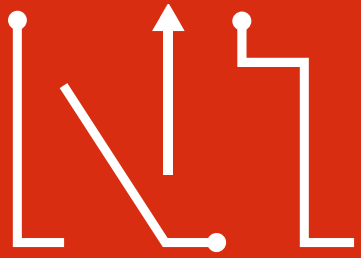




KONFERENCA

PORTOROŽ, 15. DO 17. MAJ 2017

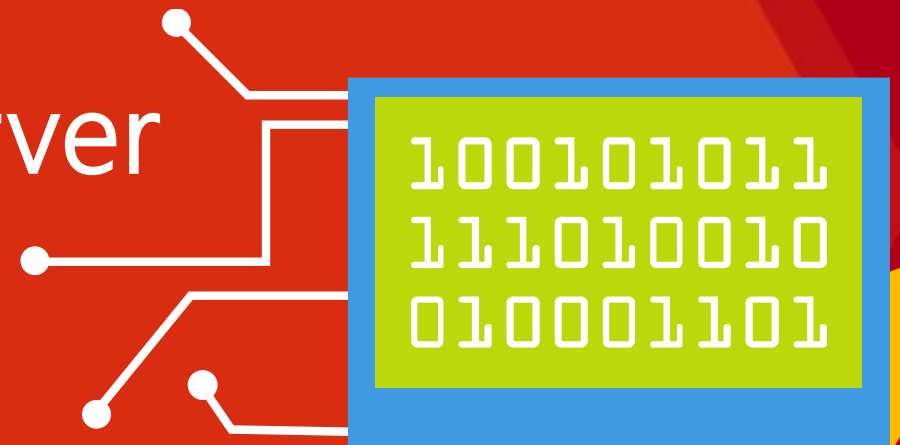




KONFERENCA

Deja Vu:

Virtual Databases for SQL Server



Urh Srečnik <urh.srecnik@abakus.si>

ORACLE®

Certified Professional

Oracle Database 12c
Administrator

ORACLE®

Certified Associate

Java SE 8 Programmer

TEHNOLOGIJA



Abakus Plus d.o.o.

- History

- From 1992
- ~20 employees

- Applications

- Special
 - Document Management System
 - Newspaper Distribution
 - Flight Information System
- Oracle Database:
 - ARBITER - the ultimate audit trail tool
 - APPM – Abakus Plus Performance and Monitoring Tool

- Services

- OS & Network Administration
- DBA, Programming

- Hardware

- Servers, SAN Storage, firewalls
- Backup Server

- Infrastructure

- > 20 years of experience with High Availability on GNU/Linux.

Reference



Aerodrom Ljubljana
Letališče Jožeta Pučnika Ljubljana



BANKA
SLOVENIJE



Mestna občina
Ljubljana



MILENIJUM
OSIGURANJE

NLB Vita
Življenjska zavarovalnica



MESTNA OBČINA KOPER
COMUNE CITTÀ DI CAPODISTRIA



ZAVOD ZA
ŠPORT RS
PLANICA



ISKRATEL



RAM 2



BERNARDIN GROUP
RESORTS & HOTELS



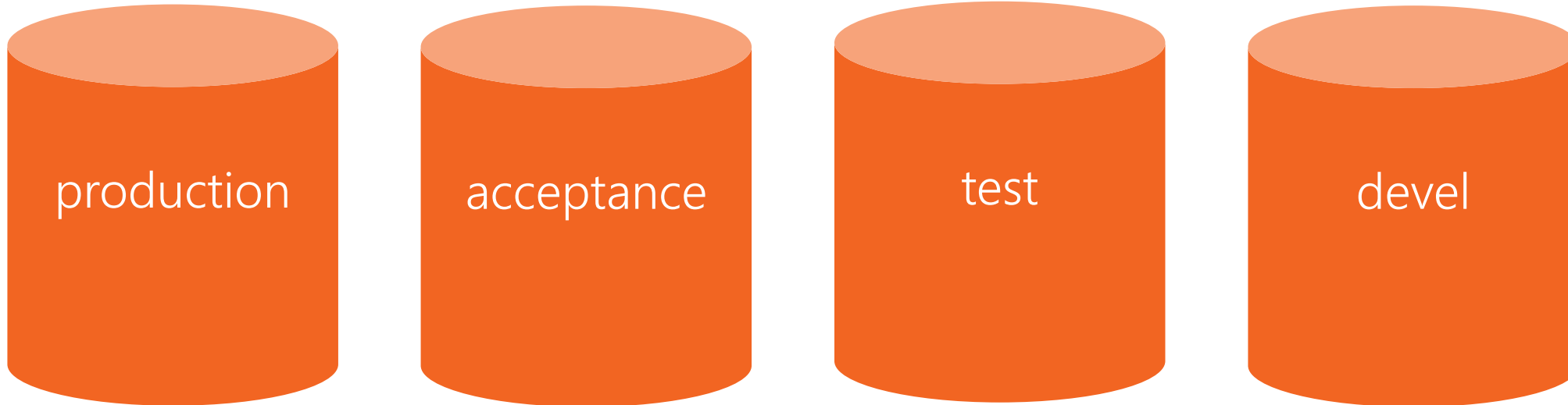
dejavu Data at your service.

Problem #1

Production Size and Growth

The Problem

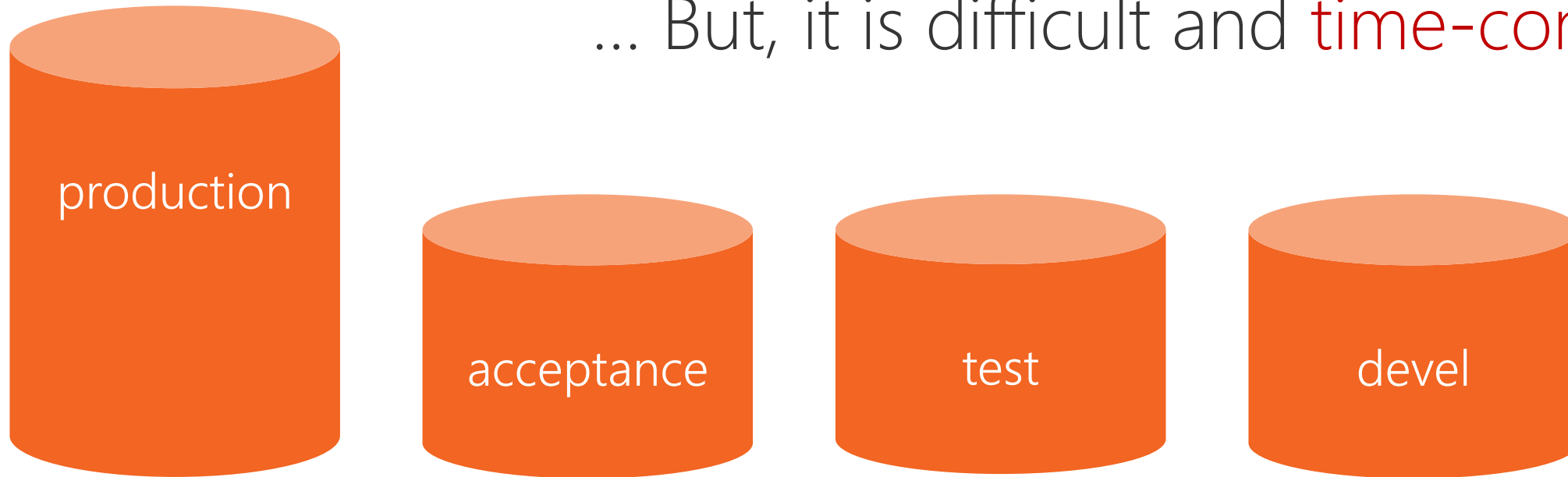
- Beyond every production environment...
 - There is test, development and other project-specific environments



The Problem

Trying to migrate the problem by copying subsets of production.

