Cloud
Your Adaptive Integration Platform?

Philip Russom
Research Director for Data Management, TDWI
September 30, 2015
Speakers

Philip Russom
TDWI Research Director,
Data Management

Madhukar Kumar
VP of Products,
Liaison
Today’s Agenda

• Cloud Directions
  – Clouds, in general
  – Data integration (DI)

• Cloud DI Use Cases
  – Focus on Data Integration
  – Other Data Mgt Disciplines

• Miscellaneous Issues in Cloud DI

• The Future of Cloud DI
Where We’re Going in the Clouds…

• Traditional and Web-based applications & IT infrastructure
  – Not going away; being joined by cloud-based platforms
• SaaS-based applications are growing in popularity
  – Salesforce.com, ADP, NetSuite…
  – Older packaged apps now available via clouds (from SAP, Oracle, Microsoft)
• Enterprise cloud platforms propagating
  – Amazon, HP, IBM; VMWare
• The Result: Mixed Platforms: Clouds, Webs, traditional…
  – Platform diversity & complexity is increasing
• Data Management faces new requirements & challenges
  – Let’s focus here today, especially on data integration…
REQUIREMENTS
Cloud Data Integration (DI)

• General Requirements for Cloud DI:
  – DI tools/platforms must run on clouds
  – DI tools/platforms must interface with tools/platforms running on clouds, plus the usual enterprise & Web tools/platforms

• Primary Requirements for Cloud DI:
  – Support interfaces & protocols of popular SaaS-based applications and generic platforms running atop clouds
  – Right-time interfaces, from offline batch to real-time on-demand
  – Leverage elastic scalability of various cloud types
  – Enable business & mildly technical users to do
  – Enable advanced DI features for technical users developing multi-platform solutions
  – Provide bullet-proof security features
USE CASE
Data Sync among Cloud-Based Apps

• A very common case today…
  – Sync’ing data between SaaS-based salesforce.com and enterprise applications for CRM, SFA, etc.

• More complex cases…
  – Sync’ing customer data across multiple applications
  – Some on-premises, others in clouds

• Simpler cases…
  – Extracting customer data for BI reports
  – Uploading leads for sales to follow in salesforce.com
USE CASE
Cloud-Based Data Hubs

• Think of types of data stores in use for years…
  – _Hubs for customer data, product data, other data domains_
  – _Repositories of shared reference data for purposes of master data management (MDM)_
  – _Data staging areas for general data integration_

• Imagine these data stores in a cloud…
  – _Closer to SaaS apps and cloud platforms_
  – _Scalability guaranteed by cloud infrastructure_
USE CASE
Elastic Scaling for DI Volumes & Loads

• Data volumes fluctuate in most DI solutions
  – *Incoming data is landed in a staging area; much is deleted after processing into target formats*

• Loads on a DI server fluctuates
  – *Server sits idle most of business day; heavy batch-processing loads kick in after hours*

• Imagine DI solutions on a cloud
  – *Cloud assists scalability and performance*
  – *Server resources reallocated instead of sitting idle*
USE CASE
Agile Data Provisioning

• DI tool made for SaaS data will...
  – Provide tools for selecting data, then moving or enhancing it
  – Present data models and interfaces of popular SaaS applications in the DI development environment, making it easy to select the data source or target
  – Be easy enough for mildly technical users to do a certain amount of self service for DI

• With these features, any user can provision and integrate data sets quickly
  – Whether moving data into or out of a cloud
USE CASE
Data Quality & Enhancement

• All forms of data integration reveal data quality issues
  – Whether problems to fix or opportunities to leverage
• Cloud DI often handles customer data
  – Which is notoriously wrong, inconsistent, out-of-date
  – Which benefits from additional data about customers
• A good cloud DI solution will…
  – Improve customer data, not just move it
  – Standardize customer data
  – Help reduce redundant records
  – Aggregate and append additional customer info
USE CASE

