



Data Center Segmentation

Peter J. Welcher

Introduction

I am glad to be back to writing and posting articles. Last month got a little too hectic and I ran out of time for article-writing.

This month's article takes a look at Data Center segmentation. This is something I am seeing more and more of. Not only that, but I expect it is something we will all be seeing more of. So I thought it would be a good idea to write about segmentation.

Some of the forms of "segmentation" I have seen so far:

- Stricter separation of Production and Test servers (Performance / Test, or other names)
- Better isolation of financial servers from other servers (and users)
- Substitute "credit card" or "patient data" for "financial" in the previous item
- Protect some or all servers from being hacked by unauthorized insiders

Some Words About Data Center Design

Data Center design used to be relatively easy. It was something we all viewed as a variant of campus switched network design. If you were small, you attached all the servers to a couple of big switches, or smaller switches at the top and bottom of each rack. Big sites had more servers and so used more layers of switches, and also built up more expertise in dealing with large numbers of servers. That used to be all there was to it.

We worried more about firewalling and how to connect to ISP's, things like BGP and transit domains and AS number. Big organizations maybe debated the best way (BGP, DNS, other) to failover to another site. Then all those niche boxes started showing up. Content switches or load balancers, SSL offload, application gateways, etc. And life got more complicated. The number of servers also started increasing rather dramatically. And life got more complicated.

If that does not apply to you, consider that somebody else's past history may be what is coming up for you.

What I have noticed is that Data Center design has been becoming a specialized area. It is no longer "campus design, web DMZ, done."

Cisco now has a bunch of SRND design guides out on the topic. And people are seeing some strong advantages to deploying devices like Content Switches, CSS's, CSM's, ACE modules, F5 BigIP load balancers, Storage Area Networks (SAN's), and so on. Among them, High Availability.

The reason I mention this is that I think there is something we all need to notice: Data Center Best Practices are rapidly changing. For a brief period, designs were including separate NIC cards and VLANs for backup networks. Now, in most cases, backup is taken directly off the SAN. Or off the converged front end network -- no need for a dedicated backup network.

The whole idea of where to use OSI Layer 2 versus Layer 3 in the Data Center is also changing. Up until recently, I had thought the trend was to push L3 down to the access layer. That is, dual-homed servers would connect to two access layer switches with a trunk between them. From there up to distribution and core would be L3 routed point-to-point links. With perhaps some legacy hold-outs, parts of the data center that were designed when VLANs ruled.

The reason for all this L3 routing is to avoid the possibility of Spanning Tree problems that take out the entire data center. This is why sites that have had VLANs in the Data Center core are usually going to at least a L3 routed data center core. With L2 to the core / distribution layer, you have a bad Spanning Tree day and whammo, a solid chunk of or all of your data center stops working, or talking to the rest of the organization. Nasty. And L2 problems have a tendency to do that to you.

However, I also see that VMWare VMotion is catching on in a big way. It currently only works across the same VLAN. So all of a sudden some designs have a one or several VLANs spanning the row of servers (not too bad) or the entire data center. The idea is to be able to quickly re-use servers anywhere and move applications around with VMotion as the need arises. That contributes to High Availability of services, and ease of adding server capacity. However, I do not want to put the entire Data Center at risk with large VLANs. That does not seem to balance the risks properly.

Get Those Servers Organized

What seems to be feeding this in part is server organization. New rows of racks with 6509 "bookends", cabling, etc. are costly. So there is an incentive to fill each row rather tightly. The consequence: you cannot guarantee there will be a spare server or room for one, in any given row of racks. So every server everywhere in the data center has to be able to connect to any VLAN any other server is in. Even if there is space, the local VMotion expert or consultant may be telling management that the VMotion Best Practice is one big VLAN everywhere in the Data Center.

My current recommendation: spend the money to buy more rack space, it is probably far cheaper than the downtime when even a single Spanning Tree moment occurs.

There has also been a lot of "rapid server consolidation" lately. As in, "let's move those servers to the new data center, as soon as we can". Or "let's pack those applications onto the servers using VMWare."

Those are both good objectives in principle. If there is not much planning, you end up with servers re-located without much rhyme or reason. Very different servers and applications may end up in the same VLAN or sharing the same access switch(es). Or very different applications can end up on the same server. To be diplomatic, why don't we call that "ad hoc server or application organization" (or "placement").

