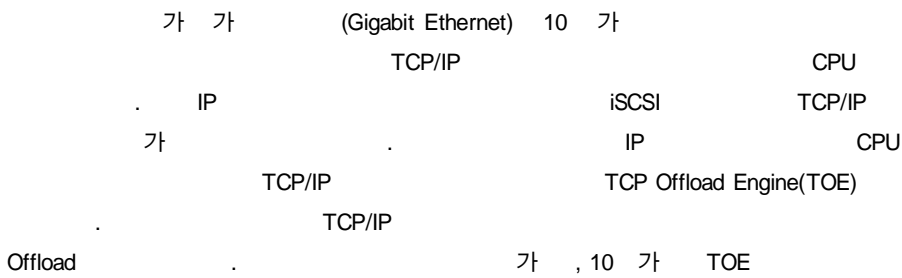


TCP Offload Engine(TOE)

* ** ***

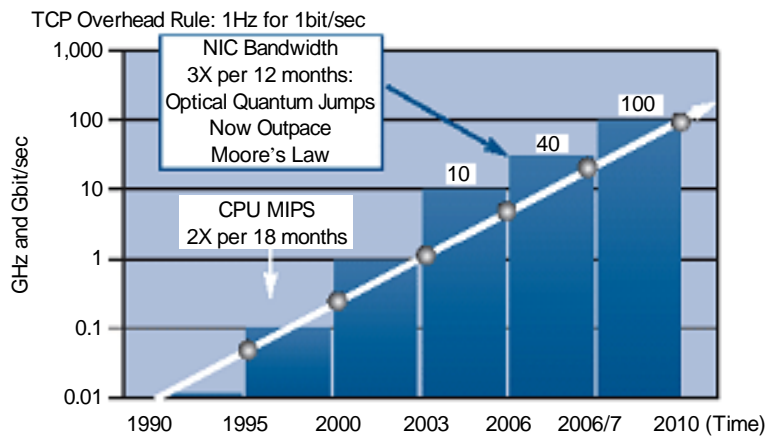


- I.
- II. TCP/IP
- III. TOE
- IV. TOE
- V.

I.
 TCP/IP 가
 TCP/IP
 CPU 가
 1bit/sec TCP/IP 1Hz
 CPU 10/100Mbps
 CPU TCP/IP
 가
 CPU TCP/IP
 (1 (a)) CPU
 18 2 가 NIC
 12 3 가
 가
 [4]. 10 가
 20GHz

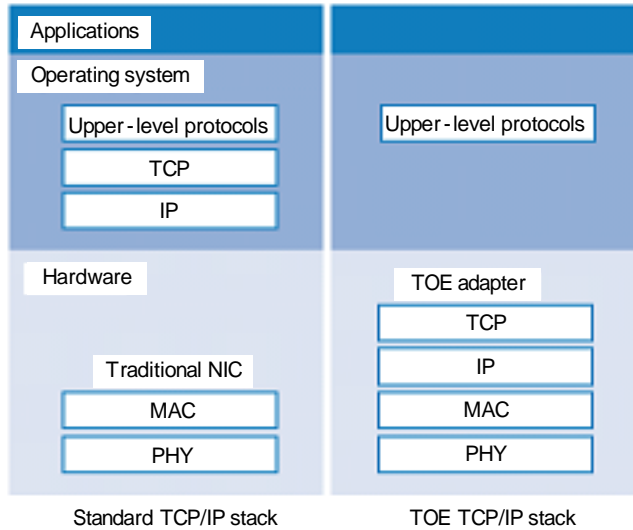
* ETRI /
 ** ETRI /
 *** ETRI /

가 . 2007 가 . TCP/IP
 SCSI iSCSI
 RDMA(Remote Direct Memory
 Access) TCP/IP 가
 TCP/IP Offload Engine(TOE) CPU TCP/IP NIC 가
 TCP/IP 가 . (1 (b)) TOE 가 Transport,



< >: Intel 2002.