Cloud Services Brokerage

• A traditional data integration platform helps you integrate data across multiple applications on multiple platforms.
• But what if some or all of the platforms are clouds?
• A cloud services brokerage is a cloud-based data integration platform that helps you integrate data across multiple applications on multiple platforms, including clouds.
• As organizations outsource more applications, data, and platforms to clouds, so grows the need for cloud services brokerages.
MISC ISSUES WITH Cloud Data Integration

- Like other forms of DI, Cloud DI is also about interfaces
  - Look for DI tools that support the usual on-premise applications and databases, as well as SaaS and cloud based ones
- Like other forms, Cloud DI demands a staging area
  - Be sure the SaaS or cloud license enables this
- A lot of Cloud DI support is mostly batch
  - You also need real-time, on-demand access
BEWARE!

Be sure your Cloud is optimized for Data Mgt

• Many clouds are optimized for specific applications and their servers
• Application servers and data servers have different requirements for cloud technical resources
  – Also different requirements for domain expertise as applied to consulting and professional services
• Look for a cloud that’s built specifically for data management
  – Otherwise, performance and functionality may be poor
  – Support and services could be light on DM
• Related Issue, when there’s Cloud and On-Premises versions
  – Be sure both are of equal functionality
  – Sometimes, cloud version is a dumbed-down subset
The Future of Cloud DI

• Usage will increase as users get comfortable with
  – Data security and privacy issues
  – Moving data in/out of clouds
• Some organizations will adopt Cloud DI due to
  – Integration with Cloud-based Apps
  – Need for Scalable Data Hubs
• Some orgs already have cloud for app servers & storage
  – At a lifecycle stage where it’s time to take DI to clouds
• Clouds add complexity to your mix of platform types
  – Expect a mixture of various cloud types, plus Web and traditional
• TDWI expects many DI & BI/DW tool/platforms on clouds
  – Should be common within a few years
  – Eventually, clouds will be permanent infrastructure
    • Even in BI/DW, DI, Data Mgt, etc.
SUMMARY and CONCLUSIONS

What to look for in Cloud & other Hosted Solutions for Data Mgt

- Products designed for cloud, not just ported to it
- Built/optimized for data and its processing (not apps)
- Data management in the cloud, but on premises, too
- Complex and advanced data management practices addressed, not just lowest common denominator
- SOA for easy embedding, custom web services, message handling, methods for orchestration
- Demand proof of large technology investment in provider’s cloud
- Smart people managing the cloud platform for scalability, high performance and high availability
- Domain experts for in-cloud help designing DM solutions
- Professional services, consulting, managed services
- Support for many data standards, especially B2B ones
- Support for common packaged applications and other platforms
ADAPTIVE INTEGRATION AND ALLOY

Madhukar Kumar

September 2015
MARKET FORCES
Data trends continue to pull the two worlds apart

- Deconstruction of monoliths
- New and emerging data sources
- Rapid innovation and democratization of Data Science

According to industry analysts- by 2017, the number of Citizen Data Scientists will grow 5x the rate of specialized data scientists
INTEGRATION IS GETTING COMMODITIZED

Two ends of the spectrum

Put the internet to work for you.

What is IFTTT?
IFTTT gives you creative control over the products and apps you love.

DO RECIPE
- Do hue
  - Turn on or off your lights with a tap

IF RECIPE
- If I post a picture on Instagram, save the photo to Dropbox

What are Recipes?
Recipes are simple connections between products and apps. There are two types of Recipes: Do Recipes and IF Recipes.

