ISSN 2347 - 3983
Volume 11. No.11, November 2023
International Journal of Emerging Trends in Engineering Research
Available Online at http://www.warse.org/IJETER/static/pdf/file/ijeter0211112023.pdf
https://doi.org/10.30534/ijeter/2023/0211112023

Design of Monitoring Tools for Data Centre Downtime Reduction

Dr. G. Prasad, A.S. Aravind, C. Pradeep
NRSC/ISRO, Hyderabad, India, gurram_p@yahoo.com
NRSC/ISRO, Hyderabad, India,aravind.excel@gmail.com
NRSC/ISRO, Hyderabad, India, pradeepcps4@gmail.com

Received Date: September 27, 2023 Accepted Date: October 21, 2023 Published Date : November 07, 2023

ABSTRACT

This paper presents a new monitoring tool and event management method for data centre compute, network and storage infrastructure based on node event processing. The uptime of highly classified data centres are not only to be maintained at the highest level of reliability and availability of the operation, but also fast, specific event identification and rectification, which altogether improves availability of the resources is important. The new method, using a tree node for each element of the data centre resources provides information about the compute, network and storage file system configuration in a specific node. Its major advantage is that in our case where a large number of heterogeneous computers are present, it helps us in monitoring all the elements of the computer resources and gives information for alerting the associated work centres before any of the identified events that might occur. By monitoring and informing apriori to the concerned work centres the state of the systems, it lowers errors in data centre physical infrastructure operating costs, improving at the same time the level of operations efficiency. This method resulted that the use of tree nodes significantly reduces the number of unexpected events, the time needed for the main event identification, and the maintenance response time to events. By using event entities processing, multilayer nodes have a significant impact on the efficient operation of data centre physical infrastructure.

In this paper, the design and development of two customised dashboards to monitor the compute, storage and network elements of the heterogeneous data centre for uptime maintenance and optimal performance is presented. The dashboards are designed, keeping in view the nature of tasks carried out and the resource requirements of various work centres in the data centre. One dashboard displays dynamically created icons for each of the compute resources in the data centre. On clicking any of the icon, complete details of the corresponding server is fetched showing the status, usage, configuration and available resources. Furthermore, a unique colouring scheme is followed wherein the icon is displayed green if the server is healthy and orange if the server is facing a resource crunch (disk, memory, etc.) and red if the server is not reachable. The dashboard GUI refreshes every 5 min (is configurable), displaying the latest status details of the servers in the data centre.

The second Dashboard is developed with the capability to monitor the storage, cloud and network infrastructure components. The dashboard collects data from different elements of the storage i.e. Meta Data Servers, Storage, Core and Edge switches etc. and processes the collected data to a customized format for display. It delivers details like availability of Storage Meta Data Servers, switches and file systems, disk space capacity monitoring, file system backup status, Monitoring of the hierarchical Storage including Tape Library and the availability of Production ESXi hosts cluster. The GUI is updated with new requirements to further fine-tune and reduce manual intervention for monitoring operations.

Key words: Dashboard, ESXi, MDS, SAN, Satellite

1. INTRODUCTION

A data centre is a building specified for hosting a facility that contains concentrated equipment to perform Store, manage, process, and exchange digital data and information. Elements of the data centre facilitate application services or management for various types of data processing, such as digital data processing and storage, web hosting and simulations to cloud based services. Data centres are the pillar of today's IT society, and many software development companies, data processing institutes relies on data centres. As the internet community is growing each day, so does the dependency on data centre storage and compute. There are different types of data centres and IMGEOS is a level 3 data centre involved in satellite data acquisition, processing and dissemination.

Remote monitoring in today's data centres has become very essential due to the large compute infrastructure and storage volumes. In order to maintain high availability of data, all components of the data centre infrastructure must carry out their assigned tasks, hence maintaining and monitoring the health of the data centre components including power, environment is of utmost importance [1]. All physical infrastructure failures

have a significant impact on the effectiveness of data centre services. Therefore, monitoring and control is important for all existing component parameters in the data centre infrastructure system, which are crucial in maintaining the Turnaround time of the products generated. In terms of power consumption, it is followed by servers, storage, and other infrastructure elements. Our data centre consumes as much energy as 240KVA. In order to induct new resources in to the data centre and to maintain the energy consumption within the limit regulation of the energy has to be carried out. The main goal is to clearly monitor all processes of data centre operation and data centre energy reduction is now regarded as another concern for data centre smooth operations [2].

