BROADCOM®

WHITE PAPER

A Scalable Approach to Gigabit Ethernet Switch Design

06/27/02

16215 Alton Parkway • P.O. Box 57013 • Irvine, California 92619-7013 • Phone: 949-450-8700 • Fax: 949-450-8710

REVISION HISTORY

Revision #	*Date*	*Change Description*
567x_569x-WP100-R	06/27/02	Initial Release

Broadcom Corporation
16215 Alton Parkway
P.O. Box 57013
Irvine, California 92619-7013

TABLE OF CONTENTS

LIST OF FIGURES

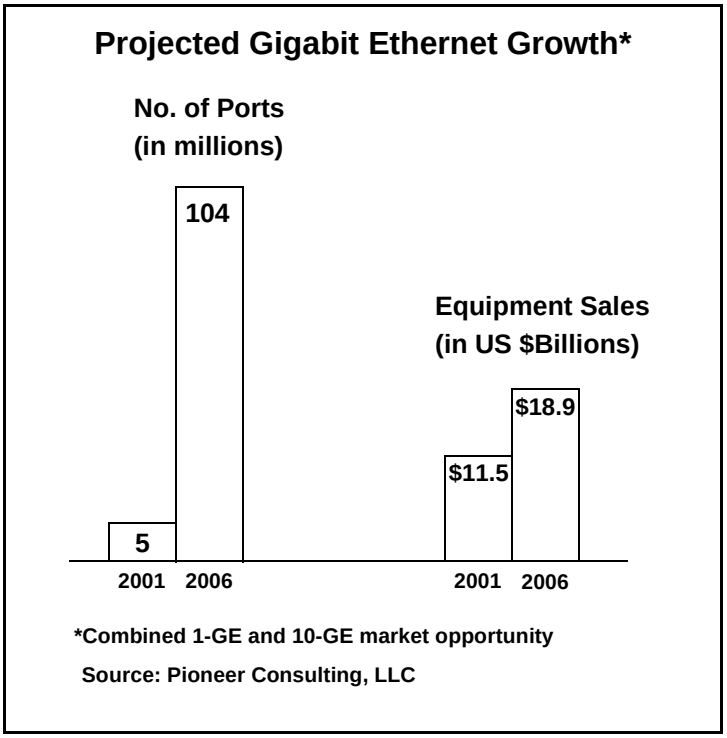
LIST OF TABLES

THE LAN SWITCHING LANDSCAPE

Enterprise network traffic volumes are accelerating quickly. As a result, organizations are seeking much faster LAN connections—to gigabit speeds and beyond—not only in their campus backbones, but also in their wiring closets for high-speed connectivity all the way to user desktops.

There are several application drivers behind the exploding volumes of enterprise traffic. Businesses are continuing to grow progressively more reliant on their corporate networks for automating their business processes. For example, Web-centric applications, often multimedia in nature, have emerged not only to reach employees across corporate intranets, but also to enable enterprises to communicate with their business partners across extranets and to serve customers via the Internet. Specifically, a few of the primary applications driving gigabit-speed network requirements include following:

- Storage-area networks, which allow multiple servers access to the same storage devices connected by a multi gigabit-speed switch
- Network-based corporate training, often involving streaming media or interactive video conferencing
- Network-based transport for PC and server data backup
- IP PBXs, which enable the convergence of voice and data on the corporate LAN



THE MARKET OPPORTUNITY

The sheer volume of traffic spilling onto corporate networks is driving Gigabit Ethernet switch requirements. The Dell'Oro Group, for example, predicts that worldwide Ethernet switch sales will increase from $11.5 billion in 2001 to $18.9 billion in 2006. The firm attributes the predicted growth to the availability of 10-Gigabit Ethernet, Gigabit Ethernet over copper, and Layer 3 Ethernet switching technologies. These market figures indicate that LAN switch makers are quickly coming under market pressure to deliver their Gigabit Ethernet systems.

DESIGN CHALLENGES

In this milieu, equipment-makers are faced with several challenges. These include being able to quickly build high-availability, high-throughput devices that can scale to meet the needs of various-sized and growing businesses. Designers also must be able to easily setup their devices in the configurations that match their target customer and application. These configurations include the following:

- Stackable switches for wiring closet applications
- Chassis-based systems for highly resilient backbone applications
- Integrated switch/servers for high-speed server interconnections
- Standalone or “fixed” switches for small/medium-sized enterprise applications
- Unmanaged switches for small environments with no IT support

Another prevalent system challenge is to minimize design, assembly, and manufacturing costs. Methods for lowering total costs are by constructing a switch using a minimum number of components, reusing chip resources and development expertise, and using a simple board assembly.

And what about system features? To satisfy end customer requirements, switch makers must consider including most, if not all, of the following capabilities in their products:

- Layer 2 switching with associated 802.1Q Virtual LAN (VLAN) support and 802.1p traffic classification capabilities
- Layer 3 routing
- Quality of service (QoS) for prioritizing certain traffic, such as latency-sensitive voice over IP (VoIP) packets
- Layer 2 and 3 multicasting for bandwidth-efficient point-to-multipoint transmissions
- Very high levels of resiliency and switch uptime to ensure network reliability
- A high-throughput backplane that enables wire-speed performance
- Scalability—including high port density in a small amount of board space—so end customers can grow their networks as needed at a reasonable cost per port
- Security, such as per-port MAC address filtering, firewall filtering, and intrusion detection capabilities
- The ability to integrate features such as QoS, trunking, and port mirroring (for diagnostics) across interconnected modules to ensure consistent performance and resiliency, end to end

All of these capabilities are supported in the *Broadcom StrataXGS™ Switch Architecture* from Broadcom Corporation. The sections that follow examine the components and attributes of this modular and highly integrated architecture. The following sections will also discuss the form factors that can be designed with the StrataXGS family—often, in conjunction with members of Broadcom's StrataSwitch II™ Fast Ethernet (10/100 Mbps) family—and the important decisions and design considerations that accompany building these devices.

StrataXGS Gigabit Ethernet Architecture

Reduced Components

The Broadcom StrataXGS architecture has been designed so that only a small number of components are required to build a variety of Layer 2/3 systems in different form factors that support a range of port densities. StrataXGS components support a high level of integration, with all memory, interfaces, packet forwarding and other capabilities supported on-chip. The integrated nature of the architecture delivers approximately a 75% reduction in the number of chips required over competing solutions, which results in lower manufacturing costs. The reduced component count also yields higher-quality products because of the decreased complexity; Broadcom estimates that the StrataXGS solution reduces mean time between failures (MTBF) by about 20%.

Interoperability with Broadcom StrataSwitch

The StrataXGS architecture is an extension to the proven Broadcom StrataSwitch II (Fast Ethernet) architecture. There are already 20 million StrataSwitch ports deployed in the marketplace today. StrataXGS is compatible with StrataSwitch in several ways. First, the packet flows from ingress to egress is the same. Secondly, memory management is consistent across the two architectures. Lastly, use of the fourth-generation *Broadcom Switch Application Programming Interface* (API) for enabling custom features is portable across both families.

The integrated nature of the two architectures enables designers familiar with Broadcom StrataSwitch to re-use their engineering and development expertise, which lowers costs and expedites time to market. Switch-makers can also use both the StrataSwitch and StrataXGS chips as building blocks in a common design to mix Fast Ethernet, Gigabit Ethernet, and 10-Gigabit Ethernet ports in standalone/fixed, chassis, stackable, server, and unmanaged switch form factors.

With the StrataXGS family, Broadcom has improved the StrataSwitch architecture in several ways:

- The depth of IP address tables and routing information has been significantly increased.
- Switches that can store more route entries can support a greater number of attached devices.
- StrataXGS family of components uses a widely used approach called *hashing*, which increases table lookup performance compared with binary search capabilities.
- Broadcom has built a special 10-Gbps chip-to-chip interface, which improves Scalability and performance.

QUALITY OF SERVICE

StrataXGS chips offer extensive support for packet classification, marking, and prioritization to enable QoS support. QoS on the LAN is increasingly a checklist item for enterprise customers, who are beginning to run latency-sensitive applications such as VoIP locally. End customers might also want to ensure the integrity of high-priority data applications and enforce access control lists (ACLs).

By leveraging the HiGig interface, class-of-service markings and QoS can be integrated across a stack of chips used to build a system so that markings are not lost when packets are passed from one component to another. All StrataXGS chip ports support eight CoS queues into which traffic with different priority weights can be placed.

HIGIG™ INTERFACE

New to the Broadcom product family is a 10-Gbps, full-duplex chip-to-chip interface that enhances system scalability and performance, called the *HiGig* interface. It adds an 8-byte header to the Layer 2 GE frame as it exits the interface. This header contains information about the packet, its source and destination ports, and port mirroring. This information hastens table lookups when two Broadcom components are connected together, thus improving overall system performance.

