

Gerenciamento de recursos humanos

Capítulo 8 – The Data Warehouse Toolkit

Vanessa Maia Aguiar de Magalhães

Orientadora: Profa. Dra. Junia Coutinho Anacleto

29/11/2009

- **Introdução e motivação**
- **Controle de transações com marcação de tempo em uma dimensão**
- **Dimensão com marcação de data/hora com fatos de instantâneos periódicos**
- **Dimensão Auditoria**
- **Dimensão Outrigger**
- **Dilema AND/OR**
- **Pesquisas por substrings**
- **Dados de questionários**
- **Conclusão**

Este capítulo analisa dados de recursos humanos, aplicando várias técnicas de modelagem dimensional a este contexto.

Neste modelo de RH: os dados dos funcionários não são acumulativos, não são números e estão sempre mudando.

Négocio problema

1. Quadro de funcionários com mais de 100 mil empregados
2. Cada funcionário com perfil detalhado, com pelo menos 100 atributos incluindo:
 - a. **Data de admissão**
 - b. **Salário**
 - c. **Data de avaliação**
 - d. **Resultado da avaliação**
 - e. **Periodo de férias**
 - f. **Formação acadêmica**
 - g. **Endereço**
 - h. **Plano de seguro**
 - i. **Entre outros**
3. Funcionários estão sempre sendo contratados, transferidos, promovidos ...

O que a Empresa deseja controlar e analisar?

Controlar e analisar os eventos de transações do departamento, devendo responder as todas as perguntas sobre o perfil de cada funcionário.

Ex: Quem foi transferido ou promovido?, mudanças de endereço...

Como fazer isso?

A partir dos históricos detalhado de transações dos perfis dos funcionários

Esboço do esquema inicial

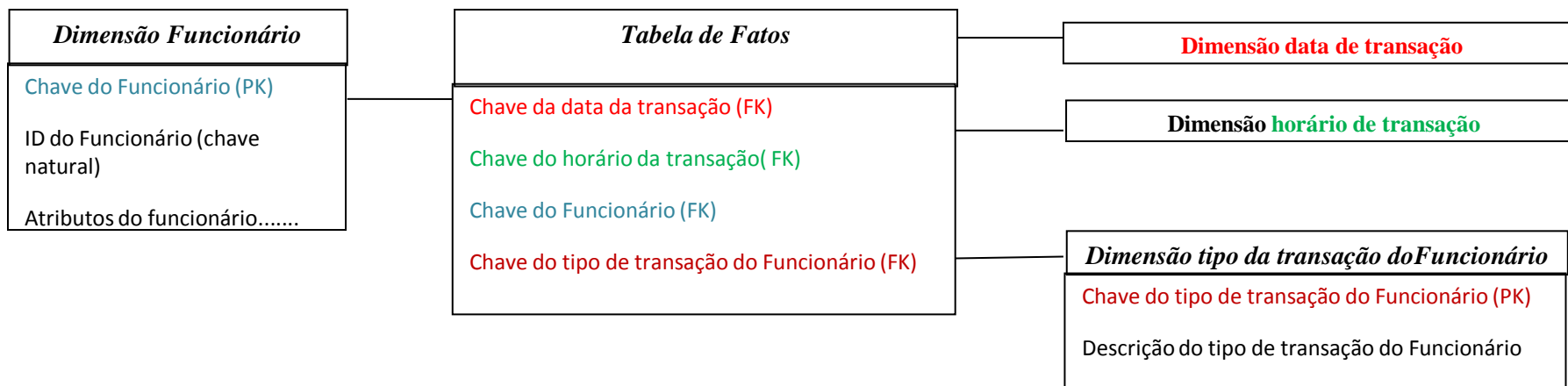


Tabela de Fatos

Não existe medidas numéricas associadas a transação da tabela de fatos, é considerada uma tabela **sem fatos**.

Medidas associadas: por exemplo, um novo endereço ou uma promoção, refére-se ao **tipo da transação do funcionário**.

Qual é o problema deste esquema?

Neste esquema, uma mudança no perfil do funcionário, resultaria em uma nova linha de tipo 2 na dimensão funcionário e uma nova linha na tabela de fatos.

Resultado desta modelagem - o mesmo números de linhas na tabela de fatos e na dimensão de funcionário.

Vamos agora simplificar o esquema

Basta colocar a tabela de fatos de lado e substituir-la para dimensão de transação de funcionário, que contém um perfil de cada mudança do funcionário.

Justificativa: Não existe uma medição numérica associada a transação de perfil, ela apenas resulta em um novo conjunto de características do perfil.

Dimensão de Transação do Funcionário

Chave da transação do funcionário
ID do Funcionário(chave natural)
Nome do Funcionário
Endereço
.....

