International conference on Intelligent Systems, Data Mining and Information Technology (ICIDIT'2014), April 21-22, 2014 Bangkok (Thailand)

Issues Related to Scheduling in Grid Computing

Sandeep Kaur

***Abstract*---**-Grid scheduling task consist of both remote and local user. The grid scheduler does not own the local hosts, therefore does not have control over them. The grid scheduler must make best effort decisions and submit the jobs to the hosts selected, generally as a user. This paper consist of some grid scheduling tasks and details of them.

***Keywords*---**Grid scheduler, Remote, Local user, Local hosts.

I. Introduction

GRID computing offers many solutions are constructed using a variety of technologies and open standards. It provides scalable, highly secure and extremely high-performance mechanisms for discovering and negotiating access to remote computing resources [5] in a seamless manner. This makes it possible for sharing of computing resources, on an unprecedented scale, among an infinite number of geographically distributed groups. This serves as a significant transformation agent for individual and corporate implementations immediate computing practices, toward a general-purpose utility approach very similar in concept to providing electricity or water types of utilities, much like Grid computing utilities are available on demand and will always be capable of providing an always available facility negotiated for individual or corporate utilization. Grid computing environment must be constructed upon the following foundations:

A. Coordinated resources

Avoid building grid system with a centralized control, and should provide the necessary infrastructure for coordination among the resources based on respective policies and service level agreements.

B. Open standard protocols and frameworks:

The use of open standards provide interoperability and integration facilities. These standards must be applied for resource discovery, resource access and resource coordination. Workflow scheduling is one of the key issues in the management of workflow execution. Scheduling is a process that maps and manages execution. The Grid computing applications to have common needs.

C. Application partitioning

That involves breaking the problem into discrete pieces.

Sandeep Kaur is M.Sc, M.Tech, in CSE from Lovely Professional University, Phagwara (Punjab).

D. Discovery and scheduling

Is the tasks and workflow.

E. Provisioning and distributing

Application codes to specific system nodes

F. Results management

It assisting in the decision processes of the environment.

G. Autonomic features

such as self-configuration, self-optimization, self-recovery and self-management.

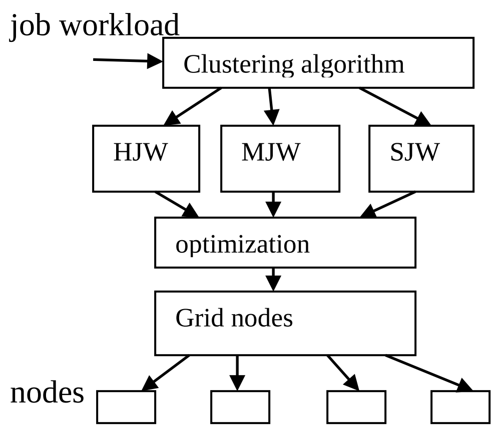

Fig. 1 Grid scheduling process[7]

Example of Grid issues related to scheduling :

NASA *Information Power* ***Grid (IPG)***

NASA's Information Power Grid (IPG) is a high-performance computational and data grid. Grid users can access widely distributed heterogeneous resources from any locations, with IPG middleware adding security, uniformity and control.

Some of the major projects undertaken by IPG are:

II. Resource Broker

A grid user has to make a resource selection from a large number and variety of resources that they could use for an application. For each potential resources .the resource selection system consider the following factors:

A. Computer system characteristics, such as amount of memory, amount of disk space, CPU speed, number of CPUs, type of operating system, available software, and so on
B. The time required for the execution of the job
C. The cost to use that resource or computer system

III. Performance Prediction

There are several types of predictions that are useful when deciding where to run applications. These include job/application execution time on different computer systems, wait time in scheduling queues before the job begins executing. and the time to transfer files between computer systems.

- **Job Manager**

It is used to reliable execute jobs maintain information about jobs. These jobs consists of file operations and execution operations.

- **Portability Manager (PM)**

Portability is a key issue with the grid environment and PM is responsible for the establishment of a suitable environment for the execution of the user applications by automatically identifying the dependencies of each user programme. Framework for Control and Observation in **Distributed Environments (CODE).**

The CODE project provides a secure, scalable, and extensible framework for making observation on remote computer system. It then transmits this observational data to where it is needed, performing actions on remote computer systems and analysing observational data to determine what actions should be taken. Observational data is transmitted using a distributed event service.

- **Test and Monitoring Service**

The IPG Test and Monitoring Service will provide a framework for examining the health of the grid, so the problem with, or degradation of, grid resources are promptly detected; the appropriate organization, system administrator[5], or user is notified; and solutions are dispatched in a timely manner.

