Musify

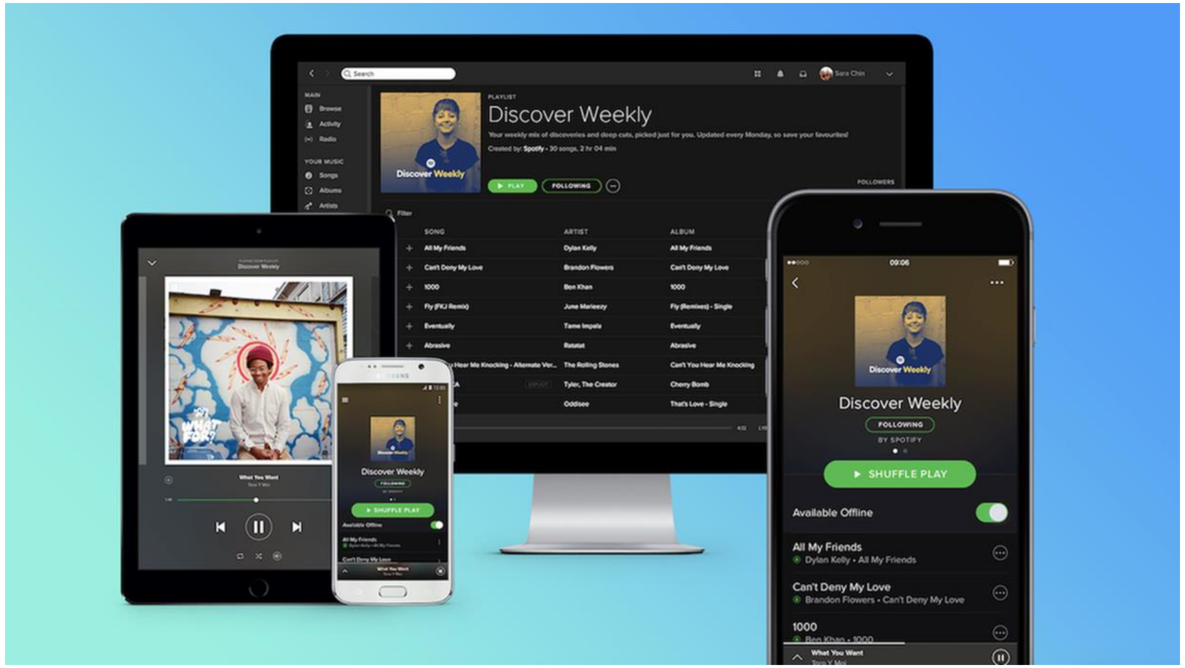
ANALYZE A CLOUD-BASED DATA CENTER ARCHITECTURE

EMIS 7352 – PROJECT 3 – GROUP 3

APOORVA JAIN

MUSIFY REFRESHER

- Musify is a web-based, digital music service providing its subscribers with access to 40,000,000 songs.
- Songs can be digitally streamed at one of 3 bit-rates
  - Low: 112 kbps
  - Standard: 192 kbps
  - High: 240 kbps
- Songs can also be downloaded to the user's device for offline playback
- All songs are Data Rights Management (DRM)-protected (encrypted) to ensure that only authorized users are able to stream or download and play them
  - The System supports permission requests that provide users with the necessary DRM keys to play the songs.



PROJECT 3 OVERVIEW

Musify

- Objectives:
  - Evaluate a Cloud-Based architecture for the Musify problem as the logical evolution from project 1 (centralized, "brick and mortar" architecture) and project 2 (Distributed brick and mortar/mobile/modular architecture)
  - Asses and determine the best architecture alternative
    - Based on total cost, security, reliability/availability, scalability, quality-of-service (QoS), etc.

- Tasks:
  - Evaluate two (2) Cloud Services Providers (CSPs)
    - Migrate the architecture from projects 1 and 2 into the cloud
    - Compare services, costs, models, service level agreements (SLAs), etc.
  - Compare the advantages and disadvantages of centralized vs distributed vs cloud
  - Make a recommendation on which architecture is the best all around alternative

PROJECT 3 OVERVIEW

- Deliverables:
  - Total Cost of the cloud-based architecture
    - Breakdown costs by storage, processing, data transfer (ingress/egress)
  - Service Level Agreements
    - Compare/contrast across the CSPs
  - Discuss the Advantages and Disadvantages
  - Sensitivities in the Architecture
  - Overall Cost Comparison between Projects
  - An overall recommendation

PROJECT 3 OVERVIEW

Musify

- Which CSPs did we evaluate
  - Amazon Web Services (AWS)
  - Google Cloud Platform (GCP)

- Reasons
  - Readily available online documentation and resources
  - Published and detailed pricing structure
  - Similar Service Offerings allow for more direct comparison
  - Relevance to similar "real-life" content delivery businesses like Netflix (AWS) and Spotify (formerly AWS and now GCP)
  - Team familiarity

PROJECT 3 APPROACH

Musify

- Migrate the architectural approach from the brick-and-mortar data center architecture
  - Identify equivalent server instances, storage services, etc. in the CSPs service list
  - Resize based on CSP service specifications (e.g., the server I/O capacities may not align)
- Explore other select alternative services to identify best cost, where it makes sense
  - Examples: alternative storage services, higher capacity server instances
- Did NOT include additional services provided by the CSPs that were not part of project 1 and 2 cost
  - Examples: load balancing, archive, security services, monitoring services, etc.

PROJECT 3 APPROACH

- Did NOT explore (design and cost) other architectural alternatives
  - We do discuss them, however
- Used published discounts and reserve pricing when it benefited (it didn't always)
  - Assumed additional discounts for capacities that exceed their published pricing tables (i.e., there are additional discounts available to high-capacity subscribers but they are not published))
- Determined capacity needs and costs by month (1 through 36)
  - Aligns with "scale-on-demand" advantage of a cloud service-based architecture
  - More accurate costs

PROJECT 3 APPROACH

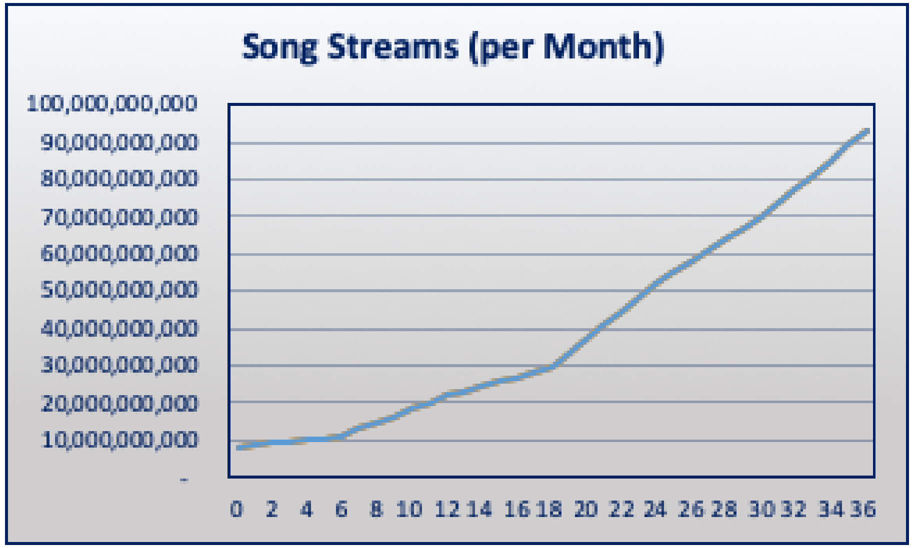
- Labor
  - We assumed a similar staffing profile and cost as was used in Project 2
    - 1 manager, 2 Admins, 2 Techs per shift x 2 shifts
  - Responsible for monitoring and management of cloud services
  - Wages based on Dallas, TX

