Piloting the Cloud: A Technical Tour with Walter Myers III

https://mindmatters.ai/podcast/ep337

Robert J. Marks:

Greetings and welcome to Mind Matters News. I'm your host in the cloud, Robert J. Marks, because today we're going to talk about the cloud where you probably have some of your files and photos stored.

The cloud might seem mysterious to the unenlightened, almost magical place. It's almost a magical place where files, photos, and programs somehow float in the air instead of being stored on a computer. But where exactly is the cloud? Is it in the sky? How does it store my files? If I can't see it, is it safe? Can I lose my data if the cloud disappears? You always need an internet connection to access my stuff. How is it different from saving files on my computer's hard drive?

The idea of files existing somewhere outside of a physical device can be confusing and mysterious and users might worry about security, privacy and how to retrieve their files if they switch devices. And I'm one of those.

I personally use Microsoft OneDrive as my cloud. It's great. I can access my pictures and images from different devices. This is not an endorsement of Microsoft OneDrive. There are many great services out there, and I use OneDrive because it was introduced to me by my son, Jeremiah, and it works great. And I got to confess, I have not taken the time to shop and compare, but the cloud has become my friend.

So to talk about the cloud, our guest today is Walter Myers III. He is a principal engineering manager, leading a team of engineers, working with management of the Microsoft Azure Cloud, and Azure is A-Z-U-R-E. I had to look up how to pronounce it. So it is Azure. He's going to educate us about the cloud. He holds a master's degree in philosophy from Biola University's Talbot School of Theology, where he is an adjunct faculty member in the Master of Arts in Science and Religion program.

He teaches on Darwinian evolution from a design-centric perspective. He is a former board member of the Lincoln Club of Orange County, where he served as programs chair from 2015 to 2018. And personally, Walter fights to reduce the power and scope of government. He's a big advocate for free markets, lowering taxes, and more freedom for all. He's a regular contributor to Discovery Institute's Evolution News and Views website, and he writes on parallels in the design of a computer system and biological organisms.

Walter, we've got to get you writing for Mind Matters News, now, from the technical side.

Walter Myers:

I'd love to, Bob.

Robert J. Marks:

Okay, that's great. Well, welcome. By the way, this is something interesting, Walter, I am Robert J. Marks II. You are Walter Myers III. And just like royalty of old, we place Roman numerals after our names. I've tried to update my name recently by calling myself Robert J. Marks 2.0. I think I'm just suggesting this to you because you could update your name to Walter Myers 3.0, so you're an advanced version of what came previously.

My parents, interestingly, didn't want to name me Robert J. Marks Jr., like we have in the news today as Robert F. Kennedy Jr. because where I came from, people started calling you junior, and I had an uncle

that died when he was 90 years old. And even when he was in his late eighties, they called him junior, which seemed kind of strange. So I'm blessed to have the second.

So what do you think of Walter Myers 3.0? Do you think that that would update your persona at all?

Walter Myers:

Well, I'm actually, I'm thinking about Walter Myers 4.0, which will be an AI, who will represent me and actually be perfect.

Robert J. Marks:

That's great. Okay. Well, if we repeat a podcast sometimes, maybe we can talk to Walter Myers 4.0.

Walter Myers:

Hopefully in the near future. Well, if Ray Kurzweil is correct, it's going to be in the very near future.

Robert J. Marks:

Well, Ray Kurzweil is not correct. It isn't going to happen. You know, what was it, 20 years ago, he wrote a book called The Singularity is Near. That was about 20 years ago. And so last year, nothing happened. It got better. It got better.

So he wrote another book called The Singularity is Nearer. And so in 20 years he's going to write another book said The Singularity is Just Around the Corner. It's just almost here. But no, I don't think it'll ever happen.

So let's talk about the cloud.

Walter Myers:

All right.

Robert J. Marks:

This is interesting, and actually, you and I have chatted at some conferences about this, and I must admit that I was uneducated about the cloud. So kind of tell us where exactly is the cloud and how does it physically exist? Where is the cloud? Is it in Detroit, or Michigan, or where is it at?

