

TCP/IP-stack extensies

Pieter de Boer

Martijn P. Rijkeboer

6 juli 2005

Onderzoeksproject TCP/IP-stack extensies

Probleemstelling

- Toename bandbreedte
- TCP/IP onveranderd
- Achterblijvende doorvoersnelheid

Doel

- Extensies en configuratie-opties TCP
- Testen extensies en configuratie-opties
- Uitkomsten ten behoeve van dienst “Kanon”

Problemen met TCP

- Window size limiet

$$\frac{1000 \text{ ms}}{5 \text{ ms}} \times 64 \text{ Kbytes} = 100 \text{ Mbits/s}$$

- Herstellen na dataverlies
- Round-trip time meten

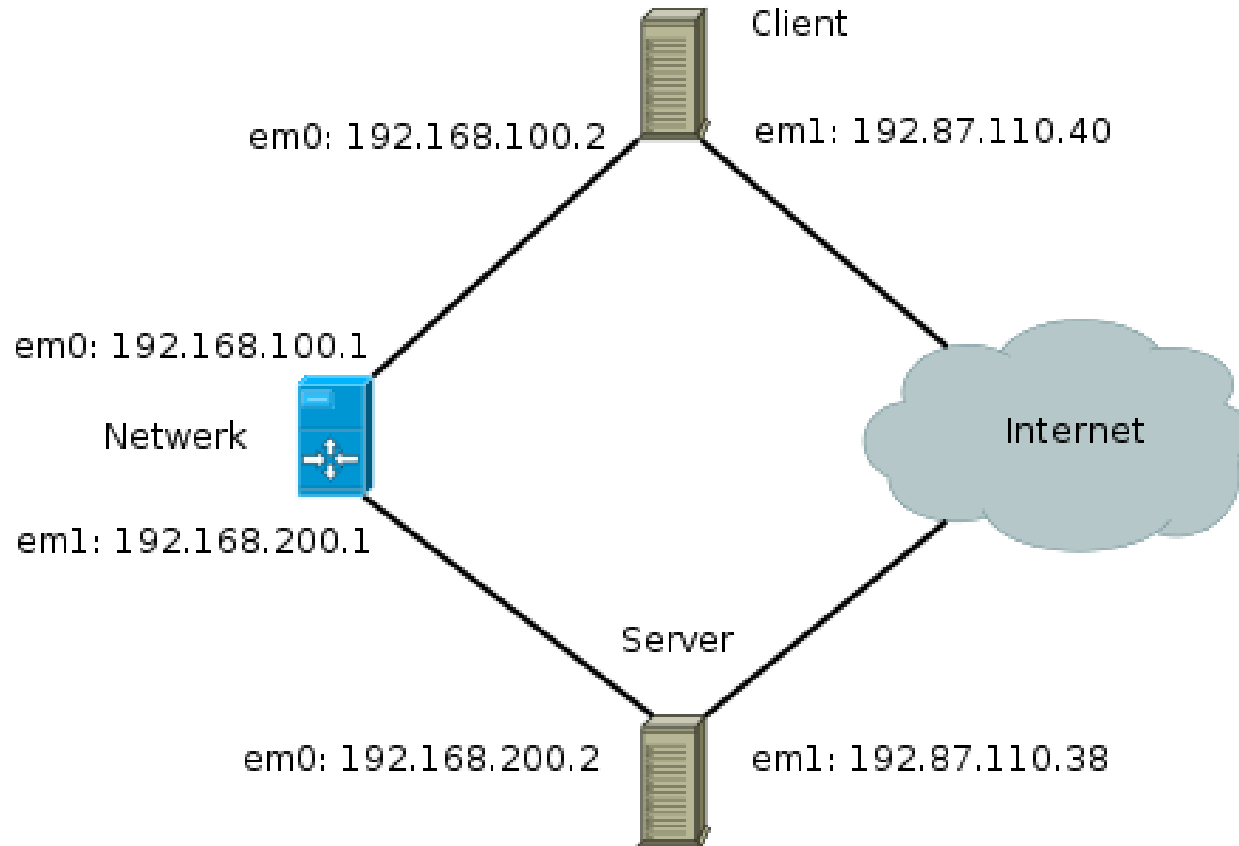
Extensies - 1

- Send en receive space
- TCP extensies for high performance
 - TCP window scale option
 - Round-trip time measurement
 - Protect against wrapped sequence numbers
- TCP selective acknowledgement options

