

Analysis of Big Data and Cloud Services with Experimental of Manually and Automated Testing

Prof.(Dr.) Rajender Kumar

Research Scholar, D.Sc Computer Science & Applications (Post-Doctoral Fellow), Desh Bhagat University, Punjab India

ABSTRACT

Big Data has increased much focus from the scholastic world and the IT business. In the advanced and figuring world, all together is created and gathered at a rate that quickly surpasses the limit go. Right now, more than 2 billion individuals worldwide are associated with the Internet, and more than 5 billion people possess cell phones. By 2020, 50 billion gadgets are relied upon to be associated with the Internet. Now, anticipated information creation will be 44 times more prominent than that in 2009. As data is exchanged and shared at light speed on optic fiber and remote systems, the volume of information and the speed of market development increment. In any case, the quick development rate of such substantial information creates various difficulties, for example, the fast development of information, exchange speed, different information, and security. In any case, Big Data is still in its outset arrange, and the space has not been checked on all in all. Distributed computing has opened up new open doors for testing offices. New innovation and social network patterns are making an ideal tempest of chance, empowering cloud to change inside tasks, Customer connections and industry esteem chains. To guarantee high caliber of cloud applications being worked on, designer must perform testing to analyze the quality and exactness whatever they plan. In this examination paper, we address a testing natural engineering with important key advantages, to perform execution of experiments and utilized testing strategies to improve nature of cloud applications.

Keywords : Big data, data analytics, data management, big data-as-a-service, analytics-as-a-service, business intelligence lease, storage cloud computing, cost-benefit analysis model.

Article Info

Publication Issue :

Volume 8, Issue 6

November-December-2022

Page Number : 224-235

Article History

Accepted: 05 Nov 2022

Published: 22 Nov 2022

I. INTRODUCTION

The Internet immersion constantly increments, as an ever-increasing number of individuals peruse the

Web, utilise email and informal organisation applications to chat with each other or access remote interactive media administrations, for example, versatile TV. Besides, a few requesting portable

system administrations are currently accessible, which require expanded information rates for particular tasks, for example, gadget stockpiling synchronisation to distributed computing servers or high-determination video. The entrance to such a worldwide data and correspondence framework, alongside the advances in computerized sensors and capacity have made a lot of information, for example, Internet, sensor, gushing or cell phone information[1]. Moreover, information examination is the reason for examinations in numerous fields of learning, for example, science, building or administration. Not at all like online enormous information, area information is a fundamental part of versatile huge information, which are bridled to upgrade and customize portable administrations. Thus, a period where information stockpiling and processing progress toward becoming utilities that are pervasively accessible is currently presented. Moreover, calculations have been produced to associate datasets and empower more advanced investigation. Since developments in information design are on our doorstep, the 'enormous information' worldview alludes to vast and complex informational collections (i.e., petabytes and exabytes of information) that conventional information preparing frameworks are insufficient to catch, store and break down, trying to gather insight from information and make an interpretation of it into upper hand. Subsequently, huge information needs all the more processing force and capacity gave by distributed computing stages. In this specific situation, cloud suppliers, for example, IBM, Google, Amazon and Microsoft, give arrange available capacity valued by the gigabyte-month and processing cycles evaluated by the CPU- hour[2].

II. BIG DATA: BACKGROUND AND ARCHITECTURE

IBM information researchers battle that the key size of enormous information are the "4Vs": volume, speed,

assortment and veracity. As extensive and little ventures continually endeavor to plan new items to manage enormous information, the open source stages, for example, Hadoop, give the chance to load, store and question a gigantic size of information and execute progressed huge information investigation in parallel over a conveyed bunch. Clump preparing models, for example, Map Reduce, empower the information administration, blend and handling from various sources[3]. Numerous enormous information arrangements in the market misuse outside data from a scope of sources (e.g., interpersonal organizations) for demonstrating and feeling investigation, for example, the IBM Social Media Analytics Software as a Service arrangement. Cloud suppliers have just started to build up new server farms for facilitating long range interpersonal communication, business, media content or logical applications and administrations. Toward this path, the determination of the information distribution center innovation relies upon a few variables, for example, the volume of information, the speed with which the information is required or the sort of investigation to be performed.[4] Theoretical huge information stockroom design is displayed. Another critical test is the conveyance of enormous information abilities through the cloud. The appropriation of enormous information as-a-benefit (BDaaS) plans of action empowers the powerful stockpiling and administration of extensive informational collections and information handling from an outside supplier, and also the misuse of a full scope of investigation abilities (i.e., information and prescient examination or business knowledge are given as administration based applications in the cloud). In this specific circumstance, Zheng et al. basically audit the administration produced enormous information and huge information as-a-benefit the proposition of a framework to give usefulness to overseeing and breaking down various sorts of administration created huge information[5]. A major information as-a-benefit structure has been likewise utilized to give

enormous information administrations and information investigation results to clients, upgrade proficiency and lessen cost[6].

