Running a Server? How Does It Score on the Server Administration Self-Assessment Scorecard?

IT Security: A Call to Action for the Education Community 3:20-4:30, November 1st, 2006, Ramada Plaza, Fargo ND

Joe St Sauver, Ph.D. 235 Computing Center University of Oregon, Eugene OR 97403 joe@uoregon.edu or 541-346-1720

http://www.uoregon.edu/~joe/sasas/

A Note About the Format of This Talk and a Disclaimer

- I've prepared this talk in some detail so that:
 - -- it can be followed by those who may not be present when the talk was originally given,
 - -- to insure that the contents of the talk are available to those in the audience who may be hearing impaired, and
 - -- to minimize the need for audience members to jot down notes.

Having a talk that's prepared in some detail also helps keep me on track.

 Disclaimer: all opinions expressed in this document are strictly my own. Independently assess and reconfirm all recommendations presented, and note that even if you follow all recommendations given here, you may still experience a security breach.

Section 0. Introduction

Living With Distributed Servers

- It's a fact of life these days that many mission critical servers reside outside central IT, in departments or in individual faculty/staff offices or labs.
- While those distributed servers are often run in a very professional manner, sometimes limited funding or other circumstances bound what's practically possible.
- Other times those responsible for the distributed servers may simply never have been trained on even basic elements of server administration and operation – they're learning what's needed "on the job," sometimes via very painful lessons.
- When you get right down to it, if a mission critical server goes down or gets hacked, it's the <u>institution</u> that gets hurt most, regardless of whether that server is run by central IT or by an individual department. It is therefore in the <u>institution's</u> best interest to encourage good sys admin practices <u>everywhere</u>.

Where, Oh Where, <u>Are</u> Our Mission Critical Systems?

- Many schools do NOT have a complete inventory of their distributed mission critical systems. They simply **don't know** <u>where</u> at least some mission critical systems live, or <u>who</u> runs them, or <u>what data</u> may be on them.
- If you don't know the who/what/where of your school's mission critical systems, you can't do a centrally coordinated risk assessment – you may need to rely on the operator of those systems to self-identify themselves and their servers.
- Unfortunately, in some cases, there may be reluctance to <u>publicly</u> acknowledge mission critical distributed systems if there's been a history of central IT opposing shadow systems or hindering the "exfiltration" of data for use on those servers.
- One answer to that problem? Produce and release a server administration self-review which can be completed <u>privately</u>.

The Art of "The Possible"

- The server self-assessment we'll be talking about today is actually the third version which we developed -- the first version, a one page checklist, was found during review and testing to be too short; the second version, a much longer and more detailed instrument, ended up being too long.
- The trick? Balancing a set of potentially conflicting practical constraints. For example:
 - -- Simple checkbox items are easy for people to go through, but many times, rather than a simple black-and-white pass or fail response, there may be a continuum of possible responses, some better and some worse, which implies the need for a scaled numeric value for responses.
 - -- There are an awful lot of important things to consider when it comes to server administration, BUT if a self-assessment gets too long, people simply won't complete it. 6

So What Key Areas Did Our Server Self-Assessment Scorecard Need to Cover?

- Ultimately we figured that there were probably seven key areas we needed to touch upon:
 - 1. The server's hardware (including the server's machine room space or equivalent)
 - 2. The server's operating system
 - 3. Accounts and passwords
 - 4. Application software
 - 5. The network
 - 6. Staffing
 - 7. Operational practices
- We'll now look a little at the questions from each of those areas. But before we do, what does the thing actually look like?

Server Administration Self Assessment Scorecard - Mozilla Firefox	
<u>File E</u> dit <u>V</u> iew <u>G</u> o <u>B</u> ookmarks <u>T</u> ools <u>H</u> elp	
A Section of the section of	
Server Administration Self Assessment Scorecard (SASAS)	-
This scorecard will guide you through a brief review of your server administration practices, and may help you to identify ways to improve your server's security and availability.	
<i>Note:</i> Completion of this self-assessment is not a replacement for a professional detailed risk analysis and operational review.	
Section I. Server Hardware	
QI-1. Server hardware:	
(5 pts) My server uses hardware that's described by the vendor as "server-class." (0 pts) My server uses hardware that's described by the vendor as "desktop-class."	1
QI-2. Server age:	
(5 pts) My server is new (one year old or less). (3 pts) My server is aging (more than one year old but less than three years old). (0 pts) My server is old (three or more years old).	
QI-3. Server capacity:	

Section 1. Hardware

QI-1. Server Hardware

(5 pts) My server uses hardware that's described by the vendor as "server-class."

(0 pts) My server uses hardware that's described by the vendor as "desktop-class."

Does It Matter If We Buy a "Server?"

- It actually may. Performance issues aside...
- Mass market desktop PCs are built for a very price sensitive market, and desktops may be economically designed with different longevity and duty cycle expectations than servers.
- Desktop PCs do not commonly have things like error correcting memory, high performance disk, hot swap or redundant components, rack mount configurations, or other features normally associated with server class systems.
- Vendors may provide different (hopefully better!) support for servers as opposed to desktops
- The capital cost of typical system's hardware today is also such a small part of the total cost of a system that it may be false economy to try to "cheap out" and make do with just desktop class hardware.

"I think you're 100% wrong about that!"

- I very well may be! The good part about this self-assessment tool is that you can modify it to suit <u>your</u> preferences.
- Hold that all systems are fundamentally alike, and the server/workstation distinctions are largely marketing hype? Go ahead, delete the offending question from your version of the self-assessment tool. Rip it out, "carpe diem"-style. :-)
- Think that it matters, but not as much as I seem to think it may? Great, tweak the scores to better suit your preferences.
- Did I overlook a key topic? Feel free to add it!
- The power of this assessment lies not so much in any single question or any particular point allocation, but in the fact that you're thinking about at least <u>some</u> of the issues, and so, hopefully, will server administrators out in your departments!
- Discussing inclusion or exclusion of an item, or suitable point allocations, can be an excellent consciousness-raising task.

