

Projeto



Observatório Nacional de Clima e Saúde

Coordenação Técnica Institucional: Fiocruz e INPE

Coordenação Nacional

CGVAM- Coordenação Geral de Vigilância Ambiental

Secretaria de Vigilância em Saúde

Ministério da Saúde

Apoio Institucional

OPAS – Organização Pan-Americana de Saúde



Pequeno Guia de Orientação Tecnológica para as Geotecnologias (TI Espaciais) Associadas a FASE 1 - Acesso a Base de Dados Distribuídas do Projeto *Observatorium*: Observatório de Clima e Saúde

Índice

Índice	2
Lista de abreviaturas e siglas	3
Introdução	4
Serviços baseados em Mediadores	4
Módulo de Integração	5
Módulo de Otimização	5
Módulo de Execução	6
Anexo – Conceitos	8
Arquitetura orientada a serviço	8
Serviços Web – padrão W3C	8
Serviços geográficos – padrão OGC	9
OGC versus W3C	9
Referências	10

Lista de abreviaturas e siglas

GML	Geography Markup Language
OGC	Open Geospatial Consortium
SOA	Service Oriented Architecture
SOAP	Simple Object Access Protocol
W3C	World Wide Web Consortium
WCS	Web Coverage Service
WFS	Web Feature Service
WMS	Web Map Service
WPS	Web Processing Service
WFS-T	Transactional Web Feature Service
WSDL	Web Services Description Language
XML	Extensible Markup Language

Introdução

Este documento apresenta a solução usada na arquitetura do Observatorium de Clima e Saúde para acessar diferentes fontes de dados com distintos modelos de dados e localizadas em servidores remotos pertencentes aos produtores dos dados propriamente ditos.

A arquitetura adotada para o Observatorium de Clima e Saúde é uma extensão da Arquitetura proposta e implementada em Xavier (2008). É essencialmente uma arquitetura orientada a serviços Web, onde estes serviços são destinados à integração de diferentes fontes de dados geográficos e um serviço particular que permite a colaboração interativa de usuários.

Essa estratégia propõe uma abordagem virtual na integração dos dados, ou seja, os dados ficam nas fontes e as informações são extraídas diretamente das mesmas quando uma consulta é processada. A principal vantagem da abordagem virtual é a não replicação de dados, o que implica que as informações recuperadas estão sempre atualizadas. A desvantagem reside no tempo de resposta às consultas, que dependendo da quantidade de fontes de dados envolvidas e da informação requerida, podem levar um tempo considerável para ser processadas. Apesar disso, a abordagem virtual é a mais apropriada quando há um grande número de fontes de dados e quando os mesmos são atualizados com frequência.

A arquitetura possui uma camada chamada Mediador, responsável pela ligação entre o usuário e as fontes de dados. Numa arquitetura mediada, existe uma camada chamada Adaptador (figura 1) que faz a ligação entre o mediador e as fontes de dados provendo o acesso aos dados e escondendo a heterogeneidade da fonte de dados para o mediador. A seguir será detalhada a camada Mediador.

Serviços baseados em Mediadores

A camada Mediador faz o acesso aos dados através de protocolos abertos e aderentes a padronização de serviços geográficos proposta pelo OGC- Open Geospatial Consortium. O mediador é composto dos módulos Integração, Otimização, Execução e Participação, figura 1. Serão detalhados os 3 primeiros módulos que estão envolvidos em uma consulta distribuída.

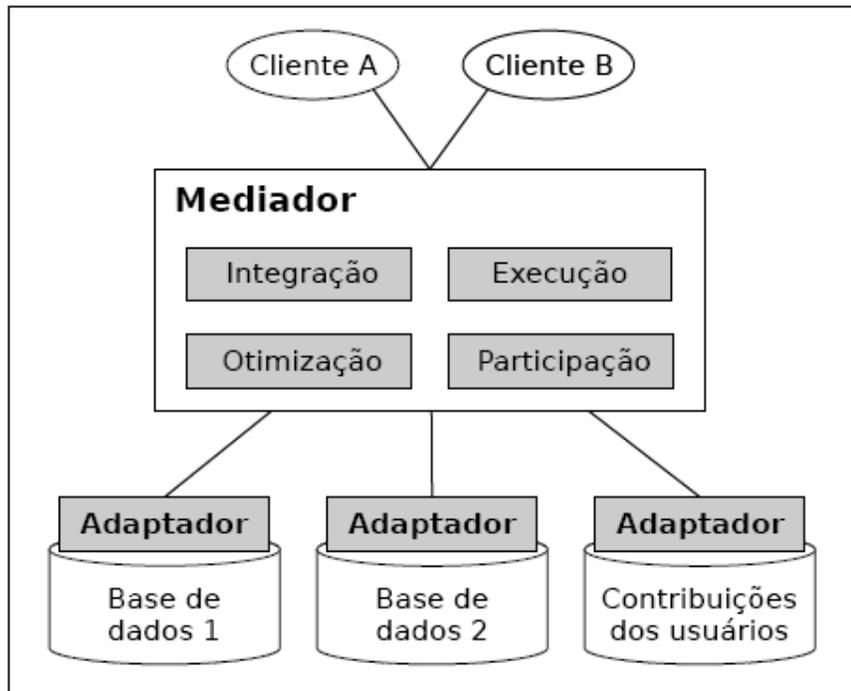


Figura 1 – Visão geral da arquitetura proposta por Xavier (2008)

