Management Platform for Disparate Cloud Ecosystem

Empowering Enterprises for Seamless and Expenditious Use of Virtual Resources

20X BUSINESS AGILITY

35% COST OF OPERATIONS

MINUTES TIME TO PROVISION

SERVICE DEPLOYMENT

POLICIES & AUTOMATION

HA & AUTO REMEDIATION

ANALYTICS & OPTIMIZATION

MEASURED SERVICES

INTEGRATION

FUTURE PROOF

Issue 1

Featuring research from

Gartner®

A Robust Platform for Seamless and Expeditious Management of Disparate Cloud Ecosystem!

Cloud360 Overview

Cognizant's Cloud360 is an innovative platform that empowers enterprises to embrace cloud-enabled environments with industry-leading platforms, thus enhancing their IT services while improving business agility and operational efficiency. Cloud360 leverages private or public environments to cost-effectively manage applications, deliver IT services, ensure enterprise compliance and governance, and more importantly, protect existing IT investments. Enterprises can gain the combined benefits of a powerful cloud infrastructure with a robust cloud management service.

Cloud Management Platform for Service and Operations Management

Delivered along with services, Cloud360 enables enterprises to transform existing IT environments to realize cloud business benefits, integrate with other eco-system components, and manage across industry-leading private and public clouds. Cloud360 enables easy deployment and management of applications across cloud environments so that users benefit from self-service on-demand access to the infrastructure and IT leaders benefit from better compliance, greater transparency, and operational governance.

Cloud360 prepares cloud environments for the future by offering a seamless service management layer between applications and infrastructure for better governance, business agility, operational efficiency, and integration across the enterprise. Being a technology-independent platform, it can orchestrate workloads across multiple cloud platforms, hypervisors, and operating systems. Cloud360 has successfully implemented cloud services for access management, service management, and service optimization tiers for enterprises.

FIGURE 1 Cloud360 Service Management & Operations
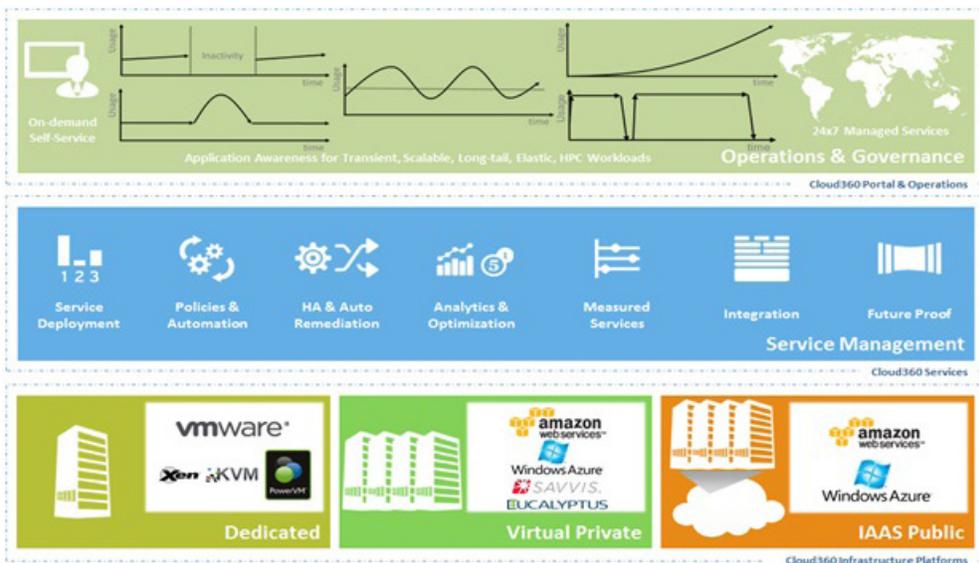
Source: Cognizant

Customers get key benefits of on-demand application roll-outs, automated deployments, auto-remediation of environment issues, SLA-driven application management, and efficient consumption metering. This helps them to adapt better to changing business needs and to reach markets faster.

Cloud360 Offerings @ Work

Cloud360 has successfully completed numerous engagements where enterprises have significantly reduced costs by deploying Cloud environments. The following are some business scenarios where Cloud360 can be put to work.

On-demand and Self-service IT

Cloud360 enhances the business agility and operational efficiency of IT by enabling on-demand provisioning, while enforcing compliance. Enterprises can scale in or scale out their resources in the cloud environment based on their needs within minutes. Cloud360, with its application profiles and instance profiles, ensures a smooth and rapid deployment of application environments. Users can perform 'one-click' deployment of application environments just in time across industry-leading platforms through the self-service portal, making IT simple and productive.

