

UNIVERSIDADE DO ESTADO DE MATO GROSSO
FACULDADE DE CIÊNCIAS EXATAS E TECNOLÓGICAS
CURSO DE BACHARELADO EM SISTEMAS DE INFORMAÇÃO

KENNEDY OLIVEIRA ROCHA

**UM ESTUDO COMPARATIVO DE SISTEMAS DE INTERCONEXÃO
PARA COMUNICAÇÃO INTERCHIP**

UNEMAT – Campus de Sinop

2019/1

UNIVERSIDADE DO ESTADO DE MATO GROSSO
FACULDADE DE CIÊNCIAS EXATAS E TECNOLÓGICAS
CURSO DE BACHARELADO EM SISTEMAS DE INFORMAÇÃO

KENNEDY OLIVEIRA ROCHA

**UM ESTUDO COMPARATIVO DE SISTEMAS DE INTERCONEXÃO
PARA COMUNICAÇÃO INTERCHIP**

Projeto de Pesquisa apresentado à Banca Examinadora do Curso de Bacharelado em Sistemas de Informação – UNEMAT, Campus Universitário de Sinop – MT, como requisito parcial para obtenção do grau de Bacharel em Sistemas de Informação, sob orientação do Prof. Dr. Ivan Luiz Pedroso Pires.

UNEMAT – Campus de Sinop

2019/1

LISTA DE TABELAS

Tabela 6.1 — Formato de <i>Frame</i> e Pacote Ethernet.....	15
Tabela 6.2 — Especificações de Normas 803.2.....	17

LISTA DE FIGURAS

Figura 6.1 — Subcamadas Ethernet 100 Gigabit	13
Figura 6.2 — Arquitetura ENoC.....	18

LISTA DE ABREVIATURAS

100GBASE-R	Família de dispositivos Ethernet que trabalham a uma velocidade de 100 Gb/s e codificação 64b/66
100GE	Ethernet 100 Gigabit
10GE	Ethernet 10 Gigabit
25GE	Ethernet 25 Gigabit
400GE	Ethernet 400 Gigabit
802.3	Grupo de trabalho, pertencente ao IEEE, responsável pela padronização da Ethernet
CSMA/CD	Carrier Sense Multiple Access with Collision Detection
DVD	Digital Versatile Disc
ENoC	Expansible Network on a Chip
FEC	Forward Error Correction
HCA	Host Channel Adapter
IB	InfiniBand
IEEE	Institute of Electrical and Electronics Engineers
LAN	Local Area Network
LLC	Logical Link Control
MAC	Medium Access Control
MDI	Medium Dependent Interface
MII	Media Independent Interface
OM3	Fibra Ótica com largura de 1500 MHz
OM4	Fibra Ótica com largura de 3500 MHz
OSI	Open Systems Interconnection
PCS	Physical Coding Sublayer
PE	Elementos de Processamento
PHY	Physical Layer Device
PLS	Physical Signaling Sublayer
PMA	Physical Medium Attachment
PMD	Physical Medium Dependent
QR	Queue Pair
RS	Reconciliation Sublayer

RS-FEC	Reed-Solomon Forward Error Correction
SFD	Start Frame Delimiter
SoC	System on a Chip
TCA	Target Channel Adapter

SUMÁRIO

1	INTRODUÇÃO	8
2	JUSTIFICATIVA	9
3	HIPÓTESE	10
4	OBJETIVOS	11
4.1	OBJETIVO GERAL	11
4.2	OBJETIVOS ESPECÍFICOS	11
5	FUNDAMENTAÇÃO TEÓRICA	12
5.1	ETHERNET	12
5.1.1	Camada Física	13
5.1.1.1	<i>Physical Coding Sublayer</i>	13
5.1.1.2	<i>Forward Error Correction</i>	14
5.1.1.3	<i>Physical Medium Attachment</i>	14
5.1.1.4	<i>Physical Medium Dependent</i>	14
5.1.1.5	<i>Medium Dependent Interface</i>	14
5.1.2	Camada de Enlace	15
5.1.3	Reconciliador	16
5.1.4	Evolução	17
5.2	<i>EXPANSIBLE NETWORK ON A CHIP (ENOC)</i>	18
5.3	INFINIBAND (IB)	19
5.3.1	Camada Física	19
5.3.2	Camada de Enlace	20
6	METODOLOGIA	22
7	CRONOGRAMA	23
	BIBLIOGRAFIA	24

