MINISTÉRIO DA CIÊNCIA E TECNOLOGIA
INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS

Computação em Grade e Aplicações na Física

Haroldo Fraga de Campos Velho

Laboratório de Computação e Matemática Aplicada (LAC)
Instituto Nacional de Pesquisas Espaciais (INPE)
E-mail: haroldo@lac.inpe.br
Web-page: http://www.lac.inpe.br/~haroldo/

Conteúdo da apresentação

- Computação em grade – preliminares
- Exemplos de projetos de computação em grade
- O projeto G-BRAMS
  - # Modelo meteorológico BRAMS
  - # Previsão meteorológica e climatologia
  - # Resultados do projeto
- Computação em grade para o projeto BRAVO

Dois princípios contraditórios

- **Na estratégia militar:**

  Dividir para conquistar!

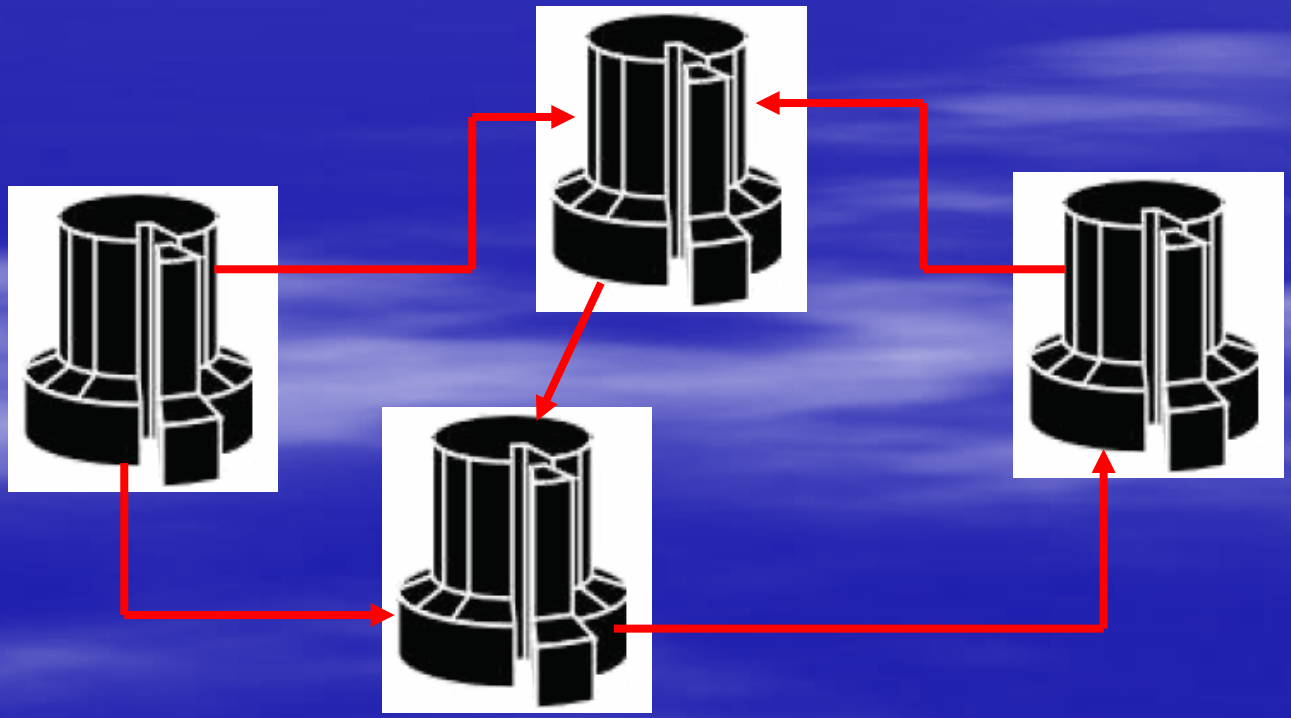
- **Em computação em grade:**

  Agrupar para conquistar!

Computação em grade - preliminares

"Grid" concept (Foster-Kesselman):

hardware and software environment with reliable, constant, and cheap access for the computer facilities



Computação em grade

artigo de Ian Foster

The Grid: A New Infrastructure for 21st Century Science

www.aip.org/pt/vol-55/iss-2/p42.html#fig1

Grid goes beyond sharing and distributing data and computing resources:

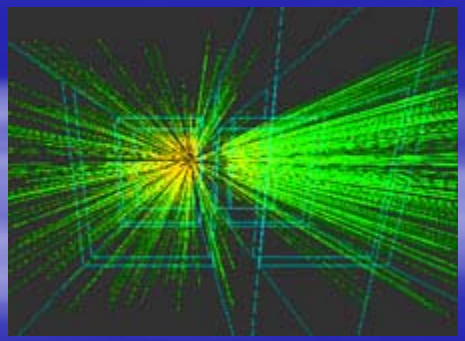
Science portals
Distributed computing
Large-scale data analysis
Collaborative work
Computer-in-the-loop instrumentation
Scientific instruments such as telescopes, synchrotrons, and electron microscopes generate raw data streams that are archived for subsequent batch processing. But quasi-real-time analysis can greatly enhance an instrument's capabilities

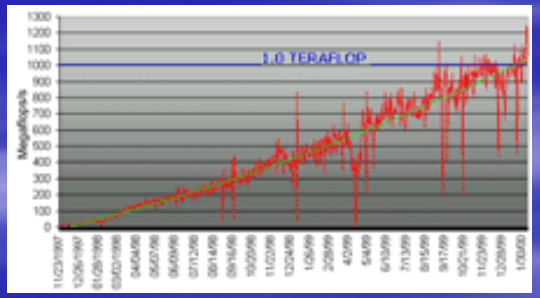
Computação em grade

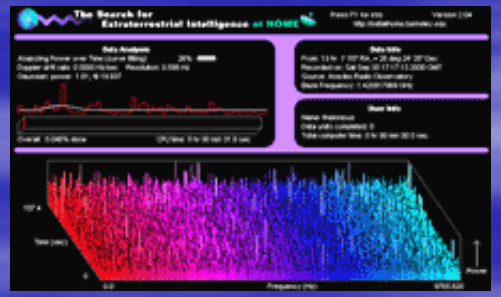
Nature – December/2000

artigo de Ian Foster

Internet computing and Emerging Grid







Projetos: CERN Entropia network Seti

Projetos em *grid computing*

- Grid Physics Network (GriPhyN)
- International virtual data grid Lab. (iVDgL)
- Data Grid (European Union)
  - # High energy physics
  - # Biology
  - # Earth observation
- Grid.org (IBM support)
  (Human proteome, folding project, cancer research)

Montage

An Astronomical Image Mosaic Engine

NASA Space Act Award Winner 2006

What is Montage?

Montage is a toolkit for assembling Flexible Image Transport System (FITS) images into custom mosaics. Key features for end users are:

Accuracy
: Preserves spatial and calibration fidelity of input images

Portability
: Runs on all common Linux/Unix platforms

Scaleability
: Runs on desktops, clusters and computational grids

Availability
: Open source code and user documentation available for download

Generality
: Supports all World Coordinate System (WCS)

News

January 5, 2007

Montage 3.0 is available download.

Learn more about how Montage is used in:

Science Product Generation
Quality Assurance
On-line Services
Analyzing Data
Education and Public Outreach

Montage

An Astronomical Image Mosaic Engine

NASA Space Act Award Winner 2006

A view of the Rho Oph dark cloud constructed with Montage from deep exposures

Onde buscar informação?

- **Grid computing info centre (GRID Infoware):**

  http://www.gridcomputing.com/

- **Wikipedia:**

  http://en.wikipedia.org/wiki/Grid_computing

- **Open Grid Forum**

  http://www.ogf.org/

- **Grid Computing Planet**

  http://www.gridcomputingplanet.com/

Grade (grid) x agrupamento (cluster)

- Conectividade
- Heterogeneidade
- Dispersão geográfica
- Compartilhamento
- Múltiplos domínios administrativos
- Controle distribuído

GBRAMS: Proposta

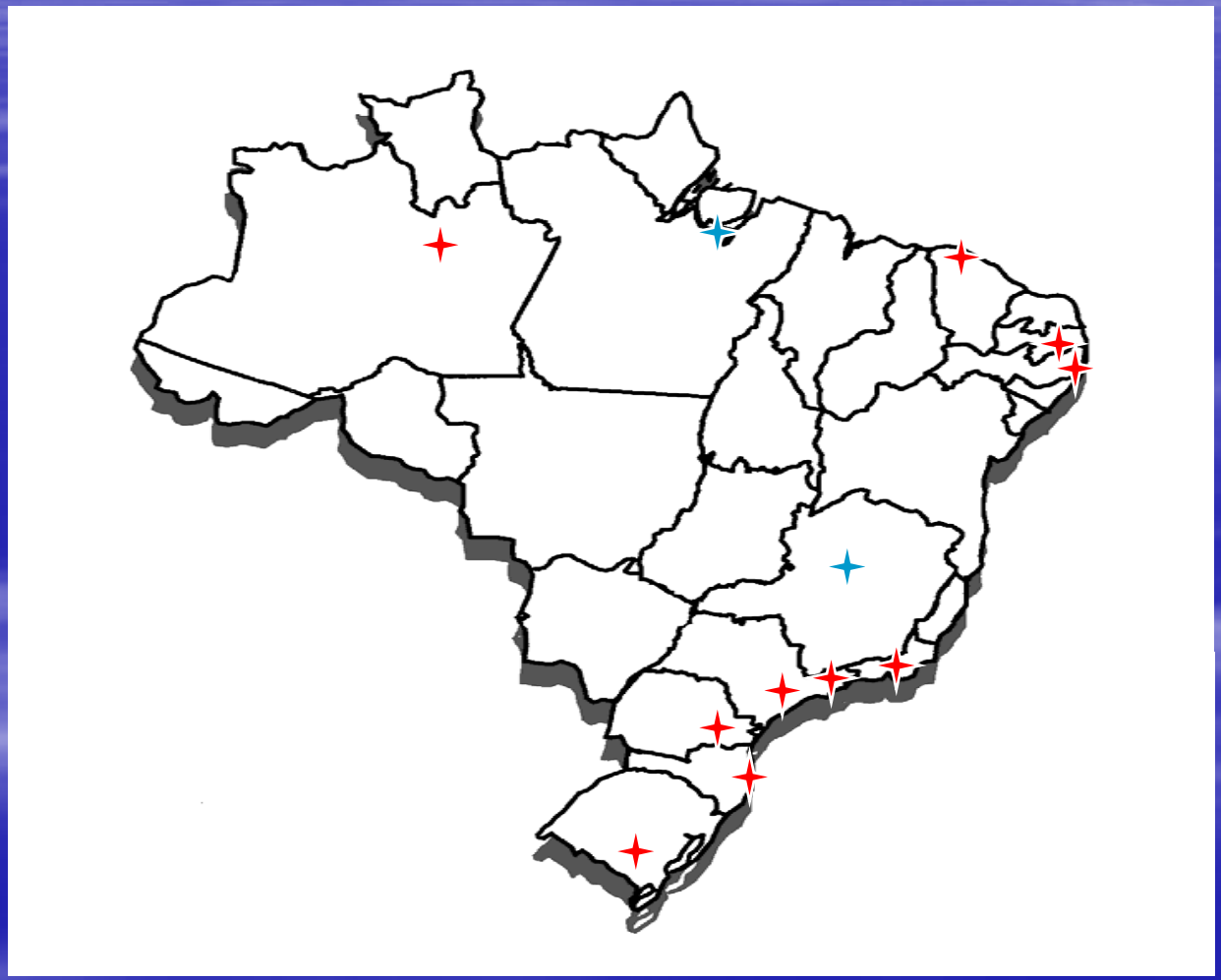
- **Objetivo:**
  - Testar a viabilidade do uso operacional de grades em previsões de clima regional
  - Limites do conceito de grade

- **Mecanismos:**
  - Três clusters (UFRGS, CPTEC, LAC)
  - Três middleware (OurGrid, Globus, OAR)
  - Portal (um middleware por vez nos três clusters)
  - Software previsor: BRAMS

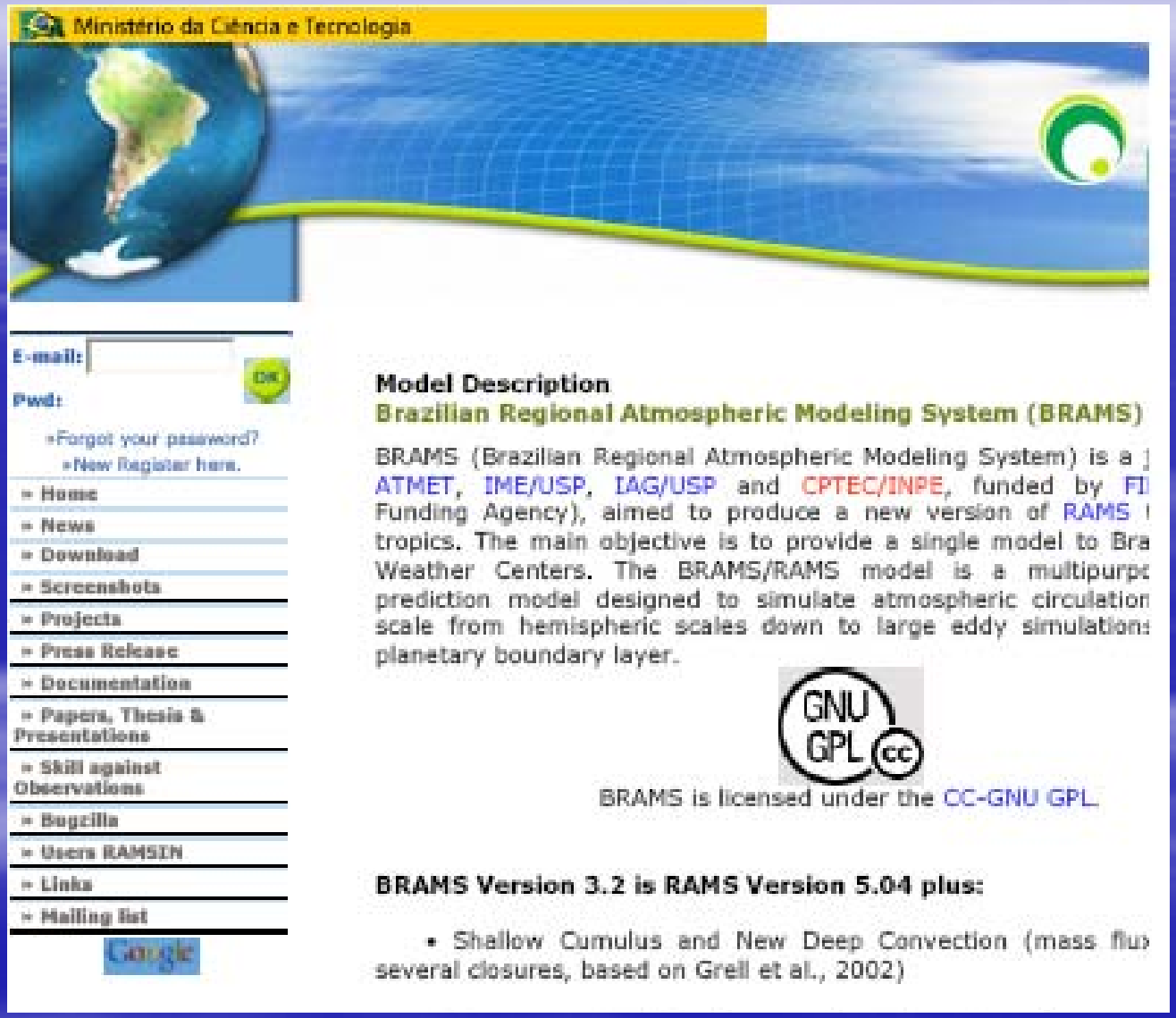
Previsão de Tempo e Clima

- **Órgãos nacionais**
  - **INPE/CPTEC e**
  - **INMET**
- **Órgãos Estaduais:**
  - **Dezenas (dados INPE/CPTEC, INMET)**
- **Iniciativa Privada:**
  - **Alguns (dados INPE/CPTEC, INMET)**