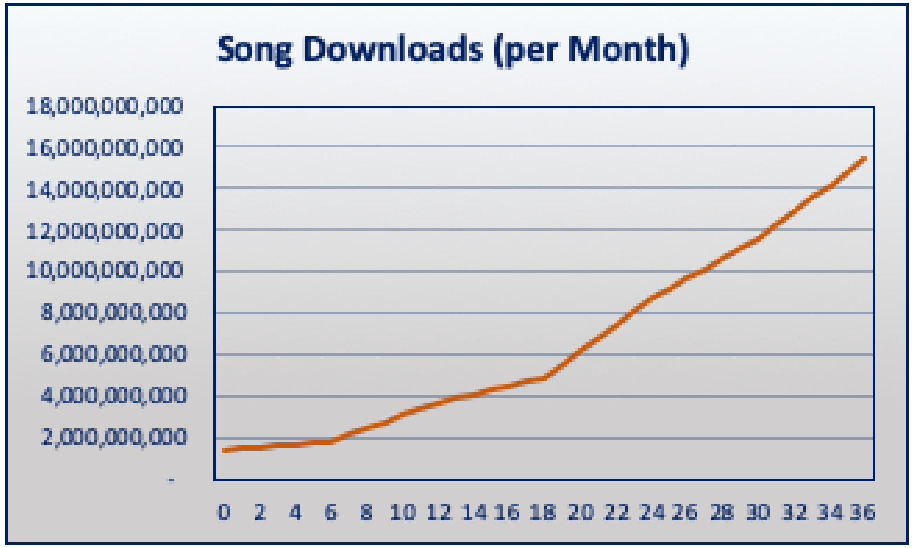
- Facility Space
  - We leased some Tech Office Space to house the Data Center Admin staff
  - Similar approach to pricing as Project 2
  - Used Dallas, TX numbers

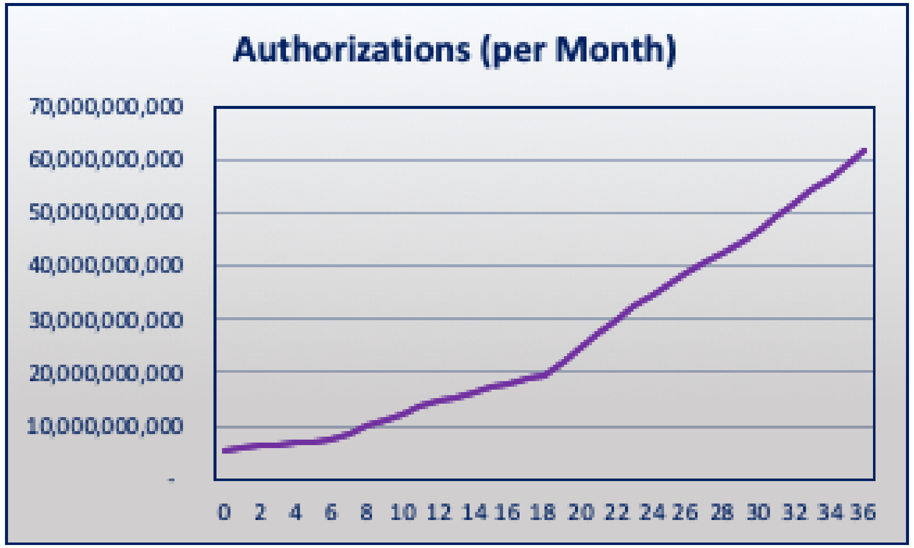
- Labor and Facility Costs included in costs for both AWS and GCP

