A SURVRY ON FAULT TOLERANCE IN CLOUD COMPUTING

Puja Chaturvedi[1], Brona Shah[2]

[1]Student, Computer Engineering, Silver Oak College of Engineering &Technology,Gujarat, India

[2]Assistant Professor, Computer Engineering , Silver Oak College of Engineering & Technology, Gujarat, India

ABSTRACT

With the immense growth of internet and its users, Cloud computing, with its incredible possibilities in ease, Quality of service and on-interest administrations, has turned into a guaranteeing figuring stage for both business and non-business computation customers.. The dynamic environment of cloud results in various unexpected faults and failures.Failure can lead to performance degradation to user and can result in loss of accuracy and efficiency .fault tolerance is the ability of a system to react gracefully to an unexpected equipment or programming malfunction . In order to achieve robustness and dependability in cloud computing, failure should be assessed and handled effectively. In this Paper ,we survey the various type of fault its taxonomy and fault tolerance technique in cloud .

Keywords :-*Cloud computing, Fault tolerance.*

---

1. NTRODUCTION

C*loud computing* refers to applications and services that run on a distributed network using virtualized resources and accessed by common Internet protocols and networking standards. It is distinguished by the notion that resources are virtual and limitless and that details of the physical systems on which software runs are abstracted from the user.[1]

It represents a real paradigm shift in the way in which systems are deployed. The massive scale of cloud computing systems was enabled by the popularization of the Internet and the growth of some large service companies. Cloud computing makes the long-held dream of utility computing possible with a pay-as-you-go, infinitely scalable, universally available system. With cloud computing, you can start very small and become big very fast. That's why cloud computing is revolutionary, even if the technology it is built on is evolutionary.

It is one of the hottest technical topics today, with broad-ranging effects across IT, Information Architecture, Business, Software Engineering, and Data Storage. Cloud Computing is an innovative technology that is revolutionizing the way it does computing. In recent years popularity of rapid growth in processing and storage technologies, computing resources have become cheaper, more powerful and more ubiquitously available than even before which is known as cloud computing.

Cloud is also responsible for the storing the huge client data and loss of this critical datacan lead to severe calamity.therefore cloud should maintain or provide proper fault tolerance mechanism so minimum amount of fault occur and if data is loss it can be recovered .

Fault tolerance allow system to act gracefully to an unexpected equipment or programming malfunction with minimal capacity rather than shutting down and improved the performance parameter.[1]

This paper organize as follows : next section present type of fault or error .in section III present metrics of fault section IV present fault taxonomy and finally section V present different fault tolerance strategies and finally we conclude the paper.

2. TYPE OF FAULT

This section we define various type of fault arise in the cloud[1][2]

2.1 Transient
A transient error is a temporary error that is likely to disappear soon .by definition it is safe for a client to ignore a transient error and retry to failed operation on the same database server .it occurs once at unpredictable intervals.

2.2 Intermittent
An intermittent error is an error that randomly occurs .these errors are commonly difficult to resolve because of the difficulty in reproducing the error it is malfunction of a device or system that occurs at intervals usually irregular in a device or system that function normally at time .the more complex the system or mechanism involved the greater the likelihood of an intermittent fault.

2.3 Permanent
A permanent fault is one that continues to exist until the faulty component is repaired.it existing or intended to exist for an indefinite period.

2.4 Byzanite failure
In these faults, the system components fail in arbitrary ways, causing the system to behave incorrectly in an unpredictable manner. The system may process requests incorrectly and produce inconsistent outputs.

2.5 Crash failure
When crash failures occur, they cause system components to stop functioning completely or remain inactive during failures – for instance, failures due to power outages or hard-disk crashes.

2.6 Hardware failure
It refers to failure of any physical component in the cloud environment.

2.7 VM failure
VM failure means failure of any one or many virtual machines hosted on an physical machine

2.8 Application failure
It refers to failure of the application perform in the cloud.

3. METRICS OF FAULT TOLERANCE

Metrics identified how the system perform even when the failure occur ,different type of metrics are[1][2][3]

3.1 Throughput
It defines the number of tasks whose execution has been completed. Throughput of a system should be high.

3.2 Response time
Time taken by an algorithm to respond and its value should be made minimized

3.3 Scalability
Number of nodes in a system does not affect the fault tolerance capacity of the algorithm.

3.4 Performance
This parameter checks the effectiveness of the system. Performance of the system has to be enhanced at a sensible cost e.g. by allowing acceptable delays the response time can be reduced.

3.5 Availability
Availability of a system is directly proportional to its reliability. It is the possibility that an item is functioning at a given instance of time under defined circumstances.

