USDA

United States Department of Agriculture

Office of the Chief Financial Officer

1400 Independence Avenue, SW

Washington, DC 20250

TO: Subcabinet Officials

FROM: David Combs
Chief Information Officer

JUN 21 2007

Charles R. Christopherson, Jr.
Chief Financial Officer

JUN 21 2007

SUBJECT: Development of USDA Data Center Strategy

In order to address several information technology (IT) vulnerabilities and weaknesses, we have undertaken a project to develop a comprehensive data center strategy and define and certify USDA enterprise data centers. This supports the various cyber security findings as a result of the internal control work conducted during fiscal year (FY) 2006 and other recent audit findings.

It is vital that we prepare a comprehensive data center strategy that can be implemented to move USDA in a well planned manner toward compliance with industry and government data center standards while achieving operational efficiency, security and agility. The reality is that our discretionary budgets continue to decline; our USDA policy requiring the control and protection of sensitive information is stricter; and our requirement to recover critical applications in a disaster is shorter. These changes and the noted physical access audit failures require that we review our current data center environment across the Department and change our operational framework.

Although the development of a comprehensive data center strategy takes time, it is imperative that the Department move forward now on this effort. The recent incidents related to breaches of privacy information dramatically highlight the need for compliance and clearly articulated IT standards. However, before we can begin to develop a data center strategy for the Department, we must first define and communicate what a data center is and provide a process for certification of our data centers.

There are many types of data centers in existence throughout USDA. These data centers support a variety of USDA IT. These include enterprise-wide information systems, major program systems, agency-specific systems, both mission critical and non-mission critical.

We have begun reviewing the "as-is" data center environment, analyzing our data center needs for the entire Department, and are now preparing a short- and long-term strategy that will allow the Department to move toward operational efficiency and agility. The strategy will address the placement of automated information systems within USDA centers, determining which should reside in designated enterprise data centers and which should reside in specialized/designated non-enterprise data centers.

Implementation of this data center strategy will be completed in phases. Those phases are represented in the milestones and the estimated timeframes as follows:

Milestone	Estimated Timeframe for Completion
Develop USDA standard definitions for enterprise class data center and minimum standards for these data centers	Completed
Develop enterprise class data center certification application	Completed
Issue Departmental Memorandum regarding enterprise data centers and certification of enterprise data centers	Not Later Than (NLT) June 20, 2007
Agencies apply to have data centers certified as enterprise class data centers	NLT July 6, 2007
Review of agency enterprise data center applications	NLT July 24, 2007
CIO provides certification of approved enterprise class data centers	NLT July 24, 2007
Develop USDA standard definition for system and classification of systems to be hosted at enterprise and non-enterprise data centers	NLT July 1, 2007
Review existing enterprise class data center space	NLT August 20, 2007
Review existing disaster recovery space	NLT August 20, 2007
Validate candidate systems/space requirements for enterprise data centers	NLT July 20, 2007
Identify candidate systems/infrastructure and space requirements for placement in data centers	NLT August 20, 2007
Perform gap analysis between candidate systems and existing data center space	NLT August 20, 2007
Develop criteria for non-enterprise data centers	NLT November 15, 2007
Issue policy for non-enterprise data centers	NLT December 15, 2007
Develop implementation plan/transition strategy/costs for placement of candidates in enterprise and non-enterprise data centers and DR sites	NLT December 31, 2007
Approve transition strategy and proposed timeline	NLT January 31, 2008
Implement the prioritized Phase 1 migration plan (placement of systems in enterprise data centers)	March 2008 – December 2008
Develop Phase 2 plan (placement of systems in non-enterprise data centers)	NLT September 2008

Attached to this memorandum are the USDA Enterprise Data Center Definition and the minimum standards required to be approved as an enterprise data center (Attachment 1). The required application to request certification of existing centers as enterprise data centers is also attached (Attachment 2).

Agencies and organizations with data centers meeting the minimum standards must apply for certification using the application checklist attached to this memorandum. All applications for certification as an enterprise data center must be submitted to Robert Suda, Associate Chief Information Officer by July 6, 2007. If you have questions or need clarification, please contact either Stacy Riggs, OCIO, on (202) 720-2225 or Wendy Snow, OCFO, on (202) 619-7636.

Attachments:
(1) USDA Enterprise Data Center Definition
(2) Enterprise Data Center Certification Application

cc: Agency Administrators
Deputy Administrators for Management
Agency Chief Information Officers
Agency Chief Financial Officers

Attachment 1 June 19, 2007

USDA Enterprise Class Data Center Definition

The Department's intent is to provide a standard definition for a USDA Enterprise Class Data Center.

USDA Enterprise Class Data Center Definition – A professionally managed and operated, institutionally supported facility, providing convenient access to, manipulation of, and/or distribution of data (including supporting information and expertise) for a wide community of users. It has a long-term charter (not tied to the lifetime of a specific project) and is capable of hosting systems that may be department-wide, shared services or agency-specific. The facility must meet USDA specified physical standards and sustain USDA specified operational standards.

All Enterprise Class Data Centers within USDA will be certified as meeting and sustaining the following minimum physical standards:

1. Electrical Systems: Professionally managed power distribution to include power conditioning, emergency power, and modular distribution.
    a. Facility electrical systems shall be rated at a minimum of the Tier III level as defined by the Uptime Institute (reference Appendix A). The facility electrical system Tier III requirement does not include a requirement for dual Utility Switching equipment or dual power feeds for non-mission critical hosted system components.
    b. Facility electrical power conditioners shall be employed to protect the data center electrical supply from fluctuations in power quality.
    c. When used, indoor generators shall be located in an isolated area as to protect the rest of the data center facility in the event of mechanical failure.

2. HVAC Systems: Professionally managed heating, ventilation, humidity control and air conditioning (HVAC) systems.
    a. Facility mechanical systems must be rated at a minimum of Tier III level as defined by the Uptime Institute (reference Appendix A).
    b. HVAC and other mechanical systems shall provide automated climate controls and automated system management to maximize system efficiencies
    c. The data center shall have raised flooring to provide a plenum for air to circulate below the floor.

3. Fire Detection and Suppression: Fire Detection and Suppression systems meeting local codes are required for the data center.
    a. Automated fire detection and suppression systems shall be installed in the data center.
    b. Fire suppression systems shall be designed to minimize damage to equipment outside of the fire zone.

4. Water and Flood Alarms: Enterprise data centers are to have water and flood prevention plans.
    a. Data Center facilities shall not have pipes containing water or other fluids above the data center floor space.
    b. Restrooms shall have functioning drains and secondary water barriers to prevent the flow of liquids into the data center operations.
    c. Data Centers shall have alarmed water monitors located under the raised flooring to detect water intrusion.

5. Telecommunications:
    a. Data centers shall have an on-site UTN node within the data center.
    b. Telecommunication infrastructure must meet the standards published in TIA-942 (Telecommunications Infrastructure Standard for Data Centers).

6. Physical Security:
    a. Certified as a Level IV Security Facility as set forth in the physical security classifications of the Department of Justice (reference Appendix B). The data center shall also at minimum meet requirements specified in DM3510-01, Physical Security Standards for Information Technology (IT) Restricted Space.

All Enterprise Class Data Centers within USDA will be certified as meeting and sustaining the following minimum operational standards:

1. Information Systems Security:
    b. Operations shall comply with NIST and FISMA standards.
    c. The data center shall have documented and approved hardening guides for server configuration.
    d. The data center shall perform monthly vulnerability scans on all systems and shall have a patch management system in-place to address vulnerabilities.
    e. The data center must have documented disaster recovery/business continuity plans and it must have documented emergency response processes.
    f. Personnel working in the data center shall hold fully adjudicated background investigations.
    g. The data center shall have an approved, up-to-date certification and accreditation of the general support systems that complies with the Government's mandatory requirements.
    h. The data center shall have a documented and exercised incident response process and procedures. The plan is to include electronic and manual hard shutdowns if the breach can not be controlled electronically.
    i. The data center shall have firewalls and network intrusion detection systems with active monitoring to prevent, detect, and manage electronic risks and attacks.
    j. Every two years, the data centers shall have comprehensive security analysis by an independent organization which specializes in system security.
    k. Data Centers shall have a designated security professional who is responsible for ensuring the maintenance of security procedures and compliance to the security requirements.

2. Facility Management:
    a. Data Centers shall maintain access logs, maintenance records, patch and upgrade records for the data center facility and hosted equipment.
    b. Topography of the facility infrastructure and hosted systems shall be maintained by the Data Center.
    c. Data Centers shall operate a cable management program, including the labeling and maintenance of cable diagrams of all network and electrical cables.
    d. Data Center shall have all fiber and cable tested every 5 years.
    e. Data Center shall have standardized and documented commissioning and decommissioning procedures for facility components and equipment hosted in the data center.
    f. Data Centers are required to quarterly confirm with the vendor the capacity of the lines subscribed.
    g. Data Center is required to monitor the capacity and continued operation of the telecommunications lines.
    h. Data Centers shall have a roof and structure inspection every five years.
    i. The data center shall have and maintain preventive maintenance and emergency services contracts on all data center infrastructure components.
    j. The data center shall maintain and operate documented, repeatable, standard procedures for scheduling maintenance
    k. Data centers shall not use temporary power, cooling, or control systems in a permanent manner. Temporary systems are to be used no longer than 60 days.
    l. Electrical, Fire Detection, HVAC, and environmental monitoring systems shall be actively monitored 24x7 for events by data center staff, and shall provide an alarm notification to data center personnel.
    m. Upgrades to Electrical systems and HVAC for data centers shall be designed to support modern energy conservation practices
    n. Electrical systems including UPS, generators and switch gear shall be professionally evaluated annually by an independent, certified technician.
    o. Electrical power shall be tested annually to ensure a clean power source for system hardware.
    p. Backup power capabilities shall be tested monthly.
    q. HVAC systems are to be professionally evaluated quarterly by an independent certified technician for efficiency and reliability.
    r. Fire suppression systems are to be inspected annually by an independent certified technician
    s. Data center facility sewers and drainage systems are to be inspected annually by an independent certified technician
    t. Data Center shall monitor the capacity of the telecommunications lines.
    u. Data Center shall monitor the operational status of telecommunication lines with tools that will notify data center personnel upon a communications line failure.

