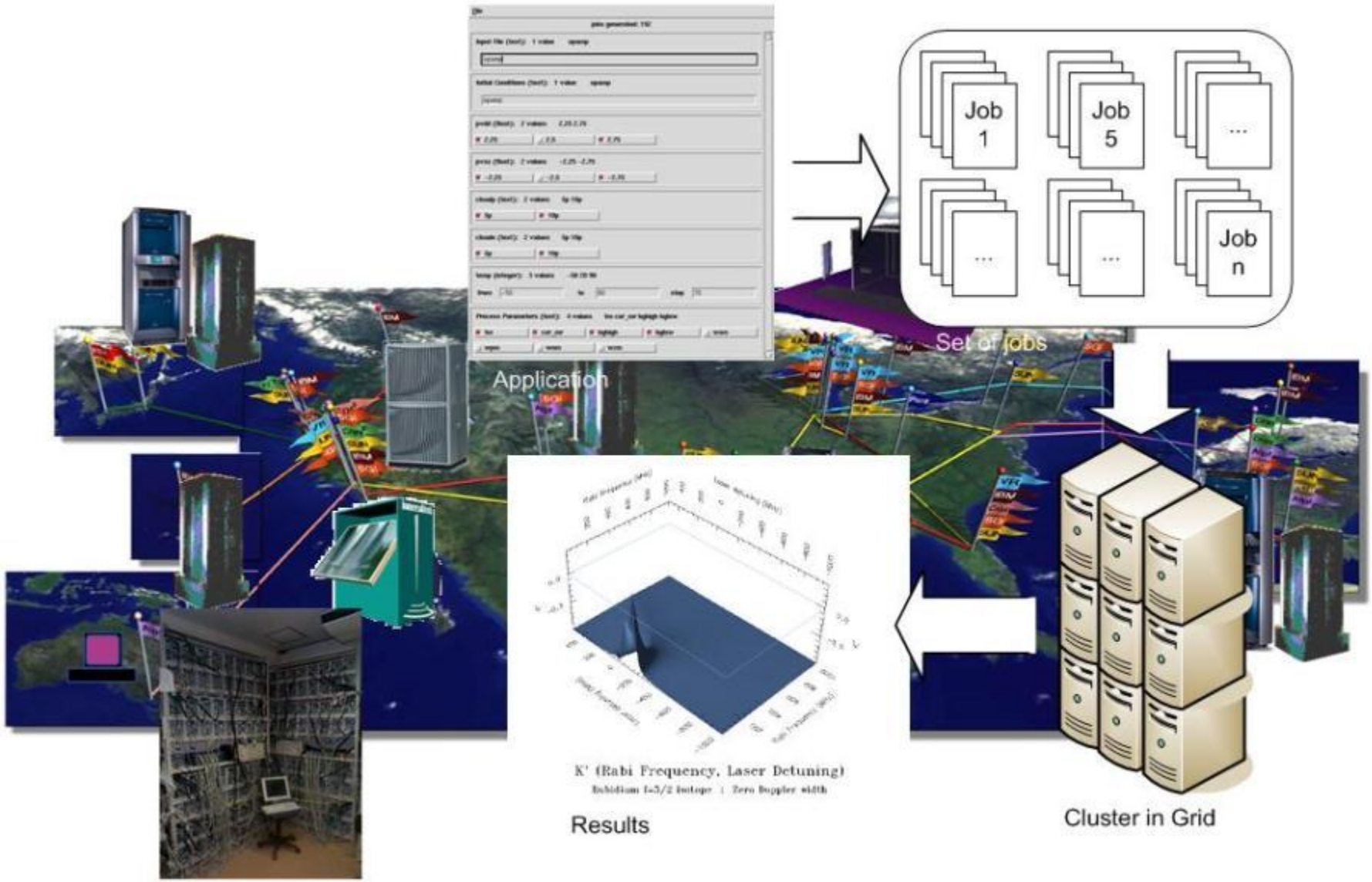
Metode şi Algoritmi de Planificare (MAP)