(a) CPU NIC [6]



(b) TCP/IP TOE TCP/IP

(1)

Network 가 . 가
 TCP/IP . TOE
 Offloading . TCP/IP
 TOE . TOE

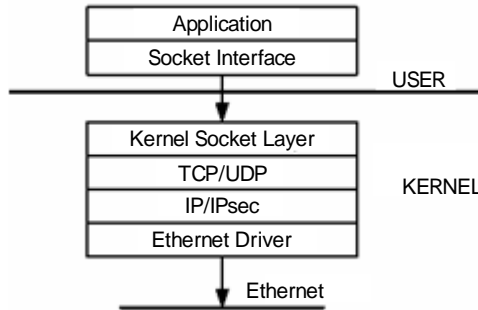
II. TCP/IP

1. TCP/IP

TCP/IP LAN WAN
 . IP IP 가 . TCP
 FTP, HTTP TCP . TCP
 (checksum)
 MSS(Maximum segment size)
 (Sequence number)
 (sliding window)
 (Flow control)
 (Congestion control)
 . TCP 가

2. Traditional TCP/IP

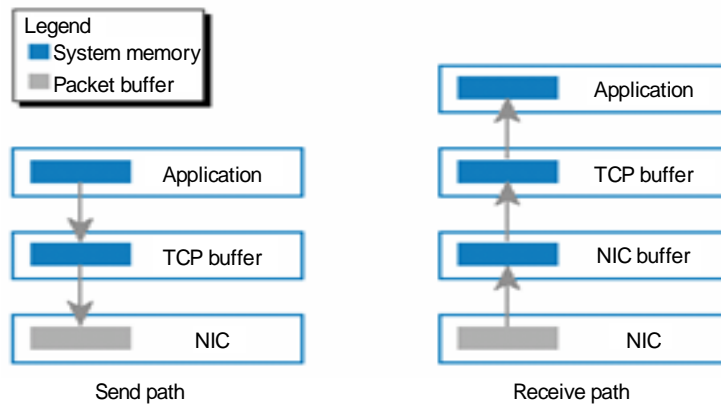
TCP/IP (2)
 OS . TCP
 (socket) 가 .
 (2)
 4~64KB TCP/IP OS



(2) BSD socket

TCP/IP [3]

- MTU(maximum transmission unit) TCP/IP 가
- OS NIC NIC TCP NIC
- DMA , (Interrupts)가 .
- TCP/IP .
- Context Switch : 가
- (user space) (kernel space)
- 가 .
- : (3) NIC .
- NIC OS 가

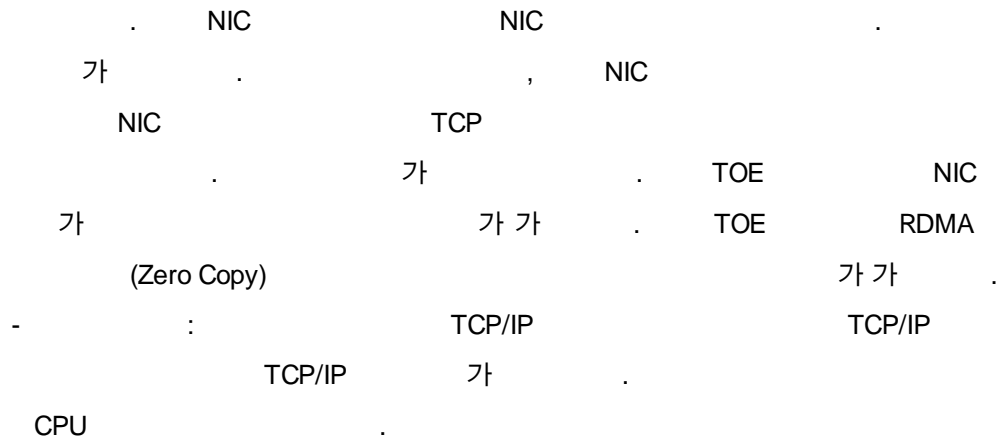


(3)

