Using ArcGIS for Server in the Amazon Cloud

Sterling Quinn, Andrew Stauffer, and David McGuire discuss how to set up ArcGIS for Server in the Amazon Elastic Compute Cloud, typical deployment patterns, and workflows.

http://video.esri.com/watch/664/using-arcgis-for-server-in-the-amazon-cloud

---

Video Transcription

00:01 Alright. Welcome. We'll go ahead and get started with Using ArcGIS Server in the Amazon Cloud.

00:07 How's the audio in the back? Good? Pretty good. Okay.

00:12 We got a variety of us here today from different areas in different departments in the company.

00:17 Andrew Stauffer is a server support technical lead, so he gets the really hard incidents and figures out how to fix them every time.

00:27 He's out of the Seattle office.

00:30 I'm a product engineer on the ArcGIS Server development team working out of Olympia, Washington.

00:35 And Dave McGuire is with us. He's with technical marketing in Redlands.

00:38 So, kind of a variety of departments and experiences.

00:41 The common thread is that we've worked with Amazon EC2, and we're excited to talk about it today and excited that you're here.

00:44 Raise your hand if you've worked with Amazon EC2 before or if you're used ArcGIS Server in it.

00:53 Just a couple here and there.

00:55 Okay. We kind of anticipated that, so we're going to start from the beginning, and this will be a review if you've used this before.

01:02 We'll start with an overview, and then we'll get into some more advanced topics that maybe you haven't done before…

01:07 …or might be useful to you to add onto whatever deployments you've done in Amazon EC2.

01:13 We'll talk about how to…especially how to scale your configuration and take advantage of the elastic properties of EC2.

01:20 EC2 stands for Elastic Compute Cloud.

01:23 So, we want to show you how to take advantage of that elasticity, the ability to grow your site...

01:28 ...and then shrink it in response to demand or the lack thereof.

01:33 David's going to present us a case study of a real site that was built using Amazon EC2 that Esri technical marketing set up.

01:42 And then at the end, we'll take some frequently asked questions.

01:44 One really hot topic with Amazon EC2 is billing and licensing, also security.

01:51 We're going to cover some of those near the end.

01:53 If you have questions throughout, please hold those and jot them down.

01:56 We'll be happy to take your questions at the end, and we'll try to leave some time for that.

02:02 So to get started, to understand Amazon EC2, it's important to understand virtualization...

02:07 ...and the ability to have a single physical machine that hosts multiple virtual machines.

02:15 You log in to those virtual machines just like you're logging in to a separate computer...

02:19 ...but really the one physical box is what hosts all of them, and it's able to divide up that...

02:24 ...and to you it looks like you're just logging into a separate operating system.

02:28 What Amazon did is that they built a very large computing infrastructure to run their website.

02:34 So, everybody at Christmastime has gone into Amazon.com and bought a package.

02:40 They had to build a lot of computing power to be able to handle those rises in traffic.

02:44 Eventually, they decided to sell out this...this hardware infrastructure as a web service.

02:49 They felt that they had gotten good enough at it that they could sell it as an efficient computing platform.

02:55 And so when you use Amazon EC2, what you're doing is you're running virtual machines on Amazon's hardware.

03:02 Somewhere in their data centers throughout the world, you're spinning up a virtual machine.

03:09 And, you can think of it this way, you might have at your work a server room that's got some physical machines in it.

03:14 There's a finite number of machines. They were all built in a certain year, and they have a certain amount of computing power.

03:20 Eventually they're going to run out or become old...

03:23 ...and you can only scale out to the extent of the number of machines you have in your room.

03:28 With Amazon EC2, the boundaries are a little more unknown.

03:33 They add computing power all the time.

03:35 They update their hardware, and you can also expand to what seems to you to be an infinite amount of...

03:42 ...infinite amount of computing power.

03:44 It's really not infinite. They know the bounds.

03:45 But to you, as a cloud user, it's not...it's not important for you to know the back-end infrastructure that they use.

03:55 They just provide you a service where you know that you can get a computer on demand when you need it.

04:03 There's a couple of real advantages to this.

04:05 One I kind of touched on is, in your server room, you're servers may become old, out of date, or fail...

04:12 ...and then you've got to deal with that on your own right away.

04:15 With Amazon EC2, they take care of that part, and they have a very robust network infrastructure.

04:21 It's not a hundred percent foolproof.

04:24 You know, things happen to them, too, occasionally.

04:27 But they're in the business of maintaining their hardware and keeping you up and running...

04:30 ...and if something goes wrong, they're in charge of fixing it.

04:33 Usually they do so very quickly.

04:37 I touched on the ability to do elastic deployments so you can add and remove machines as needed.

04:42 You pay them an hourly fee for the machine time that you use.

04:46 So, if you expand your site, you'd pay extra, but then when your traffic goes down...

04:50 ...you can destroy those extra machines you created and thereby reduce the cost of what you pay from there on out.

04:58 Amazon EC2 is really good for prototyping applications.

05:01 If you want to test this out for just a week and create some virtual machines on Amazon's cloud...

05:07 ...you can pay them the hourly fee for a week, and you can experiment with different sizes of machines...

05:13 ...and then your total bill for that week may be less than $100, and for that...

05:17 ...that's a pretty good deal in come cases for being able to prototype and see if that will work for you.

05:22 Also, if you're developing an application and maybe you don't have an extra machine sitting around as a development box...

05:27 ...Amazon EC2 is great for that as well.

05:29 In fact, I used it...I was giving a demo on server object extensions which requires me to do some development...

05:35 ...and I wanted a clean machine.

05:37 This was a couple weeks ago, and so I thought Amazon would be a great case for that...

05:41 ...where I could just get a clean machine and work with it and then terminate it when I was done.

05:46 David's going to talk about a situation where Amazon EC2 was used in an emergency deployment.

05:51 All of a sudden, the website got a lot of traffic that was unexpected, and it crashed...

05:55 ...so they moved it into Amazon EC2 so that they could easily add more computer power as they saw that they needed it.

06:02 And then a final benefit that I've listed here is the ability to set up a site that's public facing in a fairly easy manner.

06:10 When you create these instances, you have the ability to open up the firewall so that they can be seen by the world.

06:15 At the same time, they're not immediately connected to your network...

06:18 ...so some of the risks might be lower than if you had just gone into your own organization's network...

06:24 ...and tried to set up an external-facing machine.

06:27 In our organization, in Esri, if we want to get an external-facing box inside our server room...

06:32 ...we have to work through the IT staff and fill out many forms and get approval.

06:36 That's for good reason for safety so that we don't, you know, expose things on our network that we don't want to have exposed.

06:44 It's a little bit easier for us to set up a box on Amazon EC2 and set up a few web services...

06:49 ...and then when we're done, we can get rid of that.

06:51 You might have seen in the Plenary Session, if you looked closely...

06:55 ...the top of some of those demo machines that they were remoting into, those were Amazon EC2 machines.

07:00 And the reason they did that is because they just needed those machines for a short time.

07:03 They needed to expose public-facing web services, and then at the end of the conference, they could just get rid of it and clean it up.

07:12 When we talk about Amazon EC2, there's two key terms that we're going to throw out there...

07:16 ...and maybe I've already done it inadvertently that you need to understand, and we often make the mistake of mixing these up.

07:23 In fact, I explained these, I think, on Tuesday's session and then I immediately wasn't thinking and mixed them up.

07:29 So, it's important to...to know what these two things are.

07:32 The first term that we'll talk about is EC2 instance.

07:35 That is the virtual machine that you're running on Amazon.

