

Projeto e operação de Bancos de Dados - Parte 1

Prof. Dr. José de J. Pérez Alcázar
Bancos de Dados

INTRODUÇÃO

DADO X INFORMAÇÃO

Uma percepção do mundo real pode ser vista como uma série de fenômenos diferentes que algumas vezes têm alguma relação entre si.

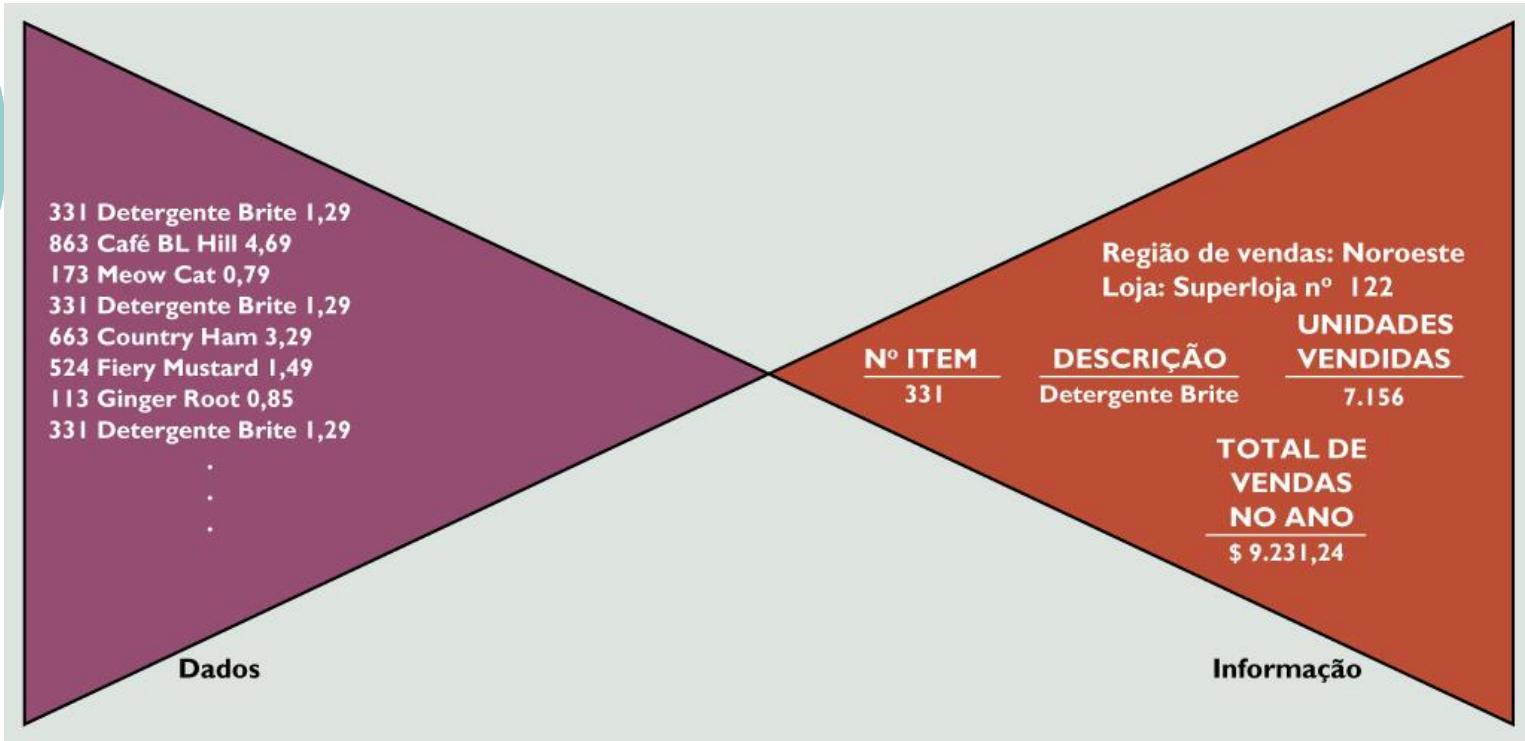
“Dado” → A descrição destes fenômenos.

Através deles obtemos informação do mundo real

Informação → Qualquer aumento do conhecimento obtido a través da interpretação e uso de dados.



INTRODUÇÃO

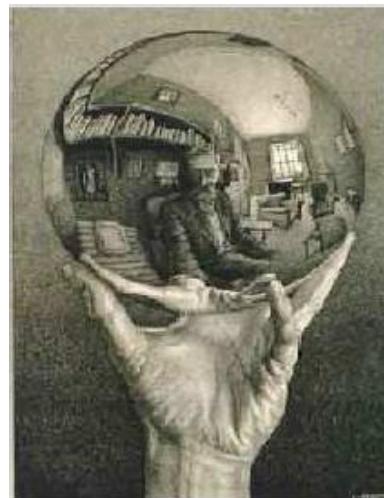


Dados e informações

INTRODUÇÃO

Conhecimento.

- Capacidade de resolver problemas, inovar e aprender baseando-se em experiências prévias; Tomar decisões, atuar.
- Esforço de investigação para descobrir aquilo que está oculto, que não está compreendido ainda.
 - » Adquirir conhecimento não é reter informação, mas utilizar estas para desvendar o novo e avançar.



INTRODUÇÃO

- Portanto, dado não é informação e informação não é conhecimento.
- Hoje em dia, organizações competem pelo domínio do **conhecimento** científico e tecnológico;
- COMO ?
 - » Armazenando, processando, acessando e disponibilizando informações por meio de redes de comunicação.