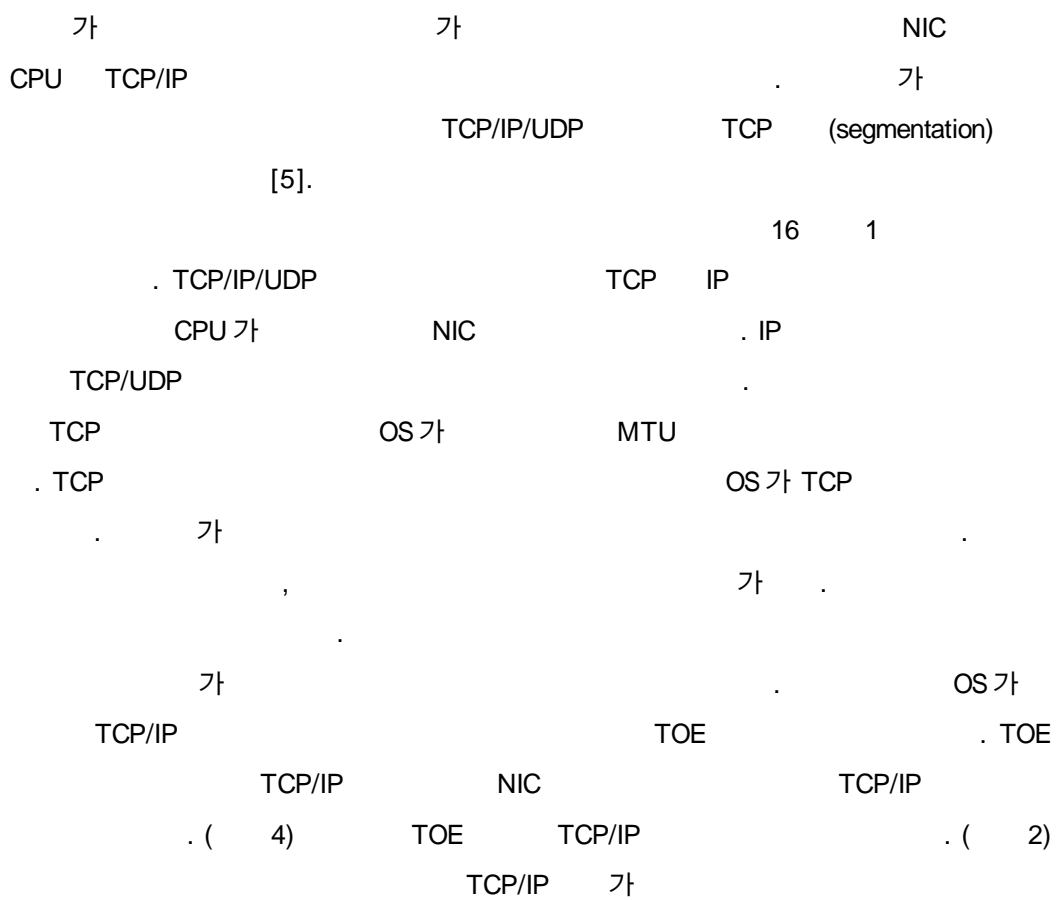
NIC

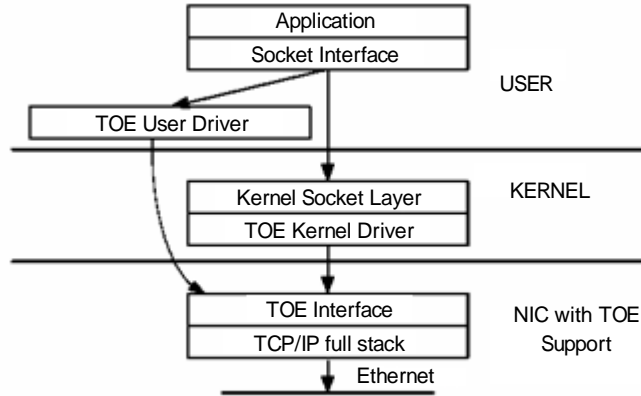
[2]

.....



