Cloud Native at the Edge: Intelligent IoT Gateways

arm

A Project Cassini Reference Implementation

White Paper

Index

Executive Summary

In the Internet of Things (IoT), due to scale, infrastructure costs, the need for greater efficiency, and new requirements for real-time applications, deployments are moving data processing from the cloud to the edge. Cloud-native development techniques facilitate this transition, by making it easier to deploy new workloads and services across a diverse, secure ecosystem.

Arm's Project Cassini is delivering the foundation that enables cloud-native development at the edge. With Project Cassini in place, edge developers will be able to take advantage of a wider diversity of underlying hardware platforms, while benefiting from the innovation that DevOps, agile software, and cloud-native approaches have brought to the data center.

Smarter, Scalable and More Flexible Deployment

The traditional IoT deployment model has been to generate data using diverse connected end points at the edge, where cost, device size, and power consumption are key, and then act on this data in the cloud, where ample resources for power, compute processing, and storage capacity can be used to make decisions and initiate responses.

There are several reasons, however, why this approach of sending data from the edge to the cloud for processing is becoming unsustainable. Transmitting data to the cloud increases power consumption, because wireless communications use energy to send and receive data, and can be expensive when using pay-by-the-kilobit cellular services. Also, as deployments grow in size and scope, the sheer amount of data needing to be transmitted can overwhelm the communication infrastructure.

Perhaps more important, though, is the fact that new and emerging applications depend on real-time services at the edge, and make use of Machine Learning and Artificial Intelligence (ML/AI) for real-time decision-making. These applications – think avionics, robotics, autonomy, logistics, energy delivery – often involve safety-critical operations that,

if performed too slowly, can cause harm to people and property. These applications simply can't tolerate the delays introduced by sending data back and forth to remote servers on the cloud.

To perform more data processing and analysis locally, deployments are putting more intelligence into IoT edge gateways. The result is increased operational efficiency, support for real-time operation, and faster decision-making at the edge.

Arm-based System-on-Chips (SOCs), with their low-power, compact hardware configurations and familiar software toolchains, are well suited for use in intelligent IoT edge gateways, but differences in hardware implementations, at the silicon and board levels, as well as variations in firmware and middleware, create inconsistencies that can complicate development, deployment, and ongoing management. These inconsistencies – which are in many ways the inevitable result of such a vibrant and diverse Arm ecosystem, created to support an exceptionally wide variety of applications – can require developers to spend time resolving issues at low levels of the stack, and this can divert attention away from other important tasks, such as adding new features and product differentiation.

What's needed is a consistent foundation on which to build, so developers can focus on the value that comes from creating and deploying new workloads and services from a single code base. A stable foundational architecture also makes it easier to apply the benefits of DevOps and agile software development to the edge. Using cloud-native techniques, developers of intelligent IoT edge gateways can benefit from the same kinds of new approaches that have transformed data centers in recent years, making it easier to port software, reuse code, enable continuous development, and deploy faster.

> Cloud-native enabled gateways facilitate real-time services at the edge and add value through ML/AI.

Accelerating Cloud-Native Edge with Project Cassini

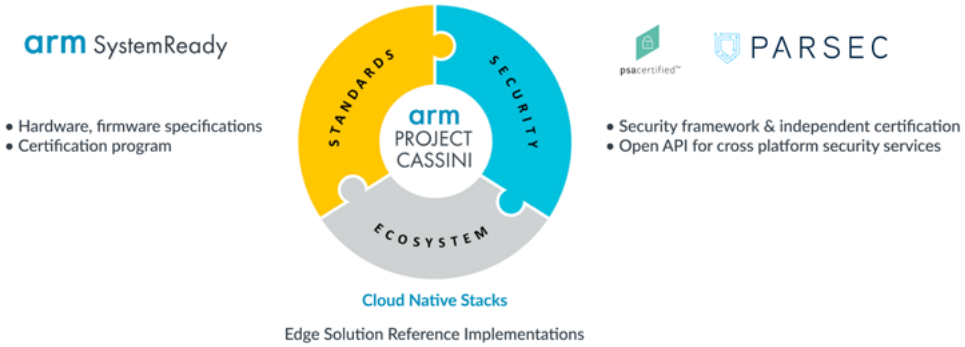
Arm created Project Cassini to meet this need. It's a comprehensive industry initiative aimed at bringing the cloud-native experience to the edge, so developers can deploy workloads and services more easily across a diverse and secure edge ecosystem.

