Are Your Critical Facilities Ready To Be Managed With Big Data?

A Technical Paper prepared for SCTE by

Dustin Boyette
Product Manager, Intelligent Distribution
EnerSys
3767 Alpha Way, Bellingham, WA, USA
+1 360 603 0672
dustin.boyette@enersys.com

Table of Contents

List of Figures

1. Introduction

Today, critical communications facilities are largely managed manually. While some facilities (headends, primaries, data centers) may have fairly extensive monitoring, the data provided still tends to end up on a 2D "dashboard" for humans to interpret and act upon as necessary. Some facilities–especially remote hubs–may have very little telemetry at all.

This paper introduces big data then presents it as key to enabling advanced management of critical facilities (CF) in the future. The importance of collecting broad and deep data–starting today–is discussed along with extant and future sources of this data from facility infrastructure. Common challenges of contemporary facility management will also be discussed in the context of how big data can help reduce the human burden posed by those challenges. Several example use cases enabled by big data–when combined with modeling, visualization, and artificial intelligence–will be presented.

2. What is Big Data?

As is typical for contemporary jargon, a concrete definition for big data is elusive; you'll find slightly differing definitions depending on what organization is using the term. In the simplest sense, big data refers to data sets too complex and/or large for traditional data processing methods to take advantage of.

An analogy might be the spiral-bound paper notebooks of every student in a school district (replete with chemistry lab notes, doodles, and phone numbers of the student at the adjacent desk). In contrast would be the perfectly alphabetized contacts list in your smartphone.

2.1. The Three *V*s

While there are many characteristics of big data, it is chiefly defined by "The Three Vs":

Volume - Lots of unstructured data

Velocity - Data delivered rapidly, real-time or near real-time with little or no pre-processing

Variety - Many kinds of data available that don't readily fit traditional, ordered database schemes

More characteristics (also conveniently starting with the letter V) have augmented the definition over time, but for the purposes of this paper the original three will suffice.

Additionally, the term "broad" can be synonymous with Variety here. The term "deep" can be synonymous with Volume over time; that is, historical data.

2.2. A Dollar a Month

The primary driver enabling big data is the ever-decreasing cost of mass storage. In the early days of computers, five megabytes of storage occupied the space of a washing machine and cost perhaps quarter of a million dollars. The idea of storing even a byte of data with unknown importance–data which may never be accessed again– was unfathomable.

Today, however, vast amounts of both mass storage and compute resources are available at trivial cost relative to the potential benefit. For example, one terabyte of infrequent-access storage through a commercial cloud storage service such as Amazon S3 is a little over ten dollars a month. And at the lowest-cost "Deep Archive" tier, a terabyte is less than one dollar a month and becoming cheaper every year.

