Collection of $10^9$ records. Which data structure is better?

**AVL tree**

Height: $O(\log_2 n)$

$\log_2 10^9 \approx 30$

**B-tree of degree $10^3$**

Height: $O(\log_{1000} n)$

$log_{1000} 10^9 = 3$

Number of comparisons to find a key

$\approx 30$
Memory Hierarchy

- **Volatile memory**
  - Registers (100s of bytes)
  - Cache (Few Mega bytes)
  - Main memory (Few Giga bytes)

- **Persistent memory**
  - Main memory (Few Tera bytes)
  - Disk (Few Tera bytes)

- **Comparison**
  - Registers to Cache: 10-100 times slower
  - Cache to Main memory: 100 times slower
  - Main memory to Disk: 10^5 - 10^6 times slower
Collection of $10^9$ records
AVL tree

Height: $O(\log_2 n)$
$\log_2 10^9 \approx 30$

Number of disk accesses: 30

B-tree of degree $10^3$

Height: $O(\log_{1000} n)$
$\log_{1000} 10^9 = 3$

So a B-tree is much better.

Make a node fit in a block