2009-2010

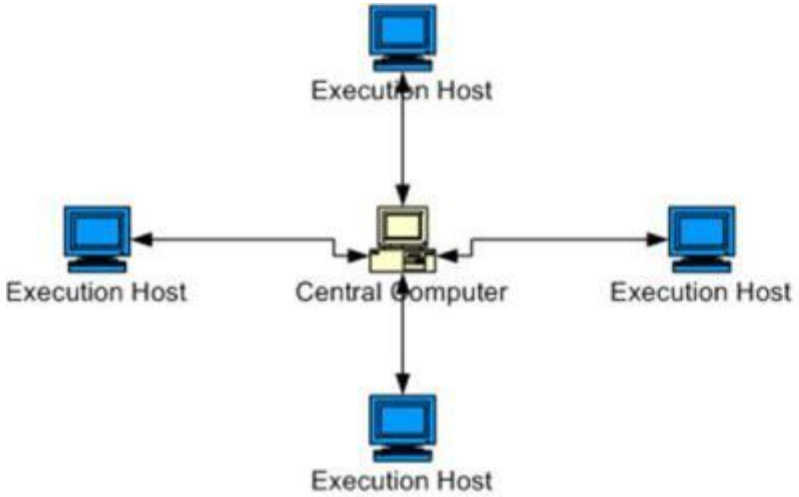
Curs 5

Arhitectura planificatoarelor

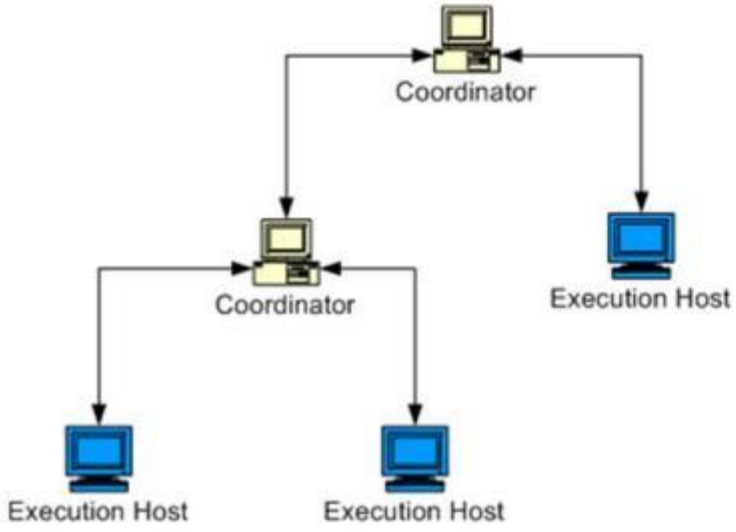
Large Scale Scheduling

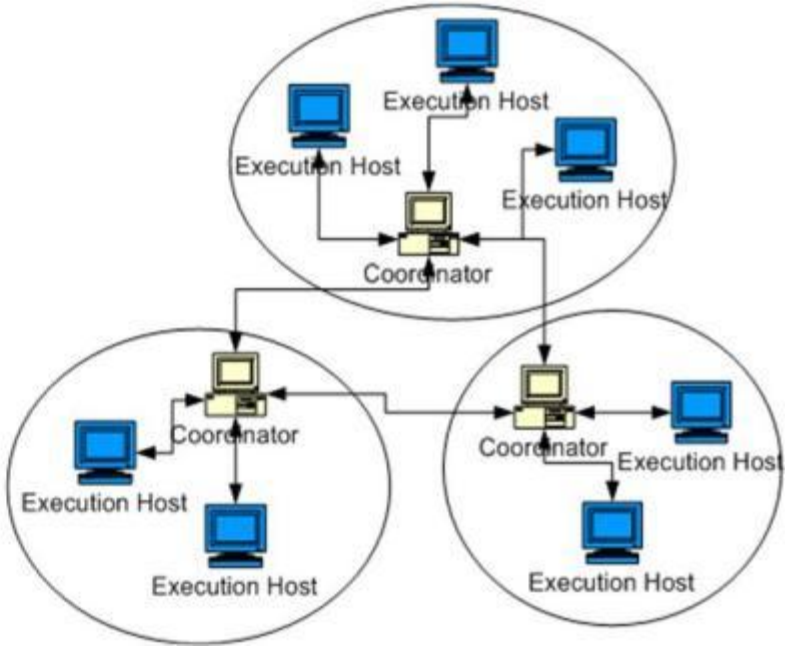


