WAN acceleration – an overview

Matthias Hammer, Dipl.-Ing. (FH)

Broadcast Networks and Servers

Institut für Rundfunktechnik (IRT)

Agenda

- Introduction
- Background
- Different types of TCP
- Tests NHK-IRT
- Which way to go?
- Improvement of TCP-Transfers
- Enhancements by “transparent” hardware
- Conclusion

Introduction

Why optimization?

- No doubt: Up-to-date networks are very powerful!
  - Layer 2 -> (NG-) SDH, RPR, DTM, ATM, 10 Gbit-Ethernet, (WIMAX)
    „Unlimited“ bandwidth anywhere!
  - Layer 3 -> IP (regardless if v4 or v6)
    Successful and also powerful...
  - Layer 4 -> Transport Control Protocol (TCP)
    Problem! Invented a long time ago (1981)... and no continuous adaptation!

    „Running standard TCP on Long Fat Networks (LFN) is like driving a Ferrari with bicycle-tires!“

    Why?

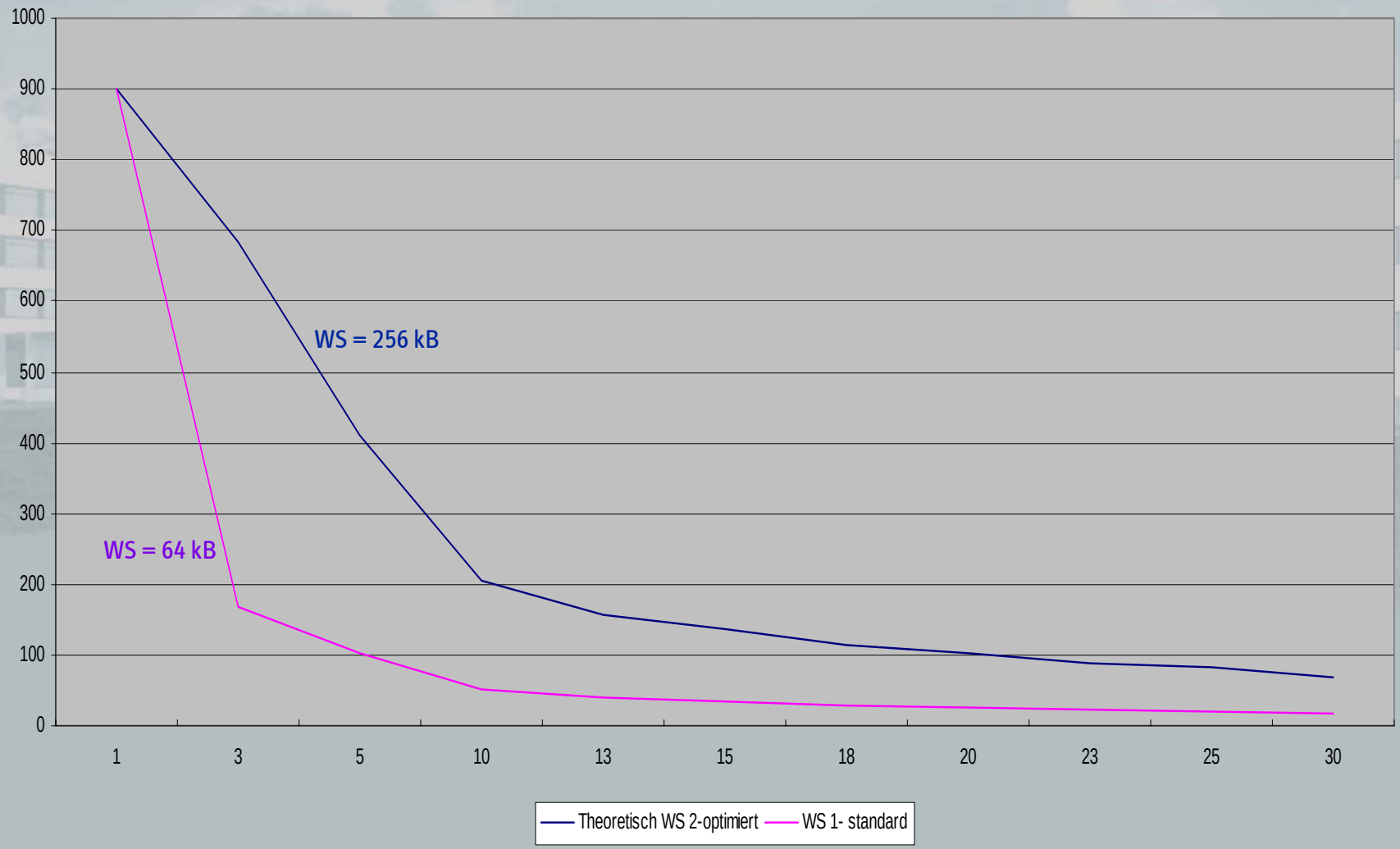
Background (1)

- IP networks entering new areas (e.g. contribution)
- Broadcasters must be defined as “heavy users”
  - Transfer of huge amount of data (24h)
  - Worldwide networking (contribution and partly also distribution)
  - Time-critical stuff (news)
- The data-delivery has to be reliable
  - parallel data-transfers, capacity bottlenecks, phys. Failures –> resulting in (packet-) loss
- Throughput results from the number of bits „on the fly“ per RTT...