1 INTRODUÇÃO

Chip é um circuito integrado que possui vários componentes elétricos e são largamente utilizados na computação cotidiana desde os modems, telefones, reprodutores de DVD, televisões e telefones a supercomputadores, servidores e data center. A junção dos elementos de um chip, onde há comunicação entre os mesmos e é definida uma rotina de funcionamento é denominado de Sistemas em Chip (*System-on-Chip* (SoC)). Tal sistema é composto de processadores, microcontroladores, buffers, dispositivos de entrada e saída de dados entre outras coisas (Greaves, 2011). Eles são amplamente implementados em ambiente de clusters em racks, contribuindo para o processamento em sistemas distribuídos, computação paralela e computação de alto desempenho. Nos clusters em racks, esses SoCs são acoplados de tal maneira que várias máquinas trabalham em conjunto para processar algo. No entanto, a demanda computacional aumenta exponencialmente através dos anos e soluções são criadas quase que concomitantemente. Nesse âmbito têm-se a Rede em chip (*Network-on-Chip* (NoC)), que possibilita esse complexo de dispositivos executarem tarefas em grupo, auxiliando na comunicação dos mesmos.

Nas soluções dentro dessa área, têm-se a ENoC, que pode expandir esses recursos de forma escalável com reconfiguração automática, porém tem como limitação o gargalo a taxa de transmissão de 25 Gb/s.

2 JUSTIFICATIVA

O ambiente de clusters em racks demanda um alto poder computacional e a ENoC é capaz de expandir esses recursos computacionais de forma escalável, provendo a possibilidade de expansão com a funcionalidade de reconfiguração, porém o gargalo ainda é o sistema de interconexão.

Ao considerar a computação de alto desempenho, a demanda da computação e poder de processamento é tão crescente quanto possível, dessa forma prover uma alternativa que mitigue o gargalo de interconexão vai ao encontro do objetivo da ENoC em atender essa necessidade. No entanto, atualmente a ENoC é limitada a uma taxa de transmissão de 25 Gb/s, porém com baixa sobrecarga de comunicação. Paralelamente, a 100GE intenta trabalhar com tecnologias para garantir a transmissão de dados em alta velocidade sem perda de dados (LAN/MAN Standards Committee, 2018), sendo implementada com uma taxa de transmissão de 100 Gb/s. Ao implementar a 100GE como sistema de intercomunicação nos clusters dentro de uma ENoC, esse estudo trará a essa arquitetura uma maneira mais célere de transmitir seus dados, o que pode colaborar para a redução do gargalo presente na interconexão.

3 HIPÓTESE

Esse trabalho parte do pressuposto que é possível utilizar o padrão Ethernet 100 Gigabit como sistema de interconexão numa ENoC voltada a ambientes de servidores de rack, em específico para computação de alto desempenho e clusters. No entanto, uma das principais características da ENoC é o protocolo de baixa sobrecarga, diferente da 100GE que possui um protocolo cuja sobrecarga é maior devido aos requisitos de sua arquitetura. Ao considerar a 100GE como objeto de interligação na ENoC, obtêm-se o ônus de maior taxa de transmissão de dados, porém com mais encapsulamentos de protocolos — pois o protocolo mínimo da ENoC é inferior ao exigido na 100GE, sendo então o motivo da baixa sobrecarga — e, justamente isso que essa pesquisa visa comparar, se ao usar um novo sistema de intercomunicação para a arquitetura ENoC, será solucionada a limitação da taxa de transmissão sem perda de desempenho, mesmo com a sobrecarga de pacote. O que se pressupõe é: mesmo com maior *overhead* (ou sobrecarga), essa implementação melhorará o funcionamento devido à maior taxa de transmissão.

4 OBJETIVOS

4.1 OBJETIVO GERAL

Fazer um estudo comparativo do desempenho e verificação de viabilidade da Ethernet 100 Gigabit como sistema de interconexão para a arquitetura de rede em chip expansível ENoC.

4.2 OBJETIVOS ESPECÍFICOS

- Estudar a Ethernet 100 Gigabit;
- Estudar sistemas de interconexão para comunicação Inter-Chip;
- Estudar soluções para computação de alto desempenho;
- Estudar conceitos de simulação para desempenho de arquiteturas de redes;
- Adaptar a ENoC para suportar a 100GE como sistema de interconexão;
- Simular o desempenho da arquitetura ENoC com a 100GE como sistema de interconexão;
- Simular o desempenho da arquitetura ENoC com a IB como base de comparação para sistema de interconexão;
- Simular o desempenho da arquitetura ENoC com outros sistemas de interconexão, além da 100GE, IB e a própria ENoC;
- Comparar os resultados das simulações de desempenho com demais sistemas de interconexão, incluindo a IB e a ENoC;
- Analisar os resultados obtidos e a viabilidade da utilização da 100GE como meio de comunicação para a ENoC;
- Publicar os resultados.

