

# DBVS realizavimas

---

- Pagrindiniai DBVS komponentai
- Duomenų saugojimas diske

## Seminaro tikslai

- Trumpai apžvelgti pagrindinius DBVS komponentus
- Detaliai nagrinėjami optimalaus duomenų dėstymo diske principai ir algoritmai.
- Seminaras skirtas besidomintiems vidine DBVS architektūra ir funkcionavimo principais.
- Klausytojai turi žinoti DBVS pagrindines sąvokas, SQL ir DDL

# Pranešimo eiga

## □ DBVS komponentai

- Pagrindiniai DBVS komponentai
- Komponentų sąveika

## □ Duomenų dėstymas diske:

- Atminties rūšys, diskinė atmintis
- Efektyvus diskinės atminties naudojimas
- Duomenų apsaugos metodai

# Pagrindiniai DBVS komponentai

Užklausų vykdytojas:

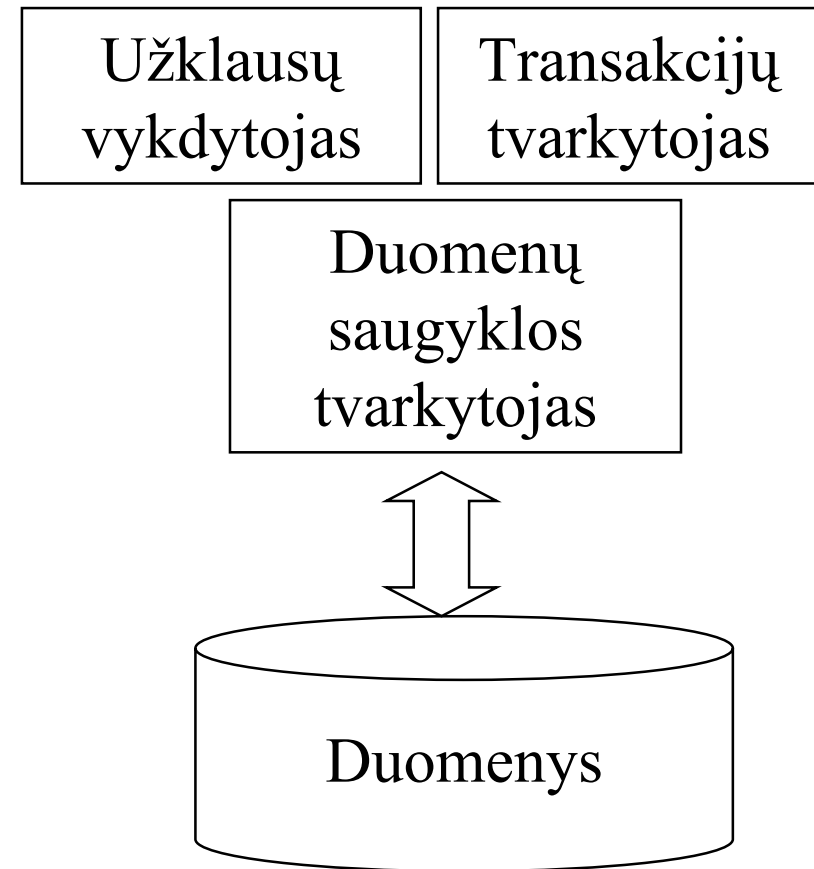
- SQL kompiliatorius
- Vykdomo variklis
- Indeksų, failų, įrašų tvarkytoja

Duomenų saugyklos tvarkytojas:

- Buferių tvarkytojas
- Disko tvarkytojas

Transakcijų tvarkytojas:

- sesijos ir transakcijos
- žurnalai ir atstatymas
- konkurencijos valdymas



# DBVS komponentų sąveika

Užklausa

SQL kompiliatorius

Užklauso planas

Vykdymo variklis

Reikalavimai indeksams, failams ir įrašams

Indeksų/failų/įrašų  
tvarkytojas

Transakcijų  
tvarkytojas

Buferiai

Užraktų lentelė

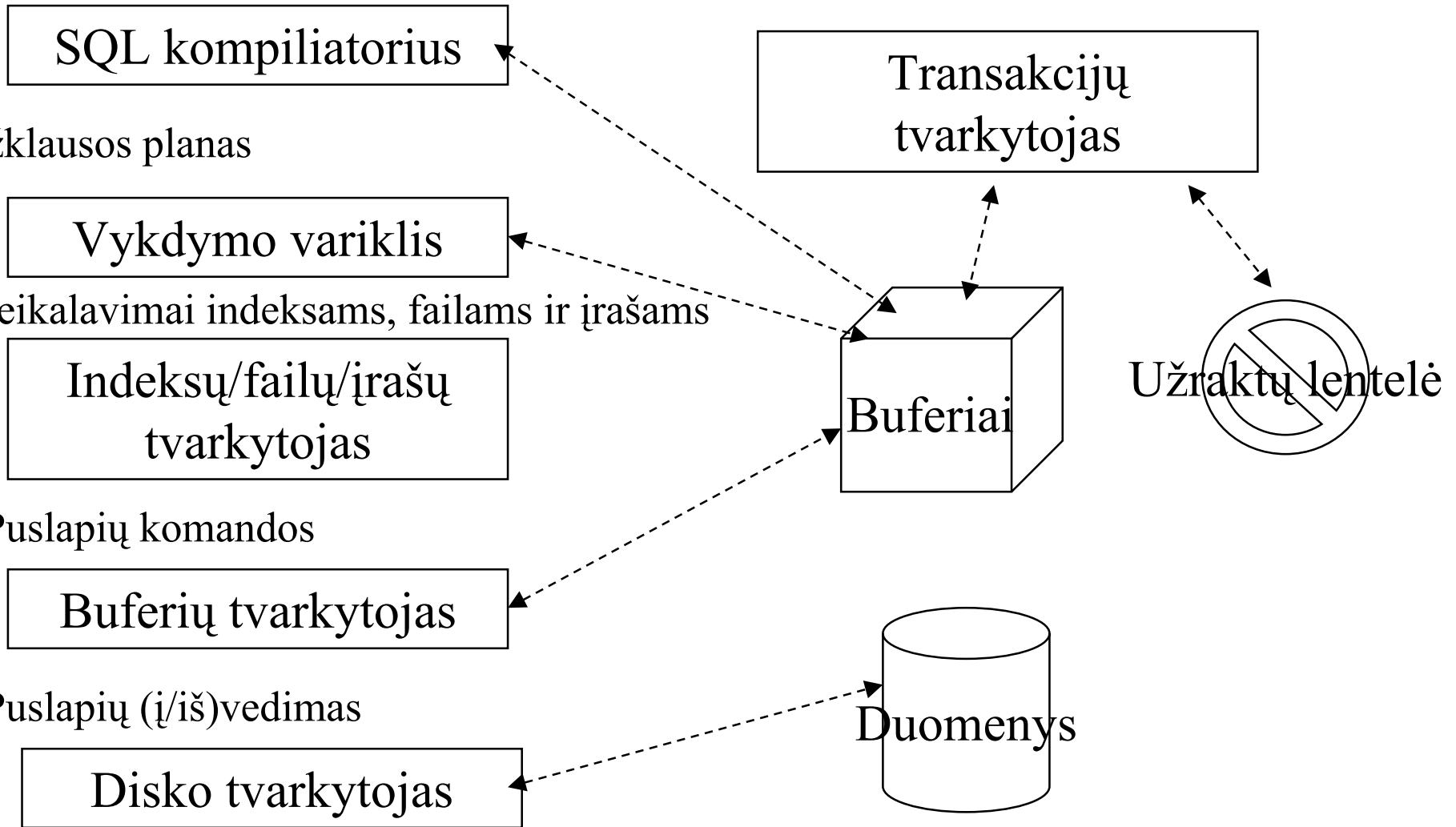
