

# WHY ALL CLOUDS ARE NOT CREATED EQUAL

ENTERPRISE CLOUD, PUBLIC CLOUD,  
CARRIER CLOUD

STRATEGIC WHITE PAPER

Cloud computing technology brings an unprecedented level of independence and liberation in deploying applications. Applications are no longer tied to dedicated hardware, yet clouds vary significantly in their capabilities and their cost. This paper helps readers understand the differences between enterprise clouds and public clouds, and explains the advances available in carrier clouds.

# TABLE OF CONTENTS

The Cloud Declaration of Independence / 1

Clouds Created Different / 2

Public Clouds / 3

Enterprise Clouds / 4

Carrier Cloud / 5

Conclusions / 6

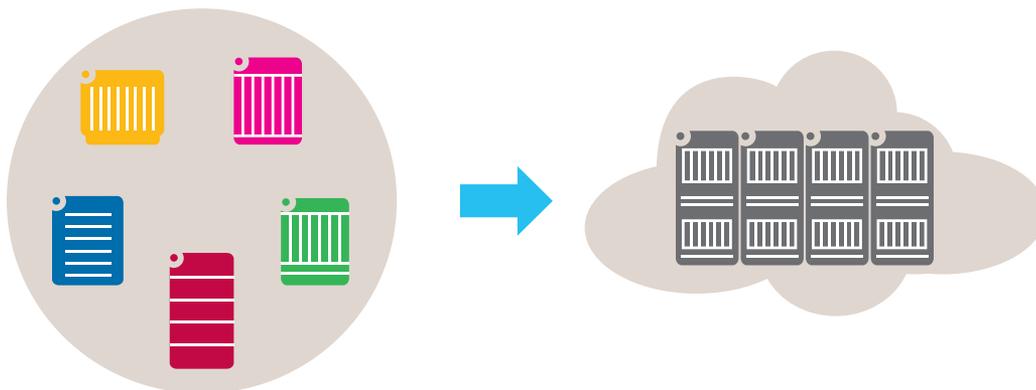
# THE CLOUD DECLARATION OF INDEPENDENCE

Cloud computing is galvanizing the information and communication industry. Enterprises expect dramatically lower cost and higher agility in their IT operations. Communication service providers (CSPs) are looking for new cloud-based architectures that allow them to become a new type of cloud provider and to virtualize their own network and IT infrastructure, which would enable the rapid introduction of services with new types of business models.

These cloud models differ from classical information and communication technology (ICT) architectures, where each application requires dedicated resources, and particular parts of the application (compute tier, database tier) are allocated to specific blades or servers.

Software upgrades require complex in-place upgrade procedures with difficulty to limit service interruptions. The dependency on particular server hardware threatens long-term business-critical applications when hardware becomes end-of-life and prevents these applications from benefiting from the ongoing technology evolution.

**Figure 1. Cloud computing overcomes the complexity and cost associated with diverse hardware dedicated to specific applications. The cloud is based on highly standardized compute and storage and network nodes.**



Cloud computing technology can overcome many of these issues. With cloud computing, applications become essentially independent from specific physical equipment [Figure 1]. In recent years, two technologies have been developed that are at the foundation of cloud computing: virtualization and high-speed networking. With virtualization, a physical computer is separated into multiple virtual machines, giving each application its own environment and operating system as if it was the sole user of the computer. Moreover, virtual machines can be migrated at run time from one physical machine to another, or even from one data center to another data center. This way, cloud applications can follow their users, optimizing service experience and resource utilization. The rapid advance of high-speed network technology is a second essential enabler for cloud computing. Due to these high-speed networks, application workloads can be placed remotely in cost-effective data centers of a cloud provider without causing performance or latency issues.

In this way, cloud computing is liberating application providers from physical hardware and geographical constraints. For many, this flexibility will be of higher value than savings from better utilization of physical data-center resources.

# CLOUDS CREATED DIFFERENT

The Declaration of Independence states that “all men are created equal”, but not so for clouds. Clouds are created different to support a spectrum of applications with very different characteristics and architectures. Applications require different levels of availability and response latencies. For example, a large simulation can use up available cloud resources but will not have stringent response time requirements. To the contrary, the success or failure of a virtual desktop application depends on such millisecond response times, and unavailability of compute resources during cloud busy hours would not be tolerable.