By enabling scale and security across a wide set of devices, each supported by an extensive ecosystem and servicing different IoT applications, Project Cassini will bring consistency and portability to the industry, while also enabling the migration of cloud-native applications to the edge.

For example, Project Cassini will make it easier to deploy software onto intelligent IoT edge gateways because hardware and firmware will use standards that enable secure, consistent operation, as well as the ability to run OSes, applications, and other off-the-shelf software, without modification. Project Cassini will also make it easier to define, test, and update software for IoT edge gateways because developers will use cloud-native techniques and portable, commodity software components that mask underlying complexity.

This standards-based approach will provide a baseline that makes it easier for OS and virtualization vendors to support multiple hardware architectures, and will unlock an entire ecosystem, built on cloud-native principles, that supports new, real-time services and leverages ML/AI workloads.

Project Cassini uses standards, security, and ecosystem to create a cloud-native experience for edge



Project Cassini is built on three foundational components: robust standards, security APIs and microservices validated by certification programs, and reference solutions that make cloud-native stacks a part of edge development.

- **Standards: Arm SystemReady**

  The Arm SystemReady program is a foundational certification program that defines a minimum set of hardware and firmware. For hardware, SystemReady defines a common Base System Architecture (BSA) and a set of market-specific supplements. For firmware, the Base Boot Requirements (BBR) describes standards-based boot recipes and implementations. Standardization across low-level firmware and hardware features makes it possible for off-the-shelf and community software distributions ("distros"), as well as other workloads, run seamlessly on diverse Arm platforms.

- **Security: PSA Certified and PARSEC**

  Together, these two initiatives ensure a baseline of security for all connected devices. The PSA Certified initiative offers a security framework and certification scheme that gives an objective assessment of the quality of the implementation for the device root of trust, thereby establishing a foundation of trust on which to build. The open-source Platform Abstraction for Security (PARSEC) initiative provides secure root-of-

trust abstraction and common runtime security services for applications in a platform and architecture-agnostic manner, for application-wide security.

- Ecosystem: Cloud-native stacks and reference implementations

  Given the consistency in hardware, firmware, and security provided by the other foundational components of Project Cassini, it becomes possible to apply the modern software development and deployment approaches, commonly used in the cloud, to the edge. For instance, Project Cassini makes it easier for developers to use cloud-native techniques, such as service-oriented architectures, container orchestration, and Continuous Integration and Continuous Delivery (CI/CD) pipelines, to streamline development and accelerate deployment. Project Cassini also makes it easier for software vendors to offer portable solutions that are easy to integrate into the larger system and run without modification on every device. This is especially important in large scale deployments, such as smart cities, where a centralized entity manages a wide range of services and applications across multiple, widely different devices, gateways or micro-servers.

  As part of Project Cassini, Arm is actively working with various ecosystem partners to build and deliver reference platforms. These platforms use Project Cassini as a foundation to address specific use cases across the edge. They provide proof of concept and, in many cases, also include actual working open-source code or new product offerings.

What That Means for IoT Deployments

With Project Cassini in place, it will be easier and more cost-effective to design, launch, manage, secure, and scale IoT deployments. Project Cassini will provide the homogeneous experience that developers who use cloud-native and agile-software methods are used to, while keeping the value of Arm's product diversity and its vibrant ecosystem.

The manufacturers of intelligent IoT edge gateways will have access to a wider selection of Application-Specific Standard Products (ASSPs) and will be able to integrate those ASSPs into their systems more easily, regardless of the OS they've chosen. Reduced friction at the hardware level, along with increased portability at the software level, will make it easier to manage products across platforms, with greater scalability, because there's less need to maintain product-specific variants in software and fine-tune integration.

Also, the consistency provided by Project Cassini will make it easier to port cloud-native stacks that have traditionally been developed on the x86 architecture onto Arm-based devices. This will, in turn, expand the Arm edge ecosystem, in much the same way that cloud-native development has expanded the options for Arm in the datacenter space.