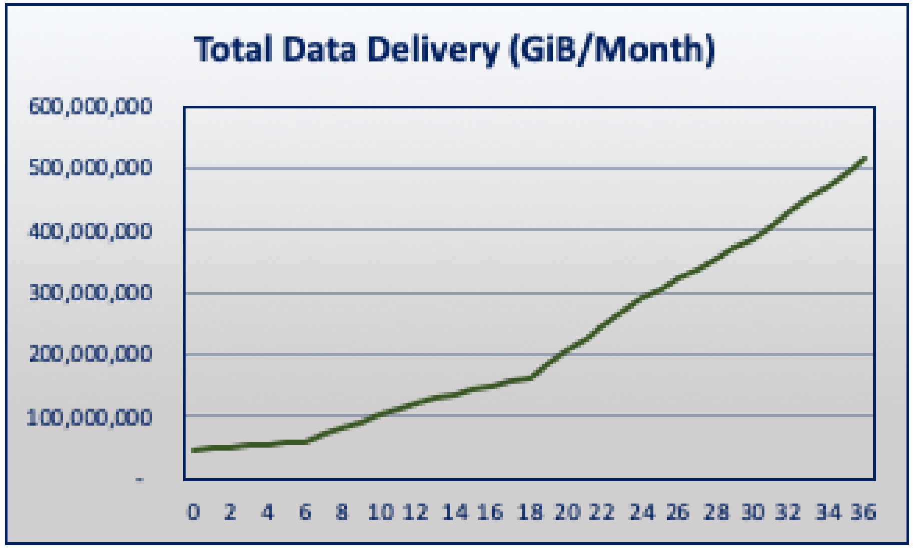
PROJECT 3 VOLUMETRICS

Musify









amazon
web services

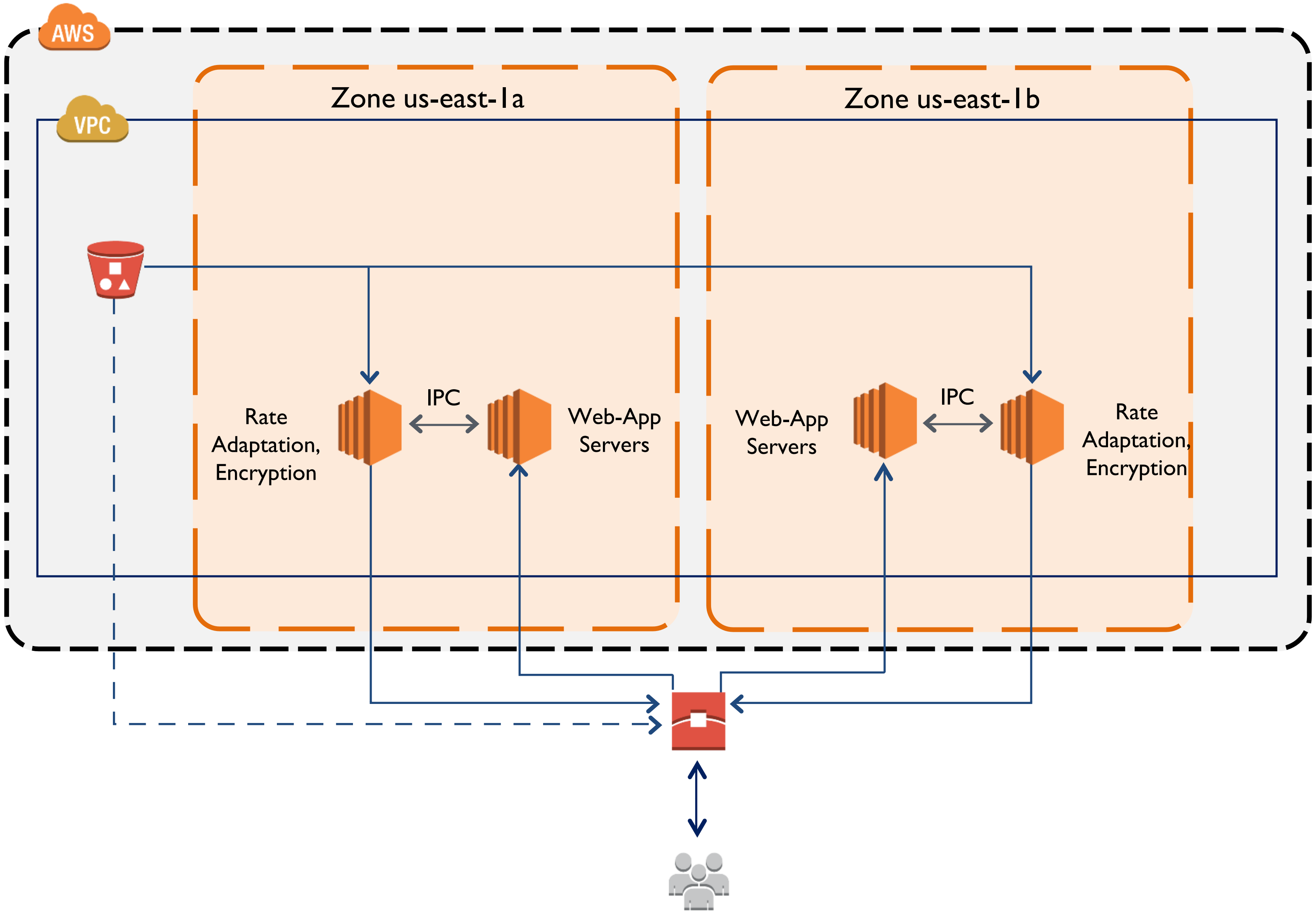
AMAZON WEB SERVICES

Musify

- Amazon EC2: virtual processing servers
  - Web-Application and authorization processing
  - Encryption/Rate Adaptation
- Amazon Elastic Block Store (EBS)
  - Direct-mounted to EC2 instances for local scratch/temp files
- Amazon Simple Storage Service (S3)
  - Object-based storage for song library
- Amazon CloudFront
  - Content Delivery Network provides content cached at edge locations throughout the US for delivery of downloadable and streaming songs

AWS ARCHITECTURE

Musify



Dual Availability Zones in us-east-1 (only region supporting CloudFront)

Each AZ has capacity to support 100% of user requests (on demand)

CloudFront caches songs from S3 to one of 3 caching regions and distributes across 60 edge locations in continental US

AWS COSTS - OVERVIEW

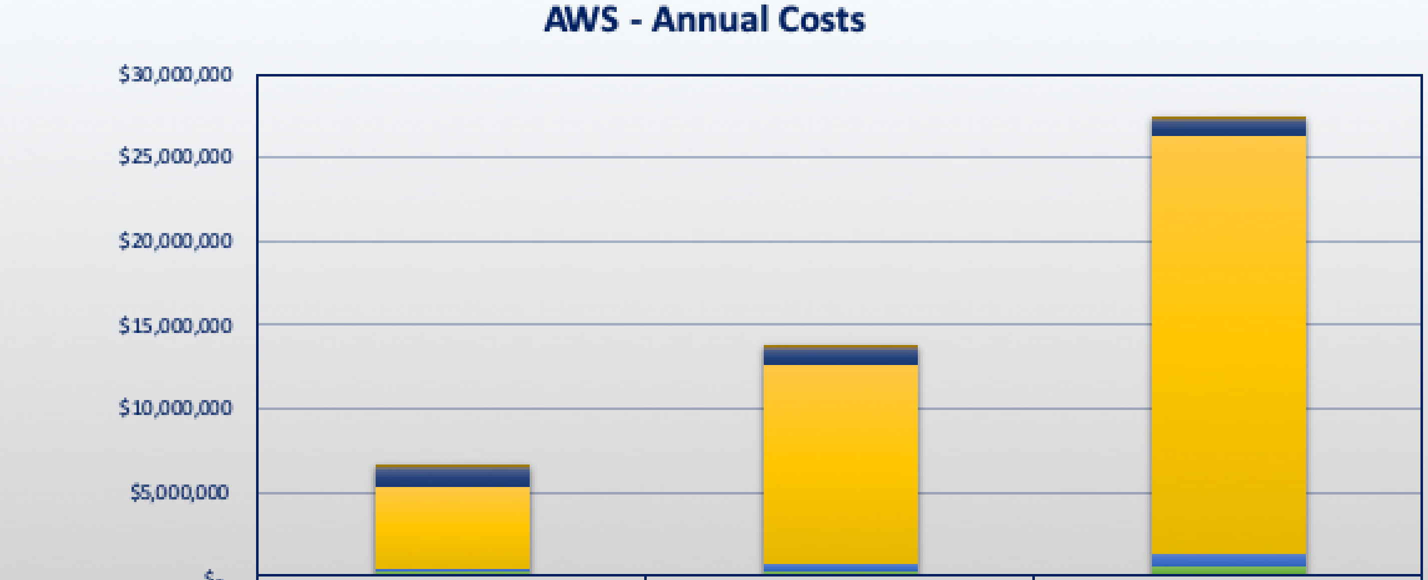
Musify

- Assumptions/Basis
  - S3:
    - Standard S3 Storage
    - Storage of High Bit-Rate and Downloadable Songs only (same as prior projects)
    - No assumed discount (not high volume and static sizing)
  - EC2:
    - Used a1.2xlarge general compute instance
      - Similar in specifications to General XL server in projects 1 and 2 (8 cores, 32 GB memory, etc.)
      - Refactored for available bandwidth (up to 7.5 Gbps usable BW per instance)
      - Used standard 1-year reserve pricing on all instances
    - Included 500GB of EBS per instance for scratch/temp space for processing requests, handling songs, etc.
  - CloudFront
    - Assumed a 75% discount on data egress costs (due to volume); not a published discount
    - Assumed 10% missed caches for pricing purposes

