

АППАРАТНАЯ РЕАЛИЗАЦИЯ ИНТЕЛЛЕКТУАЛЬНЫХ ФУНКЦИЙ УДАЛЕННОГО ПРЯМОГО ДОСТУПА К ПАМЯТИ

Я. В. Шелеметьева

Кафедра информационно-вычислительных систем, Поволжский государственный технологический университет

Йошкар-Ола, Республика Марий Эл, Россия

E-mail: yanabaikova@yandex.ru

В работе описаны основные принципы функционирования сетевых адаптеров, их назначение, механизмы передачи данных в сеть/из сети. Рассмотрены интеллектуальные сетевые адаптеры, решающие проблему загрузки центрального процессора вычислительного узла в высокопроизводительной сетевой системе. Представлен алгоритм работы технологии удаленного прямого доступа к памяти.

ВВЕДЕНИЕ

Сетевой адаптер (сетевая карта, сетевая плата, Ethernet-адаптер, NIC) - это периферийное устройство вычислительного узла, непосредственно взаимодействующее со средой передачи данных, которая прямо или через другое коммуникационное оборудование связывает его с другими узлами вычислительной системы [1].

Сетевой адаптер (СА) вместе со своим драйвером реализует канальный и физический уровни эталонной модели взаимодействия открытых систем в конечном узле сети. Точнее, в сетевой операционной системе пара адаптер-драйвер выполняет только функции физического уровня и подуровня канального уровня (управление доступом к среде, англ. – Media Access Control, MAC). Подуровень управления логической связью (англ. – Logical Link Control, LLC) обычно реализуется модулем операционной системы (ОС), единым для всех драйверов и сетевых адаптеров [2]. Сетевой адаптер совместно с драйвером выполняют две операции: передачу и прием кадра.

СЕТЕВЫЕ АДАПТЕРЫ TCP/IP

В современных компьютерных системах вычислительные узлы объединены между собой сетевыми технологиями, и значительный объем данных передается по сети. Протокольный стек TCP/IP обеспечивает надежную доставку пакетов в сетях Ethernet, и за его поддержку отвечает центральный процессор. Контроль последовательности номеров пакетов, подтверждение передачи, стратегия управления потоком данных методом скользящего окна, алгоритмы контроля перегрузки, сегментация и повторная сборка пакетов, а также вычисление контрольной суммы требуют весьма значительных процессорных ресурсов. При этом осуществляется многочисленное каскадное буферирование-копирование: сначала из буферов сетевого адаптера в буферы сетевого протокола, из него в буферы транспортного протокола, далее в буферы приложения и из

них в буферы дисков, что также подразумевает занятость процессора [5], (рис.1).

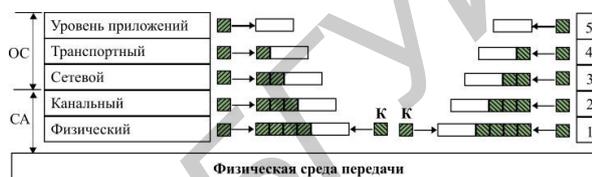


Рис. 1 – Стандартный механизм передачи данных в TCP/IP сетях, где К – кадр

Полоса пропускания устройств высокопроизводительной системы, как правило, выше, чем пропускная способность технологий Fast Ethernet или 10 Gigabit Ethernet, поэтому две или более копии в стеке протоколов и проверка контрольной суммы не столь существенно загружают центральный процессор [7]. Однако, с использованием 10 Gigabit Ethernet и технологий, позволяющих передавать данные по сети с еще большей скоростью, загрузка центрального процессора существенно увеличивается. Центральный процессор не справляется с увеличивающейся нагрузкой. Таким образом, в сети увеличиваются задержки, уменьшается скорость передачи данных, снижается эффективность вычислений отдельных конечных узлов, а значит, и вычислительной системы в целом [3].

Для решения данных проблем был предложен выход – освободить процессор от обработки TCP/IP трафика и перенести эти вычисления в сетевой адаптер, что позволит направить всю вычислительную мощность центрального процессора на выполнение критически важных приложений [7].

Альтернативой выполнению устройств ТОЕ из отдельных компонентов является интеграция соответствующих программ и оборудования на кремниевом кристалле. Речь идет о реализации ТОЕ на базе программируемой микросхемы FPGA (Field-Programmable Gate Array) или в виде специализируемой микросхемы ASIC [6].

Адаптеры ТОЕ так же называют интеллектуальными сетевыми адаптерами, так как они

выполняют функции центрального процессора вычислительного узла [6]. На адаптеры перенесены следующие функции протоколов транспортного уровня (TCP) и сетевого уровня (IP):

1. Функции протокола IP, реализуемые в адаптере TOE: маршрутизация сообщений – передача данных по конкретному маршруту, адаптация передаваемого пакета по правилам промежуточных локальных сетей, управление потоком данных с целью предотвращения блокировки памяти, расчет КПК заголовков, управление срочностью, надежностью, пропускной способностью данных.
2. Функции протокола TCP, реализуемые в адаптере TOE: установление логического канала путем обмена запросом и подтверждением на соединение, упаковка и распаковка пакетов на концах транспортного соединения, управление потоком (получатель при подтверждении правильности передачи передает сообщение о размере окна ответов – диапазон номеров пакетов, которые получатель готов принять), пометка о срочности передаваемых данных – возможность управления скоростью передачи [1].

