Design Engineering

ISSN: 0011-9342 | Year 2021
Issue: 9 | Pages: 3410 - 3420

State of the Art in Cloud Computing for Large-Scale E-Commerce

Archana Kumari, Babu Rao K
CMR University, Bengaluru, Bengaluru, India
archana.17phd@cmr.edu.in, baburao.k@cmr.edu.in

Abstract

Cloud technologies are proving to be resourceful in solving many of infrastructure & platform relatedchallenges faced in the E-Commerce industry. The continuously evolving "Cloud Computing" is pushing the E-Commerce architectures towards reimagining and rebuilding of the majority of its tech-stack from the bottoms up.The platform components such as scalable web-interfaces, application specific databases& load balancers etc. are the kind "commodity product" offered by Cloud Service Providers (CSPs) and these might be quite useful in building the cutting-edgeE-commerce platforms. In this paper, we are surveying the cloud impact on *"Large-scale E-Commerce Platform designs"*, by going through literature available in public domain (both academic research papers as well as technical materials published at CSPs portals). We will first discuss the relevant services and techniques offered by the CSPsfor E-Commerce, and then we will review how these services may have potential reach across spectrum of E-Commerce. In the end, we will cover, how the individual architectural components (like backend client-gateways, databases) can be redesigned using Cloud technologies in more scalable and efficient manner.

Index terms—E-Commerce, Cloud Service Providers, IaaS, PaaS, FaaS, Cloud Native, Scalability, Time to market

I. INTRODUCTION

History of internet-basedE-Commercedates back tomid-1990s when amazon, e-bay and others started to setup their online E-Commerce platforms[1]. Evolution of the E-Commerce ecosystem has self-evidently been linked with the evolution of internet technologies in past two and half decades. What started as purely a website-based interface to purchase and sell products, has over the period,grown into a complex world of large suite of products available online, sophisticated user interfaces and datacentres running across the globe to serve the compute requirements [2]. The journey of this growth is partly shared with the advancements in Cloud technologies, as with newdevelopments in Cloudworld, it helps theE-Commerce industry, by taking care ofsome of heavy lifting tasks like infra-structure management and thus allowing the E-Commerce owners to focus more on their products. Essentially, Cloud technologies areoffering solutions to some of the common challenges faced in the E-Commerce industry like scalability and cost of running

ISSN: 0011-9342 | Year 2021
Issue: 9 | Pages: 3410 - 3420

the plant. In next few sections, we'll discuss about those challenges and see how "Cloud computing" can address them by approaching them "bottoms up".

II. E-COMMERCE Challenges for Cloud Computing

These are few critical challenging areasfor modern E-commerce platforms and where "Cloud Computing"might be an effective technology to handle the challenges[3][4]:

- *Scalability& Durability:* With growing E-Commerce traffic, it's crucialthat theE-Commerce platform is able to "scale-up"for the increased load without any deterioration in the quality of services it offers to its users. Similarly, "Durable" platforms are thosehaving in-built handling to cope-up with any hardware or software failures, and without durability, there is a risk of being out of market in event of any system failures.
- *Growing need for sophisticated features:* Features like *Real-time recommendations*, *social media integrations* and *Personalization* etc., E-Commerce tends to become more sophisticated over time.
- *Cost of Running the Plant*: The *Profitability* just not depends on providing new features and on boarding more customers, but it also requires optimizing the cost of running and operating the plant. The cost includes, paying the bills for compute, storage and network, as well as, paying for the engineering expertise required to run the operations.
- *Time to market*: Lesser time taken to add new feature or functionality is a subtle differentiator among competing E-Commerce owners. Some architectures/designs are more flexible over others, when it comes to making system changes*(for example micro-services over monolithic, cloud managed services over deploying and managing "on premise" systems)* and thus theyoffer superior *time to market* compared to others.

III. Cloud Computing: Technology&SERVICES

Cloud computing presentsdifferent flavours of services and computation models to serve various functional and non-functional requirements of the businesses. In this section we will review some of the important services and techniques offered by CSPs which are relevant for handling the challenges described previously.

