Legacy Players Feel the Heat as AWS Storage Revenue Approaches $10B. The storage business as we know it has changed forever. On premises, external storage was once a virtually unlimited and untapped bastion of innovation, VC funding and lucrative exits. Today it is a shadow of its former self and the glory days of storage [...]

© 2020 Wikibon Research | Page 1
Legacy Players Feel the Heat as AWS Storage Revenue Approaches $10B.

The storage business as we know it has changed forever. On premises, external storage was once a virtually unlimited and untapped bastion of innovation, VC funding and lucrative exits. Today it is a shadow of its former self and the glory days of storage in the data center will not return. We believe AWS’ storage business is around $6.5-$7B today and will approach $10B annually on a run rate basis by the end of next year. While several trends are affecting the traditional storage business, but by far, AWS’ moves are the most notable.

In this Breaking Analysis we’ll lay out our premise for what’s happening in the storage industry, the impact of AWS and we’ll share some fresh insights from ETR data that supports our current scenario.

Three Decades of Tectonic Shifts

Cloud’s March on the Storage Party Continues

- Traditional storage biz as we know it is over
- A decade of steady share shifts
  a. The cloud sucking sound, led by AWS’ expanding plans
  b. Flash: The death of performance by spindles
- AWS put object on the map w/ S3
  a. Storage awakening within AWS
  b. Much more aggressively expanding its portfolio
  c. AWS storage is now a multi-billion dollar business ($8-$10B)

A simplified history of storage shows us there have been five major waves of innovation over five decades. The dominant industry model has evolved from:

- A mainframe-centric vertically integrated business model led by IBM;
- To a dis-integrated business that saw 70-80 Winchester disk drive companies rise and then fall, serving a booming PC industry - this wave was led by the likes of Seagate;
- Which supplied the emergence of an intelligent, controller-based, external disk array business that drove huge margins for function that while lucrative, was far cheaper than captive storage from system vendors - this era was led by EMC and NetApp;
- And then this business was disrupted by a flash and software-defined, led by Pure Storage and VMware.

The future of storage is being defined by cloud and intelligent data management; and is being led by AWS and a three letter company we’ll call TBD.
The Impact of AWS Cannot be Overstated

While legacy storage players are tired of talking about cloud’s impact on their business, the reality cannot be ignored. Cloud has been the most disruptive force in storage over the past ten years and we’ve reported on the spending impact extensively. But cloud is not the only factor pressuring the on-prem storage business. Flash has killed what we call “performance by spindles,” in other words, the practice of adding more disk drives to keep performance from tanking.

But drilling down into the cloud, AWS has been the most significant factor. Lots of people talked about object storage before AWS but there sure wasn’t much spending going on. S3 changed that. AWS is getting much more aggressive about expanding its portfolio of storage offerings. S3 came out in 2006 – it was the very first AWS service. Then EBS, Elastic Block Storage came 2 years later. But nobody really paid too much attention. Last fall at storage day, we saw AWS announce a number of services, many file related and this year we saw four new announcements from Amazon at re:Invent.

We think AWS’ storage revenue will surpass $6B this year and could be as high as $10B by 2022. There’s not much data out there but this would mean AWS’ storage biz is larger than that of NetApp, which means AWS is larger than every traditional storage player with the exception of Dell.

Here’s a little glimpse of what’s coming at the legacy storage business. It’s a clip of the Vice President of AWS Storage, Mai-Lan Tomsen Bukovec.

Ok, now you may say “what does that have to do with anything” and…well, we just think it’s badass that this woman boxes and runs a business that we think is approaching $10B; and the business she looks after is taking spending from traditional storage in a fairly major way.

AWS Storage Announcements at re:Invent 2020

AWS made four storage announcements this year at re:Invent.

io2 Block Express Volumes

The first is EBS io2 Block Express volumes. AWS claims this is the first storage area network (SAN) for the cloud and it offers up to 256,000 IOPS and 4,000 MB/Sec throughput and 64TB of capacity. Ok now first of all – it’s not really the first SAN in the cloud in our view, there may be others but Pure Storage announced Cloud Block Store in 2019 at it’s annual Accelerate customer conference.

