



Informatica® PowerCenter
10.5.6

Guia de Ajuste de Desempenho

© Copyright Informatica LLC 2000, 2024

Este software e a documentação são fornecidos somente sob um contrato de licença separado, contendo restrições sobre uso e divulgação. Não está permitida de forma alguma a reprodução ou a transmissão de qualquer parte deste documento (seja por meio eletrônico, fotocópia, gravação ou quaisquer outros meios) sem o consentimento prévio da Informatica LLC.

Informatica, o logotipo Informatica e PowerCenter são marcas comerciais ou marcas registradas da Informatica LLC nos Estados Unidos e em muitas jurisdições por todo o mundo. Uma lista atual das marcas comerciais da Informatica está disponível na Internet em <https://www.informatica.com/trademarks.html>. Os nomes de outras companhias e produtos podem ser nomes ou marcas comerciais de seus respectivos proprietários.

DIREITOS DO GOVERNO DOS ESTADOS UNIDOS Programas, softwares, bancos de dados, bem como a documentação e os dados técnicos relacionados, distribuídos a clientes do Governo dos EUA são "softwares de computador comerciais" ou "dados técnicos comerciais", de acordo com o Regulamento de Aquisição Federal aplicável e os regulamentos suplementares específicos da agência. Como tal, a utilização, duplicação, divulgação, modificação e adaptação estão sujeitas às restrições e aos termos de licença estabelecidos no contrato governamental aplicável e, na medida do que for aplicável pelos termos do contrato governamental, aos direitos adicionais estabelecidos no FAR 52.227-19, Licença de Software de Computador Comercial.

Partes desta documentação e/ou software estão sujeitas a copyright de terceiros, incluindo sem limitação: Copyright DataDirect Technologies. Todos os direitos reservados. Copyright © Sun Microsystems. Todos os direitos reservados. Copyright © RSA Security Inc. Todos os direitos reservados. Copyright © Ordinal Technology Corp. Todos os direitos reservados. Copyright © Aandacht c.v. Todos os direitos reservados. Copyright Genivia, Inc. Todos os direitos reservados. Copyright Isomorphic Software. Todos os direitos reservados. Copyright © Meta Integration Technology, Inc. Todos os direitos reservados. Copyright © Intalio. Todos os direitos reservados. Copyright © Oracle. Todos os direitos reservados. Copyright © Adobe Systems Incorporated. Todos os direitos reservados. Copyright © DataArt, Inc. Todos os direitos reservados. Copyright © ComponentSource. Todos os direitos reservados. Copyright © Microsoft Corporation. Todos os direitos reservados. Copyright © Rogue Wave Software, Inc. Todos os direitos reservados. Copyright © Teradata Corporation. Todos os direitos reservados. Copyright © Yahoo! Inc. Todos os direitos reservados. Copyright © Glyph & Cog, LLC. Todos os direitos reservados. Copyright © Thinkmap, Inc. Todos os direitos reservados. Copyright © Clearpace Software Limited. Todos os direitos reservados. Copyright © Information Builders, Inc. Todos os direitos reservados. Copyright © OSS Nokalva, Inc. Todos os direitos reservados. Copyright Edifecs, Inc. Todos os direitos reservados. Copyright Cleo Communications, Inc. Todos os direitos reservados. Copyright © International Organization for Standardization 1986. Todos os direitos reservados. Copyright © ej-technologies GmbH. Todos os direitos reservados. Copyright © Jaspersoft Corporation. Todos os direitos reservados. Copyright © International Business Machines Corporation. Todos os direitos reservados. Copyright © yWorks GmbH. Todos os direitos reservados. Copyright © Lucent Technologies. Todos os direitos reservados. Copyright © University of Toronto. Todos os direitos reservados. Copyright © Daniel Veillard. Todos os direitos reservados. Copyright © Unicode, Inc. Copyright IBM Corp. Todos os direitos reservados. Copyright © MicroQuill Software Publishing, Inc. Todos os direitos reservados. Copyright © PassMark Software Pty Ltd. Todos os direitos reservados. Copyright © LogiXML, Inc. Todos os direitos reservados. Copyright © 2003-2010 Lorenzi Davide, todos os direitos reservados. Copyright © Red Hat, Inc. Todos os direitos reservados. Copyright © The Board of Trustees of the Leland Stanford Junior University. Todos os direitos reservados. Copyright © EMC Corporation. Todos os direitos reservados. Copyright © Flexera Software. Todos os direitos reservados. Copyright © Jinfonet Software. Todos os direitos reservados. Copyright © Apple Inc. Todos os direitos reservados. Copyright © Telerik Inc. Todos os direitos reservados. Copyright © BEA Systems. Todos os direitos reservados. Copyright © PDFlib GmbH. Todos os direitos reservados. Copyright © Orientation in Objects GmbH. Todos os direitos reservados. Copyright © Tanuki Software, Ltd. Todos os direitos reservados. Copyright © Ricebridge. Todos os direitos reservados. Copyright © Sencha, Inc. Todos os direitos reservados. Copyright © Scalable Systems, Inc. Todos os direitos reservados. Copyright © jqWidgets. Todos os direitos reservados. Copyright © Tableau Software, Inc. Todos os direitos reservados. Copyright © MaxMind, Inc. Todos os direitos reservados. Copyright © TMate Software s.r.o. Todos os direitos reservados. Copyright © MapR Technologies Inc. Todos os direitos reservados. Copyright © Amazon Corporate LLC. Todos os direitos reservados. Copyright © Highsoft. Todos os direitos reservados. Copyright © Python Software Foundation. Todos os direitos reservados. Copyright © BeOpen.com. Todos os direitos reservados. Copyright © CNRI. Todos os direitos reservados.

Este produto inclui software desenvolvido pela Apache Software Foundation (<http://www.apache.org/>) e/ou outros softwares licenciados nas várias versões da Licença Apache (a "Licença"). Você pode obter uma cópia dessas Licenças em <http://www.apache.org/licenses/>. A menos que exigido pela legislação aplicável ou concordado por escrito, o software distribuído em conformidade com estas Licenças é fornecido "NO ESTADO EM QUE SE ENCONTRA", SEM GARANTIA OU CONDIÇÃO DE QUALQUER TIPO, seja expressa ou implícita. Consulte as Licenças para conhecer as limitações e as permissões que regulam o idioma específico de acordo com as Licenças.

Este produto inclui software desenvolvido pela Mozilla (<http://www.mozilla.org/>), direitos autorais de software de The JBoss Group, LLC; todos os direitos reservados; software copyright © 1999-2006 de Bruno Lowagie e Paulo Soares e outros produtos de software licenciados sob a Licença Pública GNU Lesser General Public License Agreement, que pode ser encontrada em <http://www.gnu.org/licenses/lgpl.html>. Os materiais são fornecidos gratuitamente pela Informatica, no estado em que se encontram, sem garantia de qualquer tipo, explícita nem implícita, incluindo, mas não limitando-se, as garantias implicadas de comerciabilidade e adequação a um determinado propósito.

O produto inclui software ACE(TM) e TAO(TM) com copyright de Douglas C. Schmidt e seu grupo de pesquisa na Washington University, University of California, Irvine e Vanderbilt University, Copyright (©) 1993-2006, todos os direitos reservados.

Este produto inclui o software desenvolvido pelo OpenSSL Project para ser usado no kit de ferramentas OpenSSL (copyright The OpenSSL Project. Todos os direitos reservados) e a redistribuição deste software está sujeita aos termos disponíveis em <http://www.openssl.org> e <http://www.openssl.org/source/license.html>.

Este produto inclui o software Curl com o Copyright 1996-2013, Daniel Stenberg, <daniel@haxx.se>. Todos os direitos reservados. Permissões e limitações relativas a este software estão sujeitas aos termos disponíveis em <http://curl.haxx.se/docs/copyright.html>. É permitido usar, copiar, modificar e distribuir este software com qualquer objetivo, com ou sem taxa, desde que a nota de direitos autorais acima e esta nota de permissão apareçam em todas as cópias.

O produto inclui software copyright 2001-2005 (©) MetaStuff, Ltd. Todos os direitos reservados. Permissões e limitações relativas a este software estão sujeitas aos termos disponíveis em <http://www.dom4j.org/license.html>.

Este produto inclui o copyright de software © 1996-2006 Per Bothner. Todos os direitos reservados. O direito de usar tais materiais é estabelecido na licença que pode ser encontrada em <http://www.gnu.org/software/kawa/Software-License.html>.

Este produto inclui o software OSSP UUID com Copyright © 2002 Ralf S. Engelschall, Copyright © 2002 e OSSP Project Copyright © 2002 Cable & Wireless Deutschland. Permissões e limitações relativas a este software estão sujeitas aos termos disponíveis em <http://www.opensource.org/licenses/mit-license.php>.

Este produto inclui software desenvolvido pela Boost (<http://www.boost.org/>) ou sob a licença de software Boost. Permissões e limitações relativas a este software estão sujeitas aos termos disponíveis em http://www.boost.org/LICENSE_1_0.txt.

Este produto inclui software copyright © 1997-2007 University of Cambridge. Permissões e limitações relativas a este software estão sujeitas aos termos disponíveis em <http://www.pcre.org/license.txt>.

Este produto inclui o copyright de software © 2007 The Eclipse Foundation. Todos os direitos reservados. As permissões e as limitações relativas a este software estão sujeitas aos termos disponíveis em <http://www.eclipse.org/org/documents/epl-v10.php> e em <http://www.eclipse.org/org/documents/edl-v10.php>.

Este produto inclui softwares licenciados de acordo com os termos disponíveis em <http://www.tcl.tk/software/tcltk/license.html>, <http://www.bosrup.com/web/overlib?License>, <http://www.stlport.org/doc/license.html>, <http://asm.ow2.org/license.html>, <http://www.cryptix.org/LICENSE.TXT>, <http://hsqldb.org/web/hsqldbLicense.html>, <http://httpunit.sourceforge.net/doc/license.html>, <http://jung.sourceforge.net/license.txt>, http://www.gzip.org/zlib/zlib_license.html, <http://www.openldap.org/software/release/license.html>, <http://www.libssh2.org>, <http://sf4j.org/license.html>, <http://www.sente.ch/software/OpenSourceLicense.html>,

fusesource.com/downloads/license-agreements/fuse-message-broker-v-5-3- license-agreement; <http://antlr.org/license.html>; <http://aopalliance.sourceforge.net/>; <http://www.bouncycastle.org/license.html>; <http://www.jgraph.com/jgraphdownload.html>; <http://www.jcraft.com/jsch/LICENSE.txt>; http://jotm.objectweb.org/bsd_license.html; <http://www.w3.org/Consortium/Legal/2002/copyright-software-20021231>; <http://www.slf4j.org/license.html>; <http://nanoxml.sourceforge.net/orig/copyright.html>; <http://www.json.org/license.html>; <http://forge.ow2.org/projects/javaservice/>; <http://www.postgresql.org/about/license.html>; <http://www.sqlite.org/copyright.html>; <http://www.tcl.tk/software/tcltk/license.html>; <http://www.jaxen.org/faq.html>; <http://www.jdom.org/docs/faq.html>; <http://www.slf4j.org/license.html>; <http://www.iodbc.org/dataspace/iodbc/wiki/iODBC/License>; <http://www.keplerproject.org/md5/license.html>; <http://www.toedter.com/en/jcalendar/license.html>; <http://www.edankert.com/bounce/index.html>; <http://www.net-snmp.org/about/license.html>; <http://www.openmdx.org/#FAQ>; http://www.php.net/license/3_01.txt; <http://srp.stanford.edu/license.txt>; <http://www.schneier.com/blowfish.html>; <http://www.jmock.org/license.html>; <http://xsom.java.net>; <http://benalman.com/about/license/>; <https://github.com/CreateJS/EaselJS/blob/master/src/easeljs/display/Bitmap.js>; <http://www.h2database.com/html/license.html#summary>; <http://jsoncpp.sourceforge.net/LICENSE>; <http://jdbc.postgresql.org/license.html>; <http://protobuf.googlecode.com/svn/trunk/src/google/protobuf/descriptor.proto>; <https://github.com/rantav/hector/blob/master/LICENSE>; <http://web.mit.edu/Kerberos/krb5-current/doc/mitK5license.html>; <http://jibx.sourceforge.net/jibx-license.html>; <https://github.com/lyokato/libgeohash/blob/master/LICENSE>; <https://github.com/hjiang/jsonxx/blob/master/LICENSE>; <https://code.google.com/p/lz4/>; <https://github.com/jedisct1/libsodium/blob/master/LICENSE>; <http://one-jar.sourceforge.net/index.php?page=documents&file=license>; <https://github.com/EsotericSoftware/kryo/blob/master/license.txt>; <http://www.scala-lang.org/license.html>; <https://github.com/tinkerpop/blueprints/blob/master/LICENSE.txt>; <http://gee.cs.oswego.edu/dl/classes/EDU/oswego/cs/dl/util/concurrent/intro.html>; <https://aws.amazon.com/asl/>; <https://github.com/twbs/bootstrap/blob/master/LICENSE>; <https://sourceforge.net/p/xmlunit/code/HEAD/tree/trunk/LICENSE.txt>.

Este produto inclui software licenciado de acordo com a Academic Free License (<http://www.opensource.org/licenses/afl-3.0.php>), a Common Development and Distribution License (<http://www.opensource.org/licenses/cddl1.php>), a Common Public License (<http://www.opensource.org/licenses/cpl1.0.php>), a Sun Binary Code License Agreement Supplemental License Terms, a BSD License (<http://www.opensource.org/licenses/bsd-license.php>), a nova BSD License (<http://opensource.org/licenses/BSD-3-Clause>), a MIT License (<http://www.opensource.org/licenses/mit-license.php>), a Artistic License (<http://www.opensource.org/licenses/artistic-license-1.0>) e a Initial Developer's Public License Version 1.0 (<http://www.firebirdsql.org/en/initial-developer-s-public-license-version-1-0/>).

Este produto inclui copyright do software © 2003-2006 Joe Walnes, 2006-2007 XStream Committers. Todos os direitos reservados. Permissões e limitações relativas a este software estão sujeitas aos termos disponíveis em <http://xstream.codehaus.org/license.html>. Este produto inclui software desenvolvido pelo Indiana University Extreme! Lab. Para obter mais informações, visite <http://www.extreme.indiana.edu/>.

Este produto inclui software Copyright © 2013 Frank Balluffi e Markus Moeller. Todos os direitos reservados. As permissões e limitações relativas a este software estão sujeitas aos termos da licença MIT.

Consulte as patentes em <https://www.informatica.com/legal/patents.html>.

ISENÇÃO DE RESPONSABILIDADE: a Informatica LLC fornece esta documentação no estado em que se encontra, sem garantia de qualquer tipo, expressa ou implícita, incluindo, mas não limitando-se, as garantias implícitas de não infração, comercialização ou uso para um determinado propósito. A Informatica LLC não garante que este software ou documentação não contenha erros. As informações fornecidas neste software ou documentação podem incluir imprecisões técnicas ou erros tipográficos. As informações deste software e documentação estão sujeitas a alterações a qualquer momento sem aviso prévio.

AVISOS

Este produto da Informatica (o "Software") traz determinados drivers (os "drivers da DataDirect") da DataDirect Technologies, uma empresa em funcionamento da Progress Software Corporation ("DataDirect"), que estão sujeitos aos seguintes termos e condições:

1. OS DRIVERS DA DATADIRECT SÃO FORNECIDOS NO ESTADO EM QUE SE ENCONTRAM, SEM GARANTIA DE QUALQUER TIPO, EXPRESSA OU IMPLÍCITA, INCLUINDO, MAS NÃO LIMITANDO-SE, AS GARANTIAS IMPLÍCITAS DE COMERCIALIZAÇÃO, ADEQUAÇÃO A UMA FINALIDADE ESPECÍFICA E NÃO INFRAÇÃO.
2. EM NENHUM CASO, A DATADIRECT OU SEUS FORNECEDORES TERCEIRIZADOS SERÃO RESPONSÁVEIS, EM RELAÇÃO AO CLIENTE FINAL, POR QUAISQUER DANOS DIRETOS, INDIRETOS, INCIDENTAIS, ESPECIAIS, CONSEQUENCIAIS OU DEMAIS QUE POSSAM ADVIR DO USO DE DRIVERS ODBC, SENDO OU NÃO ANTERIORMENTE INFORMADOS DAS POSSIBILIDADES DE TAIS DANOS. ESTAS LIMITAÇÕES SE APLICAM A TODAS AS CAUSAS DE AÇÃO, INCLUINDO, SEM LIMITAÇÕES, QUEBRA DE CONTRATO, QUEBRA DE GARANTIA, NEGLIGÊNCIA, RESPONSABILIDADE RIGOROSA, DETURPAÇÃO E OUTROS ATOS ILÍCITOS.

As informações contidas neste documento estão sujeitas a alteração sem aviso prévio. Se você encontrar quaisquer problemas nesta documentação, informe-os em infa_documentation@informatica.com.

Os produtos Informatica apresentam garantias segundo os termos e condições dos acordos em que são fornecidos. A INFORMATICA FORNECE AS INFORMAÇÕES NESTE DOCUMENTO "COMO ESTÃO" SEM GARANTIA DE QUALQUER TIPO, EXPRESSA OU IMPLÍCITA, INCLUINDO, SEM QUAISQUER GARANTIAS DE COMERCIALIZAÇÃO, ADEQUAÇÃO A UM DETERMINADO FIM E QUALQUER GARANTIA OU CONDIÇÃO DE NÃO-VIOLAÇÃO.

Data da Publicação: 2024-08-07

Conteúdo

Prefácio.....	9
Recursos da Informatica.	9
Rede da Informatica.	9
Base de Dados de Conhecimento da Informatica.	9
Documentação da Informatica.	9
Matrizes de Disponibilidade de Produto da Informatica.	10
Informatica Velocity.	10
Informatica Marketplace.	10
Suporte Global a Clientes da Informatica.	10
 Capítulo 1: Visão geral de Ajuste de desempenho.....	11
Visão geral de Ajuste de desempenho.	11
 Capítulo 2: Afunilamentos.....	12
Visão geral do capítulo Afunilamentos.	12
Usando estatísticas de segmento.	13
Eliminando afunilamentos com base em estatísticas de segmento.	13
Exemplo.	13
Afunilamentos de destino.	14
Identificando afunilamentos de destino.	14
Eliminando afunilamentos no destino.	14
Afunilamentos de origem.	15
Identificando afunilamentos de origem.	15
Eliminando afunilamentos na origem.	16
Afunilamentos de mapeamento.	16
Identificando afunilamentos de mapeamento.	16
Eliminando afunilamentos de mapeamento.	16
Afunilamentos de sessão.	17
Identificando afunilamentos de sessão.	17
Eliminando afunilamentos de sessão.	17
Afunilamentos do sistema.	17
Identificando afunilamentos do sistema.	17
Eliminando afunilamentos de sistema.	18
 Capítulo 3: Otimizando o destino.....	20
Otimizando destinos de arquivo simples.	20
Descartando restrições de índices e chave.	20
Aumentando os intervalos do ponto de verificação do banco de dados.	21
Usando cargas em massa.	21
Usando carregadores externos.	21

Minimizando deadlocks.	22
Aumentando o tamanho do pacote de rede do banco de dados.	22
Otimizando os bancos de dados de destino Oracle.	22
Capítulo 4: Otimizando a origem.	24
Otimizando a consulta.	24
Usando filtros condicionais.	25
Aumentando o tamanho do pacote de rede do banco de dados.	25
Conectando-se às origens de banco de dados Oracle.	25
Usando o Teradata FastExport.	25
Usando o tempdb para associar o Sybase ou as tabelas do Microsoft SQL Server.	26
Capítulo 5: Otimizando mapeamentos.	27
Visão geral de Otimizando mapeamentos.	27
Otimizando origens de arquivos simples.	27
Otimizando o comprimento do buffer sequencial em linha.	27
Otimizando origens de arquivos simples delimitados.	28
Otimizando origens de arquivos XML e simples.	28
Configurando a leitura de passagem única.	28
Otimizando mapeamentos de passagem.	29
Otimizando filtros.	29
Otimizando conversões de tipos de dados.	29
Otimizando expressões.	30
Fatoramento de lógica comum.	30
Minimizando chamadas de função de agregação.	30
Substituindo expressões comuns por variáveis locais.	30
Escolhendo operações numéricas versus de string.	30
Otimizando comparações Char-Char e Char-Varchar.	30
Escolhendo DECODE versus LOOKUP.	31
Usando operadores em vez de funções.	31
Otimizando funções IIF.	31
Avaliando expressões.	32
Otimizando procedimentos externos.	32
Capítulo 6: Otimizando transformações.	33
Otimizando transformações de Agregador.	33
Agrupando por colunas simples.	33
Usando entrada classificada.	34
Usando agregação incremental.	34
Filtrando dados antes de agregar.	34
Limitando conexões de porta.	34
Otimizando transformações personalizadas.	34
Otimizando transformações de unificador.	35

Otimizando transformações de pesquisa.	35
Usando drivers de banco de dados ideais.	36
Armazenando tabelas de pesquisa em cache.	36
Otimizando a condição de pesquisa.	38
Filtrando linhas de pesquisa.	38
Indexação da tabela de pesquisa.	38
Otimizando várias pesquisas.	38
Criando uma transformação pesquisa de pipeline.	39
Otimizando Transformações de Normalizador.	39
Otimizando transformações de Gerador de Sequência.	39
Otimizando transformações de classificador.	39
Alocando memória.	40
Diretórios de trabalho para partições.	40
Modo Unicode.	40
Otimizando transformações de Qualificador de Origem.	41
Otimizando transformações de SQL.	41
Otimizando Transformações de XML.	41
Eliminando erros de transformação.	42
Capítulo 7: Otimizando sessões.	43
Grade.	43
Otimização de empilhamento.	44
Sessões e fluxos de trabalho simultâneos.	44
Memória de buffer.	44
Aumentando o tamanho do buffer DTM.	45
Otimizando o tamanho do bloco de buffer.	45
Caches.	46
Limitando o número de portas conectadas	46
Localização do diretório de cache.	46
Aumentando os tamanhos de cache	47
Usando a versão de 64 bits do PowerCenter.	47
Confirmação baseada no destino.	47
Processamento em tempo real.	48
Latência de liberação.	48
Confirmações baseadas na origem.	48
Áreas de preparação.	48
Arquivos de log.	48
Rastreamento de erros.	48
Emails pós-sessão.	49
Capítulo 8: Otimizando implantações de grade.	50
Visão geral do capítulo Otimizando a implantação de grades.	50
Armazenando arquivos.	50

Arquivos do sistema de arquivos compartilhado de alta largura de banda.	51
Arquivos do sistema de arquivos compartilhado de baixa largura de banda.	51
Arquivos de armazenamento local.	51
Usando um sistema de arquivos compartilhado.	52
Configurando um sistema de arquivos compartilhado.	52
Equilibrando o uso da CPU e da memória.	52
Configurando mapeamentos e sessões do PowerCenter.	53
Distribuindo arquivos pelos sistemas de arquivos.	54
Configurando sessões para distribuir arquivos.	54
Otimizando transformações de Gerador de Sequência.	56

Capítulo 9: Otimizando componentes do PowerCenter..... 57

Visão geral do capítulo Otimizando componentes do PowerCenter.	57
Otimizando o desempenho do repositório do PowerCenter.	57
Localização do repositório e do processo do Serviço de Repositório.	58
Condições de ordem em consultas de objeto.	58
Usando um espaço de tabela do banco de dados DB2 de nó único.	58
Otimizando o esquema do banco de dados	58
Cache de Objeto para o Serviço de Repositório.	59
Otimização da resiliência.	60
Otimizando o desempenho do Serviço de Integração.	60
Usando drivers ODBC e nativo.	60
Executando o Serviço de Integração no modo de movimentação de dados ASCII.	60
Criando cache no PowerCenter Metadata para o Serviço de Repositório	61

Capítulo 10: Otimizando o sistema..... 62

Visão geral do capítulo Otimização do sistema.	62
Aprimorando a velocidade da rede.	63
Usando várias CPUs.	63
Reduzindo a paginação.	63
Usando a associação de processador.	64

Capítulo 11: Usando partições de pipeline..... 65

Visão geral de Usando partições de pipeline.	65
Aumentando o número de partições.	65
Selecionando os tipos de partição de melhor desempenho.	66
Usando várias CPUs.	67
Otimizando o banco de dados de origem para particionamento.	67
Ajustando o banco de dados.	67
Agrupando dados classificados.	68
Otimizando consultas de classificação única.	68
Otimizando o banco de dados de destino para particionamento.	69

Apêndice A: Contadores de desempenho.....	70
Visão geral do capítulo Contadores de desempenho.	70
Contador Errorrows.	70
Contadores Readfromcache e Writetocache.	71
Contadores Readfromdisk e Writetodisk.	71
Contador Rowsinlookupcache.	72
Índice.....	73

Prefácio

Consulte o *Guia de Ajuste de Desempenho do PowerCenter®* para saber mais sobre gargalos em tempo de execução e como fazer ajustes para obter o desempenho ideal.

