

Databázové systémy I

Radim Bača

Katedra informatiky, FEI

radim.baca@vsb.cz

dbedu.cs.vsb.cz

- Organizační část
- Proč databázové systémy?
- Historie databázových systémů
- Klíčové koncepty
- Relační datový model
- Úvodní SQL příkazy

Předmět

- Stránky **dbedu.cs.vsb.cz** obsahují v záložce Databázové systémy I všechny materiály potřebné ke studiu předmětu
- Na stránkách předmětu se budou kromě slidů přednášek objevovat i videa přednášek
- Pro přihlášení na stránky je nutné použít váš školní login a heslo

Hodnocení předmětu

- **Dva testy**, které se budou konat přímo na cvičeních
- První test bude zaměřen na psaní SELECT dotazů (40b / 20b minimum)
- Druhý test bude zaměřen na DML a DDL příkazy (20b / 10b minimum)
- U obou testů budou možné maximálně dvě opravy
- Na opravu se kvalifikujete jen v případě, že vyřešíte alespoň jeden příklad
- Při testech bude možné pracovat se všemi materiály
- **Domácí projekt** týkající se návrhu informačních systémů (30b / 15b minimum)
- **Zápočtová písemka** (10b / 5b minimum)

Software

- MS SQL Server 2016
- Máte již vytvořeny účty na serveru
dbsys.cs.vsb.cz\Student
- Došel vám email s heslem
- K připojení je možné využít klienta SQL Server Management Studio (SSMS) nebo DataGrip
- SSMS se dá zdarma stáhnout na MSDN

Proč SQL Server?

Decentní desktopový klient s mnoha možnostmi (SSMS)

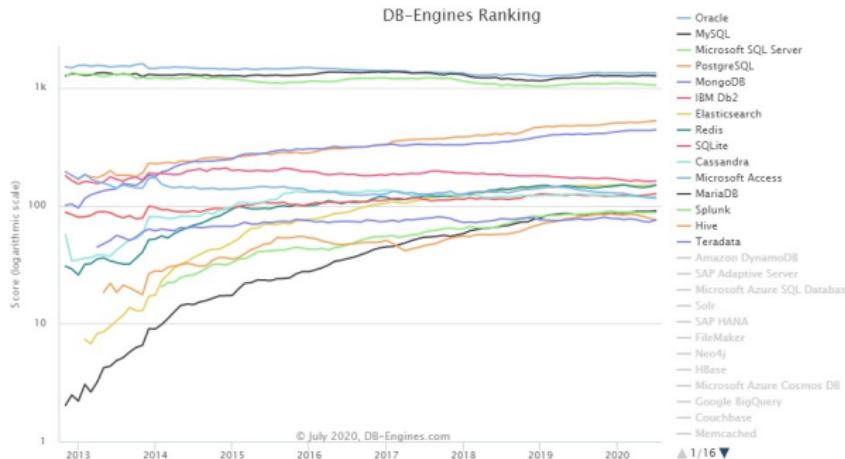


FIGURE: Srovnání relačních SŘBD¹

¹zdroj: http://db-engines.com/en/ranking_trend/relational+dbms

Literatura a další zdroje

- M.Krátký, R.Bača. Databázové systémy, url:
<http://dbedu.cs.vsb.cz> (tohoto předmětu se týkají první dvě části knihy)
- J.Pokorný, M.Valenta. Databázové systémy, ČVUT, ISBN 978-80-01-05212-9, 2013.
- H.Garcia-Molina, J.D.Ullman, J.D.Widom. Database Systems: The Complete Book, Prentice Hall, ISBN 0-13-031995-3, 2002.
- Přednášky budou zveřejňovány na youtube. Odkaz na jednotlivé přednášky nalezenete na stránkách předmětu.

Kdo je kdo



Kdo je kdo



Kdo je kdo



Friendster



Facebook

Proč databázové systémy?

- Databázové systémy jsou doslova všudypřítomné
- Skvělá databáze je taková databáze o které vedoucí projektu ani nemá tušení, že v aplikaci je
- Databázové systémy (nebo-li systémy pro řízení báze dat - SŘBD) jsou:
 - masivní
 - perzistentní
 - bezpečné
 - více-uživatelské
 - pohodlné
 - efektivní
 - spolehlivé

Proč databázové systémy?

- Databázové systémy jsou doslova všudypřítomné
- Skvělá databáze je taková databáze o které vedoucí projektu ani nemá tušení, že v aplikaci je
- Databázové systémy (nebo-li systémy pro řízení báze dat - SŘBD) jsou:
 - masivní
 - perzistentní
 - bezpečné
 - více-uživatelské
 - pohodlné
 - efektivní
 - spolehlivé

Proč databázové systémy?

- Databázové systémy jsou doslova všudypřítomné
- Skvělá databáze je taková databáze o které vedoucí projektu ani nemá tušení, že v aplikaci je
- Databázové systémy (nebo-li systémy pro řízení báze dat - SŘBD) jsou:
 - masivní
 - perzistentní
 - bezpečné
 - více-uživatelské
 - pohodlné
 - efektivní
 - spolehlivé

Proč databázové systémy?

- Databázové systémy jsou doslova všudypřítomné
- Skvělá databáze je taková databáze o které vedoucí projektu ani nemá tušení, že v aplikaci je
- Databázové systémy (nebo-li systémy pro řízení báze dat - SŘBD) jsou:
 - masivní
 - perzistentní
 - bezpečné
 - více-uživatelské
 - pohodlné
 - efektivní
 - spolehlivé

Proč databázové systémy?

- Databázové systémy jsou doslova všudypřítomné
- Skvělá databáze je taková databáze o které vedoucí projektu ani nemá tušení, že v aplikaci je
- Databázové systémy (nebo-li systémy pro řízení báze dat - SŘBD) jsou:
 - masivní
 - perzistentní
 - bezpečné
 - více-uživatelské
 - pohodlné
 - efektivní
 - spolehlivé

Proč databázové systémy?

