System Reliability of Fault Tolerant Data Center

Montri Wiboonrat
College of Graduate Study in Management
Khon Kaen University, Bangkok Campus
Bangkok, Thailand
montwi@kku.ac.com, mwiboonrat@gmail.com

***Abstract*—A single point of failure (SPOF) in system operations is a weak point of system reliability. Mean time to failure (MTTF) of system operations is equal to the shortage component's MTTF in system. A Tier IV data center is designed to eliminate the SPOF. Data center system reliability is not only depended on the MTTF of each component in the system, but also relies on the mean time to repair (MTTR) of each component. Researcher performed Tier IV DC power distribution systems (PDS) through simulating software, BlockSim7. The research question is tried to investigate how to improve system reliability. Component's inherent characteristic (CIC) and system connectivity topology (SCT) are applied to improve the system reliability of Tier IV data center. The results demonstrated an increasing PDS reliability, site plus site, of Tier IV data center and improving survival probability of system that helps for future improvement on any critical system.**

Keywords—System Reliability; Mean Time To Failure; Mean Time To Repair; Probability density function (pdf).

I. INTRODUCTION

The redundancy represents a possible approach to enhancing system reliability. In a series-parallel design methodology, serial systems reduce reliability, while more parallel systems help increase it. The redundancy scheme helps enhance the overall system reliability. It, however, costs more. A data center consists of multiple hardware components that are bound to fail sooner or later. The Tier IV data center is a fault tolerant system that is designed to eliminate a single point of failure (SPOF) [3]. System downtime adversely affects not only recovery and lost opportunity costs, but also the company's reputation and customers' confidence. The reliability and cost trade-off becomes a controversial issue among all concerned, including top management, IT managers, and financial managers [8]. Different organizations have different levels of recovery time objective (RTO) and recovery point objective (RPO) subject to system failures and power outages to varying degrees of risks [4]. This paper employs the reliability block diagrams (RBD) with reliability information obtained from IEEE 493, or the so-called Gold Book, and component vendors' field test data. The study attempts to improve the data center system reliability by integrating 2 parallel systems of TIA 942's Tier IV data center [3]. The research question comes up with, how the reliability of two parallel systems (PS) is higher than two parallel load-sharing (LS) systems and which once is the most benefits to investor subject to investment, efficiency, and system reliability. Research findings suggest the system connectivity topology, i.e., parallel topology, helps increase the system reliability of DC operations.

II. RELIABILITY BACKGROUND

A. Reliability Factors on Data Center Failures

System failure in data centers may be caused by many sources.

1. Human error: daily operation, regular planned downtime for maintenance, and unplanned downtime [5], [9].
2. Component's inherent characteristics (CIC) and system failures are dependent on mean time to failure (MTTF), the complexity of system connectivity topology (SCT), and operational conditions. A component/system failure may propagate or activate other component/ system fault, error, or failure. This process or failure is similar to a chain-reaction that is affected from component to component, component to system, or system to system [5], [7].
3. Operational conditions are other factors that relates to system failures, e.g., humidity, temperature, altitude, and dust.
4. Natural disasters; this is beyond the control of data center to handle. A design for a parallel site needs to be considered to compensate for downtime losses [8].

B. Reliability Determination

All equipment reliability data is obtained from the IEEE 493 Gold Book Standard, as shown in Table I. Fig. 1 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 I. 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 mean time between failure (MTBF) and mean time to repair (MTTR) of component's inherent characteristics (CIC) distribution, failure modes effects and criticality analysis (FMECA), and environmental effects [4], [7].

ISBN: 978-1-61208-192-2

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).

The following assumptions apply to the proposed Tier IV data center system networks, as shown in Fig. 1:

- 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 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.

III. DATA CENTER MODEL ASSUMPTION

A data center is a complex system that consists of 16 systems. Operations on business continuity strategy mode: primary and secondary sites require 2 data centers operating at the same time to back each other up, i.e., system-of-systems (SoS) or site-plus-site [3]. To limit the scope of this investigation, the researcher focused on data center Tier IV, power distribution systems (PDS) and power distribution system-of-systems (PDSS: both primary and secondary), parallel input, from incoming utility throughout loaded consumptions, e.g. server racks, storage racks, and networking equipment racks. This research assumes the external systems e.g. utility system, communication system or integrated service provider (ISP), are described to operation 100% uptime during this research.

PDS is the most sensitive system for data center downtime. Thus, the research tested fault injection by the cutting of the main utility system as human error or unplanned downtime. UPSs will take action to recharge power back to the system immediately as long as the battery can handle the loaded points. Gen-Set will activate within 15 seconds to change power back to the UPSs to resupply on loaded points [10]. By TIA 942, Tier IV, Gen-Set availability is 96 hours for consecutive operation without interruptions [3].

