WHITE PAPER

Global Data Center Engineering

Electrical Systems Load Testing: A Data Center Imperative

"With engineering, I view this year's failure as next year's opportunity to try it again. Failures are not something to be avoided. You want to have them happen as quickly as you can so you can make progress rapidly."
Gordon Moore

Contents

Version 1.4: 25 January 2015

Abstract

As the data center industry continues to evolve and adapt to a variety of demands imposed by governments, businesses, and technology advancements, ensuring that critical systems are capable of meeting required loads and maintaining long term operational sustainability are fundamental.

Data centers and other facilities with high availability power requirements each has unique designs. To ensure the integrity of the systems implemented will perform as expected for the life of the system, careful and specific testing process must be conducted. Initial commissioning testing and completion of an Integrated Systems Test (IST) is the first step in validating designs versus implementations.

In addition to IST, continuous application of reliability engineering during the life of the system is essential to reducing component failure rates, and maintaining operational sustainability.

Emphasis on achieving predictable facility and infrastructure resiliency has gained broader acceptance, and in some industries is required. The recognized global authority for classification of data center resiliency is The Uptime Institute's[1] Tier Classification standards. Design and implementation of Tier Rated data centers assures stakeholders that expected levels of availability are reached, and underpinning proof of the resiliency capability is stringent testing.

A further driver to electrical load testing occurs in tuning power systems to achieve maximum energy efficiency, especially under less than full circuit load. Efficiency tuning is a high priority driver in meeting government and self-imposed power reduction requirements. Occurrence of partial load commonality has increased, due to the proliferation of dual-corded equipment providing power redundancy to IT components.

Providing successful electrical and power generation infrastructure requires rigorous testing during discrete events, and also at fixed periodic intervals to ensure component failure rates are minimized. Further providing effective benchmarks for improvement initiatives and capacity planning to sustain the data center on an ongoing basis is aided by the use of load bank testing. Stress testing systems requires the selection of the right mix of load bank types to ensure simulation of as close to real load as possible is achieved.

The objective of this paper is to examine the events that trigger when electrical load testing should be performed, the drivers behind why load testing should be performed, and the results and benefits derived from identification of faults, improvements, and efficiencies affected through electrical systems load testing.

[1] Uptime Institute Tier Classification *Data Center Site Infrastructure Tier Standard: Topology* outlines the parameters of Tier Ratings in this standards document.

Introduction

Continuous and stable power is the key to providing successful infrastructure. With technology advancements placing steadily increasing demand on data centers critical systems infrastructure, ensuring uptime and availability of systems and services is imperative. The result of a failure within the systems can carry significant costs from loss of business generating revenue, damage to consumer confidence from loss of service, and to service providers from penalties derived from missed SLAs.

Cost of failure is realized in two forms. The first comes from the loss of revenue generated in missed sales to customers. (e.g. stock trades, banking transactions, consumer goods purchase, online auctions, etc.) In the second case, providers of data center facilities and services (Managed Service Provider or MSP) may be faced with stiff financial penalties resulting from contractual obligations that arise from failure to meet Service Level Agreements (SLA).

For many industries, such as investment banking, online auction houses, internet retail, or government sectors, the damage from outage can be measured in tens of thousands of dollars or more per minute[2].

Two Power Drivers: Capacity (Equipment Demand) and Availability (Uninterrupted Supply)

Power Capacity (Equipment Demand)

The demand for greater power capacity has continued to rise on a course similar to that of Moore's Law, which has predictable impacts on increasing energy consumption and generation of heat loads. The multicore processor was debuted by IBM in 2001 with the introduction of the Power4 CPU. By 2007, Intel had prototyped an 80 core CPU, and even by 2010 standards, 12 core[3] with lesser-known makers having up to 54 core[4] CPUs are commercially available. In addition, these CPUs are increasingly placed into servers which can accommodate multiple chips on a single motherboard, resulting in 100's of cores in a single server[5].

