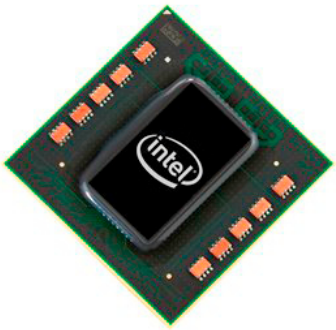
PRODUCT BRIEF
Intel® Ethernet Controller X540
Network Connectivity

Intel® Ethernet Controller X540

Integrated Single Chip 10GBASE-T Controller Simplifies 10 Gbps Ethernet Server LOM, Converged Network Adapter, and Network Daughter Card Designs



Key Features

- Industry-first single chip, dual-port 10GBASE-T Controller with Integrated MAC and PHY
- 12.5 W maximum power
- Unified Networking delivering LAN, iSCSI, and FCoE over 10GBASE-T
- Two independent 10GBASE-T interfaces with SR-IOV support
- SMBus and NC-SI interfaces with OS2BMC, MCTP and Wake-on-LAN (WoL) support
- Flexible I/O virtualization for port partitioning and quality of service (QoS) of up to 128 virtual ports
- Jumbo Frames of up to 15.5 KB
- Integrated IPsec and MACsec Security Engines
- Data Center Bridging (DCB) with FCoE stateless offloads
- PCI Express* v 2.1 with 5.0 GT/s and 2.5GT/s support for x1, x2, x4, x8 links widths
- 10GBASE-T, 1000BASE-T, and 100BASE-TX Link Modes
- Reliable and proven 10 Gigabit Ethernet technology from Intel® Corporation

10 Gigabit for the Broad Market

The Intel® Ethernet Controller X540 is Intel's latest industry-leading innovation, to reduce the cost of 10 Gb Ethernet and bring it to the broad server market. The small, 25 x 25 mm package with two low-power 10GBASE-T ports is designed using industry-leading integrated RFI filters for 100 meter cable lengths. Designed for LAN-on-Motherboard (LOM), Network Daughter Cards, and Converged Network Adapter (CNA) integration, the single-chip design provides significant BOM savings by reducing the support components (bridge chips, crystals, and EEPROMS) required when compared to other multi-chip solutions.

10GBASE-T Simplifies the Transition From 1 GbE to 10 GbE

The X540 controller provides two integrated 10GBASE-T PHYs providing 10 Gbps of throughput, which are also backwards compatible with legacy GbE switches and Cat 6A cabling. The ability to auto-negotiate between 100 Mbps, 1 Gbps, and 10 Gbps speeds provides the backwards compatibility for a smooth transition and easy migration to 10 GbE.

Next Generation Immunity

The X540 controller implements a 5th-channel receiver dedicated to the common mode signal specifically designed for RFI/EMI detection. The 5th-channel, along with a powerful cable diagnostic algorithm that accurately measures all of the TDR and TDT sequences within the group of four channels, greatly improves reliability and signal integrity.

Exciting New Data Center Models

More than simply a 10x per port increase in performance, using the X540 controller (vs. a standard 1 GbE controller) opens doors for exciting new usage models, including Unified Networking, I/O Virtualization, and Flexible Port Partitioning.

Simplified I/O Virtualization

Virtualization changes server resource deployment and management by running multiple applications and operating systems on a single physical server.

Intel® Virtualization Technology for connectivity (Intel® VT-c) delivers I/O virtualization and Quality of Service (QoS) features designed directly into the X540 controller's silicon. I/O virtualization advances network connectivity used in today's servers to more efficient models by providing FPP, multiple Rx/Tx queues, Tx queue rate-limiting, and on-controller QoS functionality that is useful for both virtual and non-virtual server deployments.

Flexible Port Partitioning (FPP)

By taking advantage of the PCI-SIG* SR-IOV specification, FPP enables virtual Ethernet controllers that can be used by a Linux* host directly and/or assigned directly to virtual machines for hypervisor virtual switch bypass. FPP enables the assignment of up to 64 Linux host processes or virtual machines per port to virtual functions. An administrator can use FPP to control the partitioning of the bandwidth across multiple virtual functions. FPP can also provide balanced QoS by giving each assigned virtual function equal access to 10Gbs of bandwidth.