Automation and Governance

Cloud360 enables users to define complex automation policies using simple, intuitive English-like language. These policies offer auto-healing of environmental issues, ability to schedule tasks and control cost. This **Cloud Automation Language (CAL)** provides users the much-needed flexibility to create policies and gain a greater control of their environments. Easy to adopt and customize, it can be quickly implemented across cloud operations for virtual entities related to applications, infrastructure, and monitoring.

For example, a manager wants to define a policy to notify of the top 5 users with highest burn rate and prevent any further provisioning requests, when the consumption metering in last 24 hours goes over 50% of the previous 30-day consumption average. The policy using CAL is simply defined as follows:

```
If

  Consumption metering ("Compute Date" >
  -24h AND Usage ("Compute Date" = -30d) >
  50)

Do

  Notify the Owner (Usage (Top 5))

  Restrict any provisioning operation
```

FIGURE 2 Business Use Cases for Cloud360
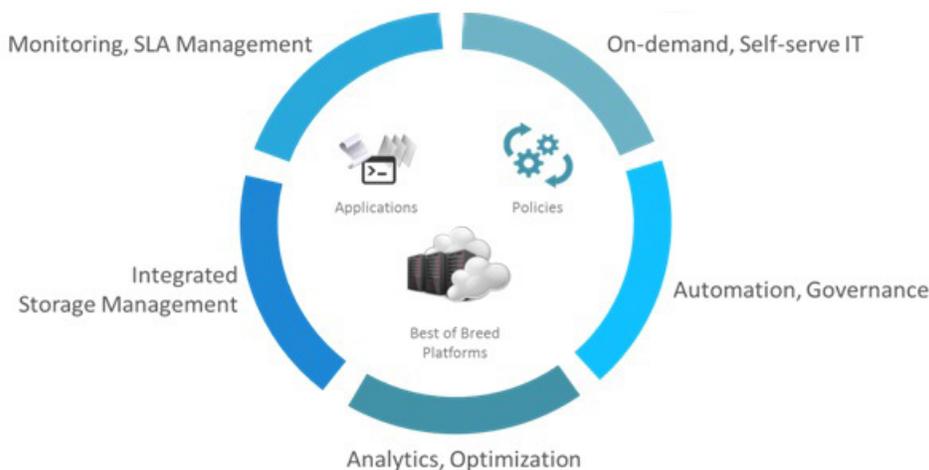
Source: Cognizant

Additionally, Cloud360 brings in the necessary standardization of cloud resources by using instance profiles and application profiles, thus enabling simpler management of cloud resources while ensuring appropriate governance.

Analytics and Optimization

Cloud360 further enhances the pay-per-use benefit of a cloud by offering consumption metering and analytics on resource usage and activity. Visibility into cost of running applications, compute instances, and storage can be determined to drive better utilization, allocation of resources, and IT budgets. Businesses can track the cost of deploying and maintaining business services/applications through the Cloud360 dashboard, thus empowering them to make better business decisions.

Monitoring and Service Level Management

Cloud360 offers a zero-touch monitoring and auto-remediation of the application environment, thus ensuring high availability and application SLA adherence. Applications, associated services, OS resources, and hypervisors can be monitored. Pre-defined application monitoring profiles can be used as is or can be customized and applied on instances with a single click, to simplify monitoring complexities and ensure compliance.

Integrated Storage Management

Cloud360 offers intelligent storage management features for provisioning and managing storage environment, such as profile-driven storage, alerts on performance issues, and storage shortages. With Cloud360, users can schedule backups, monitor the storage environment for usage, performance, and availability and generate reports as required. These capabilities boost efficiency by automating manual processes and simplifying administration tasks.

A Robust Cloud Management Platform

Cloud360's key capabilities, such as policy-driven automation, integration with existing IT environment, and right sizing of the infrastructure for application elasticity, makes it a robust Cloud Management Platform for implementing business-critical application on the cloud. Cloud360 provides comprehensive services across the following tiers:

Access Management Tier

Cloud360's self-service portal empowers users with the ability to provision resources on demand and the dashboard helps them to monitor the status of cloud resources and requests or tasks. With the Self Service Web Portal, Command Line Interface (CLI), and Application Programming Interface (API), the product is highly configurable and can be integrated with third-party and proprietary software tools and applications. The open, extensible architecture of the product allows users to integrate and work with many products and tools that the enterprises already possess. Cloud360 supports role-based authentication. It enables enterprises to create customized roles and define permissions and access rights for performing various operations in the cloud.

Service Management Tier

Cloud360 offers a seamless service management layer between applications and infrastructure, helping enterprises achieve business agility, operational efficiency, better IT governance, and tighter integration of components across the enterprise from applications and platforms. A standardized service catalog offerings through application profiles and instance profiles ensures smooth one-click deployment of application environments without compromising on the governance policies. Coupled with Cognizant's Managed Services, Cloud360 continuously monitors the application environment and manages the cloud based on the business metrics and baseline SLA to ensure high availability. Users can customize different monitoring profiles and policies for tracking the performance of cloud resources and automating remedial measures. Cloud360 also provides greater visibility into the provider billing, consumption of cloud resources, and show-back costs for every group and deployment level. Business analytics and reporting empowers management for better resource allocation, cost control, and decision-making.