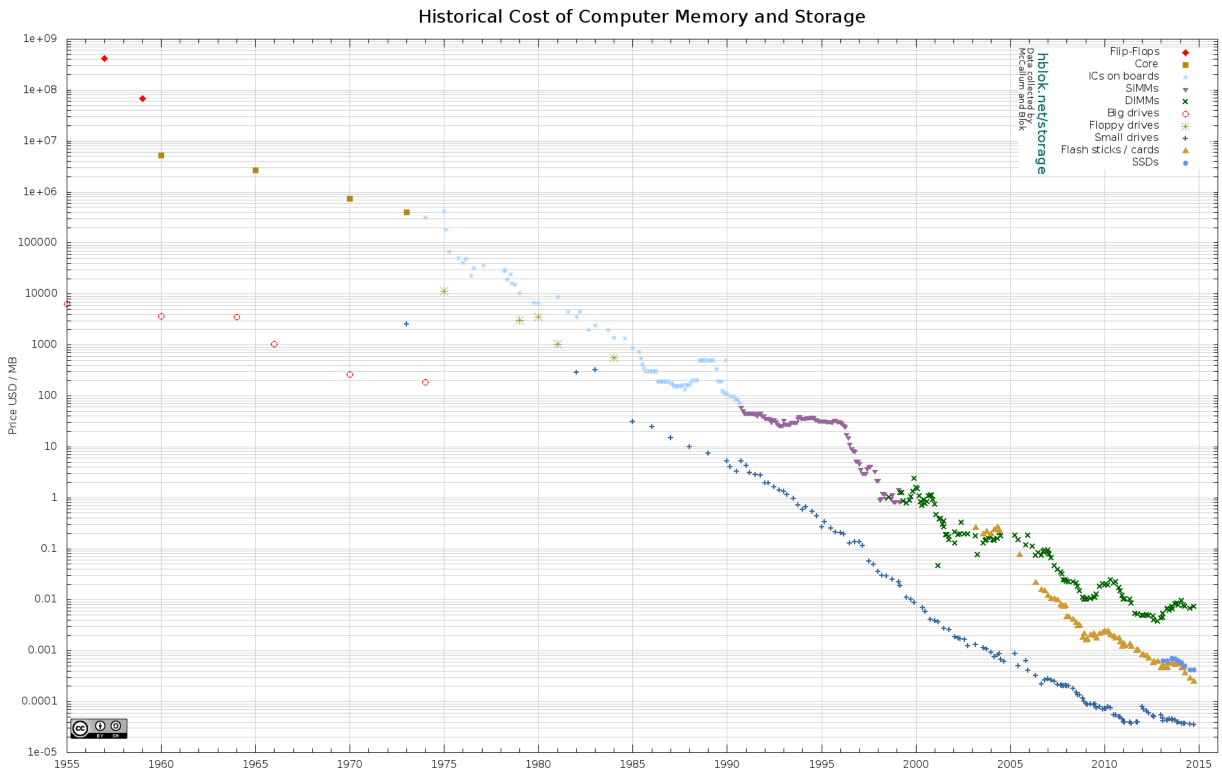

Figure 1 - Historical Cost of Data Storage

2.3. Of Warehouses and Lakes

In order to take advantage of any data sets, there first has to be place to store the data. The storage method must be reliable, extensible, and allow for quick access to subsets of data of interest.

An evolution of traditional data processing concepts, a data warehouse stores well-structured data. The schema for the data is predefined and data is transformed to fit. A customer contact database is a simple example of structured data: first name, last name, city, country, etc.

Moving this into the Critical Facility (CF) management area with which we are concerned, consider a database of facility equipment. Expected fields would be manufacturer, model, serial number, install date, rack location, and so on.

But what if data is available that doesn't fit the structure?

```
Power Input Module B Inlet Air Temperature 56°C at 03:02:01 on 2022-07-03
```

Where does that go?

It goes into a data lake, a vastly scalable centralized repository for raw data of virtually any type be it structured or not.

A data lake takes a different approach to data storage that is better suited to massive amounts of diverse data available today and in the future. The idea behind a data lake is to store data in its elemental, uninterpreted state. The data is minimally processed before storage. A value (the data type is less important than with structured data–integers, floating point, strings, even image data are welcome in the pool) along with some context–a source identifier and a timestamp–are all that are needed.

3. Data Sources in Critical Facilities

Even in a critical facility built without an aim towards monitoring/telemetry, there are many sources available for useful data. We are primarily concerned with data from the facility infrastructure itself and will expand on that below. However, it’s important to note that to fully take advantage of big data even the management channel data from the revenue-generating equipment itself is important to gather.

3.1. Power

Virtually every piece of power equipment you will find in a communications facility–from an Automatic Transfer Switch (ATS) to a Surge Protective Device to a DC Plant–is available today with a means to provide data in real time or near real time to a monitoring system via a network.

Early network-enabled power systems provided only coarse data–bus voltage, total current, alarm status, and other data at that level. Today, power systems are available with the ability to measure consumption down to the smallest branch circuit.

