Capacity Planning and System and Network Security

Internet2 Member Meeting
April 24th, 8:45-10:00 AM, Crystal City, VA

Joe St Sauver, Ph.D. (joe@uoregon.edu)
Internet2 and University of Oregon
http://www.uoregon.edu/~joe/i2-cap-plan/

Disclaimer: All opinions expressed in this document are solely those of the author and should not be taken as representing those of any other entity. These detailed slides are provided as an alternative/accessible format for those who may have difficulties hearing oral delivery of this material.

1. Introduction and Context

Why Worry About Capacity Planning Now?

- At one time, long ago in the bad old days, everyone worried:
  "Will we have enough network capacity?"
  "Will we need (and can we afford!) to buy more bandwidth?"
- Today, however, particularly with the capacity of the new Internet2 Network, worries about bandwidth sufficiency are largely a thing of the past. We now have bandwidth abundance.
- But old issues, including our ability to efficiently use that capacity, remain: end-to-end performance remains elusive for many users at many sites.
- At many sites there may also be concerns about the potential security implications associated with today's huge new pipes, and how those huge pipes may interact with campus networks and campus systems.

Our Environment is A Dynamic One…

- We're migrating to the new Internet2 backbone network
- Internet2 and NLR are merging
- There's a new ESNet network and soon much LHC traffic
- The FCC's new rural health care initiative is coming
- The Internet2 commercial peering service is active
- Internet2 has an updated AUP/COU document
- There are new and evolving applications (network video is squarely front and center these days, or so it seems)
- Middleboxes may not be keeping up
- International networking connectivity continues to evolve
- A new major desktop OS (Vista) is here with a new IP stack
- **Any ONE of those items could have a phenomenal effect on how your Internet2 connection gets used…**

Even If Your Site Isn't Intentionally Planning on Changing (Much)…

- … other sites ARE likely going through changes, and the changes they're making can (and probably will) impact your campus, whether you want them to or not.
- None of us operate in isolation, and the result can be thought of as a sort of astrophysical "N-body" problem, with each entity gravitationally affecting each other entity.
- The "N-body" problem can be a complex problem to try to solve, except via an iterative and incremental process.
- My theory, and my hope, is that the evolutionary process we're all currently engaged in will go more smoothly if we collectively pay concious attention to the operation of the network, particularly with respect to network capacity.

Some Hypothetical Capacity-Related Objectives

- Just like Goldilocks and the Three Bears, we all want the right size connection: not too big, nor too small, just right.
- We want to load that connection efficiently, so we get good value for our money, and so that our users can get their work done (and if we have excess capacity which we don't currently need, it would be great if we could use that capacity productively rather than letting it just go to waste).
- We also want to make sure that we look at capacity as an end-to-end resource, so that the local campus network infrastructure is correctly engineered to be able to support high speed flows, and our local hosts are also ready to take advantage of our high throughput connectivity.
- Finally, we also want to avoid having our connections get used as a conduit for inbound or outbound attacks

An Online Arms Race

- It is common to hear the phrase, “We’re in an arms race with the online miscreants,” typically after hacker/crackers unleash some new type of tricky-to-handle innovative type of attack. However, that sort of “victory through technical innovation” strategy is subordinate, in my opinion, to the real cyber “arms race.”
- The real cyber arms race is one of sheer capacity, where the ultimate outcome of the “war” will be determined by the capacity of the bad guys to source brute force attack traffic via bots, versus the capacity of the good guys, e.g. you, to soak up that traffic via high capacity connections and systems while continuing to do business as usual.
- **“He conquers who endures.”**
  --Aulus Persius Flaccus, 34-62 A.D.

Nation States and Terrorists

- Computer network operations ("CNO") are now an accepted part of US and foreign cyberwarfare doctrine, and because many military and governmental networks are persistently blocking large parts of the Internet, commercial service providers and colleges and universities may be some of the only meaningful online targets left (heck, they need to have SOMETHING they can still attack, right?)
- Similarly, if we believe terrorists might target tangible western interests at home and abroad, is there any reason to believe that western online activities are not also at risk?
- Do not assume that the bad guys, whether foreign governments or terrorists, have somehow managed to overlook one of the biggest and fastest networks in the world…

What About a "Laisez Fair" Approach to Network Capacity Planning?

- There are some who will suggest that our connections to Internet2 have gotten so large that the need to think carefully about network capacity planning or worry about security is over. I disagree.
- I believe that today's larger pipes, and the need to achieve a balanced end-to-end buildout, INCREASE our obligation to be good stewards of our network connections, and to actively monitor and manage that capacity. If neglected, if unattended, large, lightly loaded network connections WILL be bent and used against us, just like jet aircraft, nuclear power plants or other powerful but potentially abuseable assets.
- At the same time, those faster connections give us a potent tool to help insulate us against some types of network attack.

DDoS'ing a 10gig Connected Site: A Non-Trivial Objective

- For example, consider a **distributed denial of service (DDoS) attack,** such as when a miscreant floods your site with spoofed traffic from an army of bots. When he or she does this, clearly that person's attempting to exhaust your **network's capacity**. If you're **thinly provisioned** with little excess capacity the attacker may have an easy time rendering your connection unusable, and taking you off the air. If you have lots of headroom, as you would if you're 10gig connected, the attacker's job becomes far harder, and many more resources must be marshalled to successfully prosecute that attack. [Since there are some old school hacker/crackers out there who may actually enjoy a "challenge," let's just stipulate up front that yes, even a 10 gig connected site can be DDoS'd, so there's no need for you to demonstrate or "prove" this, thanks anyhow.]

But Even If It Is Hard to Packet Flood Your 10 Gbps Network Connection…

- … are there any less-richly connected key chokepoints on your network which the bad guys can hit instead?

- For example, what about name servers? If your domain name servers are taken out of action, even if your wide area connectivity still has capacity, you're off the air.

- So do your campus name servers represent a soft spot in your campus's security armor? Are those DNS servers connected via mere fast ethernet links? Do more than one of them live on the same subnet? If so, you're making it easy for the bad guys to take you offline, even if they can't easily flood your 10gig connection.