Tangible Benefits

These various benefits will combine to make the edge the core infrastructure needed to support new and emerging real-time services, and will open up new opportunities for edge-centric innovation. Furthermore, by being able to move seamlessly across multiple products and platforms throughout the edge, and being able to manage various devices the same consistency, developers will have more time to focus on differentiation and will be able to unlock added value in modern IoT deployments:

- **Increased Operational Efficiency**
  Since intelligent gateways bring processing and storage closer to the edge, decisions can be made faster and responses can be quicker. Being able to make decisions at the edge, instead of in the cloud, reduces latency and enables the real-time responses needed for new applications, especially in the industrial sector.

- **Support for Multi-Tenancy**
  Bringing intelligence closer to the data means that more of the data, especially data that typically wouldn't be moved to the cloud, can be made more useful. Intelligent IoT edge gateways can process this data, to gain more and better insights, and to enable multi-tenancy operation, where the same data set can be used to support multiple applications. That increases the service options and creates new revenue streams.

- **Flexible Scalability**
  Rather than rent more cloud cycles to support growth, developers can extend applications and expand coverage quickly and cost-effectively by using a deployment approach that is more localized and granular. At the same time, distributed computing means decisions can be made throughout the edge, not just in a single cloud-based location.

- **Optimized Deployment Cost**
  Using a scalable, standards-based approach to development saves time and effort at each stage of the project and allows for greater design optimization. This, in turn, lowers the total cost of ownership, with savings in hardware, services, communication, power consumption, reliability, and more.

- **Robust Device Management**
  Growth in the IoT is often held back by lack of uniformity and the need to micro-manage sub-components, from OSes, and Application Enablement Platforms (AEPS), to the devices themselves. With Project Cassini, all devices are managed the same way, with the same services, for seamless scalability. Developers can leverage a single code base across the entire Arm ecosystem, and thereby enhance their existing services and expand into new markets.

IoT Edge Gateways Deliver

- *Computing power*
- *Embedded OS & I/O*
- *Communication bridge*
- *Application and device security*
- *Multi-tenancy*

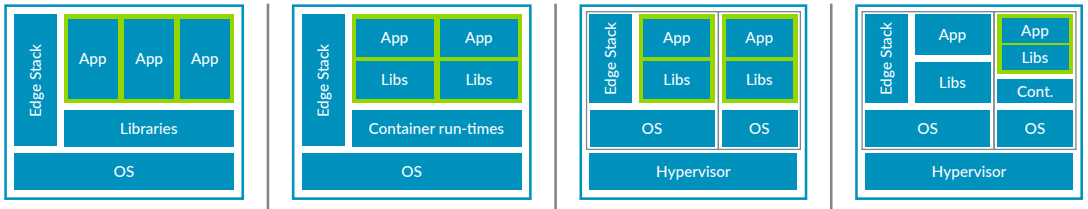
- More Uptime

  Distributed computing means tasks are shared between the cloud and the edge. The two remain connected yet independent, so a failure in one part of the edge doesn't shut down the entire operation. Moving away from centralized computing, where everything resides in the cloud, makes the deployment less vulnerable to downtime.

Software Just Works

The below figure shows how Project Cassini enable seamless portability, so software "just works." The green boxes represent the same applications being deployed across different device types. With Project Cassini providing the foundation for consistent operation, the same application, client, or device-management function can be enabled on any device, regardless of its CPU or OS hypervisor.

Run cloud native applications across the edge



Traditional model

- Reduced platform overhead and portability friction
- No need to re-architect applications
- Real-time OS and application capability

Device Gateway

Container model

- Manage application dependencies
- Enable faster update / rollback capability
- Multi-tenancy policies

IoT Gateway

Hypervisor model

- Deploy legacy applications without porting code
- Temporal and spatial isolation between VMs and OS
- Suitable for real-time and safety critical applications

Industrial Gateway

Hypervisor + Container model

- Leverage hypervisor and container deployments
- RTOS stack and applications + containerized Linux applications to consolidate nodes

