PROXY PROPOSALS FOR MEASURING DATA CENTER PRODUCTIVITY

CONTRIBUTORS:

JON HAAS, INTEL

MARK MONROE, SUN MICROSYSTEMS

JOHN PFLUEGER, DELL

JACK POUCHET, EMERSON

PETER SNELLING, SUN MICROSYSTEMS

ANDY RAWSON, AMD

FREEMAN RAWSON, IBM

 Version 1.0 Release Date: TBD

TABLE OF CONTENTS

BACKGROUND

This paper is a natural progression of the work that The Green Grid and the industry have already done to advance energy efficiency in data centers. A prior paper[1] defined the Power Usage Effectiveness (PUE) and Data Center Infrastructure Efficiency (DCiE) metrics as follows:

$$\text{PUE} = \frac{\text{Total Facility Power}}{\text{IT Equipment Power}} \quad \text{and its inverse,} \quad \text{DCiE} = \frac{\text{IT Equipment Power}}{\text{Total Facility Power}} \times 100\%$$

These metrics deal exclusively with power. The same paper also introduced the evolution of these metrics to those that measured productivity and suggested that Data Center Productivity (DCP) could be a measure of useful work, divided by total facility power.[2]

A subsequent paper[3] investigated this further and introduced the concept of a family of DCP metrics (referred to as DCxP metrics) based on a definition that measured useful work produced in a data center divided by any quantization of a resource consumed to produce that work. The paper also defined the first metric in this family, DCeP, where the resource consumed is energy. Thus, DCeP is defined as follows:

$$\text{DCeP} = \frac{\text{useful work produced in a data center}}{\text{total energy consumed in the data center to produce that work}}$$

The paper goes on to provide a detailed metric[4] for accurately measuring useful work in a data center.

BUILDING ON CURRENT METRICS AND INFRASTRUCTURE

While the DCeP metric is well defined, it can sometimes be difficult to implement. The Green Grid recognizes that there is a need for something that is simpler to implement than the DCeP metric, and that can be used to make comparisons of the work done within a data center. In addition, the industry has already made substantial investment in instrumentation to derive metrics pertaining to power and energy, and there are opportunities to leverage those investments and provide new tools which can take advantage of instrumentation already in place.

PURPOSE OF PAPER

To that end, The Green Grid is proposing that a simple indicator, or proxy, while less accurate than a metric such as DCeP, can be useful because it will be easier to implement. The goal is to find a proxy that will substitute for a difficult measurement, but that still gives a good enough indication of useful work completed. An example of an efficiency proxy might be the Environmental Protection Agency (EPA) mileage rating in the window of a new car, which gives the prospective buyer a useful indication of the vehicle's fuel efficiency, even though the buyer realizes that "Your actual mileage will vary."[5]

To define a suitable proxy, or proxies, The Green Grid is engaging the industry by offering several proxy proposals and soliciting feedback. The goal is not to pick a winner at this time but to learn from the industry which features and techniques are most useful and which are to be avoided before further research is begun.

The resulting proxy (or proxies) will not be a substitute for DCeP. The intent is to provide another tool to assist data center operators by adding simplicity and ease of use at the expense of accuracy. As such, it is referred to as a proxy because it is more an indicator than an accurate measure.

The Green Grid is proposing eight proxies for evaluation and will solicit feedback from The Green Grid members and the larger community of data center and information technology professionals. After at least 90 days, The Green Grid will begin evaluating feedback. The final choices for a proxy or proxies will then be the subject of additional research work by The Green Grid.

UNDERSTANDING THE CONTEXT FOR PROXY UTILIZATION

A less-accurate and often indirect measurement such as a proxy can only provide a useful indication if the user understands its limitations. It becomes very important, then, to understand the context within which the metric is valid, and the limits of the metric itself.

With the goal to define metrics that are useful, it is important to understand that specific metrics may need to be devised for specific kinds of work done in data centers. Useful work in a data center can take many forms. High performance computing (HPC) centers may measure work in terms of the number of proteins folded, genomes calculated, or weather models iterated. Web-search data centers might measure the number of queries served or the number of pages indexed. Corporate data centers might handle a mixture of emails, web pages, application transactions, and even voice traffic using Voice Over Internet Protocol (VOIP).

Comparisons between data centers would be valuable in a marketing or evaluation sense, but it is unlikely that any proxy for data center productivity will be comparable across multiple data centers. Rather, the primary use for a proxy will continue to be an indication of improvement over time for a single data center, and very constrained comparisons between data centers that perform the same function. The reader should keep this limitation in mind when evaluating the proposed proxies, as well as when using proxies or other metrics to assess data center performance and efficiency.

ENERGY VERSUS POWER

Note that both energy and power are discussed in the preceding paragraphs. These terms mean different things but are often used interchangeably and incorrectly. It is important to understand the difference and adhere to the correct meaning.

Power is a rate. For our purposes, it is the rate at which IT equipment or data center infrastructure consumes electricity. It is measured at an instant in time as joules per second, or watts. In a data center it is typically reported in kilowatts (kW).

Energy, on the other hand, is a total quantity, or an amount. It is power integrated over time and is typically measured in kilowatt-hours (kWh).

Additional confusion can happen when one realizes that power is typically measured not by taking an instantaneous measurement of power, but by measuring energy use over a period of time and then dividing by that time period to compute an average power. This is done because instantaneous power at a data center varies over time in a cyclic fashion 50 to 60 times a second. Thus average power becomes the most meaningful measurement.

Power capacity and energy consumption are both important in the data center. Power capacity (particularly peak power capacity) is important because both the power conditioning and distribution infrastructure and the cooling systems must be sized based on their combined peak power requirements. Energy consumption is important because a significant part of a data center's operating expense is attributed to energy consumption. One is used to assist in provisioning equipment and the other is used to measure cost of operation.

PROXY EVALUATION CRITERIA

It will be difficult for any proxy proposal to address all aspects of data center performance and productivity. Each proposal should satisfy a set of ideal criteria, some of which are described below. If additional criteria are important, be sure to describe them in feedback to The Green Grid in order to have them evaluated and potentially included in the final decision criteria. For now, however, here is a list of some of the criteria that a proxy will ideally address:

EASE OF USE:

The proxy should be easy to use and require as little setup or decision-making as possible. The results should not require highly skilled technical personnel (e.g., college-educated computer science or engineering personnel) to interpret. The results should be simple and objective.

