Securing Smart Devices

PROTECTING INTELLECTUAL PROPERTY (AI/ML) AT THE EDGE

www.sequiturlabs.com

SEQUITUR LABS

TABLE OF CONTENTS

INTRODUCTION.....3
PROTECTING IP IN A SECURE ENCLAVE.....4
PROTECTING AN AI MODEL WHILE IT IS AT REST.....5
PROTECTING AI MODELS WITH DEDICATED HARDWARE.....6
LET'S SUMMARIZE.....7
WHAT DO I DO NEXT?.....8

INTRODUCTION

Over the last several years, we've seen significant technological innovation across nearly every business sector, including manufacturing, consumer products, healthcare and more. The adoption and integration of digital technologies into products at every level has become pervasive, with interconnectivity speeding the rise of IoT (Internet of Things). Practically every modern device we use today is now connected, allowing greater automation, control and analytics-based insights. The growth of this new area in device and computing interconnectivity is not only calling for better network and internet infrastructure, but also stronger and more effective security solutions.

The primary object of attention and protection is the intellectual property in the form of rich applications that include artificial intelligence (AI) and machine learning (ML) algorithms at the edge. While the communication and management of devices from a single place has become possible, this has made it possible for IoT device vulnerabilities to be exploited—causing enormous damages to businesses worldwide if these applications are compromised. To confront this challenge, security enabled IoT platforms are leading the way to counter this threat, providing the means to securely interface with and control a wide range of sensitive connected devices and systems such as intelligent video analytics systems, healthcare devices, robotics, autonomous vehicles and machines, industrial control systems, and smart home products.

> ***The chip designer revealed that a hacker stole test files for a 'subset' of current and upcoming graphics hardware, some of which had been posted online before they were taken down.***
>
> —Engadget

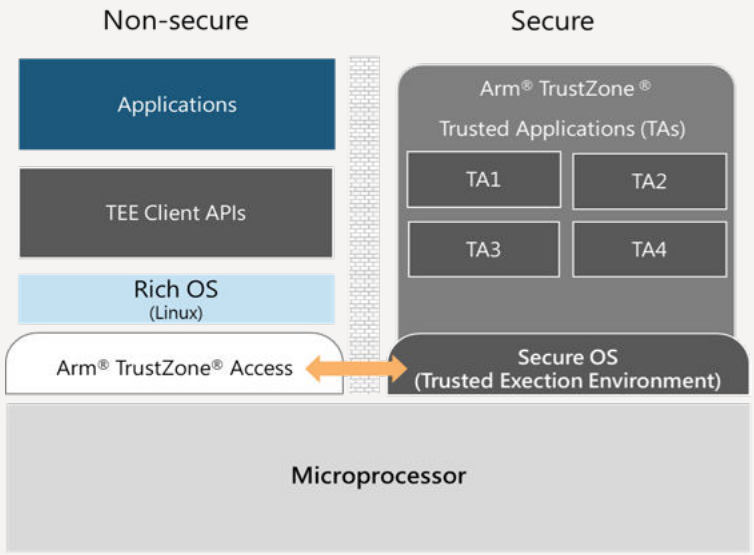
Many IoT products incorporate artificial intelligence (AI) or machine learning (ML) to conduct complicated tasks that require intelligent functionality with access to sensitive code or data sets, allowing decision making without pre-orchestrated programming. The algorithms and models that deliver this functionality represent critical intellectual property (IP) and create significant value for the products and their vendors. As there are a vast number of potential IoT security threats, manufacturers and integrators remain focused on their search for best-in-class security strategies for locking down their products—as these algorithms and models simply cannot be compromised. Such theft of an organization's intellectual property can create long-term damage to a company's revenue and brand and must be protected.

There are countless examples of IoT security attacks and their impacts on some of the most well-known brands in the world. In a case that occurred in March of 2020, criminal hackers cracked the Xbox Series X graphics code and AMD's future computer GPU's data and leaked the information on the Internet. According to one report , "The chip designer revealed that a hacker stole test files for a 'subset' of current and upcoming graphics hardware, some of which had been posted online before they were taken down."

An optimal security solution is implemented using a method that best fits the product, intellectual property protected, and company business model. Considerations include size of the code (memory), the type of attack surface involved, the hardware architecture, and the overall threat profile. Three methods can be used to provide robust protection for intellectual property.

