

# Service Oriented BI

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



European Master's Programme  
in INFORMATION TECHNOLOGIES  
for BUSINESS INTELLIGENCE



Alberto Abelló & Oscar Romero 05/07/2011 2

## 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

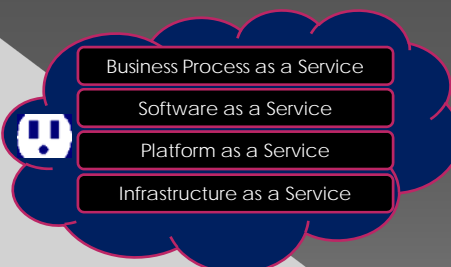
## Past seminars (Sep'09-Jul'12)

- Antoni Olivé: Què és un Servei?
- Ruth Raventós: Service Management
- Enric Mayol, Tomás 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
- Antonio Villegas: Everything you always wanted to know about Web services but were afraid to ask
- Ferran Martí: Breu Introducció a ITIL
- Jordi Pradel: SOA des de les trinxeres
- Antonio Valle: Modelo de Costes. Caso práctico de aplicación
- Jordi Torres: 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 online
- 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

Alberto Abelló & Oscar Romero 05/07/2011

5

## Service layers



Alberto Abelló & Oscar Romero 05/07/2011

6

## Outline

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

Alberto Abelló & Oscar Romero 05/07/2011

7

## Business Process as a Service

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



Alberto Abelló & Oscar Romero 05/07/2011

8

## Software as a Service

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



Alberto Abelló & Oscar Romero 05/07/2011

9

## Platform as a Service




BigTable  
SimpleDB  
ElasticMapReduce  
Amazon DynamoDB  
Amazon ElastiCache  
Amazon EMR  
Amazon Redshift  
Amazon S3  
Amazon Glacier  
Amazon CloudFront  
Amazon CloudWatch  
Amazon CloudTrail  
Amazon CloudSearch  
Amazon CloudFormation  
Amazon CloudWatch Logs  
Amazon CloudWatch Events  
Amazon CloudWatch Alarms  
Amazon CloudWatch Insights  
Amazon CloudWatch Logs Insights

Alberto Abelló & Oscar Romero 05/07/2011

10

# Infrastructure as a Service



- Amazon EC2
- IBM SmartCloud
- Google app engine
- Etc.

Alberto Abelló & Oscar Romero 05/07/2011 11

# Infrastructure as a Service IaaS

Alberto Abelló & Oscar Romero 05/07/2011 12

## Electricity as a utility



Alberto Abelló & Oscar Romero 05/07/2011

13

## Computation as a utility



Alberto Abelló & Oscar Romero 05/07/2011

14

## 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)

Alberto Abelló & Oscar Romero 05/07/2011

15

## Characteristics

- On-demand self-service
- Broad network access
- Resource pooling
- Rapid elasticity
- Measured service

Alberto Abelló & Oscar Romero 05/07/2011

16



## What's new

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

Alberto Abelló & Oscar Romero 05/07/2011

17

## A cow or bottled milk?

### Buy a cow

- High up-front investment
- High maintenance cost
- Produces a fixed amount
- Difficult to scale

