1.8 DYNAMIC CLOUD INFRASTRUCTURE

Through cloud computing, clients can access standardized IT resources to deploy new applications, services, or computing resources rapidly without re-engineering their entire infrastructure, thus making it *dynamic*. Cloud dynamic infrastructure is based on an architecture that combines the following initiatives:

1. **Service management:** Offers business transparency and automation across the pillars of business for consistent delivery.
2. **Asset management:** Maximizes the value of critical business and IT assets over their lifecycle with industry-tailored asset management solutions.
3. **Virtualization and consolidation:** Reduce operating costs, improve responsiveness, and fully utilize the resources.
4. **Information infrastructure:** Helps businesses achieve information compliance, availability, retention, and security objectives.
5. **Energy efficiency:** Offers green and sustainable energy solutions for business.
6. **Security:** Provides end-to-end industry customized governance, risk management, and compliance for businesses.
7. **Elasticity:** Maintains continuous business and IT operations while rapidly adapting and responding to risks and opportunities.

1.9 CLOUD COMPUTING CHARACTERISTICS

Cloud computing uses commodity-based hardware as its base. The hardware can be replaced any time without affecting the cloud. Cloud computing uses a commodity-based software container system. For example, an instance or service can be migrated from one provider to other service provider with zero impact.

Cloud computing also requires a virtualization engine and an abstraction layer for the hardware, software, and configuration of systems. Cloud computing also has the multi-tenant feature where multiple customers share the underlying infrastructure resources without compromising the privacy and security of their data. Clouds implement the 'pay-as-you-go' pattern with no lock-in and no up-front commitment and are elastic as the service delivery infrastructure expands and contracts automatically on the basis of the capacity needed.

1.9.1 *Cloud Computing Hurdles*

There are various hurdles in adopting the cloud for large-scale cloud deployment services. The first and foremost is security. Now because of new paradigm, the data security concern is more hyped for cloud, but in some manner it is same as securing the datacenter services, network, storage in hosted and utility-based solutions. As the services are opened and delivered over the network between the cloud service provider and the consumer, the security in this model is perceived at higher levels. Other inhibitors can be location-independent resource pooling where consumer does not know where his services are running or where his data is stored. It is also believed that multi-tenant models are somewhat less secure than dedicated models. Limited service management and monitoring capabilities in the public cloud model also add to the complexities.

The second hurdle is of regulation and compliances. There is a need of data governance models to be established in the enterprises and federating data privacy. In large organizations, IT delivery is taken with the concerns of reliability, performance, and availability. There are different levels of maturities for organizations seeking different levels of service-level agreements (SLAs) but cloud service providers are not equipped to deliver the services. There is a need of stringent recovery point objective (RPO) and the recovery time objective (RTO) with the agreed number of mins/hours down-time. There is a very less chance to take any corrective actions after any impact and it does not cover the consequences happened due to outage. It is also difficult to cater to different type of SLAs for various tenants by the service providers as it is very difficult to tailor the SLAs based on the customer interest and requirements. Also it is observed that if the service delivery is for a complex application bundle, there is a risk of poor performance. There are few components in the cloud ecosystem that are beyond the control of a consumer or a service provider such as bandwidth.

Cloud migration is the next hurdle in cloud computing. This requires the property of powerful interoperability of platforms that should identify the appropriate application that can be migrated to the cloud. It is important to identify the interdependencies and integration points with standards and interfaces that are lacking today among service providers.

Cloud migration becomes more complex if the service bundles are integrated from multiple cloud service providers. This can also become the deal breaker or the reason for downgraded performance. There can also be the licensing problem in the cloud environment. Also there are issues related to multi-geography-based application platforms deployment and implementations in the cloud-based model and get some hits with respect to the desired levels of service for multiple offerings.

Last hurdle is the workload suitability for cloud. It is also a big question whether the workload is seasonal for the cloud deployment or not. Not all the applications are suitable candidates for the cloud. It depends on the function of the business, enterprise policies, application architecture, scalability, suitability, usage patterns according to pay-per-use-model, or infrastructure requirements in the service model.

1.10 CLOUD ADOPTION

Business functions that suit cloud deployment can be low-priority business applications -- such as analytics, against partner and field service-based functions – and other low-priority business functions. Cloud favors traditional Web applications and interactive applications that comprise two or more data sources and services, and services with low availability requirements and short life spans; for example, enterprise marketing campaigns need quick delivery of a promotion that can just as quickly be switched off. It is also helpful when high-volume, low-cost analytics and disaster recovery scenarios, business continuity, and backup/recovery-based implementation are required. It is like a boon to one-time batch processing with limited security requirements, record retention, media distribution, and mature packaged offerings, such as e-mail, collaboration infrastructure, and collaborative business networks.

3. **Operational efficiency:** Sometimes conventions and configurations followed by the test and operation teams may differ from those followed by the development teams. This difference may cause the application behavior to be different from what was intended as well as result in delayed services. The template-based approach, with its solution stacks of hardware, configurable applications, and operating systems, is more transparent and can help the teams to understand the environment better.
4. **Hosted tools:** In cloud computing, the developers and testers need not to install, configure, run, or maintain tools on their systems as they can log into the network maintaining these tools from any machine.

These four benefits help the developers and testers to concentrate on their core work and retain focus without worrying about other jobs. This increases quality and productivity; and therefore results in more developer innovation, increased test quality and coverage, which are beneficial for an organization.

There are a number of major challenges that developers face today in getting started and rolling out new applications and services faster. Innovative new products and services are the lifeblood of rapidly growing companies and represent a substantial portion of corporate sales and profits. In an environment of heightened competition, the inability to roll out new applications and services quickly means declining market share and lost revenue.

A growing application backlog leaves lines of business and end-users frustrated because they consider IT as a bottleneck and look for ways to work around IT to roll out new products and services more quickly. Testing backlog is often very long and a major factor in the delay of new application deployments.

A major reason why testing takes so long, on an average it takes weeks, is because of the time taken to set up the application environments for testing and QA as well as production. The testing phase becomes longer because of the time required to procure new hardware and software, and then schedule the time with the IT to configure and set up the systems. Configuration and setup are manual processes where errors are easily introduced. An average new application takes six to nine months to be deployed. The whole process application development is affected by a number of factors ranging from poor governance to poor collaboration between business users and development to inflexible infrastructure and tools. Almost 30% of the defects are caused by wrongly configured test environments. This is a result of following manual processes without any automation to replicate testing environment along with challenges organizations face in finding available resources to perform tests in order to move new applications into production. Test environments are seen as expensive and provide little real business value.

2.4 CLOUD DEPLOYMENT MODELS

Let us talk about cloud computing and different types of cloud deployment models and services that can be delivered using these models. Cloud computing is a style of computing in which business processes, application, data, and any type of IT resource can be provided as a service to the users.

Cloud delivery models can be briefly classified into the following three types (Fig. 2.2):

1. **Public:** In a public cloud, a business rents the capability and pays for what is used on-demand.
2. **Private:** In private clouds, a business essentially turns its IT environment into a cloud and uses it to deliver services to their users.
3. **Hybrid:** Hybrid clouds combine elements of public and private clouds.

Private cloud is good when the control and customization are bigger concerns to drive the efficiency. Today public clouds are also process-oriented and easily standardized with lower risk problems. There are many standardized functions that can be moved to public cloud such as collaboration, search, CRM, and sales force management.

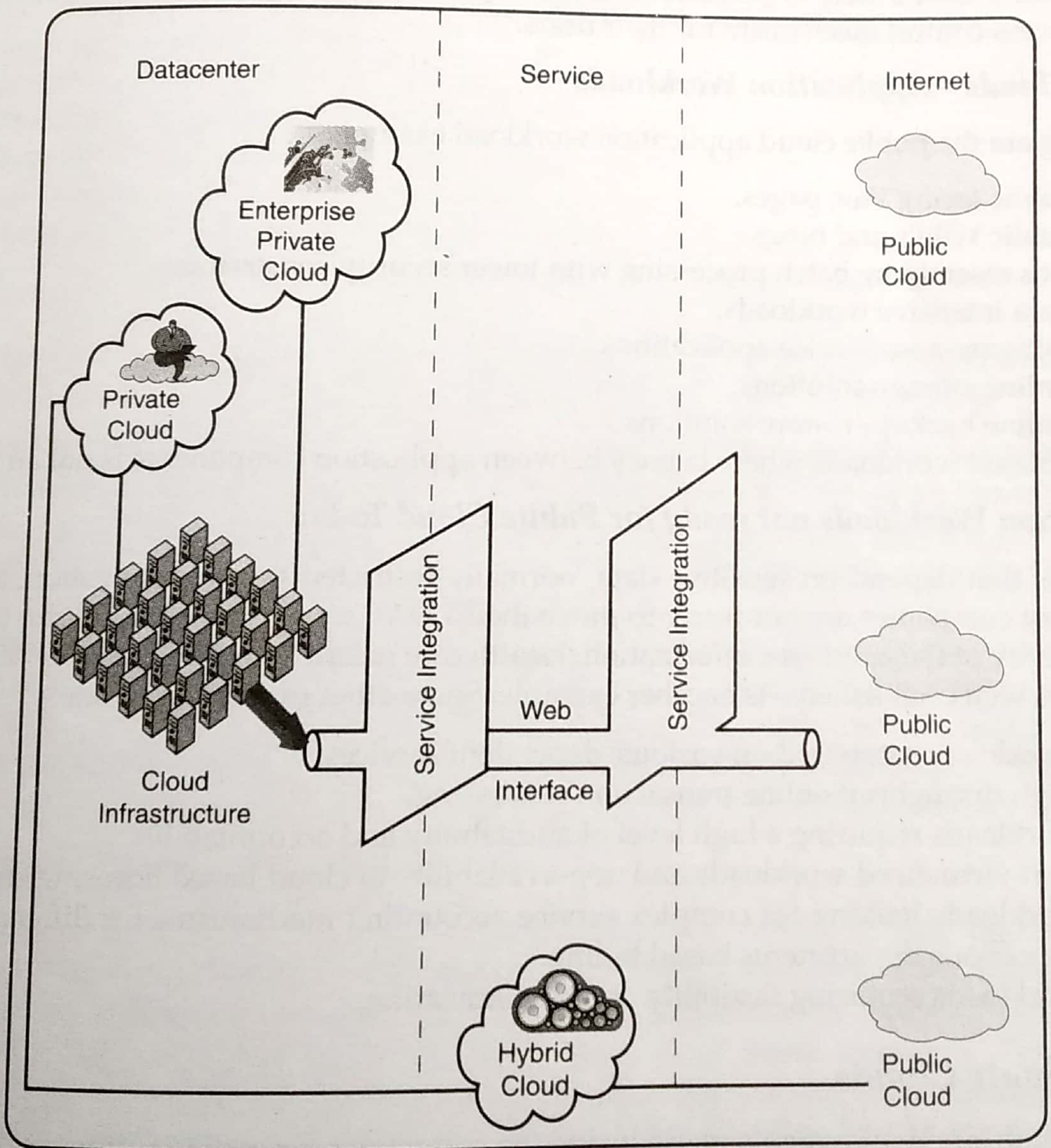

FIGURE 2.2 Private, public, and hybrid clouds.

There is no one-size-fits-all model; in a number of cases, businesses may end up using all these models eventually based on the business model for different services.

2.4.1 *Public Clouds*

Public cloud services are offered by third-party datacenter provider to end-user consumers over the Internet. Public cloud offers resource pooling, self-service, service accounting, elasticity, multi-tenancy to manage the solutions, deployment, and securing the resources and applications. Companies can use it on-demand and with the pay-as-you-use option, it is much like utility consumption. Enterprises are able to offload commodity applications to third-party service providers (hosters).

