

**CENTRO UNIVERSITÁRIO UNIVATES
CURSO DE ENGENHARIA DA COMPUTAÇÃO**

**FERRAMENTA DE COLETA E ANÁLISE DE
ESTATÍSTICAS PARA SISTEMA
GERENCIADOR DE BANCO DE DADOS
POSTGRESQL**

Fabiano Tomasini

Orientador: Evandro Franzen

Sumário

- Introdução
- Problema
- Objetivos
- Trabalhos relacionados
- Desenvolvimento
- Demonstração do sistema
- Conclusão
- Trabalhos futuros

Introdução

Sobre a Ferramenta

Auxiliar no monitoramento de uma base de dados extraindo informações do catálogo do SGBD.

Importância de monitorar o ambiente (base de dados e SGBD)

- Garantir disponibilidade aos dados armazenados;
- Garantir que informações importantes sejam obtidas de maneira fácil e rápida;
- Evitar ocorrências de impasses.

Problema

- Gerenciamento de ambiente complexo, é necessário ativar muitas ferramentas para ter um monitoramento eficiente.

The screenshot displays a complex monitoring environment with three main windows:

- ZABBIX**: A monitoring dashboard showing various metrics and graphs. The top navigation bar includes Monitoring, Inventory, Reports, Configuration, and Administration. The main content area shows a graph for 'zabbix_sender: sender load average'.
- pgAdmin III**: A database management tool showing the 'Object browser' for a PostgreSQL server. The tree view shows 'Server Groups' (Servers (2)) and 'Databases (9)'. The 'smbd' database is selected, showing its 'Catalogs (2)', 'Event Triggers (0)', and 'Extensions (1)'.
- Terminal Window**: A terminal window showing the output of the 'top' command, displaying system statistics and a list of running processes. The output includes system load, memory usage, and a table of processes with columns: PID, USUÁRIO, PR, NI, VIRT, RES, SHR, S, %CPU, %MEM.

The terminal window output is as follows:

```
top - 11:32:17 up 1 day, 23:59, 13 users, load average: 0,
Tarefas: 295 total, 2 executando, 293 dormindo, 0 parado
%Cpu(s): 1,7 us, 0,5 sy, 0,0 ni, 97,8 id, 0,0 wa, 0,0 st
KiB Mem: 8051284 total, 6012256 used, 2039028 free, 3
KiB Swap: 8261628 total, 440600 used, 7821028 free. 28

  PID USUÁRIO  PR  NI  VIRT  RES  SHR S  %CPU  %MEM
 1169 root      20   0 401640 88420 56324 S   6,6   1,1
 2341 ftomasi+  20   0 1884228 278332 47744 S   3,3   3,5
10320 ftomasi+  20   0 938056 281084 26132 S   3,0   3,5
 2636 ftomasi+  20   0 1372888 363088 80412 S   1,3   4,5
 2743 ftomasi+  20   0 738356 24972 5504 S   1,3   0,3
 23146 ftomasi+  20   0 881968 113156 23816 S   1,0   1,4
10541 ftomasi+  20   0 631360 19668 12520 S   0,7   0,2
   11 root      20   0   0      0      0 S   0,3   0,0
   29 root      20   0   0      0      0 S   0,3   0,0
 2323 ftomasi+  20   0 20232 168 100 S   0,3   0,0
 3033 ftomasi+  20   0 660856 26028 10260 S   0,3   0,3
```

Problema

- Monitoramento não é pró-ativo para itens específicos do SGBD (melhor horário para manutenção, alerta de objetos obsoletos);
- Análise de estatísticas torna-se complexa com os recursos fornecidos pelas ferramentas conhecidas, pois não fornecem informações de maneira gráfica, somente através de comandos sql executados diretamente no SGBD.

Objetivo geral

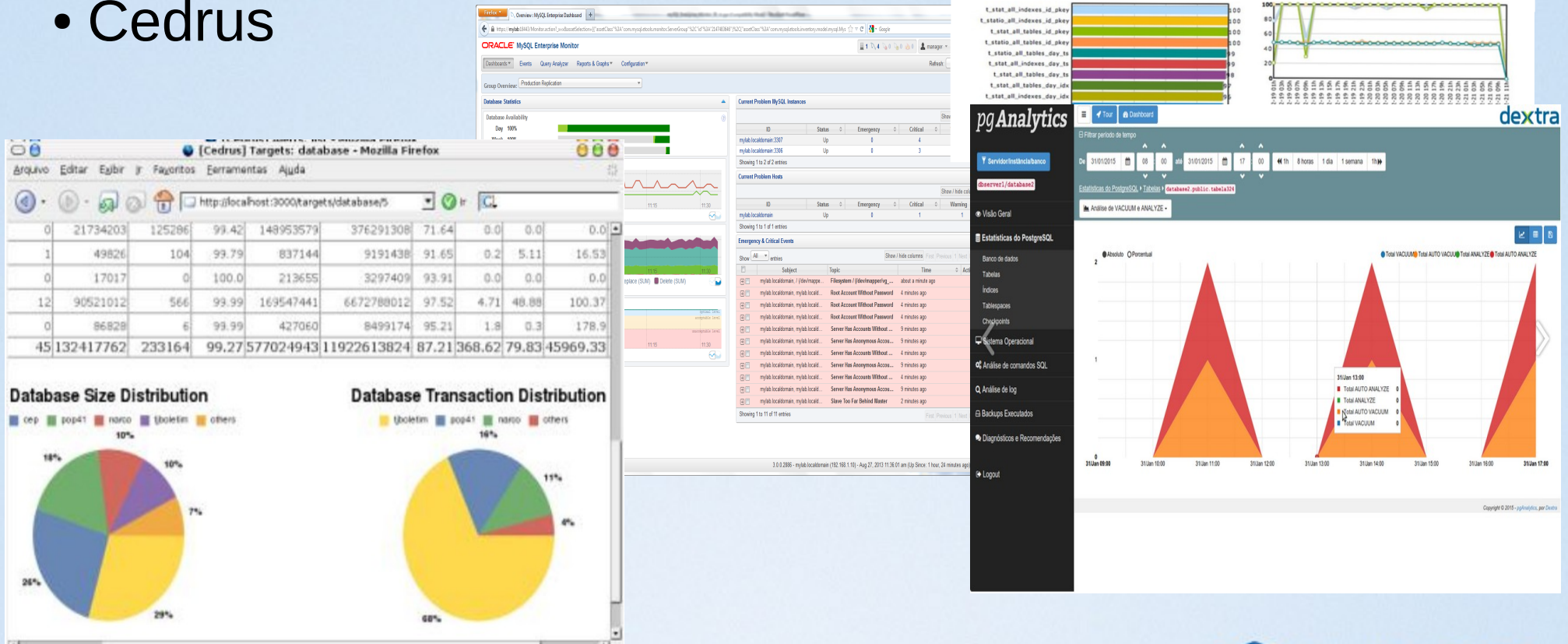
- Tendo como base a realidade apresentada, o presente trabalho, visa proporcionar ao DBA (*Database Administrator*) uma ferramenta para auxiliar na administração, monitoramento e otimização de um banco de dados que utilize SGBD PostgreSQL.

