Four Steps to Determine When a Standby Generator is Needed for Small Data Centers

White Paper 52

Revision 1

by Wendy Torell and Victor Avelar

> Executive summary

Small data centers and network rooms vary dramatically in regard to the amount of uninterruptible power supply (UPS) runtime commonly deployed. This paper describes a rational framework for establishing backup time requirements. Tradeoffs between supplemental UPS batteries and standby generators are discussed, including a total cost of ownership (TCO) analysis to help identify which solution makes the most economic sense. The analysis illustrates that the runtime at which generators become more cost effective than batteries varies dramatically with kW and ranges from approximately 20 minutes to over 10 hours.

Contents

Click on a section to jump to it

APC white papers are now part of the Schneider Electric white paper library produced by Schneider Electric's **Data Center Science Center**
DCSC@Schneider-Electric.com

Introduction

Virtually all small data centers and network rooms have UPS protection; however, the runtime varies dramatically across installations. For example, a typical UPS installed in a network room may have a full-load runtime of 20 minutes, whereas a UPS supporting a phone switch may have a runtime of 4 hours. A variety of approaches are routinely used to provide additional runtime where appropriate or necessary. These approaches include, but are not limited to:

- Adding a redundant parallel UPS or oversizing the UPS
- Adding supplemental UPS batteries
- Adding a diesel or gas powered standby generator

For many users, it is difficult to decide how much runtime is appropriate or necessary for the application, and how much additional availability is provided when runtime is extended. This paper describes a rational approach to assessing downtime risks and ensuring that required runtimes are met. This approach is comprised of four steps:

1. Understanding and limiting the consequences of downtime
2. Identifying runtime requirements
3. Maintaining cooling
4. Choosing between a generator or supplemental batteries

Step 1: Understanding and limiting the consequences of downtime

Understanding the consequences of backup time exhaustion is critical to understanding the value of additional runtime. These consequences may include:

- Abrupt termination of transactions, resulting in loss of possibly unrecoverable key data
- Leaving the protected loads in a software state where an extended and/or manual restart scenario is needed
- Abrupt termination of customer transactions , resulting in customer frustration and lost business

Related resource
White Paper 10
Preventing Data Corruption in the Event of an Extended Power Outage

Note that consequences such as data corruption may be avoided without implementing extended runtime. Unlike a utility power failure, the exhaustion of backup time is an event which the UPS can anticipate when operating under battery power. The UPS can signal applications and/or users to take actions before battery depletion which can dramatically reduce the problems associated with power termination. Specifically, the UPS can command applications and operating systems to shut down gracefully, which can dramatically simplify and shorten the restart procedure and assure a known state of the system. In addition, the UPS can provide warnings to users regarding when the system will shut down, so users can save data and close open transactions. White Paper 10, *Preventing Data Corruption in the Event of an Extended Power Outage* discusses this option in greater detail. However, in many mission critical environments, this system shutdown is simply not an option. This is where extended runtime becomes necessary.

Step 2: Identifying runtime requirements

Explicit backup time requirements *may be* provided by the business users of the protected equipment. When no requirement is provided, it is imperative for the system specifier to consult with the users, so that there is a clear understanding of the expectations for system performance. This consultation should include two key factors: the costs of downtime and the risk profile for the site.

Downtime cost profile for the business

It is important to understand the cost of downtime to a business and, specifically, how that cost changes as a function of outage duration. In many cases, the cost of downtime per hour remains constant as illustrated in **Figure1a**. In other words, a business that loses at a rate of 100 dollars per hour in the first minute of downtime will continue to lose at the same rate of 100 dollars per hour after an hour of downtime. An example of a company that might experience this type of profile is a retail store, where a constant revenue stream is present. When the systems are down, there is a relatively constant rate of loss.

Some businesses, however, may lose the most money after the first 500 milliseconds of downtime and then lose very little thereafter as shown in **Figure 1b**. For example, a semiconductor fabrication plant loses the most money in the first moments of an outage because when the process is interrupted, the silicon wafers must be scrapped.

Still others may experience little loss for a short outage (since revenue is not lost but simply delayed), but as the duration lengthens, there is an increased likelihood that the revenue will not be recovered. Regarding customer satisfaction, a short duration may often be acceptable, but as the duration increases, more customers will become increasingly upset. With significant outages, public knowledge often results in damaged brand perception, and inquiries into company operations. All of these activities result in a downtime cost that begins to accelerate quickly as the duration becomes longer. An example of this cost-of-downtime profile is shown in **Figure 1c.**

The majority of businesses would fit one of the profiles in these three figures. However, there are some rare exceptions where the profile might have a unique shape that is different from those presented here.

