Universidade Federal do Rio de Janeiro

Núcleo de Computação Eletrônica

Bruno Maggi Batista

REDES CORPORATIVAS COM VIRTUAL SWITCHING

Rio de Janeiro

2010

Bruno Maggi Batista

REDES CORPORATIVAS COM VIRTUAL SWITCHING

Monografia apresentada para obtenção do título de Especialista em Gerência de Redes de Computadores no Curso de Pós-Graduação Lato Sensu em Gerência de Redes de Computadores e Tecnologia Internet do Núcleo de Computação Eletrônica da Universidade Federal do Rio de Janeiro – NCE/UFRJ.

Orientador:

Moacyr Henrique Cruz de Azevedo, M. Sc., UFRJ, Brasil

Rio de Janeiro

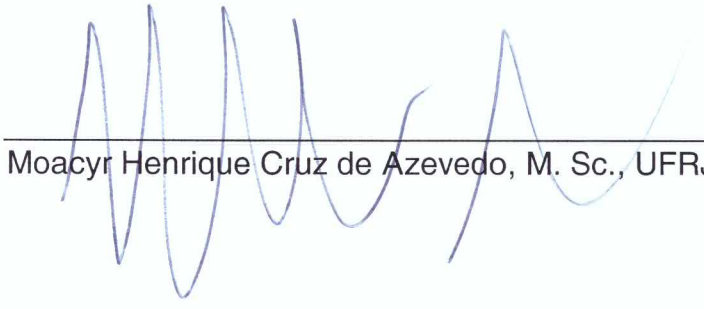
2010

Bruno Maggi Batista

REDES CORPORATIVAS COM VIRTUAL SWITCHING

Monografia apresentada para obtenção do título de Especialista em Gerência de Redes de Computadores no Curso de Pós-Graduação Lato Sensu em Gerência de Redes de Computadores e Tecnologia Internet do Núcleo de Computação Eletrônica da Universidade Federal do Rio de Janeiro – NCE/UFRJ.

Aprovada em outubro de 2010.



Moacyr Henrique Cruz de Azevedo, M. Sc., UFRJ, Brasil

Dedico essa monografia à minha madrinha Marlene Bessa Maggi e minha avó Celestina Emília Bessa Maggi por tornarem esse sonho possível.

AGRADECIMENTOS

Gostaria de agradecer aos meus familiares, principalmente meus pais, José Carlos Batista e Marilda Bessa Maggi Batista, e também à minha irmã Carla Maggi Batista, por sempre estarem presentes acreditando no meu desenvolvimento educacional e profissional.

Agradeço em especial ao meu tio Marcos Barbosa Vasques pelo incentivo ao aprendizado e por ser exemplo de que nunca devemos estacionar nosso investimento intelectual. Um guerreiro incansável na busca pelo conhecimento.

Aos colegas e amigos de trabalho, primeiramente a equipe de redes com Daniel Araújo e Haroldo Donato, pela parceria e comprometimento no dia-a-dia da Cetip, sempre juntos para a evolução do setor. Também agradeço ao meu companheiro de pós-graduação e de Cetip, Magno Araújo, pela troca de experiências durante o curso e apoio nas horas mais complicadas.

E para terminar, um enorme abraço para o grande Moacyr Azevedo, professor, orientador e diretor do curso, por sua dedicação e principalmente, pela ajuda e paciência, para orientar na elaboração dessa monografia.

RESUMO

BATISTA, Bruno Maggi. **REDES CORPORATIVAS COM VIRTUAL SWITCHING**. Monografia (Especialização em Gerência de Redes e Tecnologia Internet). Núcleo de Computação Eletrônica, Universidade Federal do Rio de Janeiro. Rio de Janeiro, 2010.

Estudo de soluções em virtualização de switches dos principais fabricantes do mercado, aplicando em redes de corporações com uso de links de alta velocidade. É dada ênfase nas tecnologias e protocolos necessários para aplicação da solução virtualizada.

Com uma análise dos três principais fabricantes do mercado: Cisco, Nortel e Juniper, será proposta uma solução para adequação na rede da Cetip S.A., cenário este que necessita de replicação entre sites e redundância total, tanto na infra-estrutura de rede como nos sistemas envolvidos.

ABSTRACT

BATISTA, Bruno Maggi. **REDES CORPORATIVAS COM VIRTUAL SWITCHING**. Coursework (Specialization in Computer Network Management and Internet Technology). Núcleo de Computação Eletrônica, Universidade Federal do Rio de Janeiro. Rio de Janeiro, 2010.

Study of virtualization solutions from the leading manufacturers of switches market, investing on corporate networks with the use of high-speed links. The emphasis is done on technologies and protocols needed to implement virtualization solution.

With an analysis of the three main manufacturers: Cisco, Nortel and Juniper, will be offered a solution to fit in Cetip S.A. network, this scenario requires that the replication between sites and full redundancy both in network infrastructure and systems involved.

LISTA DE FIGURAS

LISTA DE TABELAS

LISTA DE ABREVIATURAS E SIGLAS

ACL	*Access Control List*
ATM	*Asynchronous Transfer Mode*
CAPEX	*Capital expenditure*
CEE	*Converged Enhanced Ethernet*
CoS	*Class of Service*
DWDM	*Dense Wavelength Division Multiplexing*
FCoE	*Fibre Channel over Ethernet*
FR	*Frame Relay*
GbE	*Gigabit Ethernet*
Gbps	*Gigabit per second*
GLBP	*Gateway Load Balancing Protocol*
HSRP	*Hot Standby Router Protocol*
IEEE	*Institute of Electrical and Electronics Engineers*
IP	*Internet Protocol*
IST	*Inter-Switch Trunking*
IST	*Inter-Switch Trunking*
LACP	*Link Aggregation Control Protocol*
LAN	*Local Area Network*
MAC	*Media Access Control*
Mbps	*Megabit per second*
MEC	*Multichassis EtherChannel*
MLT	*Multi-Link Trunking*
MPLS	*Multi Protocol Label Switching*
NSF	*Nonstop Forwarding*
OLTP	*Online Transaction Processing*
PAgP	*Port Aggregation Protocol*
PNNI	*Private Network-to-Network Interface*
QoS	*Quality of Service*
RSFN	*Rede do Sistema Financeiro Nacional*
R-SMLT	*Routed Split Multi-Link Trunking*
RTM	*Rede de Telecomunicações para o Mercado*
SAN	*Storage Área Network*
SMLT	*Split Multi-Link Trunking*
SPB	*Sistema de Pagamentos Brasileiro*
SSO	*Stateful Switchover*
STA	*Spanning Tree Algorithm*
STP	*Spanning Tree Protocol*
TDM	*Time-Division Multiplexing*
VLAN	*Virtual Local Area Network*
VoIP	*Voice over Internet Protocol*
VRRP	*Virtual Router Redundancy Protocol*
VSH	*Virtual Switch Header*
VSI	*Virtual Switch Interface*
VSL	*Virtual Switch Link*
VSS	*Virtual Switching System*
WAN	*Wide Area Network*

SUMÁRIO

1 **INTRODUÇÃO**

Com a crescente utilização do conceito de virtualização, empresas passaram a aplicá-lo também na infraestrutura de rede, principalmente nos grandes data centers onde se busca simplificar o gerenciamento e aperfeiçoar a utilização dos recursos disponíveis.

Isso se traduz em economia de custos, espaço físico, energia, refrigeração e consequentemente, na redução do orçamento.

Em seu funcionamento, uma solução de redes virtualizada, oferece alocação de recursos lógicos que são direcionados a determinados grupos de hosts, subredes e serviços, ao invés de segmentação física.

Existem vários cenários onde essa tecnologia pode ser aplicada preservando o ambiente versátil, escalável e seguro, como: data centers, campus universitário, bancos e empresas que necessitam estender sua rede local em geral.

