

1010101
1010101
1010101
1010101
1010101
1010101
1010101
1010101
1010101
1010101

Alta Disponibilidade e Cluster para OpenACS

Eduardo Santos

eduardo.edusantos@gmail.com
eduardo.santos@planejamento.gov.br

www.softwarepublico.gov.br
eduardosan.wordpress.com



1010101
1010101
1010101
1010101
1010101
1010101
1010101
1010101
1010101
1010101

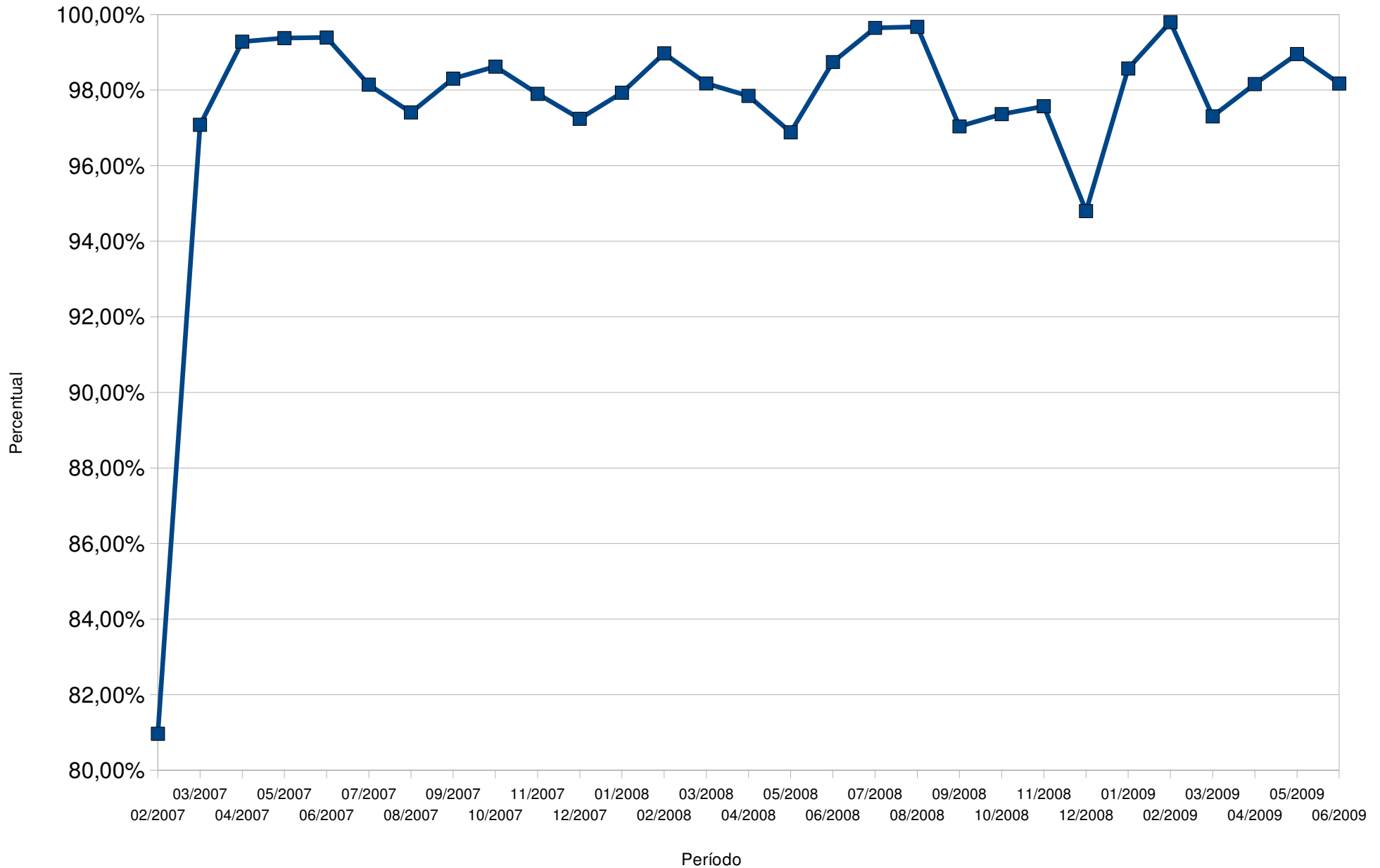
Análise do Problema

Mitos de Internet

- Em sistemas para a Internet, mais de 90% das operações são SELECT
 - Alto índice de escrita em disco;
 - Muitas transações e alta concorrência.
- Usuário mais pró-ativo: **236 tickets** cadastrados no sistema de rastreamento de erros
- Dados de acesso são realmente relevantes?

1010101
1010101
1010101
1010101
1010101
1010101
1010101
1010101
1010101
1010101

Contribuições (s/ Coordenação)

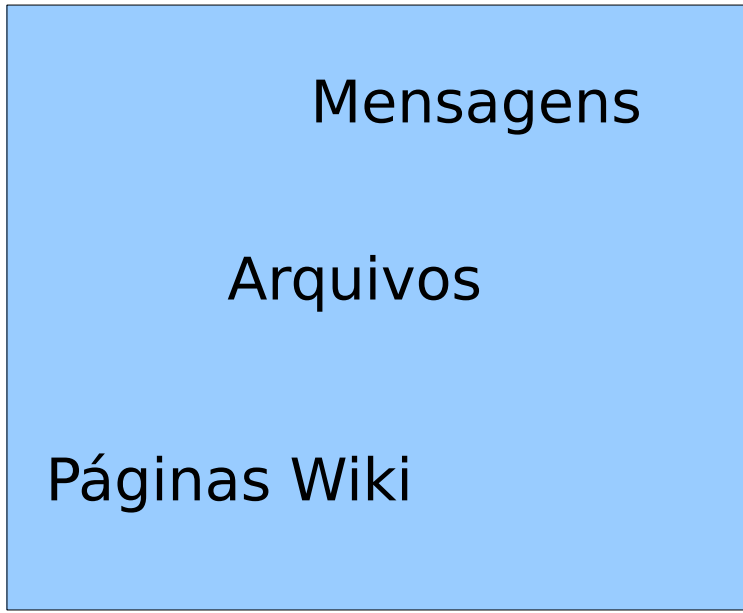




Crescimento da Base (Objetos)