Walter Myers:

The cloud's all over, Bob. All over.

Robert J. Marks:

Okay.

Walter Myers:

Well, really the cloud, it's just a representation of physical resources. In other words, data centers, being these huge data centers with literally hundreds or thousands of servers that are accessible through the internet, which itself is a global system of interconnected computer networks that allows communication between various devices, computers, mobile phones, whatever, around the world.

And sort of that concept of the cloud, why is it called the cloud? It dates back really to early network design. When engineers would map out various components of their networks, these are early days

before in the internet, what they would do is, they would communicate with these other things that were sort of outside, but they didn't know what these networks were necessarily, that they may be communicating with. Right? So they just kind of made it this sort of rough blob of undefined devices and it kind of looked like a cloud. So that's where the name was, how the name was adopted. Right? As the standard image for representing computers outside of one's own network. So outside of your home network, outside of a business network, that's the cloud. It's somewhere out there. You don't know where it is, where those data centers may exist, but you can gain access to those through the internet. So that's where the concept of the cloud comes from.

Robert J. Marks:

Okay. So where is the cloud? You say it's distributed, so is it distributed all over the country? Or are there places where the cloud is kind of concentrated, say in different cities and locales?

Walter Myers:

The cloud is distributed all over the world. In other words, anything that you're accessing, in other words, some particular service that's out, some particular... For example, a good example of the cloud accessing it is, you use Microsoft's OneDrive, or if you use a Microsoft's Outlook, right? That is a distributed set of computers around around the world. And that's not specific. That's not Azure, okay? Because Azure is our public cloud, as well, as AWS has a public cloud, Google has a public cloud.

But other than the public clouds that we'll discuss more, is that we have globally distributed systems. For example, when you authenticate into Outlook, live.com, right? That is a global distributed network of data centers, data centers that coordinate with one another, for your authentication purposes. That's not in any particular region or any place, but it's one of those 300 some odd data centers we have around the world, for our customers to be able to authenticate into our software services.

So let me maybe just talk a little bit about what that is.

Robert J. Marks:

Good. Yes.

Walter Myers:

In the cloud. So when you go to outlook.com or you go to OneDrive, that is software as a service; SaaS. Okay? So we have that concept. The other concepts we need to think about as well as on premises. So in other words, a customer, instead of having their work on, in other words, start running their servers on premises, they might say, "Hey, we want to move to the cloud." What that means is they want to move to the Azure Cloud and they want to take advantage of infrastructure as a service.

Now infrastructure as a service means you don't have to worry about the capital costs of setting up servers and all those types of things. We'll do the servers for you and you can access those servers through the internet. It's a virtual machine. It's actually a virtual machine sitting on top of a real machine, right? But it's virtual, virtualized and protected that only you can access, that infrastructure as a service, and you can access that like any other server as if it were just on your premises but it's not.

Then there's platform as a service; PaaS. Platform as a service goes that sort of extra step where you don't even worry about the hardware. You don't have to patch it. You don't do any of that stuff. You just basically, but that's where you have to sort of write to a certain programmatic constructs that we've created, where you'll be able to deploy your application, but all you worry about is your application. That's it. You don't have to worry about any underlying patching or anything of the operating system.

And then you have, again, getting back to SaaS on that sort of continuum. So there's a continuum of how much work you have to do. With SaaS, you don't have to do anything. It's already there. You do your email or you upload your files onto OneDrive. That's software as a service. You don't even have to worry about it. We handle everything for you.

Robert J. Marks:

Okay. So I guess what you're saying is that the cloud is in a bunch of different cities all over the place, and wow, that seems to be very confusing in terms of a traffic jam if you're trying to access what you're doing, but I guess you got these routers that kind of take you to the right place to get your stuff. Is that right?