On this research simulation, the research defines the fault tolerance of data center system design as reliability of 2 parallel system of Tier IV data center. The modification prototype in Fig. 2 is reproduced from original Tier IV data center from Uptime Institute, Fig.1. Each number in Fig. 2 is referred from IEEE 493 Gold Book [7]. Table I refers to characteristic of each number in Fig. 2 that identifies failure rate, MTBF, and MTTR of each number. A failure rate (λ) of Primary data center system and Secondary data center system is assumed equivalent and each system has a constant (λ).

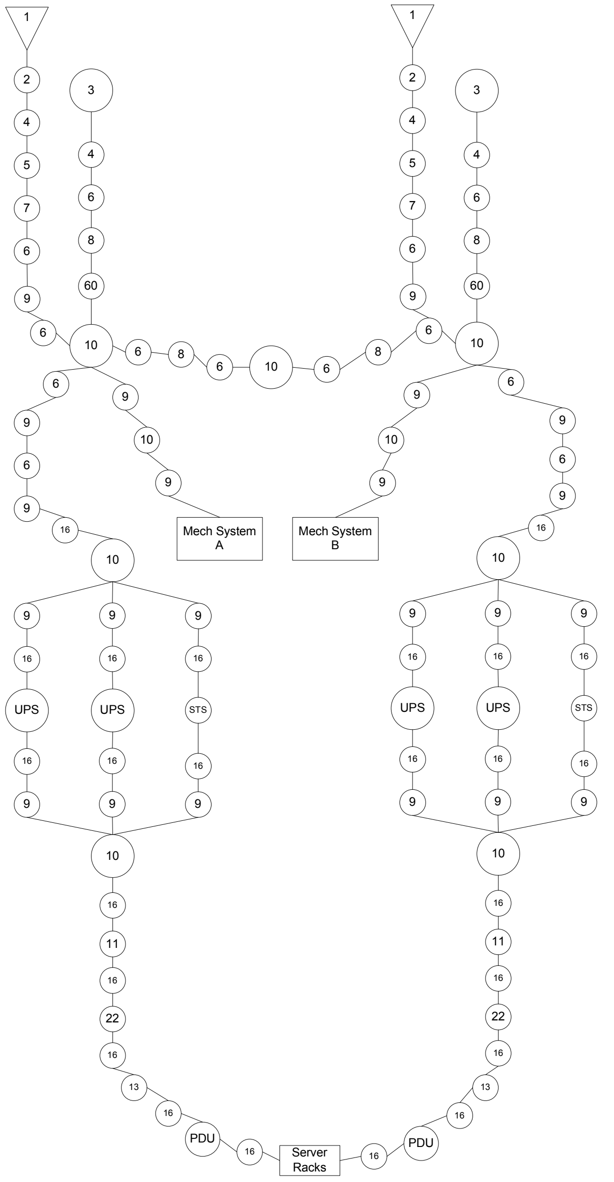


Fig. 2. PDSS of Tier IV Data Center

The reliability of data center system designs for the two parallel systems are reliable in parallel operations only when the failure of both systems results in system operation. On the other hand, for a parallel system to succeed, at least one of the two parallel systems in the whole system needs to perform successfully, or operate without failure, for the operating interval on the intended mission.

The research proposes a simulation approach applied to a reliability block diagram (RBD) by BlickSim 7. The system reliability results from Fig.1 through RBD show the MTBF of Tier IV data center is 75,434.78 hours, and the failure rate (λ) is 14.0865×10^{-6} [5]. The exponential distribution is applied in Tier IV data center reliability analysis. The *f(T)* of the exponential distribution is extremely convenient, often used to describe a steady-state hazard rate, as shown in Fig. 3 [2].

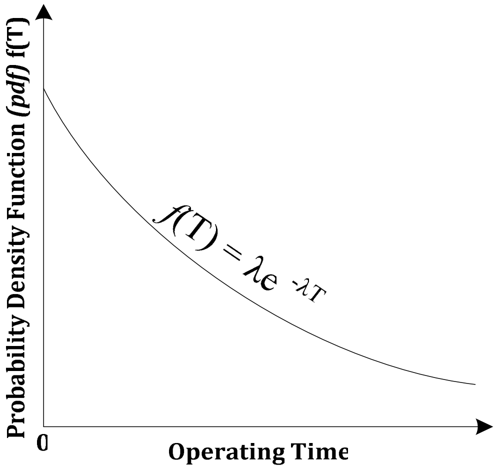


Fig. 3. Relationship of *pdf* and Operation Time [2]

