DCIM Policies: Automating Data Center Standard Operating Procedures

WHITE PAPER

from

GreenField Software Private Limited

June 2017

Introduction:

Recent high profile data center outages have again brought to fore that while a lot of equipment and facility investments have been made on redundancy and disaster recovery, there is still high reliance on manual operations. Surveys[1] have indicated that human error ranks as the second highest causal factor in data center outages. This in turn has been attributed to failure in adherence to standard operating procedures (SOPs) which are usually well defined but forgotten - or worse, not made aware to operating staff. Beyond procedures to mitigate risks of data center outages, other aspects in a Data Center SOP document relate to maintaining access and data security, energy efficiency, equipment replacements, preventive maintenance, audit trails and more – what are generally outlined as best practices.

We have seen however the pitfalls of keeping the SOP as a manual and not automating the procedures. The logical home for automated procedures is the DCIM (Data Center Infrastructure Management) which essentially is an Operations, Planning and Management software for a Data Center. **These set of operational procedures are packaged into a "DCIM Policies" framework which link into different modules of the DCIM Software such that the DCIM detects any potential violation and sends alerts**. It could also prevent an accidental mishap, such as overloading a Rack or taking down an electrical device for preventive maintenance without providing for back-up.

This paper outlines twelve key operating procedures that should be part of "DCIM Policies". We have split this into three broad categories: **Risk Management**, **Governance and Efficiency Management**.

I. **Risk Management:** This tries to mitigate a Data Center Manager's nightmare of an unplanned downtime, or worse an extended outage that disrupts business application availability, causes massive financial loss and damages an organization's reputation. [Alarm Policy, Escalation Policy, Redundancy Policy, Disaster Recovery Policy]

II. **Governance:** Streamlined governance with chain of command, checks & balance system, and audit trails are few of the universal best practices any organization adopts to ensure voluntary or statutory compliance measures. This applies to Data Centers as well. [Security Policy, Data Retention Policy, Approval Policy, SLA Policy]

III. **Efficiency Management:** The Green Grid, ASHRAE and Uptime Institute have defined number of KPIs for an energy and operationally efficient data center. It is up to each organization, based on their own business priorities, to decide which KPIs matter to them, the benchmarks they wish to maintain and accordingly decide the policies that would help them get there. [PUE Policy, Rack Load Policy, Replacement Policy, Preventive Maintenance Policy].

I. Risk Management:

1. Alarm Policy: which devices and parameters need what frequency of monitoring, and deciding their threshold levels.
   (i) Expected operating temperature and humidity range: This is a high-priority decision factor under DCIM alarm policy to prevent smoke, fire or damage to devices. Ideally, we should include the operating temperature and humidity ranges at (a) device-level: the air intake of IT devices placed on racks, (b) at rack-level: lower, middle and upper portions on rear and front, (c) row-level: for each hot and cold aisle, and (d) room-level: for general comfort of operating staff. As each level of monitoring involves additional cost, the minimum monitoring and threshold levels for DCIM Policy should be set for rack-level on front. In such a situation, we need to ensure during warranty period for new equipment purchase[2] that they can withstand the full range of 10°C to 35°C inlet temperature (50°F to 95°F) and relative humidity of 20% to 80% or 21°C (70°F) dew point. This is defined by the ASHRAE Class A2 allowable temperature and humidityrange[2]. Minimum rack temperature threshold is therefore set at 21°C (70°F) to avoid dew formation. Upper level settings are typically 26°C as warning and 28°C as critical.
   (ii) Other Alarm Policy settings depending on Data Center's business needs: PUE, fuel level in tanks, rack power load, UPS load, CPU & memory utilization of servers, preventive maintenance and Diesel Generator mock-run dates.

2. Escalation Policy: It is important to establish a clear-cut escalation process to know as to how and when alerts should be escalated in a data center. Escalation policy need to be developed and rehearsed to ensure the chain of command is informed and the appropriate resources are brought to bear as the situation develops. An escalation table can be defined in DCIM, which outlines the protocol, channels for escalating issues and contact personnel with the appropriate expertise. For example, see the table below defining the escalation procedures for a UPS related problem.

