

Um artigo técnico da Oracle
Setembro de 2009

Compressão Avançada com o Oracle Database 11g Release 2

Introdução	3
Compressão Avançada Oracle	3
Compressão para dados de tabela	4
Compressão de Tabela OLTP	4
Compressão para arquivos de dados	7
SecureFiles Deduplication	7
Compressão no SecureFiles	8
Compressão de dados de backup	9
Compressão no Recovery Manager	9
Compressão no Data Pump	9
Compressão para Tráfego na Rede	10
Conclusão	11

Introdução

As empresas estão sofrendo uma explosão no volume de dados necessários para operar seus negócios com eficácia. Essa tendência no crescimento dos dados pode ser atribuída a diversos fatores-chave. Alterações recentes nas regulamentações, como a Sarbanes-Oxley e a HIPAA, contribuem para essa tendência ao obrigar as empresas a reterem imensas quantidades de informações durante longos períodos. A distribuição em massa pela Internet de conteúdo rico e de multimídia, possibilitada pelos avanços nas tecnologias de banda larga, também contribui para o crescimento no volume geral de dados. Alimentando ainda mais a tendência exponencial no crescimento dos dados está o advento da Web 2.0, com aplicativos colaborativos que promovem enormes quantidades de conteúdo gerado por usuários. Diversas estimativas indicam que o volume de dados está praticamente dobrando a cada 2 a 3 anos.

Essa súbita explosão no volume de dados apresenta um desafio assustador de gerenciamento para os administradores de TI. Antes de tudo, estão os custos de armazenagem crescendo em espiral: mesmo com a queda dramática no custo por MB de armazenamento nos últimos anos, o imenso crescimento no volume de dados que devem ser mantidos on-line torna o armazenamento um dos maiores elementos de custo na maioria dos orçamentos de TI. Além disso, a escalabilidade e o desempenho dos aplicativos devem continuar a atender as demandas dos negócios, mesmo quando os volumes de dados explodem.

O Oracle Database 11g apresenta a opção de Compressão Avançada para ajudar os clientes a enfrentar esses desafios. As inovações nas tecnologias de compressão da Oracle ajudam os clientes a reduzir os recursos e os custos de administrar grandes volumes de dados. A introdução dessas empolgantes novas tecnologias surge em uma época na qual bancos de dados de terabytes, antes considerados uma novidade, estão se tornando prevalentes nos centros de dados das empresas.

Compressão Avançada Oracle

A opção de Compressão Avançada do Oracle Database 11g introduz um conjunto abrangente de recursos de compressão para ajudar os clientes a maximizar sua utilização de recursos e a reduzir seus custos. Isso permite aos administradores de TI reduzirem significativamente o espaço geral

de armazenamento nos bancos de dados ao permitir a compressão de todos os tipos de dados – sejam relacionais (tabela), desestruturados (arquivo) ou dados de backup. Embora a economia de custos com armazenamento seja considerada frequentemente como o benefício mais tangível da compressão, tecnologias inovadoras incluídas na opção de Compressão Avançada foram criadas para reduzir os requisitos de recursos e os custos de tecnologia de todos os componentes de sua infraestrutura de TI, incluindo memória e largura de banda na rede.

Compressão para dados de tabela

A Oracle tem sido pioneira na tecnologia de compressão de banco de dados. O Oracle Database 9i introduziu a Compressão Básica de Tabela há muitos anos, quando esses dados comprimidos eram carregados usando operações de cargas em massa. O Oracle Database 11g Release 1 introduziu um novo recurso chamado Compressão de Tabela OLTP, que permite comprimir os dados durante todos os tipos de operações de manipulação de dados, incluindo DML convencional como INSERT e UPDATE. Além disso, a Compressão de Tabela OLTP reduz a sobrecarga de compressão associada de operações de gravação tornando-a adequada para ambientes transacionais bem como OLTP. A Compressão de Tabela OLTP, portanto, estende os benefícios da compressão a todas as cargas de trabalho de aplicativo.

Deve ser observado que a Compressão Básica de Tabela é um recurso base do Oracle Database 11g Enterprise Edition (EE). A Compressão de Tabela OLTP é uma parte da opção Compressão Avançada Oracle, que requer uma licença além da Enterprise Edition.

