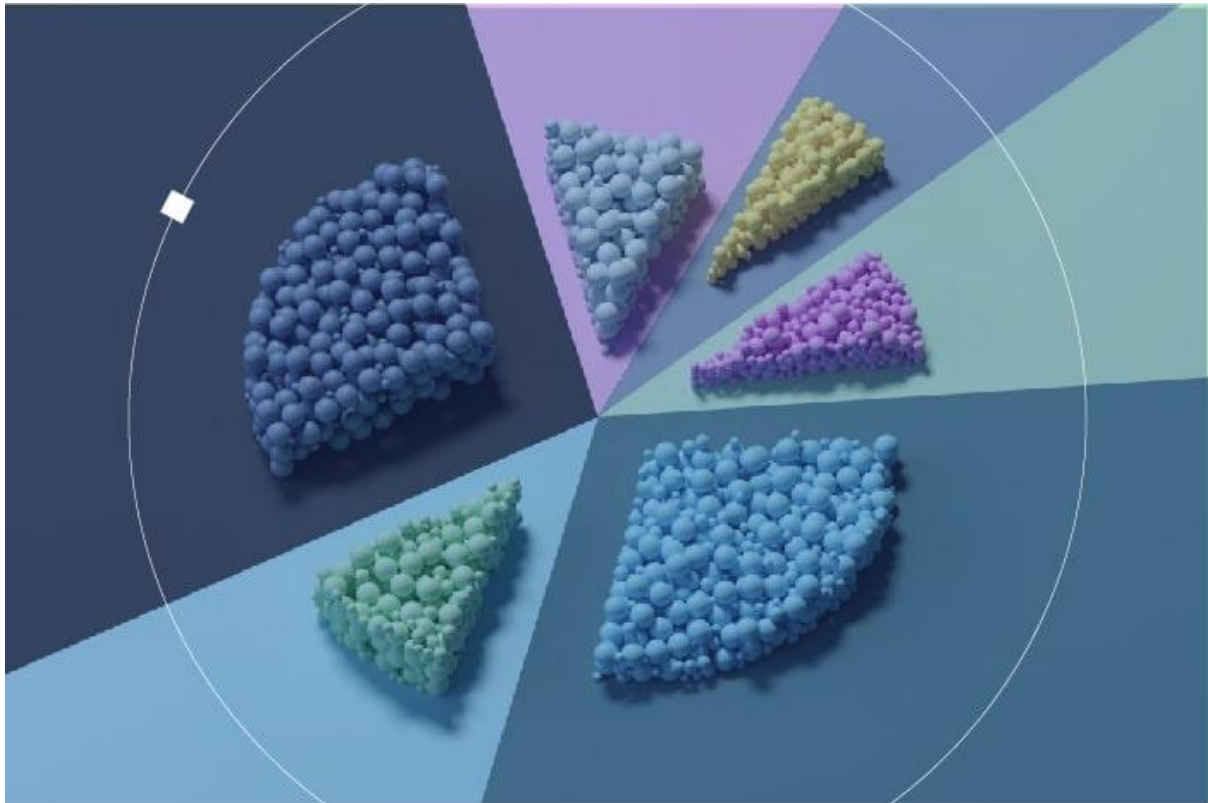


تفکیک داده ها و فواید آن برای پایگاه داده



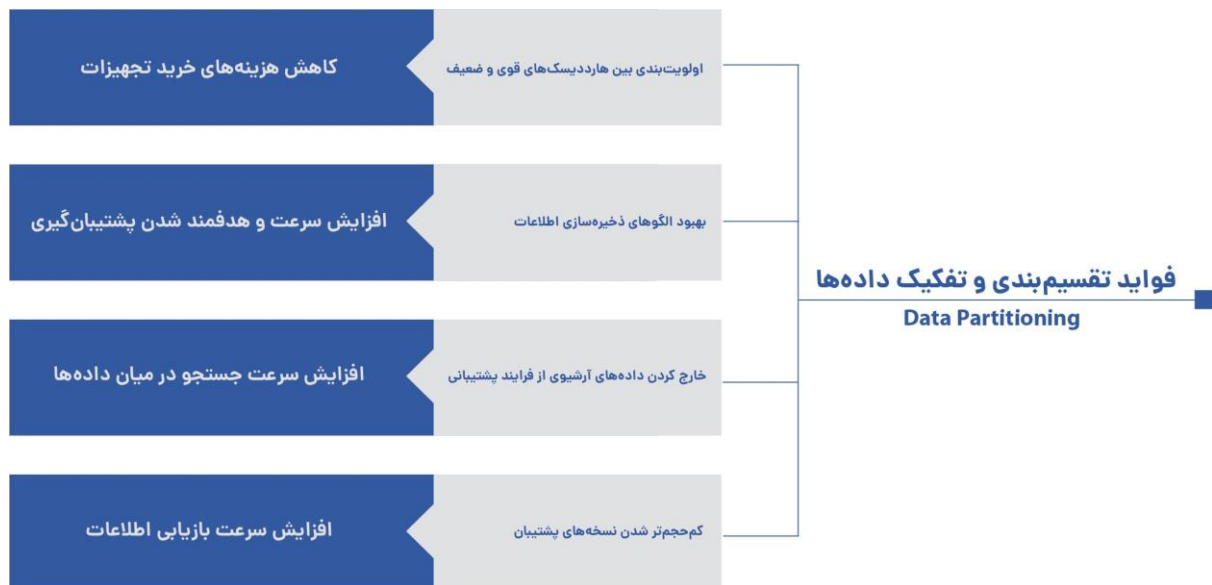
با کاهش تنوع و افزایش قیمت سخت افزارها و بالاخص تجهیزات ذخیره سازی اطلاعات، خرید تجهیزات جدید برای پایگاه داده به یک بحران تبدیل شده است. بحرانی که هم از لحاظ مالی و هم از لحاظ چگونگی تأمین کالا، باعث سردرگمی مجموعه های مختلف می شود.

خبر خوب این است که بخش قابل توجهی از نیاز به سخت افزارها می تواند با تغییر و بهبود الگوهای نرم افزاری مرتفع شود. پس در موقعیت های بحرانی به صرفه تر است که ابتدا ملاحظات نرم افزاری را در نظر بگیریم و بعد از آن اگر همچنان نیاز مدنظرمان ارضاء نشد، به سراغ تهیه سخت افزارها برویم.

در رابطه با پایگاه داده، که عملاً اصلی ترین محل ذخیره سازی داده ها به حساب می آید، مجموعه اقداماتی تحت عنوان تقسیم بندی و تفکیک داده ها با عنوان Partitioning در نظر گرفته می شود که از خدمات اصلی پایگاه داده به حساب می آید و به دلایل متعددی می تواند منجر به بهبود عملکرد کلی سیستم و به طور خاص بهبود عملکرد و سلامت پایگاه داده شود.

شاید بهترین روش برای تشخیص میزان اثرگذاری و اهمیت تفکیک داده ها و اطلاعات، این باشد که دو حالت «قبل» و «بعد» Partitioning را با هم مقایسه کنیم و آسیب های احتمالی قبل از آن را با مزایای بعد از آن در دو کفه ترازو قرار دهیم. با انجام این کار می توانیم تصمیم بگیریم که این عمل

تا چه میزان برای مجموعه‌ی ما ضرورت دارد و برای پیاده‌سازی آن نیز می‌توانیم به روش‌های مدنظرمان، اقدام کنیم.



قبل از Partitioning

هنگام جستجو در داده‌ها، برای هر جستجو، تمامی داده‌ها باید بررسی شوند و طبیعتاً زمان زیادی از کاربر می‌گیرد و فشار زیادی به سیستم وارد می‌شود.