Scheduler Architecture

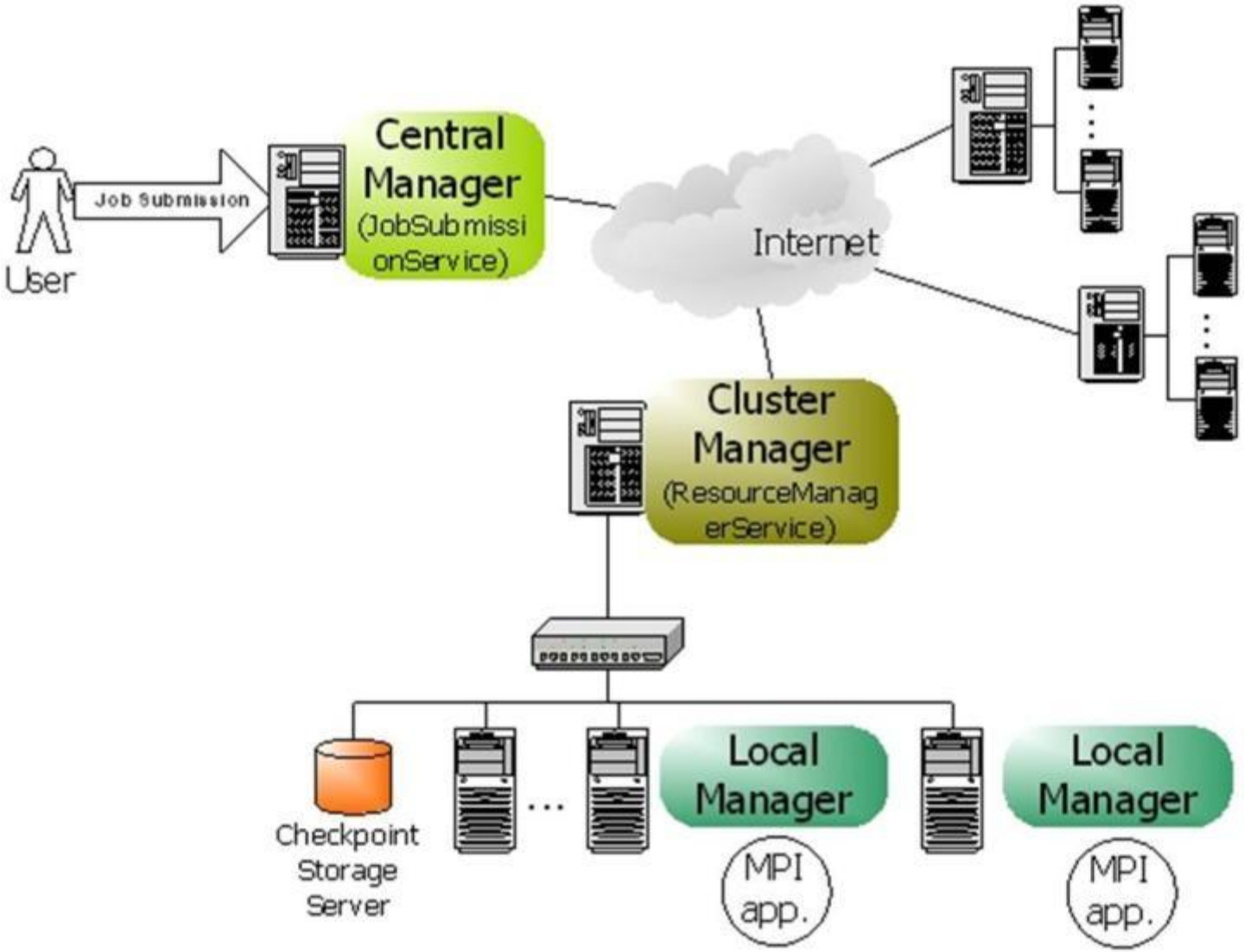


- A Centralized Scheduler
- A Hierarchical Scheduler
- A Decentralized Scheduler

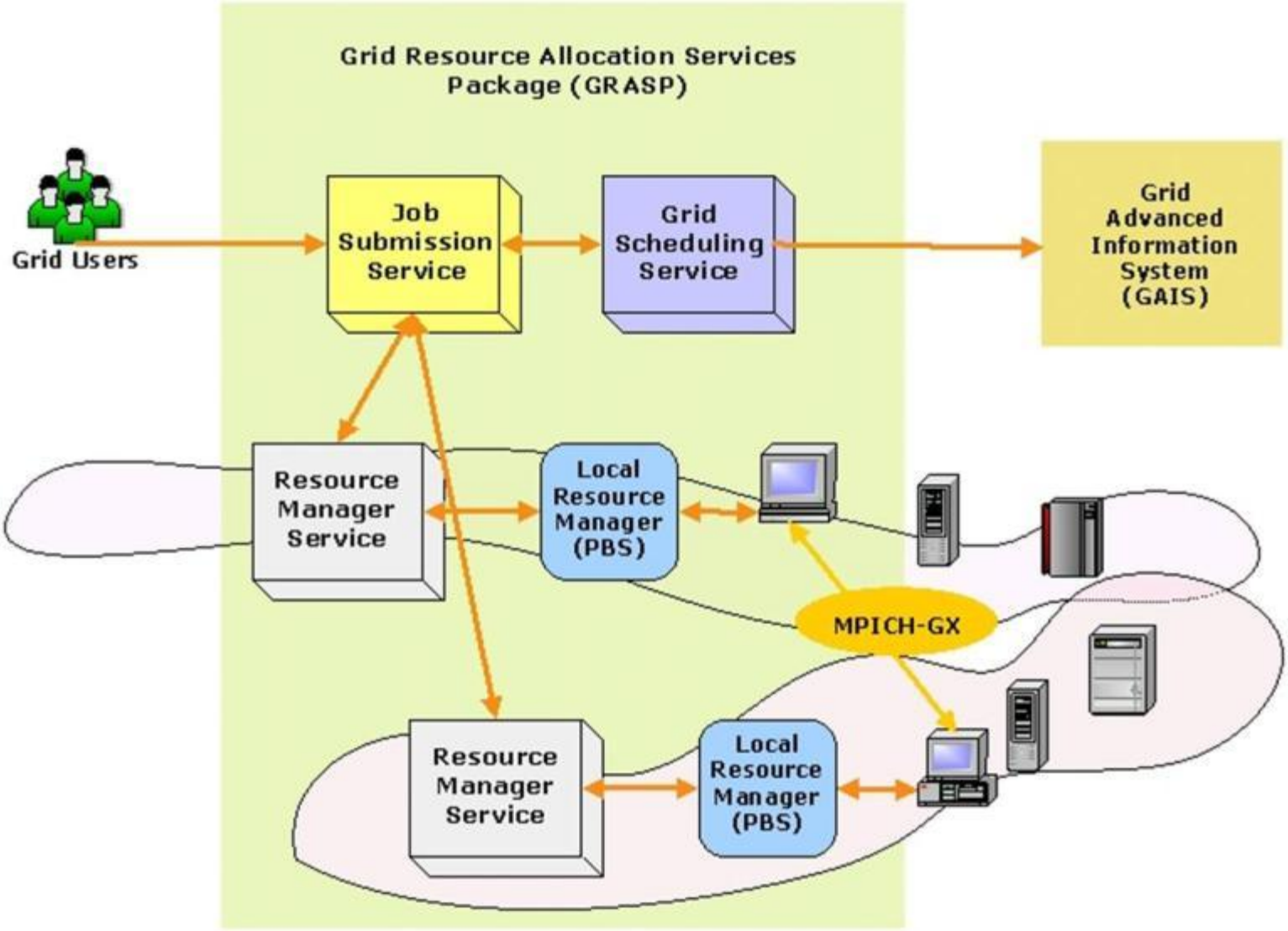




