

IBM® Analytical Decision Management
Versão 17 Release 0

*Guia de resolução de problemas e
desempenho*

IBM

Nota

Antes de utilizar estas informações e o produto suportado por elas, leia as informações gerais nos “Avisos” na página 29.

Informações sobre o produto

Essa edição se aplica à versão 17, liberação 0, modificação 0 do IBM Analytical Decision Management and to all subsequent releases and modifications until otherwise indicated in new editions.

© Copyright IBM Corporation 2010, 2015.

Índice

Prefácio v

Capítulo 1. Tópicos básicos 1

Técnicas para a Resolução de Problemas 1
Procurando em Bases de Conhecimento 2
Obtendo correções 3
Entrando em contato com o Suporte IBM 4
Trocando informações com IBM 5
 Enviando Informações para o Suporte IBM 5
 Recebendo informações do suporte de IBM 5
Inscrevendo-se para obter atualizações do suporte 6

Capítulo 2. Resolução geral de problemas 9

Resolução de problemas com lista de verificação para IBM Analytical Decision Management 9
 Coletando dados para IBM Analytical Decision Management 9

Capítulo 3. Resolução de problemas do IBM Analytical Decision Management 11

Falta de memória 11
Corrigindo a saída incompleta do gráfico de Ganhos 11
Gerenciando versões de fluxos 12
Verificando a conectividade do IBM SPSS Modeler Server 12
Regras de segmento 12
Modelos de CPLEX. 12

Capítulo 4. Melhores práticas de desempenho 17

Número de registros de origem. 17
Número de campos no modelo de dados do projeto 18
Fazendo upload de dados 18
Dimensões (hierarquias) e os membros de dimensão 18
Otimizando a geração de fluxo 19
SQL pushback 19
 Atributos derivados de modelo e SQL pushback 20
Alocação estática versus dinâmica 23
Fazendo referência a modelos 24
Fazendo referência a regras externas 25
Otimização do CPLEX. 25
 Tamanho do problema. 26
 Modelagem de problema de negócios na OPL 26
 Complexidade do modelo de OPL. 26
 Dados 27
 Algoritmo 27
 Recursos (memória e CPU) disponíveis 27
Acesso ao banco de dados ao usar o Serviço de escoragem 28

Avisos 29

Marcas Comerciais 30

Índice Remissivo 33

Prefácio

Sobre o IBM Business Analytics

O software IBM Business Analytics entrega informações precisas e consistem nas quais os tomadores de decisão confiam para melhorar o desempenho dos negócios. Um portfólio abrangente de inteligência de negócios, análise preditiva, gerenciamento de desempenho financeiro e estratégia e aplicativos analíticos fornecem insight claro, imediato e prático sobre o desempenho atual e a capacidade de prever resultados futuros. Combinado com soluções valiosas para segmentos do mercado, práticas comprovadas e serviços profissionais, as organizações de todos os tamanhos podem alcançar a maior produtividade, automatizar com confiança decisões e fornecer melhores resultados.

Como parte deste dossier, o software IBM SPSS Predictive Analytics ajuda as organizações a prever futuros eventos e agir proativamente com esse insight para melhores resultados de negócios. Clientes comerciais, governamentais e acadêmicos em todo o mundo confiam na tecnologia do IBM® SPSS como uma vantagem competitiva na atração, retenção e aumento dos clientes, reduzindo a fraude e mitigando os riscos. Ao incorporar o software IBM SPSS em suas operações diárias, as organizações se tornam empresas preditivas – capazes de dirigir e automatizar decisões para atender às metas de negócio e obter vantagem competitiva mensurável. Para obter informações adicionais ou para entrar em contato com um representante, visite <http://www.ibm.com/spss>.

Suporte técnico

O suporte técnico está disponível para manutenção dos clientes. Os clientes podem entrar em contato com o Suporte Técnico para obter assistência no uso de produtos IBM Corp. ou para obter ajuda na instalação em um dos ambientes de hardware suportados. Para entrar em contato o Suporte Técnico, consulte o website em IBM Corp. <http://www.ibm.com/support>. Esteja preparado para se identificar, sua organização e seu contrato de suporte ao solicitar assistência.

Capítulo 1. Tópicos básicos

Técnicas para a Resolução de Problemas

Resolução de problemas é uma abordagem sistemática para resolver um problema. O objetivo da resolução de problemas é determinar o motivo de algo não funcionar conforme o esperado e como resolver o problema.

A primeira etapa no processo de resolução de problemas é descrever o problema totalmente. As descrições do problema ajudam você e o representante de suporte técnico IBM a saber onde iniciar para localizar a causa do problema. Essa etapa inclui fazer algumas perguntas básicas:

- Quais são os sintomas do problema?
- Onde o problema ocorre?
- Quando o Problema Acontece?
- Sob Quais Condições o Problema Ocorre?
- É Possível Reproduzir o Problema?

As respostas a essas perguntas geralmente fornecem uma boa descrição do problema, o que pode levar você a uma resolução de problemas.

Quais são os sintomas do problema?

Ao começar a descrever um problema a pergunta mais óbvia é: Qual é o problema? Essa questão pode parecer direta; entretanto, é possível dividi-la em várias questões mais focadas que criam uma imagem mais descritiva do problema. Estas questões podem incluir:

- Quem ou o quê está relatando o problema?
- Quais são os códigos de erro e as mensagens?
- Como o Sistema Falha? Por exemplo, entra em loop, é interrompido, trava, tem uma redução de desempenho ou apresenta resultado incorreto?

Onde o problema ocorre?

Determinar onde o problema se origina nem sempre é fácil, mas é uma das etapas mais importantes na resolução de um problema. Muitas camadas de tecnologia podem existir entre a criação do relatório e a falha dos componentes. Redes, discos e drivers são apenas alguns dos componentes a serem considerados quando você está investigando problemas.

As questões a seguir o ajudam a se concentrar no local em que o problema ocorre para isolar a camada do problema:

- O problema é específico a uma plataforma ou sistema operacional ou é comum a diversas plataformas ou sistemas operacionais?
- O ambiente e a configuração atual são suportados?

O fato de uma camada relatar o problema não significa que o problema tem origem nessa camada. Parte da identificação da origem de um problema é entender o ambiente no qual ele existe. Não tenha pressa em descrever o ambiente com problema por inteiro, incluindo o sistema operacional e versão, todos os produtos de software e versões correspondentes e informações de hardware. Confirme se a execução está ocorrendo dentro de um ambiente que possui uma configuração suportada; muitos problemas podem ser investigados nos níveis incompatíveis de software que não deveriam ser executados juntos ou não foram totalmente testados juntos.

Quando o Problema Acontece?

Desenvolva uma linha de tempo detalhada de eventos que levam a uma falha, especialmente nos casos de ocorrências únicas. A maneira mais fácil de desenvolver uma linha de tempo é começar pelo fim: Comece pelo momento em que o erro foi reportado (sendo o mais preciso possível, incluindo milissegundos) e continue pelos logs e informações disponíveis. Geralmente, é necessário observar apenas tão longe quanto o primeiro evento suspeito que localizar em um log de diagnóstico.

Para desenvolver uma linha de tempo detalhada de eventos, responda a estas perguntas:

- O problema ocorre apenas em um determinado horário do dia ou da noite?
- Com que frequência ocorre o problema?
- Qual sequência de eventos leva ao momento em que o problema foi relatado?
- O problema ocorre após uma mudança no ambiente, como por exemplo, upgrade ou instalação de software ou hardware?

Responder a esses tipos de perguntas pode dar a você um quadro de referência no qual investigar o problema.

Sob Quais Condições o Problema Ocorre?

Saber quais sistemas e aplicativos estão em execução no momento em que um problema ocorre é uma parte importante da resolução de problemas. Essas questões sobre seu ambiente podem ajudar a identificar a causa raiz do problema:

- O problema sempre ocorre ao executar a mesma tarefa?
- Uma determinada sequência de eventos precisa ocorrer para que surja o problema?
- Algum outro aplicativo falha ao mesmo tempo?

Responder a esse tipo de pergunta pode ajudá-lo a explicar o ambiente no qual o problema ocorre e correlacionar quaisquer dependências. Lembre-se de que apenas porque vários problemas ocorreram ao mesmo tempo, eles não estão, necessariamente, relacionados.

É Possível Reproduzir o Problema?

Do ponto de vista de uma resolução de problemas, o problema ideal é aquele que pode ser reproduzido. Geralmente, quando um problema pode ser reproduzido, tem-se um conjunto maior de ferramentas ou procedimentos à sua disposição para ajudar a investigar. Consequentemente, os problemas que podem ser reproduzidos são com frequência mais fáceis de depurar e resolver. No entanto, os problemas que podem ser reproduzidos podem ter uma desvantagem: se o problema tiver um impacto comercial significativo, você não vai querer passar por ele novamente. Se possível, recrie o problema em um ambiente de teste ou desenvolvimento, que normalmente oferece mais flexibilidade e controle durante sua investigação.

- O problema pode ser recriado em um sistema de teste?
- Os diversos usuários ou aplicativos estão encontrando o mesmo tipo de problema?
- O problema pode ser recriado ao executar um comando único, um conjunto de comandos ou um determinado aplicativo?

Procurando em Bases de Conhecimento

Em geral é possível localizar soluções para problemas procurando bases de conhecimento do IBM . É possível otimizar resultados utilizando os recursos, ferramentas de suporte e métodos de procura disponíveis.

Sobre essa Tarefa

É possível localizar informações úteis ao procurar no centro de informações, mas às vezes é necessário procurar além do centro de informações para responder às suas questões ou para resolver problemas.

Procedimento

Para procurar nas bases de conhecimento informações necessárias, use uma ou mais das seguintes abordagens:

- Localize o conteúdo necessário ao usar o Portal de Suporte IBM. Inclua **Decision Management SPSS** na sua lista de produtos, juntamente com quaisquer outros produtos IBM que você possui.

O Portal de Suporte IBM é uma visualização unificada e centralizada de todas as ferramentas e informações de suporte técnico para todos os sistemas, software e serviços da IBM . O Portal de Suporte IBM permite que você acesse o portfólio de suporte eletrônico IBM a partir de um local. É possível padronizar as páginas para focar nas informações e recursos necessários para prevenção de problemas e uma resolução de problemas mais rápida. Familiarize-se com o IBM Support Portal, visualizando o vídeos demos sobre esta ferramenta. Estes vídeos o introduzem ao Portal de Suporte IBM , exploram resolução de problemas e outros recursos, além de demonstrar como é possível ajustar a página ao mover, incluir e excluir portlets.

