

به نام آنکه جان را فکرت آموخت



## بخش اول : مقدمه

مرتضی امینی

نیمسال دوم ۹۱-۹۲

هر سیستم نرم‌افزاری از مجموعه‌ای از داده‌های ذخیره شده استفاده می‌کند.

فرمت ثابت و از پیش تعیین شده دارد. well-formatted است،  
 ساختمند (structured)  
 نیم ساختمند (semi-structured)  
 به لحاظ ساختاری  
 ناساختمند (un-structured)

آیا نیاز به تحمیل یک ساختار در اینها داریم؟ آیا واقعا داده‌ها ساختمند داریم؟

انواع سیستم نرم‌افزاری:

بنیادی یا پایه (سیستم‌های عامل)

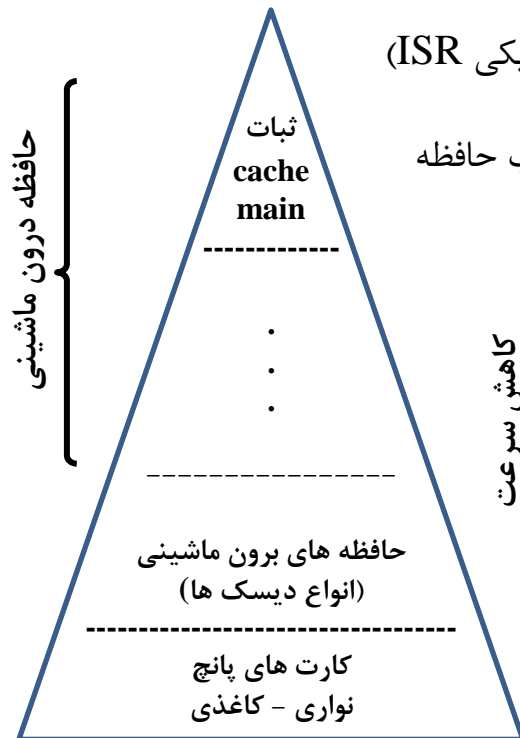
نیمه بنیادی (DBMS، DMS، کامپایلرها، اسمبلرها، و ...)

کاربردی (برنامه‌های کاربردی)

ابزاری : انواع toolها



کاهش قیمت به ازای بیت (CPB)  
 افزایش ظرفیت  
 کاهش سرعت



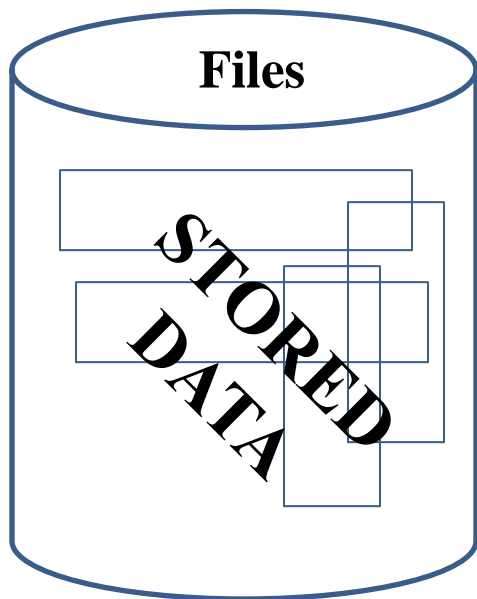


کنجکاوی: دلایل استفاده از این سلسله مراتب حافظه چیست؟

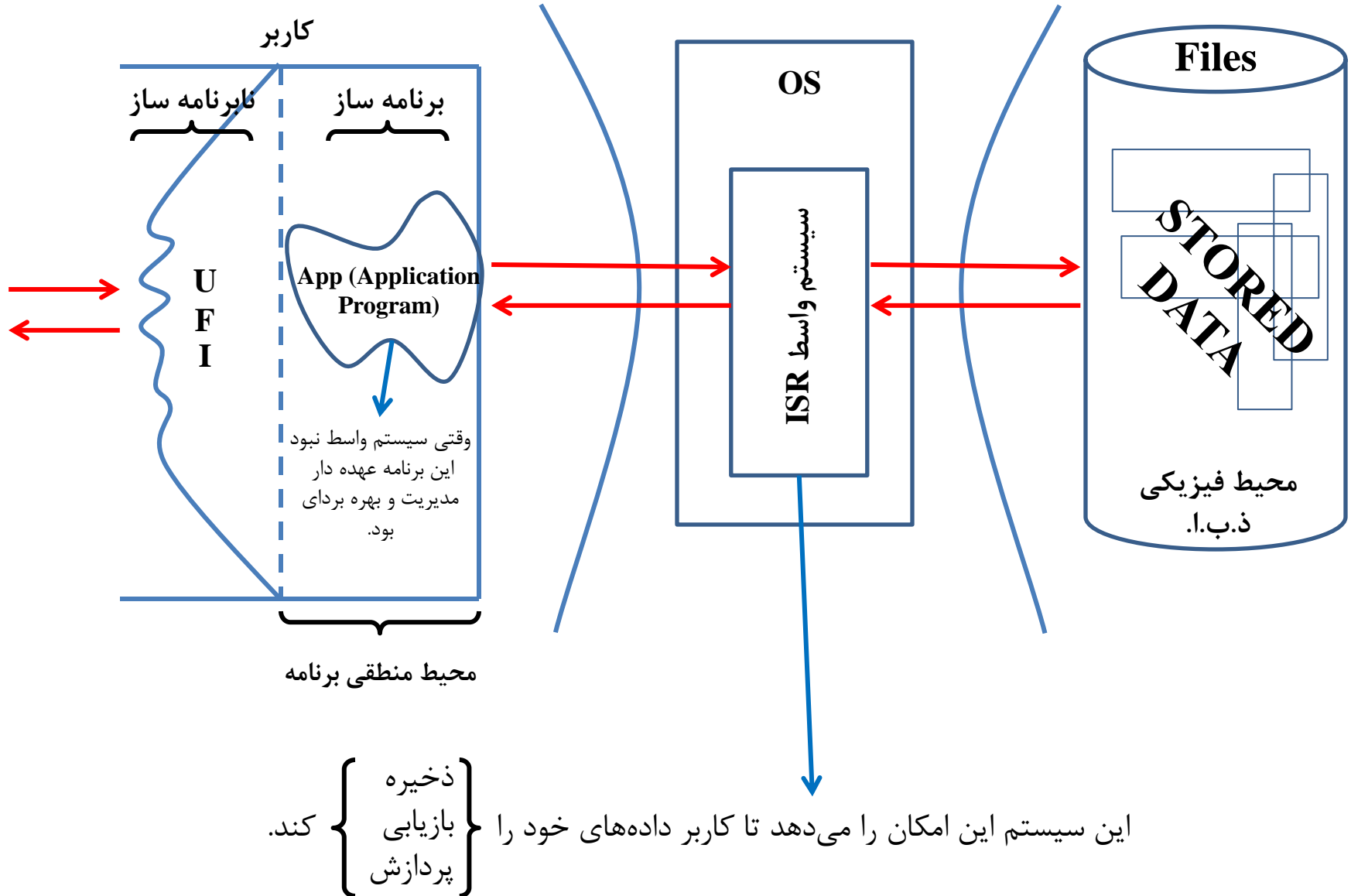
کنجکاوی: چه داده ای، برای چه مدتی، در کدامیک از مراتب سلسله مذکور قرار می گیرد؟

کنجکاوی: خصوصیات عمومی فایل ها چیست؟

□ محیط فیزیکی «ذ.ب.ا» (ذخیره و بازیابی اطلاعات) یا ISR (Information Storage and Retrieval)



ISR باید { ایجاد  
مدیریت  
بهره بردای } شود. ← نیاز به یک سیستم واسط ذ.ب.ا داریم.