A. IaaS: Cloud hostings

IaaS (Infrastructure as service) is *"Cloud based Hosting"*service from the vendors. The most popular ones are EC2 (from AWS), Google Compute Engines (from GCP) and Droplets (from Digital Ocean).These are essentially "CPU" given on rent by the cloud vendors for any general-purpose use case (e.g. running webservers, computation enginesand databases or deploying any business specific logic). Opting forthe IaaS solutions, provides several benefits over physically owning the hardware such as:

ISSN: 0011-9342 | Year 2021
Issue: 9 | Pages: 3410 - 3420

- Near real-time procurement and easy decommissioning of servers.
- Convenient Host replication by using "Installation Images" (example: AMIs/Packer/Bundles)
- Out of box health monitoring and maintenance done by the vendors
- Ability to automatically scale in case of high traffic*(Autoscaling feature)*

Due to above-described advantages, these *"Cloud hosted machines"*offerefficient infrastructure management compared to physically owned the host, having said, IaaS is considered a bit of basic offering compared to CSP managed services like databases & orchestrated runtime etc. (popularly termed as *Cloud Native*)

B. PaaS &FaaS: Cloud native compute engines and Serverless computing

Platform as a Service (PaaS)

Container Orchestration: Containersessentially provide lightweight VM like functionality without much of hostingoverheads, their orchestration, by definition, is the automation of deployment, management and scaling processes. GCP defines it as, *"Containers offer a logical packaging mechanism in which applications can be abstracted from the environment in which they actually run."[5]*

Containers have several advantages like they provide isolation with run anywhere capabilities, they provide ease of scalability of actual application deployment (compared to ease of infrastructure scalability offered by IaaS solutions) [6].Additionally, Containers are also critical to building micro-services in cloud by abstracting out the environmental dependencies, and thus helping smooth deployment of pieces of complex workflows represented by microservices.

Function as a Service (FaaS)

With FaaS, business logic developers can focus on their core product features and business logic instead of responsibilities like operating system (OS) dependencies, access control, OS patching, provisioning, right-sizing, scaling, process deployment management, process scheduling. "Cloud Functions" and "AWS Lambda" are the most popular choices for the FaaS model, which implement "Pay-As-Per-Function-Call" pricing model.

Below table (table 1) draws comparison of various cloud computing models on what is already managed by cloud service providers for the cloud users and what is not.

ISSN: 0011-9342 | Year 2021
Issue: 9 | Pages: 3410 - 3420

Table 1. "CSPs" delivered features across IaaS, PaaS &FaaS

	Physical Ownership in IaaS.	Cloud Vendor Data-center		
		Rented Servers	Containers	Orchestless compute
Business Logic & Data	✕	✕	✕	✕
Application Runtime	✕	✕	✕	✔
Application environment management	✕	✕	✔	✔
OS & Host Management	✕	✔	✔	✔
Host Procurement	✕	✔	✔	✔

Storage

CSPs typically offer *Object Store* to allow user save and retrieve their files using web-APIs. There could be variety of usage of the cloud storage, but the most popular usagesis to serve static content for the websites, i.e. webserver can serve the HTMLs with URI linked to the resources in the *Object Stores*. Doing this, reduces the burden to serve static content on the webserver and improves the overall site performance. Other popular usages of this service to store backups of databases, logs and analytical data. The most popular cloud products under this category are AWS "S3", GCP "Cloud-storage" and Digital Ocean's "Cloud Space."[7]

CDN

A content delivery network (CDN) is group of servers and data which is distributed geographically across the globe and help minimize the latency of the internet requests and improves overall user experience.As per Cloudflare, "It allows for the quick transfer of assets needed for loading Internet content including HTML pages, JavaScript files, stylesheets, images and videos." [8]

Databases

The purpose of a database is to persist and retrieve business data in a structured form. There are various types and use-cases of "managed" database services offered by CSPs, few of them are listed as below (in table 2)[9]:

ISSN: 0011-9342 | Year 2021
Issue: 9 | Pages: 3410 - 3420

Table 2. Cloud Database Services & Use-cases

Database Type	Use-cases
RDBMS	Traditional applications, ERP, CRM, e-commerce
Key-Value Store	High-traffic web apps, e-commerce systems, gaming applications
In-memory DBs	Caching, session management, gaming leader boards, geospatial applications.
Document DBs	Content management, catalogues, user profiles
Wide Columns DBs	High scale industrial apps for equipment maintenance, fleet management, and route optimization
Graph Dbs	Fraud detection, social networking, recommendation engines
Time Series Dbs	IoT applications, DevOps, industrial telemetry
Ledger Dbs	Systems of record, supply chain, registrations, banking transactions

Other services

Beside the above-described services, CSPs provides services which can be used, off the shelf, to build large scale, high throughput & distributed applications.