Walter Myers:

Yeah, and that really is the promise of the internet, Bob. Is that the fact is that, and the idea of the internet came in terms of what would happen? Originally for the internet is, what would happen if you have different networks that need to talk to one another. What would happen if there was a nuclear bomb and it blew up one particular pathway or one, you know, some wire that are running underground, right?

If you say, have a nuclear blast in Detroit and it blows up basically some data center, or it blows up some connection point, well, what happens? Well, that's the whole thing with routing in the internet, is that it'll find and it will find the closest path, right? So if you have all of these different interconnected computers with these cables running that are connecting all these things up, if one of those paths gets chopped, it'll find some other path. And so that's the wonder and the beauty of the internet is that it's fully distributed, geographically distributed that. And so if one or several paths somehow are affected, they will find some other path for your computer, to get to the other computer, that has the information that you want to access.

Robert J. Marks:

Okay but what about data? If you have a center blown up where you have all of your files stored, that's not a matter of finding paths. You've destroyed the information that you've had. I kind of assume from talking to you before, that there's a redundancy in the storage of your information in the cloud. Is that right?

Walter Myers:

Yeah. Yes, there is.

Now I know really about the redundancy with respect to Azure, and I'm sure it's very much the same as our distributed networks.

But for example, in Azure you select a region is where you deploy your application, your software, and that goes into storage. So you have storage and compute. You have storage, compute, and you have networking.

So in storage, we keep three copies of that in each of our data centers. I'll talk a little bit more about regions, but in other words, you select the region. And a region may be made up of a single data center. Those are our older models, but we're changing. I'll talk a little bit more about about how this works. It gets a little complicated.

So in our older regions, we had one data center, but we would copy your data, so it would be in triplicate. That way if a rack went down, or a rack went down in that data center, then we could access

the other two copies. So if that one goes down and say we pull that rack out, and we pop another rack in of servers, we still have those other two, and then it'll replicate when that rack gets put back in that was damaged, or that damaged rack gets replaced, then would have already copied it to some other rack, so you'll always have your data in triplicate and that's the redundancy within that data center.

Now we also have the concept of availability zones. So the way that we're building data centers now is there will be three in a region. There will be three actual physical data centers that are connected up. And so you can basically, and they'll be able to talk to one another within say, a 10 or 15 milliseconds. So it's physics, right? You have to have them close enough so they can talk fast enough so that if, say for example, we have a region in the San Francisco area. If there were, say an earthquake in San Francisco and it damaged one of those data centers, the chances are it's not going to damage all three data centers within that region.

If that data center goes down and the customer selects, you don't have to do this, but you can say, "I want to be, I want to use the availability zones. I want to be replicated across those three data centers, so that if any of those three data centers go down in that particular region, I'm up and running on another one of those regions." So then you'll have two, and then if another one goes down, another data center goes down, a second one, the third one will continue running and you get even lower likelihood that it would be that due to actual data centers or a region would go down. That's typically how it works.

We also have the concept to make it more a little bit more complicated of pairings. So for example, we have pairings between regions. So say for example, so we have a west US two region and we have an east US two region, okay? You can actually, the customer can select, I want my data to be replicated across those regions. So if a whole region goes down, I can be up and running in another, in that regional pair.

So that's another thing that we have with respect from a disaster recovery perspective, is if you go down in a particular region, you can have to make it so. You have to select this option, but then you can be running in another region entirely.

Now we will move shuttle all of your data across, and then you can actually have your application running in both places and replicating data across those regions, or you can have it sort of cold where you can say, "Hey, if I'm down in one region, that region is declared down, I can then deploy my application to the other region. The data's there because I've replicated all the data. And then I can start." I'll have some downtime, obviously in that case when you have a disaster. But I can be up in a four to six hours, all depending on what my recovery time objective, my RTO is. Whenever you hear RTO recovery time objective, it says that if there's a disaster, how much time do I have to bring my system back up someplace else so that I can continue running?

Robert J. Marks:

So I do have a question. I did an interview with somebody that was talking about EMP, Sarah Sigman. She is an expert on electromagnetic pulses, and there is this idea that if there was a thermonuclear explosion about the elevation of the space station, that it could take out most of the grid, the power grid in North America, which of course would be disastrous for the cloud, but also would it not take out some of the computers on which this stuff was stored?

