CHAPTER 2

BACKGROUND OF CLOUD COMPUTING
## CHAPTER 2

### BACKGROUND OF CLOUD COMPUTING

2.1 Market Oriented Cloud Computing  
2.2 Planning Applications on Cloud  
2.3 Understanding IT Governance in Cloud  
2.4 Security Issues in Cloud  
2.5 Aneka Platform  

**Summary**

### 2.1 MARKET ORIENTED CLOUD COMPUTING

The enables and barriers of cloud computing technology with the primary research domains including available cloud architectural models, data integration among servers and data centres, data retrieval from cloud, resource allocation and load balancing, security are discussed in detail in this chapter. The other issues with respect to the development of optimal model for effective data access in a overall cloud architecture is also explored. This chapter discusses the available literature that supports the cloud computing research followed by future directions to derive objectives for this thesis.

The one of the emerging trend in the computing world is *Cloud Computing* which was earlier named as *dujour* [38, 39] is also now known with many other jargons like Internet Computing, Market Oriented Computing etc. It is being a trend setter in IT industry, which is completely changing the traditional software development life cycle. Mobile computing is taking the best advantage of these cloud facilities like Amazon, Google Docs, etc. Using cloud based applications; mobiles needs less computing complexity when compared to desktops. Many applications like Emails, Google-apps, Microsoft-live etc., are all derived from cloud based technologies. As data access and effective resource utilization are the main issues of cloud computing, some of such techniques are mentioned below:
• **Prediction of Workload** [40] is a method of predicting futuristic workload on the servers with respect to the user requests, which needs some performance criteria to calculate the influence of some other parameters connected with cloud.

• **Reuse of Memory** [41] used to decrease the number of data transfers via VMs. It dynamically reallocates the memory and reuses the data leading to minimum processing and migration time.

• **Task and Resource Scheduling** [42] is based on many different scheduling algorithms, wherein the cloud resources are geographically dispersed on the precedence criteria and priority in clouds. Many cloud QoS parameters like demand, cost etc. are used while scheduling tasks on resources.

However, the term *cloud* came into existence in 2006 [43], but became popular in 2007 with the release of first document on the cloud by IBM. It describes the essentials of cloud of all kinds for business organizations for better gains. IBM’s model of technology adoption programme allowed its employees to utilize the services of cloud.

In PaaS, dynamic provisioning is implemented as a optimal solution which supports hybrid clouds which helps in mapping of cloud applications onto the cloud infrastructure. Aneka offers cloud provisioning service that gives the benefit of different IaaS providers for scaling up their existing cloud infrastructure based on their application requirements [44].

Community clouds provide a hybrid integration of services of various cloud computing platforms to address the concerned requirements of a community which can be an industry, a business sector or any other generic sectors. The definition of community cloud by NIST [45] is as follows:

“The infrastructure is shared by several organizations and supports a specific community that has shared concerns (e.g. mission, security requirements policy, & compliance considerations). It may be managed by the organizations or a third party, and may exist on premise or off premise”.
It may also specify the key features of the type of clouds over some common concerns that controls some business players in cloud computing and aims to combine the basic principles of digital ecosystem [47].

Digital Ecosystems are also the derivatives of open socio-technical systems which are adaptive and distributed which have some properties like scalability, sustainability and self-organization. The core aim of Digital Ecosystem is for sustainable regional development of SMEs.

As clouds facilitate data-intensive applications many programming paradigms like Map-Reduce have developed for better application development. As it is required to have some control during the application development, execution and deployment to ease the usability and effectiveness of programming many cloud management systems like Hadoop are developed. Hadoop is basically a cluster management tool developed to process large data by matching different commodity systems together to mimic distributed computation. Hadoop binds many smaller nodes together and thus forms a single compute cluster. It is required to develop programming frameworks that allow users to develop scalable and robust programming models that depend on parallel computing architecture. Further, it is required to develop APIs that work in such kind of distributed platform.