And you may also be saying – what’s the big deal. The performance a joke compared to what many on-prem products can demonstrate. As a point of comparison, there’s a startup called VAST Data. Its 4U enclosure with bundled storage and compute can do 400,000 IOPS and 40,000 MB/Sec. And that can be scaled. So yeah, we hear you.

But the point is that this feature will be great for a lot of workloads. And it’s fully integrated with AWS services. Meaning for example, it will be very convenient for AWS customers to invoke this capability for Aurora and other AWS databases through its RDS service…it’s another easy button for developers to push. This is especially important as we see AWS rapidly expanding its machine learning and AI capabilities with SageMaker and embedding ML into RedShift and driving analytics. Integration is key for its customers.

And now is Amazon retail going to run its business on io2 volumes – I doubt it. I believe they’re running on Oracle and need much better performance – but this is a mainstream service for the EBS masses to tap.
EBS Gp3 Volumes

The other notable announcement from AWS was EBS Gp3 volumes. This is essentially a service that lets you programatically set SLAs for IOPS and throughput, independently, without needing to add additional storage. Again, you may be saying things like Solidfire let me do this years ago and gave me more than 3,000 IOPS and 125 MB/Sec performance. But this is great for mainstream customers that want more consistent performance. And it's integrated into the AWS stack.

And AWS announced that Gp3 volumes is priced at 20% less than previous generation volumes, which you might say is also no big deal - and we would agree - 20% is probably not as aggressive as the average price decline per GB of any storage technology.

Automatic Tiering & Replication

Two other announcements that AWS made at re:Invent this year; one automatically moves data to colder storage tiers and a replication service. On the former, data migrates to tier 2 after 90 days of access and tier 3 after 180 days. AWS hired a bunch of folks out of EMC years ago and put them up in the Boston Seaport - so they've acquired lots of expertise in a lot of different areas and while this is not rocket science, it saves customers money. And for sure we'd like to see more policy options than just a fixed 90 and 180 day policy; and more importantly we'd like to see intelligent tiering where the machine is smart enough to promote certain data sets when needed - like at quarter end for comparisons, etc. But as NFL Hall of Fame coach Hank Stram would say - AWS is matriculating the ball down the field.

"Just keep matriculating the ball down the field boys!" Hank Stram with some classic #WednesdayWisdom for ya. @Chiefs pic.twitter.com/OcR1ZjtRvp

— NFL Films (@NFLFilms) April 11, 2018
Storage Spending in Context - Not a CIO Priority

The chart below shows spending across the ETR taxonomy. It depicts the Net Score – or spending velocity for different sectors. We’ve highlighted storage. Don’t put too much weight on the January data because the survey was just launched. But you can see storage continues to be on the back burner relative to other spending priorities. As we’ve reported – CIOs are really focused on cloud, containers, automation, productivity and other key areas like security.

The Numbers Don’t Lie - Traditional Storage is a Business in Decline

The chart below shows data for eight leading names in storage. We put storage in quotes because as we said – the market is shifting and for sure companies like Cohesity and Rubrik are not positioning as storage players. Rather they’re category creators around intelligent data management. But they’re an adjacency and they are in the ETR storage taxonomy.
As you can see above, we’re showing the year over year quarterly revenue growth for the leading storage companies as reported in their financial statements. NetApp is the big winner, growing a whopping 2% (adjusted for constant currency). They beat expectations but expectations were way down. You can see in the rightmost column we’ve added the ETR Net Score from October. A Net Score of 10% says that if you ask customers are you spending more or less with a company, there are 10% customers spending more than spending less.

For comparison, a company like snowflake has a Net Score around 75%. So 10% is in the red zone. And yet NetApp is the big winner this quarter.

Nutanix isn’t really a storage company, but again they’re adjacent to storage. Nutanix essentially created and still leads the hyper-converged infrastructure (HCI) space. And it, like many of these companies, is transitioning to a subscription pricing model. So that puts pressure on the income statement. Every company on this chart is moving to an annual recurring revenue model and that as-a-service approach will be the norm by the end of the decade (e.g. HPE GreenLake, Dell APEX, etc.).

HPE’s Nimble business has momentum but other parts of the company’s storage portfolio are softer. Dell continues to see pressure on its storage business although VxRail is a bright spot. And IBM’s mainframe storage cycle seems to have run its course.

