MAXIM Engineering Journal

Volume Sixty-Three

LETTER FROM THE CEO 2

IN-DEPTH ARTICLES

Advanced Enterprise Features Enable Next-Generation SAS Systems 3

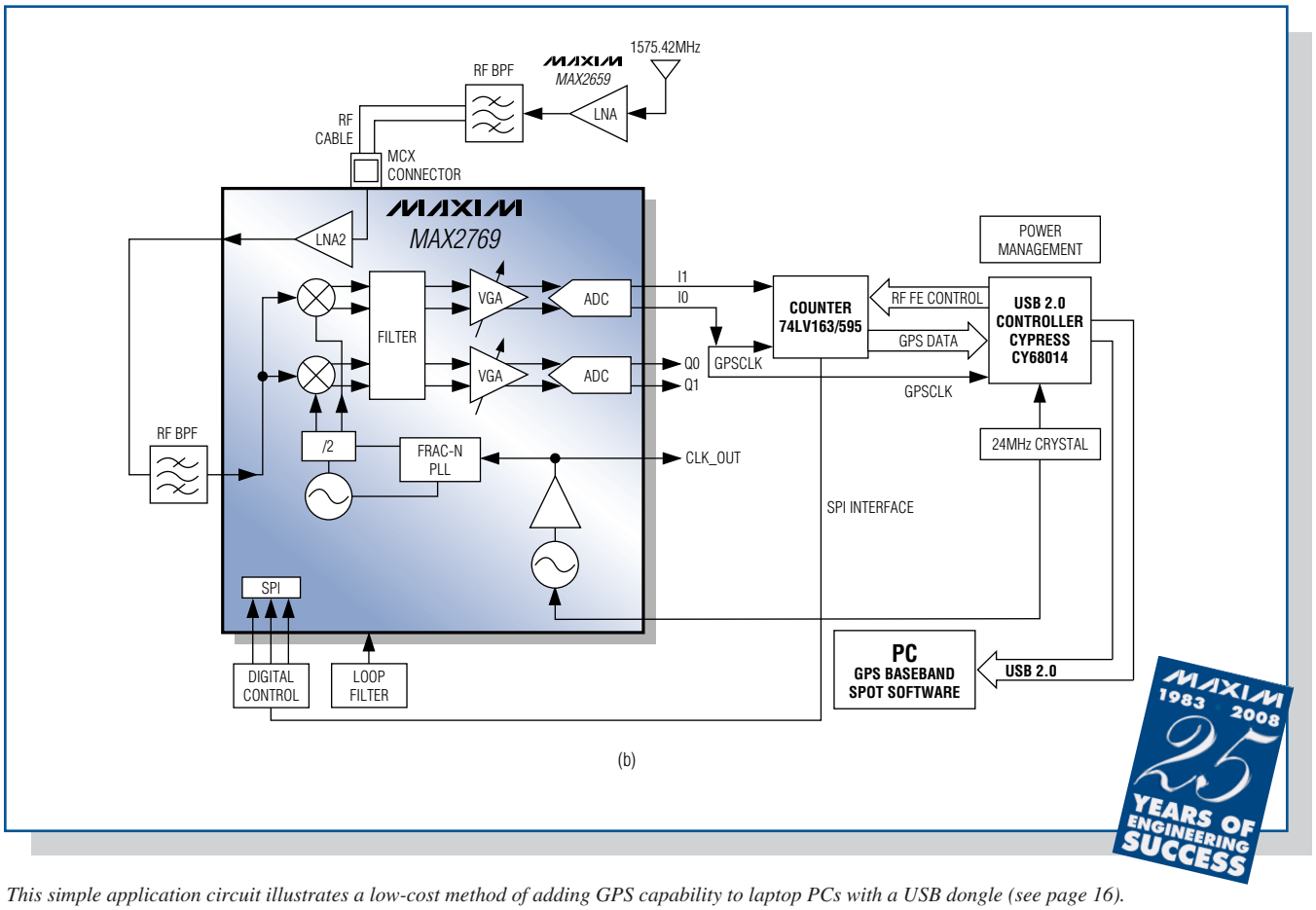
Carrier Ethernet Service Demarcation in Optical Networks 8

Universal GPS Receiver Lets You Use a Laptop PC for Soft Baseband Processing 12

DESIGN SHOWCASE

AC-Based Continuity Tester Finds Single-Ended Faults 17

Two Methods for Shutting Down a Current-Sense Amplifier 18



This simple application circuit illustrates a low-cost method of adding GPS capability to laptop PCs with a USB dongle (see page 16).

Letter from the CEO

Keeping Manufacturing Competitive to Better Serve Our Customers

In what feels like the distant past, Jerry Sanders (former CEO of Advanced Micro Devices, Inc.) made the widely quoted statement "Real men have [wafer] fabs." Although fab ownership has not proven to be a requirement for success in the digital world, in the analog and mixed-signal market, major players like Maxim do need to own fabs.

The very nature of the products that we design demands that we use many customized processes to optimize device performance. Our fabrication facilities, therefore, are valuable resources that provide our designers with an important technological edge, allowing them to create innovative products that deliver performance well beyond what a typical foundry process might allow.

We currently utilize close to 160 proprietary processes, and over 90% of all products are manufactured in our own fabrication facilities. We have no plans to adopt what the industry refers to as a fabless strategy and offload our manufacturing to foundries and other third-party suppliers.

Having said that, we do have two models for using external wafer fabs: 1) when we can use available technology, especially in deep submicron, where capital costs are too high for one company like ours to afford; 2) when we enter into strategic, exclusive agreements with a foundry to install our proprietary technologies for better flexibility and manufacturing redundancy.

Last year we did this by partnering with Seiko Epson to use their advanced, submicron 200mm fabrication plant in Sakata, Japan. This deal combines Maxim's innovative mixed-signal design and process technologies with Seiko Epson's world-class semiconductor manufacturing. As a result, we are able to deliver a new level of quality and service to our customers, with reduced cycle times and a very competitive cost structure.

We continually evaluate facility efficiency and capabilities to determine which fabs to maintain, upgrade, or shut down, and when to add new capacity to handle increased market demand. This year we did just that. We announced our decision to retire our oldest fabrication facility in Dallas, and transferred the products mostly to our larger, much more efficient manufacturing plant in nearby San Antonio, Texas. The San Antonio facility is a much newer plant that has many times the capacity of our twenty-year-old Dallas factory. The newer factory allows us to produce chips much more economically to ensure that we deliver the lowest cost product to our customers.

As production demands dictate, we will increase the capacity of the San Antonio facility and bring on-line other manufacturing facilities that are in the final stages of qualification. This is an on-going process to better serve you.

We are always at your service,

Tunç Doluca
President and Chief Executive Officer

Advanced Enterprise Features Enable Next-Generation SAS Systems

By Sam Barnett, Business Manager

As first-generation systems based on Serial-Attached SCSI (SAS) technology enter the market, leading server and storage OEMs are racing to deliver the next group of platforms. In conjunction with the release of generation I systems, Maxim has enhanced its popular NexSAS™ product family.

Delivering high-performance storage solutions for SMB (small-/medium-business) and enterprise systems, generation 1.5 of the NexSAS family includes features found only in high-availability, superior-performing, Fibre Channel systems today. The NexSAS line comprises three new high-PHY-count expanders, extensions in intelligent mux/demux devices, a new enclosure-management/SAS-backplane controller, and the industry's best rate-agile signal conditioners.

Maxim's NexSAS product family is divided into four functional areas: expander technology, SAS/SATA support devices, server/enclosure/baseboard management, and signal conditioning. This article details the technology breakthroughs behind generation 1.5 NexSAS expanders. Specifically, it discusses the NexSAS multi-affiliation STP/SATA bridge and support for SAS 2.0 zoning.

NexSAS Expander Technology

Designed for high-performance interconnect, the NexSAS expander family features a single functional design. This design approach ensures scalability from low-port-count offerings to the high-port-count devices needed for blade servers, storage enclosures, and switching/island SAN applications.

All of Maxim's expanders divide the core systems function into two separate components: the expander function and the management function. Included in the expander functional block are the expander connection manager (ECM), expander connection router (ECR), broadcast primitive processor (BPP), and the physical interfaces (PHYs) to the device (ranging in number from 6 to 36, depending on the application).

Complementing the core functions is the management functional block, which provides expander-management and vendor-specific enclosure-management functions. This block also integrates various peripheral interfaces needed for interaction in the overall system, including two-wire serial, JTAG, UART, general-purpose I/O (GPIO), and serial peripheral interfaces. The number and types of these interfaces vary with the particular expander and specific operating mode(s) employed.

Figure 1 outlines the basic functional block diagram of all expanders in Maxim's NexSAS product family. Generation 1.5 expanders offer a range of features to meet different system requirements. They also share a common set of features, including:

- **High-Performance Switching Architecture**. A low-latency, nonblocking switching matrix provides up to 54Gbps of aggregate switching capacity (in high-port-count expanders with linear scaling for lower-port-count expanders).
- **Maxim's Universal PHYs**. Self-configuring, rate-adaptable PHYs support SAS initiators and SAS/SATA targets at both 1.5Gbps and 3.0Gbps data rates. Multiple per-PHY preemphasis/deemphasis capability, and multiple per-PHY output levels leverage over 20 years of Maxim's technology development. They can be combined into SAS "wideports" of up to *n* PHYs wide.
- **Virtual Initiator and Target Capability**. SMP (Serial Management Protocol), SSP (Serial SCSI Protocol), and STP (Serial ATA Tunneling Protocol) initiator and target functions enable robust value-added services, such as self-configuration, self-discovery, SES-mastering, and rogue drive identification.
- **Integrated Enclosure-Management (EM) Subsystem**. Robust enclosure-management capabilities maximize OEM investment in legacy firmware/software/custom features, and allow flexibility in system designs.
- **Integrated Ethernet Interfaces**. Two 10/100 Ethernet MACs are included in the EM subsystem to support blade servers or other applications that require Ethernet access to the expander.
- **Multi-Affiliation STP/SATA Bridge**. Extensions of the original STP/SATA bridge provide for two concurrent (active/active) NCQ affiliations between initiators and targets using Maxim's expander technology. This innovative approach eliminates the initiator starvation issues commonly associated with the single-affiliation mechanisms in SAS 1.1.