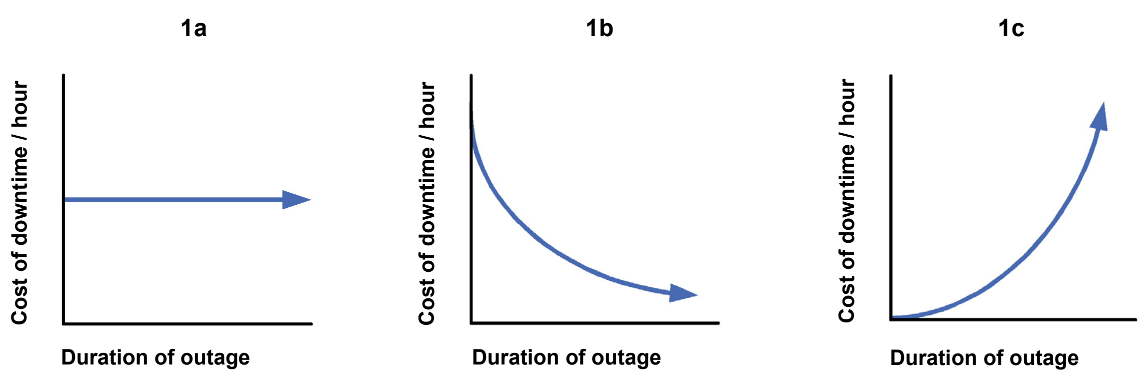


Figure 1a (left)

Constant cost of downtime

1b (center)

Decreasing cost of downtime

1c (right)

Increasing cost of downtime

Estimating the cost of downtime is often extremely difficult since total costs are a combination of direct costs and indirect costs, with indirect costs often very difficult to estimate.

Direct costs may include such things as:

- Wages and costs of employees that are not productive
- Revenues that cannot be captured because systems are down
- Wages and cost increases due to induced overtime or time spent checking and fixing systems
- Direct equipment costs for damaged equipment

Indirect costs may include such things as:

- Reduced customer satisfaction
- Customers who may have gone to direct competitors while you were down
- Damaged brand perception
- Negative public relations

Table 1 provides typical costs of downtime for a business, by industry. They are represented in dollars per employee-hour. These costs can be used as a guide to assess the financial impact for businesses in the same or similar industries – simply multiply the relevant revenue number times the number employed at the business to get the total revenue risk per hour.

Power risk profile for the site

Once the total cost of downtime is estimated, it is important to understand the risk profile for the site. A risk profile shows the likelihood of various power events for a physical location based on data collected over a period of time. Risk profiles are composed of two variables: frequency of events and duration of events. In general, short duration events occur most frequently, and as the event duration gets longer, they become less frequent. **Figure 2** illustrates this relationship.

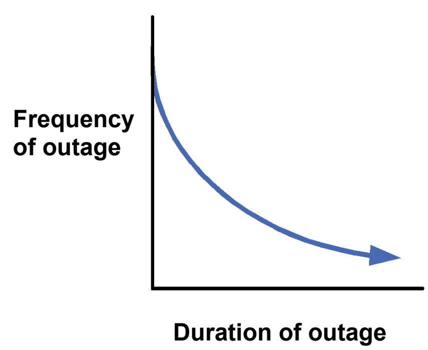
Table 1

Costs of downtime by industry

Industry sector	Revenue / employee-hour
Banking	$130.52
Chemicals	$194.53
Construction and Engineering	$216.18
Consumer Products	$127.98
Electronics	$74.48
Energy	$569.20
Financial Institutions	$1079.89
Food / Beverage Processing	$153.10
Health Care	$142.58
Hospitality and Travel	$38.62
Information Technology	$184.03
Insurance	$370.92
Manufacturing	$134.24
Media	$119.74
Metals / Natural Resources	$153.11
Pharmaceuticals	$167.53
Professional Services	$99.59
Retail	$244.37
Telecommunications	$186.98
Transportation	$107.78
Utilities	$380.94
Average	**$205.55**

Figure 2

Relationship between outage duration and frequency of occurrence



Tables 2a, b, and c represent sample sag / undervoltage and interruption profile data for New York, Texas, and Singapore, respectively. These profiles were created using an availability analysis calculator developed by Schneider Electric, based on power quality studies and weather statistics.

