Published by :
http://www.ijert.org

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 6 Issue 02, February-2017

Cloud: Build vs. Buy?

Make an informed and intelligent decision

Gaurav Shad
School of Computing,
Informatics and Decision Systems Engineering
Arizona State University
Tempe, Arizona

Riya Chanduka
B.Tech in IT
Independent Scholar

***Abstract*—Ever faced with the prospects of having to build a new data center or expand existing facilities and got stuck between the numerous options available, including continuing to "build" and operate own data center capabilities or outsource the requirements to a third party vendor, i.e. "buy"? This paper will help make an informed "build vs. buy" decision by justifying the needs of cloud, deeply analyzing the process of buying all the services and building own data center and presenting a total cost of ownership model (TCO) that covers the cost aspect of decision making. The paper will give deep insights into other aspects like ROI, risks and opportunities as well and will give the pros and cons of both, buy and build. In the end, it will discuss the current scenario of what companies are doing and help decide which road is better to take.**

***Index Terms*—Cloud Computing, Data Center, Hybrid Cloud, Return On Investment (ROI), Total Cost Of Ownership (TCO).**

I. INTRODUCTION

Given the dynamic nature of today's IT industry and the rapid changes in business requirements, companies find themselves with the need to constantly innovate. A key to a company's success is the stable infrastructure at the back end that can support the requirements.

A huge shift is visible in today's data center environment. Similar to the past emergence of such paradigms as virtualization, storage consolidation, the next big shift in today's data center begins with cloud computing. Companies are now looking forward to move to cloud and the ones that already moved are looking to expand and get the maximum out of what they have.

This paper will first lay the groundwork on what cloud is, what services it provides and then move forward into the deep research that is to be done before making a decision of build vs. buy. The analysis includes deep understanding of the whole process on how to buy the services, how to build own data center and then present a TCO (total cost of ownership) model which focus on the cost factor for this decision. The paper will, based on current trends, helps to decide what would be an informed and intelligent decision in this case.

This paper is helpful for any organization or company or an individual who wants to move its services over cloud but is stuck between the decision of "buy" or "build" and want to gain more insights in these domains.

II. WHAT IS CLOUD COMPUTING?

We have heard this term cloud and cloud computing so many times but there is no common definition when it comes to cloud computing and in fact, there has been much discussion in academia and industry about what cloud computing actually means. The US National Institute of Standards and Technology (NIST) has developed a working definition that covers the commonly agreed aspects of cloud computing. The NIST definition – "Cloud computing is a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction. This cloud model is composed of five essential characteristics, three service models, and four deployment models" [1].

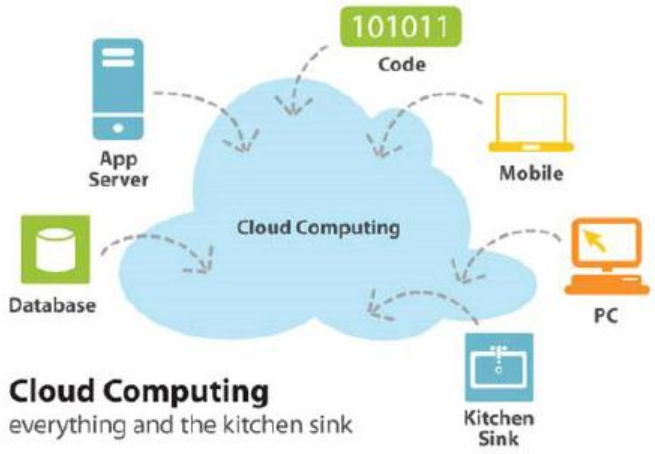

Fig. 1. Overview of Cloud Computing

A. Characteristics

1) *On-demand self-service*
A consumer can unilaterally provision computing capabilities, such as server time and network storage, as needed automatically without requiring human interaction with each service provider [1].

2) *Broad network access*
Capabilities are available over the network and accessed through standard mechanisms that promote use by heterogeneous thin or thick client platforms (e.g., mobile phones, tablets, laptops, and workstations) [1].

