Chapter 2
Applications and Technology Trend in Embedded Flash Memory

Hideto Hidaka

2.1 Embedded Flash Memory in MCU History

Compared with a stand-alone memory device, embedded memory is defined as a memory function with a memory interface embedded inside a chip and not exposed to the chip interface. In the present market, embedded SRAM is extensively utilized for working memory in MCU and SOC applications to comfortably match with the CMOS logic technology environment. In parallel, embedded flash-memory advantages listed in Fig. 2.1 provide critical solutions in embedded-systems designs. These advantageous properties of embedded flash memory come from programmability (for supply-chain innovation), non-volatility (for better power management and low power), and on-chip embeddedness (for compact optimized systems, etc.) and are particularly beneficial for embedded system designs. These three points will be discussed in Sect. 2.1–2.3.

A non-volatile memory in today's industry actually provides a memory to store information for >10 years while the power supply is turned off. Non-volatile memory has been preferred especially in remote local systems in eliminating the demands for power supply and battery back-up for system simplicity and maintainability. It also helps mitigate environmental concerns against the extensive use of batteries. Most of the non-volatile memories now in production are ROM and programmable ROM. A family of ROM technologies—including one-time programmable ROM (OTP ROM), electrically erasable and programmable ROM (EEPROM), and Flash Memory—utilizing charge (electron or hole) storage structure is classified in Fig. 2.2. The term "flash memory" implies fast erasure operations in blocks of memory cells to accelerate the erase operation in bulk. The distinction between programmable ROM and RAM lies in their performance. RAM

H. Hidaka (✉)
Renesas Electronics Corporation, Tokyo, Japan
e-mail: hideto.hidaka.pz@renesas.com

H. Hidaka (ed.), *Embedded Flash Memory for Embedded Systems: Technology, Design for Sub-systems, and Innovations*, Integrated Circuits and Systems,
DOI 10.1007/978-3-319-55306-1_2

- Lower system cost
- Fast system development
- High performance code storage
- Reliability, security, safety
- CMOS-oriented integration
- Flexible selection of memory capacity
- Low power memory interface

Fig. 2.1 Advantages of embedded flash memory in embedded systems

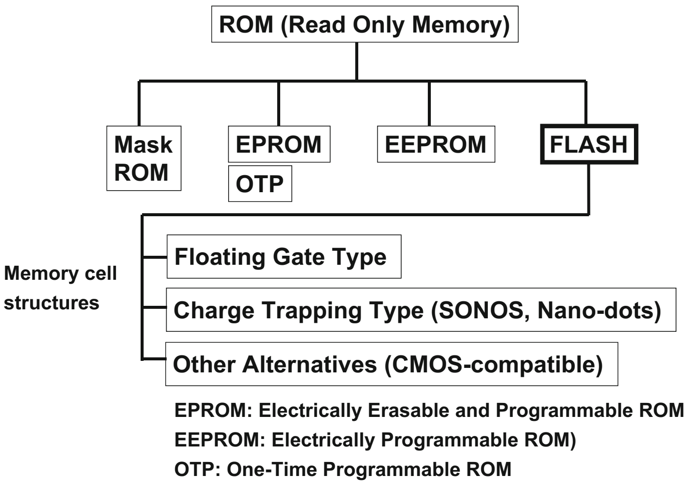


Fig. 2.2 Flash memory and related ROM devices

requires actually 10–20 years lifetime with rewrite capability with $>10^{16}$ cycles, whereas typically flash memories are limited to $<10^{6}$ cycles and much slower write operations (2–4 orders in time) because of the physical limitation in the SiO_x/SiN_y system.

In the history every ROM species listed in Fig. 2.3 has been embedded in the real products, initially for program updates after software development by OTP, and finally into the overall cost reduction by flash memory. Now flash memory has the second largest market size among embedded memory species, following only embedded SRAM. Presently the major application of embedded flash memory is for program-code storage in micro-controller (MCU) products, called "flash-MCU." Embedding flash memory on an MCU chip not only enhanced the values by embeddedness but innovated the supply chain to reduce the total cost of delivery.

MCU is an essential part in many of embedded systems to provide optimized compact system designs with real-time computing tasks by single-purpose embedded software implementation. Figure 2.3 describes the evolution of MCU products with regard to the use of embedded memory. After one-chip integration in 1970s, the embedded memory for program code storage in MCU evolved

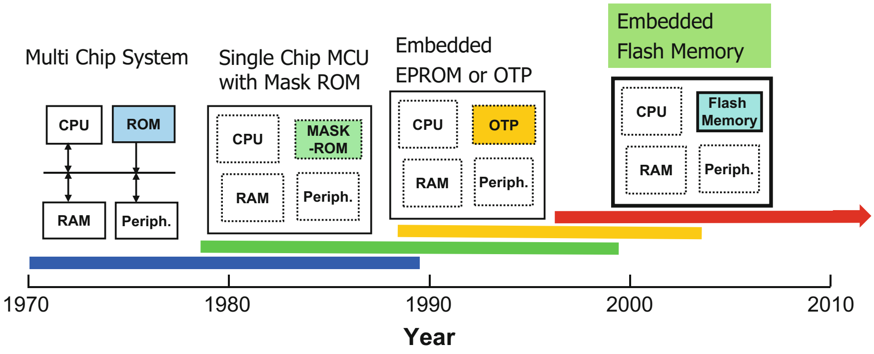


Fig. 2.3 Evolution of MCU by on-chip memory [1]

significantly every decade from mask ROM, OTP (one-time programmable ROM), and finally flash memory. In the 1990s, there were strong demands for applying the flash memory in MCU to save total cost and turn-around time of MCU system development. After the introduction of embedded flash memory in MCU, flash memory has improved its features such as optimized ROM features, single power-supply operation, and better reliability. The flash-MCU dominates the worldwide MCU market.

We have seen a great leap in the market penetration of embedded flash memory thanks to the overall cost reduction in the supply chain through design, production, and inventory control simply by programmability in flash-MCU. Here the overall cost advantage by the economy of scale exceeds the disadvantage of higher wafer-process cost for incorporating flash memory, which was a key factor for successful embedded flash-memory market development.

Integrating embedded flash memory in MCU products for programmable-code implementation has a large impact on the design and the supply chain for embedded systems as follows.

(1) Short turn-around time in system development: The embedded mask ROM scheme takes a longer time from the start of software development to the shipping of a system because program modifications and debugging require iterations of mask-ROM revisions to be shared by the system vendors and MCU vendors. In contrast, flash-MCU products are shipped with un-programmed (blank) states to customers during software development. On-chip flash memory is programmed after system assembly; hence, the developments of total system and program can proceed in parallel, thus resulting in a shortened turn-around time of system development. And the recovery time after debugging becomes shorter by recovering programs in the field. Therefore, the total cost and time of system development are reduced.

