

Operações do SPRACE: DØ

Sergio M. Lietti

I Workshop do SPRACE

27/Março/2007

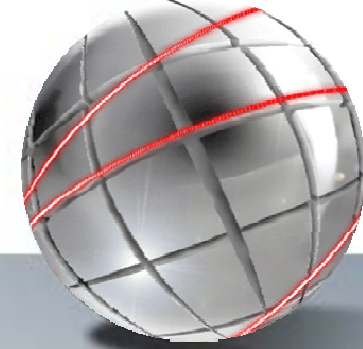


Sumário



- O Experimento DØ
- O Cluster SPRACE
 - suas três fases
- Operações para o DØ durante a primeira fase
 - Simulação de MC (mcfarm)
- Operações para o DØ durante a segunda fase
 - Simulação de MC e Reprocessamento P17 (SAMGrid)
- Operações para o DØ durante a terceira fase
 - Simulação de MC e Reprocessamento P20 (OSG)

O Experimento DØ



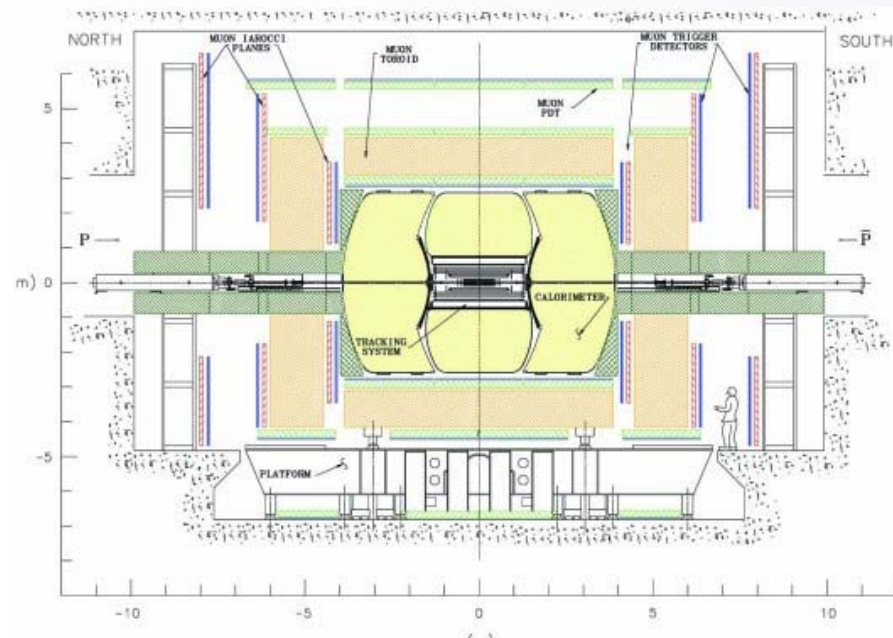
- O DØ é um detector do Tevatron, no Fermilab.
- O Tevatron é um anel de colisão próton-antipróton com energia de 2TeV no centro de massa.



O Experimento DØ



- O DØ é composto por quatro sub-detectores:
 - Detectores de trajetória e de vértice (SMT e CFT)
 - Calorímetros Eletromagnético e Hadrônico
 - Detectores de Múons
 - Dectector de Prótons Frontais (FPD)



O Experimento DØ



- Principais objetivos
 - Medidas precisas do Modelo Padrão
 - Busca do Higgs
 - Busca por novos modelos: SS, ED, etc.
- Grande quantidade de dados
- Processamento e Simulação dos dados em curto espaço de tempo
- Necessidade de clusters de computadores além dos existentes no Fermilab

O Cluster SPRACE



- São Paulo Regional Analysis Center (SPRACE)
- Cluster do Grupo de Física Experimental de Altas Energias do IFT/UNESP
- Principais operações para os experimentos DØ do Fermilab e CMS do CERN.
- Financiado por Projeto Temático da FAPESP.
- Três fases de instalação (2004, 05 e 06).

O Cluster SPRACE – Fase 1



- Instalação em março de 2004
 - 1 head node (Itautec – 4 X Intel Xeon 2.4 GHz)
 - 22 worker nodes (Itautec - 2 X Intel Xeon 2.4 GHz)
 - 1 storage head node (Dell- 2 X Intel Xeon 2.4 GHz)
 - 2 storage servers com 4 TB
 - 4 nobreaks de 3KVA cada
 - 1 switch (DLink - 24 x - Ethernet 10/100/1000 Base-T)
 - 2 racks
 - Ar condicionado: 2 x 36.000BTU
 - Instalações elétricas, piso, etc.

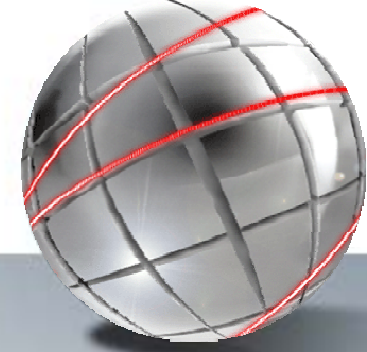
O Cluster SPRACE – Fase 1



- O cluster na fase 1

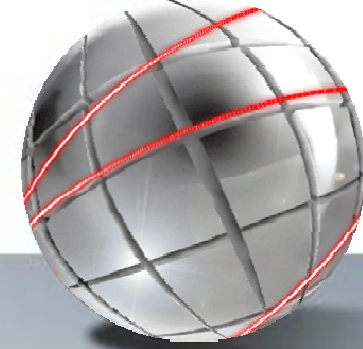


O Cluster SPRACE – Fase 2

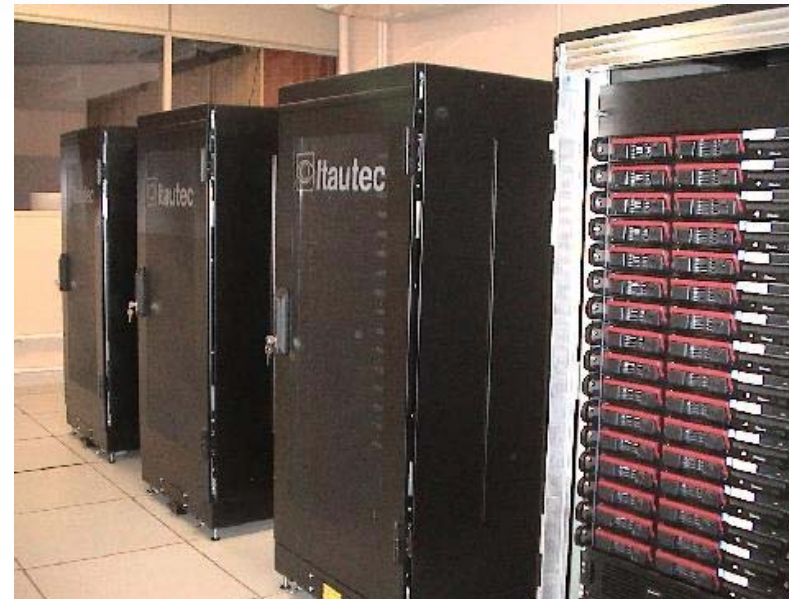


- Instalação em junho de 2005
 - 32 worker nodes (Itautec - 2X64Bit IntelXeon 3.0GHz)
 - 1 storage head node (Dell - 2 64Bit IntelXeon 3.0GHz)
 - 2 storage servers com 8 TB
 - 4 nobreaks de 3KVA cada
 - 1 switch (DLink - 24 x - Ethernet 10/100/1000 Base-T)
 - 1 switch (3com - 24 x - Ethernet 10/100/1000 Base-TX)
 - 2 racks
 - Ar condicionado: 1 x 80.000BTU

