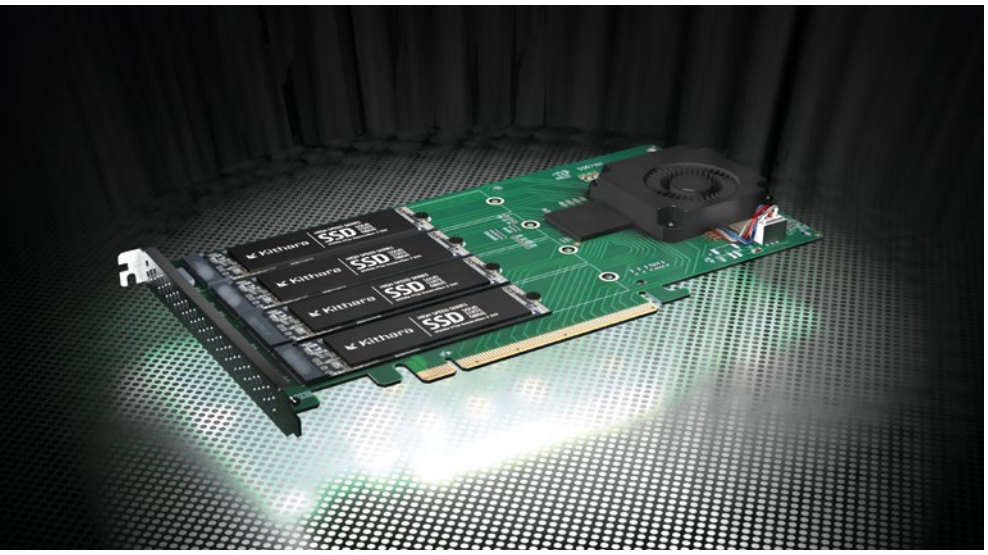
Deterministic Real-Time Measurement Data Acquisition with MDF and PCAPng

Data-driven development approaches, AI-based analyses, and digital twins today rely on high-quality, gapless measurement data. What is gaining traction in many industrial sectors has long been standard practice in measurement data acquisition—albeit with continuously growing data volumes and increasing real-time requirements.

Today, little can be achieved without large, high-quality data sets. At the same time, exponentially increasing data volumes are pushing even modern measurement applications to their limits when it comes to storing information both completely and efficiently. The solution is twofold: software-based real-time systems provide the necessary performance characteristics to acquire data deterministically and without loss, while remaining flexible enough to meet application-specific requirements. Dedicated file formats, in turn, are essential tools for storing specialized data sets in a compact and structured manner. Kithara combines both aspects.

To enable gapless acquisition of measurement data, the underlying storage medium must first be adapted accordingly. Kithara's Windows real-time extension addresses this requirement with a dedicated storage functionality. This takes the form of a specialized device driver that provides direct access to modern, high-performance SSDs with an NVMe interface, enabling higher sustained read and write speeds than during regular operation as well as real-time access. These high throughput rates are achieved by using specially configured, very large data blocks for read and write operations. The acquired raw data can be written and accessed in parallel, while data loss caused by high-frequency real-time timers is avoided thanks to low jitter.



Structure by Format

Within the real-time context, the acquired raw data is stored using specialized file formats. Kithara relies on MDF (Measurement Data Format) and PCAPng (PCAP Next Generation) to store large volumes of information in a hierarchically structured manner, make them immediately available for further use, and thus create a structured basis for automated analysis and machine learning workflows. The two formats cover different ranges of applications.