07:39 You can pick the size of this instance.

07:40 You can pick the number of CPUs.

07:42 You can pick the amount of RAM.

07:43 Amazon offers you a menu of choices.

07:46 They also offer a pricing chart for each one.

07:49 So, don't go crazy right away and pick the biggest machine because you will pay a little bit more...

07:54 ...or a lot more per hour depending on the specs.

07:56 The first time you see that chart of what's available, you may start salivating...

07:59 ...and get all ready to create a whole bunch of map caches and things with it.

08:04 But it's important to very carefully pick those things and look at your...your budget before you do so.

08:11 However, those can be very useful in cases.

08:13 The other term we want to talk about is the image.

08:16 The machine image or A-M-I or AMI, as I'll be in the habit of saying in this session...

08:22 ...defines the operating system and the software that's going to be installed on your EC2 instance.

08:28 Okay. So you use an AMI to create an instance, and from there on you work with your instance.

08:35 If you have ArcGIS Server installed on your instance, you put services on your instance and

applications on your instance.

08:43 When we...what do we mean when we say ArcGIS Server is now supported on Amazon EC2?

08:48 Well, what we mean is that Esri created an AMI, or machine image, that has ArcGIS Server on it.

08:55 It also has ArcGIS Desktop.

08:57 It's running Windows Server 2008, and you can install other software on there as you see fit.

09:04 We have another AMI that's available that has an enterprise geodatabase on it.

09:07 It's running ArcSDE and PostgreSQL.

09:11 The reason a PostgreSQL is on it is that we have the rights to redistribute that AMI with that software on it.

09:22 The way that you work with EC2 is, well, if you worked with EC2 several years ago...

09:27 ...your options would be to use the command line and type in commands or write scripts that work with it.

09:33 Now, they wanted to open EC2 to a wider variety of people.

09:36 They created an easy-to-use GUI application that runs in the web called the AWS Management Console...

09:43 ...which we'll show you in a demo in just a minute.

09:46 There's other third-party tools out there that people have built.

09:49 Before Amazon built their own console, there was an application called Elasticfox that a lot of people still like to use.

09:55 Some people feel that it has still more functionality than the Amazon-provided one.

10:00 Andrew found one just the other day, O2 Amazon EC browser.

10:04 So, these things are just popping up here and there occasionally.

10:07 You might find the one that you like that works the best for you.

10:10 It's largely a matter of personal preference.

10:14 So this is kind of the workflow.

10:15 You'll use your computer in your own office which is the one in the lower left.

10:20 Sometimes we'll call that on-premises or on-site.

10:23 That's you. Then you'll use the AWS Management Console to create an instance running in the cloud or in Amazon EC2.

10:33 So let's watch a demo of how this happens.

10:37 I've already logged into the AWS Management Console.

10:40 One thing about this application is, sometimes it's a little quirky in certain web browsers.

10:45 I found that Google Chrome tends to work the best on this, so I installed...I bit the bullet and installed Chrome...

10:52 ...and it's...it's worked out fine from here on.

10:55 You can see I have a couple instances here.

10:57 I'm working, first of all, in the US West Data Center.

11:02 If I use this drop-down, it shows all the data centers that Amazon has throughout the world.

11:07 Currently, there's five.

11:08 And you might want to pick the one that's closest to you just for the least amount of latency.

11:13 However, sometimes legal requirements will require you to put the data in a certain place.

11:17 So that's why they offer a variety.

11:23 I wouldn't be surprised to see more of these pop up in the future.

11:26 I haven't verified that with Amazon, but they seem to be adding them at a pretty fast pace.

11:31 They seem to be pretty aggressive in making these.

11:34 Right now, I've clicked...I've logged in with my Amazon account, and that's how I got here.

11:40 I clicked the EC2 tab.

11:43 So, Amazon offers a variety of web services and things are available online.

11:50 Today, we'll just be working with EC2 which is their system of renting out hardware infrastructure.

11:56 The other one that you might be familiar with up here is S3.

11:59 That's Amazon's service for offering mass storage and serving out of data in the cloud.

12:08 I've clicked instances here.

12:09 This shows the instances that I have running.

12:11 Actually I have one running now that I've created for this session earlier.

12:15 And I have one that I created yesterday down at the showcase that I was showing somebody a demo.

12:20 And I've terminated this so that I don't incur any further charges on it.

12:27 Let me show you the process of creating an instance of your own.

12:30 So you'll go to this panel and you just choose launch instances.

12:34 And the very first thing that you do is you pick the image that you want to use.

12:38 Amazon provides some vanilla images for you.

12:42 They just have an operating system installed.

12:44 There's a lot of Linux ones on here.

12:46 Originally Amazon only supported Linux instances.

12:51 To get the Esri ones, you'll click Community AMIs and you'll search for Esri.

12:56 Now, in order to see anything here, you need to contact Esri customer service and give them your Amazon account number.

13:02 They will then share the image with your account.

13:05 So not just anybody can go in here an type in Esri and spin up ArcGIS Server.

13:09 You have to...you have to work it out with customer service so that they know your Amazon account.

13:15 I see a bunch of images have actually popped up here.

13:17 We have ones for the different service packs.

13:20 This one at the bottom you may have noticed, this is for a kind of a road-ahead item, but for 10.1...

13:24 ...we're working on an _____ [Unintelligible] image.

13:26 It's not available now but hopefully in the future.

13:30 And I'm just going to pick the one for service pack 2 with ArcGIS Server.

13:33 That's the way to service pack, and I'll click Select, and it'll go onto the next area...

13:38 ...where I'll pick some of the technical details of my instance.

13:41 Now there are so many of these that we don't have time to describe all of them today.

13:45 But we will describe some of the ones where you may want to change the default.

13:50 There's a drop-down here where I can choose an availability zone.

13:53 This is a little bit different than the regions or the data centers that I showed you.

13:57 Within each region, Amazon has sub...you can think of them as subregions or zones that they keep isolated from each other.

14:05 And this is just for, again, creating a more fault-tolerant site.

14:10 There was a representative from Amazon here earlier in the week that told us that...

14:14 ...these different availability zones use different power grids and they're built on different floodplains...

14:19 ...just to be all the more safe in...in not having the situation where all of your hardware could go down at once.

14:27 And so if you're building a site that will have multiple machines in it...

14:29 ...it's a good idea to spread them out between these availability zones.

14:33 If you pick no preference, Amazon will choose a zone for you.

14:37 Sometimes that's not a bad idea because they have logic that will put your instance in the zone...

14:42 ...that is currently most free or available.

14:46 This is the place where you choose the instance type.

14:49 So you can see the different types of different computing power that's available to you.

14:54 Using these instances, you can see there's some very large ones.

14:58 There's also a very small one, Micro.

15:01 We recommend that for ArcGIS Server, you use at least the large size of instance in order to have sufficient memory.

15:12 When I click Continue, there's just more technical options I can pick.

15:16 A couple that I'll point out.

15:18 If you choose CloudWatch, that applies a small extra fee, but that allows Amazon to monitor the usage of your machine...

15:25 ...and this is required if you want to do things like autoscaling based on triggers of CPU usage and things like that.

15:33 We'll cover the autoscaling a little bit later in this session...

15:35 ...but it's important to remember that you would want to enable this option if you want autoscaling.

15:41 Also, Amazon recently introduced something called termination protection.

15:44 This is almost like insurance against you accidentally terminating your instance before you're ready to do it.

15:50 Apparently, there's a way now to get it back if you've enabled this...

15:54 ...and I believe they change a small extra fee for this option.