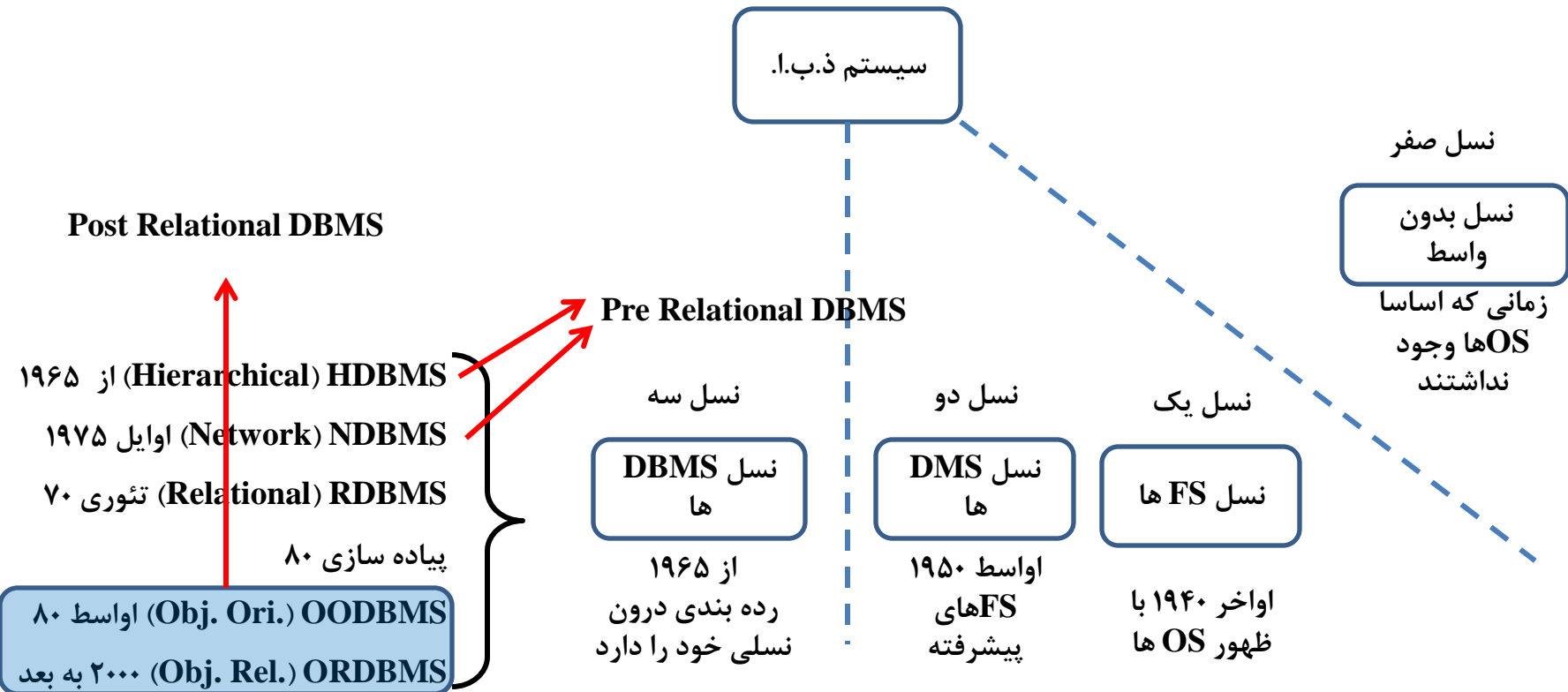




کنجکاوی: رده بندی از مفهوم کاربر ارایه کنید؟ به بیان دیگر گونه های دیگر کاربر کدامند؟

□ سیستم واسط "ISR" سیر تحول خاص خود را دارد :

□ ۶ نسل تکنولوژیک قابل بازیابی است (به طور کلی) [دیدگاه نرم‌افزاری]







- ☐ در این نسل بندی، نسل بعدی نسل قبلی را منسوخ نمی کند. نسل بعدی نسل قبلی را تکمیل می کند و از آن استفاده می کند.
- ☐ انواع نیاز های پردازشی، کنترلی، و عملیاتی سبب ایجاد نسل های سیستم «ذ.ب.ا.» شد.
- ☐ تذکر: نسل های بعد از ششم بر پایه **خرد** (Wisdom) بنا شده اند.



### داده (Data) □

□ **تعریف اول ANSI:** نمایش بوده‌ها، پدیده‌ها، مفاهیم یا شناخته‌ها به طرزی صوری و مناسب برای برقراری ارتباط، تفسیر یا پردازش توسط انسان یا هر امکان خودکار

□ **تعریف دوم ANSI:** هر نمایشی اعم از کاراکتری (نویسه‌ای) یا کمیت‌های قیاسی که معنایی به آن قابل انتساب باشد (توسط انسان یا یک مکانیسم خودکار)

### اطلاع (Information) □

□ **تعریف دقیق و جامعی از مفهوم اطلاع وجود ندارد.**

□ **تعریف ۱ [LIPS92]:** اطلاع، داده پردازش شده است.


□ **تعریف ۲ [روحا ۷۸-الف]:** معنایی که انسان به داده منتسب می‌کند، از طریق قراردادهای شناخته شده‌ای که در نمایش داده به کار می‌روند.

□ برخی داده را همان مقدار واقعا ذخیره شده و اطلاع را معنای آن می‌دانند. بنابراین اطلاع دارای خاصیت اطلاع‌دهندگی و ارتباط‌دهندگی است، در حالیکه داده مجرد این خاصیت را ندارد.






### دانش (Knowledge)

 تعریف [FROS87]: دانش عبارت است از نمایش نمادین جنبه‌هایی از بخشی از جهان واقع (جهان موردنظر یا محیط مطرح)

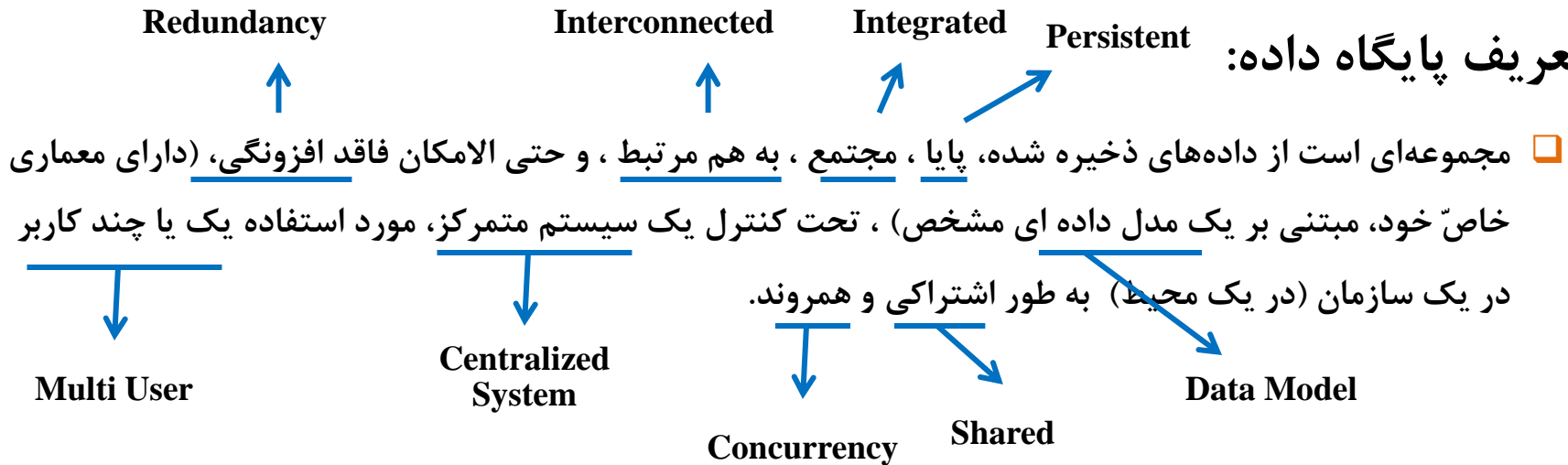
 مثال: شنبه هوا بارانی است. حسن فرزند علی است.

### خرد (Wisdom)

 خرد یا دانایی مفهومی فراتر از دانش است. اینکه چه پردازش(هایی) باید روی مجموعه‌ای از دانش انجام شود تا خرد حاصل شود، به درستی دانسته نیست. عموماً خرد را به نحوه استفاده از دانش مربوط می‌دانند.



### تعریف پایگاه داده:



## □ مثال کاربردی

### □ محیط عملیاتی: دانشگاه



بخشی از جهان واقعی که قصد ایجاد سیستم برای آن را داریم.

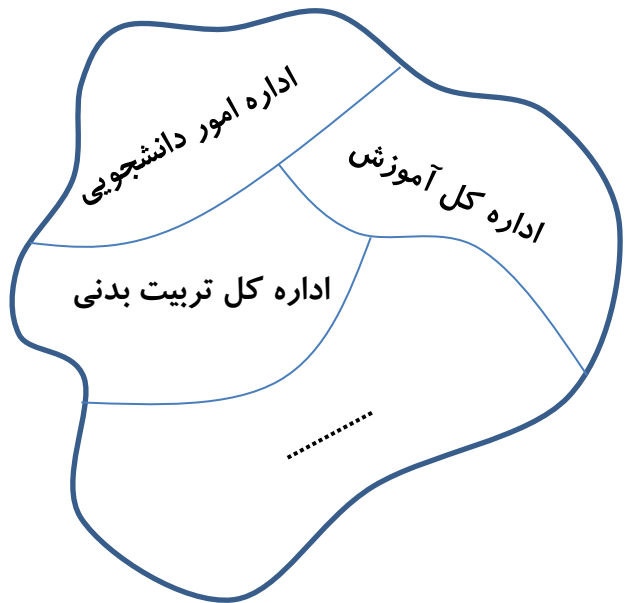


Micro Real World (خرد جهان واقع)  
Mini World  
Universe of Discourse (جهان مطرح)

□ نکته: هر محیط از تعدادی زیر محیط تشکیل شده است.

