Advanced Technology Academic Research Center (ATARC)

GSA DCOI Managing Partner PMO

DCIM/Energy/Metering Project Team

DCIM's Expanding Role: Managing Our Critical National Infrastructure of Data Centers

Cloud & Data Center Working Group Document

November 2017

Table of Contents

1 Introduction 1
2 Policy 1
2.1 DCIM tools to drive DCOI Compliance 1
2.2 Operational Efficiency 2
3 Security 3
4 Technology – Current Capabilities and Future Possibilities 5
4.1 How is Machine Learning Relevant to DCIM and Automation? 5
4.2 What is Analytics? 6
4.3 How is Analytics Relevant to DCIM and Automation? 6
5 Conclusion 7
6 Acknowledgements 7

1 Introduction

There has been recognition over the past several years that our nation's data centers should be considered as part of our critical national infrastructure because of the services they provide and the data they contain. Whether a data center is government owned or privately held they need to be defended and optimized with the best available tools and technology.

It is far too common to hear about data breaches and inefficiently run IT infrastructure. But these issues help to advance policies and procedures to refine data center optimization management. There is no major organization, private or public, that can effectively execute their core business mission without fast, reliable and secure IT functionality. No longer should there be an ordinate/subordinate relationship between the core business and IT; they are now one in the same. The foundation of which, is the data center.

No matter from where an organization accesses compute, either the cloud or their own enterprise data center, they must leverage the technologies available to optimize their physical operations and the logical delivery of their business applications. DCIM tools, we propose, can serve as the basis of data center optimization management and bring together physical whitespace management, energy management and play a key role in compute management. This convergence of monitoring and management has the potential to bring to light the importance of DCIM tools in providing a platform for the business, facilities and IT components in an organization to better manage the entirety of their compute portfolio.

We brought together individuals from the public and private sectors to propose ideas and methods that will generate discussion, be put into practice and move the needle forward for how we operate this important part of our critical national infrastructure. We categorized our thoughts into three areas: Policy, Security and Technology.

2 Policy

The Federal Data Center Optimization Initiative (DCOI) is just one attempt to provide guidance, goals and oversight for how data centers and the compute services they deliver serve the Federal Sector and ultimately taxpayers. These best practice metrics that Federal agencies strive for can also be used as a standard to measure private sector data centers.

2.1 DCIM tools to drive DCOI Compliance

The benefit of a robust DCIM tool is the ability to provide dashboard insight into the critical metrics of the data center. The DCOI metrics include PUE, server utilization and cost, virtualization insight, energy metering insight and data center facility utilization. The principal financial benefit of optimization of the data center is the opportunity to re-capitalize the infrastructure and associated costs.

Here is a breakdown of how a DCIM tool can provide this information:

- Power Usage Efficiency (PUE): A DCIM tool should capture all elements of power consumption to include the UPS, floor based distribution units, cooling units and the IT load at the cabinet level. It should also capture the power telemetry data outside the immediate white space that affects the PUE. This consists of overhead power consumers like switch gear, lighting, ancillary power, etc., which is a challenge to capture.

- Server Utilization and Cost: Critical to measuring DCOI progress and compliance is understanding the utilization of the individual rack mounted assets including: basic 1-U devices, network equipment, individual blade consumption, etc. The ability of a DCIM solution to capture individual asset utilization provides insight further up the IT stack and knowledge of power usage. As important, physical asset information, for equipment to be excessed (decommissioned, data wiped and removed) is identified and located. The enterprise can utilize this information for asset life-cycle management, modernization strategy and asset reduction/consolidation planning. DCIM can provide valuable usage trending information, charge-back cost modeling and management information that can be used to improve service catalog offerings.

- Virtualization insight: Along with physical asset utilization, DCIM can also drive modernization and consolidation strategies. Understanding both the virtualization ratio and physical asset utilization will illustrate excess capacity and provide guidance for physical and virtual consolidation feasibility. In conjunction with server utilization information, progress towards optimization can be achieved by right sizing the optimal virtualization and utilization loads. In other words, using real-time telemetry to optimize and automate production output to utilize the most efficient compute when and where it is needed at any given time.

- Energy Metering: Management tools can identify the various power and cooling chain entities that currently are metered, and as important, those that are not metered. Granular and robust metering provides insight into utilization, costs and identification of inefficiencies within the data center. The ability to identify overhead non-IT loads provides opportunities for efficiency improvements and cost reductions.

- Facility Utilization: The most significant benefit of understanding facility utilization is identifying opportunities to optimize cabinet usage, reclaim capacity and defer capital expenditures. This is an ongoing task because no data center is a static environment. It is always in a constant state of refinement due to cyclical IT workloads, hardware refresh life-cycles, and energy consumption affected by environmental and production processing parameters.

2.2 Operational Efficiency

In addition to policy compliance, DCIM tools manage data center operations and reveal efficiencies and inefficiencies which can then be used to obtain necessary budget and resources. Here are examples of ongoing benefits of DCIM:

- Efficiently design the whitespace for cooling, power, cabling and asset management.

