PowerMon 2: Fine-grained, Integrated Power Measurement

TR-09-04

Daniel Bedard, Robert Fowler,
Min Yeol Lim, Allan Porterfield

October 5, 2009

RENCI Technical Report Series
http://www.renci.org/techreports

PowerMon 2: Fine-grained, Integrated Power Measurement

RENCI Technical Report TR-09-04

Daniel Bedard
danb@renci.org
Renaissance Computing Institute

Robert Fowler
rjf@renci.org
Renaissance Computing Institute

Min Yeol Lim
mylim@renci.org
Renaissance Computing Institute

Allan Porterfield
akp@renci.org
Renaissance Computing Institute

October 5, 2009

Abstract

We describe version 2 of RENCI PowerMon, a device that can be inserted between a computer power supply and the computer's main board to measure power usage at each of the DC power rails supplying the board. PowerMon 2 provides a capability to collect accurate, frequent, and time-correlated measurements. Since the measurements occur after the AC power supply, this approach eliminates power supply efficiency and time-domain filtering perturbations of the power measurements. PowerMon 2 provides detail about the power consumption of the hardware subsystems connected to each of its eight measurement channles. The device fits in an internal 3.5" hard disk drive bay, thus allowing it to be used in a 1U server chassis. It cost less than $150 per unit to fabricate our small quantity of prototypes.

1 Introduction

Power consumption has gained significant influence as a performance criterion for server-class computer systems. As the cost of energy increases and the cost of equipment decreases, operating costs of a computer system can exceed the system's initial purchase price. Furthermore, the energy consumed by idle home computers has grown to a significant fraction of national energy use. The topic of energy consumption in computing receives an increasing amount of attention from researchers, vendors, and the end users of computing systems.

To address the power and energy problem in computing, manufacturers are developing new standards to motivate the development of more energy efficient hardware [1, 2]. The academic research community and system operators have studied the evaluation and modeling of power consumption, as well as programming and operating strategies designed to improve energy efficiency [3, 4, 5]. In order to study and improve energy efficiency most effectively, however, researchers need to measure computer power consumption. There are many devices on the market designed to measure power consumption of electronics. At the low end, devices such as the WatchDog.com PowerEgg [6] and the Watts Up? [7] line of products can be used to measure power consumption and other parameters, such as power factor. These devices are inserted in the AC line between the device and an output, presenting measurements on an integrated display, a serial or USB port, or through an integrated web server [8].

Since motherboards require DC power, a power supply is used to convert incoming AC power to DC "rails" of several different voltages. AC-to-DC power supply conversion efficiency varies with load (though this can be shown to be a roughly linear relationship for some supplies [9]). Furthermore, because the DC output is buffered and filtered for stability, fast variations in the DC load do not translate to corresponding variations on the AC side of the supply. These effects make it difficult to evaluate the energy impact of variations due to programming or compilation experiments. For example, in the case of a quad-core 3-GHz processor that can issue three instructions per clock cycle, a sensor capable of one measurement per second has a process resolution of 36 billion instructions! To evaluate the effects of programming changes, it is therefore important to be able to measure energy consumption as accurately and at as fine a granularity as possible.

Currently, the most common form factor for power supplies is known as ATX. Each system board uses a design-specific strategy to power the CPUs and other devices on the board from the supply rails. Power consumption by processors, memory, interface buses, and disk drives vary depending upon the activity required by these devices.

RENCI PowerMon was developed specifically to provide frequent, accurate measurements of power consumed by the individual subsystems of a motherboard, bypassing effects of the power supply on those measurements. The device is connected between an ATX power supply and a motherboard and reports voltage and current measurements on up to eight channels via a PC's USB port. After experiments with an initial prototype, we designed PowerMon 2 to increase measurement frequency by a factor of sixty. The new device also fits well into a 3.5 inch internal hard drive bay, and it uses a more commonly available circuit board substrate than the previous version.

The use of PowerMon 2 in a Ph.D. research project is reported in [9].

2 Measurement Hardware

PowerMon 2 allows measurement of eight individual channels. This permits the measurement of the five rails used to power the components on a motherboard, in addition to peripherals such as hard drives and supplemental supplies to video cards. PowerMon 2 is inserted in the computer system by plugging the ATX power cables into input connectors on the device and then connecting the outputs via power jumper cables to the motherboard.

An Atmel ATmega168 8-bit microcontroller is at the heart of PowerMon 2. The microcontroller communicates with the host computer, interfaces to PowerMon 2's eight current and voltage sensors, and schedules, formats, and timestamps measurements. This microcontroller was chosen for its built-in I^2C module, as well as a built-in serial USART [10]. These hardware modules permit interrupt-driven communication to the host and to the sensors, greatly simplifying the microcontroller code. While the microcontroller runs its program code using an internal 8MHz RC oscillator as a timebase, an inexpensive 32 kHz watch crystal was added to allow precise timestamping of the data.

USB was chosen as the host interface to PowerMon 2 because of its speed and ubiquity. An FTDI FT232R IC is used to provide a USB interface to the microcontroller's USART [11]. The FT232R appears as a virtual serial port on the host PC and allows direct access to the ATmega168.