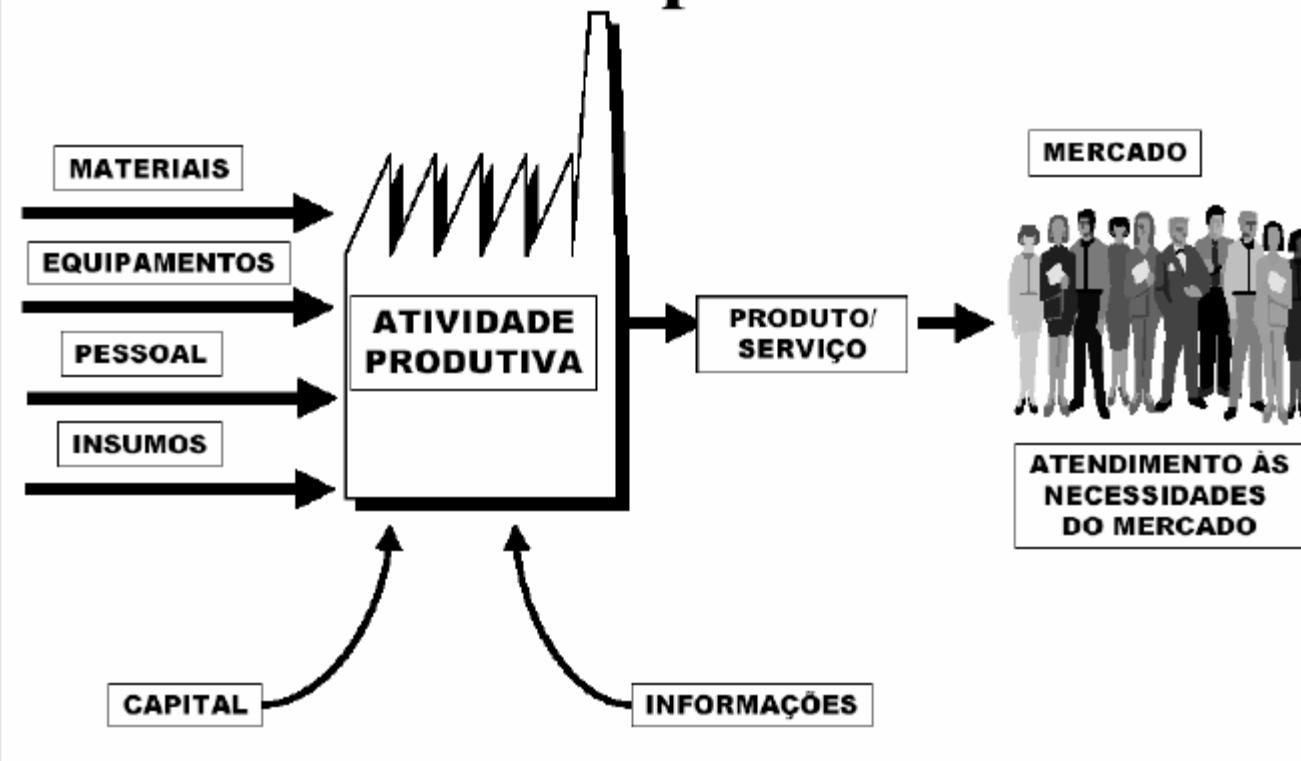
Introdução

Sistema

Arranjo de componentes que trabalham juntos para alcançar uma meta comum, ou múltiplas metas, aceitando entradas, processando elas, e produzindo saídas de forma organizada.

- Ex.: **automóvel** (objetivo: transportar passageiros e carga - subsistemas: motor, caixa de marchas, suspensão); **computador** (armazenar e processar dados – subsistemas: teclado, vídeo, CPU, ...)
- Um sistema é formado por sub-sistemas, que por sua vez são sistemas → Simples e Complexos.

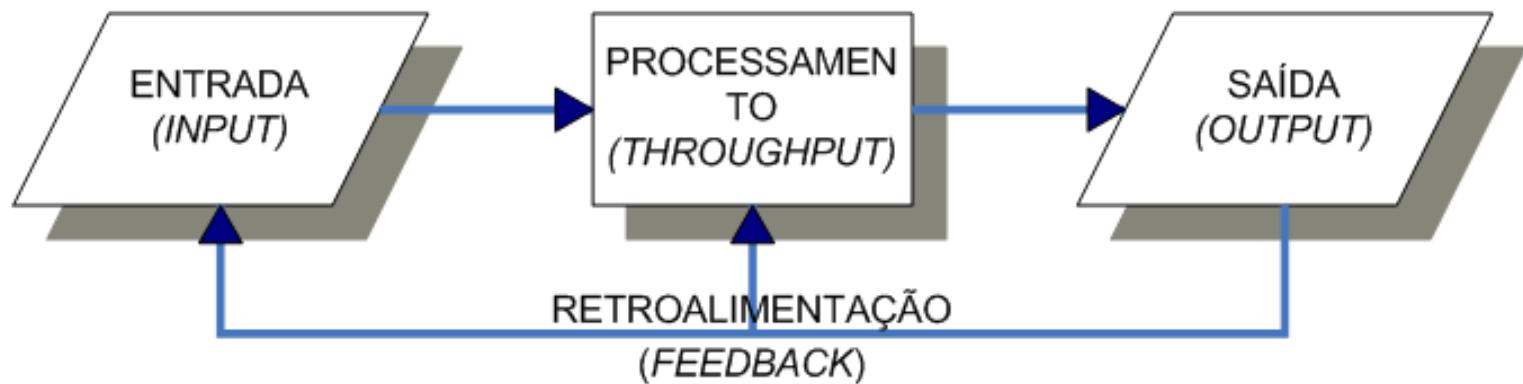
A empresa



Modelo Geral de um sistema:

AMBIENTE

Copyright © 2005 Osnaldo Pereira de Araújo



INTRODUÇÃO

A informação sempre foi um recurso fundamental no funcionamento das organizações. Ainda mais, agora que vivemos no que se chama Sociedade da Informação.

As organizações precisam integrar o uso de sua informação para o planejamento e controle.

INTRODUÇÃO

Um Sistema de Informação é um conjunto de elementos que interagem entre si. Esses elementos são de natureza diversa e normalmente incluem:

- **A infraestrutura computacional e de comunicação**
- **O recurso humano que interage com o Sistema de Informação**
- **Os dados ou informação fonte que são introduzidos no sistema**
- **Os programas (processos) que produzem diferentes tipos de resultados.**

Implantado através de um *Sistema Manual* ou um *Sistema Computacional*

CONCLUSÃO

Um sistema de informação é um meio de oferecer a informação necessária para uma organização. Um sistema de informação recebe a informação, a armazena, a processa, e oferece o seu acesso aos usuários, da forma que eles desejam.

Um sistema de informação é construído principalmente a partir de SGBD's.

INTRODUÇÃO

Dados “úteis”
para geração
de informação



COLECTADOS E
ARMAZENADOS

BANCO DE
DADOS

A informação contida no BD deve representar um instante (estado) de uma aplicação determinada. Cada mudança no BD é um reflexo dum evento (ou seqüência) que ocorre neste ambiente.

*Modelo de uma parte do Mundo Real
(Universo do discurso)*



Sistema de Banco
de Dados



Sistema de informação computacional,
cujo objetivo principal é registrar e manter
dados.

Trata com BD's

Um Sistema de Banco de Dados (SBD) tem 4 componentes:

- “Software”
- “Hardware”
- “Dados”
- “Usuários”

INTRODUÇÃO

HARDWARE Áreas de memória secundária nas quais está localizada o BD (discos, fitas, etc.), junto com os dispositivos associados (controladores, canais, etc.)

SOFTWARE SGBD. Todo acesso é feito a través do SGBD.

Facilita  Estruturação
Armazenamento
Recuperação
Manipulação
Controle de Dados

Promove  Independência de Dados
Dados compartilhados

INTRODUÇÃO

CLASSES DE USUÁRIOS:

- Programador de aplicações
- Usuário final
- Administrador do BD

**Programador da
aplicação**



Escreve programas que utilizam o BD em
C, C++, Java, Python, etc.

