Business School
UNIVERSITY OF COLORADO DENVER

Information Systems Program

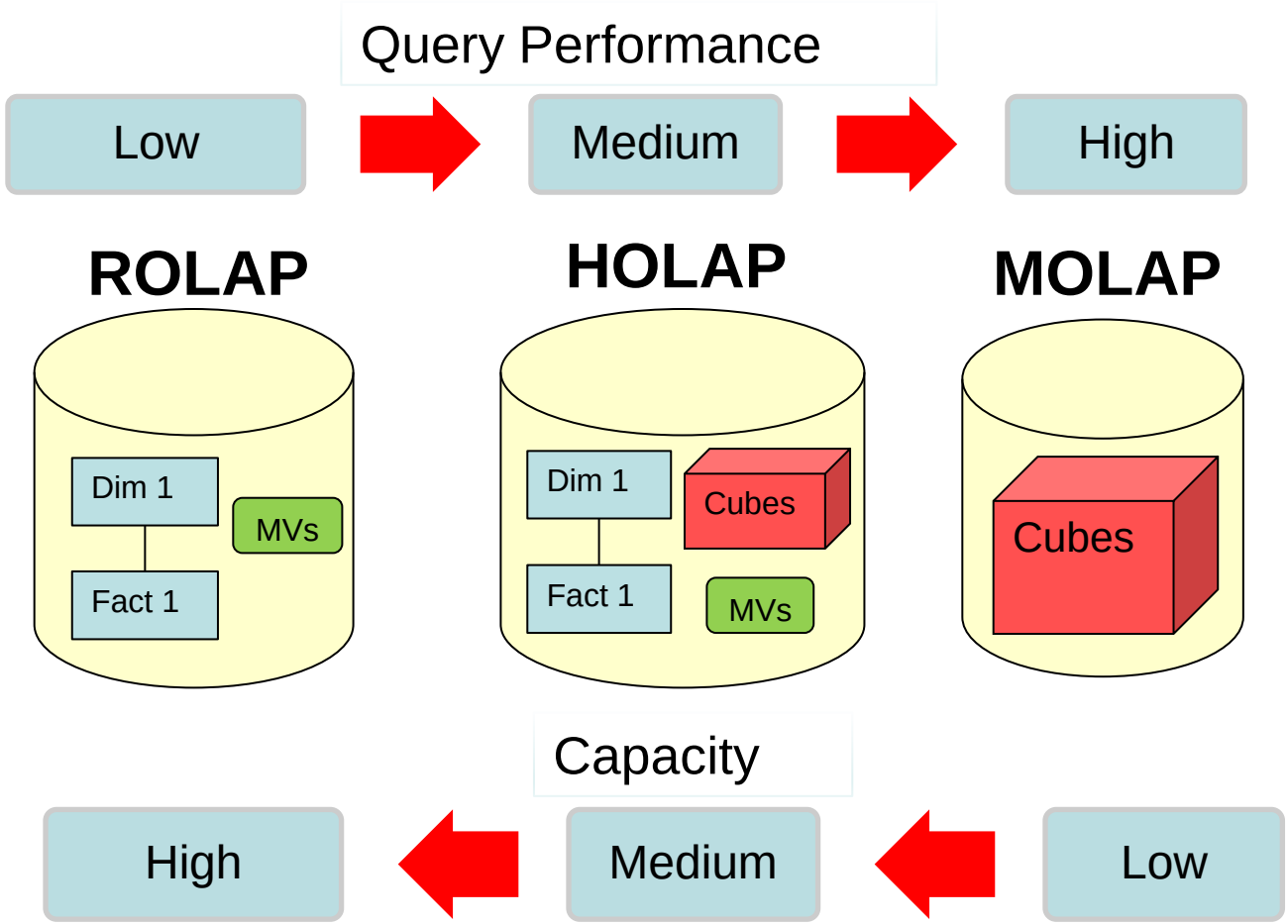
Module 5
Physical Design and Governance of Data Warehouses

Lesson 1: Storage Architectures

Lesson Objectives

- Compare and contrast physical architectures for implementing data warehouses
- Explain components of bitmap join indexes for star join queries
- Explain advantages of columnstore structures for business intelligence queries