Table 2a

Sample risk profile for New York

Duration range	Sag or undervoltage events / years	Interruption events / years	Total / year
Less than 1 second	65.56	9.01	72.57
1 second - 10 minutes	0.56	2.26	2.82
10 minutes - 1.5 hours	0.03	0.23	0.26
Greater than 1.5 hours	0.01	0.05	0.06

Table 2b

Sample risk profile for Texas

Duration range	Sag or undervoltage events / years	Interruption events / years	Total / year
Less than 1 second	115.1	16.94	132.04
1 second - 10 minutes	0.6	2.31	2.91
10 minutes - 1.5 hours	0.02	0.22	0.24
Greater than 1.5 hours	0.0	0.05	0.05

Table 2c

Sample risk profile for Singapore

Duration range	Sag or undervoltage events / years	Interruption events / years	Total / year
Less than 1 second	120	20	140
1 second - 10 minutes	1.2	4.8	6.0
10 minutes - 1.5 hours	0.15	0.54	0.69
Greater than 1.5 hours	0.025	0.09	0.115

With a rough understanding of the types, frequencies, and durations of events likely for a given area, and the costs of downtime understood, it is now important to understand the typical time for systems to recover from power events. For example, if a 2-hour outage results in a corrupted database that takes 4 hours to recover, the total outage duration for that event is actually 6 hours. Likewise, a 500ms outage would take 4 hours and 500 ms to recover. Note that a corrupted database is just one of several tasks that may be required before a complete system recovery.

Making the business case

It should be evident that cost of downtime and risk of downtime events are the two key factors in determining the length of runtime needed for a business. As with any business case analysis, other variables must also be considered before determining the runtime and equipment solution that can feasibly be implemented. These include:

- The cost of the solutions at varying runtimes (discussed in later section)
- The budget – how much is available to spend
- The internal rate of return (IRR) of this project relative to other business projects
- Indirect benefits of implementing a solution, such as creating a competitive advantage
- Past personal experiences with downtime events or those experienced by others
- Impact on the cost of business continuity insurance

Step 3: Maintaining cooling

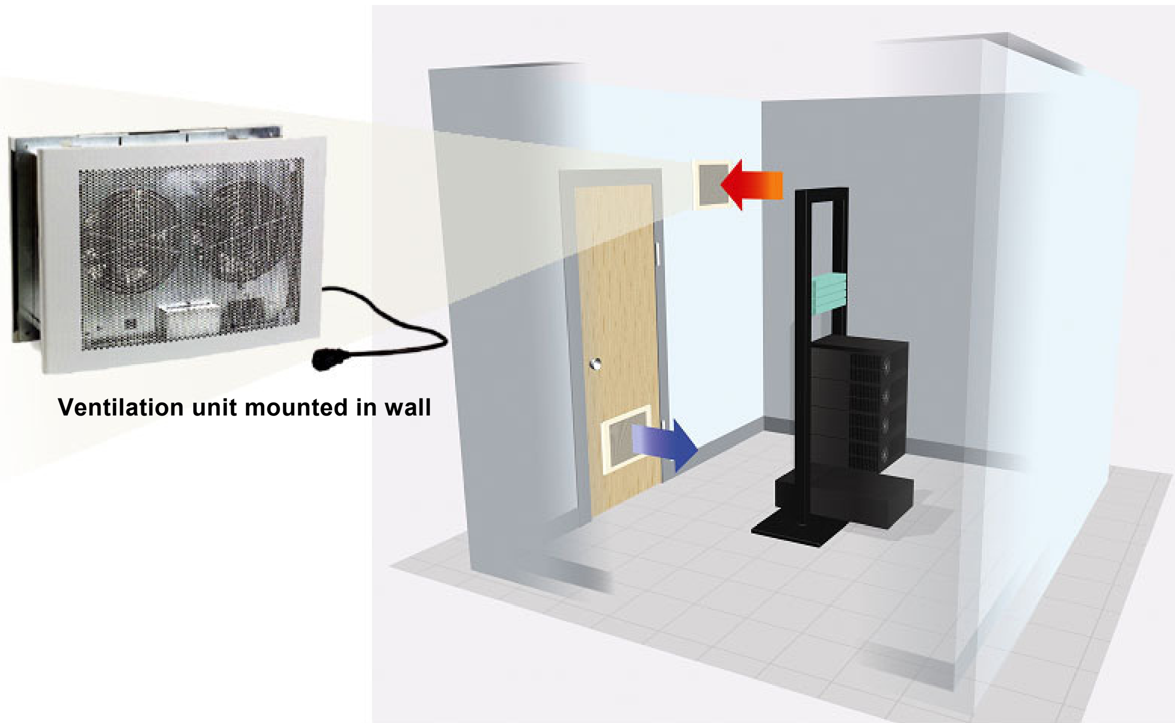
In the event of a power failure, it is necessary to remove the heat generated by the protected loads in addition to delivering power. A load that is powered but not adequately cooled can overheat and shut down – or worse, become damaged. Typically, a data center may “heat up” in less than 20 minutes, depending on various factors and the distribution of load and racks. In situations where racks are loaded with high density equipment such as blade servers, thermal shutdown could occur in less than five minutes. When the UPS operates under battery power, the room, racks, and the UPS / battery itself must be kept within the manufacturer’s operating temperature limits. If the batteries are sized to provide two hours of runtime during an outage, then there is a strong likelihood the room will heat up beyond the manufacturer’s recommendations. It is therefore critical to consider cooling of the UPS and IT equipment during outages.

