SBSI2018
XIV SIMPÓSIO BRASILEIRO DE SISTEMAS DE INFORMAÇÃO
4 A 8 DE JUNHO - CAMPUS-SEDE - CAXIAS DO SUL/RS
SBC Sociedade Brasileira de Computação

DORIC: an Architecture for Data-Intensive Real-Time Applications

Miguel Cadaviz; Kleinner Farias; Lucian Gonçales; Vinicius Bischoff

Programa de Pós Graduação em Computação Aplicada (PPGCA)

Universidade do Vale do Rio dos Sinos (UNISINOS)

Visão Geral

- Introdução
- DORIC
- Avaliação
- Considerações Finais
- Trabalhos Futuros

Introdução

- Redes conectam uma gama diversificada de dispositivos;
- Diferentes tipos de aplicações;
  - Variados tipos de dados produzidos;
    - Estruturados e não estruturados;
- DIRT
  - *Data-Intensive and Real-Time applications*;
    - Gerenciamento de grande quantidade de dados;
    - Processamento em tempo real;
- Algumas características essenciais foram identificadas para esses tipos de aplicação;
- 5 características precisam ser satisfeitas para que uma arquitetura suporte uma aplicações DIRT.

Características

Data Sources	Publisher and Subscriber	Information Processing	Analysis and Storage	Consumers
Data Centers	Apache Kafka	Apache Spark	Apache Cassandra	Health
Web Applications	Amazon Kinesis	Apache Storm	Apache Hbase	TI Operations
Air Pictures	Google Pub/Sub	Apache Stamza	Apache Druid	Retail
Machines	MapR Streams	Twitter Heron	Elasticsearch	Government
Buildings	Apache DistributedLog	Amazon Kinesis	Apache Solr	Agriculture
Vehicles		Google Dataflow	MapR-DB	Marketing
Wearables	Azure Event	Apache Apex	FiloDB	Manufacture
Smartphones		Apache Flink	memsql	Industry
Tablets		Apache Flume	Apache Kudu	Science
Internet of Things		Azure Strem Analytics	Google Cloud Storage	Financial Services
		Apache Kafka Streams	Google Cloud Bigtable	Media Agencies

Análise Comparativa

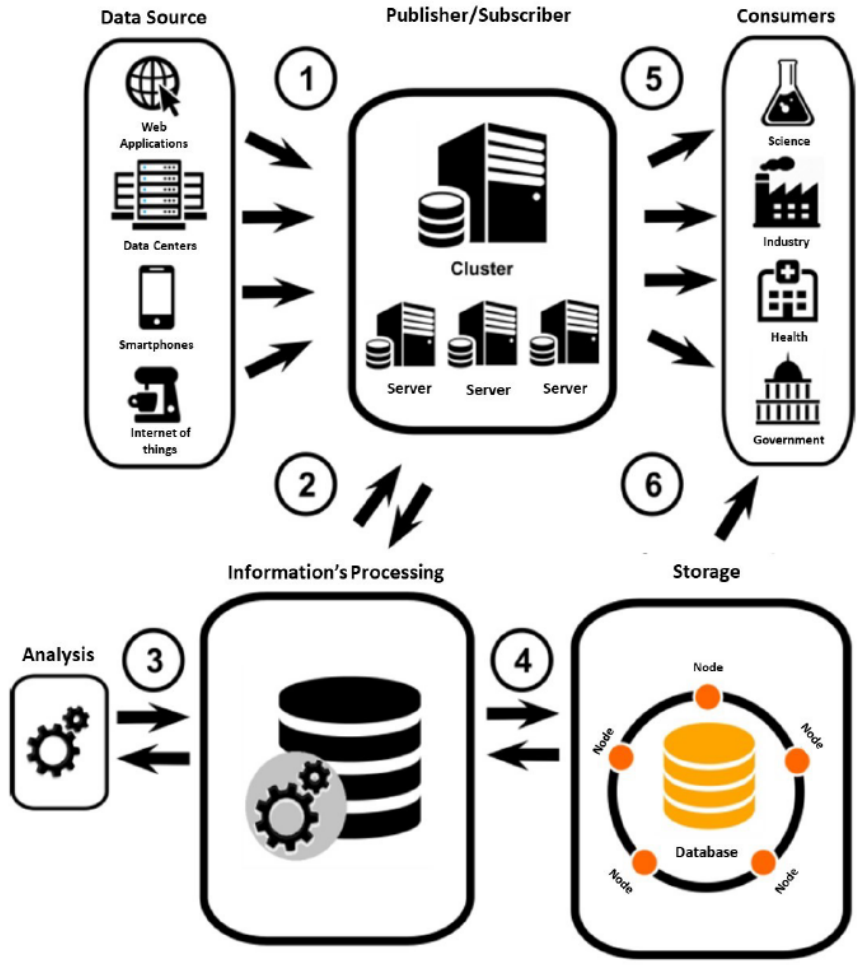
	Genérica	Data-Itensive	Large-scale	Real-Time	Extensível
Architecture for Scientific Applications	-	-	+	+	-
Architecture for Weather Stations	-	+	-	+	~
Architecture for managing Industrial Processes	-	+	+	-	-
DORIC	+	+	+	+	+

+ Suporta ~ Suporta parcialmente - Não suporta

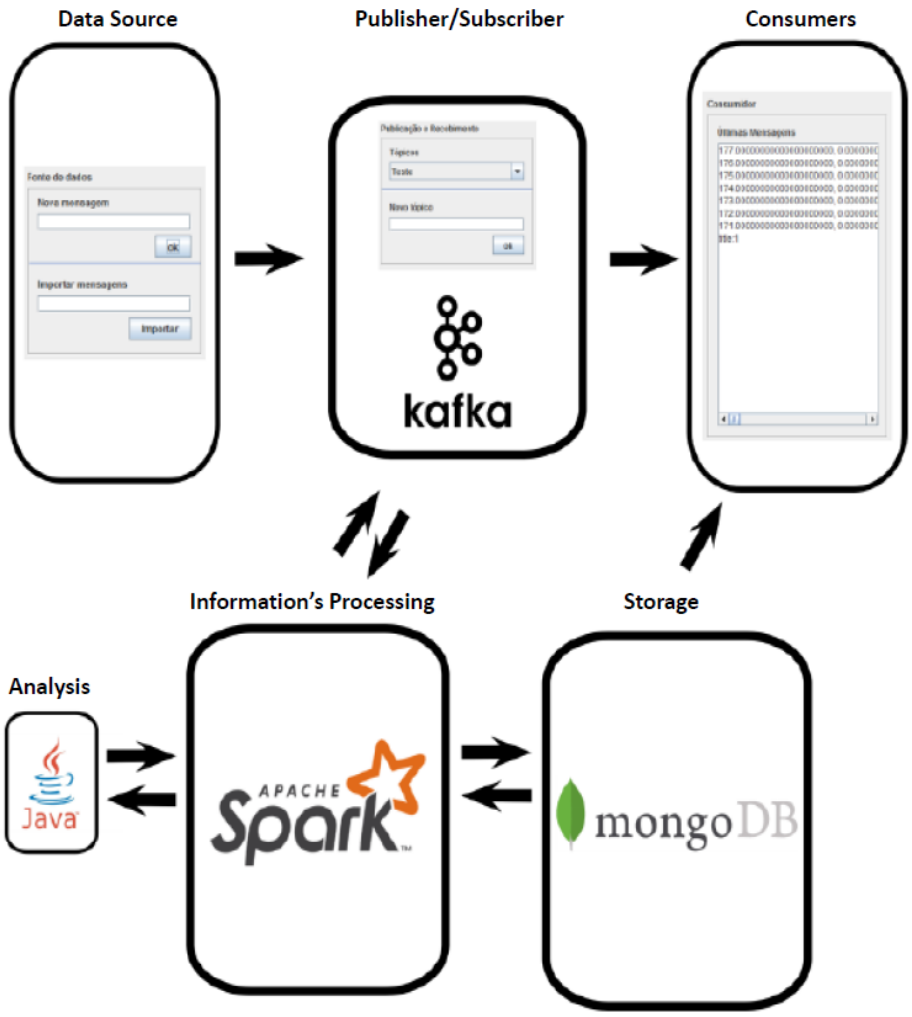
Requisitos

- A arquitetura deve:
  - **NFR01:** suportar o processamento e recepção de informações de larga escala com integridade;
  - **NFR02:** forncer respostas em tempo real.
  - **NFR03:** ser extensível.
- A combinação desses requisitos são considerados na arquitetura proposta.

