

Has the time come to "Cluster Clouds" for even greater levels of Availability?

Introduction:

Pyalla Technologies, LLC., was founded by Margo Holen and Richard Buckle amidst the turbulence of the 2008 – 2009 Global Financial Crises. The company's focus has always been on NonStop and through the posts and commentaries it has brought into focus many of the issues of the day. To suggest Margo and Richard have seen it all and been participating in almost everything that has happened would do them a disservice. For much of their business life, Margo and Richard have not just been there as members of the NonStop community but have consistently volunteered in support of the NonStop community.

The HPE NonStop system of today has journeyed far from when it was first rolled onto the data center dock at Citibank in May, 1976. Even as NonStop will shortly celebrate its fiftieth birthday that journey has seen a transformation that many in the IT industry failed to predict. Breaking free from its propriety past to first embrace blade infrastructures to where today it lives on the Intel x86 architecture, NonStop continues to demonstrate its relevance as an open and modern system. However this was only the beginning of the journey as NonStop has foregone running on traditional systems with its move onto virtual machines. Nowhere has this generated more excitement than in those enterprise data centers where today NonStop has completed its transformation and become a pure software play.

And the journey continues unabated. With the option to run on virtual machines, NonStop as software has the ability to run within clouds bringing the cloud experience to NonStop even as HPE can count on NonStop participation within the edge to cloud platform as a service. However, there is the possibility of yet one more transformation that will involve the NonStop community. The very foundation of NonStop as it made unreliable hardware fault tolerant and brought robust configurations across virtual machines possible lies in it's potential to bring a degree of robustness across different cloud offerings clustering clouds as NonStop has always done with processes, systems and machines.

For NTI the proposition that more is yet to come as NonStop continues its journey hasn't gone unnoticed. Already the integration of NonStop with external apps running in clouds is ongoing. However, what is recognized by NTI is how a fault tolerant NonStop system can be adversely affected should apps running in clouds become unavailable to NonStop. Just as networking has always been part of the NonStop proposition it is no exaggeration to view the clouds of today becoming part of the bigger NonStop. This white paper explores a number of options even as the approach it endorses is one of incremental change through following baby-steps, it's clear to NTI that the cloud can no longer be considered a resource apart from NonStop but rather as an integral part that makes support of mission critical applications possible.



Tim DunneGlobal Director Worldwide Sales
NTI









As I look back at the past year perhaps the best way to describe it was how there were many times where things were broken. I am not talking about the occasional drinking glass or dinner plate but objects that mattered to us. When things are broken their performance is hindered to where in some cases they no longer serve any useful purpose. Replacement appears to be the order of the day.

Mind you, I am not talking about the weather here in Colorado as I have lost track of how many days have passed without a meaningful snow fall. This morning may have been the exception as a little rain together with what might pass as a little sleet did fall but not for long enough to stick to anything. Enjoying December days that have climbed into the 70sF (20sC) would surely pass as unusual even if we didn't talk about weather patterns that look to be broken.

However, before developing this story line further, please be happy for us as remedies for almost everything have been found. But what was broken? This time last year Margo broke her leg badly where the remedy happened to include the insertion of metal rods and nails. In summer our Range Rover was rear-ended on the freeway and the insurance company wrote it off. That new sectional we had waited for did finally show up but the central portion of the sectional had been scratched in transport.

When it comes to IT and to the digital transformation and the pivot to everything-as-a-service, it's hard to make light of the fact that the role clouds are playing isn't proving as rock-solid as promoters would have you believe. Not for them is an outage here or there something for us to worry about, but when a major cloud services provider like Amazon Web Services (AWS) breaks then yes, we should all be concerned. In fact there were enough outages for CRN to publish **The 10 Biggest Cloud Outages Of 2021 (So Far)**. As for the tag line, it was rather long but managed to sum up the predicament of many affected at the time:



Outages can mean the end for companies, depending on their choices in design and deployment, or they can be complete non-events,' Miles Ward, chief technology officer at Los Angeles-based Google partner SADA Systems, tells CRN. 'Cloud has changed the nature of outages.





