SCTE·ISBE CABLE-TEC EXPO 2017
DENVER, CO
OCTOBER 17-20

2017 Fall Technical Forum
SCTE·ISBE • NCTA • CABLELABS

Media Processing On Cloud

Scalable, Manageable, and Cost Effective

A Technical Paper prepared for SCTE/ISBE by

Srini Akkala
Senior Architect
ARRIS
37 Whitehall Way, Bellingham, MA 02019
301-529-5230
srini.akkala@arris.com

© 2017 SCTE-ISBE and NCTA. All rights reserved.

Table of Contents

List of Figures

List of Tables

© 2017 SCTE-ISBE and NCTA. All rights reserved.

Introduction

Media processing is becoming a complicated problem. The dramatic increase in the number of viewing platforms is creating the need for multiple file renditions for each asset. Add multi digital rights management requirements into the mix and before long VOD libraries can easily expand from a few thousand to over hundreds of thousands of unique assets. This type of scale requires a new approach to business, infrastructure, and networking systems. Due to these added complexities of logic & scale, this new approach requires lot of planning and design in terms of storage, network bandwidth, and choice of tools & technologies.

On-Premises deployment starts with capacity planning (compute capacity, storage capacity, & network capacity). This model is not designed to handle unforeseen processing demands. Even with the best planning and forecasting, an on-premises solution can quickly run out of storage, network bandwidth, and/or processing power with the ever increasing asset library and growing customer base.

Given the ever changing video technology landscape, growth in consumer demand and content availability - setting up a cost effective and "future proof" media processing solution on-premises is often an unrealistic dream. Adapting to cloud based media processing solutions can make this dream a reality.

Solution

VOD processing involves different types of assets from multiple content providers publishing to varying types of screens. A typical media process workflow can be seen below. This paper discusses how each step can be designed to be scalable, manageable, and cost-effective by implementing a cloud-based solution.

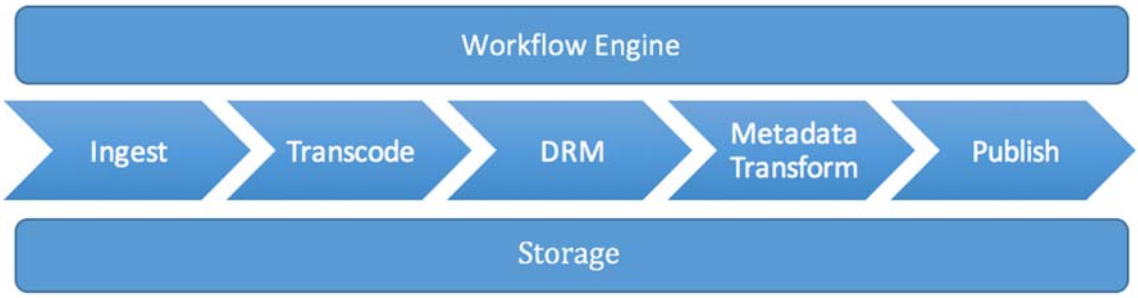

Figure 1 - Content processing workflow steps

The figure below illustrates how the above media processing steps are mapped to Amazon Elastic Compute Cloud (EC2) components.

AWS is taken as an example cloud computing platform, but the solution can easily be implemented in Google Cloud Platform or Microsoft Azure as both have dedicated media processing functionality built-in.

© 2017 SCTE-ISBE and NCTA. All rights reserved.

