UNIT I

Cloud Computing provides us a means by which we can access the applications as utilities, over the Internet. It allows us to create, configure, and customize applications online.

What is Cloud?

The term **Cloud** refers to a **Network** or **Internet**. In other words, we can say that Cloud is something, which is present at remote location. Cloud can provide services over network, i.e., on public networks or on private networks, i.e., WAN, LAN or VPN.

Applications such as e-mail, web conferencing, customer relationship management (CRM), all run in cloud.

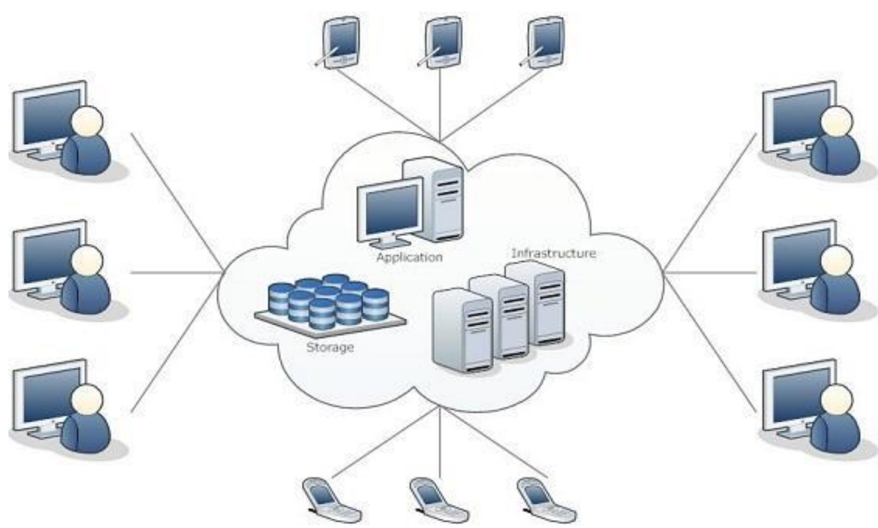
What is computing?

The Process of utilizing computer technology to complete a task is known as computing. It may involve computer Hardware and software. Eg. Sending an email, swiping a card, etc.

What is Cloud Computing? (NIST)

Cloud computing is a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction.

C*loud computing* refers to applications and services that run on a distributed network using virtualized resources and accessed by common Internet protocols and networking standards.



We need not to install a piece of software on our local PC and this is how the cloud computing overcomes **platform dependency issues**. Hence, the Cloud Computing is making our business application **mobile** and **collaborative**.

CLOUD COMPUTING IN A NUTSHELL

When plugging an electric appliance into an outlet, we care neither how electric power is generated nor how it gets to that outlet. This is possible because electricity is virtualized; that is, it is readily available from a wall socket that hides power generation stations and a huge distribution grid. When extended to information technologies, this concept means delivering useful functions while hiding how their internals work. Computing itself, to be considered fully virtualized, must allow computers to be built from distributed components such as processing, storage, data, and software resources.

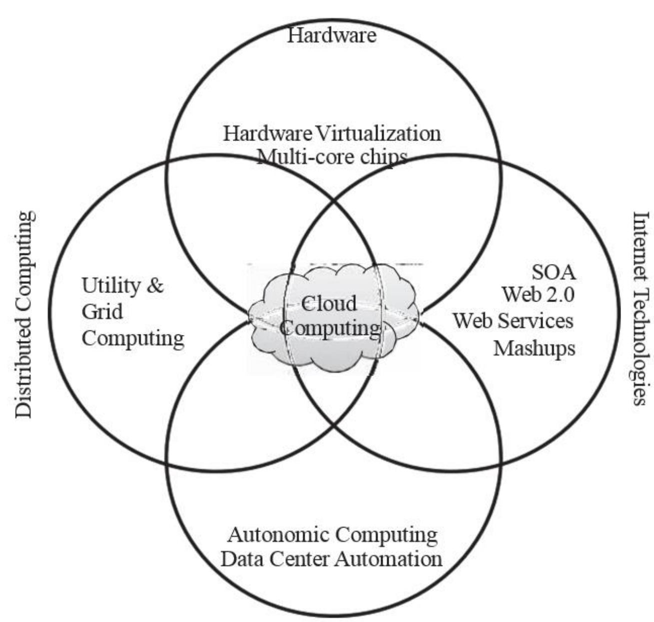
Technologies such as cluster, grid, and now, cloud computing, have all aimed at allowing access to large amounts of computing power in a fully virtualized manner, by aggregating resources and offering a single system view. In addition, an important aim of these technologies has been delivering computing as a utility. Utility computing describes a business model for on-demand delivery of computing power; consumers pay providers based on usage ("payas-you-go"), similar to the way in which we currently obtain services from traditional public utility services such as water, electricity, gas, and telephony. Cloud computing has been coined as an umbrella term to describe a category of sophisticated on-demand computing services initially offered by commercial providers, such as Amazon, Google, and Microsoft. It denotes a model on which a computing infrastructure is viewed as a "cloud," from which businesses and individuals access applications from anywhere in the world on demand. The main principle behind this model is offering computing, storage, and software "as a service."

Enabling Technology for the Cloud

Cloud enable technologies for the software, hardware & networking.

Technology	Requirements and Benefits
Fast platform deployment	Fast, efficient, and flexible deployment of cloud resources to provide dynamic computing environment to users
Virtual clusters on demand	Virtualized cluster of VMs provisioned to satisfy user demand and virtual cluster reconfigured as workload changes
Multitenant techniques	SaaS for distributing software to a large number of users for their simultaneous use and resource sharing if so desired
Massive data processing	Internet search and Web services which often require massive data processing, especially to support personalized services
Web-scale communication	Support for e-commerce, distance education, telemedicine, social networking, digital government, and digital entertainment applications
Distributed storage	Large-scale storage of personal records and public archive information which demands distributed storage over the clouds
Licensing and billing services	License management and billing services which greatly benefit all types of cloud services in utility computing

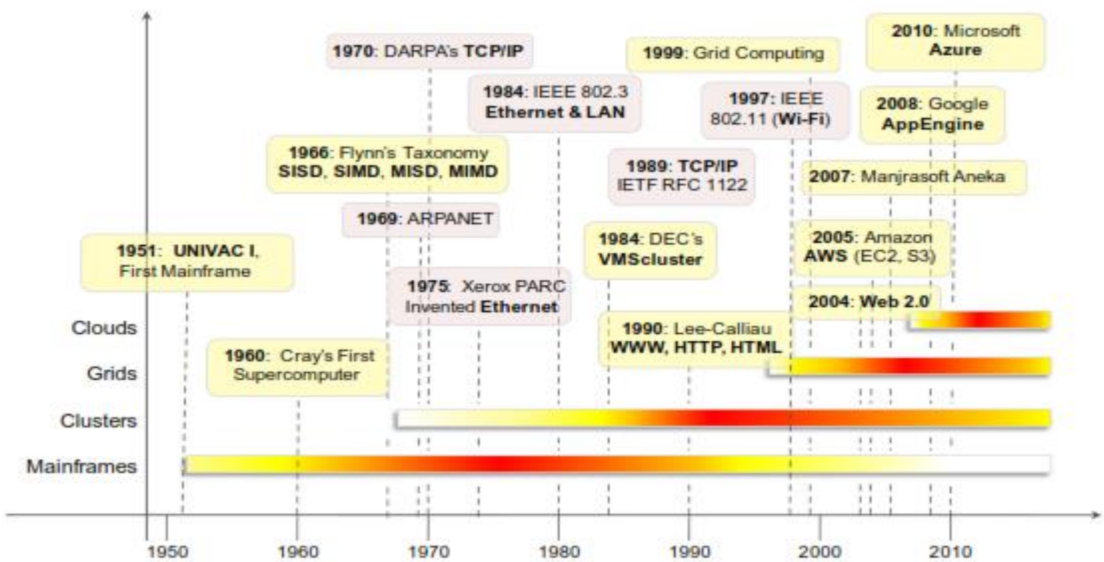
The History and Development of Cloud Computing



1 Distributed systems

Clouds are essentially large distributed computing facilities that make available their services to third parties on demand. A distributed system is a collection of independent computers that appears to its users as a single coherent system. Independent components and that these components are perceived as a single entity by users. This is particularly true in the case of cloud computing, in which clouds hide the complex architecture they rely on and provide a single interface to users. The primary purpose of distributed systems is to share resources and utilize them better. This is true in the case of cloud computing, where this concept is taken to the extreme and resources (infrastructure, runtime environments, and services) are rented to users.

In fact, one of the driving factors of cloud computing has been the availability of the large computing facilities of IT giants (Amazon, Google) that found that offering their computing capabilities as a service provided opportunities to better utilize their infrastructure.



Evolution of Distributed computing Technologies

Distributed systems often exhibit other properties such as heterogeneity, openness, scalability, transparency, concurrency, continuous availability, and independent failures.

Three major milestones have led to cloud computing: mainframe computing, cluster computing, and grid computing.

