Vol 7 Issue 3 Dec 2017

ISSN No : 2249-894X

Monthly Multidisciplinary Research Journal

Review Of Research Journal

Welcome to Review Of Research

RNI MAHMUL/2011/38595 **ISSN No.2249-894X**

Review Of Research Journal is a multidisciplinary research journal, published monthly in English, Hindi & Marathi Language. All research papers submitted to the journal will be double - blind peer reviewed referred by members of the editorial Board readers will include investigator in universities, research institutes government and industry with research interest in the general subjects.

Address:-Ashok Yakkaldevi 258/34, Raviwar Peth, Solapur - 413 005 Maharashtra, India
Cell : 9595 359 435, Ph No: 02172372010 Email: ayisrj@yahoo.in Website: www.oldror.lbp.world

Review Of Research
ISSN: 2249-894X
Impact Factor : 5.2331(UIF)
Volume - 7 | Issue - 3 | DECEMBER - 2017

DISTRIBUTED JOB PROCESSING SYSTEMS AND ITS CHALLENGES

Dr. Putti SrinivasaRao
Associate Professor and Dean of PG Studies ,
JBIET Hyderabad Telanagana India.

ABSTRACT: -

A distributed system consists of many heterogeneous processors with different processing power and all processors are interconnected with a communication channel. Today, various systems those are commercially available have not addressed several practical concerns for the end users. Even today, systems are too complex in nature for the average end-user to manage them. It is quite expensive even to enhance the systems to meet newer business requirements. this paper presents the challenges in the present distributed systems and solutions to overcome those challenges in distributed systems .

KEYWORDS: *Job processing Systems ,communication channel , heterogeneous processors .*

I. INTRODUCTION:

In a distributed system consists of many heterogeneous processors, if some processors are less loaded or idle and others are heavily loaded, the system performance will be reduced drastically. System performance can be improved by using proper load balancing.

Some of the concerns are:

- Support of the framework on newer hardware after 5 or 10 years.
- Data being portable or upgradable in future.

- Reports and other artefacts that user generates being readable in future.
- Managing the system when no further support is available from the original vendor.
- System being expanded to handle more loads in future.
- Amount of load the framework can support to the maximum.

It is of course difficult to answer all these questions given the fact that technologies will in-fact keep changing, computer hardware will have to be upgraded, operating systems will have to be upgraded to newer versions.

The solution lies in designing software applications that are simpler in design, yet does the required processing with the required level of comfort

and performance. With the basic technologies, without any advanced tools, it is indeed possible to develop systems that can solve most of the complex business problems, will be flexible enough to extend it to newer technologies, and will be able to expand to take up more processing load.

II .DISTRIBUTED COMPUTING

In order to leverage the computing power available with large number of inexpensive set of computers, it is important that these systems be interconnected through a Local Area Network or Wide Area Network.

In order to leverage the power of such a distributed and heterogeneous network of inexpensive computing environments the importance of effective distribution of tasks across the processors is undisputable. Added to this, it may be necessary to consider the importance of cost of computation, cost of network and cost of storage while distributing the tasks, so that the overall cost is reduced to the most optimum level.

The most important aspect of a distributed computing environment is effective coordination of the resources available across several nodes. There are several frameworks that are available for distributed environments.

The cost factors that are generally considered are:

a. Computational cost b. Storage cost c. Network cost
b. Cost of other resources like memory, administration & monitoring etc.

A distributed computing environment is used for processing of the following types:

a. Processing large volume of data.
b. Computation that takes considerable time when run on a single node.
c. Large number of tasks, each taking small amount of time to complete.

So, the fundamental concept of Distributed Computing is to break a large task into smaller tasks that can be performed independently. Each such smaller task is distributed to different nodes and each such node executes the task completely. Finally, the results from all these nodes, if needed, may be collated and used for subsequent processing/reporting purposes.