ACCURACY:

It is expected that accuracy may be sacrificed for ease of use. Accuracy requirements are driven by the granularity of the decisions required; as needs become more specific, accuracy becomes more important. However, the proxy must indicate correctly. It must move in the correct direction and in correct proportion when an action is implemented. For example, a ten percent increase in work performed in the data center should lead to a proportional increase in data center productivity shown by a proxy.

RESPONSIVENESS:

The proxy should react immediately to changes in the IT or facility infrastructure and not have to be explicitly re-tested. One should not have to wait more than a few hours to see if a change in operations or infrastructure results in a change in efficiency.

COST:

Where possible, a proxy should be designed to leverage existing data collection tools and infrastructure that would typically be found in the data center. Setup effort should be less than one man-week in order to begin collecting the metric. No additional capital equipment or software licenses should be required.

INVASIVENESS:

The proxy should be minimally invasive, i.e., the potential effect on production systems should be as small as possible. Ideally, no additional software will need to be developed or additional agents installed on production systems.

TIME TO IMPLEMENT:

It should be possible to implement the proxy in less than a week. However, The Green Grid acknowledges that it is impossible to establish a realistic time to implement, so this is considered a goal and not a requirement. Industry and policy-making bodies do not want to wait for feature changes to existing products, or the development of new software or measurement techniques.

COMPLETENESS:

The proxy should account for every device in the data center. The portion of the proxy that measures useful work should account for work done by all of the IT equipment. All power consumption in the facility should be counted in the energy calculations.

OBJECTIVITY:

The proxy should require as little personal judgement as possible – either in setup or interpretation. Data collected should be numeric and continuous in nature, versus textual or discrete (e.g., on/off).

UTILITY:

The proxy should provide the information required to make better decisions. Results should be immediately actionable. Once the proxy has been established, the aggregation of real-time measurements should be available within minutes, with the readings stored and trended over time to allow for seasonal and event-driven differences to be identified and adjusted.

COMMISSIONING:

The proxy should be measurable during the commissioning of new data centers. Proxy measurements should be able to be utilized in a standardized test to verify efficiency design specifications for new data centers.

OPERATIONAL ABILITY:

The proxy should be measurable without interrupting the daily operation of the data center.

PROPOSED PROXIES

The following pages present eight proxy proposals for industry evaluation. All proxies are suitable to all data center archetypes, but remember that they are only useful for comparisons within a single data center.

ENERGY PROXY VERSUS POWER PROXY

Note that the first seven proxies are energy proxies and have energy in the denominator. The last proxy is a power proxy and thus has power in the denominator. This is entirely appropriate depending on how the numerator is defined.

Productivity is defined as the ratio of useful work to the amount of resource consumed to produce the work. When useful work is expressed as an aggregate amount of work completed over a specified period of time (referred to in the proxies as the assessment window) then the appropriate quantity to which this should be compared is energy. When useful work is expressed as a rate (e.g., transactions/sec) then the physical quantity to which it should be compared is power, also a rate. In this case, both useful work and power should be assessed at the same point in time.

Keep in mind that different proxies have different utilities. An “energy” proxy is useful for measuring CO2 footprint, electric consumption, and the cost to run the data center. A “power” proxy is typically used for capacity planning.

Energy proxies have E_{DC} in the denominator. E_{DC} is defined as the total energy consumed by the data center during the assessment window over which useful work is aggregated. Power proxies have P_{DC} in the denominator. P_{DC} is defined as the power drawn by the data center at the time the useful work assessment is made. Both E_{DC} and P_{DC} are measured in accordance with practices established in a previous Green Grid paper[6].

An energy proxy can be differentiated to be a power proxy, if the user decides that would be more useful. In essence, an energy proxy can be evaluated over a very small period, thus approximating a power proxy. In particular, proxies #5 through #7 can be easily converted to power proxies by removing the variable T from the product in the numerator and changing the denominator to power, P_{DC}.

the green grid™

SHORTCOMINGS ACKNOWLEDGED

The Green Grid acknowledges that these proxies, to varying extent, do not explicitly address all data center devices and thus fall short of a complete overall measure of data center productivity. In particular, the areas of storage and networking are not directly addressed. Although some proxies measure variables that are coupled to storage or network usage, for most this is not the case. While all of the proxies account for the energy or power consumed by these devices, none of them specifically account for the value of these non-compute devices in the measurement of useful work.

With regard to storage in particular, the greatest business value may not always be in its measure of throughput, but in its retention of information. Determining the value of data retention is a complex problem that is not addressed here; not because it is not important, but because it would complicate the proxies and thus defeat part of their purpose, namely, simplicity. The contribution of storage and networking needs to be further explored before meaningful and useful conclusions can be made regarding their impact on data center productivity. The Green Grid expects that specific metrics will be developed to address these areas in the future.

Another shortcoming is that proxies which measure individual server utilization (proxies #5 thru #7), as opposed to measuring an actual transaction, do not account for interactions between the individual servers and the data center. For example, a server may show a high utilization, not because it is executing a transaction, but because it is being forced to wait because of something another server is doing.

FREQUENCY OF MEASUREMENT

Frequency of measurement is dependent on the intended use of the proxy. If the proxy is being used to evaluate changes to the configuration or set point of a data center, the assessment window and frequency of measurement should be a period of time that would allow the data center to experience its normal fluctuations in computational demand. Guidelines for frequency of measurement for DCiE have been published in a prior paper[7] and apply to these proxies as well.

It is also anticipated that a proxy might be taken before and after any change has been made that might change the value of the proxy, such as a change in workload, improvement in power supply or distribution efficiency, or if and when a sample workload is updated (much as benchmarks are periodically updated).

SIMILARITIES AND DIFFERENCES

Proxies #1 through #7are energy proxies. Proxy #8 is a power proxy.
Proxies #1 and #2 are related; proxy #2 is a specific implementation of the concept presented in proxy #1.
Proxies #2 and #3 both use a subset of servers to make measurements.
Proxies #5 and #6 are very closely related; both use CPU utilization and SPEC* benchmarks.

PROXY PROPOSAL #1 - USEFUL WORK SELF-ASSESSMENT AND REPORTING

DETAILED DESCRIPTION:

This proxy estimates the metric Data Center Energy Productivity (DCeP), as previously defined[8] by The Green Grid, and its definition is consistent with the "Desired Attributes of a Productivity Metric."[9] As does DCeP, this proxy seeks to measure the aggregate amount of useful work that a data center produces during an assessment window and compares it to the total amount of energy consumed by the data center during this time.

This proxy specifically addresses criticisms regarding the difficulty in the measurement of the numerator of DCeP, viz. useful work. The formula for useful work[10] has a number of parameters that must be defined, and the definition of what constitutes a task remains difficult. The time-based aspect of the value of a task can be ignored by setting the utility function, $U_i(t,T)$, to 1, but the user still must determine the normalization factor, V_i, for each task type.