When such processes are not effectively implemented, inconsistencies in the operation and maintenance of the systems are inevitable, leading to unexpected downtimes. Any downtime results in the increase inturnaround time of the production, which in turn does not meet the timely data requirements of the users. Our method is used to detect the location of an event in the distribution of the approximately 180 computers. For each of the system, the configuration of the system, OS file systems allocation, Storage file systems mounts, multipath etc., a data base is built with different colours on the dashboard depicting the possible causes of events or failures. This paper presents a new remote monitoring method for data centre physical infrastructure parameters, based on multilayer node and event entities processing [3]. The method highly improves information about the events in a specific node. Its major advantage is that in the case of a large number of nodes, it considerably reduces the number of identified eventsCurrent Trends in Technology and Science.

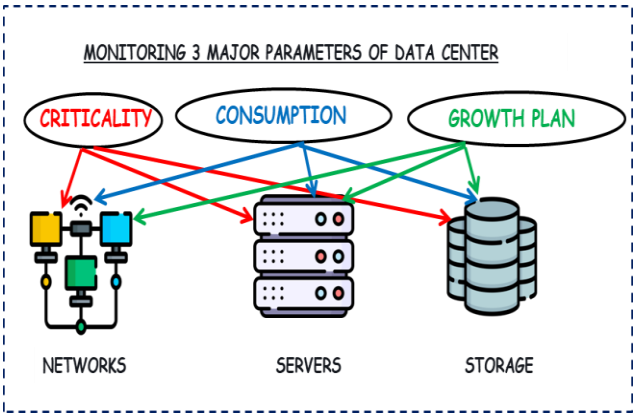

Figure 1: Major Monitoring Parameters

2. IMGEOS DATA CENTRE ELEMENTS

The Integrated Multi Mission Ground Segment for Earth Observation Satellites (IMGEOS) data centre comprises of servers, storage arrays of made of SSD, SAS, NLSAS, core switches, edge switches, Fibre optic cables, Ethernet cables, server racks and phone lines. To maintain the turnaround time (TAT) as per the IMGEOS policy, the data centre has to stay healthy meaning, providing continuous uptime [4]. However, this requirement which is essential has given rise to the complexities associated with managing a modern data centre, which encompasses a growing number of devices in heterogeneous environments and dispersed over multiple locations.

Also the data centre security solution includes authentication, authorization and data privacy. There are significant challenges in offering what many consider to be a totally secure data centre infrastructure. When considering remote management solutions, connectivity paths to networks and each element (Server, storage) inside the data centre (gateway or Wide Area Network (WAN) edge, core, distribution, and access) it must be designed with no single point of entry for unauthorized users as shown in Figure 1. In this new era of having to do more with less data centre managers, it is essential to evaluate tools that enable them to centrally manage their data centres.

Networks continue to grow in size and scope and the need for greater integration of applications and services complicates the picture [5]. Data Centre at NRSC is the substrate of the IMGEOS Operational chain viz., Data Ingest, Ancillary Data processing, level-0 processing, and various levels of Data processing to produce satellite data and then pass on the product to Data Exchange gateways where the product will be handed over to the end-user via various dissemination methods. Being the production environment, handling Data centre remains a major challenge for IT administrators to keep it operational for 24 x 7 and 365 days. Controlling Data centre is broadly classified into Operations, Administration and Management (OAM). These are processes, activities, tools and standards that support the implementation of this huge live system.

Under Operations - the system monitoring, that keeps a watch on the system health primarily by learning the system steams. IMGEOS DX Integrated Monitoring Dashboard is developed as an entirely customized solution that provides various insights of heterogeneous IMGEOS data centre. One of the obvious places to economize is infrastructure consolidation and one obvious solution is remote management. Remote management allows centrally located personnel and applications to monitor, manage, and respond to globally distributed networks and systems from a single location. With these tools, IT managers can respond to problems quickly and perform corrective actions.

