A Report of Cloud Solution for Pet Food Company

Yiyong Duan

University of Western Australia

22628103@student.uwa.edu.au

Abstract

This report introduces a cloud architecture for Pet Food Company to upgrade their on-premises systems into cloud based. To begin with, basic concepts of cloud computing and different cloud service modes are introduced. In the following, an overall cloud architecture is designed according to the factual situation of PFC. Specifically, for their in-house systems, a SaaS solution is described. Besides, Application Programming Interface is introduced for potential applications development. Moreover, in order to resolve the high concurrent issue, a scaling solution with auto-scaling and load balancers is recommended for PFC. Lastly, DevOps is briefly introduced for PFC's future microservices development.

Keywords: Cloud Computing; Data center; SaaS; Scaling; DevOps.

1. Introduction

In the context of information era, data is getting more significant for enterprises to strategize their business. For storing and analyzing data, Pet Food Company (PFC) has set up their data centers with mix of servers, networking and storage devices, which make sure they can access data directly and safely. With the demand for global expansion, PFC decides to move their systems into cloud for meeting the requirement of data computing. In this report, the cloud architecture with scaling and DevOps are customized for PFC.

2. Cloud Computing

Over the past decades, data centers (DCs) have sprung up as the increasing requirement of data computing. Two issues have occurred in this process. Firstly, most of individuals can't afford the cost of DCs. Second, redundancy resources are in idle when DC is unused. Therefore, some companies integrated the IT resources, especially the DCs and data storages, and provided a service for data computing according to needs.

In general, cloud computing provides on-demand services for multiple users, from a model consisted of shared computing resources pool[1]. In the context of PFC case, cloud computing can help PFC launch an e-ecommerce engine and agilely allocate the IT resources for their expansion of business needs[1].

According to services, cloud computing is generally divided into three parts:

Infrastructure as a Service (IaaS). IaaS provides on-demand self-service for computing, networking, storage and other services. IaaS is fit for the company with invested physical infrastructure. The mainly IaaS providers are Amazon EC2, Google Compute Engine, Microsoft Azure, etc.

Platform as a Service (PaaS). PaaS provides customization tools and software on the cloud for the users to develop applications. Mainly providers include Amazon Elastic Beanstalk, Google App Engine, Heroku, etc.

Soft as a Service (SaaS). SaaS combines both the infrastructure and software on the cloud. Users

can directly run business without operating hardware and developing apps, which reduce the cost of development and management. Mainly provides include Microsoft Office 365, Amazon Workspaces.

2.1 Prerequisites of Moving to Cloud

Before moving to cloud, the company should consider several conditions by comparing the pros and cons of cloud and traditional DCs, and decide whether the cloud computing is optimal choice.

Cost and risk. Traditional DCs are physically partitioned into portions directly for business requirement, which is convenient and safe. However, it's not agile for reassign network resources if the company needs to expand. In the process of business growth, it's hard to predict how many resources will the traditional DCs need, which may cause exceeding cost of investment. Moreover, the cost of energy of physical infrastructure can be immeasurable. Compare to traditional DCs, the cloud based on virtualized data center networks (DCNs) are more agile and flexible for company to change the scale, without the worries of investment risk as well as the cost of faculty maintenances.

The market expansion. For local companies, the traditional DCs may satisfy the business in most cases. For those companies who aim at global market, the data will be too massive for DCs to handle. Therefore, moving to cloud is necessary.

Security. The data security is about protecting data from hacking, deletion, leakage, etc. It's trusting for company to save data in their own physical DCs. As for cloud computing, it's concerning that the data can be leaked by vendors or from the unsafe data port. The company should carefully investigate the cloud tenants in advance, and decide what type of data could be stored.

The Pros and Cons of Cloud are listed below:

Characteristics of Cloud	Pros	Cons
Cost	Pay as business grow, more agile for company with expansion plan, reduces the cost of hardware iteration update.	The rental is not cost-effective in the long-term.
Agility	Easy to implement, access data wherever network is available, no hardware required.	Network issues that limit the massive data transmission, unexpected leakage may occur from unsafe network port.
Security	Save cost of security measures and crews.	Risk of data confidentiality, security is out of company's control. No backup available if cloud servers break down.

