

# EXECUÇÃO DE WORKFLOWS CIENTÍFICOS EM NUVENS COMPUTACIONAIS

Elaine Naomi Watanabe

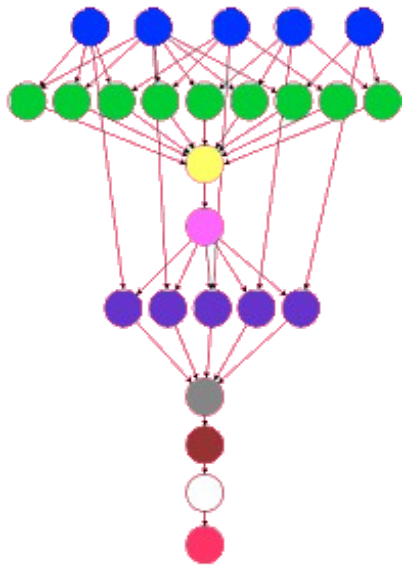
Instituto de Matemática e Estatística

[elainew@ime.usp.br](mailto:elainew@ime.usp.br)

# Workflows científicos

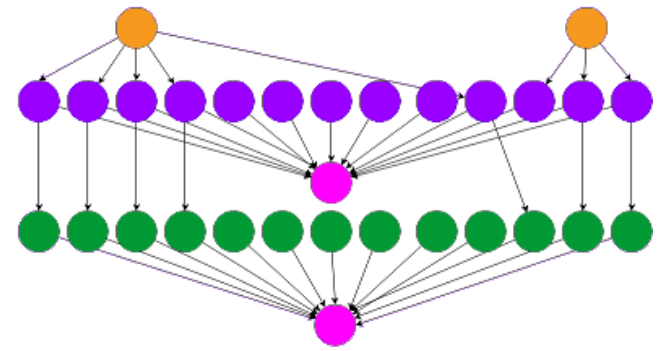
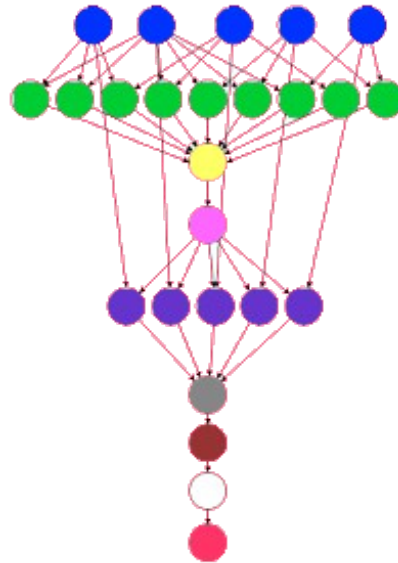
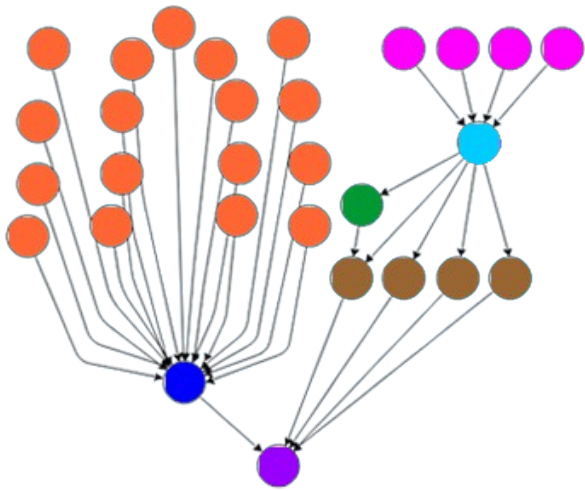
Um **workflow científico** é a automação de experimentos ou processos científicos, expressa em termos das atividades a serem executadas e das dependências dos dados manipulados.