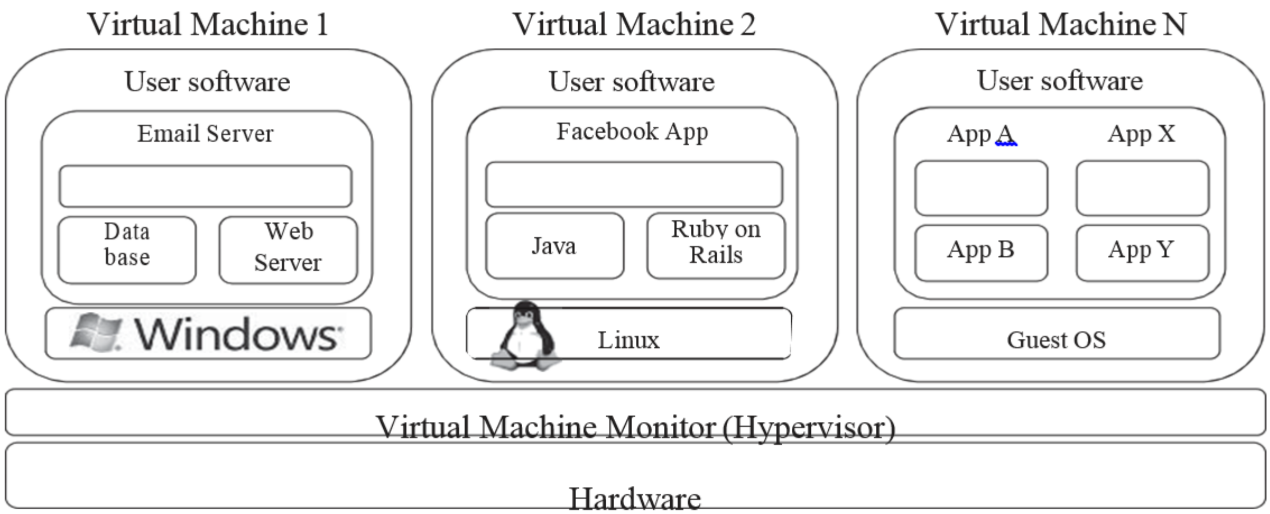
- **Mainframes.** These were the first examples of large computational facilities leveraging multiple processing units. Mainframes were powerful, highly reliable computers specialized for large data movement and massive input/output (I/O) operations. They were mostly used by large organizations for bulk data processing tasks such as online transactions, enterprise resource planning, and other operations involving the processing of significant amounts of data. Even though mainframes cannot be considered distributed systems, they offered large computational power by using multiple processors, which were presented as a single entity to users. One of the most attractive features of mainframes was the ability to be highly reliable computers that were "always on" and capable of tolerating failures transparently. No system shutdown was required to replace failed components, and the system could work without interruption.

- **Clusters.** Cluster computing started as a low-cost alternative to the use of mainframes and supercomputers. The technology advancement that created faster and more powerful mainframes and supercomputers eventually generated an increased availability of cheap commodity machines as a side effect. These machines could then be connected by a high-bandwidth network and controlled by specific software tools that manage them as a single system.

- **Grids.** Grid computing appeared in the early 1990s as an evolution of cluster computing. In an analogy to the power grid, grid computing proposed a new approach to access large computational power, huge storage facilities, and a variety of services. Users can "consume" resources in the same way as they use other utilities such as power, gas, and water. Grids initially developed as aggregations of geographically dispersed clusters by means of Internet connections. These clusters belonged to different organizations, and arrangements were made among them to share the computational power.

2 Hardware Virtualization

Virtualization is another core technology for cloud computing. It encompasses a collection of solutions allowing the abstraction of some of the fundamental elements for computing, such as hardware, runtime environments, storage, and networking. The idea of virtualizing a computer system's resources, including processors, memory, and I/O devices, has been well established for decades, aiming at improving sharing and utilization of computer systems. Hardware virtualization allows running multiple operating systems and software stacks on a single physical platform. As depicted in Figure 1.2, a software layer, the virtual machine monitor (VMM), also called a hypervisor, mediates access to the physical hardware presenting to each guest operating system a virtual machine (VM), which is a set of virtual platform interfaces



Virtualization software separates a physical computing device into one or more virtual device. Virtualization is essentially a technology that allows creation of different computing environments. These environments are called virtual because they simulate the interface that is expected by a guest.

3 Web 2.0

The web 2.0 is the 2nd generation of worldwide web, which focuses on the ability of people to share information online. The Web is the primary interface through which cloud computing delivers its services. At present, the Web encompasses a set of technologies and services that facilitate interactive information sharing, collaboration, user-centered design, and application composition. This evolution has transformed the Web into a rich platform for application development and is known as Web 2.0. This term captures a new way in which developers architect applications and deliver services through the Internet and provides new experience for users of these applications and services.

Web 2.0 brings interactivity and flexibility into Web pages, transmission from static html pages to dynamic web pages. Providing enhanced user experience by gaining Web-based access to all the functions that are normally found in desktop applications. These capabilities are obtained by integrating a collection of standards and technologies such as XML, Asynchronous JavaScript and XML (AJAX), Web Services, and others.

Examples of Web 2.0 applications are Google Documents, Google Maps, Flickr, Facebook, Twitter, YouTube, de.li.cious, Blogger, and Wikipedia. In particular, social networking Websites take the biggest advantage of Web 2.0.

4 Service-oriented computing

Service orientation is the core reference model for cloud computing systems. This approach adopts the concept of services as the main building blocks of application and system development. Service-oriented computing (SOC) supports the development of rapid, low-cost, flexible, interoperable, and evolvable applications and systems.

A service is an abstraction representing a self-describing and platform-agnostic component that can perform any function—anything from a simple function to a complex business process.

Virtually any piece of code that performs a task can be turned into a service and expose its functionalities through a network-accessible protocol. A service is supposed to be loosely coupled, reusable, programming language independent, and location transparent. Loose coupling allows services to serve different scenarios more easily and makes them reusable. Independence from a specific platform increases services accessibility. Thus, a wider range of clients, which can look up services in global registries and consume them in a location-transparent manner, can be served.
Service-oriented computing introduces and diffuses important concepts, which are also fundamental to cloud computing: quality of service (QoS).
Quality of service (QoS) identifies a set of functional and nonfunctional attributes that can be used to evaluate the behavior of a service from different perspectives. These could be performance metrics such as response time, or security attributes, transactional integrity, reliability, scalability, and availability. QoS requirements are established between the client and the provider via an SLA that identifies the minimum values (or an acceptable range) for the QoS attributes that need to be satisfied upon the service call.

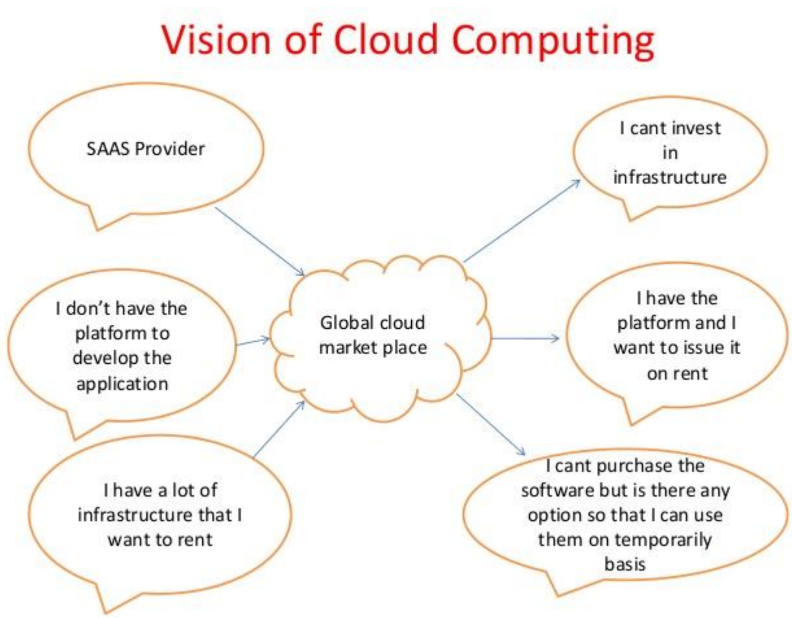
5 Utility-oriented computing
Utility computing is a vision of computing that defines a service-provisioning model for compute services in which resources such as storage, compute power, applications, and infrastructure are packaged and offered on a pay-per-use basis. The idea of providing computing as a utility like natural gas, water, power, and telephone connection has a long history but has become a reality today with the advent of cloud computing.

6 Autonomic Computing

The increasing complexity of computing systems has motivated research on autonomic computing, which seeks to improve systems by decreasing human involvement in their operation. In other words, systems should manage themselves, with high-level guidance from humans.

In this sense, the concepts of autonomic computing inspire software technologies for data center automation, which may perform tasks such as: management of service levels of running applications; management of data center capacity; proactive disaster recovery; and automation of VM provisioning.

Vision of Cloud Computing



- Cloud computing provides the facility to provision virtual hardware, runtime environment and services to a person having money.These all things can be used as long as they are needed by the user, there is no requirement for the upfront commitment.

- The whole collection of computing system is transformed into a collection of utilities, which can be provisioned and composed together to deploy system in hours rather than days, with no maintenance costs.
- The long term vision of a cloud computing is that IT services are traded as utilities in an open market without technological and legal barriers.
- In the near future we can imagine that it will be possible to find the solution that matches with our requirements by simply entering our request in a digital global market that trades with cloud computing services.
- The existence of such market will enable the automation of the discovery process and its integration into its existing software systems.
- Due to existence of global platform for trading cloud services will also help service providers to potentially increase their revenue.
- A cloud provider can also become a consumer of a competitor service in order to fulfill its promises to customers.

Five essential characteristics of cloud computing by NIST (National Institute of Standards & Technology)

• **On-demand self-service:** A consumer can unilaterally provision computing capabilities, such as server time and network storage, as needed automatically without requiring human interaction with each service provider. Customer can order the service at any time 24X7, which becomes immediately available to him.

• **Broad network access:** Access to resources in the cloud is available over the network using standard methods in a manner that provides platform-independent access to clients of all types. This includes a mixture of heterogeneous operating systems, and thick and thin platforms such as laptops, mobile phones, and PDA.

• **Resource pooling:** A cloud service provider creates resources that are pooled together in a system that supports multi-tenant usage. Physical and virtual systems are dynamically allocated or reallocated as needed. Intrinsic in this concept of pooling is the idea of abstraction that hides the location of resources such as virtual machines, processing, memory, storage, and network bandwidth and connectivity.

• **Rapid elasticity:** Resources can be rapidly and elastically provisioned. The system can add resources by either scaling up systems (more powerful computers) or scaling out systems (more computers of the same kind), and scaling may be automatic or manual. From the standpoint of the client, cloud computing resources should look limitless and can be purchased at any time and in any quantity.

• Per-Usage Metering and Billing/Measured service: The use of cloud system resources is measured, audited, and reported to the customer based on a metered system. A client can be charged based on a known metric such as amount of storage used, number of transactions, network I/O (Input/Output) or bandwidth, amount of processing power used, and so forth. A client is charged based on the level of services provided.

Advantages of cloud computing

• **Lower costs:** Because cloud networks operate at higher efficiencies and with greater tilization, significant cost reductions are often encountered.

