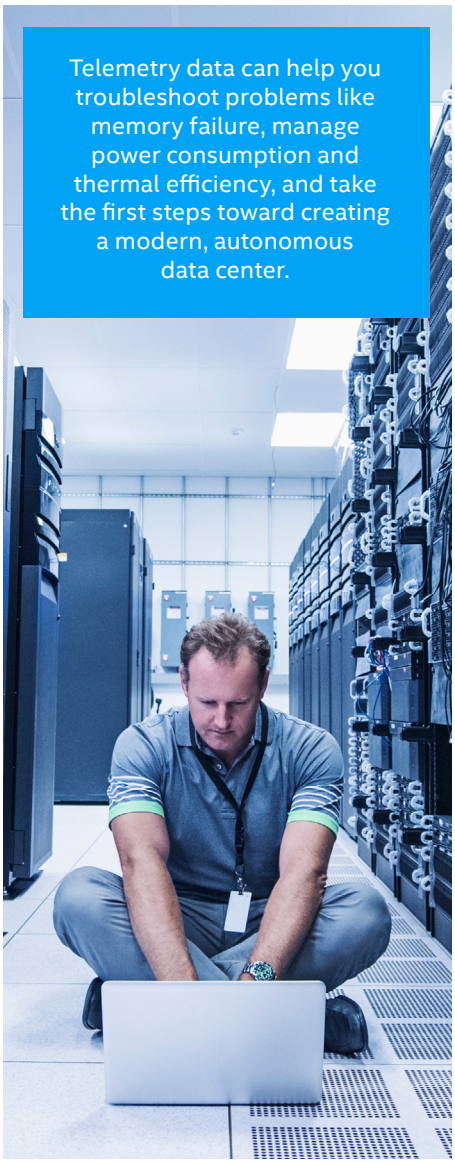
White Paper

Telemetry and Orchestration
Data Center Operations

Building Intelligence into Your Cloud and Data Center

Collect and visualize telemetry data for efficiency and reliability.



If you are in charge of maintaining and troubleshooting data center hardware for your organization, you might not be aware of what hardware telemetry can do for you. You might not even be aware that hardware telemetry exists, because it generally doesn't present itself to you unless you take steps to expose it.

Telemetry, in general, simply refers to a stream of data coming from any number of sources, including servers, software, and environmental equipment. The topic of this paper is more specifically hardware telemetry, which refers to the stream of data generated by the components on your hardware platform, such as the CPU, memory, and PCIe interface.

Intel® Xeon® Scalable processors incorporate registers for monitoring cache, CPU frequencies, memory bandwidth, and input/output (I/O) accesses. Hardware telemetry is generated from a robust set of model-specific registers (MSRs) and sensors on the Intel® platform. The MSRs contain reconfigurable and fixed performance monitoring units (PMUs), in addition to configuration information. Any metric that is derived from a PMU counts as hardware telemetry.

Software telemetry—data such as network traffic and utilization that is counted by the kernel, operating system, hypervisor, or the software running on top of those components—is also useful and important, but it is mostly beyond the scope of this paper.

This paper will show you why and how to begin collecting telemetry from your Intel architecture. You'll learn which data to collect and what tools to use to take advantage of that data for purposes of troubleshooting, orchestration, and automation. And you'll get a glimpse of Intel's larger vision for the future of the modern autonomous data center. But first, this paper will explain a few of the reasons it's worth your time to start collecting and using telemetry data.

Why telemetry?

You can achieve several benefits by exposing, collecting, and visualizing your hardware telemetry. First, this data can prove invaluable in helping you track down and fix some of the thorniest problems you might encounter, such as memory failure or overall system performance that doesn't meet your expectations. Second, beyond troubleshooting, hardware telemetry can help you manage system health elements, such as power consumption and thermal efficiency, which can help reduce the risk of hardware problems developing. Finally, by exposing your hardware telemetry and beginning to use it for data center orchestration, you are taking the first steps toward creating a modern, autonomous data center in which automation and self-healing will increasingly replace human intervention.

Telemetry for troubleshooting

When problems occur in your data center, it can be difficult and time consuming to identify exactly what's gone wrong. Sometimes you're faced with a performance problem or degradation, and you don't know whether it's caused by a failure of hardware, such as memory, or a workload imbalance that's overloading certain servers or components. As your data center grows larger in scale, these problems tend to occur more often and become even harder to troubleshoot.

Without the right information, you can find yourself constantly struggling to put out fires and not knowing what exactly is causing them.

One of the most powerful uses for hardware telemetry is to provide you with a finer-grained picture of exactly what is going on in all of your servers and their components. Unlike higher-level utilization metrics, hardware telemetry can help you answer questions like:

- How is your memory actually being used? Which applications are using it?
- Do you have a lot of resource contention in your cache?
- Do you have a properly hyper-threaded application that's being well distributed? Or are you only using one core? And what's the performance of that core?
- What are the power and temperature conditions inside each server? Is your server throttling?

