
SISTEM BERKAS



Matheus Supriyanto Rumetna & Tirsa Ninia Lina

SISTEM BERKAS

Matheus Supriyanto Rumetna & Tirsia Ninia Lina

SISTEM BERKAS

Tim Penulis:

Matheus Supriyanto Rumetna & Tirsa Ninia Lina

Desain Cover:

Fawwaz Abyan

Tata Letak:

Handarini Rohana

Editor:

Aas Masruroh

ISBN:

978-623-459-287-0

Cetakan Pertama:

Desember, 2022

Hak Cipta 2022, Pada Penulis

Hak Cipta Dilindungi Oleh Undang-Undang

Copyright © 2022

by Penerbit Widina Bhakti Persada Bandung

All Right Reserved

Dilarang keras menerjemahkan, memfotokopi, atau memperbanyak sebagian atau

seluruh isi buku ini tanpa izin tertulis dari Penerbit.

PENERBIT:

WIDINA BHAKTI PERSADA BANDUNG

(Grup CV. Widina Media Utama)

Komplek Puri Melia Asri Blok C3 No. 17 Desa Bojong Emas
Kec. Solokan Jeruk Kabupaten Bandung, Provinsi Jawa Barat

Anggota IKAPI No. 360/JBA/2020

Website: www.penerbitwidina.com

Instagram: [@penerbitwidina](https://www.instagram.com/penerbitwidina)

Telpon (022) 87355370

Kata Pengantar

Rasa syukur yang teramat dalam dan tiada kata lain yang patut kami ucapkan selain rasa syukur, karena berkat rahmat dan karunia-Nya buku yang berjudul system berkas ini telah dapat di terbitkan untuk dapat dikonsumsi oleh khalayak banyak. Umumnya berkas dan rekaman tidak dapat dilihat oleh pengguna akhir suatu sistem komputer, tetapi rekaman dan berkas tersebut merupakan komponen yang sangat penting dalam sistem database. Hal yang paling penting berkaitan dengan rekaman dan berkas adalah kinerja berkas. Data merupakan aset penting dalam organisasi. Berbagai pendapat yang sering dikemukakan oleh berbagai pakar mengenai dengan sistem komputer yang kontemporer adalah perangkat keras memiliki umur yang singkat, perangkat lunak memiliki umur yang lebih panjang. Materi ini memberikan pengenalan dasar terhadap berkas yang digunakan untuk menyimpan data dengan volume yang besar dalam sistem komputer yang kontemporer.

Minat utama dalam buku ini adalah berkas rekaman. Diasumsikan bahwa semua rekaman berada dalam sebuah berkas yang memiliki format dan ukuran yang sama dan dengan jumlah rekaman yang sangat besar sehingga tidak semua dapat ditampung di dalam pengingat komputer. Kondisi tersebut memerlukan organisasi pengingat yang hierarkis dan multilevel. Dapat diasumsikan pengingat diorganisasi secara bilevel. Pada level atas adalah pengingat utama yang digunakan untuk mengolah informasi, sedangkan pada level yang lebih rendah terdapat disk (pada umumnya disk magnetik) tempat rekaman akan disimpan. Disk pada umumnya adalah peralatan yang sifatnya diakses secara random yang memungkinkan pembacaan kembali sejumlah rekaman dengan waktu pembacaan yang tidak bergantung pada posisi dimana rekaman disimpan. Kemampuan disk untuk dapat diakses secara random tersebut memungkinkan untuk di susunnya struktur serta cara memanipulasi informasi secara efektif dan efisien.

Oleh karena itu buku yang berjudul system berkas ini hadir sebagai bagian dari upaya untuk menambah khazanah, diskusi sekaligus penguatan di bidang system berkas. Akan tetapi pada akhirnya kami mengakui bahwa tulisan ini terdapat beberapa kekurangan dan jauh dari kata sempurna, karena sejatinya kesempurnaan hanyalah milik Tuhan semata. Maka dari itu, kami dengan senang hati secara terbuka untuk menerima berbagai kritik dan saran dari para pembaca sekalian, hal tersebut tentu sangat diperlukan sebagai bagian dari upaya kami untuk terus melakukan perbaikan dan penyempurnaan karya selanjutnya di masa yang akan datang.