Usuários finais



Acessa o BD a través de um linguagem
de consulta (parte integrante do sistema)
ou a través de um programa interativo.
Ação comum a recuperção

INTRODUÇÃO

Administrador do banco de dados :

Pessoa com o controle do sistema.

Funções:

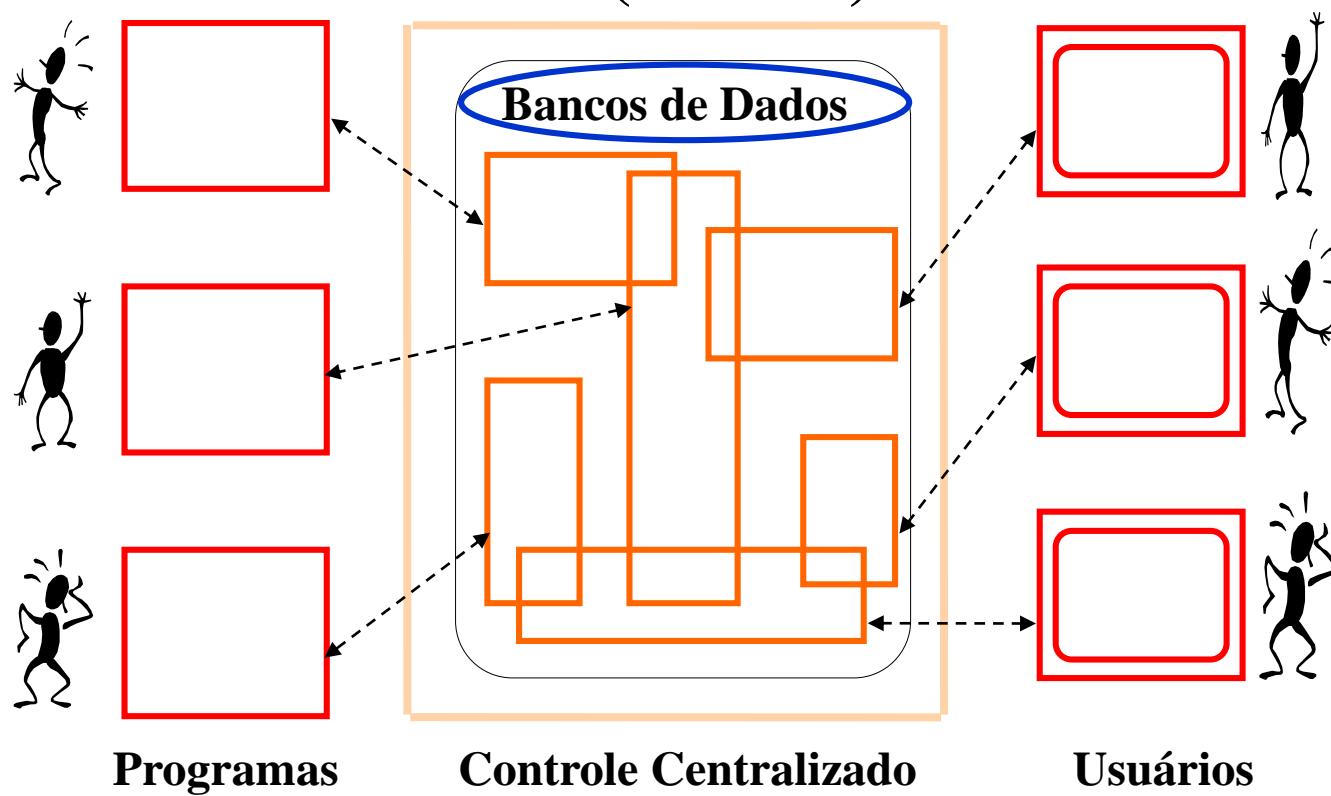
- Mudanças na organização física e lógica
- Garante diferentes tipos de acesso por parte dos usuários
- Define uma estratégia de “Backups” e recuperação, etc.

Projetistas do Banco de Dados:

São responsáveis pela identificação dos dados que serão armazenados e pela seleção das estruturas apropriadas para representar e armazenar estes dados.

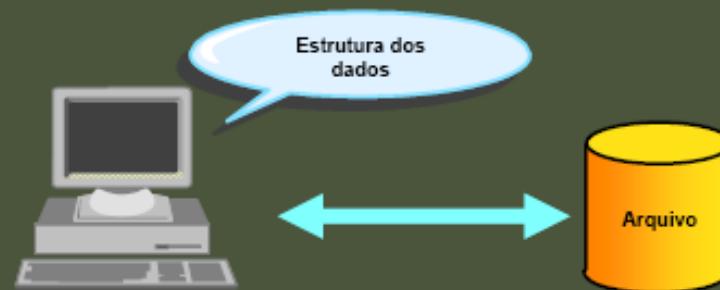
INTRODUÇÃO

Sistema de Gerenciamento de Banco de Dados (SGBD)



Uma perspectiva histórica

- ❑ No início da computação, programas tinham o único objetivo de armazenar e manipular dados.
- ❑ Esses programas gravavam seus dados em disco, segundo estruturas próprias.
- ❑ Programas que não conhecessem a estrutura dos dados não podiam utilizar os dados.



