Databases Overview

Richard Park
Databasics

What is a database
Types of databases
    Relational databases

Dyalog ↔ Database
Databasics

A database is a system for storing data

Previously:
   Filing Systems

Nowadays:
   Databases → Database Management System
   (DBMS)
Database Management System
1. A structured arrangement of data
2. Query language
   a domain specific language (DSL) for:
   - reading
   - writing
   - manipulating data
Organisation

**Relational**
- Tables

**Non-relational** (the rest)
- Time series
- Key-value
- Document / Object (e.g. JSON)
- Graph

...
**Organisation**

<table>
<thead>
<tr>
<th>PCode</th>
<th>Name</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>P234</td>
<td>Gruel</td>
<td>Food</td>
</tr>
<tr>
<td>P695</td>
<td>Water</td>
<td>Drink</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CID</th>
<th>Address</th>
<th>Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>C4124</td>
<td>2 Main St.</td>
<td>Alice</td>
</tr>
<tr>
<td>C3156</td>
<td>54 Nice Rd.</td>
<td>Bob</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>RID</th>
<th>Name</th>
<th>Address</th>
</tr>
</thead>
<tbody>
<tr>
<td>SH01</td>
<td>Shoppz</td>
<td>Shoppz.uk</td>
</tr>
<tr>
<td>SH02</td>
<td>Buyaz</td>
<td>Buyaz.com</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>TID</th>
<th>PCode</th>
<th>CID</th>
<th>RID</th>
<th>Price</th>
<th>Qty</th>
<th>DateTime</th>
</tr>
</thead>
<tbody>
<tr>
<td>354</td>
<td>P234</td>
<td>C3156</td>
<td>SH02</td>
<td>$5.04</td>
<td>3</td>
<td>2020-05-16</td>
</tr>
<tr>
<td>355</td>
<td>P695</td>
<td>C4124</td>
<td>SH01</td>
<td>$4.95</td>
<td>4</td>
<td>2020-06-03</td>
</tr>
<tr>
<td>356</td>
<td>P234</td>
<td>D3156</td>
<td>SH01</td>
<td>$2.59</td>
<td>2</td>
<td>2020-06-04</td>
</tr>
</tbody>
</table>
Let's make a database
Database creation

Efficiency

Integrity
Database creation

Efficiency

Integrity

\rightarrow \text{Reduce redundancy}
Example: Entity-Relationship Diagram
<table>
<thead>
<tr>
<th>TID</th>
<th>PCode</th>
<th>CID</th>
<th>RID</th>
<th>Price</th>
<th>Qty</th>
<th>DateTime</th>
</tr>
</thead>
<tbody>
<tr>
<td>354</td>
<td>P234</td>
<td>C3156</td>
<td>SH02</td>
<td>$5.04</td>
<td>3</td>
<td>2020-05-16</td>
</tr>
<tr>
<td>355</td>
<td>P695</td>
<td>C4124</td>
<td>SH01</td>
<td>$4.95</td>
<td>4</td>
<td>2020-06-03</td>
</tr>
<tr>
<td>356</td>
<td>P234</td>
<td>D3156</td>
<td>SH01</td>
<td>$2.59</td>
<td>2</td>
<td>2020-06-04</td>
</tr>
</tbody>
</table>

**PCodes and Types:**

- P234: Gruel, Food
- P695: Water, Drink

**Retailers:**

- SH01: Shoppz, Shoppz.uk
- SH02: Buyaz, Buyaz.com

**Customers:**

- C4124: Alice, 2 Main St.
- C3156: Bob, 54 Nice Rd.
Query Language

Relational

Structured Query Language (SQL)

Non-relational AKA NoSQL (misnomer-ish)

GraphQL

RESTful APIs

Also sometimes SQL
APL Arrays

Data fits in WS

Single-user update

Update from APL
Advantages over "just APL arrays"

Lots o' data
Data integrity
Query optimisers
Multi-user access
Access control / security
Compatibility
Relational Database Paradigms

**OLTP (OnLine Transactional Processing)**

- Inserts, updates, deletes
- Simultaneous writes from multiple users

**OLAP (OnLine Analytical Processing)**

- Write less, read more
- Better suited for vectorised operations
Options

APL Matrix
APL Inverted Table
Dataframe system
DBMS
Dataframes
Python + Pandas + Py'n'APL
github.com/Dyalog/pynapl
R + data.tables + RSConnect
github.com/kimmolinna/rsconnect
Data locality

DBMS

Matrix
APL WS
Inverted Table

MAP

Component files

Dataframes
Non-APL Programming Languages
Examples

APL

Nested Matrix
Nested Matrix

Query Language: APL
Intuitive
Convenient
Performance suffers with lots of data
Examples

APL

Nested Matrix

Inverted Table
Inverted Table

AKA column-oriented or "columnar database"

Query Language: APL

Slightly more complex DSL (might use 8⍷)

Relatively convenient

Good analytical performance
Inverted Table

Inverted Tables / Roger Hui / Dyalog '18
dyalog.tv/Dyalog18/?v=IOWDkqKbMwk

vecdb
github.com/Dyalog/vecdb

FlipDB
carlislegroup.com/products/flipdb
Examples

APL
- Nested Matrix (OLTP)
- Inverted Table (OLAP)
- Flat Inverted Table (Fast OLAP)

DBMS
- MariaDB (OLTP)
- ClickHouse (OLAP)
Usage examples

APL Arrays
SQAPL
LOADDATA.dws
SQA.Do
Low level interface
Dyalog APL SQAPL Interface Guide
Version 6.3
Application Programming Interfaces

C & C++ Connectors
The Client Library for Developing MariaDB Applications with the C Programming Language.

Java Connector
Information on the JDBC Driver for MariaDB.

.NET Connector
Connecting to MariaDB from .NET applications.

Node.js Connector
Information on javascript Node.js connector

ODBC Connector
MariaDB Connector/ODBC is a standardized ODBC database driver.

Perl DBI
Connecting to MariaDB using the Perl DBI Module.

PHP
Connecting to MariaDB from PHP Applications.
Lightweight, advanced connectors for high-performance data access and data streaming.

MariaDB Connector/ODBC is a standard ODBC 3.5 driver for building cross-platform applications on top of MariaDB Platform, complete with encrypted connections via TLS/SSL. MariaDB Connector/ODBC is LGPL.

Product | ODBC connector
--- | ---
OS | MS Windows 64-bit

Release Notes
Show All Files

ODBC Connector
/mariadb-connector-odbc-3.1.13-win64.msi

11.21 MB Download
Create a new Data Source to MariaDB

Welcome to the MariaDB ODBC Data Source Wizard!

This wizard will help you to create an ODBC data source that you can use to connect to a MariaDB server.

What name do you want to use to refer to your data source?

Name: 

How do you want to describe the data source?

Description: 
Database

A system of organised data

+ 

A Query Language
Types

Organisation & Query Language

Relational

Non-relational
Options

APL Matrix
Inverted Tables
Dataframes
DBMS
So much more to cover...

Relational databases
  Inverted tables
  Permanent storage (component files, memory mapped...)
  SQAPL and related tools

Non-relational databases
  Time series
  Key-value (dictionaries)
  Document / Object
  Graph
  ...

Non-ODBC interaction
dyalog.com/dates-for-your-diary.htm

July 15  BAA Open Session
britishaplassociation.org/webinar-schedule-2021

July 30  APL Problem Solving Competition Closes
  dyalogaplcompetition.com

August 01  APL Campfire
  apl.wiki/campfire

August 05  Error Handling Part 4
  dyalog.tv