Time Frame	Escalate To	Method
0+1 min	Facilities engineer on site	SMS
0+3 min	Supervisor	Phone
0+6 min	UPS vendor's support engineer	Phone
0+9 min	Data Center Manager	SMS& Email
0+10 min	All users impacted by UPS failure with necessary actions	Email

3. Redundancy Policy: A data center is a mission critical place, where we need at least two of everything (N+1/N+2 configuration) so that our business stays afloat and has maximum uptime, whether this applies to servers/VMs, networks or storage devices. A redundancy policy is important to be defined in DCIM depending on customer's needs i.e. whether to have an N+1, N+2 configuration. It is not just redundant components that are important but also the process to test and make sure they work reliably such as scheduled failover drills and research into new methodologies. If we cannot have two, we need to figure out how we can cobble together a replacement system if the primary equipment becomes unavailable or fails. For example: if we didn't have the budget for an official clustering or load balancing solution, we would need to develop a failover plan with a backup DCIM server.

   There should be a policy in place to document the entire failover procedure. Documenting the procedure would help any of our coworkers to follow the steps, in case of any disaster.

   When it comes to redundant components, it is important to have all devices identical in every possible way to make supporting them as predictable as we can – they should have the same manufacturer/model number, run the same operating system, have the same drivers and hot-fixes, plugged into the same ports in different switches or PDUs. Another important point in redundancy is that it gives us tremendous leverage when it comes to applying changes since we can take half of a redundant pair down to move or upgrade it, then do the same for the other half. However, we should never do this without leaving a gap of time in between to make sure the first change was successful.

4. Disaster Recovery Policy: It is crucial to have a disaster recovery plan in place with metrics of RTO (Recovery Time Objective) and RPO (Recovery Point Objective) well defined in the SLA. A data center disaster is when none of redundancy options are available: a complete power outage, for example, is a disaster. In such a situation, how quickly can we recover to get at least one section of the data center up and running (RTO). How much longer will it take to recover to the point before the initial power failure leading to complete outage (RPO).

   Every single change in the data center should have a recovery/backup plan associated with it. For example:
   - How are we going to put things back to the way they were if something fails during upgrade of an electrical equipment?
   - Should we first have a drill that both DG sets are running in case of power failure from the mains?
   - When should we inform business users to make their own contingency plans should there be any disaster arising out of the change?

   We need to document this plan and make it as elaborate as possible. DR Policy should ensure we capture notes, screenshots and other supporting evidence to build more stringent measures/policies to avoid recurrence of a similar event.

II. Governance:

5. Security Policy: Includes role-based access. This should define (a) registered user-ids with specific roles, the pages they have access to and which reports only certain individuals will receive at specified frequencies. (b) If a change needs to be rolled back, who will have approval rights. (c) Who is granted permission to scan historical/archived data if a problem has been identified, and we need to trace this through audit trail.

6. Data Retention Policy-
   A data retention policy, is an organization's established protocol for retaining information for operational or regulatory compliance needs. This would involve (a) number of days or months we retain data in production; (b) classify data according to importance and use (example: audit versus analysis); (c) archive and purge depending on retention period, importance and use; (d) define exceptions, like we may purge temperature data beyond 90 days ***except*** for those days when it breached thresholds beyond 10%; and (e) medium in which archived data will be stored.

   Given sensor data grows rapidly, it has storage and energy cost implications for the data center. It is recognized that a significant proportion of the data stored is either unnecessary or duplicated. Hence a well-structured data retention policy should be mandatory for data center operations. There are essentially three main objectives in developing a data retention policy, which can be summarized as follows:

   a. To keep important records and documents for future use or reference;
   b. To purge records or documents that are no longer needed; and
   c. To organize records so they can be searched and accessed at a later date.

