CLOUD CLOUT WITH OPEN APIS

WHAT YOU SHOULD ASK OF YOUR CLOUD PROVIDER

STRATEGIC WHITE PAPER

As cloud services become increasingly popular, more questions arise about the capabilities of cloud solutions. According to a recent Alcatel-Lucent study, performance, data security and ease-of-use are the top concerns among IT decision-makers considering a move to cloud services. The carrier cloud addresses these concerns, and open application programming interfaces (APIs) play a key role in meeting the requirements of cloud users. This paper explains the value of well designed cloud APIs, both for the data center and the cloud network. It also summarizes the critical issues cloud users and cloud providers need to consider as they look to take advantage of the cloud.

TABLE OF CONTENTS

Introduction / 1

Questions about cloud solutions / 1

Introducing open APIs / 2

Web APIs / 3

Applying APIs to the cloud / 3

APIs for the cloud network / 5

API management / 6

Recommendations and conclusions / 7

Abbreviations / 7

Resources / 8

Contact / 8

INTRODUCTION

The popularity of using cloud services for information and communications solutions is sweeping across industries and even public administrations. Every day, more consumer, business and government services are being moved into the cloud, making them accessible any place and from any device. As cloud services grow in popularity, the new technology puts existing enterprise computing, communications and storage infrastructures into question.

The cloud promises economies of scale but other benefits may be equally important. Services can be deployed within minutes or hours rather than the weeks and months that it takes to order and deploy dedicated servers. Resources can be scaled almost instantly to follow demand curves as customers adopt services and increase (or decrease) usage. Cloud application developers do not need to plan as far ahead and risk misallocation of capital.

While there is a high level of interest in cloud services from all parties, many enterprises and communications service providers are asking questions about the details of cloud solutions. They want to know how they can:

- Manage and control cloud solutions
- Assure key performance and regulatory parameters
- Integrate cloud services into their IT environment

To better understand the questions and concerns that are developing around cloud services, Alcatel-Lucent undertook an in-depth study of perceptions and attitudes toward the cloud among 3886 IT decision makers in seven countries. According to this study, concerns about performance, data security and ease of use are top-of-mind for the decision makers.

In the following section, we take a closer look at the main questions behind the decision makers' concerns. We then explore how cloud application programming interfaces (APIs), as part of the new carrier cloud, help to address these concerns.

QUESTIONS ABOUT CLOUD SOLUTIONS

Cloud users expect the same, or even higher, performance levels when applications are moved into the cloud. This certainly applies to latency-critical, real-time applications such as virtual desktops, online gaming and communications and conferencing. Communications service providers are on a trajectory to become nimble cloud telcos with much of their service infrastructure — IP multimedia subsystems (IMSs), operational support systems (OSSs) and business support systems (BSSs) — moving into the cloud. Alcatel-Lucent is even taking wireless access functionality out of physical base stations and putting it into the cloud.

Many web-based client-server applications need to provide very short response times. This is a consequence of a trend where browser interfaces replace fat-client desktop applications. In a virtualized cloud environment, these performance criteria cannot always be directly observed and influenced. As a result, an "order and pray" strategy where the cloud user simply hopes the network will provide adequate service levels is no longer an option. Cloud users need assurances that critical latency, bandwidth, availability and other parameters will meet their requirements and that these parameters can be monitored. Moreover, what good is a flexible cloud data center that can be rapidly adapted to match increasing traffic if the corresponding network links take days or weeks to reconfigure?

Cloud technology promises rapid scalability of services. But how is this actually accomplished? How do cloud users know that allocated storage, computing and network resources are running out of capacity or, conversely, are underutilized? What is the service utilization profile over the course of a day, a month or a year? Do we need staff dedicated to monitor and adapt cloud installations or can this process be automated?

For business and regulatory reasons, enterprises need to maintain a certain level of control over their cloud solutions. Compliance with privacy rules and other legal requirements mandates that cloud users supervise and monitor cloud solutions. In many cases, cloud users need the ability to allocate cloud resources that are both close to end users and within particular national jurisdictions.