O Cluster SPRACE – Fase 2



- O cluster na fase 2

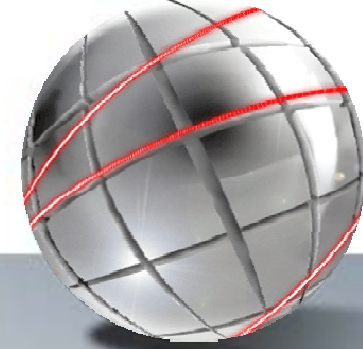


O Cluster SPRACE – Fase 3



- Instalação em outubro de 2006
 - 32 worker nodes (Itautec - 2 Intel Xeon Dual-Core 2.0 GHz)
 - 4 nobreaks de 3KVA cada
 - 2 switchs (3com - 24 x - Ethernet 10/100/1000 Base-T/RJ-45 , SFP ports)
 - 2 racks.

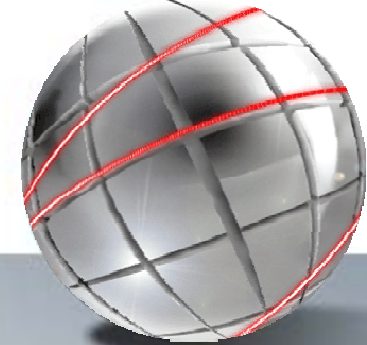
O Cluster SPRACE – Fase 3



- O cluster na fase 3



O Cluster SPRACE – Atualmente



- 244 CPU's com 1 GB RAM/CPU
 - 50 CPU's (Intel Xeon 2.4 GHz)
 - 66 CPU's (Intel Xeon 64Bit 3.0GHz)
 - 128 CPU's (Intel Xeon Dual-Core 2.0 GHz)
- 12 TB de armazenamento
- Link de 1GBit com NAP do Brasil

Operações para o DØ - primeira fase

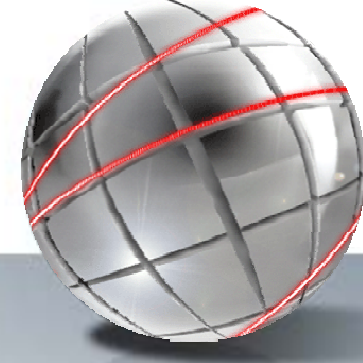


- O SPRACE iniciou suas operações para o DØ contribuindo na Produção de Monte Carlo
- Instalação do software do DØ foi necessária:
 - Uma estação do SAM (*Sequential Access Model*) foi instalada na SPRACE

O SAM é um sistema de manuseio de dados organizado como um conjunto de servidores geograficamente distribuídos que operam em conjunto para armazenar e acessar arquivos

A estação *sprace* possuía um cache de 200GB e serviu prioritariamente para enviar os dados de simulação de Monte Carlo do experimento para o armazenamento em fita no Fermilab e disponibilizá-los para o uso pela colaboração.

Operações para o DØ - primeira fase



– Instalação do McFarm

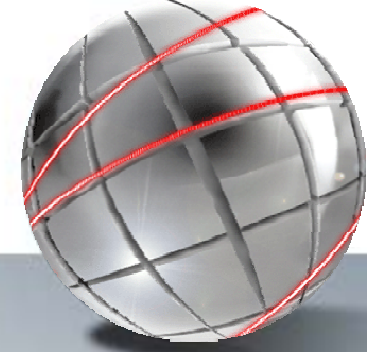
O McFarm (Monte Carlo Farm Manager) é um programa mantido pela UTA com a finalidade de gerenciar a produção de Monte Carlo do DZero em um cluster local de computadores com sistema operacional Linux.

Suas principais funções: paralelização de grandes programas, a monitoração inteligente com correção automática de problemas, a notificação de problemas por e-mail, e a disposição automática dos resultados.

Os programas executáveis são gerados pelo McRunjob (Monte Carlo Run Job), ferramenta da colaboração DØ para gerenciar a produção de dados via simulação de Monte Carlo.

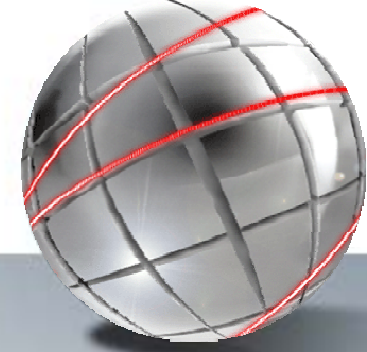
O SAM é utilizado para obter os arquivos necessários à simulação e armazenar o resultado.

Operações para o DØ - primeira fase



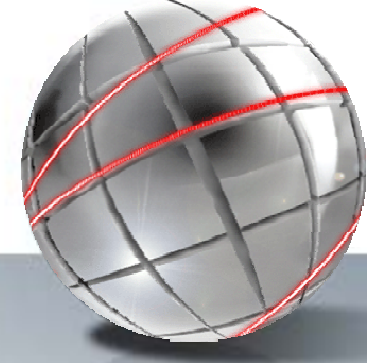
- Nesta primeira fase foram realizadas 84 simulações totalizando aproximadamente 3.5 milhões de eventos.
- O período foi de Julho de 2004 a Junho de 2005

Operações para o DØ - segunda fase



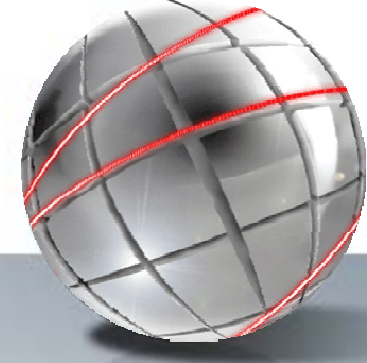
- Durante a segunda fase do SPRACE, as principais contribuições para o DØ foram:
 - Reprocessamento P17 dos dados do DØ
 - Produção de Monte Carlo
- Para tanto foi necessária a instalação do SAMGrid, primeiro projeto de Grid a ser efetivamente utilizado pelos experimentos de Física de Altas Energias.