(2) Simplified and flexible production and inventory control: Variation of the final product sets is often offered by different firm-wares in the MCU market. If mask ROM MCU is used, each set must be manufactured with each different mask ROM data for MCU, which causes complicated inventory management. Flash MCU simplifies this situation by programming the MCU just before the set assembly, thus reducing the inventory cost. Likewise, at every step of the supply chain the inventory control with respect to different mask ROM data (for different MCU product models) is eliminated, thus causing a supply-chain innovation (Fig. 2.4).
(3) HW-SW separation and co-design: Flash-MCU has successfully separated HW and SW in the embedded system design where SW development is conducted independent of HW designs. In such a system, a further advantage emerges by HW-SW co-designs. For example, a lower-power system is realized by SW-programmable power management by programmable powering-up sequence compositions and use-case adaptable control of power-delivery schemes. The potential capability of software will advance the control scheme and create new values.
(4) Paradigm shift in software development by learning mechanism: In embedded-systems development by single-chip MCU, hardware control has been replaced by software control, and the product lineup is unified in hardware. Here a new problem has arisen in the turn-around time for software development becoming longer according to complex system requirements.

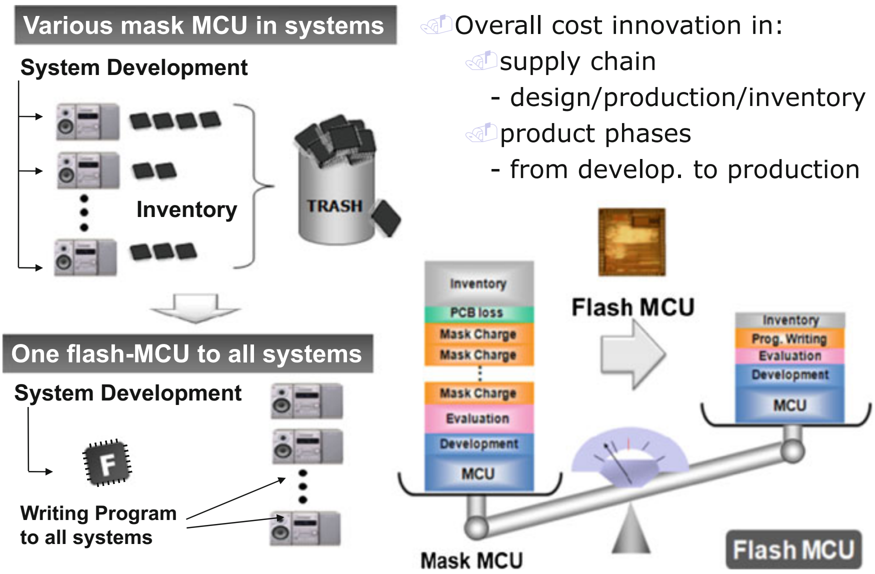


Fig. 2.4 Delivery-cost reduction in supply-chain innovation by flash-MCU

Flash MCU can alleviate some of this problem by introducing learning mechanisms for parameter updates in the system, thus providing new system values.

(5) Upgradable system for lifetime maintenance: A SW-upgradable system and lifetime maintenance scheme is enabled, thus providing another innovation in the supply chain. Especially the in-field SW update by over-the air (OTA) will contribute to actualizing such business models.

Embedded flash memory in MCU has achieved an innovation for increasing supply and demand based on the overall cost reduction in the supply chain and value creation for embedded-systems development. This is the basic rationale of embedded flash memory in MCU products for embedded systems. Technologically this hits the balance between application-specific optimization and minimum cost by programmability.

Together with the eFlash innovation in the basic cost structure previously described, finding new uses of eFlash—as well as developing eFlash technology and sub-system designs to form a standard and to meet new demands—has expanded the MCU market significantly. Figure 2.5 lists meaningful uses of eFlash in embedded systems. These are guiding concepts for eFlash applications in embedded-systems designs.

In addition to achieving values in the non-volatile storage of code and data, eFlash opened the door for design reform in wide levels of HW-SW co-designs to augment design efficiency in small embedded systems, security applications by using the hidden data on the chip, and upgradable VLSI functions. These are items to develop and fully utilize in future cyber-physical system (CPS) and IoT system designs. Finding new uses and applications of eFlash is still underway in every aspect of embedded-systems designs.

Fig. 2.5 Embedded flash-memory uses

- **Upgradable program codes**
  - **Flash innovation; cost reduction in supply chain**
  - **System boot, power on/off sequence etc.**
  - **OTA for maintenance &system update**
- **Data storage**
  - **Parameter settings - data logging**
- **Selectable functions, re-configurable**
  - **FPGA configuration - Function switches**
  - **Upgradable VLSI functionality**
- **Security - Secure data/key storage**
- **Design flexibility in development**
  - **HW-SW partitioning and co-design**

2.2 Expanding Applications of Embedded Flash Memory

Unifying the technology and design into an established standard of flash-MCU configurations has contributed to widening the applications of eFlash. In addition to the program-code storage by embedded-systems program sizes (up to multiple megabytes), in many cases of embedded-systems designs small-capacity data storage (up to 32 kB) with multiple rewrite capability (100 K–1 M cycles) to replace the externally provided EEPROM is required. This one-chip code-data storage configuration of flash-MCU (Fig. 2.6) has been established to help increase the market volume effectively. This is supported by a reasonable combination of on-chip integration utilizing the same memory cell technology for both the code and data flash memory on the chip, which has been realized by technology development.

2.2.1 Standards for General Purpose MCU

In the typical flash-MCU architecture depicted in Fig. 2.7, the on-chip flash memory first merged code ROM with data ROM, thus providing a standard small-system configuration. In parallel, by developing applications for multiple-time programmability in real-time control and achieving an overall cost advantage, all the multi-billion dollar MCU markets now focus on flash-MCU solutions. Among them major segments of the market that have increased the market size significantly are automotive control and smart-card applications.