□ در هر محیط مجموعه ای از **نوع موجودیت‌ها** وجود دارند که نیاز های } داده‌ای کاربران ناظر به آنهاست. پردازشی

(یعنی به داده هایی در مورد آنها نیاز دارند.)





❑ **نکته:** زیرمحیط های یک محیط معمولا با هم اشتراک دارند در نوع موجودیت ها (Entity Type یا Object Type)

❑ مثال : در محیط دانشگاه دانشجو، استاد، درس، کلاس، و ...

❑ مثال : نوع موجودیت دانشجو در هر سه زیر محیط مطرح است.

❑ **مسئله (خواسته) :** ایجاد سیستم(های) کاربردی برای این زیر محیط ها

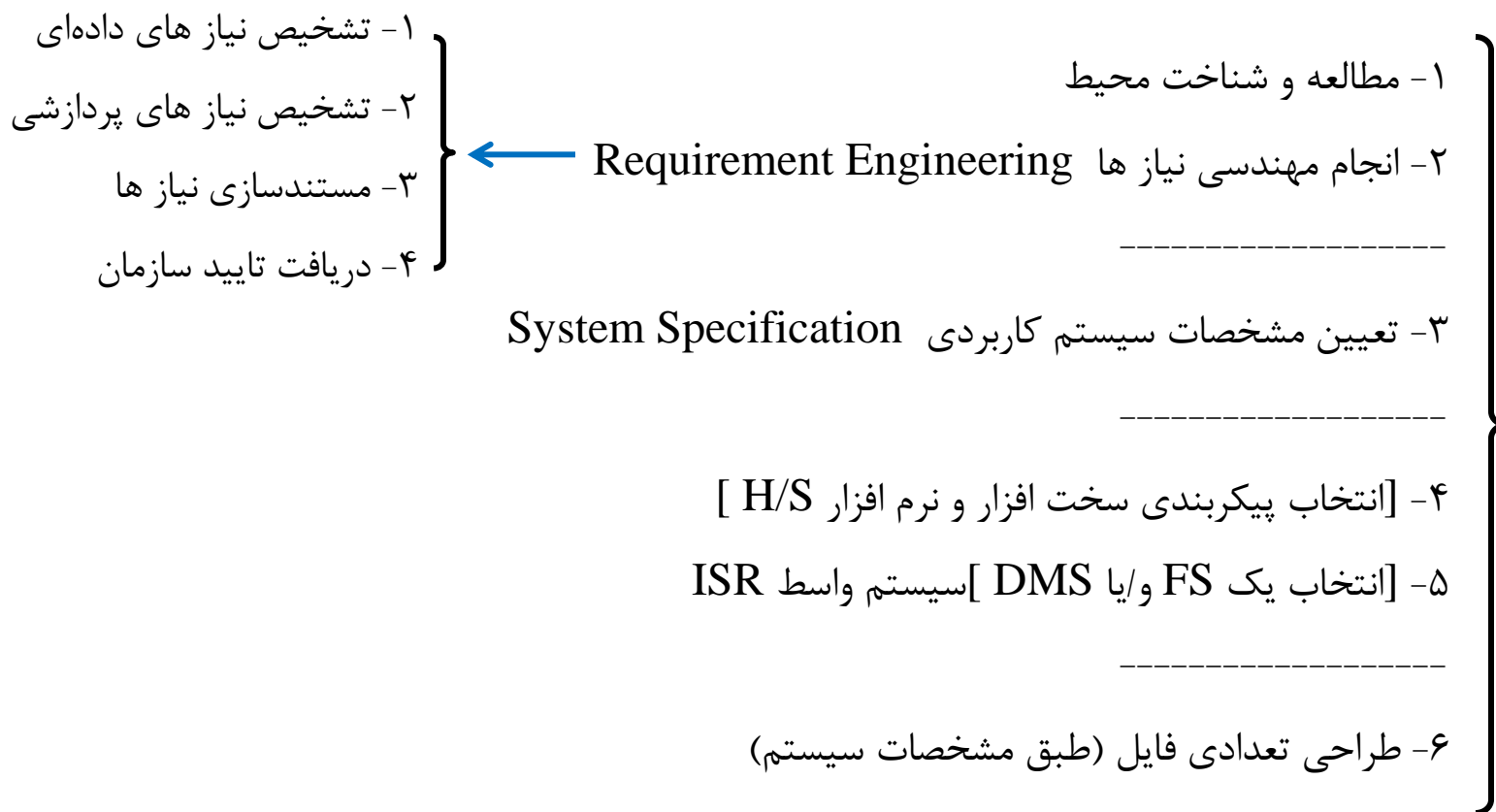
❑ برای این منظور در اساس دو مَشی-روش (approach) وجود دارد.  $\left. \begin{array}{l} \text{مشی فایلینگ [سنتی یا کلاسیک] یا ناپایگاهی} \\ \text{مشی پایگاهی Database Approach} \end{array} \right\}$

یعنی ممکن است مشی های بینابینی نیز وجود داشته باشد.



کارهای لازم در مشی فایلینگ به طور خلاصه:

توجه: این کارها معمولاً برای هر زیرمحیط به طور جداگانه انجام می شود. ← تعدادی سیستم کاربردی جدا (نامجموع) و بی ارتباط در یک محیط ...





۶-۱- تعیین فرمت رکورد

۶-۲- تعیین ساختار فایل

ساختار فایل: ساختاری که براساس آن فقره داده ها (رکوردها) در سطح منطقی [و/یا فیزیکی] با یکدیگر مرتبطند. ساختار فایل یک امکان برای نمایش ارتباط بین فقره داده‌هاست (Data Items) خواه در سطح نمایش منطقی باشد یا فیزیکی.

**کنجکاوی:** چند نوع ساختار فایل وجود دارد؟

۶-۳- نحوه دسترسی به رکوردها - استراتژی دسترسی

۶-۴- اندازه فایل ها

۶-۵- میزان گسترش چه میزان باشد

۶-۶- ارتباط با فایل های دیگر

۶-۷- عملیات مجاز در فایل ها + کاربران

□ کارهای لازم در مشى فایلینگ به طور خلاصه : (ادامه)

۷- طراحی واسطه‌های کاربری (UFI)

۸- طراحی تعدادی برنامه کاربردی (Application Program) [ضمن تعیین تراکنش (ها)]

۹- تولید برنامه‌های { ایجاد  
کنترل  
پردازش } فایل‌ها

۱۰- ایجاد محیط فیزیکی «ذ.ب.ا.» به طور آزمایشی (برای داده‌های تست)

۱۱- ایجاد محیط فیزیکی «ذ.ب.ا.» با داده‌های واقعی اما حجم محدود و انجام تست مرحله دوم

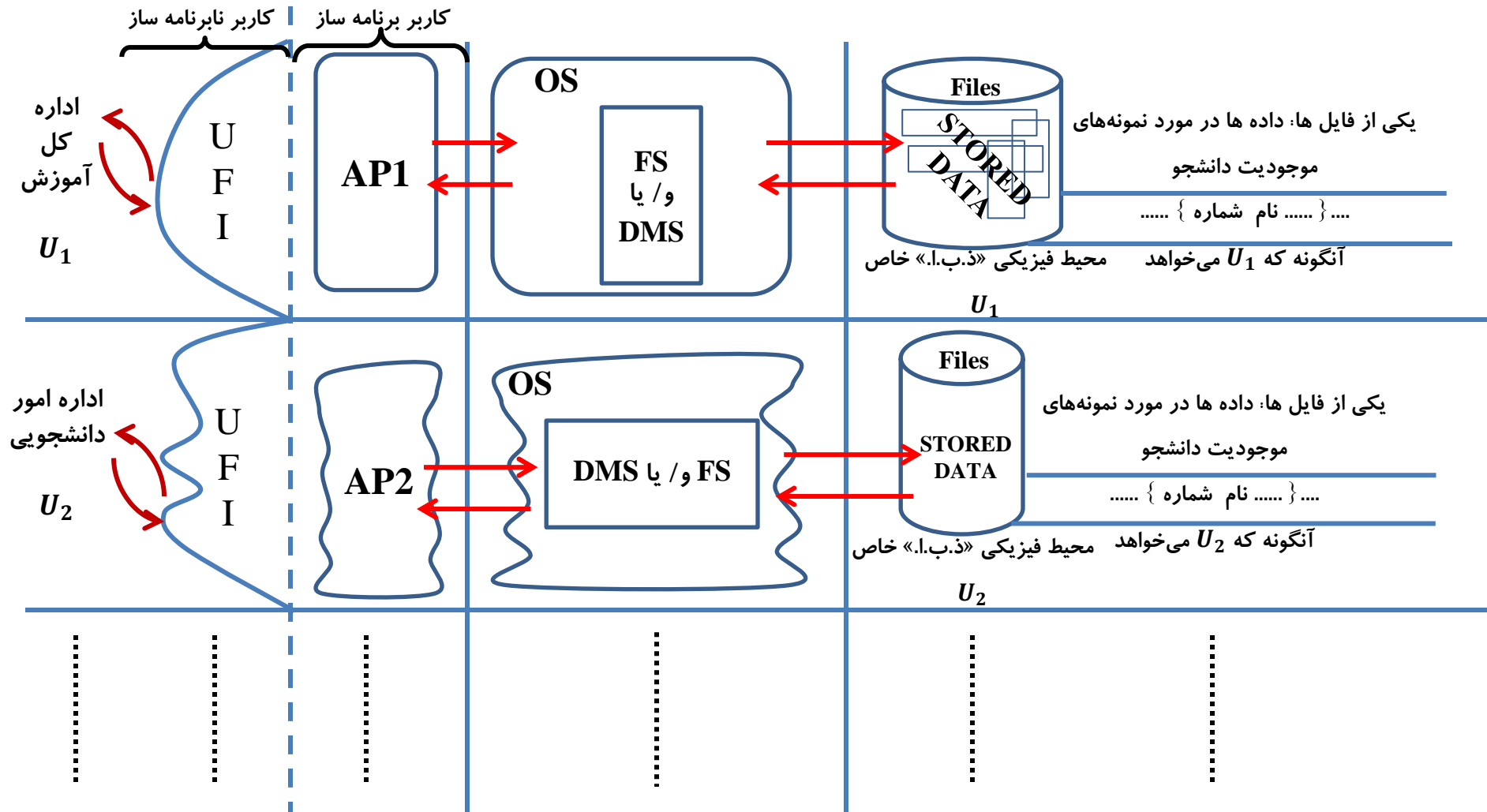
۱۲- ایجاد محیط فیزیکی «ذ.ب.ا.» با داده‌های واقعی و حجم واقعی و انجام تست مرحله سوم