The HiGig interface allows the interconnection of several StrataXGS chip modules to form systems with different port densities. Communication between the CPUs, a process called *stacking*, is performed in-band via the HiGig header using special master/slave protocols developed by Broadcom. Link aggregation, port mirroring, and class-of-service markings are fully supported across the interconnected modules at wire speeds.

INTEGRATED SERIALIZER/DESERIALIZER (SERDES)

A SerDes interface—which converts data from a parallel bus into a serial stream—is integrated onto each port of each StrataXGS chip. This design precludes connections to external SerDes chips, simplifying design.

The integrated SerDes support, CoS functions, Broadcom API, and other StrataXGS features are described in more detail in the section, “StrataXGS Features and Benefits in Depth.”

STRATAXGS PRODUCT FAMILY

The two primary chip components of the Broadcom StrataXGS Gigabit Ethernet switch architecture are the *StrataXGS BCM5670* and the *StrataXGS BCM5690*. The BCM5670 is a high-speed (160-Gbps) *switch fabric,* also called a *backplane* or *interconnect*, that has eight 10-Gbps ports. A smaller version, the BCM5671, is an 80-Gbps switch fabric with four 10-Gbps ports. The 10-Gbps ports on the BCM5670 and BCM5671 are the building blocks for interconnecting line cards to a switch fabric. A 10-slot chassis, for example, can be designed with eight line cards and two management cards.

The 10-Gbps ports on the BCM5670/71 accept traffic from the 10-Gbps uplinks in other chip components, these can be arranged in a variety of configurations to build a Gigabit Ethernet switch of the size and configuration that matches the designer's business plan. Once traffic is placed on the switch fabric, it is switched, with no delay, to its egress port.

Switching with no delay is possible because the total bandwidth available on the switch fabric is greater than the sum of the speeds of all the ports feeding traffic onto it. Therefore, no buffering queues are required that could introduce added delays. This design approach is called a *nonblocking* architecture. Combining chips expands the number of ports in a system while also retaining the nonblocking throughput characteristics of a single switch.

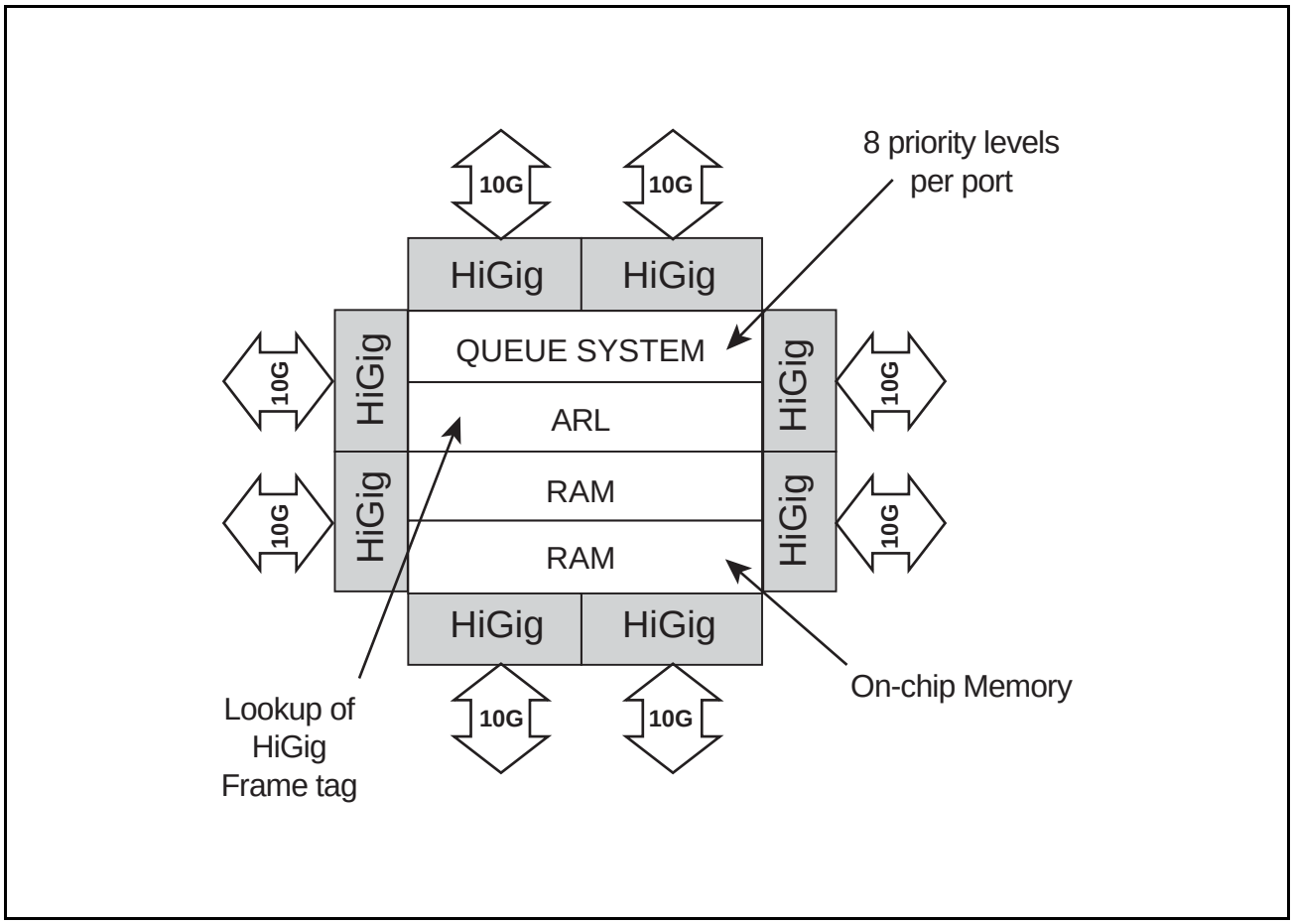


Figure 1: BCM5670 Block Diagram

The other primary StrataXGS component, the BCM5690, is a switch-on-a-chip. It supports twelve 1-Gbps ports and a one HiGig (10-Gbps) interface. The BCM5690 can be used in a standalone fashion to build a fixed 12-port switch, or it can be combined with other BCM5690s, BCM5670s, or StrataSwitch components to form different-sized configurations and to achieve higher port densities. For example, eight BCM5690s can combine to scale from 12 to 96 GE ports.

Designers can link as many as 32 XGS chips (not including the switch fabric) for a maximum configuration of 384 GE ports (ports using Broadcom components) and 64 (32 with BCM5673) 10-GE ports. One can also create a system containing 768 Fast Ethernet using X-StrataSwitch and Y-XGS components.

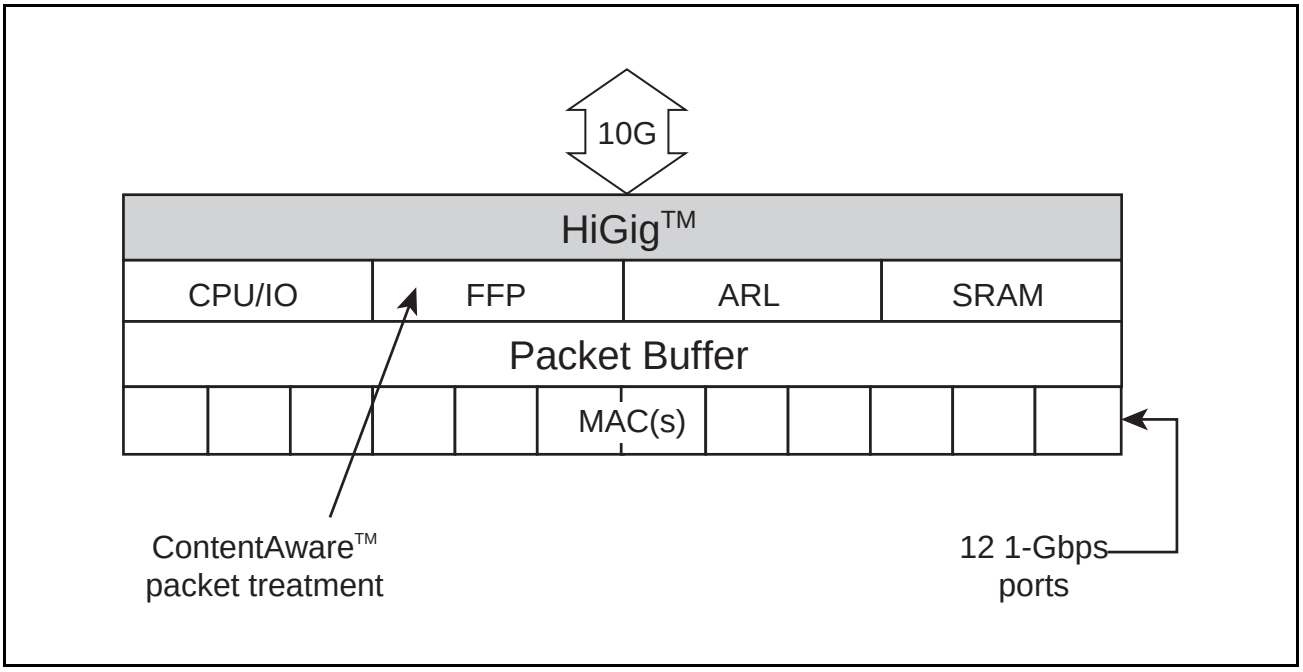


Figure 2: BCM5690 Block Diagram

Another product, the *BCM5691*, supports 12 GE ports but no uplink, for use in standalone configurations. Both the BCM5690 and BCM5691 support a standard PCI bus for moving signals to and from the CPU and, in the case of the BCM5690, to a PCI uplink.

