Challenges of Using Unstructured P2P Systems to Support Distributed Exascale Computing

Zohreh Esmaeili Bidhendi, Pouria Fakhri, Ehsan Mousavi Khaneghah

Department of Computer Engineering, Faculty of Engineering, Shahed University, Tehran, Iran, Zohrehesmaeili@Shahed.ac.ir, Pouriafakhri@Shahed.ac.ir, Emousavi@Shahed.ac.ir

*Correspondence:
Ehsan Mousavi Khaneghah, Department of Computer Engineering, Faculty of Engineering, Shahed University, Tehran, Iran, Emousavi@Shahed.ac.ir

Abstract

The formation of dynamic and interactive events in distributed Exascale systems makes it impossible to use pre-defined response structures to manage the requirements of scientific applications. The nature of the dynamic and interactive events is such that there is a need for scalability and creating new structures to meet the processes' requirements. In unstructured P2P systems, resource discovery activity and responding the process requests are normally based on the absence of a specific and pre-defined structure. Emphasizing the concept of the lack of a response structure in distributed Exascale and unstructured P2P systems, this paper tends to address the challenges of employing the manager of these types of systems in distributed Exascale systems.

Keyword: Distributed Exascale Systems, Unstructured P2P, Response Structure, Scalability

1. Introduction

The formation of dynamic and interactive events in scientific applications has led to the need for distributed Exascale systems, which, in addition to running the application in the shortest response time, should be able to manage dynamic and interactive events and their impact on the application's processes [1]. In traditional scientific applications, due to be obvious nature of the problem, and consequently, the specific requirements of the scientific application to design the computing system, the response structures of the application requests were specified at the time of designing the computing system. Definite response structure enables the system's designer to whether create an appropriate structure to answer the application requests in the existing computing system such as cluster computing systems, or to determine the response structure in a way that at the time of making a request, the system responds to requests based on using the resources that are outside the system or on other systems such as Grid computing systems. In open traditional computing systems, such as grid and P2P, the response structure is determined at the time of designing the computing system, but due to the unavailability of computing resources required by the scientific application, finding and allocating resources are carried out at the time of running the scientific application. [1, 2, 3].

As a result, in this type of computing systems, the concept of scalability is introduced as a means to continue the response process and increase the likelihood of responding to the process requests. In distributed Exascale computing systems, the

nature of the scientific application is in such a way that due to the occurrence of the dynamic and interactive nature, it is not possible to define an exact response structure for the application requirements. In this type of computing systems, proportional to the occurrence of the dynamic and interactive nature, the requirement and execution process may change in a way that the response structure intended to run the application is not able to respond to dynamic and interactive events. In this situation, depending on the nature of the request as well as the dynamic and interactive nature, the system manager may expand the system. This expansion and creating a new response structure may not be in line with the primary response structure.

Therefore, the presence of any non-flexible structure in the system to respond to application requests will make the manager of the distributed Exascale system unable to perform its activities. Taking into account this issue, at first sight, using a particular type of P2P systems known as unstructured P2P systems will be considered as an appropriate environment to running applications with a dynamic and interactive nature. Due to the lack of pre-defined structures, the manager of the distributed Exascale system will manage the dynamic and interactive events based on the scalability mechanisms in case a dynamic and interactive nature occurs.

2. Related works

To support the concept of scalability in distributed Exascale systems, some of the frameworks presented in unstructured P2P systems and resource discovery techniques in these systems have been used. Some of these techniques and frameworks are designed for responding to dynamic requests of the requesting process and to improve parameters such as: the number of associated machines in the resource discovery operation, the time required to process the information in the resource discovery operation, the processing time in the machines, and the number of faults in each request.

One of the main problems in unstructured P2P systems is the heavy traffic, which is mainly because of the high volume of responses to queries most of which are not intended by users. One of the solutions presented in [4] for this problem is using top-K query, where a score is considered for each response to a query through a scoring function and the user can receive a specific (K) number of responses related to queries via the system. A framework called fully distributed is presented to implement it. Another resource discovery technique in unstructured P2P systems is Learning Automata-based Resource Discovery which tends to find the shortest possible path that connects the user to the corresponding node for responding to a process request. This technique can reduce the problems caused by the flooding method. [5]

Another technique used in unstructured P2P systems is Improved Adaptive Probabilistic Search. This technique considers the file's type and score based on previous searches for this file type to minimize the search space and the search overhead. Improved Adaptive Probabilistic Search includes four main phases: 1- Search, 2- Selecting the neighbor, 3- Scoring, and 4- Flow control. In the search phase, if a node receives a request for a file, it will search it in its local database. If the file exists, the corresponding node will send the response, otherwise, it will send the request to its other neighbors. In the second phase, the intended neighbor is selected

based on the information stored in the nodes' index tables. In the scoring phase, the score of each file in each node increases or decreases based on the number of search requests for that file, and the flow control phase is used to control and limit the number of received messages [6].

