

Informations- & Datenbanksysteme

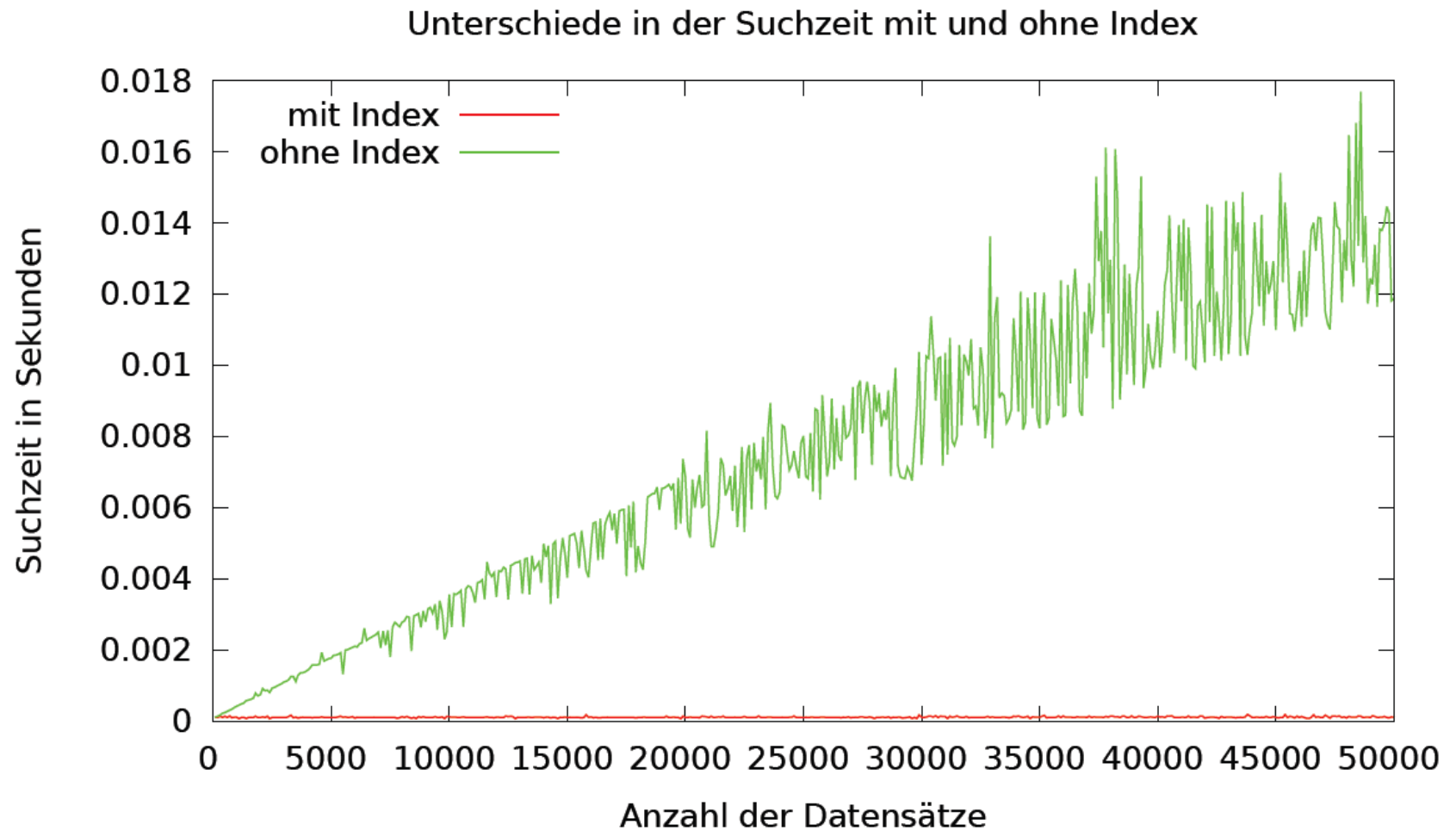
Indices
(Adams Kap. 6)

Motivation in Worten

In aller Kürze:

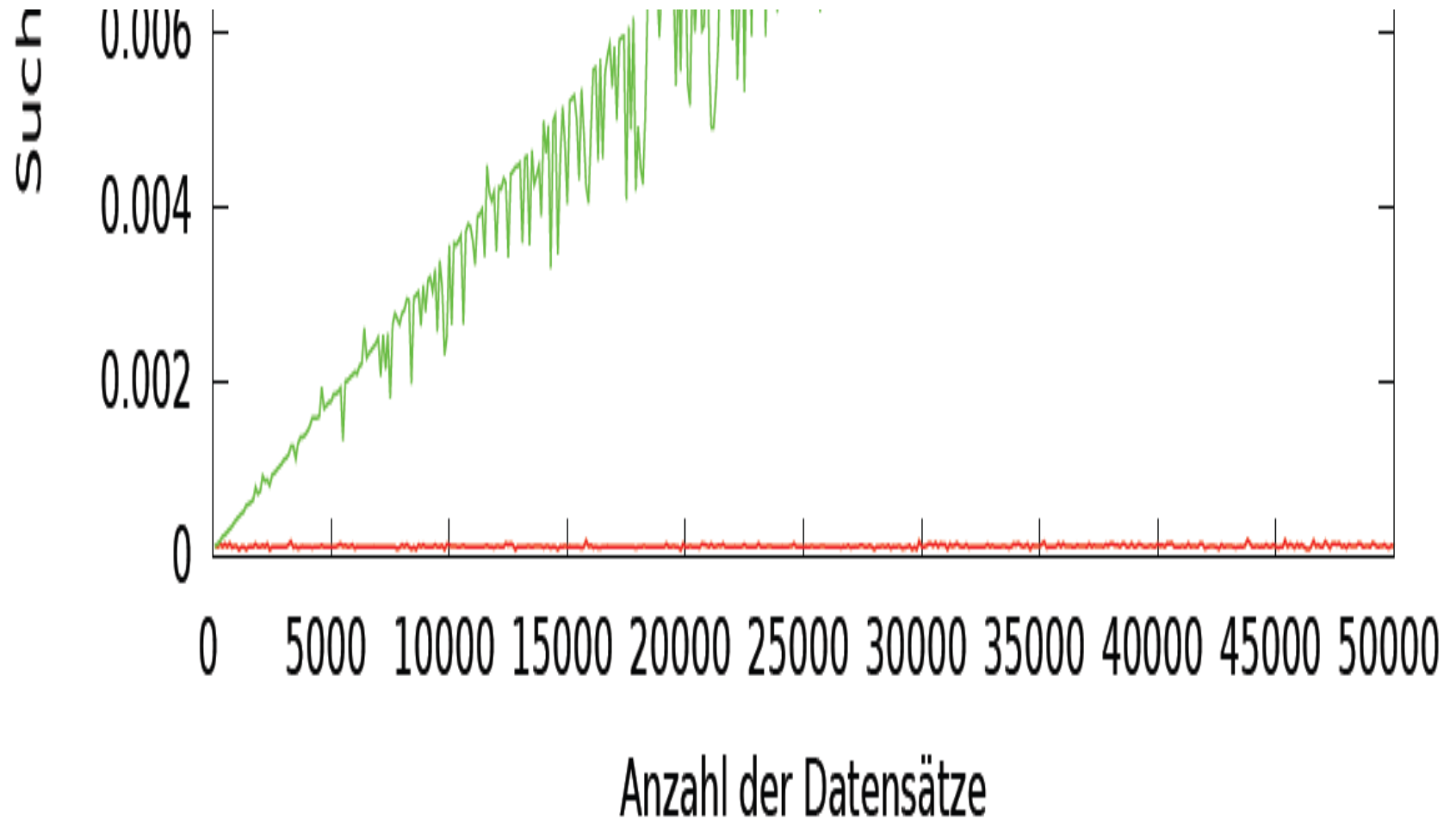
- Handling grosser Datenmengen beschleunigen
- Duplikate/Dubletten verhindern

Motivation in Bildern



Quelle: Adams

Motivation in Bildern



Quelle: Adams

Wie funktioniert das?

1. Für die effektiven Daten wird ein Fingerabdruck berechnet (z.B. ein Hash)
2. Fingerabdrücke und Verweise werden verknüpft
3. Tupel wird in effizienten Datenstrukturen abgelegt (z.B. B-Tree oder sortierte Listen)

