

I - Dados da Equipa de Investigação

1 - Nome do Coordenador do projeto: Fernando Miguel Gamboa de Carvalho

2 - Área Departamental: ADEETC

3 - Membros da equipa de investigação / Nome da Instituição e/ou Área Departamental do ISEL:

- Fernando Miguel Gamboa de Carvalho, docente 1619, ISEL, ADEETC
- António Rito Silva, IST, INESC-ID
- Diogo Poeira, aluno de mestrado número 36238, MEIC, ADEETC

II – Dados do Projeto

1 – **Acrónimo:** WebFluid

2 – **Título:** Reactive Web streams for Big Data

3 - **Data (início e fim):** 7 de Setembro de 2020 a 6 de Setembro de 2021

4 – **Financiamento solicitado:** 5000 euros

5 - **Sumário para comunicação (máximo 365 caracteres) (Âmbito e Objectivos):**

O projecto WebFluid trata da otimização da transmissão de Big Data na Web entre serviços e aplicações, através da eliminação de operações de IO bloqueantes. O objectivo principal é garantir um fluxo reactivo (não-bloqueante) desde: 1) uma fonte de dados, 2) passando por diversas transformações entre serviços Web, 3) até à produção final de uma user-interface HTML.

Síntese Descritiva (máximo 2500 caracteres) (Foco no Estado da Arte, Conhecimentos da Equipa e Resultados):

Lidar com big data em aplicações Web não trata apenas com o tamanho dos dados, mas também com os dados espalhados por diferentes nós em toda a World Wide Web, o que afeta diretamente a latência. Por outro lado, o big data está geralmente disponível em tempo real, o que exige soluções capazes de enviar dados de forma assíncrona para os clientes, em vez de ter consumidores que puxam os dados da fonte de forma síncrona.

Enfrentando estes dois desafios, alcançámos uma proposta bem-sucedida de modelos de ordem superior (HoT – higher-order templates) no nosso trabalho anterior [CD19]. A proposta de HoT pode lidar com grandes conjuntos de dados em aplicações web modernas, fornecendo um comportamento progressivo da interface do utilizador e mostrando melhor desempenho do que a concorrência em benchmarks realistas que lidam com dados reais, como por exemplo a Last.fm Web Api.

HoT é o resultado de um projeto de código aberto - HtmlFlow - que consiste numa DSL Java para escrever HTML de forma fluente e tipificada e classificado com 60 estrelas pela comunidade no maior repositório de código aberto <https://github.com/xmlnet/HtmlFlow>. A natureza funcional e composicional do HtmlFlow distingue-se da concorrência por: ser baseada na ausência de estrutura de dados de suporte; permitir chamadas encadeadas a métodos; e fornecer segurança no HTML. Estas características têm contribuído para a crescente popularidade da biblioteca HtmlFlow, além do melhor desempenho que apresenta em relação a outros mecanismos actuais em benchmarks reconhecidos, tais como o spring-comparing-template-engines.

Na presente proposta, propomos dar um passo à frente, aumentando as capacidades da abordagem HoT com suporte para processamento de dados não-bloqueante, tal como o standard de reactive streams que fornece um padrão para o processamento de fluxo de dados assíncrono em diferentes ambientes tecnológicos como JVM, Javascript e outros. Os mecanismos de templates Web (como JSP ou Thymleaf) suportam apenas resolução bloqueante do HTML, dada a natureza dos templates baseados em texto. O HtmlFlow interrompe a gênese conceptual dos templates de texto e baseia-se em templates orientados a funções de primeira classe que aproveitam ao máximo as características intrínsecas da composição de funções.

Delegaremos na infraestrutura HtmlFlow como a espinha dorsal das aplicações Web com apresentação do lado do servidor (server-side presentation), que ainda corresponde aos dias de hoje à maior base instalada em aplicações Web. No nosso trabalho mais recente [CDG20] demonstrámos como a abordagem de HoT pode ser usada em casos reais de utilização de aplicação web bem conhecidas (e.g. spring-petclinic), alavancando a homogeneidade tecnológica entre o back-end e o front-end em torno de uma única linguagem de programação de host, como o Java. Essas observações dão perspectivas promissoras sobre o uso dessa mesma metodologia noutras aplicações web.

Para avaliar os resultados deste trabalho, planeamos testá-lo num conjunto de benchmarks conhecidos e numa seleção de aplicações web do mundo real que abrangem diferentes características da carga de trabalho em cenários de Big Data.

Temos uma vasta experiência no desenho, desenvolvimento e liderança de projetos complexos de aplicações web. Essa experiência foi acumulada em mais de 20 anos de pesquisa e desenvolvimento de aplicações e frameworks Web inovadoras em projectos de desenvolvimento do Centro de Calculo do ISEL e do INESC ID, bem como na docência da unidade curricular de Desenvolvimento de Aplicações da Web. Além disso, o IR desta proposta tem larga experiência no sector como líder de projetos e desempenhando a função de manager nas empresas Altitude Software e Quatro SI. Acreditamos que o sucesso dos nossos projetos, como o HtmlFlow, são o resultado da conciliação da nossa experiência entre a academia e a indústria, combinando rigor com pragmatismo.

6 – Publicações Anteriores

CD19 2019 F.M.Carvalho, L. Duarte, "Hot: Unleash web views with higher-order templates" In: Proceedings of the 15th International Conference on Web Information Systems and Technologies - Volume 1: WEBIST,. pp. 118–129. WEBIST '19, INSTICC, SciTePress (2019).

CDG20 2020 F. M. Carvalho, L. Duarte, J. Gouesse, "Text Web Templates considered Harmful" In: Lecture Notes in Business Information Processing, Springer (em fase de publicação)