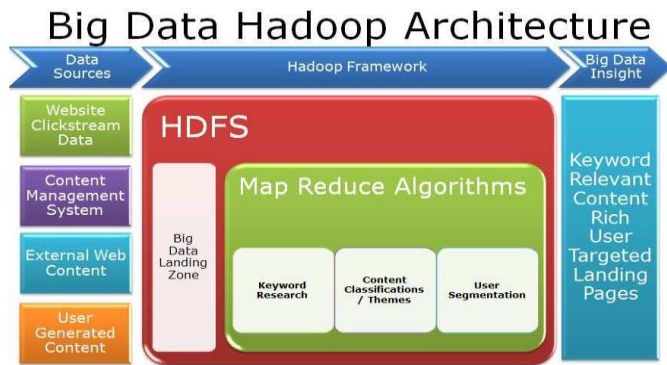


Fig. 1: A conceptual big data warehouse architecture

Data increments quickly at a rate of 10x at regular intervals [6]. From 1986 to 2007, the worldwide capacities with respect to mechanical information stockpiling, calculation, preparing, and correspondence were followed through 60 analogs and advanced innovations in 2007, the limit with respect to capacity as a rule reason PCs was 2.9×10^{20} bytes (ideally packed) and that for correspondence was 2.0×10^{21} bytes. These PCs could likewise oblige 6.4×10^{18} guidelines for every second. Be that as it may, the figuring size of broadly useful PCs increments every year at a rate of 58%.[7] In computational sciences, Big Data is a basic issue that requires genuine consideration. Up to this point, the basic scenes of Big Data have not been brought together. Moreover, Big Data can't be prepared utilizing existing innovations and strategies[8]. Therefore the age of boundless information by the fields of science, business, and society is a worldwide issue. As for information examination, for example, systems and standard apparatuses have not been intended to look and dissect extensive datasets[9]. Therefore, associations experience early difficulties in making, overseeing, and controlling huge datasets. Frameworks of information replication have likewise shown some security shortcomings as for the age of numerous duplicates, information administration, and strategy.

These approaches characterize the information that are put away, broke down, and got to. They additionally decide the significance of these information. To process unstructured information sources in Big Data ventures, concerns with respect to the versatility, low idleness, and execution of information foundations and their server farms must be tended to[10]. In the IT business all in all the fast ascent of Big Data has produced new issues and difficulties as for information administration and investigation. Five normal issues are volume, assortment, speed, esteem, and multifaceted nature as per in this examination, there are extra issues identified with information, for example, the quick development of volume, assortment, esteem, administration, and security[11]. Each issue speaks to a significant issue of specialized research that requires discourse. Henceforth, this exploration proposes an information life cycle that uses the advances and wordings of Big Data. Future research bearings in this field are resolved in light of chances and a few open issues in Big Data control. This gatherings the basic issues in Big Data into three classifications in view of the shared characteristic of the test[12].

A. Volume of Big Data

The information compose that increments most quickly is vague information. This information write is described by "human data, for example, top notch recordings, films, photographs, logical reenactments, monetary exchanges, telephone records, genomic datasets, seismic pictures, geospatial maps, email, tweets, Face book information, call-focus discussions, cell phone calls, site clicks, archives, sensor information, telemetry, therapeutic records and pictures, climatology and climate records, log documents, and content[14]. As indicated by Computer World, unstructured data may represent over 70% to 80% of all information in associations. These information, which for the most part begin from online networking, constitute 80% of the

information worldwide and represent 90% of Big Data. As of now, 84% of IT chiefs process unstructured information, and this rate is relied upon to drop by 44% soon. Most vague information are not demonstrated, are arbitrary, and are hard to break down. For some, associations, proper methodologies must be created to oversee such information.[15] Table 1 portrays the fast creation of information in different associations further. As indicated by Industrial Development Corporation (IDC) and EMC Corporation, the measure of information created in 2020 will be 44 times more prominent [40 zetta bytes (ZB)] than in 2009. This rate of increment is relied upon to hold on at half to 60% yearly. To store the expanded measure of information, HDDs must have substantial capacity limits. In this way, the accompanying segment explores the advancement rate of HDDs [16].