Related resource
White Paper 68
Cooling Strategies for IT Closets and Small Rooms

For small computer rooms or closets, battery-backed fan ventilation may be adequate during extended outages. The basic principle is that the heat in the space is vented to the ambient building air, and the room or closet air does not rise substantially above the building ambient air temperature. An example of such a system is the Schneider Electric Wiring Closet Ventilation Unit, as illustrated in **Figure 3**. This approach is effective up to approximately 2 kW. For more information on this strategy, see White Paper 68, *Cooling Strategies for IT Closets and Small Rooms*.

Figure 3

Application of a wiring closet ventilation unit



For larger spaces and / or greater power densities, precision air conditioning may be required to maintain adequate room conditions. Some users may want to power the entire air conditioning system from the UPS to ensure cooling backup during an extended outage. While at first glance this may appear to be an easy solution, it will cause more serious problems for the power system's integrity. Compressors contained in cooling units, when energized, require peak currents that may be up to six times higher than their steady-state requirements. This high inrush current may cause the UPS to transfer to bypass, defeating the purpose of having a UPS system backup for critical IT equipment. A backup generator, on the other hand, can provide power to both UPS and cooling equipment during long outages.

When using a generator, there is usually a short interruption in cooling between when the utility power fails and when the generator restores power to the cooling system. In most cases this interruption poses no threat to the critical loads of a data center. However, in cases where the rack power density approaches or exceeds 10 kW, this small period of time without cooling can cause these individual racks to overheat. In these cases it is recommended that air-conditioner blowers and air distribution fans be powered by the UPS.

Step 4: Choosing between a generator and supplemental batteries

Supplemental batteries are very easy to install and are sometimes the only choice when the installation of a generator is impractical. For instance, a site might not have space available, or there might be restrictions by environmental regulatory organizations on installing a generator. Another possibility is that the landlord of a leased space does not permit the installation of a generator.

On the other hand, if extended runtime is determined to be necessary and high density loads are present, the only realistic option to keep the systems running effectively is a generator, since cooling must be supported (as discussed in Step 3).

Supplemental batteries are more cost effective than a generator in many cases. However, when the runtime requirement increases and / or the load needing extended runtime increases, there is a point where a generator has a compelling economic advantage. **Figure 4** illustrates that when comparing the total cost of ownership (TCO) of any particular size UPS with batteries vs. a UPS with generator backup, there is a clear decision point.