But then, CRN highlights something that should warm the hearts of many in NonStop, particularly at this time of year. Consider it your early arrival of your Christmas gift:



Every cloud engineering team has seen how impossible it is for customers to engineer around these kinds of outages and is working hard to distribute, subdivide, and make fault-tolerant these central services, Ward said.

Given that this article by CRN was published back in late July so missed reporting on the big AWS outage it's worth noting that among the top three worst outages were:

In third place - **Fastly Outage in June**. "Fastly impacted bulletin board website Reddit, video streaming service Twitch and a number of news sites including CNN and The New York Times." Among the comments reported at the time by CRN was this particular gem:



Michael Goldstein, CEO of LAN Infotech, a Fort Lauderdale, Fla.-based solution provider, told CRN at the time that the global outage shows how critical it is for customers to properly architect their cloud and on-premises network.



Cloud isn't any different than on-premises-with both cloud and onpremises you need to make sure you have the right architecture,' Goldstein said. 'We make sure that when we put mission-critical applications in [Microsoft] Azure for our customers we have multiple data center regions to prevent an outage like this. You need a failsafe and a continuity plan to prevent outages.

Rising to second place and given the generalized heading of **More Microsoft Issues** this time it centered on issues to do with Microsoft Teams. Apparently, "Teams' calling service sent calls straight into some users' voicemails." Now depending on your level of tolerance of virtual meetings this may have been a blessing in disguise, but in reality, it really all came back to issues with the infrastructure, according to Microsoft via updates provided by the Microsoft 365 Status Twitter account:



Microsoft 'isolated a recent change that has caused portions of infrastructure to send some Microsoft Teams calls straight to voicemail

But then, one Microsoft partner, Amaxra, according to its president and CEO, Rosalyn Arntzen, told CRN that "over the past few years, Microsoft had gotten "dramatically better" at updating partners "as soon as they are aware of an issue and listing when they expect the issue to be solved—or at least provide a status."





Coming in with the blue-ribbon winning outage of the year (so far) was the **Akamai Outage, June 17. Remember this outage? Turns out it happened** "Nine days after the Fastly outage, (where) a system issue with Cambridge, Mass.-based Akamai Technologies caused internet outages for global airlines, banks, and stock exchanges. The company saw service disruptions for its hosting platform, which helps defend against Distributed Denial-of-Service (DDoS) attacks.

The way CRN reported this outage was to highlight that:



Downdetector.com showed spikes in complaints about service outages for websites of companies inside the U.S. as well as in a number of other countries including Australia, Germany and India

And remember among the also-runs was the outage at Verizon that reports blamed on a fiber cut in Brooklyn, but that was later confirmed as being "a software issue triggered during routine network management activities." And then there was the issue at Google when "The Google Drive cloud storage service—and associated cloud apps including Google Docs and Google Sheets—suffered multiple service issues ... While users could still access Google Drive, affected users could not create new documents and were 'seeing error messages, high latency, and/or other unexpected behavior,' according to the company."

And there you have it: The myth of the infallibility of clouds. Amazon, Microsoft and Google. Of course, it was left to Larry Ellison to capitalize on their circumstances by virtue of his claim that Oracle cloud didn't fail. Surely, you cannot be serious, Larry?







For all the upside associated with capitalizing on cloud services there is still the fundamental issue that resilience and indeed reliability of levels we associate with NonStop are simply mythical. Fail-safe continuity and indeed fault tolerance for "central services" is being openly discussed even as we know that with todays' modern languages tools and services there is a lot that can be done to deliver a kind of pseudo fault tolerance. To think that all those years ago, the original Tandem Computers understood the issues better than any other vendor.

And yet, when those cloud services' vendors, providing the underlying infrastructure and most important of all the networking and integration services get it so hopelessly wrong, how can users deploying mission critical applications know for certain that these services will always be there, 24 x 7? The reality is a lot more sobering; they cannot provide anything close to ironclad guarantees. There is a reason why NonStop continues to thrive four decades after being first introduced; it's fault tolerant in so many ways that it should be hard to ignore it's contribution to cloud computing.

I am not entering into this conversation lightly. However we aren't discussing how to best fix a broken toy of which there will be many reports over the holidays. Two opportunities come to mind that in the coming months I will be exploring in more detail. And they have to do with how we think about NonStop going forward and whether our own ideas about the role of NonStop may indeed be outdated.