The overall outcome is greater numbers of processors placed onto chips, with greater numbers of chips placed onto a single server, results in higher density of power utilization placing more computing power in a single server at risk to failure. The impact on a single server by 2010 standards could potentially shut down thousands of

[2] On August 3, 2009 (cnet news article: http://news.cnet.com/8301-1023_3-10302072-93.html) PayPal experienced an outage of about 1 hour. Estimates at the time suggested that PayPal processes about $2,000 in payments per second, an impact exceeding $7.2mm for a one hour outage. (Accessed on 10 August 2010)
18 March 2010 (SearchDataCenter.com article: http://searchdatacenter.techtarget.com/news/article/0,289142,sid80_gci1484726_mem1,00.html) eBay published in their justification for building a $334mm data center, that an outage results in a loss of $2,000 to $4,000 per second of down time, depending on what elements are unavailable. (Accessed on 10 August 2010)

[3] AMD's 12 Core Magny-Cours Opteron 6172 series.

[4] Azul's 54 Core Vega 3 series.

[5] SuperMicro, Tyan, Dell all produce 4 or more socket motherboards, and Azul Vega 3 appliance can support up to 16 CPUs for a total of 864 cores.

virtualized servers and desktops. Compound that to the rack and room level, and the costs of failures become staggering.

Figure 1 illustrates the current state of transistor growth curve mapped against Moore's Law after 28 iterations. The impact of the curve over the life of Moore's law is at the verge of significant liftoff, where the exponential growth indicates a corresponding increase in the power required to run these systems. In the next 10 years (less than half the expected service life of a data center), this demand will push existing critical systems designs to the brink.

Comparison of Moore's Law Transistor Growth to Power Dissipation Increase (Watts[6])