Table 1. Pros and Cons of Cloud

In this case, as introduced previously, PFC aims at expansion of market, including especially Asia and China. For the cost perspective, cloud DCNs are cost-efficient for the growth development. Besides, with the combination of physical and cloud IT resources, the data can be protected privately and more secure. Last but not least, the flexible cloud can provide data computing of trans-regional business flexibility and convenience.

After primarily thought of preparatory conditions, the company may now step on the road to cloud.

2.2 Move to Cloud

In general, the steps of moving to cloud include: choosing the mode of cloud, cloud vendors investigation, cloud architecture design, system test, migration deployment.

2.3.1 Choose Mode of Cloud

The existing modes of cloud are: public cloud, private cloud, hybrid cloud and multi-cloud.

Public cloud is a cloud service that customers share the computing resources from mutual server. Each customer's data is independent and hidden from the other customers.

Private Cloud is not shared with other customers. It could be provided by third-party cloud vendors or established by company itself.

Hybrid Cloud is the combination of public cloud and private cloud. Each type of cloud cooperates with others to finish the same computing task. For instance, the web application works in the public cloud, whilst the data storage works in the private cloud.

Multi-Cloud is a cloud service, which is mainly provided by different public cloud vendors (also includes private cloud), in order to finish computing tasks separately. Company usually adopts multi-cloud like a diversified portfolio to minimize the risk of unitary.

In PFC's case, since the company has the physical IT infrastructures and rented DCs, it's convenient for them to scale up on-premises infrastructures to the public cloud with configuring the Virtual Private Cloud (VPC). The hybrid cloud mode allows the company stores the confidential data in their own DCs, while cloud computing can be flexibly done in public cloud.

2.3.2 Choose Cloud Computing Service

PFC have invested the development of hardware and software. According to the cloud service introduction in 2.1, IaaS is the recommended computing service in this case. In other words, PFC is capable of transiting IT resources into cloud and operating the business in a hybrid cloud mode.

2.3.3 Choose Cloud Providers

The cloud leaders in the global are such as Amazon Web Services (AWS), Alibaba Cloud, Google Cloud, Microsoft Azure, IBM, etc.

In global market of IaaS, Amazon owned the biggest shares amounted to 45% in 2019, while Azure owned 17.9%, Alibaba lifted the shares to 9.1%, estimated by Gartner research. However, in APAC (Asia Pacific) region, Alibaba cloud owns the biggest IaaS market share. The table 2 shows the Cloud Services Leadership based on IaaS, PaaS and hosted private cloud revenues in Q2 2020, APAC region[2].

Cloud Services Leadership in APAC Region

Rank	Global	China	Japan	Rest of East Asia	South & Southeast	Oceania
1	Amazon	Alibaba	Amazon	Amazon	Amazon	Amazon
2	Alibaba	Tencent	Fujitsu	Microsoft	Microsoft	Microsoft
3	Microsoft	Baidu	Microsoft	Alibaba	Google	Telstra

Table 2. Cloud Services Leadership in APAC Region (Source: Synergy Research Group)