... But, it is difficult and **time-consuming**



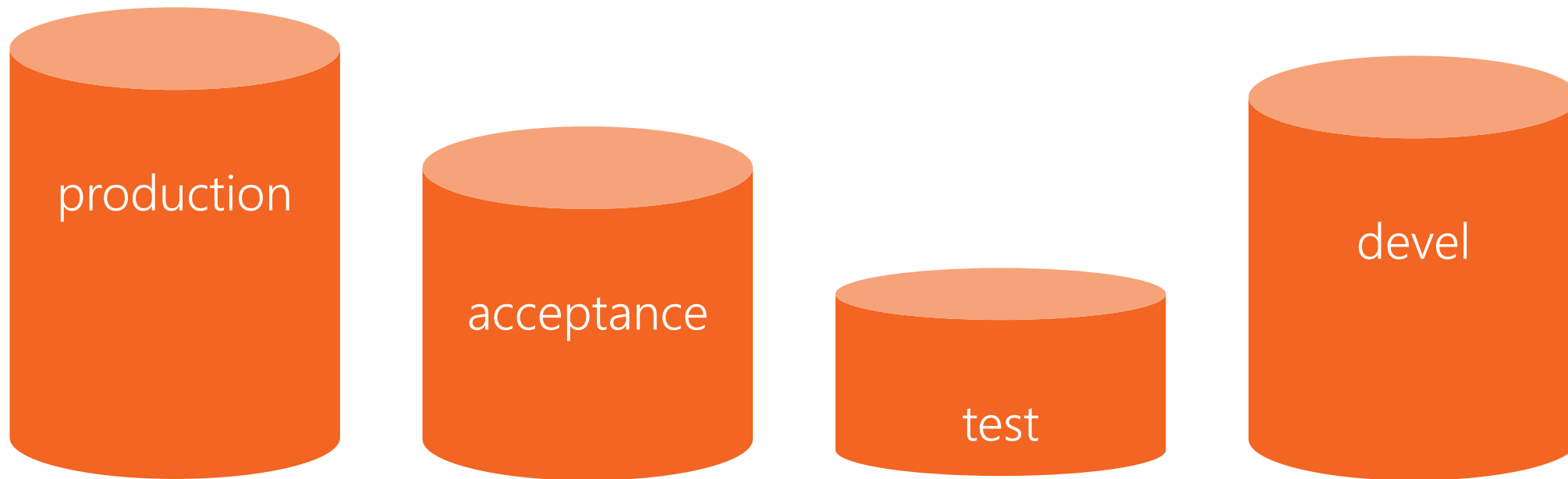
Time Consuming?

How long does it take to copy 5 TB of data?

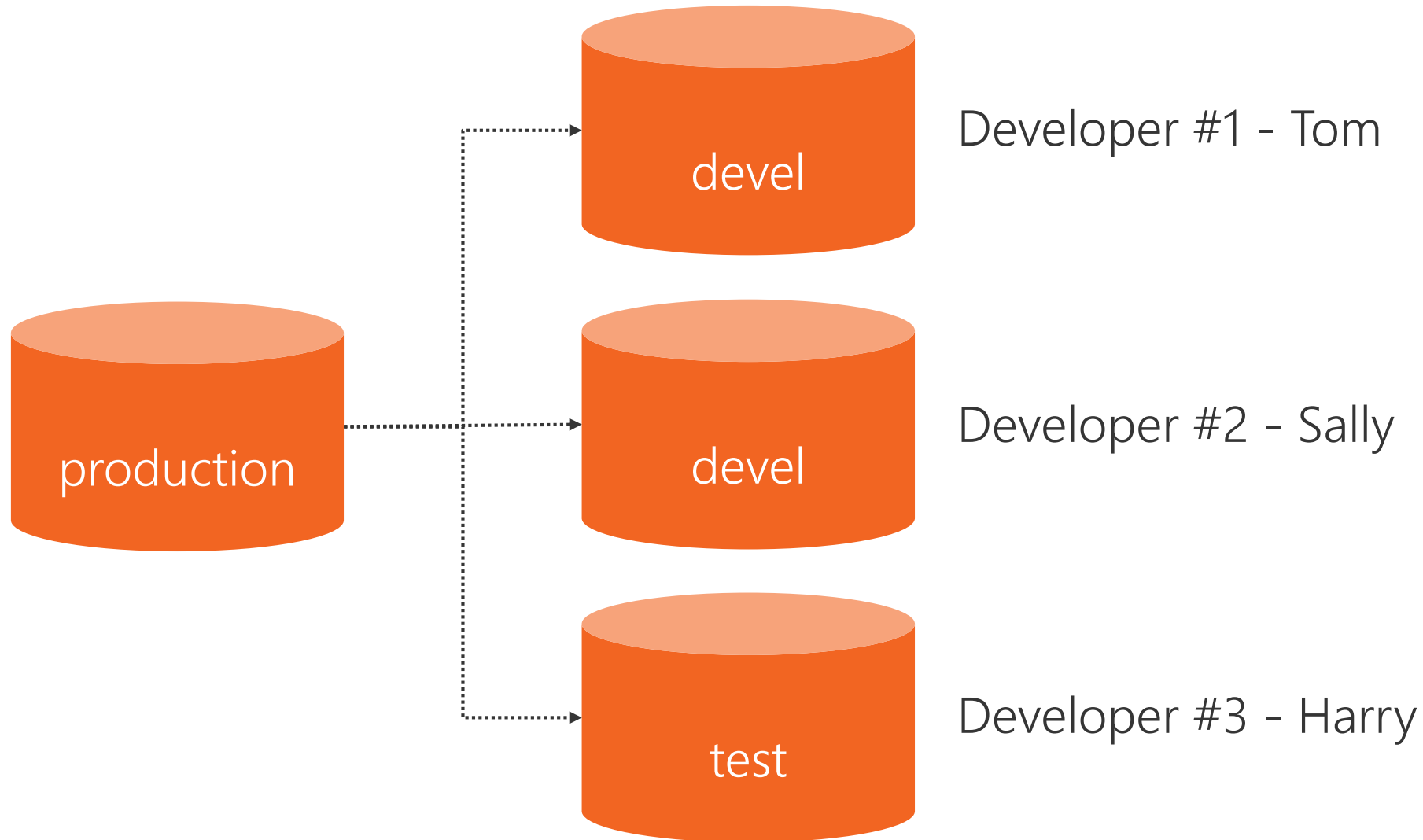
MB/S	Time Required	Comment
55	26:28:45	<i>Notebook, 1x HDD, 5400 RPM</i>
250	5:49:32	<i>Notebook, 1x SSD</i>
280	5:12:05	<i>SAN1, 10x 600 GB 15k, FC</i>
455	3:12:03	<i>SAN2, 30x 146 GB 15k, FC</i>
1727	0:50:36	<i>SAN3, 16x SSD, Infiniband 40G</i>
300	4:51:16	<i>LTO-7, tape drive</i>

The Problem

So, most give up and the environments become stale.



