



Industry Standard Tier Classifications Define Site Infrastructure Performance

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

One of the most common sources of confusion in the field of uninterrupted 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. As Internet website hosting continues to grow into a major industry, competing companies with data centers of radically different infrastructure capabilities are all claiming to deliver “high availability.”

With the explosive growth of the Internet comes an increased demand for computer hardware reliability. Information technology customers expect reliability of “Five Nines,” or 99.999%.

Unfortunately, the substantial investment a business makes to achieve “Five Nines” in its computer hardware is insufficient to protect its mission-critical computing functions. These significant investments must be accompanied by a solid understanding of how well their site infrastructure can support their availability goals.

The Uptime Institute® has developed a tiered classification approach to site infrastructure functionality that addresses the need for a common benchmarking standard. The Institute’s system has been under development for several years, and includes measured availability figures ranging from 99.67% to more than 99.99% It is important to note that this range of availability is substantially less than the current Information Technology (IT) expectations for “Five Nines.”

Over the last forty years, data center designs have evolved through at least four distinct stages, which are captured in the Institute’s classification system. Tier I first appeared in the early sixties, Tier II in the seventies, Tier III in the late eighties and early nineties, and Tier IV in 1994 with the United Parcel Service Windward project, which was the first site to assume the availability of dual-powered computer equipment. The Uptime Institute participated in the development of Tier III concepts and pioneered the creation of Tier IV.

Invention of Tier IV was made possible by Ken Brill, Executive Director of The Uptime Institute, who envisioned a future when all computer hardware would come with dual power inputs. During construction of the \$50 million Windward project, United Parcel Service worked with IBM and other computer hardware manufacturers to provide dual-powered computer hardware.

Dual power technology requires having at least two completely independent electrical systems. These dual systems supply power via diverse power paths to the computer load, which moves the last point of electrical redundancy from within the Uninterruptible Power System (UPS) down to within the computer hardware itself. Brill’s intuitive conclusion has since been confirmed by Uptime Institute research that has determined that 95% of all site infrastructure failures occur between the UPS and the computer load. Since completion of the Windward project in 1994, System+SystemSM Tier IV electrical designs have become common and the number of computer hardware products with dual inputs has grown.

The advent of dual-powered computer hardware in tandem with Tier IV electrical infrastructure is an example of site infrastructure design and computer hardware design simultaneously achieving higher



This chart illustrates Tier similarities and differences

	TIER I	TIER II	TIER III	TIER IV
Number of delivery paths	Only 1	Only 1	1 active 1 passive	2 active
Redundant components	N	N+1	N+1	2 (N+1) or S+S
Support space to raised floor ratio	20%	30%	80-90%	100%
Initial watts/ft²	20-30	40-50	40-60	50-80
Ultimate watts/ft²	20-30	40-50	100-150	150+
Raised floor height	12"	18"	30-36"	30-36"
Floor loading pounds/ft²	85	100	150	150+
Utility voltage	208, 480	208, 480	12-15kV	12-15kV
Months to implement	3	3 to 6	15 to 20	15 to 20
Year first deployed	1965	1970	1985	1995
Construction \$/ft² raised floor*	\$450	\$600	\$900	\$1,100+
Annual IT downtime due to site	28.8 hrs	22.0 hrs	1.6 hrs	0.4 hrs
Site availability	99.671%	99.749%	99.982%	99.995%

*Excludes land and abnormal civil costs. Assumes minimum of 15,000 ft² of raised floor, architecturally plain one story building fitted out for the initial capacity, but with the backbone designed to reach the ultimate capacity with the installation of additional components. Make adjustments for NYC, Chicago, and other high cost areas.

© 2001 The Uptime Institute

availability. With the significant improvements in computer hardware design currently being made, many data centers constructed even in the last five years offer only Tier I, II, or III functionality, falling far behind in their capacity to match the availability offered by the information technology they support.

Defining the Tiers

The tier classification system involves several definitions. A site that can sustain at least one “unplanned” worst-case site infrastructure failure with no critical load impact is considered fault tolerant. A site that is able to perform planned site infrastructure activity without shutting down critical load is concurrently maintainable (fault tolerance level may be reduced during concurrent maintenance). It is important to remember that a typical data center site is composed of at least twenty major mechanical, electrical, fire protection, security and other systems, each of which has additional subsystems and components. All of these must be concurrently maintainable and/or fault tolerant for the entire site to be considered concurrently maintainable and/or fault tolerant.

Some sites built with fault-tolerant System+System electrical concepts failed to incorporate the mechanical analogy, which involves dual mechanical systems.

Such sites are classified Tier IV electrically, but only achieve a Tier II level mechanically. The following list summarizes the characteristics of each Tier.