So I'm wondering, are you aware that any of these centers are covered with like Faraday Shields or something, which would protect them from an EMP?

Walter Myers:

I'm not aware of that and I don't think our data centers have gone that far with respect to being immune from EMP. So yeah, don't know about that, but I don't suspect so from what I understand of our data centers.

Now, what I will say about our data centers from a power perspective is, we don't deploy data centers where we're not virtually guaranteed that there will always be power. So for example, typically you want your data centers someplace, for example, being powered by hydropower, right?

Robert J. Marks:

Yes.

Walter Myers:

So if in fact, there's one of our data centers in Quincy, Washington, and that's fully, it gets all of its power from a hydroelectric power plant. So the thing is, you know that in this particular case, that water's going... At least the water's going to continue to flow, right? So we always try to find those places that are geographically stable and have a consistent source of power, hopefully hydropower.

And as you've probably seen in the news, there's going to be starting up a new plant on Three Mile Island. They're going to be starting up one of those reactors so it can power our data centers. And I think you're going to see a lot more of that over time, because nuclear itself, it's clean. Okay, you do have nuclear waste you have to deal with, but it is clean and it's highly efficient. So we'll probably be seeing a lot more of these data centers powered by nuclear power over the coming years.

Robert J. Marks:

Yeah, I was reading on the web, so it must be true.

Walter Myers:

It must be true.

Robert J. Marks:

It must be true. Yeah. So that Google was thinking about acquiring or building a nuclear plant, to generate the power for their AI and the other stuff that they're doing. That's really amazing the companies now are actually generating or constructing their own power plants.

Let me ask you about your batting average. Are you batting a thousand? Have there ever been files which have been lost in the cloud?

Walter Myers:

Not to my knowledge. We've had other things happen, but not to my knowledge that we've actually had customer files lost, in either Azure or in our SaaS, distributed offerings with respect to SaaS like Outlook or OneDrive. Not to my knowledge.

Robert J. Marks:

Okay. Well, that's good news. What about security? You talked about natural disasters and protecting from natural disasters. What about cyber attacks and such? I would think that America's adversaries would be very interested in attacking the servers that are in charge of the web. And I'm sure that, I would assume that you spend a lot of time making sure that that doesn't happen.

Walter Myers:

Yeah, we spend a lot of time making sure that doesn't happen. But what we do in the Azure Cloud is we have what's called a Zero Trust Model. So the Zero Trust Model strategy, it reduces the risk of data breaches by validating authenticity at every step, believing that every individual, regardless of his or her organization, is a potential threat to security unless they can prove otherwise. So basically to be able to access any data where we're authenticating through every step.

And we also have this thing called sort of defense in depth, which is sort of a seven layers. There is basically there's a physical security, you know, in terms of being able to gain access or biometric access into our data centers. There's identity and access. So Azure Active Directory are now called Intra ID. There are other authentication providers like Auth0 and Okta, where that, in order to gain access to those resources, you must go through those checks. You must be authenticated and then there will be some, there'll be authorization in terms of once you're authenticated, what are you authorized to actually to access?

On the perimeter, we have a DDoS service, the distributed denial of service, distributed denial of Service; DDoS. So that's to basically protect our customers from someone trying, some bad actor trying to just basically wear down their website so that people can access it.

The other thing we do is at the network level. So at the network level, there's security that we apply that to be able to only allow people to gain access to certain resources at the network level itself. Also with application, there's SSL TLS encryption. So basically we encrypt. We encrypt that traffic over HTTPS, that's going against that particular data resource. And also your data at rest is encrypted. So we use all the modern constructs, public key, public key security, with respect to making sure that your data is protected and only you can gain access to it.

And one thing, Bob, that people have always been worried about in working with the cloud is, can someone else access my data?

Robert J. Marks:

Yes.