Extensies - 2

- Increasing TCP's initial window
- NewReno modification

Opstelling



Hardware

Dell PowerEdge 1850:

- CPU: Dual 3.0 / 2.8 GHz Xeon met Hyperthreading
- Geheugen: 2 / 1 GB ECC DDR2-SDRAM (400MHz)
- Hdd: Seagate Cheetah 10K RPM 36,7 GB SCSI
- Netwerk: twee 1Gbit Intel netwerkkaarten

Software

- FreeBSD 5.4-RELEASE
- Dummynet
- Netem (kernel 2.6.11.9)
- Iperf 2.0.1

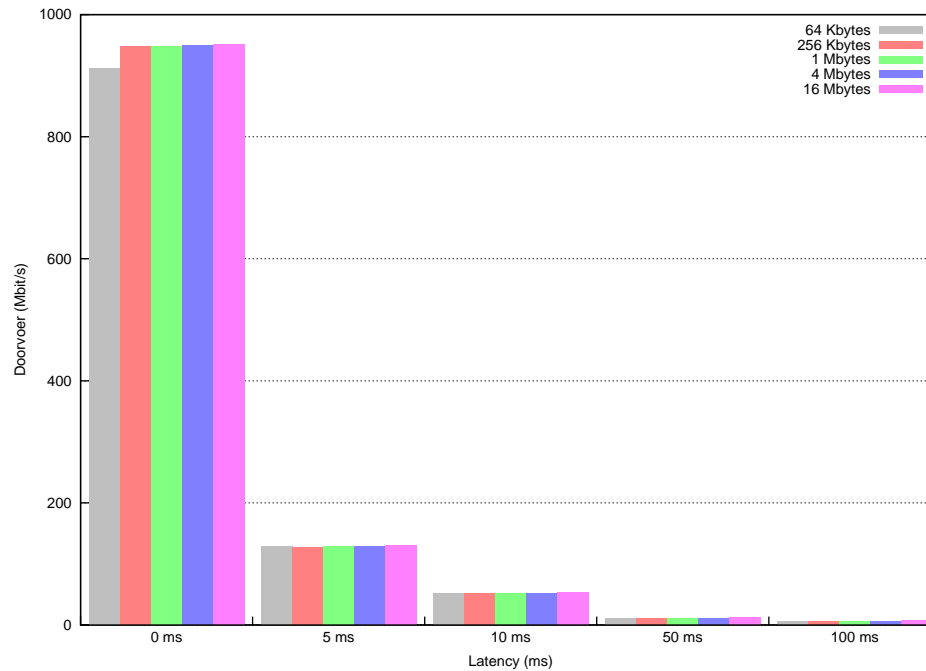
Resultaten - 1

Resultaten van het project:

- Overzicht extensies/opties
- Testscripts
- Parsingscripts
- Overzicht tuning/debugging
- Testresultaten