This proposal recognizes that the author of an application or software platform is in the best position to decide what constitutes useful work for that application and should provide the instrumentation required to measure the quantity of this work completed in a given assessment window. Each application or software platform (e.g., email/calendar software, relational database) typically has internal instrumentation that assesses its health and current level of performance, so the addition of software to accumulate units of useful work by a specific application will be a natural extension of features already included by the author.

HOW TO CALCULATE THE PROXY:

The proxy for DCeP is calculated by dividing the aggregate useful work completed during the assessment window by the amount of energy consumed during the assessment window, as follows:

$$\text{Proxy}_{DCeP} = \frac{\sum_{i=1}^{n} (N_i * W_i)}{E_{DC}}$$

where:

- n is the number of instrumented applications running during the assessment window.
- N_i is a normalization factor for each software application.
- W_i is the number of units of useful work reported by a particular instrumented application.
- E_{DC} is the total energy consumed by the data center during the assessment window.

How to measure variables in the numerator:
Wi is measured during an assessment window.

When collecting values for W_i, all applications should either be polled at the beginning and end of the assessment window and the difference calculated, or each application's internal accumulator should be reset to zero at the beginning of the assessment window and the value of accumulated useful work recorded at the end.

Units: normalized tasks / kWh

PROS AND CONS

Pros

- Consistent with the definition of both Data Center Productivity and Data Center Energy Productivity
- Accommodates the fact that what constitutes useful work is a matter of human judgment
- Automates the process of computing aggregate useful work for an entire data center

Cons

- Problem of setting the relative weighting factor of the useful work reported by different applications remains

USE MODELS:

Establish the energy productivity baseline of a data center by using the proxy to estimate DCeP for the data center. Make changes to the set points of equipment or the configuration of IT equipment in the data center and then reassess DCeP using the same set of weighting factors across applications. If DCeP increases, the owner/operator can be assured that the changes made are actually enhancing the energy productivity of the data center.

PROXY PROPOSAL #2 - DCEP SUBSET BY PRODUCTIVITY LINK

DETAILED DESCRIPTION:

This proxy proposes a solution for measuring productivity at the application level by aggregating units of useful work reported by a subset of the IT infrastructure in a data center during an assessment window, scaling this number so that it represents the entire data center, and then dividing the result by the total energy consumed by the data center during the assessment window. The units of work are reported by an API that runs in conjunction with each application running on the subset. The subset can consist of servers, routers, storage, etc.,—any device that runs an application. A software development kit (SDK) developed by Intel Corporation for implementing this API will be made publicly available by The Green Grid.

The SDK, called Productivity Link, can be adapted to any application. It will report data from software that contains native instrumentation for doing so (i.e. Microsoft* Exchange Server), as well as applications that do not. In the case of applications that do not have native instrumentation, the SDK uses other means such as polling operating system (OS) registers or providing a middleware shim to gather data.

The SDK requires the user to implement a five-function API (requiring approximately ten lines of code) which is integrated into the application and writes out counter contents in a standard manner. No external libraries or run-time software is required; it is standalone and becomes part of the application. These counters can be easily read and aggregated to report the productivity of the application. The SDK provides pre-written scripts for many applications and runs on Solaris*, OS-X*, and both 32 and 64-bit Linux* and Windows*.

HOW TO CALCULATE THE PROXY:

The proxy is calculated by multiplying the aggregated useful work completed by a subset of servers during the assessment window by a ratio designed to make the number representative of the entire data center, and dividing this product by the amount of energy consumed by the data center during the assessment window, as follows:

$$Proxy_{Prod\ Link} = \frac{\left(\frac{N_{DC}}{N_{subset}}\right) * \sum_{i=1}^{n} W_i}{E_{DC}}$$

where:

- N_{DC} is the total number of servers in the data center.
- N_{subset} is the number of servers in the subset.
- n is the number of instrumented applications running during the assessment window.
- W_i is the number of units of useful work reported by an instrumented application.
- E_{DC} is the total energy consumed by the data center during the assessment window.

How to measure variables in the numerator:

W_i is automatically reported by the installed API and reported to a central collection point. The proposal identifies two levels of useful work which can be reported and tallied – results and tasks. A reported result might be how many emails were sent by an email server, for example. In some instances it is impossible to directly report a result and tasks must be tallied instead. For example, it is not possible to count how many web pages have been served, but the API could report how many "http gets" were completed, thus providing a

Units: (units) / kWh

Application archetype specific but would be application tasks or unit of performance per kWh. For example, if you were trying to compare email solutions then the units would be emails processed for Exchange Server, Gmail, Lotus Notes, etc. per kWh.

PROS AND CONS

Pros

- Automates the process of computing useful work for an entire data center by compiling data for each of the application archetypes that comprises IT infrastructure components
- Does not prohibit the data center operator from weighting workloads and calculating a total value for the entire data center
- Could be used as a differentiator and promoted as a Green Software feature
- Consistent with The Green Grid definition of data center productivity
- Is ready as a proof of concept demo with examples and comparisons of measurements of the same applications, using both native and non-native instrumentation
- Will provide ISVs with a standard API for application instrumentation

Cons

- Will require ISVs to support the standard API defined in Productivity Link

USE MODELS:

The proxy can be used to establish the energy productivity baseline of a data center. Any changes in the application or supporting IT infrastructure will then show up as a positive or negative impact and can be quantified in a consistent manner.

PROXY PROPOSAL #3 - DCeP SUBSET BY SAMPLE WORKLOAD

DETAILED DESCRIPTION:

This proxy provides an easily obtainable metric that is similar to DCeP. It measures the useful work done in a data center per unit quantity of energy expended. It gives the data center operator a tool to determine how efficiently the data center is running, as well as a way of measuring the benefits resulting from changes in data center configuration or operations.

HOW TO CALCULATE THE PROXY:

The proxy is calculated by obtaining a "useful work number" from an instrumented subset of servers running a sample workload during an assessment window, multiplying this number by a ratio designed to make the number representative of the entire data center, and dividing this product by the total energy consumed by the data center during the assessment window, as follows:

$$\text{Proxy}_{\text{DCePsubset}} = \frac{\left(\frac{N_{DC}}{N_{subset}}\right) * W_{subset}}{E_{DC}}$$

where:

N_{DC} is the total number of servers in the data center.

N_{subset} is the number of servers in the subset.

W_{subset} is the useful work number produced by instrumented software running on a server subset.