- Procure por conteúdo sobre o IBM Analytical Decision Management usando um dos recursos técnicos adicionais a seguir:
 - IBM Analytical Decision Management notas técnicas e APARs (relatórios de problemas)
 - IBM Analytical Decision Management Website do suporte
 - IBM Redbooks®
- Procure conteúdo usando a procura de cabeçalho principal da IBM . Você pode usar a procura do cabeçalho principal do IBM , basta digitar sua sequência de caracteres de busca no campo Procura no topo de qualquer página ibm.com® .
- Procure conteúdo usando qualquer mecanismo de procura externo, como Google, Yahoo ou Bing. Se você usar um mecanismo de procura externo, é mais provável que seus resultados incluam informações que estão fora do domínio ibm.com. Entretanto, alguma vezes, é possível localizar informações de resolução de problemas úteis sobre produtos IBM em newsgroups, fóruns e blogs que não estão no ibm.com.

Dica: Inclua "IBM" e o nome do produto na sua procura se estiver procurando informações sobre um produto IBM .

Obtendo correções

Uma correção de produto pode estar disponível para resolver o seu problema.

Procedimento

Para localizar e instalar correções:

1. Determine de qual correção você precisa pela IBM Fix Central. Observe que, no momento da composição deste guia, nenhuma correção estava ainda disponível para IBM Analytical Decision Management.
 - a. Verifique as correções recomendadas para o IBM Analytical Decision Management para confirmar que seu software está no nível de manutenção mais recente.
 - b. Determine se o IBM já publicou uma correção individual para resolver seu problema. Verifique a lista de problemas que foram corrigidos na documentação *leia-me* do IBM Analytical Decision Management que está disponível para cada fix pack e pacote de atualizações listado.
2. Faça download da correção.
 - a. Abra o documento de download e siga o link na seção Pacote de download.

- b. Ao fazer download do arquivo, Assegure-se de que o nome do arquivo não seja alterado. Mudanças podem ser intencionais ou podem ser mudanças inadvertidas que são causadas por determinados navegadores da web ou utilitários de download.
3. Aplique a correção.
 - a. Siga as instruções na seção Instruções de instalação do documento de download.
 - b. Assine para receber notificações de e-mail semanais sobre correções e outras informações do Suporte do IBM .

Entrando em contato com o Suporte IBM

O Suporte IBM fornece assistência para defeitos do produto, respostas para perguntas mais frequentes e execução de redescoberta.

Antes de iniciar

Após tentar localizar sua resposta ou solução usando outras opções de autoatendimento, tais como as notas técnicas, é possível entrar em contato com o Suporte IBM . Antes de entrar em contato com o Suporte IBM , sua empresa deve ter um contrato de assinatura e suporte de software IBM ativo, e você deve estar autorizado a enviar problemas para a IBM . Para obter informações sobre os tipos de suporte disponível, consulte o tópico Portfólio de suporte no *Manual de suporte de software*.

Procedimento

Conclua as etapas a seguir para entrar em contato com o Suporte IBM sobre um problema:

1. Defina o problema, reúna informações de segundo plano e determine a gravidade do problema. Para obter ajuda, consulte o tópico Obtendo Suporte IBM® no *Manual de suporte de software*.
2. Reúna informações de diagnóstico. Consulte o tópico “Coletando dados para IBM Analytical Decision Management” na página 9 para obter mais informações.
3. Envie seu problema para o Suporte IBM de uma das seguintes maneiras:
 - Online, por meio do Portal de Suporte IBM: é possível abrir, atualizar ou visualizar todas as suas Solicitações de serviço no portlet Service Request na página Solicitação de serviço.
 - Por telefone: para obter o número de telefone que deve ser chamada na sua região, consulte a página da web Diretório de contratos mundiais.

Resultados

Se o problema enviado for sobre um defeito de software ou sobre documentação ausente ou imprecisa, o Suporte IBM cria um Authorized Program Analysis Report (APAR). Nesse APAR, o problema é descrito com detalhes. Sempre que possível, o Suporte IBM fornece uma solução alternativa que você pode implementar até que o APAR seja resolvido e uma correção seja fornecida. A IBM publica APARs resolvidos no website do Suporte IBM , assim, os usuários que experimentam o mesmo problema podem se beneficiar da mesma resolução.

Esteja preparado para trabalhar com o representante do suporte técnico da IBM usando a Assistência no Site da IBM , que é um plug-in de assistência remota que você pode transferir por download para o seu computador. O representante do suporte técnico da IBM pode usar a Assistência no Site da IBM para visualizar seu desktop e compartilhar o controle do seu mouse e teclado. Esta ferramenta pode diminuir o tempo gasto para identificar o problema, coletar os dados necessários e resolver o problema. Para obter mais informações, consulte Assistência no Site da IBM.

Trocando informações com IBM

Para diagnosticar ou identificar um problema, pode ser necessário fornecer ao Suporte do IBM dados e informações do seu sistema. Em outros casos, o Suporte do IBM pode fornecer ferramentas ou utilitários para uso para determinação de problemas.

Enviando Informações para o Suporte IBM

Para reduzir o tempo gasto para resolver o seu problema, você pode enviar informações de rastreamento e diagnóstico para o SuporteIBM .

Procedimento

Para enviar informações de diagnóstico para o Suporte IBM:

1. Abra um Problem Management Record (PMR). É possível utilizar a Ferramenta de solicitação de serviço.
2. Colete os dados diagnósticos necessários. Os dados diagnósticos ajudam a reduzir o tempo gasto para resolver seu PMR.
3. Compacte os arquivos usando formato ZIP ou TAR.
4. Transfira os arquivos para a IBM . É possível usar um dos seguintes métodos:
 - A ferramenta de solicitação de serviço
 - Métodos de upload de dados padrão: FTP, HTTP
 - Métodos de upload de dados seguros: FTPS, SFTP, HTTPS
 - Email

Todos esses métodos de troca de dados são explicados no Website do Suporte IBM.

Recebendo informações do suporte de IBM

Ocasionalmente, um representante do suporte técnico do IBM irá solicitar que você faça download das ferramentas de diagnóstico ou outros arquivos. É possível usar o FTP para fazer download desses arquivos.

Antes de iniciar

Assegure-se de que seu representante do suporte técnico do IBM forneceu o servidor preferencial que deve ser usado para download de arquivos e o diretório e os nomes de arquivo exatos que devem ser acessados.

Procedimento

Para fazer download de arquivos do Suporte do IBM :

1. Use o FTP para se conectar ao site que seu representante do suporte técnico do IBM forneceu e efetue login como **anônimo**. Use seu endereço de e-mail como a senha.
2. Acesse o diretório apropriado:
 - a. Acesse o diretório /fromibm.
`cd fromibm`
 - b. Acesse o diretório que seu representante do suporte técnico do IBM forneceu.
`cd nameofdirectory`
3. Ative o modo binário para a sessão.
`binary`
4. Use o comando `get` para fazer download do arquivo que seu representante do suporte técnico IBM especificou.

get filename.extension

5. Termine a sessão do FTP.

quit

Inscrevendo-se para obter atualizações do suporte

Para manter-se informado sobre informações importantes sobre os produtos IBM que você usa, é possível assinar as atualizações.

Sobre essa Tarefa

Ao assinar para receber atualizações, é possível receber informações e atualizações técnicas importantes para ferramentas e recursos de Suporte IBM . É possível inscrever-se nas atualizações usando uma das duas abordagens:

Assinaturas de Feeds RSS e de mídia social

Os feeds RSS e assinaturas de mídia social estão disponíveis para o IBM Analytical Decision Management:

- Centro de conhecimento do IBM Analytical Decision Management, o qual possui um feed RSS que você pode usar para se manter informado sobre as atualizações das informações

Para obter informações gerais sobre RSS, incluindo etapas para introdução e uma lista de páginas da Web com RSS ativado do IBM , visite o site www.ibm.com/software/support/rss/ .

My Notifications

Com My Notifications, é possível assinar atualizações de Suporte para qualquer produto IBM . (My Notifications, substitui My Support, que é uma ferramenta semelhante que você pode ter usado no passado). Com My Notifications, é possível especificar se você deseja receber comunicados por email diários ou semanais. É possível especificar que tipo de informação você deseja receber (como publicações, dicas e sugestões, atualizações do produto (também conhecidos como alertas), downloads e drivers). O My Notifications permite customizar e categorizar os produtos sobre os quais deseja ser informado e os métodos de entrega que melhor se adequam às suas necessidades.

Procedimento

Para assinar as atualizações do Suporte:

1. Assine o feed RSS do Centro de conhecimento do IBM Analytical Decision Management, o qual permite que você se mantenha informado sobre as atualizações regulares das informações. Para configurar o feed RSS, acesse a página de configuração feed RSS do Centro de informações.
2. Assine o feed RSS do IBM Analytical Decision Management, que permite que você permaneça atualizado com o conteúdo de Suporte de Software IBM mais recente para o produto. Para assinar, acesse a página de feeds RSS do IBM Analytical Decision Management.
3. Para assinar o My Notifications, acesse o IBM Support Portal e clique em **Gerenciar todas as minhas assinaturas** no portlet Notifications.
4. Conecte-se usando seu ID e senha IBM e clique em **Submit**.
5. Identifique o quê e como deseja receber atualizações.
 - a. Clique na guia **Subscribe**.
 - b. Clique em **Information Management** ou outra marca de software ou tipo de hardware.
 - c. Clique em **SPSS Decision Management** e selecione os nomes de quaisquer produtos adicionais que você usa. Clique **Continuar**.

- d. Selecione suas preferências de como receber atualizações, seja por e-mail, on-line em uma pasta designada ou como um feed RSS ou Atom.
- e. Selecione os tipos de atualizações de documentação que deseja receber, por exemplo, novas informações sobre downloads de produto e discussão sobre comentários do grupo.
- f. Clique em **Submit**.

Resultados

Até modificar suas preferências de feeds RSS e My Notifications, você receberá notificações de atualizações que foram solicitadas. É possível modificar suas preferências quando necessário (por exemplo, se você parar de usar um produto e começar a usar outro).

Informações Relacionadas

Feeds RSS do Suporte de Software IBM

Assine as atualizações de conteúdo do suporte My Notifications

Página de sing in do My notifications for IBM technical support

Tour de visão geral do My notifications for IBM technical support

Capítulo 2. Resolução geral de problemas

Resolução de problemas com lista de verificação para IBM Analytical Decision Management

Os pré-requisitos necessários estão instalados e em execução?

