بسمه تعالی

دانشگاه جامع علمی و کاربردی

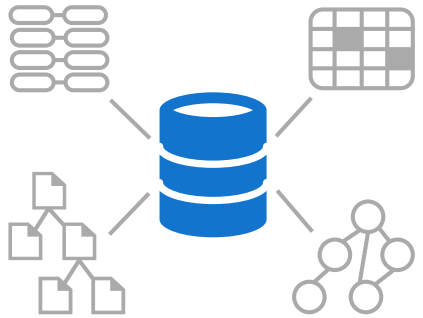
اصول طراحی پایگاه داده ها

Data Base

مدرس : نوید همراهی

www.navidhamrahi.ir

n.hamrahi@gmail.com



فهرست جلسات

جلسه اول: مفاهیم پایگاه داده‌ها

جلسه دوم: مدلسازی معنایی داده‌ها

جلسه سوم: محدودیتهای روش ER

جلسه چهارم: پایگاه داده در محیط انتزاعی

جلسه پنجم: معماری پایگاه داده‌ها

جلسه ششم: سیستم مدیریت پایگاه داده‌ها

جلسه هفتم: DBMS در یک سیستم کامپیوتری

جلسه هشتم: معماری سیستم پایگاه داده‌ها

فهرست جلسات

جلسه نهم: مدل رابطه‌ای

جلسه دهم: قواعد جامعیت پایگاه داده‌ها

جلسه یازدهم: زبان SQL

جلسه دوازدهم: ادامه زبان SQL

جلسه سیزدهم: دید در مدل رابطه‌ای

جلسه چهاردهم: طراحی پایگاه داده‌های رابطه‌ای

جلسه پانزدهم: نرمالترسازی رابطه‌ها

جلسه شانزدهم: طراحی فیزیکی پایگاه داده‌ها

اهداف این درس

در این درس به بخشی از مفاهیم مبنایی دانش و تکنولوژی پایگاه داده‌ها و اصول مدلسازی و طراحی آن پرداخته می‌شود. پایگاه داده و عناصر اصلی محیط آن معرفی می‌شود. با مدلسازی معنایی داده‌ها و محیط انتزاعی آشنا می‌شویم. سطوح معماریهای پایگاه داده‌ها بیان می‌شود. سیستم مدیریت پایگاه داده‌ها (DBMS) و اجزاء آن شرح داده می‌شود. به انواع معماریهای سیستم پایگاهی و مفاهیم اساسی مدل رابطه‌ای پرداخته می‌شود. با زبان SQL به عنوان یک زبان رابطه‌ای و همچنین با دیدهای رابطه‌ای آشنا می‌شویم و در نهایت طراحی پایگاه داده‌ها به روش بالا به پایین، سنتز و طراحی فیزیکی را خواهیم دید.

جلسه اول

مفاهیم پایگاه داده‌ها

آنچه در این جلسه می خوانید:

۱- سیستم ذخیره و بازیابی اطلاعات در معنای عام
۲- رده‌های تکنولوژیکی سیستم مدیریت پایگاه داده‌ها
۳- داده
۴- تعریف اطلاع
۵- تعریف دانش
۶- تعریف پایگاه داده‌ها

آنچه در این جلسه می خوانید:

۷- مراحل کلی کار در مشی فایلینگ
۸- معایب مشی فایلینگ
۹- مراحل کلی کار در مشی پایگاهی
۱۰- عناصر محیط پایگاه داده‌ها
۱۱- انواع سخت‌افزارهای محیط پایگاه داده
۱۲- انواع نرم‌افزارهای موجود در محیط پایگاه داده‌ها

هدفهای کلی: مقدمه و آشنایی با مفاهیم پایگاه داده‌ها

هدفهای رفتاری: دانشجو در پایان این جلسه می‌تواند:

- داده، اطلاع، شناخت، سیستم ذخیره و بازیابی و پایگاه داده‌ها را تعریف کند.
- رده‌های تکنولوژیکی پایگاه داده را بیان کند.
- رهیافتهای ایجاد یک سیستم کاربردی را ارائه کند.

اصطلاح پایگاه داده‌ها یکی از رایج‌ترین اصطلاحات در دانش و فن کامپیوتر است

در این درس دانشجویان تنها با بخشی از مفاهیم بنیادی دانش وفن پایگاه داده‌ها آشنا شده، آگاهی پایه‌ای لازم را برای مطالعه بیشتر و یا کار در این زمینه کسب می‌کنند.

سیستم مدیریت پایگاه داده‌ها یکی از سیستم‌های ذخیره و بازیابی اطلاعات است.

سیستم ذخیره و بازیابی اطلاعات در معنای عام:

هر سیستمی که به کاربر برنامه‌ساز یا
نابرنامه‌ساز امکان دهد تا اطلاعات خود را
ذخیره، بازیابی و پردازش کند.

رده‌های تکنولوژیکی سیستم مدیریت پایگاه داده‌ها

۱- سیستم فایلینگ
۲- سیستم مدیریت داده‌ها
۳- سیستم مدیریت پایگاه داده‌ها
۴- سیستم مدیریت پایگاه شناخت
۵- سیستم مدیریت پایگاه داده‌های شیئ گرا
۶- سیستم هوشمند مدیریت پایگاه داده‌ها
۷- سیستم معنایی مدیریت پایگاه داده‌ها
۸- سیستم مدیریت پایگاه داده‌های زمانبند
۹- سیستم مدیریت پایگاه داده‌های نیم‌ساختمند و ناساختمند

ادامه

رده‌های تکنولوژیکی سیستم مدیریت پایگاه داده‌ها

۱۰- سیستم مدیریت پایگاه داده‌های بی درنگ

۱۱- سیستم داده‌کاوی و کشف شناخت

۱۲- سیستم مدیریت چند پایگاهی

۱۳- سیستم اطلاعات اجرائی

۱۴- سیستم فعال مدیریت پایگاه داده‌ها

۱۵- سیستم مدیریت پایگاه داده‌های شیئ-رابطه‌ای

داده

تعریف اول- نمایش ذخیره‌شده اشیاء فیزیکی، چیزهای مجرد، بوده‌ها، رویدادها یا چیزهای قابل مشاهده که در تصمیم‌سازی بکار می‌آیند.

داده

تعریف دوم- هر مجموعه‌ای از بوده‌ها

داده

تعریف سوم- بوده‌های خام که معنای اندکی دارند مگر اینکه به صورت منطقی سازمان‌دهی شده باشند

تعریف داده از دیدگاه ANSI
(موسسه استاندارد ملی آمریکا)

نمایش بوده‌ها، پدیده‌ها، مفاهیم یا شناخته‌ها به طرزی صوری و مناسب برای برقراری ارتباط، تفسیر یا پردازش توسط انسان یا هر امکان خودکار

هر نمایشی اعم از کاراکتری یا کمیتهای قیاسی که معنایی به آن قابل انتساب باشد.

تعریف اطلاع

اطلاع به داده‌ای اطلاق می‌شود که توسط یک فرد یا سازمان برای تصمیم‌گیری بکار می‌رود

اطلاع، داده پردازش‌شده است.

اطلاع عبارت است از داده سازمان یافته‌ای که شناختی را منتقل می‌کند

تعریف دانش

دانش عبارتست از نمایش نمادین
جنبه‌هایی از بخشی از خرد جهان واقع

تعریف پایگاه داده‌ها

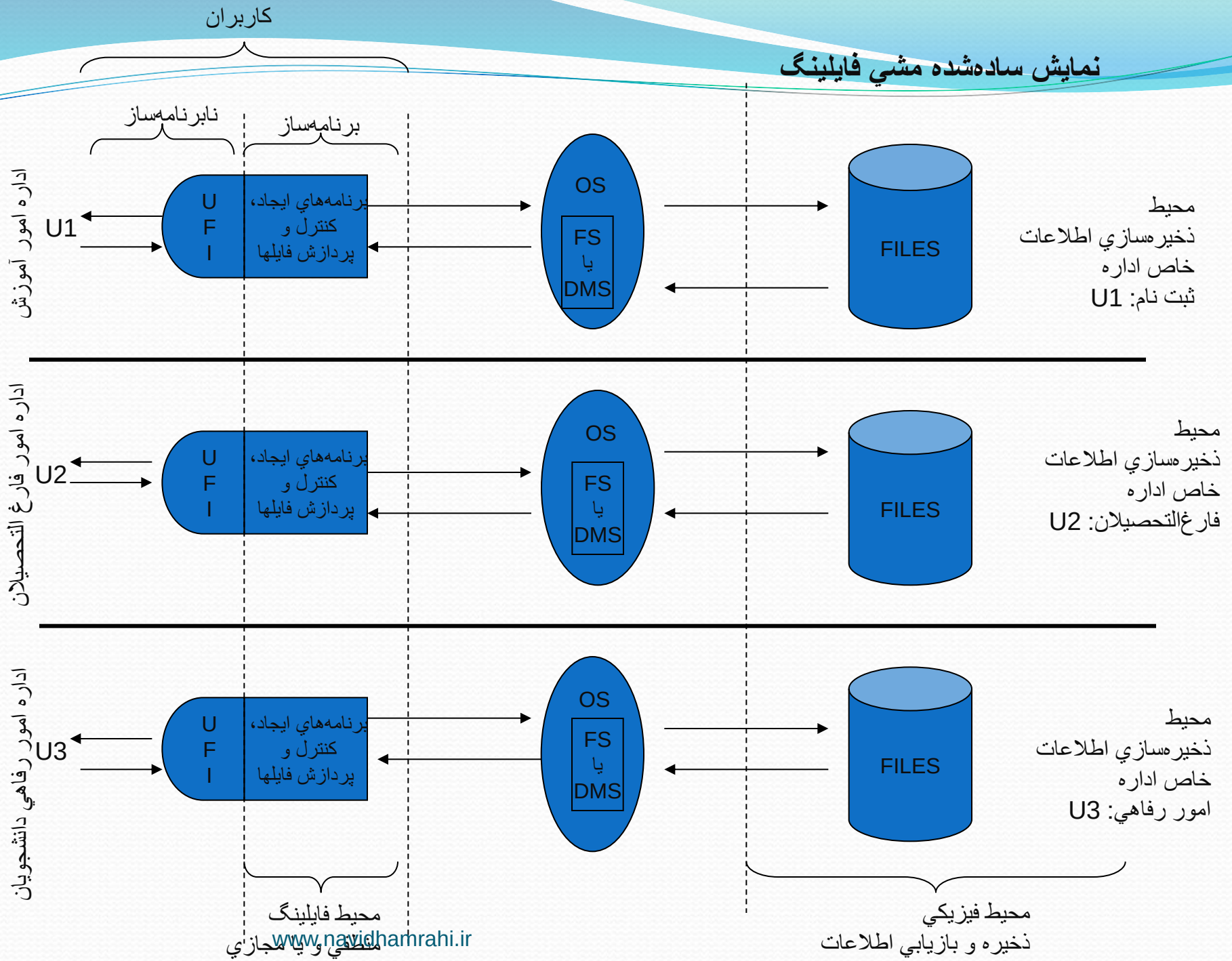
مجموعه‌ای است از داده‌های ذخیره شده و پایا، به صورت مجتمع(یکپارچه) (نه لزوما فیزیکی، بلکه حداقل به طور منطقی)، بهم مرتبط، با کمترین افزونگی، تحت مدیریت یک سیستم کنترل متمرکز، مورد استفاده یک یا چند کاربر از یک یا بیش از یک "سیستم کاربردی"، به طور همزمان و اشتراکی

برای ایجاد یک سیستم کاربردی دو رهیافت وجود دارد:

۱- رهیافت سنتی یا مشی فایلینگ

۲- رهیافت (مشی) پایگاهی

نمايش ساده‌شده مشي فايلينگ



www.navidhamrahi.ir

مراحل کلی کار در مشی فایلینگ

- تحلیل و بررسی نیازهای اطلاعاتی و پردازشی هر قسمت به طور جداگانه
- اجرای مراحل کلاسیک اولیه لازم برای طراحی و تولید یک سیستم کاربردی
- تعیین مشخصات هر سیستم و وظایف آن
- طراحی تعدادی فایل
- نوشتن مجموعه‌ای از برنامه‌های ایجاد، کنترل و پردازش فایل

مراحل کلی کار در مشی فایلینگ

- استفاده از یک پیکربندی سخت‌افزاری و نرم‌افزاری مشخص

- انجام تستهای لازم و تنظیم سیستم کاربردی

- ایجاد یک سیستم کاربردی برای هر قسمت و برپایی محیط فیزیکی ذخیره و بازیابی اطلاعات و سیستم بهره‌برداری از آن خاص همان قسمت.

معایب مشی فایلینگ

۱- عدم وجود محیط مجتمع ذخیره‌سازی اطلاعات و عدم وجود سیستم یکپارچه

۲- عدم وجود سیستم کنترل متمرکز روی کل داده‌ها

۳- افزونگی

۴- عدم وجود ضوابط ایمنی کارا و مطمئن

۵- خطر بروز پدیده ناسازگاری داده‌ها

ادامه

معایب مشی فایلینگ

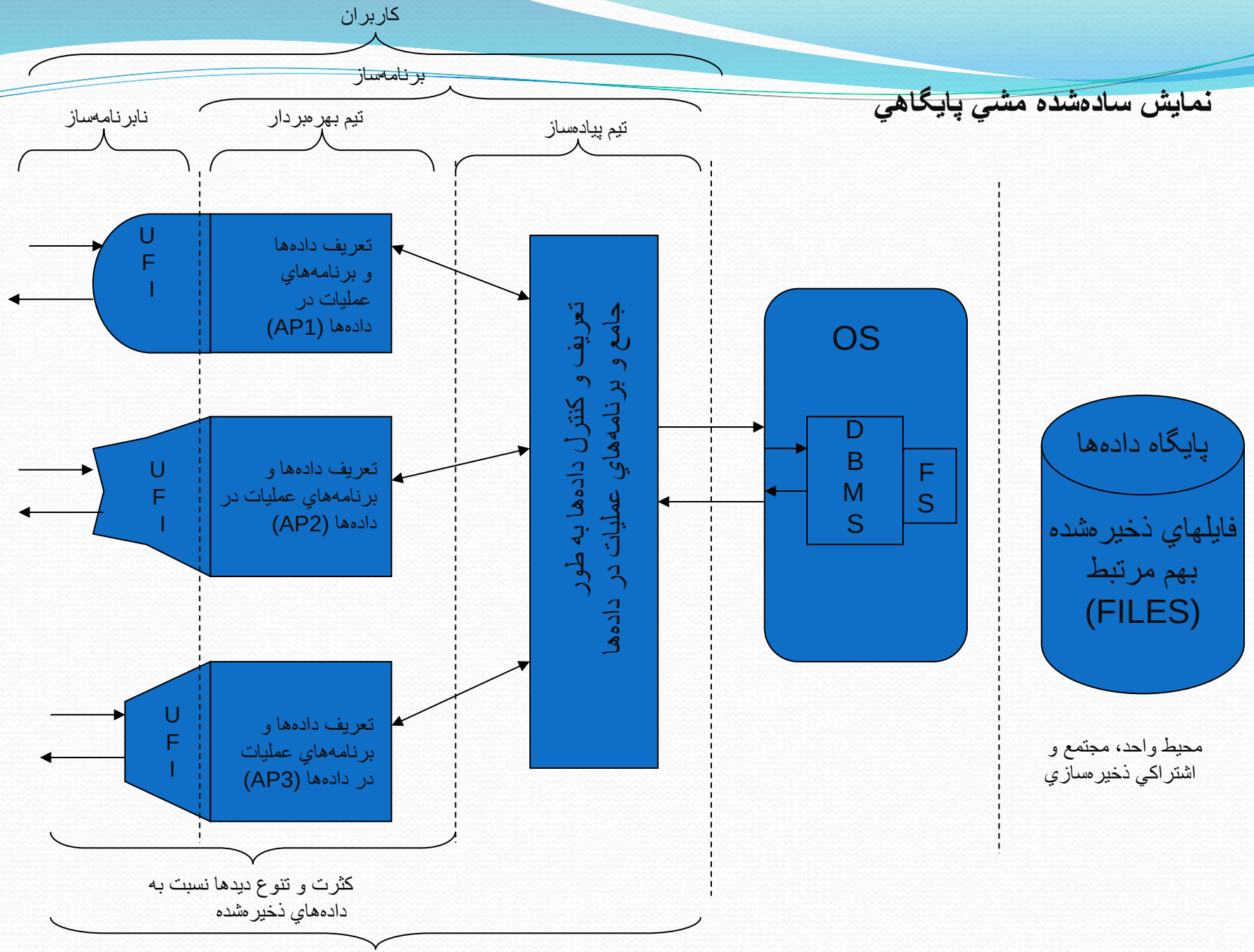
۶- عدم امکان اشتراکی شدن داده‌ها

۷- مصرف نابهینه امکانات سخت‌افزاری و نرم‌افزاری

۸- حجم زیاد برنامه‌سازی

۹- وابستگی برنامه‌های کاربردی به محیط ذخیره‌سازی داده‌ها

نمايش ساده‌شده مشي پايگاهي



مراحل کلی کار در مشی پایگاهی

- بررسی و تحلیل نیازهای پردازشی و اطلاعاتی همه قسمتها توسط یک گروه
- مدلسازی معنایی داده‌ها
- تعیین مشخصات جامع (یکپارچه) کاربردی و وظایف آن
- انتخاب یک یا چند پیکربندی سخت‌افزاری-نرم‌افزاری
- استفاده از یک یا چند DBMS
- طراحی پایگاه داده‌ها در سطوح لازم

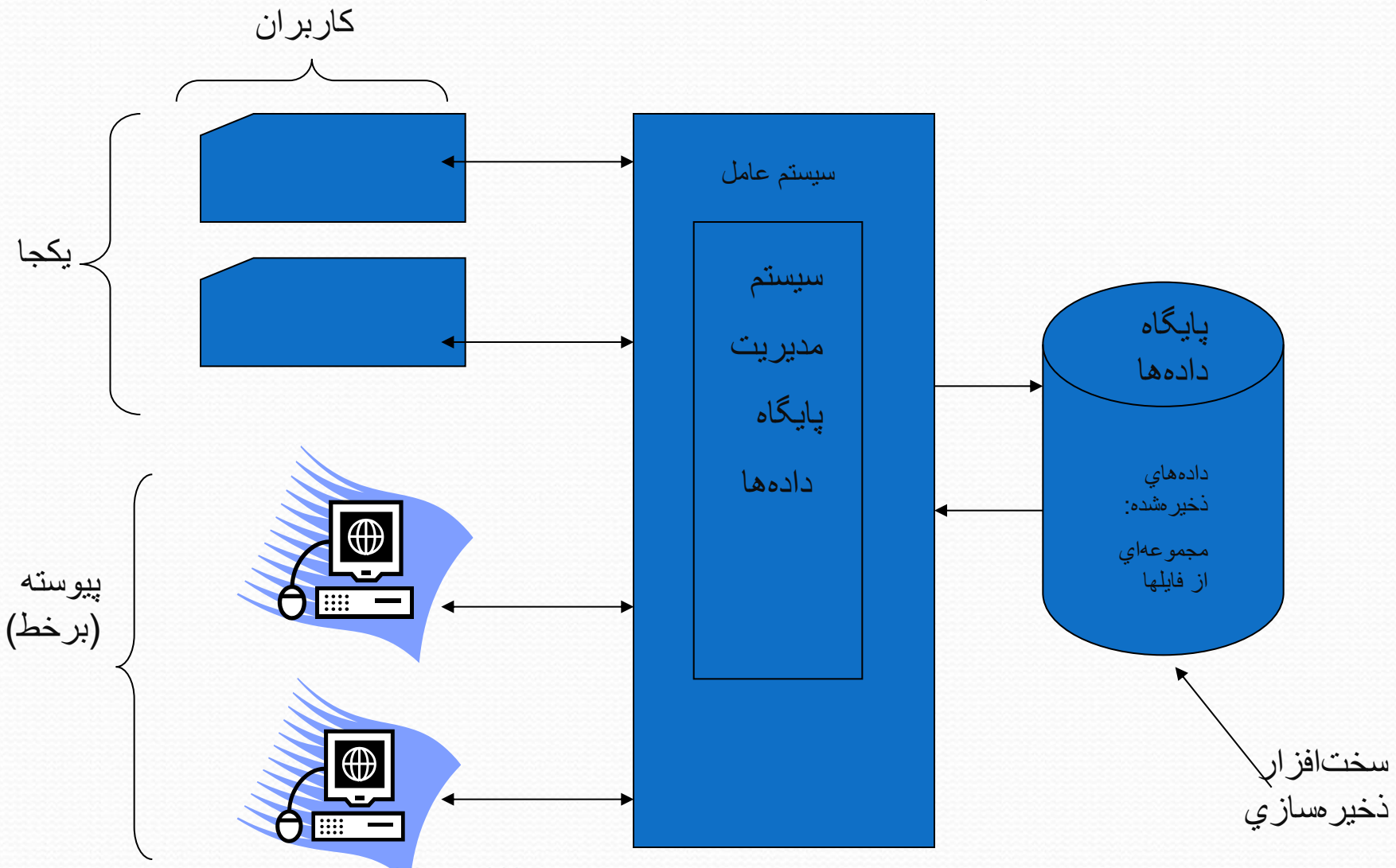
ادامه

مراحل کلی کار در مشی پایگاهی

- تولید مجموعه‌ای از برنامه‌های ایجاد و کنترل پایگاه داده
- ایجاد محیط واحد و مجتمع ذخیره‌سازی و مشترک بین کاربران
- طراحی و تولید واسطهای کاربرپسند مورد نیاز
- تعریف پایگاه داده هر قسمت توسط کاربر مربوطه
- طراحی برنامه‌های عملیات در پایگاه داده
- بهره‌برداری واقعی از سیستم پس از تستهای لازم

عناصر محیط پایگاه داده‌ها

۱- سخت‌افزار ۲- نرم‌افزار ۳- کاربر ۴- داده



انواع سخت‌افزارهای محیط پایگاه داده

۱- سخت‌افزار ذخیره‌سازی داده‌ها

۲- سخت‌افزار پردازشگر

۳- سخت‌افزار ارتباط

انواع نرم‌افزارهای موجود در محیط پایگاه داده‌ها

۱- سیستم مدیریت پایگاه داده‌ها (DBMS)

۲- برنامه‌های کاربردی قابل اجرا در محیط DBMS

۳- رویه‌های ذخیره‌شده

۴- نرم‌افزار شبکه

جلسه دوم

مدلسازی معنایی داده‌ها

آنچه در این جلسه می خوانید:

۱- مدلسازی معنایی داده‌ها
۲- انواع روشهای مدلسازی معنایی داده‌ها
۳- سه مفهوم معنایی موجود در روش ER
۴- تعریف موجودیت
۵- سه ضابطه در رابطه با تشخیص یک نوع موجودیت
۶- موجودیت مستقل و وابسته
۷- تعریف صفت

آنچه در این جلسه می خوانید:

۸- انواع صفت و تعاریف هریک
۹- ارتباط
۱۰- خصوصیات نوع ارتباط
۱۱- نمودار ER
۱۲- نمادهای رسم نمودار ER
۱۳- وضع مشارکت در ارتباط
۱۴- درجه آن ارتباط
۱۵- چندی یا ماهیت نوع ارتباط

هدفهای کلی: آشنایی با مدلسازی معنایی داده‌ها

هدفهای رفتاری: دانشجو در پایان این جلسه می‌تواند:

- مدلسازی معنایی پایگاه داده‌ها و انواع آن را بیان کند.
- روش مدلسازی ER و مفاهیم آن را تشریح کند.
- نمودار ER و نمادهای آن را رسم کند.

مدلسازی معنایی داده‌ها

داده‌های ذخیره‌شدنی در پایگاه داده‌ها
ابتدا باید در بالاترین سطح انتزاع
مدلسازی معنایی شوند.

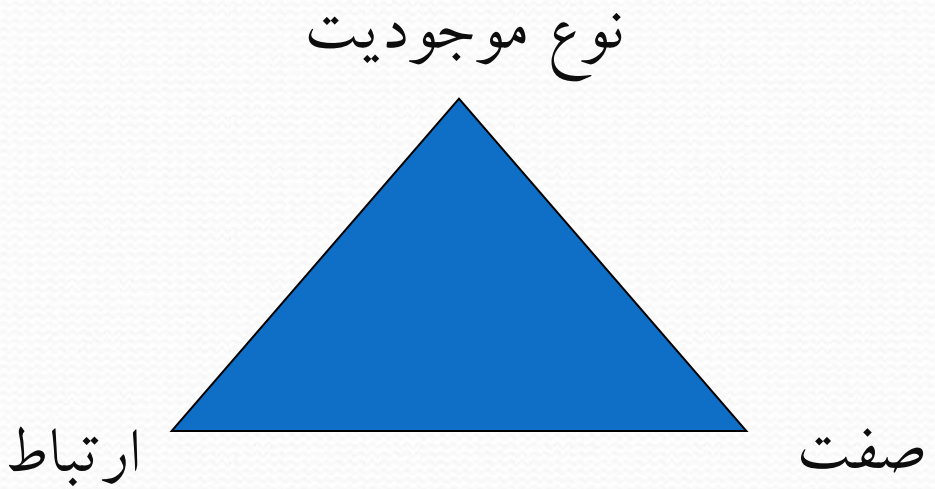
انواع روشهای مدلسازی معنایی داده‌ها

روش موجودیت- ارتباط (ER)

روش زبان عمومی مدلسازی (UML)

روش تکنیک مدلسازی شیئی (OMT)

سه مفهوم معنایی موجود در روش ER



تعریف موجودیت

مفهوم کلی شیئ، چیز، پدیده و به طور کلی هر آنچه که می‌خواهیم در موردش اطلاع داشته باشیم و شناخت خود را در موردش افزایش دهیم.

سه ضابطه در رابطه با تشخیص یک نوع موجودیت

۱- معمولا نمونه‌هایی متمایز از یکدیگر دارند.

۲- معمولا بیش از یک صفت دارد و کاربر به مجموعه‌ای از اطلاعات در مورد آن نیاز دارد.

۳- معمولا حالت کنشگری (فاعلیت) یا حالت کنشپذیری (مفعولیت) دارد.

موجودیت مستقل و وابسته

موجودیت مستقل (قوی)، موجودیتی است که مستقل از هر موجودیت دیگر و به خودی خود، در یک محیط مشخص مطرح باشد.

موجودیت وابسته (ضعیف)، موجودیتی است که وجودش وابسته به یک نوع موجودیت دیگر است.

تعریف صفت

خصیصه یا ویژگی یک نوع موجودیت است و هر نوع موجودیت مجموعه‌ای از صفات دارد. هر صفت یک نام، یک نوع و یک معنای مشخص دارد.

انواع صفت

- ساده یا مرکب
- تک‌مقداری یا چندمقداری
- شناسه یا ناشناسه موجودیت
- هیچ‌مقدارپذیر یا ناپذیر
- ذخیره‌شده یا مشتق

مقدار صفت ساده از لحاظ معنایی تجزیه‌نشدنی یا اتومیک است.

صفت مرکب از چند صفت ساده تشکیل شده است.

صفت تک‌مقداری، صفتی است که برای یک نمونه از یک نوع موجودیت حداکثر یک مقدار از دامنه مقادیر را می‌گیرد

صفت چندمقداری بیش از یک مقدار از دامنه مقادیر می‌گیرد.

صفت شناسه موجودیت که گاه به آن کلید هم گفته می‌شود، دو ویژگی دارد:

۱- یکتایی مقدار دارد.

۲- حتی الامکان طول مقادیرش کوتاه است.

صفت هیچ مقدارپذیر

هیچ مقدار یعنی مقدار ناشناخته، مقدار غیرقابل اعمال، مقدار تعریف‌نشده.

اگر مقدار یک صفت در یک یا بیش از یک نمونه از یک نوع موجودیت، برابر با هیچ‌مقادر باشد، آن صفت هیچ‌مقدارپذیر است.

صفت ذخیره‌شده و مشتق

صفت ذخیره‌شده صفتی است که مقادیرش در پایگاه داده‌ها ذخیره شده باشد.

صفت مشتق، صفتی است که مقادیرش در پایگاه داده‌ها ذخیره نشده باشد، بلکه حاصل یک پردازش روی فقره‌هایی از داده‌های ذخیره شده باشد.

ارتباط

تعریف- اندرکنش (تعامل) بین دو یا بیش از دو نوع موجودیت است و ماهیتا نوعی بستگی بین انواع موجودیتهاست

خصوصیات نوع ارتباط

۱- هر ارتباط یک نام دارد

۲- هر ارتباط یک معنای مشخص دارد و این معنا با معنای هر ارتباط دیگر متفاوت است.

۳- هر ارتباط نمونه‌هایی دارد.

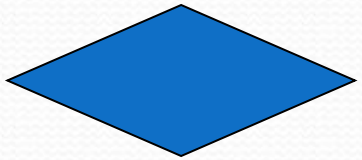
نمودار ER

نموداری است که سه مفهوم اساسی مدل ER، یعنی نوع موجودیت، صفت و ارتباط نمایش داده می‌شوند

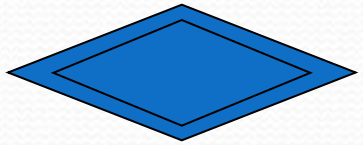
نمادهای رسم نمودار ER

نوع موجودیت

نوع موجودیت ضعیف



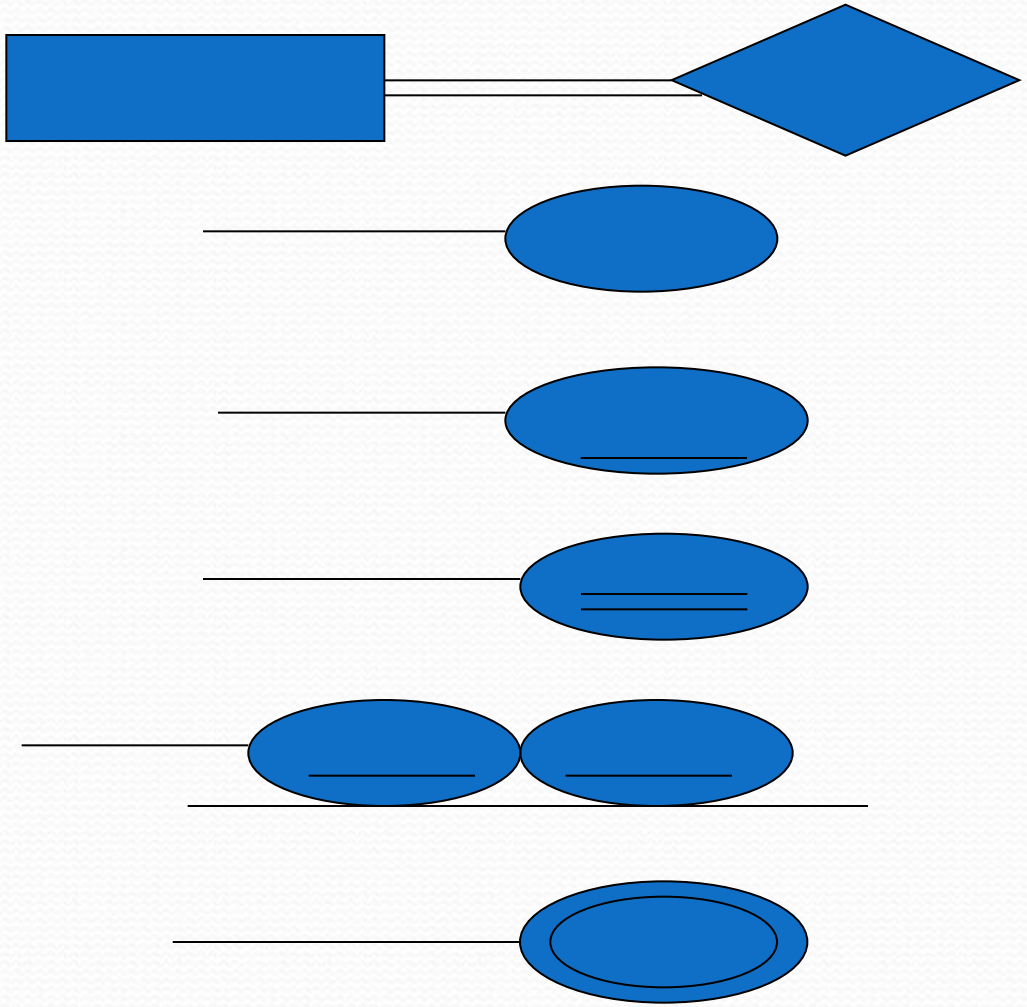
نوع ارتباط



نوع ارتباط با موجودیت ضعیف

مشارکت نوع موجودیت در نوع ارتباط

نمادهای رسم نمودار ER



مشارکت الزامی

صفت

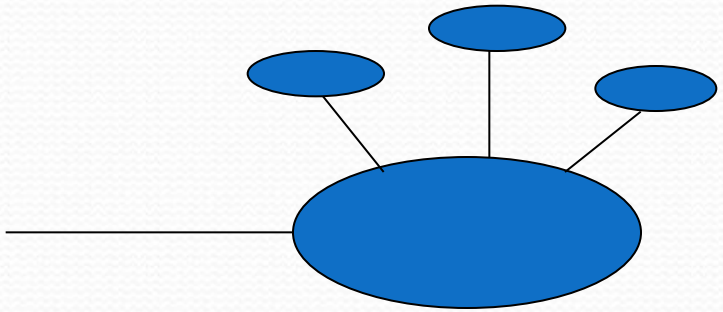
صفت شناسه اول

صفت شناسه دوم

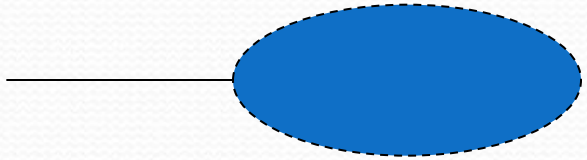
صفت شناسه مرکب

صفت چندمقداری

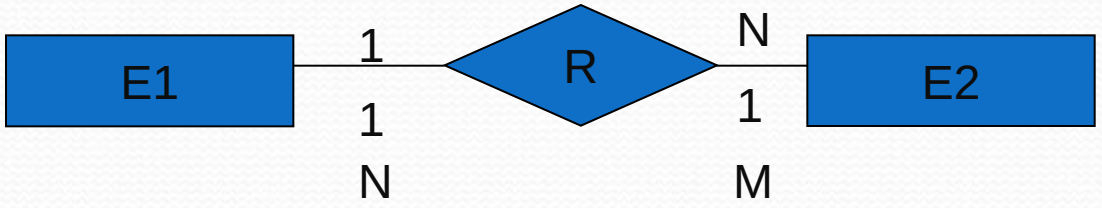
نمادهای رسم نمودار ER



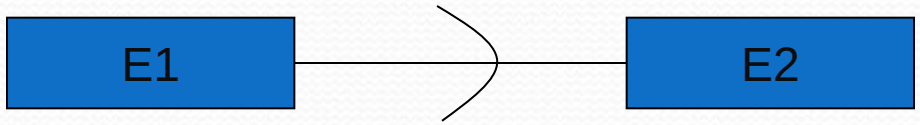
صفت مرکب



صفت مشتق



چندی ارتباط **1** به **N**
1 به **1**
M به **N**



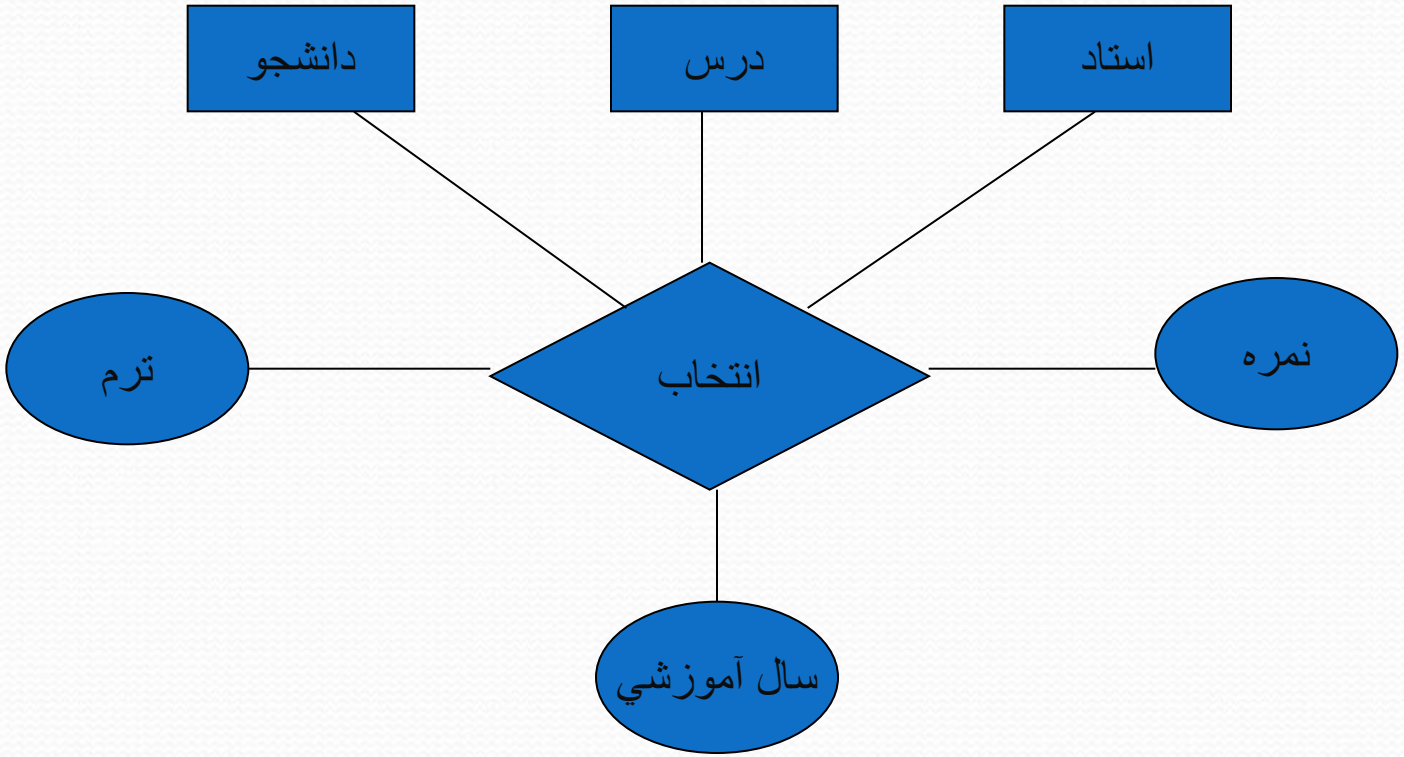
ارتباط "گونه‌ای است از"

وضع مشارکت در ارتباط

مشارکت یک نوع موجودیت در یک نوع ارتباط را الزامی گویند، اگر تمام نمونه‌های آن نوع موجودیت در آن نوع ارتباط شرکت کنند. در غیر این صورت مشارکت غیرالزامی است.

