RFCODE

WHITE PAPER

Real-Time Data, Continuous Visibility, and Predictive Analysis

The Cornerstones of Data Center Efficiency

Introduction

We live in an always-on, always-connected digital world. That is only possible because data centers enable every aspect of modern civilization from social connections to global markets. These critical facilities are, truly, the engines of commerce. They also use a lot of energy—3% of the world's electricity.[1] And the truth is that much of that energy is wasted. Data centers are expensive, both to build and to run. A company cannot afford to let its data centers, whether multimillion-dollar enterprise facilities or smaller satellite and edge facilities, operate at anything less than optimal efficiency. The way to improve agility and efficiency and ensure higher performance at a lower cost is with real-time monitoring, capacity planning, and predictive analysis technologies.

> *"A typical rack of state-of-the-art servers ... uses more that $17,000 per year in electricity. Given that data centers can hold hundreds of such racks, they constitute a very energy-intensive building type."*
>
> *- Best Practices for Data Centers: Lessons Learned from Benchmarking 22 Data Centers*

The media tends to focus on hyperscale data centers run by household names like Facebook, Amazon, and Google, the latter with more than 60 facilities each.[2] There are about 500 of these data centers, the kind that fill large warehouses or entire multi-building campuses, which is just a small percentage of the over 8 million data centers around the world.[3] Most data centers have smaller footprints, from a few hundred square feet in a building otherwise used for office space, to modular data centers housed in shipping containers, to small server closets.

1 Tsunami of data could consume one fifth of global electricity by 2025, The Guardian, December 11, 2017, https://www.theguardian.com/environment/2017/dec/11/tsunami-of-data-could-consume-fifth-global-electricity-by-2025

2 Hyperscale Data Center Count Passed the 500 Milestone in Q3, Synergy Research, October 17, 2019 https://www.srgresearch.com/articles/hyperscale-data-center-count-passed-500-milestone-q3

3 The Data Center Dilemma, Data Center Knowledge, April 19, 2019, https://www.datacenterknowledge.com/industry-perspectives/data-center-dilemma-our-data-destroying-environment

No matter the size of the installation, data centers use a lot of energy—in some cases, up to 200 times the electricity of office space.[4] Most enterprise data centers spend millions of dollars on electricity each year. That's real money in any economy, much less in today's highly competitive global landscape and troubled economy. Add in rising energy costs throughout much of the industrialized world and the increased demand for sustainability and energy efficiency to slow climate change and one would think most data center operators would consider improved energy efficiency their highest priority.

Yet this is not the case. Surveys indicate that energy efficiency has broken into the top three[5] concerns, but data center managers still consider maintenance and extending the product life cycle to be more important. The common denominator of these top two concerns is ensuring uptime. In an always-on world, a data center cannot go offline. A seemingly minor change or miscalculation and an overloaded circuit trips. A computer room air conditioner (CRAC/CRAH) unit malfunctions and unmonitored servers overheat. An uninterruptible power supply fails and the facility goes dark.

These are concerns that keep people up at night, because data center outages are expensive. In 2019, the Ponemon Institute and IBM reported that the global average total cost for a breach has increased 12% since 2014, to $3.92 million.[6] And that number includes only readily quantifiable impacts—system downtime, business disruption, lost customers, lost revenue, decreased productivity, equipment repair and the like. The cost to a business's reputation is harder to measure, but it lasts longer and affects the bottom line far more than the expense of the actual event.

Most operators use a multi-pronged approach to ensure uptime: incorporating redundant systems in their design, applying best practices for operational and maintenance procedures, and using innovative and integrative technologies such

4 Best Practices for Data Centers: Lessons Learned from Benchmarking 22 Data Centers. S. Greenberg et al. 2006. ACEEE Summer Study on Energy Efficiency in Buildings. http:// datacenters.lbl.gov/sites/all/files/aceee-datacenters.pdf

5 Report: Energy Efficiency Not Top Concern for Data Center Managers, DECEMBER 19, 2018 https://www.environmentalleader.com/2018/12/report-energy-efficiency-not-top-concern-for-data-center-managers/

6 2019 Cost of Data Breach Report, IBM and Ponemon Institute, https://www.ibm.com/security/data-breach

as data center infrastructure management (DCIM) systems to improve reliability. And while they are focusing on these areas, data centers are often unable to focus on creating or maintaining efficient operations.