3. Data Center Staff: Full-time, dedicated management and technical support staff to perform operations and maintenance functions.
    a. The data center shall provide a 24x7x365 helpdesk
    b. The data center shall provide 24x7x365 monitoring of all control, alarm, physical security, information security and communication systems
    c. On call evening, night, and weekend staff shall be provided by data center to maintain patches and upgrades or repair to IT systems and facility infrastructure
    d. Facility and equipment maintenance schedules shall be available in advance to customers. Maintenance periods/schedules shall be negotiable.
    e. Personnel are to have the skills required to maintain the service level agreements and the security required for a data center.
    f. Data Center must maintain a staffing plan for disaster and pandemic response

4. Management Processes: The data center must have mature enterprise-wide processes to include configuration management, change management, project management, problem reporting and correction, capacity management, performance management, and vulnerability management.
    a. The data center shall have a program in place to implement and maintain the service delivery of the processes as defined by the IT Infrastructure Library (ITIL).
    b. Performance reporting (metrics) shall be in place to be used in Service level agreements (SLAs) with hosted systems.
    c. Customer service level surveys shall be delivered and reviewed annually by a level of management above the data center executive. Summary shall be provided to the senior level executive that has responsibility for the operation.
    d. Every three years that data center shall be benchmarked against industry for customer service, cost, and environmental impact.
    e. Management of the data center shall have at least one person responsible for operation and one person responsible for systems security on call at all times.

DRAFT

Appendix A
Uptime Institute Data Center Tier Ratings

A synopsis of the Uptime Institute data center tier ratings are as follows:

Tier I: A single path for power and cooling distribution, without redundant components, providing 99.671% availability.
A Tier I data center is susceptible to disruptions from both planned and unplanned activity. It has computer power distribution and cooling, but it may or may not have a raised floor, a UPS, or an engine generator. If it does have UPS or generators, they are single-module systems and have many single points of failure. The infrastructure should be completely shut down on an annual basis to perform preventive maintenance and repair work. Urgent situations may require more frequent shutdowns. Operation errors or spontaneous failures of site infrastructure components will cause a data center disruption.

Tier II: A single path for power and cooling distribution, with redundant components, providing 99.741% availability.
Tier II facilities with redundant components are slightly less susceptible to disruptions from both planned and unplanned activity than a Tier I data center. They have a raised floor, UPS, and engine generators, but their capacity design is "Need plus One" (N+1), which has a single-threaded distribution path throughout. Maintenance of the critical power path and other parts of the site infrastructure will require a processing shutdown.

Tier III: Multiple active power and cooling distribution paths but only one path active, redundant components, concurrently maintainable, providing 99.982% availability.
Tier III level capability allows for any planned site infrastructure activity without disrupting the computer hardware operation in any way. Sufficient capacity and distribution must be available to simultaneously carry the load on one path while performing maintenance or testing on the other path. Unplanned activities such as errors in operation or spontaneous failures of facility infrastructure components will still cause a data center disruption.

Tier IV: Multiple active power and cooling distribution paths, redundant components, fault-tolerant, providing 99.995% availability.
Tier IV provides site infrastructure capacity and capability to permit any planned activity without disruption to the critical load. Fault-tolerant functionality also provides the ability of the site infrastructure to sustain at least one worst-case unplanned failure or event with no critical load impact. This requires simultaneously active distribution paths, typically in a System+System configuration. Electrically, this means two separate UPS systems in which each system has N+1 redundancy. Tier IV requires all computer hardware to have dual power inputs as defined by The Uptime Institute's Fault Tolerant Power Compliance Specification Version 1.2.

Tier Classifications Define Site Infrastructure Performance

By W. Pitt Turner IV, P.E., John H. Seader, P.E., and Kenneth G. Brill

Widely accepted within the uninterruptible industry, The Uptime Institute's Tier Performance Standards are an objective basis for comparing the capabilities of a particular design topology against others or to compare groups of sites. This paper defines a four Tier system providing discussion and illustrations of *each classification. Significant cautions about Tier misapplication are provided. While the paper focuses* primarily on design topology, sustainability (how the site is operated once constructed) plays a more significant role in what site availability is actually achieved. Actual site performance figures combining both design topology and sustainability are presented by Tier classification.

This white paper:

- *Equips non-technical managers with a simple and* effective means for identifying different data center site infrastructure design topologies.
- Provides IT based definitions and performance requirements for each Tier Level.
- Provides actual 5-year availability for 16 major sites by Tier classification.
- Warns that site availability is a combination of both design topology and "sustainability" with considerable optimization "art" involved.
- Warns that component/system counts or MTBF analysis plays no role in determining Tier compliance partially because each fails to include sustainability factors which account for 70% of all infrastructure failures.
- Cautions "self proclaimed" Tier claims all too often turn out to be misleading, incomplete, or wrong.
- *Outlines need for third-party validation of site* selection, design, and sustainability decisions before committing to multi-million dollar projects.
- Provides a commentary on typical Tier attributes.

Background

One of the most common sources of confusion in the field of uninterruptible uptime is what constitutes a reliable data center. All too often, reliability is in the eye of the beholder—what is acceptable to one person or company is inadequate to the next. Competing companies with data centers of radically different infrastructure capabilities are all claiming to deliver high availability.

With the continuously increasing pressure on high availability comes an increased demand for computer hardware reliability. Information technology customers expect availability of "Five Nines" or 99.999%. Unfortunately, the substantial investment a business frequently makes to achieve Five Nines in its computer hardware and software platforms is likely to be insufficient unless matched with a complementary site infrastructure that can support their availability goals. The site infrastructure includes 16 power, cooling, and other critical physical layer environmental sub-systems that must work together as a tightly integrated uptime system.

Tier History

The Uptime Institute, Inc.® (*Institute*) developed a four tiered classification approach to site infrastructure functionality that addresses the need for a common benchmarking standard. The *Institute's* system has been in use since 1995 and has become the default standard for the uninterruptible uptime industry. An early-1990s Tier predecessor outlined seven ways of distributing critical power to the computer equipment, but was not simple and all inclusive. A broader standard was required.

Creation of the *Institute's* original Tier definition was stimulated by multiple industry requests. Senior management decision makers needed a simple and effective non-technical means of conveying the differences in data center investments. Since the original pioneering work done more than 10 years ago, the Tier concept has been further developed and validated by broad industry use. The *Institute's* objective performance-based standard is very useful in ensuring a consistent framework to compare various alternatives companies may consider for obtaining data center space. These include such options as owned, leased, third party providers, and so on.

Site Availability As Actually Experienced By Information Technology

The following tier commentary includes actual measured results for site availability ranging from 99.67% to more than 99.99%.

These figures are not predictive of future site results, but do reflect actual operating experience at a specific list of sites representing the four Tiers of functionality. It is important to note that this range of availability is substantially less than the current Information Technology (IT) expectations of Five Nines. This leads to the conclusion that site availability limits overall IT availability.

Four Tier Levels Reflect Evolution of Data Center Uptime Objectives

Over the last 40 years, data center infrastructure designs have evolved through at least four distinct stages, which are captured in the *Institute's* classification system. Historically, Tier I first appeared in the early 1960s, Tier II in the 1970s, Tier III in the late 1980s and early '90s, and Tier IV in 1994. The *Institute* participated in the development of Tier III concepts and pioneered in the creation of Tier IV. Tier IV electrical power distribution systems were made possible, in part, by Ken Brill, Executive Director of the *Institute*. In 1991, he envisioned a future when all computer hardware would come with dual power inputs. This became US Patent 6,150,736. United Parcel Service's 1994 Windward data center project was the first Tier IV design. During construction of the Windward project, United Parcel Service worked with IBM and other computer hardware manufacturers to provide dual-powered computer hardware[1].

Tier IV technology requires having at least two completely independent electrical systems. These dual systems supply power through diverse power paths to the computer equipment. This effectively moves the last point of electrical redundancy from the Uninterruptible Power Supply (UPS) system downstream to a point inside the computer hardware itself. Brill's intuitive conclusion has since been confirmed by *Institute* research that has determined that of the vast majority of site infrastructure electrical failures occur between the UPS and the computer equipment. Since completion of the Windward project in 1994, System plus System℠ (S+S) Tier IV electrical designs have become common and the number of computer devices with dual inputs has grown dramatically. There are exact parallels in the mechanical systems design.

The advent of dual-powered computer hardware in tandem with Tier IV electrical and mechanical infrastructure is an example of site infrastructure design and computer equipment design working together to achieve higher availability. Even with the significant improvements in computer hardware design made over the past 10 years, many data centers constructed in the last 5 years, and even today, claim Tier IV functionality, but actually deliver only Tier I, II, or III. This constrains their capability to match the availability required by the information technology they support. The purpose of this paper is to outline what it takes to consistently meet the requirements of the different tier levels.

