WEEK 3 - PAPER 1

Large-Scale Data Storage Systems – DATA-5400 | Spring 2020

Christina Morgenstern

Estimating the costs for storing data in the cloud is important because if you know in advance how much it will cost it will save you some money. Amazon AWS has several options for storing data in the cloud, including the AWS S3 standard, the AWS EBS and the AWS S3 Glacier services, which offer a range of storage options in the cloud.

The Amazon Simple Storage Service (Amazon S3) is a scalable, durable platform that allows data to be accessed from any location using the internet for serverless computing, Big Data storage, backup and recovery. Amazon Elastic Block Store (Amazon EBS) offers local storage on Amazon EC2 for relational and NoSQL databases, data warehousing, Big Data processing, backup or recovery. Highly affordable long-term storage classes are provided by Amazon S3 Galcier services [1]. In the following, I will calculate the costs for storing 100 TB and 1 Petabyte of data on the different AWS storage services specifying the EU Frankfurt region and using the Simple Online Calculator tool [2].

Within the Amazon Elastic Block Store one can choose from four different volume types in order to balance optimal price and performance and to be used with Amazon Elastic Compute Cloud. Benefits of this type of storage include the ease in use, encryption and set-up, the failure-proof reliability, high performance, the flexibility to increase storage without impact on workload.

In order to accommodate 100 TB, I selected the Cold HDD (st1) option with 10000 GB storage each and calculated 10 volumes. The total monthly costs for this option would be $3,000 (see Fig. 1 and Fig. 4). This is substantially cheaper than the other volume types. The throughput Optimized HDD (st1) type of EBS volume is designed as a low-cost ($0.045/GB-month) HDD volume for frequent access and intensive workloads

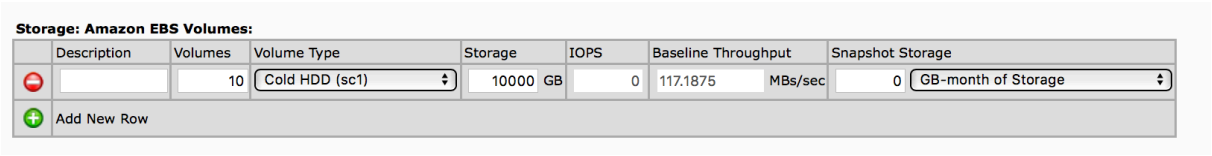

Figure 1. Estimating monthly costs for 100 TB storage on Amazon EBS Volumes.

Calculating the cost for 100 TB of storage on Amazon S3 is done using the calculator tool and selecting the S3 side tab (Fig. 2). The monthly pricing for this object storage service is approx. $2,500 (see Fig. 4).

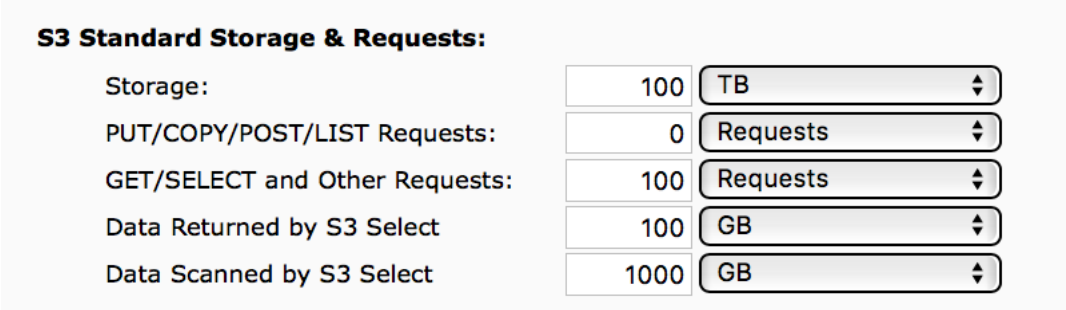

Figure 2. Estimating monthly costs for 100 TB Storage on Amazon S3.

The cheapest storage option among the three selected AWS services is the Amazon S3 Glacier storage with monthly costs of about $460 (see Fig. 3 and 4).