The parameters considered for the study are listed below. Each factor adds to the complexity in load computation and scheduling/load distribution

a.Reasonably complex computation (e.g. computing large number of Primes)
b.Computing capabilities and ability to take the load
c.Communication infrastructure
d.Storage infrastructure
e.Network / topology
f.Distribution of load

III JOB PROCESSING IN DISTRIBUTED SYSTEMS

The basic requirements of a Job Processing system are:

1.Accepting jobs from different channels
2.Scheduling the jobs
3.Dispatching to the appropriate processor
4.Ensuring the job is completed

Apart from these, a complex job processing system can also have additional requirements like:

1.Managing different job priorities 2. Handling job dependencies
3.Notification mechanisms 4. Monitoring and Control

IV CHALLENGES IN DISTRIBUTED SYSTEMS:

However, there are many challenges in this type of a system . Some of the important challenges are given below.
1.The input channels to be supported and the technique therein.
2.Ensuring that once a job is accepted, the system processes the job, completes the job and the job reaches a final state.
3.In case of crashes, it should be possible to re-start the job and it should continue from where it had left earlier.
4.For a long running job, it should be possible to track the progress of the job processing.

To start with, considering that the jobs are identical. This means, the jobs generally will have same computational requirement, identical data sets to work on and identical means of reporting back the result.

The simple design is similar to the diagram given below.

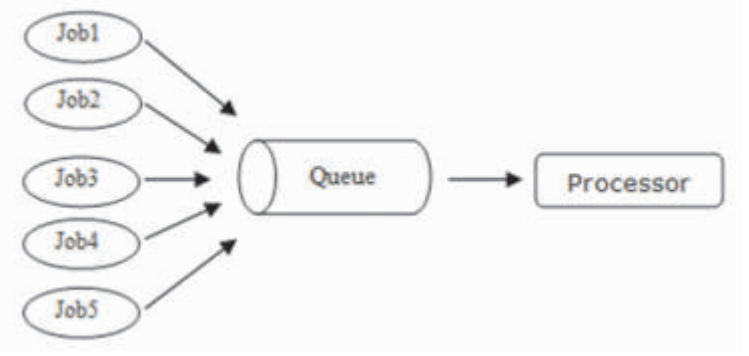

Figure 1.1: Job Processing Model (Single Processor)

Considering that it has only one processor that processes all the jobs and the jobs are submitted to a queue that the processor listens on. The total time taken to complete processing of the n jobs can be given by the following:

$$T = \sum_{i=1}^{n} w_i p_i$$

Where:
T = total time taken to process n jobs
n = total number of jobs
w = wait time of the job before the processing starts
p = processing time of the job within the processor

V Scheduling and Load Balancing in Distributed Systems

Scheduling : Scheduling is most important and pays a crucial role in improving the systems performance in distributed systems . From the research point of view the distributed scheduling algorithms can be classified in three categories, i.e. 1-Sender Initiated Algorithms 2-Receiver Initiated Algorithms and 3-Symmetric Algorithms.

The basic idea of distributed scheduling is to improve the performance. The challenge in scheduling of distributed systems is how to schedule the tasks in order to be processed effectively .

Load Balancing: Load balancing is technique to distribute load among all processors connected through network effectively. The excess load or unexecuted load is migrated from one processor to other processors to balance the load .

ADVANTAGES OF LOAD BALANCING

1.Improve the overall system performance 2.Reduce the job idle time

3.Utilize the resources effectively 4. good response time
5.High throughput and reliability

Many load balancing algorithms have been developed but no single algorithms are suitable for all applications .

VI MOTIVATION AND RESEARCH CHALLENGES

Most of the today's task execution Distributed Systems are based on Master/Slave Architecture (Centralized). The problem with this approach is poor scalability. To avoid single component as a manager it must move to Decentralized architecture.

The issues or challenges in decentralized approach is

1. Distributing tasks
2. Load Balancing among nodes

The new approach must take into account the drawbacks and limitations in the existing implementations of other job processing systems and provide an approach to eliminate and/or minimise the impact of these limitations. The new approach must provide a mechanism for improving the throughput.

The architecture and design must be such that each of the components are self-contained and do not depend on each other. The load balancing, fault-tolerance and failover recovery must be built into the system .The communication overhead for the processors in enquiring loads of peer processors must be reduced by making the load information available at a central place (with clustered and failover mechanism to avoid single point failures). We can find cost factor for Job processing by introducing network cost, storage cost and computational cost for each processor. The cost is computed as overhead percentage and the best processor is calculated based on the processor having the least cost.The overall system must have the capability to scale horizontally as well as vertically to achieve the required performance, and minimize the total cost. The System can have the option to use processor affinity and thread pools.

