Cloud Computing: A New Paradigm in IT that has the Power to Transform Emerging Markets

Article *in* International Journal on Advances in ICT for Emerging Regions (ICTer) · September 2012

DOI: 10.4038/icter.v4i2.4672

San Murugesan
BRITE Professional Services

International Journal on Advances in ICT for Emerging Regions

Special Issue On Cloud Computing

Volume 04 Number 2 September 2011

CONTENTS

GUEST EDITORIAL

RESEARCH ARTICLES

Guest Editorial

Cloud Computing: A Boon for Emerging Regions

San Murugesan

A number of converging factors have given rise to cloud computing emergence as a new IT service delivery model and IT usage model that is attractive not only to individual users but also to businesses, educational institutions, governments and community organizations. In cloud computing, applications computing resources and application development platforms are provided as a service to users accessed through the Internet or other dedicated networks. It promises to offer utility-like availability of huge computing resources at low cost, higher flexibility and better scalability. This new computing delivery model is poised for causing a paradigm shift in IT delivery and usage, and is expected to dramatically transform how we embrace IT.

According to *The Economist,* "The rise of the cloud is more than just another platform shift that gets geeks excited. It will undoubtedly transform the information technology (IT) industry, but it will also profoundly change the way people work and companies operate. It will allow digital technology to penetrate every nook and cranny of the economy and of society, creating some tricky political problems along the way". Cloud computing is particularly attractive to emerging regions.

As cloud computing now begins to move from the fringe to the mainstream, there is considerable excitement (and hype) surrounding the movement among its various stakeholders. Corporations are eagerly investing in promising cloud computing technologies and services not only in developed economies but also increasingly in emerging economies – for example in India, China, Singapore, Philippines and South Africa to name a few. Several IT companies have begun developing and deploying cloud computing platforms, applications and tools. IT departments in many enterprises are now being asked, or soon might be asked, to explore how their organizations can embrace cloud computing and to deploy and monitor applications on a cloud.

Paradigm Change

Until recently, the processing and storage of data was done by your local — personal or enterprise — computer. With cloud computing, the data, applications, and processing power are all being provided by the cloud. Cloud computing thus allows users and businesses to get away from the limitations of their local environment; all you need is an Internet connection and a Web browser to be productive from anywhere in the world. It allows you to access computation-and/or storage-intensive applications using just a light, low-end computer, since your computing/storage needs can be fulfilled by the cloud. As your data is stored at a secure data center, the risk of damage to or loss of data on your desktop computer is reduced. It should come as no surprise that cloud computing has attracted the interest of key players in the IT industry, all of whom hope to establish profitable business models and take the lead in the emerging cloud computing market.

In response to growing current interest on cloud computing among researchers, IT professionals and the IT industry, this issue focuses on cloud computing in emerging regions.

In This Issue

In the opening article, “Cloud Computing: A New Paradigm in IT that has the Power to Transform Emerging Markets,” San Murugesan presents a brief but comprehensive overview on cloud computing and its service and deployment models, and outlines the benefits and risks of cloud computing. He then examines why the clouds are particularly attractive to emerging regions.

While cloud computing is gaining momentum and many different services are being offered targeted at businesses of all sizes and in different sectors as well as individuals, in the context of wide spread deployment of clouds for core, business critical application there is concern on cloud reliability, availability and security (RAS). In the article, “Cloud Computing Reliability, Availability and Serviceability (RAS): Issues and Challenges,” Farzad Sabahi reviews issues and challenges of cloud computing’s RAS. Beginning with a brief discussion on virtualization technology, a key element of cloud infrastructure, he outlines the key RAS issues faced in adopting clouds, examines intrusion detection methods and outlines counter measures to improve cloud RAS.

Another key issue that confronts current and potential cloud adopters is, trust in cloud computing services and in cloud service providers. Trust plays a key role is widespread adoption and use of clouds. Mohamed Firdhous, Osman Ghazali and Suhaidi Hassan, in their article, “Trust

Management in Cloud Computing: A Critical Review," examines this issue. They look at what trust is and how trust has been applied in distributed computing environment, and then summarises different trust models that have been proposed for various distributed systems. Then they discuss trust management systems for cloud computing focusing on their applicability and implementability and compare trust models and trust management systems.

We conclude this issue on a positive note with a real-world case study - streamlining India's grain supply chains embracing cloud computing. In their article, "ICT Leap-frogging Enabled by Cloud Computing for Emerging Economies: A Case Study on Streamlining India's Grain Supply Chains" H.S. Jacob Tsao, Shailaja Venkatsubramanyan, Shrikant Parikh and Prasenjit Sarkar introduce their research into streamlining India's grain supply chains and how we can leverage cloud computing with very little additional investment in IT resources to successfully and efficiently managing grain supply chains in emerging economies. It illustrates the paradigm shift and mindset that is needed to embrace the potential of cloud computing to address some of the key problems in different sectors currently facing emerging economies.

IT'S TIME TO EMBRACE CLOUDS

Cloud computing, driven by economic imperatives and the promise of flexibility and convenience, will remain an area of keen interest for years to come. If the current momentum in cloud computing continues, both individuals and enterprises will be computing in one or more clouds, public or private, in a major way within the next five years. Though there are major stumbling blocks for enterprises moving some of their applications into the clouds — such as reliability, performance, bandwidth requirements, trust and security issues — these barriers are being gradually lowered or removed.

We are into a new, transformational, disruptive phase of IT in the 21st century. Like earlier transformations, we expect to see wide-ranging implications and new opportunities as we start to embrace cloud computing and the "everything-as-a-service" model.

It's time to give cloud computing serious and favorable consideration. By embracing clouds we can help reshape the IT landscape for the benefit of all - individuals or businesses of all sizes and shapes. Widespread adoption of cloud computing will also help close the prevailing digital (information) divide in underdeveloped and developing economies.

I hope you will find that the articles in this special issue present useful insights and ideas about cloud computing and how you can embrace its potential for the benefit of individuals, businesses, governments and the society in the emerging regions. I also hope this issue also motivates researchers to address the limitations of and concerns on clouds and embark on cloud-based solutions to real problems facing the businesses and community in emerging regions.

Cloud Computing: A New Paradigm in IT that has the Power to Transform Emerging Markets

San Murugesan

***Abstract*—Cloud computing offers utility-like availability of computing resources and applications via the Internet at low cost. It has the power to transform enterprises in emerging markets offering access to advanced IT infrastructure and applications that many of them couldn't afford. An informed understanding of cloud computing and its benefits, limitations and risks are the keys to successfully embrace the opportunities this new computing paradigm offers. This article presents an overview of cloud computing concepts, cloud services, cloud hosting models, and applications. It also identifies potential risks and discusses the prospects of clouds and what businesses and individuals can do to successfully embrace cloud computing.**

***Index terms*—Cloud Computing, Virtualisation, Software as a Service, Platform as a Service, Infrastructure as a Service**

I. Introduction

Computing landscape — both enterprise and personal – is changing again with the emergence of cloud computing. In its evolution over the past four decades, computing has passed through several stages -- from mainframe computers to minicomputers to personal computers to network computing, client-server computing, and distributed computing. Now, coming full circle, computing is migrating outward to the clouds, to distant computing resources reached through the Internet. In the cloud computing model, applications and computing resources are provided to users as a service through the Internet (i.e., "the cloud"), similar to utilities like electricity supplies. Users can use computing resources when and where they need them and in the amount they need, and pay for the resources used. Cloud computing offers huge computing power, on-demand scalability, and utility-like availability at low cost and is a radically new IT delivery and business model.

Cloud computing causes paradigm shift in how we deliver and use computing resources and applications and it is poised to have profound impact on businesses, commerce and individuals. It can reduce significantly information technology costs - by as much as 50 to 80 percent - and offer businesses better agility and flexibility. As a result cloud computing is gaining keen interest among IT vendors, professionals and users - business, governments and individuals. Cloud service delivery model is particularly attractive to business and individual users in emerging markets. The rise of the cloud is more than just another IT platform shift; clouds will transform not only the IT industry but also the emerging markets, and the transformation will be profound. Clouds presents many opportunities for emerging markets to profoundly improve the way people communicate, share information and work and companies operate. Emerging markets can embrace clouds offering new kinds of applications. Even a small business can suddenly serve millions of customers all around the world and the flexibility and versatility clouds enables companies to enter new markets quickly with very little overhead costs.

In this paper, we examine cloud computing concepts, cloud services, cloud hosting models, and applications. We also identify potential risks and discuss the prospects of clouds and what businesses and individuals can do to successfully embrace cloud computing. Finally, it also discusses clouds' prospects and implications to businesses and individuals.

II. Computing

The "Cloud" is an evolution of distributed computing and of the widespread adaption of virtualization and service oriented architecture (SOA). In cloud computing, IT-related capabilities and resources are provided as service, via the Internet and on-demand, accessible without requiring detailed knowledge of the underlying technology. Depending on one's perspective, clouds can be described in different ways [see Geelan 2009] and hence the term cloud computing remains fuzzy and carries several definitions.

- "Cloud computing is a model for enabling 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" [1].

This is an abstract version of the keynote presented at ICTer2010 Conference, Colombo.

San Murugesan is Adjunct Professor at the University of Western Sydney in Australia and Director of BRITE Professional Services.

He served in academia and industry for a long period and has published widely. His current areas of interest are green IT, cloud computing and IT in emerging markets. He is co-editor of the upcoming book, *Harnessing Green IT: Principles and Practices* (Wiley, 2011) and editor of *Handbook of Research on Web 2.0, 3.0, and X.0: Technologies, Business, and Social Applications* (Information Science Reference, 2009). Dr Murugesan serves as Associate Editor-in-Chief of IEEE *IT Professional* magazine, and green technology and IT in emerging markets editor for the magazine. He also serves on the editorial board of several other international Journals, IEEE Technical Steering Committee on Green Computing and the advisory board of Global Science and Technology Forum. He is a Fellow of Australian Computer Society and Senior of IEEE Computer Society.

He can be contacted: san1@internode.net

- "A style of Computing where scalable and elastic IT capabilities are provided as a service to multiple customers using Internet technologies" – Gartner.

Key benefits of computing clouds are reduced capital and operational costs, improved flexibility and agility, ubiquitous access, and easier and quicker application development and deployment [2].

A. Cloud Structure

A computing cloud is a massive network of computing resources interconnected in a grid running in parallel, harnessing the resources of each to provide huge computing and storage capacity. The cloud is accessible from a PC or a smart mobile phone, over the Internet through a Web browser. Users' request is processed by the cloud's system management software which manages the cloud resources and assigns and monitors users' processing tasks. The system software finds suitable resources to perform the requested task and calls the system's provisioning service which request and block necessary resources in the cloud and launch the appropriate application(s). Once an application is launched, the cloud's monitoring and metering functions track resource usage and bills the users for the services used. Automatic dynamic management of the cloud's tasks and monitoring and reporting the usage are key aspects of cloud computing.

B. Attributes of Computing Clouds

Computing clouds have several distinguishing attributes. They:

- have massive resources at its disposal and can support several users simultaneously.
- support on-demand scalability of users' computational needs.
- offer ubiquitous access. Stored data and applications are accessible by authorized users anywhere, anytime.
- facilitate data sharing, enterprise-wide data analysis, and collaboration.
- can self-reconfigure providing continuous availability in case of failure of its computing resources.
- offer enhanced user experience via a simplified Web browser user interface.

C. What's New?

While the clouds draw on some of the older foundations of IT such as centralized shared resources pooling, concept of utility computing and virtualization, they incorporate new mechanisms for resource provisioning and dynamic scaling. Further, it adopts new business and revenue models and incorporates monitoring provisions for charging for the resources used. Cloud computing became viable for wider offering and adoption only recently with the adoption of broadband Internet access and advances in virtualization and data centre design and operation as well as the crucial philosophical and attitude change by both the IT vendors and users.

D. A Paradigm Shift

Cloud computing represents a paradigm shift - transition from computing-as-a-product towards computing-as-a-service. It's a transition from buying hardware and software as products which we install, configure, use and maintain to using applications and computing infrastructure in the clouds as a service, paying on the go for the resources used. With cloud computing, you aren't tied to a particular computer or a specific private network to run or access your applications and data. Further, cloud computing makes it easier for group members in different locations to collaborate both synchronously and asynchronously. It also facilitates a variety of mobile computing applications which are poised for major uptake.

III. Cloud Computing Service Models

Cloud services can be classified into four broad categories: software as a service (SaaS), platform as a service (PaaS), infrastructure as a service (IaaS), and other Internet-based Services (IBS). Each service category can be used independently or used in combination with others.

A. Software as a Service

In the SaaS model, end-user applications are hosted by a cloud vendor and delivered as a service to customers across the Internet on some form of "on-demand" billing system or free as in some personal applications. This model eliminates the need to locally install and run the application on the user's computer, and thereby also alleviates the users from the burden of hardware and software maintenance and upgrades. The software license is not owned by the user. Costs to use the service become a continuous expense rather than an up-front huge capital expense at the time of purchase. Examples of SaaS include Webmail, Google Apps, Force.com CRM, Quicken online accounting, NetSuite's Business Software Suite**,** Sun Java Communications Suite and Paychex payroll management system. SaaS clouds are also called *software clouds*.

B. Platform as a Service

In the PaaS model, application development platforms and middleware systems are hosted by a vendor and offered to application developers, allowing them to simply code and deploy without directly interacting with the underlying infrastructure. The platform provides most of the tools and facilities required for building and delivering applications and services - workflow facilities for application design, development, testing, deployment, and hosting, as well as application services such as Web service integration, database integration, security, storage, application versioning, and team communication and collaboration. Examples of PaaS include Google App Engine, Microsoft Azure, Amazon's Web

Services and Sun Microsystems NetBeans IDE. The PaaS clouds are also called *platform clouds* and *cloudware*.

C. Infrastructure as a Service

In an IaaS cloud, raw computer infrastructure, such as servers, CPU, storage, network equipment and data centre facilities are delivered as a service on demand. Rather than purchasing these resources, clients get them as a fully outsourced service for the duration they need them. The service is billed according to the resources consumed. Amazon Elastic Compute Cloud (EC2), GoGrid, and FlexiScale are some of the examples of IaaS clouds. These types of clouds are also called *utility computing* or *infrastructure clouds*.

D. Other Internet-Based Services

Storage, middleware, collaboration, and database capabilities are other Internet-based services that clouds offer directly to users. With cloud storage, data is stored in multiple third-party servers, rather than on dedicated servers used in traditional networked storage, and users see a virtual storage. The actual storage location may change as the cloud dynamically manages available storage space; however, the users see a static location for their data.

Data storage is a common infrastructure use of cloud computing. Key advantages of cloud storage are cost, data safety and availability. Data stored in a cloud is safe against accidental erasure or hard-drive failures as a cloud keeps multiple copies of data across multiple physical machines continually. If one machine crashes, the data that was in that machine can be retrieved from other machine(s) in the cloud. Cloud vendors generally offer more powerful security measures than what a small business could afford. Enterprise data storage in clouds, however, raises some concerns which are discussed later.

IV. Cloud Computing Hosting Models

Based on where the clouds are deployed and by whom, computing clouds are classified into five categories: public clouds, private or internal clouds, virtual private clouds, vertical or community clouds, and hybrid clouds.

A. Public Clouds

A most common and widely known form of clouds, public clouds are open to anyone and the cloud infrastructure and applications are owned by the cloud provider – the organization that offers the cloud services. Public cloud services are usually offered on a pay-per-usage model; some applications on public clouds can be used free of charge.

B. Private or Internal Clouds

These are clouds provided and controlled by an enterprise behind its firewall. Unwilling to head into public clouds because of concerns surrounding them and compliance requirements, some enterprises have begun to deploy their own cloud computing environment for their exclusive use to gain operational efficiencies and to effectively use their existing resources.

C. Virtual Private or External Private Clouds

These are segments of a public cloud designated for a user with added provisions for meeting specific security and compliance requirements. They provide users more control over the resources they use than in a pure public cloud. An example of this type of cloud is Amazon's Virtual Private Cloud.

D. Vertical Cloud or Community Clouds

These clouds are optimized for use by a particular industry sector or a group of users meeting their specific requirements to address issues that are crucial to them. AcademyOne’s Academic Navigator aimed at academics and students and Asite Solution’s applications specifically designed for the construction industry are examples of these types of clouds.

E. Hybrid Clouds

A hybrid cloud is a combination of two or more of the above cloud hosting models. In this model, an enterprise makes use of both public and internal clouds deploying its less critical, low-risk services on public clouds and business-critical core applications on its internal cloud. A hybrid model allows for selective implementation addressing security, compliance and loss of control concerns as well as enabling adoption of public clouds that offer cost benefits and more application options.

F. Choosing Your Cloud

A major decision that IT managers and enterprises have to make is the cloud type - public clouds, internal clouds, or variations of them - that is well suited for their applications. To arrive at a better decision, they have to understand the differences between these deployments, and understand the risks associated with each in the context of characteristics and requirements of their applications. Further, they also have to consider [3]:

- performance requirements, security requirements, and cloud service availability and continuity;
- amount of data transfer between the user and the clouds and/or between the clouds;
- sensitive nature of the applications;
- control of their application and data;
- total cost involved;
- trust on the external cloud providers;
- terms and conditions imposed by the external cloud providers; and
- in-house technical capabilities.

V. Pros and Cons of Cloud Computing

Cloud computing offers several benefits to users – both individual and enterprises. But it also has limitations and poses some risks the impact of which depends on the application type and liabilities involved. In embracing cloud computing, therefore, users must understand, acknowledge, and address these limitations and risks.

A. Benefits of Cloud Computing

Clouds offer the following benefits:

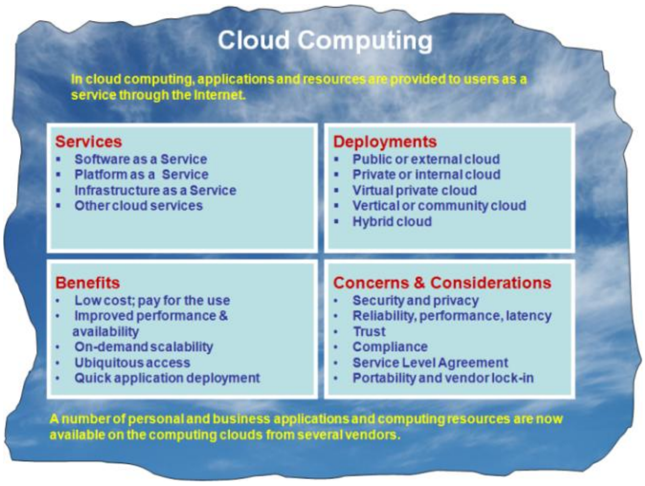
- Lower operational/service cost to users - they pay for what they use
- On-demand scalability to meet peak and uncertain computing demands
- Access to applications from anywhere, anytime, any device
- Shared access to data/application supporting collaboration and teamwork
- Ease of and quicker application deployment
- Freedom from being tied up to a single computer
- Increased data safety than most businesses can afford and manage in their own on-premise IT system

Public clouds eliminate significant capital expenses for hardware and up-front license fees for software as well as the headaches of hardware and software maintenance and upgrade by the users. Cloud applications can be deployed instantly and simultaneously to thousands of users in different locations around the world, and can also be regularly updated easily. Further, as clouds provide improved business continuity and data safety, they are particularly attractive to small and medium-size enterprises as well as enterprises in natural disaster prone areas. Application developers can use computing clouds to try their ideas without having to invest on their own infrastructure.

B. Limitations of Cloud Computing

The general limitations of cloud computing are:

- the need for a constant, high-speed network access to connect to clouds;
- the possibility of slow response at times due to increased traffic on the network or higher load on the computers in the cloud;
- the security of the data on external clouds;
- risks of unauthorized access to users' data; the loss of data due to cloud failure (despite replication across multiple machines); and
- concerns regarding reliability and continued availability of services by cloud providers.



Cloud Computing

In cloud computing, applications and resources are provided to users as a service through the Internet.

Services
- Software as a Service
- Platform as a Service
- Infrastructure as a Service
- Other cloud services

Deployments
- Public or external cloud
- Private or internal cloud
- Virtual private cloud
- Vertical or community cloud
- Hybrid cloud

Benefits
- Low cost; pay for the use
- Improved performance & availability
- On-demand scalability
- Ubiquitous access
- Quick application deployment

Concerns & Considerations
- Security and privacy
- Reliability, performance, latency
- Trust
- Compliance
- Service Level Agreement
- Portability and vendor lock-in

A number of personal and business applications and computing resources are now available on the computing clouds from several vendors.

Fig. 1. A snapshot of cloud computing

For a snapshot of cloud computing, refer to Figure 1. Table 1 highlights key technologies that support cloud computing and features and benefits of cloud computing.

TABLE I
CLOUD COMPUTING PARADIGM

Technology	• Virtualization • Provisioning • Service orientation • Metering services • Load balancing • Web services API • Redundancy and fail over • Security
Features	• Flexibility ▪ On-demand elasticity ▪ Scalability ▪ Access from any device from anywhere, location independence • Self-service IT • Better security
Business Benefits	• On-demand provisioning • Ability to easily meet evolving IT requirements • Immediacy – quick to deploy applications • Effective utilization of resources • Service level contract (SLA)
Economic Benefits	• Utility-based, pay-as-you go model • No long term contract/commitments • No capital expenses

VI. CLOUD APPLICATIONS

The most popular cloud-based applications are Web mail and document management and sharing systems such as Google Docs, calendars, and contact management, Microsoft Office Live Workspace, Apple MobileMe, and Adobe Photoshop Express. Businesses cloud applications include customer relationship management (CRM) tools, payroll processing, human relationship management and VoIP in clouds. Enterprise cloud applications and services are provided by a range of vendors including Salesforce.com, NetSuites and Fonality, EMC, HP, IBM, Oracle, SAP, Accenture, and Infosys. Clouds help smaller companies access applications and services that they couldn’t have access to otherwise but were available only to larger enterprises. One of the cloud-based services used by most internet users is online storage and backup services provided by vendors such as Carbonite, iDrive, Rackspace, and Google. Several businesses of all sizes in different sectors have used clouds for different kinds of applications and derived cost and other benefits.

For example, City of Orlando cut costs by over 66 percent by moving email and application into the cloud [4]. At BP, cloud computing solved a short-term capacity problem that its existing data centre had no cost-effective way of addressing [5]. Lilly used Amazon’s EC2 service cloud to deliver speed and flexibility to its researchers for doing data analysis, rather than investing on 50 servers needed to do the analysis the researchers were interested [5]. *The New York Times* harnessed 100 virtual cloud servers to convert 11 million documents to PDF and archived the documents on the cloud, spending only $240 [6].

Future Group, India’s largest retailer, is using cloud computing to support data warehousing and analytics for its multi-format chain of retail stores. The retailer uses clouds to manage its customer loyalty program which involves storage and analysis of millions gigabytes of data [7].

Major companies now use the cloud to provide affordable specialised IT services to small and medium enterprises. For instance, GE Healthcare has launched a software service that allows physicians to manage the medical side of their practice. Physicians can obtain and share data with other medical practitioners, access decision support tools and prepare and share medical reports about their patients. The service also lets their patients to schedule appointments and request prescriptions and access their laboratory results [8, 10].

Most businesses now use some form of cloud computing. According to Mimecast Cloud Adoption Survey [9], 51 percent of enterprises have already adopted some form of cloud computing. 66 percent of enterprises are now thinking of adopting cloud computing. Companies that used cloud said they were very satisfied with their cloud experiences: 74 percent of users claimed that they were using resources better; 73 percent reported a reduction in infrastructure costs, and 72 percent said that end-user experience had improved.

An Accenture report [10] highlights some of the others business applications of cloud. Government agencies in the US, UK and other countries are leaning toward cloud computing and software-as-a-service solutions as they try to meet the dual goals of containing costs and modernizing technology. The Whitehouse Office of Management and Budget has announced that, from now on, cloud computing would be the "default approach to IT" for US government agencies. The cloud-first policy [11] is expected to reduce over 2000 data centre infrastructure by as much as 40 percent, lowering costs, improving security and performance, and speeding up the deployment of new applications.

VII. Cloud Concerns and Remediation

Despite its promises, cloud computing’s mainstream adoption is constrained by perceived and real barriers and concerns. Security and privacy are two of top concerns in moving into clouds followed by reliability and availability of cloud services, as well as adherence to compliance requirements, where applicable. External clouds raise additional concerns about loss of control and sharing data outside the enterprise firewall, and on potential hike in cost of services by cloud vendors.

Many people think that because they don't know where their data is stored remotely and since the applications are accessed over the Internet, cloud applications are insecure, and believe that if the data and application are physically housed in computers under their control they can protect them better. But this is not necessarily the case as economies of scale allow cloud providers to offer for more sophisticated security, disaster recovery, and service reliability than an individual institution, particularly small enterprise, can deploy on its own. However, recent security breaches at Salesforce.com, epic.org and Google Docs, to name a few, attest the ongoing concerns on cloud security and serve as reminders to be cautious.

Cloud computing security concerns and requirements can differ considerably among the stakeholders – end-user service consumer, cloud service provider and cloud infrastructure provider and are determined by the specific services they provide or consume. The Cloud Security Alliance has identified seven top cloud security threats and outlined impact of those threats as well as remediation for them [12].

Many enterprise computing applications must meet compliance requirements which depend on the type of business and customer base. Although not a guarantee, to better ensure desired level of service delivery and to limit the liabilities, service level agreements (SLAs) with the cloud vendor is highly recommended when consuming cloud services. A cloud SLA specifies terms and conditions as well as expectations and obligations of the cloud service provider and the user. By careful planning and user’s requirements built into service providers’ cloud offerings, both the cloud vendors and users can reduce risk and reap the rewards of cloud-based hosted services.

VIII. Migrating to Clouds

To embrace cloud computing a new mindset is needed. To successfully use and benefit from clouds, an enterprise must prepare itself strategically, culturally and organizationally, and take a holistic view of cloud computing. It must develop its strategic plan and follow a phased, pragmatic step-by-step approach that provides a business context for its cloud adoption. It must choose a cloud option that is appropriate for the application in consideration, as outlined earlier, and manage the risks of migrating to clouds by applying safeguards and provisions available [13]. Moving into clouds is not just about technology, the cloud migration should also factor in the role of people, processes and services, and the change management process. Migration to clouds will also demand a new kind of IT management and governance framework.

IX. Cloud Prospects and Implications

Computing clouds is a powerful change-agent and an enabler. Soon the core competency for most enterprises would be on using IT services and infrastructure that cloud computing offers as hosted services, not building their own IT infrastructure. Cloud computing is already transforming the way we think about computing environments and can drastically improve access to information as well cut IT costs.

Ongoing developments -- increasing maturity of clouds, introduction of new cloud computing platforms and applications, growth in adoption of cloud computing services, and the emergence of open standards for cloud computing - will boost cloud computing's appeal to both cloud providers and users. More personal productivity applications will appear as software services.

Clouds will enable open-source and freelance developers to deploy their applications in the clouds and profit from their developments. As a result, more open source software will be published in clouds. The SaaS model of cloud computing has reached the tipping point and will be widely adopted as an online service. PaaS and IaaS models have matured for mainstream adoption.