Figure 4

Typical battery vs. generator TCO analysis

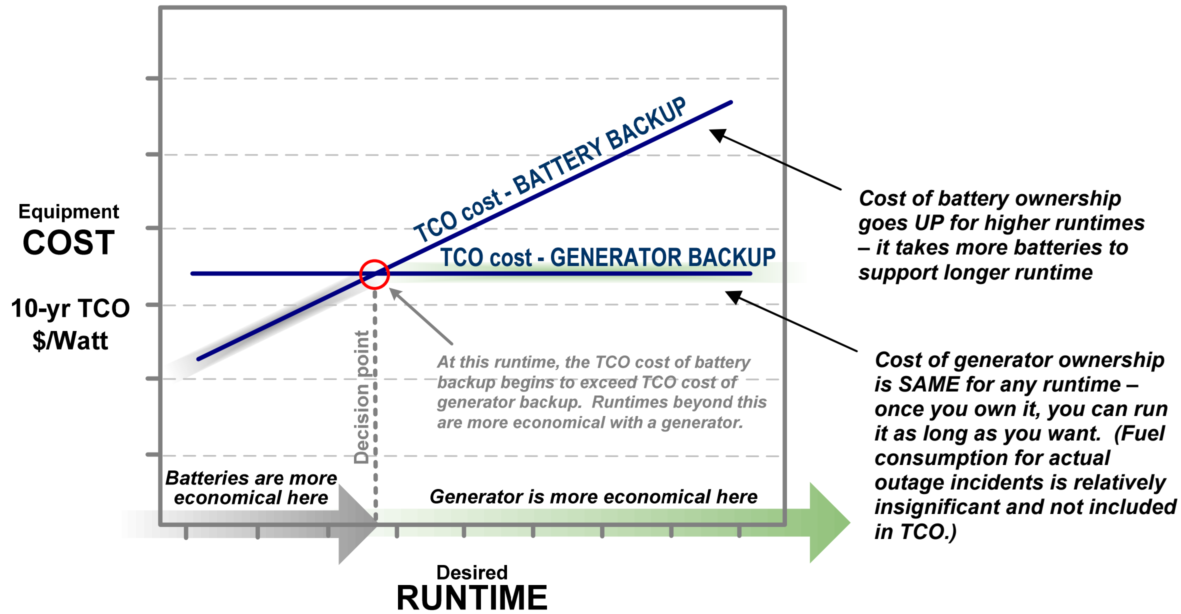
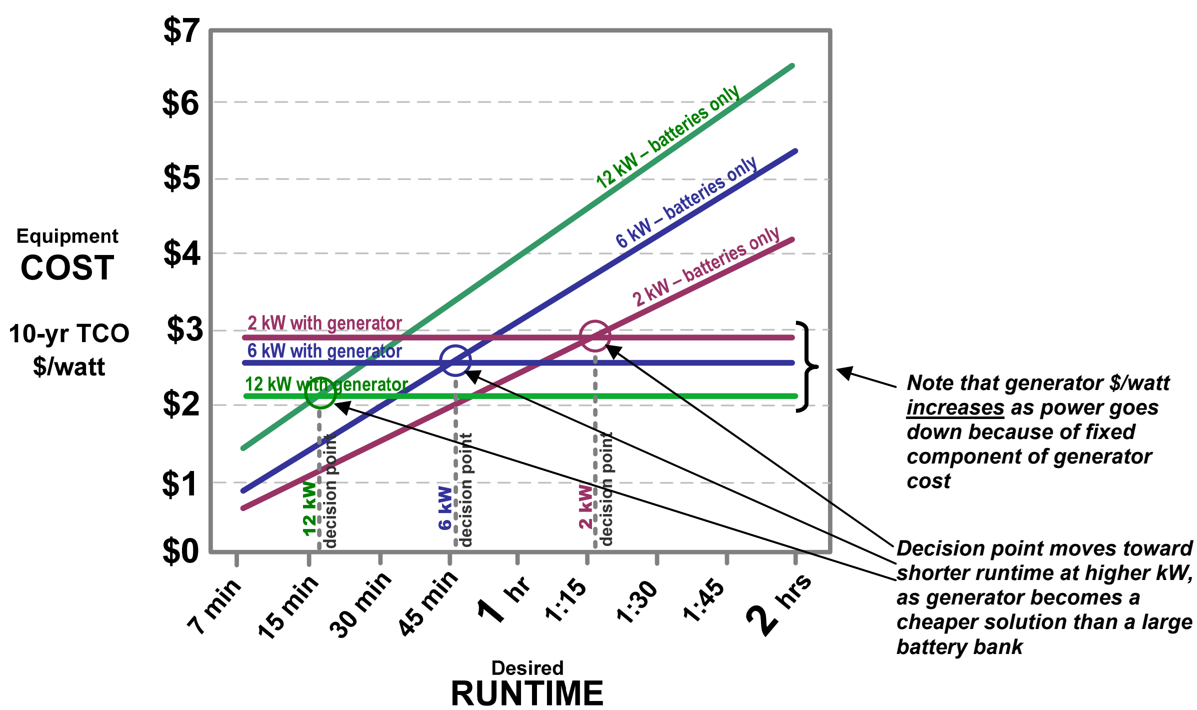


Figure 5 shows the system TCO in $ per watt as a function of runtime for loads ranging from 2 kW to 12 kW. An explanation of the data and assumptions that are the basis of this graph is found in the **Appendix**. This graph can be used as a guide when deciding which solution makes most economic sense.

