Using Agent Base Models to Optimize Large Scale Network for a Large System Inventories

Ramez Ahmed Shamseldin, Senior Design Engineer, Verizon Business and Shannon R. Bowling, Assistant Professor, Old Dominion University
shamsrdin@hotmail.com and sbowling@odu.edu

Abstract. The aim of this paper is to use Agent Base Models (ABM) to optimize large scale network handling capabilities for large system inventories and to implement strategies for the purpose of reducing capital expenses. The models used in this paper either use computational algorithms or procedure implementations developed by Matlab to simulate agent based models in a principal programming language and mathematical theory using clusters, these clusters work as a high performance computational performance to run the program in parallel computational. In both cases, a model is defined as compilation of a set of structures and processes assumed to underlie the behavior of a network system.

Introduction

The nature of digital networks, as described in [1] is comprised of non-variable bandwidth channels that transfer data. Furthermore, the growth of demand for transmitted visual data has been abruptly increased to satisfy customer needs, which resulted in the development of multi-video compression standards such as MPEG-2 [2], H.263 [3] and MPEG-4 [4].

In today's world, there are two kinds of video transmissions that have been established, one of them consists of full transmission of stored packets of video from a server to the customer's premises before playback begins; the other is a concurrent transmission which is under a certain restriction of quality of service (QoS) and serves as a real-time application.

The nodes in a network represent a video between customers who requested the service to watch certain movies. The selected video file is downloaded to the customer's computer site according to the system requested. It is also added to an inventory which can be allocated to several other sites in the future. System redundancy has been taken into account with regard to system needs for any overly excessive demands.

Agent-based models (ABM) are used in simulating social life, not only to understand environmental change and human roles, but to be attractive to many practitioners from a variety of subject areas. Human changes can happen through space and on different time scales. Many vital opinions of ABM and simulations are that numerous phenomena, even though system is complexity, dynamically or both combined dynamically complexity, can be described as autonomous agents that are relatively simple and follow certain rules for interaction.

Computer models are used for interesting research practices and testing theories within certain discipline structures. The progression fundamentals of a real-world structure are difficult to be observed and collecting data as well as controlling it under certain conditions is impossible. Assumptions based on theories for these structures can be implemented in a computer model that can perform and compare to this practical data.

Mainly, models used in this conference either use computational algorithms or procedure implementations developed by agent based models in any principal programming language or mathematical theory to underlie the behavior of a network system.

Literature Review

Storage systems occur in a variety of contexts, including manufacturing, warehousing, and the service sector. Most storage systems do not deal with dynamic complexity because they are static and are usually in the form of physical warehouses. Storage systems dealing with materials can be either continuous or discrete storage. There are three major factors affecting storage systems: depending on size of storage, storage methods and layout of the storage system.

Storage throughput has been used as a measurement to describe the number of storage that can be retrieved per time period storage/retrieval (S/R). From there the size of the storage system is powered by throughput and cost parameters of transferring materials. Storage system mission controls input/output (I/O) functionality that can be determined by storage requirements is distributed centrally over time.

The storage method contains specification of unit load, S/R and storage equipments; these methods can be handled by machine or by humans and can be automatically launched by automatic guided vehicles.

The storage system layout, by using three dimensions - height, length and width - can identify the location of storage items. In this paper, the system layout control other storage parameters throughput and storage method will be discussed and five different storage system layouts will help to understand what the differences are between these physically traditional storage types and virtual layouts for our case scenario.

Dedicated Storage Location

Every SKU (Stock Keeping Unit) is related to items in a warehouse and has a unique storage dedicated to it is location. Dedicated storage is characterized by the assignment of fixed storage locations for the items stored in the warehouse [5]. For items to be allocated as measured unites used in the warehouse, they are assigned the cube per order index (COI).

In such cases, the more popular items have to be near the I/O point in the warehouse for reduction in travel time and travel distance according to the S/R. As an example, active items have to be placed in the most convenient and accessible place; this minimizes cost effectiveness and gets item to I/O points.

Randomized Storage Location

The items in the storage warehouse are stored randomly in any available storage location. For an example, when the inbound load arrives for drop off, the item in the closest available slot is designated. This is known as first-in and first-out.

