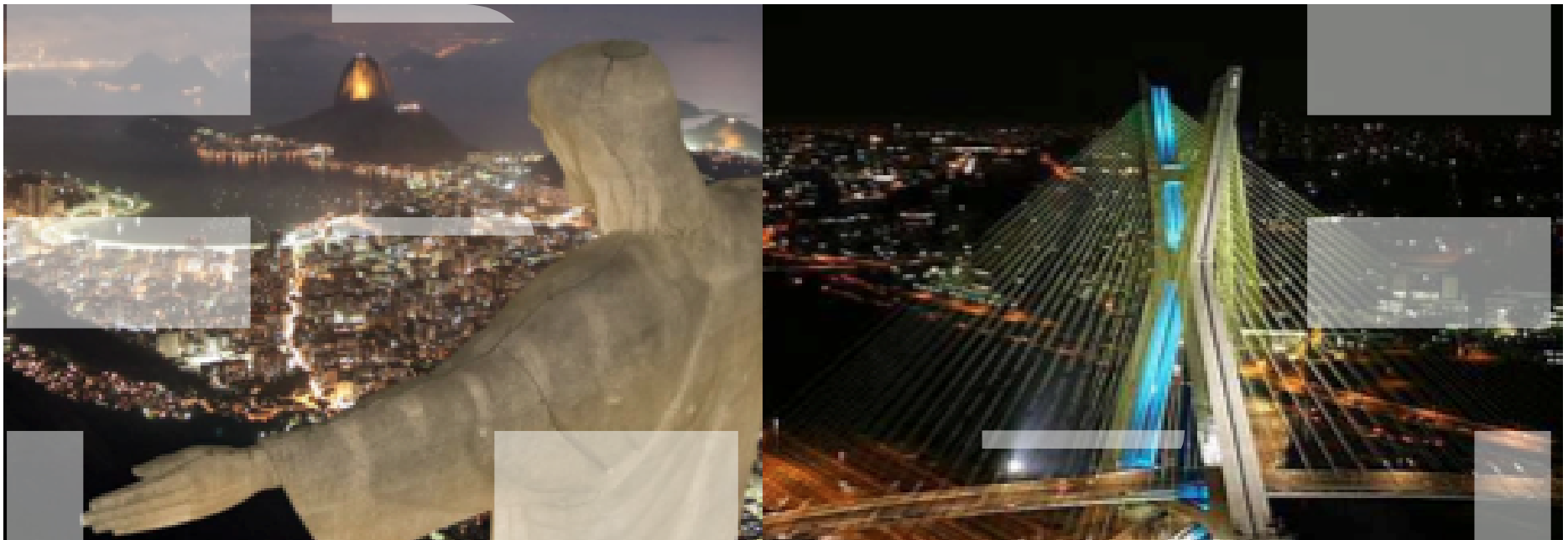




Quando algoritmos e força bruta se encontram: processamento de alto desempenho na IBM Research Brasil

Lucas C. Villa Real
Research Software Engineer
lucasvr@br.ibm.com

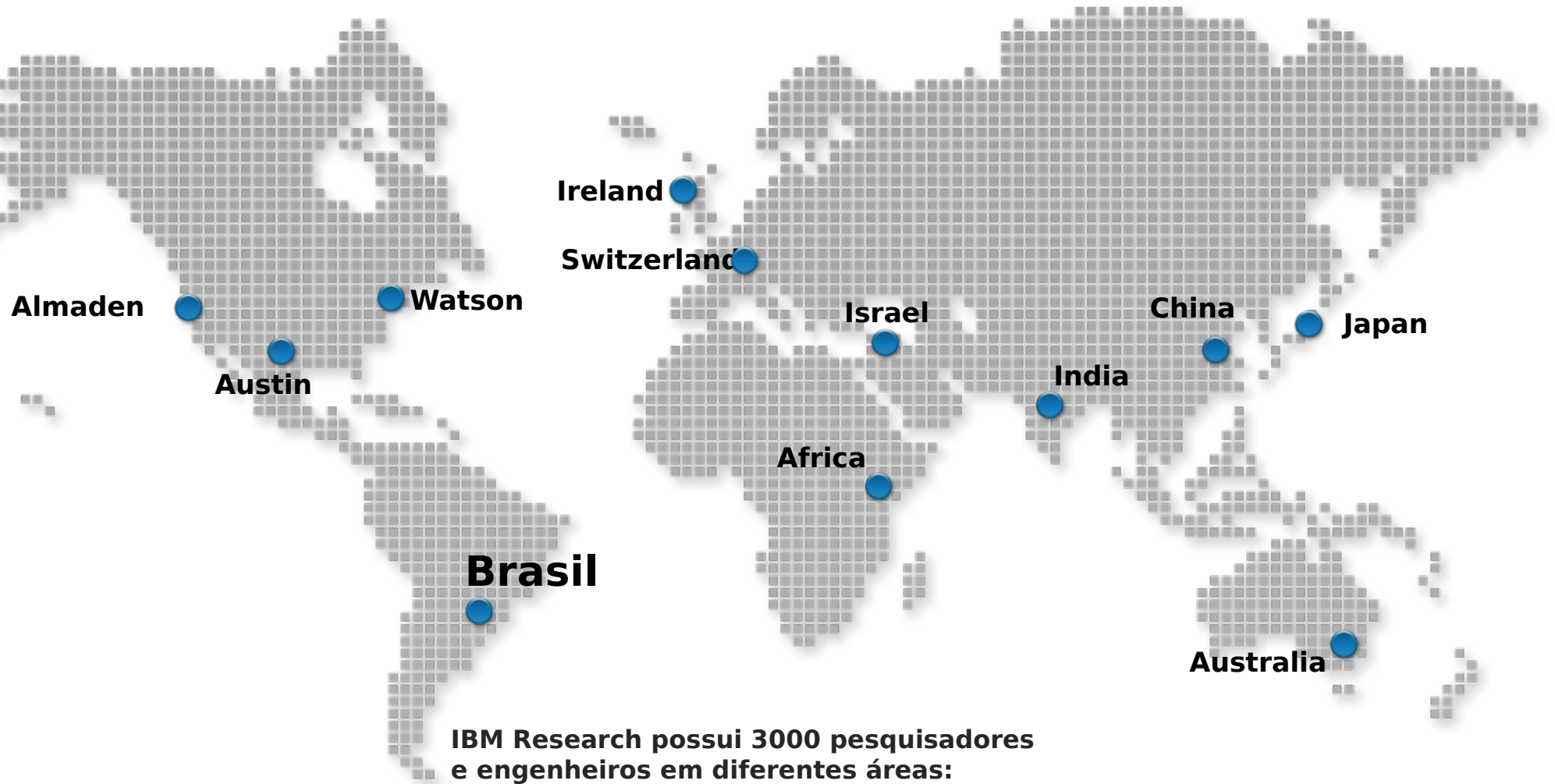


Agenda

- IBM Research – Brasil
- Problemas de larga escala
 - Simulação de alagamentos
 - Dados geo-espaciais de alta resolução
 - Formatos de arquivos e I/O
- Consórcio OpenPOWER
- Considerações finais

Nossos laboratórios no mundo

12 Labs em 10 Países



Ciência do
Comportamento



Química



Ciência da
Computação



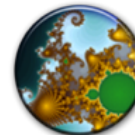
Engenharia
Elétrica



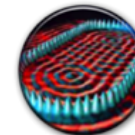
Ciência de
Materiais



Ciência da
Matemática



Física



Ciência de
Serviços



IBM Research – Brasil

Áreas de Foco

Soluções em
Recursos
Naturais

Óleo e Gás,
Mineração,
Agricultura,
Logística e
Sustentabilidade

Sistemas de
Engajamento

Cidades Inteligentes,
Engajamentos dos
Cidadãos, Integração com
Comunidades e Educação

Analíticos de
Dados Sociais

Otimização dos serviços
em larga escala,
otimização e integração
no contexto empresas
sociais.

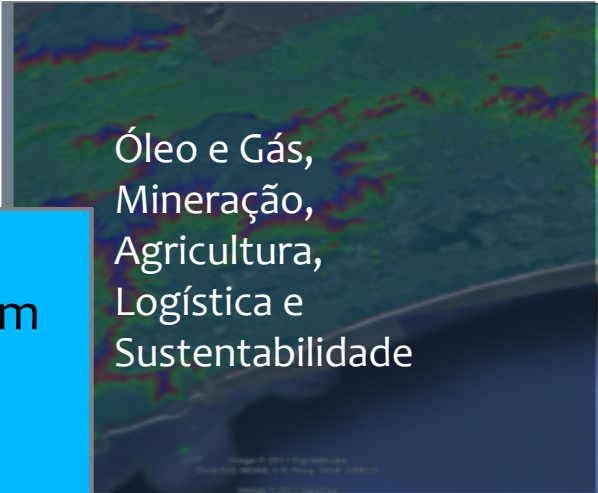
Dispositivos
Inteligentes

Micro e nano
tecnologias e matérias-
primas com o objetivo
de resolver grandes
desafios.

