

**LIGA DE ENSINO DO RIO GRANDE DO NORTE
FACULDADE NATALENSE PARA O DESENVOLVIMENTO DO RN
ESPECIALIZAÇÃO EM REDES DE COMPUTADORES**

ARMAZENAMENTO DE DADOS E PROTOCOLOS FIBRE CHANNEL

FRANCISCO RODRIGUES ROCHA

**NATAL/RN
2010**

FRANCISCO RODRIGUES ROCHA

ARMAZENAMENTO DE DADOS E PROTOCOLOS FIBRE CHANNEL

Monografia apresentada ao curso de Especialização em Redes de Computadores da Faculdade Natalense para o Desenvolvimento do Rio Grande do Norte como requisito parcial para a obtenção do título de Especialista em Redes de Computadores.

Orientador:

Aluízio Ferreira da Rocha Neto, M.Sc.

Co-orientador:

Carlos Henrique Grilo Diniz, M.Sc.

**NATAL/RN
2010**

FRANCISCO RODRIGUES ROCHA

ARMAZENAMENTO DE DADOS E PROTOCOLOS FIBRE CHANNEL

Monografia apresentada a Faculdade Natalense para o Desenvolvimento do Rio Grande do Norte e aprovada em ____ de _____ de _____, pela banca examinadora composta pelos seguintes membros:

M.Sc. Aluizio Ferreira da Rocha Neto (Orientador)

Membro examinador

Membro examinador

Natal/RN
2010

” O ser humano vivencia a si mesmo, seus pensamentos como algo separado do resto do universo - numa espécie de ilusão de ótica de sua consciência. E essa ilusão é uma espécie de prisão que nos restringe a nossos desejos pessoais, conceitos e ao afeto por pessoas mais próximas. Nossa principal tarefa é a de nos livrarmos dessa prisão, ampliando o nosso círculo de compaixão, para que ele abranja todos os seres vivos e toda a natureza em sua beleza. Ninguém conseguirá alcançar completamente esse objetivo, mas lutar pela sua realização já é por si só parte de nossa liberação e o alicerce de nossa segurança interior..”
- Albert Eistein -

Agradecimentos

Agradeço a Deus em primeiro lugar por ter-me concebido como pessoa humana.

Agradeço aos meus pais por terem me criado e dado um dos bens mais preciosos para a minha vida que foi a educação. Desde cedo entendi que eu precisava estudar.

Agradeço a minha querida e amada esposa Andrezza pela paciência e compreensão em entender a importância deste trabalho.

Agradeço a minha família e aos meus filhos Luiz Gustavo e Pedro Henrique, pequenos ainda, pela paciência de me deixar digitar o trabalho quando queriam brincar.

Agradeço ao professor orientador M.Sc. Aluizio Ferreira da Rocha Neto, primeiro por coordenar o curso de Pós Graduação em Redes de Computadores de forma harmoniosa e cuidadosa para os professores e alunos do curso e segundo por me ajudar na escolha do trabalho, por me orientar e dar-me a oportunidade de aprender mais um pouco. O professor Aluizio se assemelha a figura de um Pai cuidando de maneira zelosa pelos seus alunos.

Agradeço ao professor co-orientador M.Sc. Carlos Henrique Grilo Diniz pela valiosa ajuda no âmbito do trabalho na JFRN, sempre disposto a auxiliar, esclarecendo dúvidas, fornecendo recursos para melhor elaboração do trabalho.

Agradeço a todos os professores que participaram deste curso, pela disseminação das idéias, assuntos e conceitos relacionados, fornecendo subsídios para pesquisas futuras e trabalhos na área de redes de computadores.

Agradeço ao amigo e consultor técnico José Roberto pelas valiosas explicações e dissolução de algumas dúvidas a respeito de assuntos técnicos e tecnologias explanados neste trabalho.

Ficha catalográfica

RESUMO

Este trabalho versa sobre as tecnologias de armazenamento das informações em redes com ênfase na tecnologia que usa os protocolos *Fibre Channel* (canal de fibra). Será apresentada uma visão prática de uma rede SAN de alta disponibilidade.

Palavras-chave: armazenamento, redes, protocolos, Fibre Channel, SAN, alta disponibilidade.

ABSTRACT

The article in question deals with the technologies for storage of information in networks with emphasis on technology that uses the protocols Fibre Channel (fiber channel). Will show a practical vision of a SAN network for high availability.

Key-words: storage, networks, protocols, Fibre Channel, SAN, high availability.

SUMÁRIO

1	INTRODUÇÃO.....	12
1.1	OBJETIVOS.....	12
1.1.1	Objetivos Específicos	12
1.2	MOTIVAÇÃO.....	13
2	REFERENCIAL TEÓRICO.....	14
2.1	DIRECT ATTACHED STORAGE.....	14
2.2	NETWORK ATTACHED STORAGE	17
2.3	STORAGE AREA NETWORK.....	19
2.4	RAID	22
2.5	PROTOCOLOS FIBRE CHANNEL	25
2.5.1	Camadas do Modelo Fibre Channel	26
2.5.2	Classes de Serviços.....	32
2.5.3	Topologias da Arquitetura do <i>Fibre Channel</i>	34
2.5.4	<i>Fabric Zoning</i>	37
2.5.5	<i>Fabric</i> - Topologia Física.....	40
2.6	CONVERGÊNCIA DE REDES.....	42
3	VISÃO PRÁTICA DE AMBIENTE SAN DE ALTA DISPONIBILIDADE	48
4	CONCLUSÃO.....	60

REFERÊNCIAS BIBLIOGRÁFICAS

LISTA DE FIGURAS

Figura 1: Tecnologia <i>Direct Attached Storage</i>	15
Figura 2: Exemplo de Unidade eSATA da Micronet.	16
Figura 3: Arquitetura <i>Network Attached Storage</i>	17
Figura 4: Arquitetura de uma SAN integrada a uma LAN.....	20
Figura 5: Modelo de Camadas do Fibre Channel.	26
Figura 6: Quadro de uma arquitetura <i>fibre channel</i>	28
Figura 7: Hierarquia dos pacotes de dados do FC.	30
Figura 8: Exemplo de utilização das portas no <i>Fibre Channel</i>	34
Figura 9: Topologia <i>point-to-point</i> do FC.	35
Figura 10: Topologia <i>arbitred loop</i> do FC.	36
Figura 11: Switch IBM B16 da JFRN e suas conexões em fibra óptica.	37
Figura 12: Visão frontal do switch IBM B16 pela interface web.....	37
Figura 13: Zoneamento.....	38
Figura 14: Endereçamento de portas no fabric.....	40
Figura 15: Topologia <i>core-edge fabric</i>	41
Figura 16: Topologia <i>mesh parcial</i>	41
Figura 17: Topologia <i>mesh total</i>	42
Figura 18: Tecnologia FCIP.	43
Figura 19: Tecnologia iFCP.	44
Figura 20: Tecnologia FCoE.	44
Figura 21: Representação das pilhas de protocolos OSI, FCoE e FC.	45
Figura 22: Exemplo de ligação iSCSI em rede SAN-IP.....	46
Figura 23: Comparação entre as pilhas iSCSI e FCoE.....	47
Figura 24: <i>Roadmap</i> panorama das velocidades no FC.	47
Figura 25: Rede SAN JFRN.	49
Figura 26: SFP 4Gbps e fibra óptica multimodo de padrão LC.	50
Figura 27: Topologia mesh total – Fabric SAN JFRN.	50
Figura 28: Interface web de gerenciamento do switch B32SW01 aba <i>alias</i>	52
Figura 29: Interface web de gerenciamento do switch B32SW01 aba <i>zone</i>	53
Figura 30: Interface web de gerenciamento do switch B32SW01 aba <i>zone config</i>	54
Figura 31: Storage IBM DS4700 e unidades de expansão Exp810.....	55
Figura 32: Interface de gerenciamento do sistema de <i>storage</i> aba <i>physical</i>	56
Figura 33: Interface de gerenciamento do sistema de <i>storage</i> aba <i>logical</i>	57
Figura 34: Interface de gerenciamento do sistema de <i>storage</i> aba <i>mappings</i>	57
Figura 35: Rack com <i>Tape library</i> IBM ULT3582, Servidor TSM e <i>Tape library</i> TS3310....	59

LISTA DE TABELAS

Tabela 1: Interfaces físicas e suas velocidades.....	16
Tabela 2: Principais níveis de RAID.	24
Tabela 3: Cabeçalho do quadro do <i>fibre channel</i>	28
Tabela 4: Comparação do modelo de camadas OSI com FC.	31
Tabela 5: Tipos de portas utilizadas na conexão entre dispositivos na SAN.	34
Tabela 6: WWN.	39
Tabela 7: Novo esquema de WWN.	39
Tabela 8: Dispositivos da rede SAN <i>fibre channel</i> da JFRN.	48

1 INTRODUÇÃO

Com a diminuição da escrita em papel e o uso constante do computador, aliada à globalização das informações pelo mundo todo, houve a necessidade de armazenar estas informações com segurança e disponibilidade.

Inicialmente, no modelo convencional, o armazenamento das informações se fazia de maneira modesta através de dispositivos de armazenamento ligados diretamente aos servidores. Normalmente essa ligação se fazia através de cabos de cobre e não existia a filosofia de redes. Então um único servidor poderia acessar os dados que estivessem naquele dispositivo, ligado a ele, sem garantia de disponibilidade, nem flexibilidade, nem escalabilidade, com um mínimo de segurança dessas informações.

Precisou-se investir em infra-estrutura para sistemas de armazenamento, precisou-se investir em mídias com grandes capacidades de armazenamento e em redes que ligassem esses dispositivos aos seus servidores agora fornecendo segurança da informação, disponibilidade para os sistemas operarem 24 horas por dia e 365 dias por ano, e sobretudo, eficiência no transporte dos dados requerendo uma rede mais versátil que pudesse ligar uma gama de servidores aos dispositivos de armazenamento.

1.1 OBJETIVOS

Neste trabalho, serão abordadas as tecnologias mais empregadas como o *Direct Attached Storage* (DAS), *Network Attached Storage* (NAS) e o *Storage Area Network* (SAN) com ênfase na rede SAN e nos protocolos *fibre channel* que implementam alta performance e alta disponibilidade. No final será abordada uma visão prática de uma rede SAN em utilização para exemplificar melhor este trabalho.

1.1.1 Objetivos Específicos

Para atingir o objetivo geral deste trabalho, as seguintes etapas foram seguidas:

- Estudar os conceitos das tecnologias de armazenamento de dados;
- Apresentar as tecnologias mais utilizadas hoje em dia no armazenamento de dados com a melhor utilização dos recursos envolvidos;
- Exemplificar através de figuras, tabelas e fotos, os protocolos de transporte de dados, as tecnologias empregadas e mostrar como os dispositivos se conectam para formar as redes de armazenamento das informações.

- Apresentar recursos e tecnologias consolidadas para a manutenção da base de informações guardadas nas unidades de armazenamento e o emprego de recursos e dispositivos redundantes.
- Descrever os protocolos fibre channel e suas funcionalidades no transporte dos dados.
- Apresentar as novas soluções usando os protocolos já existentes (IP, FCP, SCSI) para convergência das redes SAN e Ethernet;
- Apresentar uma visão prática da rede SAN-FC da Justiça Federal do RN(JFRN)
- Descrever os dispositivos e conexões físicas da rede SAN-JFRN;
- Mostrar o sistema de redundância dos dispositivos de armazenamento e formas de backup das bases de dados e aplicações;
- Apresentar modificações que podem ser feitas na rede atual SAN-JFRN para melhorar os caminhos redundantes, armazenar melhor as fitas de backup e aumentar a eficiência no transporte dos dados.

1.2 MOTIVAÇÃO

Falar das redes de armazenamento de dados, mais precisamente da rede baseada em protocolos fibre channel atraiu a minha atenção, a princípio, em virtude do conhecimento que tive ao participar da montagem física da rede SAN da JFRN.

A conclusão do curso de Pós Graduação em redes pela FARN, com os conhecimentos obtidos nas disciplinas, na busca de um tema que tivesse sido pouco explorado para o trabalho de final de curso, fui ajudado pelo professor Aluizio a escolher esse assunto.

A necessidade de conhecer um pouco melhor como é feito o transporte dos dados pela rede SAN, os protocolos utilizados nesse transporte, a questão da redundância dos dispositivos de armazenamento, para melhorar o conhecimento e poder interagir melhor no gerenciamento da SAN-JFRN.

Necessidade de procedimentos e iniciativas para melhorar a performance da SAN-JFRN, me interessaram a estudar um pouco mais sobre esse assunto.

2 REFERENCIAL TEÓRICO

Neste capítulo serão apresentadas os conceitos e teorias dos principais sistemas de armazenamento disponíveis no mercado.

Os sistemas de armazenamento de dados utilizados hoje em dia nas grandes organizações e instituições governamentais requerem integridade, disponibilidade e flexibilidade de suas informações. As tecnologias mais empregadas são: *Direct Attached Storage*, *Network Attached Storage* e o *Storage Area Network* cujas especificidades são descritas a seguir.

2.1 DIRECT ATTACHED STORAGE

Segundo (MORIMOTO, 2007), abreviadamente DAS se refere ao tipo de ligação onde o computador ou o servidor é diretamente conectado ao dispositivo ou dispositivos de armazenamento externos. Como exemplo, na figura 1, vê-se um servidor ligado ao conjunto de discos, para armazenamento de dados, denominado storage. Observa-se que entre eles não existe nenhum dispositivo comutador (switch) ou interfaces que possam gerar cabeçalhos desnecessários ao transporte dos dados. Se forem usados comutadores entre os dispositivos na rede DAS, estes não podem contribuir com o aumento dos cabeçalhos de transporte e devem permitir uma ligação ponto a ponto entre as entidades comunicantes, sem compartilhamento com outros servidores. É uma das características do DAS a alta eficiência no transporte dos dados.

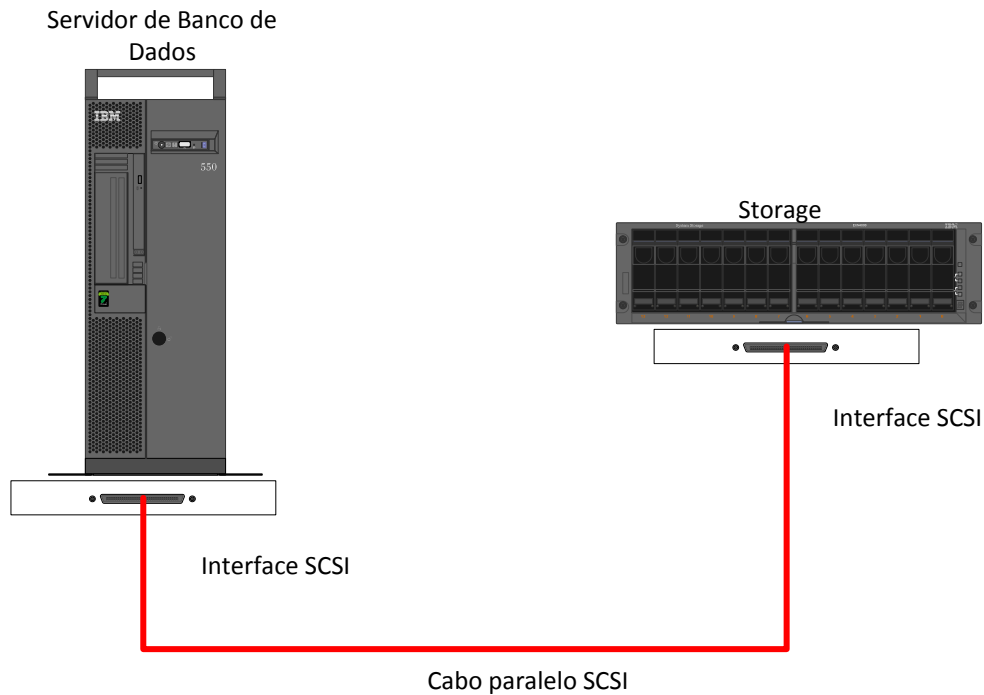


Figura 1: Tecnologia *Direct Attached Storage*.

Essa ligação direta é feita através de uma conexão que pode ser do tipo SCSI (*Small Computer Systems Interface*), FC (*Fiber Channel*), SATA (*Serial Advanced Technology Attachment*), eSATA (*external SATA*), etc. Estes padrões, juntamente com os seus protocolos, são os mais usados devido à alta velocidade de transmissão dos dados que os mesmos oferecem e baixo overhead, em alguns casos. Em microcomputadores comuns, uma modalidade de DAS bastante evidente é a conexão do próprio HD por exemplo do tipo SATA (serial ATA) com cabo e interfaces do mesmo padrão. A tabela 1 mostra as velocidades empregadas nas interfaces mais comuns.

<i>Tipo de Conexão</i>	<i>Velocidade(MBps)</i>	<i>Largura do Barramento (bits)</i>	<i>Máximo de dispositivos suportados</i>
SCSI-1	5	8	8
SCSI-2	10	8	8
Wide SCSI-2	20	16	16
Ultra SCSI	20	8	8
Wide Ultra SCSI-3	40	16	16
Ultra2 SCSI	40	8	8
Wide Ultra2 SCSI	80	16	16
Wide Ultra3 SCSI	160	16	16
Ultra320 SCSI	320	16	16
USB 1.0	12		
USB 1.1	12		

USB 2.0	60		127
USB 3.0	600		
SSA *	80		
Blu-Ray (2X – 72Mbit/s)	16		
FireWire 800(IEEE1394)	100		63
SAS*	300		128-16000
PATA (parallel ATA)	133		
SATA (Serial ATA)	150-600		
FC (fiber channel)	200-400-800		224

Tabela 1: Interfaces físicas e suas velocidades.

*SSA – Serial Storage Architecture; SAS – Serial Attached SCSI

Na modalidade *Direct Attached Storage* apenas o servidor ligado ao *storage* tem acesso aos dados. Outros servidores somente terão acesso a esses dados através desse servidor. Se ocorrer qualquer problema com esse servidor, os dados ficarão inacessíveis até que o mesmo volte a operar.

A escalabilidade e a alta disponibilidade são limitadas na topologia do DAS embora o desempenho seja bastante alto uma vez que os equipamentos envolvidos se ligam diretamente e utilizam o mesmo protocolo (SCSI, FC, etc.) para encapsular os pacotes, transmiti-los pela rede e desencapsulá-los, ou seja o *overhead* (cabeçalho a mais no encapsulamento do pacote de dados) é baixo.



Figura 2: Exemplo de Unidade eSATA da Micronet.