Operações para o DØ - segunda fase



- A arquitetura do SAM-Grid é composta por três grandes componentes: o manuseio dos dados, o gerenciamento do processamento e o gerenciamento da informação.
 - O manuseio dos dados é feito através do uso do SAM.
 - O gerenciamento do processamento é feito através dos componentes do JIM (*Job and Information Management*), desenvolvido pela equipe da divisão de computação do Fermilab.
 - A parte de monitoramento e busca de recursos é baseada em ferramentas desenvolvidas pelo Condor-G e Globus Toolkit, e utiliza informações compartilhadas em formato XML.

Operações para o DØ - segunda fase



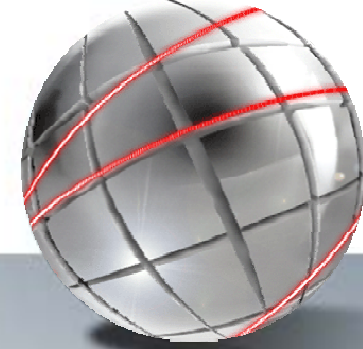
- A autenticação dos usuários habilitados a usar os recursos do site é feita através de certificados emitidos pelo DOE-Grid, agência do departamento de energia dos EUA (DOE).
- Os usuários e máquinas certificadas a usar os recursos locais são identificados através de seu registro em um arquivo presente em todas as máquinas pertencentes ao Grid.

Operações para o DØ - segunda fase

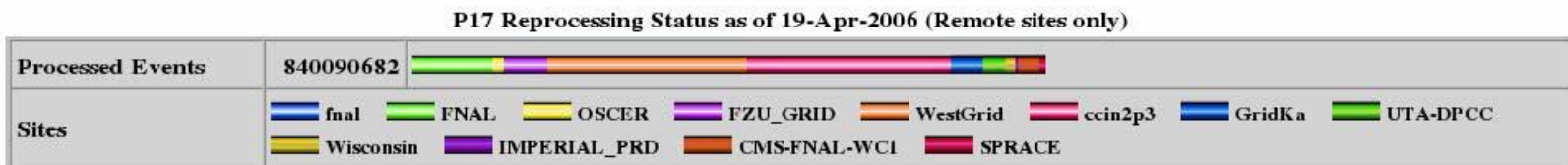
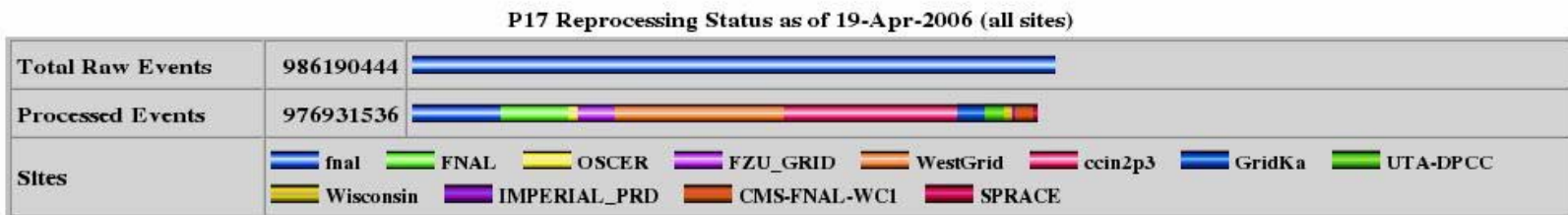


- Para o reprocessamento P17 dos dados do DØ ainda foi necessário:
 - **Certificação da produção:** Esta certificação visa confirmar que os dados reprocessados em um determinado local sejam idênticos aos dados reprocessados em outro local. Para isto um conjunto de dados padrão é processado em todos os sites e deve fornecer resultados idênticos.
 - **Certificação da consolidação:** Após serem processados, os resultados de vários eventos são consolidados em um único arquivo onde a descrição desse conjunto de eventos é armazenada. Esses arquivos consolidados (*merged*) são os usados nas análises. Neste caso, conjuntos de dados diferentes são processados em cada site e o conteúdo do arquivo consolidado é comparado com o conteúdo dos arquivos que lhe deram origem.

Operações para o DØ - segunda fase



- O reprocessamento P17 dos dados do DZero no SPRACE teve início em 18 de Agosto de 2005
- Foram reprocessados 4.253 arquivos de dados brutos, contendo ao todo 9.206.931 eventos.
- Para processar estes dados, foram transferidos cerca de 3.120 Gigabytes de dados do Fermilab para nosso cache local.

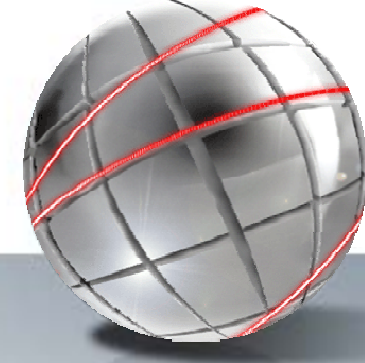


Operações para o DØ - segunda fase

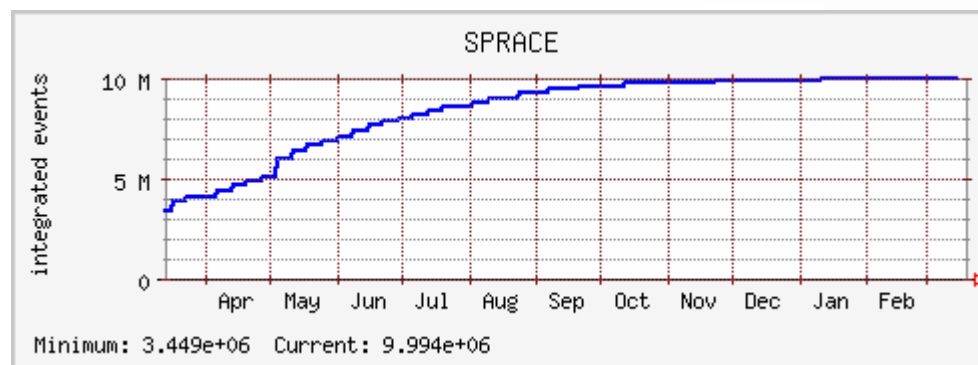
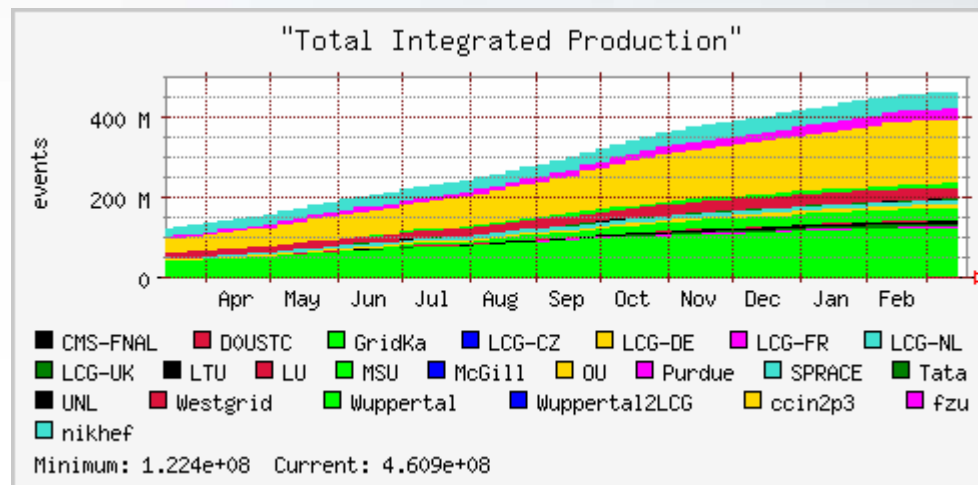