The term 'public' does not mean

1. That it is free, even though it can be free or fairly inexpensive to use.
2. That a user's data is publically visible – public cloud vendors typically provide an access control mechanism for their users.

Public Clouds – Application Workloads

Following are the public cloud application workload examples:

1. Public facing Web pages.
2. Public Wiki's and blogs.
3. Jobs resembling batch processing with lower security constraints.
4. Data intensive workloads.
5. Software-as-a-Service applications.
6. Online storage solutions.
7. Online backup/restore solutions.
8. Isolated workloads where latency between application components is not an issue.

Application Workloads not ready for Public Cloud Today

Workloads that depend on sensitive data, normally restricted to an organization, are public today. Most companies are not ready to move their LDAP server to a public cloud because of the sensitivity of the employee information. Health care record – until the security of the cloud provider is well established – is another example. Some other examples include:

1. Workloads composed on various, dependent services.
2. High throughput online transaction processing.
3. Workloads requiring a high level of auditability and accountability.
4. Non-virtualized workloads and non-availability of cloud-based licensing strategy.
5. Workloads looking for complex service accounting mechanisms for different services for various departments based billing.
6. Workloads requiring flexibility and customization.

2.4.2 *Private Clouds*

Private clouds are deployments made inside the company's firewall (on-premise datacenters) and traditionally run by on-site servers. Private clouds offer some of the benefits of a public cloud computing environment, such as elastic on-demand capacity, self-service provisioning, and service-based access. Private cloud is suitable when the traditional requirements, such as control, security, and resiliency, are more emphasized by an organization with the restricted and designated user access and authorization.

Services in Private Cloud

This section highlights the services provided by a private cloud and the services consumed from public cloud, specifically:

1. Virtualization.
2. Government and management.
3. Multi-tenancy.
4. Consistent deployment.
5. Chargeback and pricing.
6. Security and access control.

The services consumed from a public cloud are as follows:

1. Security and data privacy.
2. Ease of access.
3. Discovery of services.
4. Restful interface support.
5. Lower cost.
6. Speed and availability.

High 'Cost of Privacy'

Many experts believe that a private cloud implemented with internal hosting/running of the infrastructure makes it difficult to realize many key benefits of clouds, including

1. **Eliminating capital expenses and operating costs:** Ownership of the hardware or software eliminates the pay-per-use potential, as these must be upfront purchases. The full cost of operations must be shouldered as there is no elasticity. If the private cloud hardware is sized for peak loads, there will be inefficient excess capacity. Otherwise, the owner will face complex procurement cycles.
2. **Removing undifferentiated heavy lifting by offloading datacenter operations:** Utility pricing (for lower capital expenses and operating expenses) usually implies an outside vendor offering on-demand services. It relies on the economies of multiple tenants sharing a larger pool of resources. These higher costs might be justified if the benefits of quicker and easier self-service provisioning and service-oriented access are large.

Private Clouds Provide More Control

In traditional security models, location implies ownership which in turn implies control when security is location-specific. Then location, ownership, and control are aligned. Strong requirements for control and security usually drive a preference for a private cloud, where they own the cloud resources and control the location of those resources. For example, government may not want their applications or data to reside outside certain borders. Clouds rely on virtualization; and in the public model, this loose coupling breaks the link between location and application, and reduces the perceived ownership and control.

When we talk about the information control, it is not related to fixed geography or total ownership of the information. One example is public key encryption – the ownership of the key means control over the information without owning the rest of the infrastructure. The information control can be managed over the infrastructure that is trustful on the basis of the contracts, regulations, SLAs, standards, and imposition of the security mechanism on service providers. Compliance is difficult outside of traditional security models. As long as control through technology and contracts can be clearly demonstrated, it is possible to make a public cloud computing environment as compliant and as secure as a privately owned facility. Auditors and regulators are continuously adapting to new technologies and business models. Ownership can have multiple avenues as follows:

1. Full-implementation ownership.
2. Lack of full ownership.
3. Controlled ownership.

There are many possible approaches in between, such as partial control and shared ownership. There are also different levels of limited access – specific departmental access, industry-only access, and controlled partner access.

2.4.3 *Hybrid Clouds*

A hybrid cloud is a combination of an interoperating public and private cloud. This is the model where consumer takes the non-critical application or information and compute requirements to the public cloud while keeping all the critical information and application data in control. The hybrid model is used by both public and private clouds simultaneously. It is an intermediate step in the evolution process, providing businesses an on-ramp from their current IT environment into the cloud.

It offers the best of both cloud worlds – the scale and convenience of a public cloud and the control and reliability of on-premises software and infrastructure – and let them move fluidly between the two on the basis of their needs. This model allows the following:

1. Elasticity, which is the ability to scale capacity up or down within minutes, without owning the capital expense of the hardware or datacenter.
2. Pay-as-you-go pricing.
3. Network isolation and secure connectivity as if all the resources were in a privately owned datacenter.
4. Gradually move to the public cloud configuration, replicate an entire datacenter, or move anywhere in between.

2.4.4 *Community Clouds*

This is the cloud managed by groups of people, communities, and agencies especially government to have the common interests – such as maintaining the compliance, regulation, and security parameters – working on the same mission. The members of the community share access to the data and applications in the cloud.

2.4.5 *Shared Private Cloud*

This is a shared compute capacity with variable usage-based pricing to business units that are based on service offerings, accounts datacenters. It requires an internal profit center to take over or buy infrastructure made available through account consolidations.

2.4.6 *Dedicated Private Cloud*

Dedicated private cloud has IT Service Catalog with dynamic provisioning. It depends on standardized Service-Oriented Architecture (SOA) architectural assets that can be broadly deployed into new and existing accounts and is a lower-cost model.

2.4.7 *Dynamic Private Cloud*

Dynamic private cloud allows client workloads to dynamically migrate to and from the compute cloud as needed. This model can be shared and dedicated. It delivers on the ultimate value of clouds. This is a very low-management model with reliable SLAs and scalability.

2.4.8 *Cloud Models Impact*

Clouds will transform the IT industry and profoundly affect how we live and how businesses operate. Cloud computing

1. Offers the scalable compute model to be accessed from anywhere.
2. Simplifies service delivery.
3. Provides rapid innovation.
4. Provides dynamic platform for next generation datacenters.

Some say it is grid or utility computing or Software-as-a-Service, but it is all of those combined.

Public Clouds: Benefits

There are various benefits of public clouds. Following are some of the benefits of public clouds:

1. No big upfront investments.
2. Offer self-service for rapid-start development.
3. Deliver new pricing models for hardware, software, and service consumption.
4. Flex-up and flex-down capacity in short span.
5. Demonstrate proof of concept (POC), collaboration, workload management.

Internal Private Clouds Drive Cost Savings

There are significant cost savings in implementing an internal private cloud instead of using a usual traditional infrastructure. With a traditional infrastructure, each server typically runs a single application and the hardware is sized to meet peak demands. This setup leads to very low average hardware utilization and high software costs due to the number of servers that are deployed and the lack of resource sharing. The internal private cloud uses virtualization on larger servers and leverages advanced service management capabilities to drive efficiency. The servers can be dynamically provisioned to adjust to workload changes and end-users can request the services they need through self-service portals, which drive automation.

Significant cost savings can be achieved by leveraging these capabilities to automate test and development environments. Automation drives down IT labor cost by automatically responding to the changes in the environment and taking action before the problems occur. Virtualization coupled with service management greatly improves server utilization and reduces software license costs since fewer machines need licenses. Automated provisioning and standardization outsourced in a service-based model. So cloud computing has really taken a hold because of the fact that you take virtualization, standardization, and automation, and drive this automated delivery of services at a reduced cost.

Cloud computing introduces the concept of 'IT-as-a-Service'. To support this service, the cloud infrastructure must perform the following tasks:

1. **Abstraction:** Alleviate IT consumers from the operations of applications and allowing the end-users to focus instead on the execution of high-value activities.
2. **Virtualization:** On-demand access to services irrespective of location and resource availability.
3. **Dynamic workload management:** Workload placement, provisioning, and de-provisioning based on available resources and capacity management.
4. **Data management:** Fast, secure, reliable data access and mobility with integrated data protection and recovery management.

Because all the data reside on the same shared storage systems, effective and efficient data storage management becomes critical in a cloud deployment.

3.2 GAMUT OF CLOUD SOLUTIONS

There are a variety of cloud offerings in the broad spectrum of services (Fig. 3.1). Below are the common types under which most of the services can be viewed:

1. **Platform-as-a-Service (PaaS):** This refers to provisioning the developer's platform who writes custom application using the hardware, operating system (OS), database, and middleware. There are some software policies using which developers write and balance the scalability for application development.
2. **Software-as-a-Service (SaaS):** This refers to provisioning a purpose-built software over the Internet that uses underlying hardware, OS, and distribution capabilities.
3. **Infrastructure-as-a-Service (IaaS):** This is the provisioning of hardware or virtual computers where the organization has control over the OS, thereby allowing the execution of arbitrary software.
4. **Storage-as-a-Service (StaaS):** This is the provisioning of database-like services, billed on utility computing basis, for example, per gigabyte per month.
5. **Desktop-as-a-Service (DaaS):** This is the provisioning of desktop environment either within a browser or as a terminal server.

The distinction between these five categories of cloud offering is not necessarily clear cut. In particular, the transition from Infrastructure-as-a-Service to Platform-as-a-Service is a very gradual one.

3.2.1 *Platform-as-a-Service (PaaS)*

Instead of just offering applications over the Web in the form of SaaS, public cloud players are actually offering an entire PaaS. They provide the foundation to build highly scalable and robust Web-based applications in the same way the traditional operating systems such as