Service Optimization Tier

Cloud360 supports multiple systems (hypervisor, OS, infrastructure, and cloud platforms) to protect existing investments by the client and create a "technology independent" platform that orchestrates workloads across cloud solutions. It integrates consulting, professional, and IT services to ensure correct assessment and sufficient

customization. It also offers modules, such as **SLA-driven IT management**, **Consumption Metering**, **Policy-driven Governance**, and **User Management**, that goes beyond technical cloud management. Cloud360 enables users to define complex provisioning and automation policies, using simple, intuitive English-like language, for optimizing provisioning workflows, enabling auto-remediation of environments, and automating cloud management. Intelligent Placement Policies can be customized to provision and backup instances at the best possible location and ensure compliance.

Addressing Traditional IT Problems

Cloud360 not only eliminates the challenges faced in the traditional IT setups but also future-proofs the cloud environments against the challenges emerging due to technological advancements and changing demands.

Breaking the Silos for Application-level View

Traditional IT setups have separate management or operations layers for each type of application, platform, operating system, hardware and storage. For provisioning an environment, business users have to approach and coordinate with different teams. These teams often work in silos and have different processes, policies, compliances, and request processing time. Coordinating with all of them for setting up environments becomes a time-consuming activity and results in loss of business opportunities.

Cloud360 provides a unified IT service, thus saving the user's efforts to deal with disparate teams and also ensuring better compliance and governance within enterprises. It enables the users to get a single orchestrated view of different resources with a clear visibility of the entire application stack. It conducts end-to-end orchestration of application environment and reduces the provisioning time from weeks to minutes.

Case study: A Top 10 global Pharmaceutical major wanted to deploy 'Development and Test instances' for their Agile Research environments. Managing vast underutilized system resources and a mix of environments, high administrative costs to provision were some of the major challenges. Cloud360 enabled faster deployments of applications by providing on-demand, self-service provisioning, and management of application development, and test environments.

FIGURE 3   IT as a Service by Cloud360
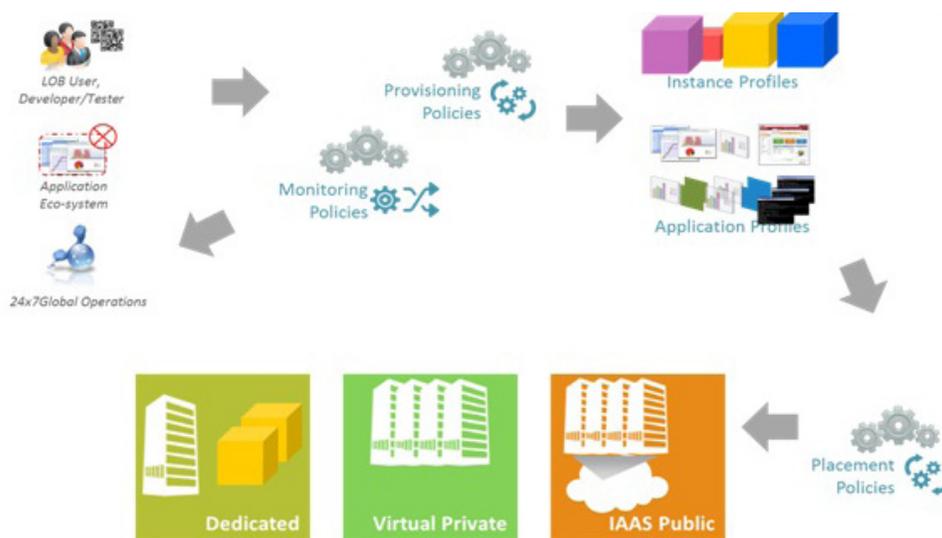
Source: Cognizant

With Cloud360, the client experienced increased utilization of computing resources and reduced management costs.

Automated Deployments to Eliminate Errors

Deploying development, test, or production environments manually involves a high risk of human error. Incorrect or incomplete deployment of an environment or even manual changes done afterwards can greatly affect the sanity of environments and its resources. Lack of standard configurations and policies can cost enterprises valuable time and resources.

Cloud360, with its application profiles and instance profiles, ensures a smooth and rapid deployment of application environments. These pre-defined, customizable profiles enable one-click deployment of environments, thus saving time and improving speed to market. Cloud360 brings in standardization of processes and workflows through its automation, provisioning, and monitoring policies, thereby reducing the possibilities of manual error.