3.6 Usability

The extent to which a product can be used by a user to achieve goals with effectiveness, efficiency, and satisfaction..

3.7 Reliability

This aspect aims to give correct or acceptable result within a time bounded environment.

3.8 Overhead associated

It is the overhead associated while implementing an algorithm. Overheads can be imposed because of task movements, inter process or inter-processor communication. For the efficiency of fault tolerance technique the overheads should be minimized.

3.9 Cost effectiveness

Here the cost is only defined as a monitorial cost.

4. FAULT TAXANOMY

Cloud is prone to faults and they can be of different types. Various fault tolerance techniques can be used at either task level or workflow level to resolve the faults .fault taxonomy help to provide robustness in the cloud[1][4]

4.1 Reactive fault tolerance

Reactive fault tolerance techniques are used to reduce the impact of failures on a system when the failures have occurred.

• **Check pointing/Restart-** The failed task is restarted from the recent checkpoint rather than from the beginning. It is an efficient technique for large applications.

• **Replication-** various replicas of task are run on different resources until the whole replicated task not to crashed.

• **SGuard-**It is based on rollback recovery and can be executed in HADOOP, Amazon Ec2.

• **Task Resubmission-**The failed task is submitted again either to the same machine on which it was operating or to some other machine.

• **Rescue workflow-**It allows the system to keep functioning after failure of any task until it will not be able to proceed without rectifying the fault.

• **Job migration -**On the occurrence of failure, the job is migrated to a new machine.

• **Retry-** This task level technique is simplest among all. The user resubmits the task on the same cloud resource.

• **User defined exception handling-** Here the user defines the specific action of a task failure for workflows.

4.2 Proactive fault tolerance

Proactive fault tolerance predicts the faults proactively and prvent the fault by replace the components which cause fault by other working components thus avoiding recovery from faults and errors.

• **Software Rejuvenation-** the system is planned for periodic reboots and every time the system starts with new state.

• **Proactive Fault Tolerance using Self-** Failure of an instance of an application running on multiple virtual machine is controlled automatically.

• **Proactive Fault Tolerance using Preemptive Migration-** In this technique a application is constantly observed and analysed preemptive migration of a task depends upon feed back loop control mechanism.

5.FAULT TOLERANCE STRATEGY

5.1 Checkpoint based fault tolerance

Check point based fault tolerance[1] records the system state periodically after a certain time limit so that, if a failure occurs, the last checkpoint state of the system is restored and the task execution is resumed from that point. checkpoint strategy as it can be expensive in terms of performance. In case of virtualized environment, such as the cloud, the checkpoint strategy becomes more challenging, where huge virtual machine (VM) images needs to saved and restored.

5.2 Adaptive fault tolerance technique

Adaptive fault-tolerance techniques[1] help the system to maintain and improve its fault tolerance by adapting to environmental changes. Adaptive fault-tolerance techniques for cloud computing monitors the state of the system and reconfigures the cloud computing system for the stability of the system if errors are detected.

5.3 Fault tolerance for real time application(FTRT)

FTFR[4][5] is the highly intensive computing capabilities and scalable virtualized environment of the clouds help the systems to execute the tasks in realtime. Most of the realtime systems require high safety and reliability. The real time cloud fault tolerance model revolves around the reliability of the virtual machines. The reliability of the virtual machines is adaptive and changes after every computing cycle. The proposed technique depends on the adaptive behaviour of the reliability weights assigned to each processing node. The technique uses a metric to evaluate reliability. The nodes are removed if the processing nodes fail to achieve the minimum required reliability level. The primary focus of the system is on the forward recovery mechanism.

5.4 Dynamic adaptive fault tolerant strategy(DAFT)

The dynamic adaptive fault-tolerant strategy[1] observes a mathematical relationship between the failure rates and the two most common fault-tolerance techniques, checkpoints and replications. Historical data about failure rates helps the cloud computing system to configure itself for the checkpoints or the replicas. A dynamic adaptive checkpoint and replication model is made by combining checkpoints and replications. The dynamic adaptive fault-tolerant strategy was evaluated in a large-scale cloud data centre with regard to level of fault tolerance, fault-tolerance overheads, response time, and system centric parameters.

5.5 Fault and intrusion tolerant cloud computing hardpan(FITCH)

The novel fault-tolerant architecture for cloud computing, FITCH[1], supports the dynamic adaptation of replicated services. It provides a basic interface for adding, removing, and replacing replicas. The FITCH interface also provides all the low-level actions to provide end-to-end service adaptability. FITCH, are easily extendable and adaptable to various workloads through horizontal and vertical scalability. The number of computing instances that are responsible for providing the service are increased or decreased through horizontal scalability.

