GRIDBYTE ™
MULTI-DIMENSIONAL THINKING

29911 Niguel Road—Box 6298
Laguna Niguel, CA 92607-6298
949.916.0799
800.634.0796
email: info@gridbyte.com

Title: Compute Infrastructure Performance Audit & Diagnostics
White Paper: GridByte-CS2007-08
Created by: Sam O. George/GRIDBYTE™
Release Date: 2007-MAY-30
URL: http://www.gridbyte.com/Resources/Documents/GridByte-CS2007-08.pdf
Keywords: Compute infrastructure performance audit, System Analysis

Case Outline: Client was a fabless semiconductor company with 70 employees. Site contained 55 Linux compute-servers, 2 NAS file-servers, 10 ancillary Linux/Solaris workstations and 130 other devices (Windows workstations, lab equipment, printers, roaming laptops, etc). Their engineering staff were experiencing a number of critical performance problems that included: (1) high "network latency" during moderate to high stress simulation and graphics rendering; (2) unexplained data-losses after simulations; (3) low to high user-satisfaction with the performance of various access methods.

GridByte™ was retained to complete an infrastructure performance audit[1] and offer best practice suggestions consistent with application of HPC in IC fabless companies. This work included assessment of enterprise scalability and providing a framework to bridge the expectations of engineering staff with IT performance.

Solution Brief: GridByte™ completed the performance audit in a record 6 weeks—a period that included major operating system upgrades on all compute servers. Our findings included a 10-point actionable list of major issues/problems. Our calculations showed that the client stood to gain 40% to 80% improvement in aggregate performance if they optimized as recommended.[2] We provided full numerical performance analysis and reported network topology implementation. Our work led to the discovery of dynamic performance bottlenecks. We completed network diagramming of all 274 IP devices on the client's network; analyzed and documented performance of client's hybrid use-model vis-à-vis opportunities for performance improvement; and demonstrated saturation points of current network implementation vs. growth projection and deployment/use of grid virtualization. Our work exposed critical

---

[1] Compute Infrastructure Performance Audit is exhaustive. It examines infrastructure performance and creates templates of current performance relative to documented expected performance. The audit examines infrastructure at all levels of computing: hardware (computers, switches, and routers), CAD and productivity software deployment and use, virtualization tools and network fabric implementation. The work also examines the performance of networks at the packet level vis-à-vis opportunities for improving application and virtualization layer performance. The effort provides numerical performance numbers for various metrics. The various metrics are examined in the context of global and localized use-models and rated for efficiency against a baseline of the infrastructure. What a performance audit is not: It is not a device counting exercise. The audit is generally not mandated to correct problems—it simply must find them. The performance audit may add hardware components to the infrastructure, but it must remain completely independent so that measurement does not affect the "network under observation."
Many clients request infrastructure benchmarks as output from the performance audit. However, the focus of a performance audit is to measure how well the infrastructure performs in multiple areas against its design metrics. The answer to the question, "How well is my infrastructure performing?" may be simple. The supporting material reflects a matrix of interdependencies. The answer depends on successful application of use-models that support the mission of the organization.

[2] For reference: In June 2006, the Sandia National Laboratories 8,960-processor "Thunderbird" Linux cluster logged 38.3 teraflops. In November 2006, it logged 53 teraflops. They accomplished this 38% performance improvement by optimizing the Thunderbird Linux software. If the lab had chosen to improve performance via hardware, they would have required 3,443 processors and an extra $5 million.—paraphrased from "Tweaking and Twiddling" by Douglas Eadline in Linux Magazine, May 2007, www.linuxmagazine.com.

29911 Niguel Road—Box 6298
Laguna Niguel, CA 92607-6298
949.916.0799
800.634.0796
email: info@gridbyte.com

bottlenecks due to NAS file-server ethernet I/O *misconfiguration* among many other discoveries.

Solution Detail: The client's compute infrastructure in Figure 1 consists of 2 main subnets—let us call these subnet 40 and subnet 30. As shown in Figure 1, subnet 40 is a "general access" network for all users in the company. Subnet 30 is a restricted high performance compute cluster we shall call the grid cluster. Subnet 30 also hosts the main NAS file-server called *fs-h*. This file-server provides NFS and CIF filesystems for the entire infrastructure. Not shown in Figure 1, the client also maintained another file-server, *fs-d*, that mirrored all the *fs-h* data via the subnet 30 switch.