One problem with "ad hoc server or application placement" is that it may be hard to get a maintenance window. There can be so many different stake-holders on servers on the access switch pair, that change control or maintenance window becomes a nightmare. Or, with VMWare, the problem would be that the server cannot ever be rebooted, for the same reason -- too many stakeholders (application owners).

The network side of this is shared server VLANs. We used to design so that each access switch pair had a couple of VLANs on it. Servers were addressed into the VLAN(s) appropriate for their location. There was no separation based on functionality. The result is that we ended up with firewall rules that do everything by server IP address, because the subnets had no logical meaning. That was fine at the time, just part of life. Over time, it has lead to some long and ugly ACLs. Not to mention, hard to maintain ACLs.

A new problem has started showing up. We are now starting to see significant re-work, because new security Best Practices require isolation of servers and information. Credit card PCI standards can be interpreted that the servers holding crucial credit card data need to be firewalled at least. Governance and audit standards are starting to be interpreted to require firewalling of servers with corporate financial records.

There is also increased awareness that most severe security incidents involve insiders. So there is a growing emphasis on keeping traffic from unauthorized users away from servers. It is not enough that an insider cannot log into the server, we now also have to prevent them from attempting exploits or denial of service against the server.

It goes without saying, that management wants it done yesterday. And nobody wants to physically move server chassis or re-address servers if they can help it.

In the rest of this article, we will look at some of what can be done about this. If you are in the midst of segmentation, I provide some ideas for "quick fixes". If you are not in the midst of segmentation yet, I hope this whole article will provide food for thought. In particular, get those servers organized when they are deployed, as it is a lot simpler to do it earlier rather than later!

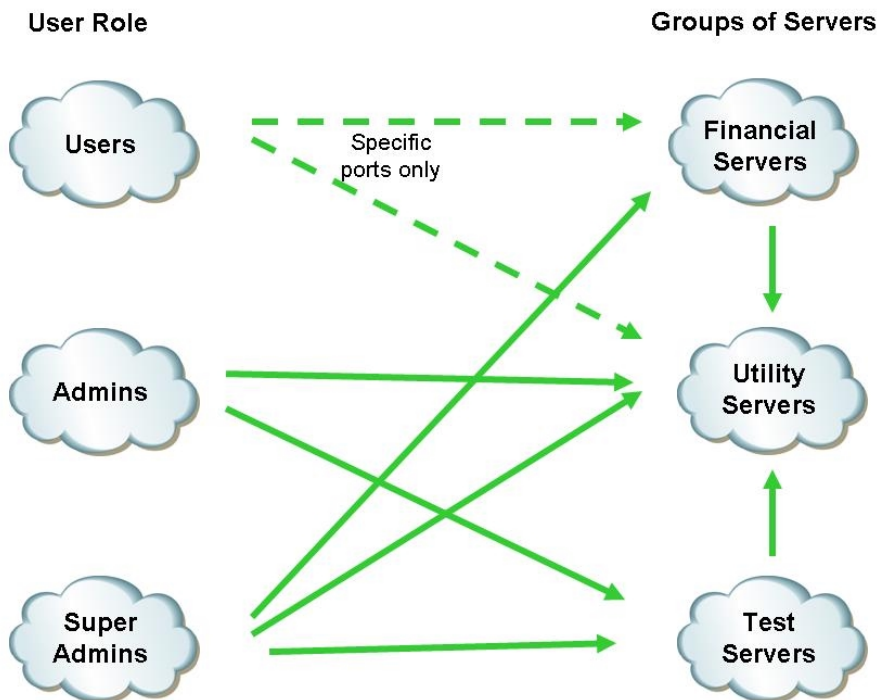
Data Center Segmentation

Data Center Segmentation is designing the data center to control access to key servers.

Generally we want to control two aspects of server access:

- Who (which users or user roles) can send packets to a server?
- Which other servers a given server can communicate with?

Here is a simplified diagram suggesting this:



Technology solutions that might help implement this include:

- Role-based access controls, to control users' access to servers
- Server to server controls

We take a look at how to implement these in the next two sections. As part of this we discuss how to put some controls in place, buying time for a more robust and scalable implementation.