O IBM Analytical Decision Management requer o software obrigatório a seguir. Quando encontrar problemas, verifique se os pré-requisitos a seguir estão instalados corretamente. Consulte a documentação de instalação para obter informações mais detalhadas sobre requisitos do sistema.

- IBM SPSS Collaboration and Deployment Services Server. Consulte a documentação do IBM SPSS Collaboration and Deployment Services para obter instruções de instalação e informações sobre resolução de problemas.
- IBM SPSS Modeler Server. É possível fazer a instalação em uma máquina diferente ou na mesma máquina que o IBM SPSS Collaboration and Deployment Services, no entanto, o IBM SPSS Modeler Server deve estar em execução no mesmo fuso horário que o servidor de aplicativos usado pelo IBM SPSS Collaboration and Deployment Services. Consulte a documentação do IBM SPSS Modeler para obter instruções de instalação.
- IBM SPSS Modeler Adapter. É fornecido com o IBM SPSS Modeler e deve ser instalado na mesma máquina em que o IBM SPSS Collaboration and Deployment Services está instalado. Consulte a documentação do IBM SPSS Modeler para obter instruções de instalação.

O IBM SPSS Collaboration and Deployment Services Server e o IBM SPSS Modeler Server devem estar em execução antes da ativação do IBM Analytical Decision Management. Aguarde alguns minutos depois de iniciá-los antes de ativar o produto.

Você configurou as credenciais e definição de servidor necessárias para IBM SPSS Modeler Server?

Antes de usar o IBM Analytical Decision Management, você deve usar o IBM SPSS Collaboration and Deployment Services Deployment Manager para configurar as credenciais e uma definição de servidor e para o seu IBM SPSS Modeler Server. Caso contrário, os usuários não poderão efetuar login no IBM Analytical Decision Management. Consulte a documentação de instalação do IBM Analytical Decision Management para obter instruções.

Você está usando uma URL customizada?

A URL padrão do IBM Analytical Decision Management ainda é a mesma (*http://hostname:port/DM*). Mas a partir da versão 6 do IBM SPSS Collaboration and Deployment Services, a URL base para ativação dos aplicativos baseados em navegador do IBM SPSS Collaboration and Deployment Services e do IBM Analytical Decision Management serão customizáveis.

Se tiver problemas para acessar a URL do IBM Analytical Decision Management ou qualquer URL do IBM SPSS Collaboration and Deployment Services, é possível que um administrador as tenha customizado. Entre em contato com seu administrador. Mais informações estão disponíveis no capítulo *Raízes de contexto do aplicativo* do *IBM SPSS Collaboration and Deployment Services Guia de instalação e configuração do repositório*.

Coletando dados para IBM Analytical Decision Management

Ao resolver problemas do IBM Analytical Decision Management e antes de entrar em contato com o Suporte IBM, os seguintes dados devem ser coletados.

- Anote as seguintes informações. Muitas destas informações podem ser coletadas usando o IBM SPSS Collaboration and Deployment Services Deployment Manager baseado em navegador (<http://hostname:port/security/login>). Depois de efetuar logon, clique em **Sobre** e, em seguida, clique em **Mostrar detalhes**. As versões de muitos componentes serão listadas. Você também pode usar os links na parte inferior da tela para fazer download de detalhes da versão, detalhes do sistema e logs.
 - Sistema operacional
 - Versão do IBM Analytical Decision Management
 - Versão do IBM SPSS Collaboration and Deployment Services
 - Versão do IBM SPSS Modeler Server
 - Servidor de aplicativos e versão
 - Banco de dados e versão
- Anote qual o aplicativo do IBM Analytical Decision Management estava sendo usado quando o problema ocorreu, juntamente com a origem de dados específica usada.
- Se a instalação do IBM Analytical Decision Management falhar, consulte os arquivos de log no diretório IBM SPSS Collaboration and Deployment Services \Server\log (por exemplo, *C:\Program Files\IBM\SPSS\Deployment\7.0\Server\log*).

Capítulo 3. Resolução de problemas do IBM Analytical Decision Management

Os tópicos nesta seção fornecem informações sobre resolução de problemas relacionados ao IBM Analytical Decision Management. Para informações sobre resolução de problemas relacionados ao IBM SPSS Collaboration and Deployment Services, consulte o *Guia de resolução de problemas do IBM SPSS Collaboration and Deployment Services*.

Falta de memória

Se você encontrar problemas de falta de memória, pode ser necessário aumentar a configuração de memória do servidor de aplicativos. Por exemplo, se a configuração padrão de memória do WebSphere Application Server 1024 MB não for grande o suficiente, aumentá-la para 2048 MB ou mais pode resolver o problema de falta de memória.

No WebSphere Application Server:

Abra o console administrativo do WebSphere Application Server e altere a definição **Tamanho máximo de heap** para um valor maior, como **2048**.

OU

Abra o arquivo *server.xml* do WebSphere (por exemplo, *C:\profiles\AppSrv01\config\cells\Config9601VM0Node01Cell\nodes\Config9601VM0Node01\servers\server1\server.xml*) e altere a configuração `maximumHeapSize` para um valor maior, como 2048.

No JBoss Application Server:

Abra o arquivo *\$JBOSS_EAP_HOME/bin/standalone.conf.bat* e altere a configuração de memória:

```
rem # JVM memory allocation pool parameters - modify as appropriate.  
set "JAVA_OPTS=-Xms1200M -Xmx1300M -XX:MaxPermSize=256M"
```

No Application Server Weblogic:

Abra o arquivo *\$domain_home/bin/setCDSEnv.sh* e altere a configuração de memória:

```
USER_MEM_ARGS="$USER_MEM_ARGS -Xms256m -Xmx1536m  
-XX:MaxPermSize=512m"
```

Corrigindo a saída incompleta do gráfico de Ganhos

Se você estiver usando o JRockit JVM, você pode ver a saída incompleta do gráfico de *Ganhos*. Ao executar uma tarefa que contém um arquivo de cenário do IBM SPSS Modeler com um Tipo de avaliação de *Ganhos*, o gráfico resultante *Ganhos* pode estar incompleto.

O JRockit produz resultados numéricos errados para expressões que incluem NaN. Para uma solução alternativa desses problemas, especifique `-XstrictFP` como um parâmetro para a JVM ou altere para uma JVM diferente.

Gerenciando versões de fluxos

Os projetos do IBM Analytical Decision Management são armazenados como fluxos do IBM SPSS Modeler no repositório. Cada vez que um fluxo é salvo no repositório, uma nova versão é criada. Fluxos com um grande número de versões podem executar mais lentamente que os fluxos com um menor número de versões. Recomendamos o gerenciamento do número de versões que existem no repositório pela exclusão de versões que não são mais necessárias. Se tiver um número muito grande de versões desnecessárias no repositório e estiver tendo problemas para fazer a limpeza delas, você pode entrar em contato com o representante do Suporte do IBM SPSS para obter um utilitário repositório que pode ajudar a automatizar o trabalho de limpeza.

Antes de limpar as versões antigas, é altamente recomendado fazer backup de seu repositório como uma precaução no caso de quaisquer arquivos necessários serem acidentalmente excluídos.

Verificando a conectividade do IBM SPSS Modeler Server

Quando o IBM Analytical Decision Management não pode se conectar ao IBM SPSS Modeler Server, uma mensagem de erro indefinida é exibida quando os usuários tentam efetuar login. Se encontrar esse problema, faça o seguinte:

- Verifique se a definição do servidor e a definição de credenciais para o IBM SPSS Modeler Server estão corretas no cliente do cliente IBM SPSS Collaboration and Deployment Services Deployment Manager e no IBM SPSS Collaboration and Deployment Services Deployment Manager baseado em navegador. Consulte o *IBM Analytical Decision Management* para obter mais informações.
- Verifique se o IBM SPSS Modeler Server está configurado corretamente. Consulte a Nota técnica a seguir para obter detalhes.

<http://www-01.ibm.com/support/docview.wss?uid=swg21488191>

Regras de segmento

Ao criar regras de segmento no IBM Analytical Decision Management, observe que uma regra como Age BETWEEN 50 and 20 não retornará resultados. É a mesma expressão $x \leq \text{valor}$ e $\text{valor} < y$, o que nunca pode ser verdade. Isto é consistente com o comportamento do SQL Server. Idade ENTRE 20 e 50 é o formato correto.

Modelos de CPLEX

Os usuários avançados podem ajustar os parâmetros do CPLEX para IBM Analytical Decision Management para obter melhor desempenho em grandes conjuntos de dados. Também recomendamos minimizar o número de alocações possíveis apresentadas para o CPLEX (o volume de dados). Em outras palavras, você deve reduzir o número de combinações (por exemplo, o número de ofertas alternativas alocadas para cada cliente).

Os administradores podem usar o botão **Opções avançadas** na guia Otimizar no IBM Analytical Decision Management para exportar os arquivos *lp* ou *.mps* do CPLEX para propósitos de resolução de problemas ou para controlar certos parâmetros do CPLEX.

A seguir está uma lista de parâmetros do CPLEX que pode ser usada para o nó do CPLEX no cliente do IBM SPSS Modeler. Estes são os nomes customizados válidos que são reconhecidos pela integração do CPLEX com IBM Analytical Decision Management.

Importante: Consulte a documentação do IBM ILOG CPLEX Optimization Studio para obter detalhes sobre essas configurações. Observe que elas podem ter efeito adverso no desempenho do IBM Analytical Decision Management.