A variant of such method has been tried by the University of Santa Barbara and California [48] which uses an ontology based method instead of traditional SaaS and departs from any thing as a service concept and make efforts to define an ontology for clouds. As per XaaS, cloud is a layer architecture containing five layers i.e., hardware, software kernel, software infrastructure and environment and finally application, wherein each layer handles the need of a different category of cloud users. This five layer architecture provides a better interaction model among different cloud entities, both on the semantic and functional levels.

The Open Virtualization Format [49] provides a common format for storing the data and also provides a uniform structure to describe the VM related meta data information.
Although the OVF provides complete specification for packaging and distributing VMs images in a platform independent fashion, this facility is provided by very few vendors.

The researchers in [50] reviewed many cloud related and supported technology areas like, IoT, data centres, and Hadoop. As data generation, acquisition, storage and analysis are the major parts of big data computing, cloud gives the necessary infrastructure support for data analytics. Many applications areas of big data are taking new phase lift in IT industry like IoT, health care, social networks, smart grid, enterprise management, etc.

The work in [9, 50] addresses major barriers towards the success of cloud and they include data transfer bottleneck, confidentiality of data, availability, data locking, data auditability, performance unpredictability, scalability of storage, quick scaling, bugs, software licensing, etc. This is also lists the major enabler technology of cloud which include VM support, APIs for software development, many cloud providers, security via encryption, better bandwidth, scalability, pay per use and offer reputation.

Majority of the software companies started providing cloud based services by the end of 2009 with their own platforms. It is very difficult for the cloud consumers to find out the differences in performance and reliability in these platforms. Peng et al., [31] briefs the major challenges and the detailed comparative study of all the major cloud platforms. Their survey demonstrates that there are no standard criteria for comparing these cloud platforms, as each claim their significance in one or other major services and also they are very much dependent on the user requirements.

Actually, now a days good number of software industries are migrating towards cloud. Also, many cloud vendors are joining to become federated cloud service providers. The major issue in federated cloud is to select optimum service provider. While, selecting reliable cloud service provider, in [51] authors have proposed a case study where in the services offered are based on both user requirement and the service provider’s capacity. This maps all the basic client requirements like memory, Disk, CPU, multithreaded CPU, encoding and compression etc. to the cloud service providers capacity in terms of servers, storage and other software features. Yet another study in [2, 3] discuss research
issues and directions towards open cloud systems, where the authors address major design issues that act as hurdles while implementing open cloud architecture [50] in virtual computing laboratory, that supports virtualization and SOA.

Cloud security issues are delt in [47, 49] that are more concerned towards data security while transmission onto the cloud. Here, the author discuss on the major design issues involving authentication, access control etc. in clouds.

SaaS also requires isolation regarding performance and hence the authors in [52] have discussed various metrics to quantify the performance related issues in isolation. It is researched that resource allocation by service providers can be dealt carefully so that performance degradation does not occur. Hence, better resource management with security concern is more required for higher performance. The different types of compute clouds are given in table 2.1 [3].

<table>
<thead>
<tr>
<th>Cloud Model</th>
<th>Benefit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Public</td>
<td>minimum capital requirement, Quick startup time, outsourcing of non-core functions to a service provider, leveraging scalability, Use of better software stack, billed by usage.</td>
</tr>
<tr>
<td>Private</td>
<td>flexibility in resource allocation; need of capital investment, on-premise data; good for leveraging existing investments that allow control of service levels, operational reporting, cost savings.</td>
</tr>
<tr>
<td>Hybrid</td>
<td>Integration of vendor and private cloud makes complex, permitting control of data, selection of provider infrastructure that are scaled as need grows and further allowing some internal control when required.</td>
</tr>
<tr>
<td>Community</td>
<td>Sharing service cost among different organizations and can be developed to permit information sharing with various organizations without any help of external network environments.</td>
</tr>
</tbody>
</table>

2.2 PLANNING APPLICATIONS ON CLOUD [22]

The following business requirements are to be considered before deploying applications to the cloud.

