

Die Performance von Gigabit-Ethernet versprechen einige Hersteller sogenannter intelligenter Netzwerkkarten mit ihren Produkten auf der Basis von 10-Base-T/100-Base-TX. Eigene Prozessoren auf den Netzwerkadaptoren sollen zudem die CPU entlasten, indem sie Operationen für den TCP/IP-Stack übernehmen oder komplexe Berechnungen für die Verschlüsselung von IPSec durchführen.

DIRK PELZER

Schnelle Schlauberger fürs Netz

Bei Grafik-Karten und RAID-Controllern gehören sie längst zum Alltag: Schnelle Prozessoren, die komplexe und ressourcenintensive Berechnungen anstellen, um die CPU des Host-Systems zu entlasten. Anders wären komplexe Multimedia-Anwendungen und schnelle durch RAID-Technologie ausfallsicher gestaltete Plattensubsystem kaum denkbar. Anders die Situation bei Netzwerkkarten. Intelligenz, die dafür sorgt, dass sich der Server oder die Power-Workstation eines Anwenders auf die ihre eigentliche Aufgabe, nämlich die Bereitstellung von Anwendungen konzentrieren kann, ist selten. Statt dessen werden vor allem bei Servern die oft knappen CPU-Ressourcen damit belastet, sich um Standardaufgaben, wie die Segmentierung von TCP-Paketen zu kümmern. Aber auch die Verschlüsselung von Datenpaketen mittels IPSec kostet die CPU eines Systems unnötig viel Rechenzeit.

TCP-Offloading entlastet CPU

Viel Raum für Verbesserungen also, auf den sich derzeit einige der etablierten Hersteller, wie die Platzhirsche 3COM oder Intel stürzen, der aber auch Newcomern im eigentlich gesättigten Segment der 10/100-Mbit-Ethernet-Karten Chancen eröffnet. So hat die kalifornische High-Tech-Schmiede Alacritech drei Netzwerkadapter für Server unter Windows-NT und Windows 2000 im Programm, die mit einem, zwei oder vier Ports ausgerüstet selbst Gigabit Ethernet das Wasser reichen sollen. Eine von Alacritech entwickelte Architektur namens SLIC (Session-Layer Interface Card) soll dabei den Datendurchsatz beflügeln. Der in einem ASIC (Application Specific Integrated Circuit) implementierte Internet-Protocol-Processor (IPP) vereint Routinen zur schnelleren Bearbeitung des TCP/IP-Stacks in sich.

Die Leistung des IPP kommt dabei allerdings erst zum Tragen, wenn der Sitzungsaufbau über TCP stattgefunden hat. Auch andere „Verwaltungstätigkeiten“ des TCP-Protokolls, wie Sitzungsabbau, Timeouts und Fehlerbehandlungen überlässt der IPP nach wie vor der CPU des Hostsystems. Der Hersteller rechtfertigt die eingeschränkte Auslagerung der TCP-Funktionen auf den Adapter damit, dass nur rund fünf Prozent des Netzwerkverkehrs auf das TCP-Management entfielen, während die Implementierung in Silizium sehr aufwendig wäre und die Kosten für einen Adapter unverhältnismäßig in die Höhe getrieben würden. Leider enden die Einschränkungen damit noch nicht, denn auch für das UDP-Protokoll bietet Alacritech keinerlei hardware-basierende Beschleunigung an. So können die Anwendungen aus dem Bereich Streaming Video, Multicast oder Voice-Over-IP keinen Nutzen aus dem auf der Netzwerkkarte vorhandenen Prozessor ziehen. Auch eine Unterstützung für IPSec-Verschlüsselung sucht man auf der Alacritech-Karte vergebens.

