International Journal of Engineering, Science and Mathematics (UGC Approved)
Vol. 8 Issue 5, May 2019,
ISSN: 2320-0294 Impact Factor: 6.765
Journal Homepage: http://www.ijmra.us, Email: ijesmj@gmail.com
Double-Blind Peer Reviewed Refereed Open Access International Journal - Included in the International Serial Directories Indexed & Listed at: Ulrich's Periodicals Directory ©, U.S.A., Open J-Gage as well as in Cabell's Directories of Publishing Opportunities, U.S.A

FUNDAMENTAL CONCEPTS OF CLOUD COMPUTING ARCHITECTURE AND SERVICE MODELS

Rakesh Upadhyay
Research Scholar
Dr. Sanjay Kumar Tiwari
Assistant Professor

Abstract

The use of cloud computing has brought about widespread mental shifts in the business and scientific communities. It offers an easy answer to the problem of IT infrastructure while reducing the costs involved. Cloud computing is the method that is described as the distribution of information technology resources on demand over the internet with a pricing model that works on a pay-as-you-go basis. One of the most difficult aspects of cloud computing is ensuring that all of the devices in use are constantly connected to a high-speed internet connection. People in all walks of life are becoming more worried about the ever-increasing amounts of data that are kept on their computers. They are looking for a solution that will allow them to store their important data, manage that data, have universal access to that data, and simply and securely share that data with others. Computing on the cloud is the one and only answer for the problems that people are having in this situation.

1. INTRODUCTION

1.1 Cloud Computing Architecture

In the introduction, it is said that the concept of offering centralised computing services through a network is not new. In the 1960s, mainframe timesharing technology was popular, but it was eventually supplanted by personal computers and architecture based on client-server relationships. Up until around ten years ago, the majority of business computer infrastructure was comprised of very powerful and expensive servers. The design of the infrastructure was a single, unified piece, and each of these powerful servers was readily capable of hosting 20–30 corporate applications.

International Journal of Engineering, Science and Mathematics (UGC Approved)
Vol. 8 Issue 5, May 2019,
ISSN: 2320-0294 Impact Factor: 6.765
Journal Homepage: http://www.ijmra.us, Email: ijesmj@gmail.com
Double-Blind Peer Reviewed Refereed Open Access International Journal - Included in the International Serial Directories Indexed & Listed at: Ulrich's Periodicals Directory ©, U.S.A., Open J-Gage as well as in Cabell's Directories of Publishing Opportunities, U.S.A

This market was controlled by only a handful of hardware vendors, such as IBM, Sun, HP, and Dec. The servers produced by these companies were prohibitively expensive to acquire and maintain, demanded a significant amount of time to install and upgrade, and, in some instances, were susceptible to server outages that could last for several hours until a vendor representative delivered proprietary replacement parts. Directly onto the hardware was the operating system installed, and the majority of the servers supported several applications inside the same operating system without providing either physical or virtual separation.

All different types of enterprises are turning to cloud computing as a means of storing their data on the cloud and ensuring that they have on-demand access to their own infrastructure. The architecture of cloud computing may basically be broken down into two categories. The first is known as the Front End, while the second is the Back End. Internet connectivity has been established between the two locations. Back end is responsible for ensuring that the information stored in the cloud is protected against unauthorised access. The back end is used by the service providers. It is responsible for managing all of the resources that are required in order to provide the services. It consists of a traffic control mechanism, a traffic management mechanism, a deployment model, a server, a high volume of data storage, a traffic management mechanism, and a security mechanism.

International Journal of Engineering, Science and Mathematics (UGC Approved)
Vol. 8 Issue 5, May 2019,
ISSN: 2320-0294 Impact Factor: 6.765
Journal Homepage: http://www.ijmra.us, Email: ijesmj@gmail.com
Double-Blind Peer Reviewed Refereed Open Access International Journal - Included in the International Serial Directories Indexed & Listed at: Ulrich's Periodicals Directory ©, U.S.A., Open J-Gage as well as in Cabell's Directories of Publishing Opportunities, U.S.A

