



IN5060

Performance in distributed systems

autumn course



UiO • **University of Oslo**

What is performance?



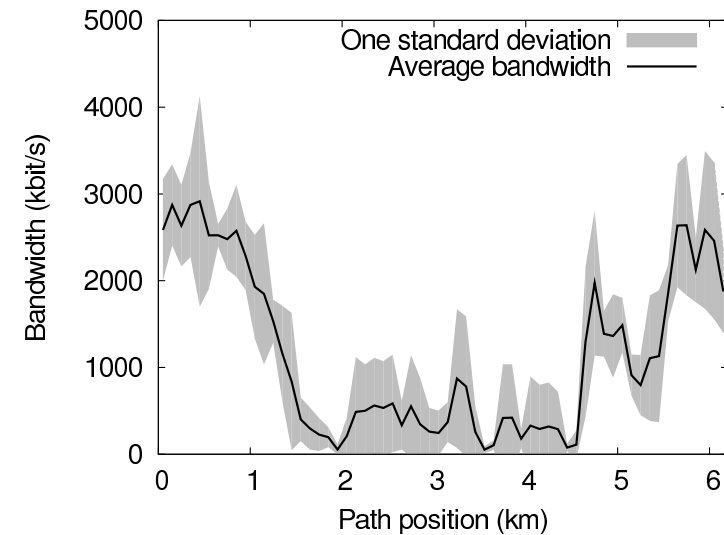
Stage performance

World Opera Production – Dec 2011 @ Tromsø



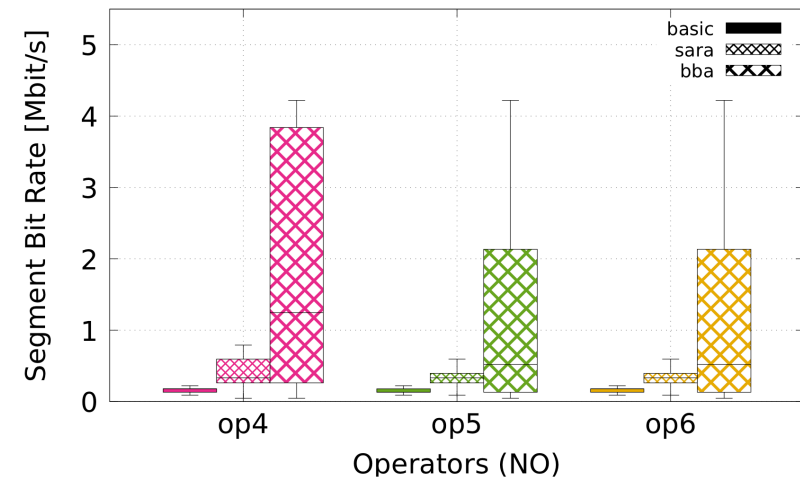
Stage performance

Third Life Project@WUK – Oct 2015 @ Vienna



Download performance by position

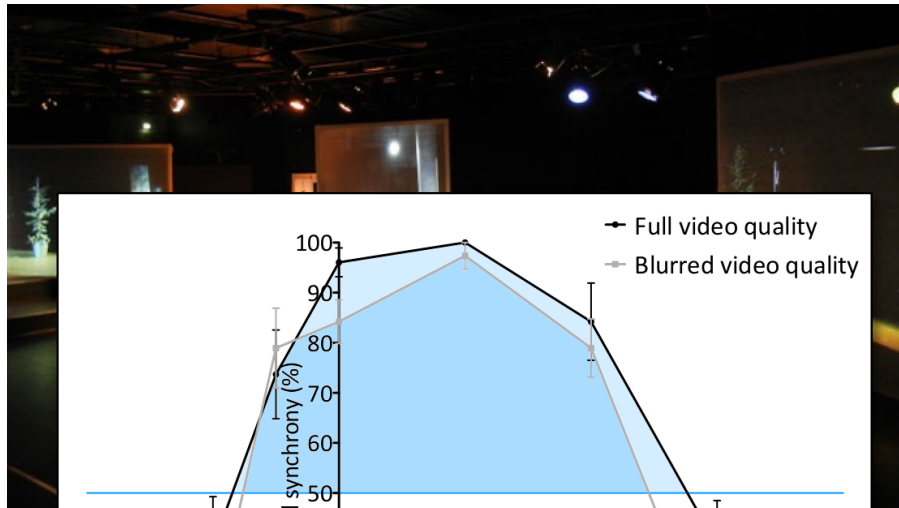
HTTP Adaptive Streaming measured on Bygdøy Ferry, 2011



Download performance by operator & algorithm

HTTP Adaptive Streaming, MONROE nodes, 2018

What is performance?

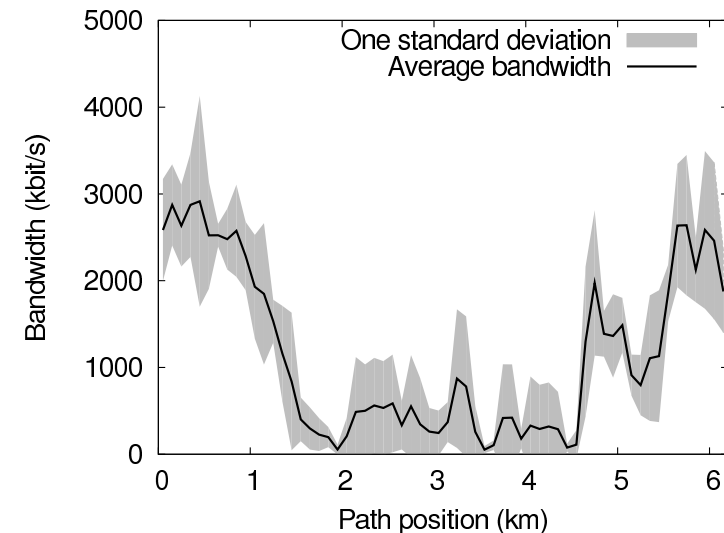


Users' perception (Quality of Experience)
Asynchrony between audio and video, 2015