- **End-to-End SAS 2.0 Zoning**. For security, control, and performance, NexSAS expanders support both SAS 2.0 PHY-resolved and address-resolved zoning.
- **Other Enhancements Beyond Generation I Devices**. Additional capabilities include EPOW, integrated A/D converters, LPC for trusted-platform module interconnect, and parity protection of the on-chip memory.

Recent additions to the NexSAS family include the VSC7156 24-PHY, VSC7157 36-PHY, and VSC7158 18-PHY expanders. Like their predecessors (the VSC7153, VSC7154, and VSC7155 generation I expanders), these devices are targeted at enterprise server, storage enclosure, and blade applications.

Generation 1.5 expanders bridge the gap between the SAS 1.1 and planned SAS 2.0 specifications. This article examines two of the key differentiating features, the multi-affiliation STP/SATA bridge and SAS 2.0 zoning, in more detail below.

Understanding the NexSAS Multi-Affiliation STP/SATA Bridge

Fairness and Performance Limitations of Prior Designs

The SAS 1.1 specification defines the STP/SATA bridge as the expander component that provides a bridge between STP-capable initiators and SATA targets (HDDs). Without an STP/SATA bridge, SAS could not support Serial ATA (SATA) drives.

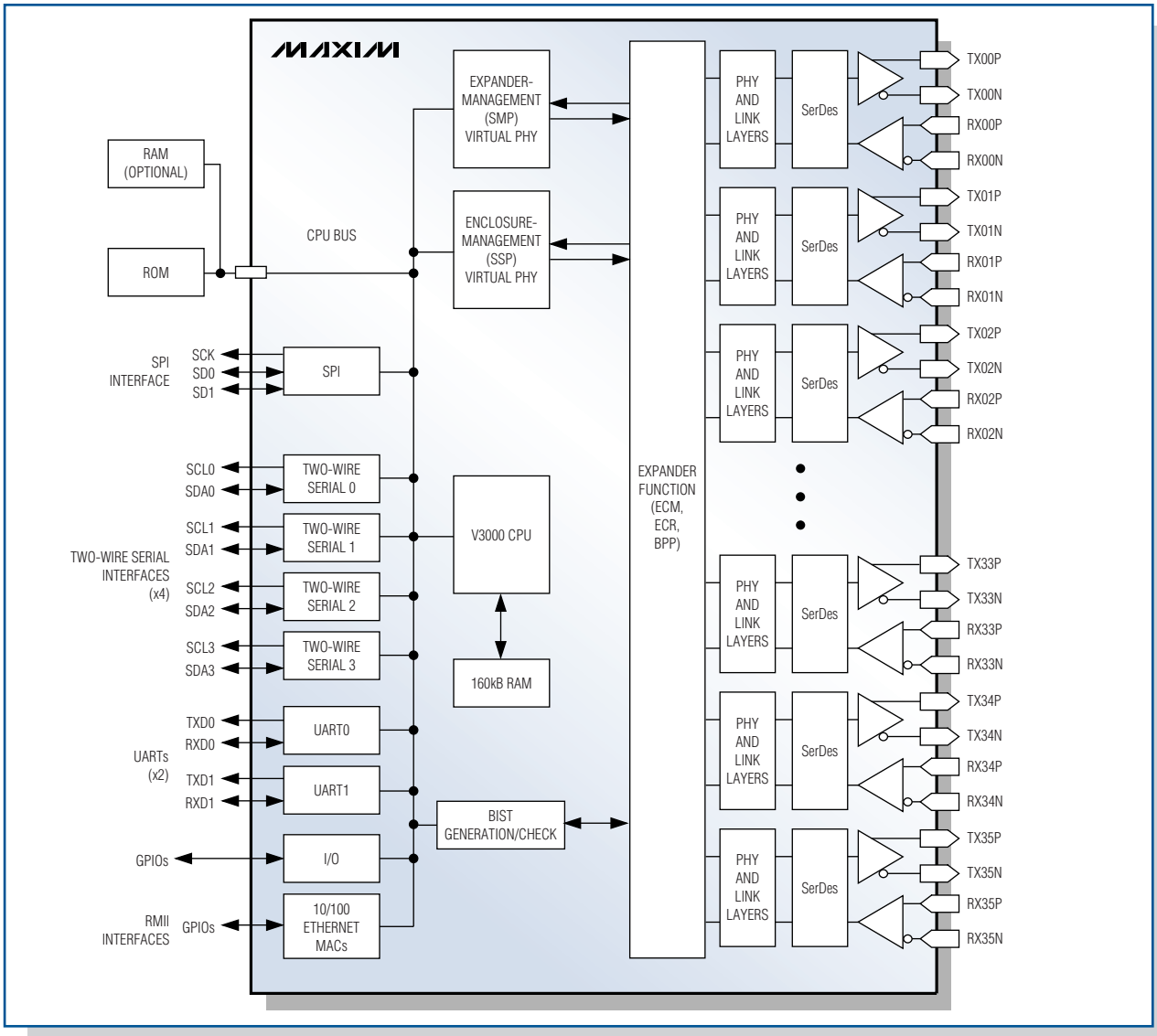


Figure 1. Functional diagram of Maxim's generation 1.5 NexSAS expander family.

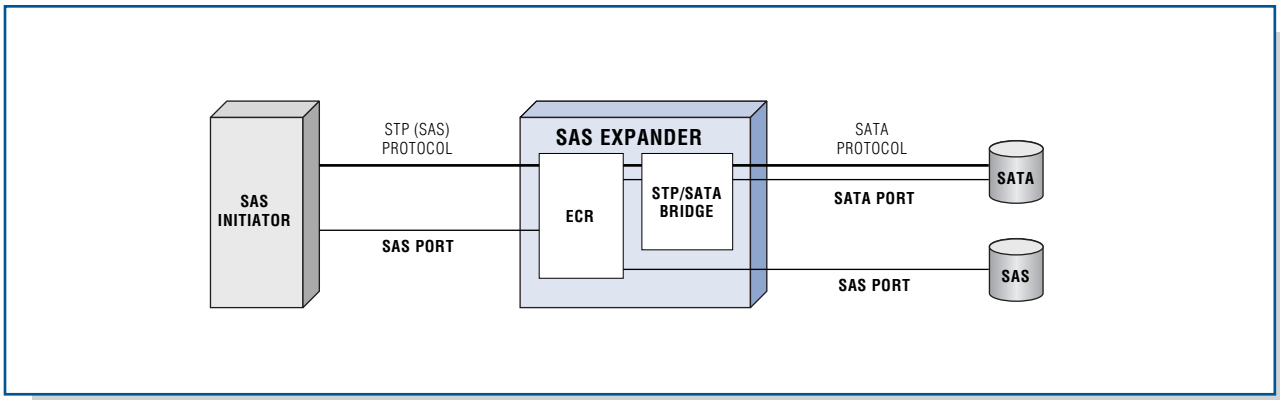


Figure 2. The STP/SATA bridge supports connections between STP-capable initiators and SATA drives.

In general, the SAS 1.1 specification defines the function of the STP/SATA bridge as follows:

- In SAS expanders, the STP/SATA bridge allows the connection of SATA drives to the SAS domain.
- The SAS protocol is used for connection setup/tear-down between the host (initiator) and the STP/SATA bridge.
- The STP/SATA bridge only passes native SATA protocol to the SATA drive during connection.
- Expanders typically have one STP/SATA bridge available in each PHY, but the bridge is enabled only when connected to the SATA device.

Figure 2 illustrates the STP/SATA bridge concept. Unlike SAS drives, SATA drives do not recognize multiple-host or multiple-initiator concurrent access. During a session, one STP initiator maintains an affiliation with the SATA target. During this affiliation, any other STP initiator attempting to gain access to the same SATA device receives an OPEN_REJECT (STP RESOURCES BUSY) and is forced to retry the connection attempt. This simple, exclusive affiliation mechanism maintains coherence by allowing commands from only one host to be active on the drive at any given time.

When all outstanding commands from a given initiator have completed, the initiator *should* send a special CLOSE (CLEAR AFFILIATION) primitive to the STP/SATA bridge, resulting in the release of the affiliation and allowing other hosts to access the drive/resource.

As you can imagine, generation I expander devices support only a single affiliation in compliance with the SAS 1.1 specification.

With any new technology, implementation tends to expose certain undesired behaviors. The original STP/SATA bridging mechanism was no different. Limitations of the originally envisioned bridge involve both fairness and performance. Manifestations of these limitations are outlined briefly below.

Fairness

- There are no formal limits imposed on the longevity of a given affiliation, meaning that an affiliation could be established forever.
- Fair access to SATA devices must be enforced by the initiators; in the absence of host-to-host communications, this may not be assured.
- Software algorithms to achieve fairness have proven painful for OEMs.
- Existing real-world STP initiators do not "play fair:" they occasionally fail to release affiliations.
- If a valid affiliation exists between the expander STP/SATA bridge and an initiator, other hosts cannot perform any inquiries to the drive, including diagnostic ones.

Performance

- Single-threaded host access to SATA disks is slow in load-balancing systems.

Figure 3 depicts the SATA affiliation concept in SAS.

New Approach Overcomes Traditional Performance Limitations

Generation 1.5 NexSAS expander products address both the fairness and performance limitations of the SAS 1.1 STP/SATA bridge. Through the updated bridge function, two hosts can simultaneously issue active commands on the same SATA drive. **Figure 4** illustrates the operation of the multi-affiliation STP/SATA bridge.

