

Migration to Open Source Databases

Jutta Horstmann

www.osdbmigration.org

whoami

- Unix/Linux sysadmin
- DBA, developer (Informix)
- DB developer (Oracle)
- Web stuff (MySQL, PostgreSQL)
- Claim to Fame: OpenUsability.org
- Comp Sci Diploma Thesis:
Migration to Open Source Databases



Agenda

What's this?

Why migrate?

Where to migrate?

How to migrate?

- About Migration
- Migration: Pros and Cons
- Open Source Databases: Comparison
- Case Studies
- Workflow: Activities & Assets
- Obstacles
- Tools

Migration?

What's this?

Why migrate?

Where to migrate?

How to migrate?

*"The **process** of changing from the use of one platform, environment, IT system, etc., to another, esp. in such a way as to **avoid interruptions in service.**"*

(Oxford English Dictionary)

Concepts

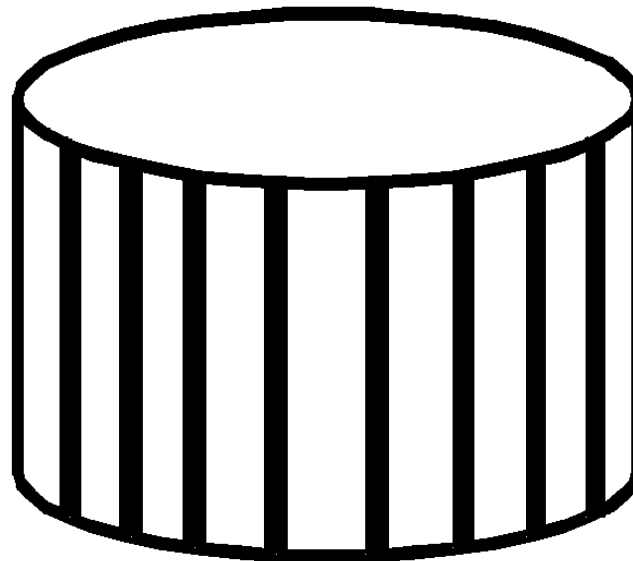
What's this?

Why migrate?

Where to migrate?

How to migrate?

Source System



Concepts

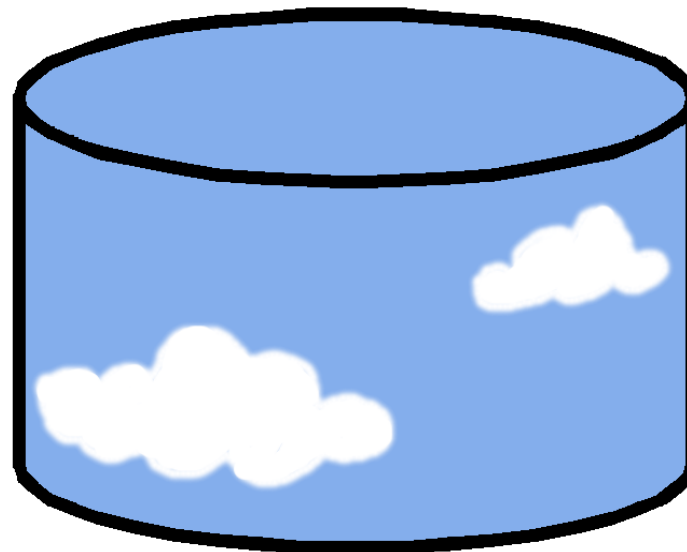
What's this?

Why migrate?

Where to migrate?

How to migrate?