3. DESIGN METHODOLOGY

The health of a data centre is ensured by the availability and accessibility of the resources and services rendered by it. Thethree major parameters that affect the health all the layers of data centre are criticality, consumption and growth plan. Traditional remote monitoring is not an online service therefore it cannot provide real-time monitoring. Instead it relies on intermittent status updates. Digital remote monitoring is online and connected to a data centre usually through a gateway which allows for real time monitoring[6].

IMGEOS data centre is equipped with around 180 servers that are being used by different work-centres in the satellite data products generation chain. These servers are maintained to ensure that they are

100%available during 24hrs X 7 day operations. The server categories in this chain carry out functions pertaining to real time Satellite Data Ingest, Ancillary Data Processing, near real time Data Products generation, Quality Evaluation and Certification, Data Exchange Gateways for dissemination and special products generation. In order to monitor the availability and performance of these servers, an application in VB. NET has been developed [7]. The utility will display icons of each server and fetch complete details of any server when corresponding icon is clicked as shown in Figure2. The details consist of server IP/hostname, CPU and make information, Linux OS version, mounted filesystems, active users, multipath information and FC/network port details. Snapshot of the information window is shown in Figure 3.

The dashboard also shows a colour-coding for the icons incorporated in the GUI. Here green indicates normal functioning, amber indicates a warning (either memory or the disk space of any file system is reaching maximum limit) and red indicates that the server is not reachable. The GUI automatically refreshes once in 10 minutes (customizable) to ensure availability of updated information. The application is particularly useful in indicating if any server is not available/reachable at any point of time, and it also indicates if resources (disk space and memory) in any server are exceeding desired thresholds.

Figure 2: Servers Monitoring Application GUI

Each icon depicts a server. The servers are categorised based on work-centres. Green colour indicates that the server is functioning normally. Orange indicates a warning, while red means the server is unreachable/switched OFF. The details of the system clicked on the GUI will display the parameters as shown in Figure 3.The dashboard application development consists of configuring at three levels as illustrated in Figure 4.

First, all the servers in the network are configured to execute a script,in the background (eg. crontab in linux, task scheduler in windows), which generates detailed system information at pre-defined intervals (every 10 min in this case).The utility next pings each server to check availability over network, creates icons (buttons) dynamically and allocates colour to the icon based on the serverstatus (red - not reachable, green – reachable and healthy, orange – reachable but needs attention).

SYSTEM INFORMATION as on Wed Jul 13 09:30:01 UTC 2022

SERVER HOSTNAME :boole
IP ADDRESS : 10.1.1.136 10.2.1.136 175.28.1.136 192.168.122.1
SYSTEM MAKE : FUJITSU
Serial Number : YM3G002218
OS and version info : Red Hat Enterprise Linux Server release 6.6 (Santiago)
Server uptime : 187 days

DF INFO

Filesystem	Size	Used	Avail	Use%	Mounted on
/dev/sda1	79G	571M	75G	1%	/
tmpfs	32G	452K	32G	1%	/dev/shm
/dev/sda8	107G	15G	87G	15%	/home
/dev/sda2	30G	18G	11G	62%	/opt
/dev/sda7	9.8G	7.9G	1.4G	86%	/usr
/dev/sda6	15G	1.8G	13G	13%	/var
ipo	52T	40T	13T	77%	/ipo

Memory INFO

	total	used	free	shared	buffers	cached
Mem:	62	6	56	0	0	2
-/+ buffers/cache:	3	59				

MULTIPATH INFO

multipathd (pid 5958) is running... MULTIPATH COUNT: 64

ACTIVE USERS

o3scat pts/0 2022-11-21 16:59 (10.1.2.15)
o3scat pts/1 2022-11-22 18:32 (10.1.2.15)
mgmt :0 2022-02-17 15:37 (:0)

FC INFO

PORT NAME	PORT STATE	MAX SPEED	SPEED	CLASS
0x100000109b9b34db	Online	16 Gbit	16 Gbit	Class 3
0x100000109b9b333c	Online	16 Gbit	16 Gbit	Class 3