Figure 2 - “Smart” DC Fuse Panel With Individual Branch Circuit Current Monitoring

Stationary batteries, a key component of high-reliability power systems, have traditionally offered little in the way of data. Monitoring systems are often seen in facilities, but they typically only offer jar-to-jar voltages and perhaps a few temperature measurement points. However, the intelligence mandated by lithium batteries in consumer and automotive applications is making its way into reserve power batteries. Imagine a large critical facility with dozens of strings of backup batteries, each cell able to provide detailed metrics on its status. There could easily be ten thousand pieces of useful data gathered from a facility’s batteries per minute.

3.2. Environmental

Many facilities have environmental monitoring capability either as separate systems or as part of the active Heating/Ventilation/Cooling (HVAC) systems. Computer Room Air Conditioners (CRAC) have been available with network connectivity for some time.

Again, like many power systems, the data from these HVAC systems is usually basic and local to the unit itself. Perhaps there would only be a dozen temperature readings directly from the HVAC systems in a facility. However, the number of temperature and airflow sensors present on all the equipment in a facility is likely in the hundreds if not thousands. The low cost of digital temperature sensors means that granular temperature readings should be available for intelligent thermal analysis of an equipment rack, row, or zone.

In addition, new low-cost infrared (IR) sensors could soon enable equipment to detect local hot spots without the need for quarterly site visit by a technician with an expensive IR camera (the data from which is quickly stale).

3.3. Security/Access

Today's advanced security systems provide data far beyond simple contact closures triggered by magnets on doors. Key card and Radio Frequency Identification (RFID) integration allows the tracking of not only access by personnel, but the presence of tools and even equipment assets. Visual and IR cameras add another.

Looking beyond the dedicated security and access control systems, however, many pieces of equipment have sensors that can detect when a physical change has been made: module insertion/removal, access door open, rack door open, etc.

4. Turning Data Into Information

In traditional CF telemetry, there is priority given to quickly turn elemental data into important (or seemingly important) information. Formulae and scripts are written to take near-term data and either present it to an operator as information, or cause some immediate action: put a rectifier module to sleep, open a low-voltage disconnect, turn on an economizer (or trigger the fire alarm!) as examples. All important things at the time, to be sure.

However, often the raw data is discarded after being processed into immediate information, and is no longer available for analysis months or even years later.

4.1. First, You Need the Data

Obviously to take advantage of big data you first need data. What may not be obvious is the importance of rich historical data, and the important of capturing that data as soon as possible.

4.1.1. The Internet of Things (IoT)

The Internet of Things, or more commonly just IoT, is a popular buzzword covering all manner of intelligent, network-connected widgets that will magically and seamlessly talk with each other and work together. Any of us who've wasted an afternoon trying to get a Hi-Fi music streamer to stay connected to wireless speakers can attest, it's not quite there yet.

However, the idea of every *thing* being connected is fascinating: from the expected like discrete current transducers and humidity sensors, to the mundane such as light switches and fixtures. If a light fixture happens to also know the temperature, why should it not report that data to the building network?

