

An Empirical Study on Data Center System Failure Diagnosis

Montri Wiboonrat

Graduate School of Information Technology, Assumption University, Bangkok, Thailand
mwiboonrat@gmail.com, montri.w@tss-design.com

Abstract

Data center downtime causes business losses over a million dollars per hour. 24x7-hour data availability is critical to numerous systems, e.g. public utilities, hospitals, and data centers. Service interruption signifies lives or deaths, higher costs, and poor service quality. This research conducted the system diagnosis of reliability assessment for Tier IV data centers (DC), employing the Failure Modes, Effects, and Criticality Analysis (FMECA) and the Reliability Block Diagram (RBD). The techniques of series-parallel, active standby, k-out-of-n, bridge, full redundancy, fault-tolerant, and multiple utilities were applied in the system failure diagnosis to provide high system availability. Component reliability data were obtained from the IEEE Std. 493 Gold Books. Simulation results from data center system failure diagnosis reveal the functional steps of data center downtime and pinpoint solutions to terminate or mitigate the data center downtime. Proposed improvements on the component's inherent characteristics (CIC) and the system connectivity topology (SCT) help reduce the failure rate by 1.1706 hours in 1,000,000 hours of operation.

1. Introduction

Reliability is the measurement that expresses the ability of a system to perform its intended function during a specified time. Reliability levels are mutually dependent with economics. Increased reliability is attained through increased investment. It also permits the consumers to reduce their downtime costs [9]. The reliability assessment is normally accomplished with such indicators as the mean time between failure (MTBF) and the mean time to restoration (MTTR). The probability of occurrences and the duration of each outage may also be adopted in reliability assessment.

FMECA represents a widely-accepted inductive method for identifying the component probability of failures and potential system reliability. It prescribes a structured and formalized step to cause-and-effect

analysis. The problems can thus be mitigated at an early stage to avoid complicated and costly corrective actions. While FMECA can be applied at an initial design stage, it is also applicable at development, validation, implementation, or operation stages [5].

A reliable equipment data source is instrumental to a systematic precision analysis [12]. The reliability data were obtained from the IEEE 493 Gold Book Standard Network and component vendor's filed test data. This paper adopted a Reliability Block Diagram (RBD) with Monte Carlo Simulation to compare the traditional Tier IV data center topology with the new reliability enhancement data center design. With parallel processing analysis, FMECA analyzes the power distribution system in sub-system connectivity topology and component reliability inheritance.

2. Definitions

Availability: A ratio describing the percentage of time a component or system can perform its required function.

Lambda (λ): An inverse of the mean exposure time between consecutive failures. It is typically expressed in either failures per year or failures per million hours.

MTBF: The mean exposure time between consecutive failures of a component or a system. The mean time between failures is usually expressed in either years per failure or million hours per failure. For some applications, measurement of mean time between repairs (MTBR) rather than mean time between failures may provide more statistically correct information.

MTTR: The mean time to repair a failed component. For a system, it is the total amount of time it is unavailable due to a failure and is expressed in hours.

Reliability: An indication of the ability of a component or a system to perform its intended function during a specified time.

Systems: A group of components connected or associated in a fixed configuration to perform a specified function.

3. Reliability Determination

Calculation of all components and systems reliability stems from two basic parameters, including Mean Time Between Failure (MTBF) and Mean Time To Repair (MTTR). Following show equations derived from both [2], [15]:

$$\lambda = \frac{1}{MTBF} \quad ; \text{Component failure rate (fault/hour)}$$

$$\mu = \frac{1}{MTTR} = \frac{1}{r} \quad ; \text{Component repair rate (repairs/hour)}$$

The system availability (A) is defined as the total operating time over the total time:

$$A = \frac{MTBF}{(MTBF + MTTR)} \quad \text{or} \quad A = \frac{\mu}{(\lambda + \mu)}$$

$$q = \frac{\lambda}{\lambda + \mu} \quad ; \text{(Steady-state unavailability)}$$

$$\omega = \lambda(1 - q) \text{ or } \omega \approx \lambda, \text{ where as } q \ll 1; (q \approx \lambda.r)$$