E_{DC} is the total energy consumed by the data center during the assessment window.

How to measure variables in the numerator:

W_{subset} is output from a subset of the data center. A small subset of computers, representative of the typical computers running in the data center, runs an instrumented sample workload chosen to represent the typical workload of the data center. The sample workload can be injected and measured on a subset of servers normally operating in the data center, or the measurement can be performed statically on a subset of servers set aside for that purpose. The Green Grid will supply a number of sample workloads, each consisting of compilable source code, and each appropriate to a particular data center archetype. The sample workload runs during a specified time period and automatically outputs the base useful work number, W_{subset}.

Units: work units / kWh

PROS AND CONS:

Pros:

• The metric is similar to polling techniques, and the statistics of polling are familiar and proven

• The technique is also very similar to the standard benchmarking procedures used with applications and systems in the data center environment

• It is straightforward to collect the necessary performance data, and the extrapolation of the data to the data center level is quite simple

• The author anticipates that in statistical studies, the proxy will be a good predictor of the actual DCeP of the data center

Cons:

- Data center operators who are comfortable working with “box” metrics, such as measuring CPU utilization, may not be familiar with benchmarking, although this is increasingly not the case
- An industry organization will need to agree upon and make available standardized compilable source code to be used as a sample workload. The code must include the instrumentation needed to collect the performance data. Like a standard benchmark, this code needs to be portable to different machines. Source code will need to be provided for each data center archetype

USE MODELS:

This metric allows for a convenient and rapid measurement of the effects of making changes to a data center to enhance performance or energy efficiency. Data center operators can obtain a direct indication of the benefit of making changes to a data center such as virtualization or processor scaling by running the metric before and after changes are made.

The metric can be used in three different ways. First, the data center operator can run the metric as part of a pilot project, prior to deployment, to estimate the impact of the changes or additions to the data center. Second, the machines running the subset can be scattered throughout the production data center. Third, some data centers already do continuous testing using injected workloads to benchmark or monitor the health of the data center environment. In this case, the measurement of the proposed proxy can be treated as just another benchmark.

PROXY PROPOSAL #4 - BITS PER KILOWATT-HOUR

DETAILED DESCRIPTION:

Productivity can be viewed as output divided by input. At the highest level, a data center consumes energy and produces a stream of output. This proxy is the ratio of the total bit volume from every outbound router on the data center network divided by the total energy consumed by the data center.

HOW TO CALCULATE THE PROXY:

The proxy is calculated by summing the bits coming out of all outbound routers in the data center during an assessment window and dividing by the total energy consumed by the data center during the assessment window, as follows:

$$Proxy_{bkwh} = \frac{\sum_{i=1}^{k} b_i}{E_{DC}}$$

where:

k is the total number of outbound routers.

b_i is the total number of bits coming out of the ith router during the assessment window.

E_{DC} is the total energy consumed by the data center during the assessment window.

How to measure variables in the numerator:

b_i is summed from traffic statistics at all routers providing outbound traffic from the data centers during the assessment window.

Units: Mb / kWh

PROS AND CONS:

Pros:

- Inexpensive setup cost: less than 1 man-week to set up, in most cases
- Small number of touch points: most data centers have one router for several hundred servers, or 2 to 20 outbound routers. Only these routers have to be monitored
- Inexpensive to measure: once set up, readings are automatic or take minutes to complete
- Easy to understand: metric behaves well when compared to intuition
- Can be measured without interrupting operations
- Accounts for all building infrastructure
- Accounts for all IT loads: servers, storage, networking
- Nothing to install on data center IT systems; probable small change on main data center routers

Cons:

- Uncertainty of "value of bits;" are all bits coming out of a data center the same, with a large enough sample?

USE MODELS:

An interesting characteristic of this proxy is that it can be used at any level of the data center or IT infrastructure. The bit stream from a job on a single computer should be consistent when the same work is completed on a different server. For example, the bit stream produced by reading and replying to 100 emails should be the same whether the mail server is a single blade or a complex of redundant enterprise servers. Similarly, retrieving a group of files from a storage device should produce the same bit stream regardless of whether the storage device is a single disk or a fully redundant Storage Area Network (SAN), or even a robotic tape library.

Even the efficiency of network devices can be measured using this proxy. A stream of bits forwarded by a small router would require less energy than the same stream of bits forwarded by a pair of large redundant routers. The small router would have a higher “bits per kilowatt-hour” metric, implying a more energy efficient system for forwarding the bit stream.

Measuring the energy each of these systems requires to produce the work (i.e., the bit stream) should give a good indication of the efficiency of the system.

The bits-per-kWh metric can be used to show the following data center changes:

- Idle server reduction: Removing idle servers lowers energy consumption without affecting the bit stream, since idle servers produce no outbound bits
- Server consolidation: No change to the bit stream, but reduced energy required to produce the same amount of work
- Tech refresh with high-efficiency power supplies: The bit stream remains unchanged, but the servers require less energy to produce the stream
- Partially filled data center that grows over time: Bits-per-kWh would increase over time as more work was performed inside the data center. More jobs performed in the data center would lead to more bits of output, and while energy consumed would increase, so would the volume of the bit stream. The key to efficient design is driving more bits of output with less IT equipment
- Infrastructure efficiency improvements: In this example, facility energy would decrease while the bit stream remained constant, so bits-per-kWh would increase with improved infrastructure efficiencies

PROXY PROPOSAL #5 - WEIGHTED CPU UTILIZATION – SPECINT_RATE

DETAILED DESCRIPTION:

This proxy produces a weighted number based on CPU utilization that indicates the amount of useful work produced in the data center. It is generic and does not distinguish between the type of work or application present on a system.

HOW TO CALCULATE THE PROXY:

The CPU utilization for each server in the data center is averaged over an assessment window and the clock speed of each server is measured. A benchmark number for each server is determined based on published benchmark results that correspond to a similar server model and configuration. It is suggested that the SPECint_rate2000* and SPECint_rate2006* benchmarks be used as the standard benchmarks for this proxy. These benchmarks should provide suitable data for the mix of servers in most data centers today. For each server, the average CPU utilization is multiplied by the benchmark number for that particular server and a scaling factor based on the actual CPU clock speed and the clock speed of the CPU used in the published benchmark result.. These products are aggregated and the sum is multiplied by the length of the assessment window. This product is then divided by the total energy drawn by the data center during the assessment window, as follows:

$$\text{Proxy}_{\text{SPECint}} = \frac{T * \sum_{i=1}^{n}\left(U_{AVG}CPU_i * B_i * \left(\frac{Clk\_CPU_i}{Clk\_B_i}\right)\right)}{E_{DC}}$$

where:

T is the length of the assessment window.

n is the number of servers being measured.

