

REPORT REPRINT

# Economics of containers and Kubernetes: COVID-19 has shown the value of paying as you go

JULY 13 2020

By Owen Rogers

In a recent Cloud Price Index study, we performed a price comparison of AWS, Google and Microsoft container platforms using a statistical model, and found that utilization should be the key decision criterion for cost-conscious enterprises.

---

THIS REPORT, LICENSED TO CARBON RELAY, DEVELOPED AND AS PROVIDED BY 451 RESEARCH, LLC, WAS PUBLISHED AS PART OF OUR SYNDICATED MARKET INSIGHT SUBSCRIPTION SERVICE. IT SHALL BE OWNED IN ITS ENTIRETY BY 451 RESEARCH, LLC. THIS REPORT IS SOLELY INTENDED FOR USE BY THE RECIPIENT AND MAY NOT BE REPRODUCED OR RE-POSTED, IN WHOLE OR IN PART, BY THE RECIPIENT WITHOUT EXPRESS PERMISSION FROM 451 RESEARCH.



**S&P Global** Market Intelligence

### Introduction

Containerization is essentially a newer alternative to the virtualization of full-system images into VMs. Workloads built upon containers pack all their user-mode dependencies, but share key operating system resources under a common kernel, meaning they are smaller and abstracted from underlying library conflicts. Since application libraries and dependencies can be updated independently for each application, and many containers can operate on the same server, there's an expectation of benefits, such as increasing application agility and reducing costs.

Cloud infrastructure providers are a natural home for containers. They provide inexpensive access to a wide array of servers, with potential for responsive and automated scaling. It wasn't until 2015 that the hyperscalers took the opportunity seriously. In that year, Google, Microsoft and Amazon Web Services (AWS) launched platforms in the form of Google Kubernetes Engine (GKE), Azure Container Service (ACS) and Elastic Container Service (ECS), respectively. Oracle has followed suit, as have IBM, Alibaba and others. But with more options, variations and pricing metrics, understanding – let alone optimizing – costs has become a nightmare for enterprises.

In our Economics of Cloud Containers and Kubernetes report, the digital economics unit and cloud-native channel at 451 Research used public website information from AWS, Google and Microsoft to build Python-based cost models using each provider's respective container platforms. A sample of the outputs of these simulations was verified manually, with the data fed into a chi-squared automatic interaction detection (CHAID) decision-tree for statistical analysis. CHAID analysis is designed to reveal the influence of various inputs in achieving a given result.

### 451 TAKE

Ultimately, the choice of container platform depends on one's technical requirements and hyperscaler relationships. However, it would be naive to say cost doesn't matter – enterprises don't have endless pockets to pay a premium for something, when it could be found cheaper elsewhere. Value is a function of both capability and cost, and enterprises can make savings today by making some informed decisions before deployment. How confident are you that you can keep utilization of your servers high? If very confident, then a server-based model can be cheapest, especially if you're confident enough to make a commitment. If not, then PAYG is less risky from a financial perspective – you can grow and shrink as needs dictate, paying only for what you use. COVID-19 has shown that even the best forecasts can't predict everything.

### Method

We used pricing gathered from provider websites to build Python functions that determine prices for running the same workload on AWS Fargate, Google Cloud Run, Google GKE, AWS ECS, AWS EKS, Microsoft AKS and Microsoft Container Instances. Using these functions, we investigated the cost per container for three container sizes (0.5vCPU + 2vGB, 0.5vCPU + 4vGB and 0.5vCPU + 16vGB). We chose these sizes because they reflect the common ratios of virtual machines designed for general, compute and memory-focused workloads, respectively. This approach means our findings would also hold true for multiples of these container sizes. For server-based models, we chose a suitable virtual machine size for each container size, such that an integer number of containers would fit in the virtual machine dimensions to reduce waste. We chose the cheapest US East/Central region offered by each provider.

## REPORT REPRINT

Our simulation priced each offering based on the different virtual machine/container configurations, size of cluster, utilization and requests. In excess of 33,000 prices were determined for 3,000 scenarios, and the numbers were processed using a Python-based CHAID decision-tree. A full demonstration of the complexities and calculations can be found in the full report.

