Mobile and Modular Data Center

Montri Wiboonrat

College of Graduate Study in Management Khon Kaen University, Bangkok Campus

25 Bangkok Insurance Bld 25thF South Sathon Rd. Bangkok 10210 Thailand

Tel. 081-1738200 Fax 02-6774145

E-mail: montwi@kku.ac.th, mwiboonrat@gmail.com

Abstract

The evolution of technologies in the 21st century are driven the next generation data center design. The new design must answer business needs such as less time deployment, high efficiency and effectiveness, and less investment. This research paper is concentrated on modified feature of mobility and modularity through data center without adversely impact of availability, capacity, efficiency, performance, and security function of data center. The case study of 21 data centers consolidation is investigated and formulated container concept of mobile and modular data center (M^2DC). The concept of M^2DC benefits to help organization reduce total cost of owner ship (TCO), project deployment, business risks, and increase equipment and facility utilization, and energy and system efficiency.

1. Introduction

The neo society in which we live is facing an "intelligence" era. With the continuing development and evolution of ICT comes a hastily escalating amount of transactions and data, and as a result the data centers responsible for creating, processing, storing, exchanging, and distributing these data have become a vital part of business and social infrastructure.

On the other hand, with the increasing amount of data, power consumption has also continued to exaggerate, and energy conservation has become a new business and social intention from the perspective of environmental-considerations. Operating the data center will become more and more complicated in the future as the requirement to available, flexible, reliable, and secured handle unpredictable vacillations in the amount of data.

CIO needs more smart solution for next generation data center for instance consolidation and virtualization. The challenge for CIO is how to minimize the total cost of ownership (TCO) for the data center infrastructure without unfavorably impacting the availability, capacity, efficiency, performance, and security of critical systems [7]. Data center consolidation is not only saving the present investment by sharing basic facility infrastructure and utilized resource to meet maximum efficiency of equipment design but also saving the future costs of electrical billing, operations management, and maintenance.

The research question came up after December 2011, since the World Bank has estimated 1,425 billion Baht in economic damages and losses due to flooding, as of 1 December 2011. Many of businesses, especially banking, telco, and retail segments, had suffered from this crisis. A ton of data cannot recover and destroy during flooding. Legacy data center must shut down operations and transfer to disaster site. What will be happened if they have only one data center? They need to transfer data or servers to other site and start operations. Sometime without facility infrastructure, it took a few weeks or a few months to recover business. This is not included business losses during data center downtime.

RECEIVED 2 January, 2013

ACCEPTED 28 June, 2013

Researcher which encourages the smart community practically urbanized the mobile and modular data center (M^2DC) which can solve various problems such as mobility, availability, security, short time implementation, reducing investment, space utilization, increasing energy efficiency and system reliability. The case study of consolidation 21 data centers was deploying the advantage concepts of mobile, modular, and consolidation to create M^2DC prototype. The research results propose comprehensive and miscellaneous solutions of M^2DC in order to meet the business needs.

1.1 Downtime

Resiliency is among the most common explanations for data center investments. Obsolete, low-quality data centers are often an unacceptable business risk. Nevertheless upgrading facilities typically isn't always the most direct or even the most effective means of making applications more available. Availability is a critical facet of data center environment. Availability is the measure of how often you data and application are ready for you to access them when you need them. At the margin, investments to improve system designs and operations may consent better returns than investments in physical facilities [12]. Downtime devastatingly begets from application and system failures, not facility outages. The Table 1 shows the amount of downtime expected for different levels of availability.

Table 1 Correlation between availability and downtime per year

Level of availability	Downtime per year
90%	36.5 days
95%	18.25 days
99%	3.65 days
99.9%	8.76 hours
99.99%	50 minutes
99.999%	5 minutes

1.2 Investment

Investments in data center capacity are a fact of business life. Businesses require new applications to interact with customers, manage supply chains, process transactions, and analyze market trends. Those applications and the data they use must be hosted in secure, mission-critical facilities. To date, the largest enterprises have needed their own facilities for their most important applications and data. How much data center capacity you need and when you need it, however, depends not only on the underlying growth of the business but also on a range of decisions about business projects, application architectures, and system designs spread out across many organizations that don't always take data center capital into account. As a result, it's easy to build too much or to build the wrong type of capacity.

Data centers are suitable for organizational investors for long-term investment. As data center is a mercantile real estate facility used to host computer systems and networking components. Data centers vary from typical business office and industrial buildings because they are designed to accomplish the mission critical by provide reliable facility infrastructure such as power, cooling, and network communication equipment.

A new model for managing data center amalgamates factory-style productivity to keep costs down and short time implemented with a more agile, innovation-focused approach to adapt to rapid change. Data center infrastructure managers believe that constant innovation represents their only chance of meeting user expectations and business demands within budgetary constraints. Data center innovation is a critical part of the direction for enterprise infrastructure functions.

