

PONTIFÍCIA UNIVERSIDADE CATÓLICA DO RIO GRANDE DO SUL
FACULDADE DE INFORMÁTICA

DIEGO CEZAR MOLINA

**ESTUDO E IMPLEMENTAÇÃO DO PROTOCOLO DE CONTROLE DE
AGREGAÇÃO DE ENLACES (LACP)**

Porto Alegre
2011

DIEGO CEZAR MOLINA

**ESTUDO E IMPLEMENTAÇÃO DO PROTOCOLO DE CONTROLE DE
AGREGAÇÃO DE ENLACES (LACP)**

Trabalho de Conclusão de Curso de graduação
como requisito para obtenção do grau de
Bacharel em Ciência da Computação.

Orientador: Prof. Dr. Fernando Gehm Moraes

Porto Alegre
2011

ESTUDO E IMPLEMENTAÇÃO DO PROTOCOLO DE CONTROLE DE AGREGAÇÃO DE ENLACES

RESUMO

O Protocolo de Controle de Agregação de Enlaces é usado para otimizar a transferência de dados entre sistemas de telecomunicações interconectados por múltiplos enlaces físicos. Este protocolo define e controla Grupos de Agregação de Enlaces. Além disso, gerencia o fluxo de dados entre sistemas, possibilitando a melhor utilização da largura de banda disponível em cada enlace através de mecanismos de balanceamento de carga e de distribuição de dados entre enlaces.

Em caso de falha de algum enlace físico ou porta, os dados que estavam sendo transmitidos por estes podem ser desviados para outra porta, aumentando assim a disponibilidade de transferência de dados entre sistemas. Para isso, o padrão 802.1AX especifica a comunicação entre equipamentos terminais de dados que são interconectados por enlaces que operam em *full-duplex* e a uma mesma taxa de transferência.

O presente trabalho tem por objetivos o estudo desse protocolo e sua implementação. Este trabalho também proporcionará experiência ao Autor em participação em um projeto de pesquisa e desenvolvimento em parceria com empresa do ramo de telecomunicações.

Palavras-chave: Enlace Agregado, Agregador, Agregação de Enlaces, Grupo de Agregação de Enlaces, Redes Locais, Gerência

STUDY AND IMPLEMENTATION OF THE LINK AGGREGATION CONTROL PROTOCOL

ABSTRACT

The Link Aggregation Control Protocol is used to optimize the data transfer between telecommunications systems interconnected by multiple physical links. This protocol defines and controls Link Aggregation Groups. Moreover, manages the flow of data between systems, enabling better use of available bandwidth on each link through mechanisms of load balancing and data distribution between links.

In case of failure of a physical link or port, the data were being transmitted by them can be diverted to another port, thereby increasing the availability of data transfer between them. For this, the standard 802.1AX specifies communication between data terminal equipment that are interconnected by links operating in full-duplex and the same throughput.

This work aims of the study of the protocol and its implementation. It will also provide experience to the Author of participation in a research and development project in partnership with a telecommunications company.

Keywords: Aggregated Link, Aggregator, Link Aggregation, Link Aggregation Group, local area network, management

SUMÁRIO

1	INTRODUÇÃO	7
1.1	Contextualização do LACP em sistemas de telecomunicações	7
1.2	Objetivos e características do LACP.....	8
1.3	Objetivos do trabalho	9
1.4	Estrutura do documento	9
2	MÓDULOS COMPONENTES DO LACP.....	10
2.1	Agregador	11
2.2	Controle de Agregação	12
2.3	Analizador e Multiplexador do Controle de Agregação (CPM).....	12
2.4	Protocolo Marcador.....	13
3	O CONTROLE DE AGREGAÇÃO DE ENLACES	14
3.1	Características do LAC.....	14
3.2	Identificadores.....	14
3.3	Controle de Agregação	17
4	INFRAESTRUTURA PARA A IMPLEMENTAÇÃO DO PROTOCOLO	21
4.1	Visão macro da operação do LACP.....	21
4.2	Variáveis de controle de estado e sua organização	22
4.3	Transmissão de pacotes de controle.....	30
5	DESCRIÇÃO DAS MÁQUINAS DE ESTADO DE CONTROLE.....	32
5.1	Modelo genérico de implementação das máquinas de estados	32
5.2	Máquina Receptora.....	33
5.3	Máquina Transmissora	38
5.4	Máquina de Transmissão Periódica	38
5.5	Lógica de Seleção	40
5.6	Máquina Multiplexadora	42
6	RESULTADOS.....	46
6.1	Simulador.....	46
6.2	Configuração de Teste	46
7	CONCLUSÃO	53

1 INTRODUÇÃO

O presente Capítulo contextualiza o Protocolo de Controle de Agregação de Enlaces (*Link Aggregation Control Protocol* - LACP) na área de telecomunicações (Seção 1.1) bem como suas características (Seção 1.2), e apresenta os objetivos deste Trabalho de Conclusão de Curso. Ao final do Capítulo apresenta-se a estrutura deste documento (Seção 1.4).

1.1 Contextualização do LACP em sistemas de telecomunicações

Uma rede de computadores, doravante denominada simplesmente de rede, é uma coleção de dispositivos, interconectados por canais de comunicação, que possibilita seus usuários compartilharem recursos e informações.

Em conexões típicas entre equipamentos de interconexão como dois *switches* interligados ou um *switch* ligado a um servidor, a troca de quadros entre estes equipamentos não possui nenhum tipo de controle em relação ao uso da largura de banda disponível através dos enlaces que os interligam, nem mesmo em relação às possíveis falhas em alguma porta ou enlace.

A quantidade de informação que é trocada entre estes equipamentos é limitada pela tecnologia que está sendo utilizada, tanto nos controladores das *interfaces*, quanto nos cabos que interligam estes equipamentos. Como essas conexões são ponto-a-ponto, a largura de banda máxima obtida através destas conexões está limitada a largura de banda de somente um enlace.

Como a largura de banda disponível cresce, historicamente, uma ordem de grandeza a cada nova geração de tecnologia [WIK11a], é necessário um mecanismo que otimize a utilização desta largura de banda para melhor proveito desta tecnologia.

Um problema comum nesse tipo de sistema é se uma porta em que um dado enlace está conectado falhar, o tráfego que era transmitido por este enlace ficará interrompido até que a conexão seja reestabelecida, fazendo com que o dispositivo perca desempenho em função de estar impedido de utilizar o enlace.

No início dos anos 90, vários fabricantes introduziram técnicas para combinar enlaces *ethernet* físicos em um único enlace lógico, sendo um exemplo a técnica denominada de *Channel Bonding* [WIK11b]. Entretanto, estas técnicas exigem configurações manuais e equipamentos compatíveis nas duas pontas do enlace.

O LACP apresenta uma solução para os problemas citados anteriormente. Este protocolo foi proposto em 2008 e descreve como utilizar múltiplos cabos/portas de rede em paralelo para criar um único cabo lógico (agregação) que funciona com a capacidade resultante da soma da largura de banda de todos os cabos, resolvendo o problema referente à taxa de transferência que antes era limitada pela velocidade de um único enlace. Como o protocolo trabalha gerenciando o fluxo de múltiplos enlaces físicos, o problema da falha de algum enlace também é resolvido, pois, se algum enlace falhar, o fluxo que antes era transmitido por este enlace será desviado para um outro que faz parte desta mesma agregação.

No exemplo da Figura 1, uma topologia do tipo *switch-switch*, interliga seis computadores de mesa e dois servidores. Os *switches* estão conectados através de quatro enlaces que operam na mesma taxa de transferência (1000 Mb/s). A elipse que circunda os enlaces simboliza uma agregação. Estes enlaces pertencem a um mesmo Grupo de Agregação de Enlaces (LAG) que sendo administrado pelo LACP, possibilita uma taxa de transferência entre ambos os *switches* de até 4000 Mb/s [SYS02].

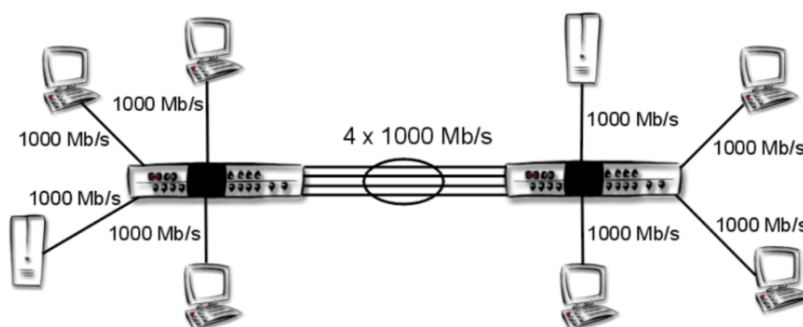


Figura 1 – Exemplo de configuração de uma agregação de enlaces.

1.2 Objetivos e características do LACP

O LACP mantém o fluxo de dados entre os enlaces balanceado, pois o tráfego de cada cliente é distribuído entre os múltiplos enlaces de um determinado LAG. Em caso de mudança na conectividade física (ex.: se algum enlace falhar) o LACP convergirá a uma nova configuração na ordem de um segundo ou menos. Esta convergência pode ser alcançada independentemente da ordem em que os eventos (ex.: conectar ou desconectar algum cabo) ocorram tornando seu comportamento determinístico [IEE08].

O LACP também é compatível com protocolos existentes, como o conjunto de padrões IEEE 802.3, não sendo necessárias mudanças para usar protocolos de camadas mais altas ou aplicações. Caso algum enlace não possa participar de um LAG (ex.: em função de suas configurações ou de alguma restrição na capacidade do dispositivo ao qual está conectado) ele irá operar normalmente sem ser agregado a nenhum outro enlace. Com relação aos quadros IEEE 802.3, o LACP não adiciona dados nem altera seus conteúdos ao manipulá-los. O LACP também especifica objetos de gerência apropriados para sua configuração e controle.

Em relação ao Modelo de Referência OSI [TAN03], a subcamada aonde o LACP está posicionado se encontra entre as subcamadas de Controle Lógico de Enlaces e de Controle de Acesso ao Meio. Por isto, sua operação é transparente para os clientes MAC pois para estes, um LAG (que pode ser composto por vários enlaces) será interpretado como um único enlace.

O posicionamento arquitetural da subcamada aonde o LACP é implementado é apresentada na Figura 2.

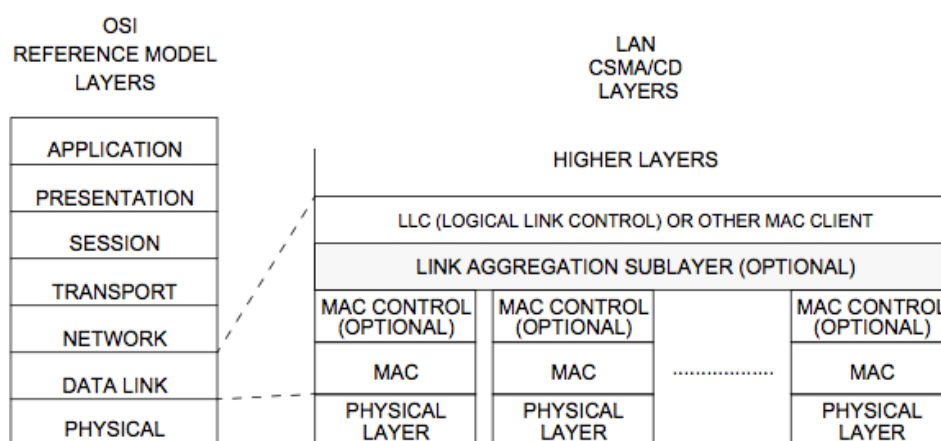


Figura 2 – Posicionamento arquitetural da subcamada aonde o LACP é implementado.

Para o protocolo, o sistema do ponto de vista local é chamado de Ator e o sistema remoto de Parceiro. O Ator mantém informações do seu estado e do estado do Parceiro pois, estas informações são necessárias para o estabelecimento e manutenção das agregações. Os sistemas trocam estas informações através de Unidades de Dados do Protocolo de Controle de Agregação de Enlaces (LACPDUs) que é composto pelas informações do Ator e pelas informações que o Ator possui sobre seu Parceiro. O LACP entra em um estado estável somente quando o Ator possuir as informações corretas sobre seu Parceiro e vice-versa. A partir deste momento a transmissão de dados é habilitada. O LACP descreve máquinas de estado que controlam a validade das informações trocadas entre os sistemas e que são responsáveis pelo controle do estado lógico das variáveis de controle utilizadas pelo Controle de Agregação (*Link Aggregation Control - LAC*). Cabe também a estas máquinas de estado tratar os LACPDUs vindos do sistema Parceiro, salvando suas informações localmente para sua posterior avaliação pelo LAC.

1.3 Objetivos do trabalho

O presente trabalho tem por objetivos o estudo e a implementação do LACP para consolidar e exercitar as diversas áreas de conhecimento que foram estudadas ao longo do curso de graduação do Autor. Além disto, este trabalho será utilizado nos equipamentos da empresa PARKS S/A que em conjunto com a PUCRS formam um convênio.

1.4 Estrutura do documento

O Capítulo 2 apresenta os principais módulos que compõe o LACP, destacando a função de cada um, quais são implementados em *hardware* e quais são implementados em *software*. O Controle de Agregação de Enlaces (LAC), módulo do LACP implementado em *software*, é o alvo do presente trabalho. As funcionalidades do LAC são apresentadas no Capítulo 3. O Capítulo 4 apresenta uma visão geral da implementação do LAC, assim como as estruturas de dados necessárias para implementá-lo. O Capítulo 5 corresponde à maior contribuição do trabalho, que é o detalhamento da implementação do LAC. O Capítulo 6 apresenta resultados referentes à implementação do LAC, e o Capítulo 7 conclui o presente trabalho.

2 MÓDULOS COMPONENTES DO LACP

Este Capítulo apresenta os módulos que compõem o LACP, responsáveis pelo controle de entrada e saída de quadros dos clientes MAC. A Figura 3, retirada do documento que descreve o LACP [IEE08], mostra em um diagrama de blocos a estrutura lógica dos módulos que o compõe: Agregador (Seção 2.1), Controle de Agregação (Seção 2.2) e o Analisador e Multiplexador do Controle de Agregação (Seção 2.3).

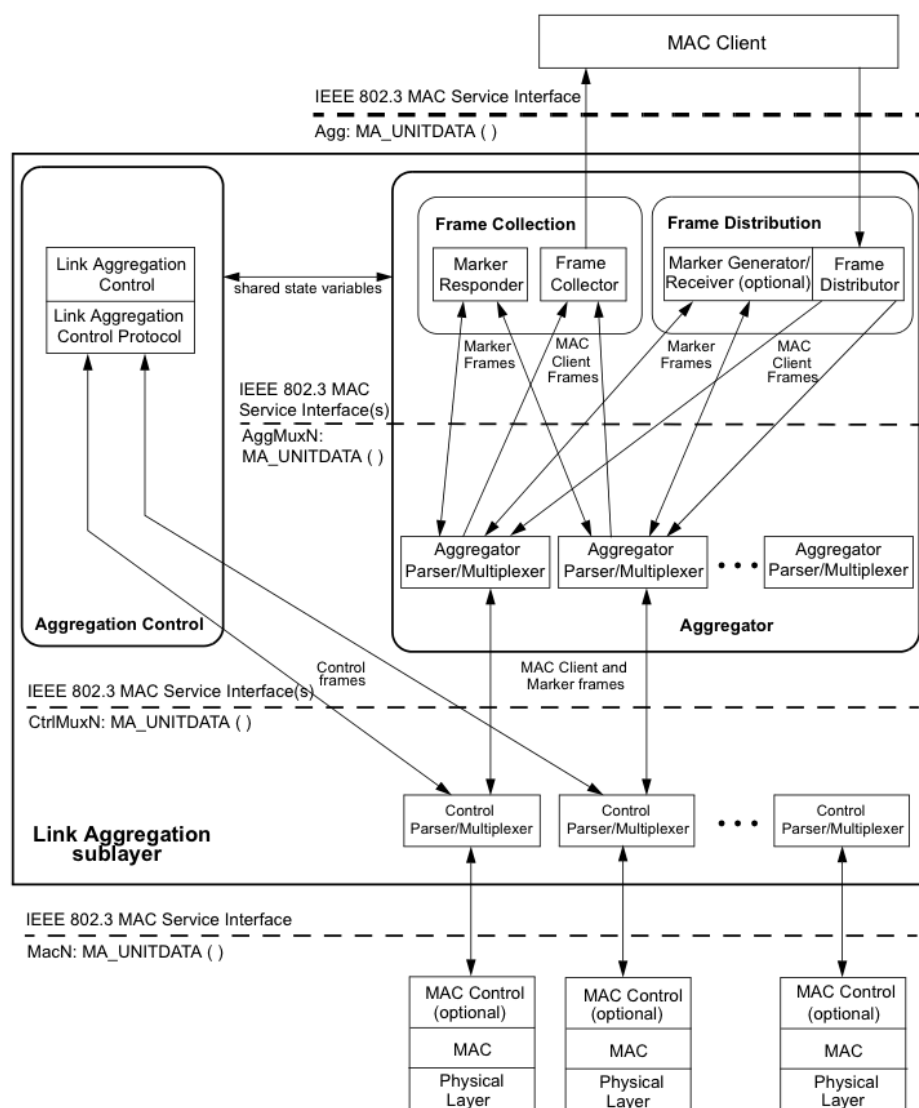


Figura 3 - Diagrama de Blocos da subcamada aonde o LACP é implementado.