نمایش مشارکت الزامی

تعداد شرکت کنندگان در یک ارتباط را درجه آن ارتباط می گویند.



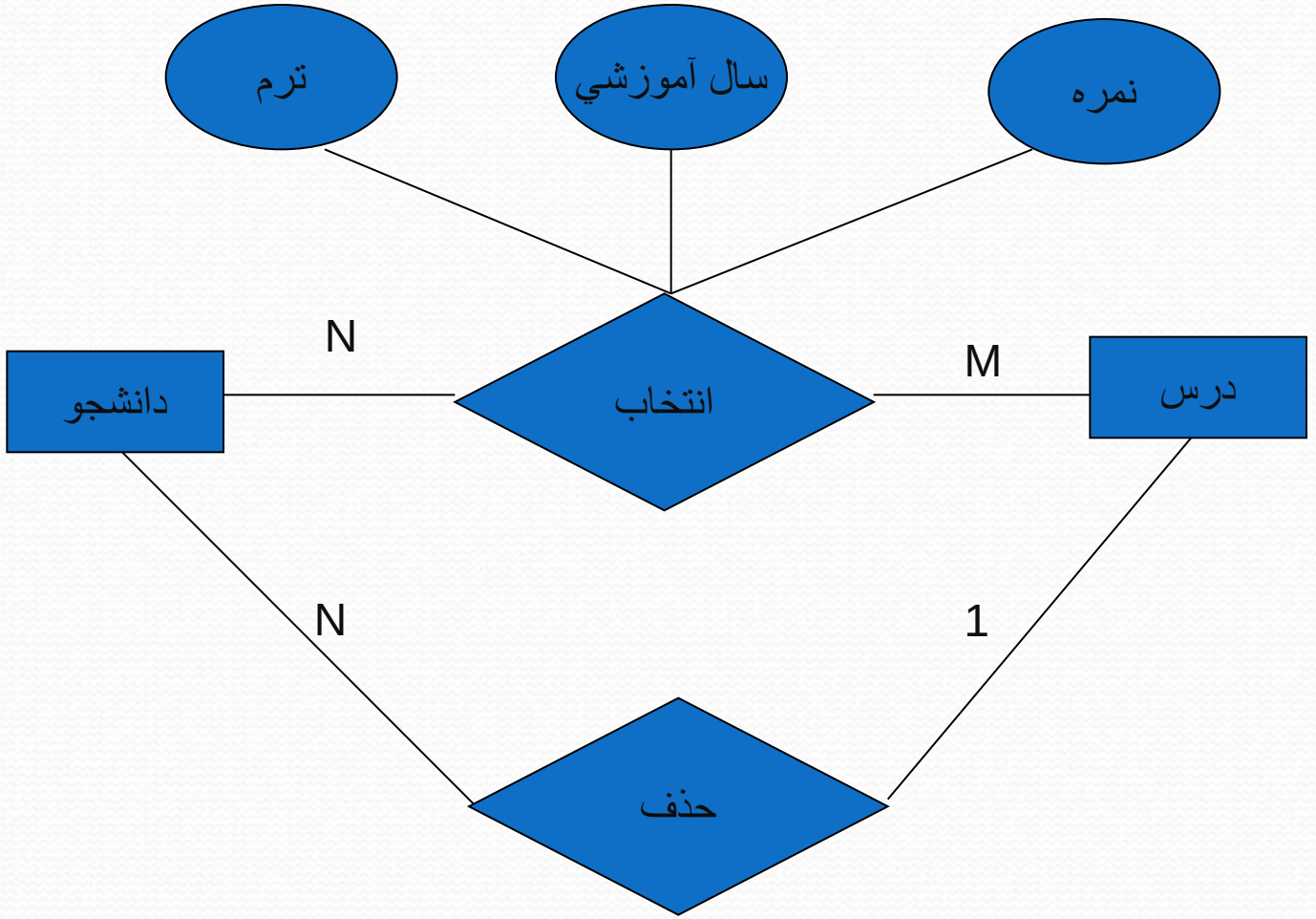
ارتباط بین سه موجودیت

چندی یا ماهیت نوع ارتباط عبارتست از چگونگی تناظر بین دو مجموعه نمونه‌های آن دو نوع موجودیت.

انواع چندی ارتباط:

- یک به یک 1:1
- یک به چند 1:N
- چند به چند N:M

نمایش چندی ارتباط



جلسه سوم

محدودیتهای روش ER

آنچه در این جلسه می خوانید:

۱- مشکلات روش ER(سه نوع دام)
۲- تجزیه و ترکیب
۳- تخصیص و تعمیم
۴- تجمع
۵- وراثت صفت
۶- دسته‌بندی
۷- مراحل مدلسازی معنایی داده‌ها
۸- روش مدلسازی UML
۹- نمادها
۱۰- خصوصیات کلی روش مدلسازی معنایی داده ها

هدفهای کلی: مشکلات و محدودیتهای روش ER

هدفهای رفتاری: دانشجو در پایان این جلسه می‌تواند:

- مشکلات روش ER (انواع دامها) را تشریح کند.
- محدودیتهای روش ER را که در روش EER برطرف شدند، بیان کند.
- مراحل مدلسازی معنایی داده‌ها را توصیف کند.
- مدلسازی UML و نمادهای به کار رفته در آن را شرح دهد.
- خصوصیات کلی مدلسازی معنایی داده‌ها را بیان کند.

مشکلات روش ER

۱- دام حلقه‌ای

هنگامی ایجاد می‌شود که با داشتن مثلا سه ارتباط دو موجودیتی، وجود یک ارتباط سه موجودیتی را نتیجه بگیریم در وضعی که این استنتاج درست نباشد

مشکلات روش ER

۲- دام چندشاخه (چتری)

این نوع دام وقتی ایجاد می‌شود که بین یک نوع موجودیت E و هریک از دیگر انواع موجودیت F، G و ... ارتباط 1:N با مشارکت الزامی وجود داشته باشد، ولی ارتباط بین مثلا F و G، در مدلسازی دیده نشده باشد.

مشکلات روش ER

۳- دام شکاف

هنگامی ایجاد می‌شود که بین دو نوع موجودیت E و F، یک ارتباط باچندی 1:N و مشارکت الزامی وجود داشته باشد، ولی F خود با نوع موجودیت G، ارتباط 1:N با مشارکت غیرالزامی داشته باشد. به دلیل غیر الزامی بودن ارتباط بین F و G، نمی‌توان همه اطلاعات دوموجودیتی در مورد ارتباط بین نمونه‌های دو نوع موجودیت E و G را بدست آورد

محدودیتهای روش ER که در روش EER برطرف گردیدند:

- تجزیه
- ترکیب
- تخصیص
- تعمیم
- تجمع
- وراثت صفت

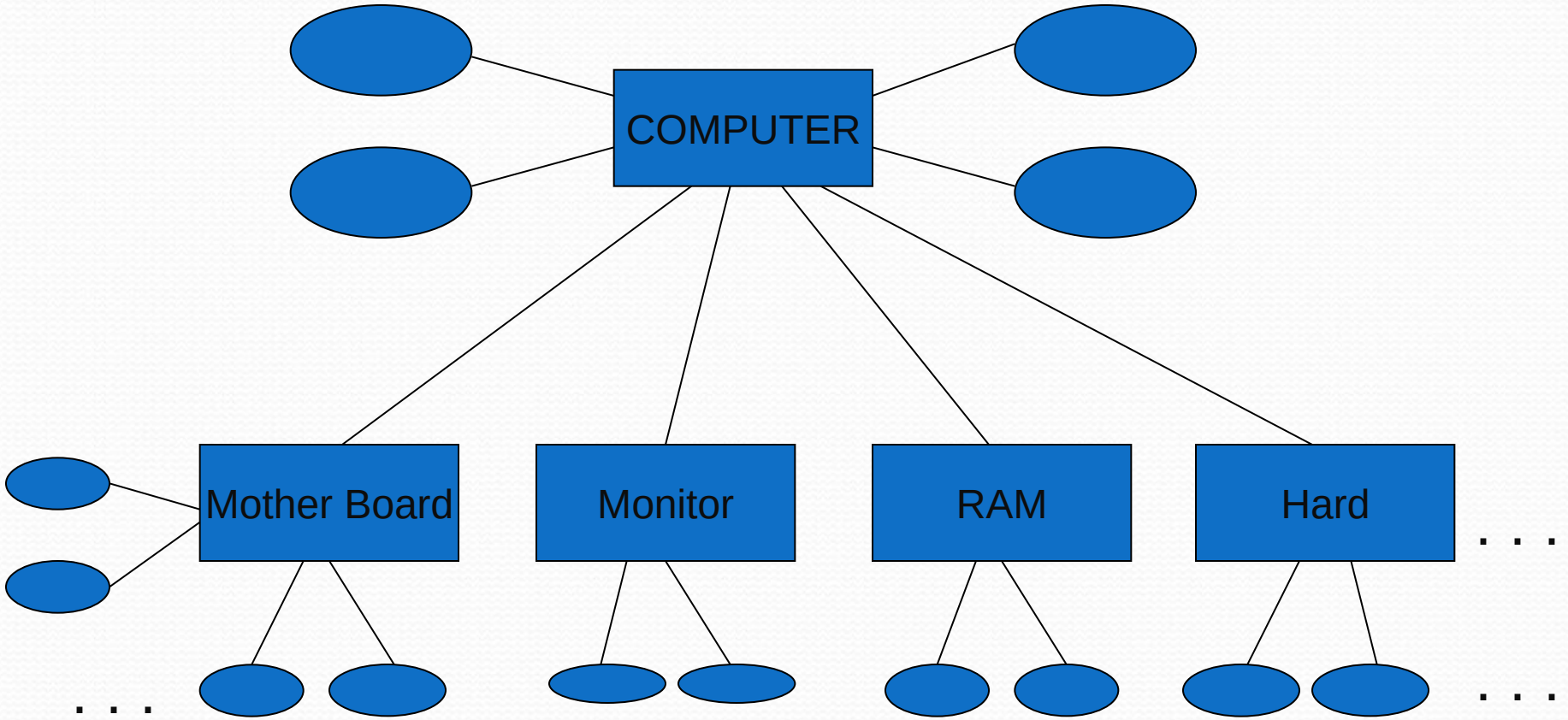
تجزیه

تجزیه یا جداسازی یعنی یک شیئ کل را به اجزاء تشکیل‌دهنده آن تقسیم کنیم. شیئ کل صفات، ساختار و رفتار خود را دارد و هریک از اجزاء نیز صفات، ساختار و رفتار خاص خود را دارند. شیئ کل شامل اجزاء خود است و بین شیئ کل و اجزایش، ارتباط شمول وجود دارد. به این نوع ارتباط در EER، ارتباط "جزئی است از ..." گفته می‌شود.

ترکیب

ترکیب، عکس عمل تجزیه است و در این عمل، با داشتن
Ei(i=1 , 2 , ...) یک نوع موجودیت E را بازشناسی می‌کنیم
به نحوی که Eiها اجزاء تشکیل‌دهنده آن باشند

مثال تجزیه و ترکیب



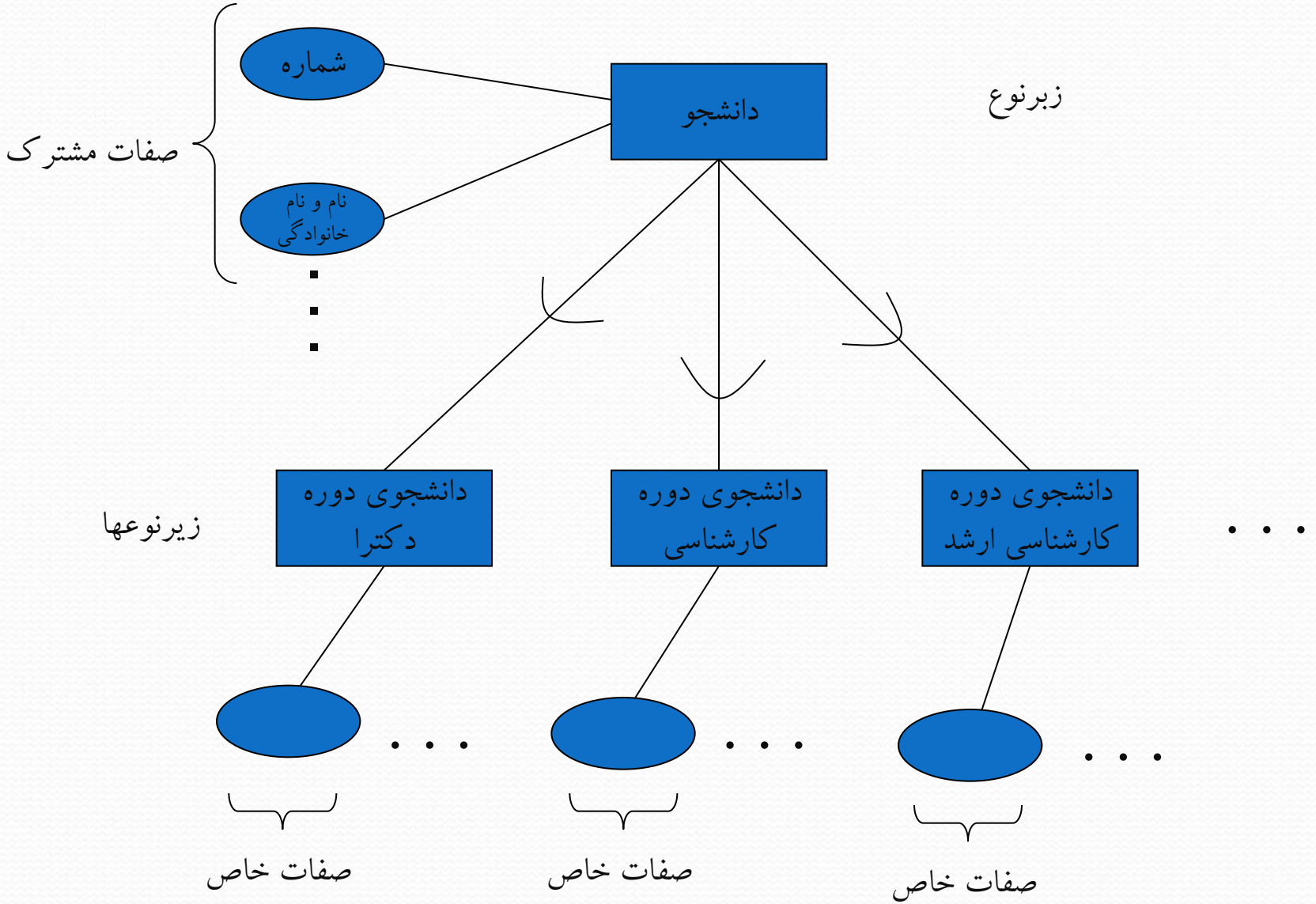
تخصیص

تخصیص عبارتست از مشخص کردن گونه‌های خاص یک شیئ براساس یک یا چند ضابطه مشخص، مثلا اگر شیئ موجود زنده را درنظر بگیریم، سه گونه خاص آن عبارتند از: انسان، حیوان و نبات. در روش EER هر یک نوع موجودیت می‌تواند خود زیرنوع موجودیتهایی داشته باشد. بین هر زیرنوع و زبرنوع ارتباط "گونه‌ای است از ..." وجود دارد.

تعمیم

تعمیم، عکس عمل تخصیص است، به این معنا که با داشتن زیرنوعهای خاص، صفات مشترک بین آنها را در یک مجموعه صفات برای یک زبرنوع موجودیت درنظر می گیریم

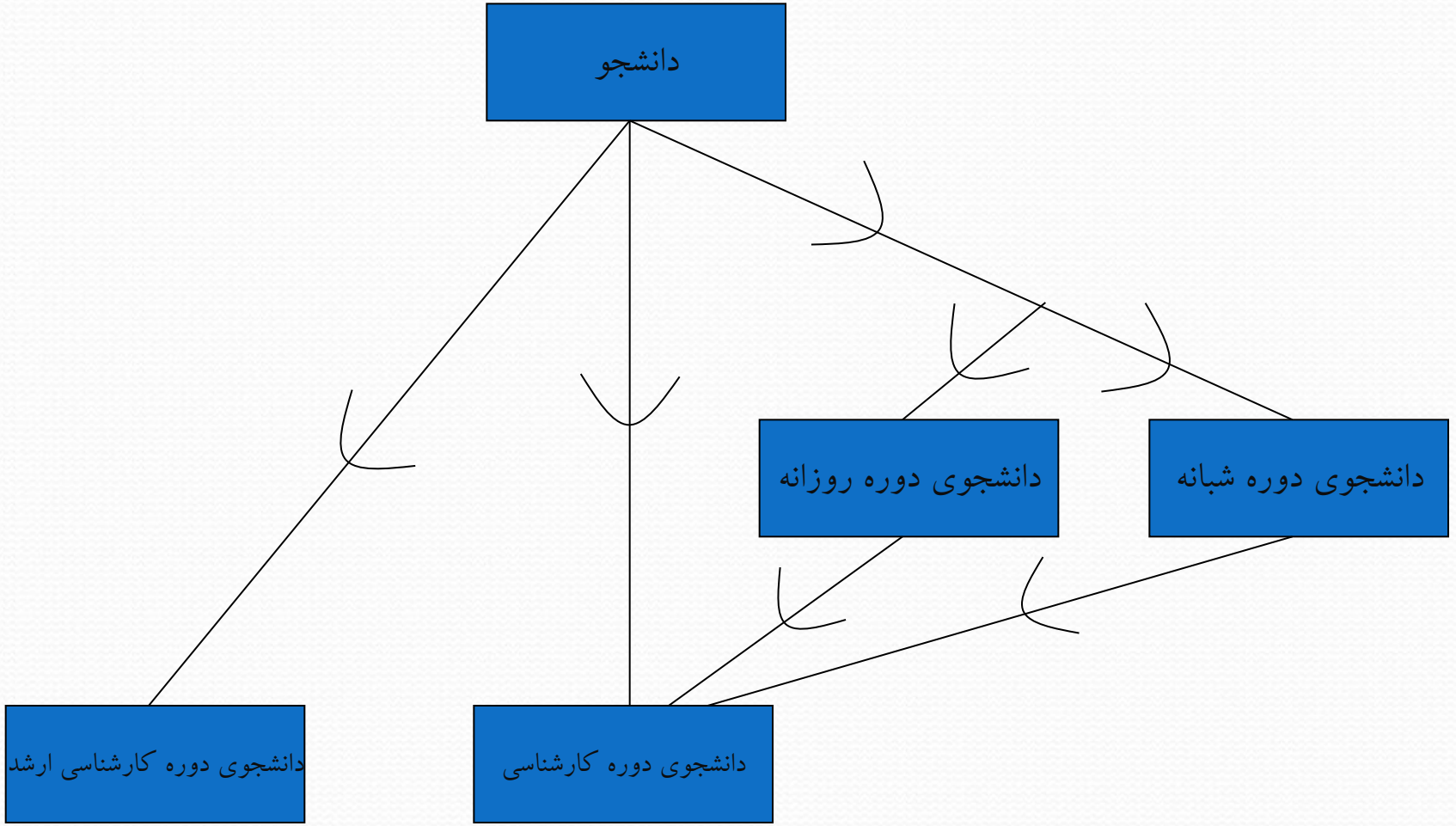
مثال تخصیص و تعمیم



وراثت چندگانه

یک زیرنوع موجودیت، می‌تواند در عین حال زیرنوع یک زبرنوع موجودیت دیگر هم باشد.
با این ترتیب می‌توان مفهوم وراثت چندگانه را در روش EER نمایش داد.

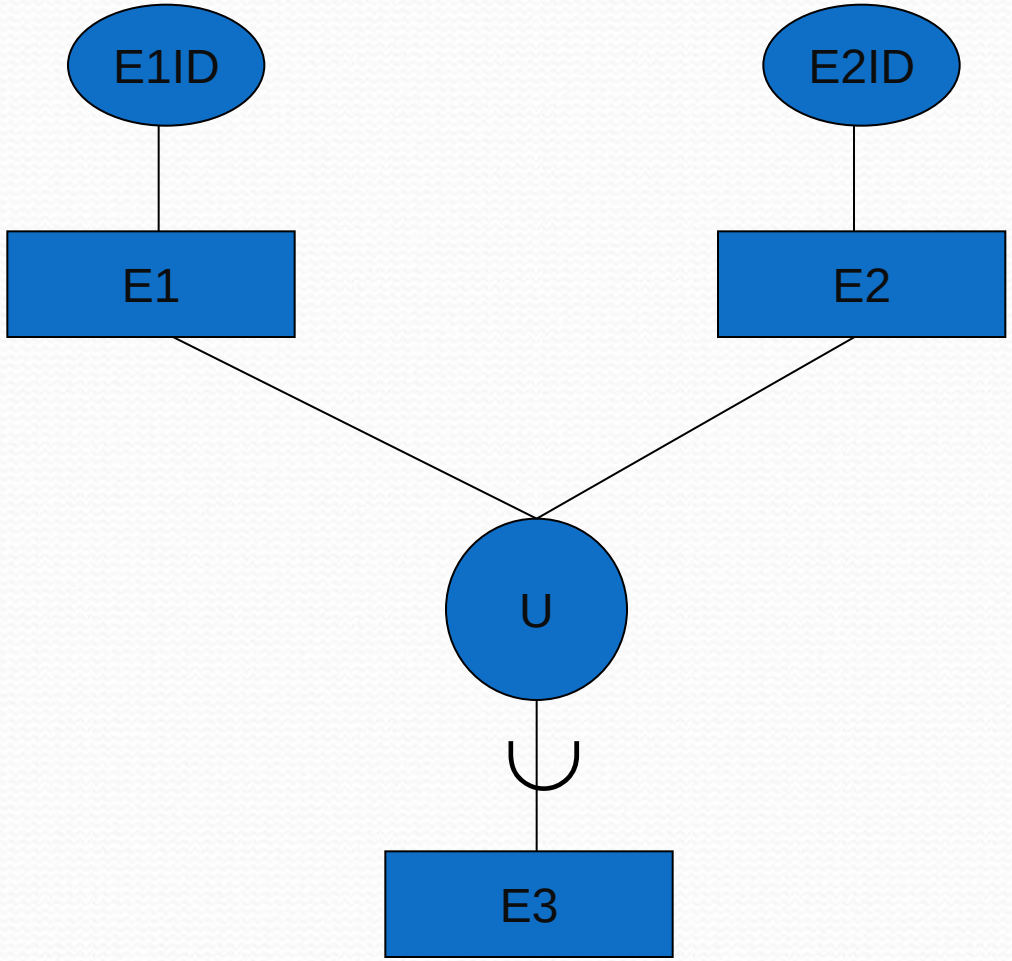
مثال وراثت چند گانه



دسته‌بندی

یک زیرنوع می‌تواند زیرنوع بیش از یک زبرنوع باشد. ممکن است زبرنوعهای این زیرنوع، از یک نوع نباشند. به این زیرنوع اصطلاحا دسته (طبقه) گویند. برای نمایش دسته، از نماد **U** استفاده می‌شود.

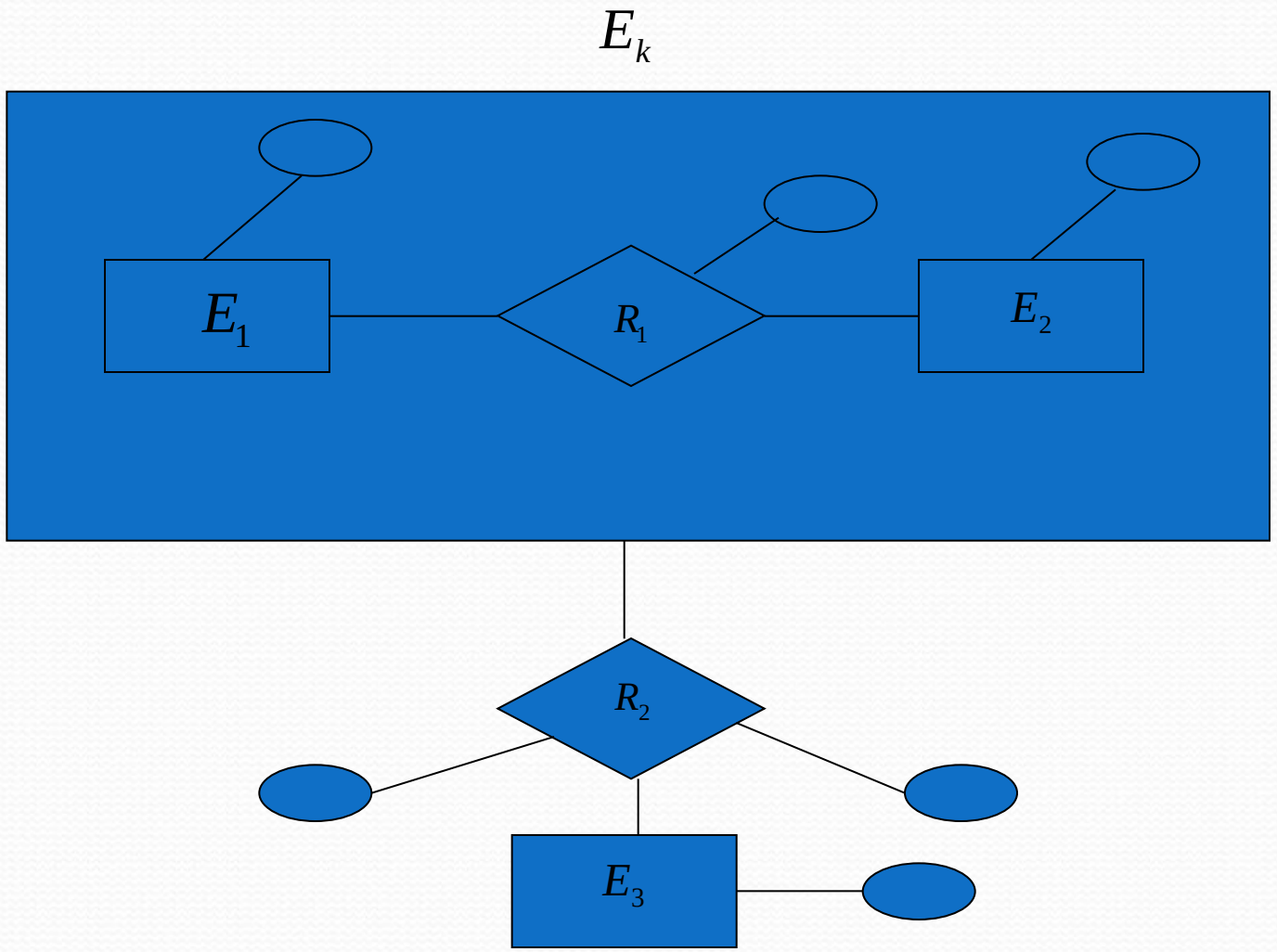
دسته‌بندی



تجمع

تجمع عبارتست از ساختن یک نوع موجودیت جدید با دیدن دو یا بیش از دو نوع موجودیت، که خود باهم در یک ارتباط شرکت دارند، به صورت یک نوع موجودیت واحد. در واقع مجموعه‌ای از موجودیتهای مرتبط را باهم مجتمع کرده و به عنوان یک نوع موجودیت واحد، درنظر می گیریم و این نوع موجودیت واحد خود می‌تواند با نوع موجودیت دیگری ارتباط داشته باشد.

نمایش تجمع



مراحل مدلسازی معنایی داده‌ها

۱- مطالعه، تحلیل و شناخت محیط

۲- برآورد خواسته‌ها و نیازهای اطلاعاتی و پردازشی همه کاربران و تشخیص محدودیتهای معنایی

۳- بازشناسی انواع موجودیتهای مطرح و تعیین وضع هریک

۴- تعیین مجموعه صفات هر نوع موجودیت

مراحل مدلسازی معنایی داده‌ها

۵- بازشناسی انواع ارتباطات بین انواع موجودیتها، تشخیص نوع مشارکت و چندی ارتباط

۶- رسم نمودار ER

۷- فهرست کردن پرسشهایی که پاسخ آنها از نمودار ER بدست می‌آید.

۸- وارسی مدلسازی انجام شده تا اطمینان حاصل شود که مدلسازی پاسخگوی نیاز کاربران است.

روش مدلسازی UML

در این روش از چند نمودار برای نمایش مدلسازی و طراحی نرم‌افزار استفاده می‌شود

مفاهیم اصلی در این مدلسازی:

- رده
- صفت
- بستگی

تناظر بین مفاهیم UML و مفاهیم EER	
مفهوم در UML	مفهوم در EER
رده	نوع موجودیت
شیئ	نمونه موجودیت
صفت	صفت
بستگی	ارتباط
پیوند	نمونه ارتباط
بستگی انعکاسی	ارتباط بازگشتی
بستگی مقید	نوع موجودیت ضعیف
میدان ساختمند	صفت مرکب
صفت پیوند	صفت ارتباط
چندی بستگی	چندی ارتباط

نمادها

- رده با یک مربع یا مستطیل نشان داده می‌شود وبه شکل زیر به سه قسمت تقسیم می‌شود:

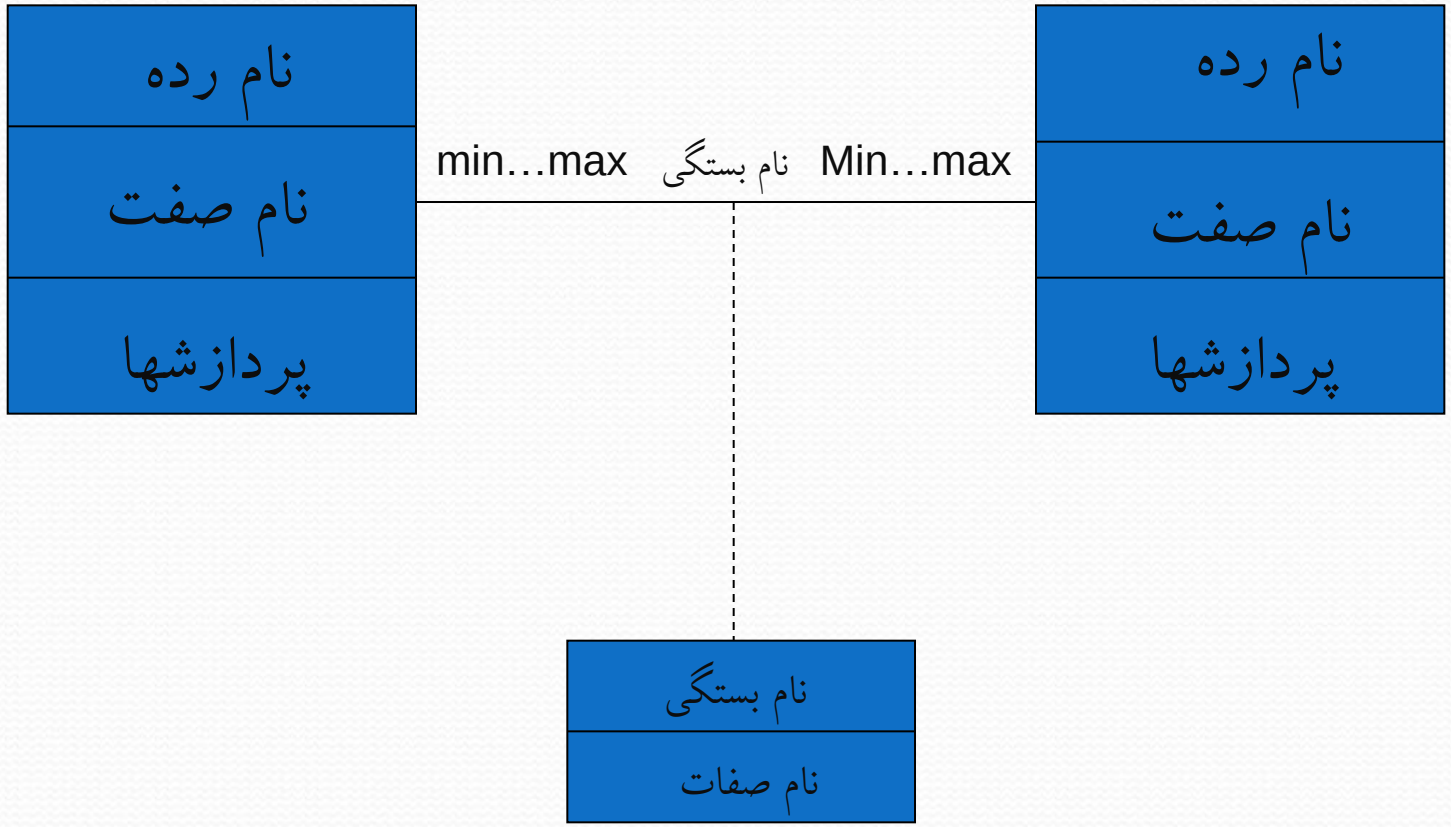
نمادها

- بستگی بین دو رده، به صورت یک خط متصل کننده دو رده نمایش داده می‌شود و نام ارتباط روی خط نوشته می‌شود.
- تجمع به صورت زیر نشان داده می‌شود:

- چندی بستگی به صورت min...max نوشته می‌شود. اگر به جای max علامت ستاره باشد، به این معنا است که مقدار max محدودیت ندارد.
- صفت چندمقداری به صورت یک رده جداگانه نشان داده می‌شود، ولی فاقد قسمت مربوط به پردازش است.

