Silicon Lifecycle Management Infrastructure Pivots to Detect Security Breaches

Adam Cron
Synopsys, Inc.
Mountain View, CA, USA
a.cron@ieee.org

Firooz Massoudi
Synopsys, Inc.
Mountain View, CA, USA
massoudi@synopsys.com

***Abstract*—The potential to use embedded sensors and supporting infrastructure to detect hardware security attacks is a good use of existing device- and system-level information sources embedded in an SoC (system on chip). These integrated circuit package-level sensors and monitors have been available for some time, but their application emphasis has usually been on performance, reliability, and safety. By coupling existing and novel sensor technology with in-system and cloud-based analytics, it is possible to detect and mitigate the efforts of hardware hackers as they try to apply their skills to uncover the secrets and capabilities stored inside today's advanced electronics. This paper will review sensors and supporting infrastructure and their capabilities to help detect and mitigate a security breach.**

Keywords— Silicon Lifecycle Management (SLM); security; AI- and ML-based analytics

I. Introduction

Chips incorporate sensors and monitors to address specific functional requirements of their existence. Typical features are added to improve yield [1] or performance [2], or to address functional safety [3] issues. Today's fastest and most reliable components incorporate sensors such as voltage monitors and temperature sensors to detect voltage drop and temperature fluctuations. Using these measurements, frequency or voltage settings can be dynamically adjusted [4][5] to improve system performance without overrunning temperature specifications or power requirements of a particular load. Other sensors available today can measure reliability degradation [6], and process variation [7]. These and other functional monitors can be used to detect unexpected behaviors or changes in measured metrics across the die and alert system control software or directly address the device issue.

Depending on the sensor type, what it is monitoring, the periodicity of use, and whether it stores or processes multiple datapoints might dictate the frequency at which it might need to be queried and whether additional data storage is needed to process and filter its sample set. Various schemes are available today which allow sensors to be attached to system busses, or to special-access, purpose-built busses customized for a particular sensor type. A dedicated non-intrusive data collection fabric and/or a processor can be inserted in the SoC (system on chip) to monitor and analyze various sensor data and take proper action when a potential attack is detected.

Once a dataset for a sensor exists, in-system processing might be used to reveal system averages or load-based ranges in which the sensor's data might reside. Machine-learning algorithms might be incorporated into the analysis of these datasets to determine if new measurements fall within acceptable ranges. Once a measurement is detected that falls outside the expected ranges, a system might try to determine if this is a security attack in progress or some other anomaly being detected for some other reason. This system can potentially detect anomalies in a certain sequence of operations, for example in power-up and power-down sequences, and check against valid voltage and temperature profiles for these sequences. Combinations of sensor measurements might be fused together to determine certain constraints of normal system use to help reduce the number of false alarms being triggered by the sensor infrastructure. Figure 1 illustrates a flow and infrastructure applicable to security attack detection.

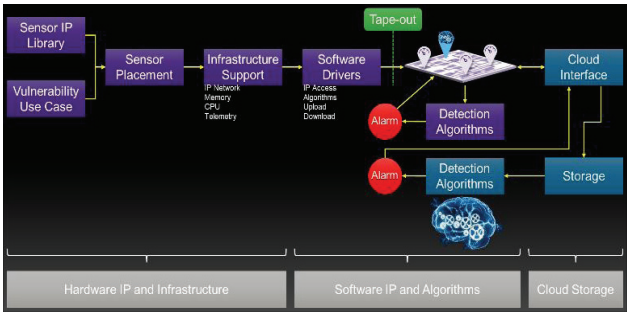

Figure 1: Three pillars of the attack detection scheme

If supported by the system in which the sensors are embedded, and as part of their system-level functionality and use model, these datasets might be shipped off to compute farms so that baseline performance models can be built for systems of the same type. Depending on the use case and what the user is trying to detect, these datasets can be compared against each other to, once again, determine if the system is being abnormally utilized which may be an indication of a security breach. The latency of such a discovery should be considered, since off-chip alarms will take time to react to by the exposed system.

