Abstract Workflow Composition in (K-Wf) Grid Environment

Tomasz Gubała, Marian Bubak,
Maciej Malawski and Katarzyna Rycerz

Academic Computer Centre
Cyfronet AGH
Krakow, Poland

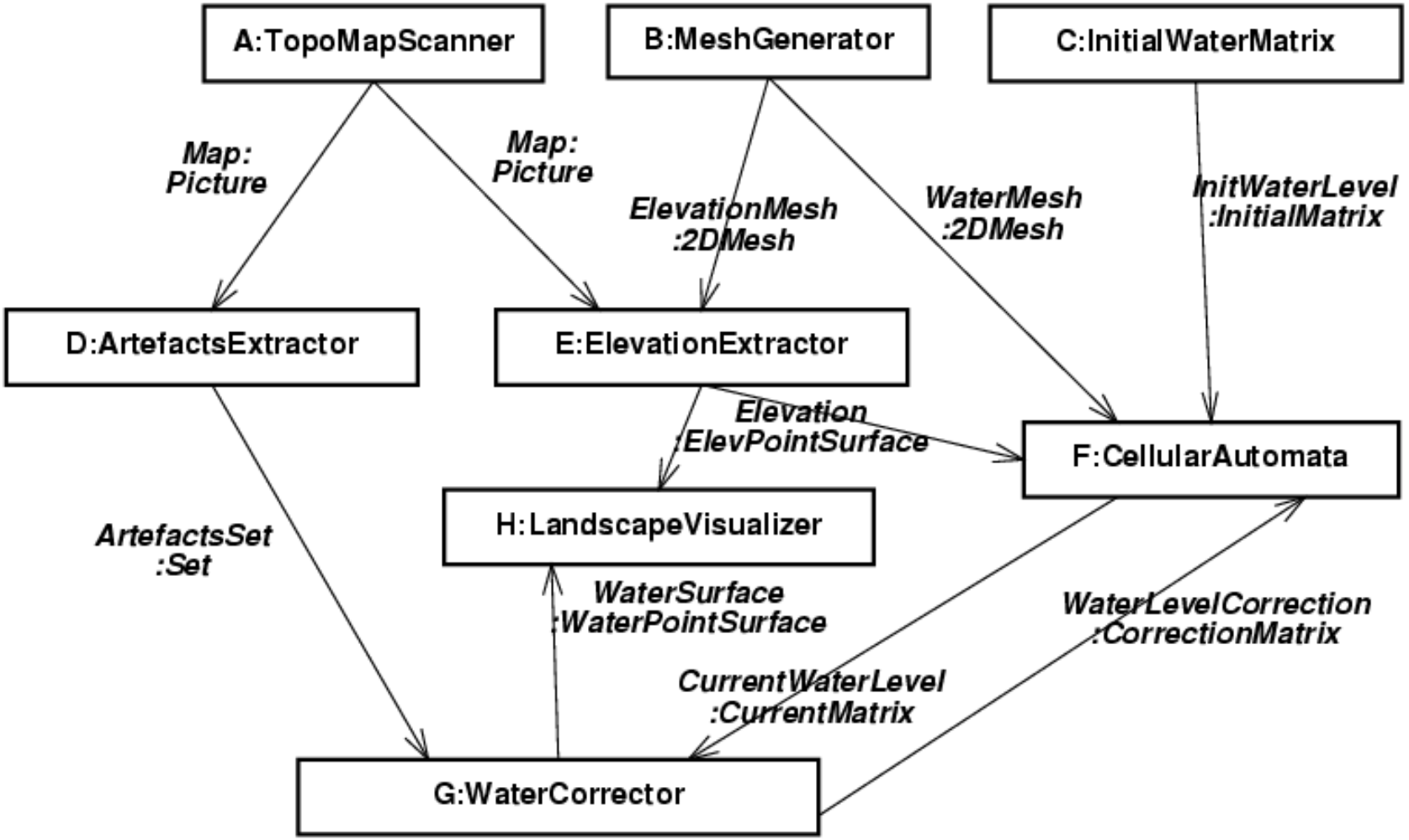
Contents

- General presentation of workflow composition problem
- How does it apply to Grid
- Our research on workflow composition so far
- Workflow Composition Tool in K-Wf Grid

Some Explanations

- Workflow
  - Directed graph (possible branches, loops, joins ...)
  - May contain other features (expressions, events ...)
- Nodes: a (computational) resources
- Edges: data or control flow links
- Workflow composition
  - A process of discovering new resources and interconnecting them in order to provide a solution for user-defined problem (application)