B. Rapid Growth of Data

The data compose that additions most rapidly is hazy data. This data create is portrayed by "human information, for instance, top quality accounts, films, photos, consistent entertainments, cash related trades, phone records, genomic datasets, seismic pictures, geospatial maps, email, tweets, Face book data, call-center dialogs, mobile phone calls, site clicks, reports, sensor data, telemetry, remedial records and pictures, climatology and atmosphere records, log archives, and substance. As demonstrated by Computer World, unstructured information may speak to more than 70% to 80% of all data in affiliations. These data, which generally begin from web based systems administration, constitute 80% of the data worldwide and speak to 90% of Big Data. Starting at now, 84% of IT bosses process unstructured data, and this rate is depended upon to drop by 44% within the near future. Most obscure data are not shown, are unpredictable, and are difficult to analyze. For a few, affiliations, appropriate strategies must be made to direct such data. Table 1 delineates the speedy age of data in various affiliations further. According to

Industrial Development Corporation (IDC) and EMC Corporation, the measure of data made in 2020 will be 44 times more critical [40 zettabytes (ZB)] than in 2009. This rate of augmentation is required to drive forward at half to 60% consistently. To store the extended measure of data, HDDs must have huge limit limits. Thusly, the going with section investigates the change rate of HDDs[19].

C. Development Rate of Hard Disk Drives (Hdds)

The enthusiasm for electronic limit is exceptionally adaptable. It can't be completely met and is controlled just by spending designs and organization capacity and utmost. Goda et al. (2002) and discuss the recorded scenery of limit devices, starting with appealing tapes and plates and optical, solid state, and electromechanical devices. Going before the automated revolt, information was commonly secured in straightforward tapes as showed by the available bits. Beginning at 2007, in any case, most data are secured in HDDs (52%), trailed by optical limit (28%) and propelled tapes (roughly 11%). Paper-based limit has dwindled 0.33% out of 1986 to 0.007% of each 2007, notwithstanding the way that its capacity has persistently extended (from 8.7 preferably compacted PB to 19.4 in a perfect world pressed PB). Figure 2 depicts the fast headway of HDDs around the globe.

III. BIG DATA MANAGEMENT

The plan of Big Data must be synchronized with the assistance establishment of the affiliation. To date, most of the data used by affiliations are lethargic. Data is logically sourced from various fields that are disturbed and tangled, for instance, information from machines or sensors and sweeping wellsprings of open and private data. Effectively, most associations were not capable either catch or store these data, and available instruments couldn't manage the data in a sensible measure of time. In any case, the new Big Data development improves execution, energizes headway in the things and organizations of plans of

activity, and gives essential administration support [8, 48]. Tremendous Data development hopes to restrain gear and taking care of expenses and to check the estimation of Big Data beforehand giving basic association resources. Suitably directed Big Data are open, tried and true, secure, and sensible. Hereafter, Big Data applications can be associated in various complex consistent controls (either single or interdisciplinary), including barometrical science, astronomy, sedate, science, genomics, and biogeochemistry. In the going with region, we rapidly look at data organization mechanical assemblies and propose another data life cycle that uses the progressions and wordings of Big Data.

A. Management Tools

With the progression of figuring advancement, enormous volumes can be administered without requiring supercomputers and high cost. Various instruments and methodologies are available for data organization, including Google Big Table, Simple DB, Not Only SQL (NoSQL), Data Stream Management System (DSMS), MemcacheDB, and Voldemort. Regardless, associations must make uncommon devices and advancements that can store, get to, and examine a considerable measure of data in close consistent in light of the way that Big Data shifts from the regular data and can't be secured in a single machine. Also, Big Data does not have the structure of traditional data. For Big Data, without a doubt the most normally used gadgets and methods are Hadoop, Map Reduce, and Big Table. These progressions have rethought data organization since they effectively process a great deal of data gainfully, cost reasonably, and in a lucky way. The going with section depicts Hadoop and Map Reduce in moreover detail, and furthermore the distinctive endeavors/frameworks that are related to and sensible for the organization and examination of Big Data.

B. Hadoop

Hadoop is engraved in Java and is a best level Apache wander that started in 2006. It complements disclosure from the perspective of versatility and examination to recognize close amazing achievements. Doug Cutting made Hadoop as a collection of open-source stretches out on which the Google Map Reduce programming condition could be associated in a scattered system. Straightforwardly, it is used on broad totals Hadoop is made out of HBase, HCatalog, Pig, Hive, Oozie, Zookeeper, and Kafka; in any case, the most understood parts and without a doubt comprehended norms are Hadoop Distributed File System (HDFS) and Map Reduce for Big Data. The Hadoop organic group, and furthermore the association of various fragments to each other.

HDFS

This perspective is associated when the measure of data is unnecessarily for a singular machine. HDFS is more personality boggling than other record structures given the complexities and vulnerabilities of frameworks. Bundle contains two sorts of centers. The foremost center point is a name-center that goes about as an expert center. The second center point compose is a data center point that goes about as slave center point. This sort of center point comes in items. Alongside these two sorts of center points, HDFS can similarly have helper name-center point. HDFS stores archives in prevents, the default square size of which is 64MB. All HDFS records are duplicated in items to energize the parallel getting ready of a great deal of data.

