



IRI CoSort
Sort, Transform & Report

V10 Product Overview



CELEBRATING
40 YEARS
2018
The CoSort Company

Uma introdução para o CoSort

Desde 1978, CoSort atende às crescentes necessidades de manipulação de dados das empresas com instalações de arquivos simples, banco de dados e data warehouse de alto volume. O CoSort também é uma solução preferida para classificação herdada e migrações de dados para Unix e Windows. O IRI trabalhou para tornar o CoSort o produto de classificação comercial mais amplamente licenciado em sistemas abertos e é fortemente focado no desenvolvimento de tecnologias relacionadas à manipulação de dados.

O CoSort agora é uma solução de melhoria de desempenho para muitas aplicações e uma passagem única para plataforma de grande escala:

Data Processing	Transformação, Migração, Limpeza, Processamento de Dados, etc.
Data Presentation	Apresentação e Relatório de Apresentação de Dados
Data Protection	Proteção de dados Mascaramento, criptografia, etc.
Data Prototyping	Teste de prototipagem de dados Geração de dados

Data Processing

O programa CoSort Sort Control Language (SortCL) pode executar transformações de dados paralelos para integrar, preparar e converter grandes volumes de dados. Em apenas um passe de E / S e script de tarefa, o SortCL pode:

selecionar, classificar / mesclar, ingressar, pesquisar, converter tipos de dados e endian, formatos de arquivo, re-mapear / reformatar, pivô, sequência, calcular, agregar, gerenciar sub-strings, limpar, criptografar, desidentificar e realizar transformações complexas

Origens e destinos incluem arquivos compactados, simples e de índice, pipes, tabelas via Open Database Conectividade (ODBC) e procedimentos personalizados.

Data Presentation

Os usuários do SortCL podem gerar os resultados dos processos acima em um ou mais detalhes e relatórios de resumo. Os usuários podem combinar junções, cálculos cruzados, pesquisas de hash e seleção condicional para gerar relatórios e subconjuntos formatados para: operações de cobrança, segmentação de clientes, captura de dados alterados, análise de dados forenses e negócios para ferramentas de inteligência. A formatação pode incluir layouts, campos e cabeçalhos especiais de arquivos, números de página, valores do ambiente, tags HTML incorporadas (para postagem na web) e a conversão de dados em CSV ou XML para transferências para ferramentas analíticas e de BI.

Data Protection

SortCL e o produto de máscara de dados derivado IRI FieldShield, podem proteger dados confidenciais no campo, com base nas regras de negócios. As funções incluem preservação de formato AES de 256 bits, técnicas de criptografia e desidentificação e mascaramento de dados para anonimizar, ofuscar, pseudonimizar ou redigir campos. Funções adicionais de criptografia ou segurança também estão disponíveis através de transformações de campo personalizadas. Protegendo dados no nível do campo (durante ou após o processamento e apresentação) é mais rápido e deixa disponíveis dados não sensíveis de arquivos, discos e bancos de dados.

Data Prototyping

SortCL e o produto de dados de teste derivado IRI RowGen, podem criar ou selecionar aleatoriamente o campo de teste dados e exibí-los em arquivos reais (de produção) e formatos de relatório. Você pode criar qualquer número, tipo e tamanho de arquivos, registros e faixas de valores necessários para simular com segurança a realidade e aplicações de teste de estresse. Os usos incluem banco de dados e população de ferramentas ETL, benchmarking, desenvolvimento de aplicativos e terceirização.

Usuários e usos do CoSort

O CoSort é um processador de dados sequenciais de alto desempenho e uso geral em uma variedade de formatos. Também serve como plataforma de migração para dados herdados e programas de classificação, além de suporte a operações de inteligência de negócios, ETL e governança de dados. CoSort reduz tempos de execução, riscos e complexidade para uma variedade de usuários de TI e de negócios.