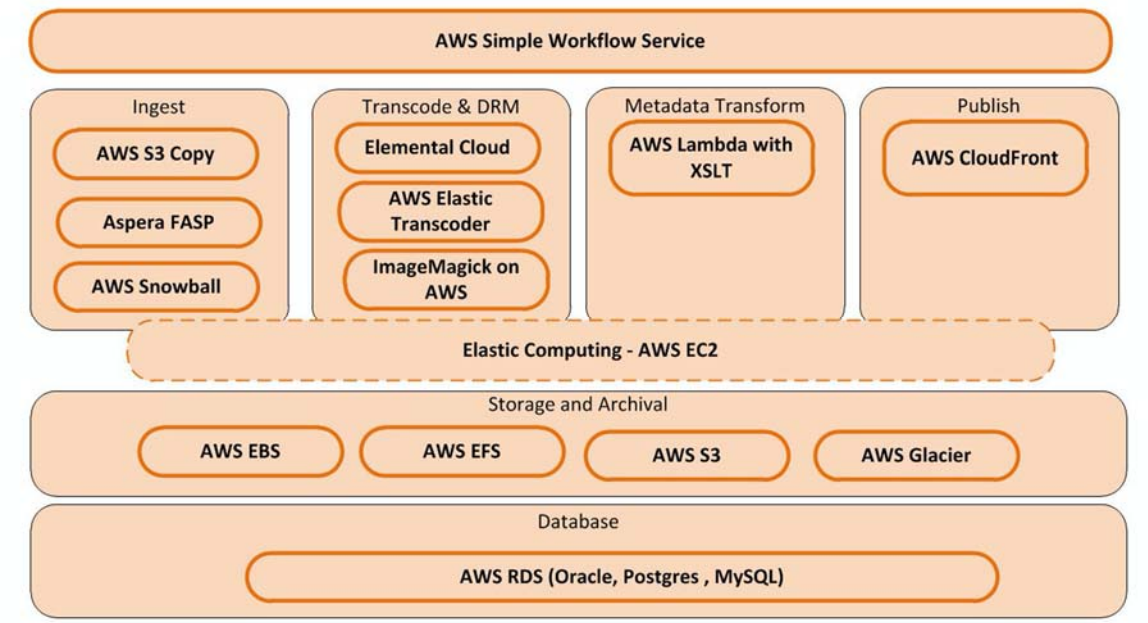

Figure 2 – AWS Components

Elastic Computing

One of the main benefits of cloud computing is elasticity. Elasticity is the ability to instantly grow in scale to meet the resources required during periods of peak or unforeseen load demands. The platform can be scaled up by adding more resources when demand increases and then removed as demand diminishes. Amazon's EC2 service provides such an elastic computing capability.

One of the primary factors to be considered for compute capacity planning, in the context of VOD media processing, is number of assets required to be processed on a daily basis (e.g.: 200+ per day). Transcoding capacity is determined based on daily loads. On occasion, Cable providers run into situations which require the re-processing of their entire library within a defined period of time. For example, "*Offering 4K content"* requires re-processing of most of the library to generate content in 4k resolution. Accommodating re-processing requirements along with daily loads is not feasible with existing transcoding capacity. In this circumstance, one of the only options is to procure new transcoders to fulfill this kind of re-processing demand, which is not cost effective. Elastic computing enables you to procure (lease) extra transcoding capacity for a certain amount of time and shrink back to the original capacity after the re-process is completed.

© 2017 SCTE-ISBE and NCTA. All rights reserved.