- نام میدان بعد از نام صفت نوشته می‌شود و بین این دو نام علامت : گذاشته می‌شود.
- صفت پیوندی در یک مربع (مستطیل) نوشته می‌شود و این مربع با خط‌چین به خط نشان‌دهنده بستگی متصل می‌شود. نام پیوند و نام صفات پیوند در دو قسمت از این مربع گذاشته می‌شود.

نمایش صفت پیوند در UML



خصوصیات کلی روش مدلسازی معنایی داده‌ها

- گویایی
- صوری بودن
- سادگی مفاهیم
- قابلیت نمایش نموداری
- ایجاز
- جامع بودن مفاهیم
- قابلیت نمایش ساختار، حالت و رفتار نوع موجودیت
- گسترش‌پذیری

جلسه چهارم

پایگاه داده در محیط انتزاعی

آنچه در این جلسه می خوانید:

۱- پایگاه داده‌ها در محیط انتزاعی
۲- سطوح محیط انتزاعی
۳- گونه‌های موجود ساختار داده‌ای
۴- مفهوم ساختار داده‌ای در سطوح مختلف پایگاه داده‌ها
۵- ساختار داده‌ای رابطه‌ای
۶- شمای پایگاه جدولی
۷- عملیات در پایگاه جدولی

آنچه در این جلسه می خوانید:

۸- عملگرهای جبر رابطه‌ای
۹- برخی ویژگیهای ساختار داده‌ای جدولی
۱۰- ساختار داده‌ای سلسله‌مراتبی
۱۱- برخی ویژگیهای ساختار داده‌ای سلسله‌مراتبی
۱۲- ساختار داده‌ای شبکه‌ای
۱۳- مجموعه کوداسیل
۱۴- برخی ویژگیهای ساختار داده‌ای شبکه‌ای

هدفهای کلی: آشنایی با پایگاه داده در محیط انتزاعی

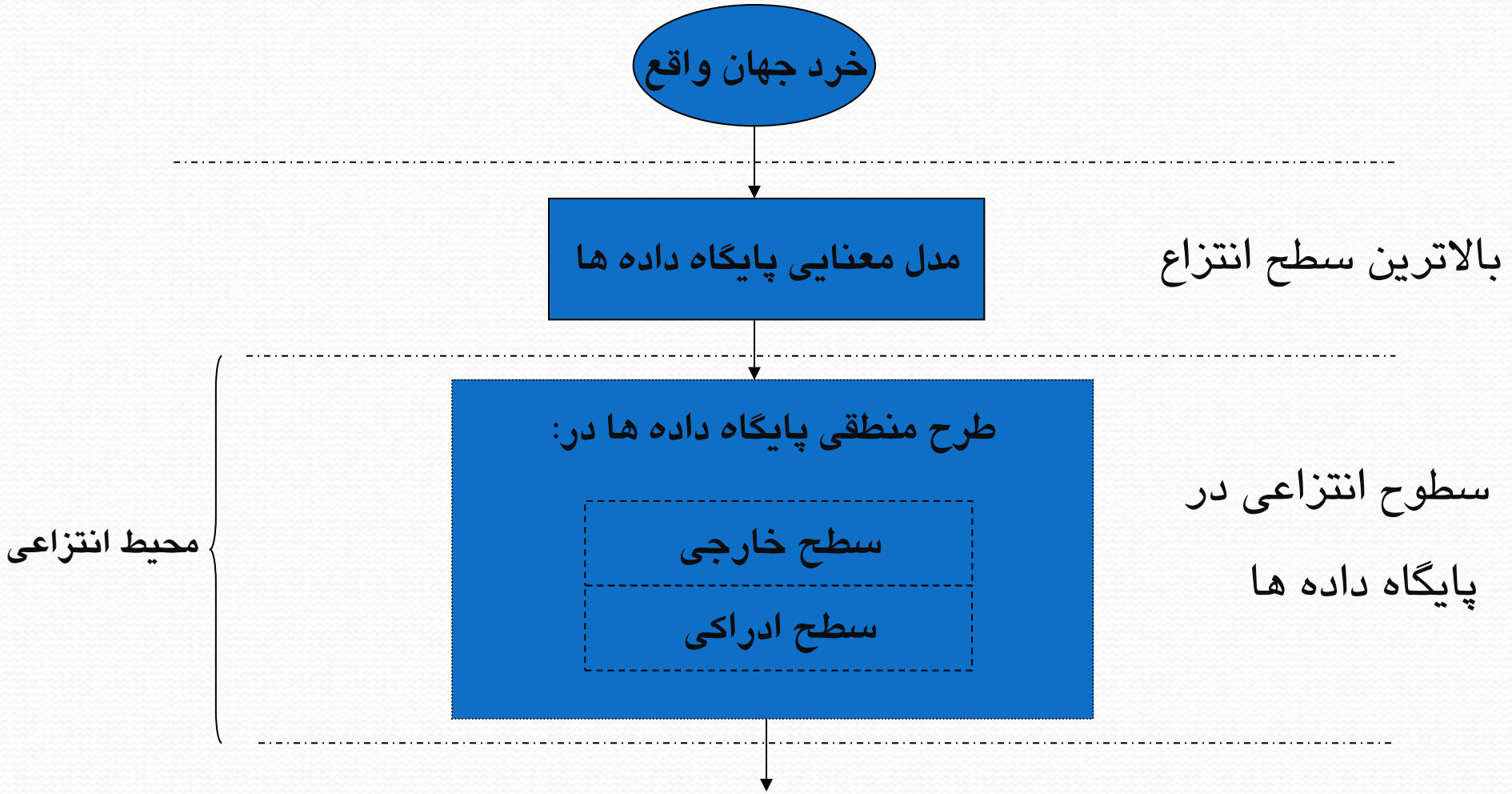
هدفهای رفتاری: دانشجو در پایان این جلسه می‌تواند:

- پایگاه داده در محیط انتزاعی و سطوح این محیط را بیان کند.
- ساختار داده‌ای رابطه‌ای و ویژگیهای آن را بیان کند.
- ساختار داده‌ای شبکه‌ای و ویژگیهای آن را بیان کند.
- ساختار داده‌ای سلسله‌مراتبی و ویژگیهای آن را بیان کند.

پایگاه داده‌ها در محیط انتزاعی

در مقوله پایگاه داده‌ها وقتی از محیط انتزاعی سخن می گوییم، منظور محیطی است فراتر از محیط فایلینگ منطقی و فایلینگ فیزیکی. مفاهیمی که در این محیط مطرح می‌شوند باید از جنبه‌های فایلینگ پایگاه مستقل و ماهیتا انتزاعی باشند.

سطوح محیط انتزاعی



ساختار داده‌ای امکانی است برای نشان دادن داده‌های در مورد انواع موجودیت‌ها و انواع ارتباطات بین آنها

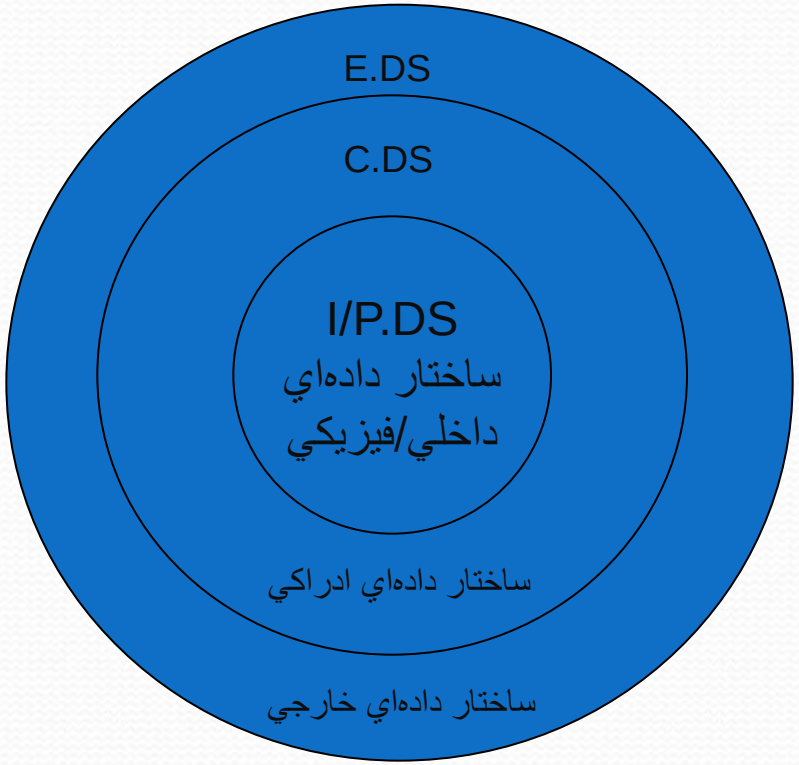
هر ساختار داده‌ای حداقل یک عنصر ساختاری اساسی دارد.

پایگاه داده‌ها در محیط انتزاعی، گردایه‌ای است از نمونه‌های متمایز عنصر (عناصر) ساختاری اساسی یک ساختار داده‌ای مشخص.

گونه‌های موجود ساختار داده‌ای

- ساختار داده‌ای رابطه‌ای
- ساختار داده‌ای سلسله‌مراتبی
- ساختار داده‌ای شبکه‌ای
- ساختار داده‌ای جعبه‌ای
- ساختار داده‌ای هایپرگرافی
- ساختار داده‌ای لیستهای وارون

مفهوم ساختار داده‌ای در سطوح مختلف پایگاه داده‌ها



ساختار داده‌ای رابطه‌ای

رابطه مفهومی ریاضی است. اما از دید کاربر، رابطه نمایشی جدولی دارد.

مفاهیم ساختار جدولی عبارتند از:

جدول سطر ستون

عنصر ساختاری اساسی در این ساختار داده‌ای، جدول است.

شمای پایگاه جدولی

شمای پایگاه داده‌ها عبارتست از تعریف (توصیف) ساختهای انتزاعی طراحی‌شده و نوعی برنامه است شامل دستورات تعریف داده‌ها و کنترل داده‌ها، و دستورات عملیات در داده‌ها در آن وجود ندارد.

عملیات در پایگاه جدولی

درج (INSERT)

حذف (DELETE)

بهنگام‌سازی (UPDATE)

بازیابی (SELECT)

مثال درج یک نمونه دانشجو

```
INSERT
INTO STT
VALUES(‘7413673’,’AHMADI’,’bs’,’Comp’,’D222’)
```

مثال حذف یک نمونه درس

```
DELETE
FROM COT
WHERE COID=‘COM777’
```

مثال حذف چند سطر

```
DELETE
FROM STT
WHERE STDEG='ms'
```

با این دستور مشخصات دانشجویان دوره کارشناسی ارشد از جدول STT حذف می‌شود

مثال بهنگام‌سازی تک سطر

```
UPDATE COT
SET COID='Com303'
WHERE COID='Com202';
```

با این دستور شماره درس Com202 عوض می‌شود.

مثال بهنگام‌سازی چند سطر

```
UPDATE COT
SET CREDIT='1'
WHERE COTYPE='Lab';
```

با این دستور، تعداد واحد تمام درسهای آزمایشگاهی، یک می‌شود.

مثال بازیابی

```
SELECT DOID
FROM STCOT
WHERE STID='76140444'
```

با این دستورشماره درسهای یک دانشجو بازیابی می‌شود.

عملگرهای جبر رابطه‌ای که برای بازیابی از محیط جدولی بکار می‌روند:

۱- عملگر گزینش (تحدید) RESTRICT: زیرمجموعه‌ای افقی (تعدادی سطر) را بازیابی می کند.

۲- عملگر پرتو PROJECT: زیرمجموعه‌ای عمودی (تعدادی ستون) را بازیابی می کند.

۳- عملگر پیوند JOIN: دو جدول را بهم پیوند می‌زند، یعنی سطرهایی که شرایط مورد نظر را داشته باشند، با یکدیگر پیوند می‌شوند.

برخی ویژگیهای ساختار داده‌ای جدولی

۱- از نظر کاربر نمایش ساده‌ای دارد.

۲- محیطش مسطح است.

۳- عنصر ساختاری اساسی آن جدول است.

۴- همه چیز با فقره داده‌های ساده نمایش داده می‌شود.

۵- ارتباطات با چندیهای مختلف در آن قابل نمایش است.

۶- منطق بازیابی آن ساده است.

۷- ساختار منطقی دستور بازیابی آن ساده است.

۸- برای پرسشهای قرینه، رویه پاسخگوی قرینه دارد.

۹- مبنای تئوریک قوی دارد.

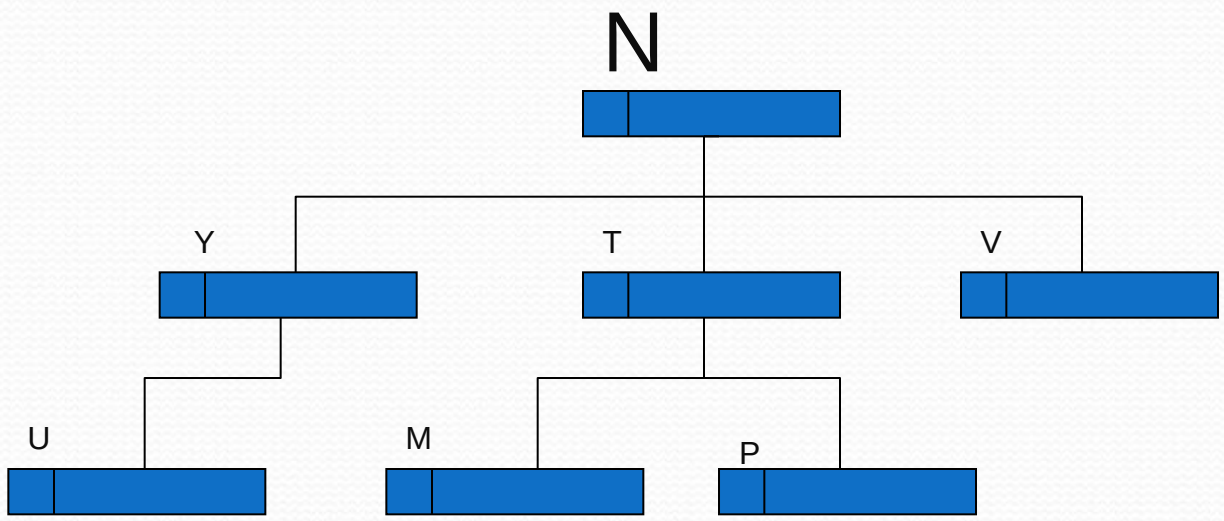
ساختار داده‌ای سلسله‌مراتبی

عناصر ساختاری اساسی

۱- نوع رکورد ۲- نوع پیوند پدر- فرزندی

نوع رکورد برای نمایش نوع موجودیت به کار می‌رود.

ارتباط بین دو نوع موجودیت به وسیله پیوند پدر—فرزندی نمایش داده می‌شود.



برخی ویژگیهای ساختار داده‌ای سلسله‌مراتبی

۱- سادگی نمایش ساختار جدولی را ندارد.

۲- مبنای ریاضی ندارد.

۳- دو عنصر ساختاری اساسی دارد.

۴- ارتباط یک به چند را نمایش می‌دهد.

۵- نمایش ارتباط چند به چند در آن دشوار است.

۶- ساخت منطقی رویه بازیابی آن به سادگی منطق رویه بازیابی در ساختار جدولی نیست.

ادامه

برخی ویژگیهای ساختار داده‌ای سلسله‌مراتبی

۷- در عملیات ذخیره‌سازی مشکلاتی دارد.

۸- تقارن ساختار جدولی را ندارد.

۹- تعدادی قاعده جامعیت ذاتی دارد.

۱۰- در مواقعی که در ذخیره‌سازی نمونه‌های فرزند، افزونگی پدید آید، پایگاه در معرض ناسازگاری قرار می‌گیرد.

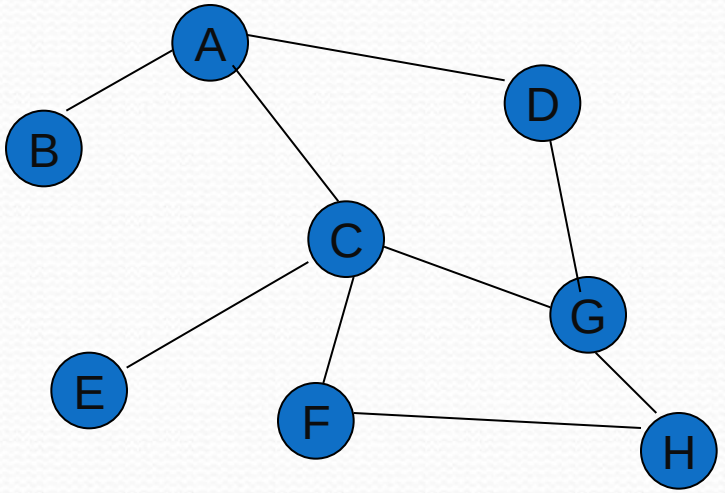
ساختار داده‌ای شبکه‌ای

عناصر ساختاری اساسی

۱- نوع رکورد ۲- نوع مجموعه

نوع رکورد برای نمایش نوع موجودیت به کار می‌رود.

نوع مجموعه برای نمایش ارتباط 1:N بین دو (چند) نوع موجودیت به کار می‌رود.



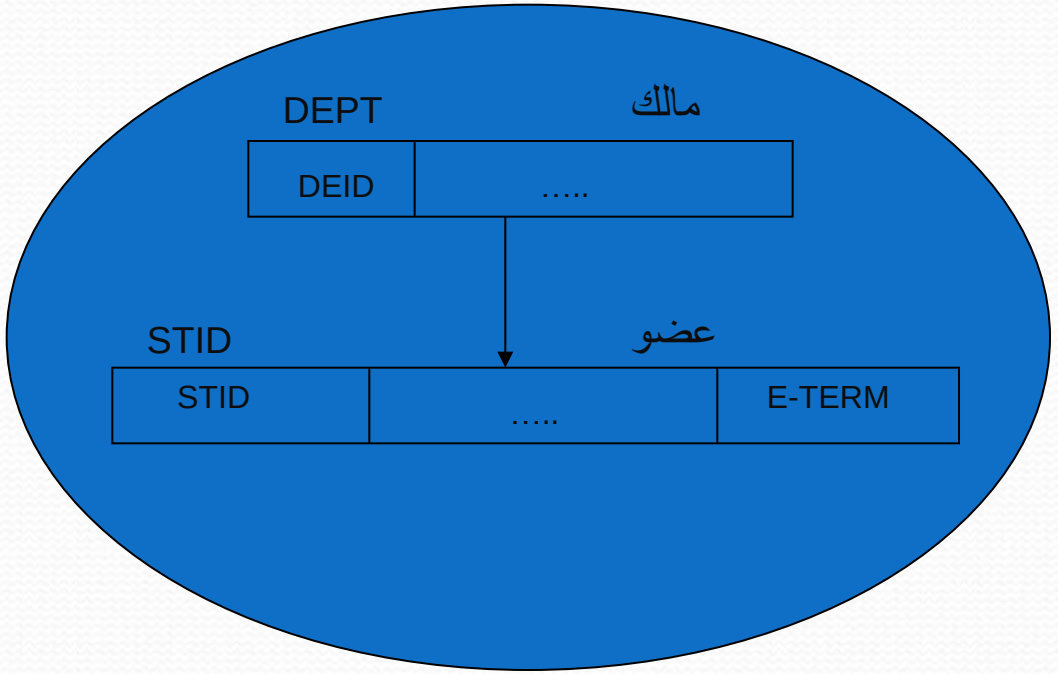
نوع مجموعه (مجموعه کوداسیل) از سه جزء تشکیل شده است:

۱- نام مجموعه

۲- یک نوع رکورد مالک

۳- یک نوع رکورد عضو

یک نوع مجموعه کوداسیل



برخی ویژگیهای ساختار داده‌ای شبکه‌ای

۱- سادگی ظاهری ساختار داده‌ای جدولی را ندارد.

۲- مبنای ریاضی ندارد.

۳- دو عنصر ساختاری اساسی دارد.

۴- ماهیتا خاص نمایش ارتباطات "یک به چند" نیست.

۵- ساخت منطقی دستور بازیابی آن پیچیده‌تر از ساختارهای دیگر است.

ادامه

برخی ویژگیهای ساختار داده‌ای شبکه‌ای

۶- مثل ساختار داده‌ای جدولی تقارن دارد.

۷- خطر بروز ناسازگاری داده‌ها نسبت به ساختار سلسله‌مراتبی، کمتر است.

۸- قواعد جامعیت ذاتی دارد.

۹- به علت حجم زیاد اشاره‌گرها، ایجاد یا اصلاح آنها می‌تواند سبب بروز فزونکاری در سیستم شود.

۱۰- بعضی آنومالیهای مدل سلسله‌مراتبی در عملیات ذخیره‌سازی را ندارد.

جلسه پنجم

معماری پایگاه داده‌ها

آنچه در این جلسه می خوانید:

۱- معماری پشنهادی ANSI
۲- دید ادراکی (مفهومی)
۳- دید خارجی
۴- دید داخلی
۵- سطوح معماری در محیط ناپایگاهی و محیط پایگاهی و نقش DBMS در ایجاد، مدیریت و پردازش فایلها
۶- کاربر
۷- زبان میزبان

آنچه در این جلسه می خوانید:

۸- نقش افزایش تعداد زبانهای میزبان مورد پذیرش DBMS
۹- زبان داده‌ای فرعی
۱۰- دستورهای DSL برای سه سطح معماری پایگاه داده‌ها
۱۱- تقسیم‌بندی زبان داده‌ای فرعی از نظر نیاز به زبان میزبان
۱۲- روند کلی مرحله کامپایل برنامه دوزبانی
۱۳- ویژگیهای زبان داده‌ای فرعی

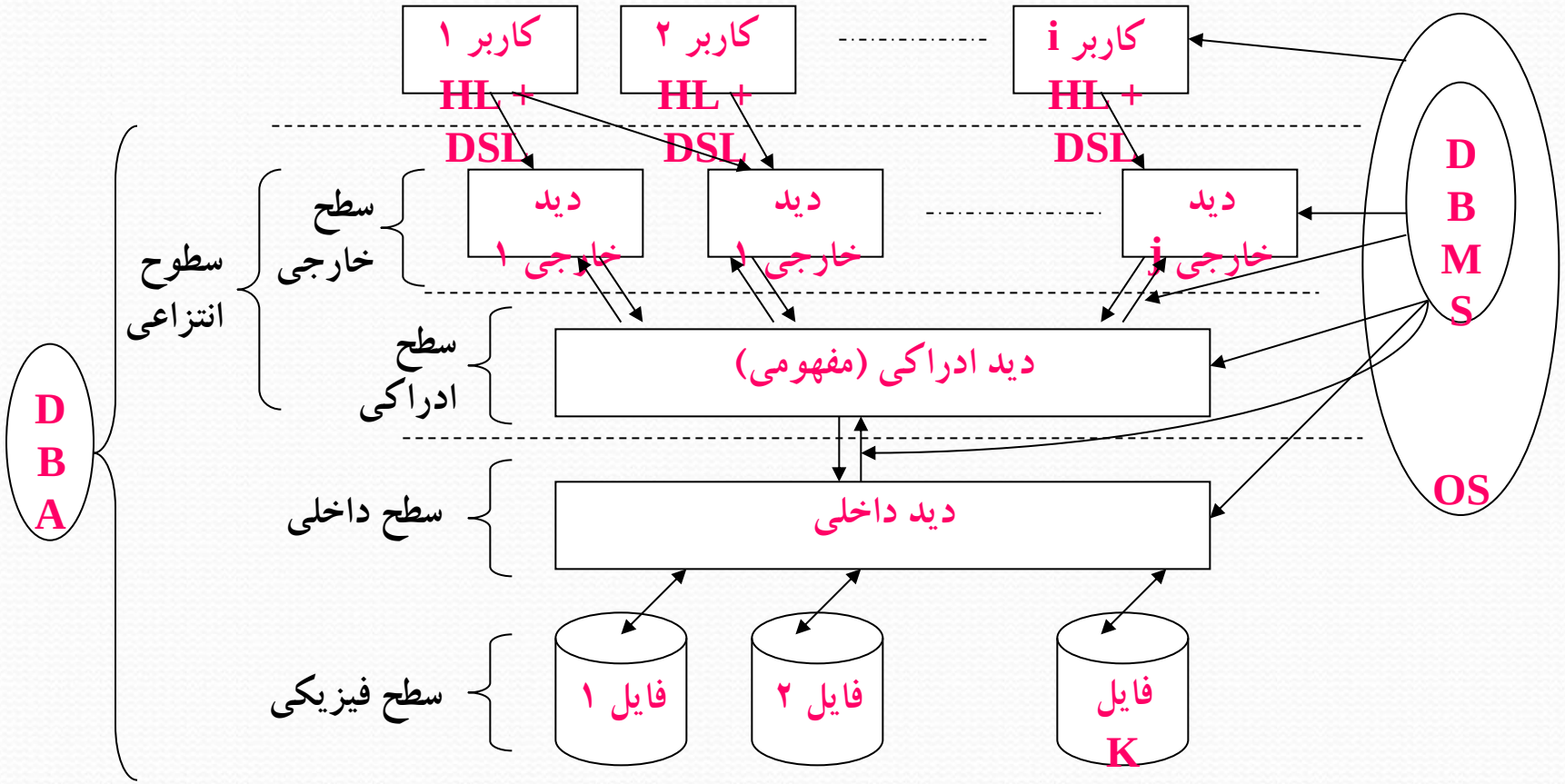
هدفهای کلی: آشنایی با معماری پایگاه داده‌ها

هدفهای رفتاری: دانشجو در پایان این جلسه می‌تواند:

- معماری پایگاه داده‌ها را تشریح کند.
- دیدهای ادراکی، خارجی و داخلی را توضیح دهد.
- سطوح معماری در محیط ناپایگاهی و پایگاهی را مقایسه کند.
- کاربر، زبان میزبان، زبان داده‌ای فرعی و ویژگیهای آن را بیان نماید.

معماری پایگاه داده‌ها

معماری پشنهادی ANSI



دید ادراکی (مفهومی)

۱- دید طراح پایگاه داده‌ها نسبت به داده‌های ذخیره‌شده است.

۲- جامع است.

۳- در یک محیط انتزاعی مطرح است.

۴- با عناصر ساختاری اساسی همان ساختار داده‌ای طراحی می‌شود.

۵- شمای ادراکی نوعی برنامه است حاوی دستورات تعریف و کنترل داده‌ها. سطح ادراکی در واقع همین شمای ادراکی است.

۶- شمای ادراکی به سیستم داده می‌شود و در کاتالوگ سیستم نگهداری می‌شود.

دید خارجی

۱- دید کاربر خاص نسبت به داده‌های ذخیره‌شده در پایگاه داده است.

۲- جزئی است.

۳- در سطح انتزاعی مطرح است.

۴- روی دید ادراکی طراحی و تعریف می‌شود.

۵- شمای خارجی نوعی برنامه است حاوی دستورات تعریف و کنترل داده‌ها در سطح خارجی که توسط کاربر این سطح نوشته می‌شود.

دید خارجی

۶- به تعریف مجموعه دیدهای خارجی کاربر، سطح خارجی گفته می‌شود.

۷- هر کاربر می‌تواند تعدادی دید داشته باشد.

۸- چند کاربر می‌توانند در یک دید مشترک باشند.

دید داخلی

۱- دید DBMS و طراح پایگاه داده‌ها است و در سطحی پایین‌تر از سطح ادراکی، نسبت به کل داده‌های ذخیره‌شده است.

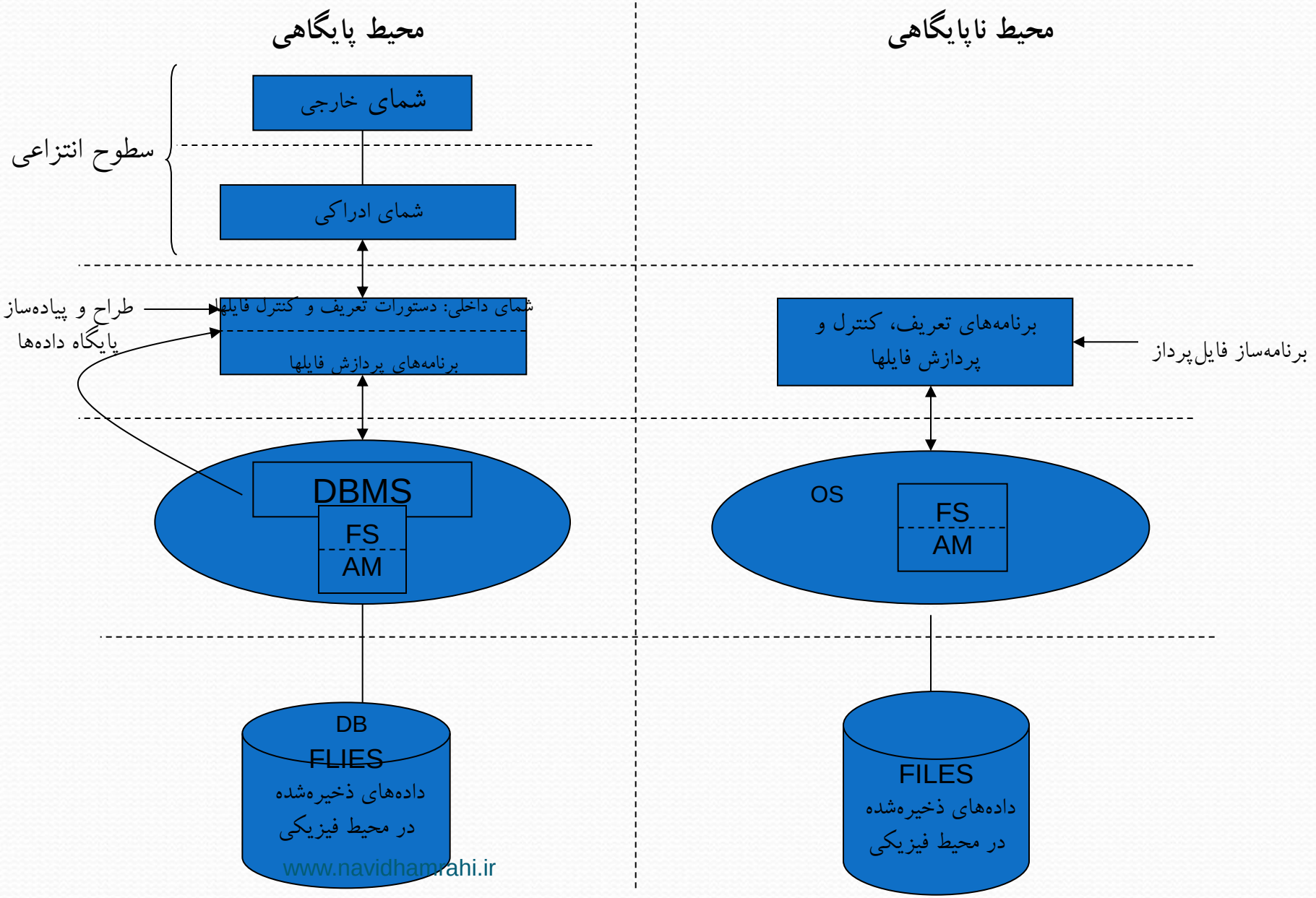
۲- در سطح فایلینگ منطقی مطرح است.

۳- مبتنی بر یک ساختار فایل است که با نظر طراح پایگاه طراحی می‌شود و به طراحی، طراحی فیزیکی موسوم است.

۴- در سطح داخلی پایگاه داده‌ها، فایلینگ منطقی تعریف می‌شود.

۵- شمای داخلی نوعی برنامه است که توسط خود DBMS تولید می‌شود و شرح فایلینگ منطق پایگاه است.

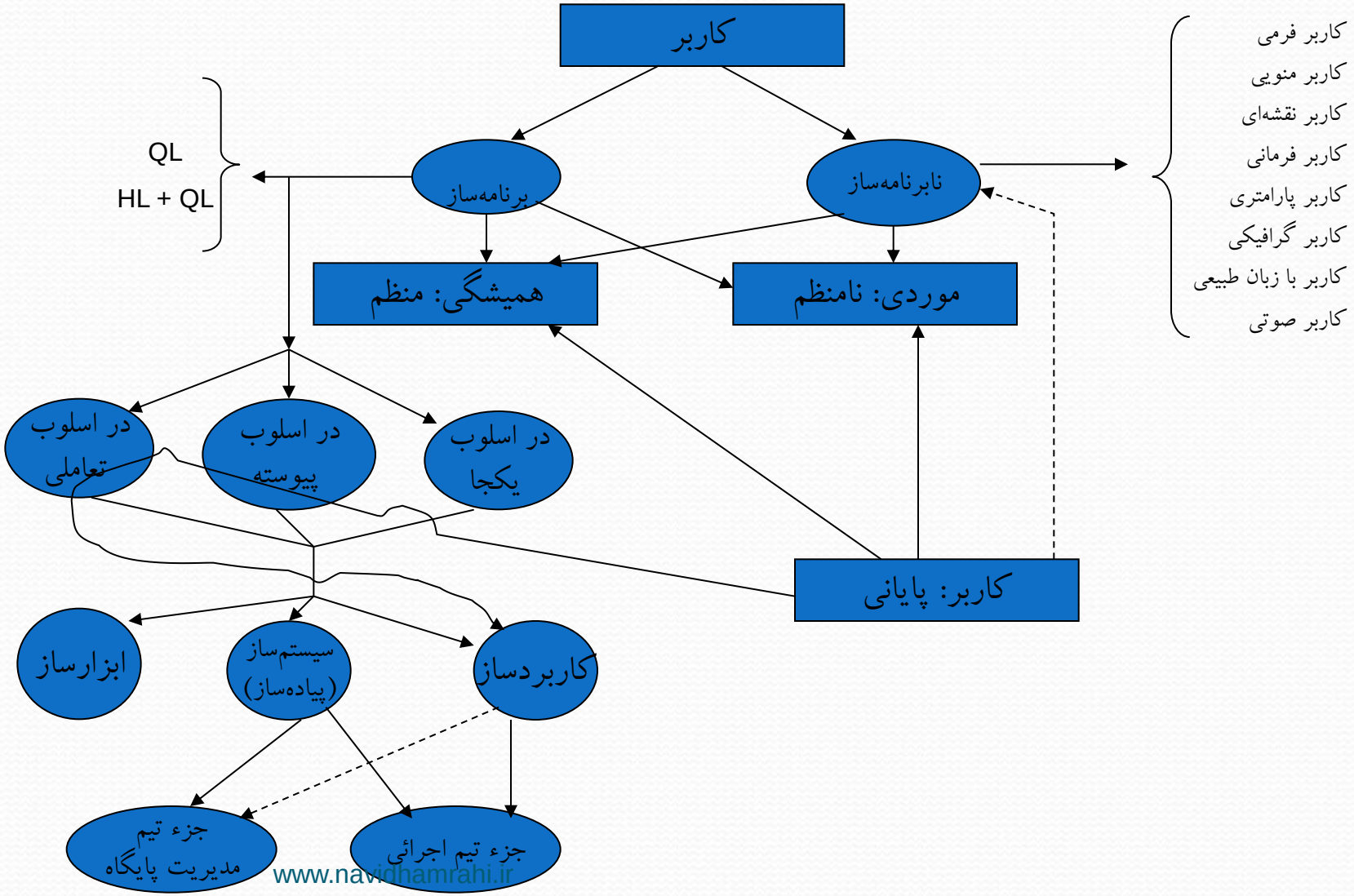
سطوح معماری در محیط ناپایگاهی و محیط پایگاهی و نقش DBMS در ایجاد، مدیریت و پردازش فایلها



www.navidhamrahi.ir

کاربر

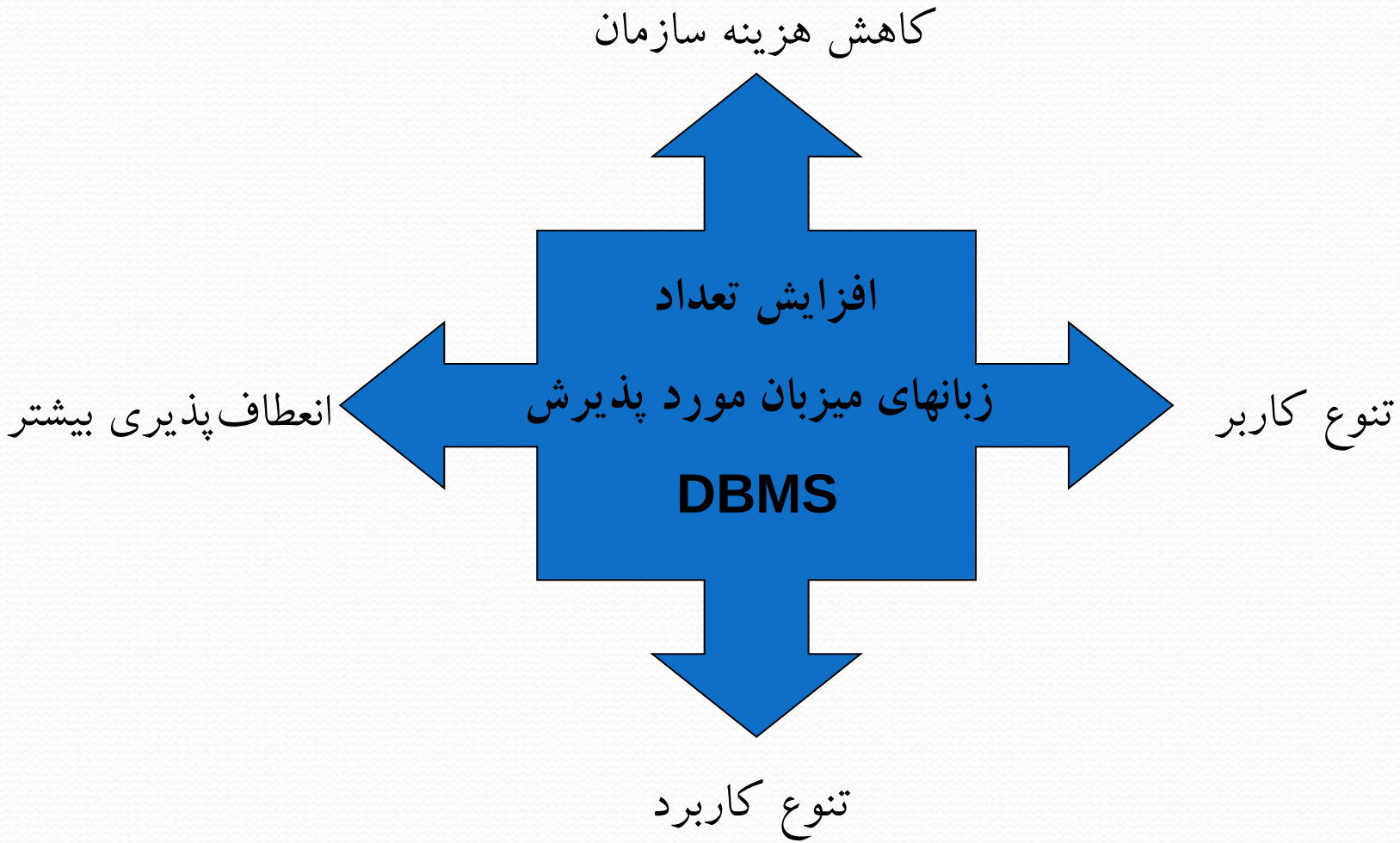
در معنای عام، هر استفاده‌کننده از پایگاه داده‌ها را کاربر گوییم.



www.navidhamrahi.ir

زبان میزبان

یکی از زبانهای برنامه‌سازی متعارف مانند کوبول، PL1، فرترن، پاسکال، C و زبانهایی مثل ADA، LISP، JAVA و نیز زبان اسمبلی است.


