**Raul Martynek:** 0:00

What a data center ultimately does. It underpins the ability for companies, organizations, for society as a whole to deploy technology, and all that technology has to go into a data center, which is why the center, you know, for the last 15 years, has had, you know, called it double digit growth. And this is a real estate business at the end of the day, right, so that's a pretty attractive dynamic for a real estate type of business. You think about Google as a search engine, right, because that's really what it is. Right, it's an advertising platform that delivers search results, generates billions of dollars of revenue. Right, it absolutely uses AI to generate those results, to figure out what to present to the people and all that stuff, right?

**Craig Smith:** 0:45

Hi, I'm Craig Smith and this is Eye On AI. A lot has been written about the enormous and increasing costs of large language models, but one aspect to that story that's gone unreported is the impact it's having on data centers, the physical server farms where all of the computation takes place. This week I spoke to Raoul Martinek, the CEO of Data Bank, one of the industry's leading data center providers, about the intricacies of data storage and the challenges of GPU intensive workloads for AI. Raoul's perspective illuminates the foundational layers of our digital world. Join us as we delve deep into the world of data centers and their pivotal role in shaping the future. But before we begin, let me mention our sponsor, netsuite By Oracle. If you're a business owner, having a single source of truth is critical to running your operations, so remember these three numbers 36,0251. 36,000 because that's the number of businesses that have upgraded to Netsuite Bi Oracle. Netsuite is the number one cloud financial system for streamlining accounting, financial management, inventory, hr and more. The number 25 because Netsuite turns 25 this year. That's 25 years of helping businesses do more with less, close their books in days, not weeks, and drive down costs. One because your business is one of a kind, so you get a customized solution for all of your KPIs in one efficient system with one source of truth. You manage risk, get the most reliable forecasts and improve margins. Everything you need all in one place. I'm not the most organized person in the world, and there's real power to having all the information you need in one place to make better decisions. This is an unprecedented offer by Netsuite. To make that possible right now, download Netsuite's popular KPI checklist, designed to give you consistently excellent performance. Download it for free at netsuite.com/eyeonai. That's E-Y-E-O-N-A-I all run together. Go to netsuite.com/eyeonai to get your KPI checklist. That's netsuite.com/eyeonai E-Y-E-O-N-A-I all run together. Again, that's netsuite.com/eyeonai. They support us, so let's support them.

**Raul Martynek:** 3:50

It's great to be here, Craig, thanks for inviting me. So I'm Roe Lortnick. I'm the CEO of Data Bank. I've been in the Internet infrastructure space for over 27 years, kind of since the advent of the commercial Internet, working with sponsors to build the foundational businesses that underpin the modern Internet. When you think about your cell phone or your desktop and you think about the physical infrastructure that's required to move those bits to you, it typically starts with, on the wireless side, some type of cell tower or a small cell, as we call it. All hits a fiber optic cable and then ultimately everything comes back to a data set. So those three links in the chain are what we consider digital infrastructure. And Data Bank is a US-based data center operator and developer. We own and operate about 65 data centers in 26 US markets, giving us the largest geographic footprint of any data center operator in the US, public or private. We're backed by a whole group of investors, including DigitalBridge, which is a leading infrastructure manager in this digital infrastructure space.

**Craig Smith:** 5:04

Yeah, and a lot of people. With the migration to the cloud and all of the talk about the cloud, I think a lot of people have forgotten that there are independent data centers out there. I mean, the cloud is a network of data centers itself, but there are data centers outside of the big cloud providers that are not necessarily cloud data centers. Yeah, yeah, and can you talk about how? And then of course, there's on-premise data centers that people use. They are very concerned about their data security and that sort of thing. But can you talk about how the existing independent data center market has evolved as the cloud has taken over and then how that plays into what's happening with AI in general, AI in particular?