- Databázové systémy jsou doslova všudypřítomné
- Skvělá databáze je taková databáze o které vedoucí projektu ani nemá tušení, že v aplikaci je
- Databázové systémy (nebo-li systémy pro řízení báze dat - SŘBD) jsou:
 - masivní
 - perzistentní
 - bezpečné
 - více-uživatelské
 - pohodlné
 - efektivní
 - spolehlivé

Proč databázové systémy?

- Databázové systémy jsou doslova všudypřítomné
- Skvělá databáze je taková databáze o které vedoucí projektu ani nemá tušení, že v aplikaci je
- Databázové systémy (nebo-li systémy pro řízení báze dat - SŘBD) jsou:
 - masivní
 - perzistentní
 - bezpečné
 - více-uživatelské
 - pohodlné
 - efektivní
 - spolehlivé

Proč databázové systémy?

- Databázové systémy jsou doslova všudypřítomné
- Skvělá databáze je taková databáze o které vedoucí projektu ani nemá tušení, že v aplikaci je
- Databázové systémy (nebo-li systémy pro řízení báze dat - SŘBD) jsou:
 - masivní
 - perzistentní
 - bezpečné
 - více-uživatelské
 - pohodlné
 - efektivní
 - spolehlivé

Proč databázové systémy?

- Databázové systémy jsou doslova všudypřítomné
- Skvělá databáze je taková databáze o které vedoucí projektu ani nemá tušení, že v aplikaci je
- Databázové systémy (nebo-li systémy pro řízení báze dat - SŘBD) jsou:
 - masivní
 - perzistentní
 - bezpečné
 - více-uživatelské
 - pohodlné
 - efektivní
 - spolehlivé

Citát

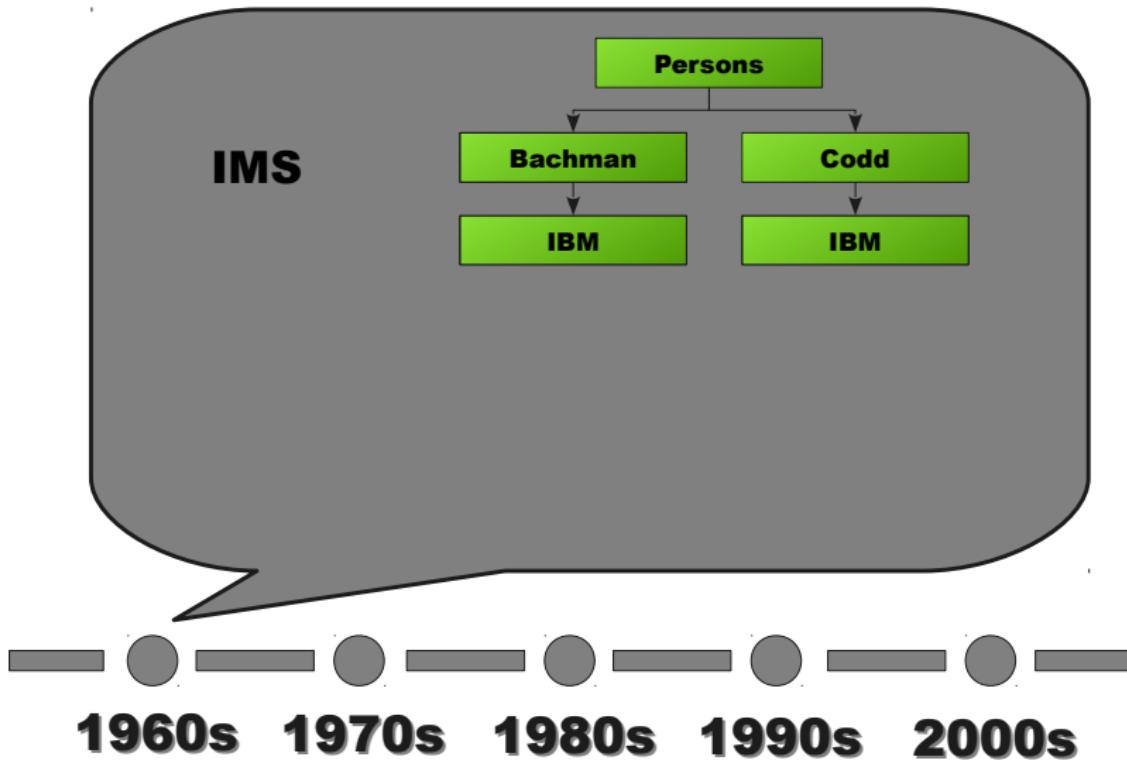
"Those who cannot learn from history are doomed to repeat it."

