

# The AXML Artifact Model

Serge Abiteboul  
INRIA Saclay & ENS Cachan & U.  
Paris Sud

[Time09]

INSTITUT NATIONAL  
DE RECHERCHE  
EN INFORMATIQUE  
ET EN AUTOMATIQUE



centre de recherche  
**SACLAY - ÎLE-DE-FRANCE**

# Context: Data management in P2P systems

Large number of distributed peers

Peers are autonomous

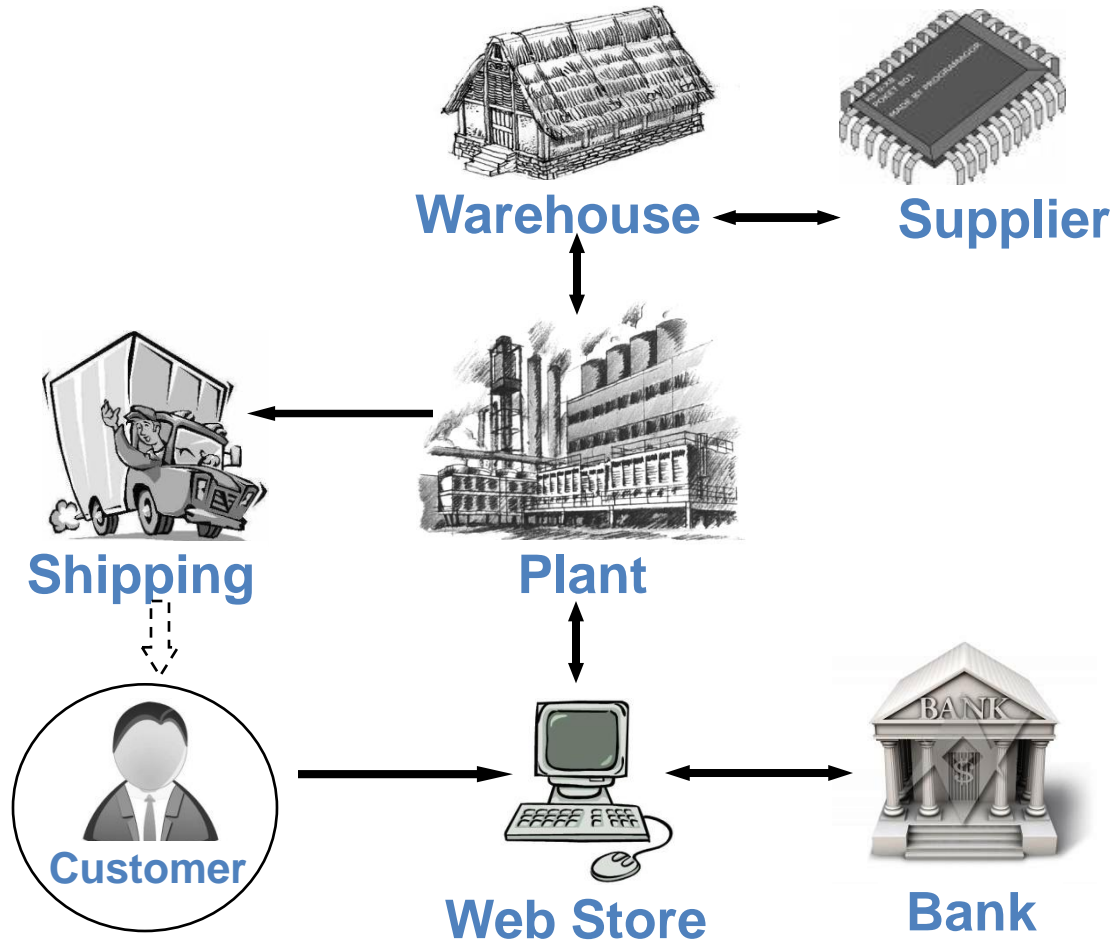
Large volume of data

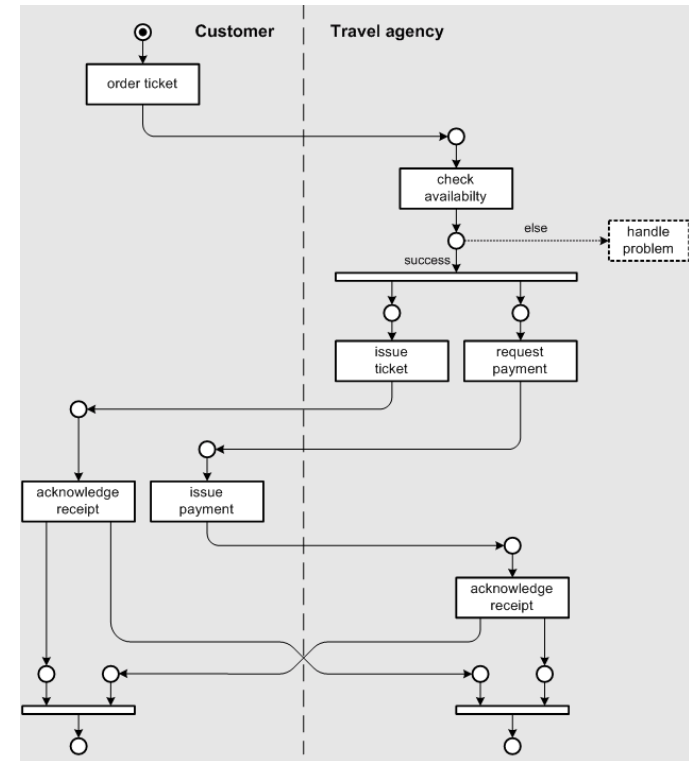