Target System



Concepts

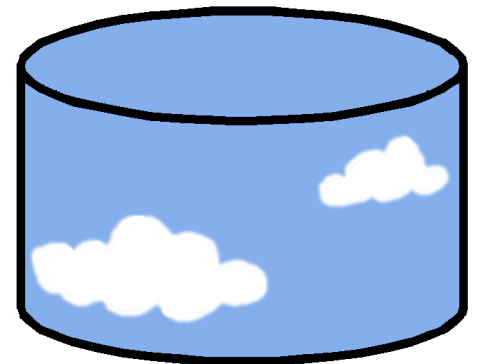
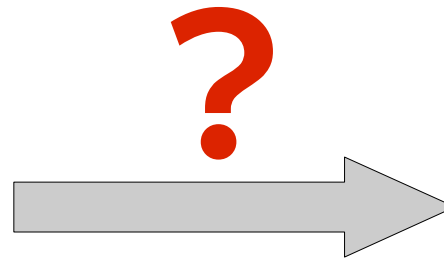
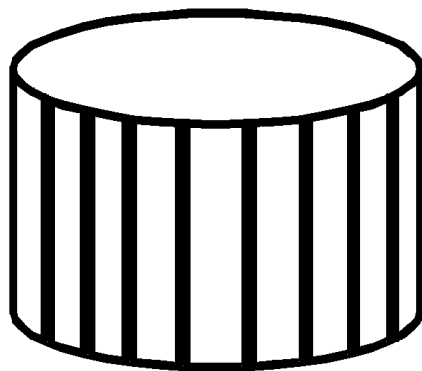
What's this?

Why migrate?

Where to migrate?

How to migrate?

Migration



Concepts

What's this?

Why migrate?

Where to migrate?

How to migrate?

Migration



Migration: Objectives

What's this?

Why migrate?

Where to migrate?

How to migrate?

The Target System:

- same functionality
- extendable
- incorporate all data
- modern hardware, software, architecture

Migration: Objectives

What's this?

Why migrate?

Where to migrate?

How to migrate?

The Migration Workflow:

- Minimize risk
- Stay on budget
- Deliver in due time
- Minimize downtime

Agenda

What's this?

Why migrate?

Where to migrate?

How to migrate?

Why migrate?

Migration: Pros and Cons

What's this?

Why migrate?

Where to migrate?

How to migrate?

No

- Cost? Time? Effort?
- Lack of roadmap
- Licensing
- ISV Support
- Maintenance
- Accountability
- Features

Yes

- Features?
- TCO
- Open Source
- Standards
- Security
- Independence

Agenda

What's this?

Why migrate?

Where to?

-- Candidats

-- History

-- Features

-- Case Studies

How to migrate?

Where to migrate?

→ 5 Open Source Databases

→ RDBMS Timeline

→ Enterprise-Level?

→ Feature Comparison

→ Performance/Scalability

→ Case Studies

Open Source Databases

What's this?

Why migrate?

Where to?