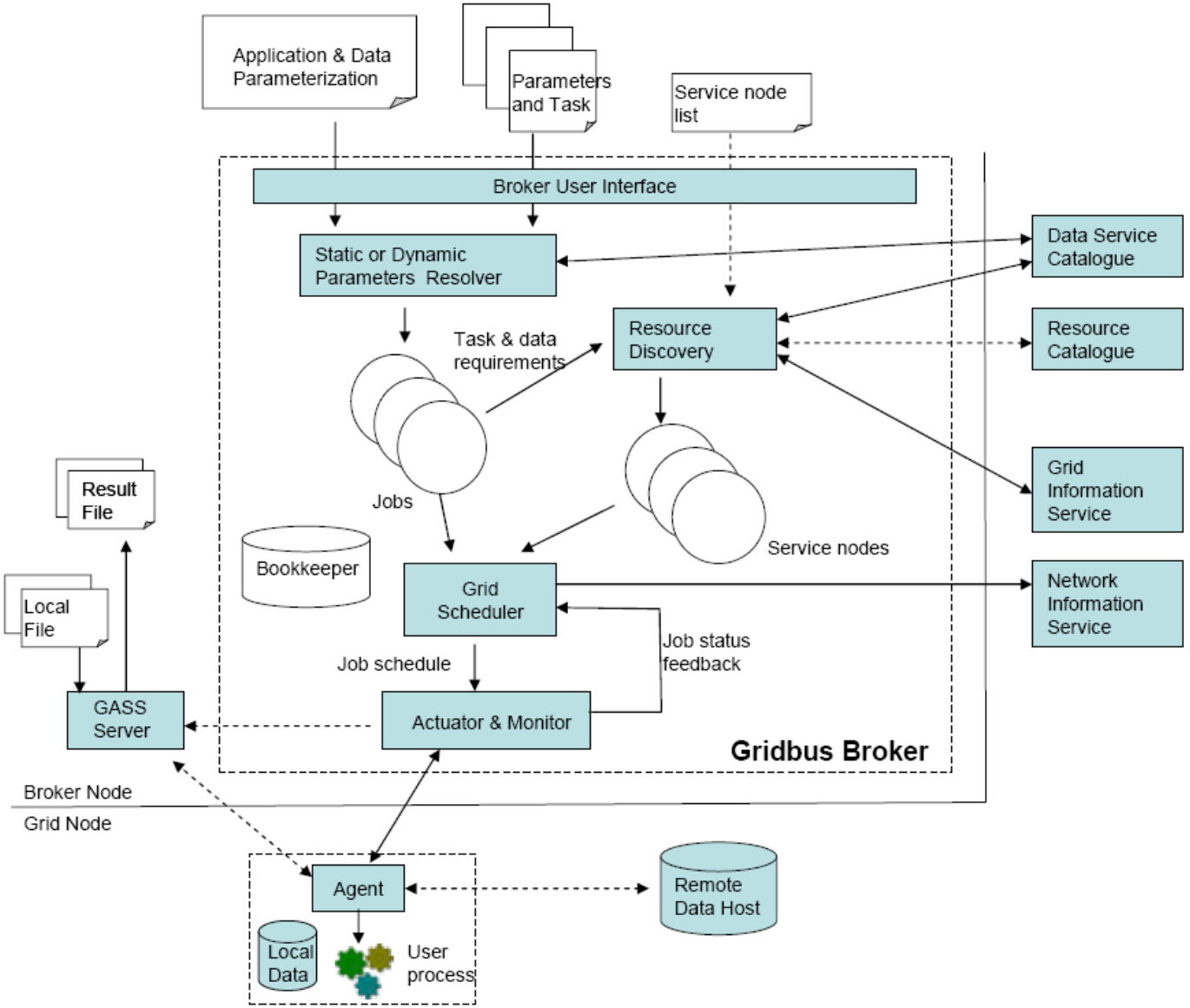
Grid Scheduling Architecture



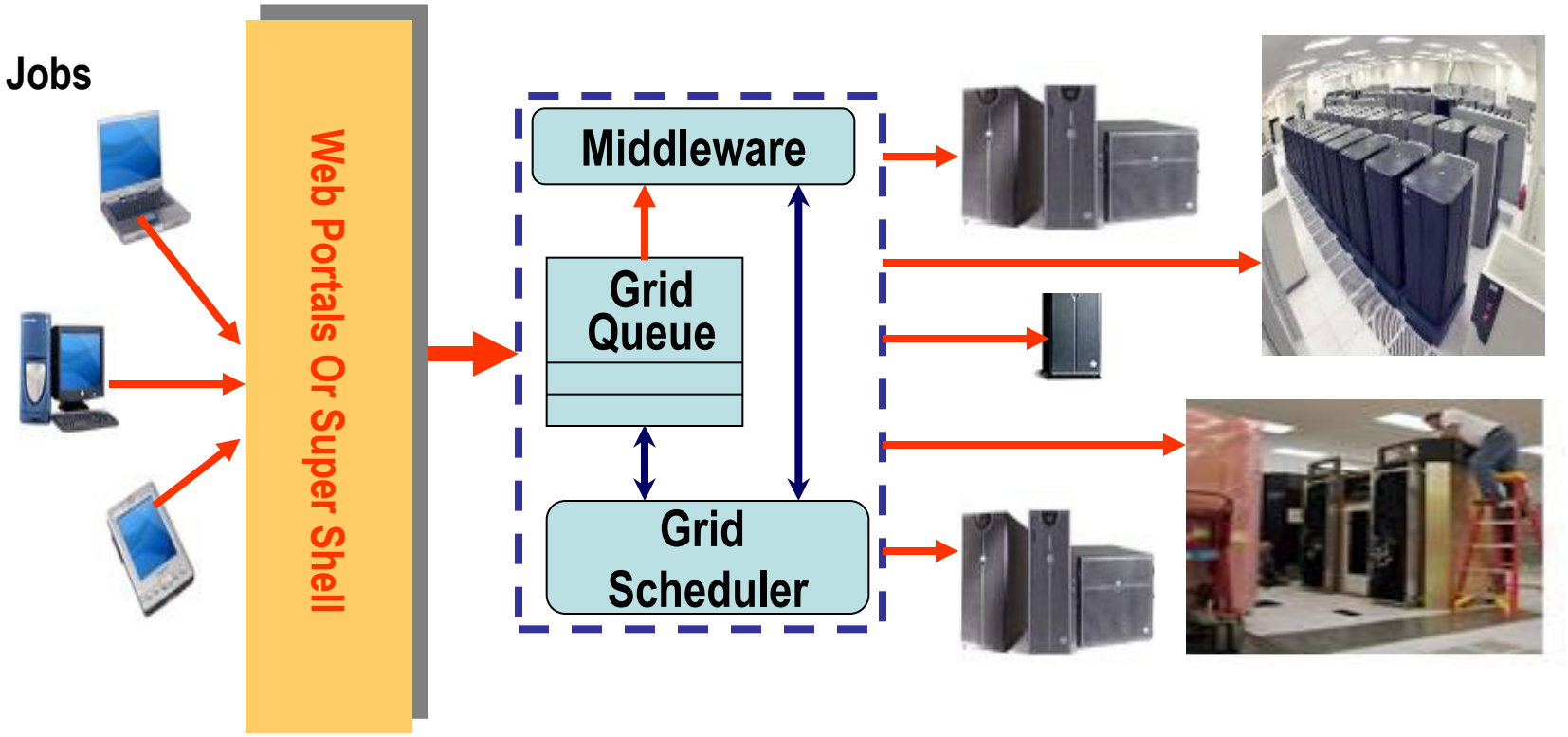
GRASP: Grid Resource Allocation Services Package



