MLD-B Case study (Zipper ahead)

Lect 1: Intro

- Main data storage, work in Mem.
- Computation
- Lot long & fast job
- Must use disk as main mem, data operations access disk with some overhead

Lect 2

- Sort: Specified system
- Focus of MLD database.

Sort Type Based System

Problems for MLD

(1) MVP (Minimum Value Product)
- Minimum Requirements
- Strategy or: Sorting Merge
  Hashing Merge
- Detailed Requirement
  - on Sorts
  - on Joins
  - Merge

(2) Extension of Sorts
- Strategy (Sorts)
- Sort of Source prop / Sort of destination
- Read Merge / Write Merge
  Read without (Read Write)

(3) Design Tools
- Sort Generator
- Sort Optimizer

(4) BSC Algo
- per-(inner, outer)
- optimize Cluster

(5) Global Strategy Implementation

Lect 3: MVP

Requirements
- Sort equipment for fast out/input
  - for fast data, easily extends set-tuple
  - for Sorter prefer property

(1) Fixed S equipment
(4) The S sort-Equipment should be the profile
(5) Sorting equipment with minimum S buffers

Detailed Requirements
(1) Operational Characteristics/ equipment
(2) Analysis & Measures

Step 1: Sort Generation

(1) Sorting
(2) Sorting Generation
  - Sorting Steps or Phasing
  - prefer: (Big S, Small, Run)

Step of time: CPU work
Step of time: Disk Access

- S Bytes: Sorted per day
- Step of time
- Are Bytes: Sorted per day

Times: 100 Bytes

Step 2: Sort Optimizer

(1) Assumptions
- Bytes
- 1000 Bytes = 1 KB
- 1000KB (1 MB or 1000 Bytes) = 10^3 Bytes. Million Bytes
- 1000 MB = 1 GB = (1000 x 1000 KB) = Billion Bytes
- 1000 GB = 1 TB

2^20 Bytes X 100 Bytes
(2^20 X 2^10 KB) = 2^30 KB / Day

In Bytes:
(2^20 X 2^10 KB) / 24 = 2^20 KB / Day

2^20 / 24 = 0.04 GB/sec

0.04 GB X 3600 sec
= 144 GB

Sort Generation: Sorting Steps or Phasing
- Sort Generation
- Sort Optimizer
- Sort Equipment
- Sort Generation
- Sort Optimizer
- Sort Equipment
- Sort Generation
- Sort Optimizer
- Sort Equipment

Sort Generation
- Sort Optimizer
- Sort Equipment

Sort Generation
- Sort Optimizer
- Sort Equipment