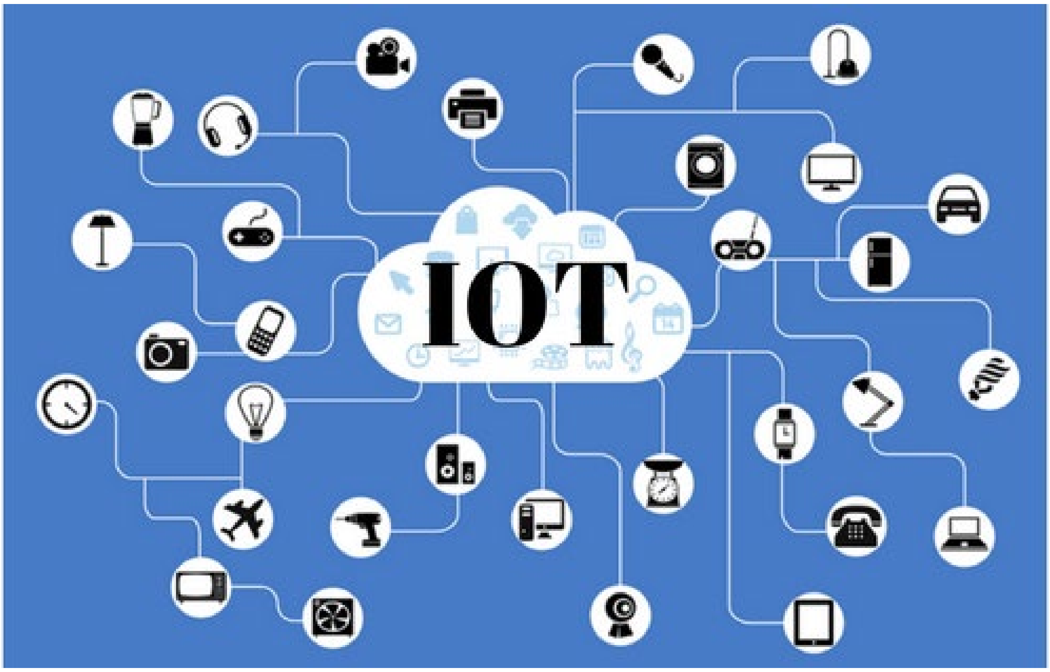

Figure 3 - IoT Example Graphic

4.1.2. "Smart" Infrastructure

The decreasing cost of sensors, microcontrollers, and ubiquitous network connectivity has made it possible for almost every piece of building infrastructure to be some degree of "smart." Today, even doorbells have microcontrollers and wireless network interfaces!

In critical facilities, however, most of the infrastructure today is still fairly dumb. That is, it is not able to provide data via a network connection. Or if the equipment is able, it is not connected, or nothing is done with the data. Indeed, some operators do not see the benefit of smart infrastructure, but they are making that choice from the perspective of the tools available to process the data today.

To take advantage of big data in the future, smart infrastructure is necessary today.

4.1.3. Retrofitting Telemetry

Operators large and small are faced with a chicken-and-egg question when it comes to smart infrastructure: "Why invest in network-enabled fuse panels and air conditioners when the majority of my existing equipment lacks the capability to tie in?"

Some may prefer the "sidewalk" solution: municipalities often require sidewalks be built whenever new construction is undertaken even if the surrounding neighborhood lacks sidewalks. While this results in a few "where the sidewalk ends" jokes on shorter timescales, longer term it results in neighborhoods with nearly contiguous walkways.

Unlike the relatively static central offices of landline heyday, today's critical facilities change, evolve, and grow quickly. When old HVAC systems and fuse panels won't meet the needs of the latest revenue-generating network gear, the new systems should be smart.

For installation conditions that don't readily permit the migration of old infrastructure to smart infrastructure, intelligence can often be retrofitted. Temperature sensors and clamp-on current sensors are readily available to augment smart infrastructure and instrument critical facilities.