IBM Research – Brasil

Áreas de Foco

Soluções em
Recursos
Naturais



Óleo e Gás,
Mineração,
Agricultura,
Logística e
Sustentabilidade

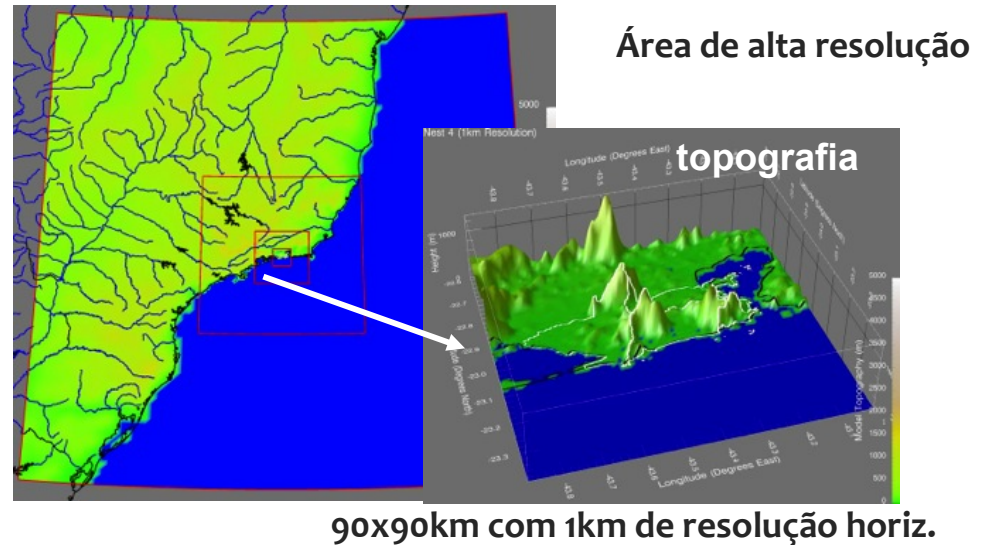
Disciplinas de Interesse:

- Matemática Aplicada
- Ciência da Computação
- Computação de Alto Desempenho
- Visualização de Dados

Problemas de larga escala

Gerenciamento de Incidentes

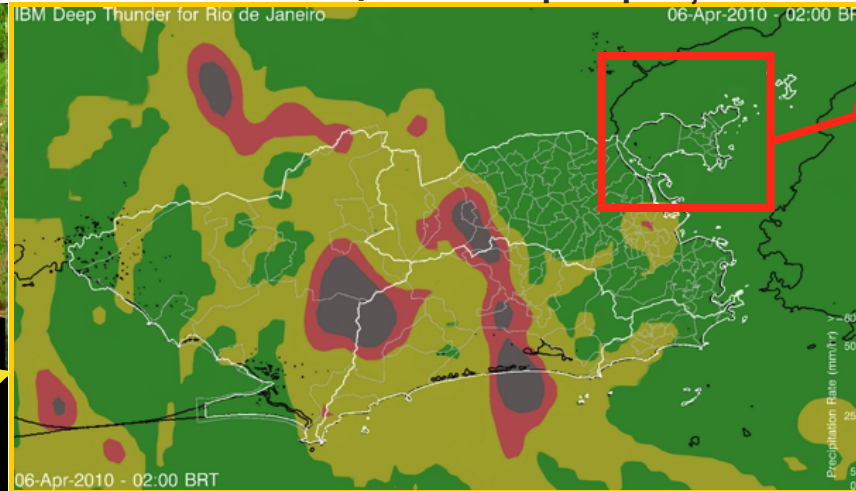
Tecnologias de pesquisa em tempo e cheias com granulação fina e modelagem de previsões



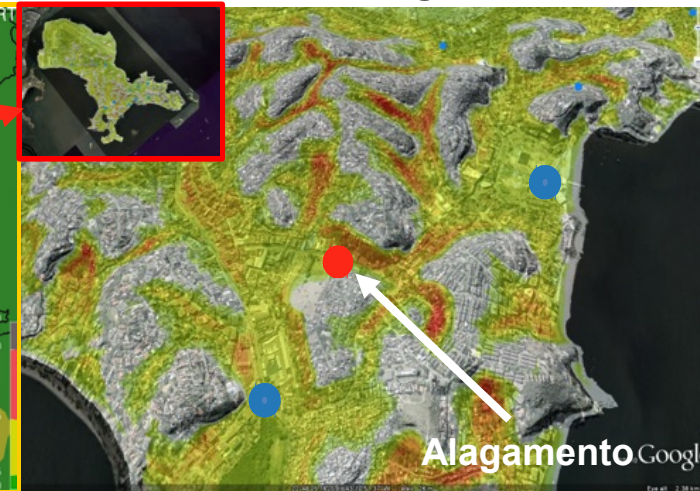
Cidade do Rio de Janeiro



Previsão de 48 horas de precipitação



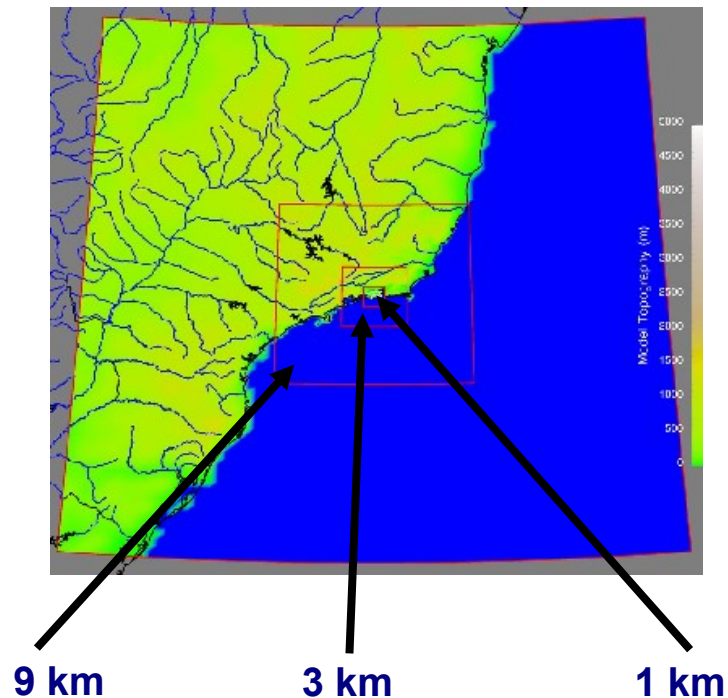
Previsão de alagamentos



Previsão Meteorológica Numérica

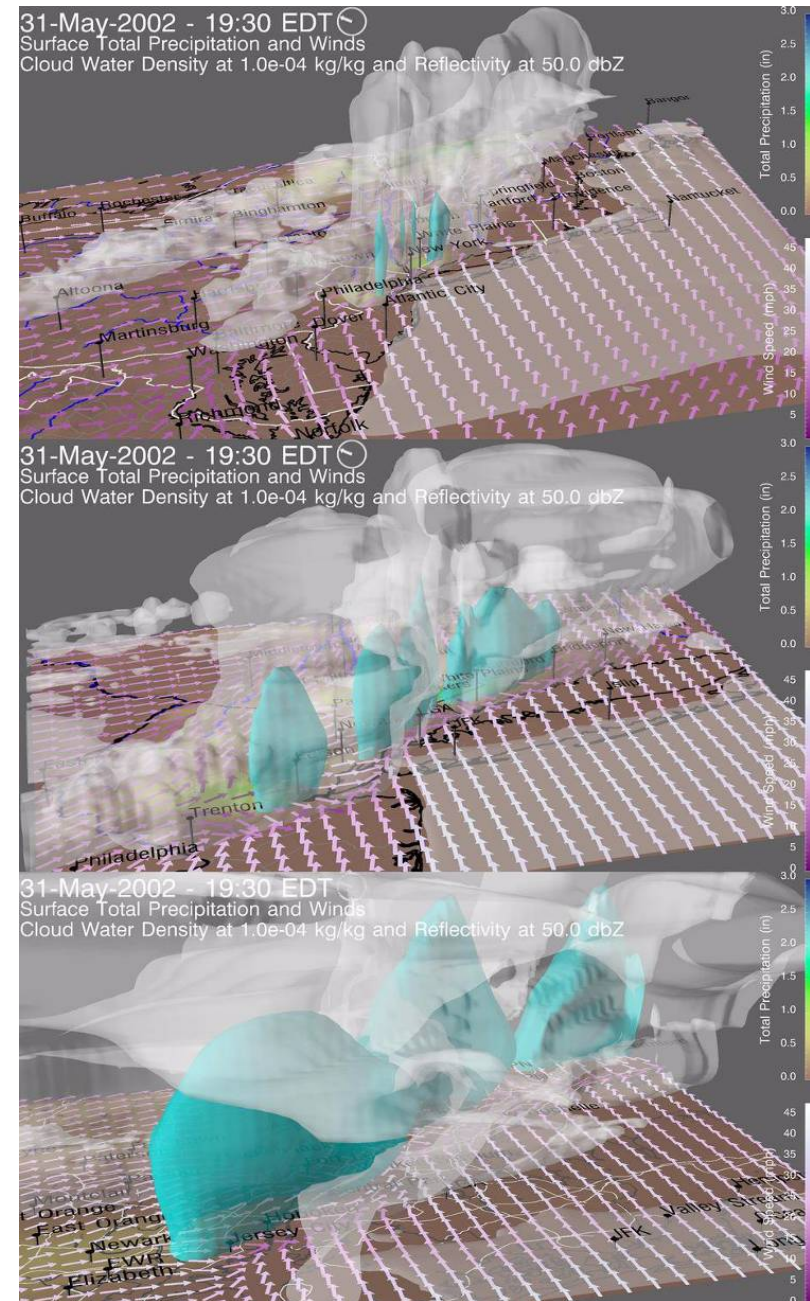
IBM Deep Thunder @ Rio:

- 65 camadas verticais
- Topografia: 90 metros
- WRF: ~4 horas*



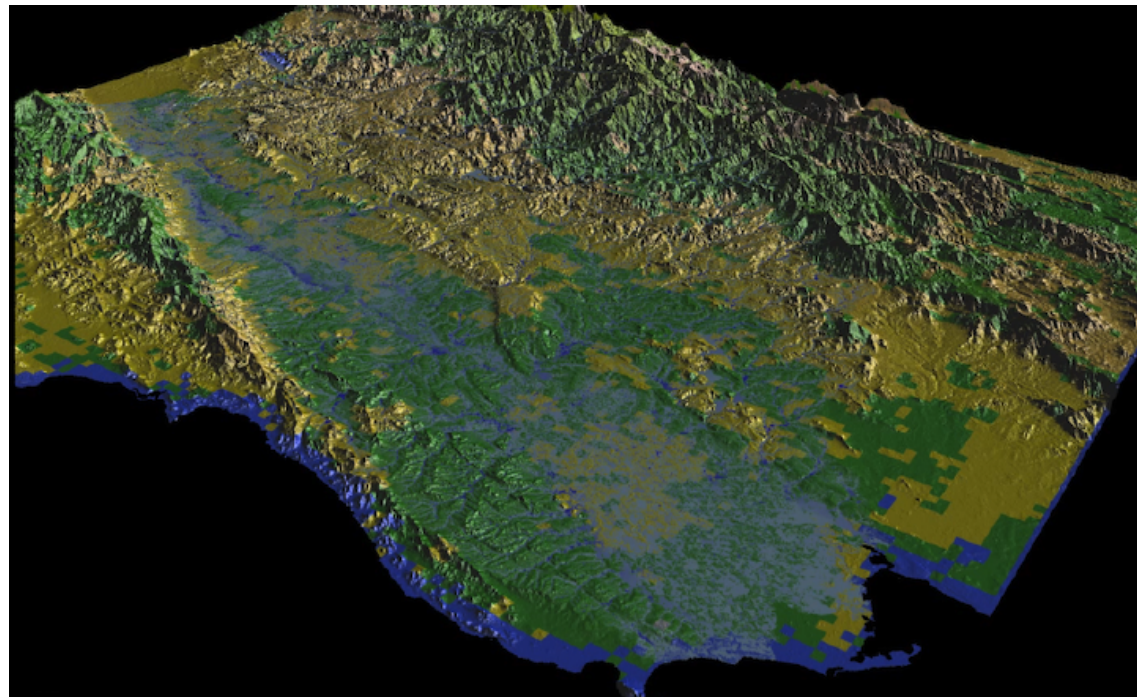
* 24, 8-way blades Power 6

192 nós de processamento + 1 nó para I/O

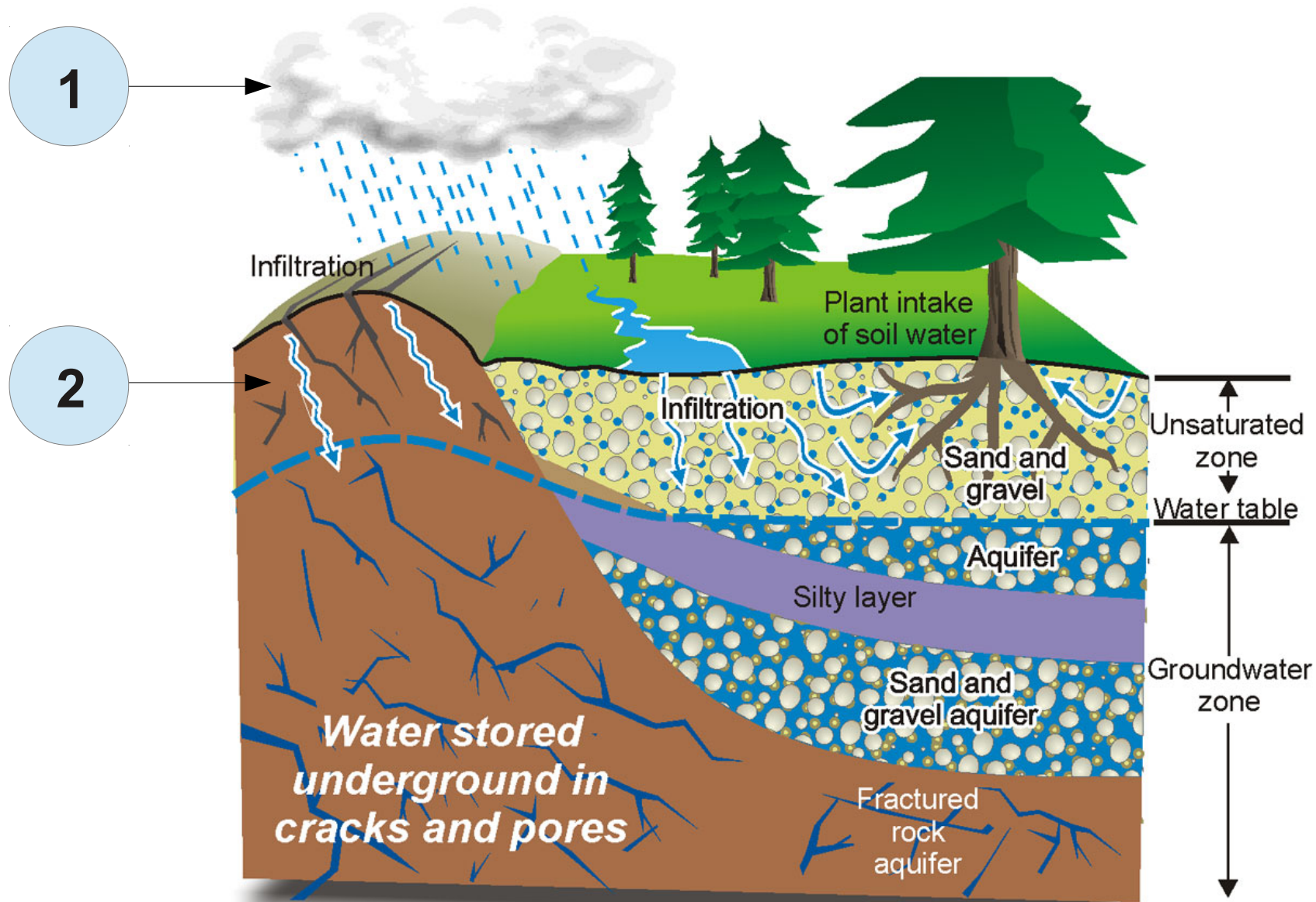


Simulação de alagamentos

- **Precipitação** é mapeada para os **dados de elevação**
- **Vegetação** potencialmente intercepta parte da chuva
- O que não é interceptado pode infiltrar no **solo**
- Água que permanece na superfície é **roteada**

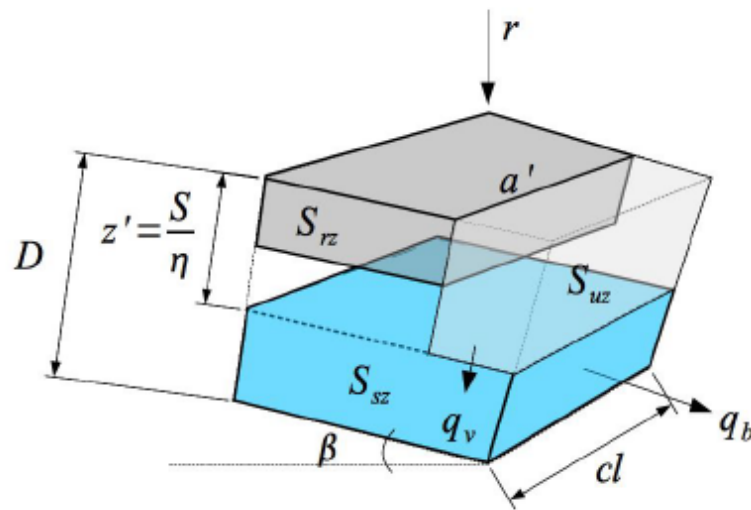


Simulação de alagamentos

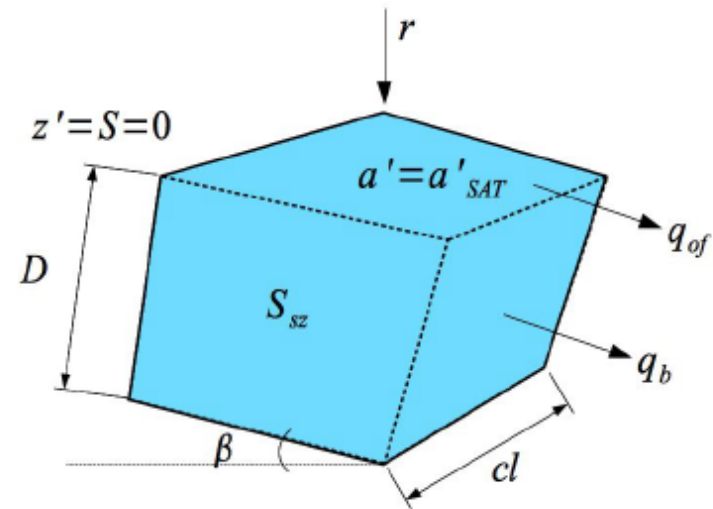


Simulação de alagamentos

Modelo de solo



- S_{RZ} : Root zone store
- S_{UZ} : Non-saturated zone store
- S_{SZ} : Saturated zone store expressed in deficit (S_i) or in depth (Z'_i)
- q_b : Sub-surface runoff in the saturated zone
- q_v : Sub-surface vertical runoff in the non-saturated zone
- q_{ofj} : Surface runoff on saturation