Cloud Computing will profoundly change the way people and enterprises use computers and their work practices as well as how companies and governments deploy their computer applications. Clouds will also help close the digital divide prevalent in emerging and underdeveloped economies. They also help save our planet by presenting a greener computing environment.

Driven by economic imperatives and the promise of flexibility and convenience, cloud computing will gain wider acceptance and adoption. Barriers to enterprise adoption of cloud computing will gradually be lowered or removed as IT professionals and the IT industry addresses these issues. Government regulations and compliance requirements will also be amended to embrace cloud computing.

Like the Internet, cloud computing is a transformational technology. It'll mature rapidly as vendors and enterprises come to grip with the opportunities and challenges that it presents. It's time to explore, experiment and embrace cloud computing as relevant.

A. Cloud Implications

The hosted computing model creates opportunities for all IT and non-IT enterprises, researchers and individuals. It creates new possibilities for businesses. There will be new investments that create new business models, new opportunities to start and form businesses. Researchers would be better able to run experiments quickly on clouds, share their data globally, and perform complex analysis and simulations. A whole new set of courses focused on cloud computing would be offered by universities and training institutions.

However, a pertinent question that arises among many IT professionals is: Will cloud computing kill IT jobs? It seems many IT professionals particularly those work on on-premise IT systems are afraid of losing their jobs because of cloud computing. While some might lose their current job, they might get absorbed in other roles. So they should be looking to evolve their roles. They will need to learn about clouds, how to deploy and manage applications on the clouds and minimize risks, and how to work with cloud providers and other functional units within their enterprise. They may also be required to help others in their enterprise understand the opportunities offered by the clouds and the limitations and risks of new model. They may need to master their people skills as cloud computing inevitably brings together different stakeholders and their interests. Also there will be need for people who understand the technical advances in clouds and the cloud-based offerings to spread the cloud around the globe, for developing new kinds of applications, and for creating innovations in IT.

X. Demystifying the Myths of Cloud Computing

There are few myths that have risen due to confusion regarding the cloud. In the following we examine and demystify them [14].

- Myth #1 - Cloud isn't secure
  Ensuring security of their cloud services is always the top priority and competitive necessity for cloud providers. Cloud providers uses a range of security measures and strategies, including physical data centre security, separation of the network, isolation of the server hardware, and isolation of storage that enterprise data centres have used for long, and a lot more including 24x7 monitoring. For instance, cloud providers such as Amazon.com maintains packet-level isolation of network traffic and support industry-standard encryption. Most companies and individuals users don't have the luxury of dedicating enough resources on security like cloud providers. In view of the scale, cloud computing provider do invest in security controls and countermeasures and provide better security than almost any small and medium size company could afford.

- Myth #2 - Cost is the only cloud advantage
  The reality is cost is just one of them, the more important advantage is the ability to move more quickly and accelerate time-to-market. For a developer in an organization to get a server to do an experiment or just expand a project, might take a few months. With a cloud he can get access to large amounts of server capacity in minutes to expedite development work and at very low cost.

- Myth #3 - You should move all infrastructure to the cloud in one step
  For a start-up it is advisable to move all infrastructure to cloud in one go. For enterprises that have new development it is easy to build it on top of the cloud and quickly take advantage of those benefits. Most major enterprises move more methodically by picking a diverse set of initial applications to try as proofs of concepts in the cloud. They run them for a few months to see how the cloud is different and understand how to operate in the cloud before moving more of their applications. Then follow it with a cloud migration plan.
- Myth #4 - One can have all the benefits of the cloud with only a private cloud
  The reality is private or internal clouds incur high capital and ongoing high maintenance costs. Special expertise is needed to implement and maintain a private cloud. Cloud computing allows companies to focus their capital and resources on innovations to accelerate their time to market, rather than running and maintaining the undifferentiating heavy lifting of infrastructure.

XI. Clouds for Emerging Markets

The ability to access software and computing power through public clouds means that many firms will not need to build proprietary systems or purchase expensive hardware. Cloud computing could give emerging market companies an opportunity to leapfrog over their rivals in developed countries. According some estimates IT professionals spend 70% of their time maintaining systems and only 30% of their time creating strategic value. Cloud computing frees the IT department from legacy issues and allows it to focus on innovative ideas to create competitive advantage.

Executives from businesses in developed countries appear more cautious about the opportunities that cloud computing bring to the businesses than their entrepreneurial counterparts in emerging regions. According to a survey, 70% of firms in the developing world are re-appraising their cloud platforms, compared with only 46% of firms in the developed world.

A. *Opportunities to Emerging Markets*

Clouds provide businesses and individuals access to advanced IT infrastructure and applications that many of them couldn't afford otherwise. They can offer productivity gains and better business continuity in the event of failures and natural disasters, besides cutting costs. They can also facilitate creation of 'Connected Enterprises' and fosters innovation supporting creating and deploying new applications on established cloud platforms quickly and cost effectively. There are now a growing number of cloud applications for virtually every aspect of a business' and individual's activities.

Cloud computing promises to speed application deployment, increase innovation, and lower costs. It also can transform the way we design, build, and deliver applications, and increase business agility. One can embrace clouds in a variety of ways to gain significant benefits and to offer applications that were not feasible otherwise. The developing world must exploit the opportunities offered by cloud computing while minimizing the associated risks. Thereby they can have access to advanced IT infrastructure, data centres, and applications, and also leverage the power clouds/IT for their benefit. In fact, developing economies could catch up with developed countries as the cloud gives them access to the same IT infrastructure, data centres, and applications.

Cloud computing could be applied in a range of areas: e-Commerce, E-Business, E-Supply Chain Management; E-Education, E-Health, E-Governance, telework/ telecommuting, Collaboration, Community building and emergency response. Although the potential of cloud computing is vast and compelling, few in the industry can fully grasp its true impact on society. Adoption of computing clouds calls for new mindset: move from P to S – product to services.

B. *Supporting Trends*

Several initiatives and developments support the emergence and adoption of clouds in emerging markets. Multinational companies are triggering the cloud industry's evolution in the developing world by establishing cloud/data centres and promoting clouds. Efforts by governments and international organizations to narrow digital divide and data divide can now better exploit the clouds. Growing middle class and IT-savvy new generations are keen on using IT and don't hesitate to compute in the clouds, like the previous generations.

Mobile Internet and Clouds form a powerful link. Two-thirds of the globe's mobile phone users (4.6 billion) live in emerging markets, and with the growth of smart phones with wireless broadband, mobile Internet and cloud based applications will become more pervasive.

XII. Conclusion

Cloud computing is positioning itself as a new platform for delivering information infrastructures and a range of computer applications for businesses, governments, charities and individuals as IT services. Cloud customers can then provision and deploy these services in a pay-as-you-go fashion and in a convenient way while saving huge capital investment in their own IT infrastructures. Clouds are evoking a high degree of interest – both in developed and emerging markets – though challenges such as security and privacy remains to be fully addressed.

The rise of the cloud is more than just another platform shift. It will transform the IT industry, but it will also profoundly change the way people work and companies operate. It will allow digital technologies to penetrate everywhere and leverage the economy and society particularly in emerging markets.

REFERENCES

[1] "NIST Definition of Cloud Computing," January 2011. http://csrc.nist.gov/publications/drafts/800-145/Draft-SP-800-145_cloud-definition.pdf.

[2] S. Murugesan, "Cloud Computing: A New Paradigm in IT," Cutter Business Intelligence *Executive Report*, Vol. 9, No. 2, 2009.

[3] B. Claybrook, "Cloud vs. in-house: Where to run that app?," *Computer World*, 1 March 2010. http://www.computerworld.com/s/article/9162542/Cloud_vs._in_house_Where_to_run_that_app_?

[4] C. Cross, "City of Orlando: cutting costs while advancing our infrastructure in just two months," March 25, 2010. http://googleenterprise.blogspot.com/2010/03/city-of-orlando-cutting-costs-while.html.

[5] D.Neal, et al., "Cloud rEvolution: A Workbook for Cloud Computing in the Enterprise,", Vol. 4, CSC Leading Edge Forum, 2010. http://assets1.csc.com/lef/downloads/LEF_2010CloudRev_Vol4_Workbook.pdf.

[6] *Introduction to Cloud Computing architecture*, Sun Microsystems, June 2009. http://eresearch.wiki.otago.ac.nz/images/7/75/Cloudcomputing.pdf.

[7] "Future Group Sets the Standard for Retail Analytics with Greenplum Database," PRWeb, 8 June 2009. www.prweb.com/pdfdownload/2505804.pdf.

[8] B.T. Horowitz, "GE SAAS Offering Eases Electronic Medical Records Management," www.eweek.com, June 15, 2010.

[9] "Loudhouse Research: Cloud Barometer Survey 2010." www.mimecast.com/barometerresearch2010.

[10] J.G. Harris and A.E. Alter, *Cloudrise: Rewards and Risks at the Dawn of Cloud Computing, Accenture, Nov 2010.* www.accenture.com/SiteCollectionDocuments/PDF/Accenture_Cloudrise_Rewards_and_Risks_at_the_Dawn_of_Cloud_Computing.pdf.

[11] V. Kundra, *Federal Cloud Computing Strategy*, Feb 2011. www.cio.gov/documents/Federal-Cloud-Computing-Strategy.pdf.

[12] *Top Threats to Cloud Computing V1.0*, Cloud Security Alliance, 2010. http://cloudsecurityalliance.org/topthreats/csathreats.v1.0.pdf

[13] I. Gotts, "Assessing the Risks of Migrating Your Offering to the Cloud," *Cloud Computing Journal*, 2010. http://www.cloudbook.net/resources/stories/assessing-the-risks-of-migrating-your-offering-to-the-cloud.

[14] "Demystifying the Myths of Cloud Computing," 2010. www.datacenterdynamics.com/focus/archive/2011/08/demystifying-the-myths-of-cloud-computing.

Cloud Computing Reliability, Availability and Serviceability (RAS): Issues and Challenges

Farzad Sabahi

***Abstract*—Cloud computing is one of today's most exciting technologies because of its capacity to lessen costs associated with computing while increasing flexibility and scalability for computer processes. During the past few years, cloud computing has grown from being a promising business idea to one of the fastest growing sectors of the IT industry. But on the other hand, IT organizations have expressed concerns about critical issues such as security that accompany the widespread implementation of cloud computing. Security, in particular, is one of the most debated issues in the field of cloud computing and several enterprises look at cloud computing warily due to projected security risks. Also, there are two other issues. They are the reliability and availability of the cloud which are as important as security. Although each of those three issues is associated with usage of the cloud, they will have different degrees of importance. Examination of the benefits and risks of cloud computing is necessary for a full evaluation of the viability of cloud computing.**

This article reviews issues and challenges of cloud computing's reliability, availability and security (RAS). Beginning with a brief discussion on virtualization technology, a key element of cloud infrastructure, it examines issues facing in cloud RAS fields. Then, it addresses the challenges and problems in cloud computing RAS. It also examines intrusion detection methods and outlines counter measures to improve cloud RAS.

***Index Terms*— Cloud computing, Virtualization, Reliability, Availability, Security, Threat, Intrusion, Countermeasure.**

I. Introduction

Cloud computing is based on virtualization technology, in which each user uses a virtual machine. Virtualization technology includes two levels of virtual machines, which are VMs (virtual machine) and hypervisors. The hypervisor has administrative rights to control VMs. But virtualization has some issues that could endanger system performance. From a cloud viewpoint, there are many important dimensions of virtualization technology to consider, but the hypervisor's Reliability, Availability and Serviceability (RAS) is an important aspect of virtualization technology and requires special attention. For example, from security viewpoint, if someone gets control of the hypervisor, he will gain full control of all VM that are under the hypervisor control. Consequently, cloud technology has some problems in RAS that it has inherited from virtualization technology. One such problem involves overflows of system due to excessive combination of VM to a physical server that affects availability and reliability. Because of these issues, cloud systems are vulnerable to traditional attacks as well as new attacks that some of them have migrated from virtualization.

Privacy is another issue which can decrease virtualization and cloud's overall performance, because the VMs are located practically in a multitenant environment, thus making it possible for a user to access a past tenant's information in the same space. Although the use of encryption algorithms could be a good solution for the user or cloud provider by making the appropriate arrangements, such as using advanced algorithms to wipe the user's data for avoidance from information leaks. But the use of encryption algorithms has problems as well, such as the inability of owners to recover their data when they lose the decoding key.

As we know in the world of network computing, there is a variety of attacks that can cause serious problems for Internet-based technologies such as cloud. This can make the cloud vulnerable to some attacks, like the DoS family (Denial of Service) which aims to make the target server inaccessible to legitimate users. The cloud can be a victim of DoS attacks, but it can also be part of the solution by allocating more resources to a user under a DoS attack in order to prevent the user from crashing. Therefore, applying countermeasures to deal with security problems in the cloud is critical, whereas one of the main countermeasures is controlling access control in the cloud. Generally, it seems the security countermeasures in the access control part of the cloud often involve prevention—for example, management of permissions for the account to determine access to different levels of virtualization in the cloud.

Besides security, cloud providers are also responsible for reliability and availability, because all users expect the highest level of QoS (Quality of Service). The cloud providers use some solutions such as partitioning to achieve maximum performance. But according to whether the cloud is based on public, private, or hybrid, the management and control of these performance parameters from RAS viewpoint will vary.

Manuscript received on March 14, 2011. Accepted July 16, 2011.

Farzad Sabahi is with the Azad University, School of Computer and Electrical engineering, Iran. (e-mail: f.sabahi@ieee.org).

This paper is organized as follows:

- Section 2 provides a general overview of cloud computing.
- Section 3 describes the virtualization technology that is the basis of cloud computing.
- Section 4 overviews information security policies in cloud computing.
- Section 5 comprehensively reviews the RAS factor in virtualization.
- Section 6 elaborates on the RAS factor with particular attention to cloud computing.
- Section 7 covers intrusion detection systems in cloud computing.
- Section 8 describes security management and countermeasures to take against intrusions.
- Finally, section 9 concludes the paper.

II. CLOUD COMPUTING: AN OVERVIEW

Cloud computing is a network-based environment that focuses on sharing computations and resources. Clouds are Internet-based and try to reduce complexity for clients by allowing them to virtually store data, applications and technologies at a remote site rather than keeping voluminous amounts of information on personal computers or on local servers. This is accomplished using virtualization technologies in combination with self-service abilities for computing resources via network infrastructure, especially the Internet. In cloud environments, multiple virtual machines are hosted on the same physical server as infrastructure. Customers only pay for what they use and avoid having to pay for local resources such as storage and infrastructure. Cloud computing, then, ultimately refers to both applications delivered as services over the Internet, and the hardware and systems software in the datacenters that provide those services. Currently, three types of cloud environments exist: public, private, and hybrid.

A public cloud is a standard model in which providers make several resources such as applications and storage available to the public. Public cloud services may be free, or may come with an associated fee. In public cloud environments, applications are run externally by large service providers, offering some benefits over private cloud environments.

For a private cloud, a business has internal services that are not available to other people. Essentially, the term “private clouds” is a marketing term for an architecture that provides hosted services for a particular group of people behind a firewall.

A hybrid cloud is an environment in which a company provides and controls some resources internally and provides other services for public use. In this type, the cloud provider has a service whereby a private cloud can be created (and is only accessible by internal staff; it would be protected by firewalls from outside access), and a public cloud environment for access by external users is also created.

Cloud is a style of computing where massively scalable and flexible IT-related abilities are provided “as services” to external customers using Internet technologies. Cloud providers offer various services in a XaaS collection that can offer the following main services.

A. SaaS

SaaS (Software as a Service): Software as a Service is a well-known service that offers network-hosted applications. SaaS is a software application delivery model by which cloud providers develop web-based software applications and then host and operate those applications over the infrastructure (usually the Internet) for use by their customers. As a result, cloud customers do not need to buy software licenses or additional equipment and they typically only pay fees (also referred to as annuity payments) periodically to use the cloud provider’s web-based software [3]. There are two major kinds of SaaS: business applications that offer software which helps various businesses perform their tasks quickly and accurately. Other type of SaaS is development tools which consist of software that is used mainly for product development and management.

B. PaaS

PaaS (Platform as a Service): In this category of service, cloud users are given a platform [4]. They can use it as their application platform independent of using their own local machine for installing those platforms.

C. IaaS

IaaS (Infrastructure as a Service): Infrastructure as a Service is a provision model in which an organization outsources the equipment used to support operations, including storage, hardware, servers and networking components. The service provider owns the equipment and is responsible for housing, running, and maintaining it. The client typically pays on a per-use basis [5]. IaaS is sometimes referred to as Hardware as a Service (HaaS).

D. Other Types of Services

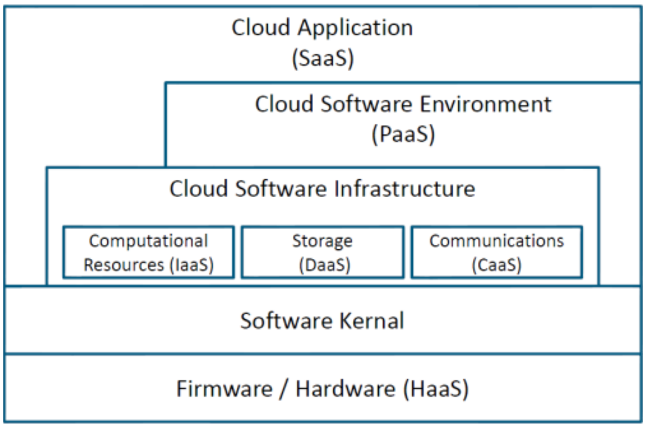

Cloud Application
(SaaS)
Cloud Software Environment
(PaaS)
Cloud Software Infrastructure
Computational Resources (IaaS)
Storage (DaaS)
Communications (CaaS)
Software Kernal
Firmware / Hardware (HaaS)

Fig. 1: Unified ontology of cloud computing.

It is important to mention that IaaS, PaaS, and SaaS are the three main categories of cloud computing services and that the other types of cloud services are subsidiary branches of these three major categories. A typical cloud computing ontology for some of these categories is illustrated in Figure 1.

Other cloud services include the following:

- DaaS (Database as a Service): Database systems provide a user friendly interface for accessing and managing data. This type of service is very useful like many financial, business, and Internet-based applications [6].
- NaaS (Network as a Service): With NaaS, providers offer customers a virtualized network [7].
- IPMaaS (Identity and Policy Management as a Service): With this service, providers deliver identity and policy management to customers [8].

E. Cloud customers

Nowadays, many IT-related clients decide to use cloud computing for their own purposes. These can be divided into three main groups: regular customers, academics, and enterprises.

1) Regular customers

This group of users merely uses the services from the cloud [1]. They are not concerned with high performance; rather, they concentrate on the service and the privacy of their data on the cloud. SaaS is the most appropriate service for this group [9].

2) Academics

Academics usually have good networks and they often prefer to use the infrastructure that they already have to improve the performance of computations and resolve grid limits. For this group, cloud computing provides convenient access to a high-performance cluster or grid-based computation infrastructure and eliminates the need to buy new hardware.

3) Enterprises

The IT industry reaps the most considerable benefits of cloud computing [10]. Many companies have decided to enter cloud-related industries or use cloud services to reduce costs and improve performance in their own (IT-related or non-IT-related) businesses.

a) Small and mid-size enterprises

Lower costs are attractive, particularly for small enterprises that simply cannot afford the cost of solutions [4]. With distributed processing, small enterprises can afford industry-standard PCs and network servers but not expensive supercomputers. In addition, they can use cloud software instead of local software or abstruse infrastructure, which can reduce the cost of purchasing and maintaining the required software. For mid-size businesses that are growing, cloud computing can also provide a cost-effective and efficient path to enterprise-grade software and infrastructure [11].

b) Large-scale enterprises

For these enterprises, lower costs are not as important as privacy. Thus, large companies often create their own clouds or are skeptical about moving to the cloud. However, privacy of information is the most important issue, and most large companies have already spent significant amounts of money on their local systems [1]. Nowadays, large-scale enterprises often collect and analyze large amounts of data to derive business insights. However, there are at least two challenges to meet the increasing demand. First, the growth in the amount of data far surpasses the growth in the computation power of uniprocessors [12]. The growing gap between the supply and demand of computation power forces enterprises to parallelize their application codes. Unfortunately, parallel programming is both time-consuming and error-prone. Second, the emerging cloud computing pattern imposes constraints on the underlying infrastructure, which forces enterprises to rethink their application architectures.

III. VIRTUALIZATION

Virtualization is one of the most important elements of cloud computing. It is a technology that helps IT organizations optimizes their application performance in a cost-effective manner, but it can also present application delivery challenges that cause security difficulties. Most of the current interest in virtualization revolves around virtual servers, in part because virtualizing servers can result in significant cost savings. The phrase virtual machine refers to a software computer that, like a physical computer, runs an operating system and applications. An operating system on a virtual machine is called a guest operating system. A layer called a VMM (virtual machine monitor), or hypervisor, creates and controls the virtual machine's other virtual systems. Figure 2 illustrates a typical virtual machine architecture foundation in a cloud environment.

A. Hypervisor

A hypervisor (see Figure 2) is one of many virtualization techniques allowing multiple operating systems, termed guests, to run concurrently on a host computer using a feature called hardware virtualization. It is so named because it is conceptually one level higher than a supervisor.

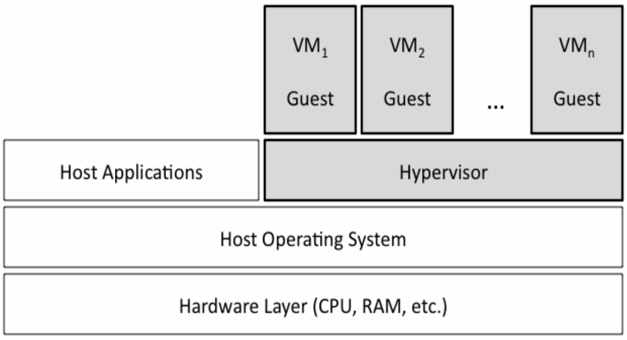


Fig. 2. Typical Virtual Machine architecture [1].

The hypervisor presents a virtual operating platform to the guest operating systems and also monitors the execution of them. Multiple instances of a variety of operating systems may share the virtualized hardware resources. Hypervisors are installed on server hardware dedicated to run guest operating systems [13].

IV. Information Security Policies

Cloud computing raises a range of important policy issues which include issues of privacy, security, anonymity, telecommunications capacity, government surveillance, reliability, and liability, among others [1]. But the most important of these issues is; security and how it is assured by the cloud provider. In addition, according to this fact that security effect on computing performance, cloud providers have to find a way to combine security and performance. For example for enterprises, the most important problem is security and privacy because they may store their sensitive data in cloud. For them, high performance processing may not be as critical as for academia users. To satisfy enterprise needs, the cloud provider has to ensure robust security and privacy more than other needs.

In cloud there are several security and privacy issues but in [14] there are the Gartner’s seven well-known security issues which cloud clients should advert are listed below:

- **Privileged user access:** Sensitive data processed outside the enterprise brings with it an inherent level of risk because outsourced services bypass the "physical, logical and personnel controls" IT shops exert over in-house programs.
- **Regulatory compliance**: Customers are ultimately responsible for the security and integrity of their own data, even when it is held by a service provider. Traditional service providers are subjected to external audits and security certifications. Cloud computing providers who refuse to undergo this scrutiny are "signaling that customers can only use them for the most trivial functions," according to Gartner.
- **Data location**: When clients use the cloud, they probably will not know exactly where their data is hosted. Distributed data storage is usually used by cloud providers, but this can cause lack of control and is not good for customers who have their data in a local machine before moving to the cloud.
- **Data segregation:** Data in the cloud typically exists in a shared environment alongside data from other customers. Encryption is effective but is not a cure-all. Encryption and decryption is a classic way to cover security issues, but heretofore it could not ensure a perfect solution. While it is difficult to assure data segregation, customers must review the selected cloud’s architecture to ensure data segregation is properly designed and available but without data leakage. Although data leakage has solution technology that named DLP.
- **Recovery:** If a failure occurs with the cloud, it is critical to completely restore client data. As clients prefer not to let a third-party control their data, this will cause an impasse in security policy in these challenging situations.
- **Investigative support:** Cloud services are especially difficult to investigate because logging and data for multiple customers may be co-located and spread across an ever-changing set of hosts and data centers.
- **Long-term viability:** Ideally, a cloud computing provider will never go bankrupt or be acquired by a larger company with new policies. However, clients must be sure that their data will remain available even after such an event.

V. Virtualization RAS Issues

In a traditional environment consisting of physical servers connected by a physical switch, IT organizations can get detailed management information about the traffic between the servers and the physical switch. Unfortunately, that level of information management is not typically provided by a virtual switch. In such a scenario the virtual switch has links from the physical switch via the physical NIC (Network Interface Card) attached to virtual machines. The resultant is lack of visibility into the traffic flows between and among the Virtual Machines on the same physical level affects security and performance surveying.

A potential problem also exists for virtualization when a provider combines too many Virtual Machines onto a physical server. This can result in performance problems caused by impact factors such as limited CPU cycles or I/O bottlenecks [15]. These problems can occur in a traditional physical server, but they are more likely to occur in a virtualized server because a single physical server is connected to multiple Virtual Machines all competing for critical resources.

Therefore, management tasks such as performance management and capacity planning management are more critical in a virtualized environment than in a similar physical environment. This means that IT organizations must be able to continuously monitor the real-time utilization of both physical servers and Virtual Machines. This capability allows users to avoid both over and underutilization of server resources. In addition, they will able to reallocate resources based on changing business requirements. This capability also enables IT organizations to implement policy-based remediation that helps them to ensure that their desired service levels are being met [16].

Another challenge with virtualization is cloud organization management of virtual machines sprawl [17]. In virtualized environment with virtual machine Sprawl, the number of virtual Machines running in it increases because of unnecessary new virtual Machines created rather than business necessity. Virtual machine sprawl concerns include the overuse of infrastructure. To prevent Virtual machine sprawl, a Virtual machine manager should carefully analyze the need for all new Virtual Machines and ensure that unnecessary Virtual machines migrate to other physical servers. In addition, by migration, an unnecessary virtual machine will be able to move from one physical server to another with high availability and energy efficiency. Determination of the virtual machine destination can be challenging; it is necessary to ensure that a migrated Virtual machine keeps the same security, QoS configurations and needed privacy policies. On the other hand, the destination must assure that all the required configurations of the migrated virtual machine are kept.

A. Virtual machine security and threats

As illustrated in Figure 2, there are at least two levels of virtualization which are virtual machines and the hypervisor. Virtualization technique which is used in the virtual machines is not as new technology. Unfortunately, it has several security issues which are now migrated to cloud technology and they are not good heritage for cloud. There are also other vulnerabilities and security issues which are exclusive or may have a more critical role in the cloud environment.

As mentioned before, in the hypervisor, all users see their systems as self-contained computers isolated from other users, even though every user is served by the same machine. In this context, a virtual machine is an operating system that is managed by an underlying control program.