Cloud solutions offer basic infrastructure, such as virtual machines and storage for data objects ("blobs"), as well as a variety of higher-level services, including Software as a Service (SaaS) and Communications as a Service (CaaS). These services need to be made accessible in a coherent universal and secure way to the applications using them. It is no longer acceptable that each service uses its own protocol. This doesn't mean that everything is standardized. But it does mean that developers — based on their experiences with Internet services — want to access these capabilities with minimal learning and easy access to developer tools.

These questions apply to public, private and hybrid clouds. After all, private cloud technology is very similar to public cloud technology. As with public clouds, private clouds are not necessarily in the same geographic area as their users, but are situated in one or more central locations that are reachable only through wide area networks.

Cloud service providers want to address these questions while differentiating their offerings. The simplicity of the application development and management process is a key criterion for an enhanced service experience. As a result, cloud providers should give their customers the right level of control for all levels of cloud functionality in a coherent form and through a single portal. APIs are central to achieving these objectives.

INTRODUCING OPEN APIS

In the past, when enterprises needed a new service or wanted to modify their service parameters, they called their service provider or sent a service order by fax. In the cloud economy, this mode of working is no longer sustainable.

The cloud provides extensive self-service capabilities, including web-based management dashboards with graphical user interfaces (GUIs) and APIs. The former simplify individual configuration tasks, but for more complex management activities the latter are needed. If we want to automate a process that is too tedious to execute manually, build a specialized GUI or integrate cloud capabilities into other applications, then APIs come into play. Software applications can invoke APIs to access the functionality of the services. Using an e-mail example, an application can send personalized messages to hundreds of recipients. Or a business social media application can use APIs to automatically set up an audio or video bridge and connect the members of a team.

For cloud users to take full advantage of APIs, they must be open; that is, they need to be well documented and the documentation needs to be easily available. In addition, the API syntax and semantics must remain stable for a reasonable amount of time.

WEB APIS

Early APIs could only be accessed directly from the system that was exposing them. That is, the application software had to be installed on the same computer or network element that implemented the APIs. Operating systems have always used APIs (system calls) to create files, execute programs and for many other purposes. However, with the advent of the web, and specifically Web 2.0, services and applications started to be distributed over the network. Application developers could take advantage of services from providers anywhere on the web. For example, a travel application could use a mapping service from one provider and an advertising service from another provider, combining them into a mash-up application running on yet a different computer.

A first style of web-based APIs, called Simple Object Access Protocol (SOAP) web services, was developed around the year 2000. More recently, a radically simplified style of web APIs, called Representational State Transfer (REST), has become prevalent in the Internet community and has been quickly adopted by many services, including Amazon's Simple Storage Service (S3).

APPLYING APIS TO THE CLOUD

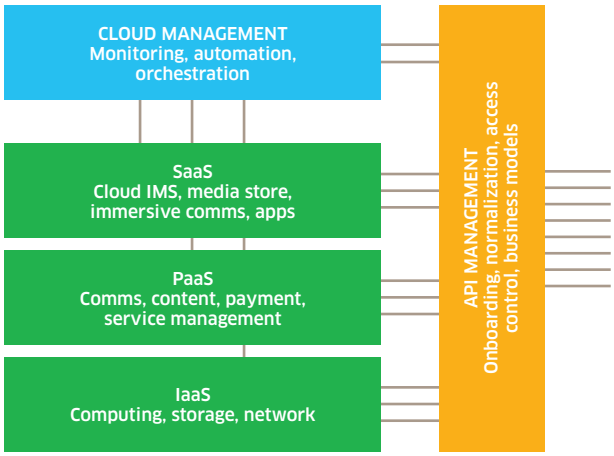
A wide variety of management tasks applies to cloud applications. For example, when a cloud application is first installed, the different types of cloud nodes need to be created, linked to each other and distributed over primary and backup data centers. Secure network links must be installed and configured. During their lifetime, applications need to be scaled according to demand and re-allocated across a set of data centers to provide the best quality of experience for the application users. After that, when a new cloud application software release must be installed, the task can be automated by allocating new virtual machines for the release, switching the load balancers over to the new machines, then de-allocating the old virtual machines. GUIs and APIs can be combined to help execute these tasks in an optimal way. Easy-to-use GUIs simplify tasks that require human intervention while APIs are essential for automating clouds and for integrating cloud capabilities into enterprise business processes. In a well defined cloud system, GUIs use these same APIs to execute the user actions on the cloud infrastructure.

