Service Oriented BI

Alberto Abelló & Oscar Romero
Dept. Enginyeria de Serveis i Sistemes d’Informació
Universitat Politècnica de Catalunya
Service definition

“Services are economic activities offered by one party to another, most commonly employing time-based performances to bring about desired results in recipients themselves or in objects or other assets for which purchasers have responsibility. In exchange for their money, time and effort, service customers expect to obtain value from access to goods, labor, professional skills, facilities, networks, and systems; but they do not normally take ownership of any of the physical elements involved.”

Lovelock & Wright

Unified services theory

“All managerial themes unique to services are founded in customers providing significant inputs into the production process.”

Sampson
Past seminars (Sep’09-Jul’12)

- Antoni Olivé: Què és un Servei?
- Ruth Raventós: Service Management
- Enric Mayol, Tomàs Aluja: Service Is Front Stage
- Ernest Teniente: Reaching the Goal
- Joan Antoni Pastor: CMMI for Services
- Pere Botella: Plataformes i iniciatives en serveis
- Albert Abelló: Understanding Service Businesses. Applying Principles of the Unified Services Theory (UST)
- Antoni Olivé: Introducció a la tecnologia RFID i a la seva aplicació en hospital
- Enric Martínez: SOA: Service Oriented Architecture
- Carles Farré: Web Usability
- Albert Abelló: Everything you always wanted to know about Web services but were afraid to ask
- Ferran Martí: Bruc introducció a MIL
- Jordi Pradel: SOA des de les trinxeres
- Antonio Valle: Modelo de Codes. Caso práctico de aplicación
- Jordi Tores: The changing role of the computer scientist in the cloud
- Carme Quer: SERVQUAL. Un instrument de mesura de la qualitat dels serveis
- Frederic Marimon: Instruments de mesura de la qualitat dels serveis on-line
- Pablo Casado: Browser + Device App Stores. Noves perspectives per al desenvolupament de la web
- Jesús Bisbal: Towards negotiable SLA-based QoS Support for Data Services
- Michael D. Myers: Conducting Critical Research in IS, and its relationship with IS Design Research
- Kazuyosi Hidaka: Service Science Research and Education in Japan
- Miquel Barceló: Per on va la recerca europea en TIC

Service layers

- Business Process as a Service
- Software as a Service
- Platform as a Service
- Infrastructure as a Service
Outline

- IaaS
  - Cloud computing
- PaaS
  - BigTable
  - MapReduce
- SaaS
  - CRM
  - SCM
- BaaS
  - SOA
  - QoS
- BI on Services
  - BPM
  - KPI

Business Process as a Service

- IBM WebSphere
- Oracle SOA suite
- webMethods
- Apache ServiceMix
- Microsoft Connected Services Framework
- Open ESB
- etc.
Software as a Service

- Salesforce.com
- Cloud9
- Cloud1
- RightNow
- Microstrategy
- Quantivo
- Oracle on Demand
- etc.

Platform as a Service

- BigTable
- SimpleDB
- SDS
- ADB
- Analytics
Infrastructure as a Service

Amazon EC2
IBM SmartCloud
Google app engine
Etc.
Electricity as a utility

Computation as a utility
Cloud computing definition

“Cloud computing is a model for enabling convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction.”

NIST (National Institute of Standards and Technology)

Characteristics

- On-demand self-service
- Broad network access
- Resource pooling
- Rapid elasticity
- Measured service
What’s new

- Illusion of infinite resources
- Elimination of up-front commitment
- Pay-per-use

A cow or bottled milk?