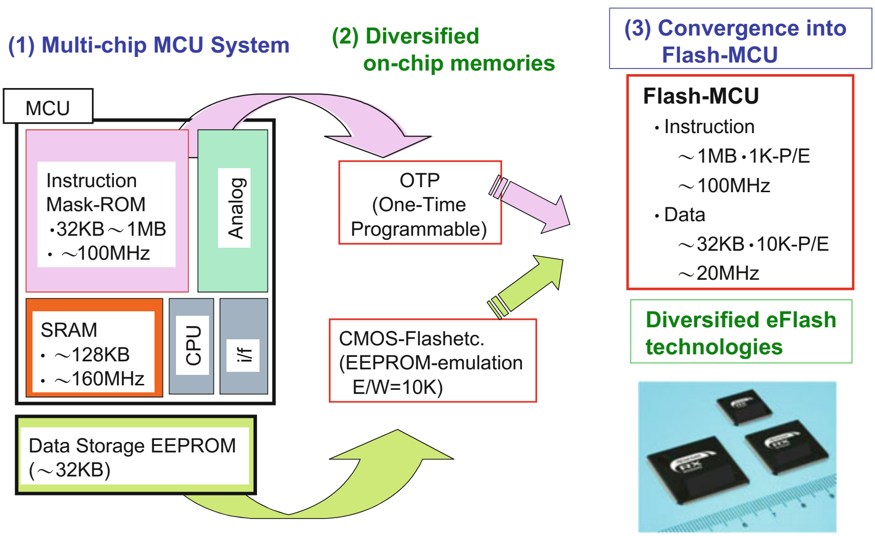


Fig. 2.6 Convergence into a standard flash-MCU configuration

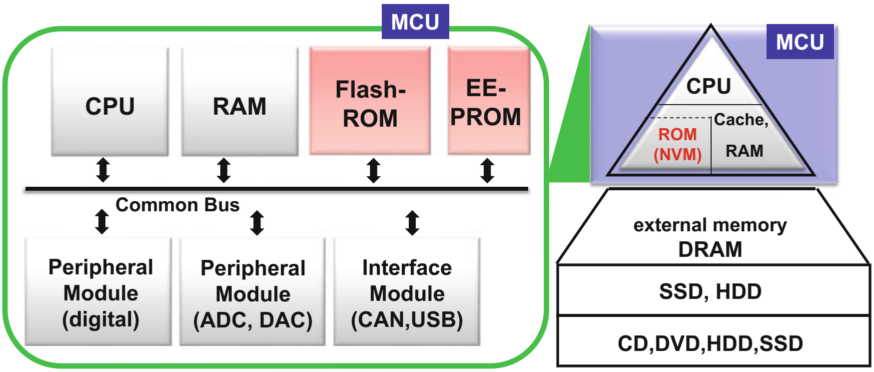


Fig. 2.7 Non-volatile memory in MCU architecture

Another important contributor to product convergence is the development of embedded-specific flash-memory technology attaining read performance competitive to mask-ROM configuration and high-temperature operation together with high reliability up to the automotive grade. The process and device technology for large-capacity stand-alone data memory such as NAND flash memory does not meet the requirements by embedded applications in read speed and reliability. These two distinct technology sets are shown in Fig. 2.8.

Accordingly, the technology map in Fig. 2.9 shows the discrepancy between stand-alone NAND flash memory and embedded flash memory in read and write performances due to requirements for fast data write in data-oriented stand-alone large-capacity memory and fast code read in embedded applications without frequent write.

The benefits of embedded flash memory come from its embedded nature, programmability, and non-volatility, as follows, compared with probable alternatives.

(1) Because in the embedded environment the off-chip memory access path and drivers are eliminated, high-speed and low-power operations are achieved, and data security is easy to implement with a less-accessible internal data bus. In addition, high-density physical packaging, enhanced reliability, reduced EMI, and lower system cost are provided by the embedded memory solutions.
It is to be noticed that 2.5D and 3D integration by stacked chips using TSV (Through Silicon Via) will provide alternatives approaching to embedded memory performance and power in large-capacity memory integrations.

(2) The flexibility of design in the embedded environment gives optimal design in memory capacity, memory interface configuration, functions, operating voltage, etc., for each application showing a tough challenge in optimizing and controlling the eFlash macro variations. Thus comes the eFlash compiler and technology/design platform approaches, now in wide usage in the industry.
An increasing number of function requirements in terms of safety, security, memory interface, etc., poses a new challenge in controlling the design and verification of the eFlash sub-system. High-level design and verification

	NAND Flash	Embedded Flash
Major Use	Data storage media	Code storage for real-time apps.
Memory Size	~16GByte	~16MByte
Operating Temp.	up to Tj=85˚C	Auto: Tj=150 ~ 170˚C Industry: Tj=125˚C
Data Access	Sequential	Random
Random Read	~1MHz	≥ 100MHz
Data reliability	Need strong ECC (Long Delay time)	SEC/DED at most (Speed Constraint)
Compatibility w/ High Speed Logic CMOS	Not required	Mandatory

SEC : Single Error Correction
DED : Double Error Detection

Fig. 2.8 Embedded flash versus stand-alone NAND flash specifications [2]

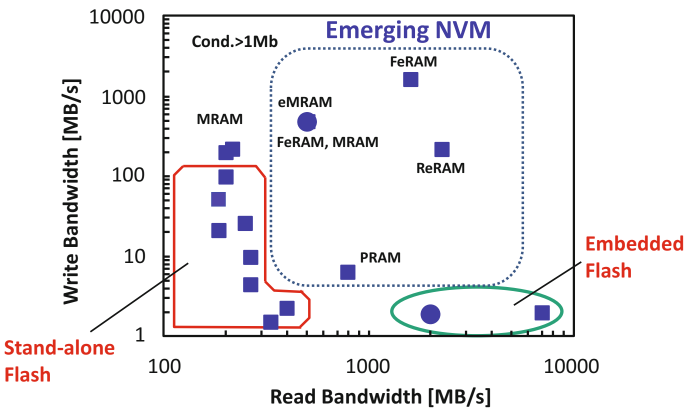


Fig. 2.9 Comparison of stand-alone and embedded flashes [3]

methodologies, such as HW-SW co-designs, will be essential in productivity management.

(3) The non-volatility attribute gives an opportunity for low-power design such as eliminating the battery-backup SRAM and achieving zero stand-by power by intermittent system operations. Attentions are focused on zero stand-by power memory because today's scaled embedded SRAM often sees a large stand-by current due to the scaled MOS FET in the memory cell.

Frequent backup of memory data to non-volatile storage necessitates great endurance as well as fast- and low-power rewrite operations at the flash memory. Emerging non-volatile memory properties are favored in intermittent operations.

(4) Field programmability will come into a new age of wireless-communicated program updates called OTA. Data-security functions and remedies for EMI reduction [2] will play important roles in the OTA era.

In contrast, the intrinsic disadvantages of embedded flash memory in terms of re-write power, large mask counts in production process, compatibility with CMOS logic process, and scalability are looming in some applications as is described in Sect. 2.3.