UAVGCPU$_i$ is the average CPU utilization for the ith server, as reported by the tool used to query the server during the assessment window while it is running a typical workload.†

B$_i$ is the benchmark result for the ith server.

Clk_CPU$_i$ is the nominal clock speed of the CPU in the ith server.

Clk_B$_i$ is the clock speed of the CPU that was used to establish Bi.

E$_{DC}$ is the total energy consumed by the data center during the assessment window.

†Note: Additional accuracy can be gained by substituting the product of instantaneous CPU speed and CPU utilization for machines with adjustable clock speed power management schemes.

How to measure variables in the numerator:
Use standard tools available for querying CPU utilization on production machines.

Units: (jobs) / kWh

PROS AND CONS:

Pros:

- Initial experimentation suggests that this approach is directionally correct
- Data is fairly easy to come by
- This proxy is easy to implement and automate
- It can possibly be used to compare relative productivity of different data center form factors with roughly the same application set

Cons:

- It does not comprehend that portion of CPU utilization that is not directly tied to the application performing the useful work (i.e. operating systems, system management agents, etc.)
- The SPECint_rate2000 and SPECint_rate2006 benchmarks may not be available for all server models to be measured. In some cases, weighting factors may have to be derived from other benchmark reports

USE MODELS:

This proxy can be used as a direct estimate of data center productivity. It can also be used as a utilization metric by comparing the actual useful work measured to the maximum possible if all systems were running at 100% utilization.

The proxy has already been used to calculate the potential benefits to two data center decisions – rate of adoption of virtualization and rate of refresh of hardware.

PROXY PROPOSAL #6 - WEIGHTED CPU UTILIZATION - SPECpower

DETAILED DESCRIPTION:

This proxy utilizes published data from SPECpower* benchmark results, in conjunction with a direct measurement of CPU utilization, to estimate the work efficiency of an individual server or groups of servers. An efficiency number for the entire data center can then be obtained by correlating CPU utilization to server models and their respective SPECpower scores for that level of utilization and calculating the total data center score by population number and mix.

As a proxy, CPU utilization can be keyed against many benchmarks and measured application workloads to develop a representation of performance. SPECpower has been established to represent this type of approach to efficiency rating using developed synthetic workloads such as the current SPEC_ssj2008[11] Java*-based workload (and additional workloads that may be in development).

The SPECpower benchmark measures corresponding energy use across ten points on the load line plus idle power and averages the results to create a SPECpower score. The number of server-side Java operations per second (ssj_ops/sec) is reported for each of the ten CPU utilization levels as shown in the table to the right, which shows a sample result.

The ten increments are defined by SPECpower to be uniform, so multiplying the 100 percent result by CPU utilization will provide an estimate of the ssj_ops/sec for any CPU utilization between 0 and 100 percent. For example, in Figure 1, multiplying 345,089 by 40% gives 138,036, which is just 0.8% lower than the 139,052 figure shown for the 40 percent load level.

The benchmark runs on a wide variety of operating systems and hardware architectures and should not require extensive client or storage infrastructure to run. Ideally there will be synthetic workloads to represent all the different application types found in a data center and additional workloads are planned.

18-NOV-07	QUAD CORE	
Box ID	R6-L3-01	
Platform	SuMic 2U	
CPU Model	5420	
CPU, GHz	2.5	
# CPUs/Cores total	2/8	
Memory	8x 4GB	
# HDD	1 SATA	
#, Type PSU	2x 1200W	
OS	WinSvr 64b	
PS Pwr Mgmt	On	
Load Level %	ssj_ops/sec	Avg watts
100	345,089	572
90	311,587	553
80	281,210	536
70	244,844	515
60	209,155	490
50	173,279	470
40	139,052	448
30	104,083	420
20	69,212	385
10	35,350	348
Idle	0	295
SPECpower Score	380.1	

Figure 1: SPECpower example

HOW TO CALCULATE THE PROXY:

The proxy is calculated by first measuring the average CPU utilization of each server in the data center and multiplying it by the published SPECpower benchmark at the 100 percent load level for that server. The result is then multiplied by a scaling factor to compensate for any difference in CPU speed between the measured server and the server that corresponds to the published benchmark. The numerator is then obtained by summing the final products for all servers and multiplying the sum by the length of the assessment window over which the average CPU utilization was determined. The final proxy is then calculated by dividing this product by the energy consumed by the data center during the assessment window, as shown here:

$$Proxy_{SPECpower} = \frac{T * \sum_{i=1}^{n} \left(U_{AVG}CPU_i * S_i * \left(\frac{Clk\_CPU_i}{Clk\_B_i} \right) \right)}{E_{DC}}$$

where:

T is the length of the assessment window.
n is the number of servers in the data center.
$UAVGCPU_i$ is the average CPU utilization for the ith server, as reported by the tool used to query the server during the assessment window while it is running a typical workload.†
Si is the SPECpower ssj_ops/sec at 100 % server utilization for the ith server.
Clk_CPU_i is the nominal clock speed of the CPU in the ith server.
Clk_B_i is the clock speed of the CPU that was used to establish Bi.
E_{DC} is the total energy consumed by the data center during the assessment window.
†Note: Additional accuracy can be gained by substituting the product of instantaneous CPU speed and CPU utilization for machines with adjustable clock speed power management schemes.

How to measure variables in the numerator:
Use standard tools available for querying CPU utilization on production machines.

Units: (ssj_ops) / kWh

PROS AND CONS:

Pros:

- The numerator of this proxy could be changed to report computer energy instead of ssj_ops to give an estimate of compute energy used at certain data center utilization
- Requires no software to be installed on measured systems
- Based on an available worldwide standard and published results

Cons:

- SPECpower results depend upon manufacturers publishing measurements for their server products

USE MODELS:

Relatively speaking, CPU utilization data required to represent a data center's average server utilization level is easy to obtain. That data can be transposed with SPECpower scores of the server types represented in the data center to create a SPECpower score for the data center when totalled.

PROXY PROPOSAL #7 - COMPUTE UNITS PER SECOND

PROXY PROPOSAL #7 - COMPUTE UNITS PER SECOND TREND CURVE

DETAILED DESCRIPTION:

This proxy is based on a trend that has been defined and published[12] for the purpose of weighting server performance based on the year the server was purchased. The trend curve was derived from insight gained from Moore's Law and data obtained from tracking IT performance over a number of years, and serves as a very useful way to compare and track changes in data center compute power. As a baseline, the performance of a server purchased in 2002 has been set to equal one MCUPS (one Million Compute Units Per Second). The trend curve, as defined, then dictates that for prior and subsequent years, the CUPS value increases seven times every five years.