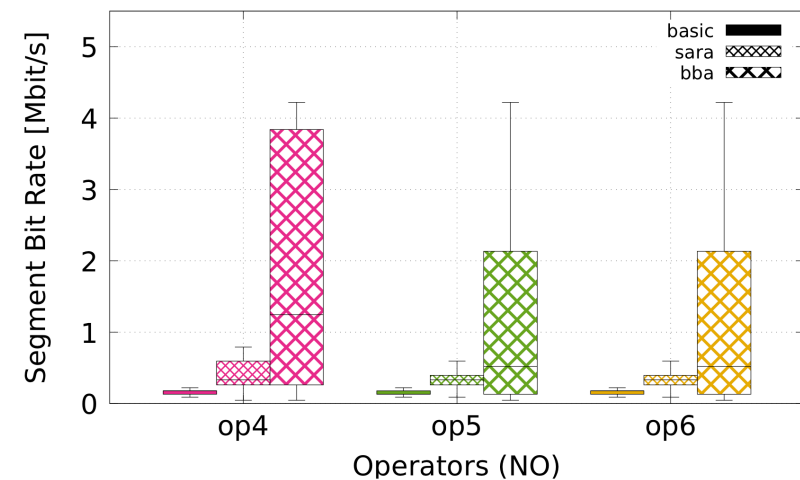
Stage performance

Third Life Project@WUK – Oct 2015 @ Vienna



Download performance by position

HTTP Adaptive Streaming measured on Bygdøy Ferry, 2011

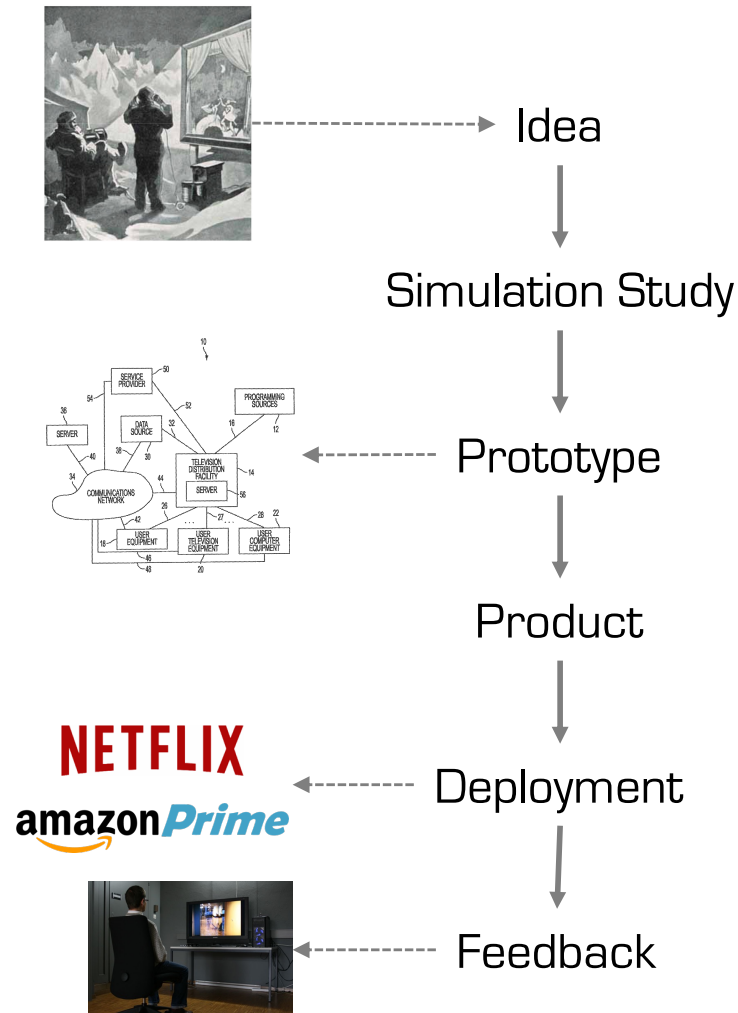


Download performance by operator & algorithm

HTTP Adaptive Streaming, MONROE nodes, 2018

Performance in Distributed Systems

Engineers and researchers solve quantifiable challenges



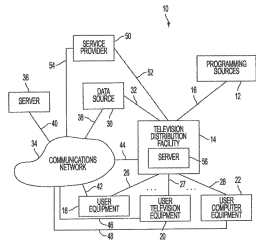
Performance in Distributed Systems

Engineers and researchers solve quantifiable challenges



Idea

Simulation Study

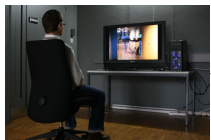


Prototype

Product



Deployment



Feedback

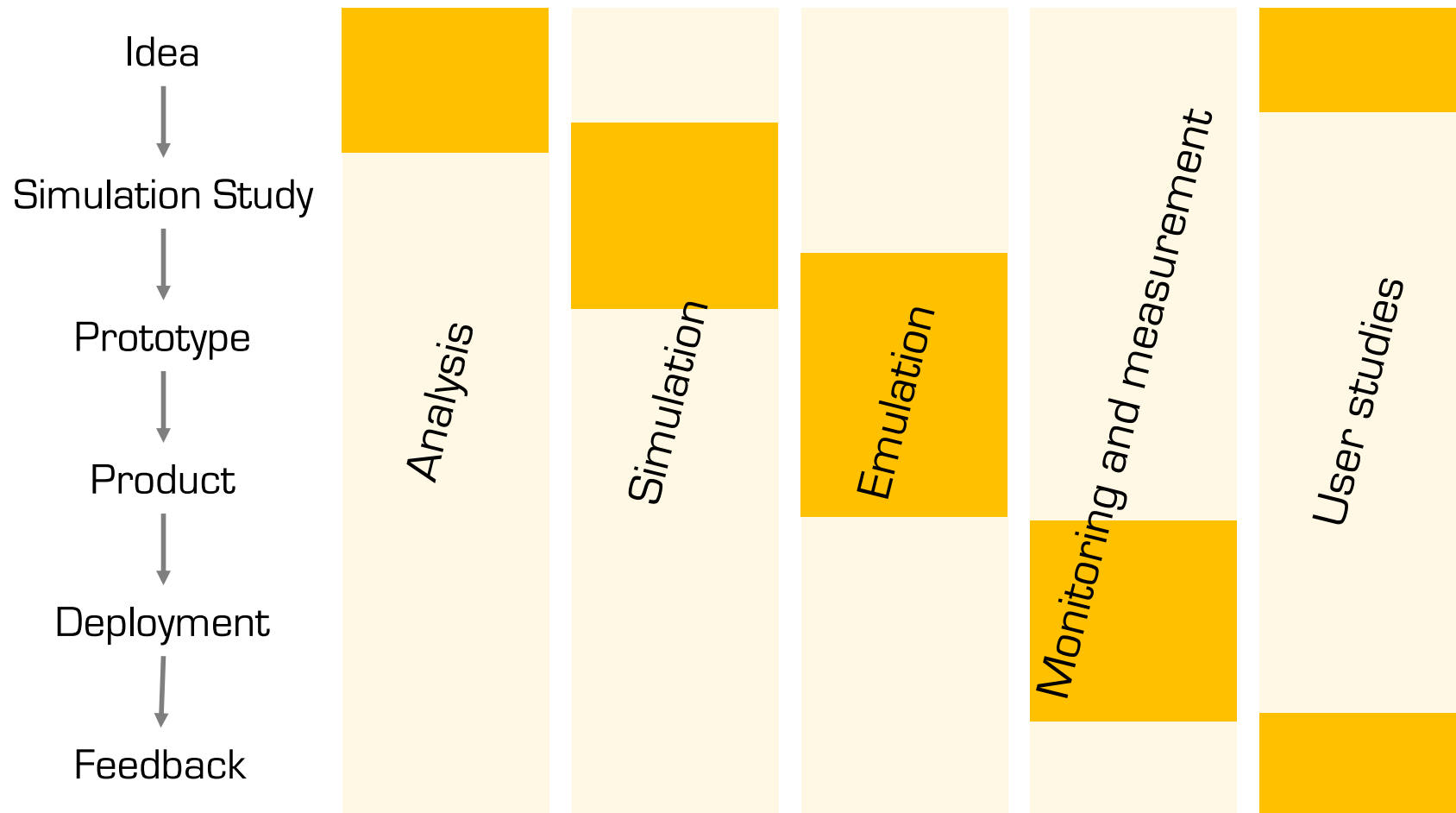
- Each step requires a performance assessment
- argue for feasibility
 - demonstrate practicality
 - study in a context
 - measure in the real world
 - assess value / success



Performance Evaluation

Performance in Distributed Systems

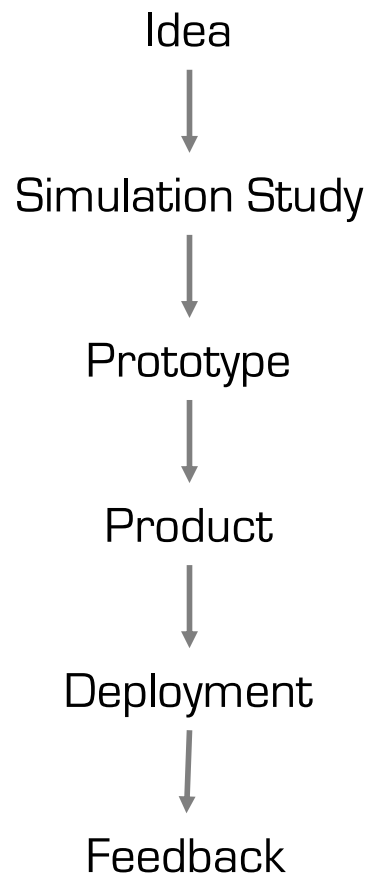
Engineers and researchers solve quantifiable challenges



Performance Evaluation

Performance in Distributed Systems

Engineers and researchers solve quantifiable challenges



IN5060: experience 3 examples



Simulation



Monitoring and measurement



User studies

Performance Evaluation

Performance in Distributed Systems

Designing and conducting studies

- pre-considerations
- avoiding bias
- measurement points and methods
- data reduction
- drawing conclusions

Presentation and reporting

- formulating a message
- selecting relevant factors
- extracting and interpreting statistics
- dimension reduction
- selecting presentation modes

Specific considerations

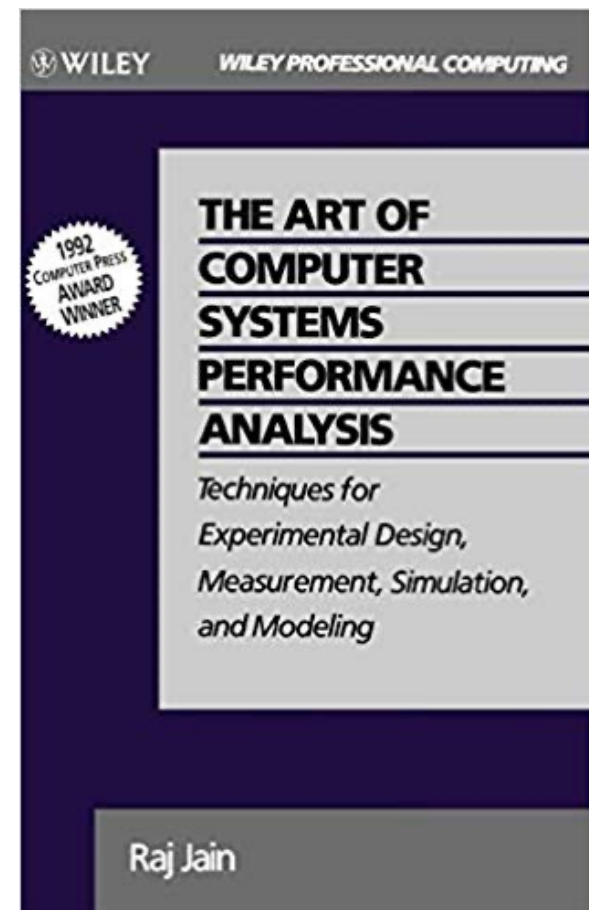
- simulation
- monitoring and measurement
- user studies

Performance in Distributed Systems

- This course is meant to provide you with a taste of the skills needed to become a good system analyst.
- It will provide you with hands-on experience in system evaluation
- It will (to some extent)
 - confront you with the tradeoffs encountered when analysing real systems
 - confront you with the error sources and red herrings encountered when analysing real systems

Performance in Distributed Systems

- The course is based on the book “The Art of Computer Systems Performance Analysis: Techniques for Experimental Design, Measurement, Simulation, and Modeling” by Raj Jain
- Reading the book is not mandatory for the course or even necessary to complete, but if you have a chance to read it in full, **do so!**



System performance analysis

Who is interested in system performance analysis?

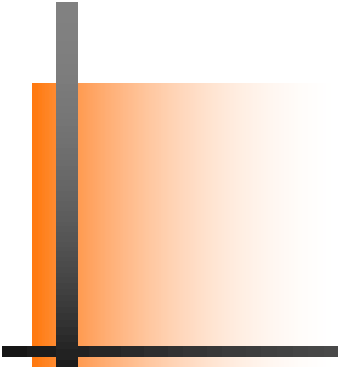
- The HW designer (company) wants to show that their system is The Best and Greatest system of All Time
- A software provider wants to show that their application is superior to the competition
- The researcher wants to publish her papers, and needs to convince the reviewers that their research improves on the state- of-the-art
- The system administrator or capacity planner needs to choose the system that is best suited for their purpose
- The enthusiast who wants to see if the newest rage from *<insert favourite multinational corporation>* is real, or fake news

System performance analysis

- How do they achieve this?
 - By providing a comparison between their own system and “the competition”
 - The results need to be *(or appear)* convincing to the target audience
 - This comparison is made through proper system performance analysis
- The techniques of models, simulations and measurement are all useful for solving performance problems
 - IN5060 will focus on experimental design, simulation, measurement and analysis
 - For modelling try for instance: MAT-INF3100 - Linear Optimisation

Theory and practice

- Theory / models will provide us with candidates for system optimisations
- Deploying them in reality may in many cases lead to unforeseen results
 - Hardware differences
 - Non-deterministic systems
 - Unexpected workloads
- Key techniques needed
 - Mathematical analysis
 - Simulation
 - Emulation
 - Measurement
 - User studies
 - Measurement techniques (monitors)
 - Data analysis (statistics and presentation)
 - Experiment design



Performance in distributed systems

Key skills of performance analysts

Key skills needed – evaluation techniques

To select appropriate evaluation techniques, performance metrics and workloads for a system

- You must choose which metrics to use for the evaluation
- You must choose which workloads would be representative

What metrics would you choose to compare:

- Two solid state disk (SSD) drives?
- Two adaptive video streaming algorithms?
- Two IaaS Clouds?

Key skills needed – measurements

Conduct performance measurements correctly

- You must choose how to apply workloads to the system
- You must choose how to measure (monitor) the system

Which type of monitor (or “probe”, hardware or software) would be suitable for measuring each of the following:

- Number of instructions executed by a processor?
- Context switch overhead on a multi-user system?
- Response time of packets on a network?

Key skills needed – proper statistical techniques

Use proper statistical techniques to compare several alternatives

- Whenever there are non-deterministic elements in a system, there will be variations in the observed results
- You need to choose from the plethora of available statistical methods in order to correctly filter and interpret the results

Which link is better?

