WiNG

A Q&A With Mohit Aron, CEO Of Cohesity

JUNE 27, 2016 MARTIN GILES

Wing is an investor in Cohesity, whose chief executive, Mohit Aron, is a pioneer of hyperconvergence, which brings together compute and storage to dramatically simplify data center infrastructure. Before launching Cohesity, Mohit was a founder of Nutanix, which is currently disrupting the world of primary storage and servers. Prior to that he was at Google, where he was involved in the development of Google File System, which is the distributed storage foundation powering most of the company's web properties.

As Cohesity celebrates the first anniversary of its public launch (http://www.cohesity.com/press/cohesity-achieves-blockbuster-growth/) , it seemed a good time to catch up with Mohit and to talk about the company's progress.

Q: Mohit, what is the core problem that Cohesity set out to solve?

A: We liken storage in a data center to an iceberg. The tip of the iceberg is primary storage, which is where people run mission-critical applications and require strict service-level agreements. The second, much bigger section of the iceberg is secondary storage, which involves applications that aren't mission critical, including backups, test/dev, and analytics. We set out to solve several problems related to this area.

The first of these is fragmentation. Different workflows are handled by products from different vendors, so customers have to juggle multiple licenses and user interfaces. Problem number two is inefficiency. There are often multiple copies of the same data throughout a data center. Furthermore, some storage infrastructure is active and other bits are idle or underutilized. The third problem is "dark data", which refers to the fact that companies often lack deep insight into what data exists in their secondary storage.

Our vision is to build a Google-like, webscale storage system that scales infinitely and consolidates all of the workflows on one platform. We refer to this vision as hyperconverged secondary storage.

Our contention is that prior approaches to implementing storage for secondary data have been underwhelming because they've only looked at point problems such as deduplication. We are the first ones to look at this whole space holistically. With Cohesity, fragmentation is gone because there's a single platform. Inefficiency's gone because the platform is active all the time and we globally deduplicate all redundant data. And dark data's lit up because we have analytics that give customers deep insights into their data. We move compute to the data, rather than the other way round.

Q: Can you explain the difference between "convergence" and "hyperconvergence"?

A: In the 1990s, EMC and NetApp promoted the idea of storage as a first-class citizen. This meant that storage needed to sit on one side of the network and compute on the other, and the separation was seen as valuable because storage admins could now focus on storage while applications could run separately. This was a great concept, but as storage has grown exponentially over time you now find you have to architect and install very expensive networking infrastructure to connect the pieces.

So some companies said instead of buying storage, compute and networking separately, buy them all from us in a single package. That's convergence. But within these packages, the different components were still separate. At my former company Nutanix, which focuses on primary storage, we collapsed the walls and said the same infrastructure can provide both compute and storage, plus have networking embedded. That's hyperconvergence. At Cohesity, we're going one step beyond. Nutanix only brought hardware together. We're bringing together software workflows in secondary storage to run on one infrastructure.

Q: Cohesity recently announced some new cloud offerings. What's the thinking behind these?

A: As I see it, the cloud is like renting a hotel room. And your data center, which you might have on-premises, is like owning a house. If you rent the room, everything might be done for you but things can get expensive over time. Some workflows will migrate naturally to the cloud. With Cohesity, we want to provide the benefit of a hybrid cloud that lets customers choose where they want data to be stored. For instance, it's a benefit to have backups living on-premises because you can recover quickly should you lose anything. But for long-term compliance, you can rent capacity in the cloud for data you won't use very often.

So keeping all this in mind, we recently announced several new features. CloudArchive lets you

take data images on the Cohesity appliance and transfer them seamlessly to cloud services such as Amazon Glacier and Google Nearline. Another feature, CloudTier, allows the cloud to be an extension of Cohesity. When data is hot, it sits on Cohesity's SSDs. When it gets a little bit colder, based on policies our customers define, it waterfalls down to hard disks. And when it gets colder still, we waterfall it down to the cloud. The third feature we announced is CloudReplicate. This allows customers to run Cohesity as a virtual appliance in the cloud and fail over to it if needed.

Q: You've just announced DataPlatform 3.0 and DataProtect 3.0, which are major releases addressing significant new workloads. Why these workloads and why now?

A: Our mission is to make secondary storage simpler and more efficient. The new releases (http://www.cohesity.com/press/cohesity-expands-data-protection-physical-servers-doubles-performance) are a significant leap forward in achieving that goal. Our large enterprise clients, such as Ultimate Software and Genex Services, are eager to see us expand our data protection coverage to include other critical storage workloads.