• **Ease of utilization:** Depending upon the type of service being offered, you may find that you do not require hardware or software licenses to implement your service.

• **Quality of Service:** The Quality of Service (QoS) is something that you can obtain under contract from your vendor.

• **Reliability:** The scale of cloud computing networks and their ability to provide load balancing and failover makes them highly reliable, often much more reliable than what you can achieve in a single organization.

• **Outsourced IT management:** A cloud computing deployment lets someone else manage your Computing infrastructure while you manage your business. In most instances, you achieve considerable reductions in IT staffing costs.

• **Simplified maintenance and upgrade:** Because the system is centralized, you can easily apply patches and upgrades. This means your users always have access to the latest software versions.

• **Low Barrier to Entry:** In particular, upfront capital expenditures are dramatically reduced. In cloud computing, anyone can be a giant at any time.

Disadvantages of cloud computing

- **Downtime:** As cloud service providers take care of a number of clients each day, they can become overwhelmed and may even come up against technical outages. This can lead to your business processes being temporarily suspended. Additionally, if your internet connection is offline, you will not be able to access any of your applications, server or data from the cloud.
- **Security:** Although cloud service providers implement the best security standards and industry certifications, storing data and important files on external service providers always opens up risks. Using cloud-powered technologies means you need to provide your service provider with access to important business data. Meanwhile, being a public service opens up cloud service providers to security challenges on a routine basis. The ease in procuring and accessing cloud services can also give nefarious users the ability to scan, identify and exploit loopholes and vulnerabilities within a system. For instance, in a multi-tenant cloud architecture where multiple users are hosted on the same server, a hacker might try to break into the data of other users hosted and stored on the same server. However, such exploits and loopholes are not likely to surface, and the likelihood of a compromise is not great.
- **Vendor Lock-In:** Although cloud service providers promise that the cloud will be flexible to use and integrate, switching cloud services is something that hasn't yet completely evolved. Organizations may find it difficult to migrate their services from one vendor to another. Hosting and integrating current cloud applications on another platform may throw up interoperability and support issues. For instance, applications developed on Microsoft Development Framework (.Net) might not work properly on the Linux platform.
- **Limited Control:** Since the cloud infrastructure is entirely owned, managed and monitored by the service provider, it transfers minimal control over to the customer. The customer can only control and manage the applications, data and services operated on top of that, not the backend infrastructure itself. Key administrative tasks such as server shell access, updating and firmware management may not be passed to the customer or end user.

Basic Components of Cloud Computing

Components in a cloud refers to the platforms, like front end, back end and cloud based delivery and the network that used. All together forms an architecture for cloud computing. With the main components like SAAS, PAAS and IAAS there are 11 more major categories in cloud computing that we will explain here.

Storage-as-a-Service: This is the component where we can use or request storage, like as we do it physically using the remote site. It is also called disk space on demand. This is the main component where even other components will have a base component as Storage-as-a-Service.

Database-as-a-Service: This component acts as a live database from remote where its functionality and other features works as though a physical db is present in the local machine. Its main objective is to reduce the cost of db using many softwares as well as hardwares.

Information-as-a-Service: Information that can be accessed remotely from anywhere is called Information-as-a-Service. Here the information will be fetched remotely. This includes, for example, live stock prices, internet banking, online news, credit card validation and so on.

Process-as-a-Service: This component combines various resources such as data and services. This happens either hosted within the same cloud computing resource or remotely. Mainly this is used for business processes where various key services and information are combined to form a process. This helps delivery on demand. For example mobile networks (internet settings are sent as soon as activated).

Application-as-a-Service: Application-as-a-Service (also known as SAAS) is the complete application built ready for use by the client. This is built to use the internet to the end users and the end users normally use browsers and the internet to access this service. This component is the ultimate front-end for end users. Some of the applications are Salesforce, Gmail, Google calendar and so on.

Platform-as-a-Service: This is the component where the app is being developed and the database is being created, implemented, stored and tested. In recent times this component allows creation of enterprise-level applications easily and is cost-effective.

Integration-as-a-Service: Integration-as-a-Service deals with the components of an application that has been built but must be integrated with other applications. It helps in medaiting between the remote servers with the local machines. Stacks from the cloud are fetched and communicated with local machines. For example salesforce has recently integrated Google maps into it.

Security-as-a-Service: This is the main component many customers require. Whoever goes for a cloud environment needs security features a lot since all the data and operations are handled remotely.There are three-dimensional securities found in cloud platforms.

Management-as-a-service: This is a component that is mainly useful for management of the clouds, like resource utilisation, virtualisation and server up and down time management. This will be like a small role like an admin point of view.

Testing-as-a-Service: Testing-as-a-Service refers to the testing of the applications that are hosted remotely, whether there is a requirement to design a working database and there is enough security for the applications and so on. This will be tested even with two or three cross clouds. This will also be a component in the development of cloud products.

Infrastructute-as-a-Service: This is called as nearly as possible the taking of all the hardware, software, servers and networking that is completely virtual. This is where all the processes and purchases of resources will take place in the cloud. Our processes will happen but we can't see what's happening at the backend. This avoids the running of multiple servers, heat, cold, temperature and so on.

CHALLENGES AND RISKS

Despite the initial success and popularity of the cloud computing paradigm and the extensive availability of providers and tools, a significant number of challenges and risks are inherent to this new model of computing. Providers, developers, and end users must consider these challenges and risks to take good advantage of cloud computing. Issues to be faced include user privacy, data security, data lock-in, availability of service, disaster recovery, performance, scalability, energy-efficiency, and programmability.

Security, Privacy, and Trust

- Information security as a main issue: "current cloud offerings are essentially public exposing the system to more attacks." For this reason there are potentially additional challenges to make cloud computing environments as secure as in-house IT systems. At the same time, existing, well understood technologies can be leveraged, such as data encryption, VLANs, and firewalls.
- Security and privacy affect the entire cloud computing stack, since there is a massive use of third-party services and infrastructures that are used to host important data or to perform critical operations. In this scenario, the trust toward providers is fundamental to ensure the desired level of privacy for applications hosted in the cloud.
- Legal and regulatory issues also need attention. When data are moved into the Cloud, providers may choose to locate them anywhere on the planet. The physical location of data centers determines the set of laws that can be applied to the management of data. For example, specific cryptography techniques could not be used because they are not allowed in some countries. Similarly, country laws can impose that sensitive data, such as patient health records, are to be stored within national borders.

Data Lock-In and Standardization

- A major concern of cloud computing users is about having their data locked-in by a certain provider. Users may want to move data and applications out from a provider that does not meet their requirements. However, in their current form, cloud computing infrastructures and platforms do not employ standard methods of storing user data and applications. Consequently, they do not interoperate and user data are not portable.
- The answer to this concern is standardization. In this direction, there are efforts to create open standards for cloud computing. The Cloud Computing Interoperability Forum (CCIF) was formed by organizations such as Intel, Sun, and Cisco in order to "enable a global cloud computing ecosystem whereby organizations are able to seamlessly work together for the purposes for wider industry adoption of cloud computing technology." The development of the Unified Cloud Interface (UCI) by CCIF aims at creating a standard programmatic point of access to an entire cloud infrastructure.
- In the hardware virtualization sphere, the Open Virtual Format (OVF) aims at facilitating packing and distribution of software to be run on VMs so that virtual appliances can be made portable—that is, seamlessly run on hypervisor of different vendors.

Availability, Fault-Tolerance, and Disaster Recovery

- It is expected that users will have certain expectations about the service level to be provided once their applications are moved to the cloud. These expectations include availability of the service, its overall performance, and what measures are to be taken when something goes wrong in the system or its components. In summary, users seek for a warranty before they can comfortably move their business to the cloud.
- SLAs, which include QoS requirements, must be ideally set up between customers and cloud computing providers to act as warranty. An SLA specifies the details of the service to be provided, including

availability and performance guarantees. Additionally, metrics must be agreed upon by all parties, and penalties for violating the expectations must also be approved.

Resource Management and Energy-Efficiency

- One important challenge faced by providers of cloud computing services is the efficient management of virtualized resource pools. Physical resources such as CPU cores, disk space, and network bandwidth must be sliced and shared among virtual machines running potentially heterogeneous workloads.
- The multi-dimensional nature of virtual machines complicates the activity of finding a good mapping of VMs onto available physical hosts while maximizing user utility. Dimensions to be considered include: number of CPUs, amount of memory, size of virtual disks, and network bandwidth. Dynamic VM mapping policies may leverage the ability to suspend, migrate, and resume VMs as an easy way of preempting low-priority allocations in favor of higher-priority ones.
- Migration of VMs also brings additional challenges such as detecting when to initiate a migration, which VM to migrate, and where to migrate. In addition, policies may take advantage of live migration of virtual machines to relocate data center load without significantly disrupting running services. In this case, an additional concern is the trade-off between the negative impact of a live migration on the performance and stability of a service and the benefits to be achieved with that migration.
- Another challenge concerns the outstanding amount of data to be managed in various VM management activities. Such data amount is a result of particular abilities of virtual machines, including the ability of traveling through space (i.e., migration) and time (i.e., check pointing and rewinding), operations that may be required in load balancing, backup, and recovery scenarios. In addition, dynamic provisioning of new VMs and replicating existing VMs require efficient mechanisms to make VM block storage devices (e.g., image files) quickly available at selected hosts.
- Data centers consumer large amounts of electricity. According to a data published by HP, 100 server racks can consume 1.3 MW of power and another 1.3 MW are required by the cooling system, thus costing USD 2.6 million per year. Besides the monetary cost, data centers significantly impact the environment in terms of CO emissions from the cooling systems.
- In addition to optimize application performance, dynamic resource management can also improve utilization and consequently minimize energy consumption in data centers. This can be done by judiciously consolidating workload onto smaller number of servers and turning off idle resources.

MIGRATING INTO A CLOUD

- Cloud migration is the process of moving data, applications or other business elements from an organization's onsite computers to the cloud, or moving them from one cloud environment to another.
- Cloud migration sometimes involves moving data or other business elements between cloud environments, which is known as cloud-to-cloud migration. The process of transitioning to a different cloud provider is known as cloud service migration. In any case, successful migration to a service provider’s environment may require the use of middleware, such as a cloud integration tool, to bridge any gaps between the vendor’s and the customer’s (or other vendor’s) technologies.