Many cloud providers already offer some form of APIs. Let's take a look at the types of APIs that apply to cloud offerings. Typically, three levels of cloud services are distinguished:

- Infrastructure as a Service (IaaS)
- Platform as a Service (PaaS)
- Software as a Service (SaaS)

Each of these levels comes with its own parameters and capabilities that need to be configured, monitored and made available to applications. Figure 1 summarizes the main cloud services delivered at each level.

Figure 1. Each level of cloud services comes with its own parameters and capabilities



IaaS is the most basic cloud offering. Typically, the concept of infrastructure encompasses virtual machines offering computing capacity, the ability to host a user-selected operating system image and storage capabilities, such as storing named unstructured data blocks. More recently, IaaS offerings have included the network resources to connect any of the resources to each other and to their users. IaaS providers may also offer configurable load balancers, firewalls and virtualized IP addresses and domain names that can be assigned to any virtual machine. IaaS APIs allow cloud users to allocate, de-allocate, configure and monitor IaaS resources, giving cloud users some control over the geographical location of the resources.

PaaS is a higher level offering providing programming languages, software and API libraries and tools. With PaaS, the cloud user does not need to manage and control the underlying operating systems, servers and storage. As application traffic increases or decreases, the platform often scales automatically without requiring the user to explicitly allocate resources.

PaaS can be viewed as the next-generation service delivery platform (SDP). By some accounts, that would be SDP 3.0. While SDP 2.0 introduced a coherent service-oriented architecture (SOA), the third generation is characterized by lightweight APIs and a separation between hardware and software through a virtualized cloud architecture.

Finally, SaaS is a cloud offering where the cloud provider runs and administers complete applications such as customer relationship management (CRM) applications, computer games or conferencing and communications applications. The applications are typically accessed through clients, including web browsers. The cloud user does not need to be aware of the underlying cloud computing and storage resources. However, the cloud user is concerned about the semantics of the application software and requires application-level APIs.

For example, a cloud-based communications application may provide APIs for routing calls depending on caller, callee or time of day. A communications application on a PC or a mobile device may use such APIs to initiate, display and control the calls of particular end users. On the backend, SaaS applications may also use other cloud services, such as identity, payment, content or advertising services. As a result, these applications become mash-ups of multiple cloud capabilities interconnected through well-defined APIs.

These APIs form part of a cloud orchestration system that facilitates the programmatic provisioning and de-provisioning of resources necessary for a cloud infrastructure. Orchestration hides the nuances of the underlying virtualization tier, exposing them only to the on-demand provisioning web interface. With the cloud orchestration system, the cloud can be integrated into a risk management and compliance architecture. In particular, if personal data are handled, there are extensive rules — the European Union Data Protection Directive is one example — for ensuring the privacy of these data. Security applies to the servers and to the network environment. Here, links can be protected with link encryption or dedicated links can be allocated.

APIS FOR THE CLOUD NETWORK

A cloud-based approach separates the application from the resources it needs to execute. Applying this principle to wide area networks creates new challenges. With a High Leverage Network™, many service providers own a rich set of network capabilities that can be monetized in differentiated cloud offerings.

Specific network resources, such as IP virtual private networks (VPNs), Layer 2 VPNs and even optical wavelength paths can be allocated to segregate cloud traffic from other traffic. Links and VPNs can be encrypted to keep data private at all times. With this approach, applications can be better protected from network attacks, such as distributed denial of service attacks. Dedicated network resources are essential for private clouds and for virtual private clouds that are hosted — that is, those with specific resources allocated at public data centers. Dedicated (virtual) network resources are also the foundation for performance objectives with dimensioned capacities, latencies and application-aware packet handling parameters.

While adding network resources to a cloud offering creates a level of quality that is essential for a variety of cloud applications, it also raises issues that must be addressed. Due to the virtual and dynamic nature of the cloud, computing and storage resources may migrate to different server farms or even different data centers. In addition, depending on demand, application traffic may change substantially. This can leave network connections under- or over–dimensioned, or even stranded.

