

# Baze podataka II

## Modul 12 – Jezik SQL

Administracija DBMS okruženja

ETL proces



# Summary

- Šta je ETL proces?
- Primjeri importovanje i eksportovanje podataka



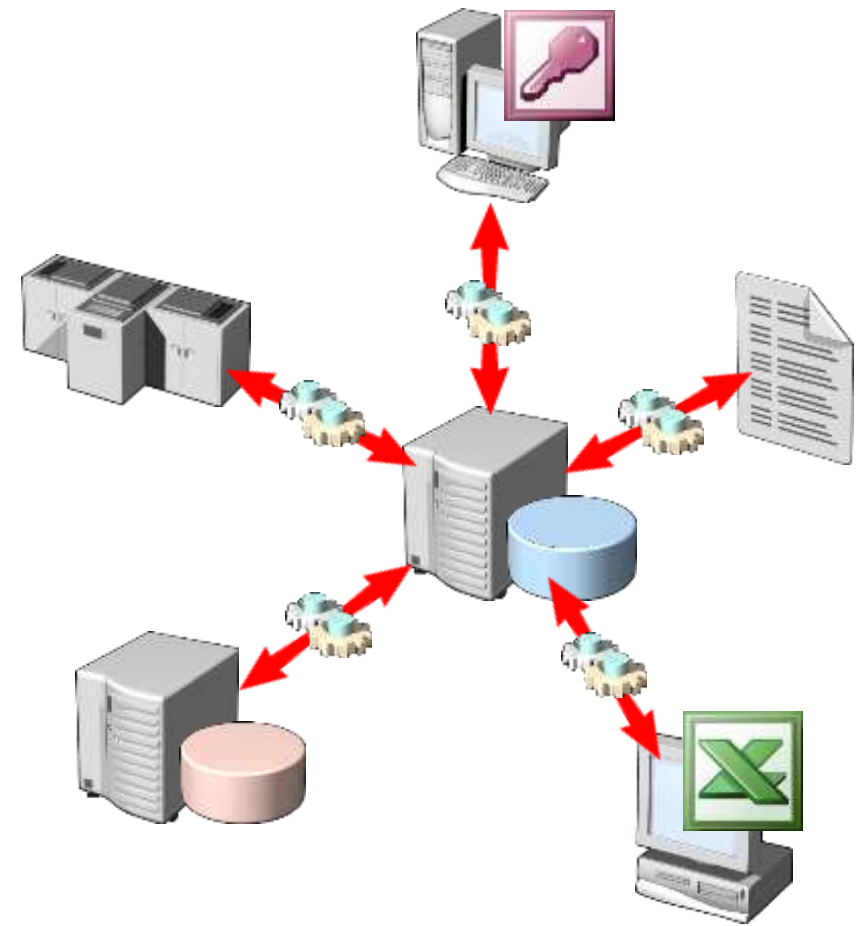
# Lekcija 1: Šta je ETL proces?

- Šta je transfer podataka?
- Šta je transformacija podataka?
- Alati i metode
- Poboljšanje performansi ETL procesa



# Šta je transfer podataka?

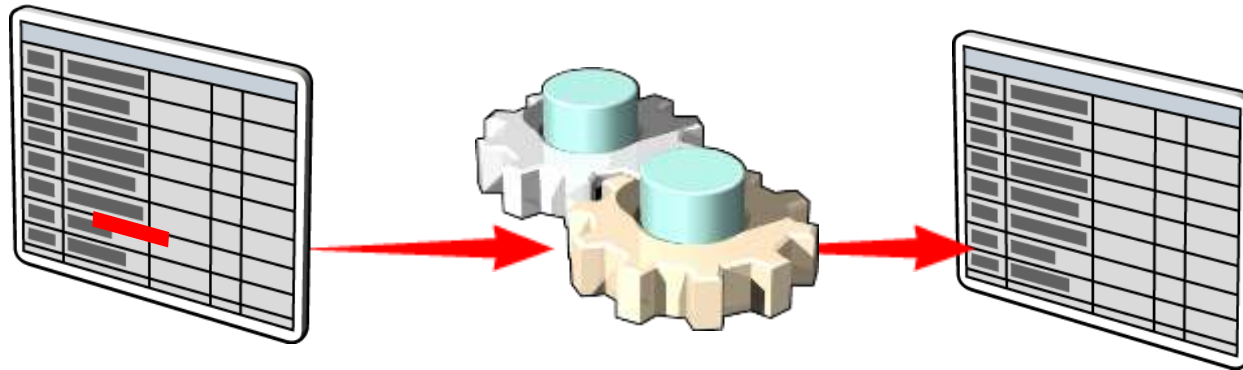
- **E**xtract
- **T**ransform
- **L**oad
- **Scenariji:**
  - Kopiranje između servera;
  - Eksportovanje rezultata upita u fajl
  - Importovanje sadržaja fajla u tabelu
  - Restrukturiranje podataka
  - Migracije
  - Data Warehouse
  - ...