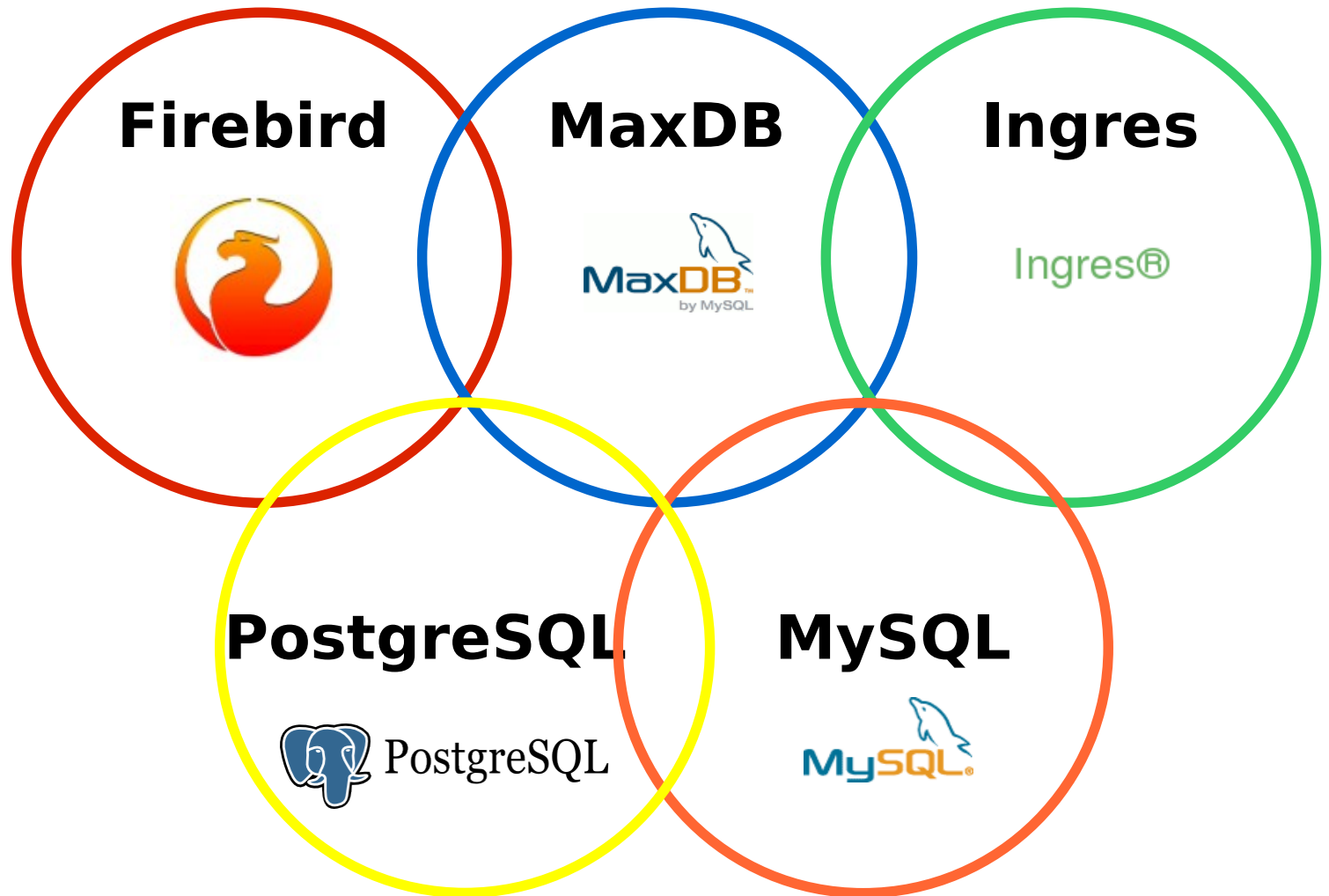
-- **Candidats**

-- History

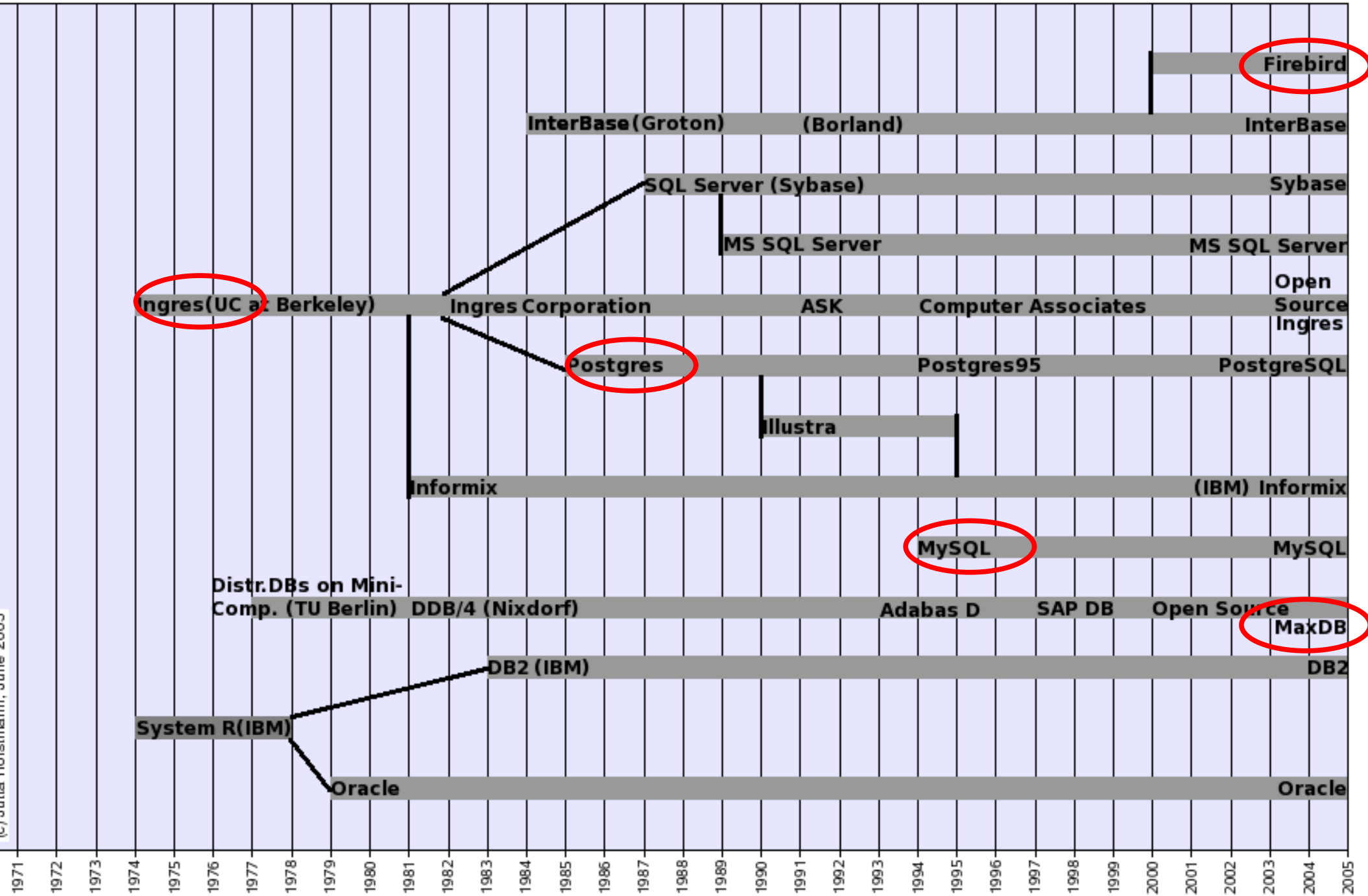
-- Features

-- Case Studies

How to migrate?



RDBMS Timeline



Requirements

What's this?

Why migrate?

Where to?

-- Candidats

-- History

-- **Features**

-- Case Studies

How to migrate?

Advanced data integrity mechanisms

+ Advanced database objects

+ Advanced SQL features

+ Advanced features, tools, support for

























- Performance and scalability

- (High) Availability

- Security

= Enterprise Level Database System

OSDB Features

	Firebird 1.5	Ingres r3	MaxDB 7.6	MySQL 5.0	PostgreSQL 8.1
Advanced Indexing				 MyISAM	
GIS support				 MyISAM	
MVCC				 InnoDB	
Two phase commit					
User defined data types					

Performance / Scalability

	Firebird 1.5	Ingres r3	MaxDB 7.6	MySQL 5.0	PostgreSQL 8.1
Tablespaces					
Table Partitioning					 Constraint Excl.
Parallelization				 Cluster	
Built-In Clustering					
Built-In Load Balancing					

Case Studies, Examples (1)

What's this?

Why migrate?

Where to?

-- Candidats

-- History

-- Features

-- **Case Studies**

How to migrate?

- Oracle → PostgreSQL
 - Assoc. of South Tyrolean municipalities
 - Lucent Technologies
- Oracle → MaxDB
 - Enercon
- Oracle → MySQL
 - Weather.com
 - NASA Acquisition Internet Service
 - Lycos Europe
- Evaluation (Oracle Data Warehouse)
 - ImmobilienScout, Berlin

Case Studies, Examples (2)

What's this?

Why migrate?

Where to?

