

Увод у релационе базе података

5



Саша Малков
Универзитет у Београду
Математички факултет
2023/2024

[PM13]
Увод у РБП
Саша Малков



Тема 2

Архитектура база података

Циљеви и претпоставке

Основни циљеви СУБП



- Систем за управљање базом података мора да обезбеди:
 - Делљивост података
 - Вишекориснички рад са подацима
 - Употребу података у сложеним мрежним окружењима
 - Испуњавање различитих потреба различитих апл. и класа корисника
 - Интегритет података
 - Поуздан рад
 - ACID својства трансакција
 - отпорност на отказе
 - Безбедан рад
 - ограничен приступ подацима

Циљеви и претпоставке

Основни циљеви СУБП (2)



- Поред тога, очекује се и:
 - Ефикасност система
 - Приступачност (тј. ефикасност / лакоћа употребе)
 - стандардизовани начини употребе
 - стандардизовани нормативи и концепти
 - Пуна и ефективна контрола и управљање (администрирање)
 - Транспарентност
 - кориснике не занимају технички аспекти



Компоненте система база података

- Хардвер
 - Рачунарски систем
 - Инфраструктура
- Софтвер
 - СУБП
 - Корисничке апликације
 - Алати за развој софтвера
- Корисници
 - Крајњи корисници
 - Развијаоци
 - Администратори



СУБП

- Систем за управљање базама података (СУБП) је најважнија софтверска компонента
- Задатак СУБП
 - Управљање организовањем и чувањем података
 - Управљање преносом података између компоненти система
 - Управљање алгоритмима и методима употребе података на ниском нивоу
 - Пружање корисничких и апликативних интерфејса за коришћење података
 - Старање о интегритету података
 - Старање о безбедности (права приступа)
 - Заштита података (у случају неисправности или отказа)



Администратори

- Администратор СУБП
 - лице одговорно за рад и перформансе целог СУБП
 - стара се расподелом ресурса и приоритета између база података
- Администратор базе података
 - лице одговорно за прављење и одржавање објекта конкретне базе података
 - стара се о исправности и ефикасности базе података
 - стара се о физичком пројекту базе података
- Администратор података
 - лице одговорно за садржај базе података
 - није технично лице у мери као претходна два
 - стара се само о концептуалном пројекту базе података



Појам архитектуре

- Архитектура система представља концептуалну дефиницију његове структуре
 - компоненте
 - функције
 - међусобне односе компоненти
 - интеракције



Архитектура СУБП

- Архитектура СУБП мора да испуни основне циљеве СУБП
- У идеалном случају, архитектура
 - испуњава све постављене циљеве
 - сакрива све елементе имплементације од корисника
- У пракси то није могуће
 - модел употребе података је тесно повезан са моделом њихове организације



Архитектура СУБП (2)

- Архитектура система за управљање базама података (СУБП, енгл. *DBMS*) обично прати предлог стандарда из 1975. године
 - позната као “*ANSI-SPARC* архитектура”
 - није формално усвојена али је заживела
 - апстрактно је дефинисана и широко у употреби

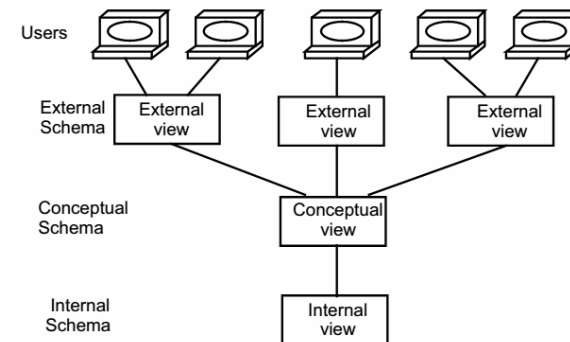


ANSI-SPARC архитектура СУБП

- *ANSI-SPARC* архитектура почива на приступу три схеме
 - Често коришћен приступ у развоју софтвера
 - Препознаје потребу
 - за различитим погледима на структуру система
 - за раздвајањем корисничке репрезентације од имплементације
- Схеме се разликују према посматрачу и природи представљања
 - спољашња схема (енгл. *external schema/view/level*)
 - концептуална схема (енгл. *conceptual schema/view/level*)
 - интерна схема (енгл. *internal schema/view/level*)



ANSI-SPARC архитектура СУБП (2)



извор: Ozsu, Valduries, Principles of Distributed Database Systems, 2.ed, Prentice Hall, 1999.



