

Web Services based Distributed Query System

Marco Fernandes . Joaquim Arnaldo Martins . Joaquim Sousa Pinto . Pedro Almeida . Helder Zagalo



Objetivos

- Criação de um módulo de acesso e pesquisa sobre um repositório distribuído, escalável e de grande interoperabilidade.
- Permitir a (des)agregação de repositórios de forma transparente de um repositório virtual único.
- Criação de um Web Service de administração para gestão remota de catálogos de informação.
- Implementação de repositórios digitais distribuídos através do módulo DisQS.
- Transparência quanto às tecnologias de indexação.

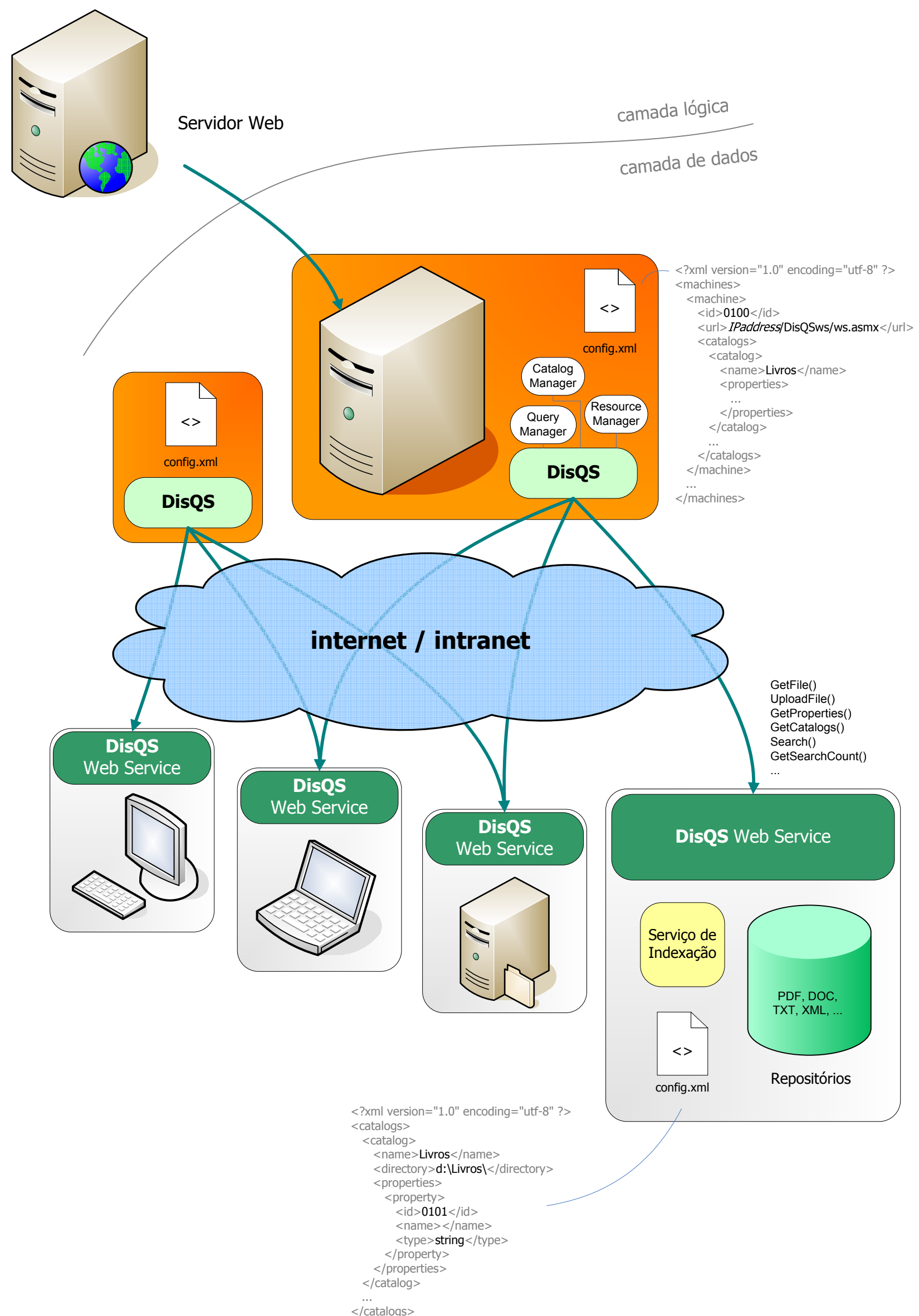
Funcionalidades

- Pesquisa simultânea em todos os repositórios da rede DisQS.
- Criação e edição remota de catálogos.
- Mecanismos de replicação de dados e balanceamento de carga.
- Implementação de caches de recursos e de resultados de pesquisas.
- Mecanismos de substituição ou remoção automática de recursos desatualizados.
- Métodos genéricos para fazer o *upload* e o *download* de ficheiros XML e ficheiros de imagens.

Por exemplo: redimensionamento e/ou conversão de formatos prévio ao pedir uma imagem.

Tecnologias

- Linguagens de programação
 - C# (.NET Framework)
- Standards
 - XML, SOAP, WSDL (Web Services)
- Serviços de indexação
 - Indexing Service (Microsoft)
 - Swish-e
 - Bases de dados relacionais



Protótipo

- Integração no SInBAD (Sistema Integrado para Bibliotecas e Arquivos Digitais)
 - Programa Aveiro Digital 2003-2006
 - POSI
- Rede DisQS constituída por diversos computadores da Universidade de Aveiro.
- Pesquisa distribuída em diversos tipos de documentos (PDF, XML, HTML, DOC, TXT, entre outros).

Resultados

- Integração em Bibliotecas Digitais e Arquivos Digitais para implementar repositórios e mecanismos de pesquisa distribuídos.
- Resultados provisórios (protótipo):
 - 6 computadores ligados à rede
 - cerca de 1 milhão de documentos armazenados no repositório
 - pesquisas duram entre centésimos de segundo (para poucas centenas de resultados) e 2-3 segundos para palavras muito comuns