Challenges

Availability and Operating Costs

After the actual computing work, the next biggest power hog in a data center is cooling.[7] Electronic equipment puts out a lot of heat and overheated devices are more likely to fail. That's why, traditionally, data centers have been kept at temperatures more like a refrigerator than a place of business. There are many techniques and technologies data center operators can employ to save energy in their facilities.

Arguably the strategies that offer the greatest potential savings in cooling costs (e.g., free cooling, chiller-free data centers) stem from recent guidance from the American Society of Heating, Refrigeration and Air Conditioning Engineers (ASHRAE) that expands the recommended and allowable temperature and humidity ranges for data center equipment operation. Today, datacenters that used to operate at 55-65°F can run at 80° or even 90°F, and with less stringent humidity limits.

The impact of these changes on energy use can be significant: a data center can save 4-5% in energy costs for every 1°F increase in server inlet air temperature.[8] Consider the example of telecommunications giant CenturyLink, which currently operates 55 data centers worldwide. In 2011 CenturyLink had an annual electricity bill of over $80 million. Starting in a pilot facility, CenturyLink implemented a monitoring program that enabled its engineering team to safely raise supply air temperatures without compromising availability. Based on the results of that pilot study, CenturyLink identified $2.9 million in potential annual savings across its portfolio.[9]

Power distribution and backup equipment also contributes to energy waste in the data center, due to conversion losses, poorly designed power chains and inefficient power supplies and cables. As with cooling, there are many strategies data center owners can employ to improve power efficiency, but the most obvious are on the compute side. Because most data centers provision for peak load—loads that may occur only a few days per year—low server utilization is the status quo in the industry and presents substantial opportunity for increased efficiency.

Experts estimate that 30% of servers are comatose—have not delivered information or computing services for six months or more.[10] At an average cost of $3,000 per server, that is a tremendous amount of data center capital not delivering return on investment. In addition to acquisition costs, these servers still draw power when idle. Additionally, every watt of electricity wasted at the device level has a cascade effect, as still more energy is needed to power the physical infrastructure that supports the device.

Increasing the density of the IT load per rack through consolidation and virtualization can offer substantial savings in not only equipment but also electricity and space—an important consideration if the data center is located where energy supply is constrained or electricity and real estate prices are high, as in most urban areas.

7 How much energy do data centers really use? 17 March 2020, https://energyinnovation.org/2020/03/17/how-much-energy-do-data-centers-really-use/
8 Raise the temperature, Energy Star, https://www.energystar.gov/products/raise_temperature
9 RF Code Delivers Millions of Dollars in Annual Power & Cooling Savings for CenturyLink. RF Code Case Study, 2014. http://resources.rfcode.com/rf-code-delivers-millions-of-dollars-in-annual-power-cooling-savings-for-centurylink
10 Report finds 30% of servers are comatose, https://www.anthesisgroup.com/report-finds-that-thirty-per-cent-of-servers-are-comatose/

32% of companies reported that their average density is currently 6-10 kW/rack, 33% reported 11-15 kW/rack, with 16% reporting higher densities.[11] These higher densities mean concentrated thermal output and modified power requirements. The only way to maintain continuous availability in high density deployments is real-time monitoring and granular control of the physical infrastructure.

Data center operators are leaving these potentially game-changing savings on the table because they are high risk without real-time monitoring and management. And because they focus on uptime and availability over almost anything else.

Capacity Planning

Asset management is a key feature of data center capacity planning. For many organizations, managing assets means recording the name and location of every piece of IT equipment in the data center in a spreadsheet or diagram. This is a labor-intensive, expensive and error-prone way to track valuable assets. Yet, a data center is a dynamic environment—equipment is moved every day, devices are taken offline for maintenance, and new equipment is deployed. Data centers that track assets manually are attempting to solve a modern problem using methodology that dates back to ancient Mesopotamia.[12]

In a data center that employs a static, manually maintained asset management system, staff must physically walk around the data center to conduct inventory audits. If a device is missing from its last recorded location or if information about the device is incomplete or conflicts with existing records, staff must investigate, reconcile discrepancies, and replace lost equipment.

11 Report: Data center rack density is rising, and heading higher, Data Center Frontier, 15 November 2019, https://datacenterfrontier.com/report-data-center-rack-density-is-rising-and-heading-higher/

12 Inventory Management History, Part One, by A. Dolinsky. Almyta Systems. http://www. almyta.com/Inventory_Management_History_1.asp

Data centers that track assets manually like this can face compromised productivity, low morale, and high costs—in terms of labor, penalties for late lease returns, unused equipment, and capital costs for purchasing what may be unnecessary IT equipment. Trying to make capacity decisions using incomplete, inaccurate data is like working in the dark—figuratively and, eventually, literally, because this is an outage waiting to happen.

