
Análise Integrada de Desempenho e Consumo de Energia em Sistemas de Armazenamento de Dados Distribuídos

Aluno: Juccelino Barros
Orientador: Gustavo Callou
Coorientador: Glauco Gonçalves
Recife 2017



Agenda

- Sistemas de Armazenamento de Dados Distribuídos
- Objetivo Principal
- Experimento
- Análise dos Resultados
- Conclusão e Trabalhos Futuros

Sistemas de Armazenamento de Dados Distribuídos

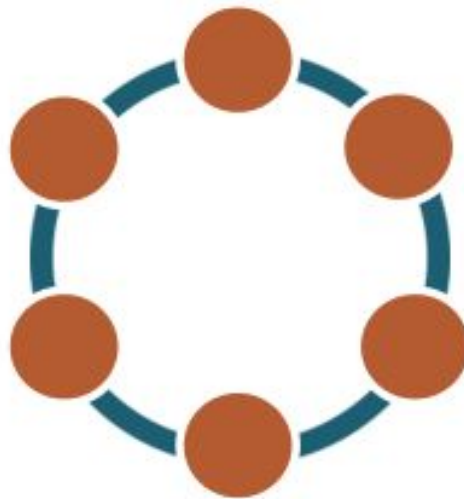
- Atender a grande demanda de dados
- Principais Características
 - **Boa escalabilidade**
 - **Elasticidade**
 - **Alta disponibilidade**
- Bancos de dados tradicionais não foram projetados para este tipo de ambiente
- Surgimento novos Bancos de Dados
- Apesar de apresentar diferentes modelos e arquitetura, eles possuem essas principais características

Objetivo Principal

- O principal fator para a escolha do sistema de armazenamento é o desempenho em operações típicas. Contudo, o consumo de energia desses ambientes também deve ser levado em consideração.
- Quais são as melhores formas de utilização desses sistemas levando em consideração o desempenho e o consumo de energia?
- O **objetivo principal** deste trabalho é fornecer insumos para ajudar os projetistas de ambientes de armazenamento de grande quantidade de dados na escolha mais apropriada levando em consideração o desempenho e o consumo de energia.

Cassandra

- Banco de Dados NoSQL orientado a coluna
- *Dynamo + BigTable*
- Armazenar grandes quantidades de dados em vários servidores
- Distribuição do tipo anel
- Utilizado em grandes empresas
 - Facebook
 - Netflix



Cluster do Cassandra

Ferramentas Utilizadas

Desempenho

- Yahoo! Cloud Serving Benchmark (YCSB)
- Gera carga de trabalho
- Calcula as principais métricas
 - Tempo de Execução
 - Vazão Média
 - Latência Média
- Pode definir quantidade de usuários simultâneos