Puslapių komandos

Buferių tvarkytojas

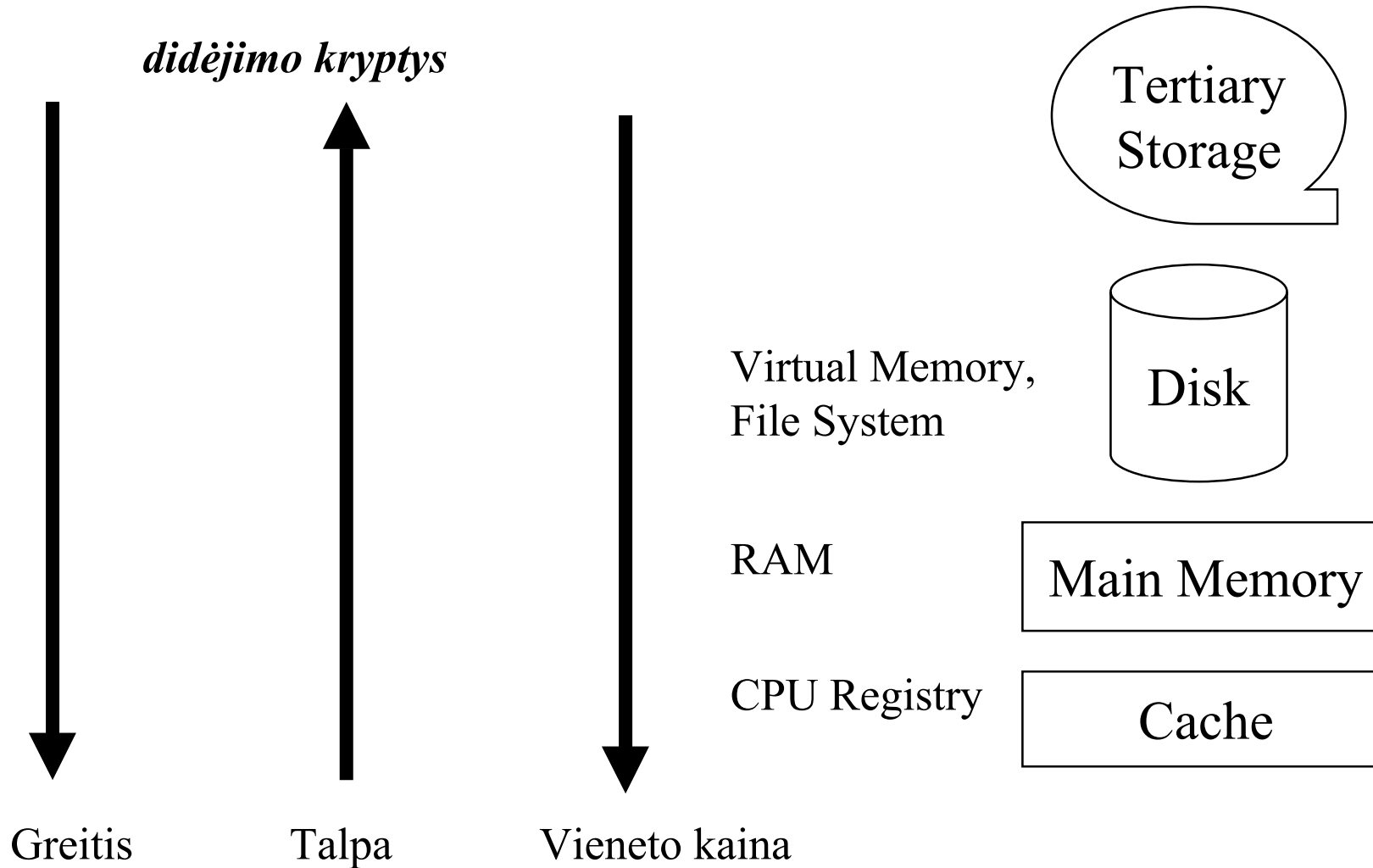
Puslapių (i/iš)vedimas

Disko tvarkytojas

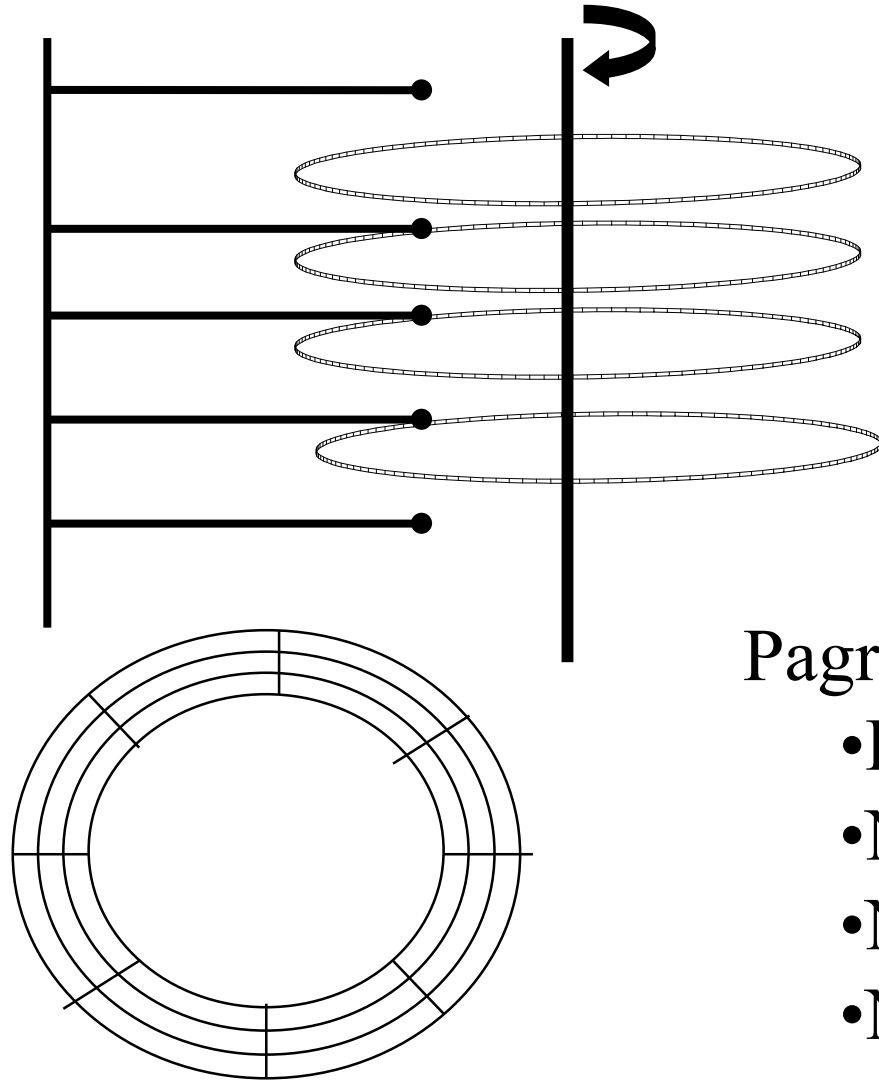
Duomenys



# Atminties tipų hierarchija



# Disko architektūra



## Pagrindiniai terminai:

- Platter, track, gap, block
- Head assembly, cylinder
- Disk controller

## Pagrindinės charakteristikos:

- Rotation speed of Assembly
- Number of Platters per Unit
- Number of tracks per Surface
- Number of Bytes per Track

## Įvedimo / išvedimo operacijų trukmė

□ Kontrolerio instrukcijų vykdymo laikas:

Paprastai labai mažas, milisekundžių dalys. Paprastai ignoruojamas.