4. Methodology

In the ideal situation of data center implementation, past system simulation and system failure diagnosis are bases for designing a new data center. Understanding of all system failure symptoms is the key to improving data center system reliability/ availability. Before initiating a DC project, problems may still be hidden because of new technology integration, lack of skills, miscommunication and interpretation, weather conditions, regulations, etc. After the project start, each problem will surface and is resolved by personnel experience or experts until project completion. Fig. 1 illustrates a relationship between the levels of technical problems and uncertainties along the project life cycle. A data center is a complex system comprising 16 sub-systems [14]. Table I shows a classification of data center complexity levels, ranging from material, component, subsystem, system, to an array of systems. A DC system diagnosis must cover all levels of data center complexity to identify possible points of failure and to determine MTBF and MTTR at component and system levels. Diagnosis through the specific root cause of problems is the key concept of FMECA [11],[18]. All equipment reliability data are obtained from IEEE 493 Gold Book Standard, as shown in Table II. Fig. 2 depicts a single line diagram of a

representative network for the Tier IV data center. The components shown in the networks are labeled with numbers, which correspond with the reference numbers in Table II. Network reliability analysis is performed with reliability data for referenced components taken from this table.

This paper investigated the system reliability/ availability of a Tier IV data center in terms of the frequency and duration of power outages. System availability depends on:

1. Reliability and maintainability of its components: including failure rate (MTBF) and repair time (MTTR) of component's inherent characteristics (CIC) distribution, failure modes effects and criticality analysis (FMECA), and environmental effects.
2. System design or system connectivity topology (SCT) (configuration or topology, dependency, and failure detection).
3. System operation behavior (operational characteristics, switching procedures, and maintenance services).

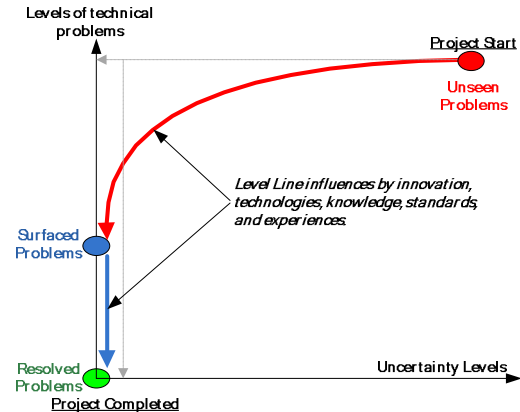


Fig. 1 A Relationship between Levels of Technical Problems and Uncertainties

The following assumptions apply to the proposed Tier IV data center system networks, as seen in Fig. 3:

- Failure rates and repair times are exponentially distributed.
- Actual power cable lengths may be indicated on the drawing. The cable failure rate is thus determined per the indicated actual cable length.
- The generators are 2N redundant.
- The power grids, generators and UPSs are 2(N+1) redundant, applicable only to Tier IV.
- The transformers, switchgears, automatic transfer switches (ATSs) and bus bars are redundant.

- There are two paths of power distribution systems.
- Terminations and splices, while normal for all systems, are not included on the drawing, and are not included in the calculations.
- The assumed breaker failure modes are 50% open and 50% short.
- Selected component's inherent characteristics (CIC) of both Tier III and Tier IV systems are similar.