2 REFERENCIAL TEÓRICO

2.1 O QUE É VIRTUAL SWITCHING

O Virtual Switching é a tecnologia que permite agrupamento de múltiplos switches em um único dispositivo virtual ou associar uma única função a vários dispositivos físicos. Sua utilização envolve várias aplicações diferentes (IP, voz, ATM) e diversos tipos de configurações, podendo cada função ter seu próprio desempenho e controles de segurança. Diversas combinações de serviços são criadas gradativamente para atender novas demandas sem precisar adicionamento de novo hardware. Toda essa flexibilidade é organizada em uma arquitetura multicamada onde são definidos planos para melhor organização do seu entendimento e funcionamento.

Utiliza tecnologia *EtherChannel* agrupando canais de alta velocidade (10Gbps) que garantem sincronismo entre os equipamentos, comutação de pacotes na sua essência, e o mecanismo de funcionamento a prova de falhas.

2.2 ESTRUTURA

O *Multiservice Switching Forum* dividiu o funcionamento do *virtual switching* em 5 planos. São eles:

a) Adaptação

Interface física para o dispositivo de rede, podendo ser um usuário ou qualquer outro elemento. Permite que a porta do switch seja particionada para ser utilizada por planos de controle diferentes. Esse plano fica ativo em todos os membros da solução virtual.

b) Switching

Representa o funcionamento do switch de acordo com as portas que estão sendo utilizadas. Pode trafegar simultaneamente serviços diferentes controlados por diferentes planos de controle. Esse plano fica ativo em todos os membros da solução virtual.

c) Controle

Gerencia os eventos de serviços da rede, controlando os dois planos anteriores. O equipamento pode conter vários planos de controle associados às interfaces virtuais. Porém apenas um switch comanda os planos de controle enquanto os demais ficam em standby

d) Aplicação

Provê serviços que usam os recursos do plano de controle.

e) Gerenciamento

Também provê serviços utilizando os recursos do plano de controle.

A figura 1 exemplifica a divisão desses planos:

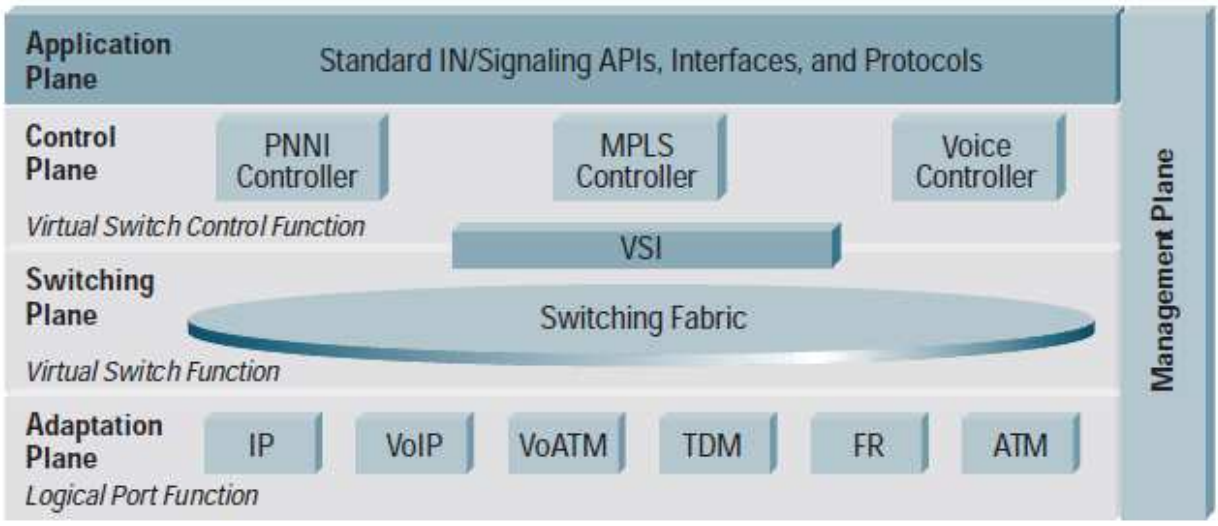


Figura 1 – Arquitetura do Virtual Switching

2.3 VANTAGENS

A tecnologia de Virtual Switching permite a formação de um *pool* de switches físicos criando um único switch virtual. Os principais ganhos em uma rede são:

- Aumento de eficiência operacional

Ao se criar uma estrutura virtual agrupando switches, a rede corporativa é simplificada, principalmente quanto ao gerenciamento. A configuração é feita através de um host único. Elimina a necessidade de protocolos para soluções redundantes como: HSRP, VRRP e GLBP.

Outro protocolo bastante difundido que também deixa de ser obrigatório é o Spanning Tree Protocol (STP). Em seu lugar, entra o Multichassis EtherChannel (MEC) como forma de evitar loops na rede. Na implantação do virtual switching a empresa ainda pode optar pela escolha de instalar dois elementos em sites diferentes que, dependendo da característica da fibra ótica, poderão estar até a 40 km de distância.

- Comunicações ininterruptas

Mesmo com a falha de um switch virtual, as aplicações que dependem de informações do estado da rede não sofrem interrupções. Utiliza EtherChannel (802.3ad ou Port Aggregation Protocol (PAgP)) para recuperação do link da camada de enlace, eliminando a dependência de Spanning Tree Protocol para a recuperação da ligação.

- Aumento da banda de transmissão disponível

Possibilidade de ativar vários canais de 10 Gbps. O balanceamento de carga do virtual switching é mais otimizado que o aplicado no STP com VLANs porque é baseado em aspectos mais granulares, com informações de camada de enlace, rede

e transporte. Além disso, elimina o tráfego antes desnecessário de roteamento entre sites, já que os membros se encontram na mesma LAN.

- Utiliza arquitetura multilayer switching

A arquitetura em camadas permite ao administrador de rede uma melhor compreensão sobre os elementos que envolvem a solução total, podendo estender sua rede em distâncias consideráveis gerando versatilidade na comunicação intra-sites.

2.4 MODELO MULTICAMADAS

O modelo multicamadas é bastante difundido para separar três tipos de segmentos que uma rede corporativa pode ter em sua funcionalidade: *core*, distribuição e acesso. Nem sempre todas as empresas possuem essas 3 camadas bem definidas, principalmente quando se trata de redes de menor porte. Hierarquicamente o Virtual Switching pode ser utilizado nas camadas core, distribuição e acesso, sendo mais comum no core e distribuição. Cada ponto apresenta vantagens específicas, podendo inclusive ser implementado conjuntamente em todas as camadas.

2.4.1 Acesso

Essa camada controla o acesso do host final aos recursos de rede. Em uma solução com Virtual Switching, essa camada é habitada por servidores que fazem parte do negócio da empresa. Geralmente esses recursos estão localizados localmente sem precisar de roteamento para outra subrede. Quando há esta necessidade, entra em cena a camada de distribuição.

2.4.2 Distribuição

A camada de distribuição é o principal ponto de comunicação entre as camadas de acesso e core. A principal função da camada de distribuição é fornecer roteamento, filtragem e acesso WAN, além de poder determinar como os pacotes podem acessar o core, se necessário. A camada de distribuição deve determinar a maneira mais rápida que as solicitações de serviços de rede são tratadas, por exemplo, como um pedido de arquivo é enviado para um servidor. A camada de distribuição é o lugar recomendado para implementar políticas para a rede. Ela possui uma flexibilidade na definição da operação da rede. Há vários itens que podem ser aplicados à camada de distribuição, tais como:

- ACL, filtros de pacote e de filas
- Políticas de segurança
- Redistribuição de tabelas de roteamento
- Roteamento entre VLANs
- Definição de domínios de broadcast e multicast

2.4.3 Core

A camada core representa o *backbone* da rede corporativa. Ela faz a função de *gateway* para a Internet. Responsável pelo transporte de grande volume, sua função é entregar o tráfego o mais rápido possível com a menor latência possível. Por isso, o uso de ACLs, roteamento e filtragem de pacotes não são recomendados, assim como o uso de protocolos de roteamento com alto tempo de convergência. Um fator crítico dessa camada é a tolerância às falhas. A presença de links e equipamentos redundantes é de extrema importância nessa camada.