Unified Networking

The Intel® Ethernet Controller X540 enables servers to use existing Cat 6 or Cat 6A cables to unify NFS/CIFS, iSCSI and FCoE with LAN traffic instead of forcing the use of Twinaxial copper or Fiber Optic SFP+ solutions.

Unified Networking Principles

Intel's Unified Networking solutions are built on the principles that made us successful in Ethernet:

- Open Architecture integrates networking with the server, enabling IT managers to reduce complexity and overhead while enabling a flexible and scalable data center network.
- Intelligent Offloads lower cost and power while delivering the application performance that customers expect.
- Proven Unified Networking is built on trusted Intel Ethernet technology, enabling customers to deploy FCoE or iSCSI with the same quality used in their traditional Ethernet networks.

Intel's Unified Networking solutions are enabled through a combination of Intel Ethernet products along with network and storage protocols integrated in the operating systems. This combination provides proven reliability with the performance that data center administrators around the world have come to expect from Intel.

Proven iSCSI SAN Connectivity

The Intel® Ethernet Controller X540 provides iSCSI support without the need for complicated firmware, driver, and proprietary software combinations. Native OS iSCSI initiators work in conjunction with intelligent offloads built-in to both the X540 controller and Intel® Xeon Processor-based servers to provide performance with proven reliability.

10GBASE-T Supports Open-FCoE

For the first time, Open-FCoE is supported on 10GBASE-T. As 10GBASE-T switches come to market enabled with FCoE support, the X540 controller enables using cost-effective Cat 6 and Cat 6A cabling for converged networking. The Open-FCoE architecture uses a combination of hardware intelligent offloads and FCoE initiators in VMware* ESXi 5.0, Microsoft* Windows*, and Linux* operating systems to deliver proven high-performance FCoE solutions.

This approach enables IT managers to simplify the data center and standardize on 10 GbE for LAN and SAN connectivity. The Intel® Ethernet Controller X540 is designed to fully offload the FCoE data path to deliver full-featured, converged network adapter (CNA) functionality as a LOM or adapter without compromising on power efficiency and interoperability.

DCB Delivers Lossless Ethernet

The Intel® Ethernet Controller X540 supports Ethernet enhancements such as Data Center Bridging (DCB), a collection of standards, for additional QoS functionality such as lossless delivery, congestion notification, priority-based flow control, and priority groups.

10 Gbps Performance at Low Cost and Low Power

The new Intel® Ethernet Controller X540 brings 10GBASE-T as a cost effective means to bring 10 Gbps Ethernet to server platforms as LOM or Network Daughter Card. The X540 Controller connects to Cat 6 and Cat 6A cabling and has backwards compatibility with many 100BASE-Tx and 1000BASE-T switches to provide broad network infrastructure support.

To reduce cost and power, the X540 controller is manufactured using a 40 nm process with an integrated MAC controller and two 10GBASE-T PHYs in a single-chip solution.

Integration translates to lower power with reduced per-port power consumption, which can eliminate the need for active fan heatsinks.

With lower power, passive heatsinks, and backwards compatibility, the X540 controller and 10GBASE-T are ready for broad deployment. The X540 controller provides bandwidth-intensive applications and virtualized data centers 10 GbE network performance with cost-effective network connectivity.

Network Manageability Interfaces

The X540 controller provides OS2BMC, SMBus and DMTF-defined Network Controller Sideband Interface (NC-SI) for BMC manageability. In addition, it introduces support for Management Component Transport Protocol (MCTP), a new DMTF standard, enabling a BMC to gather information about Intel Ethernet Converged Network Adapters that can include the data rate, link speed, and error counts.