This is common in the case of randomized storage results when less storage space occurs. Having small sized parts stored in a space designed for large size parts wastes storage space, and for the same scenario, storing large parts to fit randomly can be impossible, so adjustable shelves may need to be used.

In randomized storage it is assumed each item of a certain product is equally likely to be recovered when multiple storage locations exist for the product, and the recovery operation is achieved. In the case where the warehouse is pretty full, the travel distances are significantly of the same "equal likelihood" [6].

Class-Based Dedicated Storage

Class based storage is defined as a grouped of SKUs in one class. These classes are assigned to a dedicated storage spot, at the same time; these SKUs within an individual class are stored randomly and in a logical sequence.

The products are distributed according to their demand rates, among the number of classes and have a reserved a region within the storage area for each class. Accordingly, an incoming load is stored at an arbitrary available location for the same class.

We must look to the randomized storage location as a single class case of class-based storage policy where dedicated storage is counted as one class for each item.

In addition, the dedicated storage policy attempts to reduce travel times for S/V (storage/retrieval) by sorting the highest demand to the I/O point as well for class-based storage and calculate the product demand by COI [7].

Continuous Warehouse Storage

Increasing demand for continuous recording of hundreds of millions data daily, a necessary storage media should have the capability to handle data volumes and data flow rates.

These types of data could be called detailed records (CDRs) - which are commonly used by the telecommunication industry - at an individual basis for each customer. Software applications have been used to pose several challenges related to data volumes and data flow rates to data warehouses and to online analytical processing (OLAP).

Shared Storage

Shared storage is widely used within the computer networking industry and addresses the needs of corporate computing environments for storage systems that propose scalability, availability and flexibility.

Storage systems are known as storage computer systems (hosts) and are connected to multiple individual hosts while using the shared storage by these hosts and are managed independently and historically viewed (host-attached storage).

Shared storage systems enabled by networking technology can provide high bandwidth. In turn, it offers several benefits for today's businesses, for example by improving quality of service (QoS) and increasing operational efficiency.

Moreover, as growing needs for shares (files, data, etc.) become necessary, it is necessary to prevent buying mainframe computer complexes and computer clusters where a modest number of cooperating computer systems share a common set of storage devices.

As computing environments have grown in industry, computer storage systems have grown in storage size and in number as the cost of equipment becomes more reasonable in order to increase the computing environments.

The main disadvantage is that the known computer storage systems processors have failed and replacement parts can be required to get the system back for full operation which wastes time and is followed by a typically propagation delay of the restoration of the data.

Virtual Warehousing

As a physical location is not necessary to locate specific data content, data can be located within many virtual storage hosts. If a customer is looking for specific data to download, random locations can be used without specification and taking into account how many locations have been used.

The storage locations mentioned above, such as dedicated, randomized and class-based storage, can be used to benefit virtual storage warehouses with priority, size and rates of transferring data. On other hand, desired locations for data can be easily tracked and assigned to scale from the highest to the least high activities according to their demand.

At the same time, randomized storage results in a reduction in space and will be significant with regard to data travel time much less so than those traveling from a dedicated storage area.

Also, using other large-scale (shared) storage techniques, such as Internet, without specifying certain hosts is not problematic because data is already restored within different hosts. Finally continuous warehouse storage techniques use network capability and add more data to different new hosts entering to networks as well using existing hosts.

Network and Complexity

The internet-wide system is viewed as a large scale structure with an underlying physical connectivity that deploys real experimental studies to evaluate system architectures, however this is not possible. Instead, a randomly generated network connectivity structure is used and has been accepted at the beginning as a node degree distribution technique. A generator - also known as a software based solution - is used to generate network nodes which represent network autonomous systems (AS), original power laws and connectivity to the Internet.

The studies were then used to randomly generate networked topologies and provide precise analyses that show network modeling include [8]:

- Regular topology, such as liner, rings, trees, and stars;
- Well recognized topology, such as ARPANET or the NSFNET backbone;
- Arbitrarily generated topologies.

According to [9] when any two nodes have a relation, one link will be added with a probability depending on the distance between them given by:

$$p(u,v) = \beta \exp\frac{-d(u,v)}{L\alpha}$$

Where $d(u,v)$ is a distance from u to v; L is the maximum distance between two nods, $\alpha > 0$ and $\beta \leq 1$. However, this method does not obligate a large scale structure.