C. HBase

HBase is a supervision structure that is open-source, shaped, and passed on in perspective of the Big Table of Google. This structure is segment rather than segment based, which animates the execution of exercises over practically identical regards transversely finished enormous enlightening accumulations. For example, read and form assignments incorporate all lines yet only a little

subset of all segments. HBase is accessible through application programming interfaces (APIs, for instance, Thrift, Java, and legitimate state trade (REST). These APIs don't have their own request or scripting vernaculars. Normally, HBase depends absolutely on a Zoo Keeper event.

D. ZooKeeper

ZooKeeper takes care of, orchestrates, and names a great deal of data. It moreover gives coursed synchronization and social event organizations. This event engages appropriated techniques to supervise and add to each other through a name space of data registers (- centers) that is shared and different leveled, for instance, a record system. Alone, ZooKeeper is a scattered organization that contains pro and slave center points and stores setup information.

E. HCatalog.

HCatalog regulates HDFS. It stores metadata and produces tables for a ton of data. HCatalog depends upon Hive meta store and consolidates it with various organizations, including MapReduce and Pig, using a run of the mill data illustrate. With this data appear, HCatalog can in like manner develop to HBase. HCatalog reworks customer correspondence using HDFS data and is a wellspring of data sharing among instruments and execution stages.

F. Hive

Hive structures stockrooms in HDFS and other data sources, for instance, Amazon S3. Hive is a sub arrange in the Hadoop natural framework and produces its own specific inquiry lingo (HiveQL). This lingo is amassed by MapReduce and enables customer described limits (UDFs).The Hive organize is primarily in perspective of three related data structures: tables, portions, and bowls. Tables contrast with HDFS lists and can be appropriated in various fragments and, over the long haul, buckets

G. Pig

The Pig framework makes an irregular state scripting tongue (Pig Latin) and works a run-time organize that engages customers to execute MapReduce on Hadoop. Pig is more adaptable than Hive with respect to potential data orchestrate given its data show. Pig has its own specific data compose, outline, addresses semi sorted out data, including JSON and XML.

H. Mahout

Mahout is a library for machine-learning and data mining. It is apportioned into four essential social occasions: total isolating, course of action, batching, and mining of parallel visit outlines. The Mahout library has a place with the subset that can be executed in a passed on mode and can be executed by MapReduce.

I. Oozie

In the Hadoop structure, Oozie masterminds, executes, and administers work stream. It is melded into other Apache Hadoop structures, for instance, Hive, Pig, Java MapReduce, Streaming MapReduce, and Distcp Sqoop. Oozie combines exercises and driving forces Hadoop assignments using an organized non-cyclic graph (DAG). This model is routinely used for various endeavors.

J. Avro

Avro serializes data, conducts remote strategy calls, and passes data beginning with one program or vernacular then onto the following. In this framework, data are self-portraying and are continually secured in light of their own example in light of the fact that these attributes are particularly suited to scripting vernaculars, for instance, Pig.

K. Chukwa

At display, Chukwa is a framework for data aggregation and examination that is related to MapReduce and HDFS. This framework is at the present time progressing from its change arrange.

Chukwa accumulates and methods data from appropriated structures and stores them in Hadoop. As an independent module, Chukwa is fused into the apportionment of Apache Hadoop.

Table 1 : Hoodoo Components and their Functionalities

Hadoop	Component Functions
1) HDFS	Storage and replication
2) MapReduce	Distributed processing and fault tolerance
3) HBASE	Fast read/write access
4) HCatalog	Metadata
5) Pig	Scripting
6) Hive	SQL
7) Oozie	Workflow and scheduling
8) ZooKeeper	Coordination
9) Kafka	Messaging and data integration
10) Mahout	Machine learning

IV. CLOUD COMPUTING AND SERVICES

Circulated registering has transformed into another handling perspective where the cloud could give both virtualized gear and programming resources that encouraged remotely and give a usage on-ask for advantage illustrate.[22] Conveyed processing offered an ability to get to shared resources and customary structure, which give benefits on ask for completed the framework to perform exercises that meet changing business needs. It offers workplaces to customers to execute, pass on and manage their applications 'on the cloud' which includes virtualization of benefits that jam and accomplishes itself. Cloud testing uses cloud common designing for programming testing. Affiliations looking for after general testing that passing on a couple of challenges like confined test spending design, meeting due dates et cetera. To serve a quality thing, testing is the last response for any kind of issue we would look in future from customer site. Cloud testing is a sort of programming testing wherein testing is done through using resources over cloud applications under the cloud establishment. This is the place cloud testing has created as another approach to manage testing where circulated processing conditions are used to reproduce certifiable with application's execution, consistency, speed, security and helpfulness[18].