There is the potential to have NonStop play a guardian role — no pun intended. Should there be a central NonStop essentially polling the hybrid multi-cloud environment common today among enterprises so that exposure to any one cloud can be marginalized to where outages have no impact on the running of mission critical applications? This is clearly an over simplification but there are models that feature NonStop in this way that readily come to mind.

There is also the potential for NonStop itself, virtualized as we now have the option to deploy NonStop, treating the world of hybrid clouds as no different to either converged NonStop processors or as virtual machines. Consider one cloud as being CPU0 and another cloud as CPU1, etc. and you get the idea. This too is clearly an over simplification that perhaps throws a spotlight on the capabilities of the cloud services providers interconnect with each other, but the idea is still simple in principle. A single image NonStop system spanning multiple clouds, with the ability to perform its industry-leading take-over whenever a cloud misbehaves?

Once we get past the idea that yes, like real CPUs and even Virtual Machines, clouds are just as unreliable then the future of NonStop will warm to the opportunity this represents. The mere fact that one publication is already producing an annual Top 10 Outages article should be evidence enough that enterprises need to more seriously consider what the cloud experience really entails?

For Margo and me, this is just the beginning of a theme that we will revisit in 2022, so stay tuned. But again, the items that broke for us in 2021 have all been addressed and having said that, can you all say the same about your own hybrid IT and its supporting infrastructure? Even as we wish you the very best for the coming year perhaps it is time to ponder that ultimate question about NonStop: When did availability ever not be the issue of the day?



"I've got a feeling and I just can't let it go!"

This is a line from a haunting song by Los Angeles indie electro-soul band, Caught a Ghost. Not to be confused with the song from Frozen, it became the theme song of Amazon Prime series "Bosch" that was based on numerous novels by Michael Connelly.

For those familiar with the novels or those who have viewed the Prime series, you will know that the main character Bosch enjoys jazz and the choice of this song for the series title may have been a surprise and yet, it absolutely captures the indomitable spirit of Detective Bosch.

This is the same feeling I get whenever I hear news about the latest cloud offering or about the enterprise that has elected to migrate everything to the cloud. And this feeling is amplified whenever I read of news about yet another outage affecting millions of users. It is almost as if IT professionals have forgotten the first law of IT: Change. Call it cause and effect. Call it the swinging pendulum. Maybe think of it in terms of Newton's third law: For every action, there is an equal and opposite reaction. It's not surprising then that there are those who are beginning to question the validity of migration to clouds.

No matter what you might think of enterprises embracing clouds that IT constant of change suggests that something better will always come along and for a period, represents a return to stability. Just as importantly a return to reliability. As for those of us who watch these movements, then perhaps we are among those that understand the lines in that song **Can't let it go:**

- **ଔ** I feel my body in two different places
- **⊗** Still playing for both teams
- Sometimes it feels I was born with two faces

It isn't easy to ignore the current conflict within IT that has us wrestling with staying traditional or going virtual. Of staying centralized or going distributed. Of CapEx or OpEx. And yes, of doing nothing or doing both. Welcome, as vendors like to say, to the hybrid world of IT where it is OK to be supporting both teams. However there is legitimacy in clouds that we as the NonStop community need to acknowledge.







According to AWS, "Cloud computing is the on-demand delivery of IT resources over the Internet with pay-as-you-go pricing." And yet, as Computerworld in an HPE sponsored opinion column of May 31, 2019, noted,



The business and IT needs of any organization are defined by many factors according to IDC, such as the organization's specialization and business model, mix of workloads, financial health, maturity, workforce and customer base.

Furthermore there are positive outcomes that come with the cloud experience. Simply stated, Computerworld says:



The overall cloud experience has had far-reaching benefits on all groups within the enterprise.



Additionally, the pay-as-you-go model used with cloud-based IT lets the enterprise better balance capital and expenses with corporate objectives.



Even though the way applications are built and consumed is changing, IT teams can't easily abandon existing business-critical workloads. After all, the smooth operations of the business depend on the reliable and predictable delivery of business-critical workloads.

