Computer Science Department

SYNASC 2008

Decentralized Dynamic Resource Allocation for Workflows in Grid Environments

Florin Pop, *Ciprian Dobre*, Valentin Cristea

{florin.pop, ciprian.dobre, valentin.cristea}@cs.pub.ro

University POLITEHNICA of Bucharest

Computer Science Department

Background & Motivation

- **Original Grid vision**: Grid computing should enable scientists seamlessly solve large-scale problems using shared resources
- Current **Grid status**: resources are inappropriately used and are prone to errors
- **Innovative architecture**:
  - Decentralized dynamic resource allocation to optimal and accurate schedule workflows in Grid environments
  - Fault-tolerant and policy aware
- The superiority of the proposed design is evaluated using adequate modeling techniques

Computer Science Department

Outline

- Workflow Scheduling in Grid
- Decentralized architecture for Scheduling in Grids
- Policies for scheduling

- Resource allocation for PEGAF
  - Dynamic resource allocation component
  - Rescheduling component
  - Fault management component
- Validation model using simulation

- Impact for PEGAF community
- Conclusions and future work

Computer Science Department

Workflow Scheduling in Grid

- Complex problem
  - Diversity of dependencies
  - Grid dynamism and heterogeneity
- Heuristics are the usual approach to solve the workflow scheduling problem
- DAG workflows scheduling solutions:
  - List heuristics
  - Duplication based algorithms
  - Clustering heuristics
- In order to minimize communication times – schedule as many tasks as possible to a single cluster