■ **Tier I**

Single path for power and cooling distribution, no redundant components, 99.671% availability.

■ **Tier II**

Single path for power and cooling distribution, redundant components, 99.741% availability.

■ **Tier III**

Multiple power and cooling distribution paths, but only one path active, redundant components, concurrently maintainable, 99.982% availability.

■ **Tier IV**

Multiple active power and cooling distribution paths, redundant components, fault tolerant, 99.995% availability.

The availability numbers have been drawn from industry benchmarking conducted by The Uptime Institute and sites in the top 10 percent (this means only 10% of all sites performed at this level). The quality of human-factors management is the most significant element separating top sites from all others.



The Chart illustrates how these ideas are mapped over the architecture of 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 System + System	Engine generators Dual power System + System	Engine generators Dual power System + System
Mechanical Infrastructure	Redundant components Fans and pumps on UPS	Thermal Storage	Dual pipe Thermal Storage Dual plant
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

© 2001 The Uptime Institute

Tier I Data Center *Basic*

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 Data Center *Redundant Components*

Tier II facilities with redundant components are slightly less susceptible to disruptions from both planned and unplanned activity than a basic 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 Data Center *Concurrently Maintainable*

Tier III level capability allows for any planned site infrastructure activity without disrupting the computer hardware operation in any way. Planned

activities include preventive and programmable maintenance, repair and replacement of components, addition or removal of capacity components, testing of components and systems, and more. For large sites using chilled water, this means two independent sets of pipes. 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 III sites are often designed to be upgraded to Tier IV when the client's business case justifies the cost of additional protection.

Tier IV Data Center *Fault Tolerant*

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. Because of fire and electrical safety codes, there will still be downtime exposure due to fire alarms or people initiating an Emergency Power Off (EPO). Tier IV requires all computer hardware to have dual power inputs as defined by the *Institute's* Fault-



Tolerant Power Compliance Specification Version 2.0 (www.uptimeinstitute.org/spec.html).

Tier IV site infrastructures are the most compatible with high availability IT concepts that employ CPU clustering, RAID DASD, and redundant communications to achieve reliability, availability, and serviceability. The accompanying chart shows how these IT ideas relate to site infrastructure concepts.

Solving Incompatible “Five Nines” Expectations

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 can deliver over time is 99.995%, and this assumes a site outage occurs only as a result of a fire alarm or EPO, and that such an event occurs no more than once every five years. 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 additional failure is likely over five years. While the site outage is assumed to be instantaneously restored (which requires 24 x “forever” staffing), it can still require up to four hours for IT to recover information availability.

Tier IV’s 99.995% uptime is an average over five years. An alternative calculation using the same underlying data is 100% uptime for four years and 99.954% for the year in which the downtime event occurs.

Higher levels of site uptime can be achieved by protecting against accidental activation or the real need for fire protection and EPOs. Preventatives include high sensitivity smoke detection, limiting fire load, signage, extensive training, staff certification, limiting the number of “outsiders” in critical spaces, and treating people well to increase pride in their work. All of these measures, if taken, can reduce the risk of failures. Other solutions include placing the redundant parts of the IT computing infrastructure in different site infrastructure compartments so that a site infrastructure 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 four hours to restore. These operational issues can improve the

availability offered by any data center, and are particularly important in a “Four Nines” Tier IV data center housing IT equipment requiring “Five Nines” availability.

Authorship

Pitt Turner is a professional engineer, a distinguished fellow of The Uptime Institute, and a principal in Computersite Engineering®. He has guided over \$1.5 billion in site infrastructure investment for primarily Fortune 50 clients. Ken Brill is Executive Director of The Uptime Institute, and a principal in Computersite Engineering. He is the founder of the Site Uptime Network® and invented dual power distribution technology in 1991 for high availability data centers.

The Uptime Institute is a pioneer in creating and operating knowledge communities for improving uptime effectiveness in data center Facilities and Information Technology organizations. The fifty-one 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.

This white paper may be quoted, reproduced, or distributed in its entirety at no charge. The Uptime Institute exclusively reserves the right to certify and determine the tier ranking of site infrastructures as defined in this white paper. The *Institute’s* comprehensive site certification process involves significant additional criteria beyond the summary level information provided in this document. Sites reviewed and certified by the *Institute* can be seen at www.upsite.com. This white paper is posted and maintained on The Uptime Institute’s website at www.upsite.com/TUIpages/tuiwhite.html.

© 2001 The Uptime Institute



1347 Tano Ridge Rd, Santa Fe, NM 87506
Fax (505) 982-8484 Phone (505) 986-3900 E-mail tui@upsite.com