Tabela 1. Parâmetros do CPLEX

Configurações booleanas	Configurações de ponto flutuante	Configurações de número inteiro	Configurações de sequência de caracteres
CloneLog	EpMrk	AdvInd	WorkDir
PreInd	EpOpt	RootAlg	FileEncoding
MIPOrdInd	EpPer	NodeAlg	APIEncoding
MPSLongNum	EpRHS	MIPEmphasis	IntSolFilePrefix
LBHeur	NetEpOpt	AggInd	
PerInd	NetEpRHS	ClockType	
PreLinear	TiLim	CraInd	
DataCheck	DefTiLim	DepInd	
QPmakePSDInd	TuningTiLim	PreDual	
MemoryEmphasis	BfTol	PrePass	
NumericalEmphasis	CutLo	RelaxPreInd	
	CutUp	RepeatPresolve	
	EpGap	Symmetry	
	EpInt	DPriInd	
	EpAGap	PriceLim	
	EpRelax	SimDisplay	
	ObjDif	NetFind	
	ObjLLim	PerLim	
	ObjULim	PPriInd	
	PolishTime	ReInv	
	PolishAfterEpAGap	ScaInd	
	PolishAfterEpGap	Threads	
	PolishAfterTime	AuxRootThreads	
	ProbeTime	ParallelMode	
	RelObjDif	SingLim	
	CutsFactor	Reduce	
	TreLim	ColReadLim	
	SolnPoolGap	RowReadLim	
	SolnPoolAGap	WriteLevel	
	WorkMem	SiftDisplay	
	BarEpComp	SiftAlg	
	BarQCPEpComp	BrDir	
	BarGrowth	Cliques	
	BarObjRng	CoeRedInd	
	EpLin	Covers	
		MIPDisplay	
		NodeFileInd	
		NodeSel	
		VarSel	

Tabela 1. Parâmetros do CPLEX (continuação)

Configurações booleanas	Configurações de ponto flutuante	Configurações de número inteiro	Configurações de sequência de caracteres
		BndStrenInd	
		FPHeur	
		MIPOrdType	
		FlowCovers	
		ImplBd	
		Probe	
		GUBCovers	
		StrongCandLim	
		FracCand	
		FracCuts	
		PreslvNd	
		FlowPaths	
		MIRCuts	
		DisjCuts	
		ZeroHalfCuts	
		MCFCuts	
		AggCutLim	
		EachCutLim	
		DiveType	
		MIPSearch	
		MIQCPStrat	
		SolnPoolCapacity	
		SolnPoolReplace	
		SolnPoolIntensity	
		PopulateLim	
		BarAlg	
		BarColNz	
		BarDisplay	
		BarOrder	
		BarCrossAlg	
		BarStartAlg	
		NetPPriInd	
		NetDisplay	
		ConflictDisplay	
		FeasOptMode	
		TuningMeasure	
		TuningRepeat	
		TuningDisplay	
		MIPKappaStats	

Tabela 1. Parâmetros do CPLEX (continuação)

Configurações booleanas	Configurações de ponto flutuante	Configurações de número inteiro	Configurações de sequência de caracteres
		SolutionTarget	

Capítulo 4. Melhores práticas de desempenho

Os tópicos nesta seção fornecem as melhores práticas para o IBM Analytical Decision Management. Várias práticas podem ter impacto sobre o desempenho. Esta seção descreve abordagens de implementação que podem ajudá-lo a balancear o desempenho de acordo com suas necessidades. Observe que algumas das informações e recomendações podem não se aplicar ao seu ambiente específico ou uso do produto.

Número de registros de origem

O número de linhas de dados em sua origem de dados pode afetar o desempenho do IBM Analytical Decision Management. Em particular, o número de linhas de dados pode afetar quanto tempo leva para executar as seguintes operações:

- Varredura de dados
- Cenário
- Escorar

Essas operações normalmente têm uma sobrecarga fixa para iniciar a operação e para reunir os resultados. Para origens de dados de tamanho não trivial, as operações reais aumentam linearmente com a quantidade de dados processados. Se não houver nenhuma outra alteração, uma operação que processa o dobro de registros geralmente leva o dobro do tempo.

A guia Dados em um projeto do IBM Analytical Decision Management requer que os usuários selecionem uma única origem como o *modelo de dados do projeto*. O modelo de dados do projeto define os campos que o restante do projeto usa. O IBM Analytical Decision Management também permite que os usuários criem origens adicionais que são compatíveis com a origem de modelo de dados do projeto. Ao criar origens adicionais compatíveis com diferentes números de registros, o usuário pode escolher uma origem dimensionada adequadamente para reduzir os tempos de execução.

Por exemplo, suponha que as origens a seguir são definidas:

- Uma **origem principal** que contém 1.000.000 registros e está configurada com o modelo de dados do projeto.
- Uma **origem de cenário rápido** que contém 50.000 registros e é compatível com o modelo de dados do projeto.
- Uma **origem de cenário médio** que contém 20.000.000 registros e é compatível com o modelo de dados do projeto.
- Uma **origem de todos os registros** que contém 100.000.000 registros e é compatível com o modelo de dados do projeto.

Inicialmente, o usuário executa uma varredura de dados usando a **origem principal**. Nesse caso, os registros da origem principal são suficientes para definir todos os campos juntamente com seus valores categóricos.

O usuário vai para a guia Seleções globais e inicia a criação de regras. O usuário está fazendo experimentos com várias regras e deseja testá-las para ver o efeito bruto delas. O usuário escolhe a **origem de cenário rápido** para retornar resultados rápidos e preliminares.

Eventualmente, o usuário fica satisfeito com seus esforços preliminares e escolhe executar um cenário usando as origens de dados maiores (**origem de cenário médio** ou **origem de todos os registros**). Ainda que levem proporcionalmente mais tempo para serem executadas, o usuário está disposto a esperar pelo resultado mais abrangente.

Tamanho da origem de dados e efeitos em Teste

A operação Teste foi desenvolvida para operar com poucos registros. Muitas vezes, apenas os primeiros dez registros serão lidos na origem de dados, e o tamanho da origem de dados não importa.

No entanto, se diversas tabelas forem usadas (por exemplo, Clientes e Contatos), então, os registros dessas fontes deverão ser unidos e fazer parte da execução de Teste. Como a junção pode ter como resultado que todos os registros das origens sejam processados, o desempenho pode ser muito melhorado pelo uso de origens com pequeno número de registros.

Tamanho da origem de dados e projetos que usam Otimização

Alguns aplicativos do IBM Analytical Decision Management, como IBM Analytical Decision Management for Campaign Optimization e IBM Analytical Decision Management for Demand Optimization, usam otimização baseada em CPLEX. Esses aplicativos otimizam mais do que todos os registros de origem juntos, ao contrário de simplesmente um registro de cada vez. Como resultado, os tempos de execução para operações que executam a etapa Otimizar baseado em CPLEX não aumentam linearmente com a quantidade de dados processada. Um aumento no número de registros geralmente resulta em tempos de execução que são proporcionalmente muito maiores.

As operações que não executam a etapa Otimizar, como uma varredura de dados, continuam a aumentar linearmente.

Consulte “Otimização do CPLEX” na página 25.

Número de campos no modelo de dados do projeto

Assim como o número de linhas de dados pode afetar o desempenho do IBM Analytical Decision Management, o número de campos também tem impacto. Para origens de tamanho não trivial, as operações reais aumentam aproximadamente e linearmente com a quantidade de dados processados. Depois de ignorar qualquer sobrecarga de inicialização, uma operação que processa o dobro de campos geralmente levará o dobro do tempo.

Como uma melhor prática, use origens de dados que contêm apenas os campos que você realmente precisa.

Fazendo upload de dados

A guia Dados em IBM Analytical Decision Management permite que os usuários façam upload de arquivos de dados. Esse recurso foi projetado para arquivos relativamente pequenos. Evite usar este recurso para arquivos maiores que 750 megabytes, já que as transferências levarão muito tempo. Para arquivos grandes, use um método de upload alternativo, como o FTP, para transferir arquivos diretamente para o sistema de arquivos do IBM SPSS Modeler Server.

Dimensões (hierarquias) e os membros de dimensão

No geral, incluir dimensões adicionais e membros de dimensão aumentará a complexidade do fluxo e diminuirá o desempenho.

A sobrecarga de incluir um membro de dimensão depende de vários fatores:

- **Dimensão pai versus dimensão filho.** Incluir um membro de dimensão na dimensão pai (por exemplo, incluindo uma nova campanha) normalmente envolve mais sobrecarga do que incluir outro membro de dimensão filho (por exemplo, incluir outra oferta). Incluir um membro pai tem um impacto composto devido à sua regra de alocação, que, por sua vez, irá alocar membros filhos adicionais.

- **Regra de seleção.** Um membro que contém uma regra de seleção irá requerer processamento adicional se comparado a um membro que não tem uma regra de seleção. Membros na dimensão pai (por exemplo, campanha) têm uma regra de seleção padrão. Contudo, os resultados irão variar com base no número de registros realmente excluídos pelo filtro da regra. Uma regra de seleção cujo filtro exclui a maioria dos registros terá um impacto positivo no desempenho, pois resultará em menos alocações (para otimizar).
- **Variáveis.** Variáveis específicas de membro irão requerer processamento adicional para cada variável. Considere colocar as variáveis de priorização, otimização e/ou retorno no maior escopo possível. Por exemplo, considere um projeto que possui duas campanhas, e cada campanha alocará dez ofertas. Se a variável (probabilidade de responder) é a mesma para todas as ofertas alocadas pela *campaign1*, então, coloque-a em *campaign1* em vez de em cada uma das dez ofertas. Ao contrário, considere um projeto que tem dez campanhas e duas ofertas *compartilhadas*. Se a variável (probabilidade de responder) é a mesma para todas as campanhas, é melhor colocá-la no nível da oferta (nas duas ofertas em vez de em dez campanhas). Variáveis normalmente têm menos impacto do que incluir dimensões ou membros.
- **Número de regras.** O número de regras e os segmentos de regra (em uma regra de seleção ou de alocação) também podem impactar o desempenho se o número de regras se tornar grande.
- **Referências de modelo e regra externa.** Fazer referência a modelos ou regras externas (conjuntos de regras no IBM Operational Decision Management) também pode ter impacto sobre o desempenho. Consulte “Fazendo referência a modelos” na página 24 e “Fazendo referência a regras externas” na página 25.

Otimizando a geração de fluxo

Em determinadas condições, IBM Analytical Decision Management pode gerar um fluxo mais ideal que evita ramificações desnecessárias e será executado com mais eficiência.

Um fluxo mais eficiente é gerado quando todos os membros de dimensão na dimensão de nível folha (por exemplo, ofertas) satisfazem todas as condições a seguir:

- Regras de seleção não são usadas na dimensão de nível de folha (Oferta, por exemplo)
- Variáveis (de priorização/otimização, por exemplo) não são usadas no nível de folha
- Variáveis de retorno não são usadas no nível de folha
- O nível folha não está configurado como o `priorityDimension`
- Para aplicativos que usam uma regra de alocação, a dimensão pai (Campanha, por exemplo) usa as regras de alocação com a opção **Primeiro válido** selecionada (em vez de **Todos os válidos**). Além disso, a regra de alocação deve gerar um resultado único (por exemplo, **Alocar usando regras** ou **Alocar aleatoriamente** em vez de **Alocação múltipla**).
- Para aplicativos que usam uma regra de agregação, a dimensão pai (Área da reclamação, por exemplo) usa uma regra de agregação, na qual a opção restante é configurada como **Aplicar quando não houver ocorrências das regras** (em vez de **Aplicar sempre**).