کاهش هزینه سازمان
افزایش تعداد
زبانهای میزبان مورد پذیرش
DBMS
انعطاف‌پذیری بیشتر
تنوع کاربر
تنوع کاربرد

زبان داده‌ای فرعی

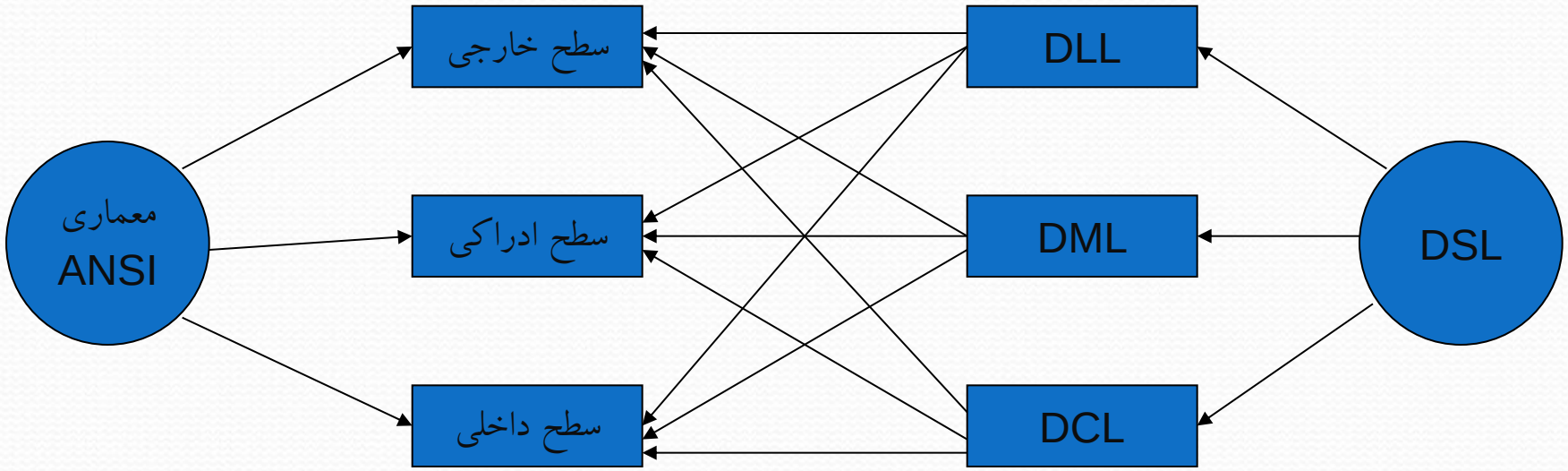
شامل سه دسته دستور زیر است:

۱- دستورات تعریف داده‌ها Data Definition Language (DDL)

۲- دستورات عملیات روی داده‌ها Data Manipulation Language (DML)

۳- دستورات کنترل داده‌ها Data Control Language (DCL)

دستورهای DSL برای سه سطح معماری پایگاه داده‌ها



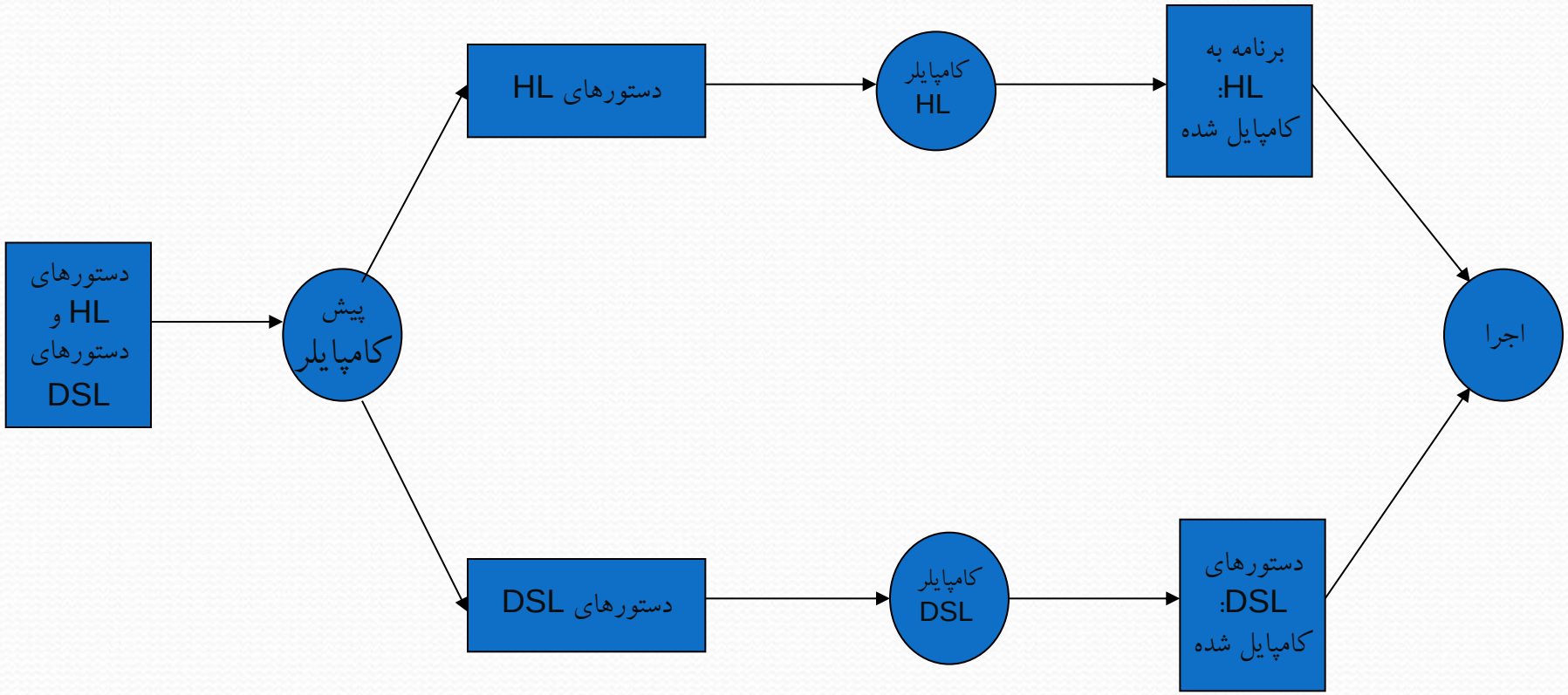
تقسیم‌بندی زبان داده‌ای فرعی از نظر نیاز به زبان میزبان

مستقل (I.DSL): به زبان میزبان نیاز ندارد و به صورت تعاملی استفاده می‌شود. در واقع یک زبان پرس وجو است.

ادغام‌شدنی (E.DSL): دستورهایش در متن برنامه‌ای به زبان میزبان به کار می‌رود و مستقلا قابل استفاده نیست.

هم مستقل و هم ادغام‌شدنی (I/E.DSL)

روند کلی مرحله کامپایل برنامه دوزبانی



ویژگیهای زبان داده‌ای فرعی

۱- تعداد دستورهایش باید کم باشد.

۲- دستورهایش باید شبه زبان طبیعی باشد.

۳- یادگیری و استفاده آن باید ساده باشد.

۴- در طراحی آن باید اصل وحدت دستور رعایت شود.

۵- دستورهایش باید مبتنی بر عناصر ساختاری اساسی ساختار داده‌ای طراحی شوند.

ویژگیهای زبان داده‌ای فرعی

۶- بهتر است نارویه‌ای (ناروشمند) باشند.

۷- بهتر است کامپایلری باشد و نه مفسری.

۸- بهتر است از نظر ساختاری کامل باشد.

۹- بهتر است از نظر برنامه‌سازی و محاسباتی کامل باشد.

۱۰- زبان باید از نظر تعداد دستورهای کنترل داده‌ها و عملکرد هردستور، غنی و قوی باشد.

۱۱- باید از نظر انواع داده‌ای و به ویژه انواع داده‌ای انتزاعی و انواع داده‌ای پیچیده غنی باشد.

جلسه ششم

سیستم مدیریت پایگاه داده‌ها

آنچه در این جلسه می خوانید:

۱- سیستم مدیریت پایگاه داده‌ها DBMS
۲- DBMS به کاربر امکان می‌دهد تا
۳- رده‌بندی سیستم‌های DBMS از نظرهای مختلف
۴- اجزای DBMS از نمای بیرونی
۵- نمای بیرونی (ساده‌شده) DBMS
۶- اجزای DBMS از نمای درونی
۷- واحدهای لایه هسته
۸- واحدهای لایه مدیریت محیط پایگاه داده‌ها
۹- ساختار یک سیستم پایگاهی

هدفهای کلی: سیستم مدیریت پایگاه داده‌ها

هدفهای رفتاری: دانشجو در پایان این جلسه می‌تواند:

- سیستم مدیریت پایگاه داده‌ها را توصیف کند.
- رده‌بندی DBMS را از دیدگاههای مختلف بیان کند.
- اجزاء DBMS از نمای درونی و بیرونی را برشمارد.
- واحدهای لایه هسته و مدیریت پایگاه داده را نام ببرد.

سیستم مدیریت پایگاه داده‌ها

DBMS

تعریف: یکی از انواع نرم‌افزارهای واسط بین محیط فیزیکی ذخیره و بازیابی اطلاعات و محیط منطقی برنامه‌سازی است.

DBMS به کاربر امکان می‌دهد تا:

۱- پایگاه داده‌های خود را تعریف کند.

۲- در پایگاه داده‌های خود عملیات انجام دهد.

۳- روی پایگاه داده‌های خود تا حدی کنترل داشته باشد.

رده‌بندی سیستم‌های DBMS

از نظر نوع ساختار داده‌ای

- سیستم رابطه‌ای
- سیستم سلسله‌مراتبی
- سیستم شبکه‌ای
- جز اینها

رده‌بندی سیستم‌های DBMS

از نظر محیط سخت‌افزاری

- وابسته به یک محیط خاص
- ناوابسته به یک محیط خاص

رده‌بندی سیستم‌های DBMS

از نظر رده کامپیوتر

- خاص محیط کامپیوترهای شخصی
- خاص محیط کامپیوترهای متوسط (Mini Computer)
- خاص محیط کامپیوترهای بزرگ (Main Computer)
- خاص محیط کامپیوترهای خیلی بزرگ (Super Computer)
- اجراشونده درچند رده کامپیوتر

رده‌بندی سیستم‌های DBMS

از نظر محیط سیستم عامل

- وابسته به یک سیستم عامل خاص
- اجراشونده در محیط چند سیستم عامل

از نظر نوع معماری سیستم پایگاه داده‌ها

- با توانش ایجاد پایگاه متمرکز
- با توانش ایجاد پایگاه نامتمرکز

رده‌بندی سیستم‌های DBMS

از نظر معماری مشتری- خدمتگزار

- با توانش ایجاد معماری چند مشتری – یک خدمتگزار
- با توانش ایجاد معماری چند مشتری – چند خدمتگزار

از نظر زبان

- سیستم دارای SQL
- سیستم فاقد SQL

رده‌بندی سیستم‌های DBMS

از نظر نوع زبان داده‌ای فرعی

- دارای I.DSL
- دارای E.DSL
- دارای E/I.DSL

از نظر ماهیت زبان داده‌ای فرعی

- با زبان رویه‌ای
- با زبان نارویه‌ای

رده‌بندی سیستم‌های DBMS

از نظر سیستم فایل

- خودکفا
- وابسته به سیستم فایل محیط سیستم عامل

از نظر نوع کاربرد

- تک منظوره
- همه منظوره

رده‌بندی سیستم‌های DBMS

از نظر قیمت

از حدود ده هزار دلار تا صد هزار دلار و گاه بیشتر

از نظر طرز برپایی

- با محدودیت برپایی یکپارچه
- دارای امکان برپایی گزینشی

رده‌بندی سیستم‌های DBMS

از نظر واسط کاربر

- با واسط زبانی
- با واسط غیر زبانی
- با هر دو واسط

رده‌بندی سیستم‌های DBMS

از نظر رفتار در قبال رویدادها

- سیستم فعال
- سیستم غیر فعال

از نظر متدولوژی زبان

- بدون متدولوژی شیئ گرایی
- دارای متدولوژی شیئ گرایی

رده‌بندی سیستم‌های DBMS

از نظر بهینه‌سازی پرسش

- دارای بهینه‌سازی متعارف
- دارای بهینه‌سازی مبتنی بر قاعده، معنایی و ...

از نظر نوع تراکنش

- پذیرنده تراکنشهای ساده
- پذیرنده تراکنشهای با مدل پیشرفته

رده‌بندی سیستم‌های DBMS

از نظر نوع پردازش

- با قابلیت پردازش بی‌درنگ
- فاقد این قابلیت

از نظر رسانه ذخیره‌سازی پایگاه داده‌ها

- با قابلیت ایجاد MMDB
- فاقد این قابلیت

رده‌بندی سیستم‌های DBMS

از نظر قابلیت تعامل بین سیستمها

- فاقد این قابلیت
- دارای قابلیت تعامل با سیستمهای همگن
- دارای قابلیت تعامل با سیستمهای ناهمگن

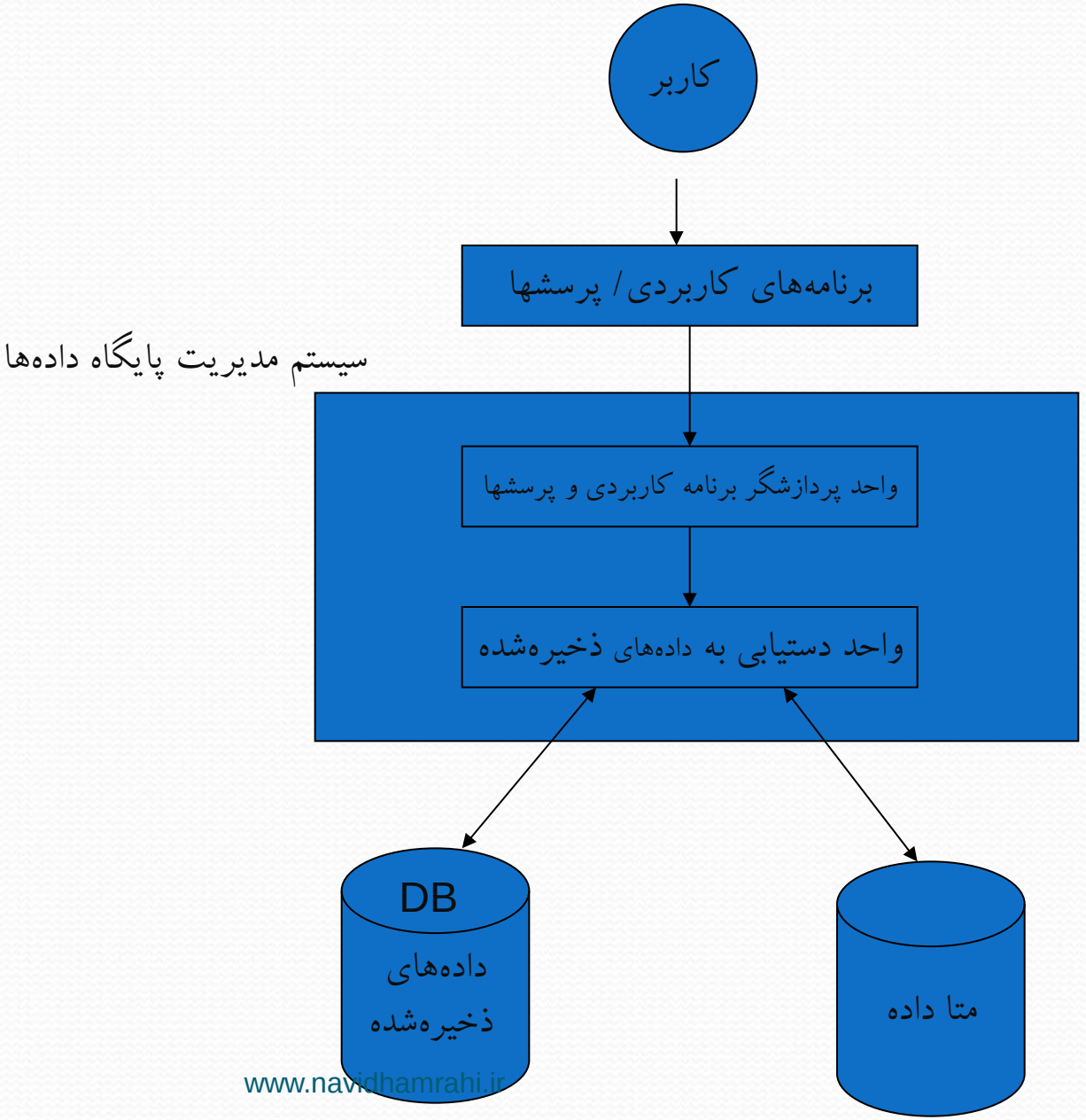
از نظر پردازش داده‌های زمانمند

- فاقد جنبه‌های یک سیستم زمانی
- سیستم مدیریت پایگاه داده‌های زمانی

اجزای DBMS از نمای بیرونی

- واحد پردازشگر پرسش‌ها و برنامه‌های کاربردی
- واحد ایجاد و مدیریت داده‌های ذخیره‌شده

نمای بیرونی (ساده‌شده) DBMS



اجزای DBMS از نمای درونی

- لایه هسته (سیستم کنترل یا موتور پایگاه داده‌ها)
- لایه مدیریت محیط پایگاه داده‌ها
- لایه تسهیلات نرم‌افزاری (ابزارها)

واحدهای لایه هسته

۱- واحد دریافت درخواست کاربر و وارسی‌های اولیه

۲- واحد تولید شماها

۳- پیش کامپایلرها برای DML

۴- کامپایلرها (پردازنده DML)

۵- پردازشگر پرسش و بهینه‌ساز پرسش

۶- واحد مدیریت سطح داخلی

واحدهای لایه هسته

۷- واحد مدیریت بافر

۸- واحد مدیریت فضای دیسک

۹- واحد ناظر زمان اجرا

۱۰- واحد مدیریت همروندی تراکنش‌ها

۱۱- واحد مدیریت انتقال داده‌ها

۱۲- واحد مدیریت کاتالوگ

واحدهای لایه مدیریت محیط پایگاه داده‌ها

۱- واحد کنترل جامعیت پایگاه داده‌ها

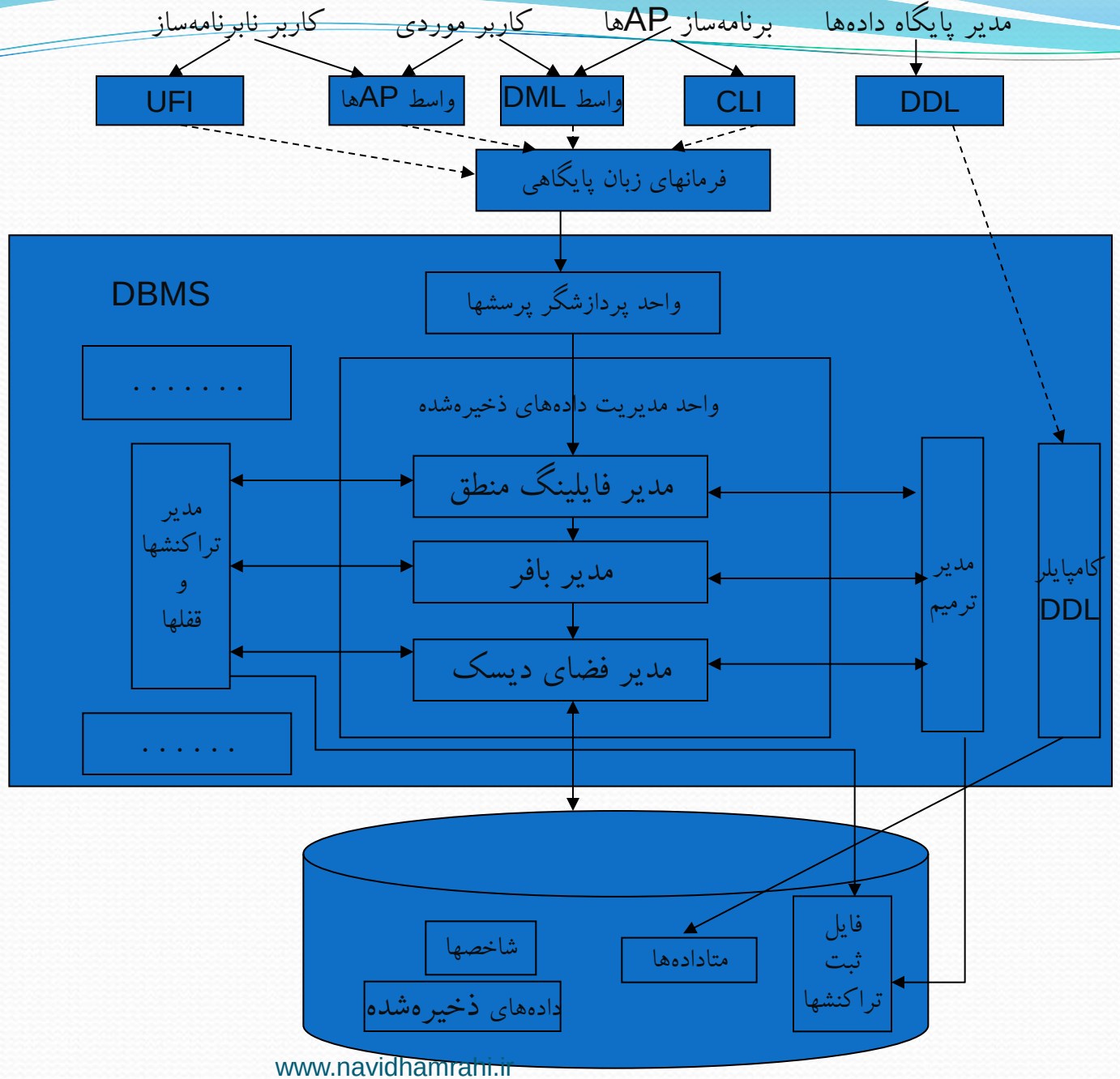
۲- واحد ترمیم پایگاه داده‌ها

۳- واحد ایمنی و حفاظت پایگاه داده‌ها

۴- واحد تولید نسخه‌های پشتیبان

۵- واحد تولید فایلهای ثبت تراکنشها

ساختار یک سیستم پایگاهی



جلسه هفتم

DBMS در یک سیستم کامپیوتری

آنچه در این جلسه می خوانید:

۱- جایگاه DBMS در یک سیستم کامپیوتری
۲- سه طرح کاربرد DBMS در برنامه‌های کاربردی
۳- اسلوبهای عملیاتی
۴- کاتالوگ سیستم و دیکشنری داده‌ها: متا داده‌ها
۵- اطلاعاتی که در دیکشنری داده‌ها نگهداری می‌شود
۶- دیکشنری داده‌ها و استفاده‌کنندگان آن
۷- پارامترهای شناخت DBMS
۸- مدیر پایگاه داده‌ها

آنچه در این جلسه می خوانید:

۹- برخی مسئولیتهای در تیم مدیریت پایگاه داده‌ها
۱۰- پنج سرمایه در مدیریت نوین سازمانها
۱۱- مزایا و معایب تکنولوژی پایگاه داده‌ها
۱۲- مزایا و معایب سیستم تک‌کاربری
۱۳- مزایا و معایب سیستم چندکاربری
۱۴- شرایط استفاده از تکنولوژی پایگاه داده‌ها
۱۵- کاربردهای جدید تکنولوژی پایگاه داده‌ها
۱۶- استقلال داده‌ای و انواع آن

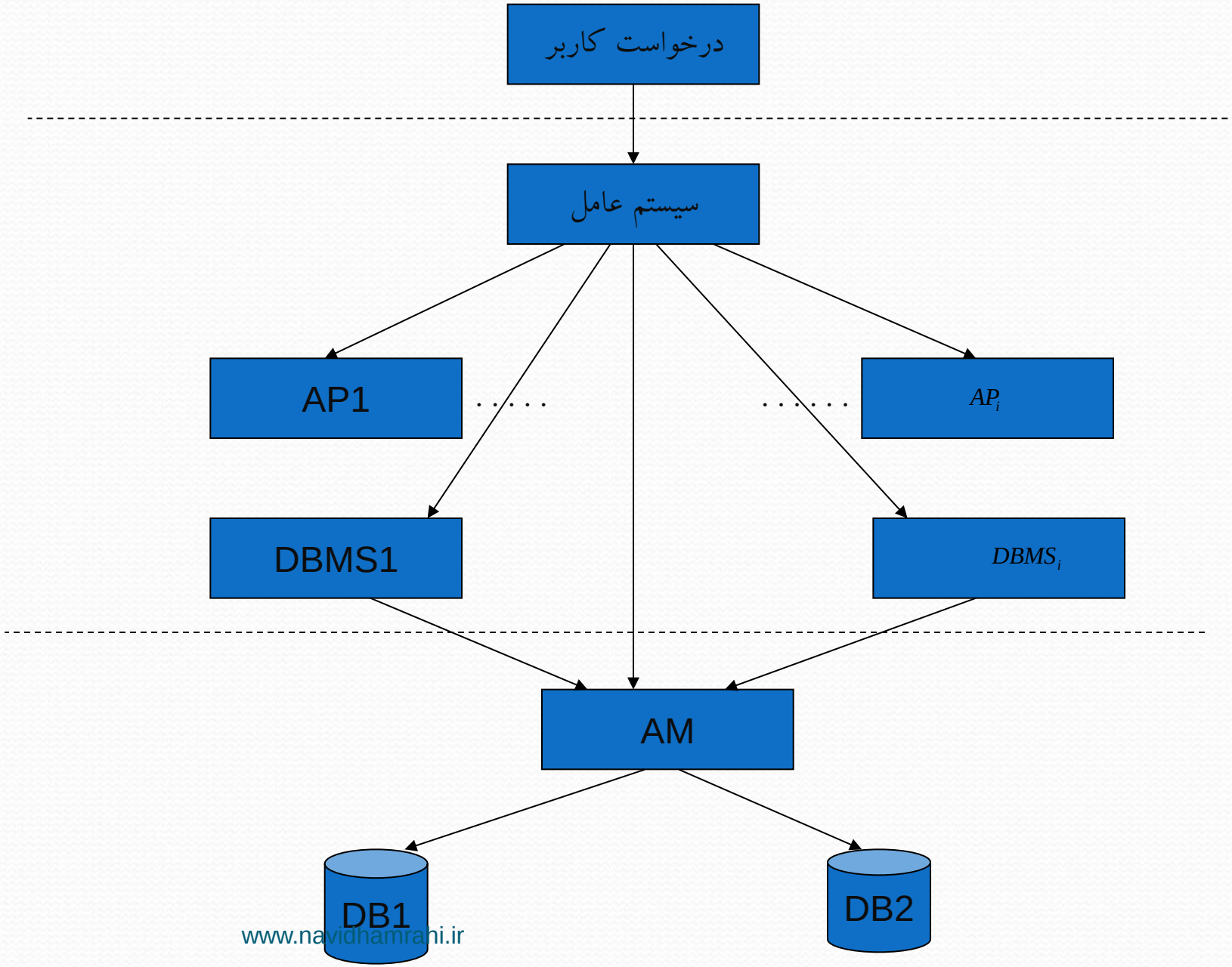
هدفهای کلی: DBMS در یک سیستم کامپیوتری و مدیر پایگاه داده‌ها

هدفهای رفتاری: دانشجو در پایان این جلسه می‌تواند:

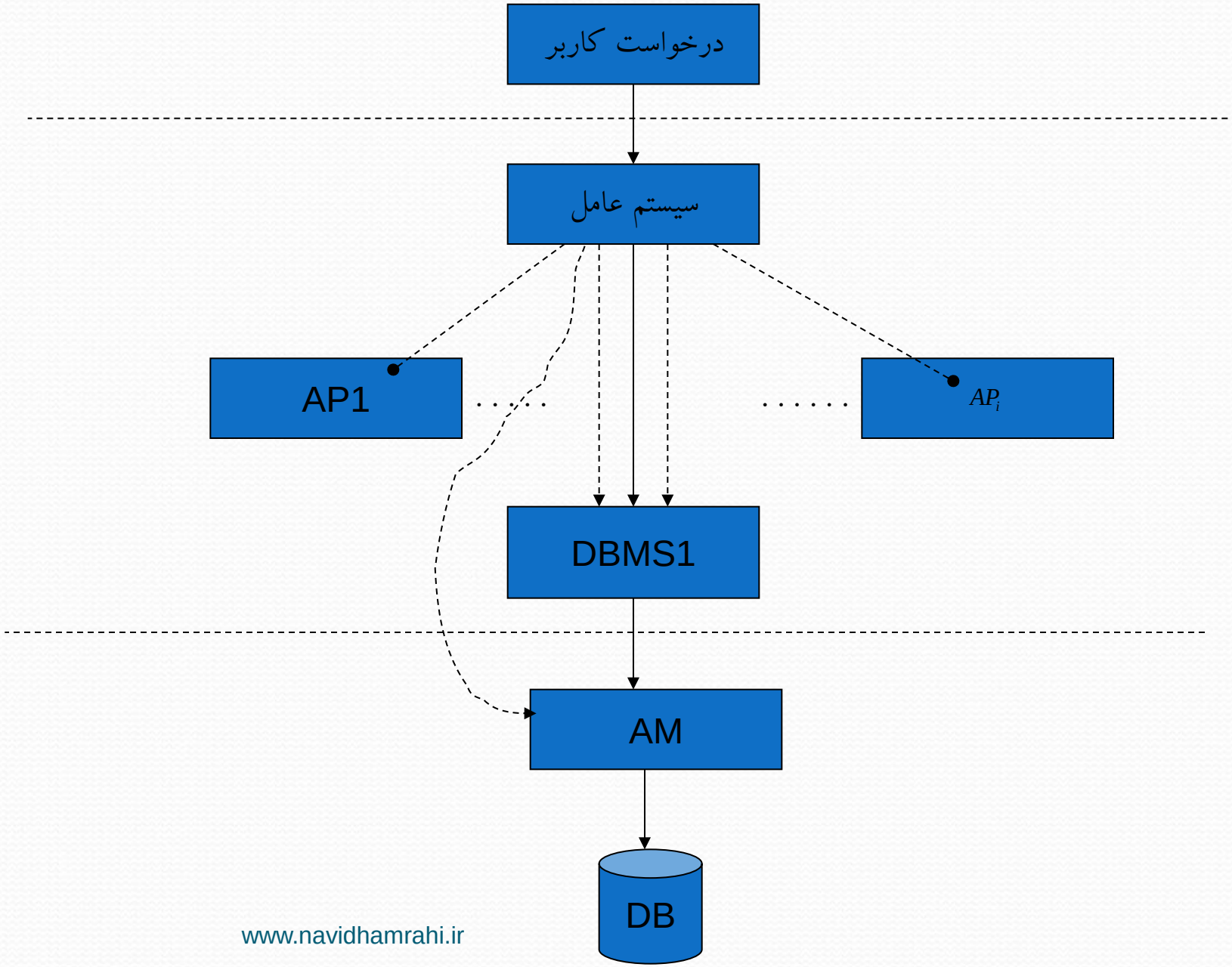
- ارتباط DBMS با سایر عناصر نرم‌افزاری را بیان کند
- اسلوبهای عملیاتی، و پارامترهای شناخت را نام ببرد.
- مسئولیتهای تیم مدیریت پایگاه داده‌ها را شرح دهد.
- مزایا و معایب سیستم تک‌کاربره و چندکاربره را توصیف کند.
- استقلال داده‌ای و انواع آن را توضیح دهد.