15:58 I think most of us may have...well, at least I have accidentally terminated my own instance before or a coworker's instance.

16:05 That's always a difficult e-mail to send to confess what you did.

16:08 But, it's not the end of the world.

16:10 You know, Amazon goes by what they call the paper plate theory which is, when working with the cloud...

16:16 ...if you get a bad paper plate or yours gets mixed up, you just throw it away and get a new one.

16:21 And that's the idea with these instances, that if you're setting one up and something goes wrong...

16:26 ...you throw it away and spin up a new one.

16:29 We'll see later in the presentation how, once you get all your apps and services installed, you can make a snapshot...

16:34 ...kind of like a backup so you can easily spin up a new instance with all your stuff on it.

16:40 Amazon now provides you a way to type kind of an alias for your instance.

16:45 Prior to this...prior to this, you had to just look at the ID that they assigned to your instance.

16:52 It was this long cryptic number, and once you got 10 or 15 of those in the list, it got very difficult to remember what instance is what.

16:59 So, you can add different name value pairs here to keep your instance straight.

17:05 You might see, if we switch to another region later on, you'll see how in this conference...

17:08 ...we've each put our name on the instances we're using among Esri staff just to keep our machines straight.

17:13 A lot of times, we put Do Not Delete in large capital letters so that nobody else goes in and messes with our instance.

17:21 In this panel, we're going to talk about something called key pairs.

17:25 This is the way that Amazon allows you to get the administrator password of your instance that you create.

17:33 When you first create your instance, you'll need to log in for the first time and, obviously...

17:36 ...you need to know the administrator password.

17:39 And in order to decrypt that password and display it on your screen, you need to have this special file called a key pair file.

17:45 If you don't have one, you can click create a new key pair.

17:49 Just type a name for it, and you'll be able to download it to your machine.

17:52 It's just a little text file with the extension .pem.

17:57 If you have one already, you can just choose it from the drop-down.

18:00 What you're going to do is save that PEM file in a safe place on your machine...

18:04 ...so that you can go retrieve at the time that you're ready to get the administrator password.

18:10 Another thing you need to do is configure the firewall that Amazon provides for your machine.

18:14 By default, Amazon puts a firewall around your machine and all traffic is blocked...all incoming traffic is blocked.

18:21 So you need to selectively open some ports in order for people to get into your machine in the way that you want.

18:29 I'll just show this by creating a new security group.

18:35 You're required to type a name and description.

18:38 And then you go ahead and add security rules.

18:40 This is like poking holes into your firewall.

18:43 One that you might want to open, if you want to expose your web services to the world, is HTTP access.

18:49 And then you're able to put in an IP address or range of IPs of people who are allowed to come in this way into your machine.

18:56 If you want everybody to come in, you use 000 like this.

19:00 This is called C-I-D-R or CIDR notation.

19:03 It's a way of specifying a range of IPs.

19:05 If this part's confusing to you or you don't know your IP address...

19:08 ...you may work with your IT staff to understand what values would be appropriate to put in these boxes.

19:14 Once I click Add Rule, you can see that it goes into the list of rules here, and it's opened up access on port 80.

19:21 If you need to open another port, you can choose to us a custom rule and then you can type in the port or port range here.

19:27 So if your organization uses a port other than 80, that's what you would do.

19:32 There's also an option for HTTPS if that's required.

19:37 One other rule that you'll want to open is Remote Desktop Access so that you can log in to the instance.

19:42 So you can choose RDP and then typically you'll type in just your IP address or an approved range of IPs.

19:57 ...and you can review all this and make sure it's okay, and I'll click Launch, and I can close this...

19:59 So I'm going to add that rule, and that will allow me to log in later.

20:02 ...and I start to see some things happen in my list here.

20:06 This shows that right now it's apportioning a space on that Amazon Data Center for my instance.

20:14 If you click this, you'll be able to see some of the things that it did.

20:20 So, it created my...it created my instance in US West 1C zone, and it's using the security group I set up.

20:30 It's got my Amazon account number here so I know who owns it, and it's got some other just basic...basic things about this instance.

20:39 Right now, it doesn't have...it didn't have a public DNS name when I first did this...

20:43 ...but if I refresh, I anticipate that I'll be able to see that now that it is actually running.

20:48 It's assigned it an address that I can use to log in.

20:54 Now even though this says it's running, it's not ready for me to log in and start using ArcGIS Server.

20:58 It needs to do a couple things.

21:00 It needs to start up Windows and get that ready to use.

21:03 Then it also needs to start up ArcGIS Server.

21:06 We put in a couple of scripts in the ArcGIS Server startup in order to make things easier for you and set up smoothly in the cloud.

21:14 There's a couple of things those scripts do.

21:15 They run the postinstall for you, and they set up all the ArcGIS SOM, SOC, and ArcGIS web services accounts.

21:22 That's a good thing. That means you don't have to worry about that.

21:25 It's all ready to go when you log in.

21:27 One thing that it also does is it handles some of the quirks of how Amazon works.

21:32 If I right-click and I stop this instance, that's akin to turning off the power on my physical machine.

21:39 That's how I stop the clock on my billing.

21:42 My instance, however, is still available, so I could right-click and choose Start.

21:46 But when you start an instance again, Amazon assigns a new IP address to it, and that can potentially mess up ArcGIS Server.

21:54 So we've put scripts inside of the instance that autocorrect for that new IP.

21:59 Now, this could still be a problem for your applications.

22:01 What if you've written a web application that's pointing at your server and it has a URL...

22:05 ...and all of a sudden the IP address or the URL changes on the background?

22:11 To remedy that, Amazon has a solution called elastic IPs.

22:14 So you can ask Amazon to give you a static IP address.

22:18 It's not going to change, and then you can choose to associate this address with any of your instances that you want.

22:25 So this way, if you stop and start your instance, your instance gets a new address...

22:28 ..you can just attach it back on to this elastic IP.

22:31 You can do that automatically through scripting so that there's minimal down time.

22:37 Later on, I'll show you a little bit about logging in to this instance and working with it.

22:43 So, while that's going on, we'll talk a little bit more about working with an instance and what you're going to do once you log in.

22:50 You're going to use Windows Remote Desktop client to get into the instance, and we'll demonstrate that.

22:57 The first thing you're going to want to do is change that autoassigned password that Amazon gave you.

23:03 When you decrypt a password, usually it has a lot of special characters and things in it...

23:08 ...for a good reason so it's not easy for somebody to guess.

23:12 But you'll log in the first time and type in your password and then you'll probably want to go into control panel...

23:16 ...and change that administrator password to something that you know or that you remember.

23:21 You also need to license ArcGIS, so you'll see the same software authorization wizard...

23:26 ...that you see when you use ArcGIS Server on premises.

23:29 And you'll need to complete that wizard and license to Esri software before you can use it.

23:34 Optionally, you can add or remove disk space.

23:37 We give you a hundred gigabyte drive that's attached.

23:39 That's also running in Amazon.

23:41 That's called an Elastic Block Store volume, or EBS volume.

23:46 So you see the word EBS in Amazon, that means a storage disk or volume that you can use.

23:52 If you want, you can remove that one and attach one that's smaller, just to incur less cost for the storage…

23:59 …or you can optionally add one that's larger or have multiple EBS volumes.

24:04 And as I mentioned, there's no need to run a postinstall so you can skip that part.

24:09 At that point, you're ready to move your data over onto the instance.

24:12 Now, this can be…this can be tricky.