Network Capacity Doesn't Just Drive Network Survivability

- Capacity also impacts routine operational capabilities.
- If you end up out of capacity and your network **congests**, TCP performance and throughput will drop as applications back off in the face of that congestion. UDP applications may exhibit jitter, dropped frames or other undesirable artifacts.
- We also know that network activity tends to exhibit substantial peaks and valleys over the course of a day, and even if you have enough capacity to handle **AVERAGE** loads, it is easy to end up with insufficient capacity to handle intraday **PEAKING** loads.
- Finally, if you don't have some headroom, you won't have much flexibility when it comes to accommodating natural **GROWTH** which will inevitably occur over time.

Capacity Can Also Affect Your Institution's Competitiveness

- When your school is competing for grants, one potentially influential factor may be the resources available to support a proposed program of work.
- Imagine a grant opportunity which involved a substantial program of network-related work. Now suppose that you are a program officer at the granting agency, and you're looking for criteria which may help you decide who should receive your award. Some applicants have made a major investment in network facilities, while other applicants have made less of an investment. Everything else being equal, where would you award that grant?
- Similarly, if you're a hot and very marketable faculty member, where are you more likely to go: an institution whose network is limping along on the edge, or one with abundant capacity?

Excess Capacity: Essential For Encouraging Innovation

- Systems with limited capacity tend to discourage experimentation simply because there's no "headroom" which can safely be "played with." What you've got is so tight it is mission critical, and you can't afford to jeopardize that make-or-break capacity.

- Systems with abundant resources, on the other hand? Take a chance, try an experiment, thereby risking failure (or possibly netting your institution a ***spectacular*** success)

- Abundance gives us the ability to dare.

Slack As A Type of Engineering “Insurance”

- In engineering, slack, or excess capacity, provides a tolerance for unexpected loads and imperfectly modeled forces. Slack provides resilience. Slack provides a margin for error and a buffer for safety. Slack is a type of engineering “insurance.”
- No one complains when a bridge is able to withstand an earthquake because it was "overengineered," or a plane is able to continue flying even after one engine stalls, or a boat stays afloat even when it is overloaded – the redundancy and excess capacity built into those systems saves lives.
- And yet, we explicitly know that slack, excess capacity -- whatever we want to call that "extra stuff" -- costs money, and our engineering compulsion to build systems with a healthy margin of slack runs smack into business realities: excess capacity is, or can be, expensive.

Capital, Projects and Priorities

- When the financial analysts and the managerial accountants take a hard look at capital expenditures, they're not doing it just because they identify with Ebenezer Scrooge. They've got their sharp red pencil out because they're trying to identify any avoidable or deferrable capital expenditures. Why? Capital is always in short supply, and if one project is done in a so-called "sloppy" way, with tons of excess capacity, that may mean that another project may not be able to be begun.
- For example, assume Internet2 had a 100Gbps connection service available (even if there's no site on Internet2 which currently needs (or can use) that level of connectivity today). If a site bought that 100Gbps service, and from a geek's perspective it would unquestionably be cool, the money that would be spent on 100Gbps service couldn't be spent on other things which might need today, such as beefed up servers, a faster campus network, more staff or better software tools.

Chasing Demand: Can We Add Capacity Just When We Need It, Perhaps?

- If we had had **infinitely flexible provisioning systems**, we could:

  -- **add capacity just as it was needed**, and then
  -- subsequently **shed that capacity** (and its associated cost) as soon as that capacity was no longer required.

- This approach of "chasing demand" is commonly seen in the service industry and some seasonal businesses, but most network providers aren't willing to sell X capacity during most of the school year, and one half X capacity on weekends or during the slow summer vacation period, sorry.

One Other Slight Problem With Chasing Demand: Lead Time

- If we could add capacity on demand with zero advance notice (or zero “lead time”) we’d have a far easier time when it comes to managing our capacity requirements. If there was no need to plan ahead, we could just add (or remove) capacity as required, adjusting our capacity to meet our empirically observed requirements.
- In Real Life™, however, we need to “pull the trigger” on orders for additional capacity in advance of the time we actually need that capacity. It takes time for orders to be approved, and for purchase orders to be cut, and for gear to be shipped and installed and burned in and configured and integrated into production. We need to plan ahead.

Can We Reduce Required Capacity By Shifting Load To Off Peak Times?

- If we had **infinitely flexible customers**, we could shift some load from prime time to periods of low demand (traditionally early morning or late night hours), thereby leveling out our load profile and reducing the amount of capacity we need to provision.

- This can be somewhat difficult to do with interactive network loads (users don't want to have to surf the web on the graveyard shift!) but there is some potential for moving non-interactive background tasks (such as running cron'd jobs synchronizing research data archives) over to non-peak periods.

Could Adaptive Pricing Help Shift Loads to Off Peak Times?

- If the economists had their way, we'd be able to use pricing to "help us" adjust the loads we face. Want breakfast at 7AM, same time as everyone else? Okay, you're going to pay regular price. Willing to eat a little earlier or a little later, say at 9AM instead? Boy, have we got a deal for you! Cell phone providers and airlines are notorious examples of this.
- Differential pricing can allow the network provider to recover the higher cost of provisioning peaking capacity, while offering flexible customers a real potential bargain.
- Unfortunately, in general in academia, users don't pay to use the network in the first place, so our ability to use discounted pricing as a demand shaping tool is limited.
- Potentially one might even envision schools PAYING users to move their peaking load traffic to an off peak time.

Will We Let Congestion Occur to Help Shift Load?

- If you fail to add capacity, and just “let things get really bad,” you may be able to encourage users to shift their load if only to avoid crumby performance (but I’m not much of a fan of this approach to capacity management).

- So what **can** we do to more professionally manage our capacity?

  Purchase an appropriate size connection.

2. What Size Connection Should I Buy?

The Internet2 IP Network "Menu"

- Options for IP network connectivity are relatively granular and limited at this time (100 Gbps connectivity for end sites, cool though it would be, is not yet an option :-)).