The Solution: Real-Time Monitoring

Real-time monitoring, capacity planning and predictive analysis technologies help data center operators improve agility and efficiency in their facilities and ensure higher performance at a lower cost.

Continuous Visibility Improves Availability and Reduces Operating Costs

Businesses don't have to choose between availability and savings—they can have both. Real-time monitoring gives businesses the information they need to manage risk, improve efficiency and decrease costs. In an intelligent data center thousands of sensors throughout the facility collect information on temperature, humidity, air pressure, power use, fan speeds, CPU utilization, and

much more—all in real time. This information is aggregated, normalized and reported in ways that allow the operator to understand and adjust controls in response to current conditions.

Consider this scenario: a technician makes an error in replacing a section of the raised floor—installs a solid tile where a perforated one was needed, or reverses a directional panel so the cool air is blowing across the aisle rather than toward the server inlets. This seemingly minor mistake can have major repercussions in terms of air flow, temperature and air pressure. The only way to catch a small issue before it becomes a large problem is through real-time monitoring.

Monitoring has benefits beyond disaster avoidance. Cloud, co-location and hosting providers can use the data collected to document their compliance with service level agreements (SLAs). Monitoring data can be integrated into the facility's building management system (BMS), allowing operators to further automate and optimize control of the physical environment. Visibility at a macro and micro level improves client confidence, streamlines decision making and increases data center availability, productivity and energy efficiency.

Capacity Planning Improves ROI and Increases Productivity

In a data center with a real-time asset management system, the user knows the exact location of every piece of equipment in the facility; can drill down to specifications, maintenance and warranty histories for every device; and may even be able to see the device on a floor plan and in context, with power paths, network connections and dependencies clearly mapped. This user fully understands the current position and status of every piece of equipment in the data center.

This enables the data center operator to correlate real-time monitoring data with asset management information to detect stranded capacity (for example, power is available in a given area but cooling is at its limit); and model "what if?" scenarios for new configurations and predict what would occur should a certain piece of equipment fail. This is a data center that is consistently available to perform today and is ready to meet the demands of tomorrow.

On a tactical level, real-time asset management means less time wasted in inventory reconciliation, fewer penalties associated with late lease returns, and smaller equipment replacement budgets. Warranty and depreciation information is readily available, audits are streamlined and change management is simplified.

On a more strategic level, asset management systems facilitate capacity planning. Building new capacity is expensive: a data center can cost $5-10 million per MW. Co-location is an option, but prices currently range from $1,000-$2,600 per rack each month, and that doesn't include electricity, bandwidth, staff, and migration costs.[13]

A company that buys new compute capacity simply because it is unable to identify that sufficient capacity already exists is making two expensive mistakes: Wasting money to buy capacity it doesn't need and taking funds from other initiatives that could drive business growth. The benefits of comprehensive, real-time asset lifecycle management extend beyond the walls of the data center to influence the financial stability of the company itself—and potentially, that of its customers.

When we think about IT security we usually think about theft—hackers breaking into a site and stealing information. The reality is that some of the most damaging IT security breaches arise from a far more mundane, and completely avoidable, cause: companies simply losing the devices storing sensitive information. Regulations like the Health Insurance Portability and Accountability Act (HIPAA) and the Payment Card Industry (PCI)

13 Colocation Pricing Trends. Data Center Talk, Oct 7, 2015. http://www.datacentertalk.com/2015/10/colocation-pricing-trends-2/

standards contain explicit rules about policies organizations must have in place to safeguard any hardware or portable devices on which the personal information of clients is stored.

The fines associated with asset management lapses are often in the millions of dollars and, when associated class action lawsuits are taken into consideration, can reach the billions.[14] With this kind of money on the line, it's clear that a robust, real-time asset management system is not just a nicety, it's a necessity.

Predictive Analysis Facilitates Business Growth

Real-time monitoring, asset management and capacity planning offer substantial benefits in themselves, but the true potential of these technologies lies in the insights that can be gleaned through detailed analysis of the data collected. Predictive analytics moves data center operations from a reactive to a proactive mode.