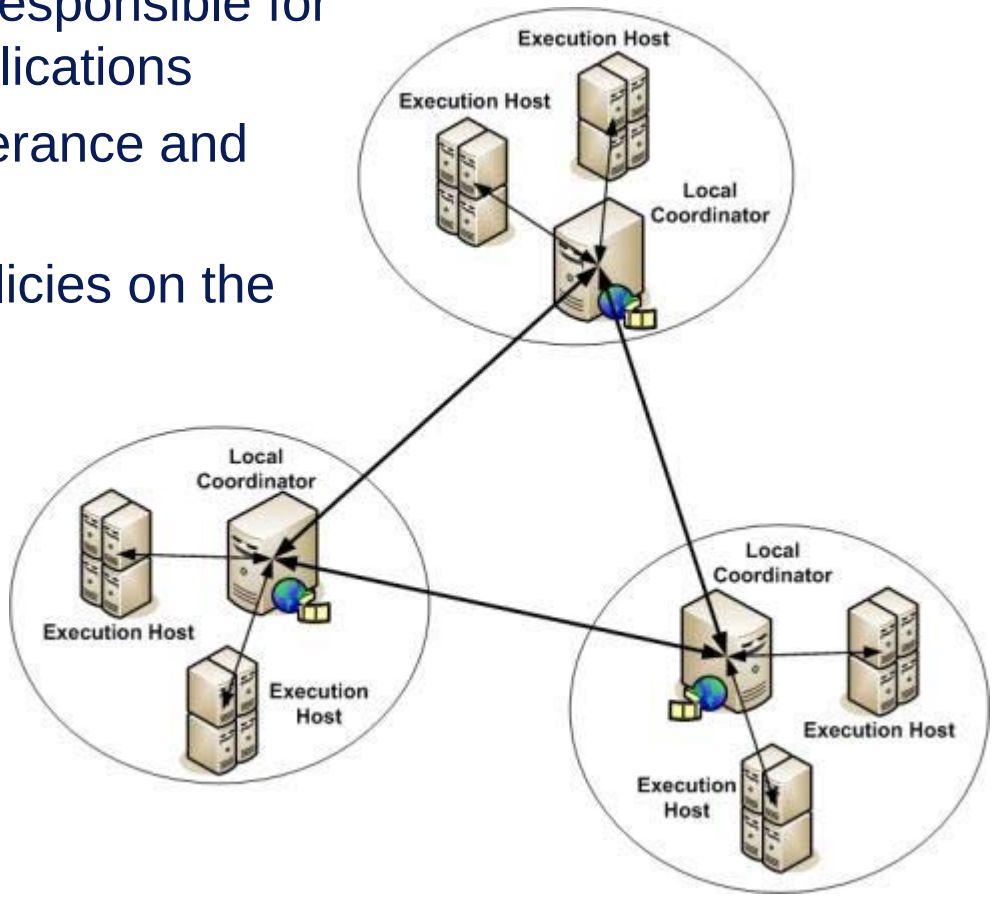
Computer Science Department

DAG Scheduling properties

GRID Communication	Node arbitrary conected ✓	Node fully conected
GRID Resources	Limited number of processors ✓	Unlimited number of processors
Fault tolerance strategy	With duplication	Without duplication ✓
Communication costs	Uniform communication costs	Arbitrary communication costs ✓
Communication model	With communication ✓	Without communication
Graph stucture	Arbitrary graph structure ✓	Restricted graph structure
Tasks dependencies	Depended tasks ✓	Independent tasks
Scheduling model	Centralized Scheduling	Decentralized Scheduling ✓

Decentralized architecture for Scheduling in Grids

- Meta-scheduling model
  - Coordination between different sites
  - The local scheduler is responsible for scheduling its local applications
  - Delivers better fault-tolerance and reliability
  - Different scheduling policies on the local sites
  - Site-autonomy