Módulo de Integração

O Módulo de Integração do Componente Mediador tem por objetivo integrar as informações das diferentes fontes de dados usadas pelo mediador. Este módulo contém uma lista dos servidores WFS que irão prover as fontes de dados para criar um esquema mediado.

Características:

- Atua como um cliente WFS
- Verifica quais servidores irão atender uma consulta
- Verifica qual é o modelo de dados
- Constrói um catálogo das propriedades de cada camada

Ao atuar como cliente WFS, o Módulo de Integração armazena as informações obtidas de forma permanente, com a possibilidade de atualizar-se a qualquer momento. Esta persistência tem por finalidade evitar que a cada consulta de um cliente para o mediador, este precise recuperar as informações junto aos servidores WFS que atuam como adaptadores para as fontes de dados.

Módulo de Otimização

Este módulo decide como as consultas serão executadas após verificar se cada propriedade listada numa determinada consulta é válida. Se a consulta envolver apenas uma camada, esta será encaminhada para o servidor que contém esse dado. Caso a consulta envolva mais de uma camada este módulo a decompõe em subconsultas, agrupadas pelo tipo de dado pedido, que são enviadas aos respectivos servidores. A figura 2 ilustra os 4 processos deste módulo.

O primeiro processo recebe como entrada uma consulta e um catálogo das propriedades válidas, e identifica a camada que é o alvo desta consulta. Se a camada pedida for inválida, uma exceção deve ser enviada ao requisitante.

No segundo processo, uma operação é criada para o tipo de dado pedido. Se não for identificado o uso de um filtro para a consulta, o Módulo de Otimização envia a operação criada para o Módulo de Execução.

Uma vez identificada uma restrição (filtro) sobre a consulta o terceiro processo é acionado sendo criada uma árvore de consulta inicial.

O quarto e último processo executado no Módulo de Otimização é otimizar a árvore de consulta inicial montada no terceiro processo. O objetivo é montar as subconsultas de forma que nenhuma operação ocorra no mediador, mas sim nos seus servidores correspondentes.

Módulo de Execução

Uma vez montadas as consultas distribuídas, estas precisam ser executadas nos respectivos servidores que contém seus dados. A estratégia de execução usada nesta arquitetura é inspirada no modelo de iteradores, tendo em vista que as operações contém a implementação de sua execução. Cada operador possui um nó de consulta, que será usado na construção de uma nova consulta a ser enviada para um servidor de dados. A execução segue da raiz para os ramos (top-bottom), executando as operações à medida que são necessárias.

