

Would you like to start off by telling me a little bit about your company, who you are, that kind of thing?

WHOA.com is a cybersecure cloud hosting platform with an emphasis on compliance workloads, focusing around HIPAA regulation and PCI regulation. Certified to ISO 27001 company (based on the 2013 standards). DCs in LV, NV and Miami, FL, by the end of this quarter they'll have a presence in Grand Rapids MI. Business established in 2013

So what were some of the challenges you faced as you started to build up the business and grow?

Storage was one of the challenges we faced. We started out on FlexPod architecture. Deployed a few NetApp arrays in our DCs and were standardising on Clustered Data OnTap. It was new in the industry at that time. It had good advantages for multi-tenancy – this was important to us. Ran into a lot of challenges in the amount of overhead it took to maintain that platform. We are a pretty technically savvy team, and **“we didn't want to be scaling by head count – we wanted to scale by technology”**. When we started to look at the processes and procedures being done on the storage infrastructure on a daily basis, and renewals came up as well, the price tag on the renewals pushed them to start looking at other vendors. Found Tintri's offering, and found their administrative and operational overheads were extremely low, to the point where we had a single, dedicated NetApp admin at one point. When we started to deprecate the NetApp platform and bring in the Tintri units, we actually just cross-trained all of our VMware admins on the Tintri arrays, and now, rather than having one person and looking to grow that team, we actually have all of our standard engineers know how to operate and use the Tintri units. In addition to that, we've tied into the APIs pretty heavily, and have built some product offerings around Tintri itself, one of them being the BC/DR product that allows a person from an on-prem or private cloud to replicate their Tintri snapshots into us. It's fully orchestrated. It starts out with basically a sign-up form. In less than 20 minutes they can begin replicating data into our data centres. Once that replication is complete, we made it very simple for them. We built a whole UI around it that allows the customer to say “Hey, I want to take this snapshot and I want to restore it on your cloud platform”. It's less than ten clicks. You can take that snapshot that's sitting in the data centre and orchestrate it up to a virtual machine in our cloud environment.

I guess you've gone from a traditional hybrid array with the FlexPod to the all-flash with Tintri, were there any other solutions you looked at? Or technology you evaluated as part of this? I guess, in my experience, a lot of service providers do this in different ways. Some of them do software-defined, some of them do cheap and deep storage and let the cache sort it out for clients. What did you guys look at as far as that goes?

We looked at some of the traditional arrays that were prevalent at that time (both AFA and hybrid arrays). We looked at Pure, Nimble, Tegile, and a few other vendors. One of the things that really drove me towards Tintri was the simplicity. Dealing with the other vendors, some have FC requirements, they don't support NFS. I'm a huge NFS fan. We **optimise our network for NFS traffic**

**and get really, really good results operating NFS in our infrastructure. Again, Tintri being an NFS-based platform, it was really an easy choice from there.** But also the API integration. Again, we are pulling API integration for snapshot information. For example, that product offering [BC/DR]. We need to calculate how much data a certain tenant is using for a certain VM to be able to showback in our portal. We're able to do that through the API, where we looked at other storage vendors and the API just wasn't robust enough for us to do those kind of operations. The other side of it too is that **I was able to eliminate an entire fibre channel fabric within my data centre. So again, paring down the employees that are required to operate that platform. That's a huge cost saving for me because at the end of the day head count is typically one of the most expensive things to operate a cloud infrastructure.**