Storage Architectures



Bitmap Index

- Useful for stable columns with few values
- Bitmap:
  - String of bits: 0 (no match) or 1 (match)
  - One bit for each row: 110010010001
- Bitmap index record
  - Column value: “USA”
  - Bitmap: 110010010001
  - Convert a bit position into a row identifier

Bitmap Index Example

SSCustomer

RowId	CustNo	…	CustCountry
1	098-55-1234		USA
2	123-45-6789		USA
3	456-89-1243		Canada
4	111-09-0245		Mexico
5	931-99-2034		USA
6	998-00-1245		Mexico
7	287-44-3341		Canada
8	230-21-9432		USA
9	321-44-5588		Mexico
10	443-22-3356		Canada
11	559-87-3211		Mexico
12	220-44-5688		USA

Bitmap Index on CustCountry

CustCountry	Bitmap
USA	110010010001
Canada	001000100100
Mexico	000101001010

Bitmap Join Index Example

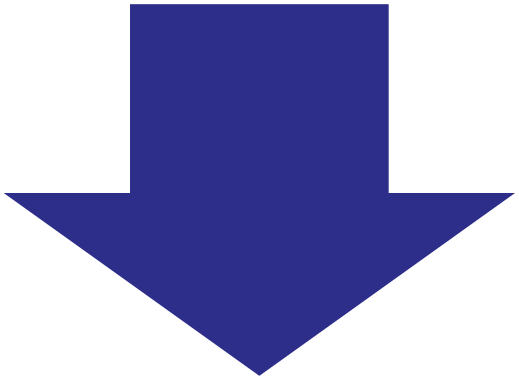
SSSales

RowId	SalesNo	…	CustNo
1	1111		098-55-1234
2	1234		123-45-6789
3	1345		456-89-1243
4	1599		111-09-0245
5	1807		931-99-2034
6	1944		998-00-1245
7	2100		287-44-3341
8	2200		230-21-9432
9	2301		321-44-5588
10	2487		443-22-3356
11	2500		559-87-3211
12	2600		220-44-5688
13	2703		098-55-1234
14	2801		123-45-6789
15	2944		456-89-1243
16	3100		111-09-0245
17	3200		931-99-2034
18	3258		998-00-1245
19	3302		287-44-3341
20	3901		230-21-9432
21	4001		321-44-5588
22	4205		443-22-3356
23	4301		559-87-3211
24	4455		220-44-5688

Bitmap Join Index on CustCountry

CustCountry	Bitmap
USA	110010010001110010010001
Canada	001000100100001000100100
Mexico	000101001010000101001010