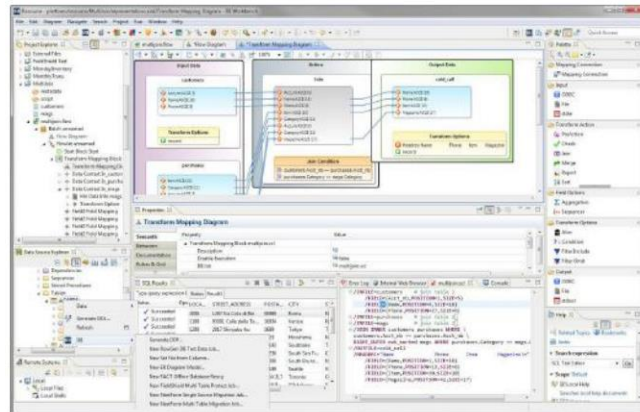
Função do Trabalho	Resultados e benefícios do CoSort
IT Manager	<ul style="list-style-type: none"> • Ferramentas de migração e modernização de software de classificação herdada • Formato universal de arquivo simples e tipo de dados, recursos de conversão • Relatório em lote de detalhes e resumo com painel opcional • Codifica e executa mais rapidamente que os trabalhos Perl, shell, SQL, ETL e COBOL • Preços acessíveis e políticas flexíveis de licenciamento
DBA	<ul style="list-style-type: none"> • Pré-classificações paralelas melhoram o desempenho de carregamento, reorganização e consulta • Classificação combinada, associação e agregação de transformações offline • Relatórios externos de lote e delta mais rápidos / fáceis que o SQL • As pesquisas de arquivos simples oferecem soluções discretas, offline e um para muitos • Metadados compartilhados com extração rápida (FACT), FieldShield e RowGen
Data Warehouse (ETL) Designer	<ul style="list-style-type: none"> • Aceleradores de classificação de plug-in para DataStage e Informatica • Transformações multithread no sistema de arquivos para armazenamento temporário de dados • Seleção complexa e lógica de expressão para integração de dados • Metadados abertos e fáceis e conversores de interface com ferramentas existentes • Proteção integrada, transformações personalizadas e procedimentos de entrada / saída
BI Architect	<ul style="list-style-type: none"> • Agregação e filtragem rápidas de dados (grandes) para ferramentas de BI

	<ul style="list-style-type: none"> • Relatório incorporado com muitas funções de formatação, incluindo PCRE • As proteções em nível de campo permitem relatórios de segmentação compatíveis • O log da Web e o tratamento de dados IPA facilitam a análise do fluxo de cliques • Relatório de alteração de captura de dados (delta) usando junções e seleção
CISO, Data Governance or Compliance Officer	<ul style="list-style-type: none"> • Anonimização em nível de campo, cancelamento de identificação, criptografia, pseudonimização • As funções de proteção podem ser executadas nos trabalhos de transformação e relatório • O log de auditoria XML pronto para consulta dos detalhes da tarefa ajuda a verificar a conformidade • Melhorias de qualidade e segurança dos dados mestre mantidas em arquivos simples • Suporte para muitos objetivos de controle COBIT
Application Developer (ISV)	<ul style="list-style-type: none"> • Bibliotecas de API seguras para threads para classificações paralelas incorporadas, transformações • Chamadas em sistema serial e paralelo ao programa SortCL • O acesso às bibliotecas de criptografia incluídas protege os fluxos de dados em tempo real • Recursos internos de geração de dados de teste (funcionalidade RowGen) • Licenciamento acessível, personalizado para modelos de negócios individuais



Aplicativos Compatíveis

Além do CoSort, muitos outros produtos IRI criam ou alavancam os metadados do CoSort Language SortCL para definir os layouts e manipulação de dados nos vários contextos adequados à finalidade descritos abaixo.



Observe que o CoSort e todos os produtos listados, também incluem e compartilham o mesmo gráfico gratuito ambiente de desenvolvimento integrado (IDE) para (Baseado em SortCL) design de trabalho e agendado implantação chamada [IRI Workbench™](#).