Software Tools and Management

Intel® Advanced Network Services (Intel® ANS), includes new teaming technologies and techniques such as Virtual Machine Load-Balancing (VMLB) for Hyper-V environments. Intel ANS also includes a variety of teaming configurations for up to eight ports, support for mixed vendors' server LOM and adapters teaming and includes support for IEEE 802.1Q VLANs, making Intel ANS one of the most capable and comprehensive tools for supporting server adapter teaming.

Additionally, Intel® PROSet for Windows* Device Manager (DMIX) and PROsetCL extends driver functionality to provide additional reliability and QoS features and configuration.

External Interfaces

PCI Express* Interface v 2.1 with 5.0 GT/s and 2.5 GT/s Support for x1, x2, x4, x8 links widths (Lanes)

Network Interfaces: Two independent Ethernet interfaces for 10GBASE-T, 1000BASE-T, and 100BASE-TX applications (IEEE 802.3an, 802.3, 802.3u, and 802.3ab)

Management Interfaces

- Pass-Through (PT) Functionality via a sideband interface
- DMTF Network Controller Sideband Interface (NC-SI)
- Intel® System Management Bus (SMBus)

BOM Cost Reduction

Features	Benefits
Single chip design	• Designed for passive heatsink thermal solutions
25 mm x 25 mm package size	• Small packaging for easier board layout and design
Integrated copper 10GBASE-T PHYs	• Single chip with integrated PHYs for lower power and simplified component placement
5th Channel Filtering and cable diagnostics	• Senses and cancels common-mode and board noise and provides advanced troubleshooting data
Intel® Lead-free technology and RoHS-compliant	• Compliant with the European Union directive (July 2006) to reduce hazardous materials
Autonomous on-die thermal management	• Monitor on-die temperature and react when the temperature exceeds a pre-defined threshold

Ethernet Features

Features	Benefits
IEEE 802.3* auto-negotiation	• Automatic link configuration for speed, duplex, flow control
Automatic Cable Diagnostics	• Powerful cable diagnostic algorithm to accurately measure all of the TDR and TDT sequences within the group of four channels
Independent port enabling and link speeds	• Each port can be configured and operated at different speeds and in different modes
IEEE 802.3x and 802.3z compliant flow control support with software-controllable Rx thresholds and Tx pause frames	• Local control of network congestion levels • Frame loss reduced from receive overruns
Data Center Bridging (DCB) support	• IEEE Compliance to Enhanced Transmission Selection (ETS), 802.1Qaz Priority-based Flow Control (PFC), 802.1Qbb Quantized Congestion Notification
Automatic cross-over detection function (MDI/MDI-X)	• The PHY automatically detects which application is being used and configures itself accordingly
IEEE 1588 protocol and IEEE 802.1AS implementation	• Time-stamping and synchronization of time sensitive applications • Distribute common time to media devices
IEEE 802.1ad (Double VLAN)	• Double-tagging can be useful for Internet service providers, allowing the use of VLANs internally while mixing traffic from clients that are already VLAN-tagged
IEEE 802.1Q (VLAN)	• Provide data separation and security between network traffic

Security and Power Management

Features	Benefits
Receive Packet Filtering	• Determine which of the incoming packets are allowed to pass to the local machine based on L2, VLAN, or management policies
Integrated MACsec, 802.1AE Security Offload Engines	• Offloads for MAC-level encryption/authentication scheme defined in IEEE 802.1AE that uses symmetric cryptography
Integrated IPsec Security Engines for offloads of up to 1024 Security Associations (SA) for each Tx and Rx	• Offloads handle a certain amount of the total number of IPsec flows on the controller in hardware
Anti-spoofing for MAC and VLANs	• Capability insures that a VM always uses a source Ethernet VLAN or MAC address on the transmit path that is part of the set of VLAN tags and Ethernet MAC addresses defined on the Rx path
Four Software-Definable Pins (SDP) per port	• Software-defined pins (SDP pins) per port that can be used for miscellaneous hardware or software-controllable purposes
Access Control Services (ACS)	• ACS Extended Capability structures on all functions
Active State Power Management (ASPM) Support	• Optionality Compliance bit to help determine whether to enable ASPM or whether to run ASPM compliance tests to support entry to L0
LAN disable function	• Option to disable the LAN Port and/or PCIe Function. Disabling just the PCIe function but keeping the LAN port that resides on it fully active (for manageability purposes and BMC pass-through traffic)
Full wake up support: • Advanced Power Management (APM) Support (formerly Wake on LAN) • Advanced Configuration and Power Interface (ACPI) specification v2.0c • Magic Packet* wake-up enable with unique MAC address	• APM - Designed to receive a broadcast or unicast packet with an explicit data pattern (Magic Packet) and assert a signal to wake up the system • ACPI - PCIe power management based wake-up that can generate system wake-up events from a number of sources
Low Power Operation and Power Management	• Incorporates numerous features to maintain the lowest power possible including PCI Express Link and Network Interface power management
ACPI register set and power down functionality supporting D0 and D3 states	• A power-managed link speed control lowers link speed (and power) when highest link performance is not required
Low Power Link Up - Link Speed Control	• Enables a link to come up at the lowest possible speed in cases where power is more important than performance