Albert et al [10] describe a system's components for a network as a complex system because of its functionality and attribute is largely to redundancy node connections. A large scale network consists of a complex communication network (CCN) along with groups of telecommunication carriers and ISPs (Internet Service Providers). It is almost impossible to analyze the infrastructure but this can be done within the limited boundaries of individual networks [11].

The redundancy of network connectivity, in other words scale-free network connections, represents an unpredicted degree of robustness for each kind of system, such as the internet, social networks or cellular (metabolic) networks. Network nodes break when faced with an extremely broken down communication rate.

Agent Based Models (ABM)

Agent based model methodology has been applied to several studies, for example, social dynamics and communication and cooperation under ecological risk [12]; complexity in artificial life applications [13]; common dilemmas for ecological economics [14]; language evaluation [15]; armed forces contradictions [16]; and human social interaction interpolating with regeneration management [17].

Huigen [18] anticipates an ABM structure, called MameLuke, which will study human environment interaction. For like structures agents are categorized according to user definitions and determinations from the objective's study, meaning that individual agent sets can fit into multiple non-divergence categories. Potential option paths (POPs) are rule-based implementation through decision making, which depends on the agent's category.

ABM was significantly used in a spatial interest group within computational mathematical organization theory (CMOT). Today, on the other hand, simulations using ABM have expanded further than the original boundaries of use and have linked up with groups of people and cover work in a variety of different disciplines such as economics, biology, sociology, artificial intelligence, physics, computer science, archaeology and anthropology.

In the last few years, growth of ABM has been significant realized especially after releasing more helpful software toolkits. This was enough to attract many practitioners from different fields to simulate numerous subject areas. Some of the better known toolkits are Swarm, Repast, AnyLogic, MASON, Ascape and NetLogo.

Gilbert et al. [19] express an example of using ABM in the artificial intelligent field for developing cellular automata. At the time, Swarm, introduced in 1996, was the only agent based modeling simulation tool available [20].

The primary characteristic of an agent is the potential to make decisions on individual bases. On the other hand, agents, in a true case, are discrete events handled individually with a set of attributes and policies that influence its actions and decision making capability. In addition, an agent may have supplementary policies that modify its policies or attributes. An agent can be purposely independent in its atmosphere and in its interactions with other agents as well itself if not over an imperfect scope of posts. An agent has objectives to accomplish (not optimize) as goal bound within its actions. Furthermore, an agent is flexible and has the ability to learn and adapt its performance over time based on ongoing skills, in other word, some form of memory.

Network Description and Functionality

The network that will be handled in this dissertation has a total of 250 nodes which represent the total number of customers carried by this network. These nodes are virtually connected by the internet and each address is recorded and knows the location of each customer. Each node in the network follows these assumptions:

- Each node is connected to the network and works online all year long with no bad connections.

- All nodes share the same bandwidth speed (uploading or downloading), and uploading bandwidth is half the speed of downloading bandwidth.
- The bandwidth speeds that will be used are limited to 128kb, 512kb, 1000kb, 2000kb and 5000kb per second.
- Uploading bandwidth and downloading bandwidth are two different streams and separated at each node.
- All nodes are spread all over the internet and connect to a separate network that can be located physically anywhere with no adverse affects on location or distance.
- All nodes can download simultaneously from the server with no affect on delay or connectivity.
- Each node can be used as virtual storage and upload any necessary file needed by another node upon request and can only to do this one node at a time.
- Each node can be downloaded from the server or from another node according to these guidelines:
  - Each node can download, at the maximum, from two locations and can be the server, the server and a single node or two nodes simultaneously.
  - Only one file can be downloaded at a time.
  - If the file exists in two virtual locations in network, the server will be exempt.
- All 250 nodes will be divided into five categories. Each category includes 50 nodes selected randomly. These categories are Actions, Crime, Comedy, Drama and Romance.
- Each node has an internal storage device and is selected randomly from a set of sizes: 50, 100, 150, 200, 250, 300, 350, 400, 450 or 500GB.

Each node will be studied throughout the year and is equivalent to 8760 hours download time and is evaluated for how many files has been selected and downloaded as these files are selected according to each node's preferences.

