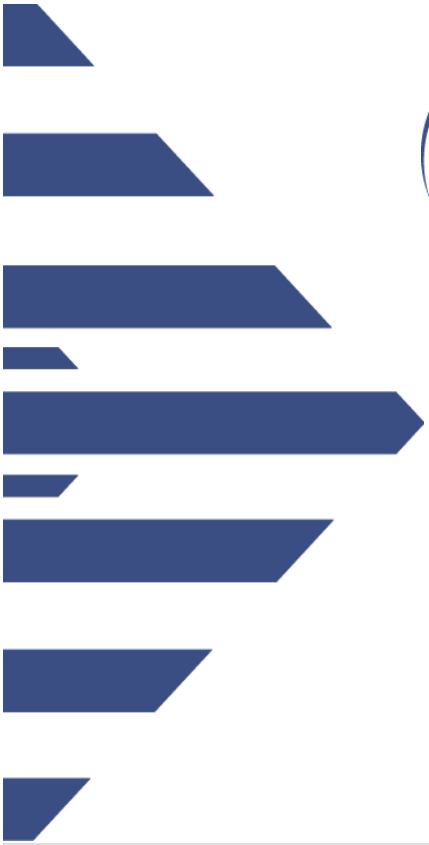
What is Cloud Computing?

Table of Contents



CERT

Cloud Computing Security

Software Engineering Institute | Carnegie Mellon

**001 Computing Security.pdf

Dennis Allen: And this module, we'll talk about Cloud Computing Security.

Overview

What Is Cloud Computing?

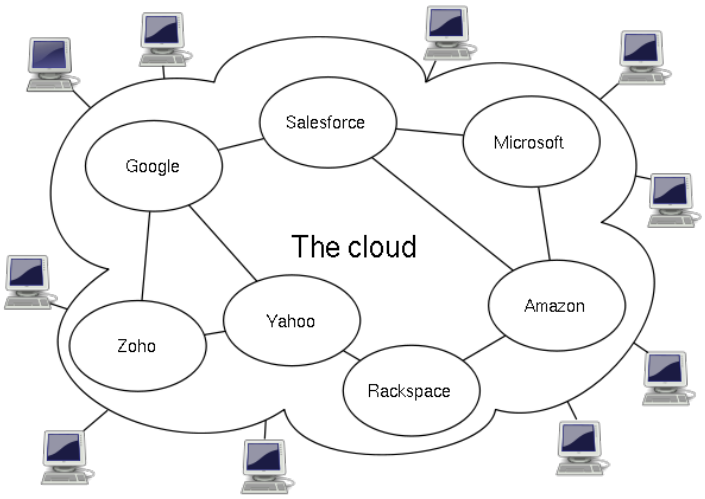
- Strengths
- Weaknesses

Technical Risks

Operational Risks

Mitigation Strategies

DISA Cloud Solutions



**003 So we'll talk a little bit about computing is. We're not going to get into a lot of the benefits and whatnot. We'll maybe cover different aspects of cloud computing as we move through the presentation, a little bit about its strengths, as I was mentioning, and certainly a lot about its weaknesses, and more than its weaknesses, things that you need to consider, things that are important for you to understand the risks involved, and how we can mitigate some of those risks, and specifically technical risks, operational risks, again, some mitigation strategies, and then, finally we'll touch upon a little bit about DISA, the Defense Information Systems Agency, some of the cloud computing

solutions that they have available to fit the government needs.

What Is Cloud Computing? -1

What Is Cloud Computing? -1

Cloud computing enables remote access, primarily through the Internet, to shared resources; (e.g., networks, servers, storage, applications, and services) typically being maintained by a third-party (the cloud provider).

Resources are often shared with other cloud provider customers.

Cloud provider infrastructures are normally virtualized and the system can usually be dynamically scaled to the customers needs.

Resources are normally allocated on a pay-per-use model.

CERT | Software Engineering Institute | Carnegie Mellon 4

**004 So what is cloud computing? Well the term, cloud computing is derived, quite honestly, very-- if you've ever done any kind of network documentation, or if you've seen a diagram of a network, quite often the internet is represented as some sort of cloud. So to be honest with you, a lot of the whole cloud computing terminology comes from just, sort of, that basis of looking at the internet as a cloud.