I/O Virtualization

Features	Benefits
Multi-mode I/O Virtualization Operations	▪ Supports two modes of operations of virtualized environments: - Direct assignment of part of the port resources to different guest operating systems using the PCI SIG SR-IOV standard (Also known as native mode or pass through mode) - Central management of the networking resources by hypervisor (Also known as software switch acceleration mode) ▪ A hybrid model, where some of the VMs are assigned a dedicated share of the port and the rest are serviced by an hypervisor is also supported
Virtual Machine Device Queues (VMDq)	▪ Offloads data-sorting from the Hypervisor to silicon, improving data throughput and CPU usage ▪ QoS feature for Tx data by providing round-robin servicing and preventing head-of-line blocking ▪ Sorting based on MAC addresses and VLAN tags
Next Generation VMDq	▪ Enhanced QoS feature by providing weighted round-robin servicing for the Tx data ▪ Provides loopback functionality; data transfers between the virtual machines within the same physical server do not go out to the wire and back in, improving throughput and CPU usage ▪ Supports replication of multicast and broadcast data
64 transmit (Tx) and receive (Rx) queue pairs per port	▪ Supports VMware* NetQueue and Microsoft* VMQ ▪ MAC/VLAN filtering for pool selection and either DCB or RSS for the queue in pool selection
Flexible Port Partitioning: 64 Virtual Functions per port	▪ Virtual Functions (VFs) appear as Ethernet Controllers in Linux OSes that can be assigned to VMs, Kernel processes or teamed using the Linux* Bonding Drivers
Support for PCI-SIG SR-IOV specification	▪ Up to 64 Virtual Functions per Port
Rx/Tx Round-Robin scheduling	▪ Assigns time slices in equal portions in circular order for Rx/Tx for balanced bandwidth allocation
Traffic Isolation	▪ Processes or VM can be assigned a dedicated VF with VLAN support
Traffic Steering	▪ Offloads sorting and classifying traffic in to VF or queues
VM to VM Packet forwarding (Packet Loopback)	▪ On-chip VM-VM traffic allows PCIe* speed switching between VM
Multicast and Broadcast Packet Replication	▪ Multicast and broadcast packets can be sent to a single queue or be replicated across a across multiple queue pools
Per-pool settings, statistics, off loads, and jumbo support	▪ Each Queue Pair or Pool has its own statistics, off-loads and Jumbo support options
Dynamic Transmit and Receive Queues	▪ Queues can be enabled or disabled dynamically
Independent Function Level Reset (FLR) for Physical and Virtual Functions	▪ VF resets only the part of the logic dedicated to specific VF and does not influence the shared port
IEEE 802.1Q Virtual Local Area Network (VLAN) support with VLAN tag insertion, stripping and packet filtering for up to 4096 VLAN tags	▪ Adding (for transmits) and ping (for receives) of VLAN tags ▪ Filtering packets belonging to certain VLANs
IEEE 802.1Q advanced packet filtering	▪ Lower processor usage
L2 Ethernet MAC Address Filters (unicast and Multicast)	▪ Enables up to 128 MAC address to be assigned to queue pools and virtual functions for controller based traffic steering
L2 VLAN Filters	▪ Enables up to 64 VLANs to be assigned to queue pools and virtual functions for controller based traffic steering
Mirroring rules	▪ Ability to reflect network traffic to a given VM or VLAN based on up to four mirroring types