The Need for Third-Party Certification Is a Growing Self-Preservation Requirement

In site infrastructure design and operation, the "devil is in the details" and the truth about a particular design topology will ultimately come out, but all too often after the warranty period has expired. When this happens, it can be a career ending event. Forensic investigation by the *Institute* into thousands of Abnormal Incidents over the last 12 years indicates that at least five and often seven interacting problems are required before a downtime failure occurs. The database upon which this analysis is built is in unique in the world.

Increasingly, senior executives desire to have their critical sites independently certified as being compliant to the Tier standards. This provides a validation that the technical details of what the designer designed and the contractor built is actually what the owner wanted. When project designers "self proclaim" a site meets a certain tier level or capacity, it is all too often inaccurate or only partly factual. The results can often be tragic involving unnecessary downtime and tens of millions in unforeseen upgrade expense.

Certification is a service performed by The Uptime Institute, who is uniquely qualified to interpret and apply the standards since the *Institute* created the underlying technology concepts that allowed the standards to develop in the first place. In addition, the *Institute* also brings awareness of emerging downtime problems and trends at least three to five years before they are commonly recognized and addressed by the rest of the industry.

Site Certification by The Uptime Institute involves two separate, interrelated activities. The first is verification of the design topology and how it complies with the

[1]There are 13 technical requirements to describe what is commonly called "dual power." The actual details and additional dual power information can be found in the *Institute's* white paper *Fault-Tolerant Power Certification Is Essential When Buying Products for High-Availability* which may be found at www.upsite.com/whitepapers.

Tier standards. The second phase is verification of site sustainability. While a particular topology design may meet the literal requirement of a Tier level, the lifecycle effectiveness of that design may be extremely limiting — typically less than five years. Sustainability includes site selection; lifecycle effectiveness of the design topology and its transparent flexibility/scalability; ease of use; staffing level and coverage, training, and skills development; management procedures and processes; metrics and dashboards; commissioning and maintenance practices; and the integration of the site infrastructure with the IT architecture. Human factors are important because 70% or more of all site failures involve people. Of these failures, 2/3 are management error and 1/3 is human error. Human sustainability factors will largely determine the actual level of site availability achieved.

Previous Tier Level Information Is Now Divided into "TIER PERFORMANCE STANDARDS" and "COMMENTARY" Sections

Responding to user questions and concerns, this white paper has been updated where appropriate and reorganized into two separate sections:

- The TIER PERFORMANCE STANDARDS are now in a totally separate section, similar to many engineering documents. The standards focus on the definitions of the Tiers and the performance confirmation tests for determining compliance to the definitions. These are 'absolute' criteria. Performance is measured by outcome confirmation tests and operational results. This is totally different than a prescriptive approach or a specific list of equipment not guaranteeing a performance outcome.
- The TIER COMMENTARY focuses on examples of the various ways to design and configure each Tier. In addition, the commentary section includes discussion and examples to aid in Tier understanding and information on common design topology failures. A comparison table of typical Tier attributes, availability and cost are provided. The commentary section also offers guidance in the comprehension, design, implementation, and the use of the Tier definitions.

Definition of Terms Used in the TIER PERFORMANCE STANDARDS and TIER COMMENTARY Sections

- Computer equipment: This is a broad phrase encompassing all information technology equipment required at a data center to perform the information processing work. It includes servers, storage, network, and all other information technology components.
- Redundant capacity components: The components beyond the number of capacity units required to support the computer equipment are referred to as redundant. If one unit of capacity is required to support the computer equipment, more than one unit of capacity is installed. Terms such as N+1 or N+2 are commonly applied.
- Useable capacity: This is the maximum amount of load that can be applied to the "N" level of capacity. Typically, the maximum amount of useable load is less than the non-redundant capacity to allow for component aging, installation errors, and to provide a contingency for unexpected demands.
- Useable capacity: This is the maximum amount of load that can be applied to the "N" level of capacity. Typically, the maximum amount of useable load is less than the non-redundant capacity to allow for component aging, installation errors, and to provide a contingency for unexpected demands.
- Site infrastructure: This comprises all of the site facility that includes the central plant plus the equipment that supports the power and cooling in the computer room. It is important to remember that a typical data center site is composed of at least 20 major mechanical, electrical, fire protection, security and other systems. Each has additional subsystems and components.
- Fault tolerant: This means that a system can sustain a worst case, unplanned event and not disrupt the end user. The fault tolerant concept originated in the IT environment. In the site infrastructure world, it means that the computer equipment will not be impacted by a facility failure. This requires multiple sources and multiple distribution paths so a failure on one source or path does not impact the other. This also requires use of computer equipment that meets the *Institute's* Fault Tolerant Compliant Power Specification. Computer equipment that does not meet that specification requires additional components, such as a point-of-use switch. During site infrastructure maintenance activity, the risk of disruption may be elevated.
- Concurrent maintainability: Originally, this was also an IT term. It means any work can be performed on a planned basis without impacting the end user. In the site infrastructure world, this means that ANY capacity component or distribution element can be repaired, replaced, serviced, tested, etc., without impacting the computer equipment.

TIER PERFORMANCE STANDARD

Tier I: Basic Site Infrastructure

The fundamental requirement

- A Tier I basic data center has non-redundant capacity components and single non-redundant path distribution paths serving the site's computer equipment.

The performance confirmation test(s)
- Any capacity component or distribution path failure will impact the computer systems.
- Planned work will require most or all of the systems to be shut down, impacting the computer systems.

The operational impact
- The site is susceptible to disruption from both planned and unplanned activities.
- The site infrastructure must be completely shut down on an annual basis to safely perform necessary preventive maintenance and repair work. Urgent situations may require more frequent shutdowns. Failure to perform this maintenance work increases the risk of unplanned disruption as well as the severity of the consequential failure.
- Operation errors or spontaneous failures of site infrastructure components will cause a data center disruption.

Tier II: Redundant Capacity Components Site Infrastructure

The fundamental requirement
- A Tier II data center has redundant capacity components and single non-redundant distribution paths serving the site's computer equipment.

The performance confirmation test(s)
- A capacity component failure may impact the computer equipment.
- A distribution path failure will cause the computer equipment to shut down.

The operational impact
- The site is susceptible to disruption from both planned activities and unplanned events.
- Redundant UPS modules and engine generators are required.
- The site infrastructure must be completely shut down on an annual basis to safely perform preventive maintenance and repair work. Urgent situations may require more frequent shutdowns. Failure to perform this maintenance work increases the risk of unplanned disruption as well as the severity of the consequential failure.
- Operation errors or spontaneous failures of site infrastructure components may cause a data center disruption.

Tier III: Concurrently Maintainable Site Infrastructure

The fundamental requirement
- A concurrently maintainable data center has redundant capacity components and multiple distribution paths serving the site's computer equipment. Generally, only one distribution path serves the computer equipment at any time.

The performance confirmation test
- Each and every capacity component and element of the distribution paths can be removed from service on a planned basis without causing any of the computer equipment to be shut down.

The operational impact
- The site is susceptible to disruption from unplanned activities.
- Planned site infrastructure maintenance can be performed by using the redundant capacity components and distribution paths to safely work on the remaining equipment.
- In order to establish concurrent maintainability of the critical power distribution system between the UPS and the computer equipment, Tier III sites require all computer hardware have dual power inputs as defined by the *Institute's* Fault Tolerant Power Compliance Specifications Version 2. This document can be found at http://www.upsite.com/TUIpages/tuifault_spec_2-0.html. Devices such as point-of-use switches must be incorporated for computer equipment that does not meet this specification.
- During maintenance activities, the risk of disruption may be elevated.
- Operation errors or spontaneous failures of site infrastructure components may cause a data center disruption.

Tier IV: Fault Tolerant Site Infrastructure

The fundamental requirement
- A fault tolerant data center has redundant capacity systems and multiple distribution paths simultaneously serving the site's computer equipment.
- All IT equipment is dual powered and installed properly to be compatible with the topology of the site's architecture.

The performance confirmation test(s)
- A single worst-case failure of any capacity system, capacity component or distribution element will not impact the computer equipment.
- Each and every capacity component and element of the distribution paths must be able to be removed from service on a planned basis without causing any of the computers to be shut down.
- In order to establish fault tolerance and concurrent maintainability of the critical power distribution system between the UPS and the computer equipment, Tier IV sites require all computer hardware have dual power

inputs as defined by the *Institute's* Fault Tolerant Power Compliance Specifications Version 2. This document can be found at http://www.upsite.com/TUIpages/tuifault_spec_2-0.html. Devices such as point-of-use switches must be incorporated for computer equipment that does not meet this specification.

- Complementary systems and distribution paths must be *physically separated (compartmentalized) to* prevent any single event from impacting both systems or paths simultaneously.

The operational impact

- *The site is not susceptible to disruption from a single* unplanned worst-case event.
- The site in not susceptible to disruption from any planned work activities.
- The site infrastructure maintenance can be performed by *using the redundant capacity components and distribution* paths to safely work on the remaining equipment.
- During maintenance activities, the risk of disruption may be elevated.
- Operation of the fire alarm, fire suppression, or the emergency power off (EPO) feature may cause a data center disruption.

Determining a Site's Tier Rating for Design Topology

Determining a site's actual Tier rating for design topology is not a complicated process, although it is one that is rarely done correctly. Figure 1 graphically illustrates the tier performance standards. For discussion of the standards, see the following commentary section.

Simply put, the Tier rating for an entire site is limited to the rating of the weakest subsystem that will impact *site operation. For example, a site with a robust Tier IV* UPS configuration combined with a Tier II chilled water system will yield a Tier II site rating.

