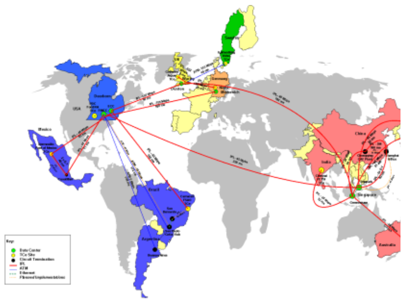
MIT Global Infrastructure System Simulator

The MIT Geospatial Data Center
jrw@mit.edu
doval@mit.edu

MIT ESD
Engineering Systems Division

Context for Building MIT's Global Infrastructure Simulator

In 2008 Walmart queries to determine total profit by store would take 6 hours to run [1]. As Stonebraker [2] noted in "One Size Fits All-an Idea Whose Time has Come and Gone", the standard central DBMS system is no longer capable of supporting global data-centric applications. Today's global information systems are highly customized and involve distributed servers and file stores and sophisticated indexing engines. "Predicting the performance of these systems is extremely difficult", says Williams, the Director of MIT Geospatial Data Center. To address these questions we have built a simulator so companies can test potential solutions before spending the money to purchase. "We built the MIT Global Infrastructure Simulator so we could accurately predict performance and detect vulnerabilities in global IT systems," says Sanchez, MIT's GDC Executive Director and Chief Architect. "The simulator has been tested by a Global 100 company with excellent results."

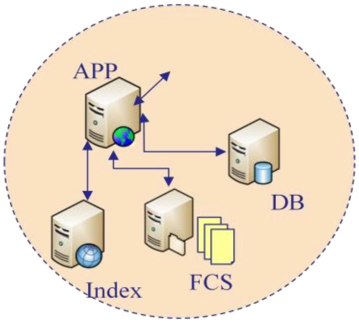


Global IT System

Overview of Global IT System Simulator

Today many companies depend on global IT systems to control their operations and supply chains around the world. These enterprise application systems may serve "real time" data to thousands of users which, in turn, consume this information in different locations concurrently and collaboratively. In analyzing a companies IT systems we have found that an enterprise's understanding of its own systems is often limited. No one person in the organization has a complete picture of the way in which applications share and move data between data centers. Furthermore, assessing the performance, security and scalability of "yet to be deployed" systems is problematic.

Here we describe a simulator which MIT has built to evaluate the performance, availability and reliability of large-scale computer systems of global scope. The goal is to provide enterprises and data center operators with a tool to understand the consequences of hardware and software infrastructure updates. These updates can include the deployment of new network topologies, hardware configurations or software applications. The simulator was constructed using a multi- layered approach and was optimized for multicore scalability. The simulator has been validated against the global systems of a Fortune 500 company. This simulation of large-scale IT infrastructures not only reproduces the behavior of data centers at a macroscopic scale, but allows operators to navigate down to the detail of individual elements, such as processors or network links. The combination of queuing networks representing hardware components with message sequences modeling enterprise software enabled reaching a scale and complexity not available in previous research in this area.



Data Center Building Blocks

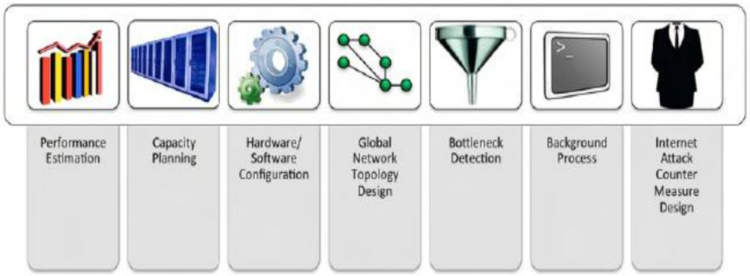
Global IT System Issues

Data sharing and collaboration capabilities have given global organizations the extensibility, agility and efficiency to operate without pause. However, these advantages have also lead to an unprecedented dependency on IT infrastructures. Kembel reported that each hour of downtime can be costly, from $200,000 per hour for an e-commerce service like Amazon.com to $6,000,000 per hour for a stock brokerage company [48]. Almost without exception, downtime is considered unaffordable and oftentimes the performance of the system and the availability of fresh information are sacrificed to keep the system operating. Unfortunately, a fully operational infrastructure cannot be left "as is" either, and three factors require making continuous adjustments to the system:

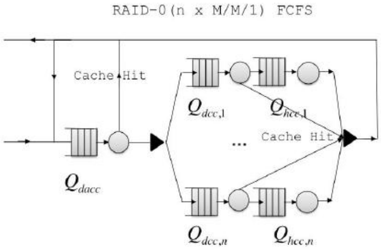
1. Continuous Innovation: Continuous integration of new features, state-of-the-art technologies or improved practices are necessary in order to maintain a competitive edge. For example, eBay deploys 300 features per quarter and adds 100,000 new lines of code biweekly [80].

