

David Grunheidt Vilela Ordine

**Comparação das tecnologias de comunicação  
entre *clusters* no processador MPPA-256  
através do CAP Bench**

Florianópolis

2019



David Grunheidt Vilela Ordine

**Comparação das tecnologias de comunicação entre  
*clusters* no processador MPPA-256 através do CAP  
Bench**

Proposta de monografia submetida ao Programa de Graduação em Ciência da Computação para a obtenção do Grau de Bacharel.

Universidade Federal de Santa Catarina  
Departamento de Informática e Estatística  
Ciência da Computação

Orientador: Márcio Bastos Castro  
Coorientador: Pedro Henrique Penna

Florianópolis

2019

# FOLHA DE APROVAÇÃO DE PROPOSTA DO TCC

<b>Acadêmico(s)</b>	David Grunheidt Vilela Ordine
<b>Título do Trabalho</b>	Comparação das tecnologias de comunicação entre clusters no processador MPPA-256 através do CAP Bench
<b>Curso</b>	Ciências da Computação/INE/UFSC
<b>Área de Concentração</b>	Ciências da Computação

## Instruções para preenchimento pelo **ORIENTADOR DO TRABALHO**:

- Para cada critério avaliado, assinale um X na coluna SIM apenas se considerado aprovado. Caso contrário, indique as alterações necessárias na coluna de Observação.

Critérios	Aprovado				Observação
	Sim	Parcial	Não	Não se aplica	
O trabalho é adequado para um TCC em CCO (relevância / abrangência)?	<b>X</b>				
O título é adequado?	<b>X</b>				
O Tema de pesquisa está claramente descrito?	<b>X</b>				
O problema/hipóteses de pesquisa do trabalho está claramente identificado?	<b>X</b>				
A relevância da pesquisa é justificada?	<b>X</b>				
Os objetivos descrevem completa e claramente o que se pretende alcançar neste trabalho?	<b>X</b>				
É definido o método a ser adotado no trabalho? O método condiz com os objetivos e é adequado para um TCC?	<b>X</b>				
Foi definido um cronograma coerente com o método definido (indicando todas as atividades) e com as datas das entregas (p.ex. Projeto I, II, Defesa)?	<b>X</b>				
Foram identificados custos relativos à execução deste trabalho (se houver)? Haverá financiamento para estes custos?				<b>X</b>	
Foram identificados todos os envolvidos neste trabalho?	<b>X</b>				
As formas de comunicação foram definidas?				<b>X</b>	
Riscos potenciais que podem causar desvios do plano foram identificados?				<b>X</b>	
Caso o TCC envolva a produção de um software ou outro tipo de produto e seja desenvolvido também como uma atividade realizada numa empresa ou laboratório, consta na proposta uma declaração (Anexo 3) de ciência e concordância com a entrega do código fonte e/ou documentação produzidos?				<b>X</b>	
Avaliação <input checked="" type="checkbox"/> Aprovado <input type="checkbox"/> Não Aprovado					
<b>Professor Responsável</b>					

# Resumo

Os processadores *manycore* de baixo consumo energético, tais quais, o *Kalray MPPA-256* (MPPA-256), vieram para solucionar o *trade-off* desvantajoso entre gasto energético e ganho em desempenho que ocorria em alguns supercomputadores. Graças a estes, diversos novos conceitos e arquiteturas foram apresentadas. Porém, devido a diversos fatores, torna-se desafiador implementar aplicações eficientes que tomam total proveito destes processadores. No MPPA-256, especificamente, a quantidade de memória limitada por *cluster* e o não compartilhamento de memória entre *clusters* são alguns destes fatores limitantes.

O objetivo deste trabalho é dar sequência ao projeto de iniciação científica denominado "Otimização do Benchmark CAP Bench para o processador *manycore* de Baixo Consumo Energético MPPA-256 ", o qual portou o CAP Bench com a biblioteca *MPPA Asynchronous Communication* (ASYNC). Além disso, mudanças serão feitas na lógica da implementação da versão antiga, a qual utilizava a biblioteca *Inter-Process Communication* (IPC), para que fique equivalente a versão nova, podendo assim comparar aspectos somente das *Application Programming Interfaces* (APIs) de cada versão.

Serão feitos experimentos no processador MPPA-256 utilizando as aplicações de cada versão do CAP Bench, comparando aspectos de comunicação das APIs de cada versão, tais quais, tempos de execução de cada *cluster*, tempos de execução total da aplicação, quantidade de bytes trocados por cada *cluster* e gasto energético total da aplicação.