Walter Myers:

And the thing is, no. No one else can access your data. When your data block is written for you, and that's marked as no one else can access your data.

Also, no one can walk into a data center, pull out a drive and say, and grab your data because your data is so smeared out across multiple hard drives in any particular data center, no one could pick up any particular hard drive and find anyone else's data because it's so smeared out across the drives, in terms of it's a 1960s technology actually, in terms of how the data is written, and how the blocks are marked so that once you delete some data, you're not writing data to the same place. You are continuously writing, and then you're marking blocks for deletion, and those blocks that are marked for deletion, some other data gets written over that. So it's very sophisticated in terms of how that data is written, how it's secured, and how it's encrypted, even within Azure storage.

Robert J. Marks:

Wow, that's fascinating. I did consulting for Microsoft when I was at the University of Washington. And first thing we did when we went in, the first day I had to sit down with lawyers, believe it or not. Lots of lawyers at Microsoft, as you probably know.

But I went into the office of the guy that I was doing the consulting for, and he had this little counter, and this counter was doing click, click, click. And I will admit to my age, this was around Y2K, when the calendar was turning around to 2000. Everybody was saying that all of the stuff in the universe is going to fall out because the way that we kept track of years before that. And I asked my friend, a guy named Mark Caseball that worked for Microsoft, I said, "What is that clicker?" He said, "That's the number of times that people are trying to attack and breach and cyber attack Microsoft."

And they actually had a little counter up there. They counted all the times that they were being attacked. I found that really, really interesting and kind of a little bit gutsy to actually say, "Yep, we're being attacked. Nope, didn't work. Yeah, we're being attacked again. It didn't work." And while I was there, it went click, click, click. I mean, it was kind of a continual thing because all of these hackers wanted to claim coup, if you will, for breaking into Microsoft, but it never works. So what you guys are doing seems to be working pretty well.

Walter Myers:

Yeah, I mean, the thing is, software is written by humans, and humans are not perfect. So there will always, always, always, when you're writing software, there will always be vulnerabilities in software. We are always finding vulnerabilities in software in Microsoft. We have whole divisions that all those divisions do is look at what is risk to our Azure platform and we try to patch those holes. Wherever they are, if we find them or someone else finds them. We patch them as quickly as we possibly can.

So yeah, their full division. So that's all they do is really manage, in fact is manage risk. And so DDoS is really the biggest concern. Basically bad actors trying to get in and poke in and stop people, or they're trying to hack in. I mean, whether it be Russia, China, Iran, or whatever. So that happens every day. And the thing is, you have to stay a step ahead, and that's why we have the Patch Tuesdays, and we even have internal, we even have internal patching.

Robert J. Marks:

Wait, you have Patch Tuesdays?

Walter Myers:

Yes. The first Tuesday of every month is called Patch Tuesday.

Robert J. Marks:

Okay. Wow. Okay. Do you all get together and look for patches? Okay.

Walter Myers:

Yeah. We get together and sometimes we'll patch those ahead of Patch Tuesday, internally. We'll roll those out to the public. So a lot more is going on under the covers than customers may recognize, as far as us keeping our system safe.

Robert J. Marks:

Oh my gosh. Do you ever hire hackers? I'm thinking of this movie, Catch Me If you Can, where Frank Abagnale, I think his name was, went around and pretended he was an airplane pilot. He pretended he was a surgeon. He was pretending. And then once he got caught and served his jail time, the FBI got him out and used him as a consultant for minimizing bank fraud because he was an expert in it. It seems like you could do the same thing.

I just wonder if you're aware of that ever happening at Microsoft where they hired these hackers that kind of got close, but no cigar, and they hired them to come in and look at the vulnerabilities because they looked at it from the other side.

Walter Myers:

I'm sure we do, but they certainly don't let me know about the stuff like that in my-

Robert J. Marks:

They don't. Okay.

Walter Myers:

Which is good for good reason.

Robert J. Marks:

Oh yeah. Well, I imagine. I imagine. Anyway, that's fascinating.