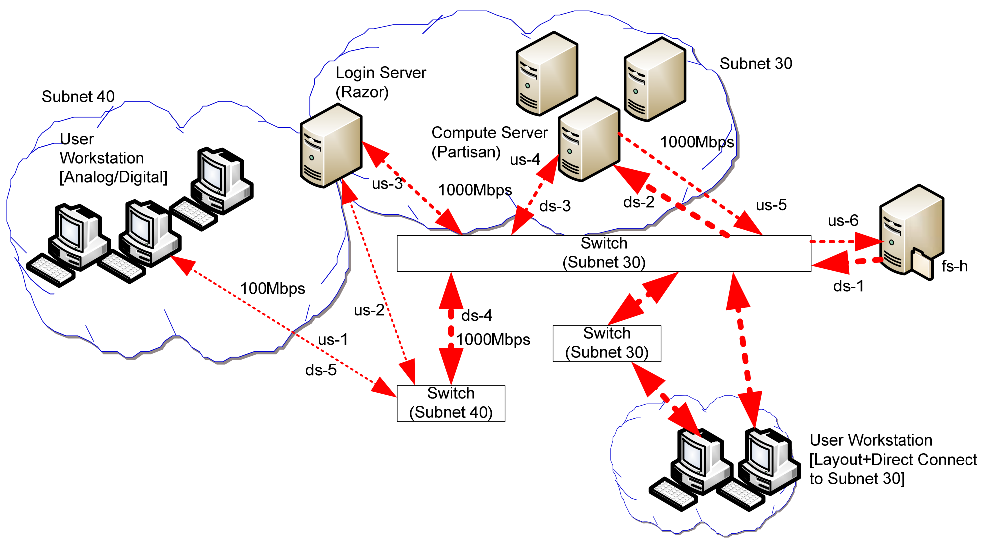

Figure 1: Physical Compute Infrastructure[3]

Network Diagramming and Topology Review

The first step of a performance audit is <u>device discovery</u> and <u>network topology assessment</u>. This usually means validation of resources and network topologies provided by the client. In this case, the client didn't have topology documentation, so we had to generate and validate the infrastructure from scratch. During audits, our GridByte™ techniques apply *agentless* discovery methods--we introduce minimal software and hardware into the client's infrastructure to minimize their downtime. In all networks, the discovery phase queries

[3] ***ds*** = downstream & ***us*** = upstream. The sizes of the link-lines indicate relative size of network traffic throughput.

GRIDBYTE™
MULTI-DIMENSIONAL THINKING

29911 Niguel Road—Box 6298
Laguna Niguel, CA 92607-6298
949.916.0799
800.634.0796
email: info@gridbyte.com

Title: Compute Infrastructure Performance Audit & Diagnostics
White Paper: GridByte-CS2007-08
Created by: Sam O. George/GRIDBYTE™
Release Date: 2007-MAY-30
URL: http://www.gridbyte.com/Resources/Documents/GridByte-CS2007-08.pdf
Keywords: Compute infrastructure performance audit, System Analysis

all compute resources (network interface speed, memory, OS configuration, MAC address, etc). This assumes that the devices are not transparent to discovery tools.[4] This particular site required application of 3 different software tools and 4 discovery methods to yield a 99% confidence level over 6 weeks. GridByte™ applied a set of scripts to exhaustively probe the infrastructure without impacting performance. Our methods successfully identified and discriminated all compute resources, switches (managed and unmanaged) and routers. We created a network diagram that identified all subnet clusters.
This site had some unique challenges. For example, only 27 out of 274 devices were SNMP enabled. Of the 27, only 11 returned useful data. Various devices responded poorly to our port-scanning tools—these devices locked-up and caused the tools to fail. Our GridByte™ methodologies applied packet injection techniques to capture critical information about each device node. These techniques rely on mathematical heuristics to determine that a device is truly a device.[5]
Our analysis of the network topology confirmed the general design of the topology in Figure 1. We found that critical components (notably *fs-h*) spanned both subnets. These spanned resources, coupled with sub-optimal network expansion created unique structural bottlenecks in the client's infrastructure. We found that because all of the switches supported layers 2 and 3 (ref. OSI model), their infrastructure faced scalability challenges. They needed upgrades to allow content switching and other traffic management efficiencies.

10-Point Actionable Level 1 Audit Detail

1. Network implementation Issues: As shown in Figure 1, the network implementation created performance bottlenecks because critical devices/services spanned both subnets or bypassed the design philosophy of the network. The problem was more complex than captured in Figure 1. In this type of engineering network, users push/pull large amounts of simulation and graphics data from workstation to compute-servers, then to file-servers and back to workstations. The network use-model was not optimal to support high sustained throughput. How does GridByte™ define sustained throughput? Let's say 30% of the users demanded multi-gigabit file-access. Can the network sustain the load without significant latency and/or loss of service? In this case, the client's infrastructure reached saturation with less than 5% bandwidth demand access. The mirroring operation across the subnet 30 production switch also created a performance bottleneck. A consequence of resource-spanning is that users logged onto the grid cluster and were generating a large amount of http traffic. Http traffic was not supposed to be part of subnet 30.