Non-TOE адаптеры – адаптеры, которые осуществляют передачу сетевых данных традиционным способом: стек протоколов TCP/IP выполняется программным обеспечением вычислительного узла, загружая процессор обработкой пакетов и передачей подтверждений.

Для интеллектуальных сетевых адаптеров TOE дополнительно разработана технология удаленного прямого доступа к памяти RDMA (англ., Remote Direct Memory Access), которая позволяет передавать данные из памяти одного компьютера в память другого компьютера без участия центрального процессора и операционной системы [4]. При этом исключается участие центрального процессора в обработке команд и необходимость пересылки данных из памяти приложения в буферную область ОС, то есть данные пересылаются напрямую на соответствующий сетевой контроллер (рис.2).

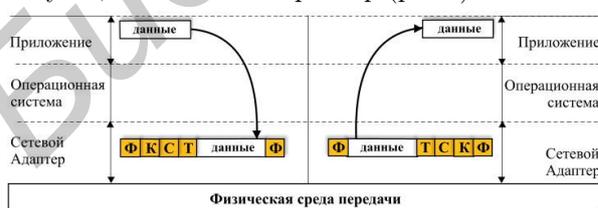


Рис. 2 – Алгоритм работы технологии RDMA, где Ф – заголовок физического уровня, К – заголовок канального уровня, С – заголовок сетевого уровня, Т – заголовок транспортного уровня

Из рисунка видно, что количество копий данных между буферами операционной системы значительно уменьшилось. Данные из буферов приложения помещаются сразу в буфер сетевого адаптера, который добавляет все требуемые за-

головки. На приемной стороне, соответственно, происходят обратные действия: адаптер, разобрав сетевые заголовки, помещает входящие данные из своих буферов напрямую в буферы приложения.

В системах Ethernet, технология RDMA входит в стек протоколов IWARP. В формате пакетов, передаваемых с помощью RDMA, указан буфер памяти приложения, в который должны быть доставлены входящие данные.

ЗАКЛЮЧЕНИЕ

1. Использование сетевых адаптеров с механизмом TOE не только обеспечивает быстрое повышение производительности серверов, но в долгосрочной перспективе (благодаря экономии серверных ресурсов) позволит уменьшить число серверов, необходимых для того, чтобы реализовать необходимое качество обслуживания пользователей [6].
2. Для того чтобы повысить производительность сетевых приложений, администраторы оснащают серверы более быстродействующими ЦП и/или устанавливая дополнительные ЦП. Сетевые адаптеры с механизмом TOE снижают нагрузку на ЦП и тем самым увеличивают ту часть его вычислительной мощности, которая может быть использована для выполнения приложений. Кроме того, применение этих адаптеров, возможно, позволит упростить конфигурацию сети за счет исключения из нее устройств, ускоряющих работу серверов [6].

1. Олифер, Н. А. Роль коммуникационных протоколов и функциональное назначение основных типов оборудования корпоративных сетей [Электронный ресурс] / Н. А. Олифер, В. Г. Олифер. – Режим доступа: <http://citforum.ru/nets/protocols/index.shtml>. – Дата доступа: 12.09.2014.
2. Олифер, В. Г. Компьютерные сети. Принципы, технологии, протоколы: Учебник для вузов 3-е изд. / В. Г. Олифер, Н. А. Олифер. – СПб.: Питер, 2008. – 958 с.: ил.
3. Mannem, V. Increasing network performance with a TCP/IP offload ASIC / V. Mannem, D. Sommers // Applied Computing. – 2002. – p. 4
4. Understanding iWARP: Eliminating Overhead and Latency in multi-Gb Ethernet Networks, NETEFFECT. pp 5p. Summer 2002.
5. Xu, J. Time for TOE. The Benefits of 10 Gbps TCP Offload / J. Xu, N. Borkar, V. Erraguntla, D. D. Clark, V. Jacobson // A Chelsio Communications White Paper. – 2013. – 7 p.
6. Friedman, D. Building firewalls with intelligent network interface adapter card / D. Friedman, D. Nagle // School of Computer Science, Carnegie Mellon University. – 2001. – 17 p.
7. Виноградова И. Л. Расчёт параметров сегмента волоконно-оптической высокоскоростной системы передачи / И. Л. Виноградова, А. А. Кашбиев, А. Х. Султанов // Вестник Поволжского государственного технологического университета. Сер.: Радиотехнические и инфокоммуникационные системы. – 2012. – № 1(15). – с. 82-91.