# Converting 85% of Dutch Primary Schools from Oracle to PostgreSQL

Martijn Dashorst
topicus.nl

# Martijn Dashorst

 @dashorst

 martijndashorst.com

 github.com/dashorst

# “I'd rather not work with databases”

Martijn Dashorst (1997, first job interview)

# Big thank you to Klaasjan Brand

# Please Give Feedback at https://2019.pgconf.eu/f

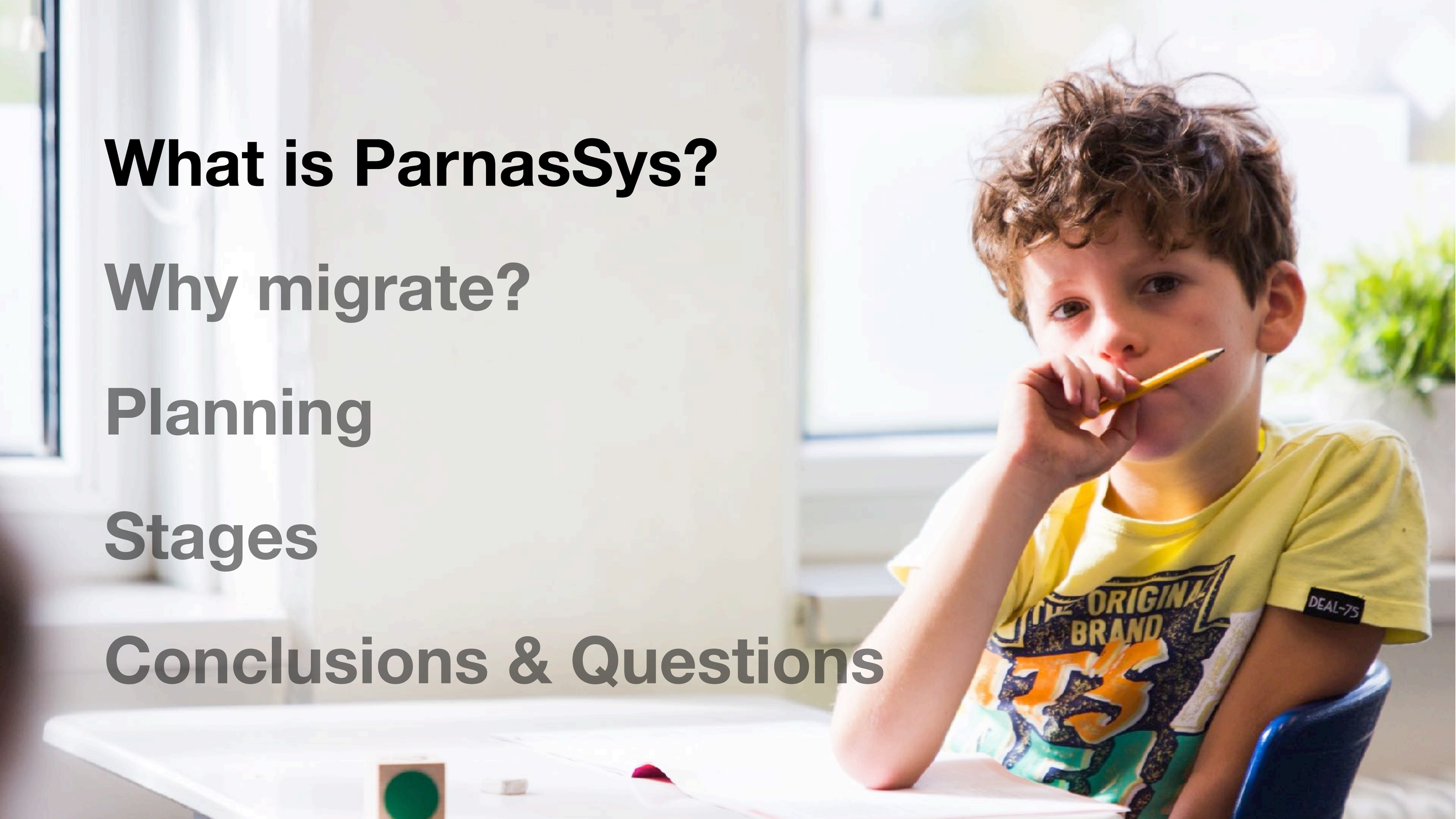
# What is ParnasSys?

## Why migrate?

## Planning

## Stages

## Conclusions & Questions



# Dutch Education System

| AGES | EDUCATION TYPE |
| --- | --- |
| 4-12 | **Primary Education** |
| 13-16/18 | **Secondary Education**<br>pre-vocational (4 yrs), pre-applied science (5 yrs), pre-research university (6 yrs) |
| 16-23 | **Vocational/University Education**<br>vocational (4 yrs), applied science (4 yrs, bachelor), pre-research university (bachelor/master) |

# Primary Education (4-12)

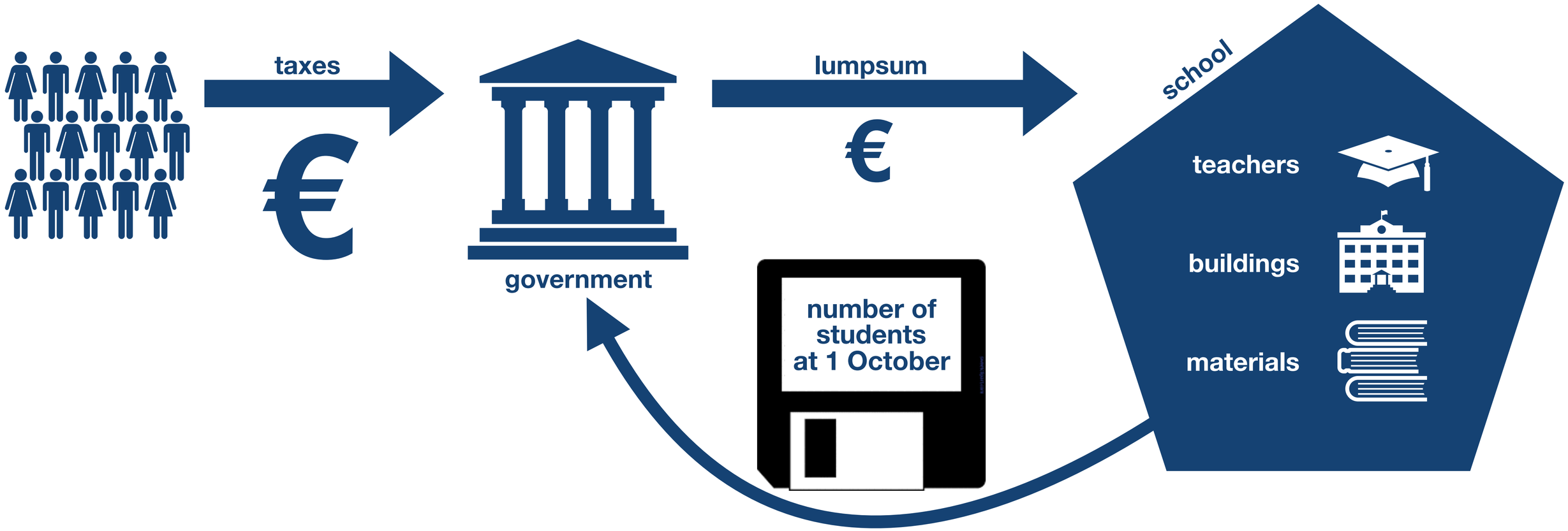
**1.5M**
**students**
ages 4-12, 2019-2020

**141k**
**employees**
2018

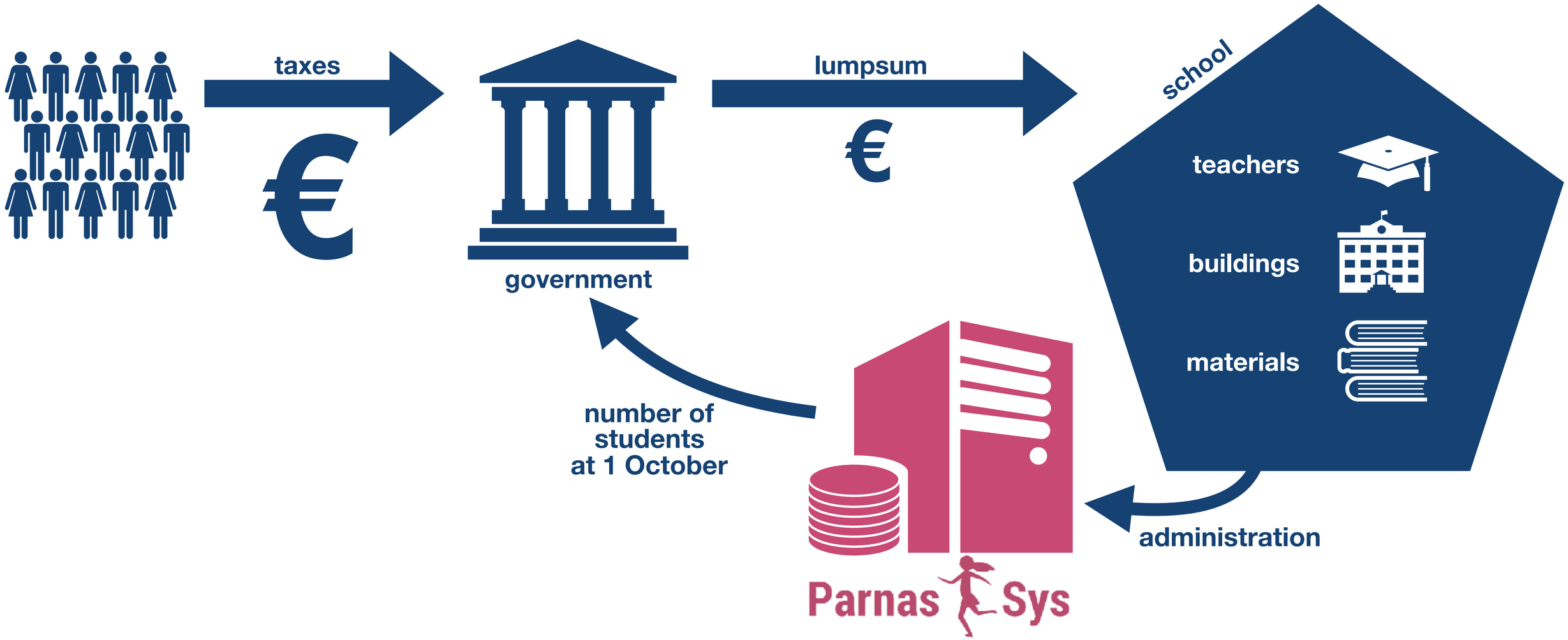
**6973**
**schools**
2018

Source: https://www.onderwijsincijfers.nl

# Financing Primary Schools



# Financing Primary Schools



**ParnasSys** ~ ***noun.***

1. A **SaaS** to run **primary schools** by keeping student records for school **financing**, student **counseling** and **guidance**, **attendance** keeping and **communicating** with **parents**

**ParnasSys** ~ ***noun.***

1. A **SaaS** to run **primary schools** by keeping student records for school **financing**, student **counseling** and **guidance**, **attendance** keeping and **communicating** with **parents**