2.5 TECNOLOGIAS ENVOLVIDAS

Serão tratadas nesse tópico primeiramente duas tecnologias bastante difundidas nas redes de empresas, mas que com as novas tecnologias de virtual switching se tornaram obsoletas. Depois será abordada a agregação de links, ponto essencial para o desempenho da tecnologia além de garantir sua redundância. E por último serão apresentados dois protocolos mais utilizados pelos fabricantes de switches de grande porte do mercado para a solução de virtualização de switches.

2.5.1 Spanning Tree Protocol

O 802.1D STP (*Spanning Tree Protocol*) foi projetado para solucionar problemas em redes redundantes. O protocolo *Spanning Tree* executa um algoritmo chamado STA (Spanning Tree Algorithm) para encontrar enlaces redundantes, negociando um caminho livre de *loops* através da eleição de um switch principal, denominado *root*. Todos os switches possuem o controle para identificar um enlace que foi alterado na topologia, e assim realizar a convergência do protocolo STP.

O fator determinante para definição do *switch root* é seu ID. Cada switch tem seu ID formado por uma prioridade que pode ser configurada e por seu endereço MAC. Primeiramente a prioridade é comparada e o switch com a maior prioridade é eleito o *root*. Caso as prioridades sejam iguais o endereço MAC do switch é comparado e o switch com menor MAC é eleito o *root*.

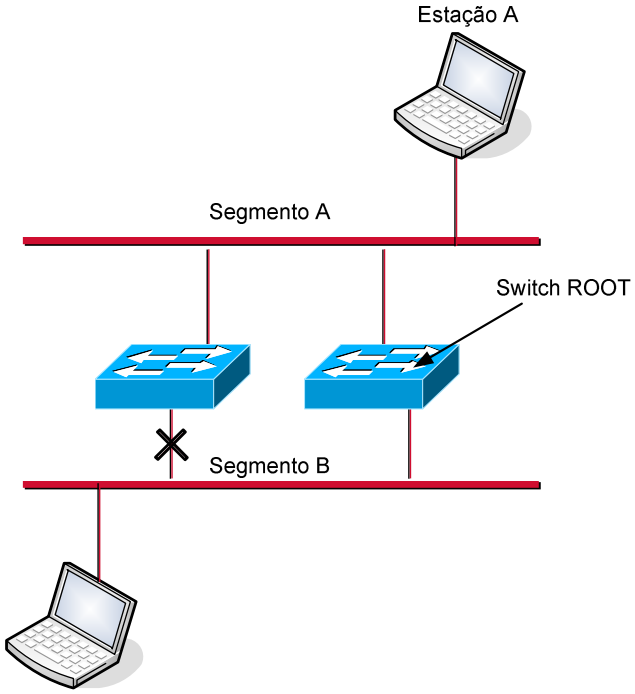


Figura 2 – Exemplo Simplificado STP

Apesar de ainda ser aplicado em muitas empresas, existem 3 problemas fundamentais com o STP que são incompatíveis com as premissas de construção de uma ambiente com *virtual switching*:

1. Convergência lenta - Dependendo do tamanho da rede, se ocorrer a falha de um link ou uma falha do switch, a rede pode levar vários minutos para voltar a operar normalmente.

2. Largura de banda ociosa – Para evitar loops, o STP bloqueia alguns links, resultando em desperdício de parte da largura de banda.

3. Design Complexo – Para evitar largura de banda ociosa, são criados múltiplos domínios de spanning tree. Isso aumenta a complexidade porque cada grupo requer uma eleição de *root*. Além disso, ainda existe o desafio de equilibrar o tráfego de rede entre as ligações existentes.

2.5.2 Virtual Router Redundancy Protocol

Existem vários protocolos para redes redundantes, para roteamento optou-se por falar do VRRP que é um padrão aberto e que pode ser combinado com algumas soluções dos principais fabricantes, apesar de sua convergência levar alguns segundos.

VRRP é um protocolo não-proprietário de redundância descrito na RFC 3768 destinado a aumentar a disponibilidade do *gateway* em uma subrede. A confiabilidade é obtida pela publicação de um roteador virtual como gateway padrão para os hosts, ao invés de um roteador físico. Dois ou mais roteadores físicos, então, devem estar configurados para divulgar o roteador virtual, com apenas um deles fazendo o encaminhamento real a todo o momento. Em caso de falha do roteador físico que está tratando o roteamento de dados em nome do roteador virtual, um acordo será realizado para outro roteador físico desempenhar seu papel automaticamente. O roteador físico que é responsável pelo encaminhamento dos dados em nome do roteador virtual é chamado de *master*. Os demais são chamados de *backup*.

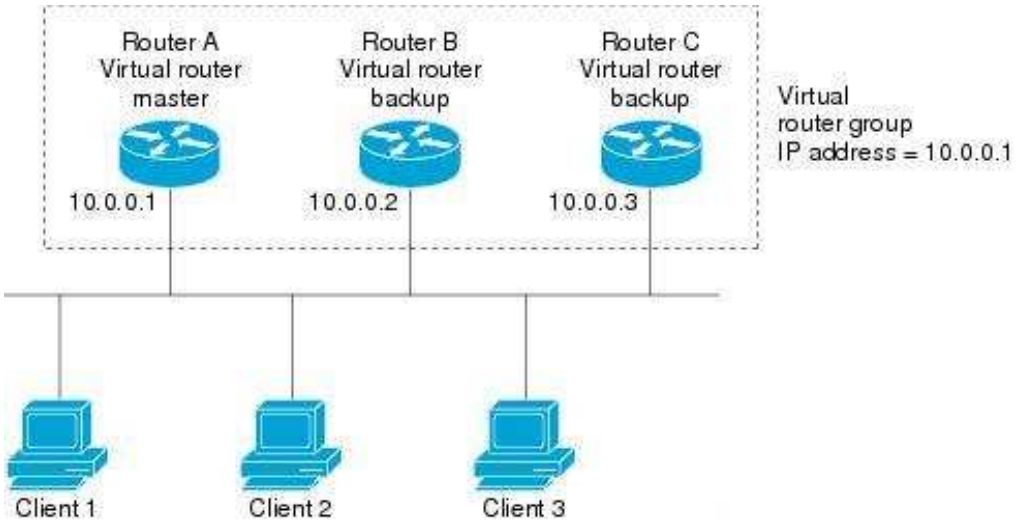


Figura 3 – Exemplo VRRP

Caso o roteador *master* fique incapaz de receber um pacote VRRP *multicast* dentro de um período superior a três vezes o tempo da divulgação, os roteadores *backup* supõem que o roteador *master* não está funcionando. Então o processo de eleição é iniciado através de mensagens *multicast* dos roteadores *backup*. O roteador *backup* com maior prioridade passa a ser o *master*, aumentando sua prioridade para um valor acima do antigo. O novo *master* passará a ter a responsabilidade de encaminhamento dos pacotes enviados para o *gateway* virtual. Caso todos os roteadores *backup* tenham a mesma prioridade, o roteador com maior endereço IP passa a ser o *master*.