QI-2. Server Age

(5 pts) My server is new (one year old or less).

(3 pts) My server is aging (more than one year old but less than three years old).

(0 pts) My server is old (three or more years old).

Servers Are <u>Not</u> Like Vintage Wines

- As they get older, they do <u>NOT</u> tend to get better.
- Mechanical components (such as hard drives or cooling fans) suffer physical wear, and will eventually fail.
- Electronic components may last virtually forever, but after three or four years, they may have become laughably slow relative to what's become available on the market.
- Vendors will also adjust things like maintenance pricing to help incent you to consider buying more current hardware.
- There's also a psychological aspect to all this: server admins know the institution values the system if it is kept up to date, and they'll probably pay more attention to a current generation system as a result.
- Routinely scheduled system upgrades are easier to budget and cleanly fund (just don't start working on it at year three!)

QI-3. Server Capacity

(5 pts) My server normally runs at less than 50% of capacity.(3 pts) My server normally runs at less than 75% of capacity.(0 pts) My server normally runs at 75% or more of its capacity.

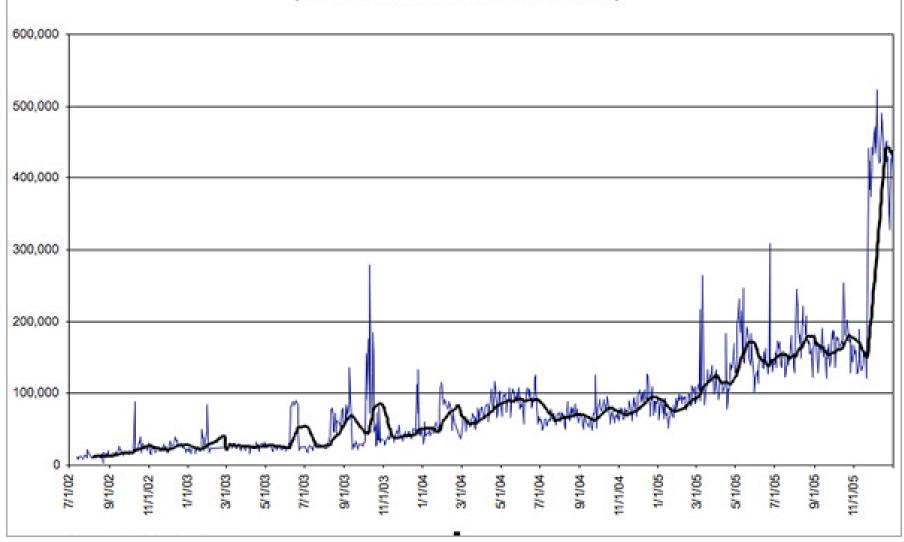
Buy Enough (And Then Buy Some More)

- Capacity planning for servers can be tough, and it is very easy to end up with an undersized system for a variety of reasons including:
 - -- data to use for numerical forecasting may be unavailable
 - -- demand for a new service may be unexpectedly strong
 - -- growth may be non-linear (an accelerating rate of growth)
 - -- unanticipated services/features may get added
 - -- a system may be targeted for attack (such as a spam flood) and additional protective measures may become required
 - new releases of the system's applications may be less sleek than former versions and may need more "go"
 - a scheduled system replacement may end up being delayed and as a result you may need to limp along on the old system for longer than you'd originally planned

-- maintenance or hardware failure may take capacity offline

An Example: Load Associated with An Internet Malware Outbreak (Sober-Y)

REJECTED SMTP CONNECTIONS PER DAY (HEAVY LINE = 28 DAY MOVING AVERAGE)



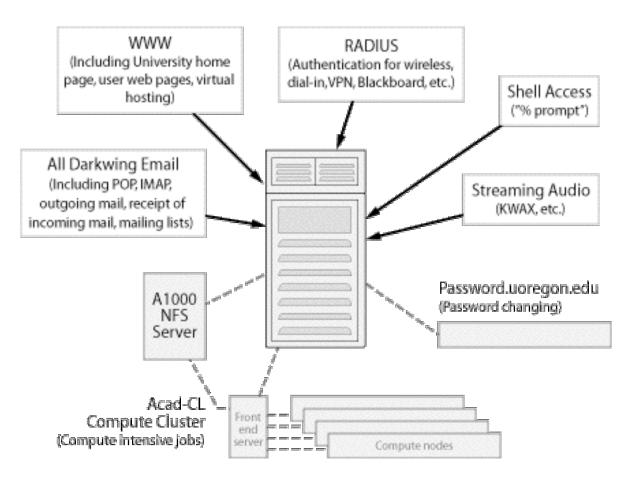
A System Architectural Digression

- Different architectures may be harder or easier to scale.
- Contrast two architectures:
 - monolithic multiprocessor system ("traditional big iron") vs.
 a stack of building block servers behind a load balancer
- Which is more expensive? (think about unit volumes sold, and the need to spread fixed development and support costs)
- Which makes it easier to incrementally scale the service?
- Are (or can) both equally easy to routinely administer?
- Are there any advantages when it comes to hardware maintenance?
- Can the load be subdivided? If so, on what basis? By application? By subsets of users? Can the load be moved around to rebalance it if necessary?

What We Had...