برای پشتیبان‌گیری و بازگردانی اطلاعات، از آنجایی که فقط از یک مجموعه داده خوانش انجام می‌شود، سرعت بسیار کم است.

اگر از چند نوع منبع ذخیره Storage متفاوت برای ذخیره‌سازی اطلاعات و داده‌هایمان استفاده کنیم، عملاً کنترلی بر چگونگی تقسیم داده‌ها در هرکدام از فضاهای ذخیره‌سازی نداریم و این ریسک را متحمل می‌شویم که در مواقعی، عملیات پردازش و خوانش اطلاعات، با سرعت کم‌سرعت‌ترین منبع ذخیره‌مان انجام شود.

بعد از Partitioning

از آنجایی که تقسیم و تفکیک داده‌ها بر اساس منطق کسب‌وکارمان انجام شده است، در جستجوی اطلاعات، دیگر نیازی به بررسی تمام داده‌های موجود در پایگاه داده نیست و فقط بخشی که مرتبط با آن جستجوی خاص است بررسی می‌شود و طبعاً جستجو سریع‌تر به نتیجه می‌رسد.

هنگام پشتیبان‌گیری و بازگردانی اطلاعات، دیگر خوانش داده‌ها از یک مجموعه داده‌ی کلی انجام نمی‌شود و تقسیم‌بندی داده‌ها موجب می‌شود که به صورت همزمان، پردازش اطلاعات در چند قسمت کوچک‌تر انجام شود و سرعت به مراتب بالاتر از حالتی است که تقسیم و تفکیک داده‌ها انجام نشده باشد. ضمن اینکه برای پشتیبان‌گیری هم می‌توان به صورت آرشیو پشتیبان‌گیری کرد. برای مثال می‌توان یک بار اطلاعات قبل از 2 سال پیش را آرشیو کرد و در هر فرآیند پشتیبان‌گیری،

فقط داده‌های دو سال اخیر را ذخیره و بازیابی کرد. این امر باعث سریعتر شدن فرایند پشتیبان‌گیری و اشغال فضای کمتر برای پشتیبان‌گیری می‌شود. همچنین تفکیک داده‌ها بر اساس زمان، باعث می‌شود که بتوانیم داده‌هایی که دیگر تغییر نخواهند کرد و بیشتر به صورت آرشیو در سیستم قراردارند را از چرخه‌ی اصلی پشتیبان‌گیری خارج کنیم و با این کار حجم زیادی از داده‌ها از فرایند پشتیبان‌گیری Backup خارج می‌شوند و علاوه بر افزایش سرعت، فضای لازم برای این عملیات نیز کاهش می‌یابد

با اولویت‌بندی بین هارددیسک‌های قوی‌تر و ضعیف‌تر، می‌توانیم اطلاعات پرکاربرد خود را روی هارددیسک‌های با کیفیت‌تر نگهداری کنیم و آن دسته از اطلاعاتی را که کمتر با آنها سر و کار داریم و معمولاً در فواصل زمانی طولانی‌تری به آنها سر می‌زنیم روی منبع‌ذخیره‌های کم‌کیفیت‌تر نگهداری کنیم. با این کار علاوه بر افزایش سرعت سیستم، باعث کاهش هزینه‌های وارده به مجموعه نیز خواهیم شد.

به نظر می‌رسد در این روزها، به کارگیری خدمات پایگاه داده و به طور خاص Partitioning می‌تواند انتخاب خوبی برای بهبود و افزایش کارایی مجموعه‌ها باشد و راهکارهایی که بهینه‌تر و ساختارمندتر باشند می‌تواند پاسخگوی بهتری به اینگونه مسائل داشته باشد. ضمناً باید در نظر داشته باشیم که Partitioning تنها زمانی مفید واقع می‌شود که بتوانیم در مرور زمان نیز آن را مدیریت کنیم و اگر نگهداشت و مدیریت مناسب انجام نشود، بعد از مدتی بی‌فایده خواهد شد.