PROTECTING IP IN A SECURE ENCLAVE

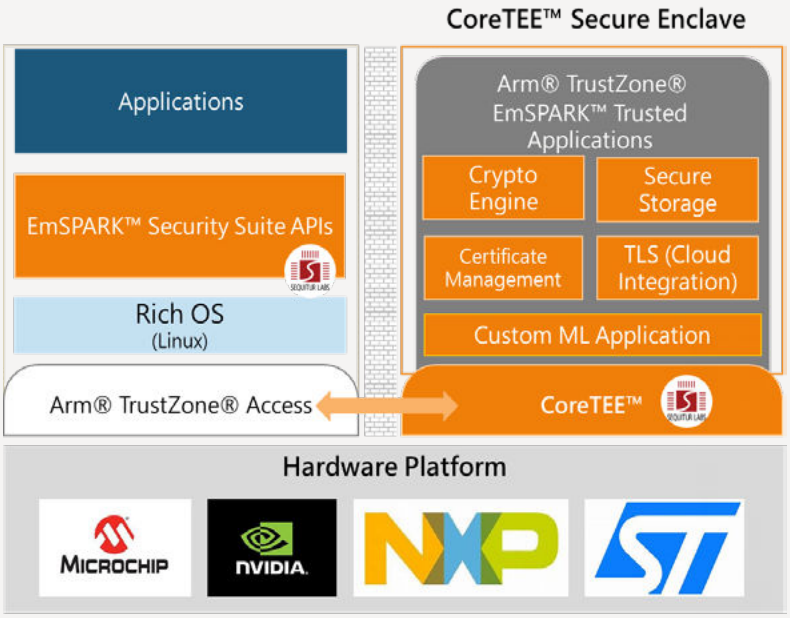
The first method for securing system applications is to isolate these applications, housing them in a secure area with restricted access. One example involves using ARM TrustZone™ architecture, where a System-on-Chip's (SOC) memory can be partitioned into a Rich (Non-secure) Environment and a Secure Environment. The Rich Environment is larger in memory size and houses known (public) software, such as Linux kernels and open source supporting applications (e.g., OpenSSL). The Secure Environment has a small memory size and houses a Trusted Execution Environment (TEE) secure operating system. Applications that need to be protected are included here along with applications that support the securing process (e.g., key / certificate management and cryptography). These are known as Trusted Applications (TAs).



In such an architecture, the secure application process works by allowing the IoT device's application, running in the Rich (Non-secure) Environment, to make a request to the Linux kernel to access the Secure Environment. The Linux kernel is suspended on one of the SoC's cores, giving access to the TEE; the TEE then resumes from suspension and invokes the requested Trusted Application. The TEE then accesses the non-secure memory (RAM) and acquires data through the shared memory between the two environments. A layer of supporting trusted applications can also reinforce security to further harden the environment, including:

+ **Cryptographic Trusted Applications:** The deployment encryption and hashing algorithms.
+ **Certificate Management for Trusted Applications:** for managing credentials.
+ **Secure Storage Trusted Applications:** for storing critical data in the Secure Environment.
+ **TLS Trusted Applications:** secure sockets for communication with external servers.

Sequitur Labs' EmSPARK™ Security Suite provide a complete set of applications, API's and development tools needed to implement this method.



This method is effective in protecting critical algorithms including Machine Learning (ML) code. A case study on this method is here: Case Study: Protecting Critical Intellectual Property with EmSPARK™ - Sequitur Labs Inc..

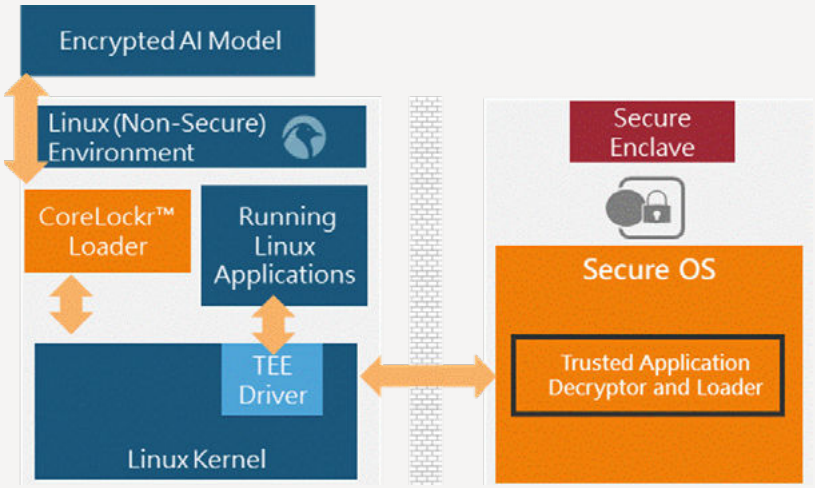
PROTECTING AN AI MODEL WHILE IT IS AT REST

While protecting IP in by housing the application in a secure enclave is a very secure approach, it requires development of a Trusted Application and generally works well for small memory footprints.