3) *Resource Pooling*
The provider's computing resources are pooled to serve multiple consumers using a multi-tenant model, with different physical and virtual resources dynamically assigned and reassigned according to consumer demand. There is a sense of location independence in that the customer generally has no control or knowledge over the

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 6 Issue 02, February-2017

exact location of the provided resources but may be able to specify location at a higher level of abstraction (e.g., country, state, or datacenter). Examples of resources include storage, processing, memory, and network bandwidth [1].

4) *Rapid Elasticity*
Capabilities can be elastically provisioned and released, in some cases automatically, to scale rapidly outward and inward commensurate with demand. To the consumer, the capabilities available for provisioning often appear to be unlimited and can be appropriated in any quantity at any time [1].

5) *Measured Service*
Cloud systems automatically control and optimize resource use by leveraging a metering capability1 at some level of abstraction appropriate to the type of service (e.g., storage, processing, bandwidth, and active user accounts). Resource usage can be monitored, controlled, and reported, providing transparency for both the provider and consumer of the utilized service [1].

B. Service Models

1) *Software as a Service (SaaS)*
The capability provided to the consumer is to use the provider's applications running on a cloud infrastructure. The applications are accessible from various client devices through either a thin client interface, such as a web browser (e.g., web-based email), or a program interface. The consumer does not manage or control the underlying cloud infrastructure including network, servers, operating systems, storage, or even individual application capabilities, with the possible exception of limited user specific application configuration settings [1].

2) *Platform as a Service (PaaS)*
The capability provided to the consumer is to deploy onto the cloud infrastructure consumer-created or acquired applications created using programming languages, libraries, services, and tools supported by the provider. The consumer does not manage or control the underlying cloud infrastructure including network, servers, operating systems, or storage, but has control over the deployed applications and possibly configuration settings for the application-hosting environment [1].

3) *Infrastructure as a Service (IaaS)*
The capability provided to the consumer is to provision processing, storage, networks, and other fundamental computing resources where the consumer is able to deploy and run arbitrary software, which can include operating systems and applications. The consumer does not manage or control the underlying cloud infrastructure but has control over operating systems, storage, and deployed applications; and possibly limited control of select networking components (e.g., host firewalls) [1].

C. Deployment Models

1) *Private Cloud*
The cloud infrastructure is provisioned for exclusive use by a single organization comprising multiple consumers (e.g., business units). It may be owned, managed, and operated by the organization, a third party, or some combination of them, and it may exist on or off premises [1].

2) *Community Cloud*
The cloud infrastructure is provisioned for exclusive use by a specific community of consumers from organizations that have shared concerns (e.g., mission, security requirements, policy, and compliance considerations). It may be owned, managed, and operated by one or more of the organizations in the community, a third party, or some combination of them, and it may exist on or off premises [1].

3) *Public Cloud*
The cloud infrastructure is provisioned for open use by the general public. It may be owned, managed, and operated by a business, academic, or government organization, or some combination of them. It exists on the premises of the cloud provider [1].

4) *Hybrid Cloud*
The cloud infrastructure is a composition of two or more distinct cloud infrastructures (private, community, or public) that remain unique entities, but are bound together by standardized or proprietary technology that enables data and application portability (e.g., cloud bursting for load balancing between clouds) [1].

III. NEED OF A CLOUD PLATFORM

Nowadays, companies are moving their services over the cloud and most of the well established companies have already done that. No matter which deployment model they follow, whether it is hybrid, private or public, there is some flavor of cloud computing all over the world. But why this came as a rapid shift is more interesting to analyze. Here are the few factors which will justify the needs of such system i.e. cloud and why it is important in current industry.

A. Rapid Launch
Companies are looking to diversify their services and products and they need a solution which allows them to configure and launch new services rapidly.

1) *Availability*
The users these days are using the services 24x7 and even a slight down of services can bring lives to a halt. This is one the main reasons of using cloud. The solution companies have should be up and running, available to the user all the time without fall [1].