A parte superior da figura apresenta o fluxo de informação entre um cliente MAC e um Agregador. A parte inferior apresenta o fluxo de informações entre os enlaces conectados ao sistema Parceiro e o LAC. No centro da figura é apresentada a relação entre os módulos do LACP, que serão descritos no decorrer deste Capítulo, e o fluxo dos quadros que passam por eles. Além disto, apresenta um diagrama de blocos que exemplifica a configuração lógica de uma agregação. Esta agregação é representada pela entidade Agregador (*Aggregator*), localizado no maior bloco do diagrama na parte superior direita, que compartilha variáveis de estado com o bloco de Controle de Agregações (*Aggregation Control*) localizado na parte superior esquerda.

2.1 Agregador

O Agregador gerencia o fluxo de dados de um LAG e é implementado em *hardware* devido a requisitos de desempenho (alta vazão de dados). É composto pelos submódulos Coletor de Quadros (*Frame Collector - FC*), Distribuidor de Quadros (*Frame Distributor - FD*) e o Analisador e Multiplexador do Agregador (*Aggregator Parser and Multiplexer - APM*) descritos nas próximas subseções.

2.1.1 Analisador e Multiplexador do Agregador (APM)

Na recepção dos quadros, o Analisador do Agregador verifica o quadro recebido. Se este for válido, o encaminhará para o FC, caso contrário irá descartá-lo. Na transmissão dos quadros, o Multiplexador do Agregador recebe os quadros do FD e os encaminha à porta especificada, caso ela esteja habilitada para distribuição. Caso contrário, os descarta.

O APM é implementado em *hardware*.

2.1.2 Distribuidor de Quadros (FD)

A distribuição dos quadros através do conjunto de enlaces que compõe o LAG é realizada pelo FD (localizado na parte superior direita do Agregador na Figura 3) que é o módulo responsável por receber os quadros vindos dos clientes MAC.

Este módulo é implementado em *hardware* e utiliza um algoritmo de distribuição de quadros que é responsável por escolher o enlace a ser utilizado na transmissão dos quadros recebidos. Este algoritmo deve garantir que a entrega dos quadros seja feita de maneira ordenada e sem duplicação.

A operação deste módulo é descrita pelo diagrama da Figura 4. Este diagrama possui dois estados, WAIT FOR TRANSMIT (espera para transmissão) e PASS TO PORT (retransmitir para uma porta). O primeiro estado fica esperando um quadro. Quando este chega, o mesmo é transmitido através de um dos enlaces do LAG, de acordo com o algoritmo de distribuição mencionado anteriormente.

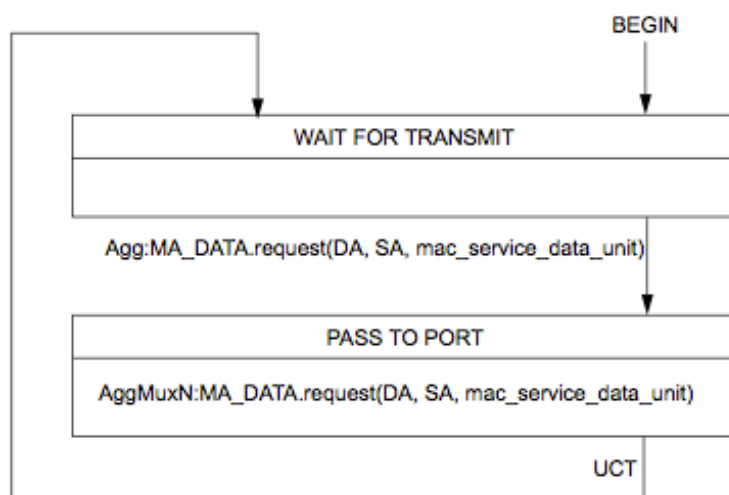


Figura 4 - Diagrama de estados do Distribuidor de Quadros [IEE08].

2.1.3 Coletor de Quadros (FC)

A recepção dos quadros vindos dos enlaces do LAG é realizada pelo FC (localizado na parte superior esquerda do Agregador da Figura 3) que os encaminham ao cliente correto. A Figura 5 demonstra o diagrama de estados do Coletor de Quadros.

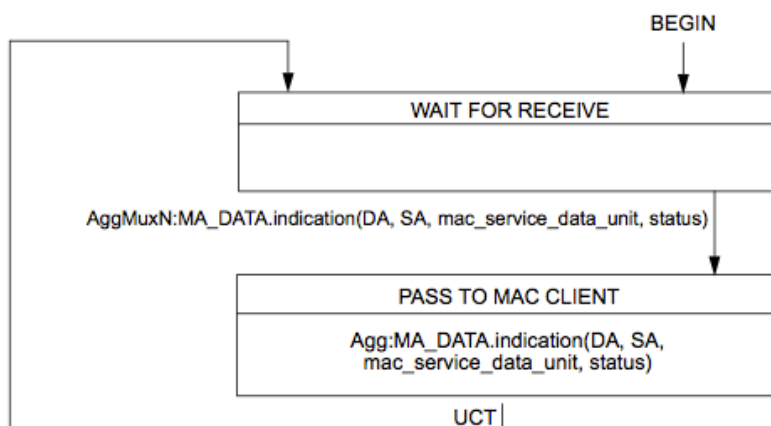


Figura 5 - Diagrama de estados do Coletor de Quadros.

O FC recebe quadros de todos os enlaces deste LAG e os encaminha ao cliente MAC. Este módulo opera de forma inversa ao FD. O estado WAIT FOR RECEIVE descreve o comportamento desta entidade que aguarda a chegada de um quadro de um dos enlaces do LAG ao qual faz parte, quando este quadro é recebido, o diagrama passa para o estado PASS TO MAC CLIENT que reencaminha o quadro ao cliente MAC correto.

Este módulo é implementado em *hardware*.

2.2 Controle de Agregação

É a entidade responsável por configurar e controlar as agregações de enlaces. Este bloco é formado pelo LAC que configura o LACP para comunicação com o sistema Parceiro.

A sua implementação é em software, sendo o módulo que será descrito no presente trabalho (Capítulo 3).

2.3 Analisador e Multiplexador do Controle de Agregação (CPM)

O Multiplexador do Controle de Agregação recebe quadros vindos dos Agregadores e do LAC e os envia à porta apropriada para transmissão ao sistema Parceiro.

O Analisador do Controle de Agregação recebe quadros vindos dos diversos enlaces de um LAG. Quando o quadro for um LACPDU, o mesmo é enviado para o LAC, caso contrário, será enviado ao Agregador correto. A Figura 6 apresenta um diagrama de estados que representa a operação do CPM. Esta entidade é implementada em *hardware*.

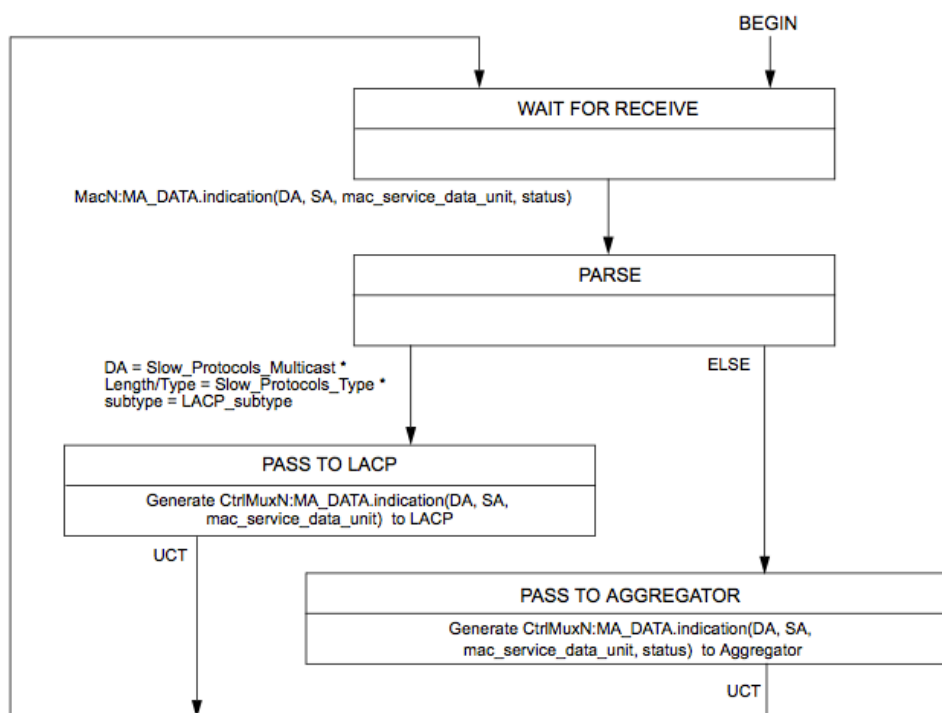


Figura 6 - Diagrama de estados do Analisador do Controle de Agregação.

No estado WAIT FOR RECEIVE, o CPM fica esperando por quadros que podem vir tanto do sistema Parceiro através dos enlaces de algum LAG quanto de um Agregador que retransmite os quadros que recebe dos clientes MAC. Após receber um quadro, o diagrama passa para o estado PARSE que decodifica o quadro. Se este quadro for um LACPDU, o diagrama passa para o estado PASS TO LACP que simboliza o encaminhamento do quadro para o LAC. Se este quadro não for um LACPDU, quer dizer que é um quadro de dados de um cliente MAC e, portanto, será enviado ao agregador responsável pelo controle deste cliente MAC.

2.4 Protocolo Marcador

O Protocolo Marcador (*Marker Protocol*) tem a função de gerar e responder pacotes de controle próprios, os quais são responsáveis pela comunicação entre dois sistemas, a fim de colocar em sincronia as funções de coleta e distribuição de quadros do LACP.

Quando um sistema deseja ativar seu FD ele solicita ao Protocolo Marcador que avise esta ação ao sistema Parceiro e para isso é gerado e enviado um *Marker PDU* (MPDU). O sistema Parceiro, ao receber um MPDU, ativa o FC, gera um pacote do tipo *Marker Response PDU* (MRPDU) e o envia ao seu sistema Parceiro (no caso, para o sistema que enviou o MPDU). Ao receber o MRPDU, o sistema ativa seu FD e dá seguimento à execução do LACP [BGF11].

A descrição e a implementação do Protocolo Marcador faz parte do Trabalho de Conclusão de Bondan L., Gobbi R. e Fochi V., intitulado "O Protocolo de Controle de Agregação de Enlaces LACP: Um Simulador e o Protocolo de Marcação de Pacotes" que será integrado ao presente trabalho.

3 O CONTROLE DE AGREGAÇÃO DE ENLACES

Este Capítulo descreve as características do LAC (Seção 3.1), os identificadores utilizados por ele (Seção 3.2) e descreve as funções que as máquinas de estado executam (Seção 3.3).

3.1 Características do LAC

O módulo de software responsável por configurar e controlar toda a operação do LACP denomina-se Controle de Agregação de Enlaces. O LAC configura as variáveis de sistema que mantêm as informações do Ator e do Parceiro.

Para cada enlace físico que participa do LACP, o LAC mantém informações de configuração em uma estrutura lógica chamada porta, tais como propriedades inerentes a cada enlace (ex.: velocidade) e definições estabelecidas pelo administrador da rede (ex.: definir um dado enlace como individual).

O LAC é responsável por criar e remover os LAGs e Agregadores do sistema, vincular e desvincular portas a um Agregador, verificar quais os enlaces de um sistema podem ser agregados e vinculá-los ou desvinculá-los de um LAG.

Para garantir a integridade dos grupos, o LAC monitora a configuração de cada enlace e com isso assegura a validade da agregação. Para garantir essas propriedades é necessário o estabelecimento de um **identificador** único (do ponto de vista do Ator) e global (do ponto de vista de todos os sistemas) para cada sistema que implementa o LACP. É preciso também um meio de identificar o conjunto de recursos que cada porta e cada agregador possuem, pois estes determinam sua possibilidade de agregação.

Para a correta execução e configuração do LACP não são necessários eventos de inicialização ou intervenção manual, pois as mudanças de estado do sistema são monitoradas continuamente. Por isto, a necessidade de reconfiguração é detectada automaticamente, sendo executada a partir da análise das capacidades de cada porta/enlace bem como de sua conectividade física.

O sistema converge rapidamente e para isto pode ser necessário a troca de somente três LACPDU's. A sobrecarga de quadros de controle entre os sistemas é considerada baixa [IEE08].

Nas seções a seguir, são descritos os identificadores (Seção 3.2), assim como o LAC e seus objetivos (Seção 3.3).

3.2 Identificadores

As entidades lógicas gerenciadas pelo LAC (sistema, portas, agregadores, etc.) possuem identificadores para representá-los, estes são descritos a seguir.

3.2.1 Identificador do Sistema

Cada porta do sistema possui um endereço MAC e o administrador definirá um destes para ser o MAC do sistema.

O identificador do sistema deve ser único e global e é formado pela concatenação do endereço MAC escolhido pelo administrador com o valor do atributo Prioridade do Sistema, também definido por ele.

Quando for necessário fazer comparações numéricas entre os identificadores de sistemas, deve ser considerado que cada um seja um número binário sem sinal com o tamanho de oito octetos. Os dois octetos mais significativos do identificador do sistema são o valor do atributo prioridade do sistema, que é um número binário sem sinal. Os outros seis octetos recebem o endereço MAC atribuído ao sistema.

A Figura 7 exemplifica o identificador de sistema para um sistema com MAC 00:11:22:33:44:55 e prioridade do sistema 66:77:88:99.

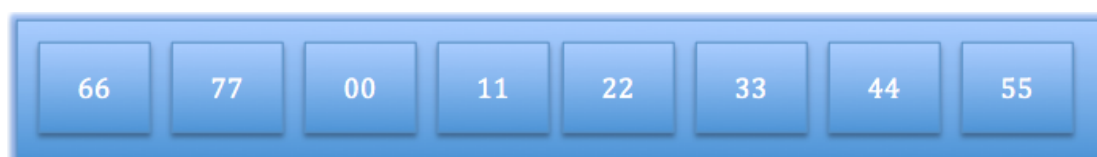


Figura 7 – Exemplo de Identificador de Sistema.

3.2.2 Identificador do Agregador

Cada Agregador deve ter um endereço MAC único que será definido a partir de um dos endereços MAC das portas que estão a ele vinculadas. Assim, nenhum Agregador pode ter um endereço MAC de uma porta que está vinculada a outro Agregador. Um Agregador deve ter também um valor inteiro que será usado pelo LAC para identificá-lo no sistema.

3.2.3 Identificador da porta

O identificador de cada porta do sistema é formado pela concatenação do atributo Prioridade da Porta com o atributo Número da Porta. Quando for necessário executar comparações numéricas entre os identificadores das portas, estes devem ser considerados números binários sem sinal formados por quatro octetos.

Os dois octetos mais significativos são derivados dos dois octetos mais significativos do atributo Prioridade da Porta, o terceiro e quarto octetos do identificador da porta são compostos pelos dois primeiros octetos do valor do atributo Número da Porta.

Supondo o valor 00:11:22:33 para o atributo Prioridade da Porta e 00:00:00:10 para o Número da Porta, o Identificador da Porta é definido como apresentado na Figura 8.

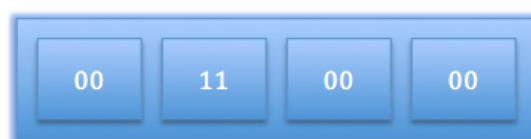


Figura 8 – Exemplo de Identificador de Porta.

3.2.4 Identificação das Características das portas

As portas físicas do sistema possuem características, e é através destas que o protocolo identifica quais podem ser agregadas.

A principal característica de uma porta, do ponto de vista de formar agregações, é denominada Chave. Se duas ou mais portas possuem a mesma chave, então elas devem ser agregadas fazendo parte do mesmo LAG. Se uma porta possui uma chave única então esta deve fazer parte de um LAG formado somente por ela.

Outras características das portas são:

- Taxa de transferência;
- Tipo de comunicação (*duplex*, *full-duplex* e *half-duplex*);
- Restrições configuradas pelo administrador da rede.

Cada porta deve ter duas chaves associadas à ela, uma Chave Operacional (OPER KEY) e uma Chave Administrativa (ADMIN KEY). A OPER KEY somente é avaliada com o propósito de formar agregações, enquanto que a ADMIN KEY permite ao administrador manipular os valores de ambas as chaves.

O valor de ambos tipos de chave somente tem significado para o sistema local, sendo que a atribuição de uma mesma OPER KEY a um conjunto de portas possibilita a formação de uma agregação do ponto de vista local, composta por qualquer sub-conjunto (incluindo o conjunto inteiro) destas portas. Entretanto, o conjunto que efetivamente formará uma agregação será composto pelas portas do Ator que estiverem conectadas às portas de um mesmo Parceiro e somente se estas tiverem o mesmo valor de OPER KEY das portas do Ator. Em um sistema, o conjunto de portas que possuem a mesma OPER KEY é chamado de Grupo de Chave.

Aos agregadores associa-se uma OPER KEY e uma ADMIN KEY que respectivamente desempenham as funções de formar agregações e possibilitar ao administrador alterar estas chaves. Portas que compõem um mesmo Grupo de Chave somente podem ser associadas a um Agregador que possui o mesmo valor de chave que este grupo. As chaves são representadas por identificadores de dezesseis bits sendo zero o único valor inválido.