Objetivos específicos

- Compreender os princípios de coleta e utilização de estatísticas de um banco de dados;
- Permitir a coleta de estatísticas através de rotinas automatizadas;
- Contribuir para uma melhor administração de dados através da ferramenta desenvolvida.

Trabalhos relacionados

- MySQL Enterprise Monitor
- Pgwatch
- Pganalytics
- Cedrus

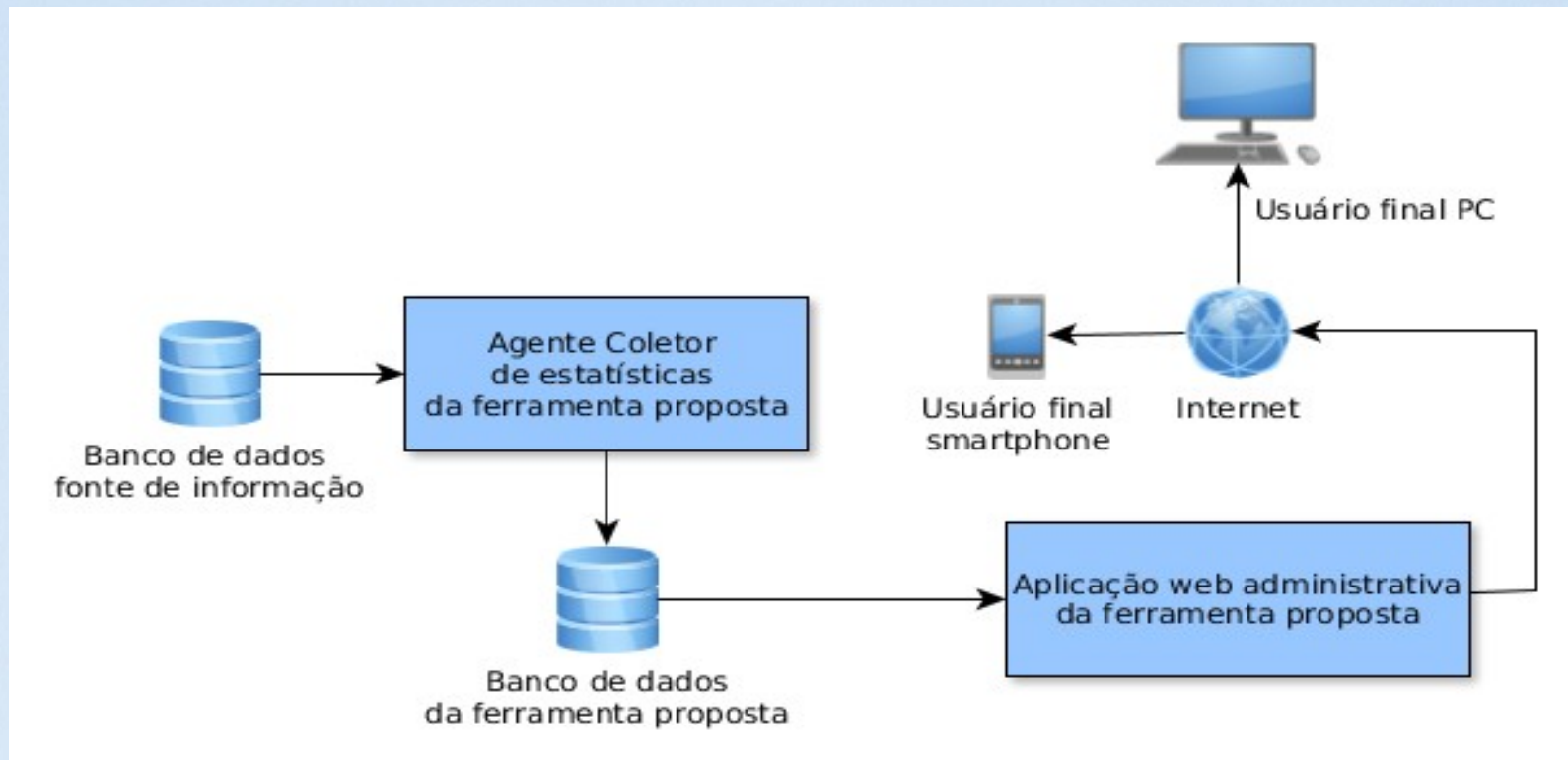


Comparativo entre as ferramentas

Características	MySQL Enterprise Monitor	Pgwatch	Pganalytics	Cedrus
Administração de múltiplos servidores	Sim	Sim	Sim	Sim
Open source	Sim	Sim	Não	Sim
Compatibilidade com versões atuais de SGBDs	Sim	Não	Sim	Não
Alarmísticas	Sim	Não	Sim	Sim
Gráficos pre definidos	Sim	Sim	Sim	Sim
Multiplataforma	Sim	Sim	Sim	Não
Geração de relatórios	Sim	Sim	Sim	Sim

Ferramenta desenvolvida

- A ferramenta contempla um agente coletor que se comunica com a base de dados a ser monitorada e envia os dados para a aplicação desenvolvida que exibe as informações de maneira gráfica e de fácil entendimento.