StrataXGS chips are highly programmable, in that system designers can add value and differentiate their companies' products by activating certain features with minor development work in software. Using the Broadcom API, new features are easily added to the existing chip.

Table 1: StrataXGS Chips

Product	***Function(s)***	***No. and Type of Ports Supported***
BCM5670	160-Gbps switch fabric. It is nonblocking, allowing all ports to operate concurrently at 100% utilization.	Eight 10-Gbps ports
BCM5690	• Switch-on-a chip • Can be combined back-to-back with another BCM5690 or to a BCM5670(s) to build bigger systems and achieve higher port densities • Eight BCM5690s can combine to scale from 12 to 96 GE ports • Supports both Layer 2 and Layer 3 capabilities	• Twelve 1-Gbps ports, each running at 10/100 full- and half-duplex or at 1 Gbps full-duplex speeds • One 10-Gbps uplink
BCM5671	80-Gbps switch fabric. It is nonblocking, allowing all ports to operate concurrently at 100% utilization.	Four 10-Gbps ports
BCM5691	• Switch-on-a chip • Supports both Layer 2 and Layer 3 capabilities	Twelve 1-Gbps ports, each running at 10/100 full- and half-duplex or at 1 Gbps full-duplex speeds

WHAT TO BUILD AND HOW TO BUILD IT

There are many types of Gigabit Ethernet switches that a designer might elect to build, depending on the enterprise applications the switch vendor wishes to support. A switch vendor might focus on building a particular type of system that addresses the needs of a specific customer environment or application. A vendor might design and manufacture a family of switches to satisfy a range of business customer applications, such as a large enterprise's wiring closet, campus backbone, and small branch office.

TYPES OF SYSTEMS

List below, are some of the types of switches that a switch vendor can design with the XGS Product Family

Stackable Switches. These connect user desktops to the high-speed campus backbone. These highly scalable switches contain some number of Gigabit Ethernet or Fast Ethernet "user" ports. Traffic from these ports is aggregated over a HiGig uplink for connectivity to corporate data resources in servers and, increasingly, to call servers or *IP PBXs* for VoIP applications. Stackables are often found in the wiring closet network segment to accommodate a growing user base. As additional users join the network, more ports can be added by "stacking" additional switches on top of one another with a common external interconnect; in effect, creating one big integrated switch.

Chassis-based Switches. This form factor is generally deployed in highly resilient backbone applications. These systems contain slots that can be used for redundant line cards, including system buses, power supplies, and security. They can be configured with redundant components to the degree deemed appropriate for the redundancy and manageability required for the installation.

Integrated Switch/Servers. This configuration combines application and data servers with network switches, thus allowing very high-speed server interconnections. In this configuration, *server blades* are generally interconnected across a GE switch backplane. SerDes support in StrataXGS chips plays a large role in this application, enabling high-speed connections across the backplane. The integrated switch/server form factor generally appeals to end customers looking for very high-speed server connectivity while wishing to collapse functions into a single device to reduce capital costs and management complexity.

Standalone or "Fixed" Switches. These are most often used in small/medium-sized enterprise applications, where high growth is not expected. These systems arrive with a fixed number of ports. In order for the customer to accommodate additional users beyond the number of ports in that switch, additional systems must be purchased.

Unmanaged Switches. This form factor represents the simplest of the devices and operates only at Layer 2, which reduces software development and makes use of low-cost CPU interfaces. These switches are generally installed in small environments with no IT support and where packet classification/prioritization for VoIP or high-priority data applications is not required.

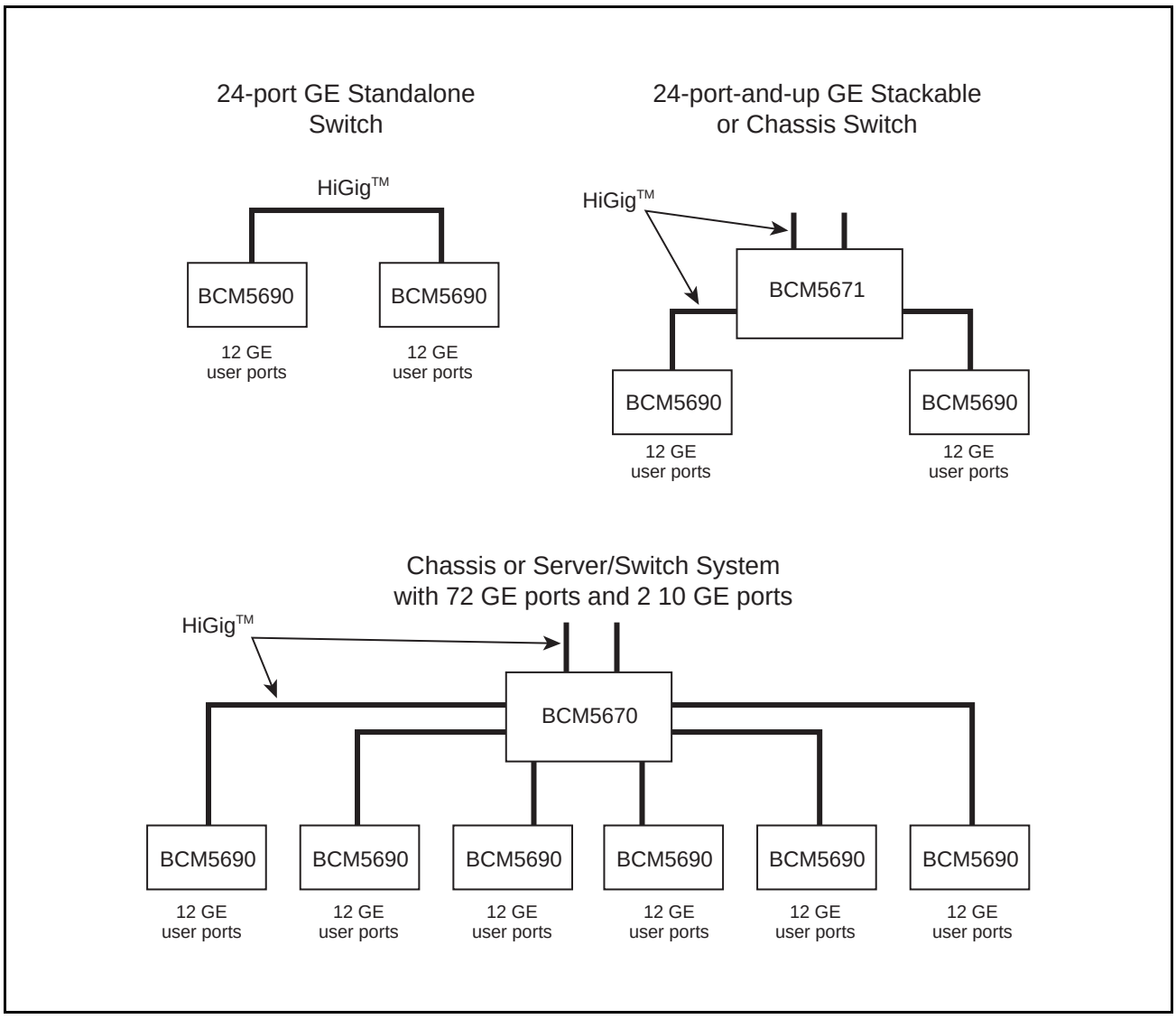


Figure 3: Enterprise LAN Applications: Scalable Architecture for Building Different Types of Systems

Note **Having common building blocks that can be used to build different types of switches provides a high ROI on software investments and an expedited time to market.**

INTERCONNECT DECISIONS

For each Gigabit Ethernet switch form factor, designers require GE chips, physical-layer interfaces (including cabling and connectors), a CPU management system, a power supply, and an operating system for activating and customizing features. From there, some of the decisions the switch vendor/designer must make include whether the switch should be a single-chip or multiple-chip solution, this largely depends on the number of ports to be supported. For multiple-chip solutions, designers must decide what type of interconnect should be used to link the components and pass packets between them.