Descrição do tipo de transação
Data transação do funcionário
Horário transação do funcionário
Data do vencimento da transação do funcionário
Horário do vencimento transação do funcionário
Indicador da transação do funcionário mais recente

Considerações:

Neste esquema, a dimensão de Transação do funcionário com marcação de data e hora pode responder a várias perguntas interessantes sobre RH, como:

- Analisar detalhadamente a sequência de transações de qualquer funcionário
- Traçar o perfil da comunidade dos funcionários em qualquer período específico
- Perguntar em um determinado período, quantos funcionários existiam e qual era o seu perfil

Tabela de dimensão de transação de funcionários não seria um tipo de tabela de fatos porque ela possui data e hora?

Resposta: Apesar de tecnicamente isso ser verdade, esta tabela possui: Valores textuais é a fonte de restrições de consultas e rótulos de relatórios

Portanto, ela é o ponto de entrada para as tabelas RH

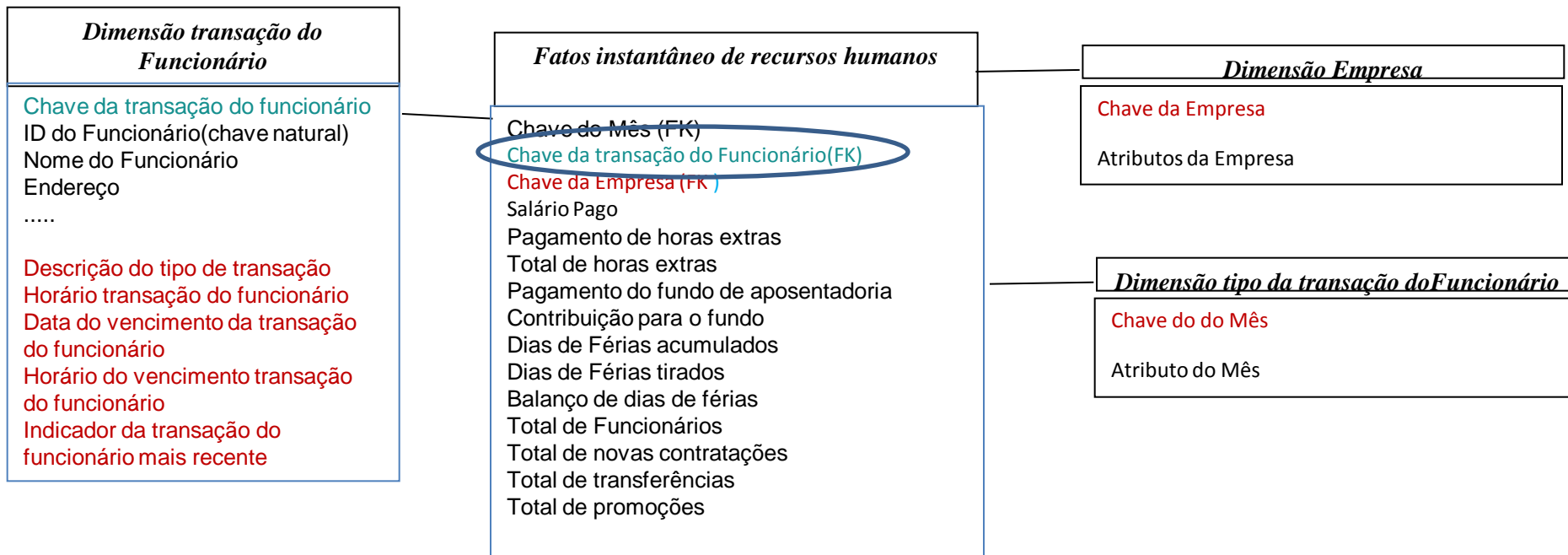
Novo foco: Reportar o status resumido dessa base de dados com a Frequência mensal regular.

Na verdade, estamos interessados em:

- Em contagens;
- Dados estatísticos e totais;
- Números de funcionários;
- Salário total pago ao mês;
- Total de dias de férias tirados
- Dias de férias acumulados
- Números de novas contratações
- Números de novas promoções

Exigência da empresa:

Pretende analisar os dados por todas as vertentes possíveis, inclusive no tempo e no espaço. Precisamos acessar totais no fim do mês, mesmo que não tiver havido transações.



Esta tabela de fatos os totais e contagens mensais respondem a todos os questionamentos de gerenciamentos relacionados aos dados estatísticos mensais. Todos os fatos são aditivos em todas as dimensões ou atributos de dimensões, exceto os atributos identificados como balanço

Por exemplo ao criamos uma linha de uma tabela de fatos, sabemos que:

- O sistema de origem que forneceu os dados de fato;
- A versão do software de extração que criou a linha
- A versão lógica de alocação, se houver uma, que foi usada para criar a linha
- Se determinada coluna de fato "Não aplicavel" é conhecida, é possível, esta danificada ou ainda não está disponível
- Se determinada fato foi alterado depois da carga inicial e, em caso positivo, por que foi alterado
- Se a linha contém fatos com mais de 2, 3 ou 4 desvios padrão na média ou, de modo equivalentes, for a dos limites de confiança derivados de outras análises estatísticas

Os 3 primeiros itens descrevem o tipo de linha de tabelas de fatos, ou seja de onde elas vêm?