- Used in conjunction with an RFID system, a thorough management of the physical plant can be realized.

- Increases cybersecurity resilience through faster responses to security alerts by revealing the location, configuration, and owner of most valuable assets (MVA).

- Provides auditors detailed system information to determine software and hardware license compliance.

- Assist and plan migrations/implementations with a better understanding of both physical and application requirements illustrating dependencies that affect production operations.

- Management of human resources can use the information from DCIM tools to manage overtime costs and augment staff for cyclical increases in workloads.

- Provide valuable information for ongoing development of a services catalog by tracking what services are highly utilized and thus could provide revenue.

- Analyze DCIM data in conjunction with information received from a CMDB for compute rationalization. This can more efficiently leverage assets and investments to determine, what can go to the cloud, what can be discarded or subsumed into another more efficient business process and what can be moved to a more efficient data center.

- Show where IT modernization efforts should be focused through the combined information of application performance, uptime and energy usage.

- By showing and monitoring energy efficiency and usage, both the CIO and CFO components will see the relationship between their efforts and the effect each has on the operating cost of day-to-day operations.

- Convergent monitoring of disparate systems and organizational stovepipes leads to the cooperation and co-management of a data center by IT, Facilities and Physical Security. Through cross-component understanding of mission and change management, the data center can be operated and concurrently maintained, thus reducing operational cost and IT production down time.

3 Security

The Trump Administration's agenda is to identify the "risk and magnitude of the harm that would result from unauthorized access, use, disclosure, disruption, modification, or destruction of IT and data." To support this, we need to leverage the data and benefits provided by a robust DCIM implementation to secure data center assets, manage risks, and manage the level of redundancy and resiliency a data center offers.

Data Centers are not subject to just cyber-attacks but those that start as a physical incursion as well. Roughly speaking, about 20% of data breaches occur following a physical instigation of events. Illegitimate entry can be gained to retrieve equipment that contains data that was slated to be excessed or destroyed. It could also be generated from a false problem/repair ticket that allows a person posing as a technician to gain physical access to the whitespace and thus would be inside the cyber perimeter and gain access to other systems both physically and digitally. A denial of service to disrupt services could be generated by a fault in the energy supply system. To counteract this, the IT operational component needs to coordinate scheduled site access with the physical security and facilities management components. All parties should have access to a change management system which leverages DCIM management capabilities as a check and balance for determining whitespace access.

Also, regular audits of infrastructure resiliency, through simulations, controlled failover tests, power infrastructure firmware updates, should be done regularly. This ensures an accurate picture of risk to the infrastructure is fully understood and included in the existing security assessment workflow.

DCIM can be used to validate tiered data centers. All of the proper inputs could be provided in the DCIM dashboard (based on TIA-942 ANSI) to quickly identify sustainable and secure data centers in our country's critical data center infrastructure. The higher the tier rating the more sustainable and concurrently maintainable a data center becomes. Lower tiered data centers may not have the physical security presence to properly thwart a targeted physical incursion. They also do not have the facility redundancy to keep the data center operational in case of an energy supply incident.

DCIM affords the opportunity for the industry to adopt better cross-talk between disparate monitoring systems; a lingua franca for data centers. This would be a step in removing communication barriers between data centers in a disaster recovery situation making recovery easier and more manageable, especially in a national attack. Today, DCIM bridges the gap between IT and Facilities and in the future, it can facilitate the communication of critical compute capabilities in a national cloud based data center D/R network. The information DCIM provides eventually could be used to direct and coordinate local, state and Federal efforts to find appropriate compute resources quickly and bolster emergency fuel rationing decisions.

Other challenges involve the actual instrumentation and access to the critical telemetry data required for effective data center management. If an agency wishes to understand its own PUE, it must have the ability to access critical power data not just within the whitespace, but also the mechanical and "grey" space. Often these systems are within the management of a different organization with a different management tool. These mechanical systems may even be on a separate TIC (Trusted Internet Connection) creating significant visibility challenges. Further, to understand the level of virtualization or asset utilization, data center managers require access to the server virtualization information as well as to power/CPU utilization. All three co-management components of a data center: IT (including cyber-security), Facilities and Physical Security all must work together to supply the needed system accesses and to protect those same access vectors. DCIM can help pull these areas together in a coordinated management effort.

4 Technology – Current Capabilities and Future Possibilities

DCIM tools will be a force in and in some cases drive IT modernization. Currently data center optimization is mainly geared towards facilities enhancements and is at various levels of readiness in its ability to provide real-time telemetry of the data center environment. We can measure PUE and temperatures and use it to make IT production processing decisions. With increased monitoring and measuring of the entire data center ecosystem, we will be able to also analyze and identify applications and business processes that use energy inefficiently and thus point development to those applications and processes for modernization.

This will only come about with a modernization in thinking. Data center engineers, CIOs, CFOs and CEOs have an opportunity to change what defines their roles and the data center's place in the organization. As this paradigm shift occurs, new possibilities for automation, control, and artificial intelligence in data center optimization management will become possible.

