

Power Measurement Tutorial for the Green500 List

R. Ge, X. Feng, H. Pyla, K. Cameron, W. Feng

June 27, 2007

Contents

1	The Metric for Energy-Efficiency Evaluation	1
2	How to Obtain $\bar{P}(R_{\max})$?	2
2.1	The Definition of $\bar{P}(R_{\max})$	2
2.2	Deriving $\bar{P}(R_{\max})$ from Unit Power	2
2.3	Measuring Unit Power	3
3	The Measurement Procedure	3
3.1	Equipment Check List	4
3.2	Software Installation	4
3.3	Hardware Connection	4
3.4	Power Measurement Procedure	5
4	Appendix	6
4.1	Frequently Asked Questions	6
4.2	Resources	6

1 The Metric for Energy-Efficiency Evaluation

This tutorial serves as a practical guide for measuring the computer system power that is required as part of a Green500 submission. It describes the basic procedures to be followed in order to measure the power consumption of a supercomputer.

A supercomputer that appears on The TOP500 List can easily consume megawatts of electric power. This power consumption may lead to operating costs that exceed acquisition costs as well as intolerable system failure rates. In recent years, we have witnessed an increasingly stronger movement towards energy-efficient computing systems in academia, government, and industry. Thus, the purpose of the Green500 List is to provide a ranking of the most energy-efficient supercomputers in the world and serve as a complementary view to the TOP500 List.

However, as pointed out in [1, 2], identifying a single objective metric for energy efficiency in supercomputers is a difficult task. Based on [1, 2] and given the already existing use of the “performance per watt” metric, the Green500 List uses “performance per watt” (PPW) as its metric to rank the energy efficiency of supercomputers. The “performance per watt” metric is defined as:

$$PPW = \frac{\text{Performance}}{\text{Power}} \quad (1)$$

Because there are many possible meanings for both performance and power in Equation (1), the Green500 List explicitly specifies that (1) performance is defined as the achieved maximal performance by the Linpack

benchmark on the entire system, denoted as R_{\max} , and (2) power is defined as the average system power consumption during the execution of Linpack with a problem size that delivers R_{\max} , denoted as $\bar{P}(R_{\max})$.

In this tutorial, we use **GFLOPS** (Giga Floating-point Operations per Second) as the unit of R_{\max} , **Watt** as the unit of $\bar{P}(R_{\max})$, and **GFLOPS Per Watt** as the unit of “performance per watt.” Thus, we have

$$\text{GFLOPS Per Watt} = \frac{R_{\max} \text{ (in GFLOPS)}}{\bar{P}(R_{\max}) \text{ (in Watt)}} \quad (2)$$

To make a submission to the Green500 List, a given supercomputer site needs to provide both R_{\max} and $\bar{P}(R_{\max})$ for their supercomputers. While the guide for measuring R_{\max} is available through the TOP500 web sites, this tutorial describes how to properly obtain $\bar{P}(R_{\max})$.

2 How to Obtain $\bar{P}(R_{\max})$?

2.1 The Definition of $\bar{P}(R_{\max})$

As mentioned in §1, we define $\bar{P}(R_{\max})$ as the total power consumption of all compute nodes of the supercomputer averaged over the execution of Linpack with a problem size delivering the peak performance R_{\max} . For this definition, we emphasize the following points:

- All references to power in this tutorial mean AC RMS (Root Mean Square) power.
- We limit the system power to the total power consumed by all compute nodes involved in computing the benchmark, excluding the power consumed by the storage nodes and front nodes, if applicable.
- As the instant power of computer systems varies over time, a relative long LINPACK execution time is required in order to minimize the measurement variance.
- The ratio of **GFLOPS Per Watt** may change with system scale (i.e, number of nodes used), benchmark, and problem size. Therefore, we explicitly specify that the $\bar{P}(R_{\max})$ is for the same system scale and the same Linpack problem size which the supercomputer delivers the maximum Linpack performance R_{\max} .

2.2 Deriving $\bar{P}(R_{\max})$ from Unit Power

A supercomputer on the TOP500 List will typically consist of a large number of identical units, where a unit is defined as a node, chassis, rack, cabinet, or even the entire supercomputing system. Unless the total system power is available from a centralized power management system, one can expect that directly measuring the total power consumption of an entire supercomputing system would be a challenging task. Thus, a simple solution is to derive the total system power $\bar{P}(R_{\max})$ from the power consumption of a single unit. Here the unit refers to the minimum device that is powered by an external power strip. Assuming there are N identical units involved in computing the Linpack benchmark and each unit consumes $\bar{P}_{unit}(R_{\max})$, then we can derive an estimate for $\bar{P}(R_{\max})$ as

$$\bar{P}(R_{\max}) = N \cdot \bar{P}_{unit}(R_{\max}). \quad (3)$$

Equation (3) assumes the following: (1) the computational workload during the Linpack benchmark is well-balanced across all units, and (2) all units are identical and consume the same amount of power for the same workload. These two assumptions hold for most situations, though it is a good practice to measure several units and use their average value as the $\bar{P}_{unit}(R_{\max})$.