۱۳- رفع اشکال‌ها در هر مرحله

۱۴- ایجاد محیط فیزیکی واقعی با نصب، پیکربندی و ورود داده‌های اولیه (Data Entry)

۱۵- آغاز بهره‌برداری و نگهداری سیستم

۱۶- رفع معایب و بهینه‌سازی سیستم







## ❑ برخی از معایب مشی فایلینگ:

❑ وجود سیستم های نامجتمع در يك سازمان [محيط] و نامرتبط به هم

❑ عدم وجود يك سیستم کنترل متمرکز روی كل داده های سازمان

❑ وجود افزونگی زیاد

❑ خطر بروز ناسازگاری داده ها (Data Inconsistency) ← **کنجکاوی**: جنبه های بروز ناسازگاری کدامند؟

❑ عدم امکان اعمال ضوابط حفظ امنیت داده ها (Data Security)

❑ عدم امکان اشتراکی شدن داده ها (Data Sharing) [یا در حداقل و یا با دشواری]

❑ مصرف نابهینه سخت افزار (به ویژه سخت افزار ذخیره ساز)

❑ وابسته بودن برنامه ها به جنبه های فایلینگ محیط ذخیره سازی، به گونه ای که اگر قرار باشد در فایلینگ

تغییراتی ایجاد شود، برنامه ها هم متناسباً باید تغییر یابد. (به طور مثال فرمت ساختار یا نحوه دسترسی

(Access Strategy) را تغییر دهیم)



□ توضیح مفهوم افزونگی:

□ افزونگی در معنای محدود (یعنی درون فایلی - intrafile redundancy - در مباحث فایلینگ)

■ عبارت است از تکرار ذخیره سازی مقادیر (value) یک صفت یا بیش از یک صفت در فایل داده‌ای یا فایل کمکی آن.

نمایه سازی چیست؟

□ این نوع افزونگی گونه هایی دارد:

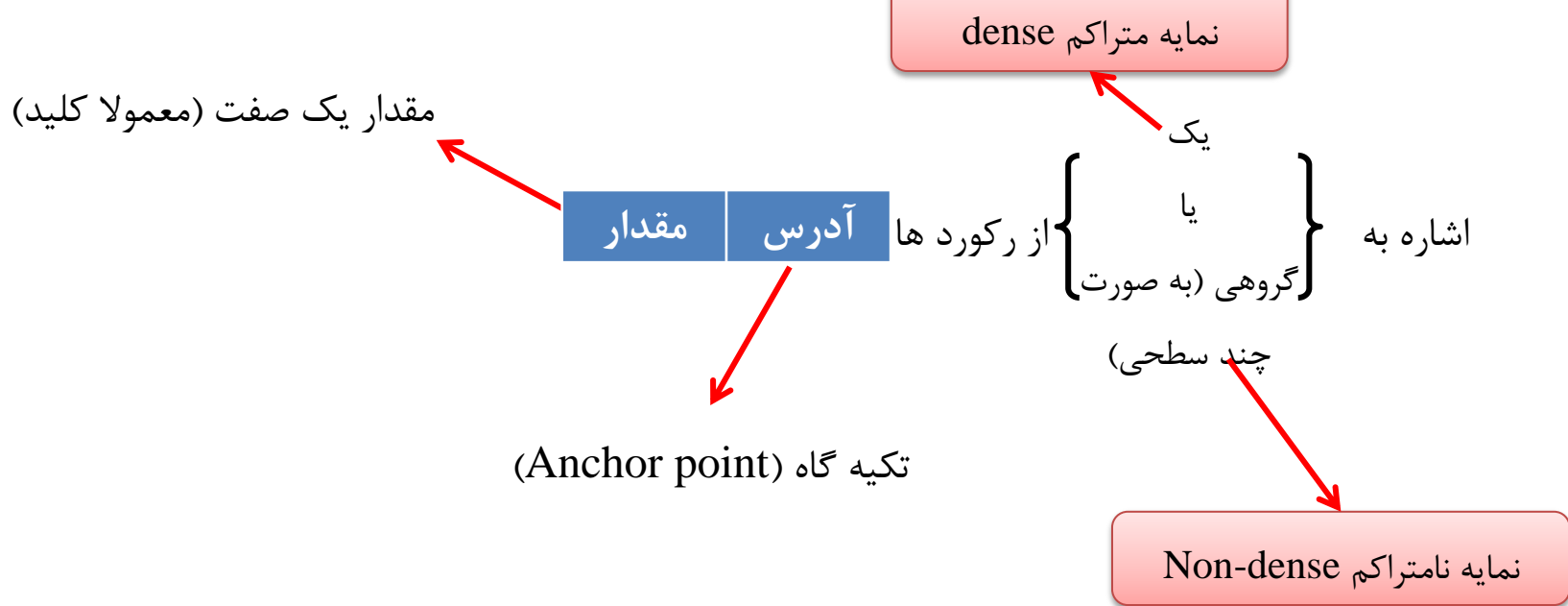
■ ۱- **طبیعی**: ناشی از ماهیت داده های محیط (مثل صفت رشته دانشجو که برای دانشجویان مختلف می تواند یکسان و در نتیجه تکراری باشد)

■ **کنجکاوی**: برای کاهش مصرف حافظه در حالت افزونگی طبیعی چه باید کرد؟

■ ۲- **تکنیکی**: ناشی از استفاده از یک تکنیک معمولا برای افزایش سرعت (مثل نمایه سازی [شاخص بندی  
[ Indexing])



تشکیل شده از تعدادی درایه (مدخل-entry)





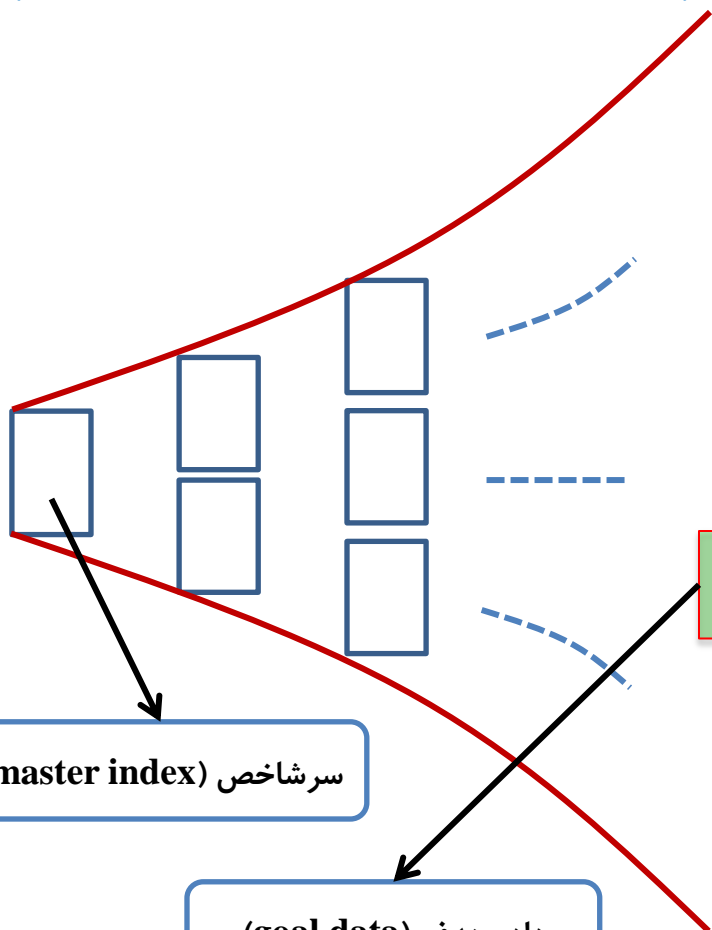
نمایه نامتراکم

نمایه متراکم

فایل نمایه سازی شده

(چندسطحی معمولاً با ساختار B-Tree یا B<sup>+</sup>-Tree)

