

More Power to Google

April 7 2007

Google is seeking the optimal energy efficiency for its large data centers, and it is counting on its top engineers to help deliver it.

Luiz Barroso, a distinguished engineer at Google, discussed the company's projects to reach optimal energy efficiency in a talk entitled, "Watts, faults and other fascinating 'dirty' words computer architects can no longer afford to ignore," at the company's complex here on April 5.

Barroso, a former Digital Equipment engineer with a history of delivering load balancing software for large-scale systems and for working on the design of the core Google infrastructure, summarized two projects he has been working on.

One, a power provisioning study, will be formally released in a paper this summer, Barroso said.

Two main points arose from the power provisioning study, he said: "Maximizing usage of available power capacity is key," and "systems are typically very power-inefficient on nonpeak conditions."

Moreover, Barroso said, "Power/energy efficiency and fault-tolerance are central to the design of large-scale computing systems today. And technology trends are likely to make them even more relevant in the future, increasingly affecting smaller-scale systems."

Barroso acknowledged that Google is building data centers where there is hydroelectric power and "engineers are squeezing every little watt out

of every card."

Indeed while circuit designers have to worry about things like temperature and other issues, "we worry about the affordability of building data centers," Barroso said.

He noted that it costs between \$10 and \$22 per watt to build a data center, while the U.S. average energy cost is only 80 cents per watt. So "it costs more to build a data center than to power it for 10 years," Barroso said.

"You want to get as close as possible to optimal usage," because unused watts cost money, he said.

So for the power provisioning study, Google looked at how much energy its machines were using over six months.

The example for the study covered only 800 machines of the thousands Google employs, and one of the findings was that "you spend 60 percent of your time at or below your peak, and racks of machines are never at peak at the same time."

Moreover, "the data center as a whole is never going above 70 percent of capacity, and that shows we could have deployed 40 percent more machines."

Barroso highlighted two hot areas of computer design made famous in the '90s that have proven to be flawed. One is the acceleration of single-thread performance, which he referred to as the megahertz race. The other is the building of big, distributed shared memory systems, which he called the DSM race.

The theory behind the DSM race was that large-scale computing systems

should use a shared-memory programming model because it was familiar to programmers and facilitates sharing of expensive resources, among other things. But the undoing of the DSM race was fault containment, Barroso said.

"A single fault can bring down the entire shared memory domain," Barroso said. "It's a very hard problem to solve ... and most of the solutions are inadequate."

Meanwhile, in the megahertz race, where even unmodified software simply gets faster by itself because of some computer architectural tricks; "the megahertz race crashes into the power wall," Barroso said.

He said that every year enterprises can buy faster servers for about the same price, "but much more energy is being used so systems become power-inefficient."

Joked Barroso: "When you get to the point where power costs more than servers, you'll have a situation like the cell phone industry model where utility companies might say, 'I'll give you these servers for free if you sign this energy contract.'"

Barroso also mentioned H.R. 5646, a congressional bill signed into law last year to promote the use of energy-efficient computer servers in the United States.

"There are a lot of things you can do to reduce energy conversion losses, like go to single-voltage rail power supply units [PSUs]," Barroso said. "You can get up to a four times reduction in conversion losses."

Moreover, Barroso said Google is "working with - its - partners to create open standards for higher-efficiency PSUs." He later said the list of partners includes Intel and AMD.

Meanwhile, new technologies such as multicore processors and increasing parallelism offer promise. "But there's a catch," Barroso said. "Are there enough threads? Can we expect programmers to build efficient/concurrent programs?"

Indeed, with more data it is easier to do parallelism. "At Google we're interested in problems where there's a truckload of data, so it might be a little easier for us," Barroso said.

However, fault-tolerant software is powerful, but it is not enough, Barroso said. Large-scale systems also need additional monitoring.

Google employs what it calls its System Health Infrastructure, which talks to every server in the system frequently and collects health signals and activity information, Barroso said

Asked if Google might consider open-sourcing this technology, Barroso said "We've been looking at open-sourcing some of the code for some time." However, "some of this is infrastructure and we build it so intertwined with other software we have that it's hard to pull things apart."

In addition, Google uses self-monitoring, analysis and reporting technology, or SMART, to do early detection of problems. And it found that disk drives with scan errors are 10 times more likely to fail than those with no errors, Barroso said.

However, the company found that more than half of the drives that failed showed no signals, he said. Indeed, 56 percent had no strong signals at all, he said.

"It's fairly easy to predict something if you give a long enough time frame," Barroso said. "I predict we're all going to die," he quipped.

In addition, Barroso said the Google study found that temperature was not shown to be a significant factor in disk failures - slightly warmer temperatures did not cause any more failures than cooler ones.

"If the variability of temperature is not that great then data center designers have a lot more flexibility" in designing more energy-efficient facilities, Barroso said.

Copyright 2007 by Ziff Davis Media, Distributed by United Press International

Citation: More Power to Google (2007, April 7) retrieved 26 April 2024 from <https://phys.org/news/2007-04-power-google.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.