Recursos da Informatica

A Informatica oferece uma variedade de recursos de produtos através da Rede da Informatica e outros portais on-line. Use os recursos para obter o máximo de seus produtos e soluções da Informatica e para aprender com outros usuários da Informatica e especialistas no assunto.

Rede da Informatica

A Rede da Informatica é a porta de entrada para muitos recursos, incluindo a Base de Dados de Conhecimento da Informatica e o Suporte Global a Clientes da Informatica. Para acessar a Rede da Informatica, visite <https://network.informatica.com>.

Como membro da Rede da Informatica, você tem as seguintes opções:

- Pesquisar por recursos do produto na Base de Dados de Conhecimento.
- Visualizar informações sobre disponibilidade de produtos.
- Criar e revisar seus casos de suporte.
- Encontrar a sua Rede de Grupo de Usuários da Informatica local e colaborar com seus colegas.

Base de Dados de Conhecimento da Informatica

Use a Base de Dados de Conhecimento da Informatica para encontrar recursos de produtos, como artigos de instruções, práticas recomendadas, tutoriais em vídeo e respostas a perguntas frequentes.

Para pesquisar na Base de Dados de Conhecimento, visite <https://search.informatica.com>. Em caso de dúvidas, comentários ou ideias sobre a Base de Dados de Conhecimento, entre em contato com a equipe da Base de Dados de Conhecimento da Informatica em KB_Feedback@informatica.com.

Documentação da Informatica

Use o Portal de Documentação da Informatica para explorar uma extensa biblioteca de documentação para versões de produtos atuais e recentes. Para explorar o Portal de Documentação, visite <https://docs.informatica.com>.

Em caso de dúvidas, comentários ou ideias sobre a documentação do produto, entre em contato com a equipe da Documentação da Informatica em infa_documentation@informatica.com.

Matrizes de Disponibilidade de Produto da Informatica

As Matrizes de Disponibilidade de Produto (PAMs) indicam as versões dos sistemas operacionais, os bancos de dados e tipos de fontes e destinos de dados com os quais uma versão de produto é compatível. Veja as PAMs da Informatica em <https://network.informatica.com/community/informatica-network/product-availability-matrices>.

Informatica Velocity

O Informatica Velocity é uma coleção de dicas e práticas recomendadas desenvolvidas pelos Serviços Profissionais da Informatica e baseada em experiências reais de centenas de projetos de gerenciamento de dados. O Informatica Velocity representa o conhecimento coletivo dos consultores da Informatica que trabalham com organizações em todo o mundo para planejar, desenvolver, implantar e manter soluções de gerenciamento de dados bem-sucedidas.

Encontre os recursos do Informatica Velocity em <http://velocity.informatica.com>. Se você tiver dúvidas, comentários ou ideias sobre o Informatica Velocity, entre em contato com os Serviços Profissionais da Informatica em ips@informatica.com.

Informatica Marketplace

O Informatica Marketplace é um fórum onde você pode encontrar soluções que ampliam e aprimoram suas implementações da Informatica. Aproveite as centenas de soluções dos desenvolvedores e parceiros da Informatica no Marketplace para melhorar sua produtividade e agilizar o tempo de implementação em seus projetos. Encontre o Informatica Marketplace em <https://marketplace.informatica.com>.

Suporte Global a Clientes da Informatica

Você pode entrar em contato com um Centro de Suporte Global por telefone ou por meio da Rede da Informatica.

Para descobrir o número de telefone local do Suporte Global a Clientes da Informatica, visite o site da Informatica no seguinte link: <https://www.informatica.com/services-and-training/customer-success-services/contact-us.html>.

Para encontrar recursos de suporte on-line na Rede da Informatica, visite <https://network.informatica.com> e selecione a opção eSupport.

CAPÍTULO 1

Visão geral de Ajuste de desempenho

Este capítulo inclui os seguintes tópicos:

- [Visão geral de Ajuste de desempenho, 11](#)

Visão geral de Ajuste de desempenho

O objetivo do ajuste do desempenho é otimizar o desempenho da sessão eliminando os afunilamentos de desempenho. Para ajustar o desempenho da sessão, primeiro identifique o afunilamento do desempenho, elimine-o e identifique o próximo afunilamento de desempenho até que esteja satisfeito com o desempenho da sessão. Você pode usar a opção de carregamento de teste para executar sessões quando ajustar o desempenho das sessões.

Se você ajustar todos os afunilamentos, será possível otimizar o desempenho da sessão aumentando o número de partições de pipeline na sessão. A adição de partições pode aprimorar o desempenho por meio da utilização maior de hardware do sistema ao processar a sessão.

Como determinar a melhor maneira de melhorar o desempenho pode ser complexo, altere uma variável de cada vez, e limite a sessão tanto antes quanto depois da alteração. Se o desempenho da sessão não melhorar, convém retornar à configuração original.

Execute as seguintes tarefas para melhorar o desempenho da sessão:

1. **Otimize o destino.** Habilita o Serviço de Integração para gravar nos destinos eficientemente.
2. **Otimize a origem.** Habilita o Serviço de Integração para ler os dados da origem eficientemente.
3. **Otimize o mapeamento.** Habilita o Serviço de Integração para transformar e mover os dados eficientemente.
4. **Otimize a transformação.** Habilita o Serviço de Integração para processar as transformações em um mapeamento eficientemente.
5. **Otimize a sessão.** Habilita o Serviço de Integração para executar a sessão mais rapidamente.
6. **Otimize as implantações de grade.** Habilita o Serviço de Integração para executar em uma grade com o desempenho ideal.
7. **Otimize os componentes do PowerCenter.** Habilita o Serviço de Integração e o Serviço de Repositório para funcionarem da forma ideal.
8. **Otimize o sistema.** Habilita os processos do serviço do PowerCenter para execução mais rápida.

CAPÍTULO 2

Afunilamentos

Este capítulo inclui os seguintes tópicos:

- [Visão geral do capítulo Afunilamentos, 12](#)
- [Usando estatísticas de segmento, 13](#)
- [Afunilamentos de destino, 14](#)
- [Afunilamentos de origem, 15](#)
- [Afunilamentos de mapeamento, 16](#)
- [Afunilamentos de sessão, 17](#)
- [Afunilamentos do sistema, 17](#)

Visão geral do capítulo Afunilamentos

A primeira etapa no ajuste do desempenho é identificar afunilamentos de desempenho. Os afunilamentos de desempenho podem ocorrer nos bancos de dados de origem e destino, no mapeamento, na sessão e no sistema. A estratégia é identificar um afunilamento de desempenho, eliminá-lo e, em seguida identificar o próximo afunilamento de desempenho até que você fique satisfeito com o desempenho.

Pesquise os afunilamentos de desempenho na seguinte ordem:

1. Destino
2. Origem
3. Mapeamento
4. Sessão
5. Sistema

Use os seguintes métodos para identificar os afunilamentos de desempenho:

- **Execute sessões de teste.** Você pode configurar uma sessão de teste para ler dados de uma origem de arquivo simples ou gravar em um destino de arquivo simples para identificar afunilamentos de origem e destino.
- **Análise os detalhes do desempenho.** Analise os detalhes do desempenho, como contadores de desempenho, para determinar onde o desempenho da sessão diminui.
- **Análise as estatísticas do segmento.** Analise as estatísticas do segmento para determinar o número ideal de pontos de partição.
- **Monitore o desempenho do sistema.** Você pode usar as ferramentas de monitoramento do sistema para visualizar a porcentagem de uso da CPU, esperas de E/S e paginação para identificar os afunilamentos do sistema. É possível também usar o Workflow Monitor para visualizar o uso do recursos do sistema.

Usando estatísticas de segmento

Você pode usar as estatísticas do segmento no log da sessão para identificar afunilamentos de origem, de destino e de transformação. Por padrão, o Serviço de Integração usa um segmento de leitor, um de transformação e um de gravador para processar a sessão. O segmento com a porcentagem de ocupação mais alta identifica o afunilamento na sessão.

O log da sessão fornece as seguintes estatísticas do segmento:

- **Tempo de execução.** Quantidade de tempo em que o segmento é executado.
- **Tempo inativo.** Quantidade de tempo que o segmento fica ocioso. Ele inclui o tempo que o segmento aguarda o processamento de outro segmento no aplicativo. O tempo ocioso inclui o tempo em que o segmento está bloqueado pelo Serviço de Integração, mas não o tempo em que ele está bloqueado pelo sistema operacional.
- **Tempo de ocupação.** Porcentagem do tempo de execução que o segmento fica ocupado de acordo com a seguinte fórmula:

```
(run time - idle time) / run time X 100
```

É possível ignorar as altas porcentagens de ocupação quando o tempo de execução total é baixo, como menos de 60 segundos. Isso não indica necessariamente um afunilamento.

- **Tempo de trabalho do segmento.** A porcentagem de tempo que o Serviço de Integração leva para processar cada transformação em um segmento. O log da sessão exibe as seguintes informações do tempo de trabalho do segmento de transformação:

```
Thread work time breakdown:  
  <transformation name>: <number> percent  
  <transformation name>: <number> percent  
  <transformation name>: <number> percent
```

Se a transformação levar uma quantidade de tempo pequena, o log da sessão não incluirá isso. Se o segmento não tiver estatísticas precisas porque a sessão é executada por um período curto de tempo, o log da sessão reportará que as estatísticas não são precisas.

Eliminando afunilamentos com base em estatísticas de segmento

Conclua as seguintes tarefas para eliminar afunilamentos com base nas estatísticas do segmento:

- Se o segmento do leitor ou do gravador está 100% ocupado, tente usar tipos de dados de string nas portas de origem e destino. As portas que não são de string exigem mais processamento.
- Se um segmento de transformação está 100% ocupado, tente adicionar um ponto de partição no segmento. Ao adicionar pontos de partição ao mapeamento, o Serviço de Integração aumenta o número de segmentos de transformação que ele usa na sessão. No entanto, se a máquina já está em execução próxima ou na capacidade total, não adicione mais segmentos.
- Se uma transformação requer mais tempo de processamento que as outras, tente adicionar um ponto de partição de passagem para a transformação.

Exemplo

Quando você executa uma sessão, o log da sessão lista as informações de execução e as estatísticas do segmento da seguinte forma:

```
***** RUN INFO FOR TGT LOAD ORDER GROUP [1], CONCURRENT SET [1] *****  
Thread [READER_1_1_1] created for [the read stage] of partition point  
[SQ_two_gig_file_32B_rows] has completed.  
Total Run Time = [505.871140] secs  
Total Idle Time = [457.038313] secs  
Busy Percentage = [9.653215]
```

```

Thread [TRANSF_1_1_1] created for [the transformation stage] of partition point
[SQ_two_gig_file_32B_rows] has completed.
    Total Run Time = [506.230461] secs
    Total Idle Time = [1.390318] secs
    Busy Percentage = [99.725359]
    Thread work time breakdown:
        LKP_ADDRESS: 25.000000 percent
        SRT_ADDRESS: 21.551724 percent
        RTR_ZIP_CODE: 53.448276 percent
Thread [WRITER_1_*_1] created for [the write stage] of partition point [scratch_out_32B]
has completed.
    Total Run Time = [507.027212] secs
    Total Idle Time = [384.632435] secs
    Busy Percentage = [24.139686]

```

Nesse log da sessão, o tempo de execução total do segmento de transformação é de 506 segundos e a porcentagem de ocupação é de 99.7%. Isso significa que o segmento de transformação nunca esteve ocioso por 506 segundos. As porcentagens de ocupação do leitor e do gravador foram bastante menores, aproximadamente 9.6% e 24%. Nesta sessão, o segmento de transformação é o afinilamento no mapeamento.

Para determinar qual transformação no segmento de transformação é o afinilamento, visualize a porcentagem de ocupação de cada transformação na divisão de tempo de trabalho do segmento. Neste log da sessão, a transformação RTR_ZIP_CODE teve uma porcentagem de ocupação de 53%.

Afunilamentos de destino

O afinilamento de desempenho mais comum ocorre quando o Serviço de Integração grava em um banco de dados de destino. Os intervalos de ponto de verificação curtos, tamanhos pequenos do pacote de rede de banco de dados ou problemas durante operações de carregamentos pesadas podem causar afinilamentos de destino.

Identificando afinilamentos de destino

Para identificar um afinilamento de destino, conclua as seguintes tarefas:

- Configure uma cópia da sessão para gravar dados em um destino de arquivo simples. Se o desempenho da sessão aumentar significativamente, você terá um afinilamento de destino. Se a sessão já grava dados em um destino de arquivo simples, provavelmente você não tem um afinilamento de destino.
- Leia as estatísticas do segmento no log da sessão. Quando o Serviço de Integração gasta mais tempo no segmento do gravador do que nos segmentos do leitor ou de transformação, você tem um afinilamento de destino.

Eliminando afinilamentos no destino

Conclua as seguintes tarefas para eliminar os afinilamentos de destino:

- Permita que o administrador do banco de dados aprimore o desempenho otimizando a consulta.
- Aumente o tamanho do pacote de rede do banco de dados.
- Configure restrições de índice e chave.

TÓPICOS RELACIONADOS:

- [“Otimizando o destino” na página 20](#)

Afunilamentos de origem

Os afunilamentos de desempenho ocorrem quando o Serviço de Integração lê os dados do banco de dados de origem. A consulta ineficiente ou os tamanhos do pacote de rede do banco de dados podem causar afunilamentos de origem.

Identificando afunilamentos de origem

Você pode ler as estatísticas do segmento no log da sessão para determinar se a origem é o afunilamento. Quando o Serviço de Integração gasta mais tempo no segmento do leitor do que nos segmentos do gravador ou de transformação, você tem um afunilamento de origem.

Se a sessão lê dados de uma origem relacional, use os seguintes métodos para identificar afunilamentos de origem:

- Transformação de Filtro
- Mapeamento de teste de leitura
- Consulta de banco de dados

Se a sessão lê dados de uma origem de arquivo simples, provavelmente você não tem um afunilamento de origem.

Usando uma transformação de filtro

Você pode usar uma transformação de Filtro no mapeamento para medir o tempo que ela leva para ler os dados de origem.

Adicione uma transformação de Filtro após cada qualificador de origem. Defina a condição do filtro como falso para que nenhum dado seja processado após a transformação do Filtro. Se o tempo decorrido para executar a nova sessão continuar praticamente o mesmo, você tem um afunilamento de origem.

Usando um mapeamento de teste de leitura

Você pode criar um mapeamento de teste de leitura para identificar afunilamentos de origem. Um mapeamento de teste de leitura isola a consulta de leitura removendo a transformação no mapeamento.

Para criar um mapeamento de teste de leitura, conclua as seguintes etapas:

1. Faça uma cópia do mapeamento original.
2. No mapeamento copiado, mantenha somente as origens, os qualificadores de origem e quaisquer associações ou consultas.
3. Remova todas as transformações.
4. Conecte os qualificadores de origem a um destino de arquivo.

Execute uma sessão no mapeamento de teste de leitura. Se o desempenho do sessão é similar ao original, você tem um afunilamento de origem.

Usando uma consulta de banco de dados

Para identificar afunilamentos de origem, execute a pesquisa de leitura diretamente no banco de dados de origem.

Copie a consulta de leitura diretamente do log da sessão. Execute a consulta no banco de dados de origem com uma ferramenta, como o isql. No Windows, é possível carregar o resultado da consulta em um arquivo. No UNIX, é possível carregar o resultado da consulta em /dev/null.

Meça o tempo de execução da consulta e o tempo que ela leva para retornar a primeira linha.

Eliminando afunilamentos na origem

Conclua as seguintes tarefas para eliminar os afunilamentos de origem:

- Defina o número de bytes que o Serviço de Integração lê por linha se ele ler de uma origem de arquivo simples.
- Permita que o administrador do banco de dados aprimore o desempenho otimizando a consulta.
- Aumente o tamanho do pacote de rede do banco de dados.
- Configure restrições de índice e chave.
- Se houver um longo atraso entre as duas medidas de tempo em uma consulta do banco de dados, você poderá usar uma dica do otimizador.

TÓPICOS RELACIONADOS:

- [“Otimizando a origem” na página 24](#)

Afunilamentos de mapeamento

Se você determinar que não tem um afunilamento de origem ou de destino, será possível que tenha um afunilamento de mapeamento.

Identificando afunilamentos de mapeamento

Para identificar afunilamentos de mapeamento, conclua as seguintes tarefas:

- Leia as estatísticas do segmento e as do tempo de trabalho no log da sessão. Quando o Serviço de Integração gasta mais tempo no segmento de transformação do que nos segmentos do leitor ou o gravador, você tem um afunilamento de transformação. Quando o Serviço de Integração gasta mais tempo em uma transformação, existe um afunilamento no segmento de transformação.
- Analise os contadores de desempenho. Os contadores errorrows e rowsinlookupcache altos indicam um afunilamento de mapeamento.
- Adicione uma transformação de Filtro antes de cada definição de destino. Defina a condição do filtro como falso para que nenhum dado seja carregado nas tabelas de destino. Se o tempo decorrido para executar a nova sessão continuar igual ao original, você tem um afunilamento de mapeamento.

Eliminando afunilamentos de mapeamento

Para eliminar afunilamentos de mapeamento, otimize as configurações de transformação nos mapeamentos.

TÓPICOS RELACIONADOS:

- [“Otimizando mapeamentos” na página 27](#)

Afunilamentos de sessão

Se você não tiver um afunilamento de origem, de destino ou de mapeamento, será possível que tenha um afunilamento de sessão. O tamanho pequeno de cache, a memória baixa de buffer e os pequenos intervalos de confirmação podem causar afunilamentos de sessão.

Identificando afunilamentos de sessão

Para identificar um afunilamento de sessão, analise os detalhes do desempenho. Os detalhes do desempenho exibem informações sobre cada transformação, como o número de linhas de entrada, linhas de saída e linhas de erro.

Eliminando afunilamentos de sessão

Para eliminar afunilamentos de sessão, otimize a sessão.

TÓPICOS RELACIONADOS:

- [“Otimizando sessões” na página 43](#)

Afunilamentos do sistema

Após ajustar a origem, o destino, o mapeamento e a sessão, tente ajustar o sistema para evitar afunilamentos do sistema. O Serviço de Integração usa os recursos do sistema para processar transformações, executar sessões, ler e gravar dados. O Serviço de Integração também usa a memória do sistema para criar arquivos de cache files para transformações, como Agregador, Associador, Pesquisa, Classificador, XML e Classificação.

Identificando afunilamentos do sistema

Você pode visualizar o uso dos recursos do sistema no Workflow Monitor. É possível usar as ferramentas do sistema para monitorar os sistemas Windows e UNIX .

Usando o Workflow Monitor para identificar afunilamentos do sistema

Você pode visualizar as propriedades do Serviço de Integração no Workflow Monitor o uso da CPU, da memória e de permuta do sistema quando estiver executando processos de tarefa no Serviço de Integração. Use as seguintes propriedades do Serviço de Integração para identificar problemas de desempenho:

- **CPU%.** A porcentagem do uso da CPU inclui outras tarefas externas em execução no sistema.
- **Uso da memória.** A porcentagem do uso da memória inclui outras tarefas externas em execução no sistema. Se o uso da memória está próximo de 95%, verifique se as tarefas em execução no sistema estão usando a quantidade indicada no Workflow Monitor ou se há uma perda de memória. Para

solucionar o problema, use as ferramentas do sistema para verificar o uso da memória antes e depois da execução da tarefa e, em seguida, compare os resultados do uso da memória ao executar a sessão.

- **Uso de permuta.** O uso de permuta é um resultado da paginação devido as possíveis perdas de memória ou um número alto de tarefas simultâneas.

Identificando afunilamentos do sistema no Windows

Você pode visualizar a guia Desempenho e processos no Gerenciador de Tarefas para obter informações do sistema. A guia Desempenho no Gerenciador de Tarefas oferece uma visão geral do uso da CPU e do total de memória usada. Use o Monitor de Desempenho para visualizar mais informações detalhadas.

A tabela a seguir descreve as informações de sistema que você pode usar no Monitor de Desempenho do Windows para criar um gráfico:

Propriedade	Descrição
Tempo do processador em percentual	Se você tiver mais de uma CPU, monitore cada uma delas para tempo do processador em percentual.
Páginas/segundo	Se as páginas/segundo for maior do que cinco, talvez você tenha pressão de memória excessiva conhecido como sobrecarga.
Tempo em percentual dos discos físicos	O percentual de tempo que o disco físico está ocupado executando solicitações de leitura e gravação.
Comprimento da fila dos discos físicos	O número de usuários aguardando para acessar o mesmo dispositivo de disco.
Total de bytes do servidor por segundo	O servidor enviou e recebeu da rede.