NETWORK PORTS INFORMATION

Port_ID	IP Address	Max Speed	Curr Speed	Link Detected
em1	---None---	1000baseT/Half	Unknown!	no
em2	10.1.11.77	1000baseT/Full	1000Mb/s	yes
em3	10.2.11.77	1000baseT/Full	1000Mb/s	yes

CPU INFORMATION

model name : Intel(R) Xeon(R) Gold 6140 CPU @ 2.30GHz
Core(s) per socket : 18
Socket(s) : 4

Figure 3: Detailed Server Information, displayed on clicking corresponding server icon in the GUI

Thirdly, action listeners are added to each of the buttons for cursor-hovering and cursor-click. When a user clicks on any of the server icon, the application accesses the corresponding server and fetches/displays the complete system information to the user. The utility also contains logic to automatically send SMS to pre-defined system administrators whenever any issue with server performance/reachability is encountered.

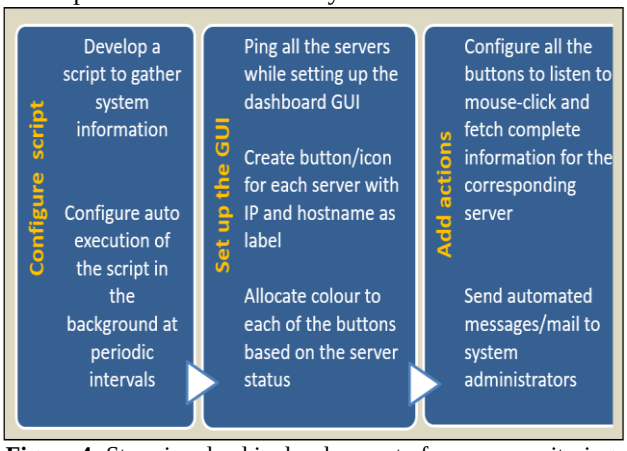

Figure 4: Steps involved in development of servers monitoring dashboard

The process of aggregating, analysing, and visualizing metrics from various components through the pipeline can be complex and time consuming. The

storage Dash Board automates the process for monitoring; keep track of activities visualizing health and availability of critical infrastructure components. This solution supports ingestion, processing, analysis, and visualization of data from sources, the calculated datasets stored in database tables that serves as a repository. Dash board displays many important parameters such as activities including filesystems, storage disks and three-tier information, which are currently available at that instance of log collection. Hence, this solution helps administrators to measure the impact of their operational initiatives and make data-driven decisions to drive continuous improvement in their day-to-day activities.

4. IMGEOS DX INTEGRATED MONITORING DASHBOARD

This solution runs the following work flow. Script running at Metadata server clusters 1 and 2 runs at a periodic interval specified by the administrator.
The shell script collects the information on Meta Cluster-name of server, local disk status, HBA status, filesystem(name, availability,total capacity, used capacity, available capacityandpercentage of used capacity,

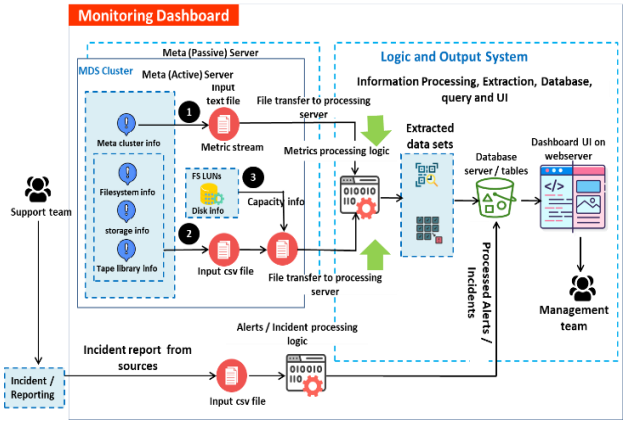

Figure 5: Architecture of storage Dashboard

backup status, count on storage LUNs (management and data), users, files and directories. Storage tiers (tier1-3) – available free LTO media count, individual drive status, alias name, ACSLS identifier, SL Console Identifier, device alias, LUN information from admin as shown in Figure 5.The information gathered is then written in a data text file and the file is transferred to the information processing server. The information processing server applies extraction logic programs on the data text file and extract the required information and write them into respective .csv files. The .csv files are then loaded into respective database tables created in the remote database server. The webserver that runs multiple webpages to obtain query results from the database and build dashboard visualizations. The Operational Processes view concerns capturing activity based behavior and flows. It describes the activities that are normally conducted in the course of achieving the goals that support the automated capability for monitoring critical data center components [8]. It describes operational activities, their inputs/outputs, operational activity actions and flows between them.

