AWS re:Invent

STG349

Optimize video processing using Amazon FSx for Lustre

Andrew Crudge
Sr. Product Manager
Amazon Web Services

Tushar Saxena
Principal Product Manager
Amazon Web Services

Mac Moore
CEO
Conductor Technologies

AWS re:Invent

aws

Agenda

Amazon FSx for Lustre overview and deep Amazon Simple Storage Service (Amazon S3) integration

Video effects rendering with Amazon FSx for Lustre

Conductor Technologies: Supercharging rendering with Amazon FSx for Lustre

Q&A

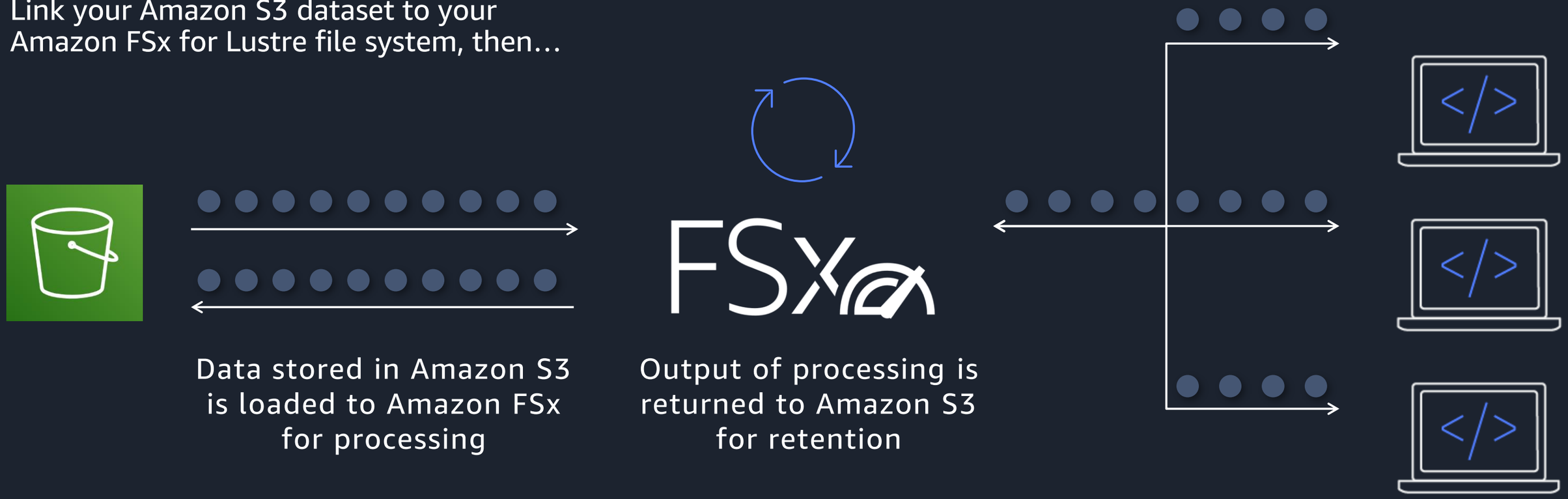
Making your workloads run faster & cheaper with Amazon FSx for Lustre

AWS re:Invent

aws

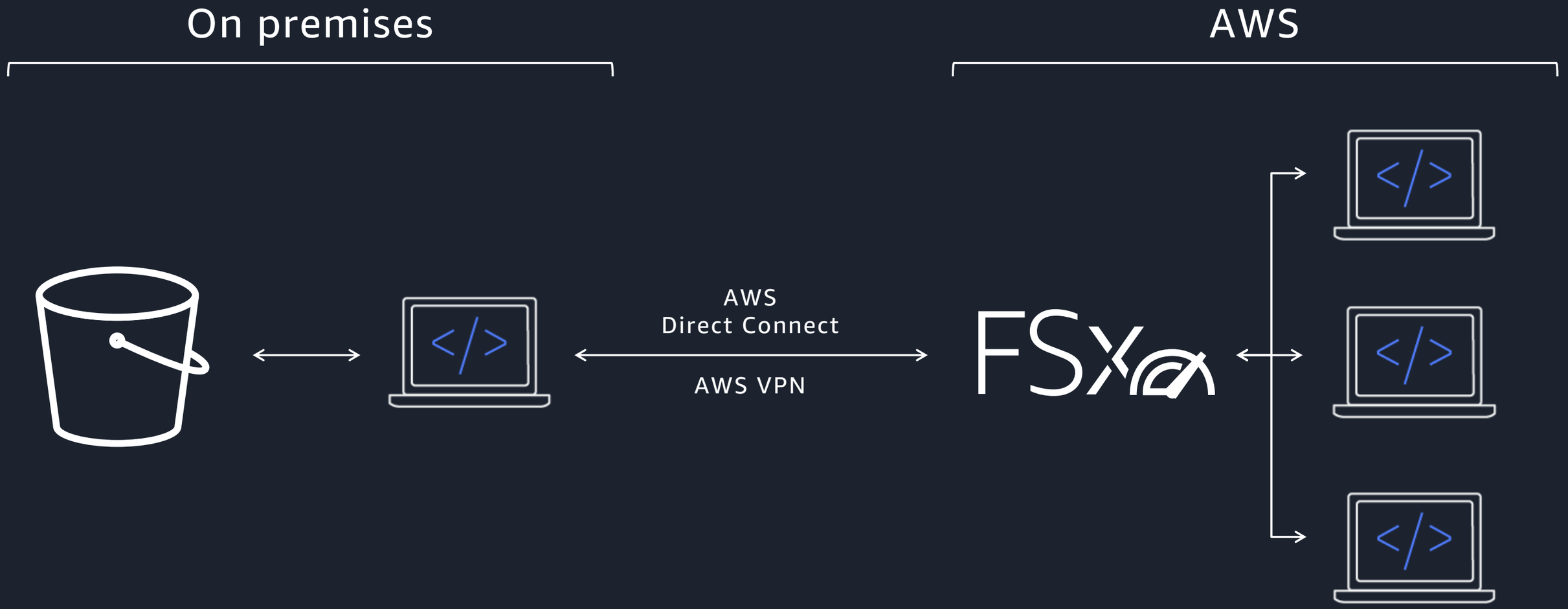
Amazon FSx for Lustre is designed for these data processing workflows in the cloud