Details of the new bridge operation are summarized below.

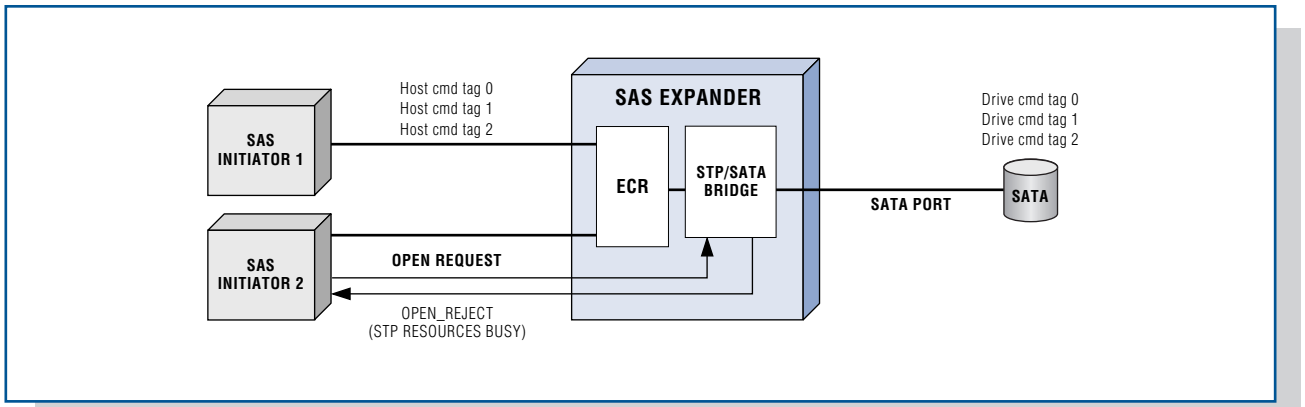


Figure 3. The SATA affiliation concept in SAS 1.1 only supports single affiliations, resulting in unanticipated fairness and performance limitations.

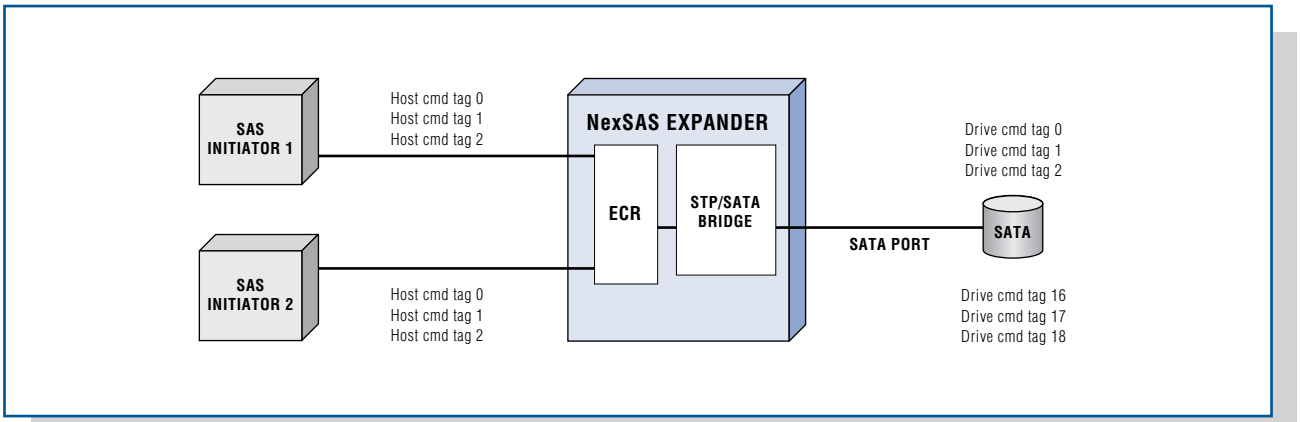


Figure 4. Generation 1.5 NexSAS expanders support multi-affiliation STP/SATA bridge operation, allowing two hosts to simultaneously issue active commands on the same SATA drive.

Intelligent Connection Management

- Flexible, nonexclusive disconnect policies (overlaps are allowed) include:
  - Timed disconnect (multiple modes and ranges)
  - Disconnect after any FIS transfer
  - Disconnect after data FIS transfer
  - Disconnect after interlocked operations
- Automatic callback of the proper initiator (host) based on queue tag/context

NCQ and PIO Command Support

- Allows utilization of all 32 NCQ command tags (or fewer by configuration) on the SATA drive
- Manages interleaving of NCQ and PIO commands from different hosts
- Tags are transparently managed and remapped to/from the drive to avoid tag number collisions between initiators (no special host software or firmware is required)

The multi-affiliation STP/SATA bridge employed by all generation 1.5 NexSAS expanders enhances fairness, performance, and system design convenience. Similarly, it benefits any SAS domain using a SATA disk that is shared by two or more hosts. It features an architecture that can scale for future products requiring even more active affiliations, while remaining completely compliant to the SAS 1.1 and proposed SAS 2.0 specifications.

Understanding the Basics of SAS 2.0 Zoning

As storage OEMs and integrators embarked on the introduction and deployment of SAS-based systems, many longed for the days of Fibre Channel and some of the enterprise features it offers for traffic management (performance) and security.

To address those demands, the concept of end-to-end zoning was introduced into the proposed SAS 2.0 specification (refer to www.t10.org for the latest on the zoning specification).

The SAS protocol, as defined, supports up to 16,384 devices (initiators, targets, expanders, and/or virtual devices) in a given domain; however, practical implementation of such large topologies is quite problematic. In essence, end-to-end zoning provides the mechanism necessary to manage these large physical topologies in smaller logical (segmented) groups. SAS zoning provides this function as well as the mechanisms to control access among and between groups.

The proposed SAS 2.0 specification divides zoning into two areas: PHY-resolved zoning and address-resolved zoning. Both approaches use identical access-control methodologies, but differ in zone group allocation.

PHY-Resolved Zoning

In a PHY-resolved zoning implementation, each PHY of a zoning-capable expander is assigned to a zone group; any device attached to that PHY (or port, in the case of a PHY grouping for wideports) becomes part of that zone group. A permission table in the expander contains access-control permissions that outline the protocol for sessions between devices in differing zone groups (by default, all devices within a zone group can interact with each other).

PHY-resolved zoning is ideal for smaller topologies, blade servers, or clustering applications.

Address-Resolved Zoning

Address-resolved zoning provides a layer of abstraction not present in PHY-resolved zoning. In an address-resolved scheme, self-discovering expanders interpret each device address and populate a table. This table contains routing information and zone permissions. Permissions in the table and zone group assignment are combined to determine what level of interaction any two devices may have with each other.

Typically, address-resolved zoning is used for larger topologies or where more granular control is needed.

Conclusion

One size never fits all, be it SAS or something else. Proper solutions differ by application, connectivity requirements, scalability, performance, and price sensitivity—all functions of innovation. For a list of all the products in the NexSAS storage products family—the server-series expanders, the enclosure-series expanders, the SAS/SATA support devices, and the backplane/enclosure-management/baseboard-management controllers—go to: www.maxim-ic.com/products/storage

Carrier Ethernet Service Demarcation in Optical Networks

By Art Harvey, Business Manager

The Metro Ethernet Forum (MEF) has now standardized a number of services for wide-area Ethernet connectivity, collectively known as Carrier Ethernet Services. The MEF's goal in defining these services is to encourage the ubiquitous adoption of Ethernet by promoting the interoperability of high-quality Ethernet services among service providers. To ensure that consumers of these services can compare features across service providers, the MEF has also defined measurable attributes for each service. These attributes must be meaningful to consumers. Therefore, they must be representative of the observable service characteristics at the point where a customer physically connects their network to the service provider's network. The physical interface where this connection occurs is known as the demarcation point.

Service demarcation has been common practice for telephone service providers ("Carriers") for decades. The most familiar form of demarcation is the small box mounted outside almost every suburban residence. This box connects the local Carrier's phone network to the wiring of the house to provide wireline phone services to the resident. This small box demarcates where the customer's responsibilities end and the Carrier's responsibilities begin.

For telephone services, the functional requirements of the demarcation unit are minimal. With Carrier Ethernet services, however, the requirements of the demarcation unit are much greater. The Carrier normally provides the customer with a contractual service-level agreement (SLA), which defines attributes of the service such as committed information rate (CIR), committed burst size (CBS), service availability, frame delay, frame jitter, frame-loss rate, and fault-recovery time. A Carrier Ethernet demarcation unit plays a critical role in ensuring that the actual service attributes adhere to the SLA.

Requirements of Demarcation

At the very minimum, an Ethernet demarcation unit provides a physical connection and measurement point: either an RJ-45 jack or an optical connector. The physical-layer interconnect has been defined by the IEEE® and is included in the MEF specifications by reference. The demarcation unit must accept standard IEEE 802.3 Ethernet frames from the subscriber and prepare them for transport across the service provider's network. The functionality beyond these minimal requirements varies by application.

The customer-interfacing part of the demarcation function is called the user network interface (UNI). The MEF is standardizing a range of functionality for the UNI, from the most basic UNI Type 1 up to the autoconfiguring UNI Type 3. The MEF recently approved a specification called *MEF 13 UNI Type 1 Implementation Agreement* and the associated certification-testing process. UNI Type 2 and Type 3, outlined in MEF 11, are expected to be expanded upon by the MEF in the future.

MEF 13's UNI Type 1.2 requires that demarcation units process certain layer-2 protocols arriving at the UNI from the customer's network. This requirement forces the demarcation unit to have layer-2 visibility and filtering. Furthermore, the ITU and MEF requirements for UNI Type 2 and Type 3 to perform certain layer-2 management protocols force the demarcation unit to have full layer-2 processing capability—at least, at a low level of throughput. These requirements impact technology decisions when architecting a demarcation unit design. One beneficial side effect of these requirements is that most demarcation unit designs have some processing power available for adding valuable higher-layer application functionality.