- Security and Privacy Requirements
• Cost / Budget Requirements
• Cloud Types like public, private or hybrid
• Backup / Recovery requirements
• Training and Capacity Building
• Reporting via Dashboard / User Interface requirements
• Access control requirements
• Data export/import requirements

There is a need for planning to meet all of these requirements and the phases involved in planning are given in figure 2.1.

![Diagram showing planning phases and key concepts in cloud computing]

Figure 2.1: Cloud Computing Planning

**Planning the Strategy:** In this phase, the issues related to the customer are analysed via two steps, i.e., (i) Value Proposition and (ii) Strategy Planning.

**Cloud computing value proposition:** In Cloud computing value proposition step, the factors influencing the customers upon the adoption of cloud are analysed along with other key issues to be solved by clients. The key factors include
(i) IT management simplification

(ii) reducing the cost of operation and maintenance

(iii) Innovation in business mode

(iv) outsourcing.

Based on the analysis of the above mentioned steps, the strategy is establishment and the corresponding document is prepared as per the client interactions.

Planning Cloud Computing Tactics: Here, to meet the business goals of the clients, the problems and risks involved in the cloud deployment are analysed. This involves development of (i) Business Architecture (ii) IT Architecture (iii) Requirements on QoS and (iv) Transformation Plan.

While developing business architecture, the risks that may be caused by deploying the cloud application are analysed in a business perspective. As a second step, the applications that support the complete business process and the technology requirements are analysed to build IT architecture that enable enterprise applications and data systems. Similarly, to develop the requirements regarding the QoS all the non-functional requirements such as security, disaster recovery, etc. are to be analysed for the better success. In the final steps proper plans to be formulated to transform the current business to cloud computing mode and there by a proper stage can be finalized to move towards cloud deployment. The cloud deployment involves two steps namely, (i) Cloud Provider and (ii) Maintenance and Technical Service. In the first step, the cloud provider is selected based the SLA and the technical service along with maintenance are provided by cloud service provider ensuring defined QoS. It is understood that in clouds, virtualization technique is used for better optimal access of resources without violating SLAs [4].

2.3 IT GOVERNANCE OF CLOUD COMPUTING

The major issue in cloud data centre management is to comply with IT governance and hence governance defines estimation of good accountability. IT governance include applying policies to both cloud users and providers, which is in line with ordinances of
the state and organizational IT policy. It also deals with different concepts and principles of cloud computing [57], which helps in managing complex cloud resources. A cloud data centre contains all the information of its remote facilities offered to its customers. It is not possible to deploy using different service providers, because the data transfer cost from one service provider to another is very huge. But, in case of federated clouds [58], this is possible. In this case, resources are shared / leased among the service providers with some agreements and affordable prices. In such case, the existing policies in resource provisioning service discriminate between different providers by mapping to various IaaS providers for a better solution.

2.4 SECURITY ISSUES IN CLOUD

As cloud is internet based different users log on to the cloud, access services, exchange information, use storage. This may lead to some scope for security breach and also the sanctity of the information can be questionable. Figure 2.2 illustrates different levels of security issues in cloud environment [59].

**Figure 2.2:** Levels of Key Security Concerns

Each level has its own key points to be addressed and plays its importance at different levels of security issues from user to infrastructure. Figure 2.3 highlights different issues requiring security along with underlying concerns.
Network Security: This deals with the security issues of underlying communication protocol over which the entire cloud service stack works. Many of the traditional threats related to network security are still valid in cloud computing. Hackers may steal the data over the communication network.

Interface Security: This deals with the issues related to the customer interface among cloud the user and service provider. The threats mainly deal with user authentication and authorization. Typically, VM interface influences the security features. Usually, better interface security features are provided by Linux based operating systems when compared to other commercial counterparts. Therefore, the interface offered by the cloud service provider must have a secured OS [60].