Role-Based Access Controls

User or role-based access controls are intended to control what traffic a user can send to a server. They represent going a step beyond using user logins to control server access. The use of login authentication, administrative groups, and server or domain rights limits what someone can do when logged into a server. Controls based on user logins exercise no control whatsoever over what the user can do when **not** logged into a server. So it is up to the network security tools to control what the user can do, even when not logged into a server.

Ideally all servers were tightly locked down (hardened), with a solid patching program. Realistically, hardening rules change over time, and do not always get applied consistently in the first place. Patching or hardening production servers is always problematic, since quick patch application is very much at odds with a controlled test and release cycle. Production servers almost always have a 3-6 month lag in being patched.

The idea of user segmentation is to leverage a network-based login, so that only authorized administrators can send packets (or selected types of traffic) to servers. The intent is to narrow any security exposure.

There are a couple of ways to tackle this.

One solution I have heard of is from Apani. It consists of IPsec from client machines to servers. What concerns me about it:

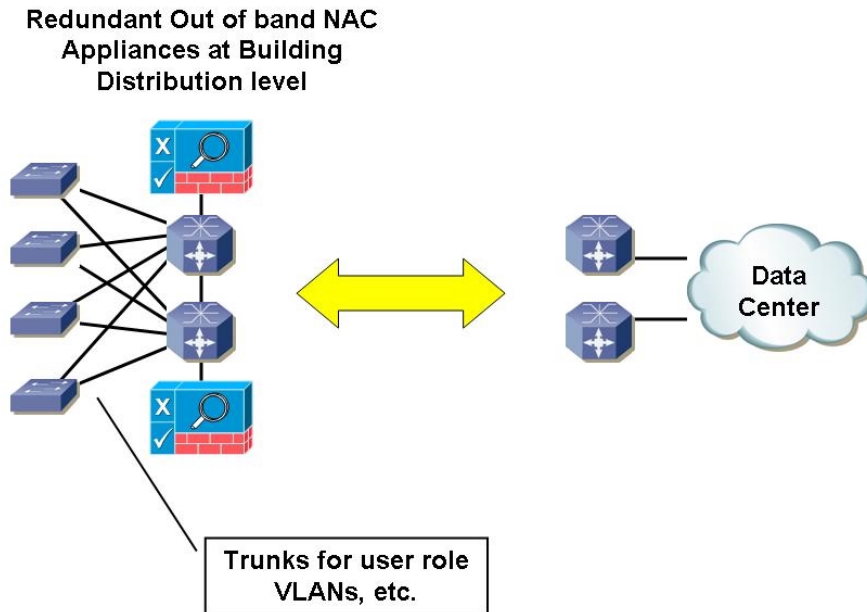
- End-to-end IPsec traffic is opaque for IDS/IPS, QoS classification, and troubleshooting purposes
- It defeats the purpose of owning a Network General Infinistream, if you have one
- IPsec traffic probably imposes a high CPU burden on servers for any high-volume transfers of information, which could be a hidden cost of this approach: you might need more servers to do the same tasks
- The VPN client is apparently incompatible with the Cisco VPN client. This breaks some current "bird flu" ("site access denial") Disaster Recovery (DR) or Continuity of Operations (CoOP) plans!

Another approach: use Cisco NAC, either NAC Appliance or NAC Framework.

Cisco NAC Appliance ("CNACA" for short) allows for in-band or out-of-band use. (We don't have time to go over the difference right now.) If you are using CNACA in-band, you can apply IP access lists to all traffic passing through the appliance, based on user role. Thus "deny ip any any" for most groups, and permit statements for the few groups of administrators. By the way, you want people to belong to single groups for simplicity, so WindowsAdmin and UnixAdmin groups probably means you need a Windows+UnixAdmin group for those who can access both.

CNACA in out-of-band mode or Cisco NAC Framework ("CNACF" for short) allow you to dynamically assign users to VLANs, based on role. (I like to call this the "VLAN shazaam" -- let's not get into that right now.) That lets the user role be reflected in their source address. You can then apply access lists (ACLs) in a conveniently placed switch or firewall. Router ACL's (RACL's) in a L3 switch or Cisco FWSM (Firewall Services Module) between the user and the servers is a convenient place.

