Energy Saving Potential of Idle Pacman Supercomputing Nodes

Brahm Lower, Jeremiah Dabney, Kurt Carlson, Liam Forbes, Oralee Nudson

Arctic Region Supercomputing Center, University of Alaska Fairbanks, Fairbanks, Alaska USA

bplower@alaska.edu

Introduction

To determine the energy saving potential of suspending idle supercomputing nodes without sacrificing efficiency, my research involved the setup of a compute node power usage monitoring system. This system measures how much power each node draws at its different levels of operation using an automated Expect[1] script. The script automates tasks with interactive command line interfaces, to perform the power measurement readings. Steps required for the power usage monitoring system include remotely logging into the Pacman Penguin compute cluster power distribution units (PDUs), feeding commands to the PDUs, and storing the returned data. Using a Python[2] script the data is then parsed into a more coherent format and written to a common file format for analysis. With this system, the Arctic Region Supercomputing Center (ARSC) will be able to determine how much energy is used during different levels of load intensity on the Pacman supercomputer and how much energy can be saved by suspending unnecessary nodes during levels of reduced activity.

Power utilization by supercomputers is of major interest to those who design and purchase them. Since 2008, the leading source of worldwide supercomputer speed rankings has also included power consumption and power efficiency values[3]. Because digital computers utilize electricity to perform computation, larger computers tend to utilize more energy and produce more heat.

Pacman, an acronym for Pacific Area Climate Monitoring and Analysis Network, is a high performance supercomputer designed for large compute and memory intensive jobs. Pacman is composed of the following general computational nodes:

- 256 four-core compute nodes containing two dual core 2.6 GHz AMD Opteron processors each
- 20 twelve-core compute nodes containing two six core 2.6 GHz AMD Opteron processors each
- 88 sixteen-core compute nodes containing two eight core 2.3 GHz AMD Opteron processors each

Method

To gain an accurate representation of potential energy savings, an average energy utilization value per compute node needed to be determined during both node idle and shut down states. In the node shut down state (as opposed to a powered down state), the compute nodes continue to draw energy unlike most home computers. This is because several low level system operations like power supply cooling fans and network interfaces must continue to run.

To measure node idle power usage, we first confirmed the nodes were not occupied with user jobs then executed an Expect script to connect to each PDU using a secure shell remote login. The power utilization data was then measured and recorded.

The same Expect script was then run on the compute nodes in a shut down state and the power measurements were recorded. Overall, the data collection process took 40 minutes total with data values being collected every 20 seconds.

The measurement procedures described above were conducted on the Pacman four-core, twelve-core, and sixteen-core node types because each consumes different amounts of energy. Power consumption averages for each node type were then determined and compared to find the change in power utilization.

While reviewing the data collected on PDUs providing power to the sixteen-core nodes, it was observed that the recording instruments lacked a resolution capable of providing accurate results. The amps for each node was recorded as zero. However, we know the sixteen-core nodes continue to draw power even though they are shut down. Because of this, further calculations and results exclude data recorded from the sixteen-core nodes.

Results

Figure 1 shows the levels of energy consumed for a Pacman four-core node running idle vs. being turned off. For comparison, Figure 2 represents the same values for a Pacman twelve-core node in corresponding states.