Anatomy of a Demarcation Unit

Figure 1 shows the basic functional blocks of an Ethernet demarcation unit. The functional blocks that interface with the customer's network are shown on the right side of the diagram under "UNI," while the functional blocks that interface with the Carrier's transport network are on the left under "network interface." The group of network interface functions are called the network interface device (NID). Although the term NID has also been used to refer to a stand-alone piece of equipment, it is used here to describe a functional part of demarcation. The MEF currently has a standard for NID functionality in draft form, but has not yet approved it. The NID function may, or may not, be in the same piece of equipment as the UNI. In practice, the dividing line between the UNI and NID varies. To further confuse the issue, the MEF has defined a range of UNI subtypes. Specifically, a demarcation unit normally contains the UNI-network (UNI-N) interface or external network-to-network interface (E-NNI).

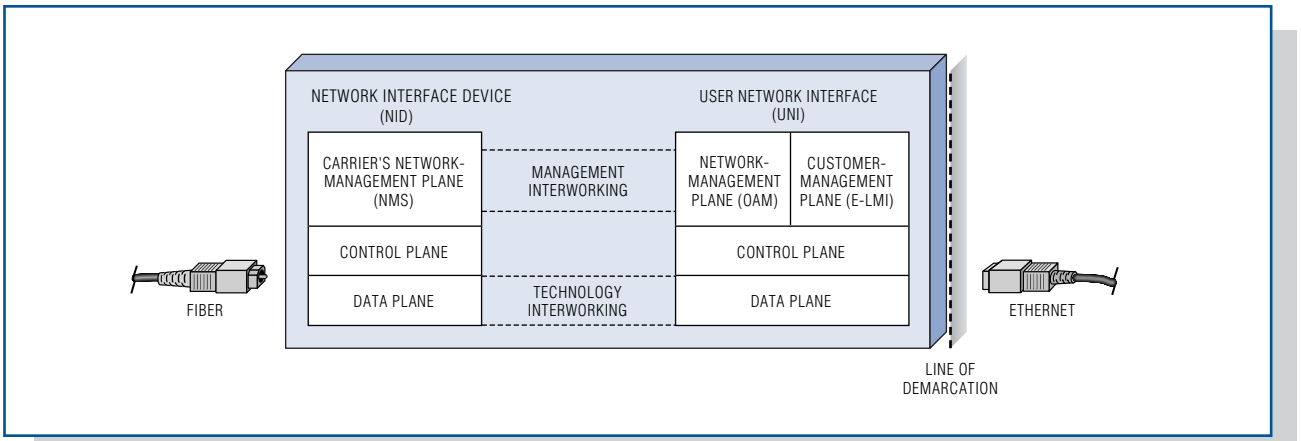


Figure 1. A functional block diagram is shown for an Ethernet demarcation unit. UNI comprises the functional blocks that interface with the customer's network, while the functional blocks under NID interface with the Carrier's transport network.

The hardware of the UNI and NID data planes are coupled very closely to the network technologies connected to them. The UNI's physical interface must consist of an IEEE-compliant electrical or optical Ethernet interface, while a Carrier's optical network connected to the NID will often be SONET/SDH. These differing technologies require some translation for the customer's traffic to traverse the Carrier's network. The translation, or interworking, function exists between the two data planes. The data-plane interworking function is performed by an integrated circuit called an Ethernet mapper, a network processing unit (NPU), or a customized field-programmable gate array (FPGA). In the SONET/SDH example, an integrated Ethernet-over-SONET/SDH (EoS) mapper is usually the best solution to perform the data-plane interworking function. The data plane is responsible for the transport and tagging of customer traffic, as well as flow control or traffic shaping.

The UNI and NID typically have separate control-plane software that handles low-level configuration and status monitoring. Examples of activities handled by the control plane are: detection of a cable connection or disconnection, status and performance monitoring, and the handling of interrupt events. The control plane is responsible for ensuring that the service is provided as instructed by the management plane. The control plane is normally implemented in software running on a local microprocessor that configures and controls the data-plane hardware. In stand-alone demarcation units, it is common for the control planes of the UNI and the NID to run on the same processor. The control and data planes can sometimes be implemented in one NPU, although doing so can lead to complex data- and control-plane corruption issues. Architectures that maintain a clear separation between the data and control planes are usually easier to implement.

Management planes are normally implemented in software at the line-card or chassis level. The UNI and NID both have management planes for fault recovery, SLA performance monitoring, etc. The management plane's primary purpose is to handle issues that concern the network operator. The MEF has defined the structure for the network operator's management information in *MEF 7 EMS-NMS Information Model*. Optionally, the UNI may also have a customer-facing management plane in addition to its network-facing management plane. The MEF has defined a basic structure for this customer-facing management interface in *MEF 16 Ethernet Local Management Interface* (E-LMI). The management planes for the UNI and NID can exchange information, but it is not a requirement.

The UNI management plane also uses a special protocol for Ethernet management defined by ITU-T and IEEE, called Ethernet OAM (operation, administration, and maintenance). IEEE 802.3ah OAM monitors the operation and health of a single point-to-point link and improves fault isolation. ITU-T Y.1731 OAM increases the scope to include advanced multiple-link operations, such as link tracing, connectivity checking, and automatic protection switching. Using OAM, the network operator can perform management tasks that were unavailable on Ethernet networks only a few years ago.

UNI Technologies

The network transport technology options for the UNI are limited by the scope of the MEF standards to full-duplex 10Mbps, 100Mbps, 1Gbps, or 10Gbps IEEE-compliant electrical or optical Ethernet. To effectively cover a range of physical connections, some demarcation units simply allow a small form-factor pluggable (SFP) module to be

inserted and, thus, configure the UNI's physical interface at the time of installation. The rest of the required UNI data-plane functionality is then implemented independent of the physical interface by using an Ethernet mapper, NPU, or FPGA. The alternative to the SFP approach is to use an on-board PHY or optical module. This alternative sacrifices some flexibility for a lower total solution cost.

NID Technology Driven by Location in the Network

Figure 2 illustrates two common demarcation situations. At site A (on-ring demarcation), direct optical-network access is available at, or very near to, the customer's premise. This situation is common in metropolitan networks and in intercarrier handoffs known as E-NNI. When the demarcation point is on the optical edge, the NID must connect either directly to the optical network or to a lower-rate tributary made accessible by a piece of optical-networking equipment. The choice of NID technologies for applications on the optical edge spans a wide range: traditional SONET/SDH, Ethernet-over-PDH-over-SONET/SDH (EoPoS), provider backbone bridge (PBB), transport multiprotocol label switching (T-MPLS), passive optical networking (PON), dense-wave division multiplexing (DWDM), resilient packet ring (RPR), optical Ethernet, and hybrid fiber-coax (HFC).

At site B (off-ring demarcation) in Figure 2, the optical network does not reach the customer's premise, and a "last-mile" technology must be used to reach the demarcation point. This situation is common in lower-density urban, suburban, and rural areas. NID technologies for demarcation away from the optical edge include: Ethernet-over-PDH (EoPDH), digital subscriber line (DSL), Ethernet first mile (EFM), and data-over-cable service (DOCSIS).

The use of a particular transport technology for a given NID is dictated by what will be the most cost-effective interface with the Carrier's existing network at the point where the service needs to be delivered. It is beneficial when architecting a demarcation platform to use a modular design to cover a range of expected UNI and NID transport technologies.

The State of the Art in Ethernet Mapping

Typically, an integrated Ethernet mapping device performs the UNI-NID interworking function in an Ethernet demarcation unit. EoS and EoPDH mappers are the two most common mapping technologies. EoS mappers have been prevalent for almost ten years, but EoPDH mappers have been available for less than two years. The recently announced DS33X162 family of EoPDH Ethernet mapping products from Maxim optimally meets the needs of today's demarcation designs. The product family includes nine devices ranging from one to sixteen PDH links, and is the industry's only solution that covers the complete range of ITU-standardized EoPDH mapping possibilities with a single software/hardware footprint. All devices in the product family maintain critical control- and data-plane separation, while still allowing flexibility for valuable higher-layer protocol implementations.

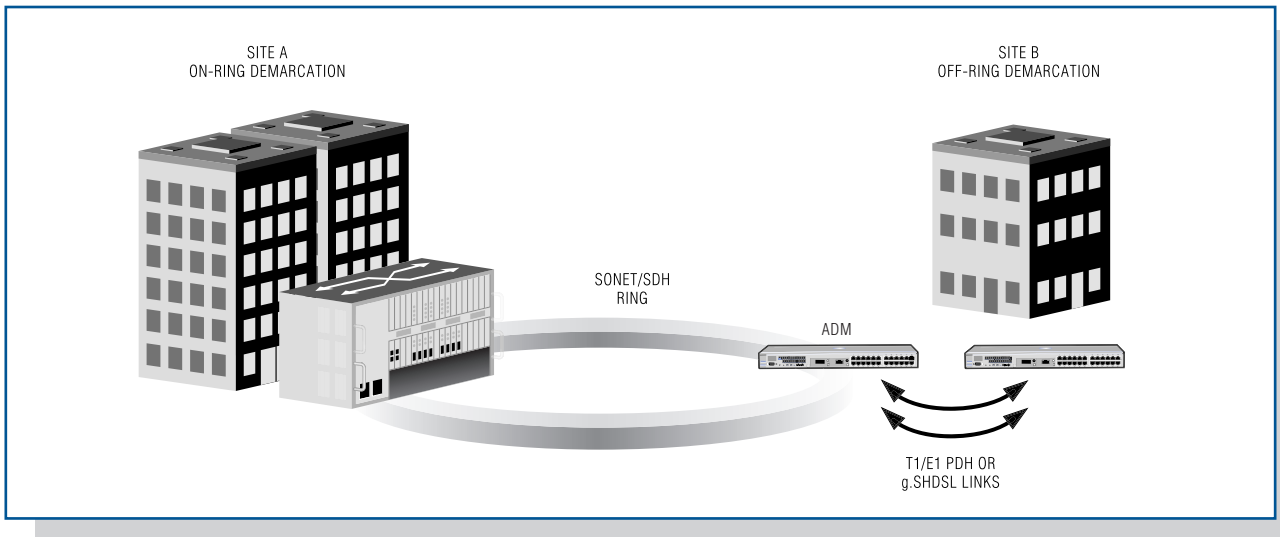


Figure 2. Two common demarcation situations are shown. Site A represents a common metropolitan network in which direct optical-network access is available at, or very near to, the customer's premise. Site B illustrates a common suburban network in which a "last-mile" technology is employed to extend the optical network to the customer's premise.