The choices of interconnects include *cascade*, *matrix*, and *daisy chain* options. StrataXGS cost-effectively supports all these configurations, which are described in more detail below.

Cascade. This is the simplest interconnect choice, most often used for low-end standalone/fixed systems. In this configuration, chips are connected in a shared-ring fashion; the transmitter of one module is connected to the receiver of the next module, and so on. The transmitter of the last module is connected to the receiver of the first module.

The merits of this interconnect choice are that it is the least expensive and uses few resources; ports, backplanes and other resources are all shared.

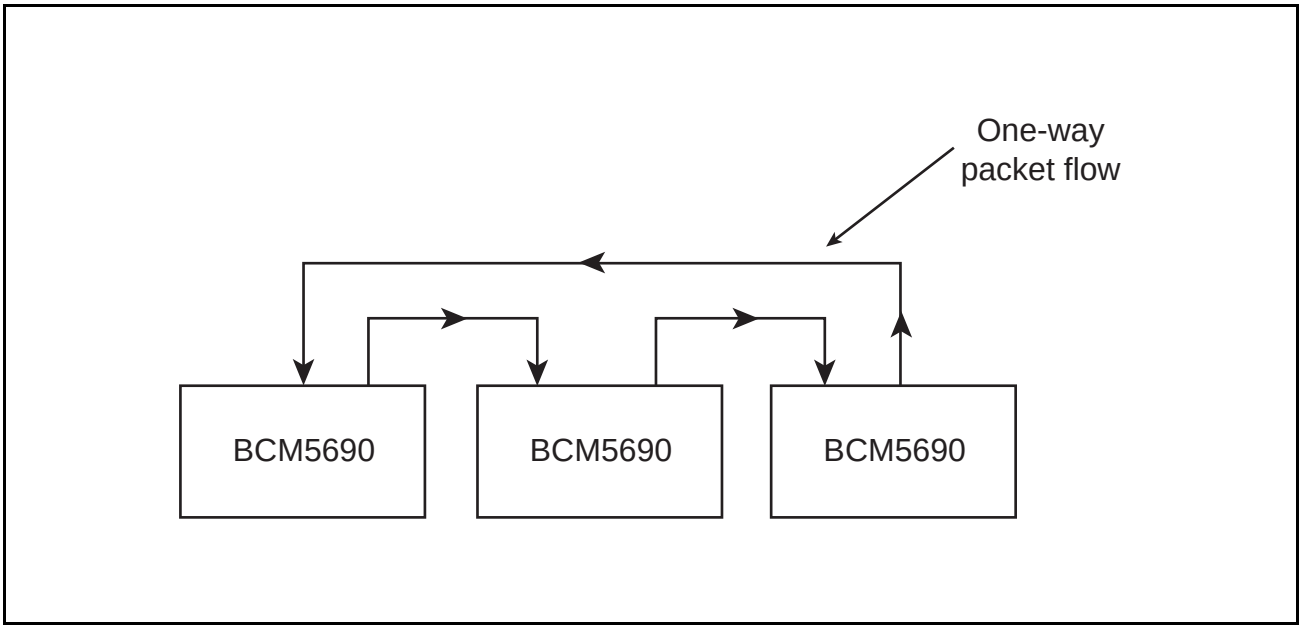


Figure 4: A Simplex Interconnect

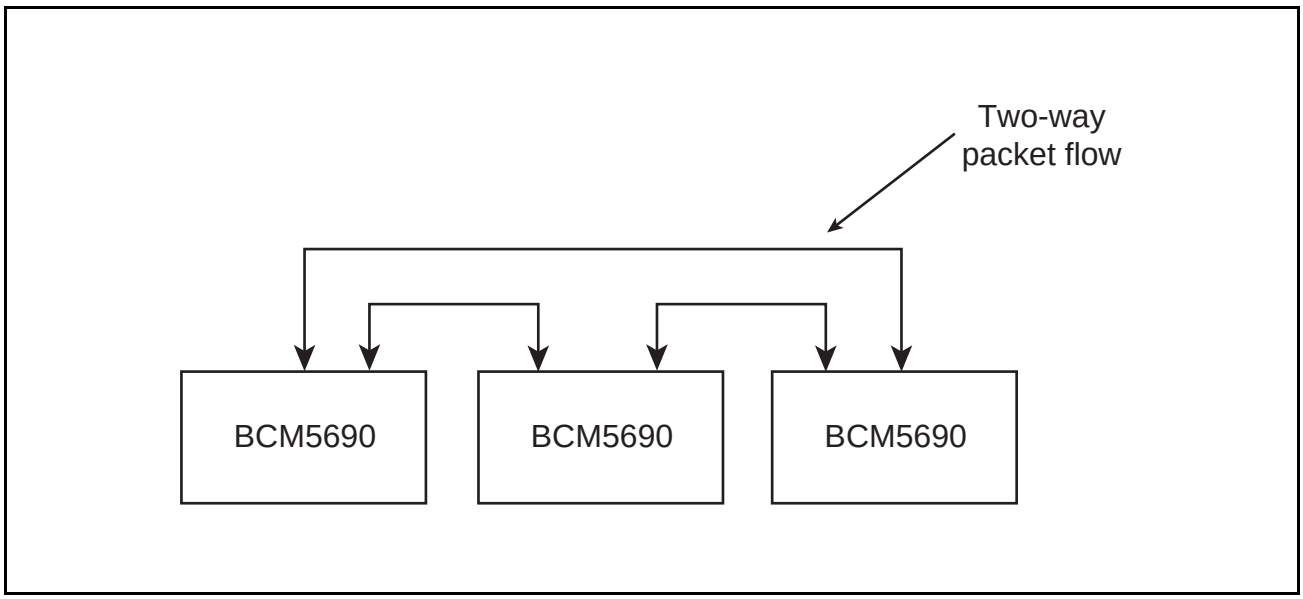


Figure 5: A Full-duplex Interconnect

Matrix. This type of interconnect is typically used for chassis-based systems and higher-end fixed systems. In this configuration, switching components are one hop away from the central processor that controls all the switches. Here, a switching fabric (BCM5670 or 5671) serves as a backplane to multiple BCM5690s that are connected to it. For resiliency, a matrix-based system can be configured with a redundant central processor, with switching chips interconnected to both processors in a full-mesh configuration. This interconnect approach adds reliability to GE systems.

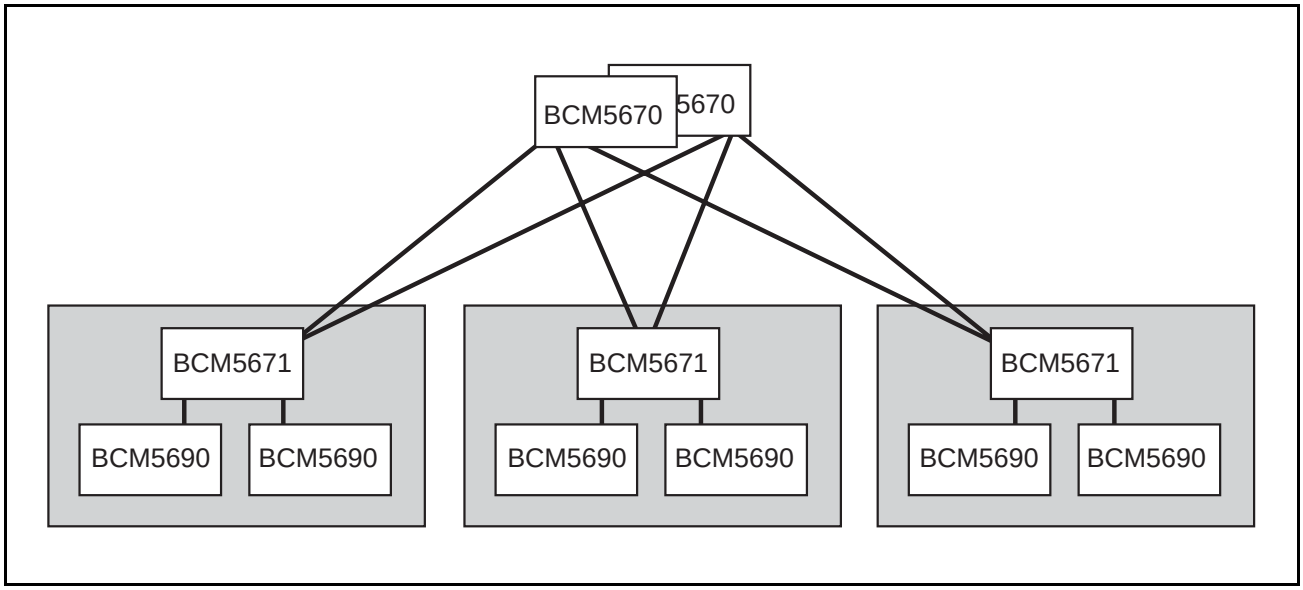


Figure 6: A Matrix Interconnect

Daisy Chain. In a daisy chain configuration, switches are interconnected to one another in a series. This modular configuration is highly scalable, as there is virtually no limit to the number of incremental ports and central management CPU

power that can be added. These characteristics are most appropriate for wiring closets in enterprise organizations where growth is expected or unpredictable. Each daisy chain-connected switch is typically built using BCM5670s with multiple BCM5690s connected to them.