Um fator que impede o amplo uso dos storages na tecnologia DAS, em relação às outras tecnologias, é a limitação do número de máquinas servidoras ligadas aos mesmos.

Os *Storages* (unidades de armazenamento) contribuem hoje em dia, de forma significativa, para aumentar a área útil dos servidores ligados a ele, seja na forma de um punhado de discos (JBOD – *Just a Bunch Of Disks*), seja na forma de volumes de discos (LV – *logical volumes*) com controladoras para realização de RAID (*Redundant Array of Independent Disks*) e interfaces gráficas para manipulação desses discos de alta performance.

2.2 NETWORK ATTACHED STORAGE

Segundo (Brooks, Byrne, Higuera, Krax, Kuo, 2006), abreviadamente NAS é a arquitetura de armazenamento e compartilhamento de dados composta por um conjunto de dispositivos armazenadores (*storages*) que se ligam aos servidores através de uma infraestrutura de rede usando algum protocolo. Como exemplo, a figura 3 mostra a concepção de um ambiente NAS. Uma diferença em relação ao DAS é a presença do comutador (*switch*) não implicando de regra que o DAS não possa usar *switches* em suas conexões.

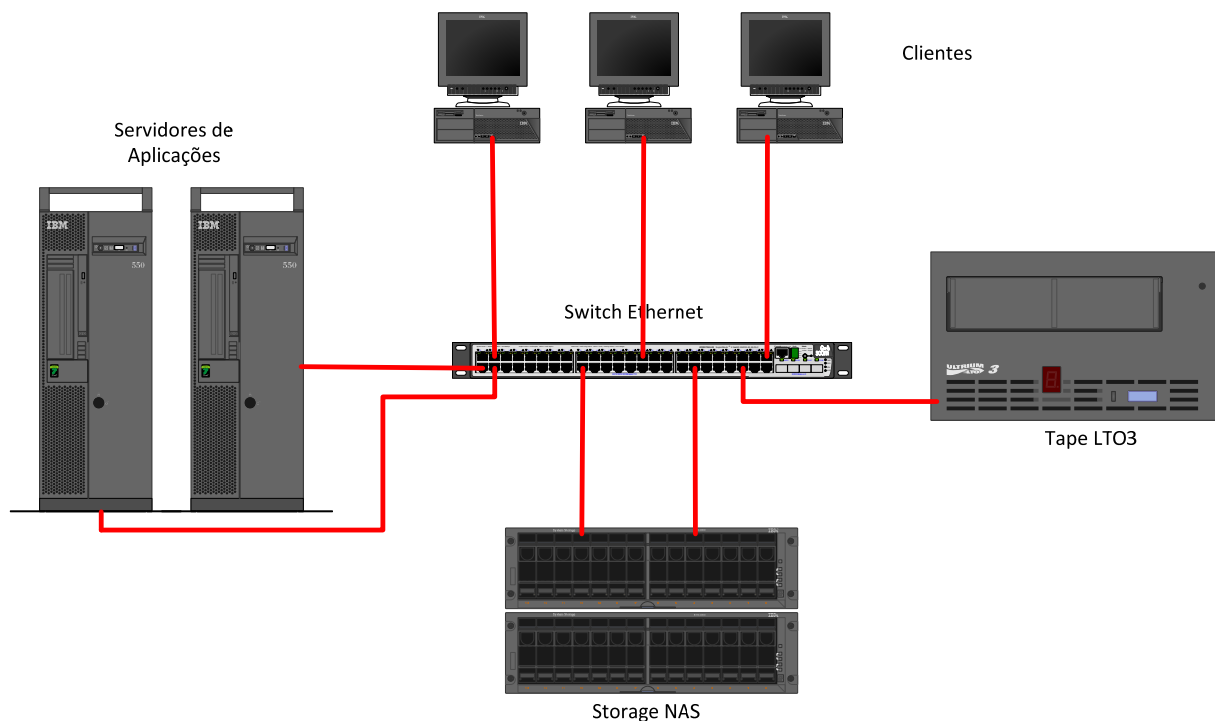


Figura 3: Arquitetura Network Attached Storage.

Nas dependências da infra-estrutura de transporte de dados, o NAS usa a rede local LAN (*Local Area Network*) e os protocolos da rede Ethernet ou Gigabit Ethernet além dos protocolos TCP/IP¹ para o transporte dos dados. Para grandes distâncias pode ser usada a rede WAN (*Wide Area Network*) baseada também nos protocolos Ethernet.

Além dos protocolos de rede, a tecnologia NAS usa também os protocolos de compartilhamento de arquivos, evidenciando a maior função desta tecnologia, estabelecendo melhor escalabilidade e disponibilidade dos dados aliada à eficiência no gerenciamento das informações. Todos os servidores ligados à rede agora têm amplo acesso à gravação e leitura dos dados, ao sistema de backup e podem fornecer aos clientes os seus serviços de forma mais eficiente.

Compartilhamento de arquivos requer provisão de arquivos pelo servidor, acesso a esses arquivos remotamente pelo cliente e gerenciamento desses arquivos. Os protocolos envolvidos neste processo são comumente o Network File System (NFS) e o *Common Internet File System* (CIFS).

O NFS é o sistema de arquivos, comumente usado pelos sistemas operacionais UNIX, que foi desenvolvido originalmente em 1984, pela Sun Microsystem, permitindo a um servidor ou computador remoto acessar áreas de disco em outro servidor como se fosse localmente. Para isso, servidor e cliente se comunicam através de RPC (Chamada Remota de Procedimentos).

A empresa Sun Microsystems lançou três versões do protocolo NFS onde o NFSv1 - leia-se NFS versão 1 - somente foi destinado ao uso interno da Sun. O NFSv2 foi lançado em 1989 e operava inicialmente em cima do protocolo UDP (*User Datagram Protocol*) especificado pela RFC 1094. RFC quer dizer *Request for Comments* - é um documento que define de maneira rigorosa os detalhes e especificações dos protocolos de Internet. O NFSv2 tem como característica também ser *stateless* ou seja, não guarda o estado das conexões entre o cliente e o servidor. Já o NFSv3 foi especificado pela RFC 1813 em 1995 e conta com suporte a arquivos de 64 bits de tamanho, uso do TCP/IP como protocolo de transporte, melhoria na eficiência do processo de escrita, dentre outras melhorias. O NFSv4 foi lançado em 2000 pelo *Internet Engineering Task Force* (IETF) e definido pela RFC 3010 com atualização dada pela RFC 3530 no ano de 2003. O NFSv4 trouxe melhorias no desempenho, melhorias na questão da segurança dos dados e passou a ser ao invés de *stateless*, *stateful*,

¹ TCP/IP: *Transmission Control Protocol/Internet Protocol*

conseguindo agora guardar estados das conexões no servidor. No começo deste ano, inovações aconteceram no modelo NFSv4 gerando a NFSv4.1 com intuito de dar suporte a sistemas em cluster e fornecer escalabilidade a arquivos distribuídos entre vários servidores.

Segundo (MORIMOTO, 2008; BARREIROS, 2010; EVANS, 2010), CIFS (*Common Internet File System*) é o sistema de arquivos usado pelos sistemas operacionais Windows. Normalmente quando se fala em CIFS, fala-se também no protocolo SMB (*Server Message Block*).

Em 1988 a empresa Microsoft juntamente com a empresa Intel, lançaram o protocolo SMB Core que foi amplamente reformulado nos anos seguintes dando origem ao CIFS. O CIFS/SMB é um protocolo de camadas superiores dependendo então do protocolo de transporte TCP juntamente com o NetBios originando o protocolo NBT.

O NetBios, que significa *Network Basic Input/Output System*, é uma API (*Application Programming Interface*) criada pela empresa International Business Machines Corporation (IBM) em 1984 para prover de forma simples o compartilhamento de arquivos na rede local. Em 1985 foi criado o protocolo *Enhanced User Interface* (NetBEUI) integrado ao NetBios que permite que cada máquina possa se estabelecer na rede com um nome composto por até quinze caracteres.

Os protocolos NFS e CIFS representam atualmente o padrão das requisições de acesso a dados remotos dos servidores para intranet e internet tornando o NAS uma tecnologia bastante flexível.

2.3 STORAGE AREA NETWORK

Segundo (LONG, 2006; COSTA, 2008), abreviadamente SAN é uma das tecnologias mais versáteis na transmissão dos dados onde o acesso às informações é realizado com alto desempenho e utiliza uma infra-estrutura de rede própria fundamentada nas conexões de fibra óptica e protocolos FC (*fibre channel*). Como exemplo, a figura 4 mostra a concepção de um ambiente SAN.

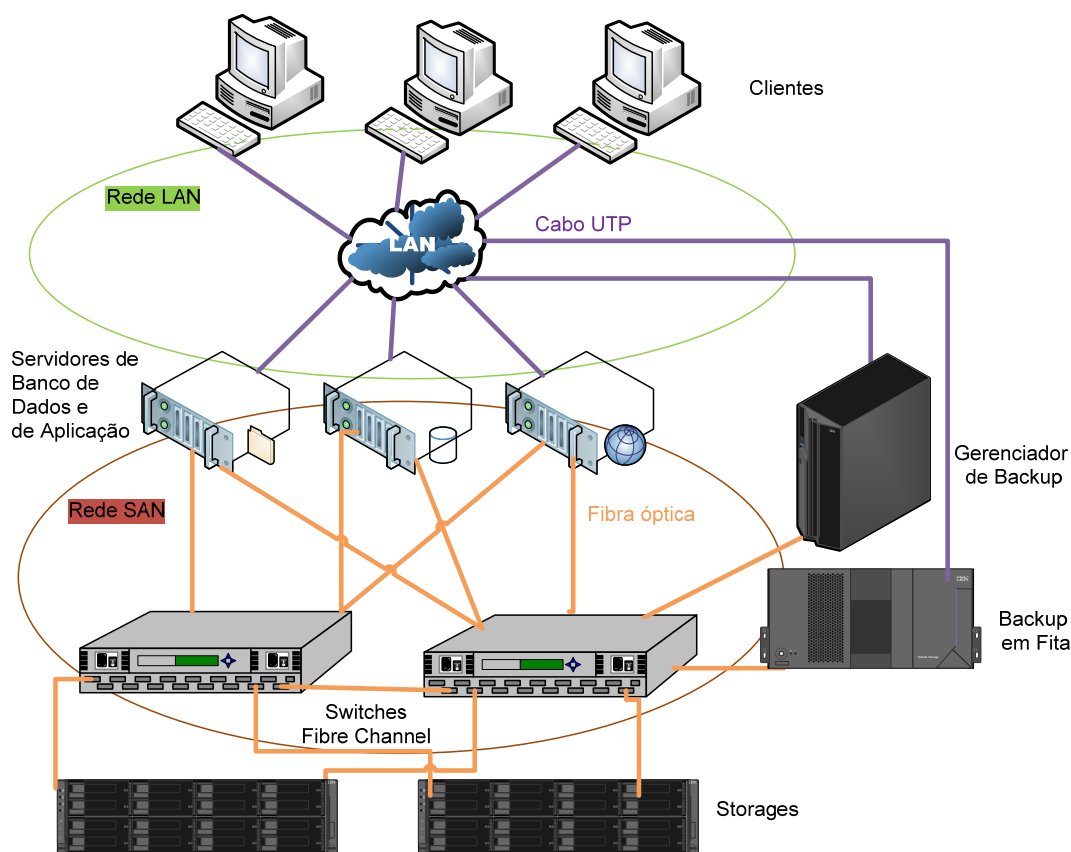


Figura 4: Arquitetura de uma SAN integrada a uma LAN.

Contextualizando de forma simples: um cliente que está na rede LAN usa indiretamente a rede SAN, de forma transparente, conforme a figura acima mostra. Através do seu cabeamento Ethernet, usando o protocolo TCP/IP, o cliente interage com um servidor de aplicação enviando o seu pedido de requisição de informações que pode ser, por exemplo, uma consulta ao banco de dados. Este servidor de aplicação se comunica com o servidor de banco de dados, que por sua vez, irá usar os benefícios da rede SAN. O acesso à rede SAN se faz então entre o servidor de banco e o *storage* que estão numa infra-estrutura de rede diferente da rede LAN composta por *switches* e conexões de fibra óptica. As configurações, tabelas e dados requisitados do banco estarão no *storage* e o transporte desses blocos de dados ao servidor de banco se dará através de um conjunto de protocolos *fibre channel* que viabiliza o alto desempenho segundo o qual essa rede pode operar.

O modelo SAN é um sistema de armazenamento compartilhado onde os dados são transportados em forma de blocos de dados diferentemente do modelo NAS onde os dados

são transportados em forma de arquivos. Para isso a SAN utiliza a rede de protocolos SCSI ou de protocolos FC enquanto o NAS utiliza a rede Ethernet ou Gigabit Ethernet.

Em termos de infra-estrutura de rede, a tecnologia *fibre channel* pode ser concebida para utilizar a própria estrutura cabeada Ethernet ou Gigabit Ethernet para o transporte dos dados, bem como a rede pode ser baseada em fibra óptica, com os equipamentos ativos inerentes a cada um desses modelos. Ela pode funcionar tanto via protocolos *fibre channel*, como em conjunto com o TCP/IP para o transporte dos dados. Estudos devem ser feitos dentro da empresa ao se implementar qualquer uma dessas concepções no intuito de minimizar custos ou priorizar alto desempenho. Algumas implicações de desempenho e TCO (Custo Total de Propriedade) devem ser levadas em conta sabendo, se o que deve ser priorizado, são requisições de I/O (entrada e/ou de saída) ou apenas leitura do banco de dados aleatoriamente.

Uma das grandes vantagens da rede SAN é o sistema de backup bastante funcional, com extrema rapidez e flexibilidade no processo de guarda e recuperação dos dados. Hoje em dia, robôs de backup ou *tapes library* fazem todo o trabalho de gravação das informações em fitas magnéticas de alto desempenho e que podem alcançar a faixa dos terabytes com compressão de dados. Nesse sistema de backup, normalmente existe um software gerente responsável pelas políticas de armazenamento que coordena os trabalhos entre os clientes e o robô de backup gerenciando também um conjunto de regras que estabelecem o tempo de retenção e quantidade de versões que os dados devem ocupar nas fitas. Todo o trabalho de gravação dos dados é feito de forma automatizada cumprindo o ritual proferido pelos scripts de backup e conforme as regras estabelecidas. A intervenção humana na recuperação dos dados somente deverá ser feita em caso de desastre nos equipamentos de armazenamento, perdas de dados, apagamentos de dados pelos clientes, etc.

O transporte dos dados do servidor de aplicação até o servidor de backup pode ser feito pela rede Gigabit Ethernet ou de forma mais eficiente pela rede SAN utilizando a fibra óptica. Na prática um cabeamento categoria 5e ou 6 pode transmitir os dados a Gigabit por apenas 100 m enquanto pela rede de fibra podemos chegar a casa dos 15 km.

Da mesma forma que o DAS e o NAS, a tecnologia SAN prevê o uso de funcionalidades de segurança da informação a nível das controladoras do *hardware* e dos dispositivos armazenadores no que se refere a proteção dos dados. Esta segurança se baseia na replicação dos dados feita localmente pela modalidade RAID ou remotamente pela cópia das informações para um equipamento secundário.

2.4 RAID

Segundo (Brooks, Byrne, Higuera, Krax, Kuo, 2006), RAID (*Redundant Array of Independent Disks*) é uma tecnologia que agrupa um conjunto (*array*) de discos normalmente de mesmo tamanho e velocidade numa matriz que confere maior grau de tolerância a falhas viabilizando a questão da segurança dos dados armazenados. Esta tecnologia foi criada na Universidade da Califórnia em Berkeley (EUA) por David A. Patterson, Garth A. Gibson e Randy H. Katz no final da década de 1980. Há vários níveis de RAID a serem empregados dependendo do grau de confiabilidade que se queira adotar ou se a pretensão maior é performance de I/O(input/output - entrada/saída) dos dados. Os padrões mais usados são os seguintes:

- **RAID 0** – Distribui os dados de maneira uniforme em todos os discos. Não há redundância das informações nem tolerância a falhas mas promove a melhor performance de I/O. Usa a totalidade dos discos do storage para o armazenamento dos dados, contribuindo para uma solução mais econômica.
- **RAID 1** – Distribui os dados simultaneamente para dois conjuntos de discos. Este padrão é normalmente chamado de espelhamento(mirror). Há redundância de informações e pode ser perdido até um disco no array (coleção de discos) sem causar parada dos sistemas. Realiza um acesso a leitura mais rápido devido a simultaneidade de leitura dos discos. Oferece um custo mais alto uma vez que a capacidade total de armazenamento é a quantidade de discos no storage dividido por dois,ou de outra forma, sincronismo de dados sempre duplica o gasto com discos. Normalmente empregado quando se exige desempenho e alta disponibilidade dos dados.
- **RAID 3** – Os dados são distribuídos nos vários discos sincronamente e paralelamente em faixas chamadas stripes(listras). Cada pedaço da informação, lista de dados, é escrito na mesma posição física de cada disco e paralelamente em discos diferentes. A redundância da informação concentra-se nas informações de paridade a serem escritas para cada stripe de dados em um disco dedicado, e desse modo, os dados podem ser reconstruídos se um dos discos falhar. A atuação do padrão RAID 3 se destaca na requisição de grandes quantidades de informações aos discos envolvidos porque cada unidade destas tem a sua contribuição nesse sistema uníssono. É melhor empregado em operações que processam grande

volume de dados como arquivos de imagem, técnicas gráficas computadorizadas, ou seja, aplicações que trabalham com grandes blocos de dados.

- **RAID 5** – Os dados são guardados de maneira uniforme pelos discos do array juntamente com as informações de paridade. A paridade é processada através dos algoritmos de cálculos matemáticos complexos usados pela controladora. Diferentemente do RAID 3, a paridade não é mais gravada em um disco dedicado, ela é disseminada em stripes próprios em cada disco. Permite apenas a perda de um disco no total de discos do array o que sugere um custo menor de implementação. Oferece um desempenho bom em relação aos outros padrões. Se caso um dos discos falhar, apesar das aplicações inerentes não pararem, haverá uma possível degradação no sistema. Requer um mínimo de três discos e o total da capacidade do conjunto é a quantidade de discos menos um. O RAID 5 se adequa melhor a ambientes onde os acessos aleatórios a pequenos blocos de dados são constantes.
- **RAID 1+0** – É também chamado de RAID 10. É uma junção dos RAID 1 e RAID 0, ou seja, realiza inicialmente um espelhamento dos dados e logo em seguida distribui esses dados em faixas(stripes) de dados em todo o array. É preciso pelo menos quatro discos para implementar este padrão e metade dos discos fica comprometido com a questão da redundância. É a modalidade destinada aos sistemas que requerem alta disponibilidade agregada ao bom desempenho. A capacidade total de discos fica reduzida pela metade o que leva a um custo mais alto de implementação. Ex.: sistemas de bancos de dados usam esse padrão por ser mais performático.
- **RAID 0+1** – Realiza inicialmente o stripe dos dados e logo após o espelhamento. Apresenta uma reconstrução de RAID mais complexa que o RAID 10 e custo e capacidade total semelhantes. Requer pelo menos quatro discos para ser implementado.