Dia	Percentual	Dia	Percentual
04/05/2009	0,0654%	20/05/2009	0,0667%
05/05/2009	0,0766%	21/05/2009	0,1035%
06/05/2009	0,0552%	22/05/2009	0,0510%
07/05/2009	0,0398%	23/05/2009	0,0209%
08/05/2009	0,0359%	24/05/2009	0,0203%
09/05/2009	0,0241%	25/05/2009	0,0671%
10/05/2009	0,0157%	26/05/2009	0,0615%
11/05/2009	0,0563%	27/05/2009	0,0562%
12/05/2009	0,0643%	28/05/2009	0,0747%
13/05/2009	0,0656%	29/05/2009	0,0365%
14/05/2009	0,0600%	30/05/2009	0,0076%
15/05/2009	0,0550%	31/05/2009	0,0208%
16/05/2009	0,0266%	01/06/2009	0,0501%
17/05/2009	0,0221%	02/06/2009	0,0417%
18/05/2009	0,0623%	03/06/2009	0,0506%
19/05/2009	0,0766%	04/06/2009	0,0199%

Variação em um mês:
1,56%



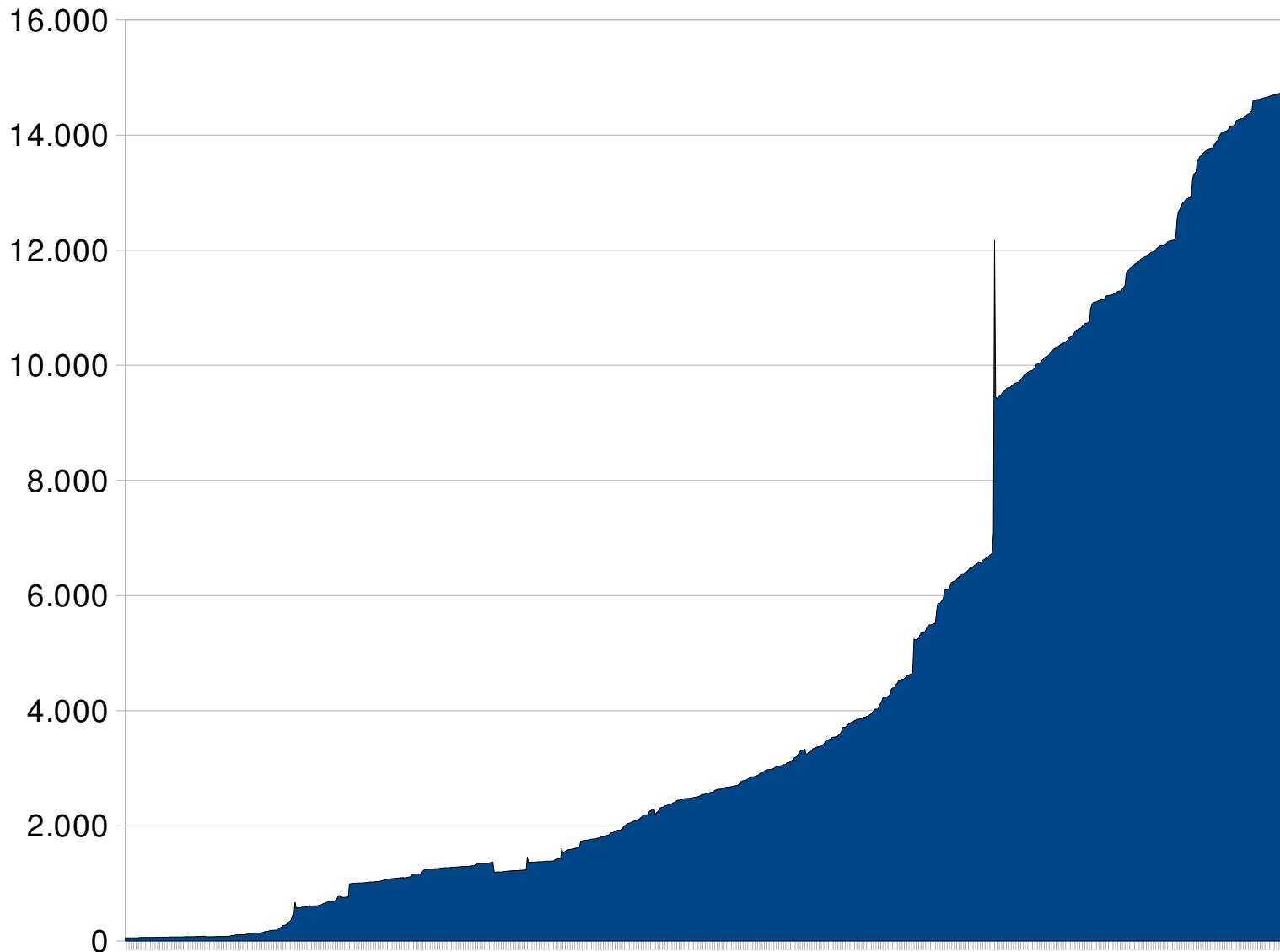
0,97
Novos Objetos/minuto
(Média)

Problemas

- Indisponibilidade causa muitos problemas
 - Insatisfação dos administradores;
 - Reclamação dos usuários de todo o país.
- Em alguns momentos sofremos de lentidão
 - Navegabilidade comprometida;
 - Pior experiência do usuário.
- Perda de **credibilidade**

1010101
1010101
1010101
1010101
1010101
1010101
1010101
1010101
1010101

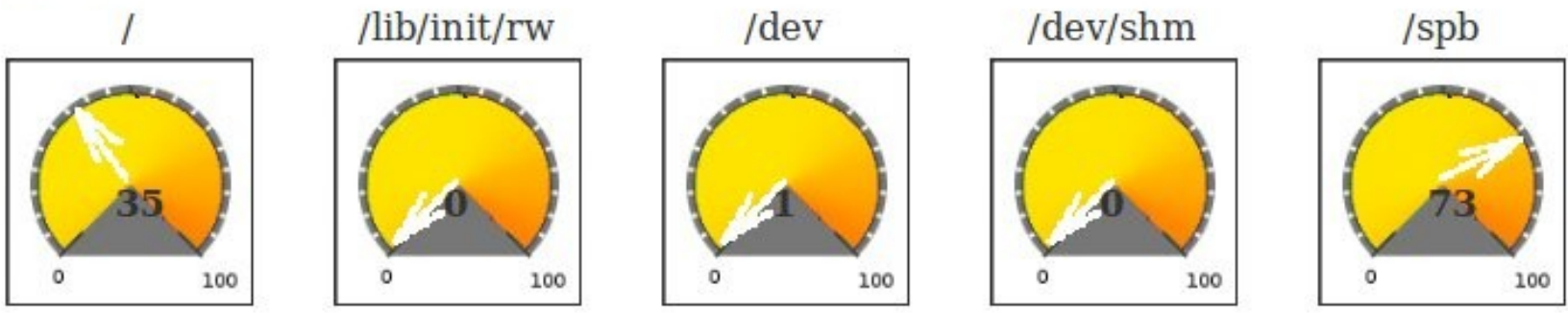
Crescimento da Base (MB)