**Security**



### Buy bottled milk

- Pay-per-use
- Lower maintenance cost
- Easy to scale
- Tolerant

**Low Cost**

Daniel Abadi analogy

Alberto Abelló & Oscar Romero 05/07/2011

18

## Providers' Challenges

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

Alberto Abelló & Oscar Romero 05/07/2011

19

## 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

Alberto Abelló & Oscar Romero 05/07/2011

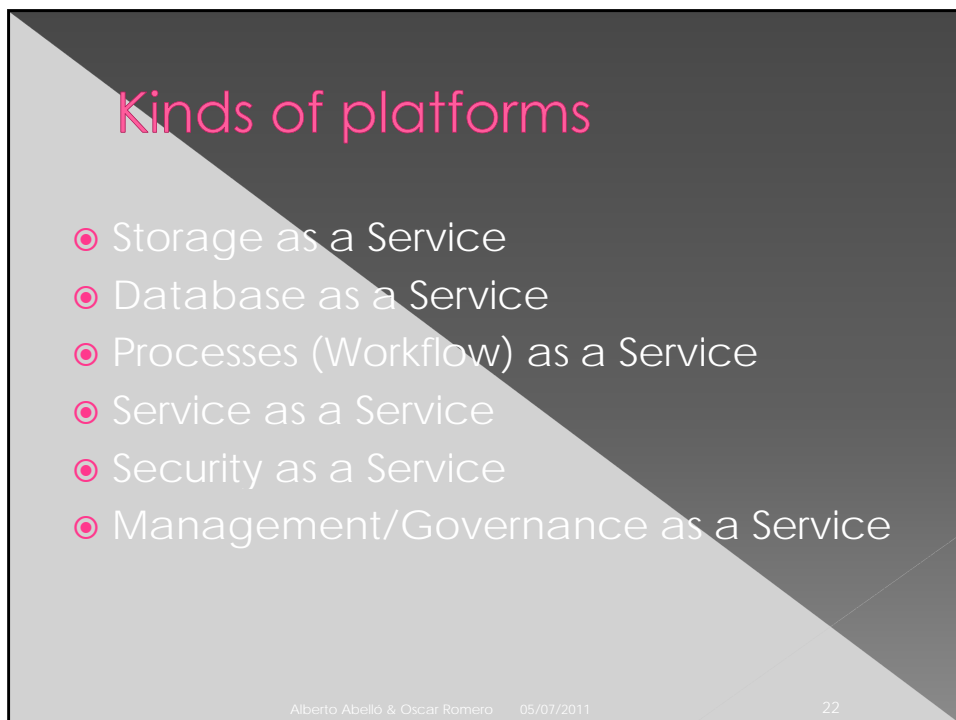
20



Platform as a Service  
PaaS

Alberto Abelló & Oscar Romero 05/07/2011 21

This slide features a dark grey background with a diagonal split into a lighter grey section on the left. The title 'Platform as a Service PaaS' is centered in a blue, 3D-style font. At the bottom, there is a small footer with the authors' names, the date, and the slide number.

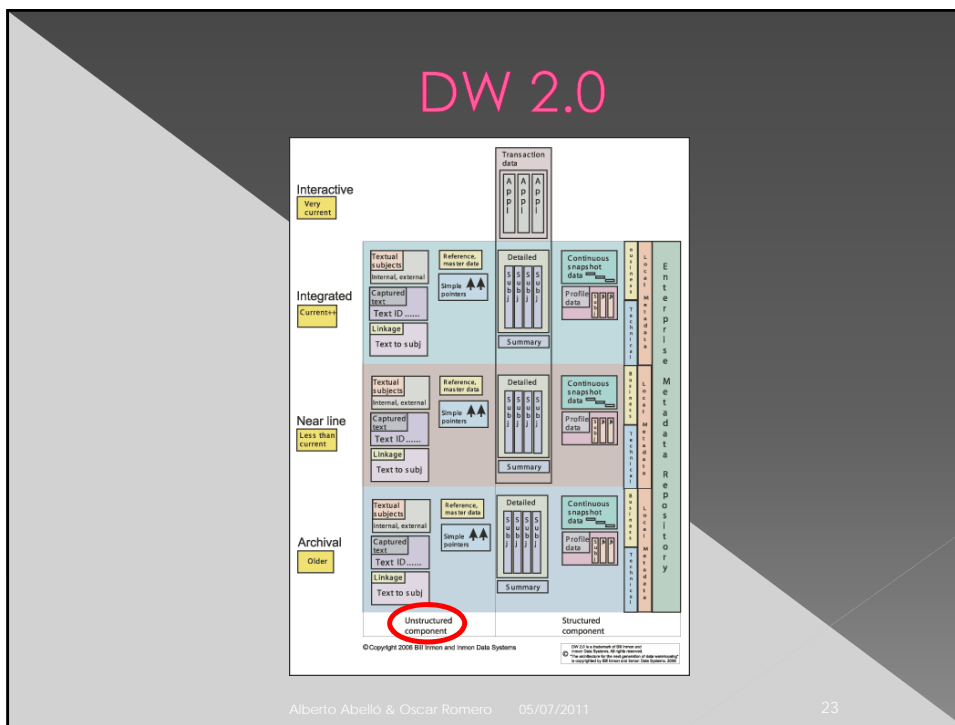


Kinds of platforms

- ◉ 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

Alberto Abelló & Oscar Romero 05/07/2011 22

This slide has the same diagonal background as slide 21. The title 'Kinds of platforms' is in a pink font. Below it is a list of six items, each preceded by a pink circle with a white dot. The footer at the bottom contains the authors' names, the date, and the slide number.



- ## Kinds of software
- ◉ Software in the cloud
    - > DBMS
    - > Workflow management
    - > Versioning
  - ◉ Cloud software
- Alberto Abelló & Oscar Romero    05/07/2011    24

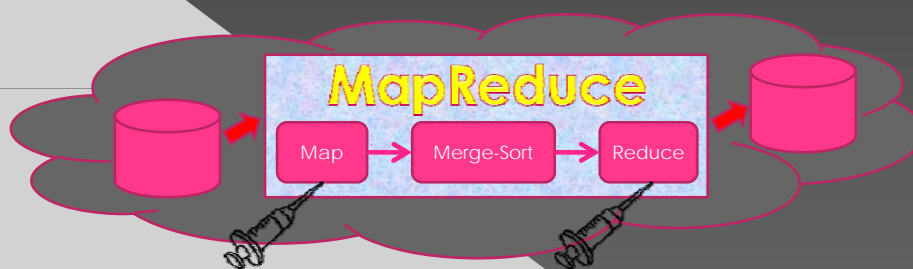
## Open source platform

- Hadoop
  - > Based on Google MapReduce (2004)
- Hbase (also Dynamo and Cassandra)
  - > Based on Google BigTable (2006)

