

Sistemas de Gerência de Bancos de Dados

- Módulo 3 -

Sistemas de Armazenamento
de Alto Desempenho



Introdução

Capacidade e Desempenho

processador	Kiloflop	Megaflop	Gigaflop	Teraflop
armazenam.	Megabyte	Gigabyte	Terabyte	Petabyte
rede	Kilobit	Megabit	Gigabit	Terabit
memória	Kilobyte	Megabyte	Gigabyte	Terabyte
	70	80	90	2000

Introdução

- Motivação:

- ▶ Investigação de fenômenos da natureza:

- coleta de dados+simulação = imenso volume de dados

- ▶ Exemplos:

- Meteorologia,
Física de Alta Energia,
Geociências,...

Introdução

- Objetivos:

- ▶ explorar o problema de armazenamento de grandes volumes de objetos muito longos, sujeitos a operações complexas
- ▶ ilustrar a discussão com exemplos concretos

Introdução

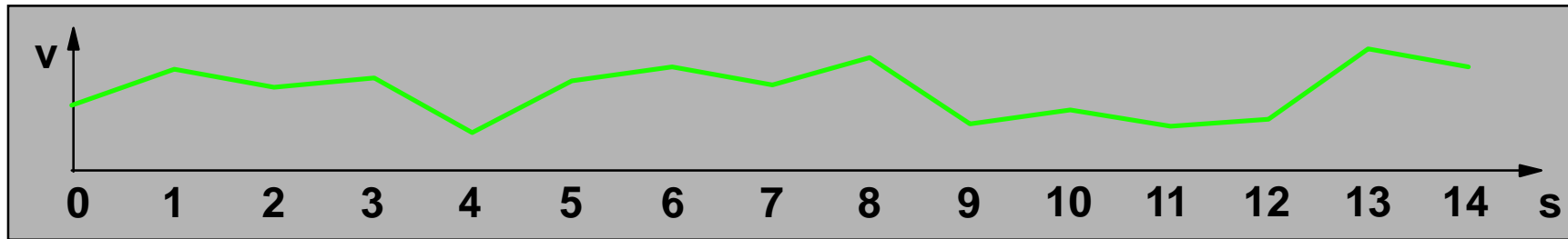
- Situações abordadas:
 - ▶ Armazenamento de Objetos Longos
 - ▶ Sistemas Paralelos de Armazenamento
 - ▶ Cooperativas de Dados

Armazenamento de Objetos Longos

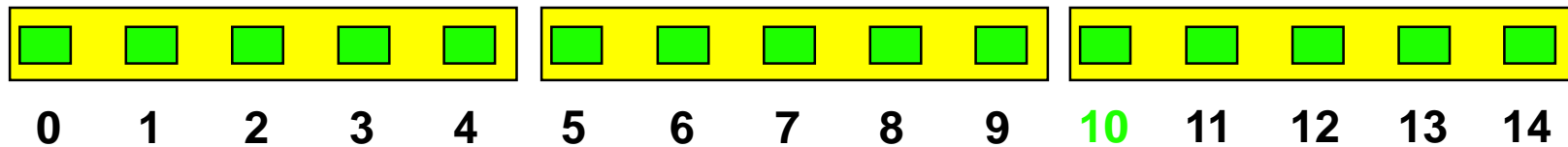
- Objetivo:
 - ▶ projetar métodos de armazenamento especiais para objetos longos que explorem as características das principais operações
- Exemplos:
 - ▶ seqüências de observações
 - ▶ imagens de sensoriamento remoto

Armazenamento de Objetos Longos

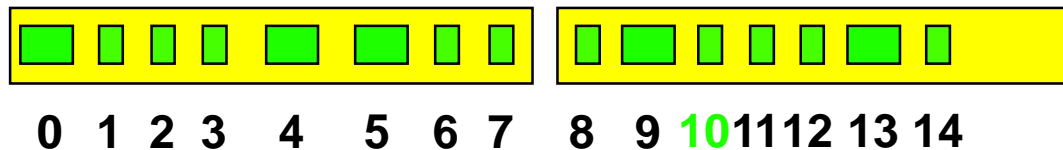
Armazenamento Convencional de Seqüências