Classic Darkwing (Simplified)

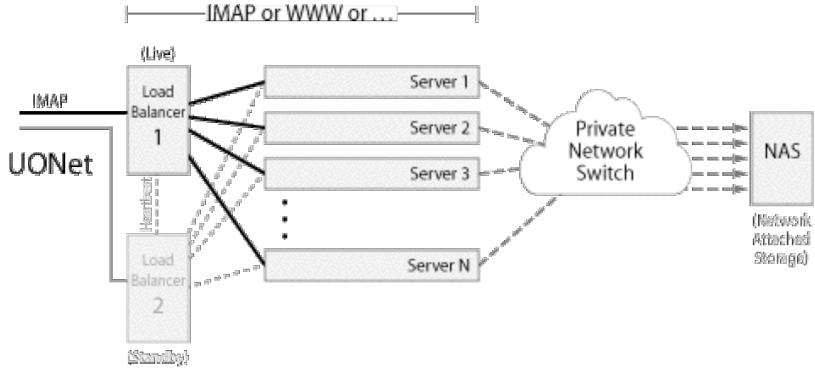
Everything ran on one box: A tangled, interconnected ...



What We Moved To...

Sample Service Node, New Model

Each major service (or group of services) runs on its own set of redundant, load banced hosts



QI-4. Server Redundancy

- (5 pts) We have a fully redundant/load balanced server configuration.
- (4 pts) We have a "hot spare" server that we can use if our primary server fails.
- (3 pts) Our server has some redundancy features, such as dual power supplies or mirrored disks.
- (0 pts) Our server is non-redundant/unprotected.

"If You've Got Two, You've Got One. If You're Got One, You've Got None."*

- In this life, um, "stuff" happens and systems fail, including mission critical systems. To avoid undesirable impacts you need to accept that those failures <u>will</u> occur and <u>plan</u> for those failures to occur and build in redundancy to all the mission critical systems you deploy. Examples:
 - -- jet airplanes can continue to fly even if an engine fails
 - -- parachutists carry a reserve chute in case their primary chute gets tangled or fails to deploy
 - -- networks get built with redundant links, so that if one link goes down traffic can fail-over onto the other link
- Distributed server administrators at your site need to become similarly conscious of the need to "bake in" redundancy.

^{*}Adage reportedly common in the military spec ops community.

"You Can Buy Just As Many 9's As You Want..."

- At the same time that some folks don't worry enough about failure of mission critical systems, sometimes people go nuts when it comes to worrying about them, and demand that <u>everything</u> be built out to "five nines" reliability, e.g., no more than about five minutes down time/year, 365*24*60*0.00001
- While you can generally buy as much reliability as you may want by adding greater and greater levels of protection and redundancy, you need to recognize that it gets harder and harder (==> "increasingly more expensive") to buy each additional "9"'s worth of protection.
- There also may be some systems where it *IS* okay for systems to be occasionally down, and making those sort of systems "bulletproof" may simply be wasting money.

QI-5. Server Maintenance

- (5 pts) We have a vendor hardware and software maintenance contract.
- (4 pts) We have extensive spare parts and are prepared to do server self maintenance as may be needed.
- (0 pts) We handle server maintenance on an informal/ad-hoc basis.

Fixing It When It Breaks

- This is another example of an item where men and women of good faith can honestly disagree, and where the "best" choice may be a function of:
 - -- the cost of replacing the broken unit rather than repairing it
 - -- what your vendor has for local maintenance vs. what you can field in house,
 - -- how much it costs to let the vendor handle it vs. carrying your own spares and doing your own work,
 - -- your experience with a particular vendor's gear (if the gear you've bought virtually never breaks, picking a per-incident plan may make more sense than buying an "insurance policy")
- Regardless of what you decide to do, you may want to make sure you've made <u>some sort of plan</u> for fixing a system that's down before such an incident actually occurs.

QI-6. Server Access Control

(5 pts) Our server is in a secure data center.

- (4 pts) Our server is in a physically secure area other than a data center.
- (0 pts) Our server is in a generally accessible or otherwise insecure area.

Servers Pop Up in the Darnedest Places

- In offices under desks or beside desks (often a "dust bunny"-• rich environment!)
- In closets (dedicated or perhaps shared with custodial staff)
- In hallways ("Hmm, no one ever pushed any of the server's • buttons before...")
- As phone or plant stands, or as makeshift ottomans •
- You name it...
- Everything else being equal, it would be nice to treat each • mission critical server like the important asset it truly is... I'm not saying that every server needs to live in a central machine room (I dislike having "all the eggs in one basket"), but SOME consideration should be given to where and how distributed servers get housed and protected. 27

QI-7. Server Power

- (5 pts) Our server has dual power supplies, and each of those power supplies is fed via diverse power sources. An uninterruptible power supply ("UPS") protects each of those power sources. Loss of one or the other power source will not result in the remaining circuit becoming overloaded.
- (3 pts) Our server is behind an uninterruptible power supply ("UPS").
- (1 pt) Our server is protected by a surge suppressor.
- (0 pts) Our server uses normal wall power w/o any special protection.

QI-8. Server Air Conditioning

- (5 pts) Our server is in an area that has dedicated (24x7) air conditioning and that air conditioning keeps our server sufficiently cool.
- (3 pts) Our server is in an area that has normal building air conditioning. That air conditioning usually keeps our server cool, but that building air conditioning is subject to interruption on weekends or outside of normal business hours (but that doesn't appear to cause a problem).
- (1 pts) Our server is in an area that is not air conditioned, but overheating does not appear to be an issue.
- (0 pts) Our server is in an area that is not air conditioned, and sometimes it may get too hot.