Buy a cow
- High upfront investment
- High maintenance cost
- Produces a fixed amount
- Stepwise scaling

Buy bottled milk
- Pay-per-use
- Lower maintenance cost
- Linear scaling
- Fault-tolerant

Daniel Abadi analogy
Providers’ Challenges

- Deployment
  - Localization
  - Routing
  - Authentication
- Tuning
  - Placement
  - Resource partitioning
  - Service level objectives
  - Dynamically varying workloads

Obstacles/Opportunities

- Availability of service
- Data lock-in
- Data confidentiality
- Data transfer bottlenecks
- Performance unpredictability
- Scalable storage
- Debugging
- Scaling quickly
- Reputation fate sharing
- Software licensing
Platform as a Service (PaaS)

Kinds of platforms

- Storage as a Service
- Database as a Service
- Processes (Workflow) as a Service
- Service as a Service
- Security as a Service
- Management/Governance as a Service
Kinds of software

- Software in the cloud
  - DBMS
  - Workflow management
  - Versioning
- Cloud software
Open source platform

- Hadoop
  - Based on Google MapReduce (2004)

- Hbase (also Dynamo and Cassandra)
  - Based on Google BigTable (2006)

MapReduce

- Processes pairs [key, value]
- Hides parallelization, fault-tolerance, data distribution and load balancing
WordCount code example

```java
public void map(LongWritable key, Text value) {
    String line = value.toString();
    StringTokenizer tokenizer = new StringTokenizer(line);
    while (tokenizer.hasMoreTokens()) {
        write(new Text(tokenizer.nextToken()), new IntWritable(1));
    }
}

public void reduce(Text key, Iterable<IntWritable> values) {
    int sum = 0;
    for (IntWritable val : values) {
        sum += val.get();
    }
    write(key, new IntWritable(sum));
}
```

WordCount execution example

![Diagram of MapReduce workflow]

Map: 1
Reduce: 1
The: 57631
**Benefits of MapReduce**

- Programming model simple yet expressive
  
  \[
  \text{map}(\text{key } k, \text{value } v) \mapsto [(i k_1, i v_1), \ldots, (i k_m(k,v), i v_m(k,v))]
  \]

  \[
  \text{reduce}(\text{key } i k, \text{vset } i v) \mapsto [o v_1, \ldots, o v_r(i k, i v)]
  \]

- Able to process structured or unstructured
- Elastically scalable
- Fine grained fault tolerance

**MapReduce performance problems**

- Does not benefit from compression
- Writes intermediate results to disk
  - Reduce tasks pull intermediate data
- Defines the execution plan on the fly
  - Schedules one block at a time
BigTable

(row:string, column:string[, time:int64]) → string

Physical implementation

- Supports single row transactions
- Compression per block can be enabled
- The schema determines the locality of data
Just another point of view?

Relational

Child → Parent → 1NF

NOSQL

Child ← Parent → NF²

NOSQL drawbacks

- No ACID
- No standard
- Low-level query

Michael Stonebraker
Do It Yourself
• Expensive
• Ad hoc development

Off the Shelf
• Economies of scale
• Concrete functionalities

… But to really unlock the power of Hadoop, you must be able to efficiently extract data stored across multiple (often tens or hundreds) of nodes with a user-friendly ETL (extract, transform and load) tool that will then allow you to move your Hadoop data into a relational data mart or warehouse where you can use BI tools for analysis. “

Ian Fyfe
Pentaho
Reference Architecture

ETL (Extraction, Transformation and Load)

BigTable

MapReduce

WWW

Trademark

- Expensive
- Many functionalities
- Mature

Open source

- Free
- Simple functionalities

Big elephant or little elephant?

Huge

Adaptive
Gartner’s market analysis

<table>
<thead>
<tr>
<th>Service %</th>
</tr>
</thead>
<tbody>
<tr>
<td>35%</td>
</tr>
<tr>
<td>30%</td>
</tr>
<tr>
<td>25%</td>
</tr>
<tr>
<td>20%</td>
</tr>
<tr>
<td>15%</td>
</tr>
<tr>
<td>10%</td>
</tr>
<tr>
<td>5%</td>
</tr>
<tr>
<td>0%</td>
</tr>
</tbody>
</table>