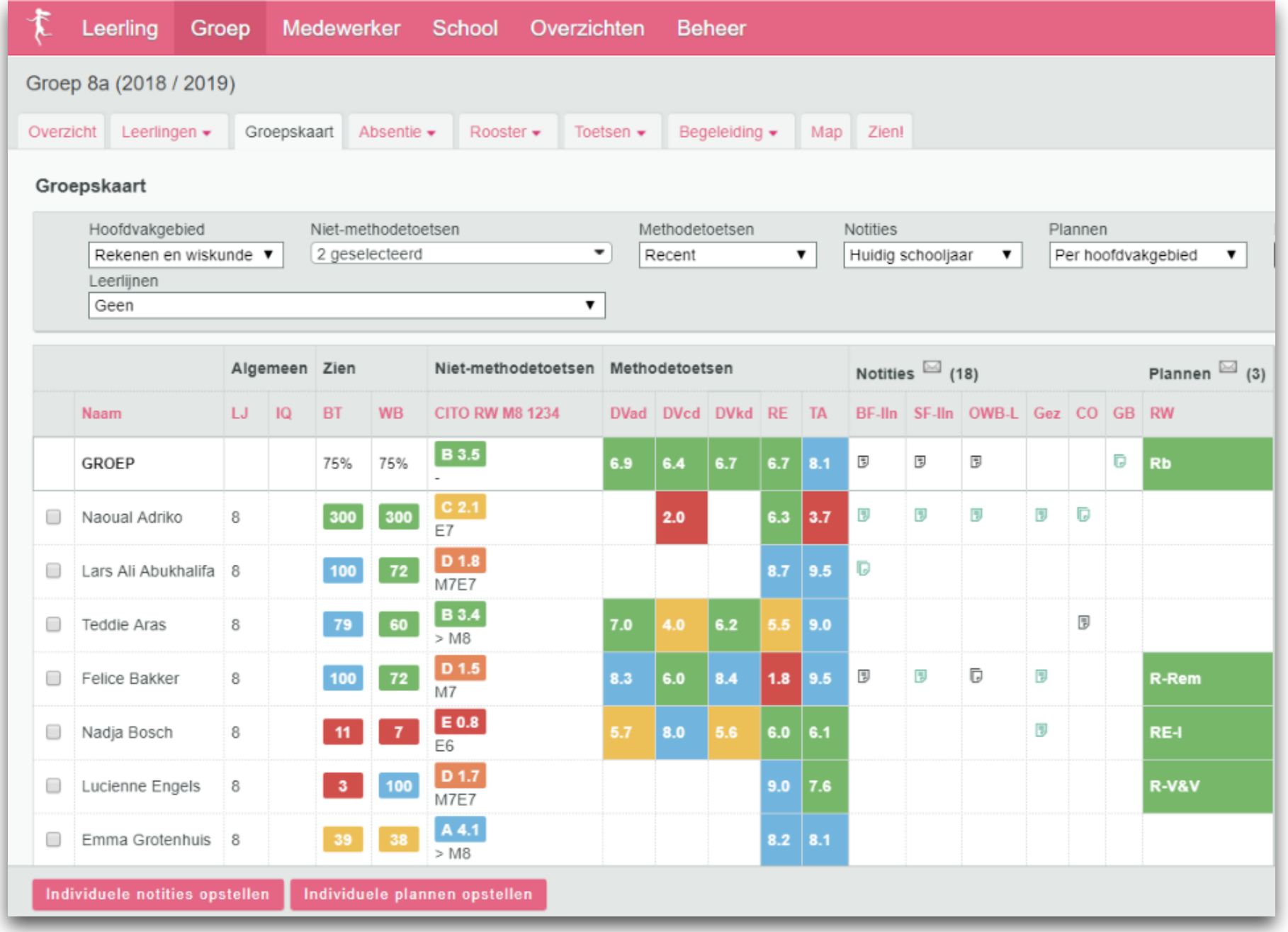
2. The only ~~pink~~**magenta** administration system in the world

# CITO Rekenen-Wiskunde

Midden 2018 - 2019 (1 november t/m 31 maart)

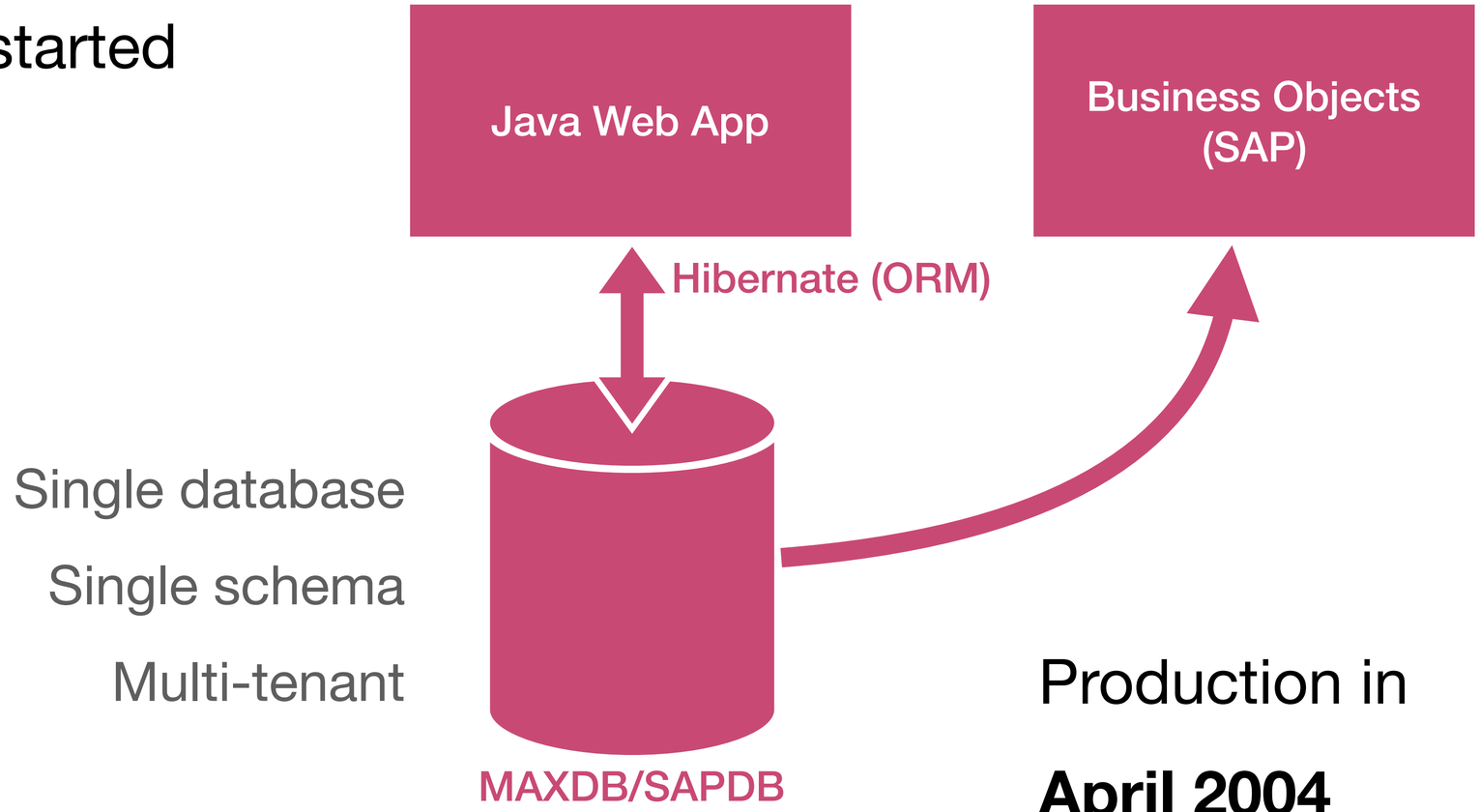
Weergaveopties

| | | | | Totaal | | | | |
|---|---|---|---|---|---|---|---|---|
| | Leerling | Leerjaar | Toets | VS | I-V | RN | FN | |
| | Simone Boskamp | 6 | 3.0 E5 | 195 | IV | < 1F | E4M5 | |
| | Kay Broek | 6 | 3.0 M6 | 199 | V | < 1F | M5 | ✓ |
| | Amara Diepeveen | 7 | 3.0 M7 | 259 | II | 1F | E7 | ✓ |
| | Julian Heesters | 6 | 3.0 M6 | 232 | III | < 1F | M6E6 | ✓ |
| | Maaike van Lieshout | 6 | 3.0 M6 | 211 | IV | < 1F | M5E5 | ✓ |
| | Lucas Tan | 6 | 3.0 M6 | 197 | V | < 1F | E4M5 | ✓ |
| | Gijs Vorstermans | 7 | 3.0 M7 | 257 | III | 1F | M7E7 | ✓ |


Leerling
Groep
Medewerker
School
Overzichten
Beheer
Groep 8a (2018 / 2019)
Overzicht
Leerlingen
Groepskaart
Absentie
Rooster
Toetsen
Begeleiding
Map
Zien!
Groepskaart
Hoofdvakgebied
Rekenen en wiskunde
Niet-methodetoetsen
2 geselecteerd
Methodetoetsen
Recent
Notities
Huidig schooljaar
Plannen
Per hoofdvakgebied
Leerlijnen
Geen
Algemeen
Zien
Niet-methodetoetsen
Methodetoetsen
Notities (18)
Plannen (3)
Naam
LJ
IQ
BT
WB
CITO RW M8 1234
DVad
DVcd
DVkd
RE
TA
BF-lln
SF-lln
OWB-L
Gez
CO
GB
RW
GROEP
75%
75%
B 3.5
6.9
6.4
6.7
6.7
8.1
Rb
Naoual Adriko
8
300
300
C 2.1
E7
2.0
6.3
3.7
Lars Ali Abukhalifa
8
100
72
D 1.8
M7E7
8.7
9.5
Teddie Aras
8
79
60
B 3.4
> M8
7.0
4.0
6.2
5.5
9.0
Felice Bakker
8
100
72
D 1.5
M7
8.3
6.0
8.4
1.8
9.5
R-Rem
Nadja Bosch
8
11
7
E 0.8
E6
5.7
8.0
5.6
6.0
6.1
RE-I
Lucienne Engels
8
3
100
D 1.7
M7E7
9.0
7.6
R-V&V
Emma Grotenhuis
8
39
38
A 4.1
> M8
8.2
8.1
Individuele notities opstellen
Individuele plannen opstellen