A tabela 2 mostra algumas figuras dos principais níveis de RAID para ilustrar os diversos casos.

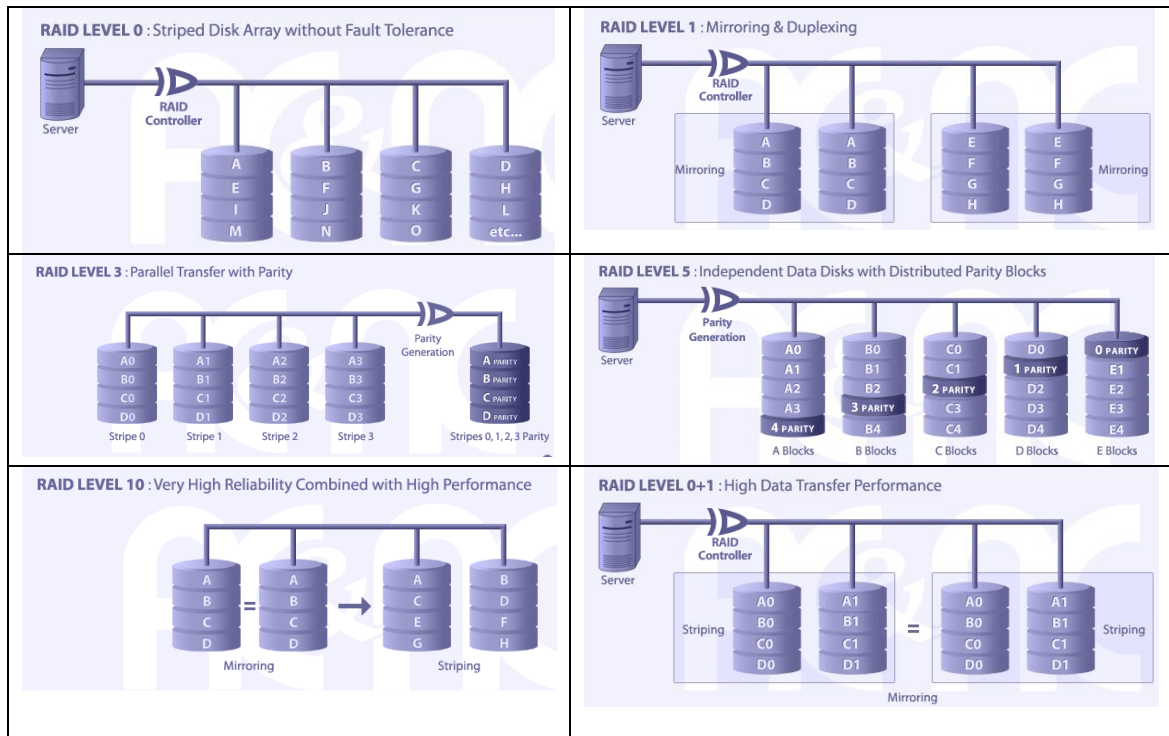


Tabela 2: Principais níveis de RAID.

Há outros tipos de RAID não comentados mas também de uso mais restrito que os propostos aqui. Juntamente com o RAID, quando se fala de *storage*, há também o conceito de *hot-spare* que é um disco reserva que fica esperando a possível quebra de um dos discos do *array*. Então ele é requerido automaticamente pelo subsistema de discos passando a integrar o *array* conforme o padrão de RAID utilizado. O *hot-spare* tem que ter um tamanho igual ou maior que os discos do array uma vez que o sistema de raid, ao usá-lo, implementar neste as mesmas características dos discos componentes deste array.

Além das formas de proteção aos dados mencionadas anteriormente, existe a proteção remota que implica em fazer *mirroring* (espelhamento) à distância copiando de um *storage* para outro, sincronamente ou assincronamente, toda a informação contida nos discos do primeiro. Requer uma conexão de alta velocidade entre os dois *storages* para que o servidor não sinta nenhuma perda de desempenho. Normalmente o segundo *storage* é colocado num outro prédio ou filial da empresa com intuito de preservar as informações em caso de incêndio, danos elétricos ou quebra de *storage*. Deve estar previsto também que se um *storage* parar, o outro assuma de forma automática ou no menor tempo possível, as requisições de I/O promovendo transparência e alta disponibilidade do acesso às informações.

2.5 PROTOCOLOS FIBRE CHANNEL

Segundo (LONG, 2006) e (TATE, LUCCHESI, MOORE, 2006), o protocolo *fibre channel* ou FCP é um protocolo de alto nível que, juntamente com o protocolo SCSI, proporcionam um transporte de blocos de dados entre entidades da rede de forma segura e com alto desempenho.

Por que *fiber* ou *fibre channel*? *Fibre channel* foi originalmente designado para utilizar o cabeamento de fibra óptica. Quando o suporte ao cobre foi aderido por esta tecnologia, o comitê decidiu manter o nome “*fibre*” a princípio, respeitando a ortografia inglesa quando se estiver referindo ao padrão. De outra forma, quando se referir genericamente ao cabeamento de cobre ou de fibra pode-se utilizar a designação “*fiber*” dada pela ortografia americana.

O *Fibre Channel* ou canal de fibra óptica é a tecnologia empregada em redes de alto desempenho como a rede SAN. Essa tecnologia inicialmente foi concebida para funcionar no campo dos supercondutores e depois foi amplamente adotada pelas redes compartilhadas de alta velocidade.

Historicamente, o *fibre channel* apareceu em 1988 e foi aprovado pela ANSI (*American National Standards Institute*) em 1994. Um grupo de trabalho denominado ANSI X3T9.5 formado em 1980, tinha como propósito desenvolver uma rede de transporte de dados de alto desempenho utilizando fibra óptica. Com o intuito de simplificar uma outra tecnologia da época, o HIPPI (*High Performance Parallel Interface*), nasceu o FDDI (*Fiber Distributed Data Interface*) que adotou a estrutura do projeto IEEE 802 para redes locais. O HIPPI era um projeto oneroso que usava na sua concepção 50 pares de cabos, conectores enormes e somente atingia pequenas distâncias para fornecer em princípio velocidades na casa dos 800 Mbps, além de ser simplex (transmissão unidirecional). Com a tecnologia *fibre channel*, esses problemas foram bastante minimizados uma vez que as perdas de sinal são mínimas, as velocidades e as distâncias podem ser bem maiores, não existe preocupação com interferências eletromagnéticas e os dispositivos e conexões utilizados têm tamanho bem reduzido em relação aos do HIPPI. Mais tarde deu suporte aos protocolos de alto nível como o SCSI, IP e ATM (*Asynchronous Transfer Mode*) sendo mais utilizado o padrão SCSI, bem como permissão de um número maior de dispositivos na rede.

Em 1980 a rede LAN era baseada nos protocolos Ethernet e os dados trafegavam a velocidade de 10 Mbps. Com o desenvolvimento dessas novas tecnologias, essas novas redes,

como a SAN, passaram a trafegar dados utilizando o método de acesso em baixo nível (camadas inferiores do modelo *fibre channel*) o que lhes confere superior desempenho com relação ao transporte de dados. Esse tráfego se assemelha ao praticado internamente pelos discos ATA e SCSI, por exemplo, onde os blocos ou segmentos de dados são requeridos pelo servidor e recebidos através da rede.

2.5.1 Camadas do Modelo Fibre Channel

O modelo OSI (*Open Systems Interconnection*) usa um mapeamento ordenado composto por sete camadas: a física, a de enlace, a de rede, a de transporte, a de sessão, a de apresentação e a de aplicação, esta última, no topo das camadas onde estão os protocolos de nível superior.

Já o modelo *fibre channel* é definido como uma arquitetura de cinco camadas hierárquicas, conforme a figura 5.

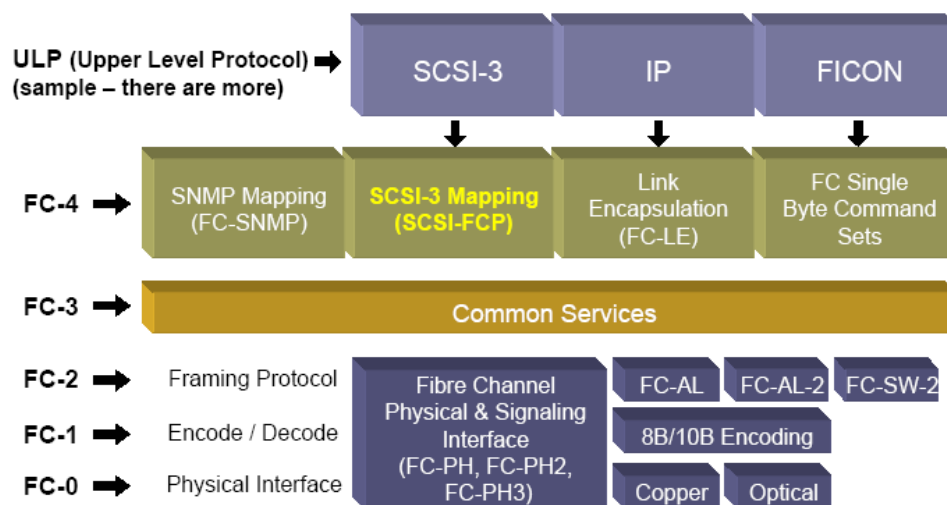


Figura 5: Modelo de Camadas do Fibre Channel.

A camada FC-0 é a camada de interface física responsável pela sinalização de transferência dos dados, incluindo parâmetros ópticos e elétricos, pelo tipo de cabeamento e conectores utilizados na infra-estrutura de rede conferindo segurança e redundância dos caminhos utilizados pela rede. FC-0 também especifica o OFC (*Open Fibre Control system*) que é um sistema de segurança utilizado para controlar o nível de potência óptica no link (linha) de dados quando ocorre uma condição de fibra aberta. Esta característica de segurança

é necessária porque os níveis de potência óptica em um sistema deste tipo excedem os limites definidos pelas normas padrões de segurança do laser. Sempre que ocorre o fato de uma fibra óptica se abrir, a partir da porta de envio, a porta receptora detecta os seus pulsos de laser em um ciclo de baixo nível o que interfere nos requisitos de segurança do laser. O receptor na outra porta detecta o sinal de pulsação e envia impulsos dentro do intervalo especificado pelos padrões de segurança do laser. Quando ocorre a restauração da fibra aberta, os receptores de ambos as portas, passam a receber os sinais de pulsação mais fortes o que irá resultar em um procedimento de duplo *handshaking* para restaurar a transmissão normal depois de alguns segundos. FC-0 estabelece também a extensão permitida para o uso do cabeamento em cobre, que pode chegar a algumas dezenas de metros, ou cabeamento em fibra, que pode chegar a dezenas de quilômetros. O modelo FCP (*Fiber Channel Protocol*) para rede SAN emprega normalmente o cabeamento óptico pelas vantagens de distância mencionadas e pela flexibilidade de ligação dos dispositivos ativos e máquinas servidoras.

A camada FC-1 define como os caracteres são codificados e decodificados para a transmissão, faz detecção de erros, sincronização de bytes e alinhamento de palavras. FC-1 utiliza codificação 8b/10b desenvolvida pela IBM como método de codificação e decodificação das informações a serem enviadas e recebidas, facilitando o controle de erros na transmissão.

A camada FC-2 possui um maior nível de complexidade por definir como as informações, blocos de dados, vindas das camadas superiores, serão segmentadas em uma sequência de quadros e entregues a essa própria camada de transporte. FC-2 é responsável pela sinalização básica no mecanismo de transmissão dos dados, sem colisões ou perdas, é responsável pelas classes de serviços suportadas e pelo gerenciamento do transporte de dados e controle de fluxo dos quadros transferidos entre o *Initiator* e o *Target*.

A camada FC-2 define os seguintes formatos para os dados: *Ordered Set* (Conjunto Ordenado), *Frame* (Quadro), *Sequence* (Seqüência) e *Exchange* (Troca).

Os dados são organizados em blocos e por sua vez em pacotes de dados denominados quadros, estes constituem as seqüências que, por sua vez, formam as *exchanges*, para serem enviadas pela rede. A organização dos dados, na arquitetura *ibre channel*, de forma ordenada para o transporte, recebe o nome de *Ordered Set*.

O *Ordered Set* é dividido basicamente em três categorias que são: *start of frame* (SOF - início do quadro), *end of frame* (EOF - fim do quadro) e sinais primitivos. SOF e EOF também indicam a classe de serviço que será utilizada. Os sinais primitivos indicam ações ou

De acordo com a tabela 3, o cabeçalho do quadro fibre channel é formado por seis palavras endereçadas de 0 a 5 e cada palavra composta por quatro bytes. Essa estrutura do cabeçalho do quadro divide-se em campos contendo funções de controle do frame que são explicadas a seguir:

- **R_CTL** (*routing control*) é o campo composto por um byte que identifica o tipo de informação contida na unidade de dados e para qual porta de destino o quadro deve ser encaminhado.
- **D_ID** (*Destination_ID*) é o campo composto por 24 bits que contém o endereço de destino do quadro a ser enviado.
- **S_ID** (*Source_ID*) é o campo composto por 24 bits que contém o endereço de origem do quadro a ser encaminhado.
- **Type** é o campo composto por um byte que identifica o protocolo utilizado na transmissão do quadro, como por exemplo o SCSI, e controle dos quadros enviados.
- **F_CTL** (*Frame Control*) é o campo composto por 24 bits que contém informações de controle do conteúdo do quadro.
- **SEQ_ID** (*Sequence Identifier*) é o campo formado por um byte que implementa um identificador de sequência atribuído exclusivamente para o par D_ID e S_ID.
- **DF_CTL** (*Data Field Control*) é o campo formado por um byte que especifica se existem cabeçalhos opcionais presentes no início do campo de dados.
- **SEQ_CNT** (*Sequence Count*) é o campo da tabela acima formado por 16 bits e tem a função de identificar a posição de um quadro em uma sequência incrementando de um (1) os quadros subsequentes dentro da mesma sequência.
- **OX_ID** (*Originator X_ID*) é o campo composto por 16 bits usado para identificar o conjunto de sequências denominado *exchange* atribuído pela origem.
- **RX_ID** (*Responder X_ID*) é o campo composto por 16 bits usado para identificar a *exchange* pelo lado do destino.
- **Parameter** é o campo de 32 bits que especifica a posição relativa dos frames de dados ou informação específica de controle dos dados no link.

Frame payload contém a unidade de dados cujo tamanho varia de 0 a 2112 bytes correspondendo a própria informação a ser transportada entre os ativos envolvidos nessa comunicação.

CRC (Cyclic Redundancy Check) é o campo de quatro bytes responsável por verificar a integridade dos dados contidos nos quadros. Esse processo de detecção dos erros é realizado e logo após ocorre a codificação desses bytes de CRC juntamente com o restante do quadro.

EOF (End of frame delimiter) é o campo composto por quatro bytes responsável por indicar o final do quadro. O EOF também diz ao receptor em que posição está o quadro na sequência de quadros em questão.

A figura 7, conforme CISCO-SAN1501, mostra a hierarquia dos pacotes de dados, ou seja, a organização dos frames (quadros) enviados pelo *initiator* (servidor) ao *target* (dispositivo de armazenamento).

Conforme a figura 7, o conjunto de frames correlacionados forma uma sequência, e por sua vez, o conjunto de sequências desassociadas forma uma *exchange* (troca).

Na transmissão entre o *initiator* e o *target*, múltiplas *exchanges* são iniciadas sendo identificadas pelo endereço de origem (OX_ID) e pelo endereço de destino (RX_ID) e controladas também pelo identificador de sequência (SEQ_ID). A figura 7 mostra as estruturas de *exchange*.

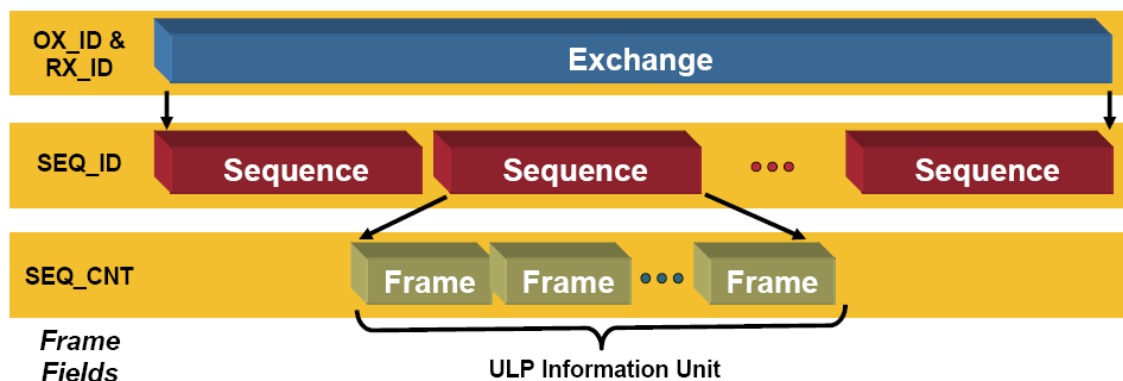


Figura 7: Hierarquia dos pacotes de dados do FC.

Durante o transporte dos dados, se ocorrer um erro em algum ou alguns quadros, o protocolo *fibre channel* não utiliza nenhuma técnica de correção de erro a nível de quadro, mas sim, verifica qual sequência foi corrompida e a reenvia novamente. Essa prática permite a minimização dos *overheads*, uma vez que basta enviar um cabeçalho por sequência de *frames*, aumentando o desempenho do sistema.

A camada FC-3 é responsável pelas funções auxiliares como o *Striping*, *Hunt Groups* e *Multicasting*. O *striping* usa múltiplas portas para aumentar a largura de banda paralelamente. O *hunt group* permite que mais de uma porta responda a um mesmo endereço através da denominação desse grupo de portas como um alias identificador. Durante a transmissão qualquer porta do grupo pode ser usada para envio de dados objetivando uma redução da latência gerada pela espera por uma porta de envio. O *multicasting* é o serviço que usa transmissões de quadros para várias portas de destino na rede.