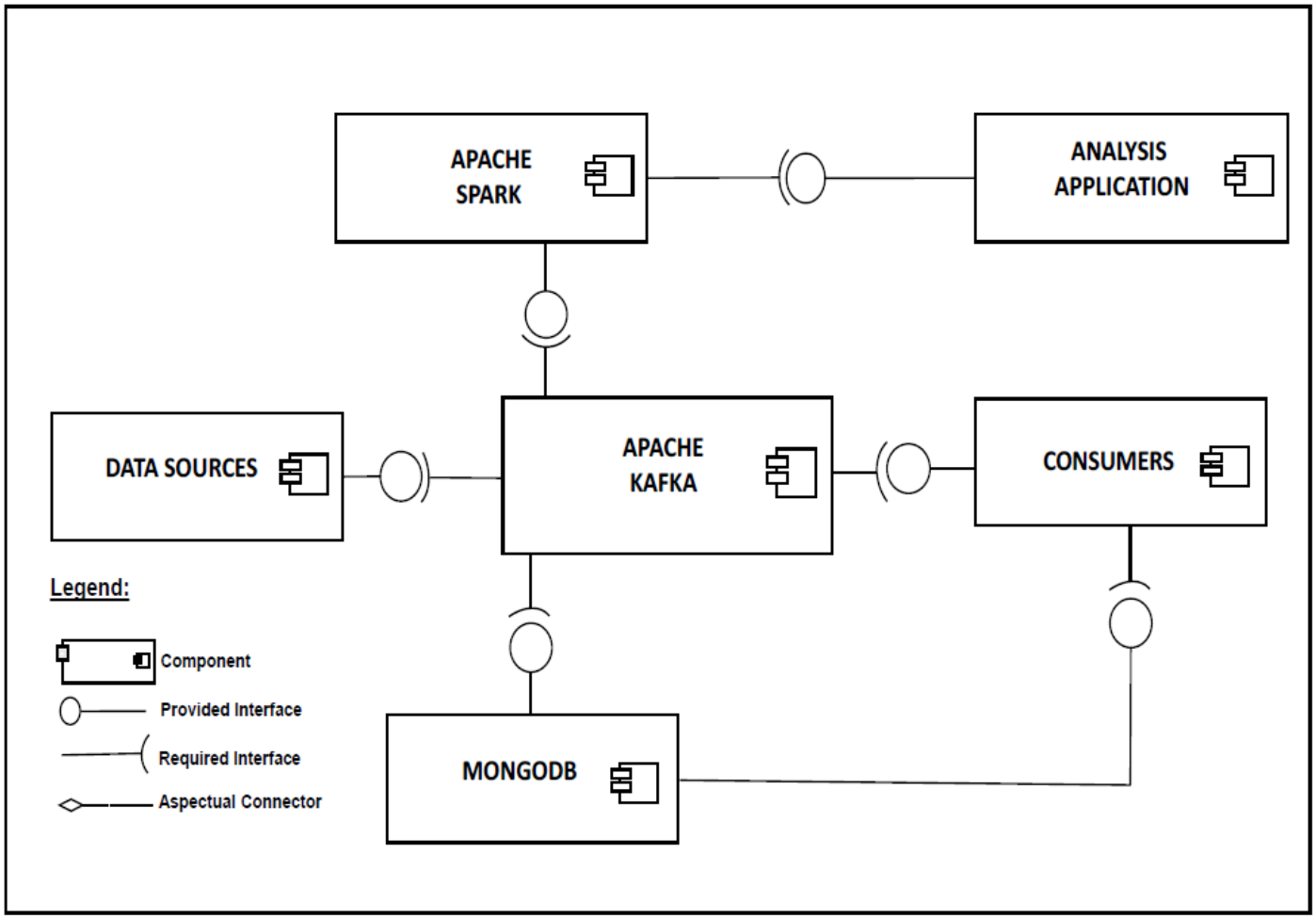
DORIC


Data Source
Publisher/Subscriber
Consumers
Web Applications
Data Centers
Smartphones
Internet of things
1
Cluster
Server
Server
Server
5
Science
Industry
Health
Government
2
6
Information's Processing
Storage
Analysis
3
4
Node
Node
Node
Node
Node
Database

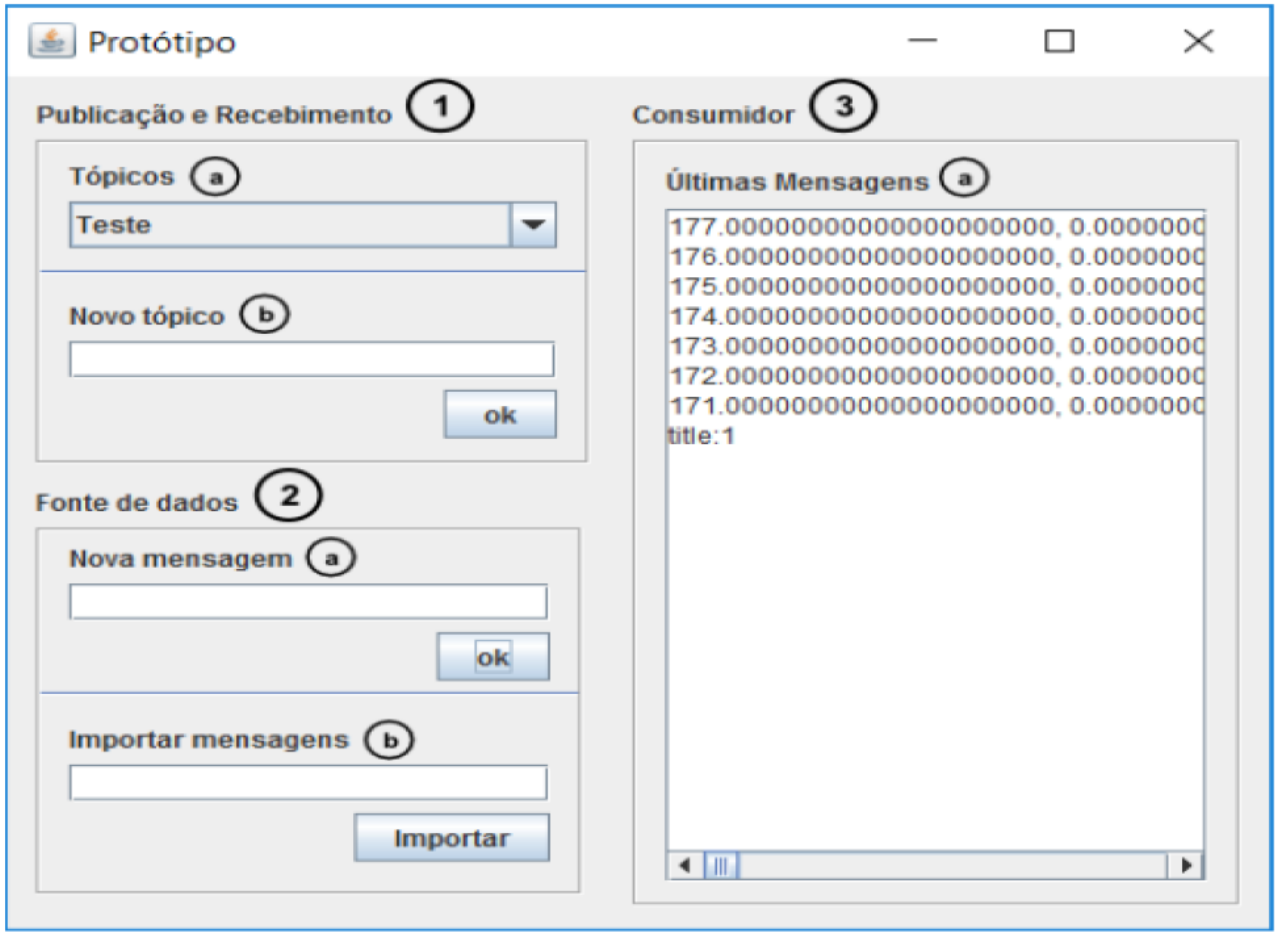
DORIC - Implementação



DORIC - Implementação



DORIC - Implementação

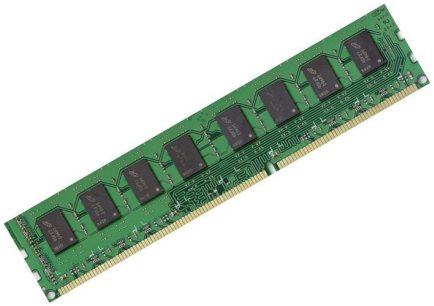


Método de Avaliação

- Verificar o desempenho do protótipo da aplicação;
- Demonstrar que a arquitetura suporta gerenciar informações em larga escala e tempo real;
- Protótipo foi executado em uma máquina virtual:

Windows 10

Core i7
2.7 Ghz



4GB DDR4
2133 Mgz

30GB
Hard Disk

Método de Avaliação

- Coleta indicados mentais de um EEG Portável

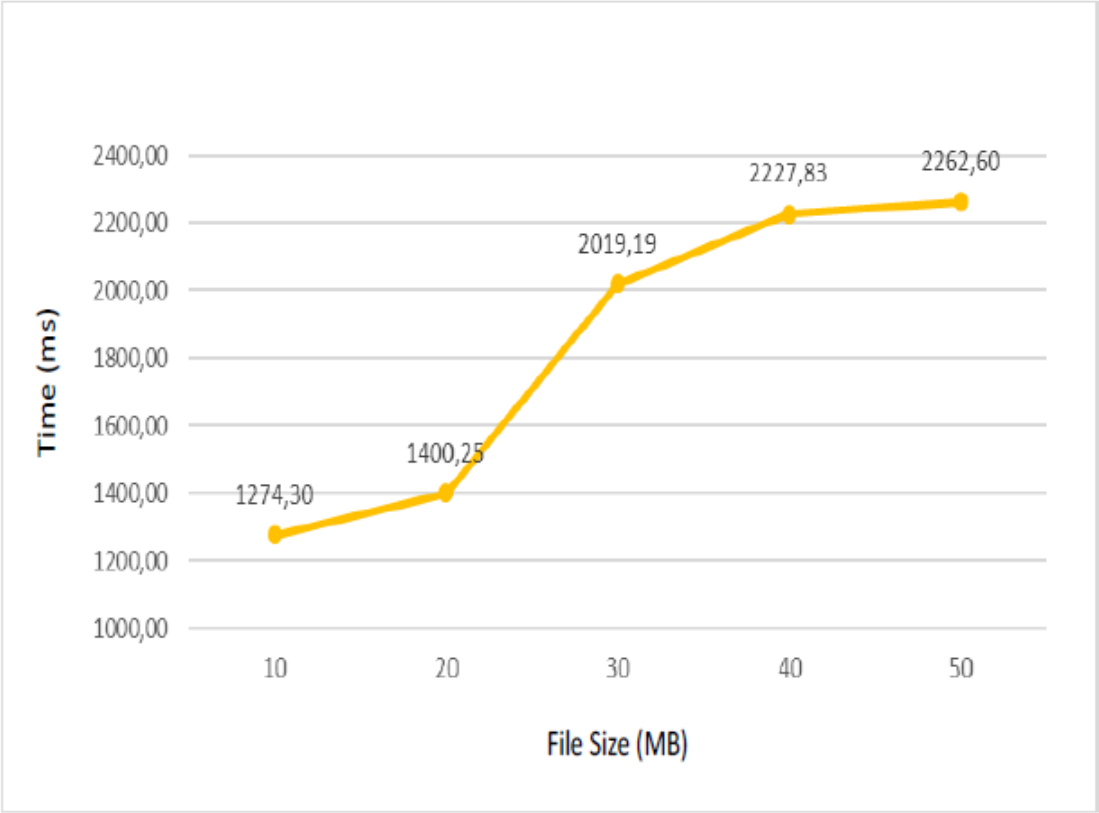
Emotiv
EPOC

AF3	F7	F3	FC5	T7	P7	O1	O2	P8	T8	FC6	F4	F8	AF4
4208.205078125000000000	4178.461425781250000000	4184.615234375000000000	4214.358886718750000000	4170.256347656250000000
4198.461425781250000000	4207.692382812500000000	4216.922851562500000000	4210.769042968750000000	4173.333007812500000000
4207.692382812500000000	4180.512695312500000000	4180.512695312500000000	4215.384277343750000000	4169.743652343750000000
4200.000000000000000000	4208.717773437500000000	4218.461425781250000000	4206.666503906250000000	4170.256347656250000000
4208.205078125000000000	4184.102539062500000000	4178.974121093750000000	4216.410156250000000000	4172.820312500000000000
4198.974121093750000000	4209.230468750000000000	4218.461425781250000000	4203.076660156250000000	4168.205078125000000000
4209.230468750000000000	4187.179199218750000000	4180.000000000000000000	4216.922851562500000000	4178.461425781250000000
4195.384277343750000000	4210.769042968750000000	4217.436035156250000000	4199.487304687500000000	4167.692382812500000000
4208.717773437500000000	4188.717773437500000000	4181.538574218750000000	4217.948730468750000000	4184.615234375000000000
4191.794921875000000000	4213.333007812500000000	4214.871582031250000000	4196.410156250000000000	4168.717773437500000000
4207.692382812500000000	4190.769042968750000000	4182.563964843750000000	4220.000000000000000000	4188.205078125000000000
4190.769042968750000000	4216.922851562500000000	4211.794921875000000000	4194.871582031250000000	4170.256347656250000000
4208.205078125000000000	4193.333007812500000000	4183.076660156250000000	4222.051269531250000000	4188.205078125000000000
4194.358886718750000000	4218.461425781250000000	4209.743652343750000000	4195.384277343750000000	4170.769042968750000000
4210.769042968750000000	4196.922851562500000000	4184.615234375000000000	4224.102539062500000000	4187.179199218750000000
4200.000000000000000000	4217.948730468750000000	4209.743652343750000000	4197.948730468750000000	4171.281738281250000000
4212.820312500000000000	4201.025390625000000000	4187.179199218750000000	4226.666503906250000000	4186.153808593750000000
4203.589843750000000000	4216.410156250000000000	4212.307617187500000000	4199.487304687500000000	4171.281738281250000000
4213.846191406250000000	4203.076660156250000000	4189.743652343750000000	4229.743652343750000000	4186.666503906250000000
4203.076660156250000000	4215.897460937500000000	4215.384277343750000000	4197.948730468750000000	4172.307617187500000000
4212.307617187500000000	4204.102539062500000000	4191.281738281250000000	4232.820312500000000000	4187.692382812500000000
4198.974121093750000000	4217.948730468750000000	4216.922851562500000000	4193.333007812500000000	4172.307617187500000000
4210.769042968750000000	4205.127929687500000000	4191.281738281250000000	4235.384277343750000000	4187.692382812500000000
4194.358886718750000000	4221.025390625000000000	4216.410156250000000000	4187.692382812500000000	4171.794921875000000000
4209.743652343750000000	4208.717773437500000000	4189.743652343750000000	4235.897460937500000000	4187.179199218750000000
4192.307617187500000000	4223.589843750000000000	4214.871582031250000000	4184.615234375000000000	4170.256347656250000000
4209.743652343750000000	4213.846191406250000000	4188.717773437500000000	4235.384277343750000000	4187.692382812500000000
4192.820312500000000000	4223.589843750000000000	4214.358886718750000000	4185.127929687500000000	4169.230468750000000000
4210.769042968750000000	4217.948730468750000000	4189.743652343750000000	4233.846191406250000000	4189.743652343750000000
4194.358886718750000000	4222.563964843750000000	4215.897460937500000000	4188.205078125000000000	4169.743652343750000000
4210.256347656250000000	4217.436035156250000000	4190.769042968750000000	4232.820312500000000000	4193.846191406250000000
4195.384277343750000000	4222.051269531250000000	4217.948730468750000000	4190.256347656250000000	4170.256347656250000000
4207.692382812500000000	4212.307617187500000000	4190.256347656250000000	4232.307617187500000000	4197.436035156250000000
4193.333007812500000000	4221.025390625000000000	4217.436035156250000000	4190.769042968750000000	4169.230468750000000000
4203.076660156250000000	4204.102539062500000000	4187.692382812500000000	4231.281738281250000000	4198.974121093750000000
4190.256347656250000000	4218.974121093750000000	4212.307617187500000000	4187.692382812500000000	4164.615234375000000000
4197.436035156250000000	4197.948730468750000000	4183.589843750000000000	4228.717773437500000000	4197.436035156250000000
4187.179199218750000000	4214.871582031250000000	4204.615234375000000000	4182.051269531250000000	4157.436035156250000000
4193.846191406250000000	4195.384277343750000000	4181.025390625000000000	4225.127929687500000000	4193.333007812500000000
4186.666503906250000000	4210.256347656250000000	4196.922851562500000000	4176.410156250000000000	4151.794921875000000000
4192.307617187500000000	4196.922851562500000000	4182.051269531250000000	4221.538574218750000000	4190.256347656250000000
4188.205078125000000000	4207.179199218750000000	4192.820312500000000000	4173.333007812500000000	4151.281738281250000000