File Size	Packets lost on Link A	Packets lost on Link B
1000	5	10
1200	7	3
1300	3	0
50	0	1

Key skills needed – do not measure forever

Design measurement and simulation experiments to provide the most information with the least effort

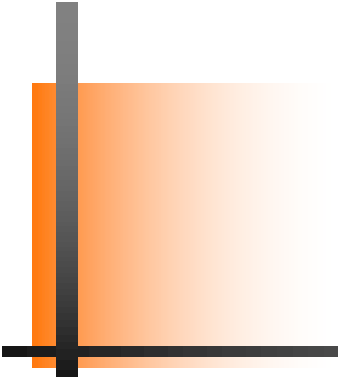
- You must choose the number of parameters to investigate
- You must make sure you can draw statistically viable conclusions

The performance of a system depends on the following factors:

- Garbage Collection Technique used: G1, G2, or none
- Type of workload: editing, computing, or machine learning
- Type of CPU: C1, C2, or C3

How many experiments are needed?

How do you estimate the performance impact of each factor?



Performance in distributed systems

Statistics 101

Why do we need statistics?

1. Noise, noise, noise, noise, noise!

445 446 397 226
388 3445 188 1002
47762 432 54 12
98 345 2245 8839
77492 472 565 999
1 34 882 545 4022
827 572 597 364



$\bar{x} = \dots$

2. Aggregate data into meaningful information.

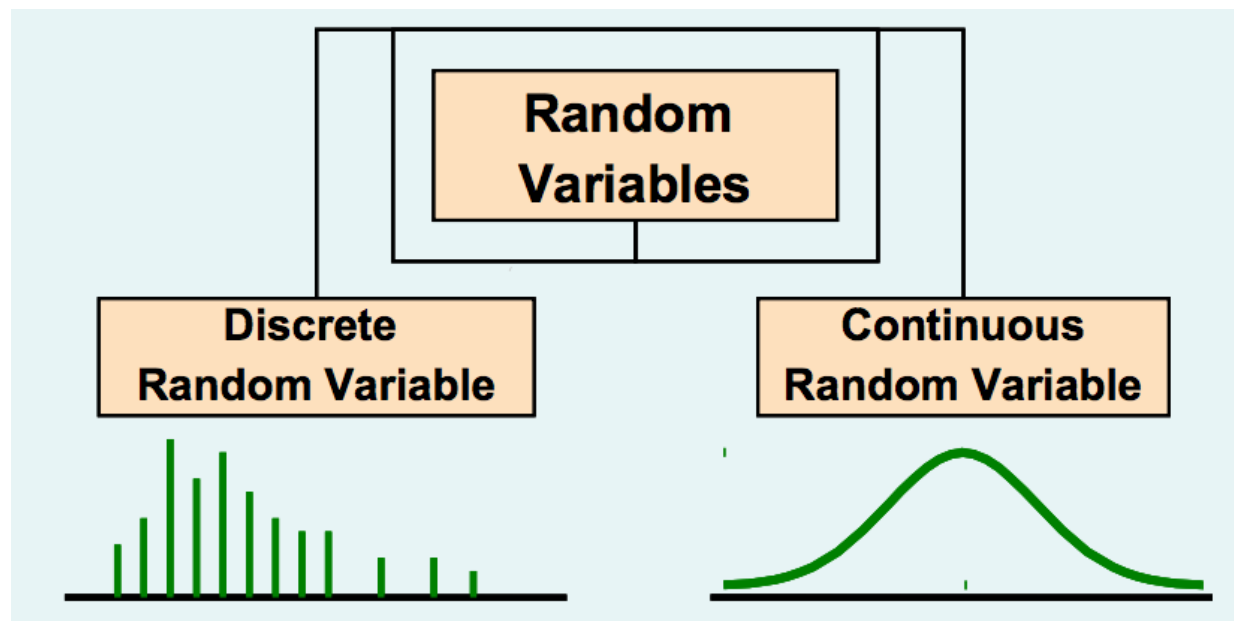
“Impossible things usually don’t happen.”

- Sam Treiman, Princeton University

Statistics helps us quantify “usually.”

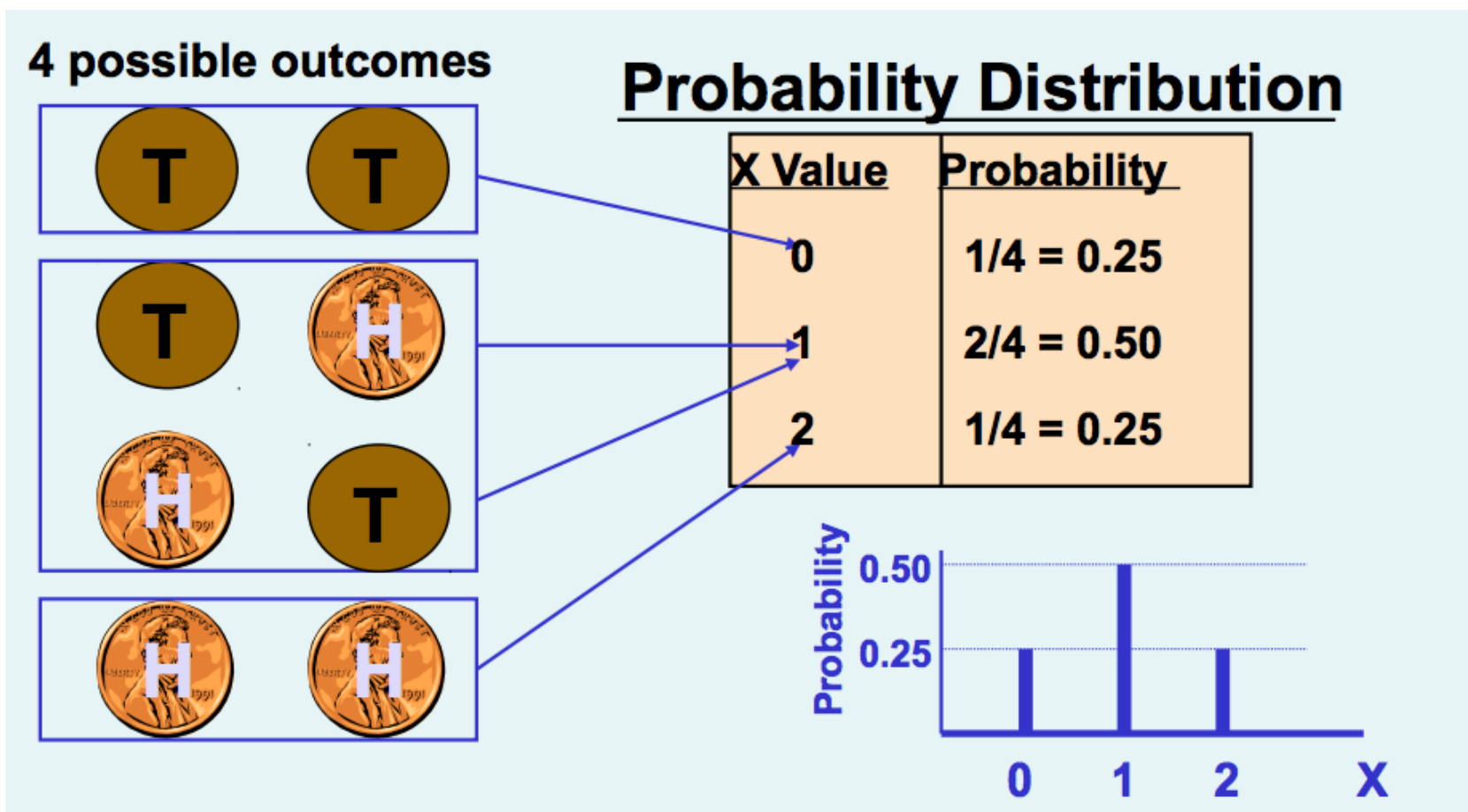
Basic Probability and Statistics Concepts

- Independent Events:
 - One event does not affect the other
 - Knowing probability of one event does not change estimate of another
- Random Variable:
 - A variable is called a random variable if it takes one of a specified set of values with a specified probability



Discrete Random Variable Probability Distribution

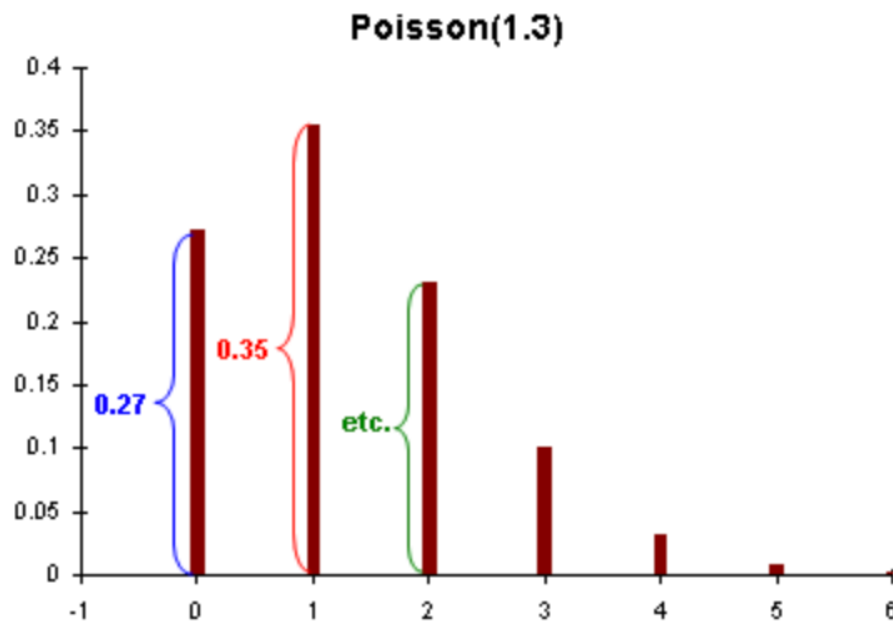
Experiment: Toss 2 Coins. Let $X = \#$ heads



