S1 EP4 - Which Fabric for NVMe Storage Area Network?

Wednesday, October 20, 2021 · 21:10

Todd Owens, Field Marketing Director discusses NVMe over Fabrics with podcast host Chris Banuelos. Though NVMe is not new, moving it outside the server into storage fabric is. The efficiency and multi-thread capability of the NVMe language when compared to SCSI (small computer system interface) makes transitioning to NVMe over Fabrics a no-brainer. The complex challenge is which type of fabric approach is best. Arrays exist today that support NVMe over RoCE and FC-NVMe. However, there is much discussion around new arrays with NVMe over Fabric TCP support. Learn more about the three different approaches as well as the pros and cons.

Speaker
Todd Owens
Field Marketing Director

Host
Christopher Banuelos
Senior Manager of
Global Social Media Marketing

Christopher Banuelos 00:04
Welcome to the Marvell essential technology podcast. I'm your host, Chris Banuelos. On today's episode I discuss "what fabric are you planning on using for your NVMe storage area network?" with Todd Owens, Field Marketing Director hear insights from Todd on his career, significant changes within the storage industry and what solutions Marvel offers in providing optionality around NVMe over fabric. Todd, you know, I remember having a conversation with you a while back, and you were talking about how there's been so many significant changes over the last 30 years, specifically in the storage industry, why don't you give our listeners today, just a brief introduction of your career and how you got started.

Todd Owens 00:47
The cool thing is, I've been in the storage business pretty much all my career. And here at Marvell today, I'm focused on driving education, awareness, and demand for our products and technology to a specific OEM customer that we have, Hewlett Packard Enterprise. The reason for that is I worked for that company for about 20 years so I have a lot of connections within HPE and I kind of know how they work and operate, and I can help them sell more of our technology in their products and solutions as a result of that, you know, that experience that I have. Here at Marvell, it's a great organization to work with, because I get exposed to a lot of different business units, because we're trying to deploy a lot of different technology into HPE. And I get to work across a lot of different organizations within Marvell itself, because I rely on other people within the organization like our corporate marketing teams, for a lot of the back end support that we need to actually market our products and technology to and through our customers.

Christopher Banuelos 02:01
Todd, going back to the beginning of your career, what's a really interesting story that sticks out to you?

T **Todd Owens 02:06**
Ford Aerospace was my first real job. And I worked on a thing called HPE 1002. To replace the disk drives, it took two of us to lift the disk out of the big it was honestly, like a washing machine or you know, you had to pull this drawer out. And he takes this big thing, the disks were probably, I want to say, probably 32 inches in diameter. And there were these, they were called platters back then, and you had six, eight, ten platters in each one of these assemblies. And it stored a whopping I think, 50 megabytes of storage or something like that, you know, and now we're looking at terabyte thumb drives that we have. So it's been a pretty interesting ride and where the technology has gone, has come from and where it is today.

C **Christopher Banuelos 03:01**
Todd, I think you could provide our listeners with a very interesting perspective in that describing what the storage industry was, like 30 years ago, what was that environment like?