Voltage and current are detected using an Analog Devices ADM1191 digital power monitor IC on each power rail [12].The ADM1191 contains a 12-bit analog-to-digital converter, a current-sense amplifier, and an I^2C transceiver. Connected across a 5-milliohm sense resistor on each power rail, this IC provides a digital measurement of a channel's voltage and the current traversing the resistor. The eight ADM1191s communicate as slaves on a common two-wire I^2C bus mastered by the ATmega168. Each chip's address is determined by configuring its two address pins with a combination of resistors.

3 Hardware Design Considerations

Several important design decisions factored into the development of PowerMon 2. The ADM1191 uses a low-valued resistor as its current sense element. We also considered using Hall effect sensors, which pass the current through a magnetic field and measure the voltage induced by charge deflection orthogonal to the current flow and magnetic field. Hall effect sensors have the advantage that they do not introduce a significant resistance into a circuit, so drawing a large current does not result in a large voltage sag [13]. We experimented with Hall effect sensors, but they required external amplifiers and manual trimming for electrical and magnetic offsets. Since the ADM1191 provides a pre-trimmed amplifier, an ADC, and an I^2C transceiver, the resultant savings in parts count and complexity made it a more appropriate fit for the project. The sag induced by PowerMon 2's five-milliohm resistor at the maximum current output of the tested power supply is still well within the ATX specification.

The desired form factor for this project drove several design decisions. To make the device functional while installed in a 3.5" drive bay, it was designed so all of its connectors would be mounted on one edge of the circuit board. This was accomplished by using surface mount connectors of a higher density than the standard 24- and 12-pin ATX connectors used by the power supply. Input and output connectors were mounted on the top and bottom of the board, with large vias connecting power traces.

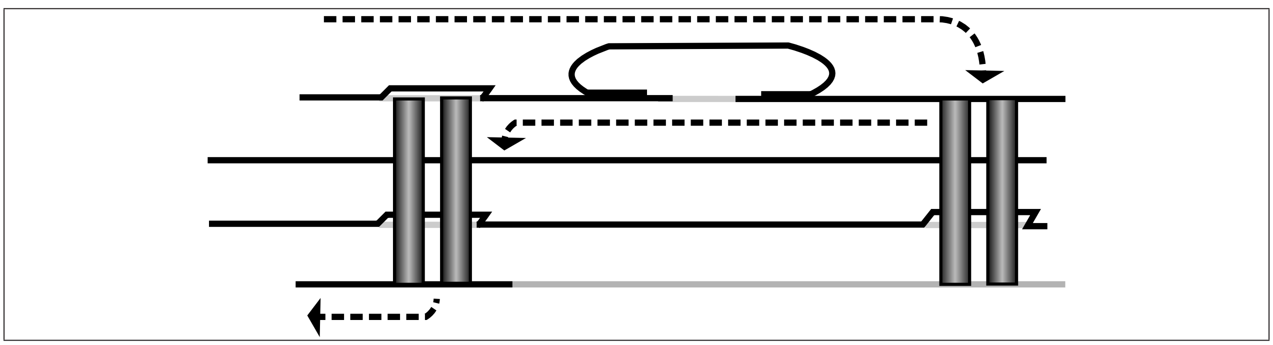


Figure 1: *Cross section of an individual power rail. Current flow is represented by dashed arrows. Current enters the board through a connector (not shown), traverses a power resistor, flows through a via to an internal copper layer, then follows another via to the bottom layer of the board, where it exits through another connector (not shown).*

To save on fabrication cost, it was highly desirable to fabricate the board from standard FR4 substrate with a standard 1 oz. / sq. ft. copper thickness. In order to accommodate currents of up to 10 Amps without generating too much heat, traces had to be relatively large. The traces were designed to be .220" wide on external surfaces, permitting currents of up to 11.3 Amps with a 20C rise. However, the width of the connectors did not permit enough space for 8 current sense resistors stacked side-by-side, so the traces were alternated between top and bottom of the board, and it was necessary to run return traces on inner layers of the board. The inner traces are .490" wide, allowing 10.1 Amps with a 20C rise. The final design contains four layers of 1-oz. FR4 substrate. The calculated net resistance of each channel, including the 5-milliohm resistor in series, is 8.31 milliohm.

4 Firmware

In the initial prototype, PowerMon, the host PC encapsulated I^2C transactions inside USB transactions. The additional latency incurred while decoding multiple USB transactions, and the need for the kernel on the host PC to handle measurement scheduling and USB encoding, prohibited accurate correlation between processor events and measurement results. For PowerMon 2, the firmware was completely rewritten to maximize measurement throughput while maintaining much of the flexibility, in terms of polling configuration, of direct I^2C calls. In PowerMon 2, the entire I^2C transaction is performed in the hardware of the ATmega168. Device-host interaction consists of a configuration and an output mode. In configuration mode, the active sensors and sample rate are configured using a command set made up of ASCII characters. In output mode, PowerMon 2 returns a series of 32-bit words containing sensor data and timestamps. Timestamps are interlaced with the sensor data each complete second during data collection. More detailed information on the command set and measurement format is contained in Appendix A and Appendix B.