5.6 Byzanite fault tolerance cloud(BFTCloud)

The BFTCloud[4][7] is a fault tolerant architecture for voluntary-resource cloud computing In voluntary-resource cloud computing, infrastructure consists of numerous user-contributed resources, The architectures operate on five basic operations: primary node selection, replica selection, request execution, primary node updating, and replica updating. The primary node is selected based on QoS requirements. The request for the service is handled by the primary node. The primary node also selects the 3f+1replicas from the pool based on QoS requirements. All the replicas and primary node perform the operation on the request and send back the result to the primary node. Based on the result, the primary node decides to update the other primary node or update the replicas. In primary updating, one of the replicas is updated to primary node. In replica updating, the faulty replica is replaced with a new one.

.

5.7 Low latency fault tolerance(LLFT)middleware

Low-latency fault tolerance (LLFT) middleware[1][2] uses the leader / follower replication . The middleware consists of a low-latency messaging protocol, a leader-determiner membership pro-tocol, and a virtual determiner framework. The low-latency message protocol provides a reliable and ordered multicast service by communicating a message ordering information. The ordering is determined by the primary replica in the group. The technique involves fewer messages than earlier fault-tolerance systems. A fast reconfiguration and recovery service is provided by the membership protocol. The reconfiguration service is required whenever the fault occurs at the replica or some replica joins or leaves the group. The membership protocol is faster as it finds the primary node deterministically based on the rank and the degree of the backups. The virtual determiner framework takes the ordering information from the primary replica and ensures that all the backups receive the same ordering information. The LLFT middleware provides a high degree of fault tolerance and achieves low end-to-end latency.

5.8 International data fault tolerance(IFT)

IFT [1]Intermediate data is the data that is generated during the parallel dataflow program .The technique considers the intermediate data as of high priority (first-class citizen). The other techniques either use the store-local approach or distributed file system (DFS) approach If there is a failure of a server that stores the intermediate data, this results in the re-execution of the tasks. In the DFS approach the data is replicated, but causes too much network overhead. The network overhead results in the delay of jobs completion time.

5.9 Adaptive anomaly detection system for cloud computing infrastructure(AAD)

An adaptive anomaly detection (AAD)[1] system for cloud computing infrastructure ensures the availability of the cloud . The framework uses cloud performance data to discover future failures. Predicted possible failures are verified by the cloud operators. The failures are marked as true or false failures on verification. The algorithm recursively learns and improves future failure prediction based on the verified data. The framework also takes into account the actual failures that were not previously detected.

5.10 Dynamic fault tolerant scheduling(DFTS)

DFTS[8] incorporate a backup overlapping mechanism and efficient VM migration strategy for designing novel Dynamic Fault Tolerant Scheduling Mechanism for Real Time Tasks in cloud computing. Author proposed model aims to achieving both fault tolerance and high resource utilization in the cloud. It is based on dynamic resource expansion for creating new Vm when task arrived and dynamic resource contraction for rejected idle resource.

5.11 Virtualized failover strategy(IVFS)

IVFS[9], critically analyze this model and proposed a model that tolerate faults based on the reliability of each computing node or virtual machine, removing these from the availability list if the performance is not optimal.provide high availability in the cloud system and fulfil client requests on service performance and completion time as defined by the SLA. Present a working model of the strategy and a mathematical relationship to represent the fault tolerance model system using the concept virtualization and FT checkpoint/replay scheme. The checkpoint/replay model is developed using the Reward Renewal process (RRP) theory a backward recovery is performed and the VM is immediately restarted and recovered from the last successful checkpoint.

5.12 Cloud computing fault net(CFN)

CFN[4][6] is used to precisely model the different components of cloud computing, such as service resources, cloud module, the detection and failure process ... etc. It is used to create the different components of Cloud Computing which gets integrated dynamically into CFN model. Based on CFN model, the properties of the components are analyzed developing a fault detection strategy at each level which dynamically detects the faults in the execution process.

5.13 Fault tolerant scheduling for real time scientific workflow with elastic resource provisioning in virtualized cloud(FASTER)

FASTER[10] employs a backward shifting method to make full use of the idle resources and incorporates task overlapping and VM migration for high resource utilization it applies the vertical/horizontal scaling-up technique to quickly provision resources for a burst of workflows it uses the vertical scaling-down scheme to avoid unnecessary and ineffective resource changes due to fluctuated workflow requests.

