

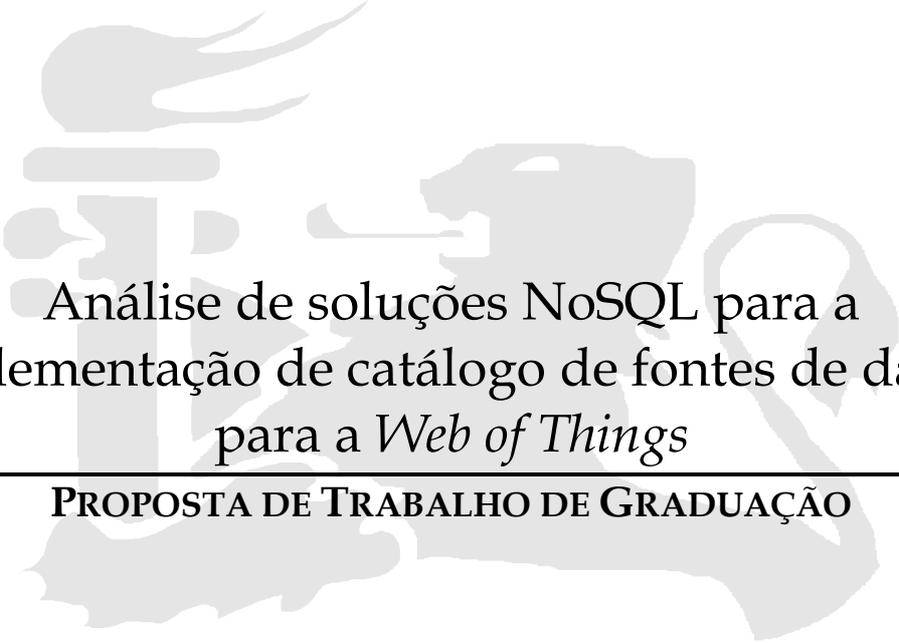


UNIVERSIDADE FEDERAL DE
PERNAMBUCO

GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

CENTRO DE INFORMÁTICA

2015.1



Análise de soluções NoSQL para a
implementação de catálogo de fontes de dados
para a *Web of Things*

PROPOSTA DE TRABALHO DE GRADUAÇÃO

Aluno: Alberto Trindade Tavares (att@cin.ufpe.br)

Orientadora: Bernadette Farias Lóscio (bfl@cin.ufpe.br)

Co-orientador: Marcelo Iury de Sousa Oliveira (miso@cin.ufpe.br)

Recife, ____ de Abril de 2015

Sumário

1. Resumo.....	3
2. Contextualização	3
3. Objetivos.....	4
4. Cronograma.....	4
5. Possíveis Avaliadores	5
6. Referências	5
7. Assinaturas	6

1. Resumo

O advento da *Web of Things* (WoT) trouxe novos desafios para a coleta, integração e processamento de dados, como o volume massivo, heterogeneidade e autonomia das fontes de dados. Uma alternativa é o uso de uma arquitetura de integração baseada em objetivo/consultas na qual as visões de integração seriam criadas de forma incremental e dinamicamente. Esta arquitetura demanda um catálogo de fontes de dados que deve ser flexível, para lidar com a heterogeneidade das fontes, e escalável, para lidar com o potencial volume massivo de fontes e consumidores de dados. Este trabalho propõe uma solução de catálogo de fontes de dados para WoT que faz uso de indexação automática e soluções de banco de dados NoSQL.

2. Contextualização

O volume de dados produzidos e compartilhados através da *Web* está crescendo rapidamente. Esses dados incluem desde informações textuais até conteúdo multimídia, em uma grande variedade de plataformas. Um dos tipos de dados que vem se destacando são os relacionados a observações físicas do mundo real. O aumento da coleta de dados do mundo físico se deve aos avanços recentes das tecnologias de comunicação com o objetivo de fornecer acesso ubíquo a Internet, assim como a redução do custo de sensores e dispositivos móveis. A extensão da Internet atual, permitindo a comunicação entre objetos e dispositivos físicos, é descrita pelo termo *Internet of Things* (IoT) [1, 2].

Nesse contexto, surge o conceito de *Web of Things* (WoT), que vai além do foco da Internet como forma de trocar de dados, ao trazer todos os recursos e interações envolvendo dispositivos, dados e pessoas para a *Web* [3]. O principal objetivo da WoT é aplicar a arquitetura da *Web* para conectar recursos físicos, tais como sensores, com outros recursos da *Web*. Dessa forma, objetos do mundo real passarão a ser tratados da mesma forma que qualquer outro recurso *Web*, sendo identificados por uma URI (*Universal Resource Identifier*) e acessíveis através de protocolos de comunicação, como o HTTP (*Hypertext Transfer Protocol*) [4].

