DR: [00:00] NVIDIA had a bunch of big time operators like AT&T, T-Mobile, Comcast, and Spectrum, all announcing they're building something that they're calling AI grids. So what's an AI grid?

Kanika Atri: [00:12] Telcos already have millions of small, medium, large data centers spread across a country. They've got this prime real estate that's already connected through fiber. They also have Spectrum. And now AI as a workload on top of it is just the most obvious thing to do.

Announcer: [00:38] This is "Telco in 20," a podcast that helps telco execs achieve a competitive advantage with AI and the public cloud. It is hosted by Danielle Rios, also known as DR. Today we're talking to Kanika Atri, Senior Director of Telecoms at NVIDIA.

DR: [00:57] Hi guys, I'm DR. NVIDIA just surveyed a thousand telco professionals for its state of AI and telecommunications report. More than 90% of respondents said AI is already delivering cost savings or revenue growth. The biggest ROI, generative and agentic AI for autonomous networks. That's a clear signal that operators aren't experimenting anymore, they're investing. But what really caught my attention were the telco announcements NVIDIA's been making at MWC and GTC, its big annual conference. They weren't spotlighting AI use cases and customer support, they were laying out plans to turn telco networks into AI compute platforms. And they made three big moves. First, AT&T, T-Mobile, Comcast, and Spectrum, all announced they're building something called AI grids. Their plan is to turn millions of existing cell sites, central offices, and switching centers, into distributed AI compute infrastructure. Second, there's a group of operators and vendors aligning to build 6G on AI-native software defined platforms.

[02:02] They want to ditch the old model where new network capabilities show up in hardware once per decade, and move it into software. And third, there's traction with sovereign AI. 24 operators worldwide are building sovereign AI factories to become AI service providers for entire countries. For an industry that's famous for going slow, that's a lot. Today I'm talking to Kanika Atri, Senior Director of Telecoms at NVIDIA. We're going to dig into where telcos are seeing real ROI with AI, how AI grids could turn telco infrastructure into the world's largest distributed compute network, and my big question, who's going to build the software that makes all of this work? So let's take

20. Kanika Atri is Senior Director of Telecoms at NVIDIA. Hi Kanika. Welcome to "Telco in 20."

Kanika Atri: [02:53] Hi DR. It's great to be here.

DR: [02:55] I'm so excited to have you on the podcast. You've been in telco for a while. You've worked on the vendor side at places like Juniper Networks, Nokia, and Siemens, and now you're at NVIDIA. And recently NVIDIA published their state of AI and telecommunications report, and you guys surveyed a thousand telecom professionals. What were the big takeaways that you guys highlighted in that report?

Kanika Atri: [03:18] For me, the big takeaway was every technology cycle actually goes through phases. It starts with investment, then adoption rises. And eventually the third phase is adapt. People adapt how they do business, how do they consume their technology? They adapt their lifestyle around it. So when I look at AI adoption and investment in telcos, this survey told us that a huge progress has been made across all tracks.

[03:51] The generative AI and agentic AI was a standout, particularly for the use cases of running networks autonomously. We call that autonomous network. And that's where the maximum ROI is being seen. Then on the sovereign AI side, we saw a lot of new business growth that telcos are seeing thanks to new AI service provider businesses that they are building. And then finally on the AI-native 5G and 6G networks, we saw much of the telcos in the investment and assessment phase. And I was actually quite surprised to see that many of them see AI-RAN coming and being part of a deployment cycle even before 6G. So basically AI as a technology is reaching that point of maturity where the benefits are clear, and that's driving more and more investment. And I hope that next year, it will become a way of life.

DR: [04:50] Yeah. Telcos are famous for moving slowly and not really embracing change. And we're in a time of very, very rapid change. And so it's going to be really interesting to see how the telcos change their culture and really apply this new technology. You guys at NVIDIA just wrapped up your GPU Technology Conference, or GTC for short. This is your big conference. And this year you had a bunch of big time operators like AT&T, T-Mobile, Comcast, and Spectrum, all announcing they're building something that they're calling AI grids. So what's an AI grid? Tell me about that.

Kanika Atri:

[05:26] Well, the next frontier of AI infrastructure build out is going to be distributed. And an AI grid embodies that idea. Think of it this way. Today, many of the hyperscalers, neoclouds, they have got these mega token factories.

[05:44] These are centralized, massive rack scale data centers. As the kind of applications grow and that adoption and adaption increases, we're going to see a lot of AI-native applications that are distributed that are in our handhelds, that are in robots, drones, sensors. That data ideally does not need to go all the way back to a centralized system to get computed. And that's where the opportunity for distributed computing is. When we talk about telcos, they are beautifully suited to this opportunity. They already have millions of small, medium, large data centers spread across a country. And these are all sovereign. They're within the boundaries of the country. And many of these today serve as what we call mobile switching offices. Many of them serve as central offices, PoP sites, and of course cell sites.