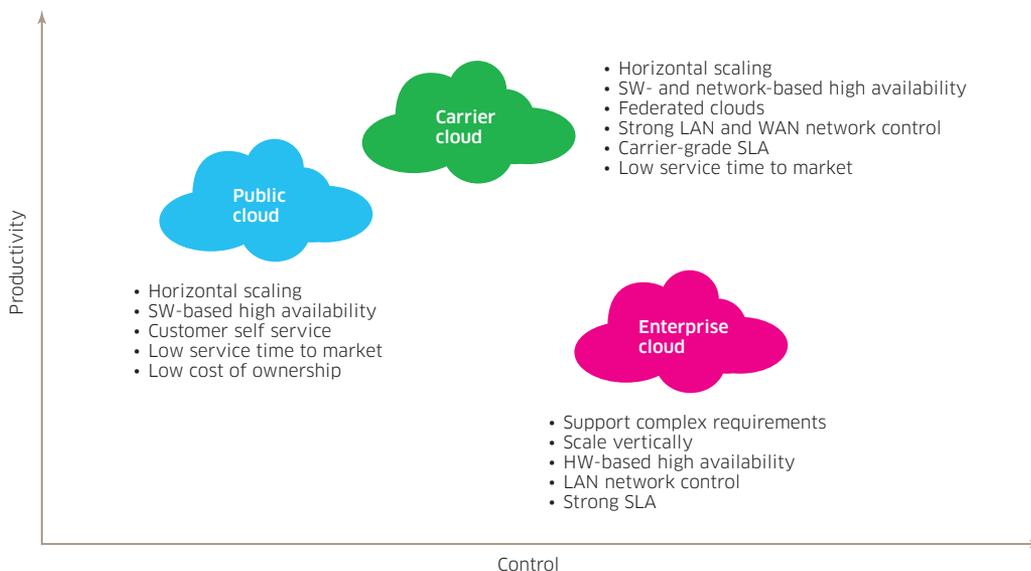
Some recent applications are architected in a way that they can easily be ported to a cloud architecture. They can be scaled horizontally by adding or removing servers and storage, and the software can recognize this pool of servers. The software uses only generic features of the hardware architecture, often an Intel x86 architecture. The operating system should be proven to work under the selected hypervisor.

Most legacy applications have been designed with a much more static software and network infrastructure in mind. Developers have exploited operating systems (or even hardware options) to optimize their applications, and this makes the journey to the cloud not easy. Re-architecting such applications is often not possible or prohibitively expensive, as underlying software and hardware is end-of-life or the teams with the right knowledge of the application simply no longer exist. Moving such applications into a cloud is particularly challenging and requires a high degree of versatility of the cloud infrastructure. As demand evolves over time, the applications cannot easily be scaled. When the capacity limit of an installation is reached, a more powerful server must be purchased and the application is migrated to the new hardware (vertical scaling).

Beyond the technical requirements, different industries need to follow specific rules and regulations (e.g., about security, privacy, certification, and auditing) that raise new questions in a cloud-based deployment.

For these reasons, cloud providers have developed a range of offerings with different levels of productivity and control to match their customer expectations [Figure 2]. Mass-market public clouds are highly standardized and highly productive. Enterprise clouds give cloud users a great deal of control but cannot match public clouds in productivity. An enterprise cloud supporting specific requirements can cost ten times as much as a standardized public cloud. Carrier clouds are similar to public clouds but provide strong network capabilities and service level agreements.

Figure 2. Different types of cloud provide various levels of productivity and control.



## PUBLIC CLOUDS

The massive public clouds (from the likes of Amazon Web Services and Rackspace) have been highly publicized. Google has been reported to be running 900,000 servers as of 2011. Public cloud users can spin up hundreds of virtual machines within minutes to match demand spikes. These clouds are optimized to deliver the highest compute capacity for the money but not the highest compute power per individual server. That is, server designs are often based on low-cost consumer PC technologies. Data centers can be run at higher temperatures, saving energy and expense for cooling with minimally higher hardware-failure rates. Public clouds are designed for failure. High availability is delivered not using hardware but through software architectures that take advantage of multiple servers where an individual failure is not critical. The data center has self-healing capabilities to work around such failed components, and failed resources are often not replaced for a long time until it becomes economical to send a worker for replacement.