HOW TO CALCULATE THE PROXY:

To calculate this proxy, servers are grouped by year of purchase. A product is then calculated for each year group by multiplying the MCUPS output for servers of that particular year times the number of servers in that year group times the average utilization of those servers over an assessment window. These products are then summed and the summation is multiplied by the length of the assessment window. The final proxy is then calculated by dividing this product by the total energy consumed by the data center during the assessment window, as follows:

$$\text{Proxy}_{\text{CUPS}} = \frac{T * \sum_{i=m}^{n} \left(7^{\left(\frac{i-2002}{5}\right)} * \text{Num}_i * \text{Util}_i \right)}{E_{DC}}$$

where:

- T is the length of the assessment window.
- m is the year of purchase of the oldest server in the data center.
- n is the year of purchase of the newest server in the data center.
- Num_i = number of servers in data center that were purchased in year i.
- $Util_i$ = average server utilization during the assessment window of the servers in the data center purchased in year i.
- E_{DC} is the total energy consumed by the data center during the assessment window.

As an alternative method of calculating the proxy, data center operators can utilize a simple table-based tool that has been published[13] for this purpose and is shown on the following page. This table provides a means for documenting the calculation of the summation contained in the numerator of the above equation.

For each year, the product D is calculated from A*B*C, the product contained in the summation parenthesis above. The summation is obtained by summing the entries in column D. This summation is then multiplied by T, the length of the assessment window, and divided by E_{DC}.

Year of Server Purchase	Estimated MCUPS per Server	Number of Servers	Average Server Utilization (%)	Total Output (MCUPS)
	A	B	C	D
2000	0.50			
2001	0.75			
2002	1.00			
2003	1.50			
2004	2.25			
2005	3.25			
2006	4.75			
2007	7.00			
2008	10.25			
2009	15.25			
2010	22.50			
			Summation of entries in column D	

Figure 2. Data Center Efficiency Tracking – Sample Template

How to measure variables in the numerator:
Use standard tools available for querying CPU utilization on production machines. If you have implemented virtualization, increase the utilization rates for physical servers hosting virtual machines as appropriate.

Units: MCompute Units / kWh

PROS AND CONS

Pros

- No additional software or hardware needs to be deployed
- Does not require the use of an industry benchmark
- Time to deploy is minimal
- The same scheme can be applied to storage and network devices, as well as servers

Cons

- There is no third party validation for the curve used to project CUPS values
- There is no way to compare servers between manufacturers
- Curves for storage and network devices are not yet available
- Granularity is lower in certain instances; for example, the difference in performance between servers purchased months apart, but in different years may be minimal if anything

USE MODELS:

It is useful to keep a running inventory of all IT assets that are online. This inventory can then be matched to CUPS curves so that CUPS/kW for the data center can be constantly reported. This allows the result of operational changes to the data center to be seen immediately. It also can assist in determining whether assets that have been 100% depreciated are worth keeping online.

PROXY PROPOSAL #8 - OPERATING SYSTEM WORKLOAD EFFICIENCY

DETAILED DESCRIPTION:

This proxy estimates the efficiency of a data center by calculating the number of operating system instances per kilowatt of power. It is a point-in-time measurement, and since the proxy does not directly measure business throughput (e.g. web pages served, ERP transactions performed), it is not a direct measure of the business productivity of the data center. However, it does provide an estimate of the efficiency with which a data center provides a common IT resource – in this case, OS instances.

The proxy is easy and inexpensive to calculate as it requires only two things - the total facilities power used to calculate DCiE (kilowatts) and the total number of OS instances hosted in the data center at the time the facility power is measured. Note that this does not exclude using more tightly scoped measurements to evaluate the energy efficiency of other aspects of the infrastructure - e.g. web server instances/kilowatts, or virtual machines/kilowatts. It is simply that these measures do not evaluate the efficiency of the whole data center and are therefore not discussed here.

HOW TO CALCULATE THE PROXY:

The proxy is calculated by dividing the total number of OS instances running in the facility at the time of the assessment by the total facilities power, as follows:

$$Proxy_{OSW} = \frac{Count_{OS}}{P_{DC}}$$

where:

$Count_{OS}$ is the total OS instance count.

P_{DC} is the average power drawn by the data center at the time of the assessment.

How to measure variables in the numerator:

$Count_{OS}$ is the count of all OS instances running in the data center at the time of the assessment, including those running inside virtual machines. This count should include support workloads, as they may be a substantial part of the total energy usage of the data center.

Units: OS instances / kW

PROS AND CONS

Pros:

- It is a useful high-level estimate of the total energy efficiency of a data center
- It is easy and quick to implement and leverages existing DCiE measurements
- It addresses a key issue facing most IT and facilities managers – the proliferation of underused servers
- It helps model and measure the impact of efficiency initiatives – e.g. server refresh, application consolidation, OS virtualization, and facility modernization
- It encourages energy and materials conservation

Cons:

- It does not extend all the way up the IT stack – i.e. to the application or business level
- It does not stop the proliferation of OS instances - it simply measures how efficiently they are used
- An accurate OS instance count can be difficult to obtain
- It is not valid to compare across deployment models - e.g. consolidation within OS instances cannot be compared to OS virtualization

USE MODELS:

Modeling: This proxy can be used to help with capacity planning. For example, if the proxy value is 0.5 OS instances per kilowatt, and the facility is adding 500 OS instances per year, the net power requirement of the facility will be increasing at 250kW per year. If the facility can only support an additional 200kW, it's clear that something will have to change.

Measurement: This proxy can be used to help assess the energy efficiency of a data center, track its improvement over time, evaluate its response to changes, and improve the accuracy of one's capacity planning models. For example, one can measure the change in energy efficiency before and after a server refresh or virtualization initiative, and compare the predicted savings of an initiative to the actual savings.

HOW TO EVALUATE PROXY PROPOSALS

At this point in the development of the productivity proxy proposals, most evaluations will be subjective in nature. The Green Grid is very interested in eliminating proposals that the industry feels do not show promise before subjecting the remaining candidate(s) to additional research and development.