Um projeto que é construído desta maneira pode ser considerado como um projeto de "regra de alocação somente" e pode frequentemente ser adequado nos casos em que a velocidade de execução máxima é essencial. Um exemplo é se o projeto for implementado com o IBM SPSS Analytics Toolkit for InfoSphere Streams.

SQL pushback

SQL Pushback é uma técnica que o IBM SPSS Modeler Server usa para melhorar o desempenho para origens de banco de dados. As junções do banco de dados e outros cálculos são automaticamente integrados às instruções SQL geradas. O processamento é enviado por push para o banco de dados em que pode ser manipulado mais eficientemente. Isso também pode minimizar a quantidade de dados que o IBM SPSS Modeler Server deve ler na memória, o que também melhora o desempenho.

O uso de SQL pushback pode resultar em desempenho dez vezes melhor (ou mais).

Nos fluxos gerados do IBM Analytical Decision Management, o SQL pushback é, geralmente, mais bem alavancado quando são usadas origens da Visualização de dados analíticos.

No geral, o SQL pushback não funciona por meio de referências de modelo. Referências de modelo impedem que o SQL pushback funcione em todos os nós de recebimento de dados. Desde que não existam referências de modelos para evitar, o SQL pushback pode ser usado em regras de Seleções globais e nas regras de seleção nos membros de dimensão da primeira dimensão (por exemplo, Área de reclamação ou Campanha).

Atributos derivados de modelo e SQL pushback

Na subguia **Derivado** da guia **Dados**, em IBM Analytical Decision Management, é possível criar atributos adicionais usando um modelo para agregar dados a partir de uma origem de dados secundária. Ao usar origens ODBC (seja diretamente ou por meio da Visualização de dados analíticos), a junção pode ser enviada por push de volta para o banco de dados. A melhoria resultante no desempenho pode ser dramática. Em alguns casos, pode ser por um fator de dez ou mais (SQL pushback leva um décimo do tempo ou menos em comparação com o arquivo simples de não pushback).

No entanto, a junção será enviada por push de volta apenas para a primeira tabela secundária. A referência de modelo impede que a junção para referências de modelo subsequentes e tabelas secundárias subsequentes faça o pushback.

O nó de origem da Visualização de dados analíticos deve ser usada para conseguir o SQL pushback com diversas referências de modelo e/ou diversas tabelas secundárias.

SQL pushback usando mais de uma tabela secundária

Para usar o SQL pushback com mais de uma tabela secundária, a Visualização de dados analíticos deve ser usada. Em vez de usar um modelo para derivar campos de uma tabela secundária, integre a junção para a tabela secundária juntamente com a lógica de derivação na Visualização de dados analíticos.

Por exemplo, suponha que você tem duas tabelas de banco de dados: Customer e ContactHistory. Para usar um modelo para derivar campos de uma tabela secundária no IBM Analytical Decision Management, você precisa definir ContactHistory como tabela secundária:



Figura 1. Secondary table

Na guia **Dados**, acesse a guia **Derivar** e use um modelo para derivar NumContacts.

Source		Derived	
Additional Fields			
Field name	Source	Measurement	Value
1	Contact_Count	CountContacts	Continuous [1,8]

Figura 2. Additional fields

O modelo *CountContacts* referenciado pode ser parecido com este:



Figura 3. Example *CountContacts* model

Esse processo pode ser repetido pela inclusão de tabelas secundárias e modelos de referência adicionais na guia **Derivar**. Por exemplo, você pode ter uma tabela *ProductsPurchased* e você pode ter um fluxo que agrega (conta) o número dos produtos comprados anteriormente por cada cliente.

Por comparação, para realizar o mesmo usando uma Visualização de dados analíticos, crie um fluxo que define a Tabela de visualização de dados. Ele fará a junção e a(s) agregação(ões). Em um caso simples com uma tabela secundária única, ele pode ser semelhante a este:

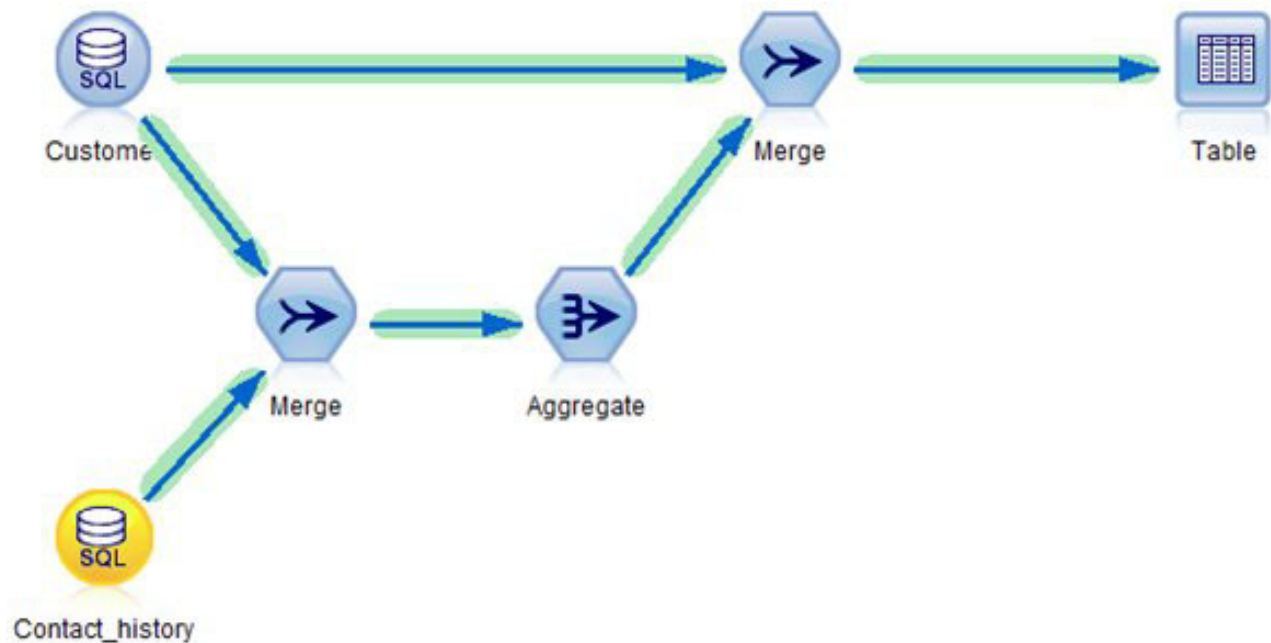


Figura 4. Example using an Analytic Data View

Se atributos derivados são necessários em múltiplas tabelas secundárias, o fluxo que é usado para definir a Visualização de dados analíticos pode ser semelhante à esta:

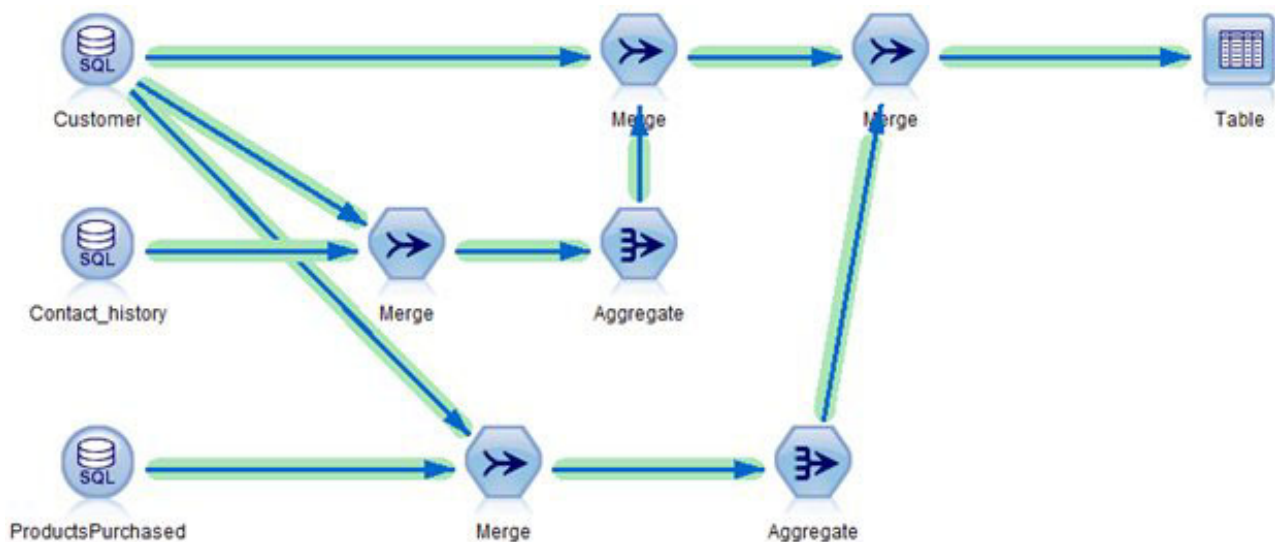


Figura 5. Example using multiple secondary tables

A definição da Visualização de dados analíticos no cliente do Deployment Manager parece de alguma forma com esta:

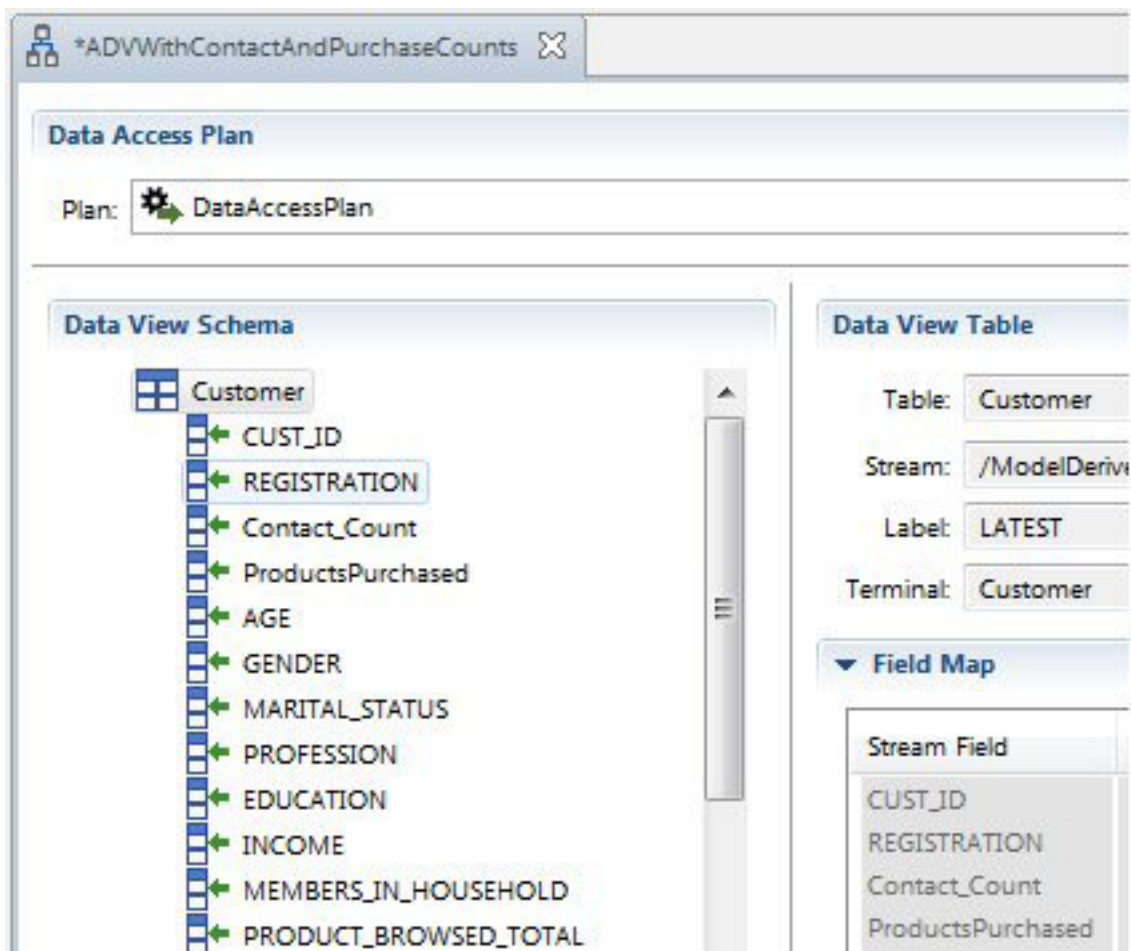


Figura 6. Deployment Manager client

Observe que o esquema de Visualização de dados analíticos contém os campos `Contact_Count`, `ProductsPurchased` e os campos da tabela `Customer`. Como esses campos estão na Visualização de dados analíticos, eles estarão presentes no modelo de dados do projeto do IBM Analytical Decision Management, e as tabelas secundárias não serão definidas na interface com o usuário do IBM Analytical Decision Management. Novamente, esta abordagem pode usar o SQL pushback e pode ser dez vezes mais rápida (ou mais) que as alternativas que não podem ser enviadas por push de volta para o banco de dados.

Usar expressões derivadas com o editor da Visualização de dados analíticos é outra abordagem alternativa que também faz uso do SQL pushback quando as origens de dados são usadas e o fluxo de origem digita um campo como o "ID de registro". Consulte a documentação da Visualização de dados analíticos do IBM SPSS Collaboration and Deployment Services para obter mais informações.

Derivando diversos atributos da mesma tabela secundária

Se você optar por derivar atributos usando modelos na subguia **Derivar** da guia **Dados**, geralmente é mais eficiente combinar a derivação de atributos de uma tabela secundária determinada em um único modelo referenciado. Por exemplo, suponha que `ContactCount` e `MostRecentContactDate` são ambos calculados usando a tabela `Contact_History`. É mais eficiente combinar a lógica em um único fluxo referenciado que produz duas saídas (`ContactCount` e `MostRecentContactDate`) em vez de fazer referência a dois fluxos separados na guia **Derivar**.

Alocação estática versus dinâmica

O IBM Analytical Decision Management suporta dois tipos de métodos de alocação na guia **Definir**:

- **Alocação estática.** Todos os membros de dimensão (`myOffer1` e `myOffer2`, por exemplo) são especificados na árvore de dimensão na guia **Definir** e se tornam parte do fluxo do projeto.
- **Alocação dinâmica** Alguns membros de dimensão são lidos em uma origem de dados.

O desempenho varia entre essas duas abordagens. No geral, a alocação estática tem melhor execução porque os nomes dos membros são locais para o projeto, e o processamento pode evitar a sobrecarga de entrada/saída da leitura dos mesmos nomes em uma origem externa.

Alocação estática

Já que todos os membros de dimensão (`myOffer1` e `myOffer2`, por exemplo) são parte do fluxo do projeto, não existem sobrecarga de entrada/saída adicional na junção e/ou leitura de arquivos secundários. No entanto, o fluxo geralmente requer uma ramificação por membro de dimensão. Essas ramificações adicionais podem ter um impacto negativo no desempenho.

Alocação Dinâmica

Por alocação dinâmica, os membros afetados devem ser lidos em um arquivo de origem. Geralmente, os membros estão em um arquivo de origem secundário. Uma das duas abordagens pode ser usada:

- A tabela secundária é vinculada à tabela primária
- A tabela secundária não é vinculada

Se a tabela secundária for vinculada à tabela primária, então, uma junção deve ser executada. Na construção do fluxo atual, essa junção não é elegível para SQL pushback, e os dados devem ser lidos no IBM SPSS Modeler Server. O desempenho varia de acordo com o tamanho das tabelas primárias e secundárias.

Se a tabela secundária não for vinculada à tabela primária, então, todos os registros são lidos a partir da tabela secundária, resultando em um membro alocado para cada registro na tabela. Nesse caso, o número total de registros na tabela secundária deve ser relativamente pequeno. Mesmo que a junção seja evitada, a alocação estática ainda é mais rápida. No entanto, a alocação dinâmica oferece melhor escalabilidade

porque os membros não precisam ser digitados manualmente na interface com o usuário do IBM Analytical Decision Management.

Fazendo referência a modelos

Em vez de usar um modelo para derivar um campo de uma tabela secundária, o IBM Analytical Decision Management também permite que modelos sejam usados em várias condições, tais como:

- Ao fazer referência a um modelo de uma regra
- Ao fazer referência a um modelo de uma variável (de priorização/otimização)
- Ao fazer referência a um modelo de um valor de campo de retorno
- Ao fazer referência a um modelo da guia de dados para definir um novo atributo da origem primária de modelo de dados do projeto (não da origem secundária)

Nessas condições, um modelo é chamado e opera em uma única linha de dados. O modelo não está agregando diversas linhas de dados. Como o modelo opera em uma única linha de dados, chamá-lo requer sobrecarga mínima de dados. Isso está em contraste com o uso de um modelo na subguia **Derivar** da guia **Dados** para agregar dados de uma tabela secundária em que os dados devem primeiro ser unidos.

Antes de fazer a escoragem de quaisquer linhas, o modelo deve ser preparado para execução. Esse processo envolve a leitura do modelo no repositório e a construção das estruturas necessárias de dados na memória. Essa sobrecarga ocorre independente do número de linhas de dados escoradas. Isso pode afetar os tempos de execução para **Cenário**, **Teste** e as operações de escoragem de lote (**Escorar agora**) ao usar conjuntos de dados pequenos. Isso também pode afetar o tempo que é necessário para a criação de uma configuração de escoragem no cliente do Deployment Manager.

O custo de execução por linha da referenciação de um modelo depende da complexidade do modelo. Por exemplo, chamar um modelo simples que executa um cálculo trivial tem uma sobrecarga mínima de desempenho. Chamar um modelo complexo que contém muitos nós e várias ramificações tem um impacto negativo mais significativo no desempenho.

Para modelos implementados por meio do Serviço de escoragem:

- Se o modelo for referido diretamente e não for referido por uma regra, então, para uma determinada solicitação de escoragem (uma determinada entidade como *cliente*), um modelo referenciado será escorado apenas uma vez e apenas se necessário para determinar a escoragem. O resultado é armazenado em cache e reutilizado, se for necessário novamente durante a mesma solicitação de escoragem (por exemplo, o mesmo modelo referenciado por duas variáveis de priorização diferentes na guia **Priorizar**).

O modelo não é executado se o resultado não for necessário. Por exemplo, suponha que um modelo é usado para derivar um campo na guia **Dados**, e esse campo é, então, referenciado por uma regra na guia **Definir**. Se uma solicitação de escoragem resultar na exclusão da entidade da guia **Seleções globais**, então, o processamento para (em seleções globais), e o modelo não é executado.

Para escoragem no IBM SPSS Modeler Server (**Cenário** e **Escorar agora**):

- No tempo de execução, ao escorar uma determinada entidade (registro de entrada), os resultados não são armazenados em cache. Se várias referências são feitas ao mesmo modelo (em duas variáveis diferentes, por exemplo), então, o modelo será executado várias vezes – uma vez para cada referência. O impacto é amplamente determinado pela rapidez com que o modelo é executado (modelo simples versus um modelo de mineração de texto grande, por exemplo).
- Se o mesmo modelo é referenciado várias vezes, você pode melhorar o desempenho, basta fazer referência ao modelo na guia **Dados** (subguia **Derivar**) para derivar um novo campo. Em seguida, faça referência ao campo derivado em todo o projeto. Ao fazer isso, o modelo sempre será chamado exatamente uma vez. Observe que a referência ao modelo evita o SQL pushback de recebimento de dados (impede que SQL pushback nas seleções globais).

- Se um modelo é raramente usado (por exemplo, é referenciado em uma variável para um membro como *oferta* que quase nunca é alocado), então, pode ser melhor colocar a referência na variável do membro (*oferta*) em vez de na guia **Dados**. Se o membro não for alocado, então, o modelo não é executado.

Fazendo referência a regras externas

Fazer referência a regras externas (como conjuntos de regras do IBM Operational Decision Management) tem as mesmas implicações e se comporta da mesma forma que fazer referência a modelos. Consulte “Fazendo referência a modelos” na página 24. Além disso, regras externas podem operar em dois modos de execução:

- **Serviço da web.** Neste modo, uma chamada de serviço da web remoto é executada cada vez que o conjunto de regras é chamado. Geralmente, essa abordagem é mais fácil de configurar. No entanto, o desempenho é o mais lento das duas opções. O desempenho depende, em parte, do desempenho da rede e do desempenho do serviço da web.
- **Local (em processo).** Neste modo, as regras executam em processo. Os dados são transmitidos na memória, e a sobrecarga de rede é eliminada. Essa abordagem é a mais rápida das duas.