Uma perspectiva histórica

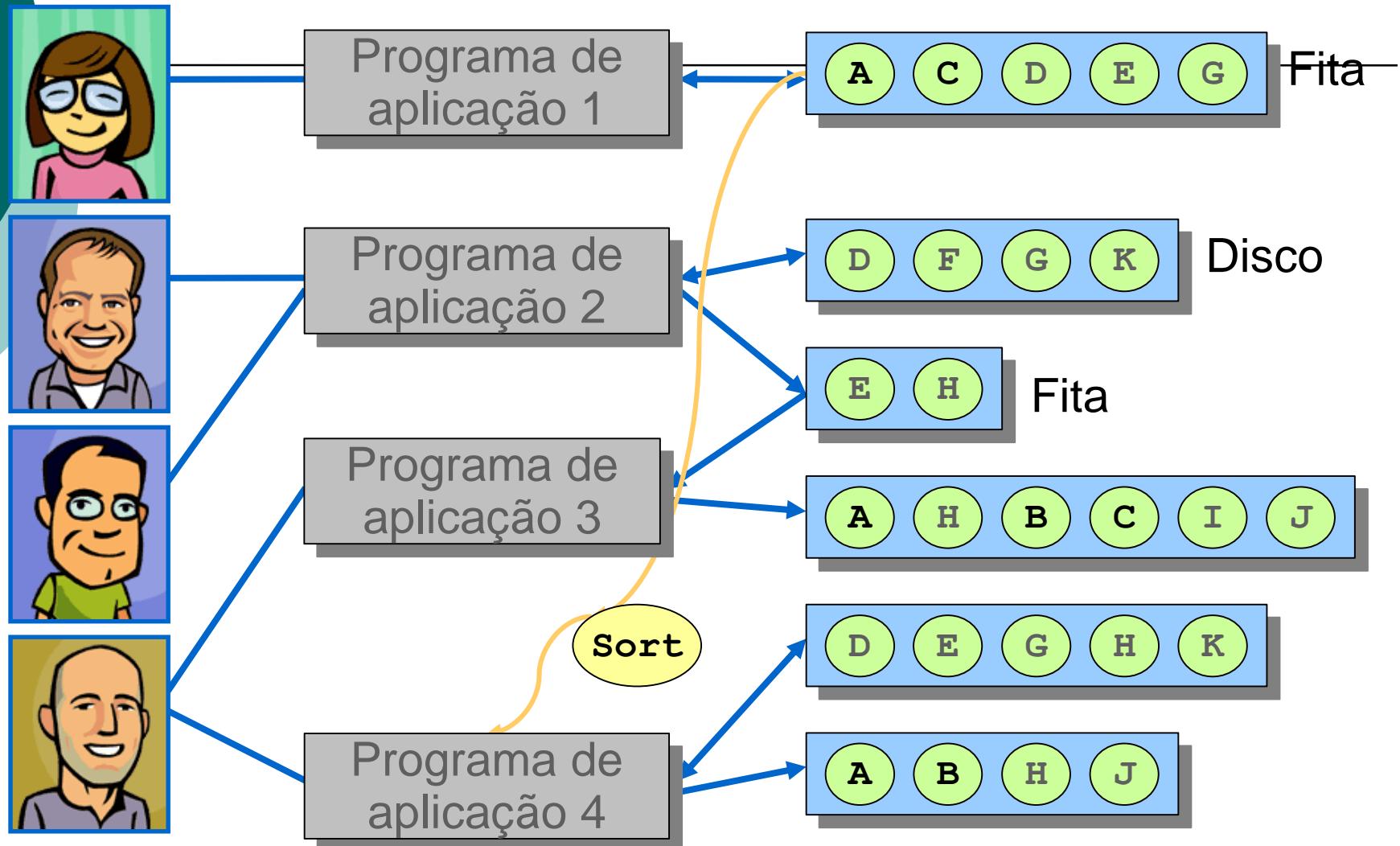
- Se vários programas precisassem compartilhar os dados de um mesmo arquivo, todos os programas teriam que conhecer e manipular as mesmas estruturas.



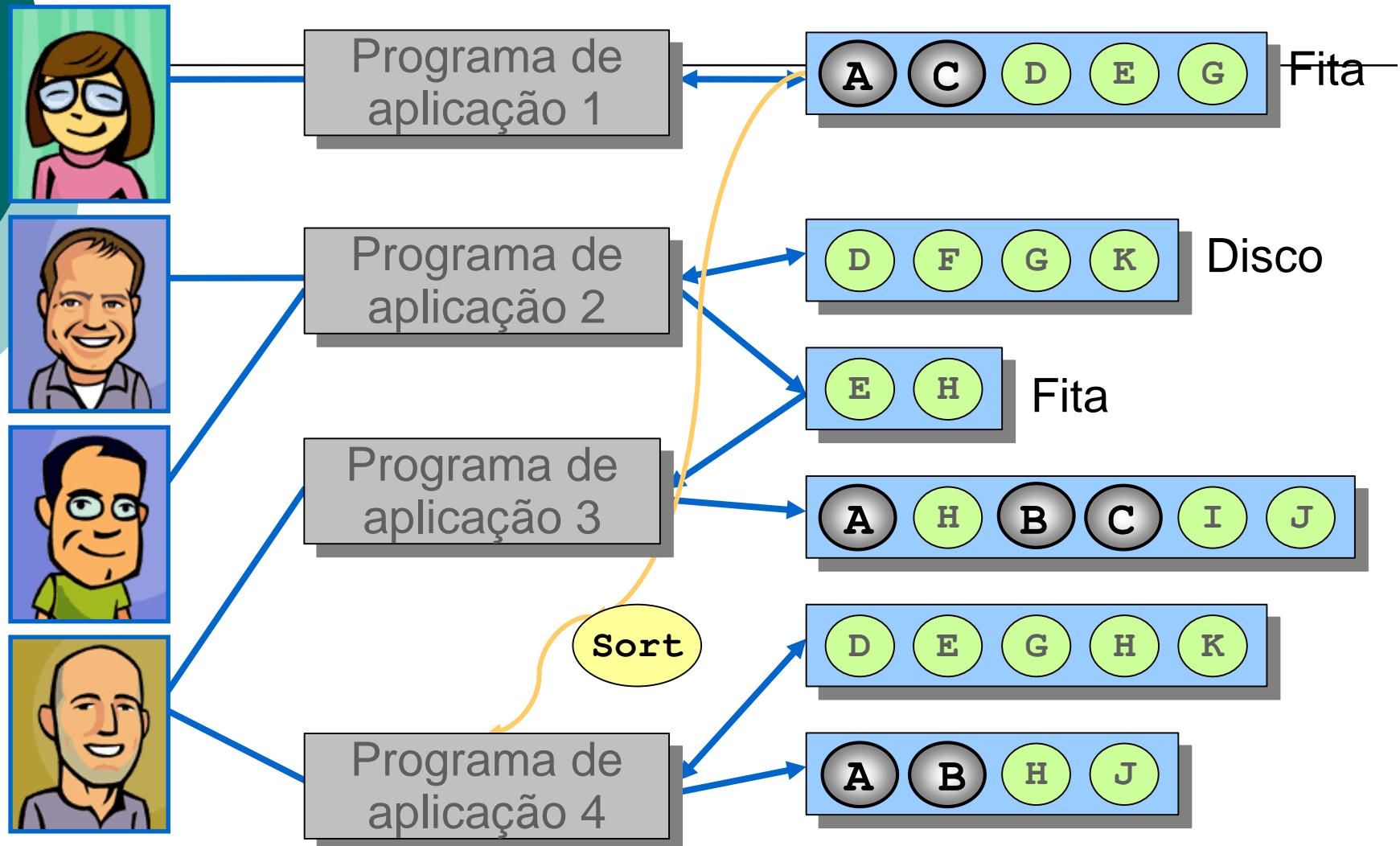
Uma perspectiva histórica

- Se algum programa precisasse realizar alguma mudança na estrutura de dados,
 - Todos os programas que acessam esse mesmo arquivo tinham que ser alterados, mesmo que a alteração ocorresse em dados não manipulados pelos programas.
- Isso gerava um grande problema:
 - Garantir a unicidade das estruturas de dados entre os diversos programas devido à existência de redundâncias.

