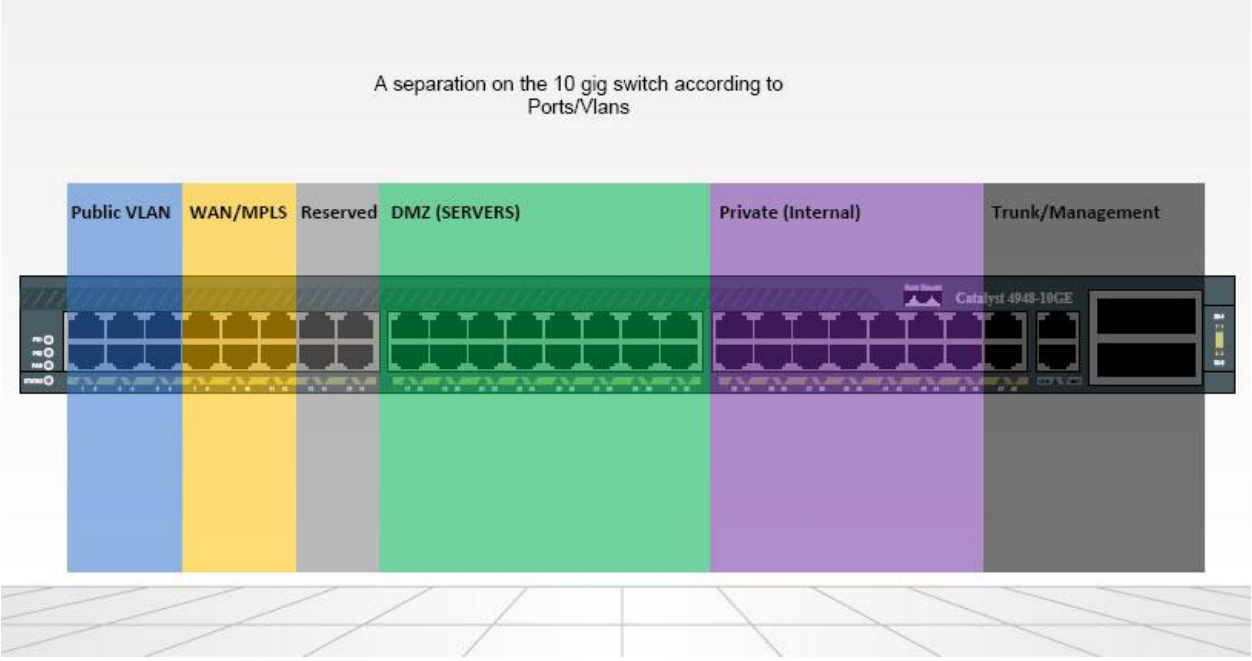
طراحی یک دیتا سنتر

پروژه جدیدی داریم و من مسئول طراحی و پیاده سازی شبکه ای تماما Cisco برای یک Datacenter تقریبا متوسط شده ام؛ این شبکه نقطه مرکزی دیتای شرکتی سرمایه گزاری برای حوزه خاورمیانه فعالیت های خود است که بالغ بر هزاران ترابایت دیتا در SAN هایش در جریان و پهنای باندی ششصد مگابیتی با 4-Core فیبر به اینترنت (تنها برای شانزده سرور) دارد هرچند که اجزای داخلی با سرعت 10Gig به هم وصلند.

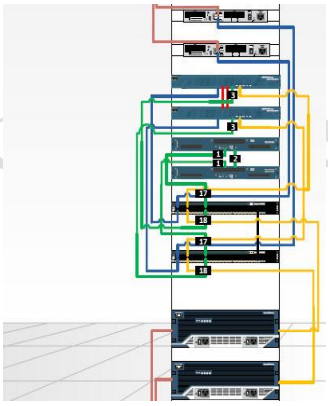
در این پروژه از دستگاه های Cisco ACE برای Load-balancing و ASA های Redundant برای امنیت و IPS برای جلوگیری و ثبت حمله ها استفاده کردیم. سوئیچ های Datacenter از سری 4900-10G است که سیسکو برای Datacenter ها طراحی کرده و از تکنولوژی Catalyst IOS استفاده میکند. دسترسی به اینترنت با کمک BGP و اتصال به دو سرویس دهنده از طریق فیبر و از دو مسیر مختلف برقرار شده است. برای اتصال به دفاتر دیگر و Partner ها از دو 3845 استفاده کردیم تا به شبکه MPLS متصل شوند.

