AWS

SUMMIT

Get the Most Bang for Your Buck #EC2 #Winning

Joshua Burgin
General Manager, EC2 Spot
Amazon Web Services

June 28, 2017

amazon web services

Amazon EC2 purchasing options

On-Demand

Pay for compute capacity by the hour with no long-term commitments

For spiky workloads, or to define needs

Reserved

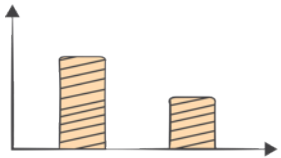
Make a 1 or 3 year commitment and receive a significant discount over on-demand

For committed or baseline utilization

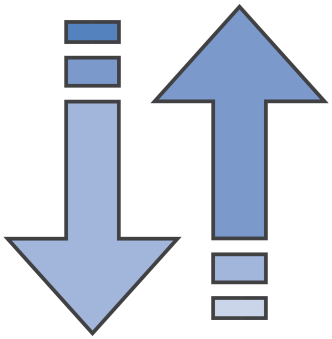
Spot

Pay market price for unused compute capacity at a steep discount over on-demand

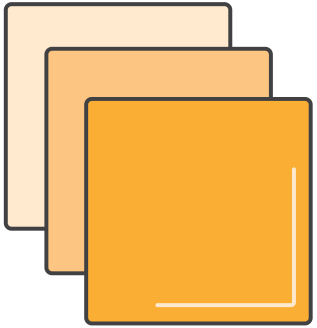
For fault tolerant, time-insensitive or transient workloads



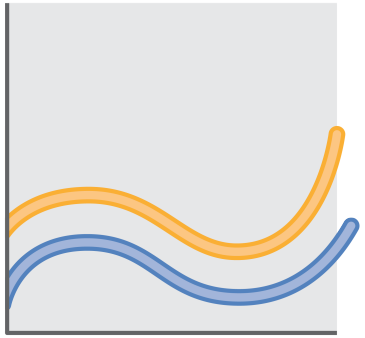
Pillars of performance and cost-optimization



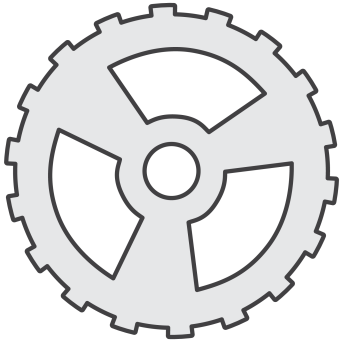
Right sizing



Purchasing options

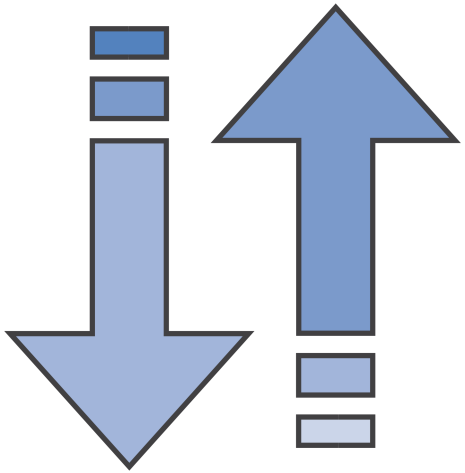


Increase elasticity



Measure, monitor, & improve

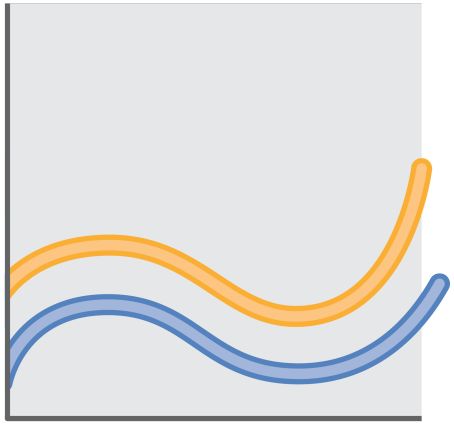
Right sizing



Right sizing

- Selecting the cheapest instance available while meeting performance requirements
- Looks at CPU, RAM, storage, and network utilization to identify potential instances that can be downsized

Increase elasticity



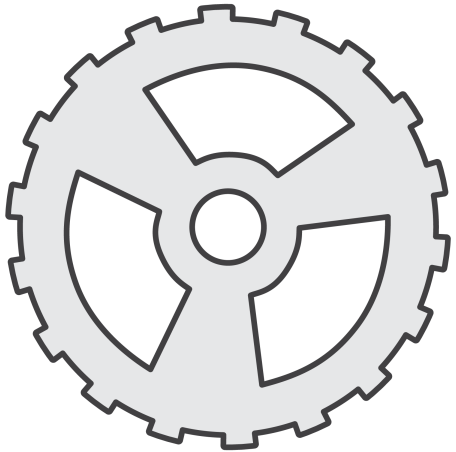
Turn off non-production instances

- Look for dev/test, non-prod instances that are running always-on and turn off

Auto-scale production

- Use Auto Scaling to scale up and down based on demand and usage (for example, spikes)

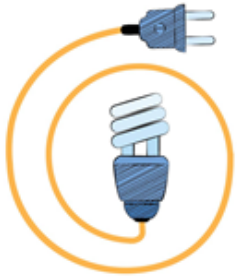
Measure, monitor, and improve: Uncover the cost-optimization opportunities



Auto-tag resources
Identify always-on non-prod
Identify instances to downsize
Recommend RIs to purchase
Dashboard our status
Report on savings

AWS pricing principles

No up-front investment



Pay as you go

Pay less when you reserve

Pay less when AWS grows

Novartis: Acceleration of pre-clinical R&D

> We completed the equivalent of 39 years of computational chemistry in just under 9 hours for a cost of around $4200.

Steve Litster
Global Head of Scientific Computing, Novartis

- Existing infrastructure to screen 10 million compounds in a computational model not available
- New infrastructure would have cost approximately $40 million to build

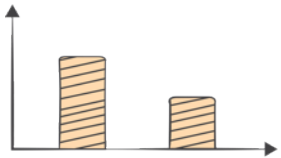
Novartis used AWS for HPC computational chemistry

Amazon EC2 purchasing options

On-Demand

Reserved

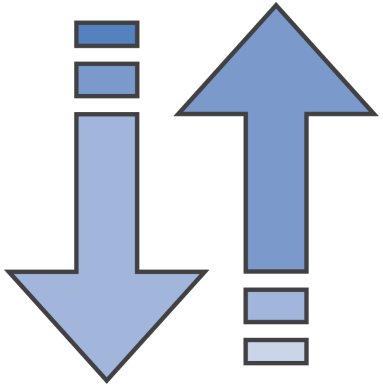
Spot



EC2 On-Demand pricing

Low cost and flexible

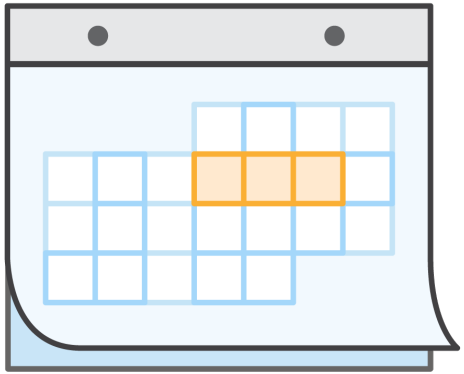
Develop and test



Short-term, spiky and unpredictable

EC2 Reserved pricing

Steady state



Optional capacity reservation

New

Upfront payments to reduce costs

Reserved Instance Overview

Determining the best RI option for your needs

	Standard	Convertible
Savings Potential	Up to 75%	Up to 56%
Change Instance families, operating system and tenancy	No	Yes
Lease Terms	1 or 3 Years	3 Years Only
Change AZ, Instance size (Linux), networking type	Yes	Yes
Payment	No Upfront Partial Upfront All Upfront	No Upfront Partial Upfront All Upfront
Regional benefit?	Yes	Yes

Convertible Reserved Instances

With a Convertible RI you **can**:

Convert to a new instance family e.g. R3 to C3 to T2 to M4

Convert to a new instance price e.g. if AWS reduces the public rate of your instances