3.2.5 Identificador de um Grupo de Agregação de Enlaces

Um LAG pode ser formado por um único ou por um conjunto de enlaces. Este LAG interconecta um par de sistemas, cujas portas em ambas extremidades possuem a mesma Chave de Grupo. O identificador único de um LAG (LAG ID) é construído a partir de:

- Identificador do Sistema;
- OPER KEY atribuída às portas do LAG;
- Identificador da Porta caso o enlace seja individual e zero caso contrário.

Os valores locais para estes identificadores devem ser diferentes de zero. Se o Ator não

conseguir resolver os valores destes identificadores do Parceiro através da troca de informação executada pelo protocolo, então valores administrativos devem ser considerados para formar o LAG ID. O LAG ID formado somente pelos identificadores do sistema e valores das chaves é suficiente para identificar um LAG formado por enlaces agregáveis. Entretanto, esse não é suficiente para identificar um LAG formado por um único enlace individual, onde o Identificador do Sistema e a Chave Operacional do Parceiro são zero (ex.: quando o Ator assume valores padrão para estes atributos do Parceiro).

Mesmo que esses valores não sejam zero podem existir múltiplas combinações de enlaces individuais com o mesmo identificador de sistema e chaves operacionais, portanto, se faz necessária a inclusão de identificadores de porta para gerar LAG IDs únicos.

Sejam S1 e S2 identificadores de sistema (Prioridade e Identificador), K1 e K2 chaves operacionais atribuídas a um LAG por S1 e S2, respectivamente, e P1 e P2 identificadores de portas (prioridade e número) se o enlace vinculado a esse LAG for individual, e zero se o LAG for composto por um ou mais enlaces agregáveis.

A forma geral do LAG ID é [(S1,K1,P1), (S2,K2,P2)]. Para facilitar a comparação de LAG IDs, estes devem estar ordenados pelo valor do identificador do sistema. A Tabela 1 apresenta um exemplo real de configuração de um LAG ID.

Tabela 1 - Exemplo de configuração de LAG ID.

	Sistema S1	Sistema S2
Parâmetros do Sistema	Prioridade do Sistema: 0x8000 Identificador do Sistema: AC-DE-48-03-67-80	Prioridade do Sistema: 0x8000 Identificador do Sistema: AC-DE-48-03-FF-FF
Parâmetros de Chave	Chave: 0x0001	Chave: 0x00AA
Parâmetros de Porta	Prioridade da Porta: 0x80 Número da Porta: 0x0002	Prioridade da Porta: 0x80 Número da porta: 0x0002

O LAG ID para esta configuração, caso o enlace fosse individual, seria assim:

[(8000, AC-DE-48-03-67-80, 0001, **80, 0002**), (8000, AC-DE-48-03-FF-FF, 00AA, **80, 0002**)]

Caso o enlace fosse agregável, o correspondente LAG ID seria assim:

[(8000, AC-DE-48-03-67-80, 0001, **00, 0000**), (8000, AC-DE-48-03-FF-FF, 00AA, **00, 0000**)]

A diferença entre as representações é que para um enlace agregável os componentes de identificação das portas estão zerados.

3.3 Controle de Agregação

As configurações administrativas possibilitam controlar a maneira como as agregações são construídas. Essas configurações permitem ao administrador identificar e modificar os parâmetros de configuração do LACP, as chaves associadas às portas e aos agregadores do sistema e identificar enlaces incapazes de serem agregados.

Cada enlace é monitorado para confirmar a consistência das informações de configuração em ambos os lados do enlace. Se o processo de monitoramento detectar mudanças que alterem o

LAG ID de um enlace, então será necessário sua remoção do LAG do qual participa e posterior inclusão em um novo LAG.

O LAC mantém informações a respeito das portas do sistema, assim como o identificador do LAG a que estão vinculadas, e o identificador do Agregador que está vinculado a este LAG. Estas informações de estado são compartilhadas e acessadas diretamente pelo Agregador.

Para cada Agregador, o LAC mantém informações sobre o estado de seu FD (*frame distribution*) e FC (*frame collection*). Se o FD estiver habilitado significa que pelo menos uma das portas que estão vinculadas ao seu Agregador está com status de distribuição habilitado. Também, se pelo menos uma destas portas estiver desabilitada, então o FD correspondente também estará desabilitado. O mesmo vale para o FC.

3.3.1 Funções do Controle de Agregação

O LAC é responsável por configurar as agregações de enlaces e para tal, são necessários um conjunto de operações. Estas operações fazem parte da Lógica de Seleção, descrita na Seção 5.5 e são descritas a seguir.

3.3.1.1 Seleção de um Grupo de Agregação de Enlaces

Cada porta é selecionada para fazer parte de um LAG exclusivamente através de seu LAG ID e sua definição inicial deve ser atrasada para permitir a chegada de informações das diversas portas do Parceiro. Pode-se assim, selecionar múltiplas portas de uma só vez para participar de um LAG. Caso as informações da porta não cheguem, devem ser assumidos valores padrão para seus parâmetros operacionais.

Quando um enlace for individual, não é necessário esperar pelas informações da porta no ponto remoto do enlace para formar o LAG ID, pois como é individual, será o único membro de seu LAG.

3.3.1.2 Concordância de participação em um Grupo de Agregação de Enlaces

Antes que os quadros sejam distribuídos e recebidos de um enlace, o LAC do Ator e do Parceiro precisam concordar com o LAG para este enlace. O LACP permite que cada entidade verifique o LAG ID do Parceiro para saber se ambos possuem as mesmas configurações. Caso contrário, é feita a troca de informações até que ambos possuam as mesmas configurações.

Quando um enlace for individual, o Ator avisará o Parceiro. Isto acelera o uso do enlace, pois quando o mesmo tiver esta característica não será preciso atrasar a definição do LAG ID, pois não existem mais enlaces que participarão deste LAG.

3.3.1.3 Vinculação de um enlace a um Agregador

Uma vez que um enlace tenha selecionado um LAG, o LAC pode vinculá-lo a um Agregador somente se sua OPER KEY for a mesma da porta.

Enlaces que não podem fazer parte de agregações (ex.: enlaces que estiverem vinculados a

dispositivos que não permitem agregações ou configurados manualmente para serem não-agregáveis) estão habilitados para operar como enlaces IEEE 802.3. Na perspectiva do LACP, enlaces não-agregáveis não são um caso especial e sim uma agregação com no máximo um enlace.

3.3.1.4 Desvinculação de um enlace de um Agregador

Uma porta será desvinculada de um Agregador usado por um LAG caso haja mudanças em seu estado, como alterações na chave operacional ou devido a alguma restrição do sistema (i.e. falha de dispositivo). Estes eventos ocasionarão mudanças no LAG ID do enlace e por isso será necessário que o LAC desvincule este enlace de seu Agregador e o mova para um novo LAG (se possível). Quando estas mudanças são detectadas, as funções de coleção e distribuição de quadros são desabilitadas. A função de distribuição de quadros é informada que o enlace não faz mais parte do grupo a que ela pertence e as informações que foram atualizadas são comunicadas ao Parceiro para então comunicar a função de coleção que o enlace não faz mais parte de seu grupo. O enlace, após ser removido de seu Agregador, pode selecionar um novo LAG e então ser vinculado a um Agregador compatível.

Qualquer conversação (troca de quadros entre clientes MAC) que seja realocada a um enlace diferente como resultado de sua desvinculação de um Agregador terá a ordenação de seus quadros preservada. Isto implica o uso do Protocolo de Marcação (Anexo 1) para assegurar que o sistema remoto não receba quadros desta conversação no grupo ao qual o enlace fazia parte depois de este grupo ser movido para um novo LAG.

3.3.1.5 Sinalização de habilitação para transferência de dados

Uma vez que um enlace esteja vinculado a um Agregador compatível, através da concordância do LAG, as entidades do LAC sinalizam prontidão às suas portas para transferir e receber dados do Agregador do cliente MAC.

Quando ocorrer alguma alteração no LAG ID da porta (e.g. quando o LAG ID da porta mudar) e esta alteração ocasionar a remoção desta porta de um LAG e inclusão em outro, o LAC somente sinalizará prontidão para transmitir dados do novo LAG nesta porta, somente quando os dados do LAG anterior acabarem de serem transmitidos. O LAC não sinaliza prontidão até que seja certo que quadros que fazem parte de um LAG anterior ainda estejam sendo transmitidos no enlace. Esta certificação envolve o uso do Protocolo de Marcação que assegura que não existam mais quadros a serem recebidos antes da reconfiguração.

3.3.1.6 Habilitação da Coleta e Distribuição de quadros

Todos os agregadores podem habilitar e desabilitar a coleta e distribuição de quadros de cada porta que está vinculada a eles. Inicialmente, a coleta e distribuição estão desabilitadas. Uma vez que o Ator e seu Parceiro sinalizaram prontidão para transferência de dados em um enlace, o processo de habilitar o enlace para coletar e distribuir quadros pode ser efetuado. É habilitado então o coletor, que prepara o enlace para receber os quadros enviados pelo distribuidor do Agregador do Parceiro e esta ação lhe é comunicada através do envio de um LACPDU. Quando o parceiro receber

esse comunicado, saberá que o coletor do Agregador remoto está habilitado e com isso habilitará o seu distribuidor.

Se pelo menos uma porta que faz parte de um LAG estiver com a função de coleta habilitada, então o estado de coleta do Agregador correspondente deverá ser habilitado. Também, se pelo menos uma porta do LAG estiver com a função de distribuição habilitada, então o estado de transmissão do Agregador correspondente deverá ser habilitado.

4 INFRAESTRUTURA PARA A IMPLEMENTAÇÃO DO PROTOCOLO

Neste Capítulo é apresentada uma visão geral do funcionamento do LACP (Seção 4.1). No final do Capítulo é explicado como funciona a transmissão dos pacotes de controle (Seção 4.3).

4.1 Visão macro da operação do LACP

O LACP funciona através da execução de máquinas de estados concorrentes cuja implementação é detalhada no Capítulo 5. A funcionalidade básica de cada máquina de estado é descrita a seguir.

- Máquina Receptora (*RX*): controla o recebimento de LACPDUs enviados pelo sistema Parceiro e verifica o estado que este sistema possui do sistema Ator;
- Máquina Periódica (*Periodic*): controla o envio periódico de LACPDUs para o sistema Parceiro a fim de manter as agregações;
- Lógica de Seleção (*Selection Logic*): conjunto de funções que relacionam portas a agregadores no sistema;
- Máquina Multiplexadora (*MUX*): controla a vinculação e desvinculação de uma porta a um agregador e também habilita e desabilita o envio e recebimento de quadros em uma dada porta;
- Máquina de Transmissão (*TX*): controla o envio de LACPDUs para o sistema Parceiro de maneira periódica ou sob demanda de alguma outra máquina de estado.

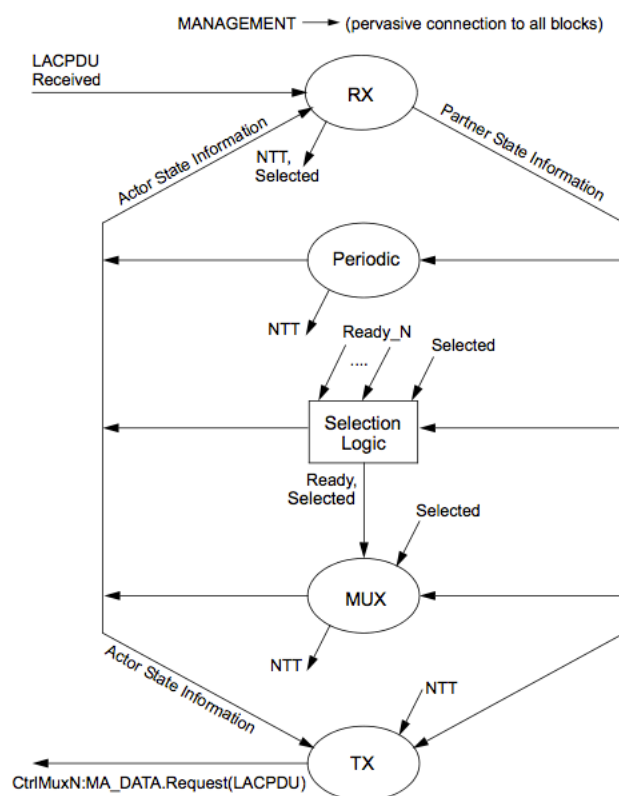


Figura 9 - Relação entre as máquinas de estado que compõe o LACP [IEE08].

A Figura 9 apresenta a relação destas máquinas de estado. Os LACPDUs são recebidos pela Máquina de Recepção (Seção 5.2) que disponibiliza as informações para as outras máquinas executarem. De maneira periódica, a Máquina de Transmissão Periódica "avisa" a Máquina de Transmissão para enviar um LACPDU, que contém informações atualizadas sobre o sistema Ator, para o sistema Parceiro.

Para cada porta no sistema, existe uma instância da Máquina de Recepção, Máquina Transmissora Periódica e Máquina Multiplexadora. A Máquina Transmissora e a Lógica de Seleção são funções chamadas sob demanda das máquinas de estados citadas anteriormente.

4.2 Variáveis de controle de estado e sua organização

Para cada entidade lógica do LACP (ex.: Agregador, Porta, LAG, etc.) foi definida uma estrutura de dados para o armazenamento e consulta de suas informações. Foi definida também uma estrutura genérica que armazena grupos dessas estruturas.

Entidades como Agregadores e Portas precisam de uma estrutura para organizá-los. A primeira idéia foi usar listas para gerenciá-los, mas ao analisar o comportamento do LACP foi percebido que esta estrutura não era ideal, pois como a lista seria acessada por vários processos leves (*threads*) seria necessário definir toda a lista como área crítica ou seja, se algum outro processo leve quisesse acessar apenas um nó desta lista, a lista inteira ficaria bloqueada. Entretanto, utilizando vetores podemos definir somente a área de memória referente à posição que queremos acessar como área crítica.

Esta Seção apresenta as variáveis de controle utilizadas para a implementação do LACP. Estas variáveis são agrupadas de acordo com as entidades que as utilizam.

4.2.1 Agregador

O Agregador é responsável por controlar o fluxo de quadros vindo das portas a ele relacionadas. Os quadros vindos dos Clientes MAC são recebidos e distribuídos entre os enlaces que estão relacionados a ele, enquanto que, os quadros vindos dos enlaces devem ser entregues aos Clientes MAC corretos. Abaixo são descritos os campos que a estrutura de dados utiliza para representar um Agregador.

- *Aggregator_MAC_address*: endereço MAC que corresponde ao endereço MAC da porta de menor índice que está vinculada a ele;
- *Aggregator_Identifier*: identificador do agregador no sistema;
- *Individual_Aggregator*: indica se o enlace ao qual este agregador está vinculado é individual ou não;
- *Actor_Admin_Aggregator_Key*: valor padrão da Actor_Oper_Aggregator_Key, definido pelo administrador do sistema;
- *Actor_Oper_Aggregator_Key*: valor da operacional da chave, definida pelo sistema Ator;

- *Partner_System*: endereço MAC do Sistema ao qual Agregador está conectado. Se o número de portas vinculadas a este Agregador for zero, então esta variável deve ser zerada;
- *Partner_System_Priority*: prioridade do Sistema a qual este Agregador está conectado. Se o número de portas vinculadas a este Agregador for zero, então esta variável deve ser zerada;
- *Partner_Oper_Aggregator_Key*: valor operacional da chave atribuída a uma agregação pelo Sistema Parceiro ao qual este Agregador está conectado. Se o número de portas vinculadas a este Agregador for zero, então esta variável deve ser zerada;
- *Receive_State*: se pelo menos uma porta relacionada a este Agregador estiver com a variável Actor_Oper_Port_State.Collecting em TRUE, indica que este Agregador está recebendo quadros. Não recebe quadros caso contrário;
- *Transmit_State*: se pelo menos uma porta relacionada a este Agregador estiver com a variável Actor_Oper_Port_State.Distributing em TRUE, indica que este Agregador está transmitindo quadros. Não transmite quadros caso contrário;
- *LAG_Ports*: vetor que possui os índices das portas que este Agregador esta gerenciando.

Do ponto de vista da implementação a estrutura Agregador é codifica como mostra a Figura

10.

```

struct Aggregator {
    u_int8_t    aggregator_MAC_address[ETH_ALEN];
    u_int32_t   aggregator_identifier;
    bool_t      individual_aggregator;

    u_int32_t   actor_admin_aggregator_key;
    u_int32_t   actor_oper_aggregator_key;

    u_int8_t    partner_system[ETH_ALEN];
    u_int32_t   partner_system_priority;
    u_int32_t   partner_oper_aggregator_key;

    bool_t      receive_state;
    bool_t      transmit_state;

    int8_t      lag_ports[NPORTS];
};

```

Figura 10 - Codificação da estrutura lógica Agregador.

4.2.2 Porta

A **Porta** é a estrutura de dados utilizada para armazenar o estado de cada porta do Sistema Ator e Parceiro. Esta estrutura pode ser dividida em duas grandes partes sendo a primeira, a parte que armazena informações da porta do Ator e a segunda armazena informações do Parceiro. Abaixo

são descritos os campos que a estrutura de dados utiliza para representar uma porta.