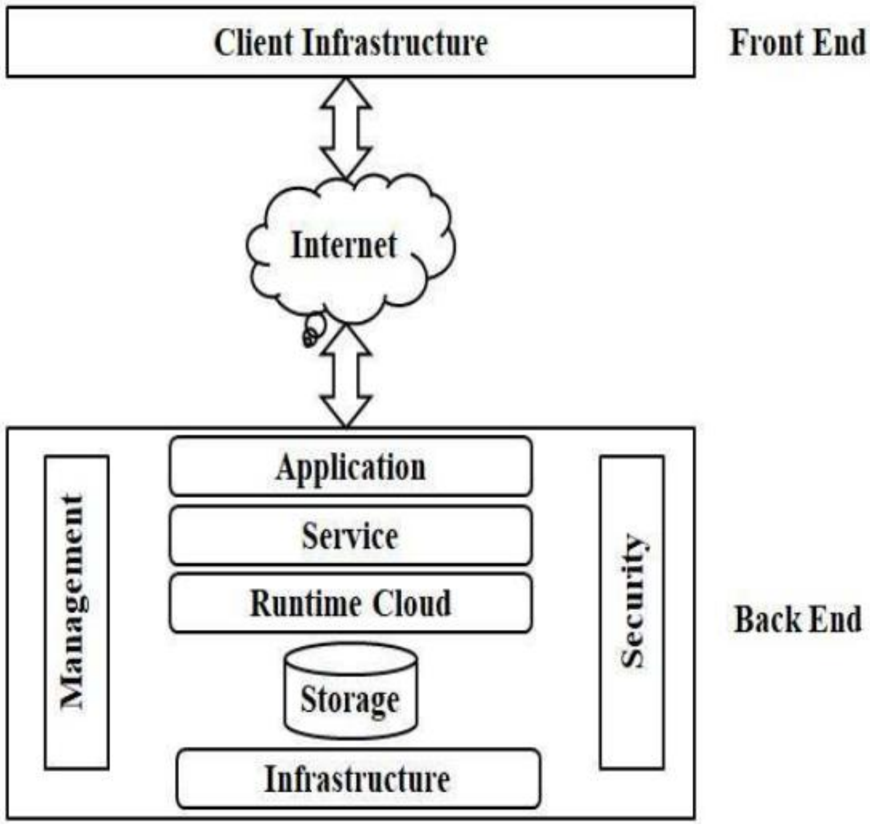

Fig. 1 Cloud Computing Architecture

On the other side, users interact with the front end of the application. Client-side interfaces and apps are required in order to get access to the front end of cloud computing. Tablets, web servers, and mobile devices are all included in this category. Access methods for cloud storage are distinct from those used for traditional storage since the cloud stores many different types of data belonging to many different users. The maximum providers are responsible for the operation of multiple access systems.

The following is a list of components that make up the architecture of cloud computing:

- Client Infrastructure: It is an element that is believed to be on the front end. The client infrastructure supplies the Graphical User Interface that users need in order to communicate with the cloud.

- Internet: It acts as a conduit via which parties on either end may communicate with one another.
- Application: It might be any kind of programme or platform to which the user wishes to get access. Service: It offers the following three categories of services:
- SaaS, PaaS and IaaS.
- Runtime Cloud: The runtime cloud provides virtual machines with an execution and runtime environment so that they may do their tasks.
- Storage: One of the most essential components of the overall architecture of cloud computing. It provides a significant amount of storage space on the cloud, which may be used for archiving and managing data.
- Infrastructure: The application level, the host level, and the network level are all covered by the services it provides. It has both hardware and software components in its make-up.
- Management: It is used to handle all of the components of the back end, including the application, service, runtime cloud, storage, infrastructure, and security concerns. In addition to this, it creates cooperation among them.
- Security: It has a built-in security mechanism that works behind the scenes.

2. CLOUD SERVICES

As was noted before, the cloud is able to supply each and every one of these technologies in the form of a service. This is the primary distinction between the cloud and "conventional" IT infrastructure, which can provide precisely the same technologies. This service may be accessed using a cloud management interface layer, which may either give access through a REST/SOAP API or a management console webpage.