Compressão de Tabela OLTP

A compressão de tabela OLTP da Oracle usa um algoritmo de compressão exclusivo, projetado especificamente para funcionar com aplicativos OLTP. O algoritmo atua eliminando valores duplicados em um bloco do banco de dados, mesmo em várias colunas. Os blocos comprimidos contêm uma estrutura chamada de tabela de símbolos que mantém os metadados da compressão. Quando um bloco é comprimido, os valores duplicados são eliminados, primeiro adicionando uma cópia do valor duplicado à tabela de símbolos. Cada valor duplicado é então substituído por uma referência curta à entrada apropriada na tabela de símbolos. Com esse design inovador, os dados comprimidos são auto-contidos no bloco do banco de dados enquanto os metadados usados para converter os dados comprimidos para seu estado original são armazenados no bloco. Quando comparado com algoritmos de compressão de concorrentes, que mantêm uma tabela de símbolos do banco de dados global, a abordagem exclusiva da Oracle oferece benefícios significativos de desempenho ao não incluir E/S adicionais ao acessar os dados comprimidos.

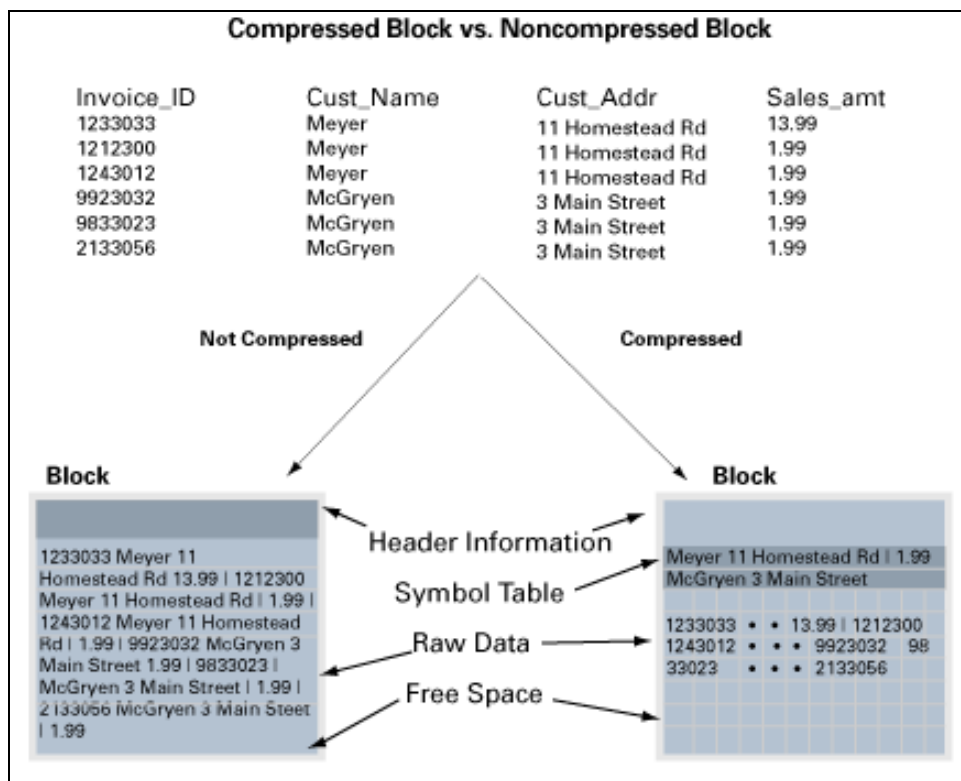


Figura 1: Bloco Comprimido versus Bloco Não Comprimido

Benefícios da Compressão de Tabela OLTP

A taxa de compressão alcançada em determinado ambiente depende da natureza dos dados sendo comprimidos e especificamente da cardinalidade dos dados. Normalmente, os clientes podem esperar reduzir o consumo de seu espaço de armazenamento por um fator de 2 a 3 usando o recurso de Compressão de Tabela OLTP. Ou seja, o total de espaço usado pelos dados sem compressão será duas a três vezes maior que o dos dados comprimidos. Os benefícios da Compressão de Tabela OLTP vão além da simples economia de espaço em disco. Uma vantagem significativa é a capacidade do Oracle de ler blocos comprimidos diretamente sem ter de descomprimir o bloco primeiro. Logo, não ocorre uma degradação mensurável do desempenho ao acessar dados comprimidos. De fato, em muitos casos o desempenho pode ser aprimorado devido à redução de E/S, pois o Oracle precisará acessar menos blocos. Além disso, o cache de buffer se tornará mais eficiente armazenando mais dados sem ter que adicionar memória.

