

Architektura BI řešení

Jiří Neoral

BI Data Architect Dixons Carphone

MVP: Data Platform

jiri@neoral.cz

www.neoral.cz

 @JiriNeoral

Osnova

1. Základní pojmy
2. Tradiční vs. Self Service
3. Příklad řešení
4. DWH a ETL/ELT
5. Semantická vrstva proč a jak
6. Prezentační vrstva
7. Nasazení, sdílení
8. Monitoring

Co je BI

- BI = Business Intelligence
- Oblast, jejímž cílem je pomoci uživatelům k lepšímu rozhodování na základě dat
- Oblasti využití
 - Business
 - Cokoliv jiného, co chcete aby fungovalo lépe a máte pro to data
 - Takže by se to dalo zkrátit na Intelligence 😊

Proces

- Sběr požadavků a analýza (nezávislé na technologii)
- Design řešení
 - mapování zdrojů (volba mastera)
 - dostupnost dat, nároky na frekvenci
 - high level design řešení (jaké vrstvy řešení budou potřeba)
 - nároky na dostupnost řešení (Disaster Recovery, availability)
 - IT schopnosti uživatelů
- Tvorba Datového skladu a ETL/ELT
- Tvorba semantické vrstvy implementující business logiku
- Prezentační vrstva
- Průběžně
 - testování
 - monitoring funkce, optimalizace, monitoring použití (kvůli zpětné vazbě)

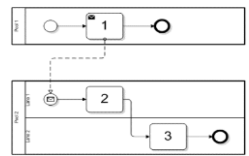
Tradiční vs Self Service BI

- Tradiční (řeší BI tým)
 - příprava dat
 - analytické modely
 - reporting
- Self service (něco z výše zmíněného si uživatelé řeší sami)
 - motivace může být rychlost dodávky
 - uživatel může narazit na nějaké limity (přístupy, technologie, pochopení systému, ...)

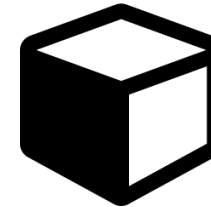
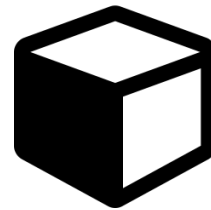
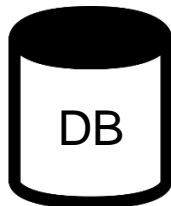
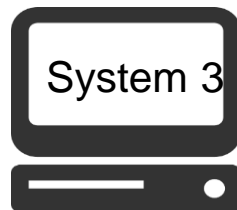
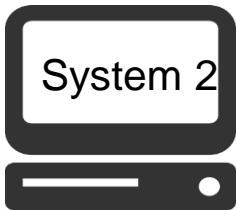
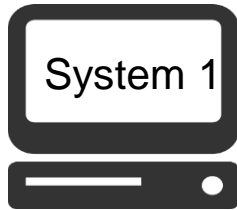
Uživatel se ptá



Jak si stojíme jako firma v porovnání prodejů s loňským rokem?