This is driven by the need to manage perception in senior *management, as well as to factually report actual site* capabilities. If a site is advertised within an organization as being fault tolerant and concurrently maintainable (Tier IV), it is intolerable to shut the site down at any time in the future—regardless of what subsystem may *have required the shut down.*

There are no partial or fractional Tier ratings. The site's Tier rating is not the average of the ratings for the 16 critical site infrastructure subsystems. The site's tier rating *is the LOWEST of the individual subsystem ratings.*

Similarly, the "Tier" cannot be imputed by using calculated Mean Time Between Failure (MTBF) component statistical reliability to generate a predictive *availability and then using that number to "match"* the actual measured availability results shown later in Figure 2. Even if statistically valid component values existed (and they don't because product life cycles are getting shorter and shorter and no independent, industry-*wide database exists to collect failures), this approach* fails to include people which consistently are involved in 70% of all site failures. A calculated reliability of 0.9999 *which ignores human interaction does NOT define a site* as being Tier IV. The only way to determine Tier Level is to

Figure 1:
Performance Standards by Tier Level

Tier Requirement	Tier 1	Tier II	Tier III	Tier IV
Source	System	System	System	System + System
System Component Redundancy	N	N+1	N+1	Minimum of N+1
Distribution Paths	1	1	1 normal and 1 alternate	2 simultaneously active
Compartmentalization	No	No	No	Yes
Concurrently Maintainable	No	No	Yes	Yes
Fault Tolerance (single event)	No	No	No	Yes

objectively determine a site's ability to respond to planned and unplanned events.

TIER COMMENTARY

The *Institute's* STANDARDS Are Outcome Based

The requirements used in the *Institute's* Tier Performance Standard are necessarily and intentionally very broad to allow innovation in achieving the desired level of site infrastructure performance, or uptime. The individual Tiers represent categories of site infrastructure topology that address increasingly sophisticated operating concepts, leading to increased site infrastructure availability. The performance outcomes defining the four Tiers of site infrastructure are very straight forward. Recent initiatives by several groups to replace the *Institute's* Tier concepts with component counts and checklists has lost focus that ultimately counts is uptime performance. Most designs that will pass a checklist approach will absolutely fail a performance requirements approach. What this means is that there is still considerable "art" to the science of uptime and how sub-systems are integrated (or not integrated).

Tier Functionality Progression

Tier I solutions acknowledge the owner/operator's desire for dedicated site infrastructure to support IT systems. Tier I infrastructure provides an improved environment compared to an office setting and includes: a dedicated space for IT systems; a UPS to filter power spikes, sags and momentary outages; dedicated cooling equipment that won't get shut down at the end of normal office hours; and an engine generator to protect IT functions from extended power outages.

Tier II solutions include redundant critical power and cooling capacity components to provide an increased margin of safety against IT process disruptions from site infrastructure equipment failures. The redundant components are typically an extra UPS modules, cooling units, chillers, pumps, and engine generators. Loss of the capacity component may be due malfunction or to normal maintenance.

Owners who select Tier I and Tier II solutions to support current IT technology are typically seeking a solution to short-term requirements. Both Tier I and Tier II are *tactical* solutions, usually driven by first-cost and time-to-market more so than life cycle cost and uptime (or availability) requirements. Rigorous uptime requirements and long-term viability usually lead to the *strategic* solutions found in Tier III and Tier IV site infrastructure. Tier III and Tier IV site infrastructure solutions have an effective life beyond the current IT requirement. Strategic site infrastructure solutions *enable* the owner to make strategic business decisions concerning growth and technology, unconstrained by current site infrastructure topology.

Tier III site infrastructure adds the concept of concurrent maintenance to Tier I and Tier II solutions. Concurrent maintenance means that any component necessary to support the IT processing environment can be maintained without impact on the IT environment. The effect on the site infrastructure topology is that a redundant delivery path for power and cooling is added to the redundant critical components of Tier II. Maintenance allows the equipment and distribution paths to be returned to "like new" condition on a frequent and regular basis. Thus, the system will reliably and predictably perform as originally intended. Moreover, the ability to concurrently allow site infrastructure maintenance and IT operation requires that *any* and *every* system or component that supports IT operations must be able to be taken offline for scheduled maintenance without impact on the IT environment. This concept extends to important subsystems such as control systems for the mechanical plant, start systems for engine generators, EPO controls, power sources for cooling equipment and pumps, and others.

Tier IV site infrastructure builds on Tier III, adding the concept of fault tolerance to the site infrastructure topology. Just like concurrent maintenance concepts, fault tolerance extends to *any* and *every* system or component that supports IT operations. Tier IV considers that any one of these systems or components may fail or experience an unscheduled outage at any time. While the Tier IV definition is limited to consideration of a single system failure, Tier IV requires that the effect of such a failure is considered on other site infrastructure systems and components. For example, the loss of a single switchboard will affect the operation of all the equipment fed from that switchboard: UPS systems, computer room cooling equipment, controls, etc.

The progressive nature of functionality from Tier I through Tier II and Tier III to Tier IV is demonstrated in the schematic illustrations found at the end of this paper. The examples show the addition of components and distribution paths, as described above. Although the illustrations shown are not recommended design solutions for any particular set of requirements, the four electrical topologies are illustrative of the Tier classification concepts. Mechanical system functionally

progresses through the increasing Tiers similarly. Consistent, across-the-board application of Tier concepts for electrical, mechanical, automation and other subsystems is absolutely required for any site to satisfy the Tier standards.

Over the last few years, site infrastructure has been occasionally described by others in the industry in terms of fractional tiers (i.e. Tier 2.5), or incremental Tiers (Tier III +, or Enhanced Tier III, or Tier IV light). Fractional or incremental descriptions for site infrastructure are not appropriate. A site that has an extra UPS module, but needs all the installed computer room air handlers running to keep the UPS room temperature within limits does not meet *site* redundancy requirements for Tier II. A switchboard that cannot be shutdown without affecting more than the redundant number of secondary chilled water pumps is not concurrently maintainable (Tier III).

IT Availability Success Is Dependent upon Successful, Fully Integrated Operation of All Site Infrastructure Systems

The Tier classifications were created to consistently describe the site-level infrastructure required to sustain data center operations, not the characteristics of individual systems or sub-systems. Data centers are dependent upon the successful operation of over 16 separate site infrastructure subsystems. Every subsystem and system must be consistently deployed with the same site uptime objective to satisfy the distinctive Tier requirements. The most critical perspective owners and designers must consider in making tradeoffs is what impact the decision has on the integrated impact of the site infrastructure on the IT environment in the computer room.

The *Institute* has measured the actual availability, or performance, of 16 data centers having site infrastructure topologies meeting the four Tier definitions and has has established availability values representative of each classification. In practice, representative site availability, stated as a percentage of annual operating time, is associated with each of the *Institute's* standard Tier classifications. These empirically determined values include sustainability and human factors over a period of up to 10 years with uptime measured from the perspective of the IT client's operations in the computer room. This "real world" site availability is strikingly different than the probability of system failure that is often calculated using values from the Institute of Electrical and Electronics Engineers (IEEE) Gold Book for recommended practices for reliable power systems or guidelines from the IEEE Orange Book for emergency and standby power. A representative site infrastructure availability of 99.95% (about 4.4 hours of "downtime" per year) is not equivalent to a statistical reliability of 0.9995 (1 in 2,000 chance of a failure). Similarly, as outlined earlier, a calculated statistical reliability of 0.9995 does not indicate a site is "better than Tier III."

The *Institute* defines site availability from the perspective of a user of IT. Any site incident or event that affects information availability as experience by end users detracts from site infrastructure availability. The site downtime clock starts running from the moment IT operations were first affected until they are fully restored. Thus, site downtime is not the 15 seconds of a utility power failure, but the total time users were down until IT availability was restored. For Tier I and Tier II topologies, downtime for site infrastructure maintenance (which includes the time to bring IT systems down, perform the site maintenance, and restore IT availability) typically has a bigger availability impact than a UPS system failure. Based on operating experience of monitored sites, the typical maintenance outage at Tier I and Tier II sites is 12 hours. The time for IT to recover from a typical outage such as momentary power loss is 4 hours at sites of any tier.

Tier I sites typically experience two separate 12-hour, site-wide shutdowns per year for maintenance or repair work. In addition, on average, across multiple sites and over a number of years, Tier I sites experience 1.2 equipment or distribution failures each year. The annual impact of maintenance and unplanned outages is 28.8 hours per year, or 99.67% availability.

Operations experience shows that, on average, Tier II sites schedule three maintenance windows over a 2-year period and have one unplanned outage each year. The redundant components of Tier II topology provide some maintenance opportunity leading to just one site-wide shutdown each year, and reduce the number of equipment failures that affect the IT operations environment. The annual impact of maintenance and unplanned outages is 22 hours per year, or 99.75% availability.

Tier III topology is concurrently maintainable, so annual maintenance shutdowns are not required, which allows an aggressive maintenance program improving overall equipment performance. Experience in actual data centers show that operating better maintained systems reduces unplanned failures to a 4-hour event every 2.5 years, or 1.6 hours on an annual basis. Tier III sites demonstrate 99.98% availability.

Tier IV provides robust, fault tolerant site infrastructure, so that facility events affecting the raised floor are

empirically reduced to one 4-hour event in a 5-year operating period, or 0.8 hours on an annual basis. Individual equipment failures or distribution path interruptions may still occur, but the effects of the events are stopped short of the IT operations environment. Tier IV sites consistently demonstrate 99.99% availability.

The representative availability percentages are a characteristic of the operating experience of multiple sites within each Tier classification. A site with a measured infrastructure availability of 99.90%—midway between Tier II (99.75%) and Tier III (99.98%)—has an operating experience consistent with sites having Tier II topology, but does not achieve the availability of Tier III sites. Availability does not determine the Tier classification. Even more importantly, infrastructure with a statistical probability of failure of 0.9990 cannot be represented as a 'Tier 2.5' site, since the impact of the failure on overall availability is not represented by the *likelihood* of a system failure.