CONCLUSION:

The idea presented here uses the currently available technology to implement a scalable, high performing, cost effective job processing system that can replace monolithic batch processing systems with complex scheduling logic for managing jobs. This technique can further be enhanced to include a monitor to dynamically upgrade and downgrade the processors in a time-bound manner or on a real-time monitoring of the system.

REFERENCES:

[1] Zeng Zeng, and Bharadwaj Veeravalli "Design and Performance Evaluation of Queue-and-Rate-Adjustment Dynamic Load Balancing Policies for Distributed Networks", IEEE Transactions on computers, VOL. 55, NO. 11, NOVEMBER 2006.

[2] Abhijit , Rajguruand and S.S. Apte "A Comparative Performance Analysis of Load Balancing Algorithms in Distributed System using Qualitative Parameters", International Journal of Recent Technology and Engineering (IJRTE) , ISSN: 2277-3878 Volume-1, Issue-3, August 2012.

[3] Iman Sadoop, Sandeep Palur and Ioan Raicu "Achieving Effiecient Distributed Scheduling with Mesage Queue in Cloud Computing for Many-Task Computing and High –Performance Computing".

[4] Andrei Radulescu, Arjan J.C. and van Gemund "Low-Cost Task Scheduling for Distributed-Memory Machines" , IEEE Transactions on Parallel and Distributed Systems VOL. 13, NO. 6, JUNE 2002.

[5] K.Q. Yan, S.C. Wang, C.P. Chang and L.Y. Tseng "The Anatomy Study of Load Balancing in Distributed System", Proceedings of the Seventh International Conference on Parallel and Distributed Computing, Applications and Technologies (PDCAT'06). 2006.

[6] "Kevin Barker, Andrey Chernikov, Nikos Chrisochoides, and Keshav Pingali "A Load Balancing Framework for Adaptive and Asynchronous Applications", IEEE Transactions on Parallel and Distributed Systems, VOL. 15, NO. 2,

FEBRUARY 2004.
[7] Jorge E. Pezoa "Decentralized Load Balancing for Improving Reliability in Heterogeneous Distributed Systems ", 2009 International Conference on Parallel Processing Workshops.
[8] Jorge E. Pezoa and Majeed M. Hayat. "Reliability of Heterogeneous Distributed Computing Systems in the Presence of Correlated Failures" . IEEE TRANSACTIONS ON PARALLEL AND DISTRIBUTED SYSTEMS, VOL. 25, NO. 4, APRIL 2014.
[9] Young Joon Lee, Geon Yong Park, Ho Kuen Song, and Hee Yong Youn "A Load Balancing Scheme for Distributed Simulation Based on Multi-Agent System".2012 IEEE 36th International Conference on Computer Software and Applications Workshops.
[10] Deng Huafeng, Zhong Linhui, and Ye Maosheng "An Efficient Load Balancing Algorithm in Distributed Systems". 2010 International Forum on Information Technology and Applications.

Publish Research Article
International Level Multidisciplinary Research Journal
For All Subjects

Dear Sir/Mam,

We invite unpublished Research Paper,Summary of Research Project,Theses,Books and Books Review for publication,you will be pleased to know that our journals are

Associated and Indexed,India

- Directory Of Research Journal Indexing
- International Scientific Journal Consortium Scientific
- OPEN J-GATE

Associated and Indexed,USA

- DOAJ
- EBSCO
- Crossref DOI
- Index Copernicus
- Publication Index
- Academic Journal Database
- Contemporary Research Index
- Academic Paper Databse
- Digital Journals Database
- Current Index to Scholarly Journals
- Elite Scientific Journal Archive
- Directory Of Academic Resources
- Scholar Journal Index
- Recent Science Index
- Scientific Resources Database

Review Of Research Journal
258/34 Raviwar Peth Solapur-
413005,Maharashtra
Contact-9595359435
E-Mail-ayisrj@yahoo.in/ayisrj2011@gmail.com