Hence there are various threats and attacks in this level, but some of them are important than others that mentioned below:

- **Virtual machine-level attacks:** The hypervisor or virtual machine technology used by cloud vendors are potential problems in multi-tenant architectures [18]. These technologies involve “virtual machines,” remote versions of traditional on-site computer systems, including the hardware and operating system. The number of these virtual machines can be expanded or contracted on the fly to meet demand, creating tremendous efficiencies [19].
- **Cloud-provider vulnerabilities:** These could be platform-level vulnerabilities, such as SQL-injections, or cross-site scripting vulnerabilities that exist in the cloud service layer and cause insecure environments.
- **Expanded network-attack surface:** The cloud user must protect the infrastructure used to connect and interact with the cloud, a task complicated by the cloud being outside the firewall in many cases [4].
- **Authentication and authorization:** The enterprise authentication and authorization framework does not naturally extend into the cloud. Enterprises have to merge cloud security policies with their own security metrics and policies.
- **Availability of the cloud provider:** Cloud providers guarantee that their servers’ uptime compares well with cloud users’ own data centers and cloud providers ensure the clients which providers can handle their applications. An enterprise must be assured that a cloud provider is faithfully running a hosted application and delivering valid results [4]. Scheduled and unscheduled maintenance is another availability factor that exists and it can harm the availability ratio of the cloud provider. Although regularly scheduled maintenance does not count as downtime, unscheduled maintenance increases downtime and affects availability [20].
- **Lock-in:** There seems to be a great deal of anxiety regarding lock-in in cloud computing. The cloud provider can encrypt user data in a particular format if a user decides to migrate to another vendor or a similar situation arises [21].
- **Data control in cloud:** For mid-size businesses used to having complete visibility and control over their entire IT portfolio, moving even some components into the cloud can create operational “blind spots,” with little advance warning of degraded or interrupted service [11].

B. Hypervisor Security

In a virtualization environment, there are several virtual machines that have independent security zones that are not accessible from other virtual machines that have their own zones. In a virtualization environment, a hypervisor has its own security zone and is the controlling agent for everything within the virtualization host. A hypervisor can touch and affect all of the virtual machine’s actions running within the virtualization host [22]. There are multiple security zones, but these security zones exist within the same physical infrastructure that, in a more traditional sense, generally only exists within a single security zone. This can cause security issues, as if an attacker is able to take control of a hypervisor, then the attacker has full control of all the works within the territory of the hypervisor. Another major virtualization security concern is “escaping the virtual machine” or being able to reach the hypervisor at the virtual-machine level. This will become an even greater concern in the future as more APIs (Application Program Interface) are created for virtualization platforms [23]. Thereupon, so undamaged controls are to disable the functionality within a virtual machine, and this can reduce performance and availability.

1) Confronting against hypervisor security problems

As mentioned before, hypervisors are management tools, and the main goal of creating this security zone is building a trust zone. Other available virtual machines are under the approval of the hypervisor, and they can rely on it, as users are trusting that administrators of system will do what they can to do tasks properly. As for security characteristics, there are three major levels in the security management of hypervisors:

- **Authentication:** Users have to authenticate their account properly using the appropriate standard and available mechanisms.
- **Authorization:** Users must receive authorization, and they must have permission to do what they are trying to do.
- **Networking:** Using mechanisms that assure a secure connection to communicate by using available

administration applications that most likely launch and work in a different security zone than that of users.

Authentication and authorization are some of the most interesting auditing aspects of management because there are so many methods available to manage a virtual host auditing purpose [24]. The general belief is that networking is the most important issue in transactions between users and the hypervisor, but there is much more to virtualization security than just networking. Networking plays a critical role in security, but it is not solely significant for ensuring security. It is just as important to understand the APIs and basic concepts of available hypervisors and virtual machines, and how those management tools work [22]. If a security manager can address authentication, authorization, virtual hardware, and hypervisor security as well as networking security, cloud clients are well on the way to a comprehensive security policy [1, 22]. If a cloud provider at the virtualization level does not, or just depends on network security to do the tasks, then the implemented virtual environment is at risk and has poor security capability. It is a waste of money if a cloud provider spends too much money on creating a robust secure network and neglects communication among virtual machines and the hypervisor, as this can cause several problems for the provider as well as for the users.

VI. Cloud RAS Issues

Using cloud means, that applications and data will move under a third-party control. The cloud services delivery model will create clouds with virtual perimeters as well as a security model with responsibilities shared between the customer and the cloud service provider. This shared-responsibility model will bring new security management challenges to the organization’s IT operations staff [25]. Basically, the first question an information security officer must answer is whether he/she has adequate transparency with cloud services to manage the governance (shared responsibilities) and implementation of security management processes (preventive and detective controls) to ensure the business that the data in the cloud is appropriately protected. The answer to this question consists of two parts: what security controls must the customer provide over and above the controls inherent in the cloud platform and how should an enterprise’s security management tools and processes adapt to manage security in the cloud. Both answers must be continually reevaluated based on the sensitivity of the data and the service-level changes over time [25].

A. Data Leakage

Basically, when moving to a cloud, there are two changes for customers’ data. First, the data will be stored away from the customer's locale machine. Second, the data is moved from a single-tenant to a multitenant environment. These changes may raise an important concern called, *data leakage*. This has become one of the greatest organizational risks from the security standpoint [26]. Virtually every government worldwide has regulations that mandate protections for certain data types [26]. The cloud provider should have the ability to map its policy to the security mandate the user must comply with and discuss the issues.

1) DLP

Nowadays, there is an interest in the use of data leakage prevention (DLP) applications to protect sensitive data with the appearance of cloud computing. To prevent data leakage, some companies have thought of DLP products. DLP products existed before cloud computing. These products aim to ensure data confidentiality and detect unauthorized access to data, but they are not intended to be used for ensuring the integrity or availability of data. As a result, experts don’t expect from DLP products to address data’s integrity or availability in any cloud model.

If data is stored in a public cloud, because of its nature, using DLP products is worthless to protect the confidentiality of that data in all types of clouds. Generally in SaaS and PaaS, because cloud clients do not have control over security management used by the cloud provider, discovery of the client’s data with DLP agents is not possible except when the provider puts this capacity into its service. However, it is possible by embedding DLP agents into virtual machines in IaaS to achieve some control over associated data.

In private clouds, the customer has direct control over the whole infrastructure; it is not a policy issue whether DLP agents are deployed in connection with SaaS, PaaS, or IaaS services. However, it may very well be a technical issue whether DLP agents interoperate with SaaS or PaaS services as architected [27]. In a hybrid cloud, if service is IaaS, the client could embed DLP agents for some control over data.

B. Privacy

Cloud clients’ data stores in data centers that cloud providers diffuse all over the globe within hundreds of servers that communicate through the Internet have several well-known potential risks within them. Because cloud services are using the Internet as their communication infrastructure, cloud computing involves several kinds of security risks [26]. Cloud providers, especially IaaS providers, offer their customers the illusion of unlimited computer, network, and storage capacity, often coupled with a frictionless registration process that allows anyone to begin using cloud services [28]. The relative anonymity of these usage models encourages spammers, malicious users and other hackers, who have been able to conduct their activities with relative impunity [29]. PaaS providers have traditionally suffered most from such attacks; however, recent evidence shows the hackers have begun to target IaaS vendors as well [28].

As is clear in cloud-based services, a user's data is stored on the third-party’s storage location [1]. A service provider must implement sufficient security measures to ensure data privacy. Generally, data encryption is a solution to ensure the privacy of the data in the databases against malicious attacks. Therefore, encryption methods have significant performance implications on query processing in clouds. Integration of data encryption with data is useful to protect the user's data against outside malicious attacks and to restrict the liability of the service provider.

It seems protection from malicious users who might access the service provider's system is the final goal, but this is not enough when clients also prefer privacy protection from a accessing to their data by provider. Any data privacy solution will have to use particular encryption, but this causes another availability issue: data recovery [30]. Assume a user's data is encrypted with a user-known key, and the user loses his/her key. How can the provider recover his/her data when it doesn't know what the key is? If the user gives the provider authority to know the key, then this makes privacy by using a user-known encryption key useless. The simple way to solve this problem is to find a cloud provider which users can trust. This way is acceptable when data stored in the cloud is not very important. This method seems useful for enterprises with the maximum size of a small company which may decide to find trustable providers rather than finding a solution for the data recovery problem. For medium-sized companies to large-sized companies, it is more critical to develop techniques and methods that enable query processing directly over encrypted data to ensure privacy from cloud providers [30]. If the service providers themselves are not trusted, protecting the privacy of users' data is a much more challenging issue. However, for those companies it seems using a private cloud is a wise solution.

If data encryption is used as a wise solution for the data privacy problem, there are other issues in this context. One of the most important issues is ensuring the integrity of the data. Both malicious and non-malicious users can cause compromise of the integrity of the users' data when this happens and the client does not have any mechanism to analyze the integrity of the original data. Hence, new techniques have to be applied to provide methods to check the integrity of users' data hosted at the service provider side [8].

All encryption methods rely on secure and impressive key management architectures. One of the problems that can occur in an encrypted environment is encryption key management in the cloud. In the cloud environment several users may use their own encryption method, and managing these keys is another issue to address in the context of encrypted data. For example, if the cloud provides database service (DaaS), the cloud provider faces more challenges in key management architectures, such as generation, registration, storage, and update of encryption keys.

1) RAS issues in Database-based service: An example

Cloud systems provide an extremely attractive interface for managing and accessing data and have proven to be widely successful in many financial, business and Internet applications. However, they have several serious limitations in database-based service such as the following which are mentioned in [6]:

- **Database systems are difficult to scale:** Most database systems have hard limits beyond which they do not easily scale. Once users reach these scalability limits, time-consuming and expensive manual partitioning, data migration and load balancing are the only recourse.
- **Database systems are difficult to configure and maintain:** Administrative costs can easily account to a significant fraction of the total cost of ownership of a database system. Furthermore, it is extremely difficult for untrained professionals to get good performance out of most commercial systems.
- **Diversification in available systems complicates selection:** The rise of specialized database systems for specific markets complicates system selection, especially for customers whose workloads do not neatly fall into one category.
- **Peak provisioning leads to unnecessary costs:** Database workloads are often tandem in nature hence they provision for the peak often results in an excess of resources during off-peak phases and thus causes unnecessary costs.

C. Data Remanance

Data remanence is the residual physical representation of data that has been in some way erased. After storage media is erased, there may be some physical characteristics that allow data to be reconstructed [31]. As a result, any critical data must not only be protected against unauthorized access, but also it is very important that it is securely erased at the end of the data life cycle. Basically, IT organizations that have full control of their own servers use various available tools that give them the ability to destroy unwanted and important data for privacy and safety purposes. But when data is migrated to a cloud environment, they now have virtual servers that are controlled by a third party.

As a solution, IT organizations must choose cloud providers that can guarantee that all customer erased data is erased immediately and securely. A traditional solution to deleting data securely is overwriting, but this technique does not work without the collaboration of the cloud provider [4, 30]. In a cloud environment, customers can't access the physical device or the data level. Thus, there is only one solution: those customers encrypt their data with a confidential key that prevents reconstruction of the erased data from residual data.

D. Cloud Security Issues

As mentioned before, the Internet is the communication

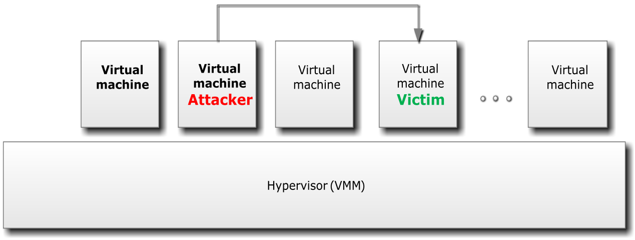


Fig. 3. Attack scenario within cloud.

infrastructure for cloud providers that use the well-known TCP/IP protocol, which uses IP addresses to identify Internet users. Similar to a physical computer in the Internet which has an IP address, a virtual machine in the Internet also has an IP address. A malicious user, whether internal or external, like a legal user who exists in network, can find these IP addresses as well. In this case, a malicious user can find out which physical servers the victim is using, and implant a malicious virtual machine at that location from which to launch an attack [28]. Because all users use the same infrastructure as the virtual machine, if a hacker steals a virtual machine or takes control of it, he also inherits the data within it. The hacker can then copy the data into his/her local machine before the cloud provider detects that the virtual machine is out of control; then the hacker can analyze the data, and may find valuable data afterward.

1) Attacks in cloud

Nowadays, there are several kinds of attacks in the IT world. Basically, the cloud can give service to legal users, but it can also give service to users who have malicious purposes. A hacker can use a cloud to host a malicious application to achieve a task, which may be a DDoS (Distributed Denial of Service) attack against the cloud itself, or arranging an attack against another user in the cloud. For example, an attacker knows that his victim [30]. This situation is similar to this scenario in that both the attacker and the victim are in the same network, but with the difference that they use virtual machines instead of a physical network (Figure 3).

a) DDoS Attacks Against Cloud

DDoS attacks typically focus a high quantity of IP packets at specific network entry elements; usually any form of hardware that operates on a Blacklist pattern is quickly overrun. In cloud computing, where the infrastructure is shared by a large number of clients, DDoS attacks have the potential of much greater impact than they do against single-tenant architectures. If the cloud does not have sufficient resources to provide services to its customers, the cause may be undesirable DDoS attacks [30]. The traditional solution for this event is to increase the number of such critical resources. But a serious problem occurs when a malicious user deliberately performs a DDoS attack using bot-nets.

Most network countermeasures cannot protect against DDoS attacks, because they cannot stop the deluge of traffic, and typically cannot distinguish good traffic from bad traffic. IPS (Intrusion Prevention Systems) are effective if the attacks are identified and have pre-existing signatures, but are ineffective if there is legitimate content with bad intentions [27]. Unfortunately, similar to IPS solutions, firewalls are vulnerable and inefficient against DDoS attacks because an attacker can easily bypass firewalls and IPSs, because they are designed to transmit legitimate traffic, and attacks generate so much legitimate like traffic from so many distinct hosts that a server, or a cloud's Internet connection, cannot handle the traffic [27].

It may be more accurate to say that DDoS protection is part of the Network Virtualization layer rather than Server Virtualization. For example, cloud systems that use virtual machines can be overcome by ARP spoofing at the network layer; DDoS protection is really about how to layer security across multivendor networks, firewalls, and load balances [32].

b) Cloud against DDoS attacks

DDoS attacks are one of the most powerful threats available in the world, especially when launched from a botnet with huge numbers of zombie machines. When a DDoS attack is launched, it sends a heavy flood of packets at a Web server from multiple sources. The cloud may be part of the solution; it's interesting to consider that websites experiencing DDoS attacks, which have limitations in server resources, can take advantage of using a cloud that provides more resources to tolerate such attacks [30]. Cloud technology also offers the benefit of flexibility, with the ability to provide resources almost real-time as necessary and almost instantaneously to avoid site shutdown.

VII. Intrusion Detection in Cloud

As we know, IDS have been used widely to detect malicious behaviors in several types of network. IDS management is an important capability for distributed IDS solutions, which makes it possible to integrate and handle different types of sensors or collect and synthesize alerts generated from multiple hosts located in the distributed environment. Facing new application scenarios in Cloud Computing, the IDS approaches yield several problems since the operator of the IDS should be the user, not the administrator of the Cloud infrastructure. Extensibility, efficient management and compatibility with virtualization-based contexts need to be introduced into many existing IDS implementations. Additionally, the Cloud providers need to enable possibilities to deploy and configure IDS for the user. Within this paper, we summarize several requirements for deploying IDS in the Cloud and propose an extensible IDS architecture that is easily used in a distributed cloud infrastructure [33, 34].

A. Intrusion detection at service level

1) IDS in SaaS

Attacks on networks are a reality in the world. Detecting and responding to those attacks is considered due diligence. The reality is that in SaaS, users will have no choice except to trust their providers to perform Intrusion Detection properly. Some providers give their users the option of getting some system logs and users can use custom application for monitoring those data, but in reality, most Intrusion Detection activities must be done by the provider and the user can only report suspicious behavior for analysis.

2) IDS in PaaS

In PaaS, similarly to SaaS, most of the Intrusion Detection activities must be done by the Cloud provider but with a little difference. If Intrusion Detection systems are outside of the users' application, they have no choice and must rely on the provider to implement IDS. But PaaS configuration is more flexible than SaaS and users may have the choice to configure the security parameters of platforms that log on to a centralized place and users can incorporate Intrusion Detection performance [34, 35].

3) IDS in IaaS

IaaS is the most flexible service for Intrusion Detection implementation. But the most important challenge in constructing a secure cloud-computing infrastructure is Transparency. Without it, the user cannot know if the cloud provider meets significant security requirements or not. Moreover, the user cannot properly design application architecture to mitigate any risks that may exist.

B. Intrusion Detection Placement

For operating Intrusion Detection in the Cloud properly, the user must identify the possible and also proper places for hosting IDS. In the traditional network, using Intrusion Detection allows the user to monitor, detect and alert about traffic that passes over the traditional network infrastructure [9, 36]. Generally, there are some places in the network with more traffic than in other place (hotspots). Placement of IDS in the physical network part of the Cloud is similar to a traditional network, because hotspots in both of them are the same.

1) In the virtual machine and network layer

Using Intrusion Detection in the virtual machine layer allows the user to monitor the system and detect and alert about issues that may arise. In addition, using Intrusion Detection to monitor the virtual network allows the user to monitor the network traffic between the virtual machines on the host, as well as the traffic between the virtual machines and the host. It should be noted that this network is different from traditional networks and that traffic never hits it [35].

2) In the Hypervisor layer

As said before, the hypervisor presents to the guest operating systems a virtual operating platform and monitors how the guest operating systems are running [8]. Deploying Intrusion Detection in the hypervisor allows the user to monitor everything that passes between the virtual machines.

As illustrated in Figure 4, the HyIDS runs inside the hypervisor. Because the hypervisor interposes on all accesses between the guest kernel and the hardware, ISIS can monitor all operating system events and data structures for intrusions.

C. Intrusion Detection Techniques Performance

As is well known, Intrusion Detection has three well-known main groups: Host-based, Network-based, and hybrid. This section discusses performance issues of Intrusion Detection techniques in the Cloud, some of which are traditional solutions and some of which are special rectification solutions for use within the Cloud.

1) Traditional IDS solutions in cloud

a) Host-based intrusion detection

The first choice in Intrusion Detection is the traditional HIDS (Host-based Intrusion Detection), which examines events and transmissions such as what file was accessed and what application was executed. This type of IDS can be used on virtual machines as well as in the hypervisor level of Cloud environments. Using Intrusion Detection in the virtual machine layer allows the user to monitor the activity of the system and detect and alert about issues that may arise. At this level, the user can use an HIDS and have control over it. This type of IDS can detect intrusion against his/her Virtual machine. The provider may also deploy an HIDS in the hypervisor layer but only the provider is authorized to manage and configure it. In the hypervisor level, HIDS can also monitor traffic between virtual machines.

The HIDS on the virtual machine would be used by the user of the Cloud but the HIDS on the hypervisor level is for provider control; if the user wants to use the hypervisor Intrusion Detection data in his independent IDS, he would have to coordinate with the provider. This issue is likely to pose difficulties because most Cloud providers prefer not to share such data with customers due to privacy policies [26, 34]. While HIDS on the hypervisor level would be under the responsibility of the Cloud provider, deploying and managing an HIDS on the virtual machine would be the user's responsibility.

b)Network-based intrusion detection

A Network Intrusion Detection System (NIDS) is another traditional solution for performing security policies in computer networks. NIDSs work by examining network traffic but with this characteristic, only the cloud provider can deploy it. Unfortunately, in cloud, because of the nature of NIDS, this type of IDS has limitations. For example, it is unable to detect attacks within a virtual network that runs completely within the hypervisor. Also, NIDS is useless in encrypted environments. This type of placement of IDS is useful in detecting some attacks on the VMs and hypervisor but it does have three important constraints. The first is that it is not useful when it comes to malicious activities within a VM, which is fulfilled completely in the hypervisor level. Secondly, it has limited visibility into the host itself. Thirdly, if the network traffic is encrypted by users, NIDS cannot decrypt the traffic for analysis [20, 34]. Even if NIDS has all encryption keys used in the Cloud, NIDS needs more computation resources to perform the decrypting. Moreover, analyzing these data results in an increased cost of detection.

2) Performance of traditional IDS

It seems that NIDS works better than HIDS but it must be considered that HIDSs are easy to implement while NIDS are difficult or at times impossible to fulfill in the Cloud environment. In addition, in the Cloud, NIDS falls completely into the area of the provider to operate and control. This paper has shown that Cloud users need to think more about moving toward the Cloud and also that Cloud providers should give more attention to security matters.

3) Hypervisor-based intrusion detection system

Another Intrusion Detection method is to use IDS, which launches at the hypervisor layer but is not strictly a HIDS for the hypervisor, which is called Hypervisor-based IDS (HyIDS) [34] or ISIS IDS (Intrusion Sensing and Introspection System) [36]. One of the promising technologies in this method is the use of VM introspection. This type of IDS allows users to monitor and analyze communications between VMs, between the hypervisor and VM and within the hypervisor-based virtual network. The advantage of the hypervisor-based ID is the availability of information, as it can see everything. The disadvantage is that the technology is new and users really need to know what they are looking for [21, 34]. There is a special type of Intrusion Detection in the hypervisor because of the level of accessing it contains and it has a good potential for improving the performance of intrusion detecting. As illustrated in Figure 4, the HyIDS runs inside the hypervisor. The hypervisor can interpose on all accesses between the guest VM kernel and the hardware, while HyIDS can monitor all operating system events and data structures for intrusions. Like NIDS, control and implementation of HyIDS is done entirely by the Cloud provider [34, 37].

VIII. COUNTERMEASURES

There are several traditional solutions to mitigate security problems that exist in the Internet environment and the cloud infrastructure, but the nature of clouds causes some security problems that exist especially in cloud environments. On the other hand, there are traditional countermeasures against popular Internet security problems that may be usable in clouds, but some of them must be improved or changed to use in cloud environments.

A. Access Control

To ensure the accessibility of authorized users the prevention of unauthorized access to information systems, formal procedures should be in place to control the allocation of access rights to services. The procedures should cover all stages in the lifecycle of user access, from the initial registration of new users to the final de-registration of users who no longer require access to information systems and services. Special attention should be paid, where appropriate, to the necessity to control the allocation of privileged access rights, which allow users to override system controls [38]. The following are the six control statements [38]:

- Control access to information.
- Manage user access rights.
- Encourage good access practices.
- Control access to network services.
- Control access to operating systems.
- Control access to applications and systems.

In the SaaS model, the cloud provider is responsible for managing all aspects of the network, server, and application infrastructure. In that model, since the application is delivered as a service to end users, usually via a web browser, network-based controls are becoming less relevant and are augmented or superseded by user access controls, e.g., authentication using a one-time password [27, 38]. Hence, customers should focus on user access controls (authentication, federation, privilege management, provisioning, etc.) to protect the information hosted by SaaS [39].

In the PaaS delivery model, the Cloud provider is responsible for managing access control to the network, servers, and application platform infrastructure. However, the customer is responsible for access control to the applications deployed on a PaaS platform. Access control to applications manifests as end user access management, which includes provisioning and authentication of users [28].

IaaS customers are entirely responsible for managing all aspects of access control to their resources in the cloud. Access to the virtual servers, virtual network, virtual storage, and applications hosted on an IaaS platform will have to be designed and managed by the customer. In an IaaS delivery model, access control management falls into one of the following two categories. Access control management to the host, network, and management applications that are owned and managed by the Cloud provider and user must manage access control to his/her virtual server, virtual storage, virtual networks, and applications hosted on virtual servers [30, 38].

B. Incident Countermeasure and response

One of the important issues in cloud security, similar to other IT fields, is finding problems and vulnerabilities that exist, but a more important issue is that the cloud provider has appropriate responses against all problems that it finds. Basically, the cloud systems are built on a collection of storage and process engines, driven by a configurable distributed transaction coordinator. To achieve some important parameters such as flexibility, scalability and efficient usage of resources, cloud providers must face major challenges in the area of adaptability and workload.

One of the main requirements of the cloud is the ability to be flexible; in the context of a cloud service, flexibility means dedicating resources where they are most needed [6]. This is particularly challenging in a database environment where there are large amounts of data that may need to be moved in order to reconcile data [6].

To allow high performance workloads to scale across multiple computing nodes, it is important for cloud provider to divide their data into partitions that maximize service performance. The main idea behind partitioning is to lessen the probability that a typical transaction has to access multiple nodes in cloud to compute its query.

In migration, available methods must be able to predict adaptation time and try to avoid cloud node overload by some procedure, such as partitioning, fragmenting, breaking big data packets in smaller pieces, and maintaining the ability to execute transactions while movement occurs [36].

To balance workloads on virtual machines properly, it is necessary to analyze and classify cloud providers resource requirements to decide how these can be allocated to virtual machines.

C. Security Management in the Cloud

The relevance of various security management functions available for each cloud delivery model is dependent on the context of deployment models. As mentioned before in the introduction, there are several important parameters in cloud security management: availability management, access control management, vulnerability and problem management, patch and configuration management, countermeasure response, and cloud system use and access monitoring. Thus, according to the type of service provided, the customer or the provider must manage some or all of them independently, or perhaps partially [30]. Thus, if a cloud is a private cloud, then the cloud provider generally manages all mentioned functions. But if a cloud is a public or hybrid cloud, then who manages which aspect depends on the type of cloud and the service provided. For example, if a cloud is SaaS, then the customer must partially manage access control and monitor system use and access, and also must manage incident response, and the cloud provider must manage the other functions. In other types of clouds (PaaS and IaaS), the functions are limited to customer applications deployed in PaaS or IaaS.

IX. Conclusion

As outlined in the article, cloud computing helps IT enterprises to optimize and secure application performance in a cost effective manner. Cloud-based applications are based on network software running on a virtual machine in a virtualized environment. In view of the vital role of the hypervisor in a virtualization system, security at this level of virtualization needs special consideration. Generally, a virtual application relieves some of the management issues in enterprises because most of the maintenance, software updates, configuration and other management tasks are automated and centralized at the cloud provider's datacenter. But this way for decentralized application and access creates its own set of challenges and security problems. There are, however, risks and hidden costs in managing cloud compliance. Cloud providers often have several powerful servers and resources that provide appropriate services for their users, but the cloud is at risk to a degree similar to that of other Internet-based technologies. Unfortunately, there are some attacks for which no perfect defense exists such as a powerful DoS attack. But as paper discussed in occurrence of DoS attacks, cloud may be a good solution or mitigation because cloud providers can use mirrors or devote more resources to protecting against attacks. However this solution's performance depends on provider facilities.

Issues introduced by this paper are the main reasons for the precaution exercised by many enterprises and even some ordinary users to the adaptation of cloud computing. But benefits of using cloud have caused some of enterprises to have a plan in which cloud computing is used for less-sensitive data, but they may have local machines to store data which are of greater sensitivity. Should cloud providers wish for clients to store greater amounts of sensitive data in the cloud computing environment, improving security (and also, of course, client perception of that security) is paramount.

Whilst cloud computing is an important trend that keeps transforming and will continue to transform the IT industry, it doesn't mean that all business IT needs should move to the cloud computing model. The key to successful cloud computing initiatives is achieving a balance between the business benefits and the hidden potential risks lurking on the path to implementation.