As listed in Fig. 2.10, embedded flash memory provides functions of code and data storage, backup storage, system boot, and trimming-information storage for memory and analog parts of the chip, etc. By applications we find code and data storage dominant in micro-controller unit (MCU), updatable coefficient parameter storage in DSP (digital-signal processor), data storage in smart IC cards and RF-D tags, and configuration storage in field-programmable gate array (FPGA) and re-configurable logic, etc. By uses in the stages from proto-typing to volume production, embedded flash memory acts as easy-to-change ROM storage for easily verifying system concepts, fixing program bugs, and supporting program updates as well as production and inventory control by unified product lineup.

Expanding applied products benefiting from previously mentioned standardization and convergence of eFlash configurations and benefits include the following:

– Micro-controller unit (MCU) for code and data storage
– Smart IC cards and RFID for secure data and key storage

● **By functions**
- **Code storage; system boot, user program, firmware, look-up table**
- **Data storage; EEPROM emulation, shadow storage, frequently updated parameters and coefficients, state before power down**
- **Add-on parameter storage for trimming information etc.**

● **By applications**
- **MCU (micro controller unit); for code and data storage**
- **DSP (digital signal processor); for coefficient storage**
- **Smart-IC cards ;EEPROM data storage**
- **RF-ID ; data storage**
- **Reconfiguration register; FPGA(Field Programmable Gate Array) etc.**

● **By uses**
- **proto-typing ; verify system concepts**
- **system development ; program debug and updates**
- **early productions ; program updates**
- **volume productions ; production and inventory control**
- **lifetime control of VLSI**

Fig. 2.10 Embedded flash-memory applications

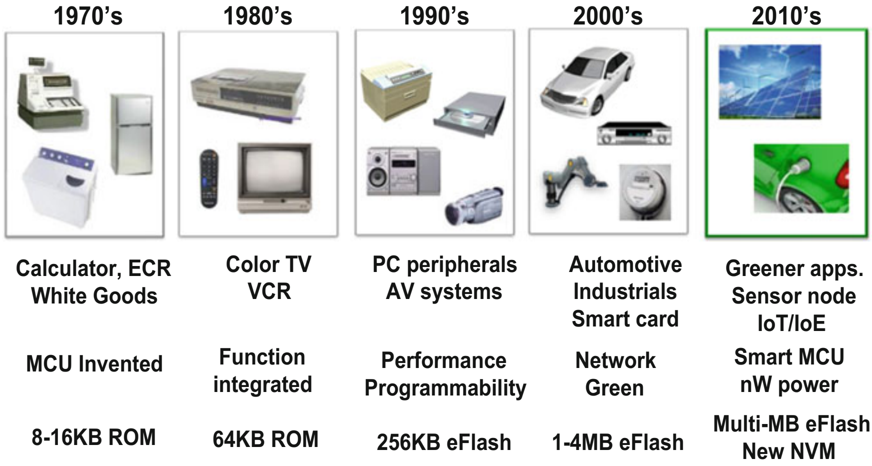


Fig. 2.11 MCU market expanded by new applications

– Analog, power, FPGA, and SoC for add-on uses for power management, analog/RF circuit tuning, memory-redundancy program, re-configurations, etc. [4, 5].

In these established fields of applied products, demands for eFlash are increasing based on the growing number of software-upgradable engine control systems for minute controls of combustion as required by CO_x regulations, over-the-air (OTA) program upgrades, calibrations of analog circuitry, power management with programmable power-up sequences, security functions such as certifications, and learning functions for artificial intelligence processing, etc (Fig. 2.11).

2.2.2 *Automotive Applications*

In Sect. 2.2.2 and 2.2.3, representative flash-MCU applications in automotives and smart cards, which have significantly expanded the flash-MCU market since the 2000s, are described to illustrate how applications have driven the development of embedded flash-memory technology.

Automotive applications of flash-MCU illustrated in Fig. 2.12 indicate that today's electrically equipped car extensively uses MCUs for local real-time electrical control functions for engine control, body/chassis control, and peripheral functions, most of which employ embedded flash memory for locally storing the control programs, control parameters, and measured data. In modern automotive engine control, several sensors for crank angle, air-flow, and knocking phenomenon are connected to flash-MCU through application-specific ICs for pre-processing the sensed data. According to measured data from these sensors, the MCU gives feedback to control fuel injection, timing of the ignition plug and throttle motor,

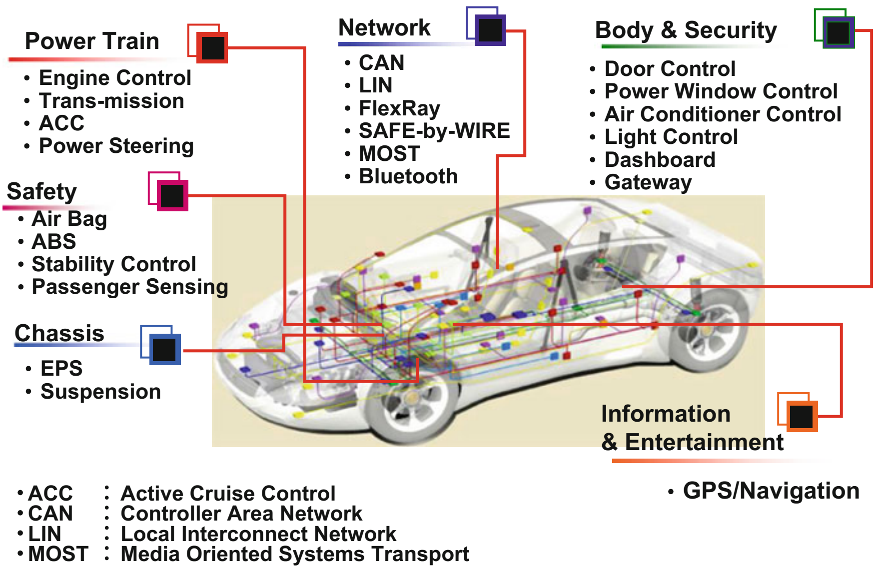


Fig. 2.12 Flash-MCU applications in a car

etc., to maintain engine operation under the most suitable conditions. Distributed real-time control like this in many embedded systems dominates MCU applications.

The trend in today's automotive control is strongly influenced by enhancing the combustion efficiency for environmental regulations, safety and security requirements, and getting more connective to outer information. These demands by new functions in a car add to uses of flash-MCU resulting in ≥ 200 units of flash-MCUs being used in a single car in some cases. By introducing electrical control in every part of the automotive control, even in combustion engine cars, and by introducing the IT region, the number of semiconductor parts used in a car amounts to 1–10 K. The electric car will use vast variety of flash-MCUs more extensively for motor control and associated power electronics.