□ Cilindro paieška (*seek time*):

Pakankamai didelis. Pajudėti per vieną takelį reikia 1-2 milisekundžių. Pajudėti per visą diską 10-40 milisekundžių.

□ Sektoriaus cilindre paieška (*rotational latency*):

Dabartiniai diskai apsisuka maždaug kartą per 10 milisekundžių.

□ Sektoriaus skaitymo greitis (*transfer time*):

Dabartiniai diskai apsisuka maždaug kartą per 10 milisekundžių. 1 00 000 baitų takelyje. Taigi 10 MB/s.



# Rūšiavimas diskinėje atmintyje

- Paprastas Merge-Sort
  - Atskirų segmentų rūšiavimas  $O(n \log n)$
  - Segmentų suliejimas  $O(n)$

## Two-Phase, Multiway Merge-Sort:

- Pirma fazė:
  - Nuskaitomas maksimalus įrašų skaičius į pagrindinę atmintį. Rūšiavimas pagrindinėje atmintyje.
  - Surūšiuotas posekis talpinamas į naują vietą diskinėje atmintyje.
- Antra fazė:
  - Visų posekių pirmas blokas nuskaitomas į atmintį.
  - Blokai suliejami. Naujas blokas įrašomas į diską.
  - Sekantis blokas nuskaitomas pirmam paibaigus.

## Disko įvedimo/išvedimo operacijų optimizavimas

- Duomenų apjungimas į cilindrus ir takelius
- Keleto diskų panaudojimas
- Diskų veidrodinis atspindys
- Išankstinis skaitymas/rašymas (read-ahead, prefetching)
- Lifo (elevator) algoritmas

## Diskų gedimai

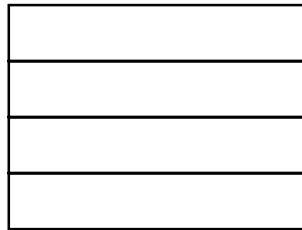
- Nenuolatinės (intermittent) sektoriaus klaidos.
- Nuolatinė (permanent) sektoriaus klaidos.
- Viso disko lūžis (disk crash)

# Apsauga nuo duomenų praradimo: kontrolinės sumos

- Parity check
- Stable storage
- RAID

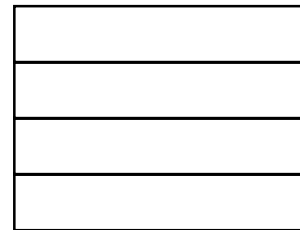
# RAID lygiai

☐ 1 lygio RAID (mirroring):

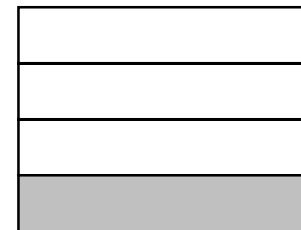
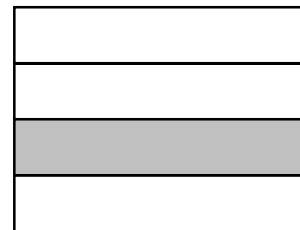
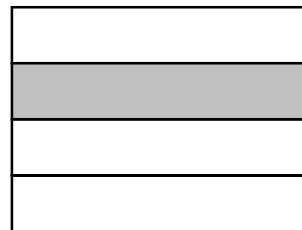
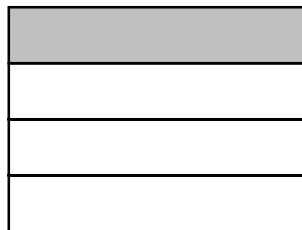


- ☐ Duomenys
- Papildoma kontrolė

☐ 4 lygio RAID:



☐ 5 lygio RAID:



## Kito seminario tema

- Fiksuoto ir kintamo ilgio duomenų tipų realizavimas
- BLOB realizavimas
- Duomenų struktūros įrašams ir blokams