(زمان جستجو بالاست)



سرشاخص (master index)

داده هدف (goal data)

شماره	نام	رشته	...
۱۰۰	۱	نرم افزار	
۱۰۱	۲	نرم افزار	
۱۰۲	۳	سخت افزار	
۱۰۳	۴	نرم افزار	
۱۰۴	۵	سخت افزار	
۱۰۵	۶	نرم افزار	
⋮			
k		نرم افزار	
⋮			
۹۹۷	۸۹۸	سخت افزار	
۹۹۸	۸۹۹	سخت افزار	
۹۹۹	۹۰۰	نرم افزار	

### افزونگی در معنای گسترده (در مباحث پایگاه داده):

عبارت است از تکرار ذخیره‌سازی داده‌ها در مورد نمونه‌های یک یا بیش از یک نوع موجودیت از یک محیط.

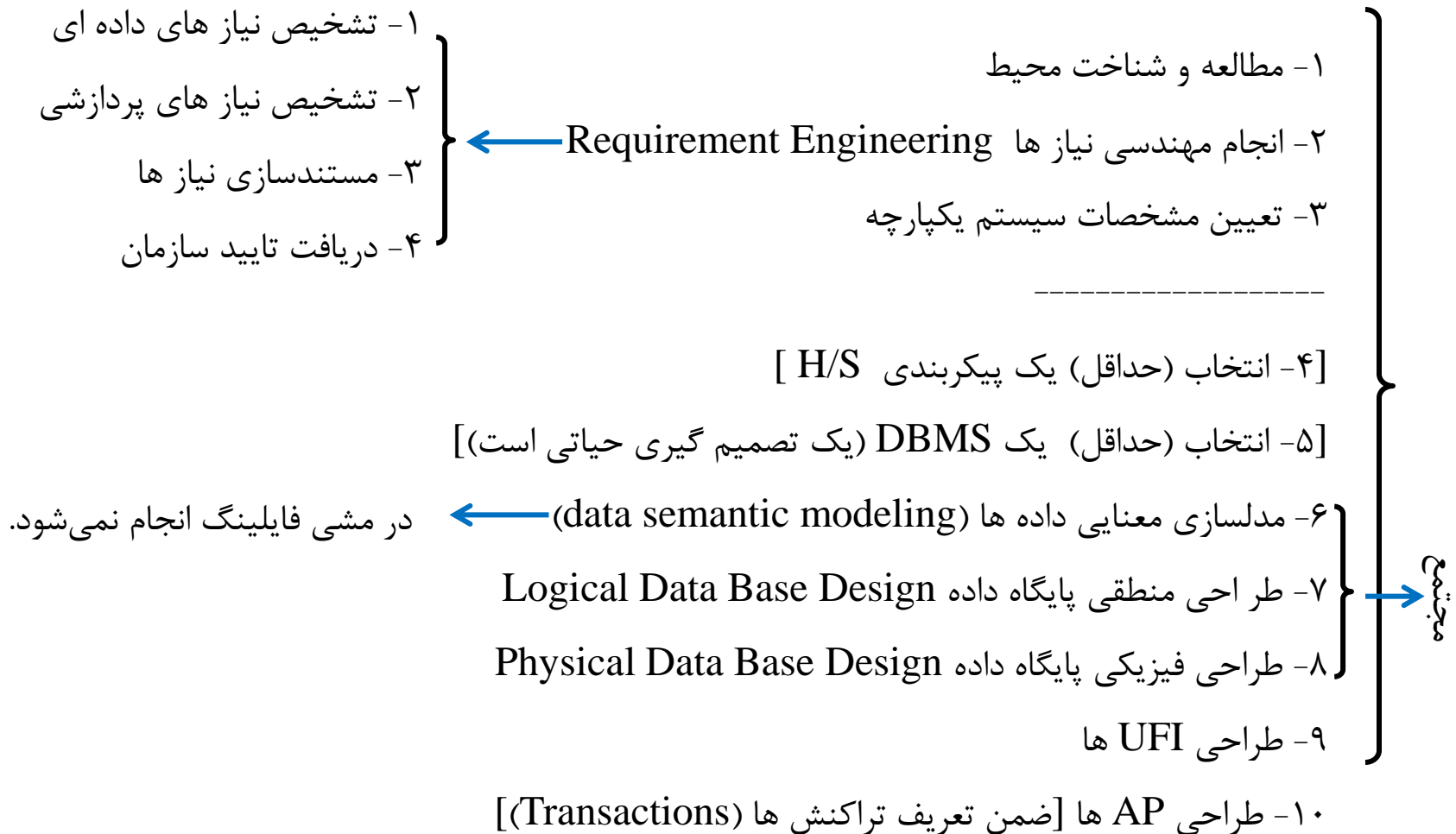
این نوع افزونگی نه از نوع طبیعی و نه از نوع تکنیکی است بلکه ناشی از رهیافت انتخاب شده برای طراحی و تولید سیستم‌های کاربردی است.

**نکته:** افزونگی از نوع طبیعی و تکنیکی در پایگاه داده هم می‌تواند وجود داشته باشد.

دلایل بروز افزونگی در سیستم‌های ISR به ویژه سیستم‌های پایگاهی کدامند؟



□ کارهای لازم در انجام یک «پروژه پایگاهی»: (فعلا نه در جزئیات)



ادامه: ...

مزایا و معایب جداسازی این دو دسته برنامه



تعریف و کنترل و عملیات در داده ها چیست؟

۱- از دیدگاه عملیات در داده ها

۲- از دیدگاه زبان های برنامه سازی

۱۱- تولید برنامه های تعریف (ایجاد) و کنترل DB

۱۲- تولید برنامه های عملیات در داده ها (پردازش داده ها)

۱۳- ایجاد محیط فیزیکی «ذ.ب.ا.» با داده های تستی و رفع اشکال ها (تست مرحله اول)

۱۴- ایجاد محیط فیزیکی «ذ.ب.ا.» با داده های واقعی اما حجم محدود و انجام تست مرحله دوم

۱۵- ایجاد محیط فیزیکی «ذ.ب.ا.» با داده های واقعی و حجم واقعی و انجام تست مرحله سوم

۱۶- تنظیم سیستم پایگاهی (Data Base System Tuning) ← به طور مثال به منظور افزایش کارایی

۱۷- آغاز بهره برداری و نگهداری از سیستم

۱۸- گسترش سیستم ← یکی از ویژگی های DBMS گسترش پذیری سیستم است.

۱۹- رفع معایب و بهینه سازی سیستم



## تراکنش Transaction:



□ دنباله ای از عملیات («قطعه برنامه») که معمولاً یک عمل تغییردهنده (درج، حذف، به روزرسانی) در محیط

ذخیره‌سازی داده‌ها انجام می‌دهد و یا باید به تمامی اجرا شود و یا اجرا نشده تلقی می‌شود

□ دارای خواص ACID (Atomicity Consistency Isolation Durability)



BEGIN TRANS

READ (A)

$A = A - 50$

UPDATE (A)

READ (B)