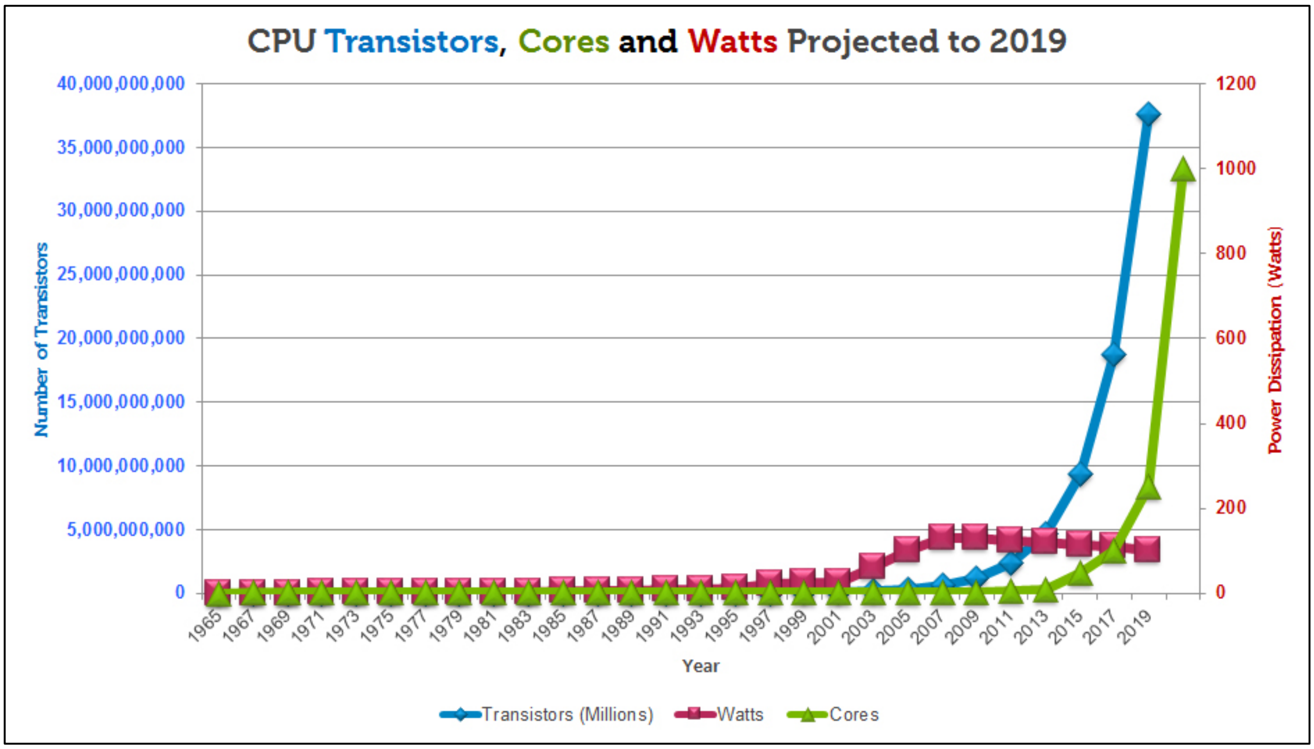

Figure 1 – Moore's Law Transistor Growth's Impact on Power Increase Curve

[6] CPU growth comparison with Intel CPU, from 1971 with 4004 through to 2010 i7/Xeon processors. Watt projections are based on trends in power dissipation, and take into consideration CPU clock speed, CPU voltage, number of cores, and are not fixed directly to Moore's Law as transistor increases.

Power Availability (Uninterrupted Supply)

In addition to the increasing power demands, reliability and resiliency of uninterrupted power supply are further drivers of data center power capacity. Availability of total data center "Up Time" is calculated in terms of the "five 9's". This term representing the number of 9's in the percentage of up time. Figure 2 illustrates the percentage of up time and expresses the amount of time spent in outage per year at each percentage of availability.

Percentage Availability	Outage Time Per Year
99%	**88 Hours**
99.9%	**8 Hours 46 Minutes**
99.99%	**52 Minutes 36 Seconds**
99.999%	**5 Minutes 16 Seconds**

Figure 2 – Outage Impact

It may be surprising to realize how much time is lost even at 99.999% availability. When comparing this to the potential cost of down time, damage to reputation suffered due to outage, or possible penalties suffered in missed SLAs, the need for assurance that systems will not fail under load becomes critical. Achieving these levels of availability is the focus of another driver, the Uptime Institute.

Uptime Institute: Tier Classification and Operational Sustainability

Tier Classification and Impact on Availability

Emphasis on availability is borne not only from lost revenue, but also as a result of focus on obtaining the highest probability of continuous operation. The Uptime Institute, established in 1993 is the recognized certifying body for Tier Ratings of data center facilities. While the rating of a facility is not mandatory, Tier Rating has become the benchmark against which data centers are assessed, and serves to define gaps between facilities and their probability of outage. There are two main components to the long-term viability of a data center to operate without outage. They are:

1. Level of redundancy for critical systems build (Tier Design Rating)
2. Level of implementation of operational sustainability best practices

Facilities are assigned a Tier rating from Tier I to Tier IV. Tiers I and II fall into a lower level of redundancy, and require at least some part of the facility to undergo a shut down even under normal circumstances (i.e. planned outages). These Tier levels are acutely sensitive to a broader outage if an unexpected failure in a critical component occurs as there is little or no redundancy. The cost of infrastructure build in a Tier I and II facility are about half that of Tier III and IV, as there are double or more the components needed in the higher Tier levels resulting from redundant pathways.

Figure 3 explains the various Tier Ratings as defined by The Uptime Institute and provides a brief description of the level of resiliency of each. By comparison to the "five 9's" diagram in Figure 2, the Tier Ratings are roughly equivalent to achieving the various levels of the "five 9's", though implementation variations may affect the level of availability actually achieved. It is relevant to note that the design of a data center does not always match the actual outcome, if modifications during construction or implementation have occurred.