These kinds of information and insights can speed and simplify problem resolution, freeing your time up for more interesting and proactive work, like infrastructure modernization.

Manageability and orchestration

Come for the troubleshooting, but stay for the opportunities. Your immediate need might be to put out fires more quickly; but once you have telemetry monitoring in place, you'll discover that everything about managing your data center starts getting easier and more efficient.

When you can see where the bottlenecks and overloads are occurring over time, you can start fine tuning your processes and orchestrating your workload placement based on real-time online workload profiling to avoid those performance bottlenecks. This helps to protect network performance and balance the load across the infrastructure.

Telemetry can uncover more than just workload congestion. You can use it to check configurations and gain insights into resource utilization, power efficiency, and general system health. In a well-orchestrated data center, you'll be better able to identify anomalies that might be early indicators of possible points of failure. Your focus can shift from constantly putting out blazes to early identification—or even prediction—of the first signs of a smoldering problem.

Greater efficiency and fewer overloads can reduce the frequency of failures and improve key performance indicators (KPIs), including total cost of ownership (TCO), reliability, security, performance, and power consumption.

Getting started with telemetry

If you are new to telemetry, a good place to start is with the Intel Telemetry Collector (ITC). ITC gives you access to the same collection of tools that Intel's own performance engineers use in performance reviews. With ITC, you can ingest and visualize data from various sources and multiple machines.

ITC is a reference collector that quickly gives you an introduction to the different metrics available, including power and thermal statistics, performance counters, process activities, threads, and operating-system-level disk, network, and memory statistics. The ITC visualization highlights typical pain points such as memory bandwidth, NUMA imbalance, and interrupt request (IRQ) affinity issues. You can identify system imbalances, frequency inefficiency, and memory and I/O issues.

ITC is a good first step into telemetry, but it is not a fleet solution. It can give you your first deep look into what's going on in your data center by looking at a limited number of your servers, but it's not designed for telemetry at scale for hundreds or thousands of servers. When you're ready to monitor all your server telemetry, then you'll want to start building a scalable open source telemetry software stack, as described in the next section

Building your software stack for monitoring

To build a scalable telemetry monitoring system, the essential software you'll want to set up will include a telemetry agent to expose the data, a database to collect the data, and a visualizer for monitoring the data and triggering notifications. The easiest and best supported tool chain for this purpose is the Prometheus stack. Prometheus is a database you can use to collect the data from all the collectd or Telegraf telemetry agents. Then you can use Grafana to build a dashboard for viewing the data.

Intel supports the Prometheus tool chain by adding exporters for Intel hardware components, either through collectd or Telegraf exporters or through direct exporters. The stack includes data collection using exporters, integration with Kafka, a time-series Prometheus database, and a visualization tier for Grafana. This Prometheus stack supports the open telemetry standard, is containerized, and is not difficult to set up. It can be supplemented by additional Intel hardware collectors to expose data from components such as Intel® Optane™ technology and the Data Plane Development Kit (DPDK) for networking data.

The collectd or Telegraf telemetry agent must be installed on each node, but the resource requirements are minimal, and the installation and configuration can be scaled by using software-configuration-management tools such as Ansible, Puppet, Salt, or Chef.

Once the telemetry agent is installed and configured, it can expose multiple streams of data to Prometheus for analysis. Figure 1 shows an example of a telemetry dashboard built using Grafana. As you add more types of data to your collection, your dashboard can expand to include different kinds of information. Notice that your dashboard can include alerts for situations that require attention.