24:14 Just a couple things you need to remember is that you are moving data across the Internet…

24:19 …and depending on your bandwidth and your ability to do that, that may be a slow process.

24:25 There's a couple strategies you can use for doing that, too…

24:28 …and your organization may require you to do one thing or the other based on security requirements.

24:34 You know, FTP is an option, secure FTP.

24:37 Windows Remote Desktop offers copy and paste functionality from one machine to another…

24:43 …or some people like to use a web locker space or a personal web space to do it.

24:49 So, a site like Dropbox.com gives you a certain amount of apportion space in the cloud, so to speak.

24:56 So you could upload your stuff to Dropbox and then download it onto your instance.

25:00 I've seen a lot of different strategies for doing this.

25:03 In any case, you may need to plan some time to do this and test which way is the fastest for you…

25:08 …and you may need to work with your IT staff if your data is sensitive to understand how to most securely move that data…

25:14 …especially if you need to encrypt it.

25:17 If you have extraordinary amount of data, Amazon offers a service where you can mail them a drive…

25:23 …and they will load it directly onto the cloud from their area.

25:28 We call this sneakernet, which is you take a disk and you walk across the room and plug it into the other one.

25:34 That's essentially what you're doing when you mail this drive to Amazon.

25:37 That service is called AWS Import/Export.

25:46 Once you get your data onto the instance, then you can go ahead and start working with ArcGIS Server.

25:52 And in this session, we won't really have a lot of demos of working with ArcGIS Server on Amazon EC2.

25:58 The reason being is because, by this point, it's pretty much the same as what you do now on premises.

26:03 You have Manager; you have ArcCatalog; you can go through the wizard and publish services.

26:08 The only difference is the web address of those.

26:10 It's pointing at an Amazon machine.

26:13 You can also write web applications and put them on your instance.

26:17 The instance is running IIS web server so you can copy a web application onto there and host it just like you would on premises.

26:27 With these instances in Amazon, you can do some fun types of things.

26:32 Map caching is one thing that can be facilitated by EC2.

26:35 Suppose that you need to serve a service on Amazon that has a big map cache.

26:41 I would almost argue that it might be faster to rebuild your cache in EC2 than it would be to copy that large cache over.

26:47 That's because EC2 gives you potentially a lot of computing power.

26:51 So, you might rent a high-powered instance from Amazon for just a day or two....

26:56 ...and use it to crank out your cache tiles as fast as you can.

26:59 And then when you're done with that instance, you can switch back to a smaller one that would cost you less...

27:05 ...but, you've got all those tiles that you made, so now you're ready to go.

27:11 I mentioned that Amazon EC2 is good for prototyping and testing.

27:14 So, If you want to try this and see how many tiles you can generate quickly, you can try it with the different instance sizes.

27:20 And there were some folks that we know at Esri that did this, and they used all the different instances to capture the same cache.

27:26 And you can see the differences in build time that occurred between those instance sizes.

27:33 They also calculated tiles per second.

27:35 In my mind, when you're building a very large cache, that's a good number to know so you get an idea of about...

27:41 ...that you can project about how long it's going to take you to build the cache.

27:44 They also calculated the relative cost.

27:46 Now this is...this should not be construed as the exact cost, because Amazon charges you per partial hour.

27:52 So even if you were to build a cache for 10 minutes on an instance that was 48 cents per hour, they would charge you 48 cents.

27:59 So, this relative cost only takes effect if you were to cache for many, many hours.

28:00 It really, if you set up the same instance and just...or the same AMI and install it on various instances of different sizes...

28:04 But, you can see how some of the instances are more efficient than others.

28:08 You may need to trade cost for speed.

28:10 For example, if you need to get those tiles built now, you might try m2.4Xlarge.

28:17 It costs a little more relatively, but it creates the cache the fastest.

28:23 On the flip side of that, if cost is a concern for you and doesn't matter quite how much time...

28:28 ...you may choose one of these more efficient instances.

28:31 So, I encourage you to run your own test like this if you're thinking about caching in the cloud.

28:46 ...you can make a chart just like this with your own data, and you would have a pretty good idea of which one to use.

28:50 This might only take a day or two and cost you...cost you a little bit of money to find out.

28:56 Once you have got your services, application, data all ready to go on your instance...

29:02 ...then you want to take a snapshot of it and save it.

29:05 And so Amazon allows you to save a custom AMI or image, and you'll use that later if you want to scale out...

29:12 ...because you'll use this AMI to create clones of your machine, and Andrew will get into more detail about that later on.

29:20 So let's log in to one of these instances and take a look at what's there and what you'll see when you first log in to it.

29:29 So I'm going to log in to this one that I already created...

29:32 ...just because I'm not sure if...if this one I just created would be completely ready yet.

29:38 But let me show a few things just before I do that of how you can tell the status of your instance.

29:42 If I right-click here and choose Get System Log, I'll see some messages that appear.

29:47 Initially, this will be blank, but now on this one I created, it actually says Windows is ready to use.

29:54 So that means Windows got initialized.

29:56 After this, you need to wait 5 or 10 minutes longer for ArcGIS Server to start up.

30:00 Unfortunately, we don't know an exact time.

30:03 It may vary depending on the health of the cloud that day and size of your instance, perhaps.

30:12 Another thing you'll need to do is you'll need to right-click your instance and choose Get Windows Password.

30:17 This is how you get the password for the first time.

30:21 So it will show you the encrypted password and then it says, hey, in order to show you the real password and decrypt this...

30:27 ...I need you to provide that key pair file that you created back when you were going through the wizard.

30:32 So you'll have this .pem key pair file on your machine, and you'll just go ahead and browse to it here.

30:38 This opens a file browser dialog.

30:41 And when you browse that key pair, it will display your admin password on the screen...

30:44 ...and that's when you're going to jot it down or copy it into Notepad so that you can then log in to your instance.

30:53 This particular instance that I've created, I created that elastic IP for it.

30:59 I also want to show you some of the volumes on disk, so that instance uses two drives.

31:05 It's got the hundred-gigabyte one for storage, and it's go the 30-gigabyte C drive.

31:10 So when you click volumes, you'll be able to see all your storage space.

31:13 Now whether your instance is turned on or off, you always pay for this amount of storage.

31:18 The storage is a relatively small cost compared to the computing power cost, the CPU costs.

31:24 So, just be aware that you will be paying a small fee even once you stop your instance because of this.

31:31 So let's go ahead and use this elastic IP, and we'll remote desktop into my instance.

31:37 So it's already to go.

31:38 You'll choose...you'll need to put administrator in here, and you'll go ahead and connect...

31:46 ...and it's prompting me for the administrator password for this machine...

31:50 ...which I have actually gone ahead and changed to something I know.

31:54 What you'll see when you first come in is the software authorization wizard...

31:57 ...and this will pop up and will prompt you to authorize ArcGIS Server.

32:02 You must license any software...Esri software that you cause to run inside of EC2.

32:07 And towards the end, we'll talk a little bit more about the licensing requirements.

32:11 I've actually already completed this wizard, and the reason I know that is because on this Welcome screen that pops up...

32:21 And this gives me some shortcuts that I can use to get into Manager or the services directory.

32:28 ...it says authorized.

32:32 If I look at the Windows Explorer for this machine, I can see the two drives.

32:39 Let me go up to the computer level.

32:42 So the C drive of 30 gigabytes that I showed you and then the D drive of about 100 gigabytes.

32:48 We recommend you put all of your data on D.

32:51 The reason is, if you terminate this instance accidentally or otherwise, the C drive goes away, but the D volume sticks around.

32:58 It's still available in that list of volumes that I showed.