Case study: One of the Top 5 Internet search firm required Cloud360 team to deliver a public cloud-based advertising services platform in only six weeks, despite a lack of accurate application benchmarking or process definitions for going into the engagement. The Cloud360 team led with a flexible architecture design that allowed application development, testing, and proof-of-concept on the cloud, monitoring of every failure point, including threads and concurrency against critical response parameters. The platform has a business SLA of 180 millisecond response time per ad request with automated deployment of Web instances to handle the usage.

On-Demand Provisioning for Faster Time-to-market

As businesses expand or new projects are planned, the need for acquiring new infrastructure arises and with that follows an entire gamut of activities, such as identifying the latest infrastructure, comparing costs, planning the capital expenditure budget, analyzing the requirement, and capacity utilization for all phases of the project, and acquiring the infrastructure accordingly. This planning also involves the cost of maintenance and overheads incurred in periods of low demand. This lengthy process eventually delays the application roll-outs.

One-click provisioning is one of the core abilities that Cloud360 provides, to help users in provisioning compute instances on demand. Cloud360 enables businesses to scale up or scale down their resources in the cloud environment in a matter of minutes whenever required. Users do not have to spend a lot of time planning for purchasing new infrastructure and resource allocation. With Cloud360, they can provision environments just in time and lower their TCO and increase their business agility.

Case study: A leading U.K. university wanted to roll out an application on cloud for an estimated 200 users. However, in the first eight weeks, they received 12,000 paid users for their application. Scaling the environment, meant to handle 200 users, to 12,000 users in a very short time was achieved successfully using Cloud360, with no people involvement and no compromise on the page load SLAs.

Auto-remediation for High Availability

Traditional IT setups require dedicated operations and maintenance teams to ensure that all resources in development or test environments such as servers, instances, and networks work smoothly. Such people-based infrastructure management incurs high costs for businesses.

Cloud360 brings in monitoring and auto-remediation of the application environment, thus ensuring high availability. Cloud Automation Language (CAL), an intuitive policy-defining English-like language, built by Cloud360, provides users the flexibility to customize automation, provisioning, and monitoring policies as per their requirements. It brings in standardization of processes and workflows through these policies. Cloud360 provides application awareness for transient, scalable, elastic, long-tailed, and HPC workloads along with managed services. It eliminates the people and process errors and failures that cause more than 40% of all mission-critical IT service outages.

Case study: A leading Engineering Design Company, suffered from lack of insight into application availability. Performance of services and application scalability was questionable for its P2P SaaS Aggregator for Industry and Social Resources. Cloud360 provided SaaS enablement, business growth, and zero-notice to onboard new users with increased application visibility and

support for the entire application stack along with SLA-based services.

Consumption Metering for Cost Control

Business users often do not have control over their own infrastructure. Resources scaled up during peak development periods remain underutilized for most part of the year. As businesses expand, it becomes difficult to track the utilization of resources. Managers lack visibility of capacity planning resulting in locking of capital in underutilized resources.

Cloud360 Self-Service Portal Dashboard offers a clear visibility of the application stack and allocation of resources in the cloud to the managers. Its flexible and configurable policy framework allows them to set cost and resource consumption limits for application deployments, locations, users groups, or departments. With analytics and reporting, they can view the utilization trend of resources such as memory, CPU, storage, and network and control their consumption across the various environments accordingly.

Case study: A leading U.S. insurance major was evaluating options for hosting its development and test environments on a public cloud environment, but with consistency across the environments and tracking the consumptions of resources across business units. Cloud360 monitored the consumption across the various environments, suggested an optimized usage scenario and eventually reduced the provisioning timelines from months to minutes with transparency on the usage metrics.

Automated Disaster Recovery

As traditional IT infrastructure is located at one location, protecting data and recovering it in case of any disaster becomes tricky. The time taken to fix issues and get a down environment up and running again is very high. The environment downtime can continue for hours, thus affecting productivity and SLAs.

Through Cloud360's monitoring policies and profiles, users can set threshold limits for various parameters responsible for the performance of applications, instances, OS, and networks. If any of these cross the limits set, automatic triggers are created to auto-heal the environments and stabilize them in very less time. Configurable Intelligent Placement Policies can be created to provision instances and store and backup data in the best possible geographical locations. This level of automation facilitates proper disaster recovery without much loss of productivity.

Delivering Enterprise IT Services with Governance and Compliance

Cloud360 empowers users by delivering IT services on demand. Cloud360 future-proofs the clouds, improve business agility by 20X, and reduce costs of operations by 35%. The following unique services offered by Cloud360 equip enterprises to achieve greater business agility and improved operational efficiencies.