Alberto Abelló &amp; Oscar Romero 05/07/2011

25

## MapReduce



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

Alberto Abelló &amp; Oscar Romero 05/07/2011

26

## WordCount code example

```
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));
}
```

Alberto Abelló &amp; Oscar Romero 05/07/2011

27

## WordCount execution example

```
The Project Gutenberg eBook of The Outline of Science, Vol. 1 (of 4), by
J. Arthur Thomson

This eBook is for the use of anyone anywhere at no cost and with
almost no restrictions whatsoever. You may copy it, give it away or
re-use it under the terms of the Project Gutenberg License included
with this eBook or online at www.gutenberg.org

Title: The Outline of Science, Vol. 1 (of 4)
A Plain Story Simply Told
Author: J. Arthur Thomson
Release Date: January 22, 2007 [EBook #20417]
Language: English
Character set encoding: ASCII
*** START OF THIS PROJECT GUTENBERG EBOOK OUTLINE OF SCIENCE ***

Produced by Brian James, Leonard Johnson and the Dallas
Distributed Proofreading Team at http://www.pgdp.net

[Illustration: THE GREAT SCARLET SOLAR PROMINENCES, WHICH ARE SUCH A
NOTABLE FEATURE OF THE SOLAR PHENOMENA, ARE IRENEUS OUTBURSTS OF FLAMING
HYDROGEN RISING SOMETIMES TO A HEIGHT OF 500,000 MILES.]

THE
OUTLINE OF SCIENCE
A PLAIN STORY SIMPLY TOLD

EDITED BY
J. ARTHUR THOMSON
REGIUS PROFESSOR OF NATURAL HISTORY IN THE
UNIVERSITY OF ABERDEEN

WITH OVER 100 ILLUSTRATIONS
OF WHICH ABOUT 40 ARE IN COLOUR
```

Map

Reduce

```
The 1
Project 1
Gutenberg 1
eBook 1
of 1
The 1
Outline 1
Of 1
Science 1
Vol. 1
1 1
(of 1
4), 1
by 1
```

The

57631

Alberto Abelló &amp; Oscar Romero 05/07/2011

28

## Benefits of MapReduce

- ◉ Programming model simple yet expressive

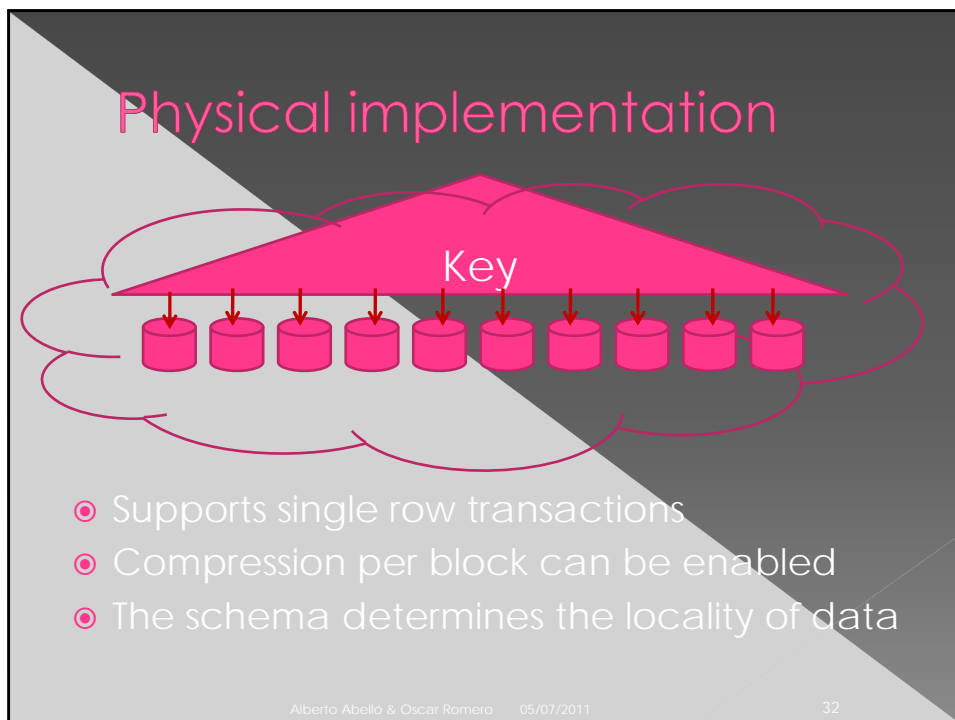
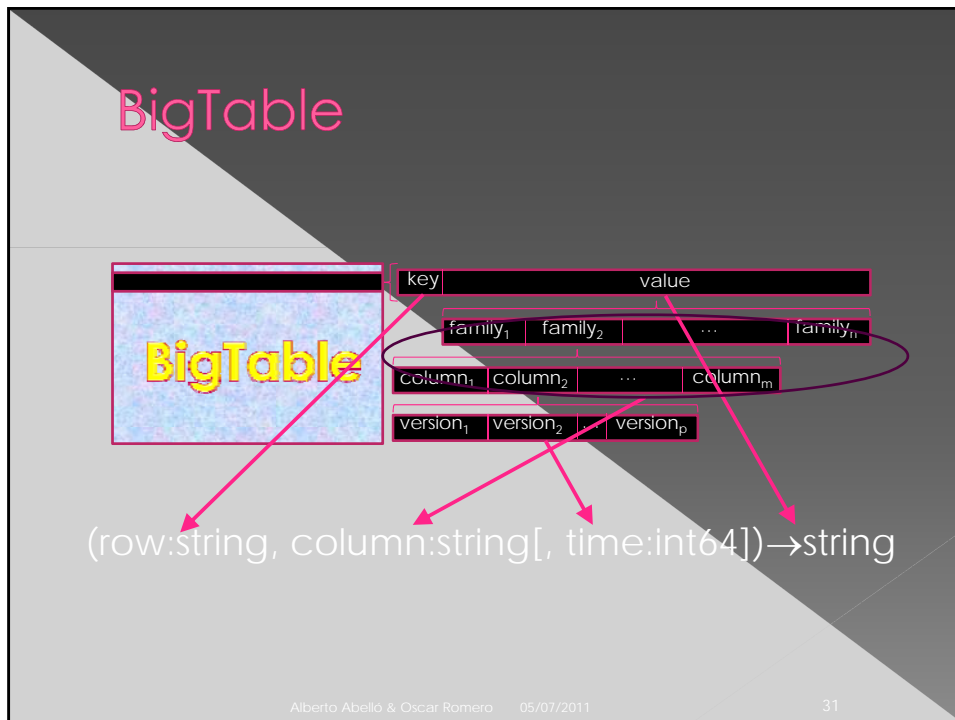
`map(key  $k$ , value  $v$ )  $\mapsto$   $[(ik_1, iv_1), \dots, (ik_{m(k,v)}, iv_{m(k,v)})]$`