Selected files will be chosen randomly and according to each node's preferences. These files have the following characteristics and assumptions:

- The network will handle files of different sizes having different time durations, and is limited to 10,000 files and all files can be downloaded from the server.
- The 10,000 files will be divided to five categories. Each category includes 2,000 files ranked from the highest priority to the lowest according to the power law degree distribution $P(k)\sim{k}^{y}$ with an exponent y range between 2 and 3. These categories are Actions, Crime, Comedy, Drama and Romance.
- The files can be downloaded from server, two nodes or a node and server at the same time by splitting the file's size to two batches - each batch contains half of the file.
- The file's batches will be downloaded either simultaneously by dividing the downstream bandwidth in half or downloading individually as the second batch will not start till the first one is completely downloaded.
- These files do not have expiration time but rather are replaceable inside the network's virtual storage which is located at the nodes. If the node's storage device reaches 75\%, the files will be deleted according to the file's priority from low to high with the exception of the server. In this case, it will be remain stored as a reference for future requests.

All nodes will be able to download any files from a server at any time with no delay. All files are ranked according to its priority and stored in the server in the five different categories. Any node can search for any file across the entire network and download it in another in order to overcome network load and reach an optimum for the network.

Networks and Their Dynamic Complexity Purpose

Networks act as huge virtual storage warehouses that are dynamically changed over a period of time. The address of the nodes will be constant, but a file's location will be changed from node to other node with time and determined priority.

The duration of this study is equivalent to 8760 hours over an entire year. This study will follow several procedures to highlight and identify the purposes of this research. In addition, it will also simulate the generated data not only to show the output results but to also understand how the network works with layers of dynamic changes as the files flow across the network.

A network's complexity is represented by nodes and a server that are interactively and laterally ordering files from the server, other neighbor nodes or both at the same time. Also, the simulation of each of these nodes requires further study. The criteria and procedure will follow:

- Gathering data of inter-arrival time which was observed for each node during the 8760 hours.
- Gathering information about what type of file category customers were interested in as well as how many files per node were accessed.
- Gathering information from where the files were downloaded by each node.
- Calculating the arrival time by each node.

- Calculating the inter-arrival download time for each file by each node.
- Calculating the arrival download time for each node.
- Setting up groups of files preferences for each customer at the time of ordering and calculating the watched time at an individual basis for each file with download time to calculate the penalty time that occurred, each group will include random numbers of files between one and six files sets at a time which represent a set ordering one time group.

A network is a set of nodes connected virtually by an intranet at all times. These nodes share their contents at the same time and evaluate best practices to reach an optimal scenario as an ideal network which can change periodically over space and time.

Optimizing a network has a set of fundamentals that are required to help to simulate this kind of network and evaluate the results. In order to do that, the next section describes a set of identifying metrics that have to be fully gathered and understood in order to direct these types of networks to the second stage, which is ready to be integrated as an optimal large scale dynamic complexity network.

Conclusions and Results

As noted before this data will be analyzed in different measures that are suitable for the type of experimental run and described in this paper. There is numerous data that will be impracticable to present here, but important ones will be presented as a key to show the differences and similarities and as the aim of this paper to show how networks can be dynamic and change complexity in time and space which then can be optimized based on specified performance measures.

Time Series was conducted on a Base Line, as shown on figures below, for different bandwidth loaded on the server with no nodes sharing the load with the server: the hard drive was not a study factor in this experiment and each node using the same value of bandwidth download speed from the server at the same time of analyzes.

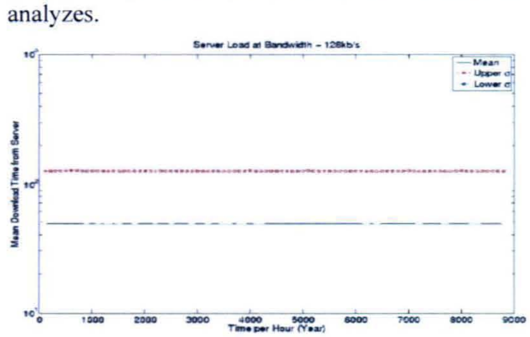

Figure 1: Server load for Base Line case scenario at download bandwidth of 128kb/s.