T **Todd Owens 03:10**
I started as an aerospace engineer in the early 80s. In the late 80s, I moved into the computer business and started in the storage group at Hewlett Packard. And at that time, we had a lot of proprietary interfaces for storage. So the storage devices were big as your washing machine. They didn't store very much. You've got more storage on your iPhone today. Then you could fill a roomful of storage devices back then. But everything was kind of proprietary. And then in the middle of the 80s, this thing called SCSI was developed, and it started to become mainstream early in US in the late 80s and really, in the early 90s, and SCSI (Small Computer System Interface) became the standard for connecting a computer or a server to an external peripheral. And I say peripheral because it wasn't just disk drives, like we think of today, back then you could connect that that computer to a scanner, a printer, a tape drive, or a disk drive. And over time, this this concept of client server computing came into being in the early 90s. And that precipitated the need for multiple servers to talk to the same peripheral, primarily tape drives because we were doing backup, you know you so yeah, it's just like we do today that hasn't changed. You've got a whole bunch of devices that have critical information on them and you want to copy that and save it away in case something gets destroyed, and you can rebuild it. So back then you had all these, these need for external storage connected to these servers and primarily driven by back up, but then people decided, hey, I can connect disk drives and I can save the amount of storage I have. And so SCSI was the big deal. That's the protocol. And then you had basically a transport. How do you how do you transport that protocol across a wire to those storage devices? Back then Fibre Channel was developed specifically for that it was purpose-built to connect servers to shared storage devices. So that was, you know, the 90s. And then in the middle of the 90s, a thing people decided, you know, we got these other wires running around the lab, these Ethernet things that connect all of our computers together, and so why can't we do this with Ethernet? So iSCSI came along. They call it internet SCSI, and it was based on running SCSI commands over the Ethernet protocol. And that was done primarily because you had a lot of people in the IT department who knew all about Ethernet. It wasn't a mystery because it connected everything together, right? That's how we connect to the World Wide Web, right? The Fibre Channel tended to be sort of a science fiction movie, you had to have special skills and things to make it work. And but over time, that's all moved. So if we, if we move to just maybe five years ago, the Fibre Channel industry had started, you know, they keep bumping speeds up, you know, with bandwidth, but they also started making significant improvements in manageability, and how you set up and configure the environment. And fundamentally the Fibre Channel guys had made their infrastructure as simple to set up as anything you would do with Ethernet. In fact, in many cases, it's actually easier. And then you had the continued development of iSCSI, and higher bandwidths, you know, 10 gig, 25 gig, etc.

C **Christopher Banuelos 06:56**
You know, Todd, one of the things I think could be helpful is to understand what exactly SCSI is, could you give us a brief description?

T **Todd Owens 07:03**
So SCSI is a language, if you will, that allows a host system, which is referred to as an initiator to talk to a peripheral today, primarily disk drives, and send commands like I want to read from a particular portion of that drive, or I want to write to a particular portion of that dry. It's a command set that is very prescriptive in terms of has headers and payloads and error correction, coding and all kinds of good stuff in it. But just think of it as a language on how compute devices, servers, talk to storage devices.

C **Christopher Banuelos 07:44**
What are some of the trends that are happening in the industry?

T **Todd Owens 07:47**
The other thing that's happened in the industry is the advent of solid state storage, or SSDs, we call them right, that's not something that's new. We've got those in our laptops, they're now prevalent in servers. And we're seeing them in storage arrays, well, that precipitated the need for a new protocol called NVMe, or non-volatile memory express. And that's used inside your laptop, that's used inside your server to talk to these really high speed, you know, solid state drives, these SSDs. Where we are today is people want to take advantage of that high performance storage, and be able to connect to it over a network. And that's created this thing called NVMe over Fabrics. And that's kind of the big buzz in the industry today that Marvell has been playing a major role.

C **Christopher Banuelos 08:45**
What are some of the current pain points that storage administrators are experiencing as they transition to the latest generation of storage technologies?