2.5.3 Link Aggregation

Também conhecido como EtherChannel, termo padronizado pela Cisco, permite a criação de um canal envolvendo diversos links físicos (par trançado ou fibra) que somam suas larguras de banda, funcionando como um link lógico. Os principais ganhos são a tolerância às falhas e links de alta velocidade. Geralmente essas conexões são usadas para interconectar equipamentos das camadas de distribuição e *core*, porém também podem ser usadas para conectar estações.

O Link Aggregation se tornou um padrão aberto a partir de 2000 quando o IEEE oficializou o padrão 802.3ad. Todo funcionamento é controlado pelo Link Aggregation Control Protocol (LACP) permitindo que o dispositivo de rede negocie um agrupamento automático, enviando pacotes para o outro ponto da conexão.

Quando está em operação, todas as portas que fazem parte da solução compartilham o mesmo endereçamento de camada 2 (MAC *address*) porque visualizam uma única conexão lógica.

Se um link falhar, automaticamente o tráfego é redistribuído através dos links restantes. Esta recuperação automática leva menos de um segundo e é transparente para aplicativos de rede e do usuário final. Isto torna a solução muito resistente e essencial para aplicações críticas. O balanceamento do tráfego é representado pela tabela abaixo, de acordo com o número de interfaces:

Nº de Portas	Balanceamento
8	1:1:1:1:1:1:1:1
7	2:1:1:1:1:1:1
6	2:2:1:1:1:1
5	2:2:2:1:1
4	2:2:2:2
3	3:3:2
2	4:4

Tabela 1 – Distribuição de tráfego do Port Aggregation

Seus componentes são:

- Links – Utiliza o padrão definido pelo IEEE 802.3, incluindo todos os sub-padrões. Todos os link de um link aggregation devem ter a mesma velocidade.
- Compatibilidade – As duas pontas da solução precisam ser compatíveis com o padrão IEEE 802.3, assim como no caso do etherchannel, onde todos os switches da linha Cisco, além do LACP, também podem realizar o canal lógico através do protocolo padrão da Cisco PAgP (Port Aggegation Protocol).
- Configuração – A configuração geralmente é feita de forma manual, em cada porta envolvida na solução, que deverá conter um comando de ativação do *link aggregation*. Outra forma, por exemplo, seria em dispositivos Cisco onde o

PAgP detecta que a outra ponta deseja formar o EtherChannel, funcionando de forma passiva no estabelecimento do canal.

2.5.4 Multichassis EtherChannel

As duas tecnologias descritas acima, juntamente com o *Multichassis EtherChannel* (MEC), complementam a arquitetura de virtual switching como uma solução sólida na tolerância às falhas e no seu desempenho. O MEC é usado em soluções Cisco para, além de prover a distribuição de carga no canal VSL (*Virtual Switch Layer*) demonstrado na figura 4, funcionar como solução para links redundantes.

A figura 4 mostra uma representação lógica dos elementos envolvidos na topologia MEC. São dois chassis físicos formando uma única entidade. O MEC representa a continuidade da operação, esteja o dispositivo de rede conectado ao chassi 1 ou 2.

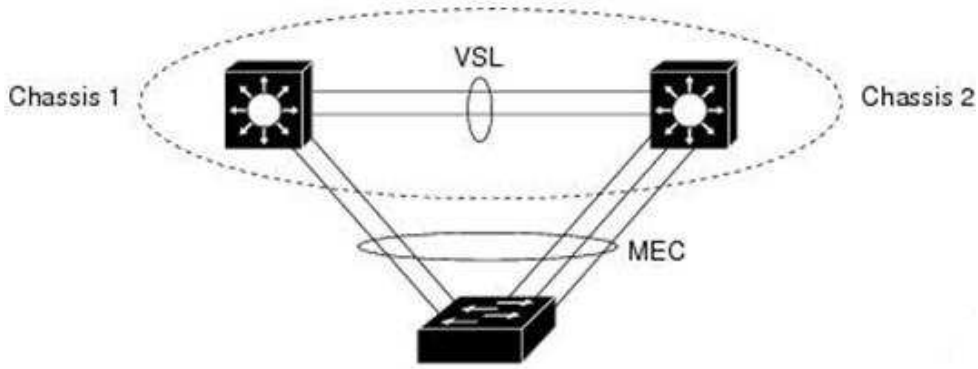


Figura 4 – Exemplo MEC

Apenas um dos chassis fica ativo. Entre eles existe a necessidade do VSL. Para estes dois componentes agirem como um único elemento lógico, eles precisam compartilhar informações e controlar o tráfego de dados. O VSL é um link especial

que leva o controle e o tráfego de dados entre os dois chassis da solução VSS (*Virtual Switching System*), como mostrado na Figura 4. O VSL é implementado como um EtherChannel com até oito links, priorizando o tráfego de controle em relação ao tráfego de dados para que mensagens de controle não sejam descartadas. O tráfego de dados é balanceado entre os links VSL pelo algoritmo de balanceamento de carga do *EtherChannel*.

Se um link do MEC falhar, ele redistribui a carga entre os links operacionais. No caso de falha de todos os links do chassi ativo, o chassi em *standby* passa a funcionar como um EtherChannel normal. Existem outras combinações que serão tratadas mais à frente quando serão abordados os principais fabricantes.

2.5.5 Split Multi Link Trunking

O Split multi-link trunking (SMLT) é um protocolo aberto utilizado por diversos fabricantes, desenvolvido pela Nortel. Permite múltiplas conexões de rede física entre dois switches de rede e outro dispositivo, podendo ser outro switch ou até mesmo um servidor. Os links são tratados como um único link lógico e é feito balanceamento de carga do tráfego em todos os links disponíveis. Para cada pacote que precisa ser transmitido, uma das ligações físicas é selecionada com base em um algoritmo de balanceamento de carga, normalmente envolvendo uma função *hash* que opera nos endereços MAC origem e destino.

O SMLT remove a limitação de que todas as portas físicas do link aggregation devem residir no mesmo switch, permitindo o balanceamento de carga entre links de switches diferentes e mantendo a alta disponibilidade da solução.

A figura 5 é bastante semelhante a figura anterior onde dois chassis operam como um único switch lógico. Nessa representação, a comunicação entre eles é feita por 4 canais de 10Gbps, formando o *Inter-Switch Trunking* (IST).

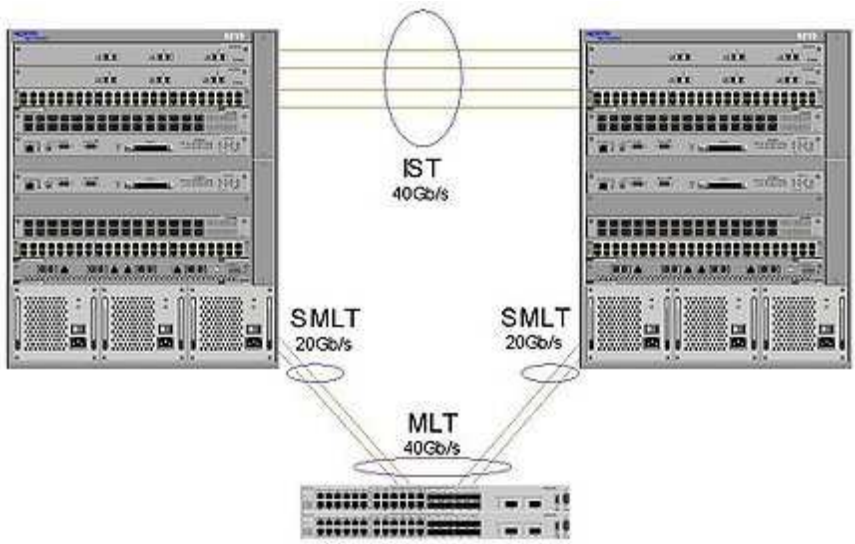


Figura 5 – Exemplo SMLT

A chave para o funcionamento do SMLT é o IST que é uma conexão MLT (*MultiLink Trunking*) entre os switches principais que permite a troca de informações sobre o encaminhamento do tráfego e do estado individual dos links SMLT. Através deste canal são compartilhadas informações de estado das conexões e tabelas de encaminhamento de diferentes swiches.