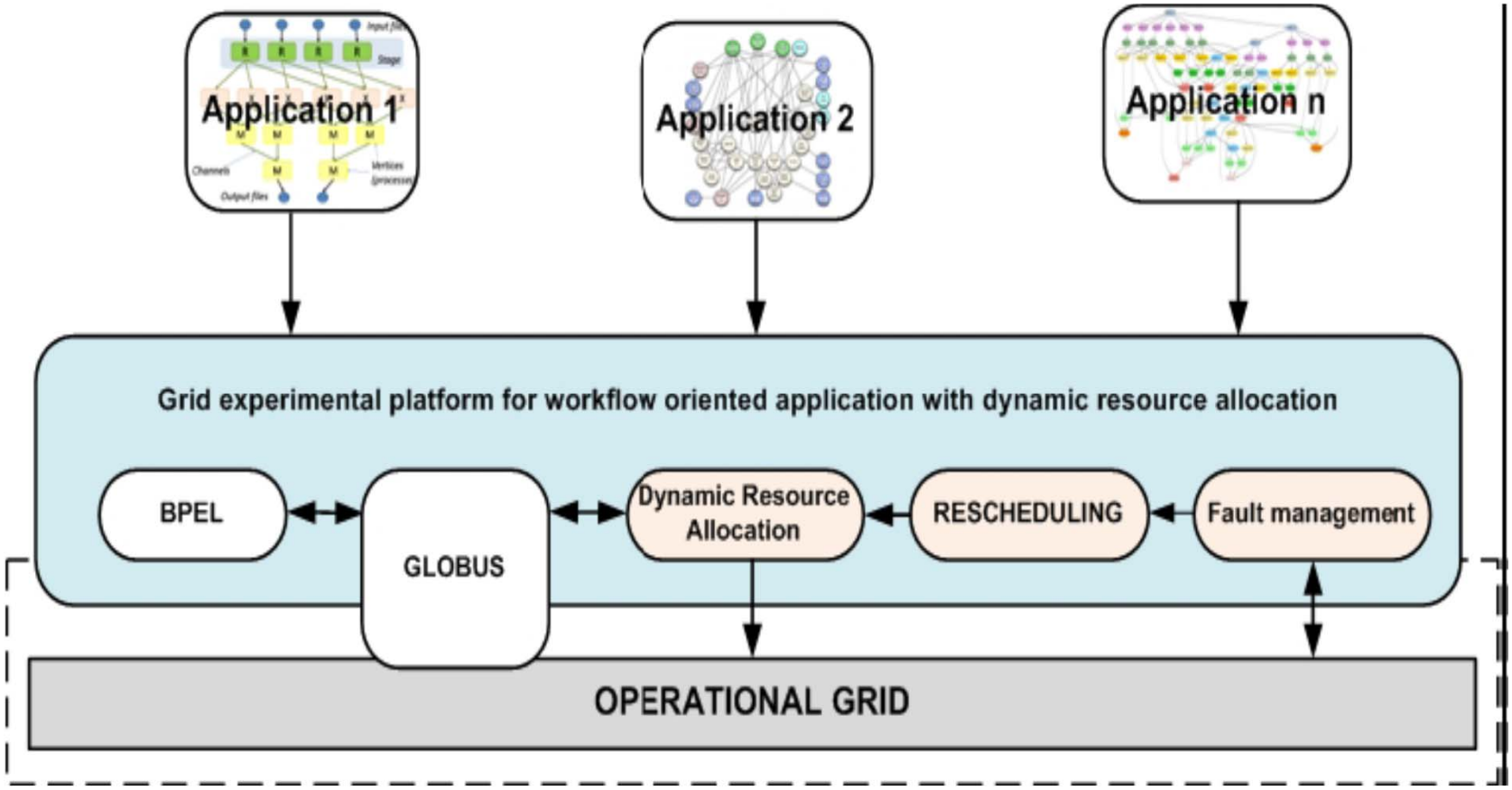
Computer Science Department

Policies for scheduling

- Scheduling policy
  - Fixed
    - System oriented – different resources requirements: CPU, Memory
    - Application oriented – completion with deadlines, QoS constrains, user requirements
  - Extensible
    - Ad-Hoc: used for dynamic resources allocation
    - Structured: used for heuristics approaches
  - Business model
    - Combined fixed and extensible policies
    - Define Grid economic model
    - Could be dynamically configured by users
    - Is recommended for workflow management

Computer Science Department

Resource allocation for PEGAF



PEGAF: framework to help scientists easily deploy large-scale workflows

Computer Science Department

Dynamic Resource Allocation Component

- Functionality
  - Collaborate with workflow execution system (BPEL)
  - Collect resources state using a monitoring service
    - List of available resources
  - Uses resources description for scheduling
  - Used in real-time environment for workflow management

- Perform service invocation on the fly as the application executes
- Useful when it is impossible to estimate the execution time for a services in a workflow

Computer Science Department

Rescheduling Component

- Fault recovery mechanism for dynamic workflow scheduling
- Rescheduling can be costly when DAGs have extra data dependencies among tasks compared to independent applications.

- Re-Scheduling
  - *Periodic batch*: group resource requests and system events which are then processed at intervals that may be periodic triggered by certain system events
  - *Event driven online*: performs the rescheduling as soon the system receives the resource request

Computer Science Department

Faults managements Component

- Execution context
  - Scheduler to support fault tolerance
- Application code
  - Use of Grid checkpointing strategies => improvements of rescheduling performance
  - The fault tolerance level provides transparent check-pointing support for Grid services using an execution coordinator
- Use of the fault-tolerance support provided by:
  - A specialized RMI layer
  - Condor DagMAN