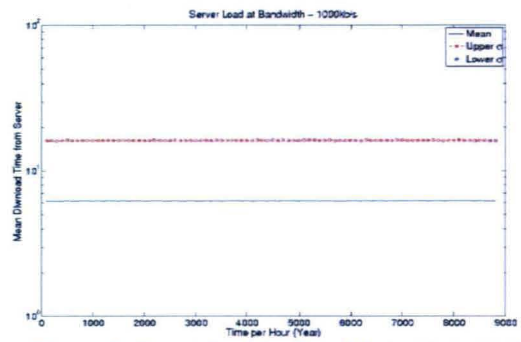

Figure 2: Server load for Base Line case scenario at download bandwidth of 1000kb/s.

Time Series was conducted on a H1, as shown on figures below, for different hard drive sizes and different download bandwidth speed loaded on the server with other nodes sharing the load with the server; the hard drive had different effects on the study in this experiment and each node using the same value of bandwidth download speed from the server and same value of hard drive size at the same time of analysis.

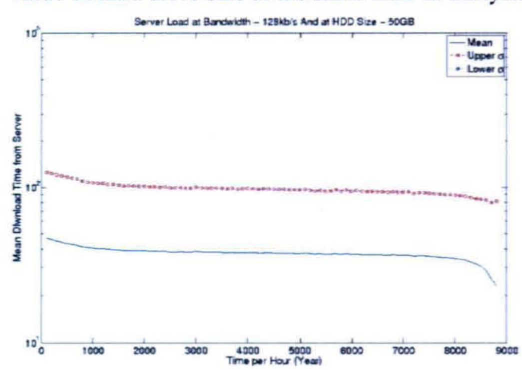

Figure 3: Server Load for H1 case scenario at hard drive size of 50GB and download bandwidth of 128kb/s.

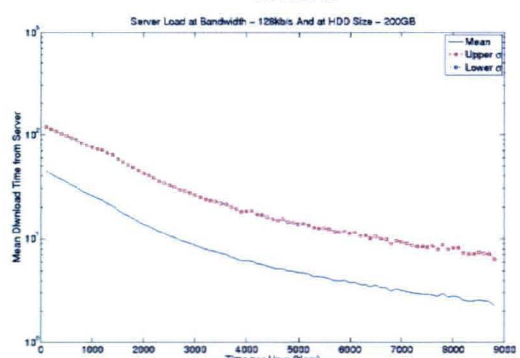

Figure 4: Server Load for H1 case scenario at hard drive size of 50GB and download bandwidth of 1000kb/s.

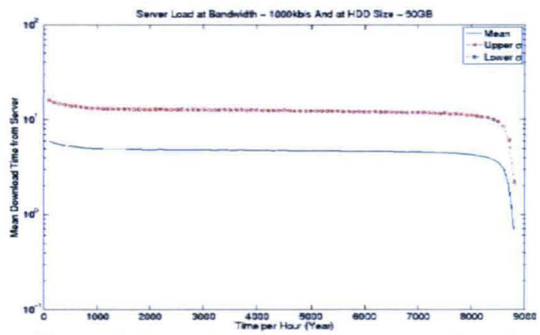

Figure 5: Server Load for H1 case scenario at hard drive size of 50GB and download bandwidth of 1000kb/s.

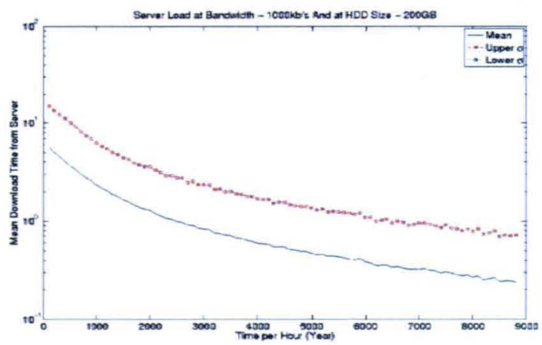

Figure 6: Server Load for H1 case scenario at hard drive size of 200GB and download bandwidth of 1000kb/s.

References

1. Militzer, M. and Suchomski, M. and Meyer-Wegener, K. (2009) "Improved p-domain rate control and perceived quality optimizations for MPEG-4 real-time video applications," *Proceedings of the eleventh ACM international conference on Multimedia, ACM Press, New York, NY, USA*, pp. 402-411.