III. TOE





(4) TOE TCP/IP [3]

TOE 가 .
 TOE 가 . TCP/IP
 (Partial offloading)
 (Full offloading) . (Data path offloading)
 TCP/IP .
 TCP/IP . ACK
 TCP/IP .
 TOE NIC (ASIC) .
 가 가 가 .
 TOE TCP/IP TOE

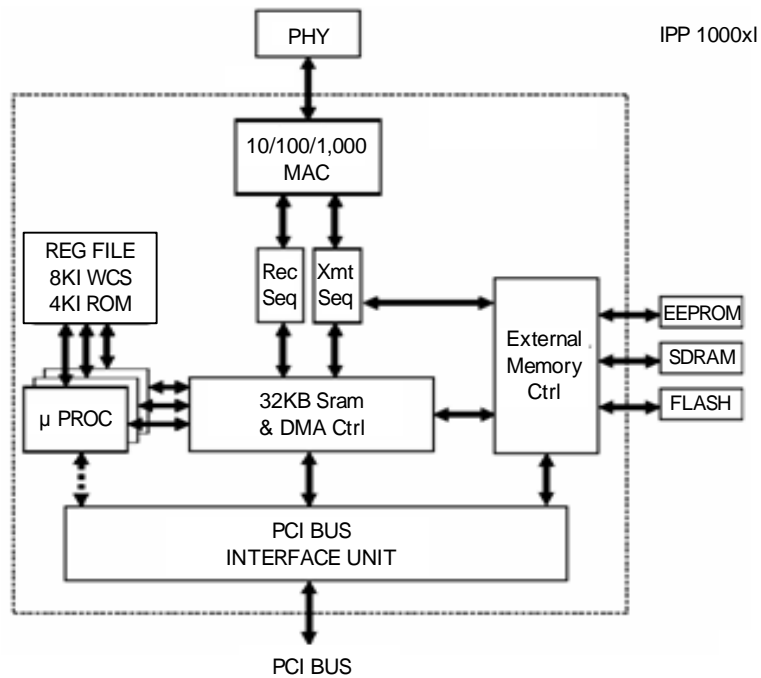
IV. TOE

TOE iSCSI RDMA
 TCP/IP . 가
 2004 10 가 TOE . TOE
 , 10 가 TOE .

1. 가 TOE

가. Alacritech

Alacritech SLIC(session-layer interface control)
 TOE . IPP(Internet Protocol Processor) TOE
 TCP/IP , 가
 . (5) Alacritech TOE IPP 1000x1 .
 10/100/1,000Mbit/s MAC PCI .
 , (Context) 32MB SDRAM
 , ROM FLASH 가 .
 Linux Window OS , 802.3ad Link Aggregation iSCSI
 . IPP 1000x1 가 ACK, Slow-start



(5) Alacritech IPP 1000x1 [6]

. Adaptec

Adaptec NAC-7711 TOE . (

6) Adaptec TCP/IP . No Offload

TCP checksum TCP reassembly 가 가 NIC 가 .