A camada FC-3 ainda está em desenvolvimento e futuramente poderá incluir funções de criptografia e compressão de dados.

A camada FC-4 está no mais alto nível da hierarquia do *fibre channel* chamada ULP (*Upper Layer Protocol*). FC-4 provê os protocolos de aplicações específicas e é responsável pelo encapsulamento das várias camadas. Como exemplo de protocolos interfaceados por esta camada temos o SCSI-3 (*Small Computer System Interface-3*), o IP (*Internet Protocol*), o HIPPI (*High Performance Parallel Interface*), o ATM (*Asynchronous Transfer Mode*), ESCON/FICON² (*Enterprise System Connection/Fiber Connectivity*). O FCP (*Fiber Channel Protocol*) permite que na mesma interface física sejam transportados tanto pacotes de dados da rede IP quanto dados SCSI sobre o protocolo *fibre channel* simultaneamente. Isso acontece devido às regras de mapeamento da camada FC-4 que descrevem essa inter-operação dos processos da camada superior ULP.

De acordo com (LONG, 2006), comparando o modelo *Fibre Channel* ao modelo OSI (*Open Systems Interconnection*), pode-se associar a grosso modo o mapeamento das camadas de acordo com a tabela 4.

Modelo de Referência OSI		Modelo Fibre Channel
Aplicação		FC-4
Apresentação		
Sessão		FC-2, FC-3
Transporte		
Rede		
Enlace		FC-2, FC-3
Física		FC-1
		FC-0

Tabela 4: Comparação do modelo de camadas OSI com FC.

² ESCON (*Enterprise System Connection*), marca registrada da IBM, é o canal de fibra óptica anterior ao padrão FICON da IBM. FICON (*Fiber Connectivity*) é um protocolo de alta velocidade de I / O. É uma interface para conexão de computador mainframe para dispositivos de armazenamento. Utiliza o fibre channel como meio físico para transmissão de dados.

2.5.2 Classes de Serviços

As aplicações do protocolo *Fibre Channel* podem requerer níveis apropriados de serviços que garantam a conectividade, entrega das informações postas a transmissão e banda disponível para maior eficiência na comunicação entre os elementos ativos do transporte de dados. Algumas aplicações, como o backup em fitas, necessitam de uma banda dedicada durante o tempo em que houver a transmissão dos dados com integridade e eficiência no transporte das informações. Outras aplicações como transmissões de vídeo não precisariam de conexões dedicadas mas certamente a eficiência na transmissão será levada em conta.

Por essas necessidades certas classes de serviços estão disponíveis e conforme (TATE, LUCCHESI, MOORE, 2006), há sete classes de serviços comentadas abaixo:

- **Classe 1** – Provê serviço com conexão dedicada estabelecida entre o *initiator* e o *target* durante o processamento das informações com reconhecimento e conexão orientada ao transporte. Ela garante recebimento dos dados pelo dispositivo *target* na mesma ordem do envio pelo *initiator* e reserva capacidade de largura de banda total para esses dois dispositivos. Por isso é essa classe atende as aplicações que usem elevado consumo de banda como backup e transmissões multimídia. A classe 1 é pouco usada devido a bloquear outros dispositivos concorrente a largura de banda que ora está sendo usada pelos dispositivos iniciais até que estes acabem sua transmissão de dados.
- **Classe 2** – Fornece reconhecimento de serviço sem efetuar conexão e faz melhor uso da largura de banda disponível. Provê conexões não dedicadas, com isso compartilhando a banda e permitindo multiplexação das várias mensagens de dados. Os quadros transmitidos podem tomar diferentes rotas sem garantia da ordem com a qual chegarão a sua porta de destino. É mais utilizada em aplicações onde o tempo e a ordem de entrega dos quadros não são críticos como os serviços de FTP (*File Transfer Protocol*) e transporte de pacotes IP (*Internet Protocol*).
- **Classe 3** – Não há conexão dedicada e os quadros recebidos no destino não são reconhecidos. A classe 3 é também conhecida como serviço de *Datagram Connectionless* (datagrama sem conexão). Ela otimiza o uso dos recursos da rede fiber channel mas precisa dos protocolos da camada superior para garantir que os quadros sejam recebidos na mesma ordem que são enviados, e possa haver a retransmissão, pelo dispositivo fonte, dos quadros perdidos e não detectados pelo destino. Em geral, o

serviço da classe 3 é usado pelas redes *fibre channel* mais comumente pela FC-AL (conectividade loop arbitrário do *fibre channel*) que será falado mais adiante.

- **Classe 4** – Provê serviço orientado à conexão e estabelece um canal virtual entre as duas portas de envio e recepção garantindo QoS (Qualidade de Serviço) e largura de banda. Apresenta como principal diferença em relação a classe 1, alocação de apenas uma fatia de consumo de banda disponível permitindo que outras conexões possam desfrutar dessa mesma banda. Também nesta classe é garantida a entrega dos quadros de dados com reconhecimento pelo nó de destino e ordenadamente excetuando-se que a multiplexação dos frames é responsabilidade do dispositivo comutador (*fabric*) do *fibre channel*. A classe 4 é normalmente usada para transmissão de aplicações multimídia como vídeo e outras aplicações em determinados setores da empresa que requeiram largura de banda confiável.
- **Classe 5** – É chamada de serviço isócrono e é destinado a aplicações que requerem entrega imediata dos quadros de dados sem a utilização do buffer. Este serviço não é ainda claramente definido e não tem registro nos documentos do *fibre channel*.
- **Classe 6** – Deriva da classe 1 e é conhecida como classe de serviço *multicast* (transmissão das informações simultaneamente a vários usuários da rede). Ela provê conexões dedicadas com multicast confiável. O serviço de *multicast* na rede SAN, a nível de portas dos nós do *fabric*, é estabelecido com reconhecimento entre origem e destino. Após estabelecida uma conexão entre origem e destino, esta é mantida até o iniciador desta conexão terminar da transmissão dos dados, é o que garante a classe 6. A classe 6 é mais utilizada em aplicações que transmitem áudio e vídeo com funcionalidades *multicast*.
- **Classe F** – É definida para uso entre switches *fibre channel* em sua comunicação ISL (*inter-switch link*). O serviço de classe F é estabelecido entre as portas do *fabric* da SAN sem conexão garantida e sem notificação de entrega dos *frames*. Ele é usado para funcionalidades de controle, coordenação e configuração entre os *switches* do *fabric* da SAN. É semelhante à classe 2 com o diferencial das portas usadas por cada serviço (classe 2 e classe F) e pelo fato desta ser usada para controle e administração do *fabric*.

No modelo *Fibre Channel*, as conexões físicas são realizadas através de portas físicas, segundo o modelo SCSI, interligadas por fibra óptica, comumente, ou cabeamento de cobre e para cada conexão entre o fonte e o destino, existirá um tipo de porta inerente aos dois. Pode-

se ver na tabela 5 os tipos de portas mais comuns usados para a interligação entre os dispositivos.

Nome da porta	Significado	Função
Node Port	Porta Nó ou Porta_N	Interliga dispositivos numa conexão ponto a ponto com portas do tipo nó ou interliga dispositivos com porta nó a switches fibre channel, nestes, a uma porta F.
Fabric Port	Porta Fabric ou Porta_F	Portas proprietárias dos switches fibre channel que servem para interligar-se à porta N de outros dispositivos.
Expansion Port	Porta de Expansão ou Porta_E	São usadas para conexão entre dois switches fibre channel.
Generic Port	Porta Genérica ou Porta_G	São portas dos switches fibre channel com possibilidade de simular uma Porta_E ou Porta_F.
Bridge Port	Porta Bridge ou Porta_B	Faz a conexão entre switches do fabric em ligações de dispositivos da rede LAN e WAN.
Loop Port	Porta Laço ou Porta_L	A letra L é acrescentada às letras das portas F, G, N que estiverem numa configuração do tipo laço arbitrado constituindo as denominações Porta_FL, Porta_GL, Porta_NL.

Tabela 5: Tipos de portas utilizadas na conexão entre dispositivos na SAN.

Um exemplo da utilização dessas portas do *fibre channel* se encontra na figura 8.

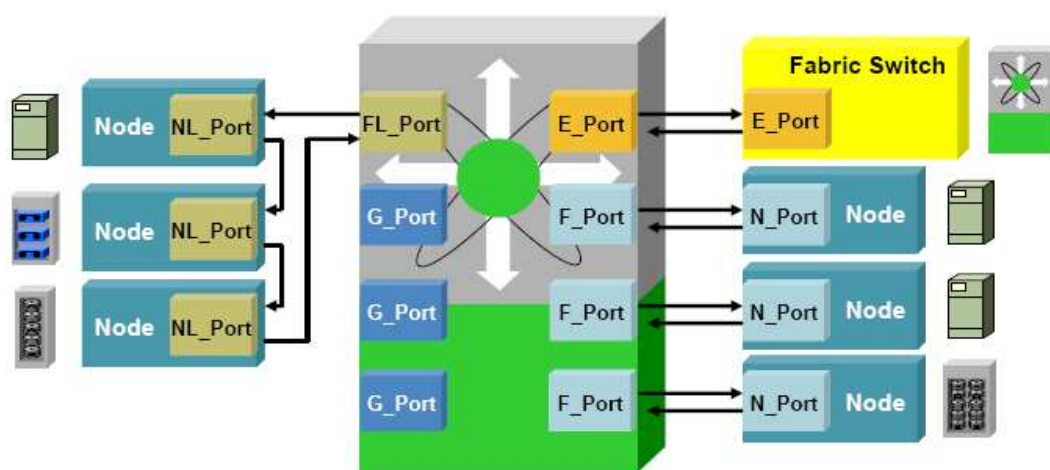


Figura 8: Exemplo de utilização das portas no Fibre Channel.

A conexão entre dois switches através das portas de expansão do tipo porta_E é chamado de *inter-switch link* (ISL).

2.5.3 Topologias da Arquitetura do Fibre Channel

A conectividade do protocolo *fibre channel*, nas redes SAN, pode ser implementada através de três topologias principais que sejam a *point-to-point* (ponto a ponto), a *arbitrated loop* (loop arbitrário) e *switched fabric* (trama de switches). Dessas três, a mais utilizada é a

switched fabric devido aos requisitos de escalabilidade, flexibilidade e disponibilidade que este tipo de conexão proporciona. Abaixo, está a explicação de cada topologia:

- **Point-to-Point** – Como o próprio nome diz, especifica uma conexão do tipo ponto a ponto, onde o *initiator* (fonte) e o *target* (destino), através das portas específicas (portas tipo N), são ligados um ao outro. Esta topologia não permite compartilhamento, permitindo o uso simultâneo do canal de dados por apenas os dois alvos envolvidos. A transmissão de dados é feita de maneira *full duplex* onde, por exemplo, numa conexão de dados a 200 MBps, as informações trafegam simultaneamente nos dois sentidos, a essa velocidade, o que daria no final, do somatório de 200 MBps + 200 MBps, uma conexão a 400 MBps. Da mesma forma, para velocidades mais baixas ou mais altas, o comportamento é o mesmo. A figura 9 ilustra essa conexão.

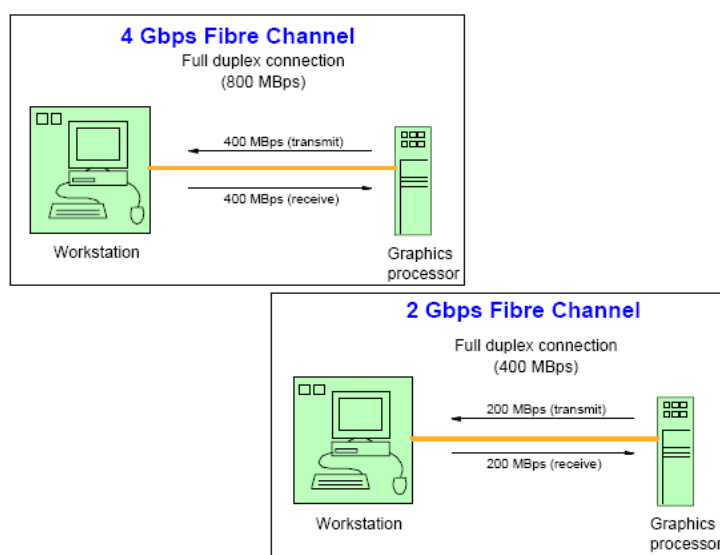


Figura 9: Topologia *point-to-point* do FC.

- **Arbitrated Loop** – Significa laço arbitrário. É o tipo de conexão que proporciona maior flexibilidade e escalabilidade à rede. Os dispositivos são ligados serialmente em um anel físico através das portas de tipo NL onde o tráfego de dados será feito unilateralmente entre essas portas. Sendo assim, existe um caminho contínuo entre quaisquer duas máquinas que estejam nessa configuração de laço. Um inconveniente para esse tipo de ligação é que se ocorrer uma desconexão em algum nó, o loop inteiro deixará de funcionar. Por esta causa desenvolveu-se uma variação dessa topologia chamada *arbitrated loop hubs* onde o *initiator* fecha uma

conexão ponto a ponto com o *target* através do *hub* para transferência dos dados. Nessa topologia, se um laço for danificado, existirá um outro caminho a ser usado através do hub. No loop arbitrário é permitida a conexão de até 127 dispositivos. A figura 10, a seguir, ilustra esse tipo de conexão.

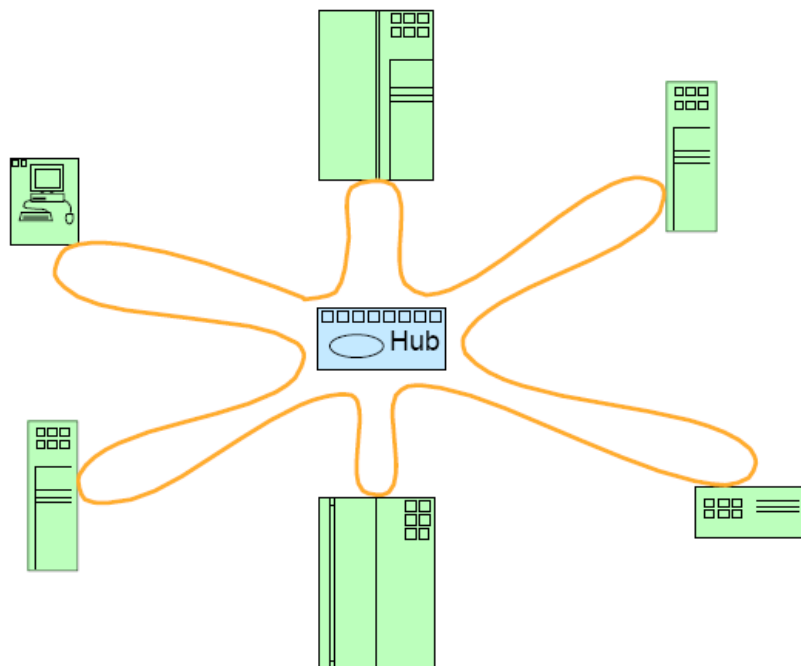


Figura 10: Topologia *arbitred loop* do FC.

- **Switched Fabric** – É a topologia onde os dispositivos ativos estão interligados na rede de armazenamento de dados – a SAN (*Storage Area Network*), através de *switches fibre channel*. Nessa topologia é visível o desempenho obtido, a escalabilidade e o melhor uso dos recursos além de se obter melhor largura de banda disponível. Na teoria, pode suportar até 2^{24} nós de dispositivos interligados aos switches, correspondendo a mais de 16 milhões de máquinas. Com relação ao número de *switches fibre channel* na rede SAN, pode-se ter até 239 switches endereçados como limitação de capacidade máxima para o modelo *fibre channel*. Uma diferença bastante significativa do *switched fabric* em relação ao *arbitrated loop*, é a questão da largura de banda total que é provida pelo *fabric* em relação a largura de banda compartilhada pelos dispositivos no loop arbitrário. Cada dispositivo agregado ao *fabric* provê um *throughput* (taxa de transmissão de dados) relativamente maior enquanto no loop arbitrário, cada entrada de

dispositivo ao anel, representa uma divisão maior da largura de banda e conseqüente diminuição do *throughput* de dados. As figuras 11 e 12 a seguir mostram as portas e conexões em fibra óptica de um *switch* IBM B16.



Figura 11: Switch IBM B16 da JFRN e suas conexões em fibra óptica.



Figura 12: Visão frontal do switch IBM B16 pela interface web.

2.5.4 Fabric Zoning

O zoning (zoneamento) é o serviço que segmenta os switches do fabric de maneira a se criar zonas para restringir ligações entre nós (membros) e a utilização de recursos na rede SAN. Apenas os dispositivos integrantes de uma mesma zona podem se comunicar conforme os direcionamentos implementados na zona impedindo que membros de outras zonas tenham acesso a essa mesma. O zoneamento organiza as comunicações entre dispositivos dentro do

fabric com a criação de zonas que interligam, por exemplo, a porta ou interface HBA³ de um servidor a uma das portas do storage e, em outro exemplo, a mesma porta do servidor à unidade de *backup* de fita constituindo uma outra zona. Portanto, um mesmo nó ou dispositivo pode estar em várias zonas ao mesmo tempo. A seguir, pode-se ver a figura 13 ilustrando o zoneamento:

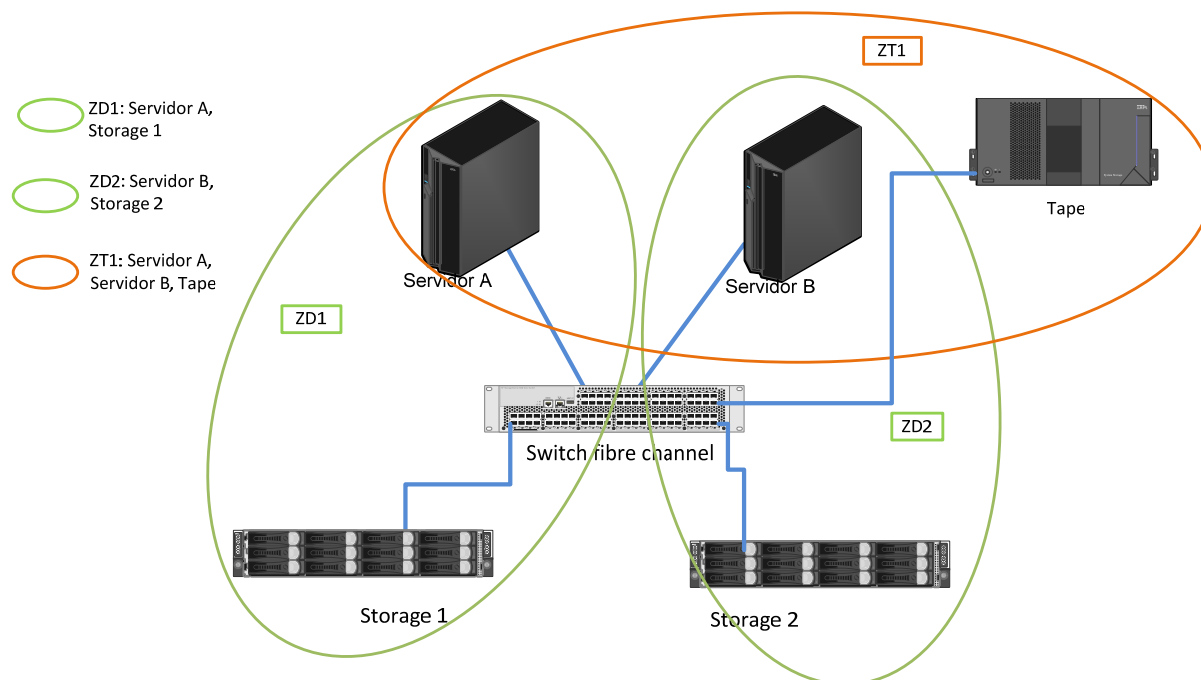


Figura 13: Zoneamento.

