



MINISTÉRIO DA CIÊNCIA E TECNOLOGIA
INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS



APERFEIÇOAMENTO E DESENVOLVIMENTO DE FERRAMENTAS DO CONTROLE DA AUTOMATIZAÇÃO DO BACKUP E RECUPERAÇÃO DOS DADOS DE SATÉLITE EM FITA

RELATÓRIO FINAL DE INICIAÇÃO TECNOLÓGICA
(PIBIT/CNPq/INPE)

Matheus da Cruz Oliveira dos Santos (Fatec Cruzeiro, Bolsista PIBIT/CNPq)

E-mail: matheus.santos@cptec.inpe.br

Daniel Alejandro Vila (DSA/CPTEC/INPE, Orientador)

E-mail: dvila@ina.gov.ar

Julho de 2012

1. RESUMO DO PLANO INICIAL

O objetivo principal deste projeto é desenvolver uma ferramenta para agilizar e facilitar o procedimento de backup e recuperação dos dados de satélite em fitas LTO e DLT.

Hoje a necessidade por dados e imagens de satélites é de grande importância para diversos setores socioeconômicos, com grande número de aplicações ambientais nos campos da Meteorologia e Clima, Agricultura, Ciências Florestais, Hidrologia, Oceanografia, Defesa Civil etc. No entanto, poucas são as instituições que disponibilizam continuamente essas informações. Dados e imagens de Satélites são processados por sistemas especializados para tal e em seguida registrados em meio magnético para armazenamento e uso atual. Com o surgimento de novos satélites, há a necessidade de automatização do processo de armazenamento dos dados. Ai entra o objetivo principal citado, durante o desenvolvimento também deverá ocorrer o backup manual.

Diversas instituições, públicas e privadas, vêm adotando o uso de software livre, aumentando o investimento em tecnologia com baixos custos. As principais ferramentas que serão utilizadas no desenvolvimento do sistema são: Java e Shell Scripts. O trabalho será desenvolvido utilizando apenas ferramentas livre que permitem aplicar a um ambiente Linux/Unix, mesmo com um grande volume de dados, permitindo uma grande redução nos custos.

Dentre os objetivos específicos do trabalho, pode-se destacar:

- 1- Aprofundar os conhecimentos no ambiente Linux/Unix, PostgreSQL e Javaserwer Pages(JSP);
- 2- Desenvolver Shell Scripts para armazenar e recuperar os dados armazenados em mídias DLT/LTO;
- 3- Aprofundar os conhecimentos em linguagem de programação Java Web e Shell Scripts;
- 4- Desenvolver interface para automatizar o processo de backup armazenamento de dados no banco de Dados de Imagens (BDI).

2. RESUMO DAS ETAPAS DE DESENVOLVIMENTO

Etapa 1.

- Estudo da linguagem Shell adotando a apostila “Introdução ao LINUX e Programação em Script-Shell”¹.
- Leitura e execução dos exercícios contidos na apostila.

Etapa 2

- Análise do funcionamento das rotinas diárias referentes a gravação, cadastro e atualização de mídias LTO e DLT.
- Gravações diárias dos dados das unidades de armazenamento SAMFS1 e SAMFS2 em fitas LTO para armazenamento e liberação do espaço ocupado.
- Criação de etiquetas para catalogação das novas fitas.

Etapa 3

- Desenvolvimento e testes de scripts em linguagem CShell para analisar e gerar listagens que exibem o espaço ocupado pelos subprodutos dos satélites Meteosat8, Goes12 e NOAA nas unidades de armazenamento SAMFS1 e SAMFS2.

Etapa 4

- Estudo da linguagem Java EE adotando a apostila “Java para Desenvolvimento Web”².

3. MATERIAL E MÉTODOS

3.1. Etapas Desenvolvidas

O diagrama a seguir (fig. 01), ilustra o caminho percorrido pela imagem desde sua geração pelo satélite até a unidade de armazenamento (storage). As imagens ficam disponíveis no website.