So how are these services being presented to the user? Can you access these applications, or access your

infrastructure over the internet? It's somehow outsourced, and it's out there in the internet somewhere, or in the cloud.

And we'll talk about the different types of cloud computing here in a second, and the different kinds of services. But essentially, there's something out there being managed by some cloud provider, some third party provider that's providing these types of services for you. We'll talk about infrastructure services, application services, etcetera.

The idea that they're shared, often shared with other customers, certainly comes into play. You can-- there are certainly options for private clouds, in which you would be managing your own infrastructure. The big thing, really, the difference between, you know, a classic computing thing, versus a cloud computing thing, if you will, is the virtualization aspects of it. So we're used to the collocation, the managing your own services, whether they're physically onsite at some collocation facility, but the whole idea with a lot of these things, with cloud computing, comes into the whole rapidly being able to deploy new machines, new infrastructure, new applications, taking advantage of certain virtual technology.

Is that always the case? Depends, there's lots of different variations of cloud computing, and what actually a cloud is, what a grid is, and how all those things play together, but for what we're talking about, we're generally talking about some sort of virtualized infrastructure that's out there that you can leverage immediately,

or close to immediately through some sort of service provider.

The pay-per-use model, in a traditional collocation facility, you're usually paying on a monthly basis. You have, like, a flat fee for what you're paying to have a server. You're paying for space, per rack unit, amount of bandwidth, perhaps, just, you know, for 1Mb bandwidth, it's x amount of dollars. In a cloud computing model, most of these services, like, Amazon, GoGrid, Rackspace, actually has collocation, and cloud computing modules.

You're paying more on the pay-per-use model. So it's how much data am I actually using on disk, how much CP utilization am I averaging on my processors that I'm using, what kind of bandwidth, in terms of people accessing, pulling down data, versus uploading data, so the model changes a little bit, and that allows people to get into that environment a lot easier. There's no capital expenditure, there's no need to come up with upfront costs for infrastructure to build out servers and put together networking equipment and have the internet connection, you can simply log on, the accounts are pretty much free, because you're going to log into a resource, provision an instance, and then based off of the type of instance it is, there may be licensing requirements if it's a Windows instance, etcetera, and then there will be various bandwidth or utilization pay-per-use models.

So that's from an infrastructure standpoint, but from a software standpoint, it's a little bit different.

What Is Cloud Computing? -2

Three main cloud computing service models

Software as a Service (SaaS)	Platform as a Service (PaaS)	Infrastructure as a Service (IaaS)
Customer uses provider's applications over a network • Google Apps • Microsoft Hotmail • IBM LotusLive • SalesForce.com • Zoho	Customer deploys their own applications to a cloud •Microsoft Windows Azure •Google AppEngine •LAMP	Customer rents processing, storage, network capacity, and other basic computing resources •Amazon EC2 •Rackspace •GoGrid

**006 There are some free things, when we talk about software as a service, this is the ability to actually log onto a website and leverage an application.

Salesforce.com has a few different applications in regards to contact management and CRM, Something like Hotmail or even Yahoo! Mail can be considered software as a service. It's providing that e-mail capability to you.

Google Apps, Google Docs, we mentioned Google code in our threats module, in terms of people accessing or leveraging a free resource to store malicious code in some cases. But it's-- software as a service is essentially, let

me go out there, find an application that suits my needs and pay for it on a per-use basis, and that one might be based off a per-account, or utilization of some other--

Has anybody heard of Zoho? All right, well you can do some different things with Zoho, too. Think of it as web based project management. It will also do document management, some other things, too, but if you were looking for a free or pay-per-use model for your project management, you can do some of those things in there, project management, change management, you can use that application.

The platform as a service, this is a little bit-- we'll come back to that one. We'll start with infrastructure as a service. We'll come to that one next.