O advento da WoT trouxe novos desafios para a coleta, integração e processamento de dados do mundo físico na *Web*. Entre os desafios, destacam-se o volume massivo, heterogeneidade e autonomia das fontes de dados. Além destes, a ausência de esquemas globais traz uma dificuldade adicional. Os recursos da WoT podem ser ubíquos e geralmente são limitados em termos de capacidade de armazenamento, processamento e comunicação [1]. Uma alternativa seria usar uma arquitetura de integração baseada em objetivo/consultas na qual as visões de integração seriam criadas de forma incremental e dinamicamente [5].

Esta arquitetura demanda um catálogo de fontes de dados da WoT, que podem ser desde dispositivos físicos até *datasets* da *Linked Open Data*. O catálogo de fontes deve ser flexível, para lidar com a heterogeneidade das fontes, e escalável, para lidar com o potencial volume massivo de fontes e consumidores de dados. Usualmente, os catálogos são desenvolvidos através de Sistemas de

Gerenciamento de Banco de Dados (SGBD) para a manipulação e persistência de seus dados. Desta forma, as características e configurações de SGBDs podem impactar diretamente o desempenho geral do catálogo. Em especial, o uso de mecanismos de indexação, particionamento, visões e cache podem ser fundamentais para o atendimento de níveis ideais de desempenho.

3. Objetivos

Este trabalho de graduação tem como objetivo principal o estudo de soluções de banco de dados NoSQL para o armazenamento de metadados sobre fontes da WoT, implementados por meio de um catálogo. Os sistemas de gerenciamento de banco de dados NoSQL se adequam melhor aos requisitos de escalabilidade e a não obrigatoriedade de utilização de um esquema comum para todos os registros.

Nós realizaremos uma análise comparativa do desempenho do catálogo com diferentes soluções NoSQL e com diversas configurações do SGBD, tais como o uso de índices e subcoleções. Um objetivo secundário desse trabalho é a implementação de um mecanismo para a criação automática de índices sobre o banco de dados NoSQL que apresentar os melhores resultados de desempenho para o armazenamento do catálogo de fontes de dados.

4. Cronograma

Nesta seção, é apresentado o cronograma previsto, na Tabela 1, para o desenvolvimento desse Trabalho de Graduação.

Atividade	Março	Abril	Mai	Junho	Julho
Formulação da proposta e revisão da literatura	X	X			
Estudo de soluções NoSQL para catálogo de fontes de dados		X	X		
Implementação de geração automática de índices			X	X	
Testes e experimentos		X	X	X	
Análise dos resultados			X	X	X
Elaboração do relatório			X	X	X
Preparação para apresentação e defesa					X

Tabela 1: Cronograma do Trabalho de Graduação

5. Possíveis Avaliadores

Os possíveis avaliadores para o resultado a ser obtido ao final de todas as etapas descritas neste documento são:

- Bernadette Farias Lóscio
- Kiev Santos da Gama

6. Referências

- [1] BARNAGHI, P., SHETH, A., HENSON, C. **From data to actionable knowledge: Big data challenges in the web of things**, Intelligent Systems, IEEE, vol. 28, no. 6, p. 6–11, 2013.
- [2] BISCHOF, S., KARAPANTELAKIS, A., SHETH, A., MILEO, A., NECHIFOR, S., BARNAGHI, P. **Semantic Modelling of Smart City Data**, W3C Workshop on the Web of Things, Berlim, Alemanha, 2014.
- [3] GUSTAFSON, S., SETH, A. **The Web of Things**, Computing Now, vol. 7, no. 3, 2014. Disponível em: <http://www.computer.org/portal/web/computingnow/archive/march2014>. Acesso em: 25 de Abril de 2014.
- [4] WILDE, E., MICHAELLES, F., LÜDER, S. **Leveraging the Web Platform for the Web of Things**, Position Paper for W3C's Web of Things Workshop, Berlim, Alemanha, 2014.
- [5] OLIVEIRA, M. I. S., GAMA, K. S., LOSCIO, B. F. **Waldo: Serviço para Publicação e Descoberta de Produtores de Dados para Middleware de Cidades Inteligentes**, XI Simpósio Brasileiro de Sistemas de Informação, 2015.

7. Assinaturas

Alberto Trindade Tavares
Orientando

Bernadette Farias Lóscio
Orientadora