Parte da estrutura que armazena informações da porta do **Sistema Ator**:

- *Actor_Port_Number*: número da porta atribuído à porta pelo administrador;
- *Actor_Port_Priority*: valor da prioridade atribuída a esta porta. Utilizada para as mudanças dinâmicas de chaves;
- *Actor_Port_Aggregator_Identifier*: identificador do Agregador ao qual esta porta está vinculada;
- *NTT (need to transmit)*: variável que indica que as informações de estado desta porta foram atualizadas e portanto, é preciso avisar o sistema parceiro enviando um LACPDU com informações novas sobre esta porta;
- *Actor_Admin_Port_Key*: valor padrão para a variável *Actor_Oper_Port_Key* desta porta. Atribuída pelo administrador;
- *Actor_Oper_Port_Key*: valor de chave operacional desta porta. Atribuída pelo Ator;
- *Actor_Admin_Port_State*: valor padrão para as variáveis de estado da porta. O estado da porta é descrito através de um subconjunto de variáveis, todas binárias. A Figura 11 apresenta este subconjunto de variáveis que é explicado a seguir.

BIT							
0	1	2	3	4	5	6	7
LACP_Activity	LACP_Timeout	Aggregation	Synchronization	Collecting	Distributing	Defaulted	Expired

Figura 11 - Representação do subconjunto de variáveis *Actor_Admin_Port_State*.

1. *LACP_Activity*: indica a preferência do sistema em participar ativamente ou passivamente no LACP. Assume dois valores, *LACP_Active* e *LACP_Passive*;
 2. *LACP_Timeout*: indica o tempo de validade das informações da porta (Seção 5.2.1);
 3. *Aggregation*: indica a capacidade do enlace de ser agregado;
 4. *Synchronization*: indica se os sistemas estão em sincronia ou não, ou seja, ambos sistemas possuem informações válidas ou inválidas sobre seu Parceiro;
 5. *Collecting*: indica se a coleta de quadros nesta porta está habilitada ou desabilitada;
 6. *Distributing*: indica se a coleta de quadros nesta porta está habilitada ou desabilitada;
 7. *Defaulted*: indica se esta porta está utilizando as informações padrão de configuração de porta ou informações recebidas em um LACPDU;
 8. *Expired*: indica se a RX está no estado EXPIRED (Seção 5.2.4.5) ou não.
- *Actor_Oper_Port_State*: variável de estado da porta. Definida pelas máquinas de estado

Parte da estrutura que armazena informações da porta do **Sistema Parceiro**:

- *Partner_Admin_System*: valor padrão para o componente endereço MAC do identificador de sistema do Parceiro. Utilizado quando esta informação é desconhecida ou sua validade tenha expirado;
- *Partner_Oper_System*: valor operacional para o componente endereço MAC do identificador de sistema do Parceiro. O Ator preenche esta variável de acordo com um LACPDU recebido ou com o valor de *Partner_Admin_System*;
- *Partner_Admin_System_Priority*: valor padrão para o componente prioridade do sistema do identificador de sistema do Parceiro. Utilizado quando esta informação é desconhecida ou sua validade tenha expirado;
- *Partner_Oper_System_Priority*: valor operacional para o componente prioridade do sistema do identificador de sistema do Parceiro. O Ator preenche esta variável de acordo com um LACPDU recebido ou com o valor de *Partner_Admin_System_Priority*;
- *Partner_Admin_Key*: valor padrão para a chave desta porta. Utilizado quando esta informação é desconhecida ou sua validade tenha expirado;
- *Partner_Oper_Key*: valor operacional para a chave desta porta. O Ator preenche esta variável de acordo com um LACPDU recebido ou com o valor de *Partner_Oper_Key*;
- *Partner_Admin_Port_Number*: valor padrão para o componente número da porta do identificador de porta do sistema Parceiro. Utilizado quando esta informação é desconhecida ou sua validade tenha expirado;
- *Partner_Oper_Port_Number*: valor operacional para o componente número da porta do identificador de porta do sistema Parceiro. O Ator preenche esta variável de acordo com um LACPDU recebido ou com o valor de *Partner_Admin_Port_Number*;
- *Partner_Admin_Port_Priority*: valor padrão para o componente prioridade da porta do identificador de porta do sistema Parceiro. Utilizado quando esta informação é desconhecida ou sua validade tenha expirado;
- *Partner_Oper_Port_Priority*: valor operacional para o componente prioridade da porta do identificador de porta do sistema Parceiro. O Ator preenche esta variável de acordo com um LACPDU recebido ou com o valor de *Partner_Admin_Port_Priority*;
- *Partner_Admin_Port_State*: mesmo que *Actor_Admin_Port_State*;
- *Partner_Oper_Port_State*: mesmo que *Actor_Oper_Port_State*;
- *Port_Enabled*: variável que indica que a porta está operável e sua conexão com o sistema Parceiro foi estabelecida. Esta informação vem da camada física e é dependente de implementação.

Do ponto de vista da implementação, a Figura 12 apresenta a codificação da entidade lógica Porta.


```

struct Port {
    u_int8_t    actor_port_mac_address[ETH_ALEN];
    u_int32_t    actor_port_identifier;
    u_int32_t    actor_port_number;
    u_int32_t    actor_port_priority;
    u_int32_t    actor_port_aggregator_identifier;
    bool_t      NTT;
    u_int32_t    actor_admin_port_key;
    u_int32_t    actor_oper_port_key;
    u_int8_t     actor_admin_port_state;
    u_int8_t     actor_oper_port_state;
    u_int8_t     partner_admin_system[ETH_ALEN];
    u_int8_t     partner_oper_system[ETH_ALEN];
    u_int32_t    partner_admin_system_priority;
    u_int32_t    partner_oper_system_priority;
    u_int32_t    partner_port_identifier;
    u_int32_t    partner_admin_key;
    u_int32_t    partner_oper_key;
    u_int32_t    partner_admin_port_number;
    u_int32_t    partner_oper_port_number;
    u_int32_t    partner_admin_port_priority;
    u_int32_t    partner_oper_port_priority;
    u_int8_t     partner_admin_port_state;
    u_int8_t     partner_oper_port_state;
    bool_t      port_enabled;
};

```

Figura 12 - Codificação da estrutura lógica Porta.

4.2.3 Unidade de dados do Protocolo de Controle de Agregação de Enlaces

O LACPDU é a estrutura de dados que comunica os sistemas que executam o LACP. Esta estrutura carrega todas as informações necessárias para a configuração dos sistemas.

De maneira abstrata, o LACPDU pode ser subdividido em duas sessões principais que se diferenciam por conter informações de sistemas diferentes. A primeira sessão contém informações de estado do sistema local e da porta pela qual o LACPDU será enviado. A segunda sessão do quadro é similar à primeira, entretanto, contém os dados do sistema remoto.

Os principais campos dessa estrutura são apresentados na Tabela 2.

Tabela 2 - Principais informações do LACPDU.

Informações do Sistema Ator	Prioridade do Sistema
	Sistema
Informações da Porta do Ator	Chave Operacional
	Prioridade da Porta
	Identificador da Porta
Estado da Porta do Ator	Atividade do LACP
	Timeout do LACP
	Agregação
	Sincronização
	Recepção
	Distribuição
	Padronizado
	Expirado

Informações do Sistema Parceiro	Prioridade do Sistema
	Sistema
Informações da Porta do Parceiro	Chave Operacional
	Prioridade da Porta
	Identificador da Porta
Estado da Porta do Parceiro	Atividade do LACP
	Timeout do LACP
	Agregação
	Sincronização
	Recepção
	Distribuição
	Padronizado
	Expirado

A Figura 13 apresenta a codificação do LACPDU.

```
typedef struct lacpdu_tag {
    struct ether_header eth_h;
    u_int8_t subtype;
    u_int8_t version;
    u_int8_t TLV_type_a;
    u_int8_t actor_info_len;
    u_int16_t actor_sys_prio;
    u_int8_t actor_system[ETH_ALEN];
    u_int16_t actor_key;
    u_int16_t actor_port_prio;
    u_int16_t actor_port;
    u_int8_t actor_state;
    u_int8_t reserved_a[3];
    u_int8_t TLV_type_p;
    u_int8_t partner_info_len;
    u_int16_t partner_sys_prio;
    u_int8_t partner_system[ETH_ALEN];
    u_int16_t partner_key;
    u_int16_t partner_port_prio;
    u_int16_t partner_port;
    u_int8_t partner_state;
    u_int8_t reserved_p[3];
    u_int8_t TLV_type_c;
    u_int8_t collector_info_len;
    u_int16_t collector_max_delay;
    u_int8_t reserved_c[12];
    u_int8_t TLV_type_t;
    u_int8_t terminator_len;
    u_int8_t reserved[50];
}lacpdu_t;
```

Figura 13 - Codificação da estrutura de dados LACPDU.

4.2.4 Identificação do Sistema

Cada Sistema do ponto de vista global, deve ser identificado de maneira única. O LACP define duas variáveis para que essa identificação seja possível.

- *Actor_system_identifier*: combinação das variáveis *Actor_System* e *Actor_system_Priority*. Descrita na Seção 3.2.1.
- *Actor_System*: endereço MAC do sistema. Endereço MAC de alguma porta deste sistema que é definida pelo administrador.
- *Actor_System_Priority*: prioridade do Sistema. Definida pelo administrador do sistema.

Do ponto de vista da implementação, a Figura 14 apresenta a codificação da estrutura de dados que armazena as informações do sistema.

```
struct System {
    u_int64_t  actor_system_identifier;
    u_int8_t   actor_system[ETH_ALEN];
    u_int32_t  actor_system_priority;
};
```

Figura 14 - Codificação da estrutura de dados que armazena as informações do sistema.

4.2.5 Variáveis de controle das máquinas de estado

Todas as máquinas de estado são inicializadas ou reinicializadas de acordo com o valor da variável *BEGIN* que é definida como TRUE quando o sistema é inicializado. Quando o LACP é reinicializado esta variável é definida como FALSE. A variável *BEGIN* é única no sistema e é compartilhada por todas as máquinas de estado.

De acordo com o funcionamento do LACP as máquinas de estado controlam variáveis que estão relacionadas às portas. Cada porta possui um conjunto de variáveis que são definidas pelas máquinas de estados. Esse conjunto de variáveis é descrito a seguir.

- *LACP_Enabled*: quando o valor desta variável é TRUE, indica que o enlace físico está operando em modo *full-duplex* e está sendo gerenciado pelo LACP. Caso o enlace passe a operar em modo *half-duplex*, o valor da variável deve ser mudada para FALSE;
- *actor_churn*: esta variável assume o valor TRUE quando a máquina *Churn*, que controla o estado do Ator, detecta que uma porta falhou ao ser configurada dentro de um tempo específico e que será necessária intervenção do administrador.
- *partner_churn*: esta variável assume o valor TRUE quando a máquina *Churn* que controla o estado do sistema Parceiro detectar que uma porta remota falhou ao ser configurada dentro de um tempo específico e será necessário intervenção do administrador;
- *Ready_N*: cada porta no sistema possui uma variável *Ready_N* que assume o valor TRUE para indicar à Lógica de Seleção que o *wait_while_timer* expirou e portanto a porta está no estado WAITING. Isto significa que a estrutura lógica da porta em questão está esperando

para ser anexada a um Agregador. Seu estado é FALSE caso a porta já esteja anexada a um Agregador.

- *Selected*: cada porta lógica no sistema local possui uma variável Selected que pode assumir três valores. Quando foi definido um Agregador compatível para esta porta, então seu valor muda para SELECTED. Enquanto esta porta não estiver vinculada a nenhum Agregador, então seu valor será UNSELECTED. O valor STANDBY significa que embora a Lógica de Seleção tenha selecionado um Agregador apropriado, existem restrições na agregação que impossibilitam a porta de fazer parte do LAG (e.g.: a chave da porta correspondente no sistema Parceiro não é a mesma que no sistema Ator). A Lógica de Seleção altera o estado dessa variável somente para SELECTED e STANDBY enquanto que a Máquina de Recepção pode alterá-la para UNSELECTED.
- *port_moved*: o valor desta variável é mudado para TRUE quando a máquina de recepção responsável por esta porta está no estado PORT_DISABLED, e a combinação das variáveis *Partner_Oper_System* e *Partner_Oper_Port_Number* que estão configuradas nesta porta forem recebidas em um LACPDU em uma porta diferente. Seu valor é FALSE uma vez que a máquina de recepção desta porta está no estado INITIALIZE.

A variável *Ready* está relacionada a um Agregador e seu estado depende da combinação da variável *Ready_N* de cada porta controlada por este Agregador. Quando a *Ready_N* de todas as portas estiverem com estado TRUE, então o valor de *Ready* é definido como TRUE. A Lógica de Seleção é responsável pela definição do estado dessa variável.

Do ponto de vista da implementação, a Figura 15 apresenta a estrutura de dados utilizada para o armazenamento destas informações.

```
struct StateMachines {
    bool_t    lacp_enabled;
    bool_t    actor_churn;
    bool_t    partner_churn;
    bool_t    ready_N;
    selected_t selected;
    bool_t    port_moved;
};
```

Figura 15 - Codificação da estrutura que armazena variáveis de controle das portas que são manipuladas pelas máquinas de estados.

4.2.6 Identificador de Grupo de Agregação de Enlaces

Este grupo de variáveis descreve o Identificador de Grupo de Agregação de Enlaces (*Link aggregation Group Identifier* - LAG ID) e armazena as variáveis para a composição do LAG ID (Seção 3.2.5) que identifica de forma única e global os LAGs do sistema. A Lógica de Seleção (Seção 5.5) utiliza as informações desse conjunto de variáveis para criar as agregações de enlaces.

A estrutura de dados que armazena esses valores é apresentada na Figura 15.

```

struct lag_id {
    u_int64_t    actor_system_identifier;
    u_int8_t     actor_system_mac_address[ETH_ALEN];
    u_int32_t    actor_system_priority;
    u_int16_t    actor_oper_port_key;
    u_int32_t    actor_port_priority;
    u_int32_t    actor_port_number;
    u_int64_t    partner_system_identifier;
    u_int8_t     partner_system_mac_address[ETH_ALEN];
    u_int32_t    partner_system_priority;
    u_int16_t    partner_oper_port_key;
    u_int32_t    partner_port_priority;
    u_int32_t    partner_port_number;
};

```

Figura 15 - Codificação da estrutura que armazena as informações do LAG ID de um enlace.

4.2.7 Informações do Enlace

O LAC mantém um conjunto de variáveis para cada enlace físico no sistema. Este grupo de variáveis é descrito a seguir.

- *lag*: referência para a estrutura LAG ao qual este enlace faz parte;
- *aggregator_identifier*: identificador do agregador que controla o fluxo de quadros neste enlace;
- *distributing*: variável que indica se este enlace está com a distribuição de quadros habilitada;
- *collecting*: variável que indica se este enlace está com a recepção de quadros habilitada.

A Figura 16 apresenta a codificação da estrutura utilizada para armazenar as informações descritas acima.

```

struct link_info {
    struct lag_id *lag_id;
    u_int32_t     aggregator_identifier;
    u_int8_t      distributing;
    u_int8_t      collecting;
};

```

Figura 16 - Codificação da estrutura que armazena informações de cada enlace.

4.3 Transmissão de pacotes de controle

O LACP mantém informações do estado de cada porta que participa do protocolo. Se a porta participa do protocolo LACP, ela está habilitada a transmitir LACPDU's.

A porta que não participa do protocolo somente envia respostas aos LACPDU's encaminhados a ela. Se as portas de cada lado de um enlace estiverem participando do LACP, ou seja, habilitadas, ocorrerão transmissões periódicas de LACPDU's. As informações contidas nestas

transmissões terão validade curta (*Short Timeout*) ou longa (*Long Timeout*), de acordo com a opção do Parceiro. Além destas transmissões periódicas, o protocolo também transmite LACPDU's quando houver necessidade de transmitir informação ao parceiro (NTT). Por exemplo, quando as informações do sistema tenham sofrido mudança.

O LACP não implementa explicitamente a detecção de perda ou retransmissão de quadros. Entretanto, se o Ator receber informações desatualizadas sobre si vindas do Parceiro ou se a próxima transmissão periódica estiver vencida, o Ator irá enviar um LACPDU que atualizará corretamente as suas informações para o Parceiro.

5 DESCRIÇÃO DAS MÁQUINAS DE ESTADO DE CONTROLE

O LACP é controlado por um conjunto de máquinas de estados (Seção 4.1). Estas máquinas controlam as variáveis de estado das portas e a validade de suas informações.

Algumas destas máquinas possuem somente uma instância para o algoritmo inteiro enquanto que, outras máquinas possuem uma instância para cada porta do sistema. Nessas máquinas, as transições de estado são causadas por eventos e geram ações a serem executadas.

5.1 Modelo genérico de implementação das máquinas de estados

Foi desenvolvido um modelo genérico para a implementação das máquinas de estados, que operam segundo a norma 802.1 AX. Este modelo foi adotado na implementação da Máquina Receptora, Máquina de Transmissão Periódica e Máquina Multiplexadora. Este modelo é apresentado na Figura 17.

```

1.  state <- STATE_1;
2.  WHILE(true)
3.      SWITCH(state)
4.      case STATE_1:
5.          // executa comandos
6.          IF(condicao_de_saida_2)
7.              state <- STATE_2;
8.              break;
9.          END IF
10.         IF(condicao_de_saida_N)
11.             state <- STATE_N;
12.             break;
13.         END IF
14.         break;
15.     case STATE_2:
16.         // executa comandos
17.         IF(condicao_de_saida_1)
18.             state <- STATE_2;
19.             break;
20.         END IF
21.         break;
22.     case STATE_N:
23.         // executa comandos
24.         IF(condicao_de_saida_1)
25.             state <- STATE_1;
26.             break;
27.         END IF
28.         IF(condicao_de_saida_2)
29.             state <- STATE_2;
30.             break;
31.         END IF
32.         break;
33.     END SWITCH
34. END WHILE

```

Figura 17 - Modelo genérico de implementação das máquinas de estados.