Terakhir, ucapan terimakasih kami sampaikan kepada seluruh pihak yang telah mendukung dan turut andil dalam seluruh rangkaian proses penyusunan dan penerbitan buku ini, sehingga buku ini bisa hadir di hadapan sidang pembaca. Semoga buku ini bermanfaat bagi semua pihak dan dapat memberikan kontribusi bagi pembangunan ilmu pengetahuan di Indonesia, khususnya terkait system berkas.

Desember, 2022

Penulis

DAFTAR ISI

KATA PENGANTAR	iii
DAFTAR ISI	v
BAB 1 PENDAHULUAN	1
A. Pengertian	1
B. Algoritma	3
C. Matriks.....	4
D. Media Penyimpanan.....	4
BAB 2 ORGANISASI BERKAS PRIMER	13
A. Latar Belakang	13
B. Organisasi Berkas.....	13
C. Medan Data	14
D. Rekaman Data	15
E. Berkas Data.....	15
BAB 3 BERKAS SEKUENSIAL	17
A. Latar Belakang	17
B. Pencarian Biner (Binary Search)	19
C. Pencarian Interpolasi.....	23
BAB 4 ORGANISASI BERKAS LANGSUNG	27
A. Kunci Sebagai Alamat Rekaman Yang Unik	27
B. Konversi Kunci Rekaman Menjadi Satu Alamat Yang Unik.....	28
C. Menentukan Alamat Dengan Konversi Kunci	29
BAB 5 MANAJEMEN KOLISI	35
A. Latar Belakang	35
B. Resolusi Kolisi	36
C. Coalesced-Hasing.....	37
D. Lich dan Eisch.....	42
E. Progressive Overflow.....	51
F. Penggunaan Buckets	55
G. Pembagian Linear	59
BAB 6 ORGANISASI BERKAS SEKUENSIAL BERINDEKS	67
A. Pendahuluan.....	67
B. Struktur Dasar.....	68
C. Penyisipan Rekaman.....	71
D. Menghapus Rekaman	73

BAB 7 PENGURUTAN REKAMAN	77
A. Pengurutan Gelembung	77
B. Pengurutan Dengan Penyisipan	80
C. Pengurutan Dengan Cepat	81
D. Pengurutan Lomuto	82
E. Pengurutan Dengan Heap (Deret)	83
F. Pengurutan Dengan Heap (Pohon Biner)	84
DAFTAR PUSTAKA	88
PTOFIL PENULIS	89



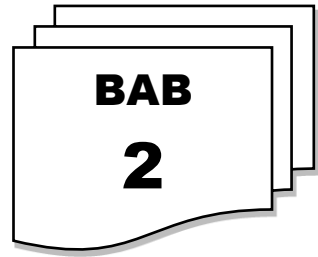
PENDAHULUAN

A. PENGERTIAN

Data merupakan aset penting dalam organisasi. Berbagai pendapat yang sering dikemukakan oleh berbagai pakar mengenai dengan sistem komputer yang kontemporer adalah perangkat keras memiliki umur yang singkat, perangkat lunak memiliki umur yang lebih panjang. Materi ini memberikan pengenalan dasar terhadap berkas yang digunakan untuk menyimpan data dengan volume yang besar dalam sistem komputer yang kontemporer.

Minat utama dalam buku ini adalah *berkas rekaman*. Diasumsikan bahwa semua rekaman berada dalam sebuah berkas yang memiliki format dan ukuran yang sama dan dengan jumlah rekaman yang sangat besar sehingga tidak semua dapat ditampung di dalam penganget komputer. Kondisi tersebut memerlukan organisasi penganget yang hierarkis dan multilevel. Dapat diasumsikan penganget diorganisasi secara bi level. Pada level atas adalah penganget utama yang digunakan untuk mengolah informasi, sedangkan pada level yang lebih rendah terdapat disk (pada umumnya disk magnetik) tempat rekaman akan disimpan. Disk pada umumnya adalah peralatan yang sifatnya diakses secara random yang memungkinkan pembacaan kembali sejumlah rekaman dengan waktu pembacaan yang tidak bergantung pada posisi dimana rekaman disimpan. Kemampuan disk untuk dapat diakses secara random tersebut memungkinkan untuk disusunnya struktur serta cara memanipulasi informasi secara efektif dan efisien.