Gartner’s considerations

<table>
<thead>
<tr>
<th></th>
<th>On-premises</th>
<th>Service-based</th>
</tr>
</thead>
<tbody>
<tr>
<td>Customization</td>
<td>+</td>
<td>+/-</td>
</tr>
<tr>
<td>Implementation time</td>
<td>+/-</td>
<td>+/-</td>
</tr>
<tr>
<td>Application shut-off</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>Hidden fees</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Security of data</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>Process integrity</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>Guarantee of quality</td>
<td>+</td>
<td>-</td>
</tr>
</tbody>
</table>
Business processes as a Service (BaaS)

- Business processes as a Service are focused on providing existing business processes through a cloud. If there is an existing process with steps that are known it can be provided as a service within the catalog. This allows the Cloud Service Provider to automate any steps within the process while leaving the changes transparent to the Cloud Service Consumer.

- Software Services allow a Cloud Service Consumer to select a specific software instance that they want created without the need to be aware of where and how it will be hosted. This allows the Cloud Service Consumer to focus on the characteristics of the application and gives the Cloud Service Provider the freedom to fulfill the request with any resources that will meet the need.

SaaS vs BaaS

- NIST (National Institute of Standards and Technology)
SOA principles

- Reusability
- Loose coupling
- Contract
- Abstraction
- Composability
- Autonomy
- Statelessness
- Discoverability

Paradigm

<table>
<thead>
<tr>
<th></th>
<th>Distributed components</th>
<th>SOA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Design</td>
<td>Functionality</td>
<td>Process</td>
</tr>
<tr>
<td>Designed to ...</td>
<td>Last</td>
<td>Change</td>
</tr>
<tr>
<td>Development cycle</td>
<td>Long</td>
<td>Interactive and iterative</td>
</tr>
<tr>
<td>Centered on ...</td>
<td>Cost</td>
<td>Business</td>
</tr>
<tr>
<td>Coordination</td>
<td>Blocks</td>
<td>Orchestration</td>
</tr>
<tr>
<td>Coupling</td>
<td>Tight</td>
<td>Loose (agile and adaptive)</td>
</tr>
<tr>
<td>Technologies</td>
<td>Homogeneous</td>
<td>Heterogeneous</td>
</tr>
<tr>
<td>Programming</td>
<td>Objects</td>
<td>Messages</td>
</tr>
<tr>
<td>Encapsulation</td>
<td>Partial</td>
<td>Full (contracts)</td>
</tr>
</tbody>
</table>
Service composition

- Primitive activity
- Complex activity
  - Atomic transaction
  - Business activity
    • Orchestration
    • Choreography

Technological Challenges

- Business Process Management
- Service Composition
- Service Infrastructure and Management

Business Processes

Business Services

Utility Services

DB | MapReduce | SCM | CRM | MDM | ERP
Implementation

- Client
  - findService
  - description (WSDL)
  - request (SOAP)
  - response (SOAP)

- Service broker (UDDI)
  - register

- Service provider
  - Business object

Engineering challenges

- Service Engineering and Design
- Service Adaptation and Monitoring
- Service Quality
Reference Architecture

ETL (Extraction, Transformation, and Load)

WWW

Business Intelligence on Services
Customer Relationship Management

- Turns manufacturers into services
- Becomes a barrier for competitors
- By coding customers allows to:
  - Instruct staff
  - Manage queues in call centers
  - Target offers
  - Sell data to other firms

SOA lifecycle

[Diagram showing SOA lifecycle with steps: Modelar, Gestionar, Ensamblar, Desplegar, Gobierno y mejores prácticas]
Key Performance Indicators

“You cannot control what you cannot measure”
IBM WebSphere architecture

IBM e-Xtend World Bank demo
Users in e-Xtend World Bank

Loan manager reviews loan status
Loan officer creates a new loan request

Senior underwriter works in loan request
Loan manager checks updates

Conclusions

- BI can benefit from services at four levels
  - IaaS
  - PaaS
  - SaaS
  - BaaS

- Services benefit from BI
  - KPI and Balance Scorecards
  - Process mining
Bibliography