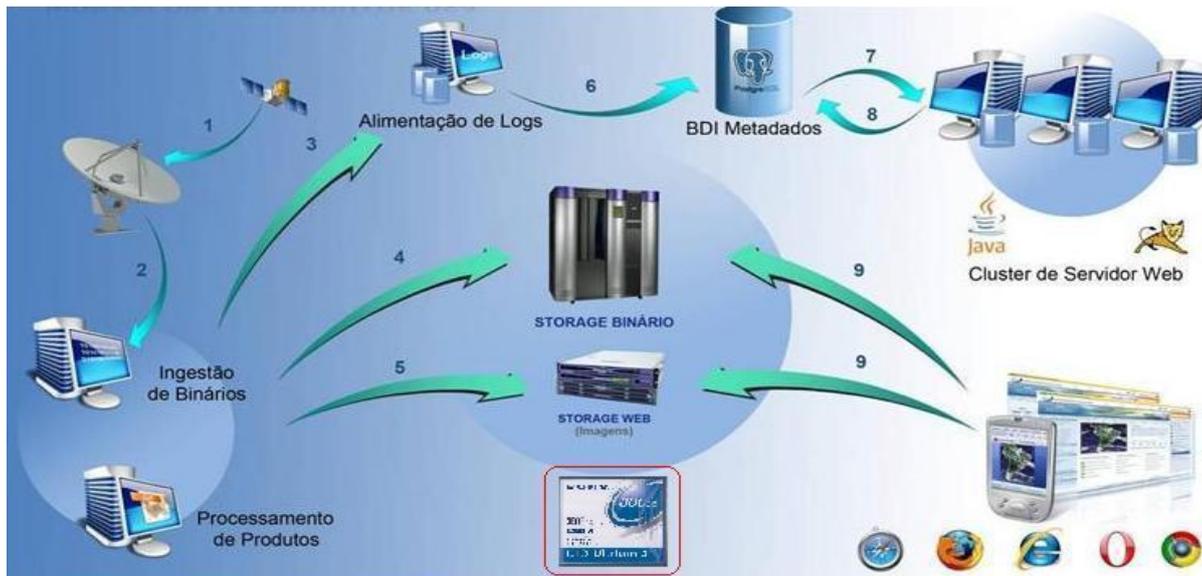


Figura1. Caminho das imagens de satélite, no site do DSA é disponibilizado o BDI com as imagens de diferentes satélites para serem consultadas, sendo essas sempre atualizadas.

Resumo das etapas:

- 1- Recepção do satélite
- 2- Recepção do arquivo binário
- 3- Geração de log para a alimentação do banco.
- 4- Armazenamento em storage binário.
- 5- Armazenamento do produto em Storage web
- 6- Alimentação dos Metadados (Log de Produtos)
- 7- Aplicação web busca imagens no banco de dados
- 8- BDI retorna link físico da imagem/binário
- 9- Aplicação acessa o local correto da imagem/binário selecionando o Storage Web ou Storage Binário

A próxima etapa é onde começa o backup manual em si. A máquina receptora é acessada remotamente, então é localizado o diretório em que se encontram os dados a serem armazenados. É preciso verificar o tamanho do diretório para designar também o tamanho da listagem e determinar a quantidade de arquivos (em GIGABYTES), que serão armazenados. Após isso a listagem é gerada e apendada ou gravada na fita LTO 2 ou LTO 3. Então é dado o comando para iniciar a gravação, através de script ou linha de comando dentro do terminal Linux.

Caso a gravação seja bem sucedida, os dados são apagados automaticamente pelo script, que verifica a listagem criada com todos os dados gravados e os dados armazenados em disco, caso estejam iguais são apagados da unidade storage.

A finalidade é gerar mais espaço para receber novas imagens, que são constantes. Essas imagens caem de minutos em minutos então se imagina o volume destas. Todos os passos de backup manual são realizados através de scripts, desde a verificação do espaço em disco, até a exclusão dos arquivos.