Com relação ao modo de operação do *zoning*, este pode ser configurado estaticamente através da configuração por portas ou dinamicamente através do *wwn* (*world wide name*).

A configuração por portas estabelece uma ligação direta dos dispositivos aos switches do *fabric* e as zonas vinculam as portas do switch às portas dos dispositivos pertencentes às zonas em questão. Se um servidor ligado a determinada zona, nesta modalidade, for mudado de porta, ele fica fora da zona, que só funciona para as portas previamente configuradas.

Na configuração dinâmica por *wwn* (*world wide name*) o dispositivo ligado a zona é reconhecido por este endereçamento, o que implica na facilidade de se mudar as entidades ligadas ao *switch*, de porta, sem afetar o zoneamento. Essa modalidade tem mais flexibilidade que o modelo por porta mas apresenta vulnerabilidades a intrusões.

³ HBA (*Host Bus Adapter*) – é um dispositivo de hardware que faz a tradução de sinais elétricos ou de sinais ópticos para sinais elétricos, entre as interfaces de entrada/saída dessa mesma placa para permitir a conversação entre os protocolos de dados nas redes SAN-IP ou SAN-FC e pode usar cabeamento de cobre ou fibra óptica.

O WWN foi especificado pelo subcomitê ANSI T11 em conjunto com o IEEE e corresponde a uma identificação única para cada dispositivo hba semelhante ao MAC (Media Access Control) das placas Ethernet. O wwn é expresso em notação hexadecimal composto por 64 bits, com um campo inicial composto por 2 bytes escritos como 10:00 ou 2X:XX onde os Xs representam informações do fornecedor da placa hba, mais 3 bytes de identificação e outros 3 bytes que compõem o número de série da hba. A tabela 6 exemplifica o endereçamento para o wwn que é também chamado de wwid (world wide identifier).

Caracteres iniciais 10:00 ou 2X:XX	Identificador do fornecedor	Número de série do fornecedor
10 : 00 :	07 : 00 : 9c :	f4 : 23 : 2a
25 : 11 :	07 : 00 : 9c :	f4 : 23 : 2a

Tabela 6: WWN.

De acordo com um novo esquema proposto pelo IEEE para endereçamentos da SAN, o wwn é um endereçamento de 64 bits onde o primeiro nibble (metade de um octeto) é o número hexa 5 ou 6; os próximos 3 bytes correspondem a identificação do fornecedor e os 36 bits restantes correspondem a outras especificações do fornecedor como o número de série. A tabela 7 exemplifica esse novo padrão de wwn.

	Identificador do fornecedor	Número de série do fornecedor
5	0 : 07 : 9c : f	4 : 23 : 2a : 89 : 1c
6	0 : 07 : 9c : f	4 : 23 : 2a : 89 : 1c

Tabela 7: Novo esquema de WWN.

Essa quantidade de bits (64) de wwn trafegando na rede provoca impactos na eficiência de roteamento dos pacotes e, devido a isso, um outro esquema de endereçamento pode ser usado nas redes fibre channel: é o endereçamento por porta que é composto por 24 bits onde os bits de 23 a 16 identificam o domínio (contém o endereço do switch fibre channel), os bits de 15 a 8 qualificam a área (grupo de portas internas do switch fibre channel) e os bits de 7 a 0 correspondem ao endereço físico da porta do nó ligado ao *switch fibre channel* conforme (TATE, LUCHEESE, MOORE, 2006) fornecendo a figura 14 a seguir:

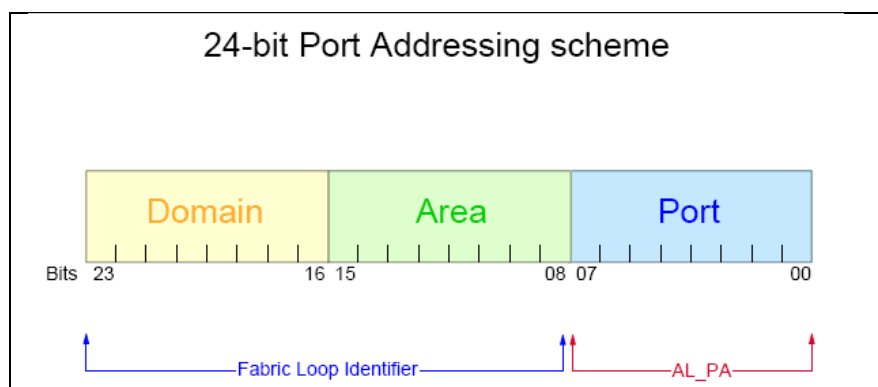


Figura 14: Endereçamento de portas no fabric.

Devido ao número de bits nesse endereçamento pode-se ter até 256 possibilidades para o domínio (domain), e considerando-se os serviços de broadcast e outras reservas, esse endereçamento que corresponde ao número de switches no ambiente SAN cai para 239 ou seja: pode-se ter até 239 switches interligados nesse ambiente teoricamente.

O campo Area provê também até 256 partes de endereço que identificam individualmente o grupo de portas_FL da topologia de loop arbitrário ou o grupo de portas_F do switch que se liga às portas_N de uma hba com mais de uma porta. Assim, cada grupo de portas tem um diferente número de Area mesmo que o grupo contenha somente uma única porta.

O campo Port provê também 256 partes de endereços que identificam as ligações entre as portas_N dos nós de dispositivos.

Pode-se calcular o número de endereços disponíveis multiplicando-se o Domain x Area x Port obtendo-se $239 \times 256 \times 256 = 15.663.104$ endereços.

2.5.5 Fabric - Topologia Física

Conforme (VERAS, 2009), o *fabric* é a configuração de múltiplos *switches* baseados no protocolo *fibre channel* para trabalharem juntos em arranjos do tipo *Core-edge fabric* ou em *Mesh fabric*.

Core-edge fabric, como o próprio nome presume, é o arrando composto de duas camadas, a *core* e a *edge* que pressupõe a utilização de *switches* convencionais na camada *edge* (borda) e *switches* do tipo *director* (dispositivos com mais portas que aumentam a disponibilidade do *fabric*) na camada *core* (núcleo). A figura 15 ilustra esse arranjo.

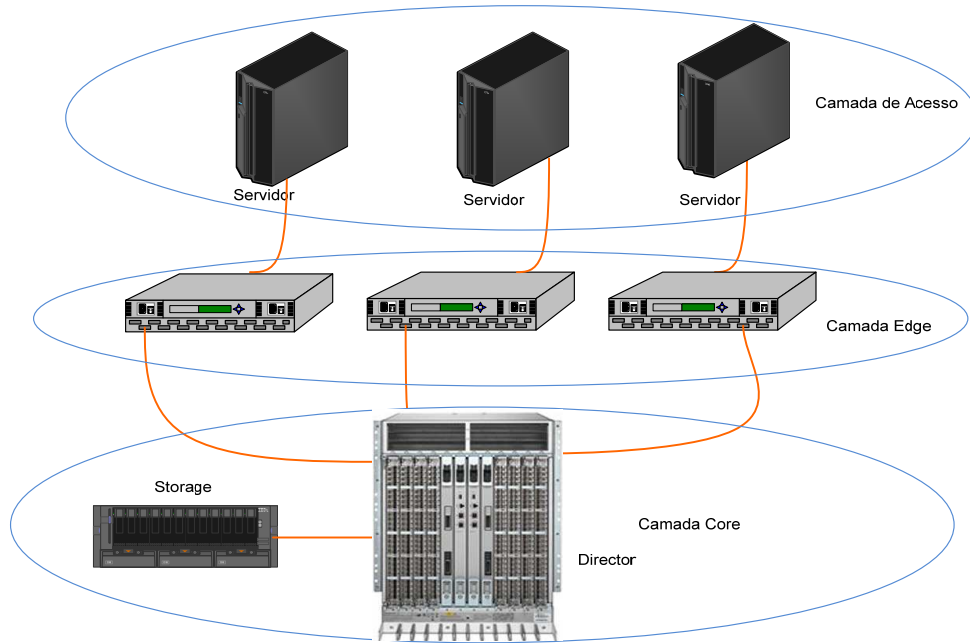


Figura 15: Topologia core-edge fabric.

Mesh fabric é o arranjo físico entre *switches* conectados diretamente um ao outro através das portas de expansão (porta_E) constituindo o ISL (ligação entre *switches*). Essa topologia se divide em dois tipos: a topologia *mesh* parcial e a topologia *mesh* total.

Topologia *mesh* parcial possui uma redundância menor, por isso com custo melhor que a *mesh* total, é usada quando se tem um número maior que quatro *switches* na rede e pode ser vista na figura 16 a seguir.

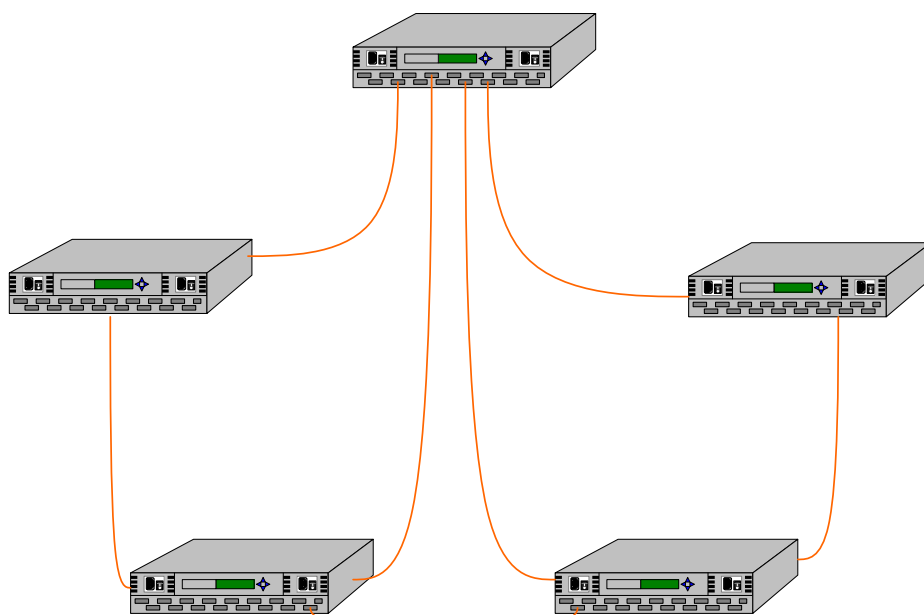


Figura 16: Topologia mesh parcial.

Topologia *mesh* total apresenta maior disponibilidade que a topologia *mesh* parcial através da conexão entre todos os switches em detrimento de uma menor escalabilidade e é melhor empregada na conectividade de redes pequenas até quatro switches interligados. A figura 17 mostra esse tipo de arranjo.

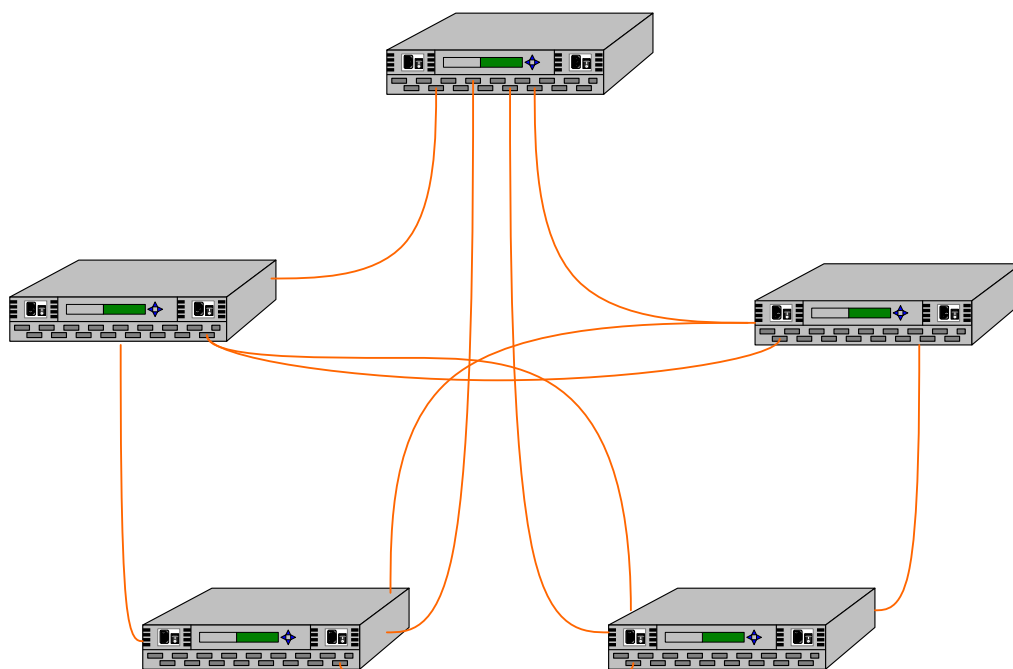


Figura 17: Topologia *mesh* total.

2.6 CONVERGÊNCIA DE REDES

Atualmente há uma tendência da infra-estrutura de redes, com intuito de flexibilidade e escalonamento, aliada ao melhor aproveitamento da base comum de rede já instalada, de utilizar juntamente com a rede SAN, as redes Ethernet, puramente, ou baseadas no protocolo de rede IP (Internet Protocol).

Com isso, surgiram os protocolos FCIP, IFCP, FCoE e iSCSI que também fazem parte dos protocolos de alto nível (ULP) descritos a seguir:

FCIP – *Fibre Channel* sobre IP – conforme (VERAS, 2009), é um protocolo de tunelamento que possibilita que múltiplas SANs baseadas no protocolo FC possam ser interligadas através de uma rede padrão TCP/IP. A RFC 3821 publicada em julho de 2004 descreve o protocolo FCIP.

A característica básica dessa rede é a formação de um túnel ponto a ponto que interliga duas ilhas de SANs que podem estar em localidades geográficas diferentes para formação de um ISL virtual. Com isso a visão da rede SAN entre as duas localidades, que podem estar usando a WAN (*Wide Area Network*), é como se os *fabrics* estivessem no mesmo local.

A desvantagem dessa rede é a necessidade de garantir alta largura de banda e o *overhead* maior que o do protocolo *fibre channel*. A figura 18 mostra a tecnologia FCIP.

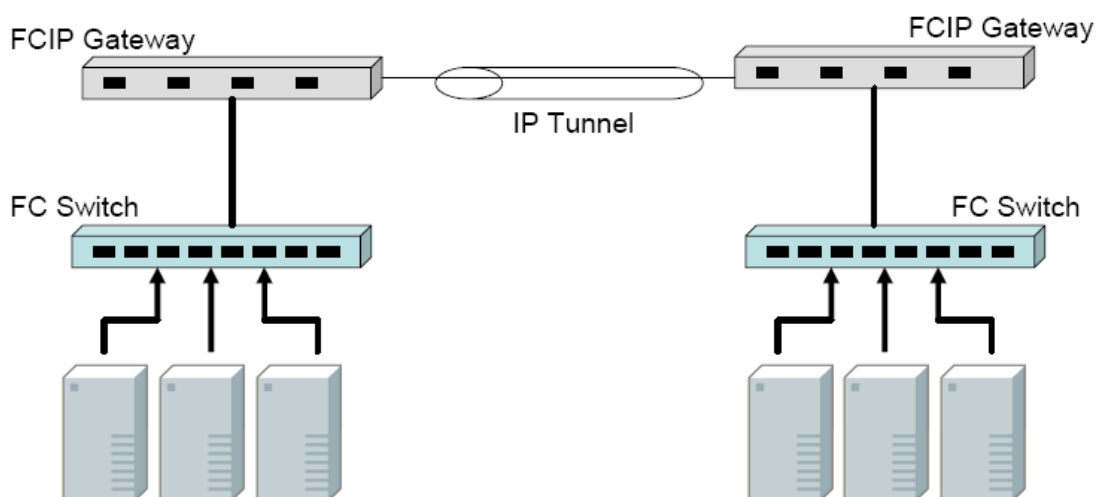


Figura 18: Tecnologia FCIP.

iFCP – *internet Fibre Channel Protocol* – é o protocolo de interligação de dispositivos fibre channel utilizando os benefícios do TCP/IP. A RFC 4172 publicou este protocolo em setembro de 2005.

A implementação do protocolo iFCP é semelhante aos encapsulamentos de *frames* usados pelo FCIP onde o controle dos congestionamentos e detecções de erros de transmissão de dados fica por conta dos protocolos TCP/IP.

Na questão da infra-estrutura de redes, há necessidade de gateways para fazerem as conversões de portas entre os switches FC e os roteadores da rede IP, conforme a figura 19 mostra. Por isso essa modalidade de encapsulamento do fibre channel pode se tornar mais cara que o implementado pelo protocolo FCIP. A figura 19 ilustra a tecnologia iFCP.