Intelligente Architektur beschleunigt Durchsatz

Der Hersteller führt aber neben dem Prozessor noch drei weitere Technologien ins Feld, mit denen er Performancesteigerungen verspricht. So haben die Ingenieure des Unternehmens den Datenverkehr auf dem Speicherbus des Hostsystems optimiert, um Ineffizienzen zu beseitigen, die durch die protokoll-abhängige Behandlung konventioneller Netzadapter entstehen. So soll etwa ein einziger Transfer eines 64 Kilobyte großen Datenpaketes in der Lage sein, den Inhalt des kompletten Level-1-Caches und etwa 40 Prozent des Level-2-Caches ungültig zu machen. Schuld daran soll die Vorgehensweise anderer Adapter sein, die eingehende Datenpakete zunächst von der Netzwerkkarte in einen Puffer innerhalb des Netzwerkkartentreibers transferieren. Anschließend werden diese Daten zusammen mit dem Inhalt des TCP/IP-Puffers in den Level-2-Cache des Systems

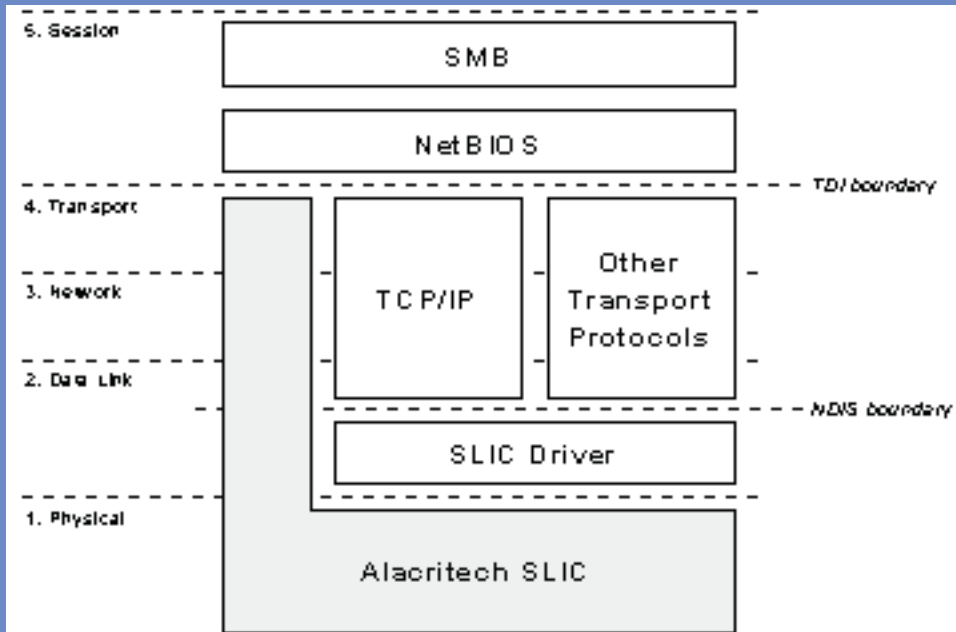
eingelezen. Von dort landen die Daten schließlich bei der jeweiligen Applikation. Für ausgehende Pakete kehrt sich die Vorgehensweise entsprechend um. Dieses umständliche Vorgehen umgeht die Alacritech-Architektur und tauscht Datenpakete unter Umgehung von Puffern und Caches direkt zwischen dem Netzwerkadapter und der jeweiligen Zielapplikation aus. Dieselbe Architektur trägt auch dazu bei, die Zahl der Interrupts, die der Prozessor aufgrund Datentransfers abzuarbeiten hat zu reduzieren. Da jede Interruptbehandlung mit Aufwänden des Betriebssystems, wie zum Beispiel dem Umschalten von Kontexten einher geht, sollte eine Reduktion der Interrupts zur Verbesserung der Gesamt-Performance beitragen. Alacritech rechnet bei einem Transfer eines 64 Kilobyte großen Server-Message-Block- (SMB) Frames bei herkömmlichen Architekturen mit circa 2140 Interrupts pro Sekunde. Dieser Wert kommt dadurch zustande, dass der 64 Kilobyte große SMB-Frame in 44 einzelne TCP-Frames zu je 1500 Bytes zerlegt wird. Konventionelle Adapter bearbeiten etwa vier TCP-Frames pro Interrupt, woraus sich die angegebene Rate von 2140 Interrupts pro Sekunde errechnet. Die Alacritech-Architektur hingegen benötigt laut Hersteller für alle 44 Frames lediglich zwei Interrupts, nämlich einen zu Beginn des Datentransfers und einen am Ende. Daraus resultiert eine gegenüber dem vorherigen Wert um 82 Prozent reduzierte Interrupt-Rate von 390 pro Sekunde.

