Knowledge Assimilation for Performance Prediction of Grid Services for optimal Workflow Execution

Zoltán Balogh
Institute of Informatics
Slovak Academy of Sciences
Presentation Outline

- Problem Definition
- Our solution using IBL
- Monitoring and getting relevant data
- Refinement of Case Retrieval
- Using historical data of one WS to predict run-time of another WS
- Future Enhancements
Initial Assumptions

Complex Applications are being created as workflows of Web/Grid Services

- as opposite to “traditional Grid applications”: input + script submission/execution + output
- SOA – Service Oriented Architectures

We assume:

- Dedicated Grid Resource Environment
Workflow Orchestration and Execution

User Request

“There is my data!”

Abstract Operation

Web Service Classes

Web Service Operations

Grid Resources

Workflow Orchestration and Execution

Abstract Grid Job
Web Service Class
Web Service Instance

WCT

AAB Scheduler

GWES

Grid Workshop GCCP 2005
Bratislava, 30.11.2005
Problem Definition

The problem can be defined for two components:

1. **Scheduler’s Requirement**
   To execute “the best” WS deployment (resp. one of it’s operations) as a part of a workflow.

2. **WCT’s Requirement**
   To use “the best” WS Class from several alternatives.

“The best” is meant in terms of predefined utility.

**Utility Function**
- minimize run time,
- minimize cost,
- ensure availability (that the WS will be maximally available),
- ensure robustness (that WS will not be interrupted, so it would need to start again),
- and others
**Example Scenario:** There are 40 deployments of the same WS deployed on different Grid Resources. Several implementations of the same WS Class

Abstract Request: Do Meteorological Simulation for me
Service Class: Meteorological Model
Service Instance: MM5
Service Deployment: http://cluster.somewhere.com/wsrf/mm5
Instance Based Learning

Case Based Reasoning applied
1. Case Representation
2. Case Indexing
3. Case Retrieval
   - Nearest-neighbor retrieval (euclidean or other)
   - Locally (linear or polynomial) Weighted Regression
4. Case Adaptation
5. Case-Base Maintenance

Instance-based learning deficiencies
- Its major disadvantage is that it requires a large amount of historical data.
- Problem with qualitative features, where exact match of feature values must be made.

We need to know in **advance** about each Grid Service Class's:
- It’s feature vector
- Results we want to predict for the Grid Service Class
**WSOperationInvocation Concept**

- **used to assimilate data from several sources**
- **represents a single invocation of a WS Instance operation**
- **Data assimilated from:**
  - Events,
  - Generic Ontology,
  - Domain Specific Ontology,
  - OWL-S Service Descriptions
Monitoring can be enacted by:

- Monitoring WS Operations
- Code Instrumentation
- WS-Notification

Statefull vs. Stateless WS

- Statefull maintain state, stateless do not
- WSRF (Web Service Resource Framework)
  - WS-Resource, WS-Life Time, WS-Notification and others
Case Retrieval must be improved – especially when we do not have well populated Case Base

We propose refinement of case retrieval for instance-based learning through semantic description of input data

- Input parameter are ontologically modeled.
- Ontology allows us to build semantic structures where concepts are derived from general to more specific.
- This allows to deduce more general concepts of ontology, thus enlarging the number of cases retrieved for reasoning, thus providing more suitable base for prediction of WS behavior.

A XML-XSD validation Web Service
- One of it’s features is file type
We do not have historical information about GS2, but have a lot of historical information about GS40.

GS2 and GS40 have deployed Grid Service and run on very similar Grid Resources.

We can use prediction model from GS40 to predict GS2.
Main Objective
- to develop an agent for the maintenance of K-Wf Grid knowledge base, extraction of knowledge from the monitoring results, workflow execution reports, user input and other sources.

Extracts knowledge
- Extracted knowledge used to predict WS instance behavior patterns
  - WS runtime prediction,
  - others: WS operation latency, WS reliability prediction, WS availability prediction
- Usable during scheduling of workflow executions

Consumes
- Historical (monitored) information and
- Semantical description of web services and data
KAA – Current Implementation Status

- **EventStore**
  - a stateless WS (Axis)
  - Stores XML Events into RDBMS (MySQL)
  - DB must be first initialized using XSD (using Torque)
  - When EventStore receives a new Event, it should initialize the update of the model for a WS in GOM (CONTEXT >> RESULT)

- **WS Behavior Predictor**
  - Takes cases from a DB (CASE = CONTEXT + RESULT)
  - Predicts RESULT for the given CONTEXT
  - Uses instance based learning and CBR to predict the run time
  - Can be adopted to predict WS performance measures (recently predicts runtime of WS depending on input parameters)

- **KAA**
  - Retrieves RESULT for a given CONTEXT
  - Visualization of KAA operation is visualized through the Logger portlet
<?xml version="1.0"?>
<request>
  <features>
    <param name="NS" value="2"/>
    <param name="EW" value="4"/>
    <param name="PROC" value="4"/>
    <param name="SIM_T" value="360"/>
  </features>

  <!-- Class of WS and it's Operation. -->
  <wscontext class="MM5" operation="run"/>

  <!-- WS Deployment Alternatives. One of them must be chosen. -->
  <wsalt>
    <ws uri="http://cluster.ui.sav.sk/mm5"/>
    <ws uri="http://cluster.cyfronet.pl/service/mm5"/>
    <ws uri="http://cluster.softeco.it/service/mm5"/>
  </wsalt>
</request>
<?xml version="1.0"?>
<result>
    <runTime wsuri="http://cluster.ui.sav.sk/mm5">
        <value>125</value>
        <status>OK</status>
        <message>159 cases used to predict the result.</message>
    </runTime>

    <runTime wsuri="http://cluster.cyfronet.pl/service/mm5">
        <value></value>
        <status>NORESULT</status>
        <message>Do not have enough cases to predict the runTime for this WS.</message>
    </runTime>

    <runTime wsuri="http://cluster.softeco.it/service/mm5">
        <value></value>
        <status>NORESULT</status>
        <message>Do not have enough cases to predict the runTime for this WS.</message>
    </runTime>
</result>
Future Plan: Supporting WCT

Same Interface implemented in several WS Class.

Example Scenario: There are 3 WS Classes of the same Interface.
Thank You

Zoltan Balogh
balogh@savba.sk
Institute of Informatics, Bratislava