We use list-pricing and do not include free tiers – for large-scale consumption across enterprises, the impact of free tiers would likely dwindle to nothing. We also do not use special offers or benefits such as Microsoft’s hybrid-use benefit. However, we do show prices for Reserved Instances (RI) – we’ve chosen three-year commitments with all up-front payment since this yields the greatest discount. We’ve not included spot instances because they – by their very nature – are unreliable. Where a provider doesn’t support a container configuration, we round up to the closest size. If a capacity limit is reached, they are excluded from the analysis. We’re also assuming a server only holds one size of container, to avoid the mathematical complexity associated with the so-called ‘bin packing problem.’

### Results

Microsoft AKS using Reserved Instances came out cheapest in 47% of scenarios, with AWS Fargate the cheapest in 30% of scenarios and AWS ECS with Reserved Instances being the cheapest in 23%. This is not reflective of enterprises’ experience because we don’t know which scenarios are most common. It is just a statistical analysis of all the (very reasonable) scenarios we investigated.

What factor most determines which cloud provider will be cheapest? The most dominant factor is utilization. Below 30% utilization of a server-based service (e.g., ECS or AKS), Fargate is the cheapest provider. Above that figure, AWS ECS RI or Microsoft AKS RI is the cheapest. The next question: What determines when ECS or AKS will win out when utilization is above 30%? Actually, the difference is razor thin between AWS and Microsoft, but the primary difference is due to the size of the VM. For the 4GB VM, AWS generally comes out on top, but for the 8GB and 16GB VM, Azure wins out. However, this difference is too small to be meaningful, and for this purpose we can say AWS RI and ECS RI are similar. This decision tree below shows what percentage of scenarios are cheapest by utilization and memory. A full analysis can be found in the full report.

Decision Tree Showing Cheapest Provider by Utilization and Memory

Source: 451 Research's Economics of Cloud Containers and Kubernetes



### Recommendations

Enterprises are unlikely to switch providers simply because a competitor's container service is cheaper than their existing one. Containers are part of cloud providers' overall portfolios – even if your provider is expensive for one service, it is likely to be cheaper for another. Don't throw the baby out with the bathwater. However, if you have other reasons to move providers, then the cost of containers can have a big impact. If you are generally happy, then you might decide that paying more isn't necessarily a problem.

Moreover, container services are different, and each buyer will have different requirements beyond cost. These will likely have a significant impact on decision-making.

The most important financial characteristic of a container platform is utilization. What percentage of the investments you've made is being used to add value? If you invest in an RI for three years, you need to sweat that asset – if the asset is used throughout its lifetime for value-adding purposes, the unit costs for container server models are likely to be lower than using a PAYG model. If using a VM for an hour, the same issue applies – you have to ensure you are using it as much as possible.

This isn't as simple as you may think. You need someone or something to constantly track container usage and scale up or down as required. Remember that scaling, even auto-scaling, must start before the server is full – if the server is already densely packed, then there are likely performance issues already extant. There are further challenges in ensuring you're using the cheapest VMs as new generations are released. If you make a forecast of utilization, but then fail to achieve this level, you have no recourse if you've made an up-front investment. At least with PAYG models, you can flex up and then down. When the unexpected happens (like global coronavirus pandemics), you are in a position to be more flexible.

Our analysis also shows that slight variations in VM and container sizes could slightly change findings. For enterprises, this means choosing the perfect VM for a container host is important for reducing TCO – again, something you wouldn't have to worry about using a PAYG model.

Our view is that, unless you are confident you can achieve a high level of utilization, most enterprises should use PAYG if cost is of paramount importance. There are pros and cons of both server and PAYG models, but at least with PAYG you have the freedom to scale up and down as required without the overhead of managing capacity. In uncertain times, this freedom is of critical importance. Fargate Savings Plans are a good compromise here, giving freedom plus a discount – but only if you are certain you can spend the commitment, even if not on Fargate. If you spend all your AWS savings plans dollars and achieve 100% utilization, you can secure prices similar to RIs.

Here, we have analyzed only one line-item on the bill. Understanding the financial impact of the bandwidth, storage, other services and support packages that will be required will help prevent runaway costs. No cloud service is an island; they are all part of the mainland.