



Presenting your data

- Data visualization for yourself to understand your data, to "listen to your data", to exploit your data, ...
- 2. Presenting it to your audience so that you can explain what you have found in the data, so that you can facilitate their understanding the data, so that they can make use of your data to build upon it, ...



Presenting information with images

"A picture is worth a thousand words."

-- Popular saying

Pictures, graphs, flow charts, UML, state machines, ... can convey an enormous amount of information if used well.

Consider "a wink" at a party

PRESENTING YOUR DATA

II2202, FALL 2015

SLIDE 3



Why use graphical presentations?

- Very compact you can present at lot of data in a small space – in contrast to a table
- To bring out difference and make comparisons
- To help abstract a general (abstract) "picture" (conception) from the data
- Many people are good at seeing **patterns** in visual scenes
- To achieve clarity and objectivity
- To support your text (i.e., to help you tell your story)



A graph is a encoding, when you look at it you need to visually decode it

"When a graph is made, quantitative and categorical information is encoded by a display method. Then the information is visually decoded. This visual perception is a vital link. No matter how clever the choice of the information, and no matter how technologically impressive the encoding, a visualization fails if the decoding fails. Some display methods lead to efficient, accurate decoding, and others lead to inefficient, inaccurate decoding. It is only through scientific study of visual perception that informed judgments can be made about display methods. The display methods of Elements rest on a foundation of scientific enquiry."

From the preface of William S. Cleveland's "The Elements of Graphing Data" [Cleveland 1989]

PRESENTING YOUR DATA

II2202, FALL 2015

SLIDE 5

[Cleveland 1989] William S. Cleveland, *The elements of graphing data*, 10.[print.] ed. Monterey, Cal: Wadsworth, 1989, ISBN: 978-0-534-03729-1.

[Cleveland 1993] William S. Cleveland, *Visualizing data*. Murray Hill, N.J.: [Summit, N.J: At&T Bell Laboratories; Published by Hobart Press, 1993, ISBN: 978-0-9634884-0-4.

The R code for the figure is the book and the data tables can be found from http://www.stat.purdue.edu/~wsc/visualizing.html



Edward Tufte's books

Examples of how to present information well and even beautifully:

- Beautiful Evidence [Tufte 2006]
- The Visual Display of Quantitative Information [Tufte 2001]
- Visual Explanations: Images and Quantities, Evidence and Narrative [Tufte 1997]

II2202, FALL 2015

• Envisioning Information [Tufte 2008]

http://www.edwardtufte.com/tufte/index

PRESENTING YOUR DATA

SLIDE 6

[Tufte 2008] Edward Rolf Tufte, *Envisioning information*, 12. printing. Cheshire, Conn: Graphics Press, 2008, ISBN: 978-0-9613921-1-6.

[Tufte 1997] Edward R. Tufte, Visual explanations: images and quantities, evidence and narrative. Cheshire, Conn: Graphics Press, 1997, ISBN: 978-0-9613921-2-3.

[Tufte 2006] Edward R. Tufte, *Beautiful evidence*. Cheshire, Conn: Graphics Press, 2006, ISBN: 978-0-9613921-7-8.

[Tufte 2001] Edward R. Tufte, *The visual display of quantitative information*, 2nd ed. Cheshire, Conn: Graphics Press, 2001, ISBN: 978-0-9613921-4-7.



Measuring a FASP file transfer

Inspired by National Center for Biotechnology Information's 'Aspera Transfer Guide' [NCBI 2014]

Downloaded and installed Aspera Connect software from: http://downloads.asperasoft.com/connect2/

Transferred a 1Gbyte test file – while collecting data using: tcpdump –I eth0 –w /tmp/xxxxxxx

PRESENTING YOUR DATA

II2202, FALL 2015

SLIDE 7

[NCBI 2014] National Center for Biotechnology Information, 'Aspera Transfer Guide', National Center for Biotechnology Information, U.S. National Library of Medicine, 16-Apr-2014. [Online]. Available: http://www.ncbi.nlm.nih.gov/books/NBK242625/. [Accessed: 18-Aug-2015]