1010101
1010101
1010101
1010101
1010101
1010101
1010101
1010101
1010101
1010101

Sistema de Arquivos

Espaço em disco :29/10/09 15:01

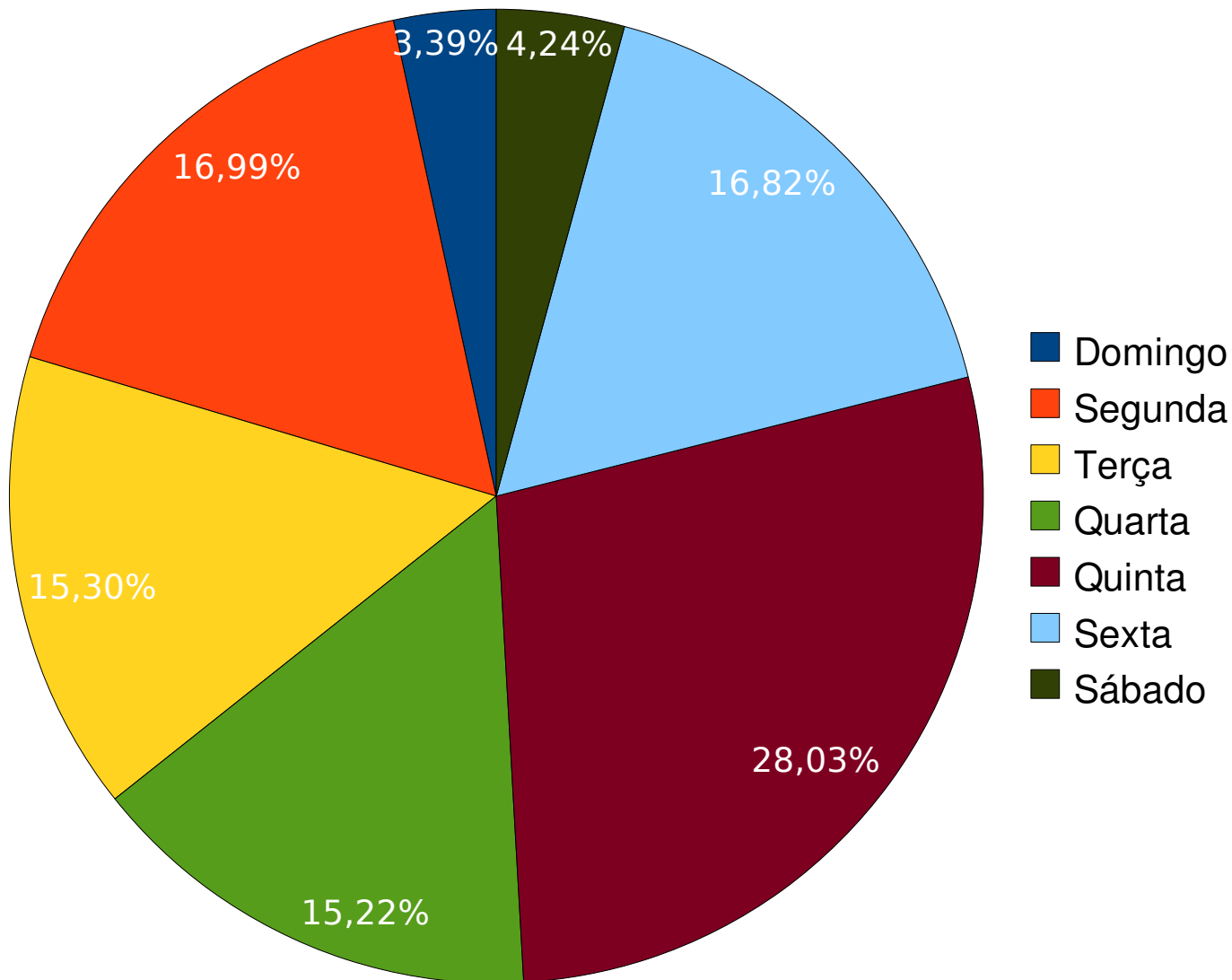


Ponto de Montagem	Sistema de arquivos	Tamanho	Usado	Disponível	% usado
/	/dev/sda2	7.4G	2.4G	4.6G	35%
/lib/init/rw	tmpfs	3.9G	0	3.9G	0%
/dev	udev	10M	52K	10M	1%
/dev/shm	tmpfs	3.9G	0	3.9G	0%
/spb	/dev/sda3	127G	87G	33G	73%





Contribuições / Dia Semana



Dia em que há mais contribuições: **Quinta-feira (28%)**

Resumo

- Usuários contribuem sempre, até mesmo na madrugada do final de semana
- Quanto maior a resposta do Portal, maior o nível de contribuição
- Serviços de Internet causam problemas quando estão fora do ar
- Grande questão a ser analisada:
disponibilidade

1010101
1010101
1010101
1010101
1010101
1010101
1010101
1010101
1010101
1010101

Alta

Disponibilidade e

Arquitetura de

Cluster

1010101
1010101
1010101
1010101
1010101
1010101
1010101
1010101
1010101
1010101

Alta Disponibilidade

“Um sistema de alta disponibilidade é um sistema informático resistente a falhas de software e energia, cujo objectivo é manter os serviços disponibilizados o máximo de tempo possível.”