2. Background

Most data center design is based on TIA 942: Telecommunications Infrastructure Standard for Data Centers, 15 April 2005 and Uptime Institute standards which the common references as Tier 1, Tier 2, Tier 3 and Tier 4 [1]. They is a new standard for data center best practice called BICSI just announce in June 21, 2010 [13].

Table 2 Compare system standard of Tiers

	Tier 1	Tier 2	Tier 3	Tier 4
Performance				
Availability (%)	99.671	99.749	99.982	99.995
Annual downtime (Hours)	28.8	22	1.6	0.4
Months to implement	3	3-6	15-20	15-20
Architecture				
Power source	Single system	Single system	Dual system	Dual system
Component redundancy	N+1	N+1	N+1	N+1
Power distribution paths	1	1	1 Active + 1 Standby	2 Active
Compartmentalisation	No	No	Yes	Yes
Concurrently maintainable	No	No	Yes	Yes
Fault tolerance to a single point of failure	No	No	Yes	Yes

This standard are covered, requirements, recommendations and additional information that should be considered when planning and building a data center e.g. site selection, layout analysis, thermal systems, power systems and security. (In this research we are discussion only Tier 2 and Tier 3)

Tier is categorized by system reliability standard, as seen in Table 2. Reliability is accomplished by creating redundancy, to prevent a single point of failure, in power supplies, backup power supplies, fiber optic communication connections, environmental controls, and security devices.

The single line diagram of power distribution system of Tier 2 and Tier 3 is shown in Figure 1 and 2 respectively [2].

2.1 Power usage effectiveness (PUE)

A Power Usage Effectiveness (PUE) is measurement metric for data centers in order to benchmark and compare energy consumption between data centers (Information Technology: IT) and to baseline consumption of facilitated infrastructure against, as shown in Figure 3., which improvements in energy efficiency may be measured [3].

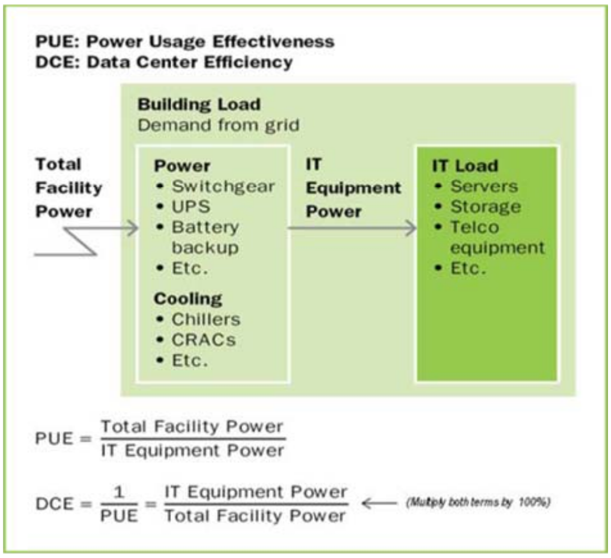

Figure 3 How to measure PUE

Most data centers operate at PUE ≥ 2 and some at PUE > 3. PUE = 3 means that for every watt consumed by the computers, 2 watts are being consumed by data centre (or computer room) air-conditioning (CRAC), lighting and power distributed system. The goal is thus to get PUE to approach 1.