Here is a diagram suggesting typical deployment of NAC Appliance in the distribution layer switches "near the user edge" (that is, not centralized deployment).



It helps immensely in writing the ACLs if there is a wildcard pattern to the subnets you use for, say, UnixAdmin. You do NOT want a list of random subnets. The gotcha is that convenient wildcard patterns generally mean re-addressing. Nobody likes re-addressing their user address space. Worse, the "obvious" trick is 10.(L3 switch).role.x /24, but that is almost guaranteed to conflict with what you are already doing with network 10. Still, re-addressing is the best way to go. By the way, if you think of "voice VLAN" as another role (for phones, membership possibly based on MAC address), then phone security and QoS fit into this re-addressing scheme quite nicely.

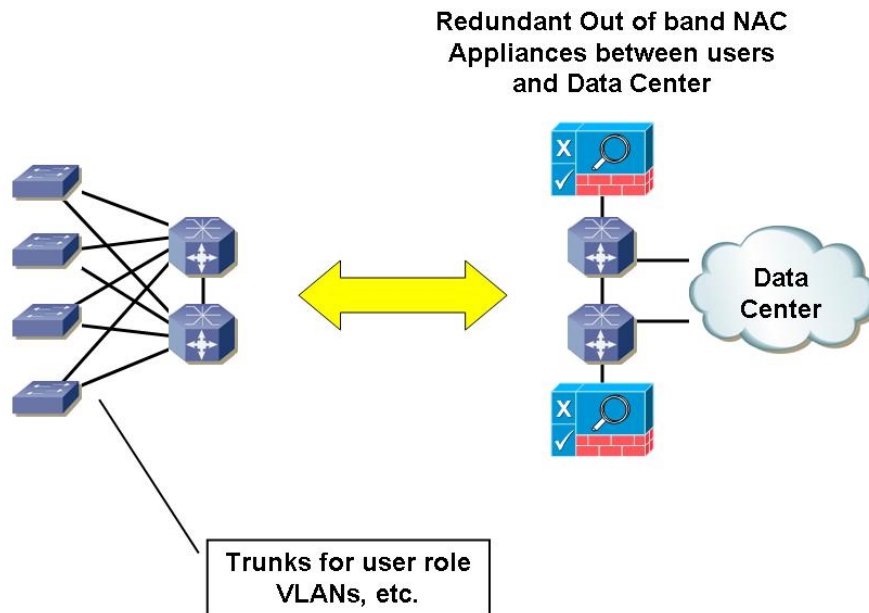
Suppose you need results quickly. There is a temporary workaround that might work for you. It is a bit ugly, so I will definitely emphasize the words "workaround" and "may not work in your circumstances".

The workaround is based on the idea that user roles either have to be externally visible and enforced as different user subnets, or they have to be used for role-based ACLs in an in-band CNACA. That suggests the workaround: use one or multiple CNACA appliances in-band between users and servers. The obvious place to do that quickly is at the data center, typically in the path from users to the data center. CAUTION: the number of users and bandwidth / throughput are definitely scaling factors to watch if considering this. This workaround scales to some extent by using more CNACA appliances.

Suppose the appliance(s) are attached at the data center combined core/distribution layer switch pair. Connect by trunks in routed mode. You can then use Policy-Based Routing (PBR) to selectively force traffic from specific subnets through the CNACA. This allows you to gradually phase in CNACA use while observing performance. You can also use PBR to spread different user subnets across different CNACA appliances, to divide up the workload.

Suppose the quick fix workaround is to be supplemented by a more typical out-of-band edge deployment at building distribution layer switches, using re-addressing. Then as a building is deployed, you can remove the new source subnets from the PBR rules, since enforcement can be by role-based subnet and RACL or FWSM ACL rules, rather than by in-band CNACA ACLs.

Thus, if the scaling works for you, this approach allows fairly smooth phase-in and phase-out of the centralized in-band CNACA deployment in favor of a deployment nearer the edge that may require more time to complete.