3 **PRINCIPAIS TECNOLOGIAS DISPONÍVEIS**

Nesse capitulo serão destacados os principais fabricantes do mercado de virtualização de switches e seus produtos. Todos eles enfatizam os seguintes pontos:

- Redução de custos - Aplicado desde o investimento até o custo operacional e de manutenção, como consumo de energia, alocação de espaço físico para os equipamentos.

- Simplificação da rede – Resumo das camadas core, distribuição e acesso(data Center) em um número reduzido de dispositivos.

- Aumento do desempenho – Utilização de links de alta velocidade e equipamentos de grande porte com capacidade de processamento avançada são cruciais para empresas que possuem aplicações críticas de negócios. Diminuição do tempo de resiliência da rede, além da tolerância à falhas.

- Versatilidade – Capacidade de estender a rede da empresa a quilômetros de distância, continuando na mesma LAN.

3.1 JUNIPER 3-2-1 DATA CENTER NETWORK ARCHITECTURE

A Juniper tem como solução para virtualização de switches uma arquitetura que simplifica a topologia de rede, reduzindo o modelo em camadas. O termo 3-2-1 faz menção as três camadas existentes e sua ordem decrescente explica o objetivo de redução das camadas em um gerenciamento e operação mais simples. Promete redução de até 35% no CAPEX, expressão designada para mensurar os custos despendidos em investimentos para a empresa. O conceito de aumentar a rede com novos switches é colocado de lado e concentra-se em uma única solução que abrange o Data Center como um todo.

Os principais produtos desta arquitetura são os switches EX4200, com interfaces de 1GbE, e o modelo EX4500, que se baseia em redes 10GbE, podendo suportar até 48 interfaces. Ambos suportam tecnologias CEE (*Converged Enhanced Ethernet*) e FCoE (*Fibre Channel over Ethernet*), comuns em implementações ethernet de SANs em grandes data centers. Os grandes parceiros dessa solução são empresas como Dell e IBM.

Para realizar a interconexão dos equipamentos existem 2 opções. Através de links de alta velocidade (10GbE) ou através de um cabo proprietário do próprio fabricante.

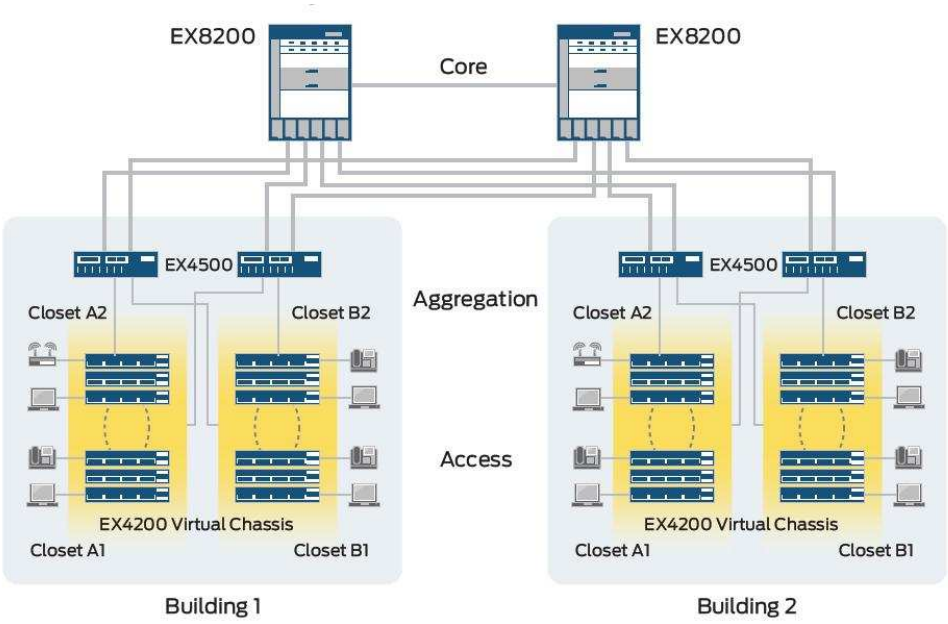


Figura 6 – Solução JUNIPER

Na figura 6 vê-se um exemplo de implementação entre dois *sites* de uma empresa. Os switches da camada *core* estão conectados de forma redundante aos dispositivos da camada de distribuição (chamada de *aggregation* pela Juniper), os switches EX4500. Ao contrário de exemplos dos outros fabricantes, essa solução é comandada pelos switches da camada de distribuição.

3.2 NORTEL SPLIT MULTI LINK TRUNKING

Tendo sido criadora do SMLT, a Nortel também possui sua solução de *virtual switching*. Além de evitar loops na rede, o protocolo tenta redirecionar o acesso a partir de uma falha na rede, recuperando a rede em menos de um segundo na maioria dos casos. Trata-se de uma extensão, ou melhoria, do *link aggregation* onde existe o controle do estado do link. Com sua utilização se torna dispensável o uso do STP que além de ter o tempo de convergência maior, ainda realiza bloqueio de links de rede.

Para criação do *virtual switching*, é necessário um IST entre os switches sobre o qual trocam informações, permitindo a rápida detecção de falhas e modificação do caminho de encaminhamento.

Outra modalidade desse protocolo é R-SMLT (*Routed Split Multi-Link Trunking*). É uma extensão para camada de rede que utiliza protocolos de roteamento conhecidos, como OSPF, BGP e RIP.

O principal produto deste fabricante é o chassi ERS8600 que está na camada *core* da rede e realiza todo processamento da solução virtual.

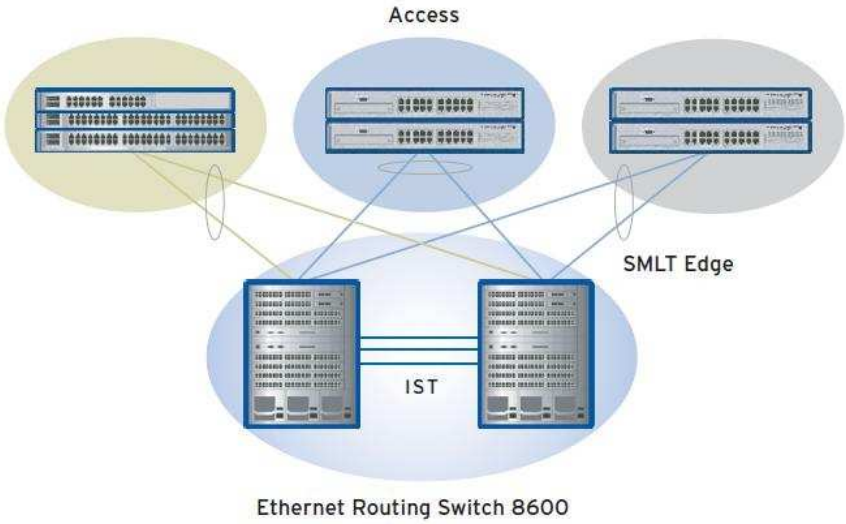


Figura 7 – Solução NORTEL SMLT

Na figura 7, vê-se um exemplo de aplicação da solução da Nortel. Todos os dispositivos da camada de acesso possuem conexões redundantes com a estrutura lógica. Em caso de falha em um dos links os switches de acesso continuam se comunicando com ambos os switches core.

Nesse exemplo o SMLT é aplicado nas camadas de acesso, garantindo todos os benefícios citados acima para servidores, aplicações VoIP, videoconferências e demais aplicações críticas.