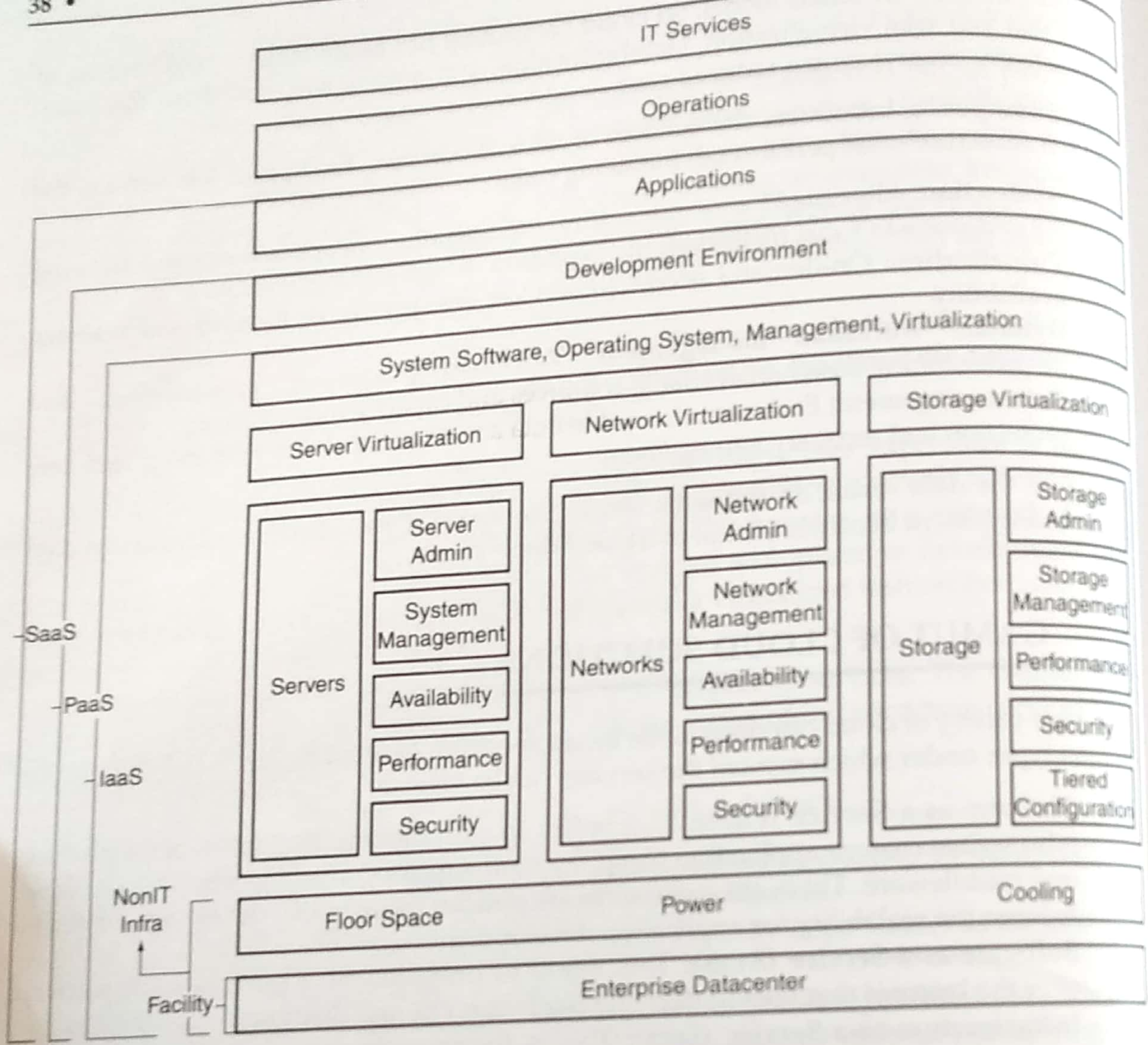

FIGURE 3.1 Cloud@datacenter.

Windows and Linux have done in the past for software developers. What is very different about this model is that the platform itself is no longer 'sold' to the customer who is then responsible for running and maintaining it. This model becomes more attractive with its operation on the hosting platform capabilities and with more capabilities in self-service model and billing. So PaaS has far-reaching implications on the business models of both PaaS vendors and their customers. One can use private clouds to speed up application deployments within minutes. It will help in tracking the usage for the chargeback and will give the option of cost-effective and secure appliance.

In order to optimize deployments, many organizations are looking to extend SOA to cloud services.

Cloud capabilities can improve the productivity of development and testing teams to roll out new applications and SOA services faster, and reduce the application backlog. The cloud provides a catalog of virtual images and patterns all ready for immediate use and these patterns define a cluster of servers working together.

PaaS saves costs by reducing upfront software licensing and infrastructure costs as well as ongoing operational costs for development, testing, and hosting environments.

PaaS enhances the capability of development platforms and helps to remove the problem of integrating the development software and tools such as OS, database, middleware, security mechanism, frameworks, and virtually shared infrastructure. PaaS also reduces the cycle time of delivery and development iterations by providing development, testing, QA, and staging on the same platform. It brings collaboration between the developers and operation team. It actually helps to achieve bigger goals of project management and remove the gaps between the development and operation communities.

Currently, PaaS is in the maturity stage. There is a requirement to bundle the application services with the platform to make it more portable and create a standard platform across vendors to choose the services. The collaboration and component models should be adopted by all vendors and should support the legacy application on the same platform for the same cloud-based services.

There are different set of users for SaaS, PaaS, and IaaS. SaaS looks for the users who want to simplify the provisioning process of business services using hosted applications. PaaS is meant for the development community to support rapid development, testing, and delivery. IaaS is a managed services platform where infrastructure is provisioned to provide compute, storage, and network services.

3.2.2 *Software-as-a-Service (SaaS)*

SaaS reduces the investment by removing development effort, software licenses configuration, installation maintenance, infrastructure cost, delivery and release, and reducing operational cost and administration.

The SaaS platform takes lesser time to build the service compared to the traditional way of doing it. It becomes easier as all the tedious work of management, software maintenance, and support is passed to a vendor. This gives time to workers to focus more on the business' value-based features.

SaaS application does not need extensive customization and should not look for complex business-process applications. So, it is not well-suited for service-based deployment.

There are some issues in SaaS related to migration also. While moving from one on-premise network infrastructure to publically available network infrastructure, SaaS requires an upgrade in the infrastructure. It also requires upgradation of software versions and this can become a big problem for the vendors offering these services.

3.2.3 *Infrastructure-as-a-Service (IaaS)*

IaaS eliminates the cost by handling the provisioning process with the automation during the peak loads. So IaaS helps the server to flex-up and flex-down to meet the demands by

The cloud application services should be mapped to the business processes that they consist of. These business processes could be at different levels. Drill down the processes to the levels that make business sense.

These processes – along with the business strategy, the modular business alignment model, and the process scenarios to be supported – form the input to Service-Oriented Modeling and Architecture (SOMA) methodology. SOMA is applied with a meet-in-the-middle approach. Business processes and the business strategy along with business goals stated are used to arrive at the service portfolio using SOMA. While deriving the services, a bottom-up approach is also followed, taking into consideration the existing assets, which consist of existing application components and services available on public clouds and internal to the cloud, as well as industry cloud component maps.

3.5.1 *Architecture Overview*

The purpose of the architecture overview artefact in a cloud-based implementation is to communicate to the sponsor and external stakeholders a conceptual understanding of the architectural goals of the cloud implementation. It offers a layered conceptual model of the application services to be cloud-enabled and provides a high-level vision of the cloud architecture and its scope to developers. It is easy to explore and evaluate alternative architectural options for the cloud implementation. This stage enables early recognition and validation of the implications of the cloud-based architectural approach and facilitates effective communication between different communities of stakeholders and developers.

This developed document will help to understand the conceptual points of the cloud implementation and its dependencies on internal or external cloud ecosystem. The document may include both the cloud and non-cloud infrastructures such as offerings, components, relationships, external systems, and elements for support systems.

This artefact can also provide key stakeholders with the first high-level view of the cloud architecture landscape of their transformed cloud-based implementation.

Typically, the artefact is produced after multiple iterations and as the project moves through the solution definition and design phases, the conceptual models get clearer and this document is kept up-to-date and governed to always have the current conceptual elements.

An architecture summary depiction represents the governing ideas and candidate building blocks of a cloud-based offering and enterprise architecture. This depiction also provides an overview of the main conceptual elements and relationships in the architecture. The main purpose of such a depiction is to communicate a simple, brief, clear, and understandable overview of the target IT system.

At the enterprise-level, the architecture summary depiction is often produced as a part of an overall IT strategy to move to a cloud-based offering model. At this instance, it is used to describe the vision of the business and IT capabilities required by an organization that needs the offering to be hosted on the cloud. The architecture summary depiction provides an overview of the main offerings and the relationships with other offerings, external systems, components, nodes, connections, users, and systems, and a definition of the key characteristics and requirements.

At the system- and component-level, the architecture summary depiction is developed very early in the project (possibly pre-proposal) and influences the initial cloud hosting vision with the component model and operational model. It is intended that design commitments should be based on this conceptual overview, while the (more detailed) component model and operational model are developed and validated. Subsequently, the component model and the operational model are the primary models used for implementation activities, while their compliance with the architecture summary depiction is maintained continually. Changes to the architecture summary depiction are made using a governance process.

For most infrastructure cloud engagements, the project scope is something less than the client's enterprise architecture. Therefore, the depiction will represent the governing ideas and candidate building blocks of the offerings that are to be hosted on the cloud, and the focus of the engagement. There can be, then, multiple views of the cloud-based IT environment represented by architecture summary depiction – the current and future view, views by waves of transformation, or views by types of clouds – for each alternative solution.

3.6 CONCEPTUAL CLOUD MODEL

So far, we have discussed different service paradigms for cloud deployment. Now let us discuss the conceptual cloud model based on cloud services.

This model gives the structure of offerings in the cloud on the basis of the components of hardware and software with their relevant roles, interdependencies, and interfaces to provide the functionality. The cloud component model for implementation planning is specified at the conceptual level.

The highest level of the conceptual model is the set of cloud-based offerings that make up the cloud-based business solution. These highest-level conceptual components are derived on the basis of business intent and business functionality. These conceptual components, which form offerings, align tightly to the business intent that initiated the creation of cloud-based offerings.

Conceptual offerings that provide the cloud-based implementation of the solution can be further broken down and depicted in a layered composition. The next level of conceptual component model broken down below the offering consists of the following elements:

1. High-level service components that form the services provided by the offerings.
2. Resources that support cloud services.
3. Technical components that provide technical underpinnings of the cloud service and support non-functional needs of the cloud application.
4. External and internal services that are leveraged by cloud application services.

The high-level conceptual component model is the main artefact that provides an abstract view of the design of a cloud application to business stakeholders. This abstract view describes how the business needs will be met by the cloud application components without delving into their technical details. Components identified can be decomposed into further layered conceptual component structures to convey further details of the respective components (Fig. 3.3).

is termed as a service product with the deployment capability in various environments and comes with the benefit of choosing suitable services that fit best in the business function.

This means that an organization is not tied to one solution for cloud computing, but can rather use multiple solutions. So, depending on business needs, the application, and what the application has to offer, cloud requirements help in selecting proper cloud environments.

Cloud-based environments come handy especially when used to develop, test, and run your application for the following reasons:

1. Easy availability in your own private cloud environment or on the public cloud.
2. Rapid access to a configurable development and test environment to speed time to market.
3. A self-service portal for multi-tenancy, provisioning, and account management.
4. Pay-as-you-go pricing, with capacity and workload management.
5. Security with compliance and regulations to protect the environment.
6. A rich service catalog with images.
7. Faster provisioning to reduce the cycle time.

4.3 CLOUD BUSINESS PROCESS MANAGEMENT

Business process management (BPM) looks after business processes across the organization which may be cross-functional and directly or indirectly customer focused. BPM is implemented by the strategy based on the core business objectives and deployment of resources across the organization to create customer-centric values. It integrates all vertical-based core functions such as supply chain and product development in an integrated way within multiple line of businesses (LOBs). In this way, BPM is distinguished from other traditional approaches.

BPM is based on the principle of continuous improvement to increase value for market sustenance of the organization. It defines and aligns operations, IT, and organization. Therefore, the cloud can help in the following ways:

1. **Integration of core business:**
   - Is holistic.
   - Crosses organizational functions and boundaries (height and breadth).
   - Includes business and technology.
2. **Value-focused efficiency:**
   - Customer-centric perspective.
   - Bottom-line success.
   - Return on investment (ROI) delivery speed.
   - Performance measurement.
3. **Continuous:**
   - Based on longer period of intervals pertaining to cloud business.
   - Continual improvement.
4. **Cultural:**
   - Cultural considerations of the organization
   - Geographical and cultural factors at the time of due diligence of requirement.

4.3.1 *Identifying BPM Opportunities*

This section discusses the opportunities required for successful cloud BPM and the characteristics of cloud deployment offerings. The answers to the following questions can help you identify cloud opportunities better:

1. Are the strategic value proposition and capabilities defined for your organization?
2. How does your overall strategy drive the design and execution of your business processes? Is there a traceability of execution to goals?
3. How do you manage your core business processes?
4. What are your current process initiatives?
5. What are your current process governance facilities?
6. Are your existing organizational structures aligned to enable efficient process operations?
7. How do your customers measure and assess the performance of your processes?
8. How does your process performance compare with that of your competitors?
9. How effectively does your current technology (information, systems, tools, machines, etc.) enable the enterprise's core business processes?
10. What risks and challenges does your current technology present for current and future process capabilities?
11. What type of products does your organization have?
12. What are the notable pieces of your IT portfolio?
13. Has your organization adopted service-oriented architecture (SOA)?
14. How are the processes currently modelled in your organization? What is included in this model?
15. What design/development tools are currently used in the organization?
16. What testing tools are currently used by the organization? What are the strengths/weaknesses of these tools?
17. Please describe the different business processes that are automated in the organization.