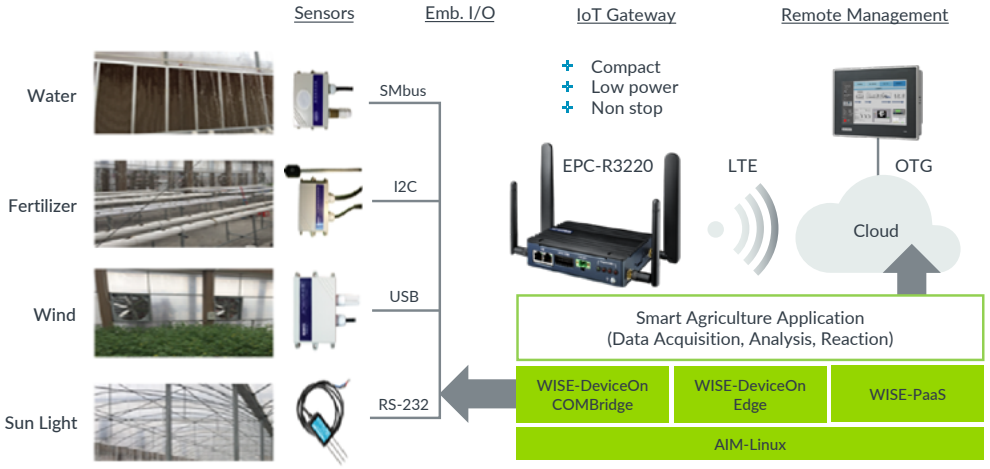
Workload Consolidation Gateway

Project Cassini enables a cloud-native development approach for IoT workloads.

Real-World Example: Smart Agriculture with Advantech

Developers can build on Project Cassini using industry-specific solutions provided by Arm ecosystem partners, such as Advantech, a leading supplier of embedded and automation solution platforms for IoT system integrators. The diagram shows a sample use case, with an IoT edge gateway used to manage data acquisition, analysis, and reaction in a greenhouse.

Use Case: Smart Agriculture



Data Acquisition, Analysis, and Reaction in a Greenhouse

The gateway obtains data from multiple different sensors through different I/O interfaces, converts data into an appropriate protocol for further data analysis and then takes the necessary actions to control the growing environment. Transmissions to the cloud are made more efficient by only including what's relevant to the cloud.

The Advantech approach simplifies development and saves time, because so much of the upfront work is already done. Advantech provides an Arm Cortex-A "box" with I/O for connecting to sensors and a performance-optimized BSP with pre-integrated I/O and drivers for LTE cellular communication.

Application software builds on AIM-Linux, a Yocto-based platform for creating an optimized Linux core, plus peripherals, for each silicon partners. AIM-Linux also includes an application add-on service that makes it easy to extend the app with useful features, such as diagnostics and device management. For example, Advantech's Edge AI software package supports the addition of AI capabilities, such as deep learning, computer vision, multimedia, and imaging, so the deployment can easily add AI-driven features, such as security cameras with video analysis to recognize when people enter or exit greenhouse.

"Project Cassini reduces the barriers between customers' product development and integration through a standardized hardware and software architecture. It has become the key solution for accelerating the implementation of Arm-based computing technology in AI and industrial IoT applications across our product portfolio."

Aaron Su Advantech AVP

The entire Advantech solution builds on the Project Cassini foundation, so developers can move quickly from prototype to production, using any Arm implementation they choose. Advantech supports a number of industry-specific applications, including transportation, medical, and automation, and middleware options, such as networking, security, and video acceleration.

Where the real value of Project Cassini comes to light is with portability, because the complete software stack can be applied across the entire Advantech portfolio. With Project Cassini devices, the same solution can be ported onto Advantech's entry-level, low-cost gateways, to support fewer than a dozen sensors, on ported onto their high-end product lines, which connect hundreds of sensors and other gateways though mesh networks. With Project Cassini, a small, family-owned farm can use the same IoT technology as the biggest agricultural outfit.

Tell Us What You Think

We're proud to count some for the most influential leaders in technology among Project Cassini's supporters. Feedback from our partners in critical. We welcome responses and seek commentary on the information we're providing. We look forward to receiving thoughts and ideas on the edge and how it can best be secured and leveraged. To share your feedback with us, we invite you to email project-cassini@arm.com.

Arm is dedicated to collaboration.
We welcome your insights on Project Cassini.

Take the Next Step

To learn more about Project Cassini and how it relates to IoT edge gateways, visit arm.com/project-cassini.

arm