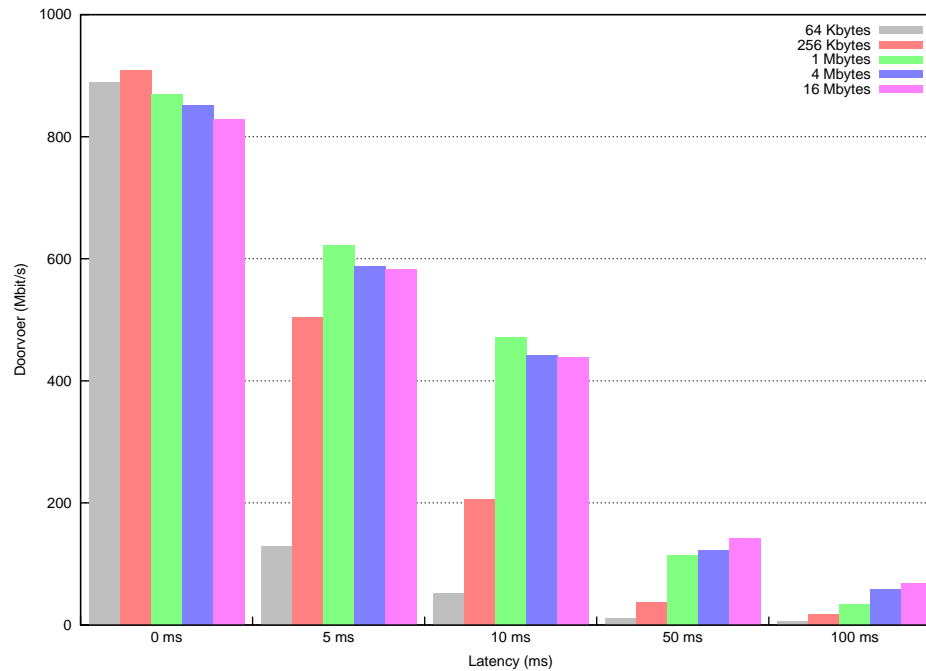
Resultaten - 2

Bandbreedte maal RTT product in de praktijk:



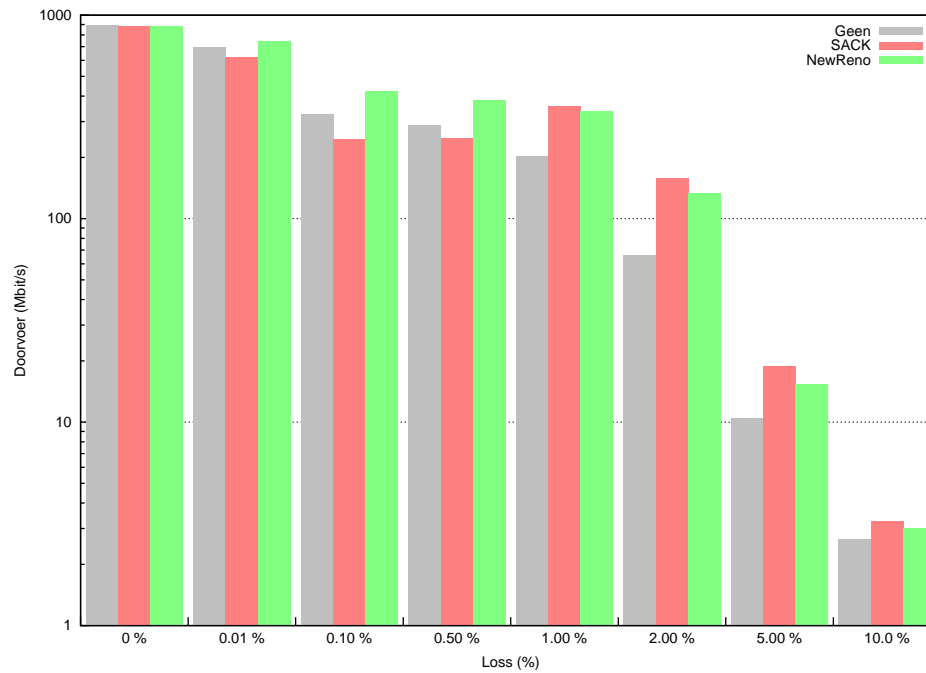
Resultaten - 3

Bandbreedte maal RTT product in de praktijk:



Resultaten - 4

Packet loss:



Problemen - 1

- Tegenvallende performance bij latency
- Prime suspect: Dummynet
- Tuning/debugging om problemen te lokaliseren

Problemen - 2

- Polling
- IP input queue
- Dummynet queue(s)
- Netwerk buffers (mbufs)
- Interface en TCP/IP errors
- HZ

Problemen - 3

Client blijkt een bottleneck te hebben:

- Packet loss
- Extreme latency
- Toch nog CPU-tijd vrij
- Zonder SMP beter
- Meer onderzoek nodig

Conclusie

- Test framework werkt
- Testresultaten tonen nut extensies
- FreeBSD vooralsnog niet geschikt
- Verder onderzoek nodig naar DummyNet ..
- .. en de performanceproblemen
- Test framework uitbreiden met meer metingen