*RTT** Bandwidth = TCP-window-size

or: Max throughput = TCP-window-size / *RTT*

Background (2)



Different types of TCP

Tested by IRT and NHK

- TCP Reno (standard for all OS)
- TCP Reno, tuned (RFC 1323)
- High Speed TCP (HSTCP, RFC 3649)
- Scalable TCP (IST: DataTAG project)
- Fast TCP (no RFC, commercial)
- TCP Vegas (experimental)
- TCP Westwood (experimental)
- BIC TCP (experimental)

(Currently all listed TCP-types, except TCP Reno, are only working/available under Linux)

TCP-Tests NHK – IRT (1)

Contact: Takeuchi.s-js@nhk.or.jp

- 3 TCP-versions tested under Linux (Kernel 2.4)
  - TCP-Reno (current standard, RFC 1323-optimised)
  - High Speed TCP (HSTCP)
  - Scalable TCP
- Throughput against:
  - Delay (5 ms to 30 ms)
  - Bit Error Ratio (BER, up to 10^{-5})
  - Receive buffer sizes (64 kB to 1 MB)
- Lab environment
  - Gigabit-Ethernet network
  - Spirent AdTech network simulator (delay, impairment)
  - Chariot IP measurement tool

TCP-Tests NHK – IRT (2)

Contact: Takeuchi.s-js@nhk.or.jp

Results:

- Most influenced by receive buffer (RB)
  - Standard RB 64 kB – no big differences
  - RB greater than 256 kB – same (higher) throughput for all versions
  - Optimum around 256 kB – performance at least 3 times higher
- Performance by different Bit-Error-Ratio (BER)
  - Reno TCP – only best by high error-rates at standard RB (64 kB)
  - HSTCP - highest throughput (RB 128 to 512 kB and BER < $2*10^{-6}$)
  - Scalable TCP – slightly best by high BER (> $5*10^{-6}$)
- Difficult to implement (Kernel (2.4) compiling, use case...)

Which way to go?

- Changes in software/ applications
  - Not always possible for most of-the-shelf software
  - Transfer via parallel TCP-sessions
  - Changing the transport protocol (e.g. UDP-based solutions)
  - Overlapped I/O, providing asynchronous TCP-send & receive/buffer
- Changes in operating system (OS)
  - Tuning of the OS TCP-Reno stack (Window-scaling, RFC 1323)
  - Adaptation of new TCP-type (e.g. HSTCP, but only Linux...)
- Changes in infrastructure
  - Hardware acceleration by “Proxy”-hardware
    - High effective acceleration of different applications
    - One machine for hole network-segments (branch offices, at least one pair)
    - Should be transparent for un-accelerated protocols
  - Decentralised communication...? (Peer to peer, P2P)