T **Todd Owens 08:53**
Well, you know, the thing that makes me really happy is solving people's problems, right? So if I can make somebody else's life a little bit better, that's a good day for Todd. And, you know, when you look at some of this technology that we have here at Marvell, it actually does that--it helps people solve problems. And so when you look at the connectivity technology we're talking about here, the storage administrators and the system administrators, they kind of share the same pain points. And it's not necessarily new pain points because of all this new technology. It's kind of the same pain points they've always had. It centers around a couple things. One is compatibility and manageability. They really want whatever they deploy the latest and greatest stuff. They want it to work in their existing environment. It's like if you were going to go buy a new car, but oh, by the way, it doesn't work on the road in front of your house. You have to go to this special road you use it like it's a NASCAR and you got to run it on a track. Maybe that's not exactly what I'm looking for. They want the standard stuff that they have today to work with a new stuff that's coming in tomorrow, they also need an ecosystem, right? The new technologies have to work with an operating system, they have to work with the infrastructure they have today. And they have to work with storage devices, right? So all of those things have to line up before any of this stuff becomes a value to those administrators. And they need to work as seamlessly as possible. They need to coexist with their existing environment. You know, if you were to go out and buy something new and say, well, by the way, you have to throw out all this old stuff if you're going to bring this new stuff in, and probably not, you know, good on the old return on investment calculator, right. So you know, it's got--the things have to work. And the beautiful thing about what the industry is doing, and what Marvell specifically is doing, in this area of connectivity, is that these new protocols, and I mentioned NVMe over RoCE, NVMe over Fibre Channel, and NVMe over TCP. These are the three protocols that are the next generation that are going to be replacing SCSI, and moving that NVMe protocol outside of the server, into the network. What we're doing in that space, all of these are going to interoperate with the existing environment. So the cool thing is our FC NVMe solutions with our QLogic Fibre Channel HBAs-They actually have drivers today that it's a single driver that works with both SCSI and NVMe. And it allows you to, to have these new NVMe native storage devices, and your legacy SCSI devices on the same storage area network. And that's really important to these administrators.

C **Christopher Banuelos 11:59**
It seems like there's a variety of options around NVMe over fabric. Can you talk about the pros and cons of each approach? And Todd, what, what options serve networks best?

T **Todd Owens 12:10**
Well, that's a great question. And I got an easy answer for you on the last part of that is, is the answer is it depends, and depends on the customer. And I'll address that here in just a second. Let's go through the three

different protocols for this, you know, NVMe over fabric environment. The first is NVMe over RoCE. Now, RoCE stands for RDMA over converged Ethernet. Now that converged Ethernet is kind of a unique little identifier, it's a special kind of Ethernet. It's not necessarily compatible with a network you and I are running on right now, which is a standard TCP IP network. NVMe over RoCE or RoCE itself requires this conversion Ethernet, which is also known as a lossless Ethernet. That means it needs to be a special Ethernet network with special capabilities. They're a bunch of fancy three letter acronyms, like DCB for Data Center Bridging and a whole bunch of other good stuff. But let's suffice it to say, you have to build a superhighway Ethernet network for this NVMe over RoCE to work on. And that means there's extra steps as a network administrator you have to take. And it doesn't scale as much as a standard Ethernet network that runs on TCP does. So one of the, you know, comparing NVMe over RoCE to what we do today, you're going to have a special network setup, it's going to be a little bit more complex, maybe than what you're using today. And it's only going to scale probably within a rack of servers, it's probably not going to scale across your data center. But the reason NVMe over RoCE is talked about today is because it is absolutely the lowest latency of the three protocols we're going to talk about. So if you really need the high absolute highest performance between a server or a cluster of servers, and in particular storage device, man, this is the this is the protocol for you. Okay, so that's why I said it depends. If your number one priority is performance, NVMe over RoCE is is probably the way to go. It's going to cost you a little bit in terms of manageability and scalability, but it's going to give you great performance. The other one that's available today and has a has probably the most robust ecosystem is in Fibre Channel over NVMe or FC NVMe as we call it. Now, it's got a very mature ecosystem. It's supported by VMware Linux Windows today, all the tier one guys support FC NVMe. There's a plethora of storage vendors out there that have FC NVMe native arrays and all of the HBAs and fibre channels, switches and directors that you need to connect all those things together-they already support FC NVMe. So it's very mature. And the cool thing is, for those customers who are using Fibre Channel today, it's an absolutely seamless transition to move to FC NVMe. I mentioned earlier, the drivers we have here at Marvell, for our QLogic HPAs. They are universal drivers they support both SCSI and NVMe. So the same adapter you're using today, to connect to your SCSI based storage will, in fact, connect to these new NVMe storage devices. So the third protocol that people are talking about in this NVMe over Fabric world is the kind of the new kid on the block. That's not new. And it was the standard was developed in 2018. And it's called NVMe over TCP. Now, what this does is allows me to use this new language NVMe and transport it over the network you and I are talking on right now: the standard TCP IP network. So no, no special tricks in the network. I don't have to create the superhighway. But I can connect servers to storage devices over the standard Ethernet network. Now, if we if we look at where storage has been for the last couple years, and where it's heading with this new NVMe over fabric,