BROAD APPROACHES TO MIGRATING INTO THE CLOUD

Migrating into the cloud is poised to become a large-scale effort in leveraging the cloud in several enterprises. “Cloudonomics” deals with the economic rationale for leveraging the cloud and is central to the success of cloud-based enterprise usage.

At what IT costs—both short term and long term—would one want to migrate into the cloud? While all capital expenses are eliminated and only operational expenses incurred by leveraging the cloud, does this satisfy all strategic parameters for enterprise IT? Does the total cost of ownership (TCO) become significantly less as compared to that incurred when running one’s own private data center?

Why Migrate?

There are economic and business reasons why an enterprise application can be migrated into the cloud, and there are also a number of technological reasons. Many of these efforts come up as initiatives in adoption of cloud technologies in the enterprise, resulting in integration of enterprise applications running off the captive data centers with the new ones that have been developed on the cloud. Adoption of or integration with cloud computing services is a use case of migration.

- At the core, migration of an application into the cloud can happen in one of several ways: Either the application is clean and independent, so it runs as is; or perhaps some degree of code needs to be modified and adapted; or the design (and therefore the code) needs to be first migrated into the cloud computing service environment; or finally perhaps the migration results in the core architecture being migrated for a cloud computing service setting, this resulting in a new architecture being developed, along with the accompanying design and code implementation. Or perhaps while the application is migrated as is, it is the usage of the application that needs to be migrated and therefore adapted and modified.
- In brief, migration can happen at one of the five levels of application, code, design, architecture, and usage.With due simplification, the migration of an enterprise application is best captured by the following:

$$P \rightarrow P'_C + P'_l \rightarrow P'_{\mathrm{OFC}} + \mathrm{P}'_l$$

Where **P** is the application before migration running in captive data center, **P'c** is the application part after migration either into a (hybrid) cloud, **P'l** is the part of application being run in the captive local data center, and **P'OFC** is the application part optimized for cloud. If an enterprise application cannot be migrated fully, it could result in some parts being run on the captive local data center while the rest are being migrated into the cloud—essentially a case of a hybrid cloud usage. However, when the entire application is migrated onto the cloud, then **P'l** is null. Indeed, the migration of the enterprise application **P** can happen at the five levels of application, code, design, architecture, and usage. It can be that the **P'c** migration happens at any of the five levels without any **P'l** component. Compound this with the kind of cloud computing service offering being applied—the IaaS model or PaaS or SaaS model—and we have a variety of migration use cases that need to be thought through thoroughly by the migration architects.

Deciding on the Cloud Migration

In fact, several proof of concepts and prototypes of the enterprise application are experimented on the cloud to take help in making a sound decision on migrating into the cloud. Post migration, the ROI (Return on investment) on the migration should be positive for a broad range of pricing variability. Arriving at a decision for undertaking migration demands that either the compelling factors be clearly understood or the pragmatic approach of consulting a group of experts be constituted. We use the following technique to make the decision:

A questionnaire with several classes of key questions that impact the IT due to the migration of the enterprise application is posed to a select audience chosen for their technology and business expertise. Assume that there are M such in the context of the entire questionnaire. Assume that in the M classes of questions, there was a class with a maximum of N questions. We can then model the weightage-based decision making as MxN weightage matrix as follows:

$$C_l \leq \sum_{i=1}^{M} B_i \left(\sum_{j=1}^{N} A_{ij} X_{ij} \right) \leq C_h$$

where C_l is the lower weightage threshold and C_h is the higher weightage threshold while Aij is the specific constant assigned for a question and x_{ij} fraction between 0 and 1 that represents the degree to which that answer to the question is relevant and applicable. Since all except one class of questions do not have all N questions, the corresponding has a null value. The lower and higher thresholds are defined to rule out trivial cases of migration. A simplified variant of this method can be presented as a balanced scorecard oriented decision making.

THE SEVEN-STEP MODEL OF MIGRATION INTO A CLOUD

Typically migration initiatives into the cloud are implemented in phases or in Stages. A structured and process-oriented approach to migration into a cloud has several advantages of capturing within itself.
In a succinct way, Figure 3 captures the essence of the steps in the model of migration into the cloud, while Figure 4 captures the iterative process of the seven-step migration into the cloud.

- Cloud migration assessments comprise assessments to understand the issues involved in the specific case of migration at the application level or the code, the design, the architecture, or usage levels. In addition, migration assessments are done for the tools being used, the test cases as well as configurations, functionalities, and NFRs of the enterprise application. This results in a meaningful formulation of a comprehensive migration strategy. Proof of concepts or prototypes for various approaches to the migration along with the leveraging of pricing parameters enables one to make appropriate assessments.

- The next process step is in isolating all systemic and environmental dependencies of the enterprise application components within the captive data center. This, in turn, yields a picture of the level of complexity of the migration.

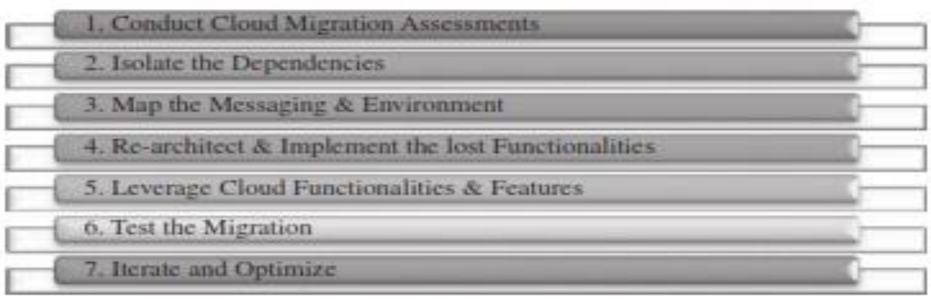


Fig. 3 The Seven-Step Model of Migration into the Cloud

- After isolation is complete, one then goes about generating the mapping constructs between what shall possibly remain in the local captive data center and what goes onto the cloud.

- Perhaps a substantial part of the enterprise application needs to be rearchitected, redesigned, and reimplemented on the cloud. This gets in just about the functionality of the original enterprise application. Due to this migration, it is possible perhaps that some functionality is lost.

Fig. 4 The iterative Seven-step Model of Migration into the Cloud.

- ➢ In the next process step we leverage the intrinsic features of the cloud computing service to augment our enterprise application in its own small ways.
- ➢ Having done the augmentation, we validate and test the new form of the enterprise application with an extensive test suite that comprises testing the components of the enterprise application on the cloud as well. These test results could be positive or mixed.
- ➢ In the latter case, we iterate and optimize as appropriate. After several such optimizing iterations, the migration is deemed successful.

Assess	Isolate	Map	Re-Architect	Augment	Test	Optimize
• Cloudonomics • Migration Costs • Recurring Costs • Database data segmentation • Database Migration • Functionality migration • NFR Support	• Runtime Environment • Licensing • Libraries Dependency • Applications Dependency • Latencies Bottlenecks • Performance bottlenecks • Architectural Dependencies	• Messages mapping: marshalling & de-marshalling • Mapping Environments • Mapping libraries & runtime approximations	• Approximate lost functionality using cloud runtime support API • New Usecases • Analysis • Design	• Exploit additional cloud features • Seek Low-cost augmentations • Autoscaling • Storage • Bandwidth • Security	• Augment Test Cases and Test Automation • Run Proof-of-Concepts • Test Migration strategy • Test new testcases due to cloud augmentation • Test for Production Loads	• Optimize– rework and iterate • Significantly satisfy cloudonomics of migration • Optimize compliance with standards and governance • Deliver best migration ROI • Develop roadmap for leveraging new cloud features

Fig. 5 details of the iterative Seven-Step Model of Migration into the Cloud

It is best to iterate through this Seven-Step Model process for optimizing and ensuring that the migration into the cloud is both robust and comprehensive. Figure 5 captures the typical components of the Seven-Step Model of Migration into the Cloud.

MIGRATION RISKS AND MITIGATION

Even with all the impressive statistics against it name; it cannot be denied that there are quite a few drawbacks of cloud computing as well. This is because the technology, although robust, is still in nascent stages of being perfected and still has a few kinks which need to be ironed out. Some of these risks include –

Interoperability: Interoperability is the one of the big challenges in cloud migration. This is concerned with the ability of devices to communicate to each other. In cloud community technology, it means the ability to develop the code which works with number of cloud service providers simultaneously irrespective of the differences between the cloud providers. So, if we migrate the business to the cloud and want to be part of cloud environment, then it should be compatible with more than one service provider.

Cloud security and privacy: Security of the cloud is the biggest threat in the current market. When we decide to adopt the cloud technology, we are migrating business sensitive information to the cloud servers and then there is less control over the data. It is then the responsibility of the cloud provider to ensure the data security and privacy. As the cloud services are accessed through internet, there is risk of attack by the hackers to steal in and cause the damage. It is, therefore, necessary to carefully choose the cloud provider and verify the methods and tools that are being used by the cloud to ensure the security and privacy to our data. Company should ensure some regulations and agreement with the cloud provider to enforce the data security.

Sensitive data: When these organizations adopt cloud computing, they need to migrate their sensitive data to cloud and if this data leaks, this causes big damage to the business. So, the migration of sensitive data to cloud is a big challenge and it needs great skills to handle and manage things.

Portability: Portability is the challenge in cloud computing which means that the applications should migrate easily from one cloud to another cloud. The software that we wish to move should be potable with the other cloud environment. This portability is difficult to achieve as the cloud providers use different standard languages for their respective platforms.

Vendor lock-in: It is a major barrier for most of the companies to migrate to cloud computing because of the lack of standardization. In vendor lock-in, customer wants to switch the vendor are bound to pay for changing the vendor. Vendor Lock-in is noticed in every technology till now despite number of efforts like open system and platform independent language.

Organizational aspects: The business organization adopting the cloud technology will undergo lot of changes in the internal working, mission, funding and staff etc. As the business is moving to the cloud, there will be new challenges for administration and the operational staff. There will be new security and privacy challenges as the data is stored on remote server. The employee of the company should be happy and confident of the new technology to explore the advantages and overcome the issues. The overall organizational setup will change and the employees at different levels should welcome this positive impact to achieve the new business heights [20].