AWS COSTS

Musify

- Web Servers and Rate Adaptation/Encryption Applications Hosted on EC2 instances
- Costs are dominated by CDN Cloudfront CDN (Data Egress charges)
  - Includes an assumed 75% discount due to volume



	Year 1	Year 2	Year 3
Other Expenses	$9,500	$9,500	$9,500
Facility/Capital Expenses	$232,020	$120,600	$120,600
Labor	$972,392	$1,001,564	$1,031,611
Network/Data Egress	$5,028,462	$11,897,525	$24,955,496
Compute Instances	$145,300	$344,544	$722,150
Storage	$224,259	$326,074	$540,349

Total 3-Year Cost: $47.7M

AWS – OTHER OPTIONS CONSIDERED

Musify

- Alternatives Investigated
  - VPC and NAT Gateways (our original architecture
  - Direct User to EC2 and S3 connectivity
- In both cases, data egress costs were prohibitive (2-3x CloudFront, even with similar discount)
- And they didn't provide the lower latency and built-in redundancy afforded through the CloudFront edge locations

AWS SERVICE LEVEL AGREEMENTS

Musify

- SLAs for the services we used/priced (EC2, EBS, S3, CloudFront) follow a similar credit based system based on monthly uptime percentage for the instance, volume, and/or service:

Monthly Uptime %	Service Credit %			
	EC2	EBS	S3	CloudFront
99.0% – < 99.90%	10%	10%	10%	10%
95.0% – < 99%	30%	30%	25%	25%
< 95%	100%	100%	100%	100%

- Subscriber must submit a claim for service credit. (i.e., not automatic)
  - Must include ALL relevant evidence of outage (dates, times, region, resource ids, service, and logs that substantiate the claim)
  - If substantiated, credit issued in one billing cycle
- Exclusions for force majeur, anything originating outside the demarcation point of the service; anything the subscriber or their equipment/software/technology causes

Google Cloud Platform

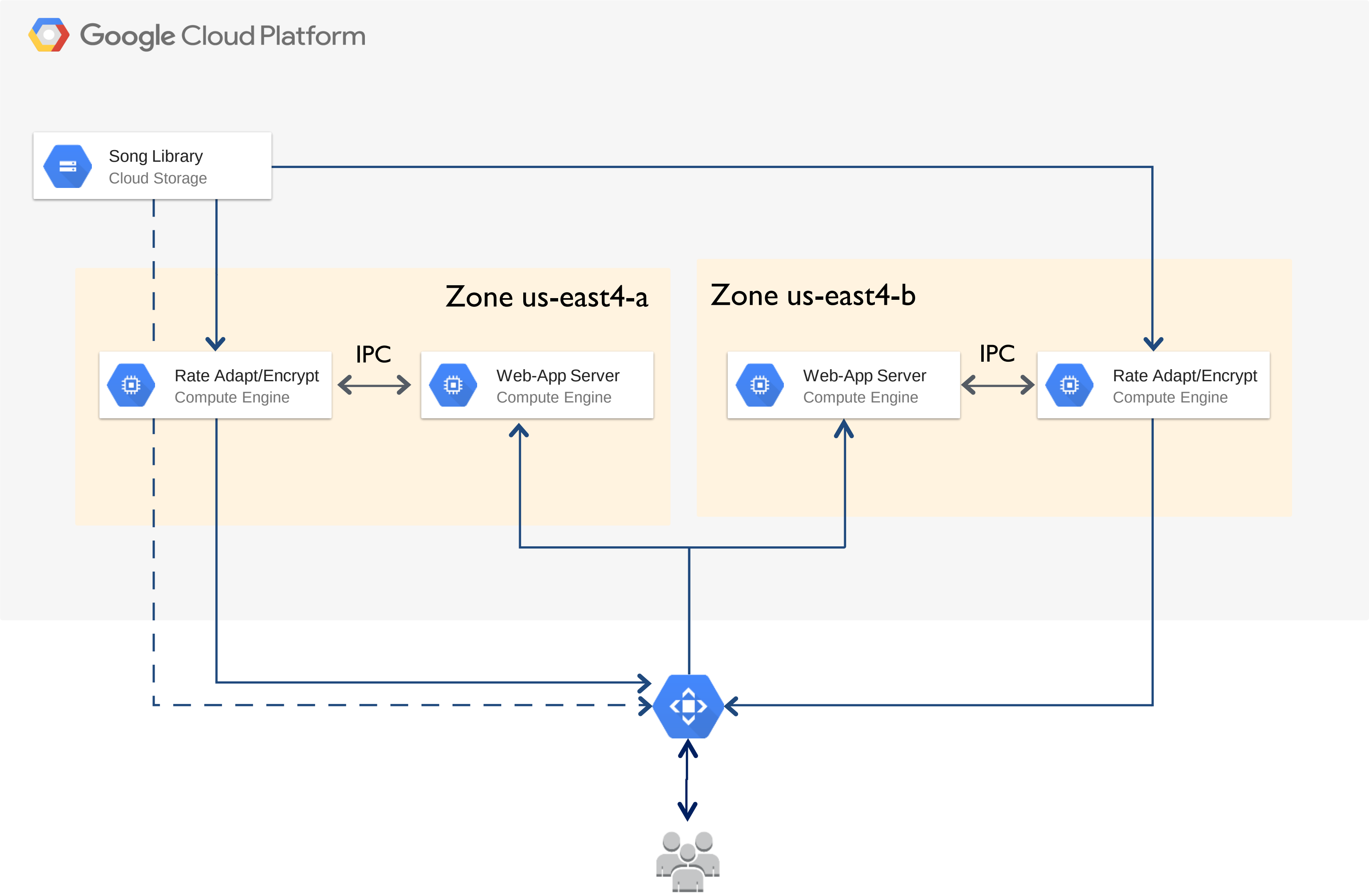
GOOGLE CLOUD PLATFORM

Musify

- Compute Engine
  - Web-Application and authorization processing
  - Encryption/Rate Adaptation
- Cloud Storage
  - Object-based storage for song library
- Cloud CDN
  - Content Delivery Network provides content cached at edge locations throughout the US for delivery of downloadable and streaming songs

GCP ARCHITECTURE

Musify



Dual zones in us-east4

Each zone has capacity to support 100% of user requests (on demand)

Cloud CDN caches songs from Cloud Storage across its US regions and distributes across 23 edge locations in continental US