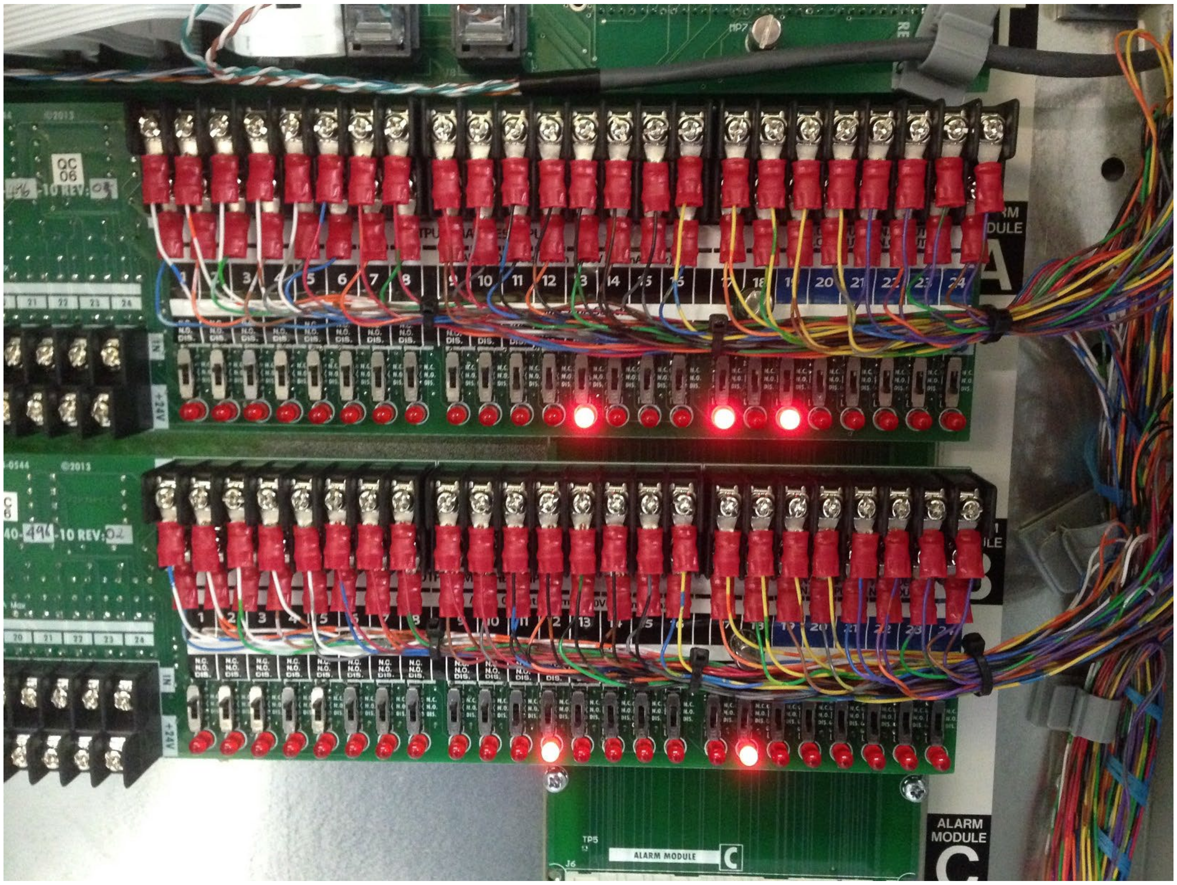

Figure 4 - Traditional Dry Contact Telemetry

4.2. Hands Off! Reducing the Human Burden of Data Management & Analysis

The aim of big data in most industries is to *reduce* the amount of manpower required to evaluate, plan, and act to accomplish a business objective. Unfortunately, as more data is made available, operators often find it more difficult to digest and act upon.

4.2.1. Modeling

The first step in making big data useful for managing critical facilities is to move ahead from the "Mission Control" dashboard. Facilities can be surveyed and physically modeled in 3D and presented virtually. Then, with sufficient data sets, the electrical and thermal behaviors of the facility can be modeled as well.

Instead of sorting through out-of-date photos, looking at rack inventory spreadsheets, or sending crews to survey sites, it's conceivable to view a current *model* of the facility. Operators will be able to virtually walk down a rack row and see what equipment is installed, where floor space is available, and even where fiber trough and cable tray is ready for growth.

4.2.2. Data Visualization

Layered on top of modeling, big data allows the ability to overlay important data directly over the virtual facility. Today, most of us use similar capabilities in everyday life without realizing it: the satellite view overlay of a map application, per-zone temperature overlay on a smart home, and so forth.

Now imagine being able to view the inlet air temperature gradient across the entire front of a critical equipment rack. Or go back in time and see what that same gradient looked like when an air conditioner failed last summer.