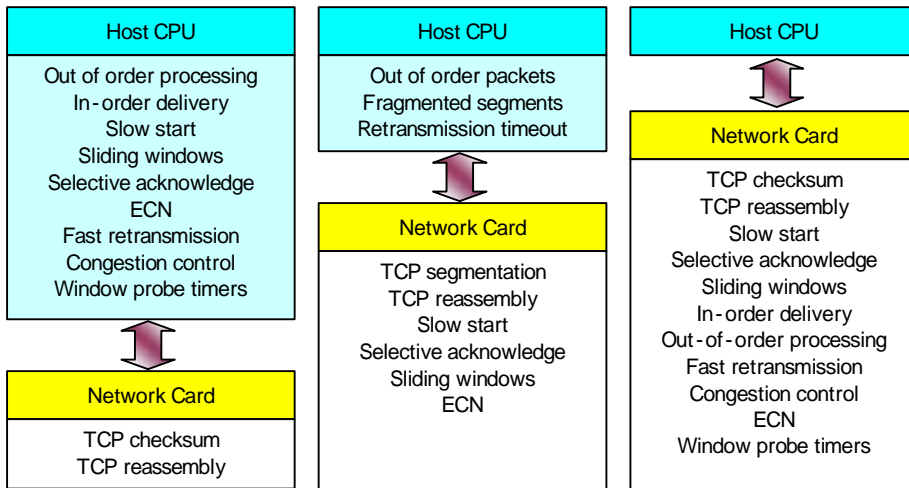
Incomplete Offload Alacritech . Complete

Offload Adaptec TCP/IP .

NAC-7711 가 MAC , 64MB SDRAM

가 1,024 TCP .

OS [8].



(6) [9]