References

[1] "Securing Virtualization in Real-World Environments," White paper2009.

[2] P. Coffee, "Cloud Computing: More Than a Virtual Stack," ed: salesforce.com.

[3] *Software as a Service*. Available: http://www.wikinvest.com/concept/Software_as_a_Service

[4] G. Reese, *Cloud Application Architectures*: O'Reilly Media, Inc.,, 2009.

[5] N. Mirzaei, "Cloud Computing," 2008.

[6] "Database as a Service," MIT-CSAIL-TR-2010-014.

[7] M. Riccuiti. Stallman: Cloud computing is stupidity. Available: http://news.cnet.com/8301-1001_3-10054253-92.html

[8] N. Antonopoulos and L. Gillam, *Cloud Computing*: Springer-Verlag London Limited, 2010.

[9] K. JACKSON, "Secure Cloud Computing: An Architecture Ontology Approach," Defense Information Systems Agency2009.

[10] R. Raja and V. Verma, "Cloud computing: An overview," Research Consultant, IIIT Hyderabad.

[11] D. Rowe. (2011, The Impact of Cloud on Mid-size Businesses. Available: http://www.macquarietelecom.com/hosting/blog/cloud-computing/impact-cloudcomputing-midsize-businesses

[12] S. Hanna, "Cloud Computing: Finding the Silver Lining," Juniper Networks2009.

[13] *Cloud Computing*. Available: http://en.wikipedia.org/wiki/Cloud_computing

[14] J. Brodkin. (2008). *Gartner: Seven cloud-computing security risks*. Available: http://www.networkworld.com/news/2008/070208-cloud.html

[15] J. Metzler. (2009, Virtualisation can make application delivery much, much harder - but you can fight back! Available: http://searchnetworking.techtarget.com.au/articles/33471-Virtualisation-can-make-application-delivery-much-much-harder-but-you-can-fight-back-

[16] "Virtualization: The next generation of application delivery challenges."

[17] (2011). *What is Cloud Sprawl and Why should I Worry About It?* Available: http://www.cloudbusinessreview.com/2011/06/08/what-is-cloud-sprawl-and-why-should-i-worry-about-it.html

[18] R. Chow, P. Golle, M. Jakobsson, E. Shi, J. Staddon, R. Masouka, and J. Molina, "Controlling Data in the Cloud: Outsourcing Computation without Outsourcing Control," presented at the CCSW'09, Chicago, Illinois, USA., 2009.

[19] D. Talbot. (2009). *Vulnerability Seen in Amazon's Cloud-Computing*. Available:

http://www.technologyreview.com/printer_friendly_article.aspx?id=23792

[20] J. W. Rittinghouse and J. F. Ransome, *Cloud Computing Implementation, Management, and Security*: Taylor and Francis Group, LLC, 2010.

[21] P. Sefton, "Privacy and data control in the era of cloud computing."

[22] Texiwill. (2009). *Is Network Security the Major Component of Virtualization Security?* Available: http://www.virtualizationpractice.com/blog/?p=350

[23] D. E. Y. Sarna, *Implementing and Developing Cloud Computing Applications*: Taylor and Francis Group, LLC, 2011.

[24] t. Ristenpart and e. al, "Hey, you, get off of my cloud: exploring information leakage in third-party compute clouds," 2009.

[25] S. K. Tim Mather, and Shahed Latif, *Cloud Security and Privacy*: O'Reilly Media, Inc., 2009.

[26] C. Almond, "A Practical Guide to Cloud Computing Security," 2009.

[27] *Cloud Security*. Available: http://cloudsecurity.trendmicro.com/

[28] N. Mead, E. Hough, and T. Sehny, "Security quality requirements engineering (SQUARE) methodolgy," Carnegie Mellon Software Engineering Institute.

[29] K. K. Fletcher, "Cloud Security requirements analysis and security policy development using a high-order object-oriented modeling," Master of science, Computer Science, Missouri University of Science and Technology, 2010.

[30] F. Sabahi, "Analysis of Security in Cloud Environments," presented at the International Conference on Computer Science and Information Technology, Chengdu, China, 2011.

[31] P. R. Gallagher, *A Guide to Understanding Data Remanence in Automated Information Systems*: The Rainbow Books, 1991.

[32] (2009). *Cloud Computing*. Available: http://groups.google.com/group/cloud-computing/browse_thread/thread/21e585b137125554

[33] S. Roschke, F. Cheng, and C. Meinel, "Intrusion Detection in the Cloud," presented at the Eighth IEEE International Conference on Dependable, Autonomic and Secure Computing, Chengdu, China, 2009.

[34] F. Sabahi, "Intrusion Detection Techniques performance in Cloud Environments," presented at the International Conference on Computer Design and Engineering, Kuala Lumpur, Malaysia, 2011.

[35] P. Cox, "Intrusion detection in a cloud computing environment," 2010.

[36] L. Litty, "Hypervisor-based Intrusion Detection," Master of Science, 2005.

[37] L. Ponemon, "Security of Cloud Computing Users," 2010.

[38] (2010). *Security Management in the Cloud*. Available: http://mscerts.net/programming/Security%20Management%20in%20the%20Cloud.aspx

[39] (2010). *Security Management in the Cloud - Access Control*. Available: http://mscerts.net/programming/Security%20Management%20in%20the%20Cloud%20-%20Access%20Control.aspx

International Journal on Advances in ICT for Emerging Regions 2011 04 (02) : 24 - 36

Trust Management in Cloud Computing: A Critical Review

Mohamed Firdhous, Osman Ghazali and Suhaidi Hassan

***Abstract*—Cloud computing has been attracting the attention of several researchers both in the academia and the industry as it provides many opportunities for organizations by offering a range of computing services. For cloud computing to become widely adopted by both the enterprises and individuals, several issues have to be solved. A key issue that needs special attention is security of clouds, and trust management is an important component of cloud security.**

In this paper, the authors look at what trust is and how trust has been applied in distributed computing. Trust models proposed for various distributed system has then been summarized. The trust management systems proposed for cloud computing have been investigated with special emphasis on their capability, applicability in practical heterogonous cloud environment and implementabilty. Finally, the proposed models/systems have been compared with each other based on a selected set of cloud computing parameters in a table.

***Index Terms*—Cloud Computing, Trust, Trust Management, Trust Models**

I. Introduction

Distributed systems like peer-to-peer systems, grid, clusters and cloud computing have become very popular among users in the recent years. Users access distributed systems for different reasons such as downloading files, searching for information, purchasing goods and services or executing applications hosted remotely. With the popularity and growth of distributed systems, service providers make new services available on the system. All these services and service providers will have varying levels of quality and also, due to the anonymous nature of the systems, some unscrupulous providers may tend to cheat unsuspecting clients. Hence it becomes necessary to identify the quality of services and service providers who would meet the requirements of the customers [1].

In this paper the authors take a look at the trust and trust management systems along with the trust models developed for distributed systems. Then a critical look at the trust development and management systems for cloud computing systems reported in literature in the recent times has been taken with special reference to the pros and cons of each proposal.

II. Cloud Computing

Cloud computing has been called the 5th utility in line of electricity, water, telephony and gas [2]. The reason why cloud has been nomenclature with such a name is that cloud computing has been changing the way computer resources have been used up to now. Until the development of cloud computing, computing resources were purchased outright or leased in the form of dedicated hardware and software resources. Cloud computing has brought a paradigm change in how computing resources have been purchased. With the advent of cloud computing, users can use the services that have been hosted on the internet without worrying about whether they have been hosted or managed in such a manner that the customers have to pay only for the services they consumed as in the case of making use of other services.

Cloud providers host their resources on the internet on virtual computers and make them available to multiple clients. Multiple virtual computers can run on one physical computer sharing the resources such as storage, memory, the CPU and interfaces giving the feeling to the client that each client has his own dedicated hardware to work on. Virtualization thus gives the ability to the providers to sell the same hardware resources among multiple clients. This sharing of the hardware resources by multiple clients help reduce the cost of hardware for clients while increasing profits of providers. Accessing or selling hardware in the form of virtual computers is known as Infrastructure as Service (IaaS) in the cloud computing terminology [3]. Once a client has procured infrastructure from a service provider, he is free to install and run any Operating System platform and application on it.

Other kinds of services that are made available via the cloud computing model are Platform as a Service (PaaS) and Software as a Service. Figure 1, shows the architecture of a typical cloud computing system.

Under PaaS, the development platform in the form of an Operating System has been made available where customers can configure the environment to suit their requirements and install their development tools [5]. PaaS helps developers

Manuscript received on February 27, 2011. Accepted July 16, 2011.

Mohamed Firdhous is a Senior Lecturer attached to the Faculty of Information Technology, University of Moratwua, Sri Lanka. (email: firdhous@itfac.mrt.ac.lk & mfirdhous@internetworks.my)

Dr. Osman Ghazali is a Senior Lecturer and the Head of School of Computing, College of Arts and Sciences of the Universiti Utara Malaysia. (email: osman@uum.edu.my)

Prof. Suhaidi Hassan is an Associate Professor and the Asst. Vice Chancellor of the College of Arts and Sciences of the Universiti Utara Malaysia. He is the team leader of the InterNetWorks Research Group. (email: suhaidi@uum.edu.my)

develop and deploy applications without the cost of purchasing and managing the underlying hardware and software. PaaS provides all the required facilities for the complete life cycle of building and delivering web applications. Thus PaaS usually offers facilities for application design, application development, testing, deployment and hosting as well as application services such as team collaboration, web service integration and marshalling, database integration, security, scalability, storage, persistence, state management, application versioning, application instrumentation and developer community facilitation.

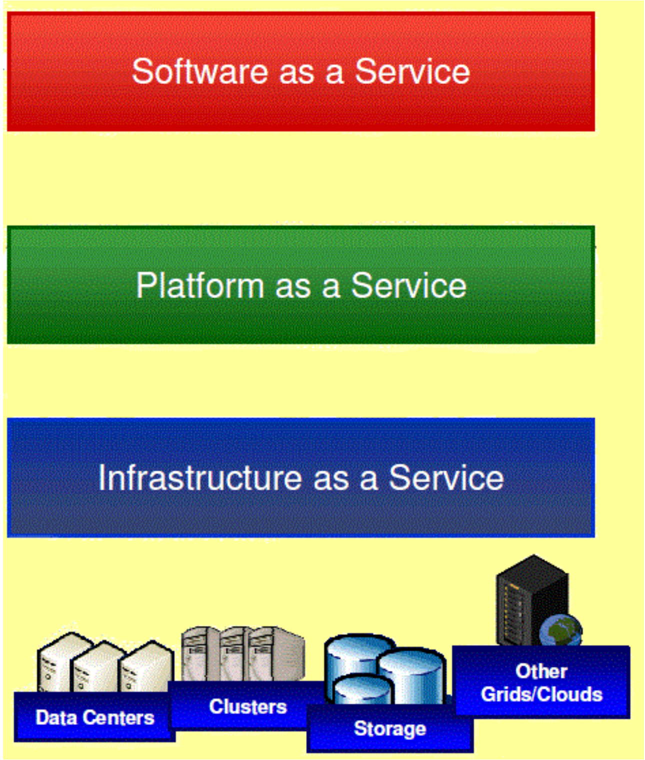


Fig. 1. Cloud Computing Architecture

SaaS is the cloud model where an application hosted by a service provider on the internet is made available to users in a ready to use state. SasS eliminates the requirement of installation and maintenance of the application in the user's local computer or server in his premises [5]. SaaS has the advantage of being accessible from any place at any time, no installation or maintenance, no upfront cost, no licensing cost, scalability, reliability and flexible payment schemes to suit the customer's requirements.

III. Trust and Trust Management

The trust and reputation have their origin in the social sciences that study the nature and behavior of human societies [6]. Trust has been studied by researchers in diverse fields such as psychology, sociology and economics [7]. Psychologists study trust as a mental attitude and focus on what happens in a person's mind when he/she trusts or distrusts someone [8]. Based on this notion, several cognitive trust models have been developed [9-12]. Sociologists approach to trust as a social relationship between people. Social context of trust has been commonly employed in multi agent systems and social networks [7,13-14]. The similarity between multi agent system and a social network are exploited in these works as agents and people behave in a similar fashion interacting with, gathering information from and modeling each other for developing trust in each other. Economists perceive trust in terms of utility [15]. Game theory has been one of the most popular tools used by experts in the computer field to study how users develop trust using different strategies [16-17]. The prisoner's dilemma is the commonly used scenario to study this scenario [18-19].

Researchers in computer sciences have exploited the benefit of all these studies as they provide vital insight into human behavior under various circumstances [13, 20-21]. The role of trust and reputation in open, public distributed systems such as e-commerce, peer to peer networks, grid computing, semantic web, web services and mobile networks have been studied by several researchers [22-25].

Although the rich literature available on trust from diverse fields is of great benefit to computer scientists, it has the drawback of presenting a complex and confusing notion for trust. This is mainly due to the reason that there is no common agreement of a single definition for what trust is? It can be seen that different researchers have defined trust as attitudes, beliefs, probabilities, expectations, honesty and so on.

Even if different disciplines and researchers look at trust from different angles, it is possible to identify some key factors that are common to everything. They are;

- Trust plays a role only when the environment is uncertain and risky.
- Trust is the basis based on which certain decisions are made.
- Trust is built using prior knowledge and experience.
- Trust is a subjective notion based on opinion and values of an individual.
- Trust changes with time and new knowledge while experience will have overriding influence over the old ones.
- Trust is context-dependent.
- Trust is multi-faceted.

McKnight and Chervany have identified 16 characteristics of trust and grouped them under five groups. They are,

- Competence; competent, expert, dynamic
- Predictability; predictable
- Benevolence; good (or moral), good-will benevolent (caring), responsive
- Integrity; honest, credible, reliable, dependable
- Other; open, careful (or safe), shared understanding, personally attractive [8].

De Oliveira and Maziero have classified trust relations into hierarchical trust, social groups and social networks.

Hierarchical trust considers all relationships in a hierarchical manner and represented by a tree organization where nodes represent individuals and edges represent the trust degrees between the pair of nodes. Any two nodes can define a trust degree between them through transitivity through other nodes [26].

Zhang et al., have classified the trust functions based on the following four dimensions [27].

- Subjective trust vs. Objective trust
- Transaction-based vs. Opinion-based
- Complete information vs. Localized information
- Rank-based vs. Threshold-based

Capability of an entity's trustworthiness being measured objectively against a universal standard, results in objective trust. If the trust being measured depends on an individual's tastes and interest, the resulting trust is called *subjective trust*. Decisions made based on the individual transactions and their results is known as *transaction based trust*, whereas the trust built based on just opinion of the individuals, is *opinion based trust*. If the trust building operation requires information from each and every node, it is called, complete information and it is known as either *global trust function* or *complete trust function*. If the information collected only from one's neighbors, it is called, *localized information trust function*. If the trust worthiness of an entity is ranked from the best to worst, it is *rank based trust* whereas the trust declared yes or no depending on? Preset trust threshold is known as *threshold based trust.*

IV. Trust Models

Several models have been developed by researchers for the purpose of building practical trust systems in distributed systems. This section takes a brief look at some of the commonly used trust models.

A. CuboidTrust

CuboidTrust is a global reputation-based trust model for peer to peer networks. It takes three factors namely, contribution of the peer to the system, peer's trustworthiness in giving feedback and quality of resources to build four relations. Then it creates a cuboid using small cubes whose coordinates (x,y,z) where z – quality of resource, y – peer that stores the value and x – the peer which rated the resource and denoted by Px,y,z. The rating is binary, 1 indicating authentic and (–1) indicating inauthentic or no rating. Global trust for each peer has been computed using power iteration of all the values stored by the peers [28].

B. EigenTrust

EigenTrust assigns each peer a unique global trust value in a P2P file sharing network, based on the peer's history of uploads. This helps to decrease the downloading of inauthentic files. Local trust value S_{ij} has been defined S_{ij} = *sat(i,j)* − *unsat(i,j)*, where *sat(i,j)* denotes the satisfactory downloads by *i* from *j* and *unsat(i,j)* is the unsatisfactory downloads by *i* from *j*. Power iteration is used to compute the global trust for each peer [29].

C. Bayesian Network based Trust Management (BNBTM)

BNBTM uses multidimensional application specific trust values and each dimension is evaluated using a single Bayesian network. The distribution of trust values is represented by beta probability distribution functions based on the interaction history [30].

Trust value of peer *i* is given by,

$$\tau_i = \frac{\alpha_i}{\alpha_i + \beta_r}, (i \in \{G, L, C\}, \bar{\imath} \in \{\bar{G}, \bar{L}, \bar{C}\}) \tag{1}$$

Where $\alpha_i = r_i + 1 \; and \; \beta_r = s_r + 1$ and r_i and s_r are number of interactions with outcome i and $\bar{\imath}$ respectively. $G, L \; and \; C$ represent shipping goods, shipping lower quality goods and not shipping any goods and $\bar{G}, \bar{L} \; and \bar{C}$ represent the converse.

D. GroupRep

GroupRep is a group based trust management system. This classifies trust relationships in three levels namely, trust relationships between groups, between groups and peers and only between peers [31].

Trust of Group *i* held by Group *j* is given by:

$$T_{r\,G_iG_j} = \begin{cases} \dfrac{u_{G_iG_j} - c_{G_iG_j}}{u_{G_iG_j} + c_{G_iG_j}} \; if \; u_{G_iG_j} + c_{G_iG_j} \neq 0 \\ T_{rG_iG_j}^{reference} \; if \; u_{G_iG_j} + c_{G_iG_j} = 0 \; and \; \exists \; Trust_{G_iG_j}^{path} \\ T_{r\,G_iG_{strange}} \; otherwise \end{cases} \tag{2}$$

Where $u_{G_iG_j} \geq 0$ and $c_{G_iG_j} \geq 0$ are utility and cost respectively assigned by nodes in group *j* to nodes in group *i*.

$T_{rG_iG_j}^{reference}$ is defined as the minimum trust value along the most trustworthy reference path.

E. AntRep

AntRep algorithm is based on swarm intelligence. In this algorithm, every peer maintains a reputation table similar to distance vector routing table. The reputation table slightly differs from the routing table in the sense that (i) each peer in the reputation table corresponds to one reputation content; (ii) the metric is the probability of choosing each neighbor as the next hop whereas in the routing table it is the hop count to destinations. Both forward ants and backward ants are used for finding reputation values and propagating them. If the reputation table has a neighbor with the highest reputation, a unicast ant is sent in that direction. If no preference exists, broadcast ants are sent along all the paths [32].

Once the required reputation information is found, a backward ant is generated. When this ant travels back, it updates all the reputation tables in each node on its way.

F. Semantic Web

Zhang et al., have presented a trust model which searches all the paths that connect the two agents to compute the trustworthiness between those two agents. For each path the ratings associated with each edge are multiplied and finally all the paths are added to calculate the final trust value [33].

The weight of the path *i* (w_i) is calculated using;

$$w_i = \frac{\frac{1}{D_i}}{\sum_{i=1}^{N} \frac{1}{D_i}} \tag{3}$$

Where N – No. of paths between agents P and Q

D_i – No. of steps between P and Q on the i^{th} path.

m_i – Q's immediate friend or neighbor on the i^{th} path. (M – set of Q's friends or neighbors)

This gives a higher weight to shorter paths.

If agent *P* and agent *Q* are friends then $P \to Q$, or neighbors then $P \leftrightarrow Q$ then *P*'s trust in *Q* can be computed directly. Otherwise,

$$T_{P\to Q} = \sum_{i=1}^{N} \frac{T_{m_i \to Q} \times \prod_{i\to j \cup i \leftrightarrow j} R_{i\to j} \times \frac{1}{D_i}}{\sum_{i=1}^{N} \frac{1}{D_i}}$$

$$= \sum_{i=1}^{N} T_{m_i \to Q} \times \prod_{i\to j \cup i \leftrightarrow j} R_{i\to j} \times w_i \tag{4}$$

Where reliability factor $R_{i\to j}$ denotes to which degree *i* believes in *j*'s words or opinions.

G. Global Trust

Several authors have presented methods that compute an improved global trust value for selecting trusted source peer in peer to peer systems [34-36].

The global trust value for node *i*, t_i is defined as:

$$t_i = \sum_{k} c_{ki}\, t_k \tag{5}$$

Where c_{ki} is the local trust value from peer *k* towards peer *i* and t_k is the global trust value of peer *k*.

H. Peer Trust

This is reputation-based trust supporting framework. This includes a coherent adaptive trust model for quantifying and comparing the trustworthiness of peers based on a transaction-based feedback system. It introduces three basic trust parameters namely feedback a peer receives from other peers, the total number of transactions a peer performs, the credibility of the feedback sources and two adaptive factors that are transaction context factor and the community context factor in computing trustworthiness of peers, then it combines these factors to compute a general trust metric [37].

I. PATROL-F

PATROL-F incorporates many important concepts for the purpose of computing peer reputation. The main components used in computing peer trust are: direct experiences and reputation values, the node credibility to give recommendations, the decay of information with time based on a decay factor, first impressions and a node system hierarchy [38].

It uses three fuzzy subsystems:

1. The first is used to set the importance factor of an interaction and related decisions. To decide and choose which data is critical or indispensable, or which data is needed more quickly, is a concept close to humans that fuzzy logic can model.
2. Then there is the region of uncertainty where an entity is not sure whether to trust or not (when the reputation of a host between the absolute mistrust level φ , and the absolute trust level θ). Fuzzy techniques are effectively applied in this region.
3. Finally, for the Result of Interaction (RI) value, fuzzy logic can be used to capture the subjective and humanistic concept of four level *"good"* or *"better"* and *"bad"* or *"worse"* interaction. RI is the result of several concepts effectively combined to produce a more representative value. The decay factor τ is calculated based on the difference of a host's values of RIs between successive interactions.

J. Trust Evolution

Wang et al., have presented a trust evolution model for P2P networks. This model uses two critical dimensions, experience and context to build trust relationships among peers. It builds two kinds of trust: direct trust and recommendation trust quantifies trust within the interval *[0,1]* [39].

Direct trust (DT) between two peers is computed using the last *n* interactions between those entities. Recommended trust is calculated using recommendations from other peers and the previous interactions with the recommending peers.

K. Time-based Dynamic Trust Model (TDTM)

TDTM is an ant colony based system that identifies the pheromone and the trust and the heuristic and the distance between two nodes. The trust value calculated by this model depends on the frequency of interaction where the trust value increases with frequent interactions and lowers as the interactions goes down [40].

Trust-pheromone between nodes *i* and *j* at time (*t +1*) is defined as:

$$\tau_{ij}(t+1) = \rho\tau_{ij}(t) + \sigma\tau_{ij}(t) \tag{6}$$

Where ρ is the trust dilution factor and $\sigma\tau_{ij}(t)$ is the additional intensity at each inter-operation between entities.

$\sigma\tau_{ij}(t)$ is defined as:

$$\sigma\tau_{ij} = \begin{cases} \dfrac{1}{\frac{1}{1-\tau_{ij}(t)}+1} & \text{if } i \text{ and } j \text{ interact at time } t \\ 0 & \textit{otherwise} \end{cases} \tag{7}$$

If the trust value $p_{ij}(t)$ between nodes *i* and *j* at time *t* is greater than a certain threshold *R*, they can validate each other's certificate, otherwise not.

L. *Trust Ant Colony System (TACS)*

TACS is based on the bio-inspired algorithm of ant colony system. In this model pheromone traces are identified with the amount of trust a peer has on its neighbors when supplying a specific service. It computes and selects both the most trustworthy node to interact and the most trustworthy path leading to that peer. Each peer needs to keep track of the current topology of the network as every peer has its own pheromone traces for every link. Ants travel along every path searching building the most trustworthy path leading to the most reputable server [41].

Ants stop the search once they find a node that offers the service requested by the client and the pheromone traces belonging to the current path leading to it are above the preset threshold, otherwise they would follow on further selecting a neighbor that has not been visited yet.

M. *TRUMMAR (TRUst Model for Mobile Agent systems based on Reputation)*

TRUMMAR is a general model for the calculation of reputation values and the determination of trust decisions. TRUMMAR identifies three types of nodes from who it can receive trust values. They are neighbors, friends and strangers. Neighbors are the trusting other hosts on its own network that are under the same administrative control, friends are the hosts from different networks that are under different, but trusted administrative control and strangers are the hosts that are willing to volunteer information but not neighbors or friends [42].

The trust value for Y in X is calculated as follows:

$$\frac{repY}{X(0)} = A\frac{repY}{X} + B\frac{\sum_i \frac{\alpha_i repY}{X_i}}{\sum_i \alpha_i} + C\frac{\sum_j \frac{\beta_i repY}{X_j}}{\sum_j \beta_j} + D\frac{\sum_l \frac{\delta_l repY}{X_l}}{\sum_l \delta_l} \tag{8}$$

Where

$\frac{repY}{X(0)}$ represents the reputation value being calculated.

$\frac{repY}{X(0)}$ represents the reputation value last calculated, modified to account for the time lapsed.

$\frac{\sum_i \frac{\alpha_i repY}{X_i}}{\sum_i \alpha_i}$ weighted sum of reputation reported by neighbors.

$\frac{\sum_j \frac{\beta_j repY}{X_j}}{\sum_j \beta_j}$ weighted sum of reputation reported by friends.

$\frac{\sum_l \frac{\delta_l repY}{X_l}}{\sum_l \delta_l}$ weighted sum of reputation reported by strangers.

α_i , β_j and δ_l are weighing factors which depend on the reputation of the individual neighbors, friends, and strangers in the host space, respectively.

A, B, C, and D are weighing factors for the respective reputation of with respect to self, neighbors, friends and strangers in the agent space and A > B > C > D.

Reputation values are restricted to values between 0 and k, i.e $0 \leq \frac{repY}{X}$:

N. *PATROL (comPrehensive reputAtion-based TRust mOdeL)*

PATROL is a general purpose reputation based trust model for distributed computing. PATROL is an enhancement over TRUMMAR. This model is based on multiple factors such as reputation values, direct experiences, trust in the recommender, time dependence of the trust value, first impressions, similarity, popularity, activity, cooperation between hosts, and hierarchy of host systems. The decision to interact with another host depends on two factors namely, the trust in the competence of a host and the trust in the host's credibility to give trusted advice. The trust in the competence of a host is calculated from the direct interactions and this is the confidence that the other host would be able to complete the intended task to the initiator host's expectations. The trust in a host's ability to give trusted advice is the confidence that the host gives consistent and credible advice and feedback. The overall trust value is a combination of the weighted values calculated for different factors calculated independently [43].

The operation of the model is as given below:

1. Host X wants to interact with host Y.
2. X calculates the time since it interacted last with Y, if this time is smaller than a predetermined threshold, it will decay the stored trust value compare against a predetermined threshold. If larger than the threshold, it will interact with Y, otherwise not.
3. If the last interaction time was larger than the threshold, it will involve other trusted hosts in its calculation of trust value for Y. If not,
4. Queried hosts will decay their stored trust value for Y and send it along with their reputation vectors.
5. X will calculate the trust for Y and check against the threshold. If the trust value is greater than the threshold, it will interact with Y, otherwise no interaction.