Als weiteren Knackpunkt zur Leistungsoptimierung haben die Alacritech-Ingenieure den PCI-Bus identifiziert. Da bei diesem Lesezugriffe etwa fünfmal langsamer vonstatten gehen, als über den Speicherbus der CPU, versprechen sich die Kalifornier von der Reduktion dieser Zugriffe einen spürbaren Performancegewinn. Konventionelle Netzwerkkarten benötigen üblicherweise pro Frame drei PCI-Bus-Transaktionen, in der sie zunächst Daten aus dem Descriptor-Ring des Treibers zur Netzwerkkarte transferieren, anschließend den Frame von der Netzwerkkarte in den Puffer des Kartentreibers übertragen und daraufhin den aktualisierten Descriptor-Ring von der Netzwerkkarte aus überschreiben. Da der Alacritech-Adapter die protokollrelevanten Aktivitäten selbst durchführt, entfallen die genannten Aktionen und der IPP übergibt der Anwendung die Daten in einem Zug direkt von der Karte.

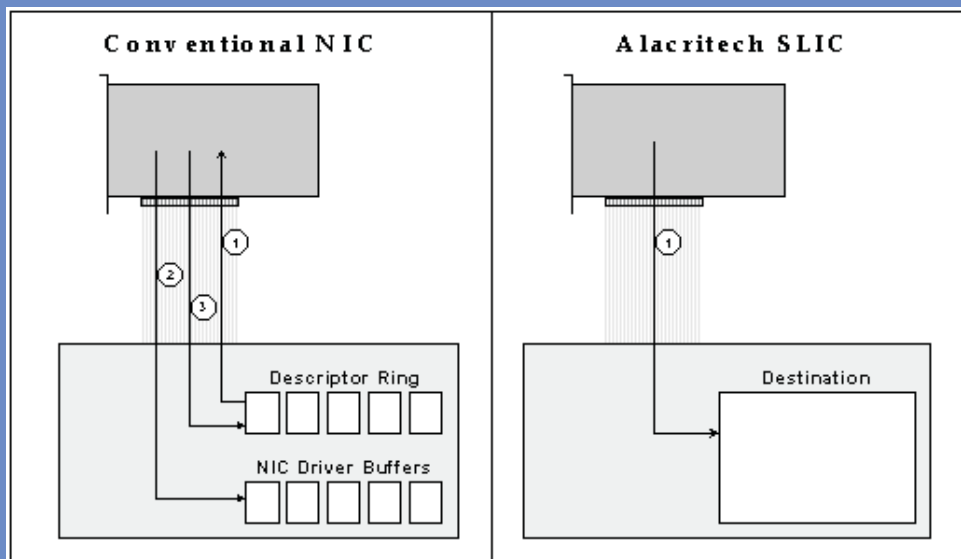
3COM bietet IPSec- und TCP-Offloading

Einen ähnlichen Weg, wie Alacritech beschreitet auch 3COM mit seiner neuen Generation von Netzwerkkarten vom Typ 3CR990 für Server. Im Gegensatz zum Newcomer Alacritech bietet 3COM mit seinen Netzadaptern aber nicht nur die Möglichkeit, die Host-CPU von TCP-Operationen zu entlasten, sondern auch die zur Verschlüsselung notwendigen Berechnungen für IPSec durchzuführen. Hierfür hat 3COM neben einem 3XP-Prozessor genannten ASIC, der einen 10/100-Mbit/s-Media-Access-Control-Baustein und einen ARM9 RISC-Prozessor vereint, auch einen Verschlüsselungschip implantiert. Dieser ist in der Lage, verschiedene Verschlüsselungs- und Hash-Algorithmen, wie 3DES, DES, MD5 und SHA-1 mit 56 Bit beziehungsweise 168 Bit in Silizium zu berechnen. Die Hardwareimplementierung von IPSec wurde von 3COM in Zusammenarbeit mit Microsoft im Rahmen der Entwicklungsarbeiten für Windows 2000 realisiert. Unter Benutzung der dort vorhandenen Programmierschnittstellen verspricht der Kartenhersteller eine Entlastung von bis zu 33 Prozent der CPU, wenn die Karte in einem Windows-2000-System eingebaut ist. Für die Bearbeitung der TCP-Segmentierung durch die 3COM-Netzwerkkarte gibt der Hersteller eine Einsparung von etwa 13 Prozent der CPU-Belastung an. Derzeit ist allerdings ein simultanes Offloading von TCP-Segmentierung und IPSec unter Windows-2000 ausgeschlossen. Es geht entweder das eine oder das andere. Neben Windows-2000 unterstützt 3COM auch andere Betriebssysteme, wie Windows 9x, Novell Netware oder Linux.