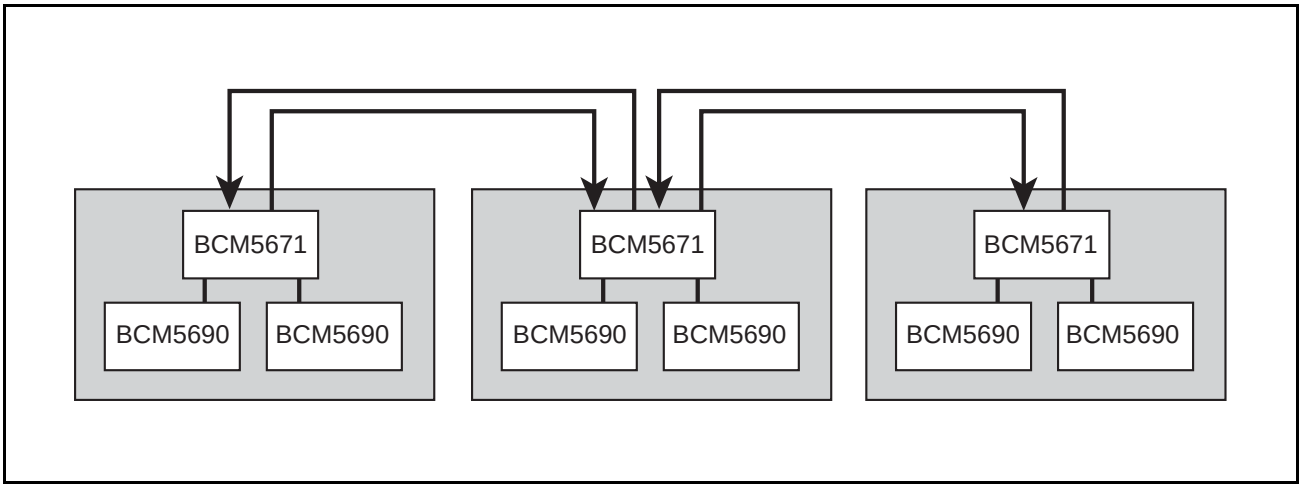


Figure 7: A Daisy Chain Interconnect

Table 2: Summary of System Interconnect Considerations

	System Candidate(s)	***Performance***	***Comparative Cost***	***Benefits***
Simplex	Fixed, low-end	Low	Low‘	Low-cost, simple
Duplex	Chassis	Low, but could boost with the addition of a BCM5680 in the transmission path	Low	Low-cost, simple, with path redundancy for added resiliency
Cascade	Fixed, low-end	Low	Low	Cost-effective
Matrix	Stackable, Fixed, Chassis	High	High	Very high resiliency
Daisy Chain	Stackable	High	Medium	Very scalable, nonblocking

Switch vendors have a variety of choices for an interconnect configuration, depending on the type of system being built and the vendor’s priorities for performance, management, and cost.

OTHER SYSTEM DESIGN DECISIONS

In addition to deciding among the various types of component interconnects, other decisions must be made for each type of switch to be developed. The primary ones are outlined below.

FIXED SYSTEMS

The designer needs to make design choices if the system is to have more than 12 ports. If there are 12 or fewer ports in the system, the designer can simply use a BCM5691 as a standalone switch.

To achieve larger port densities, the BCM5690 can be combined with additional BCM5690s or BCM5670 switch fabrics. At this point, the designer must decide if the switch should be optimized for cost or for performance.

If the designer decides that the system cost is the most important factor, a cascade interconnect for linking the chip componentry should be chosen, and the BCM5690s will be linked without the benefit of the BCM5670 switch fabric in a shared-bandwidth configuration.

If the designer decides that system performance is the most important factor, a matrix interconnect for linking the chip componentry should be chosen. In addition, a BCM5670 (or BCM5671) switch fabric will likely be included in the configuration so that packets are transmitted over a switched, rather than a shared, backplane.

STACKABLE SYSTEMS

Here, the designer must decide on a per-switch port count, then choose among the optional modules for the features and capabilities to be activated.

For the highest-redundancy levels in the stack, designers are likely to choose a matrix interconnect; for highest performance, a daisy-chain interconnect is recommended.

Which cabling medium to use is another decision to be made when building stackable switches. System vendors can choose fiber cabling or various categories of copper cabling, and gigabit interface converters (GBICs) are available for converting electrical copper signals to fiber-optics. Designers also have the option of using a backplane trace to connect the switches to one another directly.

MODULAR CHASSIS

When designing these backbone devices, a switch vendor must decide how to balance the redundancy of the backplane for resiliency and uptime with the performance of each slot in the system. For redundancy, designers can choose to connect the two HiGig connections in each blade to two separate switch fabrics in a master/slave configuration. Alternatively, the designer can connect each slot in a full mesh configuration, rather than using a switching fabric. This is a highly redundant setup; however, it does cause some performance degradation on each slot. From a performance perspective, if the system vendor elects to configure two switch fabrics in a master/slave setup, Broadcom recommends running both switch fabrics concurrently, rather than leaving one to sit idle in standby mode for failover.

From a management perspective, in the chassis system, the switch vendor must decide whether to run a centralized management blade or to distribute management capabilities across individual blades. The latter choice is often more scalable, as switch makers can add more complex management features incrementally to the system as required.

INTEGRATED SWITCH/SERVERS

In these systems, the maximum port density is usually not greater than 24 ports. The decisions here are the same as those of a fixed system – whether to interconnect switching chips with a cascading or matrix interconnect.

UNMANAGED SYSTEMS

These systems are generally fixed systems that are installed in slow- or non-growth environments. Again, the decisions here are port count and which interconnect type to use.

SYSTEM DIFFERENTIATION

Depending on how the vendor programs the chips, system designers can derive different values out of the StrataXGS family. For example, they can choose to activate and customize a number of routing, security, and voice features. To reduce componentry and development time, designers can use an Advanced Function Module, a network processor that is software-upgradable, to activate server load balancing, WAN interfaces, firewall functions and VoIP call server capabilities across all the components in a given system (see Figure 12, in the “Sample System Configurations” section).

BUILD OR BUY?

Another decision facing system vendors is whether they should develop their own software operating system for customizing and activating features or whether they should buy it. Those who wish to purchase third-party software can acquire Broadcom API-compatible operating systems from Broadcom partners RADLAN and Wind River Systems Inc.

If the system vendor elects to build the software, the Broadcom Switch API reduces development time and effort by shielding upper-layer network applications and operating systems from the Layer 1 interfaces. Using the Broadcom API enables easy migration between switch products and generations (see “Broadcom Switch Application Programming Interface (API)” on page 14).

STRATAXGS FEATURES AND BENEFITS IN DEPTH

As mentioned, the StrataXGS architecture is modular, yet integrated, to allow system vendors to develop Gigabit Ethernet switch products optimized for different market segments using the same components and familiar expertise. For additional flexibility, the StrataXGS integrated circuits can be mixed and matched with Broadcom StrataSwitch Fast Ethernet chips, which support 24 Fast Ethernet (10/100-Mbps ports) with two GE uplinks.