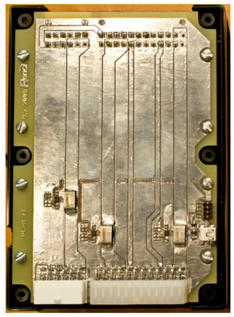
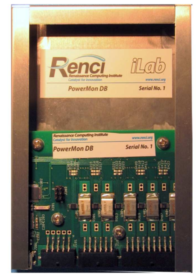
PowerMon is configured to communicate over the USB port at 1 Mbps. It is capable of collecting and transmitting over 3000 readings per second.

5 Software

A sample application has been written to demonstrate PowerMon 2's functionality using the Linux termio API. The application takes as arguments a serial port ID, a sensor mask, a sample period, and a number of samples to collect. The application communicates with PowerMon 2 and applies the necessary coefficients to convert the sensor readings to voltages and currents. It outputs the results to the console, which can be easily redirected to a file on-disk.

6 Results

PowerMon 2 implements current and voltage sensing at a rate of up to 3000 samples per second for eight individual 10-Amp DC rails in the form factor of a 3.5" hard disk drive. The total cost is less than $150 each when fabricated in quantities of 5 or more. PowerMon 2 is a useful tool for understanding workstation power usage, as it demonstrates a need for and a mechanism to provide power measurement closer to the point of load on PC motherboards than previously possible.

	PowerMon	PowerMon 2
		
Dimensions (L x W x H)	17.5cm x 12.4cm x 3.81cm	15.2cm x 10.2cm x 2.5cm
Measurement Channels	6	8
Voltage Accuracy	±0.90%	±0.90%
Current Accuracy	-6.6% / +6.8%[1]	-6.6% / +6.8%[2]
Timestamping	No	Yes
Measurement Frequency	up to ~50 Hz, not hardware-configurable	up to 1024 Hz per channel, up to 3072 Hz aggregate
Parts Count	21	26
PCB Fabrication Cost	$19.65[3]	$64.80[4]
Total Cost	$102.34	$147.07

[1]Worst case presented. Current accuracy varies depending on the current sensed. See ADM1191 datasheet for details.
[2]Worst case presented. Current accuracy varies depending on the current sensed. See ADM1191 datasheet for details.
[3]Batch quantity 8.
[4]Batch quantity 10.

References

[1] Calwell, Chris. "80 Plus: A Strategy for Reducing the Inherent Environmental Impacts of Computers." IEEE International Symposium on Electronics and the Environment (2005): 151. Web.

[2] ”Climate Savers Computing Initiative - www.climatesaverscomputing.org”, a white paper on smart and efficient computing by Intel and Google. http://www.intel.com/technology/docs/whitepaper_climate-savers-intiative.pdf

[3] Isci, C., and M. Martonosi. "Runtime Power Monitoring in High-End Processors: Methodology and Empirical Data". Proceedings of the 36th annual IEEE/ACM International Symposium on Microarchitecture. google. Web.

[4] Lien, Chia-Hung. "Web Server Power Estimation, Modeling and Management." Proceedings - 2006 IEEE International Conference on Networks, ICON 2006 - Networking-Challenges and Frontiers 2 (2006): 541. Web.

[5] Lien, Chia-Hung. "Estimation by Software for the Power Consumption of Streaming-Media Servers." IEEE transactions on instrumentation and measurement 56.5 (2007): 1859. Web.

[6] IT Watchdogs PowerEgg. http://www.itwatchdogs.com/DataSheets/PowerEgg(8.5x11)033006.pdf

[7] Electronic Educational Devices. https://www.wattsupmeters.com/secure/products.php

[8] IT Watchdogs. http://www.itwatchdogs.com/

[9] Min Yeol Lim, Daniel Bedard, and Allan Porterfield. Fine-grained and accurate power measurement.

[10] Atmel ATmega168 datasheet. http://www.atmel.com/dyn/resources/prod_documents/doc2545.pdf

[11] FTDI FT232R datasheet. http://www.ftdichip.com/Documents/DataSheets/DS_FT232R_V204.pdf

[12] Analog Devices ADM1191 datasheet. http://www.analog.com/static/imported-files/data_sheets/ADM1191.pdf

[13] Allegro Hall-Effect Based Current Sensor ICs. http://www.allegromicro.com/en/Products/Design/Allegro%20Hall-Effect%20Based%20Current%20Sensor%20ICs.pdf

A Command Set

PowerMon 2 is configured with ASCII commands sent while the device is in configuration mode. Each command consists of a lower case letter, followed by a space, followed by an argument. In some cases, a space and a second argument follow the first argument. In every case, the arguments are input in decimal base. Each command must be terminated by a newline character (0x0A). PowerMon 2 acknowledges each command with a command-specific response, followed by “OK”, followed by a carriage return and line feed (0x0D 0x0A). The following commands are accepted:

t <*32-bit time*><*linefeed*> Sets internal clock to *time*. The timestamp will be offset from this setting. This could be used to synchronize PowerMon 2 with the host’s system clock.
Returns “T=<*time*><*carriage return*><*linefeed*>”.