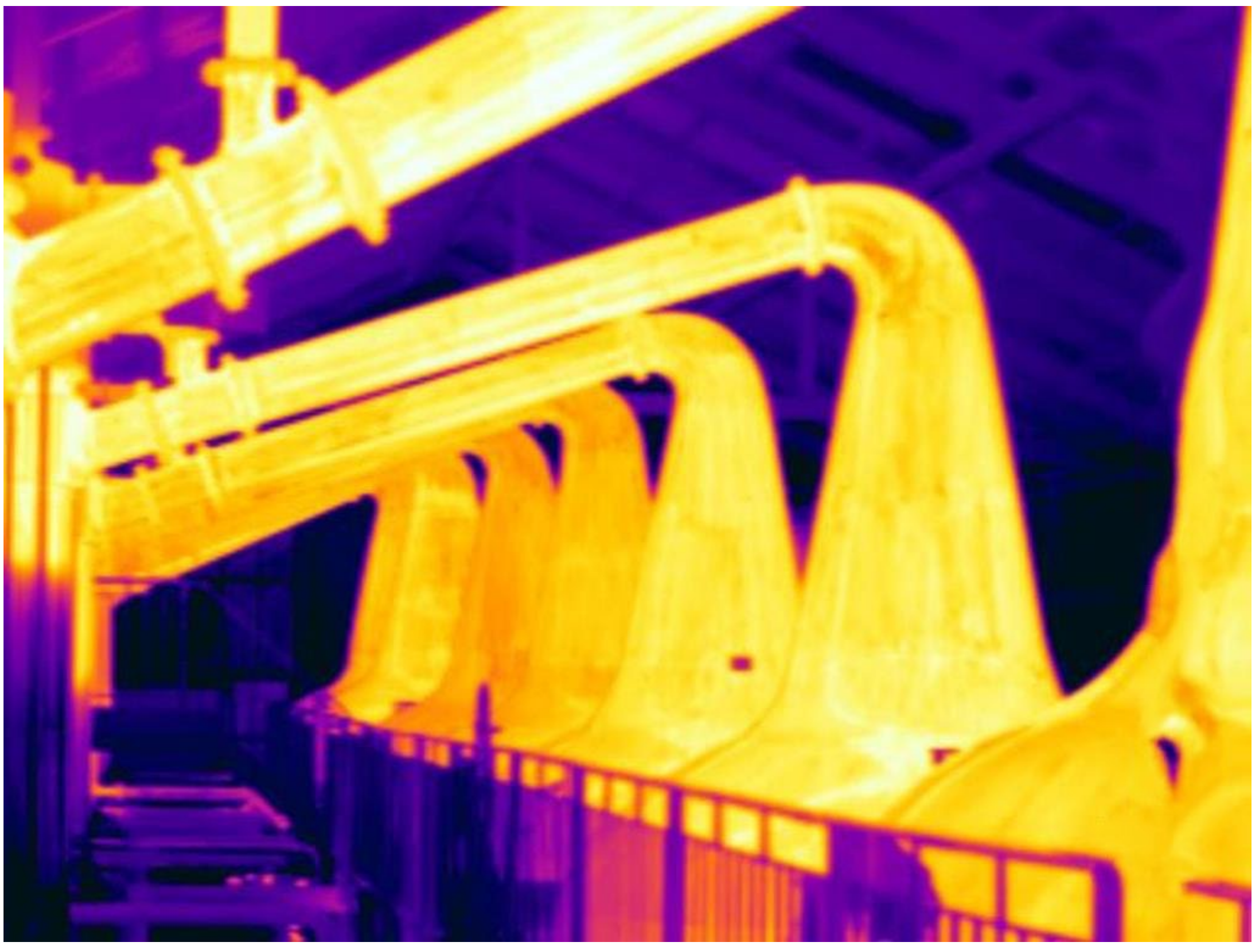

Figure 5 - Process Temperature Overlay in a Factory

4.2.3. AI

Despite being misunderstood and maligned in countless science fiction books and screenplays, artificial intelligence is at work today helping businesses and improving the lives of people.

What AI can already do today is impressive–from creating striking images prompted by just a few words (DALL-E) to finding the most efficient paths for rovers to explore other planets. But far more exciting is what AI will be able to do five and ten years from now.

In the CF space, it is readily conceivable that AI will allow us to "see the future". In our example above, instead of viewing the rack inlet temperatures from last summer, what if we could view the rack inlet temperatures for next summer? After adding another 10kW of revenue-generating equipment in the row! AI makes that possible and can do so better than the current "snapshot" predicting through the use of massive amounts of historical data.