2) *Scalability*
Companies needs a solution that can be scaled effectively so as to handle all the future needs as company grows.

3) *Cost Efficiency*
Companies need a robust technology or solution that is east to implement and cost efficient.

4) *Broad Network Access*
This is the feature listed in characteristics of cloud computing. The solution needs to be accessible by a broad range of thick or thin client devices, i.e. mobile, laptop etc.

This is some of the reasons which justify the trends we see in today's industry and why companies are moving to cloud.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 6 Issue 02, February-2017

IV. EVALUATE CURRENT SYSTEM CAPABILITIES

To check whether the current system or technology in use needs to be upgraded or replaced needs it to be analyzed on these factors:

1) *Change in needs*
The solution should be able to accommodate the dynamic changes that can affect the service. Example – a more computing resources are required by some specific application or process, it should be allocated to it without much human effort.
2) *Technical Support*
The platform or solution should have a 24x7 technical support at the back so that the operations can run smoothly.
3) *Upgrade*
The solution should also take care of the future requirement that can arise like more computing resources, expansion etc.

If the current solution does not satisfy above mentioned conditions, it needs to be replaced. A shift to cloud is needed. Then arise the question, whether to "buy" the data center services or build own data center?

V. "BUY" VS. "BUILD"

Making an informed buy or build decision requires a deep analysis of all the steps of how that can be done, costs involved, risks, opportunities and ROI. This section will give some insights on the process of buy or build and then gives a comparison on the basis of cost and Return on Investment (ROI).

A. Buy

The buy scenario is relatively easy. A data center service provider can help a company identify its needs, design the architecture of how the solution is going to be hosted and will give an estimation of the cost. In this case, the required resources will be allocated to the company in the vendor's data center and their access will be provided over the web. And depending on the type of service the company is using, whether it is SAAS, PAAS or IAAS, the control rights will be given on what it can modify and configure.

- Pros

1) *Implementation:*

This method is comparatively easy to implement.

2) *Scalability*:
The data center needs to have a lot of computing resources so as to satisfy its purpose to serve a large number of customers and scaling up or down is easy. In the case of scaling up, the customer will be allocated more resources with the minimum effort from customer and without any thinking of how it is going to be done.

3) *Operating costs*:
Cloud has this feature of measured service. That means the cost is calculated on pay per use model. And the cloud service provider will cover all other operating costs incurred due to maintenance and infrastructure support itself.

4) *Uptime*:
Uptime is a very important factor in cloud services. A good data center service vendor will have a high uptime so as to make sure all the customer services won't fall and also to maintain its image in the market.
5) *Future needs*:
The customer does not need to worry on how to accommodate the future requirements as the vendor will take care of all the upgrades and replacements that is to be done with time.

- Cons

1) *Security*:
It may turn into one of the critical issues with this scenario. The customer can have critical data stored in the data center and any leak or security breach on the side of vendor can have huge loss or repercussions on the customer. This trust issue hinders people from moving completely to cloud based systems. Example – customer can be a bank having its account holder's information.
2) *Dynamic costs*:
The cloud service has the so called feature of "Rapid elasticity" which is beneficial but can create problems sometimes. A large number of service requests or some error on the side of customer will lead to a large number of resource allocations and customer will incur more costs. This is a common issue faced by many organizations when they move to cloud based systems and buy all the services.
3) *Long term cost*:
If a well established company has all its services on cloud, deep analysis on ROI shows that in long term ,buying all the resources will cost more as compared to building own data center.