Emergency Power and Air Conditioning

- If you're feeling chilly, some of today's systems make fine functional replacements for traditional space heaters. We've also gotten very good at packing 'em into crowded data centers via 1U form factor systems, blade servers, etc.
 Because we've also gotten good at building survivable power supplies, even if we lose normal power, those systems <u>will</u> often keep soldiering along just fine.
- But what about your air conditioning? Will the A/C keep running if you lose normal power? In at least some cases, air conditioning will go off when normal power goes off, and come back on when air conditioning comes back on. What's the rate of temperature increase you see if that happens? Does it matter if you have three days worth of emergency generation capacity for your <u>servers</u> if your server <u>space</u> is heating up at a rate of four or five degrees F/hour?

QI-9. Fire Detection and Suppression

(5 pts) Our server is protected with either an inert gas ("Halon(tm)") fire suppression system or a dry-pipe preaction fire suppression system.

- (3 pts) Our server is protected by a normal water-based fire suppression system.
- (1 pt) Our server is protected by a normal fire detector.
- (0 pts) Our server doesn't have a fire detection/suppression system.

Section 2. Operating System

QII-1. Operating System Version

(5 pts) Our server is running the current (vendorrecommended) version of its operating system.

- (3 pts) Our server is running an older (but still supported) version of its operating system.
- (0 pts) Our server is running a no-longer-supported version of its operating system.

When Does MS Windows 2000 Server Go "End of Life?"

😂 Microsoft Support Lifecycle - Mozilla Firefox	
<u>F</u> ile <u>E</u> dit <u>V</u> iew <u>Go</u> <u>B</u> ookmarks <u>T</u> ools <u>H</u> elp	
🗘 🔹 🎝 😸 🛞 🖓 🔟 http://support.microsoft.com/lifecycle/?LN=en-us&x=1	2&y=10&p1=7274 💽 💽 Go 🔀
United States	Microsoft.com Home
Microsoft [®]	<u>Search Microsoft.com</u>
Help and Support	

Help and Support Home | Select a Product | Search Knowledge Base

Microsoft Support Lifecycle

Product support lifecycle information by product

Microsoft provides support on the current service pack, and in some cases the immediately preceding service pack, for the products listed below. See the <u>Lifecycle Support Policy FAQ</u> for details on the Service Pack Support Policy. As a convenience to you, Microsoft is providing a list of the currently supported service packs: <u>Supported Service Pack List</u>. If you have any questions regarding support for a product, please contact your Microsoft representative. If you need technical support, visit the <u>Contact Microsoft</u> Web site.

The information on this page is subject to the <u>Microsoft Policy Disclaimer and Change Notice</u>. Return to this site periodically to review any such changes.

Products Released	<u>General</u> <u>Availability</u> <u>Date</u>	<u>Mainstream</u> <u>Support</u> <u>Retired</u>	Extended Support Retired	<u>Service Pack</u> <u>Retired</u>	<u>Notes</u>
Windows 2000 Server	3/31/2000	6/30/2005	7/13/2010	Not Applicable	

"What's The Difference Between Mainstream and Extended Support?"

🐸 Microsoft Support Lifecycle Policy FAQ - Mozilla Firefox	
<u>F</u> ile <u>E</u> dit <u>V</u> iew <u>G</u> o <u>B</u> ookmarks <u>T</u> ools <u>H</u> elp	
• • • • • • • • • • • • •	🔹 💿 Go 💽

3. What is the difference between mainstream support, extended support, and online self-help su

Support provided	Mainstream support phase	Extended support phase
Paid support (per-incident, per hour, and others)	Х	Х
Security update support	Х	Х
Non-security hotfix support	х	Requires extended hotfix agreement, purchased within 90 days of mainstream support ending.
No-charge incident support	Х	
Warranty claims	Х	
Design changes and feature requests	Х	
Product-specific information that is available by using the online Microsoft Knowledge Base	х	х
Product-specific information that is available by using the Support site at Microsoft Help and Support to find answers to technical questions	х	Х

Note A hotfix is a modification to the commercially available Microsoft product software code to address sp

QII-2. Operating System Patch Status

- (5 pts) We've applied all vendor-recommended critical patches to our server.
- (2 pts) We've applied some critical patches, but have not applied other critical patches (for whatever reason).
- (0 pts) We've not patched our operating system with vendor recommended critical patches.

QII-3. Automatic Patch Application

- (5 pts) Our server has been set to automatically apply critical new patches.
- (3 pts) We are automatically notified of new patches, which we then manually review and apply as appropriate.
- (0 pts) We apply patches on an ad hoc/informal basis.

Patching Praxis

- Patching is another "religious" area where you'll often see distinct camps about how to handle the patching of production systems. If you've somehow managed to miss patching related fisticuffs, you may want to read the classic article, "Patch and Pray" from <u>CSO Magazine</u>, August 2003, http://www.csoonline.com/read/080103/patch.html
- I'd also encourage you to review the status of the products you currently use at http://secunia.com/product/ since in many cases, even with all patches and work arounds applied, serious vulnerabilities may remain.
- All that said, I've gotten to the point where I believe, on balance, that you're better off routinely "taking your chances" with automatic patching rather than risking rapid miscreant exploitation of unpatched vulnerabilities which have been publicized worldwide.

QII-4. Unneeded Network Services

(5 pts) All unneeded network-based services have been disabled on our server.

(0 pts) Our server offers the full set of default network services.

QII-5. File Sharing

(5 pts) File sharing has been disabled.

(3 pts) File sharing is enabled, but has been carefully limited.

(0 pts) File sharing is enabled.

Manual System Hardening Guidance



their systems.

Semi-Automating Server Hardening

http://www.bastille-linux.org/

🔰 Bastille Linux



The Bastille Hardening program: increased security for your OS

Resources | News & Updates | Credits & Sponsors | Quotes | Screenshot | Undoing Bastille | Assessment | Jay's Linux Security Page

Running Bastille On | Development Resources

What Is It?