   Drivers for a Data Retention policy:
   - Cost savings through data storage reduction;
   - Simplified, less expensive data management; and
   - Regulatory compliance/Governance (legal discovery, protection of privacy)

   Non-editable archiving to secondary storage and purging must be automated in the data retention policy, which could also include Workflow approvals and Move-Add-Changes for future audit tracking.

7. Approval Policy for Provisioning and MACs: Provisioning of power, space, cooling and network ports when adding more customers, applications and IT devices can be a contentious one as there are conflicting demands of finite amounts of these resources. An approval process with linkages to Power and Network Chains ensures that one has not over provisioned or under provisioned any section that can lead to

a power or network trip. A somewhat similar situation arises out of Move-Add-Change (MAC) – an approval process ensures that everyone knows about, agrees upon, and supports the proposed change(s). The changes and the associated approvals should be retained per the data retention policy so that one can trace back to events as well as analyze if any change resulted in an improvement or otherwise.

8. SLA Policy: An SLA in a data center contract serves 3 main purposes:
   - Establishes specific levels of availability that are guaranteed by the data center.
   - Sets communication protocol for any issues or uptime-impacting events that may arise.
   - Lays out policies and procedures revolving around planned maintenance events by the data center (timing of such events, the communication procedures, etc.)

   These agreements typically contain numerous measurable components that all revolve around meeting these key objectives.

   Below are examples of Data Center SLA Policy that cover three infrastructure service metrics like power, temperature, and network availability.

   1. Power- If the business ordered redundant Power i.e. power from dual source, a client will generally be offered a certain uptime commitment by the data center for redundant power. This should be well laid down in the SLA policy defined DCIM.
   2. Temperature- Subject to certain conditions, a facility generally will maintain over a 24-hour period, an average room temperature at the premises of 72 degrees Fahrenheit + or -5 degrees (the "Temperature Range").
   3. Bandwidth - Many data centers have a guaranteed uptime on network availability to their clients. This would be well defined under the SLA policy.

   Automatic Alerts to customers have to be generated depending on allowed variance on each SLA component, which may be measured on daily, weekly, monthly or quarterly/annual basis.

III. Efficiency Management:

9. PUE Policy: The Power Usage Effectiveness (PUE) metric is an industry standard for reporting energy performance of data centers. Data Centers strive for a PUE of 1.0, which represents a hypothetical efficient data center where energy is used exclusively to power IT, and there is no energy loss or overhead in the system. Data center power and cooling are the two major factors affecting power efficiency today. Organizations need to take several measures to ensure better PUE. It is thus vital to

lay down the PUE policies and standard operating procedure, to improve efficiency and save costs. By definition,

$$\text{Power Usage Effectiveness (PUE)} = \frac{\text{Total facilities power}}{\text{IT equipment power}}$$

PUE policies in DCIM would be as follows:

1. PUE range values: A data center may define maximum acceptable average annualized PUE depending on external temperature conditions. Alerts would be sent accordingly. Newer data centers (or where DCIM has been recently implemented) which do not have a year's PUE values maintain a daily/weekly/monthly/quarterly average.
2. UPS load: Matching UPS load to the system load improves PUE. If the UPS is only loaded to 30% capacity, efficiency will be much lower. Hence, we may define a lower threshold level of UPS load which should generate alert. Of course, an upper level load must also be defined to maintain balance of power load of the downstream devices connected.
3. Carbon Usage Effectiveness (CUE): Green Grid, the authors of PUE have also defined another metric, CUE which is dependent on PUE. Sustainability-conscious organizations, maintain CUE as anoter metric and may ask for this to be included as well for generating alerts.