Possible Solution – *in a perfect world*



Possible Solution – *in the real world*



Attempting to Solve the Problem

- It is not easy for developers or testers to share environments simultaneously.
 - Changes made by each user conflict with one another.
- Serialize access to the limited number of environments.
 - Each user queues at the resource – their project takes more time.
 - Each user destructively changes the resource.
- We need a way and time to reset or cleanup environment between each use.
 - Reset or clean up takes time and resources.

Problem #2

Debug on production?

Test Environments

- The best place to test is on production.
 - Right data, right stats, right hardware.
 - End-users may disagree.
 - Competition for resources is a threat.
 - Cannot test code on »history« data.
- Next best is a »good« copy of production.
 - How good, how often, how quickly?
 - How do you supply a terabyte sized database to five different developer teams **without** 5TB of disk space for each copy?

The Problem





- > ----- Original Message -----
- > From: xxxxxxxxxxxxxxxxxxxxxxxx
- > To: "Urh Srecnik" <urh.srecnik@abakus.si>
- > Sent: Tuesday, 8. March 2016 11:43:55
- > Subject: The database from Friday
- >
- > Urh,
- > a procedure that ran over the weekend went wrong. Can you restore the
- > Friday's database somewhere where we could repeat the procedure
- > and debug it?

The Problem

- Last night report took 30 minutes instead of 2!
- Developer: I believe I can make this part of the job run many times faster - where can I test it?
- Different versions in PROD and DEV or TEST databases

Possible Solution

- Test on actual data

Mon	Tue	Wed	Thu	Fri	Sat	Sun
1	2	3	4	5	6	7
8	9	10	11	 12 <i>debug</i>	13	14
15	16	 17 <i>test</i>	18	19	20	21
22	23	24	25	 26 <i>debug</i>	27	28
 29 <i>production</i>	30	31				

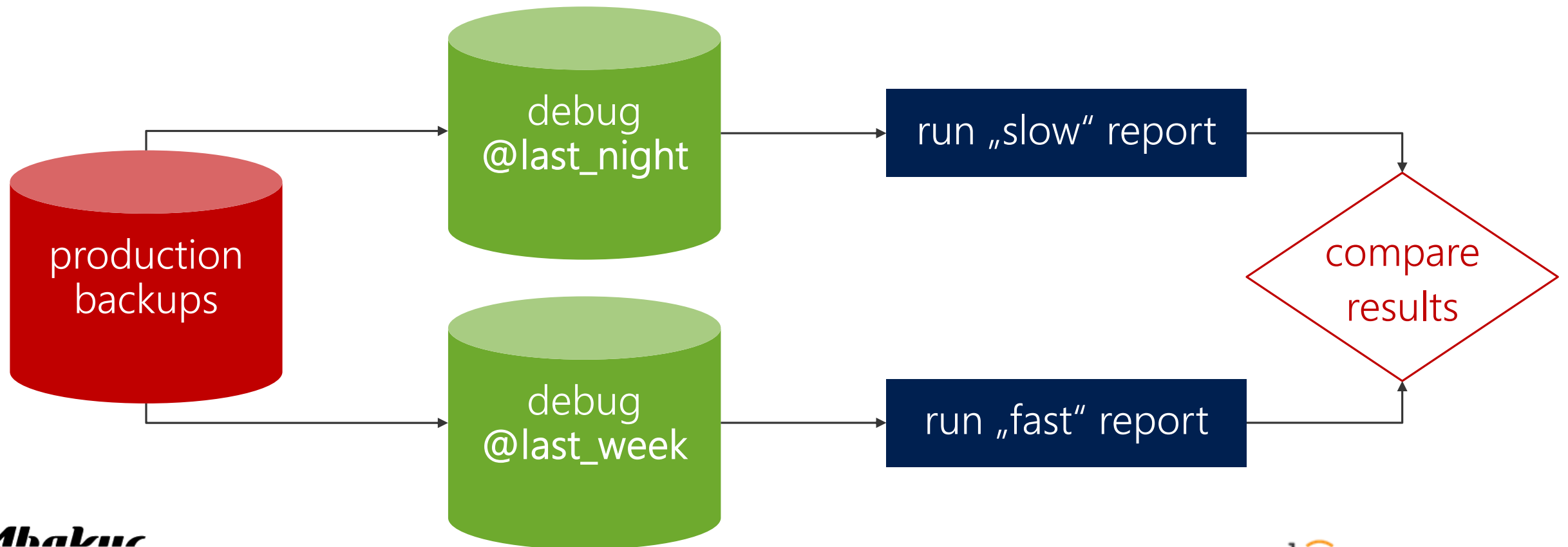
The Problem

„Last night report took 30 minutes instead of 2. Why? Make sure it doesn't do it again tonight.“

- data change,
- statistics might change,
- is execution plan the same as yesterday?