Row Storage versus Column Storage

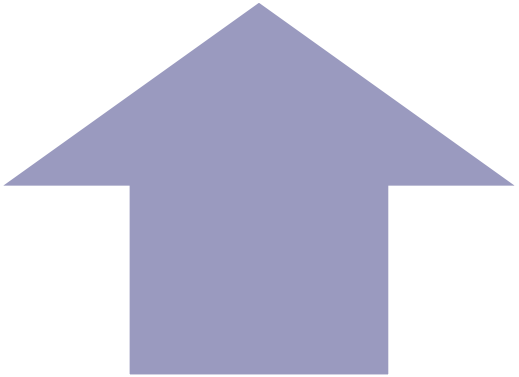


Row Storage

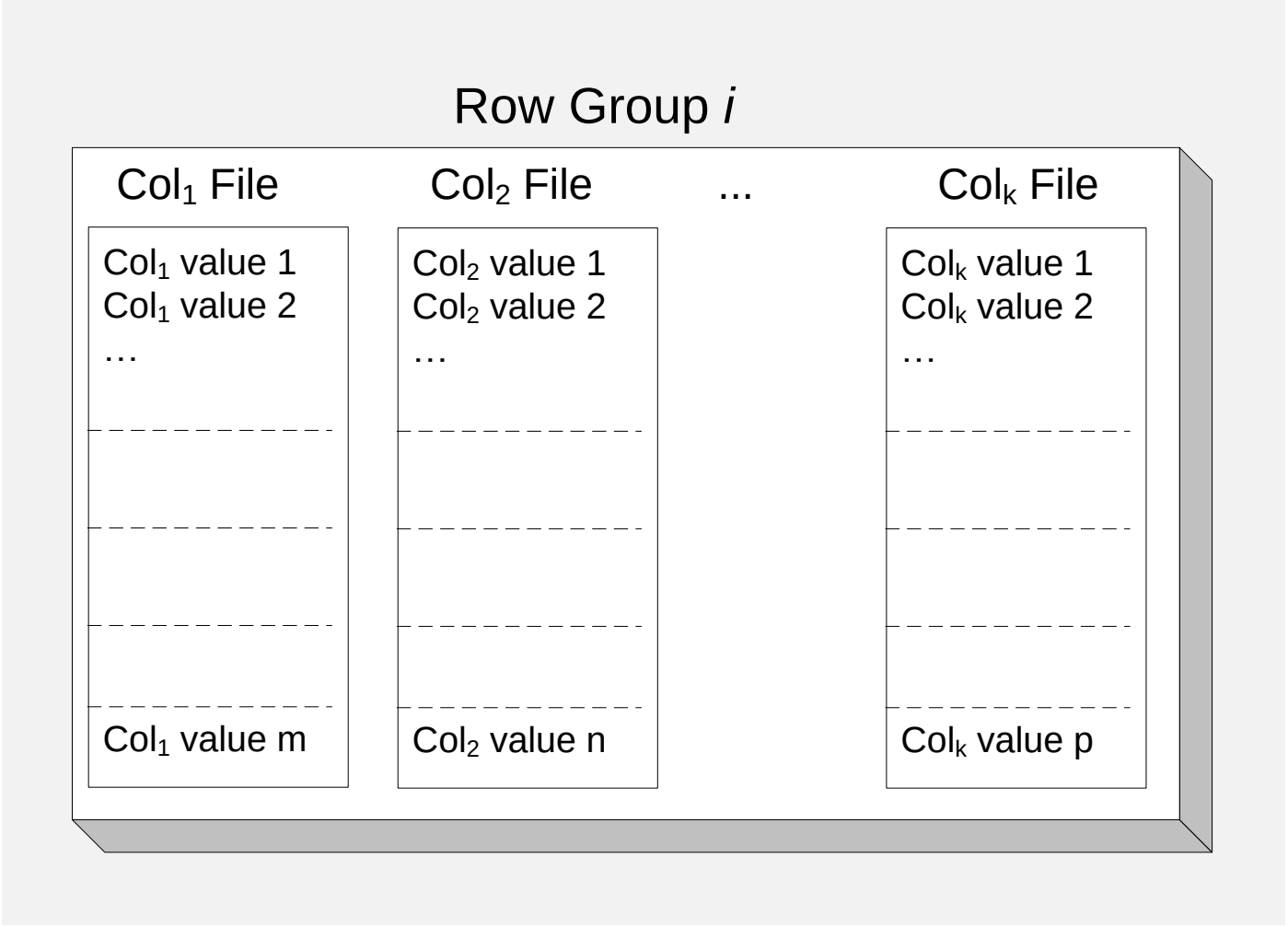
- Entire rows in physical records
- Full table scans for business intelligence queries
- Foundation storage approach for most enterprise DBMSs

Column Storage

- Specialized storage structure
- Eliminates table scan for typical business intelligence queries
- Optional storage approach in enterprise DBMSs



Columnstore File Structure



Columnstore Index

- Store columns in physical records
- Compression for duplicate values
- Optional sorting

Sales fact row group

SalesNo	SalesUnits	SalesAmt
1111	10	1200
1122	20	900
1133	50	850
1144	10	1400
1155	20	1150
1166	50	1400

Columnstore indexes

SalNo
1111
1122
1133
1144
1155
1166

SalesUnits
10 × 2
20 × 2
50 × 2

SalesAmt
850
900
850
1150
1400 × 2

Summary

- Basic architectures: MOLAP, ROLAP, HOLAP
- Bitmap indexes for stable columns with few values
- Columnar storage becoming common for large fact tables