# Workflows científicos

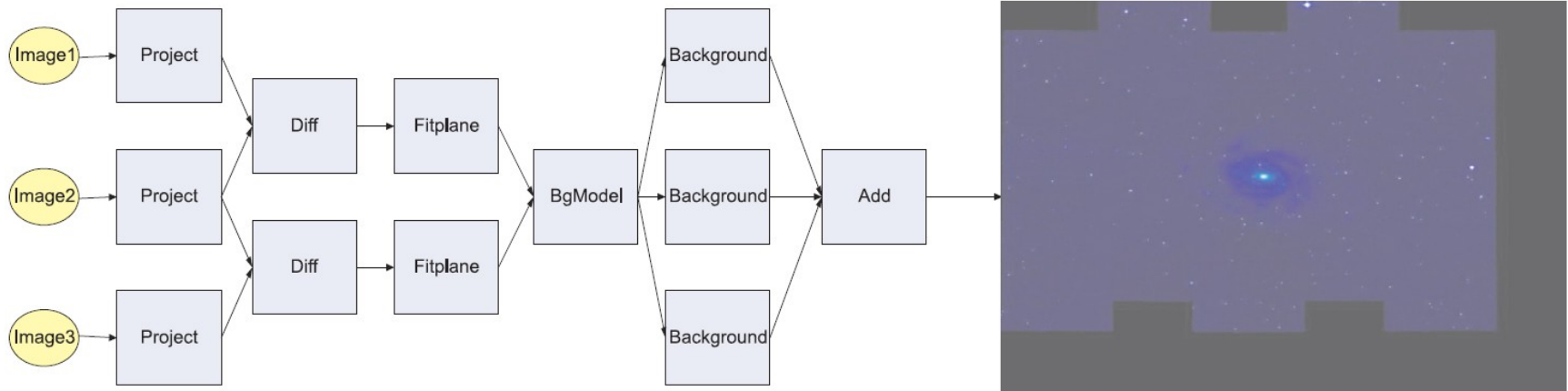


- Conjunto de **tarefas interligadas**
  - Aquisição
  - Processamento
  - Análise
  - Visualização
  - etc
- Grafo direcionado acíclico (DAG)
- **Grandes quantidades de dados**

# Workflows científicos



# Workflows científicos



Modelo do workflow científico Montage.

Fonte: Ewa Deelman. Grids and Clouds: Making workflows applications work in heterogeneous distributed environments, 2010.

# Aplicações

- Astronomia, medicina, bioinformática, visão computacional, geofísica, neurociência, etc.
- Por exemplo...
  - A fase piloto do projeto 1000 genomes, comparando o genoma de 629 indivíduos, gerou:  
7,3 terabytes =  $7,3 \times 10^{12}$  bytes

## **Sistema Gerenciador de Workflows Científicos (SGWfC):**

É um sistema que permite a descrição de um experimento científico como um conjunto de tarefas a serem executadas de forma automática, utilizando, para isso, a infraestrutura computacional disponível.

# Exemplos de SGWfC

Kepler *<http://kepler-project.org/>*

Taverna *<http://www.taverna.org.uk/>*

Askalon *<http://www.askalon.org/>*

Pegasus *<http://pegasus.isi.edu/>*

SciCumulus *<http://sf.net/projects/scicumulus>*



# Exemplos de Simuladores de Workflows Científicos para Nuvem

- WorkflowSim
- SimpleWorkflow
- DynamicCloudSim
- CloudSim-DVFS

Fonte: <http://www.cloudbus.org/cloudsim/>

Necessidade de um ambiente de **alto**  
**desempenho...**



- A **computação em nuvem** tem se mostrado uma boa alternativa:
  - Com **investimento menor** que o tradicional
  - Com possibilidade de **alocação sob demanda**

## **Escalonamento:**

seleção do recurso e do momento de execução das atividades, otimizando objetivos como o tempo total de execução.

# Motivação

- **Escalonamento** para execução em nuvem:
  - considerar **informações** como restrições de tempo e orçamento.
- **Sistemas Gerenciadores de Workflows Científicos** atuais:
  - não exploram completamente as nuvens computacionais
  - linguagens de modelagem possuem pouca informação de apoio ao escalonamento.
- **Informações de apoio ao escalonamento:**
  - podem ser inseridas no modelo de workflow

# Trabalhos em andamento

- **Modelagem e escalonamento** de workflows científicos em plataformas de computação em nuvem
- **Provisionamento dinâmico** de recursos para execução de workflows científicos em nuvens
- **Gerenciamento de dados** de workflows científicos em nuvens computacionais

# Trabalho concluído

**Economia de energia na execução de workflows científicos em nuvens computacionais:**

- Novos escalonadores de tarefas
- Simulador CloudSim-DVFS (licença LGPL)
- Conseguiu economia de energia na maioria dos casos, chegando até 22,7% [2].