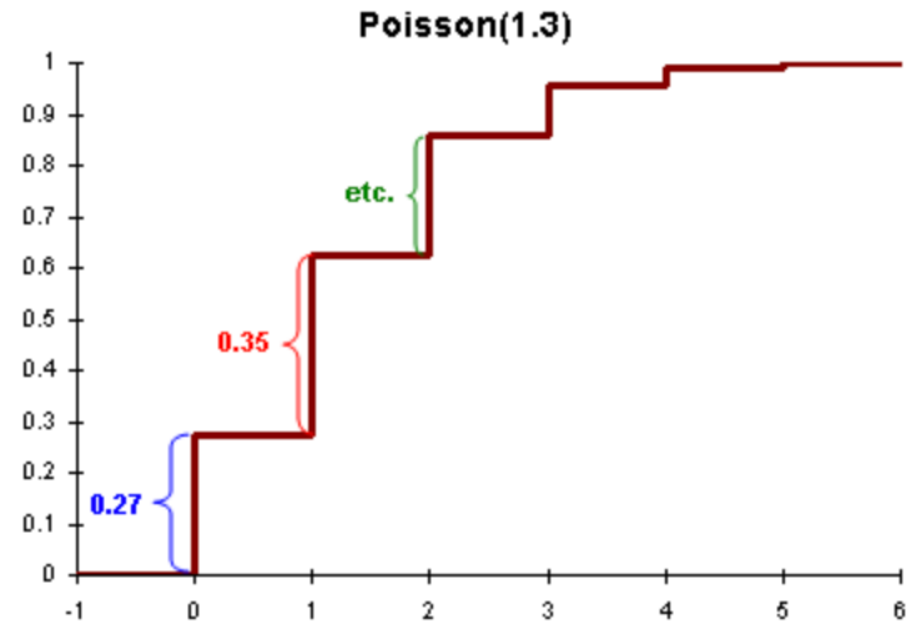
Histogram and Cumulative Distribution Function

Histogram: $f(x_i) = p_i$

Cumulative Distribution Function: $F_x(a) = P(x \leq a)$



Histogram



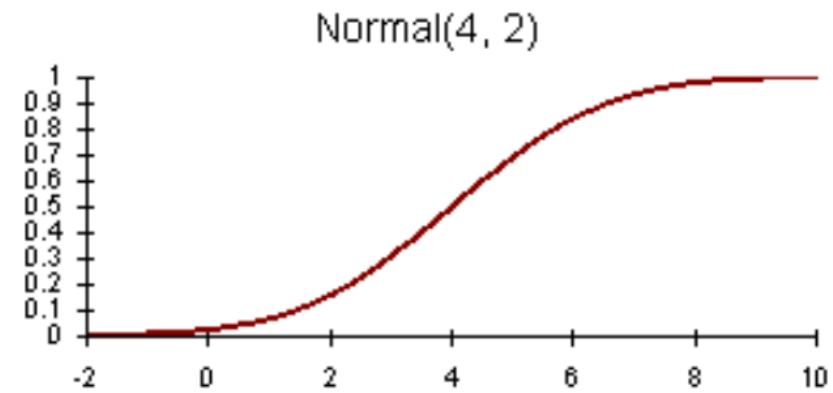
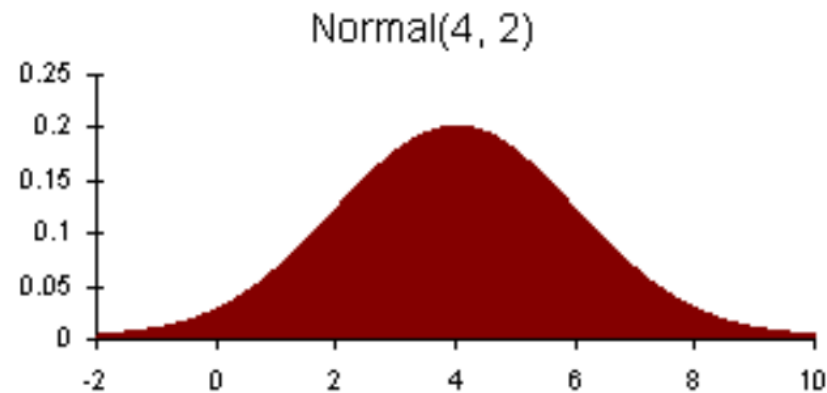
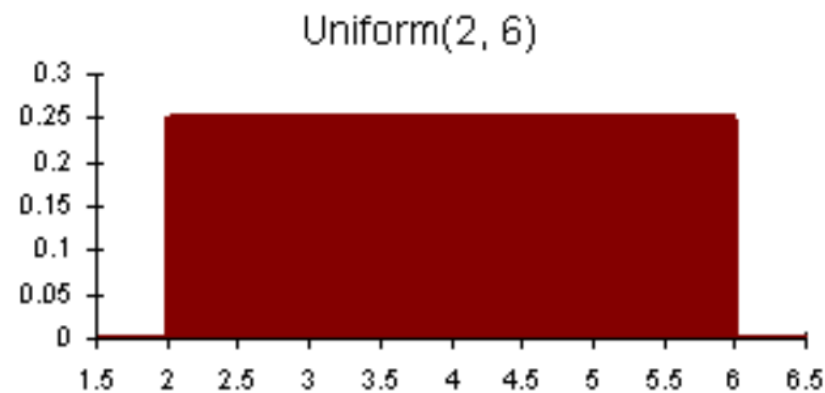
Cumulative Distribution Function (CDF)

Continuous Random Variable Probability Density Function

The probability density function, *pdf*, as $f(x)$.

$$F_x(a) = P(x \leq a)$$

The cumulative distribution function, *cdf*, as $F(x)$.

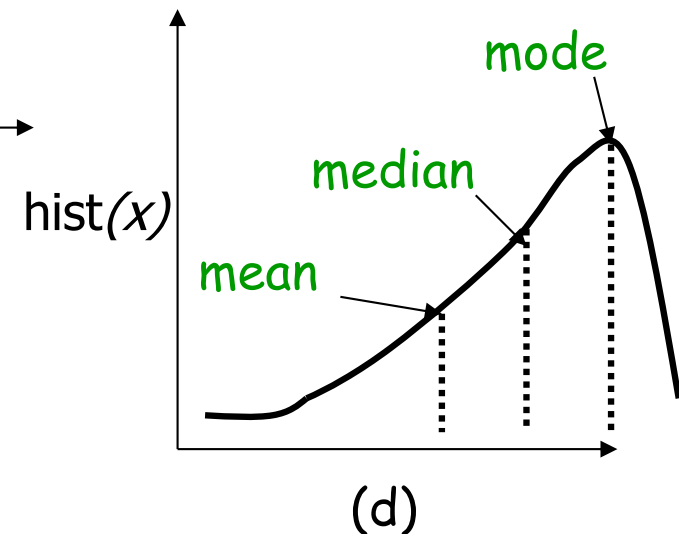
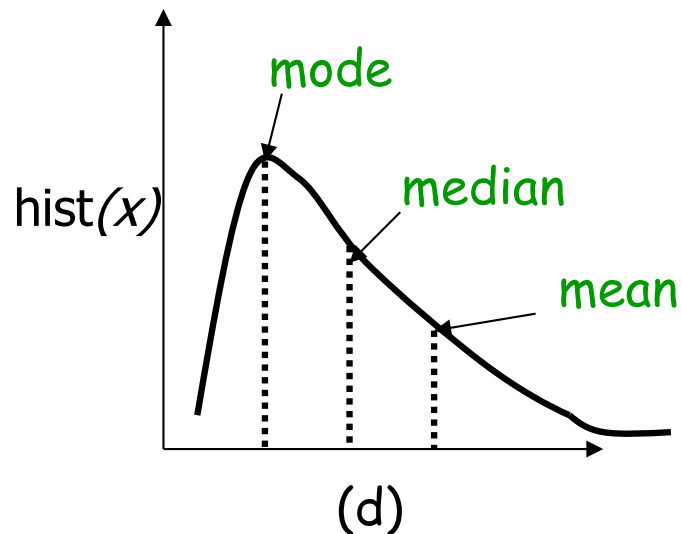
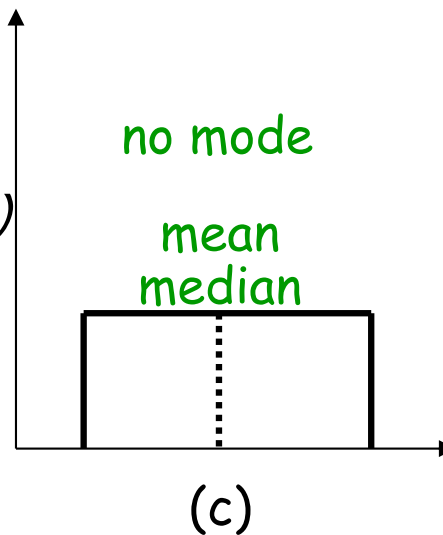
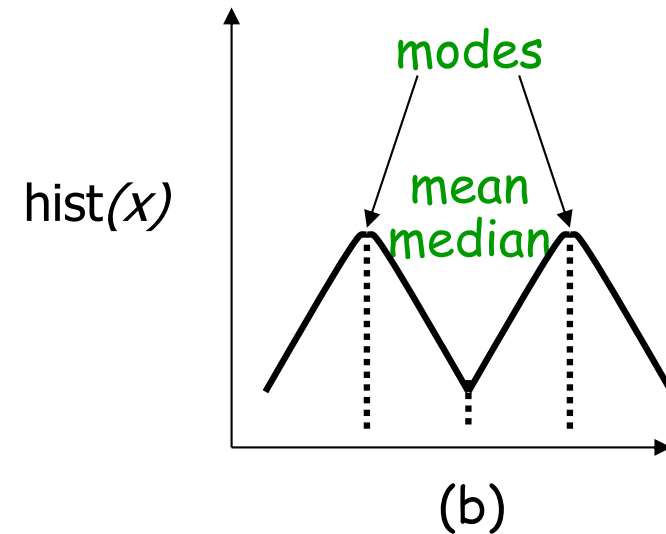
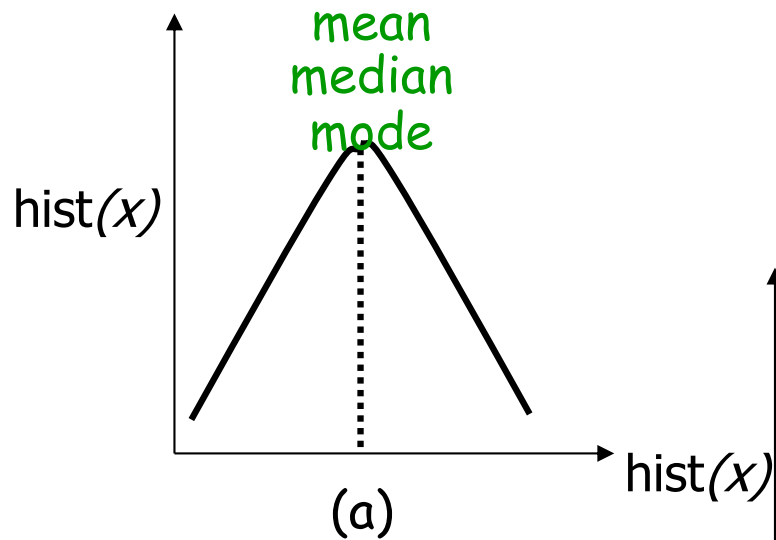


Indices of central tendency

Summarizing Data by a Single Number

- **Mean** – sum all observations, divide by number
- **Median** – sort in increasing order, take middle
- **Mode** – plot histogram and take largest bucket
- Mean can be affected by outliers, while median or mode ignore lots of info
- Mean has additive properties (mean of a sum is the sum of the means), but not median or mode