On the horizon are game-changing technologies with automation, robotics, artificial intelligence, self-orchestration, power, cooling, data storage and hardware size. All of these areas need to be coordinated, managed and optimized. The next generation of DCIM will need to incorporate these enhancements into their toolkits to continue to refine data centers and their business processes.

For example, machine learning is the ability of computer-based algorithms to learn from data and make predictions and is a sub-set of artificial intelligence. There are three general kinds of machine learning algorithms:

- Supervised learning works by presenting an algorithm with known input-output data sets that it uses to adjust an underlying predictive model. Linear regression is a form of supervised learning.

- Unsupervised learning identifies patterns in data without being presented with known input-output data sets and can be used to identify features such as clusters in data sets without knowing the context or meaning of the clusters.

- Reinforcement learning uses a reward system to learn, especially in the context of performing a task. When the algorithm performs the task well, it is rewarded with positive feedback.

4.1 How is Machine Learning Relevant to DCIM and Automation?

Machine learning, especially supervised learning, enables predictive models to be calibrated based on real operating data, and it allows computer-based systems to adapt to a changing environment. Predictive models that have been calibrated with machine learning can then be used to assist people with decision-making, to make optimized decisions automatically, or to automatically take control actions that produce an optimized outcome.

For example, machine learning can be used to calibrate a model that is designed to predict the impact of cooling controls, such as ON/OFF status, airflow rate, and temperature setpoints, on the values of temperature sensors located in a data center. Once the model has been calibrated via machine learning, the expected influence of control actions on data center temperatures can be used to execute optimized control actions. Such a machine learning algorithm can be executed repeatedly, so that the predictive model can adapt to changes in the data center that may impact the influence that cooling equipment has on the data center.

Machine learning can also be used to predict the need for maintenance. Historical data from control systems or equipment sensors, combined with historical records from a maintenance management system, can form a training set for a supervised learning system. Once trained, the underlying predictive model can be presented with new sensor data so that it can determine whether or not an event necessitating maintenance is likely to happen. Such a system may even be able to provide recommendations for specific corrective actions that should be taken.

Another use case for machine learning in data centers is the calibration of models that predict server or storage resource usage or the likelihood of a service level agreement violation. These models can then be used to optimize the scheduling of resources to increase revenue and reduce operating cost while meeting service level agreements for requirements such as latency.

4.2 What is Analytics?

Analytics is the detection, diagnosis, and communication of patterns or relationships in data, often through the use of statistics, predictive modeling, and machine learning. The presentation, display, and visualization of data, whether measured or derived, are key aspects of an analytics system or tool. Current forms of analytics focus on visualization and reporting. We should now look to a period of shared analytic data feeds from all over the organization. These feeds take consideration of analytics that may reside outside the data center, but may be useful in forming better predictive and learned models. This data may include financial, human resource or even weather information. Predictive modeling and machine learning enable the visualization of derived data in the form of what-if or what could be scenarios.

4.3 How is Analytics Relevant to DCIM and Automation?

One of the common uses of analytics in data centers is to display and visualize metrics from servers, switches, and storage devices. Analytics are also used to display and visualize data from power, cooling, and building management systems. Analytics of this sort are becoming increasingly important as more and more data center infrastructure assets are embedded with communicating microprocessors, enabling the Internet of Things.

With the growth of predictive modeling and machine learning, analytics are increasingly being used to run what-if scenarios to analyze the impact of passive events such as equipment failure and extreme weather on data centers. They are also being used to run what-could-be scenarios to analyze the impact of active events such as what will happen if a cooling unit is turned off for maintenance or if a workload is moved from one host to another.

Analytics are also being used to assist with optimizing IT workload decisions. For example, analytics can help planners and operators understand the cost and performance implications of running workloads in the enterprise versus in the cloud so that they can make the best choice between these two options.

5 Conclusion

As the data center becomes understood as part of the enterprise's critical infrastructure, it is clear that a comprehensive DCIM solution is a fundamental requirement. Moving forward, DCIM enables the opportunity for further innovation. Innovation in technology, process, and perhaps most importantly, in the way people think about what is possible for efficient and secure data center operations.

DCIM tools will provide the necessary information through an aggregated dashboard of all the systems (IT-production performance, CMDB, Change Management, Facilities-BMS, Physical Security-plant access) to gain total insight into the day-to-day management of data center operations. This convergent monitoring will enable optimization management for production performance, energy efficiency and help keep the plant and its contents physically secure.

6 Acknowledgements

The Advanced Technology Academic Research Center and the DCOI Managing Partner PMO appreciate the contributions of the following members of the ATARC Cloud & Data Center Working Group's DCIM/Energy/Metering Project Team:

- Jeffrey Chabot, Schneider Electric
- John Esparza, Schneider Electric
- Tim Harvey, ATARC
- Steve Naumann, GSA DCOI Managing Partner PMO
- Bob Thronson, Vigilent
- Pete Tseronis, Dots and Bridges
- Anthony Vicinelly, Nylte Software