- Load balancers
- Queueing services
- Authentication and Authorization services
- End users' notifications
- Infrastructure monitoring and alerting services

ISSN: 0011-9342 | Year 2021
Issue: 9 | Pages: 3410 - 3420

C. E-CommerceTypes and Overview of Cloud Adoptions

There can be arguably different ways to classify the businesses of E-Commerce, though the most popular ones are based on participants, comprising producing and consuming ends [10]. We are going to review the types and we will briefly cover the suitability of cloud technologies in making these more scalable, easing the workflow and overall making the E-Commerce operation more efficient.

Business - to- BusinessE-Commerce (B2B)

B2B E-Commerce takes place between businesses. Popular commercial examples which fall under this category are:

- CRM/Salesforce
- Online ticketing
- Cloud service providers interfaces used by business

B2B E-Commerce tend to have complex domain-specific workflows and implementations. The cutting-edge implementations of the same follows micro services-based design and they take advantages of cloud offerings such as IaaS/PaaS products, Relational & Ledger databases, and Queueing & Notification services. Essentially, using cloud can result into better *Time to market* for implementing new features and cutting down the operational costs.

Business - to - ConsumerE-Commerce (B2C)

Business to consumer E-Commerce is the most voluminous E-Commerce activity, where business produces goods and services for end consumer to consume. Typical examples are:

- Online ticketing for hotels and transportations*(e.g.,Airbnb, Google flights)*
- Online retail E-Commerce*(e.g.,Amazon, Flip Kart)*
- Over the top (OTT) media services*(e.g.,Netflix, YouTube)*
- Online payments services such as UPI.

These volumes are purely because of the vast user base,something that can potentially go beyond several millions [11]. Additionally, these businesses tend to have intermittent high volumes, for example high activity period during festive seasons, more people tend to travel during Christmas and New Year season. Essentially, these platforms should be built with ability to horizontally scale while keeping the cost of operation less. In order to achieve the same, new generation E-Commerce platforms must take advantages of many of cloud services, such as, load-balancers, Serverless computes, all sort of databases, Queuing and Notifications services.

ISSN: 0011-9342 | Year 2021
Issue: 9 | Pages: 3410 - 3420

Consumer - to - Consumer E-Commerce (C2C)

In this model of E-Commerce, consumers or individuals play the role of both producer and receiver. A good example of this is E-Bay where individuals can buy and sell the items. From the architecture and scalability point of view, these systems face similar challenge like B2C models and can reap similar benefits from Cloud Adoptions.

Other types

There are other types of E-Commerce like consumer to business (C2B), Business to Governments etc. We believe that, from architectural or design considerations, these follow patterns similar to earlier described types and differ mostly in domain specific use-case.

D. E-Commerce Components: Cloud Solutions & Architecture

In this section of the paper, we are going to discuss some of the components of the E-Commerce platforms which are frequently used across different types and use cases of E-Commerce.

Customer UI

UIs are the entry point for end-user to interact with the E-Commerce platforms. Websites and mobile app are the two kinds of UIs, typically used by the end clients. Customer UIs run on user devices and hence these are agnostic to backbends which may run on clients“on premise” or “cloud-infrastructure”.

Web hosting/Web APIs/Load Balancing

“Web Server” are client’sgateways tothe E-Commerce backends. It’s extremely likely that, in-case of large-scale platform, there are several Webservers parallelly catering to the client requests. The web-servers are typically placed behind the load balancer(s) engines, which receive requests from client and route to any of the available web-servers. Load balancers are off the shelf solution provided by CSPs. It’s understandable for this feature to be hosted on IaaSenvironments, like AWS EC2, where user decides to run LAMP, MEAN, Python’s Django/Flask, Java’s Tomcat/Jetty stack. This approach is especially quite popular across the cloud users who are migrating from their “on-premises” infrastructure to cloud based one. There are CSPs who provide webhosting as PaaS, for example GCP provide, out of box managed PaaS for LAMP & MEAN Stacks[12]. When using this approach for web-hosting, users typically use “auto-scaling” for scalability and durability, a feature which is not available for non-cloud solutions.

The “Cloud-native” way to build web-application is based on web-APIs. Typical architecture for the same is below:

ISSN: 0011-9342 | Year 2021
Issue: 9 | Pages: 3410 - 3420

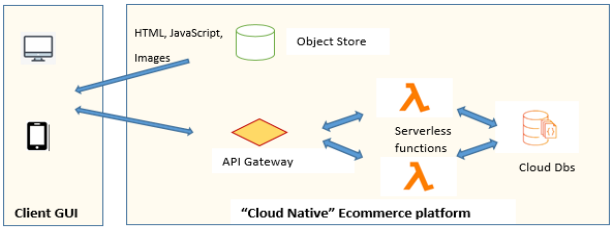

Figure 1. Web API based "Serverless"web-hosting

In the above diagram, Figure 1, Client GUI receives HTML, JavaScript and other static content from object stores (like AWS S3). Requests for dynamic contents are routed to an API gateway which is backed by a farm of Serverless computes (FaaS) and Cloud databases. This approach provides scalabilityand durability which is transparent to the developers,and fully managed by CSPs, compared to IaaS, in which a user needs to be very cognizant to scaling needs (even when onboarded with Autoscaling like feature).