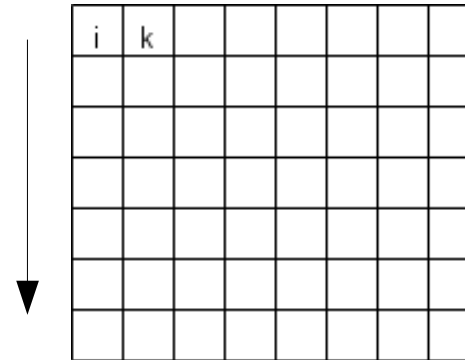
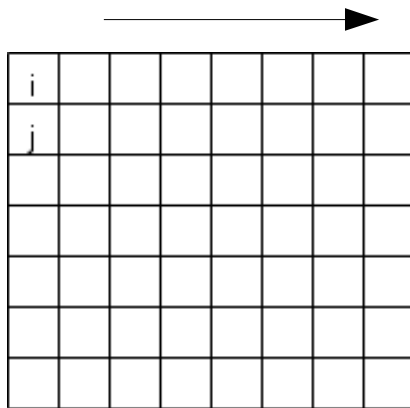
- r : rainfall
- a' : cell area
- b : cell declivity (slope angle)
- cl : width of the cell's contour
- h : soil porosity

Simulação de alagamentos

Equação de difusão da onda

H = altura da água
 S = altura da superfície
 Δd = tamanho da célula (ex.: 90, 30)

OLR = roteamento da água
 N = coeficiente de rugosidade



$$res = ((H[i] + S[i]) - (H[j] + S[j])) / \Delta d$$

if $res < zero$:

$$OLR[i] = \frac{\sqrt{|res|}}{N[j]} * \Delta d * H[j]^k$$

else :

$$OLR[j] = \frac{\sqrt{|res|}}{N[i]} * \Delta d * H[i]^k$$

$$res = ((H[i] + S[i]) - (H[k] + S[k])) / \Delta d$$

if $res < zero$:

$$OLR[i] = \frac{\sqrt{|res|}}{N[j]} * \Delta d * H[j]^k$$

else :

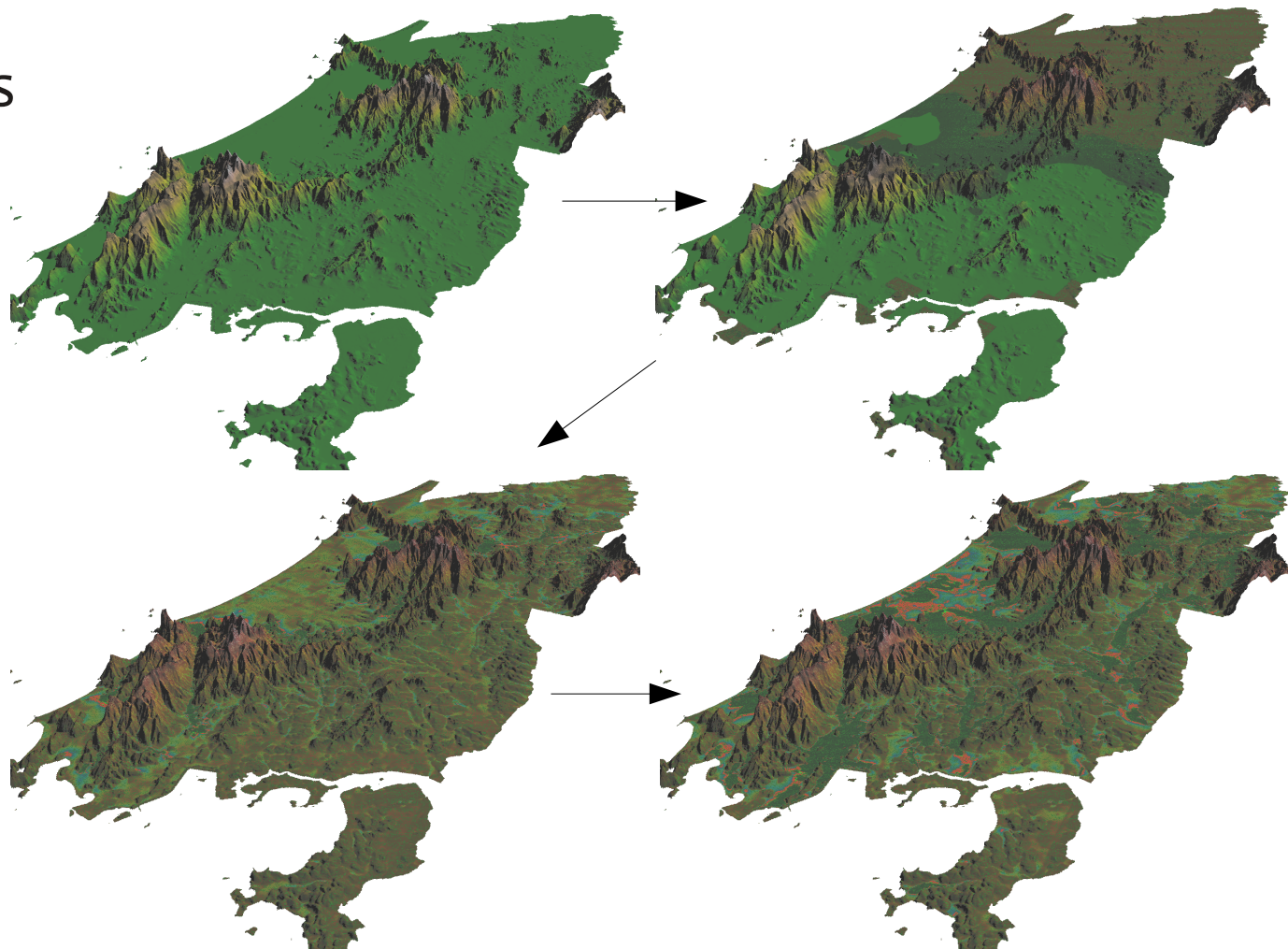
$$OLR[k] = \frac{\sqrt{|res|}}{N[k]} * \Delta d * H[k]^k$$

Simulação de alagamentos

Resultado esperado

Problema:

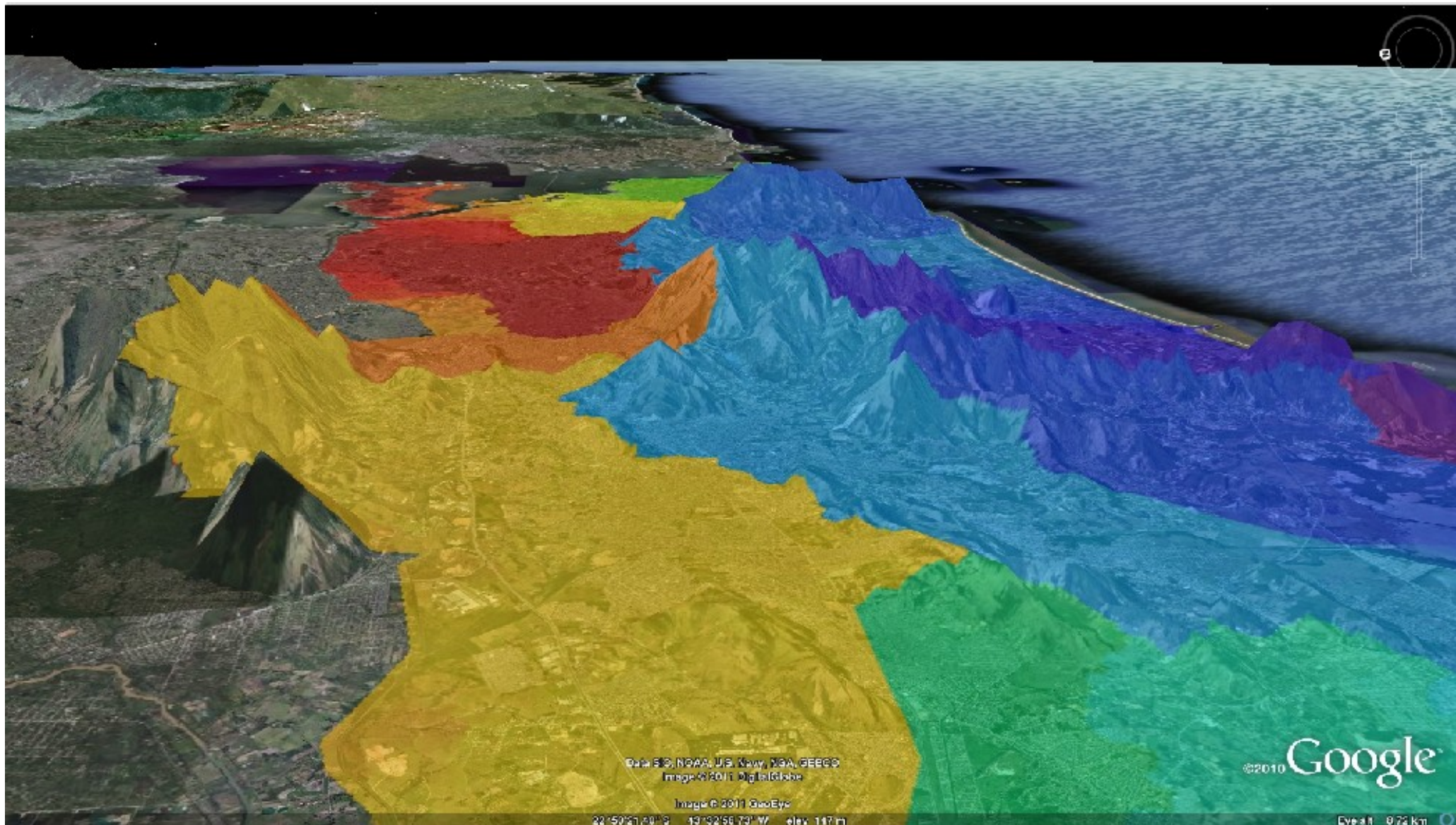
- 4-6 horas de processamento*
- 2.6 bilhões de células