The Bastille Hardening program "locks down" an operating system, proactively configuring the system for increased security and decreasing its susceptibility to compromise. Bastille can also assess a system's current state of hardening, granularly reporting on each of the security settings with which it works.

Bastille currently supports the Red Hat (Fedora Core, Enterprise, and Numbered/Classic), SUSE, Debian, Gentoo, and Mandrake distributions, along with HP-UX. **Full Mac OS X is in beta, ready for download today.** Bastille's focuses on letting the system's user/administrator choose exactly how to harden the operating system. In its default hardening mode, it interactively asks the user questions, explains the topics of those questions, and builds a policy based on the user's answers. It then applies the policy to the system. In its assessment mode, it builds a report intended to teach the user about available security settings as well as inform the user as to which settings have been tightened.

QII-6. Firewalls

(5 pts) A host-based software firewall has been installed on our server, and is configured to deny all traffic except that which has been explicitly permitted.

(3 pts) A host-based software firewall has been installed on our server, but is configured to permit all traffic except that which has been explicitly forbidden.

(0 pts) Our server does not use a host-based software firewall.

"But If They Use a Firewall <u>*/*</u> Won't Be Able To Scan Them Anymore!"

- True, and this is actually a good thing. The goal is (or should be) to harden your institutional servers to ALL external parties, including security scanners running elsewhere at the local site.
- If you're currently relying on active probes (such as from Nessus) as a material part of your security toolbox, you should be planning to augment or replace that active approach with passive network monitoring with intrusion detection products such as Snort or Bro.
- Another alternative is to deploy an agent-based system which can periodically "phone home" with information about the status of critical systems (note, however, that some agentbased systems can be quite expensive).

QII-7. Checksumming Of Critical System Files

(5 pts) A checksumming program (to detect unauthorized changes to critical files) is being run on our server.(0 pts) A checking program is not being run on our server.

QII-8. Antivirus/Antispyware

- (5 pts) Our server doesn't run Windows, or if we do run Windows, we have current antivirus/antispyware software installed.
- (0 pts) Our server runs Windows, but we do not run antivirus/antispyware software on it, or that software does not have up to date definitions.

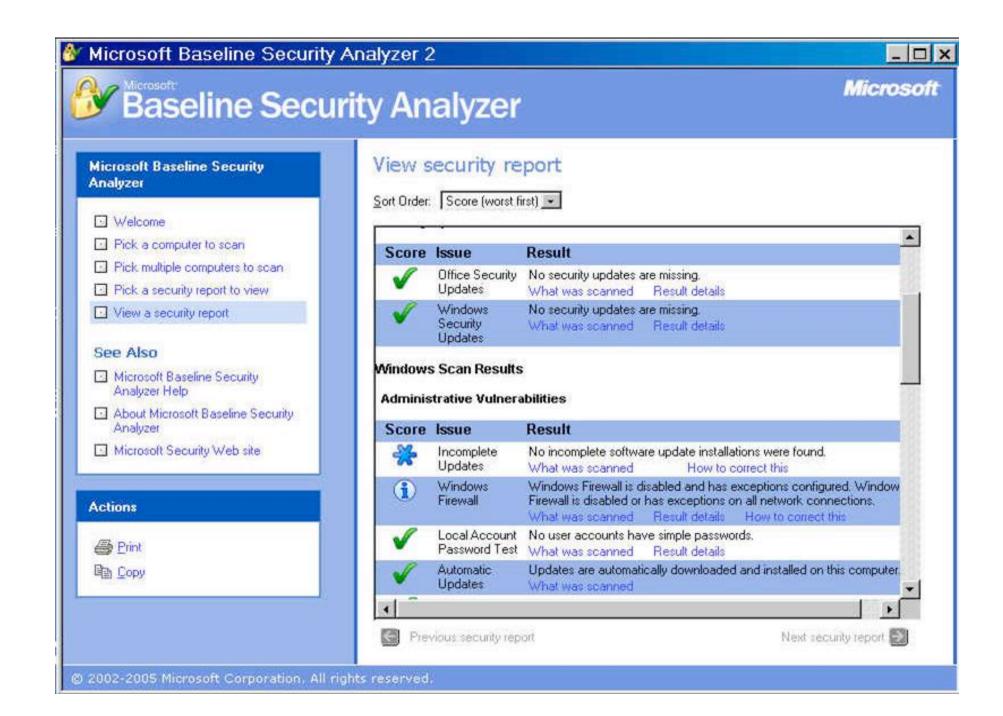
AV Software: Protecting The Server or Filtering Content Sent <u>Via</u> the Server?

- Sometimes users wonder if antivirus and antispyware requirements for servers reflects an attempt to protect the server itself from infection, or if the goal is to block dangerous content from successfully transiting the server (and thus potentially infecting other servers and workstations).
- Ideally, antivirus and antispyware protection would be applied to **both** the server itself and to screening the content that is transiting the server, but for the purpose of **this** question, we're primarily worried about the server itself.
- Not explicitly mentioned, but obviously important: servers should not be used both as servers and as general purpose workstations (e.g., don't surf the web from a server or read email from a server, do those sort of tasks on a separate non-mission critical system).

QII-9. MS Baseline Security Analyzer v2

(5 pts) Our server doesn't run Windows, or if we do run Windows, when we run Microsoft Baseline Security Analyzer v2, no issues are flagged.

(0 pts) We run Windows and Microsoft Baseline Security Analyzer v2 flags one or more security issues.



Section 3. Accounts/Passwords

QIII-1. Account Creation and Deletion

- (5 pts) Only authorized users have an account on our server; no-longer-needed accounts get removed.
- (0 pts) Account creation and deletion is handled on an ad-hoc basis, accounts are shared, or we may potentially have unauthorized users or no-longer needed accounts on our server.