s <*16-bit sample period*> <*32-bit number of samples*><*linefeed*> Sets a trigger. Once the trigger is enabled (see “e” below), samples will be collected at an interval of period / 1024. Therefore, for a period of 1, samples are collected every 1/1024 seconds. For a period of 2, every 1/512 seconds. The number of samples collected before the trigger expires is reflected in the second argument. This can be set to zero to allow for the indefinite collection of samples. Returns “S=<*period*>,<*samples*><*carriage return*><*linefeed*>”.

m <*16-bit sensor mask*><*linefeed*> Sets collection mask. This is a 16-bit integer with a one in the bit field of each sensor for which data will be collected. The board has eight sensors, numbered 0-7. For example, to collect data from all the sensors, the mask would be 0000 0000 1111 1111, in binary, or 255 in decimal (which is the appropriate syntax). To collect data only from sensors 0 and 2, the user would input “*m 5*”. Returns “M=<*mask*><*carriage return*><*linefeed*>”.

e<*linefeed*> Enables data collection. After this command is executed, the device collects data from the sensors and outputs the data. When the output is complete, the device re-enters configuration mode, returning “OK<*carriage return*><*linefeed*>”. Data output is terminated once the requested number of samples has been collected, or once a “d” is transmitted during output mode.

d Terminates output mode. Sending a “d” (no newline required) terminates output mode, entered by the “e” command. The device will empty its sensor data queue and return to interactive mode, returning “OK<*carriage return*><*linefeed*>”.

B Measurement Format

During data collection mode, the device returns a series of 32-bit words representing sensor data and timestamps. Timestamps are interlaced with the sensor data each complete second during data collection. Data is transmitted in network byte order (most significant bits first). The format for the data follows:

bit 31 Overflow flag: if more sensor data is collected than can be processed for transmission, the overflow flag is set to 1. The flag is not cleared until the end of the output sequence.

bit 30 Timestamp flag: this flag is set to 1 for timestamps only *if the timestamp flag is set*

bits 29..0 These thirty bits represent the number of seconds since the device was powered on, or the sum of an offset and the number of seconds since an offset was set. This allows a maximum uptime of ~34 years before timestamps roll over.

if the timestamp flag is not set

bit 29 Done flag: this flag signals that the data collection buffer of the device is empty and the trigger is no longer active, (all lower bits should be 0).

bit 28 0 (zero).

bits 27..24 The bus number of the sensor from which a reading was collected (i.e., sensor number 0 through 16).

bits 23..16 Eight highest order bits of the voltage reading.

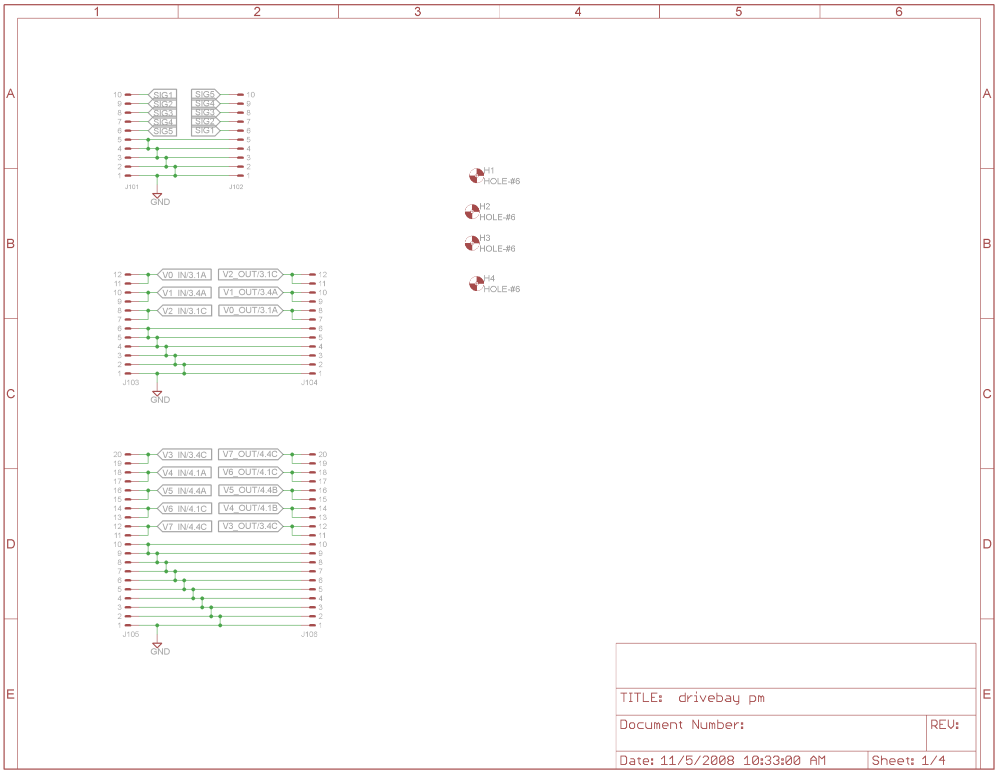
bits 15..8 Eight highest order bits of the current reading.