Convert to a new operating system e.g. Windows to Linux

Convert to a new instance size

Convert tenancy e.g. from Dedicated Instances to default

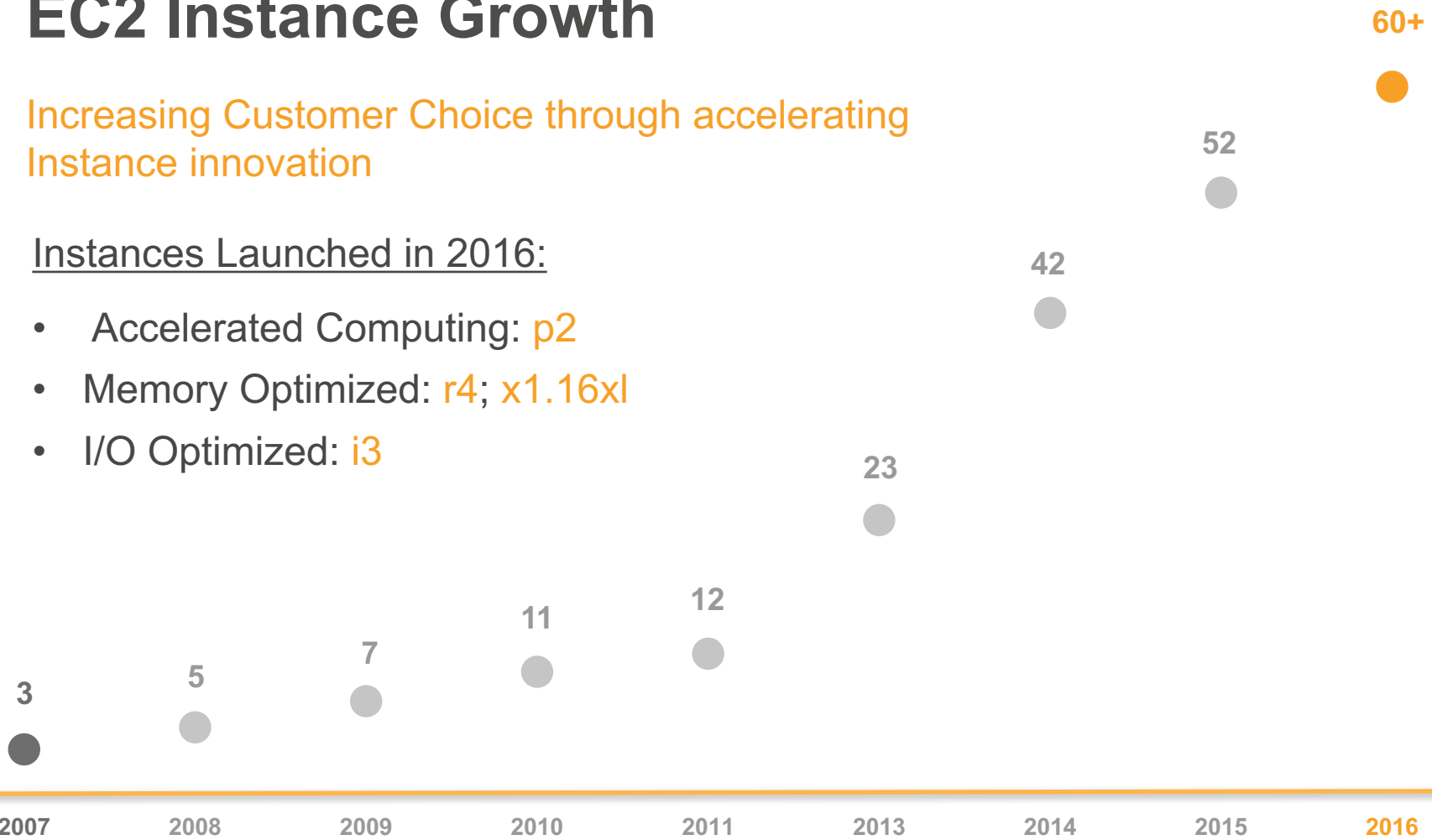
Convert to a different payment option e.g. No Upfront to Partial Upfront

EC2 Instance Growth

Increasing Customer Choice through accelerating Instance innovation

Instances Launched in 2016:

- Accelerated Computing: p2
- Memory Optimized: r4; x1.16xl
- I/O Optimized: i3



Reserved Instance Overview

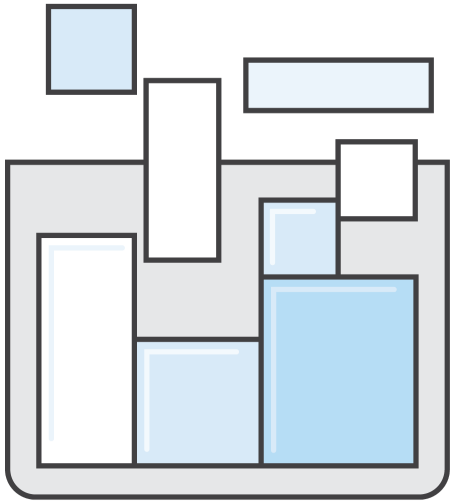
Determining the best RI option for your needs

	Standard	Convertible
Savings Potential	Up to 75%	Up to 56%
Change Instance families, operating system and tenancy	No	Yes
Lease Terms	1 or 3 Years	3 Years Only
Change AZ, Instance size (Linux), networking type	Yes	Yes
Payment	No Upfront Partial Upfront All Upfront	No Upfront Partial Upfront All Upfront
Regional benefit?	**Yes**	**Yes**

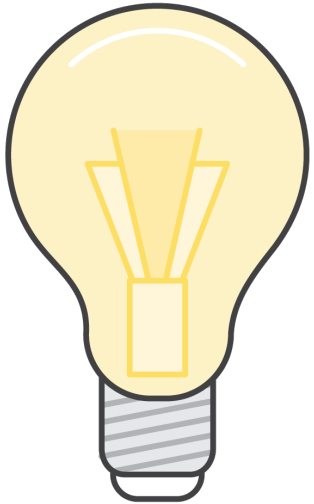
Regional Benefit Overview

	Capacity Reservation Benefit?	Automatic discount across AZs?	Automatic discount across instances sizes?	Sellable on the RI Marketplace?
Zonal	Yes	No	No	Yes
Regional	No	Yes	Yes	No

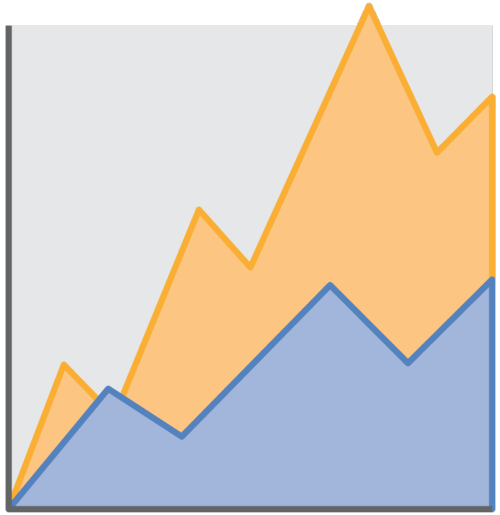
EC2 Spot pricing



Time or instance flexible



Experiment and/or build cost-sensitive businesses



Users with urgent computing needs or large amounts of additional capacity

Spot Instance details

90% savings!*

Options

- Spot fleet to maintain instance availability
- Spot block durations (1-6 hours) for workloads that must run continuously

Commitment level

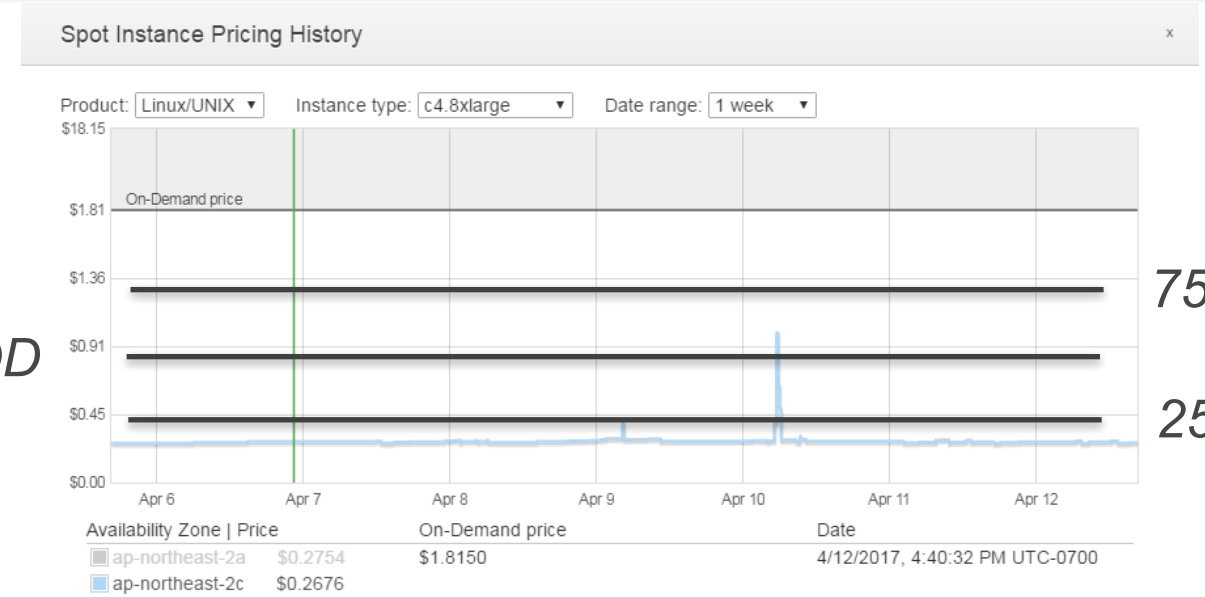
- None

* Compared to On-Demand price based on specific EC2 instance type, region, and Availability Zone