5 FUNDAMENTAÇÃO TEÓRICA

5.1 ETHERNET

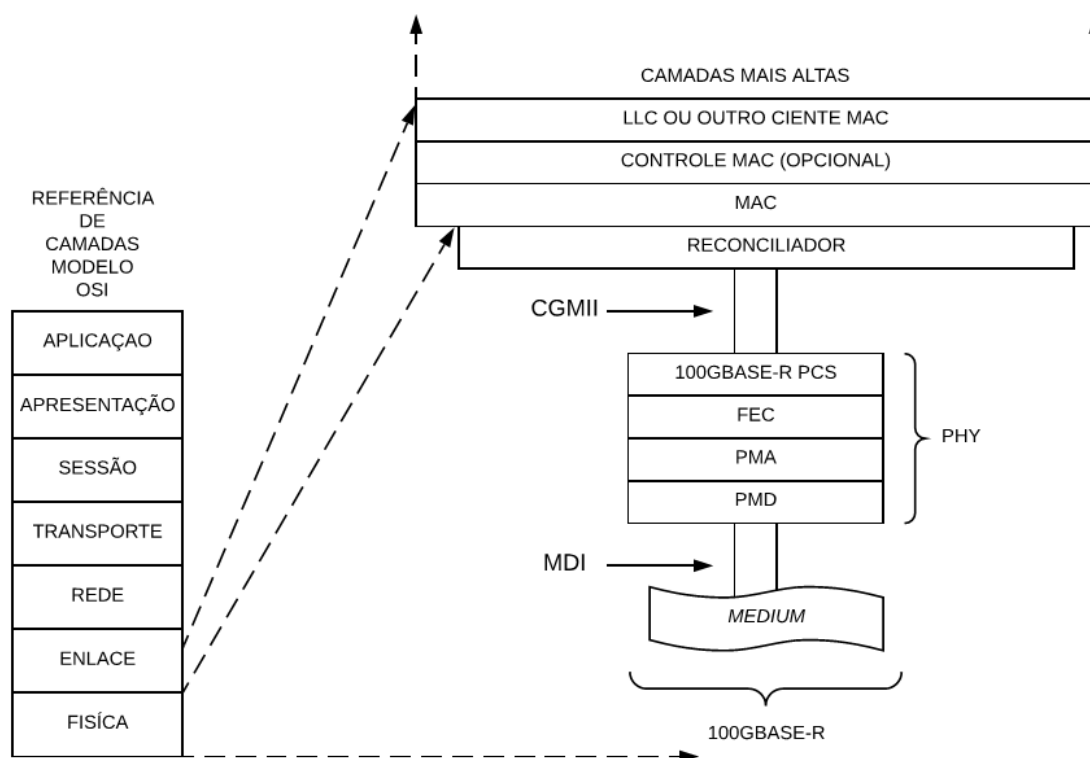
A Ethernet é um conjunto de normas e padrões de rede que define regras numa Rede de Internet Local (*Local Internet Network* (LAN)) para a transmissão de dados, implementando o algoritmo de Acesso Múltiplo com Detecção de Transporte e Controle de Colisão (*Carrier Sense Multiple Access with Collision Detection* (CSMA/CD)) para acesso a dados e o Controle de Acesso ao Meio (*Medium Access Control* (MAC)) para controle de acesso ao meio.

Esse protocolo é atualmente padronizado pelo IEEE 802.3, um grupo de estudo pertencente ao *Institute of Electrical and Electronics Engineers* (IEEE), cuja a responsabilidade é estudar e padronizar esse modelo de rede, tal qual atua na camada física e de enlace de dados no modelo *Open Systems Interconnection* (OSI). Os padrões são especificados por velocidade, ou seja, para cada velocidade há uma normatização. As funcionalidades dessa tecnologia tratada nesse estudo, estão especificadas em (LAN/MAN Standards Committee, 2018).

Dentro da camada física do Modelo OSI, a Ethernet define padrões de cabeamento, dispositivos (*switches* e *patch panels*), faixas de envio de dados e estruturas para que a velocidade desejada seja atingida. Já na camada de enlace, é usado um controlador de *link* lógico para destinar os dados de forma mais eficiente e também o MAC, que define *frames* de dados e garante que cada dispositivo conectado à rede tenha um endereço único, evitando o envio e processamento desnecessário de informações. Para interligar essas duas camadas foi desenvolvido o reconciliador e o *Media Independent Interface* (MII). Esses componentes da 100GE são ilustrados na Figura 5.1 e são especificados nos subcapítulos adiante.