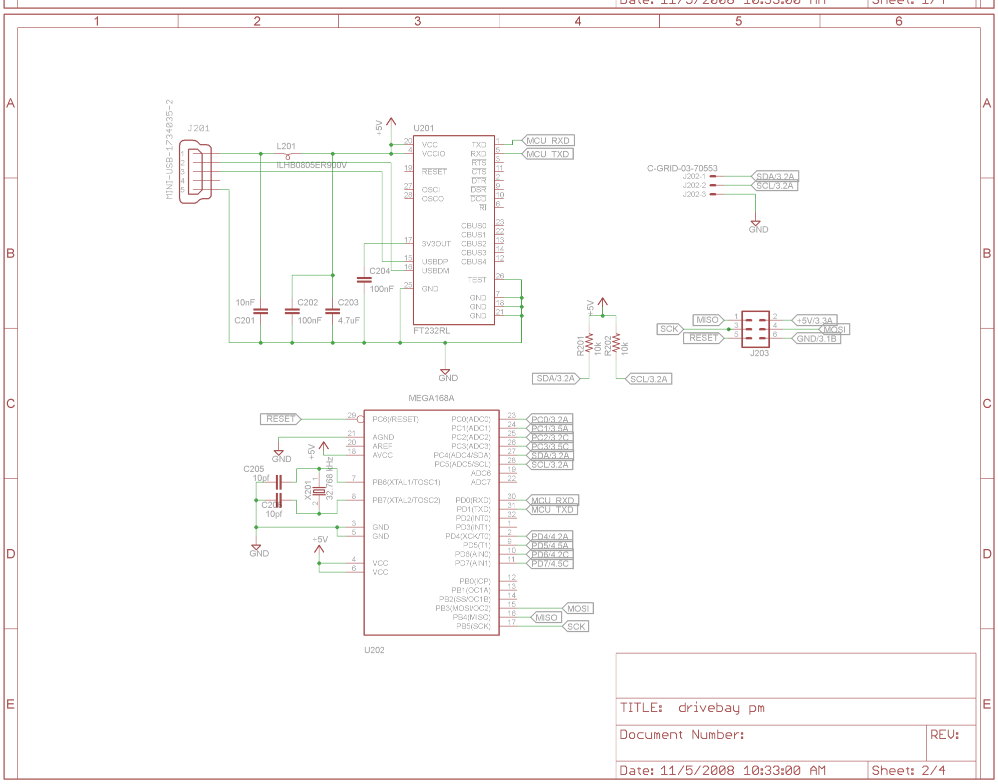
bits 7..4 Four lowest order bits of the voltage reading.

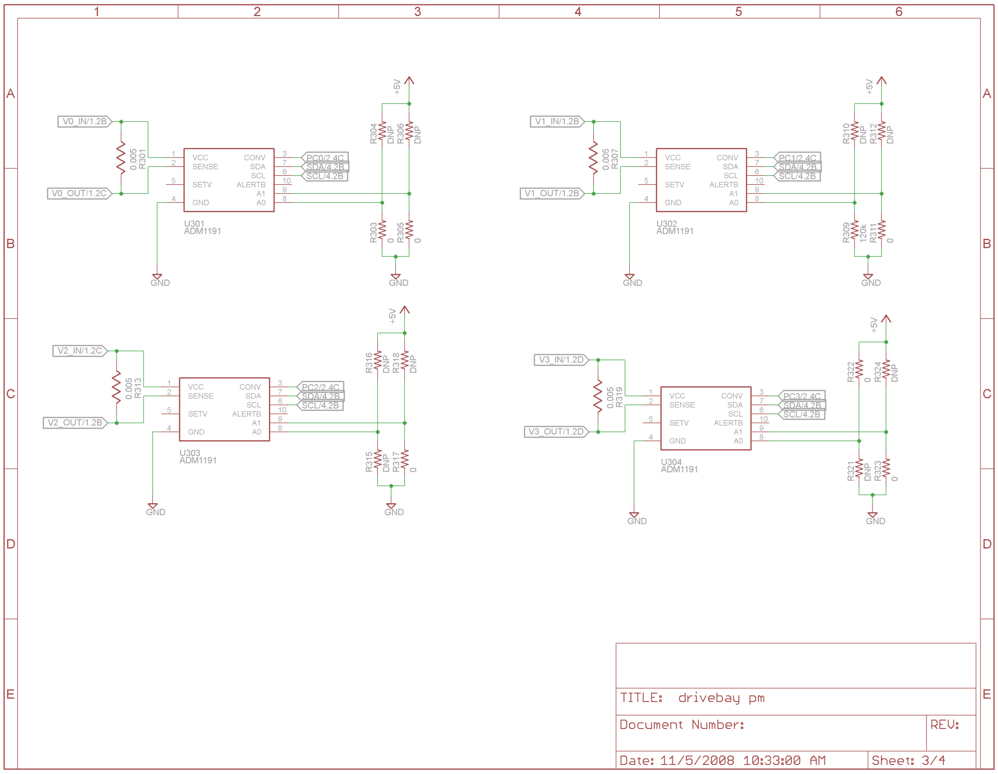
bits 3..0 Four lowest order bits of the current reading.