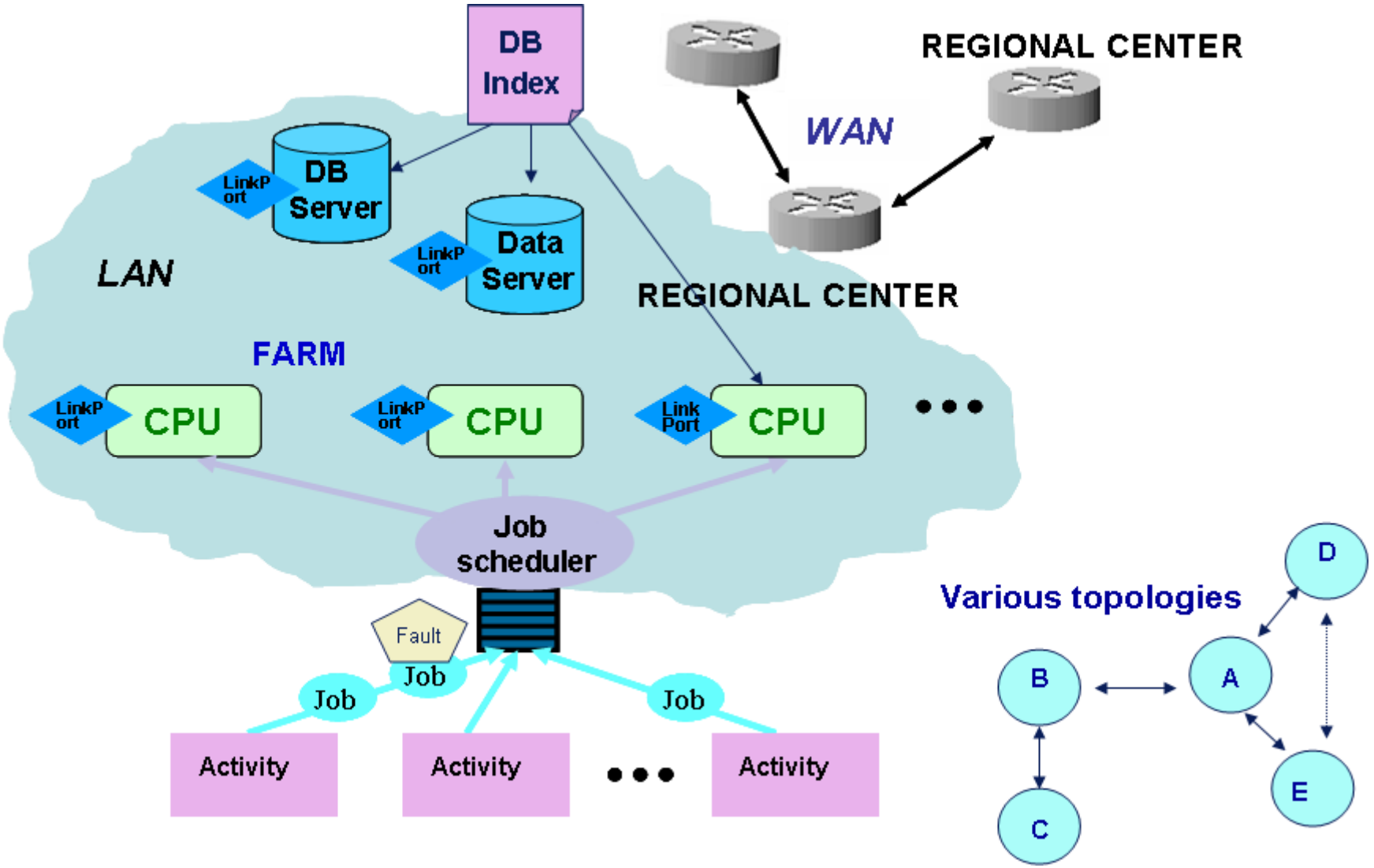
Validation model using Simulation

- The simulation model proposed by MONARC 2 includes the necessary components for simulating various distributed systems technologies, with respect to their specific components and characteristics

- The model includes the necessary components to describe various actual distributed system technologies, and provides the mechanisms to describe concurrent network traffic, evaluate different strategies in data replication, and analyze job scheduling procedures

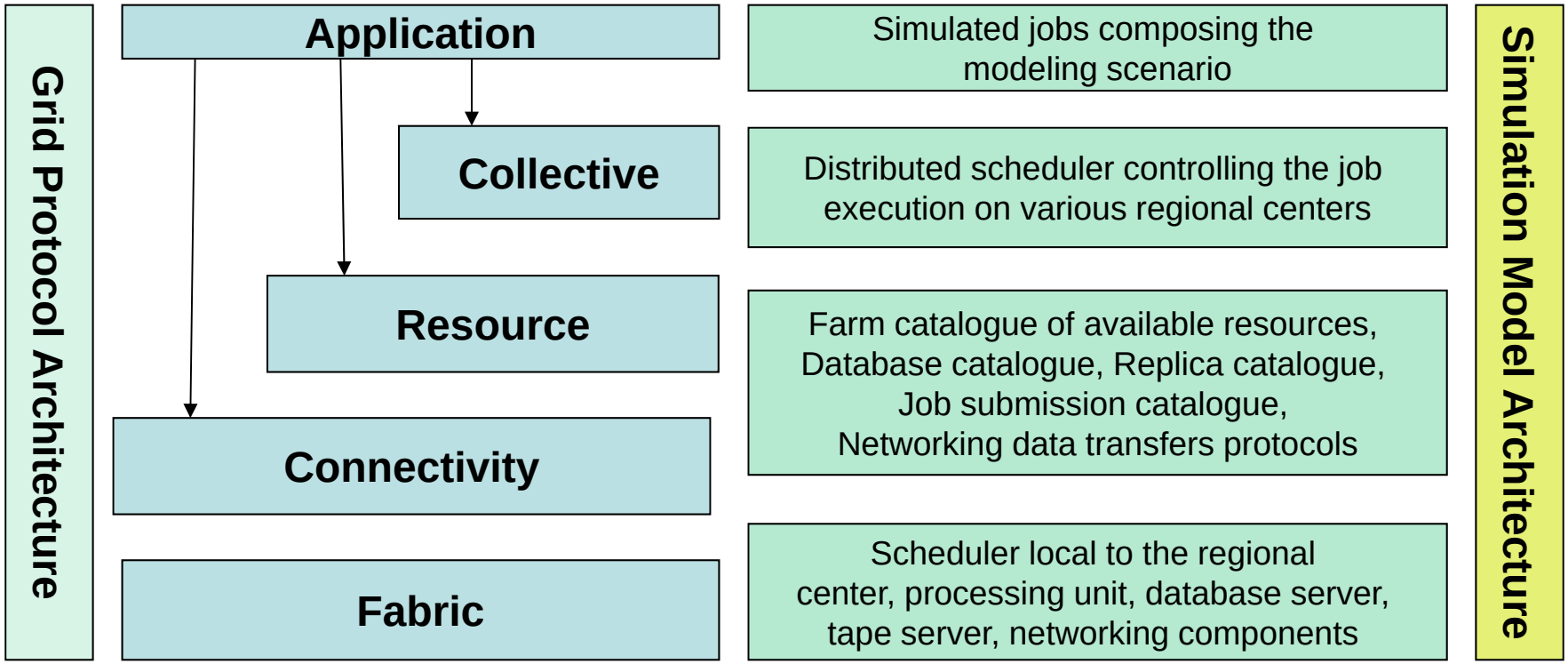
- One important characteristic of the proposed simulation model is its generality