Infrastructure as a service, now we're looking to actually provision our machines from scratch. We can log into a web interface, we can select a template, say, "Hey, I want a Linux machine, I want a Windows machine. I can start that up, and get just a vanilla operating system. And once I get that vanilla operating system, I can configure it as I need to. I may have an IP address automatically assigned, or I might be provided with the IP addresses and I might have to configure it, depending on your environment. But essentially it's bare bones operating system in capabilities, and then you can build from there. Amazon EC2, they're-- Elastic Compute Cloud, I believe is what it's called, provides you the computing resources to do something like that. They also have an

S-3, a Simple Shared Storage, where you can actually-- or storage service, I should say. Let me make sure I get the right terminology here. I'm always messing up my acronyms with these things. But either way, you have a storage service that's available with Amazon as well.

Rackspace has, like I mentioned, the collocation facility, the traditional, build your servers up for you, or you can provision things in a virtual environment. So if you want to take something in between with that, now I want to leverage an existing platform, and I have a platform as a service, you can take the middle ground there with a platform as a service option.

Does anybody know what that LAMP stands for? It's sort of---

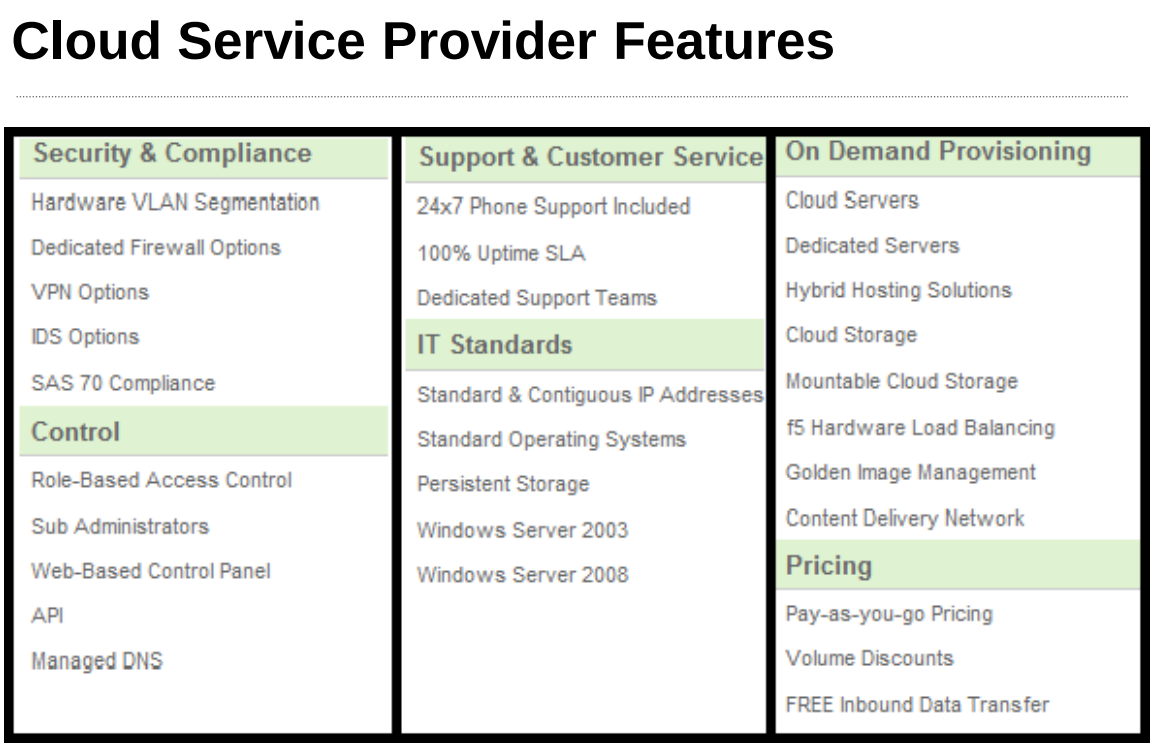
Student: Linux, Apache, MySQL, PHP?

Dennis Allen: Yep, you got it. Linux, Apache, MySQL, PHP, So essentially, it's sort of a core configuration that you can build upon. If you wanted to have a web based application that uses a database, you can get a generic LAMP instance set up, and then you just configure it with your own code or something like that. So now we're talking about having, you know, a platform already set up for you to just basically come in and throw your application on. Google's app engine has some different capabilities within that as well.

So those are really the three different things, software as service, platform as service, infrastructure as service when

you're looking at some of the cloud computing models.