Scheduling Models - Central Broker

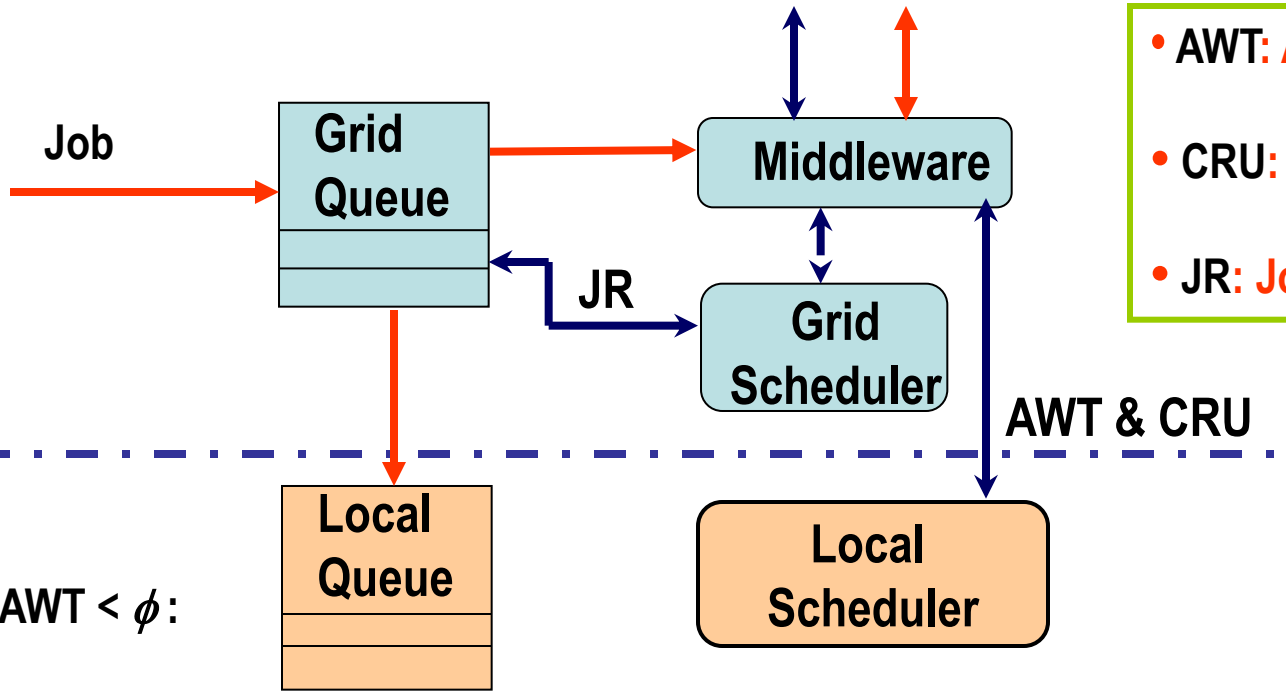


Centralized Architecture



Advantages: Global View
Disadvantages: Single point of failure, Scalability

Interaction between Grid and Local Schedulers



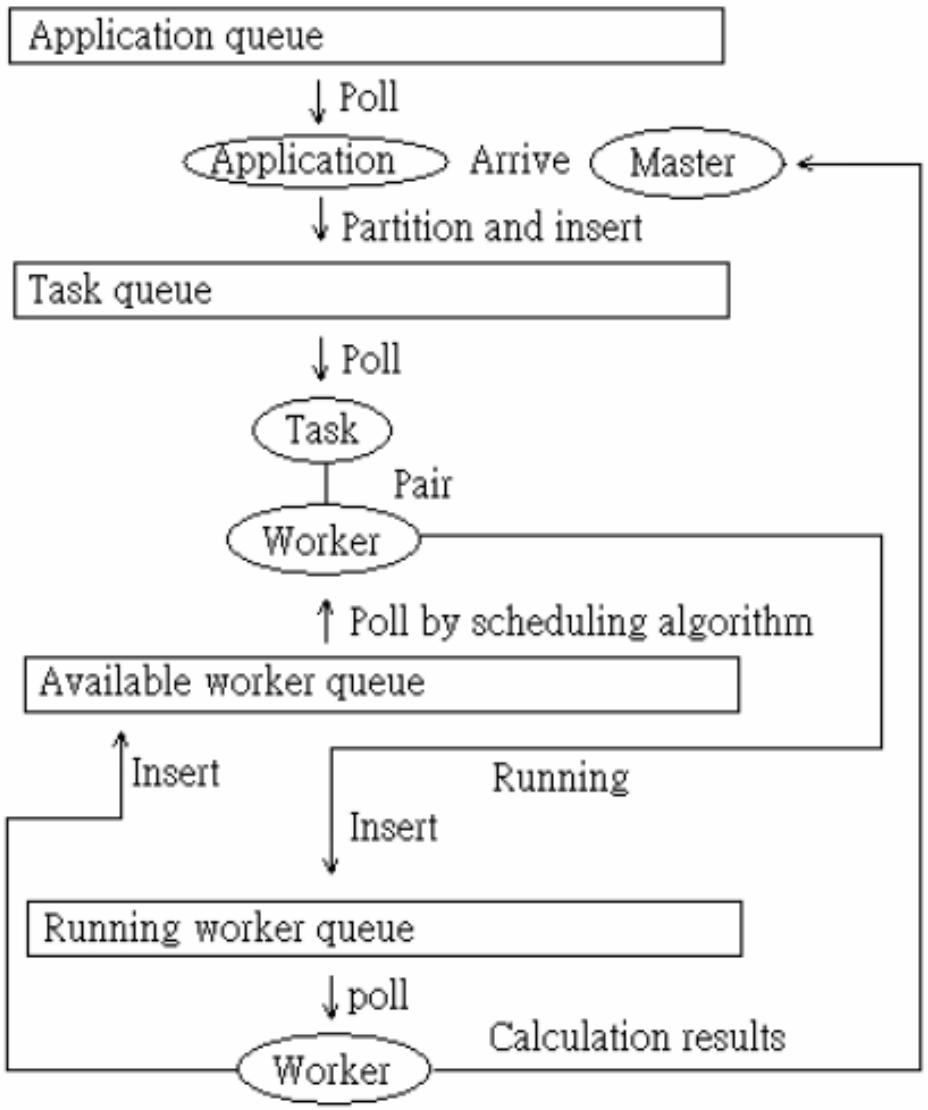
- AWT: Approximate Wait Time
- CRU: Current Resource Utilization
- JR: Job Requirements