$B = B + 50$

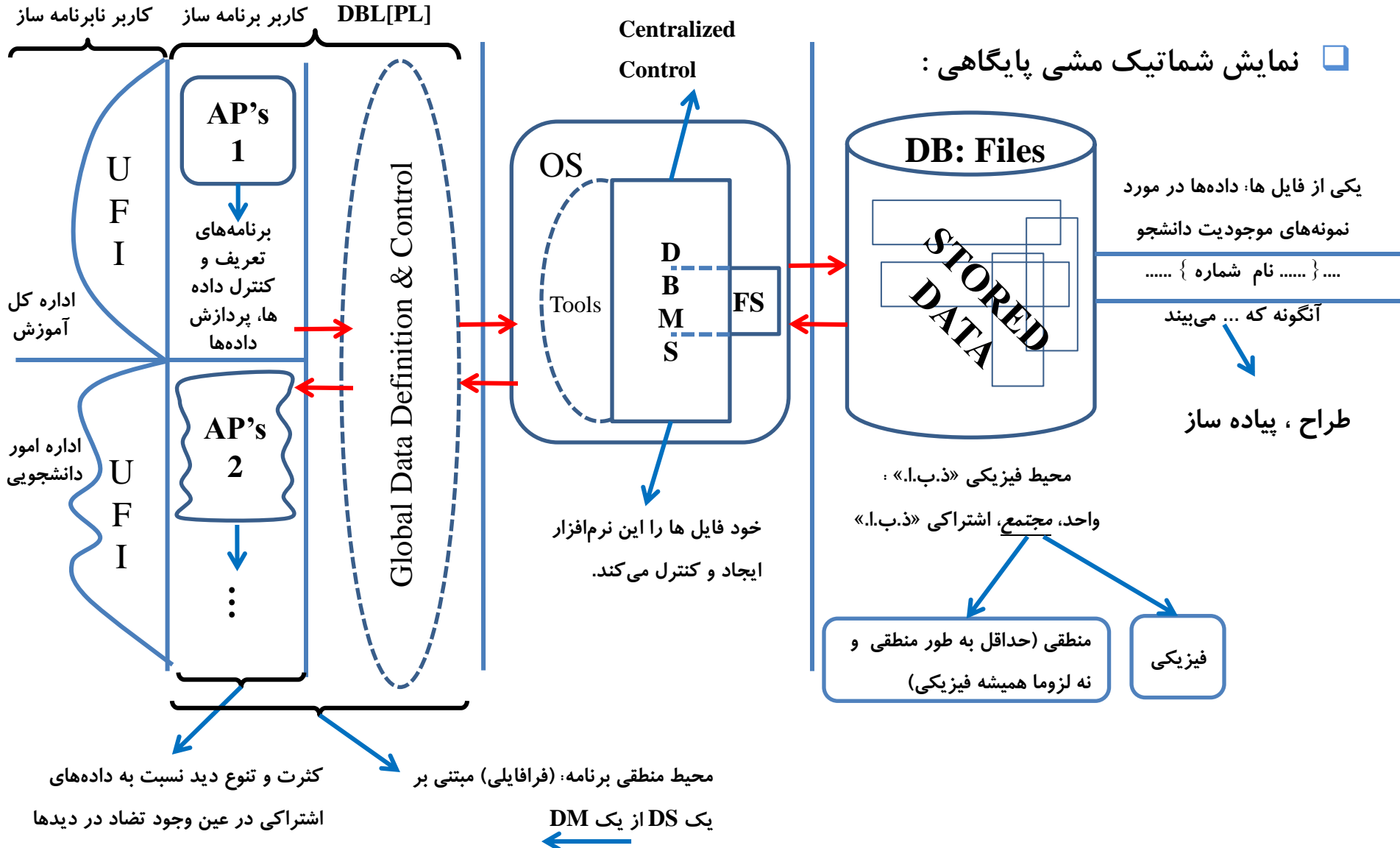
UPDATE (B)

END TRANS

شرط سازگاری پایگاه داده در این مثال :  $A+B$  ثابت باشد









۱- خود نرم افزار DBMS




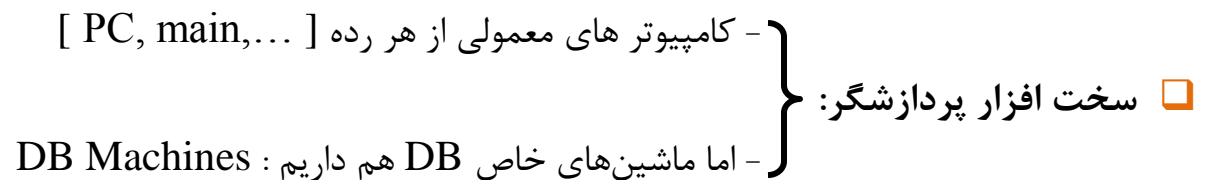
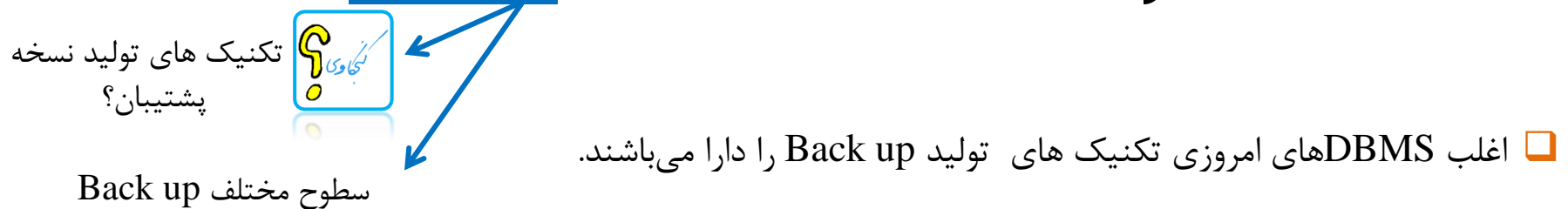
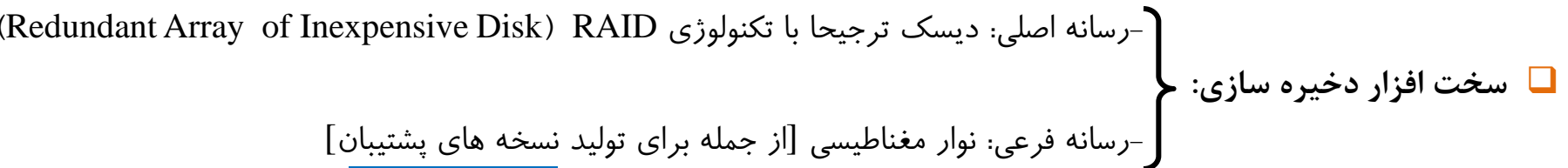
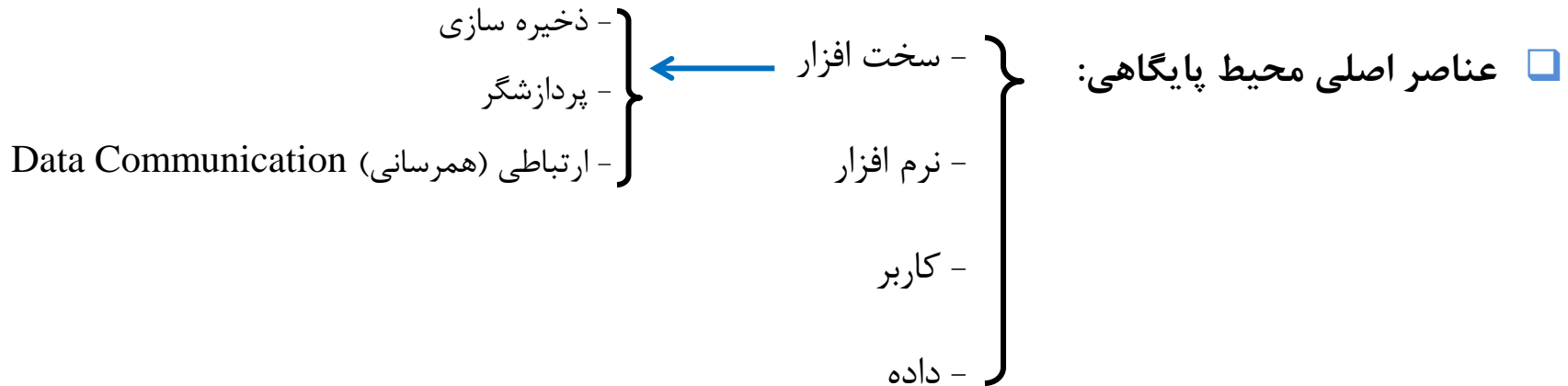
چگونه از این کثرت دید می توان به آن «وحدت» رسید؟



۲- معماری پایگاه داده

تمرین: مزایای مشی پایگاهی چیست؟  (طبق معلومات فعلی: عکس معایب مشی فایلینگ)

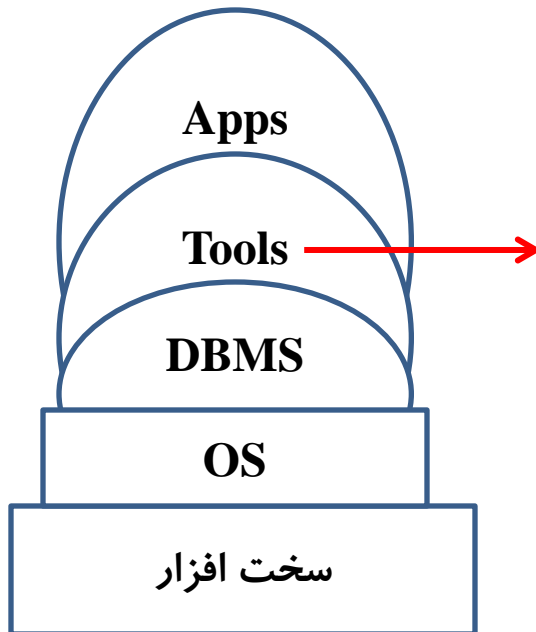
تمرین: چند سطح تعریف داده داریم؟ 





- سخت افزار ارتباطی (همرسانی):
- امکانات محلی: برای ارتباط دستگاه های جانبی با پردازنده
  - امکانات شبکه‌ای: برای ایجاد شبکه در سیستم پایگاهی نامتمرکز