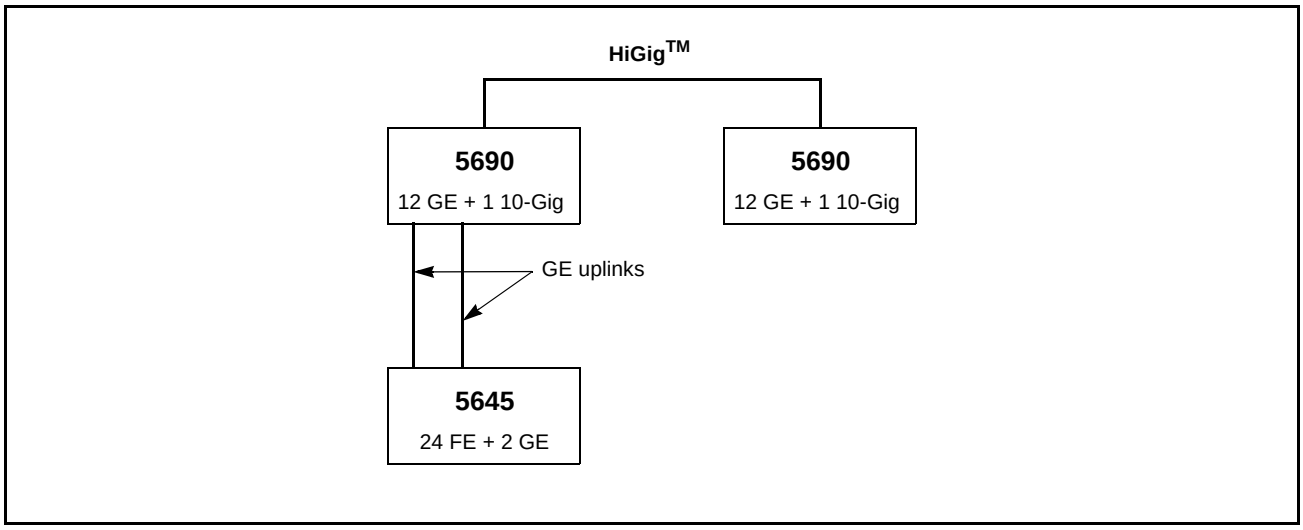


Figure 8: 24 FE Port + 22 GE Port Applications

EXPEDITED TIME TO MARKET

The architecture of the StrataXGS product family as described earlier was developed to enable system designers to bring feature-rich systems to market quickly. The architecture enables this by giving designers a wide range of flexibility in the type and size of systems they design while at the same time keeping the number of components required to do so small. Reducing the chip count required for system design is a function of a high level of integration on each chip. Memory, SerDes, XAUI, packet forwarding, and other capabilities are all collapsed onto a single component, reducing the complexity of board design. The architecture also simplifies development tasks with the structured and portable Broadcom Switch API, described below.

BROADCOM SWITCH APPLICATION PROGRAMMING INTERFACE (API)

The Broadcom Switch API, which comes free of charge with the architecture, enables systems designers to layer applications and their own software operating systems over the API without having to create special interfaces for each of the underlying chip components. The structured API is a system-level software interface that separates function calls from the driver component, so drivers do not have to be modified when a new function call is added. Developers, can then write one command that can be driven to multiple components, while retaining the flexibility to exert chip-level control. Chip components can thus be used ubiquitously, saving significant development time. System designers can quickly activate or disable the features using the API, which is portable to Unix, Linux, and Wind River Systems' VxWorks operating systems.

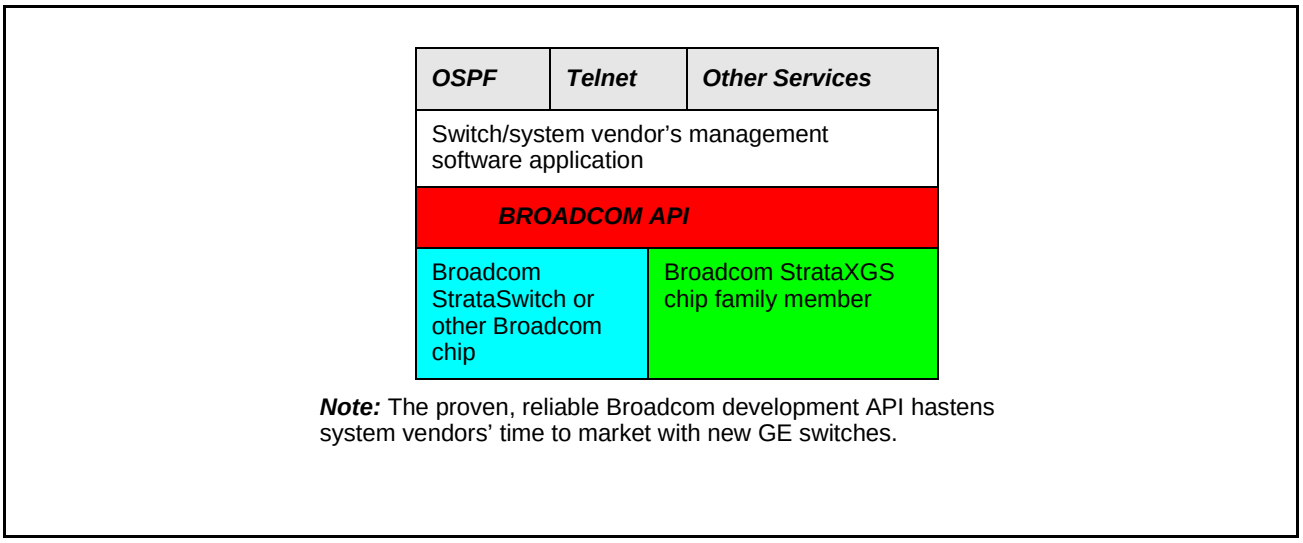


Figure 9: The Broadcom API at Work

INTEGRATED APPROACH TO STACKING

As mentioned, StrataXGS products provide a stacking solution for connecting modules across a HiGig interface in a tightly integrated way that extends a number of powerful features across all the components so that no functionality is lost. The XAUI-based HiGig interface extends the reach of features at wire speeds across both switch fabrics and cables. These features, for example, include trunking, port mirroring, VLAN membership, QoS, and a single IP address for unified management. In addition, multicast packets are distributed efficiently across the stack without replicating the packet at the source.

By contrast, stackables built with competing products today have loose connections across the set of switches and do not offer the integration of these services across the modules. For example, in stackable switches built on competing architectures, QoS mechanisms enforced at the interface level on a switch are lost when the frame is placed on the stack media. This is not the case with StrataXGS.

The integration is achieved by the 8-byte HiGig tag, which carries information relevant to trunking and port mirroring across the stack and is added to each packet that traverses the stack across HiGig interfaces. A single CPU and command line interface (CLI) are used for management across a stack.

In addition, new advanced discovery and management protocols are available for stackable systems. These protocols, Master Discovery Protocol (MDP) and Reliable Data Protocol (RDP), designate a master device for control and management in a stack of chips. They also define a backup master, should the primary master fail. Both MDP and RDP run on top of the Broadcom API.

RESILIENCY/REDUNDANCY AT THE PORT LEVEL

Enterprise customers are requiring near-100% uptime in their network equipment now, as data applications grow increasingly mission-critical and voice traffic (for which businesses have virtually no tolerance for downtime) joins the LAN. There are several ways that the StrataXGS architecture addresses this requirement. First, the BCM5670 has a redundant link so that system vendors can connect different ports across multiple paths.

Designers can also build a dual-chip configuration and trunk ports between two StrataXGS chips. When both connections are up and running, traffic can be shared between the two connections. However, if one connection should fail, traffic will automatically fail over to the live chip. This function is in full compliance with the IEEE 802.3ad specification for trunking.

Port mirroring is also supported; activity on a particular port can be copied onto another designated port, to which a sniffer can be attached for debugging and running diagnostics. This enables problems to be addressed without having to take any components out of commission while doing so, contributing considerably to system uptime.

PERFORMANCE

The StrataXGS architecture, as mentioned, allows for nonblocking, wire-speed performance because the aggregate bandwidth in the backplane is greater than the sum of the incoming ports. Activating any number of intelligent features, such as QoS, does not affect performance, which is 32 million packets per second (pps) across the backplane.

Another feature that contributes to high-end performance is support for dual trunking, which enables traffic load sharing. StrataXGS also prevents head-of-line (HOL) blocking to boost throughput. To prevent HOL blocking, when an egress port is congested because of Class of Service (CoS) markings, packets destined to that port are dropped at ingress.

Flow control mechanisms are also at work to enhance performance. The system can send out “pause” frames when the number of packets for a given port exceeds a predefined threshold. These features also are contributing factors to QoS.

TRAFFIC PRIORITIZATION AND MANAGEMENT

IETF Differentiated Services (DiffServ)-compliant CoS capabilities enable systems built on the StrataXGS architecture to mark Differentiated Services Code Point (DSCP) bits in order to classify a frame, then take action on the frame based on that class. The Broadcom *ContentAware*™ classification engine in the StrataXGS chips allows for wire-speed Layer 2 through Layer 7 classification and management. This means that treatment of packets through the switch can be determined based on the application type.

The Broadcom-patented *Fast Filter Processor™ (FFP)* enables switches to take action on the frames according to “if/then” scenarios. For example, a specific packet could be dropped, have its priority changed, or be steered to a specific port number in the event that certain network conditions exist. All the packet inspection, filtering, trapping, modification, and steering performed by the FFP takes place at wire speed. FFP-enabled packet treatments require some customization on the part of the switch designer.