Independent of site infrastructure experience, IT organizations often describe data center availability objectives as Five Nines, or 99.999% of uptime. This is a very aggressive goal, especially if compared to the observed consequences of a single site outage. While the site outage is assumed to be promptly restored (which requires "24 by forever" staffing), it can still require up to 4 hours for IT to recover information availability and restore end user functionality, even if the likelihood of a data base corruption or a server power supply failure are set aside. In reality, facility failures often reveal previously unknown IT architecture, hardware, or software issues.

If a momentary power outage results in a 4-hour end-user disruption, how relevant is an objective of 99.999% availability? Based on a single site outage of 4 hours, it will take 45.6 years of 100% uptime to restore cumulative site availability back to the 99.999% objective. (4 hours x 60 minutes an hour ÷ 5.26 minutes per year = 45.6 years.)

Even a fault tolerant and concurrently maintainable Tier IV site will not satisfy an IT requirement of Five Nines (99.999%) uptime. The best a Tier IV site hope for 100% uptime for a string of multiple years. Figure 2 of Typical Tier Attributes uses 99.995% for representative Tier IV site availability, but this assumes a site outage occurs not more than once every 5 years. With a properly designed Tier IV configuration, the single event exposures that can result in a site failure are the results of a fire alarm or the unintended operation of the EPO feature. Only the top 10 percent of Tier IV sites will achieve this level of performance. Unless human activity issues are continually and rigorously addressed, at least one failure is likely over 5 years.

Typical Tier Attributes

Tier I sites have their roots in the mainframe environments of the 1970s. Tier IV became possible with the advent of dual-powered computers in the 1990s. Tier II and Tier III facilities were widespread in the 1980s; Tier III is the most common site infrastructure currently being implemented although most are designed for future transparent upgrade to Tier IV. Most owners find it fairly difficult to upgrade by more than one tier level from what they previously had. A responsible approach to site infrastructure investment is to understand clearly the availability objectives necessary to support the owner's current and future business requirements, then to consistently design, build, and operate the site to conform to those needs.

The following chart (Figure 2) depicts various attributes commonly associated with a particular Tier classification, but the attributes are not requirements of the Tier definitions. For example, the presence of a raised floor or any particular floor height are not criteria for any Tier. (The recommended height of raised floors, when used, is most directly correlated to power density.)

Integration of IT Architecture and Topology with Site Architecture and Topology Helps to Ensure Achieving Uptime Objectives

There are many opportunities within the Information Technology architecture to reduce or minimize the impacts of these unfortunate site infrastructure failures. These steps may include placing the redundant parts of the IT computing infrastructure in compartments served by different site infrastructure systems so that a single event cannot simultaneously affect all IT systems. Another alternative is focusing special effort on business-critical and mission-critical applications so they do not require 4 hours to restore. These operational issues can improve the availability offered by any data center and are particularly important in a "Four Nines" data center housing IT equipment that requires "Five Nines" availability.

The four Tier Standard classifications address topology, or configuration, of site infrastructure, rather than a prescriptive list of components, to achieve a desired operational outcome. For example, the same number of chillers and UPS modules can be arranged on single power and cooling distribution paths resulting in a Tier II (Redundant Components) solution, or on two distribution

Figure 2:
Typical Tier Attributes

	Tier I	Tier II	Tier III	Tier IV
Building Type	Tenant	Tenant	Stand-alone	Stand-alone
Staffing	None	1 Shift	1+Shifts	"24 by Forever"
Useable for Critical Load	100% N	100% N	90% N	90% N
Initial Build-out Gross Watts per Square Foot (W/ft²) (typical)	20-30	40-50	40-60	50-80
Ultimate Gross W/ft² (typical)	20-30	40-50	100-150[1,2,3]	150+[1,2]
Class A Uninterruptible Cooling	No	No	Maybe	Yes
Support Space to Raised Floor Ratio	20%	30%	80-90+%[2]	100+%
Raised Floor Height (typical)	12"	18"	30-36"[2]	30-36"[2]
Floor Loading lbs/ft² (typical)	85	100	150	150+
Utility Voltage (typical)	208, 480	208, 480	12-15 kV[2]	12-15 kV[2]
Single Points-of-Failure	Many + human error	Many + human error	Some + human error	None + fire and EPO
Annual Site Caused IT Downtime (actual field data)	28.8 hours	22.0 hours	1.6 hours	0.8 hours
Representative Site Availability	99.67%	99.75%	99.98%	99.99%
Typical Months to Implement	3	3-6	15-20	15-20
Year first deployed	1965	1970	1985	1995
Construction Cost (+ 30%)[1,2,3,4,5] Raised Floor Useable UPS Output	$220/ft² $10,000/kW	$220/ft² $11,000/kW	$220/ft² $20,000/kW	$220/ft² $22,000/kW

[1] 100 W/ft² maximum for air-cooling over large areas, water or alternate cooling methods greater than 100 W/ft² (added cost excluded).
[2] Greater W/ft² densities require greater support space (100% at 100 W/ft² and up to 2 or more times at greater densities), higher raised floor, and, if required over large areas, medium voltage service entrance.
[3] Excludes land; unique architectural requirements, permits and other fees; interest; and abnormal civil costs. These can be several million dollars. Assumes minimum of 15,000 ft² of raised floor, architecturally plain, one-story building, with power backbone sized to achieve ultimate capacity with installation of additional components or systems. Make adjustments for NYC, Chicago, and other high cost areas.
[4] Costs are based on 2005 data. Future year costs should be adjusted using ENR indexes.
[5] See *Institute* White Paper entitled Dollars per kW plus *Dollars per Square Foot Is a Better Data Center Cost Model than Dollars per Square Foot Alone* for additional information on this cost model.

paths that may result in a Tier III (Concurrently Maintainable) solution. Compare the Tier II and Tier III diagrams at the end of this paper. Both topologies contain the same N+1 capacity redundancy for engine generators and UPS modules, but the alternate distribution paths define the Tier III example.

Applying the Standards

The Tier Performance Standard provides objective criteria to consistently evaluate the implementation of the selected operational concepts in a design or existing site infrastructure. The standard does not direct the specific design solution or technology the owner or design team must use to reach the site performance objective. Owners are free to choose any number of UPS configurations, products, or manufactures—as long as the result can meet the target Tier classification. Moreover, the use of static or rotary UPS systems, fuel cell technologies, direct expansion cooling, or air or water cooled chillers are left to the owner. The Tier Standards have attained wide acceptance because they allow the owner to include such concerns as first cost, operations complexity, and product availability as appropriate, while still focusing on the desired operational outcome of the completed facility.

In addition to availability, other owner requirements must be addressed in infrastructure design. Protection of data or physical assets is independent of the site infrastructure Tier classification. The increasing power densities of IT equipment required other considerations than the redundancy in the power and cooling systems. Project elements like video surveillance and gaseous fire suppression are frequently necessary to meet an owner's regulatory or insurance requirements, completely separate from IT availability objectives. The key understanding required for a successful data center operation is to differentiate between Tier Performance Standard criteria, owner risk and cost tolerance, and Industry Best Practices.

Consideration of cost, risk tolerance, and Best Practices clearly point to a wider number of site infrastructure characteristics than Tier classification, alone. Experience with the Tier Standard since its inception indicates that Sustainability characteristics become an important factor over time. Investments in Sustainability characteristics account for much of the variance within individual Tier solutions, often leading to increased availability. Typically, Sustainability characteristics decrease the cost or risk of completing maintenance, or speed the recovery from site infrastructure incidents. Less costly and less risky maintenance means the work is more likely to be completed, keeping the equipment in better condition and calibration. More operations-centric designs make operations easier, so fewer mistakes are made.

Illustrative Examples

Some examples can illustrate site infrastructure characteristics that impact sustainability, while not affecting the overall Tier classification of the solution.

- A topology that can switch the power source for all mechanical components so they continue running when any electrical panel is shut down eliminates an operations constraint to maintenance. Procedures that require critical cooling equipment to be shut down during recurring electrical system maintenance may not be allowed if another chiller is out of service for repairs. Missed maintenance leads to decreased reliability.
- A design that mounts critical components in difficult to reach areas or limits access space in the central plant may increase the time required to maintain important systems. The increased time window may eliminate the ability to schedule the maintenance activity.
- Installing engine generators and switchgear inside the facility (with adequate access space) eliminates the effects of weather and time-of-day on safe maintenance and repair activities.
- In order to improve stability, the combined load on a critical system is often limited to 90% of non-redundant nameplate over a sustained period of time.
- Compartmentalization, a Tier IV requirement, provides benefits for Tier III sites. The effects of evacuation requirements for areas affected by refrigerant leaks can be limited to the number of redundant chillers by careful Compartmentalization. Chillers that are necessary to keep the computer room cool can continue to operate while those in a separate compartment are shut down to purge the refrigerant.
- Compartmentalization of the primary and maintenance electrical distribution paths also provides a major advantage to a site. If an arc flash or electrical fire (an "unplanned event") occurred in a Tier III site, the site could be disrupted. However, if the maintenance path is physically separated from the normal path, compartmentalization would permit the site to rapidly recover on a power path through a completely different space than where the fire occurred.