Process A



Cubes



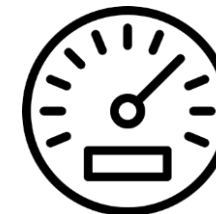
Report



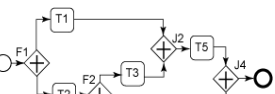
Ad Hoc



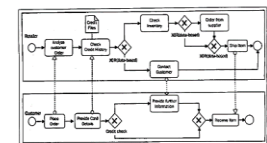
Regular



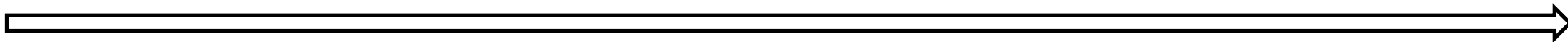
Dashboard

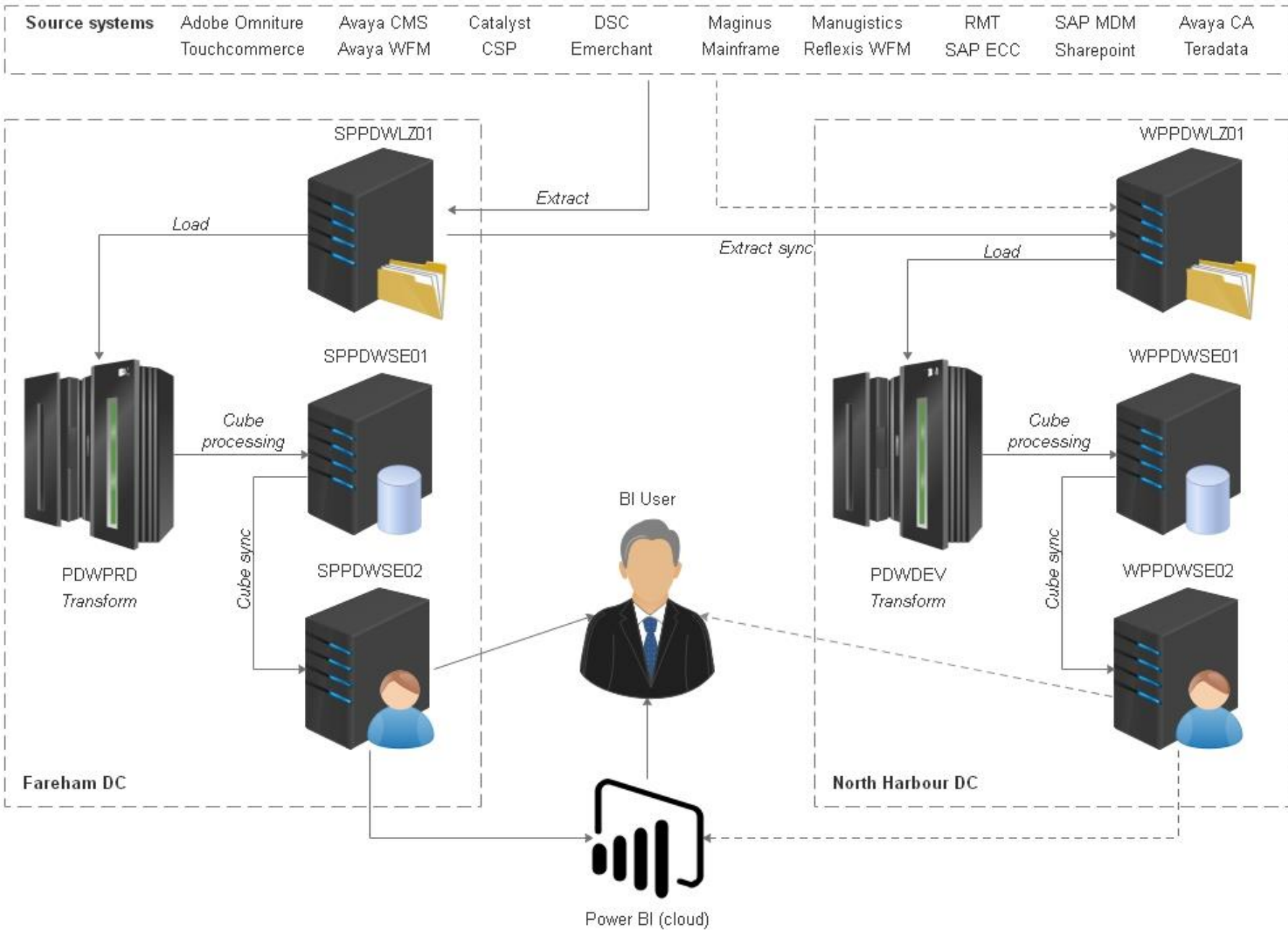


Process B



Process C

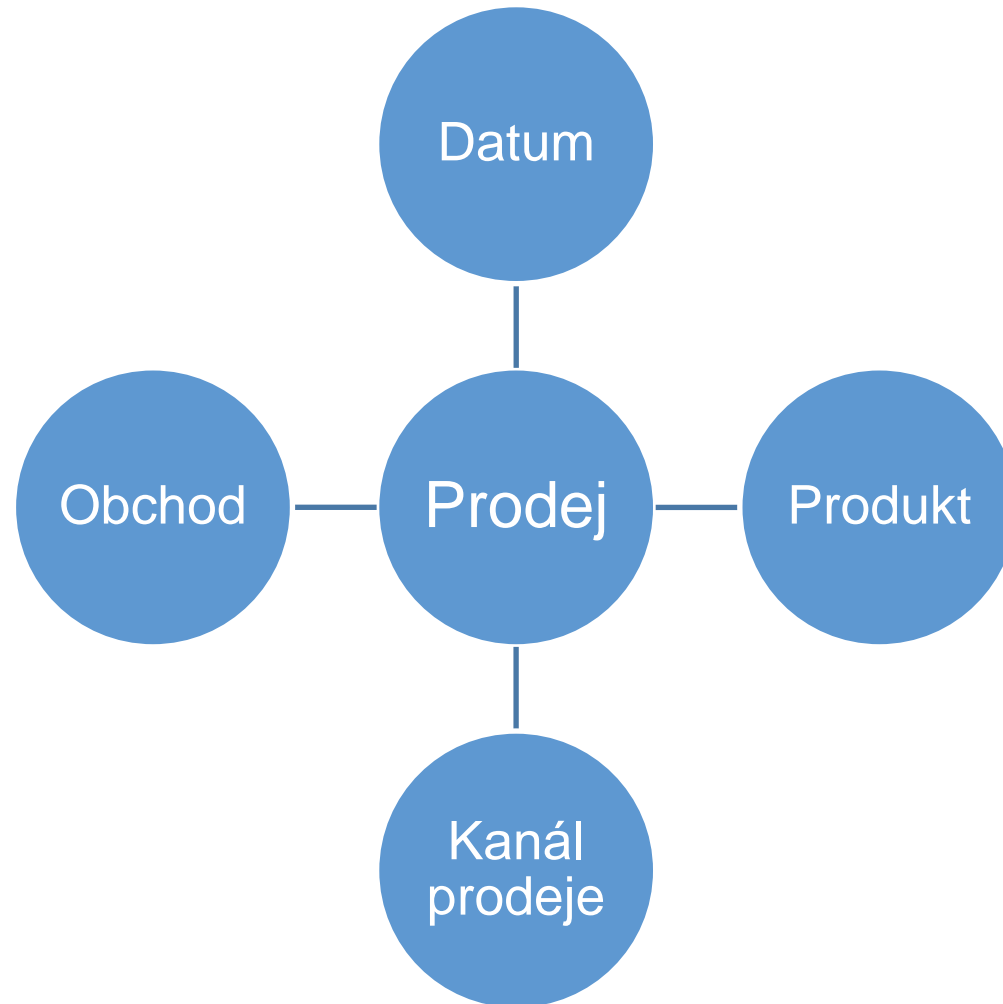




DWH řeší

- Propojení dat z více zdrojů
- Úprava schématu, aby bylo vhodné pro analýzu
- Sledování dat v historických souvislostech
- Čištění dat

Jednoduché schéma



Jak na dimenze

- Dimenze obsahuje
 - Business key
 - Surrogate key
 - Atributy
 - Atributy mohou tvořit hierarchie
- Verzování
 - SCD1
 - SCD2
 - https://en.wikipedia.org/wiki/Slowly_changing_dimension

Jak na faktové tabulky

- Dimenzionalita
- Granularita
- Typ tabulky
 - Transakční
 - Snapshotová
- Typ měřítek
 - Aditivní
 - Neaditivní
 - Semiaditivní

Benefity semantické vrstvy

- Tvorba byznys logiky
- Metadata pro opakované použití (hierarchie, metriky)
- Zpřístupnění možností pro ad-hoc analýzu
- Odstínění zátěže od DWH
- Urychlení určitého typu dotazů

Produkty společnosti Microsoft pro semantickou vrstvu

- SQL Server Analysis Services
 - Multidimensional (starší 2005, agregace)
 - Tabular (2010, in memory)
- Azure Analysis Services (Tabular)
- Power BI (Tabular)
- Power Pivot pro Excel

Demo

- Multidimensional
 - current performance demo (ad hoc analýza)
 - main (může to být i složité)
- Tabular
- Power BI

Prezentační vrstva

- Interaktivní reporty (Power BI)
- Reporty optimalizované pro tisk a export (SSRS)
- Mobilní reporty (Power BI)
- Ad hoc analýza (Excel)

- Cloud (Power BI služba)
- On premises nasazení (Power BI report server/SSRS)

Demo prezentační vrstva

- Tvorba Power BI reportu
- Tvorba Paginated reportu

- Nasazení do Power BI Service
- Sdílení

Monitoring

- Používání reportů
- ETL procesů

Dotazy po přednášce

Jiří Neoral

BI Data Architect Dixons Carphone

MVP: Data Platform

jiri@neoral.cz

www.neoral.cz

 @JiriNeoral