<https://github.com/brianfrankcooper/YCSB>

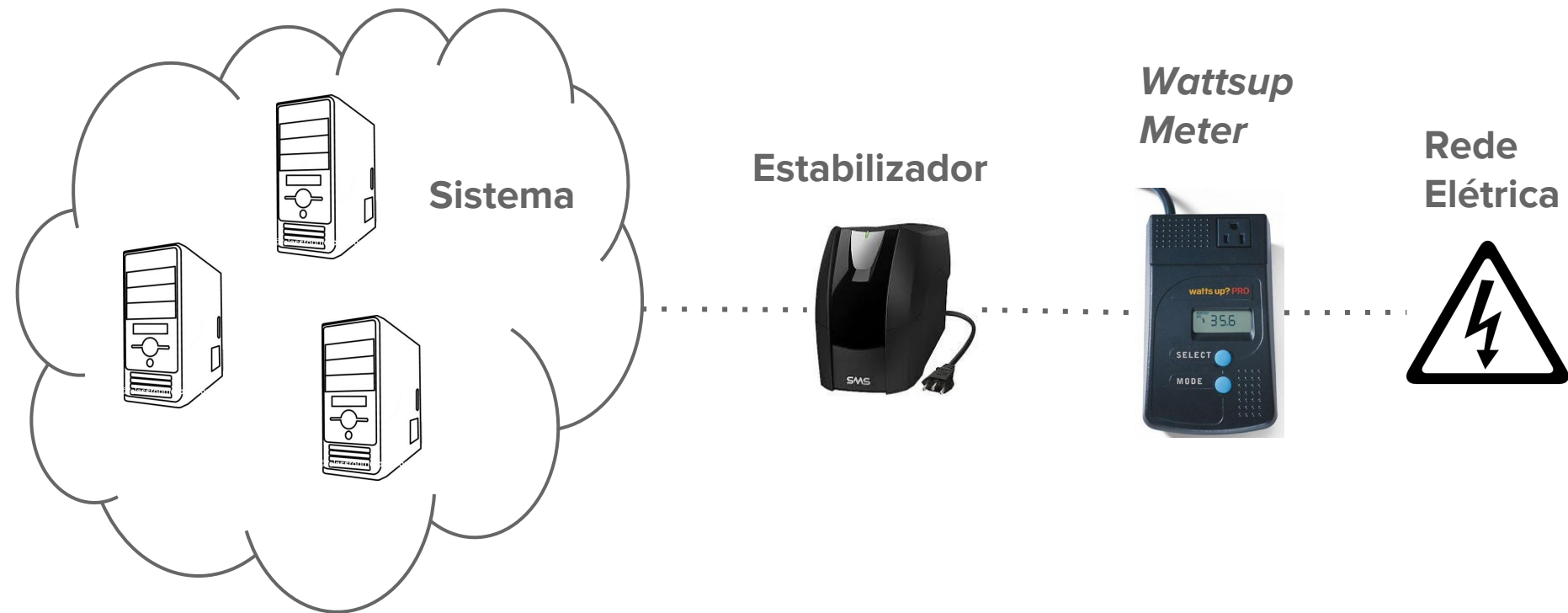
Consumo de Energia

- WattsUp Meter
- Localiza-se entre a rede elétrica e o sistema analisado
- Captura as principais métricas de consumo de energia