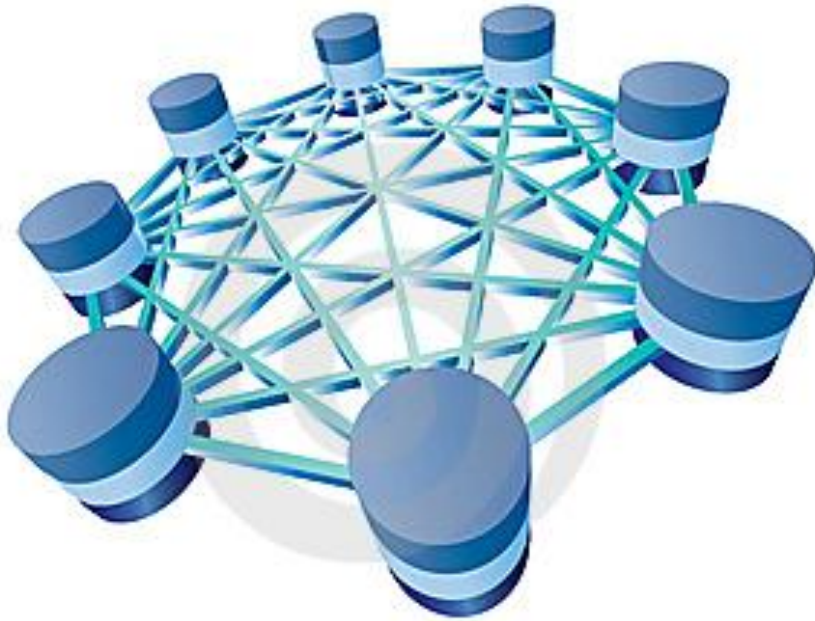
Very dynamic

- Large throughput of updates
- Intense communication between peers
- Sometimes rapidly changing network
- Sometimes moving peers

S. Abiteboul – INRIA Saclay

# Example: Dell's Supply Chain





# Data & Control

S. Abiteboul – INRIA Saclay

# Issues – compile time

## Modeling and design

- Declarative specifications, data & control (& semantics), views
- Synthesis code, existing workflow reuse, HMI...