Catalogue management

Catalogue provides a consistent view of items offered by a retailer and allows a customer to search and place an order using a client GUI. [13]. Cloud native solutions offer "Document DBs" as managed services, as Document DBs are quite efficient and effective for storing catalogue information. Cloud offered "Document Dbs" are quite convenient and auto-scalable compered to owning and managing one (e.g.,Mongo DB) on premise datacentre.

Order management

E-Commerce client orders are OLTP (Online transaction processing) data. End users, through their Client GUI send the request to purchase (and also Sell in C2C/C2B E-Commerce) goods and services. From *receiving order* to *fulfilment* to *completion*, order management is all about managing the state of the order and triggering various action depending on the state of the order. Since the nature of this data is OLTP, Relational databases are the most popularkind of databases to implementthis. Most of the CSPs provide "fully-managed" & "State of the art" relational DBs, which can be utilized for this purpose.

Historically the E-Commerce developers have been building the *order management* as monolithic state management engines, which typically, by design, are hard to scale. In past few years, a trend has started to break the complex monoliths into simpler microservices. Cloud computing offers off the shelf notification and queuing services, which can be used to implement popular "event-sourcing design pattern" based microservices[14]. This improves the overall throughput by distributing and parallelizing the work, while easing out the responsibility to build and maintain one's own queuing mechanism (like home grown Kafka).

ISSN: 0011-9342 | Year 2021
Issue: 9 | Pages: 3410 - 3420

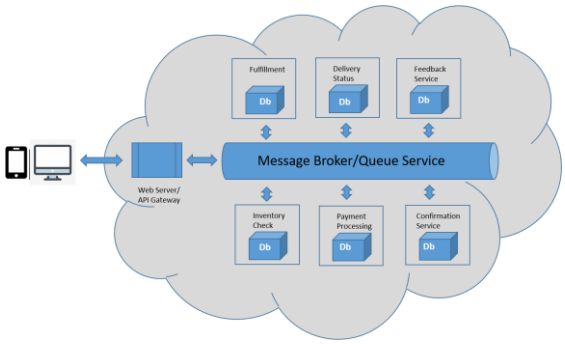

Figure 2. Event Sourcing- Micro Services Pattern for Order Management

Above diagram (Figure 2) depicts a typical "microservice" based order management on Cloud. Different functionalities like "Inventory Check", "Payment Processing", and "Fulfilment" etc. are simplified into independent services. They interact with each other using a messaging broker. The entire architecture can be built natively on cloud by utilizing containers, queues, databases and notification services.

Data Warehousing

According to [15], an E-Commerce Data Warehouse is a critical component for the growth of the brands. E-Commerce businesses operate at varying degrees of maturity when it comes to infrastructure, data management, management reporting, operational reporting, analytics, etc. Cloud Service Providers like Google (GCP) and Amazon (AWS) provide "managed Data Warehouse service" called Google Big Query& RedShift respectively for this purpose. Use of these services eliminates all the need for managing a data warehouse, eliminates all the complexities associated with hiring or contracting a database administrator to manage the data warehouse.

Real-time Analytics

Real-time Analytics in the field of E-Commerce is a relatively recent trend, which is gaining popularity to meet following requirement:

- Just in time personalization
- Dynamic pricing of goods and services
- Smart Real-time recommendations
- Fraud detection
- Smart predictions for web-traffic and need for infrastructure scale-up prediction.

Processing of "Streaming data" is at the core of real-time analytics activities. The stream data can come from various data sources like, customer activities, geo-location services, news feed,vendor actions or competitor generator events. Performing this analysis can be at times very compute intensive as volume of streaming data can be very unpredictably high. In his work, [16],

ISSN: 0011-9342 | Year 2021
Issue: 9 | Pages: 3410 - 3420

author advocates Serverless computing for this purpose withthe following argument "cloud stateless functions represent a viable platform for the analytics, with pay-as-you go model and by eliminating cluster management overhead, and fulfilling the promise of elasticity." [17-19].

