VIDEO BLOG EPISODE 1

FEATURING

JASON RABBETTS
FOUNDER
UNION SOLUTIONS

CREATING AN EFFECTIVE CLOUD STRATEGY

[00:00:07] Geoff Bourgeois: Hi, Jason. Thanks for joining me today to talk in this first edition of the *HubStor Video Blog*. My understanding is you are one of the co-founders of Union Solutions. Can you start by just giving us an introduction to yourself and telling us a little bit about Union Solutions?

[00:00:22] Jason Rabbetts: Yes, sure. Thanks for inviting me along. We are an infrastructure services company. We build transformation programs for clients to exploit cloud models, predominantly hybrid cloud models and we integrate hyper-converged infrastructure with software-defined capability and extend people's data center services into the public cloud platform.

Hopefully, what we try and deliver are outcomes based around improved service levels, improved agility, and reduced cost and risk. Typically, customers we deal with are large corporates, large enterprises, anything between 2,000 and 10,000 seats. We have some large customers above that and that's what we've been doing for 20 years now.

So, we've come through the open systems and consolidation phases of SAN and centralized backup through the virtualization age and now into, obviously, the cloud-enabled era.

[00:01:30] Geoff: That's fantastic. So, you have a front row seat to a lot of the sort of inner workings of the data center and the challenges that companies are faced with when they're trying to wrestle with the cloud and figure out how to use it.

What are some of the easiest, quickest wins that you're seeing with customers today when they're looking at cloud adoption? They've already made this massive investment on-prem. Where are you seeing them start with cloud?

[00:01:58] Jason: Well, I think most-- if we talk about our main customers, which are enterprise corporates, I think the model most of those guys are adopting is a hybrid model. They recognize they can't move everything to the cloud; certainly not overnight. So, they appreciate, it's going to be a number of years before they get to maybe their optimum point.

Most of the customers we're speaking to are addressing that really with a initiative whereby they're doing what we call skinnying [sic] down their DCs. So, they're trying to get as much density into the existing DCs they can to accommodate whatever workloads they want to retain on-prem. But, at the same time, moving what workloads they can that economically makes sense into the public cloud, but doing it in such a way where they are essentially, virtually, extending their own data center perimeter.

So, they're still trying to do it in a way that retains the enterprise-class service levels they need to deliver. I think that's the biggest challenge that we're helping customers through is most businesses demanding of their IT department that they provide them with the agility, the cloud economics, the scale, and the self-service provisioning type of benefits that cloud delivers and most are, therefore, dictating a cloud-first policy.

But most of our customers, whilst they have a cloud-first policy, don't necessarily have a cloud-first strategy. I think it's linking those two things and filling that gap that's probably the biggest value we deliver.

A good example is businesses are demanding that their IT departments use the cloud, but what they're really saying is, "We want the economics, the scale, the agility. But at the same time, we want the enterprise-class service in terms of performance, availability, recoverability, and availability that you build in a classic, well-tuned, well-optimized, high-class enterprise service stack."

So, it's getting people to understand that, of course, cloud economics and cloud models are possible. It isn't necessarily just a case of moving to the public cloud. Hybrid's probably the most likely three-year journey they're going to make in the enterprise corporate space.

Really, benchmarking, what it is they deliver now in terms of a service, how much it costs them, what the risk profile of that service is, and whether it is economically better to move it or not? Because that's the other thing that customers often discover, particularly in the enterprise space, is they do have economies of scale themselves. There are certain things that it is as economic to run in a private cloud model as it is to move to a public cloud.

So, yes, benchmarking; what service you're delivering today, what gaps exist, what does it cost you, and then delivering the best model to improve those things.

[00:05:06] Geoff: Yes. I think that's an interesting position; it's one I very much agree with. It resonates with me that hybrid cloud, for most enterprises, is going to be what's right. These are two vastly different methods of delivery. The economics are vastly different; it's an apples and oranges sort of comparison.

So, you kind of have to take a step back and go, "Okay. What should go where? What workload should go where?" And try and achieve, get the best of both worlds. I've seen personally some organizations rush into the cloud and try and force fit everything. Inevitably, they recoil from cloud and then they do this sort of overreaction where they think, "Well, cloud is bad because we're paying more. We were starting to run such and such an application and the activity costs with the cloud were really high and it was more expensive to run things in the cloud."

Well, yes, again, you have to understand what the cloud is good for and what's it's not. You have these physical issues with your network connection, and then you have to understand the pricing model of the cloud and that there are these things around activity costs and whatnot.