## Reasoning

- Verification of static/dynamic properties
- Infinite state systems

S. Abiteboul – INRIA Saclay

# Issues – runtime

## Querying

- Querying the state, searching and mining, data integration and mashups
- Querying the history and provenance

## Monitoring

- Detecting events of interest – Pub/sub,
- Reacting to changes, e.g., data synchronization
- Gathering traces
- Error detection, error diagnosis, intrusion, spamming...
- Gathering statistics towards optimization and tuning
- Scheduling, feedback, prediction

Access control , concurrency control and transaction

Optimization and automatic tuning

S. Abiteboul – INRIA Saclay

# Motivation

Goal: design and run such applications

## ***Model for distributed data management***

- Taking data into account
- Taking control/workflow into account

Basis on two concepts

- ***Artifact***
- ***Active XML***

Proposal: ***Active XML Artifact Model***

S. Abiteboul – INRIA Saclay

# Organization

Introduction

Active xml

AXML Artifacts

Two particular issues

- Verification
- Monitoring

Conclusion

S. Abiteboul – INRIA Saclay



# Active xml

I suppose you have heard about it...

# Artifact

INSTITUT NATIONAL  
DE RECHERCHE  
EN INFORMATIQUE  
ET EN AUTOMATIQUE



centre de recherche  
**SACLAY - ÎLE-DE-FRANCE**

# Artifact = Data & Control

Concept introduced by IBM Research

[Nigam & Caswell 03, Hull & Su 07]

## Data-centric workflows

- A process is described by a document (possibly moving in the enterprise)
- The behavior of an artifact is specified by some constraints on how this document should evolve

## Vs. state-transition-based workflows

- Based on some form of state transition diagrams (BPEL, Petri,...)
- Mostly ignore data

**webOrder** id=7787780  
 Customer  
     Name: John Doe  
     Address: Sèvres  
 Product: *committed*  
     Ref: PC 456  
 Factory: Milano  
 Parts: *waiting*  
 orderDate: 2009/07/24  
 Site: http:// d555.com  
 Payment: *done*  
     Bank-account ...  
 Delivery: *not-active*

S. Abiteboul – INRIA Saclay

# Artifacts in action

## In webStore

**webOrder** id=7787780  
 Customer  
     Name: John Doe  
     Address: Sèvres  
 Order selection: **on-going**  
     Ref: PC 456  
 Factory: *undecided*  
 Parts: *not-active*  
 orderDate: 2009/07/24  
 Site: http://d555.com  
 Payment: *pending*  
 Delivery: *not-active*

## In plant

**webOrder** id=7787780  
 Customer  
     Name: John Doe  
     Address: Sèvres  
 Order selection : *committed*  
     Ref: PC 456  
 Factory: Milano  
 Parts: **on-going**  
 orderDate: 2009/07/24  
 Site: http:// d555.com  
 Payment: *done*  
     Bank-account ...  
 Delivery: *not-active*

## In delivery

**webOrder** id=7787780  
 Customer  
     Name: John Doe  
     Address: Sèvres  
 Order selection : *committed*  
     Ref: PC 456  
 Factory: Milano  
 Parts: *done*  
 orderDate: 2009/07/24  
 Site: http:// d555.com  
 Payment: *done*  
     Bank-account: CEIF-4457889  
 Delivery: **on-going**  
     Address: Orsay

S. Abiteboul – INRIA Saclay

# Artifact - Different levels

Can be used as a design methodology

Can be software to support certain new functionalities

Can be a compiler from a high level spec to a running system

## IBM SOMA/BELA

- Service-Oriented Modeling and Architecture
- Business Entity Lifecycle Analysis
- Available tool: specification compiled to Websphere

S. Abiteboul – INRIA Saclay

# Advantages of state-transition-based workflows

State-transition-based workflows are very intuitive for small/centralized applications

People are used to them

Reasonable complexity for static analysis because data is ignored

- Vs, artifacts: infinite state systems

S. Abiteboul – INRIA Saclay

# Shift to data-centric - advantages of artifacts

State-transition-based workflows are not appropriate for large/distributed applications