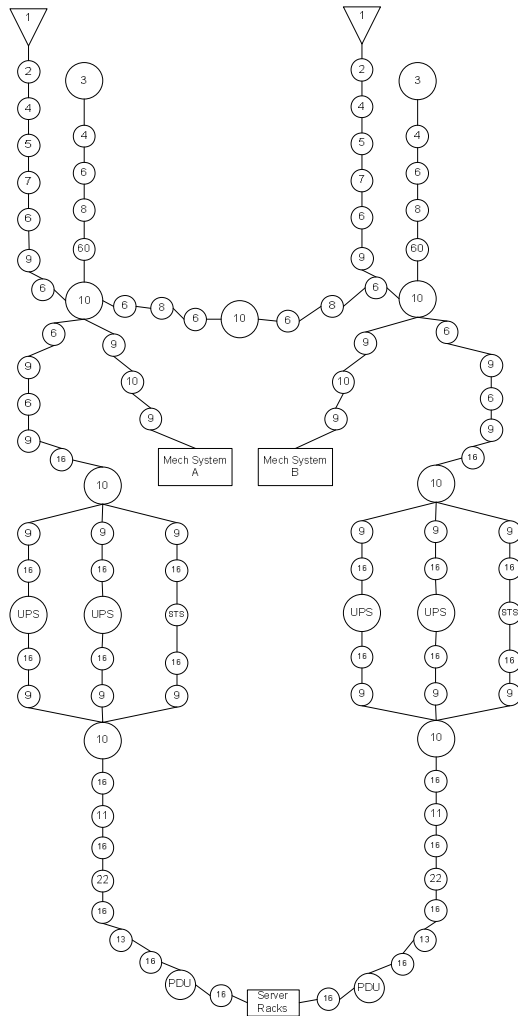


Fig. 2 Power Distribution Tier IV Data Center

This research attempts to improve the system availability of a Tier IV data center in terms of power outages by upgrading CIC and SCT levels [17]. A reliable equipment data source is essential. In FMECA, the failure modes of equipment or components data are compiled from IEEE 493 Gold Book Standard Network. Their influences on the status of equipment or components are analyzed on modes of failures,

subsequently followed by the effects of the equipment or component on the status of sub-system reliability.

A reliability block diagram (RBD) demonstrates end-to-end logical connectivity of components and subsystem levels. It depicts the power distribution in a data center from utility sources to loading points. It also shows the system connectivity topology (SCT), such as series-parallel, active standby, bridge, and k out of n design.

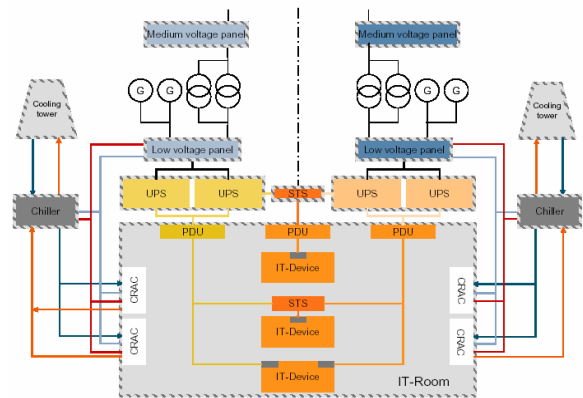


Fig. 3 A Traditional Tier IV Data Center [4]

TABLE I
LEVELS OF DATA CENTER COMPLEXITY

Levels of data center complexity		
Complexity Levels	Definition	Examples
Material	Physical substance, matter	Iron, plastic, copper, glass, silicon, fabric, etc.
Component	A fundamental element of a subsystem that never works alone	Breaker, Cable, Battery, Switch, etc.
Subsystem or assembly	A collection of components and modules combined into one unit and performing a single function of a limited scale	ATS, PDU, MDB, Rack, CRAC piping, Structure Cabling, Transformer, Camera, etc.
System	A complex collection of units, subsystems, and assemblies performing multiple functions	CRAC, Power Distribution, UPS, Generator, Fire Alarm, Water Detection, CCTV, Access control, FM200 etc.
Platform of systems	A single structure used as a base for other installed systems that are serving the platform's mission	NoC, Reception, Staging, Docking, Networking, Communication, Data Center, etc.
Array or system of systems	A large, widespread collection or network of systems functioning to achieve a common mission	Disaster recovery, iDC, Back up Site, Utility, ISP, Telco.

