Judy array

From Wikipedia, the free encyclopedia

In computer science and software engineering, a **Judy array** is a data structure that has high performance, low memory usage and implements an associative array. Unlike normal arrays, Judy arrays may be sparse, that is, they may have large ranges of unassigned indices. They can be used for storing and looking up values using integer or string keys. The key benefits of using a Judy array is its scalability, high performance, memory efficiency and ease of use.[1]

Judy arrays are both speed- and memory-efficient [*clarification needed*], and therefore they can sometimes replace common in-memory dictionary implementations (like red-black trees or hash tables).

Roughly speaking, Judy arrays are highly optimised 256-ary radix trees.[2] Judy arrays use over 20 different compression techniques on trie nodes to reduce memory usage.

The Judy array was invented by Douglas Baskins and named after his sister.[3]

Benefits [edit]

Memory allocation [edit]

Judy arrays are dynamic and can grow or shrink as elements are added to, or removed from, the array. The memory used by Judy arrays is nearly proportional to the number of elements in the Judy array.

Speed [edit]

Judy arrays are designed to keep the number of processor cache-line fills as low as possible, and the algorithm is internally complex in an attempt to satisfy this goal as often as possible. Due to these cache optimizations, Judy arrays are fast, especially for very large datasets. On data sets that are sequential or nearly sequential, Judy arrays can even outperform hash tables.[4] Lastly, because a Judy array is a tree, it is possible to do an ordered sequential traversal of keys, which is not possible in hash tables.

Drawbacks [edit]

Judy arrays are extremely complicated. The smallest implementations are thousands of lines of code.[3] In addition, Judy arrays are optimized for machines with 64 byte cache lines, making them essentially unportable without a significant rewrite.[4]

References [edit]

1. ^ http://packages.debian.org/wheezy/libjudy-dev
2. ^ Alan Silverstein, "Judy IV Shop Manual ", 2002
3. ^ a b http://judy.sourceforge.net/
4. ^ a b http://www.nothings.org/computer/judy/

External links [edit]

- Main Judy arrays site
- How Judy arrays work and why they are so fast
- A complete technical description of Judy arrays
- An independent performance comparison of Judy to Hash Tables

Wikipedia The Free Encyclopedia

Main page
Contents
Featured content
Current events
Random article
Donate to Wikipedia
Wikipedia store

Interaction
Help
About Wikipedia
Community portal
Recent changes
Contact page

Tools
What links here
Related changes
Upload file
Special pages
Permanent link
Page information
Wikidata item
Cite this page

Print/export
Create a book
Download as PDF
Printable version

Languages
Српски / srpski
ไทย
中文

Edit links

- A compact implementation of Judy arrays in 1250 lines of C code

Categories: Arrays