B. Build

Data center building is a tedious process and can take up to 6 months before it initiates all its services. A data center consists of the following components:

1. Power distribution units
2. Coalition suites
3. Heat rejection devices
4. Pump room
5. electrical system
6. UPS systems
7. Fuel oil storage tanks
8. Computer cabinets
9. Computer air handling unit (CARC)

- Data Center Model

This model will give a brief idea on what all things needs to be taken care of while building a data center and which layer will contribute how much to the total cost via a total cost of ownership model (TCO). The model will focus mainly on costs related with Real Estate, Infrastructure and Operations.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 6 Issue 02, February-2017

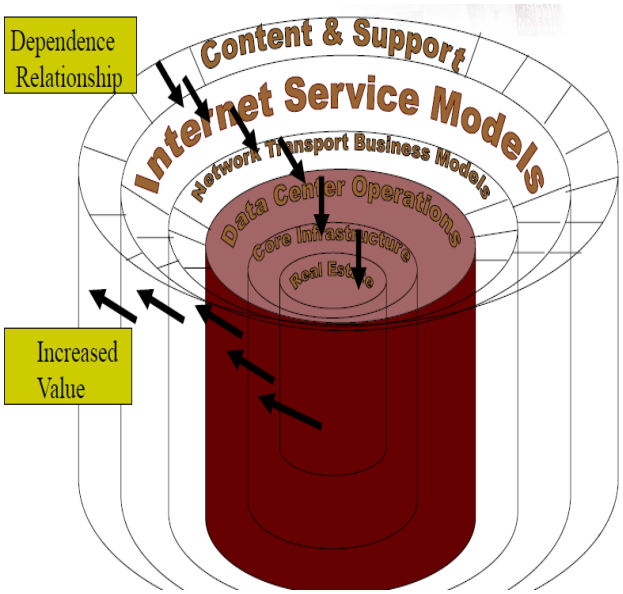

Fig. 2. Layers in a data center model

1) *Real Estate:*

The inner most layer is real estate. Setting up a data center needs space for setting up the facility. These factors needs to be carefully analyzed while looking for a space. This layer will contribute almost 50% to the total cost of the data center environment and this is the first part of TCO model.

- Scaling of the real estate in future can be very expensive if trying to setup a facility in a premium market area.
- Size is an important factor; if the chosen size is very small, data center will run out of space and if the chosen size is very big, the cost per rack will be very high.

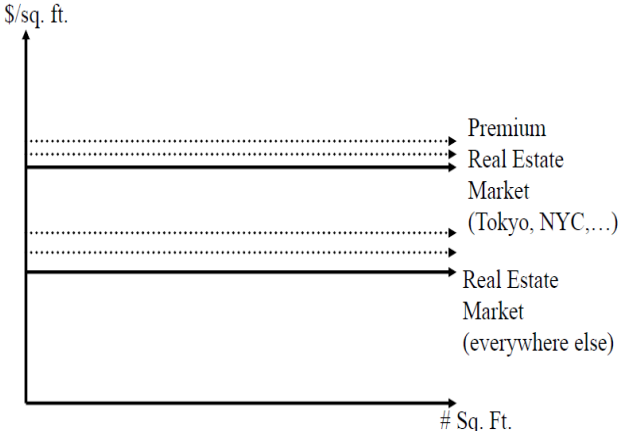

Fig. 3. Cost of scaling per Sq. Ft.

2) *Infrastructure:* This layer will include setting up the infrastructure required for a data center. This part also includes the cost incurred due to the redundancy based on the Uptime standards. The higher the level of redundancy, higher will be the cost. Setting up infrastructure includes:

- HVAC (heating venting air conditioning)

  This setup will maintain the right air temperature and the humidity. Current systems available in market are direct expansion, air cooled and chilled water system.

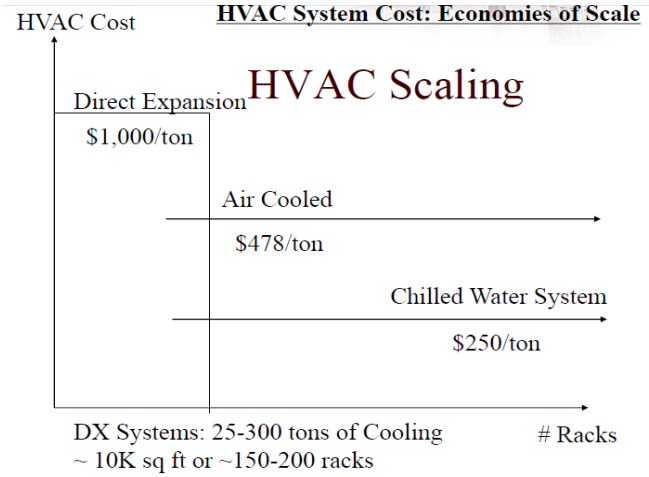

Fig. 4. Cost of HVAC scaling

- Security infrastructure

This setup includes the physical security of the setup and restricting the access only to authorized personnel.