3.2 Desenvolvimento: Backup

A listagem a seguir (fig. 02), apresenta o diretório e subdiretório que o dado a ser gravado em fita se encontra. O tamanho da listagem é designado pelo usuário, que na hora de gerar a listagem escolhe a data (no caso todo o mês 01 de 2012) ou o volume em Gigabytes a ser gravado.

```
/samfs1/Lista_met8_met8_cal_ret_ch7_bin - hmsdsa
samfsl/DSA/met8/met8_cal_ret/ch7_bin/2012/01/S11117606_201201010000.gz
samfsl/DSA/met8/met8_cal_ret/ch7_bin/2012/01/S11117606_201201010015.gz
samfsl/DSA/met8/met8_cal_ret/ch7_bin/2012/01/S11117606_201201010030.gz
samfsl/DSA/met8/met8_cal_ret/ch7_bin/2012/01/S11117606_201201010045.gz
samfsl/DSA/met8/met8_cal_ret/ch7_bin/2012/01/S11117606_201201010100.gz
samfsl/DSA/met8/met8_cal_ret/ch7_bin/2012/01/S11117606_201201010115.gz
samfsl/DSA/met8/met8_cal_ret/ch7_bin/2012/01/S11117606_201201010130.gz
samfsl/DSA/met8/met8_cal_ret/ch7_bin/2012/01/S11117606_201201010145.gz
samfsl/DSA/met8/met8_cal_ret/ch7_bin/2012/01/S11117606_201201010200.gz
samfsl/DSA/met8/met8_cal_ret/ch7_bin/2012/01/S11117606_201201010215.gz
samfsl/DSA/met8/met8_cal_ret/ch7_bin/2012/01/S11117606_201201010230.gz
samfsl/DSA/met8/met8_cal_ret/ch7_bin/2012/01/S11117606_201201010245.gz
samfsl/DSA/met8/met8_cal_ret/ch7_bin/2012/01/S11117606_201201010300.gz
samfsl/DSA/met8/met8_cal_ret/ch7_bin/2012/01/S11117606_201201010315.gz
samfsl/DSA/met8/met8_cal_ret/ch7_bin/2012/01/S11117606_201201010330.gz
samfsl/DSA/met8/met8_cal_ret/ch7_bin/2012/01/S11117606_201201010345.gz
samfsl/DSA/met8/met8_cal_ret/ch7_bin/2012/01/S11117606_201201010400.gz
samfsl/DSA/met8/met8_cal_ret/ch7_bin/2012/01/S11117606_201201010415.gz
samfsl/DSA/met8/met8_cal_ret/ch7_bin/2012/01/S11117606_201201010430.gz
samfsl/DSA/met8/met8_cal_ret/ch7_bin/2012/01/S11117606_201201010445.gz
samfsl/DSA/met8/met8_cal_ret/ch7_bin/2012/01/S11117606_201201010500.gz
samfsl/DSA/met8/met8_cal_ret/ch7_bin/2012/01/S11117606_201201010515.gz
samfsl/DSA/met8/met8_cal_ret/ch7_bin/2012/01/S11117606_201201010530.gz
samfsl/DSA/met8/met8_cal_ret/ch7_bin/2012/01/S11117606_201201010545.gz
samfsl/DSA/met8/met8_cal_ret/ch7_bin/2012/01/S11117606_201201010600.gz
samfsl/DSA/met8/met8_cal_ret/ch7_bin/2012/01/S11117606_201201010615.gz
samfsl/DSA/met8/met8_cal_ret/ch7_bin/2012/01/S11117606_201201010630.gz
samfsl/DSA/met8/met8_cal_ret/ch7_bin/2012/01/S11117606_201201010645.gz
samfsl/DSA/met8/met8_cal_ret/ch7_bin/2012/01/S11117606_201201010700.gz
samfsl/DSA/met8/met8_cal_ret/ch7_bin/2012/01/S11117606_201201010715.gz
samfsl/DSA/met8/met8_cal_ret/ch7_bin/2012/01/S11117606_201201010730.gz
samfsl/DSA/met8/met8_cal_ret/ch7_bin/2012/01/S11117606_201201010745.gz
samfsl/DSA/met8/met8_cal_ret/ch7_bin/2012/01/S11117606_201201010800.gz
samfsl/DSA/met8/met8_cal_ret/ch7_bin/2012/01/S11117606_201201010815.gz
samfsl/DSA/met8/met8_cal_ret/ch7_bin/2012/01/S11117606_201201010830.gz
samfsl/DSA/met8/met8_cal_ret/ch7_bin/2012/01/S11117606_201201010845.gz
samfsl/DSA/met8/met8_cal_ret/ch7_bin/2012/01/S11117606_201201010900.gz
```