Fonte:

http://pt.wikipedia.org/wiki/Sistema_de_alta_disponibilidade



1010101
1010101
1010101
1010101
1010101
1010101
1010101
1010101
1010101
1010101

Alta Disponibilidade

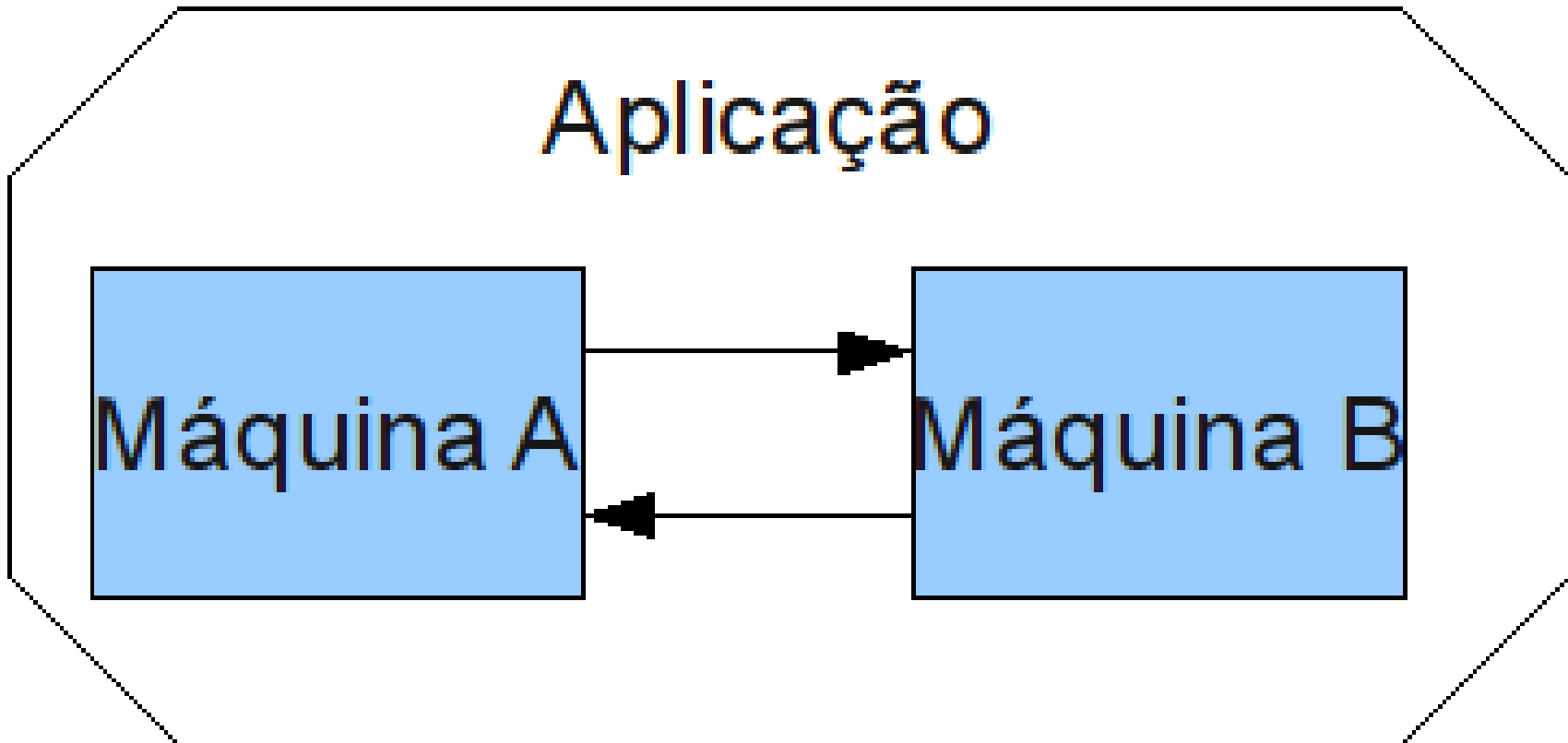
- O que é maior tempo possível?
- Qual o seu limiar de falhas, ou seja, qual serviço pode se manter sozinho ou que depende de outro?
- Qual a sua real necessidade?
- Qual o custo que a solução proposta terá?

Cluster

- Definição: divisão de processamento entre máquinas interligadas
- Normalmente relacionado a distribuição de carga entre mais de uma máquina
- Cluster serve a alta disponibilidade quando os dados/processamento são compartilhados entre mais de uma máquina

1010101
1010101
1010101
1010101
1010101
1010101
1010101
1010101
1010101
1010101

Modelo Simplório



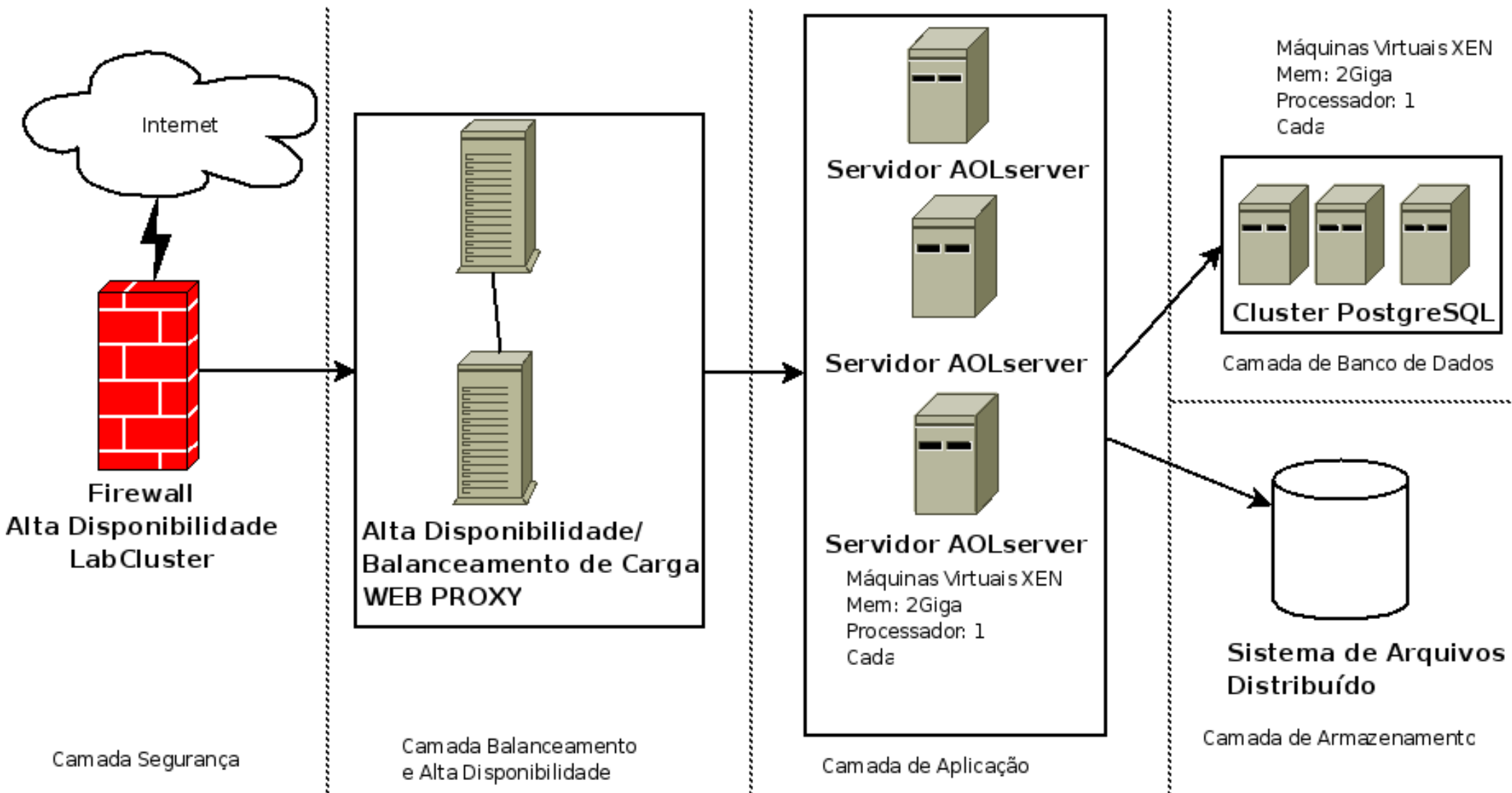
1010101
1010101
1010101
1010101
1010101
1010101
1010101
1010101
1010101
1010101