جایگاه DBMS در یک سیستم کامپیوتری

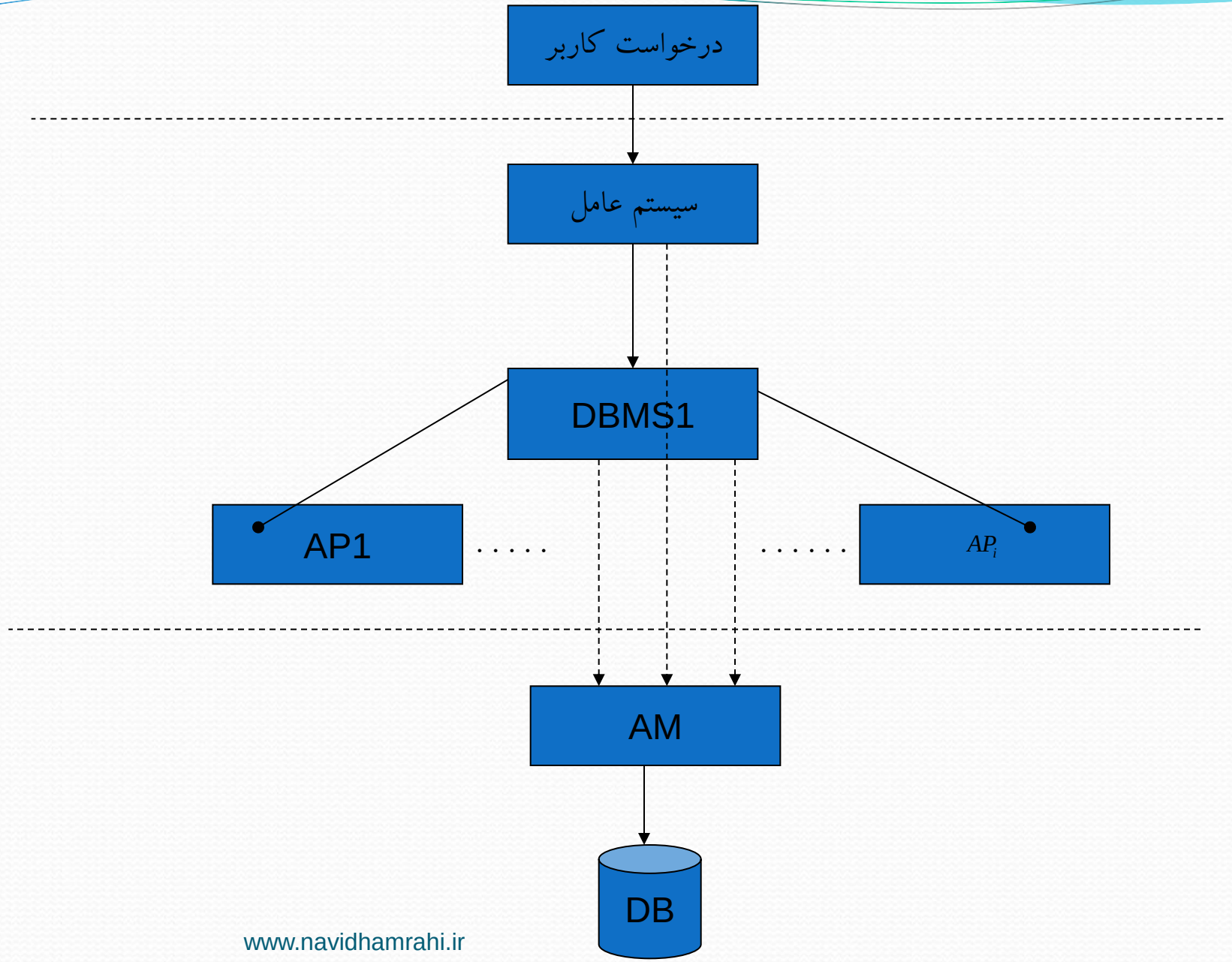
یک DBMS برای هر برنامه کاربردی



یک DBMS برای چند برنامه کاربردی تحت کنترل سیستم عامل



یک DBMS برای چند برنامه کاربردی با اجرای تحت کنترل DBMS



اسلوبهای عملیاتی

به طور کلی سه اسلوب عملیاتی وجود دارد:

- اسلوب یکجا
- اسلوب برخط
- اسلوب تعاملی

کاتالوگ سیستم و دیکشنری داده‌ها: متا داده‌ها

حاوی داده‌هایی است در مورد داده‌های ذخیره‌شده در پایگاه داده‌های کاربر و این داده‌ها به متاداده‌ها موسومند

دیکشنری داده‌ها معمولا جزئی از خود سیستم است و به دو صورت فعال و غیرفعال تولید می‌شود.

اطلاعاتی که در دیکشنری داده‌ها نگهداری می‌شود:

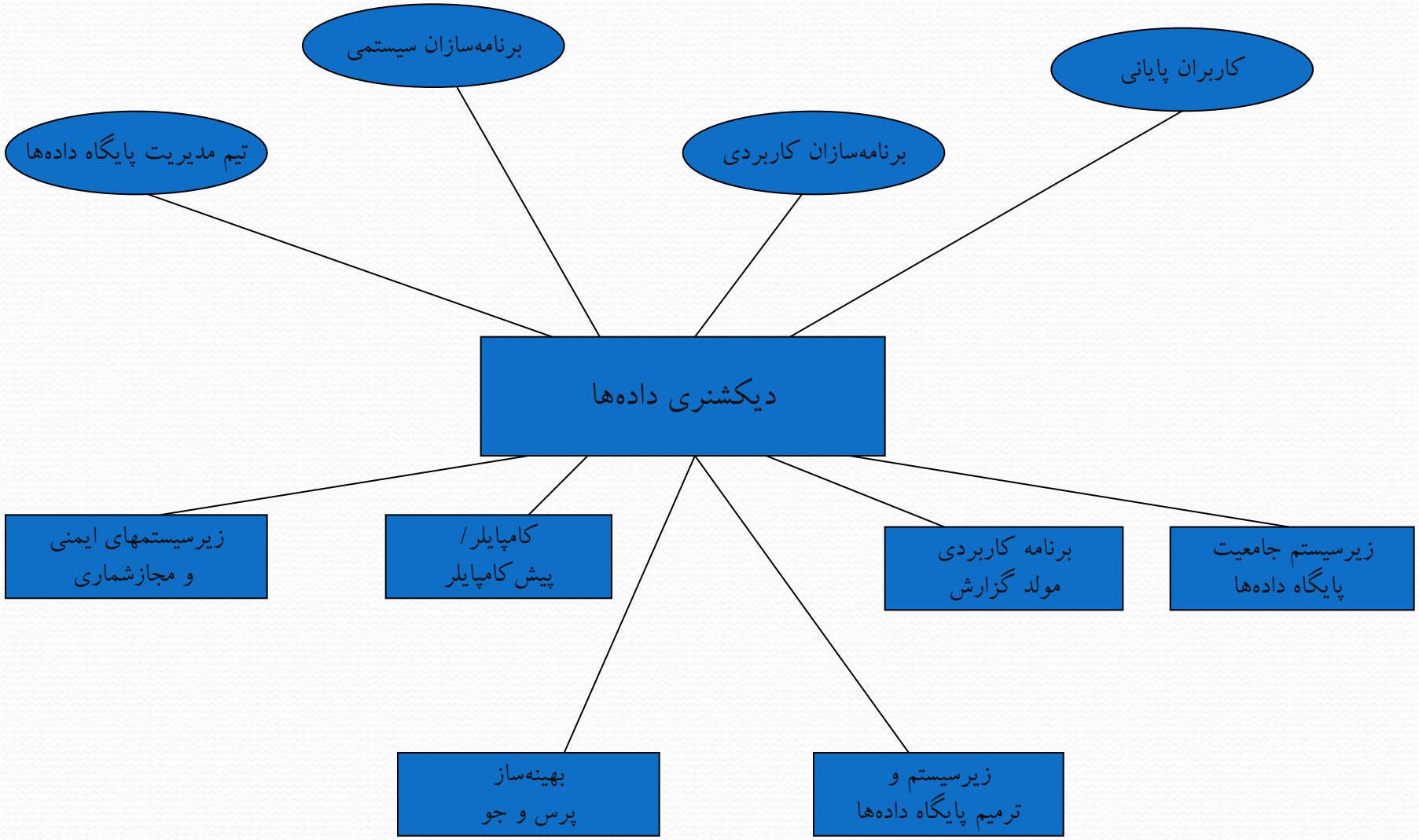
- شماهای خارجی
- شمای ادراکی
- شمای داخلی
- رویه‌های مربوط به تبدیلات بین سه سطح معماری
- شرح ساختار فیزیکی داده‌های ذخیره‌شده
- مشخصات و حقوق دستیابی کاربران به داده‌ها
- مشخصات برنامه‌های کاربردی تولید شده و ارتباط آنها با درخواستهای کاربران
- مشخصات پایانه‌های متصل به سیستم

ادامه

اطلاعاتی که در دیکشنری داده‌ها نگهداری می‌شود:

- ارتباط بین برنامه‌های کاربردی و داده‌های ذخیره‌شده
- قواعد جامعیت
- ضوابط کنترل ایمنی داده‌ها
- مشخصات پیکربندی سخت‌افزاری سیستم و رسانه‌های ذخیره‌سازی
- اطلاعات متنوع آماری پایگاه داده و کاربران
- توابع تعریف‌شده توسط کاربران

دیکشنری داده‌ها و استفاده کنندگان آن



پارامترهای شناخت DBMS

- پارامترهای مربوط به توانش‌ها و کارایی سیستم
- تسهیلات و جنبه‌های دیگر
- مشخصات کلی سیستم
- پارامترهای مربوط به معماری پایگاه داده‌ها
- پارامترهای مربوط به زبان داده‌ای فرعی

مدیر پایگاه داده‌ها

فردی است متخصص در پایگاه داده‌ها و با مسئولیت علمی، فنی و نیز اداری در محدوده وظایفی که عهده‌دار است

این مدیر همراه با یک تیم تخصصی کار می‌کند که به آن تیم مدیریت پایگاه داده‌ها می‌گویند.

برخی مسئولیتهای در تیم مدیریت پایگاه داده‌ها

- مدیر پایگاه داده‌ها
- مدیر داده‌ها
- مدیر امور پژوهش-توسعه
- مدیر سیستمهای کاربردی
- مسئول تیمهای برنامه‌سازی
- مسئول کنترل کارایی DBMS
- مسئول کنترل کارایی خود سیستم پایگاه داده‌ها
- مسئول نظارت بر عملیات روی پایگاه داده‌ها
- مسئول تماس با کاربران زیرمحیطهای سازمان
- مسئول تنظیم مستندات و وضع استانده‌ها

در مدیریت نوین سازمانها، هر سازمان دارای پنج سرمایه است:

۱- سخت‌افزار

۲- نرم‌افزار

۳- داده

۴- بودجه

۵- تخصص

مزایا و معایب تکنولوژی پایگاه داده‌ها

مزایای این تکنولوژی بستگی به نوع سیستم (DBMS) و معماری سیستم پایگاه داده‌ها و ماهیت کاربردها دارد.

سیستم تک کاربری

مزایا

۱- هر بخش از سازمان، داده‌های خود را نگهداری و پردازش می کند.

۲- با استفاده از کامپیوترهای شخصی، حجم داده‌های سیستم مرکزی کاهش می‌آید.

۳- پایگاه داده‌های ایجادشده روی کامپیوترهای شخصی معمولا کوچک و مدلسازی، طراحی و پیاده‌سازی آنها ساده است.

۴- کار با این سیستمها و برنامه‌سازی در محیط آنها ساده است.

۵- با پیشرفت کامپیوترهای شخصی، این سیستمها می‌توانند بسیاری از کارهای سیستمهای کامپیوتری بزرگ را انجام دهند.

سیستم تک کاربری

معایب

۱- وجود تعداد زیادی از این سیستمها در یک سازمان باعث بروز افزونگی، ناسازگاری داده‌ها و ناایمنی آنها می‌شود.

۲- محودیتهای سخت‌افزاری سبب محدودیت اندازه فایلها و نیز محدودیت سرعت پردازش باعث محدود شدن حجم پایگاه داده‌ها می‌شود.

۳- خود سیستم نمی‌تواند قوی و کارا باشد.

سیستم تک کاربری

معایب

۴- میزان ایمنی و حفاظت در آنها ضعیف است.

۵- امکانات تولید نسخه پشتیبان در آنها معمولا کم است.

۶- اشتراکی کردن آنها مشکلات تکنیکی جدی دارد.

۷- اعمال مجموعه واحدی از استانده‌ها در کل سازمان ناممکن است.

۸- معمولا کاربر این محیط مهارت کافی در مدلسازی و طراحی بهینه پایگاه داده‌ها ندارد.

۹- ایجاد یک سیستم جامع و یکپارچه، براساس این سیستمها، دشوار و پرهزینه است.

سیستم چند کاربری

مزایا

۱- اشتراک داده‌ها

۲- کاهش افزونگی

۳- تعدد شیوه‌های دستیابی به داده‌ها

۴- اجتناب از ناسازگاری داده‌ها

۵- تامین همروندی بهتر

۶- تسهیل پردازش تراکنشها

۷- تضمین جامعیت داده‌ها

ادامه

سیستم چند کاربری

مزایا

۸- حفظ محرمانگی داده‌ها

۹- امکان اعمال استانده‌ها

۱۰- تعدد زبانها

۱۱- کاهش حجم برنامه‌ها

۱۲- تنوع کاربران

۱۳- تسریع در دریافت پاسخ پرسشها

۱۴- استفاده بهتر از سخت‌افزار و ...

سیستم چند کاربری

معایب

۱- هزینه بالای نرم‌افزار و سخت‌افزار

۲- هزینه بیشتر برای برنامه‌سازی

۳- هزینه بالا برای انجام مهندسی دوباره به منظور تبدیل سیستم از مشی ناپایگاهی به مشی پایگاهی

۴- کند شدن اجرای بعضی از برنامه‌های کاربردی

۵- خطر آسیب‌پذیری داده‌ها

۶- تاثیرات گسترده‌تر خرابیها و دشواری بیشتر ترمیم آنها

۷- پیچیده بودن سیستم و نیاز به تخصص بیشتر

شرایط استفاده از تکنولوژی پایگاه داده‌ها

۱- نیاز به ایجاد یک سیستم یکپارچه اطلاعاتی

۲- حجم زیاد داده‌های سازمان و رشد پویای آن

۳- تغییرات مداوم در داده‌های ذخیره‌شده

۴- بالا بودن بسامد درخواستهای کاربران

۵- نیاز به اعمال کنترل متمرکز و دقیق روی کل داده‌ها

۶- وجود ارتباطات پیچیده بین داده‌ها

ادامه

شرایط استفاده از تکنولوژی پایگاه داده‌ها

۷- زیاد بودن میزان داده‌های مشترک بین برنامه‌های کاربردی

۸- مدنظر بودن صحت، دقت و سازگاری داده‌ها

۹- زیاد بودن گزارشها

۱۰- نیاز به انجام پردازشهای تحلیلی برخط

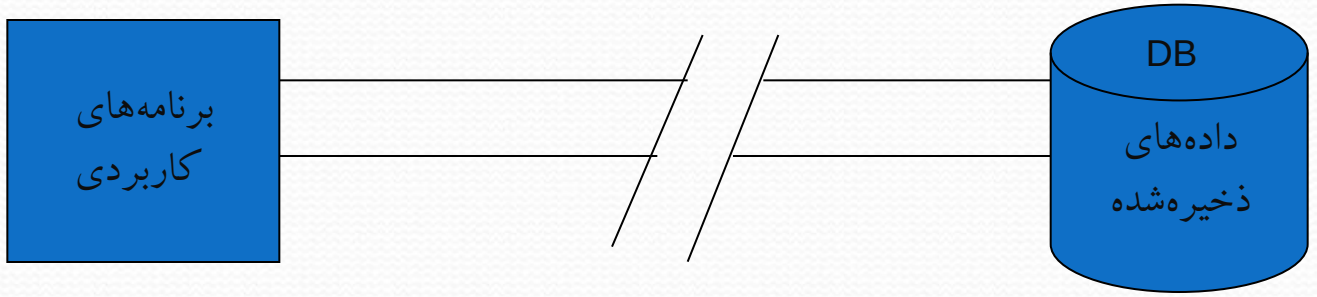
۱۱- نیاز به سیستم داده‌کاوی و کشف دانش در سازمان

کاربردهای جدید تکنولوژی پایگاه داده‌ها

- سیستم پشتیبان تصمیم
- سیستم داده کاوی چندرسانه‌ای
- سیستم انبارش داده‌ها
- حیطه‌های فضایی و جغرافیایی
- پایگاه داده‌های شخصی و همراه
- سیستم پایگاه داده‌ها در شبکه جهانی اطلاع‌رسانی
- سیستم اطلاعات اجرایی
- سیستم اطلاعات طراحی
- سیستم پردازش تحلیلی برخط چندبعدی
- ...

استقلال داده‌ای

یعنی وابسته نبودن برنامه‌های کاربردی به داده‌های ذخیره‌شده



انواع استقلال داده‌ای

- استقلال داده‌ای فیزیکی
- استقلال داده‌ای منطقی

استقلال داده‌ای فیزیکی

عبارتست از مصونیت دیدهای کاربران و برنامه‌های کاربردی در قبال تغییرات در سطح داخلی-فیزیکی پایگاه داده‌ها

چرا استقلال داده‌ای فیزیکی در سیستمهای رابطه‌ای جدید کاملا تامین است؟

زیرا:

۱- کاربران سطح خارجی در محیطی کاملا انتزاعی عمل می کنند و برنامه‌های کاربردی در این سطح با فایلینگ پایگاه داده‌ها تماس ندارند.

۲- بین سطح خارجی و داخلی، یک سطح انتزاعی دیگر واسط است و مانع تاثیرپذیری برنامه‌های سطح خارجی از تغییرات درفایلینگ پایگاه داده‌ها می‌شود.

استقلال داده‌ای منطقی

عبارتست از مصونیت دیدهای کاربران و برنامه‌های کاربردی در قبال تغییرات در سطح ادراکی پایگاه داده‌ها

تغییر در سطح ادراکی یعنی تغییر در طراحی منطقی پایگاه داده‌ها و تغییر در شمای ادراکی

این تغییر معمولا منجر به تغییر مناسب در سطح فایلینگ پایگاه می‌شود.

تغییر در سطح ادراکی دو وجه دارد:

۱- رشد پایگاه در سطح ادراکی

۲- سازمان‌دهی مجدد پایگاه در سطح ادراکی

جلسه هشتم

معماری سیستم پایگاه داده‌ها

آنچه در این جلسه می خوانید:

۱- معماری سیستم پایگاه داده‌ها
۲- انواع معماری
۳- معماری متمرکز
۴- معماری مشتری - خدمتگزار
۵- طرحهای معماری مشتری - خدمتگزار
۶- مزایای معماری مشتری – خدمتگزار در مقایسه با معماری متمرکز
۷- معماری توزیع‌شده

آنچه در این جلسه می خوانید:

۸- ویژگیهای معماری توزیع‌شده
۹- مزایا و معایب معماری توزیع‌شده
۱۰- معماری با پردازش موازی
۱۱- طرح های معماری با پردازش موازی
۱۲- معماری سیستم چندپایگاهی
۱۳- معماری سیستم پایگاههای همراه

هدفهای کلی: معماری سیستم پایگاه داده‌ها

هدفهای رفتاری: دانشجو در پایان این جلسه می‌تواند:

- گونه‌های مختلف معماری پایگاه داده‌ها شامل معماری متمرکز و نامتمرکز و انواع معماری نامتمرکز و ویژگیهای هریک را شرح دهد.

معماری سیستم پایگاه داده‌ها

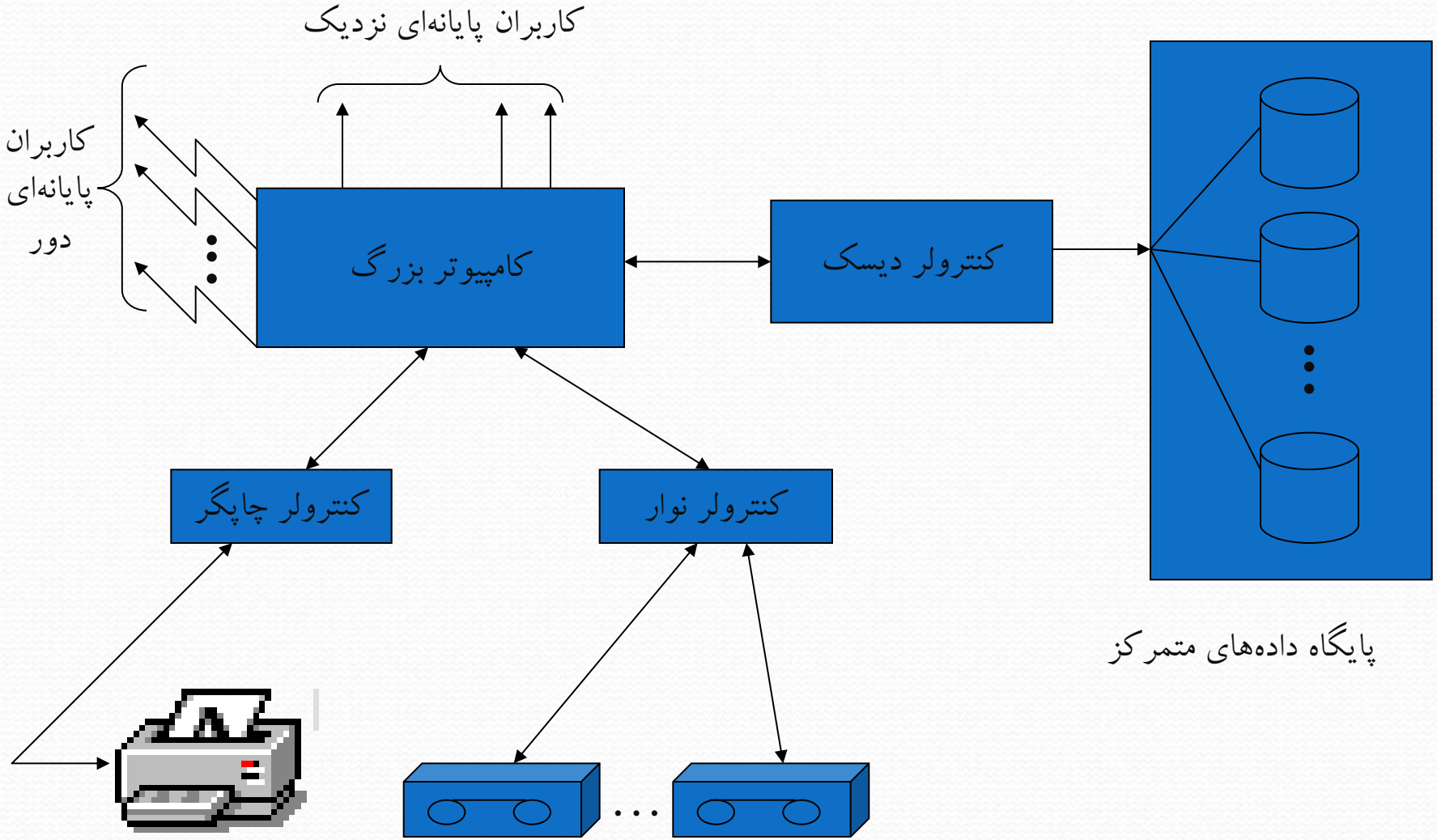
منظور، چندی و چونی اجزاء تشکیل‌دهنده سیستم و نیز پیکربندی یا طرز ترکیب اجزاء سیستم و چگونگی تعامل اجزاء با یکدیگر است. در این معماری حداقل یک پایگاه داده‌ها، یک سیستم مدیریت پایگاه داده‌ها، یک سیستم عامل، یک کامپیوتر با دستگاههای جانبی و تعدادی برنامه کاربردی و کاربر وجود دارند.

- انواع معماری
  - معماری متمرکز
  - معماری نامتمرکز
    - معماری مشتری-خدمتگزار
    - معماری توزیع‌شده
    - معماری چندپایگاهی
    - معماری با پردازش موازی
    - معماری موبایل

معماری متمرکز

در این معماری یک پایگاه داده روی یک سیستم کامپیوتری و بدون ارتباط با سیستم دیگر ایجاد می شود .

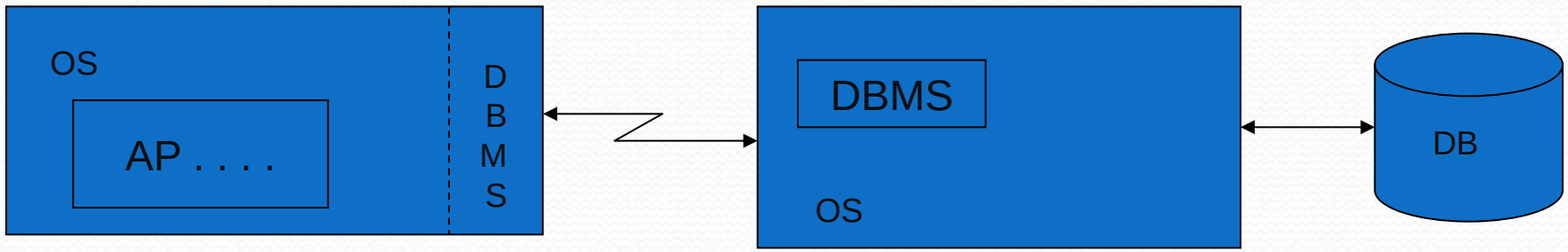
نمایش معماری متمرکز



معماری مشتری - خدمتگزار

هر معماری که در آن قسمتی از پردازش را یک برنامه، سیستم یا ماشین انجام دهد و انجام قسمت دیگری از پردازش را از برنامه ، سیستم یا ماشین دیگر بخواهد معماری مشتری خدمتگزار نامیده می شود

معماری مشتری - خدمتگزار



طرحهای معماری مشتری - خدمتگزار

از نظر تعداد مشتری و خدمتگزار

۱- چند مشتری – یک خدمتگزار

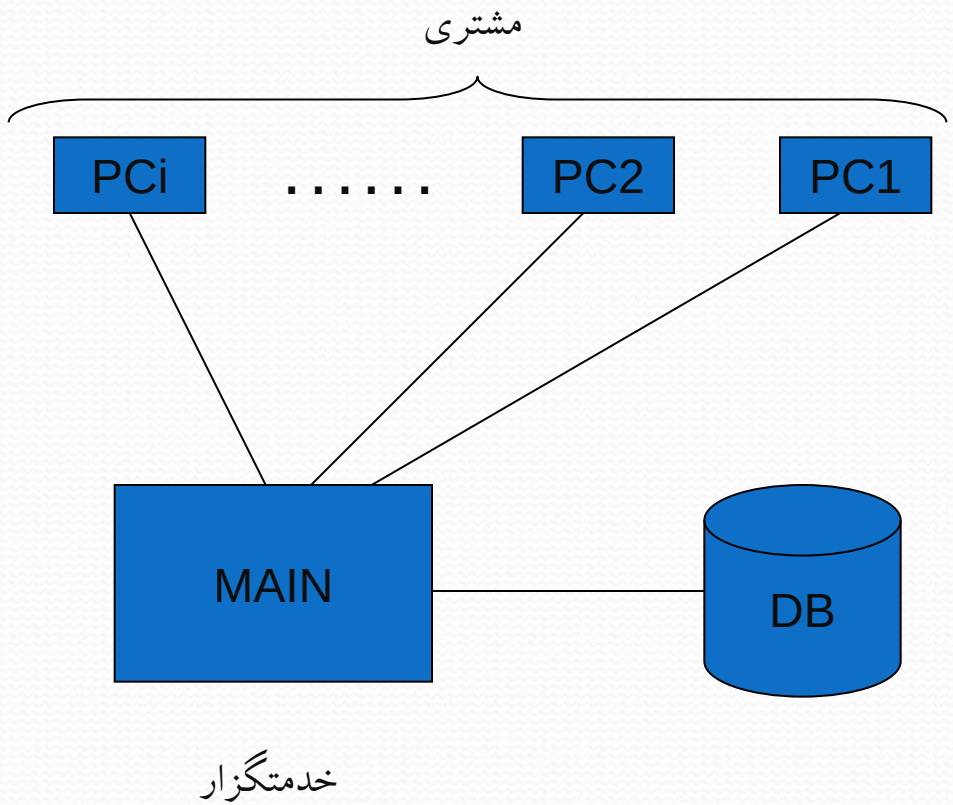
۲- یک مشتری – چند خدمتگزار

۳- چند مشتری – چند خدمتگزار

طرحهای معماری مشتری – خدمتگزار

از نظر پیکربندی سخت‌افزاری

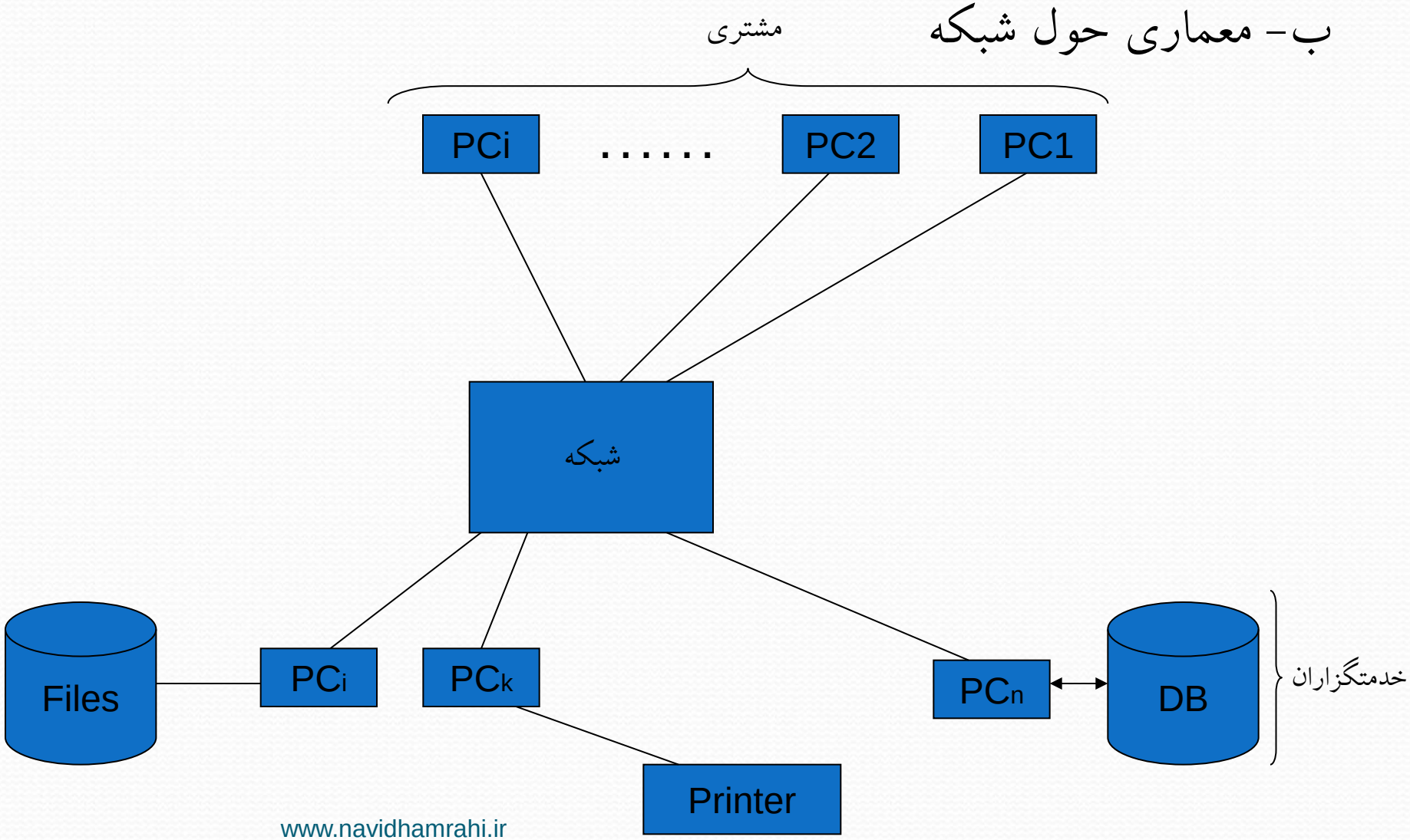
الف- معماری حول کامپیوتر بزرگ



طرحهای معماری مشتری - خدمتگزار

از نظر پیکربندی سخت‌افزاری

ب- معماری حول شبکه

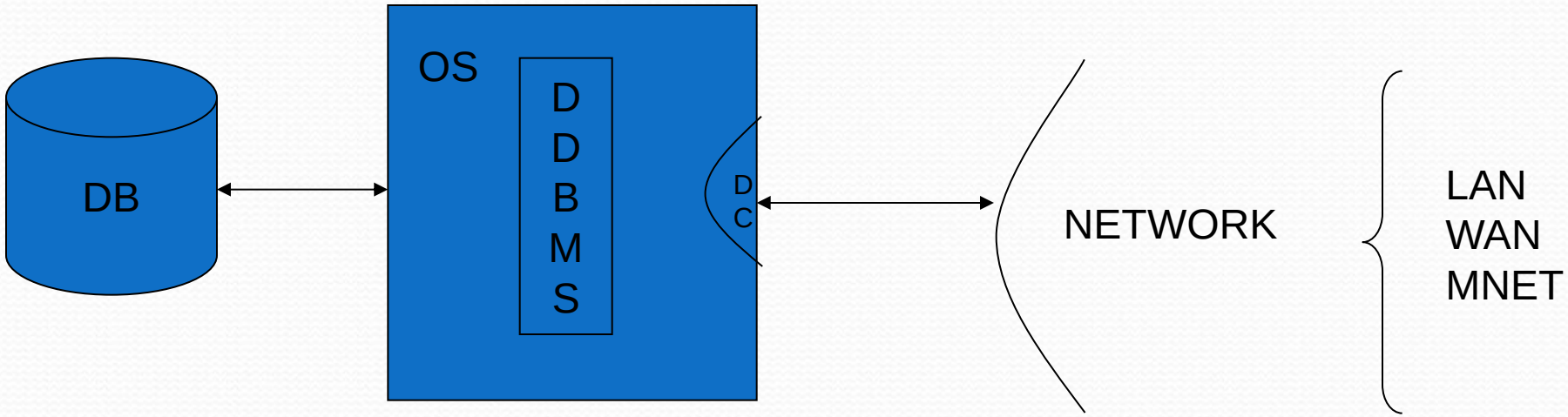


مزایای معماری مشتری-سرور نسبت به رایانش متمرکز

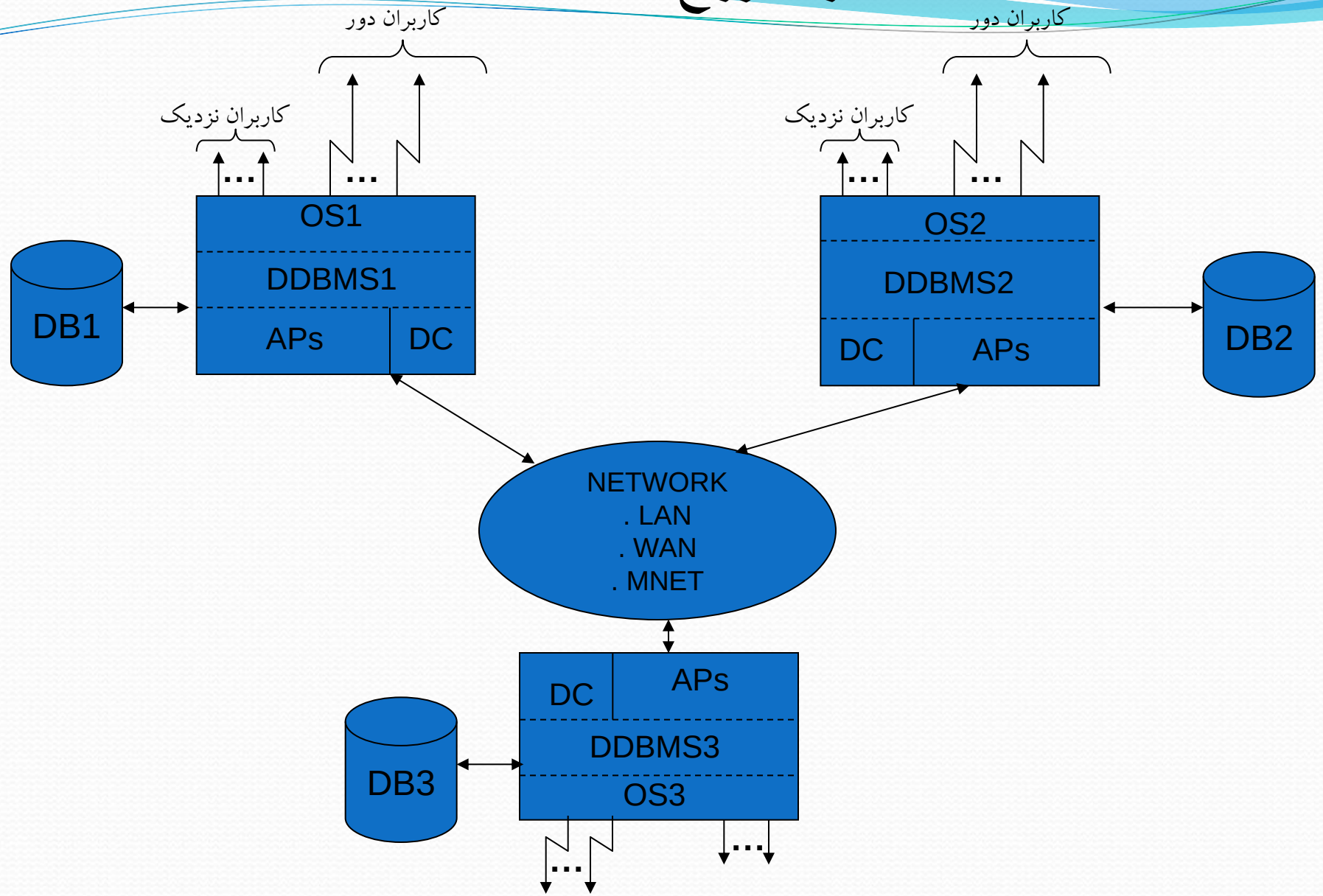
- تقسیم پردازش
- کاهش ترافیک شبکه
- استقلال ایستگاههای کار
- اشتراک داده‌ها

مجموعه ای از داده های ذخیره شده که منطقا به یک سیستم تعلق دارند ولی در مانه های مختلف یک یا بیش از یک شبکه توزیع شده اند.