The probability density function *(pdf)* of two parallel systems, that two systems are equal, when $\lambda_1 = \lambda_2 = \lambda$, is given by:

$$f_{SP}(T) = -\frac{d[R_{SP}(T)]}{dT} \tag{1}$$

The failure rate of two parallel systems is given by:

$$\lambda_{PS}(T) = \frac{2\lambda e^{-\lambda T} - 2\lambda e^{-2\lambda T}}{2e^{-\lambda T} - e^{-2\lambda T}} \tag{2}$$

The MTBF of the *PS* is given by:

$$MTBF_{PS} = \frac{1}{\lambda} + \frac{1}{\lambda} - \frac{1}{\lambda + \lambda}, \quad \text{or} \quad MTBF_{PS} = \frac{1.5}{\lambda}, \quad \text{or}$$

$$MTBF_{PS} = 1.5m \tag{3}$$

Where *m* is the MTBF of each DC-PDSS unit. It must be observed that for parallel systems even with data center PDSS units that have a constant failure rate (λ).

$$\lambda_{PS}(T) \neq \frac{1}{MTBF_{PS}} \tag{4}$$

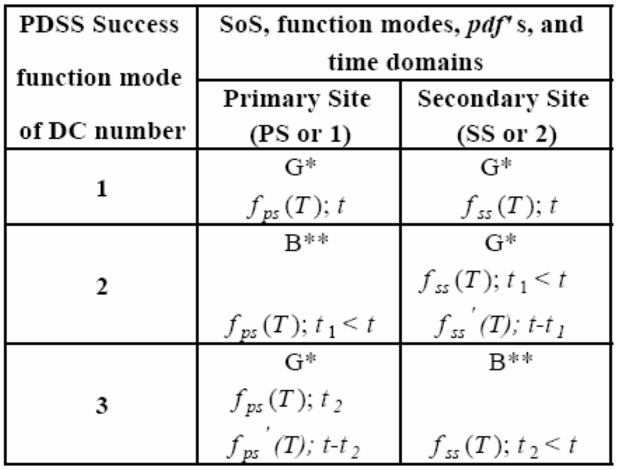
The precise determination of the SoS reliability, the change of failure rate, and PDSS reliability of the surviving system need to be properly taken into account. The reliability of the SoS of two load-sharing exponential units shown in Fig. 4 is given by Kececioglu (1991) as follows:

R(t): Probability that Primary site (PS or 1) and Secondary site (SS or 2) complete their mission successful with *pdf*'s $f_1(T)$ and $f_2(T)$, respectively, or the probability that PS fails at $t_1 < t$ with *pdf* $f_1(T)$, and SS functions till t_1 with *pdf* $f_2(T)$ and the functions for the rest of the mission, on in $(t - t_1)$, with *pdf* $f_{2'}(T)$, or the probability that SS fails at $t_2 < t$ with *pdf* $f_2(T)$ and PS functions till t_2 with *pdf* $f_1(T)$ and then functions with *pdf* $f_{1'}(T)$ in $(t - t_2)$.

This scenario is presented in Table II and depicted in *pdf* form in Fig. 4. when both of PDSSs are exponential. The system model of integration of two equal system failure rates and SCT as parallel is illustrated in Fig. 4. Fig. 4 (a) and (b) shows the overall failure rate of system integration decrease under the parallel topology concept [1]. Fig. 4 (a) is illustrated chronologically and shows sequence reaction for the total system when Primary data center fails at time t_1 after starting the operation period. The indicator of changing factor is transformed by $\lambda_2 \Rightarrow \lambda_{2'}$ on Secondary data center. And, vice versa, the changing on Fig. 4 (b) is transformed by $\lambda_1 \Rightarrow \lambda_{1'}$ on Primary data center when Secondary data center failure during time t_1 of operation period, when $\lambda_{1'} = \lambda_{2'} = \lambda'$, is given.

TABLE II

MATRIX OF SOS, FUNCTION MODES, PDF'S AND TIME DOMAINS FOR ALL PDSS SUCCESS FUNCTION MODES [1]



PDSS Success function mode of DC number	SoS, function modes, *pdf*'s, and time domains	
	Primary Site (PS or 1)	Secondary Site (SS or 2)
1	G* $f_{ps}(T); t$	G* $f_{ss}(T); t$
2	B** $f_{ps}(T); t_1 < t$	G* $f_{ss}(T); t_1 < t$ $f_{ss}'(T); t-t_1$
3	G* $f_{ps}(T); t_2$ $f_{ps}'(T); t-t_2$	B** $f_{ss}(T); t_2 < t$

G*: SoS is good throughout the designed mission.
B**: SoS fails before the designed mission