The Data Protection / Data Management Vendors are Different

And you can see the data protection, data management companies are showing spending momentum. They’re not public so we don’t have revenue data. But one has to wonder…with all the money these guys have raised and the red hot IPO and tech markets – why haven’t these guys gone public? The answer has to be that they’re not ready. Most likely, their numbers aren’t where they want them to be, they’re not predictable enough, they don’t have their operational act together…some combination of those factors. They’ll give other answers but if they had their stuff together they’d be going out right now.

Not shown are other backup, data protection and data management players like Veeam, Commvault and Veritas who are all vying for a piece of the pie.
A Deeper Look at Spending Momentum

The chart below shows the Net Score breakdown for key players in the ETR storage taxonomy. Now again - ETR here is measuring the % of respondents that are adopting new, spending more, spending flat, spending less or retiring the platform. Net Score is calculated as follows: (Adoptions, lime green + Spending More, forest green) minus (Spending Less, pink + Leaving, bright red).

Let’s look at the picture here. Cohesity leads all players in the ETR storage category with a 55% Net Score - you can see the raw data in the upper right corner, followed by Nutanix. Now they’re really not in the scope of a pure storage play; they’re an adjacency - but they do partner with many of the major storage players.

Speaking of Pure, its Net Score has come down from its high of 73% in January of 2016. It won’t likely climb up that high again; but it will be interesting to see if Pure’s Net Score can rebound in a post-COVID world. We’re watching what Pure does in terms of unifying file and object, how it’s faring in cloud and what it does with the Portworx acquisition.

Dell is doing fine with VxRail but VMware’s vSAN is well off its Net Score highs which were in the 60+% range a couple of years ago. But vSAN has definitely been a factor. HPE with Nimble we think still has room to improve and will continue to be strong.

Dell and NetApp are the big leaders in terms of presence in the data other than VMware, which has a lot of instances in the data set because of its volume. Because Dell and NetApp are so large and mature, you’d expect them to have Net Scores that are tepid and you can see a similar pattern with IBM.

Commvault, like Cohesity and Rubrik, is really a data management play - going beyond backup into business recovery, data protection, and bringing that to the cloud. We didn’t put Veeam in here and should have - they had pre-COVID Net Scores well into the 30’s and they have a steadily increasing share of the market and so we expect good things from Veeam going forward.

But it’s all relative. This is a large and mature market that is moving to the cloud and other adjacencies.
On Prem Storage Players are Pivoting

The chart below gives you another view of the competitive landscape. It’s our classic XY graph plotting Net Score on the vertical axis and Market Share on the horizontal. Market Share remember is presence in the data set.

Now think about this from the CIO’s perspective. They have their on-prem estate and they’re putting a brick wall around their core systems. What do they want out of storage for that class of workload? They want it to perform consistently. They want it to be efficient and cost effective. So what are they going to do? They’re going to consolidate vendors, consolidate storage, minimize complexity. On-prem storage arrays are not giving them the innovation and strategic advantage back when virtualization, space efficient snapshots, data deduplication and other storage services were worth taking a flyer on with a “feature product” like 3PAR for example.

Flash gave them headroom and consistent performance and so as more and more work gets pushed to the cloud, you’re seeing the bunkering in on large scale mission critical workloads. As you saw earlier, the legacy storage market is consolidating and has been for a while. It’s becoming a managed decline business where R&D is going to get increasingly squeezed.

We mentioned VAST Data earlier – they’re growing. Another company that’s growing is Infinidat and they are focused on petabyte scale. And they’re having success consolidating storage. Ironically, these are two Israeli-founder based companies that are having success. But it’s all a share shift. These two companies are tiny in the grand scheme of things and their end game is to get acquired - so they can steal share but they’re not going to reverse the trends.

It’s worth noting another point that underscores the consolidation that’s occurring in the market. Just look at the latest Gartner Magic Quadrant for primary storage. Sixty-two percent (62%) of the competitors are in the leaders quadrant!