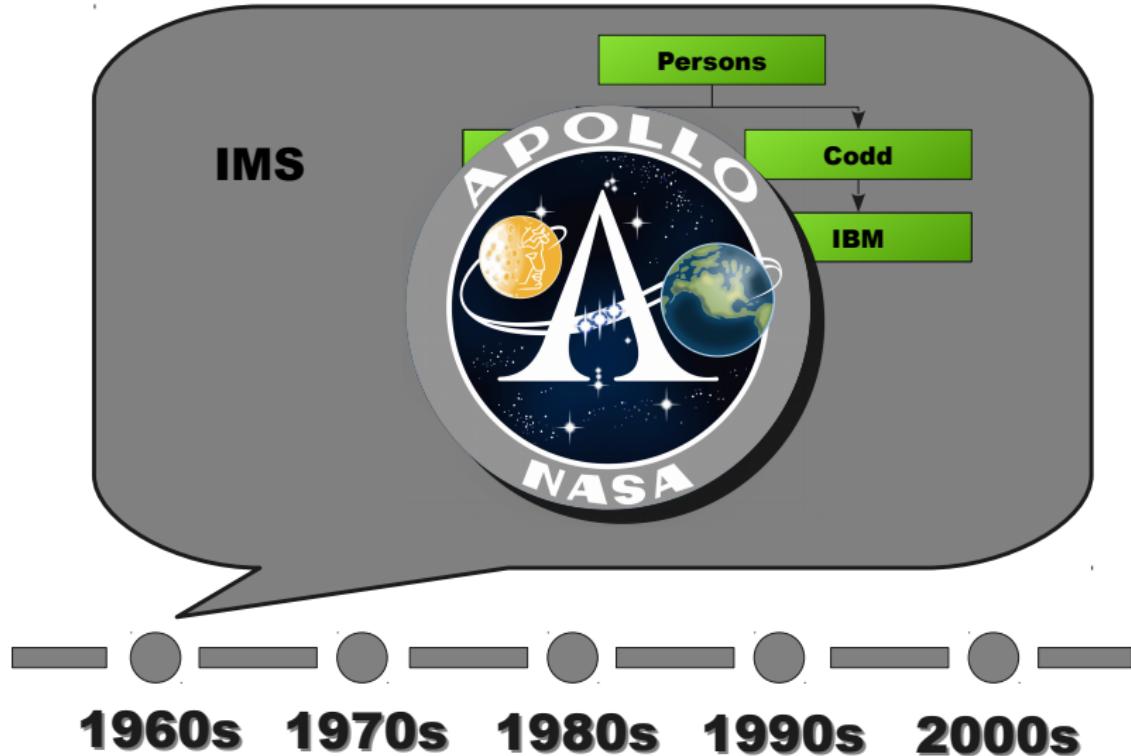
George Santayana

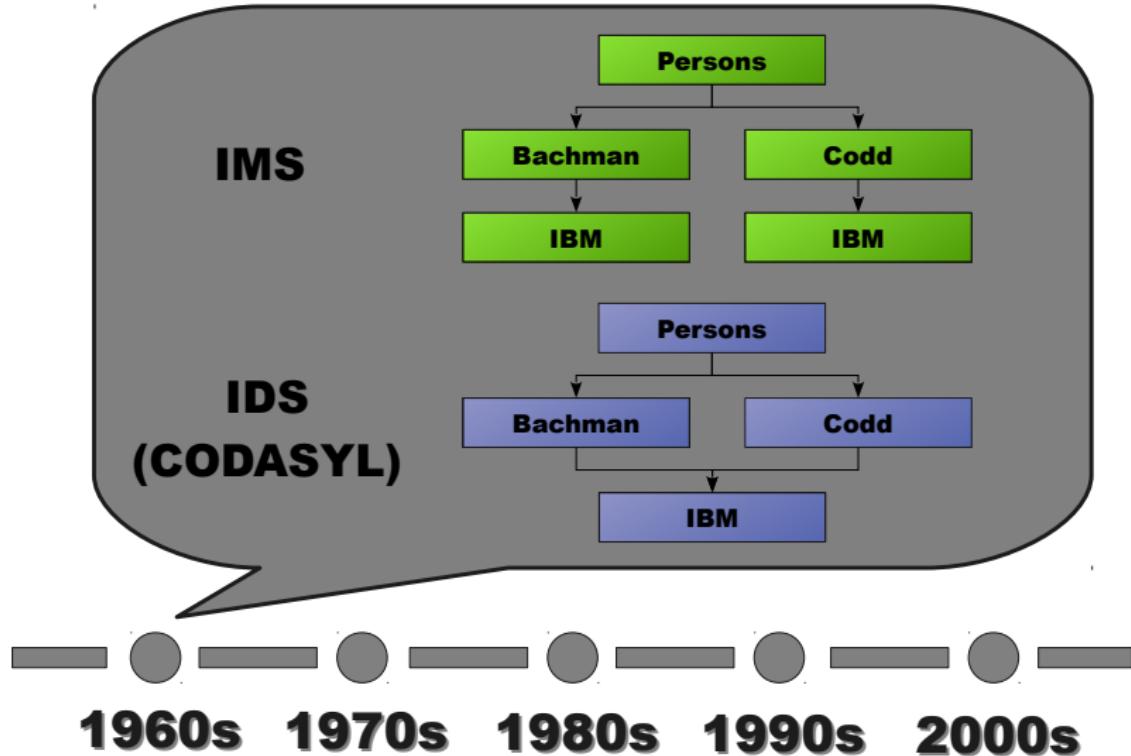


1960s 1970s 1980s 1990s 2000s









Relational data model

Person

id	name	aff.
1	Bachman	1
2	Codd	1
3	Stonebreaker	2
4	Chamberlin	1
5	Ellison	3
6	Lomet	4

Affiliation

id	name
1	IBM
2	Berkeley
3	Oracle
4	Microsoft



1960s 1970s 1980s 1990s 2000s

Relational data model

- Ingres
- System R

Person

id	name	aff.
1	Bachman	1
2	Codd	1
3	Stonebreaker	2
4	Chamberlin	1
5	Ellison	3
6	Lomet	4

