

# ОПТИМИЗАЦИЯ ПРОИЗВОДИТЕЛЬНОСТИ ВЫСОКОСКОРОСТНЫХ ВЫЧИСЛИТЕЛЬНЫХ СЕТЕЙ

В.И. ПЛАСТИНИН, В.И. МОИСЕЕВ

Пермский государственный национальный исследовательский  
университет, 614990, Пермь, Букирева, 15

Современные локальные сети строятся на протоколе Ethernet и стеке протоколов TCP/IP. Такие сети высокопроизводительны, эффективны и масштабируемы. Опыт многих лет закрепил за этими технологиями лидерство в областях ЛВС, Метро и даже WAN.

Топологии звезды или дерева самые распространенные схемы в сетях Ethernet, где принято выделять функциональные уровни доступа, распределения и ядра сети.

Такие схемы хорошо масштабируются, однако в них присутствует множество точек агрегации трафика. Если нагрузка в такой точке превышает норму, то производительность соответствующего участка сети падает. Самое медленное звено в цепочке называют *bottleneck* (с англ. узкое место).

В сетях малого масштаба и малых скоростей превышение нагрузки приводит к потерям трафика некоторых соединений. Протокол TCP имеет встроенные механизмы отслеживания таких событий и повторно передает трафик. Потеря нескольких сегментов TCP соединения в любой сети неизбежна и это не критично.

В сетях большого масштаба с высокими скоростями даже кратковременная (меньше секунды) перегрузка способна привести к потерям гигабайт информации, нарушить потоки трафика тысяч соединений.

Потеря нескольких сегментов в высокоскоростной сети может затормозить TCP соединение на nepозволительно долгий период, а в отдельных случаях приводит к нарушению ключевой функции TCP - гарантированной доставки [1].

Даже если сеть обеспечивает полосу пропускания 10-Gbps, производительность передачи файлов ограничена следующими факторами, не зависящими от сети [2]:

- Производительность операций ввода/вывода с внешней памятью ограничивает максимально достижимую производительность таких приложений, как например FTP, NFS, SMB.
- Реализация TCP стека зависит от операционной системы, новые версии ОС поддерживают различные опции улучшающие работу стека.
- Размер буфера сокета, устанавливаемый приложением, ограничивает максимальный размер окна TCP.

- Производительность CPU влияет на производительность разгрузки TCP в целом, а также для многопроцессорных систем доступны различные опции, как например RSS(Recive-Side Scaling).
- Производительность передачи по сравнению с приемом. Сетевые адаптеры обеспечивают более высокую производительность при передаче, чем при приеме, потому что передача обеспечивает больше возможностей к разгрузке(например, разгрузка большими сегментами LSO).
- Конфигурация клиента и сервера должна обеспечивать необходимый уровень производительности.

Правильный способ проверить эксплуатационные характеристики серверной системы, которая абстрагирована от сервера хранения ввода/вывода и от ограничений файловых приложений заключается в проверке из памяти такими инструментами, как Netperf и NetIQ Chariot, которые позволяют установить нужные значения размера буфера сокета. Кроме того, необходимо использовать несколько экземпляров инструментов тестирования(например, Netperf), чтобы задействовать все доступные ядра сервера.

Тесты, входящие в рекомендаций [3] включают в себя:

- Тесты производительности между виртуальными машинами (с виртуальными машинами, работающими на разных хостах VMware ESX, где каждый хост VMware ESX оснащен адаптером 10 Gigabit) с одной или более виртуальной машиной на хосте VMware ESX.
- Тесты производительности между виртуальной машиной и физическими клиентами, с одним или несколькими клиентами и одной или более виртуальной машиной на хосте VMware ESX.

Сетевые адаптеры предлагают ряд возможностей для оптимизации производительности, которые рассматриваются ниже. На представленных диаграммах строится зависимость полосы пропускания, от размера буфера сокета и размера сообщений.

TCP Segment Offload: Когда эта опция включена, то операционная система может передавать сообщения большого размера сетевому адаптеру, а сетевой адаптер будет сегментировать их на основе MSS. Эта функция освобождает центральный процессор от необходимости сегментации TCP. На рис. 1 видно, что при включении данной опции производительность передачи существенно возрастает(около 40%).

TCP checksum offload: TCP checksum опция включает на сетевом адаптере

вычисление контрольной суммы TSO при операциях передачи и приема, которая ограждает процессор от необходимости вычислять контрольную сумму. Выигрыш в производительности варьируется в зависимости от размера пакета. Маленькие пакеты получают малое увеличение производительности от использования этой опцией, в то время как большие пакеты достигают большой экономии. Экономия для максимального блока передачи (MTU) в 1500, как правило, около 5 процентов снижения нагрузки на процессор (рис. 2), а для MTU в 9000 (Jumbo Frames), экономия составляет примерно 15 процентов снижения нагрузки на процессор.