Spot Rules

Markets where the price of compute changes based on supply and demand

You'll never pay more than your bid



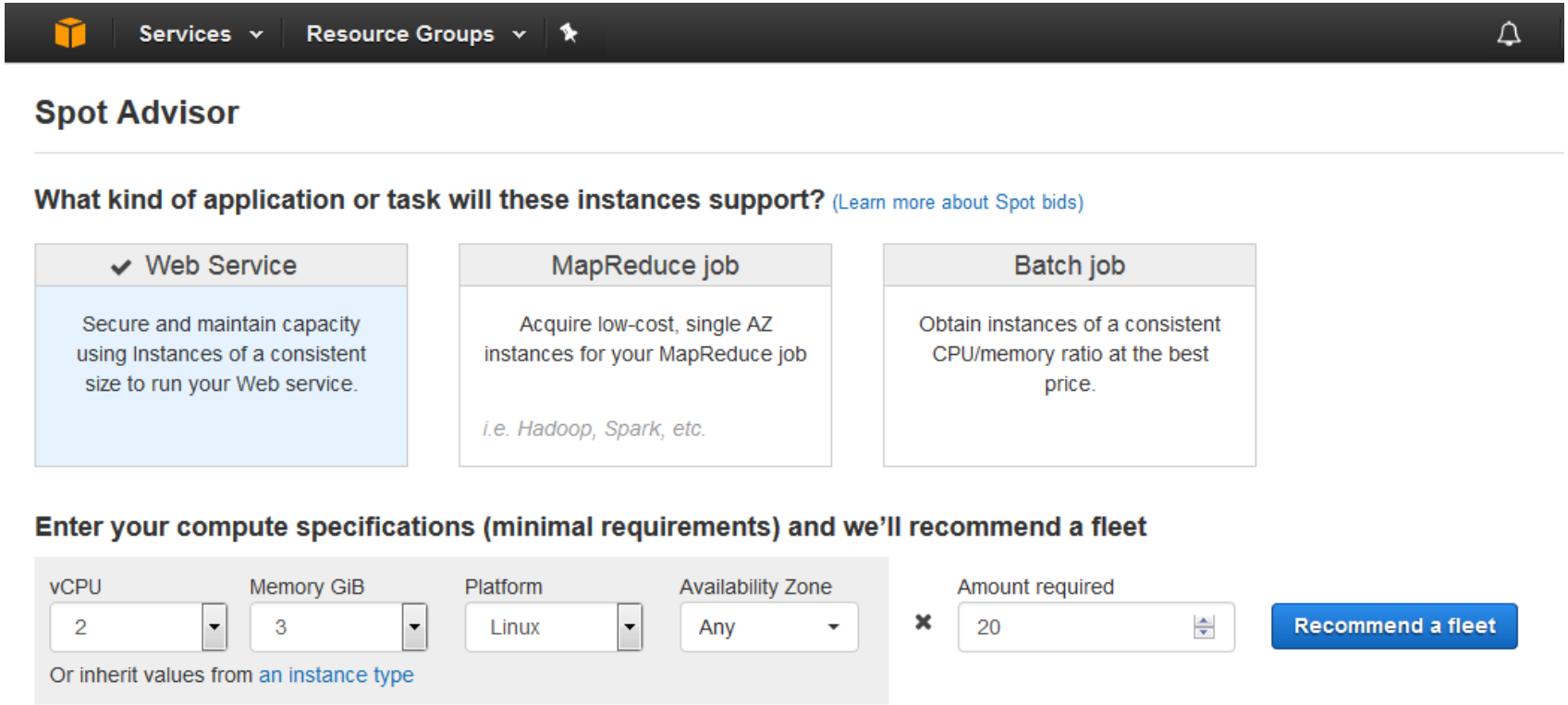
50% bid of OD

75% bid of OD

25% bid of OD

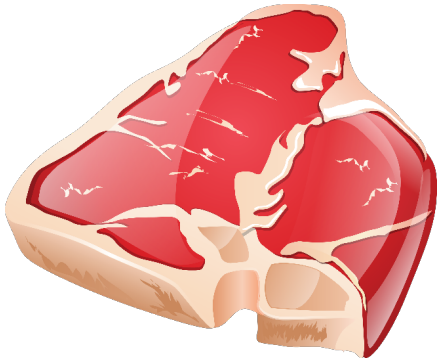
You pay the market price **85% discount!**

Spot Advisor



Which EC2 consumption model is right for me?

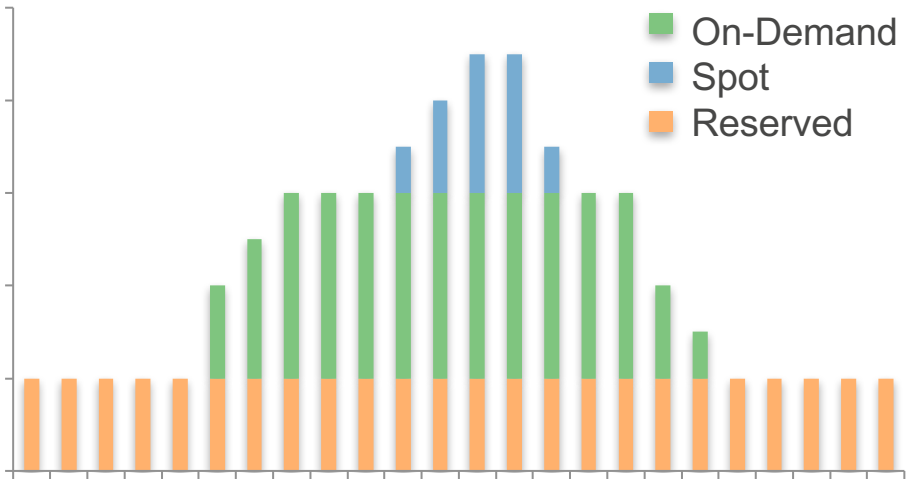
Have a balanced meal!



Why choose just one business model?

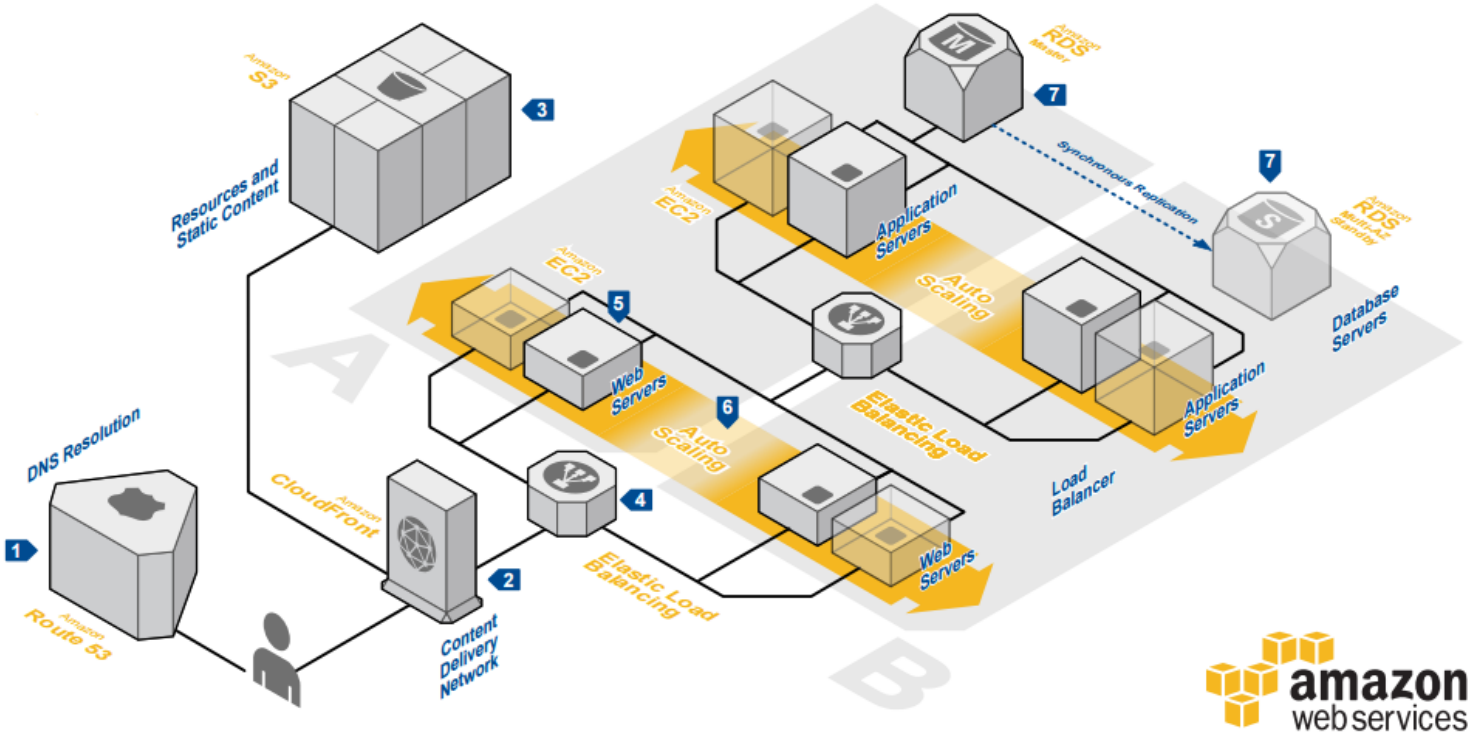
Use a combination of all three!

1. Use Reserved Instances for known/steady-state workloads
2. Set up multiple Auto Scaling groups
3. Scale using Spot, On-Demand, or both



Examples

Example: Three-tier web app

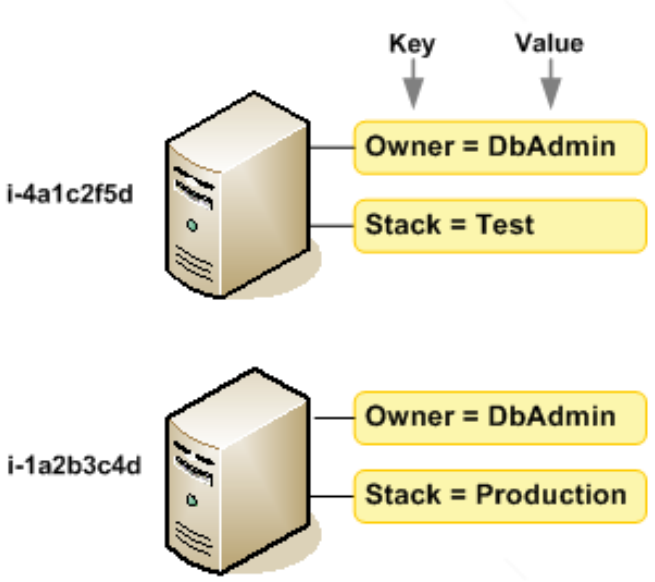


amazon web services

Tagging is essential!

- How do we **explain** the **costs**?
- How do we **allocate** the charges to the **right team**?
- How do we **save** money?