Figure 5

Representative TCO analysis for three different UPS loads

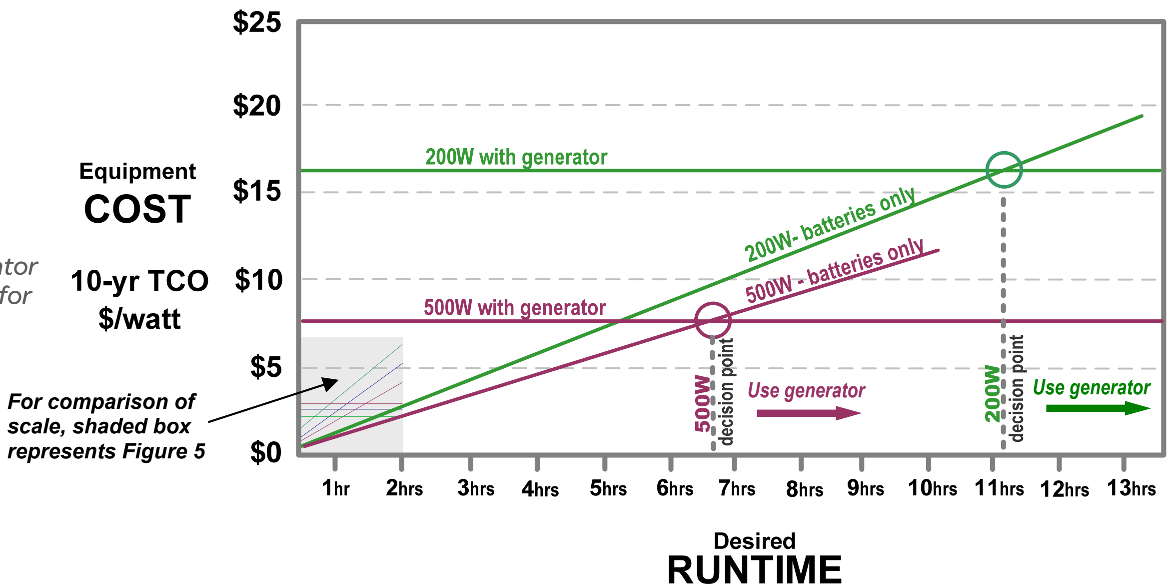


The graph shows clearly that supplemental batteries are less costly for low runtimes, but that generators begin to show a significant cost advantage as runtime requirements increase. For example, the point where a generator begins to have a cost advantage over extended run batteries for a typical 2 kW system is over an hour. When a typical 12 kW system is considered, this point becomes closer to 20 minutes.

As the loads get even smaller than the cases presented above, the point at which generators make more economic sense move to significantly higher runtimes. **Figure 6** illustrates the system TCO in $ per watt as a function of runtime for 200 watt and 500 watt loads. Again, the sample cost data and assumptions this graph is based on are found in the **Appendix**.

Figure 6

At very low power, generator is not economical except for very long runtimes



As this graph illustrates, the decision point when a generator becomes the economical choice is over 11 hours for a load of 200 watts. With such small loads, longer runtimes can easily be achieved by both oversizing the UPS capacity and adding extended runtimes. For example, an APC Smart-UPS 750 XL with two extra batteries provides a 200 watt load over nine hours of runtime at a TCO less than a generator.

Conclusion

Adding extended backup time to a UPS can be a cost effective way to increase availability in network rooms. Nevertheless, backup runtime requirements can be a primary driver of UPS system cost. The guidelines provided in this paper can be used to estimate the cost of downtime, understand the risk profile for a site, and ultimately determine the runtime requirements. When extended runtime is determined to be necessary, the appropriate decision can then be made between specifying a standby generator or supplemental UPS batteries.

About the author

Wendy Torell is a Strategic Research Analyst at Schneider Electric in West Kingston, RI. She consults with clients on availability science approaches and design practices to optimize the availability of their data center environments. She received her Bachelors of Mechanical Engineering degree from Union College in Schenectady, NY and her MBA from University of Rhode Island. Wendy is an ASQ Certified Reliability Engineer.

Victor Avelar is a Senior Research Analyst at Schneider Electric. He is responsible for data center design and operations research, and consults with clients on risk assessment and design practices to optimize the availability and efficiency of their data center environments. Victor holds a Bachelor's degree in Mechanical Engineering from Rensselaer Polytechnic Institute and an MBA from Babson College. He is a member of AFCOM and the American Society for Quality.