V. ARCHITECTURE SUPPORT FOR CLOUD COST & TESTING

Distributed computing engineering, much the same as some other application or programming, is considered into two principle segments: Front End and Back End. Front end is a customer or any application which is utilizing cloud administrations. Back end is the system of customer machines with servers having PC program and information stockpiling framework. Cloud has brought together server organization to administrate the frameworks customer, requests and so on. When client situations are produced and the test is outlined, and executed. Once the test finished the cloud specialist organization convey comes about and examination back to corporate IT experts through constant dashboards for a total investigation of how their applications and the web will perform during peak volumes[2].

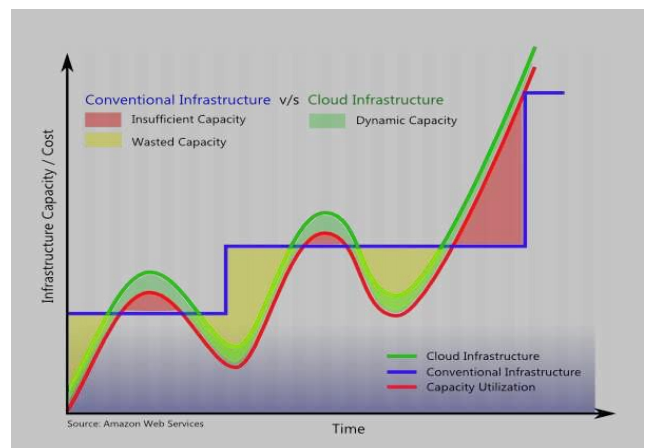


Fig. 2: Infrastructure capacity /cost

VI. VARIOUS TESTING'S TO BE PERFORMED OVER CLOUD APPLICATIONS

Cloud testing is regularly observed as just execution or load tests, be that as it may, as examined prior it covers numerous different kinds of testing. Distributed computing itself is frequently alluded to as Software as a Service (SaaS) and utility figuring. As to test execution, the product offered as an administration might be an exchange generator and the cloud supplier's foundation programming[17].

There are different trying techniques to be performed; we are here utilizing fundamental and general testing approaches:

- Stress test over Cloud application
- Load and execution test over Cloud application
- Functional testing on Cloud application
- Compatibility testing on Cloud application

- Browser Performance testing on Cloud application
- Latency testing on Cloud application
- SOASTA test over Cloud application
- Targeting framework test on Cloud application
- Failover test on Cloud application
- Capacity test on Cloud
- Soak test on Cloud

VII. FORMS OF CLOUD-BASED SOFTWARE TESTING

There are four different forms of cloud-based application or software testing. Each focus on different objectives:

A. Testing a SaaS in a Cloud

It reassures the quality of a SaaS in a cloud based on its functional and non-functional service requirements.[21]

B. Testing of a Cloud

It approves the nature of a cloud from an outside view in light of the given cloud determined abilities and administration highlights.

C. Testing Inside a Cloud

It checks the nature of a cloud from an inside view in light of the interior foundations of a cloud and determined cloud proficiencies. Just cloud sellers can execute this sort of testing since they have access to inside foundations and associations between its inner SaaS and programmed proficiencies, security, administration and screen[16].

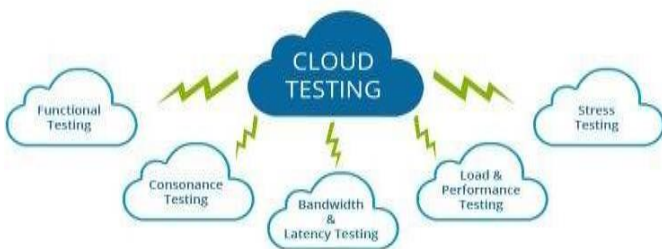


Fig. 3: Cloud based Testing

VIII. CLOUD IBASED AIR TICKET MANAGEMENT SYSTEM (ATMS MODULE)

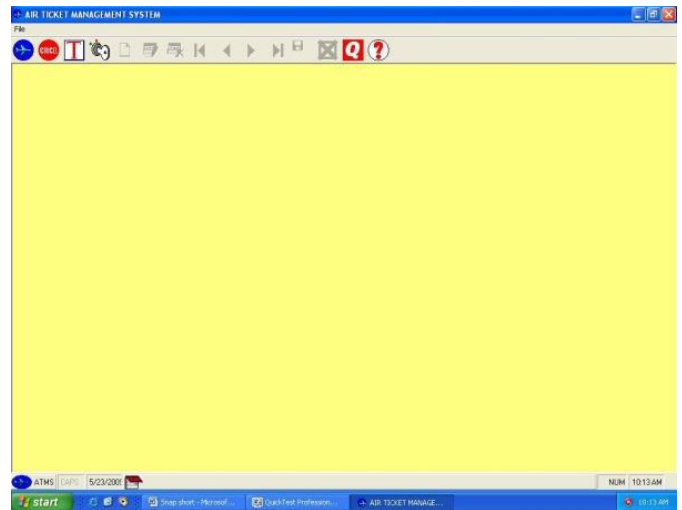


Fig 4. iCloud iBased iAir iTicket iManagement System

Description: This is Airlines Ticket management system i.e. complete module. In which researcher categorized to the unit part e.g. Airlines flight Unit, Airlines Reservation Unit system. By this Unit system, no doubt testing is done easily rather than testing to complete module. Because Unit tests are performed to prove that a piece of code does what the developer thinks it should be done. This Module was completed or implemented in high level language. These two modules were compared by manually or Automated tool i.e. QTP