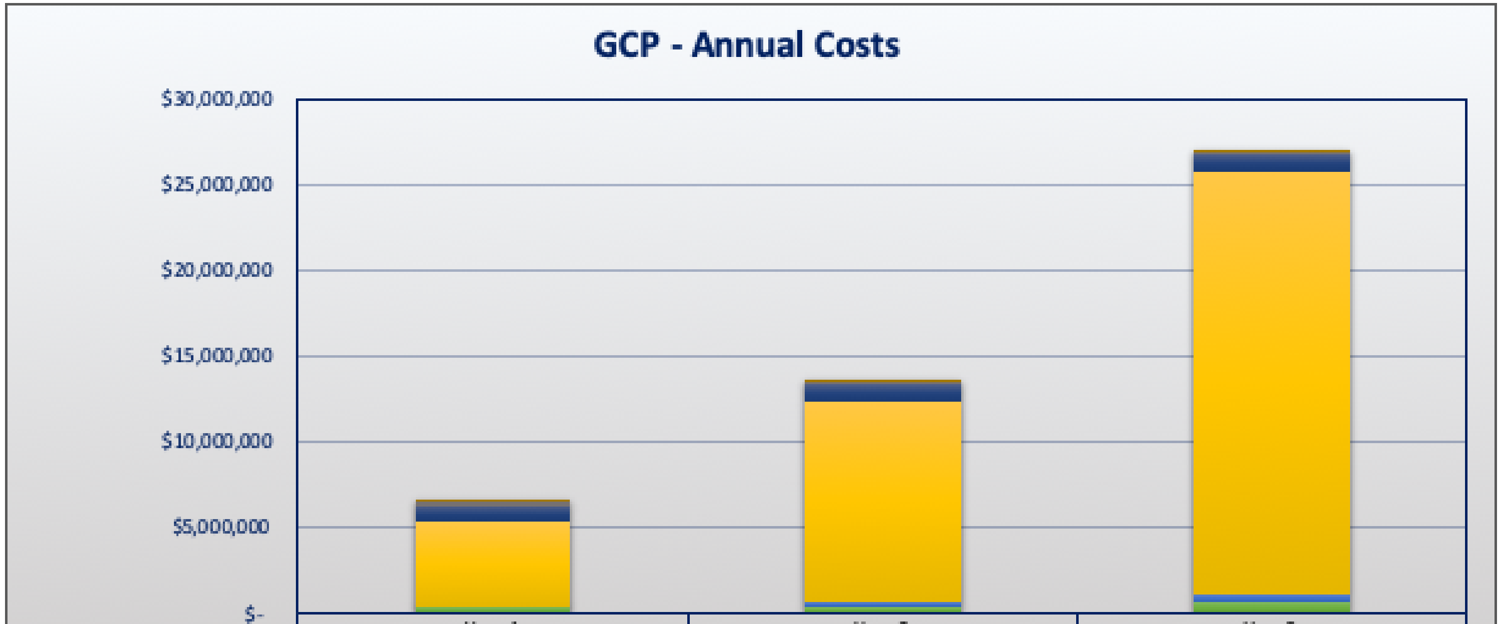
GCP COSTS - OVERVIEW

Musify

- Assumptions/Basis
  - Cloud Storage:
    - Storage of High Bit-Rate and Downloadable Songs only (same as prior projects)
    - No assumed discount (not high volume and static sizing)
  - Compute Engine:
    - Used n1-standard-8 general compute instance
      - Similar in specifications to General XL server in projects 1 and 2 (8 cores, 32 GB memory, etc.)
      - Refactored for available bandwidth (up to 12 Gbps usable BW per instance)
      - Assumed standard on-demand pricing with 25% sustained use discount
    - Included 512 GB of local attached storage per instance for scratch/temp space for processing requests, handling songs, etc.
  - Cloud CDN
    - Assumed a 75% discount on data egress costs (due to volume); not a published discount
    - Assumed 10% (of song library size) per month for missed caches

GCP COSTS

- Web Servers and Rate Adaptation/Encryption Applications Hosted on Compute Engine instances
- Costs are dominated by CDN (Data Egress charges)
  - Includes an assumed 75% discount due to volume



	Year 1	Year 2	Year 3
Other Expenses	$9,500	$9,500	$9,500
Facility/Capital Expenses	$232,020	$120,600	$120,600
Labor	$972,392	$1,001,564	$1,031,611
Data Egress	$4,894,367	$11,681,716	$24,581,556
Compute Instances	$107,596	$252,573	$527,901
Storage	$249,673	$349,532	$563,807

Total 3-Year Cost: $46.7M

GCP SERVICE LEVEL AGREEMENTS

- SLAs for the services we used/priced (Compute Engine, Cloud Storage, Cloud CDN) follow a similar credit based system based on monthly uptime percentage for the instance, volume, and/or service:

Monthly Uptime %	**Service Credit %**		
	Compute Engine	Cloud Storage	Cloud CDN
99.0% – < 99.95% (99.99% for Compute)	10%	10%	10%
95.0% – < 99.0%	25%	25%	25%
< 95%	50%	50%	50%

- Customer must submit a claim for financial credit. (i.e., not automatic)
  - Within 30 days of service loss
  - Must include ALL relevant evidence of outage (dates, times, region, resource ids, service, and logs that substantiate the claim)
- Exclusions for force majeur, alpha/beta services, anything the subscriber or their equipment/software/technology causes

CLOUD ARCHITECTURE WRAP-UP

AWS VS GCP

- Both Cloud Service Providers offer a suite of services that satisfy the Musify use case and directly translate from the on-premises architectures in Projects 1 and 2

- AWS and GCP offer similar services relevant to the Musify architecture (EC2 vs Cloud Engine, S3 vs Cloud Storage, CloudFront vs Cloud CDN)
  - As a result the logical architectures are virtually identical

- Their pricing schedules are similarly structured and competitive (for on-demand)
  - Reserve and Committed subscriptions do vary, but both offer "contact us" (higher, unpublished) discounts for high volume, longer-term commitment

- Service Level Agreements follow a similar approach (monetary credit upon request and confirmation)
  - Google has a maximum credit of 50% per month (AWS has up to 100% credit)

AWS VS GCP

Musify

CSP	Measure		Year 1		Year 2		Year 3		Total
AWS	$ 47,691,445								
	Cloud Service Costs	$	5,398,022	$	12,568,143	$	26,217,995	$	44,184,160
	Storage	$	224,259	$	326,074	$	540,349	$	1,090,682
	Compute Instances	$	145,300	$	344,544	$	722,150	$	1,211,994
	Network/Data Egress	$	5,028,462	$	11,897,525	$	24,955,496	$	41,881,484
	Operating Expenses	$	1,213,912	$	1,131,663	$	1,161,710	$	3,507,285
	Labor	$	972,392	$	1,001,564	$	1,031,611	$	3,005,566
	Facility/Capital Expenses	$	232,020	$	120,600	$	120,600	$	473,220
	Other Expenses	$	9,500	$	9,500	$	9,500	$	28,499
GCP	$ 46,716,007								
	Cloud Service Costs	$	5,251,636	$	12,283,822	$	25,673,264	$	43,208,722
	Storage	$	249,673	$	349,532	$	563,807	$	1,163,013
	Compute Instances	$	107,596	$	252,573	$	527,901	$	888,069
	Network/Data Egress	$	4,894,367	$	11,681,716	$	24,581,556	$	41,157,640
	Operating Expenses	$	1,213,912	$	1,131,663	$	1,161,710	$	3,507,285
	Labor	$	972,392	$	1,001,564	$	1,031,611	$	3,005,566
	Facility/Capital Expenses	$	232,020	$	120,600	$	120,600	$	473,220
	Other Expenses	$	9,500	$	9,500	$	9,500	$	28,499

- Total costs vary by less than $1M over 3 years
- Higher number of edge locations for CDN and more favorable SLA credits favor AWS
- Costs should vary significantly if discount assumptions on data egress are not met

SENSITIVITY ANALYSIS

- As in previous projects, Musify is so heavily driven by content delivery to customers (bandwidth), that other variables have little to no impact on the solution or the sizing
  - In most cases, a change in the variable has a proportional impact on other variables that drive bandwidth (users, # requests, request mix, etc.)