Como resultado, espera-se poder apontar, com exatidão, qual das duas tecnologias de comunicação usar em determinado contexto, no MPPA-256, a partir dos resultados colhidos com as comparações. Assim, aspira-se proporcionar dados concretos para que, futuramente, programadores que utilizem o MPPA-256 saibam quais das duas tecnologias utilizar dependendo do contexto em que sua aplicação esteja inserida.

**Palavras-chaves:** *manycores*. MPPA-256. *comunicação*. *async*.



# Sumário

	<b>Sumário . . . . .</b>	<b>5</b>
<b>1</b>	<b>INTRODUÇÃO . . . . .</b>	<b>7</b>
<b>2</b>	<b>OBJETIVOS . . . . .</b>	<b>9</b>
<b>3</b>	<b>MÉTODO DE PESQUISA . . . . .</b>	<b>11</b>
<b>4</b>	<b>CRONOGRAMA . . . . .</b>	<b>13</b>
	<b>REFERÊNCIAS . . . . .</b>	<b>15</b>





# 1 Introdução

Na década atual, foi crescente a expectativa de que os supercomputadores pudessem alcançar a computação em *exaescale*. Grande parte destes supercomputadores utilizam processadores *multi-core* do estado da arte em seus *clusters*. Porém, a comunidade científica observou uma barreira no caminho ao tentarem seguir sempre a mesma técnica para o ganho de desempenho. O aumento no número de núcleos de processamento por *chip* foi a técnica seguida na última década, o que, em determinado momento, não trouxe mais resultados significativos quanto ao *trade-off* entre consumo de energia e ganho em desempenho.

O Departamento de Defesa do Governo dos Estados Unidos (DARPA), uma das organizações mais importantes do país, também foi um dos órgãos responsáveis por alertar a comunidade científica acerca desta barreira em um de seus relatórios (KOGGE et al., 2008). Neste relatório fica claro que os supercomputadores atuais estão longe de executar a uma potência aceitável para atingir o *exaescale*, já que, enquanto o supercomputador mais eficiente do mundo utiliza cerca de 15.1 GFlops/W, o indicado como aceitável pela DARPA é de 50 GFlops/W.

Para solucionar este problema, pesquisadores realizaram diversos estudos a fim de apresentar novos tipos de arquiteturas voltadas a eficiência energética, buscando bom balanceamento entre consumo energético e desempenho, ou seja, seguindo os conceitos de *Green Computing*. Logo, o grande interesse da comunidade científica de *High-Performance Computing* (HPC) acerca deste tema foi um dos responsáveis por alavancar a produção de novos tipos de processadores, como os processadores *manycore* de baixa potência, por exemplo, o MPPA-256 (DINECHIN et al., 2013), o Adapteva Epiphany (OLOFSSON et al., 2014) e o SW26010, utilizado no supercomputador *Sunway TaihuLight* (FU et al., 2016).

Com a finalidade de avaliar o desempenho e consumo energético do MPPA-256, Souza et. al propuseram um benchmark denominado CAP Bench, o qual utiliza uma API de comunicação síncrona entre processos denominada IPC (DINECHIN et al., 2013). Esta API proporciona perda de energia ao utilizar o hardware para sincronização, além de possuir baixo nível de abstração, requerendo conhecimento prévio da arquitetura alvo para implementações paralelas eficientes. Ao realizar a otimização do *benchmark*, David et. al o otimizaram através do porte com a API ASYNC, a qual possui maior nível de abstração e maior potencial em desempenho. Além disso, alterações na lógica de implementação foram realizadas.

Portanto, faz-se necessário atualizar a versão antiga do *benchmark* quanto a lógica de implementação, para que esta se equivalha a nova implementação, criando assim um ambiente propício para comparação de aspectos puramente das tecnologias de comunicação

citadas, utilizando as duas versões do *benchmark* para isso. Após isso, necessita-se comparar ambas as implementações a fim de determinar qual tecnologia de comunicação se comporta de maneira mais eficiente e robusta no MPPA-256 em determinados contextos, para que tenhamos dados de execução de cada API numa mesma versão de placa do MPPA-256, tendo assim provas concretas na hora de escolher uma das duas APIs.

## 2 Objetivos

O objetivo deste trabalho é obter dados concretos acerca da execução de aplicações de diversos domínios de problemas no MPPA-256, utilizando as duas APIs já citadas e o CAP Bench, podendo assim, comparar as execuções de cada aplicação em cada cenário específico possível dentro do processador, obtendo identificadores precisos que, em momentos futuros, possam apontar qual das duas APIs utilizar, dependendo do domínio de problema de uma certa aplicação.