IX. MODULE TEST FOR FLIGHT CATEGORIES

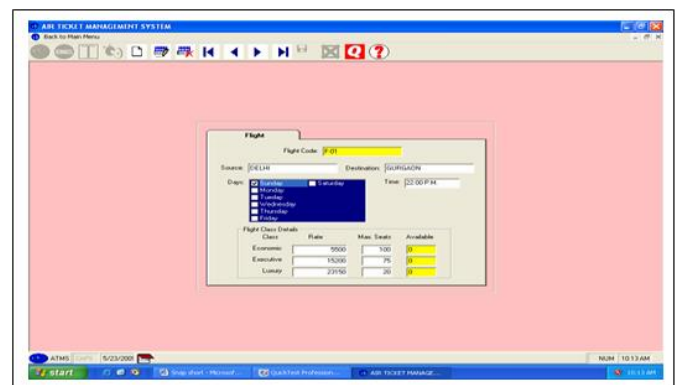


Figure : 45 iiModule iTest ifor iFlight iCategories

Description: This Module shows the Airline Flight Categories System. In this module each flight class details are mentioned i.e. economic class, executive class, luxury class etc

X. WHAT IS TEST CASE DESIGN

A test case in software engineering is a set of conditions or variables under which a tester will determine whether an application or software system is working correctly or not. The mechanism for determining whether a software program or system has passed or failed such a test is known as a test oracle. In some settings, an oracle could be a requirement or use case, while in others it could be a heuristic. It may take many test cases to determine that a software program or system is functioning correctly. Test cases are often referred to as test scripts, particularly when written. Written test cases are usually collected into test suites[10]

XI. TYPICAL WRITTEN TEST CASE FORMAT

A test case is usually a single step, or occasionally a sequence of steps, to test the correct behavior/functionalities, features of an application. An expected result or expected outcome is usually given.

Additional information that may be included:

- test case ID
- test case description
- test step or order of execution number
- related requirement(s)
- depth
- test category
- author
- Check boxes for whether the test is automatable and has been automated. Additional fields that may be included and completed when the tests are executed:
- pass/fail remarks

XII. WHAT ARE THE TYPES OF TEST CASE DESIGN TECHNIQUE

There are two types of test case design techniques they are

1. Equivalence class partition.
2. Boundary value analysis

Equivalence class partition: here the test engineer writes the valid and invalid test cases i.e. positive test cases and negative test cases. Boundary value analyses: if there is a range kind of input the technique used by the test engineer to develop the test Cases for that range are called as boundary value analyses.

XIII. Equivalence Class Partitioning

Concepts: Equivalence partitioning is a method for deriving test cases. In this method, classes of input conditions called equivalence classes are identified such that each member of the class causes the same kind of processing and output to occur. In this method, the tester identifies various equivalence classes for partitioning. A class is a set of input conditions that are likely to be handled the same way by the system. If the system were to handle one case in the class erroneously, it would handle all cases erroneously.[11]

Table: 2. Test Cases with approach of Equivalence Class Partitioning

Test Case Name	Test Case Describe	Test Steps			Test Case Status(p/f)
		STEPS	EXPECTED Result	Actual Result	
Economic Rate	Economic rate should be with in 5000-10000	1) <5000	Not Accepted	The input is accepted by the text box	Fail
		2)5000-6000	Accepted	The input is accepted by the text box	Pass
		3)6001-7000	Accepted	The input is accepted by the text box	Pass
		4)7001-8000	Accepted	The input is accepted by the text box	Pass
		5)8001-9000	Accepted	The input is accepted by the text box	Pass
		6)9001-10000	Not Accepted	The input is accepted by the text box	Pass
		7)>10000		The input is not accepted by the text box	Pass

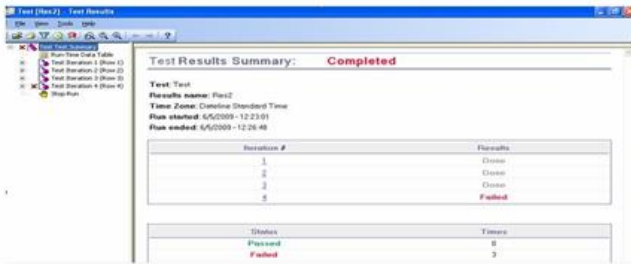


Fig 6. Test Report Summary

XIII. ANALYTICAL STUDY OF MANUAL VS AUTOMATED TESTING

Manual Testing is time consuming.

