The Deloitte On Cloud Podcast

Gary Arora, Chief Architect of Cloud and AI Solutions

Title: **Quantum computing, AI, and Security: Deloitte's Scott Buchholz on the imminent quantum leap**

Description: In this episode, Gary Arora sits down with Scott Buchholz, Deloitte's quantum computing lead, to explore the potential of quantum. Scott classifies its uses into "offensive" and "defensive" categories and talks about its impact on AI. They also discuss how quantum will change the game for cryptography. Finally, Scott warns that, although functional quantum computing isn't yet a reality, tech leaders should keep it on their radar as a shift they'll likely be forced to make.

Duration: **30:29**

Gary Arora:
Welcome back to the On Cloud podcast. I'm your host, Gary Arora, chief architect for cloud and AI solutions at Deloitte. Today's episode is a special one because it's about a technology that sounds like science fiction but is quickly becoming science reality, quantum computing. It is starting to show signs of real commercial potential and real security implications. To help us navigate all of this, I'm joined today by Scott Buchholz, who is the quantum computing lead at Deloitte, among other hats that he wears at Deloitte. Scott, thank you so much for joining the show. I know I'm going to learn a lot in this conversation and I'm pretty sure our audience will too.

Scott Buchholz:
Gary, it's always a pleasure to have the excuse to talk with you again. So, thanks for having me.

Gary Arora:
Awesome! So, Scott, I first started paying attention to quantum computing right before COVID when Google announced a major breakthrough. Their Sycamore processor achieved what they called quantum supremacy, which is solving a problem in 200 seconds that would have taken a supercomputer 10,000 years. That's 10 to the power of 4. Now, fast forward that to just a few weeks ago, Google made headlines again, their new chip, Willow, completed a benchmark calculation in under 5 minutes that would have taken one of today's fastest supercomputers 10 septillion years. That's 10 to the power of 25. That number far exceeds the age of the universe, and it's a completely insane number. So, my first question to you is why should or really when should CIO or CTO, or any tech leader start losing sleep over quantum computing? Are there any trigger points that should make them say, "Okay, quantum now matters to my road map?"

Scott Buchholz:
First of all, I don't advocate anybody losing sleep over or really over much of anything at work, if it can be avoided because sleep is so vitally important. But what I would say is this. If we think about quantum computing, there are two major aspects that people often talk about, and we can lump them, I don't love the terms, but we can sort of lump them into offensive uses and defensive uses. On the offensive uses side, we're talking about use cases, things like optimization, machine learning, simulation, whether it's chemical or Monte Carlo simulations, better simulating and understanding nature. Those are the things that fall into the offensive use cases and the idea there is quantum computers which represent a new way of performing calculations that's fundamentally different from anything we have today, actually enable us interesting ways of accomplishing those tasks.

It turns out that building a quantum computer is fiendishly difficult. There are more than half a dozen different ways of building them. A number of them involve cooling electronic circuits to temperatures 100 times colder than outer space, or it involves zapping individual atoms with lasers millions of times a

second. It's all of these really incredible technologies and ways of doing things. Because of the nature of building a quantum computer, because we're often manipulating individual atoms or very small groups of atoms, it turns out that there are certain types of errors that are commonly introduced. So, what happens is in order to make quantum computers work, we take the underlying physics, and people call these things physical qubits, and we turn them into, or we're working to turn them into, error-corrected logical qubits.

Anyway, there's a lot of explanation that sits behind that, but essentially the logical qubit count, which you'll sometimes see referenced, is the thing that we really care about when we're trying to resolve many of those use cases. Probably, when we get somewhere around 100 logical qubits of availability, which is likely to happen in the next handful of years, depending on which vendor road map you believe, is actually the point at which we'll be able to start doing things that we cannot do on supercomputers that are potentially, commercially relevant. So, next couple of years, keep an eye on logical qubit counts, likely when we get to somewhere around 100, we'll start being able to do commercially useful, interesting, relevant things, and everybody is working really hard at the moment to try to figure out where those are. So, that's the offensive side and when you should worry about that.

On the defensive side, the idea is that a sufficiently powerful quantum computer which does not exist today, and just so that people understand requires more than 1,000 logical qubits, would actually be able to decrypt the vast majority of the world's encrypted information in days and hours, rather than the lifetime of the universe, which is the current security guarantee. And as a result, people need to start planning to switch from today's encryption techniques for things like web browsing and securing data at rest, and a variety of other mechanisms, as well as digital signatures to the new standards that have been defined and are being rolled out as we speak. Those are often called post quantum cryptography. They run on today's hardware. There's a lot of nuance there, but those are the two big areas and that was a really long answer to what you probably thought, Gary, was a short question.