Sistema de Arquivos



Sistema de Arquivos



Sistema de Arquivos

Ambiente de arquivos

Por cada nova aplicação, o programador ou analista cria um novo arquivo

Uma Instalação maior possui centenas ou milhares destes arquivos com muita **redundância** de dados.

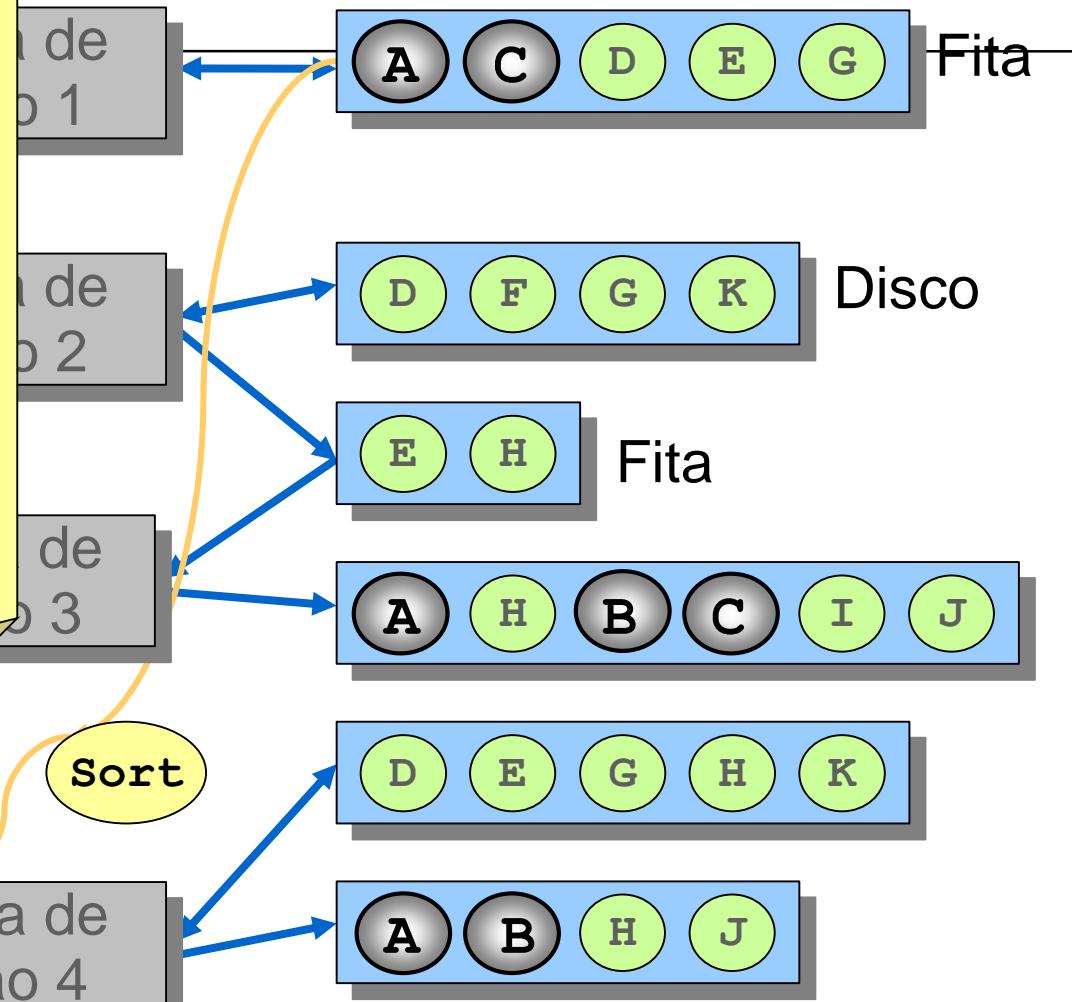
Prog. de Aplic. 5

A B C

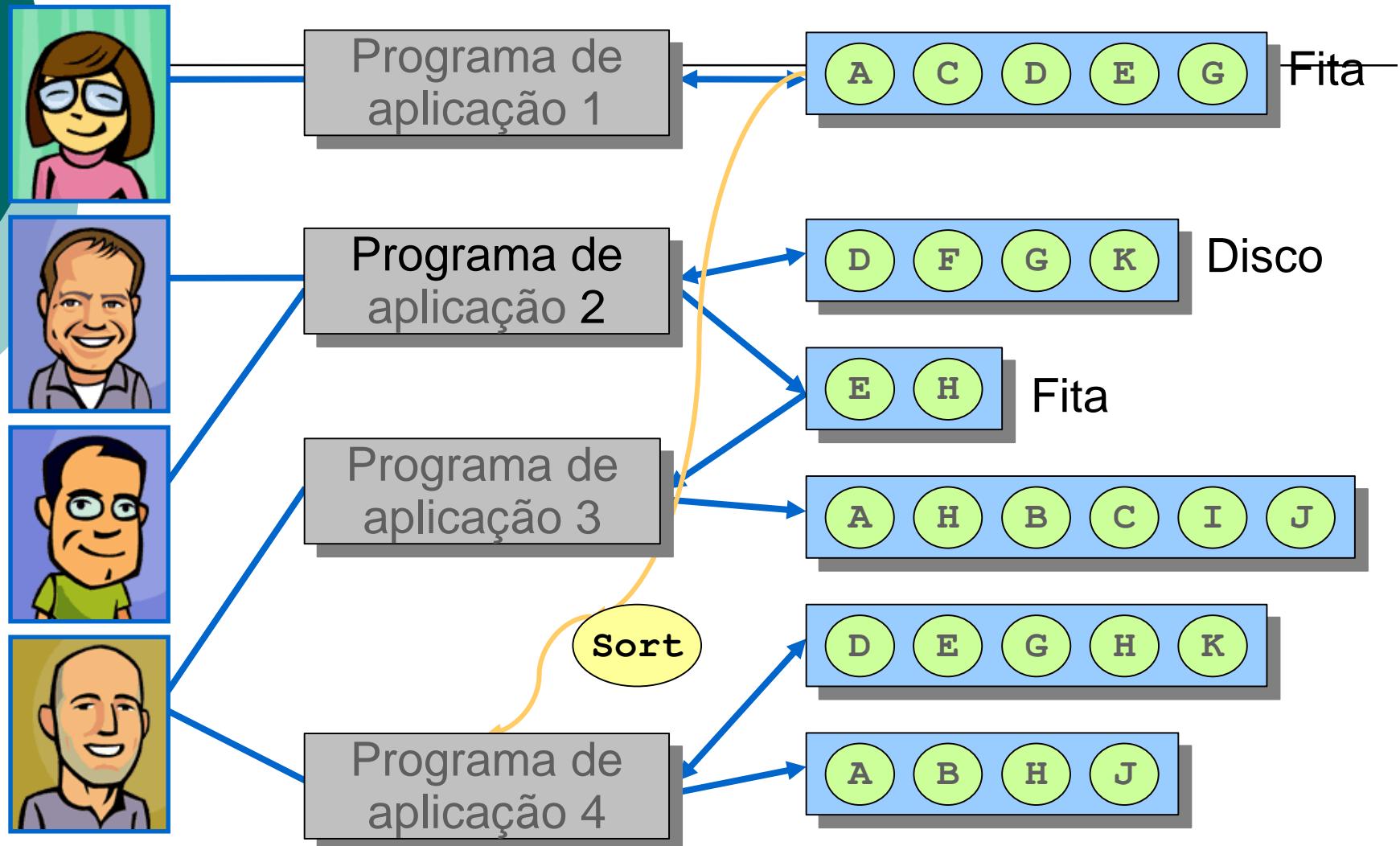


Programa de aplicação 4

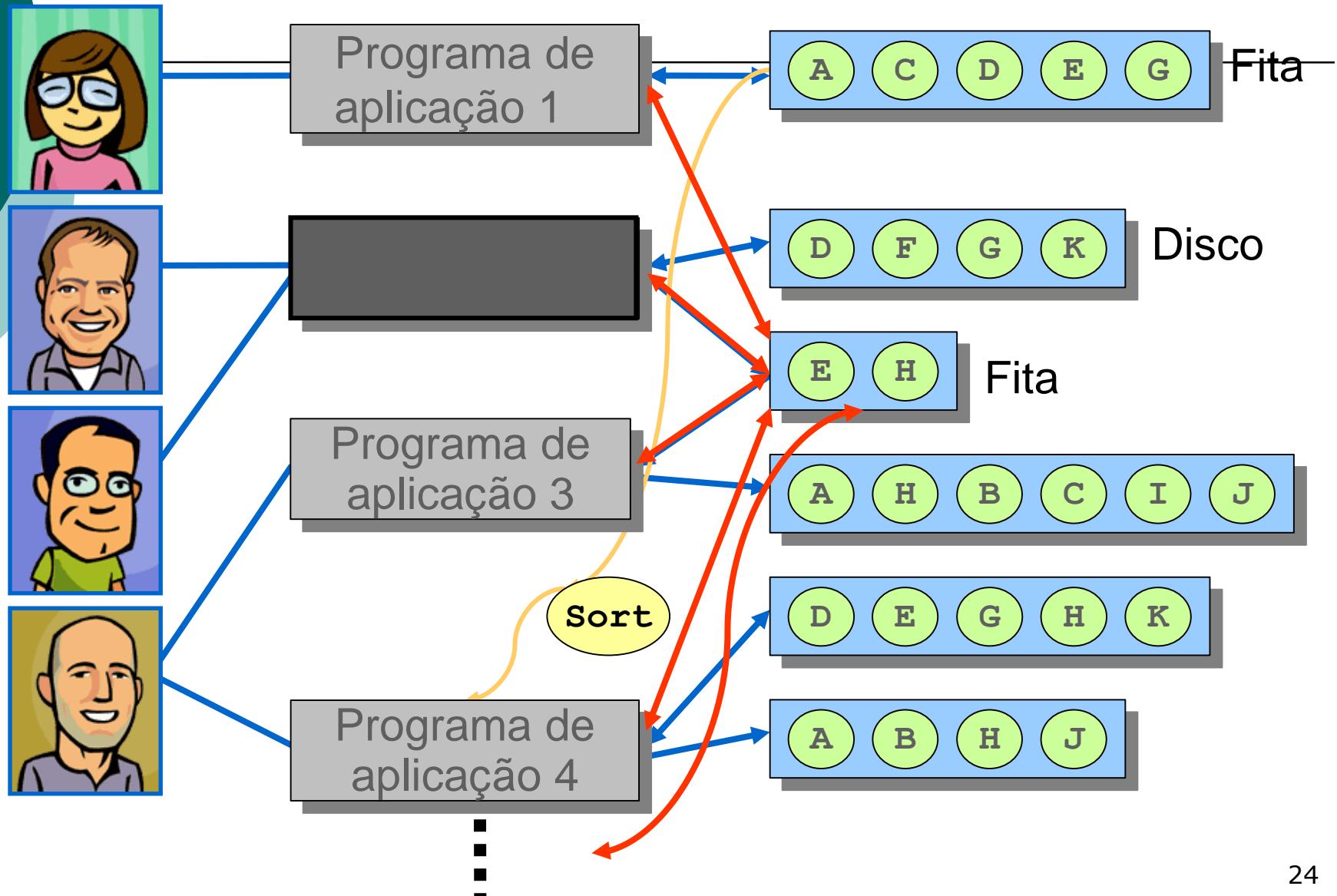
Sort



Sistema de Arquivos



Sistema de Arquivos

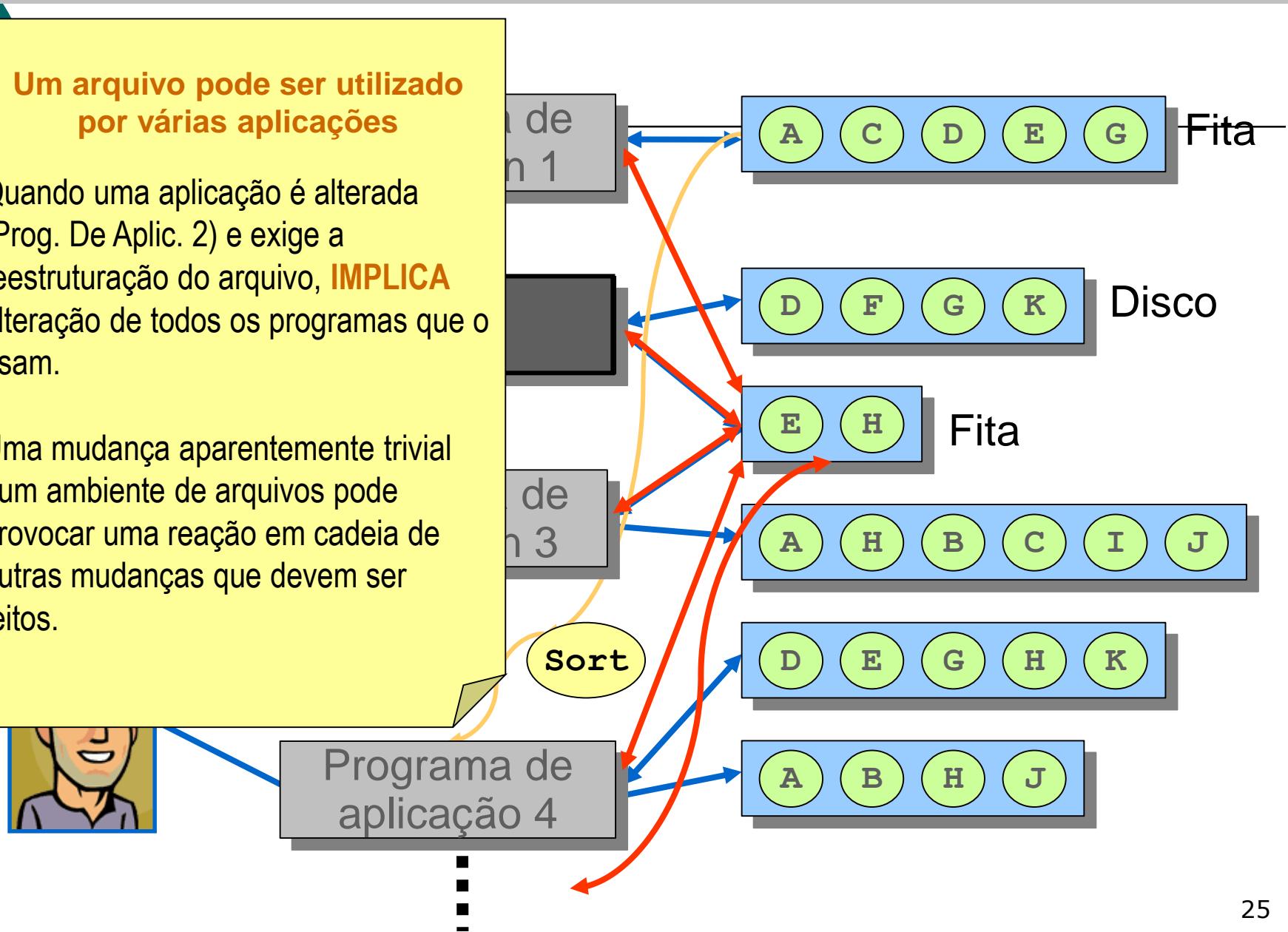


Sistema de Arquivos

Um arquivo pode ser utilizado por várias aplicações

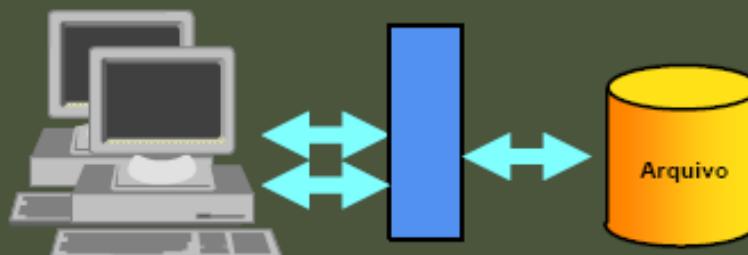
Quando uma aplicação é alterada (Prog. De Aplic. 2) e exige a reestruturação do arquivo, **IMPLICA** alteração de todos os programas que o usam.

