

# Uma Linguagem Específica de Domínio com Geração de Código Paralelo para Visualização de Grandes Volumes de Dados

**Cleverson Ledur, Dalvan Griebler, Luiz Gustavo  
Fernandes, Isabel Manssour**

Pontifícia Universidade Católica do Rio Grande do Sul - PUCRS  
Programa de Pós-Graduação em Ciência da Computação - PPGCC  
Grupo de Modelagem de Aplicações Paralelas - GMAP

**Escola Regional de Alto Desempenho - ERAD**

Abril de 2015

# Sumário

- 1 **Introdução**
- 2 **DSLs para Visualização de Dados**
- 3 **DSL-VIS**
- 4 **Estado Atual da Pesquisa**
- 5 **Conclusão**

# Análise de Big Data

- Crescimento exponencial na produção de dados.
- Informações importantes. [1].
- Técnicas para análise de dados
- Visualização de Dados [2].

## Visualização de Informações

- Desde a pré-história.
- Comunicação de forma rápida para sobreviver [3].
- Resultados mais rápidos pelo sistema de percepção humano;
- Fornecem uma opção de representação de informações textuais ou verbais;

## Linguagens Específicas de Domínio (DSL) - Definição

- Abordam um domínio de aplicação específico [4].
  - Não obrigam o conhecimento avançado de programação.
  - Interface próxima ao domínio.
- 1 Externa
  - 2 Interna
  - 3 Workbench

## Trabalhos Relacionados

- 2008 - Superconductor [5]
- 2011 - Shadie [6]
- 2012 - Diderot [7]
- 2013 - Vivaldi [8]
- 2014 - ViSlang [9]

## Comparação

- 1 Interface;
- 2 Tipo de DSL;
- 3 Tipo de Visualização Gerada;
- 4 Arquitetura;

	1	2	3	4
<b>Vivaldi</b>	Alto Nível	Interna	Volumétricas	CPU e GPUs Distribuídas
<b>ViSlang</b>	Alto Nível	Interna	Volumétricas	CPU e GPU
<b>Diderot</b>	Alto Nível	Interna	Volumétricas	CPU
<b>Shadie</b>	Alto Nível	Interna	Volumétricas	CPU e GPU
<b>Superconductor</b>	Alto Nível	Interna	Interativas	CPU e GPU

## Justificativa

- Necessidade de programação para gerar visualizações.
- Paralelização do processamento dos dados.
- Possibilidade de oferecer uma interface de alto nível para pessoas com pouco conhecimento em programação.
- Falta de uma DSL que forneça a geração automática de visualizações e um meio de processamento dos dados de maneira fácil para os usuários manipular.



## Objetivo Geral

### Objetivo

Criar um protótipo da DSL utilizando o tipo de visualização de mapas e um interpretador que crie as geovisualizações a partir de quantidades massivas de dados.

- A DSL será criada com generalizações que permitirão no futuro inserir outros tipos de visualizações também usando *Big Data*.

# Interface de Programação

```
create map

  latitude: field 4
  longitude: field 2
  markerText: field 5 + field 1 + "photo."

  pageTitle: "Flickr MAP - Canon Camera"
  pageSize: 300px;

data:

  file: "data1.csv"
  file: "data2.csv"
  file: "data3.csv"

  delimiter: ','

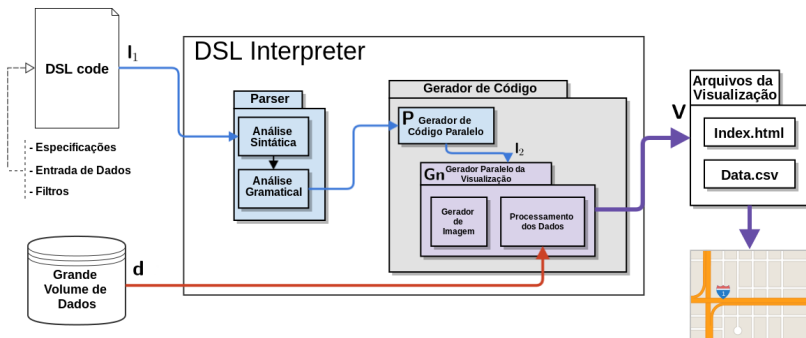
  filter: field 10 is greater than 3 AND field 1 is equal to "CANON"

  classify: by field 2

end

|
```