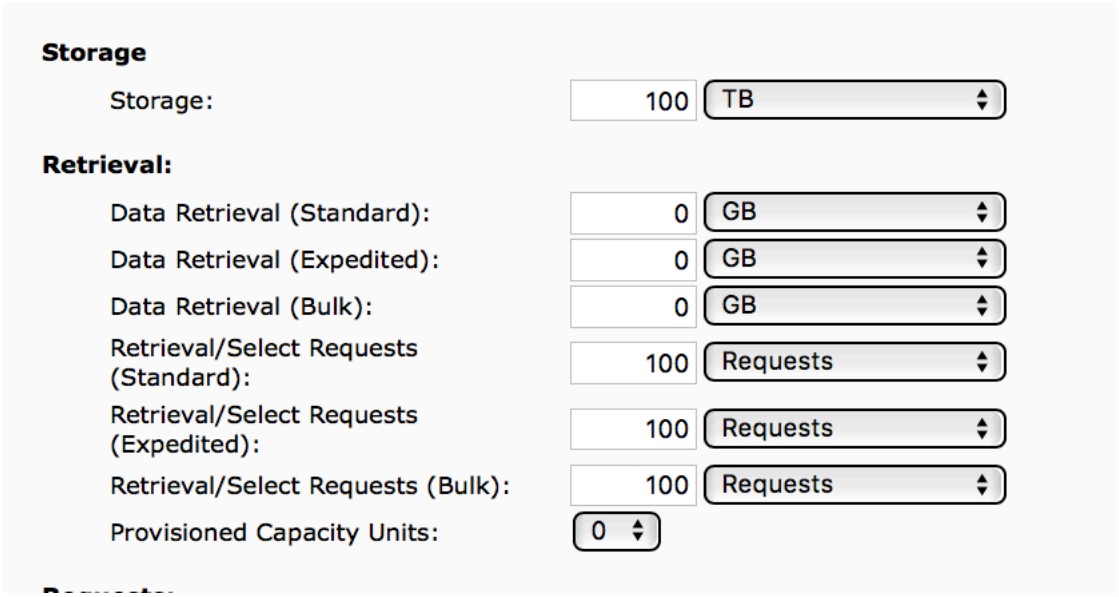

Figure 3. Estimating monthly costs for 100 TB Storage on Amazon Glacier.

⊖	Amazon EC2 Service (EU (Frankfurt))		$ 3000.00
	EBS Volumes:	$ 3000.00	
⊖	Amazon S3 Service (EU (Frankfurt))		$ 2459.93
	S3 Standard Storage:	$ 2457.60	
	S3 Standard Other Requests:	$ 0.00	
	S3 Standard Select Data Returned:	$ 0.08	
	S3 Standard Select Data Scanned:	$ 2.25	
⊖	Amazon Glacier Service (EU (Frankfurt))		$ 462.02
	Storage:	$ 460.80	
	Retrieval/Select Requests (Standard):	$ 0.01	
	Retrieval/Select Requests (Expedited):	$ 1.20	
	Retrieval/Select Requests (Bulk):	$ 0.01	
⊖	AWS Support (Business)		$ 592.19
	Support for all AWS services:	$ 592.19	
Free Tier Discount:		$	-0.13
Total Monthly Payment:		$	6514.01

Figure 4. Estimated monthly costs for 100 TB storage on AWS EC2, S3 and Glacier for EU Frankfurt region.

Increasing the storage demand to 1 PB (=1,000 TB), also increases the cost by a factor of 10 for each of the investigated services (see Fig. 5).

⊖	Amazon EC2 Service (EU (Frankfurt))		$ 30000.00
	EBS Volumes:	$ 30000.00	
⊖	Amazon S3 Service (EU (Frankfurt))		$ 24158.49
	S3 Standard Storage:	$ 24156.16	
	S3 Standard Other Requests:	$ 0.00	
	S3 Standard Select Data Returned:	$ 0.08	
	S3 Standard Select Data Scanned:	$ 2.25	
⊖	Amazon Glacier Service (EU (Frankfurt))		$ 4719.82
	Storage:	$ 4718.60	
	Retrieval/Select Requests (Standard):	$ 0.01	
	Retrieval/Select Requests (Expedited):	$ 1.20	
	Retrieval/Select Requests (Bulk):	$ 0.01	
⊖	AWS Support (Business)		$ 4421.48
	Support for all AWS services:	$ 4421.48	
Free Tier Discount:		$	-0.13
Total Monthly Payment:		$	63299.66

Figure 5. Estimated monthly costs for 1 PB storage in AWS EC2, S3 and Glacier for EU Frankfurt region.

Azure Blob storage is Microsoft´s product range for storage solutions in the cloud. It is designed for storing large amounts of unstructured data such as text or binary data. Pricing is dependent on a pay-as-you-go model with no upfront cost and termination fees. There are different options available from low-cost archive solution of Block Blobs to high-throughput Managed Disks offering persistent and secured disks with simple and scalable virtual machine deployment [4]. With Azure Blob storage you can choose between four available access tiers: *Premium* optimized for high transaction rate, *Hot* for storing data that is accessed frequently, *Cool* for storing data that is accessed infrequently and *Archive* for storing data that is rarely accessed. See Figure 6 for pricing details.

According to the pricing table in Figure 6, 100 TB in the Premium, Hot, Cool and Archive section would cost $19,500 (100,000x0.195$), $1,925 (50,000x0.0196 $+ 50,000x0.0189$), $1,000 (100,000x0.01) and $180 (100,000x0.0018$) per month, respectively.

Increasing the storage amount to 1 PB, increases the cost by a factor of 10 per month leading to the following monthly rates of $195,000, $19,250, $10,000 and $1,800 for Premium, Hot, Cool and Archive solutions, respectively.

Data storage prices pay-as-you-go

All prices are per GB per month.

	PREMIUM	HOT	COOL	ARCHIVE
First 50 terabyte (TB) / month	$0.195 per GB	$0.0196 per GB	$0.01 per GB	$0.0018 per GB
Next 450 TB / Month	$0.195 per GB	$0.0189 per GB	$0.01 per GB	$0.0018 per GB
Over 500 TB / Month	$0.195 per GB	$0.0181 per GB	$0.01 per GB	$0.0018 per GB

Figure 6. Microsoft Azure Blob storage pricing for West Europe region.

References:

[1] https://aws.amazon.com/products/storage/

[2] https://calculator.s3.amazonaws.com/index.html

[3] https://aws.amazon.com/ebs/

[4] https://azure.microsoft.com/en-us/pricing/details/storage/