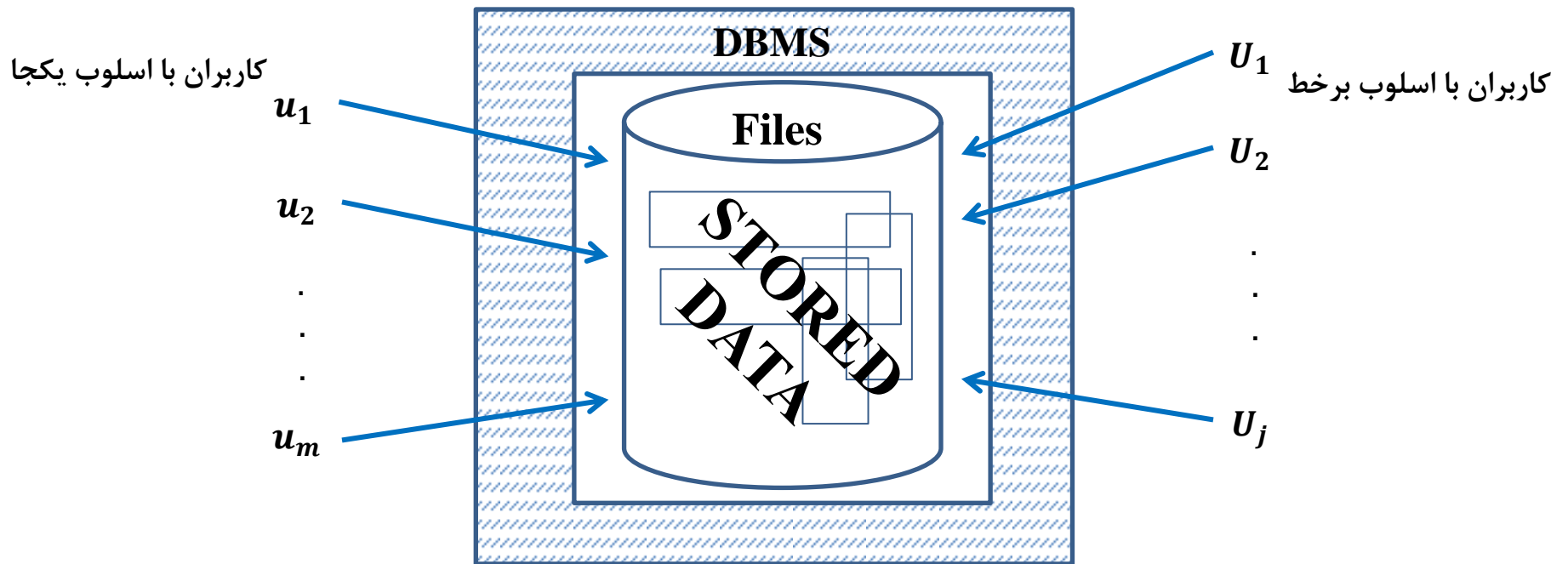
□ نرم افزار:



تسهیلات نرم افزار

یا با خود DBMS می فروشند،  
یا جداگانه خریداری می شود و به  
امکانات آن اضافه می شود.

□ در معنای عام هر استفاده‌کننده از سیستم پایگاهی را **کاربر** گوئیم که انواع مختلفی دارد.





□ انواع کاربر از نظر اسلوب عملیاتی:

□ **Batch – یکجا** (تعدادی برنامه یا پرس و جو جمع آوری می شود و به صورت یکجا به سیستم داده می شود و جواب آن بر می گردد).

□ **Online – برخط – پیوسته** (یک برنامه یا پرس و جو به سیستم داده می شود، اجرا می شود، و جوابش بر می گردد).

□ **Interactive – تعاملی** – بسته به اینکه چه جوابی داده شود عمل دیگری از کامپیوتر درخواست می شود.

▪ Online لزوماً Interactive نیست اما Interactive لزوماً Online است.

□ سیستم پایگاهی به صورت پیش فرض چند کاربره (multi-user) است.



□ داده‌های ذخیره شده در یک سیستم پایگاهی عبارتند از:

□ داده‌های کاربران

□ داده‌های سیستمی

□ مباحث مرتبط با داده در محیط پایگاهی در ادامه درس مطرح می‌گردد.



❑ سوال: می‌خواهیم یک سیستم کاربردی پایگاهی ایجاد کنیم. بر اساس کدام معماری ایجاد کنیم؟

❑ در توصیف معماری یک سیستم باید مشخص کنیم که

❑ از چه مولفه‌هایی، از هر مولفه چند عدد و با چه کیفیتی تشکیل شده است،

❑ مولفه‌ها چگونه با هم ترکیب شده‌اند (جنبه ساختاری سیستم)،

❑ مولفه‌ها چگونه با یکدیگر در تعامل هستند (جنبه رفتاری سیستم).

❑ انواع معماری سیستم پایگاهی:

❑ معماری متمرکز

❑ معماری نامتمرکز

▪ معماری مشتری-خدمت‌گزار

▪ معماری توزیع شده

▪ معماری چندپایگاهی

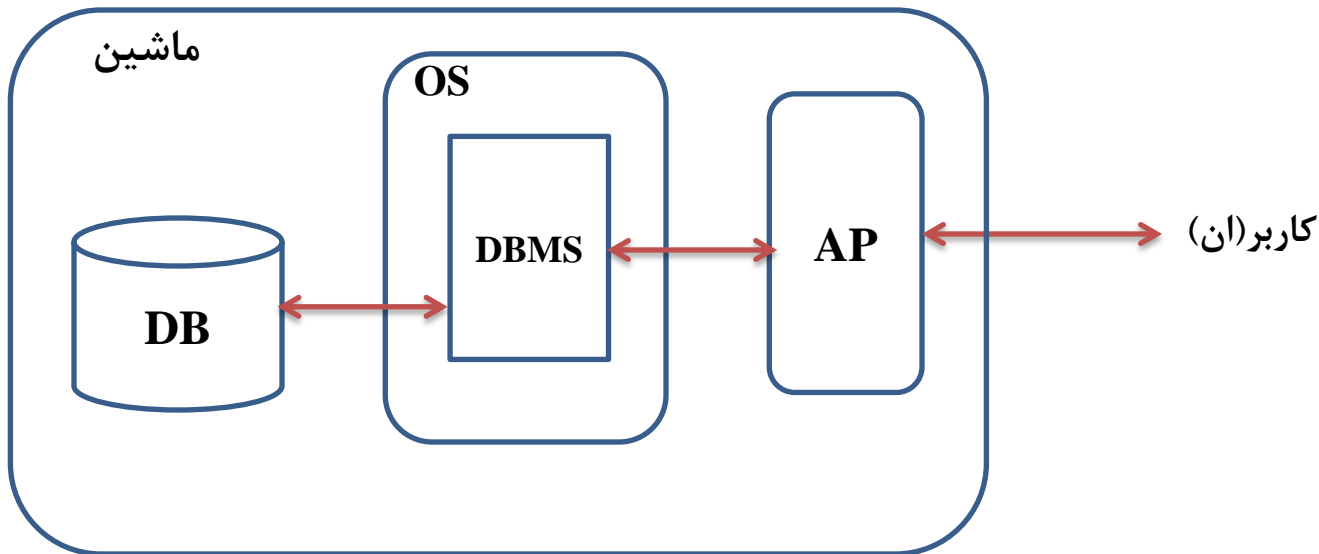
▪ معماری با پردازش موازی

▪ معماری موبایل



□ در این معماری یک پایگاه داده (متمرکز و مجتمع) روی یک سیستم کامپیوتری و بدون ارتباط با سیستم کامپیوتری دیگر ایجاد می‌شود.

□ معمولاً به صورت تک کاربری و برای کاربردهای کوچک و با امکانات محدود از این معماری استفاده می‌شود.





☐ **دلیل:** دلیل اصلی استفاده از معماری مشتری-خدمتگذار (Client-Server): تقسیم وظایف سیستم

☐ **تعریف:** هر ماشینی (فیزیکی یا منطقی) که خدمتی را به ماشین دیگر بدهد، **خدمتگذار** نامیده می‌شود.