Os diagramas são compostos por estados, ações a serem executadas nestes estados e condições de entrada e saída. O modelo descreve os estados como *cases* de um *switch* que está dentro de um laço eterno (linha 2) e as ações a serem executadas nos estados estão dentro de cada

case. Após a execução das ações são testadas as condições de saída que, quando satisfeitas, alteram o valor da variável *state* fazendo com que na sua próxima avaliação entrará no *case* de um novo estado. A Figura 18 apresenta a descrição da máquina de estados genérica descrita anteriormente.

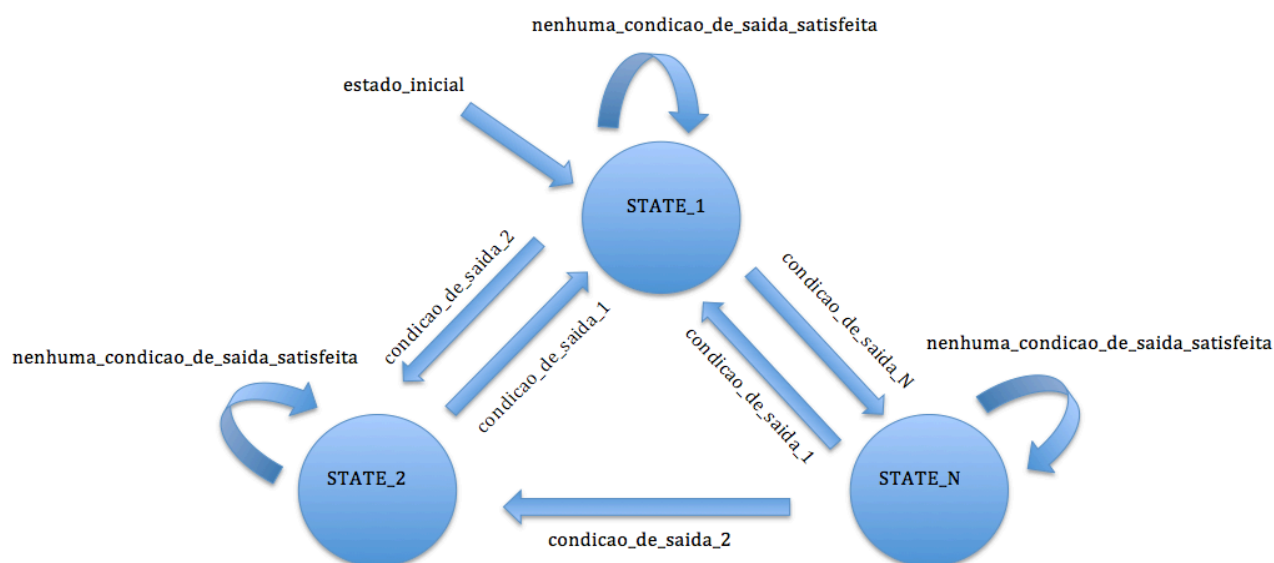


Figura 18 - Descrição em forma de autômato do modelo genérico de implementação das máquinas de estados.

5.2 Máquina Receptora

A Máquina Receptora (RX) é responsável por gravar as informações de estado da porta recebidas do sistema Parceiro e controlar o tempo com que essas informações são válidas. Para tal, a RX é descrita com seis estados.

A Figura 19 apresenta o diagrama de estados da RX, e a descrição de cada estado é detalhada na Seção 5.2.4.

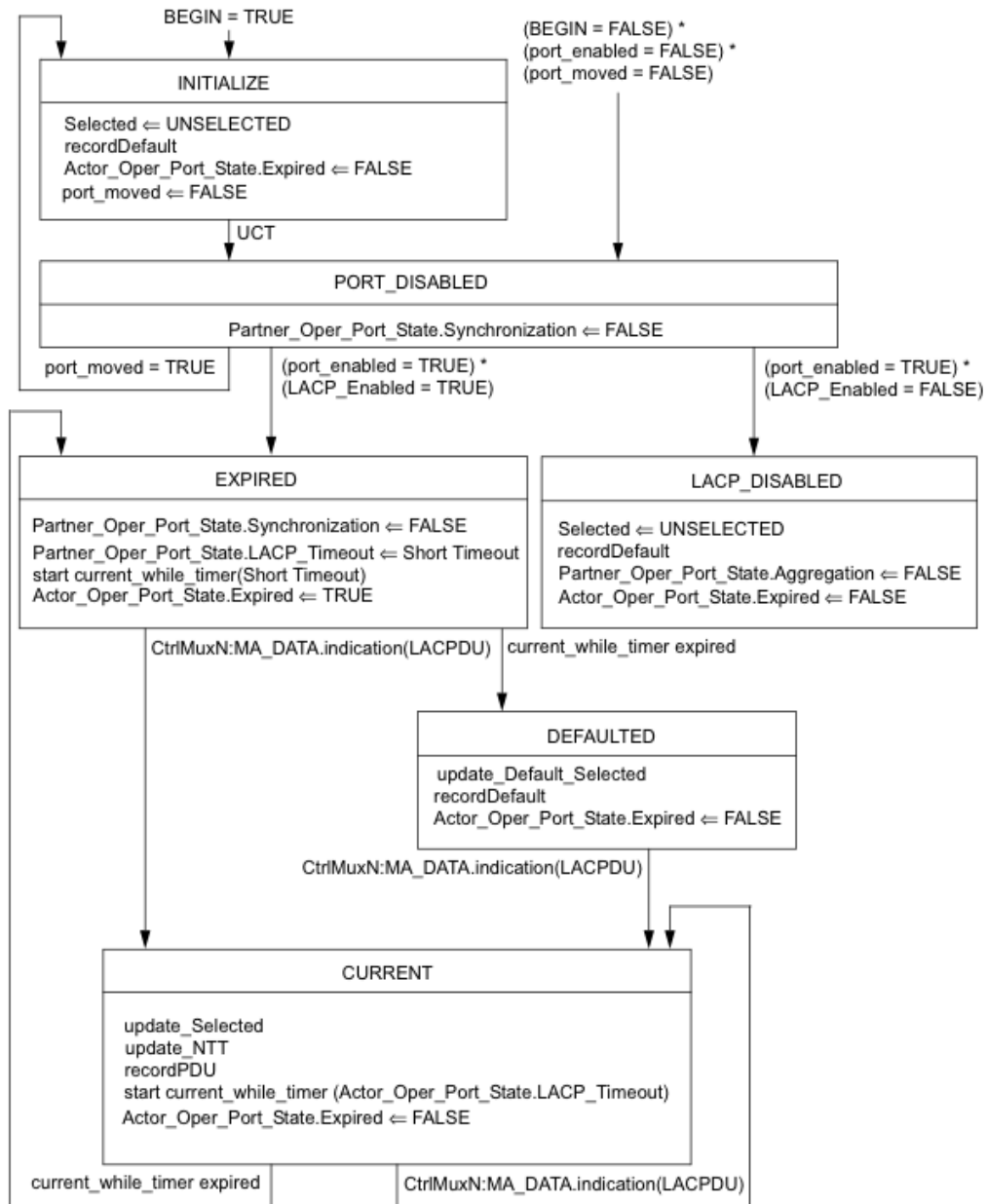


Figura 19 - Diagrama de estados da Máquina Receptora.

5.2.1 Constantes da RX

A RX adota duas constantes que são utilizadas para controle do tempo de validade das informações da porta. Essas constantes são descritas a seguir.

- *Short_Timeout_Time*: armazena o número de segundos antes de invalidar a informação de uma porta quando a mesma estiver utilizando um *Timeout* curto;
- *Long_Timeout_Time*: armazena o número de segundos antes de invalidar a informação de uma porta quando a mesma estiver utilizando um *Timeout* longo;

5.2.2 Temporizador da RX

A RX utiliza um temporizador chamado *curent_while_timer* que controla o tempo de validade das informações que são recebidas do sistema Parceiro. Se o valor do bit

Actor_Oper_state.LACP_Timeout tem o valor TRUE, então o temporizador será inicializado com o valor da constante *Short_Timeout_Time*. Caso contrário, será inicializada com o valor da constante *Long_Timeout_Time*.

5.2.3 Funções da RX

Esta máquina de estados utiliza funções para gerenciar as informações de estado da porta que está controlando. A descrição de cada função é apresentada a seguir.

- *recordDefault*: esta função define os valores padrão para o sistema Parceiro de acordo com os valores administrativos e define o bit *Actor_Oper_Port_State.Defaulted* como TRUE;
- *update_Default_selected*: esta função atualiza o valor da variável *Selected* de uma porta baseado na comparação das informações administrativas e operacionais do sistema Parceiro. Se o valor de algum parâmetro operacional for diferente do valor administrativo desta mesma variável, o valor da variável *Selected* desta porta é definido como UNSELECTED. Caso contrário, a variável *Selected* não é alterada;
- *update_Selected*: Esta função atualiza o valor da variável *Selected* da porta que recebe um novo LACPDU. A função compara os valores dos parâmetros do Ator deste novo LACPDU aos valores correspondentes dos parâmetros operacionais de Parceiro para esta mesma porta. Caso algum desses valores esteja diferente, a variável *Selected* desta porta será mudada para UNSELECTED. Caso nenhuma alteração ocorra, a variável *Selected* desta porta permanece com o mesmo valor;
- *update_NTT*: esta função atualiza o valor da variável NTT utilizando as informações contidas em um novo LACPDU. Essa função compara os valores operacionais da porta do ponto de vista do sistema Ator com os valores que o sistema Parceiro possui sobre essa mesma porta. Se as informações divergirem, quer dizer que alguma alteração de estado aconteceu e esta alteração precisa ser comunicada ao sistema Parceiro, então, o valor da variável NTT é definida como TRUE. Se os valores dos parâmetros operacionais do Ator tiverem os mesmos valores dos parâmetros operacionais do Parceiro, significa que não houve mudança no estado desta porta e, portanto, não é preciso gerar um NTT e seu valor é definido como FALSE;
- *recordPDU*: esta função grava o conteúdo dos campos referentes ao Ator de um LACPDU recebido e define o bit *Actor_Oper_Port_State.Defaulted* para FALSE pois o estado desta porta não contém mais os valores padrão. Esta função também atualiza o valor do bit *Partner_Oper_Port_State.Synchronization* através do valor correspondente neste novo LACPDU. Os valores dos campos referentes ao Parceiro são comparados aos valores operacionais do Ator e o bit *Partner_Oper_Port_State.Synchronization* é definido como TRUE se todos esses campos forem iguais e o LACP manterá o enlace ativo. No caso do enlace que enviou este LACPDU ser individual, o bit *Actor_State.Aggregation* está definido como FALSE e então o bit do Ator *Partner_Oper_Port_State.Synchronization* é

definido como TRUE e o LACP manterá o enlace ativo. Caso contrário, o bit `Partner_Oper_Port_State.Synchronization` é definido com FALSE. O LACP manterá o enlace ativo se o campo `Actor_State.LACP_Activity` do LACPDU recebido estiver definido como TRUE ou se o bit `Actor_Oper_Port_State.LACP_Activity` do Ator e o campo `Partner_State.LACP_Activity` do LACPDU recebido estiver definido como TRUE.

5.2.4 Estados da RX

A RX é composta por seis estados. Os estados INITIALIZE e PORT_DISABLED são estados de inicialização das variáveis de controle. Após a inicialização, a máquina entra no estado EXPIRED que invalida as informações de estado e fica no aguardo de um LACPDU. Uma vez no estado EXPIRED, é disparado um temporizador. Na chegada de um LACPDU antes desse temporizador expirar, a porta muda seu estado para CURRENT, caso contrário, a porta passa pelo estado intermediário DEFAULTED que define informações administrativas para a porta e fica no aguardo de um novo LACPDU. No estado CURRENT, a máquina dispara um temporizador para controlar a validade das informações atuais e fica esperando um novo LACPDU. Na chegada de um novo LACPDU, se este chegar antes do temporizador expirar, a máquina permanece no mesmo estado, caso contrário, retorna ao estado EXPIRED. O estado LACP_DISABLED é um estado final que significa que o LACP não está operando nesta porta. A descrição de cada estado é apresentada a seguir.

5.2.4.1 INITIALIZE

A porta entra nesse estado quando o sistema é inicializado ou reinicializado e também quando a porta estiver no estado PORT_DISABLED (descrito a seguir) e a condição de transição *port_moved* = TRUE for satisfeita.

Uma vez nesse estado, a máquina executa as seguintes ações:

- *Selected* <= UNSELECTED;
- *recordDefault*: descrita na Seção 5.2.3;
- *Actor_Oper_Port_State.Expired* <= FALSE
- *port_moved* <= FALSE.

Ao realizar estas quatro operações, a porta muda incondicionalmente seu estado para PORT_DISABLED.

5.2.4.2 Estado PORT_DISABLED

Existem duas condições de entrada nesse estado. A primeira é saindo do estado INITIALIZE e a segunda é quando a expressão "*((BEGIN = FALSE) and (port_enabled = FALSE) and (port_moved = FALSE))*" for satisfeita.

Uma vez nesse estado, a máquina executa a seguinte ação:

- *Partner_Oper_Port_State.Synchronization* <= FALSE

Esse estado possui três possibilidades de transição:

- *port_moved* = TRUE: condição para porta mudar seu estado para INITIALIZE (explicado na subseção anterior);
- (*port_enabled* = TRUE) and (*LACP_Enabled* = TRUE): condição para a porta mudar seu estado para EXPIRED;
- (*port_enabled* = TRUE) and (*LACP_Enabled* = FALSE): condição de transição para mudar o estado para LACP_DISABLED.

5.2.4.3 Estado LACP_DISABLED

A porta só entra nesse estado quando ao final da operação do estado PORT_DISABLED a avaliação da expressão (*port_enabled* = TRUE) and (*LACP_Enabled* = FALSE) seja verdadeira. Este estado é final e portanto, não possui condição de saída.

5.2.4.4 Estado CURRENT

A porta se mantém nesse estado enquanto recebe LACPDUs antes do temporizador *current_while_timer* expirar. Se esta porta não receber um LACPDU com informações do sistema Parceiro antes do temporizador expirar, então o estado da máquina mudará para EXPIRED.

As operações realizadas nesse estado são descritas a seguir.

- *update_Selected*: descrita na Seção 5.2.3;
- *update_NTT*: descrita na Seção 5.2.3;
- *recordPDU*: descrita na Seção 5.2.3;
- Inicializa um temporizador *current_while_timer* que conta o tempo em *Actor_Oper_Port_State.LACP_Timeout*;
- *Actor_Oper_Port_State.Expired* <= FALSE.

Após essas ações, a máquina fica esperando a chegada de um novo LACPDU do sistema Parceiro.

5.2.4.5 Estado EXPIRED

Este estado indica que as informações recebidas estão expiradas e portanto será disparado o temporizador *current_while_timer*. Se novas informações chegarem antes desse temporizador expirar, então a máquina muda de estado para CURRENT sem alterar os valores previamente recebidos. Se o temporizador expirar, então a máquina mudará seu estado para DEFAULTED.

As operações realizadas neste estado são descritas a seguir:

- *Partner_Oper_Port_State.Synchronization* <= FALSE;
- *Partner_Oper_Port_State.LACP_Timeout* <= *Short_Timeout_Time*;
- Inicializa um temporizador *current_while_timer* que conta o tempo em

Short_Timeout_Time;

- *Actor_Oper_Port_State.Expired* <= TRUE.

5.2.4.6 Estado DEFAULTED

O estado DEFAULTED é atingido caso não seja recebido LACPDU dentro do tempo contabilizado pelo temporizador disparado ao final do estado EXPIRED. Neste estado, a porta é configurada com suas informações padrão. A descrição das operações realizadas neste estado são apresentadas a seguir.

- *update_Default_Selected*: descrita na Seção 5.2.3;
- *record_Default*: descrita na Seção 5.2.3;
- *Actor_Oper_Port_State.Expired* <= FALSE.

5.3 Máquina Transmissora

A Máquina Transmissora (TX) envia LACPDUs sob demanda das outras máquinas de estado e de maneira periódica (Seção 5.4). A TX envia um LACPDU com informações de uma determinada porta quando a variável *LACP_Enabled* e *NTT* desta porta estiverem definidas com o valor TRUE. Após o envio de um LACPDU, a TX muda o estado da variável *NTT* da porta para FALSE.

Essa máquina deve garantir que no máximo três LACPDUs possam ser enviados dentro do limite de tempo definido pela variável *Fast_Periodic_Time*. Ou seja, se alguma outra máquina fizer uma requisição de envio de um LACPDU para a TX e esta já tiver enviado três LACPDUs dentro do tempo da constante *Fast_Periodic_Time*, o envio deste LACPDU será atrasado até que esta restrição de tempo não seja mais válida. O LACPDU que será enviado após a restrição de tempo, deverá ser preenchido com informações atuais e não com as informações que o Ator tinha no momento de sua requisição.

A TX é implementada através de uma função que é chamada quando alguma máquina de estados solicita o envio de um LACPDU. Esta função recebe como parâmetro o identificador da porta a qual deverá ser enviado um LACPDU.