Going Beyond the Minimum Requirements

The MEF has only standardized a minimum set of UNI functions at this time. Yet, there is plenty of room for equipment makers to integrate value-adding functions to differentiate their demarcation unit from the competition while maintaining compliance. The value-adding features fall into two categories: features that help sell the demarcation unit to the Carrier, and features that help the Carrier sell the service to their customers.

Cost of installation is a primary concern to Carriers. A demarcation unit with features for automatic provisioning and built-in, network diagnostic testing lowers total installation cost. The use of technologies that can deliver midband Ethernet while reusing existing equipment and infrastructure, such as EoPDH, also reduces costs for the Carrier. Real-time SLA monitoring and seamless integration into the Carrier's existing network-management system also add significant value to Carrier services.

Service customers, meanwhile, value convenience and reliability. An example of a convenience feature is a web-browser-based (HTML) management interface that provides a complete overview of the health of all the customer's UNIs with a single glance. The interface could be served from the UNI itself, or it could reside in a central location and collect information from all of the customer's UNIs. The same HTML subscriber interface can provide SLA-compliance reporting for each UNI over time, bandwidth-usage profile reporting, and configuration options for quality of service.

Customers also might want to utilize VLAN tagging to provide varying qualities of service to different VLANs or applications. Enterprise network administrators need automatic status monitoring using their existing network-management tools, such as SNMP, or they need to receive notifications through email alerts. Growing businesses want to be able to incrementally expand their bandwidth as their consumption increases. NID technologies that utilize VCAT/LCAS link aggregation make it easy to add bandwidth. The demarcation unit can provide for dynamic provisioning to increase the available bandwidth during periods of high demand, and decrease the bandwidth in times of low demand. These are just a few examples, but they give an idea of what is possible in the future.

Now that the MEF has defined a common starting point with MEF 13, demarcation unit architects have a common foundation on which to build tomorrow's demarcation technology.

Universal GPS Receiver Lets You Use a Laptop PC for Soft Baseband Processing

By David Weber, Strategic Applications Engineer, and Roger Bremer, Strategic Applications Engineer

Communication and navigation engineers are increasingly using software techniques for global positioning systems (GPS).[1,2] Thanks to very-large-scale integration development, powerful CPUs and DSPs can now detect and decode GPS signals in real time using software. The resulting software-based GPS receivers offer considerable flexibility in modifying settings to accommodate new applications without redesigning hardware, choosing an IF frequency, or implementing future upgrades.

With the MAX2769 GPS receiver RF front-end, a simple USB dongle or PCI Express® (PCIe®) Mini Card format can be used to add low-cost GPS capability to laptop PCs. The MAX2769 transfers raw recovered data to the PC host, which executes the baseband decoding in software, thus eliminating the baseband ASIC typically required in stand-alone GPS systems. In short, the MAX2769 enables designers to implement an adapter capable of handling GPS and Galileo signals using a single chip.

This article provides an operational overview of GPS systems and, then, details Maxim's software-based GPS receiver solution.

GPS Signaling

A GPS system consists of 24 satellites or space vehicles (each identified by a unique pseudorandom-noise code), a ground-control station, and user equipment (receivers). For civilian GPS and Galileo applications, the satellites communicate over the L1 band located at 1.57542GHz.[3] A GPS receiver requires line-of-sight visibility of at least four satellites to establish a reliable position. The acquisition and tracking of signals is very complex because each one varies with time and receiver location.

The GPS system is actually a simple spread-spectrum communication system.[4] **Figure 1** shows the signal generation block for civilian applications. First, the 50bps navigation message is repeated 20 times to produce a 1000bps bit stream. The repeated signal is then spread by a unique coarse/acquisition (C/A) code with a length of 1023 chips—a chip is the rate at which the pseudorandom-noise (PRN) code is applied. The result is a baseband signal of 1.023Mcps. Hence, the 43dB processing gain (G) of the GPS system permits it to resolve a signal well below the thermal-noise level.

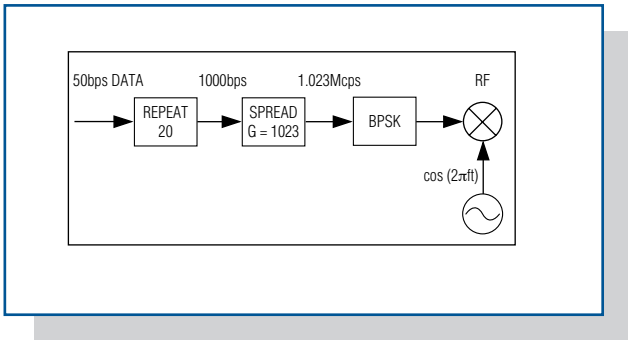


Figure 1. The signal generation block for the civilian GPS signal uses a sspread-spectrum approach for sending the signal.

Each satellite is assigned a unique C/A code, also called a Gold code.[5] Because the Gold code exhibits excellent auto- and cross-correlation properties, it is widely used in CDMA communication systems, such as WCDMA and cdma2000®. The baseband signal is modulated with binary phase-shift keying (BPSK), and upconverted to the L1 band for transmission.

Analysis of Signal-Acquisition Methods

Because GPS is a CDMA communications system, the receiver must synchronize the PRN code as a prerequisite to demodulating the data. Code synchronization is usually achieved in two steps: code acquisition for coarse-code alignment, and code-phase tracking for fine alignment.[6] More explicitly, a GPS receiver must first determine whether it has line-of-sight visibility to certain satellites or not. As we know, each satellite is distinguished by a unique C/A code. When the satellite is visible, acquisition determines the signal's frequency and code phase, which in turn establishes the corresponding demodulation parameters. The received-signal frequency varies due to the Doppler effect, which causes the frequency to deviate from its nominal value by 5kHz to 10kHz, depending on the speed of the satellite with respect to the receiver.[7]

In the receiver, the GPS signal is first downconverted to in-phase and quadrature (I and Q) components. A pair of I/Q correlators then correlate the I and Q baseband signals with the locally generated PRN sequence. After integrating over

the duration of one bit, the I-Q correlator outputs are summed to provide an output-decision variable.

Whenever the decision variable exceeds a certain threshold value, the system assumes that the corresponding acquisition was successful and proceeds to tracking mode. Otherwise, the relative phase of the PRN sequence and the oscillator frequency are adjusted to update the decision variable, and the above process is repeated.

While the simple logic structure of this serial-search method makes it feasible for implementation in an ASIC, it is not practical for software implementations because the search space is huge. Assuming that the system tolerates a 500Hz carrier-frequency offset and the Doppler frequency is 10kHz, the search space for a software implementation is roughly 2 x (10,000/500) x 1023 = 40,920. Obviously, a serial-search acquisition would be difficult in software.

A simpler acquisition method to implement with software is frequency-domain, parallel code-phase acquisition. This method combines the Doppler-frequency and code-phase searches into one search, which, after a fast Fourier transform (FFT) of the PRN code, reflects all code-phase information into the frequency domain. We then need only to search the space over the Doppler-frequency offset, thereby implementing a fast and effective software search.

The system implements this search by first multiplying the incoming signal with the locally generated sine and cosine carrier waves (the I and Q signal components, respectively). The I and Q components are then combined as a complex input to an FFT block. The result of this Fourier transform is multiplied with the conjugate of a PRN code's FFT transform (the PRN generator generates a code with zero-code phase). In practice, the FFT operation and generation of PRN code can be tabulated to reduce computation complexity.

Finally, the product of the incoming signal and local code, which represents the correction between the incoming and carrier frequencies, is applied to an inverse Fourier transform whose squared output feeds back to the decision logic. The FFT-based frequency domain has proven to be a low consumer of computational resources. For the example mentioned earlier, the complexity of acquisition is roughly 20,000/500 = 40 FFT operations.

The serial-search method has the simple logic and control architecture necessary for a convenient ASIC implementation. However, the huge search space required for this method imposes additional complexity on the software algorithm. The serial-search method, therefore, is not a good choice for software GPS receivers. In contrast, the low complexity of the parallel-code acquisition method makes it ideal for software implementation. Its logic architecture, however, is far more complex than that of the serial-search method, making it difficult to implement in an ASIC.

Tracking Refines Alignment

Acquisition establishes a coarse alignment of the GPS signal's frequency and code-phase parameters. The purpose of tracking, therefore, is to refine this alignment so that the system can demodulate the data with exact code-phase and frequency information. Tracking includes code-phase tracking and carrier-frequency tracking.

Code tracking is accomplished with the delay-lock loop (DLL) shown in **Figure 2**. The DLL circuit multiplies the incoming signal by three local replicas of the PRN code (positioned in time at ±0.5 chip); these replicas represent early, prompt, and late arrivals with respect to the incoming signal. After integration, each of these signals represents a correlation between the incoming signal and a local replica. The one with the highest correlation value is then selected and retained **(Figure 3)**.

Carrier-frequency tracking is carried out by a phase-lock loop (PLL) or Costas loop.[8] The purpose of carrier tracking is to tune the locally generated frequency to the exact frequency of the incoming signal.