Ferramenta desenvolvida

Tarefas do agente coletor

- Configuração dos tempos de coleta;
- Coleta de informações do sistema operacional;
- Coleta de estatísticas do banco de dados.

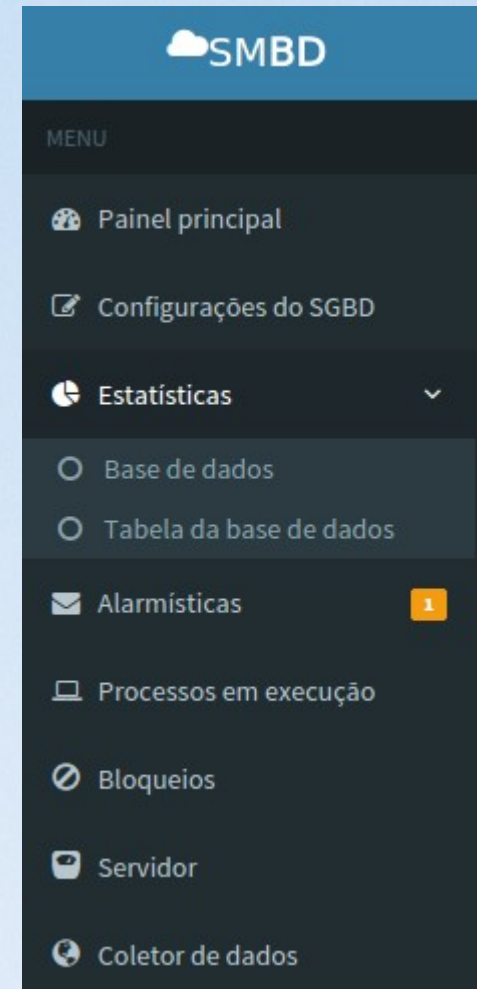
```
5  $config = array(  
6      //base de dados  
7      'host' => 'localhost',  
8      'port' => '5432',  
9      'user' => 'postgres',  
10     'password' => 'postgres',  
11     'dbname' => 'smbd',  
12     //smbd  
13     'url' => 'http://smbd.com.br',  
14     'usuario' => 'ftomasini.rs@gmail.com',  
15     //intervalo de coleta  
16     'tempo_coleta_sghd_versao' => '200 minutes',  
17     'tempo_coleta_base_de_dados' => '200 minutes',  
18     'tempo_coleta_tabela' => '200 minutes',  
19     'tempo_coleta_indice' => '200 minutes',  
20     'tempo_coleta_configuracoes' => '200 minutes',  
21     'tempo_coleta_loadavg' => '1 minute',  
22     'tempo_coleta_memoria' => '1 minute',  
23     'tempo_coleta_processos' => '1 minute',  
24     'tempo_coleta_bloqueios' => '200 minutes',  
25 );
```



Ferramenta desenvolvida

Tarefas da aplicação web desenvolvida

- Exibição de alertas;
- Exibição das configurações do SGBD;
- Exibição dos recursos da máquina em tempo real;
- Utilização dos objetos do banco de dados;
- Acompanhamento de SQLs onerosas.



Ferramenta desenvolvida

Dados coletados do sistema operacional

- Média de carga (*load average*);
- Memória.

Dados coletados do banco de dados

- Estatísticas do banco de dados;
- Estatísticas de tabelas;
- Estatísticas de índices;
- Configurações do banco de dados;
- Processos em execução;
- Bloqueios que estão ocorrendo.