One of the things I wonder is about scaling, and let me go to an example, Google, YouTube. I forget how many uploads they have a day to YouTube, and it just gets bigger and bigger and bigger. And I don't think the rate of increase is decreasing at all. I think it just keeps getting bigger and bigger and bigger and bigger. And I don't know where they're going to find all of the place to store this stuff. And I wonder if that's a problem that you are going to have as you begin to scale up. Where are you going to put all of this stuff?

I talked to one guy, he was doing kind of cutting edge, you know, think tank sort of things about the future. And he said one of the things that they're looking at right now is storage in DNA; a DNA sort of storage, biology sort of storage, because that is the most efficient way to store information that mankind knows about.

But that aside, how do you keep up with all the storage because I know that it's increasing?

Walter Myers:

Well, yeah, the storage is increasing, but DNA, at best, would be what we might call cold storage, because it can store massive quantities of data, but that data is not readily accessible.

Robert J. Marks:

It's not readable. Okay.

Walter Myers:

No, readable, but it's not readily accessible. It would take time to write, and the read times will be very slow.

So typically what we do is we look at data from the perspective of data being hot, warm, and cold. So DNA would be an example of cold data, where you're saying you're storing massive quantities of data that aren't accessed very often. Okay.

But say if you have a running website, you want it to be fast. So you're probably going to want to run your web. Now you can run your website on, you can choose different... We have different types of storage. We have storage that's solid state drives, SSDs. We have hard drive storage, and then we have some archive storage also.

So you can say, "Hey, I want this data to be cold," and we'll charge you less for it because we're using technologies that are not... The faster the technology, the more it's going to cost.

Robert J. Marks:

And I bet it's probably more efficient than those storage mines that Elon Musk just found. Right?

Walter Myers:

Oh yeah, they did. That's your government at work, right?

Robert J. Marks:

That's your government at work. That's right. That's cold storage.

Walter Myers:

Definitely cold. That's freezing storage.

Robert J. Marks:

Freezing storage. Okay. Go ahead. I'm sorry.

Walter Myers:

Yeah, so basically the hot storage would be SSD, and you pay more for that SSD. So then we have other storage that's your conventional hard drives, which are slower, but it's still pretty fast. And then we'll have other types of storage, just some of the older storage that'll be more that sort of warm, that warm storage, if you will. So, that's how.

Now the other thing is, storage density is getting greater and greater as we go along. So you think these SSDs are packing more and more data. So I don't think we've gotten to a point yet where we're saying, "Oh, we're at a crisis of, in terms of how much data we can store." Because hey, there's lots of Silicon, right?

Robert J. Marks:

Yes.

Walter Myers:

Silicon. We can definitely scale. And in fact, when people think about the cloud, we say that it's infinitely scalable. Well, nothing's infinitely scalable, but from the perspective of the people who are maybe concerned about how they're going to store all this stuff, it is in a sense, from the perspective of the user, it is in a sense infinitely scalable, even though it's not truly infinite.

Robert J. Marks:

Okay. So the storage is getting better and better. Okay.

Walter Myers:

Getting better and better, and higher and greater and greater density. And now that we're with SSDs, you know, there's no moving parts. Right? Whereas hard drives, eventually they're going to degrade over time. But really everything is moving more towards solid state drives now; SSD.

Robert J. Marks:

Yeah, I'm thinking of the computer I had a few generations ago. I turned it on and at the end of its life, it sounded like a cement mixer with all of these things going on.

Walter Myers:

Yeah, the needle moving around on the drive and it's crazy.

Robert J. Marks:

Yeah, exactly. Exactly. But today it's fast and it's efficient. Really, really amazing.

So let's talk about the reliability of storage. I'll tell you something. My wife comes from old school and I had the taxes done, and she says, "Well, print out a copy." I said, "I don't need to, honey. I have an electronic version on my computer. It's in the cloud. It's fine." She says, "Yeah, but Bob, what if that goes away?" Okay. And so there's this paranoia. So I printed it out and it's sitting over here on my little refrigerator that I have in my office, and she's never looked at it. But nevertheless, I have that in order to give her a degree of security.