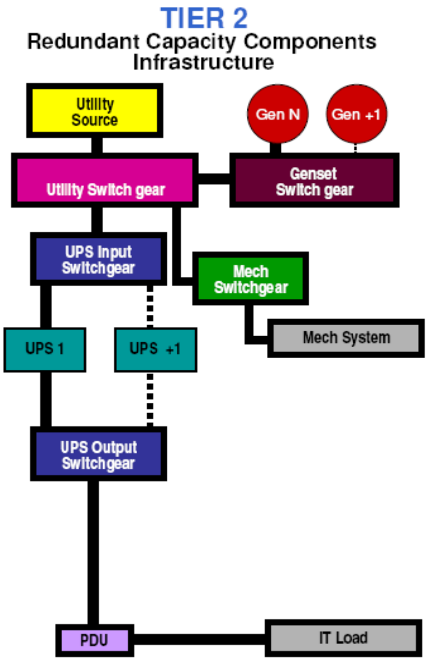

Figure 1 Tier 2 Data center

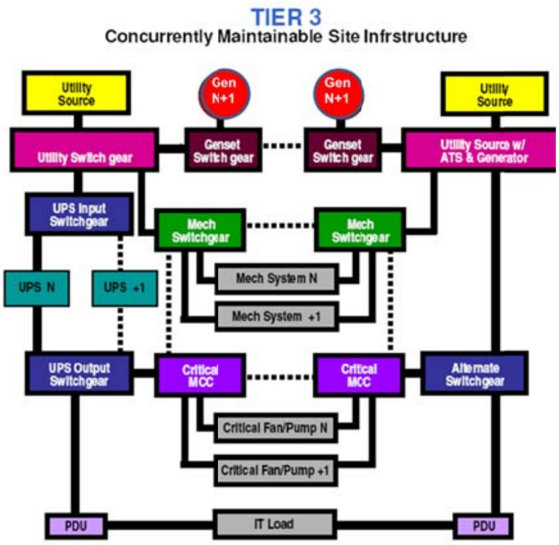

Figure 2 Tier 3 Data center

2.2 New technology vs. space limitation

The challenge of design data center has been changed since 2010. The evolution of power distribution, computer room air condition (CRAC), and network communication systems have more efficiency, smaller footprint, simple installation, and high reliability. The new challenge of design next generation data center in side 20 feet or 40 feet container is space limitation. How to optimize and utilize space is the next challenge of M^{2}DC. Table 3 is shown the general dimension design and loaded container size 20 feet. The interior conceptual design layout of M^{2}DC is demonstrated in Figure 4.

Table 3 Limitation of 20 feet container

General Purpose Containers					
External	20′	40′	Internal	20′	40′
Length	6.06m	12.19m	Length	5.90m	12.04m
Width	2.44m	2.44m	Width	2.35m	2.35m
Height	2.59m	2.59m	Height	2.39m	2.39m
Tare Weight	2,220kg	3,660kg	Max. Payload	28,200kg	26,820kg
Capacity	33.20m^3	67.70m^3	Dimensions & weight may differ marginally depending on manufacturer.		

Figure 4 Conceptual design of M^{2}DC

2.3 Legacy power design problems

The outset power capacity design is equal to the ultimate installed power capacity which is 100% built from the beginning. At the beginning of actual load will start at the estimated startup power requirement around 30-40% and rise up to the expected power requirement, which is classically equal to the design power capacity. Nevertheless, the actual startup power requirement is characteristically lower than the estimated startup power requirement. The power requirement may take a few years to reach expected power requirement or may not, which is considerable less than the design power capacity, as depicted in Figure 5.

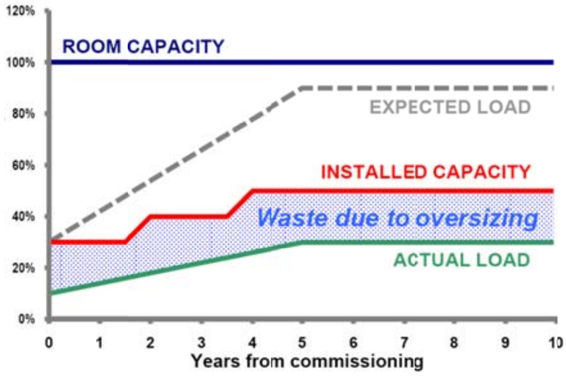

Figure 5 Legacy power capacity design

The surplus capacity design translates directly to surplus investment. In additional investment correlated with power distribution system, electrical billing, maintenance costs, and system efficiency. The average legacy data center is eventually oversized by three times in design capacity, and over four times in nameplate, manufacturer listing specific information [6].

2.4 Legacy cooling design problems

Thailand has a tropical monsoon climate and the highest average temperature of any country in the world. Temperatures in Thailand regularly stay well above 30°C throughout the year. Therefore, data center design in Thailand needs more energy for heat rejection systems.

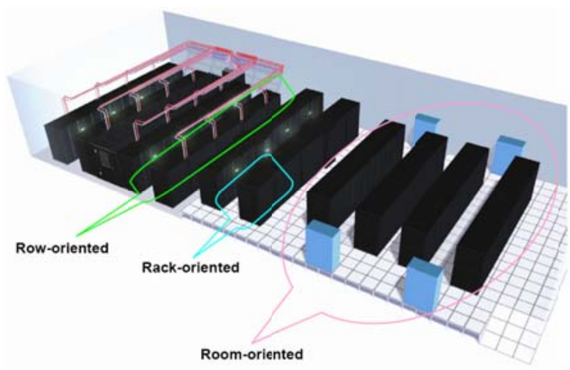

Figure 6 Layout of room row rack design

Room-based cooling is the legacy method for accomplishing data center cooling. Information and experience demonstrate that this traditional cooling system is effective when the average power density in data center is less then 10kW per rack [5]. The layout deign is presented in Figure 6. The legacy design capacity considers from each rack, row and room.