Sobrecarga mínima de desempenho

Como afirmado acima, a Compressão de Tabela OLTP não prejudica as operações de leitura. Existe um trabalho adicional realizado ao gravar dados, tornando impossível eliminar a sobrecarga de desempenho nas operações de gravação. Mas a Oracle trabalhou muito para

minimizar essa sobrecarga para a Compressão de Tabela OLTP. A solução Oracle comprime os blocos em modo de lote em vez de comprimir os dados toda vez que ocorre uma operação de gravação. Um bloco recém inicializado permanece descomprimido até que os dados no bloco alcancem um limite controlado internamente. Quando uma transação faz com que os dados no bloco alcancem esse limite, todo o conteúdo do bloco é comprimido. Subsequentemente, quando mais dados são adicionados ao bloco e o limite é alcançado novamente, todo o bloco é recomprimido para alcançar o mais alto nível de compressão. Esse processo é repetido até que o Oracle determine que o bloco não irá beneficiar-se de mais uma compressão. Somente transações que disparam a compressão do bloco irão sofrer um mínimo de sobrecarga de compressão. Logo, a maioria das transações OLTP em blocos comprimidos terá exatamente o mesmo desempenho que teria com blocos não comprimidos.

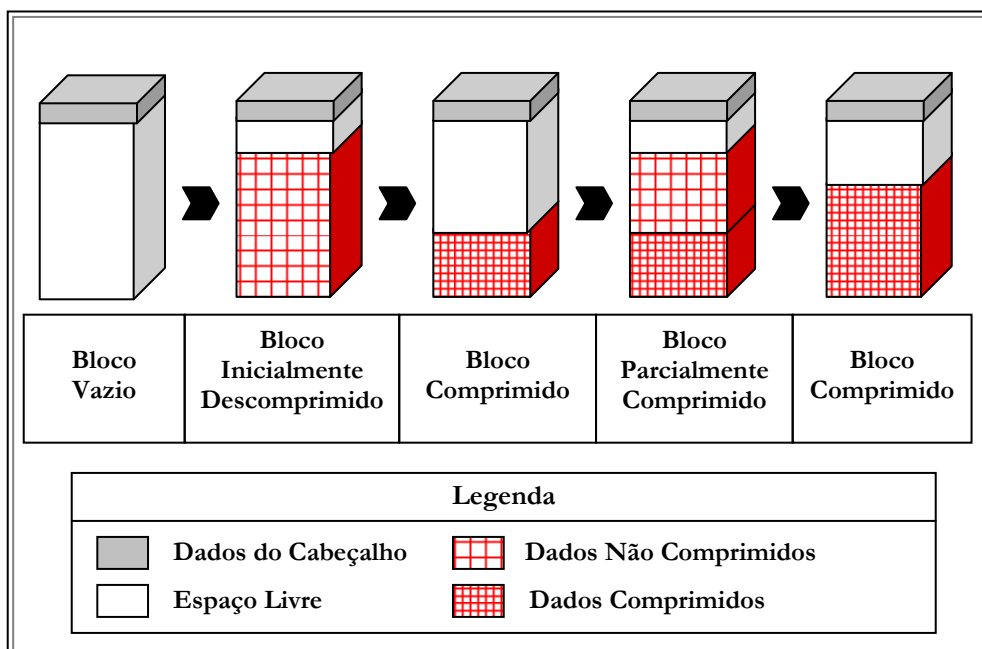


Figura 2 Processo de Compressão de Tabela OLTP

Sintaxe da Compressão de Tabela OLTP

```
CREATE TABLE emp (
    emp_id NUMBER
, first_name VARCHAR2(128)
, last_name VARCHAR2(128)
) COMPRESS FOR OLTP;
```