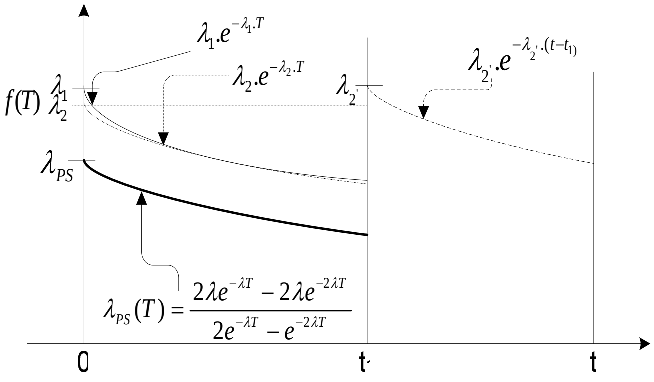
This is age and mission dependent even through the units are exponential. The first age mission $t = T$. Mathematically, the reliability of the parallel system (*PS*) is given by:

$$R_{PS}(T) = 2e^{-\lambda T} - e^{-2\lambda T} \tag{5}$$

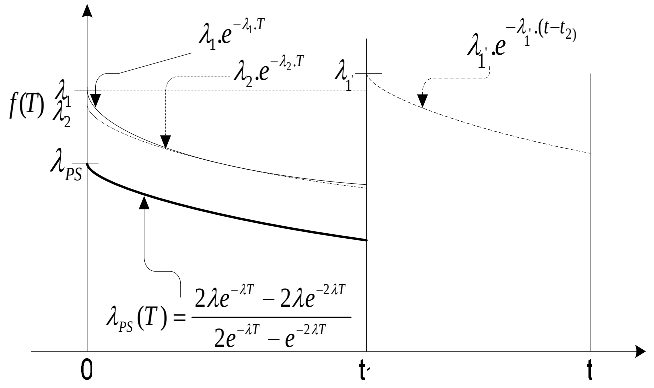
It must be observed that $\lambda_{PS}(T)$ is not constant, but a function of age, although each system of data center has a constant $\lambda_1 = \lambda_2 = \lambda$ and $\lambda_{1'} = \lambda_{2'} = \lambda'$.

System integration and operations of site plus site, normal operation of data center PDSS, of Tier IV data center is shown on Fig. 5. Utility is going to supply power throughout the PDS: Transformer, automatic transfer switch (ATS), bus-bar, UPSs, bus-bar, and loaded points. Calculation of SoS or parallel system (*PS*) of Tier IV data center is derived from Kececioglu (1991). Kececioglu defines $MTBF_{PS}$ of two systems equal to 1.5*m*, as referred to, in (3). Substitution $\lambda = 14.0865\text{x}10^{-6}$ and $T = 43{,}800$ hours of Tier IV data center to (2). The result is $\lambda_{PS}(T) = 8.88215\text{x}10^{-6}$. Substitution $\lambda_{PS}(T)$ (1), we will receive $MTBF_{PS} = 112{,}585.371417$ hours. The result from substitution $\lambda_{PS}(T) = 8.88215\text{x}10^{-6}$ and $T = 43{,}800$ hours to (5) shows the reliability *R(T)* of parallel system of Tier IV data center, SoS, during 5 years equal to *89.6128152%*.

ISBN: 978-1-61208-192-2



(a). A pdf's of two parallel DC systems when primary DC fails



(b). A pdf's of two parallel DC systems when secondary DC fails

Fig. 4. Comparing between *pdf* and operation time of DC-PDSS load-sharing system [1]

The assumption on Fig. 4 shows $\lambda_{PS} < \lambda$, we need to prove $\lambda < \lambda'$ after Primary site failover to Secondary site, as shown in Fig. 5. Since in (4) derived results contrast with (1). We are given:

$MTBF_{1_t\, or\, 2_t}$ is MTBF of each data center PDSS system before failure of Primary or Secondary site,

$MTBF_{1'_t\, or\, 2'_t}$ is MTBF of each data center PDSS after failure of Primary or Secondary site,

λ is failure rate of each data center PDSS before failure,

λ' is failure rate of each data center PDSS after failure.

$$\lambda_{PS} < \lambda < \lambda' \text{ and } MTBF_{PS} > MTBF_{1_t\, or\, 2_t} > MTBF_{1'_t\, or\, 2'_t} \quad (6)$$

Now we have $MTBF_{PS}$ = 112,585.371417 hours and substitute to (3). We received m = 75,056.91 hours. When we compare with original MTBF from Tier IV data center that equal to 75,434.78 hours, that means $MTBF_{1_t\, or\, 2_t} > MTBF_{1'_t\, or\, 2'_t}$ and vice versa $\lambda < \lambda'$ as well for (1), as depicted in Fig. 4. Hence the assumption on (6) is correct.

$$R_{LS}(t) = e^{-2\lambda t} + \frac{2\lambda}{\lambda' - 2\lambda}(e^{-2\lambda t} - e^{-\lambda' t}) \quad (7)$$

For the two parallel load-sharing (*LS*) of data center PDSS's, if the data center PDSS's are equal, i.e., they have the same *pdf*, then $\lambda_1 = \lambda_2 = \lambda$ and $\lambda_{1'} = \lambda_{2'} = \lambda'$, we will derive (7) at t = 5 years, as illustrated in Fig. 4. The reliability result is *85.3775%*.