So, particularly if you're running active data workloads, very compute-intensive, high object counts, and a lot of heavy transactions, heavy processing, like that, and then you want to pull all that data down and do stuff with it on-prem and then interact with it in the cloud. That is a higher cost profile than say using the cloud for archive backup and DR, where it's lower touch data and the economics, are probably going to make a lot more sense for that older data than they are going to be on-prem where--

We were talking with Gartner the other day and we were saying, "If you think about retention requirements, long-term retention and keeping that data on-prem, if that's done on primary storage, what are the ripple effects with backup?" And thinking about how many hardware refreshes are you going to go through for the whole lifespan of that data? You amortize the costing on that and it's ugly.

So, that's where-- I think there's a lot of-- when you look at the data management strategy, obviously, there's the compute and things like this, but just the storage strategy between on-prem and cloud, I think is where you can probably get your lowest hanging fruit opportunities to go into the cloud. Any thoughts particularly around DR strategies with the cloud and backup?

[00:07:46] Jason: Yes. They're two immediate areas that are clearly good use cases. What we're finding is certainly from a DR perspective, the economies that the cloud can deliver do exist. They're real. Particularly, again, in the enterprise-based, the typical route or the model that people have adopted, these two data centers, which means for every service they're paying twice as an insurance premium.

I think now that people are moving more to a service-oriented type model where they're trying to develop their own internal IT processes and services in the shape of a service provider, then they are examining really what service they should be providing to each application and are starting to grade those and say, "Well, that

application perhaps doesn't need replication, perhaps doesn't need this RPO or RTO, perhaps doesn't need this gold/platinum service of DR."

So, I think there's a couple of things going on. One; that people are starting to class the application and the service in a way that makes economic sense and meets really what the business is asking for and not assuming everything has to fly first-class. When they're doing that, it enables them to look at the economies of how best to deliver those services. Do they really need a second data center with everything replicated? More often than not, the answer is no, they don't.

So, I see an awful lot of our customers moving disaster recovery services into the cloud and making that part of that hybrid model. Backup to an extent; most of our customers still have a pretty acute and intensive RTO requirement and a pretty acute RPO demand.

So, there's still an awful lot of backup taking place on-prem. But, when it comes to the longer RPO and RTO service classes, definitely, again, cloud is a good place to move those. So, certainly for long0term retention. Yes.

So, yes. Backup and DR are definitely models and areas that are starting to be significantly transformed by the hybrid model. I think also unstructured data stores. I think they're a massive problem for people. They've largely gone on service door, not very well looked after. They're highly distributed, lots of content, very difficult to discover.

GDRP is demanding and a lot more in terms of people's understanding where their data is and what it is and who owns it and what it contains. I think in that process, they're also discovering what a massive proportion of their data is old, is cold, and has a mid-reference for many years in some cases.

I think this is largely symptomatic of the fact that IT, up until now, has been the scapegoat for a compliance. Whether that's statutory or regulatory, IT has been asked to deal with it rather than the business engaging properly. Therefore, IT has kept everything forever.

So, I think, certainly, what we're seeing is object storage is becoming a much more interesting subject matter for customers because it provides features and functions that could meet some of the compliance requirements and also some of the data protection requirements.

Also, I think what we're also discovering is that through that process, when they realize how much data is actually stagnant or old, what that cost actually represents to them because you then being that's DR, you link it back up, you leave it all to the copy data issues that come with it, then it's driving a lot of change in the unstructured world.

Certainly, we're starting to deploy, again, hybrid models for unstructured data, which have the knock on a-- impact onto DR and backup over the positive impact.

[00:11:56] Geoff: Yes. Those are great points that you raised there. I mean, so many different angles we can take from what you just said getting into the GDPR and the implications there, cloud backup, as well as the whole issue of IT companies keeping everything forever, how we got to this stage, and IT really bearing the burden of data management when I thought we have information management, record management guys. Where are they in this picture?

[00:12:29] Jason: Well, I think that's been a longstanding issue. I mean it's been something we've been saying for the best part of maybe 20 years, certainly very actively in the last 10 years, is a lot of the cost instituted within data center infrastructure stacks is data storage, whether that be primary or copy data.

Being a specialist in data services, particularly storage, backup, and DR, people have always assumed we're delighted with the fact that people keep all this data for so long and they're constantly maybe ordering upgrades to capacity.

That's never been the case with us. We've had what we call our smart storage philosophy, which is understanding how your data is used, who uses it, what makes up hot and cold data, what should be a primary, what should be on maybe another cost secondary tier, and making major in-loads there.

The problem we've had **[unintelligible 00:13:25]** at evangelizing and preaching this. In practice, the issue is within the customer's business themselves. The businesses will not and have not engaged and I think it's taken something like GDPR to force that agenda because technology can only do certain amount of things and that technology can only be applied to a well baked out and prescriptive set of policies and process.