Identificando afunilamentos do sistema no UNIX

Use as seguintes ferramentas para identificar afunilamentos do sistema no UNIX:

- **superior.** Visualize o desempenho geral do sistema. Esta ferramenta exibe o uso da CPU, da memória e de permuta do sistema e de processos individuais em execução no sistema.
- **iostat.** Monitore a operação de carregamento para cada disco anexado ao servidor do banco de dados. O iostat exibe a porcentagem de tempo que o disco está fisicamente ativo. Se você utiliza matrizes de disco, use os utilitários fornecidos com as matrizes em vez do iostat.
- **vmstat.** Monitore as ações de permuta do disco.
- **sar.** Visualize relatórios detalhados das atividades de sistema da CPU, da memória e do uso do disco. Você pode usar esta ferramenta para monitorar o carregamento da CPU. Ela oferece o uso em percentual com base no tempo de espera, no tempo ocioso, no sistema e no usuário. Além disso, use esta ferramenta para monitorar as ações de permuta do disco.

Eliminando afunilamentos de sistema

Conclua as seguintes tarefas para eliminar os afunilamentos do sistema:

- Se o uso da CPU é maior que 80%, verifique o número de tarefas simultâneas em execução. Tente alterar o carregamento ou usar uma grade para distribuir tarefas para diferentes nós. Se não for possível diminuir a carga, tente adicionar mais processadores.

- Se a permuta ocorrer, aumente a memória física ou reduza o número de aplicativos com utilização intensa da memória no disco.
- Se você tiver pressão de memória excessiva (sobrecarga), tente adicionar mais memória física.
- Se a porcentagem de tempo é alta, ajuste o cache para que o PowerCenter use o cache de memória interna em vez de gravar no disco. Se você ajustar o cache, as solicitações continuarem na fila e a porcentagem de ocupação do disco for de pelo menos 50%, adicione outro dispositivo de disco ou faça um upgrade para um dispositivo de disco mais rápido. Você também pode usar um disco separado para cada partição na sessão.
- Se o comprimento da fila do disco físico é maior que dois, tente adicionar outro dispositivo de disco ou fazer um upgrade no dispositivo de disco. Além disso, é possível usar discos separados para os segmentos do leitor, do gravador e de transformação.
- Tente aprimorar a largura de banda da rede.
- Ao ajustar os sistemas UNIX, ajuste o servidor para um sistema principal de banco de dados.
- Se o tempo em percentual gasto aguardando no E/S (%wio) for alto, tente usar outros discos pouco utilizados. Por exemplo, se os arquivos de cache dos dados de origem, de destino, de pesquisa, de classificação e agregado estão no mesmo disco, tente colocá-los em discos diferentes.

TÓPICOS RELACIONADOS:

- [“Reduzindo a paginação” na página 63](#)
- [“Otimizando o sistema” na página 62](#)

CAPÍTULO 3

Otimizando o destino

Este capítulo inclui os seguintes tópicos:

- [Otimizando destinos de arquivo simples, 20](#)
- [Descartando restrições de índices e chave, 20](#)
- [Aumentando os intervalos do ponto de verificação do banco de dados, 21](#)
- [Usando cargas em massa, 21](#)
- [Usando carregadores externos, 21](#)
- [Minimizando deadlocks, 22](#)
- [Aumentando o tamanho do pacote de rede do banco de dados, 22](#)
- [Otimizando os bancos de dados de destino Oracle, 22](#)

Otimizando destinos de arquivo simples

Se você usa um diretório de armazenamento compartilhado para destinos de arquivo simples, é possível otimizar o desempenho da sessão garantindo que o diretório de armazenamento compartilhado esteja em uma máquina dedicada para o armazenamento e gerenciamento de arquivos, em vez de executar outras tarefas.

Se o Serviço de Integração é executado em um único nó e a sessão grava em um destino de arquivo simples, você pode otimizar o desempenho da sessão gravando em um destino de arquivo simples que é local para o nó do processo do Serviço de Integração.

Descartando restrições de índices e chave

Ao definir restrições de chave ou índices em tabelas de destino, o carregamento de dados se torna lento nessas tabelas. Para aprimorar o desempenho, descarte as restrições de chave e índices antes de executar a sessão. É possível criar novamente essas restrições de chave e índices depois que a sessão for concluída.

Se você deseja descartar e recriar as restrições de chave e índices regularmente, use os seguintes métodos para fazer estas operações sempre que executar a sessão:

- Use procedimentos armazenados pré e pós-carregados.
- Use comandos SQL pré e pós-sessão.

Nota: Para otimizar o desempenho, use o carregamento com base na restrição somente se necessário.

Aumentando os intervalos do ponto de verificação do banco de dados

O desempenho do Serviço de Integração fica lento sempre que aguarda enquanto o banco de dados execute um ponto de verificação. Para diminuir o número de pontos de verificação e aprimorar o desempenho, aumente o intervalo de ponto de verificação no banco de dados.

Nota: Embora você obtenha desempenho ao reduzir o número de pontos de verificação, você também aumente o tempo de recuperação se o banco de dados for encerrado inesperadamente.

Usando cargas em massa

É possível usar o carregamento em massa para aprimorar o desempenho de uma sessão que insere uma grande quantidade de dados em um banco de dados DB2, Sybase ASE, Oracle ou Microsoft SQL Server. Configure o carregamento em massa nas propriedades da sessão.

Ao executar o carregamento em massa, o Serviço de Integração ignora o log do banco de dados, o que aumenta o desempenho. Entretanto, sem a gravação no log do banco de dados, o banco de dados de destino não pode realizar a reversão. Como resultado, você talvez não possa realizar uma recuperação. Ao usar o carregamento em massa, avalie a importância do desempenho de sessão aprimorado em relação à capacidade de recuperar uma sessão incompleta.

Quando você carregar em massa nos destinos do Microsoft SQL Server ou Oracle, defina um grande intervalo de confirmação para aprimorar o desempenho. O Microsoft SQL Server e o Oracle começam uma nova transação de carregamento em massa após cada confirmação. Aumentar o intervalo de confirmação reduz o número de transações de carregamento em massa, o que aumenta o desempenho.

TÓPICOS RELACIONADOS:

- [“Confirmação baseada no destino” na página 47](#)

Usando carregadores externos

Para aprimorar o desempenho da sessão, configure o PowerCenter para usar um carregador externo para os seguintes tipos de bancos de dados de destino:

- IBM DB2 EE ou EEE
- Oracle

Quando você carrega dados para um banco de dados Oracle usando um pipeline com diversas partições, é possível aprimorar o desempenho ao criar uma tabela de destino Oracle com o mesmo número de partições usado no pipeline.

- Sybase IQ

Se o banco de dados Sybase IQ é local ao processo do Serviço de Integração no sistema UNIX, você pode aprimorar o desempenho carregando os dados para tabelas de destino diretamente dos pipes nomeados. Se você executar o Serviço de Integração em uma grade, configure o Balanceador de Carga para verificar os recursos, tornar o Sybase IQ um recurso e tornar disponíveis os recursos em todos os nós da grade. Em seguida, no Workflow Manager, atribua o recurso Sybase IQ às sessões aplicáveis.

- Teradata

Minimizando deadlocks

Se o Serviço de Integração encontrar um deadlock quando tentar gravar em um destino, o deadlock somente afetará os destinos no mesmo grupo de conexão de destino. O Serviço de Integração ainda grava em destinos em outros grupos de conexão de destino.

A descoberta de deadlocks pode retardar o desempenho da sessão. Para melhorar o desempenho da sessão, você pode aumentar o número de grupos de conexão de destino que o Serviço de Integração usa para gravar os destinos em uma sessão. Para usar um grupo de conexão de destino diferente para cada destino em uma sessão, use um nome da conexão de banco de dados diferente para cada instância de destino. Você pode especificar as mesmas informações de conexão para cada nome da conexão.

Aumentando o tamanho do pacote de rede do banco de dados

Se você grava dados em destinos Oracle, Sybase ASE ou Microsoft SQL Server, é possível aprimorar o desempenho aumentando o tamanho do pacote de rede. Aumente o tamanho do pacote de rede para permitir que pacotes maiores de dados cruzem a rede de uma vez. Aumente o tamanho do pacote de rede baseado no banco de dados que você está gravando:

- **Oracle.** Aumente o tamanho do pacote de rede do servidor de banco de dados no `listener.ora` e `tnsnames.ora`. Consulte a documentação do banco de dados para obter mais informações sobre o aumento do tamanho do pacote, se necessário.
- **Sybase ASE e Microsoft SQL Server.** Consulte a documentação do banco de dados para obter informações sobre como aumentar o tamanho do pacote.

Para o Sybase ASE ou o Microsoft SQL Server, é necessário alterar o tamanho do pacote no objeto de conexão relacional no Workflow Manager para refletir o tamanho do pacote do servidor de banco de dados.

Otimizando os bancos de dados de destino Oracle

Se o banco de dados de destino é Oracle, é possível otimizar o banco de dados de destino verificando a cláusula de armazenamento, a alocação de espaço e os segmentos de reversão e de desfazer.

Ao gravar em um banco de dados Oracle, verifique a cláusula de armazenamento dos objetos do banco de dados. Certifique-se de que as tabelas estão usando valores iniciais e próximos grandes. Além disso, o banco de dados deve armazenar dados de índice e tabela em espaços de tabela separados, preferencialmente em discos diferentes.

Ao gravar em banco de dados Oracle, ele usa os segmentos de reversão e de desfazer durante os carregamentos. Consulte o administrador do banco de dados Oracle para saber se o banco de dados armazena segmentos de reversão e de desfazer em espaços de tabela adequados, preferencialmente em

discos diferentes. Os segmentos de reversão e de desfazer devem também ter cláusulas de armazenamento apropriadas.

Para otimizar o banco de dados Oracle, ajuste o log de refazer Oracle. O banco de dados Oracle utiliza o log de refazer para registrar as operações de carregamento. Certifique-se de que o tamanho do log de refazer e o tamanho do buffer são ideais. Você pode visualizar as propriedades do log de refazer no arquivo `init.ora`.

Se o Serviço de Integração é executado em um único nó e a instância Oracle é local ao nó do processo do Serviço de Integração, é possível otimizar o desempenho usando o protocolo IPC para se conectar ao banco de dados Oracle. Você pode configurar a conexão do banco de dados Oracle no `listener.ora` e `tnsnames.ora`.

Para obter mais informações sobre a otimização dos bancos de dados Oracle, consulte a documentação do Oracle.

CAPÍTULO 4

Otimizando a origem

Este capítulo inclui os seguintes tópicos:

- [Otimizando a consulta, 24](#)
- [Usando filtros condicionais, 25](#)
- [Aumentando o tamanho do pacote de rede do banco de dados, 25](#)
- [Conectando-se às origens de banco de dados Oracle, 25](#)
- [Usando o Teradata FastExport, 25](#)
- [Usando o tempdb para associar o Sybase ou as tabelas do Microsoft SQL Server, 26](#)

Otimizando a consulta

Se uma sessão associar diversas tabelas de origem em um Qualificador de Origem, você poderá aprimorar o desempenho ao otimizar a consulta com dicas de otimização. Além disso, as instruções de seleção da tabela única com uma cláusula ORDER BY ou GROUP BY podem se beneficiar da otimização, como a adição de índices.

Em geral, o otimizador de banco de dados determina a maneira mais eficiente de processar os dados da origem. No entanto, é possível aprender sobre as tabelas de origem que o otimizador de banco de dados não processa. O administrador de banco de dados pode criar dicas do otimizador para dizer ao banco de dados como executar a consulta para um conjunto particular de tabelas de origem.

A consulta usada pelo Serviço de Integração para ler os dados é exibida no log da sessão. Além disso, você pode encontrar a consulta na transformação do Qualificador de Origem. Permita que o administrador do banco de dados analise a consulta e, em seguida, crie índices e dicas do otimizador para as tabelas de origem.

Use as dicas de otimização se houver um longo atraso entre o momento no qual a consulta começa a ser executada e o momento no qual o PowerCenter recebe a primeira linha de dados. Configure as dicas do otimizador para começar a retornar linhas o mais rápido possível, ao invés de retornar todas de uma vez. Isso permite que o Serviço de Integração processe as linhas paralelas com a execução da consulta.

As consultas que contém as cláusulas ORDER BY ou GROUP BY pode se beneficiar da criação de um índice nas colunas ORDER BY ou GROUP BY. Após a otimização da consulta, use a opção de substituição SQL para aproveitar completamente estas modificações.

Você também pode configurar o banco de dados de origem para executar consultas paralelas para aprimorar o desempenho. Para obter mais informações sobre a configuração de consultas paralelas, consulte a documentação do banco de dados.

Usando filtros condicionais

Às vezes, um filtro de origem simples no banco de dados de origem pode influenciar de maneira negativa o desempenho falta de índices. Você pode usar o filtro condicional do PowerCenter no Qualificador de Origem para aprimorar o desempenho.

A utilização do filtro condicional do PowerCenter para aprimorar o desempenho dependerá da sessão. Por exemplo, se diversas sessões são lidas da mesma origem simultaneamente, o filtro condicional do PowerCenter pode aprimorar o desempenho.

No entanto, algumas sessões poderão ser executadas com mais rapidez se os dados de origem forem filtrados no banco de dados de origem. Você pode testar a sessão com o filtro do banco de dados e o filtro do PowerCenter para determinar qual método aprimora o desempenho.

Aumentando o tamanho do pacote de rede do banco de dados

Se você lê das origens Oracle, Sybase ASE ou Microsoft SQL Servers, é possível aprimorar o desempenho aumentando o tamanho do pacote de rede. Aumente o tamanho do pacote de rede para permitir que pacotes maiores de dados cruzem a rede de uma vez. Aumente o tamanho do pacote de rede com base no banco de dados lido do:

- **Oracle.** Aumente o tamanho do pacote de rede do servidor de banco de dados no listener.ora e tnsnames.ora. Consulte a documentação do banco de dados para obter mais informações sobre o aumento do tamanho do pacote, se necessário.
- **Sybase ASE e Microsoft SQL Server.** Consulte a documentação do banco de dados para obter informações sobre como aumentar o tamanho do pacote.

Para o Sybase ASE ou o Microsoft SQL Server, é necessário alterar o tamanho do pacote no objeto de conexão relacional no Workflow Manager para refletir o tamanho do pacote do servidor de banco de dados.

Conectando-se às origens de banco de dados Oracle

Se estiver executando o Serviço de Integração em um único nó e a instância Oracle for local para o nó de processo do Serviço de Integração, você poderá otimizar o desempenho usando o protocolo IPC para se conectar ao banco de dados Oracle. É possível configurar a conexão de um banco de dados Oracle no listener.ora e no tnsnames.ora.

Usando o Teradata FastExport

O FastExport é um utilitário que usa diversas sessões do Teradata para exportar, rapidamente, grandes quantidades de dados de um banco de dados Teradata. É possível criar uma sessão do PowerCenter que usa o FastExport para ler origens Teradata rapidamente. Para usar o FastExport, crie um mapeamento com um banco de dados de origem Teradata. Na sessão, use o leitor FastExport em vez do leitor Relacional. Use uma conexão FastExport com as tabelas Teradata que você deseja exportar em uma sessão.

Usando o tempdb para associar o Sybase ou as tabelas do Microsoft SQL Server

Ao associar tabelas grandes em um banco de dados Sybase ou Microsoft SQL Server, é possível aprimorar o desempenho criando o tempdb como um banco de dados na memória para alocar memória suficiente. Para obter mais informações, consulte a documentação do Sybase ou do Microsoft SQL Server.

CAPÍTULO 5

Otimizando mapeamentos

Este capítulo inclui os seguintes tópicos:

- [Visão geral de Otimizando mapeamentos, 27](#)
- [Otimizando origens de arquivos simples, 27](#)
- [Configurando a leitura de passagem única, 28](#)
- [Otimizando mapeamentos de passagem, 29](#)
- [Otimizando filtros, 29](#)
- [Otimizando conversões de tipos de dados, 29](#)
- [Otimizando expressões, 30](#)
- [Otimizando procedimentos externos, 32](#)

Visão geral de Otimizando mapeamentos

A otimização em nível de mapeamento pode demorar para ser implementada, mas ela pode aprimorar significativamente o desempenho da sessão. Enfatize a otimização em nível de mapeamento depois de otimizar os destinos e origens.

Em geral, você reduz o número de transformações no mapeamento e exclui os links desnecessários entre as transformações para otimizar o mapeamento. Configure o mapeamento de forma que o mínimo de transformações e expressões faça a maior quantidade de trabalho possível. Exclua os links desnecessários entre as transformações para minimizar a quantidade de dados movidos.

Otimizando origens de arquivos simples

Execute as seguintes etapas para otimizar as origens de arquivos simples:

- Otimize o comprimento do buffer sequencial em linha.
- Otimize as origens de arquivo simples delimitado.
- Otimize origens de arquivos XML e simples.

Otimizando o comprimento do buffer sequencial em linha

Se a sessão for lida a partir de um arquivo de origem, você poderá melhorar o desempenho da sessão configurando o número de bytes que o Serviço de Integração lê por linha. Por padrão, o Serviço de Integração

lê 1024 bytes por linha. Se cada linha no arquivo de origem for menor do que a configuração padrão, será possível reduzir o comprimento de buffer sequencial nas propriedades da sessão.

Otimizando origens de arquivos simples delimitados

Se uma origem for um arquivo simples delimitado, é necessário especificar o caractere delimitador de separação de colunas de dados no arquivo de origem. É necessário também especificar o caractere de escape. O Serviço de Integração lerá o caractere delimitador como um caractere normal se você incluir o caractere de escape antes do caractere delimitador. Você pode aprimorar o desempenho da sessão se o arquivo simples de origem não contiver aspas ou caracteres de escape.

Otimizando origens de arquivos XML e simples

Os arquivos XML são geralmente maiores que os arquivos simples devido às informações de marca. O tamanho do arquivo XML dependerá do nível de marcação do arquivo XML. Quanto mais marcas, maior o arquivo. Consequentemente, o Serviço de Integração pode levar mais tempo para ler e armazenar em cache as origens XML.

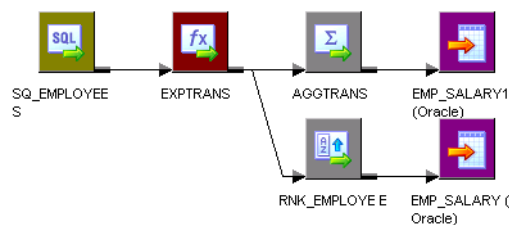
Configurando a leitura de passagem única

A leitura de passagem única permite que você preencha vários destinos com um qualificador de origem. Considere usar a leitura de passagem única se muitas sessões utilizarem as mesmas origens. Você pode combinar a lógica de transformação de cada mapeamento em um mapeamento e usar um qualificador de origem para cada origem. O Serviço de Integração lerá cada origem e enviará os dados em pipelines separados. Uma linha específica pode ser usada por todos os pipelines, por qualquer combinação de pipelines ou por nenhum pipeline.

Por exemplo, você tem a tabela de origem Compras e usa essa origem diariamente para executar uma agregação e uma classificação. Se você inserir as transformações de Agregador e de Classificação em mapeamentos e sessões diferentes, você forçará o Serviço de Integração a ler a mesma tabela de origem duas vezes. No entanto, se você incluir a lógica de agregação e de classificação em um mapeamento com um qualificador de origem, o Serviço de Integração lerá a tabela de origem Compras uma vez e enviará os dados adequados aos pipelines separados.

Quando você altera mapeamentos para aproveitar a leitura de passagem única, é possível otimizar esse recurso fatorando as funções comuns dos mapeamentos. Por exemplo, se você precisar subtrair uma porcentagem das portas de Preço tanto para as transformações de Agregador quanto de Classificação, é possível minimizar o trabalho subtraindo a porcentagem *antes* de dividir o pipeline. Você pode usar uma transformação de Expressão para subtrair a porcentagem e dividir o mapeamento após a transformação.

A figura a seguir mostra a leitura de passagem única, onde o mapeamento é dividido após a transformação de Expressão:



Otimizando mapeamentos de passagem

Você pode otimizar o desempenho dos mapeamentos de passagem. Para passar diretamente da origem para o destino sem outras transformações, conecte a transformação de Qualificador de Origem diretamente ao destino. Se você usar o Assistente do Guia Rápido para criar um mapeamento de passagem, o assistente criará uma transformação de Expressão entre a transformação de Qualificador de Origem e o destino.

Otimizando filtros

Use uma das seguintes transformações para filtrar dados:

- **Transformação Qualificador de Origem.** As linhas de filtros da transformação de Qualificador de Origem das origens relacionais.
- **Transformação de Filtro.** A transformação de Filtro filtra os dados contidos no mapeamento. A transformação de Filtro filtra linhas de qualquer tipo de origem.

Se você filtrar as linhas do mapeamento, é possível aprimorar a eficiência filtrando o início do fluxo de dados. Use um filtro na transformação de Qualificador de Origem para remover as linhas na origem. A transformação de Qualificador de Origem limita o conjunto de linhas extraído de uma origem relacional.

Se você não puder usar um filtro na transformação de Qualificador de Origem, use uma transformação de Filtro e mova-a o mais próximo possível para a transformação de Qualificador de Origem para remover os dados desnecessários no início do fluxo de dados. A transformação de Filtro limita o conjunto de linhas enviado ao destino.

Evite usar expressões complexas em condições de filtro. Para otimizar as transformações de Filtro, use um inteiro simples ou expressões de verdadeiro/falso na condição do filtro.

Nota: Você também pode usar uma transformação de Filtro ou Roteador para descartar linhas rejeitadas de uma transformação de Estratégia de Atualização caso não precise mantê-las.

Otimizando conversões de tipos de dados

Você pode melhorar o desempenho eliminando conversões de tipos de dados desnecessárias. Por exemplo, se um mapeamento mover dados de uma coluna Inteiro para uma coluna Decimal e retornar para uma coluna Inteiro, a conversão de tipos de dados desnecessária reduzirá o desempenho. Onde possível, elimine as conversões de tipos de dados desnecessárias dos mapeamentos.