Artificial Intelligence (AI) models may not fit these criteria and may need a different approach. An efficient way to protect these models is to ensure that they are encrypted while at rest. This can be addressed with virtualization.

In this approach, the following steps are taken:

+ The application is encrypted, and locked to the device, in storage (e.g. flash memory).
+ EmSPARK's CoreLockr Loader initiates the sequence of loading and running the model.
+ A Trusted Application, housed in the secure enclave, verifies, decrypts, and loads the application.
+ Interfacing with the Rich Environment's operating system (e.g. Linux), the Trusted Application loads the application directly into Random Access Memory (RAM) and runs it.



EmSPARK™ provides two mechanisms to exchange confidential information with the device:

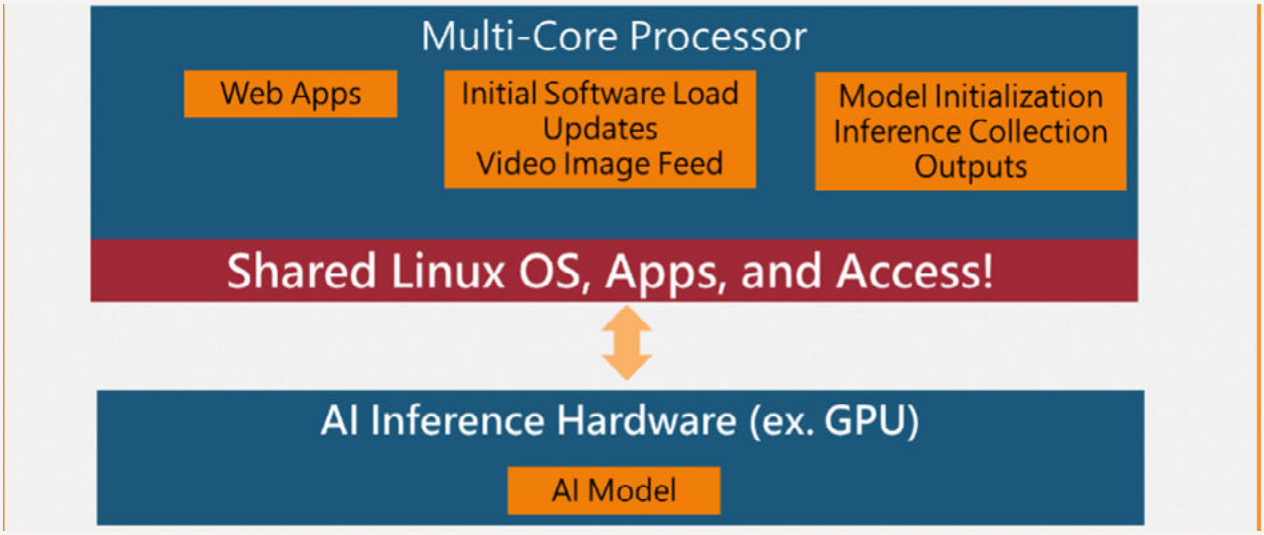
+ **Opaque Keys:** These are device-specific encrypted and signed keys to be loaded into the Trusted Execution Environment's key store.
+ **Opaque Objects:** These are device-specific encrypted and signed data, to be decrypted on the device.

In this method, the AI model is delivered as an Opaque Object, decrypted, run, and then after completing use of the model volatile memory (RAM) is cleared.

This method is an efficient way to greatly reduce the attack surface, and overall risk profile applied to running an AI model on an edge device.

PROTECTING AI MODELS WITH DEDICATED HARDWARE

Increasingly, AI models are deployed using a hardware that provides an inference platform or AI accelerator. In these architectures, the model is sharing the same operating system (e.g. Linux), applications, and developer access as supporting applications (web apps, algorithms initiating the feed and analysis of data, etc). Consider the example before, using an architecture for Intelligent Video Analytics (IVA):



In this design, the model can be accessed or corrupted by developers and applications on the device. This can be addressed using a Hypervisor model.