Cloud Service Provider Features



**007 So what this is, this is actually a couple of the features off of the GoGrid website, where they actually compare different things, compare Amazon's EC2, and Rack Space and GoGrid. And I got this in here because I want to call out some of the features. And these are the things that we'll talk a little bit about today, but it also gives you an understanding of the types of things that you would want a cloud provider to do for you.

So do you have VPN options? Can I VPN into my infrastructure? How do I access

these boxes? Usually you'll access them over SSH or remote desktop, if it's a Windows box, so you're relying on the security within those protocols. Are there options to VPN in, to maybe link my corporate infrastructure to this cloud computing infrastructure, or client VPN from a workstation to that environment? Do they allow that? Are there intrusion detection options? How much of that do I have to configure and install myself, versus how much logging and monitoring and infrastructure security infrastructure capabilities does that service provider provide? Because those are big questions and not all of them do that?

What kind of role-based access controls are there, right? Can I have different levels of users, different levels of administrators? Who's allowed to actually provision a new instance, or create a template, all right? Who's allowed to actually install software in that environment? How are those things monitored? Who logs in and what changes were made? What kind of application programming interface is available for me to actually link into that instance or provide some sort of programming? Those are big things.

Do they provide additional services, like managed DNS, or how do they manage their IP addresses? Can I have a contiguous block of IP addresses, or is it just some sort of hodgepodge IP address that they provide me?

Persistent storage, what happens when this thing gets shut down? Does it keep my data on there? Can I move that data, can I mount a cloud volume, like I

mentioned with Amazon's S3 process? Can I mount a virtual volume to multiple machines, and move it around to different machines? Those are some capabilities that you might be looking at.

Other things from a security perspective, is what happens once I go away? Once my instance is done, is that data securely wiped in some fashion? Is there the opportunity for me to have dedicated resources? Maybe I don't want to share a physical server with any other customers, I want my own dedicated server resources. Is that an option? Okay, because these are a lot of things that are concerns when it comes to security and protecting your information, that come up.

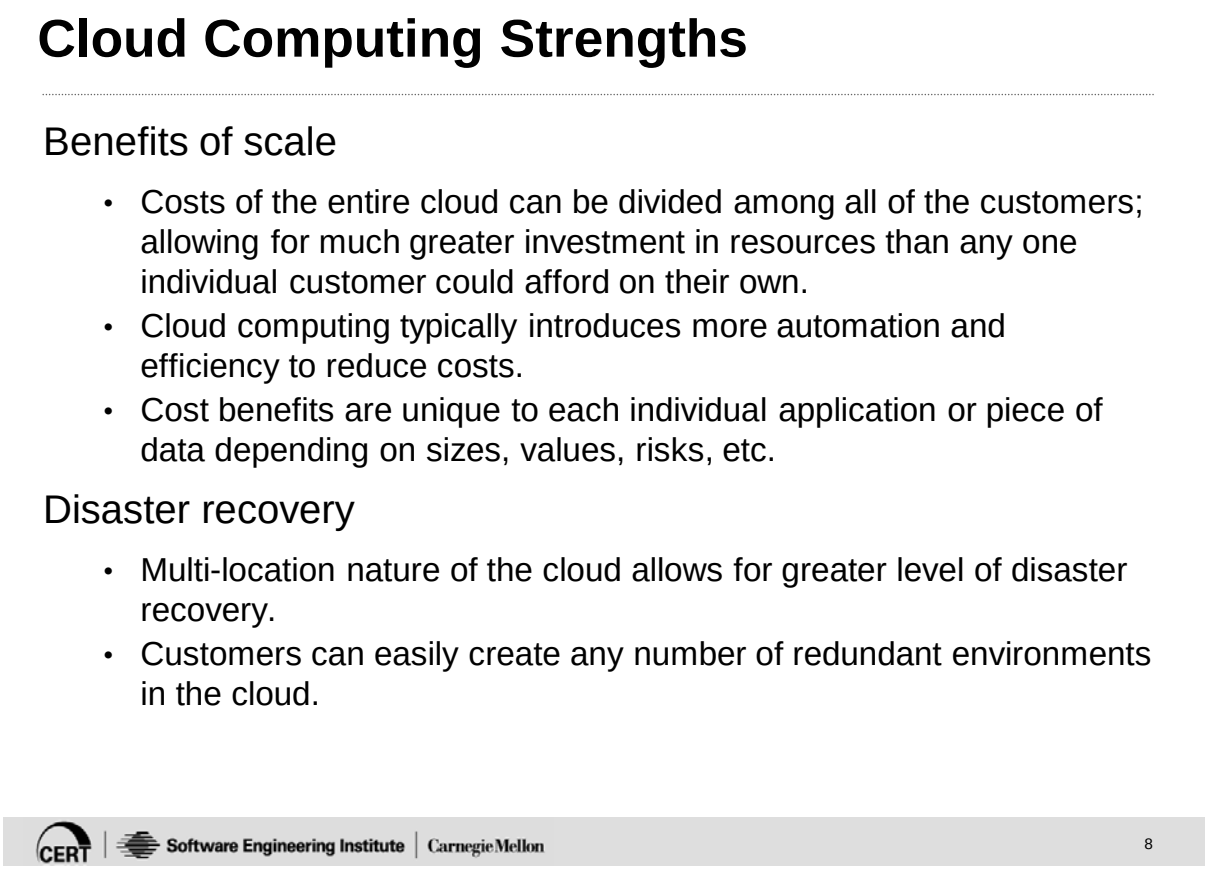
Let's see.. yeah. And then, of course, there's the support pieces of it, 24-7 support, is it e-mail support, is it phone support? What kind of engineering or dedicated team resources do I have? And what is the service level agreement? We'll talk more about SLAs toward the end of this, and what is actually going to be provided to me. Am I guaranteed 100 percent uptime? It's a pretty bold statement. I think Amazon is at 99.9, three nines, with the assumption that there's some downtime for maintenance or what the case might be.