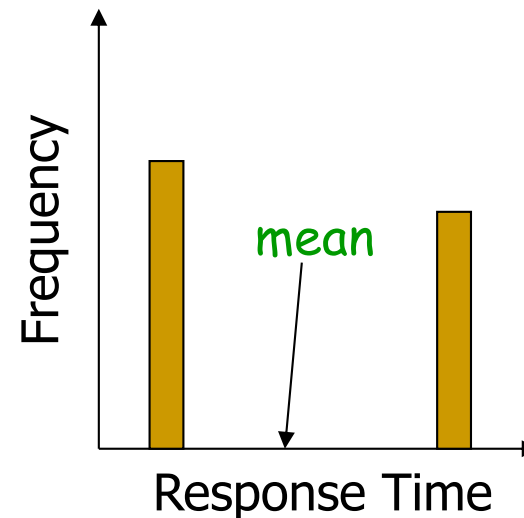
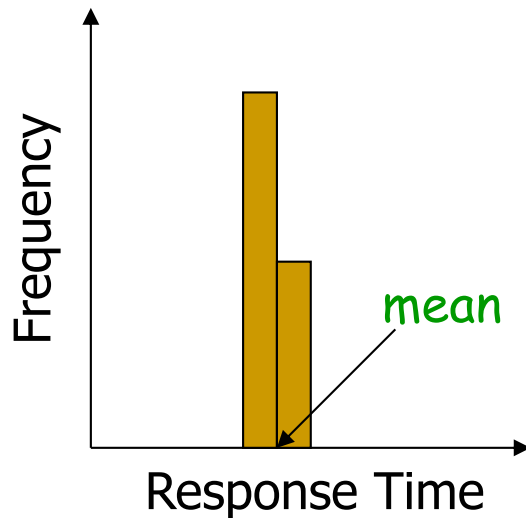
Relationship Between Mean, Median, Mode



Summarizing Variability

"Then there is the man who drowned crossing a stream with an average depth of six inches." – W.I.E. Gates

- Summarizing by a single number is rarely enough
→ need statement about *variability*



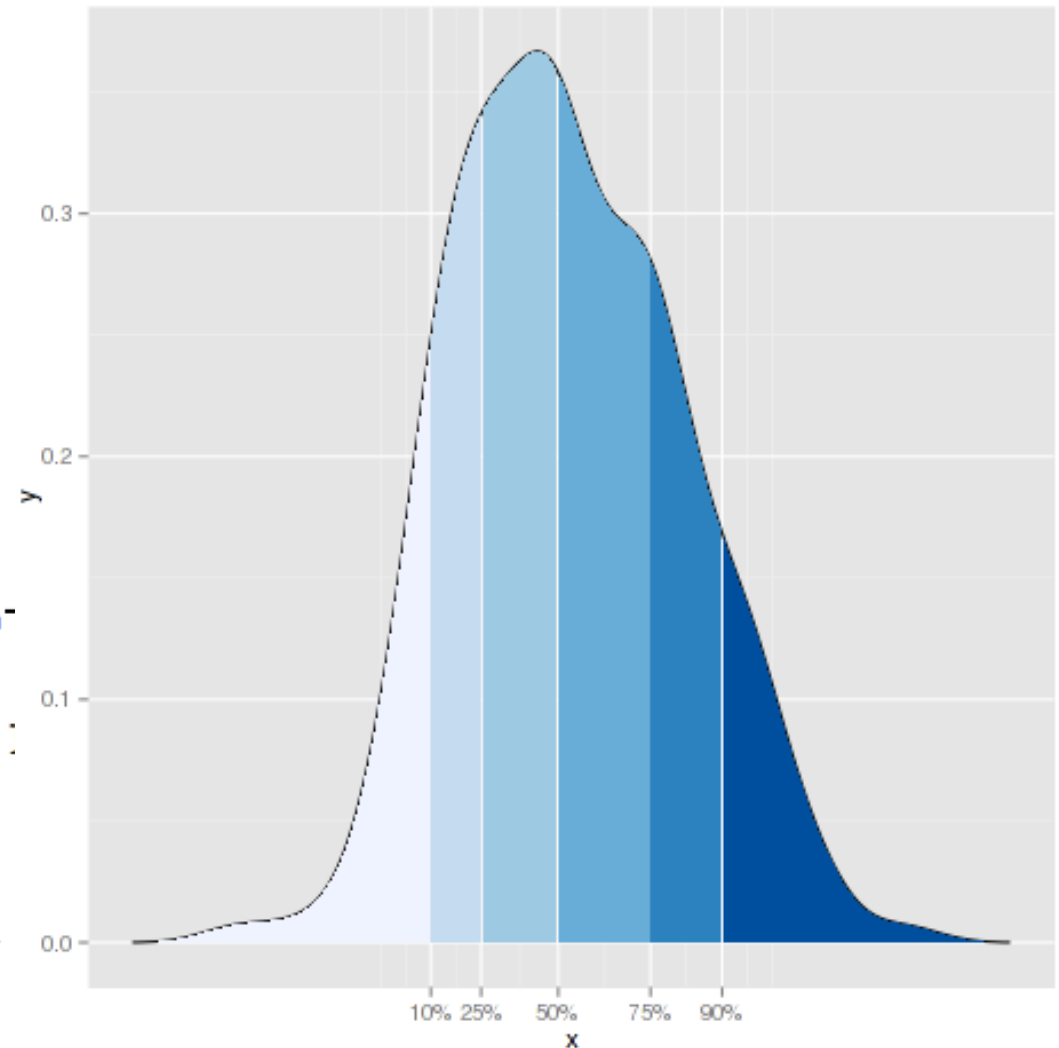
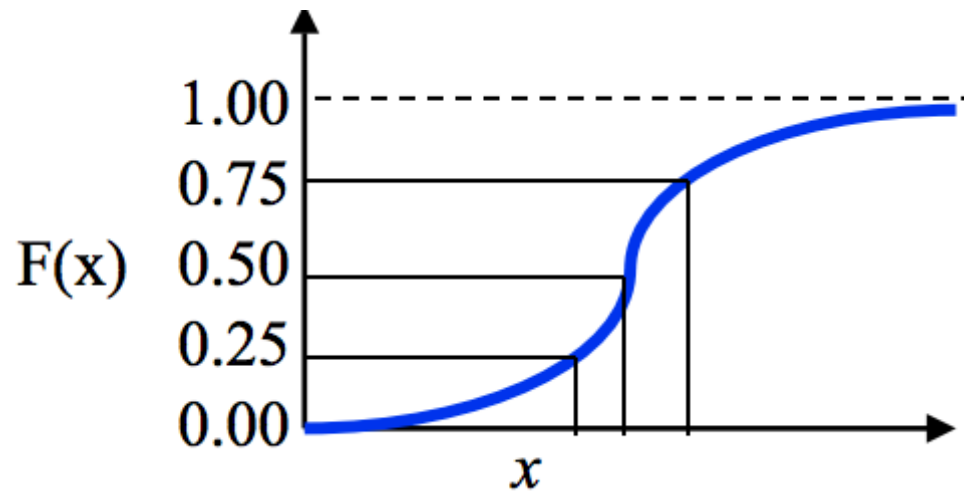
If two systems have same mean, tend to prefer one with less variability

Indices of Dispersion

- *Range* – min and max values observed
- *Variance* or *standard deviation* or *CoV*
 - Variance: Square of the distance between a set of values x_i with relative frequency p_i and the mean μ
 - $\sigma^2 = E[(x - \mu)^2] = \sum_{i=1}^n p_i (x_i - \mu)^2$
 - or, if you have exactly n samples $x_1 \dots x_n$
 - $\sigma^2 = E[(x - \mu)^2] = \frac{1}{n} \sum_{i=1}^n (x_i - \mu)^2$
 - Standard deviation, σ , is square root of variance
 - Coefficient of Variation [C.O.V.]: Ratio of standard deviation to mean: $= \sigma / \mu$
- *Percentiles*
 - The x value at which the *cdf* takes a value α is called the α -percentile and denoted x_α , so $F(x_\alpha) = \alpha$

Indices of Dispersion

- 10- and 90-*percentiles*
- [*Semi*]-interquartile range (SIQR)
 - Q1, Q2 and Q3



Determining Distribution of Data

- Additional summary information could be the *distribution* of the data
 - Ex: Disk I/O mean 13, variance 48. Ok. Perhaps more useful to say data is *uniformly distributed* between 1 and 25.
 - Plus, distribution useful for later simulation or analytic modeling
- How do determine distribution?
 - Plot histogram

For more formal testing: statistical comparison of CDF (*Komolgorov-Smirnov test*) or PDF (*Chi-square test*)
The Art of Computer Systems Performance Analysis, pp. 460-465

Comparing Systems Using Sample Data

"Statistics are like alienists – they will testify for either side." – Fiorello La Guardia

- The word “sample” comes from the same root word as “example”
- Similarly, one sample does not prove a theory, but rather is an example
- Basically, a definite statement cannot be made about characteristics of all systems
- Instead, make probabilistic statement about range of most systems
 - *Confidence intervals*