Regulatory and laws restrictions: Before migrating to cloud, the business organization needs to keep in mind the different legal and regulatory issues. These laws ensure the commitment and responsibility of both clients and tenants. In few countries of the Europe, Government regulations cannot allow the personal data of customers or other sensitive information to be physically store outside country. The need to build exclusive centers may not be feasible for the cloud providers and may prove a big challenge for them. If sensitive data is not protected and some leakage occurs, then the defined laws and regulations will decide who is responsible for the fault and how to compensate.

Solutions to overcome these challenges

Despite number of benefits of cloud computing, there is low migration rate because of different challenges faced by the companies. Following are some solutions that a company should follow to overcome the challenges of cloud migration.

Use encryption technique: The best way to avoid data from snooping while sharing between business network and cloud devices is to encrypt it. The data is always transferred from local server to cloud server in an encrypted format so that even if it is hacked appears as trash to them.

Alternate backups: Companies should keep regular backups at their local storage devices or at the alternate cloud provider. There are so many cloud providers that are offering storage as a service to their clients.

Hiring skilled and experienced cloud professionals: As the companies are migrating their business to the cloud, there should be a team of skilled cloud professionals who know how to handle this technology efficiently. Hiring such team is foremost thing for a company who has sound knowledge about the technology and the standards used in technology.

Use the services of Security Service Providers: There are number of clouds provides in the market that is providing security as a service to their clients and company should use these services to ensure the security and privacy to their data.

Invest in education: Companies should educate their employees towards the goals of new technology in order to help learners to cope with the new adopted technology. Once the users get enough knowledge regarding positive impacts of the technology, they can render their full cooperation to make it a success in all respects [19].

Proper cloud provider selection: There are number of cloud service providers in the market who are competing and the company should select a service provider after proper investigation of the track record of the provider. Organization should verify the tools and techniques used by the cloud provider and choose the best one to meet their business requirement. Before migrating data to the cloud, the company should ask number of questions to the cloud provider regarding security and privacy measures taken, firewall setup, antivirus, testing methods etc. to ensure the data security and privacy threat in future.

Cloud auditor to audit the services: Third party control or organizations should be there to audit the cloud services and other performance including security. The Cloud Auditor should create an audit plan that includes policies and procedures as a reference guide.

Selective migration of business applications: company should first decide that which applications they need to migrate to cloud and which ones to keep on local devices. Before migration, they need to verify whether the application is finance based, is it customer data or business data because all the applications may not be cloud compatible.

ETHICAL ISSUES IN CLOUD COMPUTING

An important business hosts a file or a program on a cloud network. It may be possible for unauthorized people to access or copy that file or program. And if that cloud network goes down for a while, or permanently, then the company has lost that file or program until the network goes back up - perhaps permanently - unless they also have another copy stored somewhere else. And if the company stops using that cloud network, but have left files on it, then the problem of who has the rights over those files could be a major issue.

- Cloud computing implicates several ethical obligations.
- Does the provider take precautions to protect data and guard against unauthorized attempts to access data?
- Does the provider encrypt confidential data?

Due Diligence before Using Cloud Services

Legal service providers should consult with computer security experts before using cloud computing Services. They should understand the potential security threats they face and identify both best available security measures and generally accepted commercial security standards. Legal service providers should assess the different levels of security sensitivity associated with their data and communications and develop appropriate plans to manage the different types of data and communications appropriately. They should determine which materials, if any, should also be retained in hard-copy, paper form.

Establish Appropriate Relationships with Cloud Service Providers

Substantial care should be applied in the selection of cloud service providers. Only reputable and reliable service providers should be used. Written, fully enforceable service agreements should be used to define all the terms of service. The legal service provider should fully understand all aspects of the terms of service, particularly the provisions applicable to data security, procedures associated with security breaches, upgrades and enhancements to security measures, and back-up systems and processes in the event of security breaches or other service failures. The terms of service should also provide that ultimate control over the data involved, including ownership and control of the data when the service arrangement ends, should rest with the legal service provider, not the cloud service provider. Adequate and enforceable non-disclosure provisions for confidential and proprietary material should be included in the terms of service.

Special Care for Extremely Sensitive Material

Legal service providers should take exceptional steps to secure their most sensitive data, communications, and other materials. Those special materials should be identified and evaluated specifically by legal service providers before being included in cloud computing systems. As appropriate, separate data management and security practices should be applied to the most sensitive materials. In some cases, specific client approval should be obtained prior to including highly sensitive client materials in cloud computing networks.

Continuously Review and Enhance Security Measures

Legal service providers should conduct ongoing monitoring of the security of their cloud computing activities. They should work with their cloud service providers to anticipate potential security threats and to analyze fully all security breaches they encounter. Terms of service applied to cloud computing services should be reviewed and re-examined on a regular basis. Legal service providers should remain informed of advances in cloud computing security technologies and capabilities.

Business impact of cloud

More mature markets such as the USA suggest the adoption of cloud platforms is likely to be between 50 and 75 percent of relevant ICT requirements. At 75 percent adoption, and based on current Australian GDP, this would result in opex and capex savings of 25 percent and 50 percent respectively.

The transformative impact of cloud is transformative in that it is creating new business opportunities as companies harness its power to efficiently facilitate new revenue, services and businesses. It is breaking down barriers in the supply chain and creating more effective and timely interaction between clients and suppliers.

It is delivering speed, agility and cost reduction to IT and other functional areas within the enterprise. Its transformative impact can readily be seen in areas such as HR, CRM and IT Infrastructure.

Addressing the risk equation any change requires an evaluation of the inherent risks and a strategy to mitigate them. The cloud is not different.

In addition to service delivery models and policies, a robust matrix should consider the storage of sensitive data, regulatory compliance issues, new data security and data controls and disaster recovery scenarios. Tax implications and international data access laws should also be considered.

The Economical Impact of Cloud

• **Cost of power**. Electricity cost is rapidly rising to become the largest element of total cost of ownership (TCO), 5 currently representing 15%-20%. Power Usage Effectiveness (PUE) 6 tends to be significantly lower in large facilities than in smaller ones. While the operators of small data centers must pay the prevailing local rate for electricity, large providers can pay less than one-fourth of the national average rate by locating its data centers in locations with inexpensive electricity supply and through bulk purchase agreements. 7 In addition, research has shown that operators of multiple data centers are able to take advantage of geographical variability in electricity rates, which can further reduce energy cost.

• **Infrastructure labor costs.** While cloud computing significantly lowers labor costs at any scale by automating many repetitive management tasks, larger facilities are able to lower them further than smaller ones. While a single system administrator can service approximately 140 servers in a traditional enterprise, 8 in a cloud data center the same administrator can service thousands of servers. This allows IT employees to focus on higher value-add activities like building new capabilities and working through the long queue of user requests every IT department contends with.

• **Security and reliability.** While often cited as a potential hurdle to public cloud adoption, increased need for security and reliability leads to economies of scale due to the largely fixed level of investment required to achieve operational security and reliability. Large commercial cloud providers are often better able to bring deep expertise to bear on this problem than a typical corporate IT department, thus actually making cloud systems more secure and reliable.

• **Buying power.** Operators of large data centers can get discounts on hardware purchases of up to 30 percent over smaller buyers. This is enabled by standardizing on a limited number of hardware and software architectures. Recall that for the majority of the mainframe era, more than 10 different architectures coexisted. Even client/server included nearly a dozen UNIX variants and the Windows Server OS, and x86 and a handful of RISC architectures. Large-scale buying power was difficult in this heterogeneous environment. With cloud, infrastructure homogeneity enables scale economies

BUSINESS IMPACT OF CLOUD

Cloud computing have positive impact on the business organizations as it increases their revenue and helps them to achieve the business goals. Companies prefer to use the services offered by the cloud rather than building their own infrastructure. Following are the benefits of cloud computing technology which motivates the business organizations to migrate from local infrastructure to cloud

Reduced cost: Cloud computing reduces the expenses of the company as the resources are only acquired when needed and only paid for when used as the billing model works as per usage and there is no up-front cost. The infrastructure is not purchased and thus lowering the initial expenses and maintenance cost as well. The clients are not the owners of the infrastructure but can use the cloud services online.

Unlimited scalability: This is the major benefit of cloud technology as the client has the flexibility to scale up or scale down as per the needs of the organization. The companies do not need to worry about the future demands as they can easily acquire the additional services anytime. Also, if a business grows over time, the cloud can scale effortlessly to meet the increased demand over time.

Flexibility: Cloud computing provides lot of flexibility to its clients. There is an easy testing and deployment of the services over cloud. The customers are free to decide which services they need and pay for accordingly. The cloud services can better meet the changing business demands by providing various services. If any application provided by the cloud is not getting our job done, we have the flexibility to switch to another cloud.

Better mobility: The users of the cloud can access the services of the cloud anytime anywhere from a variety of devices. Whenever they have the working internet connection, they can login and use the services. This benefit of cloud computing provides a flexible work culture to the employees and they can perform their duties from anywhere without the need to be physically present at the business headquarter.

Improved communication: Cloud computing improves communication and collaboration among employees by having access to instant messaging, conferencing and video conferencing options. They can jointly work on documents and projects ensuring higher cohesion and team work. This is possible because of data centralization and updation of cloud servers in real time.

Reliability: As the services of cloud is available all the time and can be accessed anywhere. Also, the backup and recovery management make this technology more reliable.

Increased storage: Some cloud providers offer the Storage as service to its customers. Companies can store lot more data on cloud than their local devices. If business grows and demands more storage, companies effortlessly scale up and get more storage from the cloud provider.

Easier upgrades: It is the responsibility of the cloud providers to upgrade the infrastructure and services for their customers. The new business trends and solutions are made available to the clients by cloud providers to compete in the business market by adopting latest technologies. The cloud providers maintain the system by doing different software and security updates.

Disaster recovery: Companies using cloud services need not to frame the complex disaster or failure recovery plans as the service providers take care of such issues and put the clients out of the trouble in a fast manner.

Security: The most important factor while choosing a cloud provider is the security and privacy they provide to our data. This is the main reason that these providers invest large amount on their services and infrastructure to offer better security.