- A Produção de Monte Carlo para o DØ na segunda fase do cluster iniciou-se em Julho de 2005
- Agora usando a infra-estrutura do SAMGrid ao invés do McFarm
- Produzimos no SPRACE mais de 5 milhões de eventos de Monte Carlo em 1 ano, num total de aproximadamente 300 GB de dados transferidos

Operações para o DØ - segunda fase



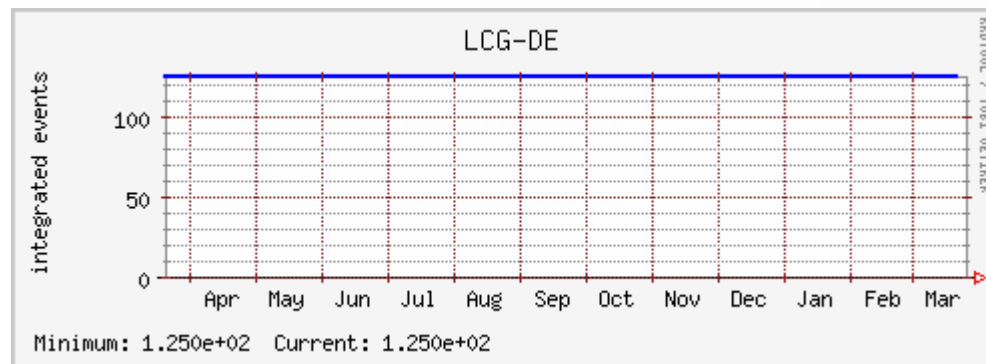
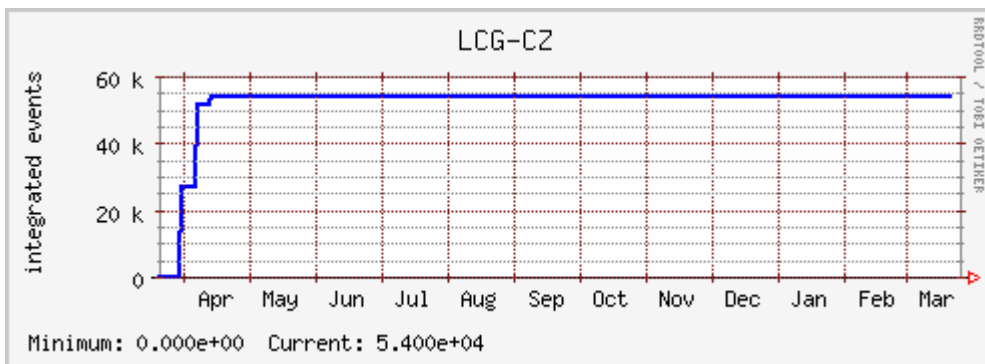
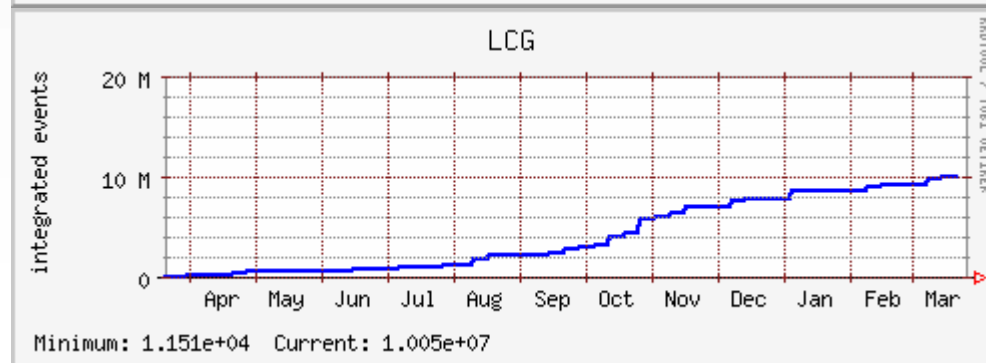
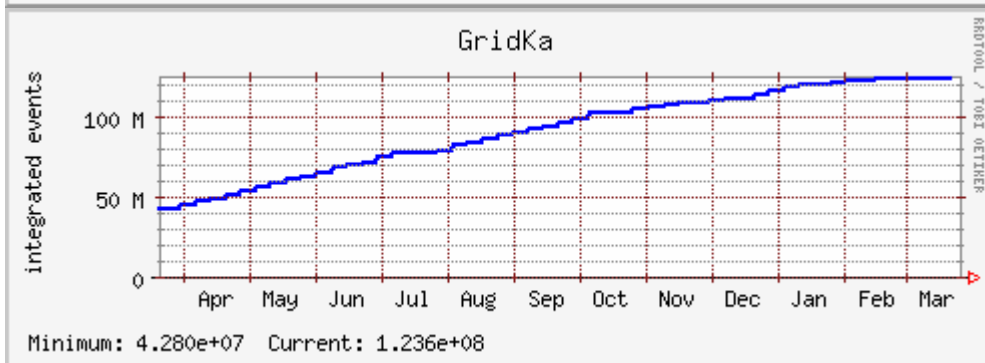
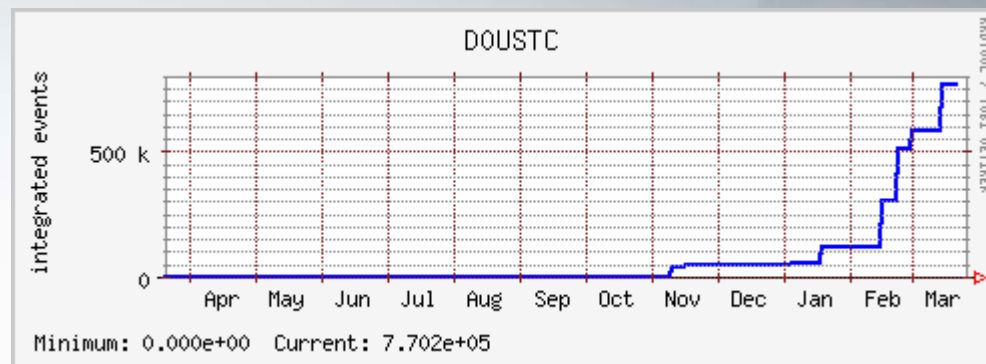
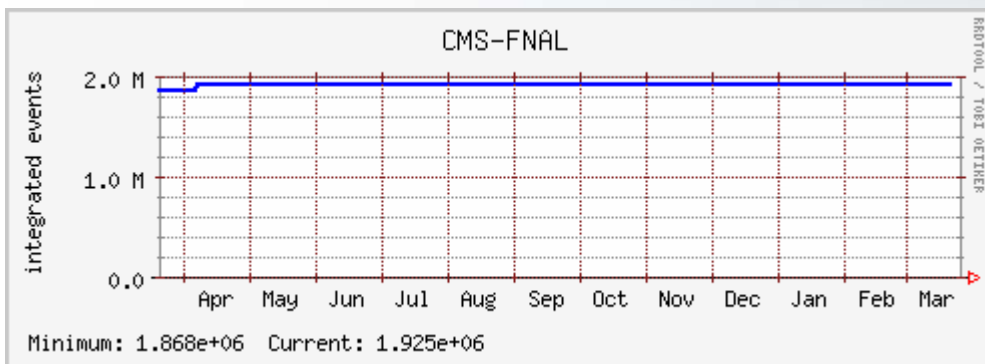
- Produção de Monte Carlo