Data is at the center of the picture

- In state-transition-based workflows, database calls are crucial & not part of the model

I believe easier to provide views of a process

- Financial service view, vendor view...
- In the style of database views – not as well understood

Facilitate interoperability/integration of workflow systems

S. Abiteboul – INRIA Saclay

# Shift to declarative - another advantage

Easier to go to rule-based workflow specifications

State-transition-based workflows – procedural

- States of the system and transitions between states
- Well-adapted to synchronous tasks & centralized activities
- Very much in the spirit of **orchestration**

Rule-based workflows – declarative

- Rules specify constraints on how the document can evolve
- Better adapted to asynchronous tasks & distribution
- Somewhat reminiscent of **choreography**

Example

1. First order, then pay, then it is delivered
2. It can be delivered only after it has been paid for

S. Abiteboul – INRIA Saclay



# AXML Artifact

# AXML Artifact

AXML was designed for modeling distributed data management applications

Can we use it as a foundation for an artifact model?

Naturally captures distribution and autonomy of artifacts

Naturally captures communication of information between artifacts

Naturally captures hierarchies of tasks (subartifacts)

Naturally captures (possibly large) collections of artifacts

Need to add state-transition-based workflows to AXML

S. Abiteboul – INRIA Saclay

# AXML Artifact – artifacts revisited

[Data state] An artifact is an **object** with a universal **identity** and a **self-describing data state** that it may be transmitted or archived

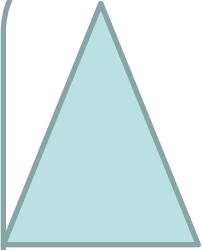
[Evolution] An artifact is created, evolves in time and space, hibernates, is reactivated or dies (archived) according to a declarative **logic**. Its evolution is constrained by some laws, **workflow** (possibly state-transition)

[Interface] An artifact publishes an **interface** and interacts with the rest of the world via **function calls**, both as **client** and **server**

[Subartifacts] An artifact may have (a collection of) **subartifacts**

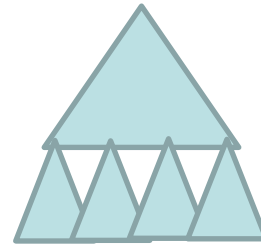
[History] An artifact has a **history** with **time, location** and **provenance** that may be recorded and queried

