Job scheduling and Workflow management for high reliable simulation

Shingo Masuda
laboratory for Internet Architecture and Systems

2006/02/24

Our research theme

- Developing the Grid system for scientific simulation
  - Provides high through put computing
    - Reduce time cost
  - User-friendliness
    - Reduce human cost

Grid vs. Ubiquitous

- Grid is the system concepts of
  - Collecting many computer resources *from many places*
- Ubiquitous is the mean of
  - There are computers *everywhere*
- Commonness
  - Users can use computer resources which they needs.
  - Users don't need consider where they are

Background

- Scientists want to
  - Get high reliable and high accurate results
  - Reduce costs (time, computer and human resources)
- Conflicted requirements and it need
  - Effective use of computer resources

Our approach

- Re-scheduling computer resource allocation
  - There are different availabilities to provide high accurate results in jobs.
  - Prioritize jobs which will be related to provide high accurate results finally .
  - Choose these jobs using process reports.
- Other jobs
  - If spend many computer resources for other jobs
  - Much time will waste. (Can not reduce time cost)

Related work

- Job scheduling
  - Jochen Krallmann: On the Design and Evaluation of Job Scheduling Algorithms
- Cluster/Grid computing middleware
  - Globus, Condor, OpenPBS
- Problem
  - Not consider real accuracy level of results
    - To improve accuracy we must dynamically change job scheduling using interim results

Meaning of word

- "Job" is
  - Job ≒ Unix (OS) process (In several cases)
    - Single process
    - Multiple process (MPI, PVM etc)
  - Processes which are performed in computing environment. (cluster, grid etc)
  - Submit job (Run job)
    - Job consists script program = Multiple processes will be executed

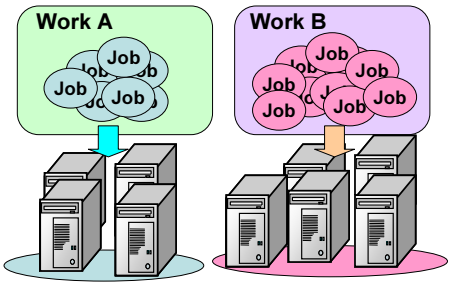
Meaning of word

- "Workflow" is
  - Relation and ordering of executing jobs
  - Work ≒ Job
    - Sometimes a work consists multiple jobs
  - Developed for business computing
    - Workflow management tools
    - Workflow describe language (BPEL, WSDL)
- Purpose of workflow tools (in scientific simulation)
  - Automation and Constructing of simulation process
  - Scientific simulation has multiple works
    - Several calculation process, produce final result, visualization

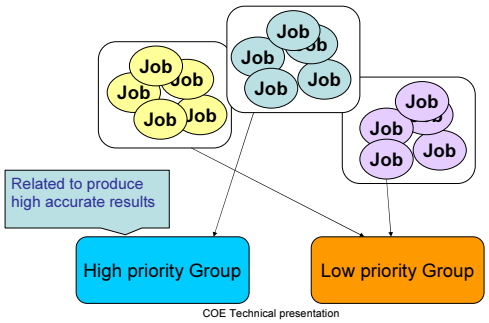
Job scheduling (Computer Resource Allocation)

- To perform multiple works concurrency
  - Each work may have multiple jobs
    - Use large number of computers
  - Divide computer resources
    - Work "A" use computer A1, A2, A3…
    - Work "B" use computer B1, B2, B3…
- Total time to get certain accurate level results depends how allocate computer resources to works => Jobs

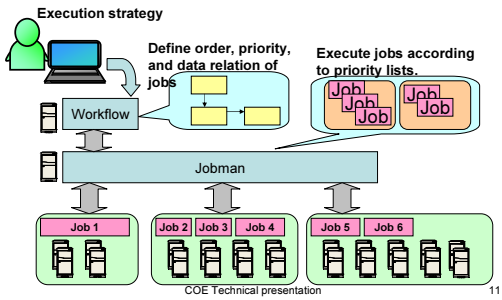
Job scheduling (Computer Resource Allocation)