The research proposes a simulation approach applied to an RBD by a "BlockSim 7" simulation software. FMECA investigates how the system detects and recovers from failures and analyzes the failure semantics, and criticality of each failure mode [1]. Fig.4 depicts a consequence of errors and failures. If an error is not dealt with properly, it could result in a system failure. The system failure shall be coped with, detected and corrected within the system tolerant time requirement. This process is of utmost importance. If power to the data center cannot be delivered to load points within milliseconds, this may result in service

disruption. Its consequence means a more serious loss of operations time and revenues.

TABLE II
EQUIPMENT RELIABILITY DATA FROM IEEE 493 GOLD BOOK [12]

Ref. #	Item Description	PREP Item #	Inherent Availability	MTTR (Hours)	Failure Rate Failure/Year	Calculated Availability	MTBF
1	Single Circuit Utility Supply, 1.78 failure/limit years, A=0.999705, Gold	NA	0.999705	1.32	1.956		4,481.60
3	Diesel Engine Generator, Package, Stand-by, 1500KW	98	0.99974231	18.28	0.1235		70,979.76
4	Manual Disconnect Switch	187	0.9999998	1	0.00174		5,037,931.03
5	Fuse, 15kV	117	0.99995363	4	0.10154		86,330.51
7	Transformer, Liquid, Non Forced Air, 3000KVA	208	0.99999937	5	0.00111		7,897,297.30
8	Ckt. Breaker, 600V, Drawout, Normally Open, > 600Amp	68	0.99999874	2	0.00553		1,585,171.79
9	Ckt. Breaker, 600V, Drawout, Normally Closed, > 600Amp	69	0.99999989	0.5	0.00185		4,738,378.38
10	Switchgear, BareBus, 600V	191	0.9999921	7.29	0.00949		923,709.17
11	Ckt. Breaker, 600V, Drawout, Normally Closed, < 600Amp	67	0.99999986	6	0.00021		41,742,857.14
13	Ckt. Breaker, 3 Phase Fixed, Normally Closed, < 600Amp	61	0.99999656	5.8	0.0052		1,685,769.23
20	Cable Aerial, < 15kV, per - 300 feet	32		1.82	0.00268	0.9999994	3,270,895.52
22	Switchgear, Insulated Bus, < 600V		0.99999953	2.4	0.0017	0.9999995	5,156,470.59
60	Cable Below Ground in conduit, < 600V < 300 feet			11.22	0.000603	0.9999992	14,537,313.43
150	Cable, Above Ground, No Conduit, < 600V, per 1000 ft.	20		2.5	0.000096	0.9999997	91,312,500.00

An integrated FMECA and RBD model is shown in Fig. 5. It helps the designer to visualize the SCT of a data center and effects of component failure modes that affect the data center's operational mode before hitting system downtime.

Normally, the data center operation is at *Step 0* meaning that power is supplied from a utility source through ATS, an uninterruptible power supply (UPS), and eventually to load points.

When the utility outage occurs, the failure mode will switch from *Step 0* to *Step 1*, the UPS-battery recharging mode, immediately. A subsequent Gen-Set mode will set a 10-to-20-second delay of ATS to allow for Gen-Set to stabilize its output power. On *Step 2*, back to UPS rectified mode plus Gen-Set mode, the Gen-Set will supply power to recharge the UPS. During this mode a tolerant time depends on the capacities of Gen-Set and UPS at load points. Once the utility resumes its normal status, an ATS will switch the UPS and Gen-Set back to normal mode (*Step 0*).

After conducting an FMECA of each component, the system component simulation by RBD will be performed. The CIC for MTBF and MTTR of each component is considered in determining system reliability. Moreover, the SCT, e.g. series-parallel, active standby, bridge, and k-out of -n will be applied through RBD to enhance data center system reliability, as illustrated in Fig. 6.