2.1 Building Scalable Architecture

The ability to scale is one of the most critical things to consider when designing the architecture of an infrastructure. Systems in "conventional" non-cloud infrastructure are often architected to be able to support the possibility for future development as well as increased demand for resources. In order to adequately prepare for future expansion, organisations need to make significant financial investments up front. This results in chronic resource overprovisioning, and as a result, systems are inefficiently underused the majority of the time. Non-cloud infrastructures do not offer elasticity, thus system resources cannot easily scale up or down. On the other hand, since the architecture of the cloud is multi-tenant, computer resources are shared across a number of different applications. This shared multi-tenant environment operates on the presumption that all hosted apps cannot ordinarily be active at the same time. When one application is not being used, it is assumed that another application is being used. Cloud service providers are able to assign resources on-demand in this fashion, which greatly improves the effectiveness of the usage of computer resources. The cloud computing infrastructure is made up of a collection of shared physical resources, including servers, storage, and networks. The most sophisticated cloud management software will automatically monitor how resources are being used and distribute them according to the requirements of the system. Even during peak hours, the cloud service provider is responsible for ensuring that sufficient computer resources are available to serve all of its clients.

2.2 Horizontal Scaling vs. Vertical Scaling vs. Automated Elasticity

Vertical scaling (scale-up) anticipates that firms will make big investments up front and will not worry about available computer resources until demand is close to reaching capacity limits (see "scale-up strategy," the blue dotted line in Figure 1.5 for more information on this method). When this limit is reached, the firm will have to make significant investments in growing (for an example, see the "Huge capital expenditure" section of Figure 1.5). On the other hand, if the firm answers too slowly and the demand ends up exceeding the capacity that is currently available, this may have a negative impact on the business (see "You just lost your clients" in Figure 1.5). This vertical scaling strategy is more common in traditional, on-premises setups than than cloud ones.

Horizontal scaling (scale-out) provides businesses with the ability to extend their environments in manageable increments and on demand (see the maroon scale-out approach line in Figure 1.5). Horizontal scaling is a popular practise because it gives a means to rapidly scale out resources without substantial capital expenditures. Despite this, horizontal scaling still requires advance planning for both capacity and demand.

Automated elasticity indicates that a cloud provider will continually monitor a client's infrastructure and will scale it according on the needs of the customer. If you look at Figure 5, you will see that there is a black dashed line with an arrow at the end that represents the anticipated level of demand. When a corporation fails to meet this need, they throw away more money than when they are closer to meeting it. The red line represents "real demand," whereas the green line depicts automated elasticity, which closely follows the red line. The use of automated elasticity results in the smallest possible divergence from the anticipated level of demand and also reduces the likelihood of losing customers.

2.3 Elasticity

One of the primary advantages that the cloud provides is elasticity. It enables the scaling of computer resources dynamically dependent on the amount of real demand. In the "conventional" infrastructure, businesses would be required to make an advance forecast of the amount of computing resources they would use, but the major objective of cloud computing is to automatically modify the amount of computing resources available in response to demand (see Figure 1.5). The cost effectiveness of your cloud infrastructure will be improved to the degree that the degree to which your resource consumption matches real demand. How hosted software applications are designed and how infrastructure components are constructed are two crucial factors that substantially influence the flexibility potential of an infrastructure. For instance, it is difficult and costly to scale-out big monolithic programmes that cannot be divided into smaller modules. These applications cannot be modularized. It is quite probable that businesses will have to incur considerable capital expenditures in order to add additional computer capacity in order to support non-modular application types. When application architects are tasked with designing cloud-based apps, one of the considerations that must be made is how an application fits into the preexisting cloud model. Components of the infrastructure, such as web servers, application servers, cache servers, and databases, need to be able to scale up or down fast, depending on the circumstances. If an organisation uses large centralised components that are difficult to scale, the organization's use of these components may need to be re-architected in order to enable compatibility with the cloud. For instance, large centralised databases may need to be denormalized and broken down into smaller segments in order to be compatible with the cloud. Additionally, information on setup as well as other data that is shared across different cloud server instances has to be readily available from newly introduced components.

2.4 Cloud Deployment Models

Private clouds, public clouds, and hybrid clouds are the three most prevalent deployment types for cloud computing. The community cloud is an alternate paradigm that may be employed, but one that is used far less often.

A private cloud is one that is constructed and maintained entirely inside a single company. Software like VMWare, vCloud Director, and Open Stack are examples of cloud functionality enabling software that companies employ in their operations.

The term "public cloud" refers to a collection of computer resources that are made available by independent companies. Amazon Web Services, Google App Engine, and Microsoft Azure are three of the most widely used public cloud platforms.

A mixture of the computing resources offered by private clouds and public clouds is what's known as a hybrid cloud.

A community cloud is a kind of cloud that pools the computing resources of many businesses into a single pool and is administered either by the IT departments of those enterprises or by independent service providers.