33:02 So it's a little bit safer to put your data on D.

33:05 We create a folder for you that's already there called GIS Data, and we also put all of the ArcGIS Server directories on D.

33:12 So we set this up automatically to have the map caches and things go into D.

33:19 You'll see, when you initially log in, the Sample Services folder...

33:23 ...and I've actually gone ahead and copied in some other data into here for services that I've published after the fact.

33:30 You might notice in Windows Explorer also that I'm pointing at my local laptop.

33:34 So I have access...I have set up Remote Desktop to give it access to my local disk drive.

33:39 And if I want, I can browse to something on my laptop and then copy it over and paste it into GIS data.

33:46 So that's pretty easy solution for getting data onto the cloud.

33:50 Now older versions of Windows had security...potential security vulnerabilities when you did this.

33:55 So, if you're concerned about this or if you're in a very highly sensitive environment...

34:00 ...you might want to work with your IT staff to understand whether this is an acceptable way of copying and pasting data to the cloud.

34:06 But, it's the method that I tend to do at work just because my data is not super sensitive, and it's the easiest way.

34:15 Some other things I want to show you is just that Manager is available on this machine...

34:19 ...and you can look at your services and manage them just like you would anywhere else.

34:24 It also times out after a while, just like anywhere else.

34:27 It's the same manager that you know and love.

34:30 Then there's really nothing else to show beyond that.

34:32 On there, it works the same.

34:33 You can also use ArcCatalog to view and preview your services if you prefer to use this way of administering ArcGIS Server.

34:41 An admin connection to your server is already created for you automatically inside of ArcCatalog...

34:46 ...so you don't have to worry about setting that up.

34:48 I've just gone ahead and renamed it here to something...something more interesting.

34:52 But if I were to go into server properties, you see that you have administrative control over your server...

34:58 ...just like you would anywhere else.

35:01 One thing I want to do just to prove that this is external facing is, I'll go ahead and use the services directory from my laptop.

35:09 So, okay, Andrew, you're supposed to remember this number so I can type it in from my laptop...1847241153.

35:18 Got it.

35:21 Let's go ahead and just use my local machine.

35:31 Anybody jot this down?

35:32 1847241153.

35:36 I thought it was saved.

35:40 Well, this is where you can all go get coffee.

35:50 7241153. Next time I'm using Notepad.

36:01 So now I'm using this machine here in San Diego to go hit the...the one that's in Amazon's Data Center...

36:07 ...the one that I chose was Northern California.

36:09 So, this would be hitting the services directory, hopefully, if I typed it correctly.

36:19 Oh, 41, sorry.

36:23 Now because I opened HTTP access for everybody through port 80, that's why I'm able to do this.

36:28 If I'd not opened that firewall, I would not be able to...I would not be able to put this address.

36:40 Okay, I messed up again.

36:43 Whoops.

36:45 It's that time of week, I tell you.

36:56 Boy, alright. There we go.

36:58 So, you'll see, these are some of the sample services that Esri sets up for you.

37:02 If you see the earth at night, that's a good sign.

37:04 That means everything's up and running.

37:06 And that's part of...part of the reason that we put these on here is just so you can verify that your service is running correctly.

37:13 And you can even have all these options such as view it in ArcGIS JavaScript or you can pull it into arcgis.com...

37:20 ...which is a cool way to test out your web services and see how they mash up over other services.

37:27 So we'll see more of this earth at night later on in another demo.

37:30 But, I've taken too much time.

37:32 I'm going to pass things up to Andrew, and he's going to get into some more advanced topics...

37:36 ...about how you would start scaling out, adding more machines, and fun things like that.

37:44 Thank you, Sterling.

37:46 Wow.

37:52 If I keep my hand on it, it stops. I'm magical. Alright.

37:58 Yeah, we will. Save everybody...

38:03 [Inaudible.]

38:05 [Inaudible.]

38:07 Alright.

38:09 And, okay. We have good sound.

38:11 Sorry about that.

38:13 Alright. So, Sterling's talked a lot about...

38:17 [Inaudible.]

38:19 Yeah, he did talk a lot.

38:20 No, Sterling's talked a lot about setting up our instance.

38:24 So we copied our map. We copied our data.

38:27 We've published some services.

38:29 Maybe we've put an application up there.

38:31 Now that we have this, we want to go ahead and talk about scaling our configuration.

38:36 So, initially, as I'm just doing this copy and paste of the data, I don't need a very large instance.

38:44 I'm still in this development state.

38:47 So as I'm in this development state and if I'm paying for compute time by the hour, I might as well choose the cheaper option.

38:54 Because, again, I'm just uploading some data, configuring services.

38:58 I'm not doing anything too intense.

39:01 Once I have all my services up and published, I put my web app together, and I know things are working...

39:07 So I'm actually testing for functionality.

39:09 I'm going to go ahead and actually take this instance and moving it forward into a staging environment.

39:16 So maybe I'm making the one instance larger.

39:20 And at this point, I've already done all my testing that proved functionality and, yes, everything does work as expected.

39:28 So now I want to whale on this thing and make sure that I'm getting the good performance.

39:33 Is there anything wrong with Amazon's Data Center, do I have any doubts?

39:37 So I'm doing all my testing here, and as I really beat on the staging server...

39:47 ...I figure out and answer some questions as far as how do I want to scale as I move forward on into production.

39:50 So it's this iterative process.

39:51 I'm just kind of running over and over to make sure, okay, I've got the right server, I've got the right size...

39:57 ...and as I move into production, you know, maybe I determine, hey, I actually need two machines here.

40:02 Now when I look at the scenario, there's actually two ways that I can go ahead and scale up.

40:09 So the first one we're going to talk about is vertical scaling.

40:12 What this means is I take that AMI that I've created and, again, once I've copied all my data over, I want to preserve my work.

40:21 So I take that snapshot or I create a custom AMI which ends up being a blueprint for future instances.

40:28 I create my development instance.

40:29 I'm testing on it.

40:30 Alright, now we're ready to do a little bit performance.

40:33 I can actually go ahead and increase the instance type, so I use more CPUs, more RAM.

40:40 One machine may not be able to handle the load of the application that I'm putting together.

40:45 So I actually want to do some horizontal scaling.

40:48 Now horizontal scaling means I'm basically putting more soldiers in my army.

40:52 I have a lot more of identical machines that are all going to be available to respond and keep my application performant under times of stress.

41:04 I'll talk a little bit more about vertical scaling.

41:07 We mentioned we can change the instance site.

41:10 So earlier in the first demonstration, Sterling showed there's an instance type that was, what was it like extra large...

41:17 ...and that's kind of the small one, right?

41:20 So, I started with that, two CPUs, a decent amount of RAM.

41:25 I determine, right, it needs to be bigger.

41:26 I can change that instance type in place.

41:29 So all I have to do is stop the instance, change it, start it back up...

41:33 ...and now that machine just turned into an eight-CPU machine with like 36 gigs of RAM.

41:39 Now, one of the things that Amazon does is when we restart instances, a new machine name gets applied.

41:49 And so, with our own custom AMI, we've actually put in a service that accommodates for this name change.

41:56 But that can be really annoying if I'm building an application.

41:58 I don't want to constantly change the URL.

42:01 So I'll actually apply a static IP or an elastic IP here.

42:05 This elastic IP allows me to switch out the machine or give me a consistent reference point.

42:11 So if we could remember all the numbers that Sterling was working with earlier, we could always go back to that machine...

42:17 ...whether it was, you know, a large instance, a small instance, or another one that I've created that has different services.