And then, of course, what sort of options are available, Windows Server 2003, Windows 2008 Server, when you get into talking about Win-- Linux it's one thing, right, because you've got open source, you've got freely available operating systems,. When you talk about Windows, you talk about SQL Server. You've got different licensing things that work out. So

your pay-as-you-go pricing is more than just bandwidth consumption, now you've got-- the cloud provider has to somehow build into it the licensing for the operating system, and how that's going to work.

The good thing about a cloud computing provider in that case, is, you don't, as a customer, have to try to figure out whether it's one CPU, two CPUs or how many different servers, and-- They'll work with you. They've got the experience in that area to understand how it is in a pay-per-use basis, pay-as-you-go basis.

Cloud Computing Strengths



**008 So there's lots of benefits to cloud computing, the first one is benefits of

scale. We can, if we need to, grow and have more servers, do load balancing, if, as our infrastructure, our needs grow, we can kind of do that per scale. We don't have to build it all at once. We don't have that initial capital expenditure to get the servers and the networking equipment to pull that off. So there's, you know, people that are starting up a business that want to get into having a hosted website or something like that, they can get into it pretty easily, with some sort of cloud computing options. If they want to access a CRM database, if they want to do some sort of contact management, there's applications that will let them do that. If they can't afford to buy Microsoft Office, or they don't want to buy Microsoft Office, there's Google Apps out there, that will let you do word processing and spreadsheets and those types of things.

The automation piece is neat. Like I said, there's a lot of them have web interfaces that will let you go through and actually point and click, and deploy a new box, without actually having to physically understand how to plug things in and deal with those types of things. I would strongly suggest anybody that's doing sort of that instance deployment, has some sort of skills and background in systems administration to understand the risks involved with that, and what's going on.

So multiple locations in the disaster recovery option, that's certainly a big deal. Certain cloud providers, Amazon, a Rackspace, a GoGrid, there's a number of larger providers out there that have not just facilities in the United States, but throughout the word, for being able to handle disaster recovery.

One of the reasons I-- I-- One of the reasons somebody I know-- once went to some cloud computing technologies, was the transference of risk. You see that, when you talk about risk management, one of the options is to transfer that risk to somebody else. I don't want to do the backups every day, I don't want to have to worry about what happens from a disaster recovery strategy. I don't want to have to manage that infrastructure piece of it. Maybe I'm only doing it part time, maybe I don't have the resources, maybe I'm a small company, a midsized company that just doesn't have the IT resources, but I need to have that sort of availability requirement. I can transfer that risk a little bit to a service provider to handle that for me.