-- Candidats

-- History

-- Features

-- **Case Studies**

How to migrate?

- Informix → PostgreSQL
 - Hanstholm Municipality, Denmark
 - US National Weather Service
 - Social Science Research Center, Berlin
- Informix → MySQL
 - Digital Library Project (DLP), Berkeley
- MS SQL → MySQL
 - AmazonTech Flight Operations

Agenda

What's this?

Why migrate?

Where to?

How to?

-- Workflow

-- Step by Step

-- Obstacles

-- Tools

How to migrate?

- Things to see and do:

Activities and Assets

- Migration workflow step by step
- Obstacles
- Tools

Activities

What's this?

Why migrate?

Where to?

How to?

-- **Workflow**

-- Step by Step

-- Obstacles

-- Tools

- Analysis
 - Analyze the source system
- Design
 - Design mappings and conversions between source and target
- Implementation
 - Implement mappings in the target
- Test

Assets

What's this?

Why migrate?

Where to?

How to?

-- **Workflow**

-- Step by Step

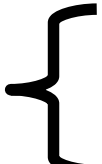
-- Obstacles

-- Tools

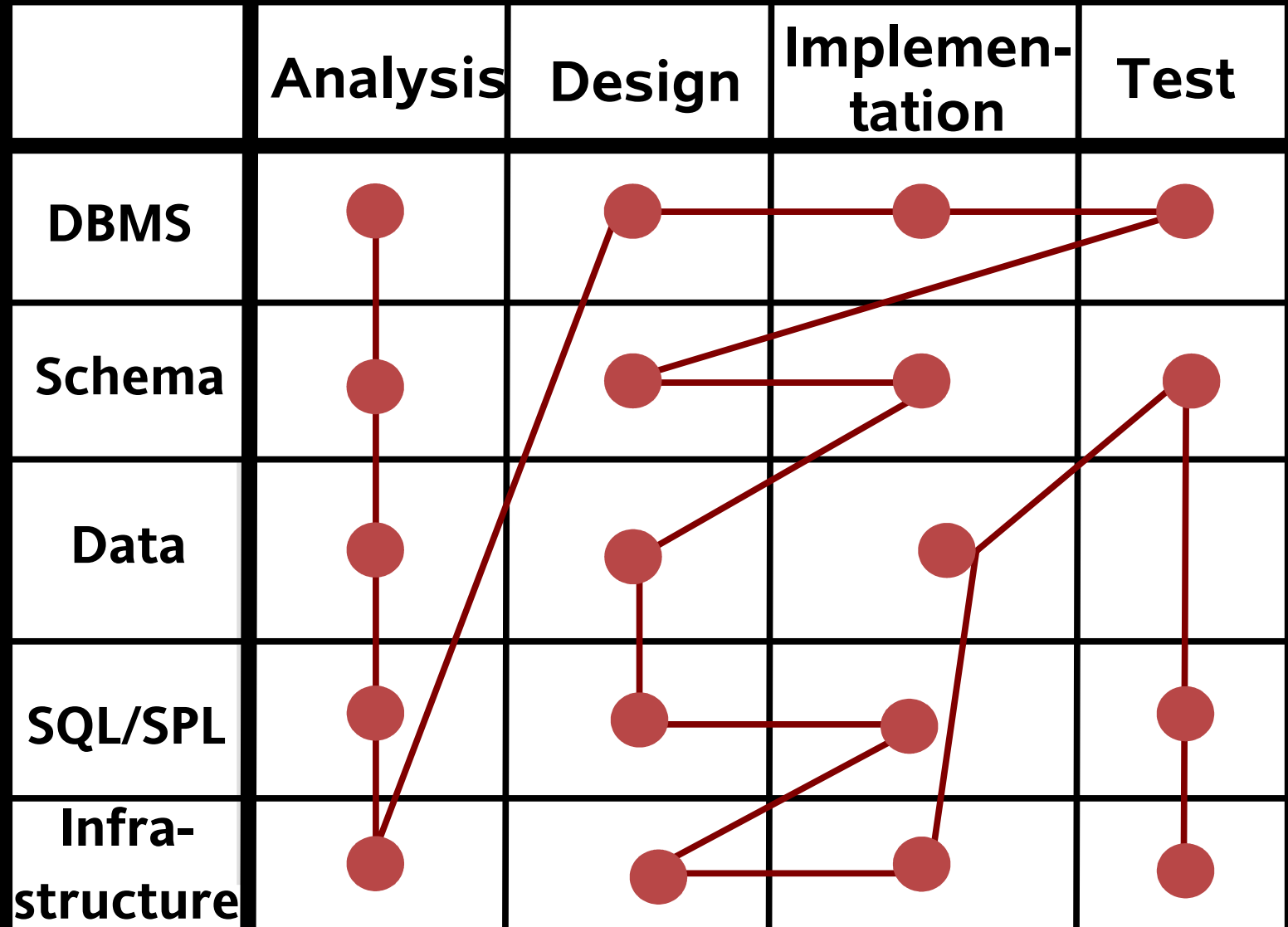
- DBMS software
- Schema(s)
 - DDL, data types, implicit constructs
- Data
- SQL / SPL
 - Client SQL, Server Stored Procedures
- Infrastructure
 - Tools, Policies, Tuning, Maintenance...