The operability of the application is made simple by mouse clicks and the metrics are displayed with related selection and clicking on them. For example under segment 2: Storage Area Network, the filesystem insights are displayed in a tabulated format and the last column of the table says about the filesystem composition. On-click the respective row opens the details of LUNs, sg alias, size and other information on that particular filesystem. Similarly, there are other displays that operate with mouse on-click the ticket opens with the ticket no selected by the radio button with the internal details of the ticket / incident.

The Code Log Collection will invoke two delivery streams to process raw data from data sources – Meta Data servers, external csv files from cloud, user end and auto generated. Code Log Processing will invoke processing scripts based on the pre-defined file patterns received from the data sources and extract the destined output files. Code upload events logic will match the output files from the pre-defined output areas and pick them and upload to the database tables and move out the uploaded file into archives area.

Code User Interface logic pulls the metrics posted in the database tables time to time through queries and displays them in a structured template rendered by a webserver at a defined refresh rate. The Integrated operational metrics and incidents monitoring dashboard is a web-based solution that is accessible from anywhere in the network and consists of. Data Centre Topical section of the dashboard which displays all the major / topical events of the data center, metrics related to the Meta data clusters, Active and passive meta-data server IP and name as shown in Figure 6. Apart from Meta information it displays the filesystem status in terms of disk space availability out of total space provisioned under each filesystem label. It also provides information about the backup status, type, media id and no. of free media available.

Operations and alerts counter section of the dashboard displays all alerts and operational incident created under various segments of the data center kept under active monitoring [9]. Storage Area Network segment displays the usage of disk based storage tiers with an elegant water fill display, with the change in the fill colour denoting the change of data water mark level respectively. The filesystem table below, shows the filesystem label, each filesystem managed by which MDS, No. of users for each filesystem, no. of files / directories in each filesystem. Tape drives information on both the clusters with drive status, SNFS ID, hardware serial no, alias name, ACSLS id, SLC id and the device path under single table view.

Filesystem Stripe Group Information segment displays the stripe group (sg) and disk information for each filesystem, on-click and selection – filesystem name, sg ID, type of disk, sg alias, sg total space, sg free space, sg free percentage, type of device, tier, status, sg read and sg write status. Ticket / Incident Call Log Creation segment as shown in Figure 7.displays the options to create an incident / ticket in case of any incident / operational activities that happens.

Incident Queue segment displays the various type of incidents created, uploaded and auto generated and pushed into the database to perform some action to be taken on the individual ticket /incident on selecting a particular ticket / incident on selecting the radio button.