According to Gartner research, the Figure 1 shows the market shares change from 2018 to 2019.
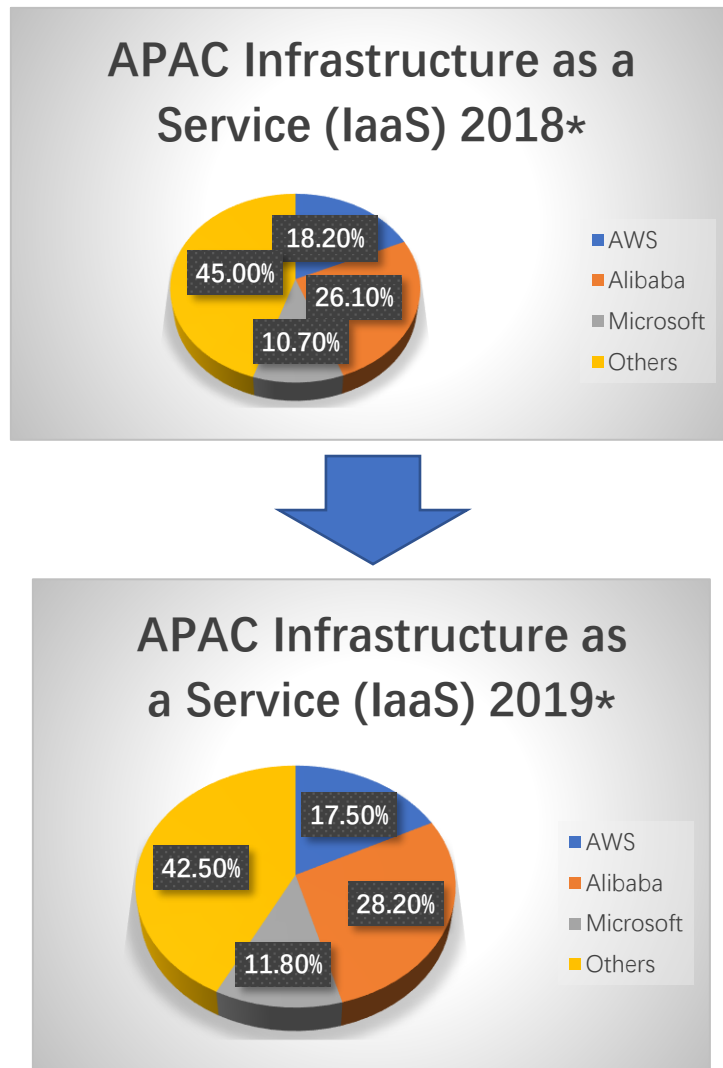
Figure 1. APAC Infrastructure as a Service 2019

*Charts created based on Gartner research. Market share: IT service, worldwide 2019[2]

In the context of PFC project, APAC market will be the primary direction of expansion. From the perspective of market share, it's recommended to adopt Alibaba Cloud or Tencent Cloud in China, AWS in rest regions of APAC. Furthermore, we can compare the differences through the particular cloud service of different cloud providers, such as Virtual Server, Container Management Service and Auto Scaling, etc. Last but not least, from the independence and security perspective, second cloud providers can be considered to form a multi-cloud mode.

3.1 Cloud Architecture

In general, a cloud architecture is consisted of the architecture design of network, application system, database and object storage. Firstly, an overall architecture is introduced for the reference of PFC's cloud architecture. In the context of PFC project, SaaS is used for the in-house systems and the potential devices. Besides, the Application Programming Interface is introduced for suppliers and vendors to simplify the process of supplication and distribution. Finally, the overall architecture of PFC is constituted.

3.2 Physical Infrastructure

Physical infrastructure layer includes server, storage, DCs, firewall devices, etc. They support the

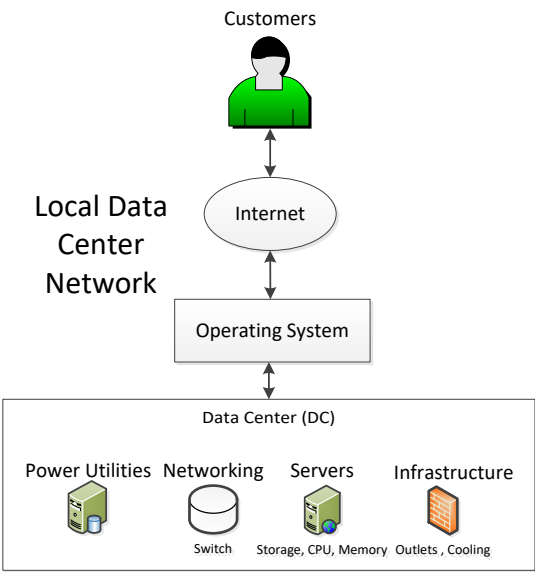data computing, storage, network and security service.