Development started

**October 2003**




Single database

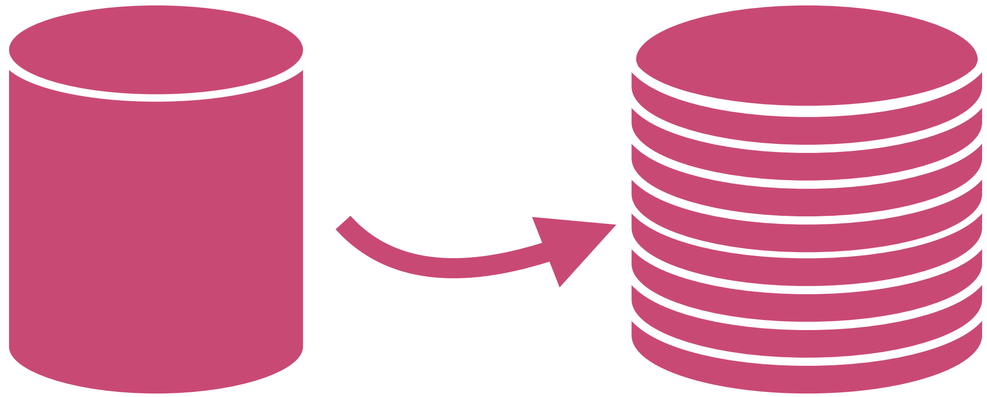
Single schema

Multi-tenant

Production in

**April 2004**

# First Database Migration in 2005




User base growing **fast**

**Unpredictable** query performance in **MaxDB**

**"Nobody got ever fired for buying Oracle"**

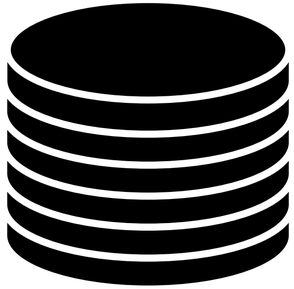
# 2019

85% of dutch
primary schools
use ParnasSys

~5M student dossiers
in our database

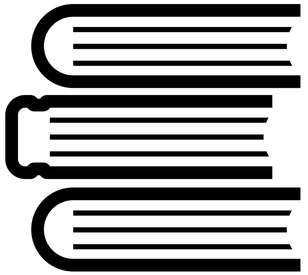
1,270,000 students in 2019-2020

800+ pages



360 tables

data & indices
~ 2TB



user files
~ 5TB

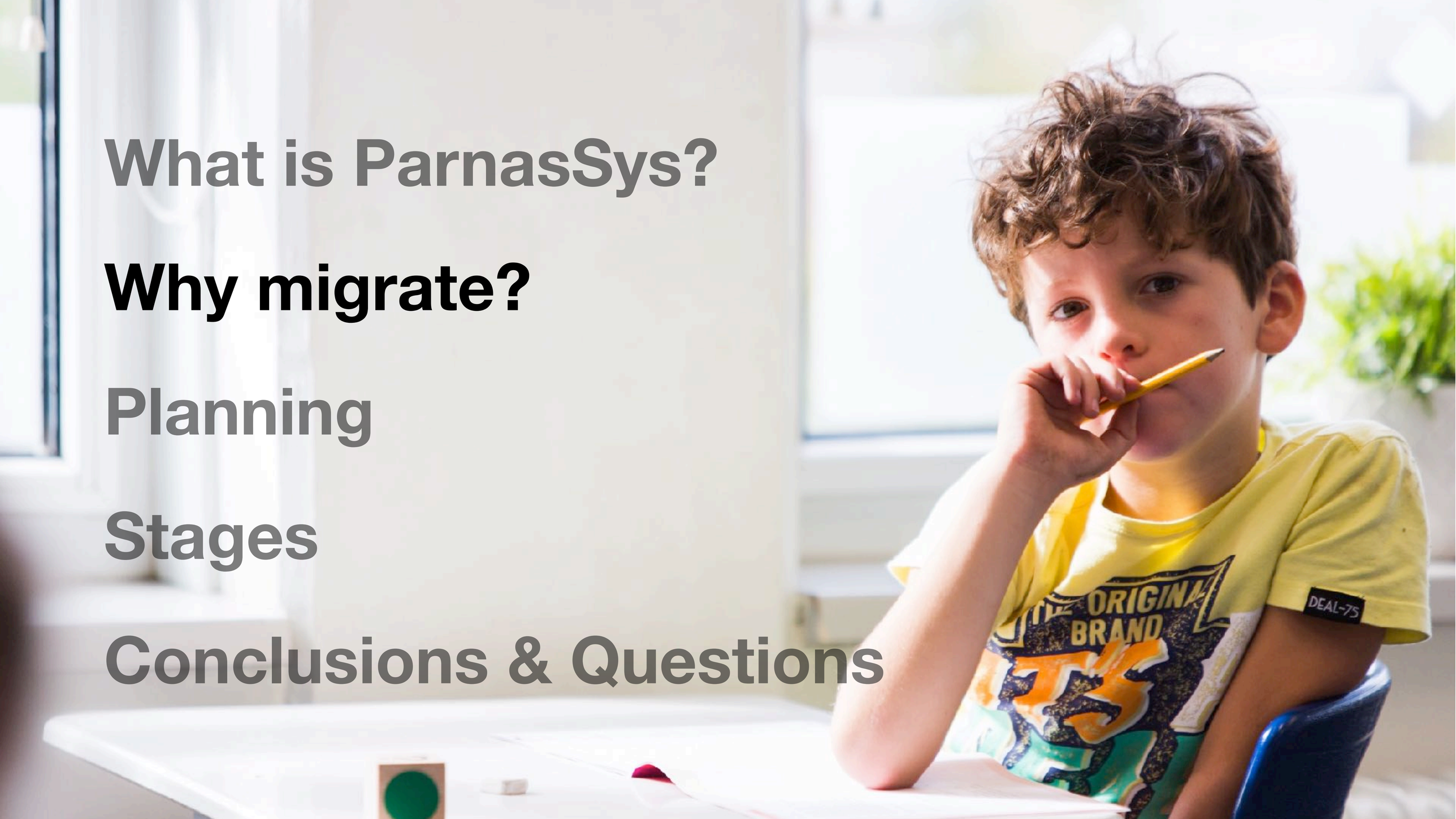
What is ParnasSys?

**Why migrate?**

Planning

Stages

Conclusions & Questions



# Mismatch Between Future of Oracle and our Product

# Cost Structure of Running Oracle Not Compatible with our Profitability

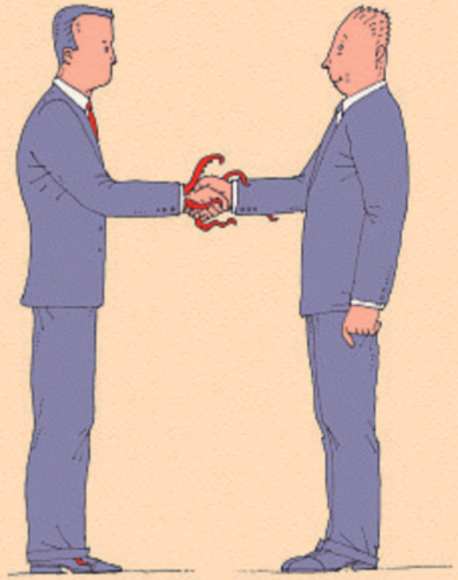
# The Oracle Migration Assistant™ Paid Us a Visit

ICT-DIENSTEN

# De wurggreep van de softwarereuzen

Zo'n 125 grote Nederlandse bedrijven en instellingen openen de aanval op beursgenoteerde IT-leveranciers. Vooral Oracle zou klanten opschepen met wurgcontracten, om vervolgens vele miljoenen per jaar aan nabetalingen op te eisen. 'Het is een verdienmodel.'

Vervolg op pagina 2



**The stranglehold of the software giants**

**[...] Especially Oracle would force unfair contracts upon clients, and claim many millions per year of late payments. [...]**

http://fd.nl/weekend/1318745/de-wurggreep-van-de-softwarereuzen

# Developer Happiness
## *Or How Do I Run My Database on My Local Machine*

# DevOps Happiness on Oracle is Incompatible with our Profitability

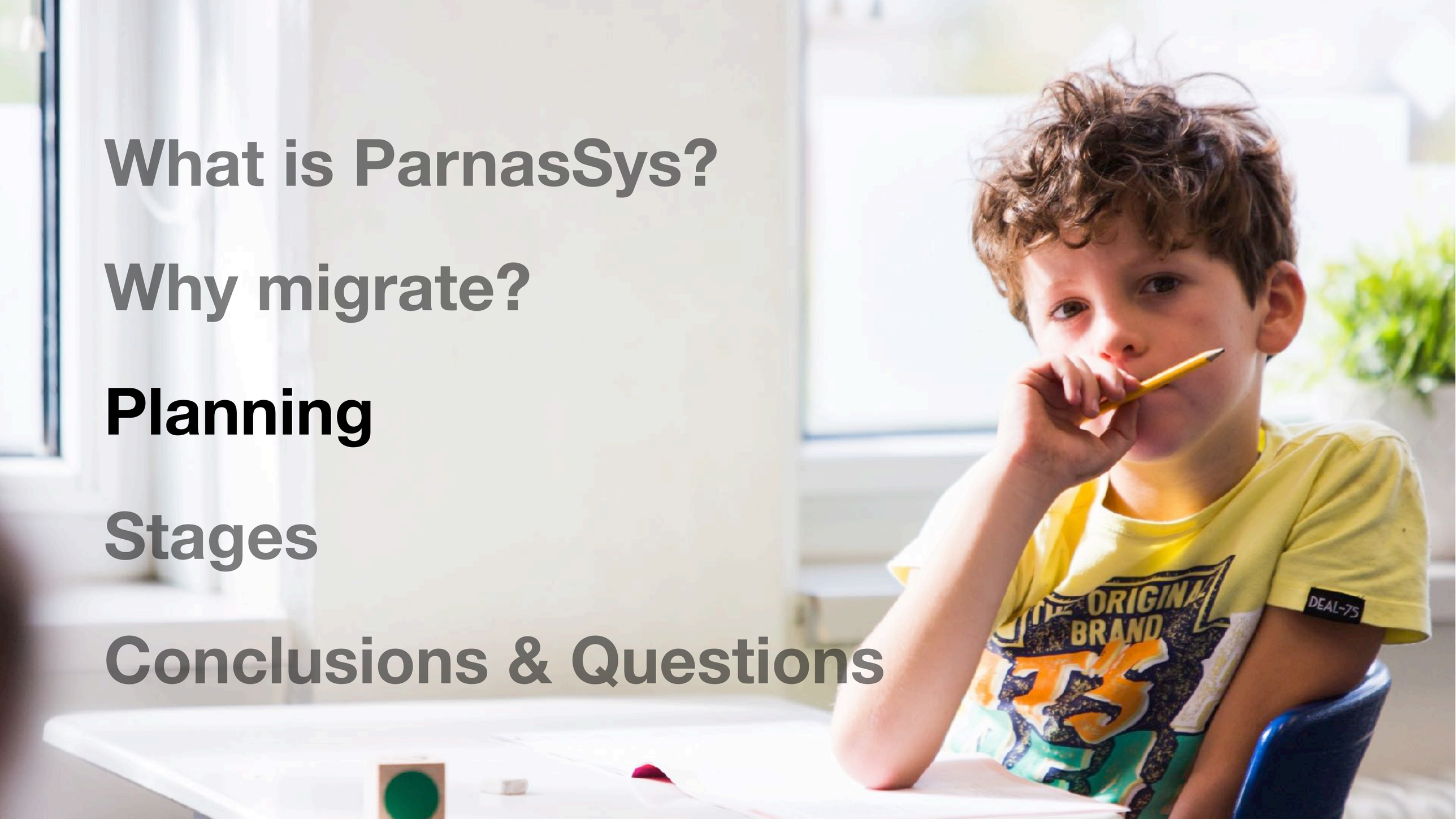
What is ParnasSys?

Why migrate?