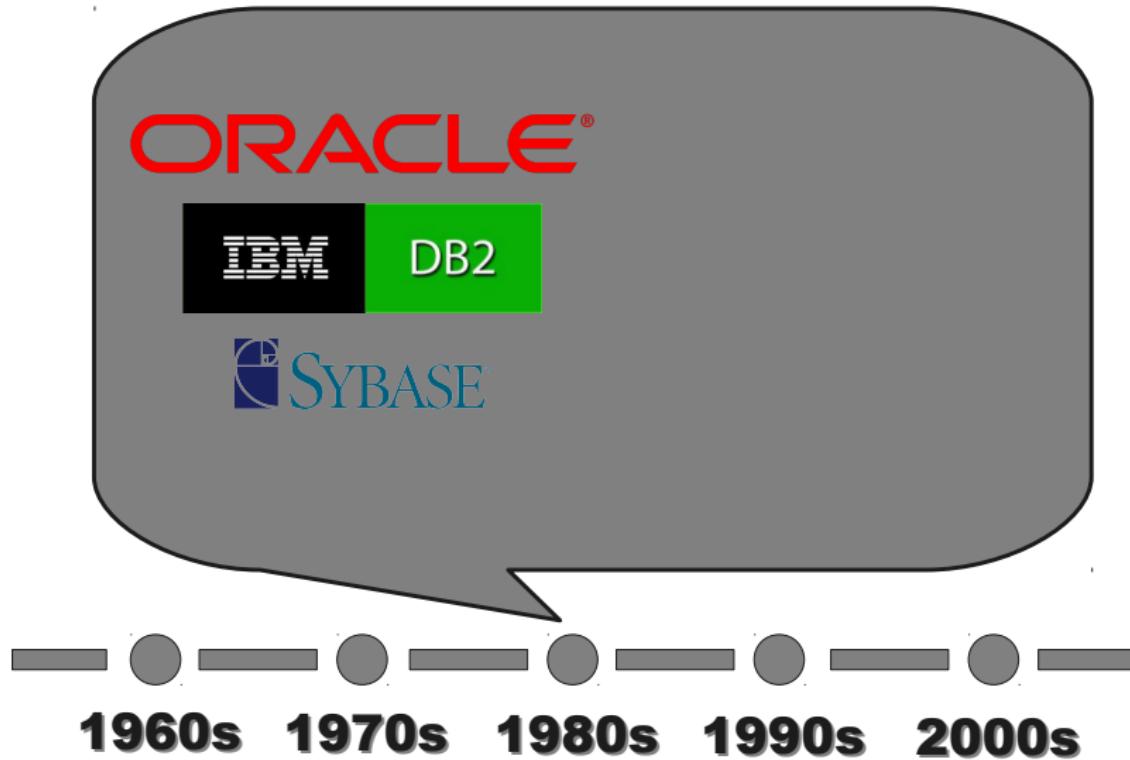
Affiliation

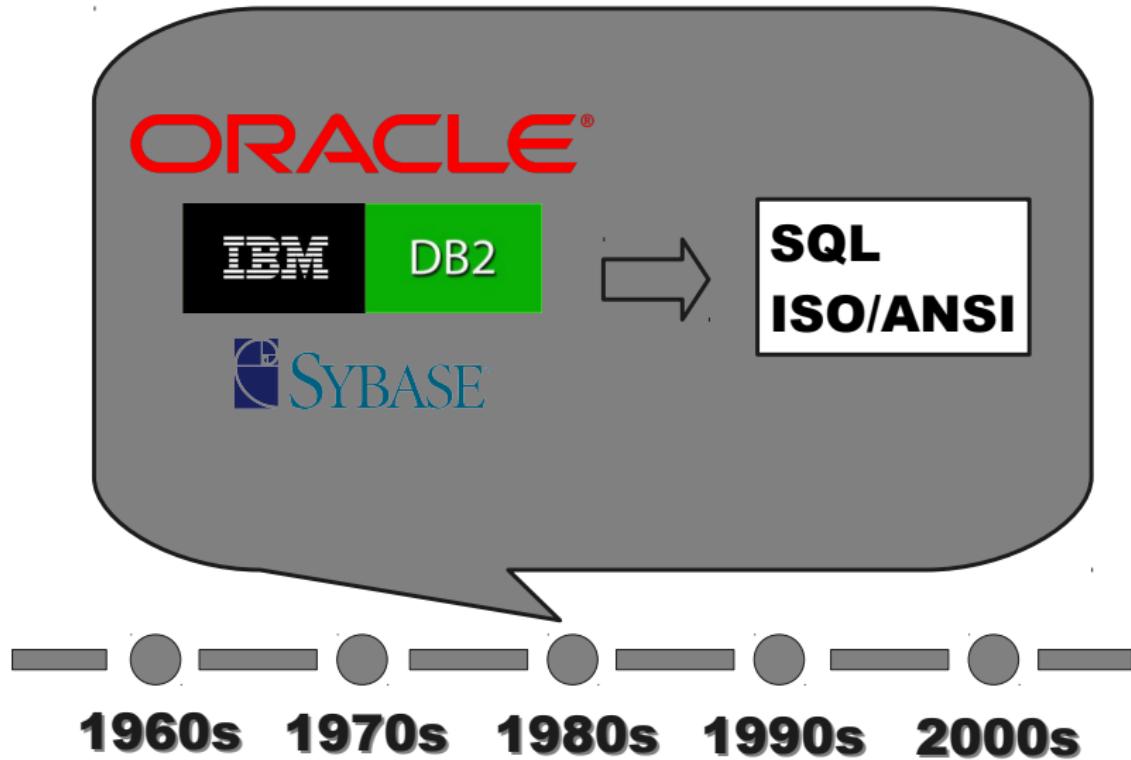
id	name
1	IBM
2	Berkeley
3	Oracle
4	Microsoft