2. Continuous Cost Reduction: As the infrastructure grows, continuously reducing the complexity of the infrastructure is a key mechanism to reduce costs. Akella et al. [2] propose out-of-the box solutions, component reuse, consolidation, standardization and interface simplification as critical efforts to succeed in the goal of reducing the complexity of IT systems.

3. Continuous Failure: Hardware failure is unavoidable. Typical data centers are composed by thousands of commodity servers that will inevitably fail, and hence, IT infrastructures needs to be designed to continuously deal with the dynamics of failure. During a year, on a cluster of 2000 nodes, Google reported 20 rack failures, 1000 machine crashes, thousands of hard drive failures among a variety of other network, configuration and power related incidents [21]. Under these circumstances, two driving forces, the need for "change" and the need "not to change", collide. Consequently, decisions affecting IT operations need to go through exhaustive reviewing processes across individuals, groups and divisions of the corporation.



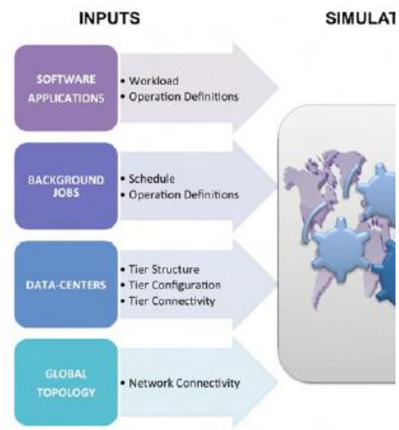
Simulator Inputs and Outputs

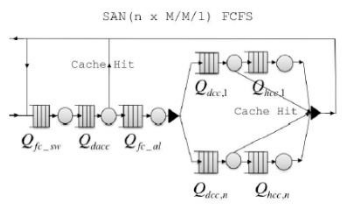


Raid Sub-Model

We believe that decisions concerning global IT systems need to be supported by quantitative analysis. We observe than human intuition about how global IT system's perform is often woefully wrong. Even the software architects of the product often have little idea how the system has been deployed, tuned and updated.

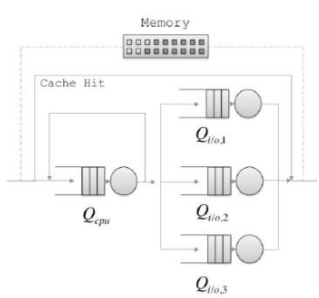
We have therefore built a simulator that we believe gives quantitative answers to questions about system performance. The simulator is an Multi-Agent Discrete Event Simulator implemented using multicore algorithms for speed.





SAN Sub-Model

Simulator Inputs and Outputs



RAM Memory Sub-Model

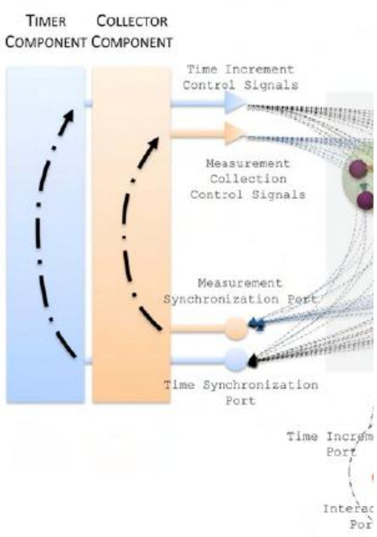
Input to Simulator

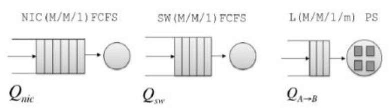
The simulator takes as input the hardware configuration deployed in each data center, the network topology, bandwidth and latency, application profile, resources allocated by individual client requests, and details on background processes. Based on the actual test performance of the application under various Use Cases as measured in the laboratory, a model of the applications performance is developed. Using this information, the simulator tracks all events and messages to produce estimates of the response time for user requests, along with measurements of the hardware allocation and network usage.