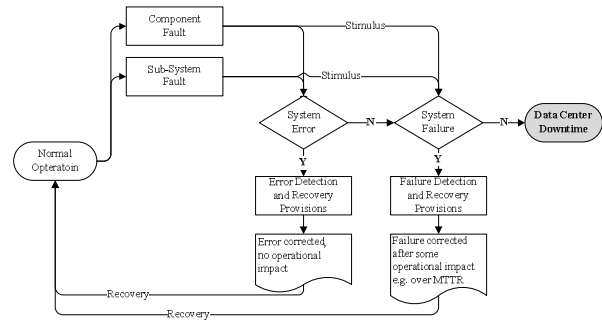


Fig. 4 FMECA System Faults

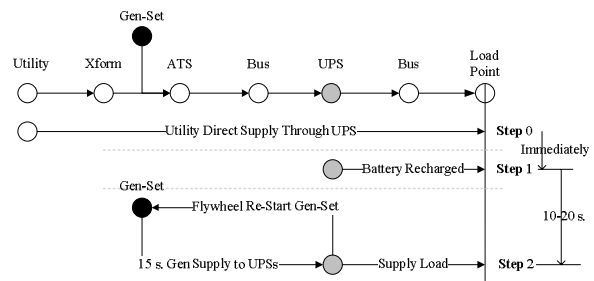


Fig. 5 Power Distribution System Failure Steps

The RBD is used to perform the Monte Carlo simulation of the data center power distribution. It demonstrates the end-to-end logical connectivity of components and subsystem levels. The logical diagram is developed to depict the sequence of power distribution flow of a data center from utility sources through the end load points.

5. Discussion

Poor power grid quality is undoubtedly one of the most significant causes of system downtime [13]. If an estimated availability of the power grid is 99.9%, it is equivalent to roughly 9 hours of downtime per year. However, Tier IV is defined as fault tolerant site infrastructure and site availability of 99.995%.

A fault tolerant Tier IV has redundant system facilities and multiple distribution paths serving the site's IT equipment concurrently [14]. System reliability enhancement of a redundant system can be accomplished through a bridge topology, as shown in Fig.6. Moreover, system reliability may be enhanced further by an active standby on ATS, shown as the third Gen-Set in Fig.6. This SCT model improves the Tier IV site availability well beyond 99.995%, as clearly evidenced from "BlockSim 7" simulation results.

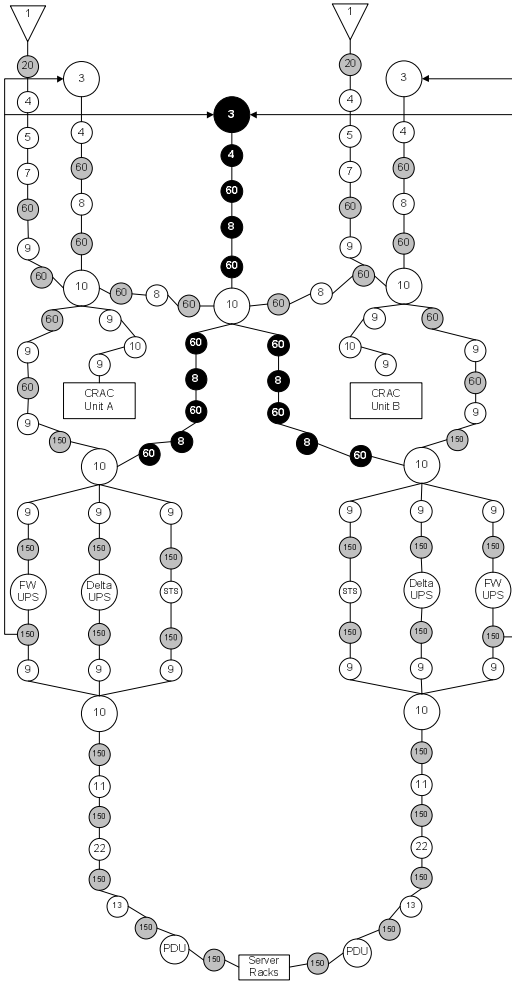


Fig.6 Enhancement of Power Distribution in a Tier IV Data Center

Redundancies on components to increase system reliability are costly. A design engineer needs to avoid over-sizing electrical equipment, since it may waste not only power consumption but also investment.