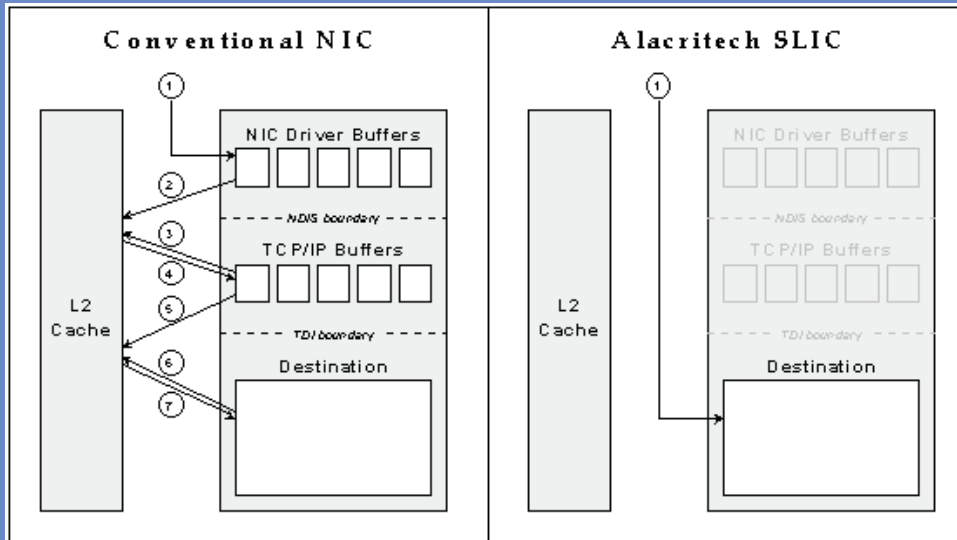
Ähnlich wie Alacritech optimiert auch 3COM die für den Datenaustausch über den PCI-Bus notwendigen Transaktionen. Die als Parallel Tasking II bezeichnete Technologie erlaubt dabei die Übertragung eines vollständigen Ethernetpakets mit einer Größe von 1514 Bytes innerhalb einer einzigen Bus-Master-Operation, vorausgesetzt der PCI-Chipsatz unterstützt derart große Raten. Damit ist anstelle von 24 Taktzyklen zur Übertragung eines Ethernet-Frames nur noch ein einzelner Taktzyklus erforderlich.