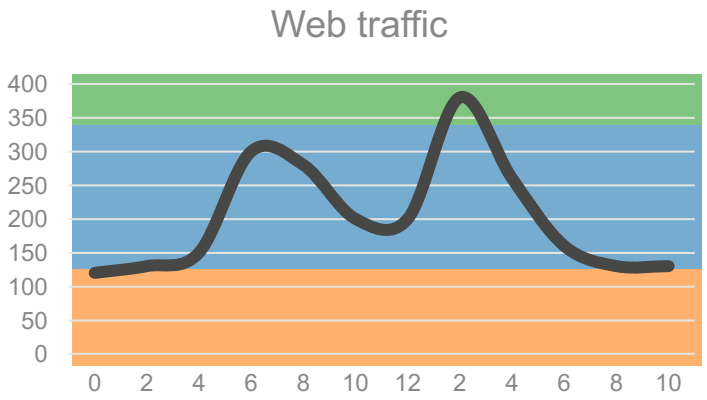
Tagging is essential!

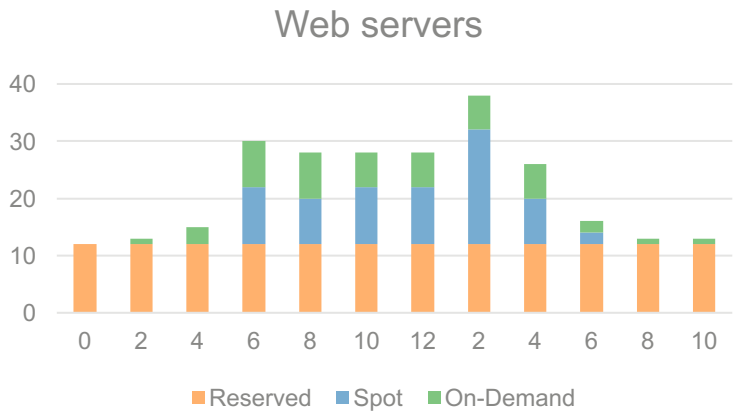


Web tier

Stateless

Scale based on demand

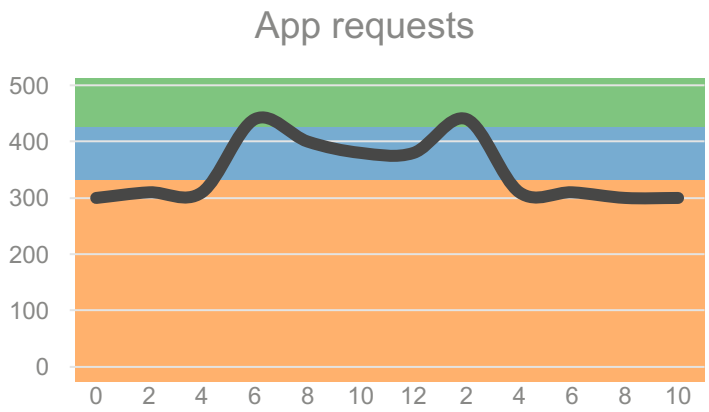


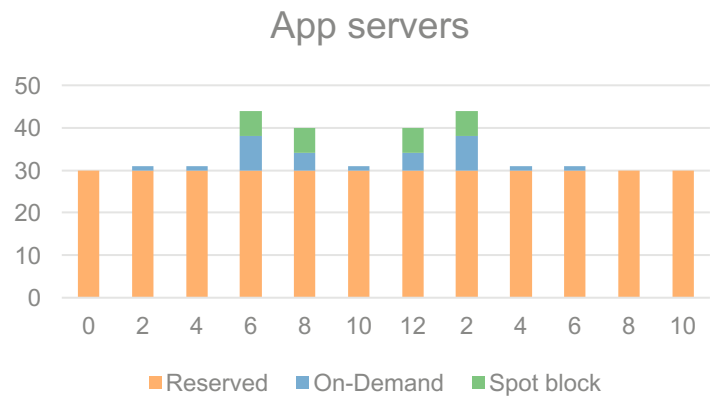


App tier

Stateful

Scale based on demand

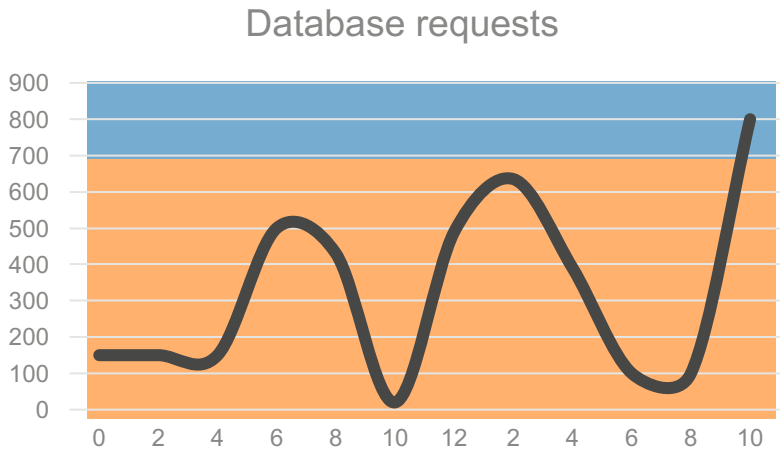


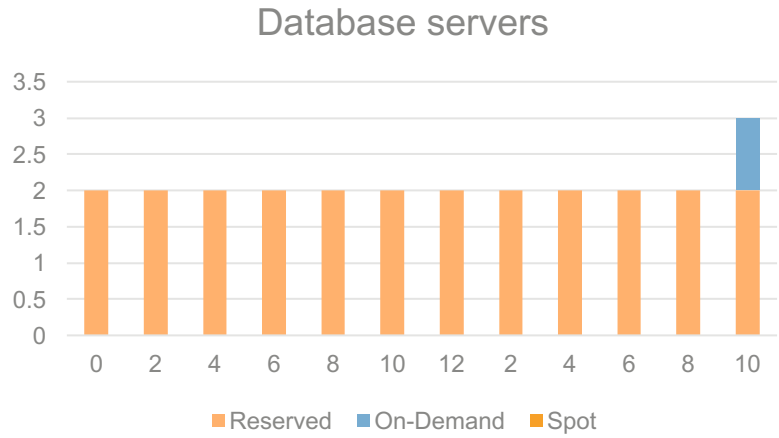


Database tier

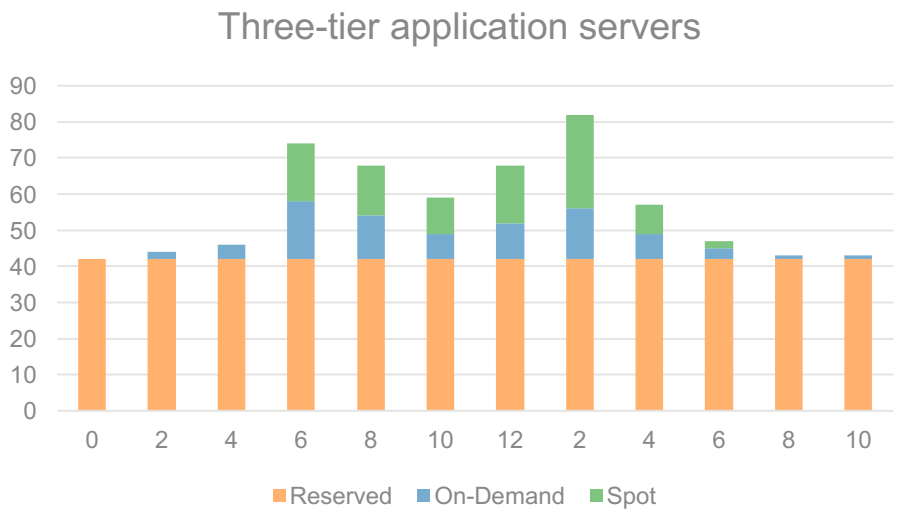
Highly stateful

Static at peak load





Summary: Three-tier web app



Summary

Have a balanced meal! Across the three tiers our meal consist of

- Spot 13%
- On-Demand 11%
- Reserved 76%

Remember!