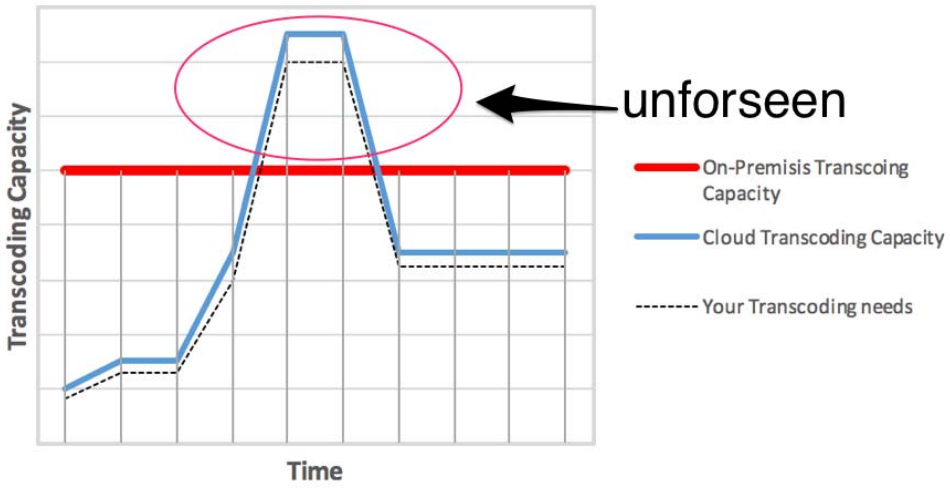

Figure 3 – Transcoding Capacity Scaling

Storage and Archival

Media content management generally requires four unique types of storage, each with its own performance requirements.

- High performance local storage for media processing (transcode, segment)
- Shared network storage (e.g. catchers)
- Fast web object storage (e.g. image files, metadata files, file segments for http delivery)
- Archival storage (mezzanine media, backups)

Amazon Web Services offers 4 types of storage -

Amazon EBS provides block-level storage that serves as a virtual hard drive for your Amazon EC2 instance. Amazon EBS is designed for workloads that require persistent storage accessible by single EC2 instances. Content is copied from EFS or S3 to EBS for high speed content processing. Processed content is copied back to EFS or S3.

Amazon S3 is object storage designed to store and access any type of data over the Internet. It is secure, 99.999999999% durable, and scales past tens of trillions of objects. S3 buckets can be configured as catchers to where providers can pitch the content.

Amazon EFS provides simple, scalable file storage for use with Amazon EC2 instances. EFS data can be accessed from multiple AWS instances.

Amazon Glacier is an extremely low-cost and highly durable storage service for long-term backup and archive of any type of data. Source content is archived in Glacier after processing.

© 2017 SCTE-ISBE and NCTA. All rights reserved.

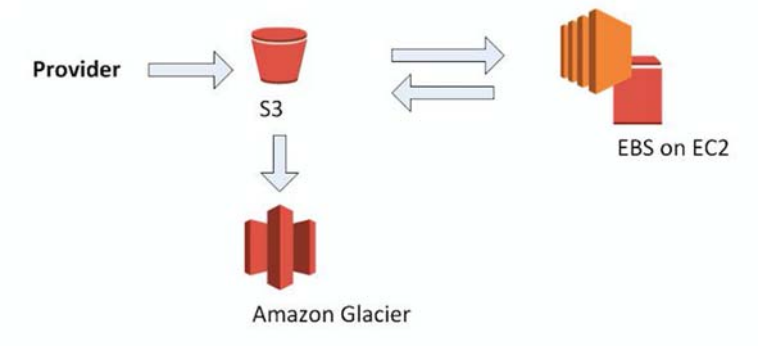

Figure 4 – AWS storage tiers

Database

Data storage requirements in the media industry are growing rapidly with each passing day (e.g.: content metadata, service usage data, viewership interaction events data etc.). Databases are getting bigger all the time. Running databases means lots of repetitive work (installation, configuration, administration, disaster recovery etc.). Amazon RDS makes it easy to set up, operate, and scale a database in the cloud. With Amazon RDS, there is no need to buy, rack, & stack hardware and no need to install software. Some of the most important benefits of using Amazon RDS include:

High Availability Amazon RDS is a highly available service which provides a SLA up-time of 99.95%

Scalability Amazon RDS offers two types of scalability features.

Vertical Scalability - Amazon enables push-button vertical scaling. This means that you can scale the size of an RDS instance [memory, CPU, PIOPS etc.] or disk, either up or down, with the click of a button.

Horizontal Scalability – The entire database is distributed across many RDS instances that will work together.

Backups: Amazon RDS offers Automated as well as Point-in-Time snapshot backups.

Available Amazon RDS engines : PostgreSQL, MySQL, Oracle, SQLServer.

Disaster Recovery

One of the main highlights of cloud computing is disaster recovery. With data centers in regions all around the world, AWS provides a set of cloud-based disaster recovery services that enable rapid recovery of infrastructure and data.