Start ascp to transfer 1G from test server

maguire@ccsser2:~/.aspera/connect/bin> ls

ascp asperaconnect asperaconnect.bin asperacrypt asunprotect plugins maguire@ccsser2:~/.aspera/connect/bin> env ASPERA_SCP_PASS=demoaspera ./ascp -L- -T -l100m aspera@demo.asperasoft.com:aspera-test-dir-large/1GB /tmp/ LOG Aspera Connect version 3.6.0.106805

LOG Alternate log directory: "-"

LOG Configuration: using v2 configuration file "/home/maguire/.aspera/connect/etc/aspera.conf", user -

LOG Initializing FASP version 3.5.4.103990, license max rate=(unlimited), account no.=1, license no.=1 product=6

LOG Configured symlink actions: create=1, follow=1, follow_wide=0, skip=0

LOG [asssh] remote host-key fingerprint f34dfcda4110604e4ecf53e6e18c6559a38cbb43

LOG [asssh] authentication succeeded, proceeding.

LOG changing session job size from 0 to 2 to match server configuration

-Llog to standard output

-T disable encryption

-I100m maximum bandwidth of request – in this case 100 Mbps

PRESENTING YOUR DATA II2202, FALL 2015



FASP session starts

LOG FASP Session Start uuid=a9063e44-f785-4bca-8e71-3eaa20a64b32 op=recv status=started source=aspera-test-dir-large/1GB (1) dest=/tmp source_prefix=- local=130.237.209.248:42132 peer=198.23.89.123:33001 tcp_port=22 os="Linux 3.7.10-1.45-desktop #1 SMP PREEMPT" ver=3.5.4.103990 lic=6:1:1 peeros="Linux 2.6.32-504.3.3.el6.x86_64 #1 SMP W" peerver=3.5.4.100392 peerlic=10:1:22001 proto_sess=20002 proto_udp=20000 proto_bwmeas=20000 proto_data=20008 LOG FASP Session Params uuid=a9063e44-f785-4bca-8e71-3eaa20a64b32 userid=0 user="aspera" targetrate=100000000 minrate=0 rate_policy=fair cipher=none resume=0 create=0 ovr=1 times=0 precalc=yes mf=0 mf_path=-mf_suffix=.aspera-inprogress partial_file_suffix=- files_encrypt=no files_decrypt=no file_csum=none dgram_sz=0 prepostcmd=- tcp_mode=no rtt_auto=yes cookie="-" vl_proto_ver=1 peer_vl_proto_ver=1 vl_local=0 vlink_remote=0 vl_sess_id=3840 srcbase=- rd_sz=0 wr_sz=0 cluster_num_nodes=1 cluster_node_id=0 range=0-0 keepalive=no test_login=no proxy_ip=- net_rc_alg=alg_delay exclude_older/newer_than=0/0 LOG Measured pMTU: 1492 Bytes, start_brtt: 174 ms
LOG datagram size 1492B, block size 1452B, path MTU 1492B



Intermediate output

LOG Receiver DS Qs ds/n/rq/ao/ap/rd/ru/no/po/pc/do=1/0/0/0/0/0/1/0/0/0 Rs i/o=1/1 mgmt backlog i/s/n =

100% 1000MB 97.3Mb/s 01:26

II2202, FALL 2015

Gerald Q. Maguire Jr. All rights reserved



FASP transfer stops

LOG FASP Transfer Stop uuid=a9063e44-f785-4bca-8e71-3eaa20a64b32 op=recv status=success file="/tmp/1GB" size=1048576000 start_byte=0 rate=96.36Mbps elapsed=87.05s loss=0.00 rexreqs=0 overhead=0 mtime="2014-04-10 19:49"

LOG Receiver DS Qs ds/n/rq/ao/ap/rd/ru/no/po/pc/do=0/0/0/0/0/0/0/0/0/0 Rs i/o=1/1 mgmt backlog i/s/n = Completed: 1024000K bytes transferred in 87 seconds

(95875K bits/sec), in 1 file.

