

Lightweight communication between domains in a single machine



Mar 9, 2007

Kangho kim
System Software Team
ETRI



Design & Implementation

Design considerations

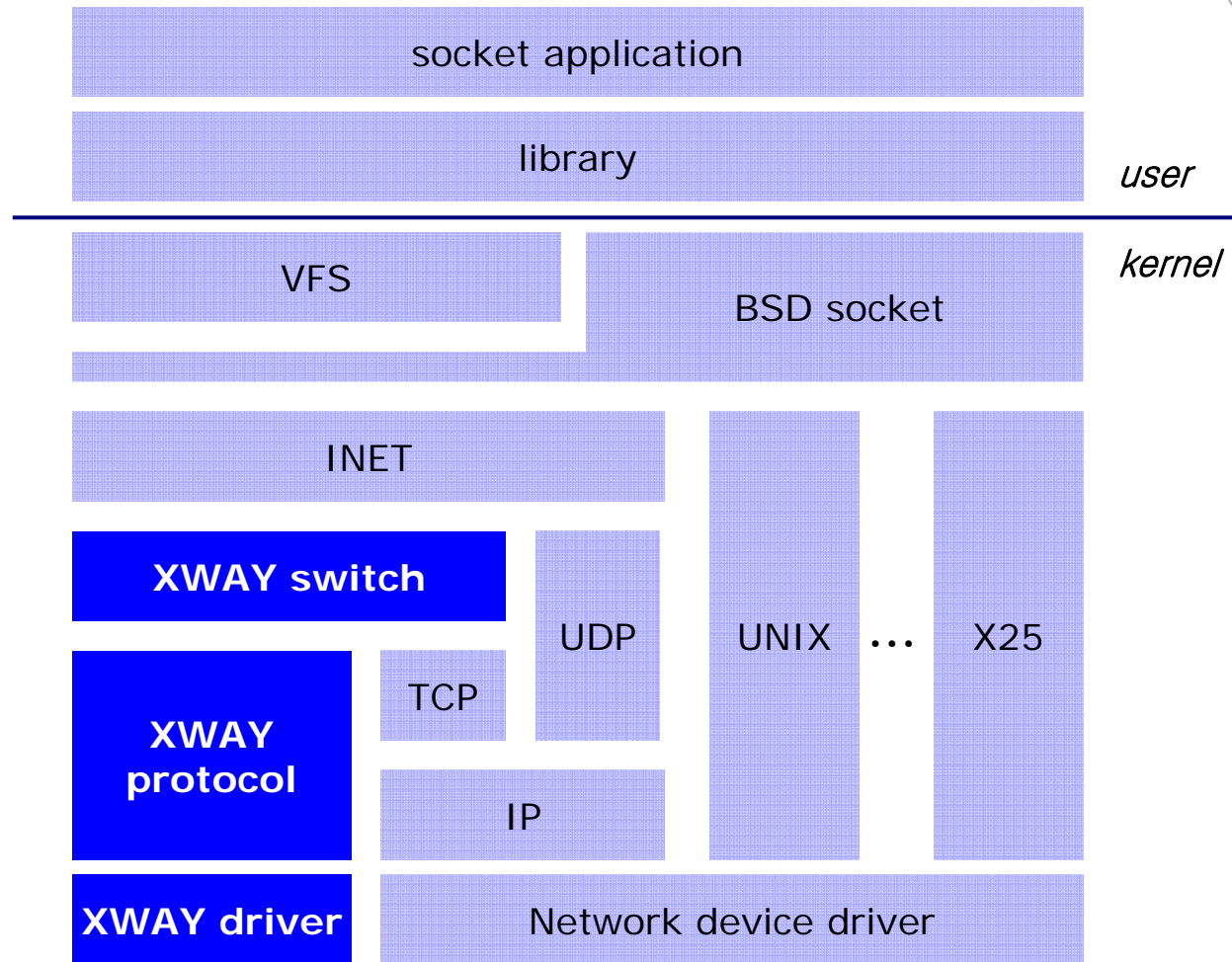
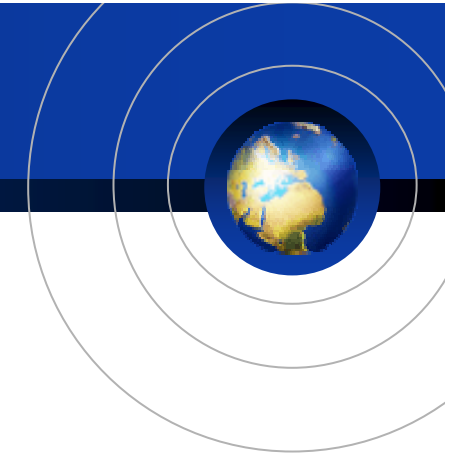