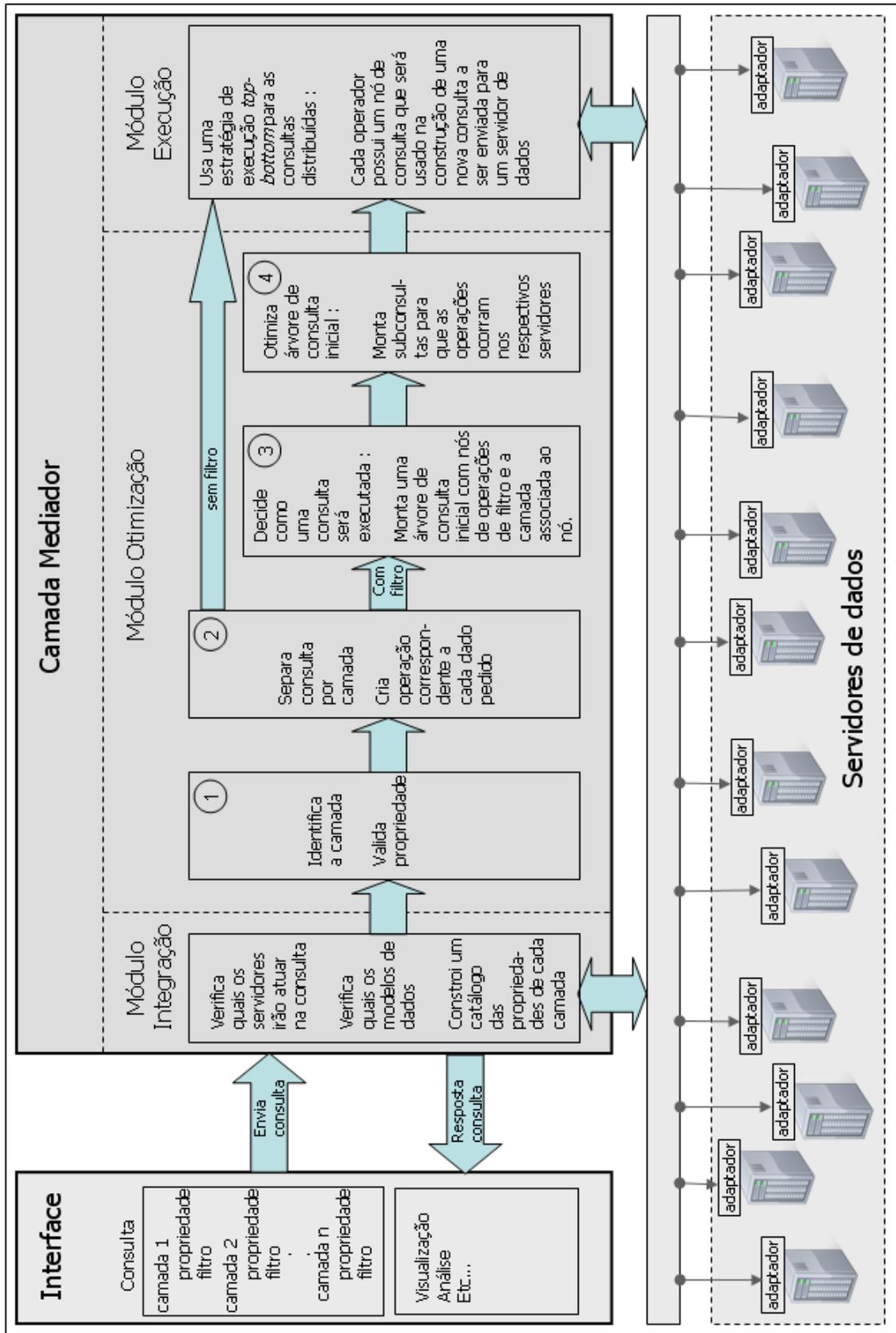


Figura 2 – Componente Mediator

Anexo – Conceitos

A seguir serão apresentados alguns conceitos envolvidos em uma arquitetura orientada a serviço na Web usando dados geográficos.

Arquitetura orientada a serviço

Uma arquitetura orientada a serviços (SOA – Service Oriented Architecture) é um tipo de arquitetura de software onde as funcionalidades implementadas pelas aplicações devem ser disponibilizadas na forma de serviços. Um serviço, do ponto de vista da arquitetura SOA, é uma função de um sistema computacional que é disponibilizada para outro sistema.

A arquitetura SOA é baseada nos princípios da computação distribuída e usa o paradigma request/reply para estabelecer a comunicação entre os sistemas clientes e os sistemas que implementam os serviços, como mostra a figura 3.

Freqüentemente estes serviços são disponibilizados por interfaces acessíveis através de web services padrão W3C ou outra forma de comunicação entre aplicações. Uma delas é a proposta pelo OGC, através de normas e especificações para serviços usando dados geográficos.



Figura 3 – Exemplo de arquitetura orientada a serviços

Serviços Web – padrão W3C

Serviços Web (Web services) são componentes de softwares que independem de implementação e plataforma e podem ser descritos, publicados e invocados através de mensagens padrão XML e que utilizam o paradigma request/reply para estabelecer a comunicação entre os sistemas clientes e os sistemas que implementam os serviços seguindo os padrões definidos pelo W3C.

A figura 4 mostra de maneira simples como se usa um serviço. Primeiramente é preciso conhecer a descrição do serviço para que se estabeleça a correta comunicação entre cliente e serviço. A descrição é feita usando WSDL (Web Services Description Language). Essa descrição é enviada ao cliente no formato XML. O cliente faz então uma solicitação seguindo a descrição e enviando conforme o protocolo SOAP (Simple Object Access Protocol) e recebe a resposta ao serviço em formato XML.



Figura4 – Comunicação entre cliente/servidor em um serviço Web

Serviços geográficos – padrão OGC

O OGC possui algumas especificações de serviços Web, os mais difundidos e implementados são WMS (Web Map Service), WFS (Web Feature Service), WCS (Web Coverage Service).

WFS é uma especificação para acesso e manipulação de dados geográficos na Web, permitindo que aplicações cliente visualizem informações geográficas na Internet. No WFS os dados são transportados na forma de GML (Geography Markup Language). Como o próprio nome indica, o GML é um markup XML para dados geográficos.