The rack-oriented cooling is shown the consistently low of electrical costs, because the CRAC units are closely coupled to the load, and right sized to the load [4]. However, this efficiency will not perform when the room size is over designed capacity of CRAC units and/or mixing among different cooling rack loaded e.g. put 10kW racks with 5kW racks. This will make disfigured flow of thermal distribution. Unnecessary of traveling airflow distance is avoided.

The effect of power density and annual electrical costs of the three cooling architectures demonstrates in Figure 7.

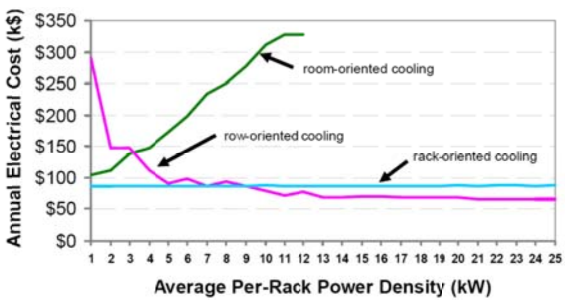

Figure 7 Room row rack oriented cooling

3. Research methodology

The research is separated into three parts of investigation and assessment the existing data centers. First, 34 data centers have been surveys subject to: number of applications; number of servers, storages, network equipments; sizing of data center; and loaded power consumption. Second, researcher was conducted direct interviews with senior or project managers. Last, product and technology assessment are to define the ideas, assess the development potential and establish the individuality.

The exploration research proposes are trying to find out more information to answer about the utilization of space, loaded power, and cooling sizing. How to optimize and utilize by applying the latest technologies for improve energy efficiency and effectiveness are expected results of this research.

4. Analysis and solution

The survey result from 34 agencies, as shown in Table 4, is intended to analysis each data center subject to space, server, and application utilization. Only 21 agencies have data centers and total data center space is 1,563 sqm. Duplication in date center investment, operations, and management is costly. Data center consolidation and optimization seem to be the best option to answer the economy and efficiency of business expectations.

Table 4 34 data centers survey

34 Agencies: Utilization Survey						
No.	**Agency**	**Sizing Data Center (sqm.)**	**No. of Servers**	**No. of Networks**	**No. of Applications**	**Ulilization Level**
1	Agency 1	152	1	1	1	3
2	Agency 2	756	10	25	35	2
3	Agency 3	-	9	-	N/A	N/A
4	Agency 4	Shared with other			N/A	N/A
5	Agency 5		17	9	10	N/A
6	Agency 6	44	4	7	4	2
7	Agency 7	Hosting outside			N/A	N/A
8	Agency 8	Hosting outside			N/A	N/A
9	Agency 9	16	1	2	1	3
10	Agency 10	54	1	2	1	3
11	Agency 11	-	3	5	1	N/A
12	Agency 12	75.65	15	12	9	2
13	Agency 13	62.4	89	38	47	1
14	Agency 14	30	39		N/A	1
15	Agency 15	No Data Center			N/A	N/A
16	Agency 16	Hosting outside			N/A	N/A
17	Agency 17	Hosting outside			N/A	N/A
18	Agency 18	36	5	3	N/A	3
19	Agency 19	31.5	23	23	15	2
20	Agency 20	31.5	13	5	8	2
21	Agency 21	9	1	1	1	3
22	Agency 22	Hosting outside			N/A	N/A
23	Agency 23	No respond			N/A	N/A
24	Agency 24	15	9	2	N/A	2
25	Agency 25	50			N/A	N/A
26	Agency 26	25	6	2	5	2
27	Agency 27	No respond			N/A	N/A
28	Agency 28	40.5	3	5	N/A	3
29	Agency 29			2	N/A	N/A
30	Agency 30	88.5	85	15	N/A	1
31	Agency 31		7	4	N/A	N/A
32	Agency 32	Not given info				N/A
33	Agency 33	16	15	4	3	2
34	Agency 34	30	4	2	1	3
	21 Data Centers	**1,563**	**363**	**High U**	**Medium U**	**Low U**

When we interviewed senior data center infrastructure executives about their innovation priorities and capabilities, many emphasized that constant innovation represented their only chance of meeting user expectations while supporting—within budgetary constraints—ever-increasing business demands for computation, data storages, and connectivity.

In most organizations, Data center began as a support function, leading to a multi-dimensional management approach. However, technology-enabled products, interactive communications, and 24x7 online information environments have thrust data center to the forefront, with critical implications for business growth and customer engagement. In addition, established practices, such as lean-management techniques, have highlighted the value of data center in reducing waste and increasing productivity and efficiency. This deeper recognition of data center's potential has given rise to a new management model category de facto data center. De facto data center comprises the bulk of an organization's data center activities, applying lessons from the production rack to room, standardization, and simplification to drive efficiency, optimize delivery, and lower unit costs.