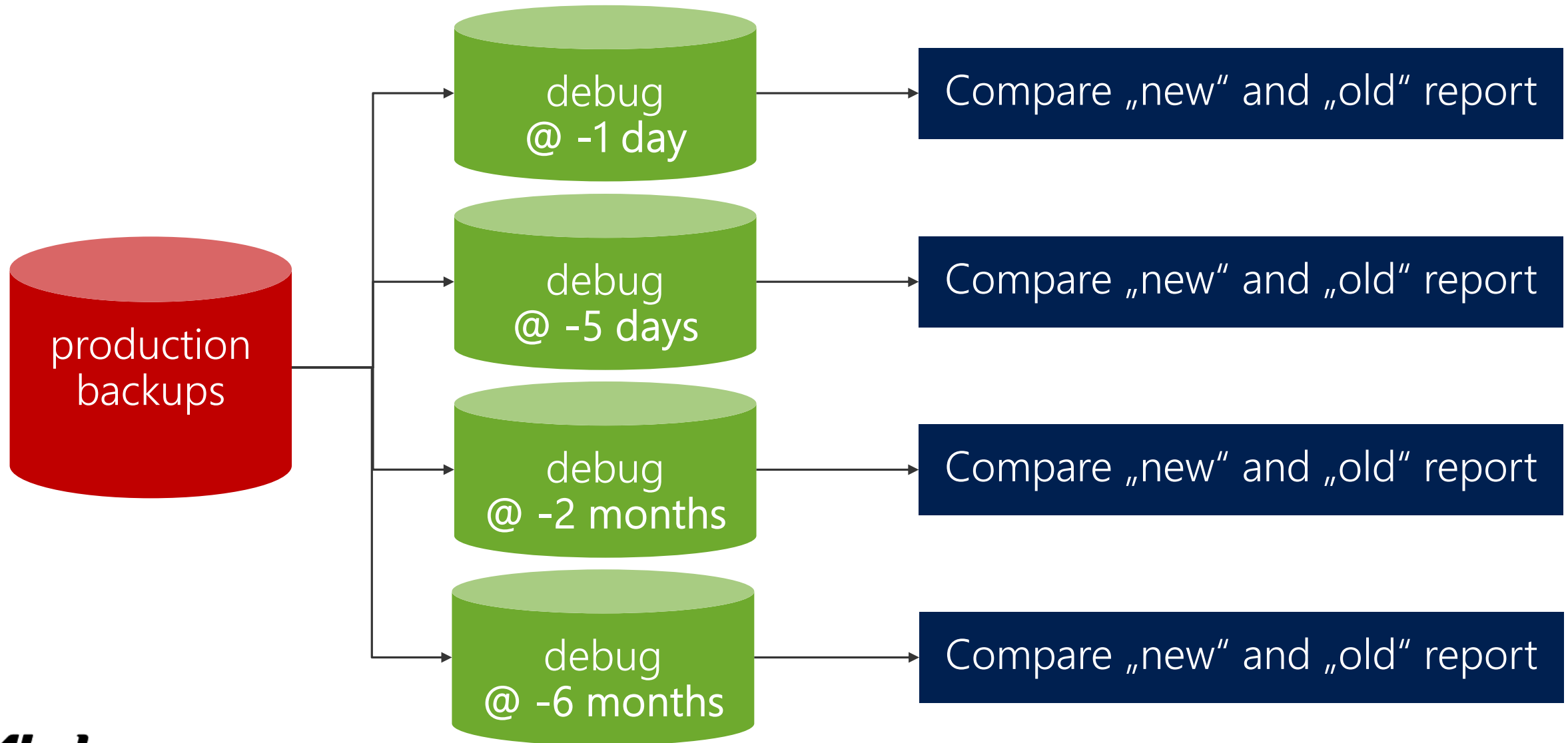
Possible Solution

- Virtual database from production backup
- Run new and old report and compare results.



Possible Solution

Developer: „I believe I can make this part of the job run many times faster - where can I test it?“