DR:

[06:46] Yep, absolutely.

Kanika Atri:

[06:47] Each of these is going to be transformed into an AI infrastructure that's capable of hosting not just the connectivity functions, but also edge AI. Collectively, we call this an AI grid. This whole substrate of geographically distributed computing sites that acts as a unified intelligence layer, such that depending on whatever workload is coming to it, this grid is intelligent enough to figure out where should I compute it. If the workload is very token intensive, perhaps it's going to get computed at an MSO location. If it's very latency sensitive, maybe it's going to get computed at the cell site location. The point being that all this land power shell exists today. Telcos have that advantage that they've got this prime real estate that's already connected through fiber. They also have Spectrum. And now AI as a workload on top of it is just the most obvious thing to do.

[07:54] So AI grid from NVIDIA is that whole reference architecture that allows telcos to transform their existing infrastructure into AI infrastructure. And we had a very exciting launch at GTC, not just new telcos adopting it. We had AT&T, Comcast, Charter, T-Mobile, IndoSat. They're building massive AI grids. We also showcased some amazing applications that benefit from this AI grid, personal voice AI agents, personalization of ad and media delivery experiences. So all of these benefit from a distributed infrastructure, whether it's

lower latency, massive concurrency, or even cost. So there's an inherent advantage to building a distributed computing infrastructure, which is the AI grid. We believe this is one of the biggest opportunities for telcos now, and they can start building it today.

DR: [08:53] Yeah, yeah. And so another thing you guys have recently announced is this massive coalition of operators and vendors like BT and Deutsche Telekom, along with Ericsson, Nokia, but this is not the full list. There's a lot of people participating in this coalition, all committing to build 6G on AI-native platforms. Jensen Huang said telecom is the next AI infrastructure build out, which I'm sure excites all the telcos. And so what does that mean for operators today?

Kanika Atri: [09:21] The biggest asset of telcos is their wireless networks.

DR: [09:25] Yeah, absolutely.

Kanika Atri: [09:26] And actually, more than 20 billion devices today connect through these wireless networks. As we look at the next generation, we all know that 6G is right around the corner. And at the same time, AI is here. When 5G was written, AI was not a thing.

DR: [09:46] Correct.

Kanika Atri: [09:46] 6G is being born in the AI era. And as we look to this next investment cycle, we believe that there are a few first principles that every network must be built on. One, they must be AI-native, because obviously they need to serve this new kind of demand. Every application is getting transformed with AI that's going to use this network, including physical AI. Two, they must be software defined, which basically means break the G cycle. Up until now, we waited every 10 years to introduce new capabilities, and these generational shifts were very closely coupled between hardware and software. As a parallel, if you look at the cloud and the software industry, in the last decade, it thrived because of the fundamental architecture that is software defined, wherein you can introduce new changes every week.

DR: [10:41] Absolutely.

Kanika Atri: [10:42] Forget every week, every day. Why can our telco networks move at that pace? And it's because that architecture

has been locked up. We want to change that. And as you do that, it's equally important to have a diverse open ecosystem to make it very secure and trustworthy. So these are the key principles. At Barcelona, we had a very big coalition of companies aligned that the next infrastructure build out will be AI-native, software defined, open, secure, and trusted. And AI-RAN is that architecture that allows them to move into that next generation.

DR: [11:19] And who's going to be on the hook for delivering that software? Telcos aren't software providers. Ericsson and Nokia, famously network hardware providers. NVIDIA, you guys aren't going to build it. So who's going to do that?

Kanika Atri: [11:32] Well, there's a massive partner ecosystem that NVIDIA has built. And of course, the telcos traditionally have purchased from the large NEPs, what we call the equipment providers and Nokias, Ericsson, Samsung and so on. And they are going to continue to have a very important role to play, which is testimony to our partnership with Nokia, NVIDIA announced a billion dollar investment. And Nokia is building CUDA accelerated 5G software that runs on any platform. What we've also believe is in openness. The whole idea is, once we open source this, we allow a lot of innovation to come in. So this unlocks the opportunity. New players can come in and write different parts of the stack, and with AI, the sky's the limit.

DR: [12:27] Well, there's a lot of software that needs to be built for the AI grid and moving workloads and orchestrating that. The hyperscalers obviously have been doing this for 20 years. Can the telcos replicate that through a vendor ecosystem? Because I do think it's potentially a good monetization opportunity for telcos, and it's tough to go compete head-to-head against the hyperscalers, and some of these ideas seem to be doing that. And so let's talk about that idea. You recently wrote about telcos building sovereign AI factories, national AI infrastructure for entire countries. And so NVIDIA has built 18 of them across five continents with different operators. Why are telcos partnering with you to do this?