Server to Server Controls

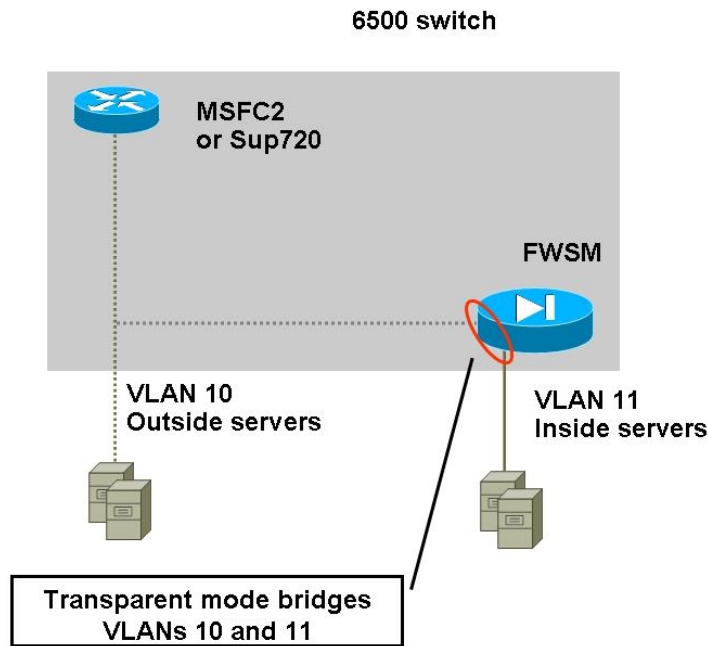
A similar issue arises with servers. Suppose they have to be segmented quickly, but they are all in shared server VLANs. There probably is not time for any rational re-addressing, moving servers to different VLANs, re-building new servers and deploying them, etc. The segmentation might mean controlling traffic between two groups of production servers, perhaps financial and non-financial servers, or production and test servers.

One way to tackle this is using VACL's, assuming the servers connect to a 6500-model switch. VACL's allow IP-based ACLs controlling traffic within a single VLAN. Unfortunately, CSM cannot be used to manage VACL's, so there is no clean object approach to managing lists of different kinds of servers. That could get rather messy to manage! You end up listing specifically what each server is allowed to communicate with, one server at a time.

One alternative is to use a FWSM in transparent mode, possibly using multiple contexts. The FWSM in transparent mode can bridge between two VLANs. Thanks to Elton Fontaine, Cisco SE, for suggesting this approach! One advantage of this is that the FWSM supports object-oriented ACLs. That could simplify the ACLs -- for example, all the server probably need DNS, NTP, and LDAP services. Another advantage of using the FWSM for this is that ASDM and CSM can be used to manage it.

The idea is to divide the servers into "outside" and "inside" groups. Say these will be put onto VLANs 10 and 11, respectively. The FWSM will control all traffic from the network or the outside server group to the inside (protected) group. Deployment will involve putting the FWSM in place with a "permit any any" rule between VLANs 10 and 11. The ports with inside servers on them will then be shifted to VLAN 11. No changes of IP address will occur. The desired ACL can then be deployed in some fashion. It will filter traffic between VLANs 10 and 11. Note that it will not control internal traffic within VLAN 11 servers.

Here is a schematic suggesting how that works:



Creating or changing access lists for production servers is generally a very touchy topic. Nobody wants those servers and applications to stop working. Yet somehow an ACL has to be deployed. And generally the server or application owner cannot tell you what ports the server needs. Even if they can, there can be the issue of how much you trust the information.

One way to reduce the degree of risk here is to collect good data on server to server traffic. A permissive ACL with logging can be used. NetFlow could also be used. The main problem is that there may be scripts or programs that only run quarterly or annually, especially on financial servers. Testing or planning will need to incorporate some way of dealing with that.

Managing Segmentation

I personally like the idea of re-addressing when doing segmentation. And no, that is not trying to create more work for consultants.

For user role-based controls, re-addressing is essential if you do not use NAC Appliances in-band (where you can keep things simple by writing role-based ACL's without specifying the source address). Otherwise you end up with long ad hoc lists of user subnets for each role, which is tedious and error-prone to administer. I have commented previously on good address schemes for IP Telephony (to simplify QoS and Security). Well, now I view IP phones and IPT gear as one more role, along with users, admins, guests, contractors-from-Company-A, consultants-from-Company-B, etc.