Modelo Simplório

- Usuário acessa a Aplicação
- A ferramenta de Cluster trata as duas máquinas como se fossem apenas uma
- Os dados são replicados entre ambas as máquinas de forma **transparente**

1010101
1010101
1010101
1010101
1010101
1010101
1010101
1010101
1010101
1010101

Proposta Inicial



Camada Segurança

Camada Balanceamento e Alta Disponibilidade

Camada de Aplicação

Camada de Armazenamento

1010101
1010101
1010101
1010101
1010101
1010101
1010101
1010101
1010101
1010101

Proposta Inicial

- Separação em camadas
- Tratar todos os serviços com modelos genéricos
- Problema da **confiabilidade** e **persistência**

Aprendizado

- Precisamos planejar não só o Portal, mas uma **Estrutura de Produção**
- Máquina virtual realmente não escala
- O SPB precisa de um servidor dedicado
- O Proxy Apache não é uma boa solução para o nosso problema (mod_proxy + AOLServer)
- Trabalhar preemptivamente

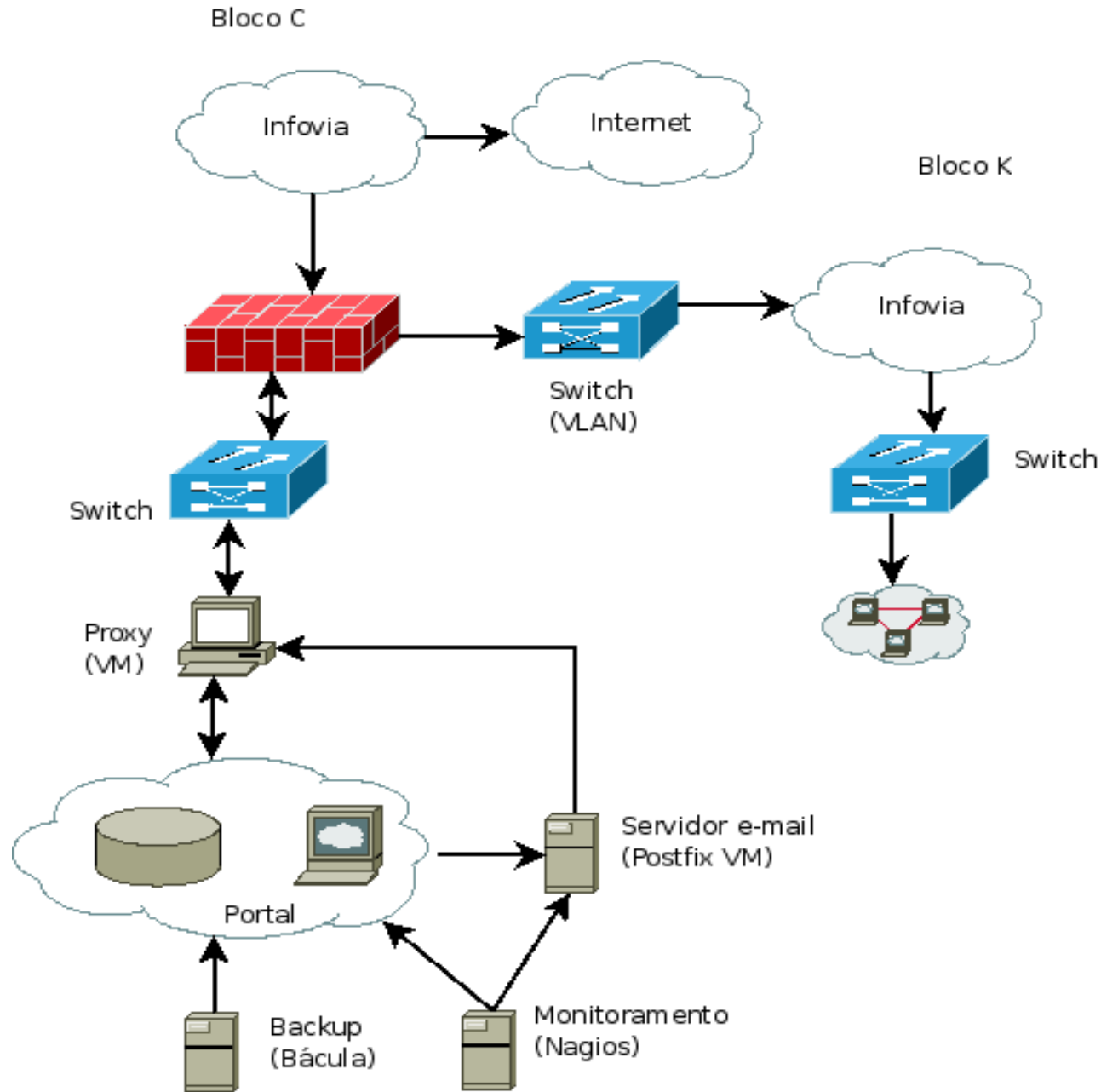
1010101
1010101
1010101
1010101
1010101
1010101
1010101
1010101
1010101
1010101