Kanika Atri: [13:05] Well, first of all, since I wrote that blog back in May or June last year, the count has gone up. We now have 24 telcos building sovereign AI infrastructure. And telcos have that existing trust and sovereignty factor that not all hyperscalers today can claim. Every country is going to build AI infrastructure. Every country will need their own AI. Every enterprise will need

to train their own data and do their own things using foundation models and other world-class applications, but tune them into what is relevant for them. Every country, every enterprise will do it. And when they do it, they want that data to stay in-country. They want to work with players that they already have deep relations and are trusted. Telcos have a huge advantage in that. And that is the reason why we're seeing a big progress and momentum in telcos building sovereign AI factories.

[14:09] Indosat in Indonesia for example, is such a beautiful case study. What they've done is they've built the full five-layer cake of AI infrastructure, as we call it, starting from the land, power shell, ships, to the infrastructure, models, and application. Indosat actually has built all the five. They've built their own models that speak Sahabat, which is the language of 270 million Indonesians. On top of that, they have 20 plus locally developed AI applications that leverage these models and take healthcare to the masses. Education, improved agriculture. So to me, it's not just the part of what they're building, it's the part of what they're doing with it, how they're enabling the nation, the enterprises. And now telcos are the natural next suppliers of intelligence. And sovereign AI infrastructure is the first step to build that centralized infrastructure, then expand to AI grids, make it distributed, and then eventually bring it through the cell sites by adopting AI-RAN and completely revolutionizing edge AI applications.

DR: [15:25] Okay, got it. That's super interesting. Well, we've talked about scaling and telco's starting to scale their AI workloads and their AI ideas and starting to implement them. You have been in telco for over 20 years. You started as one of the few women network engineers in India, and I found out that you've actually scaled a cell tower. So you've actually climbed one of these things. So tell me about that. Is it scary?

Kanika Atri: [15:51] DR, I think that was one of the most fun parts of my job. I actually started out as a RF engineer, and I actually had a nighttime job, initially was about network optimization. So we used to do drive tests sitting in a car, and you do that drive test at night because sometimes you've got to tune some things and you don't want to affect the traffic.

DR: [16:13] Certainly more efficient to drive around without traffic.

Kanika Atri: [16:15] That's right. And then during the daytime, we would actually go visit those sites where there were issues. And sometimes we had to climb towers and fix the antenna tilt. And I remember back in 2002 that as part of my engineering job, I would download data every night from every cell site. And it was huge amount of Excel sheets with billions of rows. And we had, of course, big Linux machines in which we would fit them and play with them. And I remember writing macros and small scripts to optimize handovers. Today, you can just throw this to AI, write a prompt, and say, "Optimize these handovers," and you're done.

DR: [16:56] Fix it for me while I drink my chai tea.

Kanika Atri: [16:58] Exactly. So that said, I don't miss having had that opportunity, because it's almost like you follow the path of a bit from the moment it's transmitted from that antenna and how it hits the environment. If you understand that whole air interface and how complex that is, it's almost exciting, but it gives you a foundation for life. And for those reasons, I love being in telcos, and I appreciate how much AI is actually going to help, because AI is perfectly suited for hard problems like the physical air interfaces, one of the hardest things to put into a code.

DR: [17:35] Absolutely. Well, I think the best leaders are people who have actually done the job themselves and seen lots of different jobs in the company as they've moved up. They have intimate knowledge, and it sounds like you really do. You've been on the top of a cell tower, you really know how it works. And so Kanika, this has been an amazing conversation about what's going on with NVIDIA. So thank you so much, and thanks for coming onto the podcast.

Kanika Atri: [17:57] Thank you so much, DR. It was great talking to you.

DR: [18:00] Awesome. Stick around. We end each podcast with a "Telco in 20" takeaway. I've got two minutes to tell you something you need to know. Kanika is right. The telco network is an incredible asset. Millions of sites, fiber everywhere, Spectrum nobody else has. And NVIDIA's vision is exciting. AI grids, sovereign clouds, AI-RAN. But there's a massive layer in the five layer cake that's missing and nobody wants to talk about it. All the software, you need to run it. Who's going to build that? Telcos can't. They buy and operate technology. They don't create software. And Ericsson and Nokia are hardware companies. So you've got this grand vision of distributed AI

infrastructure with a software gap no one wants to own. NVIDIA suggests open sourcing parts of the stack and letting new players write different pieces. As a startup CEO, I can tell you this sounds great in theory, but telcos have a hard time trusting unproven software.

[19:07] How are they going to get over that hump? So before you start chasing your sovereign cloud and AI-RAN dreams, make sure someone owns a software layer. Because if you don't have the software chops to make this work at scale, it's not a five layer cake, it's a recipe for disaster. The future of telco runs on software, not hardware. That's the bet we're making at Totogi. Want to talk more about it? DM me on LinkedIn or X @TelcoDR, and we'll set up a time to meet. Until then, tune into more "Telco in 20" episodes, like and follow, and leave us a five-star review. Don't forget to sign up for my no BS email newsletter on telcodr.com, and check out our awesome YouTube channel and hit that subscribe button. Later nerds.