- Power infrastructure

This setup includes monitoring of DC power, UPS and PDU's. It is also required to setup redundancy so as to decrease the chances of a system failure.

- Fire suppression

This setup is to ensure the safety from fire as the computing devices in the facility can get really hot and may lead to fire. Example – VESPA and dry pipes.

This layer is the second part of the TCO model. All the setups mentioned above will incur a huge cost I total. Building a good infrastructure is the key component for the life of a data center.

3) *Operations:*

This layer is more concerned with setting up the computing devices, their maintenance and upgrade required with time. This layer contributes to 15% of the total cost and 20-30% in yearly expenses. Major components that contribute to cost are:

- IT equipment

  This includes all the cost of initial purchase, installation and maintenance of the computing devices such as servers, storage, network components and software licenses.

- Operating costs

  This includes the cost of maintenance of the mission critical infrastructure such as HVAC, UPS and PDU's. Electricity costs are also included in this part.

- Staffing and other support

  A data center needs fully qualified technical staff to maintain the mission critical systems such as servers, HVAC, PDU's and also other non technical support staff for help desk, security etc.

  This layer is the third part of TCO model.

  The above mentioned steps need to be taken care of while building a data center.

  - Pros

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 6 Issue 02, February-2017

1) *Security*:
All the critical data and the services are present in own data center which decreases the chances of data leak or security breach.
2) *Long Term Costs:*
Deep analysis shows that building own data center will turn out to be more fruitful with time as the company grows.
3) *Dynamic Costs*:
The cost in maintaining the data center is pretty much the same whether one uses more computing resources or less computing resources. So a customer / owner do not have to worry about the dynamic costs as the requirements changes rapidly.

- Cons

1) *Investment*
Setting up a data center requires a huge initial investment and there is a risk attached with it. For most of the companies, this is the one time decision and they cannot just go over and try every possible option available in the market.
2) *Implementation*
It is more complex compared to the buying option.
3) *Scalability*
The data center may seem to accommodate all the future requirements but a time will come in future when it needs expansion and which may include a huge cost.

VI. TCO MODEL

The TCO model has four pillars in total which are mentioned below.

1. Real Estate
2. Infrastructure
3. Operations
4. Hidden Costs

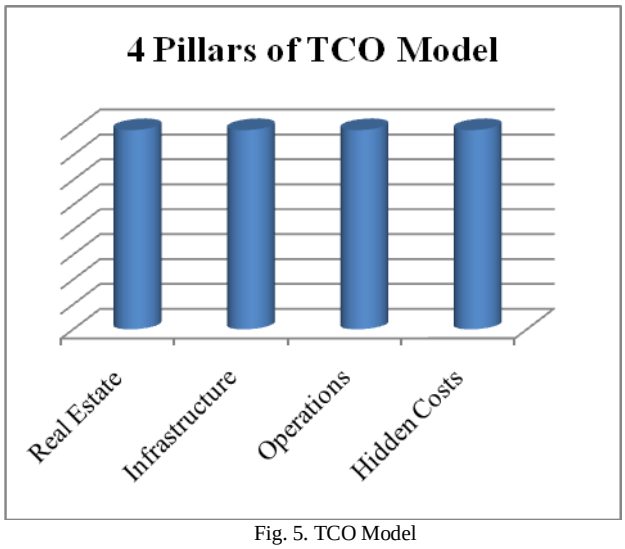

Fig. 5. TCO Model

Till now the TCO model has included the costs of buying real estate, setting up infrastructure and maintenance. This model also gives some insights on the hidden costs that can arise. Hidden costs can from many different sources including but not limited to some of the below mentioned examples:

1. *Disaster recovery planning and periodic testing*
This is the cost that often turns out to be more than initially budgeted for. One of the reasons is that any changes made in the data center environment must take into account the need to update the disaster recovery plan [2].
2. *Investments in redundant infrastructure*
Companies that build their own data center often ignore of cut down investments in redundant infrastructure which turns out to cost more lately. Example – Ignoring investment in disaster recovery planning for redundant infrastructure until they experience a disaster [2].
3. *Capital depreciation*
The investment made, especially on computing devices will depreciate over time and ROI will turn out to be less than expected if full utilization of services is not done [2].

The four pillars on which TCO is based on has covered the major portion of the cost in the data center building.

VII. COMPARISON

Based on the pros and cons in both the options, the answer to whether buy or build is not definite. It is more based on the one's requirements and trade off. The analysis of the TCO model gives the following result:

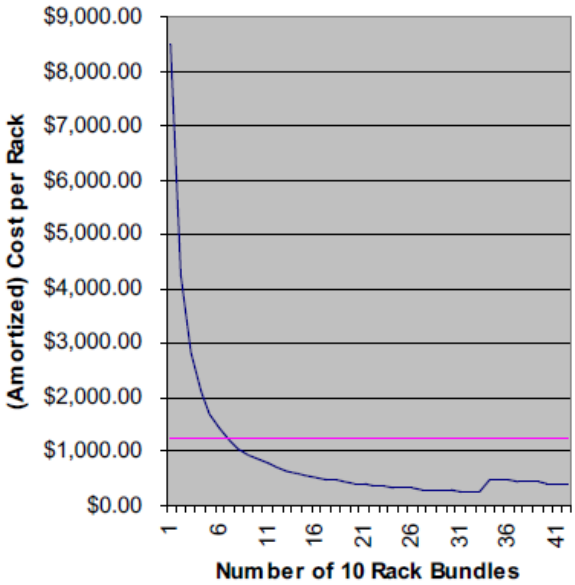

Fig. 6. The Buy vs. Build curve

The blue line in the figure shows the build curve and the pink line shows the buy curve. This model shows that initially, the cost of building would be high but as we expand, the cost of build would go lower than the cost of buy whereas the cost of buy will remain constant as it is to be paid for measured service (i.e. per resource).

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 6 Issue 02, February-2017

VIII. INFORMED DECISION

As we have seen in the previous section that there is no definite road when it comes to choose between buy and build but an intelligent decision would be to go for hybrid cloud.

A hybrid cloud is more like a mixture of buy and build. It has a private cloud (analogous to own data center) and all other services are bought from a third party vendor (analogous to buy).

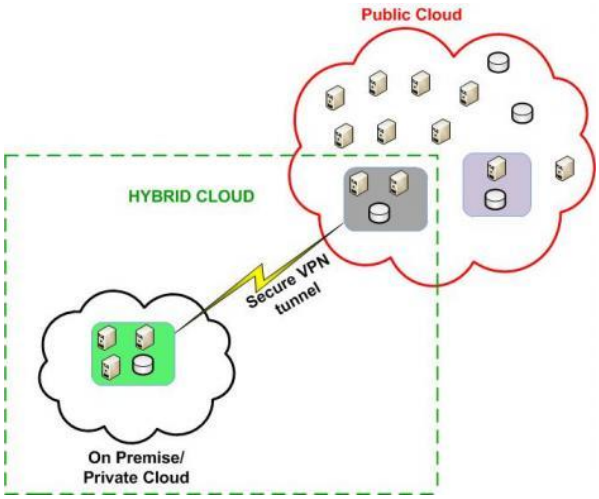

Fig. 7. Architecture of hybrid cloud

This is the latest trend in market. All the companies would start with a private cloud which then grows into a hybrid cloud and the below mentioned factors adds to its justification:

1. *Security*
   As a private cloud consists of a private cloud, it can be used to store all the critical data and perform any computations on it and the rest of the data or computing or services can be moved over to the a cloud bought from a third party vendor. Doing this will largely decrease the chances of leak or security breach. This is one the main reasons why this model has emerged in recent times.
2. *Protocols*
   The company can setup own protocols to communicate between the private and the public cloud which will enhance the security. Example – firewalls, TOR, GRE with IPSec can be used.
3. *Flexibility*
   This model is more flexible compared to buy and build as it has the aspects of both and company can literally choose which one to use and even distributes the load as required.