Ferramenta desenvolvida

Tecnologias Envolvidas



PostgreSQL



Bootstrap

HTML

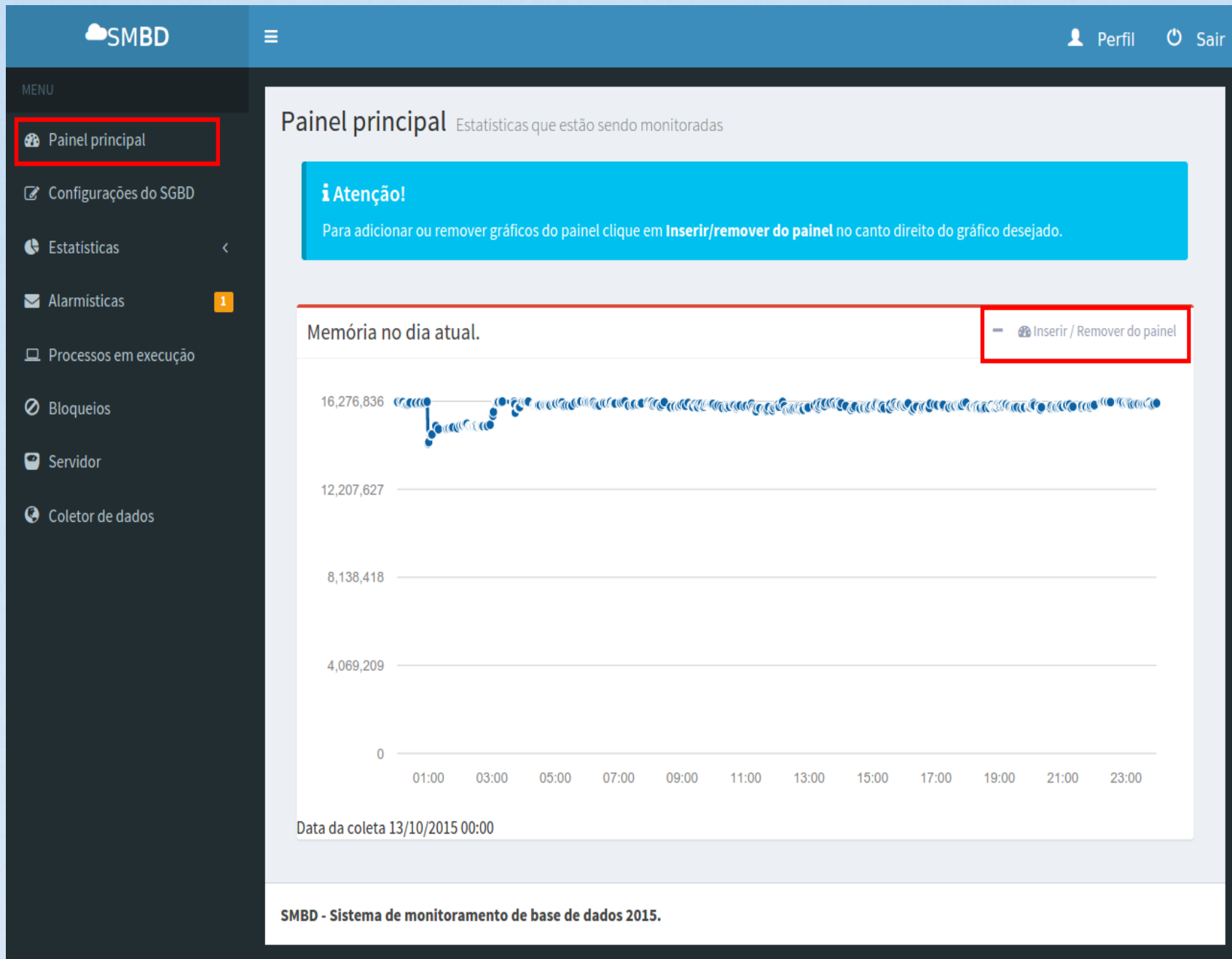


morris.js



JavaScript

Painel principal



Configurações do SGBD

SMBD

≡

Perfil Sair

MENU

Painel principal

Configurações do SGBD

Estatísticas

Alarmísticas 1

Processos em execução

Bloqueios

Servidor

Coletor de dados

Configurações do SGBD

Show 10 entries

Search:

Data coleta	Configuração	Valor	Valor padrão	Descrição
13/10/2015 07:00	archive_command	(disabled)		Sets the shell command that will be called to archive a WAL file.
13/10/2015 07:00	archive_mode	off	off	Allows archiving of WAL files using archive_command.
13/10/2015 07:00	allow_system_table_mods	off	off	Allows modifications of the structure of system tables.
13/10/2015 07:00	application_name			Sets the application name to be reported in statistics and logs.
13/10/2015 07:00	archive_timeout	0	0	Forces a switch to the next xlog file if a new file has not been started within N seconds.
13/10/2015 07:00	array_nulls	on	on	Enable input of NULL elements in arrays.
13/10/2015 07:00	authentication_timeout	60	60	Sets the maximum allowed time to complete client authentication.
13/10/2015 07:00	autovacuum	on	on	Starts the autovacuum subprocess.
13/10/2015 07:00	autovacuum_analyze_scale_factor	0.1	0.1	Number of tuple inserts, updates, or deletes prior to analyze as a fraction of reituples.
13/10/2015 07:00	autovacuum_analyze_threshold	50	50	Minimum number of tuple inserts, updates, or deletes prior to analyze.
Data coleta	Configuração	Valor	Valor padrão	Descrição

Previous

1

2

3

4

5

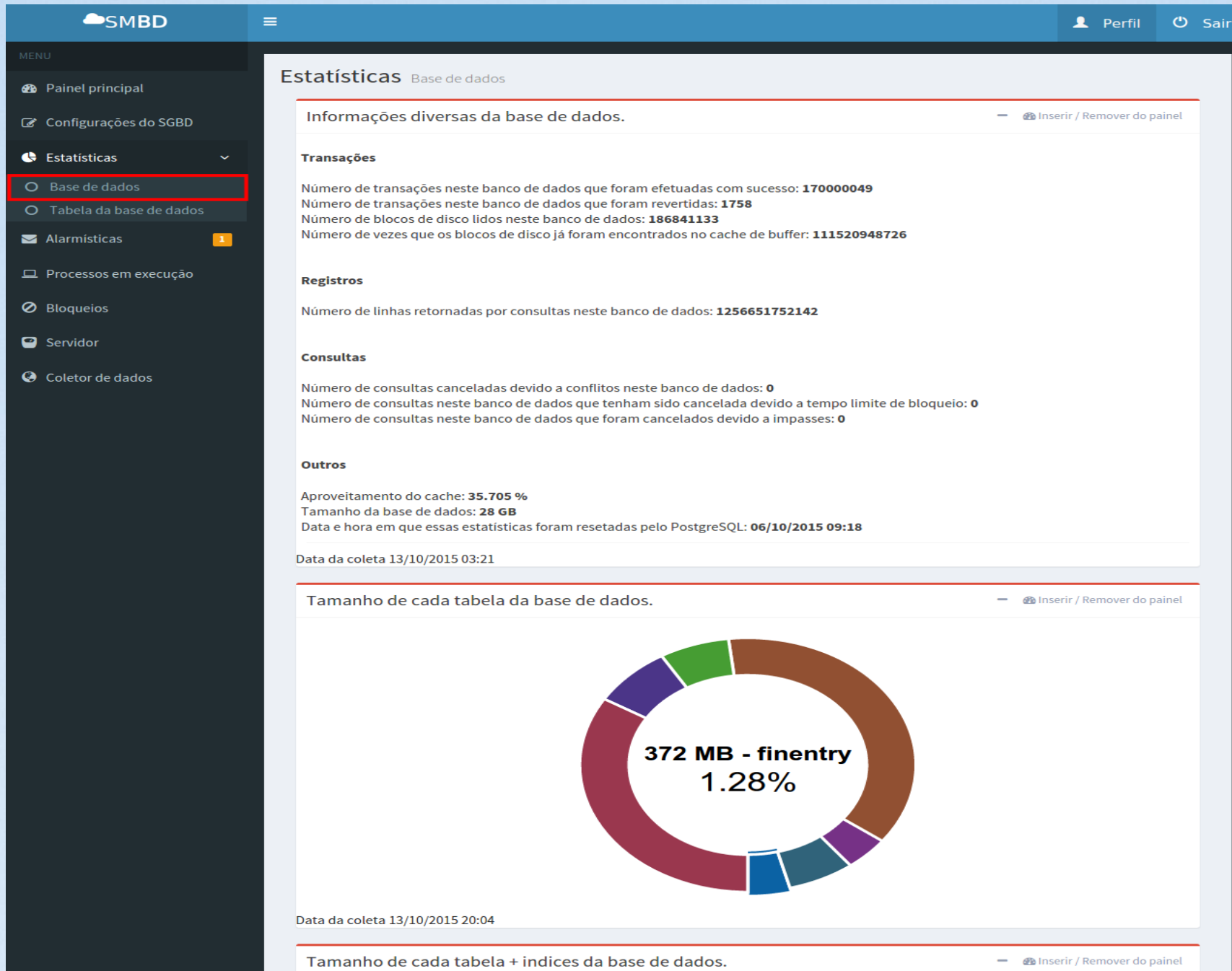
...