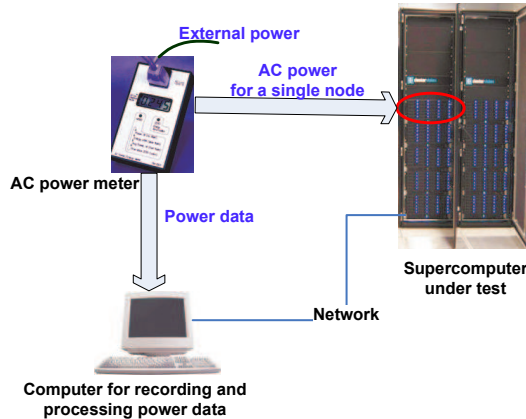


Figure 1: The power measurement for a single unit on the supercomputer under test.

2.3 Measuring Unit Power

There are multiple ways to measure the unit power consumption of a supercomputer. In practice, we have used the following three methods:

- Via a power meter
- Via a combination of current probe and multimeter
- Via a power management solution

Via a power meter. A power meter reporting the RMS power value is probably the simplest way to measure the power consumption of a single unit. Figure 1 illustrates the set-up of measuring the unit power via a power meter with PC interface. We plug a power meter between the power supply’s AC input of a selected unit and a socket connected to the external power supply system. We record the readings of the power meter using an additional PC. Optionally, we can synchronize the power meter measurement to supercomputer workload execution using certain communication protocol or scripts.

Via a current probe and a volt meter. When the unit power exceeds the upper limit of the power meter, we can use a current probe as an alternative approach. A current probe converts current to voltage that can be read by a digital multimeter. An advantage of an AC current probe is that it does not require disconnecting and reconnecting the unit power supply, i.e., no need to power off/on the system to be measured. A disadvantage is that the AC probe has to be calibrated to generate the correct current measurement. The power consumed by the unit can be calculated by:

$$P = V \cdot I \quad (4)$$

Here I is the electrical current measured with the current probe, V is the supply voltage.

Via a smart power strip. Some companies provide power strips or power distribution units that include power sensing and metering functions. These devices also provide a RS232 or Ethernet-based interface to a management interface. To measure the power consumption of one or several units of a supercomputer, we can simply replace an existing power strip with a smart power strip, and then record the power consumption of the supercomputer units under an appropriate workload.

3 The Measurement Procedure

In this section, we show the procedure of using a power meter to measure the average power consumed by a single unit $\bar{P}_{unit}(R_{max})$ of a supercomputer under test. Here, we use the *Watts up? Pro*¹ power meter

¹The product manual and user guide of *Watts up? Pro* are available online at <http://www.wattsupmeters.com>.

as an example. However, we note that the measurement procedure describe below can be applied to any kind of digital power meter that has a PC interface.

As shown in Figure 1, we plug a single unit of the supercomputer into a *Watts up? Pro* power meter and the meter will show the wattage of the power consumed by the the unit. The power data logged by the *Watts up?* can be downloaded to a Windows or Linux machine via an RS232 or a USB interface. In the following, we describe the detailed measurement procedure.

3.1 Equipment Check List

To make a submission to the Green500 List, the following equipment or devices are needed for the power measurement:

- A. **The supercomputer** to be submitted to the Green500 List;
- B. **A digital power meter with PC interface**;
- C. **A PC or workstation for data logging and processing**² with RS232 or USB interface;
- D. **The power meter software interface to the PC**;
- E. **The Linpack benchmark**.

3.2 Software Installation

3.2.1 Install the Linpack benchmark (item E) on the supercomputer (item A)

The power measurement for Green500 List requires running the Linpack benchmark on the supercomputer. For most cases, the LINPACK should be already available on the supercomputer. Otherwise, users are encouraged to consult the TOP500 web site or the Linpack web site (<http://www.netlib.org/benchmark/hpl/>).

3.2.2 Find a problem size for the Linpack benchmark that achieves R_{\max}

The problem size of the Linpack benchmark can be changed through the input file *HPL.dat*. Run the Linpack benchmark for different problem sizes to find out an appropriate size that your supercomputer can achieve the highest Linpack performance, represented as R_{\max} . Save this input file as the bench input for later power measurement.

3.2.3 Install the power meter software interface (item D) on the data logging PC (item C)

Connect the power meter to the PC via an USB or an RS232 interface. Then plug the power meter into an electric socket.

If you are using the *Watts Up? Pro* power meter, you will find the software and manual in the product package. Following the instruction to get the power meter software interface installed on the logging PC.

3.3 Hardware Connection

Connect the power meter, the measured unit of the supercomputer, and the logging PC as shown in Figure 1. **Note: Power off all affected hardware prior to connecting them.**

²We demonstrate the data logging and processing on Windows OS in this document. Linux OS can also be used with a Linux version of the *Watts up? Pro* computer interface.