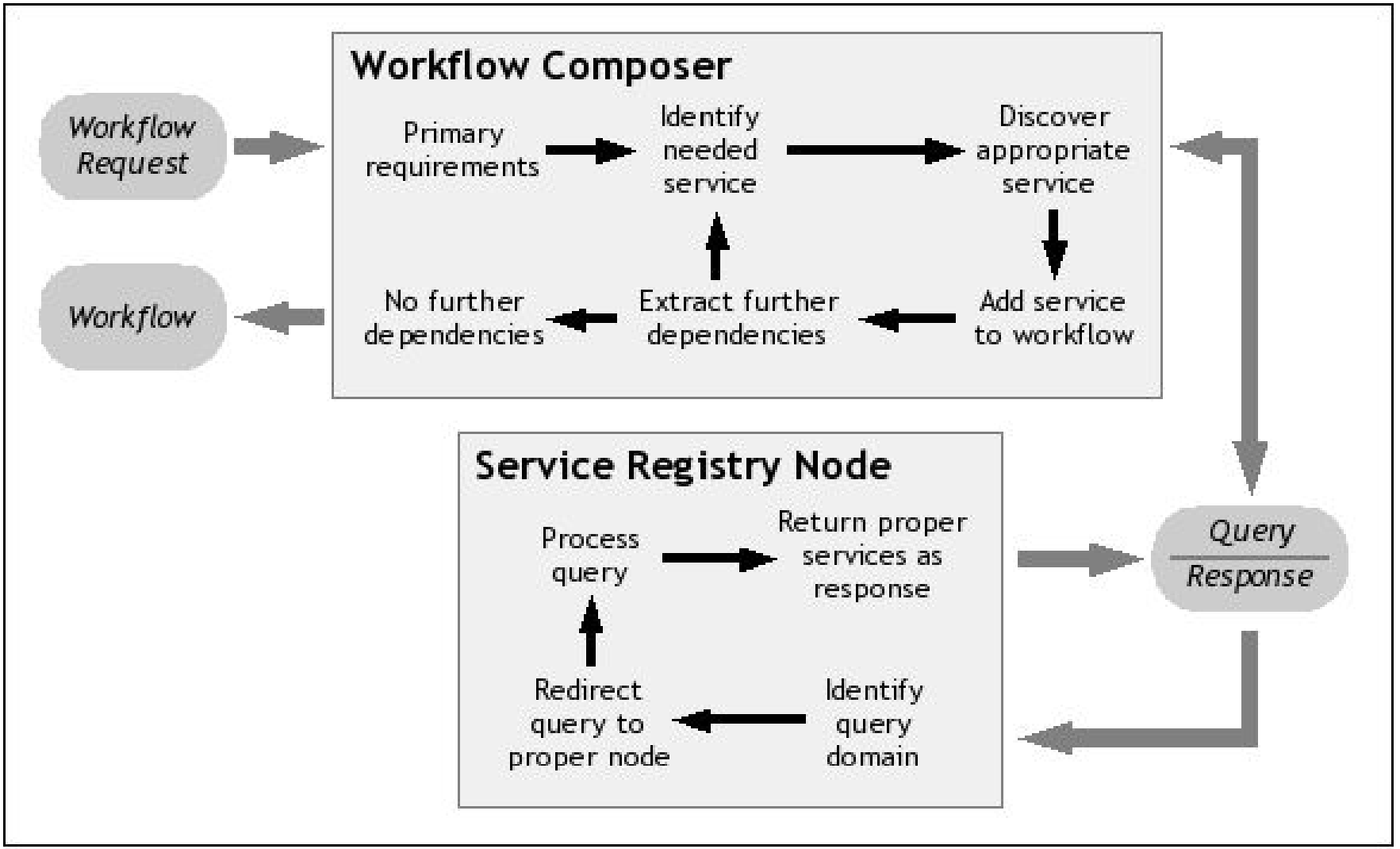
Example of Workflow



Motivation

- Workflow composition
  - User assistance in huge Grid environment
  - Semantic-based discovery of proper solutions
  - Tackling with detailed constraints of interconnected parts of application
- Where it may be useful
  - Grid programming (esp. Grid prototyping)
  - Fast software resource discovery
  - Solving a problem in many different ways

Conceptual Architecture



Composition Prerequisites

- Uniform standard of resource description
  - Like IDL/WSDL interfaces OO/WS technologies
  - Describes syntax and semantics of a resource
- Well-defined formats of input and output
  - Workflow description language on output
  - Input - incomplete workflow (even a *sketch*)
  - Possibility of multiple workflows on output due to multitude of available resources
- External Grid Service Registry

Semantical Matchmaking

- Ontologies describe meanings of particular parts of an application
  - Embedded in both data and resource description
- Reasoning engine draws conclusions using provided meaningful information
- Possibility of inexact matchmaking
  - Ontology entity similarity based on equivalence, subclassing, enclosure ...