Figure 3. Local Data Center Network

3.3 A SaaS Solution for PFC

In brief, the SaaS is a delivery model that combines all the on-premises into software and directly serves for the customers.

PFC has developed their own enterprise software and uses an on-premise architecture. When moving to SaaS business model, it allows products to be iterated more quickly, which reduces the cost of updating and maintenance.

3.3.1 Virtualization

Virtualization plays an important role in SaaS, which is responsible for delivering commands from OS (Operating System) and application to infrastructure layer, by distributed aggregation and virtualized isolation.
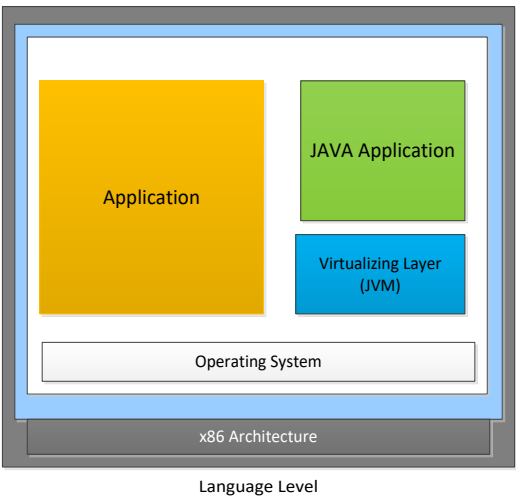

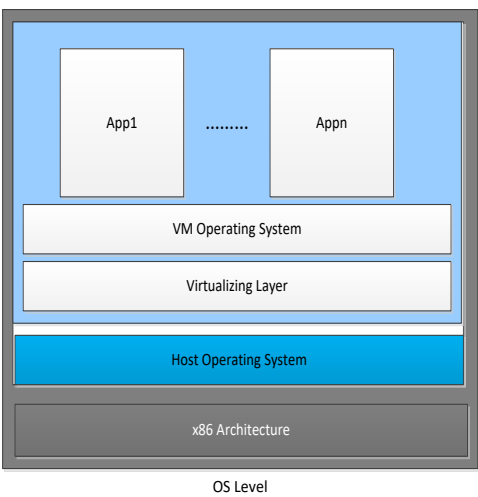Figure 4. The virtualization of language level   Figure 5. The virtualization of OS level
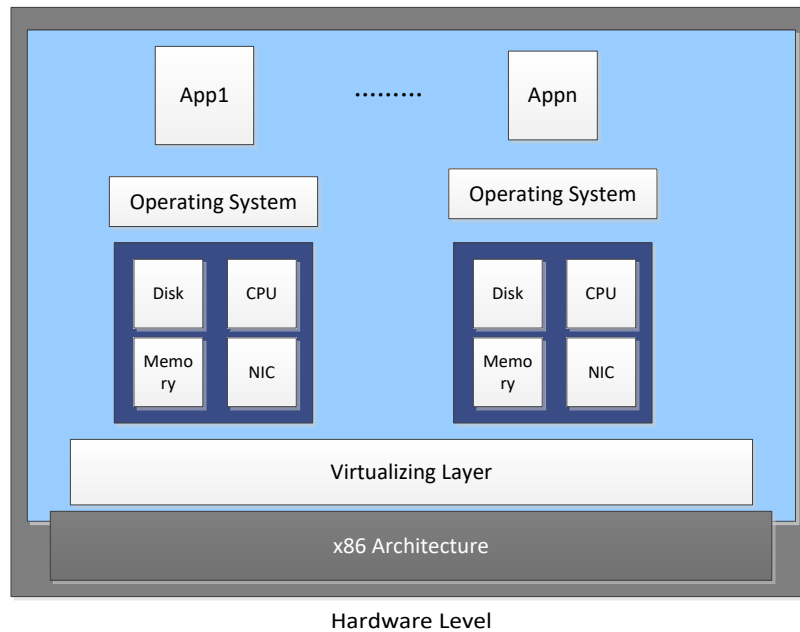
Figure 6. The virtualization of hardware level

Virtual Machines (VMs), Containers and Serverless Computing are the common approaches in virtualization.