QoS capabilities in the StrataXGS family also include metering/rate limiting to put a ceiling on the amount of network bandwidth that any one application, protocol, or user can consume. The StrataXGS components can inspect a transmission up to 80 bytes deep at wire speed and can control bandwidth on each port to a 1 Mbps granularity. Systems vendors can write software specifying what frames to recognize and what action to take on those frames. StrataXGS supports QoS-awareness of all enterprise network components, including 802.11x wireless access points and VoIP phones that may be connected to an Ethernet segment.

Packets destined for the CPU can have a separate priority than the rest of the ports; four priority classes are defined. Packets headed to CPU can be selectively enabled or disabled.

LAYER 2 FEATURES

The StrataXGS supports Layer 2 switching capabilities in hardware for greater performance. These capabilities are, auto-learning MAC addresses, address lookup, and packet forwarding. StrataXGS GE switches can support 4096 VLANs, 256 Spanning Trees, 256 Layer 2 multicast sessions, and 16384 Layer 2 switch table entries.

LAYER 3 FEATURES

At Layer 3, for increased performance, the hardware handles routing table lookup and forwarding. StrataXGS GE switches can support 4096 Layer 3 routing table entries and 512 virtual router interfaces.

SECURITY

StrataXGS supports the IEEE 802.1x authentication framework for port-based network access control. Also, per-port security is offered to avoid denial-of-service (DoS) attacks. This is achieved by denying access to certain MAC addresses by programming bits in the routing table to be rejected.

Table 3: How Features[1] are Implemented

Switch Function	***Broadcom Fast Filter Processor***	***Chip/API Support (minimal initialization)***	***Software programming by switch designer***
VoIP call server (IP/PBX)	√		√
Layer 2 switching		√	
Layer 3 switching (routing)		√	√
Billing and accounting		√	√
Bandwidth on demand		√	√
Policy management		√	√
DiffServ	√	√	√
802.1p/Q	√	√	√
Remote Monitoring (RMON)		√	√
Mirroring		√	
Trunking		√	√
IP Multicast		√	√
Per-port security		√	√

1.Basic features, noncustomized

Note **The many rich features supported in the StrataXGS architecture are activated in different ways. Most require some custom programming in addition to basic support in a StrataXGS chip.**

INTEGRATED SERDES INTERFACE

The integration of a high-speed XAUI-compatible SerDes interface on each port reduces the complexity of board layout, because it eliminates the need for a component to connect to an external SerDes chip. The SerDes interface connects to both copper and fiber media. On StrataXGS 1-Gbps ports, SerDes runs at 1-Gbps. On StrataXGS 10-Gbps ports, four serial channels, each running at 3.125 Gbps, are supported. SerDes delivers reduced signal traces, conserves board space, consumes lower power than parallel interfaces, and delivers superior electromagnetic interference (EMI) performance (8b/10b).

When StrataXGS 1-Gbps ports with integrated SerDes can connect to copper transceivers, they do so using Serial Gigabit Media Independent Interfaces (SGMII), which reduces pin count to four, down from 25, compared with using Gigabit Media Independent Interfaces (GMII). The reduced pin count lowers costs and simplifies design.

SWITCH DEVELOPMENT KIT AND EVALUATION SYSTEM

Broadcom will supply reference boards for a fully functional switch for evaluation and to jump-start a system vendor's development process. The BCM95690 SDK, for example, is available as a complete 24-port gigabit Ethernet Layer 2 – 7 SDK, and evaluation system based on the BCM5690 switch-on-a-chip.

The SDK provides a platform for accelerated software development and product evaluation. It includes a printed circuit board with a typical configuration of two BCM5690s connected back-to-back and integrated to deliver 24 GE copper ports and two optional fiber ports via GBICs. The SDK comes with the Wind River VxWorks real-time operating system, 19-inch rack-mount chassis with Compact PCI backplane and power supply, and the latest version of the Broadcom API driver source code.

Table 4 lists all the SDKs that are available for StrataXGS products.

Part Number Legend:
K= Switch Development Kit
R = Reference Design
S = Stackable
U = Upgrade Board only (no chassis or power supply)

Table 4: StrataXGS SDK Part Numbers

Part Number	*Description*
BCM95670K8	Switch Fabric Development Kit with 8 HiGig ports
BCM95670K8U	Switch Fabric Development Kit with 8 HiGig ports - Upgrade
BCM95690K24	Switch Development Kit with 24GE ports
BCM95690K24S	Switch Development Kit with 24GE and 2HiGig ports
BCM95690K24U	Switch Development Kit with 24GE - Upgrade
BCM95690K24SU	Switch Development Kit with 24GE and 2HiGig ports - Upgrade
BCM95690R24	Switch Reference Design with 24GE Ports
BCM95690R24S	Stackable Switch Reference Design with 24GE Ports
BCM95690R24U	Stackable Switch Reference Design with 24GE Ports - Upgrade
BCM95690R48S	Stackable Switch Reference Design with 48GE Ports
BCM95690R48SU	Stackable Switch Reference Design with 48GE Ports - Upgrade
BCM95691K12	Switch Development Kit with 12GE ports
BCM95691K12U	Switch Development Kit with 12GE ports - Upgrade

Note: All SDKs now available with standard PowerPC 8240 processor.
IDT 32334 or PowerPC 8245 processor available with a longer lead time.

SAMPLE SYSTEM CONFIGURATIONS

What follows are just a few sample GE switch designs that leverage the StrataXGS architecture.

