



NOVA BI

From MSBI to Pentaho

Bart Maertens



Content

- #1 Project background
- #2 Architecture: as-was, as-is
- #3 Environment
- #4 Lessons learned, caveats

#1 Project Background

About know.bi:

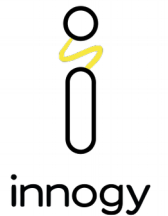
- Started in 2012
- Team of 9
- Heavily involved in community (#pcm14, #pcm16)



#1 Project Background

About Essent:

- Energy supplier in NL, BE
- Part of Innogy (daughter of RWE)
- 2.500 employees (2015)



#1 Project Background

was: **Provision**

- Microsoft-based
- On-premise

is: **NOVA**

- Open Source based
- Cloud-based (AWS)

#1 Project Background

#NOVA: UAAS

- **Utility As A Service**
- Designed for easy roll out to other sites, group companies

#Cloud:

- Scalability
- Efficiency
- Cost

#1 Project Background

#Project steps:

- Q4 2015 → Q1 2016:
 - Platform selection: PoCs with Pentaho and Jasper
- Q2-2016 → Q1-2017
 - Environment building
 - Migration
- End Feb 2017: first B2B release
- Q1 2018: B2C release

#1 Project Background

'Traditional BI' Team (9 members):

- analysts
- developers: ETL, report, dashboard
- project owner/manager

'Analytics' Team (4 members):

- data scientists
- data/software engineer

#1 Project Background

#Agile:

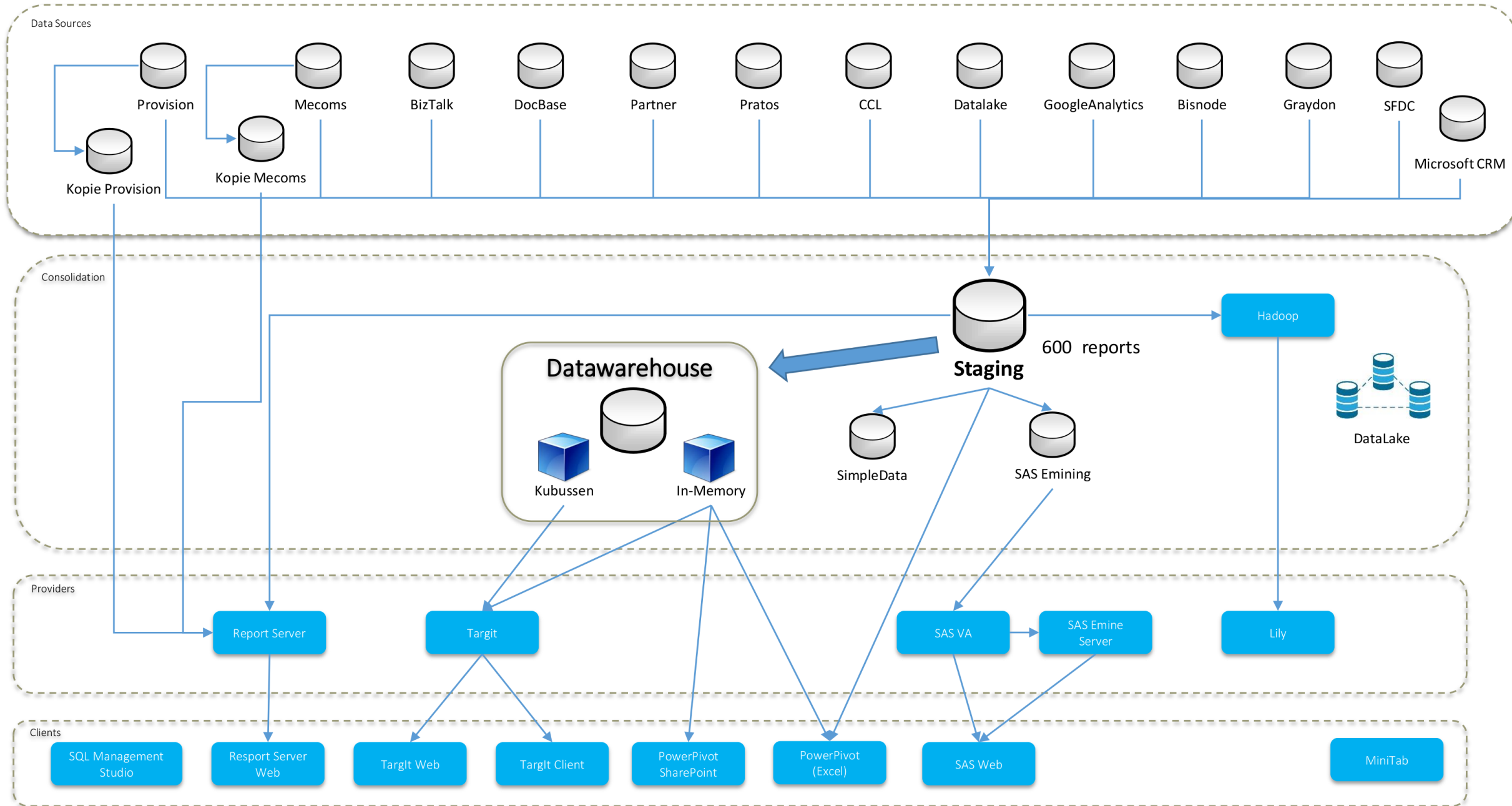
- 2 week sprints
- scrum master
- sprint planning
- refinement sessions + planning poker (story points)

