

Product Brief

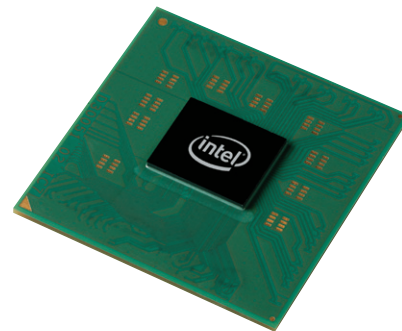
Intel® 82598 10 Gigabit Ethernet Controller
Network Connectivity



Intel® 82598 10 Gigabit Ethernet Controller

High-performance, Dual-Port 10 Gigabit Network Connectivity Designed for Multi-core Processors and Optimized for Virtualization

- High-performing, PCI Express* 1/10 Gigabit Ethernet connection ideally suited for demanding enterprise applications and embedded system designs
- Dual-port, industry-leading, energy-efficient design creating new opportunities for blades and LAN-on-motherboard (LOM) solutions
- Introduces multiple product innovations optimized for virtualization and ideal for server consolidation
- High-volume stable architecture that includes broad operating system support



Energy-Efficient, Next-Generation 10 Gigabit Performance

10 Gigabit Ethernet has moved past the early adoption stage and is rapidly becoming the mainstay for backbones within enterprise and service provider networks. The escalating deployments of servers with multi-core processors and demanding applications such as high-performance computing (HPC), database clusters, and video on demand are driving the need for a 10 Gigabit connection. The Intel® 82598 10 Gigabit Ethernet Controller is a next-generation, PCI Express* 2.0 (2.5 Gbps) controller with balanced features and energy-efficient design that meets the throughput and latency requirement of bandwidth-hungry applications. The 82598 controller is designed for multi-core processors and implements advanced features for efficient routing of packets using multiple queues as well as include load balancing of interrupts using MSI-X. The controller also supports stateless offloads such as TCP segmentation offload, header replications/splitting, and Direct Cache Access (DCA) to lower CPU utilization. In addition, the 82598 controller offers advanced features for Virtualization and Storage over Ethernet all within a very low power envelop of 4.8W (dual port).

High-Performance Design Features

The Intel 82598 10 Gigabit Ethernet Controller is a single, compact component with two fully integrated 10 Gigabit Ethernet Media Access Control (MAC) and XAUI ports. The controller supports CX4 (802.3ak*) and KX4/KX (802.3ap*) interfaces and each port contains a Serializer-Deserializer (SerDes) for backward compatibility with gigabit backplanes. The device is designed for high performance and lower memory latency. Wide internal data paths eliminate performance bottlenecks by efficiently handling large address and data words. The controller includes advanced interrupt-handling features and uses efficient ring-buffer descriptor data structures, with up to 64 packet descriptors cached on chip. A large on-chip packet buffer maintains superior performance. The 82598 controller enables network manageability implementations required by IT personnel for remote control and alerting. The communication to the Board Management Controller (BMC) is available either through an on-board System Management BUS (SMBus) port or through the Distributed Management Task Force (DMTF)-defined NC-SI.

Hardware-assisted Virtualization

As IT focuses on reducing costs by deploying virtualization and consolidating servers, this new environment also creates unique challenges for the I/O infrastructure. Recognizing these challenges, Intel has designed smart acceleration capabilities into Ethernet Controllers that optimizes networking performance on virtual servers. The Intel 82598 10 Gigabit Ethernet Controller adds Virtual Machine Device queue (VMDq) technology that offloads data sorting and data copying from the virtual machine monitor (VMM) software layer to the hardware. This improves overall throughput and CPU utilization on virtualized servers. VMDq technology also ensures transmit fairness and prevents head-of-line blocking to deliver enhanced latency performance.

Advanced Features for Storage over Ethernet

The fast growth in storage capacity coupled with server virtualization has brought the need for Storage Area Network (SAN) in the forefront. To satisfy this growing demand Intel 82598 10 Gigabit Ethernet Controller includes support for iSCSI acceleration and advanced features for unified storage connectivity. The controller enables fast and reliable networked storage with native iSCSI support with Microsoft, Linux,* and VMware operating systems as well as support for iSCSI remote boot. Advanced QoS features such as priority groups and per priority pause are also implemented in the controller. These hardware features facilitate Host I/O consolidation and enable migration for Fibre Channel over Ethernet as an alternative to expensive Fibre Channel Host Bus adapters (HBAs).

Order Code

82598EB JL82598EB

For more information, contact your Intel sales representative.