3.3.1 Virtual Machine

VM runs on the top of hypervisor, which is a software that allows us to split one physical system into distinct and quarantined virtual environments. VM can run any apps and programs like an imitation of real PC.

Because of VM's isolated environment and similar characteristics of physical computer, it's recommended for company to host their major cloud. For Amazon, it provides AWS EC2, VMware Cloud for VMs solution. For Alibaba Cloud, it provides Elastic Compute Service (ECS).

3.3.2 Container

Based on Linux system, container integrates code, runtime, system tools and other application elements into one instance. Compared to VM, container is relatively limited in size, but works flexibly on apps development.

Because of its lightness and elasticity, container is best for application deployment. Amazon provides Elastic Container Service, while Alibaba Cloud provides Container Service for Kubernetes.

3.3.3 Serverless Computing

In context of serverless computing, the company is able to run their app on the cloud server, with purely code to run a single function. All the company needs to do is uploading the code parcel, without worrying server conditions at user's end.

Because the apps are operating on the cloud server, it's perfect for the company whose network server with small bandwidth to develop applications. Amazon provides Lambda for Serverless Computing Service, while Alibaba provides Function Compute as their event-driven compute service.

3.3.4 In-house and Potential SaaS Solution

In the context of PFC scenario, the in-house systems include both hardware and developed software, which require high-level security and privacy. Therefore, it's recommended to use straight Virtual Machines with quarantined environment, which can be hosted on their servers. Alternatively, for low

latency of the cross-region communication, or other potential functional requirements, PFC can choose serverless computing with local VMs to form a hybrid mode. For financial perspective, on-demands serverless computing is better for

As for choosing cloud providers, Alibaba Cloud is applicable in China, while Amazon is the optimal choice in the rest of APAC region.