The Future of cloud computing

Cloud computing has been called the way of the future. It opens doors by making applications and technology more accessible than in previous years. Companies that would normally require enormous amounts of startup capital may only need a fraction of what was previously required to succeed.

The past of cloud computing is bright, but the future of cloud computing is even brighter. Here is what you may need to know about trends in cloud computing.

Proactive Application Monitoring

Proactive application monitoring technology is currently available, but predictive technology and software will soon make this more robust and accurate. Companies will be able to foresee disaster and avert it, mitigating damage to their systems. This will prevent downtime and make the company safer.

Technology to Ensure Uptime

Companies need uptime guarantees. With low-power processors, data centers will become more affordable, allowing companies to acquire seven to ten data centers around the world in different time zones and thereby allowing them to guarantee 99.9 percent uptime. This will keep companies from losing money and falling prey to their competitors. In the future, this concern will be near obsolete. Because of this, many small hosting companies like GloboTech Communications are offering cloud services to ensure better uptime.

Cloud Computing’s Role in Disaster Recovery and Remote Access

Cloud computing enables and enhances remote access and faster disaster recovery. When companies have an emergency information security strategy with security penetration tests, companies can maintain their competitive edge within their respective industries. With cloud computing, some companies that didn’t recognize a breach may recover within minutes instead of hours. Losing proprietary data can cripple a company and even cause doors to close. Every company should migrate to cloud computing for this reason.

The Ability to Validate Identities through Trusts

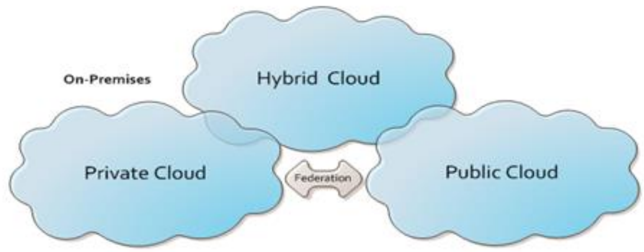
In the future, cloud security systems will be able to validate identities through a “centralized trust.” Identity-based security is thought to be more secure than current forms of security. More people will begin to trust cloud computing when this happens in future.

Centralized Data is the Future in Cloud Computing

Centralized data is expected in abundance in the future of cloud computing. This allows companies to create huge databases. Patient care can be improved with centralized data in huge databases. Better stock market decisions are also possible. Centralized data is a way of the future.

Hybrid Cloud Computing Increases Efficiency

Hybrid cloud computing is expected to help businesses become more efficient by optimizing business process performance. Businesses are excited about hybrid cloud computing because it enhances internal infrastructure and applications. The ability to scale the strengths of local networks and cloud computing is coveted by designers.



Faster Interconnects

Cloud computing is still in its infancy stage. By 2020, cloud computing is expected to be a permanent solution in many organizations. Data centers will be automated and will support scalable software architecture.

Networking Support for Cloud Computing

Cloud computing is a technique of resource sharing where servers and storage in multiple locations are connected by networks to create a pool of resources. When applications are run, resources are allocated from this pool and connected to the user as needed. The missions of connecting the resources (servers and storage) into a resource pool and then connecting users to the correct resources create the network's mission in cloud computing.

Public cloud access networking

First, public cloud access networking is most often through the Internet, though some cloud providers may be able to support virtual private networks (VPNs) for large customers. Accessing public cloud services will always create a tension between security and performance. Not all cloud computing providers will support encrypted tunnels, so information may be sent in the open on the Internet. Where encryption is available, using it will certainly increase delay and may impact performance.

Private cloud access networking

The second network consideration is private cloud access networking. Most often, enterprises will access their own private clouds using the same technology they employed for access to their data centers. This may include an Internet VPN or VPN service from a network operator. If application access was satisfactory in a "pre-cloud" configuration, a transition to private cloud computing is not likely to impact access performance.

Private intra-cloud networking

The final and most complicated issue for networking in cloud computing is private intra-cloud networking. What makes this particular issue so complex is that it depends on how much intra-cloud connectivity is associated with the applications being run. At a minimum, all cloud computing implementations will rely on intra-cloud networking to link users with the resource to which their application was assigned. Once the resource linkage is made, the extent to which intra-cloud networking is used depends on whether the application is componentized among multiple systems.

Ubiquitous Cloud and Internet of things

Ubiquitous cloud computing refers to the use of Internet resources at any place and any time for any objectives. Today, people can access the Internet via a fixed wire or a mobile wireless connection.

Cloud Service Trends

Summarizes cloud offerings from eight IT leaders and some startup cloud companies. We already studied the cloud offerings from Amazon, Google, Microsoft, GoGrid, and Rackspace. Here we examine the remaining cloud service

IBM RC2 Cloud

The computer and IT infrastructures of eight IBM Research Centers are now strongly connected to form a private cloud, called the Research Compute Cloud (RC2). RC2 is web-based, allowing more than 3,000 IBM researchers worldwide to share computing resources. It is also IBM's internal testbed for promotion of cloud technology. RC2 provides solutions to establish an autonomous computing environment on user demand.

The system supports machine virtualization, service life-cycle management, and performance monitoring. The idea is to reduce computing costs by outsourcing jobs to best-fit sites for rapid resource deployment, and thus optimal execution. The project also evaluates green energy savings and dynamic resource optimization.

Cloud System from SGI

In February 2010, SGI announced the Cyclone, a large-scale on-demand cloud computing service for high-performance computing (HPC) applications. Both SaaS and IaaS models are offered on the Cyclone. The SaaS model offers users prepackaged technical applications covering a wide range of domains.

For example, the system can be used to perform computational fluid dynamics (CFD) and finite element analysis in both the airplane and automobile industries. With the IaaS model, users can access SGI's fastest Altix servers and ICE clusters, which are networking and storage systems managed and optimized by SGI experts for specific user applications.

Some IT Companies Providing Cloud Services or Products

Company, Year	Major Cloud Offerings	Major User Groups
Amazon Seattle, 1994	Amazon Web Services (AWS); six Infrastructure-as-a-Service (IaaS) systems including EC2 for computing capacity and S3 for on-demand storage capacity	10,000 businesses and individual users including the *New York Times*, *The Washington Post*, and Eli Lilly
Enomaly Toronto, 2004	Elastic Computing Platform integrating enterprise data centers with commercial cloud offerings, managing internal and external resources with virtual machine (VM) migration	Business Objects, France Telecom, NBC, Deutsche Bank, Best Buy
Google Mountain View, CA, 1998	GAE offering Platform-as-a-Service (PaaS) capability plus office productivity tools including Gmail, calendar tools, the Postini web site creation tool, and security protection services	Small businesses, enterprises, and colleges including Arizona State University and Northwestern University
GoGrid San Francisco, 2008	Offers web-based storage and deploys Windows- and Linux-based virtual servers in the cloud with preinstalled software from Apache, PHP, and Microsoft SQL	Mostly startups, Web 2.0, and Software-as-a-Service (SaaS) companies, plus a few big names like SAP and Novell
Microsoft Seattle, 1975	Azure offering a Windows-as-a-service platform consisting of the OS and developer services that can be used to build and enhance web-hosted applications	Epicor, S3Edge, and Micro Focus using Azure to develop cloud applications
NetSuite San Mateo, CA, 1998	A business software suite including e-commerce, consumer relationship management (CRM), accounting, and enterprise resources planning (ERP) tools	Business customers such as Puck Coffee and Wrigleyville Sports
Rackspace San Antonio, TX, 1998	Mosso cloud offering a platform for building web sites; cloud files for a storage service; and Cloud Servers, an EC2-like service that provides access to virtualized server instances	Web developers and SaaS providers such as Zapproved, which uses Mosso to deliver an online productivity tool
Saleforce.com San Francisco, 1999	Offers CRM tools for sales force automation, analytics, marketing, and social networking; Force.com offers a PaaS for building web applications on the Salesforce.com infrastructure	500,000 customers in financial services, communications, and media, energy, and health care

Large-Scale Private Clouds at NASA and CERN

In 2010, two large-scale private clouds were under construction in the United States and European Union. We discuss them in this section to demonstrate the scalable growth of cloud computing platforms. The U.S. cloud, called Nebula, is developed by NASA and is designed for NASA scientists to run climate models on remote systems provided by NASA. This can save thousands of NASA users from acquiring supercomputers at their local sites. Furthermore, it enables NASA to build the complex weather models around its data centers, which is more cost-effective.

The EU cloud is built by CERN in Geneva. This is a large private cloud for distributing data, applications, and computing resources to thousands of scientists around the world. CERN deals with large data sets and throughput with a global workforce and a finite budget. Many private clouds are also under development by the IT industry, large enterprises, and other government agencies. Initially, one can see that most clouds are developed with restricted user groups. Once these clouds become mature and more secure to use to prove their claimed advantages, they could be converted to public clouds in the future.

CERN Super Cloud

CERN has long been an influential leader in solving complex IT problems. In fact, the World Wide Web owes a lot to CERN for its usefulness in getting some early work done there. Now, the focus is on cloud computing, CERN scientists plan to smash two particle beams together in an attempt to re-create the “big bang” event to uncover the enigma of many fundamental physics phenomena. The CERN scientists realize the use of cloud computing over their large data centers may be more cost-effective in performing the required simulation experiments.

Tens of thousands of scientists around the world are feeding off the huge data sets and doing all kinds of research by themselves. CERN believes the cloud project will allow it to deliver increased computing performance and offer better infrastructure services to its 10,000 researchers from 85 countries. At CERN, massive amounts of scientific data are processed and must be distributed to researchers in near real time. As a result, CERN’s cloud infrastructure has to provide the capacity necessary to support production and analysis of more than 15 petabytes of data per year, processed by 60,000 CPU cores, allowing scientists to manage workloads themselves as opposed to a centralized IT management department at CERN’s laboratory near Geneva.

Because CERN uses the platform’s Load Sharing Facility (LSF), LSF is used to balance the workload on computational server, grid and workload management solution to enable extensive scalability to analyze its vast amount of research data, the laboratory chose to partner again with the platform to explore how to more effectively utilize their resources in a virtualized cloud environment. Platform LSF and its adaptive cluster provide an open, low-cost, common platform for CERN’s scientists, allowing management of both virtual and physical servers in the cloud. In addition, scientists can manage their own application environments and control projects dynamically for maximum flexibility and efficient workload processing.