2.5 Managed Hosting

Some businesses choose to delegate the administration of their information technology infrastructure to external service providers. In this scenario, the cloud capability of the services offered by the third-party supplier is either present or absent. The primary distinction between these kinds of deployment models and private clouds is that the former are run by a third party while the latter are hosted or maintained by the original organisation. There are additional service companies who provide managed hosting services in public clouds like Amazon and Google. Examples of these are Amazon and Google.

3. LAYERS OF CLOUD ARCHITECTURE IMPLEMENTATION

The implementation of cloud architecture consists of various different levels. A user may access apps that are hosted in the cloud environment by using a browser that is running on desktops as well as mobile devices. This browser is located on top of the other layers. The following two tiers are formed by the services and applications that cloud users access via the cloud. The next layer of the cloud architecture is comprised of software platforms, such as Oracle, SAP, and.Net, among others. These services and applications operate on these software platforms. The next layer down in the architecture is called the infrastructure layer, and it contains things like servers, data bases, storage, CPUs, and so on. The examples may be seen in picture 1.5.

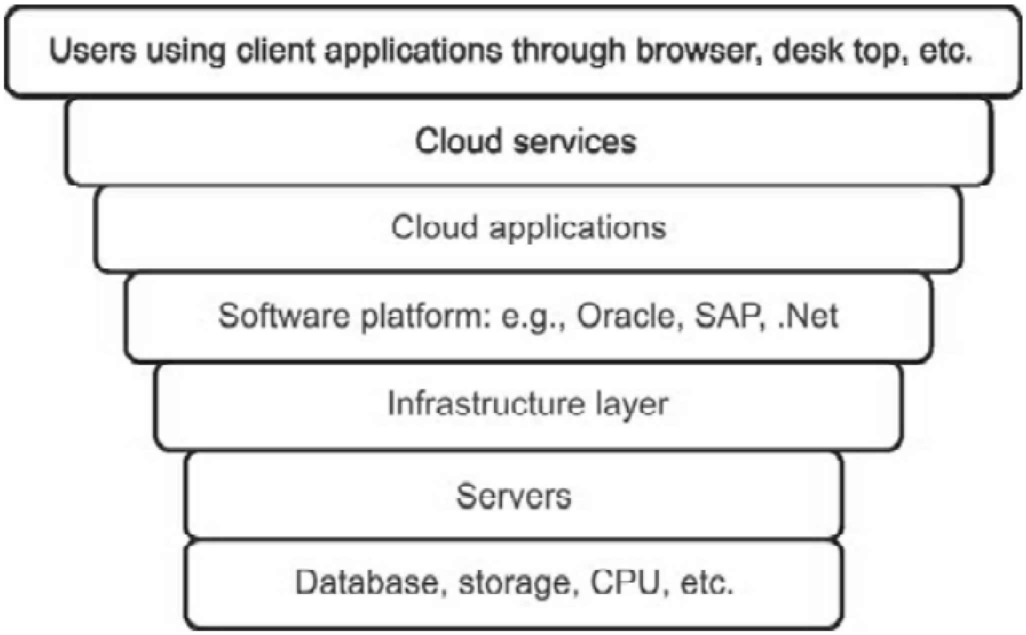

Fig. 2 Cloud computing layers.

3.1 Understanding Cloud Ecosystem

The architecture of cloud computing is realised via the collaboration of a number of different components and participants. Through the use of figure 1.6, we will get an understanding of each of them as well as describe their duties and responsibilities.'

Users or cloud subscriber: People who primarily fall into this category are those who make use of cloud services of some kind, whether it the SaaS, PaaS, or IaaS model.

Brokers: Even if the real users may desire a service, they will (or may) need an intermediary who represents their interests and integrates the needs of a number of users into a single common demand. For instance, a user may want to get driving directions, but in order to do so, he will need the assistance of an intermediary who both comprehends and satisfies these needs, or who may serve as a front end. It's possible that they are well-known websites or automobile manufacturers that offer navigational systems.