One significant trend in modern cars is a merger of a conventional physical control regime for the engine and chassis and a newly introduced IT regime to connect with the outer cyber world to provide a higher level control and/or additional values in cars. Thus, "control meets IT" in a car (Fig. 2.13) where we are beginning to find a new space of cyber-physical systems to realize cognitive perception and automotive control [6]. Together with these functional advancements, once non-functional requirements such as safety and security have become essential elements in these embedded system designs as the system becomes more integrated and more connected. The Automotive Safety Integrity Level (ASIL) by ISO-26262 defines the safety and security requirements in automotive systems. ASIL-compliant MCU and SOC with secure IPs are becoming indispensable parts of automotive systems.

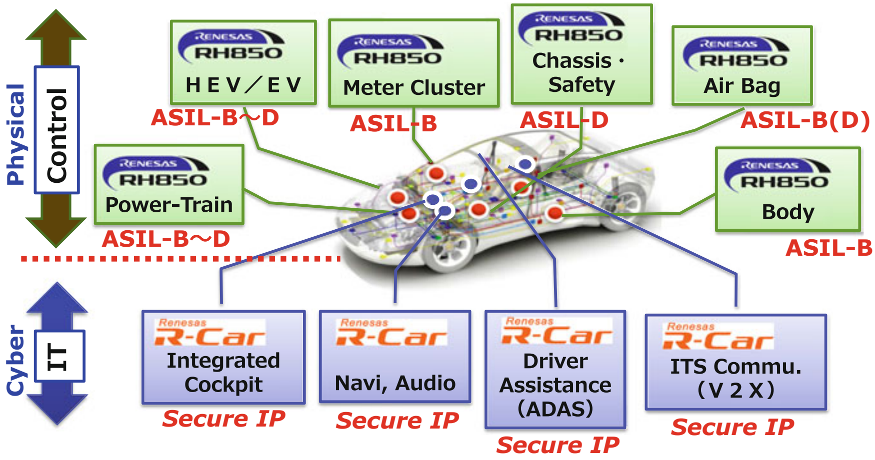


Fig. 2.13 MCUs and SOCs in a car with ASIL compliance and security

The performance and capacity trend of on-chip flash memory for automotive MCUs in Fig. 2.14 indicates ever-expanding technology requirements [2]. The overall MCU performance requirement has grown approximately by 20-fold in 10 years, 35% per year, which is supported by architecture evolution such as cache-memory usage and multi-core CPU implementation, eFlash speed enhancement, device scaling, design for reliability, etc. In parallel, the on-chip ROM capacity has grown by 23% per year to support the growth of on-chip code storage for

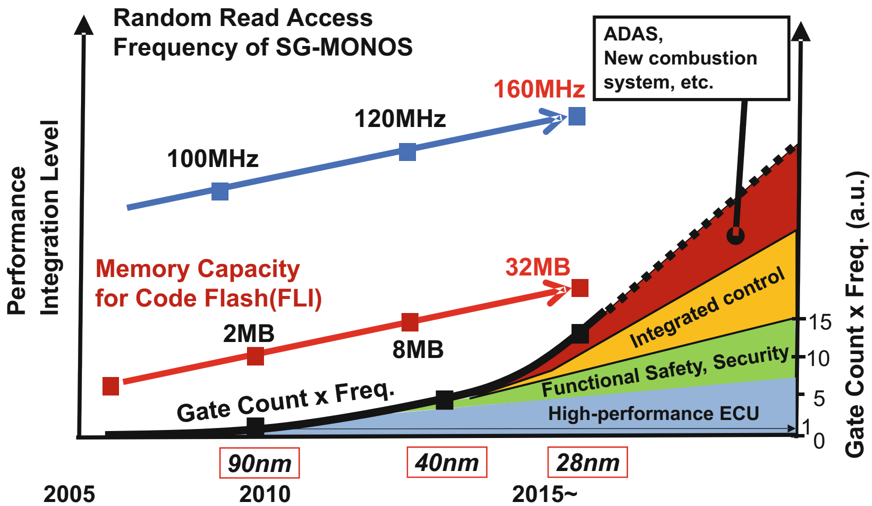


Fig. 2.14 Performance and ROM capacity of on-chip eFlash for automotive applications [2]

Fig. 2.15 Requirements for embedded flash memory in automotive uses

- **Memory band-width adapts to CPU speed**
- **Low power & low voltage operations**
- **Automotive-grade temperature operation and reliability**
  -40 to 150 C with very low failure rate
- **Offer overall supply chain cost lower than mask-ROM**
- **Safety and security functions**
- **CMOS logic embeddable technology**

program statements in automotive applications. In reality, MCU for automotive applications drives the development of scaled CMOS process integrated with embedded flash memory to achieve high-density memory as well as high-performance CMOS logic.

Examples of embedded flash-memory requirements for automotive applications in Fig. 2.15 reveal the actual challenges facing embedded flash-memory technology. From performance to match CPU to reliability under a very wide temperature range and low cost, embedded flash memory for MCU becomes the most challenging one in many aspects of today's semiconductor memory. In addition, safety and data security functions are becoming prevailing factors in the general requirements for MCU.

2.2.3 *Smart-Card Applications*

Smart cards present another application area where the capability of Flash-MCU has greatly expanded. Figure 2.16 shows an example of smart-card application in charging and paying for public transportation based on contactless card storage.

Fig. 2.16 Smart-card and applied systems [7]

Used in many everyday chip-card applications for money transactions, IDs, and data/configuration transport and storage in mobile phones, specially designed one-chip MCU called “secure-MCU” incorporates one-chip MCU with embedded EEPROM storage enclosed in logical/physical security protections. It is based on a security-intensive EEPROM with CPU or accelerator to process encryption/decryption in cryptography standards such as AES and RSA.

Embeddedness provides an advantage in data security because the memory interface is embedded on the chip, and tamper-resistant measures are easily implemented over the chip including the memory. The highest requirement for program/erase cycle is approximately 500 K-cycles, much higher than normal flash memory storage applications, because in many cases at every transaction the stored data as well as stored security key data are updated for higher security level. Thus, encrypted data and associated security keys are stored in a non-volatile memory array enclosed by security measures both logically and physically.

In the actual smart-card implementation (Fig. 2.17), near-field communication (NFC) is often adopted to provide the communication as well as power supply through a short-range wireless connection to the outer environment. Because the power-supply capacity by NFC is very limited, the current consumption by secure-MCU is required at <2 mA (max). Thus, peak power consumption is limited particularly in program/erase operations of the embedded-flash ROM.

