

NAND Rises to the Occasion in Data-Heavy IoT Applications

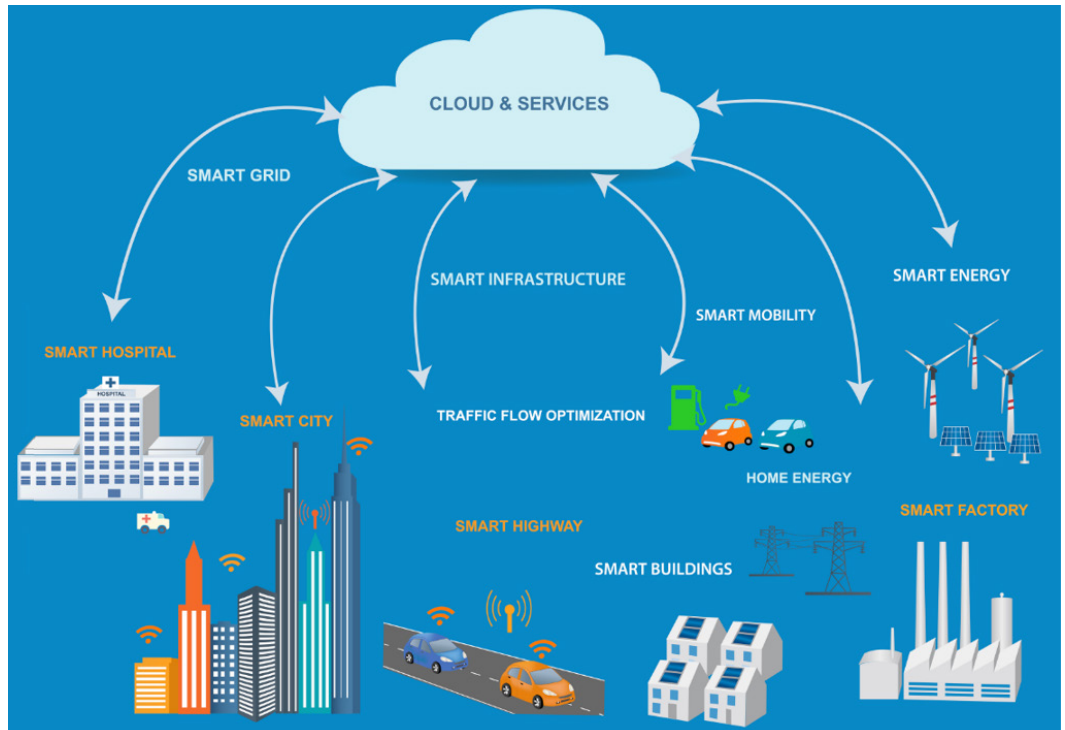
With its emphasis on collecting and analyzing mass quantities of data, the IoT requires fast, flexible non-volatile storage at every level of the hierarchy—NAND flash devices can accommodate such applications from the edge node to the cloud.

Long a staple of consumer devices such as MP3 players, digital cameras, and USB flash drives, nonvolatile NAND flash memory is raising its game to respond to the new challenges posed by data-heavy Internet of Things (IoT) applications (Fig. 1). IoT applications for NAND flash cover a breathtaking range: industrial automation and Factory 4.0, connected vehicles, wearables, healthcare, aviation, plus just about any application that begins with “smart”:

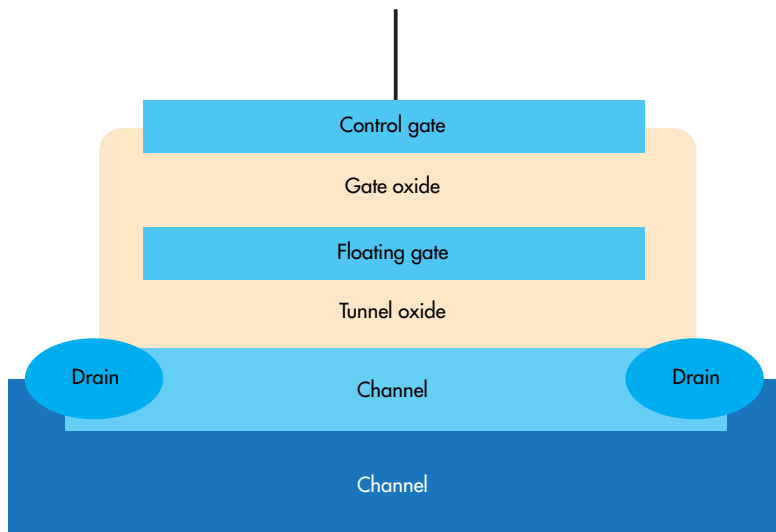
smart metering, smart homes, smart utilities, smart farms, smart logistics, and more.