ANSI-SPARC архитектура СУБП (3)

- **Спољашња схема**
 - представља поглед на базу података из угла корисника
 - различити корисници имају различите потребе
 - имају потребу да виде различите скупове података
 - и различито организоване
 - за различите кориснике може да се прави више спољашњих схема
 - често се назива и *ниво погледа* или *ниво корисника*
 - релативно висок ниво апстракције
 - углавном одвојен и од конкретне имплементације и од конкретних ограничења која уводи модел података
- Концептуална схема
- Интерна схема



ANSI-SPARC архитектура СУБП (4)

- Спољашња схема
- **Концептуална схема**
 - описује заиста имплементиран скуп података и односа међу подацима
 - прилагођен конкретном моделу података, у нашем случају релационом
 - често се назива и *логички ниво* или *ниво администратора*
 - ниво умерене апстракције
 - одвојен је од физичке имплементације, али не и од модела организације података
- Интерна схема



ANSI-SPARC архитектура СУБП (5)

- Спољашња схема
- Концептуална схема
- **Интерна схема**
 - описује имплементацију на физичком нивоу
 - непосредно, онако како су подаци имплементирани у СУБП
 - често се назива и *ниво система*
 - најнижи ниво апстракције

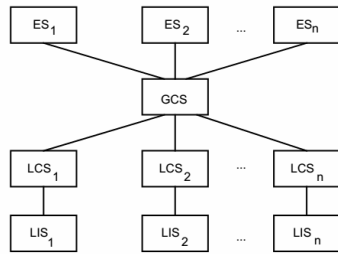


Универзалност ANSI-SPARC архитектуре

- Не зависи од изабраног модела података
- Не зависи од врсте СУБП
- Не зависи ни од нивоа дистрибуирања система
- Иако се у сваке две конкретне имплементације значајно разликују и по питању тумачења одређених појмова и по питању расподеле одговорности по компонентама



Архитектура дистрибуираних СУБП



- Ради илустрације, само да уочимо да је ова архитектура релативно флексибилна...
- Код дистрибуираних СУБП
 - и даље имамо више спољашњих схема
 - концептуална схема се дели на два нивоа:
 - *Глобална* концептуална схема је јединствена и централизована
 - чак не мора да постоји код федеративних БП
 - сваки чвор има своју *локалну* концептуалну схему
 - сваки чвор има своју *локалну* интерну схему

[PM13] Увод у РБП Саша Малков



Тема 3 Модели база података



Улога модела података

- Модел података представља релативно апстрактан стандардизован поглед на проблем организовања, пројектовања, имплементирања и употребе података



Модели података (1)

- Начин моделирања података се временом мења
- Моделирање је прошло кроз различите фазе
- Основна идеја је да се при моделирању домена подизањем нивоа апстракције остварује што је могуће “универзалније” решење



Модели података (2)

- Уочљива је тенденција да се (понекад неоправдано?) враћамо на старе идеје
 - упрошћавање и оптимизовање општих модела за неке специфичне случајеве
 - сужавање домена примене
 - конкретизација
- То се обично ради ради подизања ефикасности...
 - ...али има за последицу смањивање флексибилности система



Модели података (3)

- Употреба мање општих модела може да буде корисна...
 - ...када нас општија решења ограничавају
 - обично у погледу перформанси
- Али то често може да буде корак назад...
 - ...ако не знамо где смо били и зашто тамо нисмо остали...
 - “откривање рупе на саксији”?
 - ...може да нам се учини као добра идеја и када није...
 - ако видимо допринос а не разумемо добро цену



Модел података и архитектура

- Независно од архитектуре СУБП, модели података одређују како корисник види податке и како их употребљава
- Највеће перформансе се остварују када се на свим нивоима архитектуре користи исти модел података
- Често се на нивоу спољашње схеме (или у неким деловима спољашње схеме) користи другачији модел, ради прилагођавања корисницима
 - на пример, данас се често релационим базама података приступа помоћу библиотека које пружају програмеру објектни модел података
 - тада имамо тзв. објектно-релационо пресликавање



Најважнији модели података

- Мрежни модел
- Хијерархијски модел
- Релациони модел
- Модел ентитета и односа
- Објектно-релациони модел
- Објектни „модел“
- Модел не-релационих база података