نمای یک مانه در معماری توزیع‌شده



معماری توزیع‌شده با سه مانه



DDB = {DB1+DB2+DB3}

ویژگیهای معماری توزیع‌شده

- مجموعه‌ای است از داده‌های منطقا مرتبط و اشتراکی
- داده‌ها به بخشهایی تقسیم و در مانه‌ها توزیع شده‌اند.
- بعضی بخشها ممکن است به طور تکراری در مانه‌ها ذخیره شده باشند.
- مانه‌ها از طریق شبکه بهم مرتبطاند.
- داده‌های هر مانه تحت کنترل یک DBMS است.
- DMBS هر مانه، می‌تواند برنامه‌های کاربردی محلی را به طور خودکار اجرا کند.
- هر DBMS حداقل دراجرای یک برنامه کاربردی سرتاسری مشارکت دارد.

مزایای معماری توزیع شده

- سازگاری و هماهنگی با ماهیت سازمانهای نوین.
- کارایی بیشتر در پردازش داده‌ها.
- دستیابی بهتر به داده‌ها.
- اشتراک داده‌ها.
- افزایش پردازش موازی.
- کاهش هزینه ارتباطات.
- تسهیل گسترش سیستم.
- استفاده از پایگاه داده‌های از قبل موجود.

معایب معماری توزیع‌شده

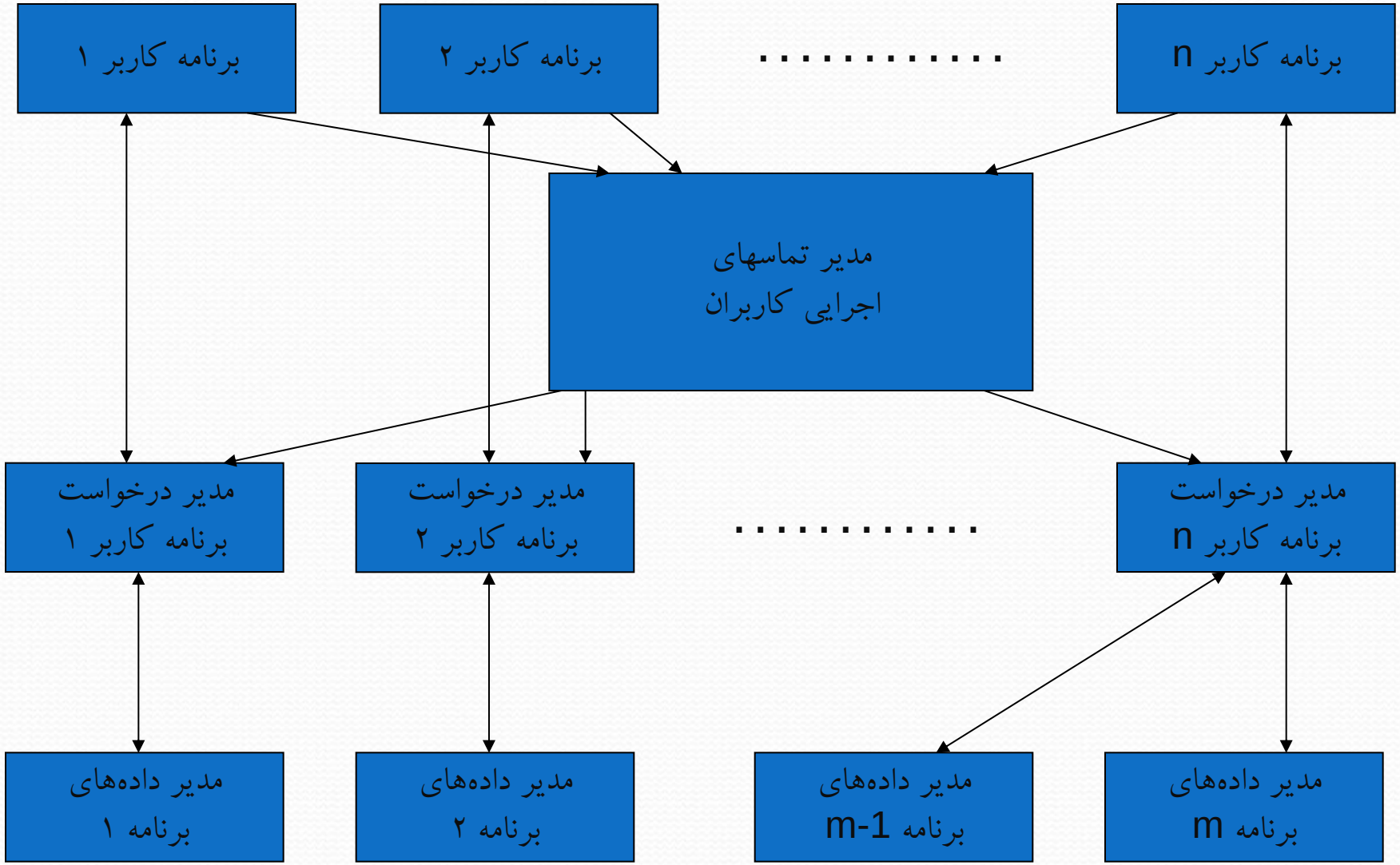
- پیچیدگی طراحی سیستم.
- پیچیدگی پیاده‌سازی.
- کاهش کارایی در برخی موارد.
- هزینه بیشتر.
- مصرف حافظه بیشتر.

معماری با پردازش موازی

در این گونه سیستمها معمولا تعداد زیادی تراکنش در ثانیه و بطور موازی اجرا می‌شود .

طرح کلی معماری با پردازش موازی

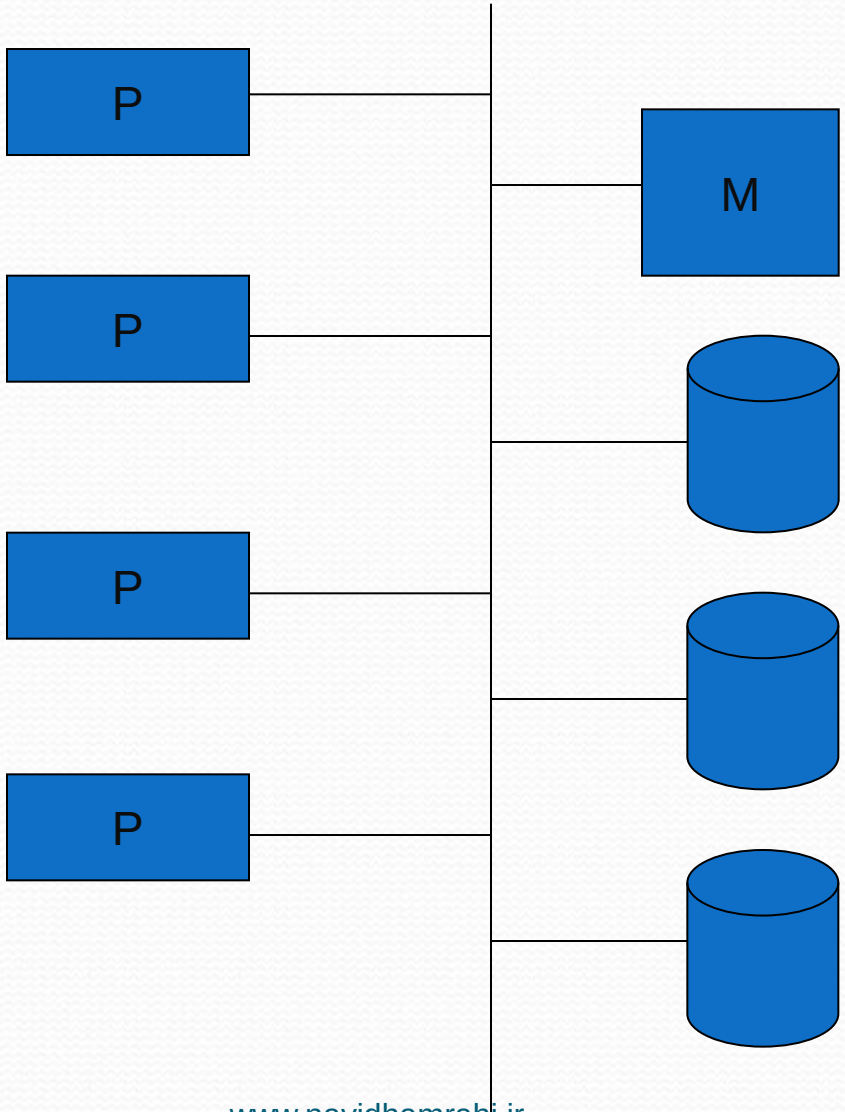
خدمتگزاران برنامه‌های کاربردی



خدمتگزار پایگاه داده‌ها

طرح های معماری با پردازش موازی

الف- معماری با حافظه مشترک



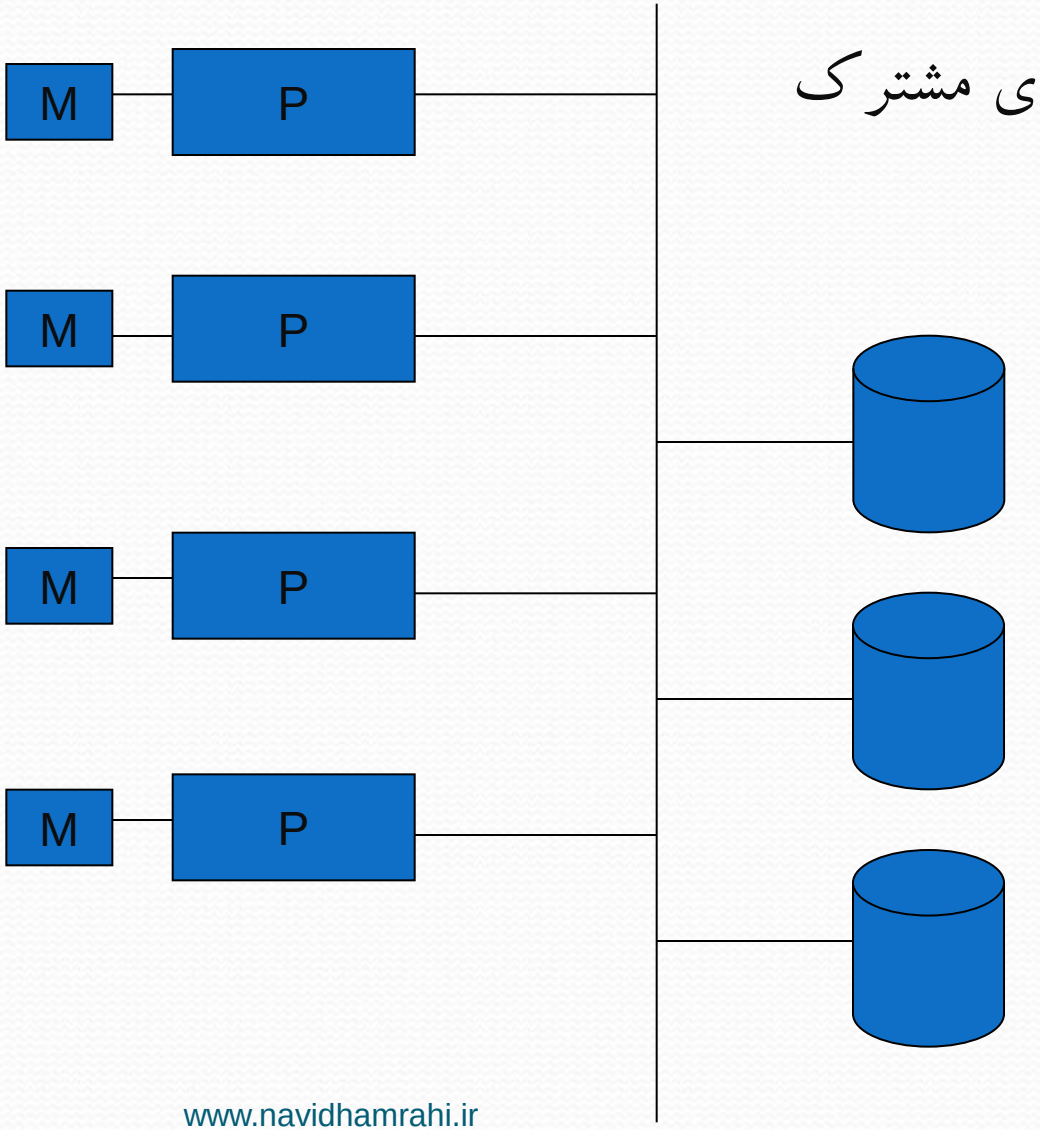
معماری با حافظه مشترک

مزیت: ارتباط بین پردازنده‌ها به طور کارا انجام می‌شود.

عیب: نمی‌توان بیش از ۳۲ یا ۶۴ پردازنده داشت.
زیرا احتمال بروز تنگنا در باسهای حافظه‌ای یا شبکه ارتباطی افزایش می‌یابد.

طرح های معماری با پردازش موازی

ب- معماری با دیسکهای مشترک



معماری با دیسکهای مشترک

مزیت:

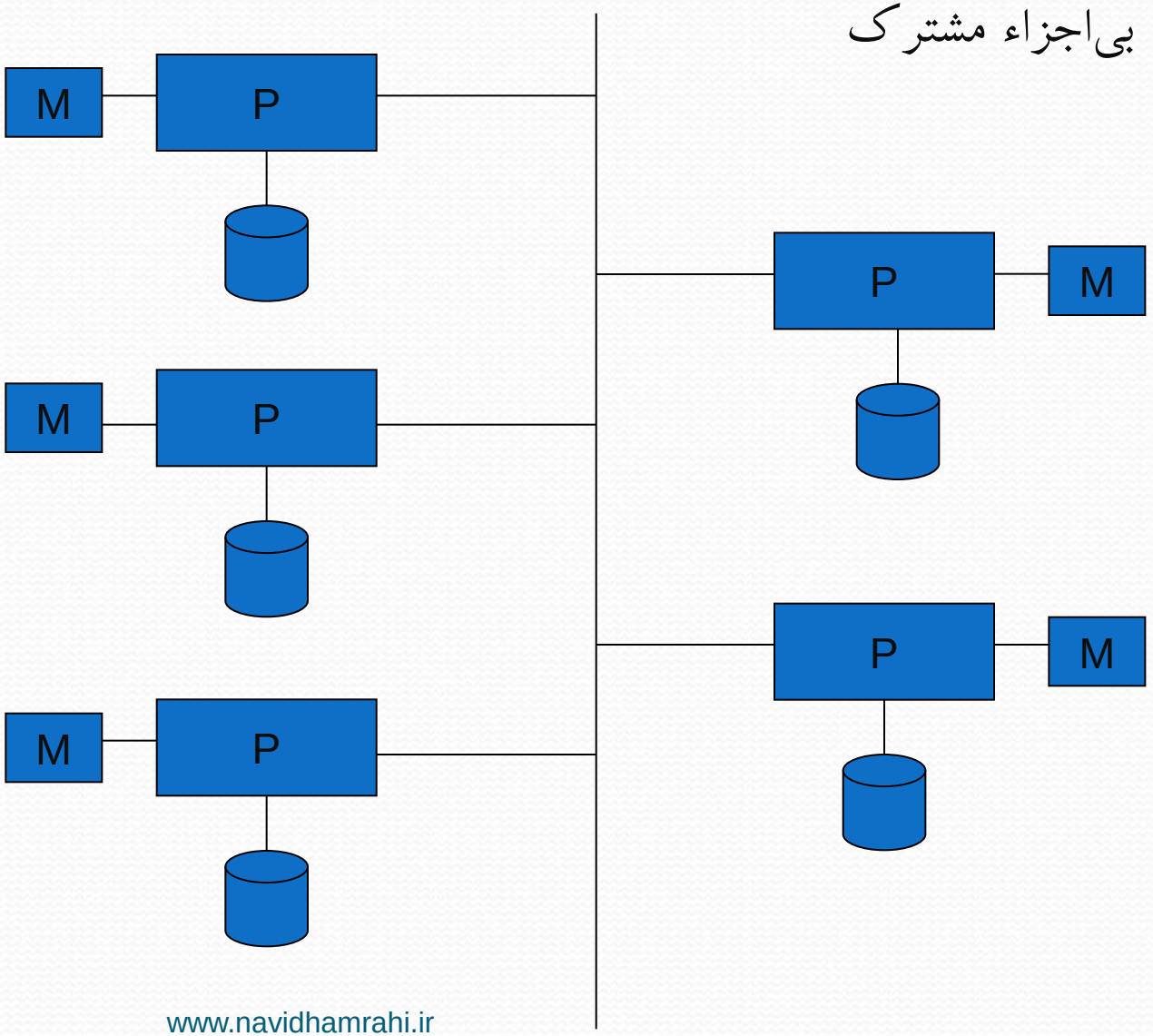
۱- عدم بروز تنگنا در باسهای حافظه‌ای

۲- تسهیل تحمل خرابی

عیب: دشواری در گسترش سیستم.

طرح های معماری با پردازش موازی

ج- معماری بی‌اجزاء مشترک



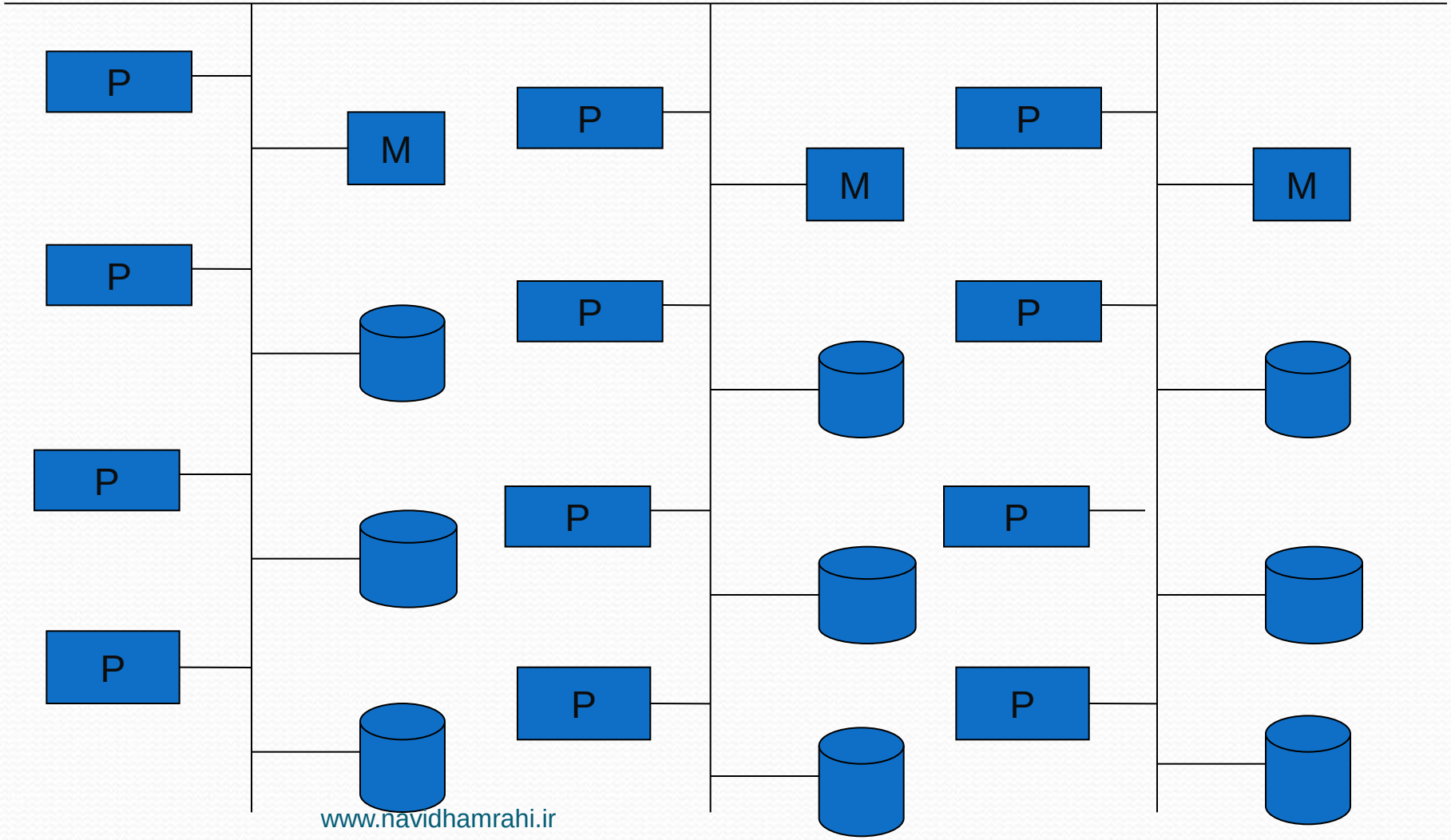
معماری بی‌اجزاء مشترک

مزیت: تسهیل گسترش

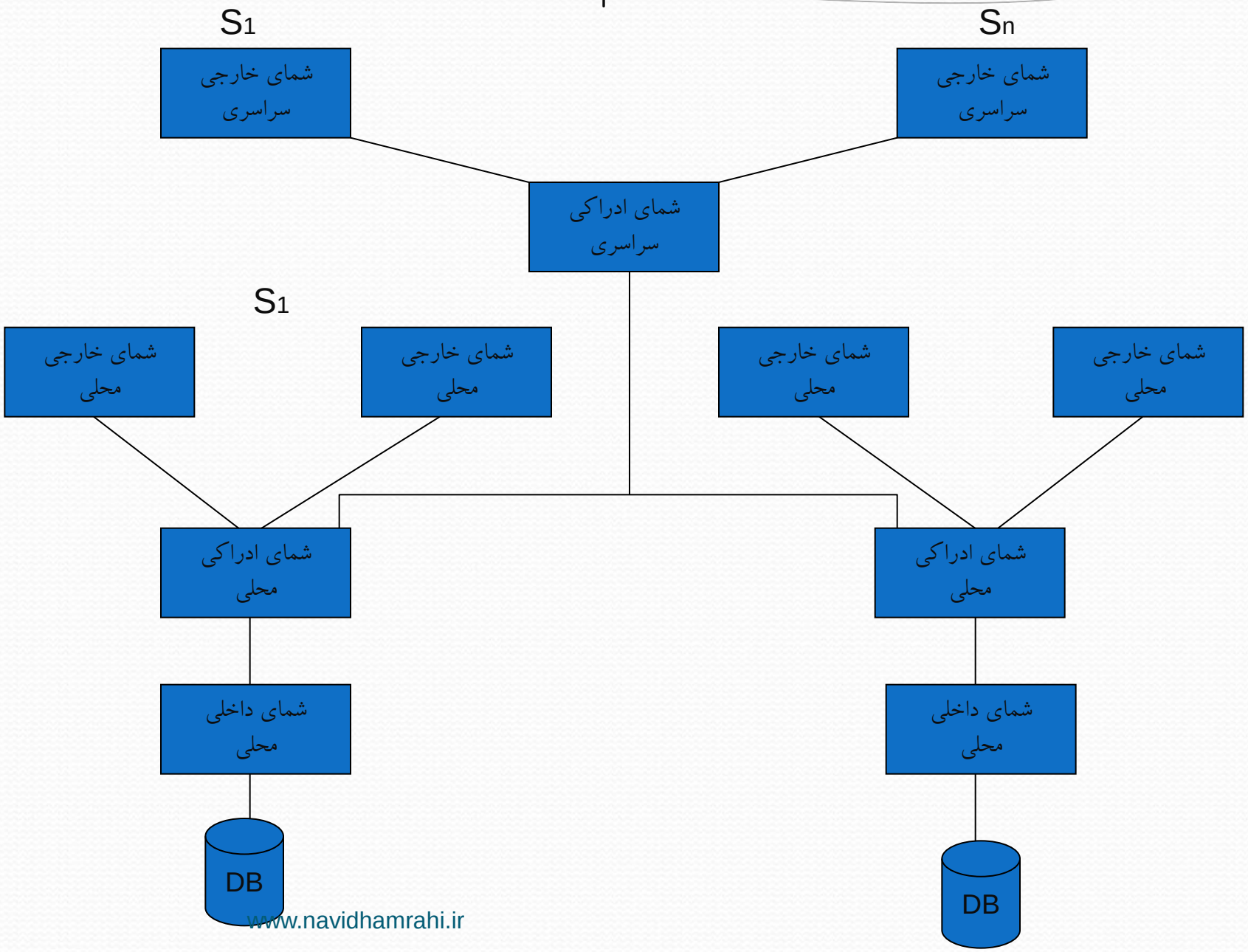
عیب: هزینه ارتباط و دستیابی‌های غیرمحلی زیاد است.

طرح های معماری با پردازش موازی

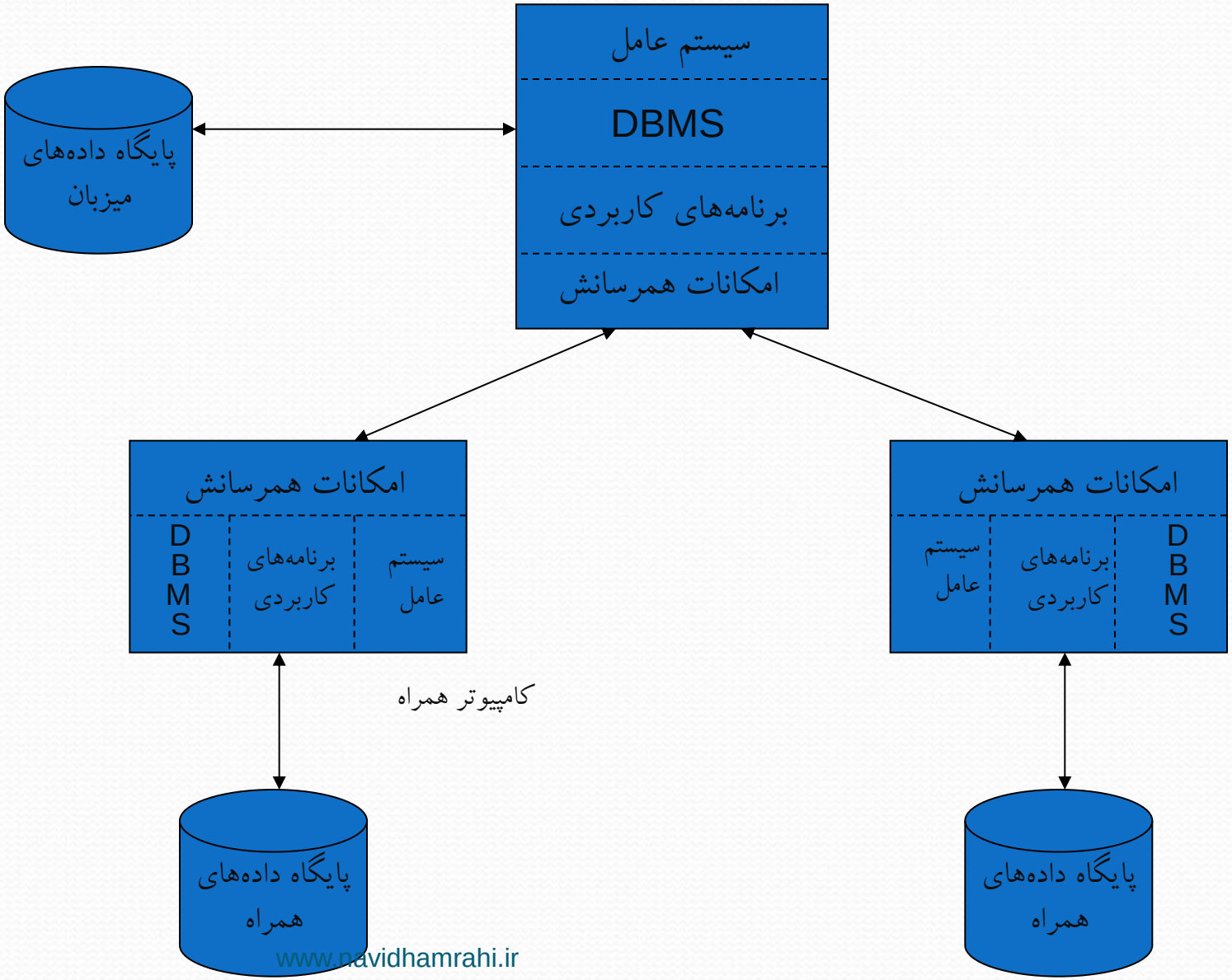
د- معماری سلسله‌مراتبی



معماری سیستم چندپایگاهی



معماری سیستم پایگاههای همراه



جلسه نهم

مدل رابطه‌ای

آنچه در این جلسه می خوانید:

۱- مفاهیم اساسی مدل رابطه‌ای
۲- بخشهای اساسی مدل داده‌ای
۳- تعریف رابطه
۴- تناظر بین مفاهیم رابطه ای و مفاهیم جدولی
۵- ویژگیهای رابطه
۶- انواع رابطه
۷- میدان (دامنه)

آنچه در این جلسه می خوانید:

۸- مزایای میدان
۹- رابطه نرمال و غیر نرمال
۱۰- دلیل لزوم نرمال بودن رابطه
۱۱- معایب رابطه نرمال
۱۲- مزایا و معایب رابطه غیرنرمال
۱۳- انواع کلید در مدل رابطه ای

هدفهای کلی: آشنایی با مدل رابطه‌ای

هدفهای رفتاری: دانشجو در پایان این جلسه می‌تواند:

- مدل رابطه‌ای و بخشهای اساسی آن را توضیح دهد.
- رابطه را تعریف و ویژگیهای آن را بیان نماید.
- انواع رابطه را نام برده و توضیح دهد.
- روابط نرمال و غیرنرمال را شرح دهد.
- کلید و انواع آن را توضیح دهد.

مفاهیم اساسی مدل رابطه‌ای

مدل رابطه‌ای در سال ۱۹۷۰ توسط کاد ابداع شد.

مدل داده ای امکانی است برای طراحی منطقی پایگاه داده‌ها، تعریف و کنترل آن و نیز انجام عملیات درآن و امکان می دهد تا این هر سه عمل اساسی در محیط انتزاعی انجام شود. بنابراین می‌توان گفت که مدل داده‌ای تامین کننده محیط انتزاعی پایگاه داده هاست .

بخشهای اساسی مدل داده‌ای

۱- بخش ساختاری

۲- بخش عملیاتی (پردازشی)

۳- بخش جامعیتی

بخش ساختاری، نشان‌دهنده عناصر ساختاری مدل است که همان ساختار داده‌ای اصلی و مفاهیم مرتبط با آن است.

بخش عملیاتی، مجموعه امکاناتی است که به وسیله آنها عملیات مورد نظر کاربر انجام می‌شود.

بخش جامعیتی، از مجموعه‌ای از قواعد و محدودیتهای جامعیتی تشکیل شده است که به وسیله آنها سیستم مدیریت پایگاه داده می‌تواند صحت، دقت و سازگاری داده‌ها را کنترل و تضمین کند.

تعریف رابطه

با فرض وجود n میدان D_1 تا D_n، نه لزوما متمایز، رابطه R از دو قسمت تشکیل شده است:

۱- سرآیند: مجموعه‌ای نامدار از n صفت به صورت $A_i:D_i$ که در آن هر A_i نام یک صفت است و هر D_i نام میدان صفت

۲- پیکر (بدنه): مجموعه‌ای است از m تاپل t به نحوی که t خود مجموعه‌ای است از n عنصر هریک به صورت $A_i:v_i$ که در آن v_i مقداری است از نوع میدان.

مقدار n را درجه (همان تعداد صفات) و مقدار m را کاردینالیتی رابطه می‌گویند

تناظر بین مفاهیم رابطه ای و مفاهیم جدولی

جدول زیر تناظر بین مفاهیم رابطه ای و مفاهیم جدولی را نشان می دهد :

مفهوم تئوریک	مفهوم جدولی
رابطه	جدول
تاپل	سطر
صفت	ستون
میدان	مجموعه مقادیر ستون
درجه	تعداد ستونها
کاردینالیتی	تعدادسطرها

ویژگیهای رابطه

ویژگیهای رابطه عبارتند از:

۱- رابطه تاپل تکراری ندارد

۲- تاپلها نظم ندارند

۳- صفات رابطه نظم مکانی ندارند

۴- تمام صفات تک‌مقداری هستند.

انواع رابطه

۱- مبنا: استقلال وجودی دارد و از رابطه‌های دیگر مشتق نیست و داده‌های ذخیره‌شده متناظر دارد.

۲- نامدار: با یک نام به سیستم معرفی می‌شود.

۳- دید: نوعی رابطه نامدار که مشتق از رابطه‌های دیگر است و ماهیتا مجازی است.

۴- لحظه‌ای: نامدار و مشتق است ولی مجازی نیست.

۵- مشتق: به کمک یک عبارت رابطه‌ای بر حسب رابطه‌های مبنا تعریف می‌شود.

ادامه

انواع رابطه

۶- عبارتی: از مجموعه‌ای از رابطه‌های نامدار و به وسیله یک عبارت رابطه‌ای به دست می‌آید.

۷- نتیجه پرسش: بی‌نام و مشتق است که حاصل اجرای یک پرسش مشخص است.

۸- بینابینی: بی‌نام و مشتق است که حاصل ارزیابی یک عبارت رابطه‌ای است که درون یک عبارت بزرگتر جای دارد

۹- مشتق: رابطه‌ای عبارتی است که به طور مستقیم و کارا ذخیره شده است.

میدان (دامنه)

از نظر ریاضی، مجموعه‌ای است از مقادیر که یک یا بیش از یک صفت از آن مقدار می‌گیرند. در مدل رابطه‌ای، این مجموعه:

- نامدار است.
- مقادیرش نوع مشخص دارند.
- مقادیرش فرمت مشخص دارند.

مزایای میدان

۱. سبب ساده‌ترشدن و کوتاه‌تر شدن شمای پایگاه داده‌ها می‌شود.
۲. تغییر در شمای پایگاه را تسهیل می کند.
۳. امکانی است برای کنترل مقداری عملیات در پایگاه داده‌ها.
۴. امکانی است برای کنترل معنایی درخواستها.
۵. پاسخگویی به بعضی پرسشها را آسان می کند.
۶. با استفاده از مفهوم میدان می‌توان امکانات و جنبه‌های موجود در سیستم مدیریت پایگاه داده‌های شیئ‌گرا را به سیستم مدیریت پایگاه داده‌های رابطه‌ای افزود.

رابطه نرمال و غیر نرمال

- رابطه نرمال :

رابطه ای است که مقادیر هیچیک از میدانهایش ، یک مقدار رابطه ای با کاردینالیتی بزرگتر از یک نباشد

- رابطه غیر نرمال :

رابطه ای است که در آن مقادیر حداقل یک صفت ، خود مقادیر رابطه‌ای هستند

(رابطه ای که حداقل یک صفت چند مقداری داشته باشد)

دلیل لزوم نرمال بودن رابطه

۱- سادگی در نمایش ظاهری رابطه (جدول با سطرهای ساده‌تر)

۲- سادگی دستورات DSL (بویژه DML و DDL)

۳- سادگی در اجرای عملیات در پایگاه داده‌ها

معایب رابطه نرمال

۱- بروز پدیده افزونگی که می‌تواند فیزیکی هم باشد.