Use as seguintes conversões de tipos de dados para melhorar o desempenho:

- **Use valores inteiros em vez de outros tipos de dados ao realizar comparações usando as transformações de Pesquisa e de Filtro.** Por exemplo, muitos bancos de dados armazenam informações de CEP dos EUA, como o tipo de dados Char ou Varchar. Se você converter os dados de CEP para um tipo de dados de Inteiro, o banco de dados de pesquisa armazenará o CEP 94303-1234 como 943031234. Isso ajudará a aumentar a velocidade das comparações de pesquisa com base no CEP.
- **Converta as datas de origem por meio de conversões porta-a-porta para melhorar o desempenho da sessão.** Você pode manter as portas em destinos como strings ou alterá-las para portas de Data/Hora.

Otimizando expressões

Você também pode otimizar as expressões usadas nas transformações. Quando possível, isole expressões lentas e simplifique-as.

Execute as seguintes tarefas para isolar as expressões lentas:

1. Remova as expressões uma a uma do mapeamento.
2. Execute o mapeamento para determinar o tempo que ele leva para executar o mapeamento se a transformação.

Se houver uma diferença significativa no tempo de execução da sessão, procure maneiras de otimizar a expressão lenta.

Fatoramento de lógica comum

Se o mapeamento realizar a mesma tarefa em vários locais, reduza o número de vezes que ele realiza a tarefa movendo-a para um local anterior no mapeamento. Por exemplo, você tem um mapeamento com cinco tabelas de destino. Cada destino requer uma pesquisa de número de CPF. Em vez de realizar a pesquisa cinco vezes, insira a transformação de Pesquisa no mapeamento antes das divisões de fluxo de dados. Em seguida, passe os resultados da pesquisa para todos os cinco destinos.

Minimizando chamadas de função de agregação

Quando você escrever expressões, fatore o máximo de chamadas de função de agregação possível. Cada vez que você usar uma chamada de função de agregação, o Serviço de Integração precisará pesquisar e agrupar os dados. Por exemplo, na expressão a seguir, o Serviço de Integração lê COLUMN_A, localiza a soma, lê COLUMN_B, localiza a soma e, por fim, localiza a soma das duas somas:

```
SUM(COLUMN_A) + SUM(COLUMN_B)
```

Se você fatorar a chamada de função de agregação, conforme abaixo, o Serviço de Integração adicionará COLUMN_A à COLUMN_B e, em seguida, localizará a soma de ambos.

```
SUM(COLUMN_A + COLUMN_B)
```

Substituindo expressões comuns por variáveis locais

Se você usar a mesma expressão várias vezes em uma transformação, é possível tornar essa expressão uma variável local. Você pode usar variáveis locais somente dentro da transformação. No entanto, se você calcular a variável somente uma vez, o desempenho melhorará.

Escolhendo operações numéricas versus de string

O Serviço de Integração processa as operações numéricas mais rapidamente do que as operações de string. Por exemplo, se você pesquisar grandes quantidades de dados em duas colunas, EMPLOYEE_NAME e EMPLOYEE_ID, a configuração da pesquisa em EMPLOYEE_ID resultará no aumento do desempenho.

Otimizando comparações Char-Char e Char-Varchar

Quando o Serviço de Integração executa comparações entre as colunas CHAR e VARCHAR, ele fica mais lento toda vez que encontra espaços em branco à direita/esquerda na linha. Você pode usar a opção TreatCHARasCHARonRead quando configurar o Serviço de Integração no Console de Administração para que o Serviço de Integração não ajuste os espaços em branco à direita/esquerda da extremidade dos campos de origem Char.

Escolhendo DECODE versus LOOKUP

Quando você usa a função LOOKUP, o Serviço de Integração precisa pesquisar uma tabela em um banco de dados. Quando você usa uma função DECODE, você incorpora os valores de pesquisa à expressão para que o Serviço de Integração não precise pesquisar outra tabela separada. Sendo assim, quando você quiser pesquisar um conjunto pequeno de valores permanentes inalterados, use DECODE para melhorar o desempenho.

Usando operadores em vez de funções

O Serviço de Integração lê expressões gravadas com operadores mais rapidamente que expressões com funções. Onde possível, use operadores para gravar expressões. Por exemplo, você tem a seguinte função que contém funções CONCAT aninhadas:

```
CONCAT( CONCAT( CUSTOMERS.FIRST_NAME, ' ' ) CUSTOMERS.LAST_NAME)
```

Você pode regravar essa expressão com o operador `||` desta forma:

CUSTOMERS.FIRST_NAME || ' ' || CUSTOMERS.LAST_NAME

Otimizando funções IIF

As funções IIF podem retornar um valor e uma ação, o que permite expressões mais compactas. Por exemplo, você tem uma origem com os sinalizadores S/N: FLG_A, FLG_B, FLG_C. Você deseja retornar os valores com base nos valores de cada sinalizador.

Use a seguinte expressão:

```

IIF( FLG_A = 'Y' and FLG_B = 'Y' AND FLG_C = 'Y',
    VAL_A + VAL_B + VAL_C,
    IIF( FLG_A = 'Y' and FLG_B = 'Y' AND FLG_C = 'N',
        VAL_A + VAL_B ,
        IIF( FLG_A = 'Y' and FLG_B = 'N' AND FLG_C = 'Y',
            VAL_A + VAL_C,
            IIF( FLG_A = 'Y' and FLG_B = 'N' AND FLG_C = 'N',
                VAL_A ,
                IIF( FLG_A = 'N' and FLG_B = 'Y' AND FLG_C = 'Y',
                    VAL_B + VAL_C,
                    IIF( FLG_A = 'N' and FLG_B = 'Y' AND FLG_C = 'N',
                        VAL_B ,
                        IIF( FLG_A = 'N' and FLG_B = 'N' AND FLG_C = 'Y',
                            VAL_C,
                            IIF( FLG_A = 'N' and FLG_B = 'N' AND FLG_C = 'N',
                                0.0,
                                )))))))

```

Essa expressão requer 8 IIFs, 16 ANDs e pelos menos 24 comparações.

Se você aproveitar a função IIF, é possível regravar a expressão como:

```
IIF(FLG_A='Y', VAL_A, 0.0)+ IIF(FLG_B='Y', VAL_B, 0.0)+ IIF(FLG_C='Y', VAL_C, 0.0)
```

Isso resultará em três IIFs, duas comparações, duas adições e uma sessão mais rápida.

Avaliando expressões

Se você não tiver certeza quais expressões estão reduzindo o desempenho, avalie o desempenho da expressão para isolar o problema.

Execute as seguintes etapas para avaliar o desempenho da expressão:

1. Cronografe a sessão com as expressões originais.
2. Copie o mapeamento e substitua metade das expressões complexas por uma constante.
3. Execute e cronografe a sessão editada.
4. Faça outra cópia do mapeamento e substitua a outra metade das expressões complexas por uma constante.
5. Execute e cronografe a sessão editada.

Otimizando procedimentos externos

Convém bloquear os dados de entrada se o procedimento externo precisar alternar a leitura de grupos de entrada. Sem a funcionalidade de bloqueio, você precisaria gravar o código do procedimento para armazenar os dados de entrada no buffer. Você pode bloquear os dados de entrada, em vez de armazená-los no buffer, o que normalmente melhora o desempenho da sessão.

Por exemplo, você precisa criar um procedimento externo com dois grupos de entrada. O procedimento externo lê uma linha do primeiro grupo de entrada e, em seguida, lê uma linha do segundo grupo de entrada. Se você usar o bloqueio, será possível gravar o código do procedimento externo para bloquear o fluxo de dados de um grupo de entrada enquanto ele processa os dados do outro grupo de entrada. Quando você grava o código do procedimento externo para bloquear dados, o desempenho melhora porque o procedimento não precisa copiar os dados de origem em um buffer. No entanto, você poderia gravar o procedimento externo para alocar um buffer e copiar os dados de um grupo de entrada no buffer até que ele esteja pronto para processar os dados. A cópia de dados de origem em um buffer prejudica o desempenho.

CAPÍTULO 6

Otimizando transformações

Este capítulo inclui os seguintes tópicos:

- [Otimizando transformações de Agregador, 33](#)
- [Otimizando transformações personalizadas, 34](#)
- [Otimizando transformações de unificador, 35](#)
- [Otimizando transformações de pesquisa, 35](#)
- [Otimizando Transformações de Normalizador, 39](#)
- [Otimizando transformações de Gerador de Sequência, 39](#)
- [Otimizando transformações de classificador, 39](#)
- [Otimizando transformações de Qualificador de Origem, 41](#)
- [Otimizando transformações de SQL, 41](#)
- [Otimizando Transformações de XML, 41](#)
- [Eliminando erros de transformação, 42](#)

Otimizando transformações de Agregador

As transformações de Agregador sempre reduzem o desempenho porque elas precisam agrupar os dados para processá-los. As transformações precisam de memória adicional para manter os resultados de grupos intermediários.

Use as diretrizes a seguir para otimizar o desempenho de uma transformação de Agregador:

- Agrupe por colunas simples.
- Use entradas classificadas.
- Use agregação incremental.
- Filtre os dados antes de agregá-los.
- Limite as conexões de porta.

Agrupando por colunas simples

Você pode otimizar as transformações de Agregador quando agrupar por colunas simples. Quando possível, use número em vez de string e datas nas colunas usadas para GROUP BY. Evite expressões complexas nas expressões de Agregador.

Usando entrada classificada

Para melhorar o desempenho da sessão, classifique os dados da transformação de Agregador. Use a opção de Entrada Classificada para classificar dados.

A opção Entrada Classificada diminui o uso de caches agregados. Quando você usa a opção Entrada Classificada, o Serviço de Integração considera que todos os dados estão filtrados por grupo. Ao passo que o Serviço de Integração lê as linhas de um grupo, ele executa os cálculos de agregação. Quando necessário, ele armazena as informações do grupo na memória.

A opção Entrada Classificada reduz a quantidade de dados armazenados no cache durante a sessão e melhora o desempenho. Use essa opção com a opção Número de Portas Classificadas do Qualificador de Origem ou uma transformação Classificador para transferir os dados classificados para a transformação de Agregador.

Você pode melhorar o desempenho ao usar a opção Entrada Classificada em sessões com várias partições.

Usando agregação incremental

Se você puder capturar as alterações da origem que afetem menos da metade do destino, será possível usar a agregação incremental para otimizar o desempenho das transformações de Agregador.

Quando você usa a agregação incremental, você aplica as alterações capturadas na origem para agregar cálculos em uma sessão. O Serviço de Integração atualiza o destino incrementalmente em vez de processar a origem inteira e fazer novamente os mesmos cálculos todas as vezes que você executar a sessão.

Você pode aumentar o tamanho do cache do índice e dos dados para armazenamento de todos os dados na memória sem paginação no disco.

TÓPICOS RELACIONADOS:

- [“Aumentando os tamanhos de cache” na página 47](#)

Filtrando dados antes de agregar

Filtre os dados antes de agregá-los. Se você utilizar uma transformação de Filtro no mapeamento, insira-a antes da transformação de Agregador, a fim de reduzir uma agregação desnecessária.

Limitando conexões de porta

Limite o número de portas de entrada/saída ou de saída conectadas para reduzir o volume de dados que a transformação de Agregador armazena no cache de dados.

Otimizando transformações personalizadas

O Serviço de Integração pode transferir uma única linha ao procedimento de transformação personalizada ou um bloco de linhas em uma matriz. Você pode gravar o código do procedimento de forma a especificar se o procedimento receberá uma linha ou um bloco de linhas.

Você pode melhorar o desempenho quando o procedimento receber blocos de linhas:

- Você pode diminuir o número de chamadas de função feitas pelo Serviço de Integração e pelo procedimento. O Serviço de Integração chama a função de notificação de linha de entrada menos vezes e o procedimento chama a função de notificação de saída menos vezes também.
- Você pode aumentar a localidade do espaço de acesso à memória para os dados.
- Você pode gravar o código do procedimento para executar algoritmos em blocos de dados, e não em cada linha de dados.

Otimizando transformações de unificador

As transformações de unificador podem reduzir o desempenho porque elas precisam de espaço adicional no tempo de execução para manter resultados intermediários. Você pode exibir as informações de contador de desempenho de Unificador para determinar se é necessário otimizar as transformações de Unificador.

Use as dicas a seguir para aprimorar o desempenho da sessão com a transformação de Unificador:

- **Identifique a origem mestre como a origem com menos valores de chave duplicados.** Quando o Serviço de Integração processa uma transformação de Unificador classificada, ele armazena as linhas no cache, cem chaves exclusivas por vez. Se a origem mestre contiver muitas linhas com o mesmo valor de chave, o Serviço de Integração deverá armazenar mais linhas no cache. Com isso, o desempenho ficará mais lento.
- **Identifique a origem mestre como a origem com menos linhas.** Durante a sessão, a transformação de Unificador compara cada linha da origem de detalhes a origem mestre. Quanto menos linhas na origem mestre, menor o número de iterações da comparação das associações que ocorrerão. Isso acelera o processo de associação.
- **Quando possível, faça associações em um banco de dados.** É mais rápido fazer uma associação em um banco de dados do que na sessão. O tipo de associação de banco de dados usado pode afetar o desempenho. As associações normais são mais rápidas que as associações externas e geram menos linhas. Em alguns casos, não é possível fazer associações no banco de dados, como unir tabelas de dois bancos de dados diferentes ou de sistemas de arquivos simples.

Para realizar uma associação em um banco de dados, utilize as seguintes opções:

- Crie um procedimento armazenado de pré-sessão para unir as tabelas em um banco de dados.
- Use a transformação de Qualificador de Origem para fazer a associação.
- **Quando possível, associe dados classificados.** Para melhorar o desempenho da sessão, configure a transformação de Unificador para usar a entrada classificada. Quando você configura a transformação de Unificador para utilizar dados classificados, o Serviço de Integração melhora o desempenho, minimizando a entrada e a saída de disco. Você percebe um melhor aprimoramento do desempenho quando trabalha com grandes conjuntos de dados. No caso de uma transformação de Unificação não classificada, indique a origem com um número menor de linhas que a origem mestre.

Otimizando transformações de pesquisa

Se a tabela de consulta estiver no mesmo banco de dados que a tabela de origem no seu mapeamento e o cache não for viável, una as tabelas no banco de dados de origem em vez de usar uma transformação Pesquisa.

Se você usar uma transformação de Pesquisa, execute as seguintes tarefas para melhorar o desempenho:

- Use o driver de banco de dados ideal.
- Armazene as tabelas de pesquisa no cache.
- Otimize a condição de pesquisa.
- Filtre as linhas de pesquisa.
- Indexe a tabela de pesquisa.
- Otimize várias pesquisas.
- Crie uma transformação de Pesquisa de pipeline e configure partições no pipeline que criem a origem de pesquisa.

Usando drivers de banco de dados ideais

O Serviço de Integração consegue se conectar a uma tabela de pesquisa com um driver de banco de dados original ou um driver ODBC. Os drivers de banco de dados originais oferecem um melhor desempenho de sessão que os drivers ODBC.

Armazenando tabelas de pesquisa em cache

Se um mapeamento contiver transformações de Pesquisa, convém habilitar o armazenamento em cache da pesquisa. Quando você habilita o armazenamento em cache, o Serviço de Integração armazena em cache a tabela de pesquisa e consulta o cache de pesquisa durante a sessão. Quando a opção não é habilitada, o Serviço de Integração consulta a tabela de pesquisa linha por linha.

O resultado do processamento e da consulta de Pesquisa é o mesmo, independentemente se você armazenar a tabela de pesquisa no cache ou não. No entanto, o uso de um cache de pesquisa pode aumentar o desempenho da sessão em tabelas de pesquisa menores. Em geral, convém armazenar em cache as tabelas de pesquisa que precisam de menos de 300 MB.

Execute as seguintes tarefas para melhorar ainda mais o desempenho das transformações de Pesquisa:

- Use o tipo de cache adequado.
- Habilite caches simultâneos.
- Otimize a correspondência de condição de pesquisa.
- Reduza o número de linhas armazenadas em cache.
- Substitua a instrução ORDER BY.
- Use uma máquina com mais memória.

TÓPICOS RELACIONADOS:

- [“Caches” na página 46](#)

Tipos de caches

Use os seguintes tipos de caches para melhorar o desempenho:

- **Cache compartilhado.** Você pode compartilhar o cache de pesquisa entre várias transformações. Você pode compartilhar um cache não nomeado entre transformações no mesmo mapeamento. Você pode compartilhar um cache nomeado entre transformações no mesmo mapeamento ou em mapeamentos diferentes.

- **Cache persistente.** Para salvar e reutilizar os arquivos de cache, configure a transformação para usar um cache persistente. Use esse recurso quando souber que a tabela de pesquisa não se altera entre as execuções de sessão. O uso de um cache persistente melhora o desempenho porque o Serviço de Integração cria o cache de memória a partir dos arquivos de cache, e não a partir do banco de dados.

Habilitando caches simultâneos

Quando o Serviço de Integração processa sessões que contêm transformações de pesquisa, o Serviço de Integração cria um cache na memória ao processar a primeira linha de dados em uma transformação de pesquisa armazenada em cache. Se houver várias transformações de Pesquisa no mapeamento, o Serviço de Integração criará os caches sequencialmente quando a primeira linha de dados for processada pela transformação de Pesquisa. Isso deixará o processamento de transformação de Pesquisa mais lento.

Você pode habilitar caches simultâneos para melhorar o desempenho. Quando o número de pipelines simultâneos adicionais é definido como um ou mais, o Serviço de Integração cria caches simultaneamente, e não sequencialmente. O desempenho melhora significativamente quando as sessões contêm um número de transformações ativas que podem demorar para ser concluídas, como transformações de Agregador, Unificador ou Classificador. Quando você habilita vários pipelines simultâneos, o Serviço de Integração não aguarda mais as sessões ativas serem concluídas para criar o cache. Outras transformações de Pesquisa no pipeline também criam caches simultaneamente.

Otimizando a correspondência de condição de pesquisa

Quando a transformação de Pesquisa corresponde aos dados de cache de pesquisa com a condição de pesquisa, ela classifica e organiza os dados para determinar o primeiro e o último valor de correspondência. É possível configurar a transformação para retornar todos os valores que correspondam à condição da pesquisa. Ao configurar a transformação de Pesquisa para retornar qualquer valor de correspondência, a transformação retorna o primeiro valor que combine com a condição de pesquisa. Ele não indexa todas as portas como faz quando você configura a transformação para retornar o primeiro ou o último valor de correspondência. Quando você usa um valor de correspondência, o desempenho pode ser melhorado porque a transformação não é indexada em todas as portas, o que pode reduzir o desempenho.

Reduzindo o número de linhas armazenadas em cache

Você pode reduzir o tamanho de linhas incluídas no cache para melhorar o desempenho. Use a opção Substituição SQL de Pesquisa para adicionar uma cláusula WHERE à instrução padrão SQL.

Substituindo a instrução ORDER BY

Por padrão, o Serviço de Integração gera uma instrução ORDER BY para pesquisas armazenadas em cache. A instrução ORDER BY contém todas as portas de pesquisa. Para aumentar o desempenho, remova a instrução padrão ORDER BY e insira uma ORDER BY de substituição com menos colunas.

O Serviço de Integração sempre gera uma instrução ORDER BY, mesmo se você inserir outra na substituição. Coloque dois hífenes '--' depois da substituição ORDER BY para remover a instrução ORDER BY gerada.

Por exemplo, uma transformação de Pesquisa usa a seguinte condição:

```
ITEM_ID = IN_ITEM_ID  
PRICE <= IN_PRICE
```

A transformação de Pesquisa inclui três portas de pesquisa usadas no mapeamento, ITEM_ID, ITEM_NAME e PRICE. Quando você inserir a instrução ORDER BY, insira as colunas na mesma ordem que as portas na condição de pesquisa. Você também deve incluir todas as palavras reservadas do banco de dados entre aspas.

Insira a seguinte consulta de pesquisa na substituição SQL da pesquisa:

```
SELECT ITEMS_DIM.ITEM_NAME, ITEMS_DIM.PRICE, ITEMS_DIM.ITEM_ID FROM ITEMS_DIM ORDER BY  
ITEMS_DIM.ITEM_ID, ITEMS_DIM.PRICE --
```

Usando uma máquina com mais memória

Para melhorar o desempenho da sessão, execute a sessão em um nó do Serviço de Integração com uma grande quantidade de memória. Aumente o tamanho do cache de índice e de dados para o máximo sem sobrecarregar a máquina. Se o nó do Serviço de Integração tiver memória suficiente, aumente o cache para que ele possa manter todos os dados na memória sem paginação do disco.

Otimizando a condição de pesquisa

Se você incluir mais de uma condição de pesquisa, coloque-as na seguinte ordem para otimizar o desempenho de pesquisa:

- Igual a (=)
- Menor que (<), maior que (>), menor ou igual a (<=), maior ou igual a (>=)
- Diferente de (!=)

Filtrando linhas de pesquisa

Crie uma condição de filtro para reduzir o número de linhas de pesquisa recuperadas da origem quando o cache de pesquisa for criado.

Indexação da tabela de pesquisa

O Serviço de Integração precisa consultar, classificar e comparar valores nas colunas de condição de pesquisa. O índice precisa incluir todas as colunas usadas em uma condição de pesquisa.

É possível melhorar o desempenho para os seguintes tipos de pesquisas:

- **Pesquisas armazenadas em cache.** Para melhorar o desempenho, indexe as colunas na instrução ORDER BY de pesquisa. O log da sessão contém a instrução ORDER BY.
- **Pesquisas não armazenadas em cache.** Para melhorar o desempenho, indexe as colunas na condição de pesquisa. O Serviço de Integração emite uma instrução SELECT para cada linha que passa pela transformação de Pesquisa.

Otimizando várias pesquisas

Se um mapeamento contiver várias pesquisas, mesmo com o armazenamento em cache habilitado e com memória de heap suficiente, as pesquisas podem reduzir o desempenho. Ajuste as transformações de Pesquisa que pesquisam os maiores valores de dados para melhorar o desempenho geral.

Para determinar que transformações de Pesquisa processam a maior quantidade de dados, examine os contadores Lookup_rowsinlookupcache de cada transformação de Pesquisa. As transformações de Pesquisa que têm um número maior no contador podem se beneficiar do ajuste das expressões de pesquisa. Se as expressões forem otimizadas, o desempenho da sessão será melhorado.

TÓPICOS RELACIONADOS:

- [“Otimizando expressões” na página 30](#)

Criando uma transformação pesquisa de pipeline

Um mapeamento que contém a transformação de Pesquisa de pipeline inclui um pipeline parcial que contém a origem da pesquisa e um qualificador da origem. O Serviço de Integração processa os dados de origem de pesquisa no pipeline. Ele transfere os dados de origem de pesquisa para o pipeline que contém a transformação de Pesquisa e cria o cache.

O pipeline parcial está em um grupo de classificação para carregamento de destino separado nas propriedades da sessão. Você pode configurar diversas partições no pipeline para melhorar o desempenho.