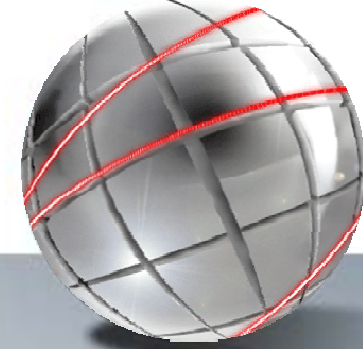
Operações para o $D\emptyset$ - segunda fase



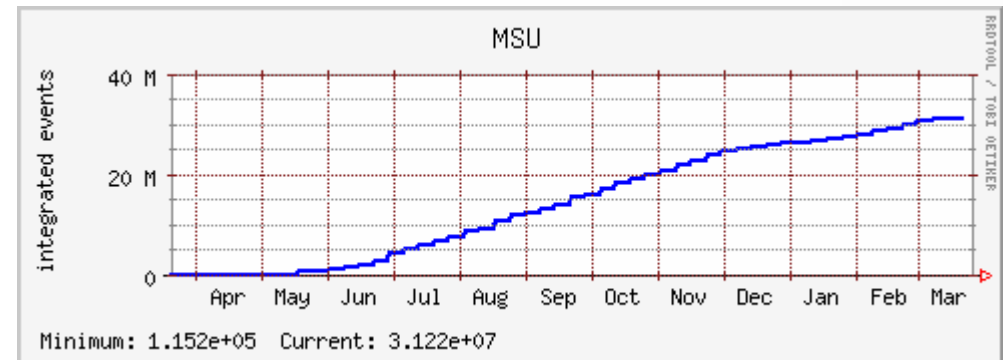
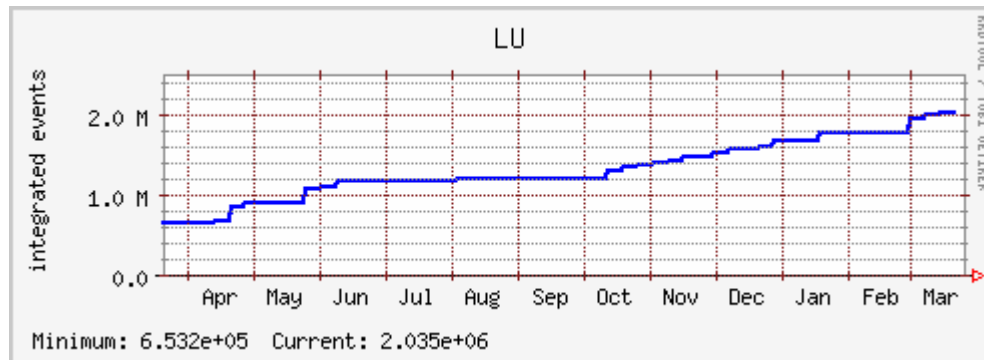
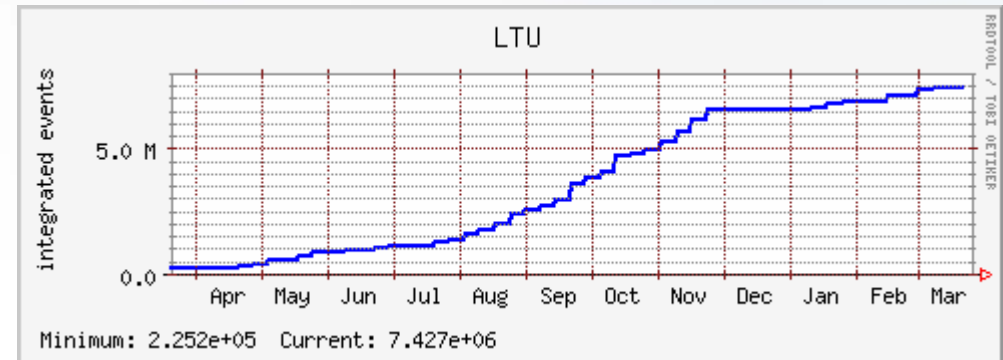
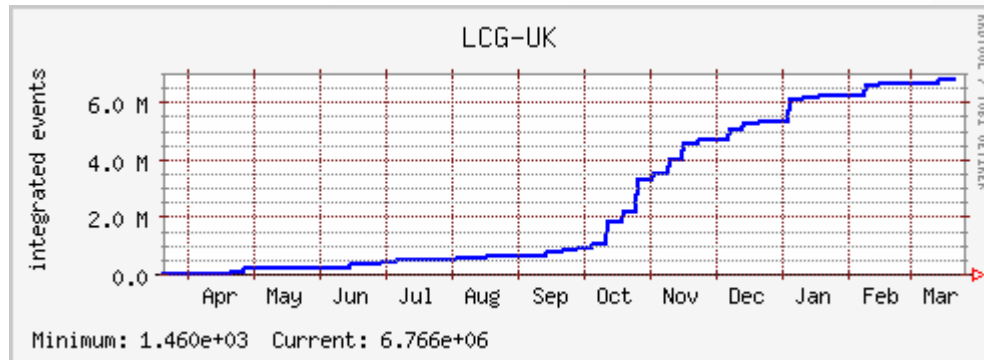
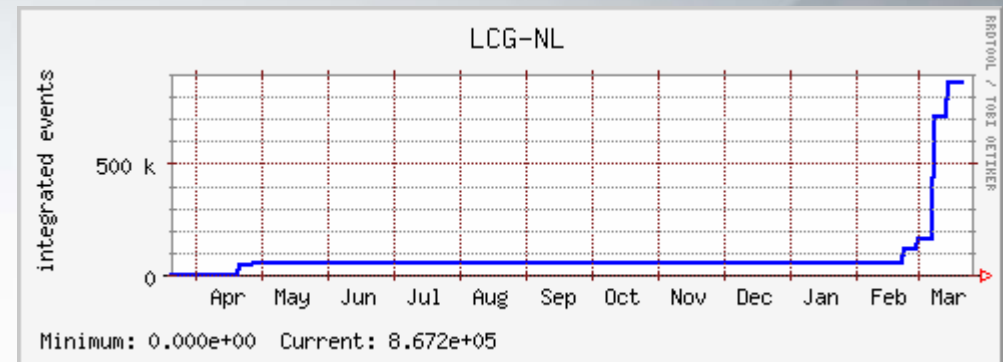
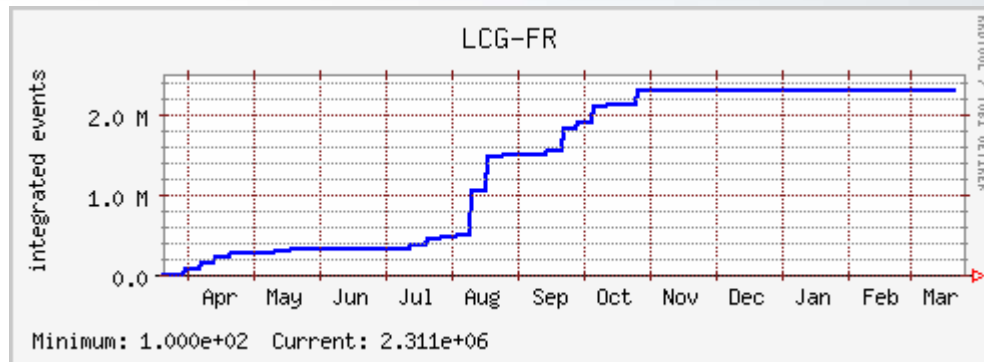
- Produção de MC em outros clusters