- Higher performance
 - Higher bandwidth, lower latency
- Binary compatibility

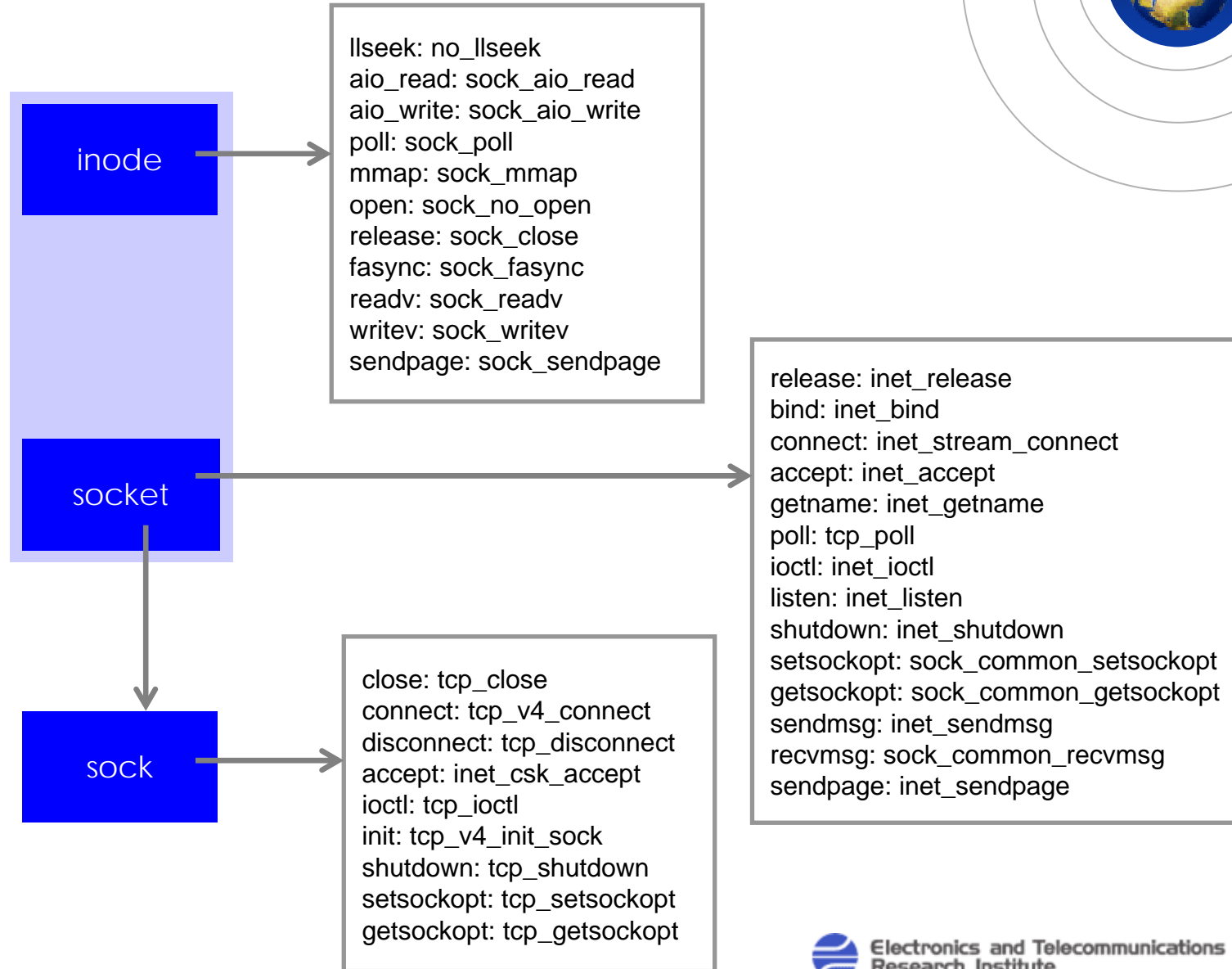
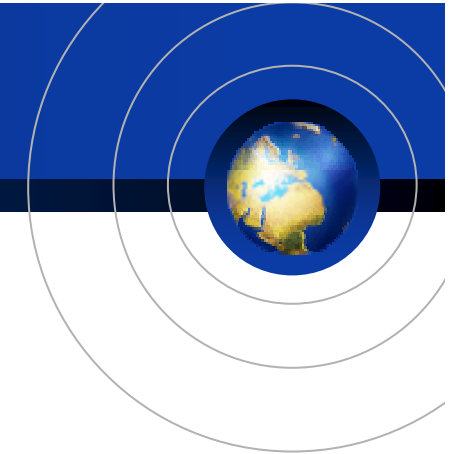
- Development expense
 - Dec. 2006 ~ Feb. 2007
 - 2 kernel programmers
- No modification to Xen if possible
- TCP only

- Minimal kernel code patch
- Xway
 - is selectable
 - don't affect native socket communication with extra-domains

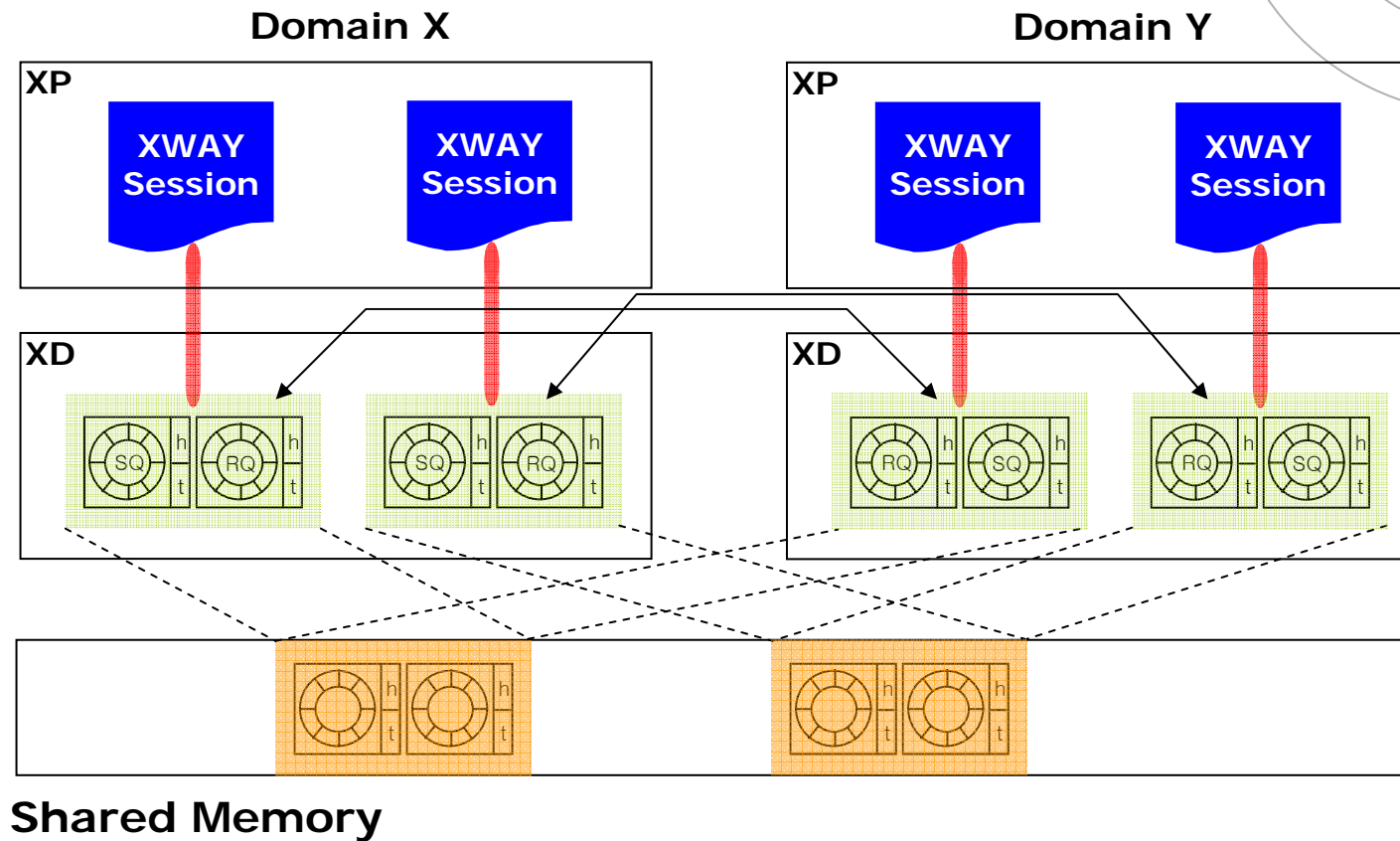
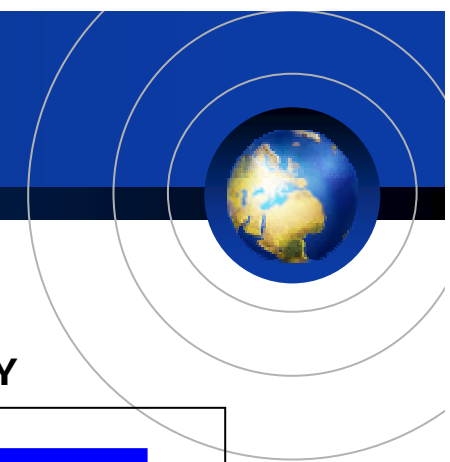
XWAY architecture