Cloud application development offerings provide

1. Cloud application reference architecture.
2. Unmatched experience of developing high-performing, secure applications across a wide range of technologies of the cloud vendor.
3. Unmatched application security expertise.
4. Leadership in cloud-related technologies such as multi-tenancy, virtualization, and pervasive computing.
5. Significant expertise with cloud business models.
6. Ability to integrate a portfolio of related cloud services.

4.3.2 *Cloud Technical Strategy*

This section describes the technical strategy on how cloud customers can enable cloud deployment.

Cloud services enable our cloud users to build middleware clouds in their datacenter and utilize public clouds, where it makes sense to provide the following cloud-enabled middleware services:

1. Infrastructure services.
2. Platform services.
3. Application services.

Cloud strategy enables our organizations to do the following:

1. Build middleware clouds in their datacenter.
2. Utilize public clouds where it makes sense.

4.3.3 *Cloud Strategy Support Areas*

The cloud strategy does so by providing support in the following areas:

1. **Cloud-enabled middleware services:**
   - Infrastructure services.
   - Platform services.
   - Application services.
   - Serving on-premise and public clouds.

So, what does it mean to develop an application service for the cloud? For one, it means product features development similar to on-premise software. This means:

1. **Enabling the software for cloud essentially implies**
   - Support for collaborative multi-tenancy.
   - Self-service registration.
   - Customers and their entitlements management.
   - Single sign-on.
   - Additional security concerns.
2. **Integration with the datacenter:**
   - Firewalls and reverse proxy configurations.
   - Certificate and domain names.
   - Management of services and patch procedures.
   - Isolation, recovery, and backup issues.

4.3.4 *Cloud Use Cases*

IaaS or Test/Development

Problem: Development teams require unpredictable amount of infrastructural facilities to get their job done. In the majority of cases, getting all of these resources in place before they are actually required in the development cycle can be quite a challenge. Purchasing the hardware consumes project budget, and procurement of it is often quite slow. Static development and testing resources require manual re-provisioning to re-purpose resources for use, or new resources need to be purchased to meet demand. In cases where project timelines are short, project managers often choose not to set up much of an environment because it depletes their budget or jeopardizes the project's delivery schedule. Actual usage of the system(s) can be quite short in terms of absolute time. Different types of projects require different kinds of development components (such as SQL server, SharePoint, and BizTalk) depending on the architecture of the solution. Besides the testing environment, the team usually needs a system that looks quite similar to the production environment. This system helps to perform various simulations for load, stress testing for various deliveries and training.

Solution: Companies can create standardized service catalog items for common infrastructure requirements and enable development and test teams to access infrastructure in a self-service model (IaaS). Overall control over the process is maintained with business policies around quotas, reservations, reclamation, and standardized offerings.

Standardized Development Platforms/Middleware (PaaS Enabled)

Problem: Developers deliver their code without involving operations into architectural decisions or code reviews. Enterprise architects recognize the large costs associated with non-standard development platforms. To simplify the ongoing maintenance and streamline development operations, many companies are creating corporate standards around development stacks that include middleware/applications. However, most companies are hesitant to an external PaaS offering due to the constraints of having to rewrite their internal applications to fit the external PaaS API sets.

Solution: Companies can create standardized development platform definitions for use by development teams to standardize and streamline their efforts. This improves corporate IT productivity by helping them build a 'private cloud' that provides a common foundation for building custom applications running securely behind their own firewall.

Application Cloud

Problem: Companies want to move beyond self-service for infrastructure and provide application owners the ability to define, instantiate, and manage complex multi-tier applications. This includes configuring the application for production usage and monitoring performance within the application for Service-Level Agreement (SLA) optimization.

Solution: End-users can access complete application definitions and manage them according to their quotas and preferences defined by cloud administrators. Applications in production can be monitored across multiple factors and automatically scaled up or down according to business policies.

Software-as-a-Service (SaaS) to End Customers

Problem: Many companies or independent software vendor (ISVs) want to deliver their applications to end-users as a service. However, creating a multi-tenant SaaS offering requires substantial development to support security, performance, and scalability needs. Due to these high costs, companies cannot offer new services based upon existing applications.

Solution: Companies can provision a unique application instance per customer, with private cloud automation capabilities. Environments are provisioned according to business policies, and unique SLAs can be delivered per customer according to the business arrangements.

The cloud engine should provide the automation to provision on-demand complex application and configuration environments required along with the dynamic application scaling. The cloud engine should also deliver unique business policy selections, including custom placement and high availability across multiple datacenters.

Cloud orchestration and provisioning engine–integrated service management is offered with network, servers, storage, services, and financing as an integrated offering for client test platforms. It

1. Delivers consistent and rapid services from a preloaded system.
2. Helps to improve the value of the business system by effective freedom of choice and faster services.
3. Reduces capital expenses and optimally utilizes the resources.
4. Based on automation and standardized platform, minimizes the chance of human error.
5. Has open architecture–based workload management.
6. Is a complete platform with infrastructure, service management, middleware and distributed system.
7. Is scalable and modular.
8. Offers virtualization for compute and storage.
9. Offers self-service–based delivery.
10. Works as 'lights-out' or 'without human intervention' operations.

Cloud orchestration and provisioning engine is offered as a service engagement which can build a solution to the client's needs, including creation of custom virtual images for dispensing. Summary points are given below:

1. Reduction in configuration and installation time.
2. Creation of new environments in minutes.
3. **Reduction in risk by codifying infrastructure:**
   - Freeze-dry best practices for repeated, consistent deployments.
   - Life cycle–based security.
4. **Simplify maintenance and management:**
   - Flexibly manage and update the components of your patterns.
   - Ensure consistency in versions across development, testing, and production.
5. Spend less time in administrating and more time in developing new solutions.

4.6 COMPUTING ON DEMAND

On-demand computing (CoD) is the need of the hour. It is very essential even in the supercomputing environment. On-demand computing can be implemented using various virtualization techniques. The cloud gives an option to leverage the computing infrastructure without actually buying the hardware. This helps to transfer the workload, if the resources are not able to support it; and at other times, others can utilize the resources that are lying idle. In this way, it is possible to use the resources in the most efficient manner. There may be many spikes in the utilization, but cloud experts can make it smoother.

The uniquely rich set of features offered by on-demand computing enables the service seeker to deploy a true utility. The platform allows users to:

1. Align cost with utilization so that users can scale costs down as well as up. This allows a workload to start with minimal upfront costs and scale the cost as the demand grows.

without paying a penalty to increase capacity. Additionally, users can benefit from not incurring the disruption to move to a larger machine.

2. Increase end-users availability significantly. As the workload can be moved dynamically, it is possible to move the workload from one server to another without any interruption, so the remedial work can be carried out if server downtime is required.
3. Balance workload dynamically across multiple servers without taking applications offline. Using workload mobility features, customers can align their costs by ensuring that the workload is deployed in such a way so as to optimize the system's resources.
4. React to short-term resource requirements almost instantly. If a workload has to be deployed at a short notice, a VM can be created on the server and resources can be allocated instantaneously using the dynamic capacity model.
5. Reduce the physical footprint in the datacenter. Consolidation of the workload on to a smaller number of servers will improve space, power, and cooling metrics.
6. Confidently increase system utilization to over 75%, without the fear of degrading performance for end-users.
7. Develop a simple charging model that reflects usage for end-users as the service delivery culture continues to mature.
8. Double the workload delivered in the power and cooling envelope.

4.6.1 *Pre-Provisioning*

For the on-demand computing requirement, pre-provisioning is a viable option as it helps the organization in meeting dynamic datacenter requirements. Organizations would like to reduce the time they take to commission servers when new workload is to be deployed. This approach is ideal when

1. The sizing and capacity planning is fully understood.
2. The workload is fairly constant, ensuring good utilization levels are being achieved.
3. There are business reasons that require physical separation of the workload.
4. The workload can be scaled horizontally.

4.6.2 *On-Demand CPU/Memory/VM Resources*

In the dynamic environment, it is important to track the requirements of CPU, memory, and VMs. It is based on the common pool concept, where resources are allocated and de-allocated as the requirement is over. This approach is ideal when

1. Workloads are trending upward so that investment can be aligned with utilization.
2. Peaks in workload are long term.
3. The workload scales vertically.
4. There is a chance to 'buy out' dynamic capacity, as it proves advantageous then.

4.6.3 *Dynamic Capacity*

Dynamic capacity is based on utility where the additional processing power for a short period of time can be shared from the resource pool. The service is measured via a service accounting tool based on the utilization and finally the billing is done on the basis of the usage. This approach is ideal when

1. The workload is very variable and multiple workloads can be hosted on to a single machine, thus leveling out the utilization.
2. The workload has short periods where system utilization increases massively, but for the majority of the time, it is not resource intensive.
3. Workloads can share a physical platform.
4. The workload is designed to scale vertically only.
5. Users want to dynamically balance workloads across servers.
6. Users want to continue to run very small workloads without incurring the overheads associated with running a physical server to support it.

Benefits

1. Partition mobility, significantly reduced power/cooling footprint, donation of unused processor cycles of VMs with dedicated processors to uncapped partitions, and, at the same time, guaranteed performance of these VMs.
2. Very short deployment time (time-to-market optimized).
3. Lowest possible cost for the deployment of small workloads.
4. Less management effort, for example, when using VMs.
5. Most granular charging scheme, for example, pay for the CPU and memory cycles when actually used.
6. Complete decommissioning of partitions, with resources being available for other purposes.
7. Flexible workload management where workloads can compensate for each other, thus reducing overall utilization.
8. Ideal for environments with identical systems management, utilities for development and testing.

Limitations

1. Dynamic capacity shows short peaks, must not exceed certain limits, and needs to be monitored (via Web interface) to ensure that the best value is obtained.
2. Utility CoD provides processor resources only to the uncapped partitions.

However, one of the most important advantages of the dynamic capacity model is the flexibility, that is, to allow the switching of CPU capacity on and off as needed, which can reduce costs significantly. This is not reflected in the calculation as it requires input from application owners.

4.6.4 Cloud Platform Characteristics Based on CoD

This section discusses cloud platform characteristics on the basis of low-end, on-demand, and dynamic-capacity–based servers.

Low-End Servers

1. Physical segregation of servers.
2. High administration cost due to the management of more physical servers.
3. Limited and complex scalability – maximum eight processors per server, slower turn-around time for server deployment.

4. Longer lead time for server deployment from ordering of servers to setting up of infrastructure.
5. Non-ideal for short product life cycle application due to fixed-cost expenditure for the hardware.
6. Wastage of hardware resources for applications that react to a volatile market.
7. Inability to share resources between applications.
8. Wastage of unused processing cycles if the application does not fully utilize the resources.
9. No hardware/application interdependency forcing downtime on application owners.
10. No capacity on demand capability. Downtime is required for adding new hardware.
11. Low price per CPU cycle purchased, but higher cost per CPU cycle actually used.

On-Demand Platform

1. Physical or logical segregation of servers or partition implementation.
2. Lower administration cost due to less physical server management.
3. Possibility to lead to quick turnaround time for new application deployment or increase the capacity according to business requirements.
4. Enhanced time to market for new product launch, with immediate availability of CPU/memory capacity.
5. Non-ideal for short product life cycle application due to fixed-cost expenditure for the hardware.
6. Wastage of hardware resources for applications which react to a volatile market.
7. Ability to share I/O, CPU, and memory resources between applications.
8. Ability to take advantage of unused CPU/memory if dynamic VM reallocation or share pool methodology is implemented.
9. For provisioning of a platform of independent hardware and application forcing downtime on application owners, careful capacity planning and management is required.
10. Capacity on demand capability – no downtime is required if COD CPU/memory is sufficient.
11. Higher price per CPU cycle, but lower cost per CPU cycle actually used.