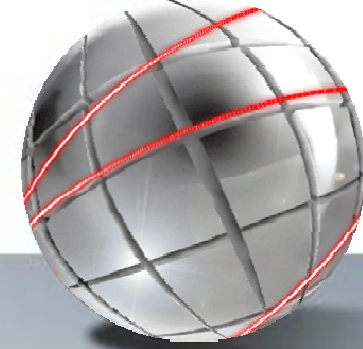
Operações para o DØ - segunda fase



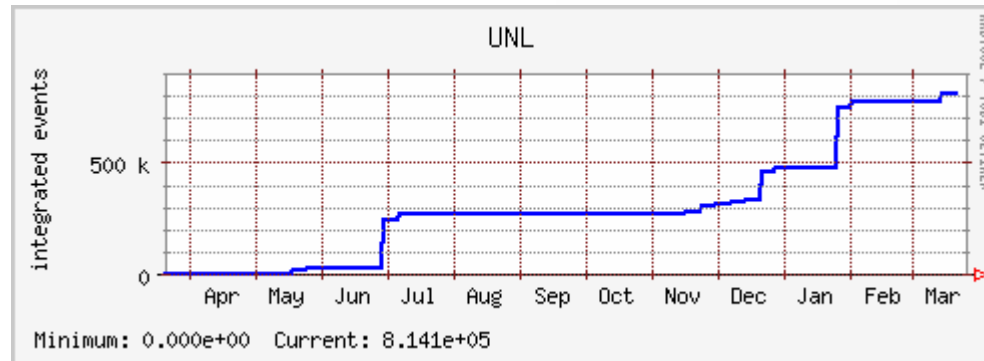
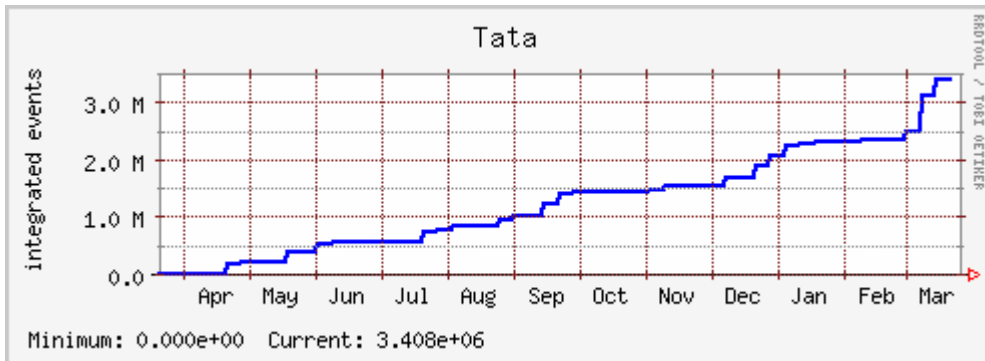
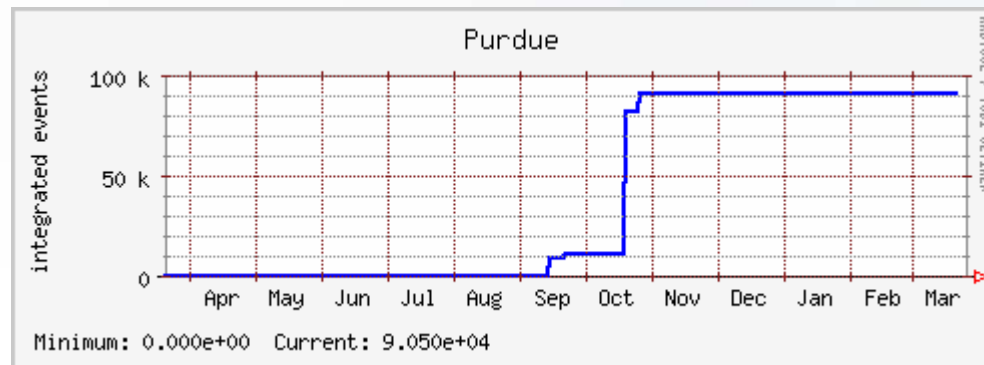
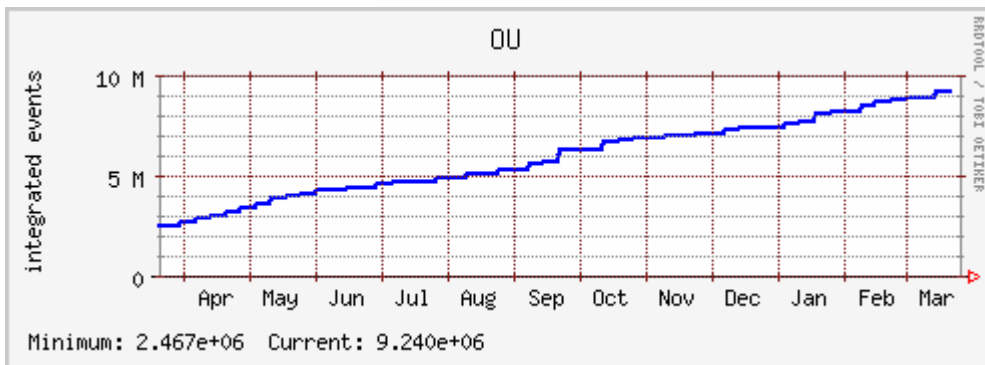
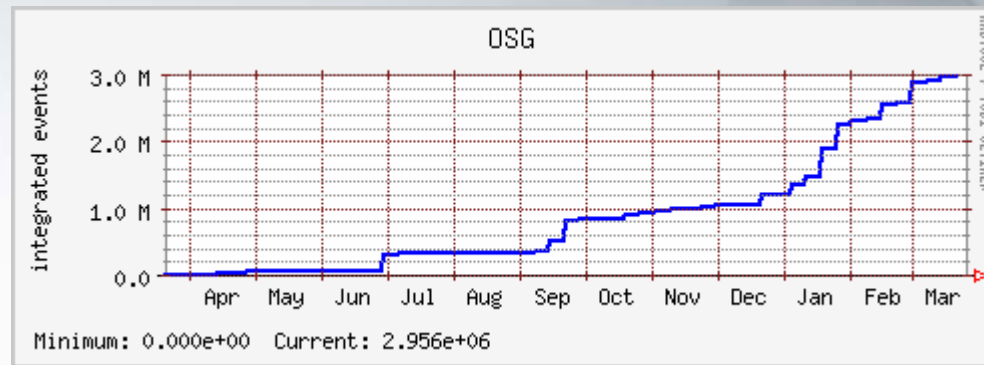
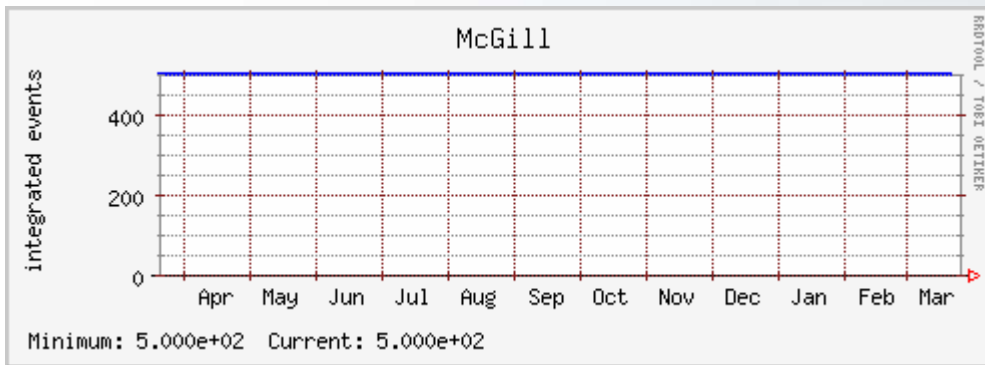
- Produção de MC em outros clusters