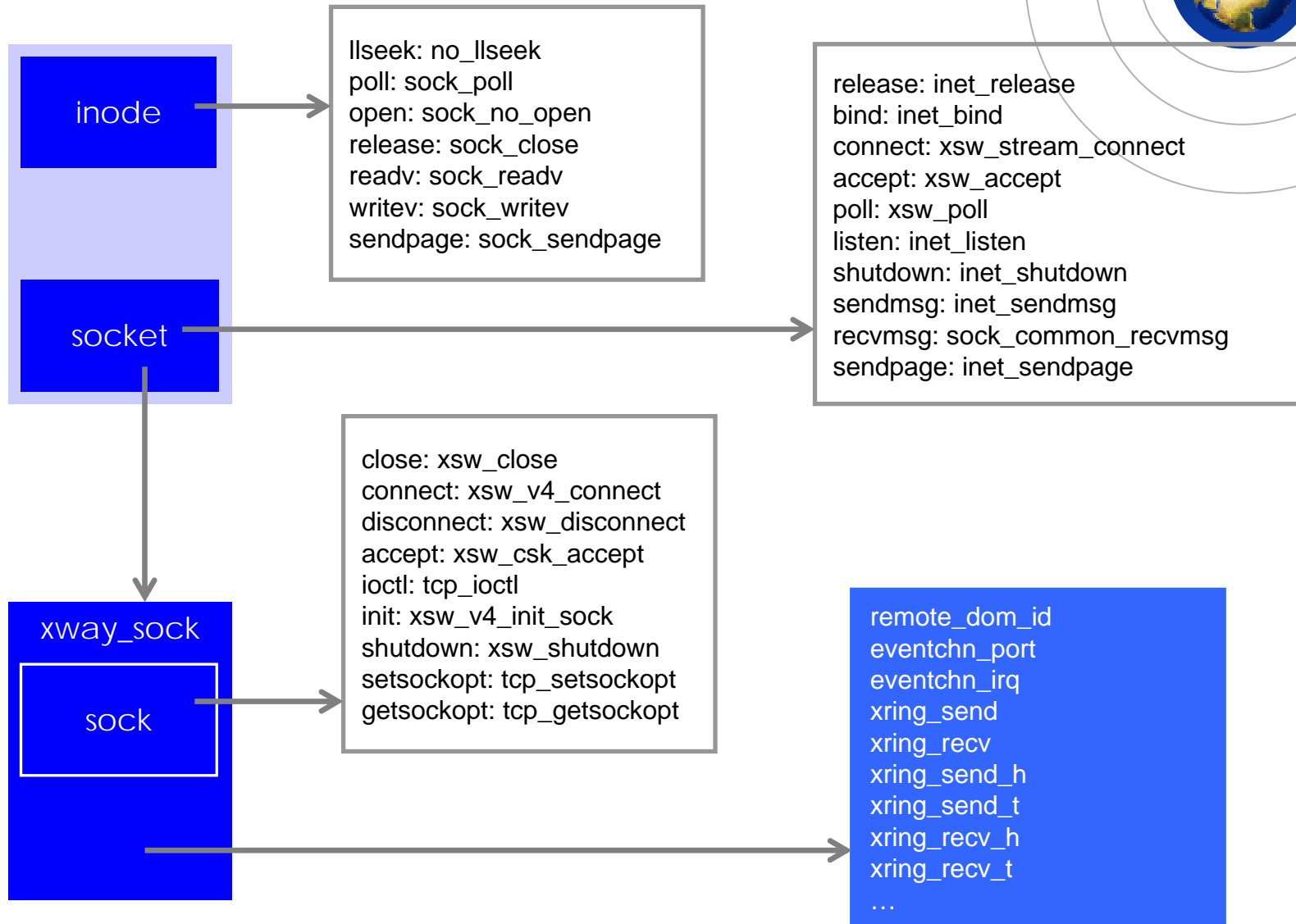
Native socket data structure



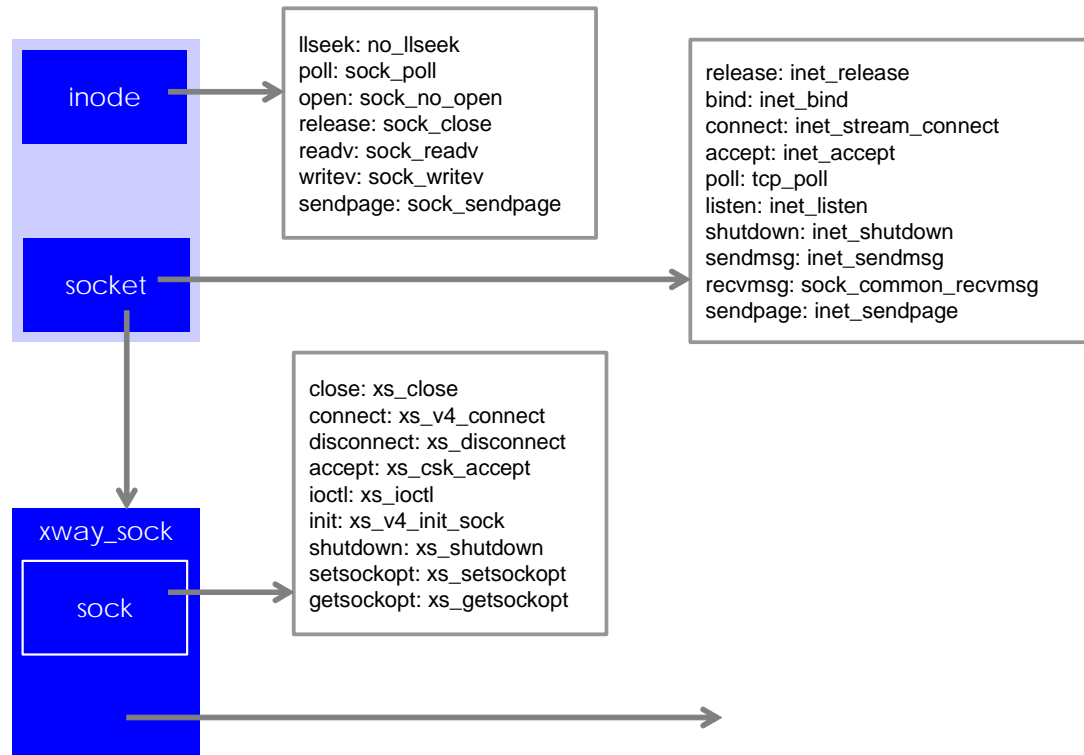
XWAY session



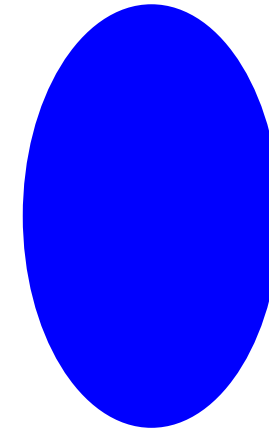