(95875K bits/sec), in 1 file.

LOG FASP Session Stop uidl=a9063e44-f785-4bca-8e71-3eaa20a64b32 op=recv status=success source=aspera-test-dir-large/1GB (1) dest=/tmp source_prefix=- local=130.237.209.248:42132 peer=198.23.89.123:33001 tcp_port=22 os="Linux 3.7.10-1.45-desktop#1 SMP PREEMPT" ver=3.5.4.103990 lic=6:1:1 peeros="Linux 2.6.32-504.3.3.el6.x86_64 #1 SMP W" peerver=3.5.4.100392 peerlic=10:1:22001 proto_sess=20002 proto_udp=20000 proto_bwmeas=20000 proto_data=20008 LOG FASP Session Params uidl=a9063e44-f785-4bca-8e71-3eaa20a64b32 userrid=0 userrid=spera" targetrate=100000000 minrate=0 rate_policy=fair cipher=none resume=0 create=0 ovr=1 times=0 precalc=yes mf=0 mf_path=- mf_suffix=.aspera-inprogress partial_file_suffix= files_encrypt=no file_cdcrypt=no file_csum=none dgram_sz=0 prepostcmd=- tcp_mode=no rtt_auto-yes cookie=""" vl_proto_ver=1 per_vl_proto_ver=1 vl_local=0 vlink=renote=0" vl_sess_id=3840 srcbase=- rd_sz=0 vl_sz=0 cluster_num_nodes=1 cluster_node_id=0 range=0-0 keepalive=no test_login=no proxy_ip=- net_rc_alg=alg_delay exclude_older/newer_than=0/0
LOG FASP Session Statistics [Receiver] id=a9063e44-f785-4bca-8e71-3eaa20a64b32 delay=176ms rex_delay=8ms ooo_delay=8ms_solicited_rex=0.00% rcvd_rex=0.00% rcvd_dups=0.00% ave_xmit_rate_98.63Mbps_effective=100.00% effective_rate=98.63Mbps_(detail:_good_blks_722160_bl_rot_722160_bl_rex_0_dup_blks_0_dup_last_blks_0_drop_blks_xmf_2) (sndr_ct:_sent_112_rcvd_112_lost_0_lost_0.00%) (rcvr_ct:_sent_879_rcvd_877_lost_2_lost_0.23%) (rex_ct:_sent_0_rcvd_0_lost_0_lost_0.00%) (progress: tx_bytes_1048576000_file_bytes_1048576000_tx_time_87494969) rex_xmit_blks_0_xmit_total_722162_rex_xmit_pct_0.00%

Completed: 1024000K bytes transferred in 87 seconds (95875K bits/sec), in 1 file

delay=176ms rex_delay=8ms ave_xmit_rate 98.63Mbps (sndr ctl: sent 112 rcvd 112 lost 0 lost 0.00%) (rcvr ctl: sent 879 rcvd 877 lost 2 lost 0.23%)

II2202, FALL 2015



Final transfer statistics

```
LOG ====== File Transfer statistics ======

LOG ------ Source statistics ------

LOG Source argument scans attempted : 1

LOG - Source path scans attempted : 1

LOG - Source path scans attempted : 0

LOG - Source path scans failed : 0

LOG - Source path scans skipped since irregular : 0

LOG - Source path scans excluded : 0

LOG - Source directory scans completed : 0

LOG - Source directory creates attempted : 0

LOG - Source directory creates attempted : 0

LOG - Source directory created or existed : 0

LOG - Source file transfers attempted : 1

LOG - Source file transfers failed : 0

LOG - Source file transfers failed : 0

LOG - Source file transfers skipped : 0

LOG - Source file transfers skipped : 0

LOG - Source bytes transferred : 1048576000

LOG ======= end File Transfer statistics ======
```