Virtual Machine Security: As many users use VM to deploy their tasks and they switch from one VM to another VM by the PaaS layer, Cloud Service Provider must provide enough security features while switching from one VM to another VM. Because of multi-tenancy feature of cloud, the same VM can be shared by various consumers at different points of time in order to have better utilization at reduced cost. But, such feature increases the scope for security breach [59]. Hence, different
consumers of same VM are to be cornered in order to protect the confidentiality of individuals.

**Compliance:** It deals with the degree of user satisfaction with respect the compliance against the slated SLAs. As SLA is the legal contract, which provides the agreed upon services to the user, there is a mandate by the cloud service provider to comply on the agreed upon SLA, failing which cloud service provider need to pay penalty to the user. But, SLAs are not standardized; however, lapses in developing standards by the service providers may lead to some security breach.

**Confidentiality/Privacy:** As cloud providers need to provide the data confidentiality and privacy to their customers against the unauthorized users. As the data being propagated geographically in cloud, ensuring confidentiality to end users is a challenging task. Various methods involving cryptography, data splitting, etc. are being in use to provide data confidentiality [59, 61]. While in cryptographic method, data is encrypted and passed on to the user, in case of data splitting the it is stored at various non-interacting host machines.

**Security Algorithms:** In [62], the issues of consistent optimization of allocation of resources and network awareness are highlighted, where the emphasis is more on the making the existing performance models more predictive and responsive. Safiriyu *et al.*, in [63] have developed a User Identity Management Protocol (UIDM) in cloud scenario, which includes both the end users and providers. It encompasses the security issues like encryption, authentication, and key management mechanism. Zhen *et al.*, in [64] have proposed a network security prototype system which uses a collaborative scheme centrally with facility of packet inspection at various levels of security that being used in multi-tenant data centers. This prototype secures the data center from variety of network attacks. The centralized security center in [64] can collect data from the networks and deploy security rules at different levels, but cannot determine violations of network policy.

Philipp *et al.*, [65] have designed a model that provides improved data processing called as Virtual Fort Knox and integrated security, which is suitable for SMEs. It basically provides security from tampering of server, administrative failure and access control mechanism.
The mechanism as in [66] is a security enhanced VM based architecture which is divided update checker and online penetration suite layers as in figure 2.4. While update checker checks for the outdated information on VM, the online presentation suite scans all the VMs and reboots them if required. Finally, the report generator generates the flaws in line of risk level, from which the errors can be easily detected and removed. The limitation of this model is that it works only in Linux environment.

![Update Checker Architecture](image)

**Figure 2.4:** Update- Checker Architecture

Sapuntzakis *et al.*, in [67] have developed a prototype that assigns VMs automatically, which prevents many security breaks, cannot update all packages of VMs. Qian *et al.*, in [68] have developed a better storage architecture for cloud that can be publicly auditable and helps to determine capabilities and expertise of owner of the data to gauge the risk of outsourcing the same to the third party for audit. The architecture consists of four components that include third party auditor, cloud server, user and the data owner. Here, the data owner will appoint the third party auditor for cloud server auditing with respect to end user requirements against its effectiveness and efficiency. But, in case of security issue, as the auditing relies on the third party auditor security cannot be completely guaranteed. Again, the issue of responsibility arises when it comes to deliver of report to the end user, whether it is the owner or the third party auditor.

This mechanisms available in [69, 70] discuss about the dynamic data driven architecture that minimizes the SLA violations by releasing resource provision. Here, the focus is mainly on resource releasing, but the security issues are not dealt properly. Cloud
Security Alliance in [71] basically meant for checking the different audit policies and government strategies to ensure proper cloud security. It also releases new policies along with ISACA and NIST.

In article [72], the authors have studied different security features in clouds. Now, majority of the organizations are having a typical master slave HDFS architecture, wherein, name node is the master and the data node is the slave. Usually, the data is triplicated and stored in data nodes and the complete access control is dealt by a single point and the name nodes can fail. Their model is depicted in figure 2.5 which contains fast recovery at the base level, privacy protection in the middle and the authentication at the top level. Here, user verification using digital signature is done at the authentication layer followed by the deployment of encryption algorithm at the middle layer followed by data recovery of data using rapid recovery algorithms.