Gary Arora:
You know for such a complex topic, I love the simplicity you brought to what quantum computing can do with your offensive and defensive categories. You mentioned optimization, which is what we, and I mean the royal collective we, have been doing over the last 100 years with computing, designing workflows and then optimizing workflows with new technologies and techniques. So, what are the kind of real-world problems that quantum computing can crack where even your biggest classical clusters of computers are simply tapping out? Beyond just the hype, what is actually in play here?

Scott Buchholz:
Optimization is an interesting problem space. When we talk about optimization as a class of problems, we are often talking about trying to schedule nurses at hospitals. We are talking about trying to figure out the sequence of manufacturing things. We are trying to talk about, if I'm a retailer and I have warehouses and a finite number of trucks, how do I make sure I get my goods from the warehouses to the stores where they're needed on time, but using the smallest amount of fuel? Those are examples of optimization problems. We've had many, many, many years, so a hundred years or more of really, really smart mathematicians figuring out ways to solve those problems.

And it turns out that today's best classical solvers are exceptional at solving these problems. Where quantum computers turn out to be useful, at least as near as we can tell, is problems where oftentimes they exist in what are sometimes called network optimization, large scale logistics problems. I'm trying to figure out the retailer example. It turns out that a number of things happen. One of them is that you may be able to classically compute an answer, but it may not actually get done in time. In some cases, you might be able to compute an answer, but it might be a good answer, not the best answer or a better answer. And so, in a number of cases, what people are trying to do is use quantum computers to either improve the initial guesses to get closer to the optimal solution to use a special class of quantum computer called a quantum annealer.

And quantum annealers have physical properties that allow them to get to very high-quality solutions very quickly, so that makes them interesting and problems where time is a major constraint. There are a lot of different options and nuances, but it turns out that those are places where we think quantum computers might be really useful. By the way, just to remember, there are a lot of optimization problems today that are solved with a pencil and a piece of paper or a spreadsheet. We're probably not replacing those things with a quantum computer. On the other hand, there are some very complex optimization problems, if you think about the world's largest retailers, they have a lot of them. And certainly, people have been looking at quantum computers for partial solutions to those, helping accelerate those things and other sorts of challenges in that regard.

Gary Arora:
And speaking of solutions where quantum computing might be useful, AI is grabbing a lot of headlines these days, and one of the limitations of AI is that vast amount of data that it needs to process, analyze, and interpret can easily overwhelm a conventional digital computer. And doing this is one of the strongest points of quantum computing. From where you sit, first of all, is this understanding correct? And second, what is the toughest engineering bottleneck that's stopping this crossover between quantum and digital computing? What is it making it so hard to solve?

Scott Buchholz:
What's funny, Gary, is when I talk with a lot of people about quantum computers oftentimes because we talk about quantum computers being able to solve problems that are intractable classically. People say, “Oh, well, that must mean that this is like a supercomputer. So, I've got all this data lying around. I can't solve it today because I don't have enough computing power, so quantum computers will come to the rescue.” Unfortunately, the answer to that is probably not or at least not in the foreseeable future.

The best simple analogy to understand the difference is that classical computers are as different from the way quantum computers work as light bulbs are to lasers. If you think about a light bulb and a laser, both of them are emitting light. But they are not substitutes for one another. It's really hard to light a room with a laser. It would be really hard to read a barcode with a light bulb. And the challenge is most people's mental models have not yet adjusted to what a quantum computer can do because we don't yet have enough of them for people to have a solid grounded intuition. Here's what we do know: Quantum computers tend to work very well when you have smaller input data that has greater variability.

Quantum machine learning—when we apply machine learning on quantum computers—appears to train to higher accuracy with less data than classical models. You could then envision that there may be interesting cases where quantum computers become useful. Incidentally, it turns out that it may be the

case that quantum generative adversarial networks or GANs, may actually be easier to implement than their classical counterparts. A generative adversarial network is where you are essentially running two machine learning models in parallel, one of them has the job of finding distinguishing true information from false information. One of them has the job of generating fake information.

And what you do is you run them in cycles and the fake information generator gets progressively better at generating fake information over time and the discriminator gets increasingly better at telling real data from fake data, and you run them to the point where the generator can essentially generate data that is indistinguishable from real data, and you have a high quality synthetic data generator. So, it turns out they may actually converge better on quantum computers. Lots of machine learning people are also starting to say, "Hey, we don't yet have a quantum computer that can run a production workload today." But we can look at how we know they're going to work in the future when they're available at scale.

Let's port some of that back to today's technology, CPUs and GPUs. And it turns out in doing that, you can actually, for instance, we've found in financial fraud data, we actually have a new novel machine learning technique that can do a better job at finding fraud and payment transactions and run in real time. Another example is there are researchers who have figured out how to use some of these quantum inspired techniques to create large language models that are almost as accurate and significantly more compact than our existing models that are running today. So, there are lots of really interesting areas of research around this idea of quantum and AI, quantum and machine learning. So, I would loosely say stay tuned because really interesting stuff is coming.