After acquisition and tracking have established the initial synchronization, the system can decode the navigation bits. Data demodulation begins by despreading the 1.023Mcps input signal to a 1000bps bit stream. Bit synchronization is then invoked to recover the 50bps information from the 1000bps stream. For bit synchronization, we first need to identify the beginning of a bit in time by finding the zero-crossing edge (at zero volts). When that edge is known, we can partition the 1000bps input stream at 20ms intervals,

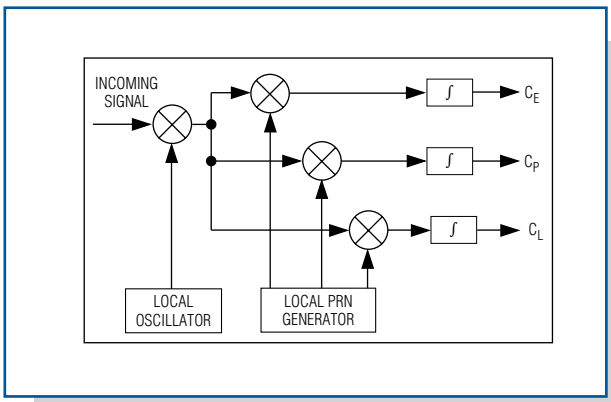


Figure 2. A delay-locked loop is used in the code-tracking phase to help refine the alignment to better demodulate the data with exact code-phase and frequency information.

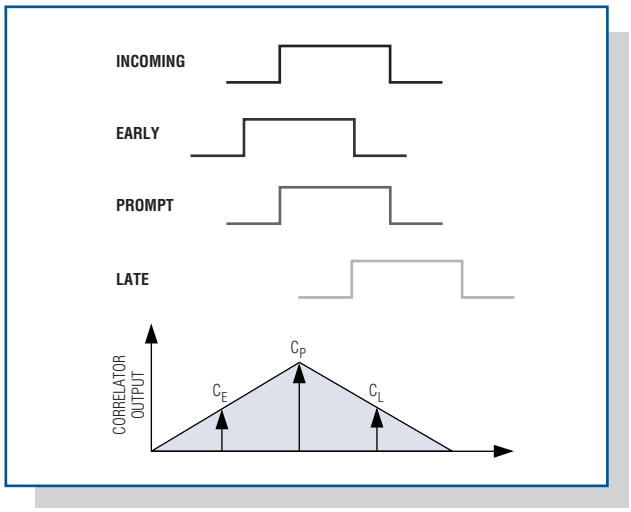


Figure 3. The DLL circuit multiplies the incoming signal by three local replicas of the PRN code (positioned in time at ±0.5 chip), which represent early, prompt, and late arrivals with respect to the incoming signal. The one with the highest correlation value is then selected and retained.

knowing that the duration of a navigation data message (50 bits) is 20ms. Finally, the bit samples in a 20ms interval are summed and averaged to decode the navigation data.

Software-Based GPS Receivers

Whereas traditional GPS receivers implement acquisition, tracking, and bit-synchronization operations in an ASIC, software-based GPS receivers offer added flexibility by implementing these functions in software. By simplifying the hardware architecture, software makes the receiver smaller, less costly, and more power efficient. The software code can be written in C/C++, MATLAB®, and other languages, and ported into all operating systems (embedded OS, PC, Linux, and DSP platforms). Thus, software GPS receivers offer the greatest flexibility for mobile handsets, PDAs, and similar applications.

For laptop computers, designers can design a USB dongle to work with any laptop with USB ports. For newer laptops with a PCIe Mini Card connector, they can put an RF front-end on the PCIe Mini Card and embed the card in the PC **(Figures 4a, 4b)**. The PCIe Mini Card interface includes a USB port, so the design of the front-end adapter is similar for both USB and PCIe Mini Cards. The main difference is in some of the power-management logic required to support PCIe and the differing DC voltages (3.3V for PCIe and 5V for external USB ports).

Figure 5b shows the block diagram of a USB dongle, demonstrating how simple the solution can be. This circuit employs a MAX2769 universal GPS receiver, a counter, and a USB interface controller to capture the signal, convert it to digital, and send it to the host PC. Software then performs all the baseband functions and displays the location on the PC screen. The notebook PC thus becomes a powerful GPS device that can support navigation and a wealth of location-based services.

The GPS front-end streams digitized IF data into the notebook through an industry-standard USB 2.0 interface. The software baseband program uses the input to calculate a position fix and, subsequently, to perform tracking. One possible source of the software is Geotate. For more information, go to: www.geotate.com.

To provide a generic interface, the software can create a virtual COM port so that it can link to a wealth of existing navigation and location applications. Most GPS software package interfaces conform to NMEA 0183 and will typically run on Microsoft's Windows® XP and Windows Vista™ operating systems. Additionally, the software should be able to handle all available assistance data, whether supplied through industry-standard protocols or proprietary customer interfaces.

The CPU inside most of today's laptop computers has plenty of performance to meet the real-time decoding needs of software GPS receivers. On a 1GHz Pentium® M system, the average processor load when tracking will be about 6%; on a 2.18GHz Core™ Duo processor, the processor load will typically be below 5% when performing updates every second. With further algorithm development, CPU usage can be reduced to under 2%.

Circuit Operation and Performance

The RF front-end of a software-based GPS receiver first amplifies the weak incoming signal with a low-noise amplifier (LNA), and then downconverts the signal to a low intermediate frequency (IF) of approximately 4MHz **(Figure 5a)**. This downconversion is accomplished by mixing the input RF signal with the local oscillator signal using one or two mixers. The resulting analog IF signal is converted to a digital IF signal by the analog-to-digital converter (ADC).

The MAX2769 integrates all these functions (LNA, mixer, and ADC), thus significantly reducing the development time for applications. The device offers a choice of two LNAs: one LNA features a very-low, 0.9dB noise figure, 19dB of gain, and -1dBm IP3, for use with passive antennas; the other LNA has a 1.5dB noise figure with slightly lower gain and power consumption, and a slightly higher IP3, for use with an active antenna. Current consumption at 2.8V is minimal, ranging from 13mA to 18mA depending on the configuration.

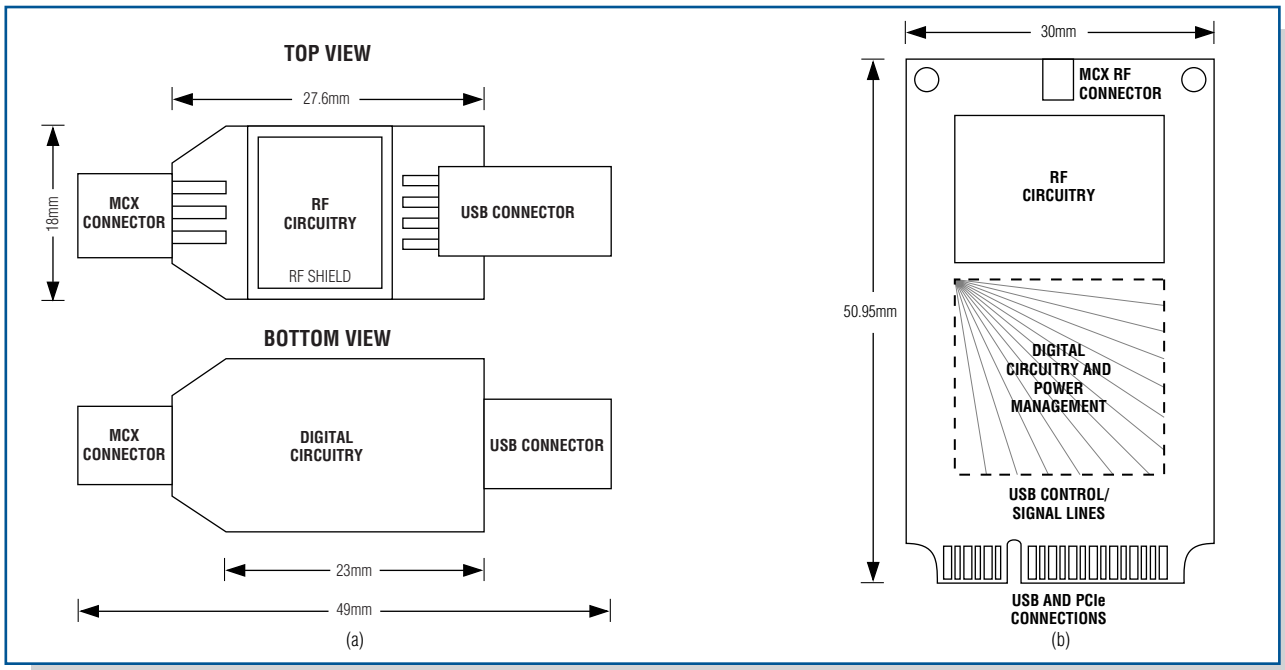


Figure 4. Typical adapter configurations for the USB dongle (a), and the PCIe Mini Card (b) show the simple, low-cost design for either option.

There is a provision for external filtering at RF after the amplifier. The signal is then downconverted directly using the integrated 20-bit, sigma-delta, fractional-N frequency synthesizer together with a 15-bit integer divider to achieve virtually any desired IF between zero and 12MHz. A wide selection of possible IF filtering choices accommodates different schemes, such as those of Galileo.

The overall gain from RF input to IF output can be tuned or automatically controlled over a 60dB to 115dB range. The output can be chosen as analog, CMOS, or limited differential. The internal ADC has a selectable output of one to three bits.

The integrated reference oscillator enables operation with either a crystal or a temperature-compensated crystal oscillator (TCXO), and any input reference frequency from 8MHz to 44MHz can be used.

A simple reference design for a USB dongle is available based on the MAX2769 with a Cypress Semiconductor USB controller, which operates from a 24MHz reference (Figure 5b). The dongle employs a single MAX8510 LDO to regulate the DC supply. A 3-wire (SPI™) digital bus is used to program the registers of the MAX2769. Alternatively, the system can be operated in one of eight hard-wired preconfigured states without SPI control.