Os três últimos itens descrevem a nossa confiança na qualidade dos dados da tabela de fatos

Dimensão transação Funcionário

Chave da transação do funcionário
 ID do Funcionário(chave natural)
 Nome do Funcionário

Fatos instantâneos de recursos humanos

Chave do Mês (FK)
 Chave da transação do Funcionário(FK)
 Chave da Empresa (FK)
 Salário Pago
 Pagamento de horas extras
 Total de horas extras
 Pagamento do fundo de aposentadoria
 Contribuição para o fundo
 Dias de Férias acumulados
 Dias de Férias tirados
 Balanço de dias de férias
 Total de Funcionários
 Total de novas contratações
 Total de transferências
 Total de promoções

Dimensão Empresa

Chave da Empresa
 Atributos da Empresa

Dimensão tipo da transação do Funcionário

Chave do do Mês
 Atributo do Mês

Dimensão auditoria

Chave de auditoria(PK)
 Data de término da extração
 Hora de término da extração
 Status da extração
 Número de registros extraídos
 Data de término da transformação
 Hora de término da transformação
 Status da transformação
 Número de registros transformados
 Número de registros transformados rejeitados
 Data de término da carga
 Hora de término da carga
 Status da Carga
 Número de registros carregados
 Número de registros carregados rejeitados

Objetivo: Complementar a dimensão Funcionário com as informações sobre qualificações técnicas.

Finalidade: Ser capaz de determinar todas as qualificações técnicas dos funcionários de TI.

Regra de negócio:

- Cada funcionário pode ter inúmeras qualificações técnicas
- Qualificações técnicas podem ser consideradas com palavras-chaves, como;
 - Linguagens de programação (Cobol, C++,...)
 - Sistemas (Linux, Unix...)
 - Ou plataformas de Banco de Dados (Oracle, MySQL, PostgreSQL...)

Outrigger do grupo de qualificações do funcionário

Chave do grupo de qualificações	Chave de qualificações	Descrição de qualificações	Categoria de qualificações
1	1	Pascal	
1	2	Cobol	
1	3	Pascal	
2	4	Linux	
2	5	Unix	
2	6	Windons	

Outrigger de grupo de qualificação de funcionário

Chave do grupo de qualificações do funcionário (PK)

Chave de qualificações do funcionário(PK)

Descrição qualificações do funcionário

Categoria de qualificações do funcionário

Utilizando o esquema anterior, teríamos um sério problema para selecionar os funcionários que são especialista em Linux ou Unix.

O uso do operador OR como por exemplo (Linux OR Unix) restringiria a consulta . E se optarmos pelo operador AND, seria difícil utilizá-lo em duas linhas, por que **a SQL não possui muitos recursos para lidar com as restrições de linhas.**

Qual seria a melhor solução para selecionar todos os funcionário que são qualificados linux e Unix, o uso do Operador OR ou AND ?

1ª solução seria o uso do UNION

Select funcionario.id, funcionario.nome

From funcionarios, qualificacao

Where funcionario.grupoqualificacao= qualificacao.grupoqualificacao and
qualificacao="UNIX"

UNION

Select funcionario.id, funcionario.nome

From funcionarios, qualificacao

Where funcionario.grupoqualificacao= qualificacao.grupoqualificacao and
qualificacao="LINUX"

2ª solução seria o uso do substring

SKILL_LIST LIKE '%\ UNIX\%' OR SKILL_LIST LIKE '%\ LINUX\%'

Questionário:

Departamento quer analisar as respostas fornecidas pelos funcionários para determinar a pontuação média do funcionário que foi avaliado

- Criamos na uma tabela de fatos com uma linha para cada pergunta no questionário respondido por um funcionário
- Duas dimensões de funcionário: uma quem respondeu e a outra quem avaliou
- Um dimensão Questionário é formada pela descritores do instrumento de pesquisa
- Dimensão pergunta - fornece a pergunta e sua categoria
- Dimensão resposta – contem as respostas e talvez a categoria das respostas
- Dimensão data do envio do questionário
- Dimensão data do recebimento do questionário

Tendo como finalidade avaliar dados do questionário



Dados de um questionário

Muito Obrigada !!

Vanessa Maia Aguiar de Magalhães

vmaia@cnpqi.embrapa.br / vmaiaufscar@gmail.com