**Raul Martynek:** 6:11

So look, just a quick little rewind, right. So if you think about it, about 15, 20 years ago, you'd walk into any office in America and there'd be a data center. It was called the computer room, right? And the internet came along and really kind of made applications portable and data portable, so to speak, right. So that's what kind of created this kind of need for, as we call it, a multi-tenant carrier, neutral data center, right. So in the mid-2000s, it's what people's mentality really switched and enterprise, it's said to themselves hey, you know what? We don't want to build and operate our own data centers. It's like an individual building and operating in their own house. So we're going to outsource that physical aspect to a third-party data center, like a data bank, right. And that's where you look at some of the early companies in the space Equinix, digital Realty which are still the two public guys. That's kind of how they started right, by meeting the requirements of enterprises that wanted to outsource their data center requirement. In the early 2010,. You had the rise of the public cloud, right. Obviously, at that time there were a lot of different companies competing for that right. It was looked at as an alternative to on-prem or to third-party co-location. You have companies like Rackspace and, if you remember, softlay or Enjoying and GoGrid. I was working for a cloud company called Boxl at the time and ultimately we were still kind of trying to tell customers hey, this is a better way to manage your IT infrastructure. Well, you know, the industry kept growing because what happened is enterprises just kept outsourcing and there was so much kind of on-prem data center demand that it just kind of propelled the sector forward. I'll tell you, in the mid-2010s, 2012, 13, 14, you know, people would ask people like us hey, how's the cloud going to impact your business? Right, because it was perceived that this way, where it's, either you're doing cloud or you're doing on-prem or third-party co-location. And fast forward to today, what's actually happened is that those are different things and ultimately, the cloud has been a huge accelerator of multi-tenant third-party data centers. Right, and it's happened in a number of ways. Number one enterprises have gotten more savvy about where they want to put their workloads, and it turns out that the cloud is really good for certain things, but for other types of workloads, especially what we call persistent workloads, it's not that great from a cost perspective. You don't get any of the benefit of that ability to burst up and burst down, right? So companies have become much more savvy about where they want to put their workloads, and that's this term that we use in IT. It's called a hybrid approach, right? Do some public cloud, do some private cloud, do some co-loc? But, more importantly, what happened is that, you know, the public cloud companies themselves grew so much and so quickly that they couldn't even satisfy their own demand. And today, if you look at the big four, right, Microsoft, Google, Amazon and Oracle, right, they probably outsourced because no one knows the exact number, but they outsourced about 50% of their data center capacity and built the other 50%. They just can't be everywhere and even their balance sheets can't support the tens of billions of dollars wired to build out all this infrastructure. And, as importantly as that, you know, when everyone talks about the cloud, we like to think about it with those four players, right, like it's some monolithic or oligopoly, right? The reality is it's not like to say there's this long tail of cloud companies. And if you think about businesses like NVIDIA or Akamai or Pure Storage or PayPal, any company that is a technology company, a lot of these companies are upset with themselves. We're going to build our technology inside co-located facilities and we're not going to build our own facilities. So the reality is, is that what a data center ultimately does? It underpins the ability for companies, organizations, for society as a whole to deploy technology, and all that technology has to go into a data center, which is why the center, you know, for the last 15 years, has had you know, call it double digit growth. And this is a real estate business at the end of the day, right? So that's a pretty attractive dynamic for a real estate type of business. And what's happened most recently is, you know, ai has appeared. Right, that is now the latest kind of technology that is highly, highly dependent on data center capacity, and what we've seen over the last six months is just really an unprecedented amount of demand to deploy AI in these data centers. And you've taken an industry that already had some pretty good tailwinds as a function of public cloud growth, as a function of enterprise outsourcing, as a function of increased technology adoption across the board, now has this new vector of growth, which is AI adoption, and then, for people like us, it's obviously a really exciting thing.

**Craig Smith:** 11:13

Yeah, and the AI and the cloud. They both kind of, from my understanding, grew up together. It's kind of chicken and egg, but that's what I've heard. I heard this from Michael Jordan, who's a very famous computer scientist, not a basketball player, who is an Amazon scholar, so he has some insight into Amazon. I also heard it from Ilya Syskover, who was one of the researchers on the original transformer algorithm paper, that the way Jordan put it is that Amazon was building these data centers for the to carry the computational load of their logistics calculations, which were AI based, and they eventually realized hey, a lot of other people need this, we can make it available publicly. And that was AWS. And the way Ilya said it to me is that the cloud existed but there was really no demand for it until people started doing AI and then there was demand for massive parallel computing, so in the early days. And that was GPU based. So in the early days it was all CPU, I imagine mostly CPU. When can you talk about the economics of the GPU based data centers to handle these AI loads, and particularly?