4.1 Data center consolidation

The survey results from Table 4, since 34 agencies are under the same roof and more than 85% of data centers are under utilization. Data center consolidation is the best option to minimize economy and efficiency. The summarization of equipments and power requirements for data center consolidation is illustrated in Table 5.

4.2 Flexible layout, expansion, and transportation

Mobile data center is flexible expansion and mobilization. According to the operating situation is possible using the adjustable layout. Since the shape and size are the same standard as a container, transportation is very easy. This enables quick response to sudden transfer in case of any unexpected situation. Moreover, it is easy to expand as requirements within short time.

Figure 8 Mobile data center

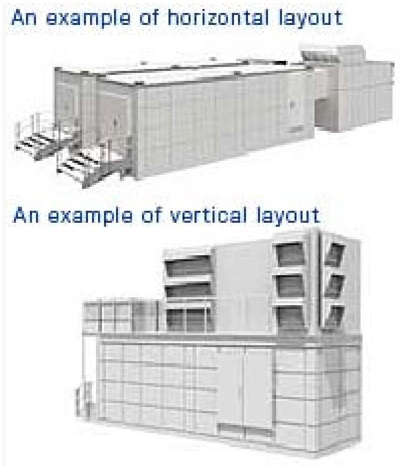

Figure 9 Modular data center

4.3 High density cooling systems

The ability to deliver redundancy in high density circumstances with fewer additional CRAC units is a strategic benefit to the in-row cooling oriented architecture (closed-loop cooling that can handle power density up to 40kW per rack, as shown in Figure 10.) and presents it a significant total cost of ownership (TCO) advantage and energy efficiency (PUE), as depicted in Table 6.

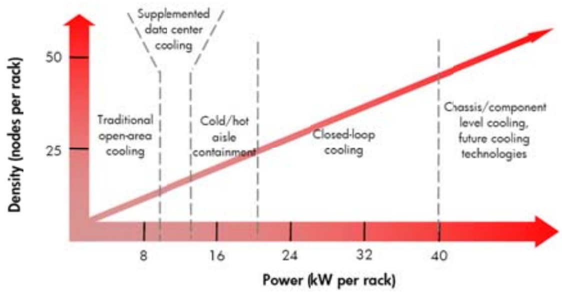

Figure 10 Cooling strategies based on server density/power per rack

Table 6 Cooling strategies: design for PUE impact

Activities	PUE Impact
Base Hot Aisle / Cold Aisle	2.75
Blanking Panels	2.73
Floor Brushes	2.70
Optimized Floor Tile Placement	2.65
CRAC / CRAH -Duct work	2.62
Drop Ceiling Plenum	2.58
Cabinet Layout - Optimized	2.55
Raised Floor - Optimized	2.50
Containment Cold Aisle	2.20
Containment - Hot Aisle	2.00
Containment - Chimney	2.00
Liquid Cooling - Stand alone Cabinets	1.75
In Row Cooling	1.50

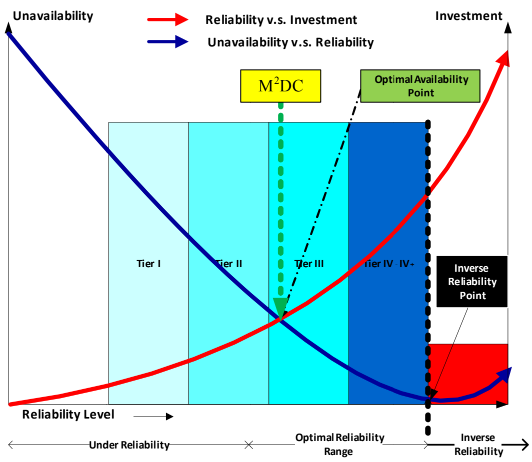

Figure 11 M^{2}DC design concerns unavailability v.s. reliability

4.4 Reliable power distribution systems

The strategic approaches to M^{2}DC infrastructure design are based on best practices of TIA 942 [8], Uptime Institute [9], BICSI [13] and effective balance among efficiency, availability & reliability, flexibility, and economy. The modification concept of M^{2}DC is concern on the idea point of complex system is increasing unavailability level [10], as demonstrated in Figure 11. As a result, the modified Tier III M^{2}DC power distribution system is trying to eliminate more connections and equipments, as illustrated in Figure 13. Especially, the M^{2}DC design is tried to eliminate the redundant equipment to fit with limitation space in 200 feet container. Generator design is N sets and UPS design is 2N.

4.5 M^{2}DC 20' for construction

Space utilization is the key of M^{2}DC design. Since M^{2}DC is deployed concept of server consolidation. Therefore, heavy draw-out power density and heat generating from server consolidation is cannot avoidable. The most common problem of data center is heat rejection. As the concept of in row cooling is the best option of PUE impact from Table 5. The case study from Table 4 requires 99,225 Watts for power requirement and 338,560 BTUs or 112,853 Watts for cooling requirements. Design for 113kW for 20 foot container is mission possible. The 3 racks of 40 kW in row cooling systems are deployed to handle 113kW heat rejection. By creating a closed-loop heat trap, it is increasing efficiency of return air flow system, as illustrated in Figure 12.