Armazenamento sem compactação

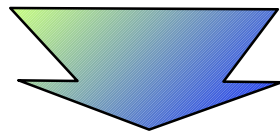
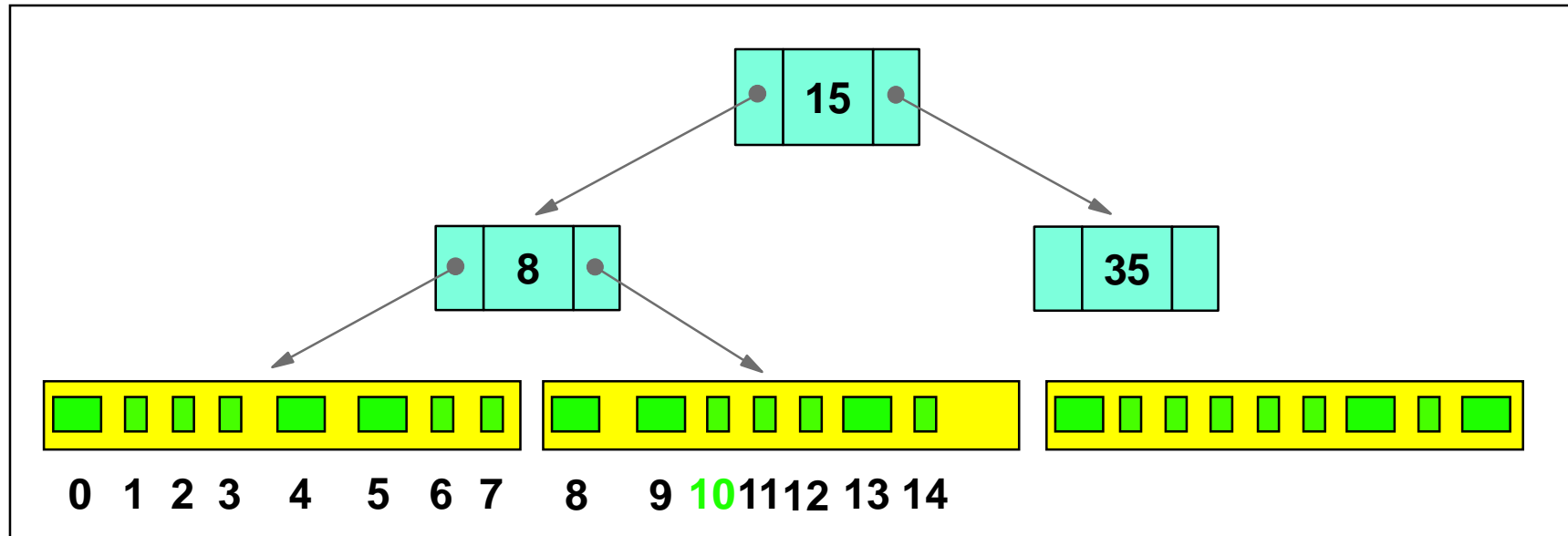