**Raul Martynek:** 13:03

And I'd like to love the comment a little bit about what you just said there about the AI in the public cloud Right, if you think about it. Yeah, I mean, you know some people that refer to the age that we were in. That's kind of the internet AI right. If you think about Google as a search engine right, because that's really what it is. Right. It's an advertising platform that delivers search results, generate billions of dollars of revenue. It absolutely uses AI to generate those results, to figure out what to present to the people and all that stuff. Right. You think about Meta, another big platform right. They used AI, obviously with their social algorithms. Right Into a recommendation Amazon's a little bit different, right. Amazon, at the end of the day, amazon was a computing platform, right. We call it IaaS, internet, our infrastructure as a service right. Amazon's proposition to enterprises is don't go buy your own servers and your own storage, just rent it from us, because we can do it better, cheaper, faster. They really didn't need AI for that. What they really needed was just great software engineers who knew how to deliver this physical infrastructure, like it was software, right. And then you had Oracle, which obviously with their databases a little different right, but I think what you know. So they were building out all their capacity to basically meet those business needs. And along comes ChatGPT, right, which is a new form of AI right. I would argue it's kind of like it's a general purpose AI. Maybe we're entering the business AI phase where that is a chat box. Obviously that's built on a large language model that gives out these eerily human-like responses. And now people are envisioning all kinds of instances to use that technology. You know, for example, bloomberg just recently announced that they're building a large language model called Bloomberg GPT, which is not all Bloomberg's financial data and financial insight to be able to output, I assume, responses around financial decisions how do you balance your portfolio and what's the risk reward, and things like that. So what we think is happening in the space today and data back, you know we have a kind of a unique perch that we sit in right Because we operate these 65 data centers with 3,000 customers and we get to see what equipment and what applications they're bringing into our data centers. I mean, they own all that equipment but we get to work with them on that right. So what we're seeing today to your question is that people are starting to say okay, I want to deploy some type of large language model or generative AI in our data center, and that's going to require a new type of processing. It's no longer going to be CPU based, it's going to be based on GPUs, and obviously we all know from reading the press about NVIDIA and how they are now in kind, of the crux of this massive movement, because their platform is now the platform of, you know, the first platform that people want to use to develop these large language models. So those are new types of workloads. I like to say they're a net new workload. It's something that has nothing to do with how the world operates today Because, again, every data center in the world today is, you know, somewhere between 70 and 100% utilized, and think about that. It's running the world today. So these are all new things that are coming into the data center because there's a new application that is useful.

**Craig Smith:** 16:34

Yeah, and those GPUs are much, much more expensive than CPUs. So how does that Impact the Ultimately the cost of inference that that end user is paying? I mean, all of these costs are built into that end cost. Is that right?

**Raul Martynek:** 16:56

That's correct. That's correct. Yeah, there's a couple layers of cost, echo that are that are pretty significant, right. As you just pointed out, the boxes themselves are just more expensive. The networking is working. This can't. This doesn't use Ethernet. This is used in finnaband, which is a very high-end networking technology that makes these clusters look like they're on the same wafer. So it seems like, right, all that stuff is a lot more expensive. Araca GPUs are a lot more expensive than Araca CPUs, right? In addition to that, these clusters consume and inordinate more, right? A CPU rat consumes five, ten, fifteen kilowatts of power per rat. We're seeing GPU rats consume ten, twenty, fifty, up to a hundred Kilowatts per rack. So you're talking about maybe ten times more power consumption Just in that same data center space, right? Ultimately, that all outputs to. You know these queries are very Intensive, right? If you think about, like, how much data this model is holding in RAM and its ability to take a prompt and then spit out a highly intelligent answer, there's a lot of computations going up, you know happening to just output that. You know that simple sentence that is a prompt, right? So Multiply that times millions and millions or tens of millions of prompts and you got a significant amount of new data center capacity that will need to be deployed if Businesses you don't want to, you know, adopt this technology and all indications are that. You know people are doing this right. We know Microsoft Because of their public announcements. Have, you know, identified a number of products where they're going to we chat cheaply and for their product set, one of them being a product called co-pilot, which allows on Microsoft Office 365 users to basically Hipchat cheap TO help them write word documents and power points and Excel files and things like that. That's gonna again, that's gonna require a hundred in our estimation, hundreds of megawatts of incremental data center capacity, and that's just one application hmm, and that's that pressure, that demand.