Also on the chip are circuits to bias an active antenna and then turn off that antenna during shutdown mode for USB compliance. The MAX2769 can detect the presence or absence of a current draw on the antenna supply and automatically switch between LNA1 and LNA2. This flexibility is useful in applications that support the insertion of a more sensitive external active antenna in place of an integrated passive antenna. The designer simply wires the external antenna port to LNA2 and the internal port to LNA1. Upon insertion of the external antenna, the MAX2769 senses the current draw and automatically switches from LNA1 to LNA2.

The MAX2769 provides a high-performance, compact solution for laptops, mobile handsets, PDAs, and automotive applications. A total nominal voltage gain of up to 115dB and a module noise figure of 1.4dB provide -143dBm acquisition and -154dBm tracking sensitivity using a commercially available GPS software package.

Conclusion

Software techniques enable simple, low-cost GPS applications. To support these possibilities, the MAX2769 provides flexibility in frequency planning for software GPS receivers, as well as traditional hardware implementations. Of course, every solution has its pros and cons—software GPS receivers require a high-performance processor and moderate amounts of memory. However, with software development and the proper selection of clocking and data update rates, memory usage can be minimized.

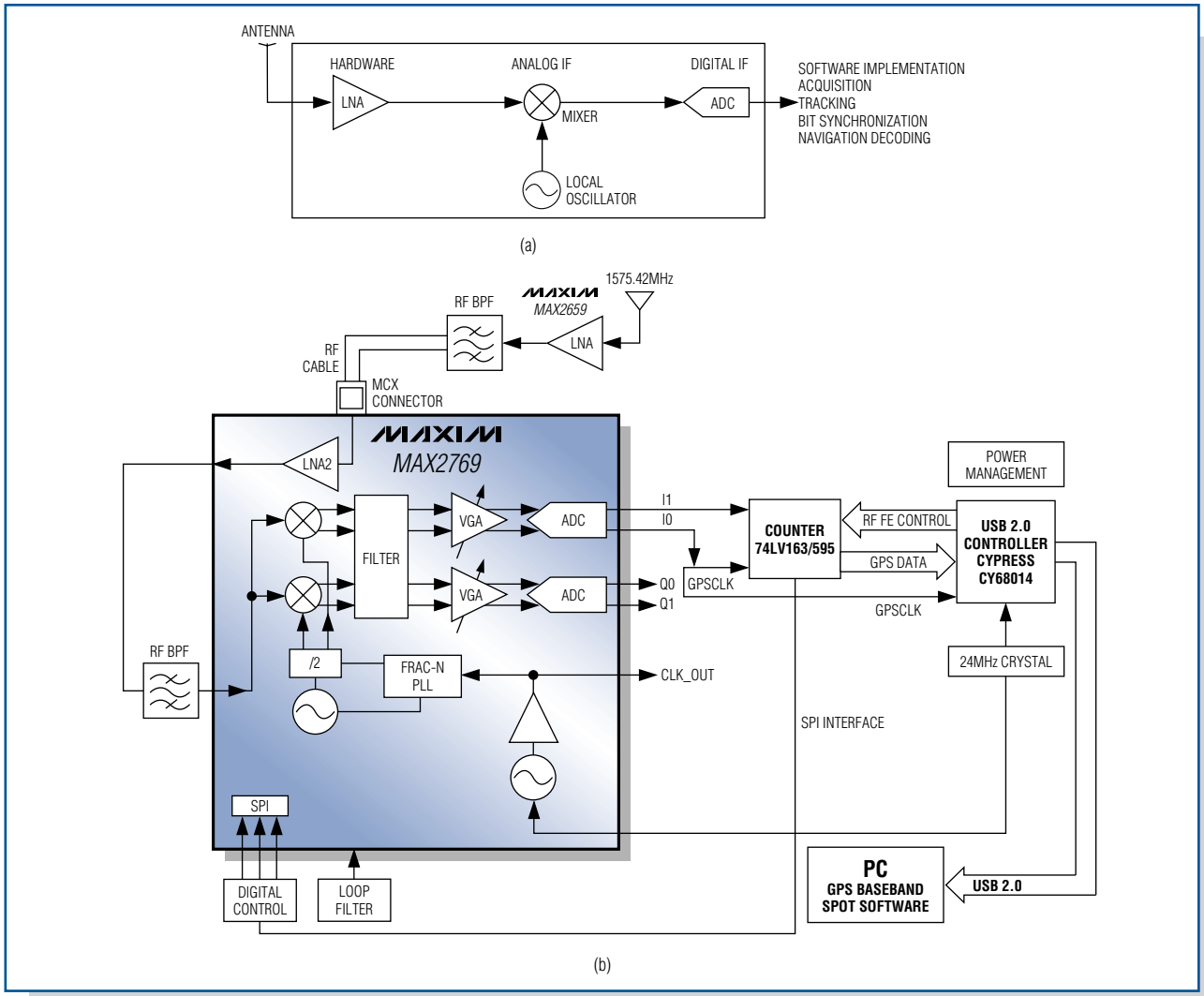


Figure 5. In a software GPS receiver the captured RF signal must be amplified, mixed down in frequency, and then digitized (a). In an actual receiver, a low-noise amplifier boosts the RF signal; then, the MAX2769 mixes it down and digitizes the signal. Finally, data travels through the counter and USB interface controller and is inserted into the USB protocol for transmission to the host PC (b).

References

1. Kaplan, E., *Understanding GPS: Principles and Applications*, 2nd ed. (Norwood, MA: Artech House Publishers, 1996).

2. Bao-Yen Tsui, J., *Fundamentals of Global Positioning System Receivers: A Software Approach*, 2nd ed. (New York: John Wiley & Sons, Inc., 2004).

3. For the decoding of navigation messages and position calculations, refer to Bao-Yen Tsui, *Fundamentals of Global Positioning System Receivers*.

4. Viterbi, A., *CDMA: Principles of Spread Spectrum Communications* (Reading, MA: Addison Wesley Longman Publishing Co., Inc., 1995).

5. Gold, R., "Co-optimal binary sequences for spread spectrum multiplexing," *IEEE Transactions on Information Theory*, Vol. 13:4 (October 1967). pp. 619–621.

6. Ziemer, R.E., and Peterson, R.L., *Digital Communications and Spread Spectrum Systems* (New York: Macmillan Publishing Company, 1985).

7. Proakis, J. G., *Digital Communications*, 4th ed. (New York: McGraw-Hill, 2000).

8. Ibid.

A similar article appeared in the December 2006 issue of Microwaves & RF.

AC-Based Continuity Tester Finds Single-Ended Faults

Kevin Bilke, Applications Engineer

This article details an AC-based continuity tester designed for frontline test and repair jobs. The circuit provides a simple GO/NO-GO test for localizing faults in multicore cables.

Open circuits are more likely at connector ends. By identifying the faulty end, the AC continuity tester allows you to open and repair the correct cable end. This action avoids the risk of damaging a good connector by opening it up. This approach is also useful for testing an installed cable with ends at different locations.

Figure 1 illustrates a circuit for a continuity tester that injects an AC signal on one cable wire and then looks for the absence of capacitive coupling on the other wires. One end of a bad cable typically shows good AC continuity, while the other end typically has one or more connector pins with no AC continuity. Because a short in the cable appears as a good connection, the operator can easily confirm that the tester is operating correctly by simply shorting its test leads together.

The left side of the circuit uses a low-power dual comparator (MAX9022) to form a relaxation oscillator operating at approximately 155kHz. It produces a peak-to-peak output signal approximately equal to the supply voltage, which is fed to a connector of the cable under test. The right side of the circuit processes any AC signal picked up by the inter-lead capacitance. That AC signal is first rectified by a pair of silicon diodes, and then integrated on storage capacitor C5. The bleed resistor (R5) provides some noise immunity and helps to reset the capacitor between tests.

The output resistor (R4) and input capacitor (C4) provide limited circuit protection. The circuit indicates OPEN for any test-cable capacitance below 100pF. (Thus, a standard 2m IEC mains test lead, whose typical lead-to-lead capacitance is 200pF, would test OK.) The circuit is also immune to false triggers caused by the 60Hz pickup from power lines.

Because the typical current draw of this low-power circuit is less than 40µA most of the time, the circuit can be powered by three AA or AAA 1.5V pencil batteries.

Many low-cost alternatives are available for the output device (a DC-activated piezoelectric buzzer in Figure 1), and most alternatives feature a suitably wide operating-voltage range. The 100nF capacitors are standard ceramic decoupling capacitors, and the circuit contains no critical passive components. The comparator's high-side drive is better than its low-side drive, so it should be used to source (rather than sink) current to the indicator device. Diodes D1 through D3 are silicon diodes.

A similar design idea appeared in the February 21, 2008, issue of EDN.

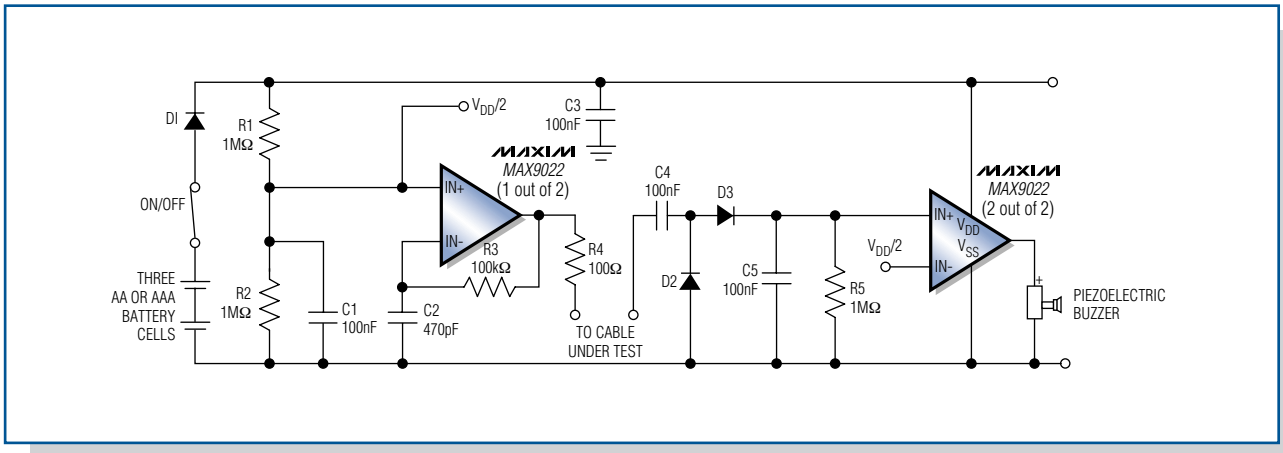


Figure 1. Based on a low-power dual comparator (MAX9022), this AC-based continuity tester locates open-circuit pins in a cable.