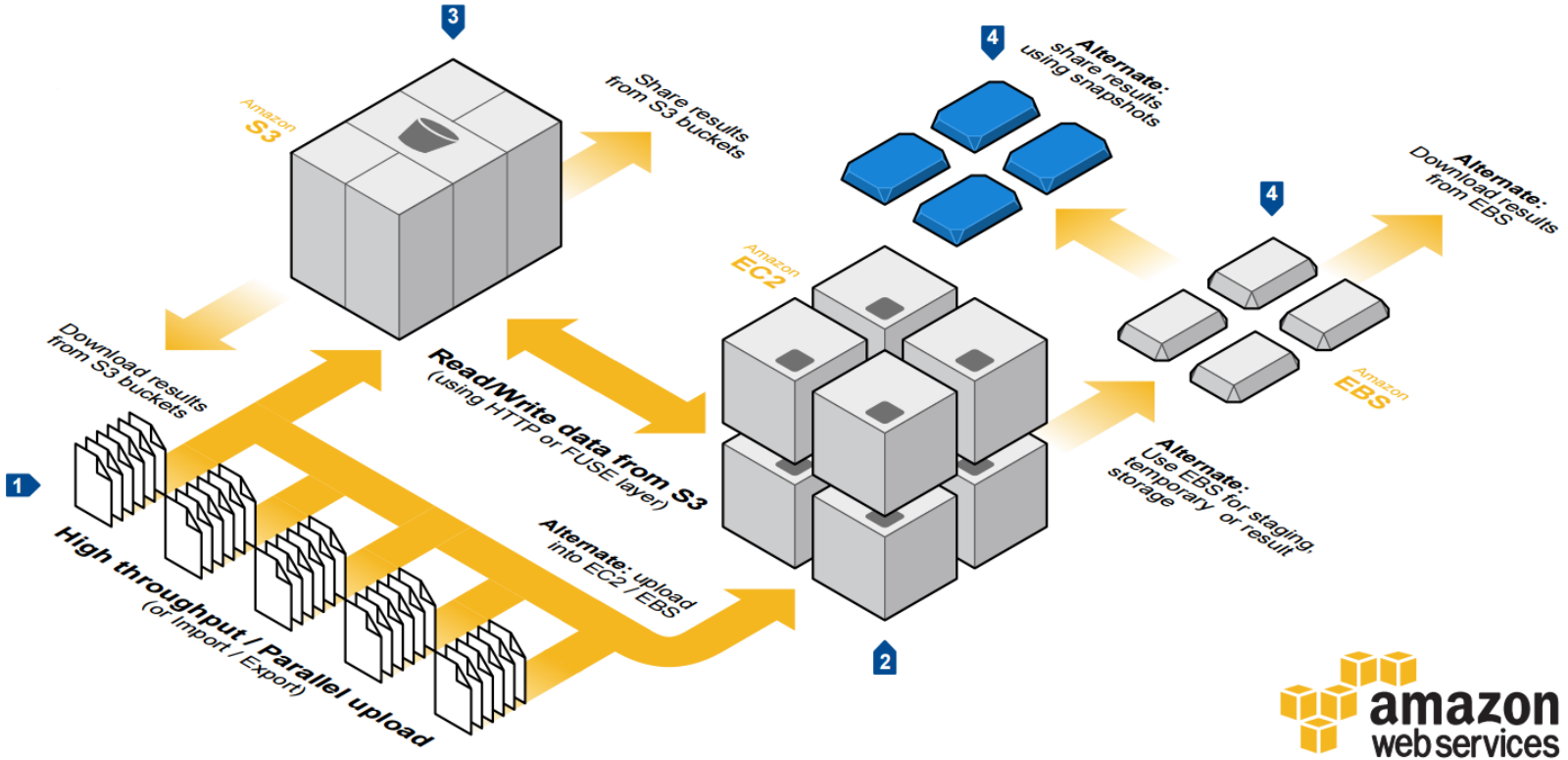
“No server is easier to manage than no server” - Werner Vogels, CTO, Amazon.com

Ubisoft uses AWS to develop and launch social games quickly

> By using the AWS cloud we were able to launch 10 social games within 18 months.
>
> **Lenin Gali**
> Senior Director, Ubisoft

- Ubisoft is a Paris-based gaming company, and creator of popular gaming titles, including Assassins Creed, Far Cry, and Just Dance
- Moving games to social and mobile platforms required capacity to scale fast; using a traditional environment would be an extensive and costly investment
- Using the AWS Cloud to optimize games at the application, caching, and data layers, improving the user experience

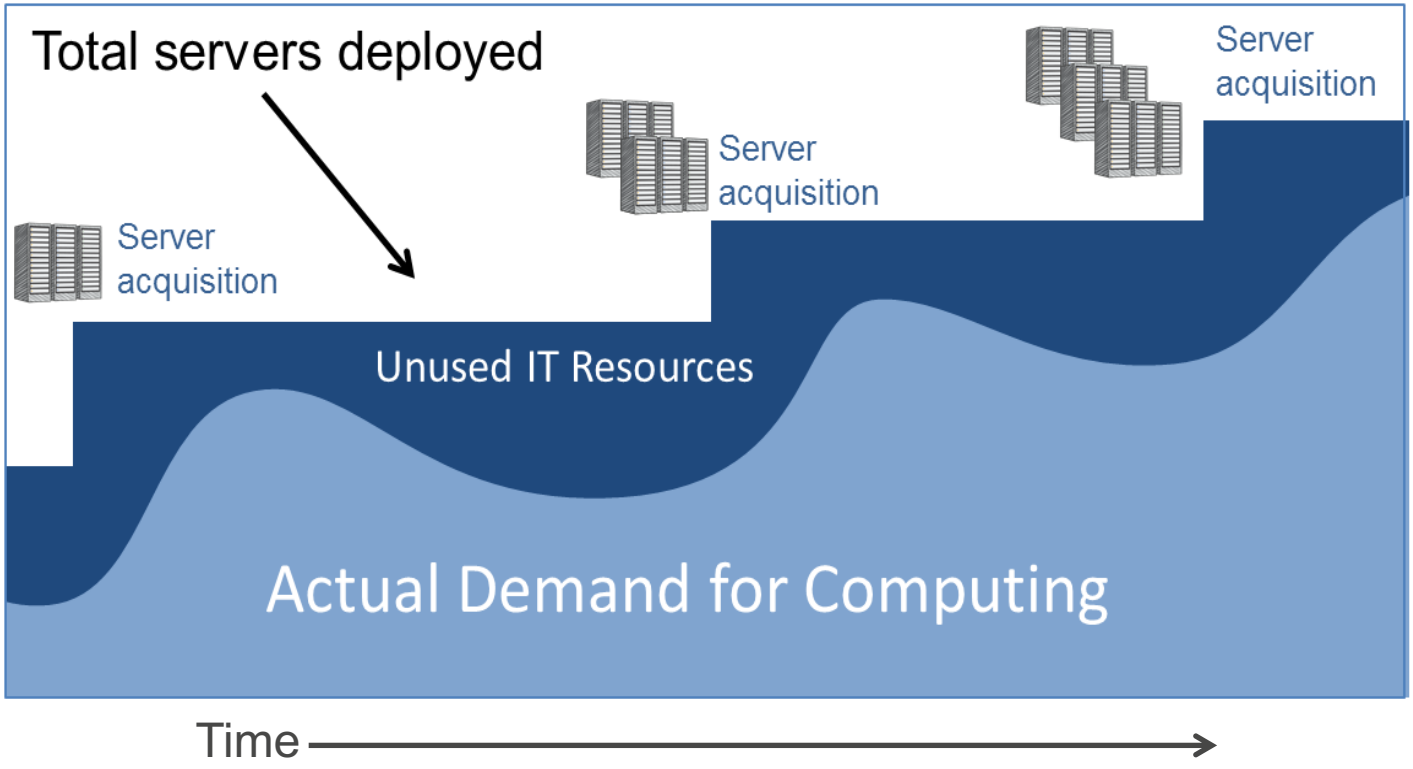
Example: Grid processing



amazon web services

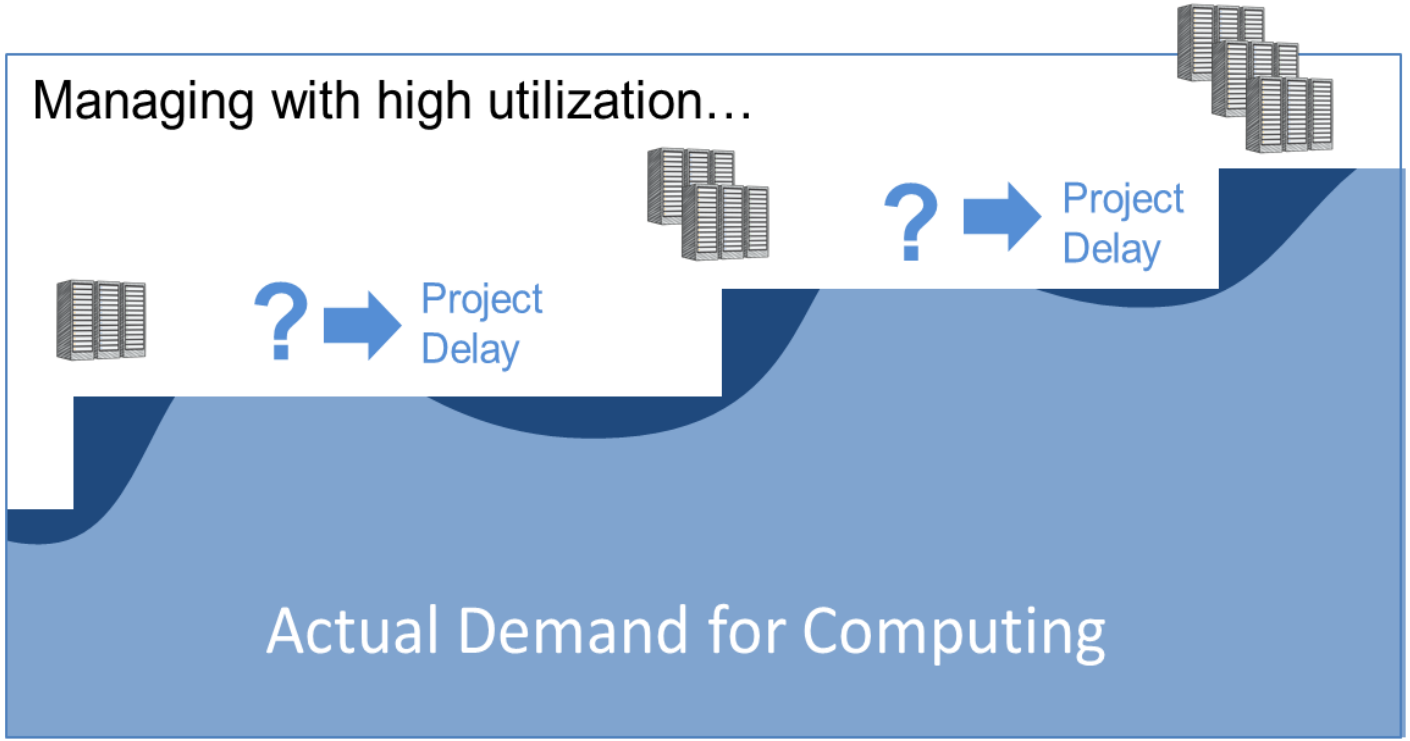
The old way: Low utilization, high costs