Otimizando Transformações de Normalizador

As transformações de Normalizador geram linhas. Para otimizar o desempenho, coloque a transformação do Normalizador o mais próximo possível do destino.

Otimizando transformações de Gerador de Sequência

Para otimizar as transformações de Gerador de Sequência, crie um Gerador de Sequência reutilizável e use-o em vários mapeamentos simultaneamente. Além disso, configure a propriedade Número de Valores em Cache.

A propriedade Número de Valores em Cache determina o número de valores que o Serviço de Integração armazena em cache de uma vez. Certifique-se de que o Número de Valores em Cache não seja muito pequeno. Considere configurar o Número de Valores em Cache com um valor superior a 1.000.

Se você não tiver valores armazenados em cache, configure o Número de Valores em Cache como 0. As transformações de Gerador de Sequência que não usam cache são mais rápidas que as que o exigem.

Quando você conectar a porta CURRVAL em uma transformação de Gerador de Sequência, o Serviço de Integração processará uma linha em cada bloco. É possível otimizar o desempenho conectando apenas a porta NEXTVAL em um mapeamento.

TÓPICOS RELACIONADOS:

- [“Otimizando transformações de Gerador de Sequência” na página 56](#)

Otimizando transformações de classificador

Execute as seguintes etapas para otimizar uma transformação de Classificador:

- Aloque memória suficiente para classificar os dados.

- Especifique um diretório de trabalho diferente para cada partição na transformação de Classificador.
- Executar o Serviço de Integração do PowerCenter no modo ASCII.

Alocando memória

Para obter o desempenho ideal, configure o tamanho do cache do Classificador com um valor menor ou igual à área da RAM física disponível no nó do Serviço de Integração. Aloque, pelo menos, 16 MB de memória física para classificar dados usando a transformação de Classificador. Por padrão, o tamanho do cache do Classificador está definido como 16.777.216 bytes. Se o Serviço de Integração não puder alocar memória suficiente para a classificação de dados, a sessão falhará.

Se o volume de dados de entrada for maior que o tamanho do cache do Classificador, o Serviço de Integração armazenará temporariamente os dados no diretório de trabalho da transformação de Classificador. Ao armazenar dados no diretório de trabalho, o Serviço de Integração requer que o espaço em disco seja pelo menos duas vezes maior que o volume dos dados de entrada. Se o volume de dados de entrada for muito maior que o tamanho do cache do Classificador, pode ser que o Serviço de Integração precise de bem mais que o dobro de espaço em disco disponível para o diretório de trabalho.

Nota: O log da sessão contém a contagem de linhas de entrada e o tamanho dos dados de entrada para a transformação de Classificador.

Por exemplo, a seguinte mensagem é exibida quando o Serviço de Integração processa a transformação de Classificador:

```
SORT_40422 Final da saída da Transformação de Classificador [srt_1_Billion_FF]. 999999999
linhas foram processadas (866325637228 bytes de entrada; 868929593344 bytes de E/S
temporários)
```

Na mensagem, o número de linhas de entrada é 999999999 e o tamanho das linhas de entrada é 866325637228.

Diretórios de trabalho para partições

O Serviço de Integração cria arquivos temporários quando classifica dados. Ele os armazena em um diretório de trabalho. Você pode especificar qualquer diretório no nó do Serviço de Integração como diretório de trabalho. Por padrão, o Serviço de Integração usa o valor especificado para a variável de processo de serviço \$PMTempDir.

Ao particionar uma sessão com a transformação de Classificador, você pode especificar um diretório de trabalho diferente para cada partição no pipeline. Para melhorar o desempenho da sessão, especifique diretórios de trabalho em discos fisicamente separados nos nós do Serviço de Integração.

Modo Unicode

Para otimizar uma transformação de Classificador, execute o Serviço de Integração do PowerCenter no modo ASCII. Se o Serviço de Integração do PowerCenter for executado no modo Unicode, a transformação de Classificador despejará dados adicionais para o disco.

Otimizando transformações de Qualificador de Origem

Use a opção Selecionar Distintos para a transformação de Qualificador de Origem se desejar que o Serviço de Integração selecione valores exclusivos de uma origem. Use a opção Selecionar Distintos para filtrar dados desnecessários no início do fluxo de dados. Isso pode melhorar o desempenho.

Otimizando transformações de SQL

Quando você cria uma transformação de SQL, você configura a transformação para usar consultas SQL externas ou consultas definidas por você na transformação. Quando você configura uma transformação de SQL para executar em modo de script, o Serviço de Integração processa um script SQL externo para cada linha de entrada. Quando a transformação executa no modo de consulta, o Serviço de Integração processa uma consulta SQL definida por você na transformação.

Cada vez que o Serviço de Integração processa uma nova consulta em uma sessão, ele chama uma função chamada SQLPrepare para criar um procedimento SQL e transferi-lo ao banco de dados. Quando a consulta é alterada para cada linha de entrada, há impacto no desempenho.

Quando a transformação for executada no modo de consulta, construa uma consulta estática na transformação para melhorar o desempenho. Instruções de consulta estática não são alteradas, embora os dados na cláusula de consulta sejam. Para criar uma consulta estática, use a vinculação de parâmetros no lugar da substituição de string no Editor SQL. Quando você usa a vinculação de parâmetros, você define parâmetros na cláusula de consulta para valores nas portas de entrada de transformação.

Quando uma consulta SQL contém instruções de consulta de confirmação e de reversão, o Serviço de Integração precisa recriar o procedimento SQL depois de cada confirmação ou reversão. Para melhorar o desempenho, não use instruções de transação em consultas de transformação de SQL.

Quando você cria a transformação de SQL, você configura como a transformação se conecta ao banco de dados. É possível escolher uma conexão estática ou transferir informações de conexão para a transformação em tempo de execução.

Quando você configura a transformação para usar uma conexão estática, você escolhe uma conexão entre as conexões do Workflow Manager. A transformação de SQL conecta-se uma vez ao banco de dados durante a sessão. Quando você transfere informações de conexão dinâmica, a transformação de SQL conecta-se ao banco de dados todas as vezes que a transformação processa uma linha de entrada.

Otimizando Transformações de XML

Você pode remover grupos não projetados em uma transformação de Analisador de XML e de uma definição de origem XML. Você não precisará alocar memória para esses grupos não projetados, mas precisará manter as restrições de chave primária-chave estrangeira.

Eliminando erros de transformação

Em grande número, os erros de transformação reduzem o desempenho do Serviço de Integração. A cada erro de transformação o Serviço de Integração é interrompido para determinar a causa do erro e para remover do fluxo de dados a linha causadora do erro. Em seguida, o Serviço de Integração geralmente grava a linha no arquivo de log da sessão.

Os erros de transformação ocorrem quando o Serviço de Integração encontra erros de conversão, lógica de mapeamento conflitante e qualquer condição configurada como um erro, como uma entrada inválida. Consulte o log da sessão para verificar onde há erros de transformação. Se os erros se concentram em transformações específicas, avalie as restrições dessas transformações.

Se você não definir o limite de erro, o Serviço de Integração continuará a processar linhas de erro e, portanto, aumentará o tempo de execução da sessão. Para otimizar o desempenho, defina o limite de erro para interromper a sessão depois de um determinado número de erros de linha.

Se precisar executar uma sessão que gere um alto número de erros de transformação, é possível melhorar o desempenho configurando um nível de rastreamento inferior. No entanto, essa não é uma solução recomendada a longo prazo para erros de transformação.

TÓPICOS RELACIONADOS:

- [“Rastreamento de erros” na página 48](#)

CAPÍTULO 7

Otimizando sessões

Este capítulo inclui os seguintes tópicos:

- [Grade, 43](#)
- [Otimização de empilhamento, 44](#)
- [Sessões e fluxos de trabalho simultâneos, 44](#)
- [Memória de buffer, 44](#)
- [Caches, 46](#)
- [Confirmação baseada no destino, 47](#)
- [Processamento em tempo real, 48](#)
- [Áreas de preparação, 48](#)
- [Arquivos de log, 48](#)
- [Rastreamento de erros, 48](#)
- [Emails pós-sessão, 49](#)

Grade

Você pode usar uma grade para melhorar o desempenho da sessão e do fluxo de dados. Uma grade é uma alias atribuído a um grupo de nós que permite que você automatize a distribuição de fluxos de trabalho e sessões pelos nós.

Um Balanceador de Carga distribui as tarefas para os nós sem sobrecarregar nenhum deles.

Quando você usa uma grade, o Serviço de Integração distribui as tarefas do fluxo de trabalho e os segmentos da sessão por vários nós. Um Balanceador de Carga distribui as tarefas para os nós sem sobrecarregar nenhum deles. A execução de fluxos de trabalho e sessões nos nós de uma grade oferece os seguintes benefícios de desempenho:

- Equilibra a carga de trabalho do Serviço de Integração.
- Processa as sessões mais rapidamente.
- Processa partições mais rapidamente.

O Serviço de Integração exige recursos da CPU para a análise de dados de entrada e formatação de dados de saída. Uma grade poderá melhorar o desempenho quando você tiver um afunilamento de desempenho nas etapas de extração e carregamento de uma sessão.

A grade pode melhorar o desempenho quando a memória ou o armazenamento temporário consistir em um afunilamento de desempenho. Quando um mapeamento do PowerCenter contiver uma transformação com

memória armazenada em cache, a implantação de uma memória adequada e de um armazenamento em disco separado para cada instância de cache resultará no aumento do desempenho.

Executar uma sessão em uma grade pode melhorar a taxa de transferência de dados porque a grade oferece mais recursos para executar a sessão. O desempenho melhora quando você executa algumas sessões na grade em momentos específicos. Executar uma sessão em uma grade será mais eficiente que executar um fluxo de trabalho em uma grade se o número de partições de sessão simultânea for inferior ao número de nós.

Quando você executa várias sessões em uma grade, as subtarefas da sessão compartilham os recursos de nó com as subtarefas de outras sessões simultâneas. Executar uma sessão em uma grade exige a coordenação entre os processos que estão sendo executados em nós diferentes. Em alguns mapeamentos, a execução de uma sessão em uma grade requer uma sobrecarga adicional para mover os dados de um nó para outro. Além de carregar os recursos de memória e de CPU em cada nó, a execução de várias sessões em uma grade aumenta o tráfego da rede.

Quando você executa um fluxo de trabalho em uma grade, o Serviço de Integração carrega os recursos de memória e de CPU nos nós sem exigir a coordenação entre os nós.

TÓPICOS RELACIONADOS:

- [“Otimizando implantações de grade” na página 50](#)

Otimização de empilhamento

Para melhorar o desempenho da sessão, transfira a lógica de transformação para o banco de dados de origem ou de destino. Baseado na configuração de mapeamento e de sessão, o Serviço de Integração executa o SQL em relação ao banco de dados de origem ou de destino em vez de processar a lógica de transformação contida no Serviço de Integração.

Sessões e fluxos de trabalho simultâneos

Se possível, execute sessões e fluxos de trabalho simultaneamente para melhorar o desempenho. Por exemplo, se você carregar dados em um esquema analítico, onde você tem tabelas de dimensões e fatos, carregue as dimensões simultaneamente.

Memória de buffer

Quando o Serviço de Integração inicializa uma sessão, ele aloca blocos de memória para manter os dados de origem e de destino. O Serviço de Integração aloca pelo menos dois blocos para cada partição de origem e de destino. As sessões que usam um grande número de origens e de destinos podem exigir blocos adicionais de memória. Se o Serviço de Integração não puder alocar blocos de memória suficientes para manter os dados, a sessão falhará.

Você pode configurar a quantidade de memória de buffer ou pode configurar o Serviço de Integração para calcular as configurações de buffer em tempo de execução.

Para aumentar o número de blocos de memória disponíveis, ajuste as seguintes propriedades de sessão:

- **Tamanho do Buffer DTM.** Aumente o tamanho do buffer DTM na guia Propriedades das propriedades da sessão.
- **Tamanho Padrão do Bloco de Buffer.** Reduza o tamanho do bloco de buffer na guia Configuração do Objeto nas propriedades da sessão.

Nota: Se o particionamento de dados estiver ativado, o tamanho de buffer do DTM será o tamanho total de todos os pools de buffer de memória alocados para todas as partições. Em sessões que contêm n partições, defina o Tamanho de Buffer do DTM para pelo menos n vezes o valor da sessão com uma partição.

Aumentando o tamanho do buffer DTM

A configuração do Tamanho de Buffer do DTM especifica a quantidade de memória que o Serviço de Integração usa como memória de buffer do DTM. Quando você aumenta a memória de buffer DTM, o Serviço de Integração cria mais blocos de buffer, o que melhora o desempenho durante períodos de lentidão momentâneos.

O aumento da alocação de memória do buffer DTM geralmente faz com que o desempenho melhore inicialmente e depois se estabilize. Se você não perceber uma melhora significativa no desempenho, a alocação de memória de buffer do DTM não será um determinante do desempenho da sessão.

Para aumentar o tamanho do buffer DTM, abra as propriedades da sessão e clique na guia Propriedades. Edite a propriedade Tamanho de Buffer do DTM nas configurações de Desempenho. Aumente a propriedade Tamanho de buffer DTM em múltiplos do tamanho do bloco de buffer.

Otimizando o tamanho do bloco de buffer

Se a máquina tiver uma memória física limitada e o mapeamento na sessão contiver um grande número de origens, destinos ou partições, poderá ser necessário reduzir o tamanho do bloco de buffer.

Se estiver manipulando um número excepcionalmente alto de linhas de dados, aumente o tamanho do bloco de buffer para melhorar o desempenho. Se você não souber o tamanho aproximado das linhas, determine o tamanho da linha executando as etapas a seguir.

Para avaliar o tamanho do bloco de buffer necessário:

1. No Mapping Designer, abra o mapeamento referente à sessão.
2. Abra a instância de destino.
3. Clique na guia Portas.
4. Adicione a precisão para todas as colunas do destino.
5. Se você tem mais de um destino no mapeamento, repita as etapas [2](#) a [4](#) em cada destino adicional para calcular a precisão de cada destino.
6. Repita as etapas [2](#) a [5](#) em cada definição de origem do mapeamento.
7. Escolha a maior precisão de todas as precisões de origem e de destino como a precisão total no cálculo do tamanho do bloco de buffer.

A precisão total representa os bytes totais necessários para mover a maior linha de dados. Por exemplo, se a precisão total for igual a 33.000, o Serviço de Integração exigirá 33.000 bytes no bloco de buffer para mover essa linha. Se o tamanho do bloco de buffer for somente 64.000 bytes, o Serviço de Integração não poderá mover mais de uma linha por vez.

Para definir o tamanho do bloco de buffer, abra as propriedades da sessão e clique na guia Configuração do Objeto. Edite a propriedade Tamanho Padrão do Bloco de Buffer nas configurações Avançadas.

Da mesma forma que a alocação de memória de buffer DTM, o aumento do tamanho do bloco de buffer deverá melhorar o desempenho. Se você não perceber um aumento, o tamanho do bloco de buffer não é um determinante do desempenho da sessão.

Caches

O Serviço de Integração usa caches de índice e de dados para os destinos XML e para as transformações de Agregador, Classificação, Pesquisa e Unificador. O Serviço de Integração armazena os dados transformados no cache de dados antes de retorná-los ao pipeline. Ele armazena as informações do grupo no cache de índice. Além disso, o Serviço de Integração usa um cache para armazenar os dados para transformações de Classificador.

Para configurar a quantidade de memória em cache, use o calculador de cache ou especifique o tamanho do cache. Você também pode configurar o Serviço de Integração para calcular as configurações da memória de cache em tempo de execução.

Se o cache alocado não for grande o suficiente para armazenar os dados, o Serviço de Integração armazenará os dados em um arquivo de disco temporário, um arquivo de cache, pois ele processa os dados da sessão. O desempenho fica mais lento todas as vezes que o Serviço de Integração realizar a paginação em um arquivo temporário. Examine os contadores de desempenho para determinar a frequência na qual o Serviço de Integração realiza a paginação em um arquivo.

Execute as seguintes tarefas para otimizar caches:

- Limite o número de portas conectadas de entrada/saída e somente saída.
- Selecione a localização ideal do diretório de cache.
- Aumente os tamanhos de cache.
- Use a versão de 64 bits do PowerCenter para executar sessões de cache grande.

Limitando o número de portas conectadas

Em transformações que usam cache de dados, limite o número de portas conectadas de entrada/saída e somente saída. Limitar o número de portas de entrada/saída ou de saída conectadas reduz o volume de dados que as transformações armazenam no cache de dados.

Localização do diretório de cache

Se você executar o Serviço de Integração em uma grade e apenas alguns nós desse serviço tiverem acesso rápido ao diretório do arquivo de cache compartilhado, configure as sessões com cache grande para serem executadas nos nós com acesso rápido ao diretório. Para configurar uma sessão para ser executada em um nó com acesso rápido ao diretório, execute as seguintes etapas:

1. Crie um recurso do PowerCenter.
2. Disponibilize o recurso para os nós com acesso rápido ao diretório.
3. Atribua o recurso à sessão.

Se todos os processos do Serviço de Integração de uma grade tiverem acesso lento aos arquivos de cache, configure um diretório de arquivo de cache local separado para cada processo do Serviço de Integração. Um processo do Serviço de Integração poderá ter acesso mais rápido aos arquivos de cache se for executado na mesma máquina que contém o diretório de cache.

Nota: Pode ocorrer redução do desempenho quando você armazenar em cache grandes quantidades de dados em uma unidade mapeada ou montada.

Aumentando os tamanhos de cache

O tamanho do cache é configurado para especificar a quantidade de memória alocada para processar uma transformação. A quantidade de memória configurada depende da quantidade de cache de memória e cache de disco que você deseja usar. Se você configurar o tamanho do cache e ele não for suficiente para processar a transformação na memória, o Serviço de Integração processará parte da transformação na memória e paginará informações a arquivos de cache para processar o resto da transformação. Cada vez que o Serviço de Integração realiza a paginação de um arquivo de cache, o desempenho é reduzido.

Você pode examinar os detalhes de desempenho de uma sessão para determinar quando o Serviço de Integração realizará a paginação para um arquivo de cache. Os contadores *Transformation_readfromdisk* ou *Transformation_writetodisk* de qualquer transformação de Agregador, Classificação ou Unificador indicam o número de vezes que o Serviço de Integração realiza paginação no disco para processar a transformação.

Se a sessão contiver uma transformação que use um cache e você executar a sessão em uma máquina com memória suficiente, aumente os tamanhos de cache para processar a transformação na memória.

Usando a versão de 64 bits do PowerCenter

Se você processar grandes volumes de dados ou executar transformações com uso intenso da memória, você poderá usar a versão de 64 bits do PowerCenter para melhorar o desempenho da sessão. A versão de 64 bits oferece um espaço com mais memória que pode reduzir significativamente ou eliminar a entrada/saída do disco.

Isso pode melhorar o desempenho da sessão nas seguintes áreas:

- **Armazenamento em cache.** Com uma plataforma de 64 bits, o Serviço de Integração não fica limitado ao limite de cache de 2 GB de uma plataforma de 32 bits.
- **Taxa de transferência de dados.** Com um espaço maior de memória disponível, o leitor, o gravador e os segmentos DTM podem processar blocos de dados maiores.

Confirmação baseada no destino

A configuração de intervalo de confirmação determina o ponto em que o Serviço de Integração confirma os dados para os destinos. Cada vez que o Serviço de Integração confirma, o desempenho é reduzido. Portanto, quanto menor o intervalo de confirmação, maior a frequência em que o Serviço de Integração gravará no banco de dados de destino e mais lento ficará o desempenho geral.

Se você aumentar o intervalo de confirmação, o número de vezes em que o Serviço de Integração confirmará será reduzido e o desempenho ficará melhor.

Quando você aumentar o intervalo de confirmação, considere os limites de arquivos de log no banco de dados de destino. Se o intervalo de confirmação for muito alto, o Serviço de Integração poderá preencher o arquivo de log do banco de dados e gerar uma falha na sessão.

Portanto, pondere sobre o benefício do aumento do intervalo de confirmação em relação ao tempo adicional que você levaria para recuperar-se de uma sessão com falha.

Clique nas configurações de Opções Gerais nas propriedades da sessão para verificar e ajustar o intervalo de confirmação.

Processamento em tempo real

Latência de liberação

A latência de liberação determina com que frequência o Serviço de Integração libera os dados em tempo real da origem. Quanto mais baixo for definido o intervalo de latência de liberação, maior será a frequência com que o Serviço de Integração confirmará as mensagens para o destino. Cada vez que o Serviço de Integração confirmar mensagens para o destino, a sessão consumirá mais recursos e a taxa de transferência será interrompida.

Aumente a latência de liberação para melhorar a taxa de transferência. A taxa de transferência aumenta quando você aumenta a latência de liberação até um certo limite dependendo do hardware e dos recursos disponíveis.

Confirmações baseadas na origem

O intervalo de confirmação baseado na origem determina a frequência em que o Serviço de Integração confirmará dados em tempo real para o destino. Para obter uma latência mais rápida, defina a confirmação baseada na origem como 1.

Áreas de preparação

Quando você usa uma área de preparação, o Serviço de Integração executa várias transferências nos dados. Quando possível, remova as áreas de preparação para melhorar o desempenho. O Serviço de Integração pode ler várias origens com uma única transferência, o que pode aliviar a necessidade de áreas de preparação.

TÓPICOS RELACIONADOS:

- [“Configurando a leitura de passagem única” na página 28](#)

Arquivos de log

O fluxo de trabalho são executados mais rapidamente quando você não os configura para arquivos de sessão de gravação e de log de fluxo de trabalho. Os fluxos de trabalho e sessões sempre criam logs binários. Quando você configura uma sessão ou fluxo de trabalho para a gravação de um arquivo de log, o Serviço de Integração grava os eventos de log duas vezes. Você pode acessar a sessão de logs binários e os logs de fluxo de trabalho no Console de Administração.

Rastreamento de erros

Para melhorar o desempenho, reduza o número de eventos de log gerados pelo Serviço de Integração quando ele executar a sessão. Se uma sessão contiver um grande número de erros de transformação e você não precisar corrigi-los, defina o nível de rastreamento como Conciso. Nesse nível de rastreamento, o

Serviço de Integração não grava mensagens de erro ou informações em nível de linha referentes a dados rejeitados.

Se precisar depurar o mapeamento e tiver definido o nível de rastreamento como Detalhado, pode haver um impacto negativo significativo no desempenho quando você executar a sessão. Não use o rastreamento Detalhado ao ajustar o desempenho.

O nível de rastreamento da sessão substituirá qualquer outro nível de rastreamento específicos da transformação contido no mapeamento. Isso não é recomendado como uma solução a longo prazo para altos níveis de erros de transformação.