**Craig Smith:** 19:16

Presumably I haven't tracked GPU prices, but presumably that puts up upward pressure on GPU prices and so maybe you can talk about that. But also is there enough our generation capacity to handle this growth, as, let's say, in the United States?

**Raul Martynek:** 19:43

Well, you know there's, there's probably enough available power. The question is, where is it? Right, because you know data centers are yes, you said they're, they're a collection, right, and they need to be Located, in many cases, in proximity to other days. So one look, our, our view is that you know, in the short term, that the data center development timeline we measure things in years. Right, because we're building these very, you know, technologically advanced buildings, the parts that go into them, what we call the MEP. The lead times on that are 50 to 60 to 70 weeks, right, so you just don't pop up a data center like you pop up, you know, you know, a bar or something that you know you might be able to construct by you know Doing the whole depot, and it wouldn't stop together, right. So, so, so, our cycle is measured in years and you have this demand that is hit like tsunami over the last six months and and we believe that, no, there isn't enough, there isn't going to be enough data-centered capacity on next year or the year after To meet all the demand that we think is going to come from businesses wanting to adopt this generative AI technology and incorporating it into their, into another business. So, yes, the GPUs, you know Nvidia is doing a really, really good job of trying to prevent a secondary market in GPUs and creating, like a sub hub For GPUs, where you know they go to the highest bidders, because obviously that doesn't help. You know them Maintain, you know good, a good supply chain and so on. And once you got GPUs, though, you got to put them in the data, right. These are not boxes that can sit in the house, can sit in an office, can sit in a warehouse. They need to go into a data center, they need to be powered, they need to be cooled, they need to be secured, because they're very expensive and they need to be, you know, connected up to the global internet, right? So, so, yeah, I think, right, that's going to be a really interesting dynamic. Don't to see what happens over the next you know, six to twelve to 18 months, when, if this demand continues at the pace that we're seeing.

**Craig Smith:** 21:55

But data center development just can't accelerate off a dime Because it's more of a multi-year type of time- yeah, are there other GPU producers that are filling in back, filling where Nvidia maybe can't meet demand and there's there any question that Nvidia is going to fall short of demand?

**Raul Martynek:** 22:20

Yeah, two good questions. I mean, yeah, there are. Look, there are other players out there, right, obviously HP is a big player. They bought a crazy supercomputer, if you remember them. So they have, and they've recently announced I think was last month they announced their kind of own internal Machine learning type of offering right, based on some chips from Intel and some chips from from the craze that are more specialized, other specialized companies like graph core, there's a couple called Cerberus, that's that's built, the largest chip that's as yet been built and they claim is, you know, from certain types of applications, is more cost-effective. Then GPUs, I mean, look, I think it's like any new technological wave, right when the internet came out, right, it was like there were I don't, I think there was, there was Cisco routers and that was it right. I mean there, it was our well fleet, I think at the time from that mistake. So I mean, so what is that? Oh, what did all that demand do? It brought in more interest, right, and it brought in more supply. Obviously we're talking about, you know, highly, highly advanced Chips. That again, not something that can be started on the dime, but look, I think there's probably enough of an ecosystem where these other players are gonna benefit right from this demand because Probably in Vinny isn't gonna be able to deliver enough GPUs to satisfy the demand. That is going back, so that people are gonna turn somewhere rather than not do these projects because, you know, this is considered obviously a technology that can create massive competitive differentiations.