42:25 And you just swap out that IP address.

42:29 Now the other type of scaling is horizontal scaling, and we've actually seen this a lot already.

42:35 One of the most common configurations that people have out there in a multiple machine deployment...

42:42 ...is where we have the web server and the SOM running on one machine and then we have additional SOC machines.

42:49 So the request comes in through the web server to the SOM.

42:52 The SOM's like a manager, and the SOM goes ahead and distributes out the work.

42:57 So it says, hey, SOC machine 1, do some work; SOC machine 2, do some work.

43:02 Now the other type of...or other way to go ahead and scale out horizontally, and some people actually do do this on premises.

43:10 It's kind of rare, but this is what we're actually doing more and more, and this is how we've architected our design here in EC2...

43:16 ...is we use an elastic load balancer to look at identical machines that have all the components installed.

43:24 So the request comes through the web, it hits the load balancer, the load balancer says, okay,

machine 1 do this...

43:31 ...and machine 1 has the web server SOM and SOC, processes the request and returns it.

43:37 And this works out well because I can go ahead and just fire up more and more machines to put behind that load balancer.

43:44 Now, how many people in here are from an IT department? Anybody? Okay.

43:50 Setting up a load balancer, adding machines, does that take a little bit more than a minute?

43:55 Yeah. The great thing is when I'm in Amazon, I can literally do this in about two to three minutes.

44:05 I go to my elastic load balancer.

44:07 There's a nice wizard that I walk through.

44:09 I create it, and with the ease of point and click, I'm adding new machines behind the load balancer.

44:16 One of the nice things that the load balancer does have is it has a health check.

44:20 So it's going to ping those servers to make sure they're okay, and I can manually configure that.

44:28 So, we've sort of stressed over and over that AMIs are the blueprints for instances.

44:34 So, after I've copied up all my data, I'm going to use this blueprint to go ahead and start up a new instance, maybe two...

44:43 ...and apply them behind that load balancer, and that's how I'm making sure that I get these identical machines all the time.

44:53 Now Sterling pointed out that I could go ahead an enable CloudWatch.

44:56 And if I'm working with the elastic load balancer, I may want to go ahead and think about implementing autoscaling.

45:03 So autoscaling's good and bad.

45:05 I can go ahead and say, alright, based on some CPU usage, I'm going to go ahead, script it out...

45:10 ...and basically authorize this to go ahead and start up new machines as I need.

45:15 A good thing about that is new machines come up and hit a load.

45:18 Bad thing is, I don't always know when that's happening.

45:21 So, wow, there's a big demand and maybe I don't want to spin up four machine or five machines...

45:27 ...but, you know, according to my rules that I've set up, it'll continue to scale out and build up

those four to five machines.

45:35 Now we'll want to go ahead and use that custom AMI, and what I'm seeing here is I have my instances...

45:41 ...we're under a period of high demand.

45:44 The CPU usage goes up to like 6, or 80 percent. I monitor that.

45:48 My custom AMI says, alright, we're busy.

45:51 We look at the autoscaling rules. We fire up the third instance.

45:54 It pings the load balancer for health, and then finally it comes back into the configuration...

45:59 ...and starts to take pressure off of those two instances.

46:03 So now our CPU usage is down, and to the end user, they're none the wiser.

46:08 They just think your application is always fast, always smooth, and you have the best services around.

46:17 Now, to do autoscaling, you have to do some scripting, and we use the Amazon API.

46:22 This is what Sterling alluded to that several years ago we didn't really have this management console.

46:26 We had to do things through command line.

46:29 Well, as we know, command line's always still there.

46:31 And so if I download this...the Amazon API, I can work with it and set up my autoscaling rules.

46:38 But I can also do some other simple procedures such as stopping and starting instances.

46:44 We have two wonderful blog posts in the ArcGIS Server blog that talk about working with the Amazon API...

46:51 ...and getting started with scripting.

46:53 I'm not really a big programmer person.

46:55 Maybe you aren't either.

46:57 But the blog post proves how easy it can be to go ahead and work with this API.

47:03 Has anybody taken a typing class where you have to read something and then type it?

47:07 Okay. Yeah, you did that.

47:09 We can do the same thing.

47:10 We read the blog post and then we can build these scripts.

47:13 The blog post actually goes into detail when we talk about, you know, some practices that

we've actually employed.

47:19 So, you know, we actually take the weekend off once in a while, so maybe on Friday night around 7...

47:26 ...we go ahead and we script the action to stop the instance.

47:29 I don't want to think about work 'til Monday morning.

47:32 So it goes ahead. It shuts the machine down, and since I'm paying by the hour, I've now saved five hours on Friday...

47:40 ...24 hours Saturday, 24 Sunday.

47:44 I script it to start at 7 in the morning on Monday, so I saved 7 hours Monday.

47:50 And what we found is actually over the course of a year, you can save about $1,500.

47:55 So, it's really worth, you know, if you're going to go and implement this in a development state...

48:02 ...actually spending a little bit of time figuring how to do this automatic starting and stopping of your instances.

48:09 You could say, yeah, I'm going to go ahead and do it manually, but like me, we all forget and...

48:13 ...the next thing you know you come back in Monday, oops, I forget to shut it down.

48:16 So the best thing we can do is just automate it.

48:20 Now we're going to go ahead and have a demonstration that actually shows a little bit more of what we've been talking about.

48:30 So I'm here, logged in to the Amazon Web Services Management Console.

48:34 I see that I have three or four instances available to me.

48:39 Now these instances, I'm going to want to go ahead and put in place behind a load balancer.

48:45 So I go down to my load balancers, and I see that I already have one set up.

48:49 I think a lot of people already see that there's not a lot of magic or trickery going on.

48:53 It's just a simple wizard to go ahead and create that.

48:57 So if I click on my load balancer, I see some of the properties that I set up through the wizard.

49:10 Next, I want to look at my instances.

49:12 So I actually see that there's only two currently right now behind the load balancer.

49:17 I see that the status is In Service.

49:20 That means it's currently healthy.

49:23 So let's go ahead and take a quick look at them.

49:26 Now what I've done is, if I look at my description of the load balancer, here's the public URL.

49:34 So if I was to copy and paste this and go to Arc...and append/ArcGIS/REST/Services...

49:41 ...the request goes into the load balancer, and I get a response from one server.

49:46 Now to best demo this, I've taken that URL and I've created a web map from it

49:54 And to go ahead and show how I'm actually spraying out request...

49:59 ...I've colored my earth at night service to be different colors on each server.

50:03 So when it first loads here, I see that one request went to the blue server, two requests went to the green server.

50:11 I pan and zoom, and then I'm starting to see the colors change.

50:15 Okay, two requests went to blue this time and only one went to green.

50:19 It's not always very consistent.

50:21 The load balancer is, you know, built by Amazon and contains the right algorithms that it just says...

50:27 ...alright, this is the best configuration.

50:29 It's not always going, okay, your turn, your turn.

50:33 It just, you know, this one's the best available machine. I'm going to send the request there.

50:37 So there's intelligence built into it that I don't even have to worry about and/or configure.

50:43 Now I said how easy it was to go ahead and add an instance into this configuration.

50:46 I'm going to go ahead and prove it.

50:50 So all I have to do is hit the plus/minus sign if I want to add or remove an instance...

50:54 ...and know that green and glue are already there.

50:57 So let's go ahead and add another color, red.

51:01 I hit Save and thank you to that nice dialog.

51:07 Now I now see that I have three instances behind the load balancer, but the status currently says Out of Service.