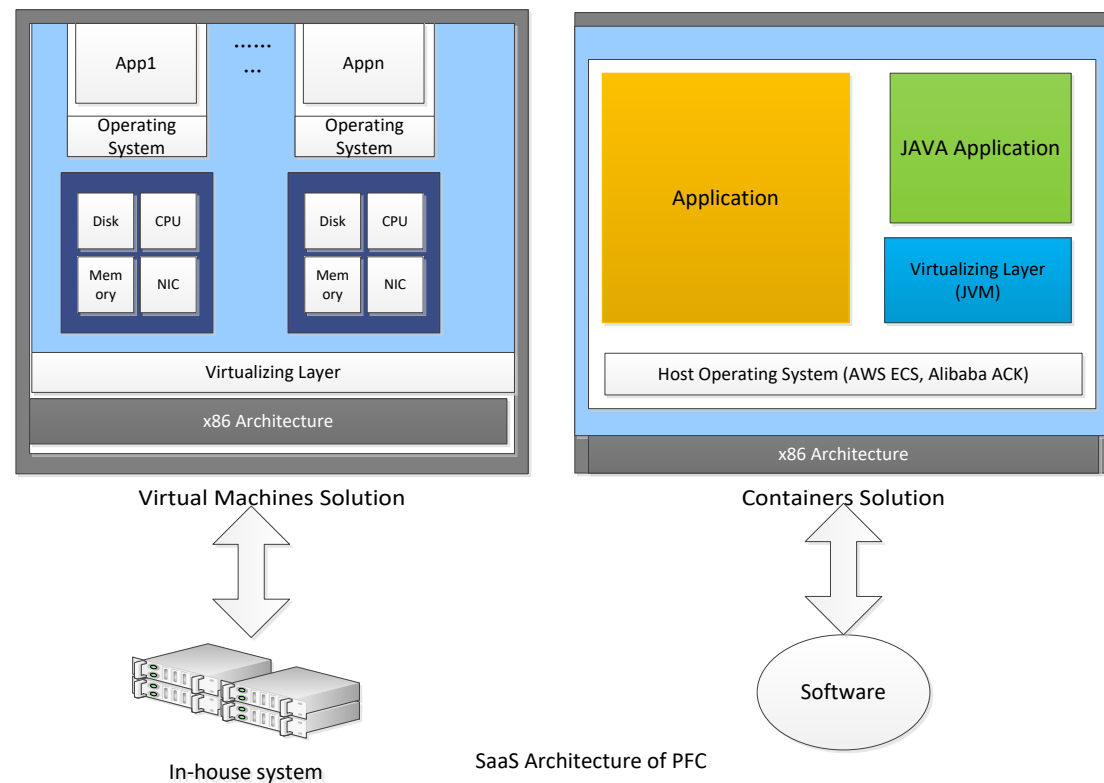### 3.3.5 SaaS Architecture of PFC

Figure 7. SaaS Architecture

3.4 Application Programming Interface (API)

The API is designed for different programs interacting in such a platform, which is defined as an interface. In this interface, the communication can be carried out between all the apps. A tool that manages API is called API Gateway, which allows developers to create, publish, monitor and secure the APIs.

3.4.1 Steps to Create APIs

In this case, we will take Alibaba Cloud for instance. According to API Gateway of Alibaba Cloud Instruction[4], the steps can be divided into four parts: Authorization, define the operation, API Gateway setup, and Debug of API operation.

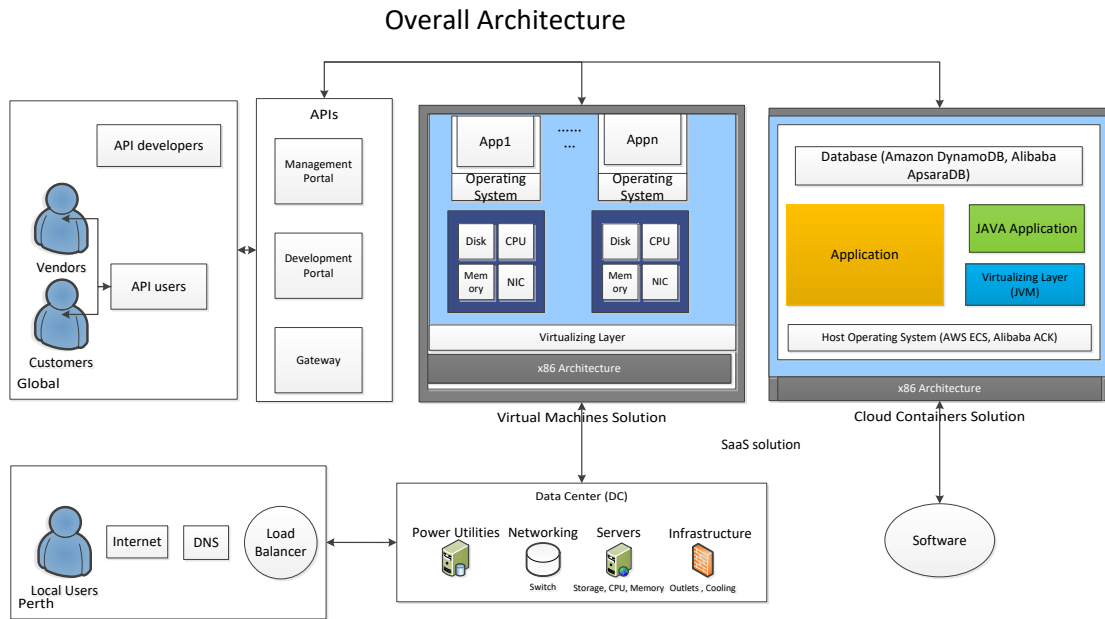### 3.5 Overall Architecture

Overall Architecture

Figure 8. Overall architecture

4. Scaling

As the pet market demand increases, the scale of the company as well as the cloud will be in an ever-expanding trend. For the server, when the volume of user requests increases like a huge surge, it's necessary to use scaling tools that prevents server from crashing and keeps it from running steady and performing well.

The two common scaling ways are: vertical scaling and horizonal scaling.

The vertical scaling suggests to increase the overall volume of load in hardware level by replacing with larger nodes. For instance, we can change a CPU with more cores or memory-chips with larger amount of memory in the PCs, the same way applies to computing infrastructures.