Armazenamento com compactação



Armazenamento de Objetos Longos

Armazenamento Indexado de Seqüências



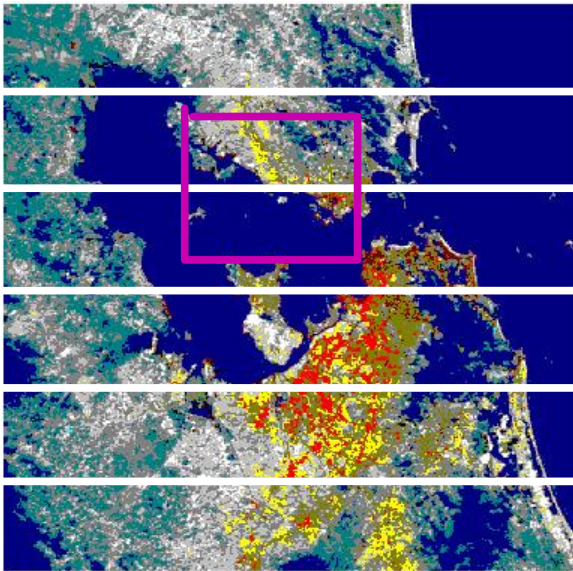
memória secundária

Armazenamento de Objetos Longos

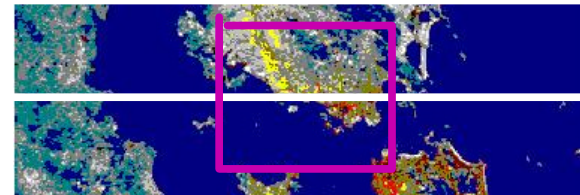
- Imagens de sensoriamento remoto:
 - ▶ Características:
 - imagem = seqüência de matrizes
 - operações: histograma, visualização, correção, segmentação
 - ▶ Estratégias convencionais de armazenamento:
 - matrizes em seqüência / linhas ou colunas em seqüência
 - entrelaçamento de linhas ou colunas
 - entrelaçamento de pixels
 - ▶ Estratégia alternativa de armazenamento

Armazenamento de Objetos Longos

Armazenamento de Imagens por Linhas



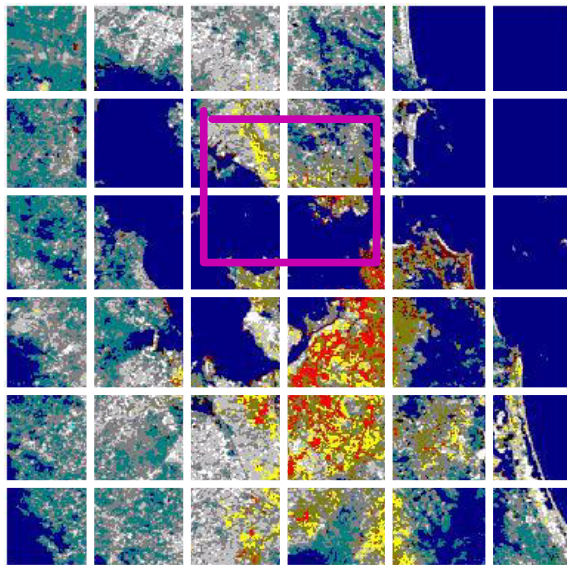
Armazenamento



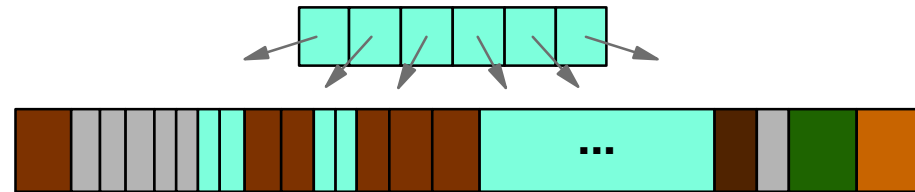
Buffer Pool

Armazenamento de Objetos Longos

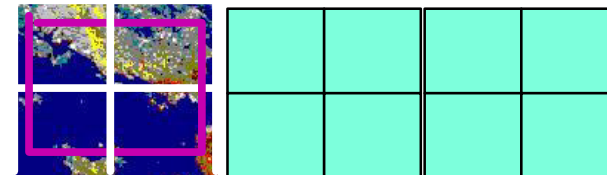
Armazenamento de Imagens em Blocos Compactados



Matriz



Armazenamento



Buffer Pool

Armazenamento de Objetos Longo

■ Resumo:

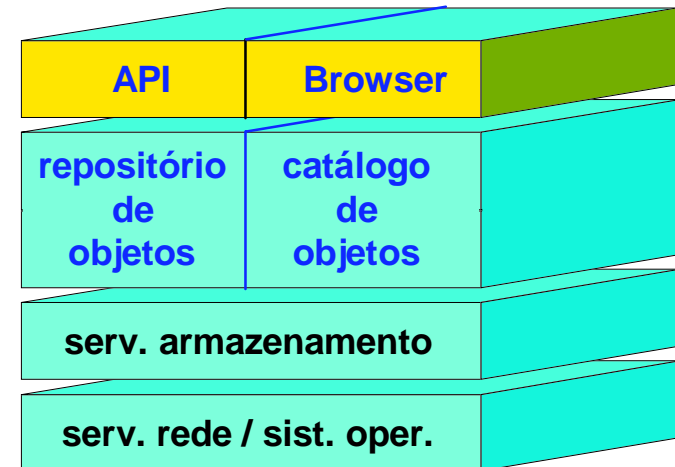
- ▶ fragmentação do conteúdo do objeto
- ▶ compactação dos fragmentos em separado
- ▶ linearização dos fragmentos
- ▶ armazenamento dos fragmentos e criação de estrutura de acesso aos fragmentos
- ▶ acesso mediado por buffer pool de fragmentos
- ▶ API com elenco de operações básicas
- ▶ implementação preferencialmente através de DBMS ou DLS *estendível*

Sistemas Paralelos de Armazenamento

- Objetivo:
 - ▶ projetar subsistemas de armazenamento especiais para grandes volumes de objetos longos
- Exemplo:
 - ▶ High-Performance Storage System (HPSS)

Sistemas Paralelos de Armazenamento

- Questões básicas:
 - ▶ construção do browser e do catálogo
 - ▶ organização dos objetos no repositório
 - ▶ otimização dos serviços de armazenamento
 - ▶ definição dos serviços de rede e sist. operacional



