



Marque a opção do tipo de trabalho que está inscrevendo:

Resumo

Relato de Caso

MANIPULAÇÃO DE DADOS AGRÍCOLAS UTILIZANDO BANCOS DE DADOS NoSQL E GEOESPACIAIS

AUTOR PRINCIPAL: Mauricio Alex Zientarski Karrei

CO-AUTORES: Felipe Giovanoni, Leandro Bonfante, Carlos Amaral Holbig

ORIENTADOR: Willingthon Pavan

UNIVERSIDADE: Universidade de Passo Fundo

INTRODUÇÃO

O armazenamento de dados vem sendo utilizado e melhorado desde os primórdios da computação, sendo importante para o sucesso de um sistema computacional. Na área agrícola, isso não é diferente. Dados sobre o clima, uso de maquinários, aplicação de produtos químicos, entre outros, possibilitam a geração de informações fundamentais para o bom manejo das culturas e seu respectivo sucesso. Porém, a existência de uma grande massa de dados, aliada a sua heterogeneidade de tipos, dificultam a coleta e o armazenamento (MURAKAMI, 2006).

Visando a melhoria no gerenciamento de dados, os Sistemas Gerenciadores de Bancos de Dados (SGBD's) tradicionais têm oferecido, cada vez mais, novos recursos para o armazenamento e o tratamento desses dados. Da mesma forma, esforços têm sido feitos na busca por soluções mais naturais, como bancos de dados NoSQL (REDMOND, 2012).

Desta forma, este trabalho visa apresentar uma plataforma para armazenamento e recuperação de dados não estruturados para a área agrícola.

DESENVOLVIMENTO:

Com o intuito de elencar tecnologias de bancos de dados capazes de suportar grandes massas de dados e armazenar tipos de dados complexos como informações geoespaciais, culminou-se inicialmente numa pesquisa bibliográfica acerca de SGBDs capazes de suportar o armazenamento desses tipos de dados complexos. Dentre os recursos elencados para o armazenamento de informações geoespaciais, constatou-se que a extensão geográfica para o

SGBD PostgreSQL, PostGIS, oferece a maior quantidade de recursos para manipulação de objetos geoespaciais, dentre outras ferramentas gratuitas como MongoDB ou Geocouch. Por outro lado, o armazenamento de informações dinâmicas e massas de dados não estruturados, direciona para a escolha entre os bancos de dados MongoDB, CouchDB e Cassandra, os quais são bases de dados NoSQL e possuem como característica a possibilidade de operar com grandes volumes de dados (REDMOND, 2012).

A partir destes estudos realizados, adotou-se para o desenvolvimento deste projeto o PostGIS para armazenamento de informações geoespaciais, por oferecer recursos de criação e armazenamento de objetos como polígonos, círculos e linhas, os quais são utilizados na criação de mapas e objetos geoespaciais. Da mesma forma, utilizou-se, também, o MongoDB para armazenamento de grandes volumes de dados não estruturados, visto que este é construído com foco na alta performance, velocidade e escalabilidade. Quanto a escalabilidade, o MongoDB dispõe de recursos para manter cópias do banco de dados em diferentes servidores, de forma a restaurar uma destas instantaneamente em caso de falhas. O MongoDB armazena dados em formatos baseados em objetos JSON (BSON - *Binary JSON*), tornando-o extremamente rápido e escalável, além de oferecer flexibilidade para operar com tipos de dados complexos (REDMOND, 2012).

Dessa forma, desenvolveu-se um modelo genérico para o recebimento de dados agrícolas utilizando MongoDB. Os dados são armazenados em uma estrutura não tabular, cuja forma pode se adaptar ao tipo de informação recebida (LAZZARETTI, 2013). Esta técnica de armazenamento de dados é útil para a necessidade deste trabalho pela sua característica de não manter as informações em tabelas, como em SGBD's tradicionais, mas sim em coleções de dados organizadas em esquemas, o que permite dinamizar a entrada de dados.

Os esquemas foram construídos de maneira a permitir a gravação de diferentes informações advindos do campo, tais como dados de estações meteorológicas ou de maquinários, sem a necessidade de adequar o modelo de banco de dados para cada tipo dado. Isto se tornou possível pela viabilidade de construir objetos BSON dinâmicos, ou seja, capazes de guardar informações independentemente do tipo do dado e da quantidade de informações que se deseja armazenar. Além disso, possibilita adequar facilmente a estrutura para a variabilidade dos dados, diferentemente do que é encontrado em modelostabulares, onde cada tipo de dado deve ser informado em uma coluna específica.

CONSIDERAÇÕES FINAIS:

Com o desenvolvimento, observa-se a potencialidade da utilização de tecnologias de bancos de dados geoespaciais e NoSQL, servindo como alternativa adequada para manipulação de grandes volumes de dados e informações heterogêneas.

O projeto apresentado pode ser incorporado em diferentes sistemas de manipulação de dados agrícolas, incluindo diferentes linguagens de programação, como Java e JavaScript.