QIII-2. Passwords

- (5 pts) Accounts use a hardware token, biometric access method, or other two factor authentication technology.
- (3 pts) Accounts have a strong and periodically changed password.
- (1 pt) Accounts have passwords, but those passwords are of unknown strength (or passwords may not be required to be periodically changed).
- (0 pts) Accounts do not require passwords, or weak passwords are known to be in use (username=password, or all account passwords are set to the same initial value with no change required, for example)

Traditional Passwords: The Walking Dead

- I kid you not, the days when traditional 6-8 character passwords will be "good enough" are over. If you're not making plans for migrating to two factor methods (something you know (like a password), plus something you have (like a hardware token) or something you are (like a biometric factor) I'd strongly encourage you to begin doing so. Don't believe <u>me</u>? How about Bill Gates? Try Googling for Bill Gates passwords dead.
- While we're talking about traditional passwords, you should also be thinking about how you'll accommodate requests for "single sign on" authentication received from departmental systems... distributed authentication via LDAP, Radius or other mechanisms can pose some unique vulnerabilities (a trivial example: think about dictionary or brute force attacks)

QIII-3. Password Encryption

(5 pts) Passwords transmitted over the network are encrypted with ssh, ssl, or similar strong encryption.

(0 pts) Passwords are transmitted in plain text (e.g., telnet is used instead of ssh, ftp is used instead of scp or sftp, or passwords are transmitted via unencrypted web pages)

Sniffing Shared Unencrypted Passwords

- If you need a concrete risk to care about use of unencrypted passwords for access to departmental systems, recognize that in many cases users will reuse a single password across multiple systems (e.g., they may use the same password on your central ERP system, which hopefully is fully encrypted, and on a distributed system which regrettably may NOT be fully encrypted). If a bad guy/gal can sniff that password en route to a departmental server, they have an excellent statistical chance of being able to access your ERP system.
- Oh yes, just for the record, you DO have sniffing exposure even if you're running on a fully switched network architecture. See, e.g.: http://monkey.org/~dugsong/dsniff/
- Also be on the lookout for local WEP "encrypted" wireless networks. Why? See "Cracking WEP in 10 Minutes," http://weblog.infoworld.com/udell/2005/06/08.html 55

QIII-4. Acceptable Use Policy

- (5 pts) All users have been informed of applicable acceptable use policies and have affirmatively indicated that they've been informed of those policies and consent to them (for example by signing a copy of that AUP).
- (3 pts) All users are informed of applicable acceptable use policies by email or by posting of the acceptable use policies on the unit's web site, but affirmative consent is not required as a condition of access.
- (0 pts) There is no acceptable use policy, users are not informed of it, or the AUP is not enforced.

Section 4. Application Software

QIV-1. Software Licensing

- (5 pts) All applications are properly licensed, current, and vendor supported, or open source.
- (1 pts) Some applications are dated, or are no longer vendor supported nor open source.
- (0 pts) One or more application may be improperly licensed (please note that it is University policy that all University units will respect copyrights and properly license all software used).

Why **Doesn't Software Get Inventoried?**

- At least at some schools, virtually every tangible object gets an inventory sticker (including \$300 bargain basement PCs).
- At those same colleges and universities, however, no systematic attempt is made to track software assets, even in cases where those programs cost thousands of dollars. Is it because you can't put an inventory sticker on actual software (as opposed to software media or accompanying manuals)? This strikes me as odd (and probably bad) thing.
- I would urge you to be a rebel. If you use commercial software, I encourage you to personally document that software on each server whether accounting requires it or not. This doesn't need to be anything elaborate – for each server simply list all the software running on it, including the version, product serial number and associated purchase order number and date – you may find that can really help!

QIV-2. Locally Developed Applications

(5 pts) Locally developed applications aren't in use.

- (4 pts) Locally developed applications are in use, but have been written in a common, easily maintained programming language.
- (2 pts) Locally developed applications are in use and are written in an uncommon programming language, however we have multiple local programmers who are fluent in that language.
- (1 pt) Locally developed applications are in use and are written in an uncommon programming supported by only a single local programmer.
- (0 pts) We have one or more locally developed application written in an uncommon language that none of our current staff know or use.

A Big Potential Nightmare

- Choice of program development language may seem like an odd thing to fixate on, but trust me, it can be a huge issue if a key system is developed in an obscure programming language by a clever person who subsequently leaves or is fired.
- It can be really hard to tell a brilliant programmer the hard truth that no matter how powerful and wonderful Ada or Pascal was as a programming language, unfortunately that's not where the current momentum lies, and regrettably there's no one else on the staff who's using those languages so could you please use Java or C or C++ or Perl or PHP or Visual Basic or C# instead?
- Be sure to watch the libraries that end up getting used, too.

QIV-3. Change Control

- (5 pts) Application modifications are made via a formal change control process, and are documented, tested, reviewed and approved before deployment on production systems occurs.
- (3 pts) Application modifications are made via an informal change process, or documentation, testing, review and approval is handled on an ad hoc basis.
- (0 pts) Application modifications are made to live production systems on an "as needed" basis without formal change control.

Separation of Duties

- Change control, to a first approximation, is all about separation of duties and insuring that a second set of eyes gets a chance to look at code before it is rolled into production use.
- The fundamental problem when it comes to locally developed code running on distributed systems? There may be only one programmer writing code for an entire project, and his/her manager may not be technically capable of assessing the impact of proposed changes to production code.
- A key question to ask: "So how <u>do</u> you handle tracking revisions to the code you're developing?" If you don't hear about some sort of formal revision control system (RCS or SCCS or Subversion or whatever), a flag should go up...
- For that matter, when was the last time you saved an archival copy of the source code for a key locally developed app?³

Section 5. The Network

QV-1. IP Address

(5 pts) Our server has a static IP address, and contact information for that IP address is up-to-date.