Typical server utilization rates are low due to need to deploy for peak needs…

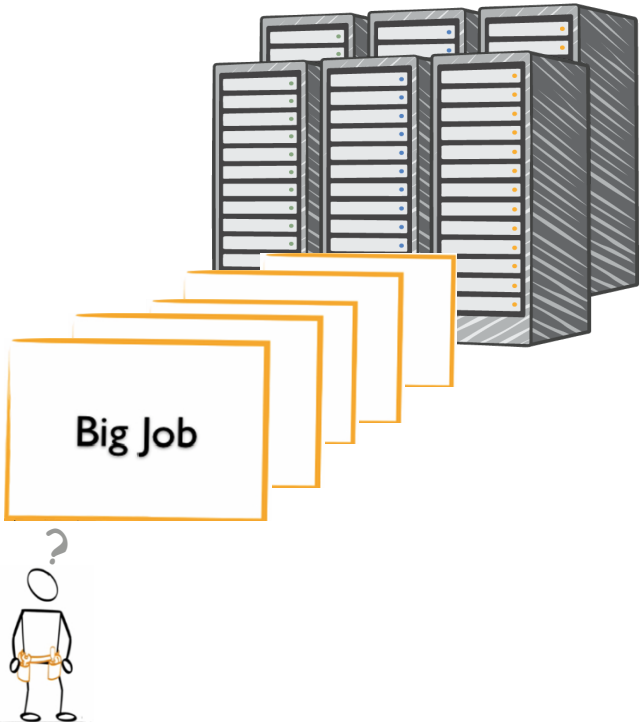


The old way: Managing utilization with grids

Higher grid utilization rates result in hidden costs: longer queue wait times and delayed results

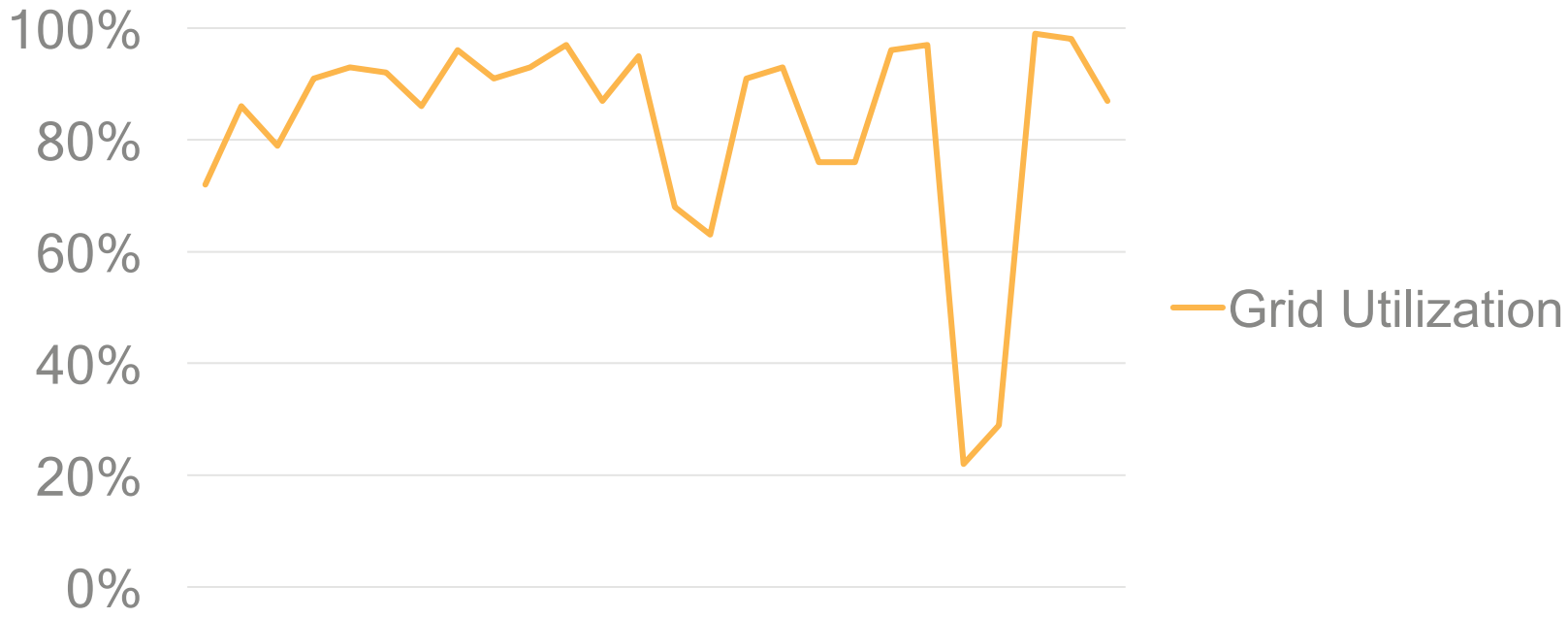


The old way: In the cloud!