Tier Rating[7]	Description
Tier I	**Basic Site Infrastructure**
Tier II	**Redundant Site Infrastructure Capacity Components**
Tier III	**Concurrently Maintainable Site Infrastructure**
Tier IV	**Fault Tolerant Site Infrastructure**

Figure 3 – Uptime Institute's Tier Rating Defined

The variances between data center Tiers I and II versus Tiers III and IV results in the need for different approaches to testing. As a result of the higher levels of redundancy required by Tier III and Tier IV facilities, including redundant power pathways, combined with an increased regularity of dual corded devices as a standard on most modern servers and network appliances, daily operational loads are generally in the range of 40% capacity.

Since a full load test is required for each circuit, and each redundancy point, there is an added emphasis to ensure optimal energy efficiency is realized when lower loads are placed on the circuits. In Tier III and Tier IV data centers, this additional consideration can have a significant impact on energy efficiency during daily operations under less than full loads.

Operational Sustainability

The Tier rating system is established to ensure that the design of data center infrastructure is in the best position to meet the demands placed on critical systems availability in accordance with the "five 9's". However, this only addresses initial design. The ongoing operational functioning of the data center environment, which is essential to ensuring the long-term viability of operational systems, must be addressed. To that end, The Uptime Institute has established a set of best practices for Operational Sustainability.

The Operational Sustainability best practices are comprised of four components:

- Staffing and Organization
- Maintenance
- Training and Planning
- Coordination and Management

The maintenance section stipulates specific Preventative Maintenance and component Life Cycle Planning, adding to the need for periodic testing. Load testing as part of preventive maintenance and component stress testing is indispensable to finding and remediating weakened components.

Reliability Engineering

[7] As defined in *Data Center Site Infrastructure Tier Standard: Topology* published by The Uptime Institute.

During the full life of the system, failure rate of components is not constant. The objective of reliability engineering is to reduce the rate of failure as close to constant as possible. Figure 4 illustrate the “Bathtub Curve”, and the result of reliability engineering brought about by load testing to achieve a lower, constant failure rate.

A reduced constant failure rate is achieved through electrical load testing throughout the life cycle of the system. Load testing, particularly stress testing, during the system life cycle will eliminate early failures when performed during commissioning, and will confirm weakening components as they reach end of life. Each of the test events involved in the component lifecycle will benefit from load testing performed periodically.