Jumbo фреймы: Jumbo фреймы (9-КВ пакеты) значительно повышают максимально достижимую производительность (рис. 3). Jumbo фреймы включены в виртуальном коммутаторе VMware Distributed Switch vNetwork (VDS)

Помимо возможностей предлагаемых сетевыми адаптерами, можно добиться существенного прироста производительности оптимизировав TSO стек гостевой ОС (рис. 4).

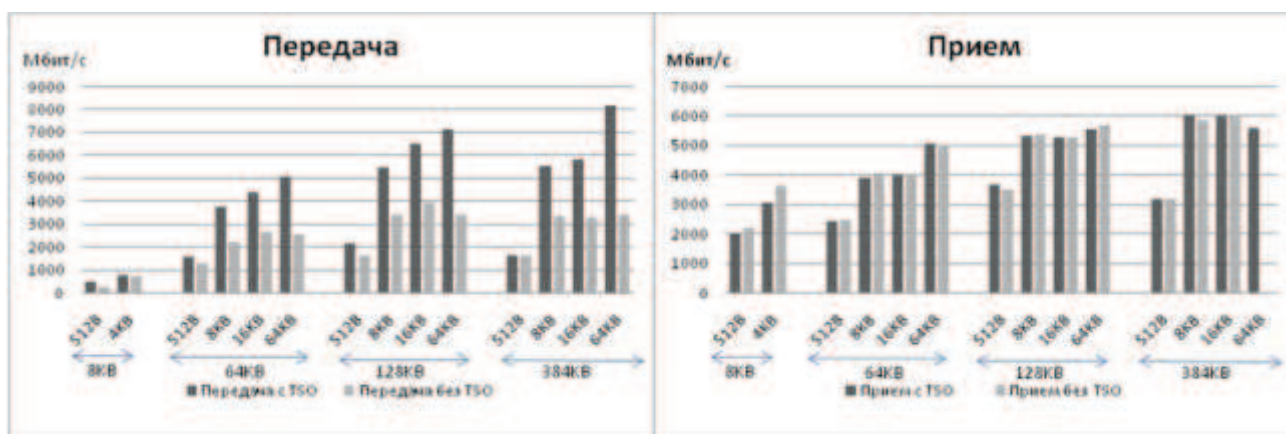


Рис. 1. Влияние TSO на сетевую производительность

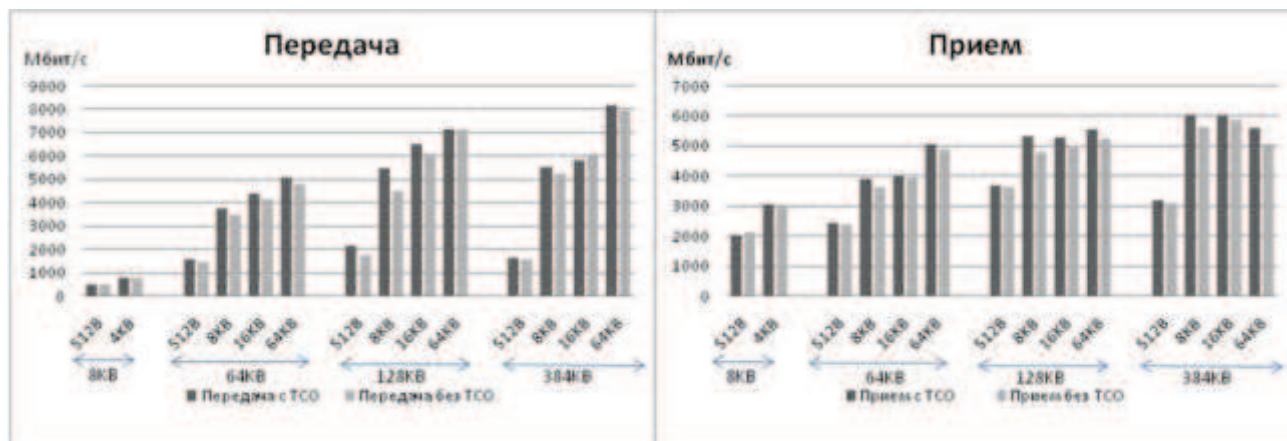


Рис. 2. Влияние TSO на сетевую производительность

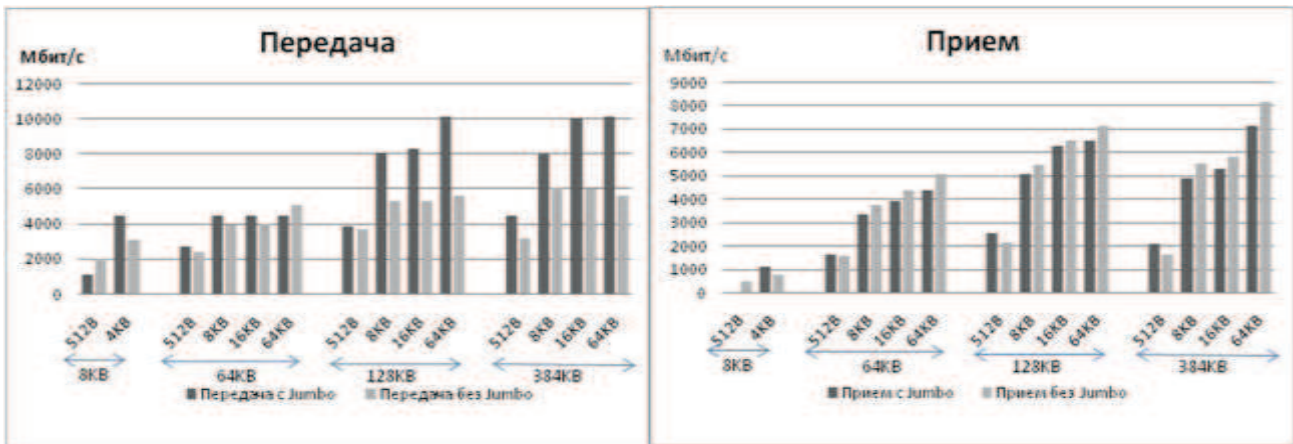


Рис. 3. Влияние Jumbo фреймов на сетевую производительность

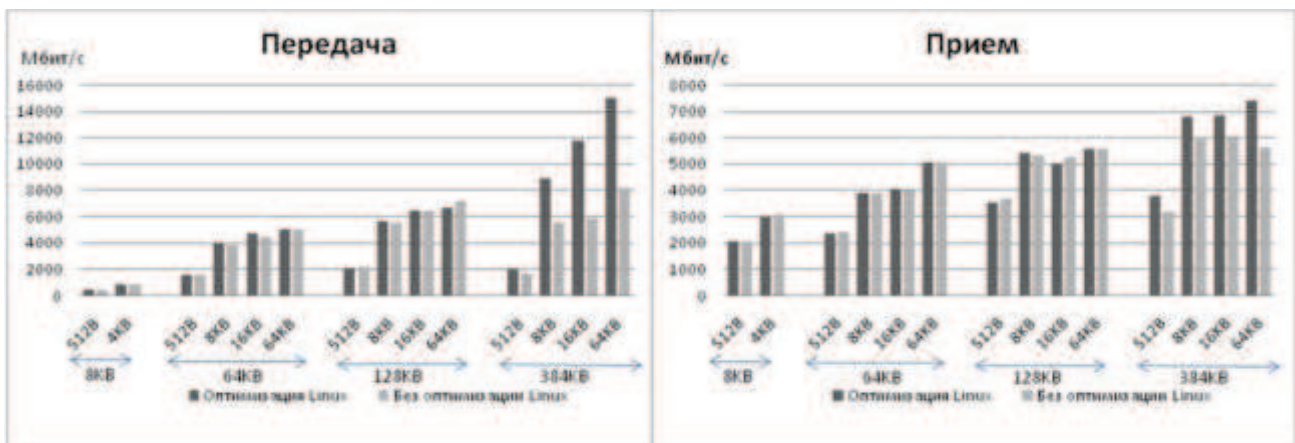


Рис. 4. Влияние оптимизации TCP-стека гостевой ОС на сетевую производительность

#### СПИСОК ЛИТЕРАТУРЫ

1. Nasif Ekiz, Transport Layer Reneging [URL] <http://www.eecis.udel.edu/~amer/PEL/poc/pdf/EkizPhDdissertation.pdf>
2. Cisco corp. Testing Virtual Machine Performance with VMware vSphere 4 on 10 Gigabit Networks Design Guide
3. VMware corp. 10Gbps Networking Performance [URL] [http://www.vmware.com/pdf/10GigE\\_performance.pdf](http://www.vmware.com/pdf/10GigE_performance.pdf)
4. VMware corp. Performance Best Practices for VMware vSphere 4.0 [URL] [http://www.vmware.com/pdf/Perf\\_Best\\_Practices\\_vSphere4.0.pdf](http://www.vmware.com/pdf/Perf_Best_Practices_vSphere4.0.pdf)
5. Tierney Brian L. TCP Tuning Techniques for High-Speed Wide-Area Networks [URL] <http://gridmon.dl.ac.uk/nfn/>