* 24 Xeon E5645 cores @ 2.40GHz
128GB RAM
Multi-Threaded

Simulação de alagamentos

Paralelismo



Simulação de alagamentos

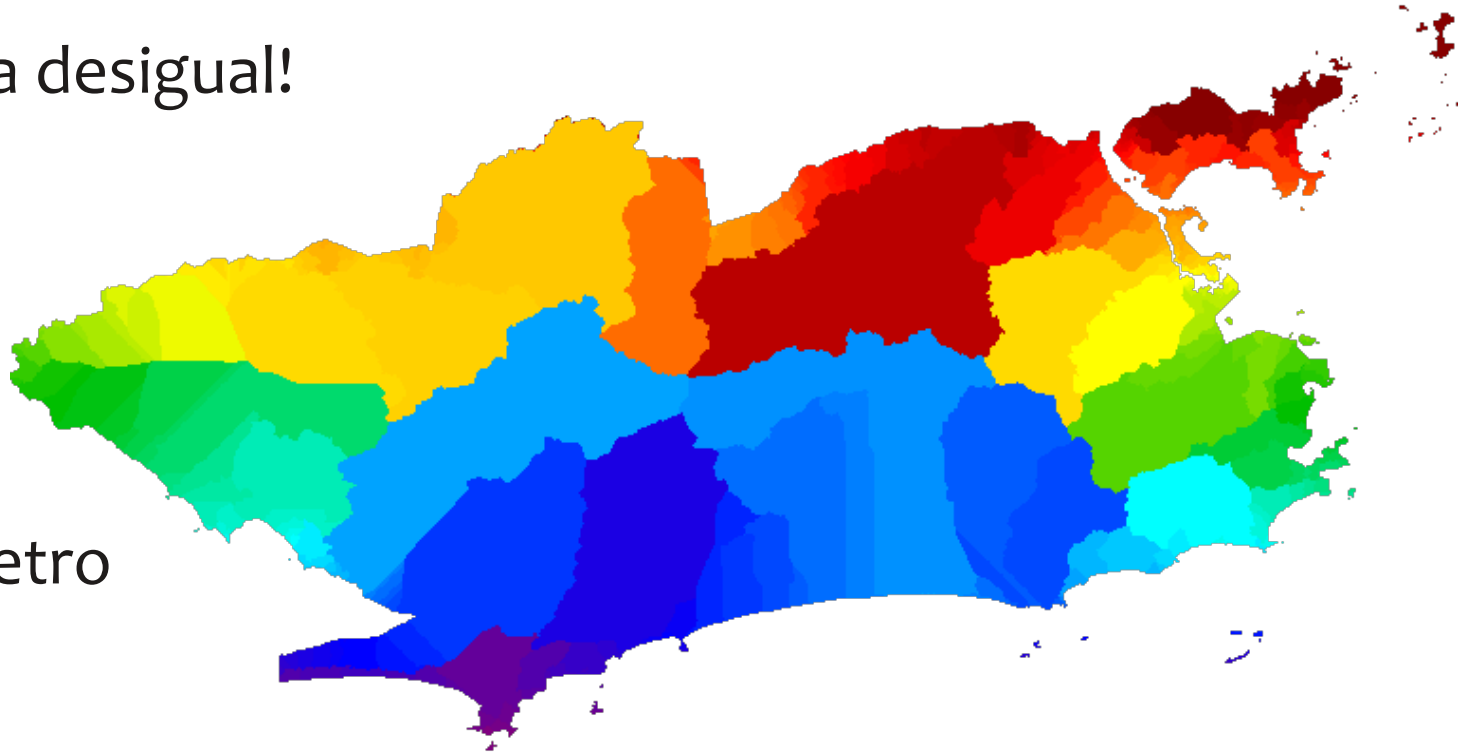
Paralelismo

Bacias hidrográficas

- Determinam um primeiro nível de paralelismo
- Problema: carga desigual!

Rio:

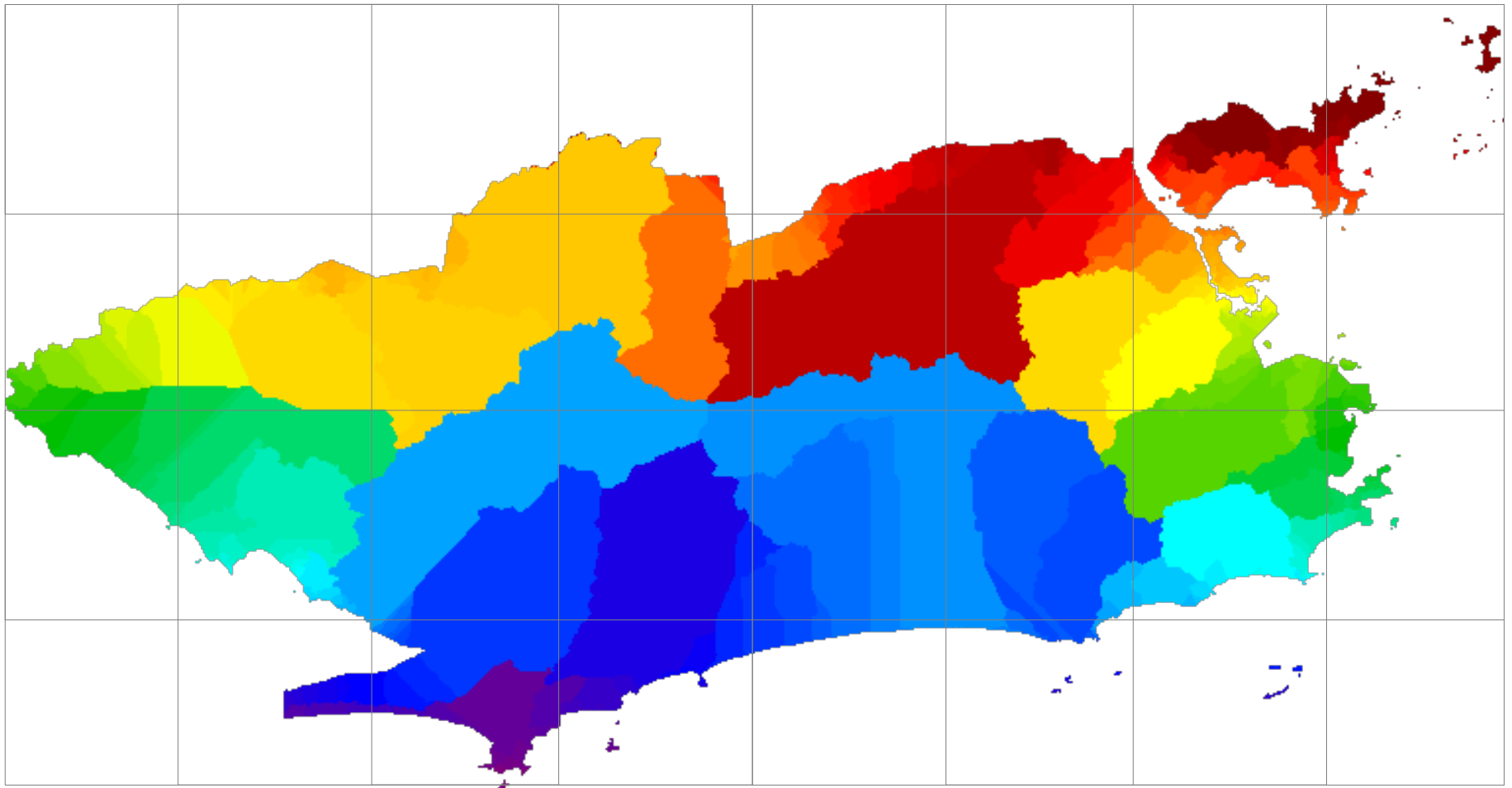
- 70k bacias
- Topografia: 1 metro



Simulação de alagamentos

Paralelismo

Decomposição de domínio (intra-bacias ou no domínio inteiro)