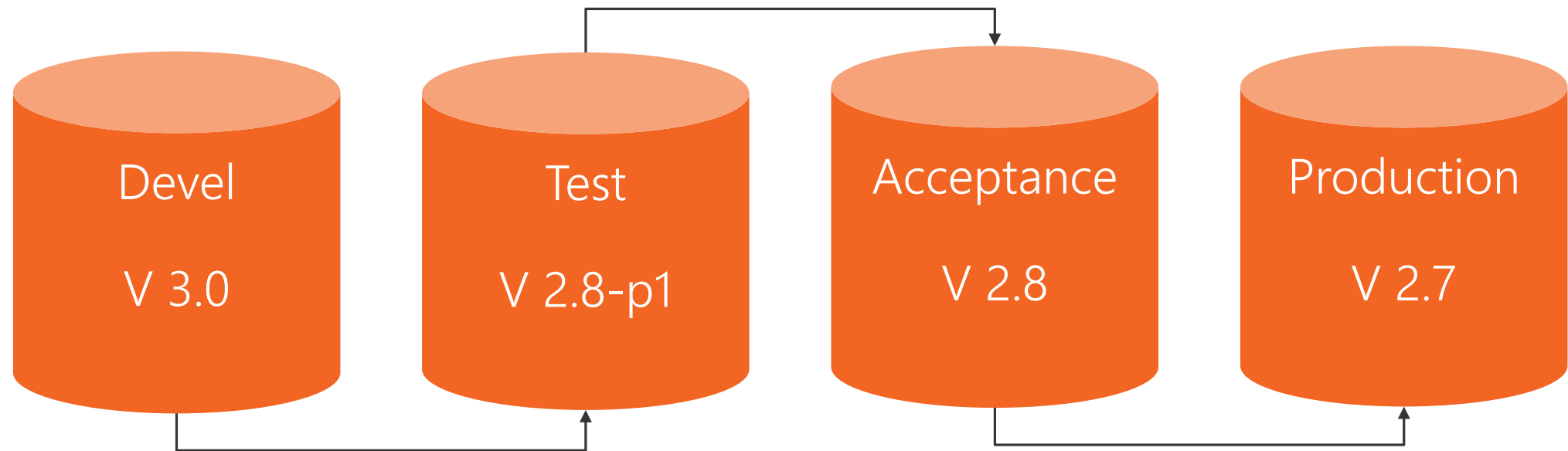


Problem #3

Environment versions

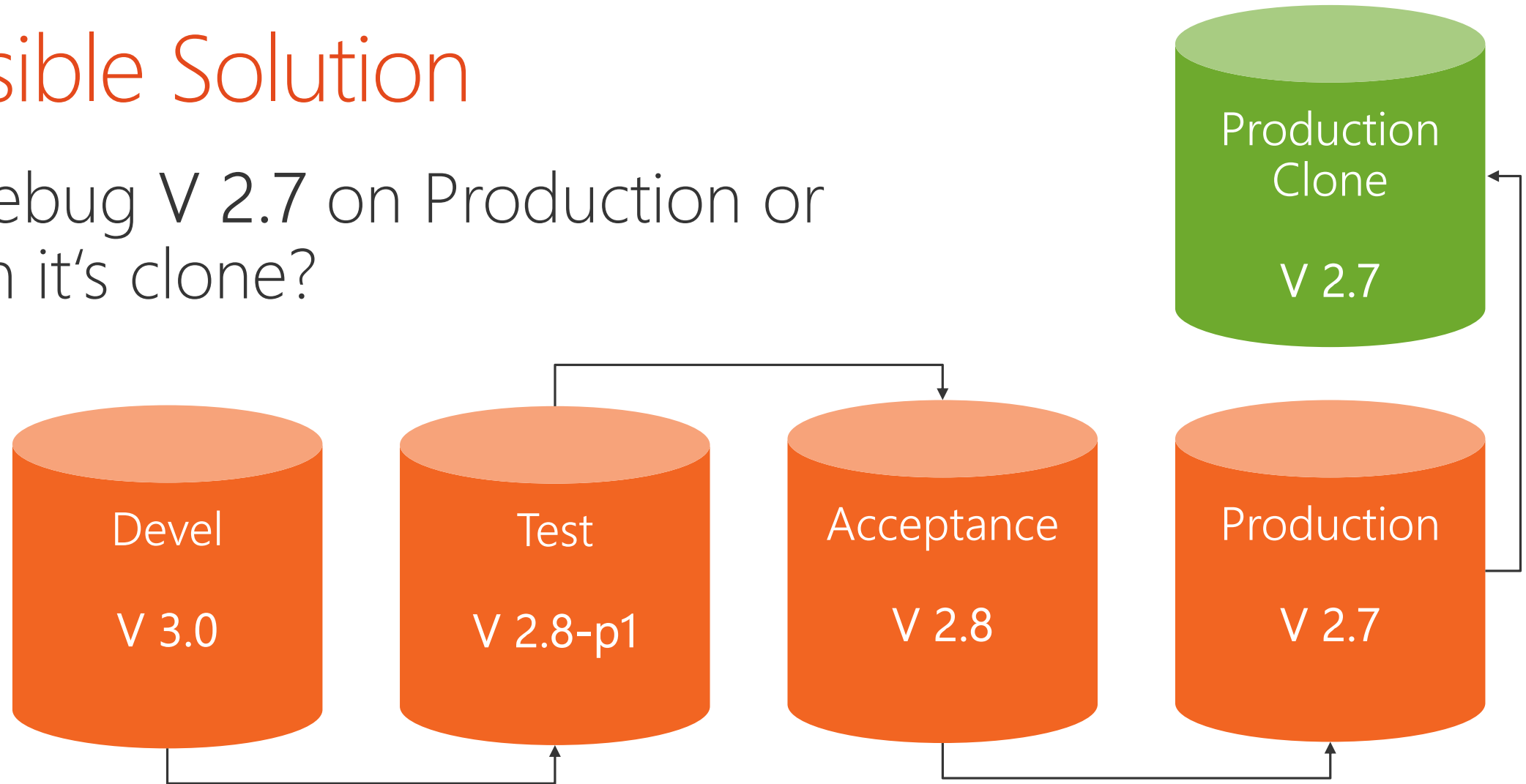
The Problem

Different versions on different environments.

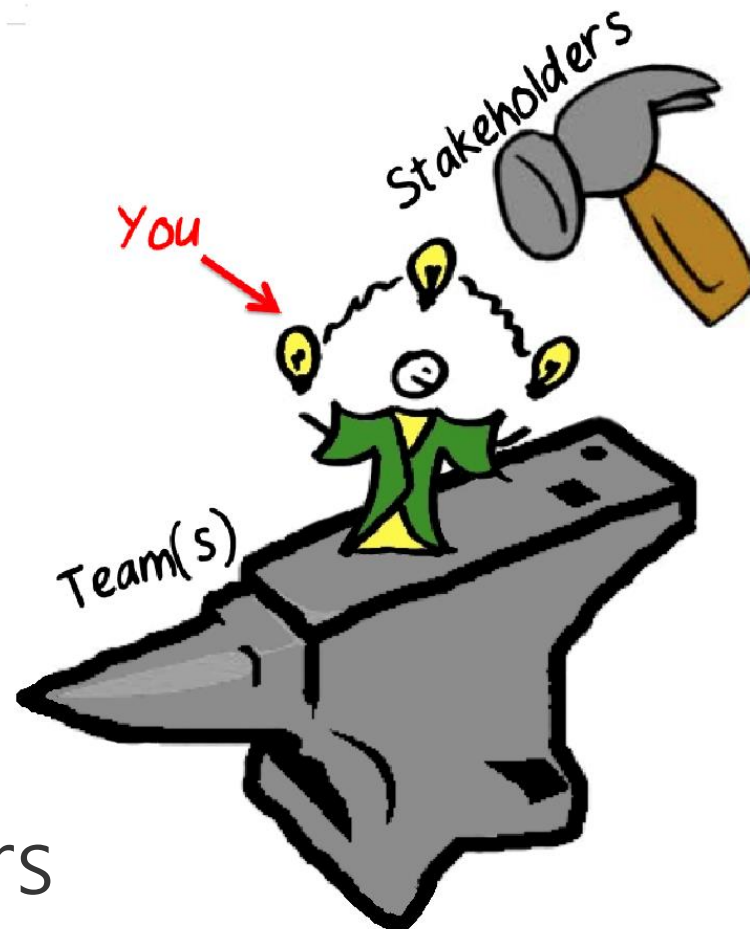


Possible Solution

- Debug V 2.7 on Production or on it's clone?



DBA



management

developers & testers

Problem #4

Solution(s) requirements

Solution Requirements and Cost

- **Server** to run the instance
- **Disk space** to host the database
- **Time** to find appropriate backups / prepare restore script
- **Time** to restore backup
- **Time** to apply transaction logs

- Standby databases? Snapshots?

Could we use...

- Single server,
- Minimal disk space by means of **block-level deduplication**
- Let **automation decide** which backup to use
- Apply **at most 4 hours** of transaction logs to get to...
- ... **Any point in time** in last 6 months?

- Without touching the production server

Mission: Impossible



Mission: Impossible



Deja Vu

- Connect to Deja Vu
- Open the database in required point-in-time.

Deja Vu: Resources



Backup Server *bks-master*

498 backups provide **65 days** of history since 2017-02-01.