Arquivo CVS

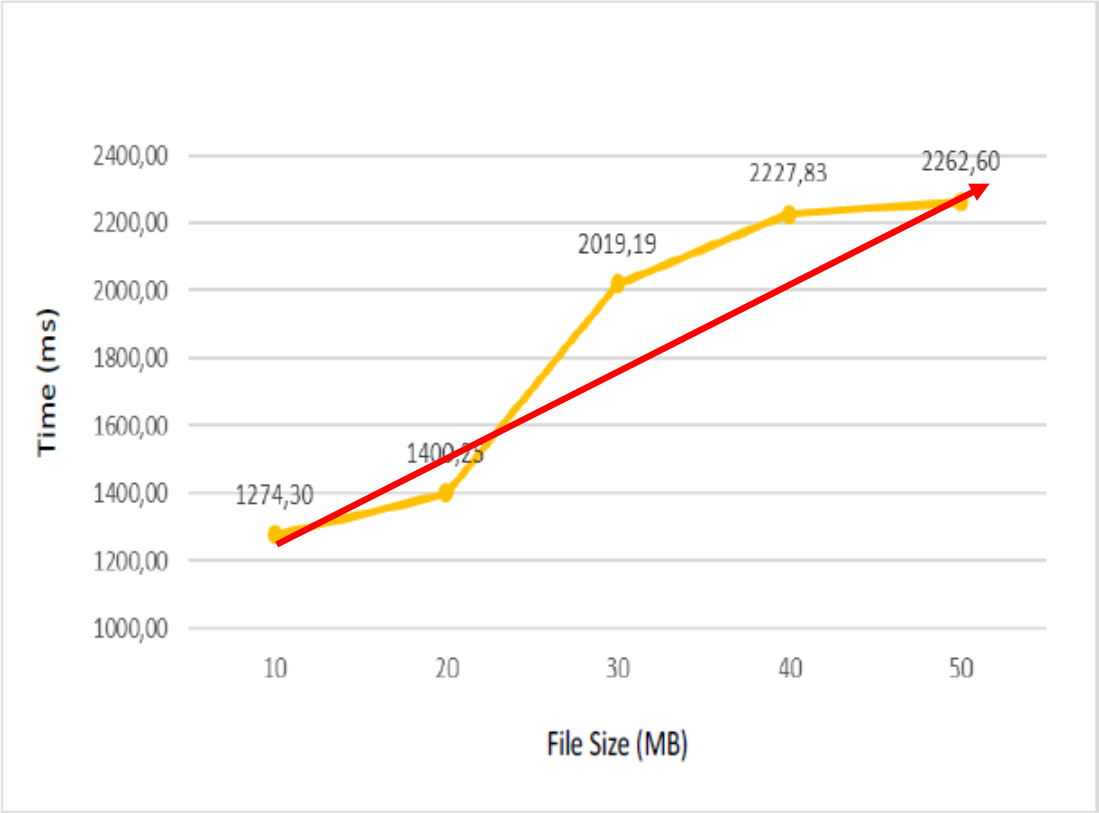
Resultados

- Data-intensive



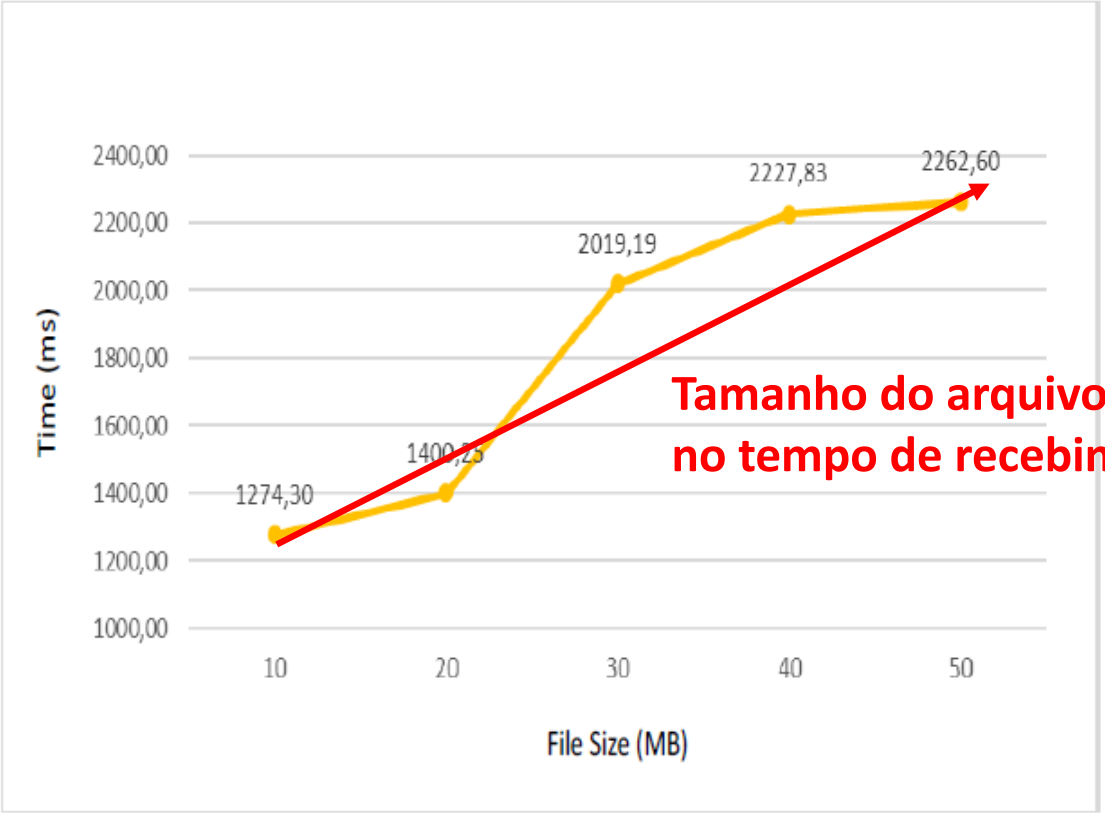
Resultados - Experimento 1

- Data-intensive



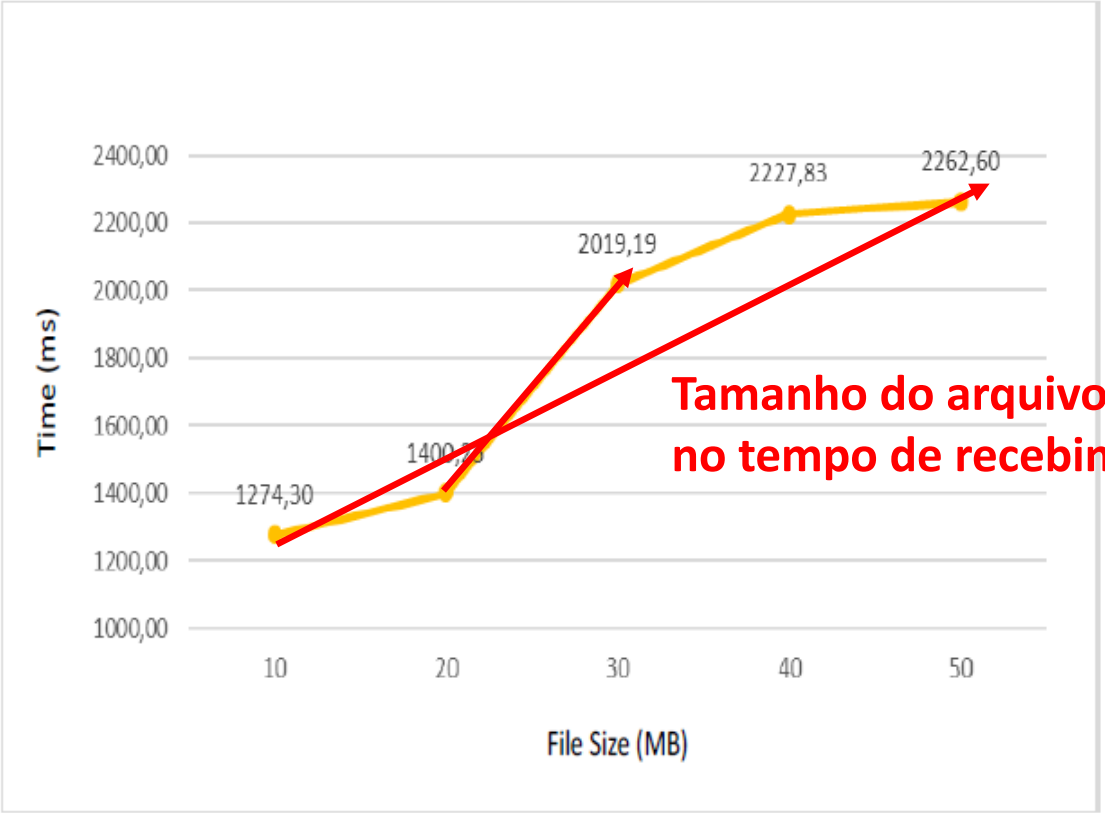