- One variable/assumption that could drive to higher resource/sizing demands would be transactions per request assumptions
  - We have assumed 1 transaction per request thus far and 3,600 TPS/server in determining the number of servers to provision
  - For TPS to overtake bandwidth as the driving sizing factor, the combination of transactions per request and TPS/server would have to increase >1350% in the architecture
    - Even more in the on-premises architectures that require more servers to satisfy bandwidth demands

- The driving cost-reducing assumption in the cloud costs is the 75% discount on data egress
  - Smaller discounts will quickly drive up costs (every 25% discount is another $40M in total cost)

ARCHITECTURE COMPARISON

COMPARISON OVERVIEW

Musify

- Revisit and Discuss Updates to Project 1 and Project 2 Results
  - Error Correction
  - Modifications to align the approaches for equitable comparison
- Compare Costs and Other Factors (PROs and CONs)
- What Would Spotify Do?
- Final Recommendation

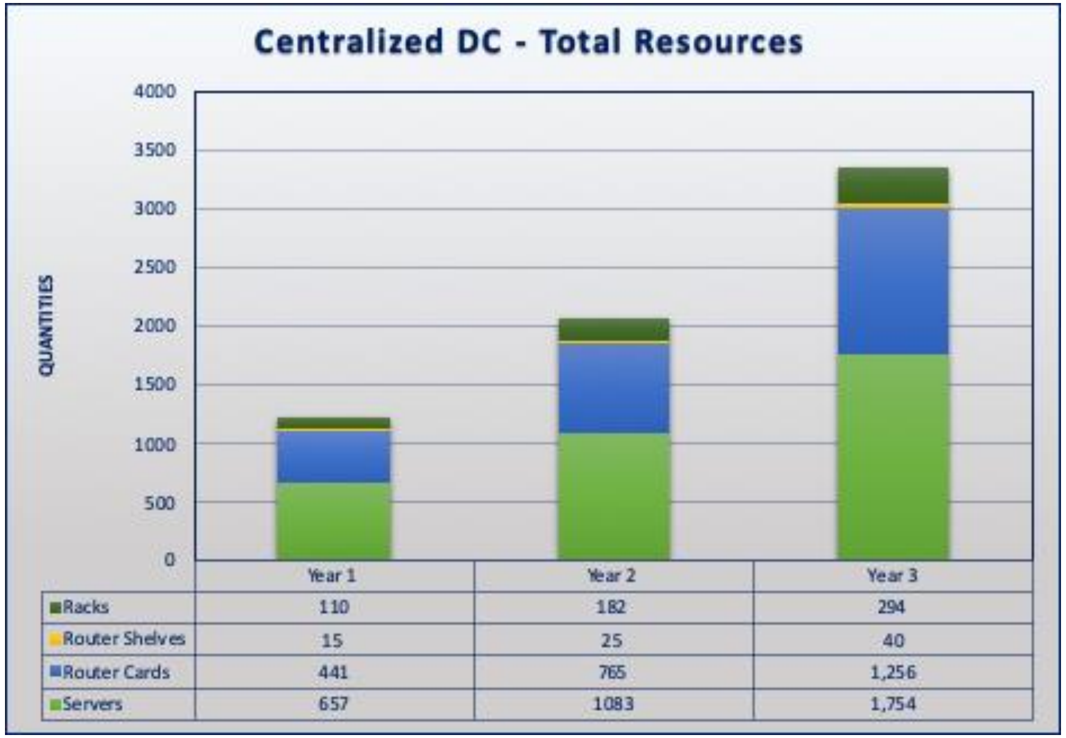
ON-PREMISE VS CLOUD

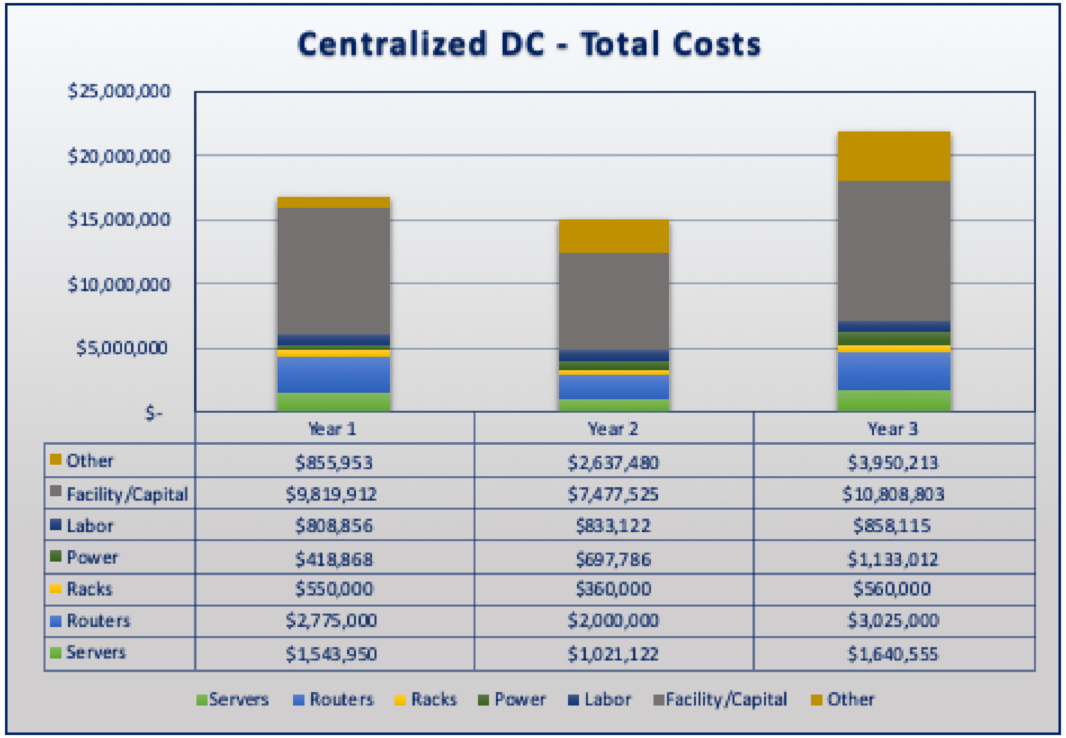
Architecture	Pros	Cons
Cloud	• No capital expense • High availability, backup/archive • Scalability and growth • Pay for only what you need • New/improving services • "Guaranteed" service levels • Tax benefits (Capex vs Opex)	• Potential 3rd party access to data (security vulnerability • Access to data dependent on internet • Potential to lose control of cost if services not used efficiently and monitored • Compliance concerns • Data egress costs can be prohibitive • Network latency
On-Premise	• Physical control over HW • Critical data secure and stored on premises • Not reliant on the internet • Standards/Regulatory compliance	• Needs dedicated IT support • Capital investment • No availability guarantees • Data loss • Scalability potentially costly and time-consuming