Resources

Click on icon to link to resource

Preventing Data Corruption in the Event of an Extended Power Outage
White Paper 10

Cooling Strategies for IT Closets and Small Rooms
White Paper 68

White Paper Library
whitepapers.apc.com

TradeOff Tools™
tools.apc.com

Contact us

For feedback and comments about the content of this white paper:

Data Center Science Center
DCSC@Schneider-Electric.com

If you are a customer and have questions specific to your data center project:

Contact your **Schneider Electric** representative

Appendix – TCO analysis of UPS with supplemental batteries vs. UPS with standby generator

For the ten year TCO analysis found in **Figure 5** and **Figure 6**, five different loads were considered – 200 W, 500 W, 2 kW, 6 kW, and 12 kW. For each, a typical UPS model was chosen and a generator sized accordingly. Estimated resale pricing for each system was applied; and industry averages for labor costs were applied. All UPS-related costs are based on estimated resale price or ERP, provided by Schneider Electric. All generator-related costs are based on estimated resale price, provided by various industry sources. Note that the generator solutions require a minimal amount of battery runtime to allow for the transfer delay to and from utility power. The TCO, in $ / watt, for each system is assumed to be flat, since a full tank of fuel is assumed to be on hand, and a fixed amount of monthly generator testing is required, regardless of target runtime. Although electricity is commonly added to a TCO analysis, for the purpose of this comparison, it has been excluded, since the electrical cost difference is negligible. Since UPS batteries have a typical lifespan of 3-5 years, the analysis assumes that batteries will need to be replaced at years 4 and 8 of the ten-year TCO. **Table A1** describes the systems and their TCO breakdown for the case of a UPS with 7 minutes of runtime and a standby generator.

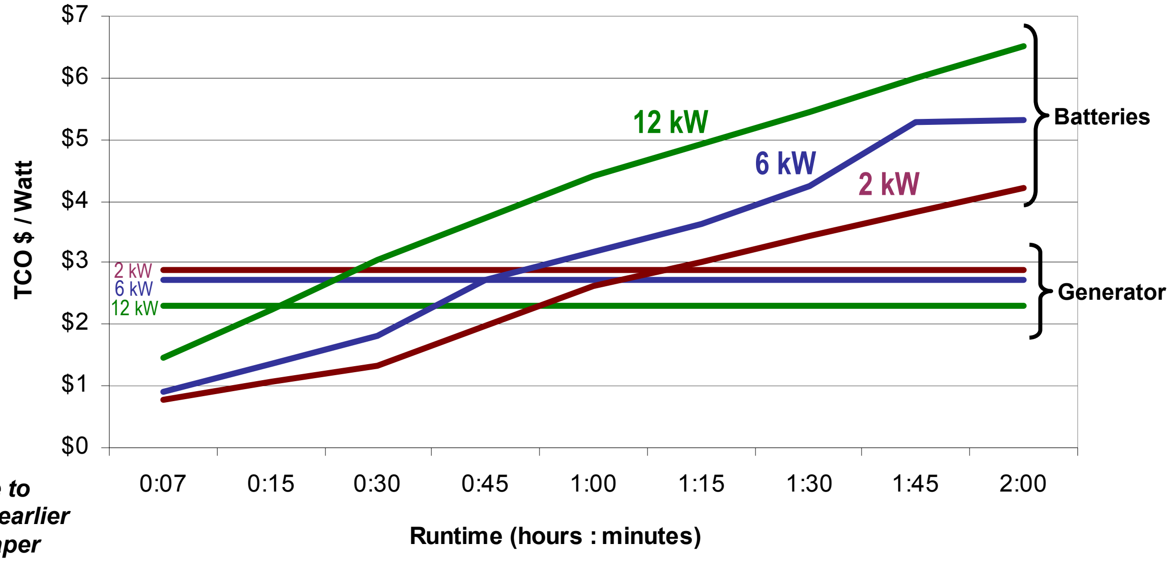
Table A1

Systems used in analysis and their TCO breakdown for 7 minutes runtime

TCO variables (based on estimated retail pricing)	200 Watt UPS & 1 kW generator	500 Watt UPS & 1 kW generator	2 kW UPS &4.5 kW generator	6 kW UPS & 12 kW generator	12 kW UPS & 20 kW generator
APC UPS model	Smart-UPS SC 420 VA	Smart-UPS SC 1000 VA	Smart-UPS 3000 VA	Smart-UPS RT 7500 VA	Symmetra LX 16kVA
APC battery model	RBC2	RBC33	RBC55	RBC44 x 2	SYBT5 x 4
UPS capacity (watts)	200	500	2400	6000	12000
Installation cost of generator	$1,000	$1,000	$2,000	$2,000	$2,000
Installation cost of UPS	$0	$0	$0	$500	$500
Maintenance of generator	$800	$800	$1,500	$3,200	$4,000
Maintenance of UPS	$0	$0	$0	$0	$0
Generator system cost	$300	$400	$750	$2,750	$4,000
ATS system cost	$600	$600	$600	$750	$1,000
Fuel for 10 years	$360	$360	$500	$801	$1,188
UPS system cost	$120	$330	$1,100	$4,225	$11,800
Battery (refresh 1) cost (year 4)	$30	$70	$240	$1,078	$1,520
Battery (refresh 2) cost (year 8)	$30	$70	$240	$1,078	$1,520
Total system TCO	$3,240	$3,630	$6,930	$16,382	$27,528
TCO in $ / watt	$16.20	$7.26	$2.89	$2.73	$2.29