The types of alarms a sensor structure with on- and off-chip analytics might provide could be as varied as the algorithms and sensors themselves. Low voltage warnings might be detected by drifting ring oscillators which could be an indication of an attacker trying to corrupt memory or state data using laser fault injection, for example [8]. An out-of-band clock frequency detection could be an indication of some sort of clock-glitching attack [9]. The convolution of a collection of sensors in combination with machine-learning algorithms might yield the most interesting alarms and error codes.

II. Sensors

To be able to detect operational anomalies in complex SoCs, a rich collection of data may be required [10]. It is critical to

have as complete a view of the silicon as possible to build the database around which analytics can be employed to detect unusual behavior at the functional (software operational) level or silicon level. Several sensors and monitor types will be discussed.

A. Temperature Sensor

Figure 2 depicts a DTS (distributed temperature sensor) which enables monitoring temperatures at multiple locations on the silicon. Each DTS can monitor 16 locations on the die. For larger dies, multiple DTS structures can be employed for finer resolution of temperature gradation across the silicon or package. Temperature sensor monitoring and correlation with software functional loads can help detect security attacks in progress [11].

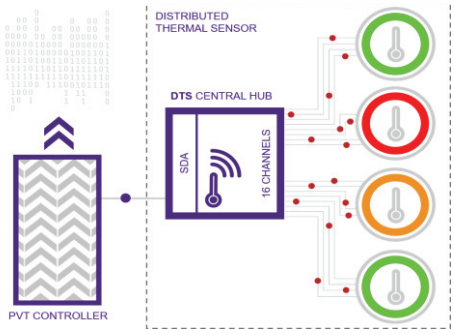

Figure 2: Distributed temperature sensor example

B. Voltage Sensor

Figure 3 shows a system of VMs (voltage monitor) which can keep track of voltage values across multiple supply lines. A single VM measures 16 independent voltage lines with an accuracy of +/-0.06% with an input range of 200mV to 1V when calibrated. [12] relates some academic work indicating that an unintended drop in voltage, if detected, could trigger a response to a security attack.

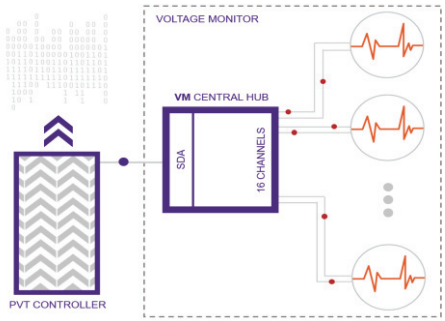

Figure 3: Voltage monitor infrastructure

C. Fast Voltage Monitor

An FVM (fast voltage monitor) and associated infrastructure, as depicted in Figure 4, can detect voltage droop by sampling the power rail at 2GHz and generating an alert within a few nano seconds. Any side channel attack that can induce a supply drop to generate an internal reset or disrupt normal functionality of the device, for example during boot stage or other critical times, can be detected by a FVM. An example security breach using a power glitch side-channel attack is provided in [13].

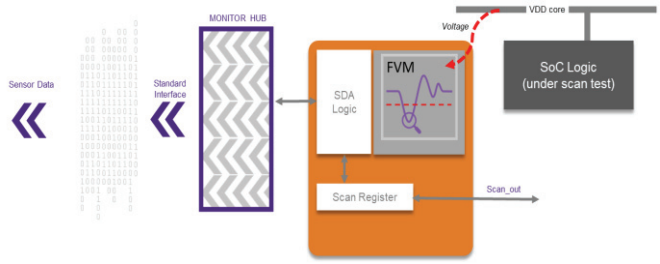

Figure 4: Fast voltage monitor infrastructure

In this case, as shown in Figure 5, the nominal supply voltage is attacked, causing it to dip below a critical threshold [14]. This voltage glitch could then cause a bit-flip in logic controlling the speed of the functional clock (for example).