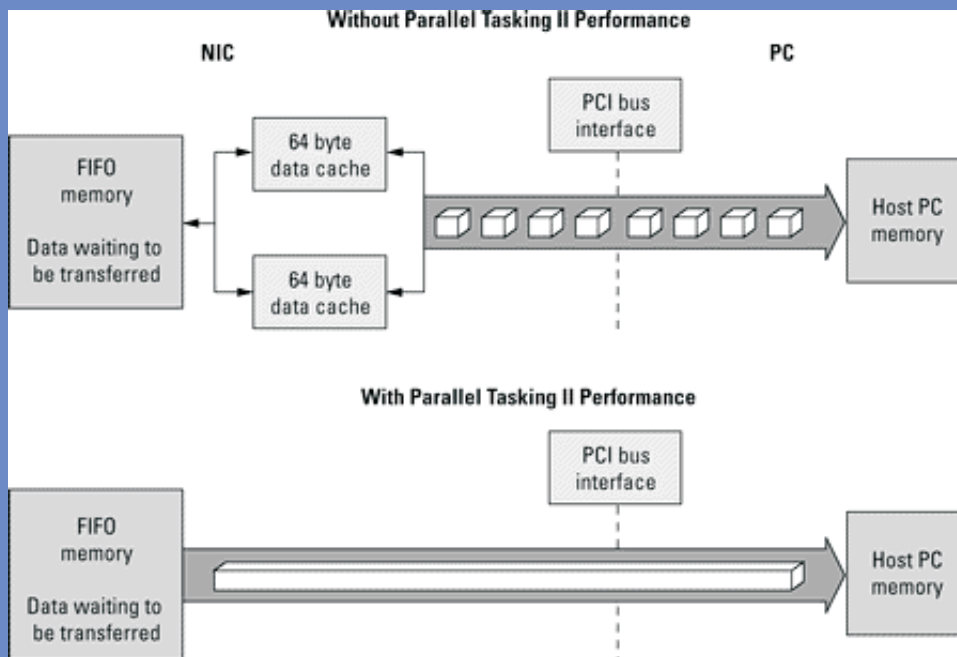
OSI.GIF: Der Aufgabenbereich der Alacritech-Karte erstreckt sich bis zur vierten Schicht des OSI-Referenzmodells



PCI.GIF: Der Alacritech-Adapter reduziert die Zahl der Transaktionen über den PCI-Bus auf ein Minimum



Memory.gif: Die Architektur der Alacritech-Adapter vermeidet aufwendige Aktionen auf dem Speicherbus



COMParallelTasking: Auch 3Com optimiert die Architektur seiner Netzwerkkarten, um Reibungsverluste auf dem PCI-Bus zu minimieren.

Intel nur mit IPSec-Offloading im Rennen

Etwas weniger Engagement als die Konkurrenz zeigt derzeit Intel im Bereich intelligenter Netzwerkkarten. Der Chip-Gigant konzentriert sich derzeit voll und ganz auf das Auslagern von IPSec-Funktionalität auf die Karte. Die Intel PRO/100 S-Adapter haben dazu einen 82550 Fast-Ethernet-Controller implantiert bekommen, der mit einem Co-Prozessor für die IPSec-Verschlüsselung ausgestattet ist. Damit arbeitet der Adapter problemlos unter Windows 2000 und einer als Packet-Protect bezeichneten Zusatzsoftware auch unter Windows-NT und Windows-98. Die Performancevorteile, die Intel mit dem IPSec-Offloading erzielt, liegen laut einer von Intel in Auftrag gegebenen Studie auf ähnlichem Niveau, wie die des Mitbewerbers 3COM.

Neben dem Prozessor, der die Intelligenz aller beschriebenen Netzwerkkarten ausmacht, verfügen die jeweiligen Produkte natürlich auch über zahlreiche Features ihrer weniger intelligenten Geschwister. So erlauben beispielsweise alle Karten das Load-Balancing von zwei oder mehr Ports in einer Karte oder über mehrere Karten hinweg. Damit lässt sich nicht nur der Datendurchsatz erhöhen, sondern auch die Ausfallsicherheit eines Serversystems erhöhen, denn die Wahrscheinlichkeit, dass mehr als ein Port ausfällt, ist doch verhältnismäßig gering. Auch Features wie die Unterstützung mehrerer VLANs über einen Link stellen für die meisten dieser Karten eine Selbstverständlichkeit dar.

