

如果嵌入式系统的存储容量无限大将是怎样的境况？

Patrick Mannion

数据采集系统一直面临着存储容量和处理能力的基础限制，但数据库结构的新思维可以让开发人员摆脱传统的桎梏。具体来说，在边缘的高保真度数据存储与现代数据访问 API 相结合，可让架构师开发更灵活、更方便使用的系统。

受局限的思维问题

常见的数据采集和分析难题包括确定：

- 感知内容
- 感知频率
- 数据存储位置
- 存储时长
- 按需分析的访问方式

出现这些问题的部分原因在于设计人员的思维受到经典设计考虑因素的限制，例如可用于存储数据的内存有限、网络带宽有限、异构网络之间的连接不可靠、功耗问题和处理能力等。

这种受局限的思维比比皆是，它有三种负面影响。首先，一旦做出决定，系统刚性就开始形成。当新思维方式、技术和分析要求出现时，现在有意义的物联网节点部署可能将在三到五年内过时。

其次，设计人员致力于以最低的频率对特定数据输入进行采样和记录，以节省有限的内存资源。这样一来，他们将面临错过重要事件的风险（图 1a）。采样间隔越短，检测到重要事件的几率将越高（图 1b）。

最后，从分析和流程改善的角度来看，可能最令人沮丧和造成损害的就是高分辨率、高保真度的历史数据缺失。虽然记录了低分辨率数据，并且可以用于将来的分析，但未记录的大部分数据都将永久丢失。

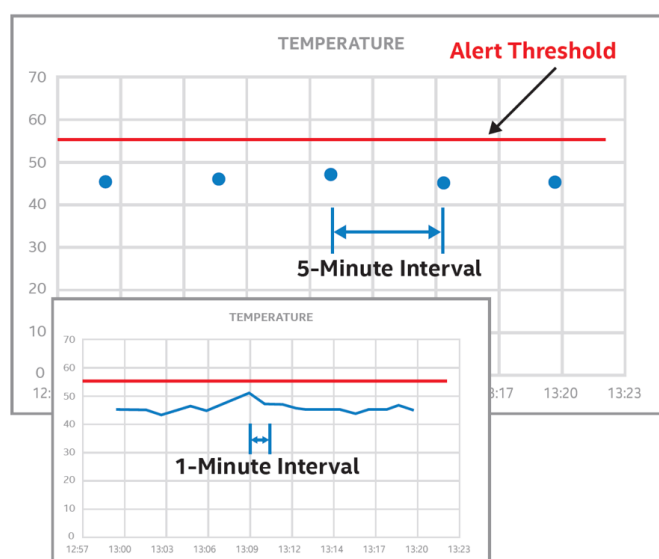


图 1a. 低频采样可能无法记录确定故障原因所需的数据。（图片来源：Exara）。

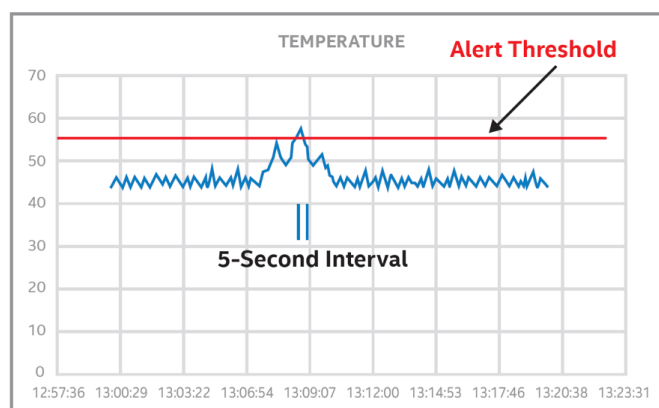


图 1b. 间隔越短，检测到事件的可能性越高。（图片来源：Exara）



这些未记录的数据可能对发现导致资产故障的异常起重要作用。如果使用正确的算法，这些数据也可能对提高系统可靠性、效率和业务本身极其有用。数据丢失实际上就是机会流失。

存储容量无限大是否可能？

如果设计人员不受约束，并且可以采用 Hz 频率而不是 dHz 频率收集所有机器的数据，将会怎样？如果他们不再需要长时间投入到数据收集和分析中，而是可以随时调整，又会怎样？有时候，这听起来好得有些不真实，但是这可能就是真的。

通过将诸如英特尔凌动® 处理器家族等现代 64 位 CPU 与一些有关数据库的创新思维结合起来，Exara 正在采取措施消除传统数据采集和分析系统设计的传统约束框架。它通过传感器和操作技术之间的抽象层，以及应用或企业层，在网络边缘进行适当的操作。

Exara 的创始人和首席技术官 Eric Kraemer 所言极是，他说：“现实世界非常混乱，共识会发生变化、设备会不断更新换代，节点也是如此。”每天都有很多不可预测性，但是刚性这一因素却随之产生。

Kraemer 表示“在我们作出假设之后，我们就创造了刚性。”他和 Exara 团队所开发的“机器数据服务”旨在消除网络边缘的刚性，并利用软件定义的物联网生态系统帮助推动行业整体向完全数字化迈进。

随时随地任何数据

Exara 的解决方案深入到传感器输入聚合点（如 PLC 或 RTU）所在的边缘网络（图 2）。在此，网络被视为南向或者北向。南向网络是运营技术 (OT) 领域，包括通过广泛的机器级协议感知、监控和控制的资产（Exara 支持最多 300 个协议）。北向网络包括 API 和应用层，它们能够以应用就绪、“及时”的方式提取和呈现所需数据。