Services for improving Business Agility

- One-click and Zero-touch provisioning to shorten the time-to-market and application provisioning from weeks to minutes

- Self-service portal, Web Services API, and CLI for a "do-it yourself" rich user experience

- Instance Profiles and Application profiles for standardizing and simplifying automation of application deployments

- Plug-ins and Integration for seamless business integration to retain previous IT investments and integrate existing tools with cloud

Services for improving Operational Efficiency

- Policies and Flex-ops to enforce business rules for cost, control, and operations in cloud

- Application Awareness for holistic management of environments where the business needs drive application sizing on the fly

- Consumption Metering and Chargebacks for accountability

- Analytics and Right sizing enabling optimizations for higher utilization and lower costs, hour by hour

Cloud360 supports multiple industry leading platforms with off-the-shelf defaults and plug-ins for rapid application deployment. It manages application resources across these platforms in an integrative and cohesive manner. Users can switch between different clouds, add, or remove clouds, and change platforms smoothly for holistic cloud management. Cloud360 supports all major operating systems, such as Linux, Windows, and AIX, and can be accessed using popular browsers and devices.

Cognizant Cloud360 Approach

Cloud360 provides an end-to-end cloud migration and management service with its **Transform**, **Integrate**, and **Manage** approach. The Cloud360 team conducts a thorough assessment of the existing infrastructure, compares cloud alternatives, creates baseline for applications, and performs migration analysis to help transform the customer's requirement in to a strong cloud strategy. It seamlessly integrates Cloud360 platform with the customer's existing IT set-up and enables on-demand provisioning of infrastructure and applications. With its global delivery network, it manages the cloud's day-to-day operations including remote monitoring, tracking

resources consumption, capacity right sizing and optimization.

Conclusion

As enterprises are making swift strides in cloud adoption to quickly respond to business demands and market opportunities with the need of a cost-effective strategy, Cloud360 presents a viable proposition through its accelerated deployment and flexible management of resources without vendor lock-in. Its key capabilities including Service Automation, Policy based orchestration, On-Demand self-service and Application Awareness brings the greatest benefits to the customers in both Business Agility as well as Operational Efficiency. Combining this with its underlying technology neutral approach and support for hybrid cloud allows for maximum flexibility and reduces significant risk to the customers in deploying their cloud solutions. Cloud360 has successfully redefined the way top enterprises across industries, such as life sciences, software development, engineering, airline, luxury apparels, online search, media and academics adopt and operate in the cloud. With on-demand provisioning, customizable automation policies and a unified way to manage multiple cloud

FIGURE 4   Cloud360 Approach

Transform
Review, Quantify & Initiate Transition

Integrate
Deploy, Integrate IT & Enable

Manage
Monitor, Control & Optimize

Source: Cognizant

platforms, Cloud360 provides enterprises a clear visibility and control of their cloud environments, thus increasing operational efficiency and lowering expenditures.

With Cognizant's deep experience in SLA-driven management and monitoring, Cloud360 helps enterprises of all sizes to deliver higher application availability to its customers 365 days and 24 hours a day. With the rapidly changing market of cloud service providers and Cloud360's periodic feature updates, enterprises find it easier to adapt to the dynamic business environmental shifts. Cloud360 offers enterprises the combined benefits of a powerful cloud infrastructure with robust cloud managed services, thus improving business agility by 20X and reducing costs by 35%. With comprehensive end-to-end cloud enablement and management services, Cloud360 becomes the ideal partner for enterprises planning operations in the cloud. Cloud360 is committed to ensure a smooth transformational journey of enterprises to the cloud.

Source: Cognizant

How to Build an Enterprise Cloud Service Architecture

Fully architected enterprise private or hybrid cloud offerings and the associated management functionality should consist of five logically separated layers. A cloud service should have independent layers that maximize deployment flexibility, while enabling multivendor substitution opportunities.

Overview

Here, we define the logical architecture of an enterprise cloud service and its management. We would expect to find five layers in a complete cloud service solution, but acknowledge that there could be wide variations in typical deployments. Gartner defines a cloud management platform as the upper layers of the cloud solution.

Key Findings

- An enterprise cloud service architecture has five functionality tiers: access management, service management, service optimization, resource management and the underlying resource tier.

- Ideally, a cloud service's layers should be logically independent of each other, they should maximize deployment flexibility and they should enable the potential for multivendor substitution.

- Cloud services may have reduced requirements for a service management tier — for example, in nonproduction, where fewer formal IT management processes are generally applied.

- Enterprises are likely to end up with many sets of cloud resources spread across internal and external resources. Unless they enable the shifting of resources, resource-sharing efficiency could be hampered.

Recommendations

- Ideally, you should assess vendor cloud solution claims against the five components independently, but realize that vendors will want a greater proportion of your architecture, which will increase lock-in and stickiness.

- Choose a cloud management platform (CMP) that enables management of hybrid resources and services, those operating both internally and externally.

- Not all cloud deployment scenarios will require the complete "stack" of capabilities, but for mission-critical production clouds, all layers should be present for at least a minimum set of functionality.

- Build cloud services with real-time infrastructure (RTI) architectures where there is a need for dynamic service optimization.