Wireshark: UDP conversion

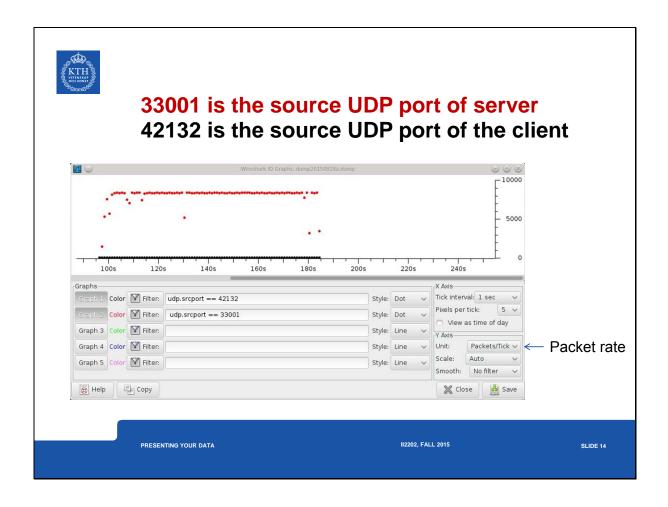
client server

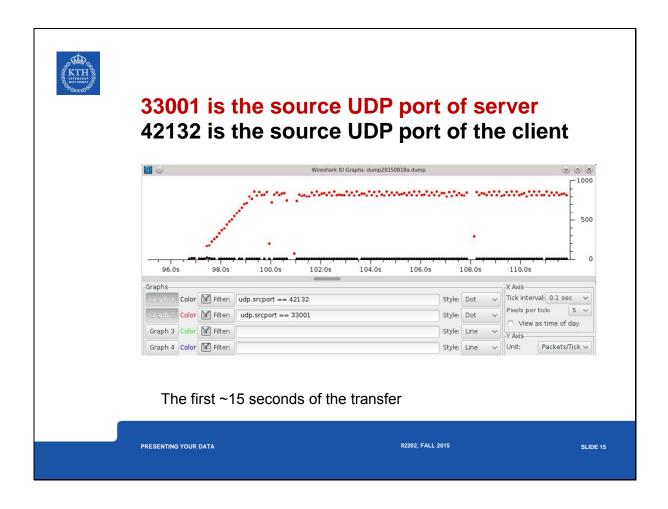
A: 130.237.209.248:42132 ↔ B: 198.23.89.123:33001