Intelligenz macht sich nicht immer bezahlt

Um den Versprechungen der Hersteller im Bereich intelligenter Netzwerkkarten auf den Zahn zu fühlen, haben unsere amerikanischen Kollegen den Vier-Port-Adapter von Alacritech genauer untersucht und Performancemessungen durchgeführt (siehe www.nwfusion.com/reviews/2001/0115rev2.html). Die Ergebnisse waren allerdings eher ernüchternd und von der versprochenen Performance im Gigabit-Ethernet-Bereich weit entfernt. So ermittelten sie je nach Anwendung einen Performancegewinn zwischen zweieinhalb und 16 Prozent. Der größte Gewinn ergab sich dabei erwartungsgemäß bei der

Übertragung großer Dateien, da hier die Segmentierung der TCP-Frames einen erheblichen Einfluss auf den Datendurchsatz hat. Wenn hingegen viele kleine Datenbank-Transaktionen gefragt waren, konnte der Alacritech-Adapter keine Punkte sammeln und lag gleich auf mit einer zum Vergleich getesteten Konfiguration bestehend aus vier Standard Intel Pro100+ Netzwerkkarten. Der wahrscheinlichste Grund für dieses Verhalten dürfte darin bestehen, dass bei vielen kleinen Transaktionen die Zahl die zum Auf- und Abbau einer TCP/IP-Session notwendige Rechenzeit den größten Anteil ausmacht, während der Anteil der TCP-Segmentierung kaum ins Gewicht fällt. Da der Alacritech-Adapter das Management der TCP/IP-Sessions aber der CPU des Hosts überlässt, fielen die Unterschiede entsprechend gering aus.

Die Ergebnisse waren zudem, wenn auch marginal, von der Anzahl der in dem Serversystem vorhandenen CPUs abhängig. So lag der mit der Alacritech-Karte erzielbare Performancegewinn mit einem Vier-Prozessor-System bei 14 Prozent, während er bei einem Server mit nur einer CPU 16 Prozent betrug. Anders als bei der Messung des reinen Datendurchsatzes stellt sich die Situation bei der CPU-Belastung des Serversystems dar. Hier konnten die Alacritech-Adapter sowohl bei der Ein- als auch bei der Vier-Prozessor-Konfiguration mit einer gegenüber den Intel-Karten 70 Prozent geringeren CPU-Auslastung glänzen.

Fazit

Ob sich die Mehrkosten für intelligente Netzwerkkarten lohnen, hängt stark vom jeweiligen Einsatzgebiet und der auf dem Serversystem vorhandenen CPU-Auslastung ab. Bei sogenannten General-Purpose-Servern, also Systemen, die für File- und Printsharing ebenso zuständig sind, wie für das Messaging oder eine Datenbank, dürfte im Mittel kein merklicher Performancevorteil spürbar sein. Ebenso dürften reine Datenbanksysteme, auf denen primär kleine Transaktionen in der Größe einiger weniger Ethernet-Frames ausgeführt werden, kaum von der Intelligenz auf der Netzwerkkarte profitieren. Hier lohnt sich eher die Investition in eine zusätzliche CPU, um das Gesamtsystem zu entlasten. Reine

File-Server hingegen, auf denen sich viele große Dateien tummeln haben aufgrund der TCP-Segmentierungs-Problematik erwartungsgemäß den größten Performancegewinn zu erwarten. Generell empfehlenswert sind intelligente Netzwerkkarten, wenn die CPU-Auslastung eines Servers bedingt durch die darauf laufenden Applikationen bereist sehr hoch ist und sehr viele Benutzer auf das System zugreifen. Hier kann die durch intelligente Netzwerkkarten erzielbare Reduktion der CPU-Belastung auch bei Mehrprozessorsystemen zu einer spürbaren Leistungssteigerung beitragen.

Der Einfluss von TCP-Sgmentierung auf die Performance

Die Segmentierung von TCP-Frames, die notwendig ist, wenn ein Datenblock die maximal zulässige Ethernet-Rahmengröße von 1513 Bytes überschreitet, belastet die CPU eines Host-Systems in erheblichem Maße. Bei der Segmentierung ist es für jedes zu übertragende Datensegment notwendig, IP-Header-Informationen zu duplizieren und eindeutige TCP-Headerinformationen zu generieren, um die Pakete am Zielort wieder in der richtigen Reihenfolge zusammen setzen zu können. Demzufolge muss die Host-CPU speziell bei großen Dateien eine Vielzahl von Operationen für die Segmentierung durchführen. Die dafür aufgewendete Rechnerzeit fehlt dann entsprechend für die Ausführung von Applikationen.