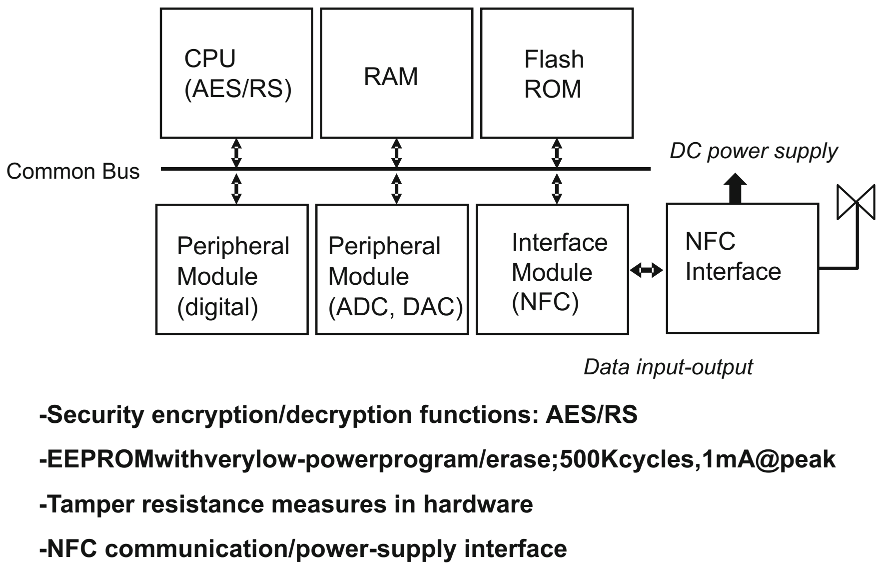


Fig. 2.17 Secure-MCU organization for contactless smart card

2.2.4 Summary of Product Requirements

Figure 2.18 summarizes the maximum requirements for embedded flash memory in various MCU applications. From high performance under high temperature, as required in stringent automotive and industry conditions, to macro-area efficiency for small memory capacity for cost-sensitive applications in PC/OA and for the consumer, the requirements vary significantly. Particularly varied are the requirements for the memory capacity at 16 KB–8 MB, program/erase endurance at 1K–500 K cycles, and read performance from 10 MHz to 200 MHz inevitably force a number of optimal technology/circuit/sub-system designs to best fit the varied market. Even only in one technology we must offer a wide variety of ROM/RAM capacity, which is a basic parameter for MCU product line-ups to be efficiently deployed in a platform manner.

In Sect. 2.1 and 2.2, we have described how the favorable properties of embedded flash memory—such as embeddedness, programmability, and non-volatility—have contributed to innovative embedded-systems solutions. From prototyping uses to overall cost reduction for code storage in MCU, real-time control parameter updates, and security data memory, the Flash-MCU market has been expanding steadily. It also has found new market drivers in automobiles and smart cards. All together, flash-MCU has won one of the most successful businesses in embedded memory applications, second only to embedded SRAM in terms of market size. In the next Sect. 2.3, the future prospects of eFlash and new demands to be met by emerging non-volatile memory in the future will be overviewed.

		Automobile		Industry	PC/OA	Consumer	Secure-MCU
		Power Train	Body				
MCU	Performance (max. freq.)	~300MHz	150~ 200MHz	~300MHz	25~ 50MHz	20~ 100MHz	10~ 50MHz
	Power	0.5mA /MHz	0.5mA /MHz	1mA /MHz	0.5mA /MHz	0.25mA /MHz	0.1mA /MHz
	Temp.(Ta)	- 40 ~ 125 C		max 85 C	- 20 ~ 85 C		
FLASH	Density (max.)	8MB	2MB	1MB	2MB	1MB	1MB
	P/E cycle	Program Area : 1K / Data Area : 100K					100K-500K (EEPROM)
	Small Cell	✓					
	Small Macro				✓	✓	✓
	Fast Access	✓		✓			

Fig. 2.18 Requirements for flash-MCU technology for different applications with important factors in each application marked

2.3 Challenges and Prospects of Embedded Flash Memory in Embedded Systems

We have observed the major trend in embedded flash-memory technology for MCU applications as driven by current application requirements:

(1) Convergence into standard flash-MCU and add-on solutions.

- code/data combination in design platform
- add-on eFlash solutions.

(2) Multiple eFlash technologies are converging into selected species for device scalability and for low-cost eFlash macro, seeking for:.

- basic memory cell scalability
- integration compatible with advanced CMOS such as high-k metal gate, FD-SOI, and FinFET.

(3) Embedded flash memory design incorporates multiple functions:

- low-power compact eFlash macro
- byte-wide access, data security and safety functions, EMI reduction, etc.

Although the MCU market has converged into standard designs in flash-MCU solutions, other market segments—such as SoC, analog, and power products—are beginning to incorporate simple and small-capacity add-on flash memory for chip-ID and security functions etc. [4, 5]. Further convergence into scalable technologies and new standards for on-chip IP designs are now under investigations.

In contrast, some recent requirements by application trends exceed the intrinsic limitations of flash-memory properties.

(1) Slow and power-consuming in program/erase operation with limited numbers of endurance, flash memory is not suitable for applications with frequent data re-write such as driving recorder applications etc.
(2) Embedded flash-memory process costs relatively high in CMOS integrations and is often not compatible with underlying CMOS technology.

Hereafter in Sect. 2.3 some of future requirements to eFlash are discussed—specifically regarding functionality limitations, scalability, power consumption, and future system requirements—in Cyber-Physical System (CPS) designs.

Flash-memory technologies for embedded uses have come to deviate from stand-alone NAND flash-memory technology significantly (Figs. 2.8 and 2.9) because of the embedded-specific requirements such as host-logic CMOS compatibility, performance, lower cost with smaller capacity of on-chip memory, and reliability quite different from those for stand-alone data flash memory. For instance, multi-level storage, now commonly used in NAND flash memory, does not offer reasonable embedded solutions because the data reliability is degraded, and timing/area-consuming strong ECC cannot be accommodated in embedded uses for low-cost, small-capacity memory solutions. Diversified flash memory technology may require universal technology solutions arising from emerging non-volatile technologies as shown in Fig. 2.19. These candidates include ferro-electric polarization (FeRAM), magnetization orientation (MRAM), amorphous/crystalline phase change (PCRAM), filament formation by atomic motion (ReRAM), carbon-nanotube switch (NRAM), etc. In terms of functionality, the device properties of emerging non-volatile memories fill the gap between current ROM and RAM as well as that between stand-alone and embedded flash memories, thus not only partly unifying the technology but also ushering in new memory applications.