Nesse âmbito, a 100 Gigabit, ou 100GE, é um conjunto de normas e tecnologias de rede para transmissão de dados numa velocidade de 100 Gb/s (LAN/MAN Standards Committee, 2018).

Figura 5.1 — Subcamadas Ethernet 100 Gigabit



Fonte: LAN/MAN Standards Committee, 2018, Seção 6, p. 83.

5.1.1 Camada Física

Nesse padrão, inicialmente são determinadas as especificações da camada física *Physical Layer Device* (PHY) para a transmissão desses dados, tal qual é dividida em subcamadas, são elas: *Physical Coding Sublayer* (PCS), *Forward Error Correction* (FEC), *Physical Medium Attachment* (PMA), *Physical Medium Dependent* (PMD) e o *Medium Dependent Interface* (MDI).

5.1.1.1 *Physical Coding Sublayer*

A primeira subcamada física PCS provê o serviço de codificação/decodificação dos dados em blocos de 66 bits (64b/66b), é responsável por distribuir os dados em diferentes faixas, compensação de diferença de taxas entre o reconciliador e o PMA, determinar quando uma conexão foi estabelecida informando então ao gerenciador quando o dispositivo está pronto para uso.

5.1.1.2 *Forward Error Correction*

Já na segunda subcamada física o FEC age com o objetivo de evitar a perda de dados através da redundância no envio de bits, adicionando bits ao *streaming* de dados pelo algoritmo Reed-Salomon, sendo então nomeado como Reed-Solomon *Forward Error Correction* (RS-FEC). Em cada especificação o RS-FEC trabalha de uma forma e, em sua implementação na 100GE, é necessário exatamente quatro faixas de envio e outras quatro para recebimento, sendo indispensável o mapeamento 10:4 quando trabalha com o PMA possuindo 10 faixas, pois tal PMA opera com 10 faixas para envio e outras 10 para recebimento.

5.1.1.3 *Physical Medium Attachment*

A terceira subcamada, o PMA, fornece o serviço de intermediação entre um PMA e um cliente, podendo esse cliente ser um PCS, FEC ou outro próprio PMA. Entre esses serviços têm-se a adaptação dos sinais das faixas dos PCS para o número de faixas físicas ou abstratas do cliente, ou seja, ele pode receber 10 faixas de *stream* de dados e transformá-la em 4 faixas de *stream* de dados. O PMA faz o direcionamento de bits de dados para que todos os bits de uma *stream* vão e voltem pela mesma faixa. Ainda na terceira camada, quando há a comunicação entre dois PMAs, pode-se usar especificação elétrica de módulos plugáveis com dez faixas a 10.3125 Gbd ou utiliza-se especificação elétrica de módulos plugáveis e pontos combinados com quatro faixas a 25.78125 Gbd.

5.1.1.4 *Physical Medium Dependent*

A quarta subcamada PMD provê o serviço de intermédio entre o PMA e o MDI controlando o envio e recebimento dos dados entre os mesmos, traduzindo o código recebido do PMA de *streamings* de bits para *streamings* elétricas ou *streamings* de bits para *streamings* de sinais óticos e o contrário também, onde o PMA trabalha com bits e o MDI com sinais elétricos e/ou óticos. Também na implementação do PMD é decidido qual modo de comunicação/conexão usar, exemplo: Fibra ótica em *Single-Mode*, *Multi-Mode* ou também cabos de cobre.

5.1.1.5 *Medium Dependent Interface*

Relacionado ao PMD, tem-se ainda o *Medium Dependent Interface* (MDI), que é a interface de comunicação entre o dispositivo PMD e o *Medium*, podendo o *Medium*

ser entendido como meio de comunicação (fibra ótica, cabo de cobre, *backplane*). Essa interface pode ser compreendida de outro modo como o receptor e/ou transmissor acoplado ao dispositivo PMD e varia conforme a normativa.

5.1.2 Camada de Enlace

Já na camada de enlace, tem-se também as divisões de especificações e como principais entidades há o *Logical Link Control* (LLC), o MAC e também o *MAC Control* com CSMA/CD, que na implementação da 100GE não é necessário.

Entre as entidades, inicialmente há o MAC, que provê o serviço de transferência de dados entre MACs, onde sua semântica de transferência é constituída de: endereço de destino (que pode ser um MAC ou um grupo), endereço de origem, unidade de serviço de dados MAC e sequência de checagem de *frame*. Tais semânticas trabalham através de *frames* e pacotes, sendo os *frames* encapsulados em pacotes pelo MAC e cada elemento é especificado conforme Tabela 5.1.