#2 Architecture: as was

(Mainly) Microsoft based:

- **ETL:** SSIS
- **DWH:** MS SQL Server
- **Visualizations:** SSAS, SSRS, Targit
- **Analytics:** SAS
- (+ Excel everywhere!)

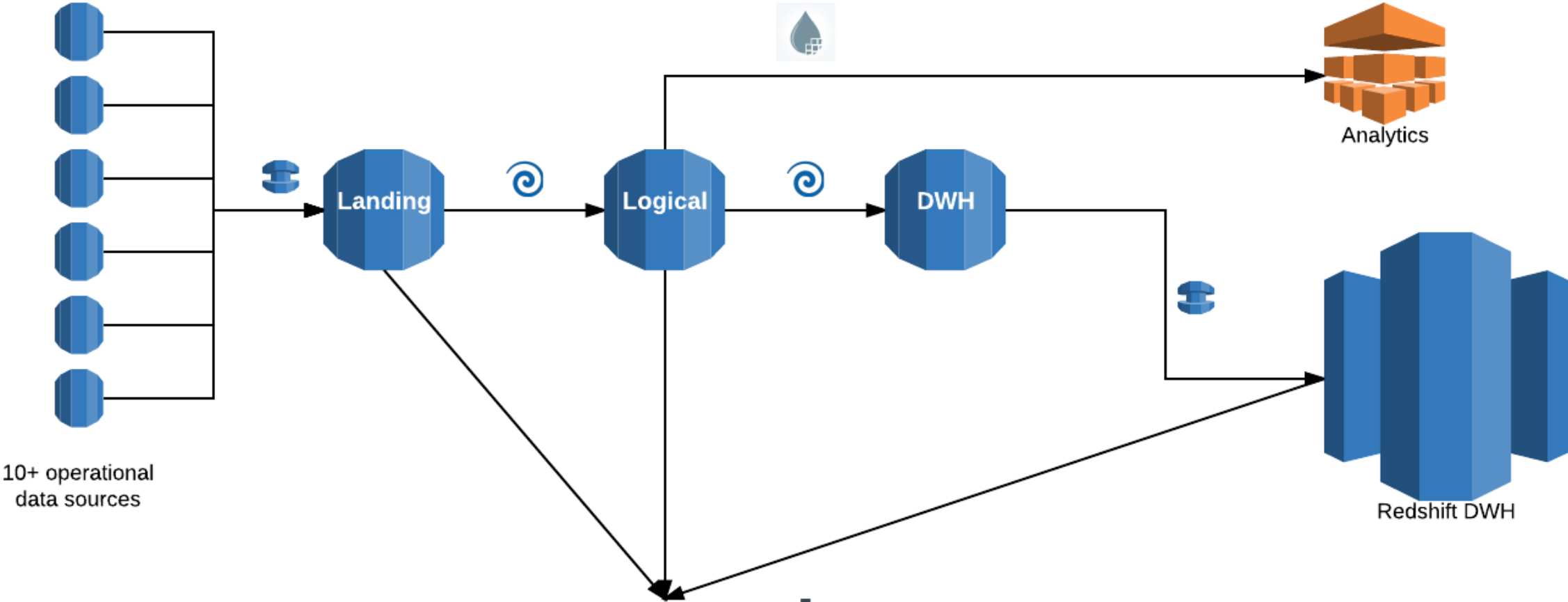
BI Landschap Essent.be AS-IS



#2 Architecture: as is

AWS/Pentaho based:

- **DMS**: read-only data source copies
- **ETL**: PDI
- **Logical** Layer: unified business layer
- **DWH**: PostgreSQL, Redshift
- **Visualizations**: Pentaho (Analyzer, Dashboard Designer, PRD)
- **Analytics**:
 - Read from logical layer through Nifi
 - R, Spark



#3 Technical Environment

- # 3 DEV, 1 UAT, 1 PRD
- # Bitbucket (git)
- # IAC: Puppet
- # Deployment:
 - Jenkins
 - deployment framework (know.bi)

#4 Lessons, caveats

- # starting BI development without real-life data is hard
- # (some) AWS (services) can bite
- # Agile (when done right) pays off



Thank you!



www.know.bi

info@know.bi

[@know_bi](#)