Umumnya berkas dan rekaman tidak dapat dilihat oleh pengguna akhir suatu sistem komputer, tetapi rekaman dan berkas tersebut merupakan komponen yang sangat penting dalam sistem database. Hal yang paling penting berkaitan dengan rekaman dan berkas adalah kinerja berkas.



ORGANISASI BERKAS PRIMER

A. LATAR BELAKANG

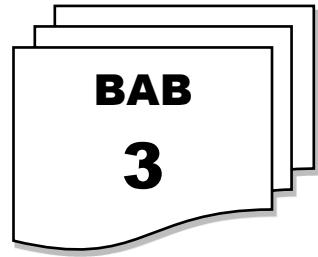
Bab ini membahas tentang cara mengorganisasi dan mengolah informasi dalam bentuk rekaman yang disimpan dalam sebuah berkas. Lebih fokus lagi akan mempelajari struktur yang tepat untuk mengorganisasi berkas dengan tujuan :

- Memungkinkan untuk di libatkannya data dalam ukuran volume yang sangat besar
- Memungkinkan dicapainya pengolahan terhadap data yang disimpan dalam berkas tersebut dengan level kinerja tertentu
- Memungkinkan dilakukannya pengolahan yang berbeda terhadap data yang disimpan dalam berkas untuk keperluan yang berbeda pula
- Menyediakan informasi yang memungkinkan dilakukannya tugas-tugas khusus

Dalam materi ini pembahasan tidak hanya berhenti pada pemilihan struktur data yang sesuai dengan kebutuhan, melainkan juga yang memenuhi kriteria efisiensi baik dilihat dari sisi waktu maupun ruang penyimpanan. Lebih jauh lagi berbagai ketentuan yang disyaratkan bagi metric suatu berkas akan digunakan sebagai pertimbangan utama.

B. ORGANISASI BERKAS

Terdapat tiga organisasi berkas utama yaitu sekuensial, langsung dan sekuensial berindeks, yang masing-masing memiliki cara memproses atau mengakses berkas tersebut.



BERKAS SEKUENSIAL

A. LATAR BELAKANG

Bab ini akan membahas tentang berkas sekuensial, sebuah tipe organisasi yang sudah sangat akrab di kalangan para mahasiswa. Dalam berkas sekuensial, rekaman yang ke $i+1$ akan diletakkan tepat sesudah rekaman ke i , sebagai contoh :

1	2	3	i	$i + 1$	$i + 2$	$n-1$	n
---	---	---	-------	-----	---------	---------	-------	-------	-----

Gambar 3.1 Susunan berkas sekuensial

Akses

Sesuai dengan namanya, berkas sekuensial sangat cocok untuk akses yang sekuensial, misalnya dalam aplikasi dimana sebagian besar atau semua rekaman akan diproses. Sebagai contoh adalah membuat daftar semua mahasiswa dalam sebuah jurusan. Melakukan akses secara sekuensial berarti bahwa proses akan berpindah dari satu rekaman ke rekaman berikutnya secara langsung.

Berkas sekuensial juga dapat diproses secara tunggal dan langsung, jika diketahui sub skripnya. Tetapi bagaimana kalau subskrip yang dimiliki bukan identitas utama rekaman, misalnya seperti "Nama mahasiswa" dalam contoh gambar 2.2. Pembacaan harus dilakukan secara sekuensial atau linear, rekaman demi rekaman, sampai "Nama mahasiswa" yang sesuai ditemukan. Untuk membaca rekaman dengan "Nama mahasiswa" = Dewi, diperlukan probe (akses terhadap lokasi yang berbeda) sejumlah 5 kali.



ORGANISASI BERKAS LANGSUNG

Meskipun teknologi berkembang sangat pesat, pengguna teknologi pada umumnya dan teknologi informasi pada khususnya merasa bahwa waktu yang diperlukan untuk memperoleh suatu informasi masih terlalu lama. Metode pencarian biner maupun interpolasi masih belum dapat mengimbangi ketidaksabaran manusia terhadap penyediaan informasi yang cepat dan akurat. Perkembangan media perekam data dalam hal kemampuan untuk menampung volume data yang sangat besar harus diimbangi dengan peningkatan teknik pencarian kembali data yang tersimpan.