Para criar esse mesmo artifício para as camadas core e distribuição, é necessário criar um R-SMLT entre os chassis, como mostra a figura abaixo:

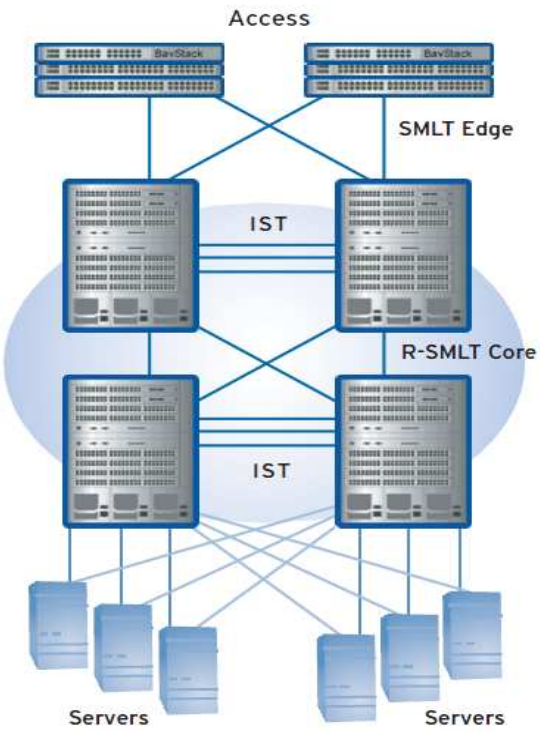


Figura 8 – Solução NORTEL R-SMLT

Dessa forma podemos visualizar um cenário mais abrangente onde os sites possuem 2 chassis cada e além de estenderem a LAN para dar mais versatilidade à operação, realizam um roteamento em comum aos 4 elementos da camada *core*.

3.3 CISCO VIRTUAL SWITCHING SYSTEM

A tecnologia VSS (*Virtual Switching System*) da Cisco não difere das demais na filosofia de criar uma entidade lógica a partir de 2 equipamentos físicos. Porém, apenas um elemento fica com o controle e ambos realizam o encaminhamento de pacotes, conforme mostra a figura 9. Em caso de falha de falha do chassi ativo, o *standby* entra em funcionamento. O modelo indicado para o VSS é o chassi 6500.

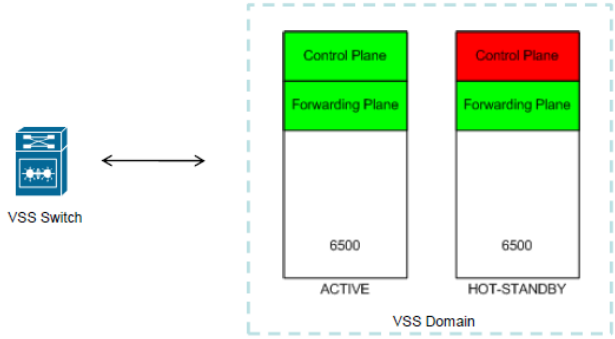


Figura 9– Controle do domínio virtual

O MEC (*Multichassis EtherChannel*) entra no lugar dos *spanning tree protocol* para evitar loops em caso de queda de conexões ativas. Com isso, elimina-se a necessidade de balanceamento de trafego de VLANs entre os equipamentos e eleição de switch *root*.

Uma premissa para o uso do VSS é a utilização de links de 10Gbps para comunicação entre os dois dispositivos. Com esses links são criadas as VSL (*Virtual Switch Link*), onde são agregados os canais 10 Gbps.

Então, o cenário mínimo para por em prática essa solução contempla dois chassis 6500, dois módulos de supervisão para o controle do VSS e dois módulos de switching, existentes com 8 ou 16 portas de 10Gbps, conforme mostra a figura 10.

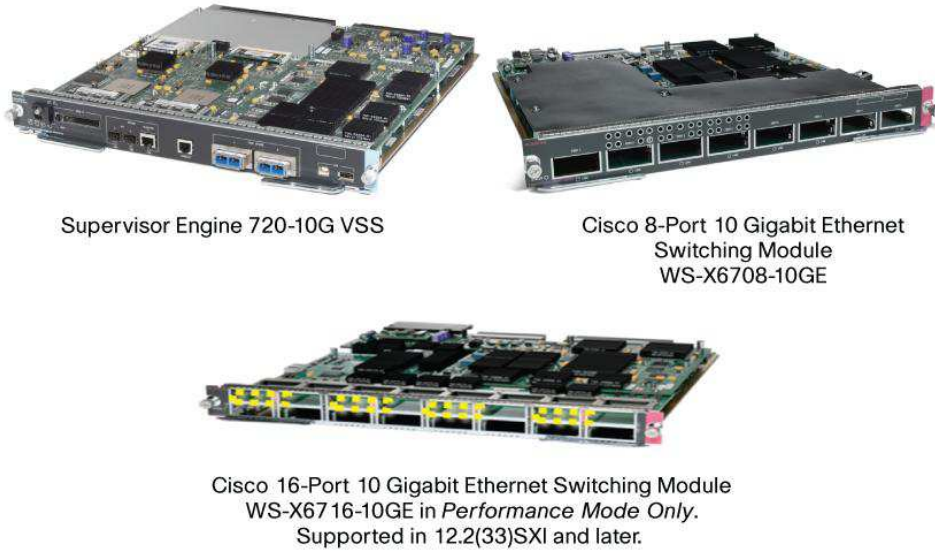


Figura 10 – Solução CISCO

Todo tráfego do VSL é encapsulado em um pacote com cabeçalho de 32 bytes. Esse cabeçalho é chamado de VSH (Virtual Switch Header) contendo:

- índice da porta de ingresso
- índice da porta de destino
- VLAN
- CoS

Pelo VSL passa tanto o tráfego de dados entre dois servidores que estão conectados em cada chassi, por exemplo, como também o tráfego de gerenciamento da solução que auxilia na formação da entidade lógica e controle do switch que está ativo.

Para garantir a redundância, existem os recursos NSF (*Nonstop Forwarding*) e SSO (*Stateful Switchover*), garantindo o menor tempo possível de inatividade no caso de falha. Enquanto o modulo supervisor ativo realiza todo controle do VSS e o encaminhamento do tráfego de dados, o módulo *standby* faz o sincronismo com o ativo na espera de alguma falha, além de encaminhar os pacotes destinados ao seu par.

Um cenário que pode ocorrer é falha geral dos links que formam o VSL. Esse cenário é chamado de *Dual Active*, ou seja, o conceito de switch standby não existe mais e os 2 switches passam a ficar ativos, respondendo para os dispositivos diretamente conectados em suas portas locais. Nesse caso os 2 switches ficam ativos e respondem pelo mesmo MAC address e endereço IP, o que pode causar instabilidade na rede.

Para evitar este cenário, deve-se configurar o VSL como *multiple-link port channel* em todos os mecanismos de supervisão e módulos disponíveis dentro do chassi.

4 **ESTUDO DE CASO**

4.1 INTRODUÇÃO

Todas as tecnologias avaliadas no capítulo anterior são soluções de grande difusão e confiabilidade. Todas atendem aos requisitos para se ter uma rede confiável, convergente e de alto desempenho. Em alguns momentos as soluções podem até parecer idênticas, porém alguns aspectos devem ser levados em consideração antes de realizar um investimento tão significativo que envolve os principais elementos da topologia da rede.

Dentre os tópicos apresentados como benefícios dessa nova forma de tratar uma rede corporativa, resolveu-se aplicar um cenário real e altamente crítico, em que essas soluções podem agregar valor ao negócio final da empresa, a partir de necessidades pesquisadas no dia-a-dia de sua operação.