So, unless the business itself recognizes what its compliance requirements are whether regulatory or whether statutory or even whether just within their own business what they would like to dictate in terms of retention policies and lifecycle management of data.

Until the business is engaged and prescribe what their policies are, which would in turn change process in how the data is stored and where it's stored, then IT will always struggle to then support that policy in process with technology and you'll always be stuck in this place with the story and everything forever and probably as a safe running first class because they don't know what it contains or who owns it and what they should do with it.

So, I think GDPR as a piece of statute coming in certainly within Europe is forcing that conversation and that conversation is leading to more innovative technology solutions to support a business that's a last engaged in a policy conversation around data lifecycle.

[00:15:16] Geoff: Yes. That makes a ton of sense. I think for those not familiar with the GDPR, that's the General Data Protection Regulation and I think specifically, the part that's going to be a real catalyst for change in data management strategies and underlying technology infrastructure is this whole right to be forgotten, that you're

going to get a phone call from somebody that says, "I want all my data that you have, it's got to be expunged."

I think what's concerning with GDPR is, of course, the proportionality issue that exists where the penalty is for non-compliance are significant. I mean I think the **[unintelligible 00:15:55]** of revenue--

[00:15:57] Jason: No. Yes, it can be up to 4% of global revenue. So, it's not 4% of the revenue of the business territory within perhaps the compliance breach has been recognized, it's 4% of global revenue.

Now, if you take that for a large European organization, which is going to be dictated by these laws, then that's massive.

[00:16:26] Geoff: It is. Now, I think where this-- the whole issue of data management strategy, I like how you talk about flying everything first class, all these data sitting on primary.

I was talking with George Crump at Storage Switzerland a month or so ago and he said, "Archiving is really the most undone thing. We talk about all the time but it's very much undone in the enterprise."

[00:16:53] Jason: Absolutely.

[00:16:54] Geoff: I think when you look at a lot of the secondary storage options that are out here, they're-- I kind of refer to them as dumb storage. I don't mean to be too harsh, but there's not a lot you can do with it, it's like, "Let's go put our data offline to go die and it's dark data." We really don't have great hooks to manage it.

Certainly, when you think of putting things off into tape libraries and shipping them off to Iron Mountain or something like that, you really start to see a picture of what will-- how do I do discovery on this if I do have a GDPR request?

I've had a front-row seat at previous companies where we did discovery work with a lot of our clients, and they would get hit with investigations or litigation, "Here's a shipment of 17,000 backups, tape backup," and the data restoration team is going to basically restore all of that and get it into a searchable format on disk. But, I mean, we're talking about millions of dollars to do that.

That's just not practical with GDPR. We are talking about potentially a couple times each day where you're going to get these requests to be forgotten and you need to go and be able to query your storage.

So, that whole issue that you're talking about having a smart data management strategy, a smart storage strategy, I think really comes into play now with things like GDPR. I think it puts a lot of pressure on the way things have been done. Everything is just accumulating on primary; it's not indexed. We don't have any object storage capabilities around it. The secondary tier is dumb. Tape is arduous, cumbersome.

You're going to get in a lot of trouble if you're sticking with that kind of model of things with GDPR.

So, I think what you're talking about is very relevant. The question, of course, is well, how? I think people want to know what kind of technology we should be looking at. You mentioned object storage. We're talking about cloud. So, if you are concerned with this, you want to look forward and move in that right direction. What kind of technologies do you recommend customers to look at?

[00:19:08] Jason: I think there's two issues that need to be addressed that then open up the conversation about more appropriate technology. I think the first one we've already dealt with which is, first of all, the businesses have to engage with IT and set out what their policies are going to be. They have to understand what their regulatory/statutory demands are and then decide whether they're going to keep things for X long, Y long, what data is important to them, what's not. That's point one.

I think the other point is symptomatic of the fact they haven't engaged up to now, which is most people's archiving strategy is actually a backup strategy, which means predominantly is offline. There's an awful lot of data stored and it's very difficult to search it, all of those issues.

So, I think until people separate their thought process. When I talk to people and say, "Have you got an archiving strategy?" "Yes. Yes, we back it up every day and we keep it for seven years." That's not an archiving strategy, that's a backup strategy, and a **[unintelligible 00:20:06]** for one.

The chances are of you needing to recover anything beyond maybe 30 days is limited and the argument always comes back, "Yes, but what if my compliance officer or my data protection officer or my CEO needs data that was there a year ago?" Well, that 99% of the time is a recovery process you're on staking for some statutory or regulatory compliance need and it's different.