As regras do IBM Operational Decision Management que são executadas localmente podem resultar em um tempo de execução geral do projeto aproximadamente 20% mais rápido em comparação a projetos que usam regras do IBM Operational Decision Management executadas por um serviço da web. Os resultados irão variar de acordo com muitos fatores, tais como:

- A quantidade de processamento de conjunto de regras em relação ao processamento geral do projeto
- Sobrecarga de rede ao usar o modo de serviço da web

Otimização do CPLEX

Os aplicativos do IBM Analytical Decision Management podem ser configurados para incluir a guia Otimização e usar otimização baseada em matemática complexa (CPLEX). Um exemplo desses aplicativos é o IBM Analytical Decision Management for Campaign Optimization.

A escoragem em um aplicativo baseado em CPLEX ocorre em duas fases:

1. A fase 1 processa as regras e os modelos. Isso é o mesmo que qualquer outro projeto que não use CPLEX.
2. A fase 2 chama CPLEX para resolver a otimização. Essa fase é exclusiva para projetos do IBM Analytical Decision Management que usam CPLEX.

Projetos que usam a otimização baseada em CPLEX geralmente não aumentam linearmente com o número de registros escorados. Se não houver nenhuma outra alteração, o aumento do número de registros em dez vezes, geralmente, resulta em tempos de escoragem que são significativamente maiores que dez vezes (os tempos de escoragem vão crescer provavelmente em escala exponencial). Isso ocorre porque a fase de resolução de CPLEX não aumenta linearmente. *Na verdade, o tempo de resolução não é previsível. É possível que problemas aparentemente maiores sejam resolvidos mais rapidamente do que problemas menores.* Isso ocorre porque a otimização determina a melhor combinação de todas as entidades tomadas como um todo, em vez de determinar o melhor resultado para cada entidade tomada individualmente. Os algoritmos que são usados não aumentam em escala previsível conforme aumenta o tamanho do problema.

Requisitos de desempenho e recurso (memória) são amplamente indeterminados. No geral, quanto mais combinações e quanto maior a complexidade, mais recursos e mais tempo são requeridos. Os valores de dados reais impactam a velocidade da otimização. O desempenho não aumenta em escala linear e não é previsível. Alguns problemas não podem ser resolvidos.

O desempenho de otimização baseada em CPLEX é influenciado por muitos fatores, tais como:

- Tamanho do problema
- Modelagem de problema de negócios na Optimization Programming Language (OPL)
- A complexidade do modelo de OPL
- Os dados (valores) sendo otimizados
- Recursos (memória e CPU) disponíveis

Tamanho do problema

O tamanho do problema se refere ao número de casos que servem como entrada no processo de otimização. No IBM Analytical Decision Management, o tamanho do problema é o número de casos (registros) que são transmitidos da etapa Definir para a etapa Otimização. Em termos de CPLEX, o tamanho do problema é o número de tuplas de entrada. O aumento do número de casos geralmente pode aumentar o tamanho do problema e resultar em tempos de execução maiores.

O número de casos que servem como entrada para a fase de otimização pode ser calculado por:

$$\text{Number of entity records} * \text{average number of allocations per entity}$$

Para ilustrar isso, considere um projeto do IBM Analytical Decision Management for Campaign Optimization que possui um arquivo de entrada de 2 milhões de entidades (clientes). Suponha que as Seleções globais excluam metade delas, deixando 1 milhão de entidades. Na guia Definir, suponha que, em média, cada cliente é elegível (com base nas regras de Seleção de campanha) para dez campanhas diferentes. Além disso, suponha que a regra de alocação para cada campanha é configurada para *parar na primeira ocorrência* – resultando em uma oferta por campanha. A dimensão de canal é usada, mas há apenas um canal definido. Isso resulta em:

$$\text{million customers} * 10 \text{ campaigns} * 1 \text{ offer} * 1 \text{ channel} = 10,000,000 \text{ cases}$$

Para ver o impacto dos canais, suponha que um cliente pode ser alocado em um de oito canais:

$$1 \text{ million customers} * 10 \text{ campaigns} * 1 \text{ offer} * 8 \text{ channels} = 80,000,000 \text{ cases}$$

Além dos canais incluídos, a regra de Alocação da campanha não está configurada para parar na primeira ocorrência, mas para executar todas. Em média, para cada campanha, quinze ofertas são alocadas para cada cliente.

$$1 \text{ million customers} * 10 \text{ campaigns} * 15 \text{ offers} * 8 \text{ channels} = 1,200,000,000 \text{ cases}$$

A tarefa do otimizador é determinar a combinação ideal das 1.200.000.000 tuplas atendendo aos limites que são impostos pelas restrições.

Modelagem de problema de negócios na OPL

A maneira como o problema de negócios é modelado no modelo de Optimization Programming Language (OPL) pode ter grande impacto sobre o desempenho. Ao criar aplicativos customizados, o editor de telas deve ter atenção na composição da OPL para que os modelos possam ser resolvidos de maneira eficiente. Diferentes abordagens no modelo de OPL podem resultar em desempenhos diferentes.

Complexidade do modelo de OPL

A complexidade do modelo de Optimization Programming Language (OPL) influencia o desempenho. A seguir estão alguns fatores que influenciam a complexidade do modelo:

- **A função Objetivo.** Uma função objetivo mais complexa aumenta o tempo que leva para produzir a solução ideal. Uma função objetivo linear é menos complexa que uma função objetivo quadrática. A função objetivo influencia os algoritmos que o CPLEX pode escolher. Objetivos quadráticos devem ser resolvidos usando algoritmo de barreira, já objetivos lineares permitem que o CPLEX use mais algoritmos. Em IBM Analytical Decision Management, a função objetivo é determinada pelo aplicativo que estiver usando (por exemplo, o IBM Analytical Decision Management for Campaign Optimization usa o Mixed Integer Programming).

- **Número e colocação de restrições.** O aumento do número de restrições aumenta a complexidade. Observe que um único gesto de restrição no IBM Analytical Decision Management pode ser refletido em diversas restrições para o CPLEX.
- **Modelagem das restrições.** Novamente, o modo como um problema de negócios é modelado na OPL pode impactar o desempenho. Isso é verdade na modelagem de restrições. Ao criar aplicativos customizados, o editor de telas deve ter atenção ao modelar restrições de negócios na OPL.

Restrições globais e de entidade

Por exemplo, uma restrição no IBM Analytical Decision Management de `Max_Offers = 1` colocada na entidade pode ser considerada como uma restrição única para o usuário do IBM Analytical Decision Management. No entanto, durante o processo de otimização, isso realmente significa que cada uma das entidades tem uma restrição `Max_Offers` de 1. Com efeito, no exemplo anterior, isto é convertido em 1.000.000 de restrições no CPLEX porque deve ser indicado em cada entidade.

Compare isso com uma restrição `Total Budget`, que é de fato uma restrição global única conforme considerada pela otimização do CPLEX. `Total Budget` tem como resultado uma única restrição no CPLEX. A restrição de orçamento faz a soma de todos os "locais" em que o orçamento pode ser gasto e pode ser expressa por uma soma única.

Restrições de membro de dimensão

Um exemplo de uma restrição em um membro de dimensão é `Offers Available`. Se houver cem ofertas diferentes, então, o CPLEX as vê como cem restrições separadas. O CPLEX considera como uma única oferta compartilhada (uma oferta que é compartilhada entre várias campanhas) como uma oferta exclusiva com uma restrição. Dez campanhas que alocam, cada, uma única oferta compartilhada têm como resultado uma restrição `Offers Available`, enquanto dez campanhas que alocam uma única oferta distinta por campanha têm como resultado dez restrições `Offers Available` (uma por oferta exclusiva).

Dados

Os valores de dados reais podem ter impacto na velocidade de otimização.

Os valores de dados têm impacto sobre como o processamento avança e quanto tempo leva. Isso é conhecido como *variabilidade de dados*. Dois problemas do mesmo tamanho podem diferir consideravelmente em relação à quantidade de tempo (memória e CPU) necessária para resolver o problema. Isso ocorre devido a diferenças na estrutura do modelo e/ou nas variáveis de acordo com as quais o algoritmo toma as decisões (uma solução representa todas as decisões em todas as variáveis do modelo).

Dependendo dos dados, alguns problemas podem não ser resolvidos. Por exemplo, considere um problema com 10.000 entidades, cada entidade deve receber uma oferta, cada oferta custa mais de U\$ 1, e o orçamento é U\$ 9.000.

Algoritmo

Com base na função objetivo e outras entradas, o CPLEX escolhe um algoritmo para resolver o problema. Algoritmos diferentes têm diferentes capacidades e executam de forma diferente.

Os aplicativos que requerem Mixed Integer Programming (MIP) podem levar um longo tempo para serem resolvidos. O motivo é que nenhum algoritmo eficiente (pela medida de métrica polinomial) existe atualmente para Programação inteira mista. O que não acontece com a Programação linear, na qual os algoritmos eficientes já existem.

Recursos (memória e CPU) disponíveis

Os recursos disponíveis (memória e CPU) influenciam o tempo que leva para resolver um problema de CPLEX. Os resultados variam. A memória é usada para manter a configuração inicial do modelo em termos da "matriz". Em seguida, a memória é usada para manter a árvore de procura. Uma vez que a

RAM foi esgotada, mas a resolução não foi concluída, é possível migrar a árvore de procura para o disco rígido (uma configuração de CPLEX existe para determinar se isto é feito ou não). Como o tempo de acesso ao disco é muito maior do que à RAM, a procura perde velocidade significativamente.

O CPLEX usa vários encadeamentos sempre que possível. Por padrão, o número de encadeamentos criados é determinado pelo número de CPUs no sistema. Executar em um sistema com vários processadores pode reduzir o tempo que leva para produzir a solução otimizada. Os encadeamentos processam as ramificações na árvore de procura os mais paralelamente possível. Por padrão, o algoritmo paralelo do CPLEX é usado.

Acesso ao banco de dados ao usar o Serviço de escoragem

Esta seção inclui considerações para implementar projetos do IBM Analytical Decision Management no IBM SPSS Collaboration and Deployment Services Serviço de escoragem para escoragem em tempo real.

