



عنوان مقاله: دیتابیس Time-Series چیست؟ ویژگی ها و نحوه کار پایگاه داده سری زمانی

نویسنده مقاله: تیم فنی نیکآموز

تاریخ انتشار: ۲۳ اردیبهشت ۱۴۰۳

منبع: <https://nikamooz.com/time-series-database/>

دیتابیس Time-Series نوعی پایگاه داده است که به طور خاص، برای ذخیره سازی و استخراج آسان داده‌هایی طراحی شده است که بر اساس Timestamp ثبت شده‌اند. این پایگاه داده، برای مدیریت کارآمد داده‌هایی بهینه‌سازی شده است که در طول زمان تغییر می‌کنند. در این مقاله، به بررسی دیتابیس سری زمانی، مزایا، معایب، ویژگی‌ها و مثال‌هایی از آن می‌پردازیم و شرح می‌دهیم که چرا این دیتابیس حائز اهمیت است.

داده های Time-Series

داده‌های Time-Series مجموعه‌ای از نقاط داده جمع‌آوری شده در بازه‌های زمانی هستند که به ما توانایی ردیابی تغییرات در طول زمان را می‌دهند. داده‌های سری زمانی می‌توانند تغییرات را در طول میلی‌ثانیه، روز یا حتی سال ردیابی کنند. در گذشته، دیدگاه ما از این داده‌ها ثابت‌تر بود. به عنوان مثال، بالا و پایین رفتن روزانه دما، ارزش باز و بسته شدن بازار سهام، بستری شدن روزانه یا تجمعی در بیمارستان به دلیل کووید-۱۹، نمونه‌هایی از آن هستند.

برای درک بهتر، به مثال زیر توجه کنید:

اگر من ۱۰ دلار برای شما بفرستم، یک بانک اطلاعاتی سنتی از حساب من پول برداشته و به حساب شما اعتبار می‌دهد. سپس، اگر ۱۰ دلار برای من بفرستید، همان فرآیند به صورت معکوس انجام می‌شود. در پایان این فرآیند، موجودی بانک ما یکسان به نظر می‌رسد؛ بنابراین، بانک ممکن است فکر کند «هیچ چیز در این ماه تغییر نکرده است.» اما، در دیتابیس Time-Series، بانک می‌بیند که دو شخص مدام ۱۰ دلار برای یکدیگر ارسال می‌کنند، احتمالاً یک رابطه عمیق‌تری در اینجا وجود دارد. با ردیابی این تفاوت، موجودی حساب آخر ماه ما معنای بیشتری پیدا می‌کند.

این مثال نشان می‌دهد که چگونه داده‌های Time-Series مدرن با آنچه در گذشته می‌شناختیم، متفاوت هستند. تجزیه و تحلیل داده‌های سری زمانی بسیار عمیق‌تر از نمودار دایره‌ای یا کار اکسل با ستون‌هایی از مجموع خلاصه شده است.



دیتابیس Time-Series چیست؟

دیتابیس Time-Series یا همان TSDB، یک سیستم کامپیوتری است که به منظور ذخیره سازی و بازیابی سوابق داده ای طراحی شده است. این سیستم، بخشی از دامنه گسترده تری به نام سری های زمانی است که مجموعه ای از نقاط داده را که با مُهرهای زمانی یا Timestamps مشخص شده اند، شامل می شود. تایم استامپ های زمانی به عنوان یک زمینه حیاتی برای هر نقطه داده و ارتباط آن با داده های دیگر عمل می کنند. داده های Time-Series اغلب یک جریان پیوسته از داده ها مانند اندازه گیری های حسگرها و قیمت های سهام در روز است. دیتابیس Time-Series به شما اجازه می دهد حجم زیادی از داده های Timestamp زمانی را در قالبی ذخیره کنید که امکان درج و بازیابی سریع را برای پشتیبانی از تجزیه و تحلیل پیچیده روی آن داده ها فراهم می کند.

مزایا و معایب دیتابیس Time-Series

در این بخش، به بررسی نقاط مثبت و منفی دیتابیس Time-Series می پردازیم.