5.14Fault tolerant elastic scheduling algorithm for real time task in virtualized cloud(FESTAL)

FESTAL[11] model base on Resource scaling up and scaling down for scheduling the task on the existing resource and eliminate the idle resource respectively author also incorporate BB overlapping for overlap the primary copy on computing and I-AEP technique for scheduling primary and backup to reserve sufficient time laxity for finish the primary as early as possible

5.15 Virtualization and fault tolerance(VFT)

VFT [4][12] is a reactive fault tolerant technique; it consists of a Cloud Manager (CM) module and a Decision Maker (DM) which are used to manage the virtualization, load balancing and to handle the faults. The first step involves virtualization & load balancing and in the second step fault tolerance is achieved by checkpoint and fault handler. The virtualization includes a fault hander. Fault handler finds these unrecoverable faulty nodes and It also helps to remove the temporary software .

5.16 Automatic self healing of server application in a virtual machine environment (SHELP)

SHelp [4][13] is an error handler which are run in different VMs hosted on one physical machine. It uses the Checkpoint/Restart as the checkpoint and rollback tool. Authors introduced two new techniques, namely, weighted rescue points and two-level rescue point database. In weighted rescue point, each rescue point is assigned a weight value which is initially set to zero; the associated weight value is incremented each time a fault is found in a function. When a fault occurs, the application is rolled back to a latest checkpoint, and first uses error virtualization at a rescue point which has the largest weight value among the candidate rescue points, then at the rescue point with the second largest weight value and so on until the fault is bypassed.

6. CONCLUSION

Ascloud become the popular in the recent year in field of computational technology and it has vast advantage over the data and resource management ,on demand service ,reduce cost , backup and recovery ,versatile compatibility that make it unique from other technologies.to maintain the reliability and availability of the service there is need of efficient fault tolerance method to avoid fault.in this paper concentrate on the different type of fault or error ,its metrics and also describe various type of fault taxonomy and strategies. There are various numerous fault tolerance methods proposed by research experts in this field.

7. REFERENCES

[1]Kashif Bilal, Osman Khalid, Saif Ur Rehman Malik, Muhammad UsmanShahid Khan, Samee U. Khan and Albert Y. Zomaya "*Fault Tolerance in the Cloud" ,ENCYCLOPEDIA OF CLOUD COMPUTING*, 2016

[2]Soma Prathiba,S. Sowvarnica,"Survey of Failures and Fault Tolerance in Cloud",IEEE,2017

[3]Zeeshan Amin, NishaSethi, Harshpreet Singh, "Review on Fault Tolerance Techniques in Cloud Computing", IJCA, 2015

[4]Salma M. A. Ataallah, Prof. Salwa M. Nassar,Prof. Elsayed E. Hemayed,"Fault Tolerance in Cloud Computing ,",IEEE,2015

[5]S. Malik, and F. Huet. "Adaptive Fault Tolerance in Real Time Cloud Computing. " ,IEEE, 2011.

[6]G. Fan, H. Yu, L. Chen, and D. Liu. "Model Based Byzantine Fault Detection Technique for Cloud Computing. ",IEEE, 2012.

[7]Y. Zhang, Z. Zheng, and M. R. Lyu. "BFTCloud: A byzantine fault tolerance framework for voluntary-resource cloud computing. " , IEEE, 2011.

[8]J.Soniya, J.AngelaJennifaSujana, Dr.T.Revathi, "*Dynamic Fault Tolerant Scheduling Mechanism for Real Time Tasks in Cloud Computing* ", IEEE, 2016

[9] Bashir Mohammed, Mariam KiranandIrfan and UllahAwan, Kabiru.M. Maiyama, "*An Integrated Virtualized Strategy for Fault Tolerance in Cloud Computing Environment* ", IEEE, 2016

[10]Xiaomin Zhu, Ji Wang, HuiGuo, Dakai Zhu, Laurence T.Yang, and Ling Liu "*Fault-Tolerant Scheduling for Real-TimeScientificWorkflows with Elastic Resource Provisioning inVirtualized Clouds*", IEEE, 2015

[11]Ji Wang, WeidongBao, Xiaomin Zhu, Laurence T. Yang, and Yang Xiang, "*FESTAL: Fault-Tolerant Elastic Scheduling Algorithm for Real-Time Tasks in Virtualized Clouds*" ,IEEE, 2014

[12]P. Das, and P. M. Khilar. "VFT: A virtualization and fault tolerance approach for cloud computing,IEEE,2013.

[13] G. Chen" H. Jin, D. Zou, B. B. Zhou, W. Qiang, and G. Hu. "SHelp: Automatic Self-healing for Multiple Application Instances in a Virtual Machine Environment. ", IEEE ,2010