Let's talk about the reliability of different storage. You have your flash drive sort of storage. You have your CDs and DVDs; they're writable, read-write CDs. And then you have the cloud. And then you have your computer. And I will tell you that I am not a fan of read-writeable CDs. I stored a bunch of stuff on read-writeable CDs, and I just went back, you know, like a year later and it was gone. I don't know where it went, but it's just not a reliable place of storage.

So can you comment on that? You know, as far as storage on the cloud, on your computer, on flash drives, on hard copy, et cetera?

Walter Myers:

Well, I mean, we don't use any CDs or DVDs, so I can tell you that, in cloud.

Robert J. Marks:

For good reason. Yeah.

Walter Myers:

For good reason. Well, I mean DVDs are, it's estimated that DVDs themselves can store up to a hundred years. That's an estimate. CDs, it's typically more like 10 to 20 years that you can trust them.

Robert J. Marks:

Really? Why are DVDs more reliable?

Walter Myers:

I don't know. Well, I think it has more to do with the type of materials that are used in the DVD versus a CD. And I think what happens over time is you get, oxidation is probably the biggest thing. Right? It's oxygen. Right? And so it all depends on what that storage is actually, what elements or molecules that storage is actually made of. Right? So oxidation is probably the biggest thing.

Also, if it's out in the sun, which is not a good idea because then you can, ultraviolet can actually affect DVDs or CDs. So if you're storing it somewhere in some nice dark place, it still can oxidize over time, so that's probably the biggest enemy of that type of storage.

But cloud storage today is, it's hard disk drives and every day, and we're moving more and more towards solid state drives. I don't believe 10 years from now there'll be any hard disk drives. I believe it'll be all solid state drives. And of course the density, the density of the memory, will continue to increase over time.

Robert J. Marks:

I believe that there's probably also degradation on the cloud, and I suspect that you probably have some sort of reliability and maintainability sort of program in place, to make sure that the electronics don't wear out.

Is that something which is addressed at all?

Walter Myers:

Well, yeah, I mean, the storage and the computer are continually monitored. And the way it works now is now you do have individual units. So you might, for example, in a rack, you might have say what we call a 1U. Right? So within a chassis. Okay? So within a chassis, the 1U just means one unit. So within a chassis, you might have 12 units. So let's say something goes bad in that chassis. We may just pull the whole chassis out and pop the chassis back in.

Then you have a rack. And a rack might have three or four actual chassis with 12 units in it. So that's 12 times 4; 4 times 12; 48. So that's how we sort of have that sort of density. But they'll be sharing power and they'll be sharing power. And they'll be sharing the fans, the cooling, right? So that'll be in the rack, or in the chassis itself.

So we can pop out. In other words, so we have a rack that has like four chassis in it that has, what, 48 units. Sometimes it can be up to 96. You can pull that whole rack out. So it used to be on a rack boundary. So basically when Azure first started, you would say if you wanted to make things sort of your redundancy was, say if I had a piece of your data on rack one, I would copy it to racks two and three. Right? So if rack one went down, I would guarantee you that I would keep it across those three racks. And there was lots of racks in the data center. So if one rack went down, right, well, I've got your data in two other racks that have been replicated, and I'll pop that rack out and pop another rack in.

So yes, things do degrade, and when they go down, you know from the monitoring that we see, we just pop that out and pop something back in, and it's all of the data centers. Everything is pretty much modular in there. You pop chassis, or pop racks in as units, as opposed to going down to an individual hard drive or individual unit, individual compute and storage unit. You don't have to do that. You can just pull whole sections out because everything is replicated and it knows how to piece those things back together, as far as the replication goes. Right?

A rack goes out that has your data on it. Then when we pull that out, the system understands that and says, "Oh, we only have two copies now. We need to go ahead and replicate on another rack."