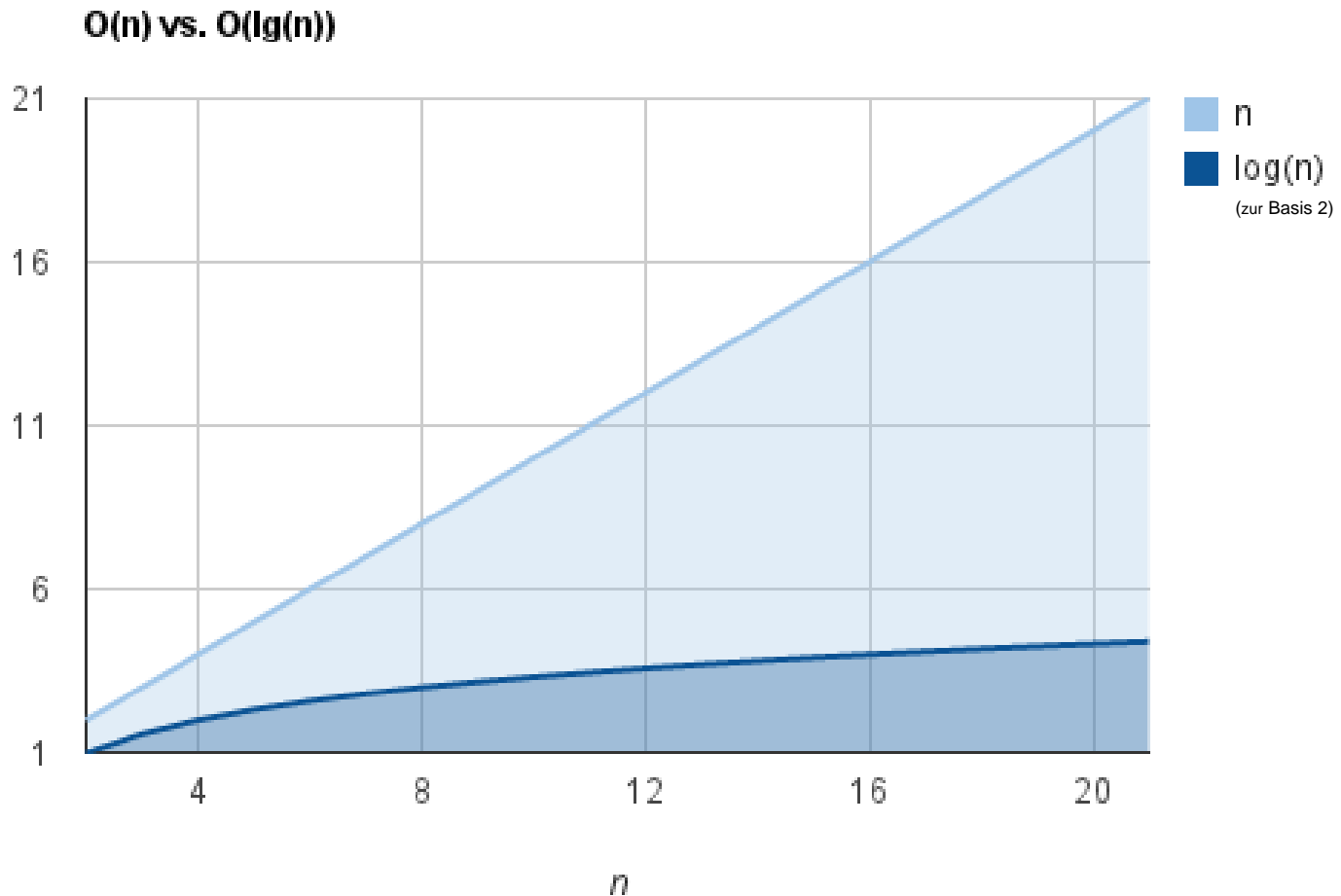
MySQL verwendet B-Tree für InnoDB und MyISAM

<http://dev.mysql.com/doc/refman/5.7/en/innodb-physical-structure.html>

Warum wird das schneller?