- Start your hybrid cloud journey by assessing requirements for what needs to be processed internally and what services can be processed on the outside. This will help enable policy-based service delivery.

What You Need to Know

Private and hybrid cloud services will require a cultural and political change inside the IT organization to make the role of operations more proactive. This will require predefined planning, policies, service levels and automated actions to take on the runtime environment, as opposed to the manual initiation of scripts or workflows. The application of the concept of "running IT like a business" will become mandatory in this environment, because IT organizations will have to learn and refine such areas as service catalog, capacity management, financial management and, more generally, service orientation. This will require new skills and transformation. Ideally, the architecture of a cloud service will not be monolithic (or tightly integrated) to enable more IT organizational flexibility and choice. However, selecting your components means that the IT organization must become its own system integrator (SI).

Alternatively, many IT organizations will choose more functionality from fewer vendors to reduce the SI effort. We recommend starting with a cloud computing strategy, so that you know how you want to operate your IT business model.

For educating yourself on cloud computing, we recommend that you start small, with a limited span of control, but build a private cloud and consider RTI architectures for expandability — more services, broader governance and more shared resources. That way, when you put the strategy together with what you've learned from starting small, you should be able to gradually develop your new cloud service architecture and its management, rather than taking a big-bang approach.

Analysis

To deliver on the five defining attributes of cloud services (service-based, scalable and elastic, shared, metered by use and Internet technologies), a cloud service architecture has five tiers of integrating components (see Figure 1). Enterprises can use these requirements to evaluate market solutions for private and hybrid cloud services — infrastructure as a service (IaaS), platform as a service (PaaS) and software as a service (SaaS) — and their management. Moreover, these requirements will aid in evaluating a CMP that can manage both internally and externally sourced services and resources based on policies. A CMP must include the top three layers of Figure 1; however, it does not need to implement the resource tiers, because it could integrate with the resource management layers and be a type of manager of managers.

Not all vendor offerings will adhere to this model. Clients do not need to deploy this entire stack for every scenario; some layers may be thin in functionality or missing altogether. For example, test/development environments are likely to need less-sophisticated service management capabilities and may, in fact, delegate service management functions to the customer. However, Gartner believes that, to achieve the fullest degree of benefit, while minimizing risk for mission-critical environments, a cloud service requires all the capabilities shown in Figure 1.

Public cloud service providers typically provide one service, or a small number of services, on a large scale. Enterprises interested in private cloud services usually have many standard, but smaller-scale, service offerings. This could lead to many cloud silos in an enterprise, with less efficiency in resource usage. We recommend that enterprises build as few silos as possible to accommodate

an expanding service portfolio, enable broad governance and maximize the sharing of infrastructure resources. This means using fewer compute/storage resource pools and using more intelligent policies to direct the placement of services onto the resources and to optimize them. In a hybrid cloud model, IT organizations can manage service delivery cohesively across internal and external pools of resources.

The fewer the overall number of pools, the more likely each can be optimized to meet overall policies and constraints. That being said, business and political trade-offs will need to be made regarding resource efficiency, which may suboptimize it to optimize business value.
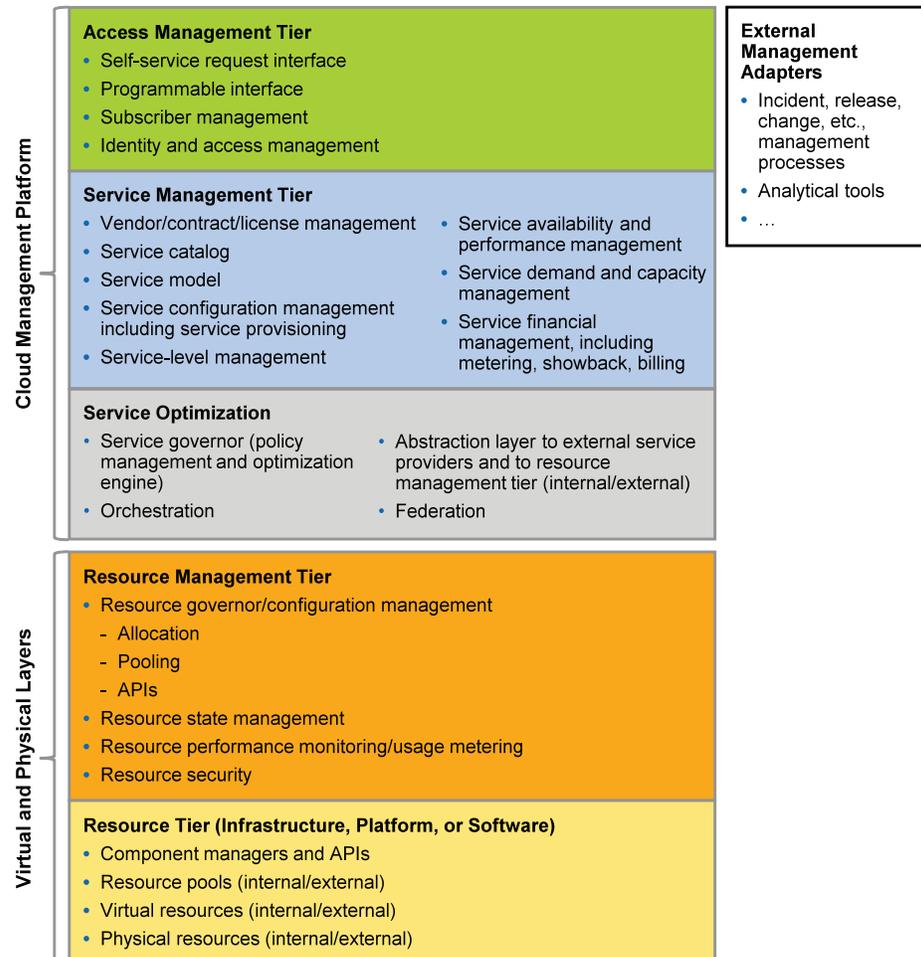
Access Management