Lawrence Livermore National Lab

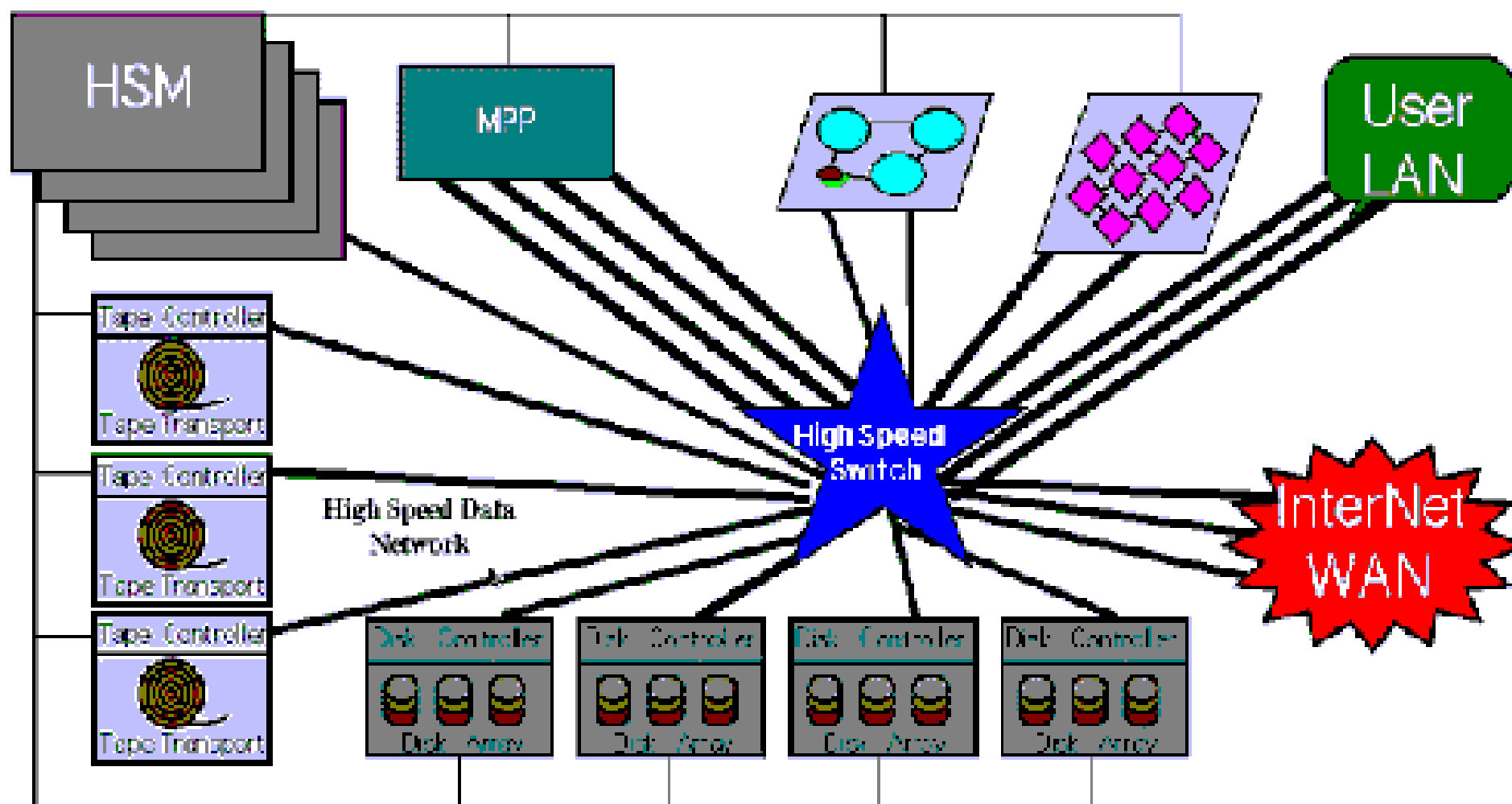
- **Livermore Computing:**

- Missão:

prover um ambiente computacional para os programas prioritários do DOE - especificamente o **ASCI** - *Accelerated Strategic Computing Initiative Program*

processamento	teraflops
transmissão	terabytes
armazenamento	petabytes

Configuração de Hardware



LLNL - National Storage Laboratory

- **Objetivos:**

- desenvolver sistemas distribuídos de armazenamento de alto desempenho
- testar, avaliar e integrar produtos de HW e SW comerciais e facilitar a comercialização de sistemas experimentais