**Craig Smith:** 23:57

I yeah, and this is a bit of a dog leg on the conversation, but I'm interested in China. I've spent a lot of time in China and China is Scrambling because of the sanctions and certain entities from buying GPUs or advanced semiconductors. How, how can China Compete, how can it keep up with data center demand if it doesn't have access to Nvidia GPUs?

**Raul Martynek:** 24:34

that's a. That's an excellent question and in fact there was a book written called AI superpowers, which kind of talks, yeah. I think carefully , yeah, I would wage war between us and China. And I think, I mean, I think that is a. I mean I think the previous administration and the current administration's efforts to cut China away from Advanced chips number one and then, as you know, even advanced manufacturing, like the A of SL In Amsterdam, right, so I mean they're gonna, you know. So obviously that was China's reaction to try to accelerate their own super. It suits their chip semiconductor industry. They have, you know, programs where they're investing tens of billions of dollars to try to do that. But I listen, right now I would say the US is in the, in the Western powers are in a really good position that they want to press through advantage. Right, because there really aren't any substitutes for the most advanced chips, which are really the ones that you kind of need to drive or you know these, these types of technologies.

**Craig Smith:** 25:41

Yeah, on the cost side of your, you're filling new data centers with GPUs. Has the price Changed, I mean, and are you passing price increases on?

**Raul Martynek:** 25:58

Yeah, yeah, pricing is going up. I mean, look, you already had, you had the beginnings of that last year because the data center space, because there's a mention, we're basically technical real estate. Our pricing is based, broadly speaking, on two major things. One is our cost of capital right to his construction cost, and last year what we saw is because of inflation, because of interest rates, right, because of supply chain, not only cost. The capital went up, right. We used to be able to borrow at an absence, using Securitized that. Now we got to do it at 7%, right. So the cost of capital has gone up and the cost of construction has gotten up. So last year you already started seeing an increase in cost that was being driven by inputs, right. But now you're seeing even more dramatic increase in costs being driven by the man right that folks are. Again, there's just so much demand. You know we have been pricing deals 30 to 35 percent more in Q2 than we did in Q4 of last year, so six months ago, and we're getting 30 to 35 percent more per kilowatt Because of the input cost.

**Craig Smith:** 27:10

It is because of well, yeah, and just to clarify, you're not buying GPUs, You're providing the rock space for check, use yeah. Or do you also provide GPUs?

**Raul Martynek:** 27:26

Yeah, we do not know. We're sick , do you know what technology people deploy? That's one of the beauties that we like about our business, Wow right.

**Craig Smith:** 27:35

Well, where do you think the power consumption is headed? And there's a lot of concern. I mean, there's certainly a lot of research being done on making these systems more efficient, making training more efficient, but Ultimately there is a concern that, as this works its way into the economy I'm talking about generative AI that the percentage of National or global a power output Is of that's consumed by these data centers is going to grow. Have you guys looked at that analysis? Are we very far away from that being a real issue?

**Raul Martynek:** 28:24

Yeah, no, I think that is going to be an issue down the road. And look at this, is this good? We need to have a dialogue in this country about, you know how we're gonna, how we're gonna face that right, because we're obviously not Consuming power because we're just deciding to run, you know, in eight things. We're doing it because other consumers say, hey, we want this capacity, right. So I think there's a lot of questions about AI. I mean. Look, the US power grid, as you know, is going through, just like power grids all over the world, through a dramatic Transformation around. You know, fossil fuel to renewable right. I mean I'm based here in the state of Texas and you know the haircut power system, which is the Texas power system, I mean, has been extremely progressive at delivering new generation. You know, and a lot of that People might be amazed that Texas is the number one state in the country. Or solar and renewable right, because there's wide swaths of land in Texas and then the business climate is very, very you know, promoting that, those types of investments, right. So you know there are, and you know the main utility here in Texas on core is part of that. So there are utilities that are very progressive about, okay, working with data center operators to, you know, figure out how to bring more power to locations that make sense from a data center development. I mean, what I think is happening you saw it in Ashburn in Washington DC area last year, where the minion power ended up with, you know, in a shortage, right, and basically, how to tell data center developers we're not going to be able to deliver the power that we told you we thought we could, and what's that did? What that is done is it's lawn, it's pushed that demand into other markets like Dallas, like Atlanta, like Chicago, like Phoenix, like Reno, right, so it's moved them into other markets around the country. So there's gonna be a lot. I think what we'll see is that you know it is gonna be new, as we call it, data center clusters that appear. You know Ashburn market is the largest data center cluster in the world. It's probably got three gigawatts of installed data center capacity, right. The next biggest one would be, like Dallas, with, you know, five, six hundred. I think what ends up happening is you end up with many, many more regions in America where there's a gigawatt or a gigawatt and a half of data center capacity and by spreading out that load across the entire country. You know I think we'll do well, you know we'll do a good job at meeting that demand. But you know, ultimately it's hard to predict, right? I mean, we're in, we're in the first six months, seven months, of this phenomenon and, and certainly you know, these things tend to accelerate and they certainly tend to play out over, you know, a moldy year period.