`reduce(key  $ik$ , vset  $ivs$ )  $\mapsto$   $[ov_1, \dots, ov_{r(ik,ivs)}]$`

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

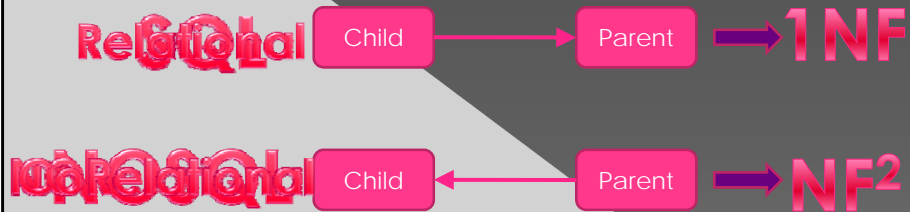
## MapReduce performance problems

- ◉ 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





## Just another point of view?



Alberto Abelló & Oscar Romero 05/07/2011

33

## NOSQL drawbacks

- No ACID
- No standard
- Low-level query

Michael Stonebraker

Alberto Abelló & Oscar Romero 05/07/2011

34

## Brewery or bottled beer?

**Do It Yourself**

- Expensive
- Ad hoc development

**Off the Shelf**

- Economies of scale
- Concrete functionalities

Florian Waas analogy

Specificity

Low Cost

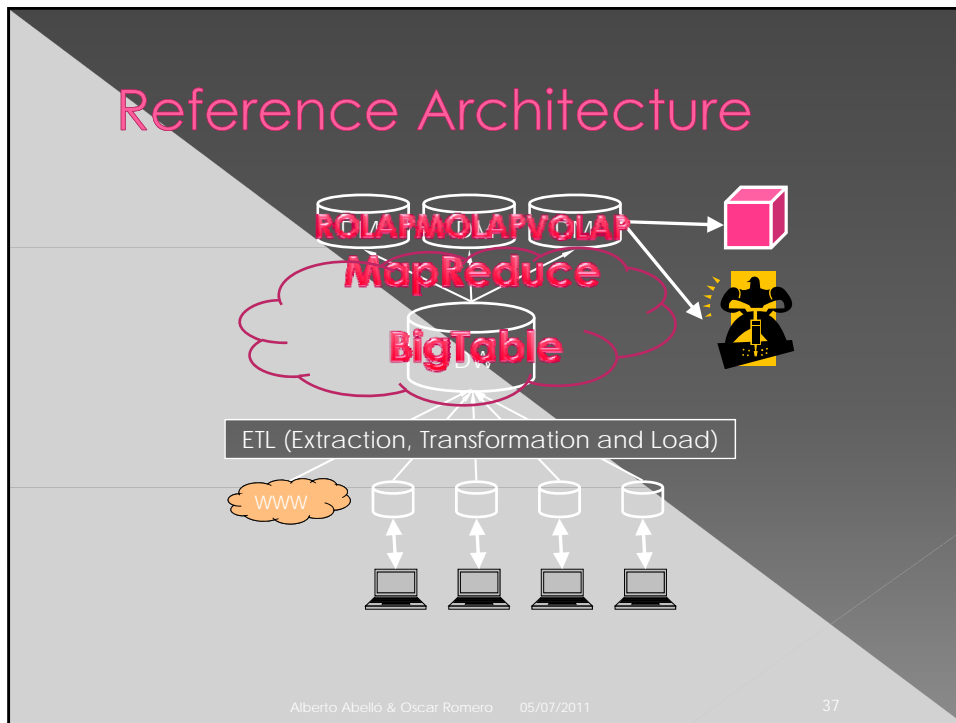
Alberto Abelló & Oscar Romero 05/07/2011 35