A future hybrid cloud might outsource some of the features of public cloud like IAAS and SAAS and use it in conjunction with company's internal infrastructure.

CONCLUSION

The analysis done in this paper shows that there is no definite answer when it comes to "buy" vs. "build" and this is more like a tradeoff between the two. A more informed and intelligent decision would be to with a hybrid cloud as it has the good aspects from both and it the emerging model of the future.

REFERENCES

[1] Peter Mell, Timothy Grance, The NIST definition of Cloud Computing, http://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecialpublication800-145.pdf

[2] Laura Stevenson, Data center: "Build vs. Buy", BlackIron Data

[3] Kent Christensen, IT and the cloud: Buy, Build or both?, http://www.datalink.com/getattachment/d31a0344-4d19-484d-8013-5ccf32103e75/IT-and-the-Cloud-Buy,-Build-or-Both.aspx

[4] Cloud services, AT&T, https://www.business.att.com/enterprise/Portfolio/cloud/

[5] Dingledine, Mathewson, Syverson, Xinwen Fu, and Yinglin Sun, Chris Zachor, Anonymizing Network Technologies, https://www.cse.unr.edu/~mgunes/cpe401/cpe401sp11/student/tor.ppt

[6] Intel Cloud Builder Guide, http://www.intel.com/Assets/PDF/general/icb_ra_cloud_computing_canonical_ubuntu.pdf

[7] Darpan Saini, To Build or to Buy?, http://blog.lendit.com/wp-content/uploads/2015/03/Cloud-Lending-Solutions-Build-v-Buy.pdf

[8] Intel Data Center Manager, https://software.intel.com/sites/datacentermanager/

[9] Ubuntu Enterprise Cloud, https://www.ubuntu.com/cloud

[10] Amazon Web Services, https://aws.amazon.com/

[11] IT Service transformation with hybrid cloud, http://www.principledtechnologies.com/EMC/hybrid_cloud_build_vs_buy_0316_v3.pdf

[12] Data center build vs. buy: How to Decide?, https://www.greenhousedata.com/whitepaper/build-vs-buy

[13] Dom Nicastro, One Year Later, Build vs. Buy Your Marketing Cloud, https://www.greenhousedata.com/whitepaper/build-vs-buy

[14] Extreme Networks: Data Center Networking, http://learn.extremenetworks.com/rs/extreme/images/Data-Center-Design-Guide.pdf

[15] J. Staten, S. Yates, F. Gillett, W. Saleh, and R. Dines, Is cloud computing ready for the enterprise? , Technical report, Forrester Research, March 2008

[16] Albert Greenberg, James Hamilton, David A. Maltz, Parveen Patel, The Cost of a Cloud: Research Problems in Data Center Networks, http://research.microsoft.com/en-us/um/people/dmaltz/papers/dc-costs-ccr-editorial.pdf

[17] Eric Brewer, Lawrence Ying, Lawrence Greenfield, Robert Cypher, and Theodore Ts'o, Disks for Data Centers, http://research.google.com/pubs/pub44830.html

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 6 Issue 02, February-2017

BIOGRAPHIES

Gaurav Shad is a Graduate Computer Science Student with the School of Computing, Informatics and Decision Systems Engineering at Arizona State University. His interest includes Big Data processing and analysis, Visualization and Machine Learning.

Riya Chanduka has done her Bachelors in Information Technology from Jaypee Institute of Information Technology, Noida. Her interest includes informatics and management information systems.