

Badger

Performance of MySQL on modern computer architectures

In the context of Gelato:

1. Analyzing Performance on Itanium-2/Linux
2. Improving Performance (when possible)
 - a. How to best configure/tune MySQL?
 - b. New tuning parameters (e.g., prefetch size)
 - c. New mechanisms (e.g., read ahead prefetching)

Rough Analysis

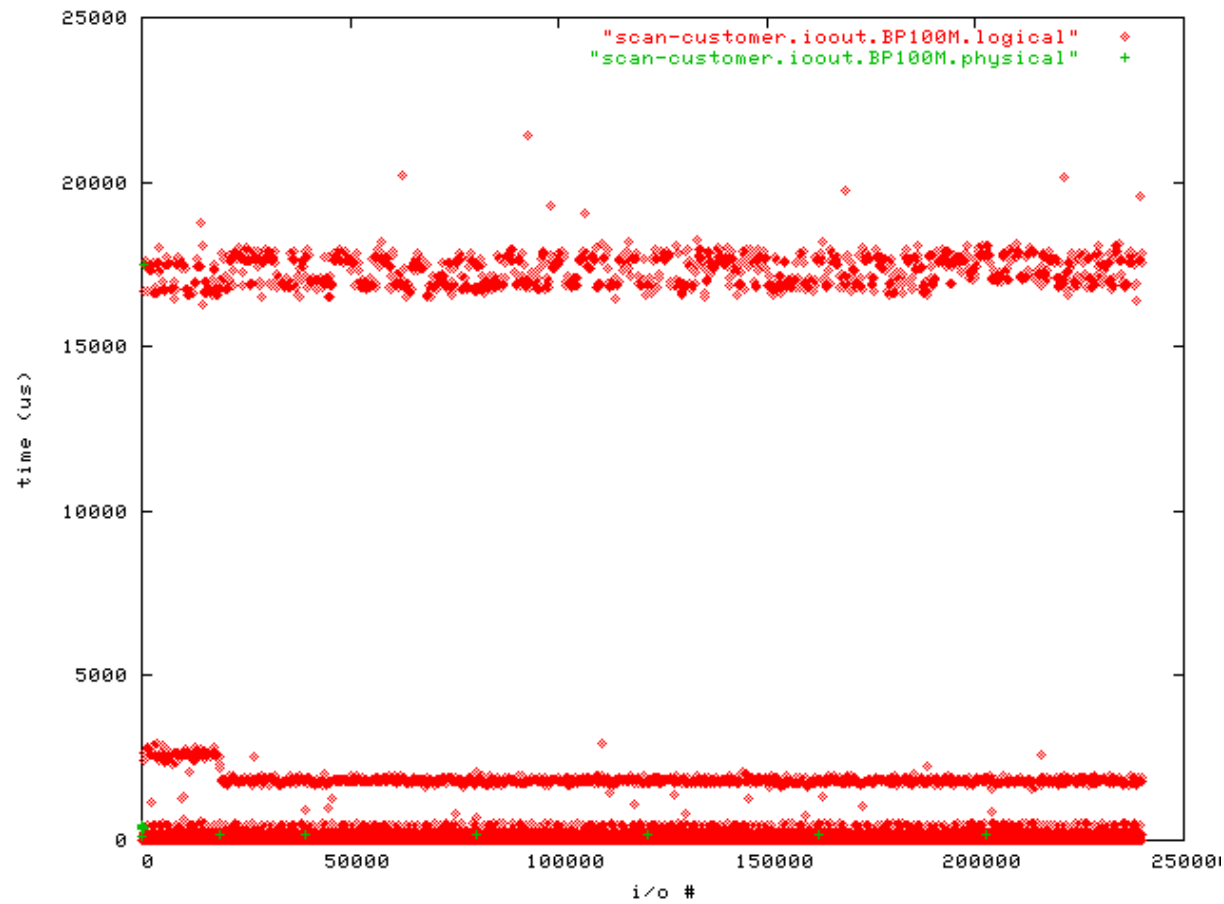
1. CPU is fast thus,
Disk IO Performance is critical
(*seeks are death*)
2. RAM is large,
L3 Cache is large thus,
In-memory Performance is critical