Assim, podem ser definidos como objetivos específicos desse trabalho:

1. Investigar a viabilidade do uso do MPPA-256 para a computação de alto desempenho;
2. Estudar aspectos das APIs de comunicação existentes no MPPA-256, como a ASYNC e IPC;
3. Implementar, em ambas as versões do CAP Bench, novas métricas na obtenção dos dados de execução das aplicações, relacionados a memória *cache* de cada *chip*;
4. Avaliar os custos e benefícios do MPPA-256 em relação ao desempenho e gasto energético, assim como sua utilidade para a Computação Sustentável (*Green Computing*);
5. Comparar as tecnologias ASYNC e IPC de comunicação entre processos do MPPA-256, a fim de prover métricas precisas para escolha de uma das duas na hora de implementar uma nova aplicação para o processador.

### **Restrições:**

- Acesso somente virtual ao processador MPPA-256.
- Documentação limitada de ambas as tecnologias de comunicação citadas.

### **Premissas:**

- Disponibilidade da versão antiga do CAP Bench;
- O orientador estará disponível para reuniões periódicas;
- Computador disponível;
- Acesso remoto ao processador MPPA-256;
- Processador disponível para execução dos experimentos;

- Disponibilidade de água, luz e energia;
- Acesso à Internet.

**Marcos:**

- Entrega do resumo em TCC I: 2ª semana de Junho/2020;
- Primeira entrega da monografia em TCC II: 1ª semana de Novembro/2020;
- Defesa da monografia: 2ª semana de Novembro/2020;
- Segunda entrega da monografia em TCC II (versão final): 1ª semana de Dezembro/2020.

**CrITÉrios de aceite:**

- Aprovação da banca avaliadora;
- Aprovação do orientador;
- Aprovação do professor responsável;
- Conformidade com as normas definidas pela instituição;
- Prazos cumpridos;

### 3 Método de Pesquisa

O presente projeto será desenvolvido em sequência aos trabalhos de pesquisa do orientando, orientador e coorientador deste trabalho, David Grunheidt Vilela Ordine, Márcio Bastos Castro e Pedro Henrique Penna, acerca do MPPA-256. Todas as contribuições desenvolvidas serão implementadas no repositório aberto do CAP Bench, disponibilizado em <<https://github.com/cart-pucminas/CAPBenchmarks>>.

O desenvolvimento do trabalho será feito nas linguagens *C* e *Python*, sendo subdividido em seis etapas: (i) Estudo acerca dos processadores *manycore* de baixa potência do estado da arte e as tecnologias de comunicação entre processos implementadas nestes; (ii) Melhoria da versão antiga para equiparação a nova versão do CAP Bench; (iii) Estudo da nova API disponibilizada pela *Kalray* para coleta de métricas relacionadas a memória *cache*; (iv) Implementação de novas métricas quanto a execução das aplicações em ambas as versões; (v) Implementação dos scripts de coleta de todas as métricas já implementadas; (vi) Coleta dos dados de execução de cada aplicação.

A segunda e terceira etapa do trabalho consistirão em intensa consulta as documentações providas pela *Kalray* acerca das APIs ASYNC e IPC. Além disso, uma nova API para coletar dados acerca da memória *cache* de cada *core* também será utilizada, sendo assim, serão consultadas, constantemente, 3 documentações. Já a primeira etapa prevê extensa leitura de artigos que refletem o estado da arte acerca do tema tratado, através de diversas plataformas como *IEEE* e *ACM*. Simplificando, todas as etapas consistirão na continuação do projeto de pesquisa de iniciação científica do orientando, a qual portou o CAP Bench com a API ASYNC.

A plataforma que contém o MPPA-256 está localizada em Grenoble (França) e o acesso remoto é proporcionado por uma parceria entre o *Distributed Systems Research Laboratory* (LaPeSD) e o *Laboratoire d'Informatique de Grenoble* (LIG). O orientando terá acesso ao laboratório LaPeSD, localizado na Universidade Federal de Santa Catarina (UFSC), porém, o desenvolvimento do trabalho será feito em seu computador pessoal, o qual terá acesso remoto a plataforma onde localiza-se o MPPA-256.