O primeiro elemento (preâmbulo), ajuda na sincronização da camada de sinalização (*Physical Signaling Sublayer* (PLS)) com o tempo do pacote e serve para avisar que um *frame* está a caminho. O SFD é a sequência de dados fixada (10101011) que antecede o *frame*, ou seja, depois dela o receptor saberá que será os bits do *frame*. Os campos de endereço possuem 48 bits cada e o endereço de destino pode ser um MAC único, um grupo ou todos os endereços da LAN. O campo de Tamanho indica o número de bytes dentro do próximo campo (Dados Cliente MAC).

Tabela 5.1 — Formato de *Frame* e Pacote Ethernet

		Quantidade de Bytes	Campo
Pacote	Frame	7	Preâmbulo
		1	SFD
		6	Endereço de Destino
		6	Endereço de Origem
		2	Tamanho
		46 a 1500	Dados Cliente MAC (<i>PayLoad</i>)
		4	Sequência de checagem de <i>frame</i>

Depois de encapsulado, o *frame* é enviado e na recepção é considerado inválido quando uma das condições a seguir é atendida:

- Seu tamanho não é condizente com o especificado no elemento de tamanho;
- Se o *frame* não possuir a quantidade de bits múltipla de 8, pois deve ser uma cadeia de bytes;
- O FCS calculado não coincidir com o valor FEC recebido.

O MAC *Control* com CSMA/CD não se faz necessário na 100GE pois essa funcionalidade com tal algoritmo não é útil nesse âmbito, visto que ela opera somente em modo *full duplex*, logo não há risco de colisão de dados.

Ainda na camada de enlace, porém acima do MAC, tem-se o LLC que facilita, através de mecanismos de multiplexação e demultiplexação, o trânsito e coexistência de vários pacotes num meio de rede com vários pontos. Isso é possível pois ele guarda o endereço de cada MAC dentro da rede e faz todos se enxergarem como um, ou seja, enquanto o MAC guarda a informação dos dados e dispositivos para mostrar a origem e destino do pacote, o LLC mostra o melhor caminho a ser percorrido para esse pacote chegar ao objetivo.

5.1.3 Reconciliador

Esses conceitos tecnológicos (PHY, MAC e LLC) se referem as duas primeiras camadas físicas do modelo OSI e para interligar as duas o 802.3 também padroniza o reconciliador (RS). Opcionalmente o 802.3 também padroniza as *Media Independent Interface* (MII), que provê a interconexão lógica entre o MAC e o PHY, atuando então embaixo do RS. O MII foi desenvolvido para que a camada de enlace de dados e o meio físico trabalhem de forma independente e é especificado na 100GE como CGMII.

Em suma, o RS converte a *stream* de dados fornecida pelo MAC para dados (sinais) paralelos do CGMII e também o mapeamento dos sinais providos pelo CGMII para as primitivas do MAC, já CGMII é o facilitador de transmissão e recebimento de dados entre o RS e o PHY.

5.1.4 Evolução

Todas essas definições são padronizadas pela IEEE para a 100GE e vários fatores foram essenciais para o alcance de tal velocidade, isso fica claro ao compará-lo com outros padrões como 10GE, 25GE e 400GE na Tabela 5.2, sendo eles conjuntos de normas para a velocidade, respectivamente, de 10 Gb/s, 25 Gb/s e 400 Gb/s, todos eles definidos pelo grupo 802.3.

Tabela 5.2 — Especificações de Normas 803.2

	10GE	25GE	100GE	400GE
Bloco de Dados no RS (Bits)	32	32	64	64
Faixas	1	1	10 ou 25	16
Velocidade por Faixa (Gb/s)	10	25	25 para 4 faixas	25
			10 para 10 faixas	

O primeiro dado se refere aos blocos de bits transmitidos através do RS, a qual se observa um aumento para o dobro do tamanho, 32 para 64 bits. A importância desse item é vista quando se calcula a velocidade de transmissão com 10 faixas transmitindo a 156,25 Mhz:

$$10(\text{faixas}) \times 64(\text{bits}) \times 156,25 = 100 \text{ (Gb/s)}$$

Na segunda têm-se a quantidade de faixas e a velocidade por faixa. Inicialmente, em 2010, a 100GE foi padronizada com 10 faixas operando a 10 Gb/s por segundo, logo após, em 2014, a 802.3 iniciou uma força tarefa para alcançar a velocidade de 25 Gb/s de transmissão numa única faixa, tal objetivo foi atingido em 2016 quando foi aprovado esse padrão. A partir desse ponto foi normatizado a 25GE com uma faixa 25 Gb/s, 100GE com 4 faixas a 25 Gb/s, 200GE com 8 faixas a 25 Gb/s e a 400GE com 16 faixas a 25 Gb/s.