Now the trick is, what are the service level agreements? What sort of policies and things do I need to adhere to, as a business, and ensure that my service provider adheres to those same sorts of policies.

So the customers can easily create a number of redundant environments. I've got a quote for you, right out of some-- the Amazon guides, the security guides. And this is interesting. This is sort of a risk thing as well. So I've got three or four quotes, so I'll read to you. "Network and application level security is your responsibility." I like that one. So they'll handle some of it. There are some security groups that can be configured so you can-- that are mandatory and you have to allow certain ports, or destinations inbound on the firewall features that you can configure through the web interface, but in terms of host-

based security, application security, if you've got a web server on there, making sure that it's secure code, or you have a web application firewall or something like that, is your responsibility. That's kind of neat, right? They kind of said, it's all you.

So there are multiple availability zones, all right, in the Amazon world. Like I said, there's US stuff. They've got locations in Germany and Ireland, in England, in London, specifically, Singapore, Hong Kong, they've got data centers throughout, for doing different things, but you have to manually choose to deploy to one of those different things, and oh, by the way, "your data is not proactively replicated across regions, unless you-- unless it's done by the customer."

So if you want to take advantage of some of these disaster recovery scenarios in multiple locations, you can't assume that this is something that's magically happening by the cloud provider. It's something that you have to configure in your environment. You have to provision multiple instances, you have to schedule the backups of your data, the synchronization of your data, whatever the case might be.

"All traffic between regions," for instance, between the US and the European Union, "is over the public internet." So that's a little caveat there. So in the US, we have multiple data centers, and this is-- again, this is specifically Amazon related things, but-- so in the United States, there's several data centers. If you were going to replicate between different data centers, or have instances there, it's going over Amazon infrastructure, and it's semi-

protected in that regard. If you wanted to replicate to some other data center overseas, it has to leave Amazon, go on the internet, come back into Amazon, so then we need to talk about encryption, right? We need to make sure, if there's sensitive data in there, that that data needs to be encrypted and protected. So there's a bunch of little things that could happen when you start to talk about these disaster recoveries and multiple locations.

You can't assume that everything is secure. I talk about that risk transference, you can't assume that everything is just magically secure because somebody else is taking care of it. And that's one of the big issues that you need to address when you're going to a cloud computing environment.

Cloud Computing Weaknesses

Interoperability

- Organizations may face challenges moving data between the local business and cloud provider.
- They may also face challenges moving data to other cloud providers.

Standards are lacking

- As is the case with the introduction of many other types of technology, standards have not yet been defined.

Technical risks

Operational risks

**009 So interoperability, we start to talk a little bit about some of the weaknesses. We'll talk in depth about the risk pieces, technical and operational risks, so we'll talk just generally right now about some of the weaknesses with interoperability. So the idea here is, if I'm using Amazon services and I want to go to one of those other ones that I mentioned, like GoGrid, or Rackspace, or whatever it might be, how do I do that? Am I locked into Amazon? How do I transfer my VMware instance, or my Xen instance? Is it possible?

There are things out there that are looking to develop some middleware that will let you move images between-- not

necessarily images, but we'll say configurations and deployments between different providers. But just in general, how can I do that? There's definitely challenges with moving the data, all right? So we're on a pay-per-use model, right, so we're on a bandwidth model as well. How-- if I've got to upload a couple of hundred meg, that might not be a big deal. If I've got to try to upload a ten gig database to this cloud computing data center, that is a big deal, right? Some of them have some sneakernet features for you. You can send them a disk, and they can install it, and you still have to manage that whole, you know, security of that disk, whether it's encrypted, or how much data it is, and what procedures it takes to actually implement that. But that is definitely an issue.

The standards are lacking-- they're starting to get there. There's definitely some initiatives out there to try and develop some standards, whether it's from an API middleware programming perspective, or data handling perspective, or from a data center perspective as well. And there are some standards that generally collocation facilities or data centers have to go by, our SAS-70 certifications, or ISO, what is it, 27001, certifications, there are some industry best practices that should be taking place, whether it's collocation facilities or cloud computing facilities.

Notices