K-WfGrid: Grid Workflows with Knowledge

Ondrej Habala
II SAS, Slovakia
Contents

- Project consortium
- Objectives and architecture
- Workflow construction and refinement
- Monitoring infrastructure
- Ontologies for knowledge management
- Knowledge assimilation
- Knowledge sharing between users
- Pilot applications
  - Flood prediction
  - Enterprise resource planning
  - Coordinated traffic management
  - Testbed
- Fraunhofer FIRST (Berlin, Germany)
- UIBK (Innsbruck, Austria)
- IISAS (Bratislava, Slovak Republic)
- CYFRONET (Cracow, Poland)
- LogicDIS S.A. (Athens, Greece)
- Softeco Sismat SpA (Genoa, Italy)
K-Wf Grid Objectives

- Integrating services into coherent application scenarios
- Enabling automatic construction and reuse of workflows with knowledge gathered during operation
- Involving monitoring and knowledge acquisition services in order to provide added value for end users

Technologies: service-oriented Grid architecture, workflows in Petri nets, ontologies, dynamic instrumentation
K-Wf Grid concept

- Execute workflow
- Monitor environment
- Analyze information
- Capture knowledge
- Reuse knowledge
- Construct workflow
Process of Workflow Refinement

User Request

“There is my data!”

Abstract Operation

Web Service Classes

Web Service Operations

Grid Resources

I want this data

There is my data!”
Performance Monitoring and Analysis for Grid Infrastructures and Applications

- **Focus on workflows that are Web/Grid services**
  - Workflow components can be multi-lingual

- **Event infrastructure**

- **Performance and dependability metrics**
  - Define and measure important performance and dependability metrics associated with Grid workflows

- **Monitoring and performance analysis in K-WfGrid**
  - Service-oriented distributed architecture
  - Unified monitoring and performance analysis system covering infrastructure and applications
  - Standardized data representations for monitoring data and events
  - Adaptive and generic sensors, distributed analysis, performance bottleneck search

- **Support for knowledge-based workflow construction and execution**
Ontologies are separated into several layers and classes.
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.

- Used to predict WS instance behavior patterns (run time prediction) for scheduling workflow executions and workflow composition based on:
  - Historical (monitored by MIS/PAS) information-WP3 and
  - Semantically described web services and data resources-WP4

Three main components: KAA WS (web interface for predictor), WS Behavior Predictor (Instance based learning and Case-based Reasoning), Event Store WS (Cases stored as OWL instances)
User Assistance: Knowledge Sharing Based on Text Notes

1. Matching User Context or Problem

2. Displaying Relevant Notes

3. User take Decisions or Actions

4. User feedback and knowledge update

(4) Enter Note

(4) Pattern Detection Annotation Note with Context

(1, 4) Compare Current User Context And Note Context

(1, 4) Displaying context List

(4) User approve and submit context of note

(2) Notes are displayed to user in detected user context

(4) Vote on note, Updating relevance
Collaboration among users
Knowledge Sharing and Recommendation
User Problem specification
Representation of Experience or Knowledge

- **Text Notes**

  - **User Assistant Agent**

  - **Current Research:**
    - Experience (Knowledge) Management
    - Knowledge, Semantic, Ontologies
    - Text Processing & Semantic Annotation
    - User Interaction
    - Knowledge Relevance, Problem detection

- **Proactive knowledge provision**
- **Works with any domain ontology**
Pilot applications (PAs)

- **PAs as domain-related instances of a typical K-Wf Grid-based application system**
  - Common horizontal features / Specific vertical customization
    - Flood Forecasting Simulation Cascade (FFSC, PA1)
    - Enterprise Resource Planning (ERP, PA2)
    - Coordinated Traffic Management (CTM, PA3)

- **Focus on specific aspects**
  - demonstration of K-Wf Grid solutions and value added
  - enhancement of functionalities
  - system testing & validation
  - real end users’ active involvement and contribution
    - production of dissemination material
Flood Forecasting Simulation Cascade

- a cascade of simulation models for weather forecasting, hydrological and hydraulic simulations, able to produce predictions of possible floods in target areas

- **Input**
  - map of area, precipitation, river water level etc.

- **Output**
  - water flow in target area

- **Suggestions / user support**
  - applicability of different models to specific scenarios (QoS)
  - comparison between stored predictions and real data
  - future dispersion forecasting
Flood Forecasting Application

(PA1-FFSC)

- Flood forecasting based on a series of simulations
- Several simulation models for:
  - Meteorology
  - Hydrology
  - Hydraulics
- 2D/3D visualization
- Implementation using GT4 WSRF
- Being extended:
  - More models
  - More services – towards risk management
PA1 – FFSC + K-Wf Grid

- **Improvement of application reliability**
  - Creation of a Grid of modular (scalable) dedicated Web Services

- **General improvement of performances**
  - CPU time-demanding meteorological and hydraulic simulation models
  - gigabytes/month transferred and stored datasets

- **Knowledge management support**
  - multi-user, extensible multi-module system (with potentially tens of thousands of existing datasets) needing sophisticated collaboration tools to enable cooperation in a multidisciplinary context
Enterprise Resource Planning

- product line management for different targeted market segments

- enterprise financial & accounting systems,
- manufacturing systems
- supply chain management
- electronic commerce support

- multi-level parametrical system setup based on simple or complex business processes to post or extract data and information
PA2 – ERP
PA2 – ERP + K-Wf Grid

- **Grid-enabled big corporate resource management**
  - transparent access to corporate resources and financial services throughout a VO composed by e.g. all affiliated companies
  - location-independent analysis and data mining possibilities

- **General improvement of performances**
Coordinated Traffic Management

- workflow creation and management in handling complex traffic planning and management tasks

  - management of different traffic situations
    - by composing and operating suitable simulation chains involving distributed resources, data and services
  
  - three classes of scenarios
    - air pollution emission forecasting / calculation
    - mid / long term planning
    - immediate emergency reaction

  - scenarios analysis and description
    - resource (services + data) availability
    - performance requirements
      - priority to accuracy and to speed
PA3 – CTM

K-Wf Grid Project - IST 511385

Bratislava, Slovakia, November 2005
From a fragmented “manually-integrated” system to a truly coordinated & distributed Grid-enabled environment

- transparent access to data repository, simulations and calculations tools, within a VO composed by all “traffic management-related” authorities

Alternatives scenarios analysis

- possibility to configure parallel simulations and forecasting tests with different boundary conditions, composing workflows and exploiting resources accordingly

General improvement of performances
Thank you!

Ondrej Habala
II SAS, Slovakia

www.kwfgrid.net