51:14 I just added it in, and the health check actually needs to be performed a few times to actually determine that this service...

51:21 ...or this instance is healthy so it will be added into the configuration.

51:27 Now, here's the parameters for my health check...

51:30 ...and one of the things that I could do with my health check is just have it simple ping to, you know...

51:35 ...maybe the default IIS start-up page.

51:38 But, I want to actually test the GIS server a little bit.

51:41 I want to ping this service and make sure that that service is actually alive.

51:47 So once the service has been pinged a few times and it's determined that it's healthy...

51:52 ...go ahead and run a refresh here, and I see that red server is now in service.

51:59 So if I go back to my application here, and we do a pan, wow, now I'm getting the red map with blue.

52:07 Let's do another request, red with green.

52:11 And so I see that, you know, as it goes to the load balancer...

52:13 ...the load balancer just has that intelligence to put the request whichever instance can best handle it at that moment.

52:24 Okay. So, next, we're actually going to go back over and see how we put this into practice or into production...

52:33 ...and David McGuire is going to talk about a case study.

52:36 Hi. I'm David McGuire. I work in...at Esri Redlands in the technical marketing department.

52:42 Mostly we make demos for things like the UC or any of our conferences that we show on stage.

52:48 But for the last year, we've also been building web maps for the esri.com site to deal with any sort of disasters that go on to give information to the public.

53:00 So, earlier this year, late one Thursday night I was watching TV, and I flipped through the channels and came to CNN...

53:06 ...and saw that there were tsunamis hitting the shores of Japan, and there was a giant earthquake...

53:13 ...and I knew that things were about to get really busy at work.

53:16 So the next morning when I got in, we started to quickly develop web maps from templates we'd built for other disasters...

53:24 ...to get information out to the public as quickly as possible.

53:27 The...the map that we deploy first usually is called our new media map.

53:32 What that does is we have a map of the area so, in this case, a map of Japan that we host with a couple services.

53:39 USGS provides a lot of information really quickly out to the public through KML files...

53:45 ...and we read those files and produce web maps out of them.

53:48 So, a couple of the services that we use are an earthquake point service that has a bunch of points...

53:55 ...that tell you the size and location of earthquakes.

53:58 And so on top of the 8.9 that hit, there's actually a bunch of smaller ones that hit over the previous week...

54:05 ...and the following couple weeks after that.

54:09 And some of those were quite big, so there's a lot of 6 and 7 strength aftershocks that hit right around there.

54:16 Another thing that they provide is a shake map which is a polygon feature class that shows you how heavy the shaking was over an area.

54:23 So, you'd have right on top of the point there'd be the 8.9 and then it gets weaker as it goes out.

54:29 You could see how heavy the shaking was throughout Japan, and actually most of Japan could feel that quake.

54:34 On top of that, we also add social media to it like Twitter, YouTube, Flickr... 00:54:45

54:48 ...or at least that's when I first discovered it.

54:50 But what it does is it lets users crowdsource in information about a disaster.

54:55 So, if a street's closed or a building's collapsed or they think there's somebody trapped in a building...

54:59 ...they can put those points on a map and responders can look at those points and go and try to help those people.

55:04 So, we provide that service...or we...we load those points onto our maps, too, so people can see what's going on in the area.

55:12 In addition that, whenever you're going to use new media on a corporate website like Twitter or YouTube...

55:19 ...you've got to be careful to filter out questionable content.

55:24 So what we did is we actually created a feature class that allows you to report Tweets to our servers...

55:31 ...if anything's inappropriate and then we can filter out those users so that they don't continue to pollute the website...

55:36 ...and so that required a feature class in an SDE database in the back end.

55:41 So we got that map out an running on local machines that Friday, and I went home for the night.

55:47 Little bit after I left, the...somebody at Esri decided to send out press release that, oh, hey, we have this great map on esri.com...

55:55 ...you should look at it.

55:56 And Saturday morning, CNN found our press release that, hey, this is a great map...

56:00 ...we're going to link to it from cnn.com.

56:03 CNN gets a bit more traffic than esri.com does, so the traffic started to pour in.

56:08 Our poor local servers just couldn't handle the load, and they started to die, and we needed to scale up fast.

56:15 So, if you look at our local layout, we just had one little ArcGIS Server and the SDE behind it...

56:20 ...and our users were coming in and pounding those servers, and it just couldn't take it.

56:25 So, we needed to move to the cloud.

56:28 I took the two AMIs that Esri provides from scratch, spun then up, authorized server and SDE.

56:37 I also created a load balancer to get ready for the staging environment.

56:42 So, the good thing about load balancer is they provide an elastic IP, so when you're going to do scaling...

56:47 ...you can set your web apps to go through that elastic load balancer, you don't have to worry about changing the IP address...

56:51 ...or applying elastic IPs to every single one of your ArcGIS Server instances.

56:55 The SDE instance, on the other hand, we also had to create an elastic IP...

56:59 ...because all those servers need to be able to find the same SDE instance.

57:03 So, once we got everything done, we needed to figure out how to scale it.

57:08 We created AMIs out of the two instances, and then we needed to figure out what we want to do.

57:13 And, like I said, with the SDE instance, actually, so what we decided to do was not just scale horizontally...

57:20 ...we also needed to scale vertically.

57:22 We decided to do both.

57:24 The server instances were fine.

57:26 All of the earthquake feeds and shake maps all used a file geodatabase so they could be replicated across any number of servers...

57:33 ...and it would always be the same.

57:34 With the feature service that we were using to report points, they needed to be in an SDE database.

57:41 And I'm not much of a database guy, so I don't know how to set up replication across databases, and I'm in a hurry...

57:47 ...so I just decided to scale that machine vertically, and we applied elastic IP to it.

57:52 So after everything was staged, we created our AMIs, we launched into the...or the production machines.

58:00 Originally we launched two ArcGIS Server instances out of the server instance, and we just scaled up the SDE instance to an extra large.

58:09 Once those were launched, all in a load balancer, we redirected the HTTP from our old machine to point to...

58:16 ...from our local machine to point that load balancer and esri.com picked all that up automatically.

58:23 So, when they'd go to esri.com, it would automatically redirect them to our new load balancer.

58:27 And that allowed the servers to...or didn't allow...so it took the drain off of our servers, put it in the cloud.

58:35 Amazon could handle the load.

58:36 Actually, two servers started to bog down.

58:38 We decided to spin up an additional two servers on Saturday, so we were up to four ArcGIS Servers...

58:43 ...and that took us about two hours from start to finish to get into the cloud, all of our services up there and running...

58:48 ..and off of our servers.

58:52 So, that survived for the rest of the weekend.

58:55 Monday morning we came in and Al Jazeera had also found our web map and decided not only are they going to link to it...

59:01 ...they're going to embed it directly into their page.

59:04 So, load spiked again.

59:06 We increased the size of the servers to six and handled the load, no problem.

59:11 We doubled the size of the SDE server.

59:13 I think we were up to eight cores by the end of it, just to deal with all the requests.

59:17 So, it worked out really well.

59:19 We used Amazon to solve our needs.

59:22 You might be wondering why we didn't just use automatic scaling.

59:27 Well, as Hugh was showing you with the health checks on the load balancer, he was able to look at just one service...

59:32 ...and make sure it was running.

59:33 We had multiple services running.

59:34 We wanted to make sure they were all running. We wanted to make sure they were all running.

59:36 So we actually developed a script in-house to go out and ping each individual service, tell us if it's running.

59:41 If it's not, it will actually text message me to tell me it's down.

59:46 Or, you could set it up to actually try to log in and restart services for you.