# Fluxo Geral



**Figura:** Fluxo de Trabalho da Proposta

## Processamento de Dados

- Seleção de algoritmos para seleção dos dados.
  - Ex: Busca Linear, Busca Binária,...
- Análise
- Escolha e Justificativa
- Paralelização usando DSL-POPP [10] [11]:
  - Multicore
  - Padrões de programação paralela (Master-Slave e Pipeline)
  - Desenvolvido no grupo de pesquisa

## Conclusão

- Benefícios:
  - Linguagem intermediária e focada no domínio
  - Paralelismo totalmente abstraído para o usuário
  - Não é vinculada a uma linguagem hospedeira
- Limitações:
  - Expressividade limitada
  - Aprendizado de uma nova linguagem
- Trabalhos Futuros:
  - Definição da Interface
  - Desenvolvimento do Gerador de Visualizações
  - Desenvolvimento do Processador de Dados
  - Desenvolvimento do Parser
  - Testes

## Referências I

- [1] Jinsong Zhang, Yan Chen, and Taoying Li. Opportunities of innovation under challenges of big data. In *Fuzzy Systems and Knowledge Discovery (FSKD), 2013 10th International Conference on*, pages 669–673. IEEE, 2013.
- [2] M Ghanbari. Visualization overview. In *System Theory, 2007. SSST'07. Thirty-Ninth Southeastern Symposium on*, pages 115–119. IEEE, 2007.
- [3] Matthew Ward, Georges Grinstein, and Daniel Keim. *Interactive data visualization: foundations, techniques, and applications*. AK Peters, Ltd., 2010.
- [4] Marjan Mernik, Jan Heering, and Anthony M Sloane. When and how to develop domain-specific languages. *ACM computing surveys (CSUR)*, 37(4):316–344, 2005.
- [5] Leo A Meyerovich, Matthew E Torok, Eric Atkinson, and Rastislav Bodik. Superconductor: A language for big data visualization, 2013.
- [6] J. Wolfgang G. Chen HASAN, M. and H. Pfister. Shadie: A domain-specific language for volume visualization. 2010.
- [7] Charisee Chiw, Gordon Kindlmann, John Reppy, Lamont Samuels, and Nick Seltzer. Diderot: a parallel dsl for image analysis and visualization. In *ACM SIGPLAN Notices*, volume 47, pages 111–120. ACM, 2012.
- [8] Hyungsuk Choi, Woohyuk Choi, Tran Quan, David GC Hildebrand, Hanspeter Pfister, and Won-Ki Jeong. Vivaldi: A domain-specific language for volume processing and visualization on distributed heterogeneous systems.
- [9] Peter Rautek, Stefan Bruckner, ME Groller, and Markus Hadwiger. Vislang: A system for interpreted domain-specific languages for scientific visualization. 2014.
- [10] Dalvan Griebler, Daniel Adornes, and Luiz G. Fernandes. Performance and Usability Evaluation of a Pattern-Oriented Parallel Programming Interface for Multi-Core Architectures. In *The 26th International Conference on Software Engineering & Knowledge Engineering*, pages 25–30, Vancouver, Canada, July 2014. Knowledge Systems Institute Graduate School.

## Referências II

- [11] Dalvan Griebler and Luiz G. Fernandes. Towards a Domain-Specific Language for Patterns-Oriented Parallel Programming. In *Programming Languages - 17th Brazilian Symposium - SBLP*, volume 8129 of *Lecture Notes in Computer Science*, pages 105–119, Brasilia, Brazil, October 2013. Springer Berlin Heidelberg.

[◀ Voltar para Capa](#)

# Dúvidas