IV. Conclusion

In this paper we have reviewed some of the building blocks of modern E-Commerce application using the latest cloud technologies. We discussed multiple cloud alternatives for components like Web-interfaces, Serverless computing, Databases and other relevant services as we understand that there is no one size fits all solution for kinds of problems faced in the E-Commerce world. Though, it's relatively easy to build a scalable and a maintainable E-Commerce using "state of the art" cloud offerings, yet before opting for a cloud service forany specific use-case, factors like the total cost of implementation and the cost of ownership must be evaluated carefully.Having said that, it's reasonable to conclude that adoption of cloud computingis apromising way forward towardssolving some of the key challenges likescalability, durability and simplification of large-scale software architecture. Not just that, overall "time to market"to build new features like personalization&large-scale real-time analytics is expected to be less when built on top of the cloud services. Additionally, cloud adoption can also greatly simplify how E-Commerce platforms are deployed and their runtimes are handled. As we know, both E-Commerce and cloud computing are growing at fast rate and we can expect newer things like introduction of AR/VR (Augmented and Virtual realities), Voice E-Commerce etc. to become popular in the E-Commerce domain. Although, these new areas need further work, one thing is quite certain that, for any new trend to become part of mainstream, it mustbe available on cloud platforms in form of "cloud service".

References

[1] Joe McFerrin, "The History of E-Commerce: How Did it All Begin?"*https://www.iwdagency.com/blogs/news/the-history-of-E-Commerce-how-did-it-all-begin*, extracted in 2020

[2] "Global Infrastructure" in *https://aws.amazon.com/about-aws/global-infrastructure/?p=ngi&loc=1,*extracted in 2020

[3] Arya M Nair, "CLOUD COMPUTING IN E-COMMERCE", *InternationalResearch Journal of Engineering and Technology (IRJET)*vol. 07 issue: 04, pp. 2936- 2940, Apr 2020

[4] Hukam Chand Saini et.al, "Benefits of Cloud Computing for Business Enterprises: A Review", *International Conference on Advancements in Computing & Management (ICACM-2019)* pp. 1003- 1003, Apr 2019

[5] CONTAINERS AT GOOGLE, *https://cloud.google.com/containers*, extracted in 2020

[6] "Amazon Elastic Container Service", *https://aws.amazon.com/ecs/*, extracted in 2020

[7] "Cloud file Storage", *https://aws.amazon.com/what-is-cloud-file-storage/* , extracted in 2020

ISSN: 0011-9342 | Year 2021
Issue: 9 | Pages: 3410 - 3420

[8] "What is a CDN? | How do CDNs work?"*https://www.cloudflare.com/learning/cdn/what-is-a-cdn/*, extracted in 2020

[9] "Databases on AWS",*https://aws.amazon.com/products/databases/*, extracted in 2020

[10] Rania Nemat (2011), "Taking a look at different types of e-commerce", *World Applied Programming*, vol. 1, no. 2, pp 100-104| June 2011. 100-104 ISSN: 2222-2510

[11] "Eight Largest Ecommerce Companies", *https://anova.co/blog/eight-largest-ecommerce-companies*, extracted in 2020

[12] "GOOGLE CLOUD MARKETPLACE" *https://cloud.google.com/marketplace*, extracted in 2020

[13] Gautham K. Acharya et al. "Cloud Customer Architecture for e- Commerce", *Cloud Standards Customer Council*, 2016

[14] "Event Sourcing pattern", *https://docs.microsoft.com/en-us/azure/architecture/patterns/event-sourcing*, extracted in 2020

[15] "Why an E-Commerce Data Warehouse is essential for growing brands?" *https://sarasanalytics.com/blog/why-E-Commerce-data-warehouse-is-essential*,Apr 2020

[16] Qifan Pu, "Towards Practical Serverless Analytics", PhD dissertation, Computer Science at University of California, Berkeley, 2019

[17] Avesh, Mohd, and Rajeev Srivastava. "Modeling simulation and control of active suspension system in Matlab Simulink environment." In *2012 Students Conference on Engineering and Systems*, pp. 1-6. IEEE, 2012.

[18] Avesh, Mohd, and Rajeev Srivastava. "Passenger car active suspension system model for better dynamic characteristics." *National Academy Science Letters*43, no. 1 (2020): 37-41.

[19] Mohd, Avesh, and Rajeev Srivastava. "Parametric optimization to design a passenger car suspension system for better dynamic performance."*European Transport/TrasportiEuropei*71 (2018): 1-14.