The scalability issue in eFlash is mostly attributed to the compatibility with the underlying advanced CMOS structure. New structures—such as high-k metal gate, FD-SOI, and FinFET—will have an impact on eFlash structures and CMOS compatibility. Thin-film storage structures—such as SONOS, nano-dot, and thinned floating gate—seem advantageous in technology nodes beyond 28 nm. In contrast, back-end integration in emerging memory, such as MRAM and resistive-oxide-RAM, will make the structure almost free from the underlying CMOS device structures and will find advantages if they realize the resistance against the thermal treatment by metal/inter-dielectric layer formations.

Significantly some of the emerging memories offer much improved performance over conventional flash memory in re-write operations as shown in Fig. 2.20. Achieving faster speed by 2–4 orders under lower voltage in re-write operations, these provide 2–3 orders lower energy than conventional flash memory. In addition,

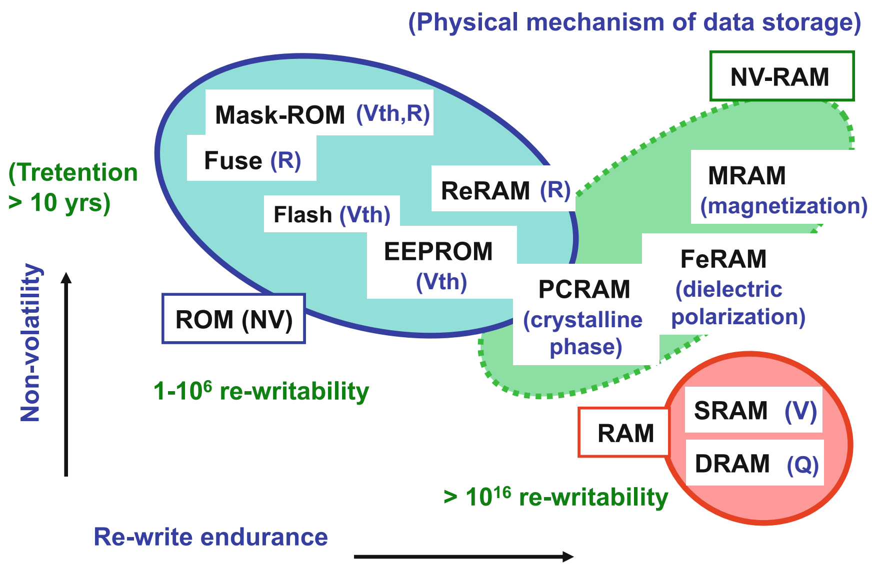


Fig. 2.19 ROM and RAM mapped with emerging non-volatile memory

they may provide lower cost because of simpler device structures. A key question remains as to whether these new physical memories will come into reality for embedded uses led by major innovation items like flash-MCU innovation in the IoT/IoE and AI era.

Memory hierarchy was originally intended for memory system cost and performance optimization. Recently it is seeking for lower-power systems, which has attracted great attention in computing system designs. In lower-tier memory hierarchy with almost-always stand-by state, non-volatile memory contributes to low stand-by power, and it can be also used in higher-tier memory hierarchies with intermittent power control for power-saving as shown in Fig. 2.21. Power-constrained designs in embedded systems for smaller-form factors and longer battery life in IoT-sensing nodes are expected to greatly benefit from this

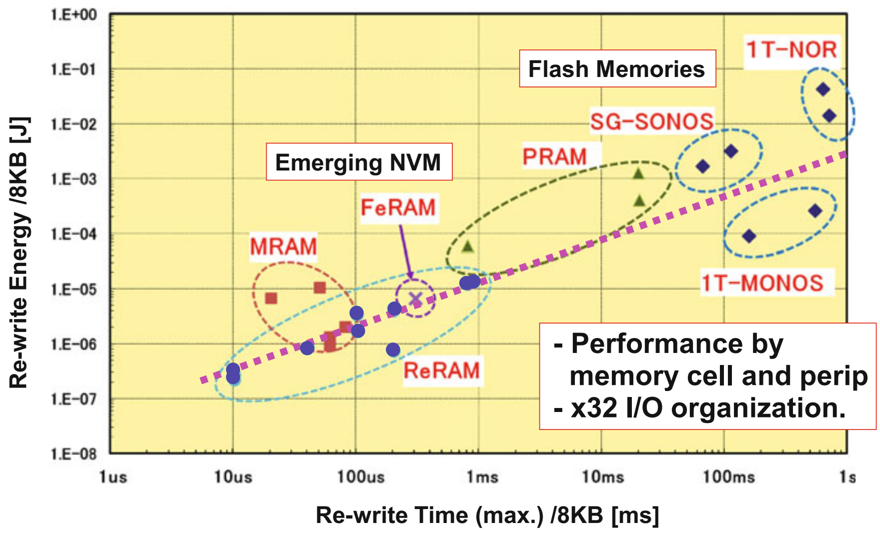
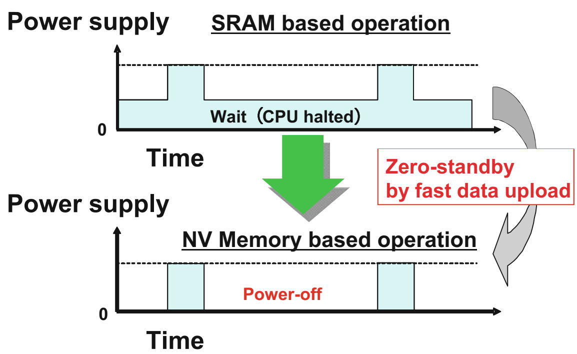


Fig. 2.20 Re-write energy in flash and other non-volatile memories [1]

Fig. 2.21 Power-saving by intermittent operations



dynamic power scaling by intermittent power control. In terms of cost, CMOS-compatibility, and re-write performance (speed, power, and re-write endurance), emerging non-volatile memories provide attractive solutions in future highly-integrated systems.

In modern moving apparatus—such as automobiles, drones, etc.—we often see requirements for merged control and information systems, i.e., cyber-physical systems (CPS). CPS designs have several essential points in common to many fields of applications such as automotive, industry, and IoT/IoE. Emerging mobility requirements in ADAS include safety, security, connectivity, low power with advanced sensing and feedback capability, which have impacts on the requirements for eFlash [6]. Together with conventional sensor-to-reaction feedback, we see a vast range of feedback control in multiple-path feedback loops as depicted in Fig. 2.22. From short-term fast memory to long-term, low-power memory adapting to learning mechanisms, embedded non-volatile memory will contribute to many aspects of CPS designs. Non-volatile memory will play important roles in every aspect of CPS designs in locally and asynchronously powered distributed systems including quite responsive feedback to long-term, learning-based intelligence.