Workflow

AWS SWF (Simple WorkFlow service) makes it easy to build media work flows that coordinate operations across work steps involving distributed components. Coordinating tasks across the application involves a great deal of house keeping with respect to the logical flow of the application. Amazon SWF

© 2017 SCTE-ISBE and NCTA. All rights reserved.

gives you full control over implementing these tasks and coordinating them without worrying about underlying complexities such as tracking their progress and maintaining their state.

If we take an example of a typical media process flow - large videos are uploaded to Amazon S3 in segments. The upload of the segments has to be monitored. After a segment is uploaded, it is transcoded by downloading it to an Amazon EC2 instance. The encoded segment is stored to another Amazon S3 location. Failures could occur during this process due to one or more segments encountering encoding errors. Such failures need to be detected and handled through Amazon SWF's cloud workflow management.

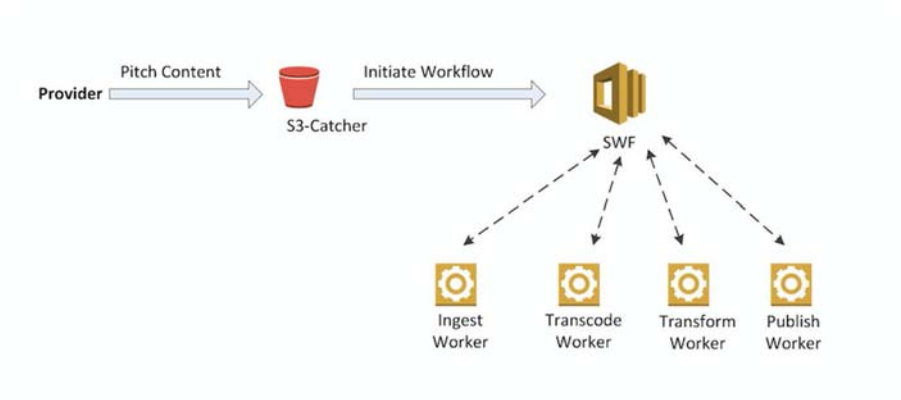

Figure 5 – Workflow Engine

Some of the most important benefits of using Amazon SWF include:

- Amazon SWF replaces the complexity of custom-coded workflow solutions and process automation software with a fully managed cloud workflow web service.
- Amazon SWF lets you write your application components and coordination logic in virtually any programming language.
- Amazon SWF seamlessly scales with your application's usage.

Ingest

The first step in a media workflow is receiving the content package from a content provider and ingesting it into the system. AWS provides multiple tools for uploading large amounts of data into S3:

S3 Copy: Command line interface tool to transfer data to and from AWS.

Aspera High Speed Transfer Service: Industry leader in high speed file transfers. Direct-to-Cloud integration with S3 APIs ensures data is written directly to S3 storage during the transfer and then immediately made available when the transfer completes. Complete data protection with in-transit & at-rest encryption, data integrity verification for each transmitted block, and automatic retry and resume from point of interruption on failure are some of the key security features.

AWS Snowball: This service is applicable for one-time petabyte-scale data transfer scenarios, for example if you are planning to move an entire library from on-premises to the cloud or a provider is planning to transfer their entire library to the cloud.

© 2017 SCTE-ISBE and NCTA. All rights reserved.

Data is transferred into and out of AWS using physical storage appliances. Simply create a job in the AWS Management Console and a Snowball appliance will be automatically shipped to you. Once it arrives, attach the appliance to your local network and copy the library to the appliance. The data is encrypted and transferred to the appliance at high speed. Once the transfer is complete and the appliance is ready to be returned, the shipping label will automatically update and you can track the job status via text messages, or directly in the Console. The data typically appears in your S3 bucket in a couple of days.

Transcode & DRM

Transcoding & DRM can be done on cloud in two ways, by installing any software based transcoders on EC2 instances or by subscribing to cloud based transcoding services. (e.g. Elemental Cloud).