Construído no Eclipse™, o IRI Workbench facilita a especificação, execução, ajuste e manutenção de Scripts de tarefas SortCL por meio de assistentes de criação de tarefas e definições de metadados, fluxo de trabalho e transformação, diagramas de mapeamento e um editor com reconhecimento de sintaxe com estrutura de tópicos de suporte ao diálogo para especificação manual. O IRI Workbench também fornece acesso, visualização e integração ao banco de dados com o SortCL e outras tarefas. Isto inclui extensões para contribuições da equipe, controle de versão da tarefa e execução e dados remotos do sistema.

IRI Voracity® Data Management	Consolida a descoberta, integração (ETL, CDC, SCD), migração, governança e análise de dados em grandes e pequenas empresas fontes no local ou na nuvem. Uma edição Hadoop do Voracity também pode converter e executar de forma intercambiável muitos CoSort tarefas SortCL no MapReduce 2, Spark, Spark Stream, Storm ou Tez. O Voracity inclui assistentes para o CoSort e todos os produtos abaixo.
IRI FACT™ Fast Extract	Descarrega tabelas de banco de dados muito grandes (VLDB) em arquivos simples em paralelo para arquivamento offline, transformação de dados, geração de relatórios, migração e recargas. O FACT pode trabalhar com metadados e tubos com o

	CoSort SortCL para executar reorganizações, replicação, operações de criptografia, ETL e BI, tudo em uma única passagem de E / S.
IRI FieldShield® Data Masking	Desidentifica informações de identificação pessoal (PII) e outros dados confidenciais em tabelas de banco de dados conectadas a ODBC ou arquivo popular formatos com múltiplas funções, incluindo: criptografia, desfoque, hash, pseudonimização, randomização e tokenização.
IRI RowGen® Test Data	Usa a sintaxe SortCL para criar dados de teste seguros em arquivos reais, relatórios e formatos de tabela para protótipos, DevOps, benchmarking etc.
IRI NextForm® Data Migration	Converte tabelas de banco de dados, formatos de arquivo, dados de campo e tipos de endian para projetos de migração de dados, banco de dados, aplicativo e plataforma.

A seguir, exemplos de produtos de terceiros com os quais o IRI mantém vários níveis de compatibilidade para aprimorar seu desempenho operacional ou recursos de proteção de dados por meio da funcionalidade CoSort SortCL:

Oracle	<p>O CoSort pode originar e direcionar tabelas Oracle via ODBC ou OCI para o arquivo sistema (usando FACT e SQL * Loader). Com esses dados, o CoSort pode:</p> <ul style="list-style-type: none"> • Transformar (classificar, unir, agregar, reformatar etc.) • Capturar dados alterados (gerar relatórios delta) • Proteger (criptografia no nível da coluna, mascaramento etc.) • Executar pré-classificações de índice (na tecla da tabela mais longa) <p>Ao pré-classificar, o CoSort pode melhorar a velocidade e a eficiência do SQL *</p> <p>Operações do Loader (e outro utilitário de carregamento RDB) e, portanto, reorganiza e consulta em tabelas pré-CoSorted. O CoSort e trabalhos relacionados também podem ser acionados ou invocados pelo Oracle Job Scheduler.</p>
IBM DB2	Tabelas DB2 de origem e de destino via ODBC ou o sistema de arquivos (usando FACT e DB2 Load), para executar as tarefas acima. Além disso, o CoSort Load Accelerator para DB2 (CLA4DB2) acelera as cargas em massa até o UDB Versão 9.5.
MS SQL Server	Executa muitas das mesmas funções que o SQL no SortCL, apenas 10x mais rápido.
MongoDB	Use SortCL para transformar, migrar, mascarar ou criar protótipo de coleções do MongoDB.