## 4 Cronograma

As atividades previstas neste projeto estão descritas abaixo:

1. **A1: Estudo de fundamentação teórica.** Revisão bibliográfica dos materiais que serão utilizados na pesquisa;
2. **A2: Estudo e alteração da versão antiga do CAP Bench.** Estudo da versão antiga do CAP Bench, assim como a API que esta porta, a IPC, para que se consiga executar e alterar esta versão de modo que equivalha-se a nova, em termos de lógica do domínio do problema.
3. **A3: Estudo e implementação das novas métricas em ambas versões do CAP Bench.** Estudo da nova API para coleta de dados acerca da memória *cache* dos *cores* através da inclusão de funções que as constituem em ambas as versões, observando os resultados intermediários e adaptando as implementações conforme o ganho de conhecimento.
4. **A4 : Escrita do relatório do TCC I.** Elaboração do relatório de TCC I contendo todos os desenvolvimentos realizados até o momento. Entrega prevista para a segunda semana do mês de Junho;
5. **A5: Implementação dos *scripts* para os experimentos.** Estudos acerca dos melhores métodos para coleta dos dados disponíveis em todas as aplicações, assim como das bibliotecas em *Python* que facilitarão a coleta e transformação dos dados.
6. **A6: Coleta das métricas implementadas.** Coleta e medição de todos os dados acerca das execuções das aplicações, após a implementação dos *scripts* em A6, e avaliação dos resultados obtidos;
7. **A7: Escrita do rascunho de TCC II.** Elaboração do rascunho de TCC II para ser defendido na banca após a coleta dos dados em A7. Entrega prevista para a primeira semana de Novembro;
8. **A8: Preparação da defesa pública.** Preparação da apresentação oral do trabalho, bem como do material de apoio a ser utilizado durante a apresentação;
9. **A9: Defesa pública.** Realização da defesa do trabalho desenvolvido. Pretende-se realizar a defesa na segunda semana de Novembro.
10. **A10: Correções e entrega da versão final do TCC.** Correção e ajustes na monografia com base na avaliação da banca para que se possa fazer a entrega final do documento. Entrega prevista para a primeira semana de Dezembro.

A figura 1 apresenta a distribuição esperada ao longo do tempo das atividades descritas anteriormente, que serão realizadas durante o primeiro e o segundo semestres de 2020.

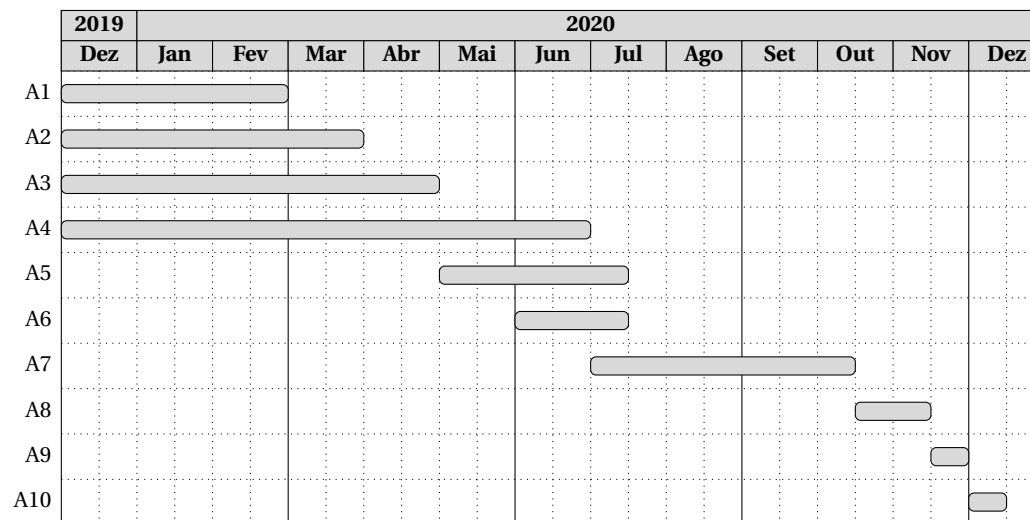


Figura 1 – Cronograma de atividades.



# Referências

DINECHIN de et al. A Distributed Run-Time Environment for the Kalray MPPA-256 Integrated Manycore Processor. In: *International Conference on Computational Science (ICCS)*. Barcelona, Spain: Elsevier, 2013. v. 18, p. 1654–1663.

FU, H. et al. The sunway taihulight supercomputer: System and applications. *SCIENCE CHINA Information Sciences*, v. 59, n. 7, p. 1–16, 2016.

KOGGE, P. et al. Exascale computing study: Technology challenges in achieving exascale systems. *Defense Advanced Research Projects Agency Information Processing Techniques Office (DARPA IPTO), Technical Representative*, v. 15, 01 2008.

OLOFSSON et al. Kickstarting high-performance energy-efficient manycore architectures with epiphany. In: *Asilomar Conf. on Signals, Systems and Computers*. [S.l.]: IEEE, 2014. p. 1719–1726. ISBN 978-1-4799-8297-4.