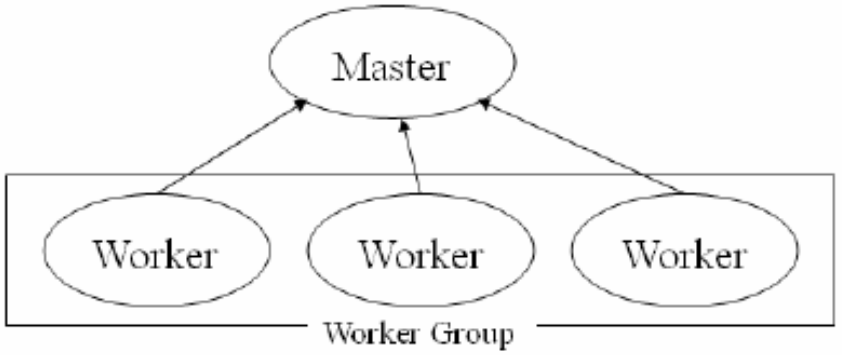
If AWT < ϕ :

Else : Considered for Migration

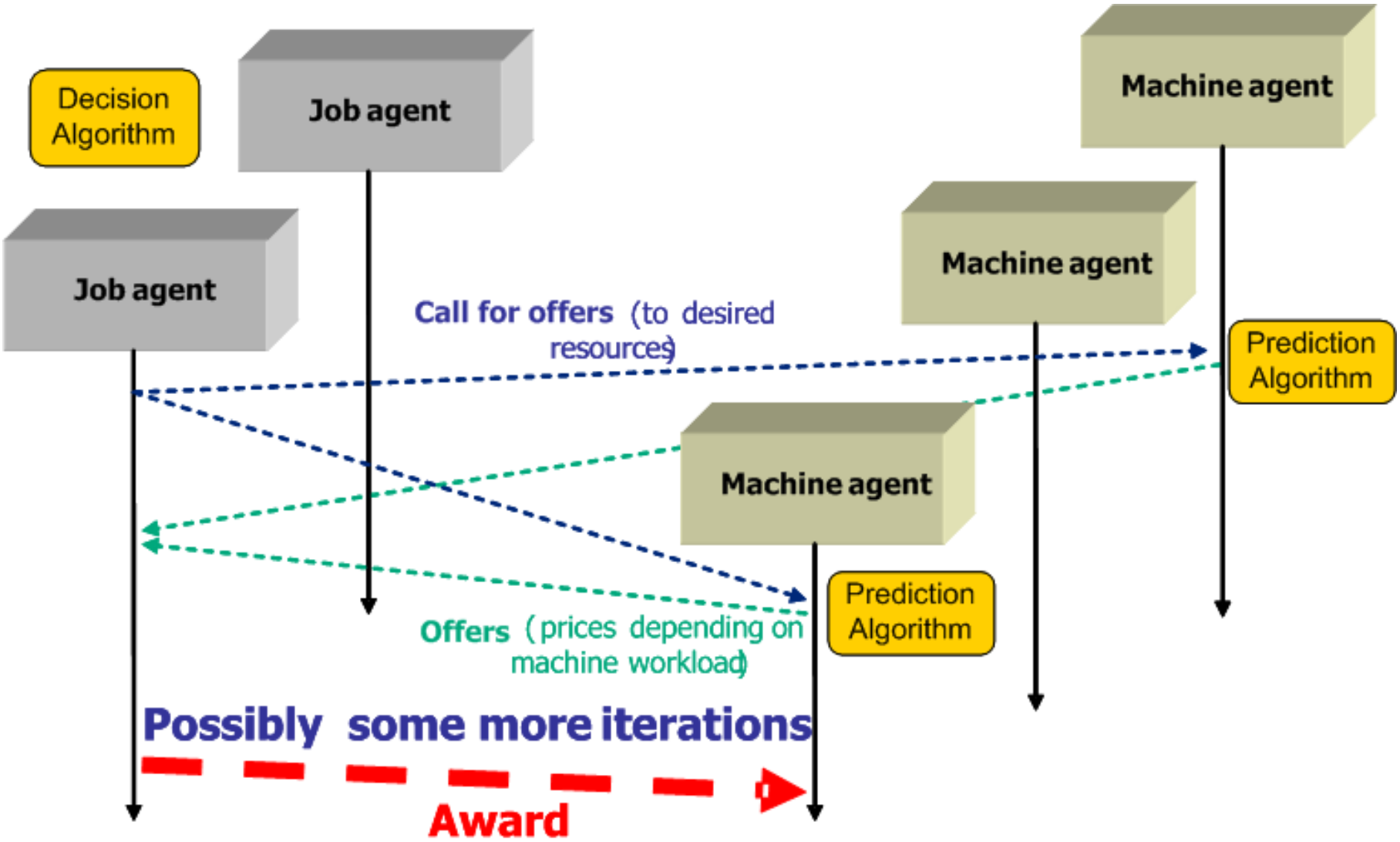
- Sender-Initiated
- Receiver-Initiated
- Symmetrically-Initiated

Scheduling Models – Replicated





Scheduling Models – Marketing model



Architecture requirements

Definition of a **Grid Scheduling Architecture** allows co-operation between local resource management systems for network, software, data, storage and processing units and additional Grid services as provided, for instance, by Grid information and monitoring systems.

Specifically, the interaction between hardware resource management and data management shall be considered in the architecture.