A cloud service presents a service-based interface to end customers in which the details such as location and the underlying multitenant infrastructure of the implementation are hidden from them. In fact, it should be possible to change the underlying technical implementation without the user knowing it, as well as easily add new offerings to the interface itself. This interface might be a self-service portal or it could be an API (for software to access the service offerings). The interface is bidirectional.

Subscriber management (including identity and access management) is the other main component of the access management tier. This is the cloud administrative interface and enables the definition of policy in terms of who can order what services, as well as detailed implementation rules and workflows for provisioning and operation.

Service Management

Gartner expects the capabilities in the service management tier to see the greatest variability in cloud offerings, because different services require different styles of management, and cloud services often delegate the service management to the customer buying them (especially in IaaS). Cloud offerings will be detailed in service catalogs, which are self-service, and may provide information for private cloud offerings, as well as approved public cloud services. The service catalog must define specifically what is included and excluded from the service offering to avoid confusion and mixed expectations.

FIGURE 5   Cloud360 Approach

Cloud Management Platform

Access Management Tier
- Self-service request interface
- Programmable interface
- Subscriber management
- Identity and access management

Service Management Tier
- Vendor/contract/license management
- Service catalog
- Service model
- Service configuration management including service provisioning
- Service-level management
- Service availability and performance management
- Service demand and capacity management
- Service financial management, including metering, showback, billing

Service Optimization
- Service governor (policy management and optimization engine)
- Orchestration
- Abstraction layer to external service providers and to resource management tier (internal/external)
- Federation

Virtual and Physical Layers

Resource Management Tier
- Resource governor/configuration management
  - Allocation
  - Pooling
  - APIs
- Resource state management
- Resource performance monitoring/usage metering
- Resource security

Resource Tier (Infrastructure, Platform, or Software)
- Component managers and APIs
- Resource pools (internal/external)
- Virtual resources (internal/external)
- Physical resources (internal/external)

External Management Adapters
- Incident, release, change, etc., management processes
- Analytical tools
- …

Source: Cognizant

User-requested SLAs are defined at this level and are used to architect the delivered service, including the management services. The service management tier provides critical input into the service optimization tier. Simple cloud services might not offer much in terms of service optimization outside of following policies on the placement of services. A complex service might require dynamic and continuous service optimization — an RTI that leverages a more functional service governor.

Many IT organizations don't charge back and don't present service pricing to their customers.

Although we recommend that they do so, a cloud service should, at a minimum, record and present metered service usage, cost estimates or list prices (based on the service levels requested) to influence end-user behavior regarding business value versus cost of service. Thus, at a minimum, a cloud service should provide a form of cost transparency or showback.

Traditionally, service management has not determined where to run services, the level of resources to apply, and when to scale the resources up or down because legacy services were static. However, in a more dynamic cloud service

architecture, service management provides insight into SLAs, service performance and root cause analysis, such that Tier 3 service optimization, can supply resources in an optimal way to meet SLAs. It does this through policies, and is executed by the service optimization tier.

More fully featured offerings will provide service design tools, enabling cloud architects and administrators the ability to design, connect and deploy service bundles onto the underlying virtual infrastructure via drag-and-drop operations (which automatically create the service model). However, care should be taken to prevent the metadata models from impeding IT organizations from using the management tools of their choosing.

The service management tier includes:

- The service catalog, which displays available services to the user, ideally with pricing information.

- The service model, which describes the taxonomy of the service with associated policies.

- Service-level management, which defines the service-level options the user can choose.

- Service availability and performance management, which provide monitoring and correlation services and integration with the broader availability and performance portfolio, if required.

