

2.3. Engine de Processamento Paralelo Unificado

O trabalho executado pelo Information Server é realizado dentro do engine de processamento paralelo. Este engine manipula necessidades de processamento de dados diversos assim como executa as análises requeridas pelos componentes InfoSphere Information Analyzer, higienização pelo InfoSphere QualityStage, e ainda transformações complexas pelo componente InfoSphere DataStage. Este engine é desenhado para:

- Paralelismo pipelining para tratar grandes volumes de dados e diminuir a janela de carga;
- Por adição de hardware escalar sem necessidade de alteração nos processos desenvolvidos.

Com o paralelismo pipelining é possível executar simultaneamente processos de transformação, limpeza e carga. Como uma linha de produção o Engine move linhas de processo a processo: ele inicia o processo de gravação enquanto o processo de leitura ainda está em execução.

Vantagens:

- Reduz a utilização de disco (áreas temporárias);
- Evita a ociosidade dos processadores;
- Suporta todas as configurações de hardware (Single Processor, SMP, MPP, Cluster, Grid);
- Aumenta a utilização de recursos pela simples adição de processadores ou nós de processamento sem a necessidade de alteração dos processos desenvolvidos.

Revision #2

Created 27 July 2022 17:54:17

Updated 5 August 2022 23:19:12