- There is nothing new to learn when one tests manually.
- People tend to neglect running manual tests.
- None maintains a list of the tests required to be run if they are manual tests.
- Manual Testing is not reusable.
- Tests have to be repeated by each stakeholder for e.g. Developer, Tech Lead, GM, and Management.
- Manual Testing ends up being an Integration Test.
- In a typical manual test it is very difficult to test a single unit.
- Scripting facilities are not in manual testing.[1]

XIV. COMPARATIVE GRAPH OF MANUAL VS AUTOMATED TESTING

Release	Manual Test	Auto Test	Manual Test Cumulative
1	10	10	10
2	10	0	20
3	10	0	30
4	10	0	40
5	10	0	50

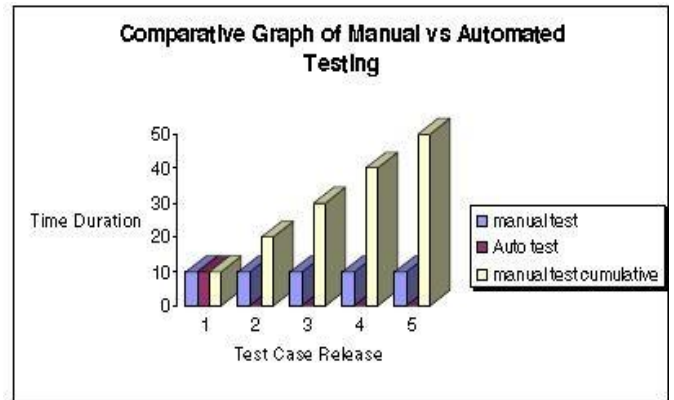


Fig 7. Comparative Graph of Manual Vs Automated Testing

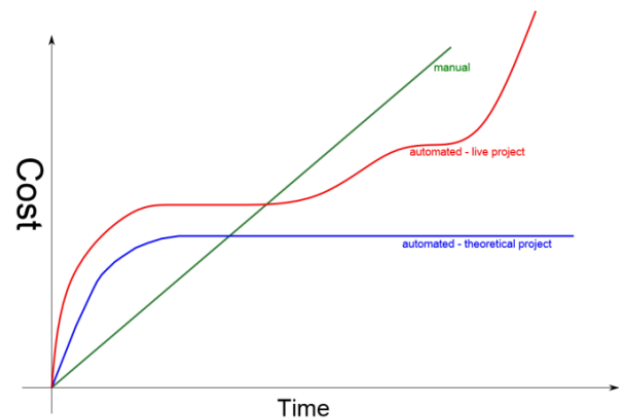


Fig 8, Comparative Graph of Cost Vs Time during Testing

XV. CONCLUSION

Since systems administration is pervasive and tremendous measures of information are presently accessible, enormous information is imagined to be the device for efficiency development, advancement and customer excess. Immense open doors identified with cutting edge enormous information examination and business insight are at the front line of research, concentrating on the examination of imaginative business-driven philosophies that can change different segments and 34 G. Skourletopoulos et al. businesses, for example, internet business, advertise knowledge, e-government, medicinal services and security . To this end, this instructional exercise paper talks about the ebb and flow enormous information research and calls attention to the exploration

difficulties and openings in this field by misusing distributed computing advancements and building new models. A money saving advantage investigation is additionally performed towards estimating the long haul advantages of receiving huge information as-a-benefit plans of action with a specific end goal to help information driven basic leadership and convey the discoveries to non-specialized partners. Cloud testing should be possible utilizing different foundations of cloud administrations and testing devices. As the progress of cloud innovation and testing as administrations, more scientists work have done to address the open issues and difficulties in distributed computing. In this paper, testing ideas used to improve the field of cloud testing. Presently a days, Diversion of programming organizations towards distributed computing because of different reasons, for example, cost decrease. Testing in the cloud, use the cloud applications, sinking the cost of registering, while expands testing viability. For look into researchers, we might want to lean toward mechanization testing utilization and presentation of computerization testing instruments for cloud applications as a future work to be execute as an expert action.