Pada bab ini akan membahas tentang berbagai teknik yang dapat digunakan sebagai alternatif dalam merancang sistem, khususnya dalam menentukan lokasi rekaman yang akan disimpan serta pembacaan kembali rekaman tersebut. Bahasan akan dimulai dengan beberapa cara untuk mengorganisasi berkas untuk pengaksesan langsung, dilanjutkan dengan berbagai teknik yang dapat dipilih untuk meningkatkan kinerja sistem.

A. KUNCI SEBAGAI ALAMAT REKAMAN YANG UNIK

Untuk mendapatkan rekaman yang diasosiasikan dengan suatu kunci primer, sangat diharapkan agar proses langsung menuju ke alamat tempat rekaman dengan kunci tertentu disimpan. Hal tersebut hanya mungkin terjadi bila kunci rekaman juga merupakan alamat lokasi rekaman. Untuk suatu aplikasi dengan rekaman berisi informasi mahasiswa, untuk 13 digit nomor identitas mahasiswa (atau kunci utama rekaman yang merupakan gabungan dari 4 digit kode fakultas + jurusan, 4 digit tahun angkatan + 5 digit nomor urut), maka diperlukan 9.000.000.000.000 lokasi sebagaimana diperlihatkan pada gambar 4.1. dengan demikian waktu pencarian akan sangat baik, yaitu 1 probe untuk setiap rekaman yang dicari. Namun teknik tersebut memiliki



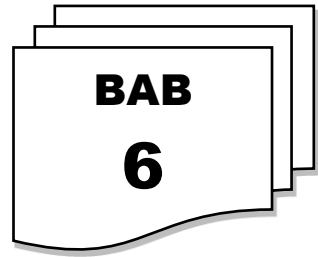
MANAJEMEN KOLISI

A. LATAR BELAKANG

Salah satu alasan di aplikasikannya fungsi hash adalah bahwa fungsi hash akan mendistribusikan kunci seperangkat data dengan lebih merata. Beberapa fungsi hash sudah dibahas pada bab sebelumnya dan salah satu diantaranya akan mendistribusikan data secara merata ke dalam berkas. Kalau tujuan tersebut tidak tercapai, salah satu strategi yang bisa diambil adalah mengkombinasikan beberapa fungsi sederhana dalam satu aplikasi. Fungsi hash yang menghasilkan banyak kolisi atau sinonim dikatakan sebagai memiliki kluster primer. Makin sedikit jumlah kolisi, makin baik fungsi hashing tersebut karena makin sedikit waktu yang diperlukan untuk melihat tempat-tempat yang berbeda dalam rangka menemukan rekaman yang diinginkan dan juga akan mempertahankan probe atau akses terhadap penyimpanan agar mendekati satu.

Meminimalkan jumlah probe pembacaan rekaman merupakan isu yang sangat penting mengingat bahwa berbagai aplikasi komersial pada umumnya memiliki volume data yang sangat besar sehingga data tersebut harus disimpan dalam penyimpanan luar (*auxiliary*). Beberapa cara yang dapat ditempuh untuk mereduksi kolisi adalah mengganti fungsi hashing atau dengan mereduksi *factor-packing*. *Factor-packing* suatu berkas adalah perbandingan (rasio) antara jumlah rekaman yang disimpan dalam berkas dengan kapasitas berkas atau dapat dinyatakan sebagai berikut:

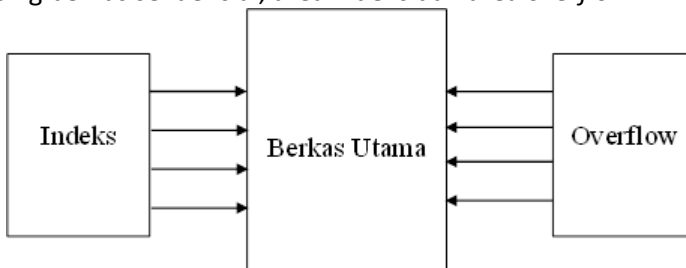
$$\text{Factor-packing (FP)} = \frac{\text{jumlah rekaman yang disimpan}}{\text{Jumlah total lokasi penyimpanan}}$$