## Market tools

“ ... 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*

Alberto Abelló & Oscar Romero 05/07/2011 36



## Big elephant or little elephant?

**Huge**

**Trademarks**

- Expensive
- Many functionalities
- Complex

**ORACLE**

**Adaptative**

**Open source**

- Free
- Simple functionalities
- Flexible

**hadoop**

Alberto Abelló & Oscar Romero 05/07/2011 38

# Friends or foes

The diagram illustrates the relationships between several technologies. At the top, three pink rounded rectangles labeled 'MapReduce', 'Java', and 'Oracle' are arranged horizontally. Below them is a single red rounded rectangle labeled 'Oracle Grid Engine'. At the bottom is a grey rounded rectangle labeled 'Hadoop Distributed File System (HDFS)'. The background features a diagonal split between light grey and dark grey.

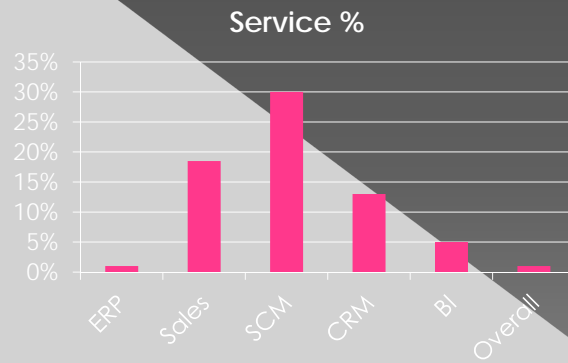
Alberto Abelló & Oscar Romero 05/07/2011 39

# Software as a Service SaaS

The slide features the text 'Software as a Service' in a blue, 3D-style font, with 'SaaS' below it in a similar style. The background is a diagonal split between light grey and dark grey.

Alberto Abelló & Oscar Romero 05/07/2011 40

## Gartner's market analysis



Alberto Abelló & Oscar Romero 05/07/2011

41

## Gartner's considerations

	On-premises	Service-based
Customization	+	+/-
Implementation time	+/-	+/-
Application shut-off	+	-
Hidden fees	-	-
Security of data	+	-
Process integrity	+	-
Guarantee of quality	+	-

Alberto Abelló & Oscar Romero 05/07/2011

42

**Business processes as a Service**  
**BaaS**

Alberto Abelló & Oscar Romero 05/07/2011 43

## SaaS vs BaaS

- Business Process as a Services 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.

NIST (National Institute of Standards and Technology)

Alberto Abelló & Oscar Romero 05/07/2011 44

## SOA principles

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

Alberto Abelló &amp; Oscar Romero 05/07/2011

45

## Paradigm

	Distributed components	SOA
Design	Functionality	Process
Designed to ...	Last	Change
Development cycle	Long	Interactive and iterative
Centered on ...	Cost	Business
Coordination	Blocks	Orchestration
Coupling	Tight	Loose (agile and adaptive)
Technologies	Homogeneous	Heterogeneous
Programming	Objects	Messages
Encapsulation	Partial	Full (contracts)

Alberto Abelló &amp; Oscar Romero 05/07/2011

46

## Service composition

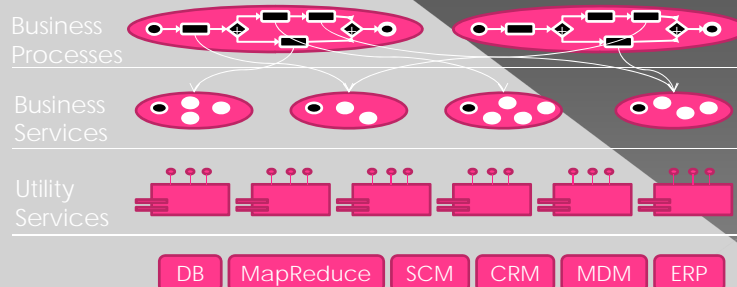
- Primitive activity
- Complex activity
  - > Atomic transaction
  - > Business activity
    - Orchestration
    - Choreography