On the software side, public clouds tend to use open source software such as OpenStack, CloudStack, and the Xen and KVM hypervisors. They often standardize on a specific hypervisor and this hypervisor may be less feature-rich and support a smaller variety of guest operating systems than more expensive alternatives. Many public cloud providers adopt open source cloud software to benefit from and contribute to a vibrant developer ecosystem, to avoid being locked in to expensive proprietary products and provide to their customers the option to choose and federate the services from multiple cloud providers for extra reliability and flexibility.

In addition to the compute and storage services, public clouds also offer virtual appliances to manage traffic and secure applications. Public clouds offer generic load-balancing services, firewalls, threat management systems, or Secure Socket Layer accelerators. The services from public cloud providers are typically limited in that a generic set of virtual appliances is offered.

Operating systems, database systems, web servers, and programming language runtimes need to be maintained, for example, to eliminate vulnerabilities as they become known. With public clouds, patching operating systems and other software platforms is typically an activity that cloud users need to take on themselves.

## ENTERPRISE CLOUDS

These clouds address a different market than public clouds. The focus here is on new or legacy applications with special architectural, performance, availability, and security requirements. To consolidate enterprise IT infrastructures, cloud solutions are needed that can support legacy applications running on older hardware and software and that are difficult or even impossible to modify and adapt to a different environment. Enterprise clouds are designed for fork lifting applications into the cloud. Enterprise clouds can also support applications that need specific storage solutions with Storage Area Networks (SAN) or Redundant Arrays of Inexpensive Disks (RAID).

Enterprise clouds can be private clouds installed on customer premises, or virtual private clouds hosted on dedicated equipment in the service provider data centers, or a combination of both variants.

Because there are often sensitive data, CIOs need to make sure that detailed company security policies and data-center best practices, as well as government regulations, are implemented (such as with ITIL, SAS 70, ISO, the U.S. Sarbanes Oxley Act or the European Data Protection Directive). The latter, for example, requires that cloud users control the location of customer data and keep it within the boundaries of the European Union. The U.S. Health Insurance Portability and Accountability Act (HIPAA) requires that all user data be encrypted, both when being transmitted and while stored. While applications designed for private data centers rarely store their data in encrypted form, cloud users may wish to do it when the data is moved to the cloud.

The cloud is fundamentally multi-tenant, that is, servers and storage devices are shared by multiple customer organizations. While data leaks due to virtualization are extremely rare, some businesses have chosen private or virtual private cloud architectures that rely on various levels of dedicated hardware. For example, enterprise clouds offer dedicated servers that are guaranteed not to run virtual machines from other customers.

Many applications are not prepared to take advantage of cloud-based scaling mechanisms. When such applications need to support more transactions and traffic, they will be installed on more powerful servers with more computing power and storage. As these applications are moved to the cloud, they need powerful servers where low-cost, lower performance servers may not be an option. Moreover, high availability is achieved by choosing highly reliable computers, disks, and switches. Even, if the original application is designed with redundant servers, fail over is assumed to be a rare event and the fail over times may be longer.

Enterprise applications often rely on highly structured network setups with multiple public or private subnets, network address translation, and special traffic routing requirements. In addition, enterprise cloud users may wish to build hybrid cloud solutions where certain highly sensitive data and processing remains in the private data center. Other tasks are run at an external cloud data center. For economical reasons, CIOs often

run the bulk of the workload on the private cloud and then burst demand peaks into a third-party cloud, thus reducing their own spare capacity and increasing the server utilization. These bursting processes need to be transparent to the applications. Workloads that run in the external cloud will remain unchanged, with the same IP addresses as in the private data center, and they will continue to interwork with servers in the private cloud. Communication links and cloud-storage resources are secured and encrypted.

