Customer Cloud Architecture for Big Data Analytics

http://cloud-council.org/resource-hub.htm#customer-cloud-architecture-for-big-data-analytics
Presenters

Manav Gupta
Cloud Technical Leader, IBM
manavgup@ca.ibm.com

Craig Statchuk
Technical Architect, IBM Business Analytics
Craig.Statchuk@ca.ibm.com

Heather Kreger
CTO International Standards, IBM
US SC38 mirror committee ANSI (Editor SC38 SOA RA)
kreger@us.ibm.com
This talk will discuss introduce the third of a series of vendor neutral Cloud Solution Architectures - the Big Data Analytics Cloud Architecture that was recently published. We will show how Open Source can be used to implement this architecture.

- What are Cloud Solution Architectures?
- Candidate Architectures
- Big Data Analytics Cloud Architecture
- Open source projects that can be used to support this
- CSCC Cloud Reference Architecture series
Cloud Customer Reference Architectures are…

- straightforward description of elements needed to implement particular application solutions using cloud infrastructure, cloud platforms, cloud software, and cloud services
- deployment neutral (public, private, hybrid) & implementable via IaaS, PaaS, SaaS
- general purpose reusable architectures as well as industry specific architectures
- vendor neutral & open

Important because they…

- enable cloud customers to understand unique features & advantages of using cloud computing
- bridge gap between understanding cloud customer needs and cloud provider offerings
- provide practical guidance on how common business applications can be realized from a cloud customer role perspective
- are stable anchors in a rapidly innovating cloud landscape
- save time, effort & money: be more productive

Useful when…

- those planning to build cloud based applications
- talking with cloud providers about their offerings
- understanding of the common elements and relationships in relevant solutions

Target audience:

- those planning on building/purchasing cloud based applications
- developers, architects, managers

Consistent with ISO/IEC 17789 International Standard Cloud Computing Reference Architecture
General
- Big Data and Analytics
- Mobile (published 5/15)
- Web App Hosting (published 12/14)
- DevOps
- Hybrid Security
- Hybrid Integration
- Social

Industry
- E-commerce
- Financial
- Travel
- Healthcare
- Gaming
- Public Sector
- Telco

Published by CSCC 7/14, Review today!
Big Data Analytics and Cloud: A high level view
Big Data Analytics Cloud Solution Architecture
Data Sources: Cloud mitigates cost/risk by providing data tools via pay per use model
Edge Services: Ensure safe flow of data from internet to data analytics processing system.
Stream Computing: Process large volumes of highly dynamic, time-sensitive continuous data streams.
Data Integration: Copy & correlate info from disparate sources to produce meaningful associations.
Data Repositories: Hosted on different infra tuned to support the types of analytics workload accessing the data.
Big Data Analytics Cloud Solution Architecture

Actionable Insight: Relate new data with existing info in an understandable manner.
Transformation & Connectivity:
Enables secure connection to enterprise systems; Filter, aggregate, or modify data.
Enterprise Apps: Consume Cloud Provider data & analytics to produce business results.
Enterprise Data: Systems of record, for example, transactional data or data warehouses.
Enterprise Directory: Provides storage for and access to user security information.
Big Data Analytics Cloud Solution Architecture

Security:
ID & access management; App & data protection; Security intelligence
Info Governance:
Provides policies and capabilities that enable analytics env. to move, manage and govern data.
Big Data Analytics Cloud Solution Architecture

User: Enterprise or third party user: data analysts, data scientists, business users, solution architects.
Big Data Analytics Cloud Solution Architecture
Big Data Analytics Cloud Solution Architecture: Details
Big Data Analytics Cloud Solution Architecture: Fraud & Identity Theft Scenario

1. User access
   - User
   - Edge Services

2. New data added
   - Data Sources

3. Verify & normalize
   - Data Integration

4. Real time events
   - Streaming Computing

5. Discover, align & analyze
   - Data Repositories
   - Actionable Insight

6. Report & share insight
   - Transformation & Connectivity

7. Transform & deliver
   - Enterprise Applications

8. Decide & take action
   - Enterprise Data
   - Enterprise User Directory

Legend:
- Security
- Information Governance
- Scalable Infrastructure
Big Data Analytics Cloud Solution Architecture: Open Source Mapping

- **PaaS** – Cloud Foundry, OpenShift
- **Streaming** – Apache Storm, Pulsar, Tigon
- **File Repository** – Static Contents – OpenStack Object Storage (Swift), OpenStack Block Storage (Cinder), NFS
- **Node-RED** – Transform / Connect – ActiveMQ, Tyk, Mobicents
- **Paxata** – Refinement and Data Preparation
- **OpenRefine** – Refinement and Data Preparation
- **CDN** – Apache Web Server, OpenStack Object Storage
- **PaaS** – Cloud Foundry, OpenShift
- **Actionable Insight** – Birt, Pentaho, Mahout, Jupyter, R, D3, Sci-kit
- **Database** – MongoDB, CouchDB, MySQL, Postgres
- **User Directory** – OpenLDAP, OpenStack Keystone
- **Security** – Identity & Access Management, Data & Application Protection, Security Intelligence
Call to Action

- **Join the CSCC Now!**
  - To have an impact on customer use case based standards requirements
  - To learn about all Cloud Standards within one organization
  - To help define the CSCC’s future roadmap
  - Membership is free & easy: [http://www.cloud-council.org/application](http://www.cloud-council.org/application)

- **Get Involved!**
  - Join one or more of the CSCC Working Groups
    - [http://www.cloud-council.org/application](http://www.cloud-council.org/application)

- **Review all of the CSCC Cloud Reference Architecture papers**
Additional Resources

- **Customer Cloud Architecture for Big Data and Analytics**

- **Customer Cloud Architecture for Mobile**
  - http://bit.ly/1cGs5Xj

- **Web Application Hosting Cloud Solution Architecture**

- **Practical Guide to Cloud Service Agreements, V2**

- **Public Cloud Service Agreements: What to Expect & What to Negotiate**

- **Practical Guide to Cloud Computing, V2**
  - http://bit.ly/1MwD9mZ

- **Security for Cloud Computing: 10 Steps to Ensure Success, V2**
  - http://bit.ly/1L3D9gZ

- **Cloud Security Standards: What to Expect & What to Negotiate**

  - http://bit.ly/1Fg7lkk

- **Migrating Applications to Public Cloud Services: Roadmap for Success**

- **Impact of Cloud Computing on Healthcare**
Thank You