Utilidade Pública: BRAMS é insumo básico



BRAMS é software livre



Tempo e Clima

- **Previsão de Tempo:**
  - Próximos dias (máximo 15; típico 5)
  - Informações Absolutas
- **Previsão de Clima:**
  - Próxima estação (4 a 6 meses)
  - Informações Relativas
- **Previsão por conjuntos:**
  - Múltiplas execuções variando dado inicial
    - membros de conjunto
  - Tempo:
    - Perturba o dado, avalia dispersão da previsão (probabilidade)
  - Clima:
    - Múltiplas datas iniciais, aumenta a segurança da média

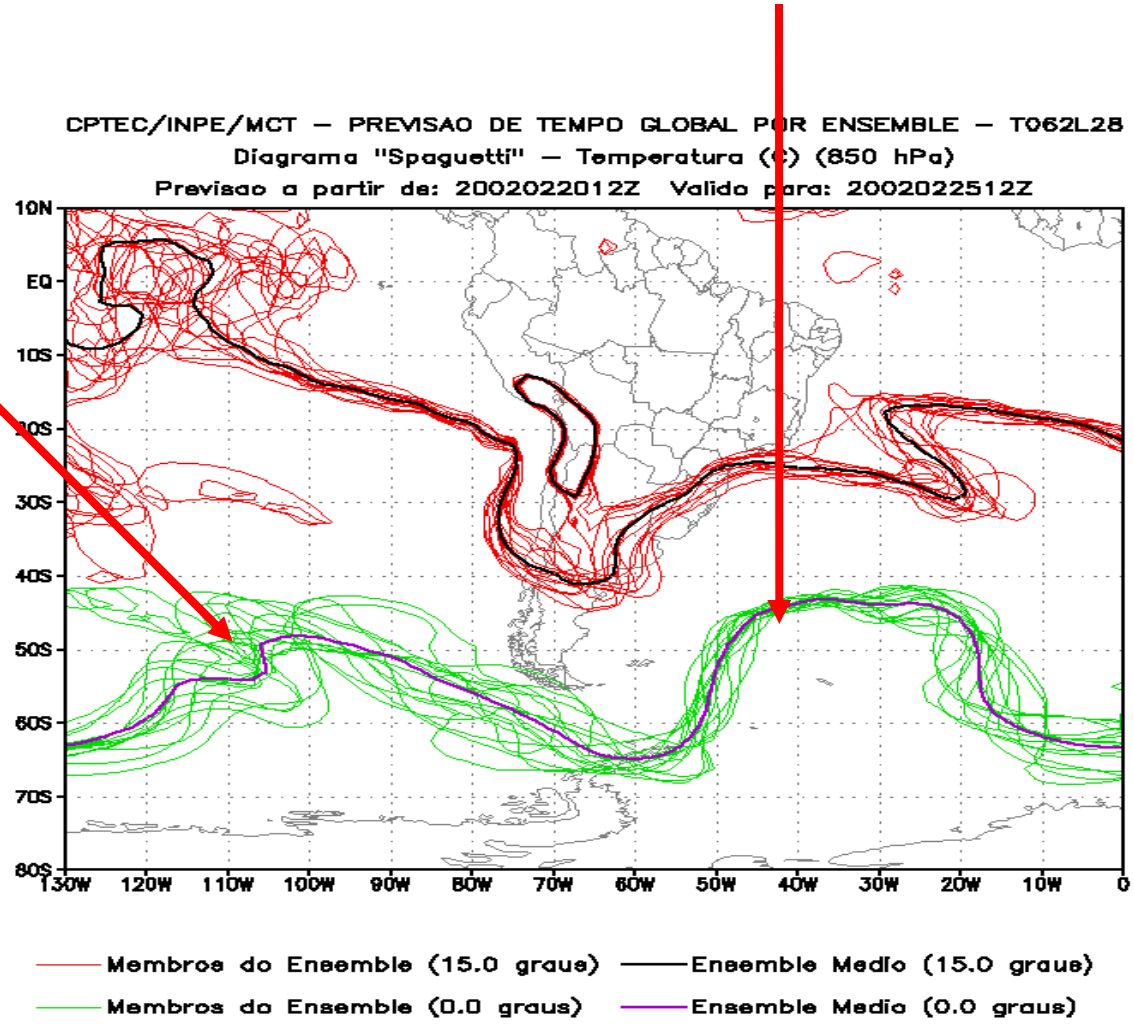
Ensemble prediction in climatology

Example of a “predictable case”

Low predictability

Operational product from Center for weather Prediction and Climate Studies - BRAZIL

Initial condition perturbation

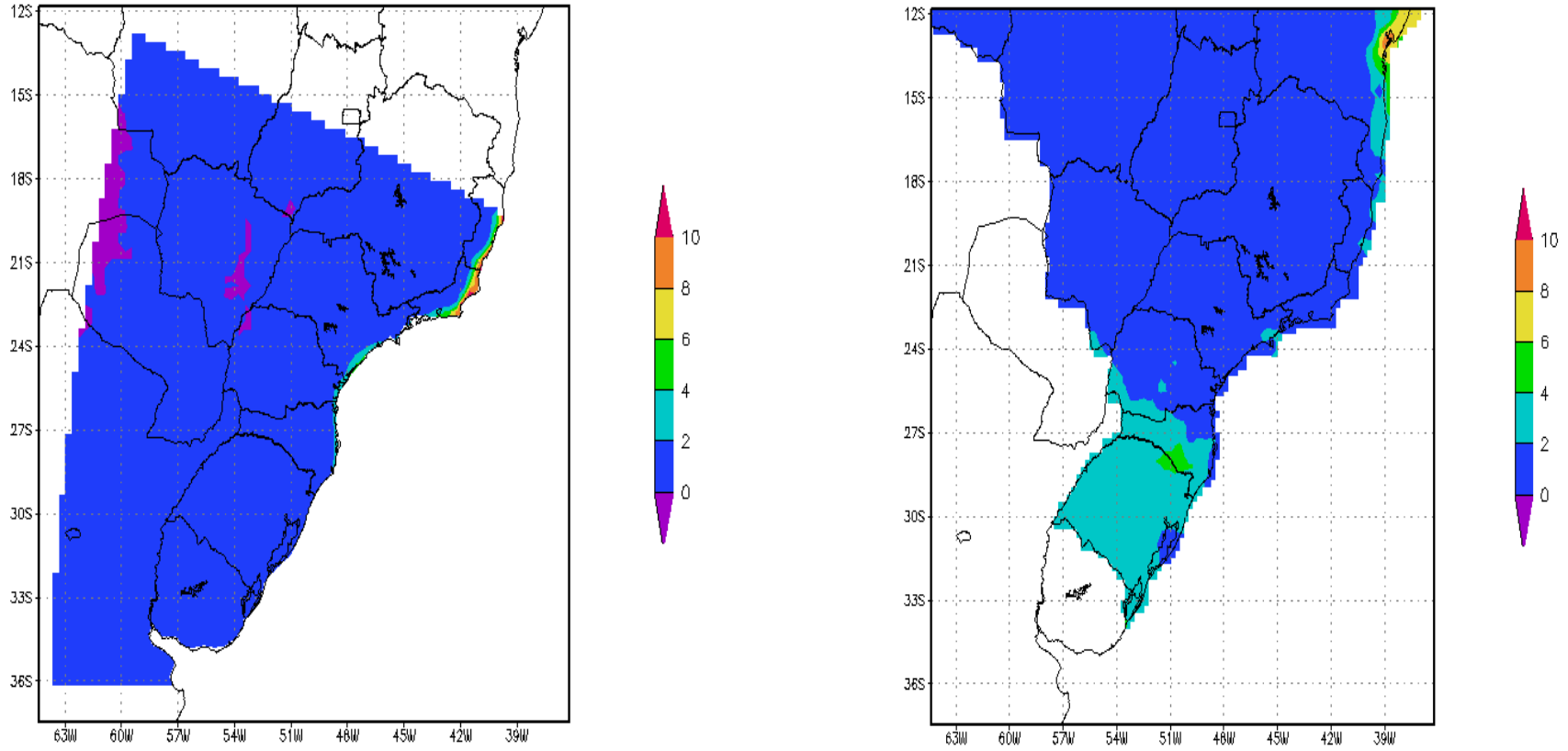


Errors of the day tend to be localized (Patil et al, 2001)

Climatologia

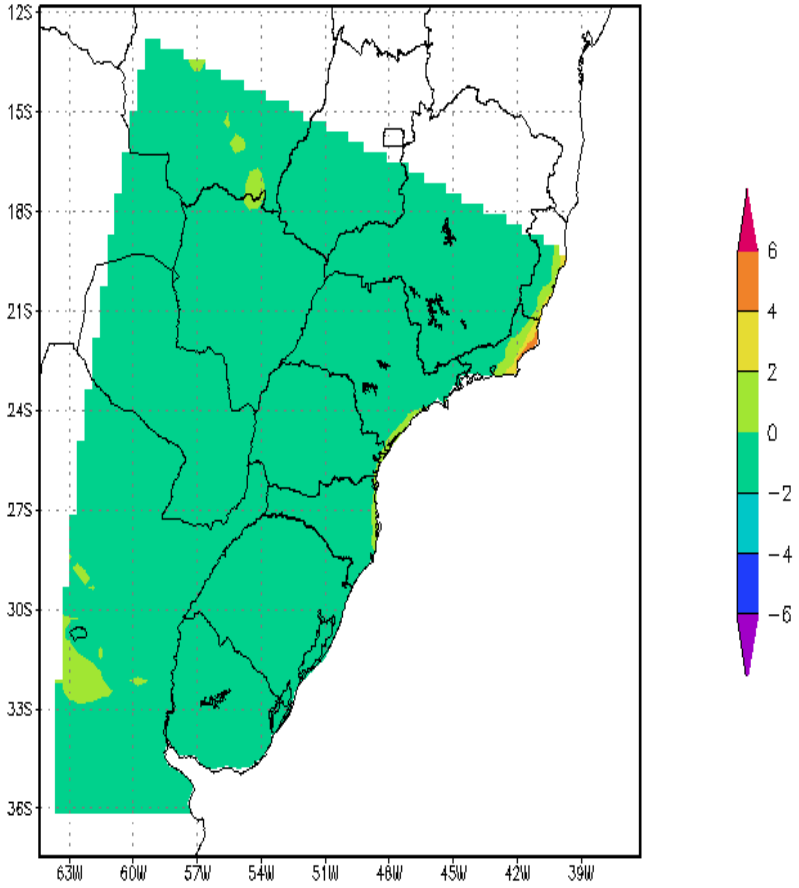
- **Previsão de Clima é reportada por desvios com relação à média pluri-anual da estação**
  - **A média é denominada climatologia**
  - **O desvio é denominado anomalia**
- **Previsão do modelo sem o viés do modelo**
  - **Requer climatologia do modelo (décadas)**
- **Verificada com relação ao desvio dos dados**
  - **Requer climatologia das observações (décadas)**

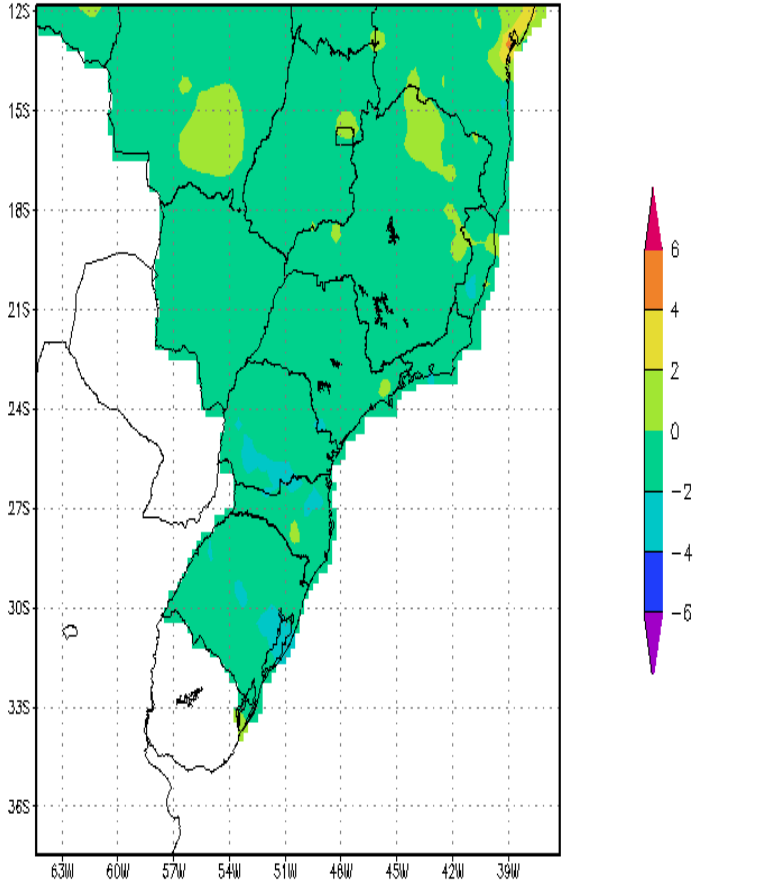
Clima: Previsão e Observação (Absolutas)



