

Friday, May 15. 2009

Some observations on the SPECpower_ssj2008

There is an good example for the interesting nature of the SPECpower_ssj2008 benchmarks. Sun benchmarked three systems. A Sun Netra X4250 with 8 GB and with 32 GB.

Unlike other loads there is only a small correlation of memory size and the number of ssj_ops. At 8 GB an Sun Netra X4250 yields 244,832 ssj_ops. The 32 GB version yields 251,555 ssj_ops

There is an parameter that's strongly correlated to the size of the memory. It's the power consumption. The system with 8 GB consumes 226W. The one with 32GB needs 294W.

Obviously this has large impact to the manager/journalist friendly single number as the result of the benchmark. The result is computed as the sum of ssj_ops at 10,20,30 and so on up to 100% utilisation divided by the used watts at each of this utilisations. When you have a way to reduce the power consumption by a large amount with only a small impact to performance, it's a sure way to improve your SPECpower_ssj2008 result. As shown above the reduction of memory size is such a method. As you don't have an large impact on performance by removing memory its quite easy to increase you benchmarking number

But there is more in you number: Look at the performance numbers: At 100% utilisation the system needs 226 watts. A system running at 0 percent utilisation needs 174 watts. This numbers are similar on every similar configured system A system doing nothing (okay, it runs the OS, another reason why we need tickless operating systems) needs 76% percent of the power consumption. So the obvious conclusion is: Run the system at 100% utilisation ... or at least try to get near to this number. But that isn't so easy. You can run multiple application on a system to do so. But running multiple application need more memory. More memory than just the 8 GB used in this benchmark. So the benchmarks shows at best, that the small configurations are nonsense, at least when you have not an application that can load 2 procs and 8 cores with just 8 GB of memory.

Another take away when you strive through this benchmarks that this benchmark is the return of single power supply configurations (the Sun configuration is one exception, all results are with doubled supplies. This is an important factor. At power supply is most efficient at high load. So benchmarkwise it's best to load a single power supply as highest as possible. At a redundant load-sharing configuration a single power supply has a smaller load and runs at a lower efficiency as the load is equally divided between the both. But for almost all jobs you have to use such a configuration as you want redundant power supplies for availability.

I really think that the SPECpower_ssj2008 needs some modification like mandating relevant memory sizes (and a load that really punishes small memory configurations) or forcing the vendors to use redundant configurations. Obviously the best way would be modification of all benchmarks (not only the ones from SPEC) to include the power consumption of the configurations. That would lead to reasonable and comparable configurations for power benchmarking.

DISCLOSURE STATEMENT: Sun Netra X4250 server 600 overall ssj_ops/watt and (226 watts, 244832 ssj_ops) @ 100% target load, (210 watts, 121828 ssj_ops) @ 50% target load, (181 watts, 24150 ssj_ops) @ 10% target load, (174 watts) @ active idle target load. Sun Netra X4250 server 478 overall ssj_ops/watt and (294 watts, 251555 ssj_ops) @ 100% target load, (226 watts) @ active idle target load. Sun Netra X4250 server 437 overall ssj_ops/watt and (296 watts, 244832 ssj_ops) @ 100% target load, (225 watts) @ active idle target load. SPEC and the benchmark names SPECpower_ssj, SPECweb, SPECjbb, SPECjAppServer, SPECComp are trademarks of the Standard Performance Evaluation Corporation. Benchmark results stated above reflect results published on <http://www.spec.org> as of March 30, 2009. For the latest SPECpower_ssj2008 benchmark results, visit http://www.spec.org/power_ssj2008

Posted by Joerg Moellenkamp in English, The IT Business at 20:17

Dem Problem mit der verschwendeten Energie durch redundante schwach-belastete Netzteile koennte man ja beikommen, wenn man mal ein bisschen mehr Augenmerk auf den Wirkungsgrad ueber eine groessere Last-Bandbreite legt in der Entwicklung. Im Heim-PC-Bereich haelt das ja langsam Einzug, dann kanns ja nimmer lang dauern bis das auch bei Servern soweit ist. Beim SPECpower-Test mit nur einem Netzteil zu fahren, ist jedenfalls IMHO Betrug, weil nicht praxisrelevant. Wenn eine Stunde Betriebsausfall mehr kostet wie mehrere Jahre ein dahinidlendendes Netzteil durchzufuettern, weiss ich fuer welchen Weg ich mich entscheide.

(Slightly unrelated: Sun sollte seine Green Computing-Initiativen mal auf tote Baeume ausdehnen. Was die uns bei jeder Bestellung an Papiermuell beschere ist legendaer.)
Anonymous on May 18 2009, 13:00