. Broadcom

2004 5 Broadcom BCM5706 TCP iSCSI,

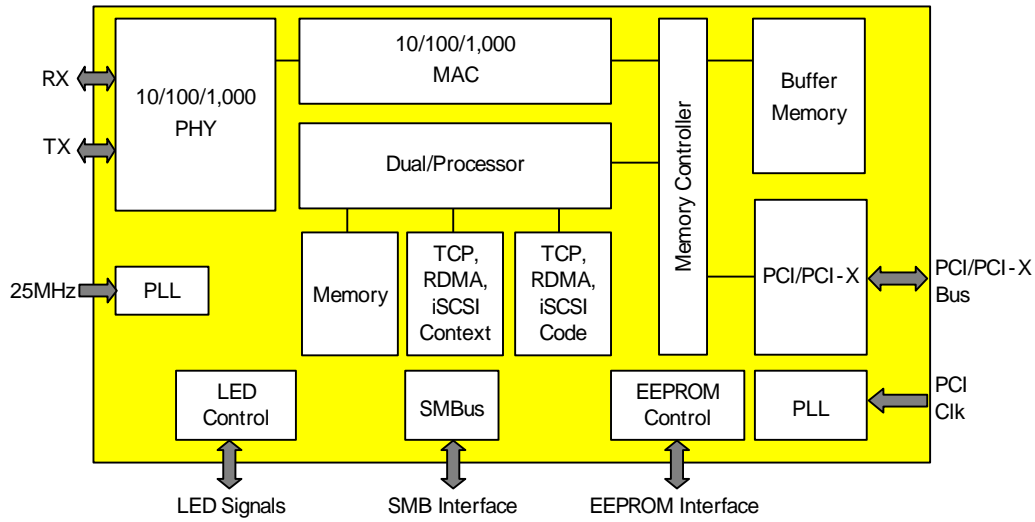
RDMA 가 PHY

PCI, PCI-X . (7) BCM5706

“Fast Path” TCP TOE Microsoft TOE Chimney

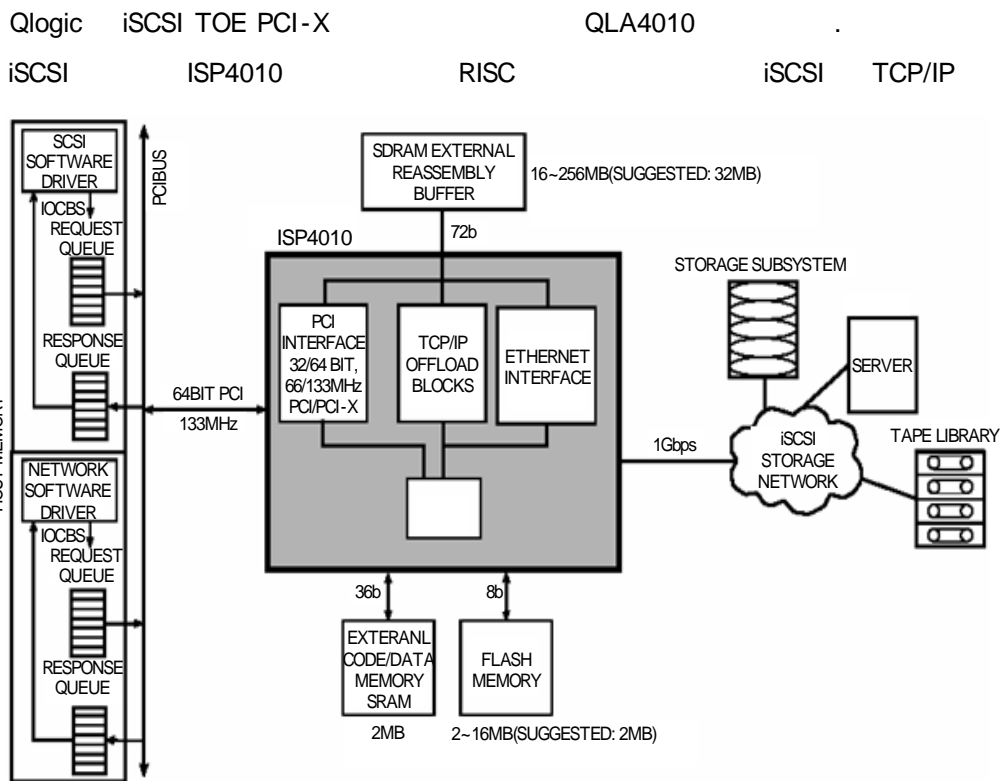
RDMA TOE SDRAM

가 Adaptive 9KB Jumbo Frames



(7) Broadcom BCM5706 [10]

. Qlogic



(8) ISP4010 [11]

. 1Gpbs PCI-X
 가 . ISP4010 (8) Remote backup,

TOE , NVIDIA iReady

< 1> 가 TOE

	Broadcom	Alacritech	Qlogic
Name	BCM5706	IPP1000x1	ISP4010
ASIC Feature	TCP Offload + iSCSI Controller, RDMA Controller +In-band Management	TCP/IP Offload	iSCSI TCP/IP Offload
Application	RDMA NIC(RNIC) iSCSI or iSER HBA TOE chimney enabled Network Accelerator	iSCSI HBA & TOE Network Accelerator	iSCSI HBA
Host Bus	PCI v2.3, PCI-X v1.0	PCI v2.2	PCI v2.2, PCI-X v1.0a
Memory()		32MB SDRAM	32MB SDRAM
MAC	(IEEE 802.3x)	(10/100/1000)	
PHY	10/100/1000 PHY		
TCP/IP Offloading	Data path TCP Offload	Data path Offload + ACK, Slow-start	Complete iSCSI & TCP/IP Offload
UDP offloading	X	X	X
IPv6 support	X	X	X
	Interrupt coalescing 9KB Jumbo Frame	Jumbo Frame	X
OS	Novell Netware, Linux, Windows, Unixware, Solaris	Linux Windows	Solaris Linux Windows
	Adaptec	iREADY(NVIDIA)	Silverback
Name	NAC-7711	ethernetMAX	iSNAP2110
ASIC Feature	Full TCP Offload	HW iSCSI, TCP/UDP/IP/ARP/RARP Offload, IPSec	iSCSI & TCP Offload
Application	TOE Network Accelerator	iSCSI HBA	iSCSI HBA
Host Bus	PCI v2.2	PCI-X	PCI v2.2, PCI-X
Memory()	64MB SDRAM	32MB SDRAM	32MB SDRAM
MAC	(100/1000)		
PHY		10/100/1000 PHY	
TCP/IP offloading	Full Offload	iSCSI Offload Engine Full TCP/UDP/IP/ARP/RARP Offload	Complete TCP Offload
UDP offloading	X	O	X
IPv6 support	X	X	Firmware Upgrade
	Interrupt coalescing 9KB Jumbo Frame	9KB Jumbo Frame DES/3DES, AES encryption	64K TCP connection
OS	Linux	Linux, Windows	Linux