از هر دستگاه چند عدد تهیه شده که هر کدام نقش Failover دیگری را دارند و هر Set از دستگاه ها در Rack مجزا قرار گرفته اند. هر رک به یک برق مستقل (Power Source) و UPS مستقل وصل است.

در اجرا تمام لینک های بین دستگاهها بصورت VLAN مجزا در 4948-10G تعریف شده و نهایتا رک ها با 20Gbps توسط سوییچ ها به هم ترانک شده (EtherChannel) تا همه اجزا به هم، Redundant-path در لایه های مختلف داشته باشند.



شکل بالا نحوه اختصاص VLAN و Private VLAN در هر یک از سوئیچ ها را مشخص میکند.
شکل زیر نحوه کابل کشی و ارتباط فیزیکی داخل یک رک را نشان میدهد:



اینترنت توسط فیبر به روتر های سیسکو متصل شده و ارتباط تا سرویس دهنده توسط BGP تا لایه 4 تست میشود و در صورتیکه مشکلی بوجود آید این رک از رک همسایه خود برای Route کردن ترافیک از سرویس دهنده دیگر کمک میگیرد. اگر برق این رک را قطع کنیم؛ 30٪ از کل دستگاه های سیسکو خاموش میشود اما تمام ارتباطات در 50ms (میلی ثانیه) به رک دوم منتقل میشود زیرا که Redundancy بین روترها، فایروال ها، Load-balancer ها و حتی Interface های سرور و خود سرورها وجود دارد. برای ارتباط با روتر های متصل به اینترنت از GLBP یا Gateway Load-balancing Protocol استفاده کردیم و توسط BGP ترافیک خروجی و ورودی Load-balance شده است و در زمان Failure با Prefix کوچکتر (از لحاظ بیتی بزرگتر) به بیرون گزارش میشود.

برای Cisco ASA5540 نیز بصورت Redundant از Failover Link و Stateful Redundant Link استفاده کردیم تا حتی ارتباط IPSEC VPN های Site-to-Site از بیرون به داخل در موقع Failure، قطع نشود. ASA ها از IP SLA (Tracking) برای انتخاب Router های خروجی به اینترنت استفاده میکنند.

IPS ها نیز کار Intrusion Prevention را بصورت redundant و Inline انجام میدهند و Signature آنها بصورت Automatic با وب سایت سیسکو از طریق SSL؛ update میشود.

از Cisco MARS برای Event Correlation استفاده کردیم تا تمامی SNMP Trap ها و لاگ های IPS و Firewall و IOS های Advanced-Security را در MARS جمع کنیم و در بازه های زمانی مشخص توسط NOC بررسی کنیم.

بعد از این که از Redundancy و Load balancing در لایه های Routing ، Switching و Security مطمئن شدیم برای هر سرویس و URL چند سرور نصب کردیم و هر سرور 4 کارت شبکه فیبر دارد. سخت افزار اصلی سرور های وب Blade Server است که داخل آن چهار Cisco Blade Switch قرار دارد و با 16 uplink به Datacenter Switch متصل شده است. روی این سخت افزار VMWARE ESX Server نصب شده تا لایه مربوط به سخت افزار ها را کنترل کند و Image های هر سرور را روی Device Pool قرار دهد و در موقع نیاز Resource و Memory بیشتری به سرور بتوان اختصاص داد.

به کمک VMOTION میتوانیم یک سرور با سیستم عاملش را از یک سخت افزار به سخت افزار دیگر در کمترین زمان منتقل کنیم و برای Recovery هر سرور از Snapshot های گرفته شده بهره ببریم.

Cisco ACE بصورت یک L4-L7 Switch/Load balancer کار میکند و ترافیک ورودی را بین سرور های یک Pool پخش میکند و بار ورودی برای سرورهای SSL بر اساس درخواست در ثانیه بین سرور ها بالانس میشود. ACE خود URL های برنامه را چک میکند و اگر سروری در زمان مناسب جواب نداد آنرا از دور (بصورت موقتی) خارج میکند. پس تا لایه Application؛ Redundancy برقرار میگردد.

در واقع یک Virtual Server تعریف میکنیم و DNS URL اینترنتی را به آن map میکنیم؛ آن Virtual IP توسط ASA ترجمه Static NAT شده به یک VIP آدرس Internal که به Load-balancer اشاره میکند میرسد. سپس درخواست به یکی از سرور های واقعی آن گروه ارسال شده و کاربر بیرونی توسط Sticky Cookie هر بار به همان سرور رجوع خواهد کرد تا Integrity حفظ شود مگر این که آن سرور به مشکل خورده و ادامه Session با سرور دیگری برقرار گردد.