Enterprise clouds also offer a great deal more options for service level agreements (SLAs). While public cloud SLAs provide quite reasonable service availability guarantees of 99.95% or even higher, definitions of what constitutes an outage are often strongly skewed in favor of the cloud provider. With enterprise clouds, SLAs go beyond bare uptime levels. SLAs can include criteria for actual availability of CPU cycles, LAN and WAN network throughput and latency, storage contention, spare resource availability for scalability, and incident and problem management. And when an SLA violation occurs, remediation of lost business revenue is in high demand. This has been proven by an Alcatel-Lucent study with more than 3800 IT decision makers in seven countries (Alcatel-Lucent 2011 Global Cloud ITDM study).

## CARRIER CLOUD

According to the same study, better performance is the top demand to improve adoption of the cloud. Current public clouds are highly centralized and users access their cloud resources via the Internet. This results in unpredictable network latencies. Public cloud providers have recognized the need to enhance their network capabilities. They are beginning to interconnect their data centers with dedicated high-speed optical links or are linking directly into collocation centers.

CSPs take advantage of their network assets to develop carrier clouds, thereby combining some of the best assets from both enterprise and public clouds and adding strong network capabilities. CSPs can offer carrier cloud services directly to their customers and also use them internally to transform their own service platforms and operations toward the cloud model with enhanced cost structure and agility.

Carrier clouds inherit from public clouds the emphasis on a competitive cost base and industrialized operation and management, including self-service. With their extensive network assets, CSPs can offer carrier clouds with stronger, carrier-grade SLAs. Forty-six percent of public cloud users cite annoying or intolerable latency of 150 to 500 milliseconds. Service providers have many points of presence throughout their geographical coverage area. With these assets, service providers can carefully establish more distributed cloud locations where their customers are and offer a level of performance that is not attainable from the highly centralized public data centers. Distributed clouds reduce backhaul networking cost and significantly reduce the impact of power failures that have repeatedly afflicted the centralized public clouds. Moreover, a service-aware network can handle cloud traffic in a much more differentiated way. Real-time traffic can be marked and queued differently than traffic with less strict latency requirements. Network resources can be allocated to specific cloud services to make sure that each service receives the bandwidth it needs, and security can be enhanced by sending the traffic over separate network links where attackers cannot normally reach via the public Internet. As an additional security measure, hardware encryption of network links secures all traffic from attacks while the data is traveling from site to site without introducing noticeable delay.

The carrier cloud is open. This allows service providers to expose web APIs to trusted in-house or third-party developers. These developers can enhance cloud services offerings with network intelligence or develop new cloud services that are customized for specific enterprise needs or industry verticals. With more creative minds contributing to cloud services, innovation accelerates. Services providers can meet customer demands for more advanced cloud offerings faster.

## CONCLUSIONS

Cloud technologies liberate service providers and enterprises from their dependency on specific physical computing equipment. This creates an unprecedented level of agility, allowing cloud users to quickly set up and subscribe to new services and adapt to floating demand. With a good mix of high-priority and low-priority services, compute, storage, and network resources can be utilized to a much higher degree than is possible with dedicated resources — and this without sacrificing performance and security.

However, not all cloud applications and cloud solutions are created equal. Cloud users should understand the different strategies of public and enterprise cloud providers. Legacy applications can be moved into the cloud — into the enterprise cloud with a certain amount of effort. Very rapidly, however, applications are developing the cloud “genes” that make them able to run on public and carrier clouds. With that, the promise of the cloud — high level of performance, flexibility, and security along with significant cost reductions — can become reality for a wide majority of applications.

Based on their powerful networks, deep experience in large scale service delivery, and extensive customer relationships, communication services providers have an excellent opportunity to offer carrier clouds and bring value to this market that cannot easily be matched by pure cloud providers. In several markets globally, CSPs have already gained the status of the most trusted cloud provider.

Further reading:

- [The Carrier Cloud: Strategic Whitepaper](#)
- [Soaring into the Cloud: Understanding the Market Opportunity for Cloud Services](#)
- [Alcatel-Lucent 2011 Global Cloud ITDM study](#)
- [Cloud Clout with Open APIs](#)