* Full Offload Complete Offload 가 .
 ** iREADY 가 .

ethernetMAX 가 PHY, TCP/IP , iSCSI IPsec
 . Siverback TOE TCP/IP
 64K TCP . < 1> 가 TOE .

2. 10 가 TOE

가. Chelsio

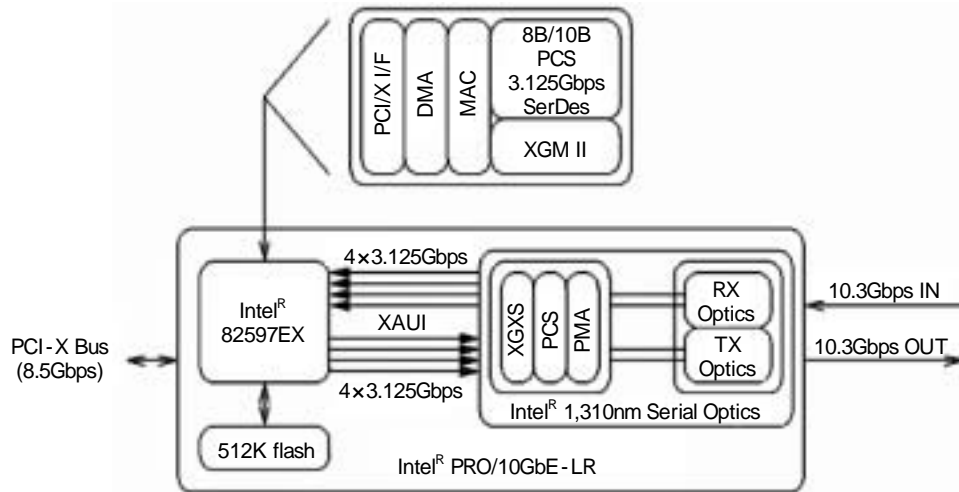
Chelsio T110 TOE 10 가 iSCSI,
 RDMA , PCI-X . T110 가



(9) Chelsio T110

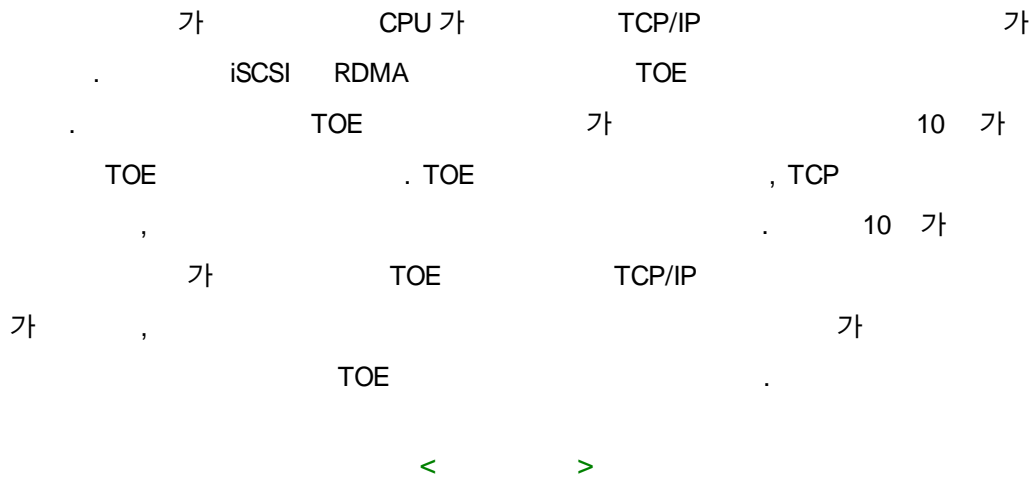
< 2> Chelsio T110 [12]