Implementação Evoluída



1010101
1010101
1010101
1010101
1010101
1010101
1010101
1010101
1010101
1010101

Estrutura Atual

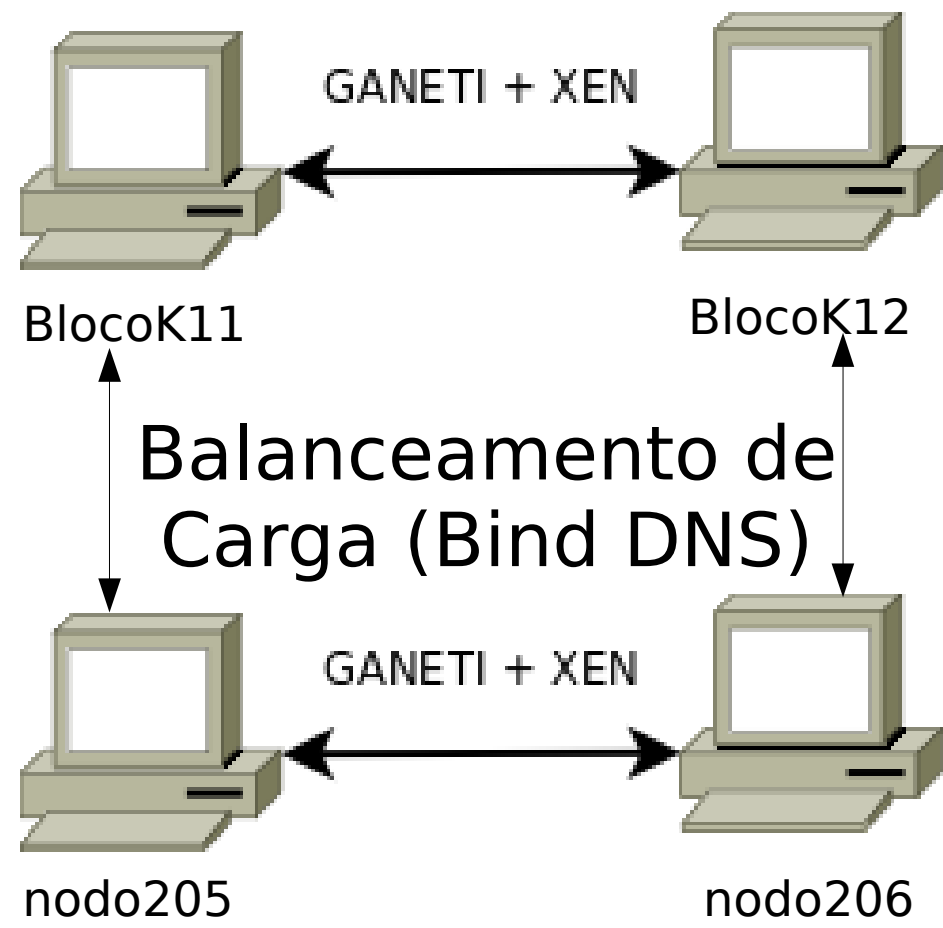


Proxy

- Remover o Apache e colocar o Nginx
- Implementar estrutura de espelhamento do Nginx, ainda com VM's
- Implementação completa do Ganeti + Xen



Proxy Cluster



No caso, os servidores estão replicados sempre e ambos funcionam ao mesmo tempo

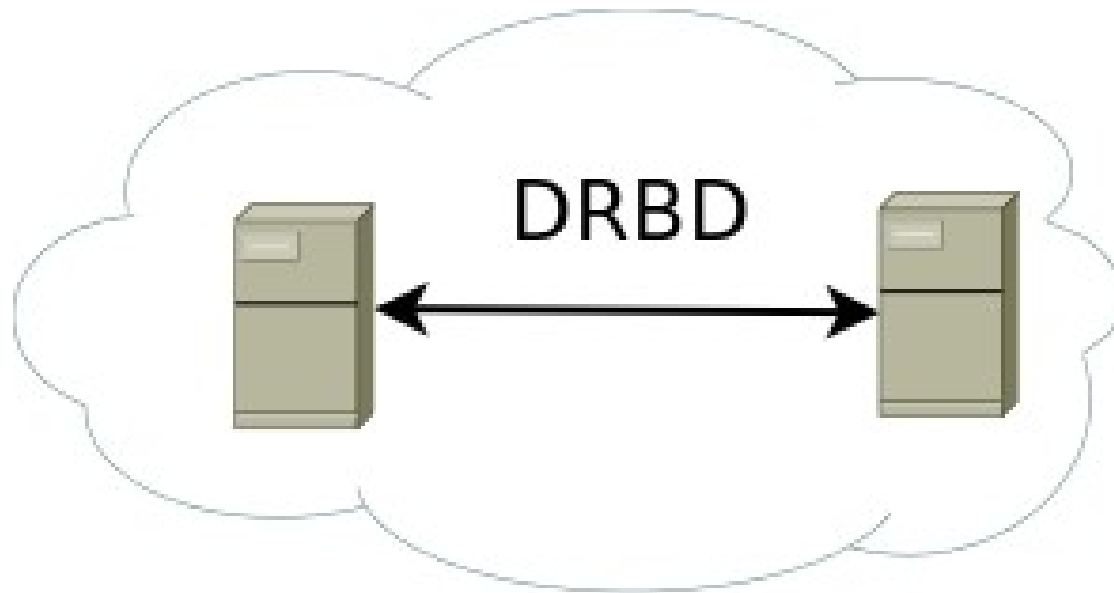


Portal

- Separar o banco de dados da aplicação
- Criar estrutura de Cluster (com VM's?)
- Servidor de arquivos separado
- Servidor com cache de imagens e download de arquivos (pool de threads)
- Alta disponibilidade em todos os serviços
 - Arquivos: DRBD
 - Servidores: Ganeti + Xen + DRBD