A Cetip é um acrônimo para Central de Custódia e de Liquidação Financeira de Títulos. Originalmente uma sociedade civil sem fins lucrativos, foi criada em 1986, pelas instituições financeiras para preencher a lacuna de um sistema eletrônico de custódia e liquidação financeira no mercado de títulos privados. Suas atividades são regulamentadas pelo BACEN – Banco Central do Brasil. É um mercado de balcão organizado para registro da negociação de títulos e valores mobiliários de renda fixa.

Em 2008 a Cetip tornou-se sociedade anônima, passando a se chamar Cetip S.A. - Balcão Organizado de Ativos e Derivativos. É uma sociedade administradora de mercados de balcão organizados, ou seja, de ambientes de negociação e registro de valores mobiliários, títulos públicos e privados de renda fixa e derivativos de balcão. É, na realidade, uma câmara de compensação e liquidação sistemicamente importante, nos termos definidos pela legislação do SPB – Sistema de Pagamentos Brasileiro.

A Cetip é depositária principalmente de títulos de renda fixa privados, títulos públicos estaduais e municipais e títulos representativos de dívidas de responsabilidade do Tesouro Nacional. Na qualidade de depositária, a entidade processa a emissão, o resgate e a custódia dos títulos, bem como, quando é o caso, o pagamento dos juros e demais eventos a eles relacionados. Com poucas exceções, os títulos são emitidos escrituralmente, isto é, existem apenas sob a forma de registros eletrônicos (os títulos emitidos em papel são fisicamente custodiados por bancos autorizados). As operações de compra e venda são realizadas no mercado de balcão, incluindo aquelas processadas por intermédio do CetipNet (sistema eletrônico de negociação).

4.2 CETIP S.A. – QUADRO ATUAL

Sua rede corporativa é contemplada por três sites, sendo um na cidade São Paulo e dois na cidade do Rio de Janeiro, onde estão localizados seus sistemas. Para suporte e comunicação intra-sites, eles são interligados com links WAN redundantes, formando um triangulo.

Seus sistemas são acessados por instituições financeiras em duas redes privadas: RTM (Rede de Telecomunicações para o Mercado) e RSFN (Rede do Sistema Financeiro Nacional). Ambas possuem estrutura redundante de links WAN utilizando protocolo de alta disponibilidade.

O prédio principal, onde estão os sistemas de produção, opera diariamente no Centro da cidade, enquanto sua estrutura de contingência é localizada no bairro da Lapa. Entre seus sites existe uma SAN que os interconecta entre uma distância de mais ou menos 2 quilômetros. Essa SAN é utilizada na replicação síncrona para o site de backup, utilizando um link de fibra ótica apagada, fornecido por uma

operadora de âmbito nacional, multiplexado por DWDM. O protocolo transitado é o *Fibre Channel* básico, sem interação com qualquer rede IP. A razão dessa infraestrutura é manter um ambiente de alta disponibilidade para um eventual sinistro no prédio principal.

O perfil típico de aplicação da Cetip é o transacional em tempo real (OLTP). São replicadas bases de dados, filas de mensageria (que permitem sua comunicação com as instituições financeiras do mercado) e aplicações.

O site de backup possui um Data Center funcionando continuamente com a mesma estrutura do principal, com servidores esperando sua ativação em caso de emergência ou testes. A única diferença fica por conta dos funcionários que se localizam apenas no prédio principal. O acesso à contingência é feito remotamente através de links WAN ou com deslocamento de algum funcionário quando necessário.

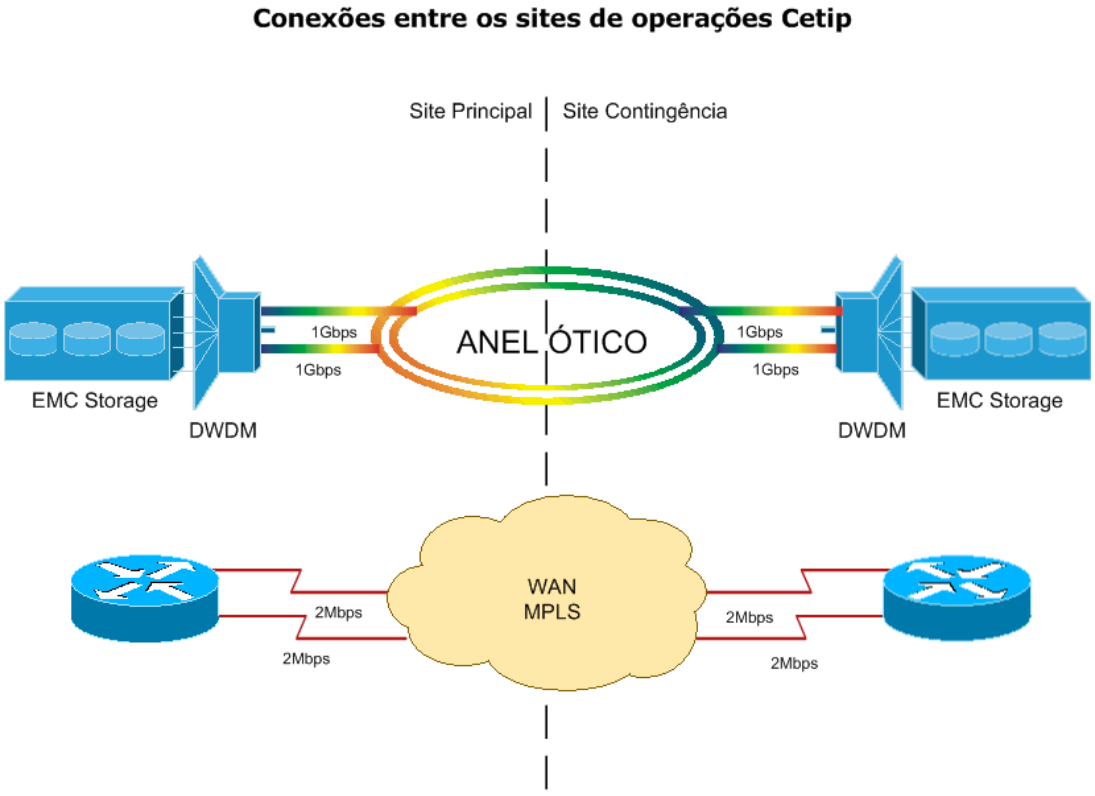


Figura 11 – Cetip Quadro Atual

A figura 11 ilustra a representação das interconexões entre os sites que abrigam os sistemas da Cetip S.A.

4.3 CETIP S.A. – MELHORIAS PROPOSTAS

Buscando trazer o máximo de benefícios possível da tecnologia *virtual switching*, foram avaliados os pontos positivos de cada fabricante em relação ao cenário acima descrito.

Em todas as tecnologias fatores como: simplificação, desempenho e versatilidade, atendem ao cenário hoje existente e poderiam ser aplicados em uma novo topologia da empresa. Porém dois pontos são bem importantes e fazem a diferença na hora de optar pela solução da Cisco. Nos dois casos a justificativa é o fato de todos os equipamentos que orbitam a rede da Cetip S.A. serem desse fabricante.

O fator custo seria o motivador principal para essa escolha, uma vez que poderiam ser realizados *trade ins* com equipamentos antigos, além de eventuais descontos de distribuidores pelo fato do parque ser 100% do mesmo fabricante.

Outro ponto bastante importante é a confiabilidade. A certeza de compatibilidade garante a interoperabilidade dos equipamentos. Switches e roteadores já existentes não teriam nenhum tipo de reconfiguração para se adaptar à nova solução, eliminando grande parte da fase de homologação da solução. Apesar de possuir protocolos proprietários, os principais fabricantes de servidores estão já adaptados aos padrões da Cisco difundidos no mercado e não existiria problema na coexistência de dispositivos da camada de acesso de outras marcas.