	USA	ENGLAND	GERMANY	INDIA	CHINA	AUSTRALIA	BRAZIL	MEXICO
USA		155	155	45	45	45	45	45
ENGLAND	155		155	45	45	45	45	45
GERMANY	155	155		45	45	45	45	45
INDIA	45	45	45		45	45	45	45
CHINA	45	45	45	45		45	45	45

AUSTRALIA	45	45	45	45	45		45	45
BRAZIL	45	45	45	45	45	45		45
MEXICO	45	45	45	45	45	45	45	

Bandwidth MB Between Sites





NIC, Net Switch Sub-Model

Holonic Multi-Agent Simulator Top Level Control

Servers and Data Centers

There are presently three different types of servers to handle data storage within the system: file, database and index servers. The former stores distributed heavyweight data and the latter two store centralized lightweight metadata. An application usually deals with two different types of metadata, and so explains the existence of two independent servers to handle this data:

- Information on location of files that is stored in the DBS.
- Information on the relations between files e.g. spatial relations between three different files, stored in the IXS.

File Server

The File Server is the main machine that provides a location for shared disk access, storing files containing the data that is used in the system. Other machines attached to the data network can connect to it to open, read, write or manipulate files. If the system data storage is distributed then apart from a central file server, every geographically dispersed site (or group of clients) has a local file server. This local file server stores data that is likely to be used by the users in that location. File servers in the system are often dedicated network-attached storage (NAS) devices with FTP server capabilities.

Central File Synchronization Server

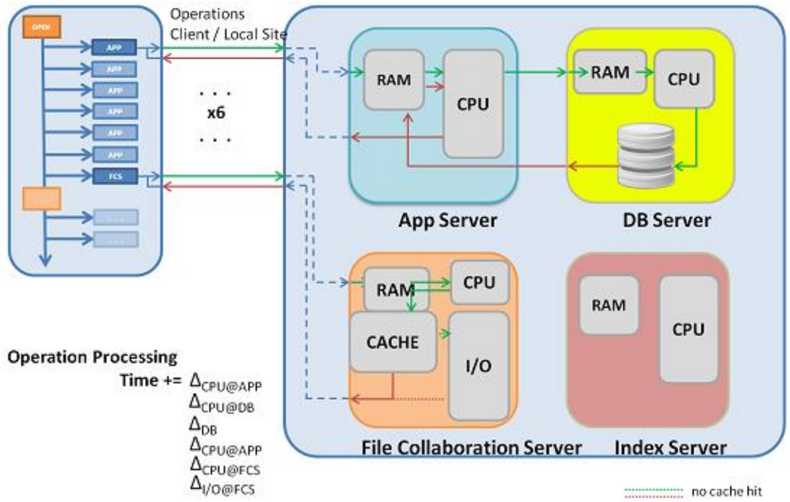
There may be an extra machine for replication purposes called a Central File Synchronization Server (CFSS). This server coordinates the movement of files around the global network.

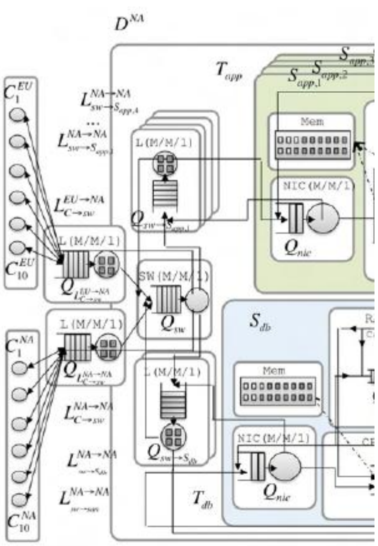
Index Server

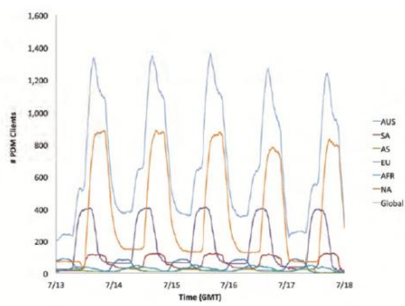
The index server stores information on the relations between the files in the different file servers within the network. For example, in computer-aided design projects, indexing allows to map geometric attributes of the models being designed to index server properties. This is necessary to know the exact geometric composition of the files within the project. In order to keep track of the spatial location of the parts of a 3D model, the index server within the system might use R-trees as an index mechanism. These sort of tree data structures split space with hierarchically nested, and possibly overlapping, minimum bounding rectangles (two-dimensional scenario) or boxes (three-dimensional scenario).

where to find a certain file at any given moment. It keeps track of last updated files, and also those with outdated information, within the network.