3. Challenges

Multiple resource discovery mechanisms in unstructured P2P systems are considered as tools of finding a resource and adding it to the system to continue responding to processes' requests. Resource discovery increases the system limits. Increasing the system limits enables the system to in addition to responding to the process request which leads to calling the resource discovery manager, increase the likelihood of responding to future requests. The nature of the scalability is such that its limits and constraints follow the constraints of the resource discovery manager. The concept of the constraints governing scalability defines the concept of scalability success. In traditional computing systems, the concept of success and failure follows the Bernoulli distribution. As a result, the scalability is either for responding to requests of the process that activate the resource discovery, or to have no impacts on the activity of the resource discovery activator process.

In these types of systems, besides the scalability from resource discovery manager activities, the scalability from voluntary membership in the system can also be considered. In this type of scalability, the elements containing the resources intended by the process, voluntarily join the system, and this membership also enhances the likelihood of responding to the requests. Since unstructured P2P systems are normally used to share the file resource, the concept of voluntary sharing in these systems can be used as a means to increase the system's capability for responding to process requests.

In unstructured P2P systems, focusing on the resource file, the absence of time and space constraints in responding to processes' requests, not changing the state of the requesting process, not changing the state of the P2P system where the resource requests are made, the independence of the resource discovery process from the frequency of outside changes of the system for resource discovery, as well as employing Bernoulli distribution for the resource discovery process, introduce the concept of system scalability as an appropriate mechanism for answering the processes' requests.

However, in distributed Exascale systems, the occurrence of a dynamic and interactive nature influences all system activities as a main element. Although several functions such as changing the resource function, scalability, and changing the procedure of responding to request from a process with a dynamic and interactive nature, can be considered as patterns of dealing with the dynamic and interactive nature, usually, using the scalability pattern and developing the system resources are employed so that the dynamic and interactive nature can be responded through them.

Unlike P2P systems, in distributed Exascale systems, scalability is not limited to finding a resource. In these types of systems, the concept of scalability is used as a solution to create a response structure and develop the concept of global activity. In these systems, unlike unstructured P2P systems, the process activating the resource

discovery manager can have a request for each of the resources in the system. Consequently, a set of constraints and limitations will govern the process of scalability. In addition to changing the computing system state, the occurrence of the dynamic and interactive nature will also change the conditions and constraints governing the request.

Due to all these factors, in distributed Exascale systems, the concept of scalability can be employed when it can support all these factors as well. Therefore, the resource discovery mechanism, and consequently, the concept of scalability for responding to the process request as well as managing the occurrence of a dynamic and interactive nature in each of the resource discovery and scalability activities need to be changed.

4. Conclusion

In general, scalability is a management tool in unstructured P2P systems, through of this issue, requests of the process with a dynamic and interactive nature, can be answered at the appropriate time. To do this, first it is checked if the machines within the system are capable of answering such requests or not. If they are not able to respond to the requests, the resource discovery manager is called which seeks to answer the requests outside the system and solves the problem through machines outside the system.

References

[1]   Elham Adibi, Ehsan Mousavi Khaneghah. "Challenges of Resource Discovery to Support Distributed Exascale Computing Environment", Azarbaijan Journal of High Pefromance Computing, Vol 1, Issue 2, 2018, pp. 168-178

[2]   Khaneghah, Ehsan Mousavi, Amirhosein Reyhani ShowkatAbad, and Reyhaneh Noorabad Ghahroodi. "Challenges of Process Migration to Support Distributed Exascale Computing Environment." *Proceedings of the 2018 7th International Conference on Software and Computer Applications. ACM, 2018.*

[3]   Mousavi Khaneghah, E., et al. "Modeling and analysis of access transparency and scalability in p2p distributed systems.*" International Journal of Communication Systems 27*.10 (2014): 2190-2214.

[4]   Akbarinia, Reza, Esther Pacitti, and Patrick Valduriez. "Reducing network traffic in unstructured p2p systems using top-k queries.*" Distributed and Parallel Databases* 19.2-3 (2006): 67-86.

[5]   Torkestani, Javad Akbari. "A distributed resource discovery algorithm for P2P grids.*" Journal of Network and Computer Applications 35.6* (2012): 2028-2036.

[6]   Navimipour, Nima Jafari, and Farnaz Sharifi Milani. "A comprehensive study of the resource discovery techniques in Peer-to-Peer networks.*" Peer-to-Peer Networking and Applications 8.3* (2015): 474-492.