Compressão para arquivos de dados

O SecureFiles, um novo recurso no Oracle Database 11g, oferece uma arquitetura com o ‘melhor de dois mundos’ para armazenar conteúdo não estruturado, como documentos, planilhas e arquivos XML. O SecureFiles foi criado especificamente para proporcionar alto desempenho para dados de arquivos, em comparação com sistemas de arquivos tradicionais, sem perder as vantagens do banco de dados Oracle. O SecureFiles é projetado como um superconjunto do tipo de dados LOB do padrão ANSI e oferece facilidade na migração de LOBs de BasicFile existentes, o precursor do SecureFiles. Com o SecureFiles, as organizações agora podem administrar no Oracle todos os dados relacionais e dados de arquivos associados, usando um só modelo de segurança/auditoria, um processo unificado de backup e recuperação e realizar recuperações perfeitas por todas as informações. A opção de Compressão Avançada do Oracle Database 11g inclui tecnologias que reduzem drasticamente o espaço de armazenamento de dados nos SecureFiles.

SecureFiles Deduplication

É extremamente comum que os aplicativos armazenem cópias exatas dos arquivos. Um exemplo típico é um aplicativo de e-mail no qual vários usuários podem receber o mesmo anexo. O SecureFiles Deduplication é uma tecnologia inteligente incluída na opção de Compressão Avançada que elimina as duplicatas de dados do SecureFiles. A solução Oracle armazena uma imagem dos dados dos SecureFiles e substitui as cópias duplicadas com referências a essa imagem. Considere um aplicativo de e-mail no qual 10 usuários recebem um e-mail com o mesmo anexo de 1 MB. Sem o SecureFiles Deduplication, o sistema armazenaria uma cópia do arquivo para cada um dos 10 usuários, exigindo 10 MB de armazenamento. Se o aplicativo de e-mail em nosso exemplo utilizar o SecureFiles com Deduplication, ele só irá armazenar um anexo de 1 MB. Isso é uma economia de 90% nos requisitos de armazenamento. Além da economia no armazenamento, o SecureFiles Deduplication também aprimora o desempenho do aplicativo. Especificamente, as operações de gravar e copiar são muito mais eficientes, pois só gravam referências à imagem no SecureFiles. Além disso, as operações de leitura podem ser aprimoradas se ainda existirem dados duplicados dos SecureFiles no cache de buffer.

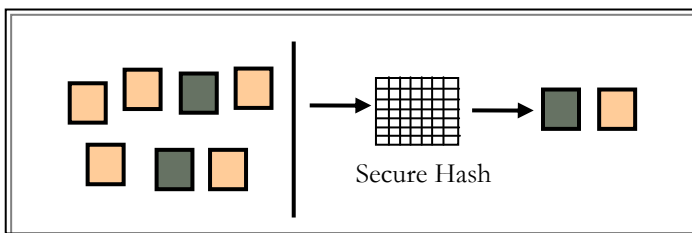


Figura 3 SecureFiles Deduplication

Sintaxe do SecureFiles Deduplication

```
CREATE TABLE images (  
  image_id NUMBER,  
  image BLOB)  
  LOB(image) STORE AS SECUREFILE  
(TABLESPACE lob_tbs DEDUPLICATE);
```

Compressão no SecureFiles

A opção de Compressão Avançada do Oracle Database 11g proporciona mais um mecanismo de controle do tamanho de seus dados nos SecureFiles. Além da eliminação de duplicatas já discutida, a compressão dos SecureFiles utiliza algoritmos padrão de compressão para minimizar ainda mais os requisitos de armazenamento de dados dos SecureFiles. Com a compressão dos SecureFiles, os arquivos típicos, como documentos ou arquivos XML, tem uma redução de tamanho de 2 a 3 vezes. Com sua inteligência incorporada, a compressão dos SecureFiles evita automaticamente comprimir dados que não irão beneficiar-se da compressão, como, por exemplo, uma instância de um documento que já foi comprimido por uma ferramenta de terceiros antes de ser inserido no banco de dados como um arquivo SecureFiles. Os aplicativos ainda conseguem realizar leituras e gravações aleatórias em dados comprimidos em SecureFiles, desde que os dados comprimidos sejam divididos em pequenos blocos de dados. Isso pode aumentar muito o desempenho em comparação com comprimir todos os arquivos antes de inseri-los no banco de dados.