As operações suportadas pelo WFS são

WFS Básico:

- GetCapabilities:
- DescribeFeatureType
- GetFeature

WFS Transacional (WFS-T):

- Transaction
- LockFeature (opcional)

A figura 5 ilustra um pedido de feição nos passos de 1 a 3. Primeiramente é preciso requisitar a versão desse serviço (GetCapabilities), é retornado um documento XML. Depois é preciso obter uma descrição do tipo de feição desejada (DescribeFeatureType), é retornado um documento XML com o esquema GML da feição. Finalmente é solicitada a feição (GetFeature) seguindo a descrição. As feições são retornadas, geralmente no formato GML.

OGC versus W3C

A criação das especificações no OGC para serviços com dados geográficos na Web tiveram início em 1997 enquanto no W3C a padronização de serviços surgiu em 2002. Por essa razão as especificações do OGC não seguem as normas do W3C para comunicação entre serviços. Por sua vez o XML, padrão de intercambio de dados na Web, não foi concebido para suportar dados geográficos, por essa razão o OGC definiu o GML, que é uma especificação do OGC cujo objetivo é codificar informação geográfica no formato XML.

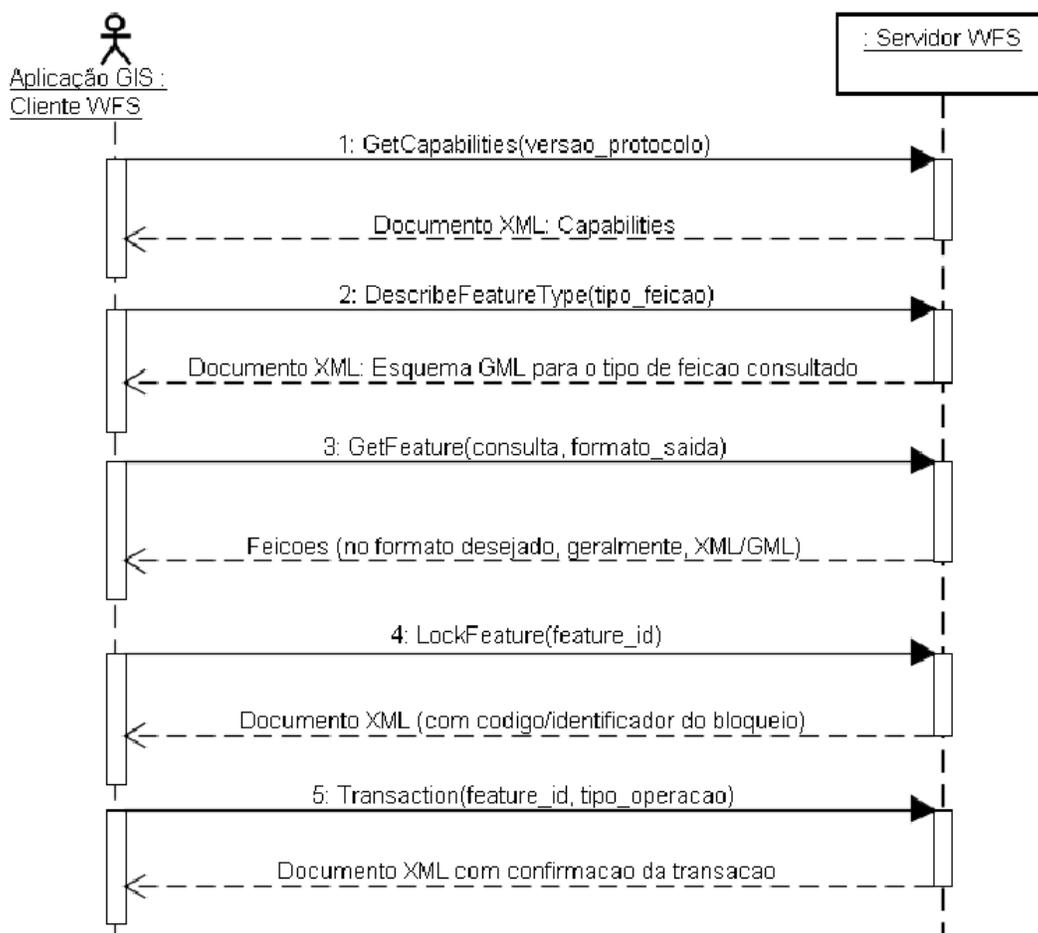


Figura 5 – Serviço WFS

Referências

Casanova, M. A.; Brauner, D. F.; Câmara, G.; Lima JR., P. O. Integração e interoperabilidade entre fontes de dados geográficos. In: Casanova, M. A.; Câmara, G.; Davis JR., C. A.; Vinhas, L.; Queiroz, G. R. (Ed.). Bancos de dados geográficos. Curitiba: MundoGEO, 2005. p. 317-352.

Xavier, E.M.A ; Serviços Geográficos baseados em mediadores e padrões abertos para monitoramento ambiental participativo na Amazônia, dissertação de mestrado, 2008