First, The Green Grid would like an assessment of whether each evaluator would consider using the proxy proposal to estimate data center productivity. A simple yes/no assessment will be asked, and the evaluator should respond as to whether each proxy would be valuable in their own operation.
Second, a more detailed assessment against predefined criteria will be gathered. A matrix of evaluation criteria and proxy proposals will be presented in survey form. Individual raters are asked to 1) state the importance of each criterion listed in the matrix, and 2) use each criterion to compare each proposal relative to DCeP as either better, worse, or the same as the baseline proposal.

DCEP AS A BASELINE FOR EVALUATION

The DCeP metric itself will be used as the baseline reference point for comparison. The following is a brief overview of the formulation for the DCeP metric. For more detail see a prior paper explaining DCeP[14].

DCeP compares the amount of useful work that a data center produces to the amount of energy it consumes to produce this work. Useful work is quantified by aggregating the computational tasks performed by all the computing equipment housed within the data center. This is expressed as:

$$\text{DCeP} = \frac{\text{Useful Work}}{\text{Total Energy Consumed}} = \frac{\sum_{i=1}^{m}\left[V_i * U_i(t,T) * T_i\right]}{E_{DC}}$$

where:

- m is the number of tasks initiated during the assessment window.
- V_i is a normalization factor that allows the tasks to be summed numerically.
- $U_i(t,T)$ is a time-based utility function for each task, where:
  - t is elapsed time from initiation to completion of the task.
  - T is the absolute time of completion of the task.
- T_i = 1 if task is completes during the assessment window, or 0 otherwise.
- E_{DC} is the total energy consumed by the data center during the assessment window.

Note that useful work is defined as the sum over i of all tasks 1 through m initiated within the assessment window, weighted by the product of a value factor and a time-based utility function $U_i(t,T)$. Only tasks that complete during the assessment are counted toward assessing the quantity of useful work completed. Note that if the time-dependent value of tasks is ignored, $U_i(t,T)$ may be set to 1. Also, as a practical matter, tasks of similar types may be assigned the same value V_i. These simplifications reduce the equation for Useful Work to:

$$\text{Useful Work} = \sum_{\text{all task types } j} \left(V_j * \text{count}_j\right)$$

where:

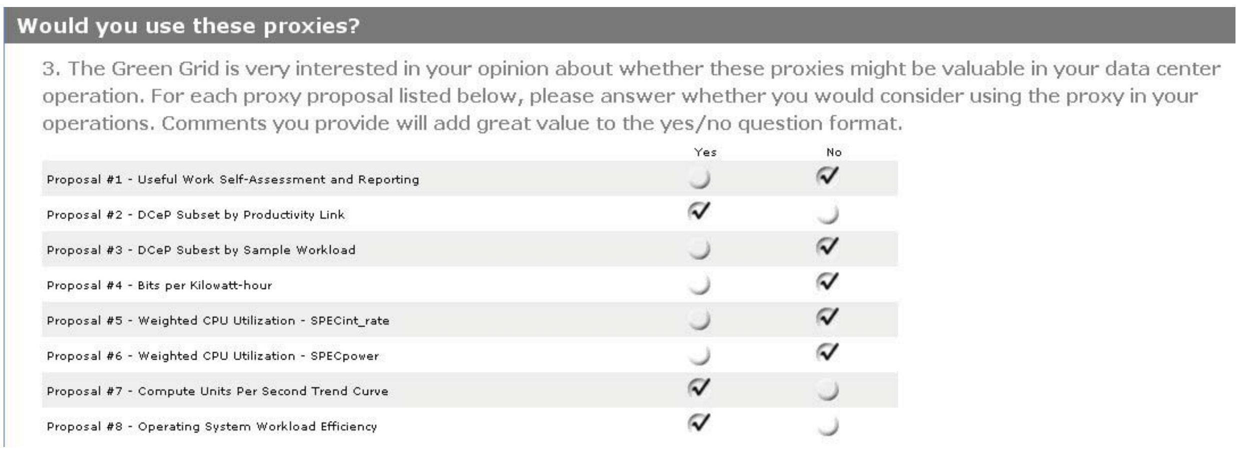
V_j is the relative value of task type j.
$count_j$ is the number of tasks of type j that complete during the assessment window.

FEEDBACK INSTRUCTIONS

The Green Grid has created a web site to collect feedback about the proxy proposals. Readers who want to provide feedback on the proxy proposals should visit:

http://www.surveymonkey.com/s.aspx?sm=JqY74L59ydrBZyDOUrB4wg_3d_3d

The site begins by giving the user the opportunity to provide optional contact information so that The Green Grid can contact the user for follow-up questions or clarification. Then the user is asked the simple question, “Would you use these proxies?” as shown below:


Would you use these proxies?

3. The Green Grid is very interested in your opinion about whether these proxies might be valuable in your data center operation. For each proxy proposal listed below, please answer whether you would consider using the proxy in your operations. Comments you provide will add great value to the yes/no question format.

	Yes	No
Proposal #1 - Useful Work Self-Assessment and Reporting		✓
Proposal #2 - DCeP Subset by Productivity Link	✓	
Proposal #3 - DCeP Subest by Sample Workload		✓
Proposal #4 - Bits per Kilowatt-hour		✓
Proposal #5 - Weighted CPU Utilization - SPECint_rate		✓
Proposal #6 - Weighted CPU Utilization - SPECpower		✓
Proposal #7 - Compute Units Per Second Trend Curve	✓	
Proposal #8 - Operating System Workload Efficiency	✓	

Figure 3. Proxy feedback – Would you use these proxies?

If no other feedback is given, this one question will still provide tremendous value to The Green Grid and the industry. Answer these yes/no questions for each proxy proposal after considering the implications for your own operations. No option for “maybe” was provided to avoid a large volume of middle-of-the-road responses.

Readers are then urged to continue further through the web site and tell The Green Grid how important each of the provided criteria for proxy evaluation is, as shown below:

Importance of Evaluation Criteria

4. How important are each of the criteria below to your selection of a productivity proxy? A description of each of the criteria is provided here (popup windows must be enabled): http://www.thegreengrid.org/news/proxypage1.org/

	Very Important 1	2	3	4	5	6	7	8	9	Not Important 10
Ease of Use	✓									
Accuracy					✓					
Responsiveness			✓							
Cost	✓									
Invasiveness			✓							
Time to Implement							✓			
Completeness					✓					
Objectivity					✓					
Utility					✓					
Commissioning				✓						
Operational ability					✓					

Figure 4. Proxy feedback – Importance of valuation criteria

And finally, the user is asked to compare each proxy to the baseline DCeP metric, based on the criteria they just commented on, as shown below for Proxy Proposal #1:

Proxy Proposal #1 - Useful Work Self-Assessment and Reporting

5. How would you evaluate the "Useful Work Self-Assessment and Reporting" Proxy compared to DCeP?

	Worse than DCeP	The same as DCEP	Better than DCeP
Ease of Use			✓
Accuracy	✓		
Responsiveness		✓	
Cost			✓
Invasiveness		✓	
Time to Implement			✓
Completeness	✓		
Objectivity		✓	
Utility		✓	
Commissioning		✓	
Operational Ability		✓	

Figure 5. Feedback – Proxy Proposal #1

User's providing feedback may want to refer to a summary table of all the proxies which is provided in Appendix A.