Sample versus Population

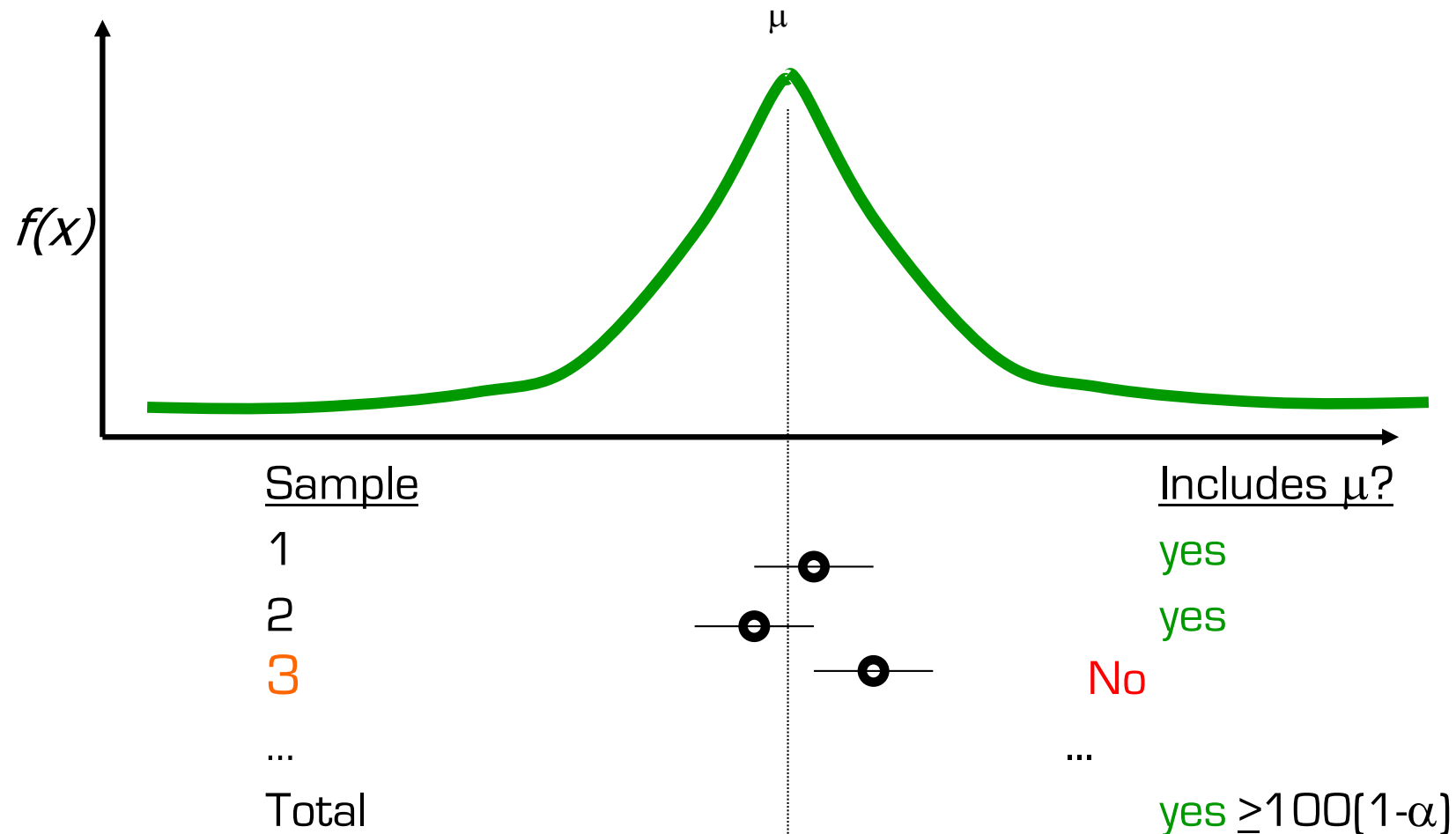
- Say we generate 1-million random numbers
 - mean μ and stddev σ .
 - μ is *population mean*
- Put them in an urn draw sample of n
 - Sample $\{x_1, x_2, \dots, x_n\}$ has mean $\bar{x}$, stddev s
- $\bar{x}$ is likely different than μ !
 - With many samples, $\bar{x}_1 \neq \bar{x}_2 \neq \dots$
- Typically, μ is not known and may be impossible to know
 - Instead, get estimate of μ from $\bar{x}_1, \bar{x}_2, \dots$

Confidence Interval for the Mean

- Obtain probability of μ in interval $[c_1, c_2]$
 - $Prob(c_1 \leq \mu \leq c_2) = 1 - \alpha$
 - $[c_1, c_2]$ is the *confidence interval*
 - α is the *significance level*
 - $100(1 - \alpha)$ is the *confidence level*
- Typically want α small so confidence level 90%, 95% or 99% (more later)
- Use 5-percentile and 95-percentile of the sample means to get 90% confidence interval

Meaning of Confidence Interval

- For a 90% confidence level, if we take 100 samples and construct the confidence interval for each sample, the interval would include the population mean in 90 cases.



What if n not large?

- Above only applies for large samples, 30+
- For smaller n , can only construct confidence intervals if observations come from normally distributed population: t-variate

$$- \left[\bar{x} - t_{\left[\frac{1-\alpha}{2}; n-1\right]} \frac{s}{\sqrt{n}}; \bar{x} + t_{\left[\frac{1-\alpha}{2}; n-1\right]} \frac{s}{\sqrt{n}} \right]$$

$\bar{x}$: sampled mean

s : sampled standard deviation

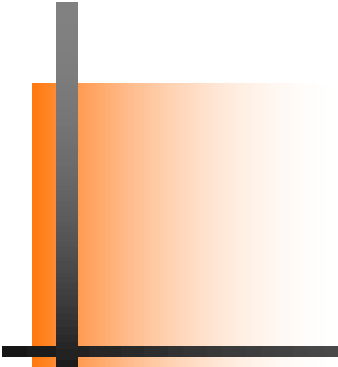
n : number of samples

- Table A.4 of Jain's book

$t_{\left[\frac{1-\alpha}{2}; n-1\right]}$: tabulated value of the student-t distribution

What Confidence Level to Use?

- Often see 90% or 95% (or even 99%), but...
- Example:
 - Lottery ticket \$1, pays \$5 million
 - Chance of winning is 10^{-7} (1 in 10 million)
 - To win with 90% confidence, need 9 million tickets
 - No one would buy that many tickets!
 - So, most people happy with 0.01% confidence



Performance in distributed systems

Performance is an art

Performance evaluation is an art

Like a work of art, a successful evaluation cannot be produced mechanically

Every evaluation requires an intimate knowledge of the system and a careful selection of methodology, workloads and tools.

Example of the need for knowledge: know your tradeoffs

- “Bufferbloat” is a term used when greedy, loss-based TCP flows probing for bandwidth fill up a large FIFO queue leading to added delay for all flows traversing this bottleneck.
- To mitigate this, aggressively dropping timer-based AQMs or shorter queues are recommended.
- What do you sacrifice by reducing the size of the queue?

Performance evaluation is an art

A major part of the analyst's "art" is:

- defining the real problem from an initial intuition, and
- converting it to a form in which established tools and techniques can be used, and
- where time and other constraints can be met

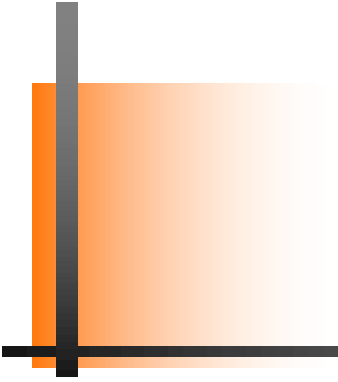
Two analysts may choose to interpret the same measurements in two different ways, thus reaching different conclusions

Performance evaluation is an art

The throughputs of two systems A and B were measured in transactions per second. The results were as follows:

System	Workload 1	Workload 2
A	20	10
B	10	20

This is called a **ratio game**. It is not appropriate for objective analysis, but useful for propaganda.

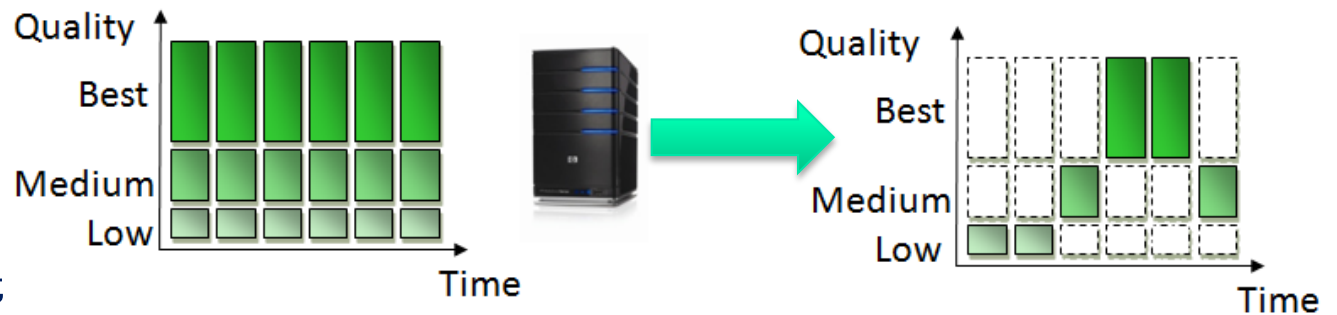


Performance in distributed systems

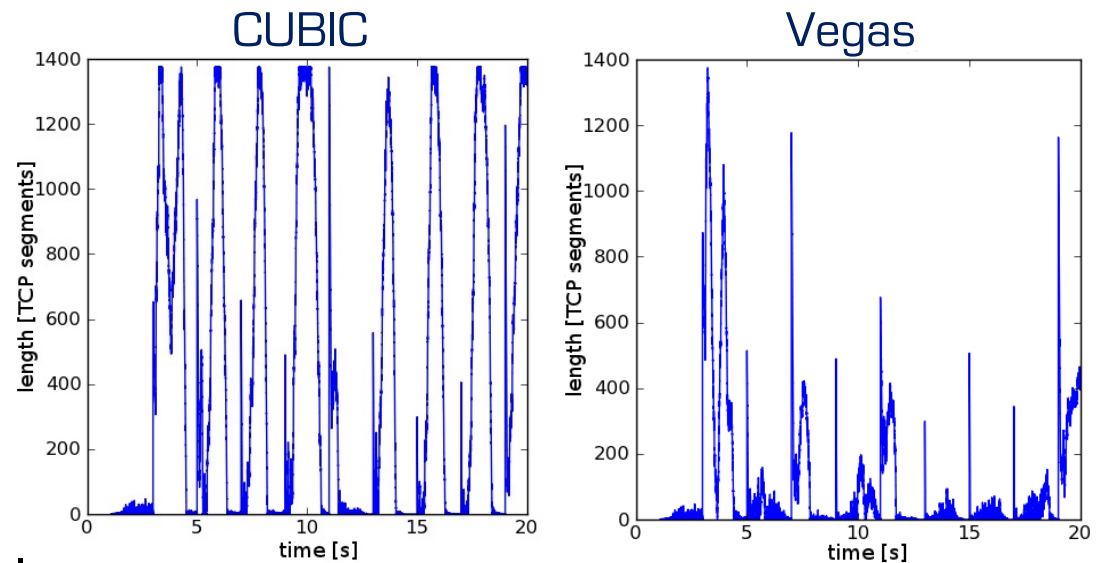
Real-world examples

Emulation study

Investigation of router queue length development in DASH streaming for different TCP Congestion Control algorithms



- Simple 2D graph showing an *independent parameter* X (time) and a *dependent parameter* Y (queue length)
- does illustrate the unstable queue length for CUBIC, but no actual distribution
- not a quantifiable result, but anecdotal