Existem três níveis de compressão de SecureFiles disponíveis: BAIXA, MÉDIA e ALTA. Por padrão, a compressão de SecureFiles usa o nível MÉDIO, que normalmente fornece uma boa compressão com uma pequena sobrecarga de CPU de 3 a 5%. A compressão de SecureFiles BAIXA, nova no Oracle Database 11g Release 2, é otimizada para melhor desempenho. A compressão de SecureFiles BAIXA na verdade mantém cerca de 80% da compressão obtida através da MÉDIA, usando 3x menos a CPU. Por fim, a compressão de SecureFiles ALTA consegue a melhor economia de armazenamento, mas resulta na maior sobrecarga da CPU.

Sintaxe da Compressão de SecureFiles

```
CREATE TABLE images (  
  image_id NUMBER,  
  image BLOB)  
  LOB(image) STORE AS SECUREFILE  
(TABLESPACE lob_tbs COMPRESS);
```

Compressão de dados de backup

Além de comprimir os dados armazenados dentro do banco de dados, a Compressão Avançada Oracle também permite comprimir dados de backup. O Recovery Manager (RMAN) e o Data Pump são as duas ferramentas de uso mais comum para backup dos dados armazenados em um banco de dados Oracle. O RMAN faz um backup bloco-a-bloco dos dados no banco de dados, também conhecido como backup “físico”, que pode ser usado para recuperação em nível de banco de dados, tablespace ou bloco. O Data Pump é usado para realizar um backup “lógico” ao carregar os dados de uma ou mais tabelas em um só arquivo. A Compressão Avançada Oracle inclui a capacidade de comprimir os dados de backup gerados por essas duas ferramentas.

Compressão no Recovery Manager

O crescimento contínuo dos bancos de dados empresariais cria um imenso desafio para os administradores de banco de dados. Os requisitos de armazenamento para manter backups de banco de dados e o desempenho dos procedimentos de backup sofrem impacto direto do tamanho do banco de dados. A Compressão Avançada Oracle inclui a tecnologia RMAN de compressão que pode reduzir drasticamente os requisitos de armazenamento dos dados de backup. Devido à forte integração do RMAN com o banco de dados Oracle, os dados de backup são comprimidos antes de serem gravados em disco ou fita e não é necessário descomprimi-los antes da recuperação, o que proporciona uma imensa redução nos custos de armazenamento.

Existem três níveis de compressão de RMAN: BAIXA, MÉDIA e ALTA. A quantidade de economia de armazenamento aumenta da compressão BAIXA para a ALTA, enquanto potencialmente consome mais recursos da CPU.

Sintaxe para configurar o algoritmo de compressão de RMAN:

```
RMAN> SET COMPRESSION ALGORITHM 'LOW|MEDIUM|HIGH';
```

Sintaxe para obter um backup comprimido RMAN:

```
RMAN> backup as COMPRESSED BACKUPSET database archivelog all;
```

Compressão no Data Pump

A capacidade de comprimir os metadados, associada com um job Data Pump esteve presente primeiro no Oracle Database 10g Release 2. No Oracle Database 11g, esse recurso de compressão foi estendido de forma que dados de tabelas podem ser comprimidos na exportação. A compressão Data Pump é uma operação em linha, de modo que o tamanho reduzido do arquivo de dump significa uma economia significativa de espaço em disco. Diferentemente do

sistema operacional ou de utilitários de compressão de sistema de arquivos, a compressão no Data Pump também está plenamente alinhada no lado da importação, não exigindo descomprimir um arquivo de dump antes de importá-lo. Os conjuntos de arquivos de dump comprimidos são descomprimidos automaticamente durante a importação sem etapas adicionais pelo administrador do banco de dados.

No exemplo de compressão a seguir do banco de dados de amostra da Oracle, os esquemas OE e SH foram exportados enquanto que simultaneamente eram comprimidos todos os dados e metadados. O tamanho do arquivo de dump foi reduzido em 74,67%.

Foram utilizadas três versões do utilitário gzip (GNU zip) e uma do utilitário de compressão do UNIX, para comprimir o conjunto de arquivos de dump de 6,0 MB. A redução no tamanho do arquivo de dump foi comparável à compressão do Data Pump. Observe que a redução no tamanho do arquivo de dump depende dos tipos de dados e de outros fatores.