Elemental Cloud is a Platform as a Service (PaaS) built on AWS cloud Infrastructure. The platform automatically provisions and dynamically scales any combination of Elemental video processing, delivery, and storage services within a secure private network. Elemental cloud can automatically scale resources to process and deliver broadcast quality video as demand fluctuates.

Elemental can output multiple container formats include 3GPP, MP4, F4V, MPEG-TS, MOV etc.). Elemental supports adaptive streaming - Adobe's HTTP Dynamic Streaming (HDS), Apple's HTTP Live Streaming (HLS), MPEG-DASH (MP4, ISO, TS), and Microsoft's Smooth Streaming (ISMV). Elemental support content protection using multiple digital rights management (DRM) technologies

Sample Total Cost of Ownership (On-Premises transcoding Vs Elemental cloud) over a five-year period of time is illustrated in Appendix-A.

Other transcoding solutions on AWS include: Amazon Elastic Transcoder, ZenCoder, encoding.com, and harmonic.

Cover art resizing is one of the most important parts of the transcoding process. Content providers usually provide a single high resolution image as part of the content feed which then needs to be resized to fit various client devices. This step is not offered in most off-the-shelf transcoders. ImageMagick is one of most popular open source software products for image resizing. ImageMagick can be installed on an AWS EC2 instance. The Amazon market place also offers multiple SaaS based image re-sizing services which are based on ImageMagick.

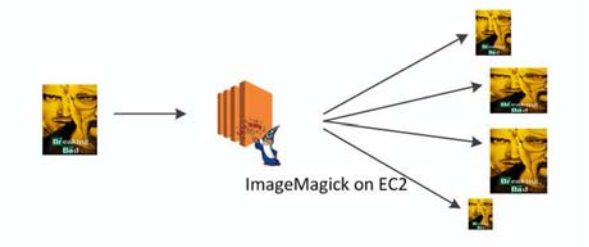

Figure 6 – Image Resizing on AWS

© 2017 SCTE-ISBE and NCTA. All rights reserved.

Metadata Transformation

Source XML (ADI) metadata needs to be transformed into various formats based on the target device or storefront. Custom software (using XSLT) needs to be developed to achieve these types of transformations. Traditionally, custom code is developed and developers need to spin up VMs (EC2) and install the correct software stack prior to uploading and running code. AWS Lambda can be leveraged for these types of custom code modules.

AWS Lambda is a server-less environment in the AWS cloud. Lambda lets you run code without provisioning or managing servers. With Lambda, you can run code for virtually any type of application or backend service, all with zero administration. Just upload your code and Lambda takes care of virtually everything needed to run and scale your code with high availability. In addition, Lambda automatically scales your application by running code in response to each trigger. Your code runs in parallel and processes each trigger individually, scaling precisely with the size of the workload.

AWS Lambda supports code written in Java, C#, Python, and Node.js

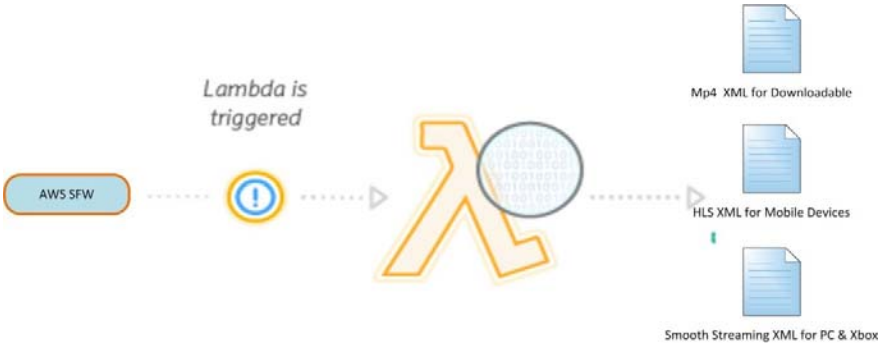

Figure 7 – Metadata Transformation Using AWS Lambda

Publish and Deliver

The final steps of the workflow are publishing and delivery. Processed content is stored in an Amazon S3 object store (origin server). Amazon also offers a global content delivery network called **Amazon CloudFront**. If you store your objects in an Amazon S3 bucket, you can either have your users get your objects directly from S3 or you can configure CloudFront to get your objects from S3 and distribute them to your users.