Emails pós-sessão

Quando você anexar o log da sessão a um email pós-sessão, habilite o registro em log de arquivo simples. Se você habilitar o registro em log de arquivo simples, o Serviço de Integração obterá o arquivo de log da sessão por meio do disco. Se você não habilitar o registro em log de arquivo simples, o Serviço de Integração obterá os eventos de log por meio do Log Manager e irá gerar o arquivo de log da sessão a ser anexado ao email. Quando o Serviço de Integração recuperar o log da sessão do serviço de log, o desempenho do fluxo de trabalho será reduzido, principalmente quando o arquivo de log for grande e o serviço de log for executado em um nó diferente do DTM mestre. Para obter o desempenho ideal, configure a sessão para gravar no arquivo de log quando você configurar o email pós-sessão para anexar um log de sessão.

CAPÍTULO 8

Otimizando implantações de grade

Este capítulo inclui os seguintes tópicos:

- [Visão geral do capítulo Otimizando a implantação de grades, 50](#)
- [Armazenando arquivos, 50](#)
- [Usando um sistema de arquivos compartilhado, 52](#)
- [Distribuindo arquivos pelos sistemas de arquivos, 54](#)
- [Otimizando transformações de Gerador de Sequência, 56](#)

Visão geral do capítulo Otimizando a implantação de grades

Ao executar o PowerCenter em uma grade, você pode configurar a grade, as sessões e os fluxos de trabalho para usar os recursos com eficiência e maximizar a escalabilidade.

Para aprimorar o desempenho do PowerCenter em uma grade, execute as seguintes tarefas:

- Adicione nós à grade.
- Aumente a capacidade e a largura de banda do armazenamento.
- Use sistemas de arquivos compartilhados.
- Use uma rede de alta transferência ao executar as seguintes tarefas:
 - Acesse origens e destinos pela rede.
 - Transfira dados entre os nós de uma grade ao usar a opção Sessão em Grade.

Armazenando arquivos

Ao configurar o PowerCenter para ser executado em uma grade, especifique a localização de armazenamento para diferentes tipos de arquivos de sessão, como arquivos de origem, de log e de cache. Para aprimorar o desempenho, armazene os arquivos em localizações ideais. Por exemplo, armazene

arquivos de cache persistente em um sistema de arquivos compartilhado de alta largura de banda. Os diferentes tipos de arquivo têm requisitos de armazenamentos diferentes.

É possível armazenar arquivos nos seguintes tipos de localizações:

- **Sistemas de arquivos compartilhados.** Armazene arquivos em um sistema de arquivos compartilhado para permitir que todos os processos do Serviço de Integração acessem os mesmos arquivos. Você pode armazenar arquivos em sistemas de arquivos compartilhados de alta ou baixa largura de banda.
- **Local.** Armazene arquivos na máquina local que esteja executando o processo do Serviço de Integração quando eles não tiverem que ser acessados por outros processos do Serviço de Integração.

Arquivos do sistema de arquivos compartilhado de alta largura de banda

Como eles podem ser acessados com frequência durante uma sessão, coloque os seguintes arquivos em um sistema de arquivos compartilhado de alta largura de banda:

- Arquivos de origem, incluindo arquivos simples para pesquisas.
- Arquivos de destino, incluindo arquivos de mesclagem para sessões particionadas.
- Arquivos de cache persistentes para pesquisa ou agregação incremental.
- Arquivos de cache não persistentes *somente* para sessões habilitadas para grade em uma grade.

Isso permite que o Serviço de Integração crie o cache somente uma vez. Se esses arquivos de cache são armazenados em um sistema de arquivo local, o Serviço de Integração cria um cache para cada grupo de partição.

Arquivos do sistema de arquivos compartilhado de baixa largura de banda

Como eles podem ser acessados com menos frequência durante uma sessão, armazene os seguintes arquivos em um sistema de arquivos compartilhado de baixa largura de banda:

- Arquivos de parâmetro ou outros arquivos relacionados a configuração.
- Arquivos de origem ou destino indiretos.
- Arquivos de log.

Arquivos de armazenamento local

Para evitar o compartilhamento desnecessário de arquivo ao usar sistemas de arquivos compartilhados, armazene os seguintes arquivos localmente:

- Arquivos de cache não persistentes que não estão habilitados para uma grade, incluindo arquivos temporários de transformação Classificador.
- Arquivos de destino individuais de diferentes partições ao executar uma mesclagem sequencial para sessões particionadas.
- Outros arquivos temporários excluídos no final da execução de uma sessão. Em geral, para estabelecer isso, configure \$PmTempFileDir para um sistema de arquivo local.

Evite armazenar esses arquivos em um sistema de arquivos compartilhado, mesmo quando a largura de banda é alta.

Usando um sistema de arquivos compartilhado

Você pode usar os seguintes sistemas de arquivos compartilhados para o compartilhamento de arquivo:

- Sistemas de arquivo de rede, como CIFS (SMB) no Windows ou Network File System (NFS) no UNIX. Embora os sistemas de arquivo de rede não tenham sido desenvolvidos para computação de alto desempenho, eles podem funcionar bem para o acesso de arquivo sequencial.
- Sistemas de arquivo clusterizados. Os sistemas de arquivo clusterizados oferecem um grupo de nós com acesso de arquivo de alta largura de banda, assim como um namespace unificado para arquivos e diretórios. O desempenho do sistema de arquivo clusterizado é similar ao do sistema de arquivo local anexado diretamente.

Nota: Se você tiver a opção Alta disponibilidade, use um sistema de arquivo clusterizado.

A configuração e o ajuste adequados são importante para o desempenho de grade pequena. Você também pode configurar mapeamentos e sessões para evitar as limitações intrínsecas dos sistemas de arquivos de compartilhados.

Configurando um sistema de arquivos compartilhado

Use as seguintes diretrizes gerais para configurar os sistemas de arquivos compartilhados:

- Certifique-se de que a rede tem largura de banda suficiente.
- Certifique-se de que o armazenamento subjacente tem largura de banda E/S suficiente.
- Configure os daemons do sistema de arquivos compartilhado, particularmente o cliente, para ter threads suficientes para acessar os arquivos com rapidez. Por exemplo, a IBM recomenda que você estime o número de arquivos que requerem acesso simultâneo e oferecem no mínimo dois threads biod para cada arquivo.

Ao executar sessões simultâneas em uma grade que usa origens e destinos de arquivo simples, fornece threads suficientes para que cada partição possa acessar os arquivos de origem e destino necessários simultaneamente.

- Configure os pontos de montagem do sistema de arquivos compartilhado com base nos requisitos de acesso. Ao executar sessões sequenciais em uma grade que usa origens e destinos de arquivo simples, evite qualquer configuração que possa diminuir a eficiência do processo de leitura antecipada e de gravação posterior. Os sistemas de arquivo otimizam o acesso ao arquivo sequencial com leitura antecipada e gravação posterior.
- Se necessário, ajuste as configurações de leitura antecipada e de gravação posterior do sistema de arquivos compartilhado.
- Analise as configurações de cache dos sistemas de arquivos compartilhados do cliente e do servidor. Aumentar as configurações pode aprimorar o desempenho.
- Configure as configurações de liberação posterior do sistema de arquivo para liberar páginas de memória depois que os dados são acessados. Caso contrário, o desempenho do sistema pode diminuir ao ler ou gravar arquivos grandes.
- Devido a diferença nos padrões de acesso, é necessário usar pontos de montagem diferentes para caches de origens, de destinos e persistentes.

Para obter mais informações, consulte a documentação do sistema de arquivos compartilhado.

Equilibrando o uso da CPU e da memória

Diferentemente dos sistemas de arquivo locais, um servidor de sistema de arquivos compartilhado pode obter ciclos extra da CPU para acessar os arquivos. Se usar um dos nós de computação como o servidor do

sistema de arquivos compartilhado para o resto dos nós, ele poderá ficar sobrecarregado e causar um afunilamento em toda grade. Quando o servidor do sistema de arquivos compartilhado está sobrecarregado, os ciclos da CPU podem aumentar, juntamente com as solicitações de transmissão e de tempo limite.

Para evitar isso, use uma ou mais máquinas como servidores dedicados do sistema de arquivos compartilhado para os nós da grade do PowerCenter. Cada máquina deve ter armazenamento, CPUs e largura de banda de rede suficientes para as tarefas necessárias.

Como alternativa, é possível montar de forma cruzada o servidor do sistema de arquivos compartilhado para distribuir a carga do servidor do arquivo entre os nós da grade. Quando os mapeamentos e as sessões do PowerCenter são configurados para serem executados com um equilíbrio do uso da CPU e de E/S, a montagem cruzada dos servidores do sistema de arquivos compartilhado pode otimizar o desempenho. Se o número de nós na grade é pequeno e você tem uma mistura equilibrada do uso da CPU e de E/S, pode não ser necessário um servidor dedicado do sistema de arquivos compartilhado.

Ao usar mais de um servidor dedicado ou montado de forma cruzada do sistema de arquivos compartilhado, tente distribuir os arquivos compartilhados entre os servidores.

TÓPICOS RELACIONADOS:

- [“Distribuindo arquivos pelos sistemas de arquivos” na página 54](#)

Configurando mapeamentos e sessões do PowerCenter

Uma das maneiras mais importantes de aprimorar o desempenho é evitar o compartilhamento desnecessário de arquivos. Quando configurados de forma adequada, os sistemas de arquivos compartilhados podem oferecer um bom desempenho para o acesso de arquivos de origem e destino. No entanto, o acesso aleatório necessário para arquivos de cache persistentes, especialmente os grandes, pode ser mais problemático.

Use as seguintes diretrizes para configurar os arquivos de cache persistentes, como pesquisas dinâmicas persistentes, para uma grade com um sistema de arquivos compartilhado:

- Quando possível, configure o tamanho do cache da sessão para manter arquivos de cache persistentes menores na memória.
- Adicione uma transformação do Classificador ao mapeamento para classificar as linhas de entrada antes da pesquisa persistente. Alterar o trabalho da pesquisa persistente para a transformação do Classificador pode aprimorar o desempenho porque a transformação do Classificador pode usar o sistema de arquivo local.
- As linhas de grupo que requerem acesso à mesma página do cache de pesquisa para minimizar o número de vezes que o Serviço de Integração lê cada página do cache.
- Quando o tamanho dos dados de entrada é grande, use as confirmações baseadas na origem para gerenciar os dados de entrada para que a classificação seja executada na memória.

Por exemplo, você tem uma pesquisa dinâmica persistente de 4 GB, que não pode ser reduzida sem a alteração da lógica do mapeamento, e tem 10 GB de dados de origem. Primeiro, adicione uma transformação do Classificador para classificar os dados de entrada e reduzir o acesso aleatório ao cache de pesquisa. Em seguida, execute as seguintes tarefas:

- Configure a sessão para executar confirmações baseadas na origem com intervalos de confirmação de 1 GB.
- Defina o escopo da transação de transformação do Classificador como Transação.
- Configure a transformação do Classificador para um tamanho de cache de 1 GB, o suficiente para a entrada de origem.

Com essa configuração, o Serviço de Integração classifica 1 GB de dados de entrada por vez e passa as linhas para a pesquisa persistente que requer acesso aos dados similares no cache.

- Se mais de um sistema de arquivo estiver disponível, configure os arquivos do cache de cada arquivo para usar diferentes sistemas de arquivo.
- Configure as sessões para distribuir os arquivos para sistemas de arquivo diferentes de mais de um sistema de arquivo estiver disponível.

Distribuindo arquivos pelos sistemas de arquivos

Distribua os arquivos em diferentes sistemas de arquivo para usar a largura de banda combinada dos sistemas de arquivo, considerando que cada sistema utiliza um subsistema de disco físico independente. Distribuir arquivos para sistemas de arquivo diferentes pode aprimorar o desempenho em uma grade que utiliza um sistema de arquivos compartilhado ou SMP (multiprocessamento simétrico).

Para obter largura de banda E/S ideal, escolha um sistema de arquivo que distribui arquivos entre diversos dispositivos de armazenamento. Se você usa um sistema de arquivo clusterizado, distribua os arquivos entre os servidores. Se possível, coloque os arquivos de origem, de destino e de cache em diferentes dispositivos de armazenamento.

Use as seguintes diretrizes ao distribuir os arquivos nos sistemas de arquivo:

- **Arquivos de origem.** Se você colocar arquivos de origem em um sistema de arquivo que permite que o Serviço de Integração leia os dados de um grande número de arquivos, ajuste a configuração de leitura antecipada do sistema de arquivo antes de armazenar os arquivos grandes em cache.
- **Arquivos temporários.** Se você colocar arquivos temporários em um sistema de arquivo que permite que o Serviço de Integração leia os dados de arquivos grandes e os grave em arquivos temporários, ajuste as configurações de leitura e gravação dos arquivos grandes.
- **Arquivos de destino.** Se você colocar arquivos de destino em um sistema de arquivo que permite que o Serviço de Integração grave arquivos grandes no disco, ajuste o sistema de arquivo para gravações em bloco grandes e simultâneas. Os arquivos de destino podem incluir arquivos de mesclagem para sessões particionadas. Como as sessões particionadas em uma grade precisam gravar arquivos no disco, ajuste o sistema de arquivo para obter desempenho ideal de bloqueio.

Configurando sessões para distribuir arquivos

Você pode configurar as sessões manualmente para distribuir a carga do arquivo. Pode ser necessário editar as sessões quando a carga é alterada significativamente ou quando novas sessões ou sistemas de arquivo são adicionados, incluindo a adição de novos nós a uma grade com um sistema montado de maneira cruzada de arquivos compartilhado.

Em vez de editar as sessões manualmente, use as variáveis de sessão para distribuir arquivos em diferentes diretórios. Isso permite que você redirecione arquivos de sessão para diferentes servidores de arquivo quando necessário.

Use as seguintes diretrizes para usar as variáveis de sessão:

- Nomeie as variáveis para os diretórios e os nomes de arquivo de sessão para refletir a lógica de negócios.
- No arquivo de parâmetro, defina cada variável para que a carga do arquivo seja distribuída igualmente entre os sistemas de arquivo disponíveis. Também é possível definir variáveis específicas de nó.

- Como opção, automatize a reconfiguração com um script para processar os arquivos de parâmetro.

Nota: Ao usar um script, use um espaço reservado no arquivo de parâmetro para que o script possa redefinir as variáveis da sessão, se necessário.

Diretrizes para scripts e arquivos de parâmetro

Ao criar arquivos e scripts de parâmetro, use as seguintes diretrizes:

- Para manter facilmente a flexibilidade e o controle das localizações do arquivo de sessão, use um script para substituir os espaços reservados no arquivo de parâmetro.
- Considere o tamanho estimado do arquivo e a capacidade do sistema de arquivo ao definir as localizações de arquivo.
- Evite organizar os arquivos de acordo com a lógica de negócios se as sessões e os fluxos de trabalho precisarem de acesso aos arquivos relacionados a negócios ao mesmo tempo. Por exemplo, se você armazenar arquivos Califórnia em um sistema de arquivo e arquivos Nova York em outro, um afunilamento poderá ocorrer se a sessão precisar de acesso a todos os arquivos ao mesmo tempo.
- Quando possível, coloque os arquivos para diferentes partições da mesma origem, destino ou pesquisa em diferentes sistemas de arquivo.

Exemplos

No seguinte fragmento de um arquivo de parâmetro bruto, o espaço reservado “{fs}” representa o sistema do arquivo onde o diretório está localizado e deve ser atribuído por um script antes de ser usado:

```
[SessionFFSrc_FFTgt_CA]
    $InputFile_driverInfo_CA={fs}/driverinfo_ca.dat
    $SubDir_processed_CA={fs}
# Session has Output file directory set to:
# $PmTargetFileDir/$SubDir_processed_CA
# This file is the input of SessionFFSrc_DBTgt_CA.
    $SubDir_RecordLkup_Cache_CA={fs}
# This session builds this persistent lookup cache to be used
    # by SessionFFSrc_DBTgt_CA.
# The Lookup cache directory name in the session is set to:
# $PmCacheDir/$SubDir_RecordLkup_Cache_CA
[SessionFFSrc_FFTgt_NY]
    $InputFile_driverInfo_NY={fs}/driverinfo_ny.dat
    $SubDir_processed_NY={fs}
[SessionFFSrc_DBTgt_CA]
    $SubDir_processed_CA={fs}
# session has Source file directory set to:
# $PmTargetFileDir/$SubDir_processed_CA
    $SubDir_RecordLkup_Cache_CA={fs}
# Use the persistent lookup cache built in SessionFFSrc_FFTgt_CA.
```

No seguinte fragmento do arquivo de parâmetro, um script substituiu o espaço reservado com os nomes do sistema de arquivo adequados, como `file_system_1` e `file_system_2`:

```
[SessionFFSrc_FFTgt_CA]
    $InputFile_driverInfo_CA=file_system_1/driverinfo_ca.dat
    $SubDir_processed_CA=file_system_2
# Session has Output file directory set to:
# $PmTargetFileDir/$SubDir_processed_CA
# This file is the input of SessionFFSrc_DBTgt_CA.
    $SubDir_RecordLkup_Cache_CA=file_system_1
# This session builds this persistent lookup cache to be used
    # by SessionFFSrc_DBTgt_CA.
# The Lookup cache directory name in the session is set to:
# $PmCacheDir/$SubDir_RecordLkup_Cache_CA
[SessionFFSrc_FFTgt_NY]
    $InputFile_driverInfo_NY=file_system_2/driverinfo_ny.dat
    $SubDir_processed_NY=file_system_1
```

```
[SessionFFSrc_DBTgt_CA]
    $SubDir_processed_CA=file_system_1
# session has Source file directory set to:
# $PmTargetFileDir/$SubDir_processed_CA
    $SubDir_RecordLkup_Cache_CA=file_system_2
# Use the persistent lookup cache built in SessionFFSrc_FFTgt_CA.
```

Otimizando transformações de Gerador de Sequência

Para aprimorar o desempenho ao executar uma sessão em uma grade com transformações do Gerador de Sequência, aumente o número de valores em cache para um número de cada linha de dados. Isso reduz a comunicação entre os processos DTM mestre e operador, e o repositório. Os DTMs mestre e operador se comunicam uma vez para cada valor em cache.

Por exemplo, você tem 150.000 de dados e sete transformações do Gerador de Sequência. O número de valores em cache é 10. Os DTMs mestre e operador se comunicam 15.000 vezes. Se você aumentar o número de valores em cache para 15.000, o DTM mestre e operador se comunicam dez vezes.

CAPÍTULO 9

Otimizando componentes do PowerCenter

Este capítulo inclui os seguintes tópicos:

- [Visão geral do capítulo Otimizando componentes do PowerCenter, 57](#)
- [Otimizando o desempenho do repositório do PowerCenter, 57](#)
- [Otimizando o desempenho do Serviço de Integração, 60](#)

Visão geral do capítulo Otimizando componentes do PowerCenter

Você pode otimizar o desempenho dos seguintes componentes do PowerCenter:

- Repositório do PowerCenter
- Serviço de Integração

Se você executa o PowerCenter em diversas máquinas, execute o o Serviço de Repositório e o Serviço de Integração em máquinas diferentes. Para carregar grandes quantidades de dados, execute o Serviço de Integração na máquina de maior processamento. Além disso, execute o Serviço de Repositório na máquina que esteja hospedando o repositório do PowerCenter.

Otimizando o desempenho do repositório do PowerCenter

Conclua as seguintes tarefas para aprimorar o desempenho do repositório do PowerCenter:

- Certifique-se de que o repositório do PowerCenter está na mesma máquina que o processo do Serviço de Repositório.
- Ordene as condições em consultas de objeto.
- Use um espaço de tabela de nó único para o repositório do PowerCenter se ele estiver instalado em um banco de dados DB2.

- Otimize o esquema do banco de dados do repositório do PowerCenter se ele estiver instalado em um banco de dados DB2 ou Microsoft SQL Server.

Localização do repositório e do processo do Serviço de Repositório

Você pode otimizar o desempenho do Serviço de Repositório configurado sem a opção de alta disponibilidade. Para otimizar o desempenho, certifique-se de que o processo do Serviço de Repositório está em execução na mesma máquina onde o banco de dados do repositório reside.

Condições de ordem em consultas de objeto

Quando o Serviço de Repositório processa um parâmetro com várias condições, ele as processa na ordem em que foram inseridas. Para receber os resultados esperados e melhorar o desempenho, insira os parâmetros na ordem em que você deseja que eles sejam executados.

Usando um espaço de tabela do banco de dados DB2 de nó único

Você pode otimizar o desempenho do repositório em bancos de dados IBM DB2 EEE ao armazenar em um repositório do PowerCenter em um espaço de tabela de nó único. Ao configurar um banco de dados IBM DB2 EEE, o administrador do banco de dados pode defini-lo em um nó único.

Quando o espaço de tabela contém um nó, o Cliente do PowerCenter e o Serviço de Integração acessam o repositório mais rapidamente do que se as tabelas do repositório existe em diferentes nós do banco de dados.

Se você não especificar o nome do espaço de tabela ao criar, copiar ou restaurar um repositório, o sistema DB2 especificará o espaço de tabela padrão para cada tabela do repositório. O sistema DB2 pode ou não especificar um espaço de tabela de nó único.

Otimizando o esquema do banco de dados

Você pode aprimorar o desempenho do repositório em bancos de dados IBM DB2 e Microsoft SQL Server ao habilitar a opção Otimizar o esquema do banco de dados para o Serviço de Repositório no Console de Administração. A opção Otimizar o esquema do banco de dados faz com que o Serviço de Repositório armazene dados de caracteres com comprimento variável em colunas Varchar (2000) em vez das colunas CLOB, sempre que possível. Usar as colunas Varchar (2000) aprimora o desempenho do repositório nas seguintes maneiras:

- **Acesso ao disco reduzido.** O repositório do PowerCenter armazena os dados do Varchar diretamente nas colunas da tabela do banco de dados. Ele armazena os dados do CLOB como referência para outra tabela. Para recuperar os dados do CLOB do repositório, o Serviço de Repositório deve acessar uma tabela do banco de dados para obter a referência e, em seguida, acessar a tabela referenciada para ler os dados. Para recuperar os dados do Varchar, o Serviço de Repositório acessa uma tabela do banco de dados.
- **Criação de cache aprimorada.** O gerenciador de buffer do banco de dados do repositório pode criar cache das colunas do Varchar, mas não das colunas do CLOB.