**Planning**

Stages

Conclusions & Questions



# Inventory

- Database contents
- Java Application Code
- Business Objects Universe
- Business Objects Reports
- Backup & Restore
- Monitoring
- Training

- No stored procedures
- No custom functions
- No triggers

# Seasonality of Education

# 2 Moments in a Year to Migrate: X-Mas & Summer Vacation

# X-Mass 2 weeks

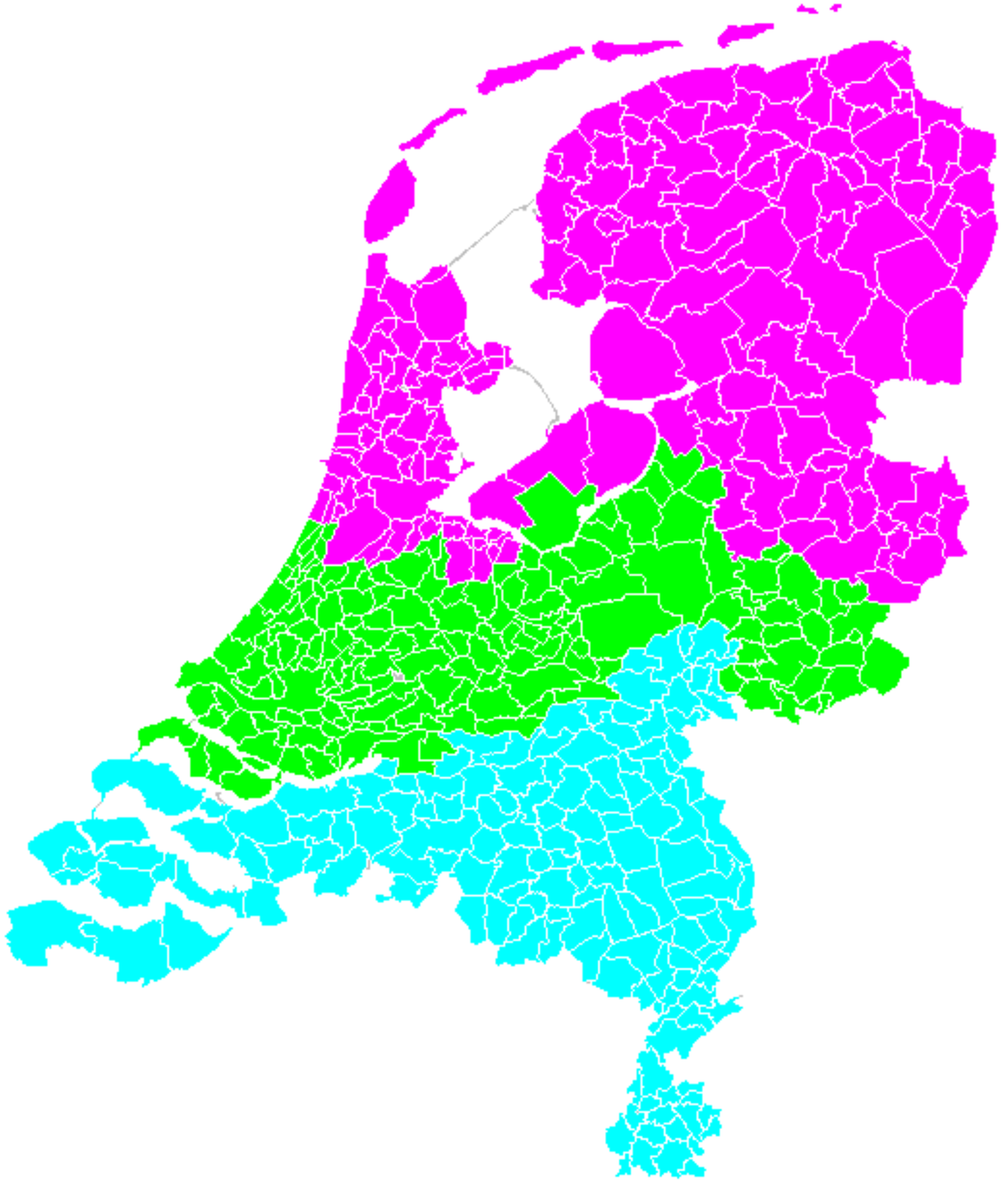
23 December 2019–5 January 2020

# Summer vacation(s)

**North**
15 juli 2019–23 august 2019

**Middle**
22 juli 2019-–30 august 2019

**South**
8 juli 2019–16 august 2019



# Ideal migration moment: start of summer vacation

**Stage 1: make it run**

incorrectly, correctly, fast

**Stage 2: migrate production**

in accordance to customer expectations

# 1 Year Migration

~~1 Year Migration~~

3 Year Migration

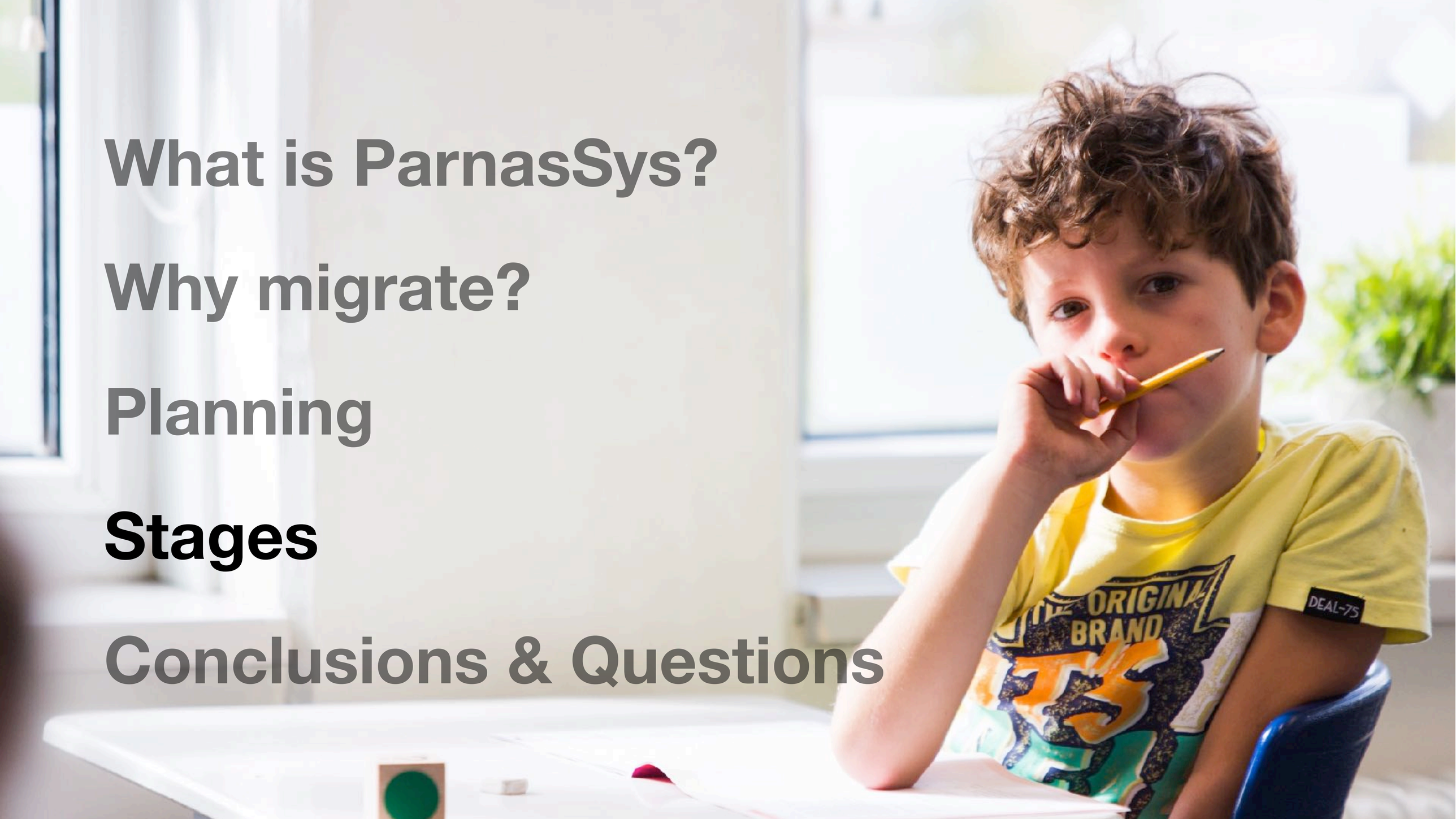
What is ParnasSys?

Why migrate?

Planning

**Stages**

Conclusions & Questions



# Stage 1: make it run


Free

Works™

Great to find issues in your application


## Perl

“The only language that looks the same before and after RSA encryption”

## Direct datatype mapping

E.G. number(1,0) does not map to boolean

## Slow

One import takes whole day

## Stable or fast (pick one)

Single-threaded: stable but slow
Multi-threaded: fast but unstable

## One-way, big bang only

# Differences between Oracle and PostgreSQL

“The good thing about standards is that there are so many to choose from.”

**— Andrew S. Tanenbaum**

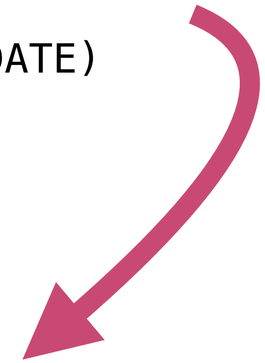
# Mapping datatypes

| Business | Java | Oracle | PostgreSQL |
|---|---|---|---|
| study year (1-8) | int | number(1,0) | number(1,0) |
| present (yes/no) | boolean | number(1,0) | boolean |
| sex (female/male) | enum Sex {<br>Female,<br>Male<br>} | number(1,0)<br>varchar2(6) | number(1,0)<br>varchar2(6) |

# ORM is 100% portable

```java
@Basic
@Temporal(TemporalType.DATE)
private Date birthdate;

@Formula("add_months(birthdate, (to_char(sysdate, 'yyyy') - to_char(birthdate, 'yyyy')) * 12)")
@Basic
@Temporal(TemporalType.DATE)
private Date birthday;
```



```sql
CREATE OR REPLACE FUNCTION verjaardag(geboortedatum IN DATE) RETURN DATE IS
BEGIN
  RETURN add_months(geboortedatum-1,
                (to_char(sysdate, 'yyyy') - to_char(geboortedatum, 'yyyy')) * 12)+1;
END;
```

# Oracle speaks in tongues

```
select 1 from dual

select nvl(null, 'n/a') from dual

select * from (select x,y from ...) where rownum between 100 and 110

sysdate

trunc
```

# Performance

tenant identifier

| ID | ORGANIZATION | NAME | BIRTHDATE | ... |
|---|---|---|---|---|
| 14342 | 1231 | JOHN SMITH | 2010-09-23 | |
| 14345 | 1231 | MARTIJN DASHORST | 2011-03-12 | |
| ... | | | | |

| ID | ORGANIZATION | NAME |
|---|---|---|
| 151231 | 1231 | Group 1A |
| 151232 | 1231 | Group 1B |
| ... | | |