![Data Security Model](image)

**Figure 2.5:** Data Security Model

Huiqi *et al*., in [73] have developed a special technique based on Random Space Perturbation and kNN that deals with the main aspects of lower processing cost, efficient query processing, query privacy and data confidentiality. Here, the experiments are conducted under threat model and confirmed on the better results at lower cost. But, the main limitation of this technique is that it suffers from data leakage and lower query privacy.

Hossein *et al*., in [74] have developed a service named as Encryption as a Service for delivery of better secured services to the customers by the service providers. Here, MAC is used for creating the private cloud for better integrity with respect to multi-
threading processes. Every single thread hits uniformly on a parallel region and after encryption it generates a team of threads. The limitation of this approach is, it mandates the programmers to develop applications using multiple threads, and otherwise the performance degrades. The whole process is as in figure 2.6.

![Figure 2.6: Encryption as a Service](image)

Xu et al., in [75] have come out with an agent based trust model that gives better credibility and reliability. As shown in figure 2.7, TAEC architecture gives better security features and works even on tiny sensor nodes using agents. Before transmitting the data from node A to node B, TACE applies encryption and ensures secured communication. In the first step, TAEC Manufacturer will provide the certificate to node A containing TAEC type, public key, and security strategy. Then, the data gets transferred to node B after digital signature verification. Even though, because of agent technology it is a platform independent model, due to digital signatures the efficiency decreases.
Wang et al., in [76] have provided a cloud auditing mechanism known as PANDA with better user revocation. Here, they use user revocation because during the revocation period if the same user approaches the model generates the private key and the data is to be re-signed with the same private key. But, this degrades the efficiency because of additional tasks involved. Authors in [77] addressed the use of virtualization for better cloud security and the same protects the integrity of both IaaS and VM. A security business flow language has been developed for a better intrusion detection system in [78], which provides a prototype for enhanced cloud security. This uses agents which are autonomous in nature and flexible SaaS architecture in their proposed model. It has been used for delivering services via UIDM in [79], as it provides better authenticity and authorization of all the users in network.

In [80], some issues and solutions related to cloud computing are covered, where in the authors have mentioned all the security aspects and also designed a software framework to cure threats and vulnerabilities. Here, all the protection schemes are migrated to third party. Orellana et al., in [56] have developed a better transparent layer that encrypts the data prior to storing in the server, where in the solution is very specific to storing documents on Google server. In this proposal, the user can opt for the security algorithm to be used and after selection the corresponding data is transformed into a cipher text and the same is stored on the Google server. It has been observed that blowfish algorithm uses smaller key size and hence faster in encryption. But, the algorithms used are all symmetric in nature and hence considered as less secured in comparison to asymmetric ones.
In [81], co-residency and network stress are used for detecting the different kinds of attacks and such attacks may harm the kernel of the cloud. To identity these attacks, literature has used Smith Waterman Genetic Algorithm in [82]. Similarly, in [83] authors have attempted to trace the kernel layer of VM, where each VM is traced in parallel. The proposed techniques can detect interaction among the threads of VM on any core. User authentication protocol in [84] is also quite promising towards ensuring security, as it is based on identity based cryptography.

Where in a critical observation at the literature enforces a need of a better model that addresses the open issues with respect to optimal resource utilization, better load balancing and a robust security mechanism in order to ensure integrity, confidentiality, security and user authorization in clouds. It shows that so far limited research has been done in area of cloud security and hence the scope for further research in this dimension. Further, as cloud computing has lately emerged as a scalable delivery platform, the scalability of cloud is reflected through its capability to increase or decrease the number of various server instances assigned to a application dynamically depending on the demand of usage of an consumer application.