Stateless Offloads and Performance Features

Features	Benefits
TCP/UDP, IPv4 checksum offloads (Rx/ Tx/Large-send); Extended Tx descriptors for more offload capabilities	▪ Improved CPU usage ▪ Checksum and segmentation capability extended to new standard packet type
IPv6 support for IP/TCP and IP/UDP receive checksum offload	▪ Improved CPU usage
Tx TCP Segmentation Offload (TSO-IPv4, IPv6)	▪ Large TCP I/O is segmented into smaller packets to increase throughput and reduce CPU overhead ▪ Compatible with large-send offload
Interrupt throttling control	▪ Limits maximum interrupt rate and improves CPU usage
Legacy and Message Signal Interrupt (MSI) Modes	▪ Enables Interrupt mapping
Message Signal Interrupt Extension (MSI-X)	▪ Dynamic allocation of up to 64 vectors per port
Intelligent interrupt generation	▪ Enhanced software device driver performance
Receive Side Scaling (RSS) for Windows environment Scalable I/O for Linux environments (IPv4, IPv6, TCP/UDP)	▪ Up to 32 flows per port ▪ Improves the system performance related to handling of network data on multiprocessor systems
Receive Side Coalescing (RSC)	▪ Merge multiple received frames from the same TCP/IP connection into a single structure
128 Tx and Rx Queues (per port)	▪ Queues provide QoS for virtualization, DCB, RSS, L2 Ethertype, FCoE Redirection, L3/4 5-tuple filters, Flow Director and TCp SYN filters
Rate Control Traffic per Traffic Class/Transmit Queue	▪ In order to guarantee each pool with adequate bandwidth, a per-pool bandwidth control mechanism is used
FCoE Tx / Rx CRC Offload	▪ Offloads receive FC CRC integrity check while tracking the CRC bytes and FC padding bytes
Large FC Receive	▪ Large FC receive includes two types of offloads that can save a data copy by posting the received FC payload directly to the kernel storage cache or the user application space

Stateless Offloads and Performance Features *(continued)*

Features	Benefits
FCoE Transmit Segmentation Offloads	• Enables the FCoE software to initiate a transmission of multiple FCoE packets up to a complete FC sequence with a single header in host memory (single instruction)
FCoE Coalescing and Direct Data Placement	• Hardware can provide DDP offload for up to 512 concurrent outstanding FC read or write exchanges
Traffic Class (TC) using 802.1p	• A specific TC can be configured to receive or transmit a specific amount of the total bandwidth available per port
Flow Director Filters: up to 32 KB - 2 Signature Filters up to 8 KB - 2 Perfect Match Filters	• The flow director filters identify specific flows or sets of flows and routes them to specific queues These filters are an expansion of the L3/L4 5-tuple filters that provide up to additional 32 K filters
Support for packets up to 15.5KB (Jumbo Frames)	• Enables higher and better throughput of data
Low Latency Interrupts	• Based on the sensitivity of the incoming data, the controller can bypass the automatic moderation of time intervals between the interrupts
Direct Cache Access (DCA) support	• Method to improve network I/O performance by placing some posted inbound writes directly within CPU cache
TCP Timer Interrupts	• Enables the software driver to read a EICR register bit set by the controller, avoiding cache thrash and enabling parallelism
No Snoop	• System logic can provide a separate path into system memory for non-coherent traffic. The non-coherent path to system memory provides a higher, more uniform, bandwidth for write requests
Relax Ordering	• When the strict order of packets is not required, the device can send packets in an order that allows for less power consumption and greater CPU efficiency
Rx Packet Split Header	• Helps the driver to focus on the relevant part of the packet without the need to parse
Descriptor ring management hardware for Transmit and Receive	• Optimized descriptor fetch and write-back for efficient system memory and PCIe bandwidth usage