5. Use Cases in Communications Facilities

With a sufficiently broad and deep set of data, what can big data do for critical facilities?

5.1. Outage Prevention

Operators spend significant time and money to ensure their networks are robust, resilient, and reliable. Redundant generators, dual DC plants, and diverse wiring are all aimed at preventing customer-affecting outages. Yet outages still occur all too frequently.

In many post-mortem investigations, sifting through logs reveals that data was available that indicated an outage was possible if not imminent. Often this data was not given enough importance to make it into any alarm logic or status dashboard. Ultimately, it's often a human who would have to react to the condition to prevent the outage. Or, if some automatic action is taken it's not granular enough and results in mitigation instead of prevention, which still results in something less than normal delivery of service.

Instead, why not allow computers to search for patterns in both current and historical data, run myriad "what if?" scenarios against the facility model, and identify where trouble may be brewing. Days, weeks, or months beforehand.

Figure 6 - Outage Postmortem

5.2. Operational Efficiency

Most operators would likely admit that their critical facilities do not operate at peak possible efficiency. Priority is understandably given to ensuring equipment is working, service is reliable, and just keeping up with customer demand. Analysis of power conversion loss, HVAC efficiency, and such is performed as time allows—usually when a major facility upgrade is planned. Such analysis is usually done on "snapshot" or coarse-grained historical data, and while the result may improve efficiency somewhat, it cannot continually *optimize* efficiency.

Big Data and advanced tools operating upon it, combined with smart infrastructure, will enable facilities to self-optimize in near real-time. Consider load-shedding, power source switching (utility to photovoltaic solar (PV), for example), economizer activation, etc.—all driven by machine learning working off broad, deep data sets.

5.3. Capacity Planning

As the worldwide demand for high-speed data grows unabated, the need to expand critical communications facilities keeps step. In these buildings exist leading-edge routers and switches and optical equipment yet planning for future growth remains largely a manual process. Obsolete equipment inventories, stale usage data, and myriad spreadsheets are used to determine whether new equipment can be added to a facility to meet growing service needs. Consider just a fraction of the simple questions that must be answered in order to add a significant new piece of equipment:

Physical: *Is there sufficient rack space available?*

Power: *Is there sufficient ampacity available at the BDCBB? At the DC plant? The service entrance? Is there even a feeder fuse or breaker position available?*

Cooling: *Can the proposed zone support the additional thermal load?*

Big data could answer these immediate questions, and much bigger ones without an operator ever needing to open a spreadsheet:

When will the utility service no longer support the rate of growth at this facility?

What would be the first point of failure during a prolonged utility outage?

Would it be cost-effective to add alternate energy sources to this facility?