Analyzing Performances

- Internal Analysis
 - Instrumenting MySQL:
Where is the time spent?
 - Physical, logical Ios
 - Locks
 - Program Path Analysis
- External Analysis:
 - IO Subsystem: Nb of pending requests, Disk Scheduler Performance
 - CPU: Inspecting performance registers (Vtune, Oprofile)
- Simple set of experiments for micro-benchmarking MySQL:
 - Cost of parsing vs. query optimization vs. query execution
 - Disk Subsystem Utilization
 - Ratio cost of physical vs. logical IOs
 - Cost of operations in-memory
 - Cost of latching / locking
 - CPU Utilization

MySQL Instrumentation



(customer table)

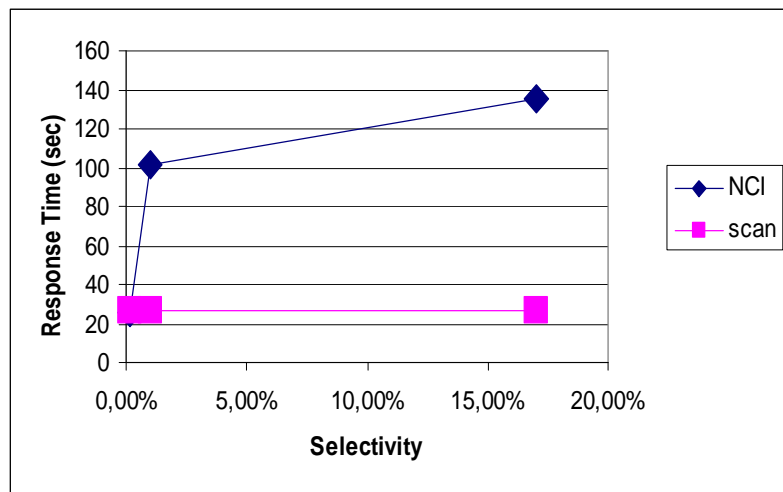
Disk Subsystem Utilization

- Asynchronous IOs on Linux. Optimal disk performance (PIII, 10000rpm SCSI):
 - Sequential ops: 4 outstanding requests
 - Random ops: 2000 outstanding requests
- Micro-benchmarking of MySQL shows:
 - Sequential reads: Nb of outstanding requests falls to 0 repeatedly. Read ahead prefetching is the culprit.
 - Random reads/writes: Nb of outstanding requests not higher than 4.
- Improving Performances:
 - Eager read ahead mechanism (with prefetch size as a tuning option)
 - Prioritized IOs in Linux to control latency while achieving high throughput when increasing the number of outstanding requests.

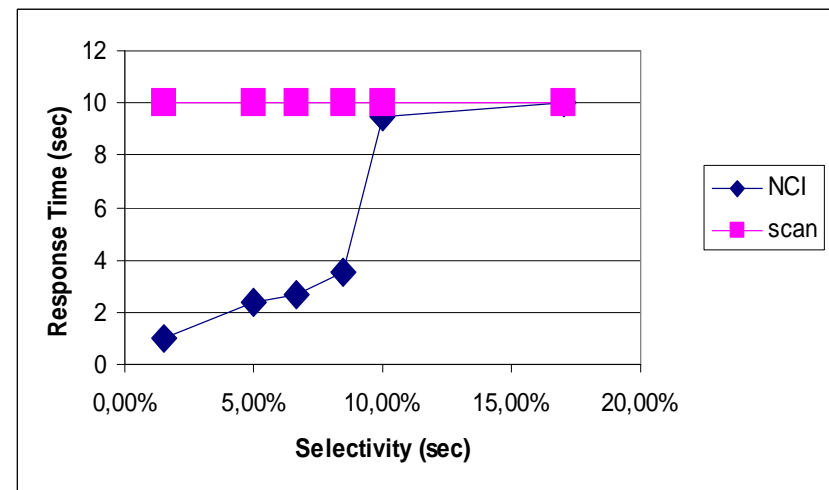
Scan vs. NCI (MySQL)

What is the query selectivity at which a non-clustered index starts to pay off for a range query?