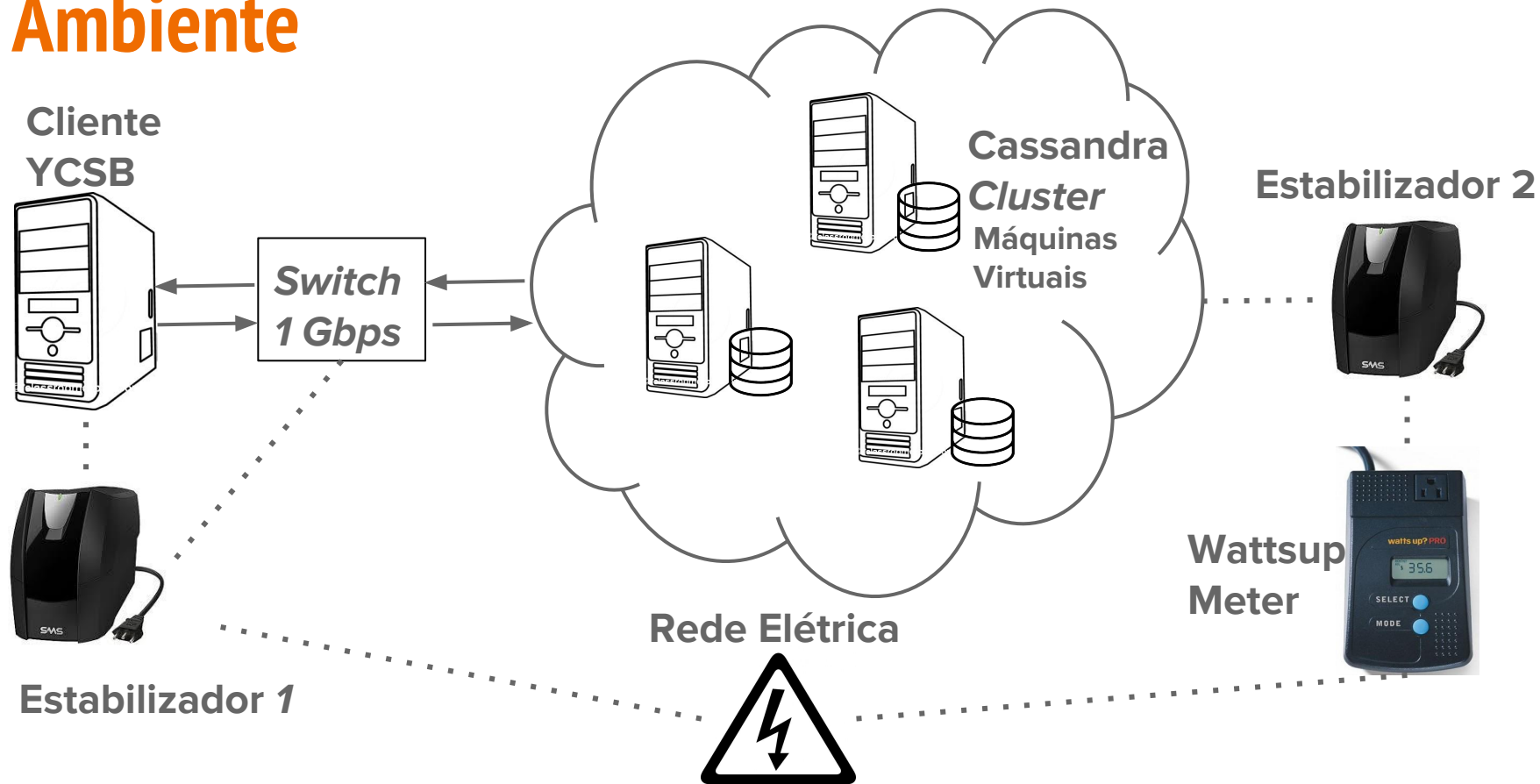


<https://www.wattsupmeters.com/secure/index.php>

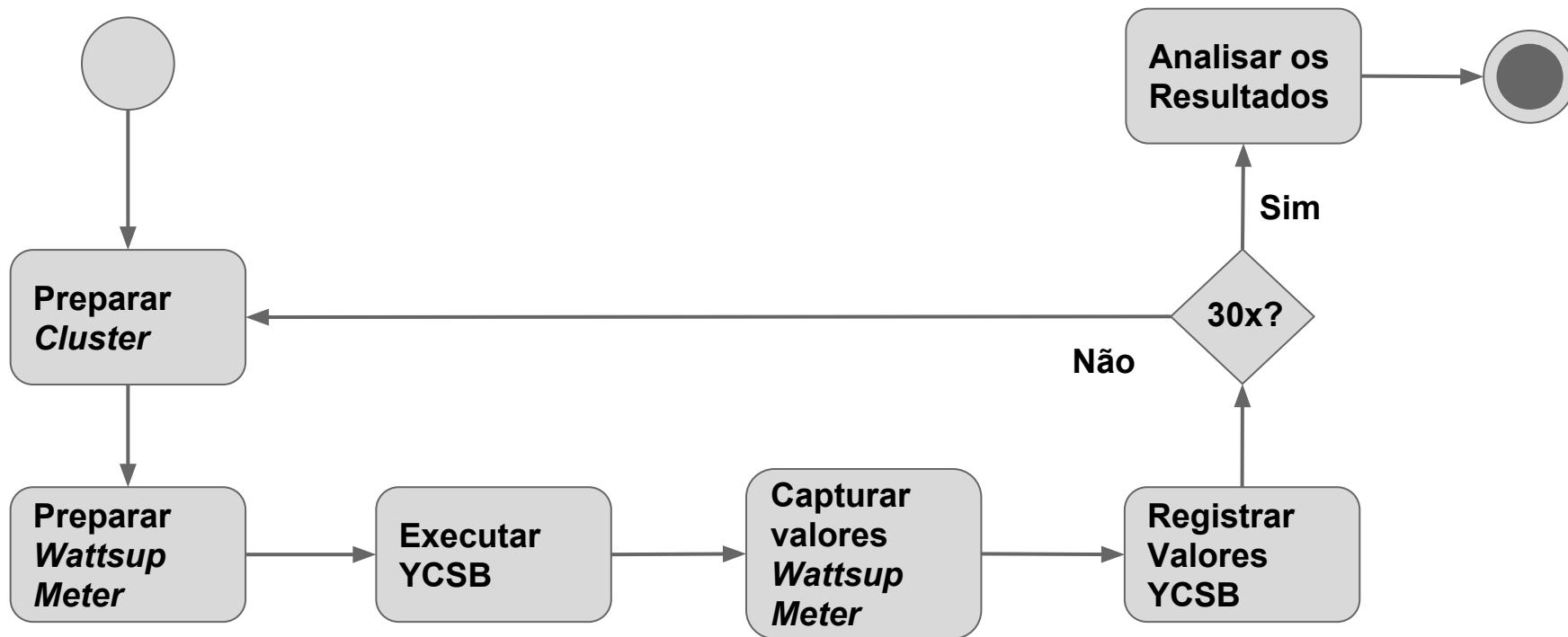
Wattsup Meter



Ambiente



Processo de Obtenção das Amostras



Características do Experimento

- Analisar o desempenho e o consumo de energia no processo de inserção de dados.
- Realizar operações de inserção no Cassandra
 - a. Os registros eram do tipo texto e tinham 1 KB cada
 - b. Inseridos em uma tabela com 10 colunas
- 2 Fatores
 - a. Quantidade de Nós
 - 1, 2, 4 e 6
 - b. Quantidade de Operações de Inserção
 - 10.000 (10K); 100.000 (100K); 1.000.000 (1M)
- 12 combinações
 - a. 30 + 3 amostras/combinção
 - b. 3 primeiras descartadas