Link your Amazon S3 dataset to your Amazon FSx for Lustre file system, then…



When your workload finishes, simply delete your file system, or keep it running for long-lived workloads

Amazon FSx for Lustre also supports cloud bursting for on-premises data repos



Amazon FSx for Lustre: Making your compute workloads faster and cheaper

Fast processing with 100+ GB/s throughput & sub-millisecond latencies

Tight integration with Amazon S3

Flexible data processing options for both short and long term

Lower TCO

Example workloads

High-performance computing

Machine learning

Media rendering and transcoding

Big data analytics

Electronic design automation

Financial modeling

Oil and gas seismic processing

Autonomous systems training

Amazon S3 and Amazon FSx for Lustre

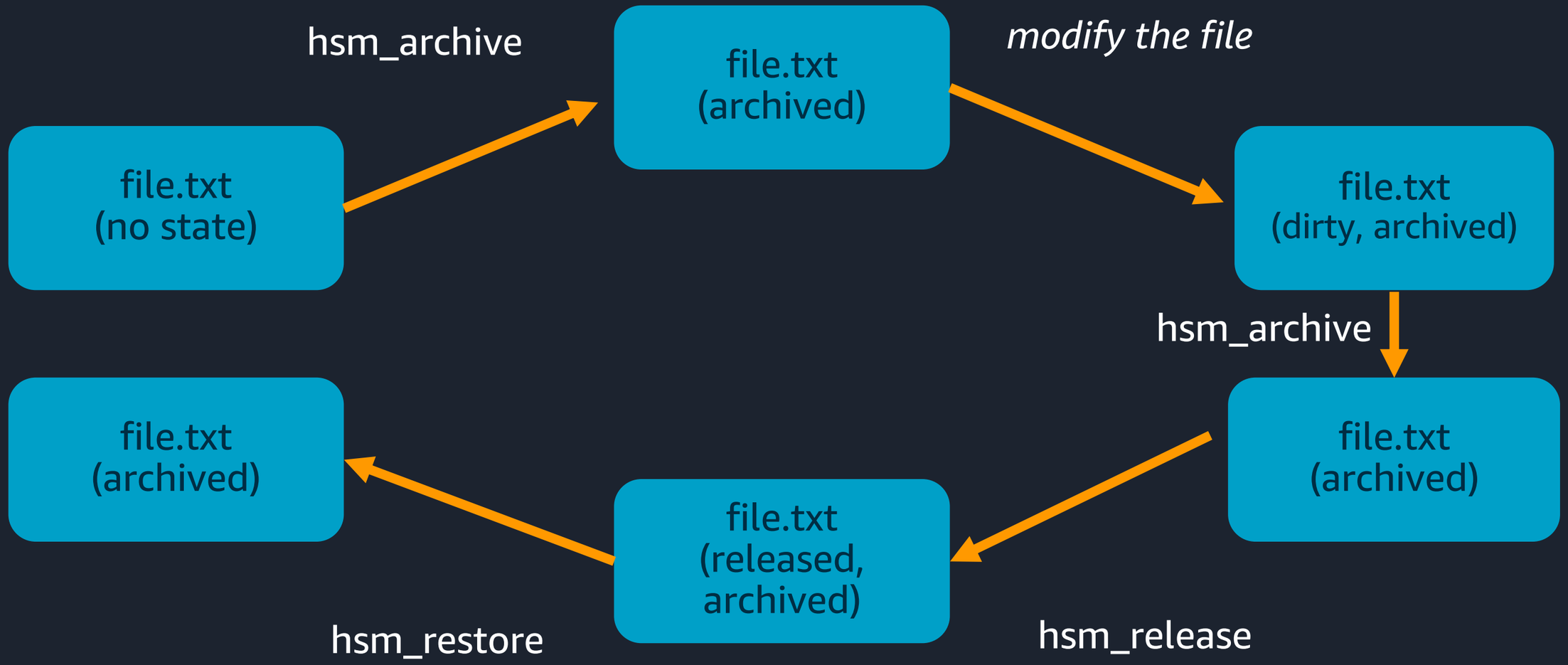
Specify an Amazon S3 bucket when you create your file system

On accessing the file system, you see all objects as files/directories

Files are moved in real time, when accessed, to the file system

Use commands to write results to Amazon S3

The life cycle of a file



Scratch file system performance

Each TB of storage provides
200 MB/s of baseline throughput,
and up to 2x burst throughput

File systems can scale
to hundreds of GB/s and
millions of IOPS

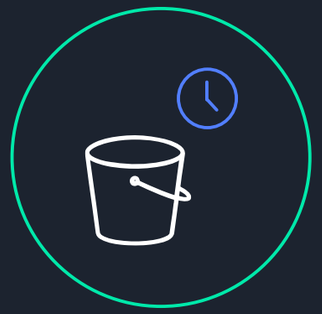
Capacity	Baseline throughput	Burst throughput
1TB	200 MB/s	up to 400 MB/s
10TB	2 GB/s	up to 4 GB/s
50TB	10 GB/s	up to 20 GB/s
100TB	20 GB/s	up to 40 GB/s
1PB	200 GB/s	Up to 400 GB/s