Figure 2.23 summarizes a history of memory-centric VLSI evolution realized by on-chip non-volatile memory, which covers programmable product functionality and VLSI design methodology to market creation and supply-chain innovation. Future development of applications by static/dynamic power scaling, advanced learning, functional safety, security, OTA, and lifetime management of VLSI systems will certainly provide the source of next innovation in the IoT/IoE era.

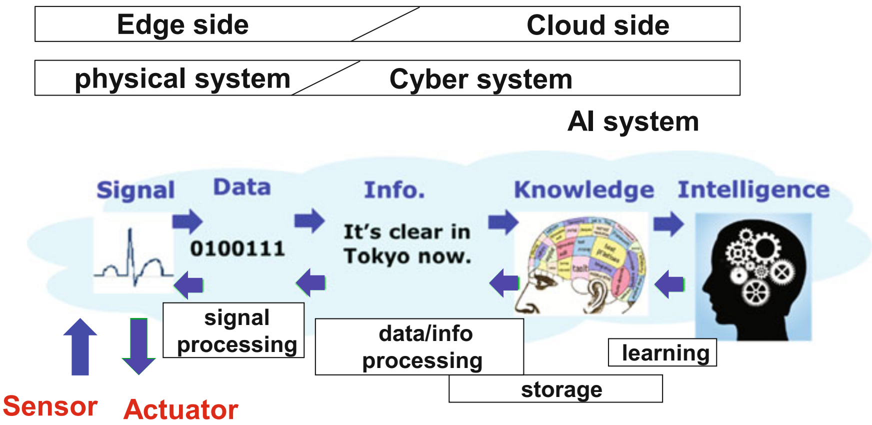


Fig. 2.22 Multiple-path feedback loops in modern control systems

	Innovation	Main Enabler	Product	Effect
1	Memory-based Logic	ROM Program Register-based computing	MPU, MCU	Program-driven Logic general-purpose logic
2	Alterable Program	SRAM/Flash	Flash-MCU FPGA	Re-configurable Supply chain innovation HW-SW co-design
3	Quasi-NVRAM Universal Mem.	NVM w/i energy-efficient re-write: (FeRAM, MRAM, ReRAM, NRAM…)	Unified-Memory/ MCU, SOC	Intermittent op. Re-usable logic Life-time management Learning, LP, FUSA/Security, OTA

Fig. 2.23 On-chip non-volatile memory triggers VLSI innovation [1]

2.4 Conclusion

In this chapter, we reviewed the history of embedded flash memory in MCU products in terms of supply-chain innovation, technology development, standardized products, expanding markets, and prospects for future technology and applications.

Embedded flash-memory technology originally inherited much from stand-alone flash-memory devices such as NOR flash memory. However, it has deviated from the main-stream stand-alone data flash-memory technology such as NAND flash memory technology due to embedded-specific requirements, and has advanced in its own path of technology development. Together with the standardization of specifications and configurations, dedicated technology has driven application development and market creation by making embedded flash memory adaptable to very broad applications and products in terms of performance, reliability, and overall cost. The rapid market penetration is attributed to the advantage in cost/value achieved by eFlash innovation since the 1990s. Embedded flash-memory technology and MCU market segment creations have advanced and expanded in interactive ways to proliferate by virtue of lower power, compact size, and high-performance integration with high reliability to contribute to embedded-systems advancements, notably in newly created automotive and smart-card applications. They will continue to serve humankind well for many years to come.

In contrast, emerging non-volatile memory technologies have been explored by introducing physical memory principles different from conventional silicon/silicon-oxide systems—such as ferroelectricity, magnetism, phase-change materials, and atomic motions—some of which have been productized in limited volumes. Although they are expected to partly solve flash-specific problems in terms of cost scalability, power consumption, and technological convergence for coming IoT/IoE and advanced mobility age, they have yet to prove real scalability compared with embedded flash memory, the ability to replace embedded flash memory and embedded SRAM in functions and performance, and the creation of new use cases and applications before coming into a "really emerged" species of semiconductor memories. Technologically they still suffer from insufficient properties under higher temperature in data retention and re-write endurance, thus limiting their scalability to a great degree. Basic research and development works are underway for solving these problems toward the next convergence in embedded non-volatile memory technology, and hopefully achieving the next innovation. This is a big challenge for humankind.

References

1. H. Hidaka, "Evolution of embedded flash memory technology for MCU," invited paper in Proceedings of the IEEE international conference on IC design and technology (2011)
2. Y. Taito, T. Kono, M. Nakano, T. Saito, T. Ito, K. Noguchi, H. Hidaka, T. Yamauchi, A 28 nm embedded split-gate MONOS (SG-MONOS) flash macro for automotive achieving 6.4 GB/s read throughput by 200 MHz no-wait read operation and 2.0 MB/s write throughput at T_j of 170 °C. IEEE J. Solid-State Circ **51**(1), 213–221 (2016)
3. Press-kit, *International Solid-State Circuits Conference*, p. 111 (2015), http://isscc.org/doc/2016/ISSCC2016_PressKit.pdf. 5 Nov 2015
4. C.C.-H. Hsu, Y.-T. Lin, E. C.-S. Yang, R. S.-J Shen, Logic non-volatile memory (World Scientific, Singapore, 2014)
5. H. Mitani, K. Matsubara, H. Yoshida, T. Hashimoto, H. Yamakoshi, S. Abe, T. Kono, Y. Taito, T. Ito, T. Kurafuji, K. Noguchi, H. Hidaka, and T. Yamauchi, A 90 nm embedded 1T-MONOS flash macro for automotive applications with 0.07 mJ/8kB rewrite energy and endurance over 100 M cycles under T_j of 175 °C in Technical digest international solid-state circuits conference, pp. 140–142 (2016)
6. H. Hidaka, How future mobility meets IT: Cyber-physical system designs re-visit semiconductor technology, Plenary Talk 3, Proceedings of the Asian solid-state circuits conference (2015)
7. http://www.smartcardalliance.org/smart-cards-intro-primer/. 30 Aug 2016

http://www.springer.com/978-3-319-55305-4

Embedded Flash Memory for Embedded Systems:
Technology, Design for Sub-systems, and Innovations
Hidaka, H. (Ed.)
2018, VIII, 247 p. 262 illus., 192 illus. in color.,
Hardcover
ISBN: 978-3-319-55305-4