Características do Experimento

- Métricas
 - a. **Desempenho**
 - Tempo de Execução (sec)
 - Vazão Média (op/sec)
 - Latência Média (ms)
 - b. **Consumo de Energia**
 - Joules (J)

Configuração do Ambiente

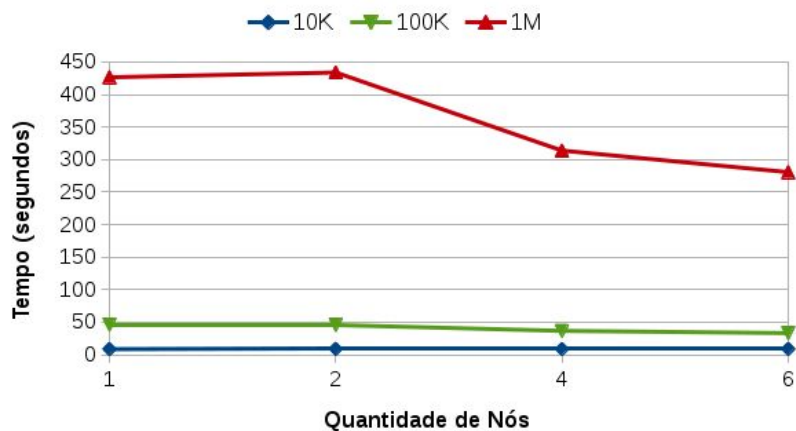
- Nó
 - Virtual Machine (VM)
 - 2GB RAM; 30G
- Cliente (YCSB)
 - 10 usuários
 - Máquina Física
 - 8GB RAM; Core i5
- Wattsup Meter
 - 1 segundo
 - Watts e Watts hour

- **Distribuição do *Cluster***

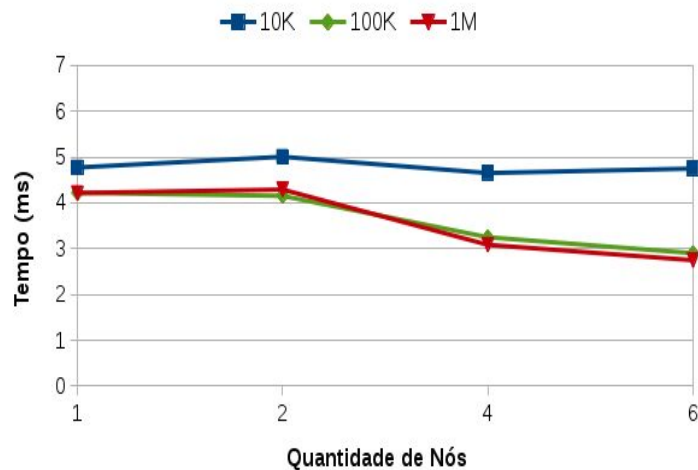
Qntd de Nós (VMs)	Qntd Máquinas	Qntd Nó/Máquina
1	1	1/1
2	1	2/1
4	2	2/1
6	3	2/1