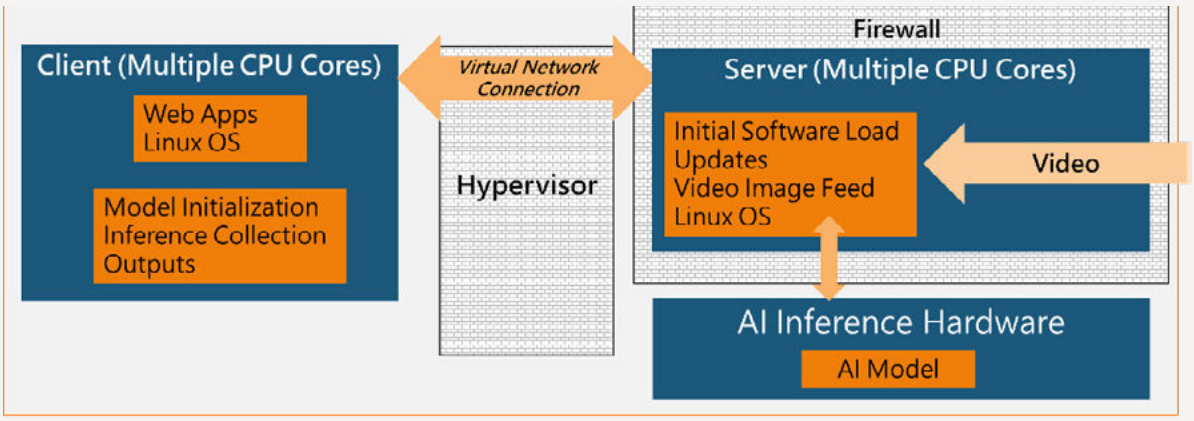
A hypervisor is defined by the following types:

+ **Type 1:** A type 1 hypervisor runs directly on the device hardware and controls the hardware, operating system, and virtual machines.
+ **Type 2:** A type 2 hypervisor runs on an existing operating system and runs it own "guest" operating system, providing subsequent virtualization.

In this method, a Type 2 Hypervisor is used.

+ The application is encrypted and locked to its dedicated hardware.
+ A type 2 Hypervisor, with virtual connection, separates the business applications (web apps, Linux operating system, Model initiation & data collection), from the model and its execution.
+ A firewall sets policies of the traffic types that are allowed to access the model.
+ A trusted application can be used to verify permissions and decrypt the model from the secure enclave.

In this architecture, the "server" cores—in this case, the instances that execute the model—are completely isolated from the business applications. This allows the model to co-exist with 3rd party applications, multiple developers, etc., without the risk of corruption or compromise.



Server Cores and AI Hardware are isolated!
Cannot be seen or accessed by the Client Cores

LET'S SUMMARIZE

Securing edge devices is simply not optional any longer. The acceleration of AI models at the network edge brings an exciting new wave of innovation to many industries, but also increases the urgency to properly secure critical intellectual property.

While the industry is becoming increasingly burdened by the growing number of threats, manufacturers and integrators are aware and incorporating next generation security into their products and solutions. This is the right move as it both enhances security and speeds time to market. These methods bring IP protection to the edge while streamlining the design of a new era of solutions and devices that are connected and secure.

WHAT DO I DO NEXT?

Evaluate EmSPARK™ for yourself. Go to www.sequiturlabs.com and get a FREE software evaluation kit and try it! If you have a specific problem you are trying to solve, drop us an email at info@sequiturlabs.com.

SEQUITUR LABS

Sequitur Labs is developing technologies for the development and management of secure and trustworthy connected devices. Sequitur's products span a range of disciplines required for trusted computing, from boot through the full device lifecycle. Sequitur's security solutions provide real business value to device makers, such as reducing BoM costs, protecting revenue by thwarting IP theft, improving product reliability and reducing liability, and improving device lifecycle management processes. To learn more about Sequitur's security platform, visit us at www.sequiturlabs.com.

Sequitur Labs Inc.,
PO Box 1127
Issaquah, WA 98027

+1 425 654 2048
+44 20 3318 1171

info@sequiturlabs.com
www.sequiturlabs.com

MWPEm-0002 Rev A. Printed in the U.S.A.