International Journal of Science, Environment and Technology, Vol. 4, No 4, 2015, 1037 – 1048

ISSN 2278-3687 (O)
2277-663X (P)

A HEURISTICAPPROACH FOR LOAD BALANCING IN SCALABLE INTERNET SERVICE DEPLOYMENT

[1]Sija Chandran S. and [2]Saumya Raj
[1]Second year MTech Student, [2]Assistant Professor
Department of Computer Science and Engineering
Caarmel Engineering College, Pathanamthitta
E-mails: [1]sijacs66@gmail.com [2]saumya.raj@bccaarmel.ac.in

Abstract - The problem of load balancing in distributed environment is directly related to the allocation of tasks among resources available in the system. The distributed system must have mechanism to deal with faults while providing efficient and reliable services to its end users. Keeping this view in mind, few schemes has been developed and presented in this work.
Index Terms - Centrality, distributed facility location, service migration, user-generated services.

Introduction

The existing system uses the CDSMA algorithm, but this algorithm is used only for the mapping of data. In the work going to propose a new algorithm CDSMA broadcast algorithm. This algorithm will not only maps the data but also finds the list of systems to which the data is going to distribute, finds the neighbor list, finds the appropriate nodes for the transferring and maintains a control structure with trusted systems to maintain security. A scalable and effective heuristic approach to deal with the complexity and limitations of their distributed placement. In this phase will develop a trust system based on processing the payment reports to maintain a trust value for each node. The nodes that relay messages more successfully will have higher trust values, such as the low mobility and the large-hardware-resources nodes. Based on these trust values, will propose a trust based routing protocol to route messages through the highly trusted nodes (which performed packet relay more successfully in the past) to minimize the probability of dropping the messages, and thus improve the network performance in terms of throughput and packet delivery ratio. However, the trust system should be secure against singular and collusive attacks, and the routing protocol should make smart decisions regarding node selection with low overhead. To overcome the scalability

*Received July 1, 2015 * Published Aug 2, 2015 * www.ijset.net*

limitations of classic centralized approaches by re-optimizing the locations and the number of facilities through local optimizations which are refined in several iterations. This work develop a scalable and distributed approach that answers our key question through an iterative reoptimization of the location and the number of facilities within network neighborhoods. Here proposing an innovative approach to migrate, add, or remove servers within limited-scope network neighborhoods by utilizing only local information about the topology and demand. This work shows that even with limited information about the network topology and demand, within one or two hops, our distributed approach achieves performance, under various synthetic and real Internet topologies and workloads, which is comparable to that of optimal, centralized approaches requiring full topology and demand information. This work also shows that it is responsive to volatile demand. Re-optimizations are based on exact topological and demand information from nodes in the immediate vicinity of a facility, assisted by concise approximate representation of demand information from neighboring nodes in the wider domain of the facility. Using extensive synthetic and trace-driven simulations, demonstrates that our distributed approach is able to scale well by utilizing limited local information without making serious performance sacrifices as compared to centralized optimal solutions. Also demonstrate that our distributed approach yields a high performance under demand and imperfect redirection. This approach leverages recent advances in virtualization technology and the flexibility of billing models, such as pay-as-you-go, towards a fully automated Internet-scale service deployment.

A. Modules:

The modules coming under this research work includes:

- Trusted Party Module
- Service Maintain
- CDSMA algorithm
- CDSMA Broadcast algorithm

In trusted Party Module the new nodes are creating, setting the neighbours, putting the nodes range, the overall control is managing and receiving the packets. In the second module called Services Maintain the services are maintaining accordingly that is compiling, save, run and keeping the services. Also to check request setting the priority values. In the CDSMA module the existing algorithm is explaining but this algorithm is used only to map the data hence a enhancement of this algorithm is going to do in the work. The new algorithm is named as CDSMA broadcast algorithm.

In the normal cases while running a program, an another program cannot compile at the same time. Here going to parallel creation, compilation, and running is maintain. While performing likewise the load will change to higher, so maintaining this load is the objective. Also in the network services the replication is a common phenomenon, this work is trying to avoid the duplication and broadcasting the services. Adaptive and fault tolerant Load balancing schemes presented in this work are promising; still there are significant scope for future research. This research work improves the understanding of distributed computing environments and advances the state-of the art through its contributions. Its investigation revealed areas in distributed system where much work remains to be done.