Migration Workflow

Activities



Assets



DBMS Software: Analysis

What's this?

Why migrate?

Where to?

How to?

-- Workflow

-- **Step by Step**

- **DBMS**
- Schema
- Data
- SQL
- Infrastructure

-- Obstacles

-- Tools

Checklist:

- ✓ Target DBMS Software: Support for current hardware / OS?
- ✓ Raw disk partition support
- ✓ Locale configuration
- ✓ Encoding
- ✓ Multithreading / -processor support

DBMS Software: Mappings

What's this?

Why migrate?

Where to?

How to?

-- Workflow

-- **Step by Step**

- **DBMS**
- Schema
- Data
- SQL
- Infrastructure

-- Obstacles

-- Tools

Example: Locale configuration

Informix:

```
setenv DB_LOCALE de_de.8859-1
```

```
setenv CLIENT_LOCALE de_de.8859-1
```

PostgreSQL:

- Server Default or
- `initdb --locale=de_DE`

DBMS: Implementation

What's this?

Why migrate?

Where to?

How to?

-- Workflow

-- **Step by Step**

- **DBMS**

- Schema

- Data

- SQL

- Infrastructure

-- Obstacles

-- Tools

- Install Open Source database on target platform
- Configure it based on the mappings

Schema: Analysis

What's this?

Why migrate?

Where to?

How to?

-- Workflow

-- **Step by Step**

- DBMS

- **Schema**

- Data

- SQL

- Infrastructure

-- Obstacles

-- Tools

Checklist:

- ✓ Quality of current design
- ✓ Constraints in schema or application?
- ✓ Indexing
- ✓ Data types (available? size? meaning?)
- ✓ Reserved words
- ✓ DDL syntax differences

Schema: Mappings

What's this?

Why migrate?

Where to?

How to?

-- Workflow

-- **Step by Step**

- DBMS

- **Schema**

- Data

- SQL

- Infrastructure

-- Obstacles

-- Tools

Examples: Data types, DDL syntax

- date range:
 - Informix: days since Dec 31, 1899
 - PostgreSQL: from 4713 BC to 32767 AD
- Exceeding maximum char length
 - ignored by Informix
 - PostgreSQL gives error
- **CREATE TABLE**
 - Informix: “username”.tablename
 - PostgreSQL: tablename

Schema: Implementation

What's this?

Why migrate?

Where to?

How to?

-- Workflow

-- **Step by Step**

- DBMS

- **Schema**

- Data

- SQL

- Infrastructure

-- Obstacles

-- Tools

- Generate DDL code based on the mappings
 - Reverse engineer: Abstract the current schema to a logical or conceptual model
 - Forward engineer: Create target schema based on abstract model
- or
- Search and replace code differences
- Load DDL code into target system

Data: Analysis

What's this?