762 GB of backup data is stored on 4 GB / **41 GB** physical volume.

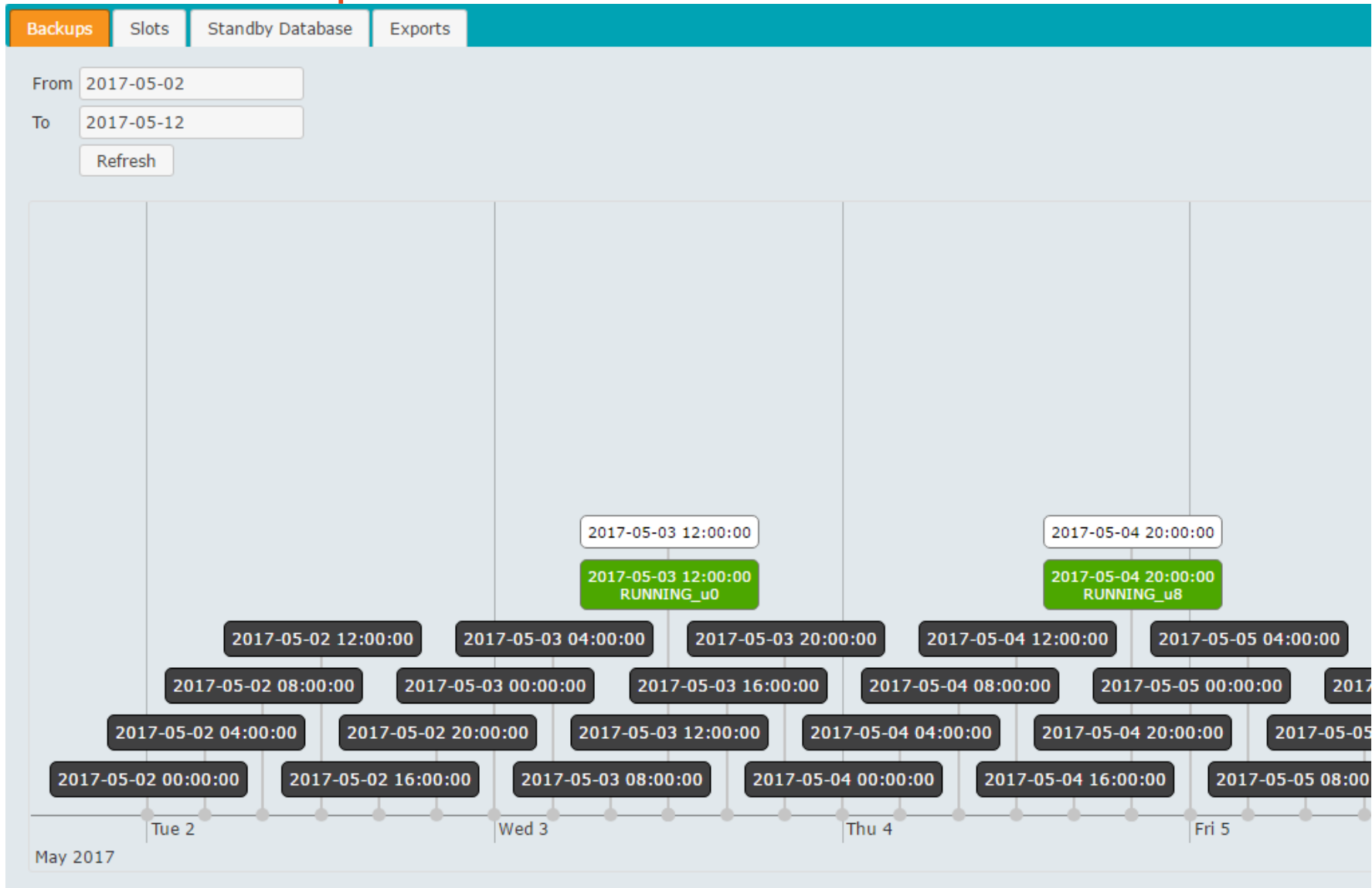
Administration My Session (BSADMIN) ▾

Resources

Type	Name	Actual Date	First Date	Last Date	Status	Monthly Growth	...
database_mssql	MSDEMO	2017-04-07 11:45:16	2017-02-11 15:00:00	2017-04-07 08:00:00	ACTIVE	495KB	
database_oracle	ORADEMO	2017-04-07 11:17:30	2017-02-28 11:09:31	2017-04-07 08:17:31	ACTIVE	20MB	

[Create Resource](#)

Deja Vu: Backups Timeline



Deja Vu: Backup Details

Backup Details ✕

Resource ID	234	MSDEMO
Backup ID	50635	/vbs/dedup_fast1/MSDEMO-2017-05-08-04-00
Actual Date	2017-05-08 04:00:00	Volume Group: SSD
Taken Date	2017-05-08 04:00:00	Origin: n/a
Persistent Until		Software: Standard Edition 13.0.1772.0
Directory Size	47MB	
Status	NORMAL	
Is Consistent?	<input type="checkbox"/>	
Is Snapshot?	<input type="checkbox"/>	
Is Succeeded?	<input checked="" type="checkbox"/> Mark as Failed	
Log File	backup_50635.log	

[Export Snapshot](#) [Create Snapshot](#) [Delete Snapshot](#) [Browse Backup](#) [Start Virtual Database](#)

Deja Vu: Open Virtual Database

Start Virtual Database ✕

Resource(s) * MSDEMO -- ▾

Slot ac ▾

Slave MSSQL-SLAVE ▾

Origin Historical Backup ▾ Perfect Recall

Origin Time 2017-05-04 08:00:00

Origin Backup #50545 Reuse Snapshot

Recover Until 2017-05-04 08:00:00 Point in Time Recovery

Open Mode READ ONLY ▾

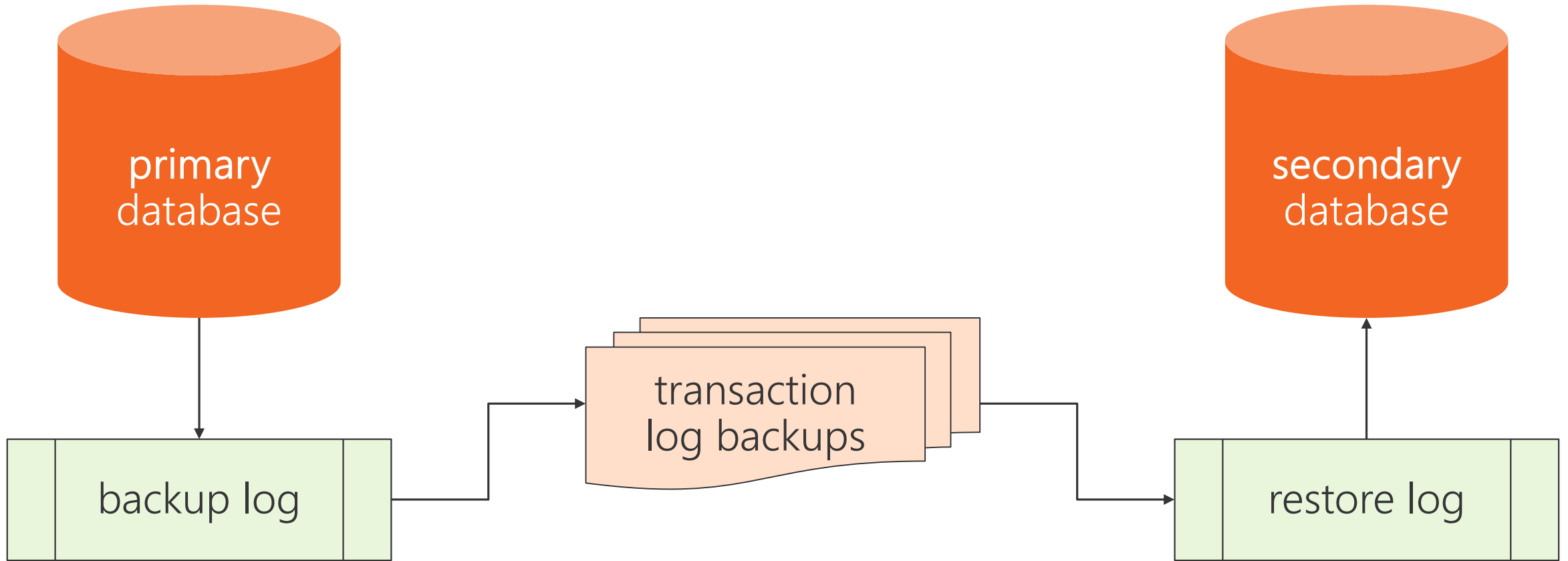
Access Method TDS - Microsoft SQL Connection ▾