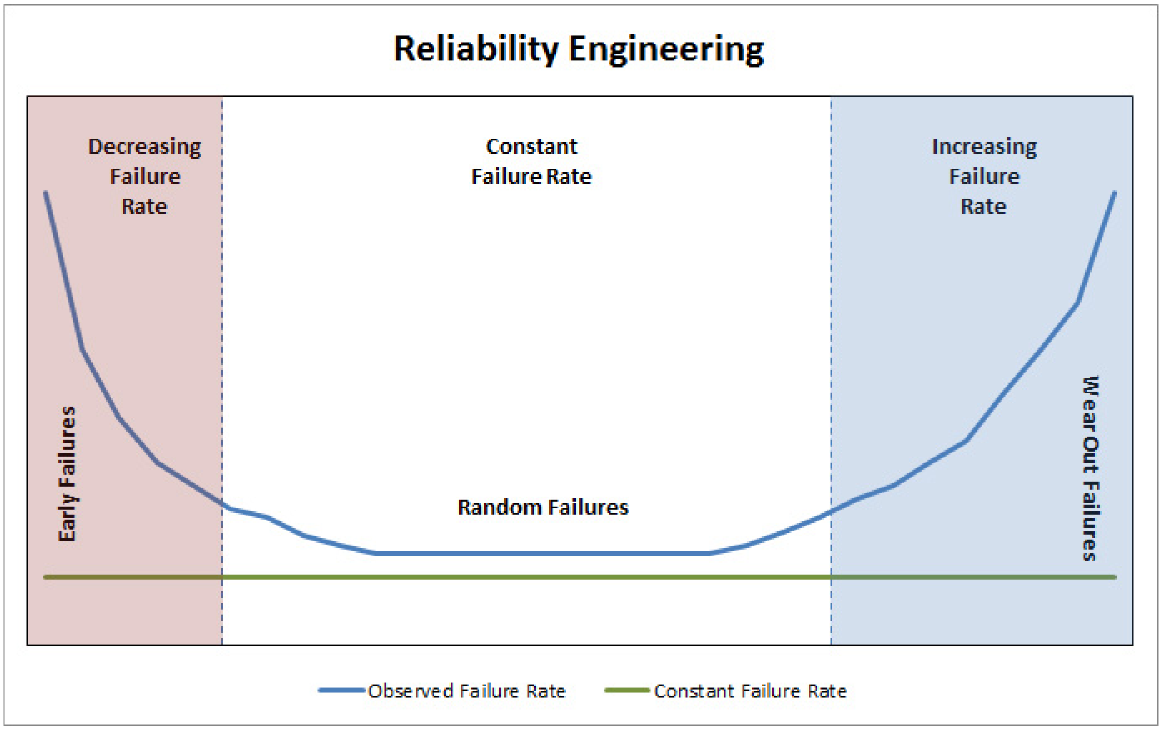

Figure 4 – Combating the Bathtub Curve

Testing Events

A number of events will generate the need for load testing. Each has a unique set of requirements, even when applied to the same facility. Load testing at defined intervals is the best strategy to counter the effects of “the bathtub curve”. The bathtub curve is a common term used to express the failure rate of equipment in early life cycle, useful life cycle, and failing life cycle. In the component life cycle, load testing plays a significant role.

New Facility (Commissioning)

The objective of load testing during commissioning (the “burn in” phase) will help to remove and replace components that fail early, before they are under production load conditions. Prior to bringing a new facility on line for service, rigorous testing must be completed. This requires diligent testing, beginning with individual components, progressing to dedicated and redundant pathways, and culminating with a full site Integration Systems Test (IST).

When new facilities are brought on line, real load will not likely be available, especially in the case of colocation facilities, where spaces are left to be customized during implementation stages based on client needs. In this instance, load banks are the only way to achieve a simulated load equivalent to what the systems will experience when the site is fully operating with demanding technology equipment loads.

Facility Expansion (New Building)

The expansion of a facility due to new build (additional build out of either previously unused floors, or expanded building space from new construction) will result in implementation of new equipment to accommodate new systems. A new series of commissioning testing will be required for new components and redundant pathways. An IST incorporating both previous and newly built systems will be required.

Facility Retrofitting (Refit/Refurbish)

A retrofit will result in some, or all of existing system components to be replaced. It may yield new implementations of improved resiliency (such as upgrades to develop the site from a Tier II to Tier III facility). It is important to ensure that new plans for testing to match the replacement fit-outs are created. Because a retrofit may utilize some existing equipment, all paths and components within an isolated path will require testing. An IST may also be required.

Component Replacement (Or Significant Repair)

Following the replacement or major repair of any critical component in the power pathway (including replacement of UPS battery sets), load testing should be conducted, and end-to-end systems test of affected components and pathways should be carried out.

Periodic Testing

The majority of testing discussed thus far has focused on ensuring that:

- New components that will fail early are stress tested and removed if found not to operate within tolerance before production loads are applied.
- Ensuring the full functionality of the designed system, including failover paths, operate to expected conditions.

During useful life cycle, periodic load testing enables measured load balancing and capacity management. As the component lifecycle progresses, load testing at end of life will assist in identifying weakened components, and enable their replacement under controlled maintenance.

Periodic testing will ensure continuous failure rate of components is minimized and remains predictable over the life of the system. Load testing will help determine where weak components exist, and the potential for failure. It will also provide critical information on component lifespan, versus manufacture specification.

Load Bank and Testing Types

Power systems with redundancy and resiliency incorporated in the design are complex. Testing all the components and pathways independently, as well as a collective system, requires different load types to ensure that accurate testing has been accomplished. To achieve this, there are two types of load banks utilized to complete comprehensive testing of data center critical power systems: Resistance load banks, and Resistive/Reactive load banks.

Resistance Load Bank

Many components in the power path such as transformers, breakers and power cabling will predominantly have a constant load applied, or at the very least a load that has little variation over shorter periods of time. The use of resistance load banks with these components is suitable and provides a solid basis for simulating loads for these elements.