# Oracle is smarter...

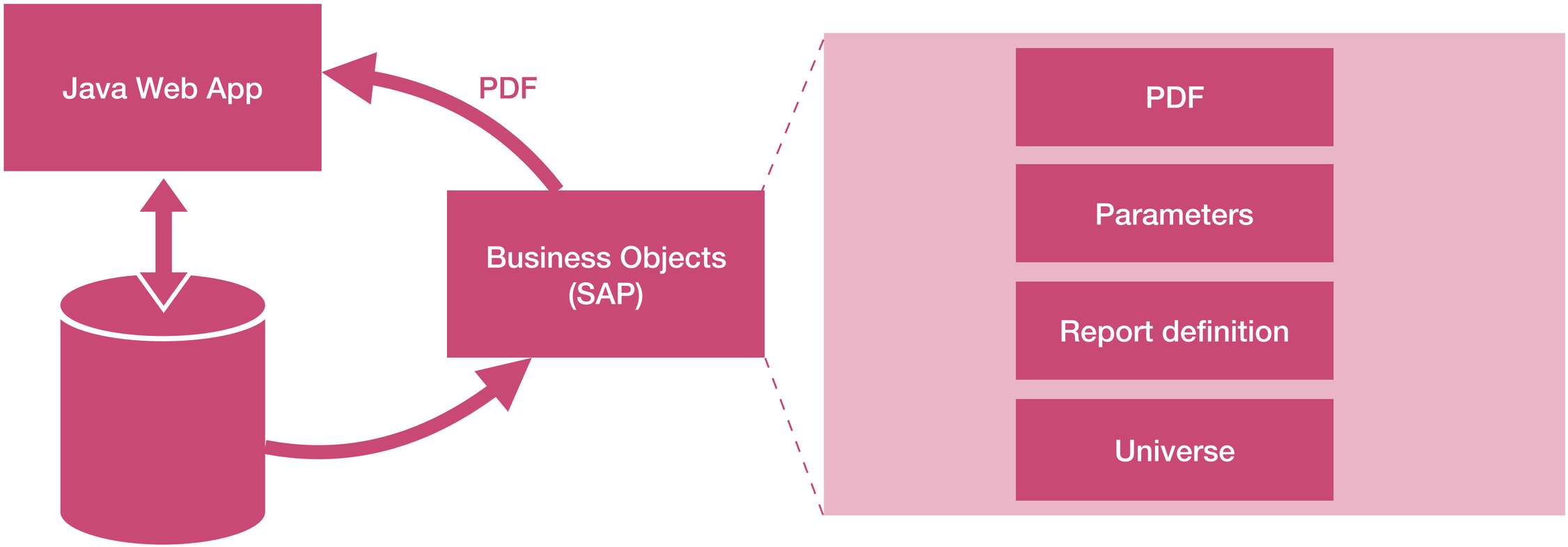
```
UPDATE
  testresult tr
SET
  tr.endresult = TRUE
WHERE
  tr.organisation = ?
  AND tr.testpart IN (
    SELECT part1.id
  FROM
    testpart part1
  WHERE
    part1.preference = (
      SELECT MIN(part2.preference)
    FROM
      testresult tr3
    INNER JOIN testpart part2 ON
      tr3.testpart = part2.id
    WHERE
      part2.preference > 0
      AND tr3.score IS NOT NULL
      AND tr3.test = tr.test
      AND ((part1.usemodulenorm IS NULL
      AND part2.usemodulenorm IS NULL)
      OR part1.usemodulenorm = part2.usemodulenorm)))
```

# Oracle is smarter...

```
UPDATE
  testresult tr
SET
  tr.endresult = TRUE
WHERE
  tr.organisation = ?
  AND tr.testpart IN (
    SELECT part1.id
  FROM
    testpart part1
  WHERE
    part1.preference = (
      SELECT MIN(part2.preference)
    FROM
      testresult tr3
    INNER JOIN testpart part2 ON
      tr3.testpart = part2.id
    WHERE
      part2.preference > 0
      AND tr3.score IS NOT NULL
      AND tr3.test = tr.test
      AND ((part1.usemodulenorm IS NULL
      AND part2.usemodulenorm IS NULL)
      OR part1.usemodulenorm = part2.usemodulenorm)))
```

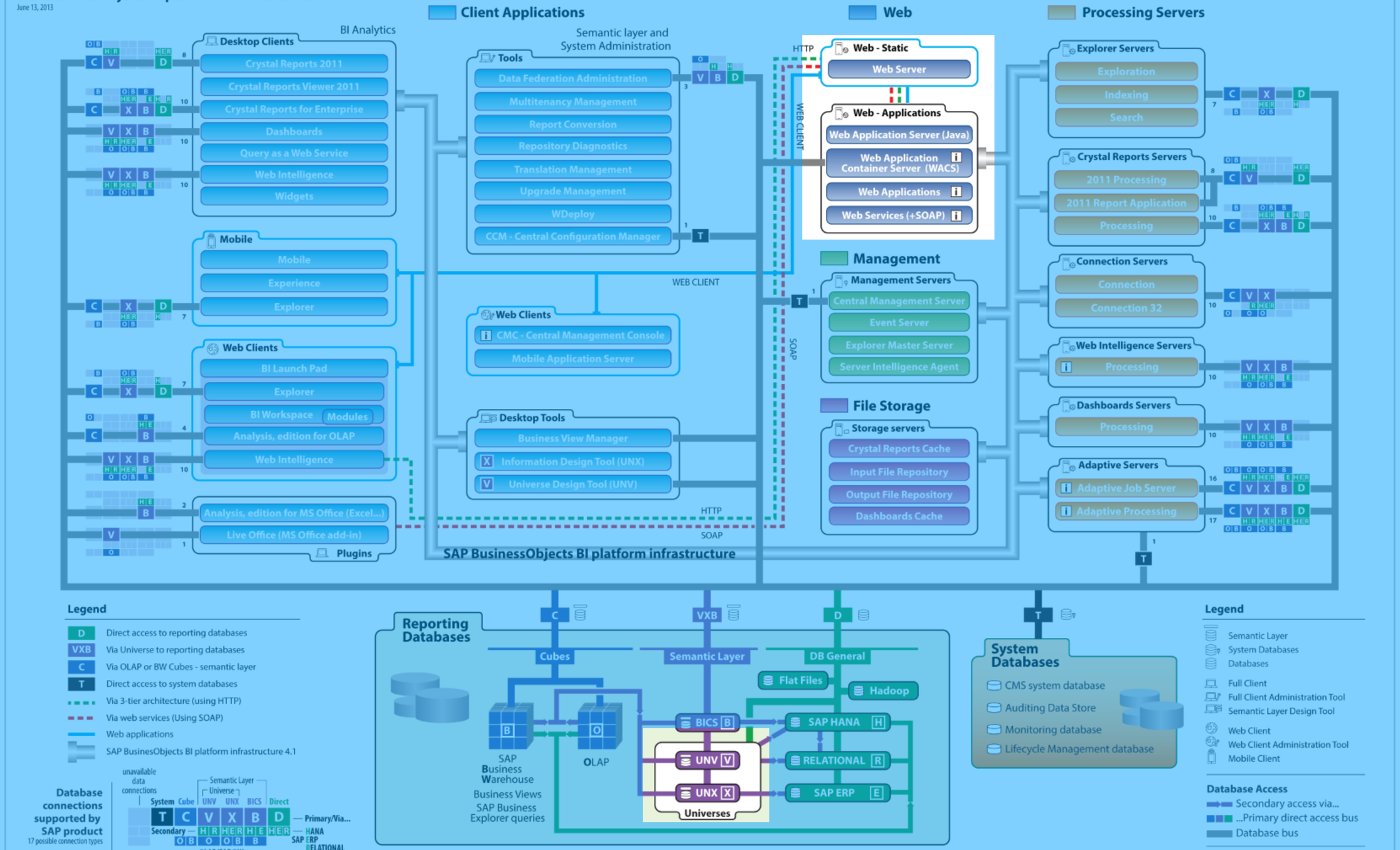
```
UPDATE
  testresult tr
SET
  tr.endresult = TRUE
WHERE
  tr.organisation = ?
  AND tr.testpart IN (
    SELECT part1.id
  FROM
    testpart part1
  WHERE
    part1.preference = (
      SELECT MIN(part2.preference)
    FROM
      testresult tr3
    INNER JOIN testpart part2 ON
      tr3.testpart = part2.id
    WHERE
      part2.preference > 0
      AND tr3.score IS NOT NULL
      AND tr3.test = tr.test
      AND ((part1.usemodulenorm IS NULL
      AND part2.usemodulenorm IS NULL)
      OR part1.usemodulenorm = part2.usemodulenorm)
      AND tr3.organisation = tr.organisation))
```

100x improvement


Java Web App
PDF
Business Objects
(SAP)
PDF
Parameters
Report definition
Universe

# SAP BusinessObjects BI platform 4.1 Architecture

June 13, 2013




## Legend

- D — Direct access to reporting databases
- VXB — Via Universe to reporting databases
- C — Via OLAP or BW Cubes - semantic layer
- T — Direct access to system databases
- Via 3-tier architecture (using HTTP)
- Via web services (Using SOAP)
- Web applications
- SAP BusinesObjects BI platform infrastructure 4.1

**Database connections supported by SAP product**

17 possible connection types










## Legend

- Semantic Layer
- System Databases
- Databases
- Full Client
- Full Client Administration Tool
- Semantic Layer Design Tool
- Web Client
- Web Client Administration Tool
- Mobile Client

**Database Access**

- Secondary access via...
- ...Primary direct access bus
- Database bus

Teachers Generate **a lot** of PDFs

but that's OK, because

Business Objects generates **a lot** of SQL

# Business Objects
# can be a bit nostalgic

# Quizz!

```
SELECT *
FROM A, B
WHERE B.column(+) = A.column
```

**1**

```
SELECT *
FROM A
RIGHT OUTER JOIN B
ON B.column = A.column
```

**2**

```
SELECT *
FROM B
LEFT OUTER JOIN A
ON B.column = A.column
```

**3**

```
SELECT *
FROM A
LEFT OUTER JOIN B
ON B.column = A.column
```

**4**

```
SELECT *
FROM B
RIGHT OUTER JOIN A
ON B.column = A.column
```

# Quizz!

```
SELECT *
FROM A, B
WHERE B.column(+) = A.column
```

**1**

```
SELECT *
FROM A
RIGHT OUTER JOIN B
ON B.column = A.column
```

**2**

```
SELECT *
FROM B
LEFT OUTER JOIN A
ON B.column = A.column
```

# Production ready?

- Application works
- Business Objects works

- Application works
- Business Objects works

**How long takes migration?**

**Behavior of application under actual load?**

- Heavy use of application side caching in Java process
- Modification of data
- Synthetic tests lie

**Can we fall back to Oracle?**

# Stage 2: Migrate Production

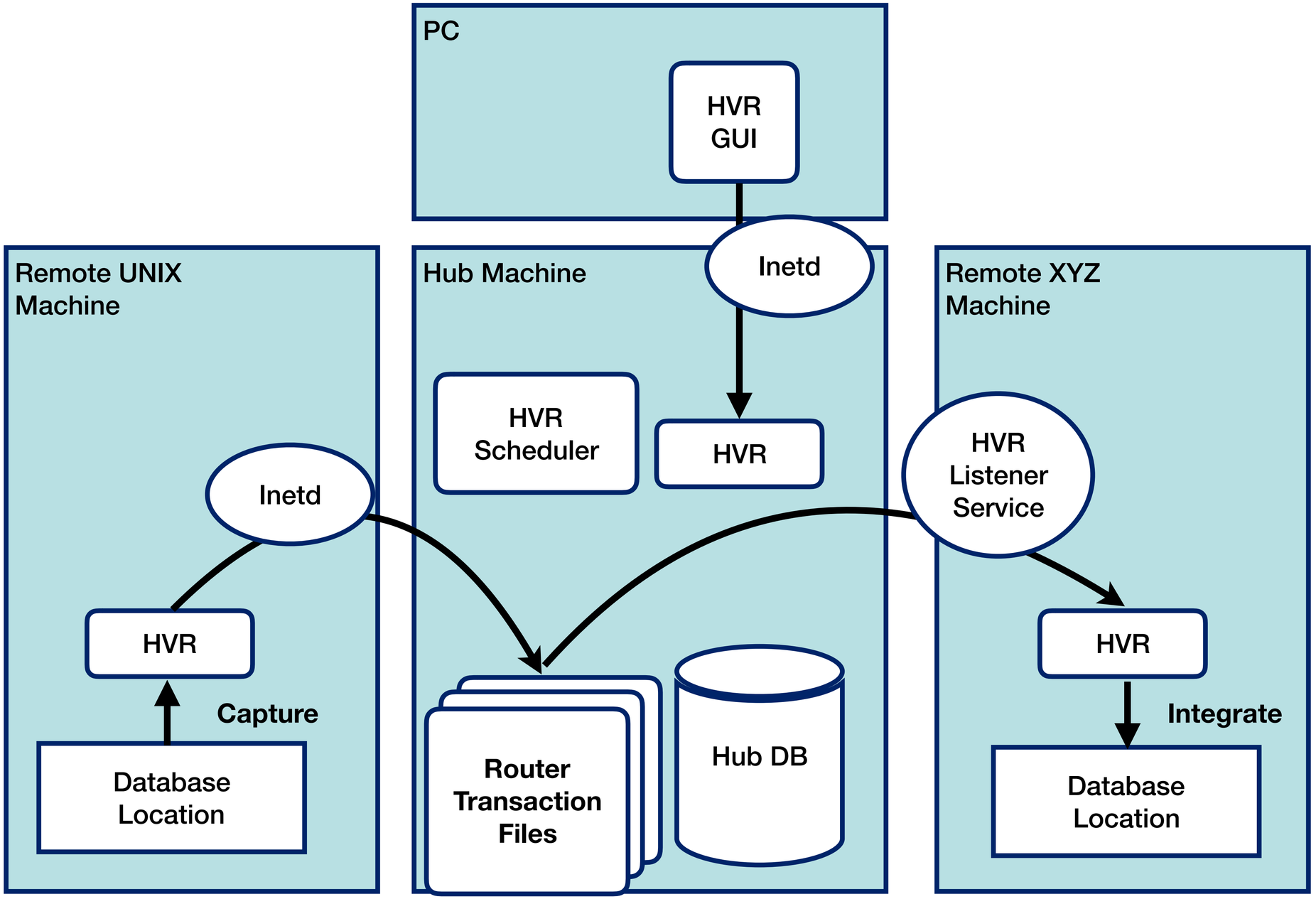
# Need help...