Amazon FSx for Lustre is cost-optimized for data processing

Optimized for processing datasets at high performance and low cost

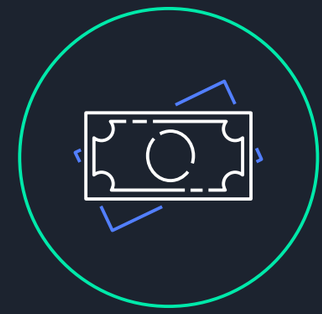
Cost-effective

Process only a portion of your dataset at one time



Long-term data stored in low-cost Amazon S3 or on premises

Spin up, spin down with scratch file systems or release with self-healing



No request costs for repeatedly accessed data

Amazon FSx for Lustre pricing (high-perf SSD): $0.14 per GB-month ($0.20 per TB-hour)

Video effects rendering with Amazon FSx for Lustre

AWS re:Invent

aws

Transcoding with Amazon FSx for Lustre

Example: You deliver on-demand movies and television content across multiple platforms over the internet

Problem: Transcoding mezzanine files to multiple formats, bitrates and resolutions, at low cost, quickly

Solution:

- Spin up high-performance Amazon FSx for Lustre file system
- Copy mezzanine files to Amazon FSx
- Spin up Spot Instances, mount FSx, run transcoder
- Write output formats to FSx
- Copy outputs to your CDN origin server
- Spin down Amazon FSx and Spot Instances

Live video rendering with Amazon FSx for Lustre

Example: A file-based application processes high-resolution video feeds from multiple cameras to create a 3D model so that video can be rendered from any perspective

Requirements: Latency-sensitive (<1ms), throughput-sensitive (>50 GB/s per event)

Solution:

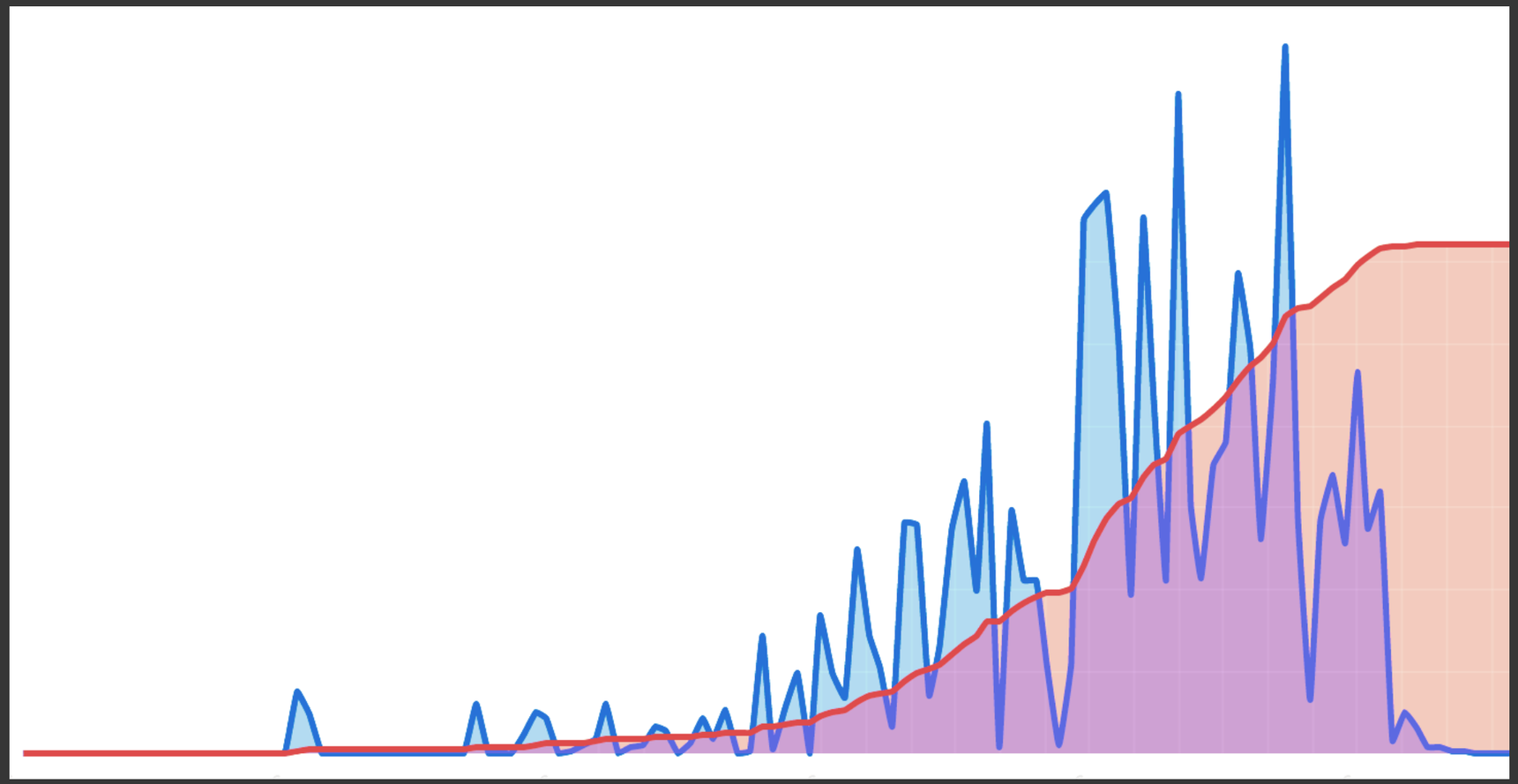
- Write the raw video frames (45 MB) to a shared FSx file system from live streams
- Read the frames from G3 instances running video rendering software
- Write intermediate and final rendered video frames to FSx