Resistive/Reactive Load Bank

The occurrence of variability in applied loads is more frequent further up the power chain (approaching the power source). When load testing battery sets, UPSs and generator sets a more interactive load simulation is required. Reactive load banks achieve this by varying the overall load demand on generators and other upstream components.

Testing Types

Complete testing of generator reactiveness to variation of load requires utilization of reactive load banks. There are two types of reactive loads: Inductive and Capacitive. Inductive load banks will create a lagging power factor, while capacitive load banks will provide a leading power factor.

Capacitive loads are best for simulating telecommunication, PC and UPS loads while inductive loads are well suited to equipment with small motors, such as condensers, air conditioners, and fan units. When testing data center environments, both are important for simulating the expected downstream loads, and a combination of load bank types should be carefully planned for best testing results.

Load Testing Benefits

Though the drivers, events and demands of the system vary over the life of the facility, the core of the benefits come from the value that load bank testing provides in reliability engineering. The process creates both specific outcomes at each stage and contributes to the overall predictability of component failure. In addition to reliability engineering benefits, additional benefits are accomplished through testing at each stage of the lifecycle.

New Build

Load testing during commissioning of a new facility provides reassurance that the quality of the electrical systems design performs as expected, and integrates as necessary with other components. It establishes the basis for SLAs and provides proof that when the load is fully applied, the system will function as designed. Testing at this point establishes a baseline for future measurement to monitor the degradation of components over time, and assist in identifying when components require replacement.

System Expansions

After system expansion, load testing provides reassurance that any additional components or updated technology introduced will operate as specified. Changes to load capacity and demands placed on the system are validated. This stage also benefits from testing in the same way as a New Build.

Retrofit/Refurbish

New loads will be established, and as with System Expansion and New Build, establish reliability of newly implemented systems within the current structure.

Component Replacement or Major Repair

Any time a major component is replaced or major repair is carried out on a key component, new load testing will ensure that the system and components will operate within expected tolerances. It is also necessary to update the benchmark performance data, since the repair effectively results in a new component in the system. Establishing expected operational data for the component, and a baseline for that component for future comparison is required.

Periodic Testing

Conducting regular testing supports quality control, capacity management and identifies failing components prior to requiring replacement. When conducted annually, periodic testing establishes the performance data against previous measurements and benchmarks, allowing for comprehensive measurement of system performance. In addition, the data generated will provide empirical evidence of initiatives to reduce energy consumption, ongoing PUE measurement, actual CO2 emissions values, as well as overall systems degradation.

Conclusion

The continual development of technology-based systems, coupled with requirements for both increased power and reliability are driving the demand for power load and availability, in the data center. The rate of this growth has continually increased during the past two decades, and shows no sign of leveling or declining in the foreseeable future.

Coupled with high demand, high transaction systems that can result in the loss of millions of dollars caused by outages, and an industry focus on achieving high availability in data centers, the need to verify the capacity and continuous load carrying capability of a data center are incontrovertible.

Conducting electrical load tests during IST, commissioning and periodic testing under conditions that simulate actual load provides the basis for ensuring system availability will be achieved when it is needed. Utilized effectively, it will provide the rationale and evidence for predicting future power trends of 5, 10 or more years. It will determine the capability and capacity of redundant and fault tolerant power systems, while establishing reliable data for measuring energy improvement initiatives. The resulting uptime and availability of critical systems is directly translated into avoidance of lost availability, providing tangible prevention of financial loss due to outage or penalty from missed SLAs.

About the Authors

Global Data Center Engineering, having completed multiple data center projects on six continents ranging in industries from financial, telecommunications, consulting and technology industries are leading experts in data center standards, including TIA-942, Tier Standard: Topology and BICSI 002. GDCE has an extensive design, operation and commissioning experience in mission critical and critical systems infrastructure Mechanical and Electrical having built data center facilities in some of the most demanding client environments. GDCE team hold the highest certifications available in the data center industry.