Figure A1

UPS with batteries vs. UPS with generator – based on actual pricing data for 2 kW, 6 kW and 12 kW loads



Compare to Figure 4 earlier in this paper

Based on this data, the decision point – the point where the generator becomes more economical – varies between 20 minutes and 1 hour 10 minutes, as shown in **Table A2**:

Table A2

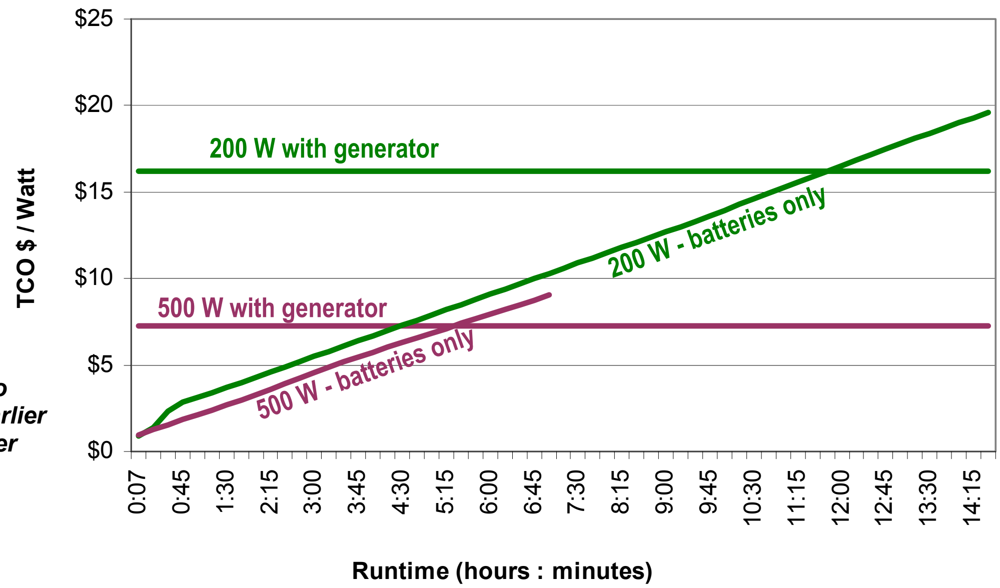
Decision Point for 2 kW, 6 kW, and 12 kW Loads

Load	Approximate decision point
2 kW	1 hour 10 minutes
6 kW	45 minutes
12 kW	20 minutes

Figure A2

UPS with batteries vs. UPS with generator – based on actual pricing data for 200W and 500W loads

Compare to Figure 5 earlier in this paper



Based on this data, the decision point – or point where the generator becomes more economical – is substantially longer in runtime than the three earlier cases, as shown in **Table A3**:

Table A3

Decision point for 200 watt and 500 watt loads

Load	Approximate decision point
200 watt	11 hours 45 minutes
500 watt	5 hours 30 minutes

An Introduction to P3 & PQU

Power Protection Products, Inc. (P3)

P3 is the industry's trusted and respected critical power, cooling and energy solutions provider.

P3 represents some of the biggest industry brands and strives to provide the top performing products. P3 stays on top of industry advances by participating in power quality studies, conferences and seminars with respect to data centers, power and loading analysis, transient voltage detection and power system harmonic analysis & mitigation, ground studies, and power monitoring systems.

P3 has designed and built a variety of data centers, IT expansion projects, wiring closet upgrades, and a host of product and equipment placements as well as provided equipment for a variety of colocation companies. We believe in providing a stable, secure IT environment to meet customers' needs.

Power Quality University

In addition to product sales, distribution, and marketing, P3 is also a leader in customer education through Power Quality University, which provides an educational environment for hands-on training, testing, and evaluation of today's power quality and data center solutions and equipment. The instructors who teach at PQU are highly qualified professionals and are all experts in the field. Upon completion of a PQU program, students can obtain Continuing Education Credits (CEU's). PQU is just one more way that P3 is affecting the electrical community in a positive way.

For more information about our power related products & services contact us at:

877-393-1223

Or visit our website.

www.p3-inc.com