Gary Arora:
Always on point with your analogies Scott. Lasers versus light bulbs, very helpful for the mental model shift. Now you mentioned a couple of examples around adversarial use cases, financial fraud, synthetic test data generation and validation. Let me ask you a security-related question. Quantum does cut both ways. Stronger machines eventually break today's encryption, with the NIST's postquantum standards now published, from a CISO standpoint, what are some of the smart moves that office can make in the next few months to stay ahead without overspending?

Scott Buchholz:
Sure. So, just to reiterate, NIST, the National Institute for Standards and Technology, which is part of the US Federal Government, sets standards for security. And one of the new standards that they have established is new encryption algorithms and digital signature algorithms that are run on classical hardware, that are called postquantum cryptography. They are algorithms that are designed to be resistant to known attacks, classically and known possible quantum attacks.

They are in the process of being reviewed for international rollout. They are already mandated as future standards for the US Federal government. So, they're, sort of, in a rollout period over time. And depending on which analyst you believe, a sufficiently powerful quantum computer is either going to exist, or will exist in the soon enough future, to warrant worrying about your existing encryption standards somewhere between, sort of 4 and 10 years from now. A lot of nuance in that, that's a whole other discussion, but what CISOs and CIOs and boards should be looking at is working to understand where cryptography exists in your organization, that is most impactful or most impacted by the idea that somebody might be able to decrypt it and get at it.

So, for instance, your customer data probably falls in that vein in many industries. We could argue whether or not credit card numbers are perhaps as important in that scheme, because credit card numbers tend to change every five years. If you sort of compare the math, you might be able to get away with not worrying about them for the time being. There's a hierarchy of importance of data and importance of securing data, but people need to have a plan to understand when those standards do become more tested, more battle hardened, etc., how are we actually going to make sure they get rolled out?

Even the hyperscalers, so the cloud providers, have said that rolling out the standards while they will implement them is going to be a joint responsibility between the client and the provider. It's not going to be that the providers are just going to snap their fingers one day and everything will be done. Individual organizations will have the responsibility for managing the transition and cryptography now exists in far more places than it used to. People just need to be prepared because this shift is coming. And although it feels like we have time, I suspect that people may wake up one morning and discover time goes by faster than ever before. At least it seems to for me, sort of sitting around and waiting to come up with a plan and get ready may not be the best strategy.

Gary Arora:
Are you seeing many projects in this space that are updating the tech stack or the applications to meet the cryptography standards?

Scott Buchholz:
The short answer is there are a few projects. So, if you look around you can see a number of the cloud providers are starting to offer this encryption as an option. Some of the cloud distribution providers are offering it. Famously, Apple is using it to encrypt iMessage data these days. There are a variety of places where the encryption standards are actually being used and implemented. In addition, people are starting to look at it, I think most organizations are currently at the thoughtful plan development stage in order to get ready for implementation. And in some cases, people might say, look, I have a legacy system, I know I'm upgrading in two years, as opposed to going in and ripping everything out and trying to fix the encryption today, let's just make updated encryption part of the standards for whatever new tool package, etc., that we select. So, it's not necessarily a rip out in every case. There may be cases where you can use existing modernization programs to actually upgrade the encryption that you've got going on.

Gary Arora:
And are there certain industries more prone to the risk than others, whether health care or financial services, or ones that are, let's say, not in cloud, yet not as technology mature industries? Or are you noticing that the urgency and the security exposure is nearly equal across all.

Scott Buchholz: It's true that health care and financial services are among the first. They tend to have either the data that is the most sensitive or the data that is sadly most likely to be stolen. Just because of the nature of what they do. If you think about it though, if the idea is a sufficiently powerful quantum computer. So, if I had one in my back pocket, somehow I could actually watch your Internet and I had a way to watch your Internet traffic, Gary, then I could

decrypt your connections with your employer or wherever else your bank and masquerade as you. And no one would be any the wiser and thus it's not just financial services and health care that need to worry, but arguably utilities need to worry. Because having somebody traipsing through your network, masquerading as one of your employees has got to be a bad thing. You can go on down the line there are lots of places where people need to worry.

Gary Arora: All right, so the urgency and the risk is here and there are actions we've got to take. Let's get tactical. Suppose I'm a developer with the credit card and a couple of hours on a long weekend. How could I roll up my sleeves and get some hands on experience to run code on a quantum computer?

Scott Buchholz: We're asking for a friend, Gary.

Gary Arora: Yes, very much so.

Scott Buchholz:
Well, a couple of things. You can absolutely go run things on a quantum computer, IBM, Amazon, Microsoft, Google, a number of others have programs where you can go and access quantum hardware today. So, depending on your credit card and your ability, you can go use an actual quantum computer. What I will caution is the following. There are a number of programming, let's call them APIs, maybe we call them languages, approaches to using quantum computers. IBM has Quiskit, Google has CERC, Microsoft has Q Sharp, and so on down the line.