24

Next

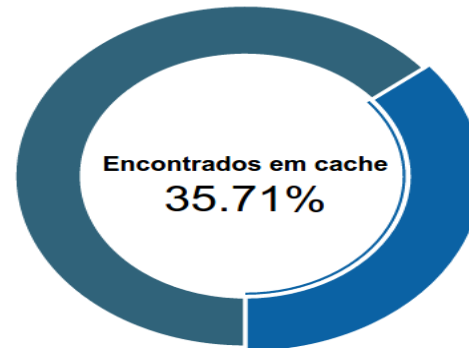
SMBD - Sistema de monitoramento de base de dados 2015.

Estatísticas da base de dados



Estatísticas da base de dados

Percentual de dados encontrados em cache em todas as consultas efetuadas na base de dados. [Inserir / Remover do painel](#)



Data da coleta 13/10/2015 03:21

Tabelas com poucas consultas que utilizaram índice. [Inserir / Remover do painel](#)

Nome da tabela	Percentual de consultas que utilizaram índices
public-acdformationlevel	21.09%
public-gtcpurchaserequest	21.21%
public-gtcsearchpresentationformat	21.49%
public-finocorrenciadoconvenio	22.22%
public-gtcloantype	22.71%
public-findadosbancariosdapessoa	24.93%
public-bascsvimportation	25.49%
public-acdfinalexaminationexaminingboard	25.58%
public-acdfinalexaminationknowledgearea	25.64%
public-ptcrequeststatus	25.98%

Data da coleta 13/10/2015 20:04

Comandos mais lentos do dia. [Inserir / Remover do painel](#)

Comando	Tempo de execução
SELECT count(numeroRegistro) FROM basDocumentoGerado WHERE numeroRegistro = NULL	00:49:35
SELECT GETPARAMETER('BASIC', 'GOOGLE_ANALYTICS')	00:49:35
SELECT order_column, person, ftime, (CASE WHEN char_length(justificativa) > 0 AND \$1::BOOLEAN IS TRUE THEN 'J' ELSE frequency END) AS frequency, upload, ordem_aluno FROM CROSS_ATTENDANCE_REPORT(\$2, \$3, TO_DATE(\$4::TEXT, getParameter('BASIC', 'MASK_DATE'))), TO_DATE(\$5::TEXT, getParameter('BASIC', 'MASK_DATE'))),	00:49:34

Estatísticas de uma tabela da base de dados

SMDB

Perfil

Sair

MENU

Painel principal

Configurações do SGBD

Estatísticas

Base de dados

Tabela da base de dados

Alarmísticas

Processos em execução

Bloqueios

Servidor

Coletor de dados

Estatísticas

Tabela da base de dados

Selecione a tabela

public-acdddegree

Buscar

Informações diversas da tabela public-acdddegree.

Vacuum

Último vacuum executado manualmente:
Último vacuum executado pelo autovacuum:
Último vacuum analyze executado manualmente: **13/10/2015 03:16**
Último vacuum analyze executado pelo autovacuum:
Nº de vezes que o vacuum foi executado manualmente: **0**
Nº de vezes que o vacuum foi executado pelo autovacuum: **0**
Nº de vezes que o vacuum analyze foi executado manualmente: **7**
Nº de vezes que o vacuum analyze foi executado pelo autovacuum: **0**

Registros

Número de linhas inseridas: **0**
Número de linhas atualizadas: **0**
Número de linhas excluídas: **0**
Número de linhas vivas: **0**
Número de linhas mortas: **0**

Varreduras

Número de varreduras sequenciais: **1880067**
Número de varreduras indexadas: **142312944**

Data da coleta 13/10/2015 20:04

Índices não utilizados da tabela public-acdddegree.

Nome	Número de varreduras nesse índice	Tamanho do índice
idx_acdddegree_degreenumber	0	120 kB
idx_acdddegree_centerid	0	120 kB


Data da coleta 13/10/2015 20:13

Índices utilizados da tabela public-acdddegree.