هر یک از Partner های خارجی این شرکت دستگاه های خود را بصورت (جفت) Redundant به Datacenter آوردند و به Stack ی از 3750 ها متصل کردند و از طریق Private VLAN برای ارتباط با سرورها از ASA عبور میکنند.

ارتباط از بیرون برای تنظیم و تغییر سرور ها از طریق MPLS و البته Redundant به دو Cisco Router 3845 برقرار میشود و به کمک SSL VPN میتوانند برنامه های خود را Upgrade کنند و شبکه را با Ciscoworks LMS؛ مدیریت کنند. توسط LMS از IOS ها و نسخ مختلف Configuration بصورت Backup – Automatic گرفته شده و تغییرات هر کاربر ثبت میشود. برای Authentication و Authorization دستورات از Cisco Secure ACS Appliance استفاده کردیم و هر دستور کاربر توسط RADIUS با سطح دسترسی او چک شده و Log میگردد.

توسط Cisco Secure ACS به هر User یک Access-list داده میشود تا تنها با IP هایی که لازم است بتواند Packet ارسال/دریافت کند.

Border Router های اتصال به اینترنت نیز قابلیت Advanced Security دارند و پس از SSL VPN میتوان به آنها SSH کرد. در آنها Access-list های Infrastructure Protection و Bogon List قرار دادیم تا کارایی Spoof Mitigation و جلوگیری از Denial of Service را در خارجی ترین نقطه از آنها بگیریم.

قسمتی از Config روتر ها جهت استفاده در هر Border Router ی متصل به اینترنت:

RFC3330 Block Bogons, Spoof Mitigation, uRPF/Static Map.
Router Filter for Ingress Traffic (Infrastructure ACL)

```
!--- Refer to RFC 3330 for additional special use addresses.

access-list 110 deny ip host 0.0.0.0 any
access-list 110 deny ip host 255.255.255.255 any
access-list 110 deny ip 127.0.0.0 0.255.255.255 any
access-list 110 deny ip 192.0.2.0 0.0.0.255 any
access-list 110 deny ip 224.0.0.0 31.255.255.255 any

!--- Filter RFC 1918 space.

access-list 110 deny ip 10.0.0.0 0.255.255.255 any
access-list 110 deny ip 172.16.0.0 0.15.255.255 any
access-list 110 deny ip 192.168.0.0 0.0.255.255 any

!--- Deny your space as source from entering your AS.
!--- Deploy only at the AS edge.

access-list 110 deny ip YOUR_CIDR_BLOCK any

!--- Permit BGP.

access-list 110 permit tcp host bgp_peer host router_ip eq bgp
access-list 110 permit tcp host bgp_peer eq bgp host router_ip

!--- Deny access to internal infrastructure addresses.

access-list 110 deny ip any INTERNAL_INFRASTRUCTURE_ADDRESSES

!--- Permit transit traffic.

access-list 110 permit ip any YOUR_CIDR_BLOCK
```


در ASA نیز از مکانیزم ساده (Health Check) تشخیص سلامت Internet بکمک ICMP استفاده کردیم:

```
route outside 0.0.0.0 0.0.0.0 192.168.1.252 1 track 252
```

```
route outside 0.0.0.0 0.0.0.0 192.168.1.253 1 track 253
!
sla monitor 1
 type echo protocol ipIcmpEcho x.x.x.x interface outside
sla monitor schedule 1 life forever start-time now
track 252 rtr 1 reachability
```

Datacenter فوق در محل Etisalat Datacenter توسط سیستم های Air conditioner و Fire Detection شرکت اتصالات دبی Host شده و برای Disaster Recovery قرار است نمونه مشابه Datacenter فوق را در یک کانینتر متحرک در Datacenter ابوظبی Host کنیم تا در موقع لزوم قابل حمل و نقل باشد در آن پلاتفرم به دلیل کمبود جا از دو 6500 با چهار (Line Card) کارت های FWSM (فایروال) و ACE استفاده میکنیم و رک های داخل Container خود؛ A/C، UPS و Fire Isolation دارند و تنها یک فیبر و یک برق سه فاز به بیرون Container لازم است.

آنچه که در حین طراحی و پیاده سازی این دیتاسنتر بیش از همه در ذهنم جذابیت داشت و به آن فکر میکردم؛ نه استفاده از دستگاه های مختلف بلکه استفاده از تکنولوژی های مختلف و بروز؛ Integration آنها و تنظیم و تست درست آن بود. پروژه ای که کارفرما در پایان به آن State of Art in Configuration گفت... شبکه ای شاید به کیفیت، سرعت، هارمونی و قدرت Lamborghini Murcielago V12...