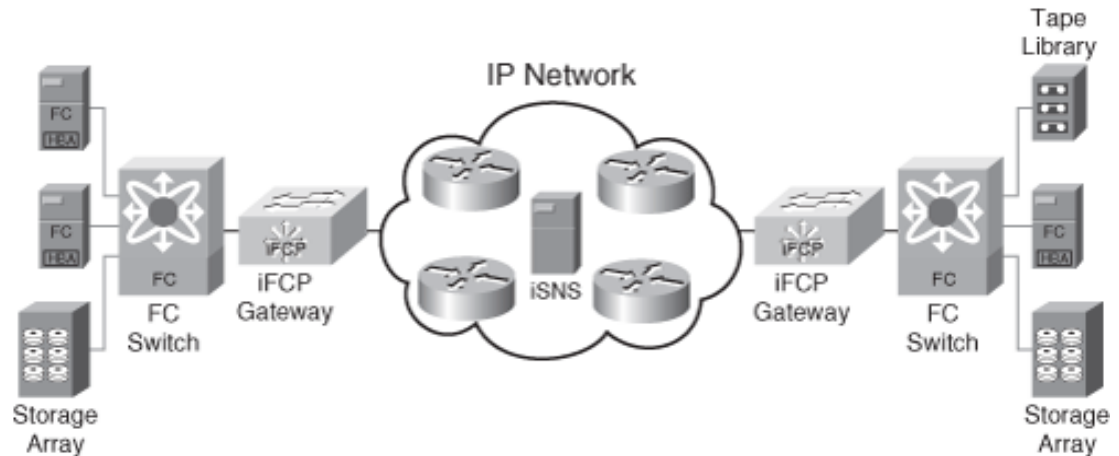


Figura 19: Tecnologia iFCP.

FCoE – *Fibre Channel sobre Ethernet* – é o protocolo que usa o encapsulamento da rede Ethernet para fazer o transporte dos quadros *fibre channel* a nível das camadas mais baixas do modelo OSI evitando assim o overhead das camadas intermediárias. Evita assim passar pelo TCP/IP, e usa dos benefícios das camadas MAC (*Media Acces Control*) e Física do modelo *Ethernet*. A figura 20 exemplifica a tecnologia FCoE.

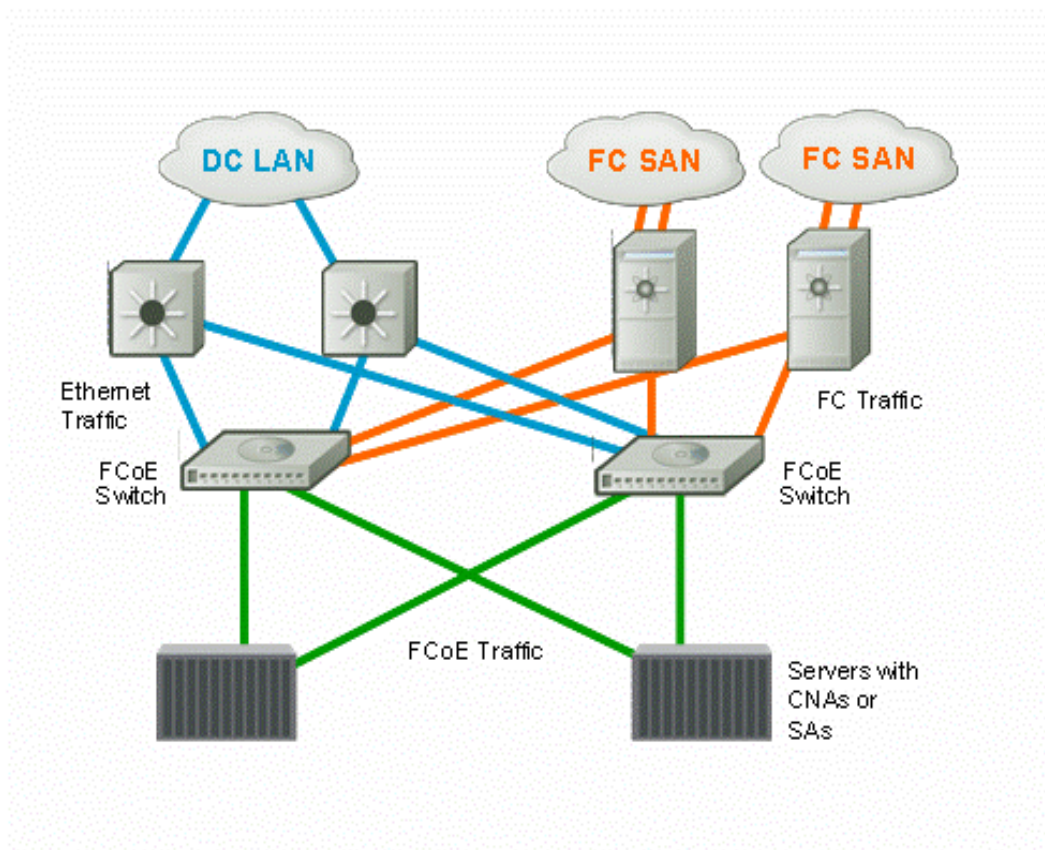


Figura 20: Tecnologia FCoE.

O modelo FCoE começou a ser desenvolvido pelo grupo de trabalho do T11 FC-BB-5 em abril de 2007. Em 3 de junho de 2009, o rascunho do desenvolvimento desse padrão foi concluído pela equipe do T11 e publicado um dia depois como um padrão ANSI.

As empresas Cisco e IBM em parceria com grandes fabricantes de *hardware* e *software*, desenvolvem essa nova tecnologia de transporte de dados juntamente com a inclusão de velocidades maiores para essas redes como o mínimo de 10Gbps para o FCoE em virtude dos gargalos e perdas de pacotes que afetam os protocolos da rede *Ethernet*. Uma representação da pilha de protocolos do FCoE, comparada ao modelo OSI e ao modelo FC, pode ser vista na figura 21.

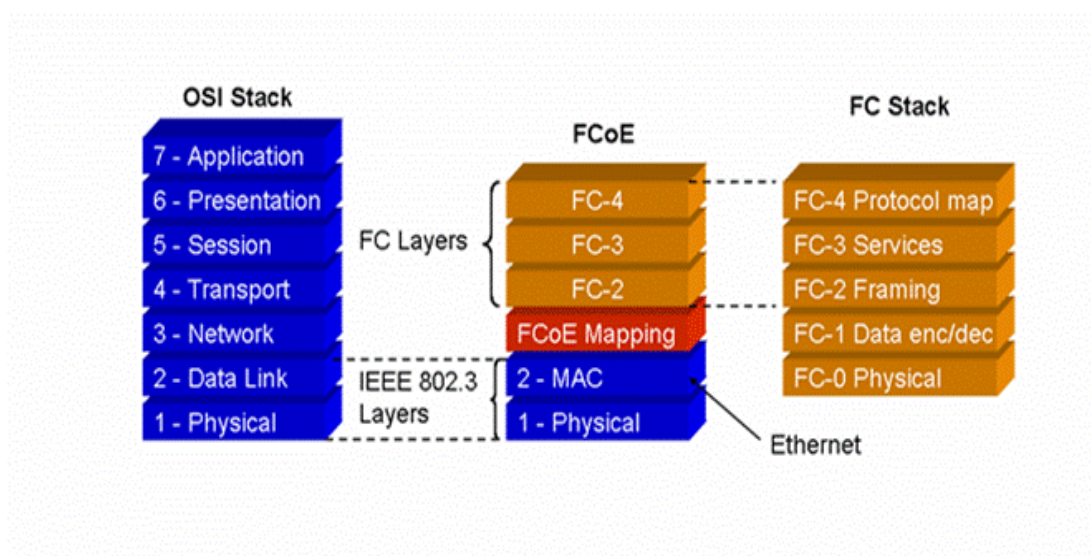


Figura 21: Representação das pilhas de protocolos OSI, FCoE e FC.

iSCSI - *internet Small Computer System Interface* – é o protocolo desenvolvido pelo IETF (Internet Engineering Task Force) para juntamente com a rede TCP/IP fazer o transporte dos dados concebidos em blocos SCSI.

O grupo de trabalhos do IETF responsável pela implementação deste padrão, teve forte ajuda de empresas de grande porte como a Cisco, IBM e HP, tendo começado no ano de 2000 e culminado com a RFC 3720 em 2003. Mais tarde, outras RFCs como a 4850 e 5048 fizeram correções e esclarecimentos sobre a RFC original.

O padrão iSCSI apresenta como componentes principais da rede de transporte de dados: o *initiator* (servidor), o *target* (unidade de armazenamento) e a rede IP. Essa estrutura de rede pode ser chamada SAN-IP onde o protocolo iSCSI tem o mesmo propósito do protocolo *fibre channel* da rede SAN-FC que é fazer o transporte eficiente do protocolo SCSI

CDB (*SCSI Command Descriptor Block*) para os dispositivos de armazenamento. A figura 22 exemplifica um modelo iSCSI.

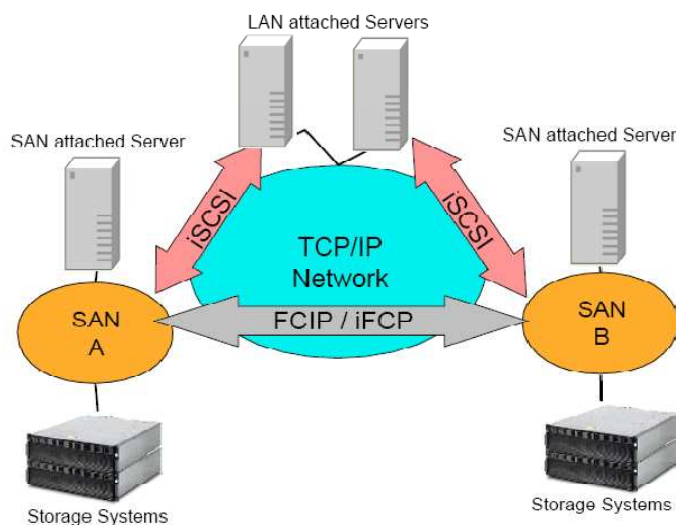


Figura 22: Exemplo de ligação iSCSI em rede SAN-IP.

Com o desenvolvimento de tecnologias e dispositivos de rede mais velozes como o 10 Gigabit Ethernet, é possível implementar na rede SAN-IP quase todos os benefícios da rede SAN-FC como eficiência, escalabilidade, interoperabilidade e compartilhamento dos recursos utilizando uma infra-estrutura já concebida para isso. O TCO (Custo Total de Propriedade) da empresa diminui, os testes, implementações e padronizações são mais simples por parte dos fabricantes de dispositivos e custos para empresa em treinamentos são menores comparados com a tecnologia *fibre channel*.

Com relação aos aspectos negativos, ainda há preocupações com a segurança do protocolo por usar TCP/IP, a largura de banda menor e instabilidades de rede, comparados a infra-estrutura de rede *fibre channel*.

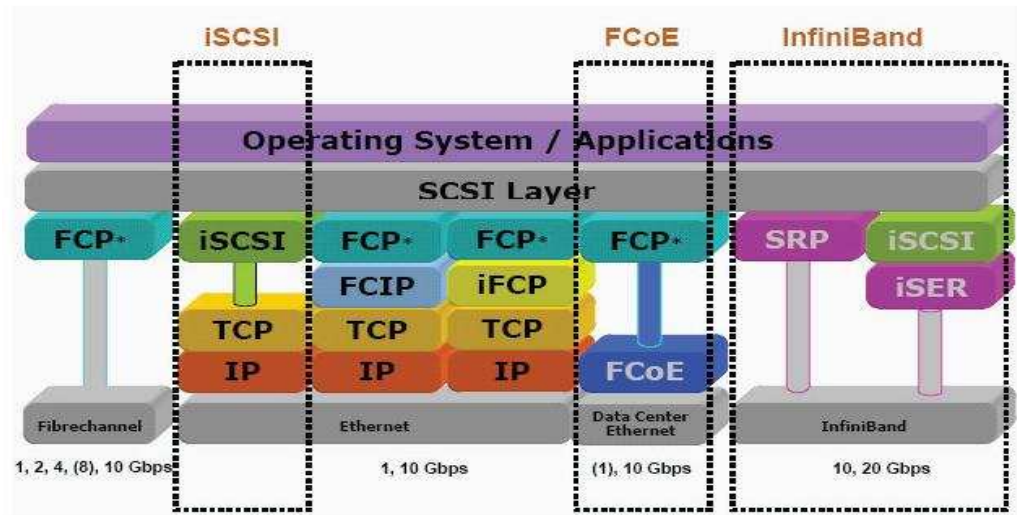


Figura 23: Comparação entre as pilhas iSCSI e FCoE.

Em termos de eficiência nas redes SAN-FC, para um futuro próximo, pode-se esperar o emprego de novas velocidades, aliadas a custos menores, conforme mostra a figura 24.

Fibre Channel Bandwidth: Fastest Speeds, Lowest Cost!

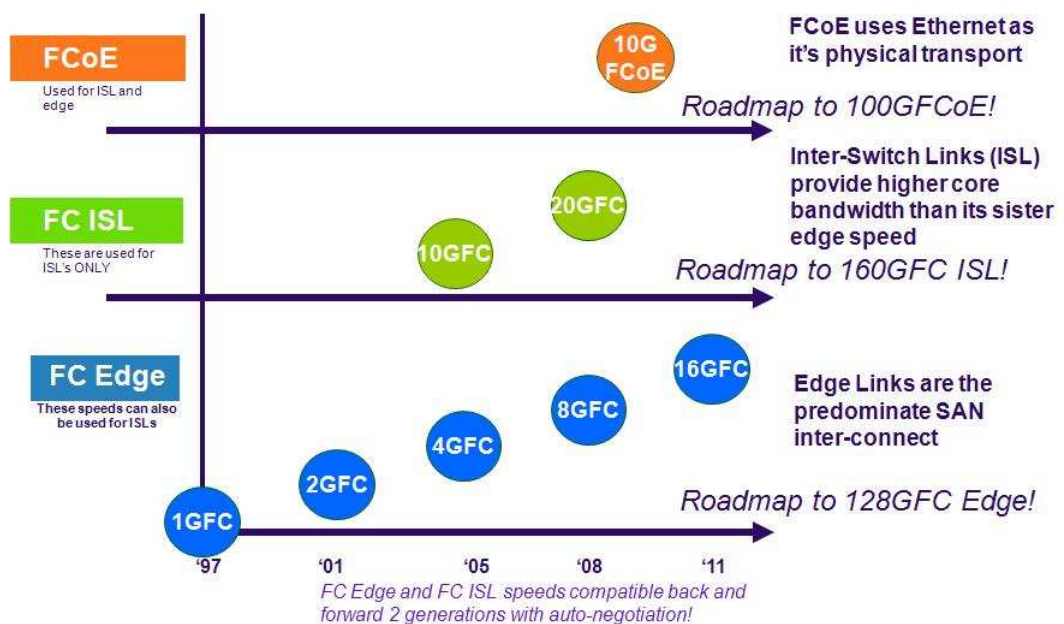


Figura 24: Roadmap panorama das velocidades no FC.

3 VISÃO PRÁTICA DE AMBIENTE SAN DE ALTA DISPONIBILIDADE

O ambiente SAN-FC descrito aqui compõe o *Datacenter*⁴ da instituição Justiça Federal do Rio Grande do Norte (JFRN).

A rede SAN provê os serviços de alta velocidade no transporte dos dados que fazem parte do acervo processual da Justiça Federal RN aliada a integridade, escalabilidade e eficiência na comunicação entre os dispositivos redundantes que precisam estar sincronizados como por exemplo os *storages* de produção e de cópia das informações.

Os dispositivos ligados a rede SAN da JFRN são mostrados na tabela 8 a seguir.

Qte.	Equipamento	Marca – Modelo	Função
01	B32SW01	IBM – B32	Switch ligado ao fabric JFRN
01	B32SW02	IBM – B32	Switch ligado ao fabric JFRN
01	B16SW01	IBM – B16	Switch ligado ao fabric JFRN
01	B16SW02	IBM – B16	Switch ligado ao fabric JFRN
01	Storage	IBM – DS4700 + 2 Exp810	Armazenamento de dados
01	Storage	IBM – DS4700 + 2 Exp810	Cópia síncrona dos dados
01	Tape Library	IBM – TS3310	Backup dos dados em fita LTO3
01	Tape Library	IBM – ULT3582	Será configurada posteriormete para ser o DRM*
08	Servidores de Banco de Dados e de Aplicações	IBM – Risc P520 IBM – X3650, Xseries336 DELL,	Bancos de Dados (Oracle, SQL Server, Postgress), Servidor de Arquivos, Sistemas Processuais, Gerenciador de Backups
09	Cluster de Vmware	HP – DL380G5	Máquinas Virtuais de aplicações como Terminal Service, Ponto Eletrônico, DNS secundário, Intranet, etc.

*DRM: *Desaster Recovery Manager*

Tabela 8: Dispositivos da rede SAN *fibre channel* da JFRN.

Os servidores de aplicação e de banco de dados (*Oracle*, *PostgreSQL*, *SQLServer*) acessam os *storages* e suas expansões segundo as configurações de *initiator* e *target* para realizarem suas conexões ponto a ponto destinadas a transferência dos dados pela rede SAN. A figura 25 mostra como os dispositivos comunicantes se ligam ao *fabric*.

⁴ *Datacenter* ou *Data Center* – Conforme (VERAS, 2009), é um conjunto integrado de componentes de alta tecnologia que permite fornecer serviços de infraestrutura de valor agregado, tipicamente processamento e armazenamento de dados, em larga escala, para qualquer tipo de organização. É o elemento central da Infraestrutura de TI (Tecnologia da Informação).

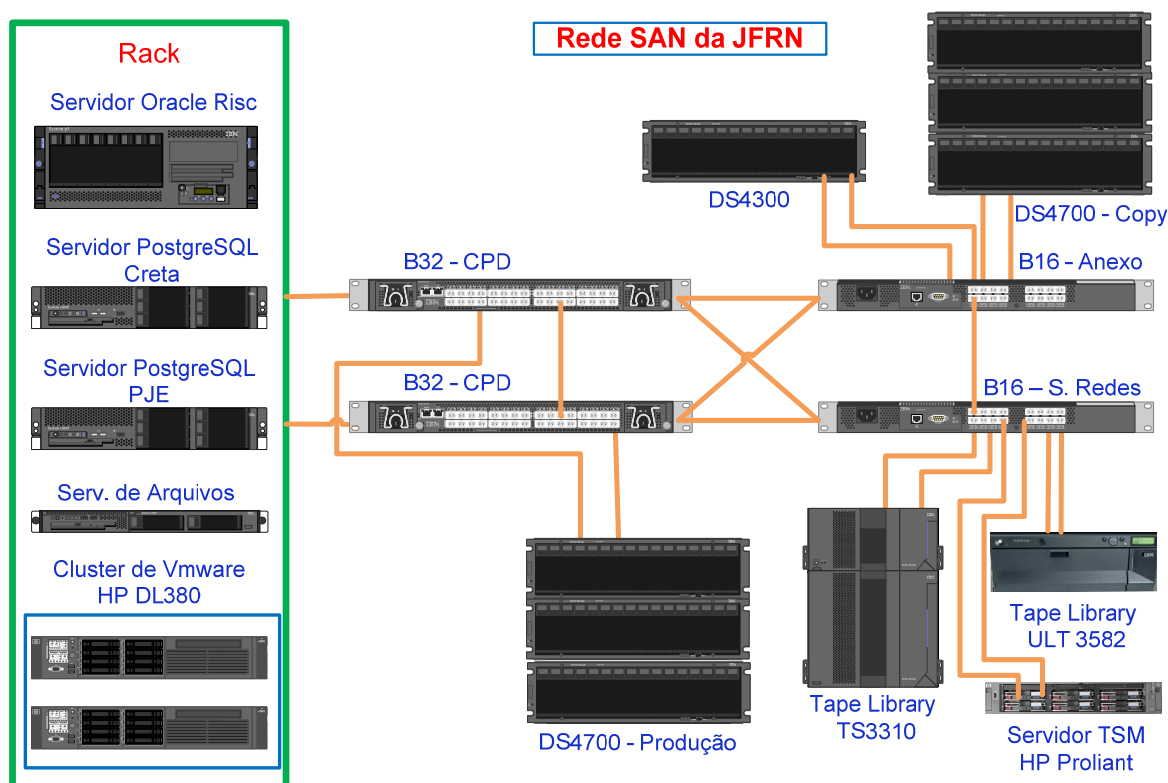


Figura 25: Rede SAN JFRN.