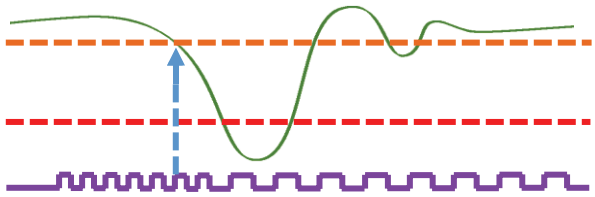

Figure 5: Voltage glitch causing critical bit-flip

D. Path Margin Monitor

Path margin monitors (PMM) are used to detect reliability or stress situations in a device. For observability of activity of a large number of signals, in the range of thousands, PMMs are used. PMM sensors are very small in area. Typically, there are many paths monitored in an SoC, potentially across a large swath of silicon real-estate. They are directly connected to the functional paths of interest. Typical applications set up a timing window in which a path or group of paths is monitored. When a path does not meet this timing window, a flag is set to alert hardware or software resources in the device to handle the event. PMMs can also detect anomalous activity on the selected paths. The data they generate can be used to detect if any part of the design that is supposed to be inactive is in fact toggling, indicating an active circuit region and potentially a security breach.

E. Process Monitor

Another method to detect reliability, stress, or process issues is to monitor a ring oscillator (RO). Ring oscillators can reflect the local performance of the silicon. However, they also vary significantly with device temperature. Performance stability also varies with voltage. If RO performance were modeled with system software loads, or silicon application loads, then divergence from these models could indicate some sort of out-of-band operation like activation of a vulnerability in the field. Device infrastructure could warn of this divergence so that an investigation could be performed to attest to proper or improper device usage.

F. Bus Monitor

PVT (process, voltage, and temperature) data can be useful to detect side channel attacks only in the context of the functionality of the SoC. Functional monitors provide information on activity levels across bus interconnect as well as between various processing complexes in the SoC. An AXI monitor, for example, can track several parameters related to bus activity and throughput at each primary or secondary port. This allows matching PVT data with functional activity levels of the circuit. Such a correlation may be able to detect a denial-of-service attack or a row-hammer style attack.

G. Signal Monitor

Signal monitors collect values on specific signals in the SoC and track sequences of operations and events. Signal monitors are designed to observe a few specific signals. They can also count several other hardware states of the device like reset events, power state changes, and clock gating signal assertions. Signal monitors have counters for these events. In addition, signal monitors can count the assertions or transitions on key signals like handshake and enable signals which can translate into abnormal activity and power usage for certain blocks in the SOC. For example, a signal monitor can track state transitions of key state machines to detect an unexpected state transition due to a voltage glitch [15].

H. Memory ECC Logic

Another level of hardware protection employed in the functional safety domain is ECC, Error (detection and) Correction Code, for memory error detection and correction and data repair. Memory ECC is also typically implemented to improve the yield and reliability of memories. These circuits can count the numbers of corrections made per unit time and thus could, themselves, be a reliability monitor, of sorts. Memory diagnostics and on-the-fly error detection and correction is recorded and reported to the sensor fabric as another piece of information to compile with all other sensor data for a complete picture of the device state.

I. Triple-Module Redundancy Elements

Similarly, triple-module redundancy (TMR) registration is also employed on specific critical bits of information that, if inadvertently flipped, would cause an unsafe or unsecure condition to occur. Counts of these bit-flipping events per unit time caught by the TMR structures can be used to detect anomalous behaviors which may indicate an ongoing security attack. Other functional safety constructs like dual-core lock-step features could also be engaged to detect security violations.

III. Sensor Control and Fabric

Each category of sensor is connected to a dedicated controller that can perform autonomous operations and control the sensors. Controllers can perform a first level of filtering and may have an alarm mechanism to signal major changes in the sensor data. Each has various programmable threshold registers to compare measured values and generate interrupts. The volume of data generated from a collection of sensors and monitors can be huge. A dedicated fabric, as shown in Figure 6, is required to interact with each sensor to collect information. Often, some filtering and hardware data processing is needed locally to generate actionable data. SLM fabric aggregates the sensor data and time stamps it to maintain correlation between different sensor data. Also, the fabric can trigger various sensors and monitors to generate data only when demanded by certain conditions. For example, if one of the sensors generates an alarm, the fabric can request and collect information from other sensors to generate a complete picture of the SoC at the time of the alarm.