Resultados - Experimento 1

- Data-intensive



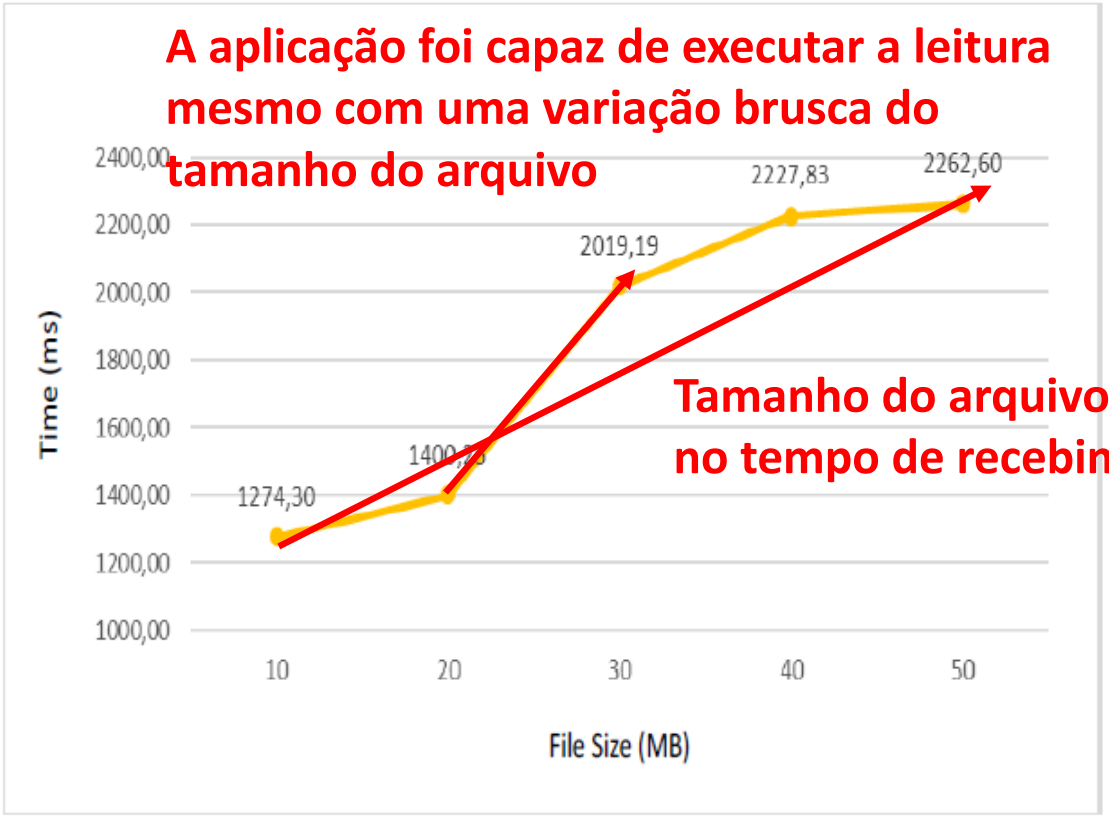
Resultados - Experimento 1

- Data-intensive



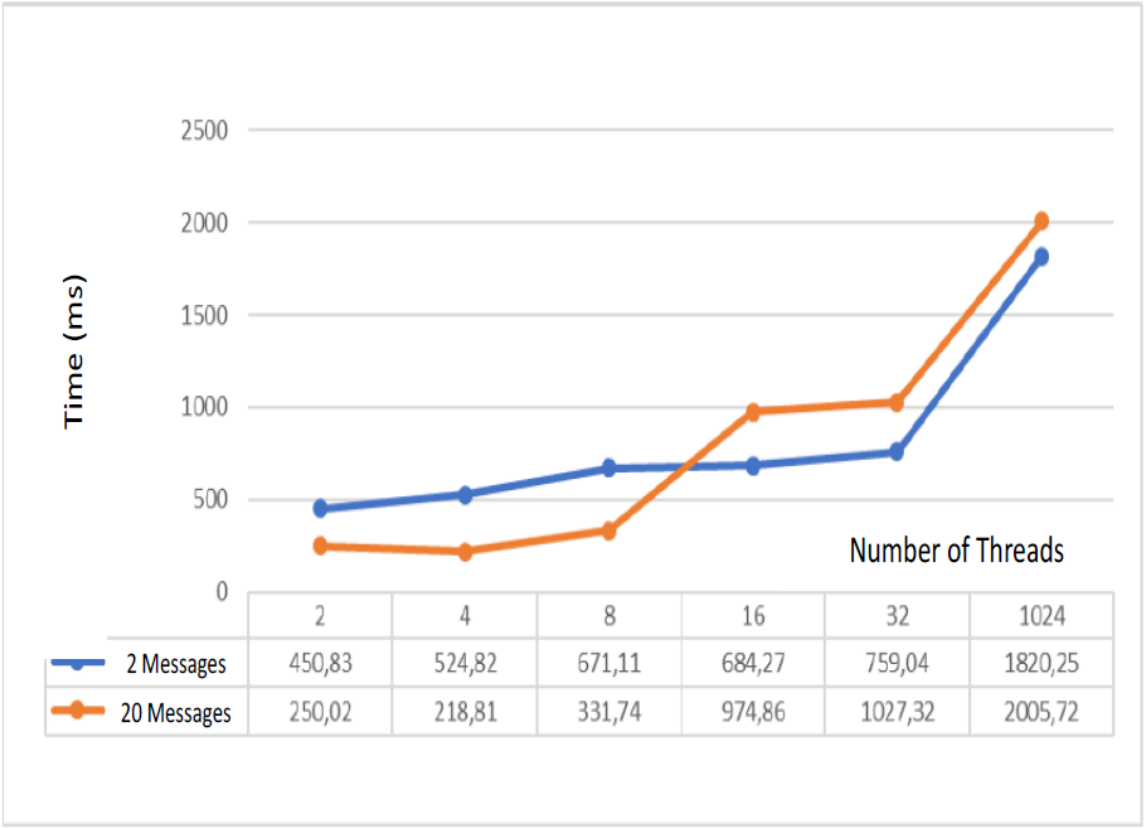
Resultados - Experimento 1

- Data-intensive.