Features

Benefits

External Interface Features

PCI Express* 2.0 (2.5 Gbps)

- Supports x8, x4, x2, x1 lanes
- Supports extended error reporting and completion timeout

Compatible extensions to PCI power management and ACPI

- Efficient power management

XAUI, CX4 (IEEE 802.3ak*) and 10GBase-KX4, 1000Base-KX (IEEE 802.3ap*)

- High-speed interfaces for backplane and PYH connections

10 Gigabit MAC Advanced Features

Intel® I/O Acceleration Technology (Intel® IOAT and Quick Data)

- Accelerated TCP I/O for improved CPU utilization

Wide, pipelined internal data path architecture

- Low-latency data handling
- Superior direct memory access (DMA) transfer-rate performance

MSI-X support

- Minimizes the overhead of interrupts
- Allows load balancing of interrupt handling between different cores/CPUs

Mechanism available for reducing interrupts generated from Tx/Rx operations

- Maximizes system performance and throughput

Low-latency interrupts

- Ability to toggle between the interrupt aggregation and non-aggregation mode based on the type of data being transferred

Optimized queues: 32 Transmit (Tx) and 64 Receive (Rx) per port

- Network packet handling without waiting or buffer overflow
- Efficient packet prioritization

Caches up to 64 packet descriptors per queue

- Efficient use of PCI Express bandwidth

Dual configurable first-in/first-out (FIFO) buffers for each port
512 Kb Transmit (Tx) and 320 Kb Receive (Rx)

- No external FIFO memory requirements
- FIFO size adjustable to application
- Error detection and correction for FIFO data

Support for transmission and reception of packets up to 16 KBytes (Jumbo Frames)

- Enables higher and better throughput of data

Programmable host memory receive buffers (256 Bytes to 16 KBytes) and cache line size (64 Bytes to 128 Bytes)

- Efficient use of PCI Express bandwidth

Auto-negotiation support as defined in IEEE 802.3ap clause 73

- Automatic link configuration for speed (10G or 1G mode)
- Improves performance and reliability

IEEE 802.3* compliant flow-control support with software-controllable pause times and threshold values

- Frame loss reduced from receive overruns
- Control over the transmissions of pause frames through software or hardware triggering

Features

Benefits

Host Offloading Features

16 Virtual Machine Device queues (VMDq)	<ul style="list-style-type: none">Allows the efficient routing of packets to the correct target machine in a virtualized environment using multiple hardware queuesEnsures transmit fairness and prevents head-of-line blocking
Direct Cache Access (DCA)	<ul style="list-style-type: none">Enables the I/O device to activate a pre-fetch engine in the CPU that loads the data into the CPU cache ahead of time, before use, eliminating cache misses and reducing CPU load
Header split and replication in receive	<ul style="list-style-type: none">Helps the driver to focus on the relevant part of the packet without the need to parse it
Tx/Rx IP, TCP, and UDP checksum offloading (IPv4, IPv6) capabilities (IPv4, IPv6)	<ul style="list-style-type: none">Lower processor utilizationChecksum and segmentation capability extended to new standard packet type
Tx TCP segmentation offload (IPv4, IPv6)	<ul style="list-style-type: none">Increased throughput and lower processor utilizationCompatible with large send offload feature (in Microsoft Windows* XP)
IPv6 offloading	<ul style="list-style-type: none">Checksum and segmentation capability extended to new standard packet type
Receive Side Scaling for Windows environment and Scalable I/O for Linux environments (IPv4, IPv6, TCP/UDP)	<ul style="list-style-type: none">Multiple Rx queues
Advanced packet filtering	<ul style="list-style-type: none">16 exact-matched packets (unicast or multicast)4096-bit hash filter for multicast framesLower processor utilizationPromiscuous (unicast and multicast) transfer mode supportOptional filtering of invalid frames
IEEE 802.1q* virtual local area network (VLAN) support with VLAN tag insertion, stripping and packet filtering for up to 4096 VLAN tags	<ul style="list-style-type: none">Ability to create multiple VLAN segments
Priority Groups (up to 8)	<ul style="list-style-type: none">Resource allocation per virtual link to provide differentiation among different traffic types (LAN, SAN, and IPC)
Per Priority Pause	<ul style="list-style-type: none">Enables finer grain control of traffic for virtual links associated with Priority group. Enables "no-drop" behavior on Ethernet for critical storage traffic