Uma mudança aparentemente trivial num ambiente de arquivos pode provocar uma reação em cadeia de outras mudanças que devem ser feitos.



Uma perspectiva histórica

- Para evitar esse problema, colocou-se um sistema intermediário:
 - Que conhece a estrutura de dados do arquivo.
 - Fornece apenas dados que cada programa precisa.
 - Armazena adequadamente os dados de cada programa.

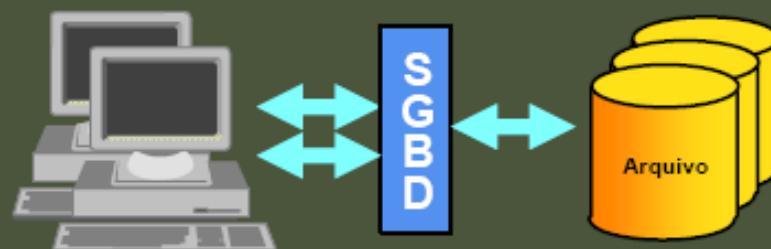


Uma perspectiva histórica

- Agora, com esse sistema intermediário:
 - Os programas "verão" apenas os dados que lhes interessam.
 - Os programas não precisam conhecer os detalhes de como seus dados estão gravados fisicamente.
 - Os programas não precisarão ser modificados se a estrutura de dados que utilizam não for modificada.
 - As alterações ficam concentradas nesse sistema intermediário.

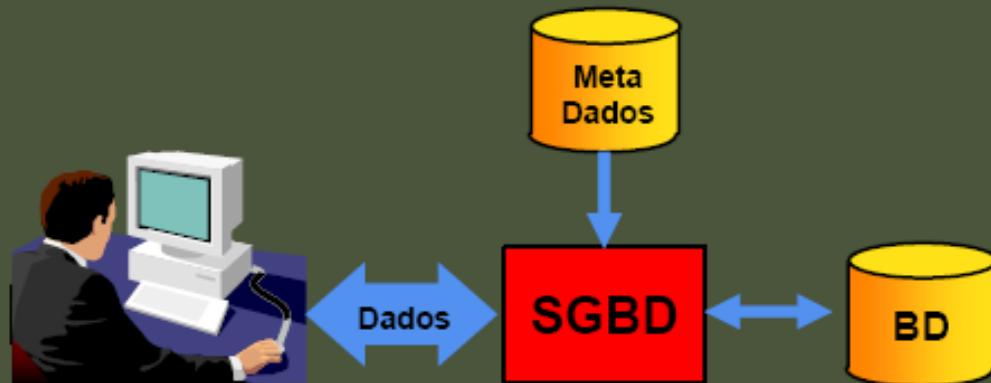
Uma perspectiva histórica

- Com o tempo, esse sistema intermediário passou a gerenciar vários arquivos.
- A essa coleção de arquivos foi dado o nome de **Banco de Dados** e o sistema intermediário recebeu o nome de **Sistema Gerenciador de Banco de Dados (SGBD)**.



Uma perspectiva histórica

- ❑ O primeiro SGBD comercial surgiu em 1960.
- ❑ Com o tempo, surgiram padrões para descrever as estruturas de dados: os **modelos de dados**.
- ❑ A descrição do banco de dados, segundo um modelo de dados é chamada de **meta dados**.