A figura 12 mostra o desenho da topologia proposta após aplicação do VSS da Cisco.

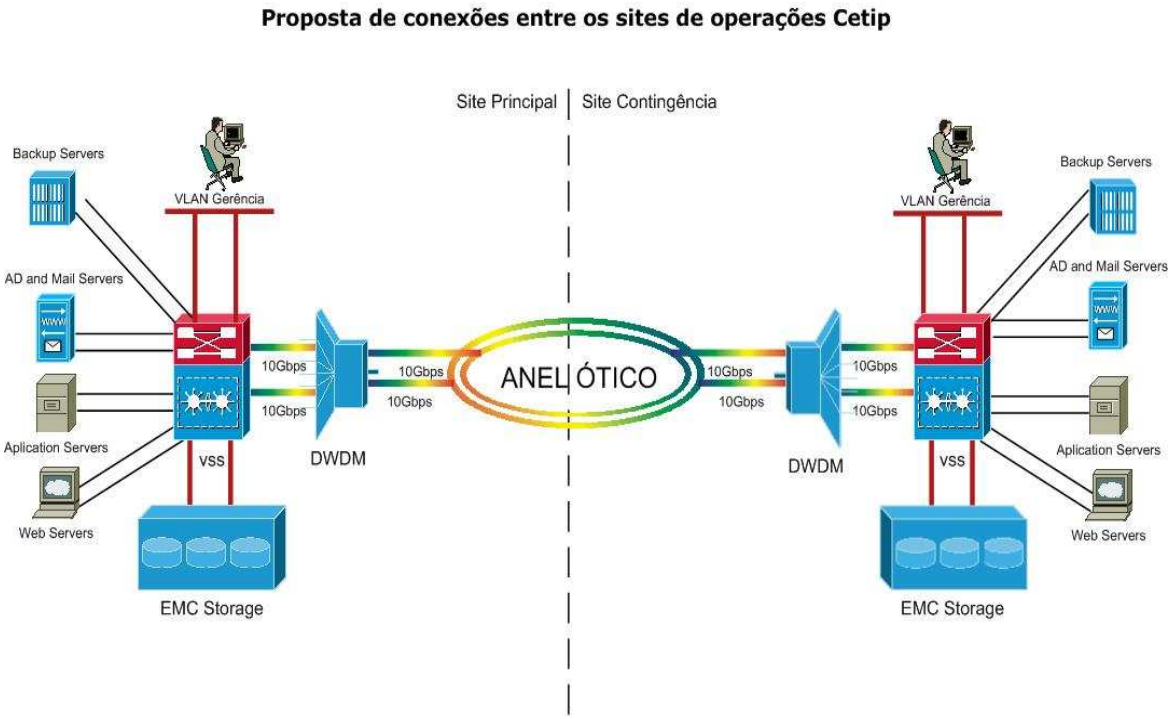


Figura 12 – Proposta Cetip

Essa proposta altera completamente o modo de utilização dos links de alta velocidade entre os sites. Com o aumento de 1Gbps para 10Gbps torna-se viável além da implementação do VSS, a eliminação dos links WAN da empresa. Já que existem links redundantes em anéis óticos que percorrem caminhos distintos entre as localidades, pode-se assumir que se tem um cenário de alta disponibilidade e desempenho.

Outro ganho considerável é passar a replicação entre *storages* para a camada de rede, podendo gerenciar o tráfego de maior volume e criticidade para a empresa.

Adicionamos a essa solução a prática de QoS, podendo restringir a largura de banda para o acesso de usuários do suporte aos servidores ou aos desenvolvedores de aplicações.

Outra funcionalidade bastante útil seria a unificação de LANs, como por exemplo, as redes de backup e de gerenciamento.

5 **CONCLUSÃO**

Além da grande reestruturação na infra-estrutura das empresas que a virtualização de servidores impôs, a área de *networking* também avançou nesse campo e acompanhou a tendência de resumir vários elementos de uma estrutura em um poderoso hardware, para oferecer vários ganhos de performance e consequentemente de negócio. O gerenciamento se torna mais fácil e a compreensão da rede também.

Neste trabalho foram apresentadas as principais tecnologias que envolvem cenários com redundância e tentam minimizar o *downtime* com alternativas de *failover* nas interligações entre os dispositivos. Primeiro os protocolos já consolidados e vastamente utilizados. Depois falou-se dos protocolos que envolvem as tecnologias de três dos principais fabricante do mercado.

Como os elementos envolvidos numa mudança de concepção da rede são dos mais diversificados, a opção melhor foi preservar a interoperabilidade dos equipamentos existentes, pois o parque da empresa escolhida para o estudo utiliza em quase sua totalidade equipamentos de rede Cisco, e também pelo fator de custo, uma vez que a reposição de equipamentos do mesmo fabricante facilitará a negociação.

Ao analisar cada solução, nota-se uma semelhança muito grande no conceito de extensão da LAN. Apenas os nomes são diferentes dentre padrões fechados ou abertos para o mercado. Na essência todos primam pela versatilidade do gerenciamento e alterações da rede.

O ponto principal que deve ser destacado, dentre a opção escolhida, é evitar ao máximo a situação de *dual-active* gerada por uma falha geral nos links que interligam os chassis. Para isso deve-se levar em conta além das configurações

recomendadas como boa prática pelo fabricante, a escolha das operadoras de telecomunicações dos links de 10Gbps e os caminhos diferentes que as fibras óticas percorrem entre o site principal e o de contingência.

De qualquer forma, a virtualização de switches preenche a necessidade que vários ambientes críticos possuem: poder operar em outro lugar de maneira mais transparente e ininterrupta possível. Isso tudo se torna viável graças à evolução da tecnologia ethernet, possuindo links cada vez mais velozes e confiáveis.

6 **REFERÊNCIAS**

[1] WEBB, Karen. Building Cisco Multilayer Switched Networks, v. 2, ed. Indianapolis: Pearson Education, 2003

[2] SOARES, Luiz Fernando Gomes; LEMOS, Guido; COLCHER, Sérgio. Redes de Computadores, v. 2ª ed. Rio de Janeiro: Campus, 2000

[3] Multiservice Switching Forum. Disponível em: <http://www.msforum.org/>. Acesso em: 17 ago. 2010.

[4] Virtual Switching. Disponível em: <http://www.networkdictionary.com/>. Acesso em: 17 ago. 2010.

[5] Cisco Virtual Switch Architecture. Disponível em: <http://www.cisco.com/>. Acesso em: 29 ago. 2010.

[6] Juniper Virtual Switch Architecture. Disponível em: <http://www.juniper.com/>. Acesso em: 29 ago. 2010.

[7] Nortel Virtual Switch Architecture. Disponível em: <http://www.nortel.com/>. Acesso em: 29 ago. 2010.

[8] Nortel Technical Brief. Disponível em: <http://www.nortel.com/>. Acesso em: 29 ago. 2010.

[9] Juniper Products and Services Literature. Disponível em: <http://www.juniper.com/>. Acesso em: 14 set. 2010.

[10] Cisco Catalyst 6500 Series. Disponível em: <http://www.cisco.com/>. Acesso em: 14 set. 2010.

[11] Split Multi-Link Trunking Frequently Asked Questions. Disponível em: <http://www.trcnetworks.com/>. Acesso em: 22 set. 2010.

[12] Virtual Switching System Q&A. Disponível em: <http://www.cisco.com/>. Acesso em: 14 set. 2010.

[13] Sobre a Cetip. Disponível em: <http://www.cetip.com.br/>. Acesso em: 14 set. 2010.

[14] Cisco Virtual Switching System. Disponível em: <http://www.cciezone.com/>. Acesso em: 27 set. 2010.