- **Precipitação JJA 2006 em mm/dia:**
  - **Prevista e Observada**

Clima: Previsão e Observação (Anomalias)





- **Anomalia da Precipitação JJA 2006 em mm/dia:**
  - **Prevista e Observada**

BRAMS: Tempo, Clima e Climatologia

- **Carga computacional:**
  - **Cluster PC com 17 processadores (1 I/O, máquina dedicada)**
  - **Máquina dedicada**

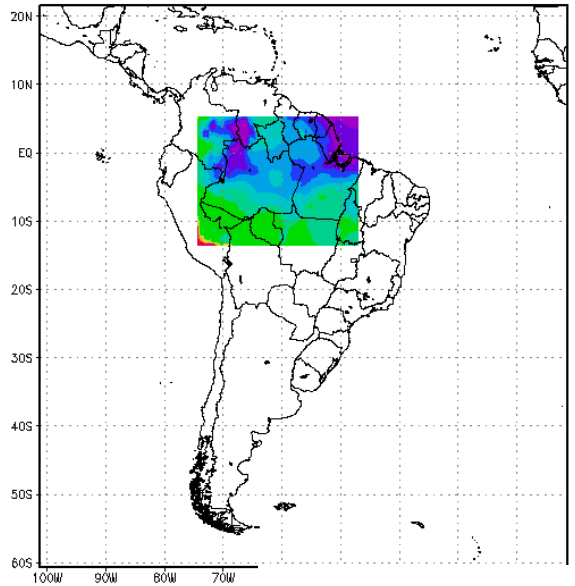
- **Carga computacional por período previsto:**
  - **Por dia: 312 segundos (5 minutos)**
  - **Por mês: 152 minutos (2,53 horas)**
  - **Por ano: 30 horas (1,25 dias)**
  - **Por década: 13 dias (0,42 meses)**

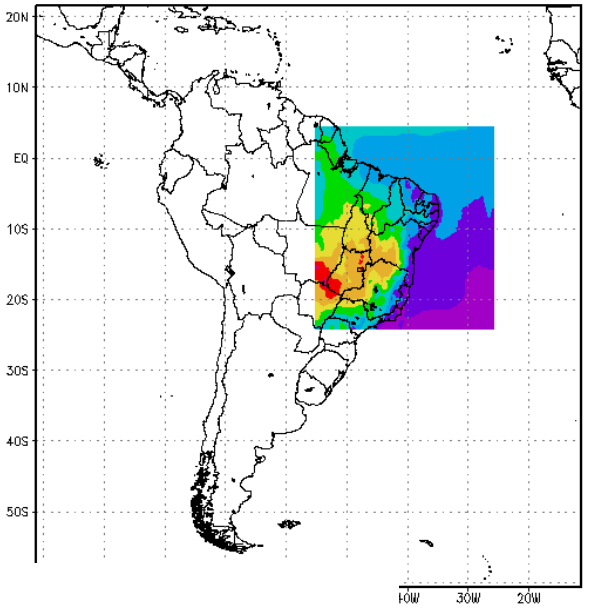
GBRAMS

- **Grade para acelerar previsão do clima:**
  - **Inviável paralelizar uma única execução do BRAMS**
  - **Paralelizar múltiplas execuções do BRAMS**

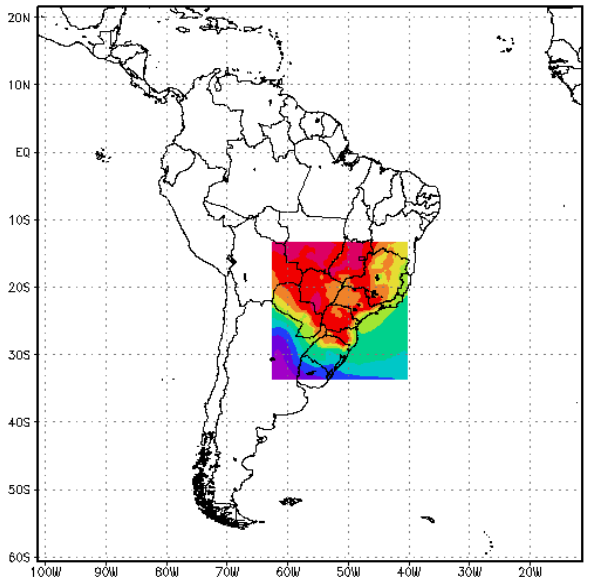
- **Alvo: Climatologia**
  - **Insumo básico para a previsão de clima dos usuários do BRAMS**
  - **Cálculo dispendioso: 0,4 mês/(década prevista x dado inicial)**
  - **Execuções sequenciais de 1 ano (checkpoint/restart)**
  - **9 instâncias de execução simultâneas**
    - **3 áreas x 3 datas iniciais**
    - **3 middleware, um por vêz**

GBRAMS: Áreas do Brasil



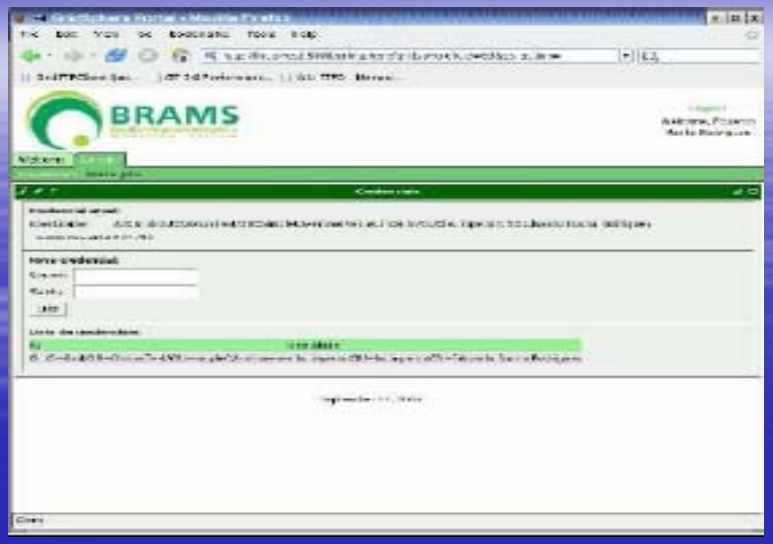


GrADS: COLA/IGES



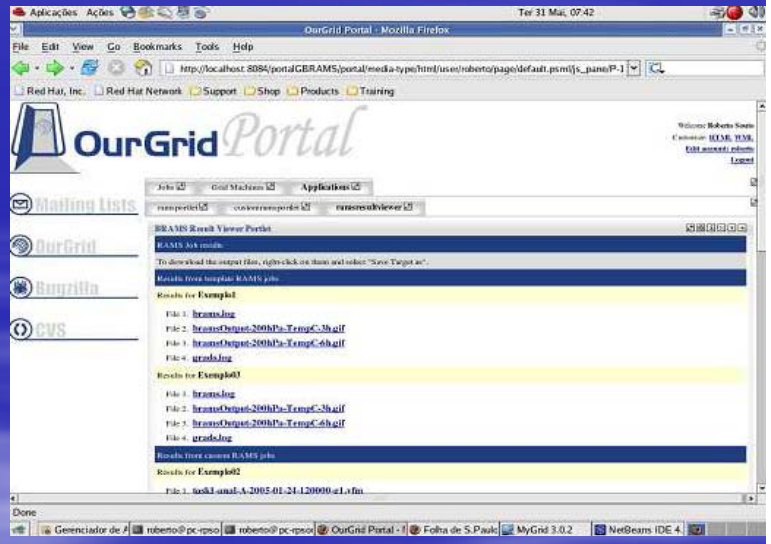
2005-09-23-10:53

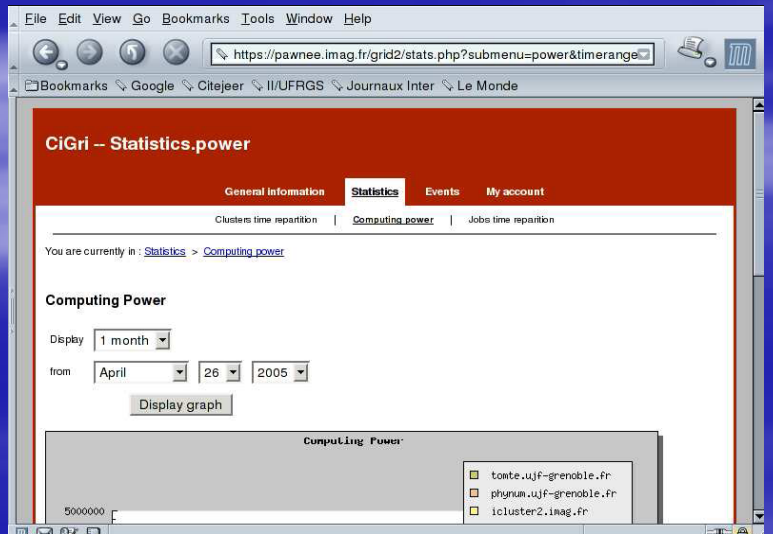
GBRAMS: Portais



Globus

OurGrid



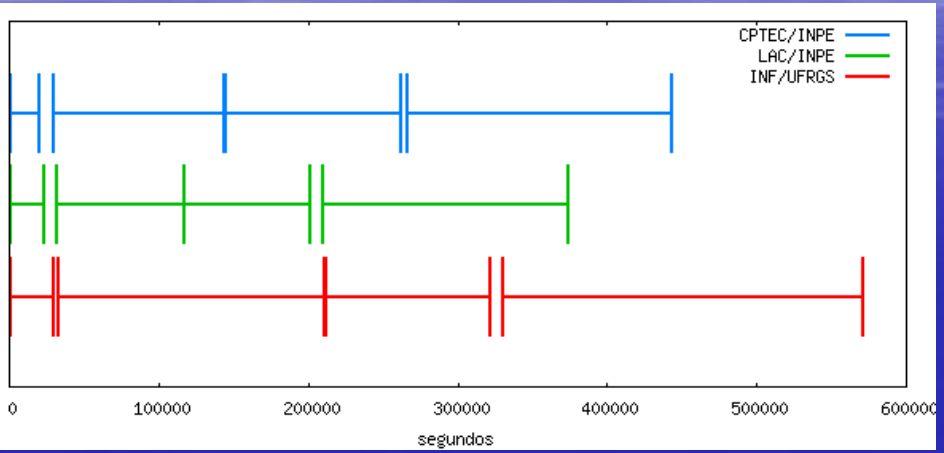


OAR

GBRAMS: Análise qualitativa

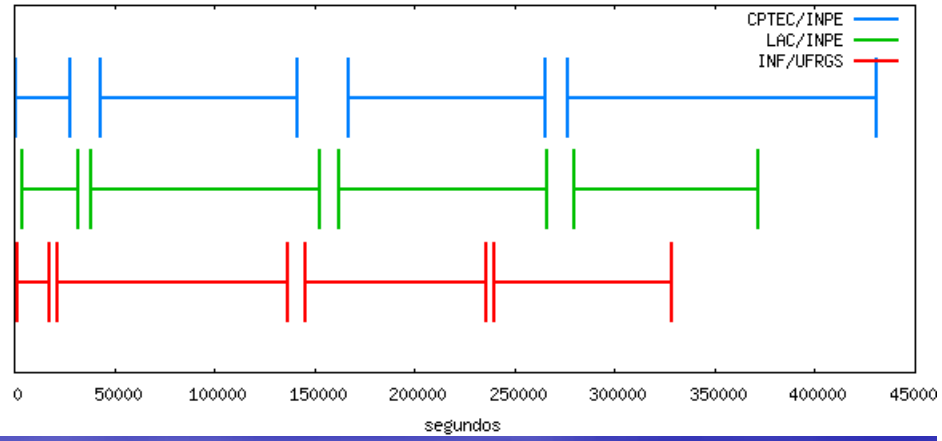
	Globus	CIGRI	OurGrid
Sw adicional	Java	mySQL, Perl	Java
Instalação	Difícil	Fácil	Fácil
Escalonamento	Nenhum	Fila de prioridades	Replicação
Segurança	GSI	Ssh	Ssh

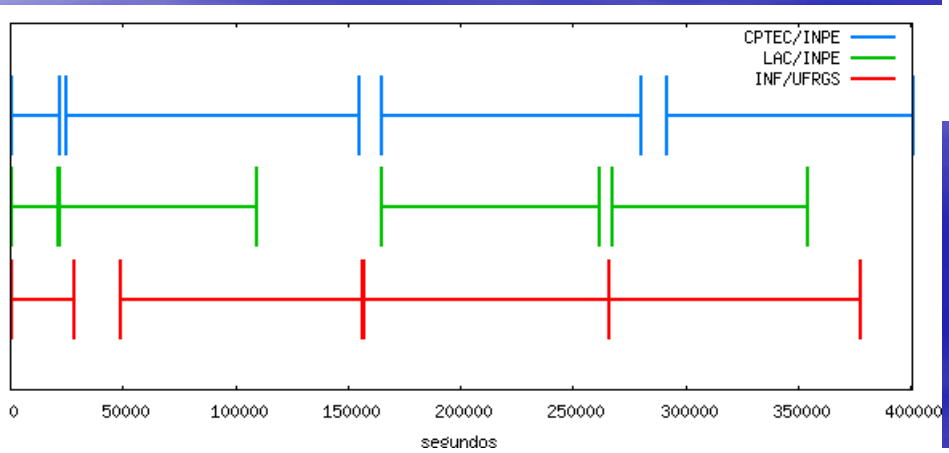
GBRAMS: Climatologia de 10 anos: 1 membro, 3 regiões, 3 middleware



Globus, 91/93 (7 dias)

OurGrid, 94/96 (5 dias)