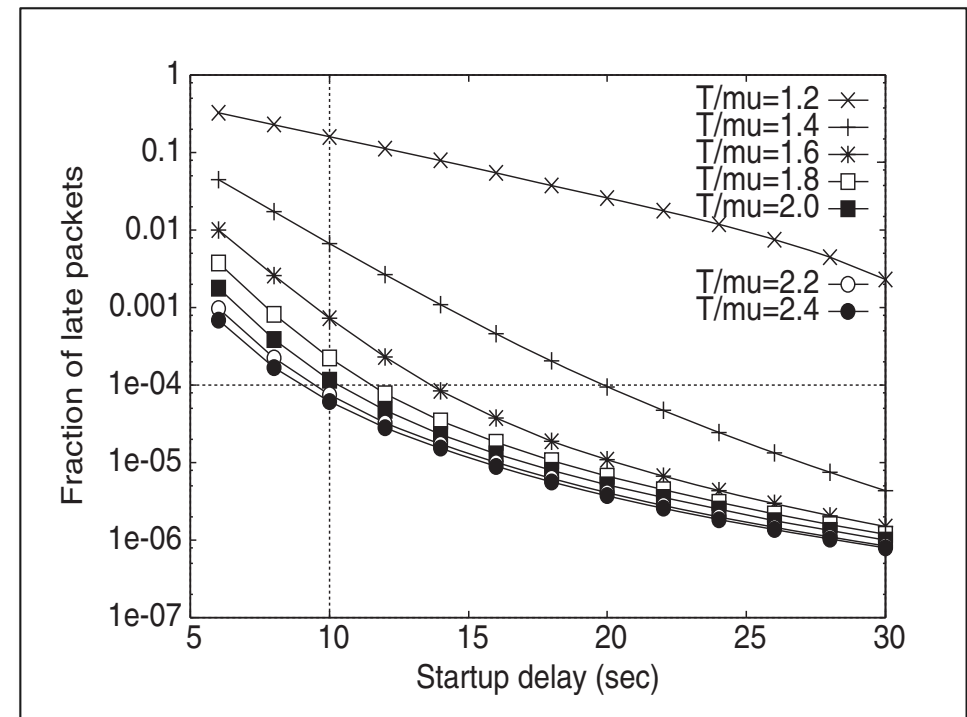


Analytical performance study

Analytical performance study to discover a relation between streaming [video] over TCP and the likelihood of stalling

Analytical graph provides deterministic, repeatable results

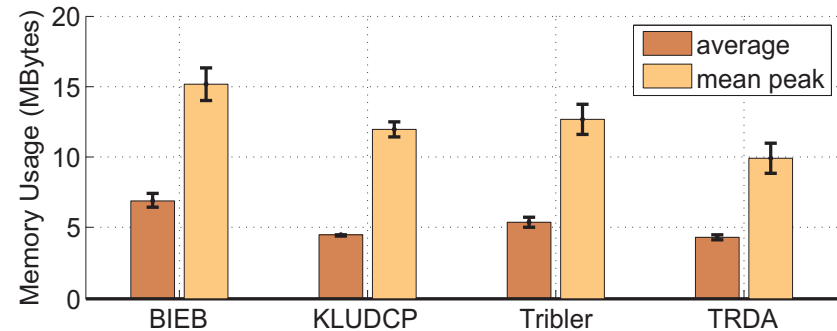
- symbols distinguish conditions
- Y-axis is logarithmic to expose differences at when very few packets are late
- note that each point is a computation with different parameters



Simulation study

Investigation of memory requirements for several DASH streaming algorithms

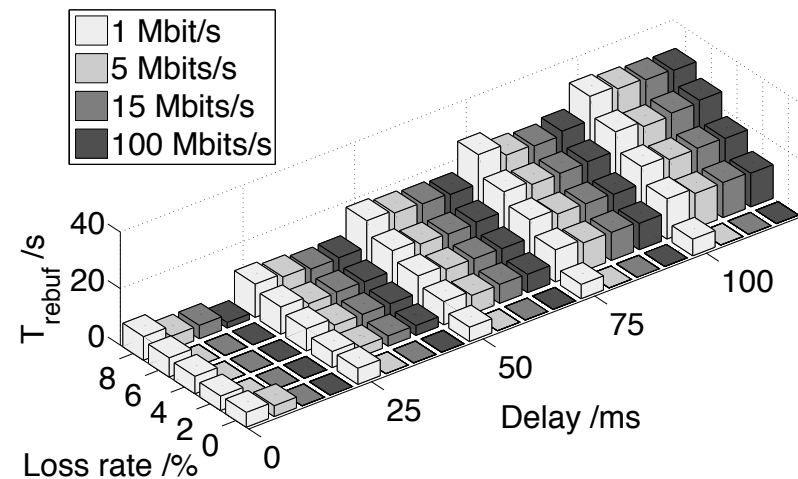
- *Block diagram* is suitable when X-axis values have no metric relation (no measure of any distance between them)
- block diagram is also better if X-values have an order but no metric relation!
- 2D graph merges 2 questions into 1 graph: average memory use and average peak memory use (average of peaks of several simulation runs) – this does not scale to many questions
- standard deviation is added for each of the averages



Emulation example

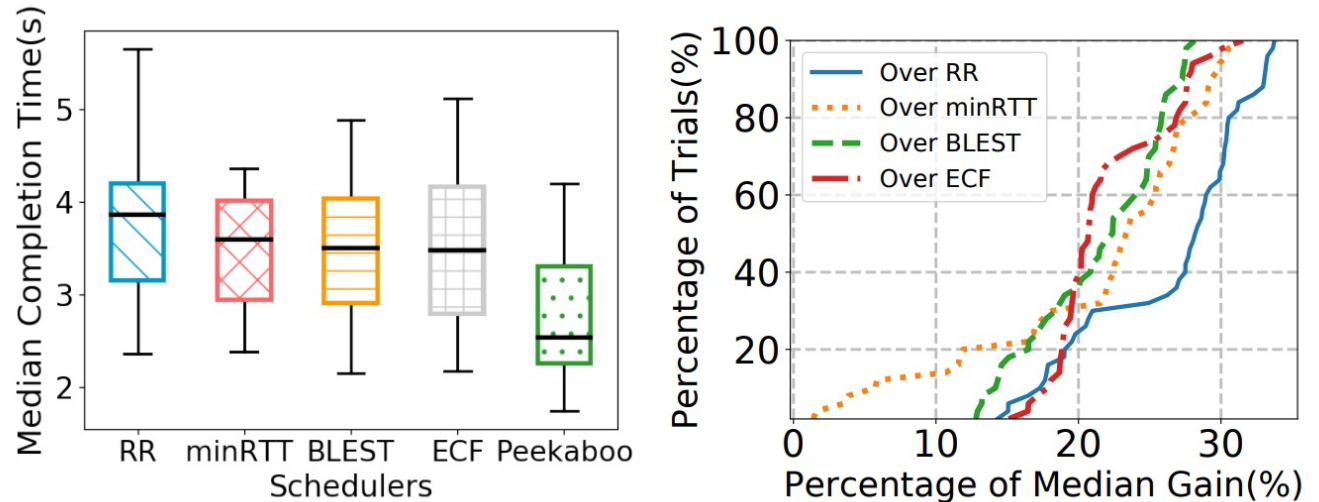
Some HTTP adaptive video streaming strategies can fail when packet loss is high and network delay is high as well. How long are the cumulative waiting times?

- *3D block diagram*
- 3 independent variables shown
- 4D information [X, Y, Z, color]
- 3 independent variables (loss rate, delay, network capacity)
- 1 dependent variable (rebuffering time)
- visually attractive
- tolerances (confidence intervals etc.) cannot be expressed
- absolute height cannot be ascertained by reader for all conditions
- does not scale to many network capacities



Emulation example

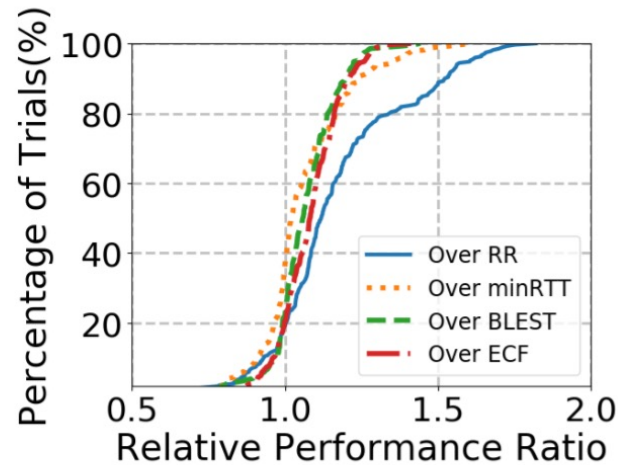
How will a learning-based multipath scheduler (Peekaboo) improve the completion time over heterogenous networks?



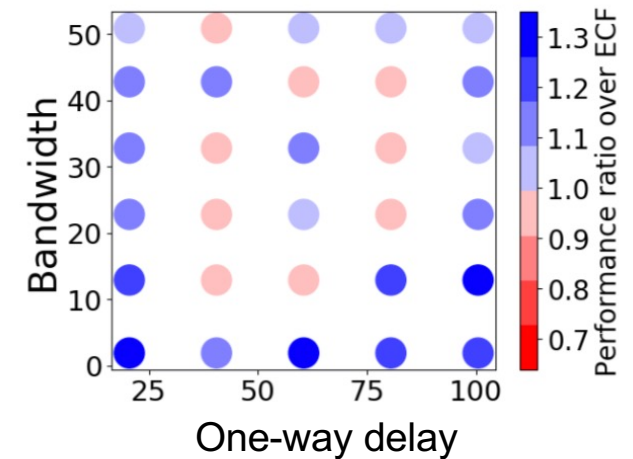
- *Boxplots and CDFs*
- Boxplots provide the statistical information both on the lower and upper percentiles
- CDF of the gain illustrates the distribution of the gain
- Quantifies both the basic characteristics of schedulers as well as the how well the proposed scheduler performs compared to others

Emulation example

How will a learning-based multipath scheduler (Peekaboo) improve the completion time over heterogenous networks?