Conductor Technologies: Supercharging video rendering with Amazon FSx for Lustre

AWS re:Invent

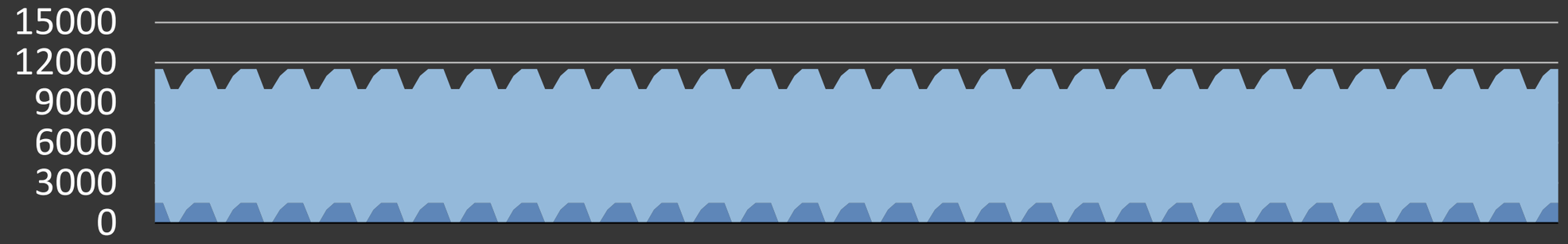
aws

Origins: Typical Project Cycles



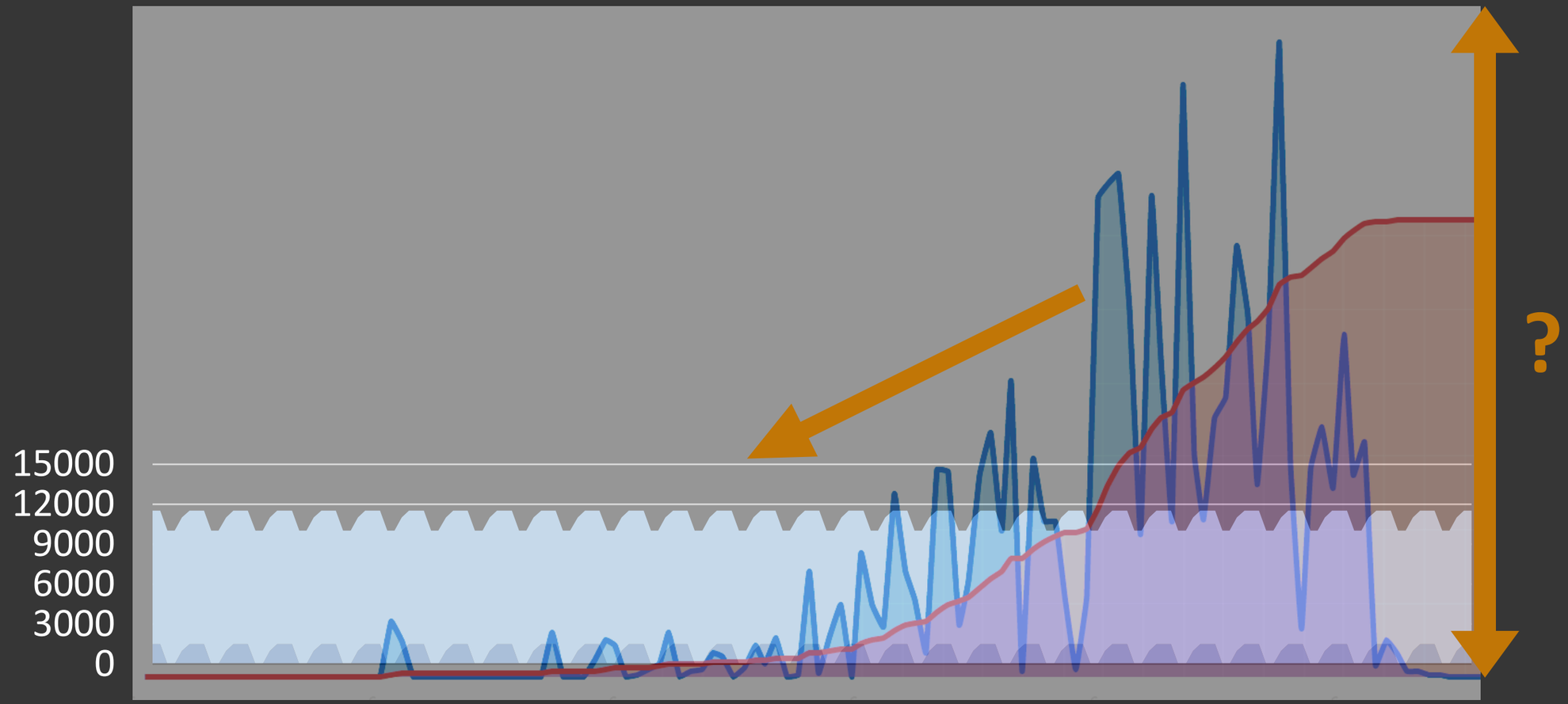
Actual rendering profile of a production (blue line = capacity | red line = spend)