2.5 ANEKA TOOL [85]

Aneka is developed by Manjrasoft Pvt. Ltd., a spinoff company of University of Melbourne. Manjrasoft Pvt. Ltd. It is developed using .Net technologies and provides good number of APIs for managing resources via IaaS layer without violating SLAs and thus enables less cost for the user. The salient features of Aneka are as follows:

- Contains Rapid Application Development framework.
- Combines with multiple VMs to provide better and faster solutions.
- QoS and SLAs are guaranteed using provision interface.
- Facilitates multiple programming environments.
- Supports simultaneous execution of multiple applications to increase resource utilization.
- Can work on Linux environment.
• Supports Task Model, Thread Model and MapReduce Model to develop parallel applications.

• It gives many security features inside its container, which is nothing but the cloud middleware.

• It has two levels namely application level and user level that work on the top of the middleware. It supports application development process, managing and monitoring the cloud environment and also provides security features.

• Aneka uses collective desktop services to make it a better parallel computer and hence reduces the total infrastructure cost. It is very useful tool to make private cloud.

Aneka provides unique functionalities covering data protection, data recovery and cloud infrastructure management.

• **Protection of User Data:** Aneka protects the user data from unauthorised access. As the cloud vendors are afraid of data security, confidentiality and theft of identity of their customers, Aneka provides lot of security APIs to prevent such damage.

• **Availability and Recovery of Data:** Typically the cloud applications are designed and developed in accordance with the bi-party agreement through SLAs, which are designed to support high degree of data availability and reliability. The tasks involved in this stage include duplication of data, monitoring of systems, periodical maintainance and failure recovery. Aneka provides many APIs to perform these tasks.

• **Cloud Management:** Aneka provides good number of APIs for cloud infrastructure management which include the issues like load balancing and scalability. Aneka dashboard provides user interface to manage cloud resources.

Currently industries are treating cloud services as an incentive towards expanding their IT related infrastructure with less cost and high ROI. As cloud computing support on demand dynamic provisioning of resources and virtualization, it can be widely used both by SMEs and larger organization to expand their computing infrastructure with very less investment. Since Aneka supports market oriented computing where in it allows trade off between cost and performance, and also flexible choice of cloud vendors depending upon the user requirement via its cloud brokering software, it is widely used as an interface between cloud application developers and cloud service providers. Since, cloud provides services based on the subscription and use of its resources the IT cost of the organization
will be reduced drastically, there is no need to invest heavily on IT capital expenditure and owning the resources on their own, the cost application development comes down because of use of such infrastructure on rental model and there will be a deep strike in their revenue. Whereas, the IT infrastructure constitute the major portion of the total software cost, taking such kind of high performance IT infrastructure on a rental basis increases the ROI of business organization. Also the delivering Software as Service i.e., develop once and give it to many is considered to be a better way of software delivery than the traditional one. However, the Internet is now matured enough to provide software as service via their web service model, and also providing software service at any time, to any device is the want of the day, cloud computing is the fit model to serve this purpose. As clouds are categorized into three types depending on the geography of physical infrastructure even the application development on these types of clouds varies. Aneka being the cloud middleware provides APIs for variety of cloud application development which include IDEs for application development, effective scheduling algorithms, monitoring and managing cloud resources, dynamic provisioning, virtualization and energy efficient resource utilization. Aneka is a .NET based cloud middleware framework that provides variety of runtime environments and the repository of APIs that facilitate the cloud application developers to develop their customized applications. Basically, Aneka uses three programming models, namely

(i) MapReduce Programming Paradigm

(ii) Thread Programming Model, and

(iii) Task Programming Model.

Apart from these, Aneka provides other services that permits the cloud users to auto-scale, control, reserve and generate invoice the resources used. Aneka supports cloud resource provisioning on public clouds such as GoGrid, Amazon EC2 and Windows Azure. It also allows application developers to deploy their applications on private clouds or the in-house clusters, where in the users can exploit the in-house parallel computing infrastructure. It provides the application developers lot of options to model their application which may include multiple authentication models, multiple programming models, distributive environment for operating system and multiple scheduling strategies. The main objective of Aneka is to facilitate open-ended programming abstractions and features for deploying applications in distributed computing scenario.
**Position of Aneka in Cloud:** Aneka is a PaaS framework, which is also known as the cloud middleware that provides Programming APIs for developing high performance distributed computing applications on the Virtual Machines. The participation of Aneka in cloud is shown in figure 2.8.