Nome	Número de varreduras nesse índice	Tamanho do índice
idx_acdddegree_learningperiodid	145750	120 kB
acdddegree_pkey	136969589	120 kB
idx_acdddegree_unique	5463526	160 kB

Data da coleta 13/10/2015 20:13

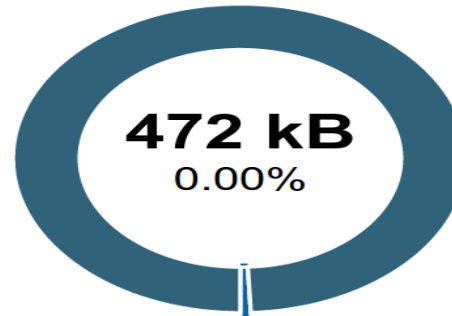
19/29

 UNIVATES

Estatísticas de uma tabela da base de dados

Tamanho da tabela public-acddegree em relação a base de dados.

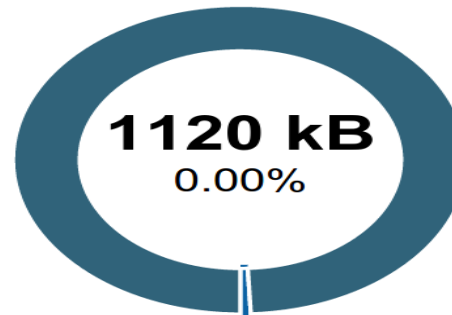
— Inserir / Remover do painel



Data da coleta 13/10/2015 20:04

Tamanho da tabela public-acddegree + índices em relação a base de dados.

— Inserir / Remover do painel



Data da coleta 13/10/2015 20:04

Percentual de dados encontrados em cache em todas as consultas efetuadas na tabela public-acddegree.

— Inserir / Remover do painel



Data da coleta 13/10/2015 20:04

Alarmísticas

The screenshot shows the SMBD (Superior Monitoring and Control) web interface. The top navigation bar includes the SMBD logo, a hamburger menu icon, and user options for 'Perfil' and 'Sair'. The left sidebar menu lists various system components: 'Painel principal', 'Configurações do SGBD', 'Estatísticas', 'Alarmísticas' (highlighted with a red box and a notification badge), 'Processos em execução', 'Bloqueios', 'Servidor', and 'Coletor de dados'. The main content area, titled 'Alarmísticas', displays a message: 'Melhor horário para manutenção' followed by 'De acordo com as informações coletadas no dia (13/10/2015) o melhor horário para efetuar alguma manutenção é 00:34 Load avarege: 0.17'.

Processos em execução

SMBD

≡

Perfil

Sair

MENU

Painel principal

Configurações do SGBD

Estatísticas

Alarmísticas

Processos em execução

Bloqueios

Servidor

Coletor de dados

Processos em execução

Memória

98

Data coleta: 13/10/2015 23:57

Memória

Memória usada: 15 GB

Memória livre: 233 MB

Memória compartilhada: 0 bytes

Memória buffers: 29 MB

Memória em cache: 15 GB

Memória swap usada: 28 MB

Memória swap livre: 7600 MB

Memória swap em cache: 3228 kB

Load average

Último min: 2.6

Últimos 5 min: 2.25

Últimos 15 min: 2.14

Data coleta: 13/10/2015 23:59

Search:

Data coleta	Usuário	Código processo	Memória utilizada	Estado	Consulta	Tempo de execução
13/10/2015 23:57	postgres	12971	283 MB	R - executável	INSERT INTO bi_rematricula SELECT COALESCE(SUM(X.matriculados), 0) AS matriculados, COALESCE(SUM(X.renovacoes), 0) AS renovacoes , COALESCE(SUM(X.reingressos), 0) AS reingressos, COALESCE(SUM(X.transferencia_de), 0) AS transferencia_de, COALESCE(SUM(X.ndas_matriculados), 0) AS ndas_matriculados, COALESCE(SUM(X.portadores_diploma), 0) AS portadores_diploma, COALESCE(SUM(X.nao_renovados), 0) AS nao_renovados, COALESCE(SUM(X.prematriculados), 0) AS pre_matriculados FROM (SELECT codigo_curso AS courseid, nome_curso AS NAME, codigo_turno AS turnid, descricao_turno AS description, (CASE WHEN situacao_no_periodo = 'Pré-matriculado' THEN COUNT(codigo_contrato) FI SF 0 FND) AS prematriculados.	00:02:42