RDMA	iWARP0.7
Storage	iSCSI initiator & target mode / iSCSI header & CRC generation & check iSCSI PDU recovery / iSCSI sessions up to 64K / DDP
TCP	TCP compliance(RFC 793, 1122, 2525, 2988)/Congestion Control(RFC 258)/High performance extensions(RFC 1323)/Increased initial window(RFC 2414)/SACK(RFC 2018)/64K connections/200K connections set-up/teardown per second
UDP	Checksum offload
IP	Path MTU discovery(RFC 1191)/Multiple IP addresses supported/Socket backlog/DiffServ(RFC 2474)
Ethernet	10Gbps(IEEE 802.3ae)/Ether II & 802.3 encapsulated frames/Priority & VLAN tagging(IEEE 802.1p/Q)/Flow control(IEEE 802.3x)/Jumbo Frames (9KB)
Security	Defense against sequence number attacks(RFC 1948)/Protection against internet attacks/Filters packets based on configurable policies
OS	Linux for TOE HBA(Kernel 2.4 and 2.6)
SDK	SDK is provided for TOE and iSCSI modes
PCI-X	PCI-X 1.0a (133MHz)
Media	T110SR (Multi-mode Fiber)/T110LR (Single-mode Fiber)



(11) 82597EX [14]

V.



[1] Bukasa Tshilombo, "Worldwide Wireless Local Loop Markets, Mobile Communications Worldwide Focus Report," Gartner Group, June 14, 1999.

[2] Sandhya Senapathi, Rich Hernandez, "Introduction to TCP Offload Engines," <http://www.dell.com/powersolutions>

[3] Venkata Krishnan, Tim Miller, "Comparing TCP/IP Ethernet with Advanced Switching for Intra-System Processor - to - Processor Communication," StarGen. Inc.

- [4] E. Yeh, H. Chao, V.Mannem, J. Gervais, and B. Booth, "Introduction to TCP Offload Engine," 10Gigabit Ethernet Alliance, April 2002.
- [5] Intel 82547EI Gigabit Ethernet Controller, <http://www.intel.com>
- [6] Alacritech IPP 1000x1 data sheet, July 2002, p.6.
- [7] Adaptec, Networking Technology paper, "Increase Performance of Network - Intensive Applications with TCP/IP Offload Engines,"
http://www.adaptec.com/worldwide/product/markeditorial.html?cat=/Technology/NAC&prodkey=nac_techpaper
- [8] Adaptec, TOE NAC 7711 data sheet, <http://www.adaptec.com>
- [9] Adptec, "What are the differences between a NIC, complete offload, and Incomplete offload?"
- [10] Broadcom, BCM5700 data sheet, <http://www.broadcom.com>
- [11] Qlogic, iSCSI Controller ISP4010, <http://www.qlogic.com>
- [12] Chelsio T110 10Gbps Protocol Engine, <http://www.chelsio.com>
- [13] Gary Gumanow, Carl Wilson, "Preparing for 2004~2005 Networking Transitions," Fall IDF, 2003.
- [14] Wu-chun Feng, Justin(Gus) Hurwitz, Harvey Newman, Sylvain Ravot, R. Les Cottrell, Olivier Martin, Fabrizio Coccetti, Cheng Jin, Xiaoliang(David) Wei, and Steven Low, "Optimizing 10-Gigabit Ethernet for Networks of Workstations, Clusters, and Grids: A Case Study," ACM/IEEE SC 2003: High-Performance Networking and Computing Conference, November 2003.