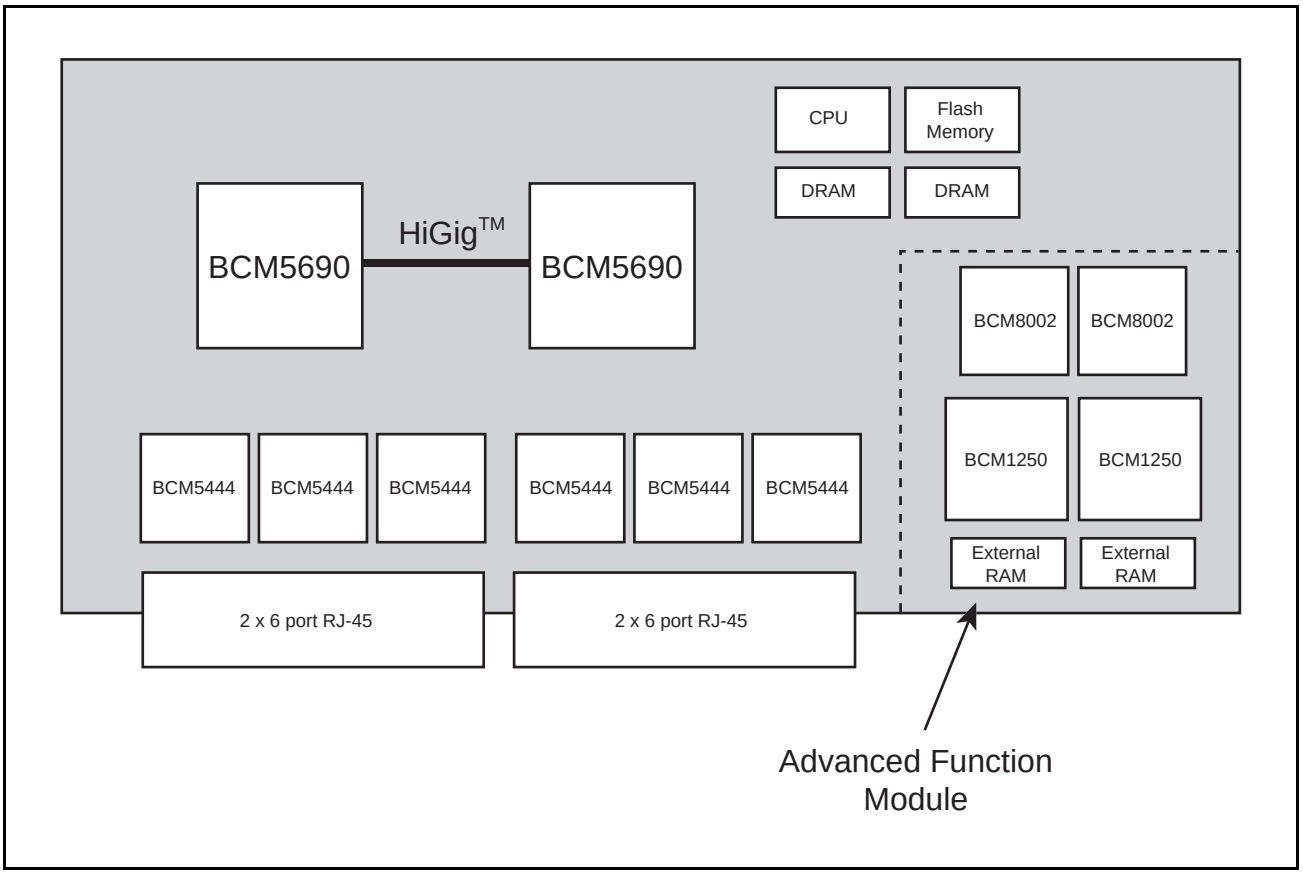


Figure 10: Fixed 24-port GE Switch with Optional Advanced Function Module

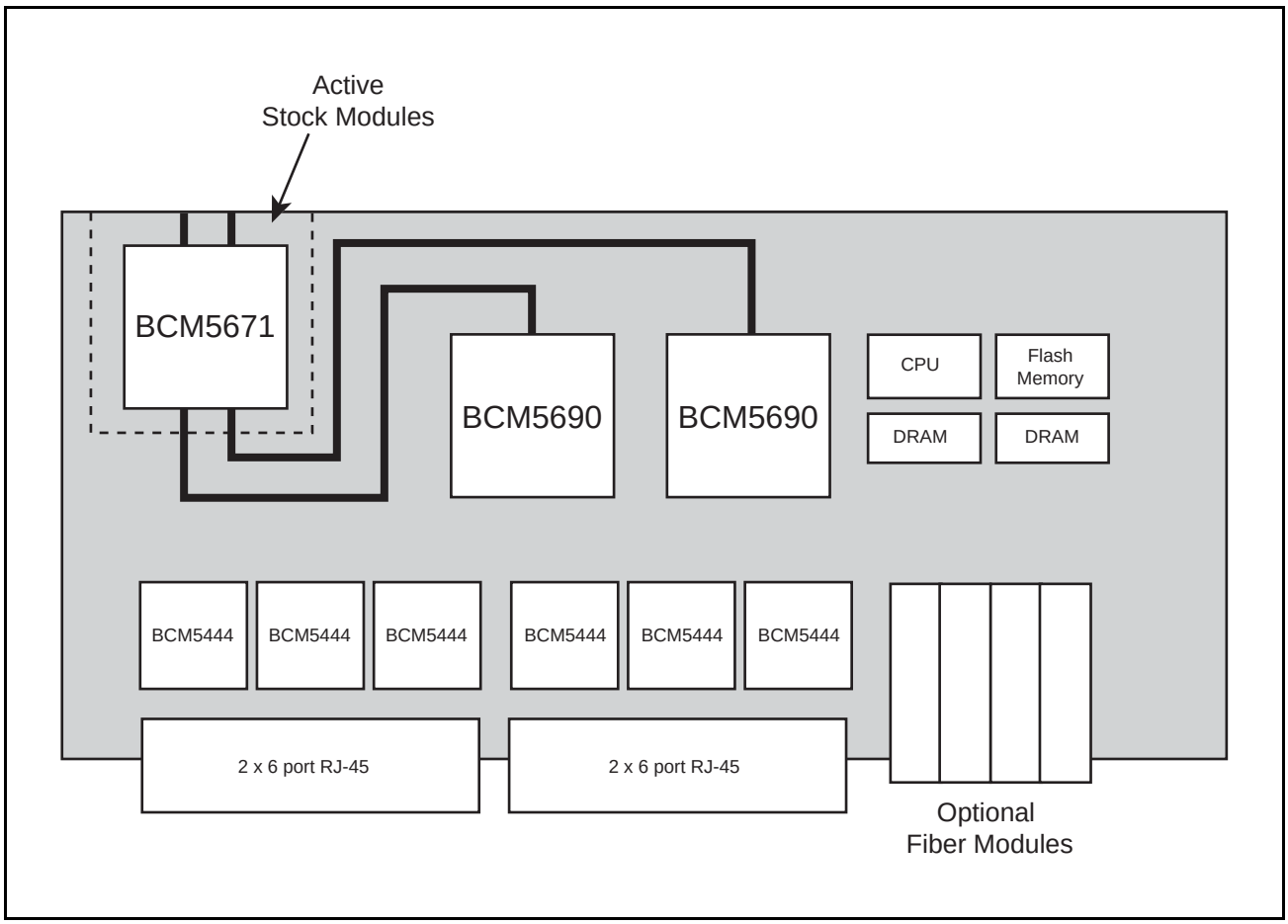


Figure 11: 24-port GE Switch with Active Stack Module

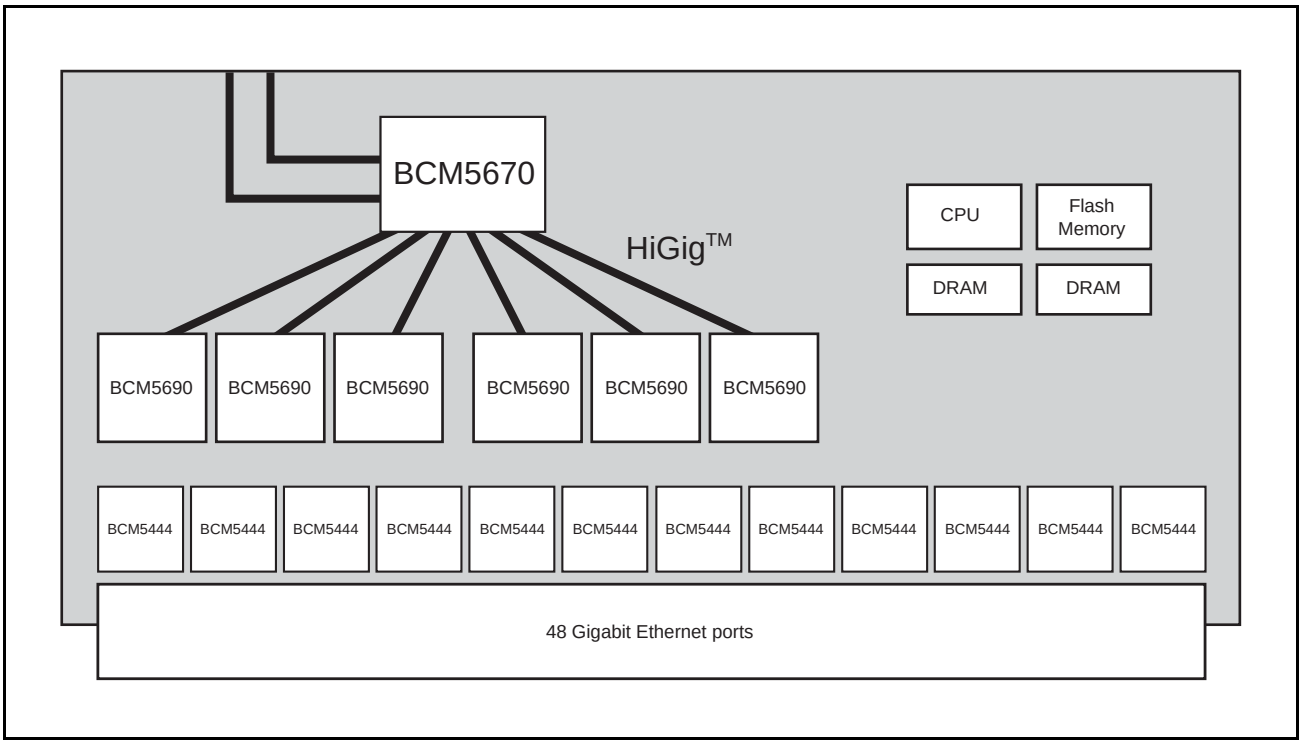


Figure 12: 48-port Nonblocking GE Switch

SUMMARY AND CONCLUSION

Businesses are become increasingly reliant on new Web-based applications to automate their businesses and to reach customers and business partners across intranets, extranets, and the Internet. Network storage, VoIP, and other multimedia applications are driving the need for much higher LAN connectivity speeds to the desktop.

The Broadcom StrataXGS Switch Architecture enables switch vendors to meet enterprise customers' high-speed connectivity needs with modular, integrated components that serve as simple building blocks. The components, with high port counts and rich feature sets, keep system design and assembly simple, which lowers assembly time and costs, while providing designers with a high degree of flexibility in the form factors they choose to build and features they choose to activate. The StrataXGS architecture allows the use of same building blocks for creating different types of systems. It is backward-compatible with proven StrataSwitch II chip architecture and the Broadcom development API so that components and expertise can be reused. This hastens a switch vendor's time to market and lowers resource requirements/costs.

Using combinations of the Broadcom BCM5670 and BCM5671 high-performance switching fabrics and the BCM5690 and BCM5691 GE switching chips, designers can scale their systems to support hundreds of GE and 10-GE ports, configured in a variety of form factors to meet the needs of their target applications.

Broadcom Corporation

16215 Alton Parkway
P.O. Box 57013
Irvine, California 92619-7013
Phone: 949-450-8700
Fax: 949-450-8710