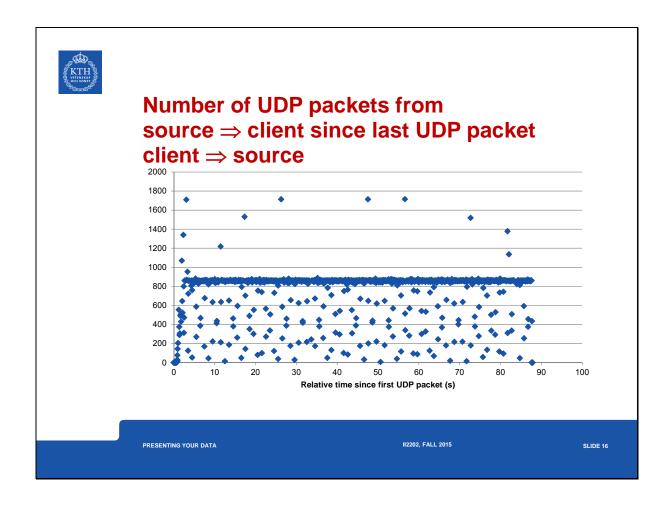
Packets: 703,728 Bytes: 1,058,302,232

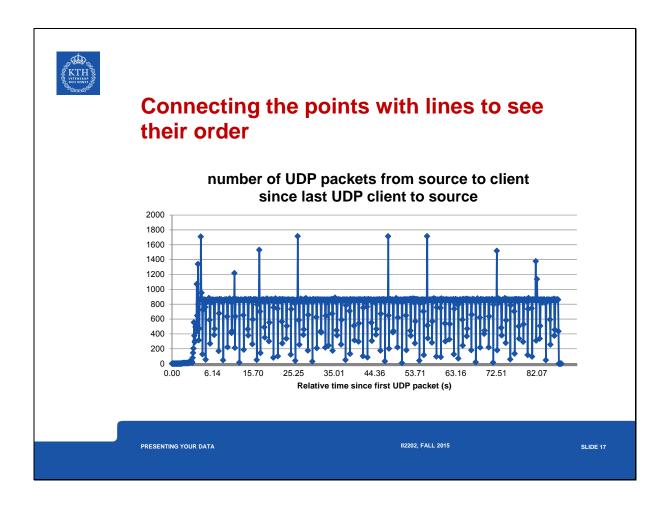
	Packets	Bytes	bps
A→B	961	88518	8,058.85
А←В	702,767	1,058,213,714	96,341,791.88