B. Content Distribution Network

The Internet has become much more than just a communication infrastructure, up to the point that some authors have defined the global network as a platform for business and society‖. This change poses new requirements on the Internet itself, which was not specifically designed to perform content distribution. In order to tackle this issue, Content Distribution Networks (CDNs), such as AKAMAI, have become a vital layer in the architecture of any content provider as they make it possible to distribute content in today‘s IP Internet in an efficient way[20]. The driving principle of CDNs is that content requests are not directly served by the origin server owned by the content provider, but they are instead mediated by the CDN infrastructure. The CDN operator owns a given number of surrogate servers, scattered all over the world, which thus perform content caching and replication, improving the Quality of Service (QoS) of the consumers by serving the requests of the clients in the neighbourhood [18] . One of the main features of CDNs is that they do not change the current key network protocols, but they rather offer countermeasures to address the peculiar characteristics of the Internet infrastructure that limit its effectiveness when performing content distribution.

Content delivery services uses CDN across the world to deliver copies of website content.CDN is made of following elements:

- Storage and Delivery nodes
- Central Controller
- Central Manager
- Request Manager
- Central Datastore

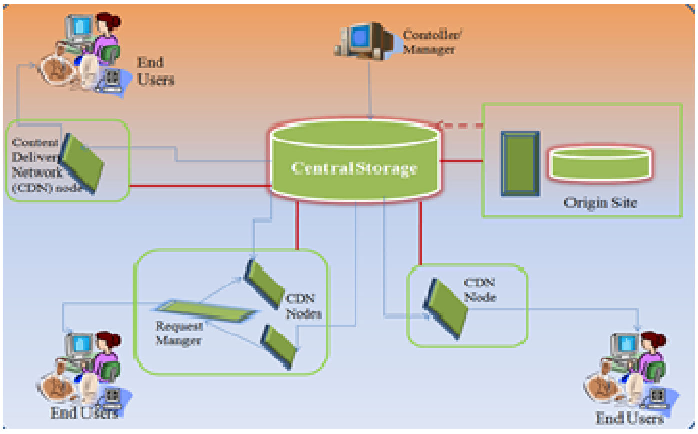

Fig: 1.2 CDN Infrastructure

II. RELATED WORKS

Out of the vast literature related to placement problems we focus on the distributed solutions and identify two main approaches: those adopting a facility location approach [10] and those drawing on a knapsack problem formulation [20]. The latter is used when constrained storage is considered. More relevant to our service deployment scenario is the former approach that attracts both theoretical and practical interest. The theoretical thread relates to the approximability of the facility location problem by distributed approaches. Algorithms are typically executed over a complete bipartite graph where m facilities communicate with n clients in synchronous send-receive rounds. With social networking sites providing increasingly richer context, user-centric service creation is expected to follow a similar growth with User-Generated Content. Service deployment is a key problem in communication networks as it determines how efficiently the user service demands are supported. This problem has been traditionally approached through the formulation and resolution of large optimization problems requiring continuous recalculation of the solution in case of network changes. The effectiveness of service provisioning in large-scale networks is highly dependent on the number and location of service facilities deployed at various hosts. Centralized approaches assume the existence of a super-entity with global topological and service demand information that has the resources to determine and realize the placements. This implies a role hierarchy across network nodes, which in many cases is missing. Finally, it is neither practical nor affordable to centrally compute a new solution each time user demand shifts or network topology changes alter the optimal service location. The disadvantage of content distribution network is that it costs money since organizations must rely on support from third-party CDN vendors, and it adds a bit of complexity for deployment procedures.

A: Problem Definition

A major change in the emerging networking landscape concerns the role of end-user, who is no more only content consumer but also generator. A major challenge for this emerging paradigm is how to make these exploding in numbers, yet individually of vanishing demand, services available in a cost-effective manner; central to this task is the determination of the optimal service host location.

The optimal placement of up to k service instances is typically treated as a k-median problem. Input to the problem is the network topology and service demand distribution across the network users. The focus instead is on the 1-median problem variant that seeks to minimize the access cost of a single service replica since it matches better the forthcoming User Generated Service paradigm [1]. This will enable the generation of service facilities in various network locations from a versatile set of amateur user service providers. The huge majority will be lightweight services requiring minimum storage resources and addressing relatively few users in the proximity of the user service provider, either geographical or social, so that the replication across the network would not be justified. Here this problem is formulated as a facility location problem and devise a distributed and highly scalable heuristic to solve it. Key to this approach is the introduction of a novel centrality metric. Wherever the service is generated, this metric helps to

- Identify a small sub graph of candidate service host nodes with high service demand concentration capacity.
- Project on them a reduced yet accurate view of the global demand distribution.
- Pave the service migration path towards the location that minimizes its aggregate access cost over the whole network.