Otimizar o esquema do banco de dados pode aprimorar o desempenho do repositório nas seguintes operações:

- Fazendo backup de um repositório
- Restaurando um repositório

- Exportando objetos do repositório
- Listando dependências entre os objetos
- Implantando pastas

Em geral, o desempenho é aprimorado conforme os tamanhos do banco de dados do repositório e da página aumentam. Portanto, otimizar o esquema do banco de dados oferece maiores aprimoramentos de desempenho em maiores repositórios do PowerCenter.

Você pode otimizar o esquema do banco de dados ao criar conteúdos de repositório ou fazer backup e restaurar um repositório existente. Para otimizar o esquema do banco de dados, o banco de dados do repositório deve atender aos seguintes requisitos de tamanho de página:

- **IBM DB2.** Tamanho de página do banco de dados de 4 KB ou maior. Pelo menos um espaço de tabela temporário com tamanho de página de 16 KB ou maior.
- **Microsoft SQL Server.** Tamanho de página do banco de dados de 8 KB ou maior.

Cache de Objeto para o Serviço de Repositório

O cache de objeto de repositório aumenta o desempenho quando você executa fluxos de trabalho. Quando o Serviço de Repositório armazena objetos na memória, o Serviço de Integração pode acessar mais facilmente os objetos em cache necessários para concluir a execução do fluxo de trabalho.

Para gerenciar o cache de objetos para o Serviço de Repositório, configure as seguintes propriedades:

Propriedade	Descrição
Cache do Agente do Repositório	Opcional. Habilita o cache do agente de repositório, o que melhora o desempenho quando o Serviço de Integração executa várias sessões repetidamente. Se você habilitar esta propriedade, o processo do Serviço de Repositório armazena os metadados solicitados pelo Serviço de Integração e os metadados que descrevem os objetos de repositório. O padrão é Sim.
Capacidade do Cache do Agente	Opcional. Número de objetos que o cache pode conter quando o cache do agente do repositório está habilitado. Você pode aumentar o número de objetos se houver memória disponível na máquina que está executando o processo do Serviço de Repositório. O padrão é 10.000.
Permitir Gravações com o Cache do Agente	Opcional. Permite a utilização de ferramentas do Cliente do PowerCenter para alterar os metadados no repositório quando o cache do agente do repositório está habilitado. Ao habilitar a opção de gravações, o Serviço de Repositório libera o cache toda vez que você salva metadados usando o Cliente do PowerCenter. Se você desabilitar a opção de gravações, você não pode salvar metadados ao repositório por meio do Cliente do PowerCenter e o cache não é liberado. Quando você desabilita a opção de gravações, o Serviço de Integração ainda poderá gravar a sessão e os metadados do fluxo de trabalho ao repositório. O Serviço de Repositório não libera o cache quando o Serviço de Integração grava metadados. O padrão é Sim.

Otimização da resiliência

Configure as seguintes propriedades de resiliência dos serviços de aplicativo para melhorar o desempenho do repositório:

Propriedade	Descrição
Limite sobre os Tempos de Resiliência	Opcional. Período máximo de tempo que o serviço mantém os recursos para fins de resiliência. Esta propriedade põe uma restrição em clientes que conectam-se ao serviço. Quaisquer tempos de espera de resiliência que ultrapassam o limite são cortados. Se o valor desta propriedade estiver em branco, o valor é derivado das configurações do nível de domínio. O padrão é em branco.
Tempo de Espera de Resiliência	Opcional. Período de tempo que o serviço tenta estabelecer ou restabelecer uma conexão com outro serviço. Se o valor desta propriedade estiver em branco, o valor é derivado das configurações do nível de domínio. O padrão é em branco.

Otimizando o desempenho do Serviço de Integração

Conclua as seguintes tarefas para aprimorar o desempenho do Serviço de Integração:

- Use drivers nativos em vez de drivers ODBC no Serviço de Integração.
- Execute o Serviço de Integração no modo de movimentação de dados ASCII se os dados de caracteres forem ASCII de 7 bits ou EBCDIC.
- Crie cache no PowerCenter Metadata para o Serviço de Repositório.
- Execute o Serviço de Integração com alta disponibilidade.

Nota: Ao configurar o Serviço de Integração com alta disponibilidade, ele recupera os fluxos de trabalho e as sessões que podem falhar devido a erros temporário da máquina ou da rede. Para recuperar dados de uma sessão ou fluxo de trabalho, o Serviço de Integração grava os estados de cada sessão e fluxo de trabalho em arquivos temporários em um diretório compartilhado. Isso pode diminuir o desempenho.

Usando drivers ODBC e nativo

O Serviço de Integração pode usar drivers ODBC ou nativo para se conectar aos banco de dados. Use drivers nativos para aprimorar o desempenho.

Executando o Serviço de Integração no modo de movimentação de dados ASCII

Quando todos os dados de caracteres processados pelo Serviço de Integração são ASCII de 7 bits ou EBCDIC, configure o Serviço de Integração para ser executado no modo de movimentação de dados ASCII. No modo ASCII, o Serviço de Integração usa um byte para armazenar cada caractere. Ao executar o Serviço de Integração no modo Unicode, ele utilizará dois bytes para cada caractere, o que pode diminuir o desempenho da sessão.

Criando cache no PowerCenter Metadata para o Serviço de Repositório

É possível usar a criação de cache do agente do repositório para aprimorar o desempenho do processo do DTM. Quando você habilita a criação de cache do agente do repositório, o Serviço de Repositório cria cache dos metadados solicitados pelo Serviço de Integração. Ao criar cache dos metadados, o Serviço de Integração lê o cache para execuções adjacentes da tarefa em vez de buscar os metadados do repositório. Somente os metadados solicitados pelo Serviço de Integração são colocados em cache.

Por exemplo, você executa um fluxo de trabalho com 1.000 sessões. A primeira vez que você executa um fluxo de trabalho com a criação de cache habilitada, o Serviço de Integração busca os metadados da sessão no repositório. Durante as execuções subjacentes do fluxo de trabalho, o Serviço de Repositório busca os metadados da sessão no cache. Isso aprimora o desempenho do processo do DTM.

CAPÍTULO 10

Otimizando o sistema

Este capítulo inclui os seguintes tópicos:

- [Visão geral do capítulo Otimização do sistema, 62](#)
- [Aprimorando a velocidade da rede, 63](#)
- [Usando várias CPUs, 63](#)
- [Reduzindo a paginação, 63](#)
- [Usando a associação de processador, 64](#)

Visão geral do capítulo Otimização do sistema

Com frequência, o desempenho se torna lento porque a sessão conta com conexões insuficientes ou com um sistema de processo sobrecarregado do Serviço de Integração. Os atrasos do sistema podem também ser causados por roteadores, comutadores, protocolos de rede e o uso por diversos usuários.

O acesso lento ao disco em bancos de dados de origem e destino, sistemas de arquivos de origem e destino e nós no domínio pode diminuir o desempenho da sessão. Permita que o administrador do sistema avalie os discos rígidos nas máquinas.

Após determinar qual das ferramentas de monitoramento do sistema apresenta um afunilamento de sistema, execute as seguintes alterações globais para aprimorar o desempenho de todas as sessões:

- **Aumente a velocidade da rede.** Conexões lentas de rede podem diminuir o desempenho da sessão. Permita que o administrador do sistema determine se a rede está sendo executada em uma velocidade ideal. Diminua o número de saltos da rede entre o processo do Serviço de Integração e os bancos de dados.
- **Use várias CPUs.** Você pode usar diversas CPUs para executar várias sessões e partições de pipeline em paralelo.
- **Reduza a paginação.** Quando um sistema operacional fica sem memória física, ele começa a fazer paginação para o disco para obter memória física livre. Configure a memória física para a máquina do processo do Serviço de Integração para minimizar a paginação para o disco.
- **Use a associação do processador.** Em um ambiente de múltiplos processadores UNIX, o Serviço de Integração pode usar uma grande quantidade de recursos do sistema. Use a associação do processador para controlar a utilização do processador pelo processo do Serviço de Integração. Além disso, se os bancos de dados de origem e de destino estão na mesma máquina, use a associação do processador para limitar os recursos usados pelo banco de dados.

Aprimorando a velocidade da rede

O desempenho do Serviço de Integração está relacionado às conexões de rede. O disco local pode mover dados de 5 a 20 vezes mais rápido que a rede. Considere as opções a seguir para minimizar a atividade da rede e aprimorar o desempenho do Serviço de Integração.

Se você usa arquivo simples como uma origem ou destino em uma sessão e o Serviço de Integração é executado em um único nó, armazene os arquivos na mesma máquina que o Serviço de Integração para aprimorar o desempenho. Ao armazenar arquivos simples em uma máquina diferente do Serviço de Integração, o desempenho da sessão se torna dependente do desempenho das conexões de rede. Mover os arquivos para o sistema do processo do Serviço de Integração e adicionar espaço ao disco pode aprimorar o desempenho.

Se você utiliza bancos de dados de origem e destino relacionais, tente minimizar o número de saltos de rede entre os bancos de dados de origem e destino e o processo do Serviço de Integração. Mover o banco de dados de destino para um sistema do servidor pode aprimorar o desempenho do Serviço de Integração.

Ao executar sessões que contêm diversas partições, permita que o administrador da rede analise a rede e certifique-se de que existe largura de banda suficiente para tratar a movimentação dos dados entre a rede de todas as partições.

Usando várias CPUs

Configure o sistema para usar mais CPUs para aprimorar o desempenho. Diversas CPUs permitem que o sistema execute várias sessões e partições de pipeline em paralelo.

No entanto, as CPUs adicionais podem causar afunilamento no disco. Para evitar afunilamento no disco, minimize o número de processos que acessam o disco. Os processos que acessam o disco incluem funções do banco de dados e do sistema operacional. As sessões ou as partições de pipeline paralelas também exigem acesso ao disco.

Reduzindo a paginação

A paginação ocorre quando o sistema operacional do processo do Serviço de integração fica sem memória para uma operação em particular e usa o disco local para obter memória. Você pode liberar mais memória ou aumentar a memória física para reduzir a paginação e diminuir o desempenho que resulta em paginação. Monitore a atividade de paginação usando as ferramentas do sistema.

É possível aumentar a memória do sistema nas seguintes circunstâncias:

- Execute uma sessão que utilize grandes pesquisas em cache.
- Execute uma sessão com diversas partições.

Se você não conseguir liberar memória, poderá ser necessário adicionar memória ao sistema.

Usando a associação de processador

Em um ambiente de múltiplos processadores UNIX, o Serviço de Integração poderá usar uma grande quantidade de recursos do sistema se você usar um grande número de sessões. Como resultado, outros aplicativos na máquina podem não ter recursos do sistema suficientes disponíveis. É possível usar a associação do processador para controlar a utilização do processador pelo nó do processo do Serviço de Integração. Além disso, se os bancos de dados de origem e de destino estão na mesma máquina, use a associação do processador para limitar os recursos usados pelo banco de dados.

Em um ambiente Sun Solaris, os administrador do sistema pode criar e gerenciar um conjunto de processadores usando o comando `psrset`. Em seguida, o administrador do sistema poderá usar o comando `pbind` para associar o Serviço de Integração a um conjunto de processadores, para que o processador execute somente o Serviço de Integração. O ambiente Sun Solaris também oferece o comando `psrinfo` para exibir detalhes sobre cada processador configurado e o comando `psradm` para alterar o status operacional dos processadores. Para obter mais informações, consulte o administrador do sistema e a documentação do Sun Solaris.

Em um ambiente AIX, os administradores do sistema podem usar o Workload Manager no AIX 5L para gerenciar recursos do sistema durante demandas máximas. O Workload Manager pode alocar recursos e gerenciar a CPU, a memória e a largura de banda E/S do disco. Para obter mais informações, consulte o administrador do sistema e a documentação do AIX.

CAPÍTULO 11

Usando partições de pipeline

Este capítulo inclui os seguintes tópicos:

- [Visão geral de Usando partições de pipeline, 65](#)
- [Otimizando o banco de dados de origem para particionamento, 67](#)
- [Otimizando o banco de dados de destino para particionamento, 69](#)

Visão geral de Usando partições de pipeline

Depois de ajustar o aplicativo, os bancos de dados e o sistema para obter o melhor desempenho de partição única, você poderá perceber que o sistema está sendo subutilizado. Nesse momento, você pode configurar a sessão para que ela tenha duas ou mais partições.

Você pode usar o particionamento de pipeline para aprimorar o desempenho da sessão. Aumentar o número de partições ou de pontos de partição aumentará o número de segmentos. Portanto, aumentar o número de partições ou de pontos de partição também aumentará a carga nos nós do Serviço de Integração. Se o nó ou nós do Serviço de Integração contiverem uma largura de banda de CPU ampla, o processamento de linhas de dados em uma sessão poderá aumentar simultaneamente o desempenho da sessão.

Nota: Se você usar um Serviço de Integração de um único nó e criar um grande número de partições ou pontos de partição em uma sessão que processe grandes volumes de dados, o sistema poderá ficar sobrecarregado.

Se você tiver a opção de particionamento, execute as seguintes tarefas para configurar manualmente as partições:

- Aumente o número de partições.
- Selecione os tipos de partição de melhor desempenho em pontos específicos em um pipeline.
- Use várias CPUs.

Aumentando o número de partições

É possível aumentar a quantidade de partições em um pipeline com o objetivo de melhorar o desempenho da sessão. Aumentar o número de partições permitirá que o Serviço de Integração crie várias conexões a origens e processe partições de dados de origem simultaneamente.

Quando uma sessão usa um arquivo de origem, você pode configurá-la para ler a origem com um segmento ou com diversos segmentos. Configure a sessão para ler origens de arquivos com vários segmentos para aumentar o desempenho da sessão. O Serviço de Integração cria várias conexões simultâneas com a origem do arquivo.

Quando você cria uma sessão, o Workflow Manager valida cada pipeline do particionamento de mapeamento. Você poderá especificar várias partições em um pipeline se o Serviço de Integração puder manter consistência de dados ao processar os dados particionados.

Use as dicas a seguir ao adicionar partições a uma sessão:

- **Adicione uma partição de cada vez.** Para obter o melhor desempenho do monitor, adicione uma partição de cada vez e anote as configurações da sessão antes de adicionar as partições.
- **Defina a memória de buffer DTM.** Quando você aumenta o número de partições, o tamanho do buffer DTM também aumenta. Se a sessão contiver n partições, aumente o tamanho do buffer DTM para pelo menos n vezes o valor da sessão com uma partição.
- **Defina os valores armazenados em cache do Gerador de Sequência.** Se uma sessão contiver n partições, pode não ser necessário usar a propriedade “Número de Valores em Cache” para a transformação de Gerador de Sequência. Se você definir esse valor como maior que 0, certifique-se de que ele seja pelo menos n vezes o valor original da sessão com uma partição.
- **Particione os dados de origem uniformemente.** Configure cada partição para extrair o mesmo número de linhas.
- **Monitore o sistema ao executar a sessão.** Se os ciclos de CPU estiverem disponíveis, você poderá adicionar uma partição para melhorar o desempenho. Por exemplo, você pode disponibilizar ciclos de CPU se o sistema estiver 20 por cento do tempo inativo.
- **Monitore o sistema depois de adicionar uma partição.** Se a utilização da CPU não for elevada, a espera do tempo de E/S aumentar ou a taxa total de transformação de dados diminuir, então provavelmente há um afunilamento de hardware ou software. Se a espera do tempo de E/S aumentar em um valor significativo, verifique a existência de afunilamentos de hardware no sistema. Caso contrário, verifique a configuração do banco de dados.

TÓPICOS RELACIONADOS:

- [“Memória de buffer” na página 44](#)

Selecionando os tipos de partição de melhor desempenho

É possível especificar diferentes tipos de partição em diversos pontos do pipeline para melhorar o desempenho da sessão. Para otimizar o desempenho da sessão, use o tipo de partição de particionamento de banco de dados em bancos de dados de origem e de destino. Você pode usar o particionamento de banco de dados em origens Oracle e IBM DB2 e destinos IBM DB2. Quando você usa o particionamento de banco de dados de origem, o Serviço de Integração consulta o sistema do banco de dados para obter as informações sobre as partições e busca os dados nas partições da sessão. Quando você usa o particionamento de banco de dados de destino, o Serviço de Integração carrega dados nos nós de partição do banco de dados correspondente.

Você pode usar várias partições de pipeline e de banco de dados. Para melhorar o desempenho, verifique se o número de partições de pipeline é igual ao número de partições de banco de dados. Para melhorar o desempenho das origens Oracle subparticionadas, verifique se o número de partições de pipeline é igual ao número de subpartições de banco de dados.

Para melhorar o desempenho, especifique os tipos de partição nos seguintes pontos de partição no pipeline:

- **Transformação Qualificador de Origem.** Para ler os dados de vários arquivos simples simultaneamente, especifique uma partição para cada arquivo simples na transformação de Qualificador de Origem. Aceite o tipo de partição padrão, a de passagem.
- **Transformação de Filtro.** Como os arquivos de origem variam de tamanho, cada partição processa uma quantidade diferente de dados. Defina um ponto de partição na transformação de Filtro e escolha o particionamento round-robin para equilibrar a carga que entra na transformação de Filtro.

- **Transformação de Classificador.** Para eliminar grupos sobrepostos nas transformações de Classificador e de Agregador, use o particionamento de chaves automáticas hash na transformação de Classificador. Isso faz com que o Serviço de Integração agrupe todos os itens com a mesma descrição na mesma partição antes que as transformações de Classificador e de Agregador processem as linhas. Você pode excluir o ponto de partição padrão na transformação de Agregador.
- **Destino.** Como as tabelas de destino são particionadas por limite de chaves, especifique o particionamento de limite de chaves no destino para otimizar a gravação de dados no destino.

Usando várias CPUs

O Serviço de Integração executa o processo de leitura, transformação e gravação de pipelines em paralelo. Ele pode processar várias partições de um pipeline dentro de uma sessão e pode processar várias sessões em paralelo.

Se você tiver uma plataforma de multiprocessamento simétrico (SMP), você poderá usar várias CPUs para processar simultaneamente dados de sessão ou dados de partições. Isso fornecerá um melhor desempenho devido ao real paralelismo atingido. Em uma plataforma com um único processador, essas tarefas compartilham a CPU, portanto, não há paralelismo.

O Serviço de Integração pode usar várias CPUs para processar uma sessão que contenha várias partições. O número de CPUs usadas depende de fatores como o número de partições, o número de segmentos, o número de CPUs disponíveis e a quantidade ou os recursos necessários para processar o mapeamento.

Otimizando o banco de dados de origem para particionamento

Você pode adicionar partições para aumentar a velocidade da consulta. Em geral, cada partição no lado do leitor representa um subconjunto dos dados a serem processados.

Execute as tarefas a seguir para otimizar o banco de dados de origem para particionamento,

- **Ajuste o banco de dados.** Se o banco de dados não for ajustado apropriadamente, a criação de partições poderá não acelerar as sessões.
- **Habilite consultas paralelas.** Alguns bancos de dados podem ter opções que precisem ser configuradas para habilitar consultas paralelas. Consulte a documentação do banco de dados para obter informações sobre essas opções. Se essas opções estiverem desativadas, o Serviço de Integração executará várias instruções SELECT de partição serialmente.
- **Separe os dados em espaços de tabela diferentes.** Cada banco de dados fornece uma opção para separar os dados em diferentes espaços de tabela. Se o banco de dados permitir, use o recurso de substituição SQL do PowerCenter para fornecer uma consulta que extraia dados de uma única partição.
- **Agrupe os dados classificados.** Você pode particionar e agrupar os dados de origem para melhorar o desempenho de uma transformação de Unificador classificada.
- **Maximize consultas de classificação única.**

Ajustando o banco de dados

Se o banco de dados não for ajustado apropriadamente, os resultados poderão não acelerar as sessões. Você pode testar o banco de dados para verificar se ele está ajustado adequadamente.

Para verificar se o banco de dados está ajustado adequadamente:

1. Crie um pipeline com uma partição.
2. Dimensione a taxa de transferência do leitor no Workflow Monitor.
3. Adicione as partições.
4. Verifique se a taxa de transferência encontra-se em escala linear.

Por exemplo, se a sessão tem duas partições, a taxa de transferência do leitor deve ser duas vezes mais rápida. Se a taxa de transferência não estiver em uma escala linear, você provavelmente precisará ajustar o banco de dados.

Agrupando dados classificados

Você também pode particionar e agrupar os dados de origem para melhorar o desempenho da transformação de Unificador classificada. Posicione o ponto de partição antes da transformação do Classificador para manter o agrupamento e a classificação dos dados dentro de cada grupo.

Para agrupar dados, verifique se as linhas com o mesmo valor de chave estão roteadas para a mesma partição. A melhor maneira de garantir que os dados estejam agrupados e distribuídos uniformemente entre as partições é por meio da adição de um ponto de partição de chaves automáticas hash ou de limite de chaves antes da origem de classificação.

Otimizando consultas de classificação única

Para otimizar uma consulta de classificação única no banco de dados, considere as seguintes opções de ajuste que permitem o paralelismo:

- **Verifique os parâmetros de configuração que realizam o ajuste automático.** Por exemplo, a Oracle tem um parâmetro chamado `parallel_automatic_tuning`.
- **Verifique se o intraparelismo está habilitado.** Intraparelismo é a capacidade de executar vários segmentos em uma única consulta. Por exemplo, na Oracle, consulte `parallel_adaptive_multi_user`. No DB2, consulte `intra_parallel`.
- **Verifique o número máximo de processos paralelos disponíveis para execuções paralelas.** Por exemplo, na Oracle, consulte `parallel_max_servers`. No DB2, consulte `max_agents`.
- **Verifique os tamanhos dos vários recursos usados no paralelismo.** Por exemplo, a Oracle tem parâmetros como `large_pool_size`, `shared_pool_size`, `hash_area_size`, `parallel_execution_message_size` e `optimizer_percent_parallel`. O DB2 tem parâmetros de configuração como `dft_fetch_size`, `fcm_num_buffers` e `sort_heap`.
- **Verifique os graus de paralelismo.** É possível que você consiga definir essa opção usando o parâmetro de configuração de um banco de dados ou uma opção na tabela ou consulta. Por exemplo, a Oracle tem os parâmetros `parallel_threads_per_cpu` e `optimizer_percent_parallel`. O DB2 tem parâmetros de configuração como `dft_prefetch_size`, `dft_degree` e `max_query_degree`.
- **Desative as opções que possam afetar a escalabilidade do banco de dados.** Por exemplo, desabilite os logs de arquivo morto e as estatísticas limitadas na Oracle.

Para obter uma lista completa das opções de ajuste de banco de dados, consulte a documentação do banco de dados.

Otimizando o banco de dados de destino para particionamento

Se uma sessão contém várias partições, a taxa de transferência de cada partição deve ser a mesma que a taxa de transferência de uma sessão de única partição. Se você não perceber essa correlação, então o banco de dados provavelmente estará inserindo linhas no banco de dados serialmente.