Combining environmental measurements with IT information like CPU utilization, server power use and fan speeds gives a clearer recognition of the relationships between compute demand and the physical infrastructure and enables more responsive, integrated control of the data center as a whole. Operators are able to use sophisticated efficiency technologies like dynamic power provisioning and load shifting with confidence. Using machine learning technologies, the complex relationships between millions of pieces of data collected over time can be analyzed to identify connections humans cannot grasp. Rules can be derived and policies integrated into management software to create more autonomous, optimized operation[15], decreasing the risk of human error and improving productivity.

Perhaps your company has more than one data center. If you've implemented real-time capacity planning and monitoring capabilities across the enterprise, predictive analysis of the data collected could facilitate shifting workloads between facilities on an intermittent basis. The ability to dynamically shift loads between facilities with confidence not only defers capital expenditures in upgrades or new facilities but also can open the door to other efficiency initiatives (e.g., onsite renewable power generation, use of "follow the moon" strategies, or participation in your utility's demand-response program). It is also a good defense against transient threats to availability (e.g., an impending storm that may affect power supplies—an increasing risk in this era of climate change).

The longitudinal data gathered through monitoring can be analyzed to identify trends as well as predict issues: if CRAC/CRAH units have a characteristic pattern of declining performance before they fail, the data center operator knows exactly when to replace a CRAC/CRAH unit that's following the same pattern. Applications may have a particular cooling and power use signature under one set of circumstances and another when

14 Big Data. Bigger Security Risks: How Data Centers Can Track, Manage, and Secure Data with Dedicated Asset Tracking Networks. RF Code White Paper, 2013. http://resources.rfcode.com/whitepaper-big-data.-bigger-security-risks-how-data-centers- can-track-manage-and-secure-data-with-dedicated-asset-tracking-networks

conditions change. Predictive analytics allows operators to understand the relationships among environmental conditions, work patterns and compute activity over time, allowing them to derive rules and develop models that assist in decision-making.

Another benefit of detailed, longitudinal data collection and analysis is that it enables more accurate measurement of the true cost of providing a service. If a business can identify precisely what application is running on a given server at a given time, track the power and cooling used to support that compute load, and incorporate data on the other resources required (e.g., staff, bandwidth, additional equipment), it can calculate precisely what to charge a specific business unit or client for that service.

This ability is nothing short of transformative in terms of business planning. Organizations can prioritize workloads and allocate costs based on demand and availability. They can develop industry-specific productivity metrics that give users visibility into the efficiency of their infrastructure and allow them to fully understand and exploit the capacity of the data center to further business growth.

Summary

The data center is the keystone of modern business. When it functions as intended, it sustains and strengthens the business. When it fails, the entire business is compromised.

In our always-on world, customers expect a data center to be continuously available and companies invest millions of dollars each year to meet those expectations. Unfortunately, inefficiencies mean that data center availability often comes at an unnecessarily high price, both in terms of capital and operational costs and environmental impact. To succeed in today's competitive business climate, data center operators must implement techniques and technologies that enable them to maintain continuous availability while optimizing efficiency. The secret to achieving both goals is real-time monitoring and management of the data center environment.

Monitoring technologies give the data center operator visibility into the environmental conditions throughout the facility and allow them to identify and address an issue before it becomes a problem. Real-time asset management systems ensure the operator has comprehensive information on every device in the data center, simplifying inventory management and facilitating capacity planning. Predictive analysis of the data collected enables more integrated, autonomous operation of the data center and informs decision making throughout the organization. A company armed with this combination of information and technology is able to fully understand and exploit the capacity of its data center to further business growth.

About RF Code

RF Code is the leading provider of critical asset tracking and protection solutions. RF Code's solution—consisting of sensors, dedicated infrastructure and powerful CenterScape software platform—replaces expensive and inaccurate manual processes with real-time intelligence, enabling organizations to reduce costs and reduce risk of downtime by improving asset utilization, increasing operational efficiency, and enhancing risk mitigation.

Our automated, real-time asset management and environmental monitoring software platform eliminates the need for costly and error-prone manual processes. With our patented, wire-free sensors, open APIs, and real-time reporting capabilities, RF Code is easily integrated with existing IT, facilities, and business systems, creating mission critical value throughout the asset lifecycle, so you can save time and money and keep track of what really matters.

RF Code solutions track more than three million sensors worldwide, and have delivered over $100 million in savings to hundreds of customers, including more than half of the internet and telecommunications companies in the Fortune 500. RF Code is based in Austin, Texas, with partners around the world. Additional information can be found at www.rfcode.com.