In each service migration step, the metric serves two purposes. Firstly, it identifies those nodes that contribute most to the aggregate service access cost and pull the service strongly in their direction; namely, nodes holding a central position within the network topology and/or routing large service demand amounts. Secondly, it correctly projects the attraction forces these nodes exert to the service upon the current service location and facilitates a migration step towards the optimal location.

B: Objective

The objective of the project is to develop a scalable heuristic for placing service facilities in distributed self-organizing environments. The approach is distributed and scalable: it solves locally and then moves the services towards their optimal location traversing an access cost-decreasing path. Moreover, it is innovative in the way it selects the nodes for the local 1 median problem. The service provision points are brought close to the demand in order to minimize communication resource consumption and enhance the Quality of Service (QoS) of the provided service. Autonomic approach facilitate the management of distributed system [5]. One of the main contribution is an autonomic replica placement module that places data copies only on servers close to clients that actually need them. The main objectives are:

- Copies are distributed so that no server become overloaded.
- Fault tolerant by using replicated servers.
- Load balancing by automatic content management.
- End user never receives outdated or stale content.
- Scalability, reliability and bandwidth saving.
- Increased speed by dispatching the requests to the closest site.

III. METHODOLOGY

In this section, the major implementation techniques of the system are described.

A: Database Design

Database design is the most important design of all design activities. Microsoft SQL 2005 Server Edition is selected to design the database system because it is compatible with Microsoft ASP.NET, the programming tool used to create the system.

B: Table Names and Functions

Here the database consists of 5 tables. The name of each table and a short description of each table‘s function is described in the following section.

1) Node- It keeps nodeid, name, username, password, ip address, user type, deleted nodes, node range, X location, Y location.

2) Nodehistory- nodehistoryid, message, sourcenodeid, destination node, deleted node, date & time.

3) Subnodehistory- subnodehistoryid, nodehistoryid, nodeid, timestamp, deleted nodes.

4) Request– Stores requesteid,serviceid,service name, nodeid, ip address.

5) Service– serviceid, nodeid, service name, service destination, ip address.

C: Table Design

Data storage is considered to be the most important part of a system. The general objectives in the design of data storage are listed below:

- Data Availability
- Efficiency of data storage
- Purposeful information retrieval
- Efficiency of data updates and retrieval
- Data integrity
- Avoid data duplication

Table:3.1 Request

Field Name	Data Type	Description
reqid	Int	Primary Key
sid	Int	Source id
Nodeid	Int	Node id
sname	nvarchar(50)	Name of source node
ipaddress	nvarchar(50)	IP address of node

Table:3.2 Service

Field Name	Data Type	Description
serviceid	Int	Primary Key
Service description	nvarchar(50)	Description of service
nodeid	Int	Node id
Service name	nvarchar(50)	Name of service
ipaddress	nvarchar(50)	IP address of node

Table:3.3 Node Details

Field Name	Data Type	Description
NodeID	Int	Primary Key
name	nvarchar(50)	Name of user
Username	nvarchar(20)	Username of user
Password	nvarchar(20)	Password of user
IPaddress	nvarchar(50)	IP address of node
Usertype	nvarchar(10)	Admin or user
Isdeleted	Bit	Node deleted or

		not
Node range	Int	Range of node
Node X	Int	X value
Node Y	Int	Y value

Table:3.4 Nodehistoryid

Field Name	Data Type	Description
Nodehistoryid	Int	Primary key
Message	nvarchar(MAX)	User id
Sourcenodeid	Int	Source node id
Destinationnodeid	Int	Destination node id
Isdeleted	Int	Node deleted or not
Datetime	Bit	Service or message transaction time

Table 3.5 Subnodehistoryid

Field Name	Data Type	Description
Subnodehistoryid	Int	Primary Key
Nodehistoryid	Int	Node history id
Nodeid	Int	Node id
Isdeleted	Bit	Node deleted or not
Timestamp	nvarchar(100)	Time in milliseconds

IV IMPLEMENTATION OF THE SYSTEM

This chapter presents the project implementation phase. The first section 4.1 describes the modules. In a network, a node is a connection point that is attached to a network, and is capable of sending, receiving or forwarding information over a communication channel. In general, a node has programmed or engineered capability, to recognize and process or forward transmissions to other nodes. So administrator firstly add new nodes in the network using relevant information. The information includes name, username, password, IP address and position of node. The position is described using the X,Y region and the range of that particular node.

