

## 数据库指标集的设计思路

原创 白鳢 白鳢的洞穴 2024年12月27日 08:31 广东



微信扫一扫  
关注该公众号

昨天和一个客户交流运维工具，谈到指标的采集，有个小伙子说他们的指标主要用于对数据库死活的监控，目前能够覆盖80%的场景。不过如果遇到了没有覆盖到的场景的时候，那才是要命的，很可能就会出现长时间无法恢复业务的情况。今早起来还在思考昨天的交流中的几个问题，于是把书中这个章节修改了一下。下面是我书中关于指标集设计思路的小节的内容。

数据库作为一种运维对象是数字化存在的，这些数据包括配置数据、运行数据、变更数据以及分析诊断数据等等，贯穿了数据库的全生命周期。从目前的阶段来看，数据库指标集是数据库数字化运维与优化的基础和核心。对于数据库运维部门而言主要关注点是运行数据，通过指标的形式进行获取，在此基础上可以产生诊断、告警之类的数据，如果配合AI技术，甚至可以进行一些预测。

数据库监控指标的发展历史悠久，实际上这是随着人们对指标的使用要求不断提升而不断改变的。最初数据库监控只是网管监控的一部分，只需要知道数据库或者还是死掉了就可以了。因此数据库监控做得都很随意，最简单只是探测死活。后来在运维实践中发现归档日志满、表空间满这些容量问题可能引发了超过60%的数据库故障。采集一下相关的指标，并设置一个告警阈值可以有效解决这些问题。后来又发现活跃会话数太多了，数据库容易出现性能问题，甚至出现严重运维故障，于是又慢慢增加了活跃会话数的采集指标。随着运维要求的增加，又发现光是采集活跃会话数还不够，还需要知道这些活跃会话数都在等什么等待事件，于是又增加了这方面的采集。

实际上这种做法是很正确的，比一下子采集了一堆数据回来不知道要用在哪里好多了。数据库的指标可以用于监控数据库的死活、性能；用于故障的提前预警；用于历史故障分析和巡检等工作。80%以上的数据库问题可能只需要十几个指标就可以解决了，不过一些深层次问题的分析与发现，可能需要更为丰富的指标才能够完成。随着技术的发展，数据库监控的需求也在不断发展，在数字化时代里，数据库监控要能够提前预警更多的故障，因此对数据库指标的要求也在不断提升。

以前的数据库监控都是给人看的，生成一张折线图就可以了。今后数据库的监控数据大部分是给AI智能体看的，AI智能体会对数据进行自动化的分析，并随时发出告警，提出优化建议。数据库指标集的构建是数据库运维自动化和智能化的基础，因此必须加以重视。在做数据库指标设计的时候，往往会受限于早些年网管系统建设思路，过度借鉴SNMP的MIB库。这些基于传统思路设计的指标集在静态配置信息上十分值得借鉴，但是动态指标往往过于简单，无法适应当前数据库运维的需要。在做数据库指标集设计的时候，要从运维工作的要点出发同时结合当前数字化运维的需求，从对数据库系统进行全面的数字化描述的角度入手，才能让数据库的指标集在实际运维工作中实用、好用。

数据库内部会存在多个组件，因此运行时需要采集的指标数量还是非常多的，少的几十个，多的能上千，为了便于管理，需要对这些指标进行分类。比如对于某个国产数据库，一般在运维时会关注数据库以下方面的运行情况：

- 配置：影响对象运行的主要参数、资源设置，有些配置是动态的，不能只做一次性采集
- 资源：运维对象涉及的各种内外资源的使用情况
- 负载：运维对象及其所有组件的活动量统计
- 性能：主要是指运维对象执行各项任务的响应时间



- 效率：主要是指运维对象在进行内部处理时最优路径的选择比率
- 并发：内部资源或数据结构的争用情况
- 集群：集群中的成员对象之间发生的活动统计
- 总体：包括总体运行状态、复制状态、备份情况、日志错误统计、安全类统计之类的指标

对数据库的指标进行梳理，可以将数据库的指标按照管理类、配置类和技术类三大类进行归类。管理类指标主要体现数据库与上下游之间的关联关系，数据库不能脱离存储、服务器等独立存在，因此必须关心其关联对象的指标，另外在运维方面有什么关注点，也应该进行指标化。比如容量、负载、资源等，要根据管理特性设计相关的指标进行管理。数据库的配置文件、配置信息等也需要指标化管理，指标化管理有助于对数据库对象进行全面的数字化描述和展示，同时一些负载和并发相关的指标的基线与波动特性也与配置有关。技术类的指标包含来自于数据库本身的指标和通过对日志分析产生的指标，日志事件指标化是对日志进行强管理的有效措施。

数据库指标集的建设实际上也是一种运维知识的积累，理解每个指标的含义，知道它们可以在哪些地方发挥作用，哪种故障可能会与哪些指标有关，哪些指标会影响到数据库运行的那个方面，这才是建设数据库指标集的最终目标。仅仅看指标的数量，而不关注指标的质量，以及指标的使用知识都是不可取的。在指标集的制定过程中，一定需要资深的DBA和专家参与才能做得更好。

对于一个数字化运维水平较高的团队而言，指标集的变化是十分快速的。有可能今天某个数据库的指标集才280个，第二天就变成300个了。因为昨天发生了一个新的故障，经过分析认为要想预警和根因定位这个故障，还需要再额外采集20个指标。

# 智能化运维 280 # 数字化转型 89 # 系统优化 333 # 数据库原理 311

智能化运维 · 目录

< 上一篇  
数据库选型是十分个性化的选择

下一篇 >  
国产数据库问题与性能诊断中的宏观分析与微观分析