Dynamic Capacity Platform

1. Choose VM or workload VM implementation for application consolidation.
2. Lower administration cost due to less physical and logical server management.
3. Can cater to quick turnaround time for new application deployment.
4. Enhanced time-to-market for new product launch, with immediate availability of infrastructure and setup.
5. Ability to scale up and down, which is ideal for applications with short product life cycle.
6. Ability to cater to applications which react to a volatile market, that is, scaling-up and -down capacity.
7. Ability to share I/O, CPU, and memory resources between applications.
8. Ability to take advantage of unused processing cycle of other applications.
9. No hardware/application interdependency forcing downtime on application owners as workload can be dynamically moved to facilitate maintenance, etc.

10. Capacity on-demand capability. No downtime is required as the machine is fully configured.
11. Higher price per CPU cycle, but lower cost per CPU cycle actually used. Average price due to the ability to optimize utilization and rapidly deploy workload.

4.7 CLOUD SOURCING

Today, we are living in an era of optimizing hardware resources and moving toward larger enterprises day by day; so, cloud computing is becoming the major ingredient of infrastructure deployments. Now you may not just need cloud computing, you also need the entire consulting, implementation, and management solutions.

The new wave that is igniting cloud deployments in the service industry is cloud sourcing - outsourcing an end-to-end solution using cloud methodology through public cloud, infrastructure, and platforms. This will be a more planned approach as it will comprise the whole service cycle of outsourcing business with cloud principles through the help of strategized connected cloud platforms matching the overall enterprise requirements.

This includes the whole cloud implementation, IT business consulting, integration, and configuration of the business. This will give the option of enjoying the benefits of service industry with the benefits of the cloud, thus providing an innovative approach of paying resources over subscription.

Real deployment of cloud sourcing will require a business model with the impact of cloud customer and cloud vendor requirements.

With respect to cloud customers, it is important to note that there is no control on the infrastructure layout of cloud deployments or over the place from where data services are offered by the cloud vendor. It is also known that cloud customers do not have to think about the operational staff for the deployments.

Thus, cloud sourcing will play a vital role in the next generation of cloud implementation. With the availability of new open-source tools, it is like icing on the cake, integrated with partner cloud solutions, platforms, and infrastructure. Also, the new charging models, such as services on both project and subscription basis, will generate a new wave to deploy and adopt cloud sourcing models. This will help to customize applications on cloud infrastructure. This will be primarily offered as a public cloud, and all these offerings will be available as managed services. These services will be prototype-based, that is, developed internally on the product and working applications. Therefore, they will provide a good opportunity to use intellectual property for developing different business vertical solutions easily.

4.8 SUMMARY

In this chapter, we have pointed out the main features of cloud orchestration and provisioning engine, BPM clouds, cloud sourcing, and requirements of service management. The next chapter will discuss different types of cloud offerings.

5.1 INTRODUCTION

The growth of information is at an exponential stage. It is growing day-by-day and is being generated by millions and millions of people with billions of smart devices, M2M, sensor networks, machines, and instruments. Today, there is a massive growth in unstructured data where the images, music, scans, and medical data are flowing more than the speed of imagination.

Until now, organizations could not fully or quickly synthesize and interpret all the information out there – they had to make decisions largely on the basis of instinct. But now, there are mechanisms that can capture, organize, and process the entire data scattered throughout an organization, and turn it into actual intelligence. These mechanisms enable organizations to make better business decisions.

5.2 INFORMATION STORAGE, RETRIEVAL, ARCHIVE, AND PROTECTION

There are various stages of an information management cycle. The information gets processed, managed, migrated, protected, retrieved, staged, and archived for business data. Organizations manage the data on the basis of characteristics of the data such as age, usage, compliance, policies, regulations, protection, disaster, and availability. It is done by the established practice of Information Lifecycle Management (ILM). ILM becomes even more important for cloud deployments as the process requires sharing data services between the cloud vendor and the subscribers.

ILM is not a latest data storage, retrieval, and protection solution; piece of hardware; or some software, but rather an approach to assess and manage the information across the enterprise. ILM is based on how data is used and how readily it must be available to the people who use it. It is focused on managing and storing data according to its value to business operations at any given point in time. It is concerned with the placement of data at the appropriate level of storage with an appropriate retention and retrieval policy. In response to these challenges, organizations are defining the following objectives to support and improve their information management:

1. **Cost reduction:**
   - Controlling demand for storage.
   - Reducing hardware/software costs.
   - Reducing storage personnel costs.
2. **Better system performance and personnel productivity (i.e., improved efficiency):**
   - Doing the 'right' storage activities.
   - Improvement in people, community, processes, and technologies to deliver storage services.
   - Defining and enforcing policies to manage the lifecycle of data.
3. **Increased effectiveness:** Defining and implementing an appropriate storage strategies to address current and future business requirements.

They are also coming up with new ways to generate, enhance, and sustain higher savings. These include:

1. **Activities for gaining initial savings:**
   - Reduce the amount of used storage as a result of initial clean-up.
   - Validate storage area network (SAN) requirements and reclaim used switches and switch ports.
   - Validate data replication requirements in order to reclaim used storage space and offset future growth requirements.
   - Develop and document information classification.
   - Develop and document classes of service.
   - Design and implement the tiered storage architecture.
   - Migrate existing information to lower-cost storage using a tiered storage architecture.
2. **Activities for maximizing savings:**
   - Reconfigure the current storage environment effectively to improve the available raw utilization.
   - Reclaim available storage that has been over allocated.
   - Enhance the information classification, classes of service, and tiered storage architecture.
3. **Activities for sustaining savings:**
   - Develop a storage model based on policies.
   - Implement changes to existing storage management processes, such as capacity planning, and provisioning to effectively improve capacity utilization on an ongoing basis.

While designing a target storage environment, the estimated financial impact is calculated on the basis of the following key cost components:

1. **Operating cost categories:**
   - ***Personnel:*** Storage support and contractors.
   - ***Facilities:*** Current floor space consumed by the storage. Telecommunication charges attributed to storage and tape vaulting services.
   - ***Storage hardware maintenance:*** Existing maintenance and incremental maintenance resulting from growth.
   - ***Storage software maintenance:*** Existing software maintenance and incremental requirements resulting due to innovation and development.
   - ***Outages:*** Cost avoidance associated with the reduction in unplanned outages.
2. **Investment cost categories:**
   - ***New hardware required:*** Typically includes disk, tape, and array cost but not the incremental cost of adding SAN fabric. Investment is either upfront or over a period of time if the client leases equipment.
   - ***New software required:*** New storage software required to support the target environment.
   - ***Hardware refresh:*** Investment required to refresh the existing hardware is often considered in the base case.

- *Transition services:* Incremental cost required to transform the current environment to the future environment. Not typically estimated until the scope of the third-party implementation services has been defined.

When more than 90% of the data stored on hard disks is not actively accessed by the users or applications, it is good to go for more intelligent management and migration to a less expensive storage. But the savings can go significantly beyond disk acquisition costs and annual hardware maintenance costs. Some points in information management are

1. **Data:** Discrete element, reasoning, discussion, or calculation of content created through the interactions between applications or between computing devices.
2. **Information:** Organized and structured collection of data.
3. **Information lifecycle management:** It is a combination of policies, processes, tools, and management practices to align the value of the information with the infrastructure needs, right from the time when the information is created till its disposition.
4. **Information taxonomy:** Data described in the context of business process requirements and lifecycle characteristics.
5. **Information classes:** Groups of information taxonomies with associated business value that provide the basis for storage management and service delivery.
6. **Value-driven data placement:** An event correlation framework that 'senses' the value of data changes and based on business policies 'responds' by moving data to the appropriate storage tier.
7. **Storage process:** A documented set of storage-related tasks and activities required to support a storage infrastructure.
8. **Storage service:** Capabilities provided to a customer base that is designed to meet their business requirements, wants, and needs, and enabled by a storage infrastructure.
9. **Enterprise class of service (COS):** A common set of storage services that are delivered to meet a corresponding set of storage requirements based upon key information management characteristics, and the features, functions, capabilities, processes, and governance required to deliver the required enterprise storage services.
10. **Storage tier:** A subset as a set of storage devices that are identified to store and/or maintain information for a predefined period of time based on key information management characteristics, such as performance, configuration, residency, retention, and value.
11. **Tiered storage infrastructure:** An organized collection of storage tiers reflecting the flow of all the information managed in the enterprise storage architecture.
12. **Utility-based service delivery:** The 'just-in-time' delivery of standardized storage processes, management, and infrastructure as a measurable service on a 'pay-as-you-go' basis.

5.3 CLOUD ANALYTICS

Cloud analytics is a new offering in the era of cloud computing. Cloud analytics will help in the consulting domain and will ensure better results. It provides users with a better forecasting technique to analyze and optimize the service lines, and provide a higher level of accuracy. Cloud analytics can help the users apply analytics principles and best practices to analyze different business consequences and achieve newer levels of optimization (Fig. 5.1).

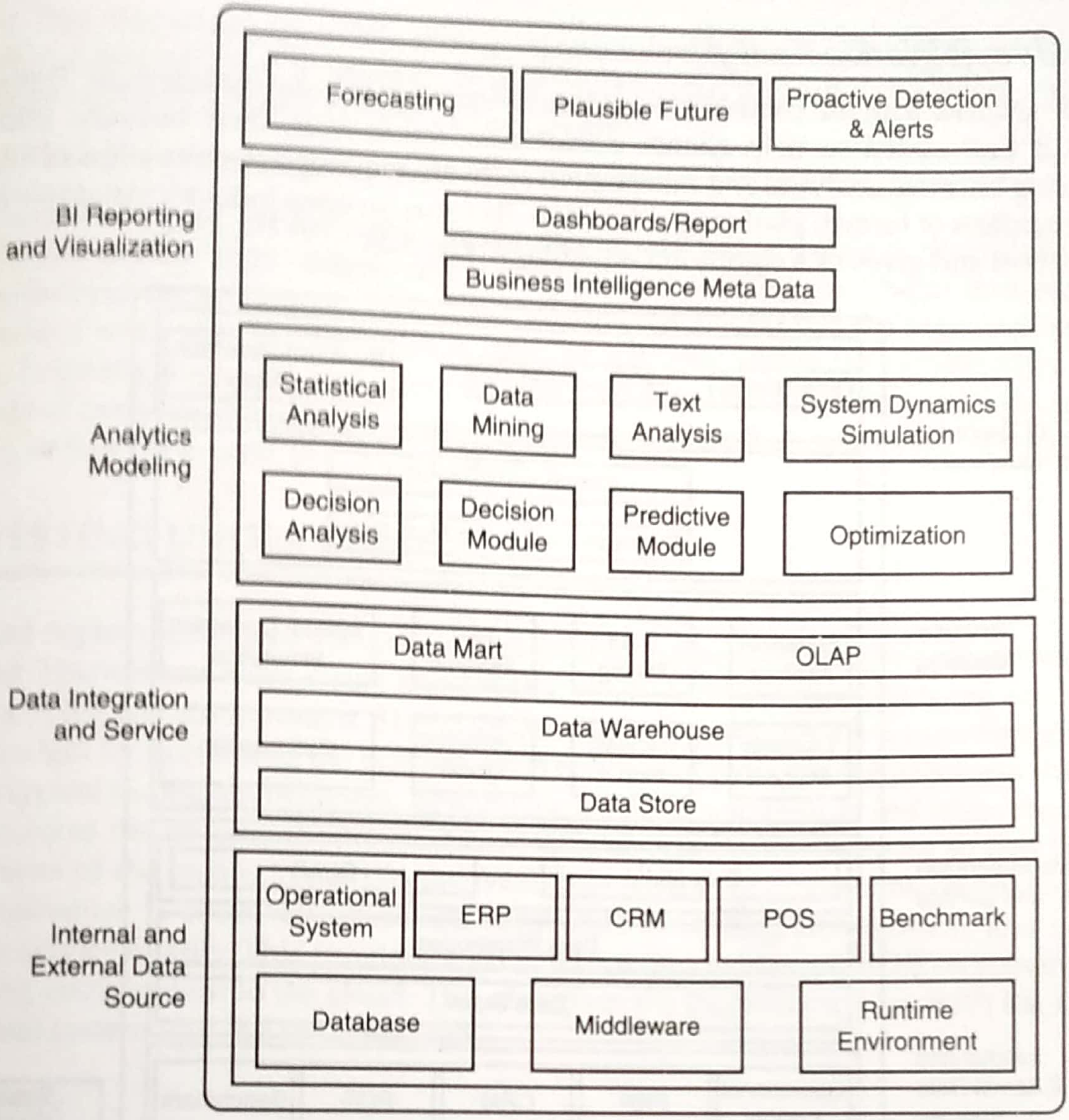

FIGURE 5.1 Cloud analytics.