مزایای پایگاه داده سری زمانی

مزیت‌های دیتابیس سری زمانی عبارتند از:

- ذخیره‌سازی بهینه برای داده‌های دارای نشانگر زمان
- عملکرد بالا برای جذب و پردازش کوئری
- قابلیت ارتقا برای مدیریت حجم داده و لود کوئری در حال رشد
- مدل داده انعطاف‌پذیر برای پشتیبانی از انواع مختلف داده
- قابلیت تحلیل براساس زمان برای تشخیص ترند، پیش‌بینی و تشخیص ناهنجاری
- نظارت و هشدار بلادرنگ براساس آستانه‌ها یا الگوهای از پیش تعیین‌شده
- نگهداری داده‌های تاریخی با مدیریت کارآمد فضای ذخیره‌سازی
- امکان یکپارچه‌سازی با ابزارها و فریمورک‌های تحلیلی
- مدیریت و تجزیه و تحلیل مؤثر جریان داده‌های اینترنت اشیا و حسگرها
- پشتیبانی از نگهداری پیش‌بینی‌کننده از طریق تجزیه و تحلیل داده‌های حسگر

معایب پایگاه داده سری زمانی

کاستی‌های دیتابیس سری زمانی عبارتند از:

- پیچیدگی پیاده‌سازی
- موارد استفاده محدود به داده‌های Time-Stamped
- مشکلات مربوط به سازگاری داده‌ها
- ابزار و پشتیبانی محدود اکوسیستم

ویژگی دیتابیس Time-Series

بارزترین ویژگی‌های دیتابیس Time-Series در ادامه فهرست شده‌اند:

- ذخیره‌سازی داده با Time-Stamp
- سازماندهی داده‌های ترتیبی
- نرخ جذب (Ingestion Rate) بالای داده
- ایندکس‌گذاری براساس زمان
- فشردگی و Downsampling
- Windowing و Aggregation ها
- سیاست‌های نگهداری داده
- انعطاف‌پذیری اسکیمای
- پشتیبانی از داده‌های جاری
- قابلیت‌های تحلیلی پیشرفته



دیتابیس Time-Series چگونه کار می کند؟

دیتابیس‌های Time-Series با گرفتن مجموعه‌ای از مقادیر ثابت همراه با مجموعه‌ای از مقادیر پویا کار می‌کنند. به‌عنوان مثال، در یک چاه نفت که در آن معیارهای بسیاری از دکل گرفته شده است، یک مجموعه از نقاط داده ممکن است دارای برچسب "Oil Pressure Rig #1" بوده و مقادیر دینامیکی مرتبط اندازه‌گیری فشار همراه با مهر زمانی باشد. این داده‌های Time-Series برای ردیابی روند فشار نفت مفید هستند که وقتی همراه با سایر معیارها، تجزیه و تحلیل شوند، می‌توانند منجر به پیش‌بینی درمورد نیازهای تعمیر و نگهداری و همچنین تصمیم‌گیری درمورد رهاشدن چاه شوند. این رکوردها در یک رسانه ذخیره‌سازی با فرمتی نوشته می‌شوند که امکان خواندن و نوشتن سریع براساس زمان را فراهم کند.

تمام رکوردها دارای مهر زمانی هستند؛ بنابراین، ترتیب نقاط داده، مشخصه اصلی داده‌ها است. استفاده از موتور پردازش جریانی می‌تواند به‌طور قابل توجهی به بهبود عملکرد و سرعت در تحلیل داده‌های Time-Series کمک کند. این موتورها به‌ویژه برای پردازش داده‌هایی طراحی شده‌اند که به‌صورت جریانی وارد می‌شوند و امکان تحویل داده‌ها به‌صورت سفارش داده‌شده به آن‌ها را دارند (به نحوی که اطلاعات به‌صورت جریان داده‌شده به آن‌ها می‌رسد). با توجه به اینکه یکی از اهداف اصلی در دیتابیس Time-Series، سرعت در ذخیره و بازیابی داده‌هاست؛ استفاده از این موتورهای پردازش جریانی به‌عنوان یک روش ایده‌آل برای دستیابی به این هدف مطرح می‌شود.

مثال هایی از موارد استفاده دیتابیس Time-Series

موارد زیر، مثال هایی قابل درک از کاربرد دیتابیس Time-Series هستند:

چاه نفت

در محیط های اینترنت اشیا (IoT)، دستگاه های مختلف به طور مداوم، داده های مربوط به معیارها و ویژگی های مختلف را ثبت می کنند. این داده ها از طریق اتصال به اینترنت به دیتابیس های Time-Series ارسال می شوند تا به منظور تجزیه و تحلیل بیشتر در دسترس قرار گیرند. مثال چاه نفت، یک مورد رایج استفاده از اینترنت اشیا است که در آن، تجزیه و تحلیل معیارهای متعدد از یک چاه نفت می تواند به پیش بینی تعمیر و نگهداری کمک کند. در این نمونه، تجزیه و تحلیل، به پیش بینی زمان خرابی تجهیزات به دلیل عوامل مختلف، منجر می شود. از دیتابیس های Time-Series برای جمع آوری حجم وسیعی از داده ها استفاده می شود، سپس برنامه ها بر روی پایگاه داده اجرا می گردند تا تجزیه و تحلیل را ارائه دهند.