The catchy tech on most people's lips is hyperconverged – is that something you evaluated or didn't look at? Did you just want to stick with traditional? We did look at the hyperconverged stuff. At the time that we started the infrastructure really, Nutanix was kind of the only market player out there (other vendors were trying to catch up). Obviously, being a Cisco shop, Cisco came in and showed us their hyperconverged platform. It was an impressive platform, but it had some growth challenges with us. We need to be able to scale individual sections of the platform. When I say individual sections I'm talking about compute vs network vs storage. In our Production clusters (and when I say production that's where tenant workloads are operating for us), a lot of times chasing RAM and CPU requirements kind of throws back to the old storage days where you're trying to hit an IO performance number or a latency number and you end up with this massive amount of spindles and a massive amount of capacity that's really waste in order to hit that one performance goal. You know see similar things happening in hyperconverged for us, as a service provider, that kind of throws those numbers out of whack. **Maybe I need to increase my CPU and RAM capacity to meet a customer's workload. Well now I'm sitting on a bunch of excess storage that I really don't want to power, I really don't want to cool, and I really don't want to manage, because it's not needed. So, we still kind of went with that "broken-out" architecture where we have storage arrays or storage appliances and we have UCS blade systems so we can increase RAM and increase CPU to the customer's workloads without having to scale out our storage at the same time where it might not be used.**

I'm curious why you went all-flash as opposed to a hybrid solution? I know the pricing is getting closer and closer every day. Was that a question of density? Economics? Performance?

Full disclosure – we have both. We have both hybrid and we have AFAs. We needed to put that in place because we needed a tiered offering. I can tell you, the shift in the market and the shift in our vision is to move to all flash only. Then at that point we can put SLAs and different parameters around IOPS and latency for customers. But **right now we have customers who are demanding that lower tier. And a lot of times they're trying to hit a price point, they're not trying to hit a performance point.** Us, as a service provider, definitely see an advantage of having an all flash array, especially if it has got a really good dedupe technology, because you can pack a lot of data in on

those SSD drives and still get and maintain that performance. Again, pulled almost an entire rack of NetApp spinning disk and reduced it down to a few RU of Tintri, so you know. I've got cooling advantages that I'm getting there, reduced power consumption, and, again, reduced head count. Those things are really important to me from a data centre operator perspective. Those are costs I'm paying close attention to.

You talked before about how Tintri's given you a reduced admin overhead which is pretty important. You managed to get your VMware techs up and running with Tintri which helps with not having to have dedicated storage bods anymore. Are there any other reasons, beyond the simplicity and API, anything else you want to talk about as far as the benefits of using Tintri? All the good stuff about it?

When you get into looking at some of the technical operations of Tintri, how Tintri works. **The secret sauce is really that per-VM management and capability. We had NetApp in place. We have a data centre in Miami and one in Las Vegas and we were looking to replicate customers between and I really ran into a bunch of challenges with NetApp because I've got to do it at the volume level, whereas Tintri I do it at the individual virtual machine level.** So, right there's a use case for data protection. Another use case is around Performance and QoS. To be completely honest with you, we are using the default, out of the box Tintri QoS settings. In other words, we're not applying QoS, we're letting the auto QoS handle that. And we've got some massive database / SQL servers that are really pounding a lot of the IO and latency out of these boxes, but Tintri stops that noisy neighbour and also gives us the visibility of those metrics right at our fingertips. That's definitely a huge plus for them and it's very attractive to me to be able to do that. We're in the process of actually taking / consuming more of that API and in our portal that we present out to our customers to manage their infrastructure we're now going to start rolling out that latency and IO throughput data back to the customers so they can see the performance of the virtual machines. The visibility is really huge. That's one of the things we pride ourselves on in our cloud platform is we give the customers full visibility of not only capacity but also performance.

Is there anything you else you wanted to cover? You've been pretty happy with the solution so far. Is there anything, moving forward? You're going to do a bit more with the API. Anything else you're after in the product?

We're really looking forward to them putting some of the more advanced features in the analytics portal. That has been something we've been playing around with quite a bit. And the fact that they're now going to be tying in to see our compute and our storage metrics. We're actually looking to dedicate our capacity management and capacity planning around that tool. It has all the metrics that we're looking for. Again, full disclosure, I'm a part of the Tintri Advisory Board, so I get to see a sneak peek of what's coming down the pipe with the analytics stuff and it's very attractive to us and we're very eager to start implementing some of that stuff in production.