PROJECT 1 REVIEW

- Updates
  - In the course of Project 3, an error was discovered in the cost model that did not properly account for the full cost of the Network Attached Storage appliances
    - Once corrected, the cost suggested that another scenario (Scenario 3 – General Servers with Rate Adaptation and Encryption) was a more cost-effective choice
    - This comparison now reflects the revised centralized architecture approach and price
  - In Project 1 we priced out building a custom brick-and-mortar data center; however, in Project 2, we elected to instead lease data center space (not colocation), but did not adjust Project 1 for comparison
    - We have update Project 1 to lease data center space as was done in Project 2
    - We also updated location-dependent variables (wages, space, power, taxes) collected during Project 2 and applied those (we used Dallas, TX as the base for the centralized DC)

PROJECT 1 UPDATED RESULTS





Total 3-Year Cost: $53.78M

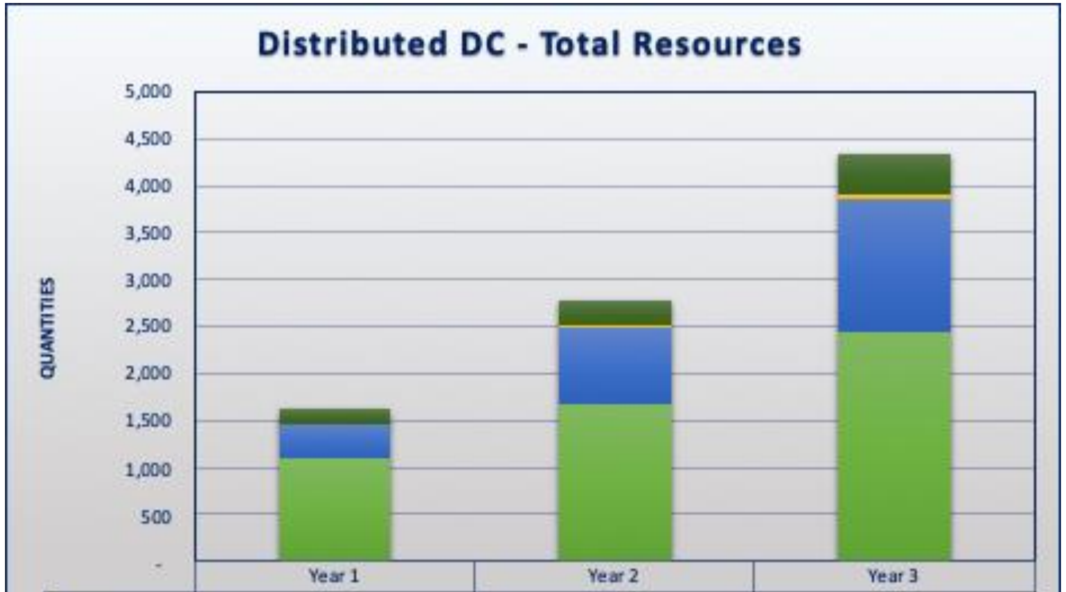
PROJECT 2 REVIEW

Musify

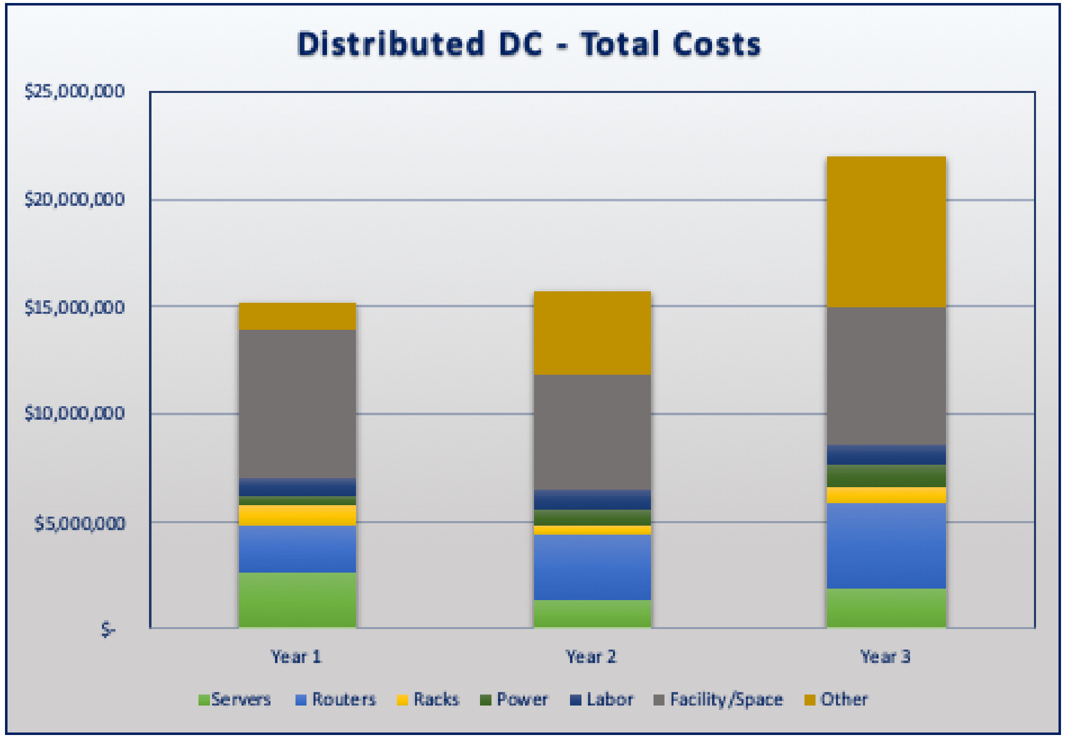
- Updates
  - Project 2 was revised so that the distributed architecture aligned with the updated project 1 architecture (General servers w/ Rate Adaptation/Encryption applications)

  - Project 2 was also updated to reflect our recommendation from the project 2 presentation
    - The "satellite" mobile DCs in the MW and S regions did not appreciably add value or quality of service to the user (i.e., reduced latencies for content delivery
    - They have been removed, so that two distributed data centers in the northeast (N. Virginia) and west (Phoenix) provide all user servers for all regions and are each capable of satisfying 100% of demand in case of data center-level outage