# Šta je transformacija podataka?

## ● Transformacija podataka uključuje sljedeće:

- Promjenu formata;
- Agregaciju;
- Očuvanje konzistentnosti;
- Validaciju;



# Alati i metode

Bulk Copy Program (bcp)



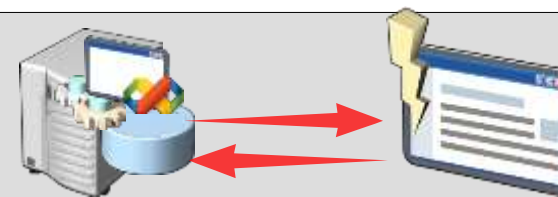
BULK INSERT



OPENROWSET(BULK)



Import / Export Wizard



XML Bulk Load



# Poboljšanje performansi ETL procesa

- **Isključiti ograničenja, indekse i okidače**
  - Nema potrebe za provjerom ograničenja za svaki zapis
  - Indeksi mogu usporiti kompletan proces
  - Provjeriti poslovne zahtjeve prije isključenja okidača
- **Minimizirati logiranje**
  - Baza bi trebala biti u BULK LOGGED ili SIMPLE modu rada
  - Provjeriti korisničke mehanizme logiranja aktivnosti
- **Minimizirati konverziju podataka**
  - Ako je moguće držati se izvornih formata podataka

# Lekcija 2: Primjeri importovanje i eksportovanje podataka

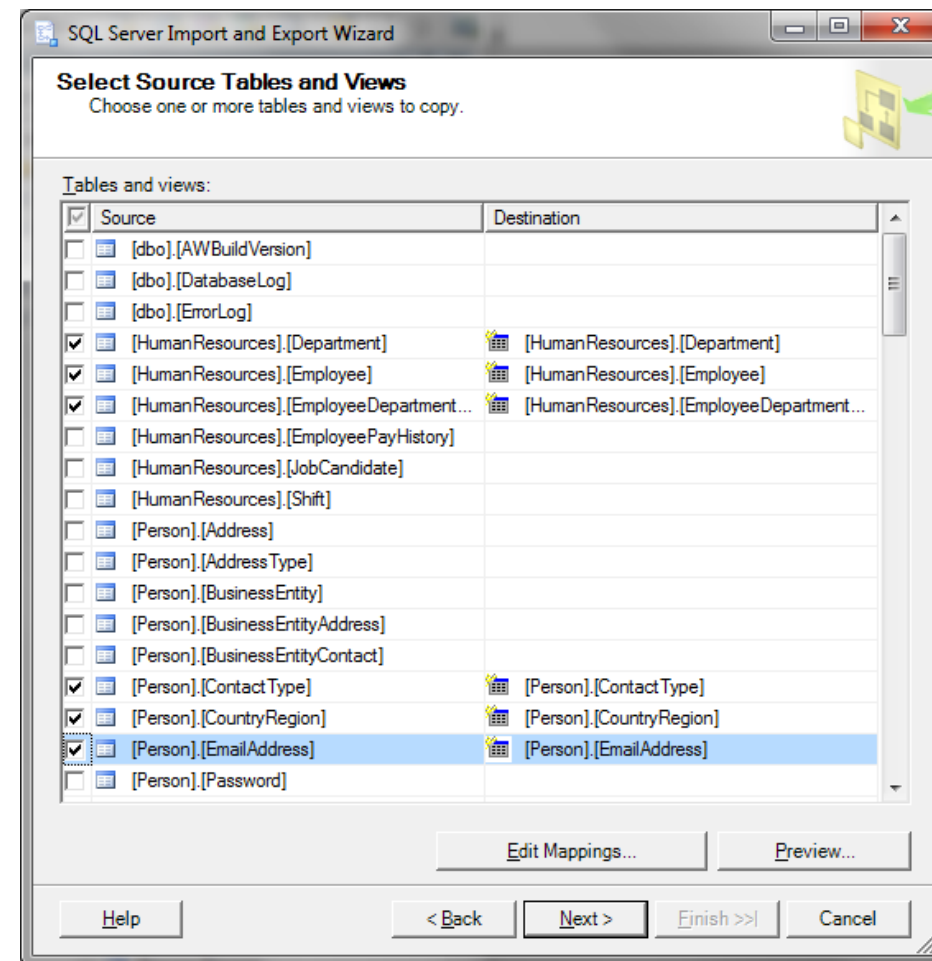
- **SQL Server Import/Export Wizard**
- **bcp alat**
- **BULK INSERT komanda**
- **Excel**





# SQL Server Import/Export Wizard

- Vrlo jednostavan alat jednostavne i srednje teške ETL procese
- Radi sa heterogenim izvorima podataka
- SQL Server Integration Services kompatibilan