2. Boncelet, C. (2000) "Handbook of Image and Video Processing," Image noise models, Academic Press: San Diego, CA, USA, Alan C. Bovik.

3. Rijkse, K. (1996) "H.263: video coding for low-bit-rate communication," *Communications Magazine, IEEE*, vol. 34, pp. 42-45.

4. Militzer, M. and Suchomski, M. and Meyer-Wegener, K. (2003) "Improved p-domain rate control and perceived quality optimizations for MPEG-4 real-time video applications," *Proceedings of the eleventh ACM international conference on Multimedia, ACM Press, New York, NY, USA*, pp. 402-411.

5. Malmborg, C.J. and Krishnakumar, B. (1989) "Optimal Storage Assignment Policies for Multiaddress Warehousing," *IEEE Transactions of Systems, MAN, AND CYBERNETICS*, vol. 19, no. 1.

6. Francis, R.L. and McGinnis, L.F. and White, J. A. (1992) "Facility Layout and Location: An Analytical Approach," Englewood Cliffs: NJ.

7. Hesket, T.J.L. (1963) "Cube-per-Order Index-D a Key to Warehouse Stock Location. Transportation and Distribution Management," vol. 3, pp. 27-31.

8. Zegura, E.W. and Calvert, K.L. and Donahoo, M.J. (1997) "A Quantitative Comparison of Graph-Based Models for {Internet} Topology," *IEEE\slash ACM Transactions on Networking*, vol. 5, no. 6, pp. 770-782.

9. Bernard M. Waxman (1991) "Routing of multipoint connections," Broadband switching: architectures, protocols, design, and analysis, *IEEE Computer Society Press*, ISBN = "0-8186-8926-9", Los Alamitos, CA, USA, pp. 347-352.

10. Albert, R. and Jeong, H. and Barabasi, A.L. (2000) "Error and attack tolerance of complex networks," *Nature*, vol. 409, pp. 371-384.

11. Claffy, K. and Monk, T.E. and McROBB, D. (1999) "Internet Tomography," *Nature*: Http://www.nature.com/nature/webmatters/tomog/tomog.html.

12. Andras, P. and Roberts, G. and Lazarus, J. (2003) "Environmental Risk, Cooperation, and Communication Complexity, *Adaptive Agents and Multi-Agent Systems*, vol. 2636, pp. 49-65.

13. Menczer, F. and Belew, R. K. (1996) "From Complex Environments to Complex Behaviors," *Adaptive Behavior*, vol. 4, pp. 3-4.

14. Jager, W. and Janssen, M. and Vries, H. D. and Greef, J. D. and Vlek, C. (2000) "Behaviour in Commons Dilemmas: Homo Economicus and Homo Psychologicus in an Ecologicaleconomic Model," *Ecological Economics*, vol. 35, pp. 357-379.

15. Bartlett, M. and Kazakov, D. (2004) "The Role of Environmental Structure in Multiagent Simulations of Language Evolution," *In Proceedings of the Fourth Symposium on Adaptive Agents and Multi-Agent Systems (AAMAS-4)*: AISB convention.

16. Cioffi-Revilla, C. and Gotts, N. M. (2003) "Comparative Analysis of Agent-Based Social Simulations: Geosim and FEARLUS models," *Journal of Artificial Societies and Social Simulation*, vol. 6, no. 4.

17. Deadman, P. and Gimblett, R. H. (1994) "The Role of Goal-Oriented Autonomous Agents in Modeling People-Environment Interactions in Forest Recreation," *Mathematical and Computer Modeling*, vol. 20, no. 8.

18. Huigen, M. G. A. (2004) "First Principles of the MameLuke Multi-Actor Modelling Framework for Land-Use Change, Illustrated with a Philippine Case Study," *Journal of Environmental Management*, vol. 72, pp. 5-12.

19. Gilbert, N. and Troitzsch, K. (1999) "Simulation for the Social Scientist," *Open University Press*: second edition, 2005.

20. Nelson Minar and Rogert Burkhart and Chris Langton and Manor Askenazi (1996) "The Swarm Simulation System: A Toolkit for Building Multi-Agent Simulations," Santa Fe Institute: Working Papers, no. 96-06-042, Http://ideas.repec.org/p/wop/safiwp/96-06-042.html.