![Figure 2.8: Position of Aneka in Cloud Computing Layers](image)

The overview of basic functionalities of Aneka is given Figure 2.9.

![Figure 2.9: Overview of Aneka Environment](image)
Aneka architecture is scalable and reliable. Many issues of cloud management like resource management at IaaS level, security etc. are handled by Aneka containers. It has a management studio for dynamic provisioning and workload management. It uses different programming paradigms and cloud models for effective mapping of clients to resources. Aneka application module has its own importance in developing real time application in the areas of health, construction, engineering and many others. Aneka basically relies on the functionalities namely Build, Accelerate and Manage.

- **Build:** Aneka combines APIs and tool studio via its SDK to build applications faster and efficient. It allows application developers to build new run-time applications over virtualized and heterogeneous network infrastructure. The APIs include (i) Task model typically used for legacy and batch applications, (ii) Thread model that use object orientation, and (iii) MapReduce model to solve applications which are data intensive and requiring analytics. Aneka even allows application developers to customize MPI and Actors (Distributive Active Objects/Agents). The tools and APIs supported by Aneka are shown in table 2.2. Run time environments formed using multiple computers can be easily build using Aneka. Different layers of Aneka are shown in figure 2.10. The supported tools include (i) Design Explorer is used for typically the parameter sweep applications and is built over the task model. It does not require any additional programming, and (ii) Work Flow applications are building over the task model, but this requires some programming efforts from the application developers to customize.

<table>
<thead>
<tr>
<th>API’s</th>
<th>Tools</th>
</tr>
</thead>
<tbody>
<tr>
<td>Task Model : batch and legacy applications</td>
<td>Design Explorer</td>
</tr>
<tr>
<td>Thread Model: object oriented applications</td>
<td>Workflow applications</td>
</tr>
<tr>
<td>MapReduce Model</td>
<td>Big Data and Data Analytics</td>
</tr>
</tbody>
</table>
Accelerate: As Aneka uses the bunch of desktop machines to create a high performance cluster and to achieve maximum local utilization, it supports rapid application development and deployment in multiple run-time environments. With the increase in demand, Aneka can make dynamic provisions of VMs through public clouds like Amazon EC2 or private clouds like Xen. The following are the steps to accelerate application development and deployment:

(i) Use Design Explorer Tool with the support of parameter sweep to achieve Rapid Deployment. The applications that are controlled by certain parameters are given as common line argument to parameter sweep to yield multiple parallel executions of the same application with varying parameter sets.

(ii) As .Net framework allows different programming languages, application running becomes faster.

(iii) As Aneka uses the philosophy of develop once and run multiple times parallel, the running time of application becomes faster and also the application becomes fault tolerant since it uses dynamic provisioning of VMs.

(iv) Aneka is optimized for cluster of multi-core machines that allows effective use of virtualization and allows the user to harness the power of multiple computers.
This reduces run-time complexity of the applications. Also, it allows executing many applications to run on the same run-time environment which increases degree of concurrency.

- **Manage:** Application development using Aneka include building GUIs and extending the APIs for monitoring and maintaining the clouds. It provides accounting module to measure and bill the users against their resource usage. It also takes into consideration the user’s priority and scalability issues on the grounds the agreed SLA for deciding on dynamic provisioning. Aneka Management Studio is a Interface that performs (i) easy installation of clouds, (ii) monitoring and tuning cloud resources, (iii) Watching CPU performance by calculating resource utilization, (iv) dynamic run-time resource allocation, (v) Billing, and (vi) SLA management.

Aneka is a middleware that also combines the computing power of multiple computing systems and the virtualization capability over a distributed network to deliver scalable and flexible physical infrastructure for the cloud users. Every individual computing system is saved as resource and stored in Aneka container and this container provides the basic services. The services are as follows:

- **Execution Services** schedule distributed threads among different processors in order to execute applications in cloud.