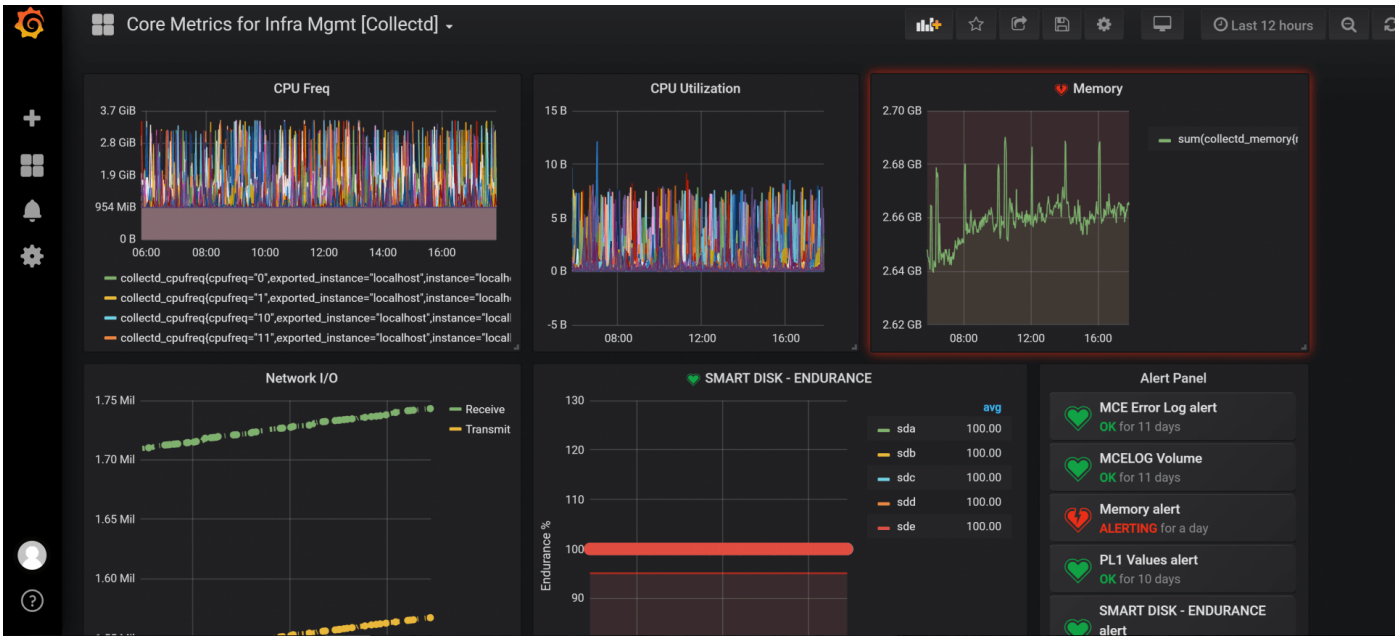


Figure 1. Example telemetry visualization using Grafana[2]

Alternatively, you can choose to use the Elasticsearch, Logstash, and Kibana (ELK) stack for data collection, analysis, and visualization.

What data to collect

As you are setting up your software stack to analyze telemetry streams, you'll face the question of how much data to collect and how often. If you collect everything every few seconds, your log files will grow exponentially, and you'll risk negatively affecting your system performance and storing a lot of raw data that you probably do not need. Therefore, it's a good idea to start slowly, beginning with just your highest priority data and gradually adding more streams. As a best practice, you can scan data less frequently until an issue is identified, and then increase the frequency until the issue is resolved. There are other techniques on how to handle the quantity of data that will be discussed in a follow-up paper. Table 1 provides some reasonable guidelines for how often to collect which data.

What Data to Collect	Why to Collect It	How Often to Collect It
System configuration, kernel version, and log messages	Troubleshooting and failure-trend detection	On reboot
Power (CPU, memory, and system) and thermal (inlet, outlet, DIMM, and CPU airflow)	Manage power and thermal efficiency; identify spikes as early indicators of failure	30 seconds to 1 minute
Performance of the CPU, cache, and memory	Target machines for workload colocation	1–10 seconds
Memory errors	Identify issues causing device degradation and early indicators of failure	1–5 minutes Trigger more frequent scans on first uncorrectable memory error (UME)
DIMM performance and health	Identify issues causing device degradation and early indicators of failure	15–30 minutes and on startup
Throughput	Failure-trend detection	30 seconds to 1 minute

Keep in mind that when it comes to telemetry data, sometimes less is more. The objective is not to collect the maximum amount of data, but to collect the right data in a useful amount.

Other tools

As you progress further into monitoring your telemetry data, you'll want to fill any gaps in tooling to ensure that you are monitoring all the important sensors in your data center. You might also want to integrate your hardware telemetry with your containerized environment. Additional tools you might find useful as your monitoring grows more sophisticated include:

- Intel Power Thermal Aware Solution (Intel PTAS) lets you measure power for each individual server, using sensors already present in the CPUs—a feature available in all modern Intel Xeon processors.
- Intel Data Center Monitor (Intel DCM) collects and analyzes the real-time health, power, and thermal details of a variety of devices in your data center.
- The Redfish API lets you measure power at each individual server, using sensors already present in the CPU—a feature available in all modern Intel Xeon processors.
- cAdvisor lets you collect data from the containers in a containerized environment and send the data to Kubernetes, so that health factors can be used when orchestrating workloads.
- Telemetry Aware Scheduling (TAS) is a Kubernetes extension that Intel built so telemetry can be used with placement rules to enable self-healing and better optimized orchestration.

The road ahead

When you expose the telemetry in your data center, you first gain an improved ability to troubleshoot failures and performance problems after they have occurred. As your monitoring improves, your telemetry dashboard can include alerts that allow you to respond immediately to situations that require attention. This allows you to address problems right when—or even before—they begin to impact the performance of your data center.