The ratio of Watt to Volt-Amp, or VA, is called the “power factor” and is expressed either as a per-unit, i.e. 0.85 or a percentage, i.e. 85% [10]. Reliability enhancement at the level of components should consider the efficiency or “power factor.” An advantage of upgrading a traditional UPS from double conversion topology to line interactive topology, or delta conversion, is its increased overall efficiency. The resultant power factor becomes close to unity and the line current distortion is restricted to near 1%. UPS upgrading provides more tolerance to voltage surges and poor power quality incidents to pass on to load points. However, the way in which the UPS mitigates poor power quality incidents often results in battery drainage, which in turn results in reduced battery life [3].

The problems with battery life and voltage surge tolerance can be rectified by a flywheel UPS. Voltage surges result from faults in parts of the electrical power system and can affect the electrical transmission and distribution system [7]. An integrated line interactive UPS coupled with a Flywheel UPS will eliminate weaknesses of each. Voltage sags and temporary voltage loss in a delta UPS will be solved by a Flywheel UPS [6]. Extra time of criticality back-up will be supported by normal batteries from a delta conversion UPS.

The most evident reliability factor for a diesel engine is the failure rate during start. A probability of 0.005, or 5 out of 1,000, attempts to start the engine was failed. To enhance the system reliability, a redundant diesel engine needs to be installed [8]. Re-supply power for Gen-Set or diesel engine from Flywheel UPS is another way to increase the success probability during start.

The result from BlockSim 7, as shown in Fig. 2, shows the MTBF of a traditional power distribution or Tier IV of 75,434.78 hours. Fig. 6 is the improved result after analysis and evaluation by the system failure diagnosis and FMECA of Tier IV, based on Table I. To enhance the system reliability, a redundant diesel engine or Gen-Set is installed. Re-supply power for a diesel engine or a Gen-Set from the Flywheel UPS is applied to the system to increase the success rate during Gen-Set starts. The MTBF of the data center in Fig.6 is 80,021.93 hours, showing an improvement in DC system reliability by 4,587.15 hours.

This result proves that the new simulation *enhances data center system reliability of MTBF around 4,587.15 hours*. During 5 years or 43,824 hours of operation, the data center system failure rate of a traditional power distribution system in Fig.2 is 14.0865×10^{-6} , while that of an enhanced power distribution system in Fig.6 is 12.9159×10^{-6} . An *improved system reliability helps reduce the failure rate of a Tier IV data center by 1.1706×10^{-6} , or only one hour of failure in 1,000,000 hours of operation*.

6. Conclusion

As a result from RBD simulation, component selection (component’s inherent characteristics) and connectivity topology (system connectivity topology) are keys to improving Tier IV data center site availability. A regular maintenance, monitoring, and control system not only prevents an overloaded system but also conserves equipment power consumption. Moreover, it helps extend the system

uptime. An RBD simulation and a real-life data center system application may be totally different. It depends on many factors, e.g. real MTBF of components, actual MTRR, defects during installation, site conditions of data center, testing and commissioning procedures, and human errors. A thorough understanding on the component's inherent characteristics (CIC) and the system connectivity topology (SCT), skills of contractor, experience of consultants and capable project managers all contribute to successful completion of a data center project implementation and achievement of a highly reliable data center system.