- **Projetos do NSL:**

- High-Performance Storage System (HPSS)
- NSL-UniTree

- **Projetos relacionados:**

- OptiMass
- integração de SGBDs e HPSS

High Performance Storage System

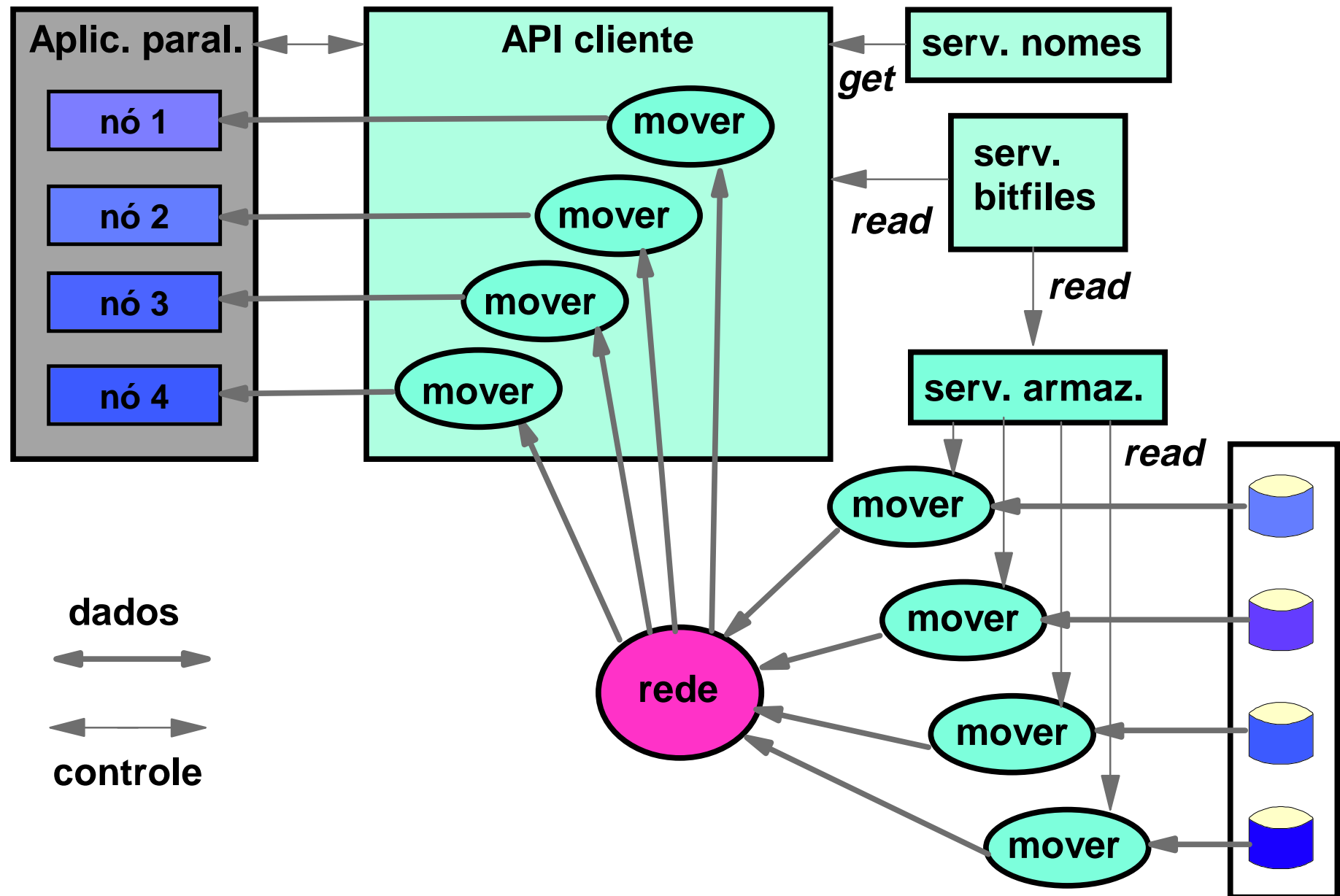
- **Descrição:**

- subsistema de armazenamento paralelo e escalonável para MPPs, bem como supercomputadores tradicionais e *clusters* de estações de trabalho

- **Características:**

- projeto baseado no IEEE Mass Storage System Reference Model, Version 5
- escalabilidade em várias dimensões: taxa de transferência, volume de dados, número de objetos, tamanho dos objetos, distribuição geográfica
- suporte a transferência de dados paralela entre periféricos e processador utilizando *data striping*

HPSS - Leitura Paralela



Cooperativas de Dados

- Objetivo:

- ▶ projetar biblioteca digital distribuída heterogênea para grandes volumes de dados

- Exemplo:

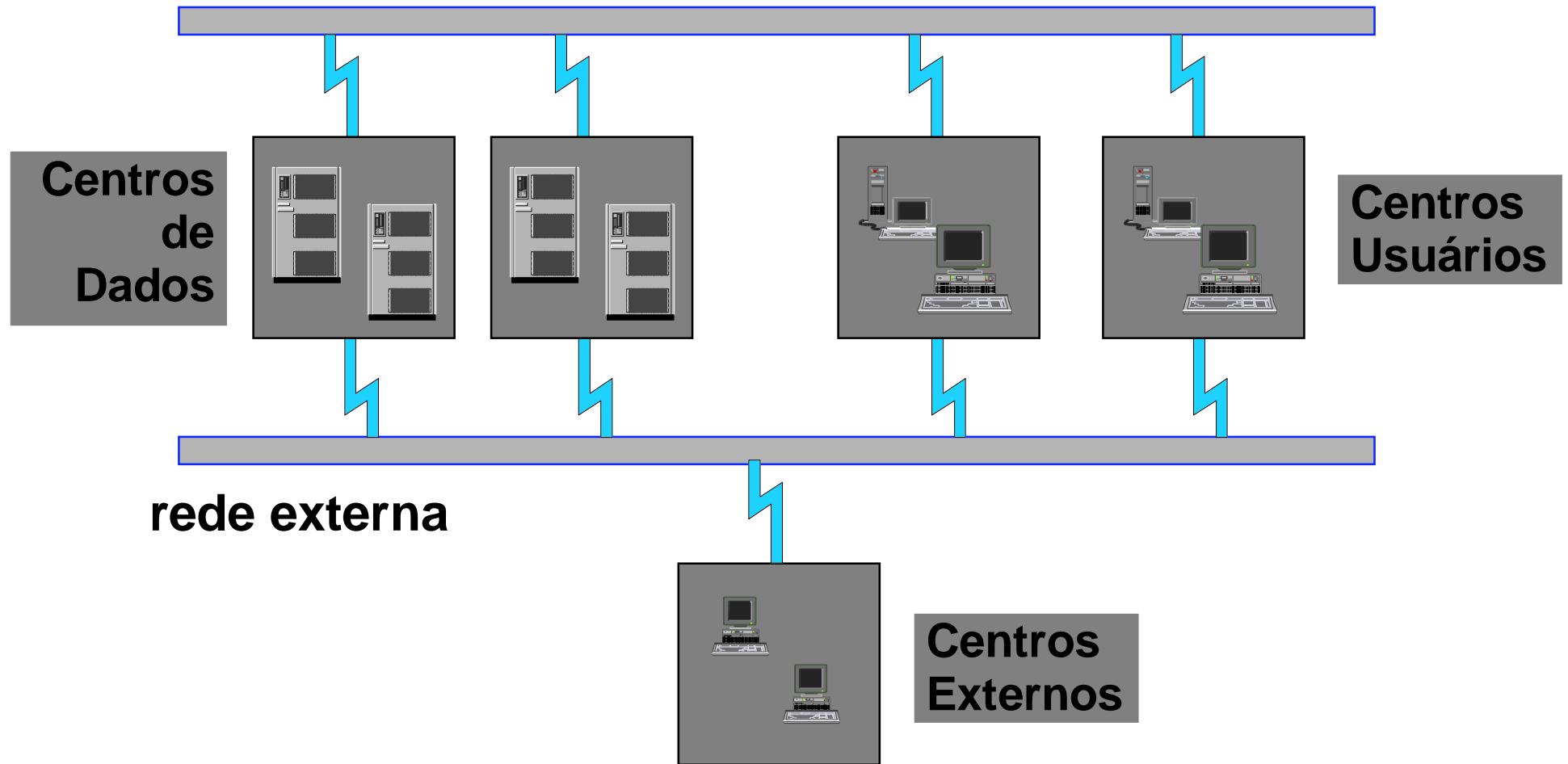
- ▶ EOSDIS - Earth Observation System (EOS) Data and Information System (DIS)

Cooperativas de Dados

- Características:
 - ▶ conjunto de centros de dados, com objetos de naturezas diversas, compartilhados entre usuários
- Abordagem:
 - ▶ bibliotecas digitais federadas
 - ▶ catálogo estendível e hierárquico
 - ▶ cada centro de dados disponibiliza:
 - catálogo próprio
 - repositório de objetos
 - repositório de SW

Cooperativas de Dados

rede interna



Earth Observation System Data and Information System

- **EOSDIS**

- parte do *Mission to Planet Earth Project*
- oferece a uma ampla comunidade de usuários acesso a dados sobre a Terra