Ao contrário da escoragem de lote, em que todos os registros de uma tabela de banco de dados são escorados, a escoragem em tempo real opera em um registro (o registro que está sendo escorado) por vez. Quando o Serviço de escoragem recebe uma solicitação de escoragem, o responsável pela chamada pode transmitir os dados de entrada na chamada ou transmitir um campo chave de modo que o Serviço de escoragem possa buscar os dados de uma tabela de banco de dados, conforme definido por um Plano de acesso de dados em tempo real de uma Visualização de dados analíticos.

Em qualquer caso, o SQL pushback não se aplica como seria o caso ao executar um fluxo de lote no IBM SPSS Modeler Server. Em vez disso, os dados que são recuperados de um banco de dados são controlados pelos filtros que são especificados no Plano de acesso de dados em tempo real. O Serviço de escoragem usa esses filtros para gerar as classes WHERE adequadas ao buscar os dados para uma solicitação de escoragem. Normalmente, esses filtros resultam em poucos registros sendo buscados.

Por exemplo, considere o exemplo anterior no qual o projeto tem Customer e ContactHistory como as tabelas de entrada. O Plano de acesso de dados em tempo real pode ser configurado para filtrar por CUST_ID. O Serviço de escoragem gera a cláusula WHERE adequada para buscar apenas os registros para este cliente de cada tabela. Para a tabela Customer, existe apenas um registro – aquele que apresenta esse cliente. Para a tabela ContactHistory, zero ou mais registros são buscados – um registro para cada vez que o cliente foi contatado. Todos os registros de contato para esse cliente são, então, transmitidos para o fluxo do IBM SPSS Modeler (referenciado a partir de um IBM Analytical Decision Management ou de dentro da Visualização de dados analíticos).

Evite situações nas quais o número de registros filtrados for grande (por exemplo, em que o cliente é geralmente contatado milhares de vezes).

Avisos

Estas informações foram desenvolvidas para produtos e serviços oferecidos nos Estados Unidos.

O IBM pode não oferecer os produtos, serviços ou recursos discutidos neste documento em outros países. Consulte um representante IBM local para obter informações sobre produtos e serviços disponíveis atualmente em sua área. Toda referência a um produto, programa ou serviço IBM não tem a intenção de declarar ou implica em que apenas esse produto, programa ou serviço IBM possa ser usado. Qualquer produto, programa ou serviço funcionalmente equivalente, que não infrinja nenhum direito de propriedade intelectual da IBM, poderá ser usado em substituição. Entretanto, a avaliação e verificação da operação de qualquer produto, programa ou serviço não IBM são de responsabilidade do usuário.

A IBM pode ter patentes ou solicitações de patentes pendentes relativas a assuntos tratados nesta publicação. O fornecimento desta publicação não garante ao Cliente nenhum direito sobre tais patentes. O Cliente pode enviar consultas sobre licença, por escrito, para:

Gerência de Relações Comerciais e Industriais da IBM Brasil
Av. Pasteur, 138/146
Botafogo
Armonk, NY 10504-1785
CEP 22290-240

Para pedidos de licença relacionados a informações de Conjunto de Caracteres de Byte Duplo (DBCS), entre em contato com o Departamento de Propriedade Intelectual da IBM em seu país ou envie pedidos de licença, por escrito, para:

Intellectual Property Licensing
Lei de Propriedade Legal e Intelectual
IBM Japan Ltd.
1623-14, Shimotsuruma, Yamato-shi
Tokyo 106, Japan

O parágrafo a seguir não se aplica a nenhum país em que tais disposições não estejam de acordo com a legislação local: A INTERNATIONAL BUSINESS MACHINES CORPORATION FORNECE ESTA PUBLICAÇÃO "NO ESTADO EM QUE SE ENCONTRA", SEM GARANTIA DE NENHUM TIPO, SEJA EXPRESSA OU IMPLÍCITA, INCLUINDO, MAS A ELAS NÃO SE LIMITANDO, AS GARANTIAS IMPLÍCITAS DE NÃO INFRAÇÃO, COMERCIALIZAÇÃO OU ADEQUAÇÃO A UM DETERMINADO PROPÓSITO. Alguns estados não permitem a exclusão de garantias expressas ou implícitas em determinadas transações, portanto, essa declaração pode não se aplicar ao Cliente.

Esta publicação pode incluir imprecisões técnicas ou erros tipográficos. Periodicamente são feitas alterações nas informações aqui contidas; tais alterações serão incorporadas em novas edições da publicação. A IBM pode, a qualquer momento, aperfeiçoar e/ou alterar os produtos e/ou programas descritos nesta publicação, sem aviso prévio.

Referências nestas informações a Web sites não IBM são fornecidas apenas por conveniência e não representam de forma alguma um endosso a esses Web sites. Os materiais contidos nesses Web sites não fazem parte dos materiais desse produto IBM e a utilização desses Web sites é de inteira responsabilidade do Cliente.

A IBM pode utilizar ou distribuir as informações fornecidas da forma que julgar apropriada, sem incorrer em qualquer obrigação para com o Cliente.

Licenciados deste programa que desejam obter informações sobre esse assunto com objetivo de permitir: (i) a troca de informações entre programas criados independentemente e outros programas (incluindo este) e (ii) a utilização mútua das informações trocadas, devem entrar em contato com:

Gerência de Relações
Comerciais e Industriais da IBM Brasil
Av. Pasteur, 138-146
Av. Pasteur, 138-146
Botafogo
CEP 22290-240

Tais informações podem estar disponíveis, sujeitas a termos e condições apropriadas, incluindo em alguns casos o pagamento de uma taxa.

O programa licenciado descrito nesta publicação e todo o material licenciado disponível são fornecidos pela IBM sob os termos do Contrato com o Cliente IBM, do Contrato de Licença de Programa Internacional IBM ou de qualquer outro contrato equivalente.

Todos os dados de desempenho aqui contidos foram determinados em um ambiente controlado. Portanto, os resultados obtidos em outros ambientes operacionais poderão variar significativamente. Algumas medidas podem ter sido tomadas em sistemas em nível de desenvolvimento e não há garantia de que estas medidas serão iguais em sistemas geralmente disponíveis. Além disso, algumas medidas podem ter sido estimadas por meio de extrapolação. Os resultados reais podem variar. Usuários deste documento devem verificar os dados aplicáveis para seus ambientes específicos.

As informações relativas a produtos não IBM foram obtidas junto aos fornecedores dos respectivos produtos, de seus anúncios publicados ou de outras fontes disponíveis publicamente. A IBM não testou estes produtos e não pode confirmar a precisão do desempenho, da compatibilidade ou de qualquer outra reivindicação relacionada a produtos não IBM. Dúvidas sobre os recursos de produtos não IBM devem ser encaminhadas diretamente a seus fornecedores.

Todas as instruções relativas as direções ou intenções futuras da IBM estão sujeitas a mudanças ou retirada sem aviso prévio, e apenas representam metas e objetivos.

Estas informações contêm exemplos de dados e relatórios utilizados em operações comerciais diárias. Para ilustrá-los da forma mais completa possível, os exemplos podem incluir nomes de indivíduos, empresas, marcas e produtos. Todos esses nomes são fictícios e qualquer semelhança com nomes e endereços utilizados por uma empresa real é mera coincidência.

Se estiver visualizando essas informações em formato eletrônico, as fotografias e ilustrações coloridas podem não aparecer.

Marcas Comerciais

IBM, o logotipo IBM e ibm.com são marcas ou marcas registradas da International Business Machines Corp., registradas em vários países no mundo todo. Outros nomes de empresas, produtos e serviços podem ser marcas registradas ou marcas da IBM ou de outras empresas. Uma lista atual de marcas registradas da IBM está disponível na web em "Copyright and trademark information" em www.ibm.com/legal/copytrade.shtml.

Adobe, o logotipo Adobe, PostScript e o logotipo PostScript são marcas ou marcas registradas da Adobe Systems Incorporated nos Estados Unidos e/ou em outros países.

Intel, logotipo Intel, Intel Inside, logotipo Intel Inside, Intel Centrino, logotipo Intel Centrino, Celeron, Intel Xeon, Intel SpeedStep, Itanium e Pentium são marcas comerciais ou marcas registradas da Intel Corporation ou de suas subsidiárias nos Estados Unidos e em outros países.

Linux é uma marca registrada de Linus Torvalds nos Estados Unidos e/ou em outros países.

Microsoft, Windows, Windows NT e o logotipo Windows são marcas registradas da Microsoft Corporation nos Estados Unidos e/ou em outros países.

UNIX é uma marca registrada da The Open Group nos Estados Unidos e em outros países.

Java e todas as marcas registradas e logotipos baseados em Java são marcas ou marcas registradas da Oracle e/ou de suas afiliadas.

Outros nomes de produtos e serviços podem ser marcas comerciais da IBM ou de outras empresas.

Índice Remissivo

C

correções
obtendo 3

D

determinação de problema
trocando informações com o Suporte
IBM 5

M

melhores práticas de desempenho 17
acesso ao banco de dados com Serviço
de escoragem 28
alocação dinâmica 23
alocação estática 23
alocação estática vs dinâmica 23
atributos derivados de modelo e SQL
pushback 20
dimensões (hierarquias) 18
fazendo referência a modelos 24
fazendo referência a regras
externas 25
geração de fluxo otimizada 19
membros de dimensão 18
métodos de alocação 23
número de campos no modelo de
dados do projeto 18
número de registros de origem 17
otimização do CPLEX 25
algoritmo 27
complexidade do modelo de
OPL 26
dados 27
memória e CPU 27
modelagem de problema de
negócios na OPL 26
restrições de membro de
dimensão 27
restrições globais e de
entidade 27
tamanho do problema 26
SQL pushback 19
derivando diversos atributos da
mesma tabela secundária 23
SQL pushback usando mais de uma
tabela secundária 20
upload de dados 18

R

resolução de problemas
assinando atualizações de Suporte 6
coletando dados 9
entrando em contato com o Suporte
IBM 4
identificando problemas 1
lista de verificação 9

resolução de problemas (*continuação*)
melhores práticas de desempenho 17
obtendo correções 3
pesquisando bases de
conhecimento 2
problemas conhecidos 11
procurando por resoluções de
problemas 2
resolução de problema sistemático 1
trocando informações com o Suporte
IBM 5

S

Suporte IBM
assinando atualizações 6
entrando em contato 4
enviando e recebendo informações 5