When one data center PDSS fails before the mission is completed in LS, and the data center PDSS is exponential. The reliability of the existing Secondary data center PDSS when Primary site fails calculates from (8), as shown in Fig. 3(a).

$$R_{2'} = (t - t_1) = \frac{R_{2'}(t)}{R_{2'}(t_1)} = \frac{e^{\lambda'_2 t}}{e^{\lambda'_2 t_1}} = e^{-\lambda'_1(t-t_1)} \quad (8)$$

Assumption, the Primary site fail at time = t_1, what is the reliability of the Secondary site during the left 3 year operation? Substitution t, t_2, and λ'; from previous calculation, and t= 43,800, t_1 = 17,520 hours to (8). We will get $R_{2'}(t) = 70.4593882\%$.

Normal operation of data center PDSS is depicted in Fig. 5 or on Primary site on left hand and Secondary site on the right hand. The utilities supply power throughout the PDSS: aerial cable, fuse, transformer, cable, automatic transfer switch (ATS), bus-duct, circuit breaker, cable, UPS, circuit breaker, cable, PDU, cable, and loaded points. After utility outage or PDSS fails on Primary or Secondary site, normal operation is degraded, when either of both sites is resume accomplishing the system will resume to normal operation, as illustrated in Fig. 6. Whenever, data center of Primary or Secondary site failure the services for external operations, end users, will not interrupt. The capacity to handle the transactions will reduce a half or increase double waiting queue or time for executing process. Absolute failure in case can be happened only when both sites are completely destroyed or malfunction at the same time [12].

IV. DISCUSSION

A reliability of the parallel Tier IV data center system for a critical mission of T duration before the first failure or 5 years on this simulation is *89.6128152%*. This *89.6128152%* implies that the probability of survival on normal operation will still perform function continuously during mission time over 5 years or *10.3871848%* of the probability of the parallel Tier IV data center system which may fail before 43,800 hours of mission operation. The reliability of two parallel systems (PS) is higher than two parallel load-sharing (LS) systems; conversely the efficiency of LS system is better than the normal two parallel systems in terms of distributing transaction, service response times, and expanding capacity and capability.

The research defines the Tier IV DC parallel system failure when both Primary data center and Secondary data center fail at the same time. MTTR and maintenance systems of each data center are the keys to keeping parallel systems more reliable and available. The critical condition, which is MTTR, must be less than the maximum tolerable period of disruption (MTPD) [11]. MTTF of equipment is depended on 4 conditions; first, selected on CICs and designed on SCT of a data center system, second, operation of data center site conditions, third, related risks on daily operation of human activities, and last, procedural maintenance [5]. The most reasons of system failures come from human omission, e.g., they do not follow the manual instruction step by step, or commission, e.g., the system crashes during installing system.