Socket data structure for XWAY



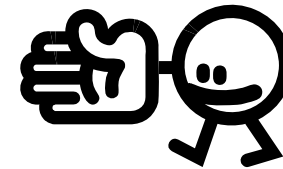
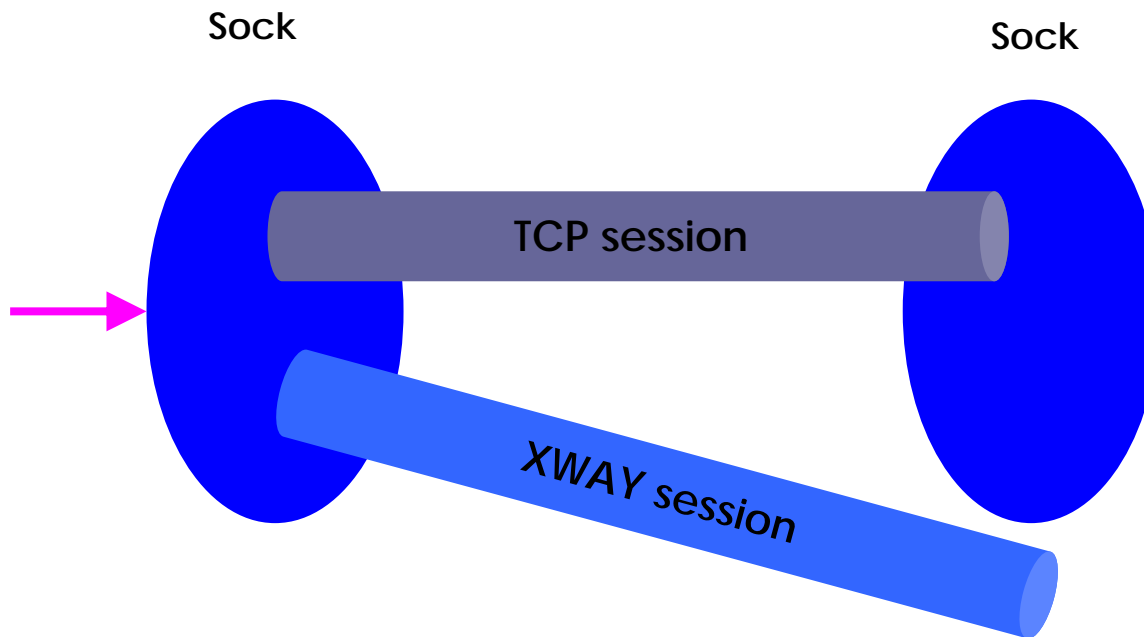
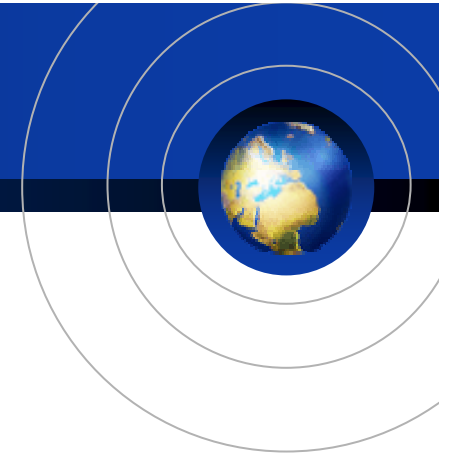
Create