S. Abiteboul – INRIA Saclay



catalogue

WEBSTORE



PLANT

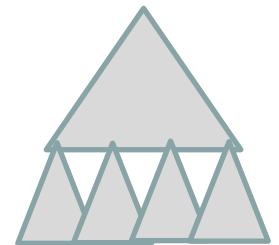
DELIVERY

# AXML Artifact model

CREDIT APPROVAL

WAREHOUSE

ARCHIVE



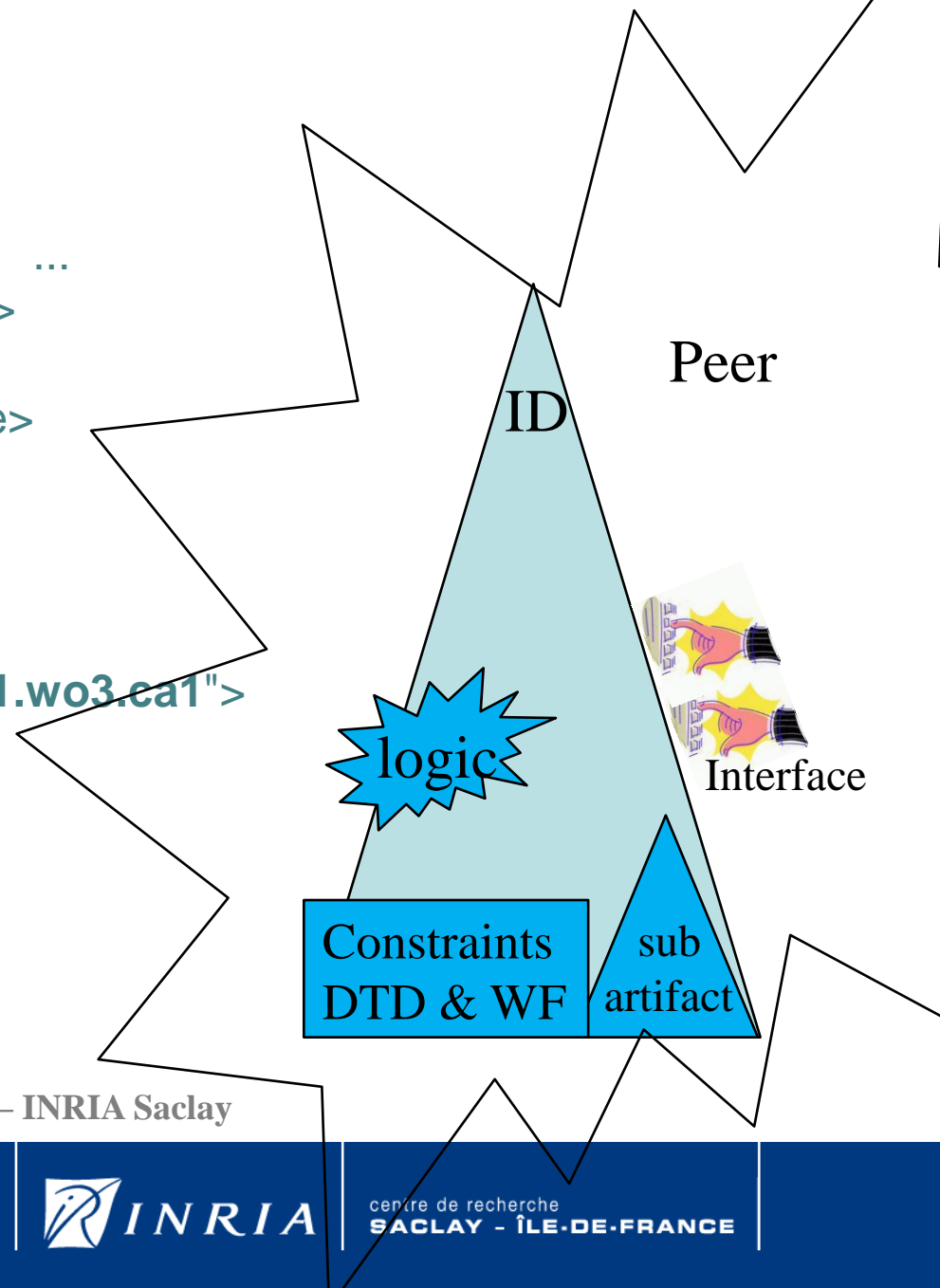
S. Abiteboul – INRIA Saclay

# AXML artifact

```

<webStore artID="store1">
  ...
  <webOrder artID="store1.wo3">
    <client>
      <name>Sue Leroux</name>
      <address> ... </address>
    </client>
    <order> ... </order>
    <order> ... </order>
    <creditApproval artID="store1.wo3.ca1">
      ...
    </creditApproval>
  </webOrder>
  ...
</webStore>

```



S. Abiteboul – INRIA Saclay

# Issue: verification

We can use all the work of verification of AXML systems

SA, L. Segoufin, V. Vianu [PODS08]

S. Abiteboul – INRIA Saclay

# Issue: monitoring distributed systems

## Importance of monitoring

- Detecting events of interest – Pub/sub,
- Reacting to changes, e.g., data synchronization
- Gathering traces
- Error detection, error diagnosis, intrusion, spamming...
- Gathering statistics towards optimization and tuning
- Scheduling, feedback, prediction

## Complex in distributed systems

Goal: monitor such systems & support active features ala active databases (triggers)

S. Abiteboul – INRIA Saclay



INSTITUT NATIONAL  
DE RECHERCHE  
EN INFORMATIQUE  
ET EN AUTOMATIQUE



centre de recherche  
SACLAY - ÎLE-DE-FRANCE