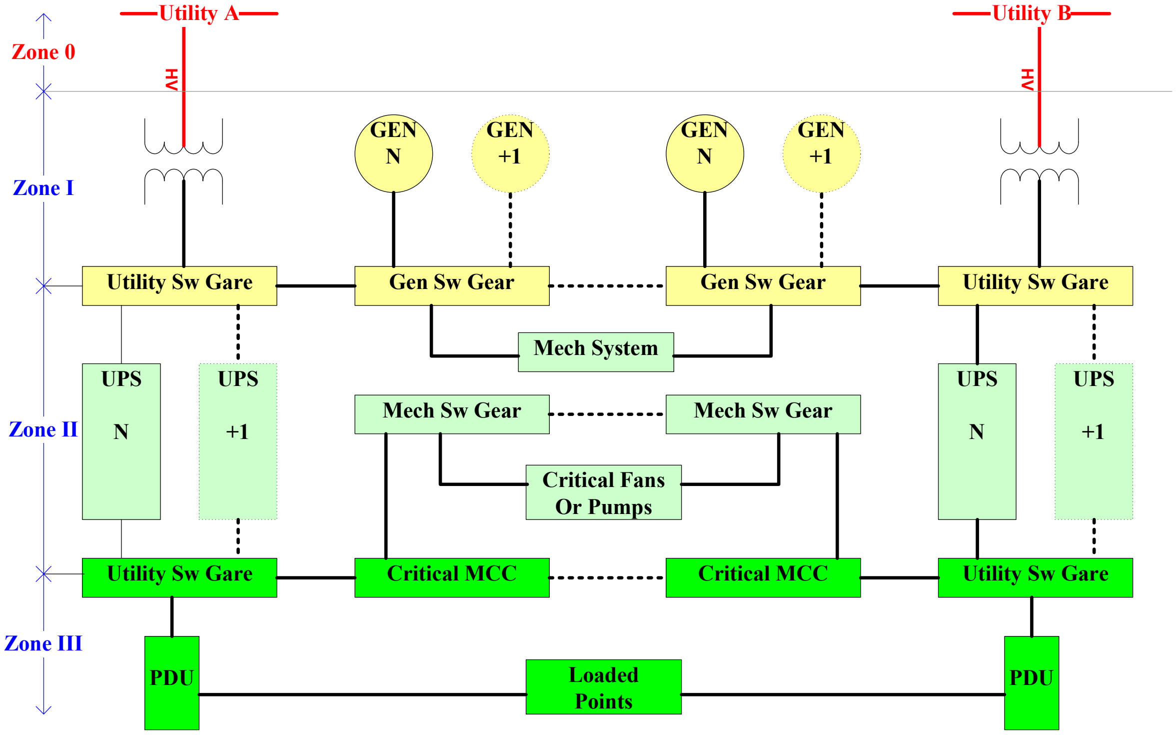


Fig. 1. Original Tier IV Data Center Diagram [3]

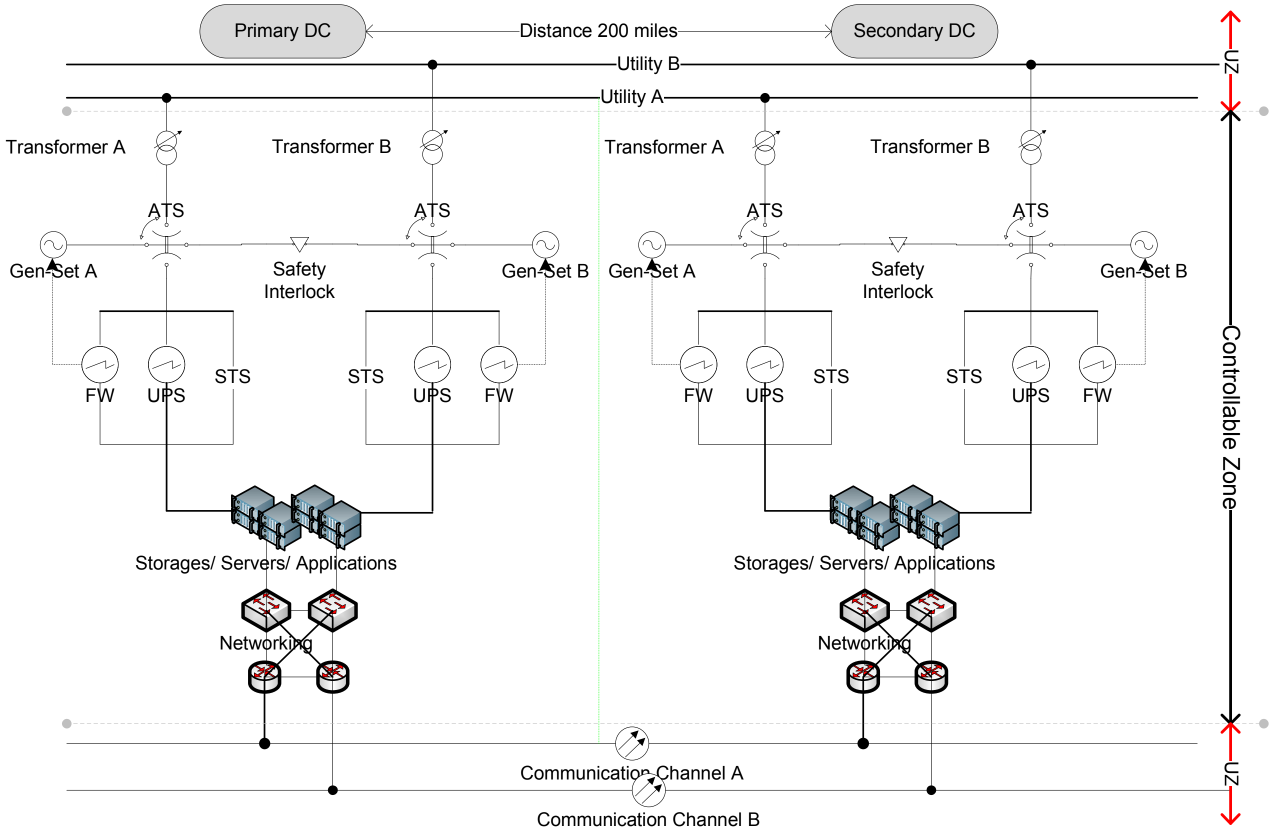


Fig. 5. Parallel System Design of Site plus Site Tier IV Data Center

TABLE I
EQUIPMENT RELIABILITY DATA FROM IEEE 493 GOLD BOOK [7]