For the NonStop user, you just have to love this last sentence. After all Computerworld reports — and yes, it is a HPE paid-for infomercial after all — "business depends on the reliable and predictable delivery of business-critical workloads." It is true that this article was written back in 2019 when little was known about GreenLake (and bringing the cloud experience to the enterprise) or the support of GreenLake by NonStop, it does highlight a potential opportunity for NonStop.

And yet there is, however, vulnerability with an "all-in on clouds" approach. Call it fragility or perhaps better still, a lessening of reliability to where outages are occurring with such frequency that enterprises are beginning to rethink their options. A number of British banks for instance have backed away from going all in with clouds. Recall the recent news coming out of Barclays and yes, there are other financial institutions following suit.

In my last post, **What's not broken and just keeps on running? NonStop delivers!** I wrote of how "Two opportunities come to mind that in the coming months I will be exploring in more detail. And they have to do with how we think about NonStop going forward and whether our own ideas about the role of NonStop may indeed be outdated. Might there be the potential to have NonStop play a guardian role — no pun intended? Should there be a NonStop essentially overseeing the hybrid multi-cloud environment common today among enterprises so that exposure to any one cloud can be marginalized to where outages have no impact on the running of mission critical applications?





When we think about the value that NonStop brings to any discussion on hybrid IT and cloud computing given the above then perhaps it is best explained albeit with a goodly amount of humor by HPE Distinguished Technologist, Keith Moore. "HPE NonStop is a mature peer-group soft system of super-paranoid, autocratic, self-centered kernels working together but with communication rules for up to 4080 enemies, all working independently and together, watching themselves and all other kernels with universal distrust."

When IBM first unveiled Parallel Sysplex for the mainframe – an overt attempt to upgrade the resilience of the mainframe when delivered as a cluster – they introduced the "coupling facility." Initially a dedicated mainframe but later updated to run virtually. The coupling facility runs no application software and has no I/O devices – it is purely in place to oversee the key elements needed by mainframe applications to ensure their continued operation. Often referred to as "clocks, locks and lists" you couldn't run a mainframe cluster with Parallel Sysplex without the coupling facility.

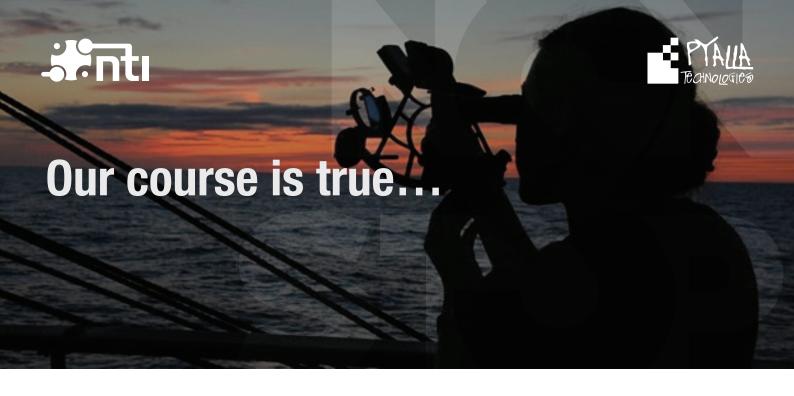
Fast forward to today; clouds and the possibility for NonStop playing a similar role. When an enterprise pursues a hybrid cloud deployments that include traditional systems, on-prem private clouds together with access to a number of public clouds, it's almost impossible to guarantee levels of availability we set down as mandatory for our business-critical workloads. But how would this work? What would it entail?

This has been the subject of a number of conversations that are at best introductory in nature even as they continue to be ongoing. When it comes to how they might work the biggest hurdle is in building out the type of resilient and redundant high-speed interconnect fabrics. Forcing a take-over of a cloud determined to be misbehaving and offloading to another resource, be that another public cloud, or something on-prem demands some level of application awareness. Maybe not check-pointing but perhaps something akin to Pathmon with the resilience it provides may be all that needs to reside on our NonStop "guardian." And then there is the modernization of applications taking place and the way we build new apps may indeed prove beneficial when it comes to easing into this approach to cloud resilience.