Problem: Real Rendering vs. Constrained Resources



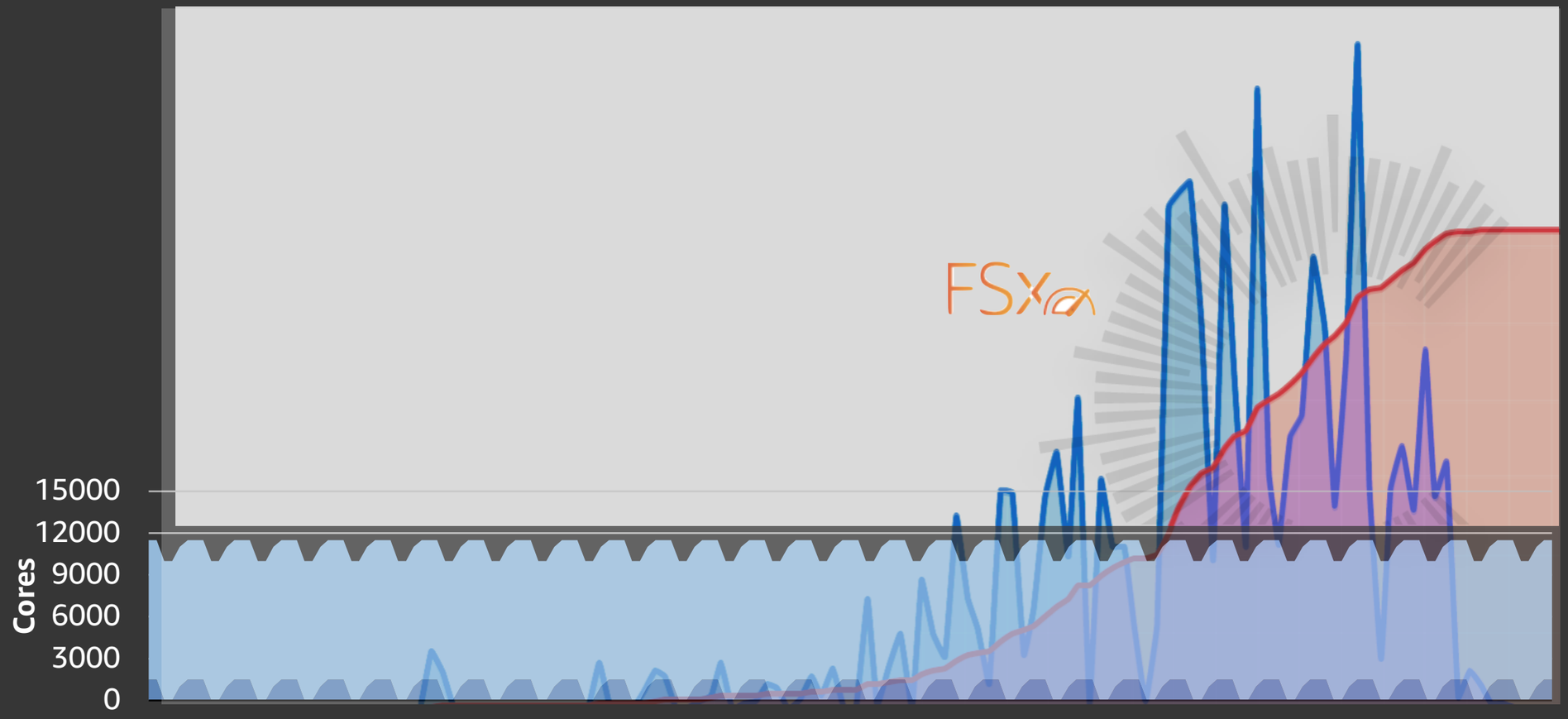
Hypothetical on-premises infrastructure capacity, typical for a growing company

Problem: Real Rendering vs. Constrained Resources



Hypothetical on-premises infrastructure capacity, typical for a growing company

Conductor Peak Shaving in the Cloud



Elasticity is key to a hybrid rendering workflow, and
Amazon FSx for Lustre performant bursting

CONDUCTOR Architecture

Availability?
Exhaustion?
Retarget?

Render job in Container on AWS Batch

Local | AWS

Submit job

CONDUCTOR

Application

- Un-synced data?
- Upload from local
- Determine farm size
- Data sync to FSx for Lustre
- Launch farm
- Monitor job progress
- Log task outputs
- Download to local
- Cleanup

Create/Maintain/Kill Instances with AWS Batch

aws

Job Info/Metrics
(databases)

Artists follow normal workflow

Local File Server

Transfer Daemon(s)

Upload Files

Download Files

Raw Logs
(streaming)

Output Files

Input Files

Data Transfer

Amazon S3 Storage

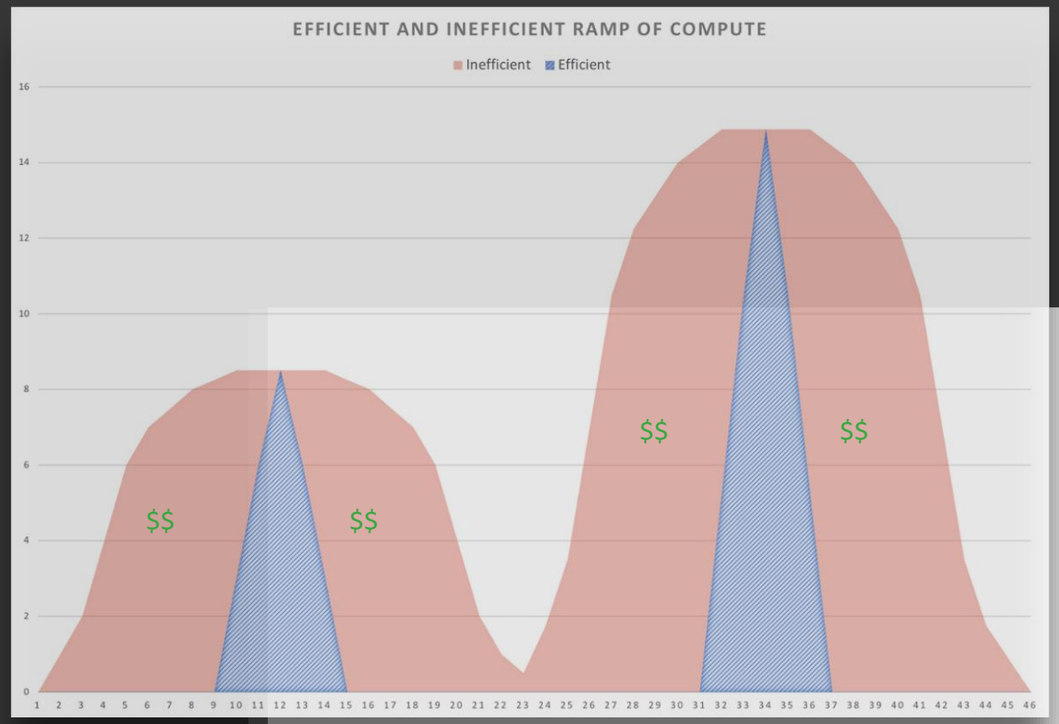
FSx

Analytics Monitoring

CONDUCTOR

User Dashboards

Why Amazon FSx for Lustre?