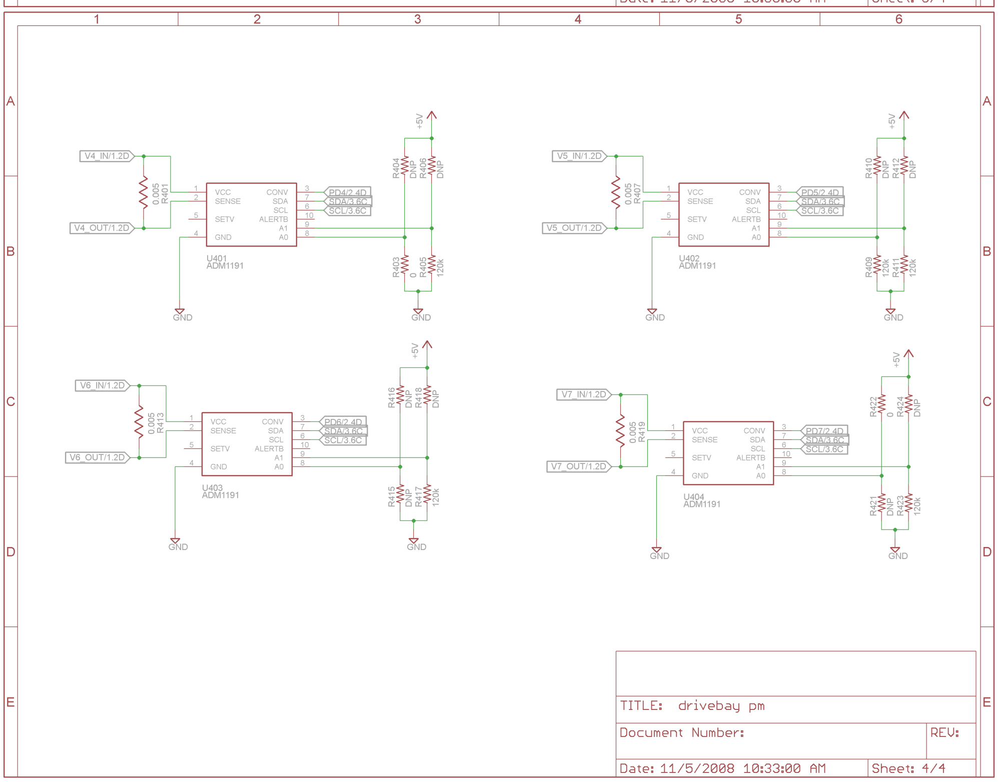
*voltage = reading / 4096 * 26.52 (Volts)*
*current = reading / 4096 * .10584 / .005 (Amperes)*