Cloud Mashups for Agility and Scalability

In web application development, a mashup is a web page or application that combines data, presentations, or functionality from two or more sources to create a new service. The main characteristics of the mashup are a combination of virtualization and aggregation. In cloud computing, the cloud has captured the computing market with dynamic resource allocation from a pool of VM resources.

Not only are AWS and GAE different in terms of their functionalities, but they also can complement each other. This has triggered the idea of mashing up different clouds to build an *intercloud* or *cloud of clouds* dynamically. We will demonstrate how to mash up GAE and AWS clouds to achieve the desired features of application agility and performance scalability. In fact, cloud mashups offer a more cost-effective solution to new startups that do not want to invest in cloud hardware and software to create their own enterprise-level data center or private cloud. A cloud mashup must be designed to support popular paradigms such as MapReduce on AWS, and yet be controllable via GAE through easy-to-operate web interfaces.

The Idea of a Cloud Mashup

Any startup in today's business world needs to keep its operational costs low in the initial years. This allows the business to recover its initial investments and post profits quickly. This, in turn, prompts venture capitalists to fund the company in hopes of a quick return on investment and the ability to supply more capital to allow the company to grow. Thus, it is critical that the company keep its initial costs low. Consider a startup company working as a social networking portal. I t needs lots of data storage space and servers, as well as associated cooling requirements and infrastructure such as a physical building to house its equipment. This leads to a moderate to large initial investment on the part of the company.

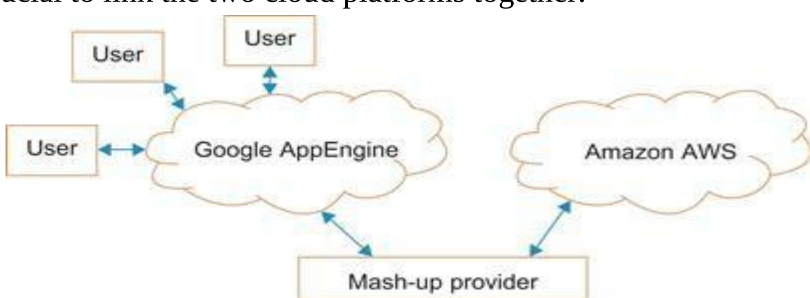
One can use the pay-per-use models from public clouds such as AWS and GAE to quickly bring the business up to speed. This can not only ensure that the loss due to time-to-market is reduced, but it can also enhance the speed with which one can unleash a new idea into the business world. Business mashups are also popular today as more and more businesses can complement each other nicely. Today's mashups are typically dataflow systems with the interaction scripted in languages such as JavaScript or PHP.

Example

A Mashup Cloud Designed on the AWS and GAE Platforms

Figure shows the concept of mashing up the GAE and AWS platforms. Essentially, one gets user input files from GAE and uses the AWS infrastructure for MapReduce. To prove the effectiveness of mashups, the performance of MapReduce on the mashup platform is experimented with scalable EC2 resources.

The mashup provider is crucial to link the two cloud platforms together.



Cloudlets for Mobile Cloud Computing

A low-cost infrastructure to enable cloud computing using mobile devices. The idea is called Cloudlet, and it offers a resource-rich portal for upgrading mobile devices with cognitive abilities to access distant clouds. This portal should be designed to be trustworthy and use VMs to explore location-aware cloud applications. The idea can be applied to opportunity discovery, fast information processing, and intelligent decision making on the road. Cloudlet makes it possible for mobile devices to access the Internet cloud easily in cost-effective mobile computing services.

The Cloudlet Idea

Figure illustrates the idea of using cloudlets for mobile cloud computing. Both mobile devices and centralized clouds or data centers have shortcomings in terms of supporting mobile computing. Mobile handsets face problems of limited CPU power, storage capacity, and network bandwidth on smartphones and tablet computers. Mobile devices cannot be used to handle large data sets. However, the distant cloud on the Internet faces the WAN latency problem. Clouds have to solve this collision problem of too many (millions) clients logging in to the cloud simultaneously.

To solve this two-sided problem cloudlets are being deployed at convenient sites, such s coffee shops and bookstores, similar to access points for WiFi services for connecting to the Internet. Widely deployed cloudlets enable distributed cloud computing and extended resource handling for these users. The idea is to use the cloudlet as a flexible gateway or portal to access the distant cloud. The cloudlet can be implemented on PCs, workstations, or low-cost servers.

Differences between Cloudlets and Clouds

Cloudlets are decentralized and self-managed, like a "virtual data enter in a box." Cloudlets operate with LAN latency and bandwidth with only a few users at a time. Clouds are centralized at large-scale data centers. They are professionally administrated all the time and they need large machine rooms with uninterrupted power and cooling. The main problem for a distant cloud to be accessed by mobile devices lies in the long latency experienced with the Internet or WANs. The problem becomes more complicated when thousands or millions of users simultaneously demand services.

Differences between Local Cloudlets and Distant Clouds

	Cloudlet	Cloud
State	Only soft state	Hard and soft state
Management	Self-managed; little to no professional attention	Professionally administered, 24×7 operator
Environment	"Data center in a box" at business premises	Machine room with power conditioning and cooling
Ownership	Decentralized ownership by local business	Centralized ownership by Amazon, Yahoo!, etc.
Network	LAN latency/bandwidth	Internet latency/bandwidth
Sharing	Few users at a time	100s–1000s of users at a time

Ubiquitous Computing

Ubiquitous computing is a post-desktop model of human-computer interaction in which information processing is integrated into everyday objects and activities. For daily activities, people may engage in using many pervasive devices simultaneously. They may not even be aware of the existence of the interactive devices. Although the idea is simple, its application is difficult. If all objects in the world were equipped with minuscule identifying devices, daily life on our planet could undergo a major transformation.

The IT cannot be realized without systems design and engineering, and user interfaces. Contemporary human-computer interaction models, whether command-line, menu-driven, or GUI -based, are inappropriate and inadequate to meet ubiquitous computing demands. The natural IT paradigm appropriate to a ubiquitous computing world has yet to emerge. Contemporary devices that lend support to ubiquitous computing include smartphones, tablet computers, sensor networks, RFID tags, smart cards, GPS devices, and others.

Development of the Internet of Things

Radio-Frequency Identification (RFID)

RFID is applied with electronic labels or RFID tags on any objects being monitored or tracked. The tagging may be applied to any objects, such as merchandise, tools, smartphones, computers, animals, or people. The purpose is to identify and track the objects using radio waves or sensing signals. Some tags can be read from tens or hundreds of meters away via a wireless reader. Most RFID tags contain at least two major parts. One is an integrated circuit for storing and processing information, modulating and demodulating a radio-frequency (RF) signal, and other special functions. The other part is an antenna for receiving and transmitting the radio signals.

RFID Tags and Device Components

There are generally three types of RFI D tags: ***active RFID tag****s* containing a battery and transmitting signals autonomously, ***passive RFID tags*** which have no battery and require an external source to provoke signal transmission, and ***battery-assisted passive RFID tags*** which require an external source to wake up the battery but have significantly higher forward link capability. Based on the radio frequency used, the passive RFI D tags operate with from *low frequency* (LF) to *high frequency* (HF), *ultra high frequency* (UHF), and the *microwave range*. In terms of functionality, there are three major components of RFID hardware:

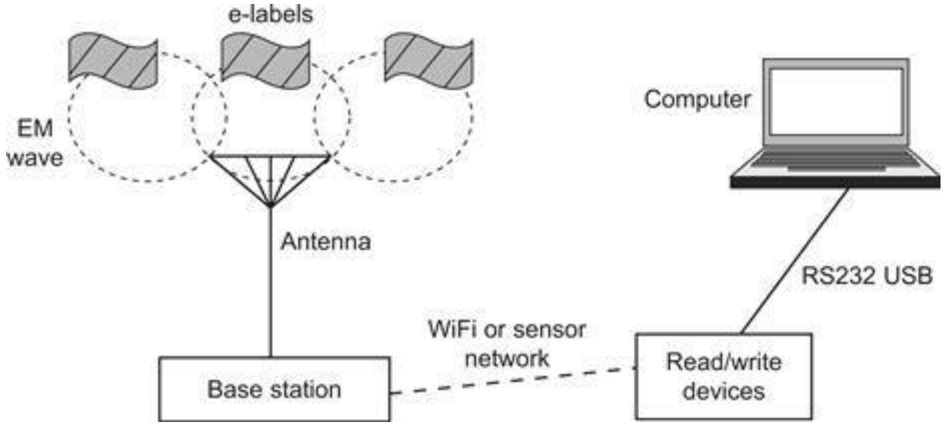
- **RFID tag** A tiny silicon chip attached to a small antenna.
- **Reader antenna** Used to radiate the energy and then capture the return signal sent back from the tag. I t can be integrated with a handheld reader device or connected to the reader by cable.
- **Reader** The device station that talks with the tags. A reader may support one or more antennae. Resembling an electronic barcode, a reader device can detect signals even without a line of sight.

Some RFI D readers can identify multiple objects concurrently. And some RFI D tag reader

How RFID Works

Innovations in RFID include *active*, *semi-active*, and *passive* RFID tags. These tags can store up to 2 KB of data and are composed of a microchip, antenna, and battery for active and semi-passive tags. The tag's components are enclosed within plastic, silicon, or sometimes glass. Data stored in the microchip waits to be read. The tag's antenna receives electromagnetic energy from an RFID reader's antenna. Using power from its internal battery or power harvested from the reader's electromagnetic field, the tag returns radio signals back to the reader. The reader picks up the tag's radio signals and interprets the frequencies as meaningful data.

There are two ways to couple schemes in RFID tags: inductively coupled or capacitively coupled. Figure shows the RFI D operations between the RFI D tags (e-labels), reading/writing devices, and backend computers. RFI D tags are still not in widespread use by ordinary people, because they are expensive and bulky. They are more often used by large businesses, moving or shipping companies, and service companies. However, RFI D tags can make our daily life or work much easier, more convenient, and accessible to visible or hidden objects surrounding us. These tags are not that expensive to produce in large quantities, and they can be made small enough to fit on almost any product or object.



