr>
<td>Compute, Storage</td>
<td>Compute, Storage and I/O Hardware</td>
</tr>
<tr>
<td>Facilities, Network</td>
<td>Physical Plant, Redundant Carriers</td>
</tr>
</tbody>
</table>

Monitoring | Management | Optimization | SLAs
### Why Virtustream

**Differentiators beyond Hadoop-as-a-service**

<table>
<thead>
<tr>
<th>Enterprise Focus</th>
<th>Experienced SAP Cloud Hosting</th>
<th>Expertise to augment HANA with Hadoop and Vora Integration</th>
<th>Enable specialized SAP Big Data use cases</th>
<th>Extend existing SAP BW/HANA landscapes with Hadoop</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Mission-Critical Systems of Record</td>
<td>• Virtustream leads SAP’s strategic direction for cloud, including HANA journey</td>
<td>• Subject Matter Expertise in SAP BW, and HANA</td>
<td>• Relevant SAP Hosting Expertise</td>
<td>• Complete coverage of Data Management – NLS, DLM, Dynamic Tiering, SDA, BODS</td>
</tr>
<tr>
<td>• Complete SAP Portfolio Coverage, including ECC, CRM, BW, HANA &amp; other SAP Applications in the Cloud</td>
<td>• Certified Hosting Services &amp; Operations</td>
<td>• Early Starter’s Advantage in Vora</td>
<td>• Operationalize SAP big data for actionable insight</td>
<td>• Co-location with SAP workloads</td>
</tr>
<tr>
<td></td>
<td>• ~300 SAP clients running production SAP</td>
<td>• Large SAP HANA deployments</td>
<td>• Fully Managed Vora Services</td>
<td>• Go to partner for SAP HEC</td>
</tr>
</tbody>
</table>