**Craig Smith:** 31:14

Yeah, the. What about internationally? I mean, are you building data centers outside the US? Are there markets that, for example in Europe, that are becoming Like the, the data center clusters for you? Yeah.

**Raul Martynek:** 31:36

Yeah, so we're more focused on the US among our different sister companies. With the digital bridge, there are companies that are building in South America, scala data centers are building in South America, atlas edge in Europe, vintage is building in Asia. We regularly dialogue with those companies so we're aware of those trends and, absolutely, you know, there's a lot of data center development, probably, on a percentage basis, actually more growth In those markets because they had started further behind than the US market, right? So from an absolute numbers perspective, the US is the largest market. It will continue to be the largest market. From a percentage group perspective, South America, Europe, Asia and now even, like the Middle East and in some parts of Africa, specifically South Africa, are starting to see, you know, significant data center absorption. So, yeah, this is gonna be a global phenomenon. I mean, you know the again, as I talked earlier, I really think it's purely about, you know, people consuming technology and when you think about that trend and the fact that there's eight billion humans on this planet, you know Most of them have a smartphone. Now there's all these mobile networks, right? So people want to interact with technology. That interacting with technology be a tick tock or be it threads or be it, whatever ends up. Driving data.

**Craig Smith:** 33:00

Gonna fall behind economically If they don't have adequate data center capacity. I mean a fall behind the US, if indeed Generative AI is what's going to be driving Economies going forward.

**Raul Martynek:** 33:22

That's a really interesting proposition. And, look, I would say, you know, if you think about it, you know you start economies, Asian economies, right, I mean a lot of them. I've been performing extremely well, of course, right, I tried to obviously have some issues now, but, you know, long through, long-term growth has been incredibly impressive. But, yeah, you could see a scenario where if you do not access the chips, and then you don't have access to, you know, data center capacity because you have an inferior power Power grid, right, and then other countries are developing very advanced forms of AI. That you know. As you know, part of the fear of AI is that it will kind of lie to our displaced workers, right? What happens if that's the case with certain of these large language models and that starts to now impact Workers in foreign countries and they don't have a method to to participate in that? So I think there's a huge amount of open questions that this Next technological revolution is going to, to, to to pose.

**Craig Smith:** 34:28

Yeah, and in terms of a data data center capacity, do you have any metrics of how it's grown in the last Six months and how you expect it to grow in the next five years? Let's say well.

**Raul Martynek:** 34:47

So you know people. They usually come out once a year on that stuff. So at the end of last year McKinsey King came out with a study that basically said that US data center capacity was gonna triple To like 17 gigawatts or I was a global number, I believe 17 gigawatts, you know, over the next seven or eight years. You know this AI thing is so new. No one has really, I think, put together a real Estimate. I did speak to one analyst that you know that is really Well plugged into the space, that you're doing some sizing and their view is that you know, from a base case perspective, ai sure out of minimal increased data center demand by 30% and it's entirely reasonable that it would double it right, so instead of 17 gigawatts We'd be going to 34 gigawatts. These are staggering figures.