10. Rack Load Policy: A data center must have a proper rack load policy in place in terms of power load, temperature, weight, U-space and ownership allocation. Threshold or procedure breaches in rack loads need to generate on-screen warnings or alerts.
    a. Rack Power: Racks are allocated power loads, say 8KW. If already loaded with devices running up to 7.5KW, then a rejection should first happen if the workflow approval request (see #7 above) had this Rack as an option to place a server of 900W. If the operator still attempts to configure the DCIM with this server, an on-screen warning would be displayed. If the operator still places the server, and the rack load has jumped beyond 8KW, then immediately a critical alert would be sent as per escalation policy.
    b. Rack Temperature: Rack temperatures are usually defined (see # 1(i) above). If temperatures exceed thresholds, alerts would be sent.
    c. Rack Weight: Depending on floor load bearing capacity, a certain weight capacity is allocated for each Rack. Alerts have to be configured accordingly.
    d. Rack U-space: Typically some U-spaces in the rack are kept free, which should be defined. If not an alert, at least an on-screen warning should appear when an operator is committing this procedure breach.
    e. Rack Ownership: Racks or even U-spaces may be allocated to a business owner. Placing a device of a different owner on this should generate a warning or alert.

    Information about a device's power rating, recommended ambient operating temperature, weight and dimensions are usually maintained in the DCIM's OEM

library. DCIM automated policy is programmed to check on this OEM library before generating the warnings or alerts.

Other considerations in the Rack load policy:

- Rack Power: Do not overload the AC supply branch circuit that provides power to the rack. The total rack load should not exceed 80 percent of the branch circuit rating.
- Rack Weight: Load heavier components first and load the rack from the bottom up. Same consideration when coupling or baying the racks: need to balance weight load between or among the racks, placing the heaviest components at the bottom.

11. Replacement Policy: In this policy, we define the life for each category of device in the Data Center.
    - Alerts can be configured when a device is coming near end of life. This helps in decommission planning. For computing devices, this would include data migration, besides procuring replacements, decommissioning old and installing new in a seamless manner.
    - Alerts could also be set-up before the actual replacement so that affected users can make contingency plans should something go wrong during the transition.

12. Preventive Maintenance Schedule Policy: As a common practice, most changes in the data center are planned after hours or during non-critical periods. Preventive Maintenance and upgrade schedules with expected downtimes can be defined in DCIM. The following can then be configured:

    - Switching off non-reachability alert during this downtime
    - If actual downtime exceeds the expected downtime by a certain margin, alert would be sent
    - Validating from the Power and Network Chains that scheduled preventive maintenance of a device does not have a cascading & catastrophic impact. If it does, an alert would be generated.

Summary:

Every process/operating procedure involved within the data center should have a policy behind it to help keep the environment maintained and managed. Deviations from acceptable range should be automatically detected for immediate corrective action and where possible even prevent a violation. Besides helping to avoid data center failures, automated policies help in better governance and driving efficiency improvements. With increased adoption of DCIM as operations, planning and management software for data centers, these Standard Operating Procedures have become their core foundation, as set of best practices similar to leading ERP software.

References:

1. Cost of Data Center Outages (2016): Ponemon Institute
2. EU Code of Conduct on Data Centers (2014): European Commission Directorate General
3. How to Prepare and Respond to Data Center Emergencies: White Paper 217 by Schneider Electric
4. Cisco Unified Computing System Site Planning Guide: Data Center Power and Cooling: White Paper
   http://www.cisco.com/c/en/us/solutions/collateral/data-center-virtualization/unified-computing/white_paper_c11-680202.pdf
5. http://searchdatacenter.techtarget.com/tutorial/Data-center-disaster-recovery-planning
6. http://searchdatabackup.techtarget.com/tip/Developing-an-electronic-data-retention-policy
7. http://policy.ku.edu/IT/data-center-standards
8. http://searchdatacenter.techtarget.com/tip/Top-ten-considerations-for-data-center-SLAs
9. www.cisco.com/c/en/us/solutions/.../data-center.../white_paper_c11-680202.pdf
10. www.apc.com/salestools/NRAN-8FL6LW/NRAN-8FL6LW_Ro_EN.pdf

GreenField Software Private Limited is an Indian venture pioneering intelligent infrastructure management solutions. The product portfolio includes GFS Crane®, a policy-driven DCIM suite, with installations in enterprise data centers of Financial Services, Telecom, Power Utilities, Media, Oil & Gas, Discrete Manufacturing and Higher Education.

For more details:

Email: sales@greenfieldsoft.com

Visit: www.greenfieldsoft.com