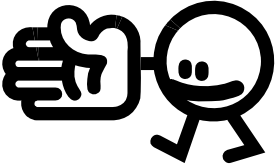
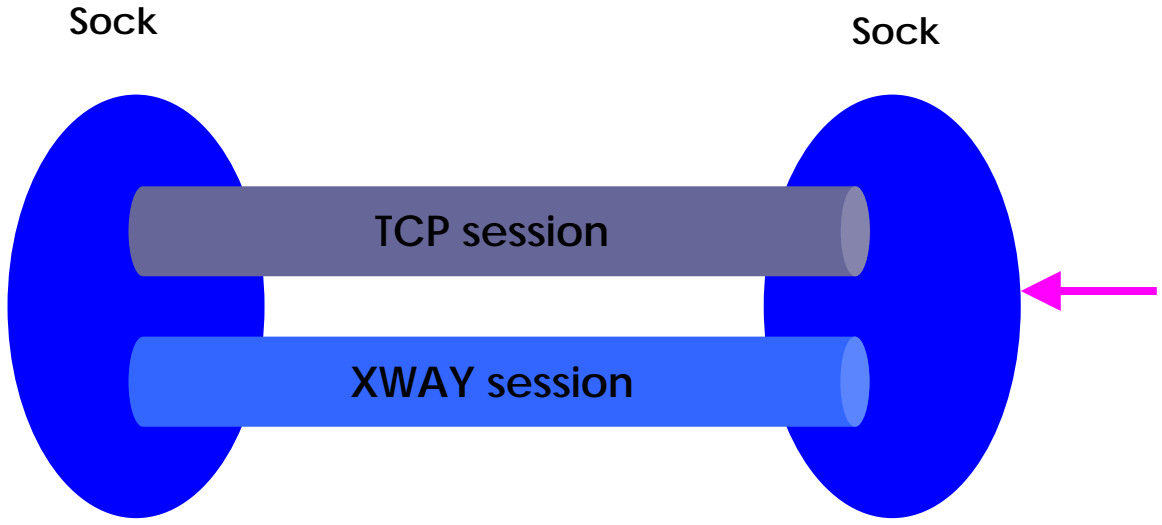
Sock



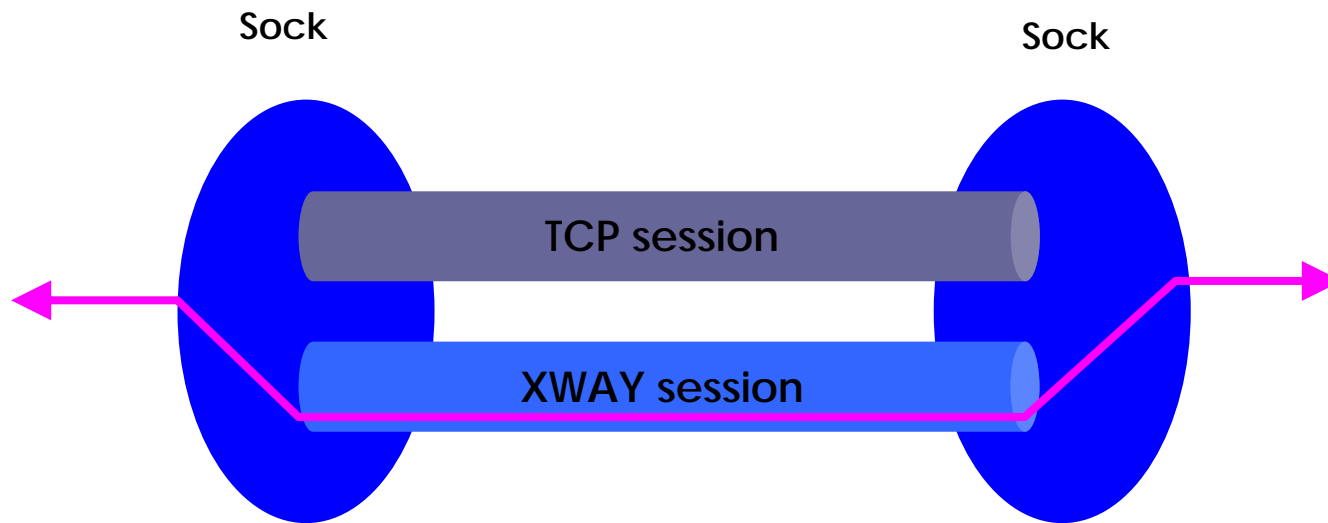
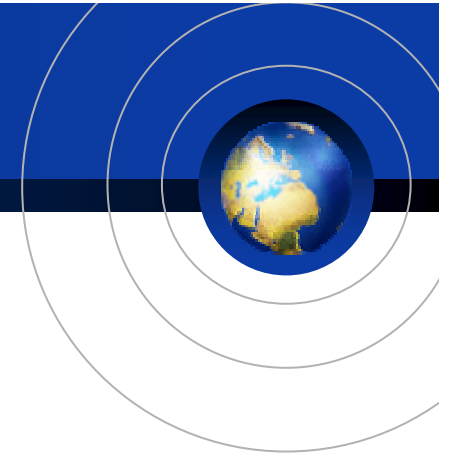
Connect