It can combine complex analytics with the newer software platforms and will lead towards the predictable business situation out of every business insight.

5.3.1 *Cloud Business Analytics Competencies*

The practice of cloud analytics requires multiple service lines to strategize how client can achieve the goals in faster manner coupled with less risk. The aligned competency features business intelligence and performance management that help increase performance by providing accurate and on-time data reporting. The next is competency comprises of analytics and optimization that provide different types of modelling techniques, deep computing and simulation techniques to check for different types of 'what if' analysis to increase performance. Another competency is enterprise information management that helps in applying different architecture related to data extraction, archival, retrieval, movement, and integration. Content management system is another competency that is required for cloud analytics and includes different service architecture, technology architecture, and processes related to capturing, storing, preserving, delivering, and managing the data. It also helps to provide access in the global environment and makes it easy to share data with stakeholders across the globe.

5.3.2 How it Works: Analytics

Analytics works with the combination of hardware, services, and middleware. This expertise makes it best suited to help clients extract new value from their business information. Delivering business analytics and information software requires a seamless flow of all forms of data regardless of format, platform, or location. Its focus on open industry standards is the key to this effort and gives us a significant advantage (Fig. 5.2).

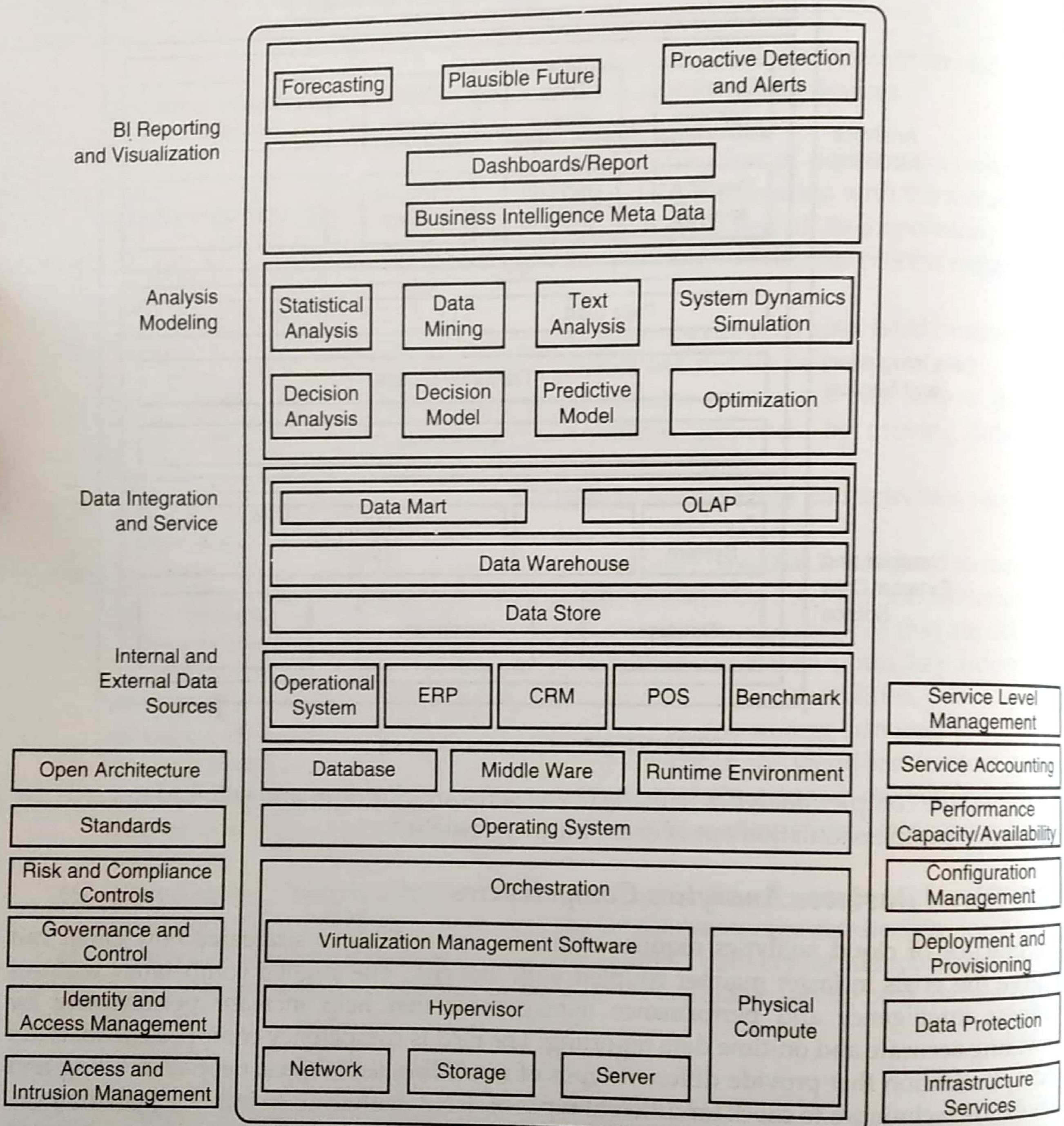

FIGURE 5.2 Cloud analytics business outcomes.

The analytics system features include the platform that provides data reporting, analytics based on text, mining activities, business intelligence dashboards, and perceptive analytics techniques. This also takes care of the storage optimization and different high-performance data-warehouse management techniques. It also involves the umbrella activity of different installation services and a highly reliable system platform.

Analytics Business Outcomes

Analytics systems help to get the right information as and when required, identify how to get it, and point out the right sources to get it. Therefore, analytics also helps in designing the policies faster as decisions are based on the information available in the organization and decision-makers work with the exploration services available within the organization. This system also helps in gauging the business results by measuring the different metrics generated with the help of analytics. Analytics give options through which the organization can increase profitability, reduce cycle time, and reduce defects.

5.4 TESTING UNDER CLOUD

Private cloud deployments need virtualized servers that can be used for testing the resources for the cloud. This ensures more secure and scalable solutions where consumers can access the IT resources in the test environment. Therefore, testing under the cloud environment gives a very good insight by decreasing the manual intervention and reducing the processes that are involved in typical testing environment. By enabling access to resources as and when required, the cloud reduces the investment on capital, as well as enables the business to handle the ups and downs of the testing requirements. With this feature, organizations can reduce the test cycles, minimize IT costs, reduce defects, rationalize the testing environment, and hence, improve the service quality. This provides a good return on investment (ROI) on moving the typical testing environment to the cloud. This also gives the flexibility to play with the surrogate of the real system without the actual risk.

5.4.1 *Benefits*

1. Cut capital and operational costs, without affecting the mission critical production applications.
2. Offer new and innovative services to clients, present an opportunity to speed up the cycles of innovation, and improve solution quality.
3. Facilitate a testing environment based on the request, and provide a request-based service for storage, network, and OS.

5.4.2 *Value Proposition*

Business test cloud delivers an integrated, flexible, and extensible approach to test resource services and management with rapid time to value. This is an end-to-end set of services to strategize, design, and build request-driven delivery of test resources in a cost-effective and efficient manner.

These test tools allow you to orchestrate and build your services and development projects, and permits to catalog and organize all of the various software assets. These can be provisioned by the administrators, development team leads, or project team members who are permitted to do the same.

3. **Allocation and run-time scaling:** Allocation is the process of instantiating a service catalog item: infrastructure, platform, or application. Run-time scaling is flexing-up or flexing-down of required resource elements to meet the SLA requirements defined by the application owner according to the standard corporate standards and business policies. The solution should be unique in providing both IaaS and runtime scaling/workload management in one product offering for the enterprise environment.

Benefits

1. **Increase agility and innovation**
   - Enable self-service delivery (within minutes).
   - Deliver on SLAs.
   - Simplify process for 'what-if' experimentation.
   - Gain control over public cloud usage.
2. **Decrease costs**
   - Increase utilization.
   - Increase operational efficiency (100 servers per admin).
   - Achieve a greener datacenter.
   - Maintain vendor choice.

Offering Key Characteristics

1. **Service layer**
   - Self-service portal for different cloud users: administrators, cloud delegates, and end-users.
   - Chargeback/billing and reporting on the basis of usage and capacity.
   - Operations management self-service portal with service management.
2. **Applications**
   - Automated application provisioning and lifecycle management.
   - Dynamic scaling to meet SLAs.
3. **Allocation engine**
   - Account-based quotas, reservations, scheduling, and approvals for resource allocation.
   - Policy based migration, movement, and failover of workloads.
4. **Resource integrations**
   - Maintenance of virtualization platforms.
   - Support for popular provisioning tools.
   - Integration for popular public cloud/external services.
5. **Datacenter integrations**
   - Role-based access.
   - Adapter-based integration to accounting, asset management, change management, entitlement, service catalog systems, and ticketing systems.

5.5 INFORMATION SECURITY

Information security risks are potential damages to information assets. Successful organizations take a risk-based approach to information security. Nothing can be 100% secure – but by knowing your current state you can take a risk-based approach. You can focus on implementing mitigating controls to address your most significant risks. The remaining minimized risk is acceptable because the likelihood of exploit and the severity of exploit versus the cost of mitigation do not have a positive cost/benefit. Risk can be quantified by the following expected (average) damage:

1. **Value of asset:** What are your valuable information assets?
2. **Vulnerabilities:** What vulnerabilities exist in your systems that can be exploited and lead to the damage of your assets?
3. **Threats:** The level of threats that aim at exploiting vulnerabilities.

Security controls are protection mechanisms to avoid information security risk and have the following features:

1. **Must be effective:** Mitigate the given risk.
2. **Should be adaptive:** Adapt to changing risks.

Three main types of security controls are as follows:

1. **Preventive:** Prevent security incidents (e.g., patching vulnerability).
2. **Detective:** Detect a security incident (e.g., monitoring).
3. **Corrective:** Repair damages (e.g., virus removal).

Successful organizations recognize risks, implement the appropriate mitigating controls, and innovate to grow their business.

5.5.1 *Expectation of Privacy*

Consumers expect that security should be built into services themselves. Over 50% of potential cloud consumers still avoid online purchases due to the fear of financial information being stolen. Expectation drives regulation, and today, vendors like automakers are expected to take a greater share of responsibility. Enterprises must shore-up their weakest supply chain partners by insuring

1. More evenly distributed security responsibilities.
2. Increased transparency from start to finish.
3. Eased burden of customer-facing unit.