Para garantir que o banco de dados insira linhas em paralelo, verifique as seguintes opções de configuração no banco de dados de destino:

- **Defina as opções no banco de dados para habilitar inserções paralelas.** Por exemplo, defina as opções `db_writer_processes` e `DB2 has max_agents` em um banco de dados Oracle para habilitar inserções paralelas. Alguns bancos de dados podem habilitar essas opções por padrão.
- **Considere particionar a tabela de destino.** Se possível, tente fazer com que cada partição grave em uma partição de banco de dados único usando uma transformação de Roteador para fazer isso. Além disso, disponha as partições de banco de dados em discos separados para impedir a contenção de E/S entre as partições de pipeline.
- **Defina as opções no banco de dados para aprimorar a escalabilidade do banco de dados.** Por exemplo, desabilite os logs de arquivo morto e as estatísticas limitadas nos bancos de dados Oracle para aprimorar a escalabilidade.

APÊNDICE A

Contadores de desempenho

Este apêndice inclui os seguintes tópicos:

- [Visão geral do capítulo Contadores de desempenho, 70](#)
- [Contador Errorrows, 70](#)
- [Contadores Readfromcache e Writetocache, 71](#)
- [Contadores Readfromdisk e Writetodisk, 71](#)
- [Contador Rowsinlookupcache, 72](#)

Visão geral do capítulo Contadores de desempenho

Todas as transformações têm contadores. O Serviço de Integração rastreia o número de linhas de entrada, de saída e de erro para cada transformação. Algumas transformações têm contadores de desempenho. É possível usar os seguintes contadores de desempenho para aumentar o desempenho da sessão:

- Errorrows
- Readfromcache e Writetocache
- Readfromdisk e Writetodisk
- Rowsinlookupcache

Contador Errorrows

Os erros de transformação afetam o desempenho da sessão. Se uma transformação tem um grande número de linhas de erro em qualquer um dos contadores *Transformation_errorrows*, é possível eliminar os erros para aprimorar o desempenho.

TÓPICOS RELACIONADOS:

- [“Eliminando erros de transformação” na página 42](#)

Contadores Readfromcache e Writetocache

Se uma sessão tiver as transformações Agregador, Classificação ou Unificador, examine os contadores *Transformation_readfromcache*, *Transformation_writetocache*, *Transformation_readfromdisk* e *Transformation_writetodisk* para analisar como o Serviço de Integração lê ou grava dados no disco. Para visualizar os detalhes do desempenho da sessão durante a execução, clique com o botão direito do mouse na sessão no Workflow Monitor e escolha Propriedades. Clique na guia Propriedades na caixa de diálogo de detalhes.

Para analisar o acesso ao disco, calcule primeiro o índice de ocorrências ou perdas. O índice de ocorrências indica o número de operações de leitura ou gravação que Serviço de Integração executa no cache.

O índice de perdas indica o número de operações de leitura ou gravação que Serviço de Integração executa no disco.

Use a seguinte fórmula para calcular o índice de perdas do cache:

$$\frac{[(\# \text{ of reads from disk}) + (\# \text{ of writes to disk})]}{[(\# \text{ of reads from memory cache}) + (\# \text{ of writes to memory cache})]}$$

Use a seguinte fórmula para calcular o índice de ocorrências do cache:

$$[1 - \text{Cache Miss ratio}]$$

Para minimizar as leituras e as gravações no disco, aumente o tamanho do cache. O índice de ocorrências do cache ideal é 1.

Contadores Readfromdisk e Writetodisk

Se uma sessão tiver as transformações Agregador, Classificação ou Unificador, examine cada contador *Transformation_readfromdisk* e *Transformation_writetodisk*. Para visualizar os detalhes do desempenho da sessão durante a execução, clique com o botão direito do mouse na sessão no Workflow Monitor e escolha Propriedades. Clique na guia Propriedades na caixa de diálogo de detalhes.

Se esses contadores exibirem qualquer número diferente de zero, será possível aumentar os tamanhos dos caches para aprimorar o desempenho da sessão. O Serviço de Integração usa o cache de índice para armazenar informações do grupo e o cache de dados para armazenar dados transformados, que são normalmente maiores. Embora os tamanhos dos caches de dados e de índice afetem o desempenho, poderá ser necessário aumentar mais o tamanho do cache de dados do que o tamanho do cache de índice. No entanto, se o volume de dados processado é maior do que a memória disponível, você pode aumentar o tamanho do cache de índice para aprimorar o desempenho.

Por exemplo, o Serviço de Integração usa 100 MB para armazenar o cache de índice e 500 MB para armazenar o cache de dados. Com 200 acessos distribuídos aleatoriamente em cada cache de índice e de dados, você pode configurar o cache nas seguintes maneiras:

- Para otimizar o desempenho, aloque 100 MB para o cache de índice e 200 MB para o cache de dados. O Serviço de Integração acessa 100% dos dados do cache de índice e 40% dos dados do cache de dados. O Serviço de Integração sempre acessa o cache de índice e não acessa o cache de dados 120 vezes. Portanto, a porcentagem de dados que são acessados é 70%.
- Aloque 50 MB para o cache de índice e 250 MB para o cache de dados. O Serviço de Integração acessa 50% dos dados do cache de índice e 50% dos dados do cache de dados. O Serviço de Integração não acessa os caches de índice e de dados 100 vezes cada. Portanto, a porcentagem de dados que são acessados é 50%.

Se a sessão executa agregação incremental, o Serviço de Integração lê os dados de agregação históricos do disco local durante a sessão e grava no disco ao salvar os dados históricos. Como resultado, os contadores `Aggregator_readtodisk` e `Aggregator_writetodisk` exibem números além de zero.

No entanto, desde que o Serviço de Integração grave os dados históricos em um arquivo no final da sessão, você ainda pode avaliar os contadores durante a sessão. Se os contadores exibem números diferentes de zero durante a execução da sessão, é possível ajustar os tamanhos dos caches para aprimorar o desempenho. No entanto, existe um custo associado a alocação e desalocação de memória. Por isso, se você souber qual o volume de dados que o Serviço de Integração processará, retenha o aumento dos tamanhos dos cache para acomodar mais volumes de dados.

Contador Rowsinlookupcache

Pesquisas múltiplas podem diminuir o desempenho da sessão. Para aprimorar o desempenho da sessão, ajuste as expressões de pesquisa para as tabelas de pesquisa maiores.

TÓPICOS RELACIONADOS:

- [“Otimizando várias pesquisas” na página 38](#)

ÍNDICE

A

- afunilamentos
 - destinos [14](#)
 - eliminando [12](#)
 - estatísticas de segmento [13](#)
 - identificando [12](#)
 - mapeamentos [16](#)
 - no UNIX [18](#)
 - no Windows [18](#)
 - origens [15](#)
 - sessões [17](#)
 - sistema [17](#)
- agregação incremental
 - otimizando transformação de Agregador [34](#)
- agrupar por portas
 - otimizando transformação de Agregador [33](#)
- ajustando
 - caches [46](#)
 - expressões [30](#)
 - mapeamentos [27](#)
 - origens relacionais [24](#)
 - rede [63](#)
 - Repositório do PowerCenter [57](#)
 - Serviço de Integração [60](#)
 - sessões [43](#)
 - sistema [62](#)
 - Transformação de Agregador [33](#)
 - Transformação de Classificador [39](#)
 - Transformação de Gerador de Sequência [39](#)
 - Transformação de pesquisa [35](#)
 - Transformação de Qualificador de Fonte [41](#)
 - Transformação de SQL [41](#)
 - Transformação de unificador [35](#)
 - Transformação do Normalizador [39](#)
 - Transformação personalizada [34](#)
 - Transformação XML [41](#)
 - transformações [33](#)
- alta disponibilidade
 - sistemas de arquivo clusterizados [52](#)
- armazenamento de arquivo ideal
 - arquivos de cache não persistentes [51](#)
 - arquivos de destino [51](#)
 - arquivos de log [51](#)
 - arquivos de origem [51](#)
 - arquivos de parâmetro [51](#)
 - arquivos temporários [51](#)
- armazenamento de arquivos
 - local [50](#)
 - sistema de arquivos compartilhado [50](#)
 - tipos [50](#)
- armazenando em buffer
 - dados [32](#)
- arquivo XML
 - comparado com arquivo simples [28](#)

- arquivos de cache
 - armazenamento ideal [51](#)
- arquivos de cache persistentes
 - armazenamento ideal [51](#)
 - diretrizes de configuração [53](#)
- arquivos de destino
 - armazenamento ideal [51](#)
- arquivos de log
 - armazenamento ideal [51](#)
- arquivos de log de fluxo de trabalho
 - desabilitando [48](#)
- arquivos de log de sessão
 - desabilitando [48](#)
- arquivos de origem
 - armazenamento ideal [51](#)
 - simples versus XML [28](#)
- arquivos de parâmetro
 - armazenamento ideal [51](#)
 - diretrizes de desempenho [55](#)
- arquivos simples
 - arquivos de origem delimitados [28](#)
 - comparado com arquivos XML [28](#)
 - comprimento do buffer [27](#)
 - localização de armazenamento ideal [63](#)
 - otimizando origens [27](#)
- arquivos simples delimitados
 - origens [28](#)
- arquivos temporários
 - armazenamento ideal [51](#)
- associações
 - no banco de dados [35](#)
- associando
 - processador [64](#)
- avaliando
 - expressões [32](#)

B

- bancos de dados
 - ajustando consultas de classificação única [68](#)
 - ajustando destinos Oracle [22](#)
 - ajustando origens [24](#)
 - associações [35](#)
 - intervalos de ponto de verificação [21](#)
 - minimizando os deadlocks [22](#)
 - otimizando destinos para particionamento [69](#)
 - otimizando origens para particionamento [67](#)
 - tamanho do pacote de rede [22, 25](#)
- buffer DTM
 - tamanho ideal do pool [45](#)

C

- cache
 - ajustando [46](#)
 - localização ideal [46](#)
 - metadados do repositório [61](#)
 - reduzir linhas armazenadas em cache [37](#)
 - tamanho ideal [47](#)
 - valores de sequência [39](#)
- cache compartilhado
 - para pesquisas [36](#)
- cache de dados
 - localização ideal [46](#)
 - portas conectadas [46](#)
 - tamanho ideal [47](#)
- cache de índice
 - localização ideal [46](#)
 - tamanho ideal [47](#)
- cache não persistente
 - armazenamento ideal para arquivos [51](#)
- cache persistente
 - para pesquisas [36](#)
- carregador externo
 - desempenho [21](#)
- carregamento em massa
 - ajustando destinos relacionais [21](#)
- chamadas de função
 - minimizando para transformação personalizada [34](#)
- compartilhamento de arquivo
 - sistemas de arquivo de cluster [52](#)
 - sistemas de arquivo de rede [52](#)
- comprimento do buffer
 - configuração ideal [27](#)
- condição de pesquisa
 - correspondência [37](#)
 - otimizando [38](#)
- consulta de banco de dados
 - afunilamentos de origem, identificando [16](#)
- consultas
 - ajustando origens relacionais [24](#)
- consultas de objeto
 - ordenando condições [58](#)
- contadores de desempenho
 - Rowsinlookupcache [72](#)
 - tipos [70](#)
 - Transformation_errorrows [70](#)
 - Transformation_readfromcache [71](#)
 - Transformation_readfromdisk [71](#)
 - Transformation_writetocache [71](#)
 - Transformation_writetodisk [71](#)
- convertendo
 - tipos de dados [29](#)
- CPU
 - múltiplos para fluxos de trabalho simultâneos [63](#)
 - várias para particionamento de pipeline [67](#)

D

- DB2
 - desempenho do repositório do PowerCenter [58](#)
- deadlocks
 - minimizando [22](#)
- descartando
 - restrições de índices e chave [20](#)
- desempenho
 - ajuste, visão geral [11](#)
 - esquema do banco de dados do repositório, otimizando [58](#)

- desempenho ()
 - latência de liberação [48](#)
 - sessões em tempo real [48](#)
- destinos
 - afunilamentos, causas [14](#)
 - afunilamentos, eliminando [14](#)
 - alocando memória de buffer [44](#)
 - identificando afunilamentos [14](#)
- diretório de cache
 - compartilhando [46](#)
- diretórios
 - caches compartilhados [46](#)
- disco
 - acesso, minimizando [63](#)
- drivers de banco de dados
 - ideal para o Serviço de Integração [60](#)

E

- email pós-sessão
 - desempenho [49](#)
- entrada classificada
 - otimizando transformação de Agregador [34](#)
- erros
 - minimizando o nível de rastreamento [48](#)
- espaço de tabela
 - tipo ideal para DB2 [58](#)
- espaços
 - à direita/esquerda, removendo [30](#)
- estatísticas de segmento
 - afunilamentos, eliminando [13](#)
 - afunilamentos, identificando [13](#)
- expressões
 - ajustando [30](#)
 - avaliando [32](#)
 - substituindo por variáveis locais [30](#)
- expressões IIF
 - ajustando [31](#)

F

- FastExport
 - para origens Teradata [25](#)
- fatoramento
 - lógica comum de mapeamento [30](#)
- filtrando
 - dados [29](#)
 - dados de origem [41](#)
- filtros
 - origens [25](#)
- fluxo de dados
 - monitoramento [70](#)
 - otimizando [70](#)
- fluxos de trabalho
 - simultâneos [44](#)
- função DECODE
 - comparada à função Pesquisa [31](#)
 - usando para otimização [31](#)
- função LOOKUP
 - comparada à função DECODE [31](#)
 - minimizando para otimização [31](#)
- funções
 - comparadas a operadores [31](#)
 - DECODE versus LOOKUP [31](#)
- funções de agregação
 - minimizando chamadas [30](#)

G

grade

- afunilamento de nó [52](#)
- desempenho [43](#), [50](#)
- desempenho do Gerador de Sequência, aumentando [56](#)
- localizações ideais de armazenamento [50](#)

I

IBM DB2

- esquema do banco de dados do repositório, otimizando [58](#)
- intervalo de confirmação
- desempenho da sessão [47](#)
- intervalos de ponto de verificação
- aumentando [21](#)

L

latência de liberação

- desempenho, melhorando [48](#)

leitura de passagem única

- definição [28](#)

M

mapeamento de passagem

- ajustando [29](#)

mapeamento de teste de leitura

- afunilamentos de origem, identificando [15](#)

mapeamentos

- afunilamentos, eliminando [16](#)
- afunilamentos, identificando [16](#)
- ajustando [27](#)
- leitura de passagem única [28](#)
- lógica comum da fatoramento [30](#)
- mapeamento de passagem, ajuste [29](#)

memória

- 64-bit PowerCenter [47](#)
- aumentando [63](#)
- bancos de dados Microsoft SQL Server [26](#)
- bancos de dados Sybase ASE [26](#)
- buffer [44](#)

memória de buffer

- alocando [44](#)

mesclagem sequencial

- armazenamento de arquivo ideal [51](#)

métodos

- filtrando dados [29](#)

Microsoft SQL Server

- banco de dados na memória [26](#)
- esquema do banco de dados do repositório, otimizando [58](#)

minimizando

- chamadas de função de agregação [30](#)

modo ASCII

- desempenho [60](#)

modo de movimentação de dados

- ideal [60](#)

N

níveis de rastreamento

- minimizando [48](#)

O

Opção Substituição SQL de Pesquisa

- reduzindo o tamanho do cache [37](#)

operações

- numéricas versus de string [30](#)

operações de string

- comparado com operações numéricas [30](#)
- minimizando [30](#)

operações numéricas

- comparado com operações de string [30](#)

operadores

- comparados a funções [31](#)

Oracle

- ajustando destinos [22](#)
- carregador externo [21](#)
- otimizando conexões [25](#)
- protocolo IPC [25](#)

ORDER BY

- otimizando para transformação de Pesquisa [37](#)

origens

- afunilamentos, causas [15](#)
- afunilamentos, eliminando [16](#)
- ajustando consultas [24](#)
- filtros [25](#)
- identificando afunilamentos [15](#)
- relacional, ajustando [24](#)

origens XML

- alocando memória de buffer [44](#)

otimização de empilhamento

- desempenho [44](#)

P

pacotes de rede

- aumentando [22](#), [25](#)

paginação

- reduzindo [63](#)

particionamento de pipeline

- adicionando partições [65](#)
- ajustando o banco de dados de origem [67](#)
- otimizando bancos de dados de destino [69](#)
- otimizando bancos de dados de origem [67](#)
- otimizando o desempenho [65](#)
- tipos de partição ideais [66](#)
- várias CPUs [67](#)

partições

- adicionando [65](#)

pipelines

- monitoramento de fluxo de dados [70](#)

portas

- conectadas, limitando [46](#)

processador

- associando [64](#)

processo do Serviço de Repositório

- localização ideal [58](#)

protocolo IPC

- origens Oracle [25](#)

R

rastreamento de erros

- Consulte níveis de rastreamento[error tracing
aaa] [48](#)

rede

- ajustando [63](#)

- rede (
 - aprimorando a velocidade [63](#))
- registro em log de arquivo simples
 - emails pós-sessão [49](#)
- removendo
 - espaços em branco à esquerda/direita [30](#)
- Repositório do PowerCenter
 - ajustando [57](#)
 - desempenho no DB2 [58](#)
 - localização ideal [58](#)
- repositórios
 - esquema do banco de dados, otimizando [58](#)
- restrições de chave
 - descartando [20](#)

S

- segmento de transformação
 - tempo de trabalho do segmento [13](#)
- segmentos
 - afunilamentos, identificando [13](#)
 - tempo de execução [13](#)
 - tempo de ocupação [13](#)
 - tempo de trabalho do segmento [13](#)
 - tempo ocioso [13](#)
- selecionar distintos
 - filtrando dados de origem [41](#)
- Serviço de Integração
 - ajustando [60](#)
 - drivers de banco de dados ideais [60](#)
 - grade [43](#)
 - intervalo de confirmação [47](#)
- Serviço de Repositório
 - criando cache de metadados do repositório [61](#)
- sessão em grade
 - desempenho do Gerador de Sequência, aumentando [56](#)
- sessões
 - afunilamentos, causas [17](#)
 - afunilamentos, eliminando [17](#)
 - afunilamentos, identificando [17](#)
 - ajustando [43](#)
 - grade [43](#)
 - otimização de empilhamento [44](#)
 - simultâneos [44](#)
- sessões em tempo real
 - desempenho, melhorando [48](#)
- sistema
 - afunilamento, identificando com o Workflow Monitor [17](#)
 - afunilamentos no UNIX, identificando [18](#)
 - afunilamentos no Windows, identificando [18](#)
 - afunilamentos, causas [17](#)
 - afunilamentos, eliminando [18](#)
 - ajustando [62](#)
- sistemas de arquivo
 - cluster [52](#)
 - compartilhado, configurando [52](#)
 - rede [52](#)
- sistemas de arquivo clusterizados
 - alta disponibilidade [52](#)
 - Consulte também os sistemas de arquivos compartilhados[sistemas de arquivos clusterizados aaa] [52](#)
- sistemas de arquivo de rede
 - Consulte os sistemas de arquivos compartilhados[sistemas de arquivo de rede aaa] [52](#)

- sistemas de arquivos compartilhados
 - alta largura de banda [51](#)
 - baixa largura de banda [51](#)
 - carga do servidor, distribuindo [52](#)
 - configurando [52](#)
 - CPU, balanceamento [52](#)
 - visão geral [52](#)
- Sybase ASE
 - banco de dados na memória [26](#)
- Sybase IQ
 - carregador externo [21](#)

T

- tamanho da página
 - mínimo para a otimização do esquema do banco de dados do repositório [58](#)
- tamanho do bloco de buffer
 - ideal [45](#)
- tempo de execução
 - estatística de segmento [13](#)
- tempo de ocupação
 - estatística de segmento [13](#)
- tempo ocioso
 - estatística de segmento [13](#)
- Teradata FastExport
 - desempenho para origens [25](#)
- tipos de dados
 - Char [30](#)
 - otimizando conversões [29](#)
 - Varchar [30](#)
- tipos de dados Char
 - removendo espaços em branco à direita/esquerda [30](#)
- tipos de dados Varchar
 - removendo espaços em branco à direita/esquerda [30](#)
- tipos de partição
 - ideal [66](#)
- Transformação de Agregador
 - agregação incremental [34](#)
 - ajustando [33](#)
 - detalhes de desempenho [71](#)
 - otimizando com conexões de porta limitadas [34](#)
 - otimizando com Entrada Classificada [34](#)
 - otimizando com filtros [34](#)
 - otimizando com grupo por portas [33](#)
- Transformação de classificação
 - detalhes de desempenho [71](#)
- Transformação de classificador
 - otimizando com alocação de memória [40](#)
- Transformação de Classificador
 - ajustando [39](#)
 - otimizando diretórios de partição [40](#)
- Transformação de Filtro
 - afunilamentos de origem, identificando [15](#)
- Transformação de Gerador de Sequência
 - ajustando [39](#)
 - desempenho da grade [56](#)
 - reutilizável [39](#)
- Transformação de pesquisa
 - ajustando [35](#)
 - otimizando [72](#)
 - otimizando a condição de pesquisa [38](#)
 - otimizando a correspondência de condição de pesquisa [37](#)
 - otimizando com caches [36](#)
 - otimizando com caches simultâneos [37](#)
 - otimizando com drivers de banco de dados [36](#)
 - otimizando com indexação [38](#)

- Transformação de pesquisa ()
 - otimizando com instrução ORDER BY [37](#)
 - otimizando com máquina de memória alta [38](#)
 - otimizando com redução de cache [37](#)
 - otimizando várias expressões de pesquisa [38](#)
- Transformação de procedimento externo
 - bloqueio de dados [32](#)
- Transformação de Qualificador de Fonte
 - ajustando [41](#)
- Transformação de SQL
 - ajustando [41](#)
- Transformação de unificador
 - ajustando [35](#)
 - criando uma origem mestre [35](#)
 - dados classificados [35](#)
 - detalhes de desempenho [71](#)
- Transformação do Normalizador
 - ajustando [39](#)
- Transformação personalizada
 - ajustando [34](#)
 - minimizando chamadas de função [34](#)
 - processando blocos de dados [34](#)
- Transformação XML
 - ajustando [41](#)
- transformações
 - ajustando [33](#)

- transformações ()
 - eliminando erros [42](#)
 - otimizando [70](#)

U

- UNIX
 - afunilamentos do sistema [18](#)
 - afunilamentos, eliminando [18](#)
 - associação de processador [64](#)

V

- variáveis locais
 - substituindo expressões [30](#)

W

- Windows
 - afunilamentos [18](#)
 - afunilamentos, eliminando [18](#)