Alberto Abelló &amp; Oscar Romero 05/07/2011

47

## Technological Challenges

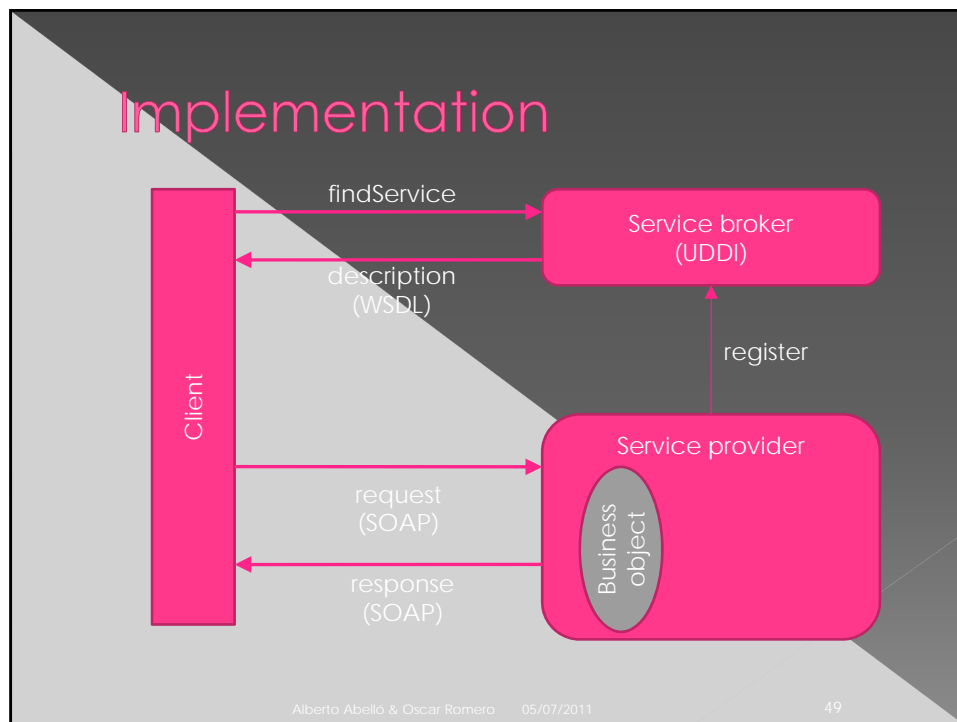
- Business Process Management
- Service Composition
- Service Infrastructure and Management

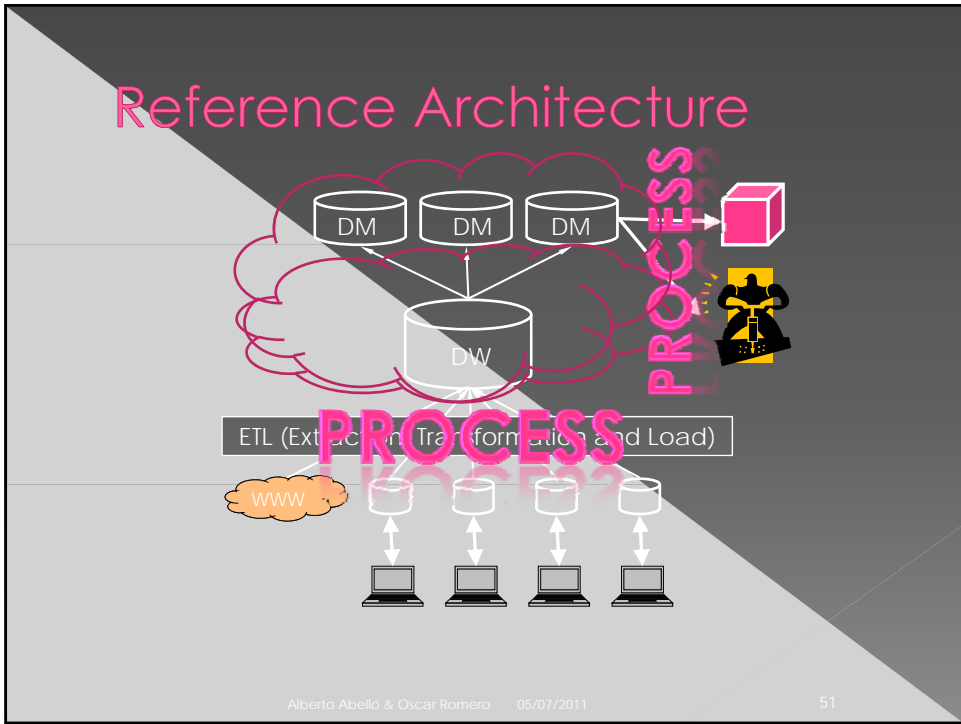


Alberto Abelló &amp; Oscar Romero 05/07/2011

48







# Business Intelligence on Services

Alberto Abelló & Oscar Romero 05/07/2011 52

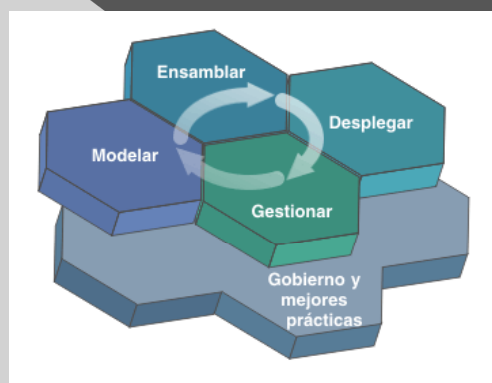
## Customer Relationship Management

- 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

Alberto Abelló &amp; Oscar Romero 05/07/2011

53

## SOA lifecycle



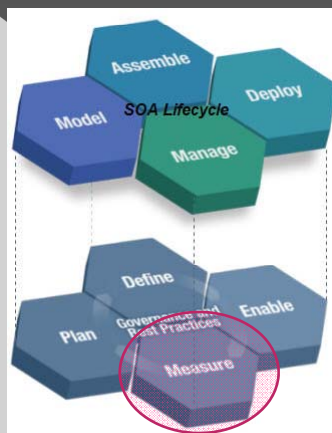
Alberto Abelló &amp; Oscar Romero 05/07/2011