REFERÊNCIAS

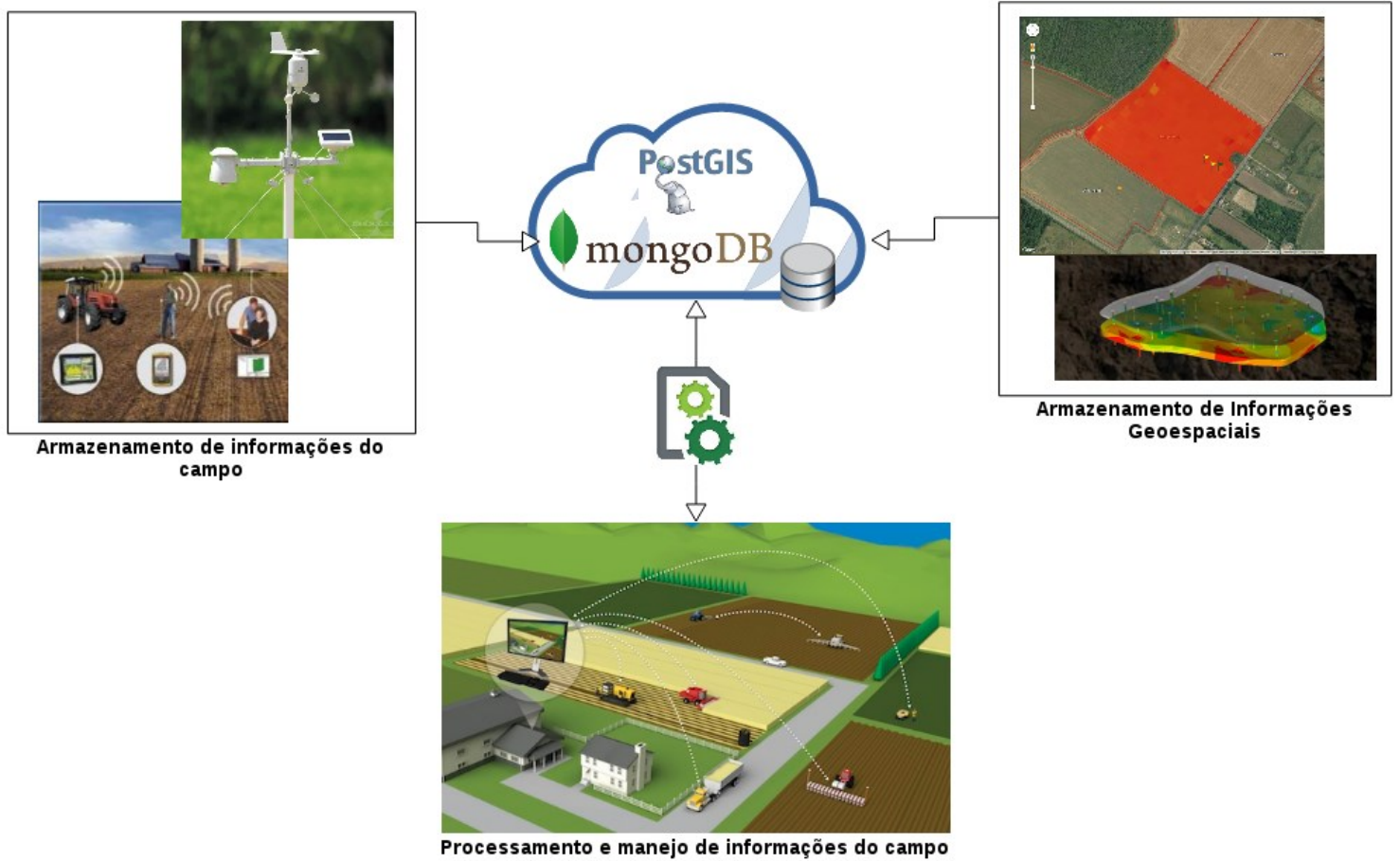
LAZZARETTI, Alexandre. T. Integração de banco de dados e modelos de simulação de culturas para estimar o impacto de mudanças do clima no rendimento de grãos e na severidade da giberela e trigo, 192 p. Tese (Doutorado). Universidade de Passo Fundo, Passo Fundo, 2013.

MURAKAMI, Edson. Uma infra-estrutura de desenvolvimento de sistemas de informação orientados a serviços distribuídos para agricultura de precisão, 192 p. Tese (Doutorado). Universidade de São Paulo, 2006.

REDMOND, Eric; WILSON, Jim R. Seven databases in seven weeks: a guide to modern databases and the NoSQL movement. Pragmatic Bookshelf, 2012.

TRIMBLE, Agriculture. Disponível em: <<https://www.connectedfarm.com/>>. Acesso em: Jun. 15, 2015.

NÚMERO DA APROVAÇÃO CEP OU CEUA (para trabalhos de pesquisa): Número da aprovação.



ANEXOS

Figura 1. Modelo arquitetural de banco de dados para centralização de informações Agrícolas