Using CloudFront can be more cost effective for popular content (most frequently accessed). At higher usage, the price for the CloudFront data transfer is lower than the price for a comparable Amazon S3 data transfer. In addition, downloads are faster with CloudFront than with Amazon S3 alone because your objects are cached closer to your end users (media consumers).

CloudFront has out of the box streaming support for Smooth Streaming, HLS, MPEG-DASH, and RTMP protocols.

Sample content distribution (with offline packaging) is show below:

© 2017 SCTE-ISBE and NCTA. All rights reserved.

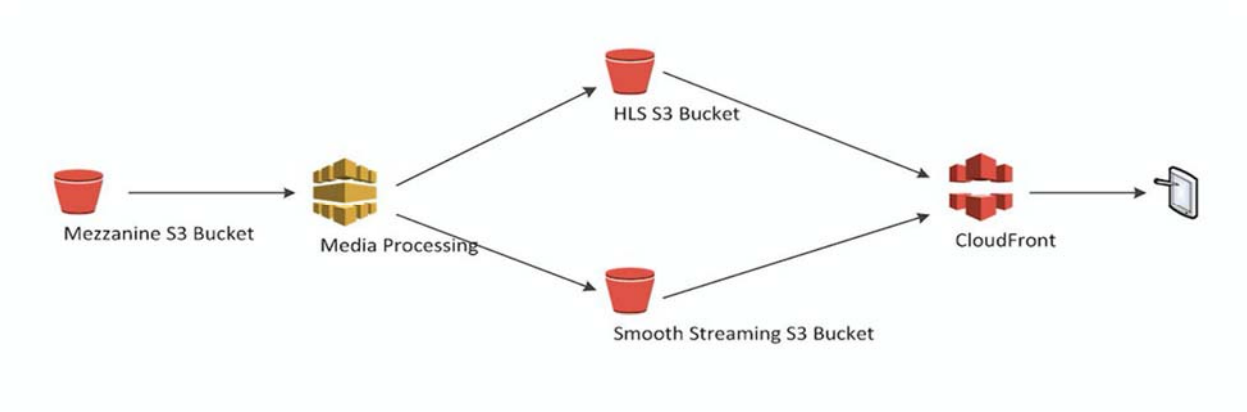

Figure 8 – Content Delivery on AWS Cloud

Conclusion

Broadcasters, content owners and other media providers are looking to take advantage of the capabilities of cloudbased media processing solutions. The primary intent is to expand/extend their existing platforms and create a unified infrastructure that is more flexible, scalable and future proof in its support for video processing and delivery. The main purpose for writing this paper is to assist these organizations with understanding the options and advantages of these technologies as well as with choosing appropriate implementation architecture.

Appendix A: Transcoding TCO

Appendix A compares the total cost of ownership (TCO) between a hypothetical on-premises solution and an AWS cloud-based solution.

Volume

Average number of assets per day	200
Average size of an asset (TV Show or Movie)	90 mins
Average size of an asset (@ 50 Mbps)	33 GB
Average growth per year	15%

On-Premises Solution

Storage

Total storage Year-1	200,000 GB
Total storage cost (Average $1/GB) Year-1	$200,000

Transcoding appliances (Elemental Server)

Number of appliances Year-1	8+8 (redundancy)

© 2017 SCTE-ISBE and NCTA. All rights reserved.