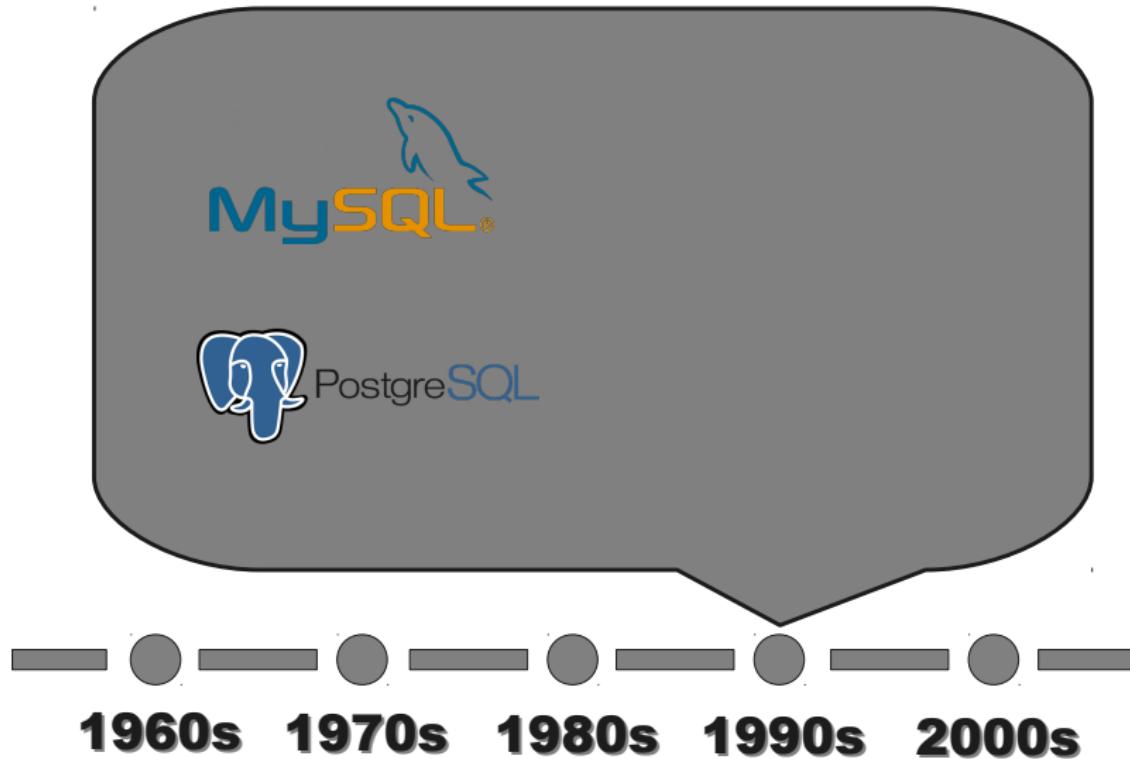


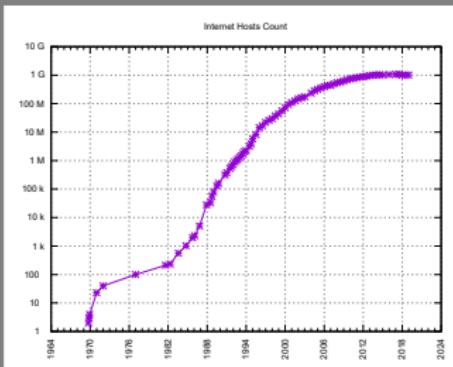
1960s 1970s 1980s 1990s 2000s



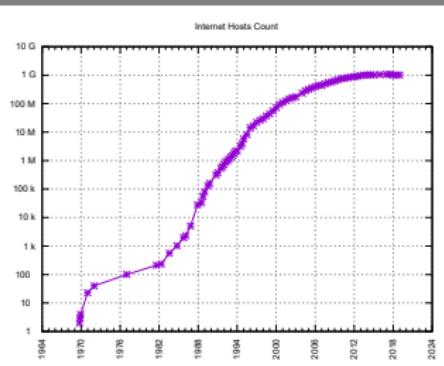








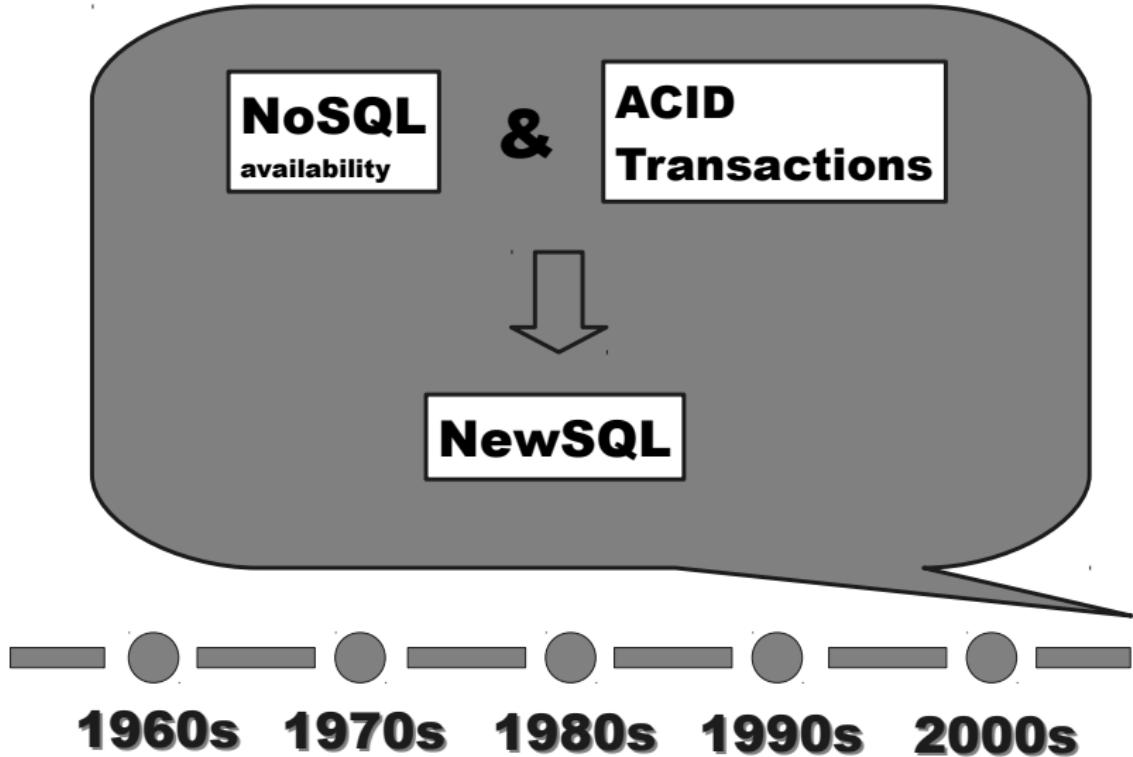
1960s 1970s 1980s 1990s 2000s



NoSQL
Not only SQL



1960s 1970s 1980s 1990s 2000s



Databázové systémy ve světě



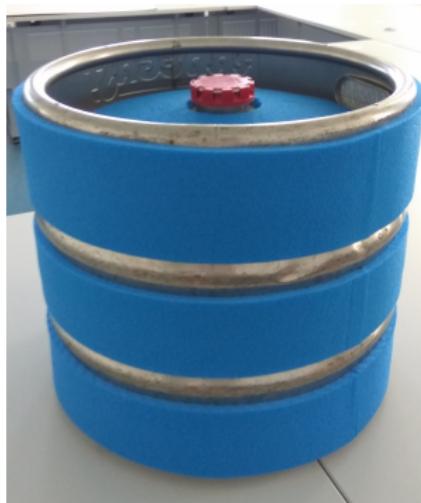
ORACLE



SAP HANA



Nové typy databázových systémů

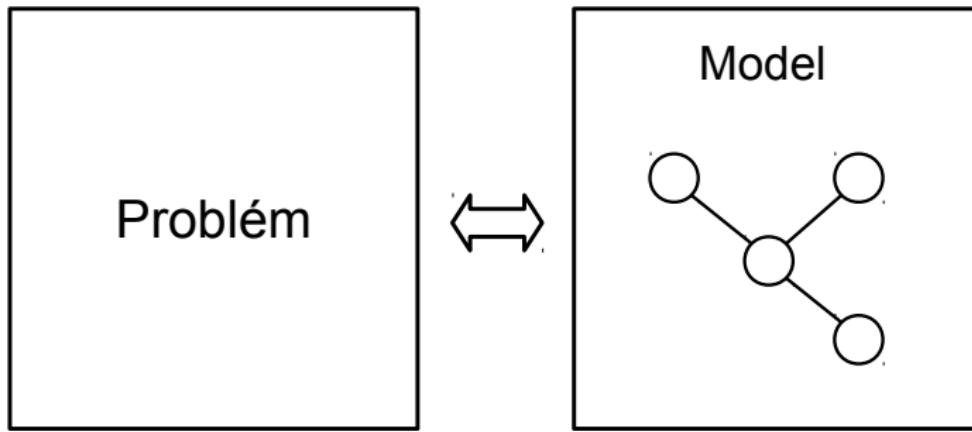


Databáze databázových systémů <https://dbdb.io>

Pojmy

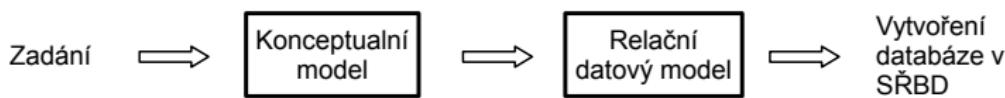
- **Databázový systém / Systém pro řízení báze dat (SŘBD)** - jedná se o aplikaci, která nabízí rozhraní pro vytvoření databáze a pro komunikaci s databází. Má většinu ze sedmi atributů definovaných dříve.
- **Databáze** - jedná se o množinu vzájemně propojených dat, které jsou uloženy v nějakém databázovém systému.

Model



Model

- Typické modely vytvořené při analýze:
 - **Konceptuální model/Logický model** - logický popis struktury databáze, který nezohledňuje použitý databázový systém
