I2 END-TO-END PERFORMANCE INITIATIVE PLANNING MEEETING

Joseph E St Sauver (joe@oregon.uoregon.edu)
University of Oregon Computing Center
Eugene, Oregon 97403 USA

0. Introduction

In this paper, we were asked to comment on three topics: (1) key aspects of end-to-end (E2E) performance that should be addressed by the proposed I2 E2E performance initiative, (2) success metrics for the initiative, and (3) criteria to be used in the E2E initiative call for participation. We cover those topics in the order they're mentioned.

1. Key Aspects of End-to-End Performance

(a) Host System Hardware and Operating System Configuration: One obvious and key aspect of end-to-end performance (which has generally received short shrift to-date) is host system performance associated with system hardware and operating system configuration.

Granted, it is true that some attention has been paid toward tuning host system TCP/IP stacks either on a manual basis (http://www.psc.edu/networking/perf_tune.html) or on an automated basis (http://www.web100.org/), however to-date the application of that knowledge to specific systems has been on a pull (system administrator-instigated) basis rather than on a push (I2-suggested) basis. There has also been no systematic effort to have those changes made where making those changes will have the greatest net effect.

***Recommendation 1:* Internet2 should work with campuses to identify the dozen or so "largest" or "most important" hosts (or cluster of hosts) at each I2 connected site, e.g., the dozen or so hosts (or cluster of hosts) which service the most users on each campus. Having identified those hosts (or clusters of hosts), I2 E2E staff should offer to work with interested system administrators to identify opportunities for tuning those systems so as to improve their end-to-end performance.**

Doing this would insure that I2's limited resources will impact the largest possible number of users as quickly as possible, thereby delivering the most "bang for the buck," while also serving to institutionalize the system tuning process by vesting tuning knowledge with professional system administrators who implicitly or explicitly set the standard for other system administrators on their respective campuses.

If Internet2 wants to encourage deployment of systems which will go fast end-to-end, they'll also need to help Internet2 users understand the limitations associated with host system hardware, such as motherboard related issues (including the importance of having 64 bit, 66MHz PCI slots available for gigabit NICs and high bandwidth SCSI controllers), the importance of providing the host with sufficient CPU to handle interrupt loads and substantive application processing loads, and the importance of deploying high performance I/O subsystems.

Recommendation 2: Internet2 should create a publicly available black book of empirical benchmarks for popular system building blocks that will help users understand the sort of throughput they can expect to receive for different NICs, different motherboards, different motherboard chipsets, different CPUs, different types of hard disk subsystems, etc.

For example, holding other system components constant, what's the effect (if any) of replacing a Celeron with a PIII of the same speed? A PIII with a Xeon? Am I better off buying dual fast PIII's or a single (slower) Xeon with lots of cache?

Or, what's reasonable throughput for Ultra 160 SCSI? How does that compare to ATA/66 IDE disk subsystems? Is there any point to using Fibre Channel drives? Am I better off buying multiple cheap spindles rather than a smaller number of more expensive faster spinning drives with more on-drive cache? Should I buy multiple SCSI controllers, or put that money into more/higher performance disks?

It is certainly true that some aspects of this sort of benchmarking will be highly application dependent, but there should be at least some general "rules of thumb" which we can document. Note, too, that in most cases, at least for end user systems, the decision making process will be financially constrained, so the street price cost associated with various alternative strategies should be explicitly included in the analysis and report.

Complementing recommendation 2, above, Internet2 should also work to address the host performance requirements of users who prefer to buy a prebuilt system rather than building their own. That is, an opportunity for profitable Internet2/industry collaboration would be in the area of specifying and working with one or more vendors to develop an I2-optimized pre-built standardized high performance desktop.

Recommendation 3: Internet2 should partner with a major system vendor (e.g., Dell, Gateway, Compaq, Micron, HP, IBM, etc.) to offer preconfigured high performance Internet2 workstations for those who'd prefer to buy such a system (rather than building it themselves); such systems should be commercially available to any interested buyer through normal vendor sales channels.