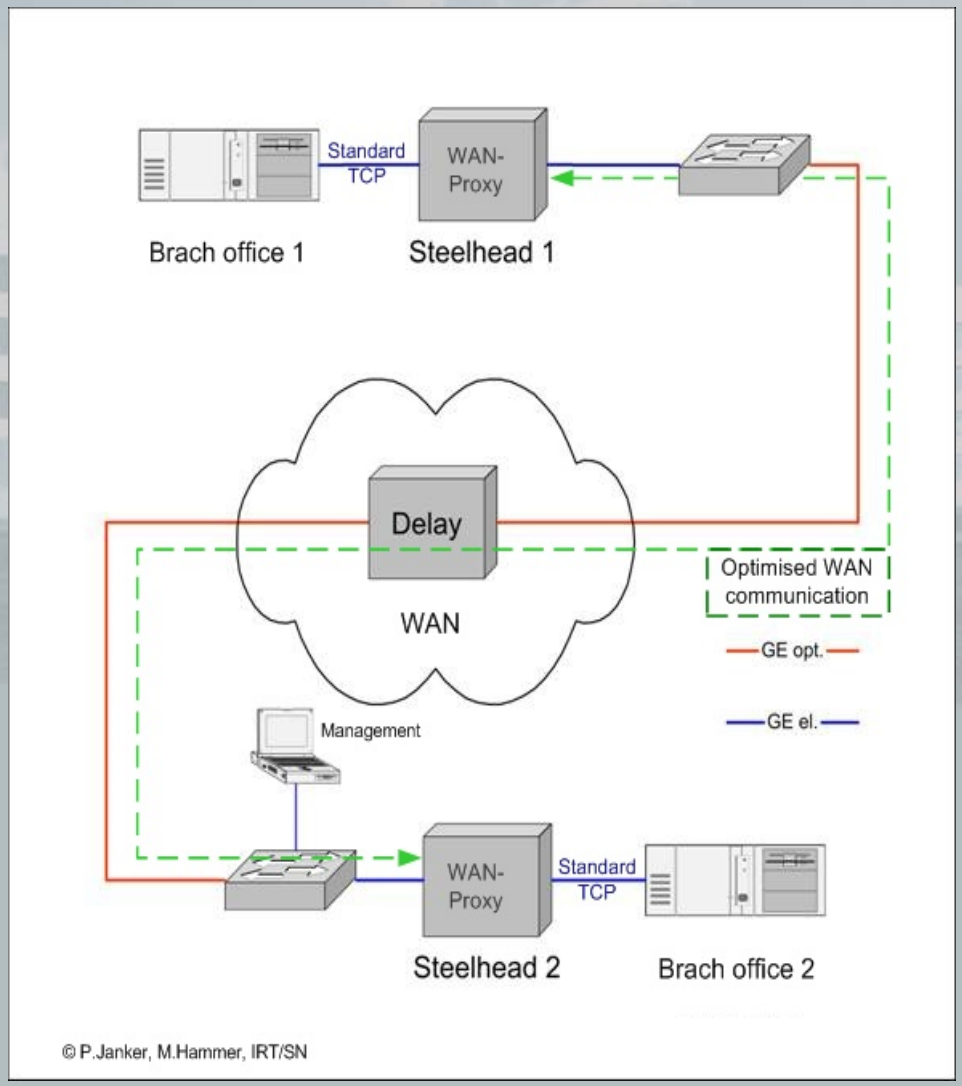
Improvement of (TCP-) Transfers

- Performance enhancement of (single) TCP-connections
  - Changes the protocol stack (concerning all TCP-sessions on the IF)
  - Necessary in all involved systems (but no changes in the applications)
  - Good compatibility against “un-tuned” (standard) systems
  - Lack of flexibility regarding changing network-conditions (new sites or networks)
- Transfer via multiple TCP-sessions (segmentation)
  - Functionality added by applicationsRTT-independent and adaptable transfer (despite using “standard” TCP)
  - Compare e.g. FTP, web-download-manager, Pro-Mpeg MDG
- Change the transport (control) protocol
  - UDP with TCP-like flow control on application level (e.g. UDT (UDP-based Data Transfer))
  - Compatibility - > Firewall difficulties?, different vendors
  - CPU-performance – No hardware-acceleration like TCP-offload engine

Acceleration-hardware (1)

- Integration
  - No changes at the software or OS (transparent inline)
  - Complexity depending on infrastructure
    - Every “island” needs at least one
  - Centralised administration and controlling
- Functionality (depending on vendor)
  - RTT-”independence” (powerful, maybe proprietary communication)
    - TCP-session are terminated locally - different protocol over the WAN
    - Additional caching and data-compression (MPEG...?)
  - Secure encryption of complete communication
  - Firewalls, troubleshooting, reliability ?
- Tested by IRT: Riverbed-Steelhead, T-Systems-TCP-Proxy

Acceleration-hardware (2)



- Example for possible integration of acceleration-HW
- Measurements in the IRT-Lab
  - Verifying the enhancement
  - Delay and bit-error
  - Transparent for
    - Clients
    - not-TCP-traffic
    - Firewalls
- Results will be published by the EBU FT-AVC Group soon

Conclusion

- If you have to work with high data rates in (IP-) WAN you are “forced” to find a (your) solution…
- Depending on your situation you have two main choices to tune:
  - *Direct* via the software/applications/OS (if possible…)
  - *In-direct* via your infrastructure, inserted hardware
- Promising progress in ongoing development of new technologies
  - New protocols (xy-TCP, UDT …)
  - Infrastructure – e.g. decentralised organisation (P2P) may be useful
- In the end there is no “one and only” solution – every user has to decide depending on his use case, his need for guaranteed future and the costs

Thank you for your attention !

Matthias Hammer

Dipl.-Ing. (FH)
Broadcast Networks and Servers

Institut fuer Rundfunktechnik
Floriansmuehlstraße 60
80939 Muenchen
Germany

Fon +49-(0)89-32399-446
Fax +49-(0)89-32399-351
E-Mail: Hammer@irt.de