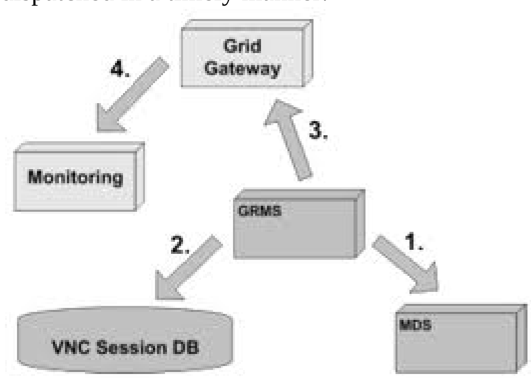

Fig. 2 Grid process [7]

- **Dynamic Accounting System (DAS)**

DAS provides the following enhanced categories of accounting functionality to the IPG community:

- Allows a grid user to request access to a local resource via the presentation of grid credentials
- Determines and grants the appropriate authorizations for a user to access a local resource without requiring a pre-existing account on the resource to govern local authorizations
- Exchanges allocation data between sites to manage allocations in a grid-wide manner instead of a site-specific manner
- Provides resource pricing information on the grid
- Collects and reports the necessary data to ensure accountability of grid users for the use of resources and to enable resource providers to better manage their grid resources [5].

IV. General Adaptive Scheduling Algorithm

- ✓ While there are tasks to schedule
- ✓ for all task I to schedule
- ✓ for all host j
- ✓ compute CT i,j = CT(task I, host j)
- ✓ end for
- ✓ compute metric i = f(CTi,l, CTi,2, ……)
- ✓ end for
- ✓ select best metric match m
- ✓ compute minimum CT m,n
- ✓ schedule task m on n
- ✓ end while[3]

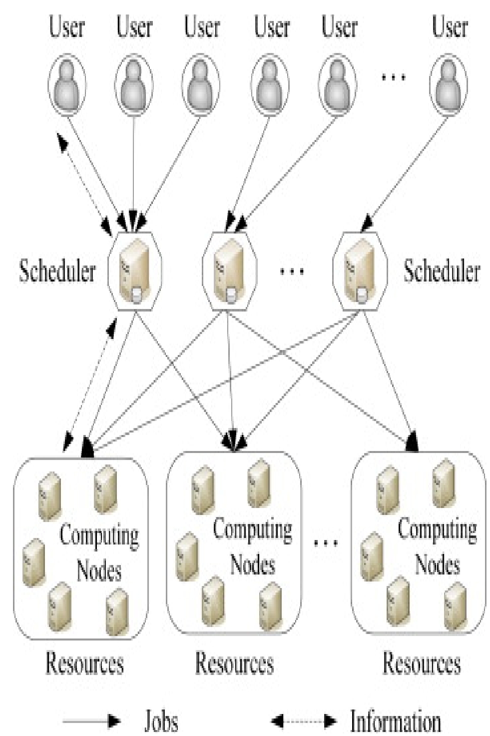

Fig. 3 Grid Scheduling process[6]

V. ACO Algorithm Techniques

It is very important how the graph of the problem is created. Another research direction is to create different heuristic based algorithms for problem arising in grid computing. ACO is basic strategies for a grid scheduling is formulated.[1]

Makespan for the execution on first free machine and ACO algorithm.[1]

Online-mode	ACO	improvement
80	67	16%
174	128	26.4%
95	80	15.8%

Scheduling algorithms is designed for distributed systems shared asynchronously by both remote and local users. Scheduling request applications, the scheduler allocates the application to the host by selecting the best match from the pool of applications and pool of the available hosts. Existing mapping heuristics can be divided into two categories:-[1]

- ***On-line-mode*:**

A task is mapped onto a machine as soon as it arrives at the map per.

- ***Batch mode*:**

Tasks are not mapped onto the machines as they arrive, instead they are collected in a set that is examined.

VI. Conclusion

Grid computing is becoming a critical component of science, business, industry.[6] A user buys a computing time on demand so in case of scheduling remote and local user. Grid are networks that include computer cluster, cluster of cluster, or special data sources.

References

[1] Stefka Fidanova and Mariya Durchova,"Ant algorithm for grid scheduling problem" published in 2006 springer-verlag Berlin Heidelberg. pp 405-412.

[2] Jia Yu, Rajkumar Buyya and Kotagiri Ramamohanarao,"Workflow Scheduling algorithm for grid computing"The university of Melbourne,VIC austrelia.

[3] Xiaoshan He,Xian-He Sun,Gregor Von Laszewski,"A Qos Guided Scheduling Algorithm for Grid computing" USA pp-1 to 15.

[4] Chunlin LI,Layuan Li,"A distributed iterative algorithm for optimal scheduling in Grid computing" published in 2007,vol. 26,pp 605-626.

[5] Joshy Joseph and Craig Fellenstein, "Grid Computing" Pearson publication by IBM press.

[6] WWW.aip.org/tip/INPHFA/vol-9/iss-4/p31.html

[7] WWW.google.co.in

Mrs. Sandeep kaur M.Sc, M.Tech in CSE from Lovely Professional University Phagwara (Punjab). have published three research papers: One is in International Conference that was held 2013 in USA and others in IJEAT. Research area is Distributed image processing, Software engineering, Data base, Software Analysis and Design, Networking.