۲- طولانی‌تر شدن کلید رابطه.

۳- عدم امکان نمایش داده‌های پیچیده.

۴- دشواری در نمایش طبیعی مفهوم سلسله مراتب.

۵- دشواری در نمایش مفهوم وراثت.

مزایا و معایب رابطه غیرنرمال

مزایا:

۱- کاهش میزان افزونگی

۲- کوتاه شدن کلید

۳- امکان نمایش داده‌های پیچیده

۴- دشواری کمتر در نمایش مفهوم سلسله‌مراتب و مفهوم وراثت

۵- افزایش سرعت عملیاتی سیستم در بازیابی اطلاعات

۶- عدم نیاز به نرمالترسازی رابطه

معایب:

۱- پیچیدگی

۲- عدم تقارن صفات

کلید در مدل رابطه‌ای

در مدل رابطه ای چند مفهوم در بحث کلید داریم که عبارتند از:

۱. ابر کلید (super key)

۲. کلید کاندید (candidate key)

۳. کلید اصلی (primary key)

۴. کلید دیگر (alternate key)

۵. کلید خارجی (foreign key)

ابر کلید

تعریف- هر زیرمجموعه از مجموعه عنوان رابطه که یکتایی مقدار در گستره رابطه داشته باشد.

کلید کاندید

تعریف- هر زیرمجموعه از مجموعه عنوان رابطه که دو خاصیت زیر را داشته باشد کلید کاندید رابطه است:

۱- یکتایی مقدار

۲- کاهش‌ناپذیری

کلید اصلی

تعریف- یکی از کلیدهای کاندید رابطه که طراح انتخاب می‌کند و به سیستم معرفی می‌شود.

ضابطه‌های انتخاب:

۱- از نظر کاربر، شناسه معمول نوع موجودیت باشد.

۲- طول کوتاهتر داشته باشد.

کلید دیگر (بدیل)

تعریف- هر کلید کاندید، غیر از کلید اصلی، کلید دیگر نام دارد.

کلید خارجی

تعریف- دورابطه R_1 و R_2 را در نظر می‌گیریم. هر زیرمجموعه از صفات رابطه R_2 که هر مقدار معلومش با یک مقدار از کلید کاندید R_1 برابر باشد، کلید خارجی در رابطه R_2 است.

کلید خارجی برای نمایش ارتباطات بین انواع موجودیتها به‌کار می‌رود.

جلسه دهم

قواعد جامعیت پایگاه داده‌ها

آنچه در این جلسه می خوانید:

۱- جامعیت پایگاه داده‌ها
۲- عوامل نقض جامعیت
۳- انواع قواعد جامعیت
۴- قواعد کاربری و انواع آن
۵- متاقواعد و انواع آن
۶- راههای اعمال قواعد جامعیت

آنچه در این جلسه می خوانید:

۷- عملگرهای جبر رابطه‌ای
۸- کاربردهای جبر رابطه‌ای
۹- حساب رابطه‌ای
۱۰- مزایا و معایب مدل رابطه‌ای
۱۱- کاتالوگ در مدل رابطه‌ای
۱۲- اشیائی که اطلاعات آنها در کاتالوگ نگهداری می‌شود

هدفهای کلی: آشنایی با قواعد جامعیت پایگاه داده‌ها

هدفهای رفتاری: دانشجو در پایان این جلسه می‌تواند:

- جامعیت پایگاه داده‌ها و عوامل نقض آن را توضیح دهد.
- انواع قواعد جامعیت را نام برده و راههای اعمال آن را شرح دهد.
- جبر رابطه‌ای و حساب رابطه‌ای را توضیح دهد.
- مزایا و معایب مدل رابطه‌ای را نام ببرد.

جامعیت پایگاه داده‌ها

یعنی: صحت، دقت و سازگاری داده‌های ذخیره‌شده در پایگاه در تمام لحظات

عواملی که سبب نقض جامعیت می‌شوند:

- اشتباه در برنامه‌های کاربردی
- اشتباه در وارد کردن داده‌ها
- وجود افزونگی کنترل نشده
- توارد تراکنشها به گونه‌ای که داده نامعتبر ایجاد شود.
- خرابیهای سخت‌افزاری و نرم‌افزاری

انواع قواعد جامعیت

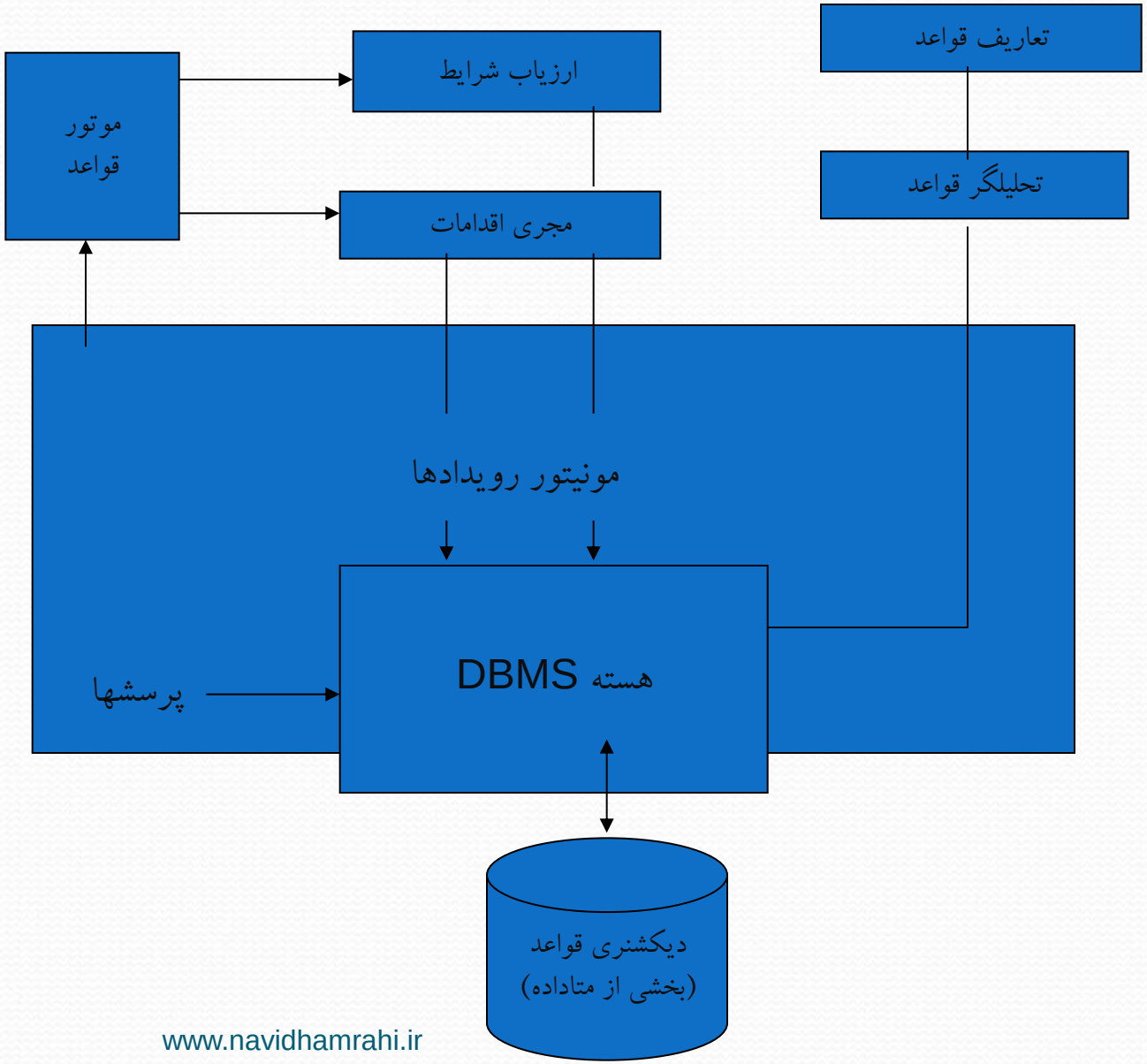
۱- قواعد کاربری (قواعد خاص)

۲- متا قواعد (قواعد عام)

۱- قواعد کاربری

قواعدی هستند که توسط کاربرمجاز تعریف می‌شوند. وابسته به داده‌های خردجهان واقع هستند، به این معنا که در مورد یک پایگاه داده خاص مطرح می‌شوند و عمومیت ندارند. به این قواعد، قواعد محیطی یا وابسته به داده و یا محدودیتهای جامعیت معنایی می گویند.

نمای ساده‌شده یک DBMS فعال



انواع قواعد کاربری در مدل رابطه‌ای

۱- محدودیت میدانی: ناظر بر یک میدان است و مقادیر مجاز آن را مشخص می کند.

۲- محدودیت صفتی (ستونی): ناظر به یک صفت است و بیان کننده نوع آن صفت است.

۳- محدودیت رابطه‌ای: ناظر بر یک رابطه است و مقادیر مجاز یک تغییر رابطه‌ای را مشخص می کند.

۴- محدودیت پایگاهی: ناظر بر دو یا بیش از دو متغیر رابطه‌ای است و به نحوی آنها را به یکدیگر مرتبط می کند.

۲- متاقواعد

قواعدی هستند که باید توسط هرسیستم رابطه‌ای در هر پایگاه داده رابطه‌ای اعمال شود، ناوابسته به داده‌های خاص هستند و عمومیت دارند.

انواع متاقواعد

۱- قاعده C_1 (جامعیت موجودیتی)

۲- قاعده C_2 (جامعیت ارجاعی)

قاعده C_1

ناظر به کلید اصلی است و چنین است:

هیچ جزء تشکیل‌دهنده کلید اصلی نمی‌تواند هیچ‌مقدار داشته باشد.

قاعده C_2

ناظر بر کلید خارجی است و چنین است:

اگر صفت خاصه A_i در رابطه R_2 کلید خارجی باشد در این صورت:

A_i در R_2 می‌تواند هیچ‌مقدار داشته باشد یا اینکه باید حتما مقداری داشته باشد که در رابطه مرجع R_1 وجود دارد. به عبارت دیگرمقدار کلید خارجی یک رابطه نمی‌تواند در رابطه مرجع وجود داشته باشد.

به قواعد C_1 و C_2 محدودیتهای ساختاری می‌گویند

محدودیتهای ناساختاری

- وابستگی‌های تابعی
- وابستگی‌های چندمقداری
- وابستگی‌های شمول
- وابستگی زمانی
- وابستگی برابری

راههای اعمال قواعد جامعیت

۱- معرفی کلید اصلی

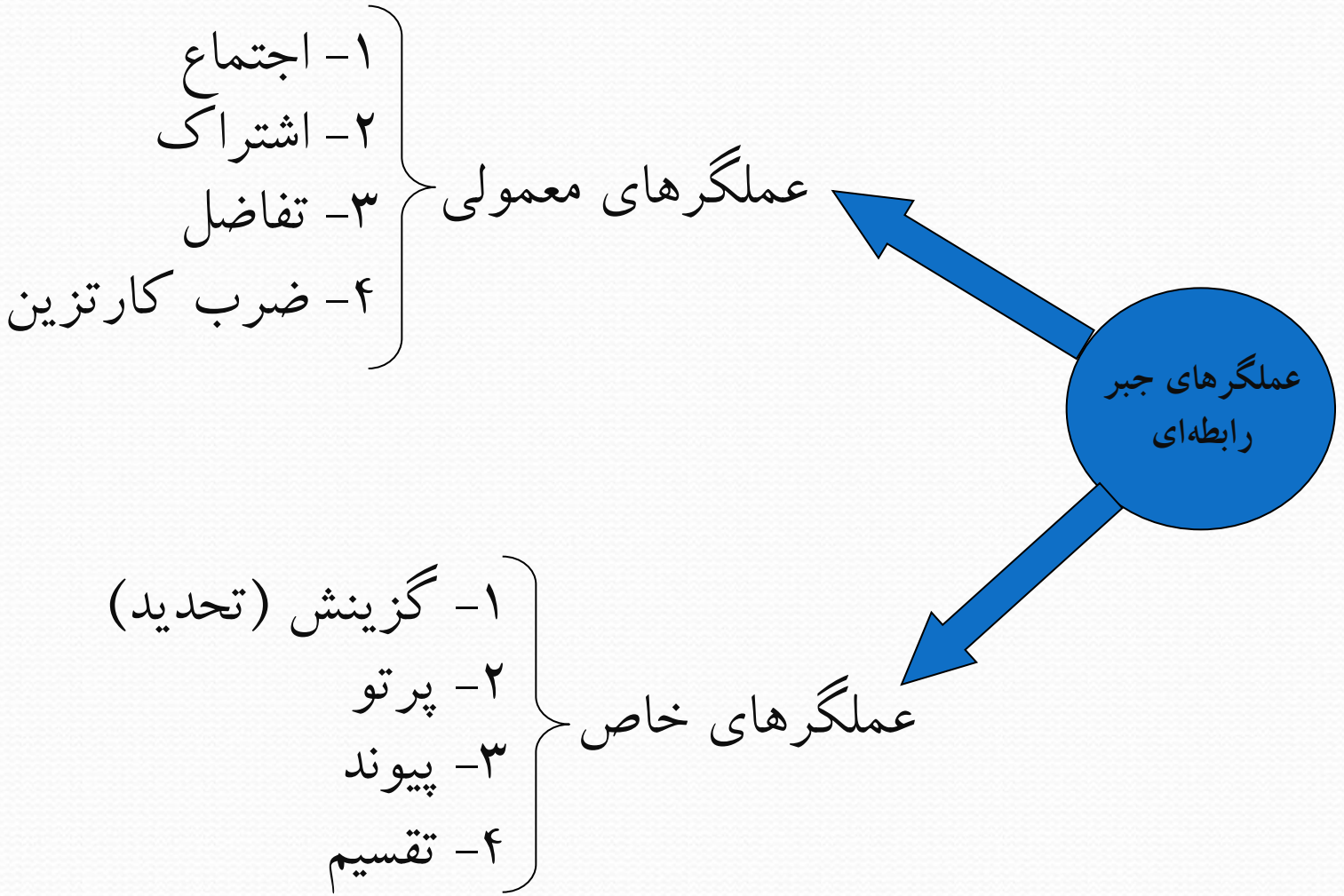
۲- اعلام هیچ مقدارناپذیری صفت

۳- معرفی کلید خارجی

۴- اعلان محدودیتهای مورد نظر، در شمای پایگاه داده‌ها

۵- نوشتن رهانا

۶- اعلان محدودیتها با استفاده از مکانیسم اظهار


عملگرهای جبر رابطه‌ای
عملگرهای معمولی
۱- اجتماع
۲- اشتراک
۳- تفاضل
۴- ضرب کارتزین
عملگرهای خاص
۱- گزینش (تحدید)
۲- پرتو
۳- پیوند
۴- تقسیم

عملگر گزینش تاپلهایی از یک رابطه را گزینش می‌کند (سطرهایی از جدول را). این عملگر تک عملوندی است و به صورت زیر نوشته می‌شود:

R WHERE cond(s)

* cond(s) شرط یا شرایط گزینش را مشخص می‌کند.

علمگر پرتو، مقادیر صفت (صفاتی) از یک رابطه را به دست می‌دهد. تک عملوندی بوده و چنین نوشته می‌شود:

PROJECT R OVER(A_1,A_2,…,A_i)

عملگر پیوند دو عملوندی است و چنین نوشته می‌شود:

R_1 JOIN cond(s) R_2

که در آن cond(s) چنین است:

$R_1.A_i$ theta $R_2.B_i$

A_i یک صفت از رابطه R_1 و B_i یک صفت از رابطه R_2 است که از یک میدان مقدار می‌گیرند و باید ناهمنام باشند.

کاربردهای جبر رابطه‌ای

۱- بازیابی داده‌ها

۲- ذخیره‌سازی داده‌ها

۳- تعریف انواع رابطه‌های مشتق

۴- تعریف قواعد برای کنترل پایگاه داده‌ها

۵- تعریف داده‌ها به عنوان حیطه بعضی عملیات کنترل همروندی تراکنشها

۶- ضابطه تشخیص کامل بودن زبانهای رابطه‌ای

حساب رابطه‌ای

حساب رابطه‌ای، با جبر رابطه‌ای منطقا معادل است، یعنی برای هر عبارت جبر رابطه‌ای، یک عبارت معادل در حساب رابطه‌ای وجود دارد و برعکس. تفاوت آنها این است که جبر رابطه‌ای، دستوری است، اما حساب رابطه‌ای توصیفی است.

حساب رابطه‌ای

- حساب تاپلی: در این حساب مفهوم مهمی به نام متغیر تاپلی وجود دارد که تنها مقادیر مجازش، تاپلهای رابطه هستند.
- حساب میدانی: در این حساب، متغیر میدانی وجود دارد که از یک میدان مقدار می‌گیرد. در این حساب یک شرط اضافی به نام شرط عضویت وجود دارد.

در حساب رابطه‌ای تاپلی دو سور وجود دارد:

۱- سور وجودی: به صورت $\exists T(f)$ نوشته می‌شود، به این معنا که حداقل یک مقدار برای متغیر T وجود دارد به نحوی که f به "درست" ارزیابی شود.

۲- سور همگانی: به صورت $\forall T(f)$ نوشته می‌شود. یعنی به ازاء تمام مقادیر متغیر T، f به "درست" ارزیابی می‌شود.

مزایای مدل رابطه‌ای

۱- نمایش ساده

۲- مبنای ریاضی دارد.

۳- تامین کننده استقلال داده‌ای است.

۴- یک عنصر ساختاری اساسی به نام رابطه دارد.

۵- کامل است.

۶- عدم نیاز به مفهوم کلاسیک نظم.

۷- برای بهترسازی طراحی، ابزار طراحی بهتر دارد.

.
.
.

معایب مدل رابطه‌ای

۱- دشواری در نمایش داده‌های پیچیده

۲- دشواری در نمایش مفاهیمی مانند وراثت، تعمیم و ...

۳- نداشتن قواعد جامعیت ذاتی

۴- عدم امکان مدلسازی رفتار اشیاء

۵- فقدان اکمال برنامه‌سازی و اکمال محاسباتی

۶- دشواری در یافتن کلید کاندید

۷- جدا بودن داده‌ها از روشها

۸- مشکلات ناشی از وجود مفهوم هیچ‌مقدار

.

.

کاتالوگ در مدل رابطه‌ای

کاتالوگ در مدل رابطه‌ای، تعدادی رابطه است: یک پایگاه رابطه‌ای که خود سیستم ایجاد می‌کند

در کاتالوگ اطلاعاتی در مورد اشیاء ذیل نگهداری می‌شود:

- میدان
- رابطه
- صفت
- دید
- قواعد جامعیت
- توابع و رویه‌های تعریف‌شده توسط کاربر
- رهاناهای تعریف‌شده توسط کاربر
- کاربر
- شاخص و ساختارهای مرتبط با آن
- ضوابط ایمنی
- پرسشها
- برنامه‌های کاربردی

جلسه یازدهم

زبان SQL

آنچه در این جلسه می خوانید:

۱- انواع زبانهای رابطه ای
۲- امکانات مهم زبان SQL
۳- دستورات تعریف داده ها، دستور ایجاد و حذف جدول و دید
۴- دستور مجازشماری
۵- دستورات پردازش داده ها
۶- توابع جمعی (گروهی)

آنچه در این جلسه می خوانید:

۷- امکان LIKE و NOTLIKE

۸- آزمون تست وجود هیچمقدار در یک ستون

۹- امکان UNION و UNION ALL

۱۰- امکان GROUP BY: (گروه‌بندی)

۱۱- امکان HAVING

۱۲- امکان BETWEEN

هدفهای کلی: آشنایی با زبان SQL

هدفهای رفتاری: دانشجو در پایان این جلسه می‌تواند:

- امکانات زبان SQL را بیان کند.
- انواع دستورات تعریف داده‌ها و پردازش داده‌ها یاد گرفته و می‌تواند استفاده کند.

آشنایی با یک زبان رابطه ای : SQL

انواع زبانهای رابطه‌ای

- SQUARE
- SEQUEL
- SQL
- QUEL
- QBE
- DATALOG

امکانات مهم زبان SQL

- دستورات تعریف داده‌ها
- دستورات مجازشماری
- دستورات پردازش داده‌ها
- دستورات پردازش داده‌ها به طور ادغام‌شدنی
- دستورات نوشتن ماژول و رویه
- دستورات کنترل جامعیت
- دستورات کنترل تراکنشها

۱. دستورات تعریف داده‌ها

۱- ۱- تعریف شما :

CREATE SCHEMA

AUTHORIZATION USER

{base – table definition , view definition , grant-operation }

۱. دستورات تعریف داده‌ها

۱- ۲- انواع داده‌ای:

CHARACTER[(length)]

INTEGER

DECIMAL [(precision[,scale])]

SMALLINT

DOUBLE PRECISION

REAL

FLOAT [(precision)]

NUMERIC [(precision[,scale])]

۱. دستورات تعریف داده‌ها

۱- ۳- دستور ایجاد جدول :

CREATE SCHEMA

AUTHORIZATION …

CREATE TABLE …

۱. دستورات تعریف داده‌ها

۱- ۴- دستور حذف جدول:

DROP TABLE table-name

۱. دستورات تعریف داده‌ها

۱- ۵- دستور ایجاد و حذف دید:

CREATE VIEW …

DROP VIEW …

۱. دستورات تعریف داده‌ها

۱- ۶- دستور مجاز شماری:

GRANT , REVOKE

با این دستور حق انجام یک یا بیش از یک عمل به کاربر (کاربرانی) داده می‌شود.

۲- دستورات پردازش داده ها

۲- ۱- دستور بازیابی (SELECT):

```
SELECT [ALL | DISTINCT] item(s)-list
FROM table(s)-name
[WHERE condition(s)]
[GROUP BY column(s)]
[HAVING conditions(s)]
```

توابع جمعی (گروهی)

- COUNT: تعداد مقادیر اسکالر را به دست می‌دهد.
- SUM: حاصل جمع مقادیر اسکالر در یک ستون را به دست می‌دهد.
- AVG: میانگین مقادیر اسکالر در یک ستون را به دست می‌دهد.
- MAX: بیشترین مقدار اسکالر در یک ستون را به دست می‌دهد.
- MIN: کمترین مقدار اسکالر در یک ستون را به دست می‌دهد.

یک مثال در مورد استفاده از توابع جمعی

```
SELECT MIN(GRADE) , MAX(GRADE)
        FROM STCOT
         WHERE TR='2'
                AND
                YRYR='78-79'
                AND
                COID='COM222';
```

بالاترین و پایین ترین نمره در درس COM222 در ترم دوم ۸۰-۷۹ را بازیابی می کند.

امکان LIKE و NOTLIKE

با این دو امکان می‌توان داده‌های مورد نظر را با دادن یک رشته کاراکتری به عنوان نشانوند جستجو و بیان شرایط مورد نظر، بازیابی کرد.

مثال: مشخصات استادانی را بدهید که نام آنها با AR شروع شده باشد.

```
SELECT *
        FROM PROF
                WHERE PRNAME LIKE ‘AR%’;
```

آزمون تست وجود هیچمقدار در یک ستون

می‌توان با امکان IS NULL وجود هیچ مقدار در یک ستون را تست کرد.

مثال: شماره دانشجویانی را بدهید که نمره آنها در درس SOC333 در ترم دوم ۷۹-۷۸ هنوز اعلام نشده است؟

```
SELECT STID
        FROM  STCOT
          WHERE COID='SOC333'
                AND
                TR='2'
                AND
                YRYR='78-79'
                AND
                GRADE  IS  NULL
```

امکان UNION و UNION ALL

با این امکان، می‌توان عملکرد عملگر UNION جبری را برنامه‌سازی کرد. حاصل این عملکرد، جدولی است که سطرهای تکراری ندارد. اگر از گزیدار ALL استفاده شود، سیستم دیگر سطرهای تکراری را در صورت وجود، حذف نمی‌کند.

امکان GROUP BY: (گروه‌بندی)

با این امکان می‌توان سطرهای جدول را حسب مقادیر یک ستون ساده گروه‌بندی کرد به نحوی که در هر گروه، مقدار آن ستون یکسان باشد.

مثال: جدول STCOT را برحسب مقادیر ستون COID منطقا گروه‌بندی کنید.

```
SELECT COID, AVG(GRADE) AS
AVGGR
        FROM STCOT
                GROUP  BY  COID;
```

امکان HAVING

با این امکان می‌توان شرط (شرایطی) ناظر به گروه سطرها اعلان کرد. نقش این امکان در واقع همان نقش WHERE در سطر است. توجه داشته باشید که HAVING همیشه با GROUP BY می‌آید.

مثال: شماره درسهایی را بدهید که در ترم دوم ۷۹-۷۸ کمتر از ۱۰ دانشجو در آنها ثبت نام کرده باشند.

```
SELECT STCOT.COID
        FROM  STCOT
                WHERE TR='2'
                    AND
                    YRYR='78-79'
       GROUP  BY  COID
       HAVING  COUNT(*)<10
```

امکان BETWEEN

شکل کلی این امکان چنین است:

Scalar-expression [NOT] BETWEEN Scalar-expression AND Scalar-expression

مثال: شماره دانشجویانی را بدهید که نمره آنها در درس HIS444 در ترم اول ۷۸-۷۷ بین ۱۵ و ۱۹ باشد.

```
SELECT  STCOT.STID
        FROM  STCOT
          WHERE TR='1'
                AND
                YRYR='78'79'
                AND
                GRADE BETWEEN '15'  AND  '19'
```

جلسه دوازدهم

ادامه زبان SQL

آنچه در این جلسه می خوانید:

۱- دستورات عملیات ذخیره‌سازی (درج، حذف و بهنگام‌سازی)
۲- جمع‌بندی در مورد امکانات نسخه SQL1
۳- دستوراتی که در نسخه SQL2 تغییر کردند یا به آن اضافه شدند
۴- امکانات جامعیتی
۵- امکانات ایمنی
۶- امتیازها

آنچه در این جلسه می خوانید:

۷- دستوراتی که در نسخه SQL3 تغییر کردند یا به آن اضافه شدند

۸- نوع داده مجرد

۹- تعریف رهانا و مزایا و موارد کاربرد آن

۱۰- تراکنش

۱۱- طرحهای اجرای تراکنشها

۱۲- تکنیکهای کنترل همروندی

هدفهای کلی: آشنایی با زبان SQL

هدفهای رفتاری: دانشجو در پایان این جلسه می‌تواند:

- از دستورات ذخیره‌سازی استفاده کند.
- امکانات اضافه‌شده در نسخه SQL2 را بیان کند.
- امکانات اضافه‌شده در نسخه SQL3 را بیان کند.

۲- دستورات پردازش داده ها

۲- ۲- دستورات عملیات ذخیره‌سازی :

برای انجام عملیات سه‌گانه ذخیره‌سازی سه دستور DELETE، INSERT و UPDATE وجود دارد.

دستور UPDATE

شکل کلی این دستور چنین است:

```
UPDATE  table-name
        SET assignment-commalist
            [WHERE Condition(s)]
```

مثال: تعداد واحد درسهای عملی را یک واحد کاهش دهید.

```
UPDATE  COT
        SET  CREDIT=CREDIT-1
             WHERE  COTYPE='p';
```

دستور DELETE

شکل کلی این دستور چنین است:

```
DELETE
    FROM  table-name
        [WHERE Cond(s)]
```

مثال: درسهای دانشجوی با شماره ۷۸۱۱۰۵۵۵ را در ترم دوم سال ۷۹-۷۸ حذف کنید

```
DELETE
    FROM STCOT
        WHERE  STID='78110555'
            AND
            TR='2'
            AND
            YRYR='78-79';
```

دستور INSERT

این دستور دو شکل کلی دارد:

```
INSERT  INTO table-name
        Values(one row);

INSERT  INTO table-name
        Subquery;
```

در شکل دوم، تعدادی سطر در جدول درج می‌شوند.

مثال: اطلاعات درسی زیر را در جدول درج نمائید.

<78110888,COM888,2,78-79,12>

INSERT INTO STCOT

VALUES <‘78110888’, ‘COM888’, 2, 78-79, 12>;

جمع‌بندی در مورد امکانات نسخه SQL1

۱- نارویه‌ای است.

۲- دارای کمال ساختاری است.

۳- زبان استانده سیستم‌های رابطه‌ای موجود است.

۴- تمام انواع داده‌ای ساده را دارد.

۵- عملگرهای بسیار قوی دارد.

۶- یادگیری آن ساده است.

۷- استقلال داده‌ای را تامین می‌کند.

۸- هم به صورت مستقل و هم به صورت ادغام‌شده قابل استفاده است.

دستوراتی که در نسخه SQL2 تغییر کردند یا به آن اضافه شدند:

۱. دستور تعریف داده‌ها
۲. دستور تعریف میدان
۳. انواع داده‌ای
۴. دستور تغییر میدان
۵. دستور حذف میدان
۶. دستور ایجاد جدول
۷. دستور تغییر جدول
۸. دستور حذف جدول
۹. تعریف جدول موقت
۱۰. دستور بازیابی
۱۱. امکانات جامعیتی
۲۱. امکانات ایمنی
۳۱. SQL پویا

امکانات جامعیتی

کاربر می‌تواند زمان وارسی جامعیت پایگاه را (براساس محدودیتها و قواعد داده‌شده) خود مشخص کند. می‌تواند درخواست کند که این وارسی بلافاصله باشد یا با تاخیر. برای این منظور از امکان زیر استفاده می‌شود:

SET CONSTRAINTS (constraint-name)

{DEFERRED | IMMEDIATE}

امکانات ایمنی

کاربر می‌تواند امتیاز دستیابی به اشیاء تعریف‌شده در شما را به منظور انجام عمل مشخص، به کاربر دیگری بدهد.

برای اعطای امتیاز از دستور GRANT و برای لغو آن از دستور REVOKE استفاده می‌شود.

امتیازها

امتیازهایی که یک کاربر می‌تواند به یک کاربر دیگر بدهد:

- USAGE: اجازه استفاده از یک میدان
- SELECT: اجازه دستیابی به تمام ستونهای یک جدول
- INSERT(X): اجازه درج یک مقدار در ستون X از یک جدول
- INSERT: اجازه درج تمام ستونهای یک جدول
- UPDATE(X): اجازه بهنگام‌سازی ستون X از یک جدول
- UPDATE: اجازه بهنگام‌سازی تمام ستونهای یک جدول
- DELETE: اجازه حذف سطر(ها) از یک جدول
- REFRENCES(X): اجازه ارجاع به ستون X از یک جدول در تعریف محدودیتهای جامعیتی

دستوراتی که در نسخه SQL3 تغییر کردند یا به آن اضافه شدند:

۱. تعریف نوع داده مجرد
۲. تعریف زیرنوع
۳. امکان تعریف زیرجدول و زبرجدول
۴. امکان ارث‌بری ستونهای جدول جدید از ستونهای جدول نامدار موجود
۵. رویه‌ای شدن زبان
۶. رهانا
۷. نوشتن رویه
۸. SQL ادغام‌شده
۹. واسط درخواست فراخوان
۰۱. دستورات کنترل تراکنشها

نوع داده مجرد

شاید مهمترین قابلیتی که در زمینه انواع داده‌ای در SQL3 افزوده شد، امکان تعریف نوع داده مجرد (ADT) باشد. با این نوع داده می‌توان مجموعه‌ای از صفات و روالها را تعریف کرد. نوع داده مجرد نمونه‌هایی دارد و هر نمونه هم به نوبه خود، همه صفات و روالهای تعریف‌شده را دارد.

ساختهای زبانی افزوده‌شده به SQL3، که در تعریف توابع و رویه‌ها استفاده می‌شود:

- احکام مرکب (تودرتو)
- احکام حلقه گردانی)
- احکام شرطی
- احکام انتساب
- احکام اعلان اشتباهات در برنامه

رهانا

تعریف- قاعده (محدودیت) یا قواعدی است که قبل یا بعد از بروز یک رویداد در پایگاه داده‌ها (معمولا یک عمل تغییردهنده داده‌ها) باید اعمال شود. این قاعده در سطح برنامه‌سازی، به صورت یک رویه از پیش تعریف‌شده است که به طور شرطی یا غیرشرطی، قبل یا بعد از انجام یک عمل در پایگاه داده‌ها، به طور اتوماتیک اجرا می‌شود.

رهانا

مزایا

۱.اجرای آن تحت کنترل متمرکز سیستم و نظارت مدیر پایگاه داده‌ها است.

۲.با وجود آن دیگر نیازی نیست که محدودیتهای مورد نظر در هر برنامه کاربردی اعمال شوند و کنترلهای لازم انجام شود.

۳.برای معماری C/S DB بسیار مناسب است.

کاربرد رهانا

۱. اعمال قواعد جامعیت.

۲. اعمال قواعد ایمنی مبتنی بر مقادیر

۳. درج رکوردهای ثبت عملیات در فایل ثبت تراکنشها

۴. تولید نسخه‌ای از داده‌های ذخیره‌شده در یک سایت دیگر.

۵. در انتشار اتوماتیک عملیات در پایگاه داده‌های توزیع‌شده.

۶. در بهنگام‌سازی دیدها در سیستم‌های رابطه‌ای

تراکنش

تعریف: به بیان ساده عملی است که تغییری در پایگاه داده‌ها ایجاد کند. این عمل که طبعا به صورت یک برنامه یا بخشی از یک برنامه است، یا باید به تمامی اجرا شود و یا اصلا اجرا نشود و می‌گوییم که تراکنش خاصیت تجزیه‌ناپذیری دارد.

خواص

۱. *تجزیه‌ناپذیری*

۲. *سازگاری*

۳. *جدایی (انفراد)*

۴. *مانایی (دوام)*

طرحهای اجرای تراکنشها

- طرح اجرای متوالی
- طرح اجرای همروند

تکنیکهای کنترل همروندی

- تکنیک قفل‌گذاری
- تکنیک زمانمهر
- تکنیک چند نسخه‌سازی
- تکنیک تایید (تصدیق)

جلسه سیزدهم

دید در مدل رابطه‌ای

آنچه در این جلسه می خوانید:

۱- دیدهای رابطه‌ای
۲- دید در SQL
۳- مزایا و معایب دید
۴- عملیات در دیدهای رابطه‌ای
۵- انواع دید از نظر پذیرش عملیات ذخیره‌سازی
۶- انواع دیدهای پذیرا
۷- انواع دیدهای ناپذیرا
۸- ویژگیهای دیدهای قابل بهنگام‌سازی

هدفهای کلی: آشنایی با دید در مدل رابطه‌ای

هدفهای رفتاری: دانشجو در پایان این جلسه می‌تواند:

- دید را تعریف کند و دستورات SQL مرتبت با آن را بکار برد.
- مزایای استفاده از دید را برشمارد.
- عملیات قابل انجام روی دید را شرح دهد.
- انواع دید از لحاظ پذیرا و ناپذیرا را نام ببرد.

دیدهای رابطه‌ای

دید در مدل رابطه‌ای نوعی رابطه است، بنابراین می‌توان آن را به کمک یک عبارت جبر رابطه‌ای یا حساب رابطه‌ای به شکل زیر تعریف کرد:

View name= Relational expression

دید در SQL

دستور ایجاد دید

CREATE VIEW Viewname [(Column-name(s))]

AS Subquery

.

.

.