(2 pts) Our server has a static IP address, but the contact information for that IP address may be out of date.

(0 pts) Our server is using a dynamic IP address.

Surely <u>All</u> Servers Use Static IPs?

- You'd think so, wouldn't you? And that contact information for the IP address would be up to date? In reality, however, that may not be the case, particularly if you have serve fairly persistent DHCP lease policies or make it hard to get a static.
- Some may wonder why I deduct so many points for missing or stale IP contact information, but the reality is that IP-toresponsible person mappings may be the ONLY clean way you have for figuring out who's responsible for a server that may be misbehaving or which may have been compromised.*

* Before anyone mentions it, yes, I <u>am</u> also aware of what some may call the "phone company method" of identifying improperly documented connections, e.g., intentionally breaking that connection and then waiting to see who calls to complain that their service is down, but we <u>can</u> do better!⁶

QV-2. Hardware Firewall

- (5 pts) Our server is sheltered from general Internet access attempts by a hardware firewall which has been configured to deny all traffic except that which is specifically allowed.
- (3 pts) Our server is sheltered from general Internet access attempts by a hardware firewall which has been configured to permit all traffic except that which is specifically forbidden.

(0 pts) Our server is not behind a hardware firewall.

Cool! We Have an Institutional Hardware Firewall in Place!

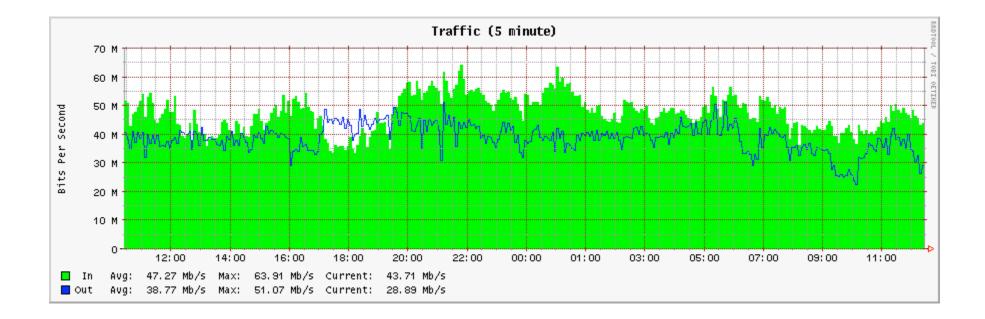
- Unfortunately, a single site-wide institutional firewall isn't really what's needed. Why? Simple: there are too many untrustworthy systems inside the trusted zone created by that firewall, and too many users with too many disparate requirements to reconcile.
- Site wide firewalls will also often interfere with things like H.323 video conferencing and people trying to go fast on high performance networks such as Internet2's Abilene network.
- Push hardware firewalls closer toward mission critical servers or strategic subnets, instead, and recognize that different policies may be required for different subnets.

QV-3. Network Capacity

(5 pts) We formally track the network traffic volume associated with our server (e.g., via MRTG, RRDtool, Cricket or similar products), and have sufficient network capacity.

- (3 pts) We have seen no indication that our server is running into network traffic capacity issues.
- (0 pts) We know or believe that we may have insufficient network capacity or network performance problems.

Sample Monitoring Network Capacity With MRTG/RRdtool



Section 6. Staffing

QVI-1. System Administration

- (5 pts) Our server is supported by a team of system administrators.
- (3 pts) We have at least one backup system administrator to provide support when our primary system administrator is unavailable.
- (1 pt) We have only a single system administrator.
- (0 pts) No one is formally responsible for our server.

QVI-2. Coverage

(5 pts) Our server is well supported 24x7.

- (3 pts) Our server is supported less than 24x7, but that coverage is sufficient because of user expectations, or because our server is not mission critical, or there are other mitigating or extenuating factors.
- (1 pts) Our server is only supported during normal business hours; there is concern that that is not sufficient given expectations for our server.
- (0 pts) Our server does not have formal coverage/staffing, or what coverage/staffing we do have is known to be insufficient/causing problems.

Salaries: The REAL Driving Cost of Running Distributed Servers

- Remember when we first started looking at the questions associated with this self-assessment, and I indicated that the hardware costs associated with a server really aren't very material? Some of you probably wondered, "So where <u>are</u> the major costs associated with running a server?"
- The answer, as always, is <u>personnel costs</u>, particularly if a mission critical system requires true round the clock coverage (rather than just "oh, I guess that I've got the on-call pager tonight" sort of casual coverage). What would a team of four good full time system administrators cost on an ongoing basis, eh?
- In some cases you may not even be ABLE to find/hire the sort of folks you need/want at any reasonable price.

Section 7. Operational Practice

QVII-1. Server Documentation

- (5 pts) Our server's configuration and routine operation is fully documented.
- (2 pts) Our server's configuration and routine operation is partially documented, or that document may be potentially out of date.
- (0 pts) Our server's configuration and routine operation is not documented or that documentation is known to be inaccurate or incorrect.

Documentation: A Vanishing Art

- While most programmers and system administrators know that documentation SHOULD be an integral part of developing code or operational processes for a mission critical system, all too often it is neglected or ignored. This fact is often discovered when a programmer or system administrator leaves for another job, at which point it may well be virtually impossible to retroactively recreate the missing documentation.
- Comprehensive and current documentation is PARTICULARLY important if a distributed project involves a small number of individuals, or if you've failed to avoid use of obscure programming languages, or if change control is spotty at best.
- Documentation needs to be backed up just like source code!