DataPlatform 3.0 addresses improved VMware backup performance. We've doubled the IOPS and throughput performance for file access, and added a new proactive monitoring service that analyzes DataPlatform installations and recommends preventative maintenance. DataProtect 3.0 extends our data protection coverage. Beyond our ability to backup virtual server environments, this release will enable enterprises to protect physical Windows and Linux servers too. With this comprehensive offering, our customers are getting faster backup and recovery times.

Q: The question of data security is on many executives' minds. How is Cohesity addressing this issue?

A: We are very software-defined in our approach. We let our customers select what needs to be encrypted in place. We can also interface with a key-management server, which means that all the keys aren't stored in a Cohesity appliance. If someone does break in, they can get access to the data, but not to the keys to decrypt it. We have role-based access controls too. I like to say that security is always a work in progress, but I'm proud of all of the security features we've already built into Cohesity.

Q: What's been happening on the customer front? And who are your main competitors?

A: Bigger companies all see the mess that exists in primary and secondary storage very clearly. So our customers tend to come from the enterprise category and the higher end of the mid-market—basically the Fortune 2000. They have all that fragmentation, all that duplicate data, and all that dark data. We've been GA since October 2015 and we already have some marquee names as clients. They include Tribune Media, GS1, Credit Acceptance, and Cvent, as well as other organizations such as the Annenberg School at the University of Pennsylvania. One big pharma company with petabytes and petabytes of data is planning to replace its existing setup with Cohesity, which will enable it to cut its spend on secondary storage by 85%.

In terms of competition, there are other companies that are innovating in this space. Actifio is one and Rubrik is another. They've recognized some of these problems and the need to integrate some of these silos. Rubrik has integrated the world of backup software with backup storage. Actifio recognized the fact that test/dev lives separately from backups, and set out to integrate them. We're saying you've got to consolidate even beyond this. You've got to bring analytics into the mix. You've got to bring in file services. So we're going above and beyond these other initiatives.

Q: You were one of the co-founders of Nutanix, which is lining itself up for an eventual public offering. Why did you leave?

A: As I said earlier, primary storage, which is what Nutanix focuses on, is the tip of an iceberg. By the time I left, Nutanix's technology had matured and my team there is capable of taking it forward and continuing to disrupt primary storage. But there was that unaddressed and very big part of the iceberg that was sitting there waiting for someone to disrupt it. That attracted me because I could see that while Nutanix was doing the right thing in hyperconvergence in the primary storage market, we still had a mess in secondary storage. So I came out to jump on that opportunity. Now between the two companies we address the whole iceberg.

Q: You were also one of the developers of Google File System, and you recently wrote a blog post about the early competition between Google and Yahoo for dominance of the web. What was the main lesson in it?

A: Systems are a complex beast. Once you build them it's very hard to change them fundamentally. And so systems should be built by thinking very carefully up front about the demands that will be placed upon them. If you only think about one requirement, you'll likely build a system that caters

to that. But then more requirements come along and you start doing patchwork after patchwork after patchwork, and you're never going to catch up.

Google recognized early on that it had a big vision. It wasn't just going to do search. It was eventually going to do email, put videos online, and do a host of other stuff. So it was already thinking about a platform that caters to all these requirements. Yahoo, on the other hand, made a quick fix. It said for the short term, NetApps can do the job for us, and then it started putting a layer of software on top and putting stuff on that. As more requirements appeared, the patchworking began. Google's platform could easily handle expansion beyond search; Yahoo's couldn't. As a result, Yahoo started falling behind in the technical race and the results are there for everyone to see.

I adopted the same approach we employed at Google when I was at Nutanix, and now I'm doing the same thing again at Cohesity. Here, I looked at the lower part of the storage iceberg—at backups, analytics, file shares, and test/dev—and then thought very carefully about the kind of platform that would be needed to consolidate all of this. Once you have the ideal platform in mind, then you think about what existing technology pieces you can leverage. You don't start with the existing pieces and then build a patchwork on top. That's the Yahoo way. The Google way is to first think hard about the ideal platform and then ask: 'OK, what can we leverage from the existing world that fits that platform?' We follow the Google approach, and that's why Nutanix and Cohesity are both successful companies winning rapid customer adoption.

---

http://wing.vc/blog/a-q-a-with-mohit-aron-ceo-of-cohesity
© 1990 — 2018 Wing, Menlo Park, CA