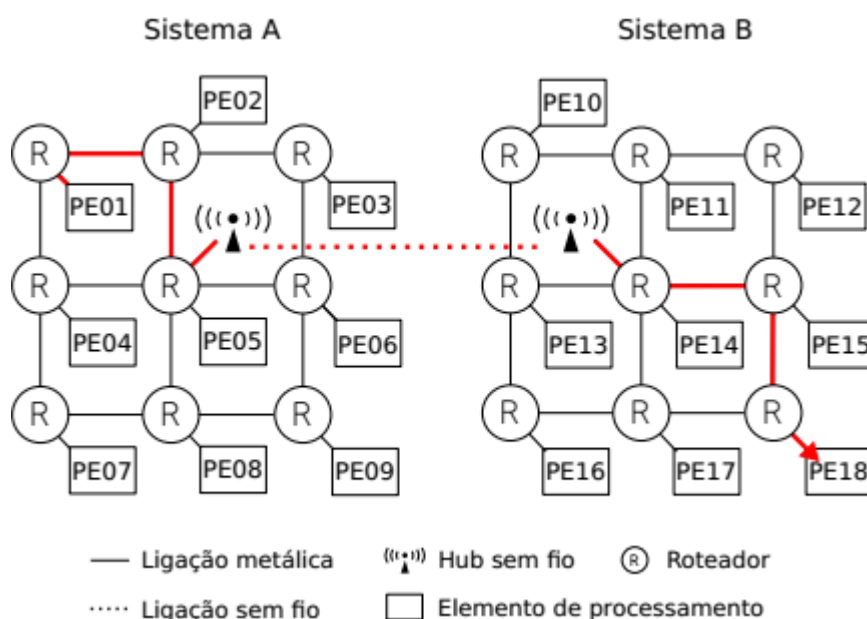
O conjunto de evolução de vários elementos como cabeamentos óticos (OM3 e OM4), cabos coaxiais, capacidade de processamento dos hardwares e aumento da demanda de dados a serem transmitidos foram responsáveis pelo avanço da Ethernet e foram elencados dois principais (Bloco de Dados no RS e Velocidade por faixa),

onde observa-se grande impacto dos mesmos no crescimento da Ethernet e mais estudos estão sendo feitos para que velocidades de 1,2 Tb/s e 800 Gb/s sejam alcançadas.

5.2 EXPANSIBLE NETWORK ON A CHIP (ENOC)

A Rede em Chip Expansível (*Expansible Network on a Chip (ENoC)*) é uma rede sugerida por (Pedroso Pires, Zanata Alvez, & Pessoa Albini, *Expansible Network-on-Chip Architecture*, 2018) para interação de Sistemas num Chip (SoCs), que permite comunicação de elementos de processamento de um chip, porém esse diálogo pode se dar tanto de elementos num chip (Intra-Chip) quanto com elementos em outro chip (Inter-Chip), conforme imagem .

Figura 5.2 — Arquitetura ENoC



Fonte: Pires, 2018, p. 31.

Na camada física, essa rede é composta por Elementos de Processamento (PE), Ligações metálicas, *buffers* e roteadores, todos eles dentro de um chip. Tais membros são dispostos numa malha bidimensional onde os PEs possuem *buffers* para armazenar suas mensagens e esses PEs são ligados a um roteador, ou seja, há um roteador para cada PE e os roteadores são também ligados a outros quatro

roteadores a sua volta. Um desses roteadores é ligado a um *hub* sem fio e o mesmo faz comunicação com outro *hub* sem fio em outro chip.

Na camada de enlace de dados, ela trabalha com roteadores, *hubs* e pacotes divididos em *flits* de 32 bits. Os pacotes são divididos e reconstruídos dentro dos PEs e enviados através dos roteadores, que possuem comunicação em barramento *full duplex*. Quando o destinatário for outro chip, o *flit* é encaminhado ao *hub* sem fio através dos roteadores, que envia o mesmo para o hub do chip de destino. O pacote é dividido em 4 bytes para endereço de destino e origem, 4 a 1500 bytes para os dados a serem transmitidos (*PayLoad*), por fim um *flit* repetindo o último *flit* do *PayLoad* para indicar o fim do pacote.