ORGANISASI BERKAS SEKUENSIAL BERINDEKS

A. PENDAHULUAN

Berkas sekuensial berindeks dirancang dengan tujuan untuk menanggulangi permasalahan pengaksesan yang dimiliki oleh organisasi berkas sekuensial tanpa mengurangi keuntungan dan tradisi yang dimiliki oleh berkas sekuensial. Berkas sekuensial berindeks memiliki 2 sifat yang mengungguli berkas sekuensial, pertama adalah indeks terhadap berkas sehingga menghasilkan pengaksesan random yang lebih baik dan kedua adalah area *overflow* untuk menyediakan ruang bila dilakukan penambahan rekaman ke dalam berkas. Gambar 6.1 memperlihatkan contoh berkas sekuensial berindeks. Pada gambar tersebut diperlihatkan tiga komponen masing-masing berkas sekuensial, area indeks dan area *overflow*.



Gambar 6.1 Komponen dalam berkas sekuensial berindeks

Berkas sekuensial berindeks memungkinkan di lakukannya pembacaan secara sekuensial rekaman-rekaman yang berada dalam area primer. Rekaman-rekaman yang disisipkan dialokasikan pada area yang terpisah yang disebut *overflow*. Rekaman-rekaman dalam area *overflow* ditempatkan



PENGURUTAN REKAMAN

Cara yang paling umum untuk mengorganisasi sebuah berkas adalah menyimpan rekaman-rekaman secara urut berdasar nilai sebuah medan dalam rekaman yang akan disimpan tersebut. Tipe pengurutan tersebut akan mempermudah dalam pencarian rekaman dengan kunci tertentu. Pengurutan atau menyusun data secara urut dengan bantuan komputer merupakan operasi yang sangat umum. Hampir semua hasil pengolahan komputer untuk konsumsi manusia selalu dalam keadaan urut.

Pengurutan data merupakan komponen dasar struktur data. Berbagai organisasi berkas memerlukan data yang urut agar dapat bekerja dengan tepat, sebagai contoh adalah pencarian biner ataupun interpolasi. Pengurutan data juga dapat dimanfaatkan untuk mengeliminasi rekaman-rekaman yang ganda. Jika terjadi duplikasi maka rekaman-rekaman tersebut akan berada pada posisi berurutan dengan demikian proses pengeliminasian dapat dilakukan dengan mudah.

A. PENGURUTAN GELEMBUNG

Salah satu prosedur pengurutan yang paling sederhana adalah pengurutan gelembung. Salah satu karakter dari pengurutan ini adalah bahwa pengurutan gelembung ini sangat mudah untuk dipahami dan diprogramkan. Tetapi dari semua prosedur pengurutan yang akan dibahas, prosedur pengurutan gelembung ini merupakan yang paling tidak efisien. Disebut pengurutan gelembung karena masing-masing kunci akan dengan lambat menggelembung ke posisinya yang tepat.

DAFTAR PUSTAKA

- Dewi Handayani, Sistem Berkas, Graha Ilmu, Kabupaten Sleman, Daerah Istimewa Yogyakarta, 2012.
- Dr.Ir. Wahyuni, M.Sc. Sistem Berkas, Andi, Yogyakarta, 2013.
- Granville Barnett, and Luca Del Tongo, Data Structures and Algorithms: Annotated Reference with Examples, First Edition, <http://dotnetslackers.com/>, 2008.
- Hariyanto, Bambang, Pengarsipan dan Akses pada Sistem Berkas, Informatika Bandung, 2000.
- Handayani, Dewi, Sistem Berkas, J & J Learning Yogyakarta, 2001.
- Ir. Wahyuni, M.Sc., Sistem Berkas, Andi Yogyakarta, 2004.
- Rosa A.S., Struktur Data Terapan Dalam Berbagai Bahasa Pemrograman: Pascal, C, C++, dan Java, Modula Bandung.
- Rumetna, M.S., Struktur dan Organisasi Data, Kabupaten Banyumas, Jawa Tengah, CV. Pena Persada, 2022.
- Wahid, Fathul, Dasar-Dasar Algoritma dan Pemrograman. Yogyakarta : Andi, 2004.