3.3.1 Select the unit of measurement

Measuring the power consumption of the entire supercomputer is difficult. So, we can measure part of the supercomputer and then multiply by a constant to derive the power of the entire system. The unit of the measurement could be one node, one chassis, or one rack. One might consider the consistency, the location, and the convenience when choosing the unit of power measurement. (Please note that the unit must be compute unit.)

3.3.2 Connect the unit of measurement to the power meter

Power off the selected unit, following the normal system administration procedure. After the unit is completely powered off, unplug the selected unit from the power strip and plug it to the outlet on the *Watts up? Pro* power meter.

When the unit has multiple power supplies or you want to combine several basic units, you can use a power strip to connect the unit to the power meter.

3.3.3 Connect the power meter to logging PC using a RS232/USB cable

Depending the type of your power meter, you could use either USB or RS232 as the communication interface between the power meter and the logging PC.

3.3.4 Connect the the power meter to an available AC outlet

At this point, you can power on all affected equipment and be ready to start the power measurement.

3.4 Power Measurement Procedure

Once both the software and hardware are properly set up, the power measurement can be done in the following seven steps:

1. **Launch the power meter (*Watts up? Pro*) data logger software** on the logging PC/workstation, and click the “receive data” tab on the opened window.
2. **Launch the Linpack benchmark** with the input file saved at 3.2.2, for which the supercomputer should deliver the highest Linpack performance R_{\max} .
3. **Start recording the power measurements samples** by pressing the “request data” button on the data logger window (on the logging PC/workstation) immediately after step 2.
4. **Stop recording the power measurement samples** by pressing “Stop” button on the data logger window (on the logging PC/workstation) when the execution Linpack benchmark is finished.
5. **Save the Linpack performance** for current run as R_{\max} .
6. **Calculate the unit average power** from the recorded power measurement data recorded between step 3 and 4. Save this value as $\bar{P}_{unit}(R_{\max})$. You can remove the first and last several samples
7. **Derive the GFLOPS Per Watt metric** $\bar{P}(R_{\max})$ using Equation 3.

To minimize the measurement variance, repeat the above measurement procedure at least three times, and submit the average value of the derived $\bar{P}(R_{\max})$ to the Green500 web site (<http://www.green500.org>).

4 Appendix

4.1 Frequently Asked Questions

1. What are the differences between $\bar{P}(R_{\max})$, rated power, and instant power?

The $\bar{P}(R_{\max})$ is an average power consumed by the compute nodes when running the Linpack benchmark with a problem size that the supercomputer can deliver its highest performance in GFLOPS. This is the power that the Green500 List requires.

The rated power is a limiting value that is shown on the equipment tag. It is usually much higher than the actual power consumed by a computer system.

Instant power is the power that an electric device consumes at a specific time point. As the instant power of a computer system varies significantly over time and/or with different workload, instant power can not reflect the power-performance efficiency of the tested computer system.

2. Why $\bar{P}(R_{\max})$ has to be measured when the supercomputer is running Linpack with a specific problem size?

There is no ideal power-performance efficiency metric for supercomputers. Since the TOP500 list rank supercomputer using R_{\max} , as a tradeoff, the Green500 list specifies $\bar{P}(R_{\max})$ as the corresponding power consumption when the system achieves its highest Linpack performance. It is easy to understand that the value of $\bar{P}(R_{\max})$ will increase when the Linpack benchmark is running on a smaller number of nodes on the same system. To make the **GFLOPS Per Watt** metric is “fair” to the top systems, we enforce the special constraints for measuring $\bar{P}(R_{\max})$.

3. Each unit of our supercomputer has multiple power supplies, how should we measure the unit power?

You can plug a multi-outlet power strip between the unit and the power meter. The power strip can aggregate the power consumption of all power supplies connected to the measured unit.

4. The Linpack benchmark finishes too quick, how to make the power measurement accurate?

The easy way is to repeat the benchmark several times to ensure the total execution time is larger than 10 minutes.

5. The power consumption of each measured unit exceeds the maximum range supported by the power meter?

You can use current probe to measure the current and then derive the power from the product of current and voltage.

4.2 Resources

- The Green500 List web site: <http://www.green500.org>.
- The TOP500 Supercomputer web site: <http://www.top500.org>.
- The EPA Energy Star program: <http://www.energystar.gov/>.
- Workshop on High-Performance, Power-Aware Computing: <http://hppac.cs.vt.edu/>.

References

- [1] Chung-Hsing Hsu, Wu chun Feng, and Jeremy Archuleta. Towards efficient supercomputing: A quest for the right metric. In *The 1st IEEE Workshop on High-Performance, Power-Aware Computing*, 2005.
- [2] Sushant Sharma, Chung-Hsing Hsu, and Wu chun Feng. Making a case for a green500 list. In *The 2nd IEEE Workshop on High-Performance, Power-Aware Computing*, 2006.