[4] Many switches and routers are left in unmanaged default configurations. This means the system administrators do not configure their IP address or enable SNMP features. This is not a problem in small networks. However, in large networks, unmanaged transparent switches hide major inefficiencies in network topology design.
[5] In highly virtualized compute infrastructures, we need to discriminate virtual IP addresses from physical ones.

29911 Niguel Road—Box 6298
Laguna Niguel, CA 92607-6298
949.916.0799
800.634.0796
email: info@gridbyte.com

2. Use Model Performance Problems—Part 1: The engineers used a number of virtualization tools and at least 5 methods to access the physical resources. Many users required *stateful* session portability. However, the various tools used were not optimized for performance and stability. Thus many user-problems with performance were related to sub-optimal use of virtualization tools. To illustrate this, consider a user accessing an X-Server application from within an RDP session in order to run a graphics/data intensive simulation. This complex scenario is just one example of the nature of the client's problems. Taken further, the absence of a cohesive use-model created further problems for the client's infrastructure. For example, they did not manage where users started X-servers. Some users started X-servers on servers not intended for such activity. Again, they saw performance problems related to the lack of an optimized use-model.

3. Use model Performance Problems—Part 2: Beyond access issues, many users routinely bypassed the grid-cluster management software that was an integral part of their use-model. This means that unsupervised workloads and interactive sessions routinely competed for the same compute-resources. Taken further, interactive sessions routinely confused the grid-engine. Consider this: an interactive session is idle while the designer attends to another task. Because the session is idle and not consuming resources, the grid-engine assigns another simulation job to this compute-server. However, when the interactive designer resumes their simulation, they now must compete with another user's simulation. This happens because the grid-engine, following its rules for submission, is not aware (and cannot be) of the interactive-user's intentions.

4. Network Link Throughput Problems: We found that the best link performance in the client's gigabit grid-cluster was 730Mb/s. The vast majority of gigabit links showed impaired throughput performance—sometimes as low as 100Mb/s. This problem was correlated to transmission payload correlated to CPU affinity features of multi-processor Linux compute servers. These problems showed up intermittently and required exhaustive performance monitoring to diagnose. Moreover, reflecting the wide range of device manufacturers, some NICs did not show this problem.[6] In short, our findings demonstrated that the client gigabit links had throughput numbers of 26% to 55% below theoretical maxima. In other words, in a real world situation, we expect a gigabit link to perform near 900Mb/s. The performance losses represented uncaptured return on investment. However, this core finding exposed shortcomings in the architecture of the client's compute use-model and switch layout topology.

5. File-server Structural Bottlenecks: Our analysis of the fileserver, *fs-h*, showed large packet losses and underutilization of 3 of 4 ethernet ports. The client was using an

[6] Some NICs use TCP Offload Engines (TCPoE or TOE). These NICs showed higher performance.

29911 Niguel Road—Box 6298
Laguna Niguel, CA 92607-6298
949.916.0799
800.634.0796
email: info@gridbyte.com

inefficient NIC load-balancing method that effectively reduced file-server throughput to a theoretical 900Mb/s (1 port) to service the entire compute infrastructure. We found that there were excessive packet-losses at one port and virtually no packet-losses at other ethernet ports. This problem was compounded because all project data was written from the compute-servers to *fs-h* and then sent back through the network to the user-workstations. Taken further, the client use-model made this problem worse. With a different system for simulation data storage, the client could reduce the workload on *fs-h* ethernet interfaces significantly. Our GridByte™ practice creates highly optimized solutions because we assess how each class of user generates/uses data. In these engineering environments, users typically generate tens of gigabytes of temporary simulation data every day. The compute-infrastructure optimization challenge requires design to maximize throughput at all levels. Unfortunately, the solution is never linear. It is a true performance loss-minimization problem.

6. Hardware Configuration Problems: The critical compute resources were not optimized for HPC engineering computing. For example, the compute-servers had default Linux installations with default disk partitioning schemes. In general, GridByte™ experience shows 50% to 100% improvement in core performance of compute-servers when these devices are fully optimized. The optimization levels include kernel tuning, NIC driver optimization and interrupt balancing.

7. Network Layer Problems: In the first week of our engagement, our inspections revealed that the client was using Linux kernel below 2.4.21—with no support for NFS v3. This means that the client could not capture significant network performance ROI; i.e., NFS v3 offers max read/write block size of 32K vs. 8K for NFS v2. This issue is directly related to application-level performance and to network throughput. Using GridByte™ multi-dimensional regression testing, we found that I/O suffered greatly under different payload conditions and different block sizes. Tuning NFS is a complex process that requires observation of software and data combined with use-model. Here, again, the client failed to capture significant ROI because the network layer was not optimized.