PROFIL PENULIS

Nama : Matheus Supriyanto Rumatna,
S.Kom., M.Sc

Afiliasi : Universitas Victory Sorong

Email : matheus.rumatna@gmail.com

id Sinta : 6683399

id Scopus : 57221417387

ORCID : <https://orcid.org/0000-0001-9474-9765>

id Publons : 3684842

id Scholar :

https://scholar.google.com/citations?hl=id&user=wpbNkOkAAAAJ&view_op=list_works&sortby=pubdate

Url : <https://matheusrumatna.com/>

Riwayat Pendidikan : - S1 Teknik Informatika (USTJ - Jayapura)

- S2 Sistem Informasi (UKSW - Salatiga)



Penulis merupakan seorang tenaga pengajar di Fakultas Ilmu Komputer Universitas Victory Sorong sejak tahun 2015. Memiliki minat penelitian yang tinggi meliputi *information system, geographic information system, cloud computing, information system audit, operation research, knowledge management, algorithms and programming*. Memiliki pengalaman sebagai pembicara dalam kegiatan Seminar (webinar) tingkat lokal, nasional dan internasional. Serta sebagai Anggota APTIKOM aktif hingga saat ini.

Nama : Tirsawp Ninia Lina, S.Kom., M.Cs.
Afiliasi : Universitas Victory Sorong
Email : tirsawp@gmail.com
Id Google Scholar : WCMj8wwAAAAJ
Id SINTA : 6683408
Id SCOPUS : 57221418898
Orcid : <https://orcid.org/0000-0002-5944-8868>
Riwayat Pendidikan : S1 Teknik Informatika (STMIK AMIKOM - Yogyakarta), S2 Sistem Informasi (UKSW - Salatiga)



Penulis merupakan seorang tenaga pengajar di Program Studi Sistem Informasi - Fakultas Ilmu Komputer, Universitas Victory Sorong, sejak tahun 2015. Minat penelitian meliputi pengembangan sistem informasi, sistem informasi geografis, web, sistem informasi manajemen, multimedia. Memiliki pengalaman sebagai pembicara dalam kegiatan Seminar (webinar) tingkat lokal, nasional dan internasional. Serta sebagai Anggota APTIKOM aktif hingga saat ini.

SISTEM BERKAS

Umumnya berkas dan rekaman tidak dapat dilihat oleh pengguna akhir suatu sistem komputer, tetapi rekaman dan berkas tersebut merupakan komponen yang sangat penting dalam sistem database. Hal yang paling penting berkaitan dengan rekaman dan berkas adalah kinerja berkas. Data merupakan aset penting dalam organisasi.

Berbagai pendapat yang sering dikemukakan oleh berbagai pakar mengenai dengan sistem komputer yang kontemporer adalah perangkat keras memiliki umur yang singkat, perangkat lunak memiliki umur yang lebih panjang. Materi ini memberikan pengenalan dasar terhadap berkas yang digunakan untuk menyimpan data dengan volume yang besar dalam sistem komputer yang kontemporer. Minat utama dalam buku ini adalah berkas rekaman. Diasumsikan bahwa semua rekaman berada dalam sebuah berkas yang memiliki format dan ukuran yang sama dan dengan jumlah rekaman yang sangat besar sehingga tidak semua dapat ditampung di dalam pengingat komputer. Kondisi tersebut memerlukan organisasi pengingat yang hirarkis dan multilevel. Dapat diasumsikan pengingat diorganisasi secara bilevel. Pada level atas adalah pengingat utama yang digunakan untuk mengolah informasi, sedangkan pada level yang lebih rendah terdapat disk (pada umumnya disk magnetik) tempat rekaman akan disimpan. Disk pada umumnya adalah peralatan yang sifatnya diakses secara random yang memungkinkan pembacaan kembali sejumlah rekaman dengan waktu pembacaan yang tidak bergantung pada posisi dimana rekaman disimpan. Kemampuan disk untuk dapat diakses secara random tersebut memungkinkan untuk disusunnya struktur serta cara memanipulasi informasi secara efektif dan efisien.



Penerbit
widina
www.penerbitwidina.com

ISBN 978-623-459-287-0



9 786234 592870