Ref. #	Item Description	PREP Item #	Inherent Availability	MTTR (Hours)	Failure Rate Failure/Year	Calculated Availability	MTBF
1	Single Circuit Utility Supply, 1.78 failure/unit years, A=0.999705, Gold Book p.107	NA	0.999705	1.32	1.956		4,481.60
2	Cable Arial, ≤ 15kV, per - mile	32	0.99999022	1.82	0.04717		185,838.46
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
6	Cable Below Ground in conduit, ≤ 600V - 1000 feet	47	0.99999743	11.22	0.00201		4,361,194.03
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
12	Ckt. Breaker, 600V, Normally Closed, < 600 Amp, Gold Book p.40	63	0.99998948	9.6	0.0096		913,125.00
13	Ckt. Breaker, 3 Phase Fixed, Normally Closed, ≤ 600Amp	61	0.99999656	5.8	0.0052		1,685,769.23
14	Ckt. Breaker, 3 Phase Fixed, Normally Open, > 600Amp	62	0.99998532	37.5	0.00343		2,555,685.13
15	Cable, Above Ground, No Conduit, ≤ 600V, per 1000 ft.	20	0.99999997	2.5	0.00012		73,050,000.00
16	Cable, Above Ground, Trays, ≤ 600V, per 1000 ft, Gold Book p.105		0.99999831	10.5	0.00141		6,217,021.28
20	Cable Arial, ≤ 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
26	Bus Duct, Gold Book p. 206, per Circuit foot		0.99981596	12.9	0.000125	0.999816	70,080.18
60	Cable Below Ground in conduit, ≤ 600V - 300 feet			11.22	0.000603	0.9999992	14,537,313.43
80	Ckt. Breaker, 600V, Drawout, Normally Open, > 600Amp	68		2	0.002765	0.9999994	3,170,343.58
90	Ckt. Breaker, 600V, Drawout, Normally Closed, > 600Amp	69		0.5	0.000925	0.9999999	9,476,756.76
110	Ckt. Breaker, 600V, Drawout, Normally Closed, < 600Amp	67		6	0.000105	0.9999999	83,485,714.29
120	Ckt. Breaker, 600V, Normally Closed, > 600 Amp, Gold Book p.40	63		9.6	0.0048	0.9999947	1,826,250.00
130	Ckt. Breaker, 3 Phase Fixed, Nomally Closed ≤ 600 Amp, Gold Book p. 40	61		5.8	0.0026	0.9999983	3,371,538.46
140	Ckt. Breaker, 3 Phase Fixed, Nomally Open, > 600 Amp	62		37.5	0.001715	0.9999927	5,111,370.26
150	Cable, Above Ground, No Conduit, ≤ 600V, per 1000 ft.	20		2.5	0.000096	0.9999997	91,312,500.00