# 1-800-KILLORACLE

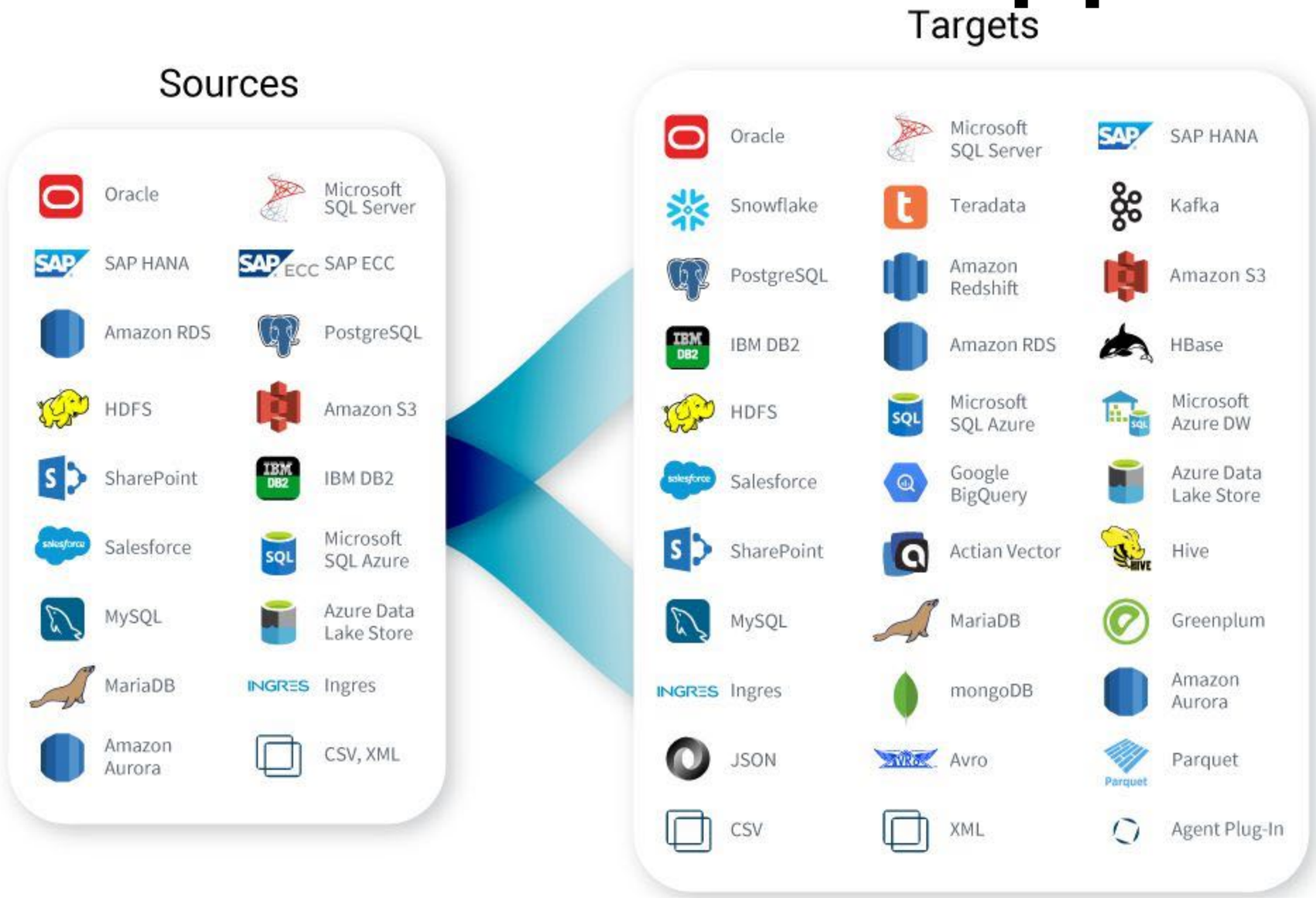
# Why don't you use HVR?

–Splendid Data

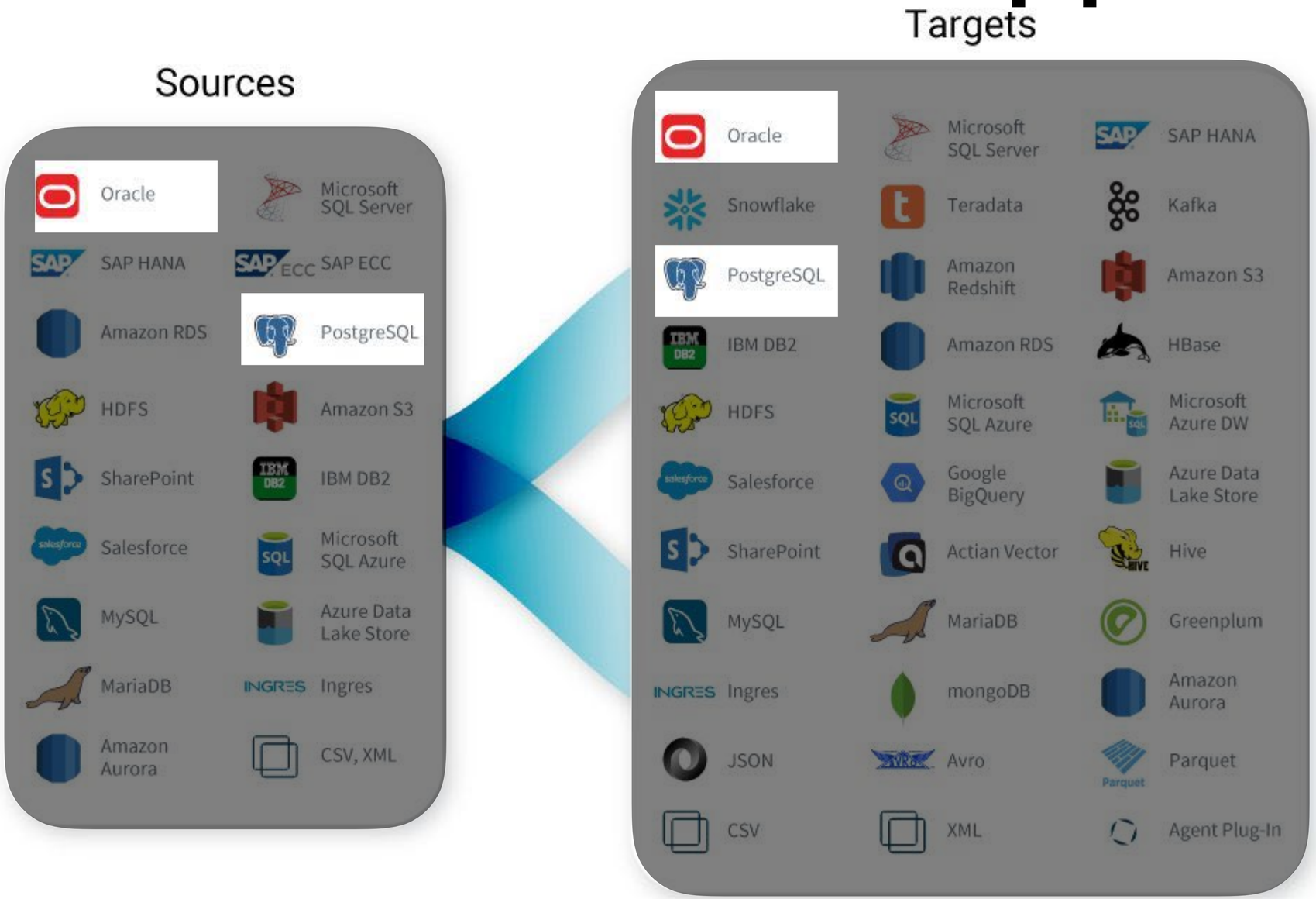

# Cross-database replication using redo logs


PC
HVR
GUI
Inetd
Remote UNIX
Machine
Hub Machine
Remote XYZ
Machine
HVR
Scheduler
HVR
HVR
Listener
Service
Inetd
HVR
HVR
Capture
Integrate
Database
Location
Router
Transaction
Files
Hub DB
Database
Location