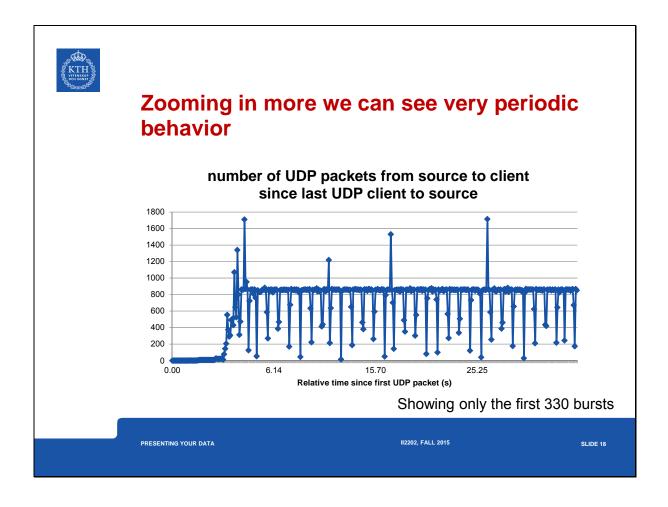
File size = 1,048,576,000 bytes

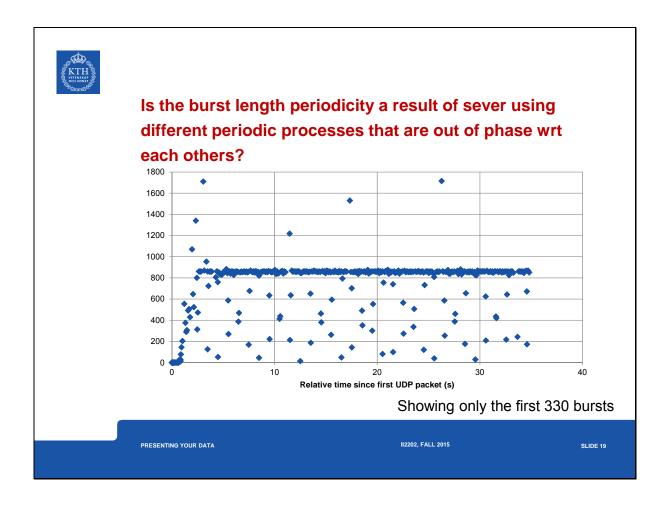


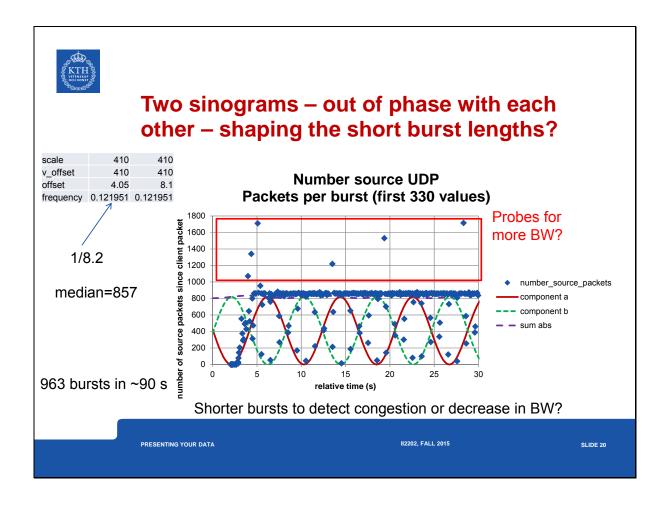




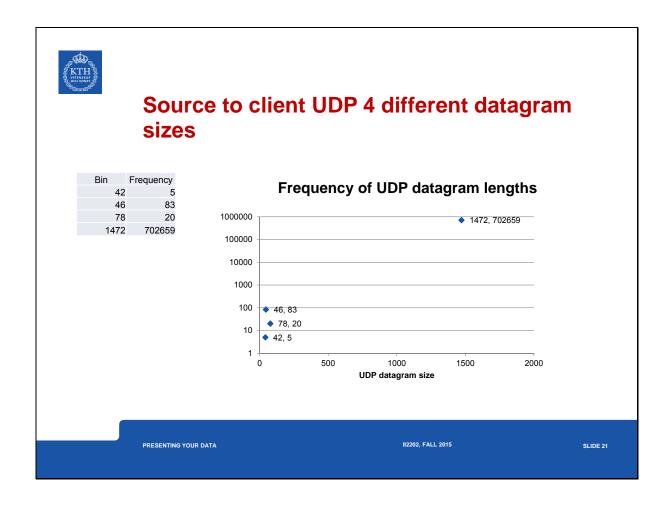


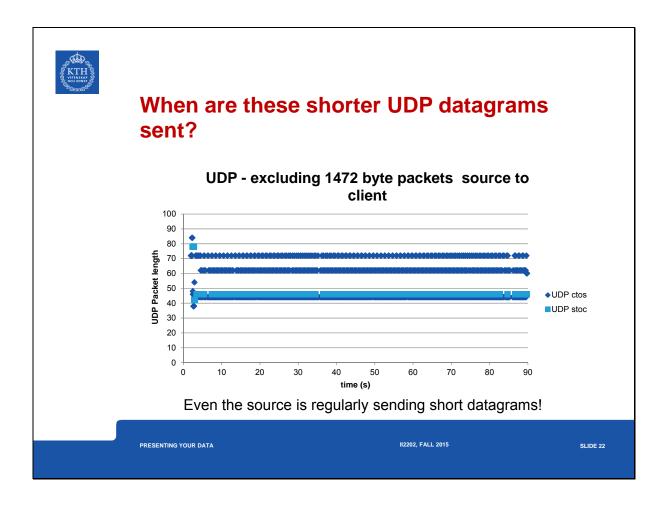


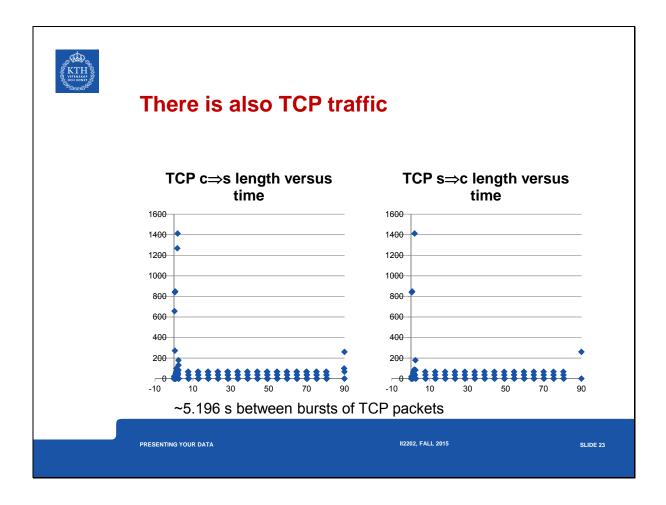


