

# پردازش پرس و جوهای توزیع شده:

برای تبدیل پرس و جوهای سطح بالا به دستورات سطح پایین از واحدی به نام پردازشگر پرس و جو (query processor) استفاده می شود.

مراحل کار:

- تبدیل پرس و جو به دنباله ای از عملگرهای رابطه ای
- تبدیل داده های مورد نیاز به صورت محلی
- اضافه کردن عملگرهای ارتباطی
- بهینه سازی برای حداقل کردن هزینه ها

# بهینه سازی پرس و جو:

یک پرس و جو را می توان به روشهای مختلف پردازش و اجرا کرد. برای یک پرس و جو گاهی چندین عبارت جبر رابطه ای وجود دارد، که بعضی بهتر عمل می کنند.

یافتن بهترین روش پردازش پرس و جو، بهینه سازی پرس و جو نام دارد که این موضوع در سیستم های نامتمرکز اهمیت بیشتری دارد.

# انواع بهینه سازی پرس و جوها:

- (1) جستجوی جامع (exhaustive search): در این روش هزینه راه حل ها را ملاک قرار داده و پرس و جویی با کمترین هزینه را انتخاب می کنیم. زمانیکه فضای جستجو بزرگ باشد، هزینه این روش بسیار بالا خواهد بود.
- (2) روشهای حسی-ذهنی (heuristic): با بکارگیری قواعدی خاص، سعی در محدود کردن فضای حالات از جوابها و رسیدن به یک راه حل تقریبا بهینه داریم.

# زمان انجام بهینه سازی:

**1 ایستا (static):** بهینه سازی به طور ایستا و قبل از اجرای پرس و جو انجام می گیرد. بهینه سازی ایستا برای روش جستجوی جامع مناسب است.

**2 پویا (dynamic):** به طور پویا و در حین اجرای پرس و جو، در هر لحظه بهترین عملگری که باید اجرا شود را مشخص می کنیم.

مزیت: اطلاعات لازم مربوط به نتایج میانی در اختیار پردازشگر پرس و جو قرار می گیرد.

معایب: به دلیل پرهزینه بودن، تنها برای پرس و جوهای تک کاره مناسب است.

**3 ترکیبی (hybrid):** هدف، بهره گیری از مزایای روش ایستا ضمن پرهیز از تخمین نادقیق می باشد. این روش ایستا عمل می کند ولی در حین اجرا بهینه سازی پویا نیز انجام می شود.

# بهینه سازی پرس و جوها:

مهم ترین اطلاعات آماری مورد نیاز در بهینه سازی پرس و جوها به شرح زیر می باشد:

- اندازه رابطه
- اندازه تاپلها
- بخشی از تاپل که در دستور پیوند شرکت دارد
- تعداد عناصر دامنه هر صفت
- تعداد واقعی مقادیر متمایز آن دامنه

# مدیریت تراکنش های توزیع شده ساده:

مدیریت تراکنش در محیط توزیع شده شامل دو بخش است:

- کنترل همروندی

- قابلیت اطمینان

مکانیزم های کنترل همروندی در محیط توزیع شده:

- کنترل همروندی توزیع شده تک نسخه

- کنترل همروندی توزیع شده چند نسخه

# کنترل همروندی توزیع شده تک نسخه:

اگر زمانبندی های محلی پی در پی پذیر هستند، اجتماع آنها که زمانبندی سراسری است پی در پی پذیر است در صورتی که ترتیب تراکنش ها همه جا یکسان باشد.

روشهای تولید مهر زمانی:

1- متمرکز

2- توزیع شده

# مهر زمانی متمرکز و توزیع شده:

روش متمرکز: هر سایت از یک شمارنده منطقی یا از ساعت محلی خود استفاده می کند.

روش توزیع شده: هر سایت با استفاده از شمارنده منطقی و یا ساعت محلی خود، مهر زمانی را تولید می کند.

مهر زمانی سراسری = (شناسه سایت ، مهر زمانی محلی)



# قانون ترتیب مهر زمانی سراسری:

برای هر دو دستور  $P_{ij}$  و  $P_{kl}$  به ترتیب مربوط به تراکنش های  $T_i$  و  $T_k$  که با هم برخورد دارند،  $P_{ij}$  قبل از  $P_{kl}$  اجرا می شود اگر و تنها اگر مهر زمانی تراکنش مربوطه اش کوچکتر از مهر زمانی تراکنش دستور دیگر باشد، یعنی  $TS(T_i) < TS(T_k)$ .

تراکنش مربوطه با مهر زمانی جدید مجددا شروع به اجرا می کند.

# انواع پروتکل های مهر زمانی:

(1) پروتکل مهر زمانی پایه: در صورت بروز مشکل، به جای منتظر نگه داشتن تراکنش آنرا شروع مجدد می کند. با این پروتکل بن بست رخ نمی دهد اما کارائی پائین می آید.

(2) پروتکل مهر زمانی محافظه کار: با هدف کاهش احتمال شروع مجدد تراکنش ها ارائه شده است و دستورات را به تعویق می اندازد.

# کنترل همروندی نامتمرکز چند نسخه:

عمده ترین مشکل در بانکهای اطلاعاتی نامتمرکز و چند نسخه، پی در پی پذیری است.

در بانکهای اطلاعاتی نامتمرکز چند نسخه، تمام نسخه های یک داده باید مقدار یکسان داشته باشند. این خاصیت همخوانی متقابل (mutual consistency) نامیده می شود.