# Lots of databases supported



# Lots of databases supported



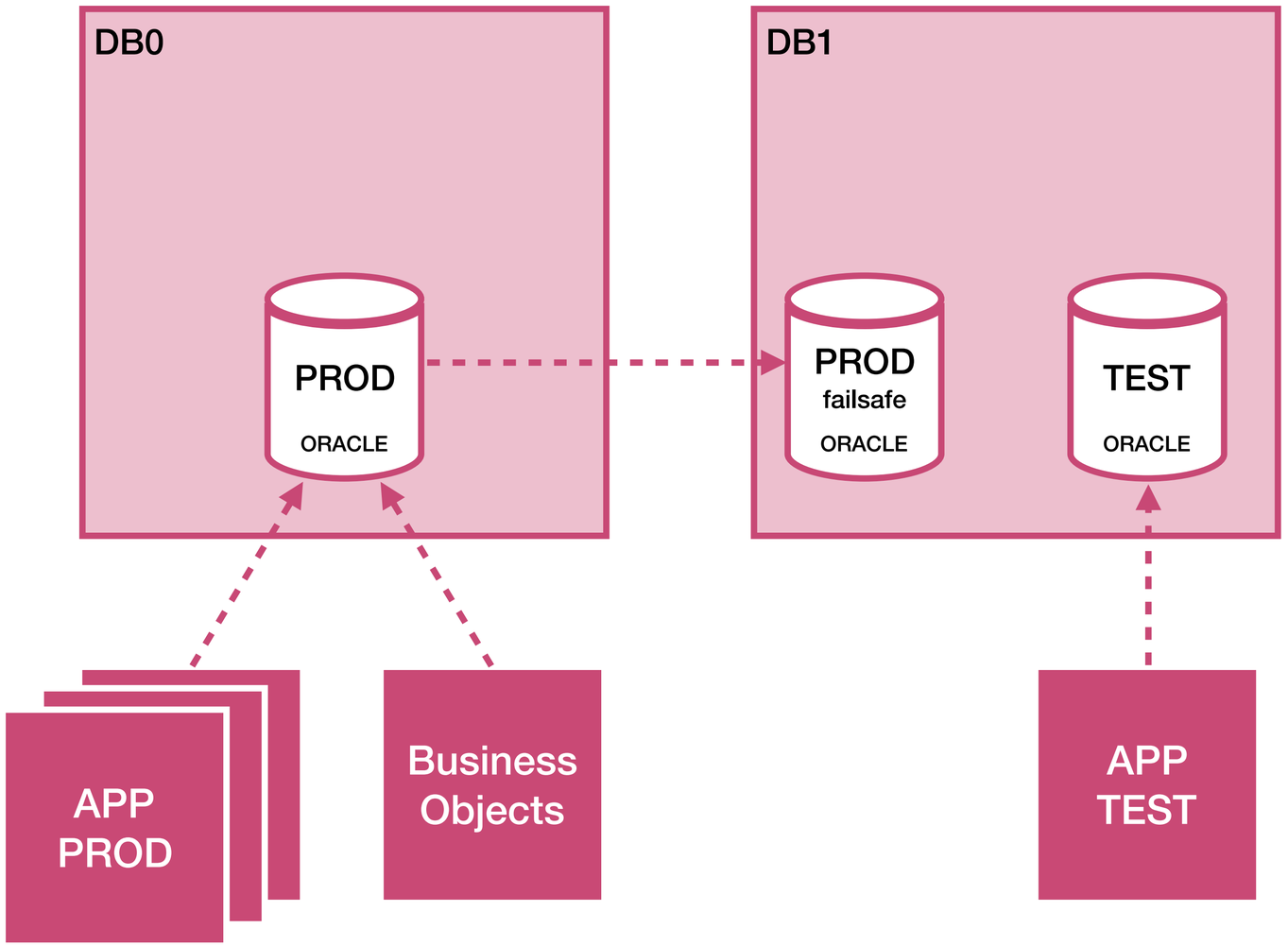
# Why HVR?

- Speed
- Real-time sync with low latency
- Options to test prior to actual migration
- Safe
- Fallback
- Expensive tool, but Special Price for Special Friend

# Noteworthy issues

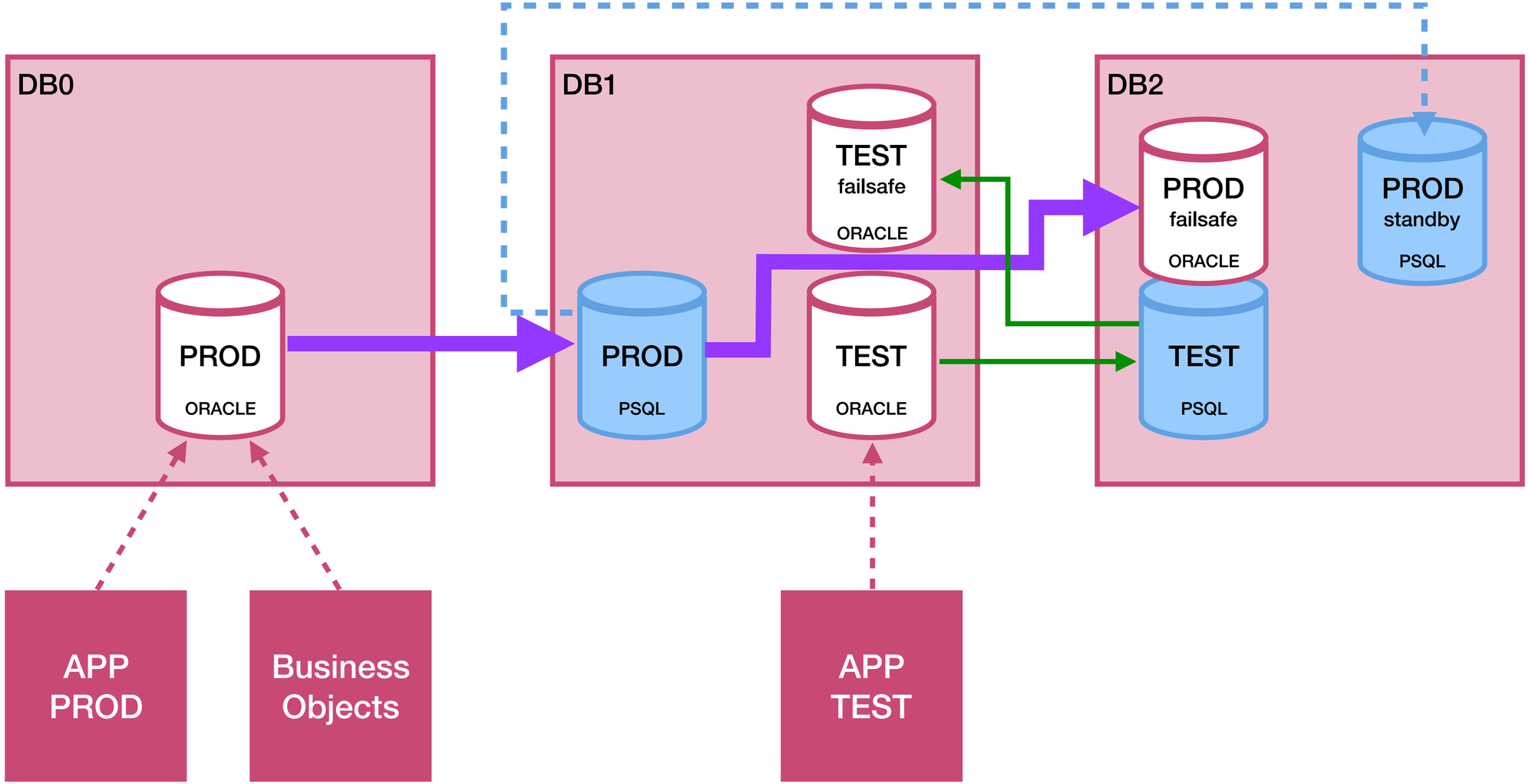
- Use primary key on each table
- Fix datatype incompatibilities because of differences between databases
  E.G. time in date field ✅ Oracle 🚫 PostgreSQL
- Disable constraint checking in target database
- Don't perform DDL changes
- Don't update too many rows in a table too often