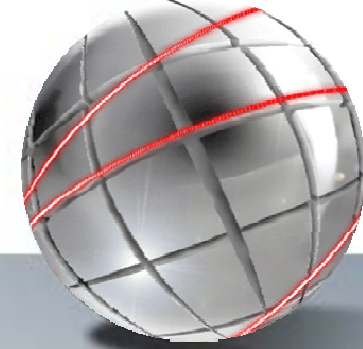
Operações para o DØ - segunda fase



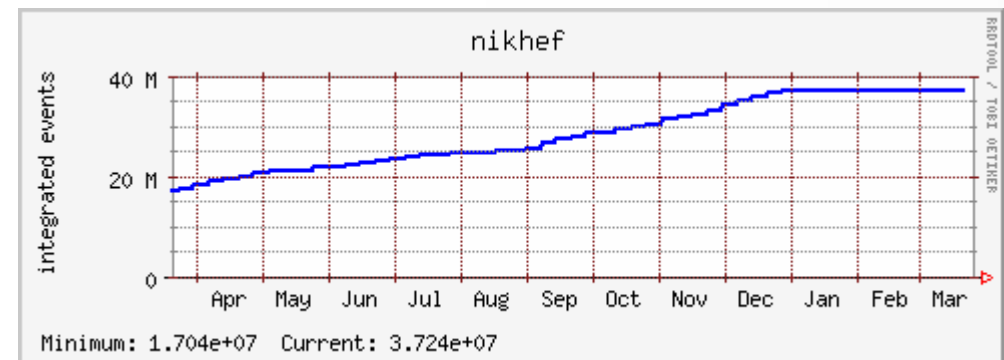
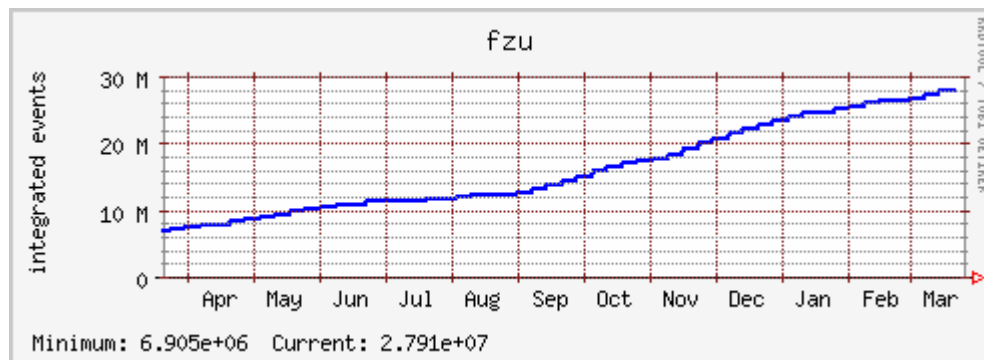
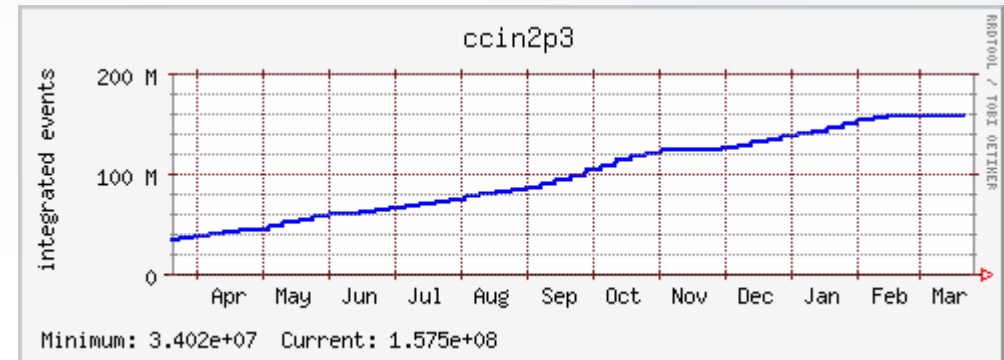
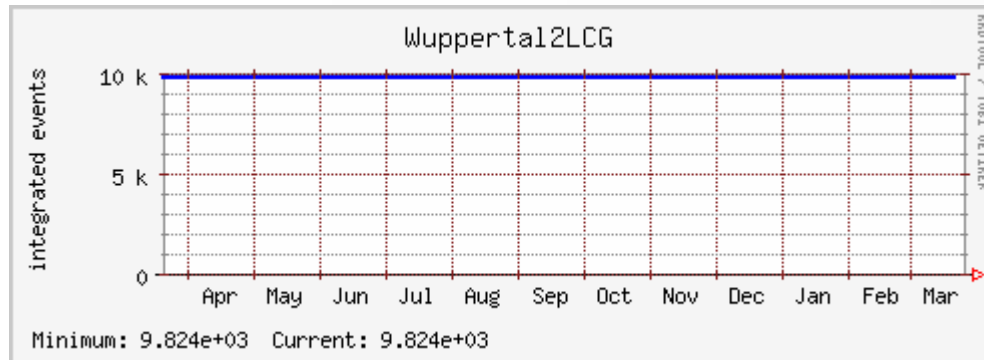
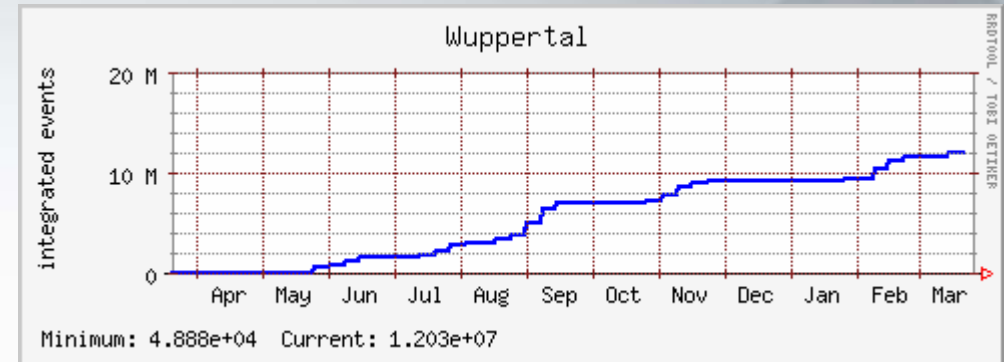
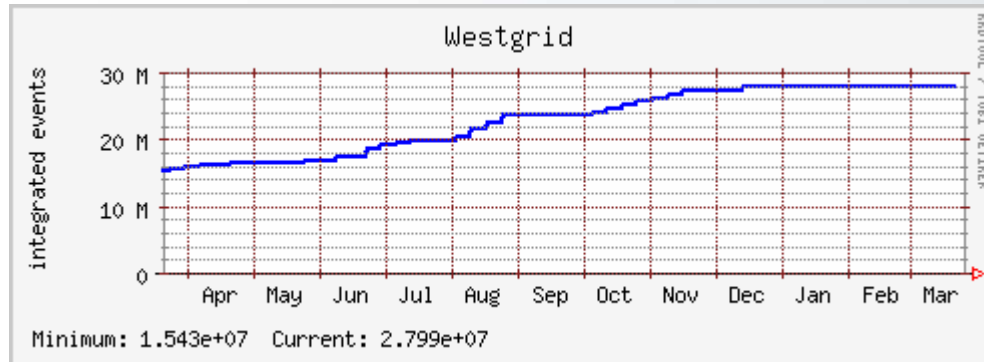
- Produção de MC em outros clusters



Operações para o DØ - segunda fase



- Produção de MC em outros clusters



Operações para o DØ - terceira fase



- Devido as operações do experimento CMS do CERN, o SPRACE integrou-se ao OSG
- Além da instalação do software do OSG, foi necessário um upgrade do SO de todas as máquinas
- As operações do DØ são possíveis graças a interoperabilidade entre o SAMGrid e o OSG
- Atualmente o SPRACE participa do Reprocessamento P20 de dados do DØ
- Após o final do Reprocessamento, o SPRACE contribuirá para a Produção de Monte Carlo

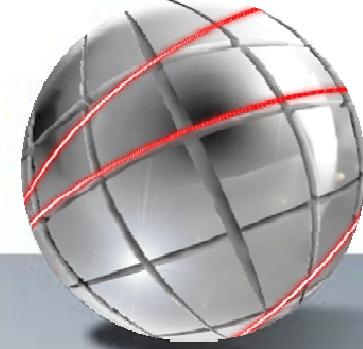
Operações para o DØ - terceira fase



- O Reprocessamento P20 dos dados do DØ está em andamento
- 500 M de eventos devem ser reprocessados até Abril.
