



AWS News Blog

Amazon EBS Update – New Cold Storage and Throughput Options

by [Jeff Barr](#) | on 19 APR 2016 | in [Launch](#), [News](#) | [Permalink](#) | [Share](#)

The AWS team spends a lot of time looking in to ways to deliver innovation based around improvements in price/performance. Quite often, this means wrestling with interesting economic and technical dilemmas.

For example, it turns out that there are some really interesting trade-offs between HDD and SSD storage. On the one hand, today's SSD devices provide more IOPS per dollar, more throughput per gigabyte, and lower latency than today's HDD devices. On the other hand, continued density improvements in HDD technology drive the cost per gigabyte down, but also reduce the effective throughput per gigabyte. We took this as a challenge and asked ourselves—could we use cost-effective HDD devices to build a high-throughput storage option for EBS that would deliver consistent performance for common workloads like big data and log processing?

Of course we could!

Today we are launching a new pair of low-cost EBS volume types that take advantage of the scale of the cloud to deliver high throughput on a consistent basis, for use with EC2 instances and [Amazon EMR](#) clusters (prices are for the US East (N. Virginia) Region; please see the [EBS Pricing](#) page for other regions):

- **Throughput Optimized HDD (st1)** – Designed for high-throughput MapReduce, Kafka, ETL, log processing, and data warehouse workloads; \$0.045 / gigabyte / month.
- **Cold HDD (sc1)** – Designed for workloads similar to those for Throughput Optimized HDD that are accessed less frequently; \$0.025 / gigabyte / month.

Like the existing General Purpose SSD (gp2) volume type, the new magnetic volumes give you baseline performance, burst performance, and a burst credit bucket. While the SSD volumes define performance in terms of IOPS (Input/Output Operations Per Second), the new volumes define it in terms of throughput. The burst values are based on the amount of storage provisioned for the volume:

- **Throughput Optimized HDD (st1)** – Starts at 250 MB/s for a 1 terabyte volume, and grows by 250 MB/s for every additional provisioned terabyte until reaching a maximum burst throughput of 500 MB/s.
- **Cold HDD (sc1)** – Starts at 80 MB/s for a 1 terabyte volume, and grows by 80 MB/s for every additional provisioned terabyte until reaching a maximum burst throughput of 250 MB/s.

Evolution of EBS

I like to think of customer-driven product and feature development in evolutionary terms. New offerings within a

category often provide broad solutions that are a good fit for a wide variety of use cases. Over time, as we see how customers put the new offering to use and provide us with feedback on how we can do even better, a single initial offering will often speciate into several new offerings, each one tuned to the needs of a particular customer type and/or use case.

The various storage options for EC2 instances are a great example of this. Here's a brief timeline of some of the most significant developments:

- [2006](#) – EC2 launched with instance storage.
- [2008](#) – EBS (Elastic Block Storage) launched on magnetic storage.
- [2012](#) – EBS Provisioned IOPS and EBS-Optimized instances.
- [2014](#) – SSD-Backed general purpose storage.
- [2014](#) – EBS data volume encryption.
- [2015](#) – Larger and faster EBS volumes.
- [2015](#) – EBS boot volume encryption.
- 2016 – EBS Throughput Optimized HDD (st1) and Cold HDD (sc1) volume types.

Workload Characteristics

We tuned these volumes to deliver great price/performance when used for big data workloads. In order to achieve the levels of performance that are possible with the volumes, your application must perform large and sequential I/O operations, which is typical of big data workloads. This is due to the nature of the underlying magnetic storage, which can transfer contiguous data with great rapidity. Small random access I/O operations (often generated by database engines) are less efficient and will result in lower throughput. The General Purpose SSD volumes are a much better fit for this access pattern.

For both of the new magnetic volume types, the burst credit bucket can grow until it reaches the size of the volume. In other words, when a volume's bucket is full, you can scan the entire volume at the burst rate. Each I/O request of 1 megabyte or less counts as 1 megabyte's worth of credit. Sequential I/O operations are merged into larger ones where possible; this can increase throughput and maximizes the value of the burst credit bucket (to learn more about how the bucket operates, visit the Performance Burst Details section of my [New SSD-Backed Elastic Block Storage](#) post).

If your application makes use of the file system and the operating system's page cache (as just about all applications do), we recommend that you set the volume's read-ahead buffer to 1 MiB on the EC2 instance that the volume is attached to. Here's how you do that using an instance that is running Ubuntu or that was booted from the Amazon Linux AMI (adjust the device name as needed):

```
$ sudo blockdev --setra 2048 /dev/xvdf
```

The value is expressed as the number of 512-byte sectors to be used for buffering.

This value will improve read performance for workloads that consist of large, sequential reads. However, it may increase latency for workloads that consist of small, random read operations.

Most customers are using Linux kernel versions before 4.2 and the read ahead setting is all they need to tune. For customers using newer kernels, we also recommend setting `xen_blkfront.max` to 256 for the best performance. To set this parameter on an instance that runs the Amazon Linux AMI, edit `/boot/grub/menu.list` so that it invokes the kernel as follows:

```
kernel /boot/vmlinuz-4.4.5-15.26.amzn1.x86_64 root=LABEL=/ console=ttyS0
xen_blkfront.max=256
```

If your file contains multiple entries, edit the one that corresponds to the active kernel. This is a boot-time setting so you'll need to reboot the instance in order for the setting to take effect. If you are using a Linux distribution that does not use the Grub bootloader, you will need to figure out how to make the equivalent change to your configuration.