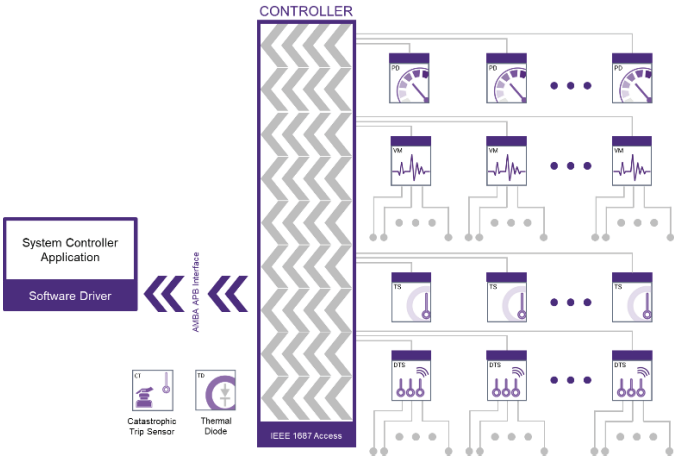

Figure 6: SoC-level sensor and monitor infrastructure

IV. Implementation and Design Automation

As the number of sensors increases in an SoC, design integration can become a challenge. Sensors and monitors are inserted in various stages of the design flow in the development lifecycle (RTL, gate, place and route), and typically during the final days of that design stage, for example after the functional RTL has been verified. If major changes are introduced for sensor integration, it can reset verification efforts and cause delays in the schedule. Design automation extends to sensor integration from RTL to synthesis and through to place and route. Sensor integration automation includes all the connectivity infrastructure between the sensors and the sensor fabric as well as integration with manufacturing test infrastructure. At the RTL level, fabric and sensor RTL and any interconnect are inserted in the design automatically by design automation tooling such as TestMAX Manager.

In addition, topological information can be included in the decision-making process concerning some of the IP implementation connectivity. For example, PrimeShield can be leveraged to select Path Margin Monitor endpoints for monitoring based on path cell types, metal layers, timing, etc.

V. Analytics

Analytics can be performed on-chip and off. On-die analytics can be used to address issues immediately. However, without the benefit of a more global perspective, it might be hard to tell what “normal operation” even means. In this case, off-die cloud-based analytics may need to be employed to get the big picture [16].

To identify any abnormal operations of an SOC due to security attacks, at times it is required to compare the operational profile of the device across a fleet of similar devices deployed in the field. This is done in two stages of analysis. On chip analytics collects and contextualizes the sensor data into a set of parameters that reflect various aspects of the chip operation. At this level, side channel attacks can be detected if certain sensor data clearly falls outside of the built model. In case a security breach is not detected at this level, cloud analytics can provide second-level detection. On-chip analytics data and parameters are sent to a central database in the cloud from all deployed devices. Then, these parameters can be compared and tabulated or graphed against a large quantity of data. At this level, outliers in any set of parameters can be detected and proper actions taken. For example, if one device's power profile stands out, or the number of internal reset sequences are out of the norm, then maybe a reset request is sent to that that device as a mitigation against a potential security breach.

VI. Concusion

Security breach detection is becoming ever more important as critical information is stored and generated on electronic devices embedded in our daily lives. In the "everything connected" world that we live in, a compromise in security of one device can open the door to access a network of devices and servers that control our finances, personal information, and national security. Investment in electronic device security is one of the highest priorities for many industries. A new generation of sensors, monitors, and analytics provides a window to silicon behavior that can take silicon security to the next level. Nefarious actors are very creative in devising new means of breaking into secure networks. Silicon manufacturers need to stay ahead of the game to keep these players at bay. Silicon lifecycle management solutions such as the Synopsys SLM family of products provide a large portfolio of hardware, software, and tools that enable software and hardware designers to build and protect secure systems.