IBM InfoSphere DataStage	Execute o programa CoSort SortCL em um estágio de arquivo sequencial do DataStage ou como uma sub-rotina antes do trabalho para trabalhos maiores de classificação, ingresso e agregação. SortCL transforma dados em uma passagem externa rápida que não requer particionamento ou malabarismo de memória. SortCL também pode combinar essas transformações com conversão de dados, limpeza, mascaramento, geração de relatórios e carregamento.
Informatica PowerCenter	Envie as maiores transformações de dados para os programas CoSort SortCL no sistema de arquivos como tarefas de comando do fluxo de trabalho. Essa abordagem é mais eficiente do que a "otimização de empilhamento" recomendada pela Informática Oracle e muito mais barato que Teradata, Ab Initio, DMExpress, etc.
BIRT	Organizar dados para o BIRT (Eclipse Business Intelligence Reporting Tool) usando CoSort no IRI Workbench. O BIRT ingere dados preparados pelo SortCL em tempo de relatório, aproveitando uma fonte de dados IRI compatível com ODA.
Splunk	Indexar Splunk imediatamente com dados preparados e protegidos por SortCL usando IRI add-on for Splunk .
Proxy Coupling	Transfira sem problemas as etapas de classificação da JCL do z / OS para SortCL no z / Linux ou outro. Plataforma de 'sistema aberto' através da tecnologia Proximal Systems "PSCsort".

Conteúdo CoSort

O pacote **CoSort** contém utilitários independentes, metadados de layout de arquivo e conversores de parâmetros de classificação, substituições de classificação de terceiros, bibliotecas de API e documentação. Os principais **programas utilitários** são:

- SortCL** - o programa de Linguagem de Controle de Classificação de quarta geração para definir dados e manipulações com sintaxe e semântica familiares à classificação de mainframe e SQL Comercial. A interface mais abrangente do pacote CoSort, SortCL combina várias funções de transformação de dados (classificação, junção, agregação, filtragem, remapeamento, etc.) com limpeza, geração de relatórios e mascaramento para: comparação de arquivos e captura de dados alterados, tipo de dados e conversões de formato de arquivo, integração e armazenamento temporário de dados (ETL), relatórios e análises de BI organização de dados, delta e resumo, além de migração de dados e

bancos de dados, replicação e federação, bem como a conformidade com as leis de privacidade de dados.

- **Sort**- um substituto para o comando de classificação do Unix, que roda mais rápido e dimensiona linearmente. Isto é executado em todas as plataformas Unix e Windows (unixsort.exe).

SortCL reconhece variáveis de ambiente e suporta pipes e fluxos de dados intermediados para permitir dados fluir entre processos sem E / S adicional. SortCL também pode ser personalizado com a saída do usuário, procedimentos para critérios especiais de entrada, saída ou comparação. Criação, modificação, compartilhamento de trabalho do SortCL e a execução é suportada no IRI Workbench.

Esses **conversores de metadados** aproveitam as informações existentes sobre o layout da fonte de dados:

- **cob2ddf** converte layouts de copybook COBOL em definições de dados SortCL
- **csv2ddf** converte Microsoft .csv. Cabeçalhos de arquivo para definições de dados SortCL
- **ctl2ddf** converte layouts de arquivo de controle do Oracle SQL * Loader para definições de dados SortCL
- **elf2ddf** converte os logs da Web no “Extended Log Format” do W3C para definições de dados SortCL
- **ldif2ddf** converte layouts LDIF em definições de dados SortCL
- **xml2ddf** converte formatos XML para definições de dados SortCL
- **odbc2ddf** converte layouts de tabela de banco de dados em um arquivo de definição de dados SortCL (.ddf)

Os utilitários acima também são suportados graficamente no IRI Workbench, que também produz DDF para Fontes JSON e MongoDB (o desenvolvimento do utilitário de linha de comando está pendente). Além do que:

- **MIMB** (do MITI) converte muitos layouts de arquivos de aplicativos em definições de dados SortCL
- **Mapping Manager** (do AnalytiX DS) cria dados SortCL e metadados de trabalho do zero, ou converte nele a partir de mapeamentos de ferramentas ETL de terceiros.

