

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/363272441>

START UP MANAGEMENT START UP MANAGEMENT

Book · March 2022

CITATIONS

0

READS

418

1 author:



[Nuzhath Khatoon](#)

Aurora PG College (Affiliated to Osmania University

24 PUBLICATIONS 54 CITATIONS

[SEE PROFILE](#)

START UP MANAGEMENT

START UP MANAGEMENT

Dr Nuzhath Khatoun



START UP MANAGEMENT

Dr. Nuzhath Khatoun



PROFESSIONAL BOOKS PUBLISHER

Head Office: 3-5-315, Street No. 7, Vittalwadi, Narayanguda,
Hyderabad - 500 029. T.S. Phone : 23227399 / 9642665303, 9848130433
E-mail : professionalbookspublisher@gmail.com
Website: www.professionalbookspublisher.com

Branch Office:
364, Kaveri Complex, 104, Nungambakam High Road, Chennai - 600 034.

₹ 135 .00

ISBN: 978-93-91679-22-4



9 789391 679224



Every effort has been made to avoid errors or omissions in the publication. In spite of this, errors may have crept in. Any mistake/error or discrepancy noted may be brought to our notice, which shall be taken care of in the next edition.

ALL RIGHTS ARE RESERVED

No part of this publication, which is material protected by this copyright notice, may be reproduced, stored in a retrieval system, or transmitted in any form or by any means -electronic, digital or mechanical, including photocopying, scanning, recording or otherwise, without prior written permission of the publisher.

First Edition - 2022
Copyright © Publisher

ISBN:978-93-91679-79-8

Price: Rs. 135.00

Professional Books Publisher

Office:

3-5-315, Street No. 7 Vittalwadi,

Narayanguda,

Hyderabad, 500029. T.S.

Tel: 040-23227399,

Mobile: 9642665303, 9848130433

Email: professionalbookspublisher@gmail.com,

Website: www.professionalbookspublisher.com

Chennai

364, Kaveri Complex, 104,

Nungambakam High Road, Chennai - 600 034.

Mangaluru

No-6-91/1, Behind kankanady Garody Temple,

49 Kankanady, D.K., Mangalore-575002

Karnataka

Mobile: 9902997473

International Conference on
Information Science and
Communications Technologies



ICISCT 2021

Applications, Trends and Opportunities

CERTIFICATE OF PARTICIPATION

This certificate is awarded to authors
Y. Nasir Ahmed and S. Pakkir Mohideen
for the contribution and research paper presentation, titled:
**Workload Characteristics in Cloud Data Centers: A Computational
Study from Alibaba Cloud**

Vice-rector for the scientific
affairs and innovations of TUIT,
Conference Co-Chair

Dr. Komil Tashev



The Ministry for Development
of Information Technologies
and Communications of the
Republic of Uzbekistan



2021 International Conference on Information Science and Communications Technologies (ICISCT 2021)

**Tashkent, Uzbekistan
3 – 5 November 2021**

Pages 1-594
 **IEEE** IEEE Catalog
Number:
ISBN:

1/2

CFP21H74-POD 978-1-6654-3259-7

Copyright © 2021 by the Institute of Electrical and Electronics Engineers,
Inc. All Rights Reserved

Workload Characteristics in Cloud Data Centers: A Computational Study from Alibaba Cloud*

Y. Nasir Ahmed

Department of Computer Applications
B. S. Abdur Rahman Crescent Institute
of Science and Technology, Vandalur
Chennai, Tamil Nadu, India
+91-8179713252
Nasir_ca_phd_18@crescent.education

Dr. S. Pakkir Mohideen

Department of Computer Applications
B. S. Abdur Rahman Crescent Institute
of Science and Technology, Vandalur
Chennai, Tamil Nadu, India
+91-9444780125
pakirmoithen@crescent.education

Abstract— Cloud computing is trending in today's digital era in every form and scaling to a peak in order to deliver the mutual expectations for Cloud consumers and providers through Cloud Data Center's. One such data center so called Alibaba Cloud's Data Cluster is analysed in this paper. The online services and Batch jobs co-exist in the same data cluster which leads to multiplexing of resource utilization in Cloud Data Center's. Therefore our findings and insights from this data production cluster allows the community and data center operators better understand the various workload patterns to improve the resource utilization and failure recovery design.

Keywords— Cloud Computing, Batch Jobs, Online Services, Data Centers, Scheduling.

Introduction:

The soul requirement of any data center is its highest performance, lowest latency, budget friendly, highly scalable. The Alibaba cloud data center which is opted for our study uses Content Network Delivery (CDN) enabling the fastest distribution of content delivery with respect to reducing the latency serving with highest speed of load times and also uses solid state drives (SSD) storage and 10-GE NICs powered network for high performance. In this paper we have tried to analyze the various computational tasks running in the cloud data center and their impact on overall performance of the system and execution policy towards running scheduled batch jobs and incoming online services. Therefore, we have chosen the latest data cluster released by Alibaba Cloud version titled as Cluster-Trace-V2018 which is released in April 2019 for researchers to better understand the modern data centers and verify their ideas. This cluster has records of 4000 machines for eight days of long running containerized services and batch jobs. Data clusters have the tasks and each task is dependent on their respective dependencies to get work done. In modern computation these are analyzed by directed acyclic graphs. The directed edge resembles its dependency and the two ends of an edge are the two compute tasks [2]. Rigorous study on cluster workload is in place but as the production trace data does not include any dependency information, so complex dependencies are not explored and managed performance attain after proper scheduler design is missed all the time. It is worth to consider the directed acyclic graphs to understand and correlate the patterns for the various task dependencies that exist in any cloud cluster [14]. Based on this we have performed a comprehensive analysis on these batch and online services to identify and better understand their scope, runtime performance and dependency structures which helps data centers operators for the best design and execution policy towards arriving computational jobs in the Production

servers [14]. Based on the analysis of the overall cloud system and close inference of the variable workload patterns a model comprising of routing the various tasks based on the interest of users is proposed, providers may be setup before actually sending the arriving tasks to perform randomly and hence the violations of service level agreements are controlled and the overall system is progressed towards managed performance. Cloud environments defined with workload variability enables to measure the realistic status with respect to productivity and key performance indicators for the researchers and providers [23].

I. RELATED WORK

Workload variances of Software, Machines have changed since 2011 hence analysis of data clusters in an data centers is essentially required to understand the various aspects such as job submission, resource usage, scheduling decisions etc. of the batch jobs and online services. An extensive analysis to be stage on resource request and resource usage and up frame their patterns so to ease the scheduling decisions within the data cluster which progress the overall health of the data center. Task Dependencies study [3] is significant to understand the various dependency schedulers in order to predict their duration and usage of resources. The dynamic behaviours with respect to various virtual machines workload from Microsoft Azure is presented by Cortez et al. [4]. Similarly, results showed [5] are dissimilar towards speedup performance. Four clusters are introduced and studied to address sensitivity of various workload characteristics of the publicly released traces [6]. Detailed dependency structured analysis studied [7] which gives a clear idea to implement the tasks and occurrences of various jobs to be placed to result in fair means of data cluster Performance Aware Fair scheduling of resources is proposed in [8], to improve the machines performance while attaining negligible loss of service level agreement towards the providers end.

The Elasticity concept in cloud computing has demonstrated in [9], where the workload forecasting is addressed and proposed the tools for this forecasting methodology for better provisioning of the resources but Shown on workloads of grid computing not on traces of cloud [14]. Although various studies incur is not cultivated in relevance to actual start time and termination time spent to complete a job through their complete life cycle. The deficient in such study might be due to the unavailability of statistical information of the jobs running in production