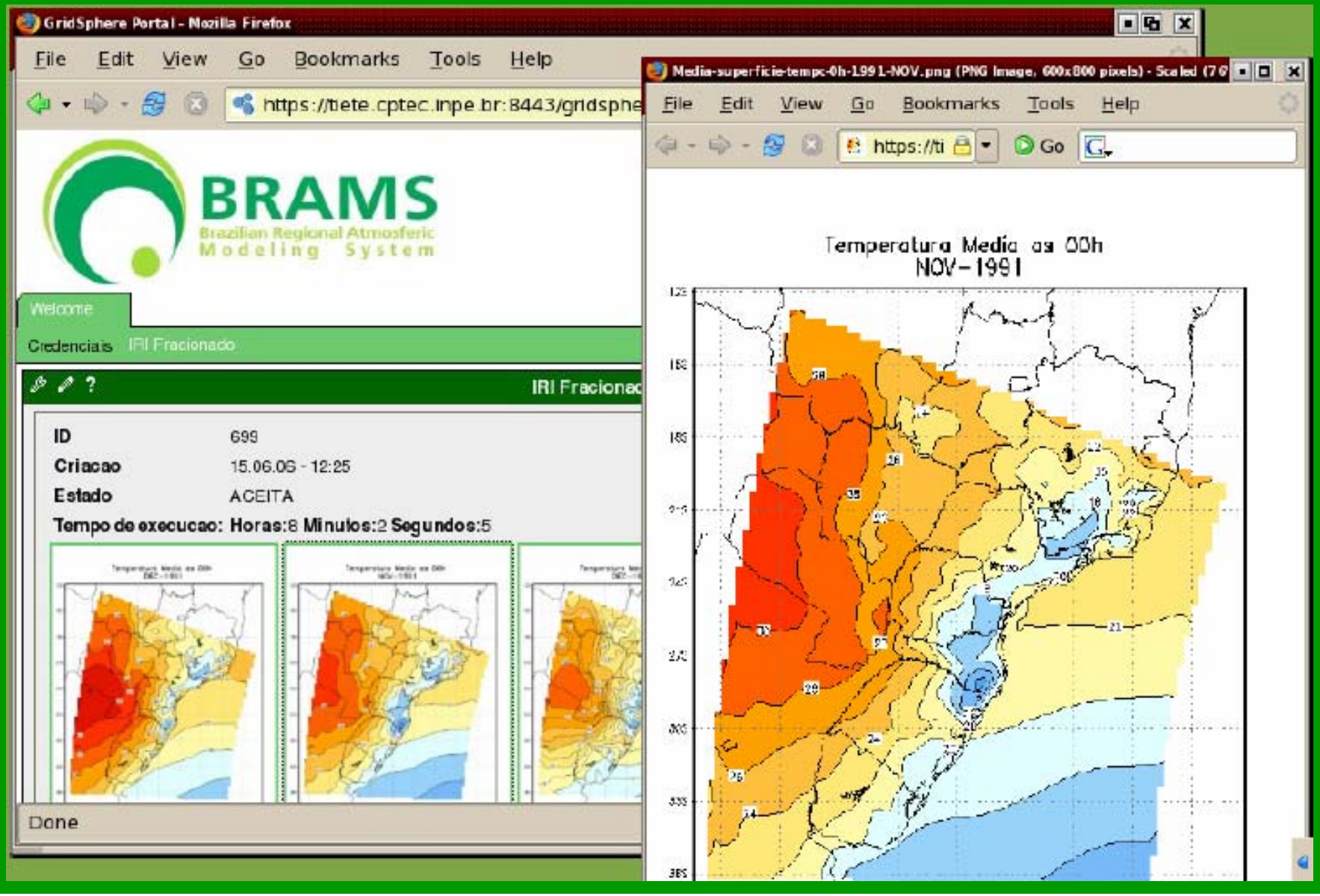
OAR, 97/99 (4,5 dias)

GBRAMS: Resultados

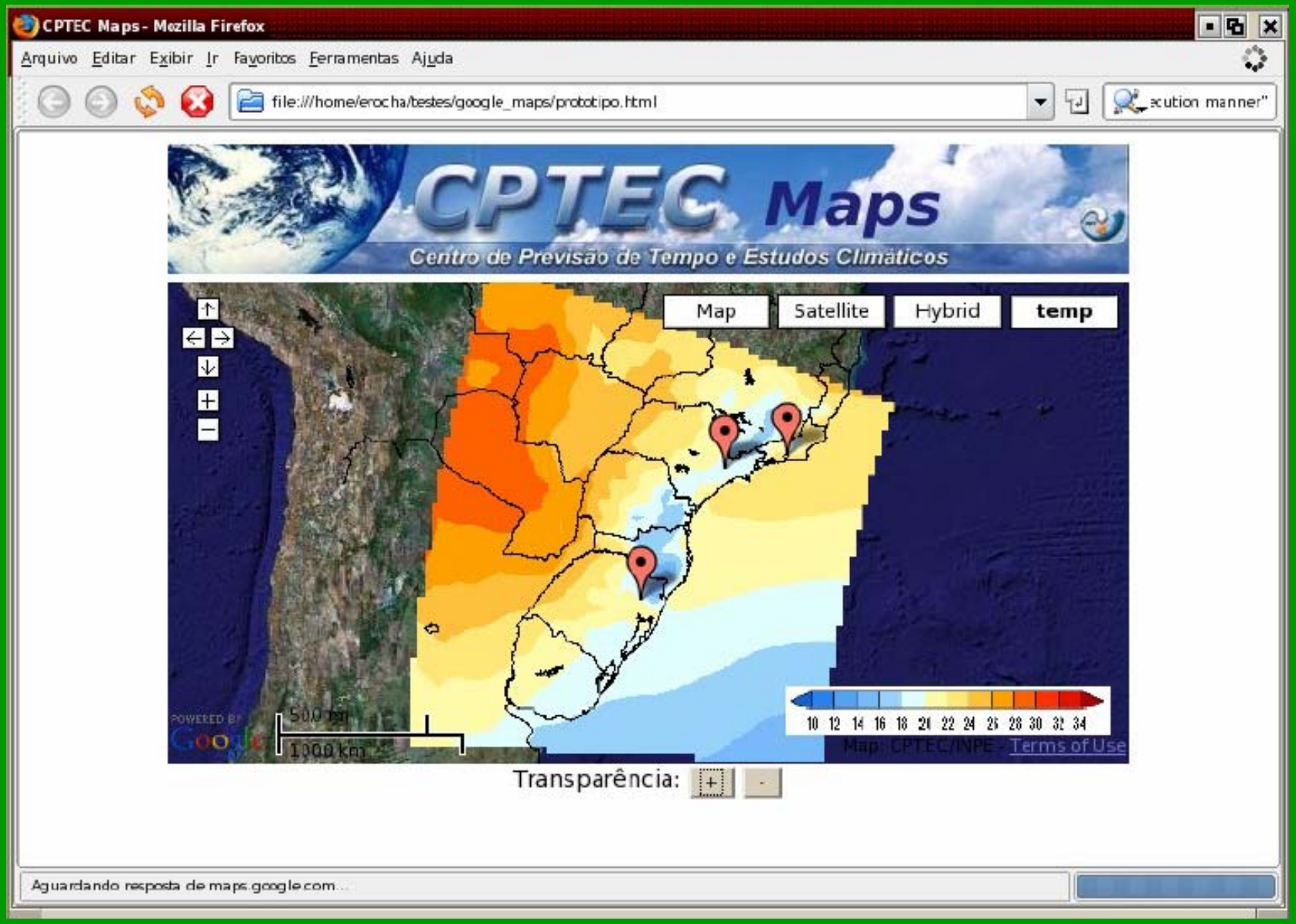
- Demonstrou a viabilidade do uso de Grade para previsão do clima
- Climatologia de 10 anos, 3 membros, 3 áreas
- Portais para 3 middleware obtiveram:

Middleware	Ocupação (%)
Globus	96
OurGrid	90
OAR-Cigri	87

G-BRAMS – sub-produto

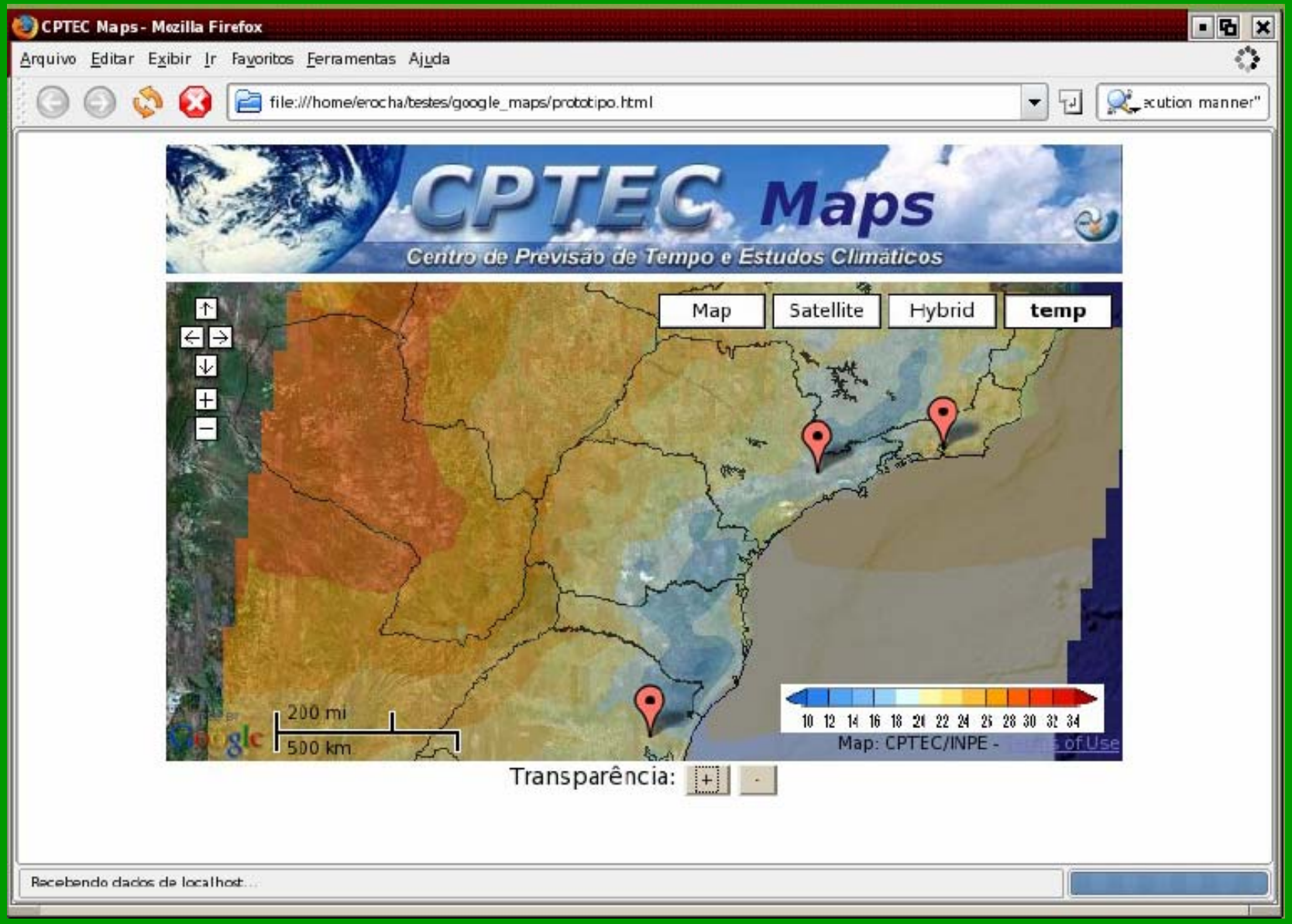


G-BRAMS – sub-produto



G-BRAMS – sub-produto

G-BRAMS – sub-produto



GBRAMS: Resultados

- **Os resultados apresentados em Conferências da área:**
- **VecPar – Julho/2006**
- **CCGRID – Maio/2007**

Processing Mesoscale Climatology in a Grid Environment

R.P. Souto, R.B. Ávila, P.O.A. Navaux, M.X. Py, N. Maillard, T.A. Diverio
Instituto de Informática/UFRGS – Porto Alegre, Brazil
{rpsouto,avila,navaux,mpy,nicolas,diverio}@inf.ufrgs.br

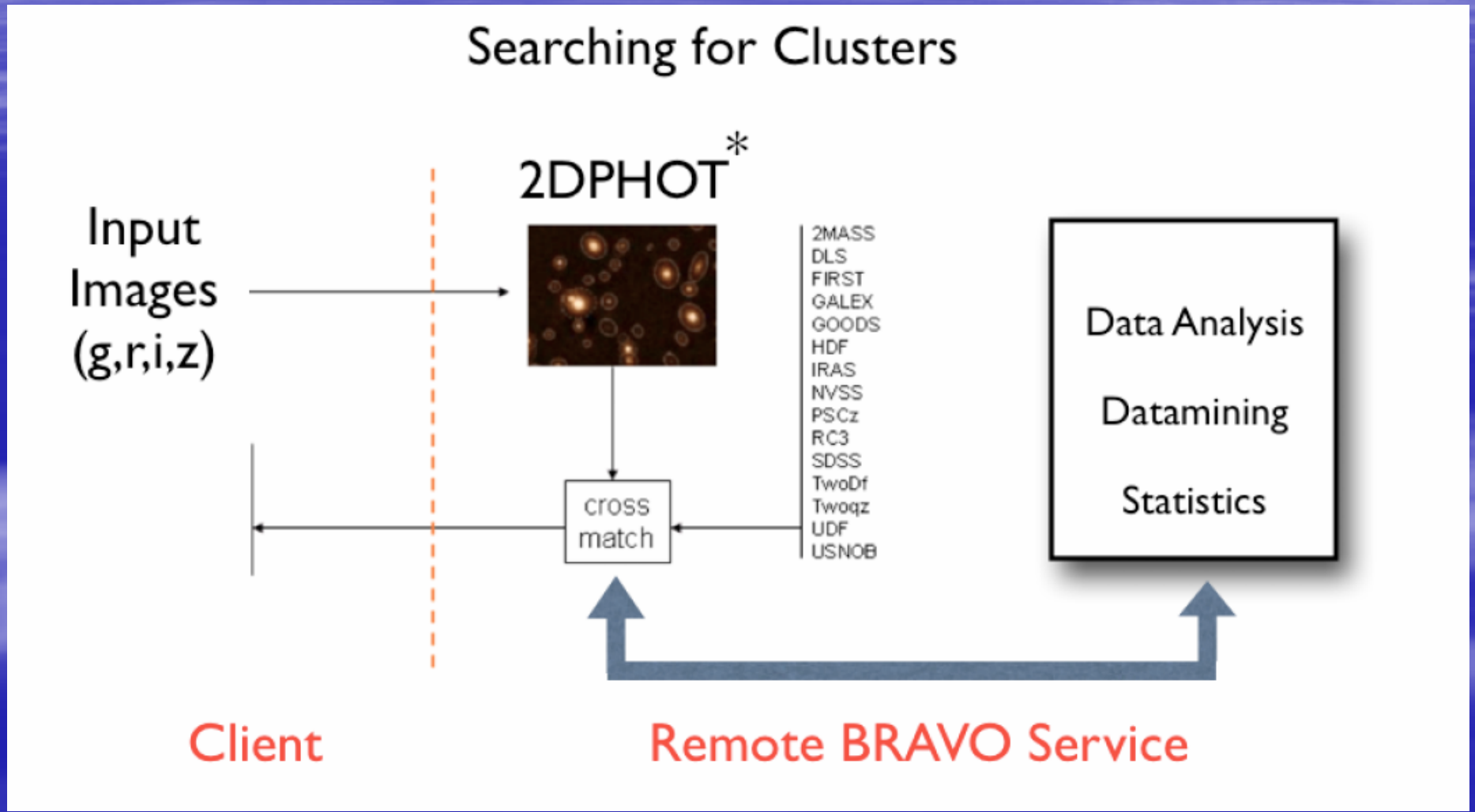
H.F. Campos Velho, S. Stephany, A.J. Preto
LAC/INPE – São José dos Campos, Brazil
{haroldo,stephan,airam}@lac.inpe.br

J. Panetta, E.R. Rodrigues, E.S. Almeida
CPTEC/INPE – Cachoeira Paulista, Brazil
{panetta,erocha,eugenio}@cptec.inpe.br

P.L. Silva Dias, A.W. Gandú
IAG/USP – São Paulo, Brazil
{pldsdias,gandu}@master.iag.usp.br

- **International Meeting on High Performance Computing for Computational Science - VecPar**
- **IEEE International Symposium on Cluster Computing and the Grid** - CCGird

Projeto BRAVO



Scientific challenges:

1. Before the XX century:

We want to know the nature laws (mechanics, thermodynamics, electromagnetism, life evolution, social behaviour, transfinite numbers)

2. During the XX century:

We know the laws (equations), but we want to solve them.
Remarkable conquest: modern numerical weather prediction!