Need to minimize uptime for cost efficiency
Need to maximize I/O performance

CONDUCTOR

Stage Two – Intelligent Rendering

Data-Driven Optimization
Machine Learning Focused

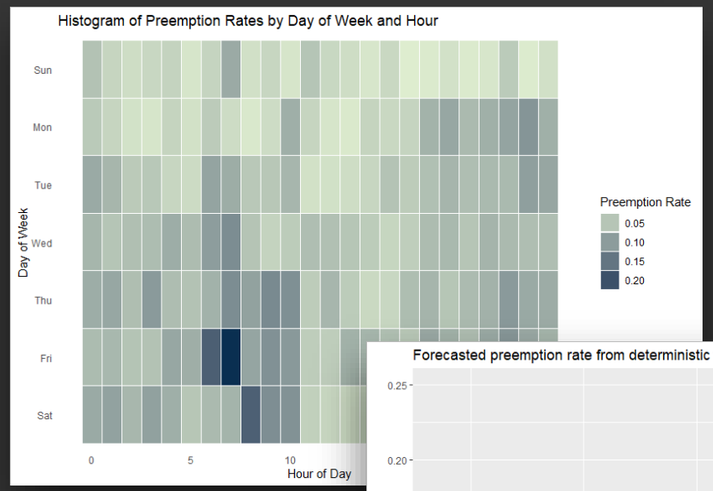
Inefficiency identification
Optimize resource use via data analytics
Automating visual validation process
Object/asset classification

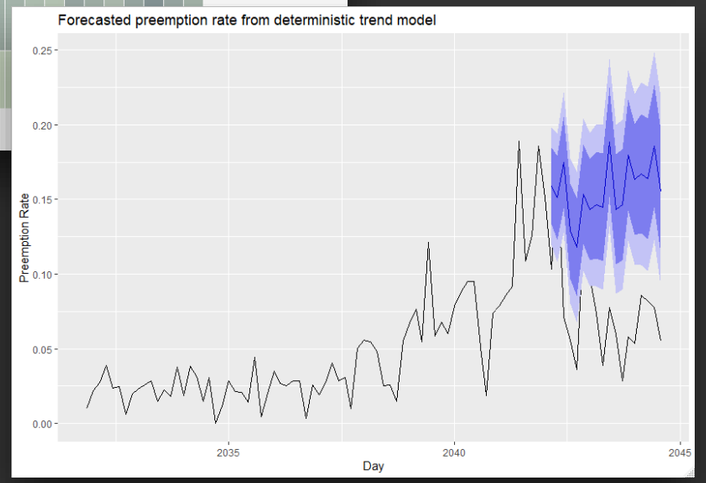
Upfront Cost Estimates
Leveraging Historical Data

Scene Complexity
Pricing Analytics
Fixed-Price Bidding Assist

Data Science & Predictive Analytics

Leveraging our 300M core hours of historical data
Understanding seasonality
Determining most economical availability
Job shifting/orchestration, upfront cost





Cloud TCO – Calculating Baseline

TCO Calculator - Cost Calculations

Monthly Core Hours Required	333,333
Monthly Capacity	600,000
Duration Between Projects	0.00
Wall clock time (months)	1.67
In-house Compute Subtotal	$ 93,600.00
In-house Storage Subtotal	$ 10,000.00
In-House Power Subtotal	$ 30,773.44
In-House Facilities Subtotal	$ 19,218.75
In-house Sys-Admin Subtotal	$ 180,000.00
In-house Render Wr Subtotal	$ -
Annual Software Cost	$ 14,312.50
Upfront Software Cost	$ 31,843.75
Annual Networking cost	$ 94,937.11
Annual Network Maintenance	$ 18,987.42
Total Annual in-House Cost	$ 493,672.97
Annual Std Conductor Cost	$ 109,246.23
Annual PreE Conductor Cost	$ 75,496.23

Exceeding In-House Farm Analysis	
Rental Space Penalty	$ 4,804.69
Software rental penalty	$ 11,539.06
Artist productivity penalty	$ 41,666.67
Core rental penalty	$ 23,400.00

Monthly

Hard Cost Assumptions

	Electricity	$/KwH/Core
4 kwh for 1,280 cores	Powering	$ 0.0078125
6 kwh for 1,280 cores	Cooling	$ 0.0039063
	Cost @ $0.10/kwh	$ 0.0011719

Compute Costs	
$262k for 1,280 cores aceca.com	HW Cost per Core-hr $ 0.0106849
	Depreciation Length (months) 1
$2,050 Monthly Colo for 1,280 Cores	**Facility/Colo**
	Monthly Space Cost/Core 1.6015625
	Storage
	DAS/SAN-HDD (cost/GB) $ 1.00

Networking	
Percentage of Total SW/HW Spend	25%
Support Percentage of Networking Spend	20%

UPS	
Based on $3.5k for 4kW	Cost/Core/Month $ 20.51

Support & Artist Assumptions	
Artist FTE	$ 125,000

Support Personnel	$/Core/hr
Loaded FTE	$ 180,000.00
Sys Admin	$ 0.02054795
Render Wrangler	$ -

Cost References (only change if known)	
Cores	2000
Purchase Price	$ 13,000
Power (kwh)	15.625
Cooling (kwh)	7.8125
Kwh Cost	$ 0.10