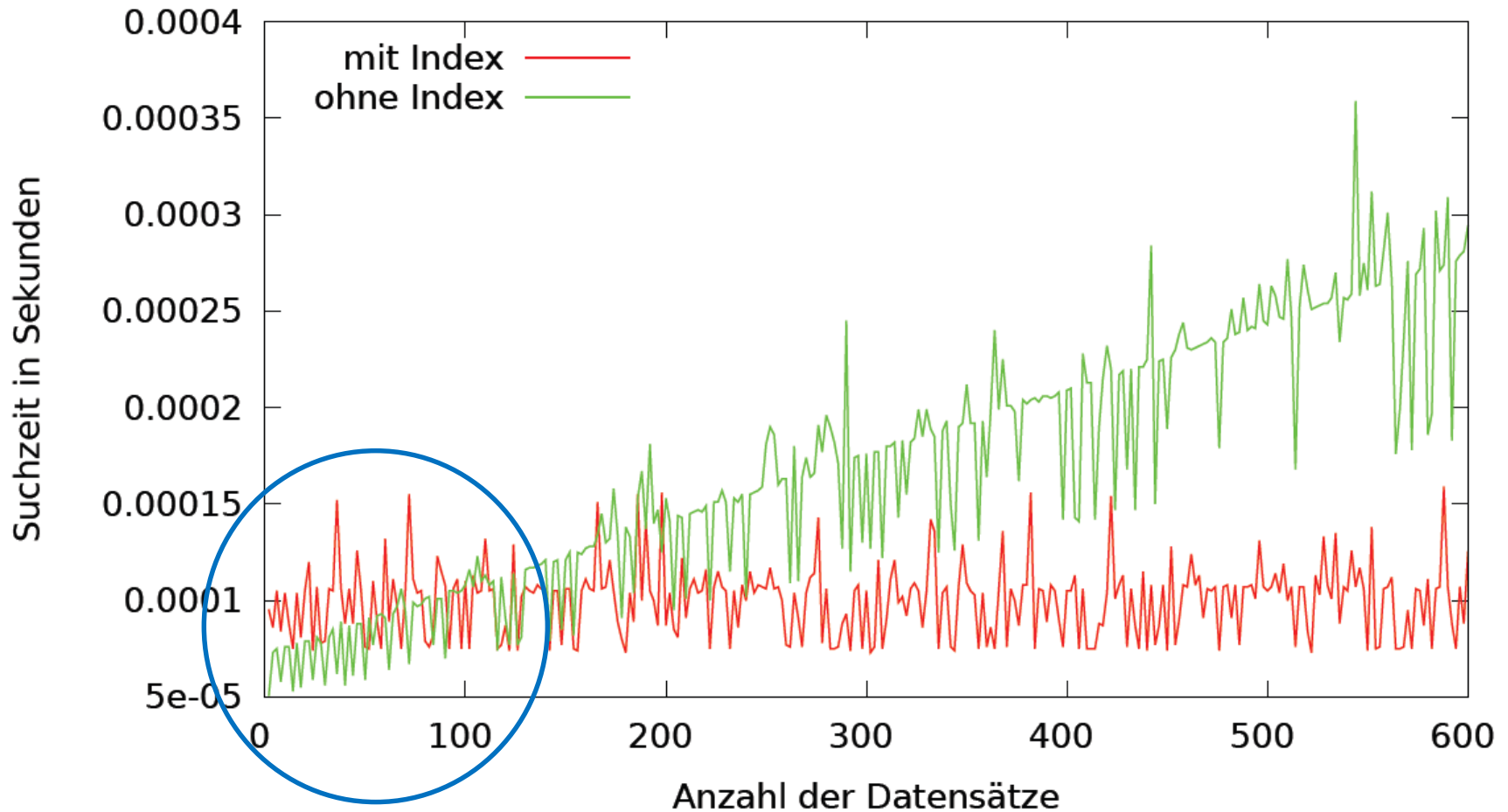
- Effektive Daten u.U. zu gross für Hauptspeicher
- Fingerabdrücke passen eher in Hauptspeicher
- Effiziente Datenstrukturen erlauben Suchen mit $O(\log(n))$
- Es müssen typischerweise nur Bruchteile der Information verglichen werden

Warum wird das schneller?



Nachteile?