3. After the XX century (our century!):

Starting this new century, data analysis is occuping a central role in the science (genomic, data mining, background cosmic radiation in microwave, data assimilation).

Matter as a fluid

$$\nabla^2\Phi = 4\pi G\rho$$

$$\frac{\partial\rho}{\partial\tau} + \nabla\cdot\left(\rho\vec{v}\right) = 0$$

$$\frac{\partial\vec{v}}{\partial\tau} + \left(\vec{v}\cdot\nabla\right)\vec{v} = -\nabla\Phi - \frac{1}{\rho}\nabla P + \nabla\cdot(\eta\nabla\vec{v})$$

Cosmic pressure

Adhesion term

$\Phi \sim 1/r$: the Newtonian gravitational potential

ρ: mass density

Two terms for estimating:

1. *Pressure* function: there is a limit condition

$$\lim_{z\to 0} \nabla P = 0$$

but $\nabla P \neq 0$ for $z \to \infty$ – and it could be similar to the radiation pressure in stelar dynamics.

2. *Adhesion* parameter (*cosmic viscosity*): there is a result for determining it.

APS/123-QED

Weakly Nonlocal Hydrodynamics and The Origin of Viscosity in the Adhesion Model

A.L.B. Ribeiro* and J.G. Peixoto de Faria†
Laboratório de Astrofísica Teórica e Observacional
Departamento de Ciências Exatas e Tecnológicas
Universidade Estadual de Santa Cruz

arXiv:astro-ph/0502580 v1 28 Feb 2005

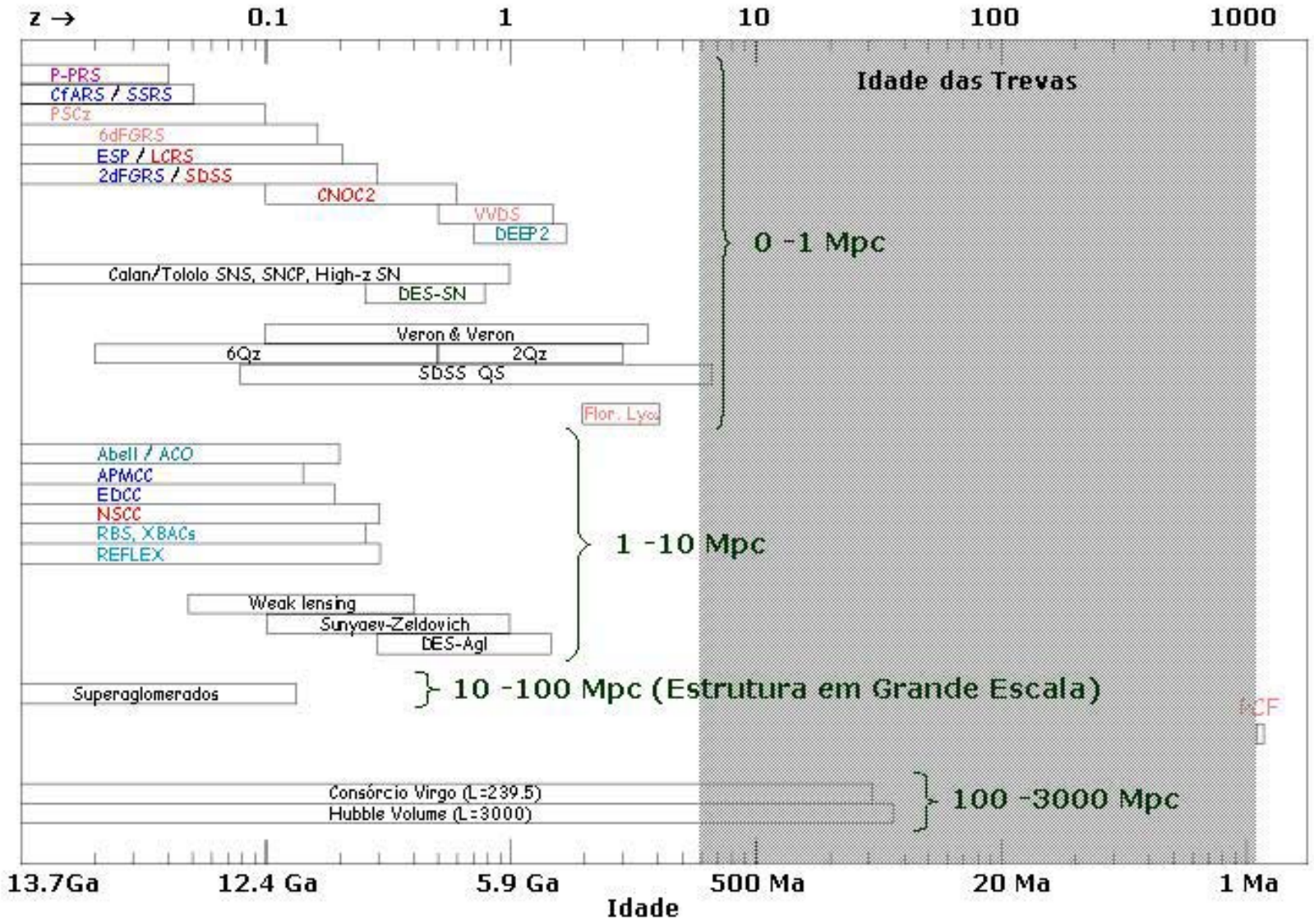
If the previous approach is correct, and similar turbulence dynamics can be employed to describe the cosmological evolution, then some features found in the turbulence in fluids must be verified.

One of this features is the -5/3 law in the inertial sub-range on the energy sprectrum.

This can be investigate considering the potential energy computed from available data (such as from the Virgo Consortium, for example).

$$\Phi = nV \int \frac{Gm^2}{2r} 4\pi n r^2 \xi(r) dr \quad ; \quad \xi(r) = \left(\frac{r}{r_0} \right)^{1.8}$$

Spectra should be computed for several redshift.


z →
0.1
1
10
100
1000
Idade das Trevas
P-PRS
CfARS / SSRS
PSCz
6dFGRS
ESP / LCRS
2dFGRS / SDSS
CNOC2
VVDS
DEEP2
0 -1 Mpc
Calan/Tololo SNS, SNCP, High-z SN
DES-SN
Veron & Veron
6Qz
2Qz
SDSS QS
Flor. Lyα
Abell / ACO
APMCC
EDCC
NSCC
RBS, XBACs
REFLEX
1 -10 Mpc
Weak lensing
Sunyaev-Zeldovich
DES-Agl
Superaglomerados
10 -100 Mpc (Estrutura em Grande Escala)
FCF
Consórcio Virgo (L=239.5)
Hubble Volume (L=3000)
100 -3000 Mpc
13.7Ga
12.4 Ga
5.9 Ga
500 Ma
20 Ma
1 Ma
Idade

Análise Preliminar

- Cubo menor: L = 120 h^{-1} Mpc

  $1.7\times10^7 \rightarrow 2.2\times10^6$ partículas

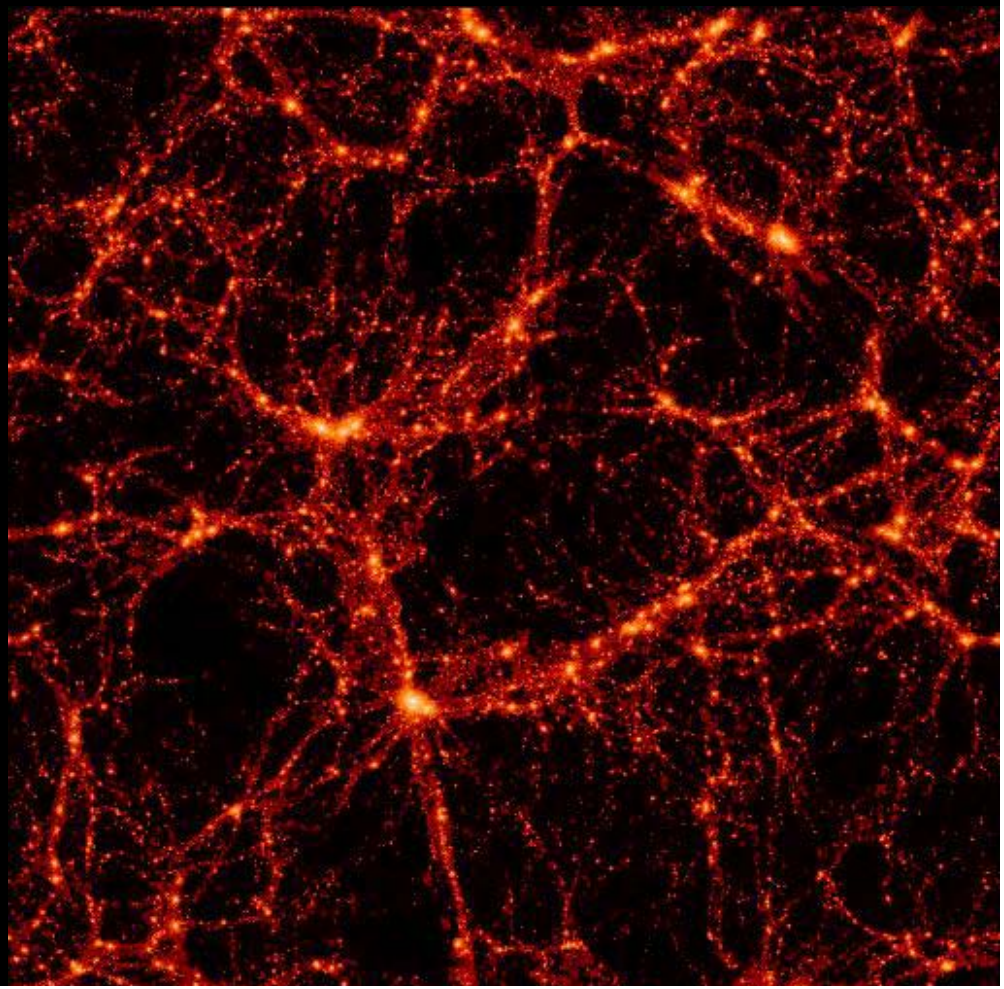
  tempo de execução ~ 60 h

- Apenas 4 instantâneos

  z = 10.0, 1.0, 0.1 e 0.0

- Efeito de borda?