Robert J. Marks:

That is really-

Walter Myers:

Interesting stuff.

Robert J. Marks:

Yeah, it is. I went to a friend's house in Switzerland; Hal Phillips. He was a, he was a guy that I've worked with in the past and really like, and he's on top of stuff. He's a guy by the way, that invented the touchscreen. And he's doing okay, but he had something in his house called RAID, if you ever heard of that. It's what you're talking about sounds a lot like RAID. It's where you have these different racks and you have different... I'm lost with the terminology, but you have these different inserts and they're redundant. And if any one of them fails, you can take it out and you can stick a new one in and it uses error correction coding in order to restore the unit that failed, that you took out and replaced. And I thought that was really fascinating, and it sounds like you're using a lot of the same sort of error correcting code in order to do that, where you look at the redundancy, there's a redundancy in the code. So that's really good stuff.

Walter Myers:

Yeah, I'm not sure, Bob, that's quite how we do it in cloud data centers, but what you're talking about is redundant array of individual drives. I'm not sure if the-

Robert J. Marks:

Yes.

Walter Myers:

But RAID. So RAID five, for example, you have to have, I believe a minimum of four drives. Okay? Because there's a three drives that it's striping the data across those three drives, and then there's the error correction drive, which is the fourth one. Right?

So if one of those drives goes out, the error correction one knows, "Oh, this is how I need to reconstitute that data that was lost, that was striped across those three drives." So you know that your data is smeared out across those three drives and the fourth drive is what's mapping and looking at what's going on here, so that if one pops out, it can then, and you pop another drive in, then it, without you doing anything, will go ahead and reconstitute that data that was lost. So that's a RAID 5 array.

Robert J. Marks:

Okay. So what you're saying is that there's really no error correction code, that the redundancy is real. There are complete copies.

Walter Myers:

Right.

Robert J. Marks:

Okay.

Walter Myers:

Well, it's not three copies, in that case. It's all the data is smeared across those three drives, but then there's that fourth drive that you add, that's with respect, that fourth drive that you add, whereas the data strike, somehow it's keeping track of the data, so that if you pull one of those drives out, you'll be able to reconstitute what you lost, what you would've lost in that one. I don't know exactly how it works, but I have a RAID 5 array and it's a beautiful thing.

Robert J. Marks:

Yeah, it's incredible.

Okay. What have we missed? Let me ask you, as we come down to the end here, is there a common misconception about cloud security that you'd like to address in talking to the people that are interested in using the Azure, the Azure cloud?

Walter Myers:

Yeah, Bob. A common misconception about cloud security is from customers, particularly during the early years of Azure, is if someone can access my data, because it is multi-tenant. Right? Lots of people are gaining, thousands of people are gaining access to this. And the answer to that question is, no.

We are using technology that's been around for decades as far as how we write blocks, how we mark blocks of data, how we mark blocks of data for deletion. So for those who are concerned, no. Your data is not next to someone else's data, and then they can go from one block, to your block, and read your data. It doesn't work that way.

We have authentication and authorization mechanisms that determine what data can actually, who can access what data, and so, no. So, no one can see your data. And even if your data is deleted, once that is marked for deletion, that data cannot be read until it's actually written to again. So you can rest assured that your Outlook emails, or your OneDrive photos and all those things, no one can gain access to those except for you.

Robert J. Marks:

Okay. Well, great. Well, thank you, Walter. I've learned a lot here, and I think people that listening have learned a lot too. So appreciate your time and sharing about the cloud with us.

Walter Myers:

Awesome.

Robert J. Marks:

Now it isn't so mysterious. We've been talking with Walter Myers III, who is a principal engineering manager. He leads a team of engineers, working with management of the Microsoft Azure Cloud. This has been Mind Matters News. And until next time, be of good cheer.

Announcer:

This has been Mind Matters News with your host Robert J. Marks. Explore more at mindmatters.ai. That's mindmatters.ai. Mind Matters News is directed and edited by Austin Egbert. The opinions expressed on this program are solely those of the speakers.