Hint: better plan for growth, when you start doing this you will be surprised at the requirements that come out of the woodwork. You will also have to exercise some expectations control. For example, micro-roles for each subtle difference in the user population is not where you want to take user role-based access controls. Network access lists as a tool are appropriate for moderate granularity of user roles.

Re-addressing servers can also help. It simplifies defining server objects in object-oriented tools. It simplifies documenting, understanding, and troubleshooting which servers are in what groupings.

Realistically, you might be able to re-address users, over time. Re-addressing servers might have to take place gradually, as older servers are replaced or consolidated. It still might be a good thing to aim for.

Tools such as ASDM and CSM should probably be used to manage segmentation, or at least the ACL aspect of it. ASDM is fine if you have one or two firewalls. For more, CSM is the way to go. CSM allows you to define a library of ACL objects, such as subnets or lists of addresses, also lists of TCP or UDP ports. You can then build ACL objects or "chunks". A specific firewall ruleset can then be defined based on those chunks and some additional entries, and deployed to the device. If the device is a PIX, ASA, or FWSM, the objects get pushed to it. If it is a router or 6500 RP / Sup720, then the ACL is expanded into a Router-style ACL (RACL) and then pushed.

Using ASDM or CSM also allows you to provide role-based GUI-driven ACL and security management access for the security team. That can help: CheckPoint administrators sometimes exhibit strong CLI-phobia when the word "Cisco" is mentioned.

Conclusion

Are you experiencing any of the trends I think I am seeing? Do you see things differently? Let me know by email!

Some good links on NAC follow in the next table:

Topic	URL

Official Cisco NAC page	http://www.cisco.com/go/nac , which currently takes you to: http://www.cisco.com/en/US/netsol/ns466/networking_solutions_package.html
NAC Framework	http://www.cisco.com/en/US/netsol/ns617/networking_solutions_sub_solution_home.html
NAC Appliance	http://www.cisco.com/en/US/products/ps6128/index.html

Some good links concerning Data Center Design follow in the next table:

Topic	URL
Cisco SRND design guides (design Best Practices)	http://www.cisco.com/go/srnd , which currently takes you to: http://www.cisco.com/en/US/netsol/ns656/networking_solutions_program_category_home.html
Cisco Data Center SRND's	http://www.cisco.com/en/US/netsol/ns656/networking_solutions_design_guidances_list.html#anchor3

For the Firewall Services Module (FWSM), here are some good links:

Topic	URL
FWSM product page	http://www.cisco.com/en/US/products/hw/modules/ps2706/ps4452/index.html
FWSM 3.1 Configuration Guide	http://www.cisco.com/en/US/products/hw/switches/ps708/products_module_configuration_guide_book09186a0080579a1e.html
FWSM 3.1 Command Reference	http://www.cisco.com/en/US/products/hw/switches/ps708/products_command_reference_book09186a00804838a0.html
FWSM 3.1 Log Configuration and Logging Messages	http://www.cisco.com/application/pdf/en/us/guest/products/ps708/c2001/ccmigration_09186a0080687693.pdf

I do hope to write about NAC, both Appliance and Framework, sometime soon. What are they, how do you design for them, how do you select which one to deploy? I have been working with a large customer and will be reviewing precisely those issues with them. I have already had several more limited discussions with customers along similar lines.

Your comments, questions, and suggestions for future articles are of course welcome! See below to decipher my email address.

Dr. Peter J. Welcher (CCIE #1773, CCSI #94014, CCIP) is a Senior Consultant with Chesapeake NetCraftsmen. NetCraftsmen is a high-end consulting firm and Cisco Premier Partner with multiple specializations, dedicated to quality consulting and knowledge transfer. NetCraftsmen has eight CCIE's, with expertise including large network high-availability routing/switching and design, IP Telephony, QoS, MPLS, IPSec VPN, wireless LAN and bridging, network management, security, IP multicast, and other areas. See <http://www.netcraftsmen.net> for more information about NetCraftsmen. Pete's links start at <http://www.netcraftsmen.net/welcher>. New articles will be posted under the Articles link. Questions, suggestions for articles, etc. can be sent to pjw <at> netcraftsmen <dot> net.

2/25/2007
Copyright (C) 2007 Peter J. Welcher