Line: 1/2968 Column: 1 Character: 115 (0x73)

Figura 2. Listagem já pronta para a gravação

O script para gravação visa o objetivo principal, mais segurança e agilidade. Como pode ser visto abaixo (fig. 03) em execução, se o usuário escolher uma unidade de gravação que não exista, o mesmo será informado que a opção é inválida e terá de tentar novamente.

```
bemtevi.cptec.inpe.br - PuTTY
-bash-3.00$ ./Script_Backup
*****
*
*      Escolha o tipo de fita a ser utilizado:      *
*
*          1) LTO2                                  *
*          2) LTO3                                  *
*          3) Sair                                   *
*
*****

* Opção: 2

*****
*
*      Escolha a unidade a ser utilizada:          *
*
*          7) Unidade 7                             *
*          10) Unidade 10                           *
*          3) Voltar e Corrigir                     *
*          4) Sair                                   *
*
*****

* Opção: 11

Opção inválida. Por favor, tente novamente...

*****
*
*      Escolha a unidade a ser utilizada:          *
*
*          7) Unidade 7                             *
*          10) Unidade 10                           *
*          3) Voltar e Corrigir                     *
*          4) Sair                                   *
*
*****

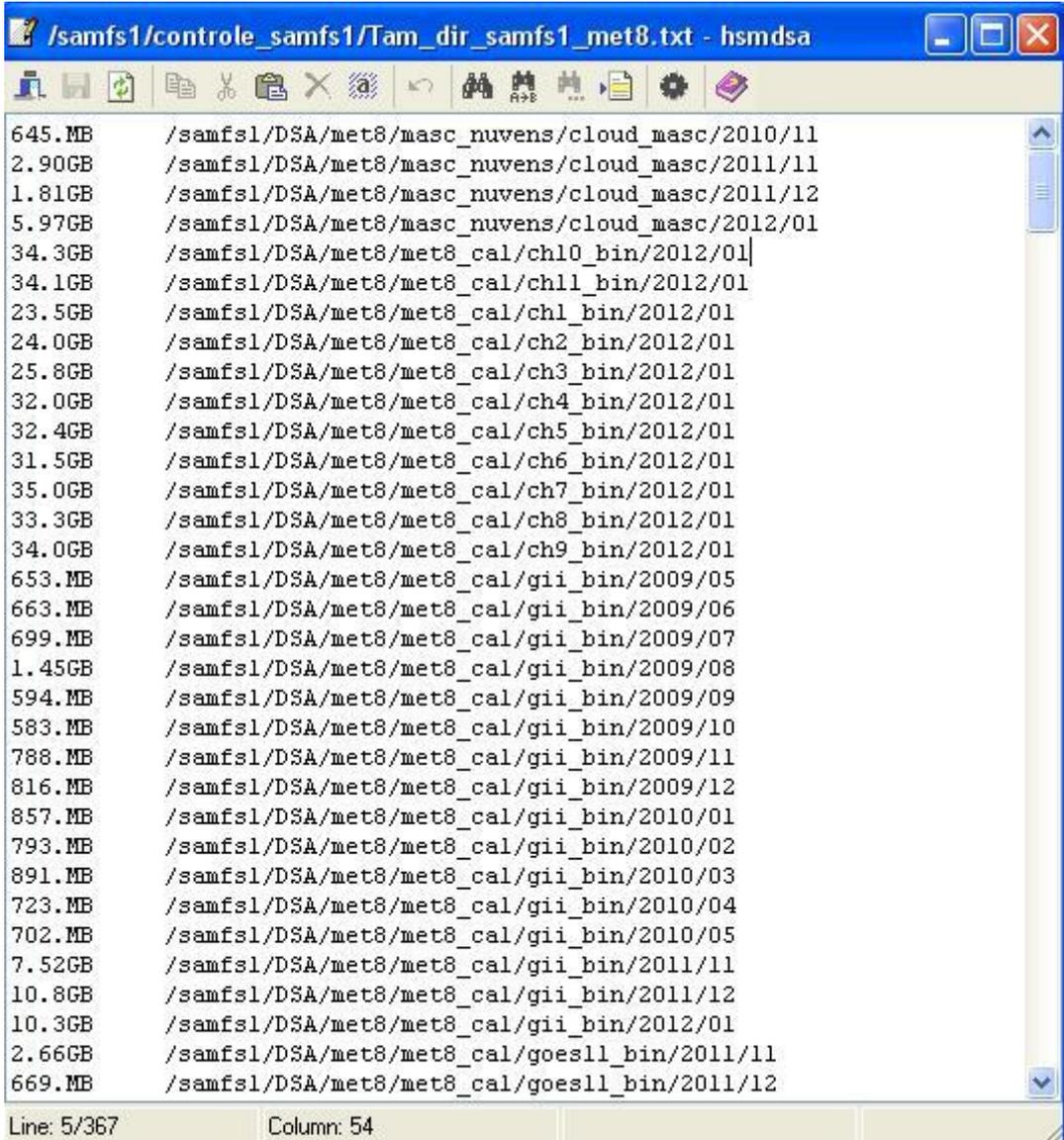
* Opção: 10
```