5.4 Máquina de Transmissão Periódica

A Máquina Transmissora Periódica (Periodic) é responsável por controlar os envios periódicos de LACPDUs para o sistema Parceiro através de requisições à Máquina Transmissora, a fim de manter uma agregação. Ela também controla o intervalo entre estas transmissões, que é determinado pela variável *Partner_Oper_Port_State.Timeout* do sistema Ator (explicadas na próxima seção). O valor desta variável indica o tempo em que o sistema Parceiro considera válidas as informações recebidas.

As transmissões periódicas ocorrem em um enlace se a variável *Actor_Oper_Port_State.LACP_Activity* (Seção 4.2.2) da porta de pelo menos um dos sistemas

esteja definida como *LACP_Active*. Essas transmissões não ocorrem quando em ambos os sistemas, essa variável é definida com *LACP_Passive*. Da mesma forma, se *LACP_Enabled* ou *port_enabled* (Seção 4.2.2) estão definidas como FALSE, pois, estas variáveis indicam que o LACP está desabilitado nesta porta ou que a porta está em um estado não operacional respectivamente. Os estados dessa máquina são descritos a seguir.

- NO_PERIODIC: neste estado, o sistema não envia LACPDUs;
- FAST_PERIODIC: neste estado, as transmissões ocorrem com o intervalo de tempo definido na constante *Short_Timeout_Time*;
- SLOW_PERIODIC: neste estado, as transmissões ocorrem com o intervalo de tempo definido pela constante *Long_Timeout_Time*;
- PERIODIC_TX: a máquina entra neste estado quando quando o timer periódico (temporizador que controla o tempo que as informações de estado do Parceiro são válidas) expira. É neste estado que uma NTT é assinalada e o estado é alterado para FAST_PERIODIC ou SLOW_PERIODIC, dependendo da preferência do sistema parceiro.

As transições entre estes estados e os eventos que os disparam são apresentados no diagrama de estados da Figura 12.

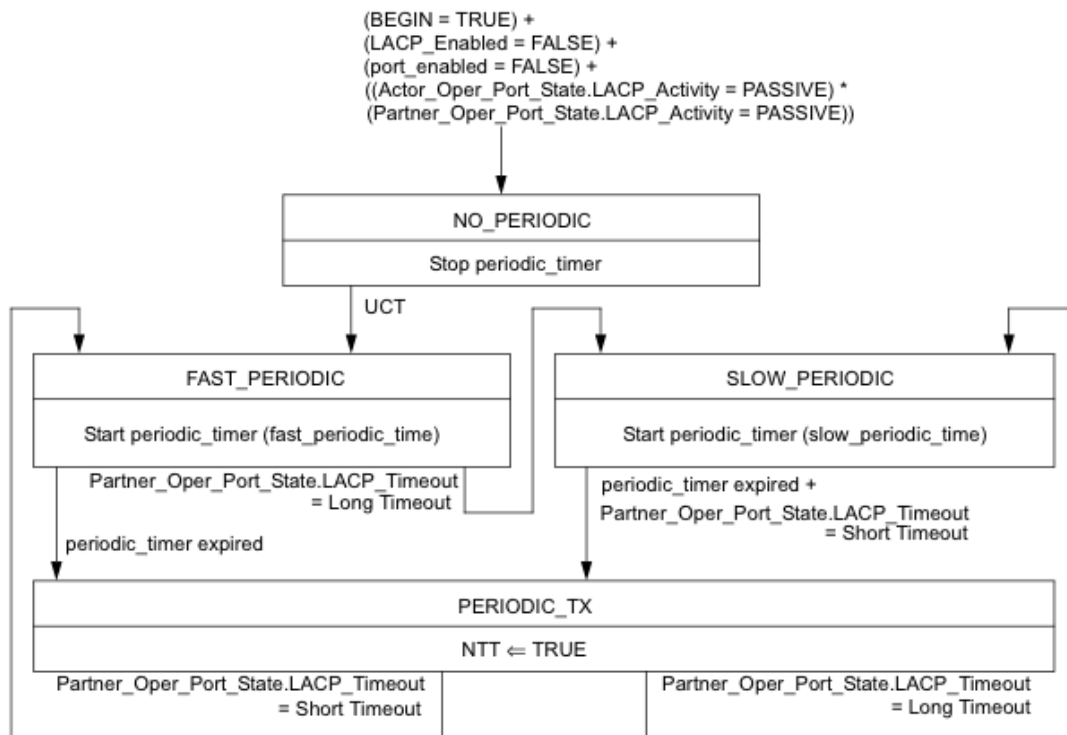


Figura 20 - Diagrama de estados da Máquina Transmissora/Periódica.

5.4.1 Constantes da *Periodic*

A *Periodic* utiliza duas constantes para o controle dos intervalos de envio de LACPDUs ao

sistema Parceiro. Este intervalo representa o tempo em que o sistema Parceiro leva para invalidar as informações que recebe. Para curtos intervalos de envio é usado o valor da constante *Fast_Periodic_Time* que possui o valor **um (segundo)**. Para intervalos longos de envio é usado o valor da constante *Slow_Periodic_Time* que possui o valor **trinta (segundos)**.

5.4.2 Temporizador da *Periodic*

A máquina periódica utiliza um temporizador para contabilizar o intervalo de envio de LACPDUs. Este temporizador se chama *periodic_timer* e recebe como parâmetro as constantes descritas na Seção anterior.

5.5 Lógica de Seleção

A Lógica de Seleção é o algoritmo responsável por selecionar um Agregador (Seção 2.1) compatível para uma porta. Esta seleção é feita através do LAG ID da porta. É responsável também por determinar se um enlace deve operar em STANDBY ou até mesmo definir o estado da variável *Selected* como UNSELECTED assim como a RX.

5.5.1 Exigências para a implementação da Lógica de Seleção

A Lógica de Seleção deve obedecer às seguintes exigências:

- Todo sistema deve ter no mínimo um Agregador;
- Cada Agregador deve ter uma Chave Operacional;
- Portas que podem ser agregadas devem ter a mesma Chave Operacional assim como as que devem operar sozinhas devem ter Chaves Operacionais distintas;
- Cada Agregador deve ter uma Chave Operacional e um identificador que o diferencie dos outros agregadores do sistema;
- Uma porta pode ser vinculada somente a um Agregador que tenha a mesma Chave Operacional que ela possui;
- Portas que fazem parte do mesmo LAG devem estar vinculadas ao mesmo Agregador;
- Portas com chaves operacionais únicas devem ser gerenciadas por um único agregador;
- Quando os parâmetros de seleção foram alterados (algum dos componentes do LAG ID)
- Uma porta não deve ser habilitada para transferir quadros enquanto não estiver **selecionada** e **anexada** a um agregador.

5.5.2 Pseudo-algoritmo da Lógica de Seleção

Esta subseção apresenta o pseudo-algoritmo da Lógica de Seleção. As funções desempenhadas por ela estão descritas na Seção 3.3.

Entrada: porta (referência da porta que será analisada)

```

1.  se esta porta estiver vinculada a algum agregador
2.      remove a porta da lista de portas do agregador
3.      define o agregador desta porta como indefinido
4.
5.      se ainda tiver alguma porta nesse agregador
6.          define o id do agregador com o mesmo id da porta de maior prioridade dentre as portas vinculadas a esse agregador
7.          para todas as portas que ainda estão vinculadas a este agregador
8.              atualiza o valor do id do agregador destas portas
9.          atualiza o valor do agregador do enlace
10.
11.      se neste agregador não tiver mais nenhuma porta vinculada
12.          remove o agregador
13.          define o agregador do enlace como indefinido
14.
15.  se esta porta não é individual
16.      preenche link_info
17.      se já existe agregador com esta chave
18.          se esta porta não está no mesmo LAG desse agregador
19.              se a prioridade desta porta é menor que a prioridade de pelo menos uma das portas que fazem parte deste agregador
20.                  define o valor do agregador desta porta como indefinido
21.                  define o valor do agregador deste enlace como indefinido
22.                  porta <- standby
23.                  return
24.              se a prioridade desta porta é maior que a prioridade de todas as portas vinculadas a este agregador
25.                  para todas as portas deste agregador
26.                      desvincula a porta do agregador
27.                      define o valor do agregador da porta como indefinido
28.                      define o valor do agregador do enlace como indefinido
29.                      porta <- unselected
30.
31.                  define o id do agregador no enlace desta porta
32.                  adiciona esta porta no agregador
33.                  define o id do agregador com o mesmo id desta porta
34.                  define o id do agregador da porta como id deste agregador
35.          se esta porta está no mesmo LAG das portas desse agregador
36.              se a prioridade desta porta é menor que a prioridade de pelo menos uma das portas vinculadas a este agregador
37.                  adiciona esta porta na lista de portas desse agregador
38.                  define este agregador como agregador desta porta
39.                  define este agregador como agregador deste enlace
40.
41.              se a prioridade desta porta é maior que a prioridade de todas as portas vinculadas a esse agregador
42.                  muda o id do agregador para o mesmo id da porta
43.                  para todas as portas com id menor que estão vinculadas a esse agregador
44.                      muda o id do agregador desta porta
45.                  adiciona esta porta na lista de portas desse agregador
46.                  define este agregador como agregador do enlace
47.                  define este agregador como agregador desta porta
48.
49.      se não existe agregador com a mesma chave desta porta
50.          cria um agregador
51.          define a chave deste agregador com a chave desta porta
52.          define o id do agregador com o mesmo id da porta
53.          adiciona esta porta na lista de portas desse agregador
54.          define o agregador desta porta com o id deste agregador
55.          define o agregador do enlace com o id desta porta
56.  se esta porta é individual
57.      cria um agregador
58.      define a chave deste agregador com a chave desta porta
59.      define o id do agregador com o mesmo id desta porta
60.      adiciona esta porta na lista de portas desse agregador
61.      define o agregador desta porta com o id deste agregador
62.      define o agregador deste enlace com o id deste agregador
63.
64.  se não existem restrições de agregação
65.      porta <- selected
66.
67.  se existem restrições de agregação
68.      porta <- standby

```

Figura 21 - Pseudo-algoritmo da Lógica de Seleção.

5.6 Máquina Multiplexadora

A MUX vincula/desvincula uma porta de um Agregador que foi selecionado pela Máquina de Lógica de Seleção e muda o estado da porta no que diz respeito a receber e transmitir quadros. A Figura 22 apresenta o diagrama de estados da MUX.

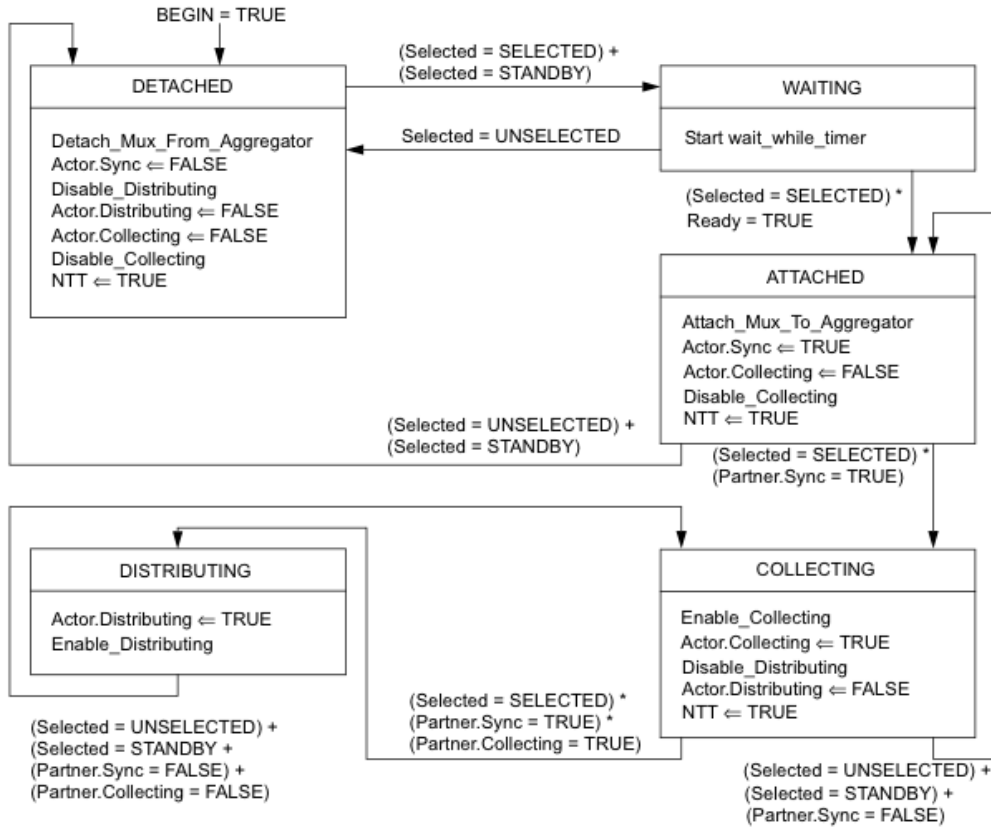


Figura 22 - Diagrama de estados da Máquina Multiplexadora.

5.6.1 Estados da MUX

A Máquina Multiplexadora controla a porta no que diz respeito a ela estar vinculada a um Agregador e também seu estado de coleta e recepção de quadros do sistema Parceiro. Esta máquina possui cinco estados que são descritos a seguir.

5.6.1.1 Estado DETACHED

Este é o estado inicial da MUX e nele a porta ainda não está vinculada a nenhum Agregador. Neste estado as seguintes ações são executadas.

- *Detach_Mux_From_Aggregator*: descrita na Seção 5.6.2;
- *Actor_Oper_Port_State.Synchronization* \Leftarrow FALSE;
- *Disable_Distributing*: descrita na Seção 5.6.2;
- *Actor_Oper_Port_State.Distributing* \Leftarrow FALSE;
- *Actor_Oper_Port_State.Collecting* \Leftarrow FALSE;
- *Disable_Collecting*: descrita na Seção 5.6.2;

- *NTT* <= TRUE.

A condição de transição desse estado para WAITING depende do estado da variável *Selected* desta porta. Esta variável deve ser SELECTED ou STANDBY para que a transição seja efetuada. A definição desta variável é feita pela Lógica de Seleção.

5.6.1.2 Estado WAITING

Uma vez nesse estado, é disparado o temporizador *wait_while_timer* descrito na Seção 5.6.4. Este temporizador serve para que as outras portas possam ser selecionadas pela Lógica de Seleção e mudem seu estado para ATTACHED para serem anexadas em conjunto ao Agregador correspondente. A máquina muda para o estado ATTACHED quando após o temporizador expirar a variável *Selected* desta porta seja SELECTED e o valor da variável *Ready* seja TRUE. Caso a variável *Selected* tenha sido mudada para UNSELECTED, a máquina irá voltar para o estado DETACHED.

5.6.1.3 Estado ATTACHED

Este é o estado aonde a MUX anexa a porta ao Agregador selecionado pela Lógica de Seleção. Neste estado, as seguintes ações são executadas.

- *Attach_Mux_To_Aggregator*: descrita na Seção 5.6.2;
- *Actor_Oper_Port_State.Synchronization* <= TRUE;
- *Actor_Oper_Port_State.Collecting* <= FALSE;
- *Disable_Collecting*: descrita na Seção 5.6.2;
- *NTT* <= TRUE.

Essa solicitação de envio de informações representada pela atribuição do valor TRUE à variável *NTT* ocasionará o envio de um LACPDU ao sistema Parceiro. A condição de saída desse estado é a porta ter o valor da variável *Selected* como SELECTED e o sistema Parceiro estar em sincronia com o Ator, ou seja, a variável *Partner_Oper_Port_State.Synchronization* = TRUE. Com essas condições satisfeitas a máquina muda para o estado COLLECTING. Se ao final da operação desse estado, a variável *Selected* for alterada para UNSELECTED ou STANDBY, então a máquina transitará para o estado DETACHED.

5.6.1.4 Estado COLLECTING

Este é o estado da MUX que sinaliza o sistema Parceiro que esta porta começou a coletar quadros. Tal ação se reflete na chamada da função *Enable_Collecting* (descrita na Seção 5.6.2) e na mudança do estado da variável *Actor_Oper_Port_State.Collecting* para TRUE.

Este é o estado da MUX que altera as variáveis de controle da porta indicando que esta porta está pronta para receber quadros - através do comando - do sistema Parceiro. Também neste estado é desabilitada a distribuição de quadros da porta através da chamada da função *Disable_Distributing* (descrita na Seção 5.6.2) e na mudança do estado da variável

Actor_Oper_Port_State.Distributing para FALSE. Ao final do estado, essas mudanças devem ser avisadas ao sistema Parceiro e isso é feito através da mudança do estado da variável *NTT* desta porta que recebe o valor TRUE.

A MUX muda seu estado para DISTRIBUTING se a variável *Selected* desta porta ainda estiver com o valor SELECTED e quando o sistema Parceiro entra em sincronia com o Ator (*Partner_Oper_Port_State.Synchronization* = TRUE) e habilita sua recepção de quadros (*Partner_Oper_Port_State.Collecting* = TRUE).

5.6.1.5 Estado DISTRIBUTING

No estado DISTRIBUTING, a MUX habilita a distribuição de quadros dessa porta chamando a função *Enable_Distributing* e define corretamente o valor da variável correspondente a esta ação (*Actor_Oper_Port_State.Distributing* <= TRUE).