- Ignoring some lower speed legacy options, at this time you've got three IP network capacity options:

  -- a gigE (1,000Mbps) at $250,000/year

  -- 2.5Gbps (2,500Mbps) at $340,000/year, or

  -- a 10gigE (10,000 Mbps) at $480,000/year

What About Buying 2<N>, etc.?

- Because of the pricing model which has been implemented, it doesn't make sense to try to buy "multiples" or combinations of the slower speed 1Gig or 2.5Gig packages.
- Why do we say that? Well:
  -- 2xGigE costs more than one 2.5Gbps circuit while delivering less capacity
  -- 2x2.5Gbps costs more than one 10gigE while again delivering less aggregate capacity
  -- 1x2.5Gbps plus 1xGige costs more than one 10gigE, etc.
- Thus, the three base options are the only IP network capacity “steps” which make financial sense. [We'll ignore redundancy and survivability issues since redundant backup connections can be provisioned at a relatively low marginal cost if they're for emergency/backup use only, and won’t be routinely passing additional production traffic.]

So How *Do* Connectors Decide What Size Pipe to Buy?

- From talking to folks, as far as I can tell, several different approaches may get used:
  -- some connectors may start with the smallest connection available, upgrading only when traffic demands that they do so
  -- as a matter of leadership, or based on anticipated future traffic requirements, or to avoid the pain of having to arrange for an upgrade, some connectors may start with the largest connection which is currently available
  -- or, some connectors may "hedge their bets" and "split the | distance," buying something in between.
- Note that unlike earlier days, the assumption/expectation today is that most connectors will be connecting at 10gig, not at a slower speed.

Thinking Through The Cost/Mbps/Month

- From a cost perspective, given known steps and costs, we can envision a usage-based cost recovery structure which will allow sites to readily move from one tier to the next…
- For instance, assume your site buys a 1Gbps connection which costs $250,000/year. If that connection were to be fully utilized, each Mbps would cost $20.83/Mbps/month ($250,000/1000/12=$20.83/Mbps/month).
- In general, however, the pipe will be less than full, and there are local costs associated with delivering that capacity, so lets assume that the billable cost/Mbps is higher, perhaps $30/Mbps/month. At that price, you'll break even at 694 Mbps, and if/when the pipe is full (at $30/Mbps/month), you'll be seeing revenue of $30*1000*12=$360,000/year
- Coincidentally, a 2.5Gbps connection costs $340,000/year, so you'd be in good shape to handle an upgrade…

The Disconnect

- Unfortunately there's a disconnect in that model: pricing for gigabit class connections would be "way too expensive" at $30/Mbps/month, at least relative to other network pricing.
  -- Consider a 10Gbps Internet2 connection: when you buy 10gig for $480,000/year, you only pay $4/Mbps/month. Even if you marked that up 250%, $10/Mbps/month is still only a third of what gigabit connection costs on a per Mbps/month basis.
  -- $30/Mbps/month is also high relative to consumer pricing: people have become accustomed to seeing for things like home cable modem service (which might offer 6 to 16 Mbps downstream) for less than $60/month, and Verizon's new FIOS service (which offers up to 15Mbps down) is only $49.99… yes, I agree that consumer broadband pricing is often cross subsidized by other services, and shouldn't matter for this sort of thing… but it still does.[27]

An Obvious Observation

- So given that you can buy capacity at a desirable price (e.g., $4/Mbps/month when you buy a 10 gigabit connection), the "trick" is:

  -- finding a group of partners with whom you can share the capacity (and cost) of a large connection, and

  -- handling the "fan out" or distribution of that capacity from the point where it is delivered to the partners who've helped to buy it

- This is the model that was the foundation for both Internet2's original gigapop architecture, and for later regional optical networks: multiple participants share a single connection.

There's No Free Lunch

- Yes, you can buy bulk bandwidth at incredibly cheap prices per Mbps, but in doing so you have taken on the role of deaggregating and distributing that bulk capacity.
- You know your approximate budget for that process: you can afford to spend the difference between what users would otherwise have to pay and your wholesale cost.
- If you have lots of nearby partners, or they already have their own fiber, this can be easy to make work.
- If you live in a sparsely populated part of the world or need to lease circuits, the budget numbers may simply not work, and you may be better off buying a smaller capacity connection directly, since those rates are largely postalized (modulo backhaul to the closest carrier point of presence).

“Tie This Back Into IT Security For Me Again, Please?”

- Happy to! Many times it is hard (or expensive, or impossible) to construct a properly architected regional network. You may be lucky to get even a single fiber path to a remote site, and there may simply nothing available which might work as a redundant path -- if you want to connect to some remote sites, it will simply have to be a spur, undesirable as that is.
- Now recall that information security has three objectives: **confidentiality**, **integrity**, and **availability**. A non-redundant network design directly goes to that third item, availability. There’s a real risk that at least some sites may end up chasing inexpensive bulk capacity at a real (if non-financial) cost, and that non-financial cost may be network availability.
- But let’s assume that a partnership has been formed, and you will be splitting a 10gig circuit.

Sharing a 10gig Packet Connection

- There are many ways that a ten gig connection could be shared.

- Just to mention two basic approaches:

  -- you could **pool access to the entire resource** (e.g., each partner could potentially use the entire 10gig connection if no one else is using it), or

  -- you could **have dedicated hard allocations** (e.g., you could create ten gig connection "shares," or 100 100Mbps “shares,” and sell one or more to each partner for their exclusive use)

Choosing a Connection Sharing Model

- Choice of one connection model over another may seem like a trivial matter, but it can have subtle and important implications.
- For example, assume that you use a pooled model, with all members of the consortia sharing the entire pool of bandwidth. If one member of the consortia experiences a distributed denial of service attack all the other members of the consortia will also suffer.
- On the other hand, when you're using a pooled model, any one participant has far more headroom than they could buy on a dedicated basis, and thus the miscreant will need to work far harder to acquire enough attack capacity to take his or her target offline.