Figura 3. Parte da execução do script para gravação

Na sequência o script pedirá para o usuário escolher se deseja gravar (caso a fita seja virgem) ou apendar (caso a fita já contenha dados), depois é necessário digitar o nome da lista de saída na qual será escrito todo o processo de gravação. Por ultimo, motivo segurança, é pedido para o usuário conferir se todos os dados estão corretos para continuar.

3.3. Scripts de controle do espaço em disco

Os Scripts foram desenvolvidos com o intuito de facilitar a visualização do espaço ocupado pelos produtos e subprodutos dos satélites NOAA, Meteosat 8 e GOES12, eles são executados semanalmente exibindo em forma de lista o endereço e o tamanho em Gigabytes e Megabytes de cada produto armazenado. Posteriormente, este script servirá como base para a geração de gráficos para auxiliar na visualização de cada subproduto. A seguir (fig 04) a listagem de saída (Meteosat 8):



```
645.MB      /samfsl/DSA/met8/masc_nuvens/cloud_masc/2010/11
2.90GB     /samfsl/DSA/met8/masc_nuvens/cloud_masc/2011/11
1.81GB     /samfsl/DSA/met8/masc_nuvens/cloud_masc/2011/12
5.97GB     /samfsl/DSA/met8/masc_nuvens/cloud_masc/2012/01
34.3GB     /samfsl/DSA/met8/met8_cal/ch10_bin/2012/01
34.1GB     /samfsl/DSA/met8/met8_cal/ch11_bin/2012/01
23.5GB     /samfsl/DSA/met8/met8_cal/ch1_bin/2012/01
24.0GB     /samfsl/DSA/met8/met8_cal/ch2_bin/2012/01
25.8GB     /samfsl/DSA/met8/met8_cal/ch3_bin/2012/01
32.0GB     /samfsl/DSA/met8/met8_cal/ch4_bin/2012/01
32.4GB     /samfsl/DSA/met8/met8_cal/ch5_bin/2012/01
31.5GB     /samfsl/DSA/met8/met8_cal/ch6_bin/2012/01
35.0GB     /samfsl/DSA/met8/met8_cal/ch7_bin/2012/01
33.3GB     /samfsl/DSA/met8/met8_cal/ch8_bin/2012/01
34.0GB     /samfsl/DSA/met8/met8_cal/ch9_bin/2012/01
653.MB     /samfsl/DSA/met8/met8_cal/gii_bin/2009/05
663.MB     /samfsl/DSA/met8/met8_cal/gii_bin/2009/06
699.MB     /samfsl/DSA/met8/met8_cal/gii_bin/2009/07
1.45GB     /samfsl/DSA/met8/met8_cal/gii_bin/2009/08
594.MB     /samfsl/DSA/met8/met8_cal/gii_bin/2009/09
583.MB     /samfsl/DSA/met8/met8_cal/gii_bin/2009/10
788.MB     /samfsl/DSA/met8/met8_cal/gii_bin/2009/11
816.MB     /samfsl/DSA/met8/met8_cal/gii_bin/2009/12
857.MB     /samfsl/DSA/met8/met8_cal/gii_bin/2010/01
793.MB     /samfsl/DSA/met8/met8_cal/gii_bin/2010/02
891.MB     /samfsl/DSA/met8/met8_cal/gii_bin/2010/03
723.MB     /samfsl/DSA/met8/met8_cal/gii_bin/2010/04
702.MB     /samfsl/DSA/met8/met8_cal/gii_bin/2010/05
7.52GB     /samfsl/DSA/met8/met8_cal/gii_bin/2011/11
10.8GB     /samfsl/DSA/met8/met8_cal/gii_bin/2011/12
10.3GB     /samfsl/DSA/met8/met8_cal/gii_bin/2012/01
2.66GB     /samfsl/DSA/met8/met8_cal/goes11_bin/2011/11
669.MB     /samfsl/DSA/met8/met8_cal/goes11_bin/2011/12
```