Simulação de alagamentos

Paralelismo

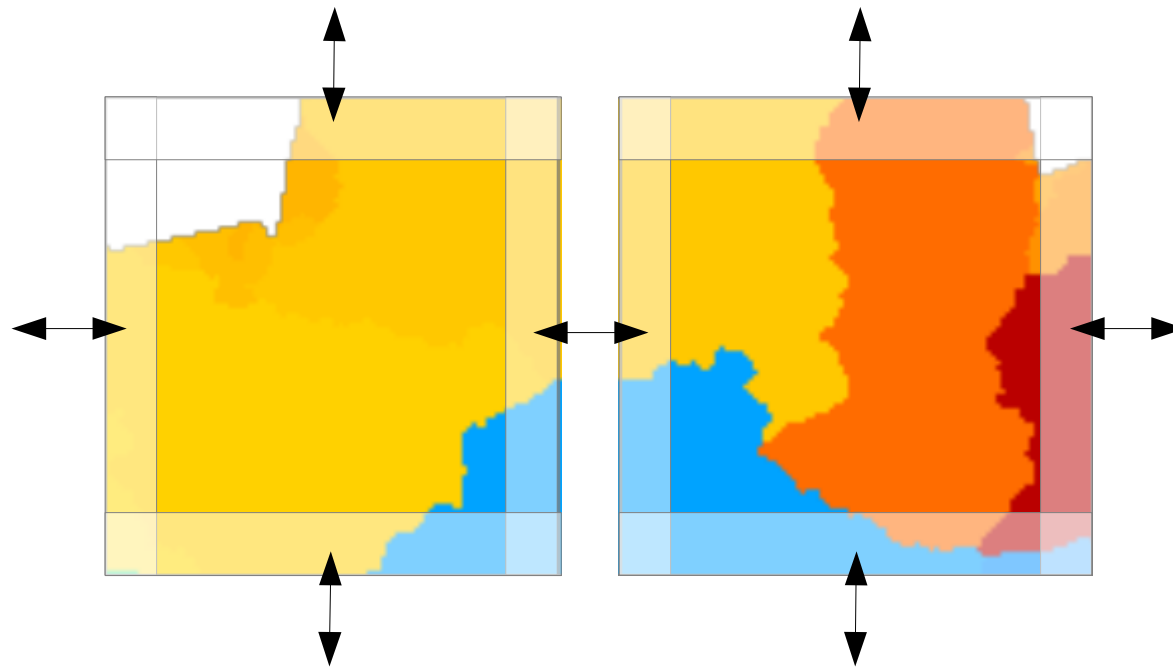
Decomposição de domínio (intra-bacias ou no domínio inteiro)



Simulação de alagamentos

Paralelismo

Decomposição de domínio (intra-bacias ou no domínio inteiro)

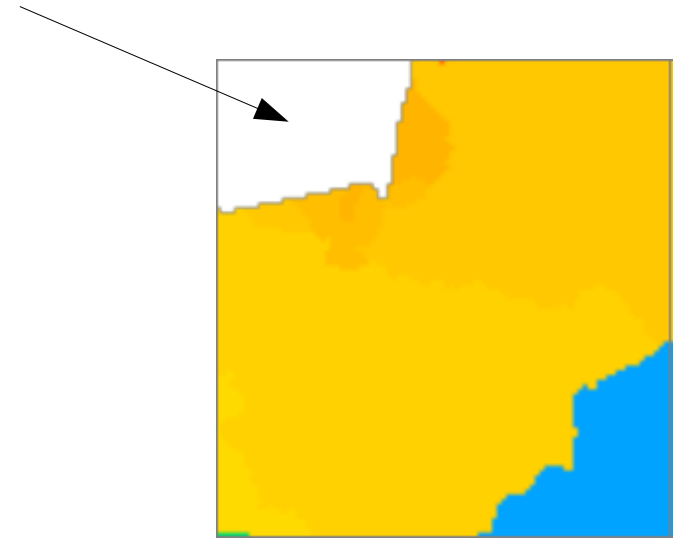


Simulação de alagamentos

Paralelismo

Tamanho dos quadrantes é difícil de ser estipulado

- Precisa-se levar em consideração número de células válidas
 - Oceano, lagoas, rios não são processados
 - Células secas
- Compreensão do problema é relevante
 - Água é transportada mais rapidamente em áreas de grande inclinação do que no plano
- Rio: quadrantes de tamanho variável: ~15 min* (MPI + OpenMP)



* IBM Blue Gene BG/P (4 PowerPC 450 cores @ 850MHz)

13.6 Gigafllops/node, 1GB memory/node

2048 nodes

Dados geo-espaciais de alta resolução

Dados geo-espaciais de alta resolução

LiDAR – Light Detection and Ranging

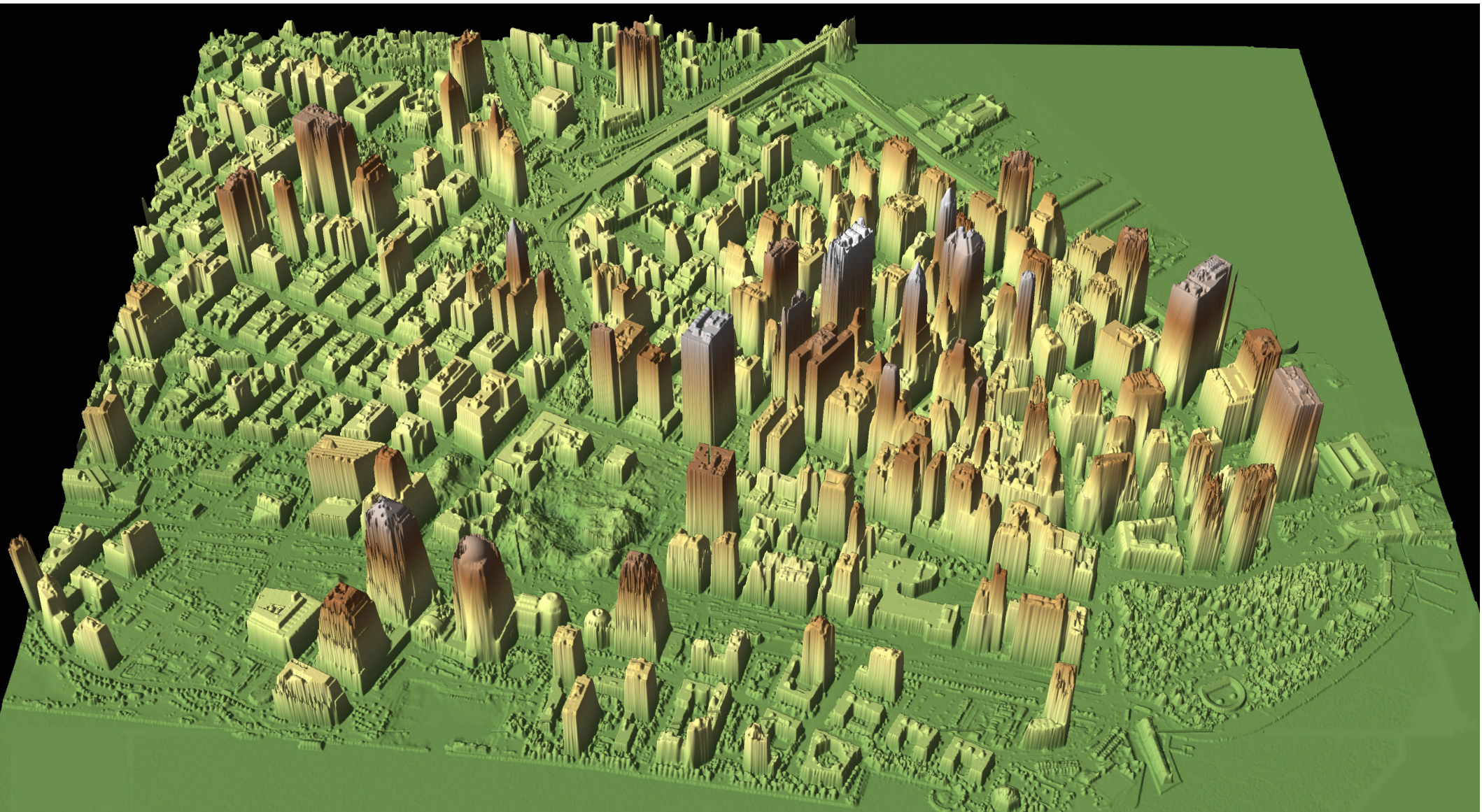
- Nuvem de pontos extremamente densa
- Múltiplos *retornos*
- Pontos podem ser *classificados*
- Pontos podem ter RGB
- Combinação com imagens
- multi/ultra-espectrais

