Disk storage failure prediction in datacenter using machine learning models

September 2021 · 13(1)
DOI: 10.1007/s13204-021-02039-4

Manikandan Ramanathan · Kumar Narayanan

Abstract

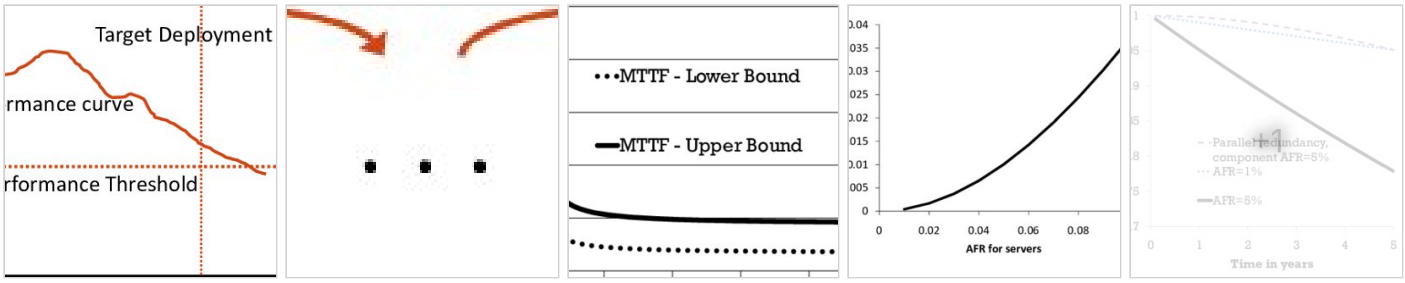
Data centers are located centralized to do computation and accessing huge amount of data by the network devices which are interconnected to form the network path. Servers are stacked, data storage is placed in them. Data server backup and server redundancies are the recovery mechanisms implemented. Data centers compute, store, distribute the data by processing them and the data center controls all the interconnected network equipment in the distributed network. In current, RAID system is implemented to avoid the service disruptions due to disk failures, the availability of system and services are achieved with this expensive model. But still the availability is lost, and service disruptions happen due to disk failures, the machine learning models to be used to predict the disk failures well in advance. Data center has increased usage of system with increased data storage, the failure in disc makes the system failed and down time increases. Analysis on the methods of problems in disk and methods of disk availability and reliability failure is the main goal. Various machine learning models are identified and discussed along with the SMART parameters for measuring the failure of the disk. Improved method of Ensembling of trees, random forest and boosting techniques are also discussed.

Public full-texts

To read the full-text of this research, you can request a copy directly from the authors.

Request full-text PDF

Similar research



Modular data centers: how to design them?

Article Full-text available

July 2009 · 1,346 Reads · 51 Citations

Kashi Venkatesh Vishwanath · Albert G. Greenberg · Daniel Reed

There has been a recent interest in modularized shipping containers as the building block for data centers. However, there are no pub-lished results on the different design tradeoffs it offers. In this paper we investigate a model where such a container is never serviced during its deployment lifetime, say 3...

Read more

View

Analysis of substantial growth on Data Center in virtualization local area network applications

Conference Paper

November 2019 · 9 Reads

Manikandan Ramanathan · Kumar Narayanan

View

Data Center Network Topologies: Current State-of-the-Art

Chapter

August 2013 · 68 Reads · 4 Citations

Yang Liu · Jogesh K. Muppala · Malathi Veeraraghavan · [...] · Mounir Hamdi

Current state-of-the-art data center networks are primarily implemented using tree-based topologies. An alternative is to use a leaf-spine configuration which is based on the folded-Clos network. While these approach meet the demands of the current DCN, they are not suited to meet future demands. This...

Read more

View

A cost comparison of datacenter network architectures

Conference Paper

November 2010 · 568 Reads · 140 Citations

Lucian Popa · Sylvia Ratnasamy · Gianluca Iannaccone · [...] · Ion Stoica

There is a growing body of research exploring new network architectures for the data center. These proposals all seek to improve the scalability and cost-effectiveness of current data center networks, but adopt very different approaches to doing so. For example, some proposals build networks entirely out ...

Read more

View

On the Mathematics of Data Centre Network Topologies

Conference Paper

August 2015 · 51 Reads · 1 Citation

Iain A. Stewart

In a recent paper, combinatorial designs were used to construct switch-centric data centre networks that compare favourably with the ubiquitous (enhanced) fat-tree data centre networks in terms of the number of servers within (given a fixed server-to-server diameter). Unfortunately there were flaws in some of th...

Read more

View