# Oracle based deployment

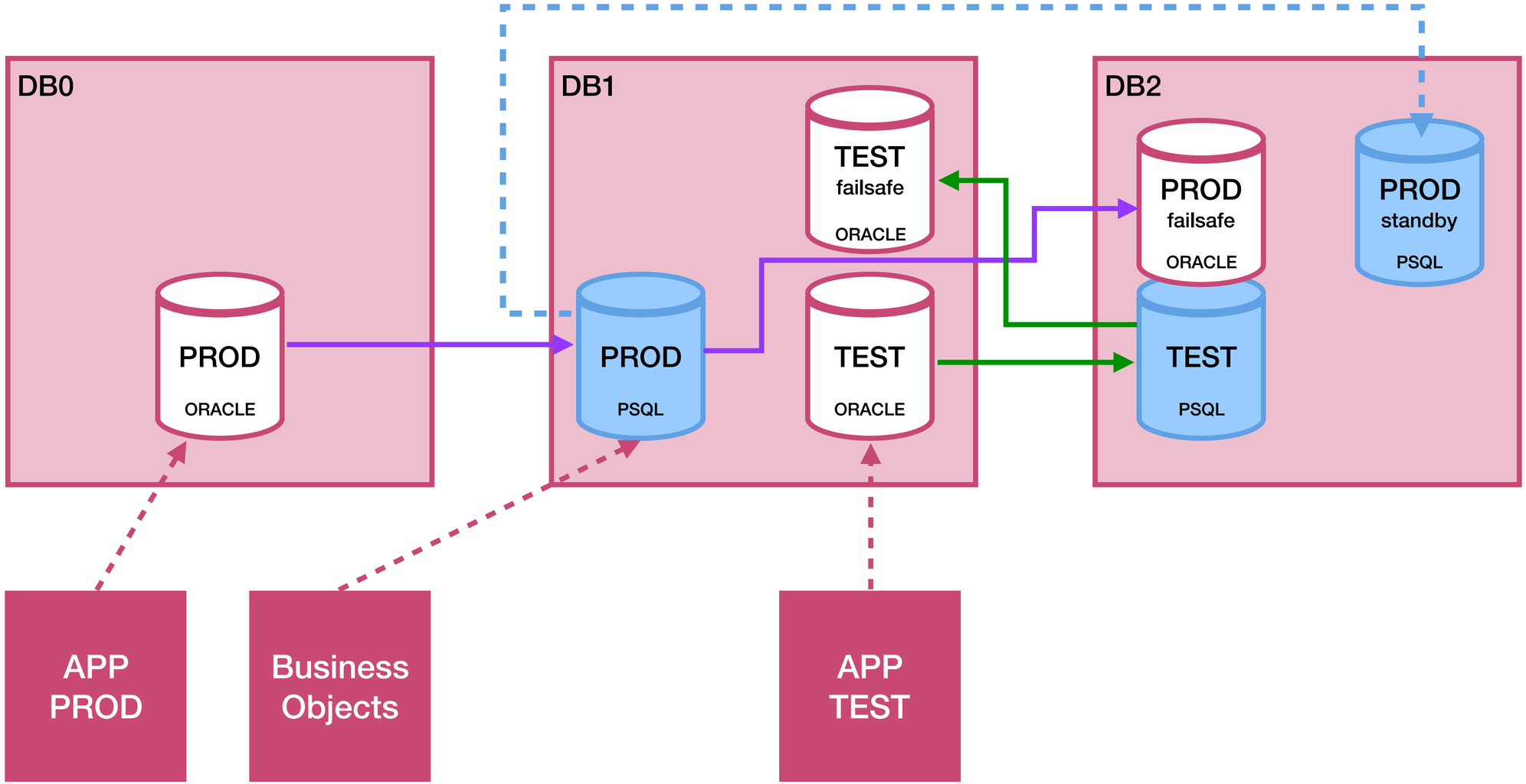



**Fail safe server**
Offline
Archive Log shipping

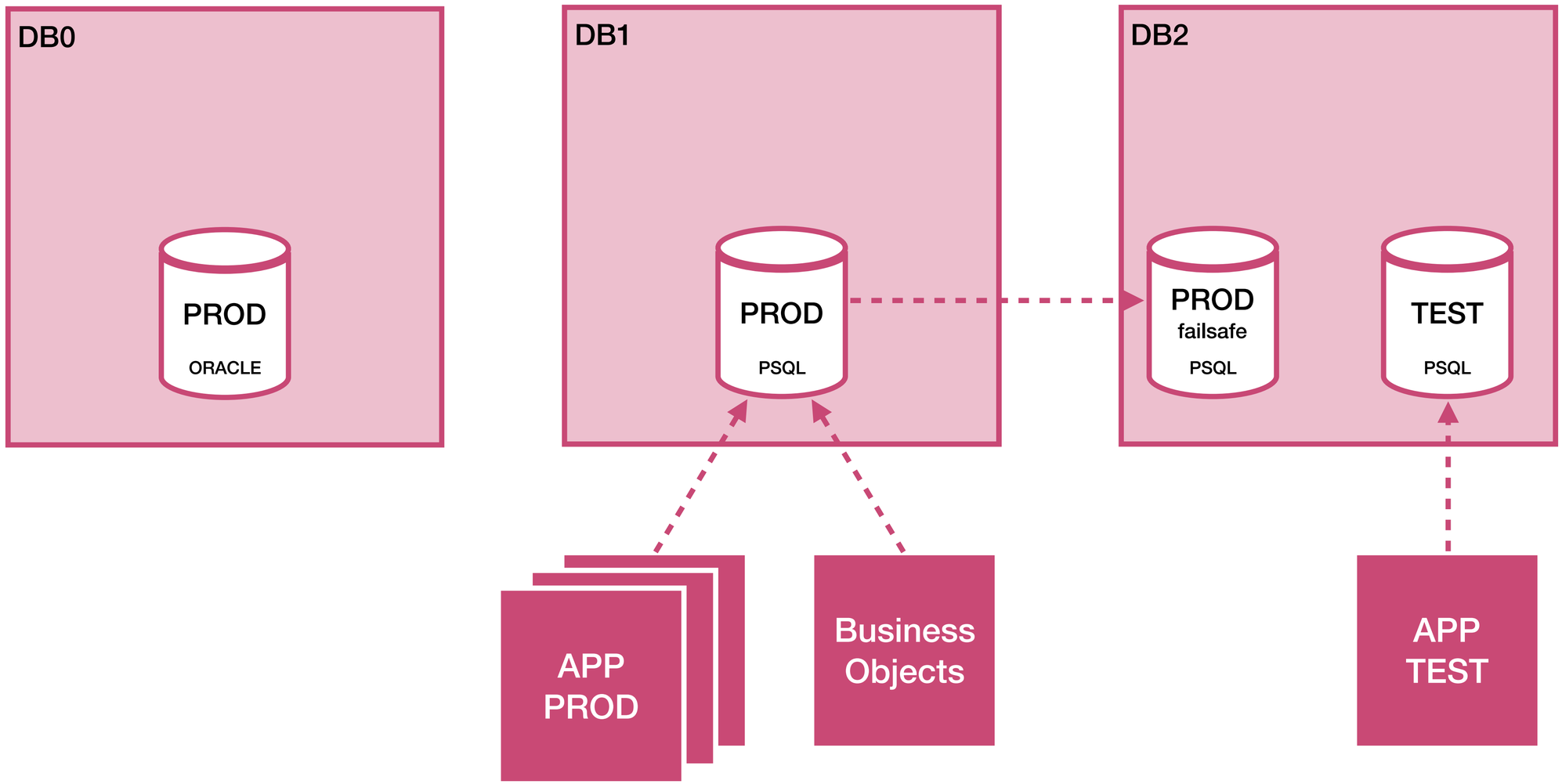
# HVR migration deployment



# Prod Testing Business Objects



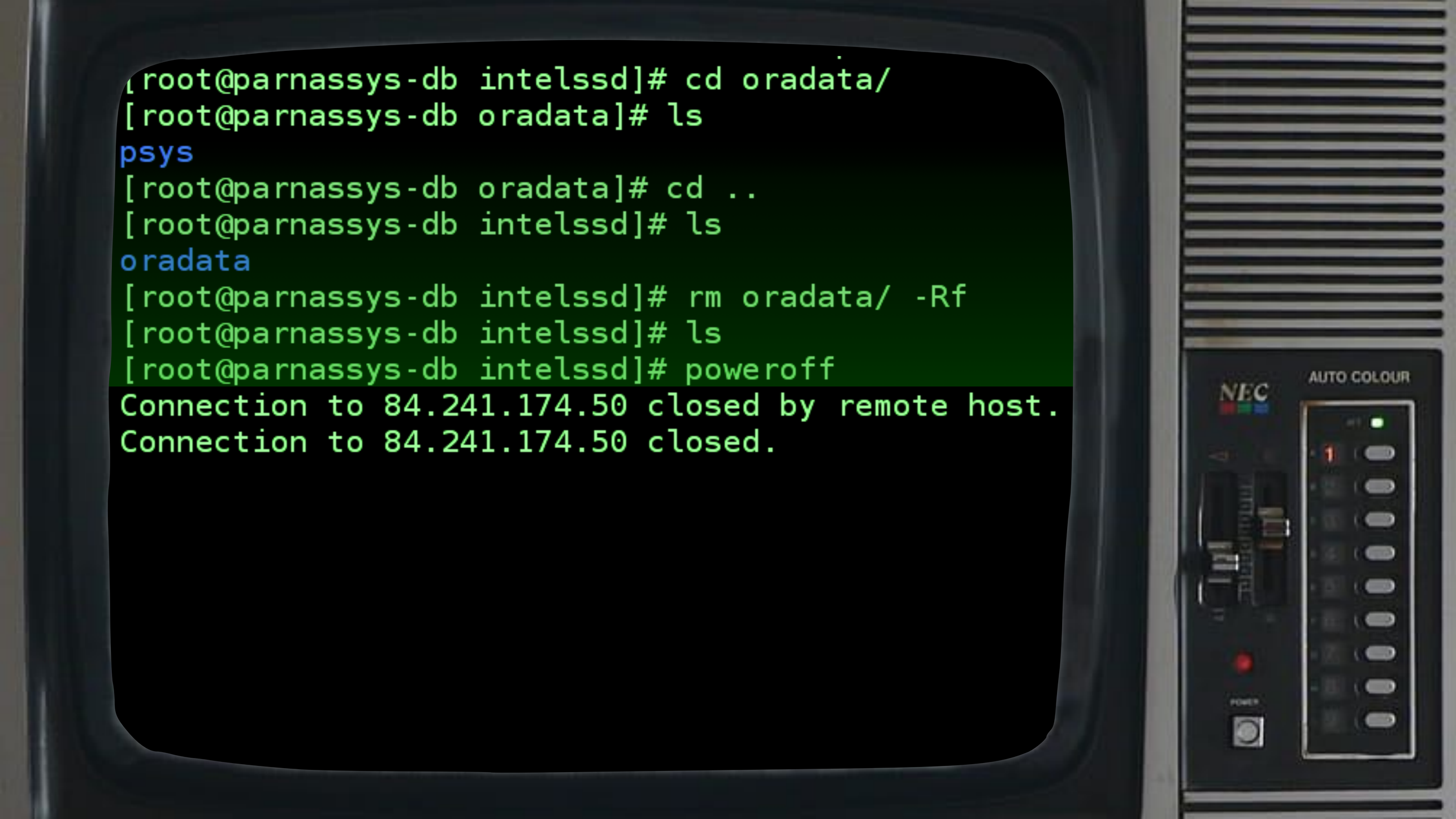
# PostgreSQL based deployment



# D-Day

de-oracle day

# 22 Juli 2019


[root@parnassys-db intelssd]# cd oradata/
[root@parnassys-db oradata]# ls
psys
[root@parnassys-db oradata]# cd ..
[root@parnassys-db intelssd]# ls
oradata
[root@parnassys-db intelssd]# rm oradata/ -Rf
[root@parnassys-db intelssd]# ls
[root@parnassys-db intelssd]# poweroff
Connection to 84.241.174.50 closed by remote host.
Connection to 84.241.174.50 closed.
NEC
AUTO COLOUR
POWER

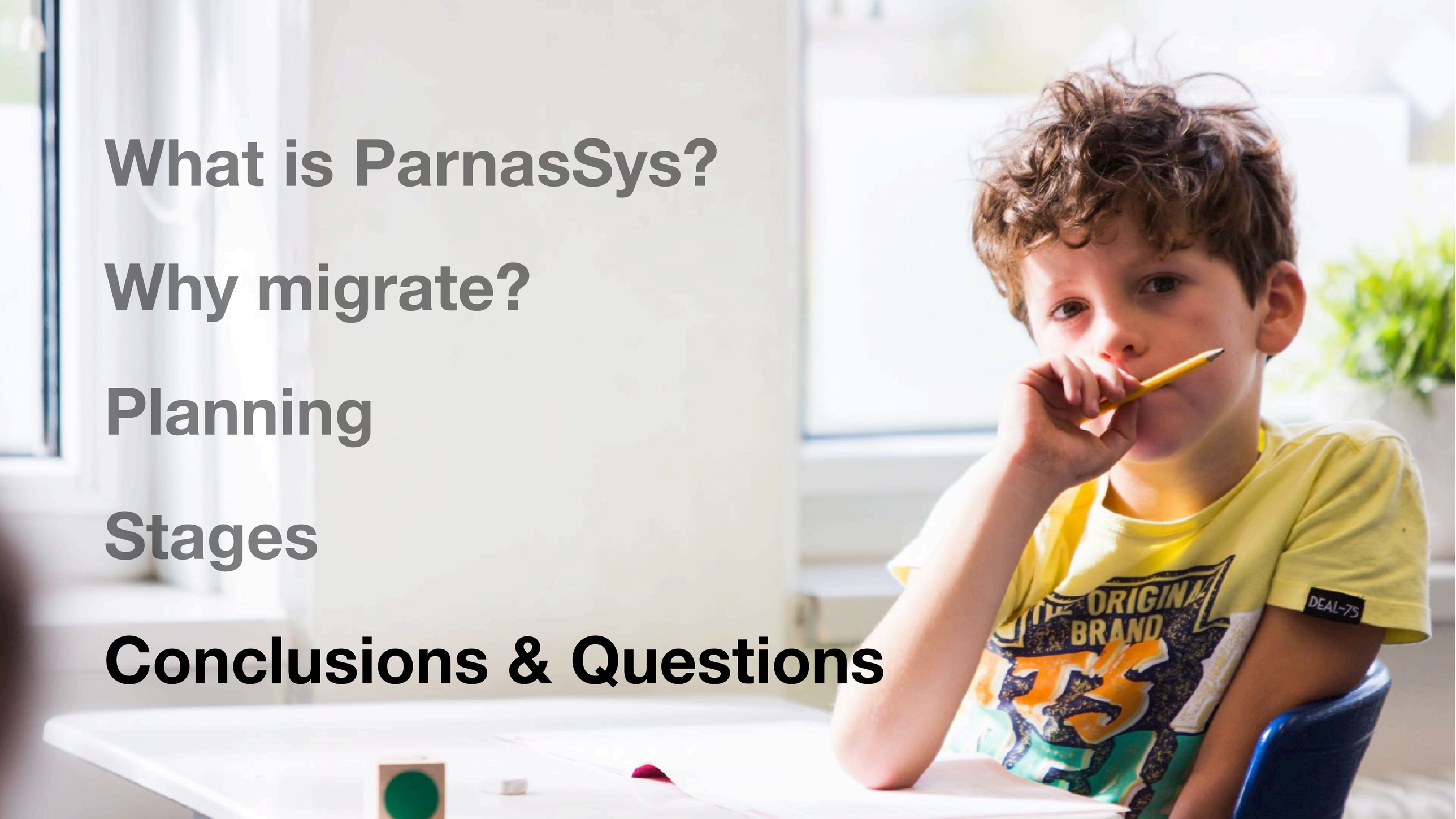
What is ParnasSys?

Why migrate?

Planning

Stages

**Conclusions & Questions**



Make a *really* good inventory of things that have to be migrated

Pick a strategy that works for you

Plan around your business to minimise risk and maximise opportunity

A great way to find bugs in your system

Don't be afraid to pay for tools and help


“Nobody got fired from migrating away from Oracle”
“Yet”

# **Migrate** our **Secondary Education** Application from Oracle to PostgreSQL

# Feedback
# https://2019.pgconf.eu/f

# Questions?

# THANK YOU!

 @dashorst

 martijndashorst.com

 github.com/dashorst