Unique 45U rack system, Figure 14, installation is required to protect data by absorbed vibration, moveable space for balancing central gravity point during life it up, and maintenance or installation from back and front sides.

Figure 14 45U racking system

4.6 Project management

M^{2}DC deployment will take less time more than a half of legacy construction data center, as shown in Figure 15.

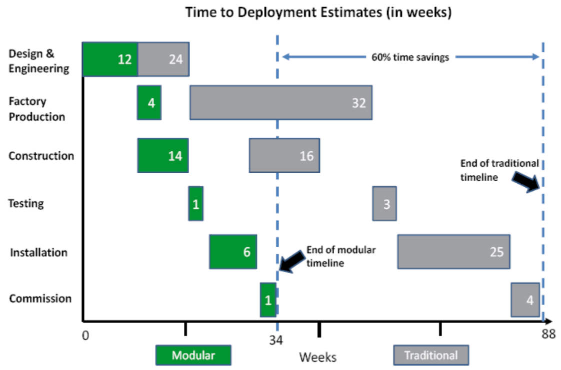

Figure 15 Project management efficiency

The report from Schneider whitepaper [11] is estimated around 54 weeks saving time during deployment containerized data center. The 54 weeks mean save more other costs such as labors, materials, operations, opportunities, and can build more M^2DCs. The estimate saving investment infrastructure (building construction) between legacy data center and M^2DC per square meter is about four times since building construction is around 200,000 baht per sqm. but M^2DC is around 50,00 baht per sqm. Moreover, M^2DC is saving electrical bill in long term operations because the different PUE ratio between legacy data center and M^2DC is 0.8 (PUE 2-1.2= 0.8) [14].

5. Conclusion

The mobile and modular data center (M^2DC) design can be used to identify cost-effective saving opportunities in operating facility infrastructure. Even though, no design guide can offer "silver bullet" to design a data center, but M^2DC offer efficient design suggestions that provide efficiency benefits in a wide variety of; space limited data center design situations; mobility incase of any natural disaster or manmade; and modular expansion (use as needed). The result from consolidation 21 data centers reveals benefits of M^2DC subject to reduce investing facility infrastructure, space occupier, electrical billing, and operations, but increase space and IT equipment utilization, and energy efficiency. M^2DC design is a relative new field that addresses a dynamic and evolving technology. The most efficient and effective M^2DC designs use relatively new design fundamentals to create the required high energy density (consolidation), high reliability environment, mobility and modularity. The above best practices capture many of the basis standard approaches used as a starting point by successful and efficient M^2DC.

References

[1] W. Pitt Turner IV et al., "Tier Classifications Define Site Infrastructure Performance", Uptime Institute, Whitepaper, 2008, pp.1-19.

[2] W. Pitt Turner IV et al., "Industry Standard Tier Classifications Define Site Infrastructure Performance", Uptime Institute, Whitepaper, 2005, pp.1-4.

[3] A. Rawson et al., "Green Grid Data Center Power Efficiency Metrics: PUE and DCIF", The Green Grid, Whitepaper, 2008, pp.1-9.

[4] K. Dunlap and N. Rasmussen, "The Advantages of Row and Rack-Oriented Cooling Architectures for Data Centers", Whitepaper, APC, 2006, pp.1-21.

[5] HP, "Cooling Strategies for IT Equipment", www.hp.com/go/getconnected
, September 2010.

[6] N. Rasmussen, "Avoiding Costs from Oversizing Data Center and Network Room Intrastructure", Whitepaper 37, Schneider Electric, 2011, pp.1-9.

[7] Emerson, "Maximizing Data Center Efficiency, Capacity and Availability through Integrated Infrastructure", Whitepaper, Emerson, 2011, pp.1-16.

[8] TIA 942, "Telecommunication Infrastructure Standard for Data Center", ANSI/TIA-942-2005, April 12, 2005.

[9] Uptime Institute, "Data Center Site Infrastructure Tier Standard: Topology", Uptime Institute, LLC, 2010.

[10] M. Wiboonrat, “An Optimal Data Center Availability and Investment Trade-Offs”, Software Engineering, Artificial Intelligence, Networking, and Parallel/Distributed Computing, 2008. SNPD '08. Ninth ACIS International Conference on, pp.712-719.

[11] D. Bouley and W. Torell, “Containerized Power and Cooling Modules for Data Centers”, Whitepaper 163, Schneider Electric, 2012, pp. 1-15.