O. META-TACS

META-TACS is an extension of the TACS algorithm developed by the [41]. They have extended the TACS model by optimizing the working parameters of the algorithm using genetic algorithms [44].

P. CATRAC (Context-Aware Trust- and Role-Based Access Control for composite web services)

Role-Based Access Control (RBAC) and Trust-Based Access Control (TBAC) have been proposed to address threats to security in single Web Service scenarios. But these solutions fail to provide the required security level in situations related to composite Web Services. CATRAC has been proposed as a security framework related to composite web services [45]. CATRAC combines both RBAC and TBAC in order to arrive at an optimum solution.

Three conditions must be satisfied to gain access to a specific web service. They are:

- Client attributes must be authenticated by the web service provider.
- Client's global role must be valid and contains the right permissions.
- Client's trust level must be equal or greater than the threshold level set for the particular service.

A trusted third party called the Role Authority issues, signs and verifies the roles assigned to the clients. Trust levels are expressed as a vector ranging from 0 to 10, indicating the fully distrusted to the fully trusted respectively. Five (5) indicates a neutral or uncertainty level which is commonly assigned to new clients.

CATRAC is made up of three entities, namely Role Authority, Servers and Clients. Clients accumulate trust points when their behavior is considered good and otherwise they lose trust points. Also, clients trust level is decayed to the neutral value gradually with time, if no interaction takes place. Trust level is decayed using the following formulae.

$$D_{TL_c} = \left((TL_c - TL_N) \times e^{\left(\frac{-t}{memo_s}\right)} \right) + TL_N \qquad (9)$$

If the current trust level is above the neutral trust level.

$$D_{TL_c} = TL_N \times e^{\left(\frac{-t}{memo_s}\right)} \quad \textit{otherwise.} \qquad (10)$$

Where

D_{TL_c} – decayed trust level for client c
TL_c – current trust level for client c
TL_N – neutral trust level
t – time elapsed
$memo_s$ – memory factor (constant)

Q. Bayesian Network -based Trust Model

Bayesian Network–based Trust Model computes trust values by combining multiple input attributes [46]. In this model, the different capabilities of providers such as the type of the file, quality of the file, download speed etc. Also, it looks at the contextual representation of trust values. That is, if two agents compute the trust values, they can trust each other's recommendation and if the agents use different criteria, they may not trust the each other's recommendation even if both are truthful.

In this system each peer identified as an agent develops a naïve Bayesian network for each provider it has interacted with. Each Bayesian network has a root node T with two branches named "satisfying" and "unsatisfying", denoted by 1 and 0, respectively. The agents overall trust in the provider's competencies represented by *p(T=1)*, which is the ratio of interactions with satisfactory results out of all the interactions with the same provider. On the other hand *p(T=0)* is the ratio of unsatisfactory results under the same criteria.

Hence: $p(T=1) + p(T=0) = 1 \qquad (11)$

Depending on the results of the previous interactions, the agent creates a conditional probability in the form of *p(File Type = "Music" | T = 1)* or *p(Download Speed = "High" | T = 1)* for each quality attribute such as file type, file quality and speed. These conditional probability values are stored in a table called the Conditional Probability Table (CPT).

Finally the provider's trustworthiness in different aspects such as *p(T = 1 | File Type ="Music" AND Download Speed = "High")* is computed by combing the conditional probability values stored in the CPT using the Bayes rule. This combined trustworthiness value is the overall trust score of the provider for the given attribute(s) or aspect(s).

The models discussed above have been proposed for different types of distributed systems such as clusters, grids and wireless sensor networks. But none of the above models has been tested on the cloud computing environment. Hence an extensive evaluation of these models needs to be carried out to understand the advantages and disadvantages of these models for use in cloud computing. The authors propose to carry out this kind of evaluation of these models in future work. Next section takes an in depth look at the trust models proposed for cloud computing.

V. Trust in Cloud Computing

Security is one of the most important areas to be handled in the emerging area of cloud computing. If the security is not handled properly, the entire area of cloud computing would fail as cloud computing mainly involves managing personal sensitive information in a public network. Also, security from the service providers point also becomes imperative in order to protect the network, the resources in order to improve the robustness and reliability of those resources. Trust management that models the trust on the behavior of the elements and entities would be especially useful for the proper administration of cloud system and cloud services.

Several leading research groups both in academia and the industry are working in the area of trust management in cloud computing. This section takes an in depth look at the recent developments in this area with the objective of identifying and categorizing them for easy reference.

Khan and Malluhi have looked at the trust in the cloud system from a users perspective. They analyze the issues of trust from what a cloud user would expect with respect to their data in terms of security and privacy. They further discuss that what kind of strategy the service providers may undertake to enhance the trust of the user in cloud services and providers. They have identified control, ownership, prevention and security as the key aspects that decide users' level of trust on services. Diminishing control and lack of transparency have identified as the issues that diminishes the user trust on cloud systems. The authors have predicted that remote access control facilities for resources of the users, transparency with respect to cloud providers actions in the form of automatic traceability facilities, certification of cloud security properties and capabilities through an independent certification authority and providing security enclave for users could be used to enhance the trust of users in the services and service providers [47].

Zhexuan et al., have taken a look at the security issues SaaS might create due to the unrestricted access on user data given to the remotely installed software [48]. The authors have presented a mechanism to separate software from data so that it is possible to create a trusted binding between them. The mechanism introduced involves four parties namely the resource provider, software provider, data provider and the coordinator. The resource provider hosts both data and software and provides the platform to execute the software on data. The software provider and data provider are the owners of the software and data respectively. The coordinator brings the other parties together while providing the ancillary services such as searching for resources and providing an interface to execute the application on the data.

The operation of the model is as follows:

Software provider and data provider upload their resources to the resource provider. These resources will be encrypted before stored and the key will be stored in the accountability vault module of the system.

A data provider searches for and finds the required software through a coordinator and then runs the software on the data uploaded to the resource provider's site.

Once the execution has started an execution reference ID is generated and given to the data provider.

When the execution of the software is over, the results are produced only on the data provider's interface which can be viewed, printed or downloaded.

Data provider will then pay for the service that will be split between the software provider and resource provider.

An operation log has been created and posted to the software provider without disclosing the data provider's identity or the content on which software was run. This helps the software provider know that his software has been used and the duration of use.

Even though the authors claim that this model separates the software and data, there is no assurance that the software cannot make a copy while the data is being processed as only the algorithm or description of the software is provided to the data owner. Without the source code, there is no assurance that the code will not contain any malicious code hidden inside. Also, since the software runs on data owner's rights and privileges, the software would have complete control over data. This is a security threat and the audit trail even if it is available, will not detect any security breaches.

The authors do not address the question of trust on the proposed platform as this would be another application or service hosted on the cloud. Both application providers and data providers need some kind of better assurance as now they are entrusting their data and software to a third party software.

Sato et al., have proposed a trust model of cloud security in terms of social security [49]. The authors have identified and named the specific security issue as social insecurity problem and tried to handle it using a three pronged approach. They have subdivided the social insecurity problem in to three sub areas, namely; multiple stakeholder problem, open space security problem and mission critical data handling problem.

The multiple stakeholder problem addresses the security issues created due to the multiple parties interacting in the cloud system. As per the authors, three parties can be clearly identified. They are namely, the client, the cloud service providers and third parties that include rivals and stakeholders in business. The client delegates some of the administration/operations to cloud providers under a Service Level Agreement (SLA). Even if the client would like to have the same type of policies that it would apply if the resources were hosted on site on the delegated resources, the provider's policy may differ from that of the client. The providers are bound only by the SLA signed between the parties. The SLA plays the role of glue between the policies. Also the authors opine that once the data is put in the cloud it is open for access by third parties once authenticated by the cloud provider.

The open space security problem addresses the issue of loss of control on where the data is stored and how they are physically managed once control of data is delegated to the cloud provider. They advice to encrypt the data before transferring, converting the data security problem to a key management problem as now the keys used for encryption/decryption must be handled properly.

The mission critical data handling problem looks at the issue of delegating the control of mission critical data to a service provider. They advice not to delegate control of this data but to keep them in a private cloud in a hybrid setup, where the organization have unhindered control. However setting up of a private cloud may not be an option to small and medium sized organizations due to the high costs involved. Hence enhancement of security of the public cloud is the only option to serve everybody.

Authors have developed a trust model named 'cloud trust model' to address the problems raised above. Two more trust layers have been added to the conventional trust architecture. These layers have been named as The internal trust layer and the contracted trust layer. The Internal trust layer acts as the platform to build the entire trust architecture. It is installed in the in house facilities and hence under the control of the local administration. ID and key management are handled under the internal trust. Also any data that is considered critical or needs extra security must be stored under this layer.

Contracted trust has been defined as the trust enforced by an agreement. A cloud provider places his trust upon the client, based on the contract that is made up of three documents known as, Service Policy/Service Practice Statement (SP/SPS), Id Policy/Id Practice Statement (IdP/IdPS) and the contract.

Level of trust required can be negotiated by parties depending on the level of security needed for the data. A cloud system thus installed is called a secure cloud by the authors.

Li et al., propose a domain-based trust model to ensure the security and interoperability of cloud and cross-clouds environment and a security framework with an independent trust management module on top of traditional security modules [50]. They also put forward some trust based security strategies for the safety of both cloud customers and providers based on this security model.

A cloud trust model based on the family gene technology that is fundamentally different from the Public key Infrastructure based trust models has been proposed by Wang et al.,. The authors have studied the basic operations such as user authentication, authorization management and access control and proposed a Family-gene Based model for Cloud Trust (FBCT) integrating these operations [51-52].

Manuel et al., have proposed trust model that is integrated with CARE resource broker [53]. The proposed trust model can support both grid and cloud systems. The model computes trust using three main components namely, Security Level Evaluator, Feedback Evaluator and Reputation Trust Evaluator. Security Level Evaluation has been carried out based on authentication type, authorization type and self security competence mechanism. Multiple authentication, authorization mechanism and self security competence mechanisms are supported. Depending on the strength of individual mechanism, different grades are provided for trust value. Feedback Evaluation also goes through three different stages namely feedback collection, feedback verification and feedback updating. The Reputation Trust Evaluator computes the trust values of the grid/cloud resources based on their capabilities based on computational parameters and network parameters. Finally the overall trust value has been computed taking the arithmetic sum of all the individual trust values computed.

Shen et al., and Shen and Tong have analyzed the security of cloud computing environment and described the function of trusted computing platform in cloud computing [54-55]. They have also proposed a method to improve the security and dependability of cloud computing integrating the Trusted Computing Platform (TCP) into the cloud computing system. The TCP has been used in authentication, confidentiality and integrity in cloud computing environment. Finally the model has been developed as software middleware known as the Trusted Platform Software Stack (TSS).

Alhamad et al., have proposed a SLA based trust model for cloud computing. The model consists of the SLA agents, cloud consumer module and cloud services directory [56]. The SLA agent is the core module of the architecture as it groups the consumers to classes based on their needs, designs SLA metrics, negotiates with cloud providers, selects the providers based on non functional requirements such as QoS, and monitors the activities for the consumers and the SLA parameters. Cloud consumer module requests the external execution of one or more services. Cloud services directory is the one where the service providers can advertise their services and consumers seek to find the providers who meet their functional requirements such as database providers, hardware providers, application providers etc.,

The authors have proposed only the model and no implementation or evaluation has been developed or described. Hence the each and every module will have to be evaluated for their functionality and the effectiveness and finally the overall model will have to be evaluated for its effectiveness.

Yong et al., have proposed a model called a multi-tenancy trusted computing environment model (MTCEM) for cloud computing [57]. MTCEM has been proposed to deliver trusted IaaS to customers with a dual level transitive trust mechanism that supports a security duty separation function simultaneously. Since cloud facilities belong to multiple stakeholders such as Cloud Service Providers (CSP) and customers, they belong to multiple security domain and server different security subjects simultaneously. The different stakeholders may be driven by different motives such as best service, maximization of the return on investment and hence may work detrimental to the other party involved. Hence cloud computing should have the capability to compartmentalize each customer and CSP and support security duty separation defining clear and seamless security responsibility boundaries for CSP and customers.

MTCEM has been designed as two-level hierarchy transitive trust chain model which supports the security duty separation and supports three types of distinct stakeholders namely, CSP, customers and auditors. In this model, CSP assume the responsibilities to keep infrastructures trusted while the customer assumes responsibility starting from the guest OS which installed by the customer on the Virtual Machines provided by the CSP. The auditor monitors the services provided by the CSP on behalf of the customers. The authors have implemented a prototype system to prove that MTCEM is capable of being implemented on commercial hardware and software. But no evaluation of the prototype on performance has been presented.

Yang et al., have studied the existing trust models and firewall technology. The authors have found that all the existing trust models ignore the existence of firewall in a network [58]. Since firewall is an integral and important component of any corporate security architecture, this non inclusion of firewall is a huge shortcoming. The authors have proposed a collaborative trust model of firewall-through based on Cloud theory. This paper also presents the detailed design calculations of the proposed trust model and practical algorithms of measuring and updating the value of dynamic trust.

The model has the following advantages compared to other models:

- There are different security policies for different domains.
- The model considers the transaction context, the historical data of entity influences and the measurement of trust value dynamically.
- The trust model is compatible with the firewall and does not break the firewall's local control policies.

Fu et al., have studied the security issues associated with software running in the cloud and proposed a watermark-aware trusted running environment to protect the software running in the cloud [59]. The proposed model is made up of two components namely the administrative center and the cloud server environment. The administrative center embeds watermark and customizes the Java Virtual Machines (JVM) and the specific trusted server platform includes a series of cloud servers deployed with the customized JVMs. Only specific and complete Java programs are allowed to run on the JVMs while rejecting all the unauthorized programs like invasion programs. The main advantage of this approach is that it introduces watermark aware running environment to cloud computing.

Ranchal et al., have studied the identity management in cloud computing and proposed a system without the involvement of a trusted third party [60]. The proposed system that is based on the use of predicates over encrypted data and multi-party computing is not only capable of using trusted hosts but also untrusted hosts in the cloud. Since the proposed approach is independent of a third party, it is less prone to attack as it reduces the risk of correlation attacks and side channel attacks, but it is prone to denial of service as active bundle may also be not executed at all in the remote host.

Takabi et al., have proposed a security framework for cloud computing consisting of different modules to handle security and trust issues of key components [61]. The main issues discussed in the paper are identity management, access control, policy integration among multiple clouds, trust management between different clouds and between cloud providers and users. The framework identifies three main players in the cloud. They are cloud customers, service integrators and service providers. The service integrator plays the role of the mediator who brings the customers and service providers together. Service integrator facilitates collaboration among different service providers by composing services to meet the customer requirements. It is the responsibility of the service integrator to establish and maintain trust between provider domains and providers and customers. The service integrator discover the services from service providers or other service integrators, negotiate and integrate services to form collaborating services that will be sold to customers.

The service integrator module is composed of security management module, trust management module, service management module and heterogeneity management module. The heterogeneity management module manages the heterogeneity among the service providers. In addition to the above modules there are other minor modules that handle small but important tasks.

In overall this is a very comprehensive framework. But the authors have not discussed the interoperability issue of each component in the framework or implemented a prototype to evaluate the function and efficiency of the components or the overall framework.

Table 1 summarizes the proposed cloud computing trust management systems under different cloud computing parameters. From this table it is evident that most of the models proposed remain short of implementation and only a few have been simulated to prove the concept. Also, there is no single model that meets all the requirements of a cloud architecture especially the identity management, security of both data and applications, heterogeneity and SLA management. Also none of these systems have been based on solid theoretical foundation such as the trust models have been discussed in Section IV.

TABLE I
SUMMARY AND COMPARISON OF CLOUD COMPUTING TRUST MANAGEMENT SYSTEMS
SUPPORT ACROSS MULTIPLE HETEROGENEOUS CLOUDS

Work	**Type**	**Identity Mgmt/ Authentication**	**Data Security**	**Cloud Layer**	**SLA Support**	**Heterogeneity Support***	**Implemented**	**Comments**
[47]	-	Discussed	Discussed	-	-	Yes	No	No concrete proposal. Only discussed the issues.
[48]	Complete Platform	No	Yes	SaaS	No	Yes	No	Only a mechanism has been proposed. No implementation or evaluation carried out.
[49]	Social security based	Discussed	Discussed	-	Discussed	No	No	No concrete proposal. Only discussed the issues.
[50]	Domain based	No	No	SaaS PaaS IaaS	No	Yes	No	Model has been tested using simulation.
[51 - 52]	Family gene based	Discussed	No	-	No	No	No	Model has been tested using simulation.
[53]	Integrated with CARE Resource Broker	Yes	Yes	-	No	Yes	No	Model has been tested using simulation.
[54 - 55]	Built on trusted platform service	Yes	Yes	IaaS	No	Yes	No	Only a model has been proposed.
[56]	-	No	No	-	Yes	Yes	No	Only a model has been proposed.
[57]	Built on Trusted Computing Platform	No	No	IaaS	No	No	Prototype Implemented	Concept has been proved with a prototype.
[58]	Domain based	No	No	-	No	Yes	No	Model has been tested using simulation.
[59]	Watermark based security	No	No	SaaS	No	No	Prototype Implemented	Concept has been proved with a prototype.
[60]	Based on active bundles scheme	Yes	No	-	No	Yes	Prototype Implemented	Concept has been proved with a prototype.
[61]	-	Yes	No	SaaS PaaS IaaS	No	Yes	No	Only a model has been proposed

VI. Conclusions

Cloud computing has been the new paradigm in distributed computing in the recent times. For cloud computing to become widely adopted several issues need to be addressed. Cloud security is one of the most important issues that has to be addressed. Trust management is one of the important component in cloud security as cloud environment will have different kinds of users, providers and intermediaries. Proper trust management will help the users select the provider based on their requirements and trust worthiness. Also, trust management would help the providers select the clients who are trustworthy to serve.

In the paper, a comprehensive survey has been carried out on the trust management systems implemented on distributed systems with a special emphasis cloud computing. There are several trust models proposed for distributed systems. These models were mainly proposed for systems like clusters, grids and wireless sensor networks. These models have not been used or tested in cloud computing environments. Hence the suitability of these models for use in cloud computing cannot be recommended without an extensive evaluation. The authors propose to evaluate these models in future work. The trust management systems proposed for cloud computing have been extensively studied with respect to their capability, their applicability in practical heterogonous cloud environment and their implementabilty. The results have been presented in table for easy reference. During the evaluation of these systems, it was found that none of the proposed systems is based on solid theoretical foundation and also does not take any quality of service attribute for forming the trust scores. Hence solid theoretical foundation for building trust systems for cloud computing is necessary. The theoretical basis required can be achieved by adapting the trust models proposed for other distributed systems.

References

[1] Sheikh Mahbub Habib, Sebastian Ries, and Max Mühlhäuser, "Cloud Computing Landscape and Research Challenges regarding Trust and Reputation," in *Symposia and Workshops on Ubiquitous, Autonomic and Trusted Computing*, Xi'an, China, 2010, pp. 410-415.

[2] Rajkumar Buyya, Chee Shin Yeo, Srikumar Venugopal, James Broberg, and Ivona Brandic, "Cloud computing and emerging IT platforms: Vision, hype, and reality for delivering computing as the 5th utility," *Journal of Future Generation Computer Systems*, vol. 25, no. 6, pp. 599-616, June 2009.

[3] Radu Prodan and Simon Ostermann, "A Survey and Taxonomy of Infrastructure as a Service and Web Hosting Cloud Providers," in *10th IEEE/ACM International Conference on Grid Computing*, Banff, AB, Canada, 2009, pp. 17-25.

[4] Christian Vecchiola, Suraj Pandey, and Rajkumar Buyya, "High-Performance Cloud Computing: A View of Scientific Applications," in *10th International Symposium on Pervasive Systems, Algorithms, and Networks (ISPAN)*, Kaohsiung, Taiwan, 2009, pp. 4-16.

[5] Michael Boniface et al., "Platform-as-a-Service Architecture for Real-Time Quality of Service Management in Clouds," in *Fifth International Conference on Internet and Web Applications and Services (ICIW)*, Barcelona, Spain, 2010, pp. 155-160.

[6] Han Yu, Zhiqi Shen, Chunyan Miao, Cyril Leung, and Dusit Niyato, "A Survey of Trust and Reputation Management Systems in Wireless Communications," *Proceedings of the IEEE*, vol. 98, no. 10, pp. 1755-1772, October 2010.

[7] Zaobin Gan, Juxia He, and Qian Ding, "Trust relationship modelling in e-commerce-based social network," in *International conference on computational intelligence and security*, Beijing, China, 2009, pp. 206-210.

[8] D Harrison McKnight and Norman L Chervany, "Conceptualizing Trust: A Typology and E-Commerce Customer Relationships Model," in *34th Hawaii International Conference on System Sciences*, Island of Maui, HI, USA, 2001.

[9] Wei Wang and Guo Sun Zeng, "Bayesian cognitive trust model based self-clustering algorithm for MANETs," *Science China Information Sciences*, vol. 53, no. 3, pp. 494–505, 2010.

[10] Mario Gómez, Javier Carbó, and Earle Clara Benac, "A cognitive trust and reputation model for the ART testbed," *Inteligencia Artificial. Revista Iberoamericana de Inteligencia Artificial (in English)*, vol. 12, no. 39, pp. 29-40, 2008.

[11] Huangmao Quan and Jie Wu, "CATM: A cognitive-inspired agent-centric trust model for online social networks," in *Ninth Annual IEEE International Conference on Pervasive Computing and Communications (Percom)*, Seattle, WA, USA, 2011.

[12] Cristiano Castelfranchi, Rino Falcone, and Giovanni Pezzulo, "Trust in information sources as a source for trust: a fuzzy approach," in *Proceedings of the second international joint conference on autonomous agents and multiagent systems (AAMAS '03)*, Melbourne, Australia, 2003, pp. 89-96.

[13] Stefano De Paoli et al., "Toward trust as result: An interdisciplinary approach," *Proceedings of ALPIS, Sprouts: Working Papers on Information Systems*, vol. 10, no. 8, 2010.

[14] Masoud Akhoondi, Jafar Habibi, and Mohsen Sayyadi, "Towards a model for inferring trust in heterogeneous social networks," in *Second Asia International Conference on Modelling & Simulation*, Kuala Lumpur, Malaysia, 2008, pp. 52-58.

[15] Ram Alexander Menkes, "An economic analysis of trust, social capital, and the legislation of trust," Ghent, Belgium, LLM Thesis 2007.

[16] Jie Zhang and Robin Cohen, "Design of a mechanism for promoting honesty in e-marketplaces," in *22nd Conference on Artificial Intelligence (AAAI), AI and the Web Track*, Vancouver, British Columbia, Canada, 2007.

[17] Jie Zhang, "Promoting Honesty in Electronic Marketplaces: Combining Trust Modeling and Incentive Mechanism Design," Waterloo, Ontario, Canada, PhD Theis 2009.

[18] Shashi Mittal and Kalyanmoy Deb, "Optimal strategies of the iterated prisoner's dilemma problem for multiple conflicting objectives," in *IEEE Symposium on Computational Intelligence and Games*, Reno, NV, USA, 2006, pp. 197 - 204.

[19] Jian Zhou, Jiangbo Wang, Rongshan Liang, and Yanfu Zhang, "Flexible service analysis based on the “Prisoner's Dilemma of service"," in *6th International Conference on Service Systems and Service Management (ICSSSM '09)*, Xiamen, china, 2009, pp. 434 - 437.

[20] Hongbing Huang, Guiming Zhu, and Shiyao Jin, "Revisiting trust and reputation in multi-agent systems," in *ISECS International Colloquium on Computing, Communication, Control, and Management*, Guangzhou, China, 2008, pp. 424-429.

[21] Lik Mui, "Computational models of trust and reputation:agents, evolutionary games, and social networks," Boston, MA, USA, PhD Thesis 2002.

[22] Mohammad Momani and Subhash Challa, "Survey of Trust Models in Different Network Domains," *International Journal of Ad hoc, Sensor & Ubiquitous Computing* , vol. 1, no. 3, pp. 1-19, September 2010.

[23] Tzu Yu Chuang, "Trust with Social Network Learning in E-Commerce," in *IEEE International Conference on Communications Workshops*

(ICC), Capetown, South Africa, 2010, pp. 1-6.

[24] Marcim Adamski et al., "Trust and Security in Grids: A State of the Art," European Union, 2008.

[25] Antonios Gouglidis and Ioannis Mavridis, "A Foundation for Defining Security Requirements in Grid Computing," in *13th Panhellenic Conference on Informatics (PCI '09)*, Corfu, Greece, 2009, pp. 180-184.

[26] Leonardo B De Oliveira and Carlos A Maziero, "A Trust Model for a Group of E-mail Servers," *CLEI Electronic Journal*, vol. 11, no. 2, pp. 1-11, 2008.

[27] Qing Zhang, Ting Yu, and Keith Irwin, "A Classification Scheme for Trust Functions in Reputation-Based Trust Management," in *International Workshop on Trust, Security, and Reputation on the Semantic Web*, Hiroshima, Japan, 2004.

[28] Ruichuan Chen, Xuan Zhao, Liyong Tang, Jianbin Hu, and Zhong Chen, "CuboidTrust: A Global Reputation-Based Trust Model in Peer-to-Peer Networks," in *Autonomic and Trusted Computing*. Berlin / Heidelberg: Springer, 2007, vol. 4610, pp. 203-215.

[29] Sepandar D Kamvar, Mario T Schlosser, and Hector Garcia-Molina, "The EigenTrust Algorithm for Reputation Management in P2P Networks," in *Proceedings of the 12th international conference on World Wide Web (WWW '03)*, Budapest, Hungary, 2003, pp. 640-651.

[30] Yong Wang, Vinny Cahill, Elizabeth Gray, Colin Harris, and Lejian Liao, "Bayesian network based trust management," in *Autonomic and Trusted Computing*. Berlin / Heidelberg: Springer, 2006, pp. 246-257.

[31] Huirong Tian, Shihong Zou, Wendong Wang, and Shiduan Cheng, "A Group Based Reputation System for P2P Networks," in *Autonomic and Trusted Computing*. Berlin / Heidelberg: Springer, 2006, pp. 342-351.

[32] Wei Wang, Guosun Zeng, and Lulai Yuan, "Ant-based Reputation Evidence Distribution in P2P Networks," in *Fifth International Conference Grid and Cooperative Computing (GCC 2006)*, Hunan, China, 2006, pp. 129 - 132.

[33] Yu Zhang, Huajun Chen, and Zhaohui Wu, "A Social Network-Based Trust Model for the Semantic Web," in *Autonomic and Trusted Computing*. Berlin / Heidelberg: Springer, 2006, pp. 183-192.

[34] Fajiang Yu, Huanguo Zhang, Fei Yan, and Song Gao, "An Improved Global Trust Value Computing Method in P2P System," in *Autonomic and Trusted Computing*. Berlin / Heidelberg: Springer, 2006, pp. 258-267.