Each Industry Has a Unique Uptime Need Driving the Site Infrastructure Tier Level Required

After careful alignment of IT availability objectives with site infrastructure performance expectations, an informed company may select a site representing any of the Tier classifications. Some considerations for selecting an appropriate site infrastructure Tier are:

Tier I is appropriate for firms such as

- Small businesses where information technology primarily enhances internal business process
- Companies who principal use of a "web-presence" is as a passive marketing tool
- Internet-based startup companies without quality of service commitments

These companies typically do not have an established revenue stream or identifiable financial impact of disruption due to data center failure. Sometimes companies with an established revenue steam will select Tier I topology because their applications have a low availability requirement, such as only during a 5.5-day business week. Other companies may select Tier I topology if they plan to abandon the site when the business requirements exceed the Tier I functionality.

Tier II is appropriate for firms such as

- Internet-based companies without serious financial penalties for quality of service commitments
- Small businesses whose information technology requirements are mostly limited to traditional normal business hours, allowing system shutdown during "off-hours"
- Commercial research and development firms, such as software, who do not typically have "on-line" or "real-time" service delivery obligations

These companies typically do not depend on real-time delivery of products or services for a significant part of their revenue stream, or are contractually protected from damages due to lack of system availability. Occasionally companies will select Tier II infrastructure if they have become burdened with impacts due to nuisance equipment outages associated with Tier I sties. A large number of institutional and educational organizations select Tier II infrastructure because there is no meaningful impact of disruption due to data center failure. Some companies have successfully used Tier II infrastructure to provide off-site electronic vaulting for offline data.

Typical applications for Tier III facilities are

- Companies that support internal and external clients 24x7 such as service centers and help desks, but can schedule short periods when limited service is acceptable
- Businesses whose information technology resources support automated business processes, so client impacts of system shutdowns is manageable
- Companies spanning multiple time zones with clients and employees spanning regional areas

Companies selecting Tier III infrastructure usually have high-availability requirements for ongoing business, or have identified a significant cost of disruption due to a planned data center shutdown. These companies are willing to accept the impact of disruption risk of an unplanned event. However, Tier III is appropriate for companies who expect the functionality requirements to increase over time and do not want to abandon the data center. Sometimes, these companies design a Tier III site to be upgraded to Tier IV.

Tier IV is justified most often for

- Companies with an international market presence delivering 24x365 services in a highly competitive client-facing market space
- Businesses based on E-commerce, market transactions, or financial settlement processes
- Large, global companies spanning multiple time zones where client access to applications and employee exploitation of information technology is a competitive advantage

Companies who have extremely high-availability requirements for ongoing business, or for whom there is a profound cost of disruption due to any data center shutdown, select Tier IV site infrastructure. These companies will know the cost of a disruption, usually in terms of both actual dollar costs and impact to market share. The cost of disruption makes the case for investment in high availability infrastructure a clear business advantage.

Making the Appropriate Tier Selection Should Be Based on Business Requirements

Selecting the site infrastructure solution based on the availability objectives required to sustain well-defined business processes with substantial financial consequences for downtime provides the best foundation for investment in data center facilities. The owners' focus during the data center design and delivery process should be the consistent application of the Tier Performance Standard, rather than allowing recurring debate over every characteristic or attribute that makes up the data center's site infrastructure.

Including criteria from a higher Tier classification, or an attribute leading to increased availability, does not increase the overall Tier classification. Moreover, deviation from the Tier standard in any subsystem will prevent a site from classification at that Tier. For example, a UPS system patterned after a Tier IV system within a site featuring a Tier II power distribution backbone will yield a Tier II site. The most significant deviations from the Tier Standard found in most sites can be summarized as inconsistent solutions.

Frequently, a site will have a robust fault tolerant electrical system patterned after a Tier IV solution, but utilize a Tier II mechanical system that cannot be maintained without interrupting computer room operations. This results in the overall site achieving a Tier II rating. Most often the mechanical system fails concurrent maintenance criteria because of inadequate isolation valves in the chilled water distribution path.

Another common oversight is the effect of shutting down electrical panels on the mechanical system the panel feeds. If more than the redundant number of chillers, towers, or pumps is de-energized for electrical maintenance, computer room cooling is impacted.

Occasionally, electrical systems fail to achieve Tier III or Tier IV criteria due to the UPS power distribution path. Topologies that include static transfer switches that cannot be maintained without affecting computer room power, fail the concurrent maintenance criteria. UPS configurations that utilize common input or output switchgear are almost always often unmaintainable without computer room outages and fail the Tier III requirements even after spending many hundreds of thousands of dollars.

Consistent application of standards is necessary to have an integrated solution for a specific data center. It is clear that the IT organization invests heavily in the features offered by newer computer equipment technology. Often, as the electrical and mechanical infrastructures are defined, and the facility operations are established, there is a growing degree of inconsistency in the solutions incorporated in a site. As shown in Figure 3, each segment must be integrated to deliver the overall data center solution. An investment in one segment must be met with a similar investment in each of the other segments if any of the elements in the combined solution are to have effect on IT availability. A well-executed data center master plan or strategy should consistently resolve the entire spectrum of IT and facility requirements.

Figure 3:
Comparing IT Solutions for Reliability, Availability, and Serviceability to Site Infrastructure

	RELIABILITY	AVAILABILITY	SERVICEABILITY
Information Technology	Clustering RAID and DASD Token Ring Console Automation Change Management	Logical Partitions Clustering Mirrored Data Hot Backup Business Continuity	Hot Pluggable Hot Microcode Updates Call Home Remote Service
Electrical Infrastructure	UPS Dual Power S+S	Engine Generator Dual Power S+S	Engine Generators Dual Power S+S
Mechanical Infrastructure	Redundant Components Fans and Pumps on UPS	Thermal Storage	Dual Pipe Thermal Storage
Facility Operations	Passive Automation Change Management MAPS/Certification Simulation	24 By Forever Staffing Compartmentalization Failure Bypass Options On-Site Spares	Work Performed during Regular Hours In-House Knowledge In-House Supervision

Lifecycle Planning

It is disappointing to observe brand new sites that received very little thought during initial design to future operations. Valves were located in inaccessible places, the access path for the addition of future components was not thought out, or sufficient capacity to simultaneously test new systems while sustaining the critical load was not provided. These details could have been addressed for no additional cost during design. This failure limits both investment value and site performance right from its initial occupancy. A more sustainable site will project future requirements and anticipate them during the initial design and construction.

Sites should be designed to anticipate increasing power requirements and tier levels. These sites provide future locations for necessary site infrastructure equipment as well as a planned means to commission them and then connect them transparently to operational systems.

Institute Site Topology and Sustainability Certification

The *Institute* exclusively reserves the right to determine Tier ranking and to certify sites as meeting Tier requirements as summarily described in this white paper. This comprehensive process involves additional criteria beyond the information provided herein. The process is similar to that used for ISO 900X certification. The ISO standard is set and maintained by the International Standards Organization who trains and certifies field inspection agencies in different parts of the world. These field inspectors inspect and validate conformance to the ISO standard before certification is granted for a limited time period. The *Institute* has licensed ComputerSite Engineering Inc., a separate but related company, to perform inspection and validation utilizing the *Institute's* Tier Performance Standards and the Institute's comprehensive database of emerging industry problems and best design practices. Sites reviewed and certified by the *Institute* can be seen at www.uptimeinstitute.org/tui_certification.html.

Conclusion

Data center owners have the responsibility to determine what Tier of functionality is appropriate or required for their sites. As such, it is a business decision to determine the Tier necessary to support site availability objectives. Part of this decision is to balance the IT operational practices with the facility practices that support the IT world. Once selected, however, the desired Tier should be uniformly implemented.

About the Authors

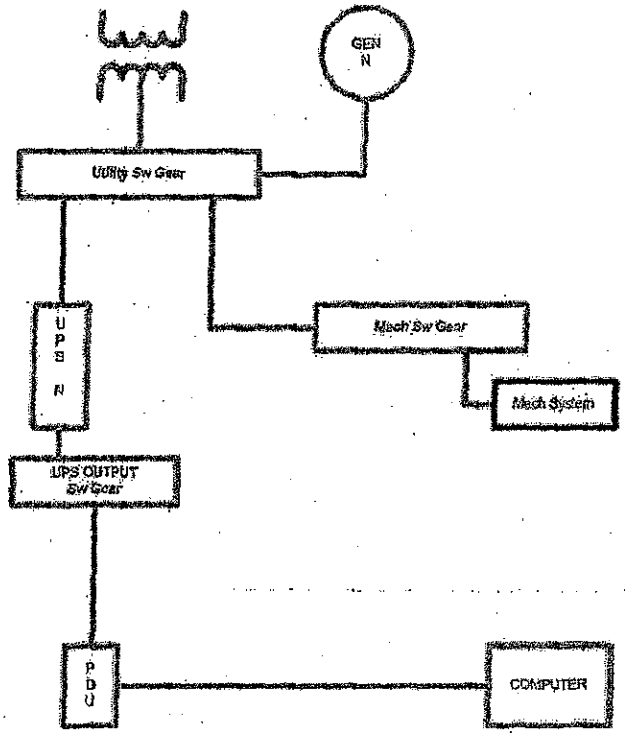
Mr. Turner is a Distinguished Fellow and Senior Certification Authority for the *Institute* and a Principal of ComputerSite Engineering, Inc. in Santa Fe, NM.
Mr. Seader is a Distinguished Fellow and Certification Authority for the *Institute* and a Principal of ComputerSite Engineering, Inc. in Santa Fe, NM.
Mr. Brill is the founder of the *Institute* and is its Executive Director. He is a Principal of ComputerSite Engineering.