Análise de Desempenho

- Tempo de Execução

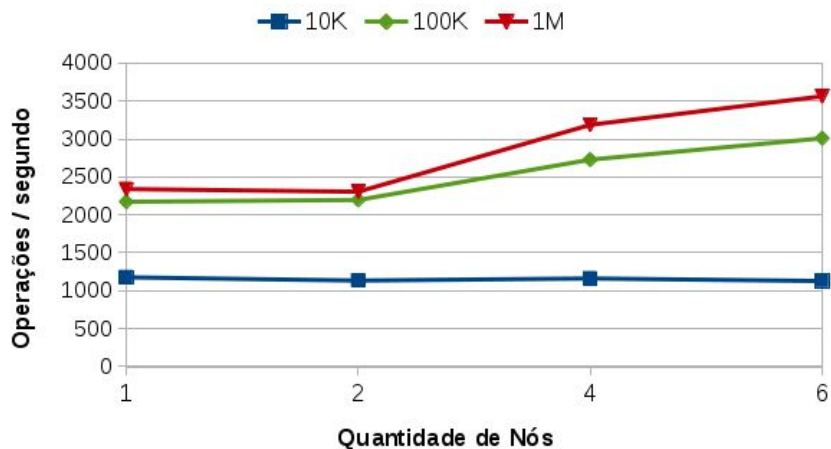


- Latência

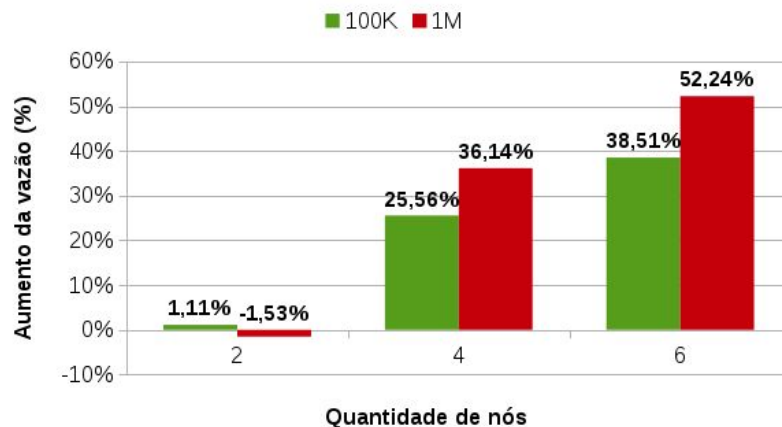


Análise de Desempenho

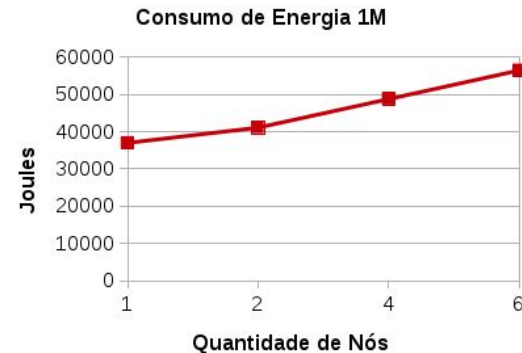
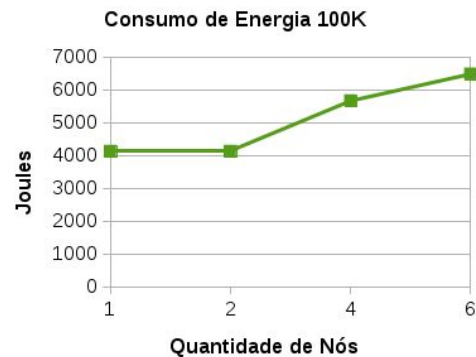
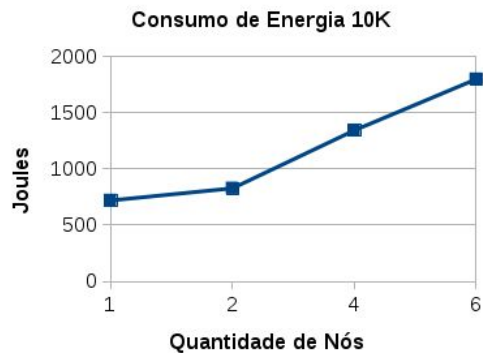
- Vazão



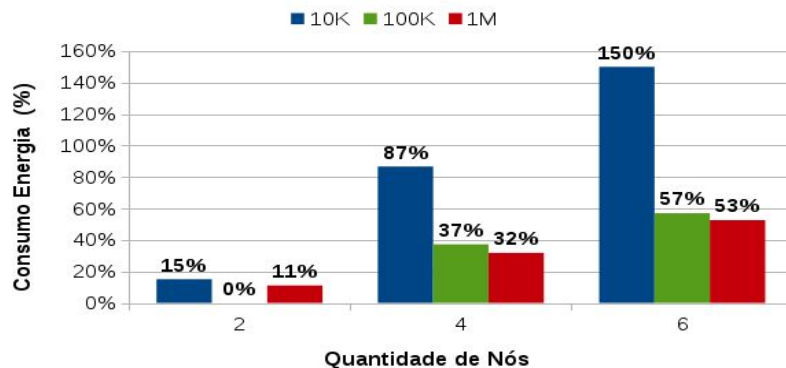
- Aumento da vazão em relação à 1 nó



Análise do Consumo de Energia



- Aumento em relação à 1 nó



Análise Integrada

Métricas	Qtd Operações	4 nós	6 nós
Tempo de Execução	100K	-20,42%	-27,88%
	1M	-26,42%	-34,20%
Vazão	100K	25,56%	38,51%
	1M	36,14%	52,24%
Latência	100K	-22,95%	-31,24%
	1M	-26,80%	-34,81%
Consumo de Energia	10K	86,67%	150,00%
	100K	37,21%	56,98%
	1M	31,77%	52,56%

Conclusão e Trabalhos Futuros

- Em algumas situações não há necessidade de aumentar os recursos computacionais, uma vez que o desempenho não apresenta melhorias significativas enquanto que o consumo de energia aumenta consideravelmente.
- Como trabalhos futuros, pretende-se desenvolver modelos formais de desempenho e consumo de energia desses ambientes.