Мрежни модел података

- *Integrated Data Store (IDS)*
 - “први” СУБП
 - 1960-те
 - *General Electric*
 - Чарлс Бекман (*Charles Bachman*)
 - 1973. добио Тјурингову награду за рад у области БП
- *Associate PL/I*
 - крај 1960-их
- *CODASYL*
 - група за стандардизацију
 - *CODASYL DBTG*, 1971.
 - иста група је раније стандардизовала *COBOL*



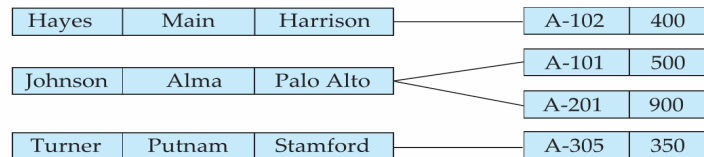
Мрежни модел података, концепти

- Налик на дијаграмско повезивање података
 - структуре података
 - налик на слоге у програмским језицима
 - слог садржи податке једне инстанце ентитета
 - састоји се од поља (налик атрибутима)
 - структуре се повезују „везама“ (*link*)
 - налик на показиваче у програмским језицима
 - слично повезивању елемената дијаграма стрелицама

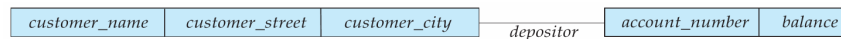


Мрежни модел података, пример

Пример података:



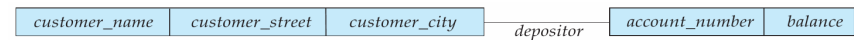
Пример модела:



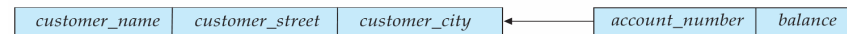
Повезујемо клијента и банковни рачун (нотација ЕР, касније детаљније):



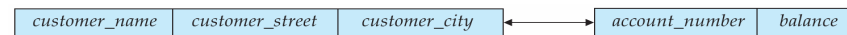
Однос више-више у мрежном моделу:



Однос један-више у мрежном моделу:

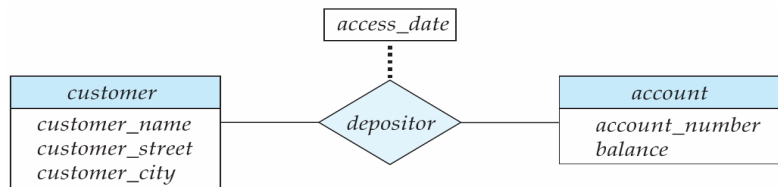


Однос један-један у мрежном моделу:

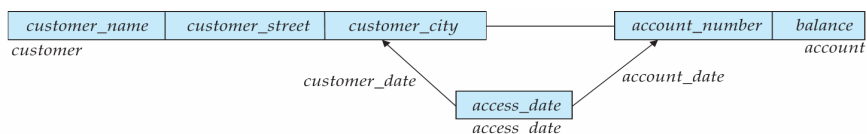


Мрежни модел података – Односи са описним атрибутима

Допуњавамо последњим датумом коришћења
када је конкретан корисник користио конкретан рачун:

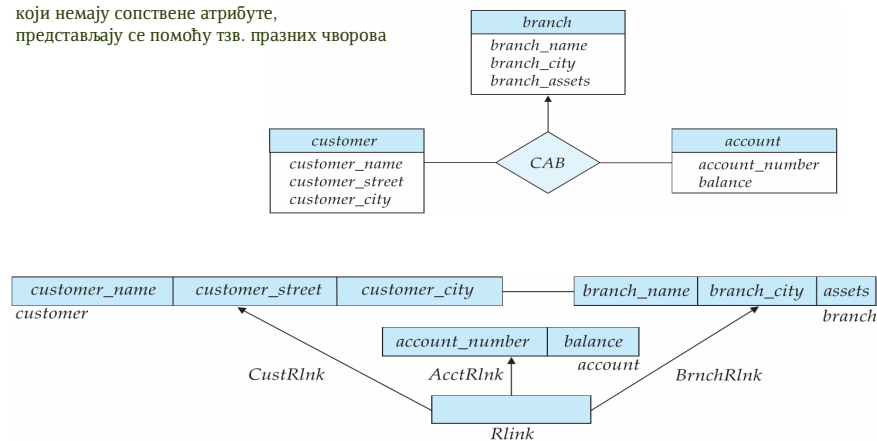


Однос са атрибутима у мрежном моделу:



Мрежни модел података – Односи са описним атрибутима

Сложени односи са више елемената,
који немају сопствене атрибуте,
представљају се помоћу тзв. празних чворова



Мрежни модел података

Мрежни модел података, стандардизација



- Разноликост веза прави проблем при употреби
- Зато стандард *CODASYL (DBTG 1971)* прописује само усмерене везе више-један
 - везе више-више се имплементирају као пар веза више-један
- *Silberschatz, Korth, Sudarshan, Database System Concepts, 6th ed. 2010*
 - (шема за семинарски раг)

Хијерархијски модел података

Хијерархијски модел података



- *Information Management System (IMS)*
 - 1968.
 - *IBM*
 - Иницијално развијен као СУБП за свемирски програм Аполо
- *System 2000*
 - *MRI*

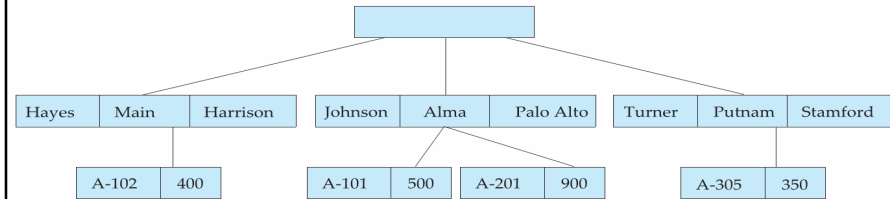


Хијерархијски модел података, концепти

- Скуп „слогова“ повезаних „везама“ тако да граде „хијерархију“
 - у основи слично мрежном моделу
 - али захтева се хијерархија
 - код мрежног модела је хијерархија била имплицирана смером веза (од један према више) а овде је експлицитна и има смер један-више
 - није допуштено вишеструко везивање чворова
 - до сваког чвора води тачно један пут од корена
 - уместо редувантности веза имамо редувантност података
 - између свака два повезана чвора постоји тачно један пут
 - повећана редувантност података
- Скуп слогова у колекцији започиње „празним“ (*dummy*) чвором

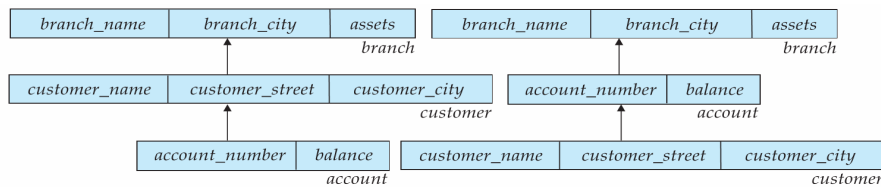


Пример података у хијерархијском моделу

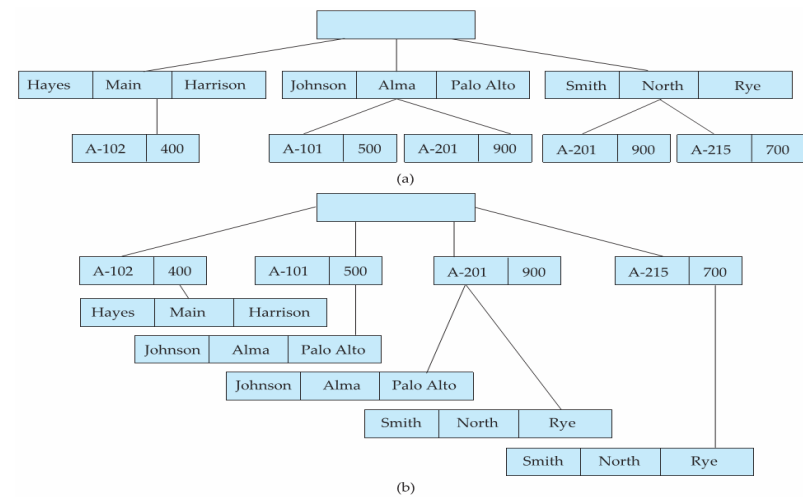


Сложени односи захтевају више хијерархија

- Сложени односи захтевају више хијерархија
 - И различити начини приступања захтевају више хијерархија
- Последица је увећана редувантност података