# bcp alat

- Konzolni alat za import/eksport podataka
- Može da koristi format fajlove

## **Pravljenje format fajla:**

```
bcp Adv.Sales.Currency format nul -T -c -x -f Cur.xml
```

## **Eksport podataka u fajl:**

```
bcp Adv.Sales.Currency out Cur.dat -T -c
```

## **Import podataka preko format fajla:**

```
bcp tempdb.Sales.Currency2 in Cur.dat -T -f Cur.xml
```

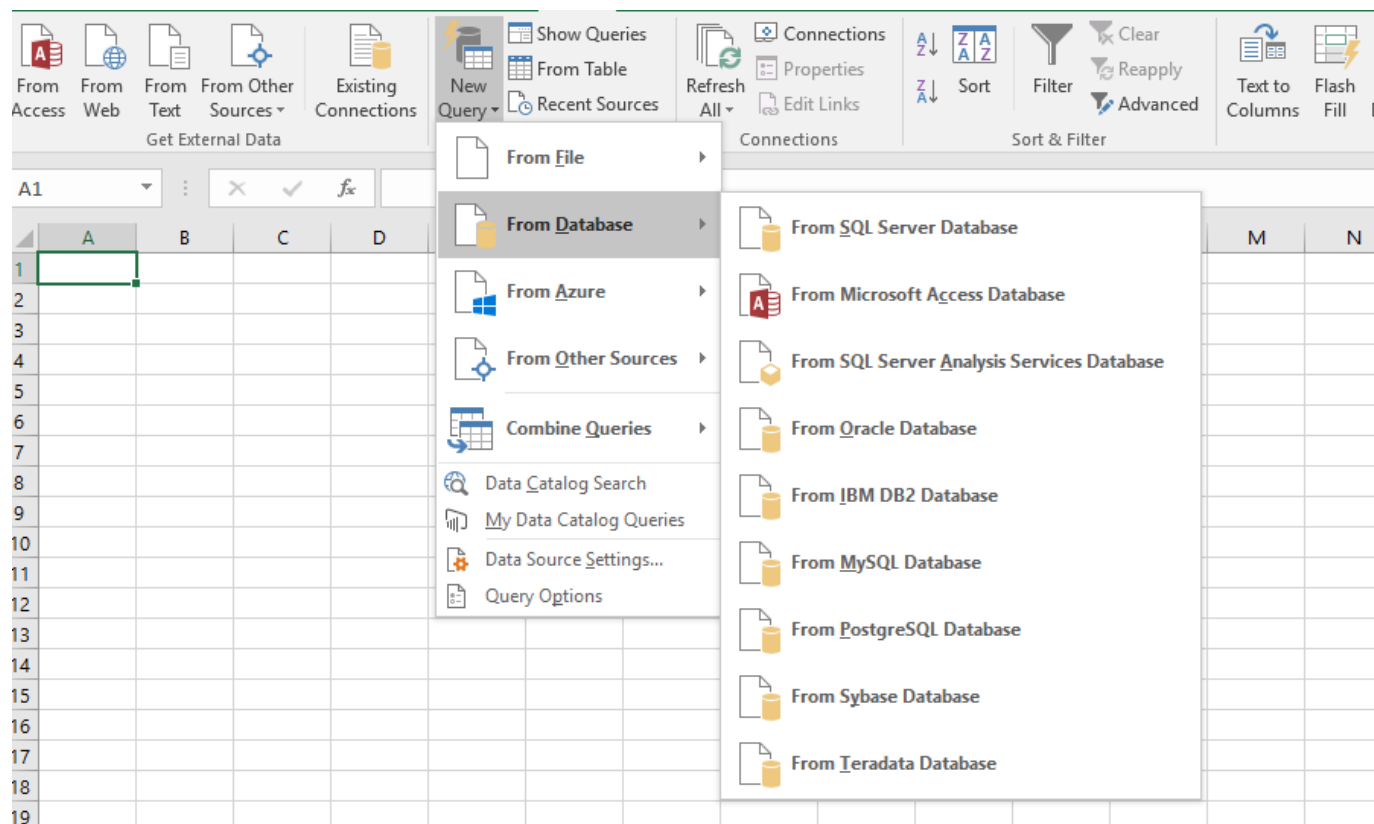
# BULK INSERT komanda

- Slično kao i bcp
- Izvršava se unutar SQL Server procesa
- Posjeduje CHECK\_CONSTRAINTS i FIRE\_TRIGGERS opcije

```
BULK INSERT AdventureWorks.Sales.OrderDetail
FROM ,C:\orders\arhiva.txt'
WITH
(
    FIELDTERMINATOR = '|',
    ROWTERMINATOR = '|\\n'
);
GO
```

# Excel

- Odličan frontend za mnoge različite formate podataka
- Kvalitetna podrška za kompletan ETL proces
- Podrška za cloud servise



# Pitanja