Esse trabalho foi fruto da colaboração de pesquisadores da comunidade científica nacional e internacional



# Software livre e a colaboração da comunidade científica

- Entre a comunidade do IME-USP
  - Pedro Campos (graduação)
  - Elaine Watanabe (mestrado)
  - Prof. Dr. Daniel Batista
    - Professor e Pesquisador da área de Redes
  - Prof<sup>a</sup> Dr<sup>a</sup> Kelly Braghetto
    - Professora e Pesquisadora da área de Banco de dados
  - Dr. Daniel Cordeiro
    - Pesquisador da área de Computação de Alto Desempenho e Teoria de Escalonamento

# Software livre e a colaboração da comunidade científica

- Entre a comunidade internacional
  - University of Melbourne (Austrália)
    - Rodrigo Calheiros (CloudSim-DVFS)
  - Université de Toulouse (França)
    - Tom Guérout (CloudSim-DVFS)
  - University of Southern California (EUA)
    - Weiwei Chen (WorkflowSim)

Grupos de pesquisam  
sobre workflows científicos...

# Alguns grupos que pesquisam sobre Workflows Científicos

- No Brasil
  - Universidade Federal do Rio de Janeiro (PESC-COPPE-UFRJ), em parceria com o Laboratório Nacional de Computação Científica (LNCC)
    - Criadores das ferramentas Chiron e SciCumulus
  - Universidade Estadual de Campinas (IC-UNICAMP)
  - Universidade de São Paulo (DCC-IME-USP)

# Alguns grupos que pesquisam sobre Workflows Científicos

- Fora do Brasil
  - University of Southern California (Information Sciences Institute)
    - Profa. Ewa Deelman
      - Criadores do Pegasus
  - University of Manchester (School of Computer Science)
    - Profa. Carole Goble
      - Criadores do Taverna, myGrid, myExperiment
  - University of Chicago (Computation Institute)
    - Prof. Ian Foster
      - Criadores do conceito de grade e do projeto Globus
  - University of Innsbruck (Institute of Computer Science)
    - Prof. Thomas Fahringer
      - Criadores do ASKALON

# Resumo

- Sistemas gerenciadores de workflows científicos são importantes para o desenvolvimento da ciência.
- Existe uma carência de escalonadores de workflows em nuvens computacionais em tais gerenciadores que explorem melhor os recursos e ambiente disponível.
- Software livre permite maior cooperação entre comunidades científicas, além de permitir reprodutibilidade, validação e novos experimentos.

MUITO OBRIGADA!

Elaine Watanabe<sup>1</sup> ([elainew@ime.usp.br](mailto:elainew@ime.usp.br))

Ricardo Oda<sup>2</sup> ([odaric@ime.usp.br](mailto:odaric@ime.usp.br))

Eduardo Cotrin<sup>2</sup> ([cotrin@ime.usp.br](mailto:cotrin@ime.usp.br))

Agências Financiadoras: CAPES<sup>1</sup>, CNPq<sup>2</sup>

Orientadora: Prof<sup>a</sup> Dr<sup>a</sup> Kelly R. Braghetto

([kellyrb@ime.usp.br](mailto:kellyrb@ime.usp.br))

# Referências

- [1] BRAGHETTO, K. R. ; CORDEIRO, D. . Introdução à Modelagem e Execução de Workflows Científicos. In: Salgado, Ana Carolina; Lóscio, Bernadette Farias; Alchieri, Eduardo; Barreto, Priscila Solís. (Org.). Atualizações em Informática. 1ed.Porto Alegre: SBC, 2014, v. , p.1-40.
- [2] WATANABE, E. N. et al . Algoritmos para economia de energia no escalonamento de workflows em nuvens computacionais. Em: Anais do 32º Simpósio Brasileiro de Redes de Computadores e Sistemas Distribuídos SBRC, 2014. Porto Alegre: Sociedade Brasileira de Computação, 2014. p. 31-44