So, I think until people separate back up process to archiving process and get that in their head and then join up the business policy, then they're always going to be running into a dead end.

Now, fortunately, people are starting to think differently now and typically, the technologies we're recommending are obvious storage-based. So, the ability to have metadata reply to any individual unstructured data object, which gives you the benefit of not having being able to apply lifecycle policies to it and data protection policies, we're also able to provide quite a sophisticated degree of context searching as well.

In real simple terms, the future, as we see it, is hybrid cloud within the data center that's hyper-converged infrastructure with some form of software-defined capability where your primary data has a high degree of latency and IO demand where we run in all flash, maybe even non-volatile RAM as we move into newer technologies in that area, but all of your unstructured data. Last year **[unintelligible 00:21:39]** the data is going to move to an object storage.

I think that's how people are going to see things. You have block storage for structured, which are largely all flash and your unstructured is going to move into an object store of some form, which gives you that compliance, gives the data protection, lifecycle management, that up until now, traditional file systems have not been able to do.

It's going to drive economists as well because inherently, the object store platform is software-defined and it is scaled out so it's cloud-enabled as well. This comes back to the very original point or discussion topic we started with, with respect to where a customer is going.

Anyone that thinks cloud is in the future is mad but when everyone says all of our IT must be cloud-based, you really have to ask yourself what they're asking. What they're really asking for is cloud-type characteristics, highly scalable, Opex-centric rather than CapEx intensive, highly agile, flexible, and delivering, as I say, cloud-style economics and agility that doesn't necessarily has to dictate where that technology sits.

[00:22:55] Geoff: Yes. Those are all great points. I was having a conversation with Garth Landers at Gartner just yesterday and we were wrestling with this whole issue of archive versus backup and cloud strategies and things like this, object storage where does it fit in.

He had great insight. He said, "I think, Geoff, we're going to stop calling it archiving." Really, what we're talking about as a secondary storage layer. A secondary storage layer needs to be really smart.

Taking things back to the issue of backup versus archive, I've seen organizations take their legacy backup strategy as is to the cloud. Let's say they cycle backups every two weeks or whatever. They just take that as is and do it in the cloud. So, the cloud is now where they're storing their backups. You look at the cost model on that and it gets pretty ugly. Right?

[00:23:52] Jason: Yes.

[00:23:53] Geoff: I think when you shift your thinking a little bit and say, "Okay. Well, we need to take a step back, revisit this whole issue of archiving versus backup, see it more leveraging the cloud as the secondary storage layer." Then you can start to think about-- okay. There's convergence of archiving and backup. When the cloud is my secondary storage layer, it becomes this forever incremental. Right?

[00:24:22] Jason: I completely agree, Geoff. I think it comes back to some of the couple of the key points. One is lifecycle management. One is smart storage. I think ultimately, backup is not a technology or a concept or process as we've traditionally known it that has a future. I don't think Storage Area Networks, as we've traditionally known it in hybrid arrays and tiered architectures, as we've traditionally known it, have a future.

I think businesses in working with IT will start to develop lifecycle policies for datatypes. They will apply that at birth. That data lifecycle policy will be automatically provisioned and executed by smart storage technologies that will automatically decide where it goes, what's the most cost-effective way of doing it.

We will move away from a backup process or an archiving process or a primary commissioning process; it will all be automated, it will all be software-defined, and it will be part of a data lifecycle policy that the customer, the consumer sets and the smart storage technologies will deal with that process.

Ultimately, I think we may be a little way from there but it's not beyond the realm of possibility within the next five or 10 years that these technologies will be smart enough to exploit the developments in the cloud marketplace and will be able to automatically determine based upon the policies set by the business where it is allowed to store it in the cloud.

Not only that, but today and for the next six months, there's an offer on with AWS that's cheaper for me to shift the next set of data to AWS rather than having it sitting in, let's say, Azure, as I've been using for the last year.

So, I think that'll become smarter to us. Now, a lot of things have to change; there has to be a marketplace; there has to be a broker system; there has to be a multi-cloud compatibility within the whole hybrid model that works data egress between platforms is going to become much more consumer economic than it is today, but that will happen.

So, you're absolutely right. Backup and archiving are separate things. There are processes that I think will just become automated and dictated by smart storage technologies that adopt whatever lifecycle policy organizations apply.

[00:27:17] Geoff: Yes. Those are all great points and I think it is very exciting when you get into a bigger picture of what's going on in the cloud. So, you imagine, "Okay. We're going to leverage this; it's our secondary tier; it's supposed to be smart storage. I don't want to get locked in. I need to be the master of the data."