5.5.2 *Security Challenges*

As the information grows day by day, datacenters and infrastructures are stretching the upper limits of these resources. They are trying to maximize the use of the power, space, and personnel; hence, CAP-EX and OP-EX are also increasing day by day. Cloud computing and virtualization give an opportunity to meet these challenges. On one hand, it gives weapons to deal with these challenges; and on the other, it also gives rise to its own problems such as veracity of the virtual environment, data integrity problems, and even security challenge.

Another area to watch is security around Web applications. An average deployed application have dozens, sometimes hundreds, of defects and a bulk of security threats today target the application layer. Companies must take a proactive action to reduce the instances for their Web applications being exploited – before a hacker even has the opportunity to compromise their business. The most feasible and economic solution for the organizations is to catch vulnerabilities as early as possible and build a secure software from the ground up and not bolt it once it has been deployed.

We are moving toward a future where enterprises will adopt the services from cloud service providers externally. The most important point is that the workloads that will use these services will be rendering the low-risk workloads. This will also account for some of the assurances for the security, and the price of the service will be the key factor to decide whether to adopt the service or not. The workloads that possess medium-to-high risk factors will adopt private and hybrid clouds. This will also cover the workloads that contain proprietary contents and also those that need more security and depth of defense. Once these services mature and settled, the latter ones will also move towards the external cloud to enjoy the benefits of the external cloud, but without compromising the security.

5.5.3 *Security Compliance*

There is a need of policies and procedures for governance and risk factors with respect to cloud security. These services should include procedures to handle change management and incident management. These services generate reports for multi-tenant environments. Therefore, one has to bank on large log and audit files to do so. Transparency is an important factor as it is very important for public clouds and it is a black box for the service users.

It is also required to conduct third-party-based checks and audits for the agreements that are breached in the process. Also, third-party-based audits can issue non-compliance to the subjected violations. This process maintains visibility in the system.

Another method is to have strong SLAs so that the flexibility can be managed for the process based on the situations that will enable the traditional outsourcing model and service management to enjoy the benefits of the cloud.

5.5.4 *Identity-Based Protection*

The cloud environment requires extra protection levels as it works with a diverse set of groups. Therefore, it is essential to have proper authentication for getting access to resources for the environment. It also requires regulated monitoring of users, details regarding the logging to the resources, and check-up for the background verifications. One of the important aspects is to maintain the access that matches with the profile of the work and gauges the risk if something goes wrong due to the improper use of resources. There can be different classes of users, such as administrators who require the access based on the work they are doing with the cloud environment.

Maintenance of the identity is required to conduct smooth operation in the cloud deployments and authenticate the real users. This is required for both internal and external purposes for the hosted applications. The biggest hurdle is to secure the confidential data. In order to do this, one has to maintain a secure protocol over the networks and activate the firewalls to ensure the security of the confidential information and the information that is sensitive but not important for the business should be destroyed.

5.5.5 *Data Protection@Cloud*

Relevant terms that dictate the protection of data are: how it is stored, how it is accessed, what the compliances are, and what audits are required as per the SLAs. It also relates to the regulation of the breach of the data and its separation on the storage infrastructure. This even includes the archived data.

This is handled by the process of encryption and is managed by encryption keys, and the data are protected in the cloud datacenter. Another point that is not taken care of most of the times is the protection of the mobile data. It should be ensured that encryption is done for the mobile data as well. One of the biggest problems in Internet-based cloud is sending a large amount of data that is not possible with the Internet-based environments. Therefore, the data should be encrypted, and both the cloud service provider and subscriber should have the keys to encrypt it.

The movement of data between different locations of the organization depends on the cloud environment, support, SLAs, and business activities. There can be violations of the intellectual property law which should be kept in mind while working with different types of data. It must be ensured that the legal teams should review all the requirements of cloud environments and the methods to control the data that are collocated in a large geographical area.

Other important thing that can be classified is the data type and its associated risks for the protection of data. The risks levels and matrix breaches can be obtained and security mechanisms based on them can be derived. The measures can be different from domain to domain; for example, public services challenges will be different from financial services data.

5.5.6 *Application Security@Cloud Deployment*

It is incorrect to think that the protection mechanism in a cloud environment works only at the application level. In fact, it is required at the image level as well. Therefore, cloud vendors should have a clear and sound way to tackle the protection mechanism by meeting the demands of the subscriber for issuing licenses for the required period of interval, destroying them after use, and making sure that the sensitive unimportant data is also destroyed at the same time.

For maintaining and supporting the security, it is important for the customer to follow the standards that cloud subscribers demand. All the Web-based requirements should be coded to match the actual requirements, and the published content on the Web should adhere to the policies of the business.

In order to work with the successfully protected virtual environments, everybody in the cloud deployment should adhere to an agreed-upon basic security policy. Cloud vendors and subscribers should have the intrusion-based polices audited and should check the prevention system placed to handle this. This exercise becomes more valuable when we work in a shared environment because different subscribers on the same cloud environment should have the agreement for security and protection policies.

So far, we have talked about the security measures on the basis of software, policies, SLAs and audits; but we should also take care of the security in physical terms as well.

These security measures can include biometric systems and closed circuit television (CCTV) monitoring. These measures can restrict any unauthenticated entry into the system.

5.6 VIRTUAL DESKTOP INFRASTRUCTURE

Virtual desktop infrastructure (VDI) provides end-user virtualization solutions. This is designed to help in transforming distributed IT architectures into virtualized, open-standards-based frameworks leveraging centralized IT services. VDI combines hardware, software, and services to connect the clients' authorized users to platform-independent, centrally managed applications, and full desktop images that run as virtual machines on the servers in the datacenter.

The VDI solution introduces a new method of delivering and managing user desktop environments. In the IT industry today, several technology vendors provide components that enable building a VDI. VDI is generically used with reference to a collection of products and infrastructure components used to form a virtual desktop solution. This solution consists of the following:

1. Portal interface.
2. Thin-client.
3. PC with client components such as
   - Browsers with client messaging.
   - Security technologies.
4. IT infrastructure.
5. Varied client devices.

Project-based services provide IT consultants specializing in virtualization technologies, assistance to assess the organization's desktop and application needs, and subsequently develop a VDI solution that best meets these needs.

5.6.1 Architecture Overview

Desktop cloud virtualization services provide several advantages to an enterprise. One of very important advantages is the reduction of cost. By moving the core function of distributed end-user devices and applications to a centralized infrastructure, the lifecycle of the end-user device is extended, and the performance requirements are moved to a centralized infrastructure. The administration of a centralized IT infrastructure is more cost efficient than the administration of a distributed one. The implementation of virtualization with VDI solutions allows businesses to simplify their IT environment, reduce costs and complexity through consolidation of physical resources, and standardization of operating environments.

VDI creates a framework that offers many advantages to the enterprise such as follows:

1. **Cost reduction:** More efficient use of resources can increase utilization.
2. **Flexibility:** Common physical infrastructure can support a variety of end-users. New desktop images can be created dynamically without a hardware-procurement cycle. Multiple types of guest OSs can run on the virtual machines so that the physical

hardware can support a wide range of end-users without any costly integration or reconfiguration of the systems between user accesses.

3. **Security:** The data remain in the datacenter with the access control.
4. **Availability:** Higher availability as VM can be quickly migrated to a different physical server in the event of a hardware failure.
5. **Efficiency:** Service delivery is more efficient when IT processes are optimized for a centralized environment.

5.6.2 *Enterprise Level*

VDI provides a set of proven integration patterns and methods for implementing a client virtualization. The VDI team works with various tools and products to help the users with an assessment of their environment to develop VDI solution requirements and a solution design. The team then develops and deploys this solution into the client environment on the basis of a common architecture that supports the four VDI solution models (shared service, virtual client, workstation blades, streaming).

VDI solution is shaped by the component selection. The base VDI architecture is designed to integrate with existing client environments and, as the name implies, provide an access method to a highly scalable virtual infrastructure to the front-end clients. The VDI server farm's back-end integrates with existing infrastructure services and legacy applications.

VDI solution is designed to reduce the dependency on distributed PCs and laptops. By placing critical applications and data in a centralized datacenter with access from a variety of client device options, VDI can significantly reduce the cost and complexity of the management, and maintenance of desktop images.

VDI end-user desktops run on virtual machines hosted in a centralized IT infrastructure in the client datacenter. The end-users access their individual desktop image or a pool of desktops through a client access device such as PC client, a thin client, or a Web-based client. The applications run on virtual machines on host servers in the datacenter from resource pools rather than on the local machine.

Additional resources can be easily and quickly added to the IT infrastructure as business requirements arise. The VDI solution provides increased security as data remain in the datacenter. Optional secure encapsulation capabilities can allow network connections to be encrypted.

VDI's virtual client solution integrates into the organization's datacenter to leverage the existing network and infrastructure services. The solution provides access services to the virtual infrastructure hosted in the VDI server farm or datacenter (Fig. 5.4).

The desktop client devices can be new or existing thin client devices, PCs with an access client, a Web browser, or various combinations depending on the client environment. The desktop client devices can vary from organization to organization, and can be provided as a part of the VDI service engagement if required.

The VDI access service or 'connection broker' provides devices and user authorization, portal integration, session management, host monitoring, application streaming, and

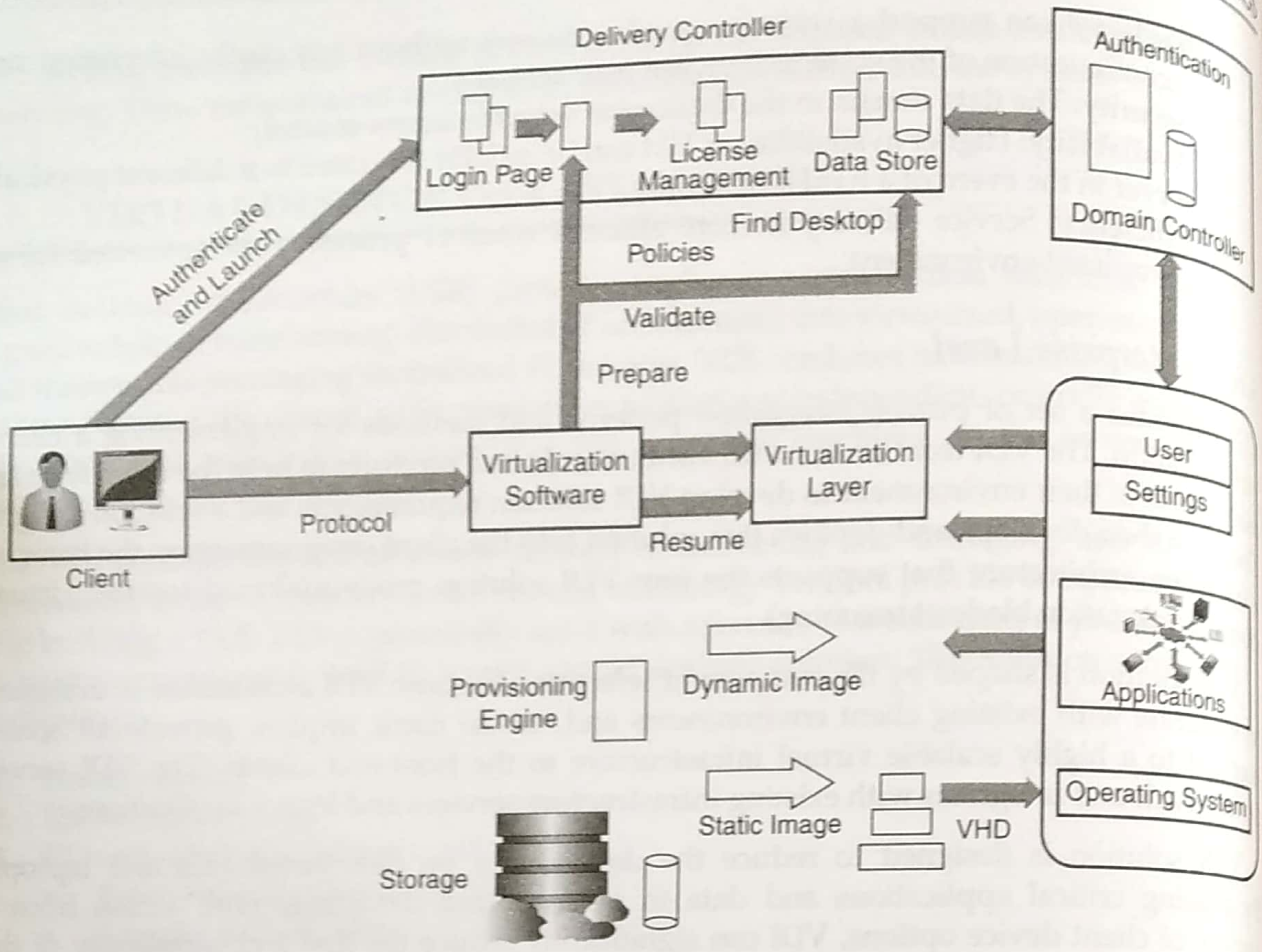

FIGURE 5.4 Virtual desktop infrastructure.

consumption-based metering. This access service also provides load balancing as needed for the front-end security servers, and for the connection servers. The security servers and connection brokers must have high availability as they are the single points of access to the virtual environment.