[12] M. Wiboonrat, “Data Center Design of Optimal Reliable Systems”, Quality and Reliability (ICQR), 2011 IEEE International Conference on, pp. 350-354.

[13] BICSI, Data Center Design and Implementation Best Practices, BICSI 2002-2010, June 21, 2010.

[14] M. Wiboonrat, “Controversial Substantiation between Disaster Risk Reduction and Hedging Investment Risk,” The 2nd International Conference on Informatics & Applications, Poland, September 23-25, 2013. (accepted for publication)

Table 5 Total data center consolidation power requirements

HARDWARE		EQUIPMENTS	PER UNIT REQUIREMENTS (SET)				TOTOAL REQUIREMENTS (SUM)			
No.	Name	Model	Rack Mount Unit (RU)	Weight / Unit (kg)	Power Consumption / Unit (Watts)	Heat / Unit (BTU/Hour)	Qty	Total Weight (kg)	Total Power Consumption (Watt)	Total Heat (BTU/Hour)
1	WAN Router	Cisco ASR 1002	2	16.75	1,120	3,821	2	34	2,240	7,642
2	Core Switch	Cisco Nexus 7009	14	136.00	12,000	40,945	2	272	24,000	81,890
3	Service Node Switch	Cisco Catalyst 6506	12	69.60	12,000	40,945	2	139	24,000	81,890
4	Fabric Interconnect	Cisco UCS 6248 Fabric Interconnect	1	15.88	700	2,388	2	32	1,400	4,776
5	SAN Switch	Cisco MDS 9124 24-Port Multilayer Fabric Switch	1	8.40	600	2,047	2	17	1,200	4,094
6	Blade Server Chassis	Cisco UCS 5108 Blade Server Chassis	6	115.66	10,000	34,121	4	463	40,000	136,484
7	Storage	NetApp FAS 3240								
		- Storage Controller Enclosure	3	67.20	956	3,262	2	134	1,912	6,524
		- Disk Shelf Enclosure	4	49.90	639	2,180	7	349	4,473	15,260
Total			43	479.39	38,015	129,709	23	1,440	99,225	338,560
									(Watts)	112,853