 نمونه‌هایی از انواع خدمتگذارها: File Server, Print Server, Message Server, DB Server

☐ انواع معماری مشتری – خدمتگذار

☐ معماری تک مشتری – تک خدمتگذار

☐ معماری چند مشتری – تک خدمتگذار

☐ معماری تک مشتری – چند خدمتگذار

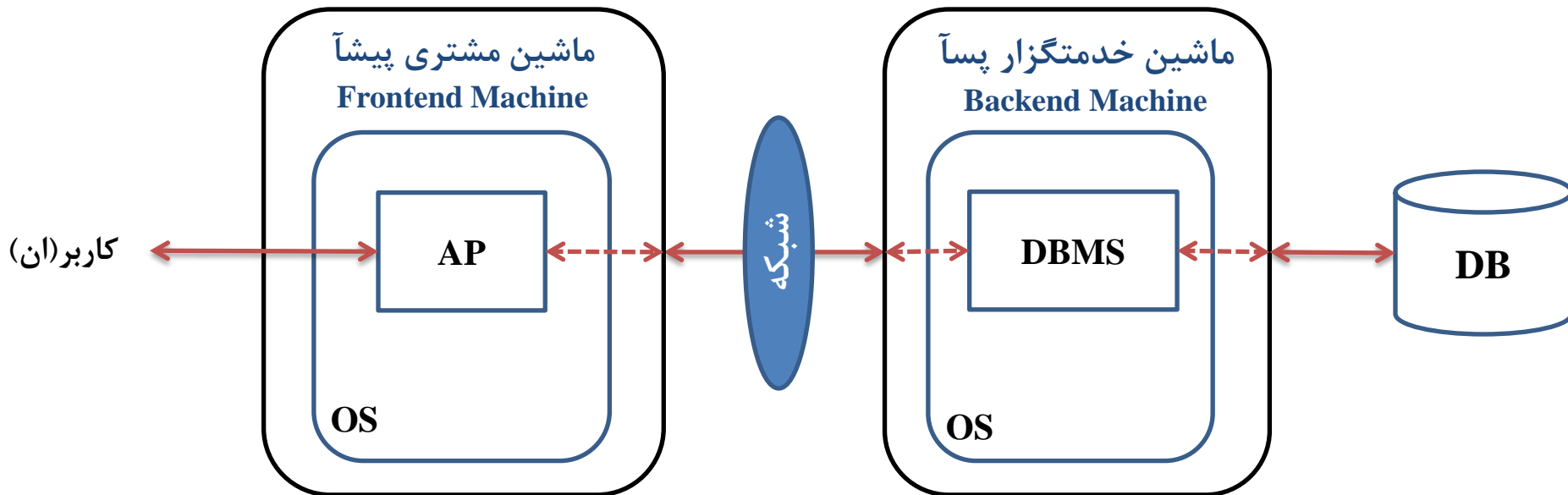
☐ معماری چند مشتری – چند خدمتگذار

معمولا شامل دو سایت:

سایت مشتری: تمام برنامه‌های کاربردی در آن اجرا می‌شوند.

سایت خدمتگذار: تمام داده‌ها در آن ذخیره می‌شوند

به این معماری، **معماری دولایه (2-tier)** نیز گویند.





ماشین‌های ساده، ارزان و حتی بدون دیسک (thin client)

□ برخی مزایای معماری سه لایه نسبت به دو لایه:

□ اطمینان عملیاتی بیشتر

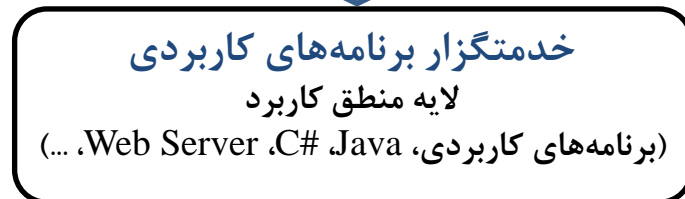
□ گسترش پذیری بهتر

□ امنیت داده‌ای بیشتر (عدم ارتباط مستقیم مشتری‌ها با کارگزار داده)

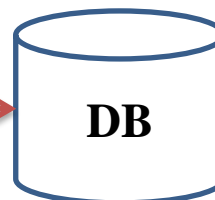
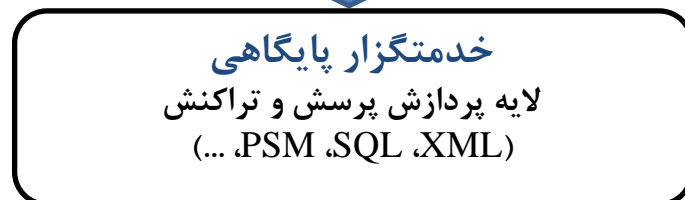
□ قابلیت کاهش هزینه سخت افزاری (با استفاده از thin client)



پروتکل HTTP



ODBC, JDBC, SQL, SQL/CLI





سیستم‌های پایگاهی همزمان با ناهمزمان ایجاد می‌شوند.



اجزای تشکیل‌دهنده سیستم‌ها (OSها و DBMSها) معمولاً همگن هستند.



برخی سایت‌ها ممکن است فقط مشتری و یا خدمتگزار باشند.

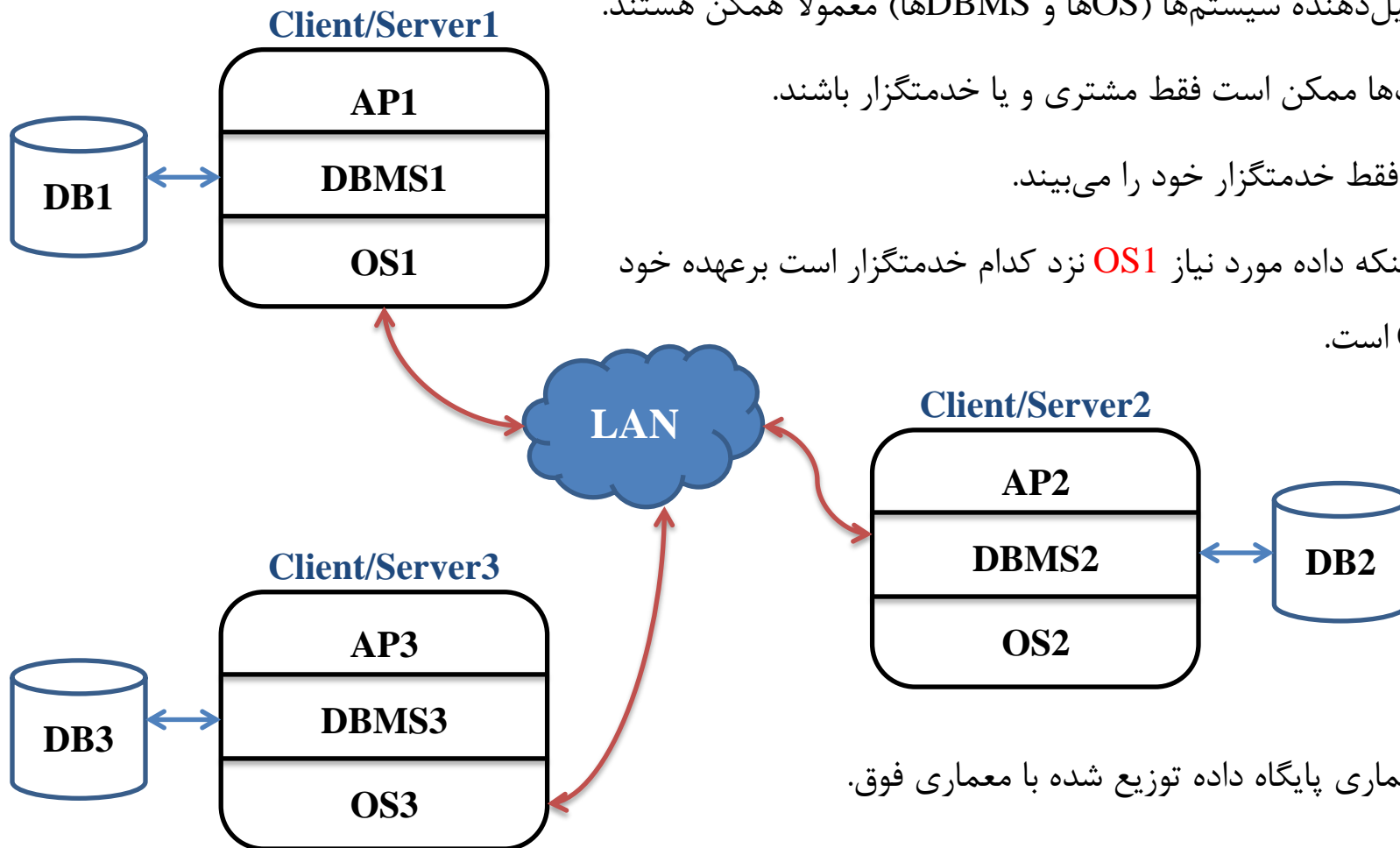


هر مشتری فقط خدمتگزار خود را می‌بیند.



مسئولیت اینکه داده مورد نیاز **OS1** نزد کدام خدمتگزار است برعهده خود

مشتری C1 است.



تفاوت معماری پایگاه داده توزیع شده با معماری فوق.



**پرسش و پاسخ ...**

**[amini@sharif.edu](mailto:amini@sharif.edu)**