Copyright © africansurveyors.com



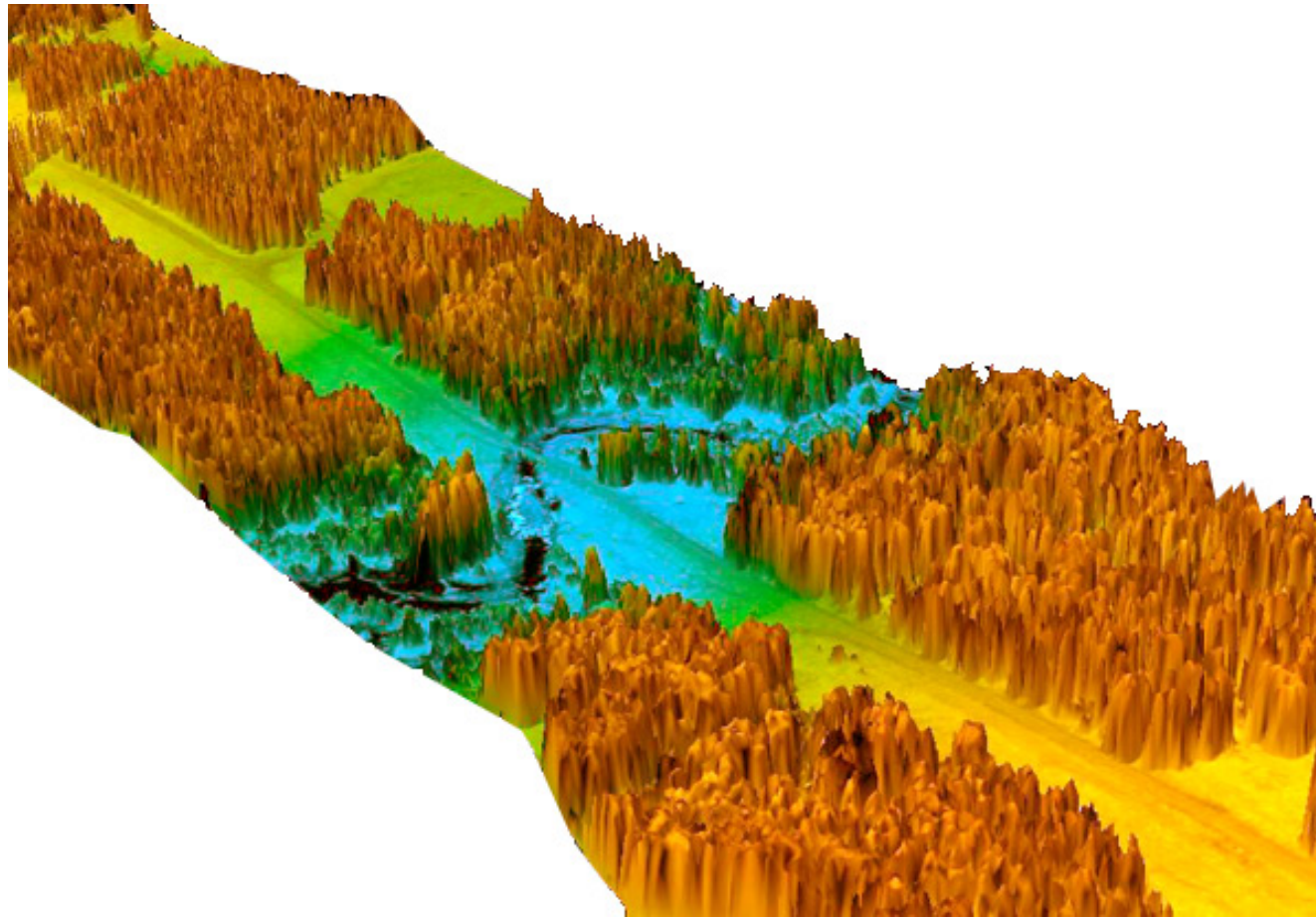
Dados geo-espaciais de alta resolução

LiDAR – Light Detection and Ranging



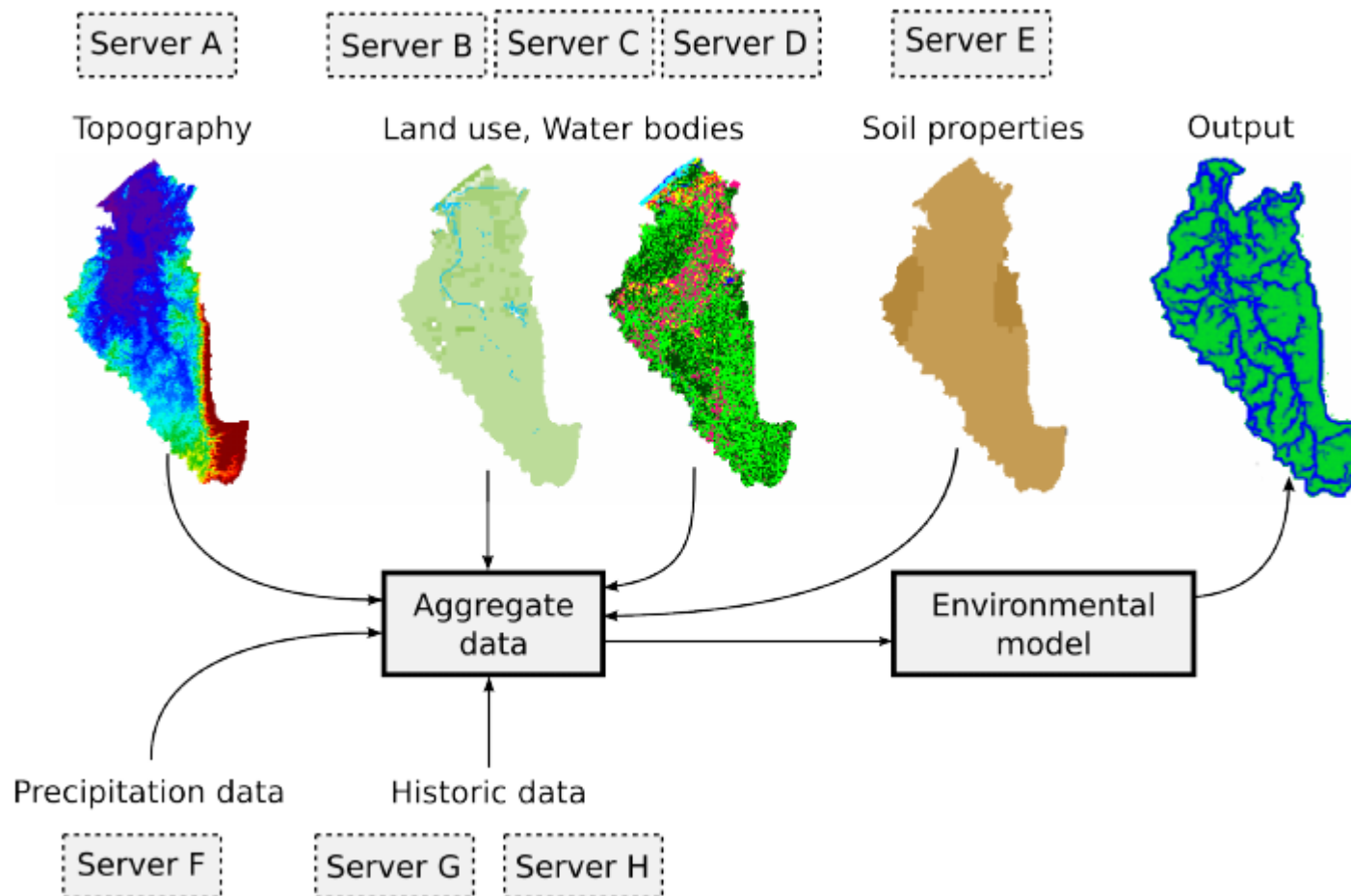
Dados geo-espaciais de alta resolução

LiDAR – Light Detection and Ranging



Copyright © gisbox.ro

Dados geo-espaciais de alta resolução



Dados geo-espaciais de alta resolução

Problema: como utilizar e visualizar dados dessa natureza?

Exemplo de um projeto:

- Raw LiDAR: 70 TB
- Dado pós-processado: 6TB

Possíveis soluções:

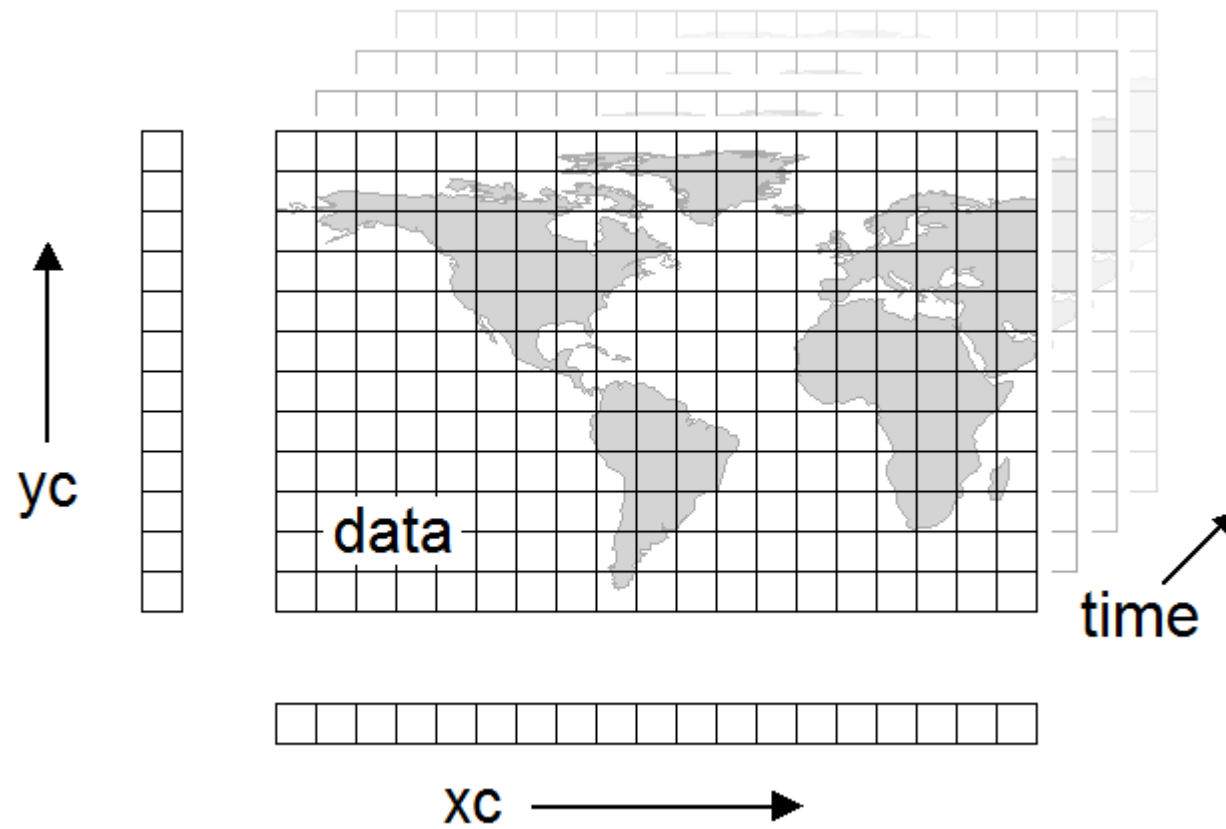
- Pirâmides com diferentes resoluções (refinamento dinâmico)
- I/O paralelo