Esses utilitários de conversão de parâmetros de classificação facilitam as migrações de classificação herdadas:

- **mvs2scl** converte cartões de classificação MVS JCL para especificações de trabalho SortCL
- **sorti2scl** converte os parâmetros SortI em especificações de trabalho SortCL
- **vse2scl** converte cartões de classificação VSE JCL para especificações de trabalhos SortCL

As seguintes **substituições** de classificação de terceiros estão disponíveis no pacote CoSort:

- **acu-cosort** - um substituto para os verbos de classificação fornecidos com o ACUCOBOL-GT
- **cla4db2** - o "CoSort Load Accelerator for DB2" substitui a rotina de classificação da IBM no UDB 5-8 carregador, além de duplicar a produtividade no Unix
- **mf-cosort** - um substituto para os verbos de classificação fornecidos com o Micro Focus Net Express, Server Express e Workbench no Unix e Windows. Os usuários do COBOL podem vincular estaticamente e dinamicamente para mfcosort e o mecanismo CoSort para acelerar a velocidade de classificação e reduzir temporariamente, classifique o espaço em novos executáveis ou em um RTS completo
- **nat-sort** - um substituto para o verbo de classificação no Software AG natural
- **proc-sort** - os usuários do sistema SAS 7-9 podem vincular dinamicamente às bibliotecas cosort() compartilhadas para substituir a função de classificação no SAS em sistemas Unix

O CoSort fornece recursos semelhantes de substituição do tipo drop-in para o Tetrad OPX e o Mainframe UniKix Gerenciador de lotes. Os ganchos CoSort também estão disponíveis para a solução ETI, Cincom Control: Manufacturing, Dynamic Information Warehouse de Kalido e suíte TeraStream ETL da DataStreams (Coréia).

Para desenvolvedores e fornecedores de software independentes (ISV), o CoSort também inclui dois aplicativos, Bibliotecas da Interface de programação (API):

- **cosort_r()** - uma versão segura da thread da API cosort() original que permite várias rotinas de operações de classificação / mesclagem para ocorrer na mesma passagem pelos dados. O mecanismo da corotina permite transferências de registros na memória entre programas e a classificação.
- **sortcl-rotina()** - a biblioteca SortCL segura para threads que permite que os programadores explorem toda a gama de comandos CoSort Sort Control Language em seus próprios programas.

As APIs do CoSort permitem definir qualquer entrada (seleção), comparação (sequência de pedidos) ou saída (relatórios) critérios, permitindo que os aplicativos realizem tarefas complexas em uma única passagem de E / S. Você pode escrever chamadas para biblioteca em qualquer idioma que possa vincular a uma biblioteca C, como C ++, COBOL, VB, Java, etc.

Por fim, o pacote CoSort também inclui a seguinte **documentação**:

- **Guia de instalação** - carregamento, licenciamento e configuração específicos da plataforma
- **Manual** - uma documentação completa do usuário e do programador para todas as interfaces acima
- **Exemplos de tarefas** - exemplo de scripts de tarefas SortCL, conversões de metadados e chamadas de API
- **Práticas recomendadas** - dicas e truques para script e ajuste de tarefas

O IRI Workbench também fornece material on-line para usuários do CoSort da seguinte forma:

- **Visão geral da plataforma** - uma cartilha básica na seção Bem-vindo para fins de orientação
- **Referência** - um conjunto de conteúdo com hiperlinks disponível em "Guias do usuário das ferramentas"
- **Folha de dicas** - um assistente interativo passo a passo de "Introdução" criado no Eclipse
- **Ajuda da caixa de diálogo** - instruções específicas do contexto para cada parâmetro do trabalho em todas as páginas