A ENoC é sugerida para ser expansível e reconfigurável, sendo que isso se dá através de sinais que o *hub* sem fio envia para informar sua presença e quando esse sinal é detectado, troca-se informações sobre seus PEs e essas informações são armazenadas dentro de cada *hub*, permitindo assim a expansividade sem necessidade de conhecimento prévio.

5.3 INFINIBAND (IB)

A InfiniBand (IB) é uma rede padronizada pela InfiniBand Trade Association destinada para computação de alto desempenho, provendo um fácil meio para transporte de mensagens direto de uma aplicação a outra aplicação, *storage* ou sistema operacional. Enquanto a Ethernet foca na transmissão de bits de dados numa rede, a IB visa criar um canal direto de comunicação, numa rede, entre elementos de uma aplicação sem necessidade de intervenção do sistema operacional (Grun & InfiniBand® Trade Association, 2010).

5.3.1 Camada Física

Na camada física, essa especificação é composta de Adaptador de Canal do Host (*Host Channel Adapter* (HCA)), Adaptador de Canal Alvo (*Target Channel Adapter* (TCA)), *Switches*, Roteadores, Cabos e Conectores.

O HCA fica num dispositivo ou computador e, fornece controle e conexão para transmissão de dados com outros dispositivos, podendo ser esse segundo um HCA, TCA ou *Switch*. Em outras palavras, o HCA é o dispositivo físico nas pontas para o

canal virtual criado entre dois pontos. O TCA promove as mesmas funcionalidades do HCA, porém de maneira mais simples, pois é feito para dispositivos com subsistemas especializados. Tal dispositivo foi, em suma, substituído pelo HCA pois este atende todas as demandas do TCA.

O *switch* é semelhante aos usados em outras redes, que é um dispositivo para multiplexação de pacotes, sendo diferenciado na maneira que é usado na implementação da camada de enlace. Já os roteadores são utilizados na segmentação de uma IB, ou seja, se há uma IB muito larga, ela pode ser dividida em sub-redes conectadas por roteadores IB.

Ainda na camada física, há os cabos e conectores, sendo os conectores o meio ao qual um sinal ótico é enviado na origem ou recebido no dispositivo de destino, ou seja, facilitam a passagem de bits elétricos para o tipo de sinal do meio. Os cabos são o meio ao qual a informação trafega, podendo ser ele uma fibra ótica, um *backplane* ou cabo de cobre. Todos eles trabalhando no sistema de velocidade IB.

A IB padroniza suas velocidades e faixas, trabalhando atualmente com 1, 4, 8 ou 12 faixas e velocidades de 2.5, 5, 10, 14.06 e 25.78 Gb/s. Tal combinação pode transmitir, por exemplo, 312 Gb/s se forem usados 12 faixas enviando 25.78 Gb/s (InfiniBand Trade Association, 2016).

5.3.2 Camada de Enlace

Na camada de enlace essa rede implementa a técnica *Flow Control*, que consiste numa coleção de providências tomadas para que um receptor não seja sobrecarregado por um dispositivo que envia numa velocidade maior. Ela é feita através de uma confirmação que o receptor envia informando ao controlador que mais pacotes podem ser recebidos.

Há vários formatos de pacotes na IB e serão descritos dois pela simplicidade de estrutura. O primeiro é o Pacote Local e é composto por 8 Bytes no cabeçalho, que contém informações como destino e origem, 0 a 4096 Bytes para *PayLoad* e 6 Bytes para cobrir os pacotes caso necessário. O outro tipo de pacote é o Global e sua diferença do pacote Local é o tamanho do cabeçalho, que é composto por 40 Bytes. O pacote Local é usado quando é preciso transportar uma informação na mesma subnet e o Global de uma subnet a outra (InfiniBand Trade Association, 2015).

Tal envio é realizado por nós entre dispositivos na rede, chamados de *Queue Pair* (QR). Os QRs são construídos em cima de canais virtuais traçados na rede, onde um tradutor de endereço - implementado na camada de transporte - traduz o endereço virtual para o caminho físico. Na camada de transporte, a IB oferece um tradutor de endereço virtuais para físicos e também implementações de mensagens/protocolos para que a comunicação com dispositivos e *storages* aconteça de forma direta e/ou facilitada.

6 METODOLOGIA

Consiste na revisão da literatura em busca de fundamentação para a Ethernet 100 Gigabit, também do estudo da ENoC para entender sua arquitetura e funcionamento. Posteriormente, adaptar-se-á a ENoC inserindo a 100GE como sistema de interconexão. Logo após, as etapas se darão conforme abaixo:

- Definição de métricas e parâmetros;
- Definição carga de trabalho;
- Definição de cenário de simulação;
- Definição de sistemas de interconexão como base de comparação;
- Realização de simulação da ENoC com a Ethernet 100GE e mais sistemas de interconexão;
- Realização de análise do desempenho a partir dos resultados obtidos;
- Conclusão da pesquisa e publicação dos resultados.