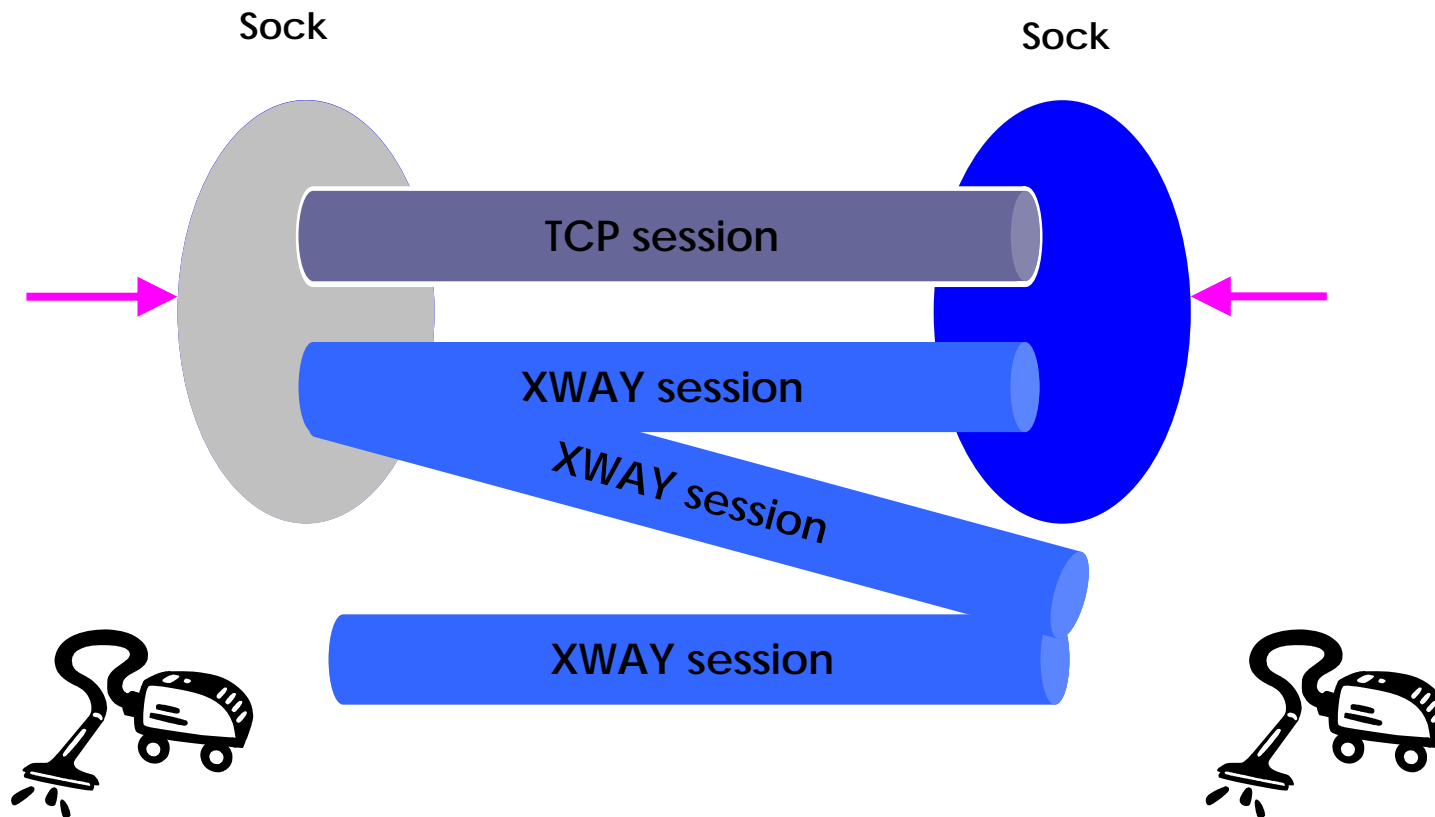
Accept



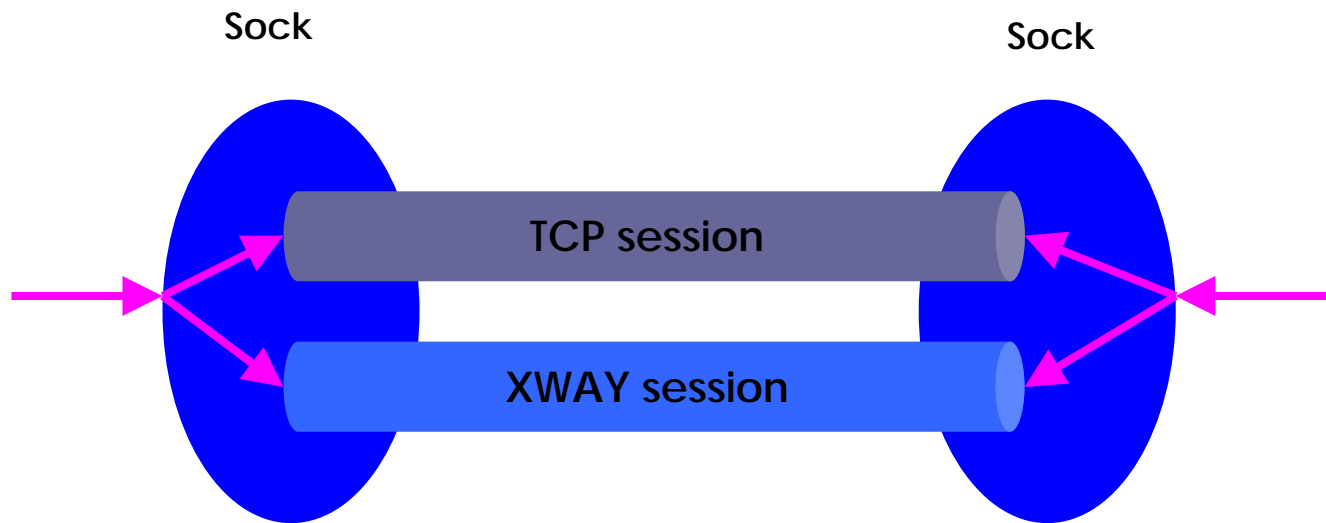
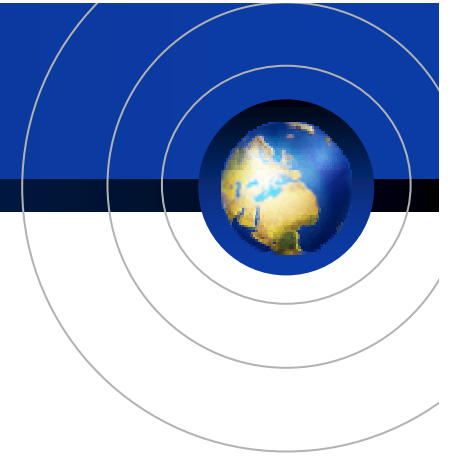
Send/Recv