Resultados - Experimento 2

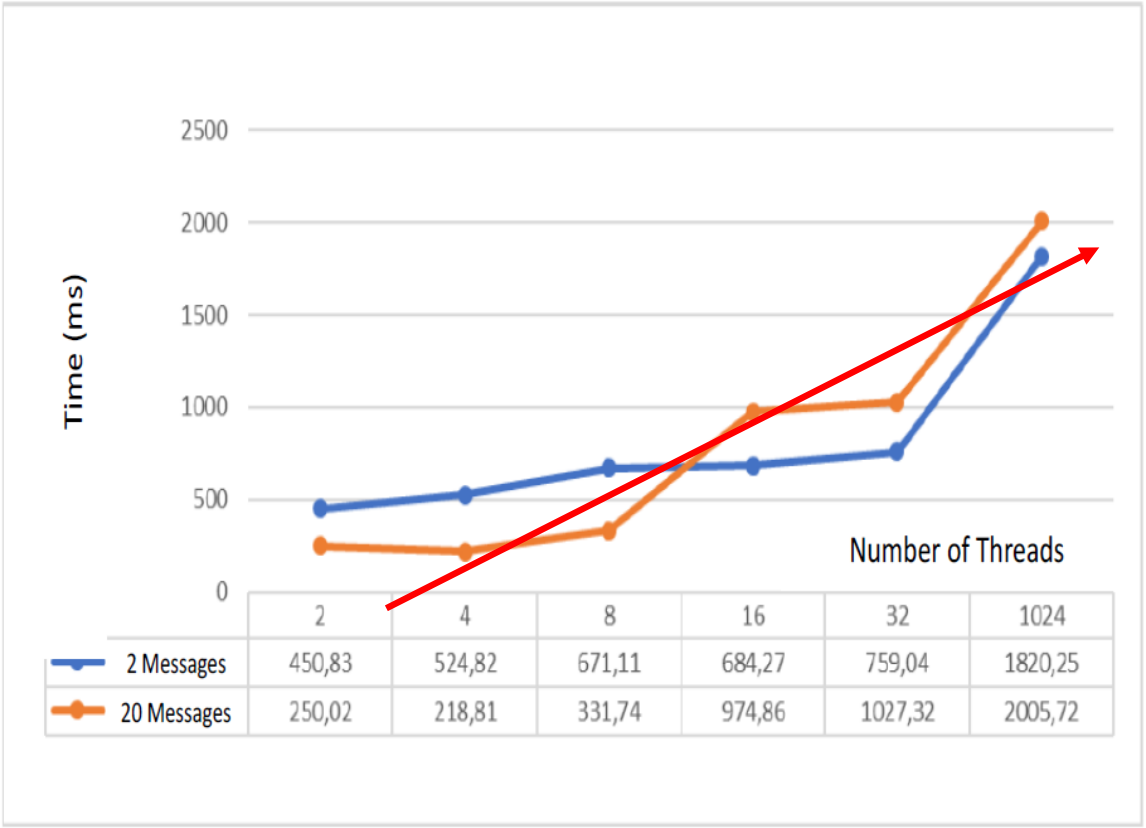
- Large-Scale.



	2	4	8	16	32	1024
2 Messages	450,83	524,82	671,11	684,27	759,04	1820,25
20 Messages	250,02	218,81	331,74	974,86	1027,32	2005,72

Resultados - Experimento 2

- Large-Scale.

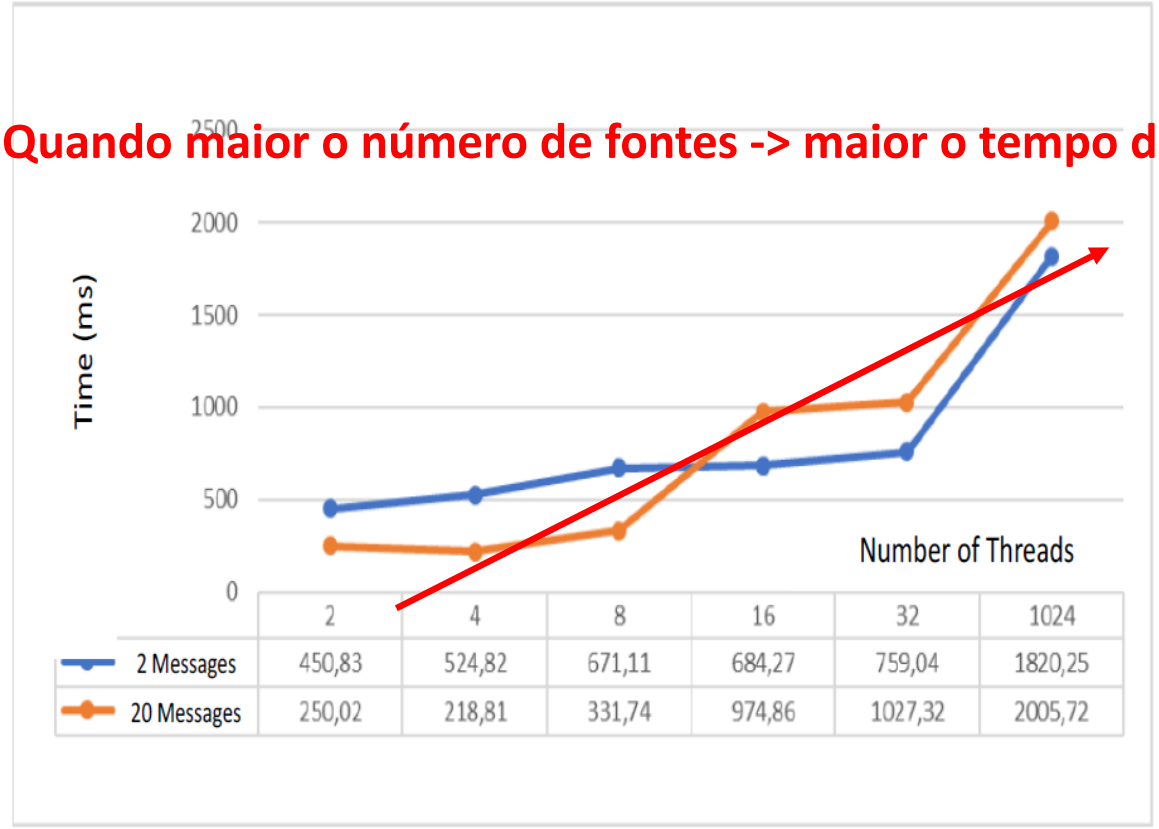


	2	4	8	16	32	1024
2 Messages	450,83	524,82	671,11	684,27	759,04	1820,25
20 Messages	250,02	218,81	331,74	974,86	1027,32	2005,72

Resultados - Experimento 2

- Large-Scale.