**Craig Smith:** 35:47

Yeah, really Wow. So where are you guys building right now? We?

**Raul Martynek:** 35:53

are building in 14 of 12 different markets right now. You know, we're building in a variety of like what we call tier one markets like Ashburn, like Atlanta, like Dallas, but we're also building in markets like San Diego, at Seattle, in Minneapolis, right. I mean, our view around AI is, you know, it's caught up in this, this other technological trend called the edge right, which is this kind of Decentralization of the internet. We think that is also a trend that is occurring and that is going to become more prevalent over the next 10 years. And that's the reason we have these large geographic footprints, because we believe that customers will look to place Really late, see sensitive workloads in metropolitan areas to service the users in that matter. Metropolitan area as opposed to today where you kind of pick one or two spots in America and you're close enough to everyone, right. But ultimately, I think the quest for more real-time applications is going to necessitate that movement of infrastructure Into, call it, the top 30 metropolitan areas in the US.

**Craig Smith:** 37:03

Yeah, the On Part of the power consumption you were talking about is For for these data centers. It is cooling. What percentage of a data center power consumption is on the cooling side and does the technology the coolant, cooling technology that you use? Very depending on where you are Geographically.

**Raul Martynek:** 37:36

Who to the question yes, so to the way that, that way that gets measured. There's actually a specific Measurement term for that. It's called PUE I percentage utilization effectiveness. Right, it really it's. It's a simple thing. We take two numbers. Let's say you build a data center, it's a ten megawatt thing, right. Let's say you're delivering ten megawatts of power To that equipment and it's consuming it, so that becomes the denominator. And then we say, hey, we look at our utility bill and say, hey, how many megawatts are we consuming? So, for delivering ten megawatts to its equipment, we're consuming, you know, ten or thirteen megawatts. Well, it's one point three divided by one. So our PUE is one point three, right, so that's the way. Sure, the extra energy that needs to go into a data center to keep the lights on, no pun intended, is also mainly to cool the data center equipment, right, and absolutely. There's different technologies for that and it does vary by climate, I mean if you know, if you're in the north, I mean, believe it or not, you know a lot of this. Technologies are advanced to the point where you can take advantage of what we call a free cooling right, where we can actually the chilling manufacturers of built systems that allow you, when the temperature falls below a certain temperature, to be able to just run fans and use the outside air to cool the liquid that then is circulating through the data center that pulls the equipment right. So we, you know, we like other players in the space to play that type of technology and it's really effective, even in places like Dallas, you believe, you believe it or not. It's obviously very hot during the summer, but In the winter it gets pretty cold, so there's a lot of hours of free cooling. Of course there's more hours of free cooling than Minneapolis, right? Ultimately, like any technology, there's different price points. You know, we, we, you know there's ways to get that PUE all the way down to, you know, 1.1. Well, guess what? All of a sudden, if you do that, you've got to increase your price by, you know, by 30% or 40%. Then customers like now, we don't want that right. So there's a balance between cost and efficiency and obviously, as a data center operator, we're constantly looking at that balance and working with our customers and our vendors to try to figure out how to improve it. It's definitely improved over time. You know, 10 years ago, 12 years ago, most data centers had a PUE of two right, so for 10 megawatts they were consuming another 10 just to cool the building and all that. Now almost all designs coming out of the gate have a design capacity of 1.2 or 1.3, so things have gotten really, really efficient and they'll continue to get a push.

**Craig Smith:** 40:30

And you were saying that. Well, what, what percentage of data center center capacity right now is to run AI or particularly generative AI systems, and very small right, because it just started right.

**Raul Martynek:** 40:51

So I think today, I think that number, today again, there's so much data center capacity installed out there today running the world as we know it, and now this AI stuff shows up and I would probably say there's only a couple hundred megawatts Of data center capacity that's been dedicated to AI today, again, I think that number, you know, goes into the gigawatts over, you know, the next couple years.

**Craig Smith:** 41:17

Yeah, and in percentage terms of capacity, that would go from what to what?