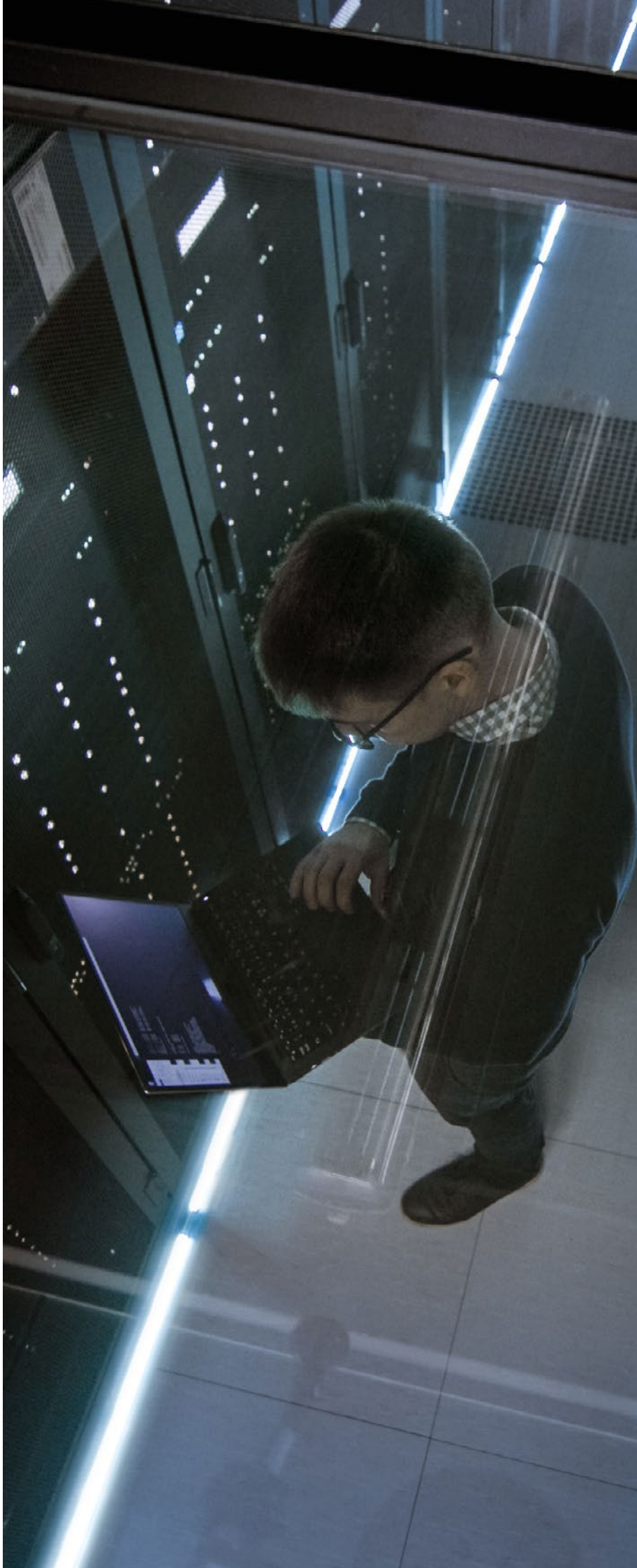
MDF is a binary file format for measurement tasks and the de facto standard in the automotive industry. Standardized by ASAM, this standard is particularly notable for its compact and flexibly configurable storage structure. Data can be efficiently compressed, sorted, and indexed, optimizing read access, while file sizes can be scaled freely—ranging up to the terabyte scale for individual files. Due to its focus on automotive applications, MDF can also be used to store raw data from automotive bus systems supported by Kithara, such as CAN, LIN, FlexRay, and Automotive Ethernet, either as bus messages or as appropriately decoded signals. The finely granular structure of the MDF format poses a particular challenge when using the large data blocks that Kithara's storage functionality relies on to achieve high read and write throughput. To avoid significant performance losses during high-speed SSD storage even when using MDF, an optimized block allocation algorithm is employed. This algorithm is capable of efficiently aggregating data of different types and structuring them within an MDF file in complex hierarchies.

By contrast, PCAPng is a much simpler format supported by Wireshark for capturing data packets within networks. This enables developers to record, monitor, and evaluate specific measurement data packets as files. Although the PCAPng format does not come close to the complexity of MDF, it offers the significant advantage that the stored measurement data files can be opened and read directly using the freely available Wireshark tool, which, for example, simplifies on-site diagnostics.

Data-Intensive Measurement Applications in Practice

While the combination of gapless real-time data acquisition and data storage using dedicated file formats is conceivable for a wide range of data-intensive applications, it was developed by Kithara specifically as a measurement data solution. Particularly with regard to hardware-in-the-loop simulations, additional real-time functions of the Kithara RealTime Suite can be incorporated.

In this context, the real-time drivers for GigE Vision and USB3 Vision are especially relevant, as industrial cameras are among the sources generating the highest data volumes in this field. To capture image data as seamlessly as possible, real-time network drivers or real-time USB drivers are used. In combination with significantly expanded buffer memories, these drivers prevent the loss of data packets, enabling the creation of high-quality data sets suitable both for traditional analyses and for the training and validation of AI models.

Since such applications are frequently found in the automotive industry, particularly in the development of driver assistance systems, the aforementioned automotive bus systems are also supported in real time.

Uwe Jesgarz, Martin Ebert