- Service demand management, which analyzes historical demand patterns as an input to the service optimization tier. Service capacity management works hand-in-hand with demand management, comparing demand and supply of resources over time, and recommending capacity increases on hybrid cloud resources.

- Service configuration management, which under the direction of the service optimization layer, executes configuration changes dynamically, as required, to provision and maintain the service.

- Service financial management, which includes license management, metering, chargeback, billing and reporting. It may also include vendor/contract management for contractual relationships with third parties (e.g., IaaS,

PaaS and SaaS). This could also be a separate function.

Many other important service management processes (and associated tools) are critical to service integrity, but reside on the outside of the cloud service. These include asset, incident, change and release management, as well as business analytics.

Service Optimization

The service optimization tier provides the runtime policy and optimization for the cloud service through the service governor. It may be simply implemented, such as only for initial provisioning placement of a service. However, even this can get complicated, with multiple resource pools spread across internal and external data centers. The service governor would perform intelligent service placement based on the policies and metadata, such as the location of the user, location of the data, mission-criticality of the service, capacity of the resource pool, or even the business unit associated with the service. Ongoing optimization may also be implemented through service governor policies, to match resources with service levels to enable the lowest cost of service delivery.

Service governor-like functionality may exist at many levels in the runtime architecture. Given the dependencies between services and resources, service governors must work together such that they don't impede each other's actions.

Contracted external service resource pools would be available for provisioning and optimization, based on established policies and through intelligence gleaned from the service management tier, particularly the service catalog and vendor/contract management. Moreover, we envision the service optimization tier to enable real-time external service sourcing or brokering as well, through a set of predefined policies and integration.

This tier may also require a transformation function to enable migrating from external to internal services or vice versa. For example, application development may occur on public cloud resources, but then be brought in-house for production deployment. If the development environment was based on a different infrastructure stack than what is used internally, the service governor would direct or orchestrate transformation services to

ensure it could migrate intact. Migration services are likely to require significant customization (to account for variations in the underlying resources), and testing would still be required to ensure that there are no incompatibilities as a result of the migration.

Resource Management

This tier includes the resource agent and controller infrastructure necessary to interface and orchestrate the actual physical and virtual resource assignments pursuant to the SLAs and policies. Unlike the service management tier, there is no level of abstraction shielding the administrator from knowledge regarding the real infrastructure that exists below. The CMP should enable real-time updates of physical resources added or subtracted from the system to talk to the virtual and physical resources, so that it can shift resources between resource pools to optimize the runtime of the environment. Orchestrating changes in physical resource configurations requires interoperability with the underlying devices and device managers themselves via their standard or proprietary APIs or interfaces.

For standardized services requested from a catalog, the service optimization layer deploys to the resource pool based on pre-established policy. Resource allocations would be performed via interfaces to private or public cloud resources, such as VMware's vCLoud or the Amazon Web Services APIs or via the cloud offering's lower-level resource governor-like capabilities. Availability features would likewise avail themselves of capabilities within the infrastructure or be provided by orchestrating spare resources to replace failed resources (as in a fabric-based infrastructure). Network driver technology may also be provided to enable an overlay network designed to meet specific isolation requirements for either security or performance reasons, should this also not be available from the underlying resources.

Service event, performance and usage data would be collected destined for the service management layer. There, it would be aggregated in terms of specific services to perform continuous service optimization via the service governor.

Resource Tier

This level represents the shared pool of resources. It consists of resource "virtualizers" — hypervisors, partitions, storage area networks (SANs), etc. — as well as the physical infrastructure fabric (servers, disk and network elements). Although it is technically not a requirement to have virtual machines as part of the IaaS cloud service implementation, it is through the resource sharing offered by this technology that we would normally expect to see it as a component of the cloud architecture to maximize shareability and its associated cost benefits. Higher-level virtualization (at the application, database management system, etc.) may also be deployed to enable dynamic service optimization.

The available virtualization technologies are relatively mature, and are a good foundation for a private cloud service. These technologies help make a cultural shift — the resources are abstracted from users, and become increasingly shared. Virtualization also introduces more-dynamic change, which requires fundamental changes to operational processes.

About Cognizant

Cognizant (NASDAQ: CTSH) is a leading provider of information technology, consulting, and business process outsourcing services, dedicated to helping the world's leading companies build stronger businesses. Headquartered in Teaneck, New Jersey (U.S.), Cognizant combines a passion for client satisfaction, technology innovation, deep industry and business process expertise, and a global, collaborative workforce that embodies the future of work. With over 50 delivery centers worldwide and approximately 150,400 employees as of September 30, 2012, Cognizant is a member of the NASDAQ-100, the S&P 500, the Forbes Global 2000, and the Fortune 500 and is ranked among the top performing and fastest growing companies in the world. Visit us online at www.cognizant.com or follow us on Twitter: Cognizant.