تجزیه و تحلیل معیارهای سیستم کامپیوتری

یکی دیگر از موارد استفاده از دیتابیس های Time-Series، تجزیه و تحلیل معیارهای سیستم کامپیوتری است. در این شرایط، وضعیت و عملکرد مختلف سیستم های کامپیوتری و نرم افزارها در دیتابیس Time-Series ذخیره می شوند و به کمک آن، متخصصان IT می توانند بر وضعیت سیستم های مختلف نظارت کنند. ردیابی معیارهایی مانند میزان استفاده از حافظه یا تعداد فرآیند، برای مشاهده اینکه آیا باید منابع کامپیوتری جدیدی مستقر شوند یا برنامه ها نیاز به تخصیص مجدد دارند، کاربردی است.

چرا به دیتابیس Time-Series نیاز داریم؟

حال این سؤال پیش می آید که چرا نمی توان از پایگاه داده سنتی و رایج استفاده کرد؟ هرچند این کار امکان پذیر است، اما بنا به دلایل زیر، دیتابیس Time-Series مورد نیاز است:

- **ذخیره سازی کارآمد:** پایگاه داده سری زمانی برای ذخیره سازی و کوئری نویسی کارآمد حجم زیادی از دیتای حاوی Timestamp بهینه سازی شده است. این نوع از دیتابیس ها، از تکنیک های تخصصی ذخیره سازی و ایندکس گذاری متناسب با داده های سری زمانی استفاده می کنند. بدین شیوه، امکان دریافت و استخراج سریع داده ها فراهم می شود.
- **عملکرد مناسب:** TSDB ها برای رسیدگی به نرخ جذب بالا و کوئری های پیچیده روی داده های سری زمانی (با کمترین تأخیر) طراحی شده اند. در دیتابیس Time-Series hc تکنیک های فشرده سازی، تجمع و پردازش موازی استفاده شده تا از عملکرد سریع کوئری ها حتی در دیتاست های بزرگ، اطمینان حاصل شود.
- **قابلیت ارتقا:** دیتابیس سری زمانی می توانند به صورت افقی Scale شود تا حجم داده های در حال رشد و بار کوئری های در حال افزایش را در خود جای دهند.
- **نظارت و هشدار بلادرنگ:** پایگاه های داده سری زمانی با جذب و آنالیز مستمر داده های جاری، امکان نظارت و هشدار بلادرنگ را فراهم می کنند.

- **تحلیل براساس زمان:** پایگاه داده سری زمانی به کسب و کارها اجازه می‌دهد تا تحلیل‌های مبتنی بر زمان، مانند تحلیل ترند، پیش‌بینی، تشخیص ناهنجاری و تشخیص الگو را انجام دهند.
- **سیاست‌های نگهداری داده:** TSDB ها از سیاست‌های انعطاف‌پذیر نگهداری داده پشتیبانی می‌کنند و به سازمان‌ها این امکان را می‌دهند که داده‌های تاریخی را تا زمان موردنیاز حفظ کرده و در عین حال، منابع ذخیره‌سازی را به‌طور مؤثر مدیریت کنند.
- **داده‌های اینترنت اشیا و حسگر:** پایگاه داده سری زمانی برای جذب، ذخیره‌سازی و تجزیه و تحلیل جریان داده‌های اینترنت اشیا به‌صورت بلادرنگ مناسب هستند.
- **اپلیکیشن‌های مالی و معاملاتی:** داده‌های سری زمانی در بازارهای مالی، سیستم‌های معاملاتی و برنامه‌های بانکی، نقش اساسی دارند و در نظارت بر ترندهای بازار، تجزیه و تحلیل الگوهای معاملاتی و تحلیل ریسک، حائز اهمیت هستند.
- **DevOps و عملیات IT:** این نوع از دیتابیس، به‌طور گسترده در **DevOps** و عملیات IT به منظور نظارت بر عملکرد زیرساخت، معیارهای اپلیکیشن و دیتای لاگ استفاده می‌شود.



موارد استفاده دیتابیس سری زمانی

توسعه دهندگان از دیتابیس‌های Time-Series برای موارد زیر استفاده می‌کنند:

- سیستم‌های نرم‌افزاری مانیتورینگ: ماشین‌های مجازی، کانتینرها، خدمات و برنامه‌ها
- نظارت بر سیستم‌های فیزیکی: تجهیزات، ماشین‌آلات، دستگاه‌های متصل، محیط زیست و خانه‌ها
- برنامه‌های ردیابی دارایی: وسایل نقلیه، کامیون‌ها، ظروف فیزیکی و پالت‌ها
- سیستم‌های معاملات مالی: اوراق بهادار کلاسیک و ارزهای دیجیتال
- برنامه‌های رویداد: ردیابی داده‌های تعامل کاربر / مشتری
- ابزار هوش تجاری: ردیابی معیارهای کلیدی و سلامت کلی کسب و کار با استفاده از [هوش تجاری](#)