2048 cores for 25kW
.3 to .5 of power to cooling
have a think on per rack power

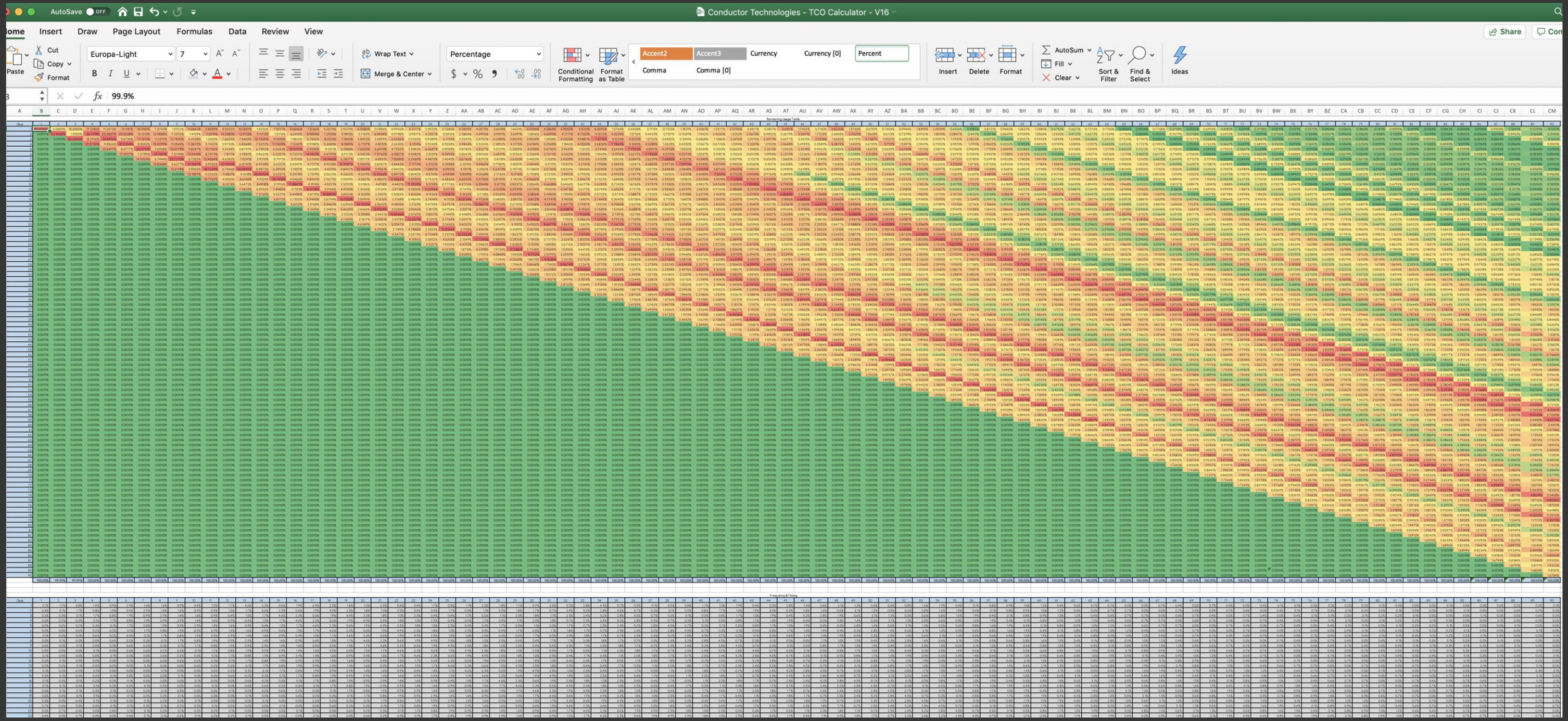
Decision tables
yes
no
on
off
1-Year
3-Year

Software	License	Subscription	Conductor
OS	$ -	$ -	$ -
VRay	$ 240.00	$ -	$ 0.0050
Katana Render Nodes	$ -	$ -	$ -
Tractor Licenses	$ 300.00	$ 60.00	$ -
Nuke Render Nodes	$ 479.00	$ 398.00	$ 0.0082
Arnold Render Nodes	$ -		$ -
Maya	$ -		$ -
Houdini Engine	$ -		$ -

	Standard	Premptible	Storage/month	Networking Overhead	Cloud Discount	Conductor Discount	YoY Cost Reduction
Conductor/core/hr	$ 0.0688	$ 0.0350	$ 0.1800	5%	10%	20%	0%

						Conductor Pricing			Hybrid Pricing		
						0%	100%	100%	0%	100%	100%
	Day	Required Render Hours	Capped Uncompleted	In-House Capacity	Actual Rendered Hours	Daily In-house Cost	Daily Conductor Standard Cost	Daily Conductor Preemptible Cost	Blended Conductor Cost	Hybrid Model (Standard)	Hybrid Model (Preemptible)	Blended Hybrid Cost
	1	9,018	-	20,000	9,018	$ 1,353	$ 788	$ 483	$ 483.34	$ -	$ -	$ -
	2	2,224	-	20,000	2,224	$ 1,353	$ 194	$ 119	$ 119.18	$ -	$ -	$ -
	3	2,419	-	20,000	2,419	$ 1,353	$ 211	$ 130	$ 129.63	$ -	$ -	$ -
	4	3,771	-	20,000	3,771	$ 1,353	$ 329	$ 202	$ 202.13	$ -	$ -	$ -

Cloud TCO – Modeling Cyclicality



Cloud TCO – Summary

Total Cost of Ownership (TCO) Dashboard

In-house Cores	1,000
Avg. Core Hours/project	500,000
# of Projects/Quarter	2
% Farm Utilization	39.8%
Storage per project (GB)	5,000
Avg Project Duration (days)	90
Over Capacity Premium	200%