Restore Point Create new restore point

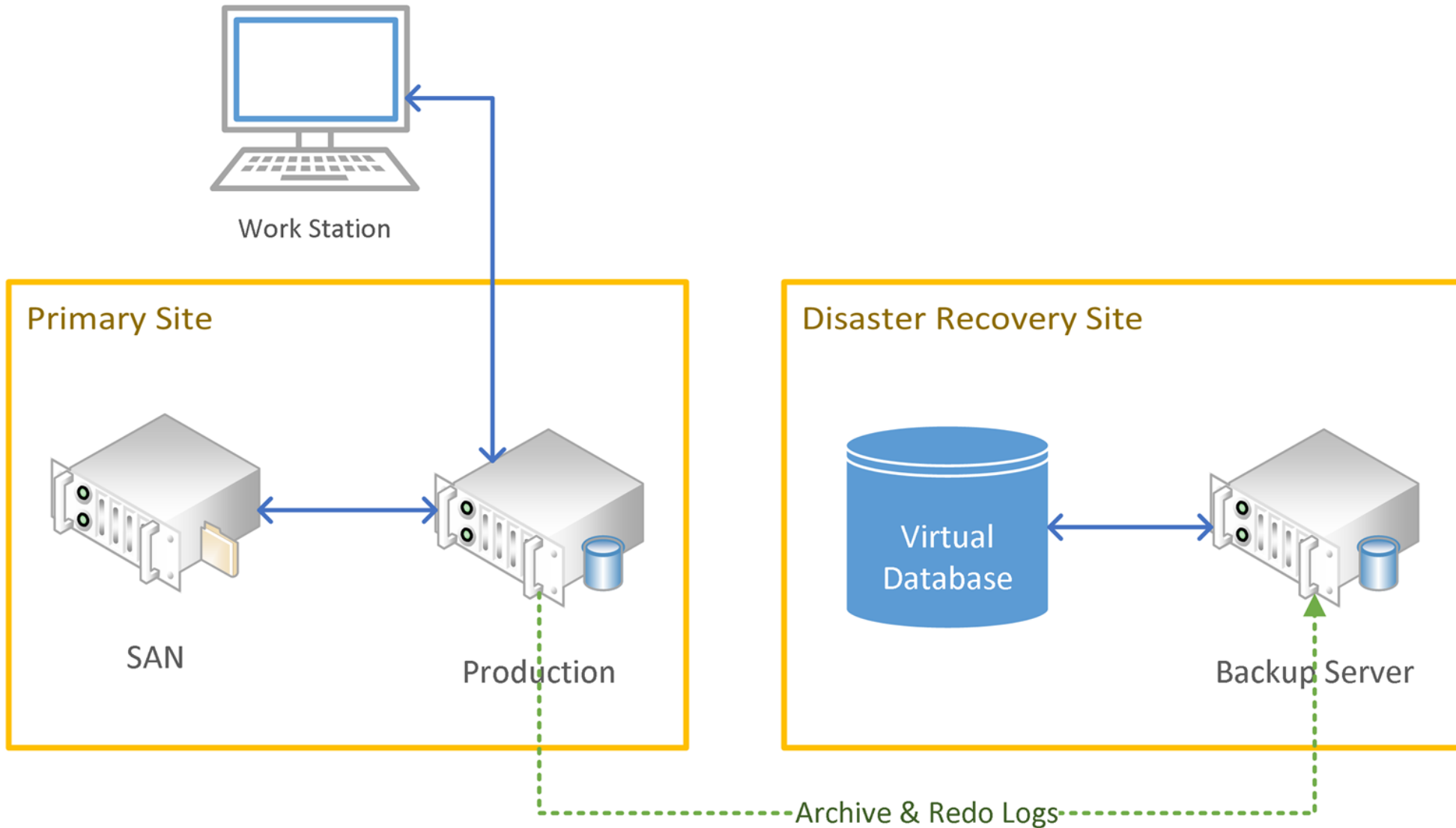
Deja Vu: Database Status

Slot Status	
Slot Name	u0
Running Server	MSSQL - SLAVE
Access Method	tds
Open Mode	READ ONLY
Actual Time	03-05-2017 12:00:00
Connect String	Server=MSSQL-SLAVE\MSDEMO;Database=u0_msdemo

Log Shipping (Physical Standby)



Deja Vu: Implementation Details



Deduplication

Backup Server *backup*

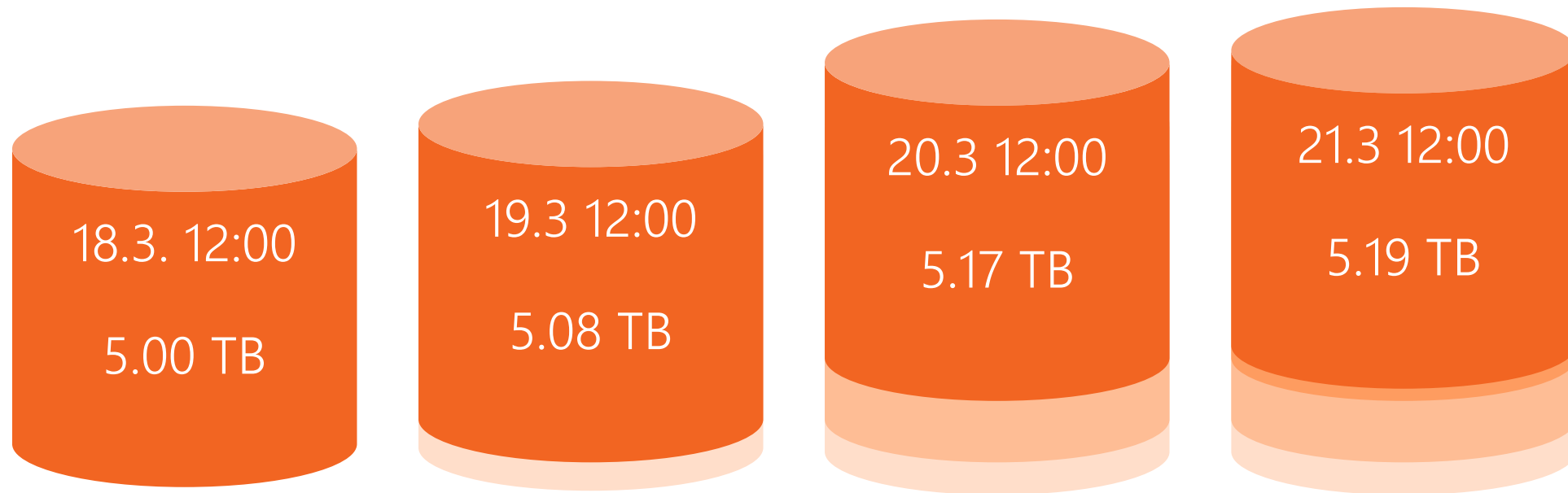
563 backups provide **577 days** days of history until 2015-02-27.

234.06 TB of backup data is stored on 3.75 TB / **4.00 TB** physical volume.

Backup Server

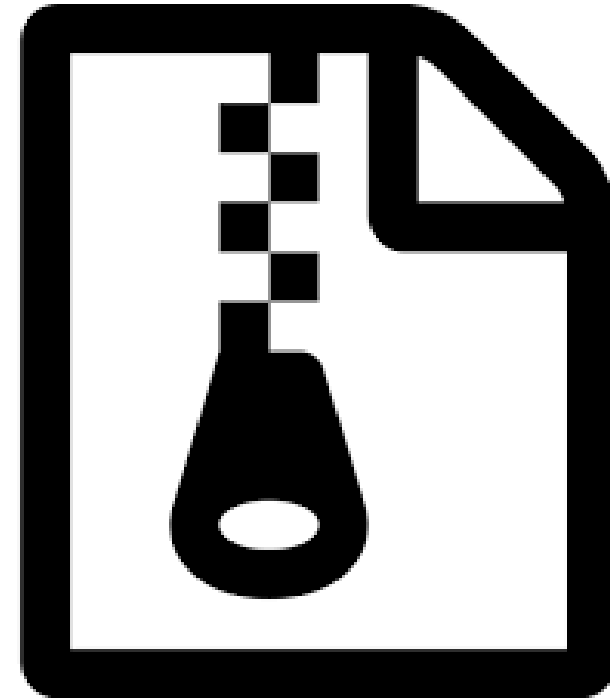
398 backups provide **113 days** days of history since 2016-06-06.

