

Q&A: Esteban Rubens on machine learning, reliability and the growing importance of flash storage in medical imaging

By [Michael Walter](#)

healthimaging.com



Esteban Rubens, Global Enterprise Imaging Principal, Pure Storage

Pure Storage is a data storage company based out of Mountain View, Calif., that specializes in cloud-based, analytics-focused solutions such as FlashBlade, which offers companies petabytes of capacity with no caching or tiering.

Esteban Rubens, Pure Storage's Global Enterprise Imaging Principal, sat down with me to discuss flash storage in medical imaging and some of the fastest growing trends affecting health IT today.

Why is flash storage so important in enterprise imaging right now?

Esteban Rubens: If you look around, flash storage is pretty much everywhere now; it's in portable devices, laptops, and so on. Flash is where storage is going, so that's where the research and innovation are happening. Hard drives were great in the 90s and in the 2000s but right now, they have plateaued. They're going up against the laws of physics. Magnetic hard drives are getting to the point where the bit density is getting too high for what the underlying magnetic media can support, so OEMs are talking about things like laser-heated heads. In terms of bit density, reliability and performance, the obvious step for storage is all flash.

Performance is key in medical imaging, and you get that from flash storage. There are radiologists reading these much larger studies now—I've seen them do 20,000- or 30,000-slice CT and MR exams in academic environments, which are just massive amounts of data—so when they are flying through those stacks and they don't have the performance they want, then they can't do their job properly. This leads to the radiologists getting upset with their IT department, and it's just a really bad situation. Ultimately, what everyone wants is smooth, consistent performance with low latency, and that's exactly what flash storage can deliver.

What are some other benefits for an imaging group that chooses flash over other forms of storage?

Another key benefit is reliability. Hard drives are not reliable, and that's where the concept of a redundant array of independent disks

(RAID) came from originally. The whole point of RAID was, hard drives fail. But we are now at the point where you have radiology and cardiology departments where organizations have hundreds of terabytes of images or even petabytes. This requires hundreds of hard drives, sometimes thousands, and they're always breaking. With flash storage, that doesn't happen. Flash is inherently more reliable; you have no moving parts.

Flash storage also uses much less power and it's much more compact. All of these things play into what people are worried about.

What changes do you see in medical imaging that could potentially impact healthcare organizations' storage plans for the next 4-6 years?

The biggest change I've seen is an increase in concurrency. When PACS first came out, it was kind of mimicking the analog world. You had light boxes for few radiologists, you had an archive. That has increased, of course, as people have wanted to get more from their investment. They want more people to access images, they want a more complete medical record, which means that everyone should be able to look at images. And anytime someone uses an image-enabled electronic medical record (EMR), images are being accessed which puts stress on the storage infrastructure. If you have thousands of users with access to patient images, you have to make sure radiologists and cardiologists get the performance they are used to. You can't say, "Hey, the radiologist has to wait when they want to open an image, because the system is being hit by a lot of people." And this is huge—concurrency is only going to increase, because you're image-enabling the EMR and you have new modalities such as digital breast tomosynthesis (DBT), which take up a lot of storage space.

I've spoken to several radiologists who have mentioned the challenges related to DBT exam file sizes. Were IT departments not prepared for these larger data sets?

When DBT first came out, it was treated as a complement to mammography by a lot of practices. But in my experience, everyone is starting to go all in with DBT and replacing old equipment as soon as they can so they can perform it. We talk to customers who talk about how 30,000 mammography exams became 30,000 DBT exams, which requires up to a hundred times more storage. And this has had some

practices totally freaked out, because they didn't think everything would move to DBT. Small facilities that usually buy 10 or 20 terabytes of storage suddenly need hundreds of terabytes, so now they have to consider going to the cloud. And even though Pure Storage is not a cloud provider, our technology, at its core, is cloud-based in terms of architecture and protocols, so we can provide that seamless transition to cloud-based storage.

There has been a lot of discussion, and even a little concern, about the impact machine learning will have on medical imaging. What impact will machine learning have on storage?

The old paradigm in storage service is that if something breaks, a ticket get opened, there's a conversation that goes on between the vendor and the customer, and then it can finally be fixed. This all takes time, and nobody likes that. At Pure Storage, we're always concerned with everyone involved in the storage management chain, from the end user to the administrator to the individuals actually purchasing the storage. So the main thing we wanted to do, in regards to machine learning, was seeing how we could make the support experience better.

We have every one of our storage devices acting as sensors in an Internet-of-Things (IoT) sense reporting any potential problems with data to us through our data lake. (And we're talking about machine data related to the status of connections, internal components, networking, that kind of thing ... not medical data.) We are alerted to potential issues in advance, and then our artificial intelligence algorithms identify other potential matches for that problem. So when we see something, we propose changes to customers that they could make to avoid the problems from ever actually happening. "Please upgrade this driver," for instance. And that's huge for these facilities, because we've proactively addressed a problem before it actually happens.

By our estimation, we've avoided hundreds of potentially bad issues, which means people are less stressed and can spend more time with their families. I've heard from users who have been on vacation and were able to proactively fix issues remotely; it saved their vacation. And as anyone in IT knows, when you're on vacation and you're responsible for IT issues, you normally have to go home when there's a problem. You can't not respond.

Healthcare organizations often find themselves being asked to do more with less. How does that impact data storage for imaging?

With our legacy competitors, you buy storage solutions à la carte and you sometimes end up buying the same thing twice. You might

buy one thing, but a few years later, you end up being asked to buy something else. With Pure Storage, you are guaranteed to always be running on the latest and greatest that we have. We don't think it's reasonable to ask people to buy the same thing again and again, and we guarantee that won't happen.

And when our customers do need an upgrade, it's not disruptive. You can upgrade on a Monday at 8 a.m. It's not, "Hey, we are going to upgrade, but we'll do it on a Friday at 11 p.m. just in case."

When you buy a Pure Storage product, you are buying a subscription to our innovation. When we innovate, every customer is going to get it. Because they are a customer, they get the best we have. And as flash storage gets faster and denser, they can ride that curve. Our technology is going to keep getting better and better, and customers will benefit from that.

Historically, there has been a transition in IT departments from specialized positions to employees who have more general job descriptions. How does Pure Storage play into this shift as lines are being blurred within organizations?

There are a lot of complexities behind the scenes with our product, but you don't need hours and hours of training and you don't need to be "a storage guy" to run them. In fact, we don't even have manuals for our products. We have, literally, a business card with all of the information you need to know. So what that means is, organizations don't need a ton of employees who are just doing storage. They can redeploy those people and have them be more productive as IT generalists.