Remote Boot Options

Features	Benefits
Preboot eXecution Environment (PXE) flash interface support	• Enables system boot up via the EFI (32 bit and 64 bit) • Flash interface for PXE 2.1 option ROM
Intel® Ethernet FCoE Boot	• Enables system boot up via FCoE
Intel® Ethernet iSCSI Remote Boot	• Enables system boot up via iSCSI
Intel Boot Agent software: Linux boot via PXE or BOOTP,Windows* Deployment Services, or UEFI	• Allows networked computer to boot using a program code image supplied by a remote server • Complies with the Pre-boot eXecution Environment (PXE) Version 2.1 Specification

Manageability Features

Features	Benefits
DMTF Network Controller Sideband Interface (NC-SI) Pass-through	• Supports pass through traffic between BMC and Controller's LAN functions
Advanced Pass Through (APT)	• Compatible Management Packet Transmit/Receive Support
Manageability and Host Packet Filtering	• Packets that pass the MAC address filters and VLAN address filters are routed to either the Host or a Management Controller
Intel® System Management Bus (SMBus) Pass-through	• Enables BMC to configure the Controller's filters and management related capabilities
Management Component Transport Protocol (MCTP)	• Baseboard management controller (BMC) communication between add-in devices within the platform
Host-Based Application-to-BMC Network Communication Patch (OS2BMC)	• Filtering method that enables server management software to communicate with a management controller via standard networking protocols such as TCP/IP instead of a chipset-specific interface
Private OS2BMC Traffic Flow	• BMC may have its own private connection to the network controller and network flows are blocked
DMTF MCTP Protocol Over SMBus	• Enables reporting and controlling information via NC-SI using the MCTP protocol over SMBus
Firmware Based Thermal Management	• Can be programmed via the BMC to initiate thermal actions and report thermal occurrences
IEEE 802.3 Management Data Input/Output Interface (MDIO Interface or MII Management Interface)	• Enables the MAC and software to monitor and control the state of the PHY
MAC/PHY Control and Status	• Enhanced control capabilities through PHY reset, link status, duplex indication, and MAC Dx power state
Watchdog timer	• The MAC and each PHY supports a watchdog timer to detect a stuck microcontroller
Advanced Error Reporting (AER)	• Messaging support to communicate multiple types/severity of errors
Controller Memory Integrity Protection	• Main internal memories are protected by error correcting code (ECC) or parity bits

Manageability Features *(continued)*

Features	Benefits
Alternative RID Interpretation (ARI)	▪ Enables an interpretation of the Device and Function fields as a single identification of a function within the bus
Device Serial Number	▪ Allows exposure of a unique serial number for each device
Vital Product Data (VPD) Support	▪ Support for VPD memory area
MACsec protected management traffic	▪ Supports a single secure channel for both host and BMC
Flexible MAC Address	▪ MAC address used by a port can be replaced with a temporary MAC address in a way that is transparent to the software layer
iSCSI/FCoE Boot Configuration via management controller	▪ Enables the configuration of iSCSI/FCoE boot code from Flash, via expansion ROM
L3 Address Filters	▪ Four L3 address filters for manageability for both IPv4 and IPv6
Flexible TCO Filters - 4	▪ The flexible 128 filters are a set of manageability filters designed to enable dynamic filtering of received packets
MCTP over SMBus	▪ Allow reporting and controlling of all the information exposed in a LOM device via NC-SI, in NIC devices via MCTP over SMBus
NC-SI Package ID via SDP Pins	▪ SDP pins of the two ports can be combined to encode the NC-SI package ID

Product Codes

Product	Production MM#	Top Marking
Intel® Ethernet Controller X540-AT2	917469	JLX540AT2

For more information on the Intel® Ethernet Controller X540, visit **www.intel.com/go/ethernet**

Printed in USA BJ/TAR/SWU Please Recycle 326917-002US