416,40 TB of backup data is stored on 21,90 TB / **39,09 TB** physical volume.



Compression

- Transaction logs should not be deduplicated
- But they are well compressed
 - About 5% - 30% of their original size



Tiered Storage

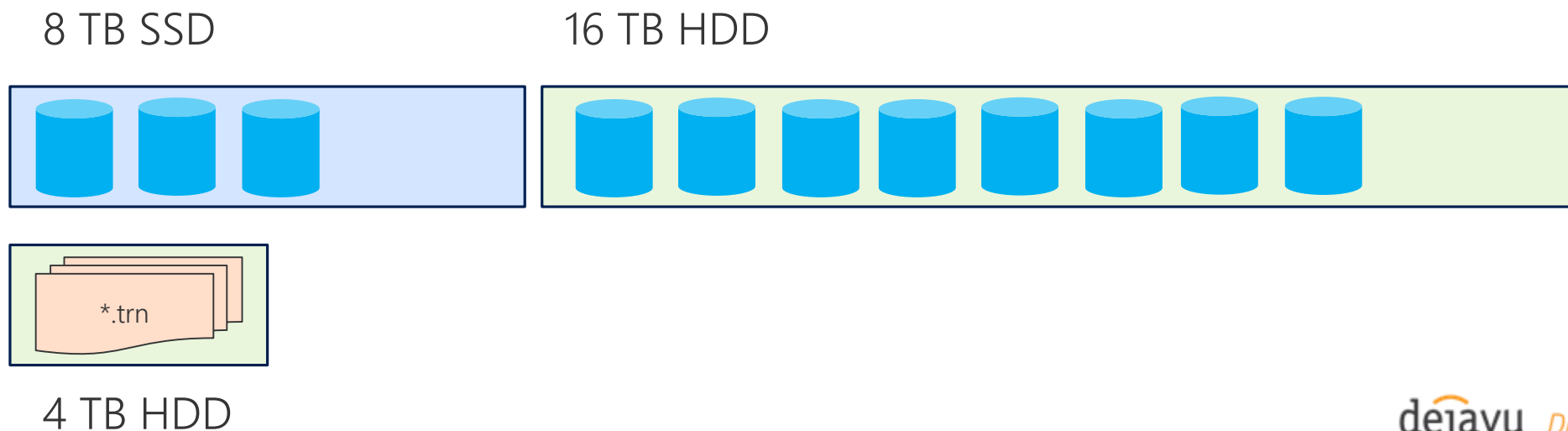
- Fast Storage (e. g. SSD)

- more expensive
- less capacity
- less history

- deduplicated areas

- Slow Storage (e. g. HDD)

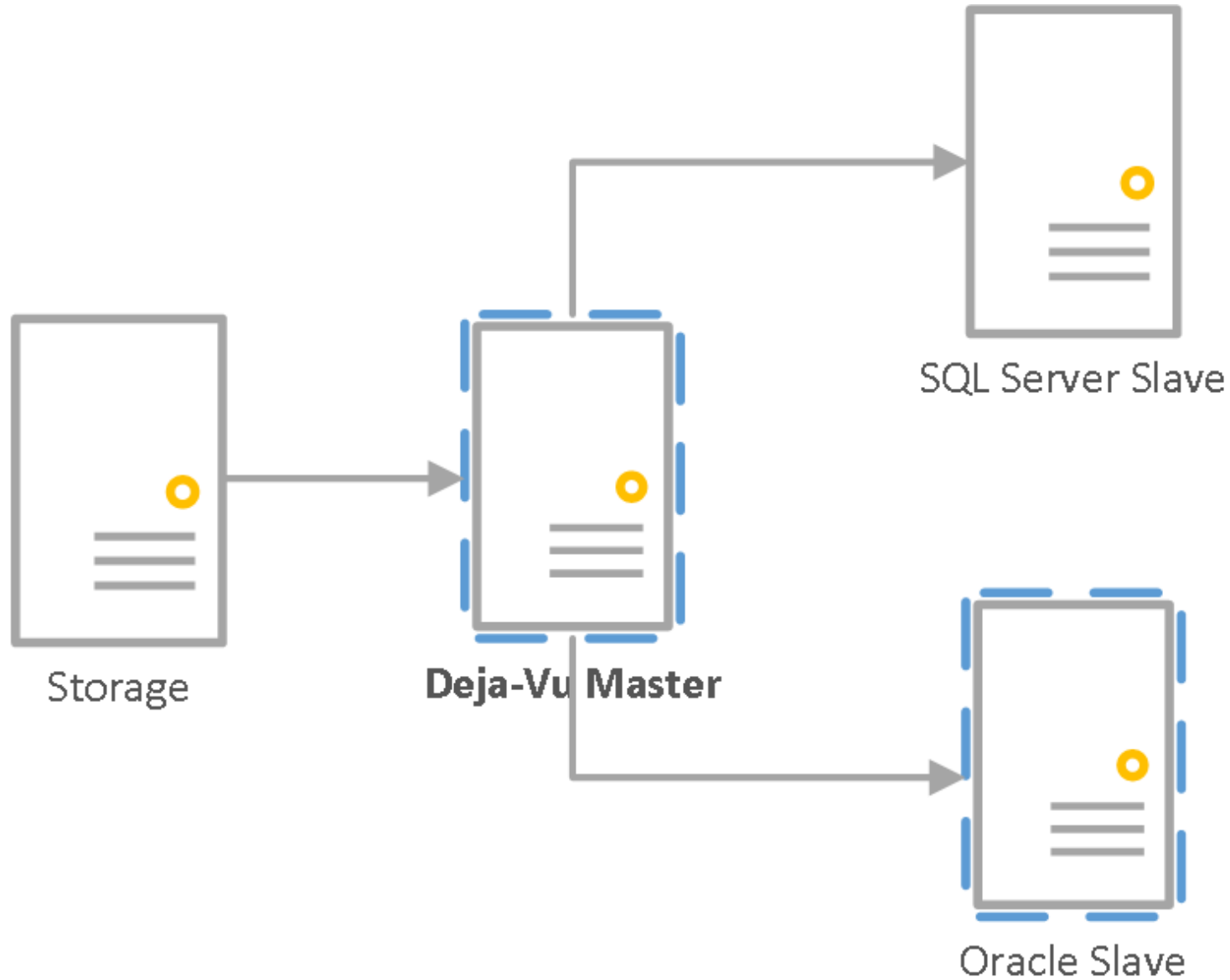
- less expensive
- more capacity
- more history
- deduplicated areas
- compression areas

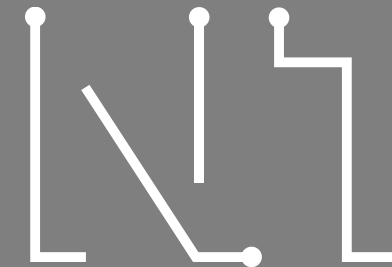


Deja Vu: Hardware



Deja Vu Deployment Options





KONFERENCA

PORTOROŽ, 15. DO 17. MAJ 2017