Do Recipes
Do Recipes run with just a tap and enable you to create your own personalized Button, Camera, and Notepad. The Do apps are available for iOS and Android.
... AND GETTING MORE COMPLEX
Major challenges in the industry

- Continued deconstruction of the ERP and monolithic systems
- IoTs, Microservices, APIs and container based approaches becoming mainstream
- Rising cost of integration and development resources
THE CHANGING LANDSCAPE OF INTEGRATION

Before

• Changes to integration every 2 to 3 years
• 50 integration patterns
• Integration and Data management as two COEs
• Integration as part of network logic (ESBs)
• Mostly Pub/Sub strategy

Today

• Changes to integration every month
• Over 200+ integration patterns
• The need to look at data as part of integration
• Microservices with smart end points and dumb pipes
• API driven
DM HAS AN INTEGRATION PROBLEM AS WELL
Data Management requires a data integration perspective

Data Governance
- Establish Rules
- Review & Resolve
- ERRORS decreasing over time

Data Stewards
- Validation
- Decisions
- REPORTS

Source Systems
- Data Extraction
- Audit & Cleansing
- Matching & Merging
- Conversion & Load
- Repeat

Data integration process steps required to make MDM a complete solution

APPLY CHANGES AND UPDATES TO SOURCE SYSTEMS
TRULY DATA CENTRIC APPROACH TO INTEGRATION
Persist the data at integration time

Typical integration architecture

dPaaS Methodology

Hybrid endpoints

Applications

Businesses
Three key components

Data Management

Built on Big Data technologies, this module uses **Polyglot persistence** and provides easy to use data virtualization layer. Also has data governance tools and data quality tools like match and merge, de-duplication, etc.

Integration

Robust messaging backbone that enables seamless service orchestration and provides open interfaces to inter connect with eco systems. Based on **microservices** architecture.

Data Visibility

Web based dashboard to provide visibility into the data flow and data heuristics like data profile. Powerful self service tools for schema on read.
BUILT USING MICROSERVICES
Built for adaptive integrations

- On demand scalability and elasticity
- Faster innovation
- Best of the technology stack
- Agnostic to consumption layer
- Modular and flexible
- Smart end points and dumb pipes
AN EXAMPLE OF USING THIS ARCHITECTURE

Using ALLOY for Data analytics pattern

End Users

BI Tool

JDBC / ODBC

Oracle DB

ALLOY App

ALLOY App

ALLOY App

API Management Layer

micro service

micro service

micro service

micro service

micro service

micro service

ALLOY App

ALLOY App

ALLOY App

Other Sources

RDBMS

CRM Apps

Raw Data

Pre-processor

Immutable Log (Kafka)

ALLOY

Graph (ArangoDB)

materialized view

Log

Workflow

Workflow

Document Log (MapR)

PostgreSQL

Raw Data

Pre-processor

Document Log (MapR)

PostgreSQL

Graph (ArangoDB)

materialized view

Log

Workflow

Workflow

ALLOY App

ALLOY App

ALLOY App

Other Sources

RDBMS

CRM Apps

Raw Data

Pre-processor

Immutable Log (Kafka)

ALLOY

Graph (ArangoDB)

materialized view

Log

Workflow

Workflow

Document Log (MapR)

PostgreSQL

End Users

BI Tool

JDBC / ODBC

Oracle DB

ALLOY App

ALLOY App

ALLOY App

Other Sources

RDBMS

CRM Apps

Raw Data

Pre-processor

Immutable Log (Kafka)

ALLOY

Graph (ArangoDB)

materialized view

Log

Workflow

Workflow

Document Log (MapR)

PostgreSQL

JDBC / ODBC

Oracle DB

ALLOY App

ALLOY App

ALLOY App

Other Sources

RDBMS

CRM Apps

Raw Data

Pre-processor

Immutable Log (Kafka)

ALLOY

Graph (ArangoDB)

materialized view

Log

Workflow

Workflow

Document Log (MapR)

PostgreSQL

Oracle DB

JDBC / ODBC

End Users

BI Tool

Oracle DB

JDBC / ODBC

End Users

BI Tool

JDBC / ODBC

Oracle DB

JDBC / ODBC

End Users

BI Tool
THANK YOU

@LiaisonTech  |  www.liaison.com
Questions?
Contact Information

If you have further questions or comments:

Philip Russom, TDWI
prussom@tdwi.org

Madhukar Kumar, Liaison
MKumar@liaison.com