There is also the problem of making the most of already deployed systems, many of which will be Windows 95/98/ME/NT/2000 systems used by non-technical people.

Recommendation 4: To maximize the end-to-end performance of existing Windows 95/98/ME/NT/2000 systems, Internet2 should make available "one click" registry tweaks suitable for I2 connected systems running at fast ethernet speeds, with the goal being to encourage end users to click on the registry tweaks to try to improve the throughput of their untuned systems. See this general approach demonstrated at www.dslreports.com/tweaks/Windows (albeit for cable modems and DSL speeds).

(b) Measurement Activities

Currently most campuses do *not* host any active measurement devices (ala Surveyor nodes); none that we're aware of have active measurement devices deployed ubiquitously throughout their campus network. Paradoxically, this means that users may have a better idea of the performance characteristics of paths between two remote campuses than they do of paths limited to their own LAN.

Recommendation 5: Internet2 campuses should be encouraged to build and deploy active measurement boxes (built to be compatible with already-deployed Surveyor measurement boxes) on all switched fast ethernet and switched gigabit ethernet subnets. Problems associated with the deployment of GPS-based time synchronization antennas may be a major obstacle to accomplishment of this initiative.

In addition to being able to actively monitor unidirectional delay and loss measurements, users also need information about available capacity on shared links, such as the available capacity on links from edge switches to aggregation routers, or the available capacity from campus core routers to the local gigapop.

Recommendation 6: Internet2 campuses should be encouraged to routinely measure the available capacity of all shared links using MRTG. That data should be made publicly available from a standardized hostname on each campus, and should be kept current. This is particularly important in the case of multiple downstream connectors aggregating into a common/shared gigapop, where Abilene only monitors the aggregate traffic presented to I2.

At one time, the vBNS had excellent passive monitoring (OCxMON) capabilities. Somewhere along the line, deployment of passive monitoring facilities became less common.

Recommendation 7: Internet2 campuses should be encouraged to routinely instrument their connections to Internet2 so as to permit passive monitoring via OCxMON.

Internet2 does not currently publish a weekly, monthly or quarterly report kin to the old vBNS monthly/quarterly reports, or the CANET3 weekly traffic report. Creation and publication of such reports such be immediately undertaken by the Internet2 NOC.

Recommendation 8: Create and disseminate weekly or monthly summary public Abilene traffic reports.

While end-to-end performance often is envisioned in terms of raw available bandwidth, latency, packet loss, and jitter metrics, another important characteristic of end-to-end performance is network availability/network reliability. We heard in San Diego, for example, how California astronomers are currently constrained when it comes to doing remote observing over Internet2 because they cannot be assured that Internet2 will be reliably available when they need it. Reliability is a crucial aspect of end-to-end performance.

Recommendation 9: Internet2 connectors should be encouraged to review their connectivity to Internet2, working to cooperatively identify and eliminate single point of failure vulnerabilities such as connection to Internet2 via a single circuit, use of a single GSR, or any other non-redundant network design elements. This is particularly critical at the gigapop level, where failure of a single gigapop circuit, router or switch can take multiple campuses completely off Internet2.

2. Success Metrics For the Suggested E2E Initiative Recommendations

All of the recommendations mentioned above articulate measurable/quantifiable goals. Internet2 management should establish phased target thresholds for distributed behaviors (e.g., behaviors which require the consent and participation of individual Internet2 member campuses, such as recommendations 1, 5, 6, 7 and 9, above), e.g.:

Phase	Time Frame	Target Compliance Level
I	90 days	50% compliance for each recommended practice
II	180 days	75% " " " " "

In the case of recommendations which are wholly within the control of Internet2 employees (recommendations 2, 3, 4 and 8), all articulated objective should be underway at the ninety day mark, and all programs should be fully active by the 180 days. Incentive compensation (e.g., bonus payments) should be contingent upon accomplishment of distributed and in-house goals by the stated dates.

3. Criteria to Be Used in the E2E Initiative Call for Participation

My recommendation would be to require participation by all I2 members in all of the recommended E2E initiatives described herein unless a request justifying exemption from participation is received and approved by Internet2.