"There should be a balance within IT where successful innovations to the business migrate to the traditional side where efficiencies, standards and cost reductions are applied to that innovation. All truly successful innovations should become foundational IT systems. I've never seen that as a goal within agile development but a successful IT project should deliver business benefits for many years," said HPE Master Technologist Justin Simonds. "Doing Agile, fast-fail new business applications in the cloud makes sense since they are quickly constructed and just as quickly deconstructed (if they fail). However I have yet to see a strategy for migration from the cloud, when a new application/service is a clear benefit to the business. When something runs 24x7x365 it is very expensive in the cloud. Applications that become critical to the business are most effectively, efficiently and inexpensively run in-house."

All things are possible and I don't think that the presence of a high-speed duplicated fabric will present an insurmount-able problem — aren't today's WAN offerings already superior to what was present on early models of Tandem Computers? Isn't Pathmon sort of already doing this (above) in a Virtual Machine? Haven't NonStop vendors already taken some baby-steps towards achieving this, whether their approach was simply for disaster recovery or better load balancing? And the monitoring vendors would be well-positioned to provide real time analysis of where your apps and data were running for all those with regulatory issues needing to be addressed?

If you consider this a little too futuristic the only ingredient missing as a real world customer requirement and as it has always happened in the past, unless we take our introductory level conversations public, how will they know that something like this is possible? It certainly lends itself to a baby-steps approach, one app at a time. Maybe you do keep the database on NonStop after all and simply execute the apps anywhere you like under the watchful eye of your NonStop guardian? As I began this post "I've got a feeling and I just can't let it go!"



For those who have hit the high seas with the intension of sailing to a faraway land, there are many elements that can contribute to missing arriving at your harbor. Sail one degree off course then for every 60 miles you sail you will miss your target by one mile. Not much, or so it seems and yet, sail around the world following the equator and you will be off course by 500 miles.

While flying between Sydney and San Francisco during the year I commuted between the Tandem office in North Sydney and the Cupertino campus, there was a time when I received an invite to enter the flight deck. Those were different times when security wasn't an issue and on the Queen of the Sky, that venerable Boeing 747.

Imagine my surprise to see the first officer pulling out his sextant, sighting a planet as it rose from the horizon and then checking with the aircraft's inertial navigation system remarking as he did so that he was right on course.

It shouldn't come as a surprise then to know that the NonStop systems of today have been tracking industry trends for decades. Not just being a part of the technology journey that continues unabated but closely tracking to real world customer requirements. If it wasn't for the uncertainty of 1970s technology then few of these customers would have wanted to invest in fault tolerant computers, but they did. And for good cause!

The world was going real time, 24 x 7 and direct business interactions with end users couldn't afford to be offline for any cause. Those early real time applications involved financial transactions and the timely movement of goods such that the business suffered whenever outages occurred. NonStop systems made the unreliable, reliable! NonStop systems successfully addressed the business requirement for greater availability when built on unreliable 1970s technology.

But change is inevitable, as I have posted throughout the past decade but what is really having an impact on business is the velocity of change. It was only a short time ago when building out proprietary server farms that virtualized everything was the fashion and today the cloud experience has entered the virtual conversation.

That isn't to say, virtualization is out of favor but rather with virtualization yet another layer of technology has arrived overarching the physical and virtual building blocks we have assembled to date. The complexity just ratcheted up further and you don't need to turn to your sextant for a better sighting of where this is leading or to understand you are now straying off course.

And this is the course correction that of itself has the potential to set NonStop towards a vastly changed horizon. There is a strong wind coming in fresh and at its center is the cloud. Yes, as cloud service providers continue to struggle to provide the guaranteed level of availability we all assume today with NonStop, there is an opportunity to move from supporting unreliable computers and even error—prone virtual machines / hypervisors, to where NonStop can treat clouds as nothing more than another grouping of processors. Think AWS as Cloud0, another AWS as Cloud1 and perhaps a private cloud as Cloud3 with Azure, the home to Cloud4.







This may be premature to consider today but in talking to HPE NonStop personnel, the day may be not too far away. When I posted earlier of there being two opportunities for NonStop with one being NonStop playing a guardian role — no pun intended - that was the subject of my previous post, **What's not broken and just keeps on running? NonStop delivers!** The other opportunity focused on treating the world of clouds as no different to either converged NonStop processors or virtual machines. If you think back to the origins of NonStop this isn't too radical an idea even as there are obstacles to doing so effectively today.