انتخاب دیتابیس سری زمانی: Timescale

انتخاب دیتابیس سری زمانی مانند Timescale، گزینه مناسبی برای ذخیره‌سازی و مدیریت داده‌های Time-Series است. دلایل انتخاب Timescale عبارتند از:

- تطابق با مدل داده و الگوی نوشتن / خواندن: Timescale به خوبی با مدل داده‌ها و الگوهای نوشتن / خواندن سازگاری دارد که توسط توسعه دهندگان به کار می‌روند.
- پشتیبانی داخلی از SQL: این دیتابیس، قابلیت‌های سری زمانی را به PostgreSQL، که یک پایگاه داده رابطه‌ای قدرتمند است، اضافه می‌کند. این یعنی، توسعه دهندگان می‌توانند از سینتکس SQL برای کوئری نویسی و مدیریت داده‌های سری زمانی استفاده کنند و از مهارت‌ها و ابزارهای موجود خود بهره‌مند شوند.
- انطباق با ابزارهای هوش تجاری Self-Service: امکان جستجو و تحلیل داده‌های سری زمانی به همراه داده‌های فراداده در Timescale فراهم است. این ویژگی با ابزارهای هوش تجاری مانند Tableau و [Power BI](#) سازگاری دارد و به کاربران امکان دسترسی سریع به اطلاعات معنادار را می‌دهد.

جایگزین های دیتابیس Time-Series

در ادامه، به بررسی گزینه‌های جایگزین برای Time-Series Database می‌پردازیم:

- سیستم مدیریت پایگاه داده رابطه‌ای (RDBMS): این [سیستم های مدیریت پایگاه داده رابطه ای](#)، که اغلب سیستم‌های پایگاه داده همه‌منظوره در نظر گرفته می‌شوند، می‌توانند برای ذخیره‌سازی و بازیابی داده‌های Time-Series استفاده شوند. با انعطاف‌پذیری سیستم‌های مدیریت پایگاه داده رابطه‌ای، می‌توان داده‌های مشابه دیتابیس Time-Series را ذخیره کرد؛ با یک تفاوت کلیدی، که نحوه نوشتن داده‌ها در رسانه ذخیره‌سازی است. از آنجایی که سیستم‌های مدیریت پایگاه داده رابطه‌ای اهداف طراحی متفاوتی نسبت به دیتابیس Time-Series دارند، برای داده‌های Time-Series بهینه نیستند و برای درج و بازیابی داده‌های Time-Series کندتر عمل می‌کنند.

• **پایگاه داده NoSQL:** نوع دیگری از **انواع پایگاه داده**، یعنی NoSQL، نیز غالباً برای ذخیره داده‌های Time-Series استفاده می‌شود. از آنجایی که **انواع پایگاه داده NoSQL** از نظر قالب داده برای هر رکورد انعطاف‌پذیرتر هستند، برای گرفتن داده‌های Time-Series از تعدادی منابع مجزا مفیدند. پیاده‌سازی یک پایگاه داده NoSQL برای داده‌های Time-Series، اغلب جایگزین مناسبی برای دیتابیس‌های Time-Series است و در عین حال، می‌تواند قابلیت‌هایی را ارائه دهد که فراتر از داده‌های Time-Series اعمال می‌شوند.

معیارهای انتخاب دیتابیس Time-Series در پروژه های مختلف

برای انتخاب دیتابیس سری زمانی مناسب، معیارهای زیر را در نظر داشته باشید:

- ارزیابی نرخ جذب و حجم داده موردانتظار
- بررسی پیچیدگی کوئری‌ها و کارایی
- بررسی نیازمندی‌های مقیاس‌پذیری
- ارزیابی انعطاف‌پذیری مدل داده
- امکان یکپارچه‌سازی با سیستم‌های موجود
- تعیین نیازمندی‌های مربوط به آنالیز و پردازش بلادرنگ
- بررسی قابلیت‌های مدیریت و نگهداری داده‌ها
- ارزیابی دسترسی به پشتیبانی کامیونیتی، منابع و مستندات دیتابیس سری زمانی
- بررسی نیازمندی‌های انطباق و ویژگی‌های امنیتی
- توجه به TCO مرتبط استقرار و عملیات Time-Series Database

سخن پایانی

دیتابیس Time-Series، پایگاه داده‌هایی هستند که برای داده‌های دارای مُهر زمانی یا Timestamp بهینه‌سازی شده است. داده‌های Time-Series نیز رویدادهایی هستند که در طول زمان ردیابی، نظارت، نمونه‌برداری و جمع‌آوری می‌شوند. در این مقاله، به بررسی آنچه درخصوص دیتابیس سری زمانی نیاز دارید، پرداختیم.