The smallest IoT applications—a remote edge node, for example—commonly use microcontrollers with embedded flash to store both code and data. Their storage requirements are quite modest, but with 30 billion connected devices forecast by 2020 and 75 billion by 2025, the data load rapidly increases at higher levels of the hierarchical IoT architecture:

1. NAND flash memory finds a home in numerous data-heavy IoT applications. (Source: Delkin Devices)



Floating gate flash memory cell



2. The structure of a typical single-level flash cell. (Source: Tech Target)

the area controller, gateway, and especially in the cloud.

Applications at the edge node aren't usually considered to be "data-heavy," but perspective is everything. An aircraft in flight can be considered a mobile edge node from the point of view of Boeing, Airbus, or the airline. However, the sensors on a single jet engine can generate 20 TB of data every hour. Even with a 5G infrastructure, the manufacturer may find it preferable to cache the information during flight and upload it while at the gate rather than risk losing data due to a dropped connection or present an opportunity to hackers.

In other edge-node applications, say, a smart factory, nonvolatile storage can help a machine analyze operational conditions and respond in real time to alerts or faults. A repair technician could also access the last few days of sensor data locally to aid in troubleshooting.

NAND flash is making inroads in the cloud, too. The NAND flash read-write speed is much faster than that of a hard-disk drive (HDD). Therefore, the initial cloud applications for all-flash storage are low-latency applications such as artificial intelligence (AI), real-time analytics, transactional databases, and machine learning.

The cost per gigabyte of flash storage is currently higher than for an HDD, although it's dropping quickly. Consequently, HDDs or hybrid flash/HDD solutions are still preferred for secondary data applications like backup storage, digital imaging, and disaster recovery.

Manufacturers are continually working to reduce the cost-per-gigabyte of NAND devices, so device life cycles tend to be quite short with frequent process feature shrinks. Longer term, industry forecasters are predicting that NAND flash will replace HDDs in almost all cloud-storage applications.

OVERVIEW OF NAND FLASH OPERATION

A NAND flash memory cell (*Fig. 2*) is similar to a standard MOSFET, but it contains two gates rather than one. The control gate turns the channel on and off, and the insulated floating gate stores charge. When a floating gate contains electrons, it's considered charged, or programmed. Otherwise, it's considered uncharged, or erased. The drive registers the cell's bit value as logical "0" when a floating gate is charged and as logical "1" when the floating gate is uncharged.

Turning on the channel requires a threshold voltage V_T to be applied to the control gate. If the floating gate is already charged, its effect is to increase the required threshold voltage to V_{T+} ; otherwise the required value of V_T is a lower value, V_T . To read a cell, the controller applies an intermediate voltage between V_T and V_{T+} . If the channel conducts, the floating gate is uncharged; otherwise, it's charged.

Unlike other memory technologies, a NAND flash device can't be reprogrammed at the individual byte level. Due to its architecture, it can only be read and programmed in pages—typically between 4 and 16 kB—and can only be erased in blocks that range from 16 to 512 kB in size. A block must be completely erased before being reprogrammed. This two-step process is called a "program/erase" (P/E) cycle. Repeated P/E cycles lead to the breakdown of the oxide layer within the floating-gate transistor cells and eventual failure. Thus, the endurance, measured in number of P/E cycles, is a key specification.