Components of the Simulation Framework



Grid architecture and its influence on the simulation model

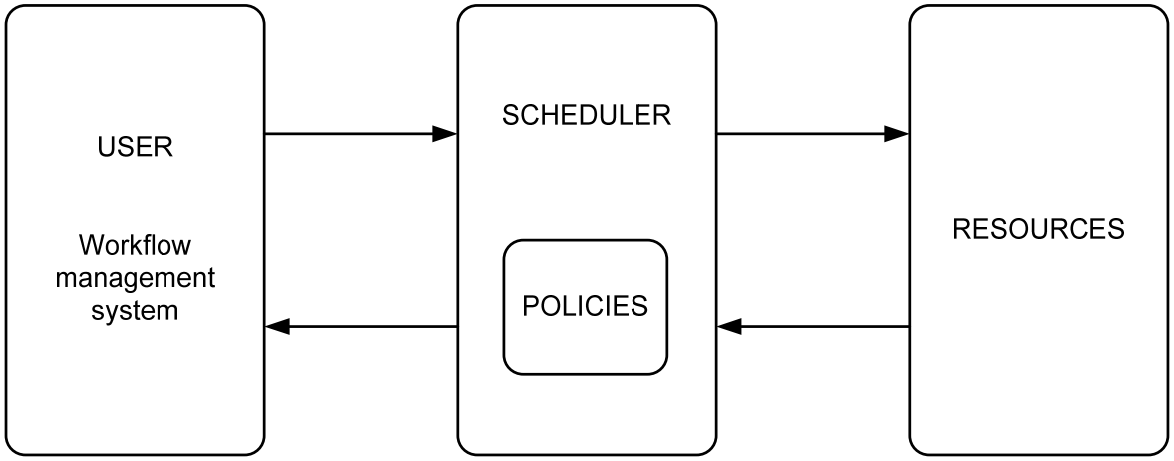
The simulation model incorporate the components and characteristics specific to the Grid layered architecture



Computer Science Department

Impact for PEGAF community

- Offers an interaction layer between large-scale distributed resources and scientific workflow management engines
- Provides a dynamic mode for resource allocation
- Considers users and resources policies
- Interacts with different types of services
- Offers a general mode for scheduling
- Supports fault tolerance using re-scheduling



Computer Science Department

Conclusions and Future work

- Innovative decentralized dynamic resource scheduling solution for large scale application workflows
- Considers the problem of dynamic resource allocation
- Proposed components are part of PEGAF
- Decentralized architecture
- Considers scheduling policies and fault management
- Modeling and simulation to demonstrate the validity of the solution

- Test the integration between existing Grid middleware and workflow engines

Questions ?

Thank you !

ciprian.dobre@cs.pub.ro

The running performance of simulation