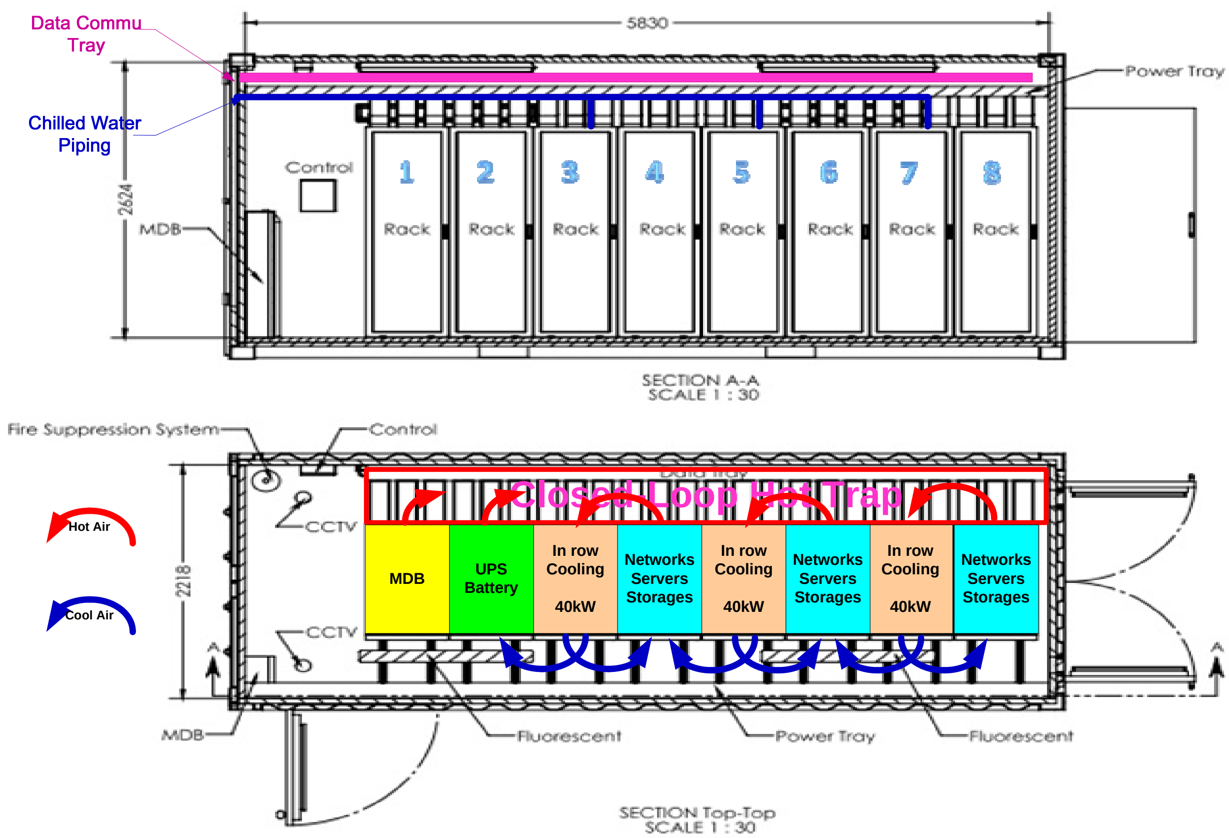

Figure 12 20 feet container design for M[2]DC

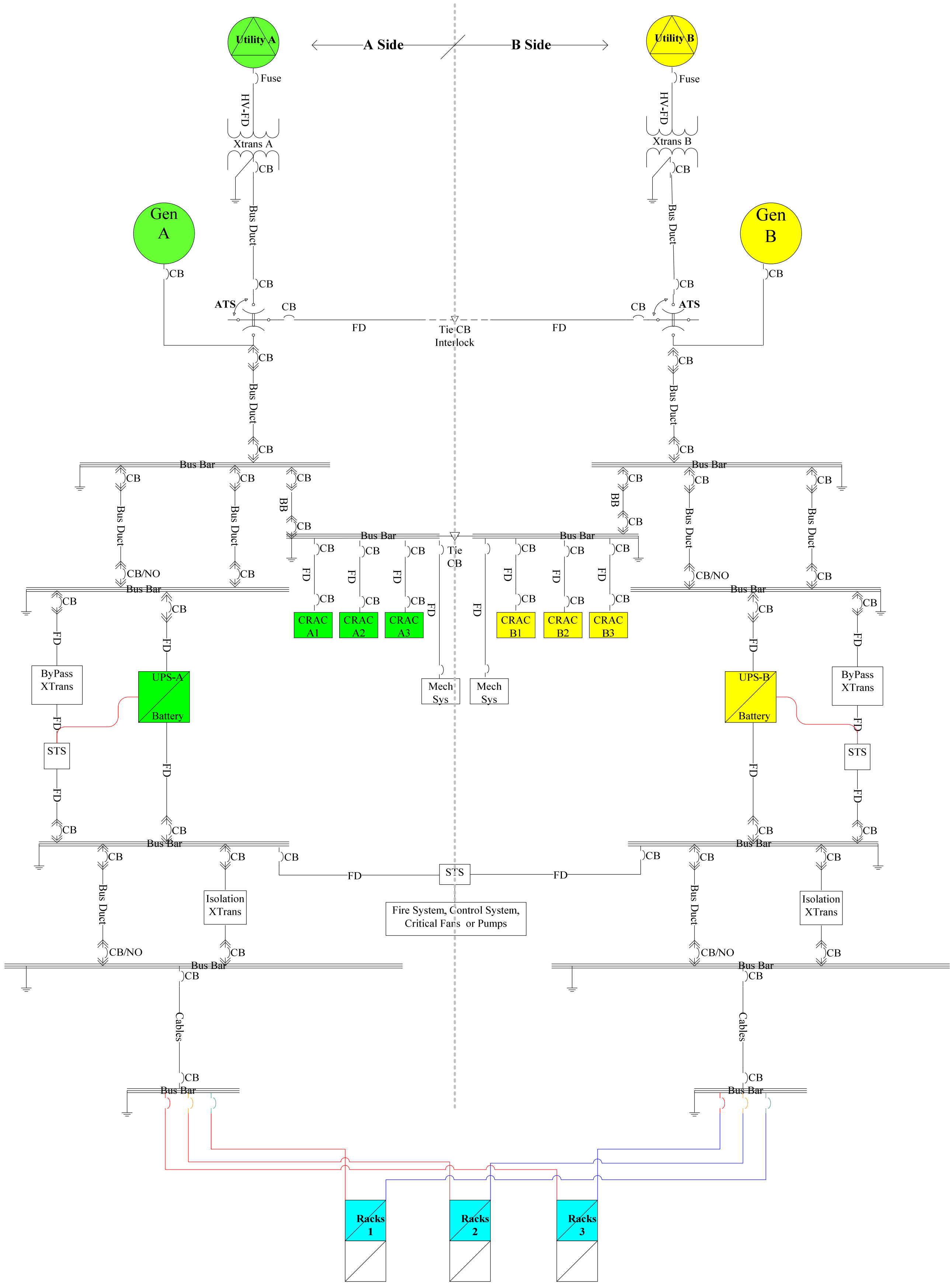

Figure13 Modified Tier III M^2DC power distribution system