As simulações serão realizadas para verificar elementos como desempenho, sobrecarga de pacotes, custo de área e tráfego de comunicação pois a simulação é uma forma eficaz de se mensurar arquiteturas e organizações de interconexões (Pires, 2018).

Como simulador, foi escolhido o Noxim pois o mesmo é amplamente utilizado no meio de pesquisa, sendo ele um simulador de redes em chip desenvolvido na Universidade de Catania (Itália), escrito usando SystemC, é desenvolvido em linguagem C++ e está disponível em repositórios online (Pedroso Pires, Zanata Alves, & Pessoa Albin, Trace-driven extension for, 2017). Este simulador foi utilizado para avaliar a ENoC e com o objetivo de se aproximar do ambiente em que a mesma foi principiada, será usado tal simulador. Esse simulador foi usufruído também em: (Kurimoto, Fukutsuka, Taniguchi, & Tomiyama, 2013), (Catania, Mineo, Monteleone, Palesi, & Patti, 2016), (yen phing, Ehkan, Ahmad, Nazri Bin Mohd Warip, & Zulkefli, 2019). Para ser capaz de atender as características da ENoC, foram realizadas modificações no simulador e corrigidos processos nele existentes. Essas alterações foram realizadas em (Pedroso Pires, Zanata Alve, & Pessoa Albin, Trace-driven and processing time extensions for Noxim, 2019), logo as simulações se darão com o simulador apossado de tais modificações.

7 CRONOGRAMA

ATIVIDADES	MÊS					
	1º	2º	3º	4º	5º	6º
Revisão da literatura						
Estudo da ENoC						
Adaptar a ENoC inserindo a 100ge como sistema de interconexão						
Definição de métricas e parâmetros						
Definição carga de trabalho						
Definição de cenário de simulação						
Definição de sistemas de interconexão como base de comparação						
Realização de simulação da ENoC com a Ethernet 100GE e mais sistemas de interconexão						
Realização de análise do desempenho a partir dos resultados obtidos						
Conclusão da pesquisa e publicação dos resultados						

BIBLIOGRAFIA

- Catania, V., Mineo, A., Monteleone, S., Palesi, M., & Patti, D. (2016). Cycle-Accurate Network on Chip Simulation with Noxim. *ACM Transactions on Modeling and Computer Simulation (TOMACS)*, Article No. 4.
- Catania, V., Monteleone, S., Mineo, A., Palesi, M., & Patti, D. (27 de 05 de 2019). *Network on Chip Simulator*. Fonte: Noxim - the NoC Simulator: <https://github.com/davidepatti/noxim>
- Greaves, D. D. (2011). *System on Chip: Design and Modelling*. Easter Term: University of Cambridge.
- Grun, P., & InfiniBand® Trade Association. (2010). *Introduction to InfiniBand™ for End Users*.
- InfiniBand Trade Association. (2015). *InfiniBand™ Architecture Specification Volume 1 - GENERAL SPECIFICATIONS*.
- InfiniBand Trade Association. (2016). *InfiniBand Architecture Specification Volume 2 - PHYSICAL SPECIFICATIONS*.
- Kurimoto, Y., Fukutsuka, Y., Taniguchi, I., & Tomiyama, H. (2013). A hardware/software cosimulator for Network-on-Chip. *International SoC Design Conference (ISOCC)* (pp. 172–175). Busan, South Korea: IEEE.
- LAN/MAN Standards Committee. (2018). *IEEE Standard for Ethernet*. IEEE Standard 802.3.
- Pedroso Pires, I. L., Zanata Alve, M. A., & Pessoa Albini, L. C. (2019). Trace-driven and processing time extensions for Noxim. *Springer Nature*.
- Pedroso Pires, I. L., Zanata Alvez, M. A., & Pessoa Albini, L. C. (2018). Expansible Network-on-Chip Architecture. *Advances in Electrical and Computer Engineering*, 61-68.
- Pires, I. L. (2018). *ENoC: Rede-em-Chip Expansível*. Curitiba: PR.
- yen phing, n., Ehkan, P., Ahmad, R., Nazri Bin Mohd Warip, M., & Zulkefli, F. (2019). Performance Analysis of the Impact of Design Parameters to Network-on-Chip (NoC) Architecture. Em *Novel FPGA implementation of EPZS motion estimation in H.264 AVC* (pp. 237-246).