- Estimou-se a necessidade de 2300 CPU's para a tarefa
- Necessidade do uso de clusters do OSG
- SPRACE e outros clusters OSG estão participando do reprocessamento

OSG Clusters	# of CPU's
Oklahoma University	200
Indiana University	250
NERSC LBL	250
University of Nebraska Lincoln	256
CMS FNAL	250
SPRACE Brasil	232
CC-IN2P3 Lyon	500
LOUISIANA LTU-CCT at LSU	250 (128)
UCSD - San Diego	300 (70)
PURDUE-ITaP	600 (?)

Operações para o DØ - terceira fase



- Status do Reprocessamento P20 dos dados do DØ até 25/Mar/2007

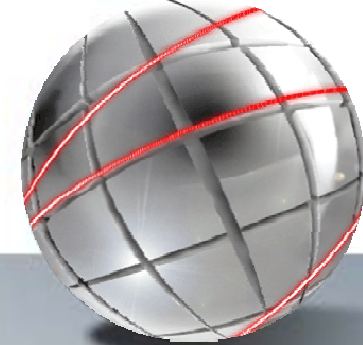
P20 (Re)processing Status as of 25-Mar-2007 (all sites)

Total Raw Events	748602675	
Merged Events	289257304	
Contribution		p20.08 p20.07

P20.07.01 Reprocessing Status as of 25-Mar-2007 (Remote sites only)

Unmerged Events	120323529	
Merged Events	85082939	
Contribution		

Operações para o DØ



- O SPRACE foi capaz contribuir de forma significativa para o experimento DØ
- Comparado aos demais clusters remotos, teve e vem tendo desempenho importante na Produção de Monte Carlo e no Reprocessamento dos dados