V. PERFORMANCE ANALYSIS

Performance analysis involves gathering formal and informal data to help users and service providers define and achieve their goals. Performance analysis uncovers several perspectives on a problem or opportunity, determining any and all drivers towards or barriers to successful performance, and proposing a solution system based on what is discovered. The accomplishment of a given task measured against preset known standards of accuracy, completeness, cost, and speed. In contrast, performance is deemed to be the fulfillment of an obligation, in a manner that releases the performer from all liabilities under the contract. The definition for performance analysis given is: A specific, performance based needs assessment technique that precedes any design or development activities by analyzing the performance problems of a work organization.

System Strength :

1. Autonomic Approach helps in managing the distributed system easily and efficiently.
2. Reduce communication and processing overhead significantly.
3. Network benefit from inherent replication of data which protects against node failures and the effects churn.
4. Applicable to a wide range of system, ranging from sensor networks and fixed office meshes.
5. Iterative solution directs the service migration towards cost effective locations.

System Limitations :

1. Nodes send messages only to neighbouring nodes.
2. High speed internet access is required.

VI CONCLUSION

A scalable decentralized heuristic algorithm that iteratively moves services from their generation location to the network location that minimizes their access cost. To tackle the problem as an instance of the facility location problems precisely, employ the 1-median formulation as more suitable for the user-centric service paradigm. Contrary to centralized approaches, where a super-entity with global information about network topology and service demand solves the problem in a single iteration, let it migrate towards its optimal location in a few hops. In each iteration, a small-scale 1-median problem is solved so that the computational load is spread along the migration path nodes. The service migration path is determined with the help of a node centrality index (metric) called weighted Conditional

Betweenness Centrality (wCBC). In view of the proliferation of user-centric service instances across the Internet and the ever increasing in-network storage capabilities, going to develop a scalable and effective heuristic approach to deal with the complexity and limitations of their distributed placement. In this phase will develop a trust system based on processing the payment reports to maintain a trust value for each node. The nodes that relay messages more successfully will have higher trust values, such as the low-mobility and the large-hardware-resources nodes. Based on these trust values, will propose a trust-based routing protocol to route messages through the highly trusted nodes (which performed packet relay more successfully in the past) to minimize the probability of dropping the messages, and thus improve the network performance in terms of throughput and packet delivery ratio. However, the trust system should be secure against singular and collusive attacks, and the routing protocol should make smart decisions regarding node selection with low overhead. Overcomes the scalability limitations of classic centralized approaches by re-optimizing the locations and the number of facilities through local optimizations which are refined in several iterations. This work develop a scalable and distributed approach that answers for key question through an iterative reoptimization of the location and the number of facilities within network neighborhoods. Here proposing an innovative approach to migrate, add, or remove servers within limited-scope network neighborhoods by utilizing only local information about the topology and demand. This work shows that even with limited information about the network topology and demand, within one or two hops, the distributed approach achieves performance, under various synthetic and real Internet topologies and workloads, which is comparable to that of optimal, centralized approaches requiring full topology and demand information. This work also shows that it is responsive to volatile demand. Re-optimizations are based on exact topological and demand information from nodes in the immediate vicinity of a facility, assisted by concise approximate representation of demand information from neighboring nodes in the wider domain of the facility. Using extensive synthetic and trace-driven simulations, demonstrates that the distributed approach is able to scale well by utilizing limited local information without making serious performance sacrifices as compared to centralized optimal solutions. Also demonstrate that the distributed approach yields a high performance under demand and imperfect redirection. This approach leverages recent advances in virtualization technology and the flexibility of billing models, such as pay-as-you-go, towards a fully automated Internet-scale service deployment.

REFERENCES

[1] Panagiotis Pantazopoulos, Merkouris Karaliopoulos, Ioannis Stavrakakis Distributed Placement of Autonomic Internet Services‖ IEEE transactions on parallel and distributed systems, Vol. 25, No. 7, July 2014

[2] C.A. La, P. Michiardi, C. Casetti, C. Chiasserini, and M. Fiore,‗Content Replication in Mobile Networks, ‛‛IEEE J. Sel. Areas Commun., vol. 30, no. 9, pp. 1762-1770, Oct. 2012.