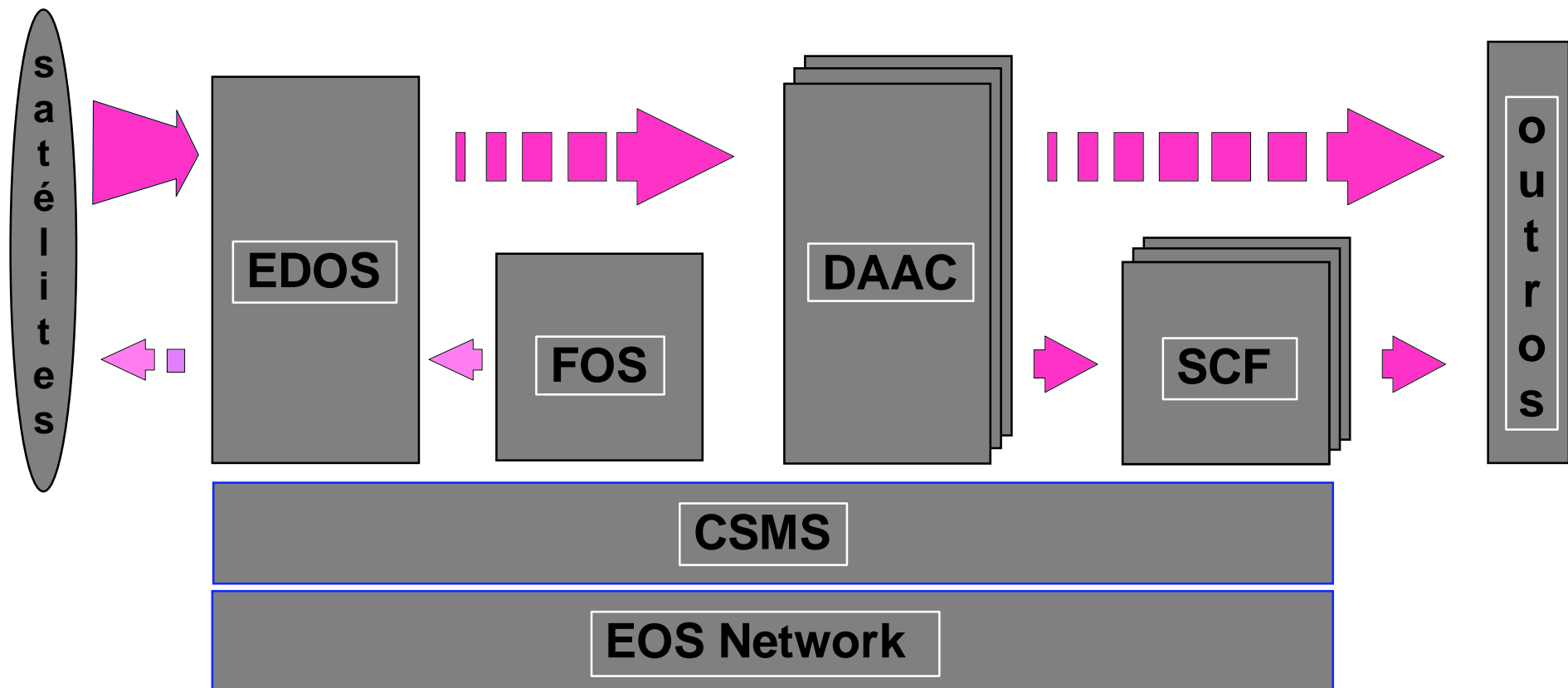
- **Serviços Principais:**

- captura dos dados dos satélites da NASA de sensoriamento remoto da Terra e processamento dos dados de telemetria
- armazenamento e distribuição dos dados gerados
- geração de produtos
- localização e acesso aos dados
- controle dos satélites e instrumentos de bordo

Earth Observation System Data and Information System

EDOS - EOS Data and Oper. System
FOS - Flight Operations Segment

DAAC - Distributed Active Archive Center
SCF - Scientific Computing Facility
CSMS - Comm. and System Mng System



Resumo

- Centro de Computação de Alto Desempenho

processamento	teraflops
transmissão	terabytes
armazenamento	petabytes

- Problemas de armazenamento:

- ▶ métodos de armazenamento para objetos longos
- ▶ sistemas paralelos de armazenamento
- ▶ construção de grandes catálogos de objetos
- ▶ reorganização de dados após captura
- ▶

Resumo

■ Referências selecionadas na Web:

HPSS Tutorial	www.sdsc.edu/hpss/hpss.html
The OPTIMASS project	gizmo.lbl.gov/optimass.html
Intelligent Archive	www.llnl.gov/liv_comp/ia/
Mission to Planet Earth	www.hq.nasa.gov/office/mtpe/
EOSDIS Overview	spsosun.gsfc.nasa.gov/EOSDIS_Over.html
ECMWF Home Page	www.ecmwf.int
IEEE Storage Systems Standards Working Group	http://www.ssswg.org/public_documents.html