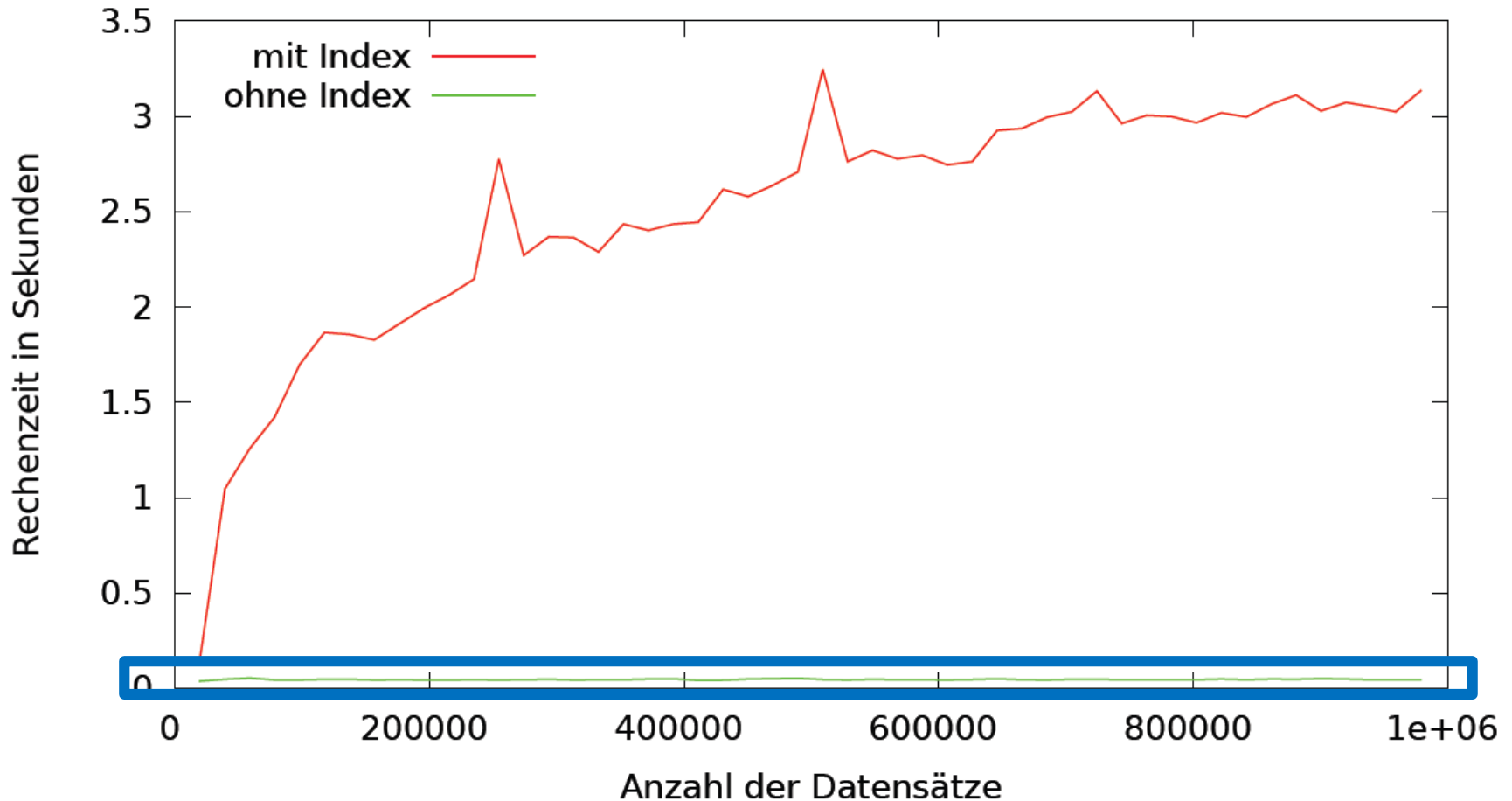
Unterschiede in der Suchzeit mit und ohne Index bei wenigen Datensätzen



Quelle: Adams

Nachteile?

Einfügen von ca. 20000 Datensätzen mit und ohne Index



Quelle: Adams

DDL?

- SHOW INDEX FROM <tabelle>
- CREATE [UNIQUE] INDEX <name> ON <tabelle>(<feld>[(<länge>)]);
 - UNIQUE erlaubt das Erzwingen von Schlüsseleigenschaften
- DROP INDEX <name> ON <tablle>;
- **Anmerkung:** Primärschlüssel werden immer automatisch indexiert.

Selektivität?

- Werden Indexe auf einen Teil des Spalteninhalts beschränkt, können Kollisionen entstehen

$$S = \frac{\textit{Unterscheidbare Einträge}}{\textit{Gesamtzahl der Einträge}}$$

- Optimal: Selektivität von 1
- Relevant bei grossen Mengen an Zeichenbasierten Indexen

Worauf wirken Indices?

- Suchen von Datensätzen
(SELECT/WHERE)
- Sortieren von Datensätzen
(ORDER BY)
- Verknüpfen von Datensätzen
(JOIN)

Fazit

- Indexe beschleunigen die Suche ab einer bestimmten Menge von Datensätzen
- Indexe können andere Operationen als die Suche beeinträchtigen
- Indexe können unterschiedlich selektiv sein
- Indexe müssen von Hand erstellt werden

Todo

- Selbstständiges Lösen der Aufgaben in Kapitel 6 von Adams (Seite 89-100)
- Weitere Übung steht auf Moodle zur Verfügung