Line: 5/367 Column: 54

Figura 4. Listagem de saída do Script de verificação do espaço ocupado pelo produto Meteosat 8

4. ATIVIDADES DESENVOLVIDAS

- Pesquisa e estudo com o objetivo de familiarizar com o ambiente Linux/UNIX e Shell script;
- Leitura da apostila "Introdução ao LINUX e Programação em Script - Shell", da Universidade Federal Fluminense;
- Leitura da revista SQL Magazine, focada em banco de dados, análise, projeto e modelagem.
- Fazendo backup manual, gerando listagens e organizando-as para serem gravadas nas unidades LTO 3 e LTO 2, através de linhas de comandos disparando scripts para gravação utilizando acesso remoto;
- Atualizações e organização de planilhas de backup;
- Gerar e aplicar identificação nas fitas LTO;
- Desenvolvimento e testes de scripts para facilitar a visualização do espaço ocupado em disco dos produtos Meteosat 8, Goes12 e Noaa;
- Organização da fitoteca para inserção correta no banco de dados para acesso via Web;
- Atualização das fitas com a nova nomenclatura;
- Início da aplicação das novas etiquetas nas fitas LTO e DLT.

5. CRONOGRAMA FUTURO

Seguindo o mesmo princípio de buscar conhecer mais o Sistema Gerenciador de Banco de Dados(SGBD) PostgreSQL e JSP, para a conclusão do projeto proposto, as etapas que irão se seguir têm como objetivo estarem mais ligadas a automatização completa do processo de backup, utilizando Java e Shell script.

Está incluso no plano inicial também a necessidade de auxiliar no desenvolvimento em Java EE, até mesmo por ser uma linguagem com muitas alternativas e vários recursos, reafirmando que o objetivo agora é aprofundar o conhecimento do mesmo, buscando alternativas mais sofisticadas para evoluir cada vez mais em relação ao plano traçado, não deixando para trás a necessidade do backup manual durante o estudo e pesquisa.

Segue abaixo as etapas futuras:

- Modelagem completa do sistema BDI x Queimadas;
- Estudo das linguagens PostgreSQL, Java EE, Hibernate;
- Auxiliar no desenvolvimento das ferramentas para backup;
- Desenvolver interface gráfica para facilitar a visualização do uso das unidades de armazenamento SAMFS1 e SAMFS2 banco de Dados de Imagens (BDI) com base nos scripts criados.

6. REFERÊNCIAS

[1] UNIVERSIDADE FEDERAL FLUMINENSE. Telecomunicações.

[2] www.caelum.com.br. Último acesso: 17/07/2012 .

Matheus da Cruz Oliveira dos Santos

Bolsista PIBIT

Daniel Alejandro Vila

Orientador