Previous Research

- Composer of workflows for applications based on Common Component Architecture (**CCA**)
- Design and *proof-of-concept* implementation using **XCAT** CCA Grid framework (**AFC**)
- First prototype of distributed service registry (**GridRegistry**) for WSDL-based services
- Design of a new tool for workflows based on Web Services with support for service semantics (to be realized in K-Wf Grid project as **WCT**)

Workflow in K-Wf Grid

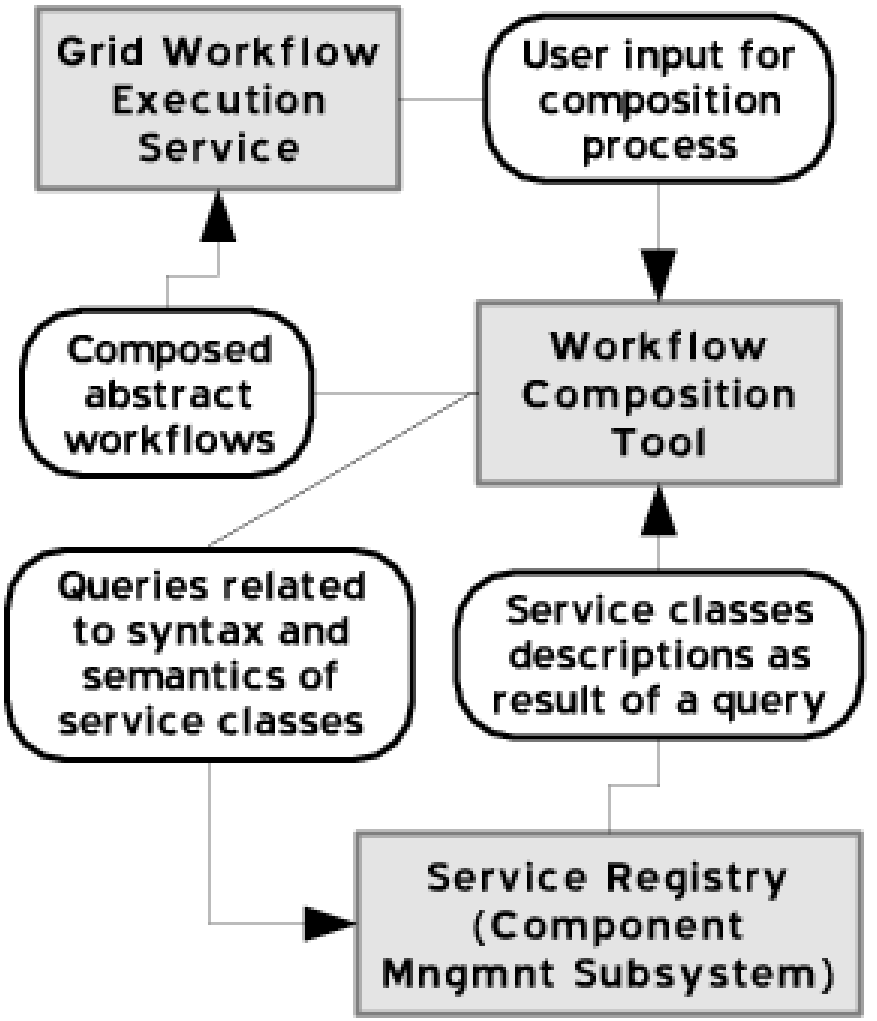
- The name
  - *Knowledge-based Workflow System for Grid Applications*
- Work Package 2
  - *Workflow Orchestration and Execution Environment*
- Task 2.3
  - *Workflow Composition Tool*

Abstract Workflow in K-Wf Grid

- Two levels of abstraction help with workflow refinement
  - **Abstract**: contains full information on how the application works
  - **Concrete**: adds information related to actual execution of that application
- Abstract workflow group ***service classes*** (which are identifiers of certain service interfaces)
- Concrete workflow knows exact ***instances*** of services (belonging to these classes) capable of conducting a computation

WCT Environment

- Execution service
  - Return flows for execution
- Service registry
  - Query for available services
- User interface
  - Ask user for assistance when needed



Technology

- Workflow description language
  - Probably based on Petri Net formalism
  - Features: loops, event handling, levels of abstraction
- Computational resource
  - Web Service (WSRF)
  - Described with WSDL and OWL-S documents

WCT Roadmap in K-Wf Grid

- Development lasts from M6 up to M28
  - First prototype: M15
  - Stable release: M24
  - Testing and refining: M28
- The most general classes and parts written
- Incremental implementation in WP2
  - Fundamental workflow elements first
  - Additional features in later stages

Thank you very much