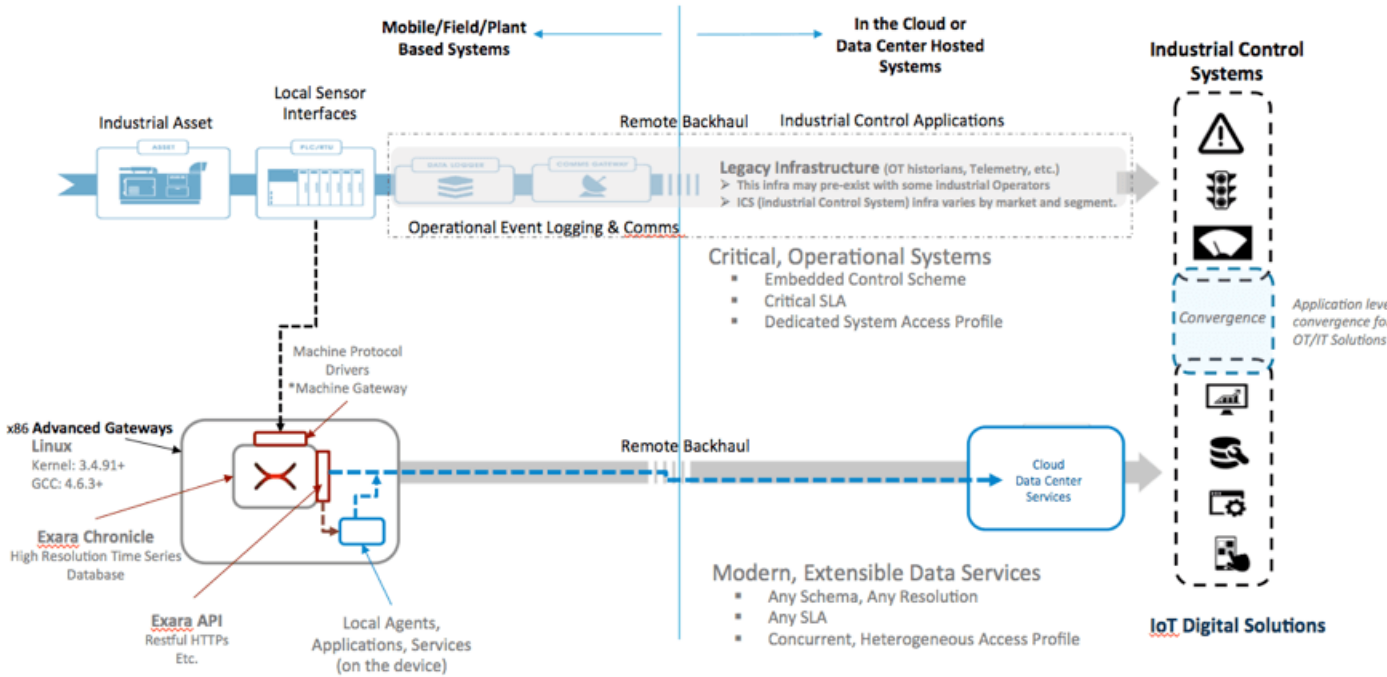


图 2. Exara 的 Chronicle 涉及本地传感器接口，以高保真度存储所有数据，并允许通过 API 访问。（图片来源：Exara）

在南向和北向的领域之间，Exara 开发了一个时间序列数据库架构，该架构可以获取并压缩来自资产的全部机器数据，然后将其以类似于缓存的方式存储在边缘中。使用标准或自定义 API，用户（例如企业 IT 部门或企业办公室）可以根据需求来提取数据。

Kraemer 说：“我们的底线可以总结为直接连接并读取本地机器协议。”通过添加更多内存，能够扩展可捕获的数据量，但是对于初学者，它可以采用 1 Hz 的平均采样分辨率采样一千个标签（传感器输入），并在 256 Gb 的固态硬盘中捕获“数百亿个事件日志”Kraemer 继续说道。“与传统的数据采集设备相比，时间序列事件记录的容量要大出几个数量级。”

更重要的是，Exara 的方法适用于并发读取访问，并通过标准 API 将该数据呈现给现代应用开发生态系统。他表示“但是您的传统行业数据采集设备并不是如此。”Kraemer 将 Exara 的 Chronicle 软件与成本为 4,000 美元的典型高端系统进行比较，他表示该设备可以记录大约 200 万个事件。

Exara 解决方案的关键推动因素在于诸如英特尔凌动处理器家族等低成本 64 位处理器的诞生。这些 CPU 能够解决较大内存空间的问题。随着内存成本下降和性能提高，诸如 SATA Gen 2 等较高带宽接口对于存储而言至关重要。Kraemer 表示“请记住，我们谈论的是处理大量数据。”

Chronicle 本身本质上是一个用 C 语言编写的高性能、时间序列的分析数据库，因此它的运行方式与内核非常接近。迄今为止，开发人员依赖于作为关系数据库而被重新利用的嵌入式事务数据库。但分析和事务数据库在性能特征、访问模式以及所获得的请求类型方面有所不同。Kraemer 说：“如果我们想要运行从行业边缘读取的应用，那么我们需要建立的核心内容是数据库基础设施。”

可扩展和安全

通过将存储和分析从物理层的变化率抽象出来，Exara 的方法可以随着网络的发展以及存储的数据量和时长而变化。典型的数据存储时间范围为三年，此架构采用“滚动窗口”缓冲方式，因此它始终可以将数据存储满三年。这通常足以分析以前未开发的数据，以确定任何异常行为的原因，或对该数据应用新的分析算法。但如果客户需要更长的存储时间，只需添加更多的内存即可。

从安全的角度来看，Chronicle 具有“只读”性质，因此它无法访问已安装的系统，数据本身也可以通过安全的现代 API 访问。

迄今为止，Exara 已经向一位客户展示了如何通过优化其使用来降低功耗，从而每年可以在一台卧式泵上节省 60,000 美元成本。