Um exemplo

- Banco de dados de uma universidade:

ALUNO	Nome	Numero	Turma	Curso_Hab
	Smith	17	1	CC
	Brown	8	2	CC

CURSO	Nome do Curso	Número do Curso	Creditos	Departamento
	Introdução à Ciência da Computação	CC1310	4	CC
	Estruturas de dados	CC3320	4	CC
	Matemática Discreta	MAT2410	3	MATH
	Banco de dados	CC3380	3	CC

DISCIPLINA	Identificador de Disciplina	Número do Curso	Semestre	Ano	Instrutor
	85	MAT2410	Segundo Semestre	98	King
	92	CC1310	Segundo Semestre	98	Anderson
	102	CC3320	Primeiro Semestre	99	Knuth
	112	MAT2410	Segundo Semestre	99	Chang
	119	CC1310	Segundo Semestre	99	Anderson
	135	CC3380	Segundo Semestre	99	Stone

HISTORICO_ESCOLAR	Número do Aluno	Identificador_Disciplinas	Nota
	17	112	B
	17	119	C
	8	85	A
	8	92	A
	8	102	B
	8	135	A

PRE_REQUISITO	Número do Curso	Número do Pre_requisito
	CC3380	CC3320
	CC3380	MAT2410
	CC3320	CC1310

Um exemplo

- Arquivo **ALUNO**: conserva os dados de cada estudante na universidade:

ALUNO	Nome	Numero	Turma	Curso_Hab
	Smith	17	1	CC
	Brown	8	2	CC

Um exemplo

- Arquivo **CURSO**: preserva os dados sobre cada curso:

CURSO	Nome do Curso	Número do Curso	Creditos	Departamento
	Introdução à Ciência da Computação	CC1310	4	CC
	Estruturas de dados	CC3320	4	CC
	Matemática Discreta	MAT2410	3	MAT
	Banco de Dados	CC3360	3	CC

Um exemplo

- Arquivo **DISCIPLINA**: guarda os dados de cada disciplina do curso:

DISCIPLINA	Identificador deDisciplina	NumerodoCurso	Semestre	Ano	Instrutor
85		MAT2410	Segundo Semestre	98	King
92		CC1310	Segundo Semestre	98	Anderson
102		CC3320	Primeiro Semestre	99	Knuth
112		MAT2410	Segundo Semestre	99	Chang
119		CC1310	Segundo Semestre	99	Anderson
136		CC3380	Segundo Semestre	99	Stone

Um exemplo

- Arquivo **HISTORICO_ESCOLAR**: mantém as notas recebidas por aluno nas diversas disciplinas cursadas:

HISTORICO_ESCOLAR	NumerodoAluno	Identificador_Disciplina	Nota
	17	112	B
	17	119	C
	8	85	A
	8	92	A
	8	102	B
	8	135	A

Um exemplo

- Arquivo **PRE_REQUISITO**: armazena os pré-requisitos de cada curso:

PRE_REQUISITO	NumerodoCurso	NumerodoPre_Requisito
	CC3380	CC3320
	CC3380	MAT2410
	CC3320	CC1310

Um exemplo

Para definir esse banco de dados, devemos:

- especificar a estrutura de cada registro, em cada arquivo;
- considerar os diferentes tipos de elementos dos dados a serem armazenados em cada registro.

Um exemplo

Cada registro ALUNO inclui os dados:

- Nome do Aluno
- Número do Aluno
- Turma
 - calouro, ou 1
 - veterano, ou 2
 - ...
- Curso/Habilitação
 - Matemática, ou MAT
 - CiênciadaComputação, ou CC

Um exemplo

É necessário, ainda, especificar os tipos de dados para cada elemento de dados em um registro, como:

- NomedoAluno em ALUNO é uma string (cadeia de caracteres alfabéticos);
- NumerodoAluno em ALUNO é um inteiro;
- Nota em HISTORICO_ESCOLAR é um caractere único no conjunto {A, B, C, D, E, F, I};

Um exemplo

Para **construir** o banco de dados
UNIVERSIDADE:

- armazena-se os dados que representam cada aluno, curso, disciplina, relatório de notas e pré-requisitos em cada registro em seu arquivo apropriado.
- Obs: os registros em diferentes arquivos podem estar relacionados.

Um exemplo

ALUNO	Nome	Numero	Turma	Curso_Hab
	Smith	17	1	CC
	Brown	8	2	CC

HISTORICO_ESCOLAR	NumerodoAluno	Identificador_Disciplina	Nota
	17	112	B
	17	119	C
	8	85	A
	8	92	A
	8	102	B
	8	135	A

O registro para “Smith” no arquivo ALUNO está relacionado a dois registros no arquivo HISTORICO_ESCOLAR

Um exemplo

A manipulação do banco de dados envolve operações de:

- consulta (*query*)
 - atualização
- Exemplos de consultas:
 - a recuperação do histórico escolar de 'Smith';
 - a relação dos nomes de alunos que fizeram as disciplinas do curso de Banco de Dados no segundo semestre de 1999;

Um exemplo

- Exemplos de atualizações:
 - mudar a turma de 'Smith' para veteranos;
 - criar uma nova disciplina para o curso de Banco de Dados nesse semestre;
 - colocar a nota A para 'Smith' na disciplina do curso Banco de Dados no último semestre.

POR QUE BANCOS DE DADOS ?

- Resposta Inicial** → Um SBD proporciona à empresa um controle centralizado dos seus dados
- Arquivos dispersos → Controle difícil

VANTAGENS

A redundância pode ser reduzida

Informação dispersa → redundância

Inconsistência pode ser evitada

Conseqüência do ponto anterior

Redundância controlada → Propagação atualização

Redundância não controlada → Inconsistência

POR QUE BANCOS DE DADOS ?

Dados podem ser compartilhados

Os dados do BD podem ser compartilhados por diferentes aplicações. Novas aplicações podem ser implementadas sem a necessidade de se criar novos arquivos.

Padrões podem ser reforçados

Deve ser seguido um padrão de representação dos dados devido ao controle.

VANTAGENS

Podem ser aplicadas restrições de segurança.

O ABD tem controle completo sobre os dados; pode:

- Garantir que as únicas vias de acesso sejam a través dos canais adequados.
- Definir a execução de verificações de acesso quando sejam acessados dados sensíveis (verificação X tipo de acesso)

A integridade pode ser mantida

Integridade  Dados corretos.

Redundância  Possíveis dados incorretos

Outros tipos de inconsistência podem existir

Exemplo: Empregado que trabalha 48 horas por dia

Salário de um empregado > Salário do chefe

VANTAGENS

Devido ao controle centralizado estas situações podem ser eliminadas com procedimentos de validação.

Tanto as restrições de segurança como os procedimentos de validação são mais importantes em BDs que em arquivos isolados.

Necessidades conflitantes podem ser balanceadas.

Conhecendo as necessidades globais da organização e as necessidades dos usuários individuais → ABD estrutura o sistema procurando o benefício global.

Independência de dados

Mudar a estrutura de armazenamento ou a estratégia de acesso sem mudar os programas (de forma dramática)

DESVANTAGENS

Custos incrementais e vulnerabilidade do sistema

- Vulnerabilidade → Natureza centralizada
- Processo de recuperação de falhas mais custoso
- Generalidade que um SGBD oferece para definir e processar dados.
- Os custos de hardware, software e programadores são mais altos.

Uso de arquivos quando:

- O BD e aplicações são simples, bem definidas e não se espera que mudem
- Altos requerimentos de eficiência em tempo real.
- Requerimentos de acesso mono-usuário.