C Schematics

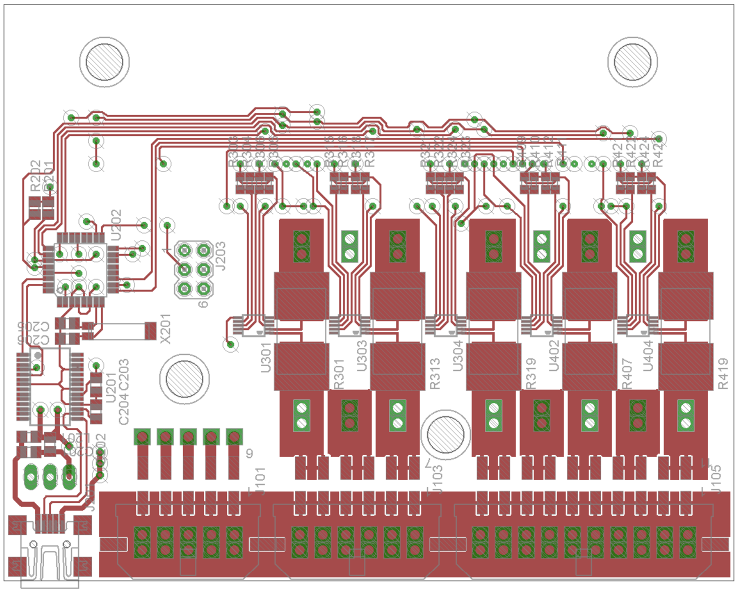




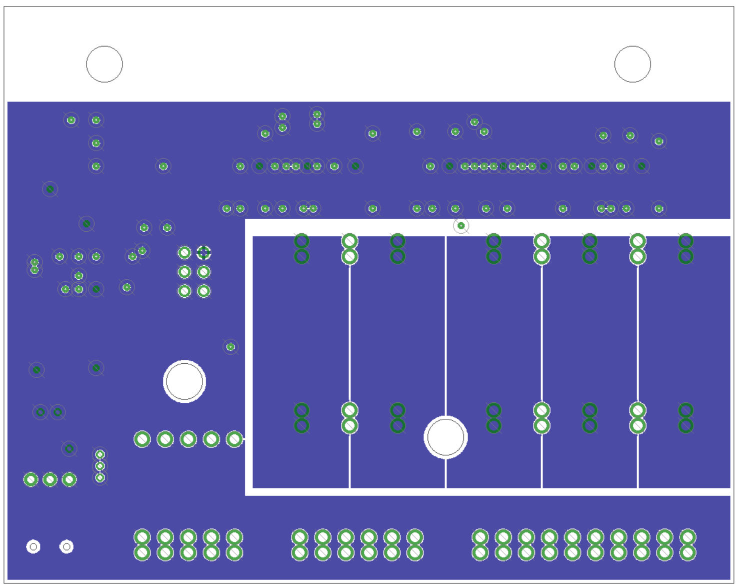

TITLE: drivebay pm
Document Number:
REV:
Date: 11/5/2008 10:33:00 AM
Sheet: 3/4


TITLE: drivebay pm
Document Number:
REV:
Date: 11/5/2008 10:33:00 AM
Sheet: 4/4

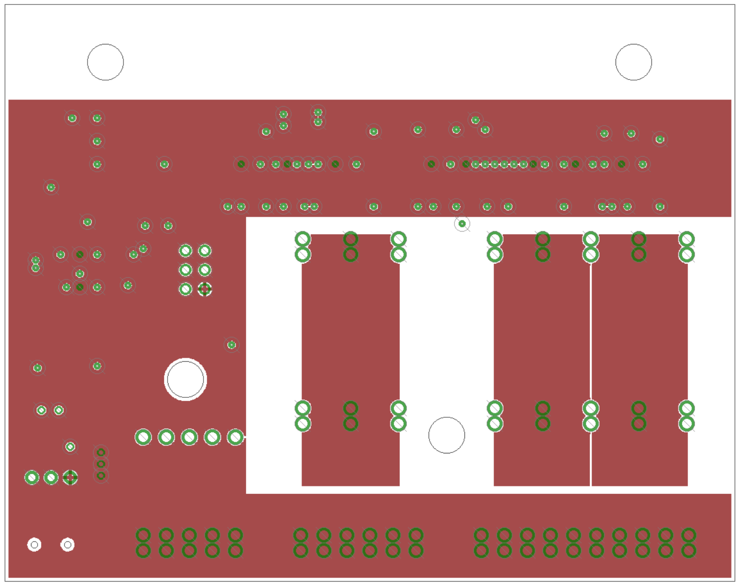
D Layout



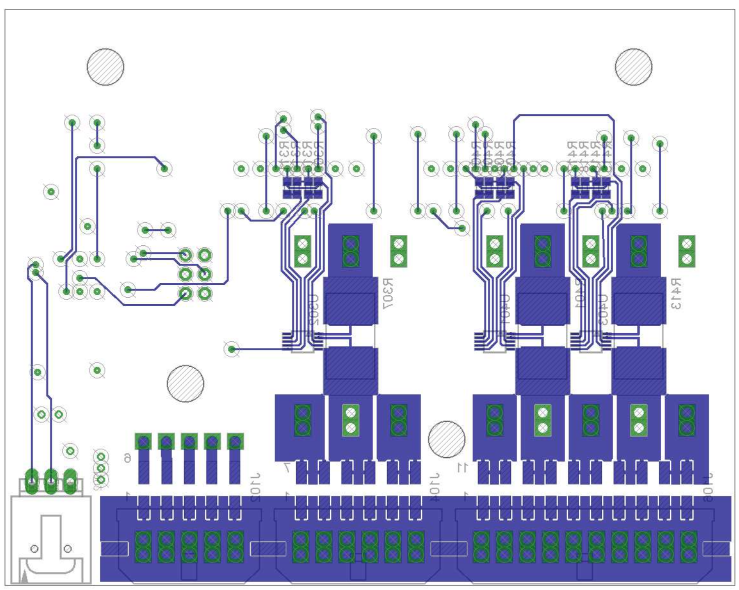
Top layer



Inner layer 1

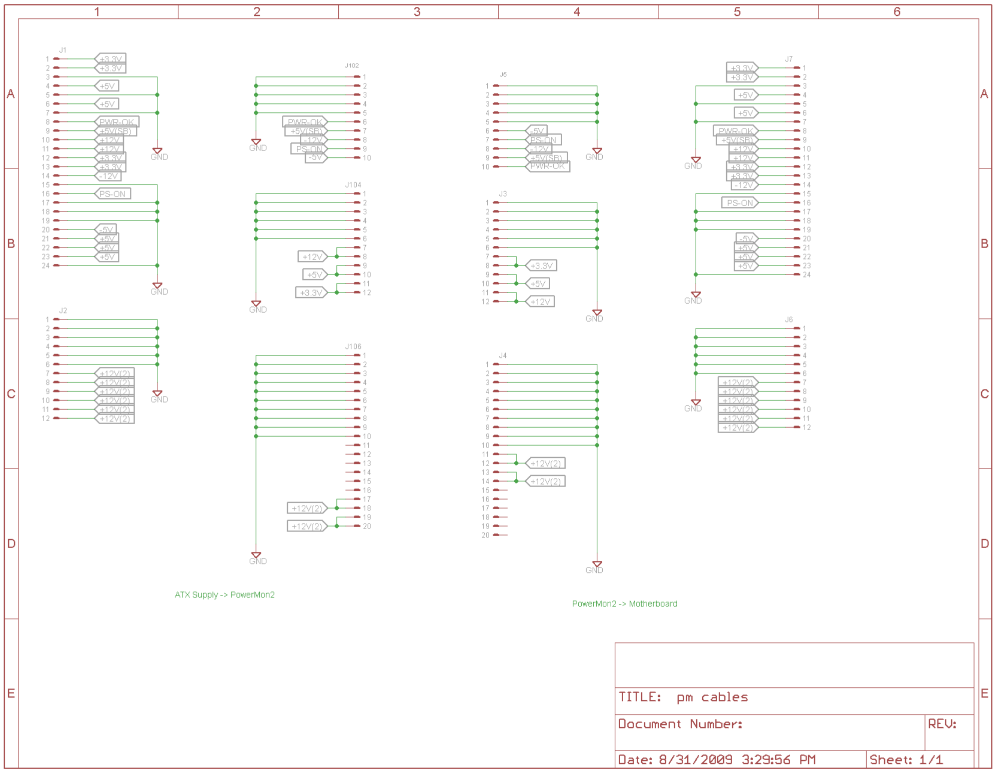


Inner layer 2



Bottom layer

E Cable Diagrams



F Bill of Materials

Identifier	Manufacturer Part Number	Description	Quantity	Unit Cost	Item Cost
C201	generic	Capacitor, 0805 ceramic chip, 10 nF	1	$0.05	$0.05
C202, C204	generic	Capacitor, 0805 ceramic chip, 100 nF	2	$0.05	$0.10
C203	generic	Capacitor, 0805 ceramic chip, 4.7 uF	1	$0.05	$0.05
C205, C206	generic	Capacitor, 0805 ceramic chip, 10 pF	2	$0.05	$0.10
J101, J102	Multicomp 2261(5561)R-10T-SM1	Connector, 10-pin Micro-Fit	2	$1.05	$2.10
J103, J104	Multicomp 2261(5561)R-12T-SM1	Connector, 12-pin Micro-Fit	2	$0.96	$1.92

Identifier	Manufacturer Part Number	Description	Quantity	Unit Cost	Item Cost
J105, J106	Multicomp 2261(5561)R-20T-SM1	Connector, 20-pin Micro-Fit	2	$1.28	$2.56
J201	Tyco 1734035-2	Connector, mini USB type B	1	$1.59	$1.59
J203	FCI 67996-206HLF	Connector, 2.54mm double row header, 6-pin	1	$0.40	$0.40