The Dedicated Share Model Eliminates The Need To Worry About Partner Usage

- The “dedicated shares” model, on the other hand, can eliminate a lot of potential hard feeling and misunderstandings which might otherwise pop up in conjunction with the pooled model. In a dedicated shares model, what you buy is what you get, and there’s no need to worry that one partner or another is using “more than their fair share.”
- The dedicated shares model also insures that if one partner does get attacked, the impact of that attack traffic is limited to just that one partner. This is a very nice feature if you’re an “innocent bystander” partner, but if you’re the attacked partner, you may suddenly find it hard to find a lot of other people interested in the fact that you (and only you) are under attack.

3. Making Good Use of Your Network Capacity

So You've Got a Connection, Now What?

- Once you have a connection, whether great or small, dedicated or shared, it's now time to use that capacity. A connection that's not used doesn't do much for your institution!

- In the following discussion, because a ten gig connection is the new Internet2 standard, we'll assume where relevant that that's what you're using.

Appropriate Utilization is NOT 100%

- While you want to use your connection, the goal is NOT to see your connection pegged or "flat topped" all the time. If you ARE seeing your connections to Internet2 in that state, that's a sign that things are **underprovisioned** and need additional capacity (I don’t think anyone is currently flattopping).
- On the other hand, it does no good to have restrictive usage policies which result in connections languishing substantially unused or underused. You can't "save" any bandwidth you're not currently using, so if you have capacity, and you have a reasonable use for that capacity, you might as well use it.
- Unfortunately, even with fast wide area connections, it is routine for some campus users at some sites to complain that they're not seeing very good throughput.

Host Tuning and End-to-End Performance

- It is well known that without proper system tuning, even a powerful workstation may not fully utilize a gigabit connection, and that same system may sometimes have trouble even fillinging a switched fast ethernet connection.
- A good quick first check that's always worth trying is NDT, a java based network diagnostic which can identify many common issues including things like duplex mismatch, etc.
- Beyond that, the classic resource for host tuning has long been PSC's "Enabling High Performance Data Transfers," see http://www.psc.edu/networking/perf_tune.html , but I'm bold enough to ask for systems which would ship from the vendor with **autotuning** network stacks, thereby eliminating the need for network "wizards" to manually tweak options and adjust buffers, etc.

Vendor Progress w.r.t. Auto tuning

- Collectively we've made some progress w.r.t. auto tunng:
  -- recent Linux kernels do autotuning, although at least in some cases the default maximum values achievable via auto tuning may still be too modest; see the discussion at http://www-didc.lbl.gov/TCP-tuning/linux.html
  -- Microsoft Vista has a new TCP kernel which is working to address some common TCP performance issues, see: http://www.microsoft.com/technet/community/columns/cableguy/cg1105.mspx
  -- Apple has a "Broadband Tuner" patch for 10.4; see www.apple.com/support/downloads/broadbandtuner10.html (not auto tuning, but at least a step in the right direction)
- But does any of this host tuning matter if most users are still running Windows XP or your servers are using ancient versions of Linux? You need to make a **conscious effort** to get users onto current versions of the OS which they're using.

Jumbo Frames

- We also all know that the default ethernet frame size, or "MTU," is too small (at just 1500) bytes to deliver optimum performance over wide area, high capacity connections.

- Although Internet2 and other high performance networking organizations have long endorsed the use of jumbo frames (see http://noc.net.internet2.edu/i2network/documentation/policy-statements/rrsum-almes-mtu.html), practical considerations have historically limited the deployed utilization of jumbo frames (see, for example, "Practical Issues Associated with 9K MTUs," http://www.uoregon.edu/~joe/jumbos/jumbo-frames.ppt or .pdf)

But We're Making Progress In This Area, Too

- I'm happy to say that despite some of those historical practical issues, progress **is** being made when it comes to getting jumbo frames deployed:
  -- RFC4821, "Packetization Layer Path MTU Discovery" was published in March 2007 by Mathis and Heffner, and as that protocol begins to be implemented we'll see an improved ability for senders and receivers to negotiate support for larger-than-normal frames
  -- a growing number of Internet2 connectors are known to support connections at or above 9000 bytes (see http://dc-2.grnoc.iu.edu/vn/analysis/connector-tech.html)
  -- jumbo frames now account for between **five and ten percent of all bulk TCP transfers**; see http://netflow.internet2.edu/weekly/longit/perc-b-jumbo-packets.png
- Are **you** taking advantage of jumbo frames at **your** site?

Routes

- A third technical factor which may contribute to underutilization of some Internet2 connectivity is the unnecessary announcement of just a very limited set of overly specific routes by an Internet2 university site.
- That is, rather than announcing **all** their address space, some schools may elect to only announce a subset of their netblocks, perhaps intentionally (but unnecessarily) limiting Abilene usage to just specific network engineering or advanced scientific facilities on campus. Announcing only a subset of routes that way will contribute to lower than optimal levels of utilization from their Internet2 connection.
- In general, unless you have a compelling legal or policy-related reason which keeps you from doing so, Internet2 universities should be announcing all their address space via Internet2, not just special labs or other facilities.

Speaking of Routing…

- Besides announcing unnecessarily specific routes, some sites with multiple network connections may find that their traffic is using the “wrong” connection.
- For example, assume you have:
  -- partners connected via a statewide or regional network
  -- commercial peering via a local exchange point
  -- Internet2 connectivity
  -- commodity transit (regular Internet connectivity) from one or more providers
- That list is probably the order you’d prefer traffic to take: traffic should go via the statewide or regional network first, if possible, then via the local exchange point, then via I2 if neither of those other options work, and then, and only then, via commodity transit if all other options aren’t a possibility.

But Some Sites Still Prefer High Priced Networks Over Cheaper Options

- Despite having a number of other options, it is not uncommon to see some Internet2 connected sites that prefer expensive commodity transit paths over cheaper paths which might also be available.
- An important part of your capacity planning is making sure that you've got sane preferences correctly instantiated.
- You should also be on the lookout for route asymetry, a condition where traffic enters your network from one path, but exits from via different one. For a variety of reasons, you'd usually prefer traffic flows to be symmetric.
- When you've looked at other issues, don't forget to check your routing preferences, too.