Clearly, to those who have been around NonStop for a very long time it all comes down to the interconnect fabric. Can I reliably deliver content between different cloud instances even when they are hybrid? Could I depend upon AWS and Azure to give me access to something that would work in this regard? Could there be a simple answer.

As one of NonStop's leading technologist said, "The answer is 'yes.' We can do inter-cpu failover across 'clouds' if you define the 'clouds' part correctly. I always look at NonStop as a closely-coupled cluster of operating systems. This means that we are actually doing this now when we span VMware vSphere bare metal instances (which we do all of the time for hardware fault resiliency). So if you define a cloud as a single instance of vSphere (VMware orchestration tool for esxi hypervisors), then we already can span clouds."

However, I am looking to go even further, even as I am hoping to attract the attention of at least one of the bigger NonStop users who has already decided to become more engaged with cloud providers. Moving the needle within the NonStop development community after all accelerates whenever there are interested NonStop users actively looking for a solution. Where I want to take the conversation is not just between instances of VMware within say a rack or server farm but rather when it involves communicating over much longer distances.

Latency together with an interconnect fabric that wont let us down under any condition. "As with all things we have done for 45 years, the level at which NonStop can provide a viable offering to fail, grow, recover, retract services, is dependent on the current latency limitations to communicate amongst the components. This has been true for all renditions of Non-Stop systems when it comes to CPU interconnects," said previously quoted NonStop technologist.

Think about this for a moment he then suggested. "Expand has a similar requirement for wide-area network (WAN), but it is much broader and more malleable in deployment. The reason NonStop has never had the intra-CPU (processor-to-processor without Expand) using a WAN is because of the latency, as has been noted. Think in terms of NonStop's ability to leverage the legacy FOX (Expand using fiber and DMA) all the way through to now where it leverages InfiniBand for NonStop clustering. All of these require low-latency guaranteed delivery. This is why NonStop currently requires a closed RoCE-enabled network for its vNSK CPU interconnect — it's about guaranteed response time and message latency."

Before you begin to sigh and think that this was a good idea while it lasted, the industry continues to move on and one of the biggest issues being addressed is the never-ending search for greater speeds over the global WANs that tie businesses together. New networks are becoming available where switching is moving toward 200Gb and 400Gb in 2022-2025 timeframes even as today, from a practical usage point of view, we max out just below 100Gb. But it's coming at some point with even faster speeds possible beyond 2025. This kind of networking would have been unimaginable to those first NonStop engineers back in the late 1970s.

However, this isn't all that is needed as there are other considerations to be made. "In order to accomplish this, NonStop would need to leverage QoS internet services requiring low-latency packet switching and also leverage some HPE technologies ... for intelligent, guaranteed packet routing," my technologist said. As for the good news? "All of this could definitely be done. It's not even something that requires new technologies".

As I just wrote, all it takes is for a NonStop user committing to clouds coming around to the idea that traditional NonStop take over can occur across disparate cloud services to where having multiple cloud vendors supporting mission critical applications, the service levels can be guaranteed to where there would be no difference to running these same mission critical applications in-house on a traditional NonStop system. Imagine that; even as these different cloud providers may not be all that ready to provide access to critical networking components, a strong enough NonStop user backed by HPE could swing a more favorable response than otherwise provided.



As a technology, NonStop has never been fixed in time. That is one reason why nearly five decades later it is still as relevant today as it was back then. We may have moved on from clusters of real CPUs to clusters of virtual machines and making that stretch play to clusters of clouds doesn't seem all that far-fetched to those who spend time considering what might come next for NonStop.

There are reasons why course corrections are made. For the most part, it is as a result of external forces. In these times when IT professionals ponder the potential for cloud support of mission critical applications, surely the time is right to steer NonStop towards supporting clouds as readily as it has real CPUs and virtual machines. When you consider networking has never proved to be a barrier over the long term, who will be the first to demonstrate the levels of availability over their cloud deployments as we have all come to expect with NonStop. Will it be you?