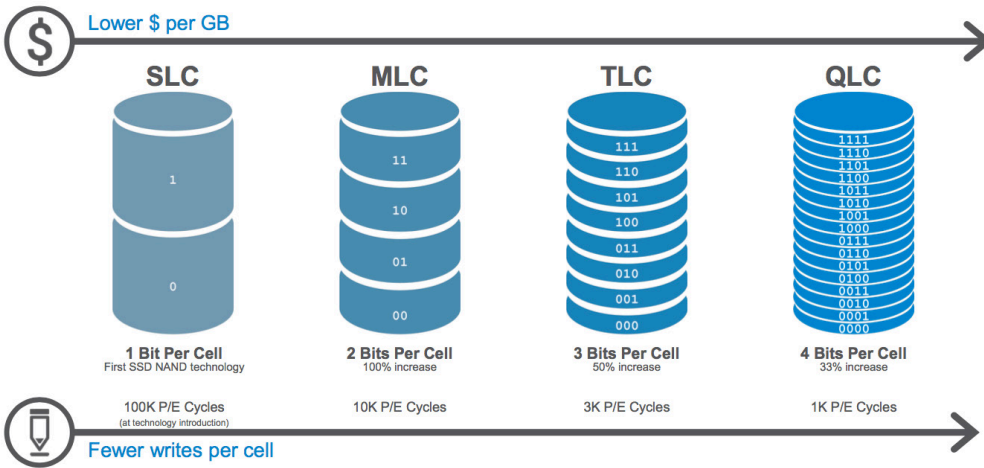
A NAND flash device requires an internal or external controller to implement error code correction (ECC), manage bad blocks, and perform wear leveling—a technique to increase endurance by counting writes and dynamically remapping blocks to spread writes more evenly. So-called "raw" NAND includes just the cells and requires an external controller. A fully managed NAND device integrates the controller.

NAND flash devices come in different varieties, classified by the number of bits that can be stored in a single cell: single-level cell (SLC), multilevel cell (MLC), triple-level cell (TLC), and quad-level cell (QLC). As *Figure 3* illustrates, there's a tradeoff between cell endurance and cell density: the higher-density cells have lower endurance.

More recently, manufacturers have begun offering three-dimensional (3D) NAND devices in MLC and TLC formats: current-generation 3D NAND devices stack up to 96 layers vertically for higher density.

Individual devices are available with capacities from 128MB to 2TB. Smaller-density devices are popular in consumer IoT applications such as wearables, smart appliances, set-top

QLC = More Density Per NAND Cell



3. The designer can choose between several varieties of NAND flash to balance between endurance and density. (Source: Micron Technology)

boxes, and smart TVs. High-density products find a home in cloud storage, plus data-heavy IoT applications such as gateways, factory automation, industrial PCs, surveillance, and drones.

There's also a choice of interfaces. The embedded multimedia card (eMMC) format uses an 8-bit parallel interface and integrates the flash memory and the controller into a 100-, 153-, or 169-ball-grid-array (BGA) package. The eMMC version 5.1 specification has a 400-MB/s transfer rate.

The Universal Flash Storage (UFS) specification employs a full-duplex LVDS interface for higher transfer speeds than eMMC. The latest (2018) UFS 3.0 revision uses two lanes to achieve a maximum total bandwidth of 2.9 GB/s.

NAND flash devices come in a variety of form factors, from single devices, to modules, to complete systems. Memory cards are used in both consumer and IoT applications; popular formats include USB flash drives, CompactFlash (CF), and secure digital (SD) cards in either standard or microSD sizes.

Solid-state drives (SSDs) combine large numbers of NAND flash devices and are popular as HDD replacements in both consumer and IoT applications. At the high end, a typical all-flash drive for cloud storage might be rack-mountable with 48 individual SSDs and a total capacity of 34 TB.

PICKING THE RIGHT NAND FOR YOUR IoT APPLICATION

The IoT space covers a lot of ground, so there's no "one size fits all" approach to picking a NAND solution. SLC NAND has the lowest density, but the highest performance and endurance. It's preferred for high-end, mission-critical applications. MLC and TLC devices feature increasing density and lower endurance, and QLC devices are suitable for high-density, cost-sensitive mass-storage applications.

Consumer-oriented IoT applications can use off-the-shelf components, but many other IoT applications have specialized requirements. For example, security is a high priority for secure boot code, firmware, keys, and other sensitive information.

Standard commodity parts don't include security features, but secure NAND flash devices offer volatile and nonvolatile block protection features. Volatile block protection locks one or multiple blocks until power is cycled; nonvolatile block protection permanently locks a fixed or variable range of blocks. Other security features include device-specific unique IDs, password protection, and AES256 data encryption.

Since some security features are computationally intensive, they're typically included in a module-level design, e.g., a secure SSD.

An extended temperature range is important for IoT edge devices that are exposed to the elements or operate in a harsh industrial environment.

Conclusion

The Internet of Things (IoT), with its emphasis on the collection and analysis of mass quantities of data, requires fast, flexible non-volatile storage at every level of the hierarchy. A wide range of NAND flash devices accommodates IoT applications from the edge node to the cloud.