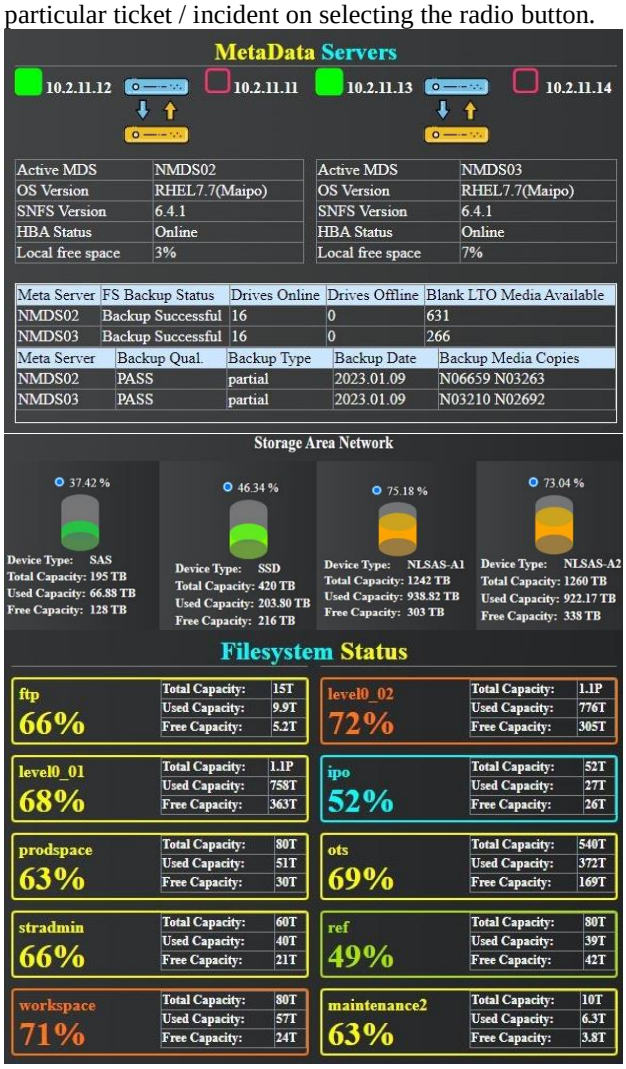

Figure 6: Data-center Topical – Metadata Servers and Storage Metrics

Incident View and Update segment displays the in depth details of each selected incident / ticket with provision to update the status, Incident handler and current summary of the ticket / incident.

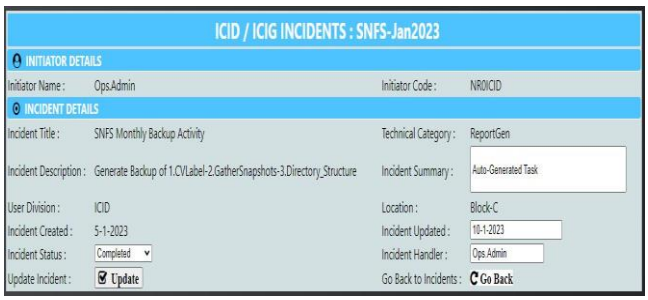

Figure 7: Tickets / Incidents view –Detailed

5. CONCLUSION

Digital remote monitoring helps to make Data centres more reliable and efficient. This method of monitoring the parameters of a physical data center's infrastructure using the multi-level approach helps in reducing the occurrence of events and event identification to the node level. Hence maintenance can be done apriori due to the alerts on the DASH board and exactly pinpoint the occurrence of event for corrective action. Monitoring allows for proactive response, and the overall good health of a Data center. While monitoring does lead to more stable and reliable Data Centre the experimental results show that the remote monitoring tool contribute to the minimization of events (alerts). The system administrators are provided with a wider and more accurate overview of the compute, storage and network elements of data center infrastructure operations.

REFERENCES

[1]. Mikhail Khovricheva, LiubovElkhovskayaa, Vladimir Foninb and Marina Balakhontcevaa, **"Intelligent Approach for Heterogeneous Data Integration Information Processes Analysis Engine in Clinical Remote Monitoring Systems"** published in the 8th International Young Scientist Conference on Computational Science 2019.

[2]. **Planning and Implementing a Secure Remote Data Centre Management Solution**. Published in the "Remote Data Centre Management" of LANTRONIX manual.

[3]. **Data Centres, Closets and Servers rooms Remote monitoring solutions 2009-2017**Monnit Corporation Hand book.

[4]. Stephen Cooper **"6 Data Centre Monitoring Tools"** White paper published on VPN_news July 28, 2022

[5]. **Networking monitoring challenges in heterogeneous environment** a white paper published in codilime on 27th October 2020

[6]. William Tschudi, TengfangXu, Dale Sartor, and Jay Stein – **E Source Building Technologies Department, Environmental Energy** Technologies DivisionErnest Orlando Lawrence Berkeley National Laboratory, University of California 94720-8134 USA Road Map for Public Interest Research on High Performance Data Centres.

[7]. Victor Avelar White Paper 237 **"Digital Remote Monitoring and Howit Changes Data Centre Operationsand Maintenance"**

[8]. Martin Eriksson Project report on **"Monitoring, Modelling and Identification ofData Centre Servers"**

[9]. VojkoMatko, Barbara Brezovec, and MiroMilanovic**"Intelligent Monitoring of Data CentrePhysical Infrastructure"** published in Appl. Sci. 2019, 9, 4998; doi:10.3390/app9234998.