Why migrate?

Where to?

How to?

-- Workflow

-- **Step by Step**

- DBMS
- Schema
- **Data**
- SQL
- Infrastructure

-- Obstacles

-- Tools

Checklist:

- ✓ How much data at all?
- ✓ How much data to extract?
- ✓ Dirty data (duplicates, unused fields, NULLs, free text)
- ✓ Formatting of data dump
- ✓ Localized formats (date, decimal)

Data: Mappings

What's this?

Why migrate?

Where to?

How to?

-- Workflow

-- **Step by Step**

- DBMS

- Schema

- **Data**

- SQL

- Infrastructure

-- Obstacles

-- Tools

Examples:

- Delimiter in data dump
 - Informix puts one at end of line which PostgreSQL does not
- Formatting of decimals
 - Informix able to use comma, PostgreSQL only period

Data: Implementation

What's this?

Why migrate?

Where to?

How to?

-- Workflow

-- **Step by Step**

- DBMS

- Schema

- **Data**

- SQL

- Infrastructure

-- Obstacles

-- Tools

- Extract – Transform – Load
 - Connect tool to both databases OR
 - Dump to CSV file, edit dump
- Cut over without taking the system down (for too long)
 - “Cold Turkey”
 - “Chicken Little”
 - “Butterfly”

SQL / SPL: Analysis

What's this?

Why migrate?

Where to?

How to?

-- Workflow

-- **Step by Step**

- DBMS

- Schema

- Data

- **SQL**

- Infrastructure

-- Obstacles

-- Tools

Checklist:

- ✓ Client connection
- ✓ Client code available?
- ✓ SQL statements isolated?
- ✓ Standard SQL?
- ✓ Stored Procedures? Triggers?
- ✓ Stored Proc. language?

SQL / SPL: Mappings

What's this?

Why migrate?

Where to?

How to?

-- Workflow

-- **Step by Step**

- DBMS
- Schema
- Data
- **SQL**
- Infrastructure

-- Obstacles

-- Tools

Examples: Where Informix differs from the ANSI SQL standard

- Non-ANSI OUTER JOIN, UPDATE
- Strings delimited by double quotes
- Aliasing without “AS”
- “MATCHES” for regular expressions

PostgreSQL: Standard compliant

SQL / SPL: Implementation

What's this?

Why migrate?

Where to?

How to?

-- Workflow

-- **Step by Step**

- DBMS
- Schema
- Data
- **SQL**
- Infrastructure

-- Obstacles

-- Tools

- Change ODBC/JDBC driver OR
 - Client source available: Migrate!
 - Provide abstraction (Standard SQL) OR
 - Convert directly to target SQL syntax
- Stored Procedures:
- Major rewrite OR
 - Use converter script OR
 - Wrapper (PL/SQL: Fyracle, EnterpriseDB)

Infrastructure: Analysis

What's this?

Why migrate?

Where to?

How to?

-- Workflow

-- **Step by Step**

- DBMS

- Schema

- Data

- SQL

- **Infrastructure**

-- Obstacles

-- Tools

Checklist

- ✓ User base and permissions
- ✓ Administrative routines
- ✓ Security policies
- ✓ Tools for admin, design, development
- ✓ Tuning strategies
- ✓ Backup and recovery

Infrastructure: Mappings

What's this?

Why migrate?

Where to?

How to?

-- Workflow

-- **Step by Step**

- DBMS

- Schema

- Data

- SQL

- **Infrastructure**

-- Obstacles

-- Tools

- Administration:

UPDATE STATISTICS ↔ ANALYSE

dbschema ↔ pg_dump --schema-only

dbexport ↔ pg_dump --data

UNLOAD/LOAD ↔ COPY TO/FROM

- Informix: column-level permissions

- Database users:

→ Informix: System users, PostgreSQL: DB

Infrastructure: Implement

What's this?

Why migrate?

Where to?

How to?