[35] Weijie Wang, Xinsheng Wang, Shuqin Pan, and Ping Liang, "A New Global Trust Model based on Recommendation for Peer-To-Peer Network," in *International Conference on New Trends in Information and Service Science*, Beijing, China, 2009, pp. 325-328.

[36] Xueming Li and Jianke Wang, "A Global Trust Model of P2P Network Based on Distance-Weighted recommendation," in *IEEE International Conference on Networking, Architecture, and Storage*, Hunan, China, 2009, pp. 281-284.

[37] Xiong Li and Liu Ling, "PeerTrust: supporting reputation-based trust for peer-to-peer electronic communities," *IEEE Transactions on Knowledge and Data Engineering*, vol. 16, no. 7, pp. 843-857, July 2004.

[38] Ayman Tajeddine, Ayman Kayssi, Ali Chehab, and Hassan Artail, "PATROL-F - A Comprehensive Reputation-Based Trust Model with Fuzzy Subsystems," in *Autonomic and Trusted Computing*. Berlin / Heidelberg: Springer, 2006, pp. 205-216.

[39] Yuan Wang, Ye Tao, Ping Yu, Feng Xu, and Jian Lü, "A Trust Evolution Model for P2P Networks," in *Autonomic and Trusted Computing*. Berlin / Heidelberg: Springer, 2007, pp. 216-225.

[40] Zhuo Tang, Zhengding Lu, and Kai Li, "Time-based Dynamic Trust Model using Ant Colony Algorithm," *Wuhan University Journal of Natural Sciences*, vol. 11, no. 6, pp. 1462-1466, 2006.

[41] Felix Gomez Marmol, Gregorio Martinez Perez, and Antonio F Gomez Skarmeta, "TACS, a Trust Model for P2P Networks," *Wireless Personal Communications*, vol. 51, no. 1, pp. 153-164, 2009.

[42] Ghada Derbas, Ayman Kayssi, Hassan Artail, and Ali Chehab, "TRUMMAR - A Trust Model for Mobile Agent Systems based on Reputation," in *The IEEE/ACS International Conference on Pervasive Services (ICPS 2004)*, Beirut, Lebanon, 2004, pp. 113-120.

[43] Ayman Tajeddine, Ayman Kayssi, Ali Chehab, and Hassan Artail, "PATROL: A Comprehensive Reputation-based Trust Model," *International Journal of Internet Technology and Secured Transactions*, vol. 1, no. 1/2, pp. 108-131, August 2007.

[44] Felix Gomez Marmol, Gregorio Mrtinez Perez, and Javier G Marin-Blazquez, "META-TACS: A Trust Model Demonstration of Robustness through a Genetic Algorithm," *Autosof Journal of Intelligent Automation and Soft Computing*, vol. 16, no. X, pp. 1-19, 2009.

[45] Cesar Ghali, Ali Chehab, and Ayman Kayssi, "CATRAC: Context-Aware Trust- and Role-based Access Control for Composite Web Services," in *10th IEEE International Conference on Computer and Information Technology*, Bradford, England, 2010, pp. 1085-1089.

[46] Yao Wang and Julita Vassileva, "Bayesian Network-based Trust Model," in *IEEE/WIC International Conference on Web Intelligence (WI 2003)*, Halifax, Canada, 2003, pp. 372 - 378.

[47] Khaled M Khan and Qutaibah Malluhi, "Establishing Trust in Cloud Computing," *IT Professional*, vol. 12, no. 5, pp. 20 - 27, 2010.

[48] Zhexuan Song, Jusus Molina, and Christina Strong, "Trusted Anonymous Execution: A Model to RaiseTrust in Cloud," in *9th International Conference on Grid and Cooperative Computing (GCC)*, Nanjing, China, 2010, pp. 133 - 138.

[49] Hiroyuki Sato, Atsushi Kanai, and Shigeaki Tanimoto, "A Cloud Trust Model in a Security Aware Cloud," in *10th IEEE/IPSJ International Symposium on Applications and the Internet (SAINT)*, Seoul, South Korea, 2010, pp. 121 - 124.

[50] Wenjuan Li, Lingdi Ping, and Xuezeng Pan, "Use trust management module to achieve effective security mechanisms in cloud environment," in *International Conference on Electronics and Information Engineering (ICEIE)*, vol. 1, Kyoto, Japan, 2010, pp. 14-19.

[51] Tie Fang Wang, Bao Sheng Ye, Yun Wen Li, and Yi Yang, "Family Gene based Cloud Trust Model," in *International Conference on Educational and Network Technology (ICENT)*, Qinhuangdao, China, 2010, pp. 540 - 544.

[52] Tie Fang Wang, Bao Sheng Ye, Yun Wen Li, and Li Shang Zhu, "Study on Enhancing Performance of Cloud Trust Model with Family Gene Technology," in *3rd IEEE International Conference on Computer Science and Information Technology (ICCSIT)*, vol. 9, Chengdu, China, 2010, pp. 122 - 126.

[53] Paul D Manuel, Thamarai Selve, and Mostafa Ibrahim Abd-EI Barr, "Trust management system for grid and cloudresources," in *First International Conference on Advanced Computing (ICAC 2009)*, Chennai, India, 2009, pp. 176-181.

[54] Zhidong Shen, Li Li, Fei Yan, and Xiaoping Wu, "Cloud Computing System Based on TrustedComputing Platform," in *International Conference on Intelligent Computation Technology and Automation (ICICTA)*, vol. 1, Changsha, China, 2010, pp. 942 - 945.

[55] Zhidong Shen and Qiang Tong, "The security of cloud computing system enabled by trusted computing technology," in *2nd International Conference on Signal Processing Systems (ICSPS)*, vol. 2, Dalian, China, 2010, pp. 11-15.

[56] Mohammed Alhamad, Tharam Dillon, and Elizabeth Chang, "SLA-based Trust Model for Cloud Computing," in *13th International Conference on Network-Based Information Systems*, Takayama, Japan, 2010, pp. 321 - 324.

[57] Xiao Yong Li, Li Tao Zhou, Yong Shi, and Yu Guo, "A trusted computing environment model in cloudarchitecture," in *Ninth International Conference on Machine Learning and Cybernetics (ICMLC)*, vol. 6, Qingdao, China, 2010, pp. 2843-2848.

[58] Zhimin Yang, Lixiang Qiao, Chang Liu, Chi Yang, and Guangming Wan, "A Collaborative Trust Model of Firewall-through based on Cloud

Computing," in *14th International Conference on Computer Supported Cooperative Work in Design (CSCWD)*, Shanghai, China, 2010, pp. 329 - 334.

[59] Junning Fu, Chaokun Wang, Zhiwei Yu, Jianmin Wang, and Jia Guang Sun, "A Watermark-Aware Trusted Running Environment for Software Clouds," in *Fifth Annual ChinaGrid Conference (ChinaGrid)*, Guangzhou, China, 2010, pp. 144 - 151.

[60] Rohit Ranchal et al., "Protection of Identity Information in Cloud Computing without Trusted Third Party," in *29th IEEE International Symposium on Reliable Distributed Systems*, New Delhi, India, 2010, pp. 1060-9857.

[61] Hassan Takabi, James B.D Joshi, and Gail Joon Ahn, "SecureCloud: Towards a Comprehensive Security Framework for Cloud Computing Environments," in *34th Annual IEEE Computer Software and Applications Conference Workshops*, Seoul, South Korea, 2010, pp. 393 - 398.

International Journal on Advances in ICT for Emerging Regions 2011 04 (02) : 37 - 51

ICT Leap-frogging Enabled by Cloud Computing for Emerging Economies: A Case Study on Streamlining India's Grain Supply Chains

Jacob Tsao, Shailaja Venkatsubramanyan, Shrikant Parikh and Prasenjit Sarkar

Abstract—**Cloud computing enables use of computing without user-side hardware, software and associated financial and knowledge requirements, except for the needs of a simple "web-terminal" and access to the internet. Emerging economies can fully capitalize on this to start a "revolution" and leap-frog developed nations by skipping completely the several major computing architectures gone through in developed nations. An analogy is the leap-frogging, that has taken place with wireless telephony. We briefly introduce our research into streamlining India's grain supply chains and then discuss in detail on how cloud computing can play a pivotal role. Although the focus is on the portion of the supply chains between wholesalers and consumers, we also discuss how cloud computing can streamline the entire supply chains and how it offers leap-frogging opportunities for the entire society. The proposed use of web-terminals for accessing the cloud for all computing needs is new, and we point out several high-impact research subjects.**

Index Terms—**Cloud Computing, web-terminal, wireless access, Leap-frogging, Grain Supply Chains, Logistics, Hubbing, Branding, Developing Nations, India**

I. Introduction

Cloud computing has been motivated in developed nations to support the new paradigms of software as a service, hardware as a service, platform as a service, etc. [1,2]. However, it may be more beneficial to emerging economies like India. This is because countries like India can leap-frog developed nations in use of computing by directly migrating to the paradigm of cloud computing and skipping the less efficient intermediate computing architectures. Note that this direct migration can virtually free users from the need to own computing hardware (stand-alone or networked personal computers) and client-side software. All such intermediate architectures also require the expertise and/or labor about their installation, operation, trouble-shooting, maintenance, upgrade, etc. These cost and knowledge requirements have been a barrier for the general population in an emerging economy to benefit from computing and the Internet. Cloud computing possesses the potential for removal of this and other barriers. Although cloud-computing researchers have expressed the opinion that emerging economies can benefit from cloud computing, they tended to seek general application areas for cloud computing, e.g., government services, or suggested further development of technologies, e.g., voice-enabled web portal, that may facilitate the adoption of cloud computing by emerging economies [3]. We have had an entirely different motivation – to streamline and improve India's grain supply chain and, in the process, discovered that cloud computing is a promising technology for solving a key part of the problem [4].

A major problem in India's grain supply chains is spoilage. Spoilage rates in these supply chains have been consistently estimated to be between 25% and 30%, with their developed-nation counterparts being approximately 3% [3,4]. The rates of price "mark-up," i.e., service charges over crop costs, have been estimated to be over 240%, with their counterparts in developed nations being between 50% and 100% [3,4]. Moreover, food prices have been steadily growing. This points to not only the importance but also the urgency of streamlining grain supply chains of India. Improving India's grain supply chains has the potentials of better food quality and lower retail prices for the consumers and higher profits for the farmers. In addition, improving profit margins for the vast number of subsistence grain farmers may reduce population migration from rural areas to large cities; such migration has been attributed as a reason for the existence or expansion of some urban slums. This research was motivated to develop service, information technology and logistics concepts that can help streamline India's grain supply chains.

We focus on the portion of the supply chains between the wholesalers and the consumers, and this portion accounts for a whopping 210% of the 240% overall price mark-up [3,4].

Manuscript received on March 01, 2011. Accepted July 16, 2011.

Jacob Tsao is Professor of Industrial and Systems Engineering at San Jose State University. He taught and conducted research at S.P. Jain Institute of Management and Research of Mumbai, India for four months during his sabbatical leave in the 2009 – 2010 academic year. (jacob.tsao@sjsu.edu)

Shailaja Venkatsubramanyan is Associate Professor of Management Information Systems at San Jose State University. (shailaja.venkatsubramanyan@sjsu.edu)

Shrikant Parikh is with S.P. Jain Institute of Management and Research. (drsparikh@gmail.com)

Prasenjit Sarkar is with IBM Almaden Research Center, San Jose. (psarkar@almaden.ibm.com)

A layer of intermediaries (i.e., middlemen) exists between the wholesalers and the retailers, and they serve three major functions: checking prices offered by a large number of small wholesalers, checking quality of grains carried by the wholesalers, and paying for the merchandizes at the time of shipping on behalf of the retailers. Cloud computing has the potential of automating the functions of price checking and advance payment. It can also allow retailers to access directly information about many existing "private, unwritten brands," which are currently understood only between the wholesalers and the middlemen, and hence has the potential for automating the function of quality checking as well. Together with other improvements in supply-chain operations, particularly the concept of "consumer visible" branding, this layer of intermediaries can be drastically reduced in importance or even completely eliminated, hence significantly reducing the supply-chain costs. In addition to cloud computing and branding, the solution we proposed in [4] to help reduce the cost of the supply chains between wholesalers and consumers has another pillar – distributed hubbing for drastically reducing the distribution cost.

Unlike how cloud computing is used in developed nations [1,2], the two primary benefits of cloud computing for streamlining India's grain supply chains and improving the operations of many other systems in emerging economies are about the access to computing, not about the computing that can be performed in the cloud. A feature of cloud computing is the requirement of a "thin client," as originally promoted in the 1990's by Sun Microsystems. This feature is actually an unsung hero from the perspective of emerging economies. This is because the minimum requirement for access to the cloud is a "web-terminal." Such web-terminals serve only the purpose of accessing the internet through a browser and require little memory and computing power. They can be very inexpensive, unlike some of the sophisticated "Smart Phones" serving as popular high-end cell phones and mobile internet-access devices in developed nations. In addition, there is no need for owning software, whose purchase and upgrading could be expensive for the majority of the population of an emerging economy; the charge of on-demand software may be much more affordable. This points to the first primary benefit of cloud computing - low cost. The second primary benefit is low knowledge requirement. With the simplicity of such web-terminals and with the computing hardware and software applications residing in the cloud, the user is no longer required to install, maintain, recover, or upgrade either hardware or software. This architecture can be referred to as 'pure cloud computing' or simply 'pure cloud'.

Cost and knowledge requirements are two major barriers for the grain retailers to access and use computing, and such barriers led to their continued complete reliance on the intermediaries and, consequently, to continued high service charges in this portion of the supply chains. (Lack of public accessible price information also led to continued incidences of price-gouging by some merchants in this portion of the supply chains). The proposed "pure cloud computing" can drastically lower the two requirements and provide unprecedented opportunities for the retailers to benefit from computing and the Internet. This type of operational concepts provides a leap-frogging opportunity for the entire Indian society in general to more quickly than otherwise benefit from the power of computing by moving directly toward (public and/or private) cloud computing, bypassing the conventional steps of stand-alone computing, client-server architecture, networked computing, software as a service, hardware as a service, etc. As will be pointed later, although such leap-frogging is quite feasible for urban and suburban area, where access to the internet is already prevalent and hence is not a big issue, significant challenges exist for such leap-frogging to take place in rural or remote areas due to low accessibility to the Internet. However, we argue that similar leap-frogging directly to cloud computing for rural and remote areas should still be the goal because it should be much less costly than going through the stage of personal computing and the other traditional computing architectures. An analogy to this ICT leap-frogging is the leap-frogging having been taking place with wireless telephony. The market penetration for cell phones has reached 50% in just 10 years while its counterpart for land-line telephony is still 18% despite of 50 years of deployment. In fact, much of the benefit enabled by access to computing via the internet can be reaped, at least during initial deployment of the proposed operational concept, with the almost ubiquitous cell-phone technology and with some basic services like Google SMS Applications, particularly Google SMS Search [5,6].

The rest of this paper is organized as follows. Section II describes the problem of streamlining India's grain supply chains. Section III proposes a new operational concept as a solution for the problem and addresses its feasibility and benefits. It focuses on two pillars of the operational concept – "consumer visible" branding and distributed hubbing for the portion of supply chains between the wholesalers and the consumers, and the discussion is brief. Section IV focuses on the other pillar of cloud computing and addresses its feasibility and its leap-frogging benefits. Section V discusses briefly how cloud computing and ICT can significantly improve the operations of the rest of India's grain supply chains. Section VI discusses the success and hindrance factors for the proposed improvement via cloud computing. Section VII discusses the promising leap-frogging role cloud computing can play in drastically lowering barriers to use of computing for the entire society of an emerging economy. Concluding remarks are given in Section VIII.

II. INDIA'S GRAIN SUPPLY CHAINS

The produce supply chains of India have long been rather organized, with heavy well-intentioned government participation through provision of physical trading facilities and legal regulations. For example, each city has a Agriculture Produce Marketing Committee (APMC) that owns and operates one wholesale market for each of several major types of produce, e.g., grains, pulses, spices and perishable produces, at which wholesalers store, display to brokers (between the wholesalers and the retailers), trade with the brokers and transship the produce. Such markets of a city were originally intended for direct trading between farmers and retailers. As scales of agricultural activities grew larger and more specialized, such direct trading became difficult, if not impossible. For example, grains are grown mostly in Punjab and other northern states of India and transported to meet the consumer needs of all parts of India. Crops are consolidated by brokers (or directly transported by farmers) for sale to grain traders at the local APMC, called Mundi at the farmer or supplier side of supply chains. Typically, another layer of brokerage lies between these traders and milling companies, which sell their products to wholesalers through yet another layer of brokerage. Between the wholesalers and the consumers are the retailers and the brokers between the wholesalers and the retailers. These supply chains are illustrated in Figure 1.

We found few articles addressing this severe problem in international journals or conferences. After summarizing India's current grain supply chains and how the spoilage and mark-up rates accumulated through the chain, Sachan et al. [7] and Sachin et al. [8] proposed, with a System Dynamics approach, cost models for three supply-chain-integration alternatives for grains (namely cooperative supply chain model, collaborative supply chain model and contract farming). Few Indian journal or conference articles about grain supply chains were found, e.g., [9,10]. Several Indian agencies and news organizations have published research reports on these and general supply chains, e.g., [11,12,13].

A. Major Causes of the Problem Pointed Out in the Existing Literature

Major root causes of the problems of high spoilage and price mark-up pointed out in the literature include [12]:

- "Lack of adequate storage and transport infrastructure at the village level and right through the supply chain, which results in loss of output to rodents, pilferage, spoilage, etc.
- Presence of a large number of intermediaries, which results in a high mark-up to the end consumer".

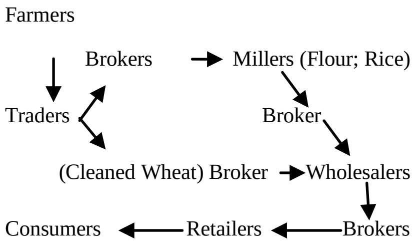


Fig. 1. India's grain supply chains

Indian government plays a big role in food supply chains. ICRA [12] states, "A plethora of laws, often overlapping in coverage, regulates the Indian food industry".

B. Implemented Partial Solutions and Current Improvement Efforts

In response to this severe problem, efforts have been made to alleviate it. A successful effort is eChoupal.com initiated and operated by ITC Ltd. of India [14]. eChoupal.com, accessed from personal computers through phone lines or satellites by trained farmers as conduits of information for farmers residing within a distance of approximately 5 kilometers, allows farmers to know the current commodity prices so that they can time the sale of their crops accordingly. ITC also purchases crops from farmers but with advanced equipment, e.g., automatic weight scale, quality-test lab, etc., and with a system of modern procurement practices, e.g., establishing quality standards based on numerical measures, quality-sensitive pricing, online price negotiation, etc. It also educates farmers on modern agricultural practices for better yield and better quality. It is important to note that much of this modernization would not have been possible without the IT infrastructure of eChoupal. Also, coverage of eChoupal expands as the IT infrastructure does. While eChoupal focuses on the farmer side of the supply chains, our focus is on the consumer side.

A small number of grocery supermarket chains have been founded in recent years and a small portion of the grain supply chains of India have been modernized. Reliance Fresh is one of them. However, accordingly to our sources, it continues to rely on wholesalers for its supply but D-Mart , another small supermarket chain, directly sources its grains from the millers or even the farmers, hence shortening the supply chains.

C. Our Focus and Improvement Opportunities

We focused on Mumbai, India for a case study and conducted a number of site visits to interview wholesalers, sales brokers for wholesalers (i.e., purchase brokers for the retailers), retailers, managers of APMC (which is the nationwide government agency regulating the grain and produce supply chains), and even purchase brokers for the wholesalers (i.e., sales brokers for the millers), among other key participants in the supply chains.

The large number of brokers involved in India's grain supply chains has been blamed for the high mark-up and spoilage rates. Although some brokers have engaged in price gouging, brokers intermediating between different pairs of other participants of the supply chains do serve multiple functions in the current supply chains. Take the retailers' perspective for example. The purchase brokers for a retailer (i.e., the sales brokers for a wholesaler) are actually needed in the current system to (i) investigate the highly variable quality levels offered by a wholesaler for one grain variety through time, not to mention the variability associated with a large number of wholesalers, (ii) search or negotiate for the best prices among many different wholesalers offering desired quality, and (iii) make advance payments to the wholesalers before collecting payments from the purchasing retailers (and bear the risk of non-payment by some retailers). The Mumbai APMC consists of three separate markets dedicated to three produce types: grains/pulses, spices and fresh produces. 600 wholesalers, 1500 brokers and 3500 trucks work at the grains/pulses market and serve all 15 million residents of Mumbai, India. Each of these 600 wholesalers occupies only a small, rectangular storage area of an approximately 1,700 square feet. The neighboring wholesale spaces share two walls; samples of up to over 50 varieties of grains are displayed on tables placed at the narrow storefront, which is approximately 20-feet long and opposite of the loading dock. The 600 stores occupy 20 virtually identical buildings.

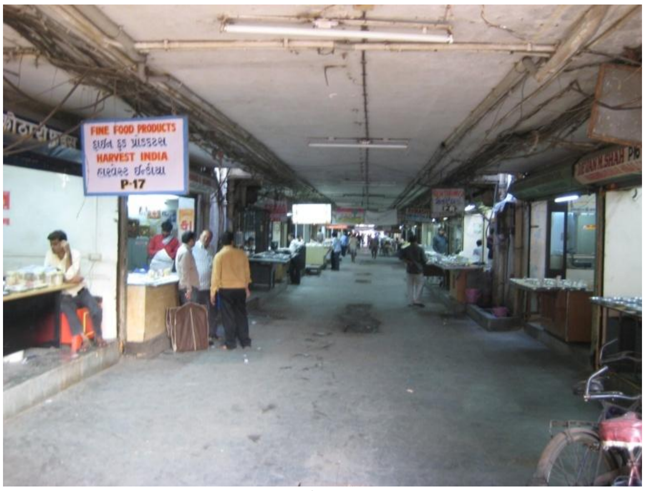


Fig. 2. 30 wholesale spaces in one of 20 identical buildings

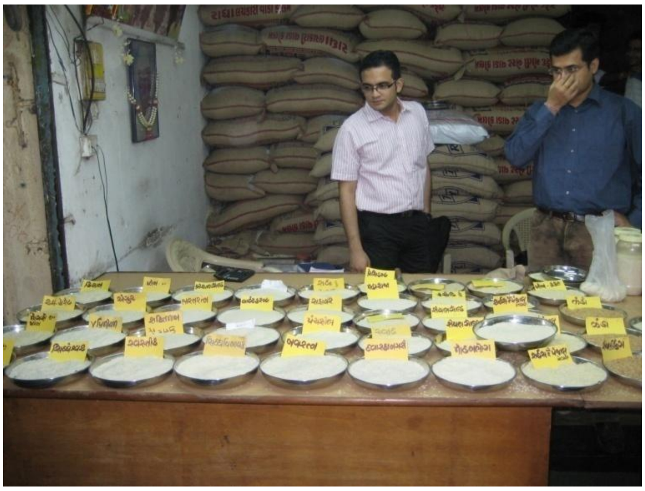


Fig. 3. A typical display of private brands at a wholesaler

A typical storefront display of sample merchandise consists of small trays labeled with “private brands.” Note that these “brands” are currently understood only between wholesalers and the middlemen (i.e., purchase brokers of retailers). Although each wholesale store is quite small, computer use is an integral part of the operation. In fact, Mumbai APMC develops software for the wholesalers. Retail operations tend to be small as well; a typical neighborhood retail store specializing in staple grains and pulses may have a storefront of 8 to 12 feet. It is safe to assume that operating a typical retailer does not require much knowledge. These motivated our search for a low-cost and low-knowledge-requirement ICT solution to help streamline the supply chains.

We note that according to the current regulations, APMC markets on the consumer side of the grain supply chains are the only locations, with few exceptions or “loopholes” where wholesale activities can be conducted and from which all the grains to be consumed by the city residents are transshipped. As cities and their populations grow larger, APMC markets outgrow their original confines, and some such markets have been relocated to city outskirts or even neighboring cities. For example, the APMC markets of Mumbai, including the grains/pulses market, spice market and fresh-produce markets, moved out of Mumbai altogether and into the neighboring city of Navi Mumbai. Serving the grain needs of a large city through a single out-of-the-city grain hub induces drastically more transportation and logistics than necessary. This motivates our service concept of distributed sub-hubs or sub-APMC markets to streamline the supply chains.

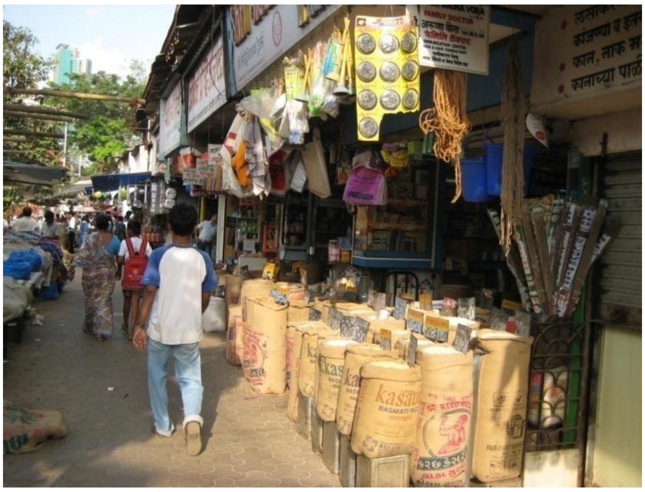


Fig. 4. A cluster of three typical grain retailers in Mumbai

III. A New Operational Concept

As mentioned earlier, the solution proposed for streamlining the supply chains has three pillars: cloud computing, distributed hubbing and “consumer visible” branding. In this section, we briefly describe an operational concept as a solution to the streamlining problem and focuses on the two non-IT pillars. The IT pillar, i.e., cloud computing, will be discussed in detail in the next section.

We observed that, except the Basmati rice (a special variety of rice), there are few “consumer visible” brands for grains, if at all. Currently, purchasing good-quality grains requires screening out low-quality products, and such a task is performed by the retailers’ purchase brokers, who poke grain bags with a sharp metal scoop to collect samples. Consumer-visible branding in India is being developed to some extent; the current burgeoning supermarket chains carry their own “brands” of “loose grains,” put them in barrels, and sell them by weight. This concept of consumer-visible branding, when implemented with other procedures and standardized practices, can help address the quality issues and eliminate the major function of quality screening currently served by the brokers between the retailers and the wholesalers. As mentioned earlier, if information about the "private brands" marketed by the wholesalers to the middlemen can be made available to the retailers or the general public by ICT, this middleman function can also be drastically reduced, if not completely eliminated.