Middleboxes

- A fourth technical factor which may constrain the utilization of a site's capacity is the presence of firewalls, packet shapers or other "middleboxes."
- These boxes often "top out" at gigabit speeds or below, acting as choke points on higher performance connections.
- Even if a middle box has interfaces rated at a particular speed (such as gigabit or ten gigabit), the ability of that box to **forward traffic** at full line rate, when configured with your production rulesets, should be explicitly tested & confirmed.
- Whenever possible, at least when performance matters, middle boxes should be eliminated entirely, or at least moved as close to the edge of the network as possible. The fewer middleboxes in the network path, the better.
- Some may wonder: "But what about peer to peer traffic????"

So What About Peer-to-Peer Traffic?

- For a number of years, peer-to-peer traffic was a tremendous concern for many schools because of the bandwidth it consumed and because of the potential copyright concerns associated with some content (e.g., see "The Case for Traffic Shaping at Internet2 Schools," I2 Joint Techs January 2002, http://www.uoregon.edu/~joe/i2-traffic-shaping.ppt)
- While a dramatic increase in capacity doesn't affect potential copyright issues associated with some peer-to-peer file sharing, at 10gig the **bandwidth issues associated with P2P traffic may now be just a historical artifact.**
- This may be a very good thing, since support for deep packet inspection and traffic shaping by commercial middleboxes often hits its limit at gig speeds (although you could try running multiple packet shaping appliances closer to the network edge if you have to).

Network Video

- While we're talking about bandwidth issues and P2P, we should also mention network video, since network video is commonly viewed as the new "P2P-like" bandwidth threat, particularly given the popularity of some network video resources such as YouTube, Google Video, etc.

- Again, at 10 gigabits per second, network video bandwidth, like P2P traffic, ceases to be a material issue at most sites simply because there's so much bandwidth available.

- In fact, the same thing can be said of pretty much **any protocol** you care to mention – at 10 gig, it's hard to think of **any** protocol (even Usenet news!) which isn't just "noise" on the wire.

QoS

- Some types of traffic (such as VoIP) can be quite sensitive to network congestion, and it is common to hear consultants urge sites to deploy QoS to prioritize and protect that VoIP traffic on converged networks which carry a mixture of data, video and voice traffic.

- An alternative to QoS is overprovisioning.

- Well, with ten gigabit connectivity, you are truly very well overprovisioned, and so along with disconnecting your packet shaping appliances, you should also stop worrying about QoS.

A Potential Chokepoint: The Campus Core (and Below)

- While wide area ten gigabit connectivity is potentially wonderful, in some cases the campus backbone will not yet have been upgraded to ten gigabit, ditto circuits to servers and workstations.
- Obviously those sort of campus choke points will limit the wide area throughput that will be seen from that campus unless/until either:
  -- the campus core is upgraded, or
  -- dedicated gigabit or ten gigabit connections are routed around the campus core and direct to the ten gig-capable border router
- Has **your** campus made the appropriate investments in your campus core and edge circuits?

Campus Web Proxy Servers

- Just as the campus core can be a chokepoint at 10gigabit, so can things other infrastructure boxes such as campus web proxy servers (assuming you're at a site that uses them).

- As was the case for other boxes, your best bet may be to work to take those campus proxy servers out of service, or at the very least you should consider upgrading and enhancing those servers to better match the capabilities of the network.

"I'd love to do all that, but at least some of my wide area links aren't 10gig!"

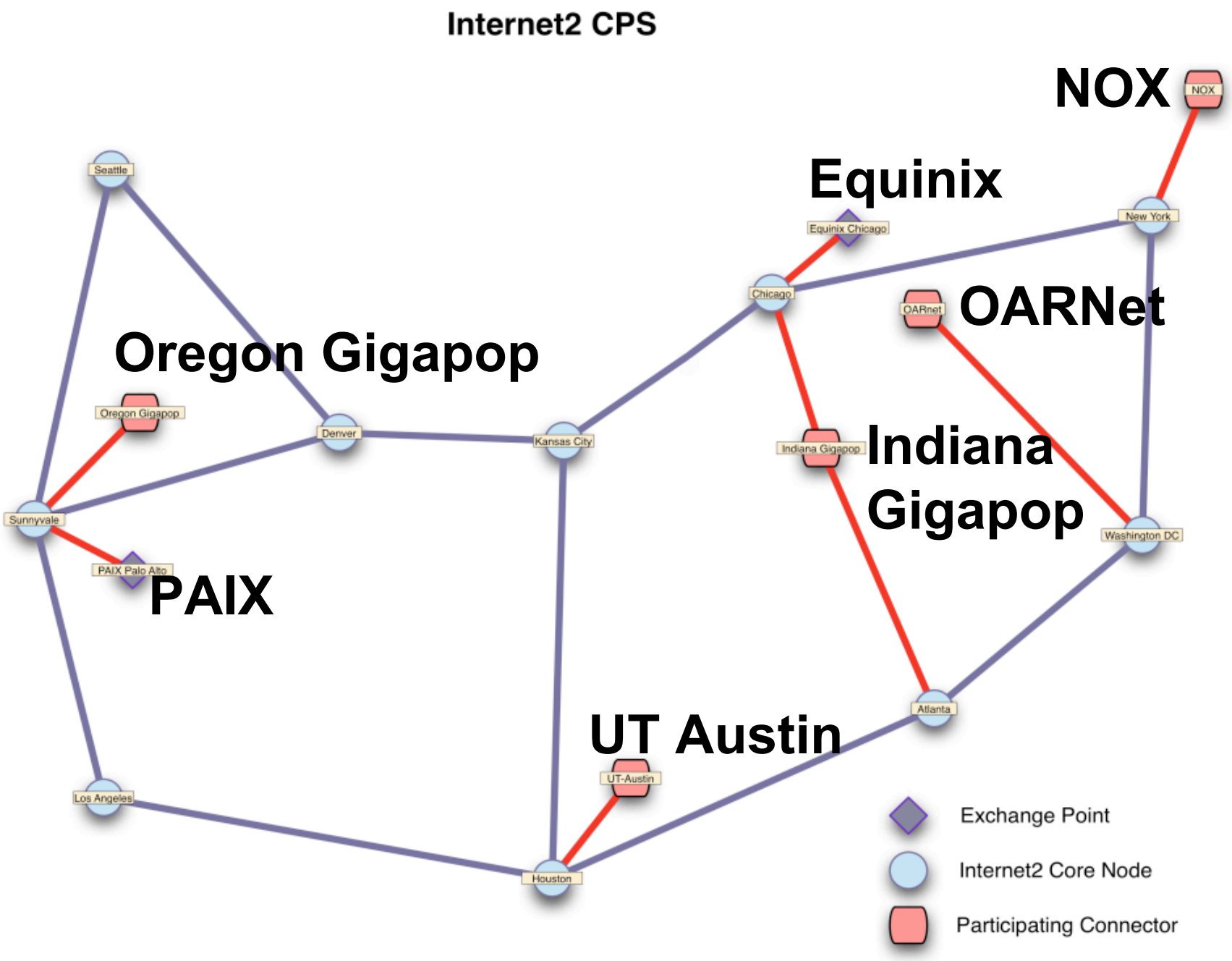
- Unquestionably, things get a bit more complicated when some connections are 10gig while other connections aren't.
- This is the contemporary version of the issue I originally raised way back in April 2000 in San Diego, at the Campus Focused Workshop on Advanced Networks (see "Going Fast(er) on Internet2," http://www.uoregon.edu/~joe/how-to-go-fast.ppt):