- **Foundation Services** monitors terms and agreements related to Aneka middleware, which in turn enable the container to accept the membership of the performing task.

- **Fabric Services** deal with physical data centres or clusters connected by a network and provide dynamic provisioning.

Other management features of Aneka include

(i) Pricing and Accounting that provide different pricing strategies to keep track of applications and reservations

(ii) Capacity Management Module to deal with elasticity issue of cloud resources depending on the need of the user
(iii) Service Oriented Architecture that allows flexible discovery of physical resources and services

(iv) Ensuring high availability of physical resources and thus ensuring better reliability to match SLAs.

The detailed view of Aneka architecture is given in figure 2.11.

![ANEKA Architecture](image)

**Figure 2.11: ANEKA Architecture**

**Programming Models in Aneka:** It provides four basic programming models for developing and deploying cloud applications.

- **Task Programming Model:** It is at the task level parallelism, where in the application is divided into task, and these tasks execute independently on each machine. Dividing the application into parallel tasks should be made in exclusive way, such that they do not overlap on issue. Typically it is kind of submit and forget programming model. In this model, Interface and Grid task APIs are
provided for users to execute their operations, but the interface can execute only one operation at a time. The grid task is operated usually on remote hosts. The Aneka middleware handles the entire task scheduling and execution. Aneka application manager builds task based applications by performing the following steps

(i) creating grid application instance
(ii) implementing interface tasks, and
(iii) submission of task.

Some of the typical example application includes convolution and Excel Grid.

- **MapReduce Programming Model** is a transformation method that uses some key values. It consists of two functions, one is *Mapping* function and another is *Reduce* function. In *Mapping* function the input along with the job is distributed to all the processes and in the *Reduce* function final result is consolidated from all the mapping function’s outputs. The typical mapping and reduce functions are given below.

  Mapper: Map<< map::<key1, value1> = list<key2, value2 >>

  Reduce: Reduce<< reduce::<key2, list value2> = list<value3> >>

MapReduce model uses many user APIs that work in distributed applications. The Aneka middleware handles the scheduling of Mappers to Processors and vice versa. It also handles the execution of Mapper and Reducer at their respective physical machines.

- **Parameter Sweeping Model** is slightly similar to task programming model, but the only difference is all the homogeneous task are scheduled on multiple machines with different parameters and the final combined values are checked out to generate a task instance. Actually, this model provides better parallelism and used on legacy applications. In this model, the user API deals with the typical lower level tasks like delete, copy and execute to compose an interface. Design explorer is used to build applications using this model via a special Wizard.
The Thread Programming Model is very similar to Java threads, which are nothing but the basic execution unit of a system. The stages of thread execution include start, stop, state query and join. It provides resources that can be easily deployed on distributed networks. Any multi threaded applications can run using this model. Grid thread is the API used for executing on remote host. Aneka middleware handles thread scheduling and maintains schedule of objects that are created using GridThread class. Aneka application manager handles the thread management at the application level. The steps involved in building applications using thread model are (i) Use Grid thread class to define the worker method, (ii) serialize the class and convert into object and instance, and (iii) Using Grid instances create and submit the classes. The typical example is Mandelbrot.

**SUMMARY**

In this chapter many challenges related to cloud computing is discussed in length. Further, an Aneka cloud model which provides platform as a service is elaborated with all its futures. As Aneka acts like a middleware of the cloud, it positions itself in between IaaS and SaaS. It facilitates cloud application developers to build software using its various programming models like task model, thread model and MapReduce model. It is developed using Microsoft .NET framework and enables users to install on multiple machines in a laboratory environment. It can be hosted on multiple physical machines and forms a distributed system on which parallel applications can be executed. Its basic functions like build, manage and accelerate are provided to users as containers, which can be customized as per their requirement. The various services like fabric services, foundation services and execution services are also provided in Aneka for application developers.