A partir deste estado, a máquina MUX somente pode voltar para o estado COLLECTING e esta transição será feita caso alguma das condições descritas abaixo seja satisfeita.

- *Selected* = UNSELECTED ou STANDBY: se a RX ou a Lógica de Seleção mudarem o estado desta variável (e.g.: se a chave desta porta mudar);
- *Partner_Oper_Port_State.Synchronization* = FALSE: se o sistema Parceiro avisar que está fora de sincronia;
- *Partner_Oper_Port_State.Collecting* = FALSE: se o sistema Parceiro parar de coletar quadros.

5.6.2 Funções da MUX

As funções da MUX são implementadas na camada abaixo desta subcamada e portanto, são implementadas em *hardware*. A descrição lógica dessas funções são descritas a seguir.

- *Attach_Mux_To_Aggregator*: esta função faz com que o CPM desta porta seja vinculado ao APM do Agregador que foi selecionado para controlar esta porta. Esta função prepara a porta para receber e distribuir quadros;
- *Detach_Mux_To_Aggregator*: esta função faz com que o CPM desta porta seja desvinculado ao APM do Agregador que controla esta porta;
- *Enable_Collecting*: esta função faz com que o APM do Agregador que controla esta porta inicie a coleta de quadros dos enlaces;
- *Disable_Collecting*: esta função faz com que o APM do Agregador que controla esta porta pare de coletar quadros dos enlaces do LAG ao qual participa;
- *Enable_Distributing*: esta função faz com que o APM do Agregador que controla esta porta inicie a distribuição de quadros desta porta para os enlaces do LAG ao qual participa;
- *Disable_Distributing*: esta função faz com que o APM do Agregador que controla esta

porta pare a distribuição de quadros desta porta para os enlaces do LAG ao qual participa.

5.6.3 Constante da MUX

Esta máquina de estados utiliza a constante *Aggregate_Wait_Time* que armazena o número de segundos que o temporizador *wait_while_timer* deve usar. Este temporizador é explicado na Seção seguinte.

5.6.4 Temporizador da MUX

A MUX utiliza um temporizador chamado *wait_while_timer*. Este temporizador atrasa a vinculação das portas ao Agregador selecionado, afim de permitir que várias portas se vinculem simultaneamente.

6 RESULTADOS

Este Capítulo apresenta a descrição de um teste de execução do LACP. Este teste é descrito através de captura das telas de depuração das portas de ambos os sistemas. Além disto, é descrito de maneira sucinta o simulador responsável pela comunicação destes sistemas.

6.1 Simulador

O LACP é um *software* que executará em uma plataforma de *hardware*. Este *software* se comunica com este *hardware* através de funções disponibilizadas pela Camada de Abstração de *Hardware* (HAL). O LACP foi desenvolvido sem esta plataforma e para simular o seu comportamento foi desenvolvido um simulador que é responsável pela comunicação entre os sistemas. O simulador e o Protocolo de Marcação (Seção **Error! Reference source not found.**), foram desenvolvidos para operar em conjunto com o LACP [BGF11]. A arquitetura desse simulador é apresentada na Figura 23.

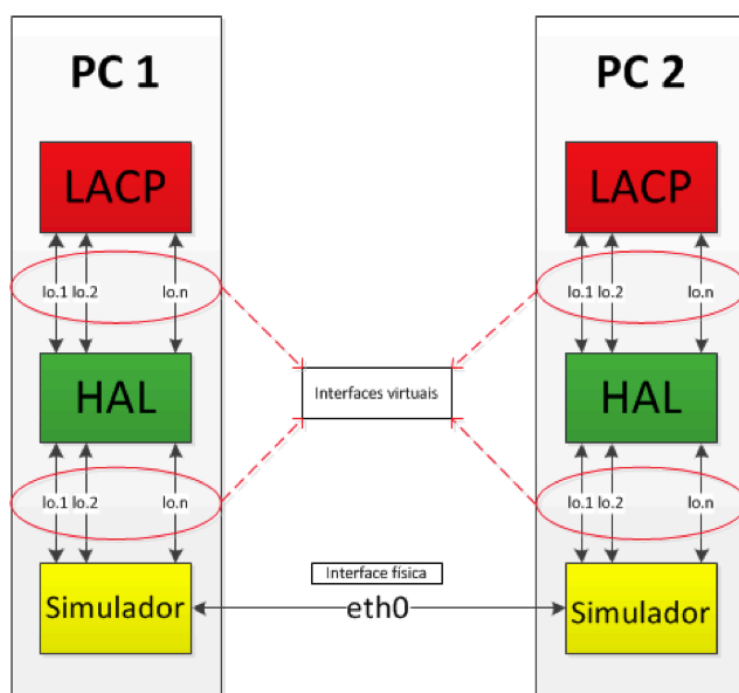


Figura 23 - Arquitetura de comunicação entre o LACP, HAL e o Simulador [BGF11].

6.2 Configuração de Teste

Para demonstrar o funcionamento do LACP, foi definida uma configuração utilizando quatro portas em ambos sistemas, chamados de Sistema A e Sistema B. O Sistema A possui quatro portas enumeradas de um a quatro, assim como o sistema B. No sistema A as portas um, dois e três possuem a mesma chave operacional, possibilitando assim, uma agregação entre elas e a porta quatro possui uma chave diferente que determina a incapacidade desta porta de ser agregada com qualquer outra.

No sistema B, somente as portas um e dois podem ser agregadas, ao passo que, as portas três

e quatro devem ser individuais e, portanto, possuem chaves diferentes das demais.

Para melhor compreensão, é apresentada a Figura 24 que representa de maneira abstrata a configuração utilizada no teste.

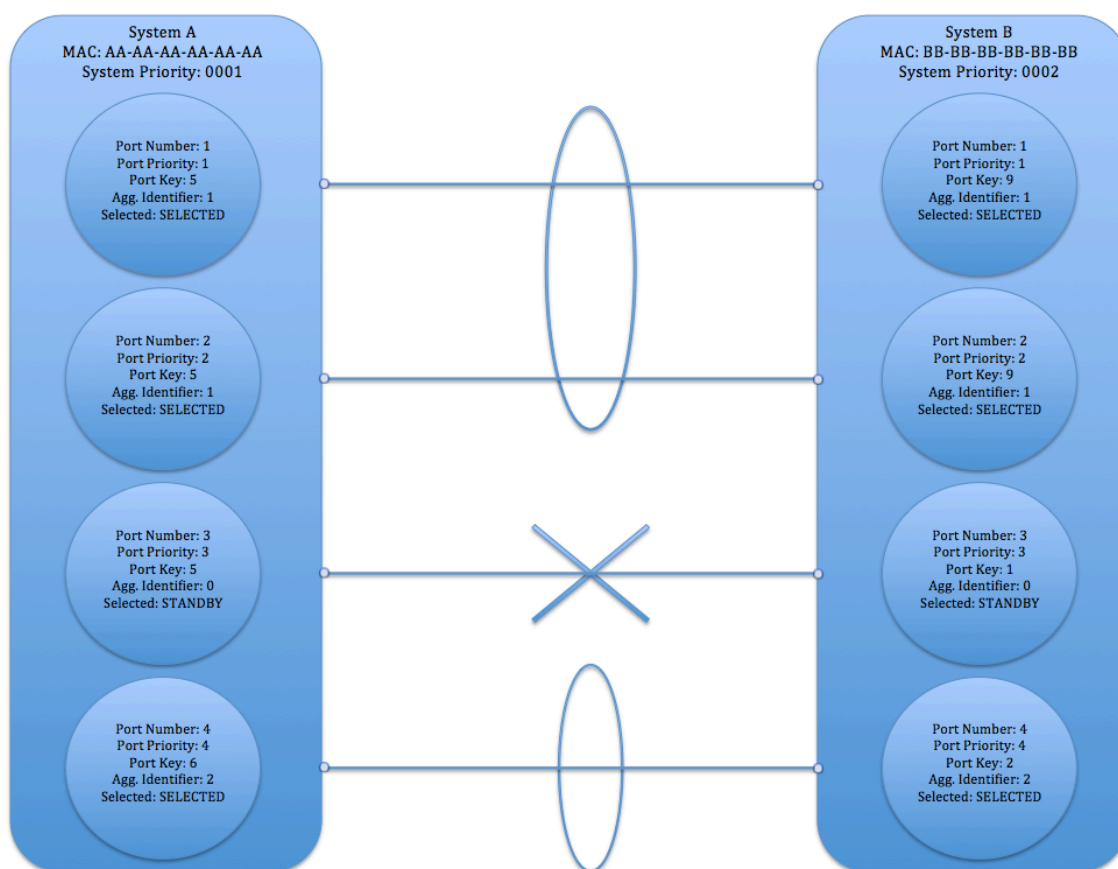


Figura 24 - Exemplo de configuração do LACP.

Para fins de depuração, as máquinas de estados imprimem informações na saída padrão. São informações sobre os estados das máquinas, o enlace (Seção 4.2.7) ao qual a porta está conectada, o LAG ID deste enlace bem como o identificador do Agregador que esta porta está vinculada, os valores de estado do Ator (*Actor_Port_State*) e o valor da variável *Selected*.

Esta configuração de teste apresenta dois tipos de agregação de enlaces, sendo a primeira formada por dois enlaces compostos pelas portas **um** e **dois** de ambos os sistemas e a segunda agregação formada pelos enlaces individuais das portas **quatro** de ambos sistemas.

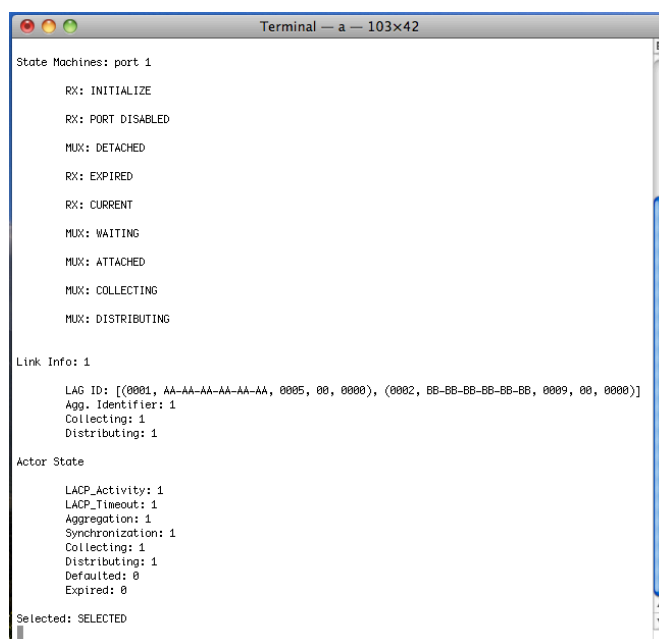
O sistema A foi configurado para suportar uma agregação composta pelas portas um, dois e três e outra individual pela porta quatro. Já o sistema B está configurado para operar utilizando uma agregação com as portas um e dois, e outras duas individuais sendo uma utilizando a porta três e a segunda a porta quatro.

Na parte superior das telas de depuração, (*State Machines: port 1*) é apresentado o fluxo do LACP através das máquinas de estados. Logo abaixo (*Link Info: 1*) são descritas informações do enlace desta porta que contém o LAG ID, o identificador do Agregador que controla esta porta e o estado das funções de distribuição e recepção de quadros nesse enlace. Por fim, é apresentado o

valor das variáveis de estado da porta (*Actor State*) e o valor da variável *Selected*.

6.2.1 Análise da configuração final do Sistema A

No Sistema A, as portas um e dois formaram uma agregação e estão prontas para trocar quadros pois, a RX de ambas as portas está no estado CURRENT e a MUX das duas portas está no estado DISTRIBUTING. A Lógica de Seleção selecionou essas duas portas para fazerem parte de uma mesma agregação que é controlada pelo Agregador que possui identificador **um**. Esta escolha foi baseada nos LAG IDs dessas duas portas que são idênticos.



```

Terminal - a - 103x42

State Machines: port 1

RX: INITIALIZE
RX: PORT DISABLED
MUX: DETACHED
RX: EXPIRED
RX: CURRENT
MUX: WAITING
MUX: ATTACHED
MUX: COLLECTING
MUX: DISTRIBUTING

Link Info: 1

LAG ID: [(0001, AA-AA-AA-AA-AA-AA, 0005, 00, 0000), (0002, BB-BB-BB-BB-BB, 0009, 00, 0000)]
Agg. Identifier: 1
Collecting: 1
Distributing: 1

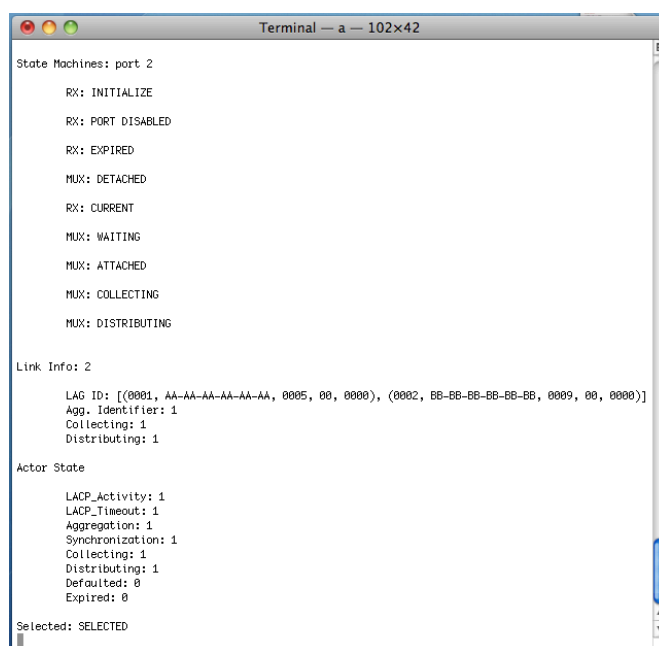
Actor State

LACP_Activity: 1
LACP_Timeout: 1
Aggregation: 1
Synchronization: 1
Collecting: 1
Distributing: 1
Defaulted: 0
Expired: 0

Selected: SELECTED

```

Figura 25 - Captura de tela de depuração da porta 1 do Sistema A.



```

Terminal - a - 102x42

State Machines: port 2

RX: INITIALIZE
RX: PORT DISABLED
RX: EXPIRED
MUX: DETACHED
RX: CURRENT
MUX: WAITING
MUX: ATTACHED
MUX: COLLECTING
MUX: DISTRIBUTING

Link Info: 2

LAG ID: [(0001, AA-AA-AA-AA-AA-AA, 0005, 00, 0000), (0002, BB-BB-BB-BB-BB, 0009, 00, 0000)]
Agg. Identifier: 1
Collecting: 1
Distributing: 1

Actor State

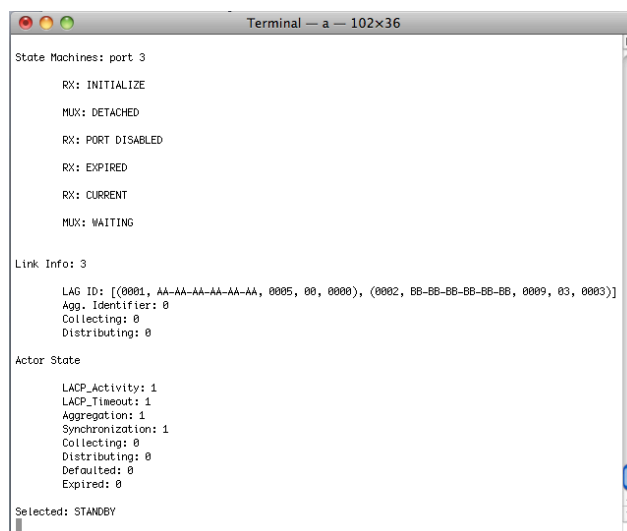
LACP_Activity: 1
LACP_Timeout: 1
Aggregation: 1
Synchronization: 1
Collecting: 1
Distributing: 1
Defaulted: 0
Expired: 0

Selected: SELECTED

```

Figura 26 - Captura de tela de depuração da porta 2 do Sistema A.

Embora exista um Agregador que controle as portas com a chave da porta **três**, ela não foi selecionada para fazer parte desta Agregação, pois seu LAG ID não é igual ao das portas **um** e **dois** que são as portas que possuem o mesmo valor de chave operacional. Portanto, a variável *Selected* desta porta está no estado STANDBY, pois a Lógica de Seleção não vinculou esta porta a nenhum Agregador compatível e, por isto, a MUX desta porta está no estado WAITING. O estado das variáveis de controle desta porta são apresentadas na Figura 27.



```

Terminal - a - 102x36

State Machines: port 3

RX: INITIALIZE
RX: DETACHED
RX: PORT DISABLED
RX: EXPIRED
RX: CURRENT
MUX: WAITING

Link Info: 3

LAG ID: [(0001, AA-AA-AA-AA-AA, 0005, 00, 0000), (0002, BB-BB-BB-BB-BB, 0009, 03, 0003)]
Agg. Identifier: 0
Collecting: 0
Distributing: 0

Actor State

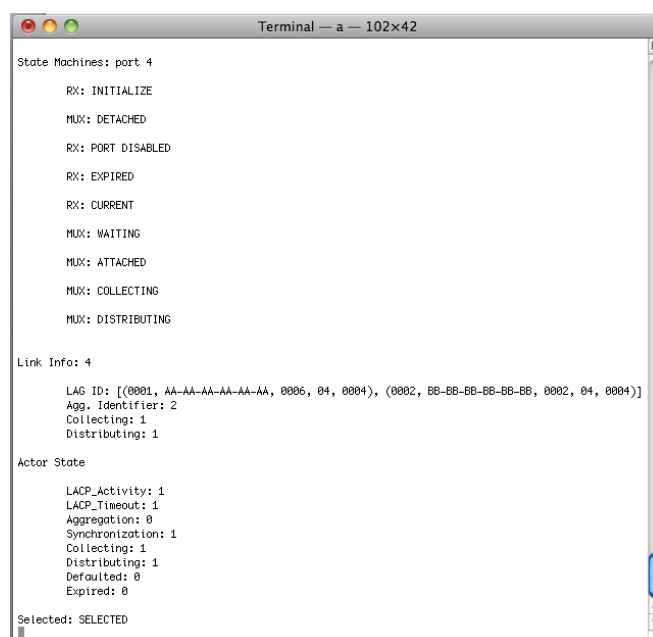
LACP_Activity: 1
LACP_Timeout: 1
Aggregation: 1
Synchronization: 1
Collecting: 0
Distributing: 0
Defaulted: 0
Expired: 0

Selected: STANDBY

```

Figura 27 - Captura de tela de depuração da porta 3 do Sistema A.