Average server cost	$50,000 per server
Average maintenance cost per server	$4,000 per server
Transcoding cost Year-1	$864,000

Staff

Number of IT employees for maintenance	1
Average IT employee annual cost	$80,000

Cloud Solution (AWS)

Storage (S3)

Total storage Year-1	200,000 GB
Total storage cost (Average $0.35/GB) Year-1	$77,000

Elemental Cloud Transcoding (Built-in Redundancy)

Average transcoding price per hour	$1.55
Total number of hours per year	109,500 hours
Transcoding cost Year-1	$169,725

Staff

Number of IT employees for maintenance	0.5
Average IT employee annual cost	$40,000

*All amounts are estimates only.

Cost Analysis

On-Premises 5-year total cost	$2,723,559
Cloud 5-year cost	$1,933,209

© 2017 SCTE-ISBE and NCTA. All rights reserved.

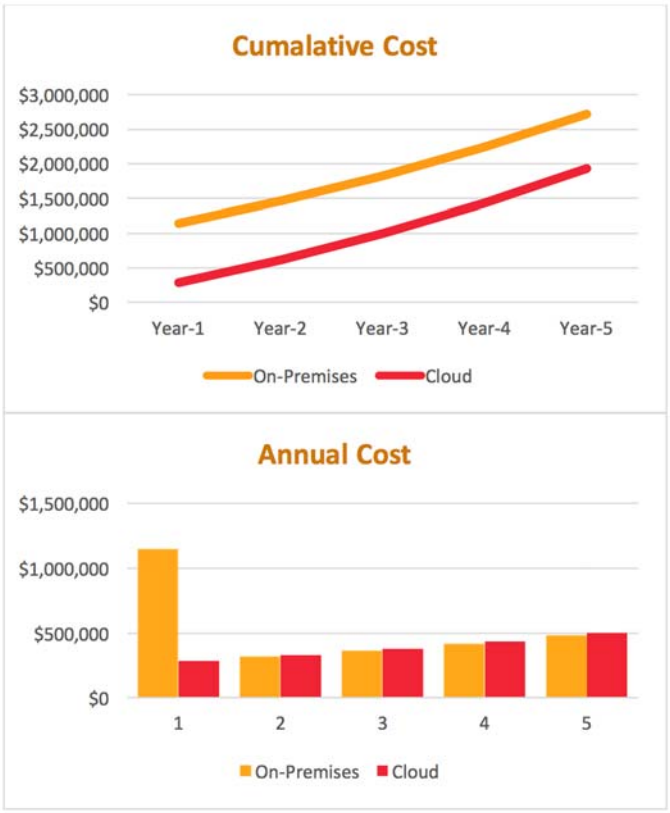

Figure 9 – TCO Analysis

Table 1 – On-Premises vs Cloud Yearly Costs

On-Premises	*Year-1*	*Year-2*	*Year-3*	*Year-4*	*Year-5*
CAPEX	$1,000,000	$152,400	$172,860	$198,789	$228,607
OPEX	$144,000	$165,600	$190,440	$219,006	$251,857
Total	$1,144,000	$318,000	$363,300	$417,795	$480,464
Cloud	**Year-1**	**Year-2**	**Year-3**	**Year-4**	**Year-5**
CAPEX	$0	$0	$0	$0	$0
OPEX	$286,725	$329,734	$379,194	$436,073	$501,484
Total	$286,725	$329,734	$379,194	$436,073	$501,484

© 2017 SCTE-ISBE and NCTA. All rights reserved.

Abbreviations

AWS	Amazon Web Services
DRM	Digital Rights Management
EBS	Elastic Block Storage
EC2	Elastic Compute Cloud
EFS	Elastic File System
HLS	HTTP Live Streaming
PIOPS	Provisioned input/output operations per second
RDS	Relational Database Service
S3	Simple Storage Service
SWF	Simple Workflow Service
TCO	Total Cost Of Ownership
VOD	Video On Demand
XSLT	Extensible Stylesheet Language Transformations

Bibliography & References

Amazon, Documentation for AWS: https://aws.amazon.com/documentation/

© 2017 SCTE-ISBE and NCTA. All rights reserved.