International Journal of Engineering, Science and Mathematics (UGC Approved)
Vol. 8 Issue 5, May 2019,
ISSN: 2320-0294 Impact Factor: 6.765
Journal Homepage: http://www.ijmra.us, Email: ijesmj@gmail.com
Double-Blind Peer Reviewed Refereed Open Access International Journal - Included in the International Serial Directories Indexed & Listed at: Ulrich's Periodicals Directory ©, U.S.A., Open J-Gage as well as in Cabell's Directories of Publishing Opportunities, U.S.A

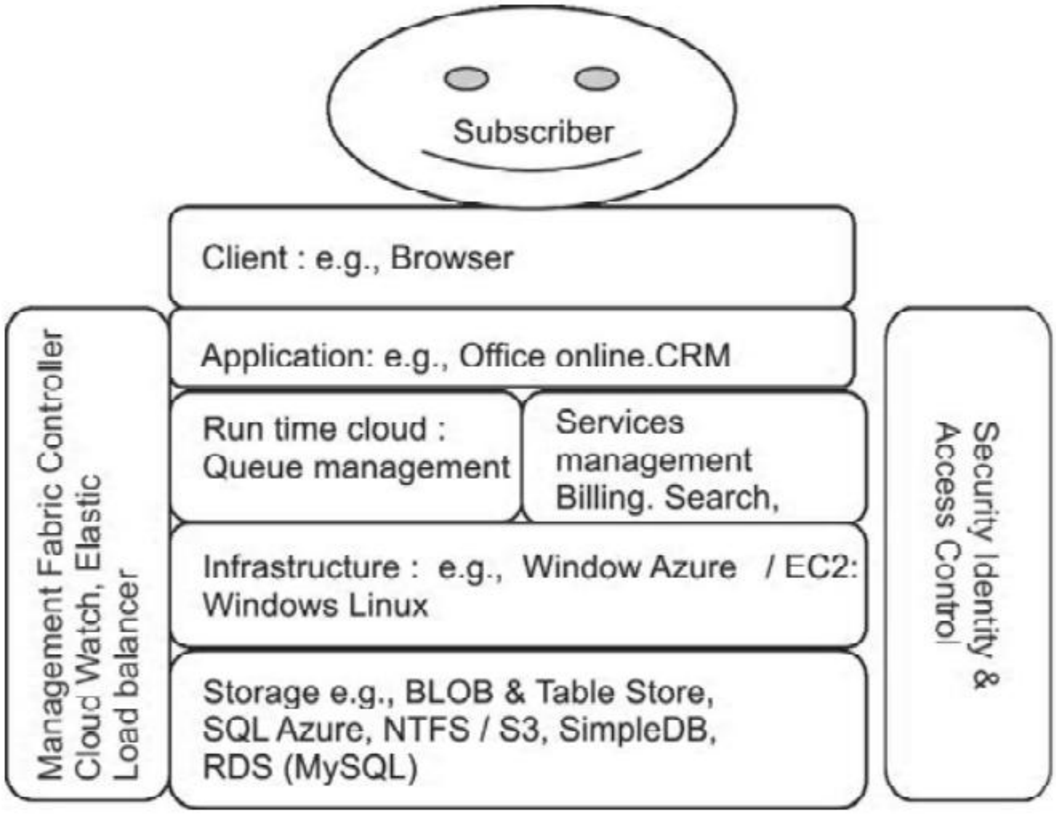

Fig. 1.6 Cloud architecture components.

A cloud broker's job include serving as a middleman between a cloud provider and an end user, with the goal of shielding the consumer from the complexities of the cloud. By acting as both a customer and a provider of services, they fulfil both functions via the activities of service intermediation, service aggregation, and service arbitrage. This organisation offers intermediate services such as identity management, performance reporting, and increased security. Additionally, this entity provides service aggregations such as the integration and transfer of data across a variety of cloud users and providers.

Service conceptualizer: These would be individuals or organisations that conceive of and construct services, then provide them to users or brokers for a charge after having done so. Therefore, this individual will create an application, which they will then market to users and brokers. This organisation also serves as the cloud environment's host for this service, which is provided by a cloud service provider.

Resources Allocator: It acts as a liaison between the cloud service provider and the consumers of the system. This body is responsible for the administration of the resources and assures the service levels.

Cloud Provider: A cloud provider is the most important of all the players in the cloud ecosystem since they are the ones who actually deliver cloud services. examples include Amazon, Microsoft, IBM, and Google among others. This organisation is responsible for providing and managing the computer infrastructure, including both hardware and software, that is necessary to make cloud services available to consumers through the internet. The kind of cloud services, such as SaaS, PaaS, or IaaS, will each need the cloud provider to fulfil a unique set of responsibilities.