No geral: Quando maior o número de fontes -> maior o tempo de processamento

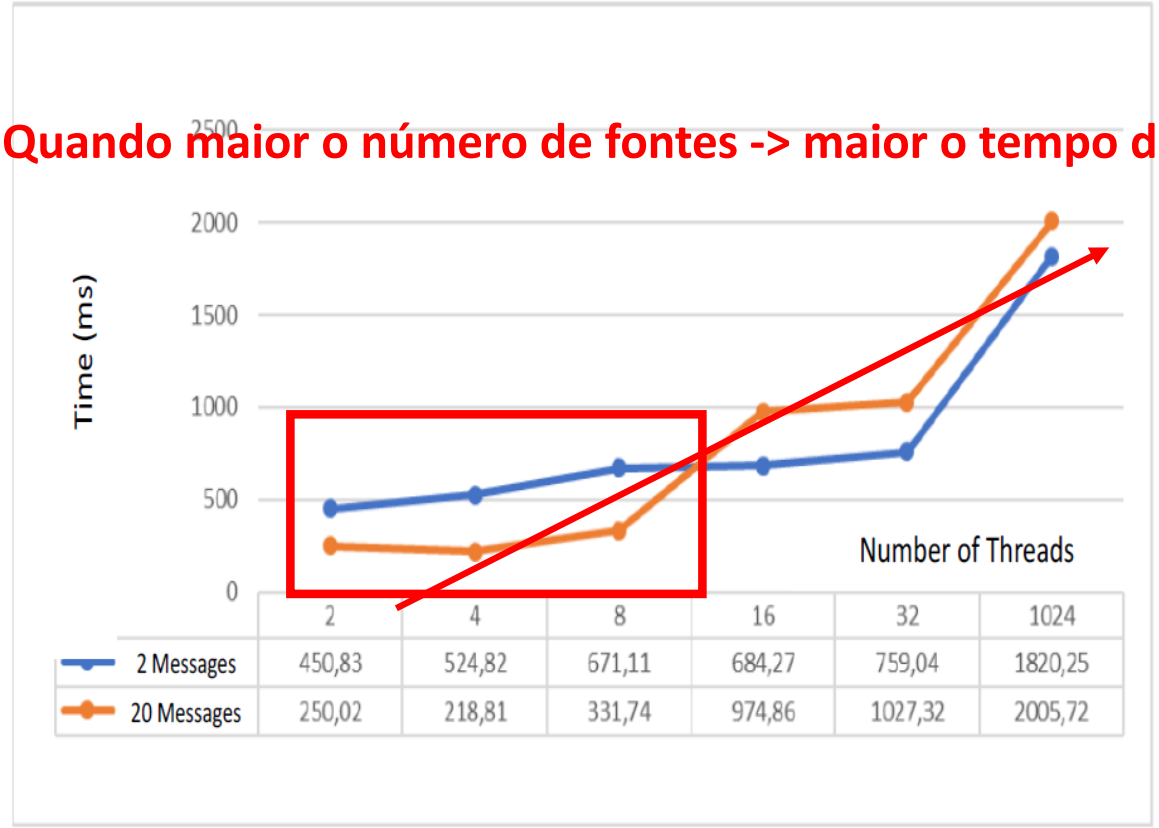


	2	4	8	16	32	1024
2 Messages	450,83	524,82	671,11	684,27	759,04	1820,25
20 Messages	250,02	218,81	331,74	974,86	1027,32	2005,72

Resultados - Experimento 2

- Large-Scale.

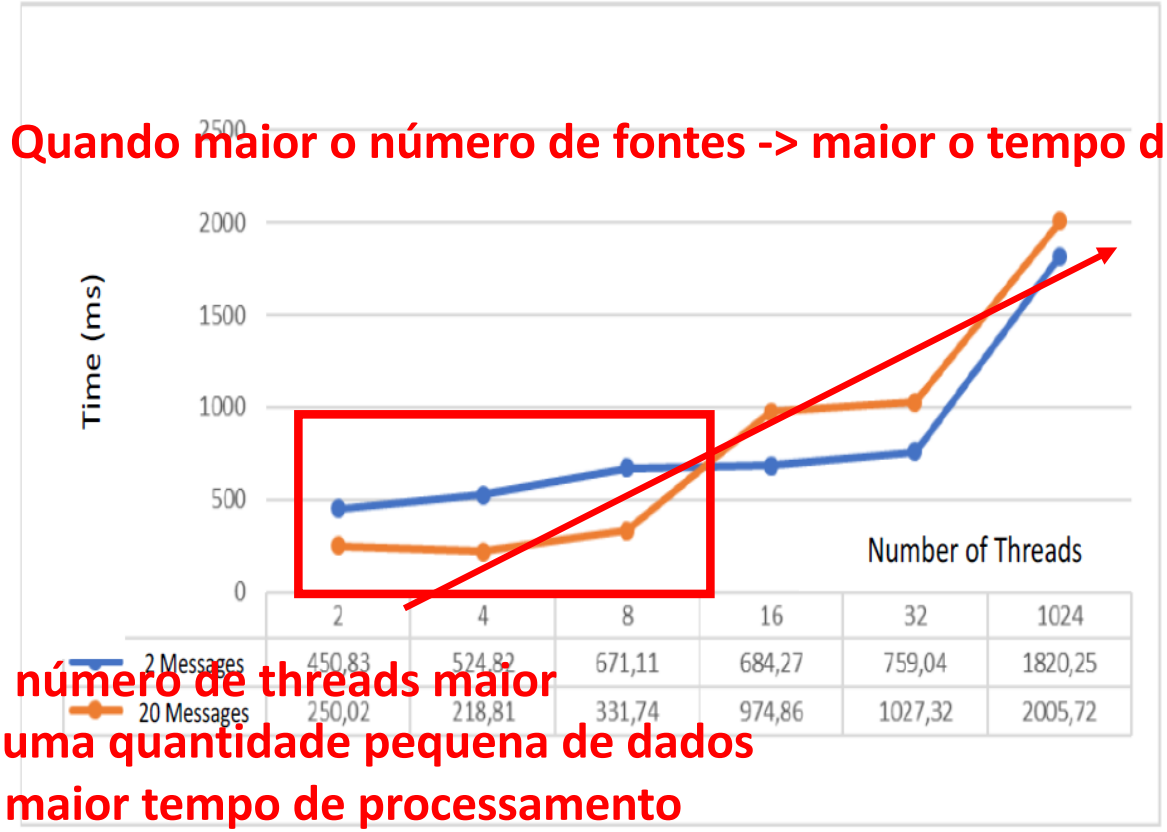
No geral: Quando maior o número de fontes -> maior o tempo de processamento



	2	4	8	16	32	1024
2 Messages	450,83	524,82	671,11	684,27	759,04	1820,25
20 Messages	250,02	218,81	331,74	974,86	1027,32	2005,72

Resultados - Experimento 2

- Large-Scale.

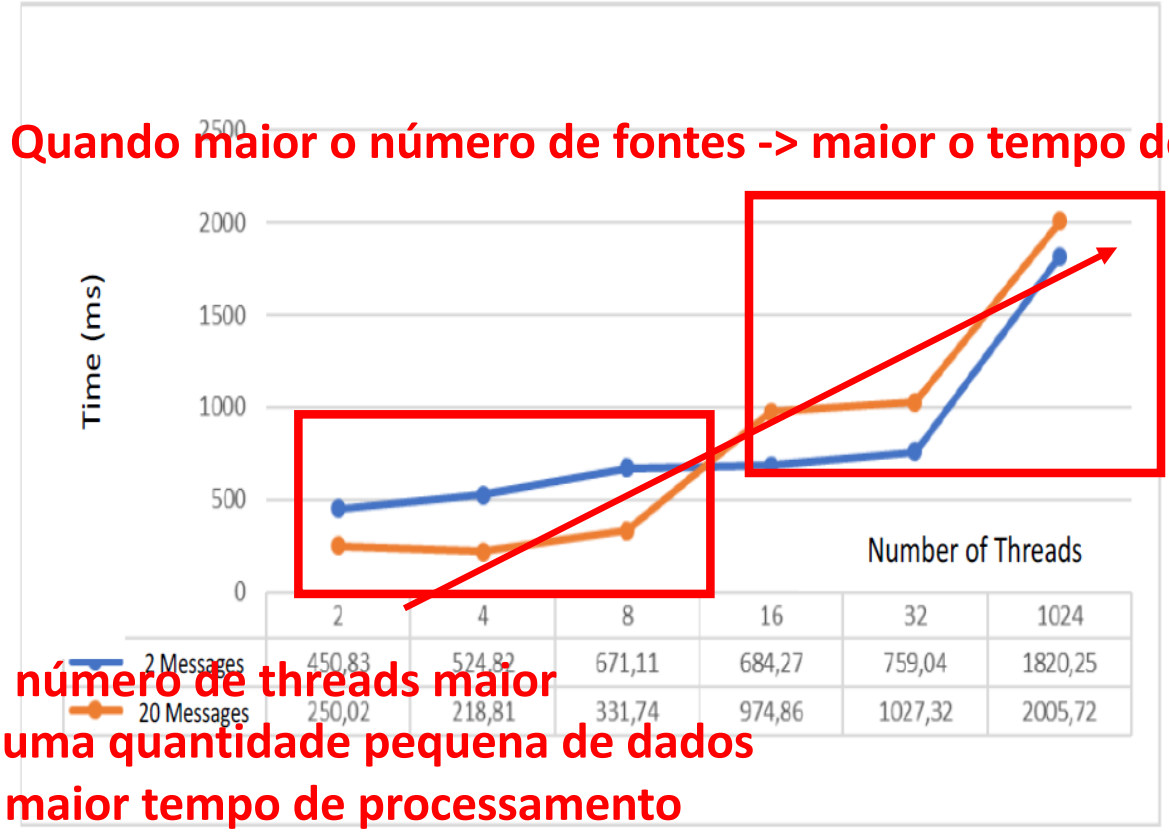


No geral: Quando maior o número de fontes -> maior o tempo de processamento

Alocar um número de threads maior
processar uma quantidade pequena de dados
Causa um maior tempo de processamento