Formatos de arquivos e I/O

NetCDF

- Formato de arquivos científico **auto-descritivo** orientado a arrays
 - Cabeçalho descreve o **layout** do arquivo
 - Cabeçalho inclui **meta-dados** na forma de nome=valor
- Múltiplas variáveis podem ser declaradas
 - Podem ter atributos (unidade, intervalo, escala, etc)
 - Podem compartilhar dimensões com outras variáveis
 - Podem ter N dimensões ou dimensões de tamanho ilimitado
 - Podem ser escalares, vetoriais, arrays
 - Devem ser tipadas: char, byte, short, int, float, double
- Formato independente de plataforma

NetCDF



Parallel-NetCDF

Implementação do NetCDF para Processamento de Alto Desempenho

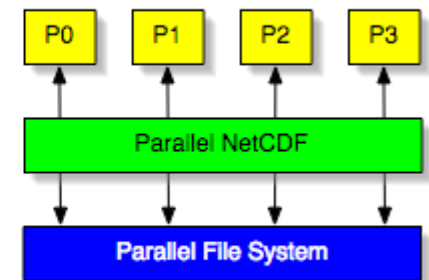
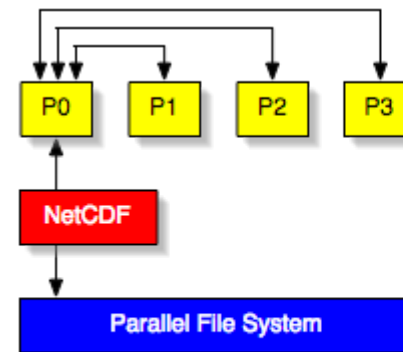
Baseada no MPI-IO

- Extensão de comunicações do MPI
- Buffering coletivo

Dois tipos de operações via API:

- Independentes
- **Cooperativas**

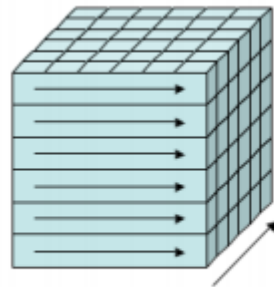
Solicitações de I/O são reordenadas para otimizar o acesso no sistema de arquivos



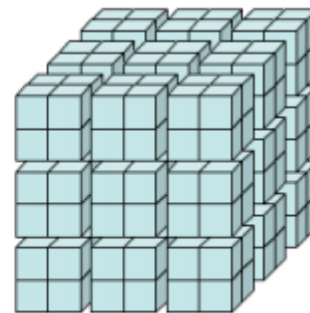
Parallel-NetCDF

Possibilita o uso de chunks:

- Ao invés de escrita sequencial, escreve-se em blocos de tamanho fixo
- Tamanho do bloco pode ser otimizado para a mídia
- Blocos podem ser comprimidos individualmente

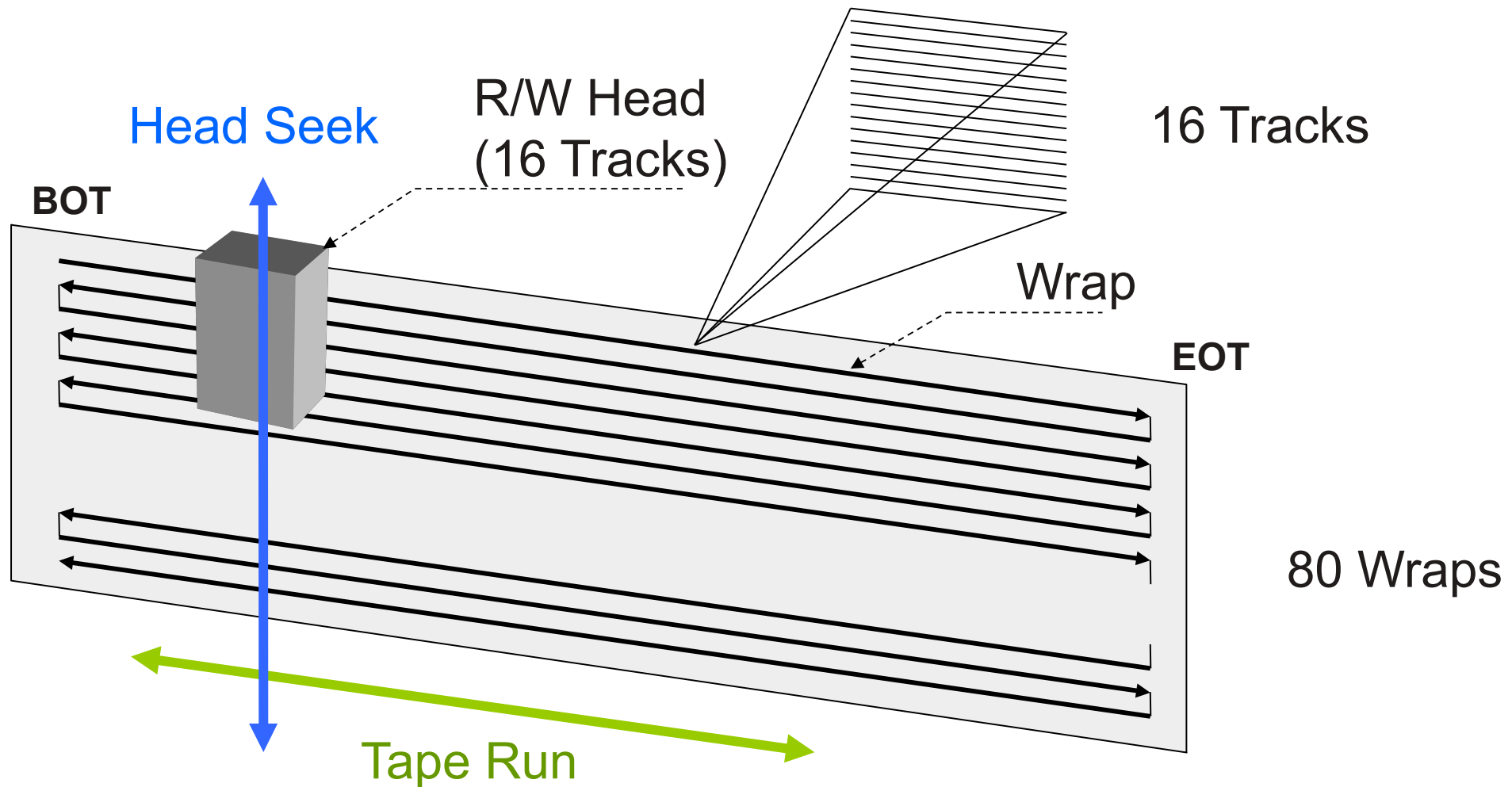


index order



chunked

I/O em Fitas (!)



LTO-6:

- Taxa de transferência: 160 MB/s (raw)
- Armazenamento: 2.5TB (sem compressão), até 1GB de cache

I/O em Fitas – LTFS

BOT

EOT

LP2

LP3

LP4

I/O em Fitas – LTFS

```
<directory>
  <name>LTFS Volume</name>
  <creationtime>2010-01-28 19:39:50.715656751 UTC</creationtime>
  <modifytime>2010-01-28 19:39:55.231540960 UTC</modifytime>
  <accesstime>2010-01-28 19:39:50.715656751 UTC</accesstime>
  <contents>
    <file>
      <name>Domain.nc</name>
      <length>10485760</length>
      <extentinfo>
        <extent>
          <partition>b</partition>
          <startblock>8</startblock>
          <byteoffset>0</byteoffset>
          <bytecount>720000</bytecount>
        </extent>
      </extentinfo>
      <extendedattributes>
        <xattr>
          <key>units</key>
          <value>meters</value>
        </xattr>
      </extendedattributes>
    </file>
    ...
  </contents>
</directory>
```

Consórcio OpenPOWER

OpenPOWER

Abertura de hardware e software da plataforma POWER

- Anúncio em Agosto/2013
- Licenciamento do POWER8 para terceiros
- Open source firmware

NVIDIA + IBM = CUDA GPU no POWER8

- Processamento de alto desempenho em clouds



Considerações Finais

Considerações Finais

- PAD é crucial para tomadas de decisões e negócios
- Múltiplas disciplinas se encontram em PAD
 - Modelagem, otimização, implementação
 - Times heterogêneos
- Dados de alta resolução em simulações são uma realidade
 - E um prato cheio pra quem busca otimizar I/O
- É possível fazer pesquisa na indústria

Obrigado!



Quando algoritmos e força bruta se encontram: processamento de alto desempenho na IBM Research Brasil

Lucas C. Villa Real
Research Software Engineer
lucasvr@br.ibm.com