Two Methods for Shutting Down a Current-Sense Amplifier

By Arpit Mehta, Strategic Applications Engineer, and Prashanth Holenarsipur, Lead Product Definer

Unlike traditional operational amplifiers, high-side current-sense amplifiers do not include an internal electrostatic-discharge (ESD) protection diode between each input pin and the power-supply pin. As a result, they can operate at common-mode voltages well above the V_{CC} supply. Furthermore, pulling the V_{CC} pin of a typical current-sense amplifier to ground places the part in shutdown mode, in which it draws no quiescent current from its input pins, only a small leakage current. Thus, the V_{CC} pin of a high-side current-sense amplifier can serve as a shutdown pin.

Consider a typical battery-operated device in which a power source such as an LDO powers several ICs on a circuit board, including a MAX4173F high-side current-sense amplifier. To extend battery life by saving power, the system frequently turns off the LDOs and, therefore, the current-sense amplifier as well **(Figure 1)**.

Typically, the inputs of the MAX4173F are connected to a current-sense resistor in the power-supply line. To simulate the affect of a shutdown signal, a differential $20mV_{P-P}$ AC signal, offset by a 20mV DC signal, rides on a 10V common-mode input voltage and is applied to the part. Loss of V_{CC} is simulated by a 0V-to-5V square wave at the V_{CC} pin. During 5V intervals at V_{CC}, the amplifier operates in its active mode. During 0V intervals, however, it goes into shutdown. Because the amplifier gain is 50, the expected output is:

$$50 \times (20mV_{P-P} + 20mV)$$

Consequently, the output is a $1V_{P-P}$ sine wave offset by 1V (**Figure 2**). As expected, the amplifier is active when 5V is applied, and it produces the expected output. When V_{CC} goes to 0V, the output also goes to 0V and the device shuts down, drawing no input or supply current.

Another way of shutting down a current-sense amplifier is to connect an nMOS transistor in the ground path (**Figure 3**), and drive it with logic-level signals capable of turning the transistor on and off. When the transistor is on, the amplifier operates normally. The drain-source drop across the transistor causes negligible offset and gain error when referred to the inputs. When the transistor is off, the amplifier shuts down because its ground is floating.

The output waveform in **Figure 4** demonstrates the expected behavior: amplifying the input signal during 5V intervals, and floating close to V_{CC} during 0V intervals. During the shutdown intervals, the leakage current measured at the V_{CC} pin is just 4µA, due to the 1MΩ input impedance of the measurement scope. When the scope probe is absent, only the nMOS transistor's leakage current is drawn from V_{CC}. Input current on the RS+ and RS- pins is just 0.3µA.

Thus, one can easily put the MAX4173F in shutdown mode either by pulling its V_{CC} pin to ground, or by opening its ground connection using an nMOS transistor. The first method depends on the availability of an LDO that can be turned off in the application. The second method requires an additional external FET. Both methods are useful for managing power in next-generation, portable multimedia devices. These schemes extend battery life, while still delivering an enriched user experience. Similar results can be expected from other high-side current-sense amplifiers.

A similar design idea appeared in the November 2007 issue of Power Electronics Technology.

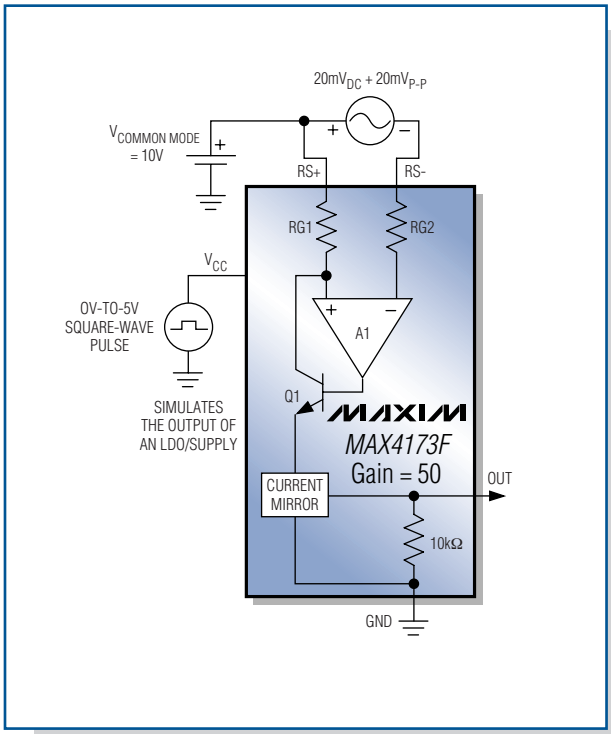


Figure 1. Zero volts on the V_{CC} pin of a current-sense amplifier (in this case, the MAX4173F) effectively shuts it down.

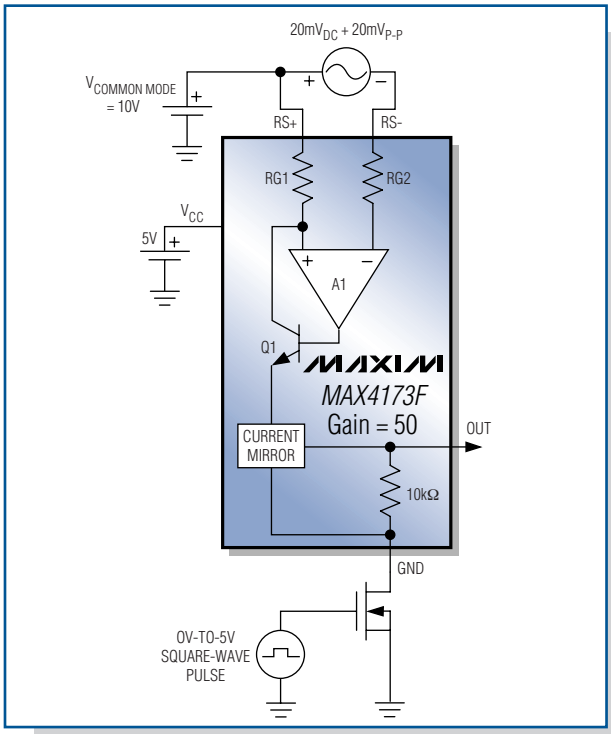


Figure 3. Opening the GND terminal of the MAX4173F also shuts it down.

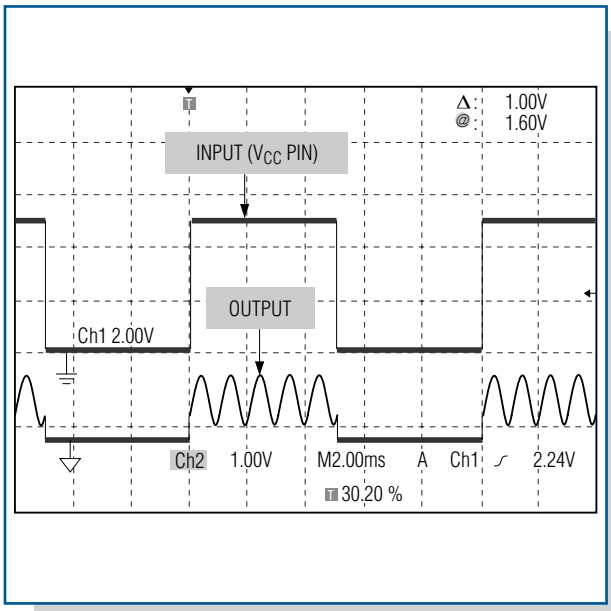


Figure 2. These waveforms illustrate the effect of shutting down a high-side current-sense amplifier using the method shown in Figure 1. The amplifier draws no quiescent current when V_{CC} is 0V.

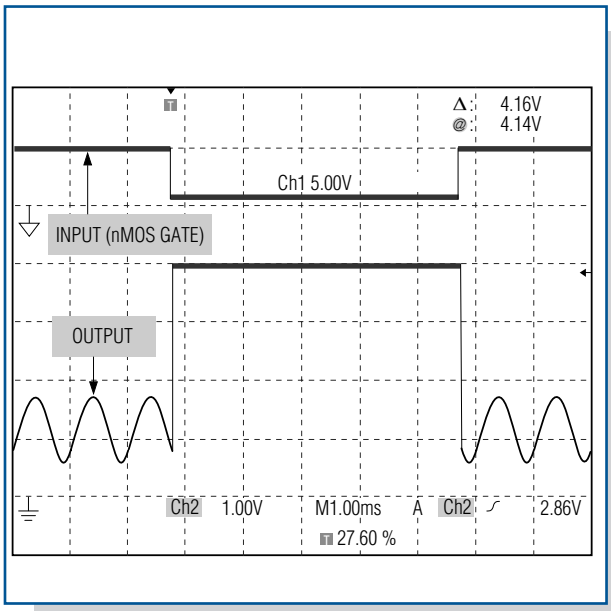


Figure 4. With the ground connection open, the high-side current-sense amplifier of Figure 3 shuts down, drawing no quiescent current.