 - **Databázové schéma/Datový model** (např. relační datový model) - popis databáze definované pro konkrétní databázový systém a pro konkrétní datový model
- Postup při vytváření databáze pak může probíhat následovně:



Datový model

- Data mohou být uspořádána (strukturována) mnoha různými způsoby:
 - Relační model (množina záznamů)
 - Semistrukturovaný model (XML, JSON, BSON)
 - Graf (RDF)
 - Nestrukturovaná data

Datový model

- Data mohou být uspořádána (strukturována) mnoha různými způsoby:
 - Relační model (množina záznamů)
 - Semistrukturovaný model (XML, JSON, BSON)
 - Graf (RDF)
 - Nestrukturovaná data

Model

- Obecněji pak postup při vytváření databáze pak může probíhat následovně:

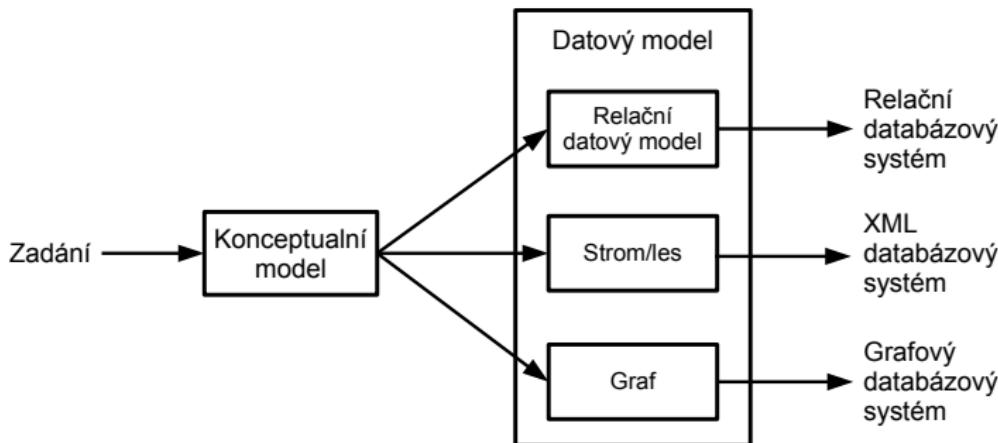


Schéma vs. data

- Schéma nám popisuje strukturu databáze
- Data jsou konkrétní záznamy, které odpovídají této struktuře
- Je to podobné jako
 - třída vs. objekt
 - datový typ vs. proměnná
 - ...

Schéma vs. data

- Schéma nám popisuje strukturu databáze
- Data jsou konkrétní záznamy, které odpovídají této struktuře
- Je to podobné jako
 - třída vs. objekt
 - datový typ vs. proměnná
 - ...