Todd Owens 16:27
The real challenge with NVMe over TCP is it's the new kid on the block and the ecosystem isn't there. So it's really Linux only today. VMware and Windows Server from Microsoft are catching up. You know, we'll see NVMe over TCP support there, you know, and probably then, you know, early 2022 timeframe. But today, you're very limited if you if you want to go that route, plus, there's no storage. Now there are a few boutique storage vendors that support NVMe over TCP today, but the big the big guns, the top tier storage providers don't yet have storage arrays that talk natively NVMe over TCP, that's going to change by the end of this year. But we'll start to see some flow out but probably a couple years away before that becomes mainstream.

Christopher Banuelos 17:17
what happens when NVMe over fabric goes mainstream. Where do we expect customers to land?

Todd Owens 17:23
Well, where we see this going is where it's already been. If you look at the SCSI world today, there are three protocols that people use with SCSI. We've talked about two of them: Fibre Channel, and iSCSI. There's another one I didn't talk about, and it's called FCoE, it's Fibre Channel over Ethernet now, was very popular about five years ago. In fact, it's gonna take over the whole Fibre Channel world and Fibre Channel is gonna go away. Well,

it didn't happen that way. Fibre Channel over Ethernet became a niche player in things like blade systems, where you have a confined environment, because FCoE was the same thing. It's Fibre Channel over converged Ethernet, it's that same thing that RoCE is using, right. And the problem is it doesn't scale and it's very complex to set up and maintain. Well, the new world is going to be just like today's SCSI world, you're going to have Fibre Channel over NVMe being used by the people who use Fibre Channel today. NVMe over RoCE is going to find a place but it's going to be a niche place, kind of where FCoE is not necessarily in blade servers, maybe. But in AI and machine learning. In those kind of environments, you're going to have a small cluster of servers of high performance servers that need these huge, large data stores. And they need to be able to access that data and compute on that data very, very quickly. So NVMe over RoCE makes a perfect type solution in that kind of, you know, defined set, but it's not necessarily going to be a wide like a wide stream type of adoption for NVMe over RoCE. And that leaves us with NVMe over TCP, which is equivalent to iSCSI. That's for the people who want to use their standard Ethernet network to not only have their network traffic, but also have their storage traffic. They're very comfortable with an Ethernet environment. And they're going to go with NVMe over TCP. And so you know, we look at all of those three things, coexisting; again, the NVMe over Fibre Channel is going to be in the data center for your mission critical you know when the data is the business, that's going to be the protocol people use. For for the SMB, or the smaller, less complex customers, you're going to see NVMe over TCP and NVMe over RoCE will be in some very specific use cases.

C **Christopher Banuelos 19:57**
Where can our listeners find more information If they're interested in learning more,

T **Todd Owens 20:02**
We have a website Marvell www.marvell.com/hpe. That will give you a good overview of all the different products and technologies that we're working with HPE on today. Specific to the Fibre Channel business that we're in here at Marvell, the other website to go to is www.marvell.com/qlogic that will take you to our Fibre Channel, adapter products page. And there you can learn about all of the cool features and capabilities that we build into our Fibre Channel connectivity solutions.

C **Christopher Banuelos 20:43**
Todd, I wanted to take the time to thank you for being on today's episode, thought we had a great conversation today. I look forward to doing a part two learning more about Fibre Channel and the technology that you and your team are working with. All righty, take care. Thank you for listening to the Marvel essential technology podcast. As always, please feel free to visit our website to learn more. And we'll see you on the next episode.