54

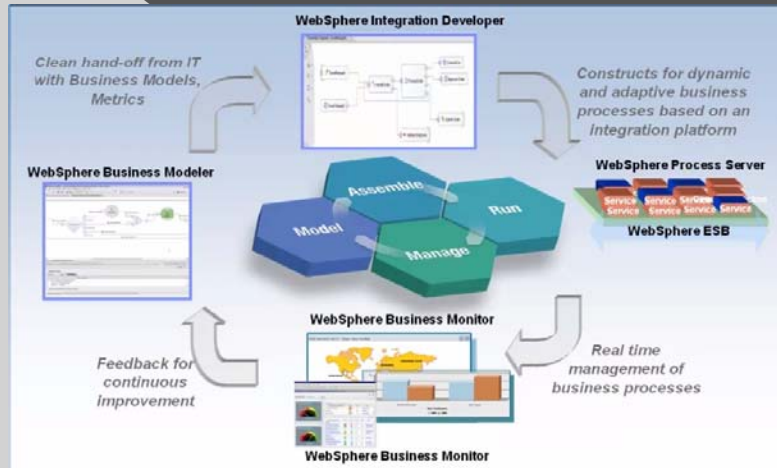
# Key Performance Indicators

"You cannot control what you cannot measure"

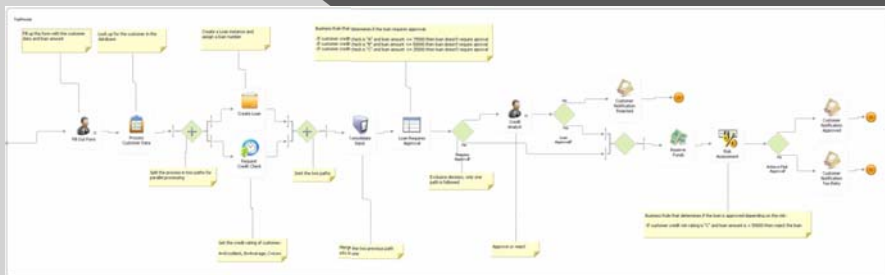
# SOA Governance



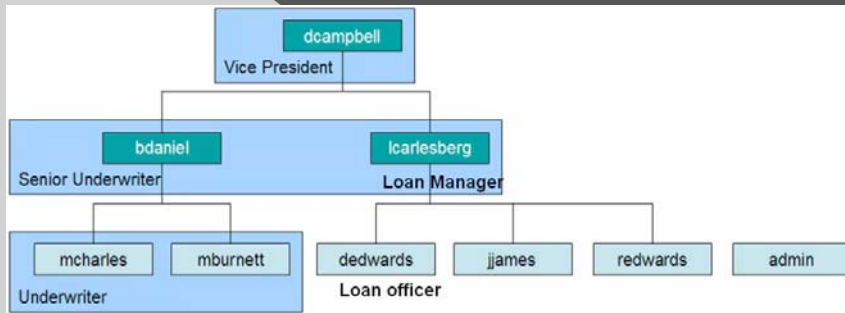
# IBM WebSphere architecture



# IBM e-Xtend World Bank demo



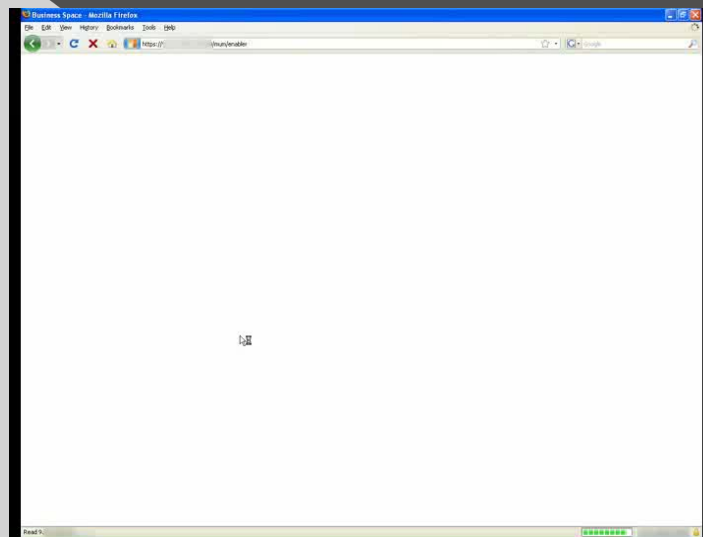
## Users in e-Xtend World Bank



Alberto Abelló & Oscar Romero 05/07/2011

59

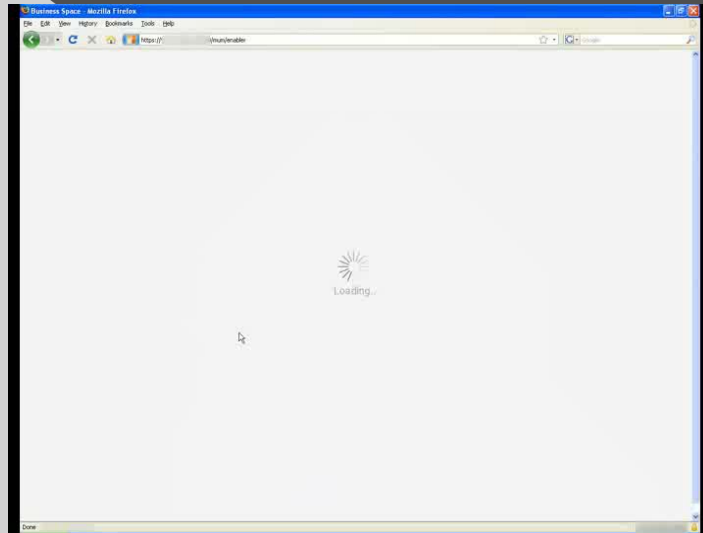
## Loan manager reviews loan status



Alberto Abelló & Oscar Romero 05/07/2011

60

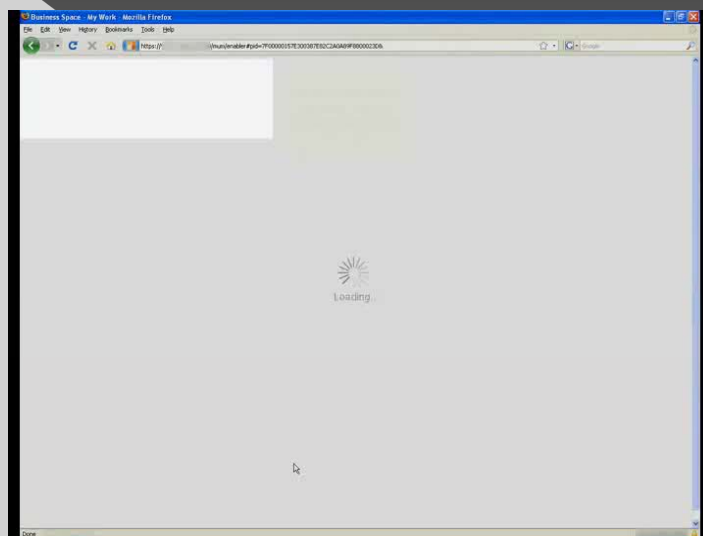
## Loan officer creates a new loan request



Alberto Abelló & Oscar Romero 05/07/2011

61

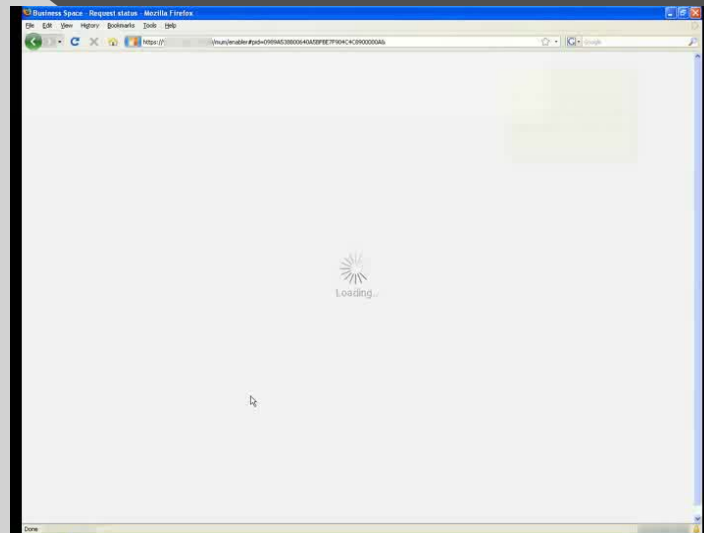
## Senior underwriter works in loan request



Alberto Abelló & Oscar Romero 05/07/2011

62

## Loan manager checks updates



Alberto Abelló & Oscar Romero 05/07/2011

63

## Conclusions

- ◉ BI can benefit from services at four levels
  - > IaaS
  - > PaaS
  - > SaaS
  - > BaaS
- ◉ Services benefit from BI
  - > KPI and Balance Scorecards
  - > Process mining

Alberto Abelló & Oscar Romero 05/07/2011

64



## Bibliography

- ◉ Mell, P., Grance, T.: The NIST Denition of Cloud Computing. Special Publication 800-145, National Institute of Standards and Technology (January 2011), draft
- ◉ Abadi, D.J.: Data management in the cloud: Limitations and opportunities. IEEE Data Engineering Bulletin 32(1), 3-12 (2009)
- ◉ Stonebraker, M., et al.: MapReduce and parallel DBMSs: friends or foes? Communication of ACM 53(1), 64-71 (2010)
- ◉ Hostmann, B.: Business Intelligence as a Service: Findings and Recommendations. Research G00164653, Gartner (January 2009)
- ◉ Erl, T.: Service Oriented Architecture. Prentice Hall (2006)
- ◉ Castellanos, M., et al.: Automating the loading of business process data warehouses. In EDBT'2009. pp. 612-623. ACM
- ◉ Josep Carmona, Jordi Cortadella: Process Mining Meets Abstract Interpretation. ECML/PKDD (1) 2010: 184-199