QVII-2. Server Monitoring

- (5 pts) The server is monitored, and an administrator is paged or otherwise automatically notified if the server hangs or crashes.
- (1 pt) The system administrator receives and responds to user-generated complaints regarding system availability.
- (0 pts) The system isn't monitored; complaints relating to system availability may be handled on a best efforts basis (or not at all).

Monitoring Performance with SmokePing



79

HOW Do Notifications Get Done?

- Email? What if email is the service that's down?
- Text messages sent to pagers or cell phones? What if phone service is down or busy or the administrator is in a region w/o coverage?
- Via changes to web pages? Do you know when someone will be looking at those web pages? Maybe you can arrange an RSS feed?
- Are there perhaps opportunities for providing integrated monitoring and notification of central systems <u>and</u> distributed systems? It is really hard to beat having a genuine human being monitoring things and insuring that the right responsible person gets notified.

QVII-3. Maintenance Windows

- (5 pts) There's a routinely scheduled maintenance window for our server.
- (2 pts) We can usually arrange a maintenance window on an ad-hoc basis when one is needed.
- (0 pts) For whatever reason, it is difficult or impossible to find time to schedule routine maintenance.

You MUST Have Maintenance Windows

- Some system administrators may not be particularly comfortable when it comes to asserting themselves, and as a result you may run into systems where users have effectively "bullied" the system administrator into never taking the server down for required work. This is very bad, and is kin to taking off in a plane that's overdue for maintenance. System administrators MUST have periods when they can take systems down for maintenance that simply can't be done live.
- If 24x7 availability is key, with some architectures it MAY be possible to take only <u>part</u> of the system down for work at any given time, but that should be viewed as a fortuitous/miraculous exception rather than an assured rule.
- Sometime we should also talk about <u>when</u> maintenance windows should be scheduled (and no, 2AM Saturday morning may NOT always be your best choice).

QVII-4. User Communication

(5 pts) We have a mailing list or other ready "push" communication channel which we can use to communicate with users about our server.

(3 pts) We have a web page or other "pull" communication channel that interested users can visit for information about our server.

(0 pts) We lack an easy way to communicate with users of our server.

QVII-5. Data On The Server

- (5 pts) Data on our server has been reviewed for sensitivity, and is appropriately controlled.
- (1 pt) We've taken some steps to address issues related to sensitive data on our server, but work remains to be done.
- (0 pts) Data on our server has not been reviewed for sensitivity, or sensitive data may not be appropriately controlled.

QVII-6. Backups

- (5 pts) The contents of our server are routinely backed up, and backups are stored at a secure off site location.
- (3 pts) The contents of our server are routinely backed up, however we do not store copies of those backups off site.
- (1 pt) We occasionally take backups of our server.

(0 pts) Our server is not backed up.

QVII-7. Disaster Recovery

- (5 pts) We have a disaster recovery plan that covers our server, we've tested that plan, and we've verified that it will prevent material long term disruption to normal business operations.
- (3 pts) We have a partial disaster mitigation plan that covers our server, but it is untested or it is likely that if we need to use it, disruptions to normal business operations will occur for a material period of time.
- (0 pts) We do not have a disaster recovery or disaster mitigation plan covering our server.

How Long Can You Afford to Be Down?

- A lot of disaster recovery scenarios seem to be "paced" for quieter, gentler times. If you were forced to actually use the methods proposed in many disaster recovery plans, how long would you be "off the air?" Multiple days? Heck, how long would it take you to just retrieve and restore multiple terabytes from tape? How long would it take for cached DNS to time out and become repointable? (http://dnsreport.com/)
- Let me be bold here, and assert that increasingly business <u>continuity</u> requirements (business <u>survival</u> requirements?) will only allow unscheduled downtime on the order of a small number of hours (let's say half a work day). Providing continuity of operations and bounding downtime at less that four hours will typically require <u>full multi site redundancy with hot systems and ongoing data synchronization</u>.

Interpreting Your Total Score

Total score (sum all items): _____ out of 185 possible total points.

165-185 points (~89% or better): Superior
145-164 points (~78% or better): Good
125-144 points (~67% or better): Weak
124 points or below: Poor

Note: depending on where you lost up to 15 points, you could have a "superior" score and still have some pretty significant issues.

Feedback From Users

- While we did <u>not</u> explicitly ask users to tell us how they did on the Server Self Assessment Scorecard, a few did so voluntarily and informally. Some paraphrased comments:
 - -- We know from talking with other distributed server administrators that we're doing better than most folks, but we still didn't score very highly.
 - -- The scorecard was helpful for us in that it gave us a lot to think about, and a lot of areas to review. We may be reevaluating whether we should really be running this sort of mission critical server in our department.
 - -- We'll be using the scorecard as a foundation for budget discussions with our department head so we can begin to work to address some of our current potential issues.

Next Steps

- It is important to realize that the server self-assessment • scorecard is **not** meant to be a replacement for a formal and thorough professional risk analysis.
- This brief scorecard is also not meant to be a tutorial on system administration procedures. If a system administrator in a department would like to know a lot more about how to professionally practice their trade, a nice book to consult is The Practice of System and Network Administration, by Thomas Limoncelli & Christine Hogan, ISBN 0-201-70271-1.
- You should also recognize that server administration is not a static once-and-done process – conditions are continually changing, and you may need to re-evaluate the condition of each server on a periodic basis, say every quarter or half year. Is documentation still current? Is capacity still adequate? Have their been staffing changes? 90

Thanks for the Chance to Talk Today!

- If you're interested, you can obtain a complete copy of the server administration self-assessment scorecard in HTML format at http://cc.uoregon.edu/serveradmin.htm
- Are there any questions?