-- Workflow

-- **Step by Step**

- DBMS

- Schema

- Data

- SQL

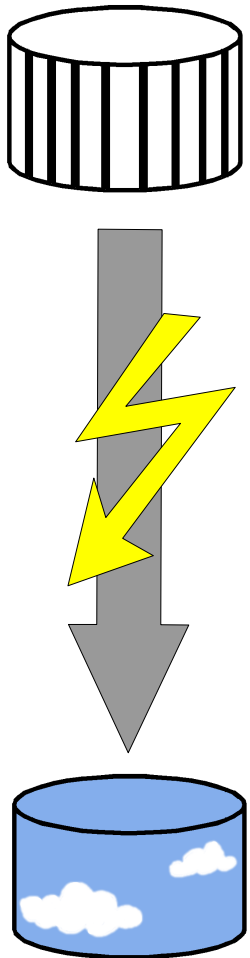
- **Infrastructure**

-- Obstacles

-- Tools

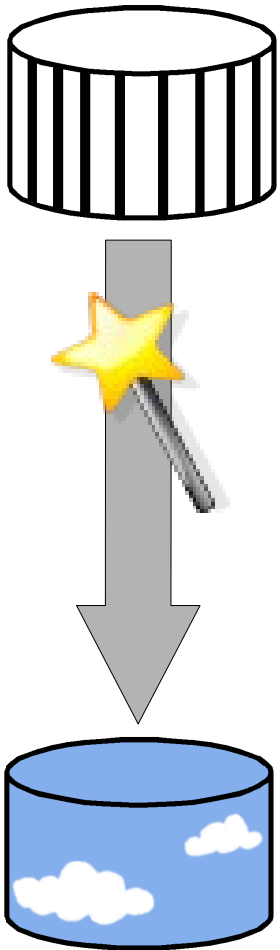
- Tools: Admin, Development, Design
 - Change ODBC/JDBC driver OR
 - Switch client
- Administrative Tasks
 - Gather policies and jobs
 - Implement them the target system way

Obstacles



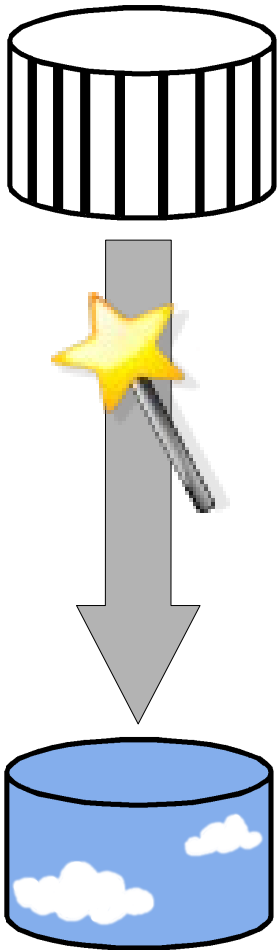
- Flawed schema design
- Missing documentation
- Non-Standard SQL implementations
- User-defined data types/functions
- Stored Procedures
- Differences in Data Types
- “Dirty Data”

Help! - Migration Tools



- Automatization
- Integrity violation detection
- Knowledge on source and target system
- Documentation
- Code Generation
- Script Scheduling
- Validation
- Speed!

Migration Tools, Examples



Closed Source:

- **SQLWays** (Ispirer). Only Windows.
Source: any, Target: MySQL, PostgreSQL
- **ProgressionDB** (Versora). Linux, Windows
MSSQL -> MySQL, PostgreSQL, Ingres

Open Source:

- Ingres Mio. \$ Challenge (CA): **shift2ingres**
- **MySQL Migration Toolkit** (MySQL AB)
- **InfToPg, OraToPg** (COSPA, Uni Bozen)

Summary

What's this?

Why migrate?

Where to migrate?

How to migrate?

- Several Open Source options
- Migration is an ambitious project
- Workflow and results depending on state of source system
- Automate the process, where possible!

Questions?



Contact:

Web: <http://www.osdbmigration.org>

Mail: jh@weltraumsofa.de