[3] G. Smaragdakis, N. Laoutaris, K. Oikonomou, I. Stavrakakis, and A. Bestavros, Distributed Server Migration for Scalable Internet Service Deployment,‛‛ IEEE/ACM Trans. Netw., doi: 10.1109/TNET.2013.2270440.

[4] P. Pantazopoulos, I. Stavrakakis, A. Passarella, and M. Conti,Efficient Social-Aware Content Placement for Opportunistic Networks,‛‛ in IFIP/IEEE WONS, Kranjska Gora, Slovenia, Feb. 3-5,2010, pp. 17-24.

[5] K. Oikonomou and I. Stavrakakis, Scalable Service Migration in Autonomic Network Environments,‛‛ IEEE J. Sel. Areas Commun.,vol. 28, no. 1, pp. 84-94, Jan. 2010.

[6] P. Pantazopoulos, M. Karaliopoulos, and I. Stavrakakis, Centrality-Driven Scalable ServiceMigration,‛‛ in Proc. 23rd ITC, San Francisco,CA, USA, 2011, pp. 127-134.

[7] K. Oikonomou, I. Stavrakakis, and A. Xydias, Scalable Service Migration in General Topologies,‖ The Second International IEEE WoWMoM Workshop on Autonomic and Opportunistic Communications (AOC 2008), Newport Beach, California, 23 June, 2008.

[8] D. Trossen, M. Sarela, and K. Sollins, Arguments for an Information-Centric Internetworking Architecture,‛‛SIGCOMM Comput. Commun. Rev., vol. 40, no. 2, pp. 26-33, Apr. 2010 *Translated J. Magn. Japan*, vol. 2, pp. 740-741, August 1987 [*Digest 9th Annual Conf. Magnetics Japan*, p. 301, 1982].

[9] M.E.J. Newman, The Structure and Function of Complex Networks,‛‛ SIAM Rev., vol. 45, no. 2, pp. 167-256, 2003.

[10] K. Oikonomou, and I. Stavrakakis, Scalable Service Migration: The Tree Topology Case,‖ The Fifth Annual Mediterranean Ad Hoc Networking Workshop (Med-Hoc Net 2006), Lipari, Italy, June 14-17, 2006.

[11] V. Jacobson, D.K. Smetters, J.D. Thornton,M. Plass,N.H. Briggs, and R.L. Braynard Networking Named Content,‛‛ in Proc. 5thACM CoNEXT, Rome, Italy, Dec. 2009, pp. 1-12.

[12] T. Sproull and R. Chamberlain, Distributed Algorithms for the Placement of Network Services,‛‛ in Proc. ICOMP, LasVegas, NV, USA, July 2010, pp. 1-8.

[13] G. Wittenburg and J. Schiller, A survey of current directions in service placement in mobile ad-hoc networks,‖ in IEEE PERCOM '08, Hong Kong, March 17-21 2008, pp. 548–553.

[14] F. Sailhan, and V. Issarny, Scalable Service Discovery for MANET,‖ 3rd IEEE Intl. Conf. on Pervasive Computing and Communications, Kauai, USA, Mar. 2005.

[15] C. Ragusa, A. Liotta, G. Pavlou, An Adaptive Clustering Approach for the Management of Dynamic Systems,‖ IEEE Journal of Selected Areas in Communications (JSAC), special issue on Autonomic Communication Systems, Vol. 23, No. 12, pp. 2223–2235, IEEE, December 2005.

[16] C. Boldrini, M. Conti, and A. Passarella, Content Place: social aware data dissemination in opportunistic networks,‖ in Proceedings of the 11th international symposium on Modeling, analysis and simulation of wireless and mobile systems. ACM New York, NY, USA, 2008, pp. 203–210.

[17] G. Pallis and A. Vakali, Insight and Perspectives for Content Delivery Networks, Communications of the ACM, vol. 49, no. 1, pp. 101–106, 2006.

[18] J. Kangasharju, J. Roberts, and K. W. Ross, ―Object Replication Strategies in Content Distribution Networks,‖ Computer Communications, vol. 25, no. 4, pp. 376–383, 2002.

[19] S. Borst, V. Gupta, and A. Walid, ―Distributed Caching Algorithms for Content Distribution Networks,‖ in Proc. of the 29th IEEE Conference on Information Communications (INFOCOM'10), pp. 1478-1486, San Diego, California, USA, 2012.

[20] S. Pandit and S. Pemmaraju, Return of the Primal-Dual: Distributed Metric Facility Location, in Proc. ACM PODC, 2009, pp. 180-189 .