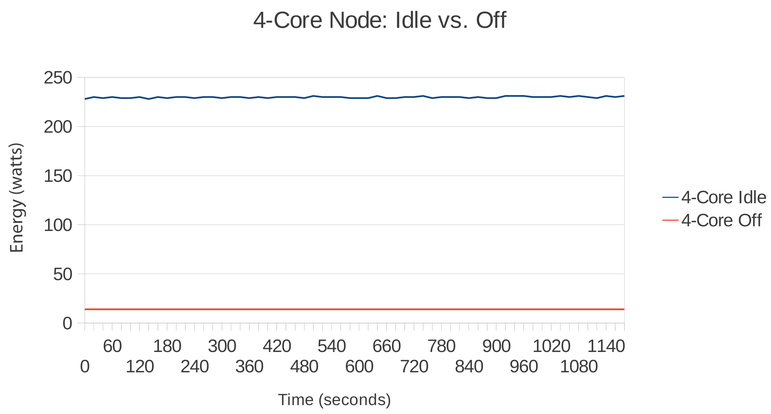

Figure 1: Power Utilization for a Pacman four-core node

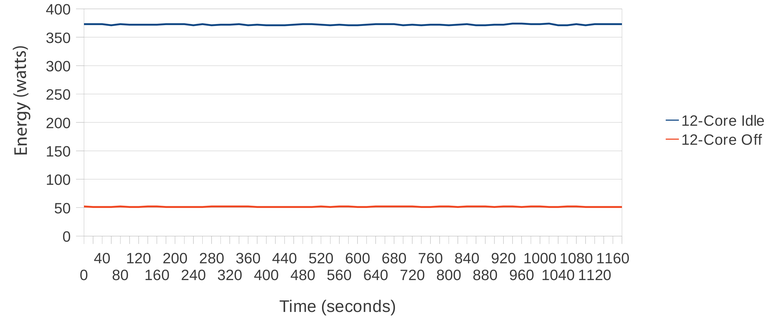

Figure 2: Power Utilization for a Pacman twelve-core node

Table 1 contains calculated power utilization averages for each of the Pacman compute node types and their change in energy utilization between the two measured states of operation: idle and shut down.

Node Type	Node State: Idle	Node State: Shut Down	Change in Measured Power Utilization
4-Core Nodes	229.56 Watts	14 Watts	215.56 Watts
12-Core Nodes	149.83 Watts	27.4 Watts	122.43 Watts
16-Core Nodes	258.56 Watts	Not Available*	Not Available*

Table 1: Power Utilization for Individual Pacman Node Types

* There is no wattage value available for 10% of the sixteen-core nodes being shut down because of measurement device limitations.

Discussion

If it is assumed that in one year 10% of the Pacman compute nodes are shut down half of the time, the total kilowatts saved during one year can be determined by comparing the power utilization of 100% of the nodes at idle to the energy used when 90% of the nodes are idle and 10% are shut down. Tables 2 and 3 along with Figure 4 show the application of this calculation.

Node Type	Total Number of Nodes Available in Pacman	10% of Total Nodes (rounded integer)
4-Core Nodes	256	26
12-Core Nodes	20	2

Table 2: Number of Nodes used in Power Utilization Calculations

Node Type	Watts consumed by 100% idle nodes	Watts consumed by 90% idle nodes	Watts consumed by 10% shut down nodes
4-Core Nodes	58,767.36	52,798.80	364.00
12-Core Nodes	5,993.30	5,393.97	54.80

Table 3: Measured Power Utilization for Pacman Compute Nodes

Total Watts Consumed by 100% Idle Nodes		Total Watts Consumed by Idle and Shut Down Nodes		Total Hours in six months		Kilowatt Hours Saved
(64,760.66	-	58,611.57)	*	4,380	=	26,933.0142

Figure 4: Total Energy Savings over a Six Month Period

For monetary calculations, the cost estimate price of a kWh is $.165517 plus $.0993 per kWh for facility usage. This means in one year, the estimated saved energy cost would be the total kilowatt hours saved times $.264817. In this case, it is calculated that 26933 kWh could be saved in the case of 10% of the Pacman compute nodes being shut down instead of remaining in an idle state over one year. This would result in a minimum savings of approximately $7132.32. It should be noted that this considers only the four-core and twelve-core Pacman nodes.

Conclusion

If the primary goal is to save energy, there are many other possibilities to be explored including the implementation of suspended states. Node suspended states continue to power the RAM but turn off the processor with the exception of a few small systems. This capability is not however, available on all machines. To apply this to Pacman, further research needs to be conducted with the particular hardware and software versions currently installed on the system. Additional areas of important consideration, further research, and exploration include the following:

- The introduction of possible increases in hardware failure rates per node due to more frequent power cycling of the idle compute nodes
- Conducting a comparison of potential savings to the time lost waiting for nodes to boot when needed
- Analyzing system utilization to gain a better estimate to replace the 10% used for this poster
- Measuring the sixteen-core nodes with more accurate equipment
- Explore the cost of implementing and managing "energy aware" scheduling software
- Consider the scalability of these power and cost savings on systems much larger than Pacman.

References

1. Expect. [Online]. Available: http://expect.sourceforge.net
2. Python. [Online]. Available: http://www.python.org
3. "Power Consumption of Supercomputers", Top 500 Supercomputer Sites, [online] 2008, http://www.top500.org/lists/2008/06/highlights/power (Accessed: 20 April 2012).
4. G. Dasgupta, A. Sharma, A. Verma, A. Neogi, R. Kothari. "Workload Management for Power Efficiency in Virtualized Data Centers", Communications of the ACM, [online] 2011, http://cacm.acm.org/magazines/2011/7/109901-workload-management-for-power-efficiency-in-virtualized-data-centers/fulltext (Accessed: 20 April 2012).