A funcionalidade plena do Data Pump está disponível quando se utiliza um arquivo comprimido. Qualquer comando utilizado em um arquivo regular também irá funcionar em um arquivo comprimido. Os usuários têm as seguintes opções para determinar quais partes de um conjunto de arquivos de dump devem ser comprimidas:

- **ALL** ativa a compressão para toda a operação de exportação.
- **DATA-ONLY** faz com que todos os dados sejam gravados no arquivo de dump em formato comprimido.
- **METADATA-ONLY** faz com que todos os metadados sejam gravados no arquivo de dump em formato comprimido. Esse é o padrão.
- **NONE** desativa a compressão para toda a operação de exportação.

Para obter mais informações sobre o Oracle Data Pump, visite

<http://www.oracle.com/technology/products/database/utilities/index.html>

Como Ativar a Compressão para o Data Pump

O exemplo abaixo mostra como habilitar a compressão para o Data Pump:

```
expdp hr FULL=y DUMPFILE=dpump_dir:full.dmp COMPRESS;
```

Compressão para Tráfego na Rede

O Oracle Data Guard proporciona a infraestrutura de software de automação, monitoração e gerenciamento para criar, manter e monitorar um ou mais bancos de dados de stand-by para proteger os dados empresariais contra falhas, desastres, erros e corrupção de dados. O Data Guard mantém a sincronização dos bancos de dados principal e de stand-by usando dados de recuperação (as informações necessárias para recuperar uma transação). As transações ocorrem

no banco de dados principal, os dados de recuperação são gerados e gravados nos arquivos de log de recuperação locais. Os Data Guard Redo Transport Services são usados para transferir esses dados de recuperação para o(s) local(is) de stand-by. Com a Compressão Avançada, os dados de recuperação podem ser transmitidos em um formato comprimido para reduzir o consumo de largura de banda da rede e, em alguns casos, reduzir o tempo de transmissão dos dados de recuperação. A partir do Oracle Database 11g Release 2, os dados de recuperação podem ser transmitidos em um formato comprimido quando a configuração do Oracle Data Guard usar tanto o transporte de dados de recuperação síncrono (SYNC) como o transporte de dados de recuperação assíncrono (ASYNC).

Para obter mais informações sobre o Oracle Data Guard, visite <http://www.oracle.com/technology/deploy/availability/index.html>

Conclusão

A explosão no volume de dados que as empresas enfrentam apresenta desafios significativos. As empresas devem adaptar-se rapidamente às mudanças no cenário dos negócios sem impactar seus resultados finais. Os gerentes de TI necessitam gerenciar com eficácia sua infraestrutura existente, para controlar custos, e continuar a apresentar um desempenho extraordinário de aplicativos.

A opção de Compressão Avançada do Oracle Database 11g proporciona um conjunto robusto de recursos de compressão que permitem aos gerentes de TI ter sucesso nesse ambiente complexo. Ao usar a opção de Compressão Avançada, as empresas podem administrar com eficiência suas crescentes exigências de dados em todos os componentes de seus centros de dados, minimizando os custos e continuando a alcançar os maiores níveis de desempenho de aplicativos.



Compressão Avançada com
Oracle Database 11g Release 2
Setembro de 2009
Autor: William Hodak
Autores colaboradores: Kevin Jernigan

Oracle do Brasil Sistemas Ltda.
Sede no Brasil
Av. Alfredo Egydio de Souza Aranha, 100
São Paulo, SP
Brasil

CNPJ: 59.456.277/0001-76
Fone: (0xx11) 5189-1000
oracle.com



Oracle is committed to developing practices and products that help protect the environment

Copyright © 2009, Oracle e/ou suas afiliadas. Todos os direitos reservados. Este documento é fornecido apenas para fins informativos e seu conteúdo está sujeito a alteração sem aviso prévio. Não há garantias de que este documento esteja isento de erros e nem que esteja sujeito a outras garantias ou condições legais, expressas ou implícitas, incluindo garantias ou condições de comercialização e uso para um propósito específico. A Oracle isenta-se de qualquer responsabilidade em relação a este documento, sendo que ele não representa qualquer obrigação contratual direta ou indireta. Este documento não pode ser reproduzido ou transmitido de qualquer forma ou através de qualquer meio, seja eletrônico ou mecânico, para qualquer objetivo, sem a permissão expressa por escrito da Oracle.

Oracle é uma marca comercial registrada da Oracle Corporation e/ou de suas empresas afiliadas. Outros nomes podem ser marcas comerciais de seus respectivos proprietários.