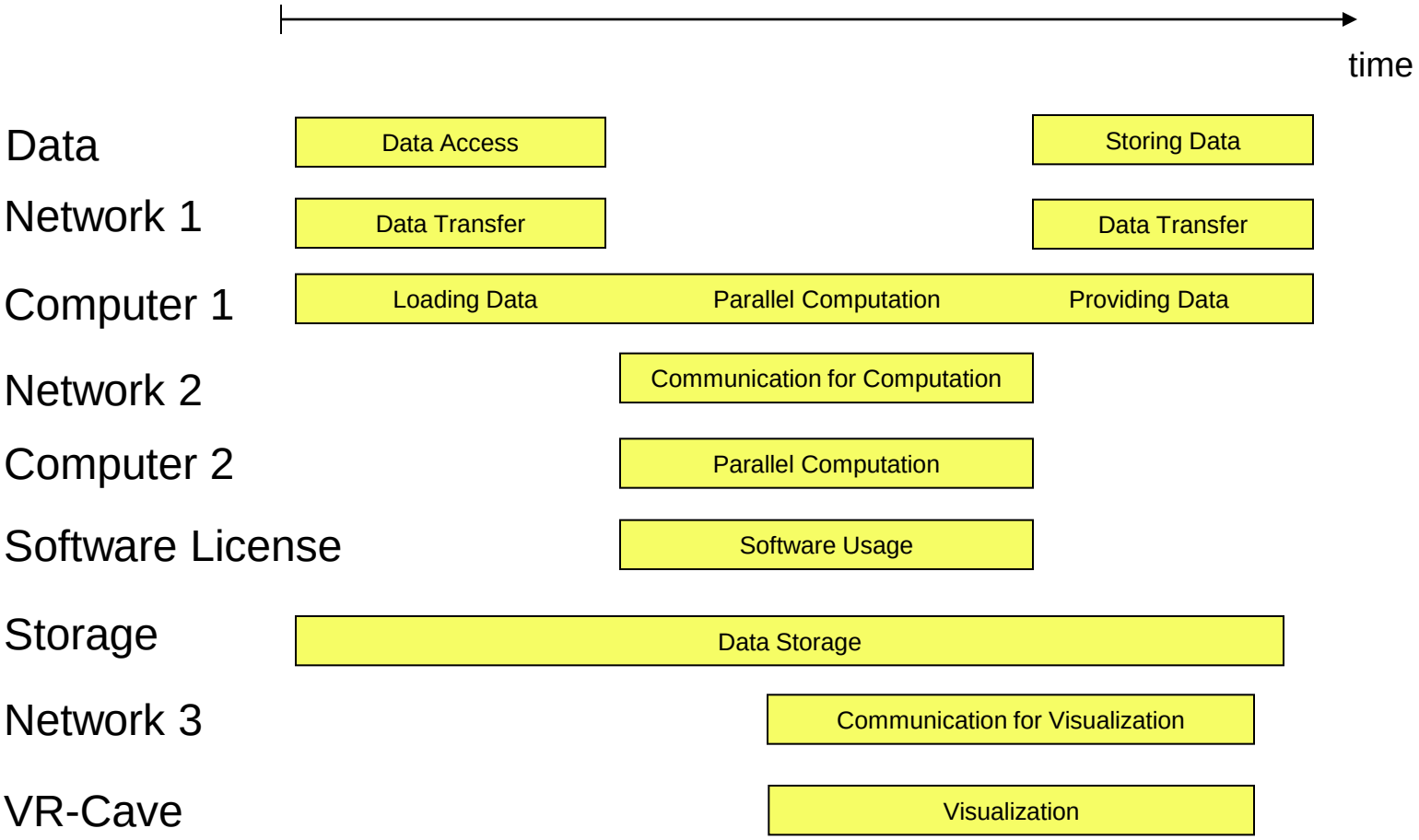
Aspects

- Define Requirements for a Scheduling Architecture
  - not algorithms
- Support for different scheduling instances
  - different local management systems
  - different scheduling algorithms/strategies
- For arbitrary resources
  - not only computing resources, also
  - data, storage, network, software etc.
- Support for co-allocation and reservation
  - necessary for coordinated grid usage (see data, network, software, storage)
- Different scheduling objectives
  - cost, quality, other
- Based on GGF standards and services
  - Use existing services, generate input for extensions

Use-Case for a Grid Job

- Required Resources:
  - needs 48 processing nodes of a specified architecture/properties for 6 hours
  - a nearby visualization device is requested during execution
- Allocation Time Requirements:
  - should be executed between 8am and 6pm the following day
- Data Requirements:
  - needs a specified data set
- Storage Requirements:
  - needs 1 GB of storage during execution
- Software Requirements:
  - utilizies a specified licensed software package
- Network Requirements:
  - a network connection with a given bandwidth between the VR device and the application is needed
- Cost Requirements:
  - The user is willing to pay at most 4$
- Objectives:
  - He prefers cheaper job execution over an earlier execution.

Possible Resource Allocation for the Job



Tasks

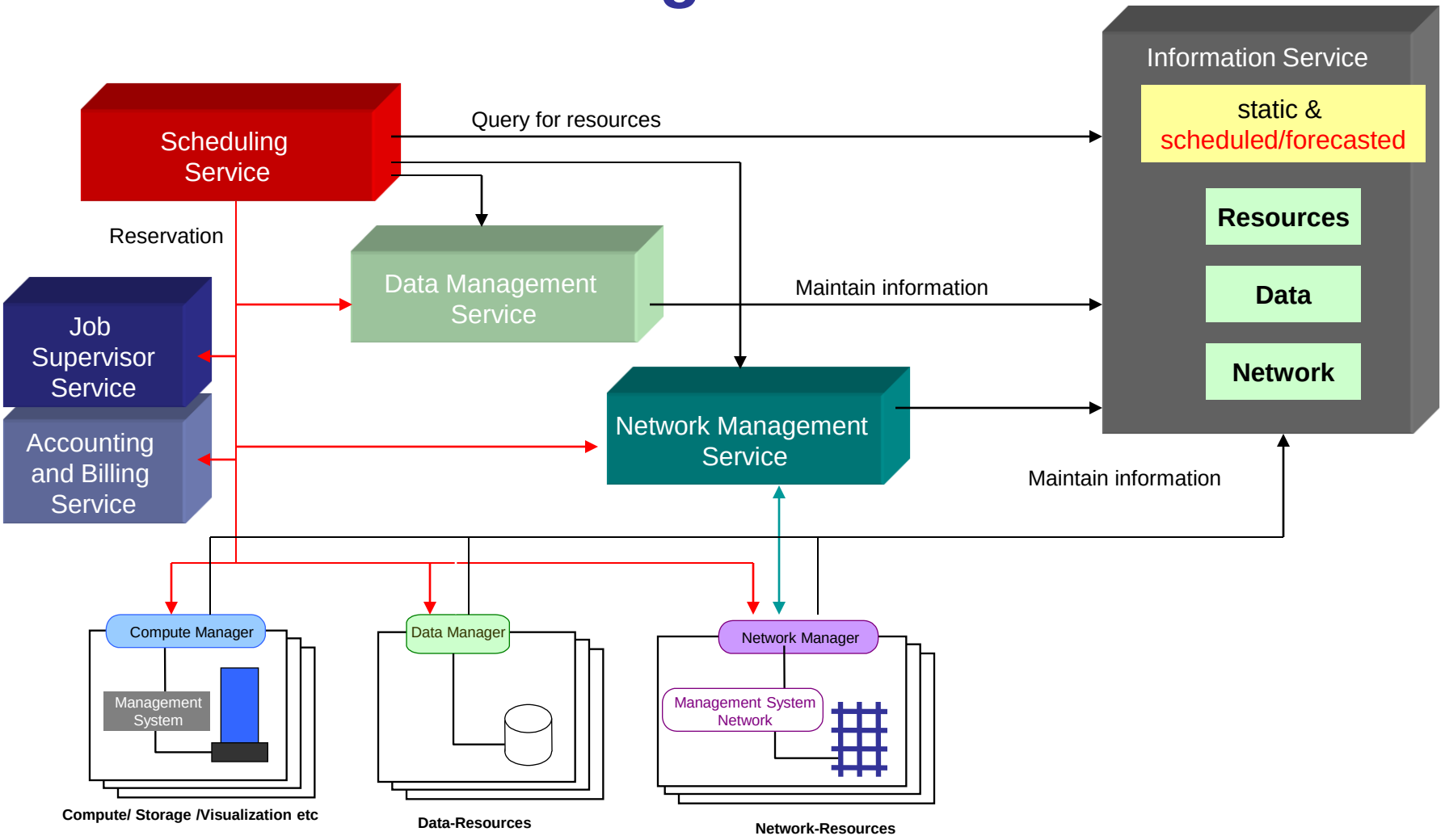
- It is the task of Grid Scheduling to determine, reserve and allocate all Grid resources that are required and available for the execution of a job.
  - support for QoS; Reservation; include data, storage, network; support workflow of jobs; …
  - ➔ 1) Define the use-cases/requirements what a Grid Scheduler should be able to support
  - ➔ 2) Define the necessary architecture to facilitate such Grid Scheduler.