About The Uptime Institute

The Uptime Institute, Inc. is a pioneer in creating and operating knowledge communities for improving uptime effectiveness in data center Facilities and Information Technology organizations. The 85 members of the *Institute's* Site Uptime® Network are committed to achieving the highest levels of availability with many being Fortune 100 companies. They interactively learn from each other as well as from *Institute* sponsored meetings, site tours, benchmarking, best practices, uptime effectiveness metrics, and abnormal incident collection and trend analysis. From this interaction and from client consulting work, the *Institute* prepares white papers documenting Best Practices for use by Network members and for the broader uninterruptible uptime industry. The *Institute* also conducts sponsored research and offers insightful seminars and training in site infrastructure management.

THE UPTIME INSTITUTE
Building 100
2904 Rodeo Park Drive East • Santa Fe, NM 87505
Fax (505) 982-8484 • Phone (505) 986-3900
tui@uptimeinstitute.org • www.uptimeinstitute.org

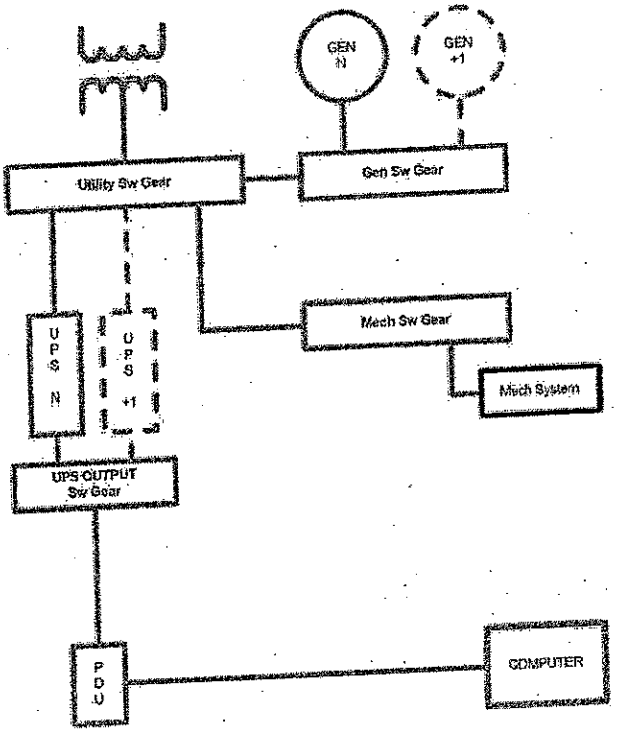
Illustrative Electrical System Topology - Tier I



Note: This diagram illustrates basic Tier I electrical distribution concepts. This diagram shall not be interpreted to represent a standard or compliant electrical system topology, or a solution fulfilling any particular set of requirements.

Site certification requires consistent application of Tier concepts to all 16 critical subsystems that comprise data center site infrastructure.

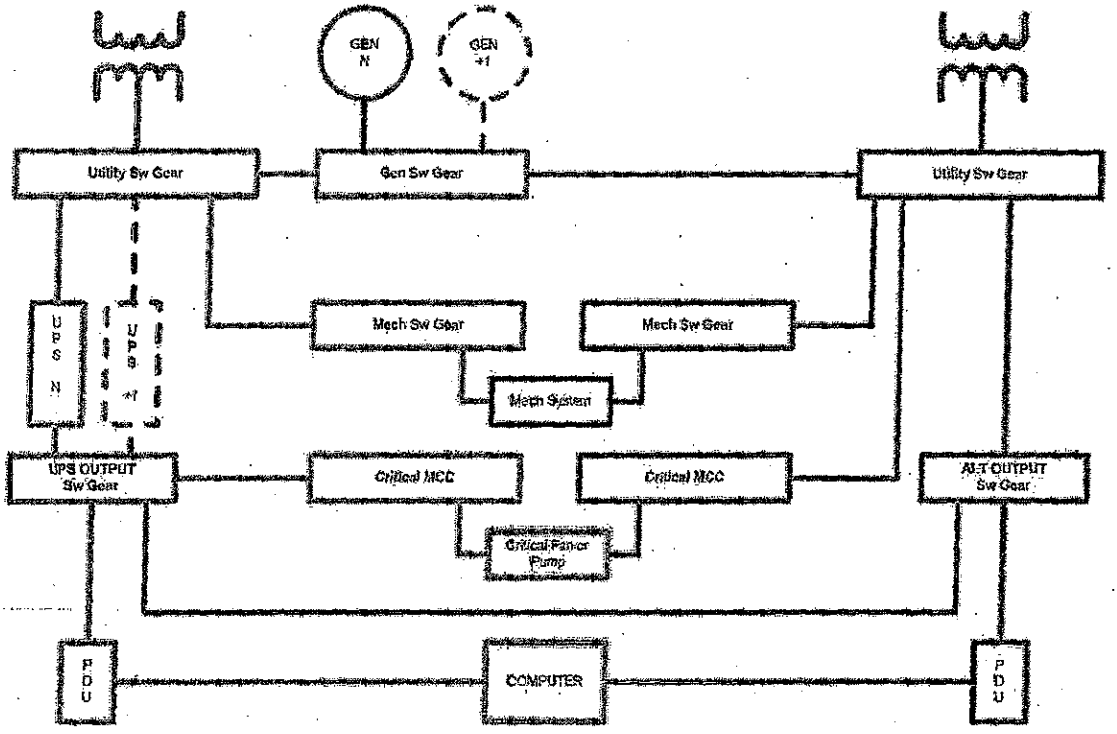
Illustrative Electrical System Topology - Tier II



Note: This diagram illustrates a basic Tier II electrical distribution concept. This diagram shall not be interpreted to represent a standard or compliant electrical system topology, or a solution fulfilling any particular set of requirements.

Site certification requires consistent application of Tier concepts to all 16 critical subsystems that comprise data center site infrastructure.

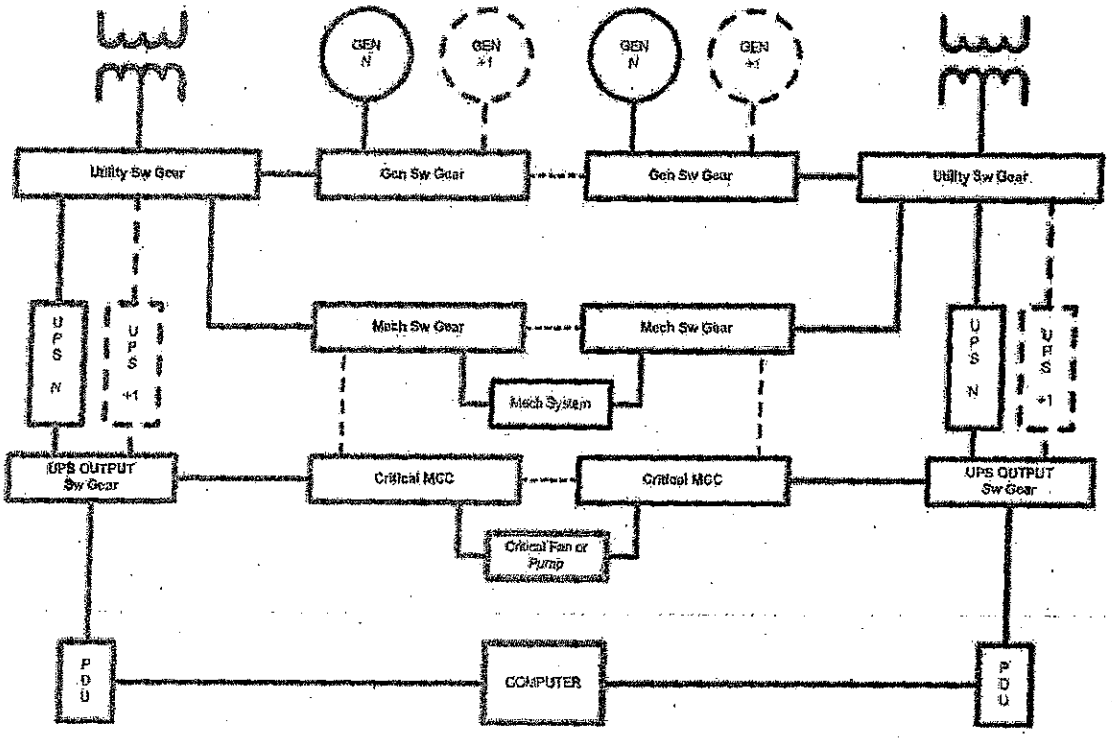
Illustrative Electrical System Topology - Tier III



Note: *This diagram illustrates a Tier III electrical distribution concept. This diagram shall not be interpreted to represent a standard or compliant electrical system topology, or a solution fulfilling any particular set of requirements.*

Site certification requires consistent application of Tier concepts to all 16 critical subsystems that comprise data center site infrastructure.

Illustrative Electrical System Topology - Tier IV



Note: This diagram illustrates a Tier IV electrical distribution concept. This diagram shall not be interpreted to represent a standard or compliant electrical system topology, or a solution fulfilling any particular set of requirements.

Site certification requires consistent application of Tier concepts to all 16 critical subsystems that comprise data center site infrastructure.