Transportation accounts for a large percentage of the supply-chain cost. It is also a major source of green-house-gas emissions. The concept and practice in India of having only one centralized location in a city set aside to enable direct interaction between farmers and retailers dates back seven hundred years. These days, this location is the APMC. Although well-intentioned, such direct interaction rarely takes place, and farmers or retailers cannot afford the time to staff a counter at such a location. In addition, since grains are mostly grown in states in northern India, there are no grain farmers in or near Mumbai or most large Indian cities, and there is never any interaction between grain farmers and retailers there. Such a centralized location has now become, by law, the only location where grain and produce wholesale can take place and the hub to which all grain shipments from millers must terminate and from which all grain shipments to retailers must start. With only one centralized hub location for wholesalers to distribute grains to their retailer customers, the total distance traveled by the distribution trucks is several times than what would be required had there been multiple sub-hub locations. In large cities like Mumbai, the APMC are most likely located in the suburb or in a neighboring city. In such a case, the total distance traveled could even be higher. Concomitant with the unnecessary distance traveled are the unnecessary fuel burn, the resulting environmental impact, traffic congestion, and all the negative consequences associated with the congestion. We conducted a study to demonstrate that a distribution system consisting of four small "sub-hubs" could drastically reduce the total distance traveled by the distribution trucks and reported the findings in [4]. In the rest of this section, we briefly summarize the findings.

Figure 2 provides a bird's-eye view of the Mumbai metropolitan area, consisting of the Mumbai peninsula and Navi Mumbai. Note that the purpose of that study was not to suggest a new distribution system for Mumbai. Rather, it was to inform the designers of possible future grain-distribution systems of other large cities so that the transportation cost can be minimized, subject to non-logistics considerations.

The total cost of transporting grains from the source to Mumbai retailers consists of long-haul cost and distribution cost. We focus on transportation cost first and then address the facility costs, particularly the fixed costs associated with construction and operations of the hub or sub-hubs. Grains are produced in northern states of India, with the state of Punjab being a major grain producer state of India and the primary supplier of grain crops to Mumbai. Since Punjab is directly north of Mumbai metropolitan area and Navi Mumbai's latitude is at the mid-point between the latitudes of the northern and southern tips of the metro area, the total long-haul transportation cost associated with the current system should be approximately the same as its counterpart associated with the alternative four-sub-hub system. The distance traveled by distribution trucks was used as the proxy for the distribution cost.

The results are summarized in Table I. Note that the large ratio 4.86 of the total distance traveled under the four-sub-hub configuration over its current one-large-hub configuration points to a potential of drastic reduction of the recurrent cost of distribution. As for the fixed costs, the issues of land availability and traffic intensity associated with accommodating a large hub in a big city like Mumbai may be much more difficult to resolve than their counterparts associated with accommodating four small sub-hubs. We hope that this drastic numerical evidence will facilitate considerations for governmental policy changes.

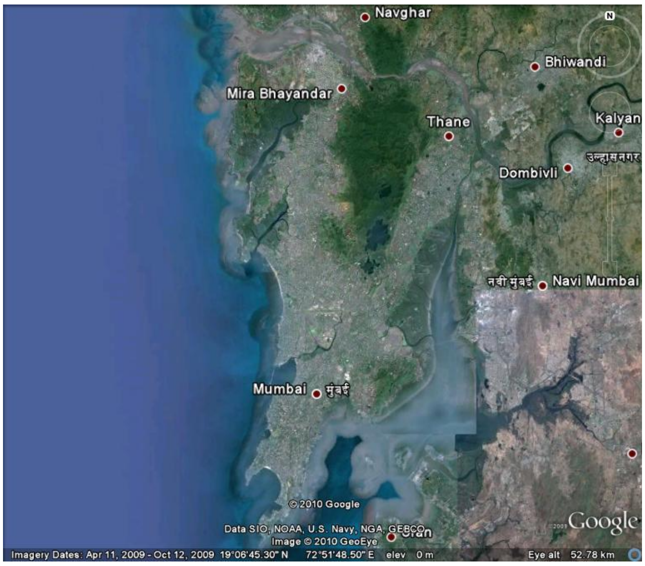


Fig. 5. Mumbai Peninsula and Navi Mumbai.

IV. ICT Leap-frogging Enabled by Cloud Computing

As mentioned earlier, a broker between a wholesaler and a retailer serves three major functions. One of them is quality checking, which was discussed in the previous section and has the potential of being automated with ICT. The two other

TABLE I
CONTRAST BETWEEN THE TOW SYSTEMS: DISTANCE TRAVELED
(TRUCKLOAD = 20 TONS)

Region	Distance to sub-hub	Distance to APMC	APMC/ Sub-hub
North West	134.5	627	4.66
North East	136.0	417	3.07
Central	65.2	479	7.35
South	74.8	473	6.32
All 4 Regions	410.5	1996	4.86
Truckload-KM	862,050	4,191,600	4.86

major functions served by the broker, namely checking prices offered by many different wholesalers and making advance payments to wholesalers, pertain directly to information technology (IT). Although India is well known for its IT industry, a large portion of the general public, particularly small merchants like grain retailers, may not have the required financial means and knowledge to acquire, install, operate, maintain, trouble-shoot and upgrade the IT hardware and software. We believe that the modern concept of cloud computing may work particularly well for India, particularly for the grain retailers. Although the concept of cloud computing originated in developed countries to deal with the issues of too much data (to be stored on personal or corporate computers), too much computing, sporadic computing needs, fast growth of computing needs, and/or to support the new paradigms of software as a service (or software on-demand), platform as a service or infrastructure as a service, it may be more beneficial to emerging economies like India. This is because countries like India can leap-frog the developed nations in use of cloud computing and avoid the financial investments and computer savvy that are required for the historically intermediate and incremental steps involving personal computing, particularly the ownership of computing hardware (stand-alone or networked personal computers) and ownership of client-side software.

Retailers in India or other small merchants can benefit from the wealth of information that could be posted on the internet, particularly the prices offered by various wholesalers, and from online payment. Note that India is on its way to catch up with the developed nations in telephony by skipping, i.e., leap-frogging, the step of land-line telephony and capitalizing on wireless communication, via cell phones. As mentioned earlier, while market penetration of land-line telephony in India is approximately 18% at this point after decades of deployment, its counterpart for cell-phone use has risen to over 50% in just 10 years. We stress that such leap-frogging may and should take place in computing or internet use in India and other emerging economies, first in individual special economic sectors and then in the larger society as a whole.

A. Service Characteristics

The key is to require user-friendly, fast and inexpensive price-checking and payment functions. From the perspective of a retailer, the services are provided via the internet. Therefore, the key business model must be at least the Application Service Providers (ASP) model, without requiring anything beyond a so-called "thin client." In fact, what is really needed is a so-called "web-terminal," which can be viewed, in terms of functionality, as a "dumb terminal" plus an internet browser. With such web-terminals, the grain retailers can access internet service applications via a land-line or wireless communication. Of course, a traditional desktop or laptop computer with internet access will serve the purpose as long as the concomitant capital and knowledge requirements pose no issues. From now on, we focus on the minimum requirement in terms of both capital cost and technical knowledge. Note, however, that a desktop- or laptop-based solution may work very well through sharing among individuals or through paid use of such hardware at a commercial facility, e.g., an internet café.

These web-terminals have also been referred to as "web-enabled" personal access devices in the US and other developed nations. Actually, the devices are not required to be "personal"; the retailers can use shared web-terminals devices to access the internet, as long as the sharing is convenient, economical and secure. To avoid confusion, we will use the term "web-terminals" to summarize the type of devices that are really needed.

From the service provider's point of view, software-as-a-service is required at the outset. As the business of such a provider grows, the provider may benefit from platform-as-a-service and infrastructure-as-a-service. In addition, as the service sector grows, such service systems must be scaled up and the services may be deployed on a private, community, public or hybrid cloud [15,16], reaping the benefit of scalability of cloud computing.

Making payment online or even through a cell phone is commonplace in India, although it is more popular among affluent and technology-savvy people. The system requirements for supporting this function are already well established and documented, and we will focus on information needed by the retailers for their own online price-checking and about their ability to recognize the private brands offered by wholesalers directly. In fact, the price is set for an individual private brand.

A wholesaler brands its merchandise according to several criteria. Although grain variety is a major and obvious criterion, there exist other major ones. Within a particular grain variety, some consumers prefer grains grown in particular regions. Some consumers prefer surprisingly rice grains that have been in storage for a significant amount of time, as long as the quality is not compromised. A rationale is that the level of water contents of such grains is reduced and therefore the rice when cooked is more "fluffy and "puffy." Such grains may also weigh less and may be perceived as less costly. Quality is of course another major criterion. The quality of the grains arriving at the storage of a wholesaler depends not only on the quality of the same grains when leaving the milling plant but also on the transportation process, through which the grains may be subject to quality deterioration caused by excessive moisture, high temperature, and insect and pest contamination. Different brands may reflect not just the varieties of the grains but also the quality levels. Information about all these brand criteria and the corresponding prices can easily coded in a database for browsing or search.

We assume that the services provided require mostly display of text and hence that the transmissions of text require small communication bandwidths. We next discuss two possible implementations, one with a land-line and the other without it but with wireless communication, regardless whether the use is shared or personal.

B. Assessment of Hardware and Software Requirements for Land-line Access to the Internet

The data-transmission rate for a typical telephone land-line in India is 46.6 kbps. It is ample for the retailers' needs for price checking and payment. Higher bandwidths are available and inexpensive. For example, the cost for broadband typically consists of a Rs. 300 monthly charge and a Rs. 3000 one-time charge ($1 ≅ Rs. 46). These should not be considered expensive by most retailers.

In the US and other developed nations, there do not seem to exist well-known brands/models of web-terminals that are used exclusively with land-line telephony. This may have been because the near-ubiquitous presence of desktop or laptop computers in businesses or even homes. The existence of such brands/models for wireless communication, as to be discussed below, may very well have resulted from the market for "mobile internet."

The basic requirements for a land-line-based web-terminal are simple. It consists of a monitor (for user interface), a communication device (hardware and software), and a browser (for display of and request for internet information). These days, computer monitors are well known to be very low-cost and highly reliable. So are the basic communication devices. Browser software is basically free. As a result, a land-line-based web-terminal should be very inexpensive and highly reliable, easing the capital and knowledge requirements to the minimum.

C. Assessment of Hardware and Software Requirements for Wireless Access to the Internet

Several well-known brands/models for wireless (mobile) web-terminals already exist in the US or other developed nations, although they are quite high-end when compared to the needs of the grain retailers. They include iPod, iPod Touch, Blackberry, Geode WebPAD, etc. Other brands/models also come with wireless voice communications, e.g., iPhone. A patent has been granted for a wireless web-terminal design that uses land-line through the base station of a cordless-phone. The rate of data transmission for typical basic cell phone in India is 10-30 kpbs. It is quite adequate for the retailers to check prices and make payment. If more bandwidth is required, a 256 kbps broadband connection can be purchased with a Rs. 3500 one-time charge and a Rs. 300 monthly minimum.

As indicated earlier, these brands/models have been developed for high-end consumers and may have much functionality that is not needed for our purpose here. The minimum requirements for a wireless web-terminal are similar to those for the land-line counterpart. One strategy is to reduce unnecessary functions of the high-end models just mentioned. However, another strategy is to capitalize on the functions already residing on a basic cell phone. One approach is to develop a "docking station" on which a basic cell phone can be placed and which is connected to a monitor for more user-friendly user interface.

D. Enabling Non-users of IT to Reap Benefit of Internet

The concepts proposed in this section may be applicable for other small merchants in India. They may also be applicable elsewhere. Given the security issues legitimately concerning IT managers of large corporations as well as the legacy-investment issues, many experts pointed out that the first adopters of cloud computing might not be big corporations but might be small to medium enterprises. Some even pointed out that a major contribution of cloud computing may be to attract current non-users of internet to become users. Russ Daniels, HP's CTO for Cloud Services Strategy, stated [15], "…we can make technology useful for a much broader group of people. It's not just consumers, but it's the people that today are nonconsumers, the people that aren't using technology because it's too complex, too expensive, too hard to get to, and that's really exciting…." Franco Travostino (Distinguished Architect of eBay) stated [15], "… I would say that entrepreneurs and small operations have to be the first beneficiaries given that clouds today don't have five-nine [99.999 percent] or seven-nine [99.99999 percent] dependability. Increasingly, we should see more of the Fortune 500 companies. They will be torn between using their own internal cloud within their own IT confines versus a real external cloud by an external provider which they do not have control over. …"

Note that internet-browsing capability has been or can be added to other common electronic devices, e.g., TV. This bodes well for web accessibility by the retailers and for more streamlined grain supply chains in India.

With the internet browsing capability of web-terminals, cell phones and the other common electronic devices discussed in this section, no conventional computers, either desktop or laptop computers, are required. This enables emerging economies' leap-frogging of developed nations in the sense that emerging economies can skip the expensive stages of stand-alone computing, client-server computing, software as a service, hardware as a service and infrastructure as a service directly to the drastically more efficient cloud computing. Privacy and some other issues associated with the use of cloud computing in developed nations may not be as important in emerging economies, at least comparatively.

V. Cloud Computing and ICT for the Entire Chains

In this section, we discuss briefly how cloud computing and ICT may significantly improve the operations of the rest of India's grain supply chains and how the potential improvement may address some critical and urgent issues facing the Indian society as a whole.

Recently, India's economic and thereby the political landscape has been rocked by high food inflation, which has been on the order of 17%+ in recent past. Several reasons have been put forward for this state of affairs. A key set of reasons revolves around inefficiencies in grain-food supply chains and low levels of visibility in the supply chains.

We now briefly discuss the portion of grain supply chains on the farmer's side. If the farmers have direct access to information about different prices offered by the traders/consolidators for the same commodity in the same or different APMCs and, for that matter, even have access to such information at the regional level, they can accordingly choose to sell the commodity in the 'right' market at a 'right' time and get a better price. This information can be made available directly to the farmers by way of providing them access to different information bases on the web. Such access to information can be made possible in a ubiquitous way by way of cloud computing, web terminals and the almost ubiquitous wireless connectivity.

In a similar fashion, the next set of traders in the supply chains and finally the very large wholesalers will stand to benefit if they have access to details about the crop planting and yield on a macro-scale. These details can include information like approximate size of area in which a particular crop (e.g., wheat) has been planted, the likely availability of the final crop to be sold at different APMCs in different parts of the country in different timeframes, and the progression of the commodities and their quantities through the supply chains. Such information in turn can be used to project availability of different commodities to be sold in different parts of country in different time frames. Another significance of this information is that the retailers and the wholesalers from which the retailers source different commodities will have a much better insight into supply of commodities and their quantities at different points in time at different relevant locations.

This transparency, i.e., the "visibility in extreme," will enable and encourage the stakeholders to trade and make commodities available at "fair market prices." Their trading decisions will be much more informed. The central point is, cloud computing, web terminals and wireless access enable this information to be universally available. Of course the information databases need to be made available on the cloud for easy access by the stakeholders.

Currently the prices at which the commodities are traded are lot more arbitrary and cannot be called "well informed". The small traders and their customers stand to lose most because of this state of affairs. Moreover, these prices are highly influenced by "dynamics" of the Commodity Exchange (MCX). It appears that a small group of traders on the exchange can have a high degree of influence on the price at which a particular commodity is traded, which in turn impacts the retail price of the commodity, hurting the customers.

A key assertion of this paper is that the unfair pricing which can be determined by improper or even speculative means can, to some extent, be neutralized by universal cloud-computing-based agricultural databases and wireless access to them.

VI. Success and Hindrance Factors

In this section, we discuss possible success and hindrance factors for implementation of the proposed use of cloud computing in India's grain supply chains. Note that although our proposal to post the price information on the internet was motivated to replace a service currently performed by the middlemen between the wholesalers and the retailers, the information is perhaps more useful for the consumer and the APMC. We first discuss factors that would facilitate the implementation and then factors that may hinder the implementation.

A. Possible Success Factors

As pointed out earlier, APMC of a city plays a pivotal role in the operation of grain supply chains. A primary goal of APMC is to ensure market efficiency in general and to prevent price gouging in particular. Mr. Sudhir Tungar, the Principal Secretary of Mumbai APMC, informed us during an interview of a recent success in detecting and resolving a price-gouging incident in the district of Colaba in South Mumbai. Local retailers told their customers about a non-existent crop shortage and charged prices that were about five times the prevailing wholesale prices charged at the Mumbai APMC approximately only 36 km away. APMC sent truckloads of the merchandise to the district immediately after detection. The public apprehension immediately subsided, and the price gouging ceased. He also informed us that the large price mark-ups might also result from retailer speculation about supply volatility. Had wholesale prices or the prevailing retail prices been posted on the internet, price-gouging incidents like this one would not have occurred. An informal survey of our colleagues and some students confirmed our conjecture that most grocery shoppers in India really want to know a fair price range for each major staple grain, pulse (e.g., beans) or vegetable. In addition, major success factors for eChoupal include the farmers' strong desire to know fair price ranges for their crop and the ability of eChoupal to provide the information [14]. We believe that APMCs' and consumers' desires for transparency in price information will help propel the realization of the proposed use of cloud computing via a web-terminal.

It should be informative to discuss successful implementations of cloud-computing-based or closely related initiatives or in India. eChoupal was developed when the client-server architecture was the state of the art in India. More importantly, it relies on expensive communication with rural farmers, via landline or microwave, for downloading educational video and other services requiring substantial bandwidth. Although client-side software may be necessary, the service can be considered cloud computing. It is conceivable that this internet-based service can be provided via a web-terminal, i.e., its resident browser, if the communication bandwidth is sufficient and advanced video technology is employed. The more important things are the value of price information to farmers and the farmers' willingness to pay for the information; they have helped propel the demand for and the success of eChoupal. Upton and Fuller [14] detailed how and how much farmers have benefited from the transparency of the price information. We believe that their counterparts on the consumer side will help propel the implementation of the proposed use of cloud computing.

Farmer lobby is very large and politically strong. Therefore, anything that can help farmers improve their livelihood becomes a favorite cause for many politicians. For example, in a recent budget, the Central Government of India pardoned a large part of rural agricultural debts. Farmers have complained about the extortionate service charges piled on their meager crop revenues for decades; so have the consumers. Streamlining any portion of the grain supply chains will benefit the farmers and the consumers. Therefore, we believe that, in terms of political support, the proposed cloud-computing-based operational concept on the consumer side will be popular for the society at large.

As for price information on the consumer side of the supply chains, Indian Harvest [17] currently provides daily spot prices of various agriculture commodities, but only for several commodity trading centers in India. Our goal is to develop a feasible and attractive operational concept that can be implemented in the near term to provide informational transparency on the consumer side and hence complement eChoupal in streamlining India's grain supply chains.

We now turn our attention to successful implementation of cloud-computing-based initiatives on the consumer side of grain supply chains in the US. This portion of the supply chains has long been fully integrated and dominated by few large supermarket chains, e.g., Safeway. The kind of wholesalers and retailers currently in operation in India have long disappeared in the US. As a result, the streamlining issue being dealt with in this paper does not exist in the US, and there are no comparable initiatives currently or in the recent past. For decades, US supermarket chains have had their own internal MIS and IT departments. With the deployment of advanced technologies like virtualization, the IT currently supporting this portion of the supply chains can be characterized as a private cloud or an "internal cloud" managed by a centralized internal authority [1].

Current US searchable/queryable agriculture databases are mostly implemented as cloud computing, with search or query requests specified and transmitted through an internet browser (without any other client-side functionality) and with the searches or queries performed in the cloud, free of charge. We give several examples for such agriculture databases to demonstrate their technical feasibility and popularity.

The internet service provider agriculture.com [18] provides daily and intraday price data; agricommodityprices.com [19] provides periodically updated commodity-specific and country-specific price data. We note, however, that other databases exist but some of them have been provided mostly for research purposes, instead of for making routine purchasing decisions. An example is the database provided by the Food and Agriculture Organization of the United Nations. In particular, FAOSTAT [20] provides time-series and cross sectional data related to food and agriculture for some 200 countries, and PriceSTAT [21] supports queries for the annual average price for a commodity, e.g., Rice (Paddy), in a country, e.g., India. The National Agricultural Statistics Service (NASS) of the US Department of Agriculture maintains a similar database, but with a focus on the US and with more details [22]. www.indexmundi.com [23] provides average monthly prices of rice and other commodities of a country, among a large amount of data constituting a country profile.

Most major supermarket chains in the US publish their prices on the internet (i.e., in the cloud) and some of them even take online orders for home delivery or pick-up at store. All these support our key assertion that the unfair pricing can, to some extent, be neutralized by universal, cloud-computing-based agricultural databases and wireless access to them.

B. Possible Hindrance Factors

We now discuss possible factors that may hinder the implementation of the proposed cloud computing initiative. Although it is possible that retailers may be unwilling to use computers or "web-terminals" to access price information posted on the internet, we believe that it is not likely. There has been no particular resistance to usage of computers in India even in rural areas - particularly when there is a good reason to use it. Computer training institutes imparting the basic computer skills (e.g. office, basic programming, basic maintenance skills) are ubiquitous even in C, D class cities and smaller towns because they are so popular and in demand.

Let us address the more general possible issue of resistance to technology adoption. There are no better examples than those involving rural populations. There are several instances where rural populations have adopted technology without much resistance, with adoption of Automated Teller Machine (ATM) being a simple example. While the fishermen of the Indian state of Kerala are returning from their fishing trips, they make cell-phone calls to check prices offered at different on-shore markets and decide on their landing spots accordingly. This also shows a strong desire for price information. In addition, basic information access is possible and inexpensive with no-frill mobile phones and of course with wireless–internet connected thin clients. The price of a mobile phone supporting GPRS-enabled internet connection is approximately Rs.3000, with an approximate GPRS (general packet radio service) connection cost of Rs.100 per month (with unlimited usage). In general, we do not see any real resistance to technology, per say. As long as the technology provides value and is easy to use and afford, it will be embraced by the target users.

Currently, cellular telephony covers over 90% of the 1 billion Indian population, with over 700 million subscribers. Even if the proposed cloud-computing-based initiative experiences initial resistance to use of web-terminals or computers, basic price information can be obtained and payment can be made via a regular cell phone. Internet searches or queries made via cell-phone text messaging have long been supported, e.g., Google SMS Applications [5]. Google SMS Search [6] can be used to get information on driving directions, sports, movies, stocks, definitions, etc; these services are free from Google but message & data rates may apply. Note that the required technology is the regular wireless telephony, not the newer and more expensive 3G or 4G technology. Note also that we are not suggesting the use of SmartPhones for accessing the internet. They are very expensive to begin with and are marketed for affluent and technology-savvy users. Also, many applications, i.e., the “Apps,” require significant client-side functions.

We now address the possible resistance to the proposed cloud-computing-based initiative from the wholesalers, the retailers and the middlemen. A retailer we visited informed us that his family and the family of his procurement agent (i.e., middleman between him and the wholesalers at the Mumbai APMC) have cooperated for generations. As long as he can make a reasonable profit, whether his agent makes a lot of profit or not does not really concern him. Although he has not felt any urgent need to find an alternative way of securing grains, he would be happy to try out the prototype technology and user interface we planned to develop. As mentioned earlier, the proposed information transparency can also prevent retailer price gouging. It can also introduce or at least encourage fair competition among local retailers. These may not be welcome by the middlemen and retailers alike. However, as discussed earlier, we believe that demand for informational transparency by the consumers, the farmers and APMC officials may entice or even force changes required for streamlining the grain supply chains of India.

Resistance from middlemen is expected. Possible replacement work for them include quality inspection, certification or assurance for the supply chains. However, the work would not be performed for individual retailers on specific retailer purchases but perhaps for APMC or a certification body. As for the threats to existing traders (not just the middlemen between the wholesalers and the retailers) and trading relationships, these problems have been overcome before, e.g. by eChoupal. ITC, the sponsor of eChoupal, has found new roles for commission agents and has successfully managed any potential resistance from them.

VII. Leap-frogging for the Entire Society in General

Technical merits of a technology alone cannot guarantee its market success; technology deployment and management of technology is a critical issue. Although cloud computing was not invented for emerging economies, e.g., for the need of overcoming barriers against internet access by people without affluence or computer savvy, it does present many opportunities for emerging economies. The transportation sectors of developed nations have experienced a similar situation in the past two decades. Advances in computing, communication and automation technologies spurred a intense level of interest in applying the technologies to improve transportation systems. A discipline called Intelligent Transportation Systems (ITS) emerged. Many application possibilities, called user services, were developed and studied. Some initiatives originally regarded as the most promising ones failed miserably, due to inattention to deployment issues. These failures as well as the necessity to streamline many concurrent research and development efforts and to minimize risks, a framework for organizing the many deployment issues at the outset of research and development was developed [24]. Developing such a framework for application of advanced computing technologies to emerging economies may be a worthy research topic.

Cloud computing can enable an entire society to take advantage of the economic benefits. This applies not only to the urban and suburban areas that have access to high-bandwidth Internet access, but also to the remainder of the developing world that have minimal connectivity to the rest of the world. In this section, we first deal with these two cross-sections of society and deal with the independent issues that are relevant to each of them. We then point out the leap-frogging opportunity offered by the concept of internal cloud to replace desktop or laptop computing or skip it all together and discuss applications of cloud computing in other sectors of the Indian society.

A. *Areas with High-bandwidth Internet connections*

Societies with high-bandwidth Internet connections, particularly the developed nations, use cloud computing with two specific aims: first, to reduce the cost of application development; and second, to avail of services that have costs that are lower than that built from traditional delivery means. The lower service costs result mainly from a new service delivery and charging paradigm for software applications. For urban and suburban areas of an emerging economy, the costs of accessing various software services can be further reduced due to reduced infrastructure and development costs, if the kind of web-terminals discussed in Section 4 are used. Note that the infrastructure costs include those incurred for needs assessment, system specification, equipment selection, purchase, installation, maintenance, upgrade, repair, security, and many other life-cycle concerns. Assessing overall cost savings achievable via this leap-frogging is a critical research issue for emerging economies. This research was never needed in developed nations and hence has not been paid any attention. As discussed earlier, the lower requirements for hardware and software knowledge associated with cloud computing are another main source of benefits that emerging economies can fully capitalize on. Although the benefit of lowered knowledge requirements on accessing the internet is much more intangible than the benefit of lower costs, assessing this benefit is also an important new research subject.

Lower cost of application development allows providers to bring applications to the market much faster than what was possible previously. This allows a larger segment of the population and even individuals to contribute applications to the ecosystem, while previously this was the domain of medium to large software houses. We can illustrate this by looking at the costs of two Cloud providers. As summarized in Table II, Google App Engine provides a lot of free resource time for application developers – these free resources can be used to mitigate the cost of application development. In addition, Google App Engine provides development tools for free.

In addition to these free resources, cloud providers such as Microsoft Azure and Google App Engine provide low development and hosting costs for cloud applications beyond the free time. See Table III for a comparison.

The next logical step is to compare these costs to that incurred by those using traditional development means. Assuming that medium to large software houses have consolidated their operations in a mid-size data center, researchers at the RAD lab in Berkeley have examined costs and found that the infrastructure costs in a large cloud data center virtualized to different application developers are five to seven times lower than that in the mid-size data centers used by the software houses. This gives a tremendous advantage to the next wave of application developers who are not burdened with the cost of maintaining these mid-size data centers.

These translate to lower cost of goods for cloud services developed as a result. Researchers have audited financial statements to examine claims that service-oriented architecture

(SOA) leads to higher profits relative to traditional software delivery models [25]. Specifically, they have examined vendors that rely on the Software-as-a-Service (SaaS) pricing model, and compare their performance to other firms that still use the traditional perpetual license model. The researchers find that, relative to their peers, SaaS firms tend to have lower costs of goods sold as a portion of revenues.