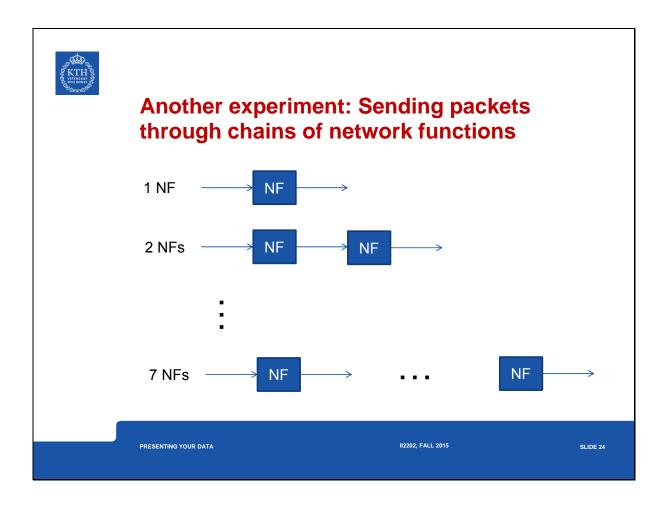


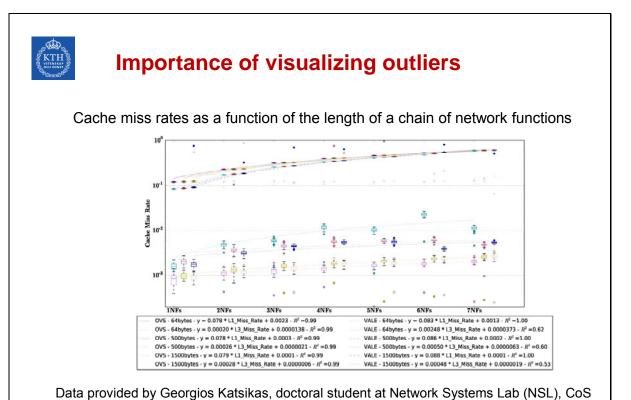
Slide 21

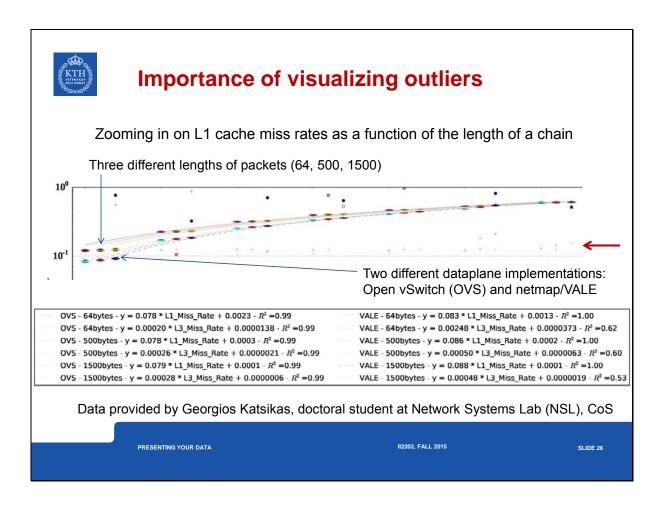


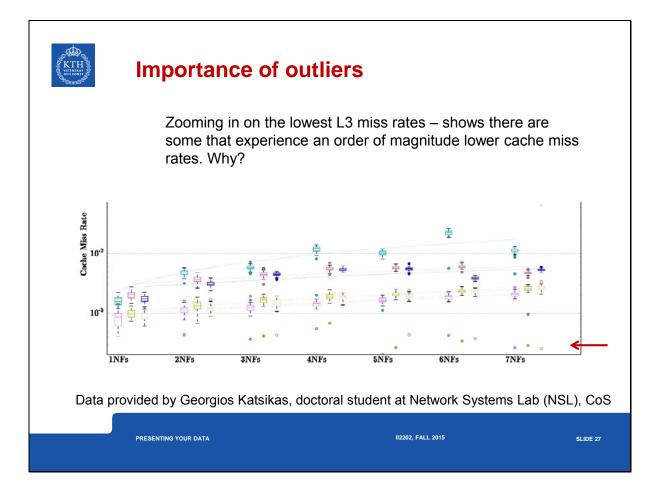














Lots of sources for more information

There are lots of different ways of presenting data graphically, see for example:

- Ray Lyons, 'Best Practices in Graphical Data Presentation', [Lyons 2010] http://libraryassessment.org/bm~doc/workshop-lyons-ray.pdf
- Dona M. Wong, The Wall Street journal guide to information graphics: the dos and don'ts of presenting data, facts, and figures [Wong 2010] (Edward R. Tufte was her thesis advisor)

PRESENTING YOUR DATA II2202, FALL 2015 SLIDE 2

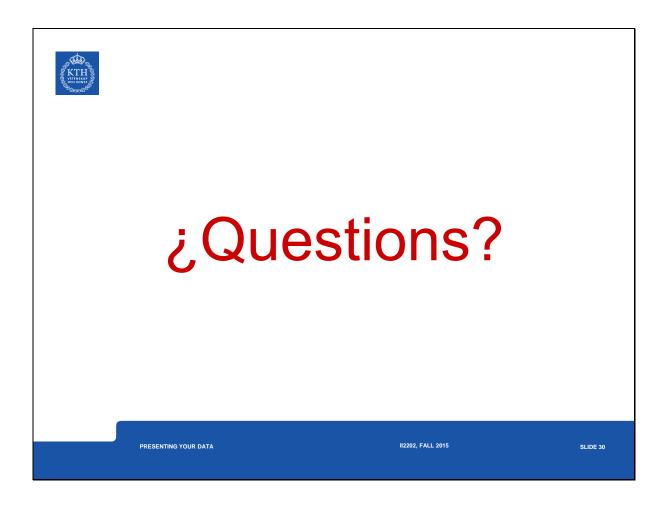
[Lyons 2010] Ray Lyons, 'Best Practices in Graphical Data Presentation', Baltimore, MD, USA, 25-Oct-2010 [Online]. Available: http://libraryassessment.org/bm~doc/workshop_lyons_ray.pdf. [Accessed: 18-Aug-2015]

[Wong 2010] Dona M. Wong, *The Wall Street journal guide to information graphics:* the dos and don'ts of presenting data, facts, and figures, 1st ed. New York: W.W. Norton & Co, 2010, ISBN: 978-0-393-07295-2.



References

[Cleveland 1989] William S. Cleveland, *The elements of graphing data*, 10.[print.] ed. Monterey, Cal: Wadsworth, 1989, ISBN: 978-0-534-03729-1. William S. Cleveland, *Visualizing data*. Murray Hill, N.J.: [Summit, N.J: At&T Bell Laboratories; Published by Hobart Press, 1993, ISBN: 978-0-9634884-0-4. [Cleveland 1993] Ray Lyons, 'Best Practices in Graphical Data Presentation', Baltimore, MD, USA, 25-Oct-[Lyons 2010] 2010 [Online]. Available: http://libraryassessment.org/bm~doc/workshop_lyons_ray.pdf. [Accessed: 18-Aug-2015] Edward Rolf Tufte, *Envisioning information*, 12. printing. Cheshire, Conn.: Graphics Press, 2008, ISBN: 978-0-9613921-1-6. [Tufte 2008] Edward R. Tufte, *Visual explanations: images and quantities, evidence and narrative*. Cheshire, Conn. Graphics Press, 1997, ISBN: 978-0-9613921-2-3. [Tufte 1997] [Tufte 2006] Edward R. Tufte, Beautiful evidence. Cheshire, Conn: Graphics Press, 2006, ISBN: 978-0-9613921-7-8. Edward R. Tufte, *The visual display of quantitative information*, 2nd ed. Cheshire, Conn: Graphics Press, 2001, ISBN: 978-0-9613921-4-7. [Tufte 2001] Dona M. Wong, The Wall Street journal guide to information graphics: the dos and don'ts of presenting data, facts, and figures, 1st ed. New York: W.W. Norton & Co, 2010, ISBN: 978-0-393-07295-2. [Wong 2010]



Gerald Q. Maguire Jr. All rights reserved