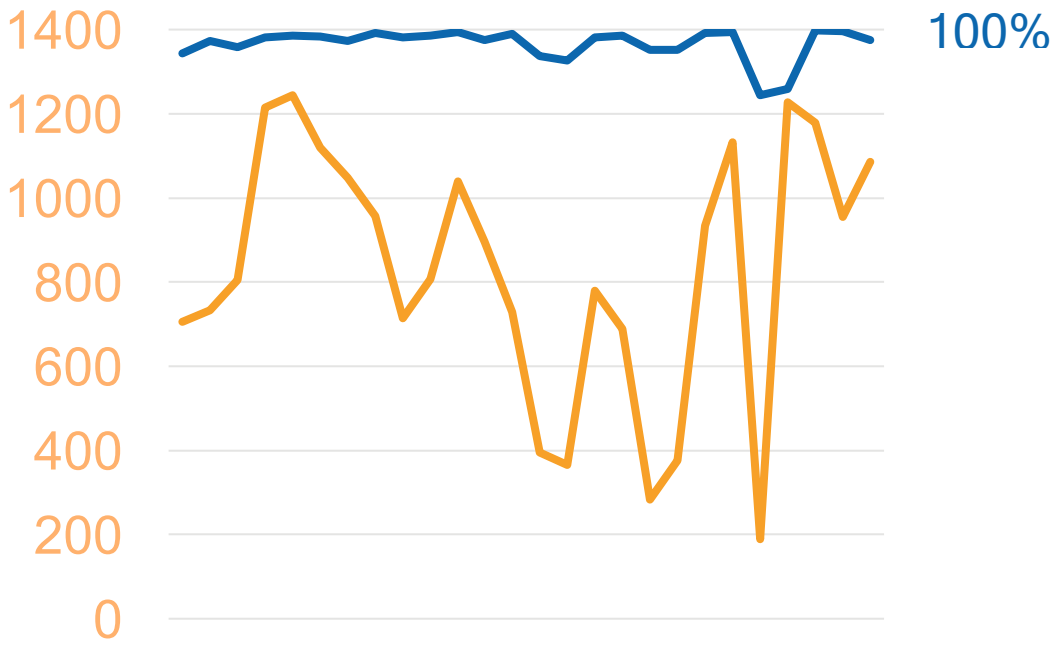
The world as seen by central IT

Higher utilization can reduce IT spending…



The world as seen by the business

But higher utilization also creates IT constraints…

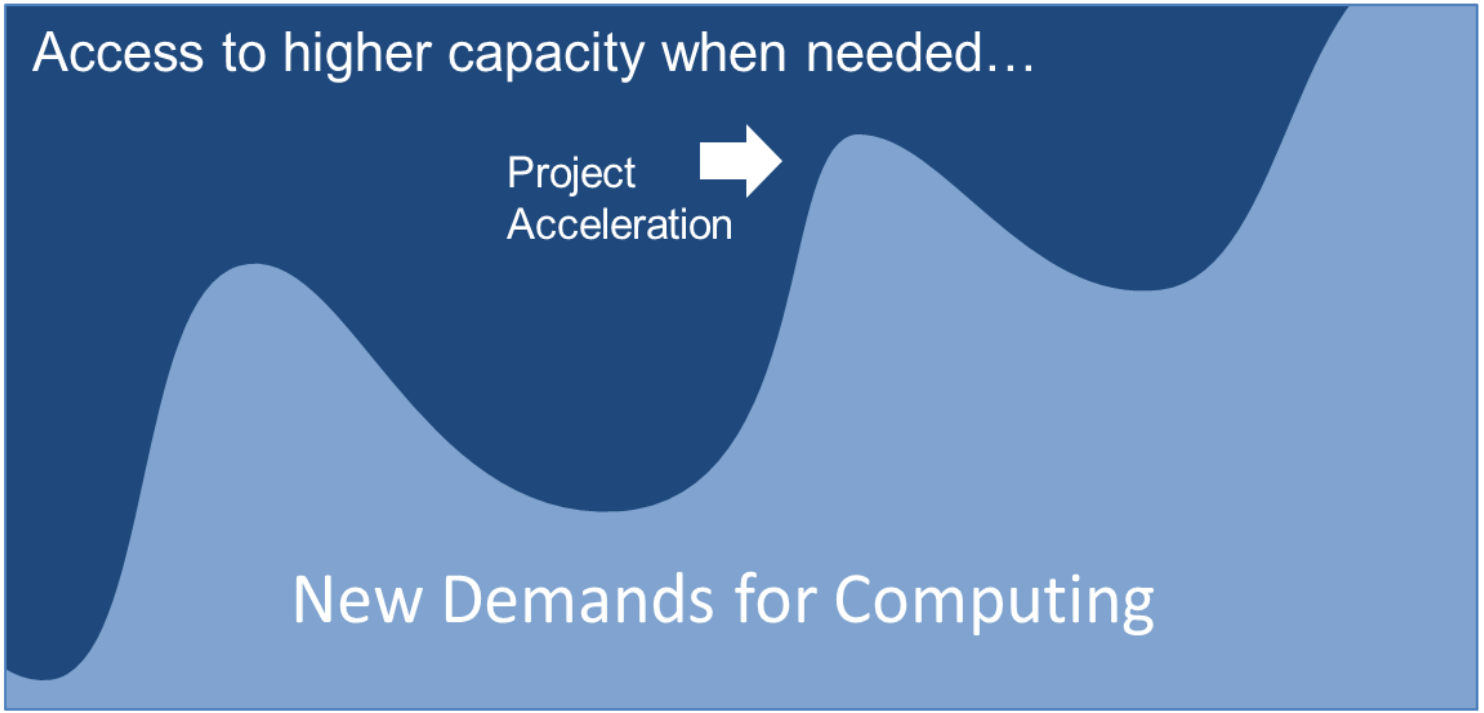


Every job in the queue represents business impact

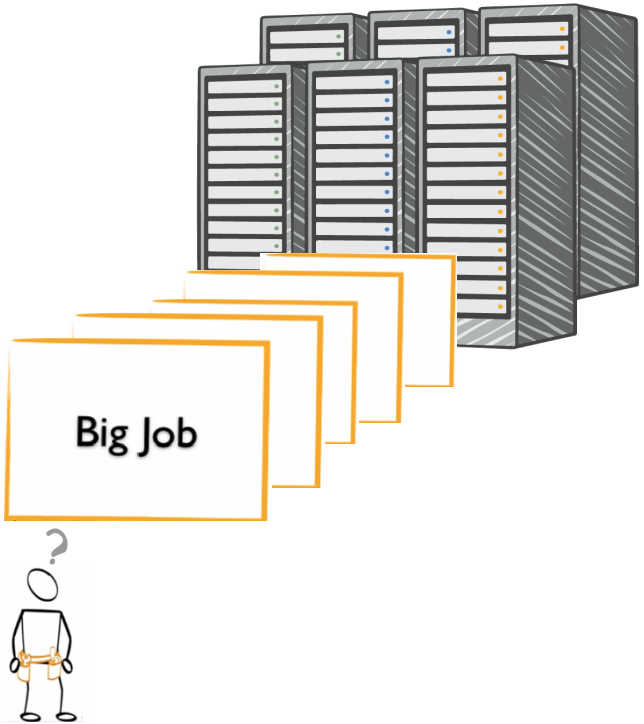
But grid utilization looks great!

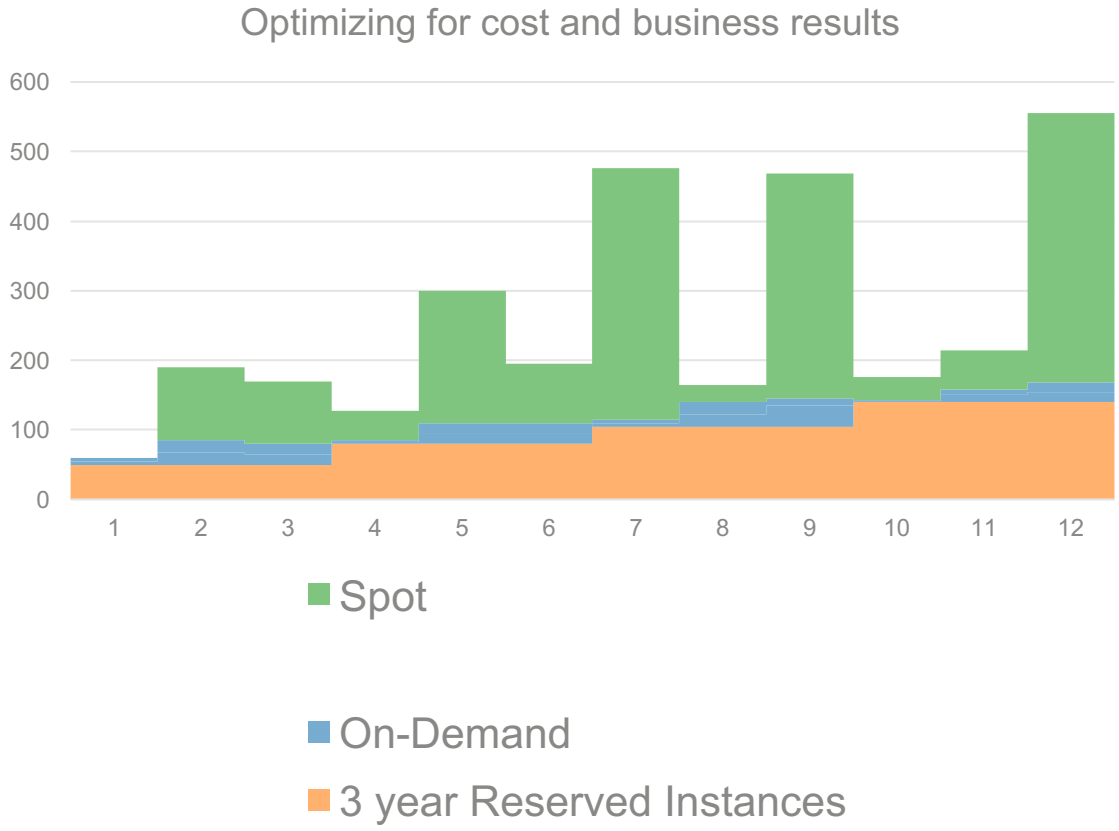
The cloud way: Scalability when needed

Scale higher to reduce time to results: shorter wait times, greater agility, faster innovation cycles

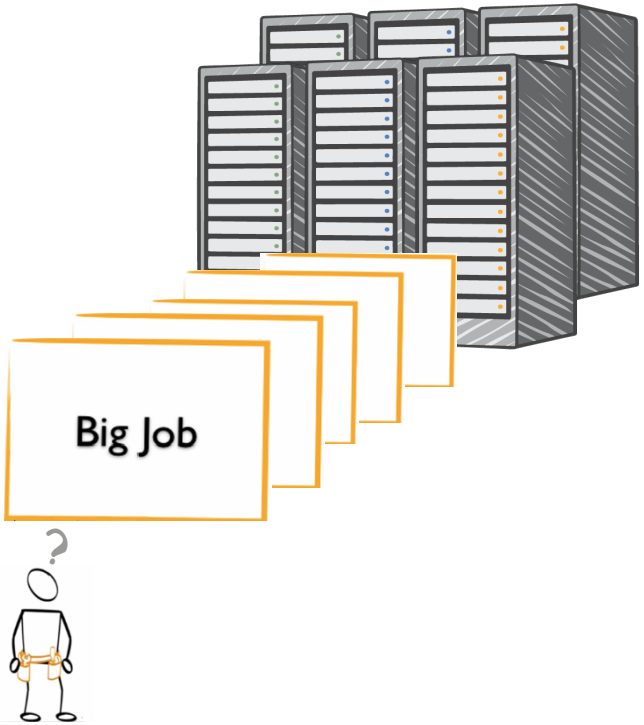


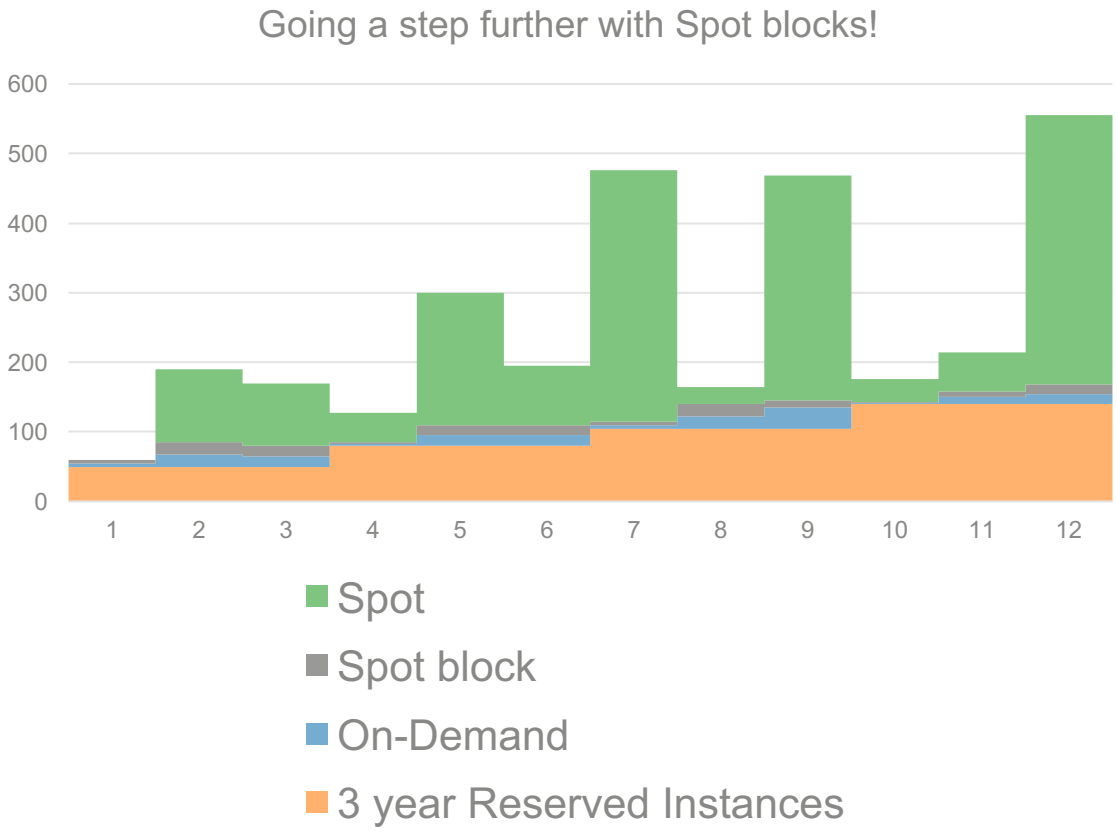
The new way: In the cloud!