L201	Vishay ILHB0805ER900V	Ferrite bead	1	$0.20	$0.20
R201, R202	generic	Resistor, 0805, 10 kiloohm	2	$0.05	$0.10
R301, R307, R313, R319, R401, R407, R413, R419	IRC / TT OARS-XP R0050F	Resistor, 5 milliohm	8	$1.57	$12.56
R303, R305, R311, R317, R322, R323, R403, R422	generic	Resistor, 0603, 0 ohm	8	$0.05	$0.40
R309, R405, R409, R411, R417, R423	generic	Resistor, 0603, 120 kiloohm	6	$0.05	$0.30
U201	FTDI FT232RL	USB transceiver	1	$4.05	$4.05
U202	Atmel ATmega168-20AU	microcontroller	1	$3.25	$3.25
U301, U302, U303, U304, U401, U402, U403, U404	Analog Devices ADM1191	power sensor	8	$3.46	$27.68
X201	Seiko SSP-T	Crystal, 32768 Hz	1	$0.90	$0.90
PCB	Sunstone Circuits	PMDB circuit board	1	$64.80	$64.80
Chassis	Hammond 1444-9	Chassis, 8”x4”x1” aluminum	1	$9.15	$9.15
24-pin Cable	generic	ATX extension cable, 12-inch	1	$4.70	$4.70
8-pin Cable	generic	ATX extension cable, 12-inch	1	$4.55	$4.55
10-pin shell	Multicomp 2260(5560)-10	Connector, 10-pin Micro-Fit	2	$0.42	$0.84
12-pin shell	Multicomp 2260(5560)-12	Connector, 12-pin Micro-Fit	2	$0.47	$0.94

Identifier	Manufacturer Part Number	Description	Quantity	Unit Cost	Item Cost
20-pin shell	Multicomp 2260(5560)-20	Connector, 20-pin Micro-Fit	2	$0.63	$1.26
Cable pins	Multicomp 2260(5560)T	Contact, Female Micro-Fit	84	$0.03	$2.52
Total Cost					$147.07