IBM Blu Acceleration

Presenter: Lavanya Subramanian

What is Blu?

- A set of optimizations to make DB2 faster
- In-memory optimized
- Leverages IBM’s competence with hardware

Key Features

- Delayed Materialization
- Exploiting Single Instruction Multiple Data (SIMD) Instructions
- Hardware-aware Parallelization
- Scan-friendly Caching

Delayed Materialization

- ***Ability to operate on compressed data***

- Compression schemes employed
  - Huffman and offset encoding

- ***Compression process preserves order***
  - Same order as the uncompressed data

- Enables optimized comparisons, scans

Exploiting SIMD Instructions

- SIMD: Single Instruction Multiple Data
  - ***Same instruction on multiple data elements***

- Blu capable of detecting SIMD hardware support

- ***Packs data into CPU registers to exploit SIMD***

Hardware-aware Parallelization

- Parallelization benefits limited by memory access latency

- Map threads to cores aware of
  - CPU-cache affinity
  - Memory affinity

Scan-aware Memory Management

- All data doesn't always fit into memory
- Problem with LRU memory management policies
  - Fall off a cliff when data does not fit in memory
- ***Employs a page replacement policy that preserves part of the working set***