Role lidí věnujících se SŘBD

- Aplikační programátor - programuje aplikační úlohy nad databázemi
- Návrhář databáze - stanoví schéma databáze
- Správce databáze - administrátor databáze a databázového systému; udržuje databázový systém v kondici a dostupnou, přiděluje přístupová práva k datům, a řeší případné výpadky.
- Programátor databázových systémů - často velké týmy lidí

Hrnečku vař

- Dcera dostala kouzelný hrneček od žebračky se kterou se rozdělila o chleba.
- Předvedla hrneček matce, která se následně po dceřině odchodu pustila do neodborné manipulace s hrnečkem.
- Výsledek všichni známe: mnoho sedláků se muselo domů prokousat skrze kaši až je z toho bolelo břicho.
- S databázemi je to podobné: neodborným přístupem k databázím můžeme způsobit bolení hlavy případným kolegům i mnoha uživatelům našich informačních systémů.

Hrnečku vař

- Dcera dostala kouzelný hrneček od žebračky se kterou se rozdělila o chleba.
- Předvedla hrneček matce, která se následně po dceřině odchodu pustila do neodborné manipulace s hrnečkem.
- Výsledek všichni známe: mnoho sedláčů se muselo domů prokousat skrze kaší až je z toho bolelo břicho.
- S databázemi je to podobné: neodborným přístupem k databázím můžeme způsobit bolení hlavy případným kolegům i mnoha uživatelům našich informačních systémů.

Hrnečku vař

- Dcera dostala kouzelný hrneček od žebračky se kterou se rozdělila o chleba.
- Předvedla hrneček matce, která se následně po dceřině odchodu pustila do neodborné manipulace s hrnečkem.
- Výsledek všichni známe: mnoho sedláků se muselo domů prokousat skrze kaši až je z toho bolelo břicho.
- S databázemi je to podobné: neodborným přístupem k databázím můžeme způsobit bolení hlavy případným kolegům i mnoha uživatelům našich informačních systémů.

Hrnečku vař

- Dcera dostala kouzelný hrneček od žebračky se kterou se rozdělila o chleba.
- Předvedla hrneček matce, která se následně po dceřině odchodu pustila do neodborné manipulace s hrnečkem.
- Výsledek všichni známe: mnoho sedláků se muselo domů prokousat skrze kaši až je z toho bolelo břicho.
- S databázemi je to podobné: neodborným přístupem k databázím můžeme způsobit bolení hlavy případným kolegům i mnoha uživatelům našich informačních systémů.

Relational data model

Person

id	name	aff.
1	Bachman	1
2	Codd	1
3	Stonebreaker	2
4	Chamberlin	1
5	Ellison	3
6	Lomet	4

Affiliation

id	name
1	IBM
2	Berkeley
3	Oracle
4	Microsoft



1960s 1970s 1980s 1990s 2000s

Části relačního datového modelu

- **Struktura** - definice relací a jejich obsahu
- **Integrita** - obsah databáze musí splňovat podmínky
- **Operace** - jak můžeme přistupovat a manipulovat s daty

Relace

- **Relace** = tabulka

Relace

Zaměstnanec (ID, jméno, RČ, funkce)

ID	jméno	rodné číslo	funkce
223	Jan Svěží	7905051111	CEO
124	Valerie První	6901112233	důležitá
154	Pavel Rubal	7105029876	dříč

- **Relace**
- **Schéma relace** - popisuje jméno a strukturu relace

Relace

Zaměstnanec (ID, jméno, RČ, funkce)

ID	jméno	rodné číslo	funkce
223	Jan Svěží	7905051111	CEO
124	Valerie První	6901112233	důležitá
154	Pavel Rubal	7105029876	dříč

- **Relace**
- **Schéma relace**
- **N-tice** - řádek (záznam) tabulky
- **Atribut a jeho doména** - sloupec tabulky a množina možných hodnot

Databáze

- **(Relační) databáze** je tedy množina relací
- **Schéma (relační) databáze** je množina relačních schémat

Integritní omezení (IO)

- Jedná se o podmínky, které definují požadované vlastnosti jednotlivých atributů
- Příklady IO:
 - *Může nebo nemůže atribut nabývat hodnoty NULL?*
 - *Splňuje hodnota datový typ a požadovaný rozsah atributu?*
- **Konzistentní databáze = všechny relace splňují IO**

Primární klíč

Zaměstnanec			
ID	jméno	rodné číslo	funkce
223	Novák	7905051111	nicka
124	Kachlička	6901112233	důležitá
154	Trier	7105029876	nicka

- **Primární klíč** - jednoznačně identifikuje řádek tabulky
- Může jít o jeden atribut

Primární klíč

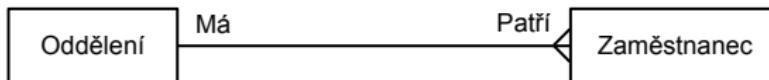
Oddělení

pobočka	jméno	ulice	obrat
22	export	Slavníkovců	15
22	import	Slavníkovců	16
26	export	Mojmírovčů	2

- **Primární klíč** - jednoznačně identifikuje řádek tabulky
- Může jít o jeden atribut, ale může být tvořen i více atributy

Vazby

- Nyní máme dvě relace Zaměstnanec a Oddělení
- Budeme chtít reprezentovat vztah, kdy zaměstnanec patří právě do jednoho oddělení
- Jinými slovy budeme chtít relačním modelem reprezentovat následující konceptuální model:



Cizí klíč

Zaměstnanec					
ID	jméno	rodne číslo	funkce	pobočka oddělení	jméno oddělení
223	Novák	7905051111	nicka	22	export
124	Kachlička	6901112233	důležitá	22	export
154	Trier	7105029876	nicka	26	export

- **Cizí klíč** - reprezentuje vazbu mezi relacemi
- Na příkladu je vidět, že cizí klíč se odvíjí od primárního klíče tabulky, na kterou se odkazuje
- Proto v tomto případě bude výhodnější vytvořit umělý primární klíč ID pro oddělení

Cizí klíč

Zaměstnanec

ID	jméno	rodne číslo	funkce	oID
223	Novák	7905051111	nicka	1
124	Kachlička	6901112233	důležitá	1
154	Trier	7105029876	nicka	3

Oddělení

oID	pobočka	jméno	ulice	obrat
1	22	export	Slavníkovců	15
2	22	import	Slavníkovců	16
3	26	export	Mojmírovců	2