PROJECT 2 UPDATED RESULTS

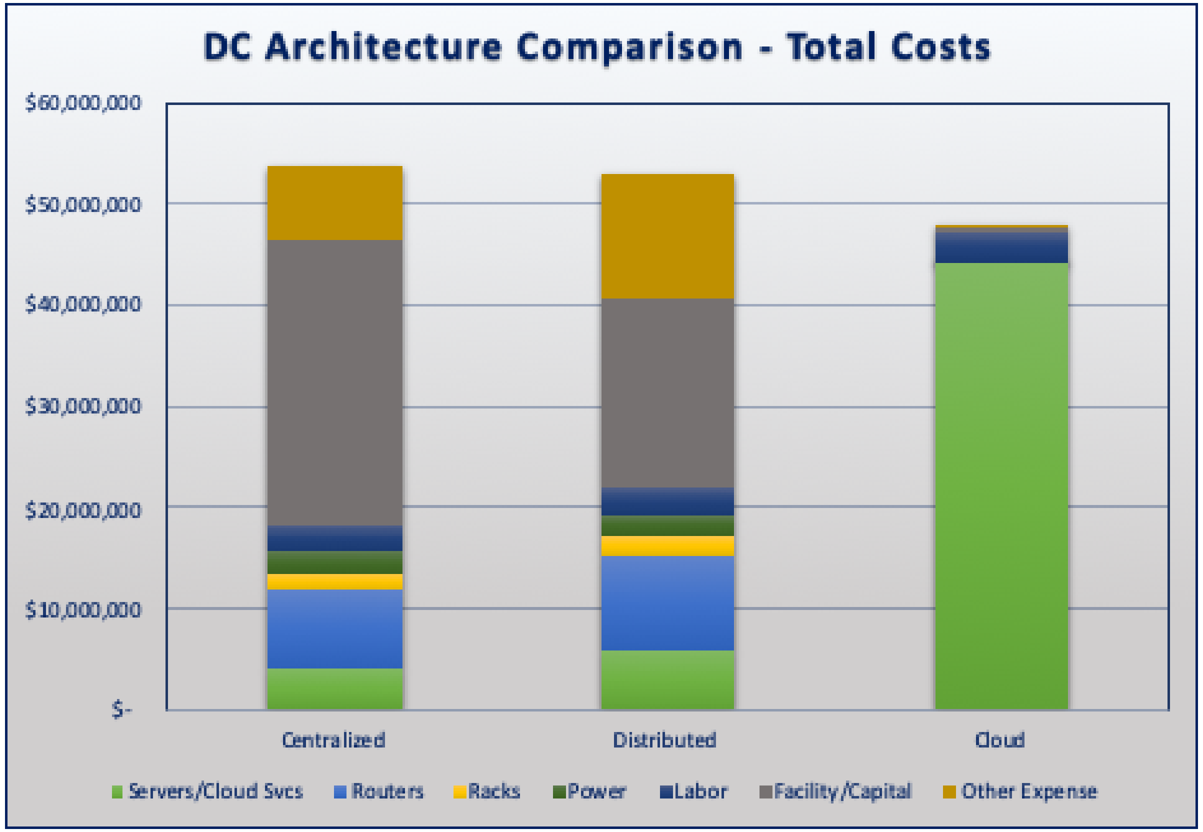


	Year 1	Year 2	Year 3
Racks	174	274	410
Router Shelves	14	32	56
Router Cards	342	806	1,426
Servers	1,106	1,674	2,438



Total 3-Year Cost: $52.97M

ARCHITECTURE "BAKE-OFF"



- At face value, the cloud architecture is the apparent choice at ~$47M; BUT …
  - This cost includes an big, unvalidated assumption about the data egress cost discount (if not met by even a few %s, the cloud cost could easily exceed the other two

ARCHITECTURE "BAKE-OFF"

Musify

Architecture	PROs	CONs
Centralized	• Direct control and access to assets and data • Standards/Regulatory compliance • Self-reliant	• Significant CAPEX and OPEX • Needs dedicated IT support • Scalable, but potentially slow and costly • "One and Done" (can't tolerate a disaster) • Expensive Tier III level system for availability
Distributed	• All the PROs of the Centralized architecture, PLUS … • Improved availability and resiliency to force majeur events • Higher availability but at lower costs (parallel systems) • Improved latency to customer	• All the CONs of the Centralized architecture, PLUS, Remote (flyaway) access to the other location(s) – or potentially extra staff
Cloud	• Minimal capital expense • High availability, backup/archive • On-demand "infinite" scalability and growth • Pay for only what you need • New/improving services • "Guaranteed" service levels • Tax benefits (Capex vs Opex) • Low-latency delivery and • SLAs that credit outages	• Potential 3rd party access to data and other security vulnerabilities • Access to applications and data dependent on internet • Cost management: potential to lose control of cost if services not used efficiently and monitored • Compliance concerns • Data egress costs can be prohibitive • Network latency into your services • SLAs that put it all on you to claim and prove

WHAT WOULD (DID) SPOTIFY DO?

Musify

- They moved from a distributed data center architecture first to AWS (in 2012) then to GCP in 2016/2017
- In Mar 2018, they announced they planned to spend ~$450M on GCP services in the next 3 years
  - Mostly for their big data analytics and other services and development
  - Notably, it's NOT using Google Cloud CDN (but it did use CloudFront with AWS)
- For their content delivery (i.e., the scope of Musify), they currently use a combination of commercial CDNs (Fastly, Verizon, Akamai) and private CDN
  - They use Cloud Storage buckets as their source, however and CDN Interconnect to connect to their CDNs
  - Netflix also developed its own private CDN (installs appliances in ISP data centers for distribution)

FINAL RECOMMENDATION

Musify

- Despite the apparent cost savings, there is significant cost risk in the assumptions associated with the cloud architecture that could easily and significantly escalate cost beyond the on-prem architectures
  - As data transmission costs continue to drop (dramatically), expect that data egress charges will follow suit, but data access and egress is the cloud's bread-and-butter, so it will likely always be the primary cost driver
  - Following a Netflix or Spotify architecture (outsourced/private CDNs) seems the way to go
    - Takes advantage of the cheap compute and storage while avoiding the costly data egress charges.

- The Centralized architecture suffers from being "centralized"
  - Has to have ultra high availability to ensure maximum uptime, but can still suffer from a disaster event without recourse

FINAL RECOMMENDATION

- The Distributed, two-center architecture from Project 2 seems to strike the appropriate balance
  - Better availability, resiliency, and latency than the centralized
  - Better control and access of assets and data, and better security over the cloud
  - Data delivery charges more manageable and deterministic
  - Not as immediately scalable as the cloud, but can be modularized and distributed across more locations

- The Final Jeopardy answer is:

What is the Distributed Data Center architecture?