Silberschatz, Korth, Sudarshan, *Database System Concepts*, 6th ed. 2010
(тема за семинарски рад)





Хијерархијски модел података, резиме

- Обично се прави једна хијерархија података као “главна”
- Помоћне хијерархије се праве са циљем омогућавања ефикаснијег приступа
 - имају сличну улогу као индекси у релационим базама података
- Неке савремене нерелационе базе података почивају на веома сличним претпоставкама и практично представљају надградњу хијерархијског модела података
 - на пример *Apache Cassandra*



Релациони модел података

- *Edgar Codd*, 1970
- Обједињено моделирање
 - и ентитети и односи се моделирају *релацијом*
 - релација се састоји од *атрибути* који имају *имена и домене*
 - скуп имена и домена атрибута једне релације представља *схему релације*
 - *шорка* је скуп именованих вредности
 - торка која има исти број вредности, њихове називе и домене као атрибути једне релације представља *инстанци домена (схеме) релације*
 - вредност (садржај) релације је скуп инстанци њеног домена
 - ...
- Интегритет се обезбеђује *кључевима*
- ...



Релациони модел података, помак

- Суштински помак који РМ остварује у односу на мрежни и хијерархијски је удаљавање од семантике програма и програмских језика, тј. од проблема имплементације
- Мрежни и хијерархијски модел раде са *чворовима и везама*
 - концепт чвора скоро потпуно одговара концепту променљиве
 - концепт везе скоро потпуно одговара концепту показивача
 - у бази података повезујемо чворове, као *места где се записују подаци*, а не *подашке*
- Релациони модел потпуно занемарује семантику имплементације и бави се непосредно садржајем



Релациони модел података

- Релациони модел података је једна од најважнијих тема курса
- Биће детаљно представљен



Модел ентитета и односа

- *Peter Chen, 1976*
- Иницијално је замишљен као алтернатива постојећим моделима
 - па чак и релационом моделу
- Није заживео као посебан модел података у области имплементације база података, али јесте као алат за моделирање база података
 - ...упознаћемо га нешто детаљније...



Објектно-релациони модел

- Проширење релационог модела
 - сложени типови података
 - кориснички дефинисани
 - дефинише се и понашање, у форми класа и метода
 - хијерархије типова (класа)
 - постојале су и у ЕР, али овде добијају већи значај, због понашања
 - прошириви типови



Објектни модел података

- Објектне базе података нису успеле да понуде конзистентан “објектни модел” – практично сваки ОО СУБП има сопствени модел
 - већина уводи тзв. *ИД* објекта
 - који има карактеристике “показивача”
 - често се не раздвајају јасно концепти *шија* и *скуја* података
 - чак се понекад и експлицитно изједначавају
- У радовима Дејта и Дарвина се разматра потпуније проширење релационог модела ОО елементима и обликовање одговарајућег упитног језика
 - (тема за семинарски рад)



Нерелационе базе

- Различити модели
 - ниска структурираност
 - већа флексибилност
 - често је цена виша редундантност
 - високе перформансе
 - не увек на пуном опсегу примене
 - једноставније и ефикасније дистрибуирање
 - сужени домени примене
 - начелно широки домени, али практично не
 - када се иде ван основних могућности, често се компликује рад или се губи на перформансама



Нерелационе базе (2)

- Најшешћи модели:
 - Каталогске базе
 - Базе садржаја
 - мултимедијалне колекције и сл.
 - Базе за веб
 - JSON
 - Научне базе података
 - вишедимензионе матрице нумеричких података
 - Графовске базе
 - ...и друго...
- Посветићемо нешто више пажње нерелационим базама података у другој половини курса

Литература за тему



- Гордана Павловић-Лажетић, **Увод у релационе базе података, 2. изг. Математички факултет, 1999.**
 - доступно онлајн: <http://poincare.matf.bg.ac.rs/~gordana/urbp-2016.htm>
- Ramakrishnan, Gehrke, **Database Management Systems, 2.ed, 2000.**
- Teorey, Lightstone, Nadeau, Jagadish, **Database Modeling and Design, 5.ed, Elsevier, 2011.**