[WITH [CASCADE|LOCAL] CHECK OPTION];

مثال

```
CREATE VIEW MAPHSTUD( STNUM, STLEV, STAREA)
        AS  SELECT  STID, STDEG, STMJR
        FROM  STT
                WHERE  STMJR='Math' OR STMJR='Phys'
        WITH LOCAL CHECK OPTION;
```

دید در SQL

دستور حذف دید

DROP VIEW Viewname {restrict | cascade}

مثال:

```
DROP VIEW MAPHSTUD CASCADE;
```

مزایای دید

۱. تامین کننده محیط انتزاعی برای کاربران سطح خارجی

۲. تامین کننده پویایی بالا در تعریف پایگاه توسط کاربر

۳. تسهیل کننده واسط کاربر برنامه‌ساز با پایگاه

۴. امکانی است برای کوتاه‌نویسی یا ماکرونویسی پرسشها

۵. تامین کننده اشتراک داده‌ای

۶. تامین کننده نوعی مکانیسم خودکار ایمنی داده‌ها

۷. تامین کننده استقلال داده‌ای فیزیکی و منطقی

۸. امکان تعریف شیئ با اندازه‌های مختلف

معایب دید

۱. ایجاد فزونکاری در سیستم برای انجام تبدیل خارجی/ادراکی و احیانا خارجی/خارجی

۲. عدم امکان انجام عملیات ذخیره‌سازی در بسیاری از گونه‌های دید و در نتیجه ایجاد محدودیت برای کاربر

عملیات در دیدهای رابطه‌ای

بازیابی از دید رابطه‌ای

چون دید خود نوعی رابطه است، پس برای بازیابی از دید هم یک عبارت جبری یا حسابی می‌نویسیم

مثال بازیابی از یک دید با استفاده از دستورات SQL

```
CREATE VIEW   V1
AS  SELECT STID, STDEG
        FROM  STT
            WHERE  STPROG='Math'

SELECT  STID
FROM    V1
        WHERE  STDEG='bs';
```

عملیات در دیدهای رابطه‌ای

عملیات ذخیره‌سازی در دیدهای رابطه‌ای

برای انجام عملیات ذخیره‌سازی در دید از همان دستورات سه‌گانه INSERT، UPDATE و DELETE استفاده می‌شود.

دیدها از نظر پذیرش عملیات ذخیره‌سازی

۱. دیدهای پذیرا

۲. دیدهای ناپذیرا

دیدهای پذیرا

۱. دیدهای گزینشی

۲. دید گزینش – پرتوی دارای کلید رابطه مبنا

۳. دید پیوندی CK-CK

۴. دید پیوندی CK-FK

۵. دید حاصل اجتماع، اشتراک و تفاضل دو رابطه

دید گزینشی حاصل عملگر گزینش در یک رابطه است.

در دید "گزینش-پرتوی دارای کلید رابطه مبنا"
علاوه بر گزینش تاپلهایی از رابطه مبنا، عملگر
پرتو نیز اعمال شده است. این نوع دید را
اصطلاحا دید دارای کلید با تاپلهای ناقص
می گوئیم.

دید پیوندی CK-CK، حاصل پیوند دو رابطه روی کلید کاندید مشترک آنها است و در عملیات ذخیره‌سازی مشکلی ندارد

دید پیوندی CK-FK، حاصل پیوند روی کلید کاندید یک رابطه و کلید خارجی رابطه دیگر است. این دید در حذف مشکل دارد، زیرا با حذف یک تاپل از این دید، در هریک از دو رابطه مبنا، یک تاپل حذف می‌شود و چنانچه کاربر بخواهد محتوای دید خود را نمایش دهد، تاپلهای دیگر هم از دید او حذف می‌شوند که درخواست نکرده است.

دید حاصل اجتماع، اشتراک و تفاضل دو رابطه در عملیات ذخیره‌سازی مشکلی ندارد، به شرط آنکه سیستم بتواند تشخیص دهد که عمل درخواست‌شده، در کدامیک از دو رابطه مبنا انجام شود.

مشکلات مهمتر دیدهای پذیرا

- بروز عارضه جانبی در خود دید
- بروز عارضه جانبی در دیدهای دیگر
- نقض قاعده جامعیت
- بروز فزونکاری در سیستم
- تغییر ماهیت عمل درخواست شده
- تعدد تبدیلات و مشکل تصمیم‌گیری

دیدهای ناپذیرا

۱. دید پرتوی یا گزینشی فاقد کلید رابطه مبنا

۲. دید پیوندی NK-NK

۳. دید پیوندی FK-FK

۴. دید حاوی صفت مجازی

۵. دید حاصل تقسیم

ویژگیهای دیدهای قابل بهنگام‌سازی

۱. عبارت تعریف‌کننده محدوده دید، یک عبارت معتبر SELECT باشد.

۲. در کلاز FROM، عبارت SELECT، فقط یک جدول وجود داشته باشد.

۳. جدول قیدشده در کلاز FROM، یک جدول مبنا یا یک دید قابل بهنگام‌سازی باشد.

۴. در عبارت SELECT، item-list ستونهای مورد نظر باید در جدول مبنا متناظر باشد.

۵. در عبارت SELECT نباید کلاز GROUP BY و HAVING و گزینه DISTINCT وجود داشته باشد.

۶. در کلاز WHERE عبارت SELECT نباید عبارتی حاوی کلاز FROM باشد، به گونه‌ای که در آن به همان جدولی ارجاع داده شده باشد که در کلاز FROM قبلی به آن ارجاع شده است.

جلسه چهاردهم

طراحی پایگاه داده‌های رابطه‌ای

آنچه در این جلسه می خوانید:

۱- مراحل طراحی پایگاه داده‌ها
۲- نمودار مراحل اساسی طراحی پایگاه داده‌ها
۳- روشهای طراحی منطقی پایگاه داده‌ها
۴- روش طراحی بالا به پایین
۵- روش سنتز رابطه‌ای
۶- روش ترکیبی
۷- خصوصیات طراحی خوب
۸- چهارده روش تبدیل نمودارER به رابطه‌ها

هدفهای کلی: آشنایی با طراحی پایگاه داده‌های رابطه‌ای

هدفهای رفتاری: دانشجو در پایان این جلسه می‌تواند:

- مراحل طراحی پایگاه داده‌ها را بیان کند.
- روشهای طراحی منطقی پایگاه را شرح دهد.
- روشهای تبدیل نمودار ER به رابطه را نام برده و توضیح دهد.
- خصوصیات یک طراحی خوب را بیان کند.

طراحی پایگاه داده‌های رابطه‌ای (روش بالا به پایین)

مراحل طراحی پایگاه داده‌ها

۱. مطالعه و شناخت خرد جهان واقع

۲. انجام عملیات مهندسی نیازها

۳. مدلسازی معنایی داده‌ها

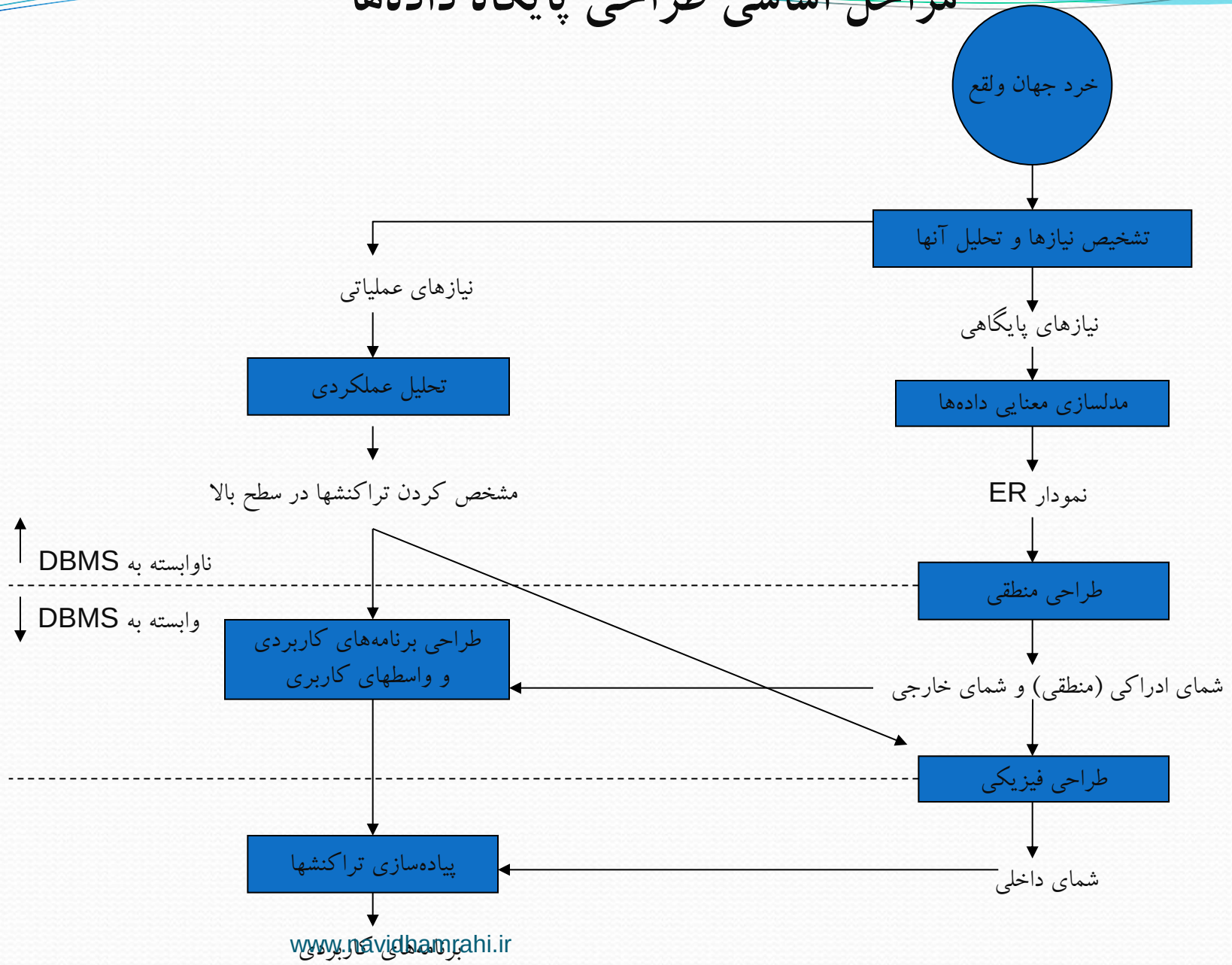
۴. طراحی منطقی پایگاه داده‌ها

۵. طراحی فیزیکی پایگاه داده‌ها

۶. انجام تحلیل عملکردی: تعیین تراکنشها

۷. طراحی برنامه‌های کاربردی و واسطهای کاربری

مراحل اساسی طراحی پایگاه داده‌ها



www.navidhamrahi.ir

روشهای طراحی منطقی پایگاه داده‌ها

۱. روش بالا به پایین

۲. روش سنتز رابطه‌ای

۳. روش ترکیبی

روش طراحی بالا به پایین

در این روش ابتدا مدلسازی معنایی داده‌ها انجام می‌شود. سپس مدلسازی را به مجموعه‌ای از رابطه‌ها تبدیل می‌کنیم. حاصل این فرایند، معمولا مجموعه‌ای از رابطه‌های خوش‌طرح است

روش سنتز رابطه‌ای

در این روش مجموعه صفات خردجهان واقع را مشخص می‌کنیم. سپس با تحلیل قواعد و محدودیتهای ناظر به صفات و تشخیص وابستگیهای بین آنها، صفات را متناسبا باهم سنتز می‌کنیم. هر گروه از صفات عنوان یک رابطه را به ما می‌دهد به گونه‌ای که حتی‌الامکان در مطلوبترین صورت نرمال باشد. هر رابطه بدست‌آمده نمایشگر بخشی از خردجهان واقع بوده، ضوابط یک طرح خوب و قابل دفاع را خواهد داشت.

روش ترکیبی

در این روش، ابتدا طراحی را با روش بالا به پایین انجام می‌دهیم و سپس با دخالت دادن برخی محدودیتهای جامعیتی که در مرحله مدلسازی معنایی داده‌ها منظور نشده‌اند و با تحلیل هر رابطه، مرحله نرمالترسازی رابطه‌ها را انجام می‌دهیم تا به مجموعه‌ای از نرمالترین رابطه‌ها برسیم.

خصوصیات طراحی خوب

۱- نمایش واضحی از خردجهان واقع باشد.

۲- نمایش صحیحی از خردجهان واقع باشد.

۳- نمایش جامعی از خردجهان واقع باشد.

۴- تمام قواعد جامعیتی که قابل اعمال در هر مرحله از طراحی منطقی باشند، در طراحی منظور شده باشند.

۵- معنای هریک از صفات از هر نوع موجودیت به درستی رعایت شده باشد.

۶- کمترین میزان افزونگی را داشته باشد.

ادامه

خصوصیات طراحی خوب

۷- کمترین میزان اختلاط اطلاعات را داشته باشد.

۸- انعطاف‌پذیری داشته باشد.

۹- کمترین دشواری در انجام عملیات ذخیره‌سازی

۱۰- کمترین میزان هیچمقدار

۱۱- هیچ اطلاع جعلی در اثر انجام عملیات در پایگاه داده‌ها پدید نیاید.

۱۲- با درنظر گرفتن طراحی فیزیکی و تاثیر آن در طراحی منطقی، بیشترین کارایی برای سیستم کاربردی پایگاه داده‌ها تامین شود.

روش تبدیل نمودار ER به رابطه‌ها

حالت اول

- تعداد نوع موجودیت: n>=2
- وضع موجودیتها: مستقل
- چندی ارتباط: N:M

در این حالت، n+1 رابطه لازم است. n رابطه برای n موجودیت مستقل و یک رابطه برای نمایش ارتباط بین آنها. اگر ارتباط بین انواع موجودیتها، صفت ساده یا مرکب چندمقداری نداشته باشد، در این صورت کلید کاندید رابطه نمایشگر ارتباط بین n موجودیت، از ترکیب کلیدهای کاندید n رابطه نمایشگر n موجودیت به دست می‌آید. سپس کلیدهای خارجی، اجزاء تشکیل‌دهنده کلید کاندید رابطه نمایشگر ارتباط هستند.

روش تبدیل نمودار ER به رابطه‌ها

حالت دوم

- تعداد نوع موجودیت: دو
- وضع موجودیتها: مستقل
- چندی ارتباط: 1:N

در این حالت، دو رابطه کفایت می‌کند: یک رابطه برای نمایش نوع موجودیت طرف یک، یک رابطه برای نمایش نوع موجودیت طرف N و نیز ارتباط 1:N. در رابطه اخیر، کلید کاندید رابطه اول، به عنوان کلید خارجی رابطه دوم، ارتباط مورد نظر را نشان می‌دهد و جزء تشکیل‌دهنده کلید کاندید رابطه نیست.

روش تبدیل نمودار ER به رابطه‌ها

حالت سوم

- تعداد نوع موجودیت: دو
- وضع موجودیتها: مستقل
- چندی ارتباط: 1:1

در این حالت، دو رابطه لازم است: یک رابطه برای نمایش یکی از دو نوع موجودیت و رابطه‌ای دیگر برای نمایش نوع موجودیت دیگر و ارتباط بین دو نوع موجودیت.

روش تبدیل نمودار ER به رابطه‌ها

حالت چهارم

- تعداد نوع موجودیت: یک
- وضع موجودیتها: مستقل
- چندی ارتباط: N:M

در این حالت، ارتباط با خود داریم. دو رابطه لازم است: یک رابطه برای نمایش خود نوع موجودیت و دیگری برای نمایش ارتباط، اعم از اینکه مشارکت الزامی باشد یا نباشد.

روش تبدیل نمودار ER به رابطه‌ها

حالت پنجم

- تعداد نوع موجودیت: یک
- وضع موجودیتها: مستقل
- چندی ارتباط: 1:N

در این حالت، هم "ارتباط با خود" داریم و یک رابطه کفایت می کند.

روش تبدیل نمودار ER به رابطه‌ها

حالت ششم

- تعداد نوع موجودیت: یک
- وضع موجودیتها: مستقل
- چندی ارتباط: 1:1

در این حالت یک رابطه کفایت می‌کند، به شرط آنکه مشارکت در ارتباط الزامی باشد. البته می‌توان با دو رابطه هم طراحی کرد.

روش تبدیل نمودار ER به رابطه‌ها

حالت هفتم
نمایش موجودیت ضعیف

موجودیت ضعیف دارای شناسه یکتا نیست، بلکه صفت ممیزه دارد. برای نمایش این نوع موجودیت در طراحی پایگاه رابطه‌ای، یک رابطه طراحی می‌کنیم که در عنوان آن، صفات موجودیت ضعیف و کلید کاندید موجودیت قوی که با آن ارتباط دارد (موجودیت شناسا)، وجود دارند.

روش تبدیل نمودار ER به رابطه‌ها

حالت هشتم
وجود صفت چندمقداری

اگر MVA، یک صفت (ساده یا مرکب) چندمقداری، EID شناسه موجودیت E و A_1 ، A_2 ... و A_i سایر صفات تک‌مقداری موجودیت E باشند، در این صورت برای نمایش این نوع موجودیت، دو رابطه لازم است:

$$R_1(\frac{EID}{C.K.}, A_1, A_2, \ldots, A_i)$$

$$R_2(\frac{EID, MVA}{C.K.})$$

روش تبدیل نمودار ER به رابطه‌ها

حالت نهم

ارتباط IS-A یا "گونه‌ای است از ..."

در این حالت یک روش این است که موجودیت سطح بالاتر با یک رابطه نشان داده می‌شود و هر نوع موجودیت سطح پایین‌تر با صفات خاص خودش همراه با شناسه اصلی موجودیت سطح بالاتر، در رابطه دیگری نمایش داده می‌شود.

روش تبدیل نمودار ER به رابطه‌ها

حالت دهم

وجود دسته (طبقه) در ارتباط IS-A

در این ارتباط وقتی زیرنوع دسته (طبقه) داریم و صفات شناسه زبرنوعها، متفاوت باشند، یک طرز طراحی این است که این زیرنوع را با یک رابطه نشان دهیم. کلید این رابطه، یک کلید جایگزین (شناسه ساختگی) است و بعلاوه یک صفت دیگر در آن درنظر می‌گیریم نشان‌دهنده نوع آن (از نوع کدام زبرنوع است). هر زبرنوع را هم با یک رابطه نشان می‌دهیم و صفت کلید زیرنوع را به هریک از رابطه‌های نشان‌دهنده زبرنوعها اضافه می‌کنیم.

روش تبدیل نمودار ER به رابطه‌ها

حالت یازدهم

ارتباط IS – A PART – OF یا "جزئی است از ..."

در این حالت، برای موجودیت کل (شامل) یک رابطه و برای هریک از موجودیتهای جزء (مشمول) نیز یک رابطه طراحی می‌شود. در رابطه نشان‌دهنده موجودیت جزء، کلید کاندید رابطه نشان‌دهنده موجودیت کل، آورده می‌شود و بقیه صفات این رابطه، خاص خود موجودیت جزء هستند.

روش تبدیل نمودار ER به رابطه‌ها

حالت دوازدهم

بیش از یک ارتباط بین دو نوع موجودیت

فرض می‌کنیم هر دو نوع موجودیت مستقل هستند. تعداد رابطه‌ها، بستگی به چندی هر ارتباط دارد. به صورت زیر عمل می‌کنیم:

- هر ارتباط با چندی N:M با یک رابطه نمایش داده می‌شود.
- نوع موجودیت سوی N و همه ارتباطهای با چندی 1:N، اگر مشارکت دو نوع موجودیت در همه ارتباطها الزامی باشد، را می‌توان با یک رابطه نمایش داد.

روش تبدیل نمودار ER به رابطه‌ها

حالت دوازدهم

بیش از یک ارتباط بین دو نوع موجودیت

- اگر مشارکت دو نوع موجودیت در یک ارتباط 1:N، الزامی نباشد، بهتر است برای چنین ارتباطی یک رابطه جداگانه طراحی شود که صفات آن، شناسه دو نوع موجودیت و در صورت وجود، صفات خود ارتباط هستند.
- در صورت وجود بیش از یک ارتباط 1:1 و الزامی بودن مشارکتها، هر دو نوع موجودیت و همه این ارتباطها را می‌توان با یک رابطه نمایش داد.
- اگر مشارکت در یک ارتباط 1:1 الزامی نباشد، بهتر است این ارتباط را با یک رابطه جداگانه نمایش داد.

روش تبدیل نمودار ER به رابطه‌ها

حالت سیزدهم

ارتباط سلسله‌مراتبی (پدر- فرزندی)

در تبدیل این نوع ارتباط بین موجودیتها به طراحی رابطه‌ای یک روش چنین است:

۱- برای نوع موجودیت ریشه، یک رابطه در نظر می‌گیریم.

۲- کلید کاندید این رابطه همان شناسه ریشه است.

۳- برای هر نوع موجودیت فرزند یک رابطه در نظر می‌گیریم و کلید کاندید هر نوع موجودیت فرزند، ترکیب شناسه آن نوع موجودیت با کلید کاندید نوع موجودیت پدرش است.

روش تبدیل نمودار ER به رابطه‌ها

حالت چهاردهم

تجمع (ارتباط با ارتباط)

در این حالت، ابتدا نمودار ER جزئی درون مستطیل نشان‌دهنده نوع موجودیت حاصل از عمل تجمیع را به رابطه‌های لازم تبدیل می‌کنیم. سپس با توجه به چندی ارتباط بین نوع موجودیت مرتبط با ارتباط و چگونگی مشارکت در ارتباط با ارتباط، طراحی را کامل می‌کنیم.

جلسه پانزدهم

نرمالترسازی رابطه‌ها

آنچه در این جلسه می خوانید:

۱- آنومالی در عملیات ذخیره‌سازی
۲- صورتهای نرمال
۳- وابستگی تابعی
۴- وابستگی تابعی نامهم (بدیهی)
۵- قواعد استنتاج آرمسترانگ
۶- مجموعه کاهش‌ناپذیر وابستگیهای تابعی
۷- وابستگی تابعی تام (کامل)
۸- رابطه 1NF
۹- رابطه 2NF

آنچه در این جلسه می خوانید:

۱۰- رابطه 3NF
۱۱- صورت نرمال بایس-کاد
۱۲- وابستگی تابعی چندمقداری
۱۳- رابطه 4NF
۱۴- وابستگی تابعی چندمقداری ادغام‌شده
۱۵- رابطه 5NF
۱۶- مزایا و معایب نرمالترسازی
۱۷- انواع افزونگی در پایگاه داده‌ها

هدفهای کلی: آشنایی با نرمالترسازی رابطه‌ها

هدفهای رفتاری: دانشجو در پایان این جلسه می‌تواند:

- صورتهای مختلف نرمال‌سازی را نام برده و هریک را توصیف کند.
- مزایا و معایب روش نرمالترسازی را شرح دهد.

طراحی پایگاه داده‌های رابطه‌ای

روش سنتز: نرمالترسازی رابطه‌ها

یک رابطه، هرچند نرمال، بازهم ممکن است در عملیات ذخیره‌سازی (درج، حذف و بهنگام‌سازی) مشکلاتی داشته باشد.

آنومالی در عملیات ذخیره‌سازی به هریک از سه وضع زیر گفته می‌شود:

- عدم امکان انجام یک عمل (که منطقا باید قابل انجام باشد)
- بروز پیامد بد پس از انجام یک عمل
- بروز فزونکاری در سیستم در انجام یک عمل

صورتهای نرمال

- صورت نخست نرمال (1NF)
- صورت دوم نرمال (2NF)
- صورت سوم نرمال (3NF)
- صورت نرمال بایس-کاد (BCNF)
- صورت چهارم نرمال (4NF)
- صورت پنجم نرمال (5NF)
- صورت نرمال میدان-کلیدی (DKNF)
- صورت نرمال تحدید-اجتماع (RUNF)

وابستگی تابعی

تعریف- فرض کنید که R یک متغیر رابطه‌ای و A و B دو زیرمجموعه دلخواه از عنوان R باشند. می‌گوییم B با A وابستگی تابعی دارد و چنین نمایش می‌دهیم: $A \rightarrow B$

اگر و فقط اگر در هر میدان ممکن از متغیر رابطه‌ای R، به هر مقدار A فقط یک مقدار B متناظر باشد.

وابستگی تابعی نامهم (بدیهی)

اگر در R(A,B,C,...) داشته باشیم: $B \subseteq A$ ، در این صورت: $A \rightarrow B$ یک وابستگی تابعی نامهم است. به بیان دیگر اگر B زیرمجموعه‌ای از A باشد، در این صورت $A \rightarrow B$ یک وابستگی نامهم است.

قواعد استنتاج آرمسترانگ

فرض: A، B، C و D زیرمجموعه‌هایی از صفات رابطه R باشند. قواعد زیر برقرارند:

۱. قاعده انعکاس: اگر $B \subseteq A$ آنگاه: $A \to B$

۲. قاعده تعدی(تراگذری): اگر $A \to B$ و $B \to C$ آنگاه $A \to C$

۳. قاعده افزایش: اگر $A \to B$ آنگاه $(A,C) \to (B,C)$

۴. قاعده تجزیه: اگر $A \to (B,C)$ آنگاه $A \to C$ و $B \to C$

۵. قاعده اجتماع: اگر $A \to C$ و $B \to C$ آنگاه $A \to (B,C)$

۶. قاعده ترکیب: اگر $C \to D$ و $A \to B$ آنگاه $(A,C) \to (B,D)$

۷. قاعده شبه تعدی: اگر $(C,B) \to D$ و $A \to B$ آنگاه $(A,C) \to D$

۸. قاعده یگانگی عمومی: اگر $C \to D$ و $A \to B$ آنگاه $A \cup (C-B) \to (B,D)$

مجموعه کاهش‌ناپذیر وابستگیهای تابعی

مجموعه‌ای از وابستگیهای تابعی R، به نام F را کاهش‌ناپذیر گوییم اگر:

۱. در F وابستگی تابعی افزونه نباشد.

۲. در سمت راست هر FD از F صفت ساده وجود داشته باشد.

۳. هیچ صفتی در سمت چپ FDهای F افزونه نباشد.

وابستگی تابعی تام (کامل)

اگر X و Y دو زیرمجموعه از مجموعه عنوان رابطه R باشند، می‌گوییم Y با X وابستگی تابعی تام دارد و چنین نشان می‌دهیم: $X \Rightarrow Y$

اگر و فقط اگر Y با X وابستگی داشته باشد. ولی با هیچ زیرمجموعه‌ای از X وابستگی تابعی نداشته باشد.

رابطه 1NF

تعریف- رابطه‌ای 1NF است اگر هر صفت خاصه آن در هر تاپل، تک‌مقداری باشد، به بیان دیگر، صفت چندمقداری نداشته باشد.

رابطه 2NF

تعریف- رابطه‌ای 2NF است اگر اولا 1NF باشد و ثانیا تمام صفات غیرکلید با کلید اصلی وابستگی تابعی تام داشته باشند. به عبارت دیگر هر صفت غیرکلید با کلید اصلی بطور کاهش‌ناپذیر وابسته باشد.

رابطه 3NF

تعریف- رابطه‌ای 3NF است اگر 2NF باشد و هر صفت غیرکلید با کلید اصلی، وابستگی تابعی بی‌واسطه داشته باشد.

صورت نرمال بایس-کاد

تعریف اول: رابطه‌ای BCNF است اگر در آن هر دترمینان، کلید کاندید باشد. این تعریف ساده‌شده و غیر صوری است.

صورت نرمال بایس-کاد

تعریف دوم: رابطه R(H)، با مجموعه وابستگیهای تابعی F، در BCNF است اگر برای هر وابستگی تابعی در F^+ به صورت $A \to B$ که در آن $A \subseteq R(H)$ و $B \subseteq R(H)$ است، حداقل یکی از دو حالت زیر برقرار باشد:

a: $A \to B$ یک وابستگی نامهم باشد یعنی $B \subseteq A$

b: A سوپر کلید رابطه R باشد.

صورت نرمال بایس-کاد

تعریف سوم: رابطه R در BCNF است هرگاه یک وابستگی تابعی به صورت $A \rightarrow B$ در مجموعه FDهای R وجود داشته باشد، A سوپرکلید R باشد.

صورت نرمال بایس-کاد

تعریف چهارم: رابطه‌ای BCNF است اگر و فقط اگر سمت چپ هر FD مهم (مطرح) و کاهش‌ناپذیر، کلید کاندید رابطه باشد.

وابستگی تابعی چندمقداری

تعریف اول- در رابطه R(X,Y,Z) با صفات ساده یا مرکب X، Y و Z می‌گوییم که Y با X وابستگی تابعی چندمقداری دارد و چنین نمایش می‌دهیم: $X \rightarrow\rightarrow Y$

اگر به یک مقدار X، مجموعه‌ای از مقادیر Y متناظر باشند.

وابستگی تابعی چندمقداری

تعریف دوم- در رابطه R، صفت Y با X وابستگی تابعی چندمقداری دارد اگر و فقط اگر مجموعه مقادیر Y متناظر با یک مقدار از جفت (X,Z) در R فقط به مقدار X بستگی داشته باشد و وابسته به مقدار Z نباشد.

رابطه 4NF

تعریف- رابطه‌ای 4NF است اگر BCNF باشد و در آن وابستگی تابعی چند مقداری مهم وجود نداشته باشد.

وابستگی تابعی چندمقداری ادغام‌شده

در رابطه R(X,Y,Z,U,V,...)، پدیده وابستگی تابعی چندمقداری ادغام شده وجود دارد اگر در پرتوی از این رابطه، پدیده وابستگی تابعی چندمقداری وجود داشته باشد. توجه داشته باشید که این پرتو باید روی سه صفت (ساده یا مرکب) باشد.

رابطه 5NF

تعریف- رابطه‌ای 5NF است اگر تمام وابستگیهای پیوندی آن ناشی از کلیدهای کاندید آن باشد. در نتیجه:

اگر بتوانیم یک وابستگی پیوندی در رابطه R پیدا کنیم که در همه پرتوهایش کلید کاندید R وجود نداشته باشد، رابطه 5NF نیست.

مزایای نرمالترسازی

- ارائه یک طراحی بهتر و واضح‌تر با کمترین اختلاط اطلاعات
- کاهش بعضی انواع افزونگی
- کاهش بعضی آنومالیها
- تسهیل اعمال بعضی قواعد جامعیت

انواع افزونگی در پایگاه داده‌ها:

۱. افزونگی طبیعی
۲. افزونگی تکنیکی ناشی از وجود کلید خارجی یا هرصفت مشترک دیگر در دو یا بیش از دو رابطه
۳. افزونگی ناشی از وابستگی بین صفات

معایب روش نرمالترسازی

۱- بروز فزونکاری در سیستم در عمل بازیابی

۲- ایجاد نوعی افزونگی از نوع افزونگی در سطح ادراکی

۳- زمانگیر بودن فرآیند نرمالترسازی به ویژه اگر محیط عملیاتی بزرگ و تعداد رابطه‌ها زیاد باشد.

۴- تصمیم‌گیری دشوار در اثر تعدد تجزیه‌ها

۵- ممکن وابستگی بین مجموعه صفات یک خردجهان، وابستگیهای به جز وابستگی تابعی و پیوندی باشد و بنابراین سبب مطرح شدن قواعد جامعیت دیگر و نیز ضوابط دیگری برای تجزیه رابطه‌ها شود و . . .

جلسه شانزدهم

طراحی فیزیکی پایگاه داده‌ها

آنچه در این جلسه می خوانید:

۱- تحلیل عوامل
۲- تحلیل پرسشها
۳- تحلیل تراکنشها
۴- تصمیمات در طراحی فیزیکی
۵- اهداف تنظیم پایگاه داده‌ها
۶- اطلاعات لازم در تنظیم پایگاه داده‌ها

آنچه در این جلسه می خوانید:

۷- مسائل مطرح در تنظیم
۸- نکات قابل توجه در بازبینی طراحی منطقی پایگاه داده‌ها پس از عملیاتی شدن آن
۹- کارهای لازم برای اصلاح شاخصهای ایجادشده در مرحله تنظیم پایگاه داده‌ها

هدفهای کلی: آشنایی با طراحی فیزیکی پایگاه داده‌ها

هدفهای رفتاری: دانشجو در پایان این جلسه می‌تواند:

- عواملی که در طراحی فیزیکی پایگاه داده‌ها موثر هستند را بیان کند.
- پرسشهایی را که در تحلیل پرسشها و تراکنش مطرح می‌شوند، بشناسد.
- اطلاعات لازم در تنظیم پایگاه داده‌ها را بیان کند.

طراحی فیزیکی پایگاه داده‌ها

طراحی فیزیکی به عواملی از جمله حجم و ماهیت داده‌ها، تعداد کاربران و از آنجا تعداد و ماهیت تراکنشها و نهایتا تعداد و ماهیت برنامه‌های کاربردی و فرکانس اجرای هر برنامه بستگی دارد.

تحلیل عوامل

تحلیل پرسشها

در این تحلیل باید به پرسشهای زیر پاسخ داده شود:

- چه رابطه‌هایی در پرسش دخیل هستند
- نوع رابطه‌ها چیست
- صفات موجود در مسند گزینش
- صفات موجود در مسند پیوند
- صفات مورد نظر در پرسش
- فرکانس اجرای هر پرسش
- اسلوب اجرای پرسش: اندرکنشی یا کامپایل‌شده

تحلیل عوامل

تحلیل تراکنشها

در این تحلیل باید به پرسشهای زیر پاسخ داده شود:

- رابطه‌های دخیل در تراکنش و نوع هر رابطه
- نوع هر عمل در تراکنش
- صفات موجود در مسند گزینش
- صفاتی که مقدارشان تغییر می‌کند

تصمیمات در طراحی فیزیکی

- روی چه فایلهایی شاخص ایجاد شود یا تابع درهم‌ساز اعمال شود.
- روی چه صفاتی از فایل، شاخص ایجاد شود.
- شاخص ساده باشد یا مرکب.
- شاخص خوشه‌ساز باشد یا نباشد.

اهداف تنظیم پایگاه داده‌ها

- تسریع اجرای برنامه‌های کاربردی
- کاهش زمان پاسخدهی پرسشها
- بهبود توان عملیاتی سیستم مدیریت پایگاه داده‌ها
- اجتناب از خرید سخت‌افزار بیشتر
- کاهش اندازه پیکربندی سخت‌افزاری
- رضایت بیشتر کاربران سیستم

اطلاعات لازم در تنظیم پایگاه داده‌ها

- اطلاعات در مورد خود DBMS
- اطلاعات اولیه مورد استفاده در طراحی منطقی و فیزیکی پایگاه
- آمار جمع‌آوری‌شده توسط DBMS و DBA در مورد رفتار سیستم مانند:
  - کاردینالیتی هر رابطه
  - تعداد مقادیر متمایز برای هرصفت و درصد هیچ‌مقدار برای هر صفت
  - فرکانس عرضه هر پرسش به سیستم و فرکانس اجرای آن
  - فرکانس اجرای هر تراکنش
  - اندازه صفحات
  - زمان اجرای پرسشها و تراکنشها
  - . . .

مسائل مطرح در تنظیم

- چگونگی اجتناب از بروز تنگنا در قفل‌گذاری روی داده‌ها به منظور افزایش همروندی اجرای تراکنشها
- چگونگی کاهش سربار ناشی از ایجاد فایلهای ثبت تراکنشها
- چگونگی بهینه‌سازی اندازه بافرها
- چگونگی بهینه‌سازی زمانبندی اجرای تراکنشها
- چگونگی بهینه‌سازی تخصیص منابع

در بازبینی طراحی منطقی پایگاه داده‌ها پس از عملیاتی شدن آن، باید به نکات زیر توجه شود:

- آیا رابطه‌هایی باید تجزیه شوند؟
- آیا سطح نرمالیتی رابطه‌هایی باید کاهش یابد؟
- آیا مجموعه دیگری از رابطه‌ها، با همان سطح نرمالیتی، باید طراحی شوند؟
- آیا صفاتی با پذیرش میزانی از افزونگی، باید در رابطه‌هایی تکرار شوند؟
- آیا در طراحی دیدهای خارجی باید تغییراتی داده شود؟

کارهای لازم برای اصلاح شاخصهای ایجادشده در مرحله تنظیم پایگاه داده‌ها

- ایجاد شاخصهای جدید
- حذف برخی از شاخصها
- کاهش تعداد شاخصها روی صفاتی که مقادیرشان مرتب تغییر می‌کنند.
- اصلاح شاخصهایی که در کنترل همروندی تراکنشها تاثیر دارند
- بازسازی شاخصها می‌تواند در کارایی سیستم تاثیر مثبت داشته باشد
- در بعضی از RDBMSها بهینه‌ساز، اگر در پرسش، عبارت محاسباتی، مقایسه یا هیچ مقدار و محاسبه تفاوت مقادیر وجود داشته باشد، از شاخص استفاده نمی‌کند.