Therefore, mechanisms are needed to adapt network resources to the evolving cloud application. Network APIs provide a fundamental layer to build mechanisms that orchestrate all types of resources, including the communications links, in a cloud solution. For performance-critical applications, availability of network resources may determine which computing and storage resources should be allocated from among several alternative data centers. Here we need efficient algorithms that can query the network status and topology using APIs, such as the Internet Engineering Task Force (IETF) Application-Layer Traffic Optimization (ALTO) protocol, to calculate optimal resource allocation strategies.

Clearly, virtualizing the network and exposing network capabilities with APIs brings risks such as configuration errors. It could even create vulnerabilities to malicious attacks. Moreover, a vast array of network technologies has been invented, and many of them have found their way into networking equipment. Cloud providers are doing themselves and their customers a favor if they focus on simple network abstractions that cloud application developers can easily grasp and incorporate into their solutions. Opening carefully designed network capabilities to the cloud user community for self-service significantly reduces cost of ownership and helps to make clouds as agile as users demand.

API MANAGEMENT

We have seen that APIs are a key ingredient to any cloud solution and that exposing cloud functionality through APIs needs to be carefully planned and designed. An API management system can help the cloud provider build a coherent interface toward cloud users at all levels.

The API management system normalizes APIs from different subsystems such as computing, storage and network, and allows building of custom APIs that abstract the details of the subsystems. Applications and application developers need to register as legitimate API users and agree to the provider's terms and conditions. They will then receive an API key for authentication in all API calls. They will also gain access to developer tools, such as sample code, documentation and a sandbox environment for testing their applications.

The API management system controls access to the cloud APIs based on the API keys and applies configurable policies. Statistics and reporting tools allow cloud providers to monitor and supervise API usage according to different criteria. They also serve as a resource for automatic or manual cloud capacity planning. The API management system generates detailed usage records that can form the basis for flexible business models, such as pay-per-use, flat-rate or revenue-share.

Cloud computing is a new technology and the industry has not yet adopted a unique standard for such APIs. This hampers the development of multi-cloud applications and the migration of cloud applications from one vendor to another. An API management system can help cloud users by providing a layer on top of proprietary cloud APIs and by normalizing these APIs to provide a common view for all cloud applications.

The API management system is typically a fully virtualized cloud-based application with a low entry threshold for cloud providers and the ability to easily scale as demand grows. Moreover, cloud providers can offer API management as a service to customers who want to offer their own application-level APIs.

RECOMMENDATIONS AND CONCLUSIONS

With cloud solutions, there is a perception that business-critical resources are no longer under the control of the application developer and user. With a carrier cloud offering that supports open APIs and API management systems, cloud users can regain control at the right level. There is no need to know the exact physical server blade and rack position where an application is being run. However, cloud users should make sure that:

- The cloud parameters that affect the performance and security of an application are readily available through both GUIs and APIs.
- APIs are available for computing and storage resources as well as for the cloud network.
- APIs are open; that is, they are well documented and stable.
- APIs from different subsystems are normalized and access to APIs is securely managed.

A carrier cloud includes the API management and cloud orchestration systems needed to build a clear API strategy. It delivers the cloud clout to meet efficiency and security requirements and to reduce total cost of ownership.

ABBREVIATIONS

ALTO	Application-Layer Traffic Optimization
API	application programming interface
BSS	business support system
CaaS	Communications as a Service
CRM	customer relationship management
GUI	graphical user interface
IaaS	Infrastructure as a Service
IETF	Internet Engineering Task Force
IMS	IP multimedia subsystem
IT	information technology
OSS	operations support system
PaaS	Platform as a Service
REST	Representational State Transfer
S3	Simple Storage Service
SaaS	Software as a Service
SDP	service delivery platform
SOA	service-oriented architecture
SOAP	Simple Object Access Protocol
VPN	virtual private network

RESOURCES

The Network-Enabled Service Provider Cloud

The Age of the Application-Aware Network

Cracking the API Code

Get more Partners with Less Effort

Application Enablement

Alcatel-Lucent cloud solutions

The NIST Definition of Cloud Computing

CONTACT

Andreas C. Lemke
Application Enablement – Solutions Marketing
Andreas.Lemke@alcatel-lucent.com

Alcatel·Lucent