Atualmente o *fabric* SAN da JFRN é composto por 4 *switches fibre channel* sendo dois do modelo IBM B32 com 32 portas *fibre channel* cada um e dois do modelo IBM B16 com 16 portas *fibre channel* cada um, com SFPs funcionando a 4 Gbps.

Os *switches* foram alocados fisicamente em salas diferentes para que os servidores de produção e de redundância pudessem participar do mesmo *fabric* SAN, convergindo para uma solução segura com separação distinta dos ambientes de armazenamento de dados.

SFP (*Small Form Factor Pluggable Media*), segundo (PANTOJA, 2010), também chamado de GBIC, é um *transceiver* (dispositivo transceptor), pertencente a camada física, que combina um transmissor e um receptor numa mesma interface eletrônica, para converter um tipo de sinal em outro. Na rede SAN-FC em questão, os SFPs convertem sinais elétricos em ópticos utilizando conectores do tipo LC (*Lucent Connector*) e fibra óptica multimodo para as velocidades já mencionadas de 4 Gbps. A figura 26, a seguir, mostra os dispositivos ópticos utilizados na SAN-FC da JFRN.



Figura 26: SFP 4Gbps e fibra óptica multimodo de padrão LC.

O ISL (interligação entre os switches) foi realizado na topologia mesh total utilizando as portas de expansão (porta_E) para a interconexão entre os switches. Esse arranjo da rede SAN pode ser visto na figura 27 mostrando as portas utilizadas na interligação e as salas onde se encontram cada switch.

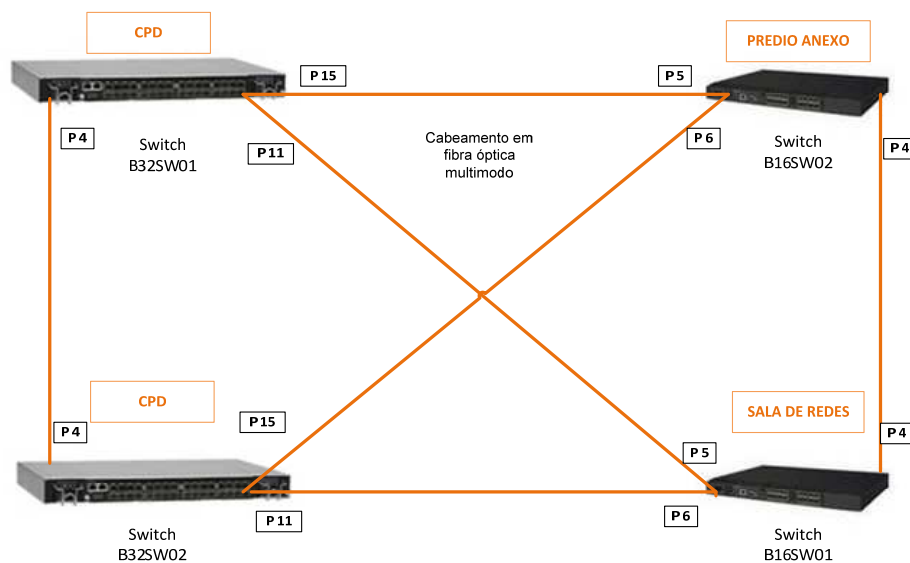


Figura 27: Topologia mesh total – Fabric SAN JFRN.

Dois *switches* B32 se encontram na sala do CPD, um B16 na sala de Redes do prédio principal e um outro B16 na sala de Redes do prédio Anexo.

As conexões são todas em fibra óptica multimodo e os SFPs (*transceivers* – mini GBICs) podem funcionar a uma distância de até 500 metros entre as salas. A comunicação é *Full-Duplex*, a velocidade na rede SAN é 4 Gbps e o gerenciamento dos switches é feito através de interface web própria do *fabric* IBM utilizando acesso pela rede *Ethernet* com cabo UTP⁵ e conexão RJ 45.

O Zoneamento na rede SAN foi implementado usando o endereçamento de 64 bits wwn por motivos de facilidade e flexibilidade das conexões dos dispositivos aos *switches fibre channel*. Para tanto, foram seguidas as melhores práticas de configuração da interface IBM.

Para melhor identificação dos membros componentes da rede, na tabela de zoneamento do *fabric* IBM SAN, foram criados *alias* baseados nos *ports* wwn inerentes às hbas de cada dispositivo. Cada entidade na rede possui normalmente duas interfaces hbas conectadas aos *switches* por questões de redundância e balanceamento de carga. Um driver de software chamado RDAC normalmente é instalado nos servidores Linux e Aix, que se comunicam com os *storages*, a fim de fazer o reconhecimento, por parte destes, ao apontamento das unidades lógicas (LUNs), que correspondem às áreas lógicas nos *storages*, e permitir funções de balanceamento de carga. As figuras 28, 29 e 30, mostram a criação dos *Alias*, a criação dos *Zones* e o *Zone Config* (zona de configuração) que foi denominado SAN_JFRN e corresponde ao pool de zonas criadas que está habilitado para funcionamento no *fabric*. Nessa arquitetura pode-se ter vários *Zone Config* mas, somente um, pode estar habilitado por vez. Os termos ZD e ZT, mostrados na terceira tela da figura 28, referem-se às Zonas de Dados e Zonas de *Tape* respectivamente.

⁵ UTP - Unshielded Twisted Pair – par trançado sem blindagem

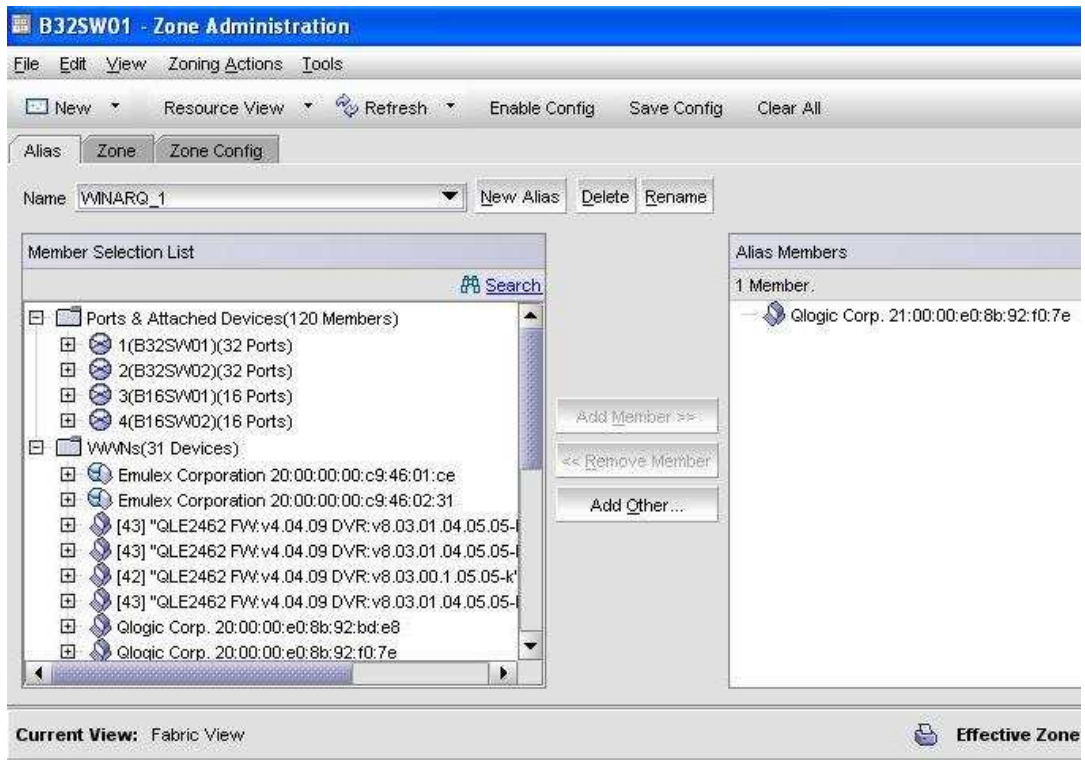


Figura 28: Interface web de gerenciamento do switch B32SW01 aba *alias*.

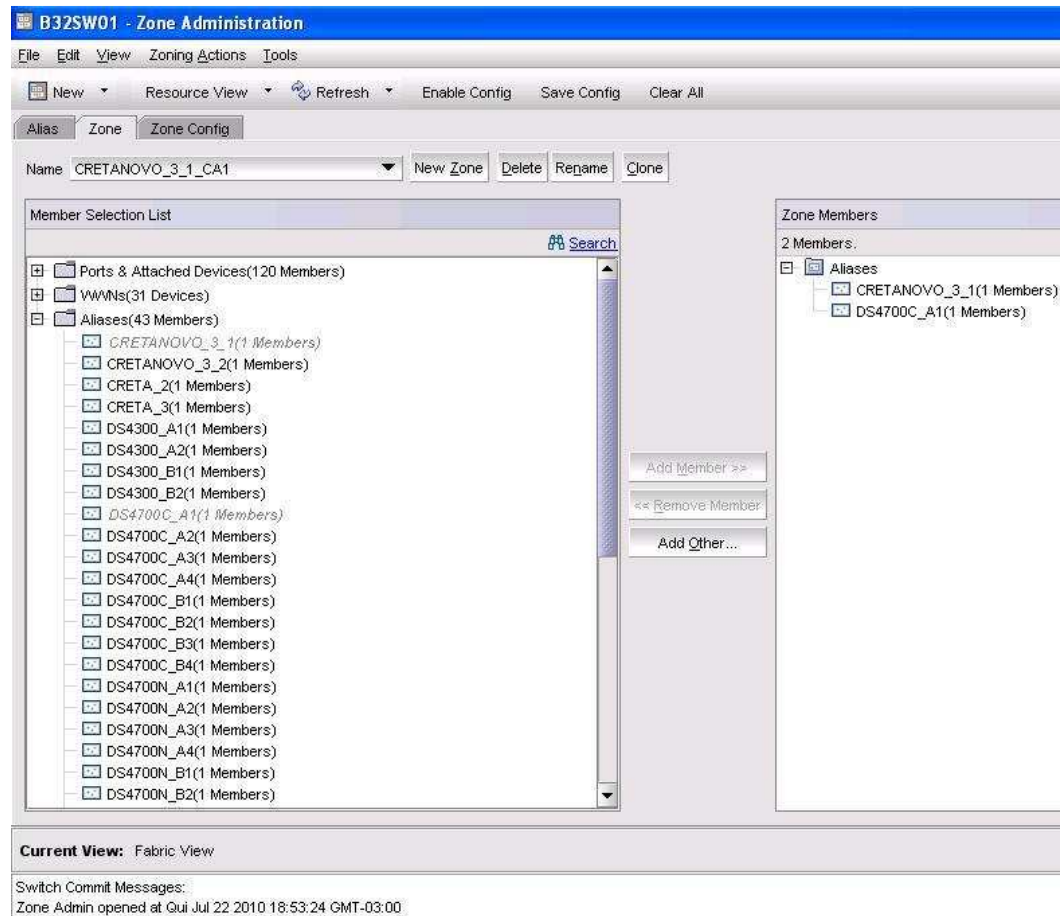


Figura 29: Interface web de gerenciamento do switch B32SW01 aba zone.

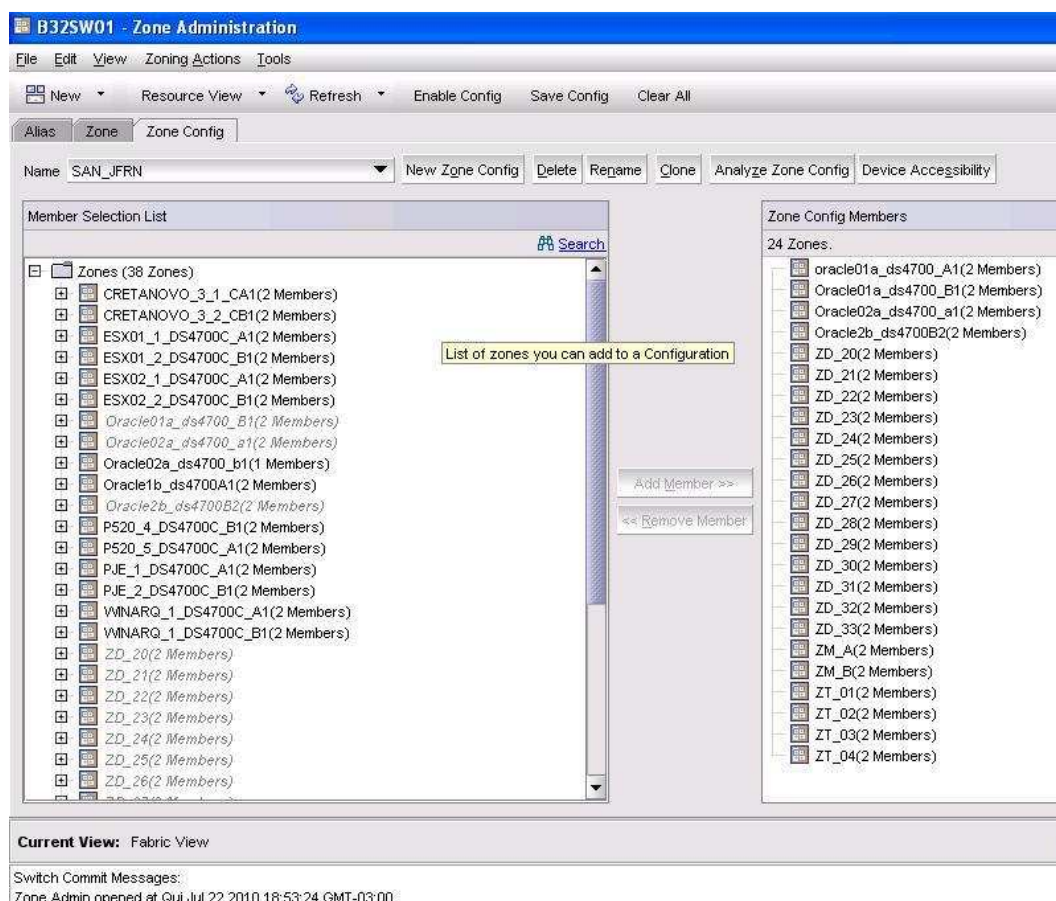


Figura 30: Interface web de gerenciamento do switch B32SW01 aba zone config.

Com o intuito de garantir alta disponibilidade das informações processuais, os *storages* foram configurados para funcionar com espelhamento síncrono onde os dados gravados no *storage* de produção são automaticamente copiados no *storage* secundário ou de redundância dos dados. Eles estão disponibilizados na rede SAN evitando qualquer perda de informação em caso de algum desastre ou problema de *hardware*. Cada *storage* possui duas unidades de expansão totalizando 48 discos em cada conjunto como pode ser visto na figura 29. As controladoras dos *storages* são ligadas ao *fabric* cada uma com duas hbas para garantir redundância e duas interfaces *ethernet* ligadas à rede LAN para gerenciamento.

As unidades controladoras implementam o padrão *raid* para segurança dos dados em caso de perda ou dano em algum HD. Os padrões *raid* 5 e *raid* 10 foram construídos em cada *array* de discos dos *storages* de produção e de cópia e em suas expansões, respeitando as melhores práticas de implementação de *raid* para servidores de banco e de aplicação. Os *storages* de produção e o de cópia dos dados, dispõem de até 3 discos *hot-spare*: um no módulo principal e os outros dois nas expansões, de cada *storage*, que ficam à espera de quebra de algum disco dos *arrays*. Neste processo de quebra de algum disco, os *storages*

também permitem a tecnologia de *hot-swapping*, ou seja: troca de disco a quente, com o sistema em funcionamento. Quando o sistema de *storage* identifica algum dano de HD (*hard disk*), automaticamente insere o disco *hot-spare* no *array* correspondente e este passa a fazer parte do padrão *raid* afetado de forma transparente aos usuários. Com finalidade de melhor performance de I/O e precaução à perda de dados em decorrência de falhas de energia, os *storages* são providos de memória *cache* de leitura e escrita protegida por bateria interna.

Recentemente foi acrescentada a segunda gaveta de discos (IBM Exp810) aos *storages* de produção e de cópia com 16 discos de 420 GB cada, interligados aos módulos principais conforme mostra a figura 31. Nestas gavetas foram criados dois *arrays* sendo 1 com 4 discos no padrão *raid* 10 e outro com 11 discos no padrão *raid* 5. O primeiro *array*, com *raid* 10, será utilizado como área de disco para instalação de dois servidores de banco de dados Oracle e o segundo conjunto de discos para outras aplicações que exijam menos performance de I/O.