A porta quatro recebeu uma chave operacional única no sistema A e, por isto, deve operar de modo individual, ou seja, não deve ser agregada a nenhuma outra porta neste Sistema. Esta porta está distribuindo e recebendo quadros pois, a Lógica de Seleção encontrou um Agregador compatível para ela e, também, porque a porta na outra ponta do enlace também está configurada para operar de maneira individual fazendo com que estas portas entrem em sincronia.



```

Terminal - a - 102x42

State Machines: port 4

RX: INITIALIZE
RX: DETACHED
RX: PORT DISABLED
RX: EXPIRED
RX: CURRENT
MUX: WAITING
MUX: ATTACHED
MUX: COLLECTING
MUX: DISTRIBUTING

Link Info: 4

LAG ID: [(0001, AA-AA-AA-AA-AA, 0006, 04, 0004), (0002, BB-BB-BB-BB-BB, 0002, 04, 0004)]
Agg. Identifier: 2
Collecting: 1
Distributing: 1

Actor State

LACP_Activity: 1
LACP_Timeout: 1
Aggregation: 0
Synchronization: 1
Collecting: 1
Distributing: 1
Defaulted: 0
Expired: 0

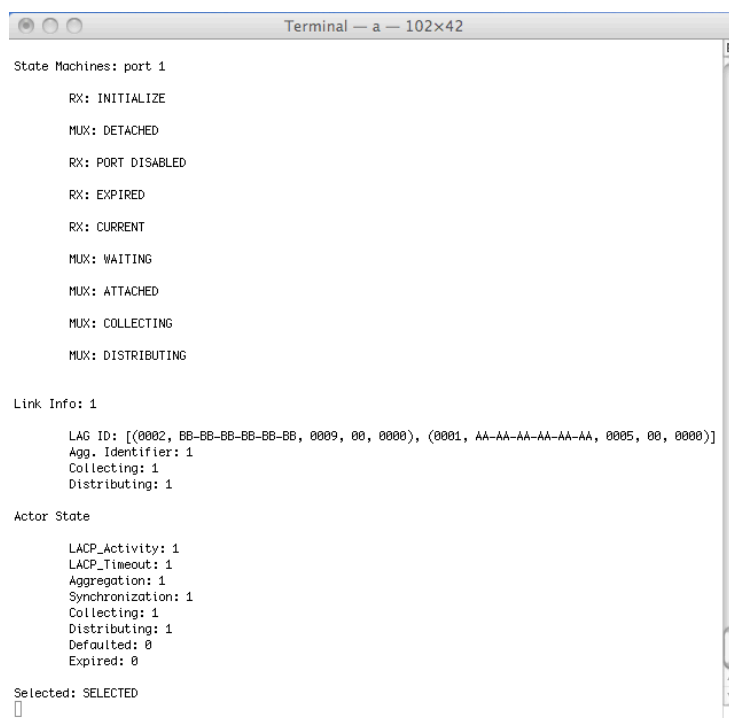
Selected: SELECTED

```


Figura 28 - Captura de tela de depuração da porta 4 do Sistema A.

6.2.2 Análise da configuração final do Sistema B

Analogamente às portas **um** e **dois** do Sistema A, as portas **um** e **dois** do Sistema B fazem parte de um mesmo Agregador e estão trocando quadros entre os enalces.



```

Terminal - a - 102x42

State Machines: port 1

  RX: INITIALIZE
  MUX: DETACHED
  RX: PORT DISABLED
  RX: EXPIRED
  RX: CURRENT
  MUX: WAITING
  MUX: ATTACHED
  MUX: COLLECTING
  MUX: DISTRIBUTING

Link Info: 1

  LAG ID: [(0002, BB-BB-BB-BB-BB, 0009, 00, 0000), (0001, AA-AA-AA-AA-AA, 0005, 00, 0000)]
  Agg. Identifier: 1
  Collecting: 1
  Distributing: 1

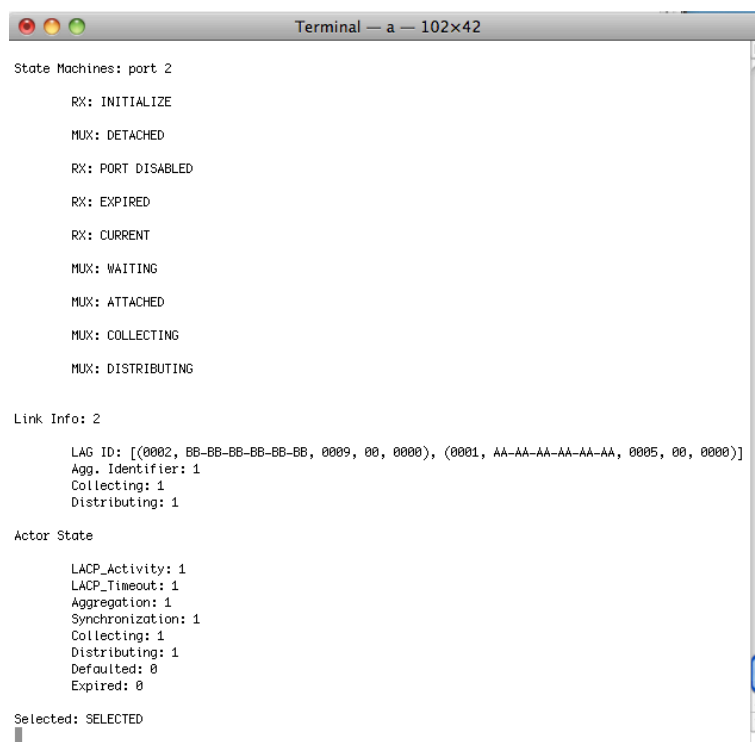
Actor State

  LACP_Activity: 1
  LACP_Timeout: 1
  Aggregation: 1
  Synchronization: 1
  Collecting: 1
  Distributing: 1
  Defaulted: 0
  Expired: 0

Selected: SELECTED

```

Figura 29 - Captura de tela de depuração da porta 1 do Sistema B.



```

Terminal - a - 102x42

State Machines: port 2

  RX: INITIALIZE
  MUX: DETACHED
  RX: PORT DISABLED
  RX: EXPIRED
  RX: CURRENT
  MUX: WAITING
  MUX: ATTACHED
  MUX: COLLECTING
  MUX: DISTRIBUTING

Link Info: 2

  LAG ID: [(0002, BB-BB-BB-BB-BB, 0009, 00, 0000), (0001, AA-AA-AA-AA-AA, 0005, 00, 0000)]
  Agg. Identifier: 1
  Collecting: 1
  Distributing: 1

Actor State

  LACP_Activity: 1
  LACP_Timeout: 1
  Aggregation: 1
  Synchronization: 1
  Collecting: 1
  Distributing: 1
  Defaulted: 0
  Expired: 0

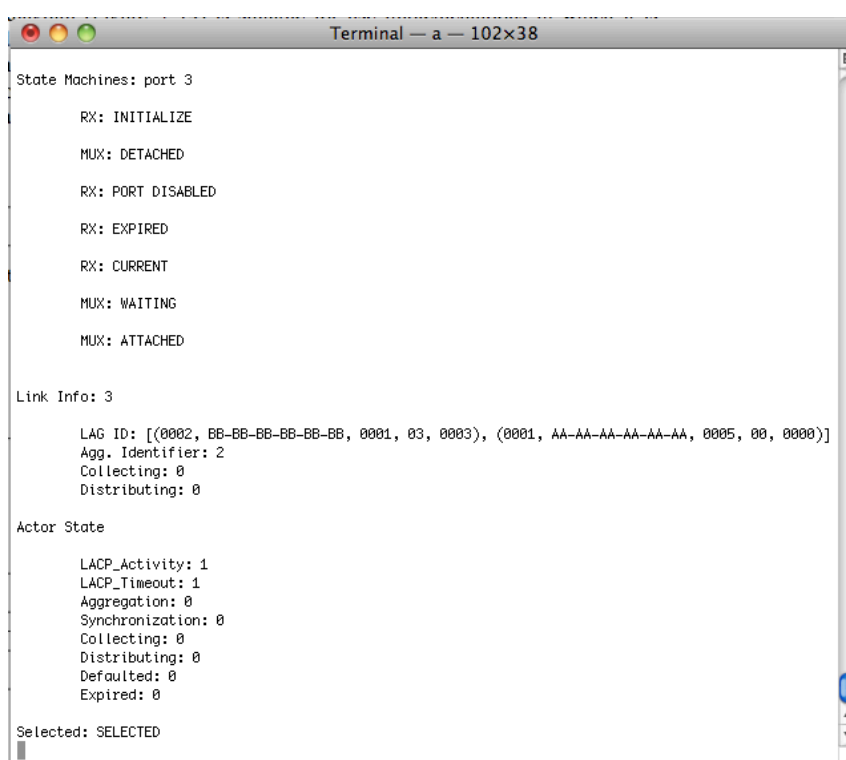
Selected: SELECTED

```

Figura 30 - Captura de tela de depuração da porta 2 do Sistema B.

A porta **três** do Sistema B possui uma chave única em seu Sistema e por isso deve operar de forma individual portanto, seu LAG ID e chave operacional são únicos no sistema fazendo com que a Lógica de Seleção associe esta porta a um Agregador. A máquina MUX desta porta está travada no estado ATTACHED e não consegue fazer a transição para o estado COLLECTING pois não consegue se sincronizar com a porta **três** do Sistema A.

Esta sincronização não acontece pois, a porta **três** do sistema A não está configurada para operar de maneira individual, logo, a variável *Actor_Oper_Port_State.Synchronization* nunca assume o valor TRUE e portanto, estas portas não estão trocando quadros.



```

State Machines: port 3

  RX: INITIALIZE

  MUX: DETACHED

  RX: PORT DISABLED

  RX: EXPIRED

  RX: CURRENT

  MUX: WAITING

  MUX: ATTACHED

Link Info: 3

  LAG ID: [(0002, BB-BB-BB-BB-BB, 0001, 03, 0003), (0001, AA-AA-AA-AA-AA-AA, 0005, 00, 0000)]
  Agg. Identifier: 2
  Collecting: 0
  Distributing: 0

Actor State

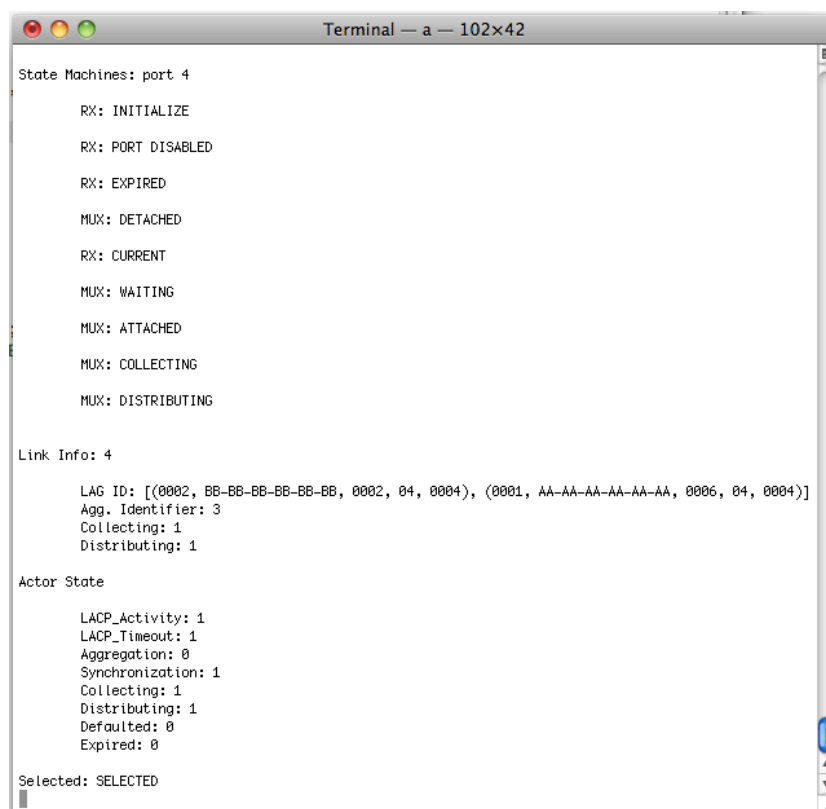
  LACP_Activity: 1
  LACP_Timeout: 1
  Aggregation: 0
  Synchronization: 0
  Collecting: 0
  Distributing: 0
  Defaulted: 0
  Expired: 0

Selected: SELECTED

```

Figura 31 - Captura de tela de depuração da porta 3 do Sistema B.

Analogamente à porta **quatro** do Sistema A, a porta **quatro** do sistema B está corretamente configurada como individual e ambas as portas estão trocando quadros. A máquina MUX da porta **quatro** está no estado DISTRIBUTING e a RX no estado CURRENT, denotando a correta configuração do LACP neste enlace.

A terminal window titled "Terminal — a — 102x42" displays network state information. The text is as follows:

```
State Machines: port 4

RX: INITIALIZE
RX: PORT DISABLED
RX: EXPIRED
MUX: DETACHED
RX: CURRENT
MUX: WAITING
MUX: ATTACHED
MUX: COLLECTING
MUX: DISTRIBUTING

Link Info: 4

LAG ID: [(0002, BB-BB-BB-BB-BB, 0002, 04, 0004), (0001, AA-AA-AA-AA-AA, 0006, 04, 0004)]
Agg. Identifier: 3
Collecting: 1
Distributing: 1

Actor State

LACP_Activity: 1
LACP_Timeout: 1
Aggregation: 0
Synchronization: 1
Collecting: 1
Distributing: 1
Defaulted: 0
Expired: 0

Selected: SELECTED
```

Figura 32 - Captura de tela de depuração da porta 4 do Sistema B.

7 CONCLUSÃO

A partir do estudo da norma que define o protocolo e de documentos que detalham suas funções foi possível entender os objetivos do padrão 802.1AX bem como seu modo de operação. Suas principais funções foram estudadas e especificadas durante a realização do trabalho.

O protocolo permite aumentar a vazão de informações, assim como distribui de forma otimizada o tráfego de dados utilizando o balanceamento de carga. Possibilita também a realocação, em tempo de execução, do fluxo de informações para enlaces diferentes.

Em relação à definição das estruturas de dados utilizadas na implementação do protocolo, percebeu-se que, por consequência destas estruturas serem compartilhadas pelas máquinas de estados, há concorrência para acessá-las. Em virtude desta concorrência, foram utilizados recursos da linguagem de programação adotada para controle dessas áreas de memória (áreas críticas).

No decorrer do desenvolvimento deste trabalho a equipe teve diversas dificuldades como a interpretação da norma que descreve o protocolo. Estas dificuldades geraram retrabalho, pois as conclusões da equipe muitas vezes estavam equivocadas.

Com relação à participação do Autor em um projeto de pesquisa, ao longo de dois semestres de estudo e trabalho foi adquirida experiência no que diz respeito a fazer parte de um grupo de pesquisa e desenvolvimento. Foi aprendida uma metodologia de trabalho que foi composta de reuniões periódicas e metas semanais, gerando assim o conteúdo intelectual que está descrito no presente trabalho.

Referências Bibliográficas

- [BGF11] BONDAN, Lucas; GOBBI, Rodrigo; FOCHI, Vinícius. “O Protocolo de Agregação de Enlaces LACP: Um Simulador e o Protocolo de Marcação de Pacotes”, Trabalho de Conclusão de Curso em Engenharia de Computação, PUCRS, 2011, 61p.
- [IEE08] IEEE COMPUTER SOCIETY. “Link Aggregation”. Acesso em: <http://standards.ieee.org/getieee802/download/802.1AX-2008.pdf>, Março 2008.
- [SYS02] SYSKONNECT. “Link Aggregation according to IEEE 802.3ad”. Acesso em: <http://legacyweb.triumf.ca/canarie/amsterdam-test/References/wp-lag-e.pdf>, Março 2011.
- [TAN03] Tanenbaum, Andrew S. “Computer Networks”. 4 ed. Upper Saddle River: Prentice Hall, 2003, 891p.
- [WIK11a] WIKIPEDIA. “Link Aggregation”. Acesso em: http://en.wikipedia.org/wiki/Link_aggregation, Março 2011.
- [WIK11b] WIKIPEDIA. “Channel Bonding”. Acesso em: http://en.wikipedia.org/wiki/Channel_bonding, Março 2011.