Service request examiner and controller: This individual (or an automated system under the supervision of an expert) is responsible for allocating and redistributing resources according to the priorities that have been established. These priorities may include resource availability and criticality.

Pricer: This individual (or an automated system that is being monitored by an expert) determines the price and accounts for consumption based on the kind of request and payment plan that is being used.

VM Monitor: The virtual machine (VM) monitor is responsible for keeping track of the virtual machines and the availability of such machines.

4. CONCLUSION

Computing on the cloud is an entirely cutting-edge new technology. It is the evolution of virtualization, utility computing, Software-as-a-Service (SaaS), Infrastructure-as-a-Service (IaaS), and Platform-as-a-Service, as well as the development of parallel computing, distributed computing, and grid computing. In addition, it is the combination and development of grid computing (PaaS). The term "cloud" is a metaphor that is used to represent the web as a location in which computing has already been pre-installed and exists as a service. Data, operating systems, applications, storage, and processing power are all available on the web and are ready to be shared. Cloud computing provides consumers with a Pay-Per-Use-On-Demand model via which they may easily access pooled information technology resources through the internet. Where the information technology resources include a network, server, storage, application, service, and so on, and they can be delivered with the least amount of administration and also contacts with service providers, and where they can be implemented in a lot more quickly and easily. Computing in the cloud offers a number of benefits over conventional forms of computing and may make the usage of information technology resources considerably simpler.

REFERENCES

[1]. Ian Foster, Yong Zhao, et al, "Cloud Computing and Grid Computing 360- Degree Compared", in the proc. of Grid Computing Environments Workshop, 2008. GCE '08, pp. 1-10, 2008.

[2]. Pawel Plaszczak and Richard Wellner Jr., "Grid Computing: The Savvy Manager's Guide", Morgan Kaufmann Publishers, Noida, 2006.

[3]. Yuan-Shun Dai and Gregory Levitin, "Reliability and Performance of TreeStructured Grid Services", IEEE Transactions on Reliability, Vol. 55, No. 2, pp. 337 – 349, JUNE 2006.

[4]. Carsten Ernemann, et al, "On Advantages of Grid Computing for Parallel Job Scheduling", in Proc. of the 2nd IEEE/ACM International Symposium on Cluster Computing and the Grid (CCGRID.02), 2002.

[5]. Hai Jin, "Virtualization Technology for Computing System: Opportunities and Challenges", in Proc. Of The 10th IEEE International Conference on High Performance Computing and Communications, 2008.

[6]. Sam Averitt, Michael Bugaev, et al, "Virtual Computing Laboratory (VCL)", in the proc. of the International Conference on Virtual Computing Initiative, IBM Corp., Research Triangle Park, North Carolina, 2007.

[7]. S. Freitag, R. Yahyapour, G. Jankowski, R. Januszewski, "Virtualization Management for Grids and SOA", CoreGRID White Paper, No. WHP-0005, August 31, 2008.

[8]. Nathan Regola and Jean C Ducom, "Recommendations for Virtualization Technologies in High Performance Computing", in Proc. of 2nd IEEE International Conference on Cloud Computing Technology and Science, 2010.

[9]. Yiduo Mei, Ling Liu, Xing Pu, Sankaran Sivathanu, and Xiaoshe Dong, "Performance Analysis of Network I/O Workloads in Virtualized Data Centers", in IEEE Transactions on Services Computing, pp. 1-16. 2011.

[10]. N. M. Mosharaf Kabir and Raouf Boutaba, "Network Virtualization: State of the Art and Research Challenges", IEEE Communications Magazine, July 2009.

[11]. "Virtualization: Benefits and Challenges", An ISACA Emerging Technology White Paper, October 2010.

[12]. M. Sasikumar, Dinesh Shikhare and P. Ravi Prakash, "Introduction to Parallel Processing", PHI, New Delhi, 2006.

[13]. Tilak Agerwala and Siddhartha Chatterjee, "Computer Architecture: Challenges and Opportunities for the Next Decade", pp. 58 – 69, 2005.

[14]. Joel Emer, Mark D. Hill, Yale N. Patt, Joshua J. Yi, Derek Chiou and Resit Sendag, "Single-Threaded Vs Multithreaded: Where Should We Focus ?", pp. 14 – 24, 2007.