Compared to the vertical scaling, the horizonal scaling works on increasing the numbers of nodes, with the same capacity of the original nodes[5].

The Amazon and Alibaba provides scaling tools such as auto-scaling and server load balancer.

4.1 Auto-Scaling

Auto-scaling serves to automatically allocate the network resources based on the volume of user requests. For instance, when a new pet product is releasing, the auto-scaling system can detect the increasing requests and automatically add more Elastic Containers for provisional demand.

4.2 Load Balancing

As the name means, load balancer is use to load balancing. When high concurrent event happens, received requests are distributed according to a configurable algorithm to a destination IP address, usually belongs to an Elastic Computing instance. Users can improve the scalability and elasticity through distributing traffic among multiple instances by load balancer. In the meantime, the possibilities of SPOFs (Single Point of Failures) are lowered to improve the availability as well as fault tolerance of system.

4.3 Scaling for PFC

Usually, we pair load balancer and auto-scaling in the same time. Auto-scaling ensures instances scale up or down by adding or removing capacity, while load balancer scales load through evenly

distributing load to anywhere available in elastic instances or computational resources servers. In other words, auto-scaling is responsible for the overall capacity volume, load balancer allocates the computing resources evenly on-demand.

In practice, the company is hard to predict the scale they need to set up the computing environment. With auto-scaling alone, the users don't have to worry the over-load issue. However, the resources are not distributed evenly and fully, in other words, users hire 1000 workers but only 500 of them are working. In other situation with load balancer only, the company have to predict how much capacity they need to keep the performance works well.

Therefore, in the context of PFC, which has ever-expanding demand on computing requirement, it's recommended to use the mix of auto-scaling and load balancer (see in Figure 9).
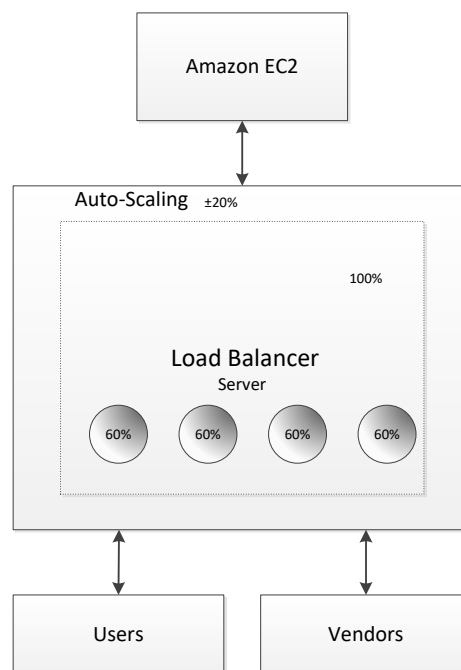

Figure 9. Auto-Scaling and Load Balancer pair solution

5. DevOps

DevOps is an integration term of development systems and IT operations, that allows development, test and production environments to be built, tested and deployed, with shorter system development life cycle and higher software quality. It provides a light communication between different independent services. Each service is independently deployed to the production environment, which is corresponding to the specific business. For instance, with DevOps, the source code is automatically constructed and tested when user submits the code to the codebase. Whenever the code is changed, a new system will be constructed to support the test process separately.

The key elements of DevOps are listed here:

● Micro scale. The complicated task is divided into several micro but complete tasks, which are convenient for developer teams to finish.

● Independence. Every microservice has its own code library and test organization, without worrying about negative interaction between other services.

● Lightness. For efficient purpose, the communication between microservices is designed to be light.

● Loose. The microservices should be loose coupling for independently deployment.

References

1. The NIST definition of Cloud Computing

2. Market share: IT Service, Worldwide 2019, 13 April 2020

3. Jiong J Gu, Cloud Computing Architecture Technologies & Practice, second edition

4. API Gateway Quick Start Document Version: 20200928 Alibaba Cloud API Gateway Quick Start

5. Huanhuan Xiong , Frank Fowley, Scalable Architectures for Platform-as-a-Service Clouds: Performance and Cost Analysis