External RF antenna devices can illuminate the tag devices similar to radar illuminating a target. However, RFID is more effective at operating at shorter ranges.

Wireless Network Support for Ubiquitous Computing

The ZigBee networks are mainly used in low-cost and low-speed monitoring and control applications, such as those used in a *wireless home area network* (WHAN). The GSM/GPRS or CDMA/1 networks are cellular mobile networks that can cover a wide area for voice and data telecommunications. The WiFi networks specified in the IEEE 802.11b standard.

WiFi is used for wirelessly accessing the Internet, reading e-mails, or conducting web page searches Bluetooth is mainly used for short-distance wireless connection of computer peripherals (e.g., keyboard, mouse, printers, etc.). In terms of data rate, WiFi is the fastest (54 Mbps for 802.11g network). The next is Bluetooth with 720 Kbps and 115 Kbps for the 2.5G GPRS mobile network. ZigBee has the slowest rate of 20–250 Kbps. However, ZigBee has high reliability and low power/cost advantages.

Global Positioning System (GPS)

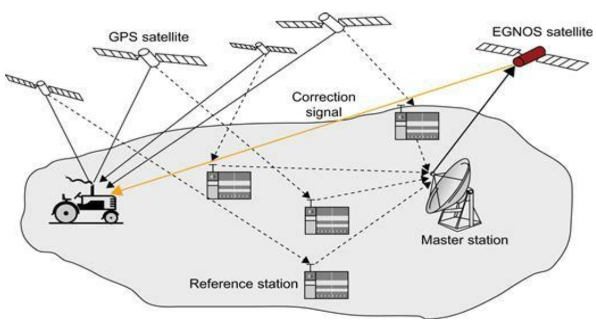
Location-based service (LBS) helps people and machines find things and determine their physical whereabouts. Sensors play a role in dead reckoning, but that approach does not satisfy practical needs for geolocation. The GPS is augmented by other signals and cellular towers. Fixed or orbiting satellite transmitters broadcast timing signals and receiving device response signals to locate the position of moving objects.

Radar, lidar, and sonar can all detect relative locations of things using electromagnetic, optical, and acoustic signals. With help from satellites, some objects can be also located by attaching active GPS devices to disclose their locations by radio, light, and/or sound.

How GPS works

GPS consists of three parts: the *space segment*, *control segment*, and *user segment*. There are 24 satellites deployed around the Earth in fixed orbits. The satellites orbit at an altitude of approximately 20,200 KM. GPS satellites broadcast signals from space, by which each GPS receiver calculates its 3D location (*latitude, longitude, and altitude*) plus the current time.

- The space segment is composed of 24 satellites in medium Earth orbit and also includes the boosters required to launch them into orbit. GPS satellites circle Earth twice a day in very precise orbits and transmit signals to Earth. GPS devices on the ground receive these signals and use a triangulation method to calculate the user's exact location. At different times, different satellites will become visible to the receiver. As shown in Figure.
- The control segment is composed of a master control station and a host of dedicated and shared ground antennas and monitor stations.
- The user segment is composed of hundreds of thousands users of the secure GPS precise positioning services. Tens of millions of civil, commercial, and scientific users are only allowed to use a degraded functionality of the so-called standard positioning services that cannot be used in hostile attack purposes.
- Essentially, the GPS receiver compares the time a signal was transmitted by a satellite with the time it was received. The time difference tells the GPS receiver how far away the satellite is. With distance measured from several satellites, the receiver can determine the user's position and display it on the unit's electronic map.



Example

Real - Time Vehicle Tracking Using Active GPS Devices

Active GPS tracking systems find their applications primarily in industries and businesses. These systems are fast becoming the standard for companies that wish to monitor fleet vehicles as well as other heavy equipment. Real time GPS tracking is practical for acquiring immediate and detailed information about large numbers of vehicles or objects that are being tracked. An example is a car rental business that provides cars for many customers; real-time vehicle tracking systems could prove beneficial in supplying current and historical data on routes taken by all the drivers. The process is divided into the following four steps:

1. GPS receivers in each car asset receive signals from a network of satellites.
2. The collected satellite information is sent to the communication center via mobile networks.
3. The control center enters calculated location information over global maps.
4. The control center sends commands to each unit to trigger alarms, stop engines, change direction, or send personal messages, and so on.

Applications of the Internet of Things

Specific Wireless Sensor Applications

Highlighted below are several applications of sensor networks:

• Military sensor networks to detect and gain as much information as possible about enemy movements, explosions, and other phenomena of interest

• Sensor networks to detect and characterize chemical, biological, radiological, nuclear, and explosive attacks and materials

• Sensor networks to detect and monitor environmental changes in plains, forests, oceans, and so on

• Wireless traffic sensor networks to monitor vehicle traffic on highways or in congested parts of a city

• Wireless surveillance sensor networks for providing security in shopping malls, parking garages, and other facilities. Wireless parking lot sensor networks to determine whether the lot is occupied or Available

Smart Power Grid

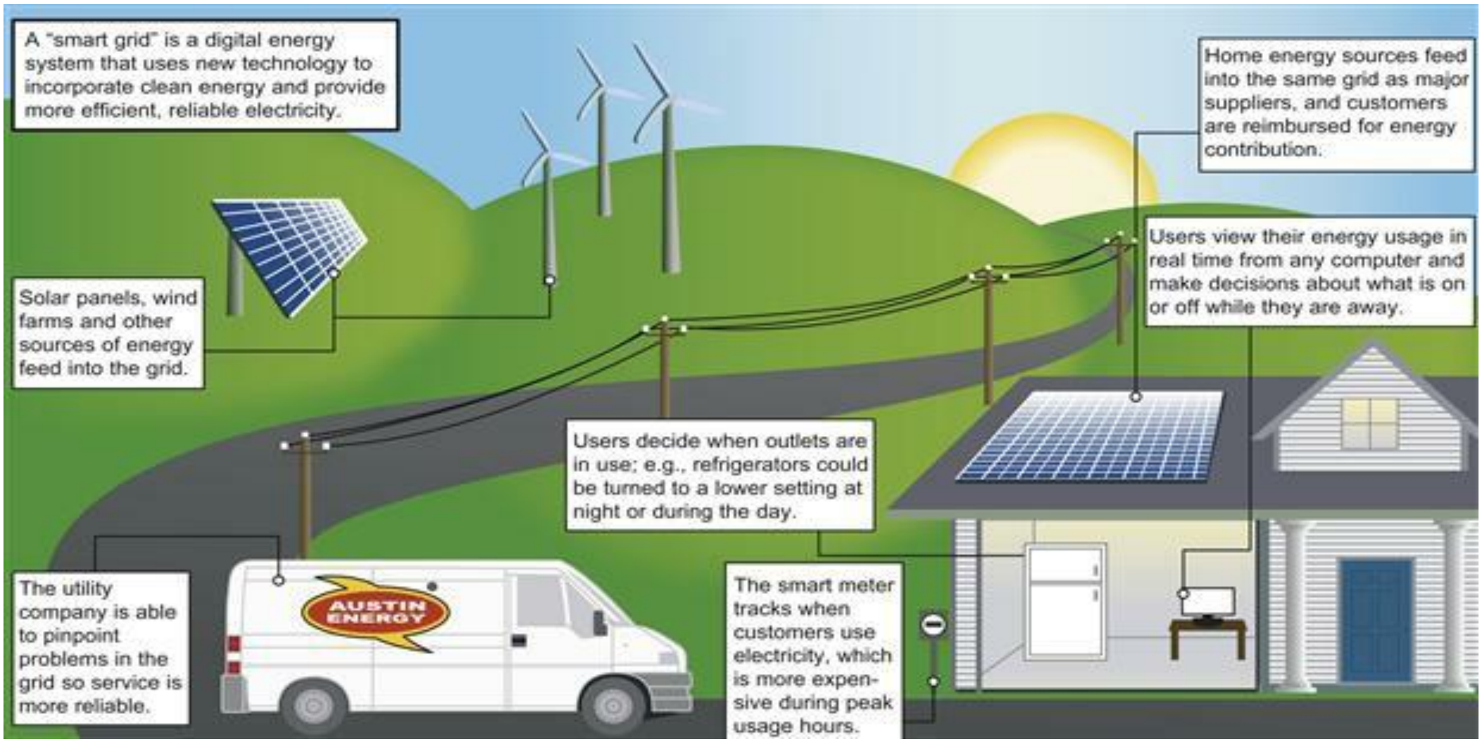
Various power companies across the United States have or are in the process of upgrading their power management and distribution systems. Various sensors at individual homes (smart thermostats) can collect information that is sent via a network to main stations (perhaps even local "hubs") that can apply complex power management and send control signals back to the grid to save energy. The smart grid is made possible by applying sensing, measurement, and control devices to electricity production, transmission, distribution, and consumption.

Example

A Smart Power Grid Supporting the Internet of Things

A smart grid includes an intelligent monitoring system that keeps track of all electricity flowing in the system. Smart meters, a digital upgrade of current utility meters, track energy usage in real time so that both the customer and the utility company know how much is being used at any given time. Energy is paid for using "time of day" pricing, meaning electricity will cost more at
peak times of use. Figure illustrates IT support of a smart power grid.

For example, when power is least expensive the user can allow the smart grid to turn on selected home appliances such as washing machines or factory processes that can run at arbitrary hours. At peak times it could turn off selected appliances to reduce demand. More involved users would be able to use the smart meter to view energy usage remotely and make real-time decisions about energy consumption. A refrigerator or air conditioning system could be turned down remotely while residents are away.


A "smart grid" is a digital energy system that uses new technology to incorporate clean energy and provide more efficient, reliable electricity.
Home energy sources feed into the same grid as major suppliers, and customers are reimbursed for energy contribution.
Users view their energy usage in real time from any computer and make decisions about what is on or off while they are away.
Solar panels, wind farms and other sources of energy feed into the grid.
Users decide when outlets are in use; e.g., refrigerators could be turned to a lower setting at night or during the day.
AUSTIN ENERGY
The utility company is able to pinpoint problems in the grid so service is more reliable.
The smart meter tracks when customers use electricity, which is more expensive during peak usage hours.