- To this end Grid Scheduler need to interact and co-operate with various services in a Grid:
  - **Local (hardware and system software) resource management/scheduling systems**
  - **Grid resource allocation**
  - **Grid data management**
  - **Grid information**
  - **Grid accounting**
  - **Grid security**
  - ➔ 3) Identify the necessary protocols and information for a Grid Scheduling Architecture

Scheduling Services

- Goal: identify necessary information and components for grid scheduling
  - no specification of implementation aspects yet
  - support for several implementations and different algorithms
  - may not be new “services“
    - current or proposed GGF components already provide services
  - based on OGS* web services
  - not single/central services:
    - e.g. several instances of Scheduling Services possible
  - higher level grid scheduling services should be as independent as possible from actual resource type
  - however, probably different for data and network scheduling!
    - **Data and network connects job parts**
    - **Both manageable, depend on resource selection**

Scheduling Architecture from a Scheduling Point of View



Scheduling Process

Scheduling Service:

1. receives job description;
2. queries Information Service for static resource information;
3. prioritizes and pre-selects resources
4. queries for dynamic information about resource availability
5. queries Data and Network Management Services
6. generate schedule for job
7. reservates allocation if possible
   otherwise select other allocation
8. delegates job monitoring to Job Supervisor

9. Job Supervisor/Network and Data Management Services monitors and initiates allocation

Example:

several resources of requested type found

Several resources selected

A subset is actually available

Determine network and data dependencies

Evaluate utility function

After several tried combination an allocation is confirmed.

Provide all resources and start job

Data Management Service

- Manages all data related actions
- provides static and forecasted information about
  - data location,
  - scheduled transfers
- answers for queries about data requirements of jobs
  - replication, migration, synchronization
  - cost evaluation
  - scheduling of data transfers (e.g. pre-fetching of data)
  - queries for resources (Scheduling or Network Service)
- initiates and monitors scheduled data transfers

Network Management Service

- Manages all network related actions
- provides static and forecasted information about
  - network connections,
  - scheduled allocations and reservations
- answers for queries about network requirements of jobs
  - determination of connection properties between resources
  - cost evaluation
  - scheduling of network transfers
- initiates and monitors scheduled network QoS features

Extending Information Services

- Information classes should include
  - Additional data information
  - Additional network information

- Information Service needs support for
  - access to scheduled and forecasted information,
  - uniform access to information,
  - delegate dynamic information requests to corresponding resource services

  - Security concept for information access!

Job Supervisor

- Monitors scheduled allocations for jobs.
- Resource failure or withdrawal before the job start causes
  - user notification
  - initiating rescheduling

- Initiates the setup and execution of job sets in cooperation with
  - resource management systems
  - data management service
  - network management service

Accounting and Billing Service

- Scheduling and resource access in the grid will require cost management.

- Trusted accounting services are necessary.
  - User approves transactions at request.
  - Billing service includes budget management.
  - Payment type is dependent on user preference.
  - no central service, trust relation ship to certified service providers.

Performance Metrics

$$AverageWaitTime = \frac{1}{N} \sum_{j \in Jobs} (StartTime_j - SubmitTime_j)$$

$$AverageTurnaroundTime = \frac{1}{N} \sum_{j \in Jobs} (EndTime_j - SubmitTime_j)$$

$$FractionOfJobsTransferred = \frac{NumberofJobsMigrated}{TotalNumberOfJobs}$$

$$FractionDataVolumeMigrated = \frac{\sum_K (InputSize_K + OutputSize_K)}{\sum_J (InputSize_J + OutputSize_J)}$$

$$DataMigrationOverhead = \frac{TotalDataMigrationTime}{\sum_J (EndTime_J - QueueTime_J)}$$

Exam's quizzes

- **1.** Cum pot fi organizate, din punct de vedere arhitectural planificatoarele? Care sunt dezavantajele fiecărui model?
- **2.** Descrieţi modelul arhitecturii descentralizate.
- **3.** Care sunt criteriile ce trebuie considerata în faza de proiectare a unei arhitecturi de planificare?
- **4.** Descrieţi pe scurt care sunt metricile de performanţă pentru o arhitectură de planificare.