The VDI server farms, or datacenters, consist of a set of physical hardware platforms that facilitates the virtual environment to host shared OS, virtual machines, and dedicated desktop clients.

The infrastructure services include existing services such as activity directory, file, print, management, network, authentication services, and storage. This solution integrates with these existing services rather than introducing redundant features not required by the business. The service will not bundle in the components that are already available in the business's environment. If additional infrastructure services are required, they can be added to the project-based service for an additional charge. The cloud service project team integrates the VDI solution with the existing infrastructure services and desktop client devices as part of the project-based service engagement. Additional offerings may be combined with the cloud service product to meet the business requirements as needed.

The virtual desktop images can be configured to access the existing infrastructure services in the client datacenter as needed. The following sections briefly discuss the components of the VDI at a higher level.

5.6.3 *Client Access*

The users use their remote desktop client device to connect to their virtual desktop. This service is provided mostly by all virtual desktop vendors and is supported by a set of products that connect remote clients to the centralized virtual desktops. End users can use existing PC-running OSs where the user initiates a remote session to a VDI resource using a remote desktop client application or a Web browser. With some third-party products, one can have local desktop icons configured to access published applications or published desktops.

5.6.4 *Desktop Virtualization Services*

VDI can be viewed as a central-server-based resource pool with components connecting end users to applications, networking, and storage resources. VDI uses vendors to centrally host and deliver a cost-efficient desktop environment from the datacenter, and uses the management server to provide the virtual desktop management.

Virtualized clients and desktop management proactively manage diverse desktop environments and VDI server-based client technology. The end user retains the features and flexibility of the traditional desktop.

Management server authenticates users, determines the pool they belong to, and using predetermined policies provisions a desktop for those end users, with complete users' specifications and privileges, and finally deploys it. These pools can be persistent or non-persistent.

Individual desktop assignment is a static, one-on-one relationship between a user and a specific virtual desktop. This configuration is good for power users where the desktop is specifically configured for a particular user. This configuration can include specific applications, data access, and resource allocations. Individual desktops give users a high degree of customization.

A virtualized IT environment helps in providing security-rich, anywhere-anytime availability and access to applications, information, and resources. VDI is a unique end-user virtualization solution that helps businesses transform their distributed IT architectures.

Desktop cloud project-based service can substantially reduce the total cost of ownership by reducing the effort required for desktop PC deployment and management, software distribution, desk-side support, and help-desk required to support and maintain desktops. Businesses can avoid significant upfront investment and continuing cost for developing and maintaining the necessary skills, knowledge, and experience in systems management and desktop virtualization technologies.

5.6.5 *Desktop Management*

VDI brings together the desirable features of traditional terminal server while retaining the important features of distributed computing.

5.6.6 *Pool Management for Virtual Desktop Infrastructure*

The role of a management server is to maintain the authentication of the user, check the pool for the user, he belongs to, the policies he is entitled for, and ultimately provision the desktop for the specific user. These pools can be persistent or non-persistent. The connection server allocates entitled users to a virtual desktop from the non-persistent pool as requested. This allocation is not retained when the user logs off and the virtual desktop is placed back into the non-persistent pool for re-allocation to other entitled users. When the user connects to the non-persistent pool on subsequent occasions, the management connection server connects the user to any virtual desktop in the non-persistent pool. This is typically a many-to-many relationship.

When a group of users is entitled to the persistent pool, every user in the group is entitled to any of the virtual desktops in the pool. The management connection server will allocate users to a virtual desktop as requested. This allocation is retained for subsequent connections.

Maximizing the use of non-persistent pools for all users who do not have a desktop customization requirement is recommended. Typically, the task users could do with a non-persistent desktop.

VDI also offers managed services for businesses that wish to derive benefits of virtual infrastructure access but lack the necessary skills and expertise required for the ongoing management of the virtual infrastructure. Businesses can avoid significant up-front investment and continuing cost for developing and maintaining the necessary skills, knowledge, and experience in systems management and desktop virtualization technologies.

5.7 STORAGE CLOUD

For any type of cloud deployment – private, public, or hybrid cloud – the environments are built using key foundation building blocks such as servers, storage, applications, and infrastructure. Storage and compute resources scale together, and the failure to manage them efficiently results in the failure of cloud services.

Storage management in the cloud can help organizations to address their challenges related to data and storage management in their clouds, for example availability of data at all times, storage resource utilization, application performance, longer restore times, higher storage costs, low productivity of storage personnel, and increased risk of data loss and downtime.

5.7.1 Value Proposition

Data and storage management within a cloud environment is a critical necessity to provide a reliable, on-demand service experience while at the same time reducing the costs and minimizing the risks. Streamlining the data to target applications plays an important role, which means the data has to be made available at all times and the storage should be provisioned rapidly to the applications built into the cloud for delivering efficient services. Often, storage administrators spend over 50% of their time on manual repetitive tasks. They find it difficult to meet stringent rules that are essential for restoring operations quickly after any disruption (database corruption, virus attack, disaster, and hardware failure) in a cloud. Failure to ensure data availability at all times can lead to a significant failure of a cloud service.

Also, the proper placement of data on different tiers of storage within the cloud, if done efficiently, helps to minimize the overall costs of hardware, software, and administration.

Cloud vendors offer technologies – storage, hardware, and software – as well as key storage services to support subscribers in their journey to leveraging cloud computing. They can assist in planning, designing, building, deploying, and even managing and maintaining storage solutions, whether on their premise or someone else's.

Cloud technology is helping organizations to build a smarter business infrastructure with immense flexibility and scalability – one that could result in improving service levels while reducing capital and operational costs. Today, many organizations in various industries including media and banking, which deal with large amounts of data, are increasingly adopting the cloud technology to address their needs of delivering faster services, protecting data in real time, seamless communication between employees, partners, and suppliers, for business continuity, and, of course, to become more energy efficient – to be a greener organization.

5.7.2 *Challenges*

Cloud services majorly focus on keeping the data and applications they are managing available at all times, and even more essential is to develop the ability to quickly restore the operations following any type of data disaster. Storage management is an important function to ensure that the data are available, capacity is provisioned rapidly, and storage resources are utilized effectively. Cloud administrators often find it difficult to meet all the following challenges they face concerning storage and data management:

1. Data availability and application performance.
2. High capital and operating costs, less ROI.
3. Utilization of storage resources.
4. Lack of automation – low productivity of storage personnel with specialists doing mundane tasks.

For customers, the drivers for adopting cloud technologies have been cited as follows:

1. Paying for only what they use.
2. Reduced costs.
3. Monthly payments instead of all up front
4. Having a standardized system
5. Always having the latest software version, since nothing is installed locally

5.7.3 *Business Drivers*

1. Need for standardization and automation of storage services.
2. Need to meet service levels consistently – provisioning on-demand computing capacity and storage capacity.
3. Need for simplified management of storage infrastructure – quick provisioning and redeployment of resources, built-in data reduction capabilities.
4. Data security and compliance issues.
5. Need to lower costs – lack of upfront capital, lower utilization of hardware resources
6. Recovery point objectives (RPOs), recovery time objectives (RTOs).

5.7.4 *Benefits*

1. Improved service levels by ensuring data availability and application performance and by quick provisioning through automation
2. Reduced capital and operational expenses by leveraging standardization and automation
3. Optimized utilization of storage resources and built-in data reduction capabilities to manage more storage with less hardware
4. Reduced hardware, software, and administration costs with policy-based data storage management.
5. Managed risk and streamline compliance through real-time data protection.

5.7.5 *Product/Solutions Overview*

Storage management software and services solutions for the cloud help in ensuring that business and IT are fully aligned and supported by integrated service management. They help in delivering a workload-optimized approach and offer a choice of implementation options for superior service delivery with agility and speed.

Cloud vendors offer second-generation storage management technology for cloud environments, delivering faster ROI. Cloud storage services include worldwide capability and capacity to provide integrated cloud service offerings to meet your storage management needs.

Cloud vendors reduce the complexity of managing cloud environments by offering a complete portfolio of automated solutions for managing data and storage infrastructure, enabling better efficiency for business resiliency, reducing costs, and improving security, while increasing visibility, control, and automation of the cloud storage infrastructure. There is a need of the broadest, most scalable, and reliable set of storage solutions available to keep the cloud services functional. These storage solutions should have the complete portfolio for protecting, managing, and virtualizing the environment.

5.7.6 *Product/Solution Description*

Cloud vendors should offer a complete portfolio of software solutions and services for the storage management in the cloud, designed to help in streamlining the storage resources in order to support cloud services, protection, and management of data, being able to virtualize the entire storage infrastructure, and offer it as a single resource to the cloud services bundle.

Cloud-based deployment and service models use new and scalable delivery models to offer cost-effective solutions. Data availability and application performance are critical factors to migrate the applications in the cloud. Additionally, a cloud environment cannot afford longer downtime as after any disruption, it is critical to restore operations quickly. Whether it is an organization managing its own private compute/storage cloud environments or a managed services provider offering cloud-based services (private or public), the key concern is how well the existing resources are optimized in their infrastructure, to improve end-user experience - whether they are able to provide flexibility, speed, reliability, and efficiency. Data and storage management play a critical role in improving on-demand service experience and reducing costs and risks in the cloud.

Often, the absence of sophisticated storage management systems in a cloud results in the lack of visibility into the storage utilization and provisioning, costs, and associated risks. Cloud administrators face difficulty in understanding how much capacity is available, where it is, which applications are accessing it, how secure they are, and whether they are able to meet stringent RTOs. Gaps that may exist between the unpredictable demand for data availability and the ability of the business to support the same in an efficient way, results in unmet service levels, additional downtime, new hardware and operational costs, and lower customer satisfaction.

5.8 SUMMARY

Today, as we are becoming interconnected, instrumented, and intelligent – the world is becoming smaller – more data is being generated within the operations of all organizations, and they are struggling to manage the complexity in their storage environments. The costs of backup and recovery, archiving, expiration, and storage resource management are exploding. It is not just about increasing the capacity but managing the data efficiently, reducing the data, ensuring adequate protection of the data, and quick restoration of the data for better business performance. We need better implementation of VDI, test cloud environments, and analytics to handle cloud offerings.