Future improvements include improved speed and accuracy of measurements, and the creation of new sensor techniques and types to cover specific security risks. Employing machine-learning algorithms is also sure to spawn specific risk mitigation techniques that can then be focused to solve a concrete issue in security. In addition, by incorporating physical-aware design views and methodologies into electronic design automation tools, the time-to-results for implementation flows incorporating security, as well as safety and test functionalities and infrastructure, will improve dramatically.

References

[1] T. Kogan et al., "Advanced functional safety mechanisms for embedded memories and IPs in automotive SoCs," 2017 IEEE International Test Conference (ITC), 2017, pp. 1-6, doi: 10.1109/TEST.2017.8242046.

[2] C. Eychenne and Y. Zorian, "An effective functional safety infrastructure for system-on-chips," 2017 IEEE 23rd International Symposium on On-Line Testing and Robust System Design (IOLTS), 2017, pp. 63-66, doi: 10.1109/IOLTS.2017.8046235.

[3] https://semiengineering.com/toward-on-chip-monitoring/

[4] https://semiengineering.com/benefits-of-in-chip-thermal-sensing/

[5] https://semiengineering.com/is-dvfs-worth-the-effort/

[6] https://semiengineering.com/if-these-chips-could-talk-actionable-insights-from-path-margin-monitors/

[7] https://semiengineering.com/monitoring-for-in-die-process-speed-detection/

[8] W. He, J. Breier, S. Bhasin, N. Miura and M. Nagata, "Ring Oscillator under Laser: Potential of PLL-based Countermeasure against Laser Fault Injection," 2016 Workshop on Fault Diagnosis and Tolerance in Cryptography (FDTC), 2016, pp. 102-113, doi: 10.1109/FDTC.2016.13.

[9] B. Ning and Q. Liu, "Modeling and Efficiency Analysis of Clock Glitch Fault Injection Attack," 2018 Asian Hardware Oriented Security and Trust Symposium (AsianHOST), 2018, pp. 13-18, doi: 10.1109/AsianHOST.2018.8607175.

[10] https://semiengineering.com/in-chip-sensing-and-pvt-monitoring-not-just-an-insurance-policy/

[11] https://www.researchgate.net/publication/349011168_Using_Digital_Sensors_to_Leverage_Chips'_Security

[12] M. M. Alam, S. Tajik, F. Ganji, M. Tehranipoor and D. Forte, "RAM-Jam: Remote Temperature and Voltage Fault Attack on FPGAs using Memory Collisions," 2019 Workshop on Fault Diagnosis and Tolerance in Cryptography (FDTC), 2019, pp. 48-55, doi: 10.1109/FDTC.2019.00015

[13] https://iacr.org/cryptodb/data/paper.php?pubkey=29258

[14] O. Bittner, T. Krachenfels, A. Galauner and J. -P. Seifert, "The Forgotten Threat of Voltage Glitching: A Case Study on Nvidia Tegra X2 SoCs," 2021 Workshop on Fault Detection and Tolerance in Cryptography (FDTC), 2021, pp. 86-97, doi: 10.1109/FDTC53659.2021.00021

[15] https://www.darkreading.com/edge-articles/glitching-the-hardware-attack-that-can-disrupt-secure-software

[16] https://semiengineering.com/debug-and-traceability-of-mcms-and-chiplets-in-the-manufacturing-test-process/

About Synopsys

Founded in 1986 in North Carolina, USA, Synopsys is now among the "Top 15" largest software companies in the world and a world leader in the areas of Electronic Design Automation (EDA), Technology Computer Aided Design (TCAD), and Software Quality, Integrity, and Security (APPSEC) tools and services. Headquartered in Mountain View, California, Synopsys employs over 19,000 engineering and support staff around the world.