7. References

- [1] J. C. Becker, and G. Flick, "A Practical Approach to Failure Mode, Effects, and Criticality Analysis (FMECA) for Computing Systems," IEEE High-Assurance Systems Engineering Workshop, Proceeding, 21-22 Oct 1996. pp. 228-236.
- [2] R. Bono, A. Alexander, A. Dorman, Y.J. Kim, and J. Reisdorf, "Analyzing Reliability – A Simple Yet Rigorous Approach," IEEE Trans. Ind. Applicat., vol.40, no.4, July./Aug. 2004, pp.950-957.
- [3] E. R. Furlong, "UPS Topology for Large Critical Power Systems (>500 KVA)", 13th Annual Power Quality Exhibition & Conference, October 30, 2002.
- [4] R. Hintemann, "Reliable Data Center (DC) Guideline" BITKOM, German Association for Information Technology, Telecommunication and New Media, 2007.
- [5] C. Kara-Zaitri, A. Z. Keller, I. Barody, and P. V. Fleming, "An improved FMEA methodology," IEEE PROCEEDINGS Annual RELIABILITY AND MAINTAINABILITY Symposium, 1991. pp. 248-252.
- [6] R. G. Lawrence, K. L. Craven, and G. D. Nichols, "Flywheel UPS," IEEE Ind. Applicat., Mag., May/June 2003. pp. 44-49.
- [7] C. J. Melhorn, T. D. Davis, and G. E. Beam, "Voltage sags in industrial systems," IEEE Trans. Ind. Applicat., vol 34, May/June 1993. pp. 397-403.
- [8] H. Nordin, and B. Lindemark, "System Reliability, Dimensioning and Environmental Impact of Diesel Engine Generator Sets used in Telecom Applications," IEEE Telecommunication Energy Conference, INTELEC' 99, 6-9 June 1999.
- [9] A. Pievatolo, E. Tironi, and I. Valade, "Sime-Markov Processes for Power System Reliability Assessment With Application to Uninterruptible Power Supply," IEEE Trans. Power Sys., vol. 19, no. 3, August 2004. pp. 1326-1333.
- [10] N. Rasmussen, "Watts and Volt-Amps: Powerful Confusion", White Paper #15, American Power Conversion, www.apc.com, 2006.
- [11] Rausand, M., System Analysis Failure Modes, Effects, and Criticality Analysis, System Reliability Theory (2nd ed), Wiley, 2004
- [12] B. Roczen, R. G. Arno, and P. S. Hale, "RELIABILITY BLOCK DIAGRAM METHODOLOGY APPLIED TO GOLD BOOK STANDARD NETWORK," IEEE Industrial and Commercial Power Systems Technical Conference, 2-6 May 2004, pp. 116-126.
- [13] S. Roy, "RELIABILITY CONSIDERATION FOR DATA CENTERS POWER ARCHITECTURES," IEE, INTELEC 2001, Conference Publication No. 484, 2001.
- [14] W. P. Turner, J. H. Seader, and K. G. Brill, Tier Classification Define Site Infrastructure Performance, The Uptime Istitute, Inc. 2006.
- [15] W. Vahlstrom, "A Practical Guide for Electrical Reliability," EC&M, October 2004, pp. 48-53.
- [16] W. Wang, J. M. Loman, R. G. Arno, E. R. Furlong, and D. Ogden, "Reliability Block Diagram Simulation Techniques Applied to the IEEE Std. 493 Standard Network," IEEE Trans. Ind. Applicat., vol.40, no.3, May/June 2004, pp.887-895.
- [17] M. Wiboonrat, "Power Reliability and Cost Trade-Offs: A Comparative Evaluation between Tier III and Tier IV Data Centers," Power Conversion and Power Management, Digital Power Forum, 10-12 September 2007.
- [18] C. Yun, T. S. Chung, C. W. Yu, C. Y. Chung, Z. Ming, and S. Xin, "Application of Reliability-Centered Stochastic Approach and FMECA to Conditional Maintenance of Electric Power Plants in China," IEEE International Conference on Electric Utility Deregulation, Restructuring and Power Technologies (DRPT 2004), Hong Kong, April 2004. pp. 463-467.