59:51 But to do that, we have to know what machines are running.

59:53 So if we turn on automatic scaling, it wouldn't be able to...or we wouldn't know when Amazon spun up a new service.

59:59 We wouldn't be able to put those monitors in place to make sure that everything was running behind the scenes perfectly.

1:00:05 And we wrote down the entire process in an ArcGIS Server blog, so you can just search for my name on the ArcGIS Server blog...

1:00:11 ...and you can see the blog I wrote.

1:00:14 It includes all the steps we went through to build the instances, links to our resource center on the exact instructions...

1:00:21 ...and a link to that script that we used to monitor.

1:00:23 So if you want to use that to monitor your own ArcGIS services...

1:00:26 ...that's free and available for everybody to use on our Resource Center as well. That's pretty much it.

1:00:33 Okay. Thanks, David.

1:00:35 To wrap up, we'll just talk about some of the common Amazon EC2 questions and issues, and then we'll take your own questions.

1:00:44 The big cost that we...or question that we're seeing is, how does the billing work?

1:00:50 There're really two separate costs that you need to consider independent of each other.

1:00:54 One is the Amazon EC2 billing.

1:00:56 Amazon bills you directly with the information that you give when you set up your Amazon account.

1:01:01 Also, in the management console, they provide a detailed summary of all the hours you've used in every region...

1:01:06 ...exactly what you've done and how your bill stacks up.

1:01:09 So that's the Amazon cost.

1:01:12 The other side of it is the Esri licensing.

1:01:16 You must license all the software that you cause to run on Amazon EC2.

1:01:22 You can use the licenses that you...you can take licenses that you ordinarily use on premises...

1:01:26 ...and you can move those and start using them in the cloud if you wish.

1:01:31 ArcGIS Server is licensed based on CPU cores, so you must not exceed that number of cores for which you are licensed...

1:01:38 ...once you add together your cloud and on-premises cores.

1:01:42 So, it's...it's a pretty simple formula.

1:01:45 If you have questions about that, you should talk to your Esri representative.

1:01:50 The ELA enterprise license agreements, are obviously the most flexible for this.

1:01:54 They allow you to use essentially an unlimited amount of software on Amazon EC2.

1:02:00 So if you were to set up the autoscaling, you could feel comfortable with that if you have the ELA.

1:02:06 Really, you need to remember the Amazon EC2 is billed like a utility.

1:02:09 If you leave the water running all the time, your water bill's going to be high.

1:02:12 Same is true with the instances, and we already talked about that.

1:02:15 One thing is, if you're going to leave the server on all the time and you know that it has to be available 24/7...

1:02:22 ...you should consider using what Amazon calls reserved instances.

1:02:25 This is a way that you pay at one time a flat fee and then you have the ability to have that instance for one year, three years, et cetera.

1:02:35 And then your hourly fee is very low after that.

1:02:38 So it's a one-time fee plus a low hourly fee that you pay from there on out...

1:02:42 ...and your...and you do it by how many years you want to have that reserved instance.

1:02:47 There...on the Amazon website, they have some cross calculators that are very good for anticipating...

1:02:54 ...or guessing how much you'll need to apportion in your budget and how much your resources will be able to get you.

1:03:02 Big question that we get asked is , well, will Amazon EC2 save me any money? Is it cost-effective?

1:03:07 And I just put this CFL lightbulb on here as a response to that.

1:03:12 I can't give you the answer yes or no, but I can talk about these lightbulbs.

1:03:15 If you use them, you could potentially save money over time. They're more energy efficient.

1:03:21 However, if you were to put hundreds of these in your house everywhere, you would not save money...

1:03:27 ...and if you left them turned on all the time, your electric bill probably wouldn't be any lower than it is...is today.

1:03:34 So you just need to use it wisely and plan well and I believe that in many scenarios it can save you money.

1:03:39 Usually organizations...well we hear the phrase, should I move to the cloud, a lot.

1:03:44 And it's not so much an issue of packing up a moving truck and taking all your stuff and putting it in the cloud.

1:03:50 Rather, you use the cloud occasionally in focused scenarios where it makes sense for you.

1:03:55 You know, at Esri, we still host a lot of stuff in our on-premises services...on our on-premises servers...

1:04:01 ...but occasionally when it makes sense to use the cloud like in the scenario David described, we go ahead and use it.

1:04:06 So it's just one more...one more thing you can draw on, a weapon in your arsenal...

1:04:11 ...that you can use when appropriate to handle your requirements.

1:04:16 A lot of people ask about security and, really...really, there's kind of two sides to this.

1:04:21 There's the things that Amazon provides and that they're responsible for.

1:04:25 And then there's the things that you are responsible for.

1:04:28 So Amazon provides very secure facilities for their data centers.

1:04:31 They don't even reveal the precise location of these data centers.

1:04:34 And if you tried to drive into one, you'd be stopped right at the front by a guard, and they are highly secure in that way.

1:04:41 I would argue that they're probably more secure than your own server room...

1:04:44 ...unless you work for the military or some institution like that.

1:04:48 We also saw how the Amazon puts a firewall on each instance that you control.

1:04:54 In order to get into that management console where you start launching instances...

1:04:57 ...you have to provide your Amazon account name and password.

1:05:01 And if you want to get even more secure, Amazon provides an option where you can have a little physical device...

1:05:06 ...where you click a button and it shows you a code, and every minute or so, that code changes...

1:05:11 ...so that's called multifactor authentication, and that's available from Amazon...

1:05:14 ...as an extra layer of security for logging into that management console.

1:05:18 You certainly don't want to give that name and password for that console out to many people...

1:05:22 ...unless...unless the people are well trained for use with Amazon EC2.

1:05:26 That could result in disaster if somebody got in and started terminating things who didn't know what they were doing.

1:05:33 It could also result in easy sabotage.

1:05:36 Finally, we saw how you have to get that key pair file in order to get the administrator password for your instance.

1:05:42 So even if you mess up and you leave the console open and you allow any IP address to use remote desktop protocol...

1:05:51 ...the people still need to know the admin password in order log in to the machine.

1:05:54 So if you can put all those factors in place, it results in better security.

1:05:58 You are responsible for keeping those names and passwords secure.

1:06:02 You can also, if you want, you can set up ArcGIS Server security on these instances...

1:06:06 ...just in the exact same way that you would set it up on-premises...

1:06:10 ...if you want to lock down access to certain services and whatnot.

1:06:16 You saw that when I was launching an instance, maybe you saw that [Unintelligible] _____ AMI.

1:06:20 That's for 10.1 that we're working on.

1:06:23 That's pretty exciting because ultimately the Linux instances can be cheaper on Amazon EC2...

1:06:28 ...so that could be a way to costs is to run on Linux.

1:06:31 Also 10.1 ArcGIS Server on Linux will be a much better...a much improved product.

1:06:37 We have new architecture on 10.1 for ArcGIS Server that's been detailed in the road-ahead sessions.

1:06:44 That allows the GIS server tier to have a little more communication with each other...

1:06:48 ...and it simplifies ArcGIS Server and ultimately we're working on making it faster, so that's available in the AMIs.

1:06:55 And then we'll...we will provide you a little utility app for building a large site on Amazon.

1:07:01 That can be tricky sometimes when you spinning up many instances.

1:07:04 You want to make sure you get all the elastic IPs in the right place and the load balancer...

1:07:08 ...and so this utility will allow you to go through it one time and put how machines you want.

1:07:12 You can even set up autoscaling rules with it.

1:07:15 So we're excited about that option.