Cyber Attack Scenario





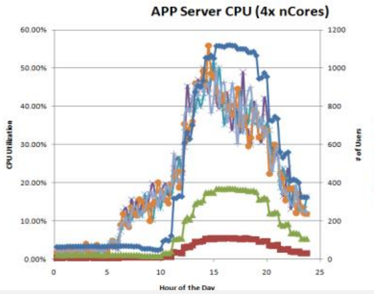


Resource Usage Against Time

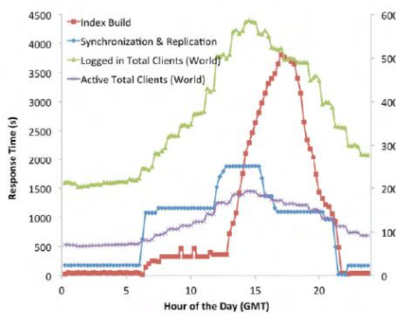
Data Center Agent Topology

Database Server

The database server keeps a log of the location of files in the network, so that the application server knows



Multicore Specification

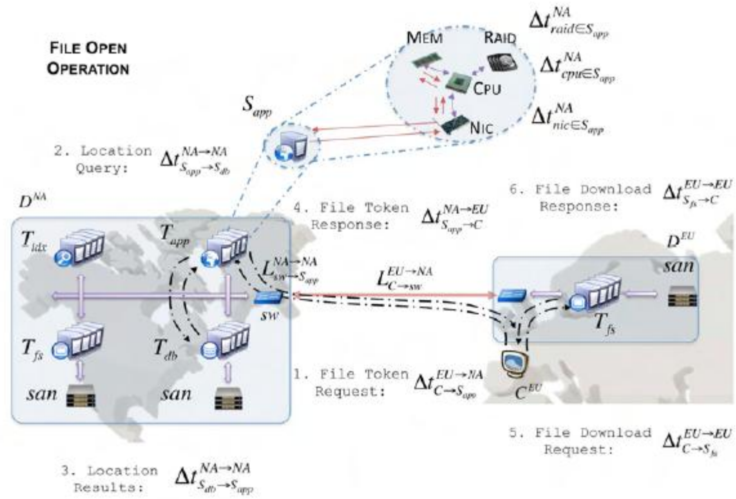


Index Build and Synchronization

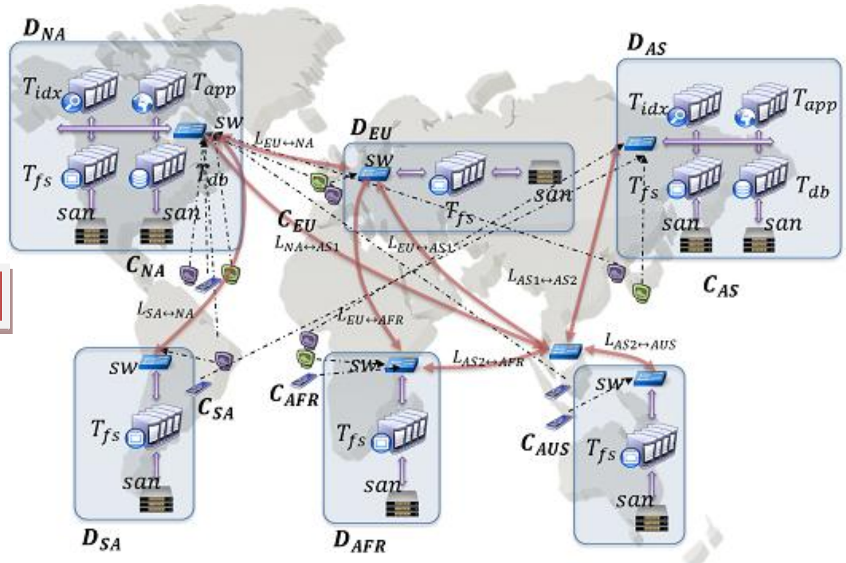
Results

The simulator outputs resource usage and user response times globally. In one study the results indicated that similar performance could be achieved with a reduction in data centers around the world. Typical cost savings can be around $1 billion.

Here we show generic results for work loads and response times for a global information system.



File Open Operation Event Cascade



Two Center Global IT Network Topology