A Survey of Tasks Scheduling Algorithms in Distributed Computing Systems

2

Nutan Kumari Chauhan
KIET Group of Institutions, India

Harendra Kumar
https://orcid.org/0000-0002-3394-5655
Gurukula Kangri Vishwavidyalaya, India

INTRODUCTION

There are various causes for using DCS. The nature of equipment may involve the utilization of a communication network which connected by some computers: for eg, data created in one site and needed in another site. There are various cases in which only one computer is required, but DCS is very helpful for practical causes. For eg., it may be extra cost-efficient to get the inclined level of performance by applying cluster of numerous low-end computers, in similarity with only high-end computer. DCS has no single point of failure, so it can offer additional reliability than a non-distributed system. DCS may be simple to expand and run than a uni-processor system. DCS have used in a large scale due to several characteristics such as communication, computational speedup, fault tolerance, increased throughput, resource sharing etc.

Tasks scheduling is very important phase in DCS, mainly it is assigning of tasks onto processors as a manner so that some efficiency actions are optimized. An effective tasks scheduling policy avoid too much inter-tasks communication and utilize the precise effectiveness of the processors. It is extremely important step in exploiting the capacity of system and it may be done in two ways: **Dynamic Scheduling** and **Static Scheduling**. The most important plan of tasks scheduling is toward decrease the running cost of a program. The cost can be characterized by time, money or some other measure of resources usage.

In DCS, resources of system can be split by users at different sites through the communication link. However in proper utilization of total computing power of DCS, a fundamental problem arises. It may happen that a certain processor has very few workloads while another processor has more workload at a given time. It is required to distribute total workload of DCS of its overall processors. This evades under use of power and further it decreases response time for deputize at more heavily loaded processor. Generally, the form of computing power sharing with the intention of getting better performance of a DCS by rearranging the workload between the available processors called load balancing. The progress in the presentation of system through redistributing the workload among processors is the main aim of load balancing policies.

The reliability of DCS is probability that system can run the whole application successfully. If system is complex then it is very difficult to system in it's entirely. The logical technique is applied to crumble the entire system into function entities compiled of subsystems and units. Each entity is either operational or failed. The sub-division generates the block diagram of system. Models are prepared to fit the logical arrangement and calculus of probability is applied to figure out reliability of system in terms of sub-division reliabilities. There are two types of reliabilities in DCS: first is processor-related reliability,

DOI: 10.4018/978-1-7998-3479-3.ch018

Table 1. Difference between serial and parallel processing

S. No.	Serial Processing	Parallel Processing
1.	There exist several tasks and processors even although only single task is active on a processor at a given time	More than one task of the program may execute concurrently for various periods during the lifetime of a program
2.	The main concern in a serial processing of programs is variability or non-uniformity of computation requirements and the overhead of communication	The main concern of parallel processing is to reduce the total time of the program by running different parts of the program
3.	The challenge of distributed processing lies in matching requirements against resources without incurring excessive communication overhead	In parallel processing, arranging the order of execution of the various tasks on the processors is a major problem
4.	Generally, serial processing is used for long distance	Generally, parallel processing is used for short distance

for example reliability of resources and reliability of computation; another is the path / link-related reliability, for example reliability of communication paths between allocated processors.

The processing of a program is of two types: serial processing or parallel processing.

Difference between serial processing and parallel processing are given in the following table-1:

Figure 1 and Figure 2 show the execution of tasks on processors in a serial and parallel way respectively. A serial program made up of five tasks is executing on a single processor in Figure 1(a), while the same serial program is executed over a two processors system in Figure 1(b). Processor P2 can execute two tasks (3 and 4) more efficiently but due to transfers of control between tasks over a communication link incur appreciable overhead.

Figure 1. Serial Processing

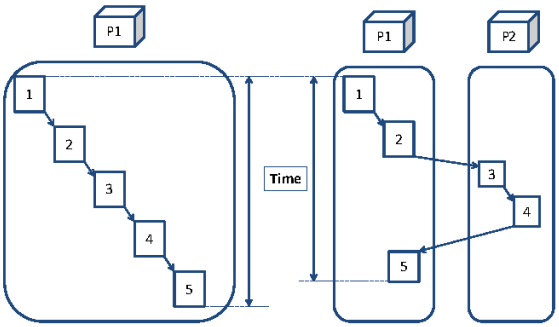
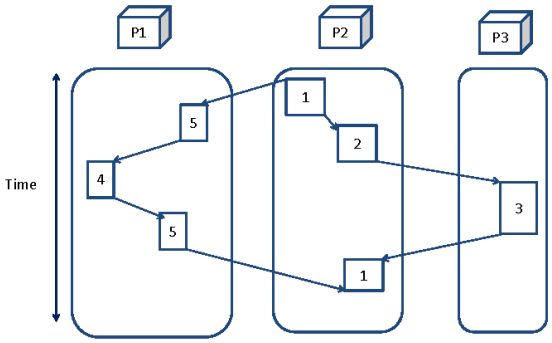


Figure 2. Parallel Processing



Further the chapter is planned as follows: Section 2 portrays the background of task scheduling techniques. Section 3 covers the application of task scheduling in DCS. Section 4 describes the future research directions. In section 5, conclusion is discussed on the basis of various factors in task scheduling techniques.