  You can't go fast JUST on Internet2 – you need to consider both your Internet2 connectivity and your non-Internet2 "regular Internet" connectivity

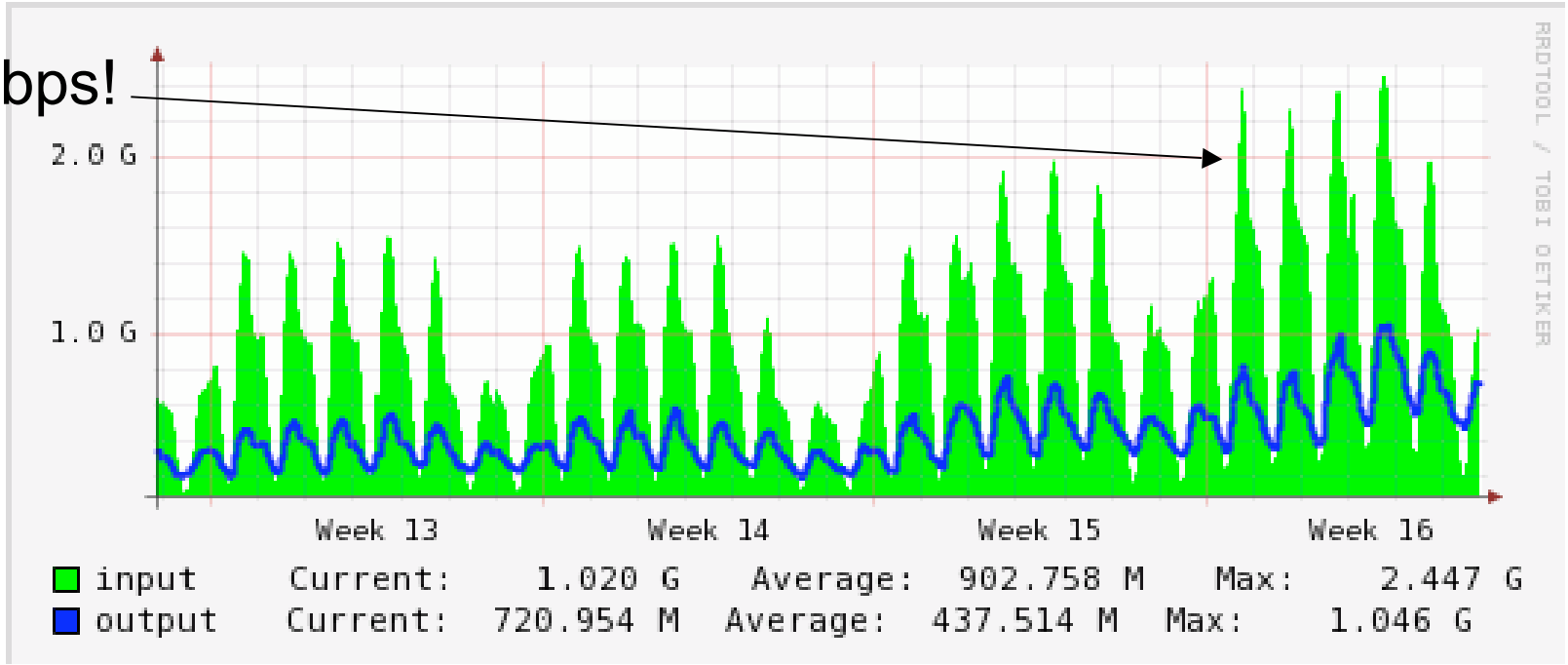
- That reality is one reason why I'm very happy to see things like the Internet2 Commercial Peering Service emerge…

Internet2 Commercial Peering Service

- The Commercial Peering Service (CPS) is a separate layer 3 routed IP service, run over the connector's existing Internet2 connection, using an MPLS Layer 3 VPN (aka "VRF")
- Depending on the connector's traffic mix and current connectivity, CPS may offload from 10-40% of the connector's regular Internet traffic, but this can vary widely from connector to connector, and may change over time.
- There's currently no additional charge when connectors add the Commercial Peering Service
- Internet2 CPS participants at this time include the Indiana Gigapop, Northern Crossroads, OARNet, the Oregon Gigapop, and the University of Texas at Austin.
- For more information see: http://www.abilene.iu.edu/i2network/commercial-peering-service.html


Internet2 CPS
NOX
NOX
Equinix
Equinix Chicago
New York
Seattle
Chicago
OARnet
OARNet
Oregon Gigapop
Oregon Gigapop
Denver
Kansas City
Indiana Gigapop
Indiana
Gigapop
Sunnyvale
Washington DC
PAIX Palo Alto
PAIX
Atlanta
UT Austin
UT-Austin
Los Angeles
Houston
Exchange Point
Internet2 Core Node
Participating Connector

Aggregate CPS Traffic At Equinix



Why Is the CPS Important?