L	z	<n>
239.5 Mpc	todos	1.22
120 Mpc	10.0	1.23
120 Mpc	1.0	1.27
120 Mpc	0.1	1.29
120 Mpc	0.0	1.30



Percolação (FoF)

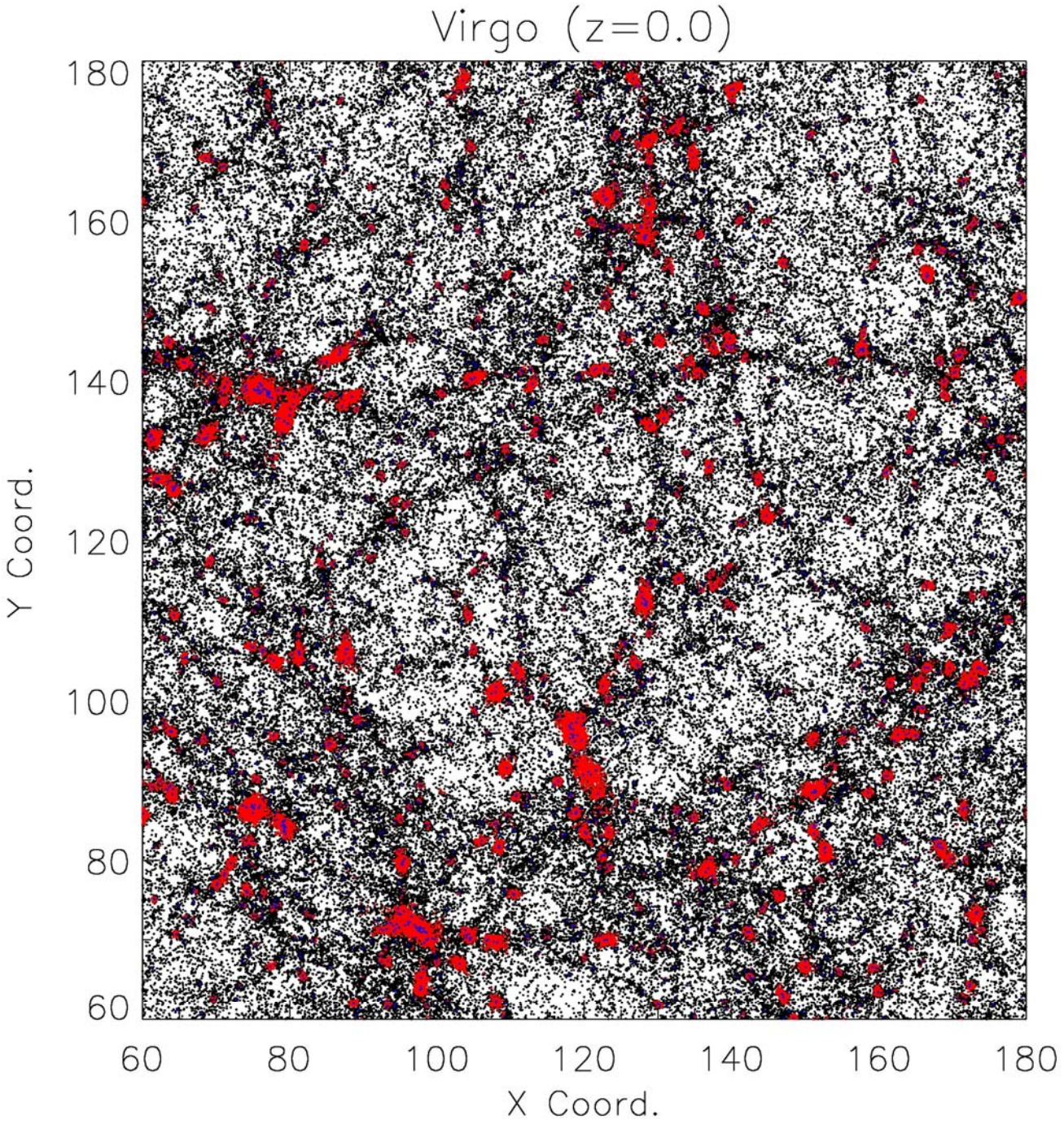
- **Raio de percolação**

  $R_{perc} = b \langle R \rangle \quad \Leftrightarrow \quad f = n / \langle n \rangle \approx 2/ b^3$

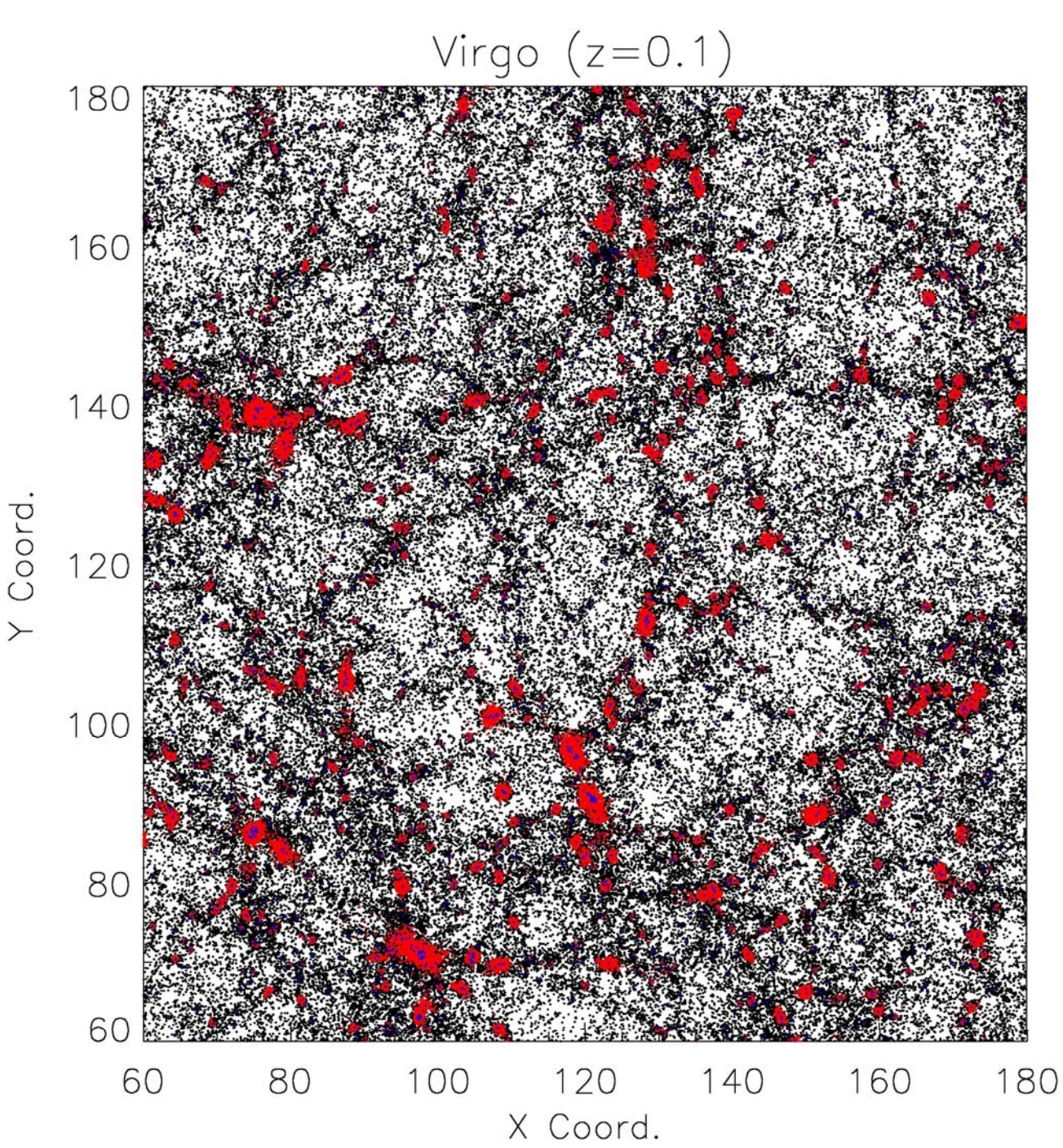
- **Escala de massas:**

$M(N_p)$	$M(M_\odot)$	classe
1	$\leq 7\times10^{10}$	galáxias pequenas e anãs
		LMC $2\times10^{10}\ M_\odot$
2-50	2×10^{11}—3×10^{12}	galáxias "normais"
		Via-Láctea $7\times10^{11}\ M_\odot$
		M87 $3\times10^{12}\ M_\odot$
50-15k	3×10^{12}—1×10^{15}	grupos ou aglomerados
		Grupo Local $4\times10^{12}\ M_\odot$
		Coma $1\times10^{15}\ M_\odot$
> 15k	$> 1\times10^{15}$	superaglomerados
		SA Local $2\times10^{15}\ M_\odot$

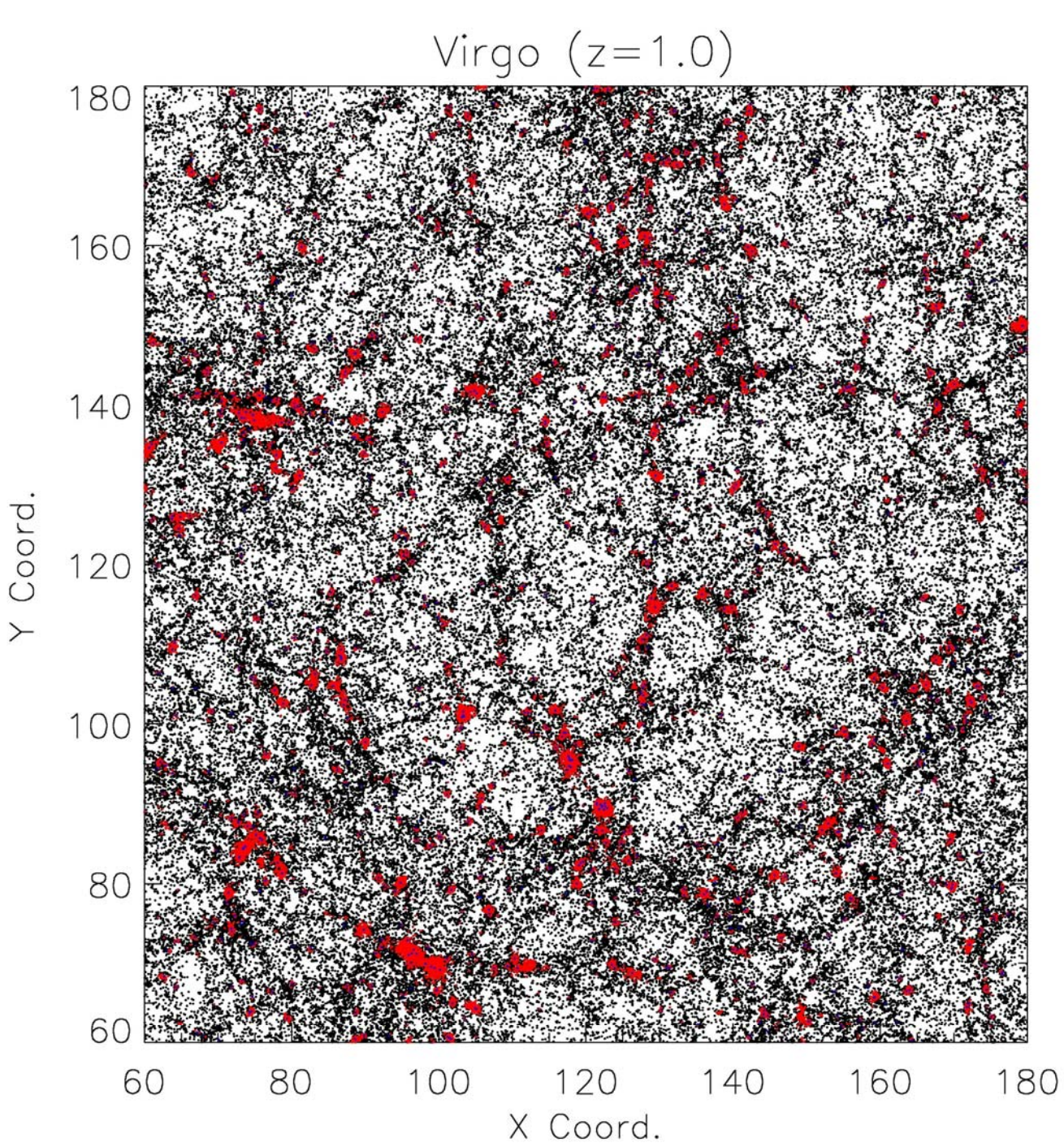
•	R_{perc}	b	f	N_p	
Gal.	0.1	0.11	1500	2-50	VL → 10 part. (R=100-150 kpc)
Agl.	0.184	0.2	250	50-15000	Gr. Local → 55 part.
SA	1.25	1.15	2	> 15000	SA Local → 30000 part.



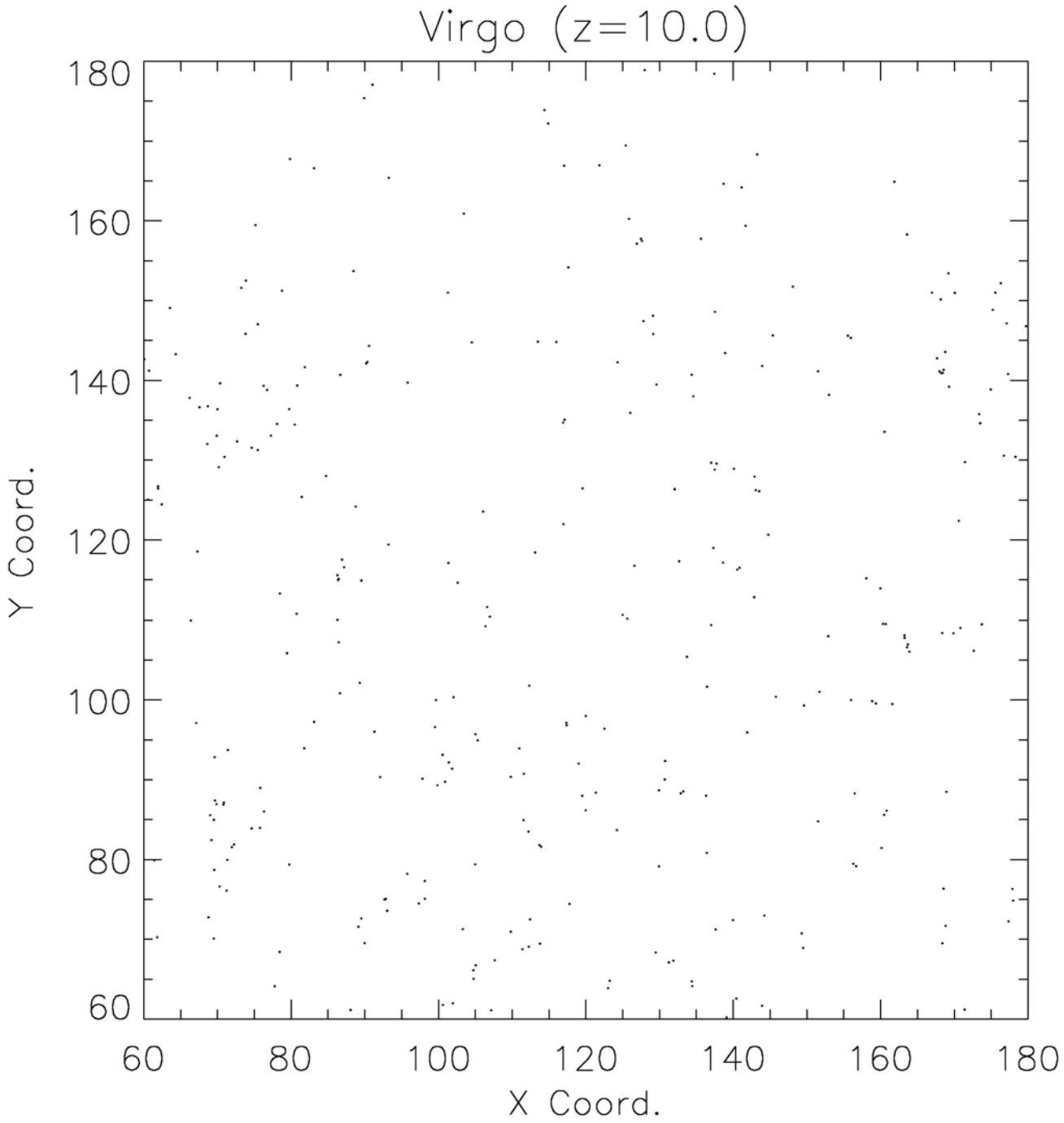
Galáxias (2-50)
Alto σ (2-50)
Aglomerados (>50)