Those are big fundamental shifts from where enterprise technology is today. Archiving products base their business on locking customers in, for the most part. You're told to be sticky and that's how they're sticky. They make it very difficult to get out. And I don't think that jives with the cloud.

I think with cloud, people have an emotional concern going in that they don't want to relive that. Typically, going to the cloud involves getting out of something. Going into the cloud, they're experiencing how painful it can be to get out of something and I think that's top of mind going into the cloud it's like, "Well, I want to know what my exit plan is in case we need to do that."

So, I think that's obviously top of mind but all the things, other things going on in the cloud, there are going to be these services that pop up out of nowhere. We're starting

to see it where "Hey, I want to do OCR indexing." TIFFs and PDFs. I do call center recordings and I want transcripts of all the audio.

So, it can't just be the sort of closed system that you're storing your data in. There might be other services that pop up in the cloud that you easily want to consume and run against targeted subsets of your secondary tier. I think whatever platform you're working with, you need to know that there is a strategy that you can get the data out, move it, but also tie in other services, do other things with that data because right now, it might seem like it's cold data but overnight, who knows? The business might find new ways to use it.

[00:29:14] Jason: Absolutely, there may be some intellectual property value to it and there's certainly be some marketing and sales information value to it. You're absolutely right; it might be cold today but it might not be tomorrow.

I think this is why object storage is such an appealing platform because it's actually a very mature one. The cloud is quite new, but object storage has actually been around as concepts and technology for either over 10 years now.

Standards have been built, the S3 standards as an example, which is portable, it's reliable. You can program your applications for its API. So, you've got a level or a degree of comfort that your data is portable or you can write your applications to a server APIs that can exploit a number of different providers of object storage technology; it's already scaled out; it's already software-defined; all these good things.

So, if you look at any growth curve in terms of consumption or revenue or whatever object storage platforms, you're probably thinking you're going to be one of the highest growth sectors within this cloud-enabled infrastructure world.

[00:30:22] Geoff: Yes. A quick question on object storage. I mean do you see people ultimately getting into the world of object storage or do they really want the value of object storage?

In other words, are we going to see systems coming forward that are fundamentally object storage but you never hear that being discussed? You don't hear things like erasure coding being discussed and scale out and stuff like that. Those are just assumed to be there. There's sort of a commoditization of object storage, if you will.

[00:30:56] Jason: Yes. I think you're right. They're just expected features and functions of object storage and they're the reason people want to move to an object storage platform.

I think what's quite interesting is we're seeing a lot of interest from our customers in object storage to fix an unstructured data problem, almost swap from a traditional NAS or a file system, a NFS file system to a more scalable intelligent platform.

That seems largely to be their motivator and driver, but actually, when they get an object store and their application developers start to understand a bit about what can

be done with it, then we're starting to see a secondary engagement whereby they're starting to develop their applications to integrate in a native format into those platforms.

That's where they're starting to see real business advantage and giving them a genuine competitive advantage in their market that can accelerate revenue as they bring new, dynamic, and exciting features to whatever it might be; their point of sale systems, their apps, whether that be mobile apps, consumer, and whatever it might be.

So, object storage isn't just about delivering a highly efficient way of storing unstructured data and being able to tick a number of articles in the GDPR self-regulations; it actually becomes an enormously enabling platform for application development and therefore, competitive edge.

[00:32:40] Geoff: Yes. Those are all great points. Well, this has been a great conversation. I've really enjoyed it. We've covered a lot of ground. Hopefully, it's been informative for the folks out there.

Just in closing, Jason. Are there any final things that you'd like to say in terms of recommendations to your customers or any IT person that's listening when you're considering going to the cloud, how they might go about evaluating it, any sort of things that they should be considerate or careful of?

[00:33:10] Jason: I think two things really. If you're considering cloud, be clear about what it is you're trying to achieve; it comes down fundamentally to what is it about the cloud that you like. Usually, it's cloud economics, cloud scale, cloud agility.

Once you're clear on those, then determine the best model for delivering that for your business because every business is going to be unique. The only way you can do that is benchmark what we do now. What do you do now? What is the service you deliver? How much does it cost? What does the business actually want that service to deliver and where are the gaps? Can it be done more cost-effectively in a different way?

That different way maybe on-prem, it might be off-prem. We have to benchmark it first. Don't just leap from one place to another because you might end up [laughs] a painful outcome; first, in terms of cost, service, and functionality.

[00:34:07] Geoff: Well, great. Thank you very much for your time.

[00:34:09] Jason: No probs. Cheers, Geoff.

[00:34:09] [END OF TRANSCRIPT]

For more information, please visit:

www.hubstor.net

www.unionsolutions.co.uk