**Raul Martynek:** 41:23

Yeah, that's another interesting thing, I know. And that one you know picks a number, right. I mean, I've seen some analysts reports that said that by 2030, 80% of workloads are going to be AI workloads. Craig, I don't know if I knew. I can't predict the future that well, but you know so. Well, look, I mean, if you think about the potential for that technology, right, if you think about the fact that all I could say is chat, gpt, that's Windows 3.1.1. That's the latest version of that technology you're ever going to account for. Right? That's frightening, right, I mean, it's exciting at the same time, right? So my point is it's entirely possible that AI cannibalizes existing workloads because it just ends up doing it, or right, but then it's gonna. But then what you're gonna do is just ship computing power over here To the AI workload, which would be a reason, or I should say, a scenario, that would say yes, in the future you're gonna have, you know, a huge amount of workloads that are AI.

**Craig Smith:** 42:26

Yeah, and you were saying that currently, the big four cloud providers offload a certain amount of their capacity to Independent data center operators. What, what percentage did you say that was?

**Raul Martynek:** 42:43

I'd say a good number on average is about 50%. 50% of where data center capacity they procure through relationships with companies like data.

**Craig Smith:** 42:54

Yeah, and are they? Are they continuing to? I mean, I know they are, but is that? Are they going to continue to build capacity themselves, or will they do both?

**Raul Martynek:** 43:08

They will do both. It's just a, it's just a way to scale. I mean an analogy. One analogy I like to use is like I think about. Think about someone like Chase Manhattan back right. They all understand scrapers in Manhattan. They probably own mortgage servicing buildings that house tens of thousands of employees. But this chase tonight and build their branch office. Yeah they are just the least right. I mean. So there's a lot, because they have a lot of different needs, from very large needs to very small needs, and it just doesn't really make sense for them to try to do it all themselves. It's just more efficient To be able to outsource that and and, kind of, you know, leverage the capacity, the balance sheet, the capability of that ecosystem.

**Craig Smith:** 43:52

Yeah, wow, this is. This has been fascinating. Is there anything I haven't asked about? How is impacting data centers, or the cost of operating data centers is impacting the inference cost of generative AI? I mean, that's really what I find most interesting.

**Raul Martynek:** 44:14

Yeah, I mean I said you know, you've asked a lot of great questions and I just think there's a lot of questions we don't even know what to ask because we really haven't entered the production, right, I mean, for the most part, what we're seeing is that a lot of this AI is still training. Right, it's people that Are building, like Bloomberg, are going to build a large language model, are going to train it and then they're going to deploy, right, and then that's when I think the business model starts to take hold. Right, because, presumably, after you start to deploy and it didn't, you understand, kind of, you know what's the revenue potential, what's the profit potential Of that software, right, I mean, I like this, the, I was like sass. When sass came out, right, where that word came out, it was kind of a strange word, no one knew exactly what it meant, right, it's like software as a service. Right, we used to buy license software. Well, that's right, you use to buy a license and you kind of owned it, right, and now you went to a month in recurring month, right, so, and think about how many saas companies have been created over the last 10, just almost up, you know, and the number one amount of companies. I think if I take that same route right that ultimately A lot of people will figure out how to apply that to just a huge number of different use cases and then they'll develop an economic model on that, and I would also imagine that you know they're. You know this is again the first version of this technology that many people are gonna be working around how to make this more efficient With inference so that it doesn't have to use so much power and so much computer. That might be another way where the industry will, will you know, kind of help us scale itself.

**Craig Smith:** 45:49

That's it for this week's podcast. I want to thank Raul for his time. If you want to read a transcript of this conversation, you can find one on our website. That's YOU find O N dot a. I also want to encourage you to visit netsuite.com/EYEONAI for an unprecedented offer by the number one cloud financial system. You can download a custom KPI checklist designed to give you consistently excellent performance. It's absolutely free at net suite dot com. Slash I on a I. That's net suite dot com. Slash E Y E O N a I all run together. So go to netsuite.com/EYEONAI to get your own KPI checklist again. netsuite.com/EYEONAI all run together. I hope you'll support them because they're supporting me. And remember, the singularity may not be near, but AI is about to change your world, so pay attention.