- Cold buffer
Data on disk



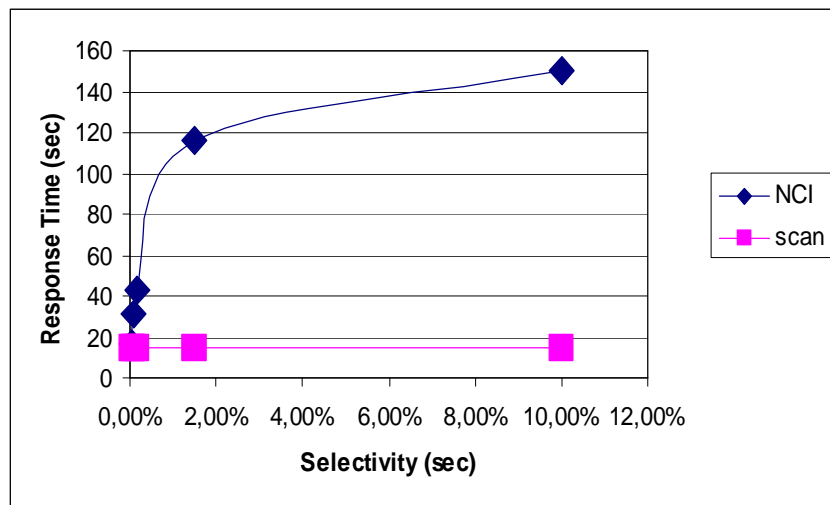
- Warm buffer
Data in memory



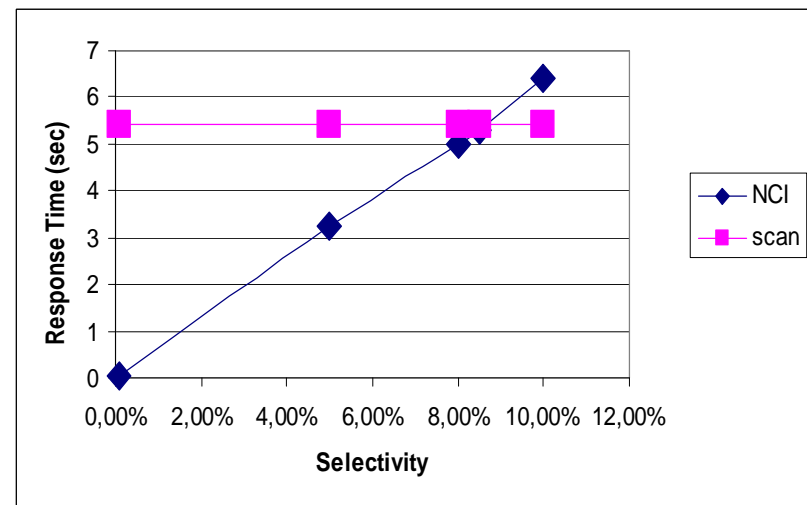
Scan vs. NCI (Oracle)

What is the query selectivity at which a NCI starts to pay off?

- Cold buffer
Data on disk



- Warm buffer
Data in memory



Badger Activities

- Profile-Driven Compilation on Itanium-2
 - Traces representing OLTP/OLAP workloads
- Disk Subsystem Utilization
 - Utilizing IO Bandwidth
 - Programming the SAN front-end
- In-Memory Performance