The new way: In the cloud!





Accelerating transformation

"We constantly understate what our capabilities are to solve problems. The biggest constraint is never the constraint of time or money, it's generally the constraint of thought."

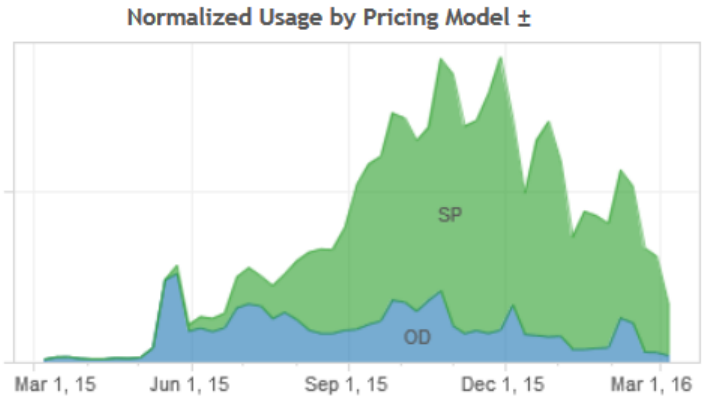
– **Jeff Smith, CEO, Suncorp Business Services**

Founded: 1996 • Employees: 15,000+ • Headquarters: Brisbane, Australia

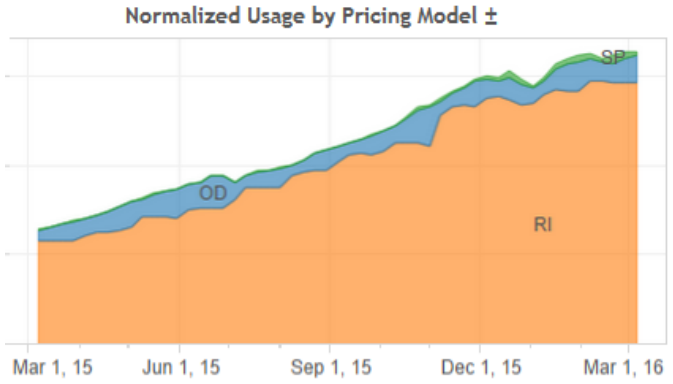
EC2 purchasing options by industry

Consumption model by industry

Web scale (e.g. Adtech) company

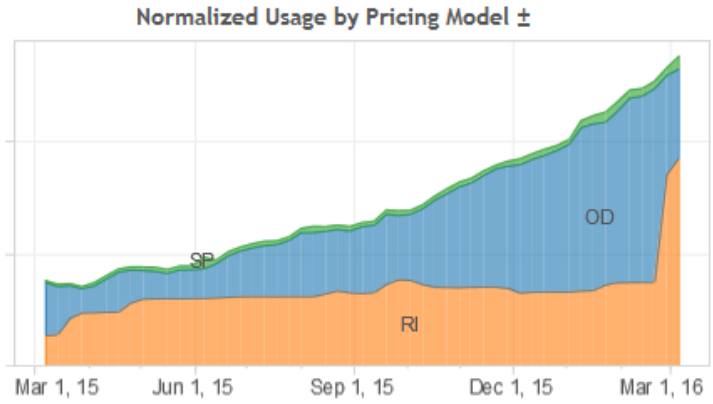


Enterprise SaaS company

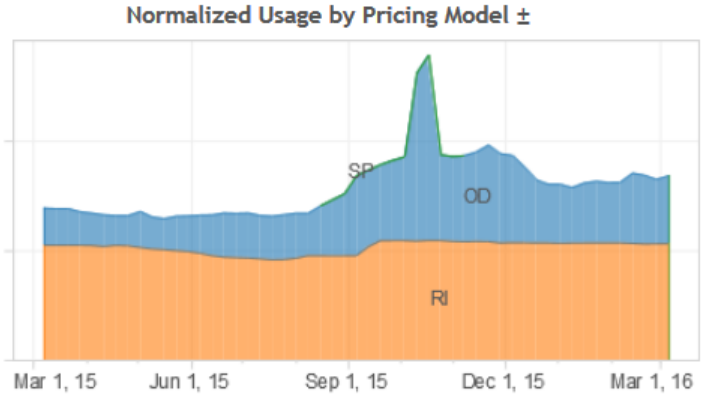


Consumption model by industry

Onboarding enterprise

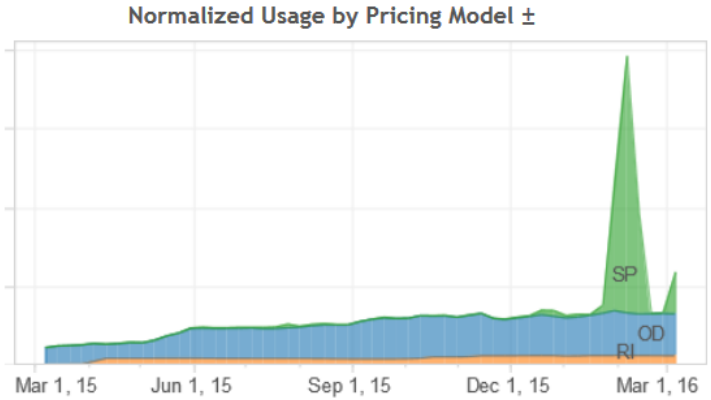


Gaming company

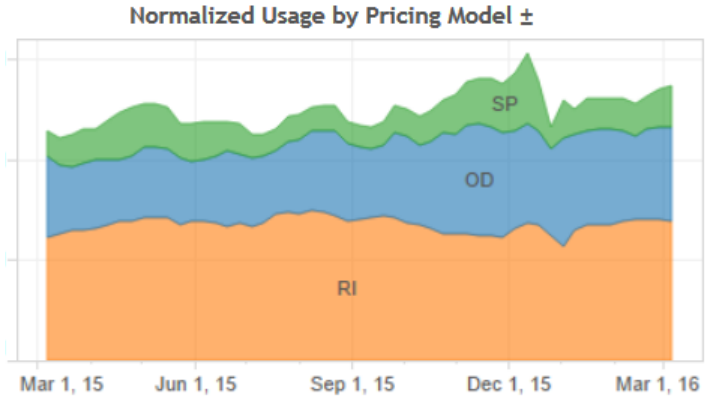


Consumption model by industry

Scientific research

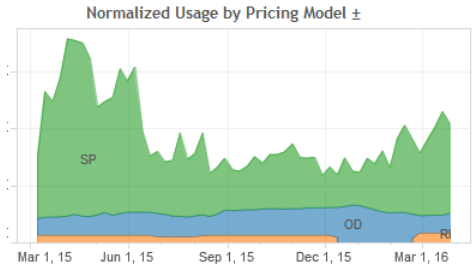


Technology company

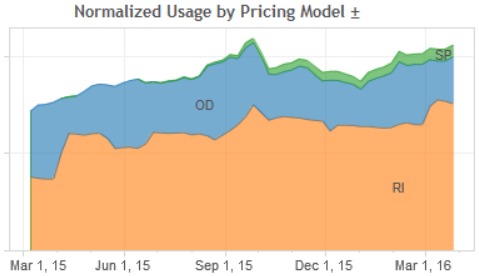


Different purchasing options in a single company

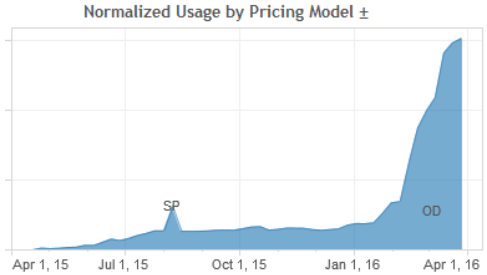
Data science



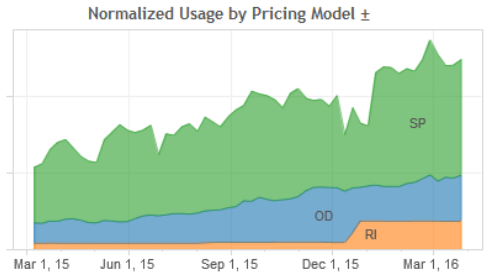
Internal IT



New app development



Test and development



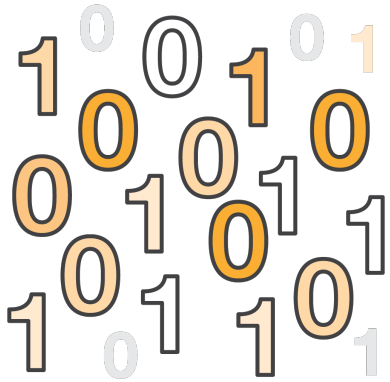
Let’s recap

- ✓ Remember the pillars of optimization
  - ✓ Right-sizing
  - ✓ Increase elasticity (turn stuff off!)
  - ✓ Measure, monitor, and improve
- ✓ Use tags to understand your services
- ✓ There are 3 core purchasing options – have a balanced meal
- ✓ Architect your workloads with performance and cost in mind

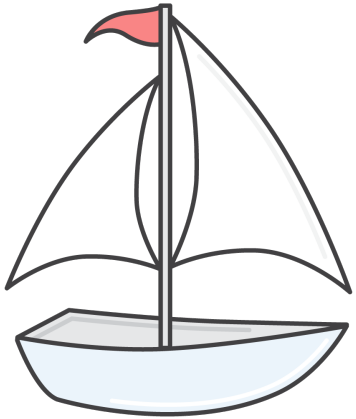
Summary

AWS is more cost-effective in both short-term and long-term than on-premises environments. By leveraging the EC2 consumptions models you gain the…

Freedom to build unfettered



Freedom to get real value from data



Freedom to say yes

AWS

SUMMIT

Thank you!

amazon web services