The point is differentiation is getting more difficult. And everyone on the ETR chart above – every on prem player – has to have a cloud strategy where they connect to the cloud, take advantage of native cloud
services and help extend their respective installed bases into the cloud...including having a capability that is physically proximate to the cloud; with a co-lo or other approach. In addition almost all players are going after multi-cloud and building subscription businesses.

AWS is Re-Defining Hybrid and the Edge

AWS wants to bring AWS to the edge – and it looks at the data center as just another edge node. This is a different view than that of the traditional on-prem players. For example at re:Invent we saw AWS’ hybrid strategy evolve. We saw smaller versions of Outposts, added Local Zones, SDKs for edge computing and other related announcements– all part of bringing AWS’ strategy to expand its presence.

We also saw a few companies – Pure, Infinidat and Veeam come to mind, that are connecting to Outposts. Also Qumulo, Clumio, Commvault and Weka made similar announcements. No doubt there are some we may have missed and more coming.

**Prognosis for the Future**

Where have all the good times gone? Look, we would never bet against data but the days of selling storage controllers that mask the deficiencies of spinning disk or add embedded hardware functions or easily picking off a legacy installed bases with flash...well those days are gone.

**Prognosis**

- Don't ever bet against data
- Hyperscale players are only going to gain momentum this decade
- Legacy business models are changing in storage
  - Circle the wagons around mission critical & low latency workloads
  - Develop a cloud-first strategy to support the installed base
    - Physically locate proximate to clouds to leverage archiving & other services (aka hybrid)
    - Includes as-a-service pricing models
  - Pivot to intelligent data management
  - Develop an edge strategy that is profitable

The narrative we hear from many on-prem vendors about repatriation isn't happening in any meaningful way. CIOs are going to rationalize their on-prem portfolios so they can invest in the cloud, AI, automation and re-skill their teams. Low latency, high bandwidth workloads with minimal jitter are the sweet spot for on-prem. It’s becoming the mainframe of storage.

CIOs are also developing a cloud-first strategy. Yes the world is hybrid. What does that mean to CIOs? It means you’re going to have some work in the cloud and some on prem. Everything that can go to the cloud will. Everything that can't won't. And they'll build an abstraction layer to connect workloads from an observability standpoint so they can maintain control and manage lock-in risk.
Breaking Analysis: Legacy Players Feel the Heat as AWS Storage Revenue Approaches $10B

Everything that doesn’t go to the cloud will likely have some type of “hybrid-icity” to it. The reverse won’t likely be the case because lots of workloads that are running in the cloud don’t need on-prem linkages.

For vendors, cloud strategies involve supporting customer migrations to the cloud. Enabling low latency hybrids, accommodating subscription models and managing the decline of a large estate, profitably. Traditional vendors are re-thinking what business they’re in – e.g. data management and developing an edge strategy that recognizes that edge workloads will require a new architecture that’s more efficient than we’ve seen built around general purpose systems.

The future is cloud, the cloud is expanding, the market is large – it’s just a lot different than what we’ve seen over the last twenty years.

Ways to Connect

Remember these episodes are all available as podcasts wherever you listen. Check out ETR’s Web site. We also publish a full report every week on wikibon.com and on SiliconANGLE.

Ways to get in touch: Email: david.vellante@siliconangle.com | DM @dvellante on Twitter | Comment on our LinkedIn posts.

Stay safe and we’ll see you next time.

Watch the full video analysis:

Photo credit: Africa Studio
Breaking Analysis: Legacy Players Feel the Heat as AWS Storage Revenue Approaches $10B

David Vellante is co-CEO of SiliconANGLE Media, as well as co-founder and Chief Analyst at The Wikibon Project, the world’s leading open source technology research community. Dave is a long-time tech industry analyst, entrepreneur, writer and speaker. As co-host of theCUBE – “The ESPN of Tech,” Vellante has interviewed over 5,000 experts since 2010. He is also a co-founder of CrowdChat, an angel funded startup based in Palo Alto using big data techniques to extract business value from social data. Prior to these exploits, Dave founded a CIO consultancy and spent a decade growing and managing IDC’s largest business unit. He lives in Massachusetts with his wife and four children where he is active in town activities including serving as the president of his town’s local “Kiddie Sports” association. Dave holds a B.S. in Applied Mathematics from Union College.

David Vellante
@dvellante
david.vellante@wikibon.org