Pro-rated vs Full Capacity Analysis		
on	Render days/month	30
	Render hours/day	20
	% Availability	82.2%

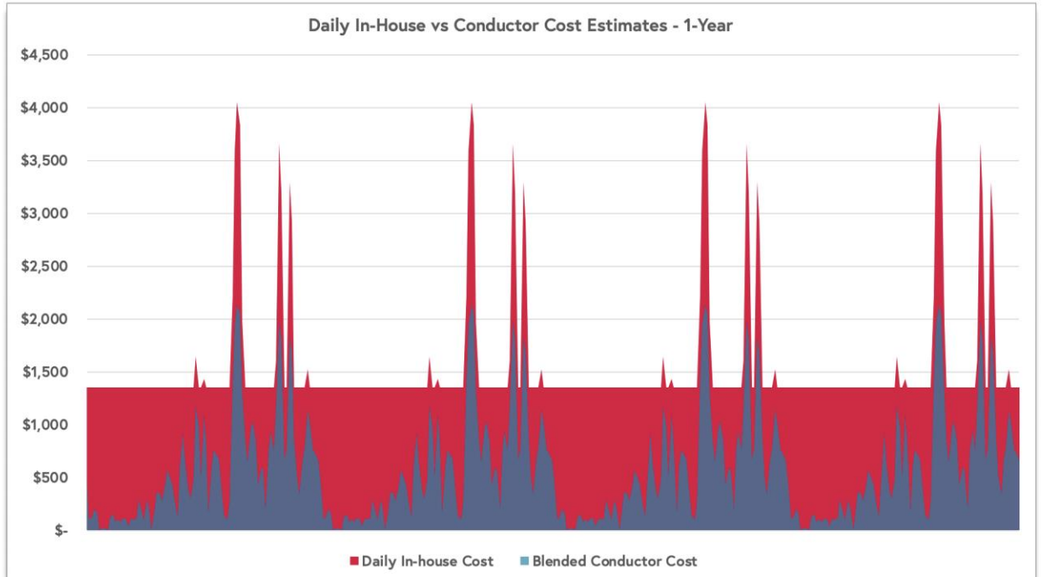
Support Personnel	
Cores per resource	1,000
Sys Admin (Per xx cores)	yes
Render Wrangler (Per xx cores)	no
# of Artists/project	40
Wasted Productivity	10%

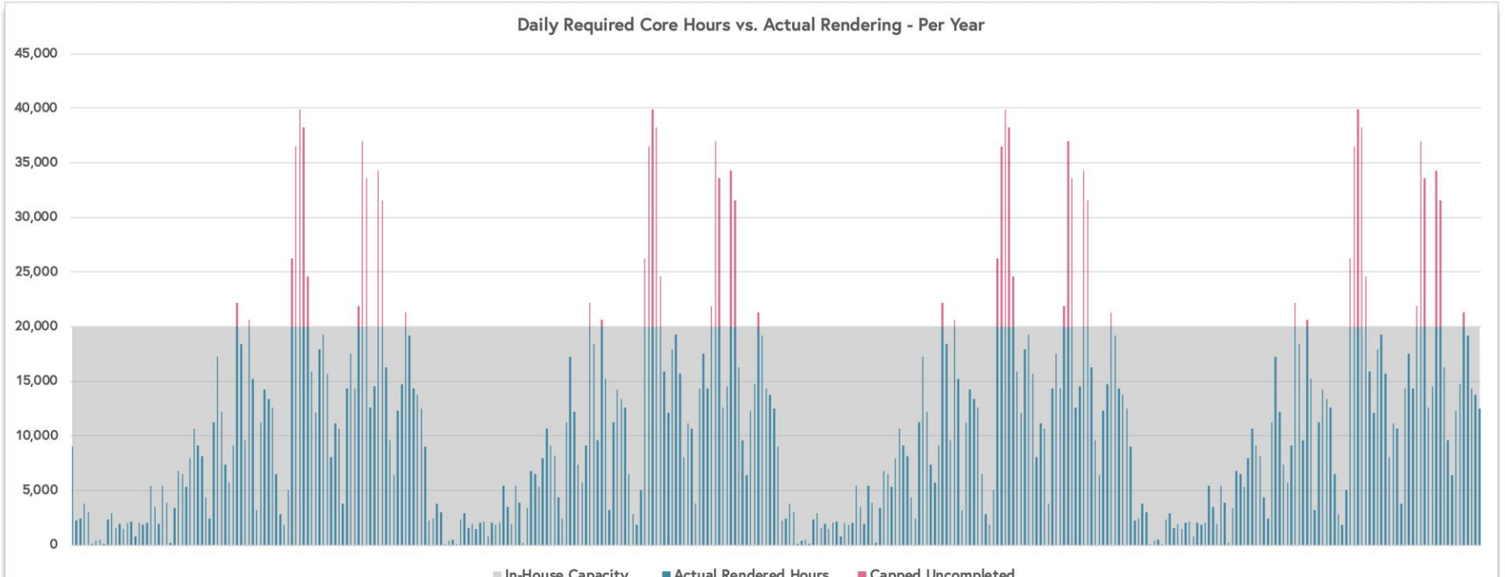
% Mix of Standard vs Preemptible	
Standard VMs	0%
Low Priority (Preemptible) VMs	100%
Cyclicality Modeling	on

Software	
OS	no
VRay	yes
Katana Render Nodes	no
Tractor Licenses	yes
Nuke Render Nodes	yes
Arnold Render Nodes	no
Maya	no
Houdini Engine	no

In-House Cost/Core-Hour	$	0.14
Conductor Cost/Core-Hour	$	0.05
Total In-House Cost	$	556,354
Total Conductor Cost	$	214,400
% Annual Savings		61%

CONDUCTOR





Thank you!

AWS re:Invent

aws

Please complete the session survey in the mobile app.