For more performance tuning tips, please read [Amazon EBS Volume Performance on Linux Instances](#) and [Amazon EBS Volume Performance on Windows Instances](#).

Comparing EBS Volume Types

Here's a table that summarizes the specifications and use cases of each EBS volume type (Although not shown in the table, the original EBS Magnetic offering is still available if needed for your application):

	Solid State Drive (SSD)		Hard Disk Drive (HDD)	
Volume Type	Provisioned IOPS SSD (io1)	General Purpose SSD (gp2)	Throughput Optimized HDD (st1)	Cold HDD (sc1)
Use Cases	I/O intensive NoSQL and relational databases.	Boot volumes, low-latency interactive applications, dev, test.	Big data, data warehouses, log processing.	Colder data requiring fewer scans per day.
Volume Size	4 GB – 16 TB	1 GB – 16 TB	500 GB – 16 TB	500 GB – 16 TB
Max IOPS/Volume	20,000 (16 KB I/O size)	10,000 (16 KB I/O size)	500 (1 MB I/O size)	250 (1 MB I/O size)
Max IOPS/Instance (using multiple volumes)	48,000	48,000	48,000	48,000
Max Throughput/Volume	320 MB/s	160 MB/s	500 MB/s	250 MB/s
Max Throughput/Instance	800 MB/s	800 MB/s	800 MB/s	800 MB/s
Price	\$0.125/GB-month + \$.065/provisioned IOPS/month	\$0.100/GB-month	\$.045/GB-month	\$.025/GB-month
Dominant Performance Attribute	IOPS	IOPS	MB/s	MB/s

You also have the option to further boost performance by using [EBS-Optimized instances](#) and RAID to create file systems that are larger and/or support more IOPS. Read about [RAID Configuration on Linux](#) and [RAID Configuration on Windows](#) to learn more.

CloudFormation Template for Testing

In order to make it as easy as possible for you to set up a test environment on a reproducible basis, we have created a simple CloudFormation template. You can launch the [st1 template](#) to create an EC2 instance with a 2 terabyte **st1** volume attached. The [st1 template instructions](#) contain some additional information.

From our Partners

Several AWS partners have been working with the new volume types ahead of today's launch. Here are their thoughts and observations:

- **Localitytics** shared their [Amazon EBS SC1 Initial Impressions](#). They created a petabyte scale data warehouse backed by a RAIDed set of **sc1** volumes. After running for a day, they saw that performance was indistinguishable from that provided by their earlier configuration. Based on their calculations the new configuration will allow them to save 10% on their platform operating costs.
- **Confluent** wrote about [Deploying Apache Kafka on AWS Elastic Block Store \(EBS\)](#). After taking a look at the new offerings, they are looking forward to creating Kafka clusters that are more cost-effective than ever before.
- **Splunk** explained how their customers can [Retain More Data at Lower Cost with new AWS Storage Volume Types](#). According to the cost model described in the blog post, Splunk users can achieve a 71% savings by moving some of their colder data to **sc1** storage.

Available Now

The new volume types are available now and you can start using them today with EC2 and EMR. You can create them from the [AWS Management Console](#), [AWS Command Line Interface \(CLI\)](#), [AWS Tools for Windows PowerShell](#), [AWS CloudFormation](#) templates, the [AWS SDKs](#), and so forth.

As you can see from the table above, this new offering gives you a unique combination of high throughput and a very low cost per gigabyte.

I am looking forward to your feedback so that we can continue to evolve EBS to meet your ever-growing (and continually diversifying) needs. Leave me a comment and I'll make sure that the team sees it.

— [Jeff](#);

PS – If you are a developer, development manager, or product manager and would like to build systems like this, please visit the [EBS Jobs](#) page.

TAGS: [Amazon Elastic Block Store](#)

Resources

[Getting Started](#)







[What's New](#)

[Top Posts](#)

[Official AWS Podcast](#)

[Case Studies](#)

Follow

-  [Twitter](#)
-  [Facebook](#)
-  [LinkedIn](#)
-  [Twitch](#)
-  [RSS Feed](#)
-  [Email Updates](#)

Related Posts

[New – TLS Termination for Network Load Balancers](#)

[Amazon WorkLink – Secure, One-Click Mobile Access to Internal Websites and Applications](#)

[AWS Backup – Automate and Centrally Manage Your Backups](#)

[New – Amazon DocumentDB \(with MongoDB Compatibility\): Fast, Scalable, and Highly Available](#)

[Western Digital HDD Simulation at Cloud Scale – 2.5 Million HPC Tasks, 40K EC2 Spot Instances](#)

[Now Open – AWS Europe \(Stockholm\) Region](#)

[And Now a Word from Our AWS Heroes...](#)

[AWS re:Invent Security Recap: Launches, Enhancements, and Takeaways](#)