8. Switch Fabric Issues: The client utilized fifteen 48-port layer 2/3 switches to service the infrastructure. The switches were not optimized along several dimensions. The IOS on core switches were at least 5 years old. This means that many routing performance improvements were available. Further, none of the switches were configured to use OSPF routing protocols—meaning that the client did not capture significant performance improvements offered by advanced routing protocols. Another limitation of their switch fabric that impacted growth-scabality is that it did not include any content services switches to address OSI Layers 4 through 7.

✉ 29911 Niguel Road—Box 6298
Laguna Niguel, CA 92607-6298
☎ 949.916.0799
☎ 800.634.0796
email: info@gridbyte.com

Title: Compute Infrastructure Performance Audit & Diagnostics
White Paper: GridByte-CS2007-08
Created by: Sam O. George/GRIDBYTE™
Release Date: 2007-MAY-30
URL: http://www.gridbyte.com/Resources/Documents/GridByte-CS2007-08.pdf
Keywords: Compute infrastructure performance audit, System Analysis

9. Monitoring Shortcoming: The client had network monitoring tools installed. However, as GridByte™ showed, these tools were not suitable for performance monitoring. Our monitoring of compute-servers over weeks created a rich dataset of performance insight for all major components; i.e., CPU, disks, buffers, RAM, ethernet interfaces, switches, etc. The data provided critical time-based insight into why the infrastructure had a host of performance problems that were difficult to diagnose and reproduce. These problems could not be captured by simple measurements based on averages. For example, the client tools reported average CPU utilization within 24-hour windows. These measurements showed CPU utilization under 20%--as is consistent with most compute resources worldwide. However, as GridByte™ showed, the average CPU utilization over a day is meaningless in performance environments. This is because performance is bursty. During the workday, utilization in short time-epochs routinely red-lines, i.e., is over 80% for several minutes. When resource red-lining occurs, users see that productivity plummets. The reasons are multiplicative. If a system is unresponsive for a few minutes, the human impulse seems to be to click more. Thus, the effect is that more redundant requests are generated. Users often become distracted—and so, the organization loses many minutes to hours of productivity.

10. Monitoring is a larger issue than fault detection: The client did not have baseline performance metrics for the compute infrastructure. GridByte™ routinely creates baseline metrics to ensure that performance deviations are captured proactively. This methodology is necessary in most organizations that want to provide a minimum level of service (or service guarantee) to each user. The method of looking externally for benchmarks does not preclude the need for baseline metrics.[7]

Conclusion: This case study highlights how GridByte™ applies multi-dimensional thinking in Level 1 performance audits. This client site showed a large number of areas of uncaptured ROI. Given their limited IT/IS staff, they chose to deal with the problems by prioritizing. However, as shown by the complexity of the case, the state of their enterprise required sustained work for a few months to create a good level of baseline performance measured across multiple dimensions of computing. Many of these problems are interrelated. Thus, various tasks must proceed in parallel.
As shown above, this work did not perform significant audit/analysis of virtualization and application tools. Our packet analysis was limited to simple protocol and packet-size discovery. Judging from level 1 findings, the client, again, had significant opportunity

[7] Compute networks are dynamic. The network compute devices, access methodologies, use-models and network architectures change every day. The proactive management style utilizes ***baselining*** to determine performance changes before users report them. For example, network performance could suffer because of a recently-added switch has not been optimized for the organizations metrics. Said differently, it is impossible to quantify, and for that matter, improve performance in tangible incremental ways. The net result is that the impact of improvements must be measured by subjective measures.

29911 Niguel Road—Box 6298
Laguna Niguel, CA 92607-6298
949.916.0799
800.634.0796
email: info@gridbyte.com

Title: Compute Infrastructure Performance Audit & Diagnostics
White Paper: GridByte-CS2007-08
Created by: Sam O. George/GRIDBYTE™
Release Date: 2007-MAY-30
URL: http://www.gridbyte.com/Resources/Documents/GridByte-CS2007-08.pdf
Keywords: Compute infrastructure performance audit, System Analysis

to capture ROI from these additional areas. However, these could not begin until the foundational work was complete.

This Level 1 audit also showed the client that their infrastructure needed significant reconfiguration to handle doubling its workforce, to say, 140 employees. Given their current ratio of human-to-machine (about 1-to-4), we would expect this client infrastructure to support just under 600 devices (140 employees)—thus, planners needed to think proactively about device growth in significant multiples of employee count. This 'little' detail is a significant source of future problems in many companies.

Citation: Sam O. George, "Compute Infrastructure Performance Audit & Diagnostics," http://www.gridbyte.com/Resources/Documents/GridByte-CS2007-08.pdf, GridByte™, 30 MAY 2007.