- The Commercial Peering Service is important for many reason, a few of which include:

  -- It is hard to beat the out of pocket incremental cost for this service :-)
  -- CPS uses Internet2 connectivity which might otherwise be idle, thereby adding value to the capacity of those connections
  -- Many resources important to the higher education community are still only accessible via the commodity Internet; the Commercial Peering Service improves access to at least some of those resources
  -- The CPS raises the credibility of the higher education networking community with the commercial networking community

The CPS Doesn't "Go Everywhere"

- Because the CPS is a peering service, rather than a transit service, it doesn't offer routes or connectivity for the entire commodity Internet -- you'll only see routes for providers who directly peer with the CPS, and then only if your site is one which is participating in the CPS service (and no more preferred route is available).
- **Q.** "How can I check if a provider is accessible via the CPS?"
  A. One way to check is by using telnet to connect to route-views.oregon-ix.net and then try doing a traceroute:

  % telnet route-views.oregon-ix.net
  Username: rviews
  route-views.oregon-ix.net> traceroute <domainname>

  If you see a reference to the Oregon Gigapop and to "MPLS Labels" in the traceroute output, traffic to that site is going via the Internet2 CPS from Oregon RouteViews.

4. What About Optical Waves?

It’s Not Just a Packet World Anymore

- In addition to traditional packet-based network connections, the new standard network paradigm also includes a dynamic optical wave connection deployed alongside the packet mode connection.
- How will the capacity of that lambda be used?

Sharing Lambdas

- Will a connector grant access to that wave…

  … on a first-come, first-serve basis?

  … on some sort of prioritized, pre-emptable basis, allowing low priority applications to use that capacity on an as-available basis, subject to pre-emption by higher-priority critical applications?

  … on a strictly scheduled basis? (when thinking about scheduling, remember that you need to schedule both ends of the circuit & have capacity "in the middle" too; scheduling can be a very painful issue to wrestle with!)

  or might that dynamic wave be **broken up** into smaller capacity sub-wavelengths, to be used in parallel?

Are There Security Issues Associated with Dynamic Waves?

- Some may worry that the capacity associated with dynamic waves might somehow pose a security risk or exposure, since the dynamic waves may bypass a connector's conventional security infrastructure.
- If anything, the dynamic wave infrastructure should generally be **MORE secure**, rather than less secure, since the wave infrastructure interconnects just one hosts (or a very limited number of hosts) on each end of a closed, circuit-like, connection.
- For example, given that closed, known population, you should NOT be seeing scans or other probe-like behaviors, which is probably good given the potential number of ports and hosts one could quickly scan via a 10 gigabit connection!

Are There Any Security-Driven Opportunities Associated With Wave Based Capacity?

- I could envision a site potentially using dynamic waves to reroute selected attack traffic to an offsite network facility for forensic analysis, perhaps even intercepting that traffic before it got all the way to the target campus.
- I could also see disaster recovery planning driving demand for static waves. How/why? Assume that realistic disaster recovery (largely constrained by acceptable recovery time objectives, system complexity, and data volumes) requires the creation and operation of a remote hot site with synchronized data. Data flowing between a primary filer and a secondary offsite filer has all the right properties: point-to-point flows, high volume, ongoing, potentially unencrypted, and institutionally important (and thus worthy of funding)[60].

5. Capacity Forecasting

Our Crystal Ball Is A Little Cloudy

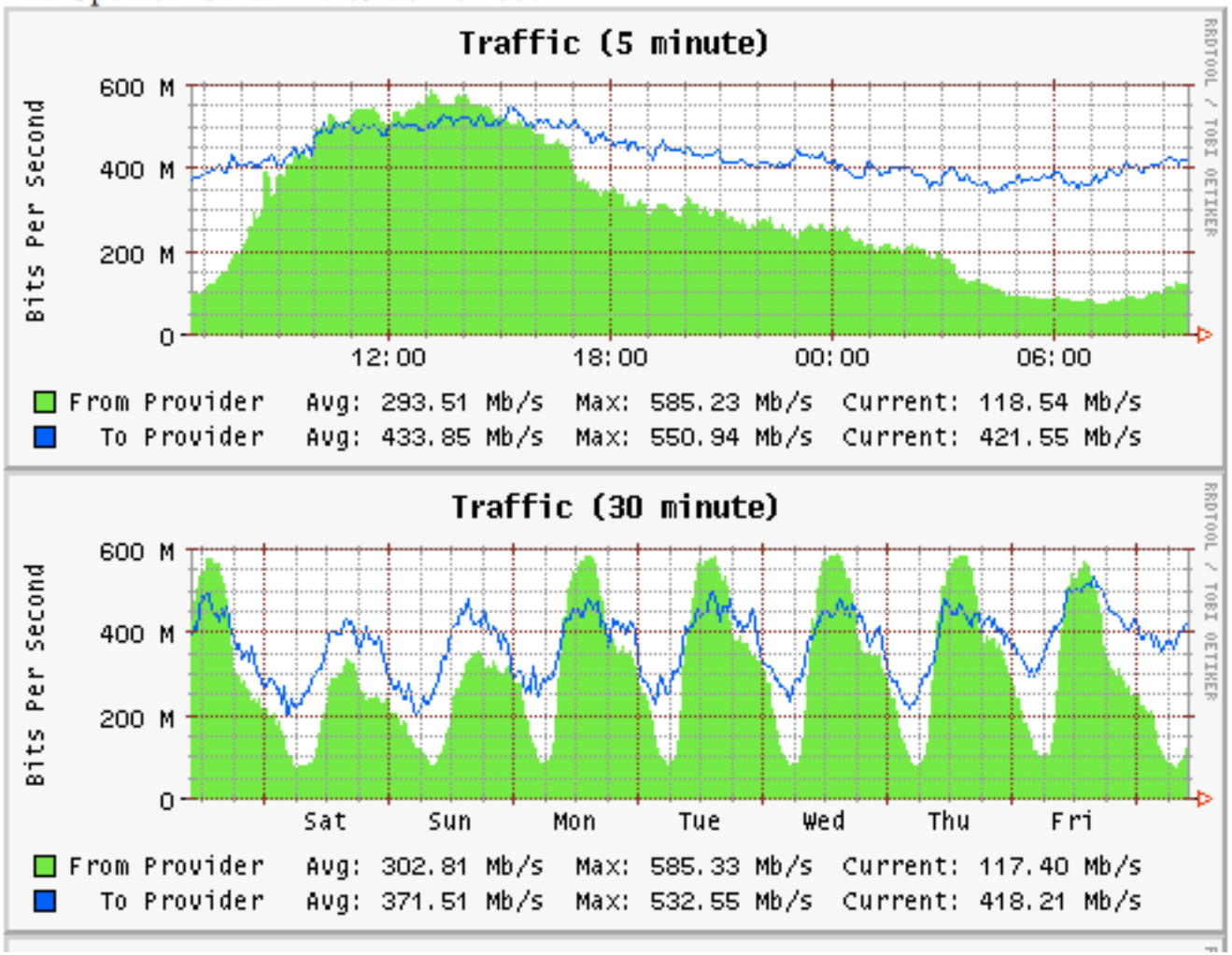
- While we know how much capacity we're using now, and we also know how much capacity we've used in the past, forecasting future capacity requirements is a little less certain.
- If we could accurately forecast the future with perfect certainty, we'd know (also with perfect certainty) the amount of capacity we'd need, and we could then very efficiently plan to deploy just that much and no more.
- In this world, though, none of us are "Svengalis" and we need to do the best we can with the limited information we have available to us.
- One thing we need to do is to make sure that we collect the information we need -- we all need to become data driven sons of guns. Most data of interest will come from SNMP.