This use of telemetry can greatly improve the infrastructure efficiency, reliability, and security of your data center. But these are also the first steps toward a grander modernization of data center infrastructure that's on the horizon. Intel is working toward the vision of a modern autonomous data center (MADC), in which the telemetry stream is harnessed for use with predictive artificial intelligence (AI) to automate the orchestration of the infrastructure and to prevent or remediate issues without human intervention.

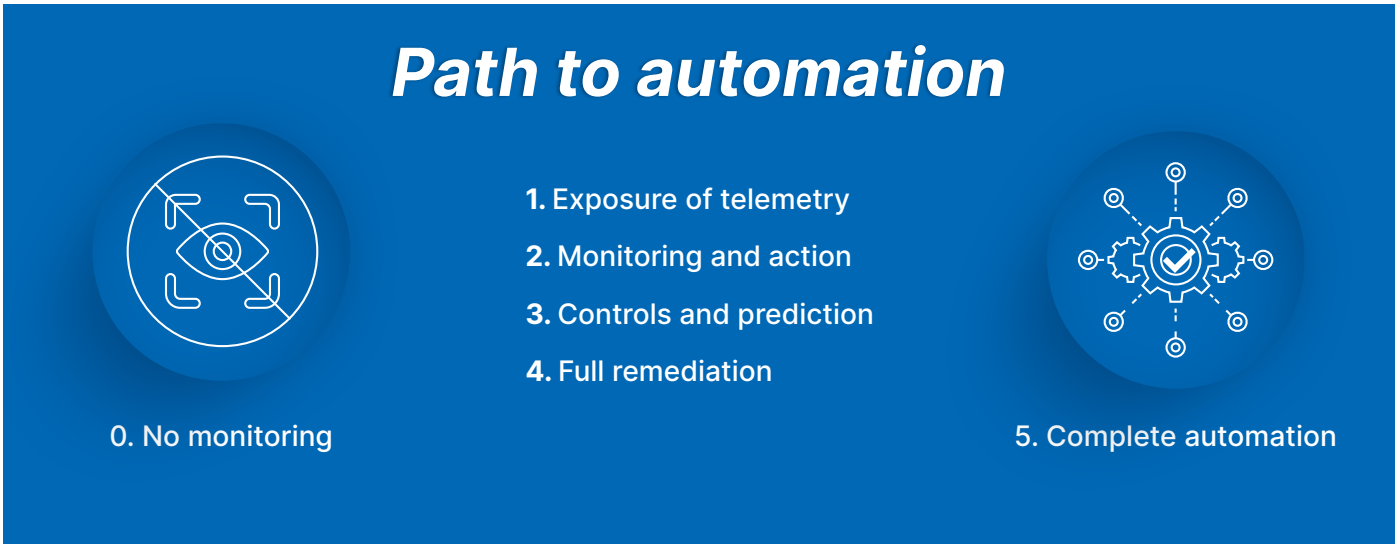


Figure 2. Five phases of data center modernization

Intelligence will arrive in the data center in five phases, as shown in Figure 2:

1. **Exposure of the telemetry data**, which is the focus of this paper.
2. **Monitoring and action**, using an open source software stack to build dashboards and implement alerts.
3. **Controls and prediction**, where control-system automations are built on a small scale using telemetry.
4. **Full remediation**, with multiple automations and control systems working together intelligently.
5. **Complete automation**, when no human intervention will be involved in self-healing systems.

While complete automation of the data center is not yet a reality, some hyperscalers working with Intel are well into phase three with predictive automation and controls. These hyperscalers understand the benefits to be gained from harnessing the telemetry in their giant data centers.

You, too, can start benefiting from deeper visibility into your data center's components and operations. By exposing and monitoring telemetry, you can troubleshoot data center problems more easily and gain improved efficiencies, reduced bottlenecks, and early alerts to warn you when problems might be developing. Then you'll be on the road to data center automation.

Orchestrating and Troubleshooting the Data Center

Collect and visualize server data for efficiency and reliability.

Learn more and get started today

Visit the Intel Cloud Hub at https://intel.com/cloud, and explore these resources to assist you in understanding and getting started with telemetry monitoring:

"Cloud Telemetry: Advancing Your IT Strategy": intel.com/content/www/us/en/cloud-computing/telemetry.html

"Optimized Resource Utilization with Cloud Orchestration": intel.com/content/www/us/en/cloud-computing/cloud-orchestration.html

"Cloud Automation: Hands-Free Workload Allocation and Balancing": intel.com/content/www/us/en/cloud-computing/cloud-automation.html

[1] Talk to your Intel representative about how to get access to ITC.

[2] Intel. "From Concept to Reality. So … Are You Ready to Rock on Telemetry?" April 2019. https://itpeernetwork.intel.com/from-concept-to-reality-so-are-you-ready-to-rock-on-telemetry/.

Printed in USA 0321/SM/PRW/PDF Please Recycle 346415-001US