Appendix B
Department of Justice
Minimum Security Standards for Federal Facilities

The following table lists a summary of the security standards for obtaining certification for the various Department of Justice (DOJ) Facility Security Levels.
Legend:
Minimum Standard = M
Standard based on facility evaluation = F
Desirable = G
Not Applicable = N/A

	Level of security				
Perimeter Security	I	II	III	IV	V
Parking					
Control of facility parking	G	G	M	M	M
Control of adjacent parking	G	G	G	F	F
Avoid leases in which parking cannot be controlled	G	G	G	G	G
Leases should provide security control for adjacent parking	G	G	G	G	G
Post signs and arrange for towing unauthorized vehicles	F	F	M	M	M
ID system and procedures for authorized parking (placard, decal, card key, etc.)	G	G	M	M	M
Adequate lighting for parking areas	G	G	M	M	M
Closed circuit television (CCTV) monitoring					
CCTV surveillance cameras with time lapse video recording	G	F	F	M	M
Post signs advising of 24 hour video surveillance	G	F	F	M	M
Lighting					
Lighting with emergency power backup	M	M	M	M	M
Physical barriers					
Extend physical perimeter with concrete and/or steel barriers	N/A	N/A	G	F	F
Parking barriers	N/A	N/A	G	F	F

Legend:
Minimum standard = M Standard based on facility evaluation = F
Desirable = G Not applicable = N/A
Source: Vulnerability Assessment of Federal Facilities, Department of Justice, June 28, 1995.

	Level of security				
Entry Security	**I**	**II**	**III**	**IV**	**V**
Receiving/Shipping					
Review receiving/shipping procedures (current)	M	M	M	M	M
Implement receiving/shipping procedures (modified)	G	F	M	M	M
Access control					
Evaluate facility for security guard requirements	G	F	M	M	M
Security guard patrol	G	G	F	F	F
Intrusion detection system with central monitoring capability	G	F	M	M	M
Upgrade to current life safety standards (fire detection, fire suppression systems, etc.)	M	M	M	M	M
Entrances/Exits					
X-ray and magnetometer at public entrances	N/A	G	F	F	M
Require x-ray screening of all mail/packages	N/A	G	F	M	M
Peepholes	F	F	N/A	N/A	N/A
Intercom	F	F	N/A	N/A	N/A
Entry control with CCTV and door strikes	G	F	N/A	N/A	N/A
High security locks	M	M	M	M	M

Legend:
Minimum standard = M Standard based on facility evaluation = F
Desirable = G Not applicable = N/A
Source: Vulnerability Assessment of Federal Facilities, Department of Justice, June 28, 1995.

	Level of security				
Interior Security	I	II	III	IV	V
Employee/Visitor identification					
Agency photo ID for all personnel displayed at all times	N/A	G	F	M	M
Visitor control/screening system	G	M	M	M	M
Visitor identification accountability system	N/A	G	F	M	M
Establish ID issuing authority	F	F	F	M	M
Utilities					
Prevent unauthorized access to utility areas	F	F	M	M	M
Provide emergency power to critical systems (alarm systems, radio communications, computer facilities, etc.)	M	M	M	M	M
Occupant emergency plans					
Examine occupant emergency plans (OEP) and contingency procedures based on threats	M	M	M	M	M
OEPs in place, updated annually, periodic testing exercise	M	M	M	M	M
Assign & train OEP officials (assignment based on largest tenant in facility)	M	M	M	M	M
Annual tenant training	M	M	M	M	M
Daycare centers					
Evaluate whether to locate daycare facilities in buildings with high threat activities	N/A	M	M	M	M
Compare feasibility of locating daycare in facilities outside locations	N/A	M	M	M	M

Legend:
Minimum standard = M Standard based on facility evaluation = F
Desirable = G Not applicable = N/A
Source: Vulnerability Assessment of Federal Facilities, Department of Justice, June 28, 1995.

	Level of security				
Security Planning	I	II	III	IV	V
Intelligence Sharing					
Establish law enforcement agency/security liaisons	M	M	M	M	M
Review/establish procedure for intelligence receipt and dissemination	M	M	M	M	M
Establish uniform security/threat nomenclature	M	M	M	M	M
Training					
Conduct annual security awareness training	M	M	M	M	M
Establish standardized unarmed guard qualifications/training requirements	M	M	M	M	M
Establish standardized armed guard qualifications/training requirements	M	M	M	M	M
Tenant assignment					
Co-locate agencies with similar security needs	G	G	G	G	G
Do not co-locate high/low risk agencies.	G	G	G	G	G
Administrative procedures					
Establish flexible work schedule in high threat/high risk areas to minimize employee vulnerability to criminal activity	F	F	G	G	G
Arrange for employee parking in/near building after normal workhours	F	F	F	F	F
Conduct background security checks and/or establish security control procedures for service contract personnel	M	M	M	M	M
Construction/Renovation					
Install mylar film on all exterior windows (shatter protection)	G	G	F	M	M
Review current projects for blast standards	M	M	M	M	M
Review/establish uniform standards for construction	M	M	M	M	M
Review/establish new design standards for blast resistance	F	F	M	M	M
Establish street setback for new construction	G	G	F	M	M

Legend:
Minimum standard = M Standard based on facility evaluation = F
Desirable = G Not applicable = N/A
Source: Vulnerability Assessment of Federal Facilities, Department of Justice, June 28, 1995.

USDA

United States Department of Agriculture

Attachment 2 Application for Classification As an USDA Enterprise Data Center

Instructions: Complete all sections of the application. Refer to Enterprise Class Data Center definition and appendices, issued as Attachment 1 of memorandum titled "Development of USDA Data Center Strategy" from David Combs, Chief Information Officer, and Charles R. Christopherson, Jr., Chief Financial Officer.

Section One: Contact Information

Agency:_____

Point of Contact:
Name:_____
Phone Number:_____
Email Address:_____

Alternate Point of Contact:
Name:_____
Phone Number:_____
Email Address:_____

Data Center Physical Address (include room number(s)):

_____
_____
_____

Section Two: Site Specific Information – Physical Standards

Electrical Systems:

1. **Uptime Institute Tier rating of your data center's electrical systems (check one):**

☐ Tier I ☐ Tier II ☐ Tier III ☐ Tier IV

2. **Backup Emergency Power Source for data center loads (including mechanical equipment):**

☐ Diesel Generator ☐ LP/Natural Gas Generator ☐ Fuel Cell

☐ Other

Electrical Systems: (continued)

3. Uninterruptible Power Supply(s) for data center computing loads:

☐ Single ☐ Isolated Redundant ☐ Parallel Redundant ☐ Distributed Redundant

☐ System + System Redundant

4. Commercial Power Feeds:

☐ Single ☐ Dual ☐ Dual-Diverse (from multiple substations or grids)

5. 24 Hour Service Contracts in-place for UPS, Generators and Batteries

☐ Yes ☐ No

6. Data Center electrical infrastructure includes a power conditioning component to ensure fluctuations in commercial power quality are isolated from the critical computing loads

☐ Yes ☐ No

Mechanical Systems:

1. Uptime Institute tier rating of mechanical systems (check one):

☐ Tier I ☐ Tier II ☐ Tier III ☐ Tier IV

2. HVAC systems provide automated climate controls and system management

☐ Yes ☐ No

3. Data center has raised floor plenum

☐ Yes ☐ No

4. Fire Suppression Systems:

☐ Wet Pipe ☐ Dry Pipe ☐ Gas ☐ Dual Source (i.e. Gas and Dry Pipe)

5. 24 Hour Service Contracts in-place for HVAC and Fire Suppression Systems

☐ Yes ☐ No

6. Presence of alarmed water monitoring system within the data center

☐ Yes ☐ No

7. Water piping and/or drains installed above the data center space

☐ Yes ☐ No

Physical Security:

1. Does the candidate data center meet all standards to qualify as a Department of Justice Level IV Facility

☐ Yes ☐ No

If yes, please attach the completed Department of Justice matrix.

2. Data Center meets the requirements of DM3510-01, Physical Security Standards for IT Restricted Spaces

☐ Yes ☐ No

Operational Standards

Information Security Systems:

1. **Guidelines for server configuration hardening are documented and approved by the Agency CIO**

☐ Yes ☐ No

2. **Monthly vulnerability scanning is performed on all devices**

☐ Yes ☐ No

3. **A patch management system is in place to address vulnerabilities**

☐ Yes ☐ No

4. **Describe the Background Investigation levels required for personnel working in the data center. (Please list your answer by job category, such as system administrator, network administrator, help desk, tape librarian, etc.)**

Description: (use additional pages as necessary)

5. **Certification and Accreditation of all data center General Support Systems are fully approved and current**

☐ Yes ☐ No

6. **Documented incident response processes and procedures are in-place and periodically exercised**

☐ Yes ☐ No

Network and Telecommunications:

1. **The data center has an UTN node on-site**

☐ Yes ☐ No

2. **Firewalls are configured for high availability**

☐ Yes ☐ No

3. **All local area networks are protected by an intrusion detection system**

☐ Yes ☐ No

4. **Describe the telecommunications and network architecture**

Description: (use additional pages as necessary)

Data Center Staff:

1. ***Is there a full time, dedicated data center staff to include both management* and technical, with on-duty operations and system security managers**

☐ Yes ☐ No

2. **Helpdesk is staffed 24x7x365**

☐ Yes ☐ No

3. **Personnel physically monitor computing systems 24x7x365**

☐ Yes ☐ No

Management Processes:

1. **Do you have a formal project in place to align data center service delivery with Information Technology Infrastructure Library (ITIL) processes**

☐ Yes ☐ No

If yes, describe the project and provide current status (use additional pages as necessary).

2. **Service level agreements and performance reporting metrics are documented and in place**

☐ Yes ☐ No

3. **Standard operating procedures for scheduling and coordinating maintenance are documented and in place**

☐ Yes ☐ No

4. **Data Center costs and customer service have been benchmarked against industry**

☐ Yes ☐ No

Disaster Recovery:

1. **Disaster recovery, business continuity, and emergency response plans/processes are documented and periodically tested**

☐ Yes ☐ No

Section Three: Additional Information

Provide any additional information that will assist the data center assessment team in the analysis or your data center. Please limit the response to 5 pages.