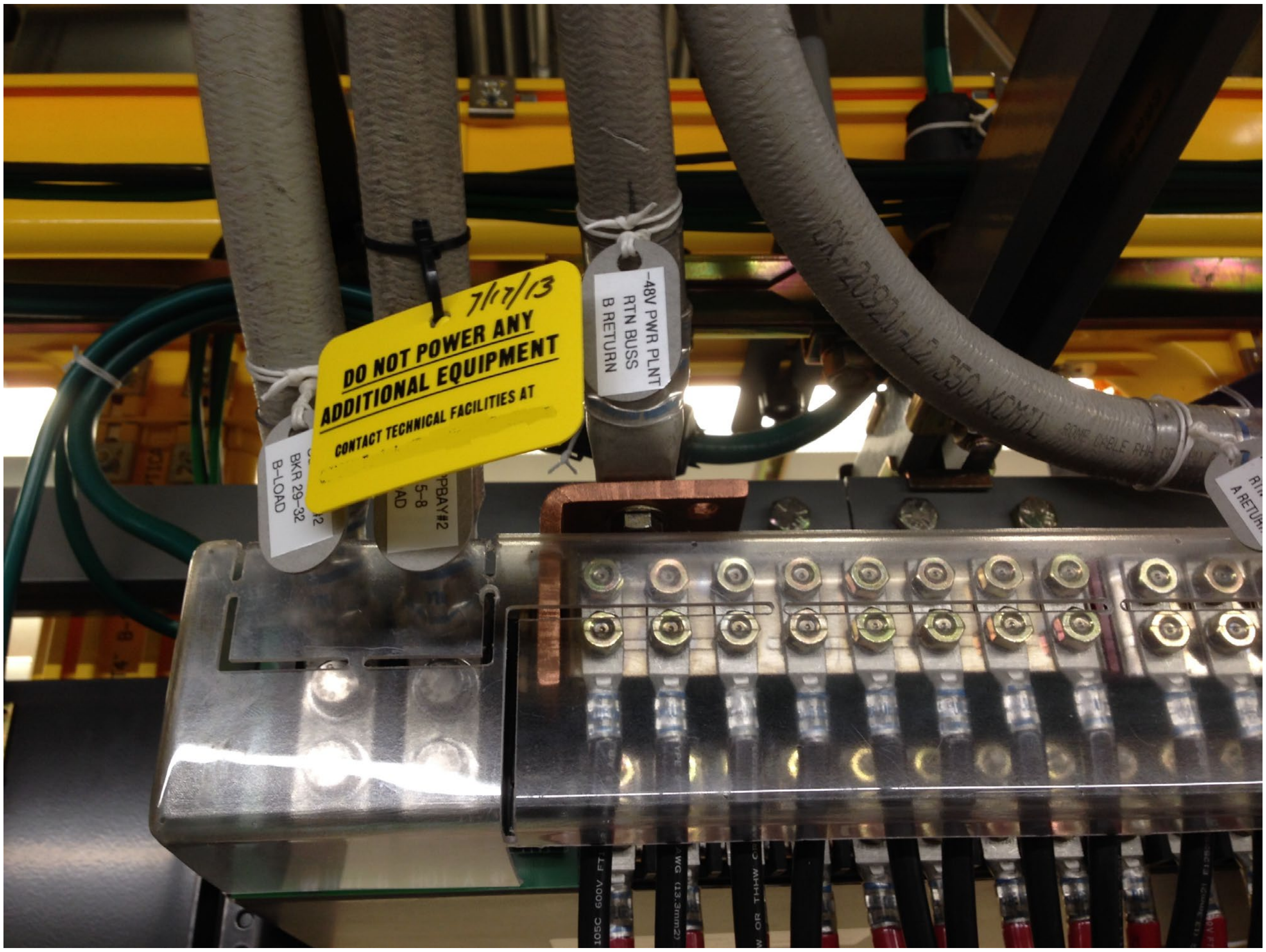

Figure 7 - Old-Fashioned Planning?

5.4. Automatic Provisioning

Taking capacity planning even further, a comprehensive CF management system could analyze customer trends, predict required service growth, then cascade the infrastructure requirements to match those needs. Imagine a system that places a purchase order for a new distribution panel along with all the properly-sized cables, circuit breakers, and mounting accessories. The order is placed taking into account delivery time of the various necessary components, and an installation technician can even be scheduled as soon as the equipment is slated to arrive.

6. Conclusion

The question posed by the title of this paper is rhetorical; your critical facilities are not yet ready to be managed by big data today, nor are the advanced modeling and AI applications yet developed enough in this field to do so. However, many facilities already have a surprisingly rich set of data available that could be valuable in future decision-making. Now is the time to collect the data already available and augment infrastructure where that data is missing. The future of critical facility management will be driven by big data.

Abbreviations

CF	critical facility
ATS	Automatic transfer switch
HVAC	Heating, ventilation, air conditioning
CRAC	Computer room air conditioner
IR	infrared
RFID	Radio frequency identification
BDCBB	Battery distribution circuit breaker bay
IoT	Internet of Things
PV	photovoltaic

Bibliography & References

Internet of Things (https://www.sas.com/en_us/insights/big-data/internet-of-things.html); SAS Institute

A Visualization is Worth a Thousand Tables: How IBM Business Analytics Lets Users See Big Data; Zif Davis 2014