Referenční integrita

- Referenční integrita - hodnota cizího klíče v záznamu musí existovat v atributu tabulky na kterou se cizí klíč odkazuje.
- Problém při škálování distribuované databáze
- NoSQL, NewSQL

Dotazování relačních databází

- Dotazy se skládají z elementárních operací nad relacemi, jejichž výsledkem je opět relace
- Příklad dotazu:

Chceme jména zaměstnanců, kteří pracují v odděleních, která mají obrat alespoň 10 milionů

Dotazování relačních databází

- Dotazy se skládají z elementárních operací nad relacemi, jejichž výsledkem je opět relace
- Příklad dotazu:

Chceme jména zaměstnanců, kteří pracují v odděleních, která mají obrat alespoň 10 miliónů

- SQL:

```
SELECT z.jméno FROM Zaměstnanec z, Oddělení o WHERE z.oID = o.oID AND o.obrat > 10
```

Dotazování relačních databází

- Dotazy se skládají z elementárních operací nad relacemi, jejichž výsledkem je opět relace
- Příklad dotazu:

Chceme jména zaměstnanců, kteří pracují v odděleních, která mají obrat alespoň 10 miliónů

- SQL:

SELECT z.jméno FROM Zaměstnanec z, Oddělení o WHERE z.oID = o.oID AND o.obrat > 10

- Relační algebra:

$$\Pi_{zamestnanec.jméno} (\sigma_{obrat>10} (Oddělení \bowtie Zaměstnanec))$$

Dotazování relačních databází

- Dotazy se skládají z elementárních operací nad relacemi, jejichž výsledkem je opět relace
- Příklad dotazu:

Chceme jména zaměstnanců, kteří pracují v odděleních, která mají obrat alespoň 10 miliónů

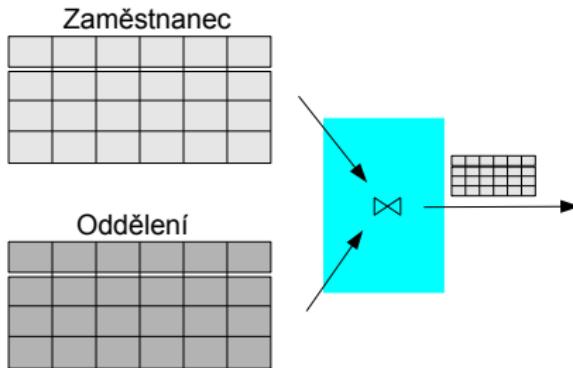
Zaměstnanec

Oddělení

Dotazování relačních databází

- Dotazy se skládají z elementárních operací nad relacemi, jejichž výsledkem je opět relace
- Příklad dotazu:

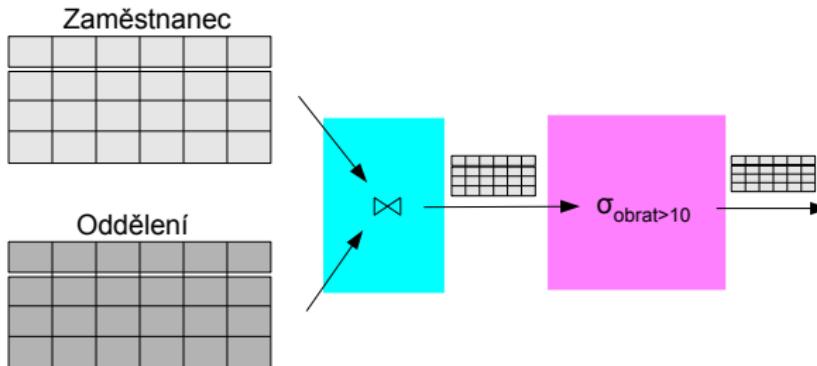
Chceme jména zaměstnanců, kteří pracují v odděleních, která mají obrat alespoň 10 miliónů



Dotazování relačních databází

- Dotazy se skládají z elementárních operací nad relacemi, jejichž výsledkem je opět relace
- Příklad dotazu:

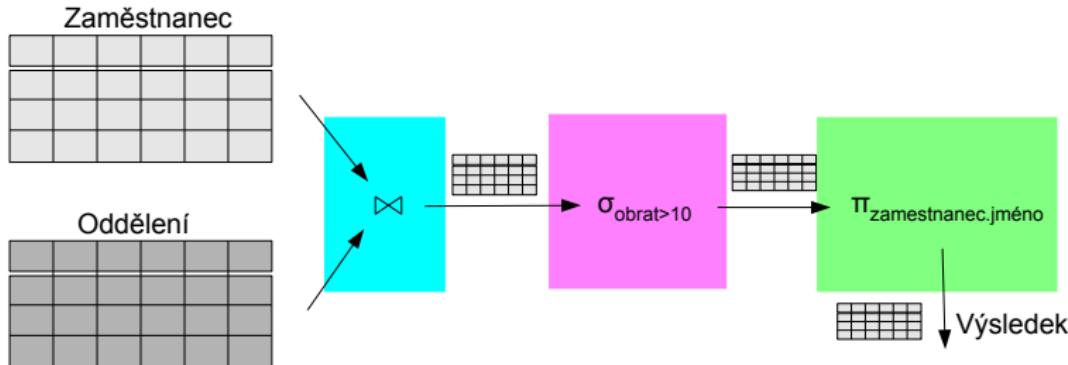
Chceme jména zaměstnanců, kteří pracují v odděleních, která mají obrat alespoň 10 miliónů



Dotazování relačních databází

- Dotazy se skládají z elementárních operací nad relacemi, jejichž výsledkem je opět relace
- Příklad dotazu:

Chceme jména zaměstnanců, kteří pracují v odděleních, která mají obrat alespoň 10 miliónů



Reference a zajímavé odkazy

- Jennifer Widom. Introduction to Databases.
<https://www.coursera.org/course/db>
- Stránky UDBS na <http://dbedu.cs.vsb.cz>
- History of databases
<https://www.youtube.com/watch?v=KG-mqHoXOXY>