A Quick & Incomplete Overview of SNMP

- SNMP is the “Simple Network Management Protocol” and has long been associated with monitoring and managing things like routers and switches (SNMP can also be used as a way to collect data from host systems).
- SNMP data is collected by polling SNMP-enabled devices via a network management system, snmpget, or other tool.
- To retrieve data, a user typically needs:
  -- the FQDN or IP address of the SNMP managed device
  -- the “community string” (or password) for SNMP access (all too often this is just “public” for read-only access)
  -- the object ID (or variable name) of the MIB (management information base) of interest, normally a series of numeric values separated by dots
- The value of that OID can then be periodically polled, and will often be graphed using MRTG or RRDtool.

Sample SNMP-Derived Graphs

Last Updated: Sat Jan 27 09:35:16 2007



SNMP Limitations

- SNMP is truly a very SIMPLE (“primitive?”) protocol. E.G.:
- SNMPv1 was not very secure (commands, data and community strings were all passed “in the clear,” and thus were easy to eavesdrop upon)
- It was tedious to walk all subelements of a MIB branch
- Counters would often roll over rapidly due to their limited range
- General access to SNMP data often has to be limited via firewalls or router ACLs, because SNMP-using devices may not have the ability to control access themselves
- Subsequent versions of the SNMP protocol addressed these issues -- but at a price of additional device complexity. As a result you may still see many simple network devices that only “speak” SNMPv1
- SNMP is still hugely popular and useful for collecting data.

Warning

- There has recently been increased security interest in network monitoring and management software, with material vulnerabilities found in some popular packages. It is extremely important that you keep all software you use on your network management station up to date, and you should harden and shelter your network monitoring and management station from miscreant attention.
- An example of the sort of thing that's being found is shown on the following slide…

Cacti Command Execution and SQL Injection V

Secunia Advisory:	SA23528
Release Date:	2006-12-28
Last Update:	2007-01-18
Critical:	 Highly critical
Impact:	Security Bypass Manipulation of data System access
Where:	From remote
Solution Status:	Vendor Patch
Software:	Cacti 0.x
CVE reference:	CVE-2006-6799 (Se

Description:
rgod has discovered four vulnerabilities in Cacti, which can be exploited by malicious people to bypass certain security restrictions, manipulate data and compromise vulnerable systems.

1) The "cmd.php" and "copy_cacti_user.php" scripts do not properly restrict access to command line usage and are installed in a web-accessible location.

Successful exploitation requires that "register_argc_argv" is enabled.

2) Input passed in the URL to cmd.php is not properly sanitised before being used in SQL

“Okay, I’ve Got Data and Graphs…”

- Graphs **may** be all you need -- they may show an interface that’s already flattopping, for example, or you may see clear evidence of trending, or total usage may be low (if that’s the case, the focus should shift from forecasting to the issues mentioned in section 3 of this talk). In other cases you may need or want to build a formal statistical model, using that to predict future demand.
- **One very important thing to note about graphs which average SNMP counters over an interval**, or graphs which average subsequent already averaged values: those values will consistently **understate the actual instantaneous load**. When you average values seen over an interval, the reported average value is ALWAYS lower than the actual peak value, assuming there’s any variation during the measurement interval at all.

“I Can’t Fit a Trend Line to This Data!”

- When non-statisticians first try working with periodic time series data like that shown in many MRTG graphs, they arere often thrown by the cyclical nature of that data -- who could fit a trend line to something that looks so wavy????
- While this is not the right forum to go into forecasting models in any depth, there is one basic trick that may really simplify the process of working with cyclical time series data to remove cyclical effects, difference your time series based on its periodicity; for example, if your data is collected on an hourly basis, subtract the value you saw 24 hours ago from the current value, and work with the differenced value…
- If you build a formal linear model using hourly data, you may want to include terms for 23 of the 24 hours, e.g.: demand=f(date, t0, t1, t2, t3, t4, t5, … + constant)

My Data Is Really "Noisy!"

- Other times the problem is not that your data is cyclical, but that it is noisy, or prone to a tremendous amount of fluctuation on an observation-to-observation basis.
- When that's an issue, you may want to consider computing a moving average, replacing each point with the average of the preceding four or five observations -- experiment with averaging different numbers of observations to see what degree of averaging helps you get a pragmatically useful set of smoothed values.
- Ultimately, however, you should have no issue predicting capacity requirements at least at the level of granularity that current decision functions require.

Thanks for the Chance to Talk Today!

- Are there any questions?