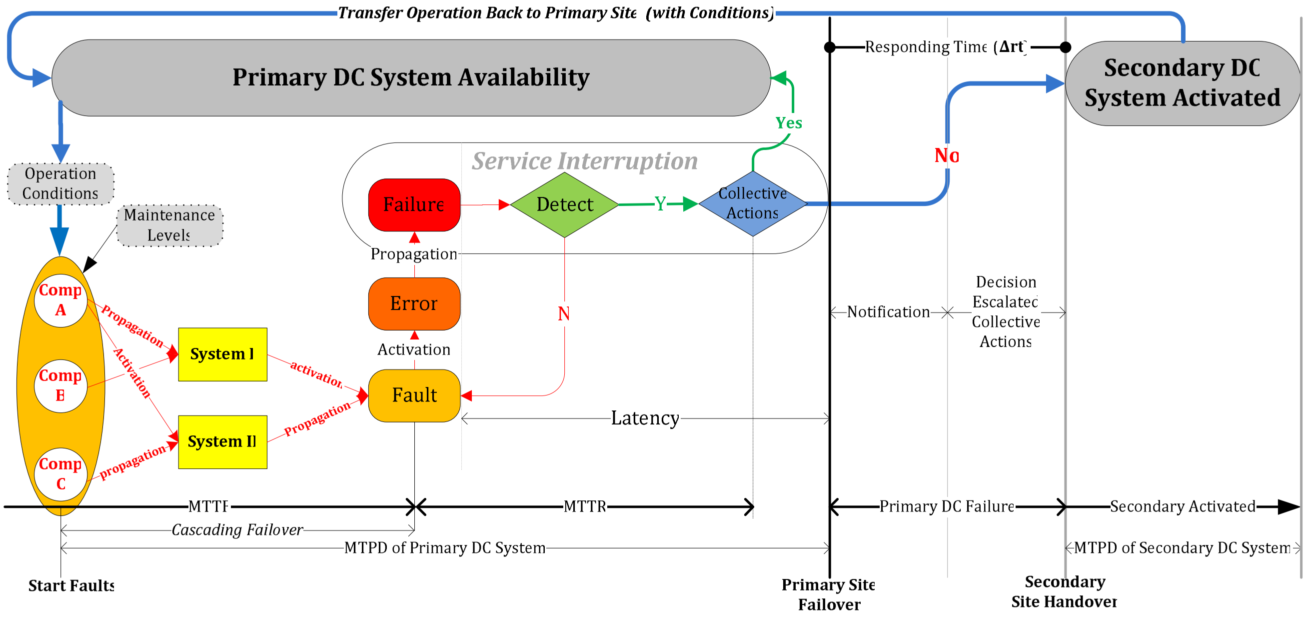


Fig. 6. System Failure Life Cycle Model.

During the design process, engineers, consultants, and designers need to understand throughout the transformation of a system failure cycle. The research results are derived from the root cause analysis of each system failure cycle to prepare the preventive actions and pre-planning for the corrective actions.

V. CONCLUSION

A simulation results from reliability block diagram (RBD) helps consultants, data center designers, IT managers, and contractors to foreseen the output of reliability system design. This helps to save time and costs from trial and error processes which in real data center operations cannot be accepted. To improve the system reliability data center designer needs to understand the MTTF, MTTR, CIC, and SCT of each type design pattern to optimize between reliability and investment. As a result from equation (6),

$$\lambda_{PS} < \lambda < \lambda' \text{ and } MTBF_{PS} > MTBF_{1_t\, or\, 2_t} > MTBF_{1'_t\, or\, 2'_t}$$

is shown the MTBF of each $\lambda_{PS} < \lambda < \lambda'$ presented reliability MTBF; *89.6128152%> 85.3775% > 70.4593882%* respectively, as in equation (6). This is implied that the reliability of two parallel systems (PS) is higher than two parallel load-sharing (LS) systems.

A regular maintenance, monitoring, and automatically control system is not only preventive and alert the system before disaster occurs but also help increase system reliability, system operations on energy conservative mode, and extending operations life-cycle of all equipments. All of these key factors are contributing to data center project operations success.

REFERENCES

[1] D. Kececioglu, “Reliability Engineering Handbook,” Volume 2, Prentice Hall, Englewood, Cliffs, New Jersey, 1991.

[2] D. J. Klinger, Y. Nakada, and M. A. Menendez, “AT&T Reliability Manual,” AT&T, Van Nostrand Reinhold, New York, 1st edition, February 3, 1999.

[3] Turner IV, W. P., J. H. Seader, V. Renaud, and K. G. Brill, Tier Classification Define Site Infrastructure Performance, White Paper, The Uptime Institute, Inc. 2008.

[4] M. Wiboonrat, “An Empirical IT Contingency Planning Model for Disaster Recovery Strategy Selection,” IEEE International Engineering Management Conference, June, 28 to 30, 2008, Estoril, Portugal, accepted for publication.

[5] M. Wiboonrat, “An Empirical Study on Data Center System Failure Diagnosis,” 3rd International Conference on Internet Monitoring and Protection, IEEE ICIMP 2008, Romania, June 29-July 5, 2008, pp. 103-108.

[6] M. Wiboonrat, “An Optimal Data Center Availability and Investment Trade-Offs,” 9th International Conference on Software Engineering, Artificial Intelligence, Networking, and Parallel/ Distributed Computing, IEEE SNPD 2008, Thailand, August 6-8, 2008, accepted for publication.

[7] M. Wiboonrat, “Beyond Data Center Tier IV Reliability Enhancement,” Power Conversion and Power Management, Digital Power Europe, November 13-15, Munich, Germany, 2007.

[8] 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, San Francisco.

[9] M. Wiboonrat, and C. Jungthirapanich, “A Taxonomy of Causal Factors and Success Criteria on Project Management Success and Project Success,” Eight Int. Conference on Opers. & Quant. Management, ICOQM 2007, October 17-20 2007.

[10] M. Wiboonrat, “Dependability Analysis of Data Center Tier III,” 13th International Telecommunications Network Strategy and Planning Symposium, Networks 2008 in Budapest, Hungary, September 28 – October 2, 2008, accepted for publication.

[11] IEEE Std. 1100-1999, (Revision of IEEE Std. 1100-1992), IEEE Recommendation Practice for Powering and Grounding Electronic Equipment, 22 March 1999.

[12] M. Wiboonrat, “Transformation of System Failure Life Cycle,” International Journal of Management Science and Engineering Management, Vol. 4, No. 2, 2008, pp. 143-152.