TABLE II
FREE DEVELOPMENT RESOURCES IN GOOGLE APP ENGINE

Category	Free Resource (Daily Rate)
CPU	6.5 hours
Bandwidth out	1 GB
Bandwidth in	1 GB
Storage (database)	1 GB
Storage Transactions	10 Million

Even non-traditional scientific applications get an economic benefit from running in the cloud, which could bolster scientific development in developing nations. Kondo et al. [26] did a survey of grid computing applications for scientific processing and tried to determine the cost-benefits of cloud computing versus volunteer computing applications. The authors calculated overhead for platform construction, application deployment, compute rates, and completion times. Given a best-case scenario, the authors found that the ratio of volunteer nodes needed to achieve the compute power of a small Amazon EC2 instance is about 2.83 active volunteer hosts to 1.

TABLE III
DEVELOPMENT RESOURCES COST IN MICROSOFT AZURE AND GOOGLE APP ENGINE

Category	Windows Azure	Google App Engine
CPU	$0.12/hour	$0.10/hour
Bandwidth out	$0.15/GB	$0.12/GB
Bandwidth in	$0.10/GB	$0.10/GB
Storage (database)	$0.15/GB/month	$0.005/GB/month
Storage Transactions	$0.01/10,000	Not Available

One potential drawback of cloud computing is that complex graphical tasks need dedicated computing power on a desktop machine with a powerful GPU. In such a scenario, these applications would not lend themselves to cloud computing and potentially increase development costs in emerging economies. However, even in such a scenario, the advent of multi-core GPUs may enable virtualization of resources for cloud computing users and allow them to avoid traditional application development costs. Lin and Wang [27] have proposed a cloud computing framework for such domains, with the end users only has a relatively inexpensive thin terminal with a high resolution screen and I/O devices.

B. Areas with Limited Connectivity

Although such leap-frogging is quite feasible for urban and suburban areas now or in the near future, where access to the internet is already prevalent and hence is not a big issue, significant challenges exist now for such leap-frogging to take place in rural or remote areas due to low accessibility to the internet. However, we argue that similar leap-frogging directly to cloud computing should still be the goal because it should be much less costly than going through the stage of personal computing and the other traditional computing architectures.

This points to a critical area for researchers to enable direct realization of the cloud computing paradigm in this section of society. We again remind the reader of the analogy between IT leap-frogging and telephony leap-frogging. With the advent of wireless-communication technology, not only is there much less need for further development of costly wired telephony, the strategy for further developing telephony and the resulting hardware/software architecture can be revamped to fully capitalize on the wireless technology. In the realm of IT, the IT development strategy and architecture for rural or remote areas of an emerging economy can be revamped in a similar way. A critical research subject is how to allow efficient access to the internet or the cloud in rural or remote areas of an emerging economy without the equipment required in the conventional architectures through which the cloud computing evolved in developed nations. More precisely, how can such access be enabled without PCs and laptops? Note that this issue has not been tackled by the developed nations because owning PCs, laptops, tablet computers, smart-phones and other advanced devices has not been and will most likely not be an issue. The reader is again reminded that in addition to costs, knowledge requirements are another major challenge for access to the cloud in rural and remote areas of an emerging economy. The examples given below tackle these issues with actual research projects in developing economies and use innovative techniques to bridge the knowledge gap.

Furthermore, looking at the past, one can look to the evolution of the PC industry as an example of how hardware costs and knowledge gaps can be overcome. The concept of personal computing did not truly exist until Apple introduced the Macintosh in 1984 – a computer that got rid of command prompts in favor of a graphical user interface. The Macintosh did not require a huge learning curve like previous DOS command-line driven machines. For the first time, computers became usable by normal people – not just the professionals. Microsoft learnt from this new concept of graphically-driven computing, following up with Windows in 1986, broadening the market so that almost everyone can afford a computer from different hardware manufacturers in developed nations. By 1995, most households in the US had a PC, and it is not unreasonable to speculate that the combination of wireless and mobile handset technologies can achieve the same for the emerging regions given that already mobile usage has crossed the 2 billion population mark.

A key inhibitor to cloud computing in emerging economies is the absence of connectivity to the cloud. Many developing regions around the world, especially in rural and remote areas, require low-cost network connectivity solutions. Traditional approaches based on telephone, cellular, satellite or fibers have proved to be an expensive proposition especially in low population density and low-income regions. In Africa, even though cellular and satellite coverage is available in rural regions, bandwidth is extremely expensive due primarily to low user densities (satellite usage cost is about US$3000 per Mbps per month). WiMax, another proposed solution, is currently also very expensive and has been primarily intended for carriers (like cellular). WiMax is hard to deploy in the “grass roots” style typical for developing regions. WiFi-based Long Distance (WiLD) networks are emerging as a low-cost connectivity solution and are increasingly being deployed in developing regions. The primary cost gains arise from the use of very high-volume off the shelf 802.11 wireless cards, of which over 140 million were made in 2005. These links exploit unlicensed spectrum, and are low power and lightweight, leading to additional cost savings. These networks are very different from the short-range multi-hop urban mesh networks. Unlike mesh networks which use omni-directional antennas to cater to short ranges (less than 1–2 km at most), WiLD networks comprise of point-to-point wireless links that use high-gain directional antennas (e.g. 24 dBi, 8 degree beam-width) to focus the wireless signal (for line of sight) over long distances (10–100 km).

To extend this connectivity to a large population of users, the developers of WiLD proposed cellular phones as a medium of connectivity. Cellular communications, including handsets and base stations, have become ubiquitous technologies throughout the developing and developed world. Roughly three billion users spend large portions of their income on these basic communications. However, the remaining half of the world currently has limited access, in large part due to lack of network coverage. Some areas do not have a high enough population density to support a traditional cellular deployment. Other areas are too far from established infrastructure to make a deployment economically feasible. This leads to many rural areas where there is no network coverage at all. To resolve this issue, the WiLD developers propose the Village Base Station (VBTS), which provides four main benefits:

- flexible off the grid deployment due to low power requirements that enable local generation via solar or wind;
- explicit support for local services within the village that can be autonomous relative to a national carrier;
- novel power/coverage trade-offs based on intermittency that can provide bursts of wider coverage; and
- a portfolio of data and voice services (not just GSM).

VBTS is essentially an outdoor PC with a software-defined radio that implements a low-power low-capacity GSM base station. Long-distance WiFi provides “backhaul” into the carrier. At around 20W, its power consumption is low enough to avoid diesel generators and the corresponding requirement for roads and fences. This also reduces the operating costs significantly. The base station can be deployed in the middle of the village, on a nearby hill, or in any other area with line-of-sight coverage. Although much of the contribution of VBTS is engineering the combination of a software radio, WiFi backhaul, and local generation, there are two main societal contributions:

- the development of a platform for a wide range of services
- the optimization of coverage versus power consumption via variable power and intermittent coverage

The connectivity has led to the development of unique solutions for the general population, one of which covers the field of education. English is the language of power in India associated with the middle and upper classes. In other developing regions, it is another language such as Spanish, Mandarin, or French which is not native to most of the population. The public school systems in developing regions face insurmountable difficulties. In India, for example, it has been consistently difficult to converse in English with those teachers responsible for teaching English in poor schools, where the overwhelming majority of children in the country struggle to learn. More important, public schooling is out of the reach of large numbers of children in rural areas and the urban slums who cannot attend school regularly, due to their need to work for the family in the agricultural fields or households. At the same time, cell phones are increasingly adopted in the developing world, and an increasing fraction of these phones feature multimedia capabilities for gaming and photos. These devices are a promising vehicle for out-of-school learning to complement formal schooling. In particular, the English learning games on cell phones present an opportunity to dramatically expand the reach of English learning, by making it possible to acquire ESL in out-of-school settings that can be more convenient than school. Games can make learning more engaging while incorporating good educational principles. More important, a large-scale evaluation with urban slums children in India has shown significant learning benefits from games that target mathematics. The developers of the English learning project believe that similar outcomes can be replicated with e-learning games that target literacy. The challenge in evaluating any language learning project, however, is that the language acquisition is a long-term process on the learner's part. Worse, with a novel technology solution that has yet to be institutionalized, there are tremendous logistical obstacles in running a pilot study over a non-trivial duration. After 3 years, in which the developers commenced with needs assessments and feasibility studies, followed by subsequent rounds of field testing interleaved with numerous iterations on our technology designs, they have established the necessary relationships with local partners for such an evaluation.

Another approach to connectivity in emerging economies is the shared usage of scarce networking bandwidth. Computer scientists in Pakistan are building a system to boost download speeds in the developing world by letting people effectively share their bandwidth. Software chops up popular pages and media files, allowing users to grab them from each other, building a grassroots Internet cache. In developing countries, almost all the traffic leaves the country. That's the case even when a Pakistani user is browsing websites hosted in his or her own country. The packets can get routed all the way through a developed nation and then back to a developing country. So a team in Pakistan is developing DonateBandwidth, a system inspired by the BitTorrent peer-to-peer protocol that is popular for trading large music, film, and program files. With BitTorrent, people's computers swap small pieces of a file during download, reducing the strain placed on the original source. DonateBandwidth works in much the same way but lets people share more than just large files. When users try to access a website or download a file, a DonateBandwidth program running on their machine checks first with the peer-to-peer cache to see if the data is stored there. If so, it starts downloading chunks of the file from peers running the same software, while also getting parts of the file through the usual Internet connection. The software could allow people in countries that have better Internet connections to donate their bandwidth to users in the developing world. DonateBandwidth also manipulates an ISP's cache. When running DonateBandwidth, a computer starts downloading part of a file, while also sending a request for other DonateBandwidth users who have access through the same ISP, and whose computers have spare bandwidth, to trigger them to start downloading other parts of the same file. The file is then loaded into the ISP's cache, so it can be downloaded more quickly. The project is similar to distributed computing schemes such as SETI@Home, which uses volunteers' spare computer power to collaboratively analyze radio signals from space, looking for signs of intelligent life. DonateBandwidth permits sharing of unused Internet bandwidth, which is much more valuable in the developing world, compared to computing cycles or disk space.

C. Potential of Internal Cloud for Non-transaction IT

In developed nations, many transaction-based enterprise IT applications involving large databases have already been or are being migrated from the client-server architecture to cloud computing, particularly internal cloud. The task of such migration is not very difficult because the uses of the applications are already centralized and the primary task is to implement the client-server interaction with the browser technology. However, many organizations in developed nations, and particularly those in emerging economies, whose primary IT needs are not database-oriented transactions can also benefit from use of internal cloud computing.

The main idea is to have all resident software applications (not supported by an external cloud) migrated to an internal cloud, i.e., centralized and virtualized servers. This way, there is no need to have a desktop computer on top of every desk. Just replace every desktop computer with a web-terminal. The web-terminals will need virtually no IT support. All the efficiency benefit achievable through an external cloud can be achieved with such an internal cloud, except for the larger economy of scale. Obviously, this internal cloud can be supplemented with access to an external cloud. For example, free software applications for office work are widely available in the internet, if they are not provided in the internal cloud. Such free applications include Google Doc, Google Spreadsheet, etc.

The proposed concept of leapfrogging by means of performing *all* computing in the cloud and accessing the cloud *only* through a *simple web-terminal* can be referred to as 'pure cloud computing' or simply 'pure cloud.' This new term of 'pure' cloud is intended to articulate a new concept of cloud computing and differentiate its minimum client-side hardware and software requirements from the higher client-side requirements of any other cloud-computing implementation. 'Pure cloud' is based not only on cloud computing but even an easier, more easily deployable and therefore more widely acceptable implementation of cloud computing where the prerequisites needed to operate the clients are almost nothing. Only a simple browser-based access to the cloud suffices, and it can be supported on thin, simple and low-cost web-clients. This 'pure cloud' concept will be particularly beneficial for emerging economies.

Migration of the non-transaction IT uses to an internal or external cloud may be beneficial for any organization in general but may be particularly beneficial for large non-profit organizations like schools and universities, hospitals, etc. When such migration is implemented in elementary through high schools, in developed nations or emerging economies, potential benefits include not only reduction of hardware, software and labor costs but also increased ability of schools to monitor or even control of proper student access to internet contents. Note that video-game playing and internet browsing unrelated to classroom learning may severely distract students. Other potential benefits include reduced opportunities for equipment theft or damage. More importantly, due to the significantly lower hardware, software and labor costs associated with the proposed internal cloud architecture, schools may finally be able to provide in-class access to computing and the internet.

D. Cloud Computing for Other Services or Countries

We discuss in this sub-section cloud-computing-based initiatives in other sectors of India or in other countries. Due to space limitation, we focus on the health industry. After a brief discussion of application of cloud computing in the US, we illustrate leap-frogging opportunities for emerging economies with one project proposal submitted to the Pakistan government and one specific start-up company in India.

It is well known that the US is experiencing a daunting healthcare crisis. Innovative IT has been regarded as a major improvement that can significantly reduce cost and increase quality. Among the better known innovations is the electronic medical record (EMR) or electronic health record (EHR). To facilitate coordination and even direct communication among different healthcare providers involved in the care of a patient, most of the EMR implementations are based on cloud computing, particularly internal cloud.

Cloud-based EMR presents an obvious leap-frogging opportunity for emerging economies. What has not been discussed in the developed nations is the leap-frogging opportunity to "pure cloud," mentioned earlier about exclusive use of web-terminals, in lieu of desktop computers, for cloud access. The first author participated in a three-partner team effort in response to a request for proposals "Establishing Center of Innovation for Use of ICT in Healthcare Sector" issued by the National ICT R&D Fund of the Ministry of Information Technology of Pakistan. The proposal features the use of cloud computing, including the concept of "pure cloud," as a main theme of the proposed research. The funding decision is yet to be made. After a nearly one-year halt of operations, the National ICT R&D Fund recently resumed operations [28].

The third author is a co-founder of an Indian company called A3 BioMed Technologies Ltd [29]. The company is in business for providing solutions and services that will enable cardiologists to remotely monitor their patients via mobile phone from anywhere in world. The solutions are based on wireless devices connecting the patient to a server. The company provides services, based on these solutions, to healthcare organizations like nursing homes, small hospitals and hospital chains. In the healthcare industry, patient's privacy, information confidentiality and data security are very important. Also, healthcare organizations are very possessive about their patients' medical records, for good reasons. As a result, each organization wants its own private server. However, many organizations are not able to support (and afford) their own physical servers on their premises and all the operational and management overhead such a private server will entail. This would impede wide adoption of the remote patient monitoring services. However, the cloud server technology comes to the rescue. A3's server software components can be installed on a cloud-based server, and each organization can have its own "private cloud" that is managed by A3 or a third party. Note that, in this architecture, a cloud-based server, including the hardware infrastructure, can be dedicated for exclusive usage by one organization at an affordable cost. So, the cloud-based server technology plays a crucial role in roll-out and adoption of this life-saving remote monitoring technology. Note that such remote patient monitoring is of particular relevance to countries like India where there is a general dearth of qualified doctors and patients residing in vast areas of the country do not have easy access to proper healthcare.

VIII. Conclusions

In the process of our seeking to streamline India's grain supply chains, we discovered cloud computing to be a promising solution. Cloud computing enables use of computing without user-side hardware, software and the associated financial and knowledge requirements, except for the need of a "web-terminal" and access to the internet. We capitalized on this opportunity and proposed a solution based on cloud computing and other operational concepts for streamlining grain supply chains in India. Design of a prototype system is underway that fulfills the functions currently provided by the middlemen between the wholesalers and retailers, adds value to the retailers, wholesalers and even consumers, and is user-friendly.

In addition to distributed hubbing, "consumer visible" branding and cloud computing, which was the main focus of this paper, many other improvement opportunities exist. Further integration of IT with the supply chains beyond the proposed use of cloud computing on the consumer side and the expanding use of eChoupal.com on the farmer side is critical, e.g., RFID, for this worthy and fertile research area. Another improvement opportunity is to increase the economy of scale. Significant reduction in the current number of 600 wholesalers may lead to significantly higher efficiency due to not only consolidated wholesale operations at the hub or sub-hubs but also coordinated distribution. IT may actually play a significant role in such a possible reduction, and this and many other IT studies are worthy subjects for future research.

Assessing overall cost savings achievable for urban and suburban areas, where access to the Internet is not a big issue, via this leap-frogging is a critical research issue for emerging economies.

Another critical research subject is how to allow efficient access to the Internet or the cloud in rural or remote areas of an emerging economy without expensive PCs and laptops.

We believe that many existing operations in other industrial sectors of an emerging economy can benefit from cloud computing in similar ways. In fact, we argued that such leap-frogging can benefit the entire society of an emerging economy.

Acknowledgment

The guest editor's and reviewers' constructive comments led to a much better paper than the original version and are gratefully acknowledged.

References

[1] P. Mell and T. Grance, "The NIST Definition of Cloud Computing," (2011). http://csrc.nist.gov/publications/drafts/800-145/Draft-SP-800-145_cloud-definition.pdf; accessed on July 29, 2011.

[2] Rimal, B.P., Choi, E., and Lumb, I. (2009), "A taxonomy and survey of cloud computing systems," Proceedings of the *Fifth International Joint Conference on INC, IMS and IDC*, p 44-51

[3] Cleverley, M. (2009), "Emerging markets. How ICT advances might help developing nations," Communications of the ACM, Vol. 52, No. 9, pp. 30 -32.

[4] Tsao , H.-S. J., Parikh, S., Ghosh, A.S., Pal, R., Ranalkar, M., Tarapore, H., and Venkatsubramanyan, S. (2010), "Streamlining Grain Supply Chains of India: Cloud Computing and Distributed Hubbing for Wholesale-Retail Logistics,"., Proceedings of *2010 IEEE International Conference on Service Operations and Logistics, and Informatics* (IEEE-SOLI 2010), Qingdao, China.

[5] Google, http://www.google.com/intl/en_us/mobile/sms/ .

[6] Google, http://www.google.com/intl/en_us/mobile/sms/search/ .

[7] Sachan, A., Sahay, B.S., and Sharma, D. (2005), "Developing Indian grain supply chain cost model: a system dynamics approach," *International Journal of Productivity and Performance Management;*Vol. 54, No. 3/4. pp.187-205.

[8] Sachan, A., Sahay, B.S., and Mohan, R. (2006) "Assessing benefits of supply chain integration using system dynamics methodlogy," *International Journal of Services Technology and Management*, Vol. 7, No. 5/6, pp. 582-601.

[9] Dharni, K. and Sharma, S. (2008),"Food Processing in India: Opportunities and Constraints," *The ICFAI University Journal of Agricultural Economics*, Vol. V, No. 3., The ICFAI (Institute of Chartered Financial Analysts of India) University Press, India.

[10] Singh, R., Singh, H.P., Badal, P.S., Singh, O.P., Kushwaha, S., and Sen, C. (2009), "Problems and Prospects of Food-retailing In the State of Uttar Pradesh (India)," *Journal of Services Research,* Vol. 8, No. 2 (October 2008-March 2009), Institute for International Management and Technology, New Delhi, India.

[11] ICRA (2001a), *Report on FMCG*, Investment Information and Credit Rating Agency, New Delhi, India, March, 2001.

[12] ICRA (2001b), *The Indian FMCG Sector*, Investment Information and Credit Rating Agency, New Delhi, India, May, 2001.

[13] ETIG (2003), *Changing Gears: Retailing in India*, Economic Times Intelligence Group, Mumbai, India.

[14] David M. Upton, Virginia A. Fuller (2003), "ITC eChoupal Initiative," HBS Case 604-016, Harvard Business School, Harvard University; also in "Corporate Information Strategy and Management" by Applegate, L.M., Austin, R.D. and McFarlan, F.W. (8th edition).

[15] Milojicic, D. (2008),"*Cloud Computing: Interview with Russ Daniels and Franco Travostino*", IEEE Internet Computing, Sept./Oct., 2008.

[16] Youseff, L., Butrico, M., and Da Silva, D. (2008), "Toward a unified ontology of cloud computing," proceedings of *2008 Grid Computing Environments Workshop*.

[17] Indian Harvest, Centre for Monitoring Indian Economy Pvt. Ltd. http://www.cmie.com/database/?service=database-products/sectoral-services/indian-harvest.htm, accessed on July 30, 2011.

[18] Agriculture.com, http://www.agriculture.com/markets/commodities, accessed on July 30, 2011.

[19] Agricommodityprices.com, http://www.agricommodityprices.com, accessed on July 30, 2011.

[20] FAOSTAT, Food and Agriculture Organization, United Nations, http://faostat.fao.org/site/291/default.aspx, accessed on July 30, 2011

[21] PriceSTAT, Food and Agriculture Organization, United Nations, http://faostat.fao.org/site/570/DesktopDefault.aspx?PageID=570#ancor, accessed on July 30, 2011.

[22] National Agricultural Statistics Service, United States Department of Agriculture, http://www.nass.usda.gov/index.asp, accessed on July 30, 2011.

[23] Indexmundi.com, http://www.indexmundi.com/commodities, accessed on July 30, 2011.

[24] Tsao, H.-S. J., "A Framework for Evaluating Deployment Strategies for Intelligent Transportation Systems", Intelligent Transportation Systems Journal (ITS Journal), Vol.6, pp. 141-173, 2001.

[25] T. Hall and J. Luter III. Is SOA Superior? Evidence from SaaS Financial Statements. *Journal Of Software*, Vol. 3, No. 5, May 2008.

[26] Kondo,D., Javadi, B., Malecot, P., Cappello, F., and Anderson, D.P. (2009), "Cost-Benefit Analysis of Cloud Computing versus Desktop Grids," *2009 IEEE International Symposium on Parallel & Distributed Processing (IPDPS)*.

[27] Lin, T. and Wang, S. (2009), "Cloudlet-Screen Computing: A Multi-core-based, Cloud-computing-oriented, Traditional-computing-compatible Parallel Computing Paradigm for the Masses," *Proceedings 2009 IEEE International Conference on Multimedia and Expo (ICME)*, p 1805-1808.

[28] National ICT R&D Fund, Request for Proposals: "Establishing Center of Innovation for Use of ICT in Healthcare Sector,"; http://www.ictrdf.org.pk/, accessed on July 30, 2011.

[29] A3 BioMed Technologies, http://www.a3biomed.com/index.htm, accessed on July 30, 2011.

ICTer COPYRIGHT FORM

To ensure uniformity of treatment among all contributors, this form shall be treated as the only copyright form for publications in "ICTer" and other forms may not be substituted for this form, nor may any wording of the form be changed. This form is intended for original material submitted to the ICTer and must accompany any such material in order to be published by the ICTer. Please read the form carefully and keep a copy for your files.

TITLE OF PAPER/ARTICLE/REPORT/PRESENTATION/SPEECH (hereinafter, "the Work"):

COMPLETE LIST OF AUTHORS:

ICTer PUBLICATION TITLE (Journal):

Copyright Transfer

The undersigned hereby assigns to the International Journal on Advances in ICT for Emerging Regions, Incorporated (the "ICTer") all rights under copyright that may exist in and to the above Work, and any revised or expanded derivative works submitted to the ICTer by the undersigned based on the Work. The undersigned hereby warrants that the Work is original and that he/she is the author of the Work; to the extent the Work incorporates text passages, figures, data or other materials from the works of others, the undersigned has obtained any necessary permission. **See reverse side for Retained Rights and other Terms and Conditions.**

Author Responsibilities

The ICTer distributes its publications throughout the world and wants to ensure that the material submitted to its publications is properly available to the readership of those publications. Authors must ensure that their Work meets the requirements as stated in section 2.7 of the ICTer Author guidelines, including provisions covering originality, authorship, author responsibilities and author misconduct. More information on ICTer's publishing policies may be found at http://www.icter.org. Authors are advised especially of the following.

1. It is the responsibility of the authors, not the ICTer, to determine whether disclosure of their material requires the prior consent of other parties and, if so, to obtain it.

2. Statements and opinions given in work published by the ICTer are the expression of the authors.

General Terms

- The undersigned represents that he/she has the power and authority to make and execute this assignment.
- The undersigned agrees to indemnify and hold harmless the ICTer from any damage or expense that may arise in the event of a breach of any of the warranties set forth above.
- In the event the above work is not accepted and published by the ICTer or is withdrawn by the author(s) before acceptance by the ICTer, the foregoing copyright transfer shall become null and void and all materials embodying the Work submitted to the ICTer will be destroyed.
- For jointly authored Works, all joint authors should sign, or one of the authors should sign as authorized agent for the others.

(1)____ __________________________

Author/Authorized Agent for Joint Authors **Date**

ICTer COPYRIGHT FORM *(continued)*

RETAINED RIGHTS/TERMS AND CONDITIONS

1. Authors/employers retain all proprietary rights in any process, procedure, or article of manufacture described in the Work.

2. Authors/employers may reproduce or authorize others to reproduce the Work, materials extracted verbatim from the Work, or derivative works for the author's personal use or for company use, provided that the source and the ICTer copyright notice are indicated, the copies are not used in any way that implies ICTer endorsement of a product or service of any employer, and the copies themselves are not offered for sale.

3. Authors/employers may make limited distribution of all or portions of the Work prior to publication if they inform the ICTer in advance of the nature and extent of such limited distribution.

4. In the case of a Work performed under any Government contract or grant, the ICTer recognizes that the said Government has permission to reproduce all or portions of the Work, and to authorize others to do so, for official Government purposes, only if the contract/grant so requires.

5. For all uses not covered by items 2, 3, and 4, authors/employers must request permission from the ICTer to reproduce or authorize the reproduction of the Work or materials extracted verbatim from the Work, including figures and tables.

6. Although authors are permitted to re-use all or portions of the Work in other works, this does not include granting third-party requests for reprinting, republishing, or other types of re-use. The ICTer must handle all such third-party requests.

INFORMATION FOR AUTHORS

ICTer Copyright Ownership

It is the formal policy of the ICTer to own the copyrights to all copyrightable materials in its journal publications and to the individual contributions contained therein, in order to protect the interests of the ICTer, its authors and their employers, and, at the same time, to facilitate the appropriate re-use of this material by others. The ICTer distributes its journal publications throughout the world by means such as hard copy, and electronic media. It also abstracts and may translate its publications, and articles contained therein, for inclusion in various compendiums, collective works, databases and similar publications.

Author/Employer Rights

If you are employed and prepared the Work on a subject within the scope of your employment, the copyright in the Work belongs to your employer as a work-for-hire. In that case, the ICTer assumes that when you sign this Form, you are authorized to do so by your employer and that your employer has consented to the transfer of copyright, to the representation and warranty of publication rights, and to all other terms and conditions of this Form. If such authorization and consent has not been given to you, an authorized representative of your employer should sign this Form as the Author.

Reprint/Republication Policy

The ICTer requires that the consent of the first-named author and employer be sought as a condition to granting reprint or republication rights to others or for permitting use of a Work for promotion or marketing purposes.

PLEASE DIRECT ALL QUESTIONS ABOUT THIS FORM TO:

Chair-ICTer Steering Committee, University of Colombo School of Computing, No 35, Ried Avenue, Colombo 7
Tel: 94-112-581245, **Fax:** 94-112-591245 **Email:** info@icter.org