It appears, if you first look at them, that it's reasonably straightforward to just pick up the code and use it. The challenge that we're finding is that knowing how to effectively program a quantum computer is a lot like learning data science from scratch. So, while you can just go run a data science algorithm, a machine learning algorithm, that doesn't actually mean that you know what you're doing, that you're getting the right results, that you would know if you were getting the wrong results, etc. And in the same way that it takes a year or two to become proficient at data science, it often takes people a year or two to become proficient in programming quantum computers. So, if you would like to go burn some spare cash on the weekend, I'm sure the vendors would love to help you out. I would also say just understand that this is a journey. This is not a "spend a couple hours of training and you're an expert." It's a whole new thing.

Gary Arora:
And I think that developers do need to be ready and be congregating in this space, given the risks that we are seeing here, the work is pretty much cut out. In terms of, let's say timelines, when do you expect the first production grade, revenue-generating workload to run on real quantum hardware or what kind of use cases could be leading the pack here that has to click before it becomes viable?

Scott Buchholz:
Well, it turns out that if you think a little bit to what I mentioned earlier around optimization. There is a special class of quantum computers that are referred to as quantum annealers. And it turns out, because they are special purpose optimization machines, they are somewhat more straightforward to construct. Not that they're easy, but somewhat more straightforward and D-Wave, which is one of the longtime vendors in that space, it has been running production workloads for a number of years. So, you can go and look. It's well known that at the port of Los Angeles are actually running on these D-Wave devices in production. Basically, when ships come in to port, they offload the containers onto the dock. They then have to have semitrucks drive down the dock and have the container put on the back of the semi.

After a lot of experimentation, they determined that using the D-Wave quantum annealer was actually the best way of deciding which semi should drive down the dock to get the container put on its back. And that happens today. When we talk about quantum computers, however, in general we're talking about what people refer to as universal gate-based machines. So, these are machines that aren't just good at optimization, but are good going to be good at a lot of things. These are the things that Amazon, Microsoft, IBM, Google, and a million startups are all building. And those devices it's hard to know exactly, but the vendors are in the process of saying somewhere around 2026, 2027, or 2028 is likely to be the time when the first production use cases start running.

I would say those use cases are not likely to be real-time use cases. So, if you think about fraud detection and financial services. If you're a bank, there might be one type of fraud detection you worry about, which is payments. So, real time in the moment. There's another type that you worry about, which are scams, and so what you see in scams is patterns of increasing payments overtime and it turns out people have used some of these quantum inspired techniques that can be used on quantum computers to actually do scam detection. And it works quite well. So, it might be that those types of use cases are actually some of the first ones from production purposes or some of the other accelerators relative to optimization and other sorts of things.

Gary Arora:
Quantum solving the problem of containers the real the OG containers is such a peak 2025 using physics that Einstein barely understood to make sure your patio furniture shows up on time.

Scott Buchholz:
Yes, that is correct.

Gary Arora:
Exactly. That's where we are going. What's your best-case win scenario that you expect to see in the next 24 months, and what headline should our listeners watch for to know that quantum just got real?

Scott Buchholz:
I mean, ideally some bright light in a garage figures something out and the world changes, but that's a dream, not a likely reality. I think the thing to watch for in the news will be, ideally, an announcement that Deloitte, working with some of its clients, has figured out how to put some of these technologies into production. And there are potential use cases around machine learning problems, as I've talked about. There are potential use cases around optimization,

so a number of people have been looking at whether or not you can better optimize bus routes and schedules using these gate-based quantum computers and using them to accelerate some of the classical technology. A lot of what people are really looking at is taking some combination of high performance computing, supercomputers, or other large-scale classical machines jointly with quantum computers together to actually solve really interesting problems.

And so, somewhere in that space is what I expect people should be looking for. I would anticipate that it's probably going to require somewhere around a hundred logical qubits. The reason that people are picking that point is because somewhere in the 45 to 50 range, logical qubit range, is actually the current limit that you could classically simulate. So, we cannot classically, simulate a computer, a general quantum computer with a hundred qubits. So, there's a certain amount of intuition that that's the scale at which capability becomes really interesting, useful, and powerful.

Gary Arora:
You know, these are certainly hard and very interesting problems to solve, such an interesting time to be in technology and contribute. That is it for the episode. Thanks for tuning in. A big thank you to our guest, Scott. If you liked this episode, please be sure to leave us a review. So, we continue to keep it real and bring you more insights from the trenches. Thanks for listening to the On Cloud podcast, until next time. I'm Gary Arora.

Operator:

Visit the On Cloud library
www.deloitte.com/us/cloud-podcast