Bloqueios

SMBD

≡

Perfil Sair

MENU

Painel principal

Configurações do SGBD

Estatísticas

Alarmísticas

Processos em execução

Bloqueios

Servidor

Coletor de dados

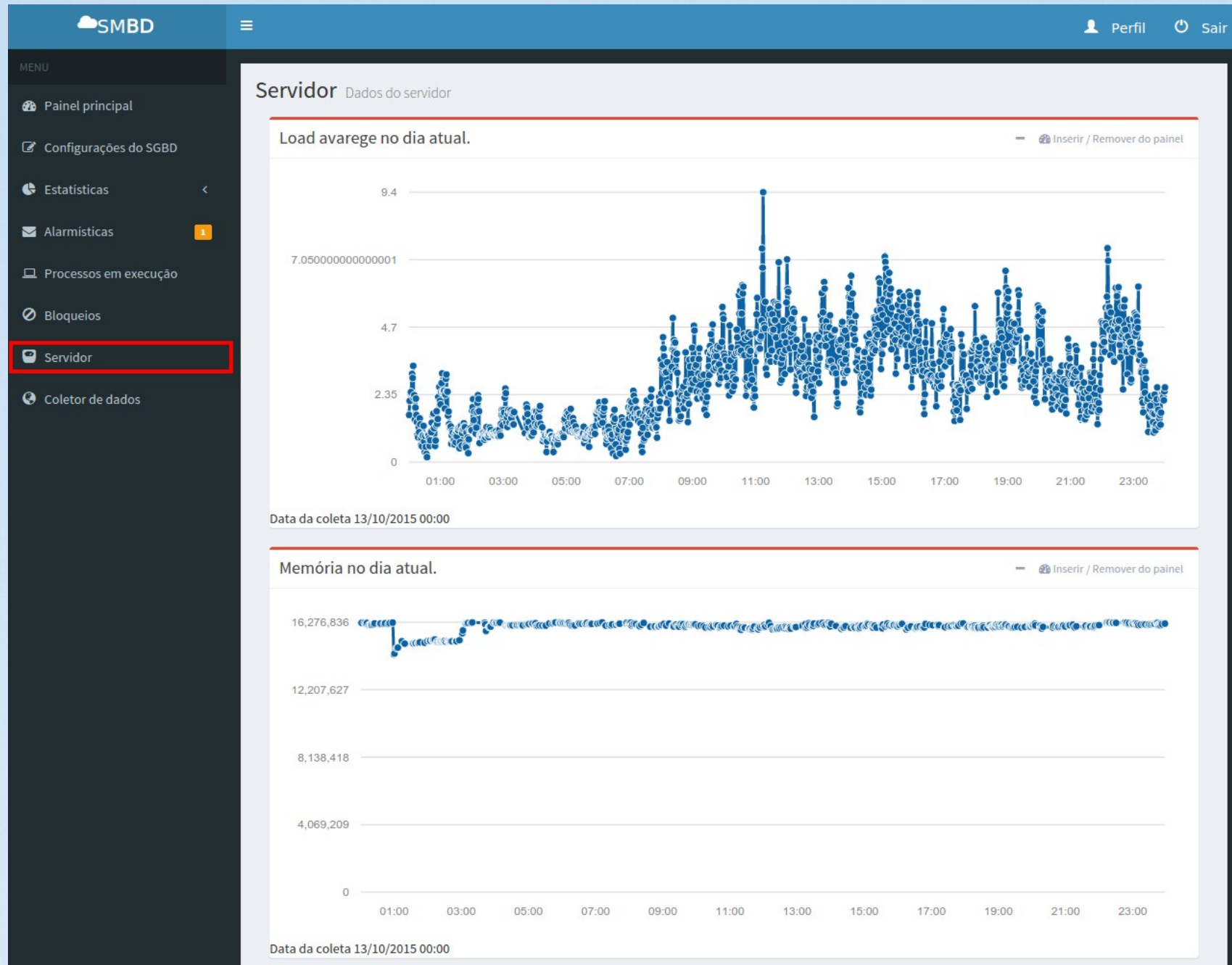
Bloqueios

↺ ↻ ↷

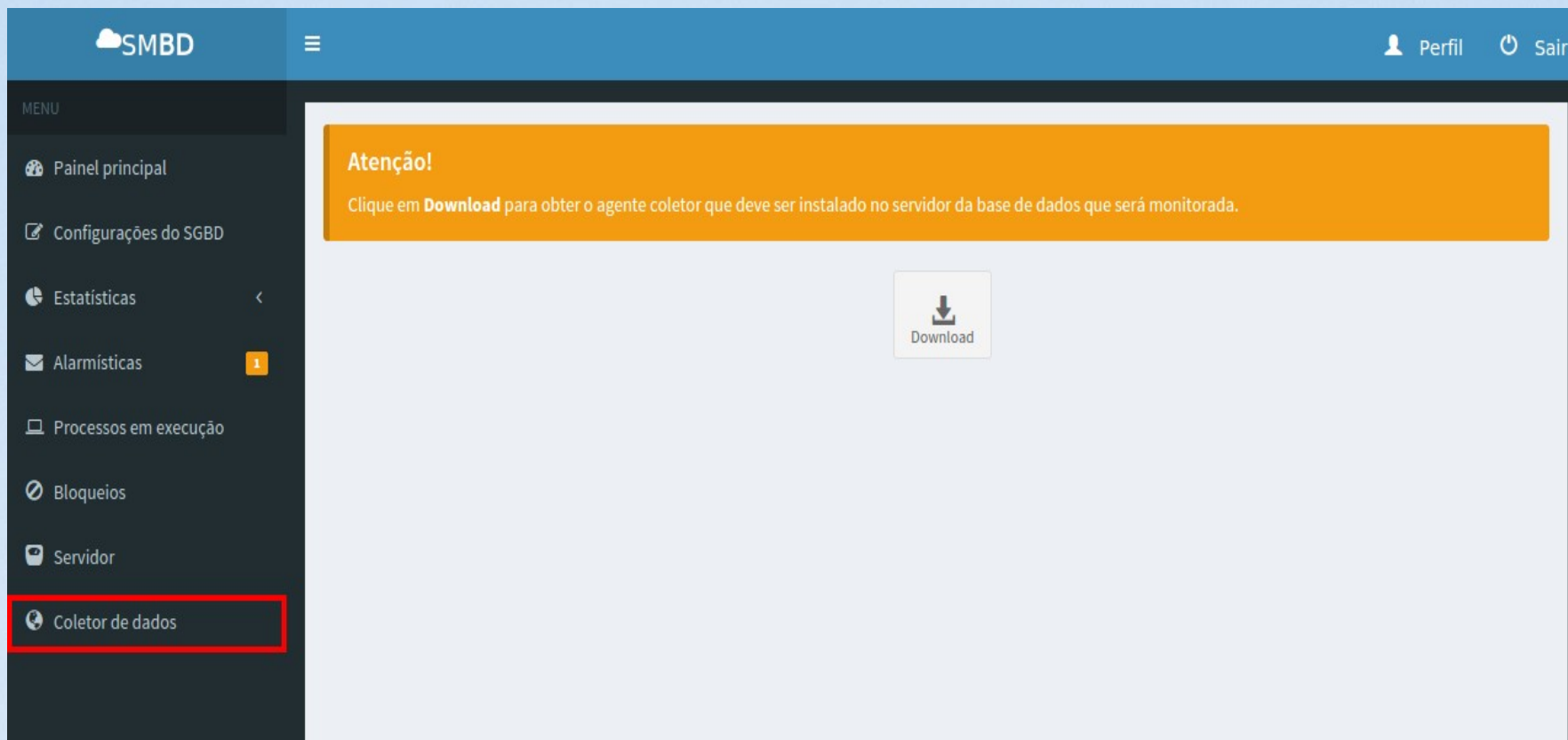
Search:

Data coleta	Usuário	Código processo	Memória utilizada	Estado	Modo	Consulta	Tempo de execução
13/10/2015 23:57	postgres	12971	283 MB	R - executável	ExclusiveLock	INSERT INTO bi_rematicula SELECT COALESCE(SUM(X.matriculados), 0) AS matriculados, COALESCE(SUM(X.renovacoes), 0) AS renovacoes, COALESCE(SUM(X.reingressos), 0) AS reingressos, COALESCE(SUM(X.transferencia_de), 0) AS transferencia_de, COALESCE(SUM(X.ndas_matriculados), 0) AS ndas_matriculados, COALESCE(SUM(X.portadores_diploma), 0) AS portadores_diploma, COALESCE(SUM(X.nao_renovados), 0) AS nao_renovados, COALESCE(SUM(X.prematriculados), 0) AS pre_matriculados FROM (SELECT codigo_curso AS courseid, nome_curso AS NAME, codigo_turno AS turnid, descricao_turno AS description, (CASE WHEN situacao_no_periodo = 'Pré-matriculado' THEN COUNT(codigo_contrato) ELSE 0 END) AS prematriculados, (CASE WHEN situacao_no_periodo = 'Portador de diploma' THEN COUNT(codigo_contrato) ELSE 0 END) AS portadores_diploma, (CASE WHEN situacao_no_periodo = 'Vestibulando' THEN COUNT(codigo_contrato) ELSE 0 END) AS matriculados, (CASE WHEN situacao_no_periodo = 'Renovado' THEN COUNT(codigo_contrato)	00:02:43

Servidor



Coletor de dados



Demonstração

<http://smbd.com.br>

Conclusão

Após a conclusão da ferramenta foi possível validá-la em um ambiente em produção. Dessa forma, foi evidenciado que a mesma cumpriu com o objetivo de disponibilizar um software para auxiliar na administração, monitoramento e otimização de um banco de dados que tenha um grande volume de informação armazenada.

Trabalhos futuros

- Acompanhamento de mais de uma base de dados ao mesmo tempo;
- Implementação de novos alarmes;
- Permitir o monitoramento de outros SGBDs;
- Implementação de novos painéis.

Referências

- BERKUS, J. **Checklist de Performance do PostgreSQL 8.0. PostgreSQL WIKI**, 2005. Disponível em: <https://wiki.postgresql.org/wiki/Checklist_de_Performance_do_PostgreSQL_8.0>. Acesso em: 06 mai. 2015.
- BLAZUS, D.O. PostgreSQL. **PostgreSQL WIKI**, 2003. Disponível em: <https://wiki.postgresql.org/wiki/Introdu%C3%A7%C3%A3o_e_Hist%C3%B3rico>. Acesso em: 03 mai. 2015.
- BLUMM, C; FORNARI, M. R. **Dúvidas frequentes sobre Banco de Dados. Revista SQL Magazine**, 28. ed., 2006.
- BORELLO, F. ; KNEIPP, R. E. **Álgebra Relacional. Revista SQL Magazine**, p. 1 - 1, 2008.
- CEDRUS. **Cedrus: PostgreSQL Manager**. Disponível em: <<http://sourceforge.net/projects/cedrus/>> Acesso em: 19 mai. 2015.
- CYBERTEC. **Pgwatch Cybertec Enterprise PostgreSQL Monitor**. Disponível em: <http://www.cybertec.at/postgresql_produkte/pgwatch-cybertec-enterprise-postgresql-monitor/> Acesso em: 19 mai. 2015.
- DATE, C. J. **Introdução a sistemas de bancos de dados**. 8.ed. Rio de Janeiro: Campus, 2004.
- ELMASRI, R; NAVATHE, S. B. **Sistemas de banco de dados**. 6. ed. Pearson, 2011.
- GIL, A. C. **Como elaborar projetos de pesquisa**. 4. ed. São Paulo: Atlas, 2002.
- MILANI, A. **PostgreSQL: Guia do Programador**. São Paulo: Novatec Editora, 2008.
- ORACLE. MySQL Enterprise Monitor. Disponível em: <<https://www.mysql.com/products/enterprise/monitor.html>> Acesso em: 19 maio. 2015.
- POSTGRESQL. Documentation. Disponível em <<http://www.postgresql.org/docs/manuals/archive/>> Acesso em: 18 abr. 2015.
- SANTOS, A. R. **Metodologia científica: a construção do conhecimento**. 2. ed. Rio de Janeiro: DP&A editora, 1999.
- SILBERSCHATZ, A.; KORTH, H. F.; SUDARSHAN, S. **Sistemas de Bancos de Dados**. 3. ed. Makron Books, 1999.
- SILBERSCHATZ, A.; KORTH, H.F.; SUDARSHAN, S; **Sistemas de Bancos de Dados**. 6. ed. Makron Books, 2012.
- SMANIOTO, C. E. **PostgreSQL. Revista SQL Magazine**, 37. ed., 2007.
- RAMAKRISHNAN, R.; GEHRKE, J. **Sistemas de Gerenciamento de Banco de Dados**. 3. ed. Mc Graw Hill, 2008.
- ROB, P.; CORONEL, C. **Sistemas de Banco de Dados: Projeto, implementação e administração**. 8. ed. CENGAGE Learning, 2011.