1010101
1010101
1010101
1010101
1010101
1010101
1010101
1010101
1010101
1010101

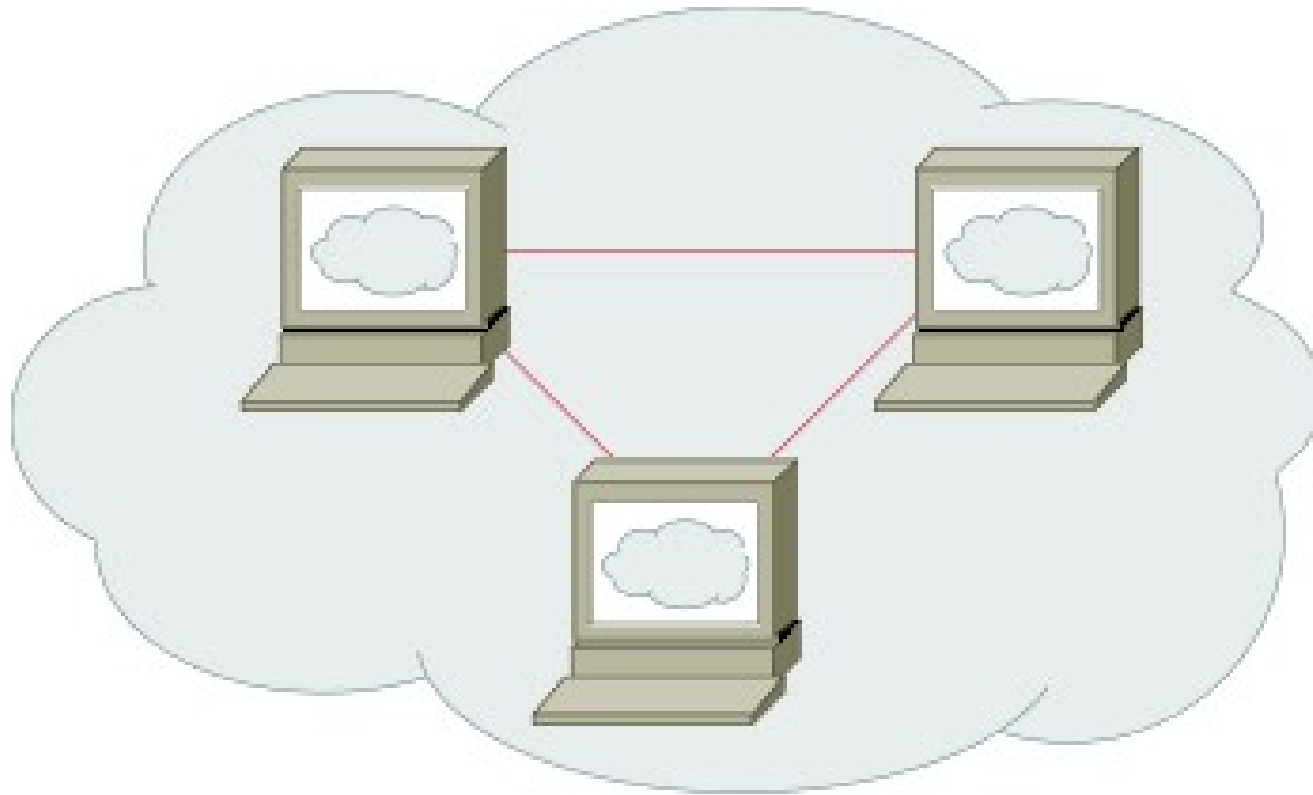
Sistema de Arquivos



Conteúdo da pasta content-repository

1010101
1010101
1010101
1010101
1010101
1010101
1010101
1010101
1010101
1010101

Servidores



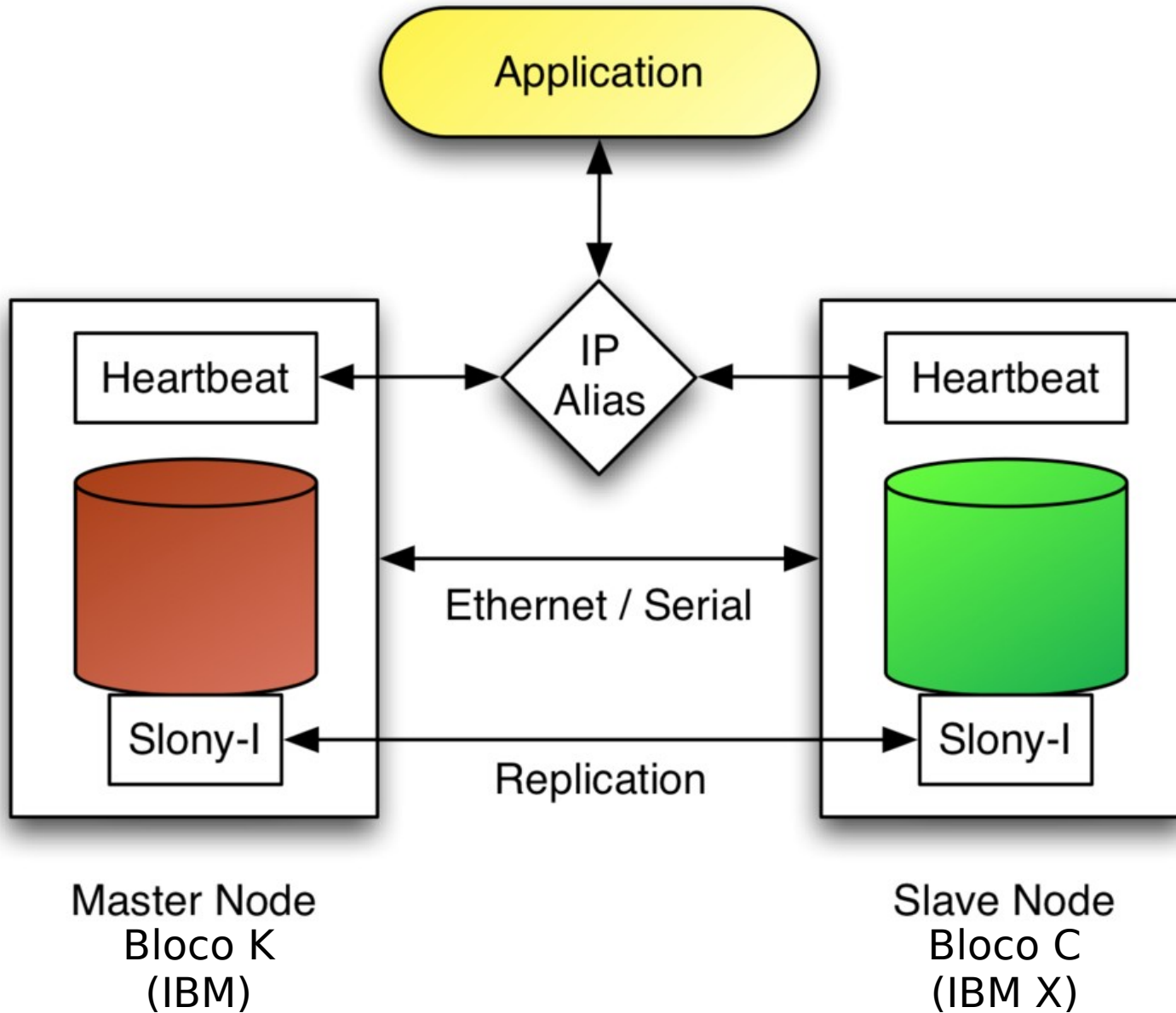
Ganeti + Xen + DRBD

PostgreSQL

- Replicação master-slave assíncrona:
 - Slony I
 - Separar os ambientes fisicamente (bloco C e bloco K)
 - Utilizar o melhor servidor que temos para tratar dos bancos (AMD Opteron 16GB de Ram 06 HD SCSI 146GB)

1010101
1010101
1010101
1010101
1010101
1010101
1010101
1010101
1010101
1010101

Slony-I

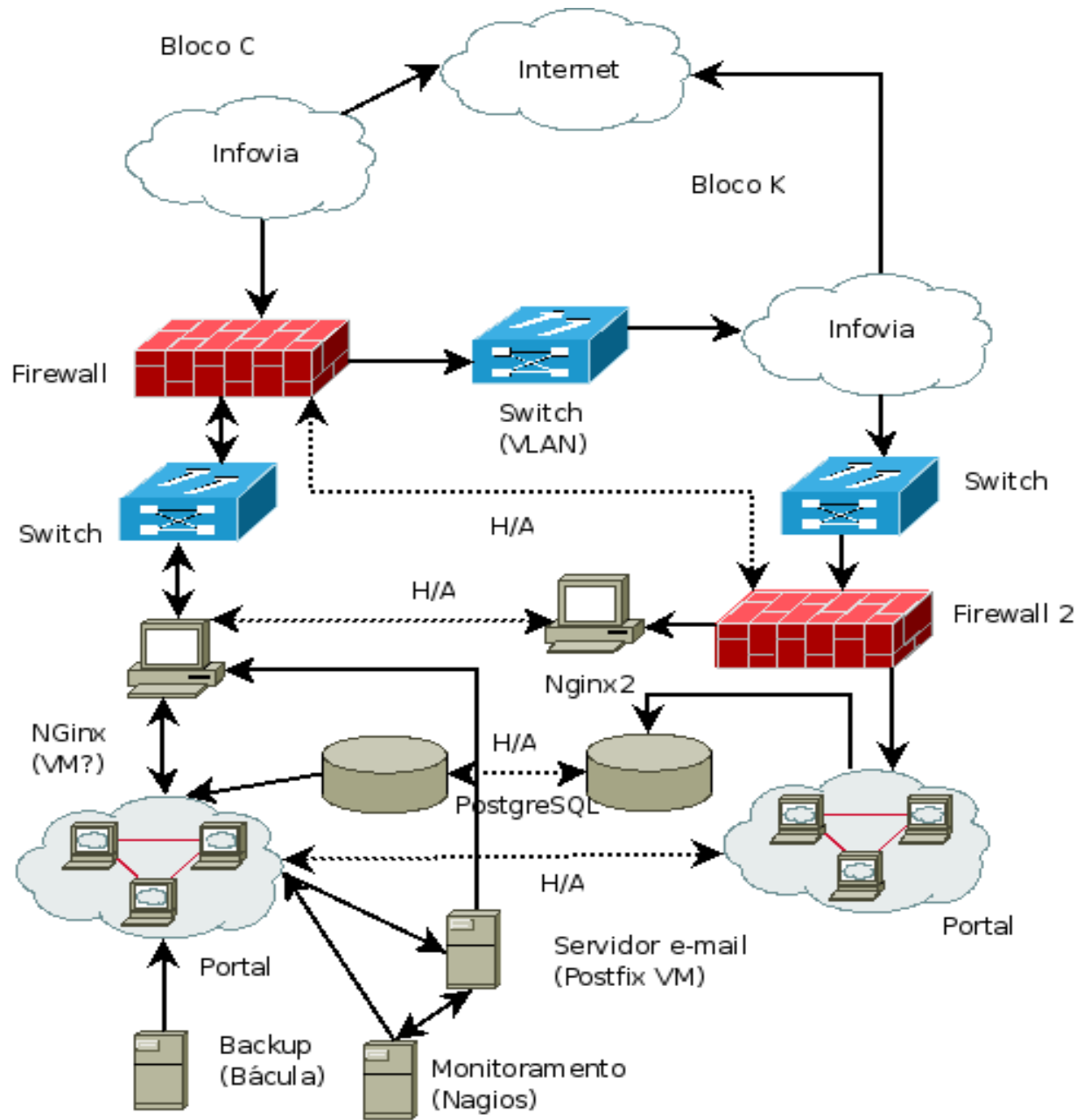


Rede

- Conseguir mais um link de rede no bloco K, para atender os seguintes requisitos:
 - Proteção em caso de falha em qualquer um dos blocos;
 - Balanceamento de carga de rede;
 - Alta disponibilidade do link de rede (Infovia Brasília/RNP)

1010101
1010101
1010101
1010101
1010101
1010101
1010101
1010101
1010101
1010101

Futura Estrutura



1010101
1010101
1010101
1010101
1010101
1010101
1010101
1010101
1010101

Resumo das necessidades

- Alta disponibilidade em todos os pontos do sistema
- Obter máximo desempenho de cada um dos componentes
- Suporte 24x7
- Garantir uma melhor experiência de navegação para os usuários

1010101
1010101
1010101
1010101
1010101
1010101
1010101
1010101
1010101
1010101

Contato

Eduardo Ferreira dos Santos
MPOG - SLTI/DSI

eduardo.santos@planejamento.gov.br
eduardo.edusantos@gmail.com

www.softwarepublico.gov.br
eduardosan.wordpress.com
catir.softwarepublico.gov.br

(61) 2020-1329