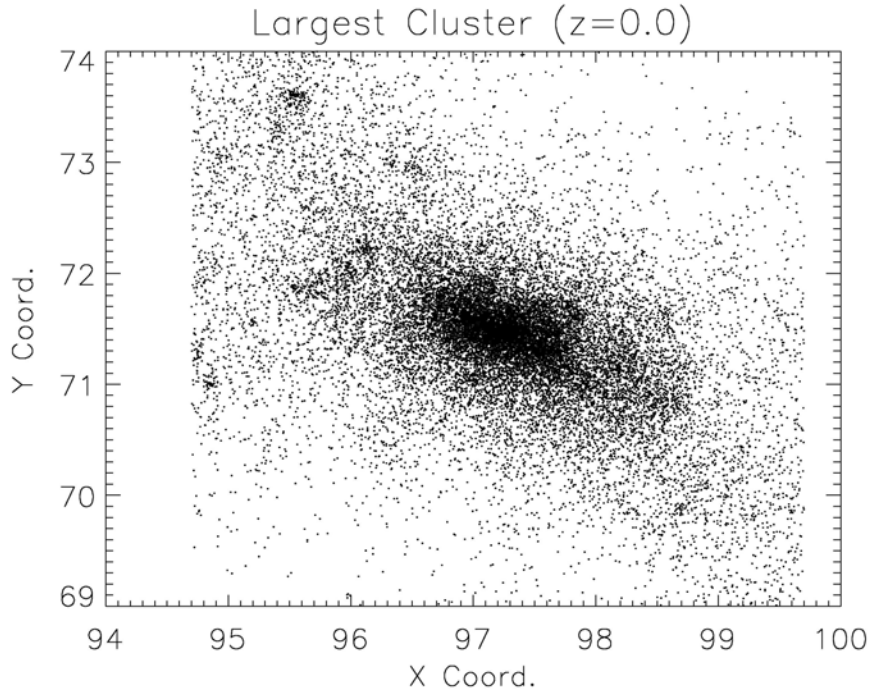
Galáxias (2-50)
Alto σ (2-50)
Aglomerados (>50)

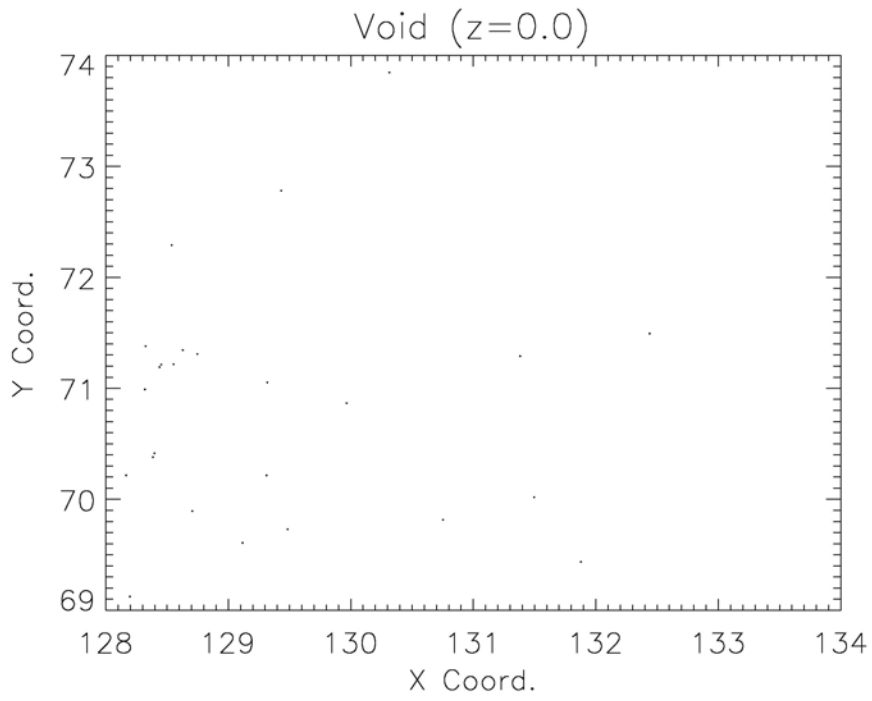


Galáxias (2-50)
Alto σ (2-50)
Aglomerados (>50)

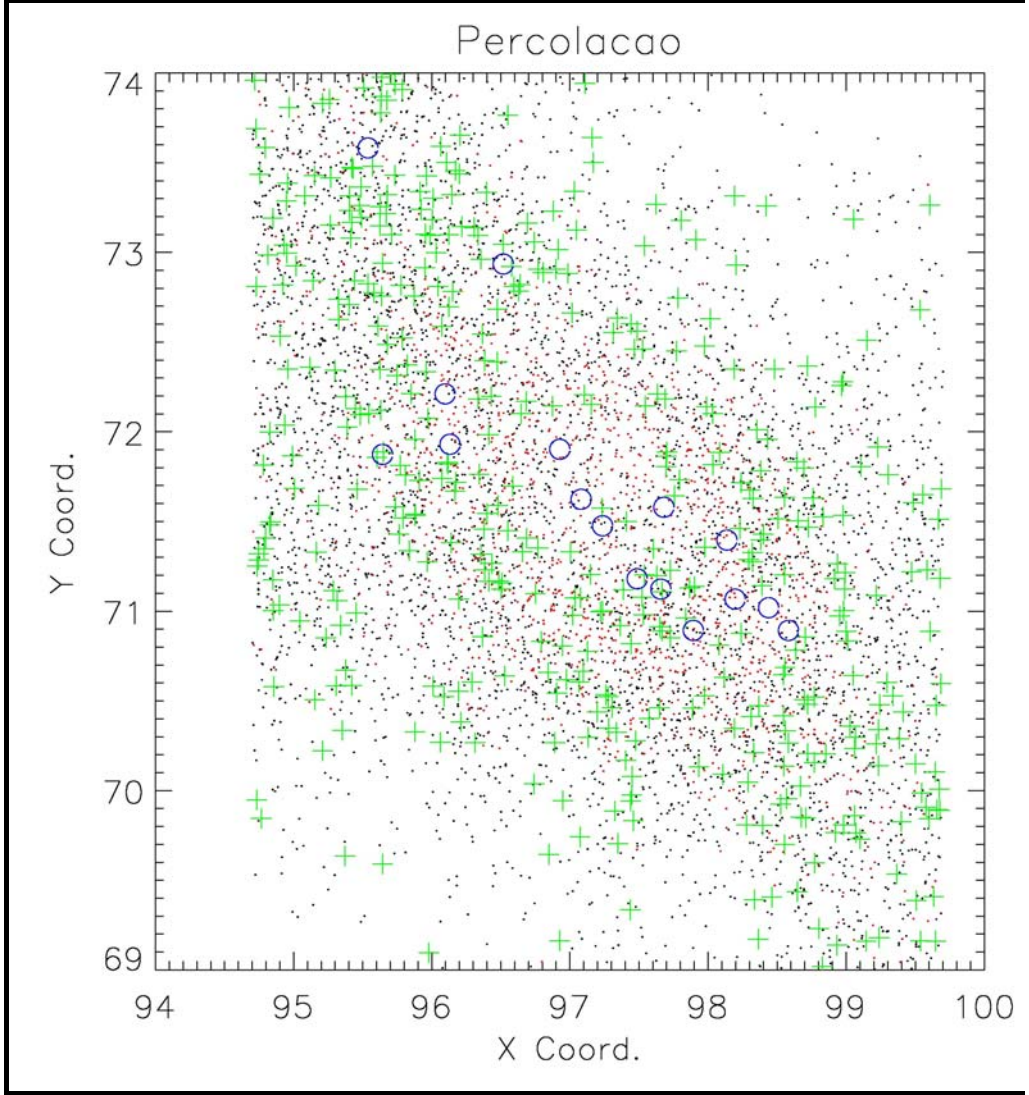


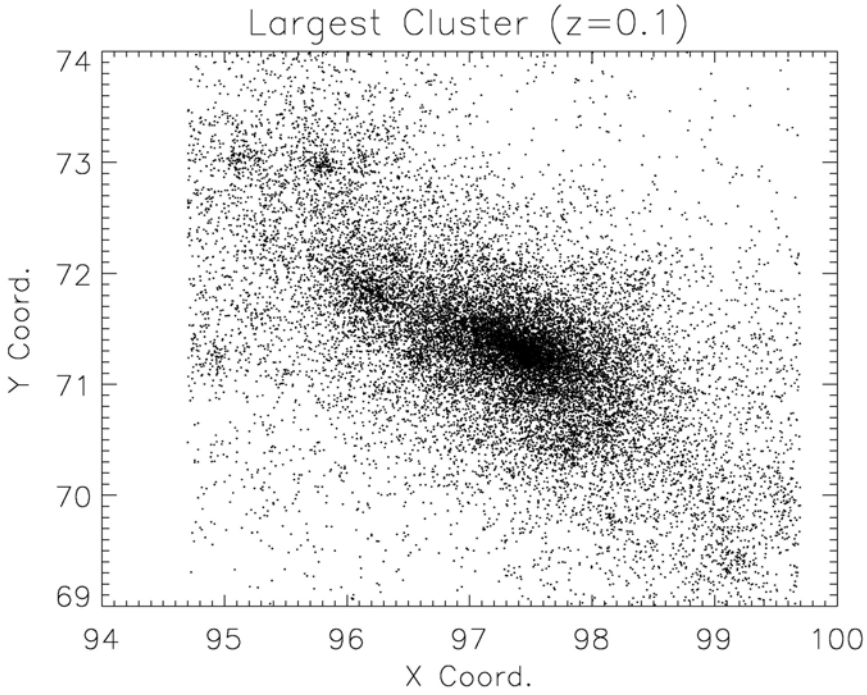
Galáxias (2-50)
Alto σ (2-50)
Aglomerados (>50)

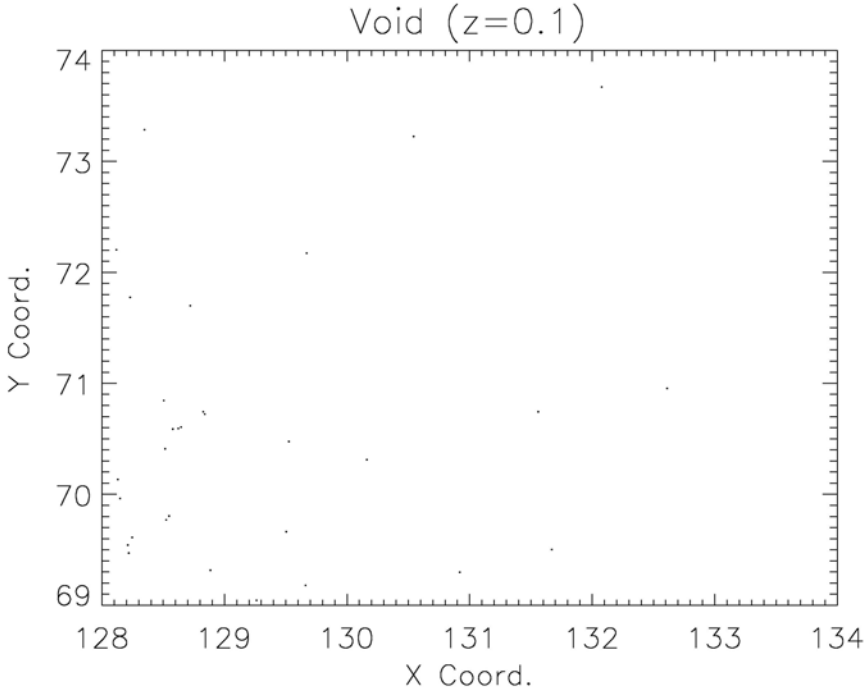

Largest Cluster (z=0.0)
Y Coord.
X Coord.
74
73
72
71
70
69
94
95
96
97
98
99
100


Void (z=0.0)
Y Coord.
X Coord.
74
73
72
71
70
69
128
129
130
131
132
133
134

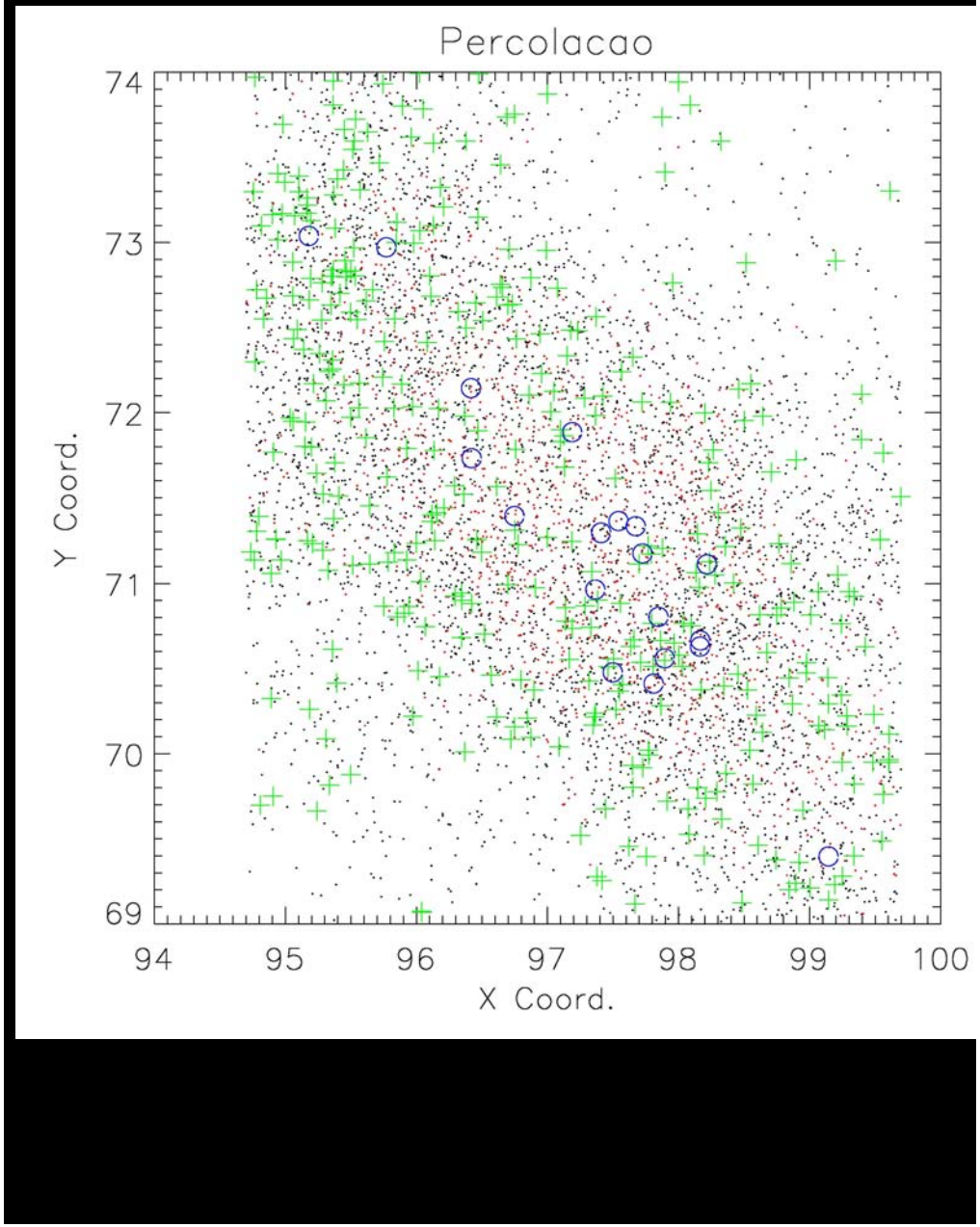

Primeiro Workshop BRAVO
DAS/LAC-INPE/MCT
5 e 6 de Fevereiro, 2007, São José dos Campos