Web-Links

<http://www.alacritech.com/home.html>:

Homepage des auf TCP-Offloading spezialisierten Herstellers Alacritech

<http://www.alacritech.com/html/whitepapers.html>:

Von Alacritech verfasste oder beauftragte Whitepapers zum Thema TCP-Offloading und Total-Cost-of-Ownership

<http://www.3com.com/products/dsheets/400517.html>:

Produktübersicht der 3COM-Adapter mit 3XP-Prozessor für TCP- und IPSec-Offloading

http://www.3com.com/technology/tech_net/white_papers/500650.html:

3COM-Whitepaper über die Parallel-Tasking-II-Technologie

http://www.intel.com/network/products/pro100s_srvr_technical.htm:

Technische Übersicht der Intel Pro/100-S-Server-karten mit IPSec-Offloading

http://www.intel.com/network/documents/pdf/intel_ipsec_final.pdf:

Whitepaper über Performancevergleich zwischen Intel Pro/100 S und einer 3COM-Karte mit 3XP-Prozessor

<http://www.nwfusion.com/reviews/2001/0115rev2.html>:

Testbericht der Network World USA über die Alcri-tech-Server-Adapter

GLOSSAR

3XP

Von der Firma 3COM entwickelter Prozessor für Netzwerkkarten, der rechenintensive Operationen für die TCP-Segmentierung und Checksummenbildung übernimmt und dadurch die Host-CPU entlastet.

ASIC

Die Application-Specific-Integrated-Circuit (ASIC) ist eine nach kundenspezifischen Bedürfnissen entwickelte Schaltung zur Realisierung komplexer Funktionen auf einem Chip

IPP

Der von der Firma Alacritech entwickelte Internet-Protocol-Processor (IPP) ist ein ASIC, der die CPU eines Host-Systems von rechenintensiven Operationen, wie zum Beispiel TCP-Segmentierung und Checksummenbildung entlastet.

IPSec

IPSec ist ein Standardisierungsvorschlag der IETF, in dem Verfahren und Protokolle für einen herstellerübergreifenden sicheren und geschützten Datenaustausch mittels des IP-Protokolls festgelegt werden.

Parallel-Tasking II

Von 3COM entwickelte Technologie zur Beschleunigung des Datendurchsatzes von Netzwerkkarten.

SLIC

Session-Layer-Interface-Card (SLIC) ist eine von der Firma Alacritech entwickelte Architektur zur Steigerung des Datentransfers über Netzwerkkarten nach dem 10-Base-T/100-Base-TX-Standard.

SMB

Server Message Block (SMB) ist ein Client-Server-Protokoll von IBM, das unter Windows-Betriebssystemen, sowie OS/2 läuft. Es dient dazu, Dateien, Verzeichnisse, Drucker und weitere Ressourcen über ein Netzwerk verfügbar zu machen.

TCP Offloading

Auslagern von TCP-Operationen, wie zum Beispiel die Segmentierung von der CPU eines Hosts auf die Netzwerkkarte.

TCP/IP

Das Transmission-Control-Protocol (TCP) ist für den Einsatz in paketvermittelten Netzen bestimmt und setzt unmittelbar auf dem Internet-Protokoll (IP) auf. Es bietet virtuelle Verbindungsdienste zur gesicherten und reihfolgegerichtigen Verbindung zwischen Computersystemen.

UDP

Das User-Datagram-Protokoll (UDP) ist ein Transportprotokoll (Schicht 4) des OSI-Referenzmodells und unterstützt den verbindungslosen Datenaustausch zwischen Rechnern. Das UDP wurde definiert, um auch Anwendungsprozessen die direkte Möglichkeit zu geben, Datagramme zu versenden und damit die Anforderungen transaktionsorientierten Verkehrs zu erfüllen. UDP baut direkt auf dem darunter liegenden IP-Protokoll auf.

Voice-over-IP

Telefonie über IP-Datennetze wie das Internet.

Zur Person

DIPL. ING. DIRK PELZER arbeitet als freier Consultant und Journalist in München. Er betreibt ein Storage Labor für verschiedene namhafte Fachzeitschriften. Zudem beschäftigt er sich mit Speichernetzen und Hochverfügbarkeit.