Close/Shutdown



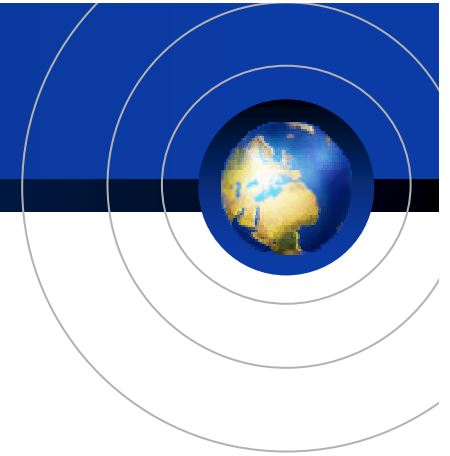
Poll





Evaluation

Environment



- Linux kernel 2.6.16.29
- Xen 3.0.3

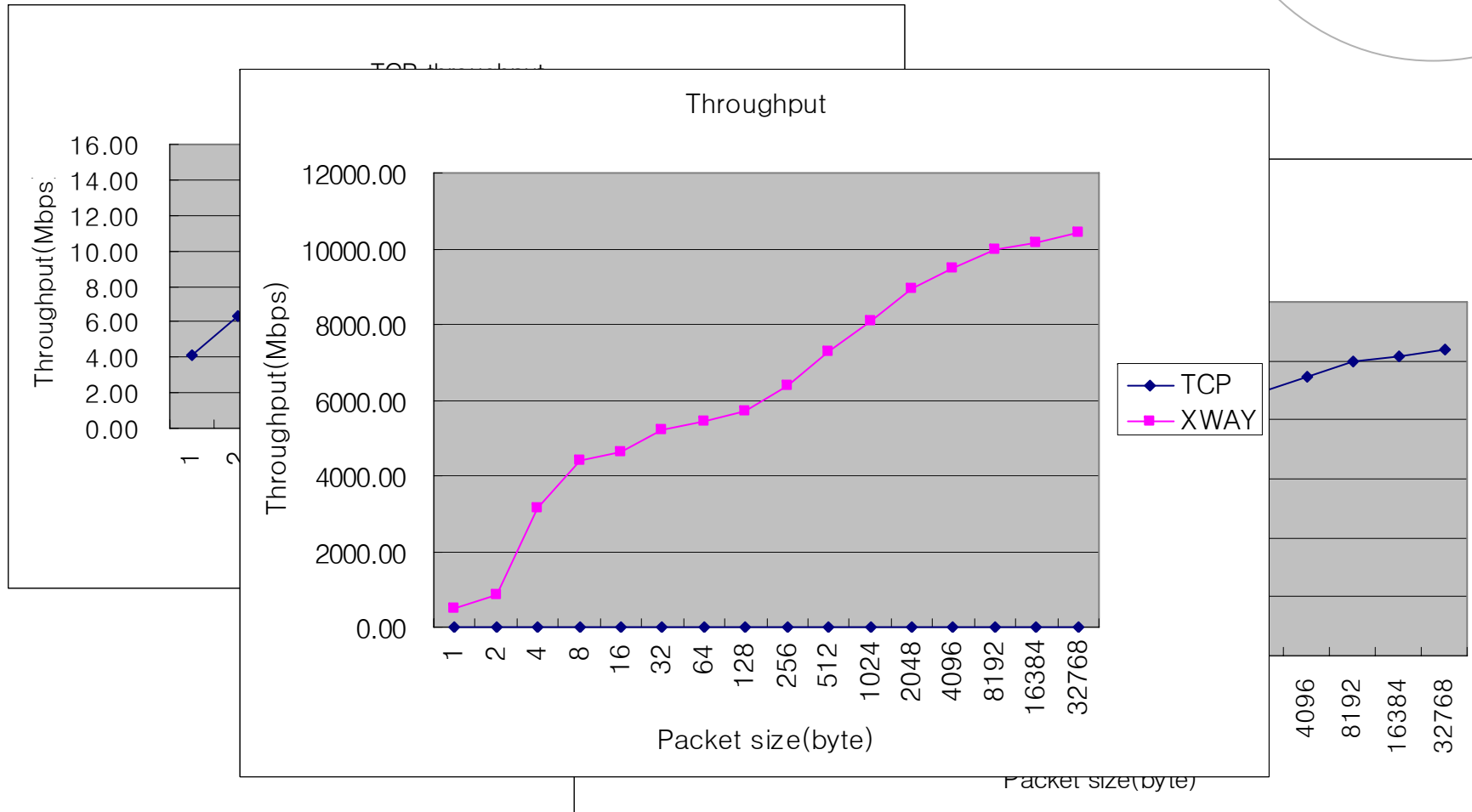
- Pentium 4, 2.53GHz
- 2GB main memory

- Domain U
 - 256 MB main memory
 - 10GB disk

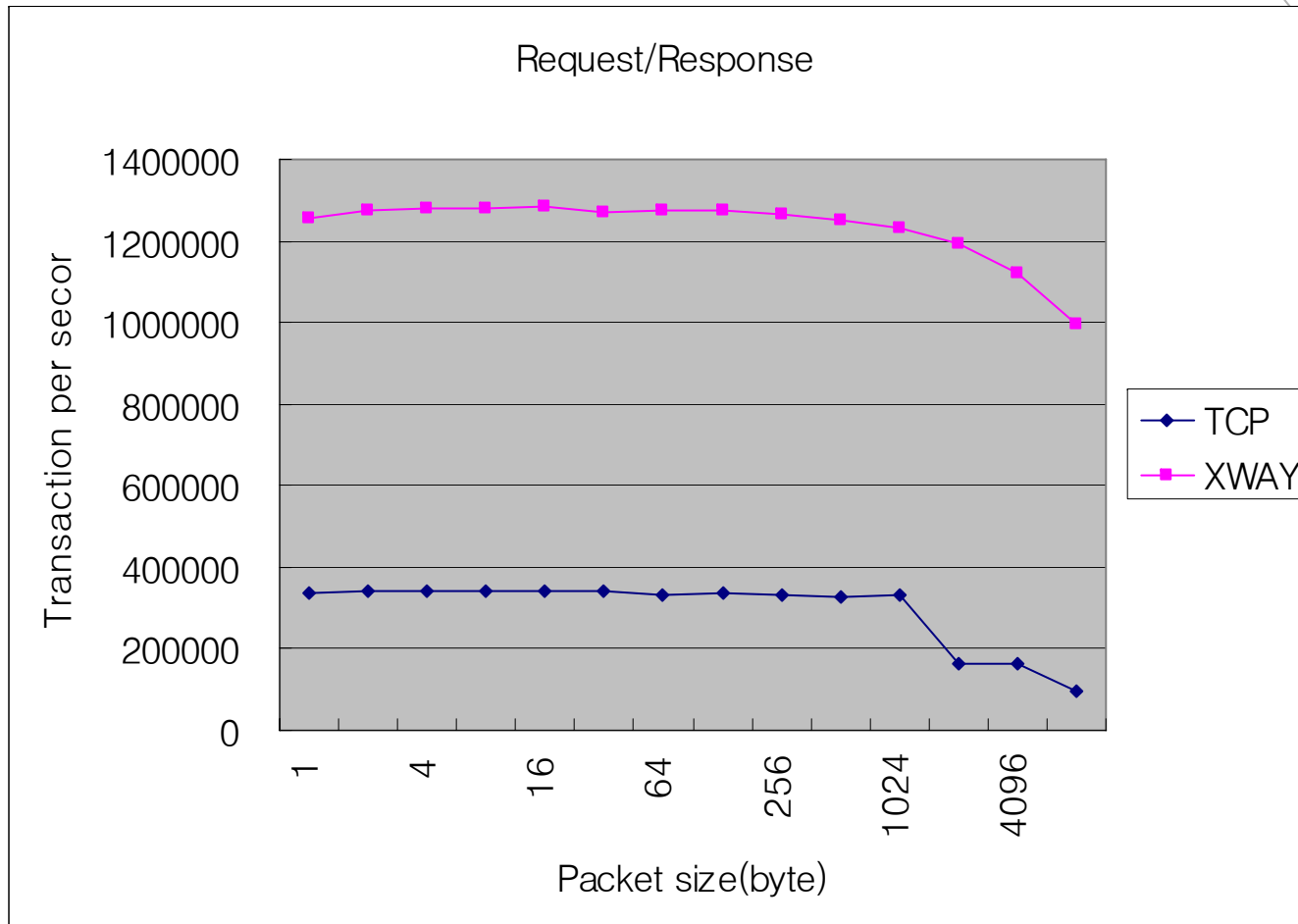
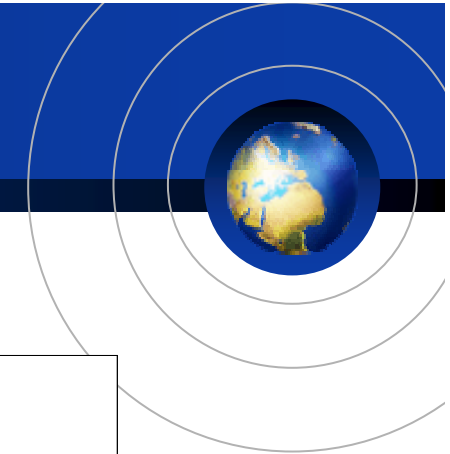
Throughput



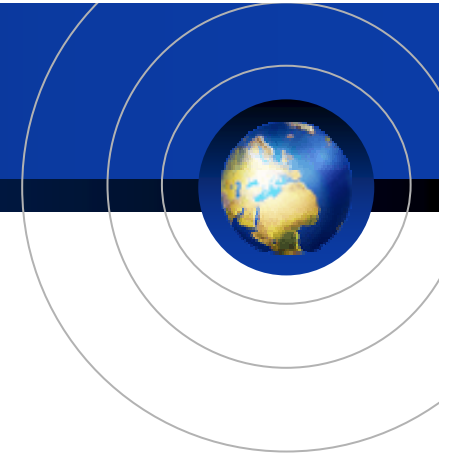
Netperf TCP_STREAM



Request/Response

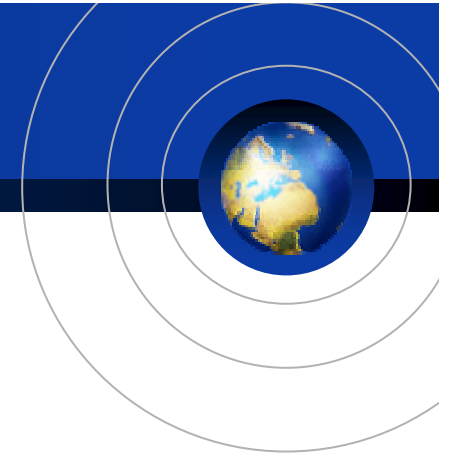


Binary compatibility



- Scp
- Ssh
- Apache – TCP_DEFER_ACCEPT
- Vsftpd – connect, accept, MSG_PEEK
- Proftpd – shutdown, connect, SO_LINGER
- Telnet
- MySQL
- Netperf
- Samba server/client

Remaining work



- Not implemented yet
 - Windows version
 - Migration
 - UDP, ...
- Performance evaluation on high-end machines
 - 32GB main memory NUMA, ...
- Analysis of the evaluation results
- Reducing CPU usage
- ...



Discussion

Comparing to XenSocket



- Common things
 - Using shared memory
 - Bypassing TCP/IP stack
 - High performance
 - Providing socket interface on the shared memory transport

- Uniqueness of Xway
 - Full binary compatibility for socket interface
 - Bidirectional socket communication
 - Dual channel approach