TIMING

The feedback form will be available on the web site for at least 90 days after the release of this paper. The final date for feedback will be published on the feedback URL a minimum of 30 days before the end of the feedback period.

FOLLOW-ON WORK

After the industry review period, The Green Grid will assess the feedback and eliminate any proxies that the industry feels do not warrant further research. Those that remain will be subjected to some form of proof-of-concept investigation and experimentation, followed by the publication of a qualitative assessment listing the pros and cons of each proxy. Further work will then involve more quantitative investigation in order to refine a proxy or group of proxies that the industry can use to assess the amount of useful work produced in a data center as a function of energy consumed or power required.

REFERENCES

1 The Green Grid, "The Green Grid Data Center Power Efficiency Metrics: PUE and DCiE," December 2007.
http://www.thegreengrid.org/gg_content/White_Paper_6_-_PUE_and_DCiE_Eff_Metrics_30_December_2008.pdf

2 Ibid., p. 11

3 The Green Grid, "A Framework for Data Center Energy Productivity," 2008, p. 8
http://www.thegreengrid.org/gg_content/White_Paper_13_-_Framework_for_Data_Center_Energy_Productivity5.9.08.pdf

4 Ibid., p. 9

5 U.S. Environmental Protection Agency, "New Window Sticker"
http://www.fueleconomy.gov/FEG/ratingsNewSticker.shtml

6 Ibid., p. 5

7 The Green Grid, "The Green Grid metrics: Data Center Infrastructure Efficiency (DCIE) Detailed Analysis," p. 4, 13
http://www.thegreengrid.org/gg_content/White_Paper_14_-_DCiE_Detailed_Analysis_072208.pdf

8 The Green Grid, "A Framework for Data Center Energy Productivity," 2008, p. 8
http://www.thegreengrid.org/gg_content/White_Paper_13_-_Framework_for_Data_Center_Energy_Productivity5.9.08.pdf

9 Ibid., p. 7

10 Ibid., p. 9

11 Standard Performance Evaluation Corporation, "SPEC Power and Performance," May, 2008,
http://www.spec.org/power_ssj2008/docs/SPECpower-Methodology.pdf

12 Emerson Network Power, "Energy Logic: Calculating and Prioritizing Your Data Center IT Efficiency Actions," 2008,
http://emerson.com/edc/docs/EnergyLogicMetricPaper.pdf

13 Ibid., Appendix B

14 The Green Grid, "A Framework for Data Center Energy Productivity," 2008, p. 8
http://www.thegreengrid.org/gg_content/White_Paper_13_-_Framework_for_Data_Center_Energy_Productivity5.9.08.pdf

APPENDIX A: SUMMARY OF PROXIES

The following table has been included to assist the user when submitting feedback.

PROXY NAME	PROXY FORMULA	VARIABLE DESCRIPTION
Proxy #1 – Useful Work Self-Assessment and Reporting	$\frac{\sum_{i=1}^{n}(N_i * W_i)}{E_{DC}}$	***n*** is the number of instrumented applications running during the assessment window. N_i is a normalization factor for each software application. W_i is the number of units of useful work reported by a particular instrumented application.
Proxy #2 – DCeP Subset by Productivity Link	$\frac{\left(\frac{N_{DC}}{N_{subset}}\right) * \sum_{i=1}^{n} W_i}{E_{DC}}$	N_{DC} is the total number of servers in the data center. N_{subset} is the number of servers in the subset. ***n*** is the number of instrumented applications running during the assessment window. W_i is the number of units of useful work reported by an instrumented application.
Proxy #3 - DCeP Subset by Sample Workload	$\frac{\left(\frac{N_{DC}}{N_{subset}}\right) * W_{subset}}{E_{DC}}$	N_{DC} is the total number of servers in the data center. N_{subset} is the number of servers in the subset. W_{subset} is the useful work number produced by instrumented software running on a server subset.
Proxy #4 - Bits per Kilowatt-hour	$\frac{\sum_{i=1}^{k} b_i}{E_{DC}}$	***k*** is the total number of outbound routers. b_i is the total number of bits coming out of the ith router during the assessment window.
Proxy #5 - Weighted CPU Utilization – SPECint_rate	$\frac{T * \sum_{i=1}^{n}\left(U_{AVG}CPU_i * B_i * \left(\frac{Clk\_CPU_i}{Clk\_B_i}\right)\right)}{E_{DC}}$	***T*** is the length of the assessment window. ***n*** is the number of servers being measured. $UAVGCPU_i$ is the average CPU utilization for the ith server. B_i is the benchmark result for the ith server. Clk_CPU_i is the nominal clock speed of the CPU in the ith server. Clk_B_i is the clock speed of the CPU that was used to establish B_i.
Proxy #6 – Weighted CPU Utilization - SPECpower	$\frac{T * \sum_{i=1}^{n}\left(U_{AVG}CPU_i * S_i * \left(\frac{Clk\_CPU_i}{Clk\_B_i}\right)\right)}{E_{DC}}$	***T*** is the length of the assessment window. ***n*** is the number of servers in the data center. $UAVGCPU_i$ is the average CPU utilization for the ith server. S_i is the SPECpower ssj_ops/sec at 100% server utilization for the ith server. Clk_CPU_i is the nominal clock speed of the CPU in the ith server. Clk_B_i is the clock speed of the CPU that was used to establish B_i.
Proxy #7 – Compute Units per Second Trend Curve	$\frac{T * \sum_{i=m}^{n}\left(7^{\left(\frac{i-2002}{5}\right)} * Num_i * Util_i\right)}{E_{DC}}$	***T*** is the length of the assessment window. ***m*** is the year of purchase of the oldest server in the data center. ***n*** is the year of purchase of the newest server in the data center. Num_i = number of servers in data center that were purchased in year i. $Util_i$ = average server utilization during the assessment window of the servers in the data center purchased in year i.
Proxy #8 - OS Workload Efficiency	$\frac{Count_{OS}}{P_{DC}}$	$Count_{OS}$ is the total OS instance count.

Figure 6. Proxy Summary