Manageability Features

DMTF NC-SI pass through	<ul style="list-style-type: none">Industry standard for BMC interfaceAllows fast data rates (up to 100 Mb/s full duplex)Better capabilities (video redirection)Extended filtering capabilities
SMBus port	<ul style="list-style-type: none">Supports pass through over the SMBus interfaceData rates of up to 400 KHzAllows serial redirection and IPMI traffic redirection to BMC
Advanced filtering capabilities (IPv4, IPv6)	<ul style="list-style-type: none">Supports extended L2, L3 and L4 filtering for traffic routing to BMCSupports MAC address, VLAN, ARP, IPv4, IPv6, RMCP UDP ports, UDP/TCP ports filteringSupports flexible header filteringAllows the BMC to share the MAC address with the host OS
Preboot eXecution Environment (PXE) flash interface support	<ul style="list-style-type: none">Enables system boot up via the LAN (32 bit and 64 bit)Flash interface for PXE image
Simple Network Management Protocol (SNMP) and Remote Network Monitoring (RMON) statistic counters	<ul style="list-style-type: none">Easy system monitoring with industry-standard consoles
Wake-on-LAN support	<ul style="list-style-type: none">Packet recognition and wake-up for LAN on motherboard applications without software configuration
iSCSI boot	<ul style="list-style-type: none">Enables system boot up via iSCSIProvides additional network management capability
MDIO – internal management interface	<ul style="list-style-type: none">Enables the MAC to monitor and control the PHY

Additional Device Features

Four outputs on each port that directly drive LEDs	<ul style="list-style-type: none">Link and activity indications on each port
JTAG (IEEE 1149.1*) test access port built-in silicon	<ul style="list-style-type: none">Simplified testing using boundary scanSupports the IDCODE instruction

Characteristics

<i>Electrical</i>	
Typical targeted power dissipation	<ul style="list-style-type: none">4.8 W @ D0 Active Link dual port 10 Gbps3.4 W @ D0 Active Link single port 10 Gbps1.6 W @ D3 Cold (wake up enabled at 1000 Mbps)
<i>Environmental</i>	
Operating temperature	<ul style="list-style-type: none">0° to 70° C (with thermal management)
<i>Physical</i>	
Implemented in 90nm complementary metal-oxide semiconductor (CMOS) process	<ul style="list-style-type: none">Offers lowest geometry to minimize power and size while maintaining quality and reliability
Package	<ul style="list-style-type: none">31 mm x 31 mm 833-pin Flip-Chip Ball Grid Array (FC-BGA) package

To see the full line of Intel® Ethernet Controllers, visit www.intel.com/network/connectivity

INFORMATION IN THIS DOCUMENT IS PROVIDED IN CONNECTION WITH INTEL® PRODUCTS. NO LICENSE, EXPRESS OR IMPLIED, BY ESTOPPEL OR OTHERWISE, TO ANY INTELLECTUAL PROPERTY RIGHTS IS GRANTED BY THIS DOCUMENT. EXCEPT AS PROVIDED IN INTEL'S TERMS AND CONDITIONS OF SALE FOR SUCH PRODUCTS, INTEL ASSUMES NO LIABILITY WHATSOEVER, AND INTEL DISCLAIMS ANY EXPRESS OR IMPLIED WARRANTY, RELATING TO SALE AND/OR USE OF INTEL PRODUCTS INCLUDING LIABILITY OR WARRANTIES RELATING TO FITNESS FOR A PARTICULAR PURPOSE, MERCHANTABILITY, OR INFRINGEMENT OF ANY PATENT, COPYRIGHT OR OTHER INTELLECTUAL PROPERTY RIGHT. UNLESS OTHERWISE AGREED IN WRITING BY INTEL, THE INTEL PRODUCTS ARE NOT DESIGNED NOR INTENDED FOR ANY APPLICATION IN WHICH THE FAILURE OF THE INTEL PRODUCT COULD CREATE A SITUATION WHERE PERSONAL INJURY OR DEATH MAY OCCUR.

Intel may make changes to specifications and product descriptions at any time, without notice. Designers must not rely on the absence or characteristics of any features or instructions marked "reserved" or "undefined." Intel reserves these for future definition and shall have no responsibility whatsoever for conflicts or incompatibilities arising from future changes to them. The information here is subject to change without notice. Do not finalize a design with this information. The products described in this document may contain design defects or errors known as errata which may cause the product to deviate from published specifications. Current characterized errata are available on request. Contact your local Intel sales office or your distributor to obtain the latest specifications and before placing your product order. Copies of documents which have an order number and are referenced in this document, or other Intel literature, may be obtained by calling 1-800-548-4725, or by visiting Intel's Web Site <http://www.intel.com/>.

Copyright © 2007 Intel Corporation. All rights reserved.

Intel, the Intel logo, Intel. Leap ahead., and the Intel. Leap ahead. logo are trademarks of Intel Corporation in the U.S. and other countries.

*Other names and brands may be claimed as the property of others.