XVI. REFERENCES

- [1]. Batalla JM, Mavromoustakis CX, Mastorakis G & Sienkiewicz K, "On the track of 5G radio access network for IoT wireless spectrum sharing in device positioning applications", *Internet of Things (IoT) in 5G Mobile Technologies*, (2016), pp.25–35.
- [2]. Agrawal D, Das S & El Abbadi A, "Big data and cloud computing: current state and future opportunities", *14th International Conference on Extending Database Technology*, (2011), pp.530–533.
- [3]. Buyya R, Yeo CS, Venugopal S, Broberg J & Brandic I, "Cloud computing and emerging IT platforms: vision, hype, and reality for delivering computing as the 5th utility", *Future Gener. Comput. Syst.*, Vol.25, (2009), pp.599–616.
- [4]. Chen CLP & Zhang CY, "Data-intensive applications, challenges, techniques and technologies: a survey on big data", *Inf. Sci.*, Vol.275, (2014), pp.314–347.
- [5]. Google, Inc, *App engine platform as a service*. <https://cloud.google.com/appengine>, (2015).
- [6]. IBM Corporation: *IBM big data & analytics hub: the four V's of big data*, (2015).
- [7]. Microsoft Corporation: *Microsoft azure: cloud computing platform and services*, (2015).
- [8]. Park K, Nguyen MC & Won H, "Web-based collaborative big data analytics on big data as a service platform", *17th International Conference on Advanced Communication Technology (ICACT)*, (2015), pp.564–567.
- [9]. Talia D, "Clouds for scalable big data analytics", *IEEE Comput. Sci.*, (2013), pp.98–101.
- [10]. Zulkernine F, Martin P, Zou Y, Bauer M, Gwadry-Shridhar F & Aboulnaga A, "Towards cloud-based analytics-as- service (CLAAaaS) for big data analytics in the cloud", *IEEE International Congress on Big Data (Big Data Congress)*, (2013), pp.62–69.
- [11]. Vakintis I, Panagiotakis S, Mastorakis G & Mavromoustakis CX, "Evaluation of a Web crowd-sensing IoT ecosystem providing Big data analysis", *Resource Management for Big Data Platforms*, (2016), pp.461–488.
- [12]. Park K, Nguyen MC & Won H, "Web-based collaborative big data analytics on big data as a service platform", *17th International Conference on Advanced Communication Technology (ICACT)*, (2015), pp.564–567.
- [13]. AVanitha K & Alagarsamy K, "Software Testing in Cloud Platform: A Survey", *International Journal of computer applications*, Vol.46, No.6, (2012), pp.21–24.
- [14]. Michael AAF & Rean G, "Above the Clouds: A Berkeley View of Cloud Computing", *Electrical*

- Engineering and Computer Sciences, (2009), pp.1-23.
- [15]. Zhang L, Xie T, Tillmann N, De Halleux P, Ma X & Lv J, "Environment modeling for automated testing of cloud applications", IEEE Software, Special Issue on Software Engineering for Cloud Computing, Vol.1, No.20, (2012), pp.1-10.
- [16]. Mahalakshmi B & Suseendran G, "Effectuation of Secure Authorized Deduplication in Hybrid Cloud", Indian Journal of Science and Technology, Vol.9, No.25, (2016), pp.1-7.
- [17]. Nathiya T, "Reducing DDOS Attack Techniques in Cloud Computing Network Technology", International Journal of Innovative Research in Applied Sciences and Engineering (IJIRASE), Vol.1, No.1, (2017), pp.23-29.
- [18]. Mahalakshmi B, "A Detailed Study on Deduplication in Cloud Computing", International Journal of Innovative Research in Applied Sciences and Engineering (IJIRASE), Vol.1, No.1, (2017), pp.1-5.
- [19]. G Ainabekova, Z Bayanbayeva, B Joldasbekova, A Zhaksylykov (2018). The author in esthetic activity and the functional text (on the basis of V. Mikhaylov's narrative ("The chronicle of the great jute"). Opción, Año 33. 63-80.
- [20]. Z Yesembayeva (2018). Determination of the pedagogical conditions for forming the readiness of future primary school teachers, Opción, Año 33. 475-499
- Science/Information Technology. He has published over 100 Research papers in SCI, Scopus, Springer and IEEE. He has 40 Patents Published and Granted in his name at National and International level. He has been guiding 08 Ph.D. scholars in his supervision, 06 have been awarded degree. In over 50s International journals, he is in editorial board or working as reviewer. He has been awarded India Top 100 Best Researcher Winner 2022 by India Prime Award and Awarded Cash Prize and appreciation certificate for Best Research paper from Hon'ble Chancellor of Desh Bhagat University, Punjab on the occasion of 05th Sep 2022.

Cite this article as :

Prof.(Dr.) Rajender Kumar, "Analysis of Big Data and Cloud Services with Experimental of Manually and Automated Testing", International Journal of Scientific Research in Computer Science, Engineering and Information Technology (IJSRCSEIT), ISSN : 2456-3307, Volume 8 Issue 6, pp. 224-235, November-December 2022. Available at doi : <https://doi.org/10.32628/CSEIT228529>
Journal URL : <https://ijsrcseit.com/CSEIT228529>

Author Details



Prof.(Dr.) R. K. Bathla is working as a Professor in Computer Science Department at Desh Bhagat University, Mandi Gobindgarh, Punjab, India. He is also Pursing D.Sc in Computer Science & Applications (Post Doctoral Fellow) from Desh Bhagat University, Punjab. He has 21 years experience in the field of Computer