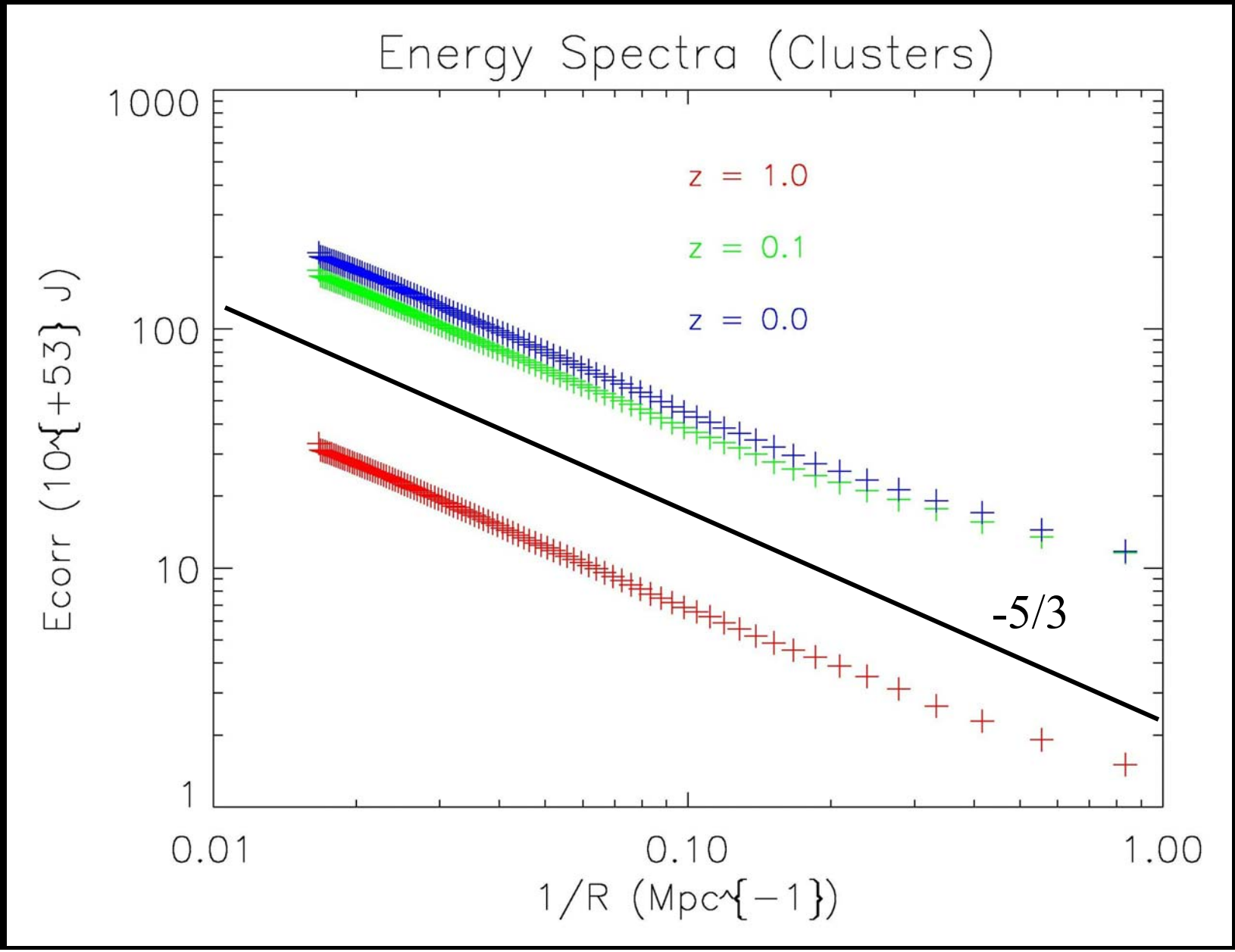

Percolacao
Y Coord.
X Coord.
74
73
72
71
70
69
94
95
96
97
98
99
100

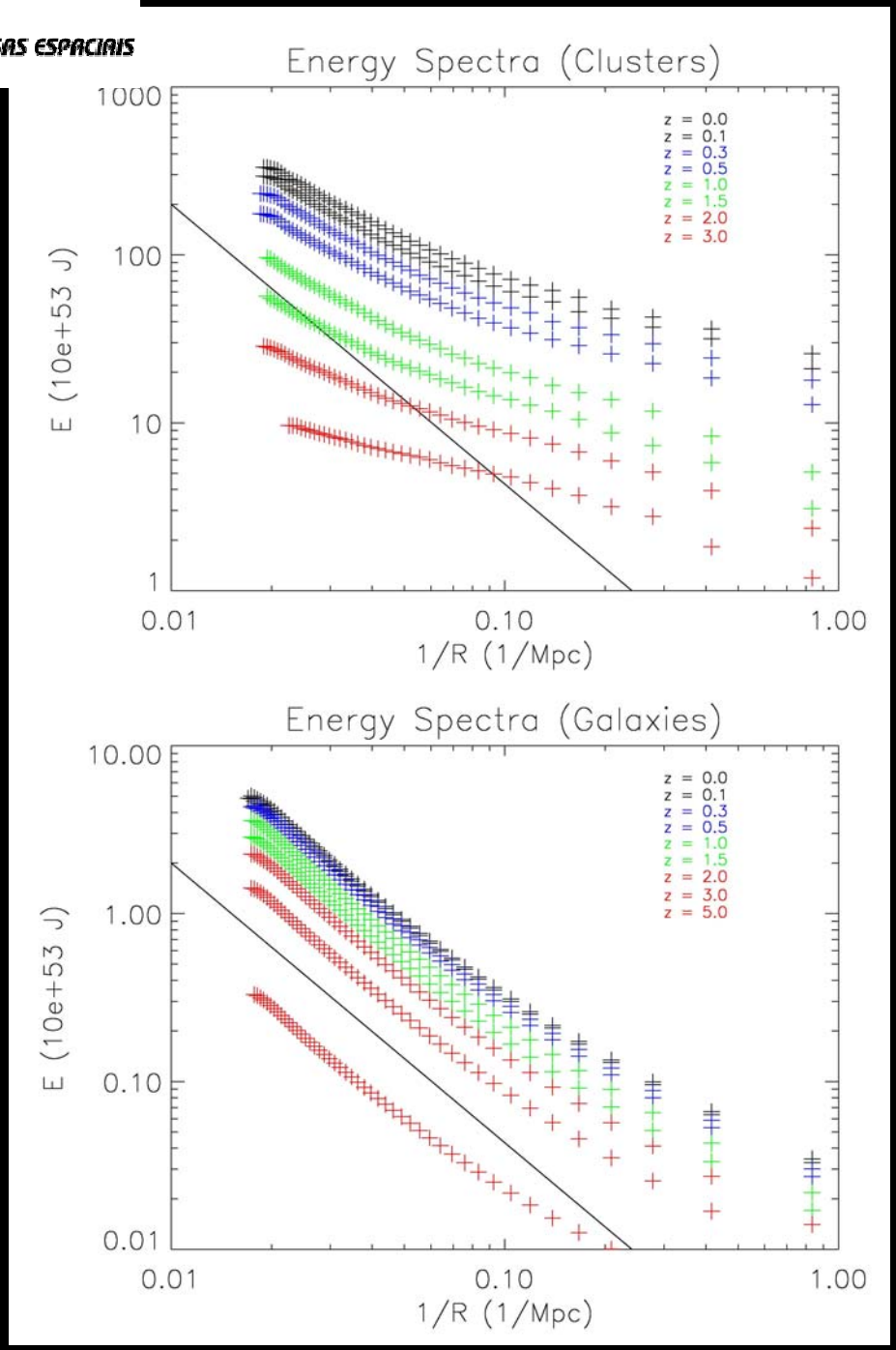

Largest Cluster (z=0.1)
Y Coord.
X Coord.
74
73
72
71
70
69
94
95
96
97
98
99
100


Void (z=0.1)
Y Coord.
X Coord.
74
73
72
71
70
69
128
129
130
131
132
133
134


Primeiro Workshop BRAVO
DAS/LAC-INPE/MCT
5 e 6 de Fevereiro, 2007, São José dos Campos

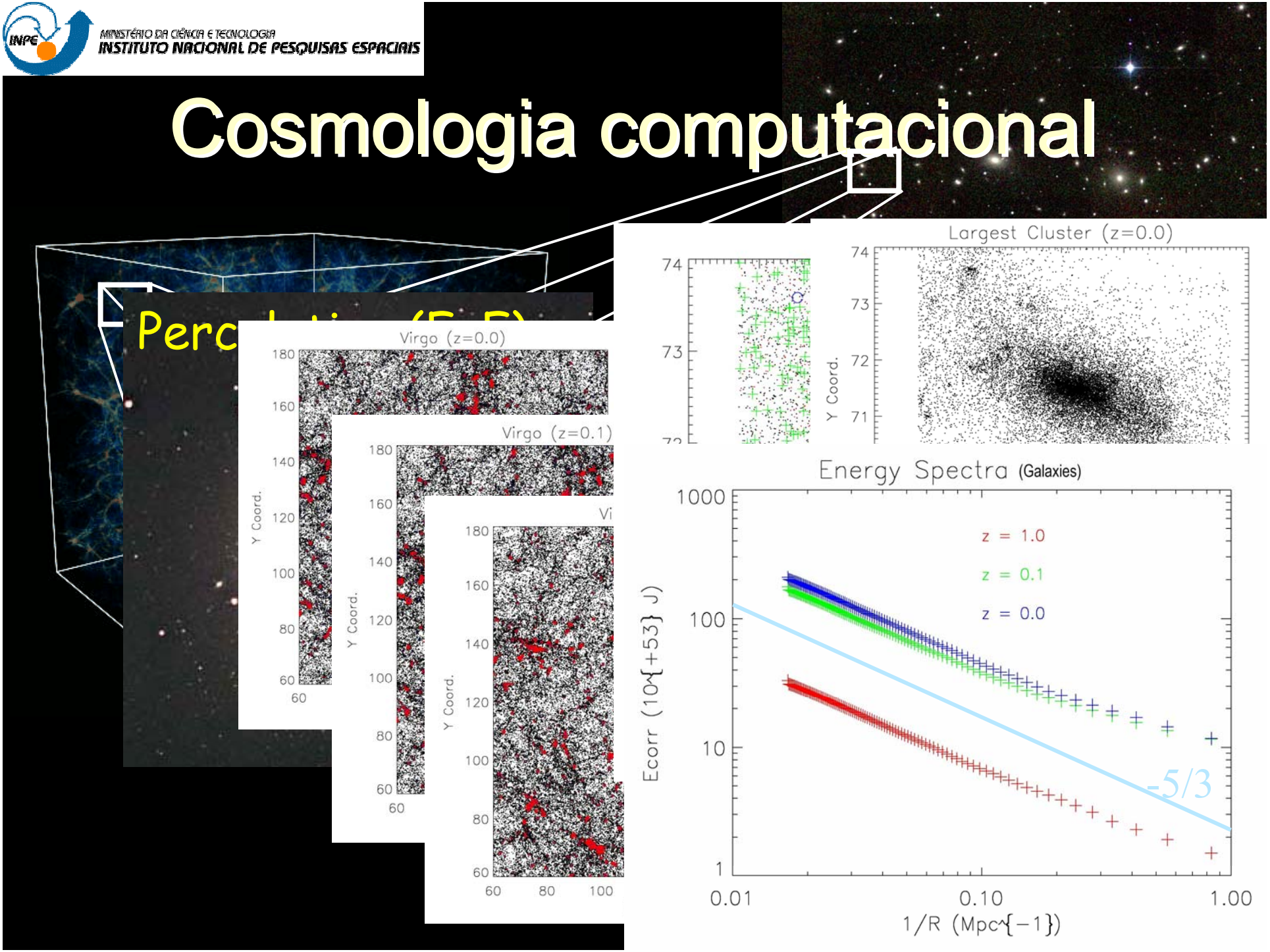

Percolacao
Y Coord.
X Coord.
74
73
72
71
70
69
94
95
96
97
98
99
100


Energy Spectra (Clusters)
1000
100
10
1
Ecorr (10^{+53} J)
z = 1.0
z = 0.1
z = 0.0
-5/3
0.01
0.10
1.00
1/R (Mpc^{-1})


Energy Spectra (Clusters)
E (10e+53 J)
1/R (1/Mpc)
1000
100
10
1
0.01
0.10
1.00
z = 0.0
z = 0.1
z = 0.3
z = 0.5
z = 1.0
z = 1.5
z = 2.0
z = 3.0
Energy Spectra (Galaxies)
E (10e+53 J)
1/R (1/Mpc)
10.00
1.00
0.10
0.01
0.01
0.10
1.00
z = 0.0
z = 0.1
z = 0.3
z = 0.5
z = 1.0
z = 1.5
z = 2.0
z = 3.0
z = 5.0

MINISTÉRIO DA CIÊNCIA E TECNOLOGIA
INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS

Cosmologia computacional



Muito obrigado!!