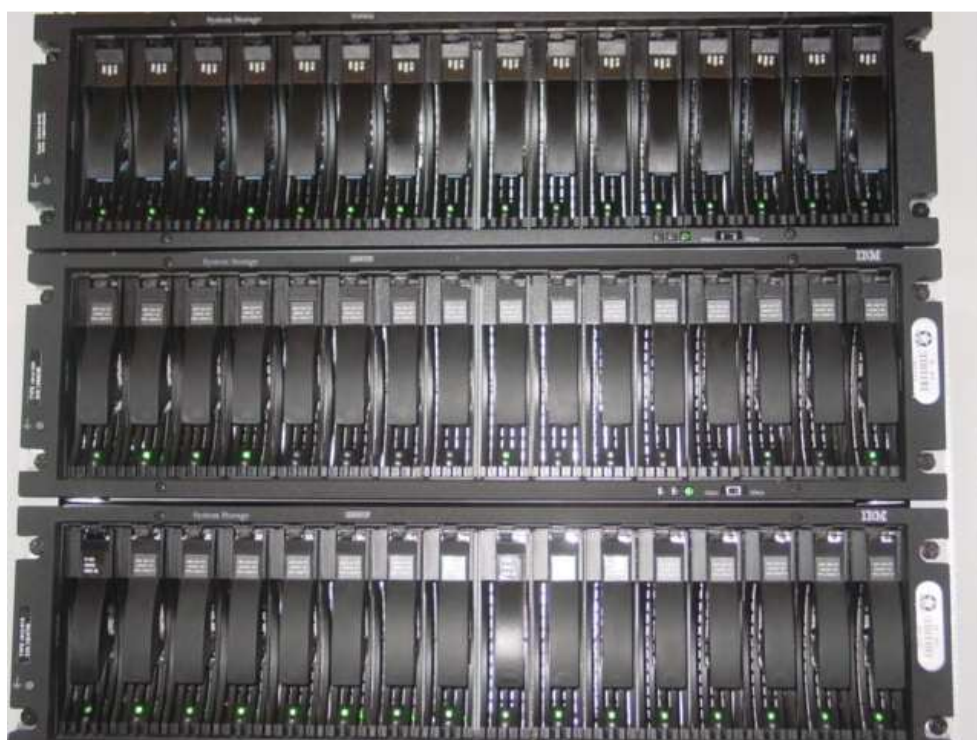


Figura 31: Storage IBM DS4700 e unidades de expansão Exp810.

O gerenciamento dos *storages* é feito através da interface de *software* da própria fabricante IBM onde se visualizam os *arrays*, padrões de *raid*, uso de *hot-spare*, monitoração de I/O, verificação do sincronismo entre os *storages*, etc. As figuras 32, 33 e 34 mostram a interface de gerenciamento dos sistemas de *storage* da JFRN.

NewDS4700 Optimal

Summary Logical **Physical** Mappings Setup Support

Physical

Enclosure 85 - Fibre FC

[View Enclosure Components](#)

(front)

(back)

A.

B.

Enclosure 0 (front)- [Show](#) [Show_FC](#)

[View Enclosure Components](#)

Enclosure 1 (front)- [Show](#) [Show_FC](#)

[View Enclosure Components](#)

Properties

Controller in Enclosure 85, Slot A

[View Complete Controller Properties](#)

Base controller properties

Status: Online

Current configuration

Firmware version:	07.60.28.00
Appware version:	07.60.28.00
Bootware version:	07.60.28.00
NVSRAM version:	N1814D47R1060V04

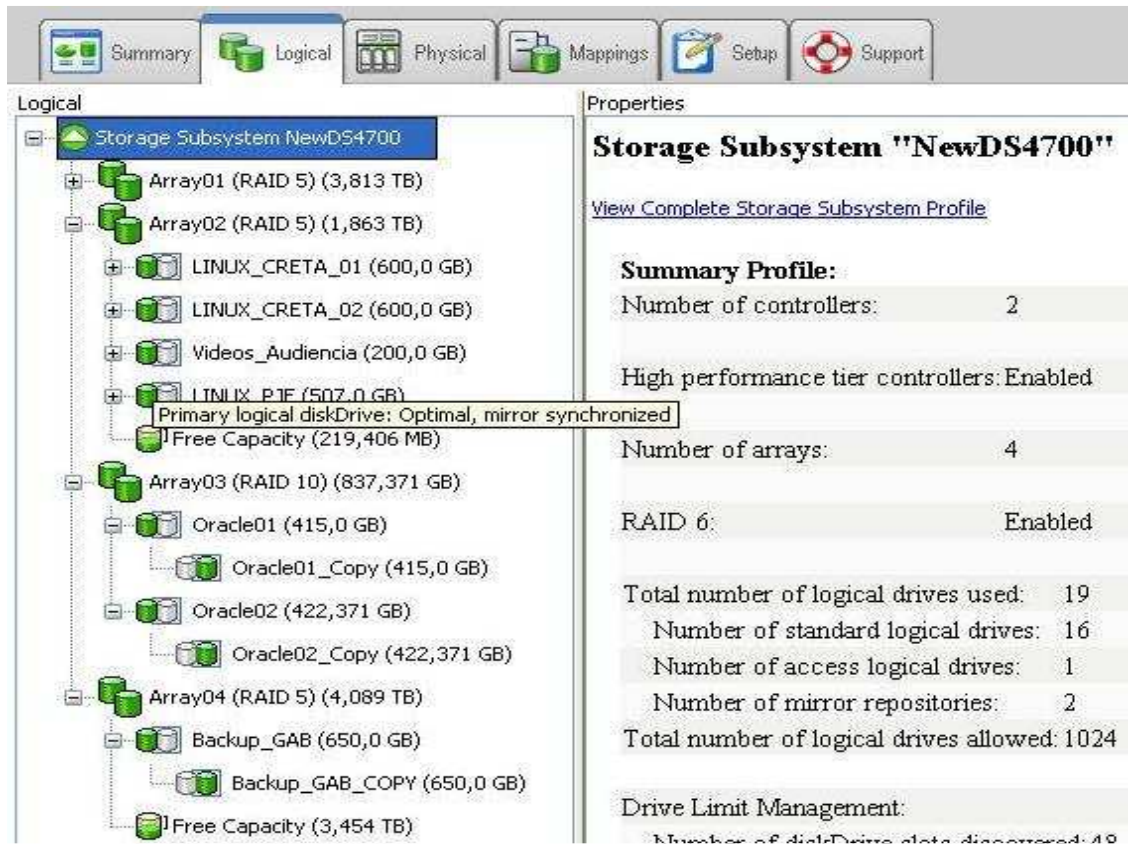
Pending configuration

Firmware version:	None
Appware version:	None
Bootware version:	None
NVSRAM version:	None
Transferred on:	None

Current ID (ALPA)

On diskDrive channel 1:	125/0x1
Replacement part number:	44X2426
Model name:	3994
Board ID:	2004

Figura 32: Interface de gerenciamento do sistema de storage aba *physical*.



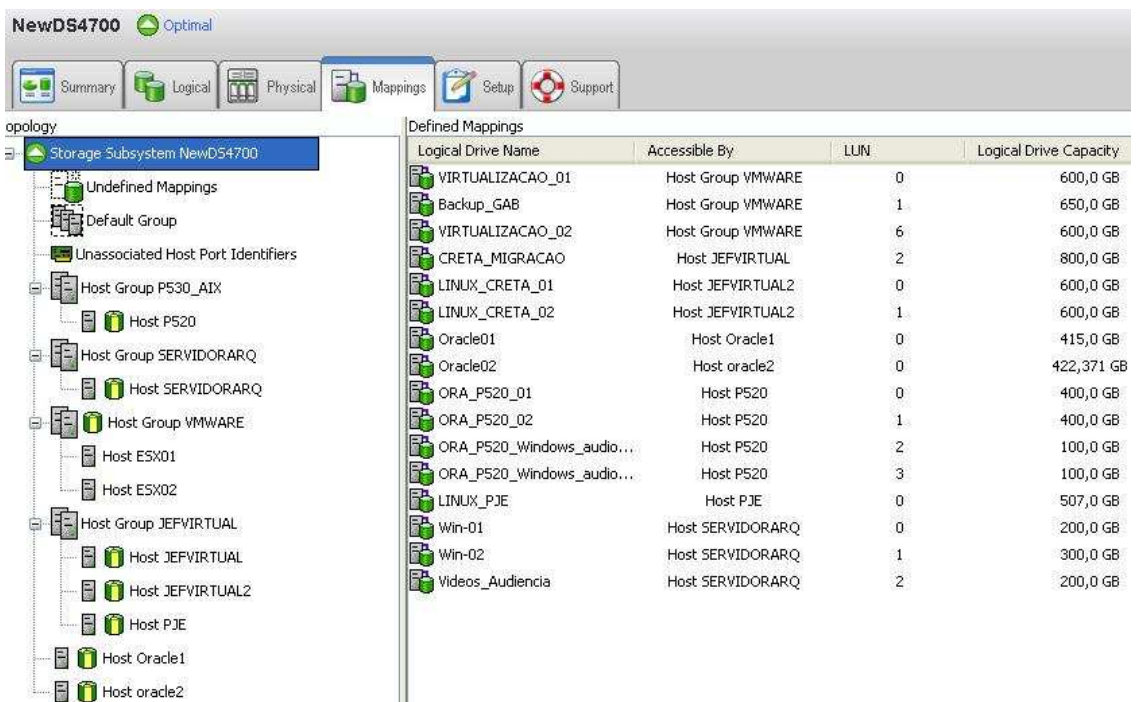
Storage Subsystem "NewDS4700"

[View Complete Storage Subsystem Profile](#)

Summary Profile:

Number of controllers:	2
High performance tier controllers:	Enabled
Number of arrays:	4
RAID 6:	Enabled
Total number of logical drives used:	19
Number of standard logical drives:	16
Number of access logical drives:	1
Number of mirror repositories:	2
Total number of logical drives allowed:	1024
Drive Limit Management:	Number of disk drives detected: 48

Figura 33: Interface de gerenciamento do sistema de storage aba *logical*.



Defined Mappings

Logical Drive Name	Accessible By	LUN	Logical Drive Capacity
VIRTUALIZACAO_01	Host Group VMWARE	0	600,0 GB
Backup_GAB	Host Group VMWARE	1	650,0 GB
VIRTUALIZACAO_02	Host Group VMWARE	6	600,0 GB
CRETA_MIGRACAO	Host JEFVIRTUAL	2	800,0 GB
LINUX_CRETA_01	Host JEFVIRTUAL2	0	600,0 GB
LINUX_CRETA_02	Host JEFVIRTUAL2	1	600,0 GB
Oracle01	Host Oracle1	0	415,0 GB
Oracle02	Host oracle2	0	422,371 GB
ORA_P520_01	Host P520	0	400,0 GB
ORA_P520_02	Host P520	1	400,0 GB
ORA_P520_Windows_audio...	Host P520	2	100,0 GB
ORA_P520_Windows_audio...	Host P520	3	100,0 GB
LINUX_PJE	Host PJE	0	507,0 GB
Win-01	Host SERVIDORARQ	0	200,0 GB
Win-02	Host SERVIDORARQ	1	300,0 GB
Videos_Audiencia	Host SERVIDORARQ	2	200,0 GB

Figura 34: Interface de gerenciamento do sistema de storage aba *mappings*.

O *backup* das informações é realizado através do TSM - Tivoli Storage Manager - que gerencia as políticas de backup, retenções dos dados e os *schedules* (agendamentos) diários. O TSM atua em conjunto com o robô de *backup* (IBM ULT-3582) que agrega uma quantidade de 128 fitas do tipo LTO3 (Fita Linear Aberta 3), fita que armazena até 400/800 GB de informações sem/com compressão.

Os servidores de banco de dados, de arquivos e de aplicações, que utilizam os recursos da rede SAN, estão instalados com o agente de *backup* do TSM e configurados conforme os *scripts* de *backup* para o transporte das informações ao pool de discos do servidor TSM. Este, por sua vez, reenvia os dados para serem gravados em fita.

Os agendamentos para os bancos de dados e servidores de aplicações prevêm *backups* diários que são realizados à noite, fora de expediente, e *backups* dos logs de transação realizados de duas em duas horas diariamente. Os *schedules* também estabelecem tempos para *backups*, em determinadas horas do dia, para o servidor de arquivos e *logs* de bancos, contribuindo para que se tenha menor perda possível de informações em caso de acontecimentos desastrosos. Nos finais de semana, são realizados *backups* a frio dos bancos de dados, ou seja, com os bancos fora do ar, e *backup* de alguns servidores de aplicação, a fim de manter em fita uma segunda opção de restauração de dados caso haja necessidade em virtude de avarias ou desastres. Há também uma política de *backup* de se guardar várias versões dos dados para dar uma garantia maior na recuperação das informações. A figura 35 mostra as *libraries* de *backup* e o servidor TSM (agente gerenciador de *backup*).



Figura 35: Rack com Tape library IBM ULT3582, Servidor TSM e Tape library TS3310.

Há algumas melhorias a serem implementadas na rede SAN e de *backup* da JFRN para uma maior eficiência no transporte e segurança dos dados como a padronização e ampliação dos *switches* que compõem o *fabric* para uma velocidade maior de 8 Gbps com instalação de mais dois *switches* nas salas de redes e prédio anexo melhorando a redundância de dispositivos nessas salas; instalação de assinaturas de software para realização de *backup* integral pela rede SAN, uma vez que atualmente é realizado em conjunto com a rede *Ethernet* em 1 Gbps; retirada de fitas do *backup* para serem guardadas em cofre existente em localidade anexa ao prédio da JFRN que hoje não ocorre; instalação de *site backup*.

4 CONCLUSÃO

O trabalho em questão reflete o uso dos protocolos *fibre channel* aliado às novas tecnologias de redes atualmente empregadas nos *Datacenters* espalhados pelo mundo. Sobretudo representa um estudo complexo de uma tecnologia versátil e segura que se utiliza de um protocolo já existente – o SCSI – para dar um salto de qualidade no transporte das informações.

Redes SANs baseadas em protocolos *fibre channel* com cabeamento e dispositivos ópticos são ainda a forma mais segura de se guardar os dados em um *datacenter* quando se quer oferecer um serviço de alta disponibilidade.

Apesar disso, novas tendências dos protocolos *fibre channel* se preocupam em utilizar os benefícios da rede Ethernet. Os protocolos FCIP, iFCP e FCoE são baseados em FCP e há uma tendência emergente: o iSCSI que usa somente os protocolos da rede *Ethernet*.

As polêmicas já começaram a respeito de qual seria o mais eficiente em transporte de dados, segurança, flexibilidade e escalabilidade. Os concorrentes mais visíveis nessa discussão são o FCoE e o iSCSI.

O trabalho mostra também uma visão prática de ambiente SAN implementado na Justiça Federal do Rio Grande do Norte. Uma rede SAN de alta disponibilidade neste caso é requerida em virtude dos sistemas processuais existentes necessitarem estar disponíveis 24 horas por dia e 365 dias por ano devido aos prazos processuais.

Esse estudo reflete também o aprendizado da Pós Graduação em Redes de Computadores que forneceu os subsídios e recursos técnicos e científicos necessários à produção do trabalho.

Os professores do curso de Pós Graduação em Redes de Computadores da FARN, merecem especial destaque pela disseminação das idéias inerentes às redes de computadores fomentando o aprimoramento dos estudos através dos trabalhos de final de curso.

REFERÊNCIAS BIBLIOGRÁFICAS

Veras, Manoel, “Datacenter: componente central da infraestrutura de TI”, Editora Brasport, 2009.

Tanenbaum, Andrew S., “Redes de Computadores”, 4ª edição, Editora Campus, 2003.

Matthews, Jeanna, “Rede de Computadores – Protocolos de Internet em Ação”, Editora LTC, 2006.

Long, James, “*Storage Networking Protocol Fundamentals*”, Cisco Press, 2006.

Wikipedia, “*Network File System (protocol)*”, disponível em [http://en.wikipedia.org/wiki/Network_File_System_\(protocol\)](http://en.wikipedia.org/wiki/Network_File_System_(protocol)), acessado em 03 de abril de 2010.

VisioCafe, “*Official Corporate Collections*”, disponível em <http://www.visiocafe.com>, acessado em junho/julho de 2010.

Morimoto, Carlos E., “Armazenamento de rede: DAS, NAS e SAN”, 2007, disponível em <http://www.guiadohardware.net/tutoriais/das-nas-san/>, acessado em abril de 2010.

NFS, “*Network File System*”, disponível em <http://www-di.inf.puc-rio.br/~endler/courses/inf1019/transp/nfs.pdf>, acessado em 03 de abril de 2010.

Barreiros, Caio C., “O Samba”, disponível em http://www.gta.ufrj.br/grad/01_2/samba/samba.htm, acessado em 03 de março 2010.

Evans, Timothy, “NetBIOS, NetBEUI, NBF, NBT, NBIPX, SMB, CIFS Networking”, 2003, disponível em <http://timothydevans.me.uk/nbf2cifs/nbf2cifs.pdf>, acessado em 04 de abril de 2010.

Cisco-SAN1501, “*Introduction-to-Storage-Area-Networking*”, 2006.

BlackBox, “*Tecnologia Fibre Channel*”, 2002, disponível em <http://www.blackbox.com.br/html/bboxexplica/fibrechannel.htm>, acessado em 10 de janeiro de 2010.

DEFINITION-ESCON/FICON, 2000, disponível em http://searchstorage.techtarget.com/sDefinition/0,,sid5_gci344040,00.html, IBM, acessado em 04 de maio de 2010.

Racherla, Sangam; Allworth, Bruce; Bagnaresi, Alessio; Bogdanowicz, Chris; Lottering, Corne; Pedrazas, Pablo; Schubert, Frank; Sexton, John; Watson, Alexander, “*IBM Midrange System Storage Implementation and Best Practices Guide*”, ibm.com Redbooks, 2010.

Tate, Jon; Edwards, Gareth; Larsen, Carsten; Myszka, Dariusz; Seiwert, Chris, “*Implementing an IBM/Brocade SAN with 8 Gbps Directors and Switches*”, ibm.com Redbooks, 2009.

Tate, Jon; Lucchese, Fabiano; Moore, Richard, “*Introduction to Storage Area Networks*”, ibm.com Redbooks, 2006.

Brooks, Charlote; Byrne, Frank; Higuera, Leonardo; Krax, Carsten; Kuo, John, “*IBM System Storage Solutions Handbook*”, ibm.com Redbooks, 2006.

FCoE, “O Futuro do Armazenamento de Rede em Canal de Fibra”, disponível em http://www.techczar.com/po/RPC_Content.aspx?CSID=2_20&NID=547, acessado em 12 de julho de 2010.

Costa, Rodrigo M., Artigo “*Fibre Channel e Storage Area Network*”, 2008, disponível em http://www.gta.ufrj.br/grad/08_1/san/index.html, acessado em março de 2010.

Pantoja, Ricardo “*TRANSCEIVER - GBIC / SFP / SFF / XFP / XENPAK*”, 2010, disponível em <http://www.pantojaindustrial.com.br/exibir.php?id=91>, acessado em julho de 2010.