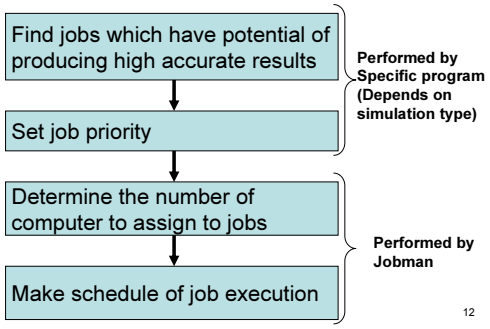
Prioritize jobs



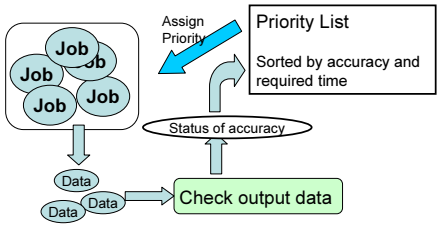
System Architecture



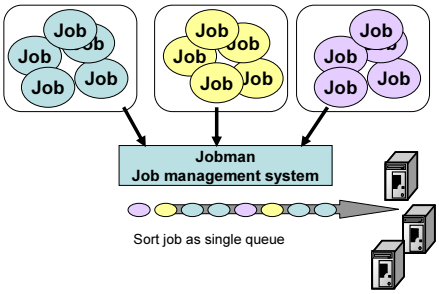
Steps of Prioritizing jobs



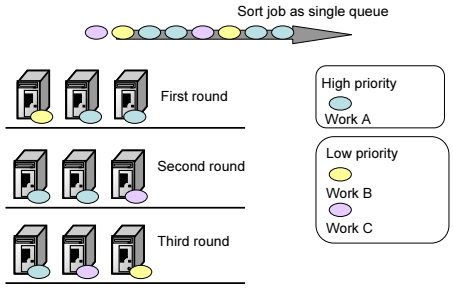
Steps of determining of job priority



Steps of applying computer resource allocation

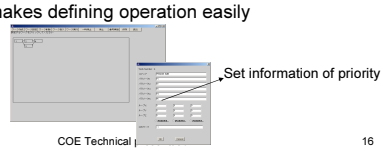


Steps of applying computer resource allocation

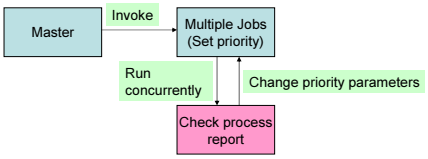


Defining tool of workflow

- Merit
  - Flexibility
    - User can select information for prioritizing jobs
    - User can change how prioritize jobs
  - User-friendly
    - GUI makes defining operation easily

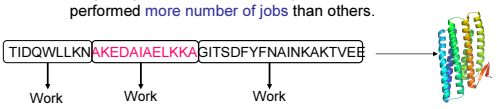


Workflow of job set priority



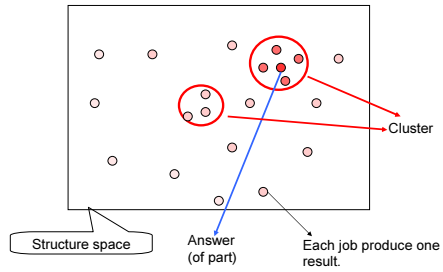
Applying to real application

- Protein structure prediction
  - Three dimensional structure from amino-acid sequence
  - Simulation consists some divided parts (called “domain”)
  - Each domain have different difficulties.
    - To get high accurate result finally some domain need to be performed more number of jobs than others.



Applying to real application

Protein structure prediction



Applying to real application

- High Prioritized jobs
  - In default, Long length part of amino-acid sequence
  - Change priority using clustering result
    - Cluster consists similar results.
    - If cluster size is smaller than others, more number of jobs of its part must be performed.
    - Set high priority to jobs of part whose clustering size is small

Result and Future work

- Result of prototype implementation
  - Our approach have availability of improving reliability and accuracy of simulation
- Future work
  - Improve accuracy of function of determining the number of resources to assign work