Resultados - Experimento 2

- Large-Scale.

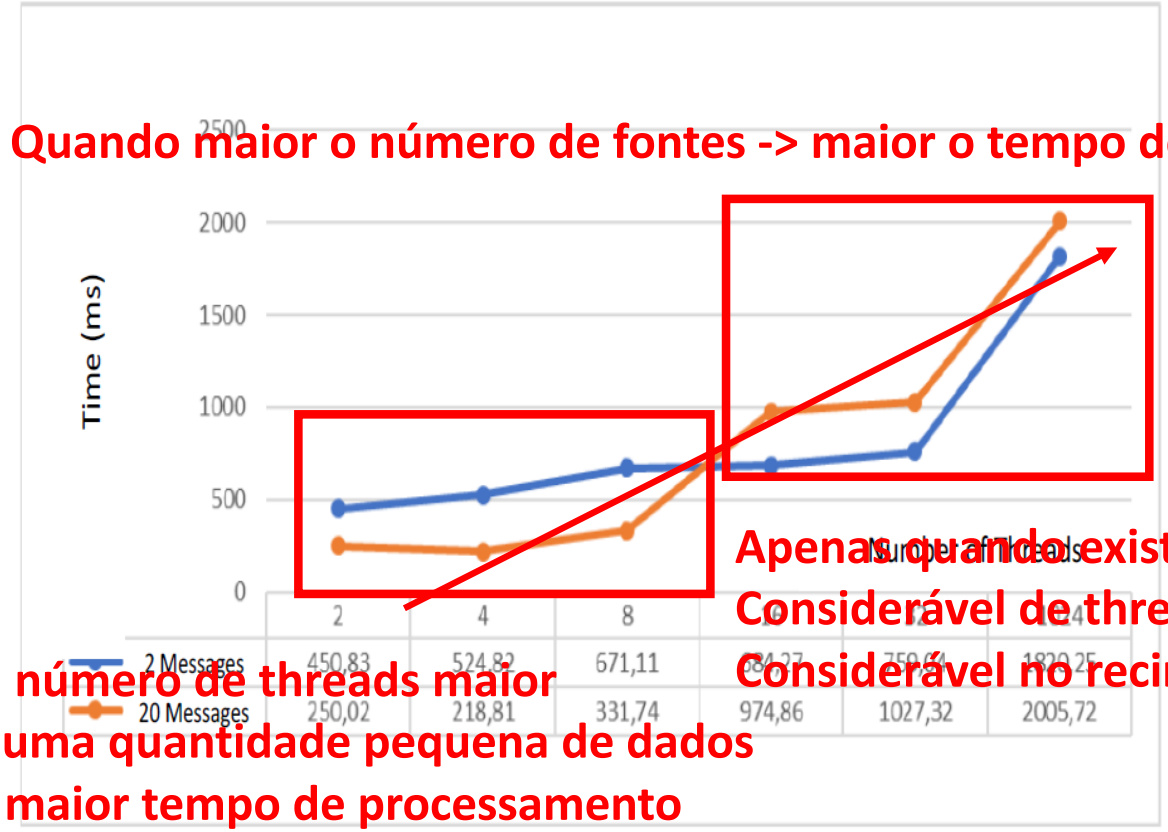


No geral: Quando maior o número de fontes -> maior o tempo de processamento

Alocar um número de threads maior
processar uma quantidade pequena de dados
Causa um maior tempo de processamento

Resultados - Experimento 2

- Large-Scale.



No geral: Quando maior o número de fontes -> maior o tempo de processamento

Apenas quando existe uma quantidade Considerável de thread existe um impacto Considerável no recimento dos arquivos.

Alocar um número de threads maior processar uma quantidade pequena de dados Causa um maior tempo de processamento

Considerações Finais

- A adoção de aplicações DIRT aumentou significativamente nos últimos anos;
- Porém soluções atuais não ofereciam soluções extensíveis a novas tecnologias;
- Nesse contexto foi proposto Doric, uma arquitetura flexível e modular para sistemas de larga escala e tempo real;
- Um protótipo foi implementado para testar a ferramenta;
- Os testes mostraram que a ferramenta foi capaz de processar um grande volume de dados em um período pequeno.

Trabalhos Futuros

- Executar uma analise comparativa com outras arquiteturas desenvolvidas;
- Testar o quão benéfico é o fator da flexibilidade em ambientes de desenvolvimento;
- Executar experimentos da implementação em um ambiente com mais recursos computacionais;
- Avaliar a arquitetura em outros cenários.

Referências

[1] R. A. Afonso, W. D. Silva, G. Tomas, K. Gama, A. Oliveira, A. Álvaro, and V. C. Garcia. Br-scmm: Modelo brasileiro de maturidade para cidades inteligentes. Simpósio Brasileiro de Sistemas de Informação, 2013.

[2] Farias, K., Garcia, A., Whittle, J., Chavez, C. V. F. G., & Lucena, C. (2015). Evaluating the effort of composing design models: a controlled experiment. *Software & Systems Modeling*, *14*(4), 1349-1365.

[3] Farias, K., Gonçales, L., Scholl, M., Oliveira, T. C., & Veronez, M. (2015). Toward an Architecture for Model Composition Techniques. In *SEKE* (pp. 656-659).

[4] L. R. Bencke, A. L. F. Perez, and O. da Costa Armendaris. Rodovias inteligentes: uma visão geral sobre as tecnologias empregadas no brasil e no mundo. iSys-Revista Brasileira de Sistemas de Informação, 10(4):80-102, 2017.

[5] J. Cao and J. Li. Large-scale real-time data-driven scientic applications. In Networking and Distributed Computing (ICNDC), 2011 Second International Conference on, pages 116-121. IEEE, 2011.

[6] W. Dubitzky, K. Kurowski, and B. Schott. Large-scale computing techniques for complex system simulations, volume 80. John Wiley & Sons, 2012.

[7] D. Evans. The internet of things how the next evolution of the internet is changing everything (april 2011). White Paper by Cisco Internet Business Solutions Group (IBSG), 2012.

[8] S. M. Kuo, B. H. Lee, and W. Tian. Real-time digital signal processing: fundamentals, implementations and applications. John Wiley & Sons, 2013.

[9] J. Menezes Jr and C. Gusmão. Intelimed:proposta de sistema de apoio ao diagnostico medico para dispositivos moveis. iSys-Revista Brasileira de Sistemas de Informação, 6(1):44-61, 2014.

[10] P. Waknis and J. Sztipanovits. Hard real-time scheduling for large scale process control applications. In Proceedings of the IEEE Workshop on Real-Time Applications, pages 71{75. IEEE, 1993.

[11] G. Wu, Z. Ao, J. Li, and Q. Zhou. A real-time receiving and distributed processing system for large-scale burst data. In 2011 Second International Conference on Networking and Distributed Computing (ICNDC), pages 111-115. IEEE, 2011.

Contato:

Lucian Gonçales

lucianjosegoncales@gmail.com