(a) Performance ratio over the whole design space



(b) Medium Low over ECF

- *Design space exploration*
- Emulations are often done over a design space of variables
- CDF of the gain illustrates the distribution of the gain but it does not tell where the gain comes from
- A heatmap over a design space can show where the gains come from
- Quantifies both the basic characteristics of gains as well as the where the gains (if any) coming from

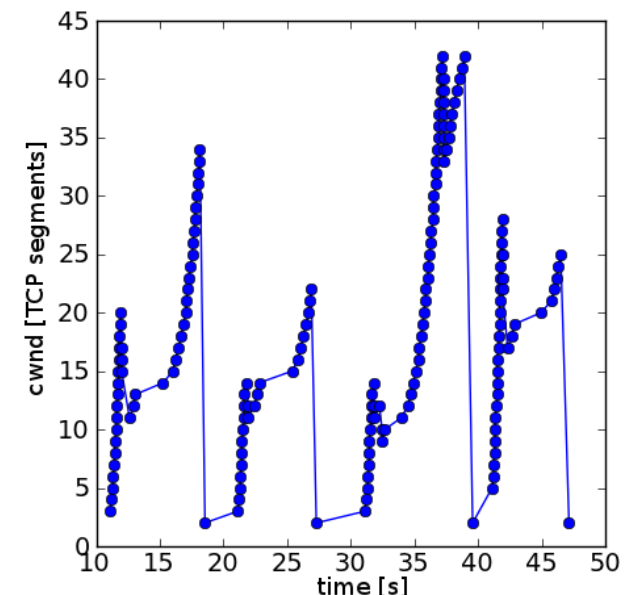
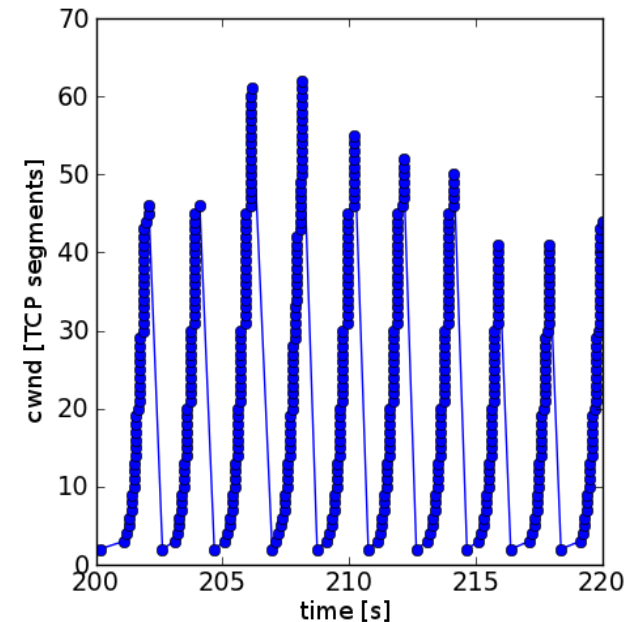
Emulation study

Investigation of sender's congestion window size in the same study.

Video segments have a duration of 2 seconds (top) and 10 seconds (bottom), the algorithm attempts to choose a quality that can be downloaded in 1 second.

- Simple 2D graph showing an independent parameter X (time) and a dependent Y (congestion window size)
- serves to illustrate that CUBIC is incapable of maintaining its congestion window between 2-second DASH segments, but enters TCP slow start
- not a quantifiable result, but anecdotal

CUBIC



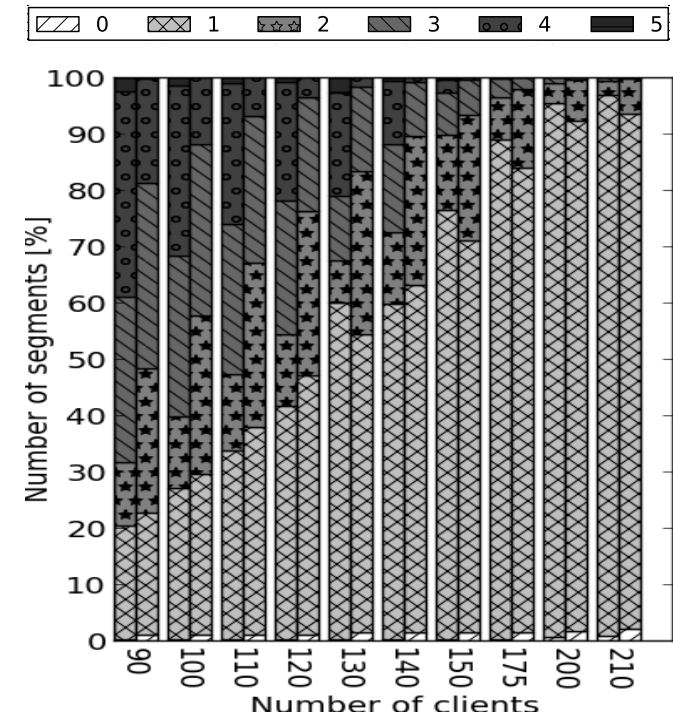
Emulation study

Investigation of the distribution of video quality in the same study.

Segments 2-sec. (left in each column) and 10-sec. (right). Patterns indicate qualities (0 stall, 5 best).

Shows the shares of qualities for entire film.

- Graph with 3 dimensions (X and segment duration independent, Y dependent)
- quite problematic
 - hard to distinguish qualities, patterns are not easily enough recognized
 - quality 1 is dominant, no visual comparison of the others
 - change of order between left and right remains hidden

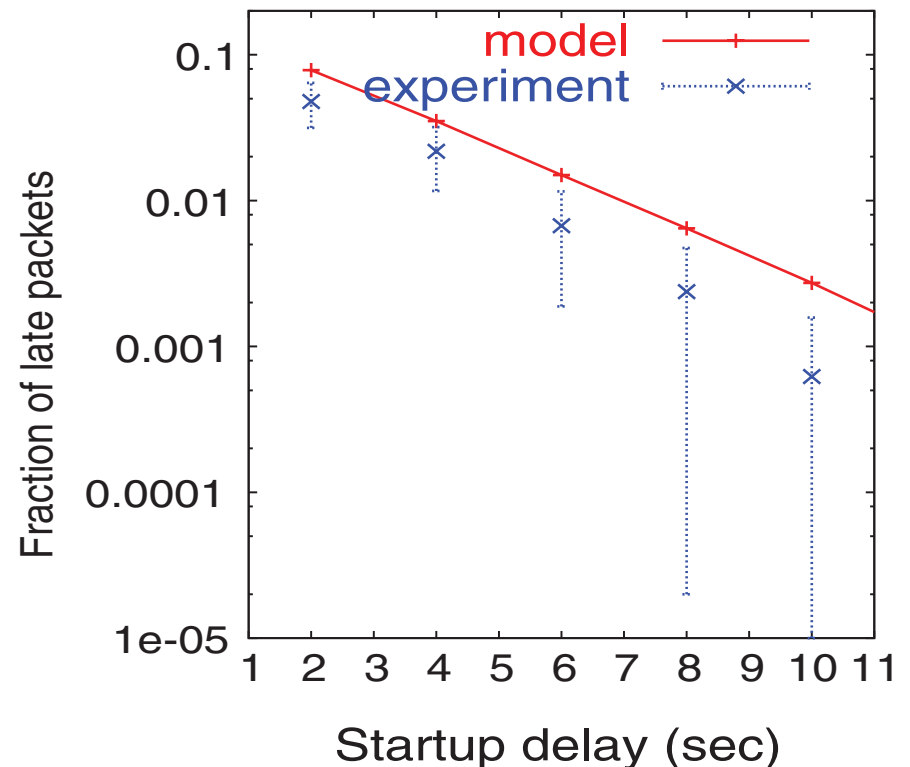


Combined study

Performance study to discover a relation between streaming (video) over TCP and the likelihood of stalling – model validated by ns-2 simulation (“experiment”)

Simulation

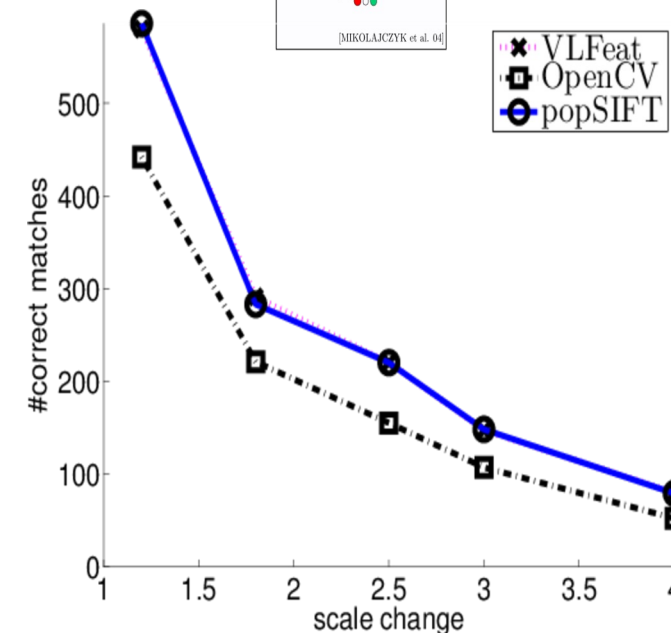
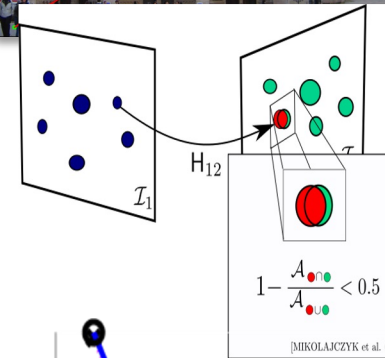
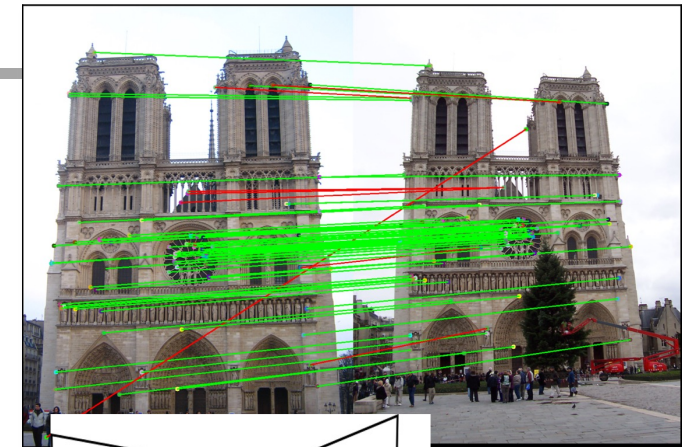
- symbols distinguish model and simulation
- Y-axis is logarithmic
- simulation is not deterministic, and error bars show the 95% confidence interval
- for the simulation, each point with error bars is derived from the results of 1000 simulation runs



Emulation example

Comparing the performance of the 3 implementations of the algorithm “Scale-Invariant Feature Transform” (SIFT)

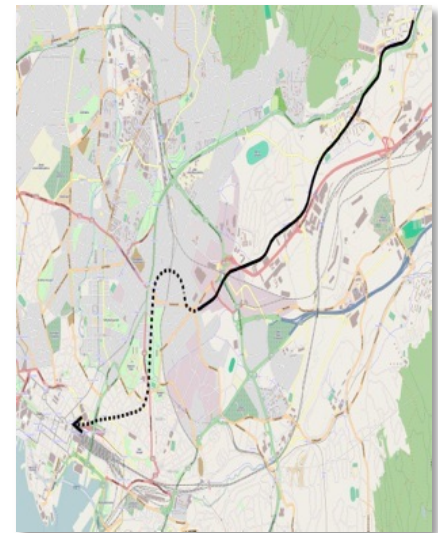
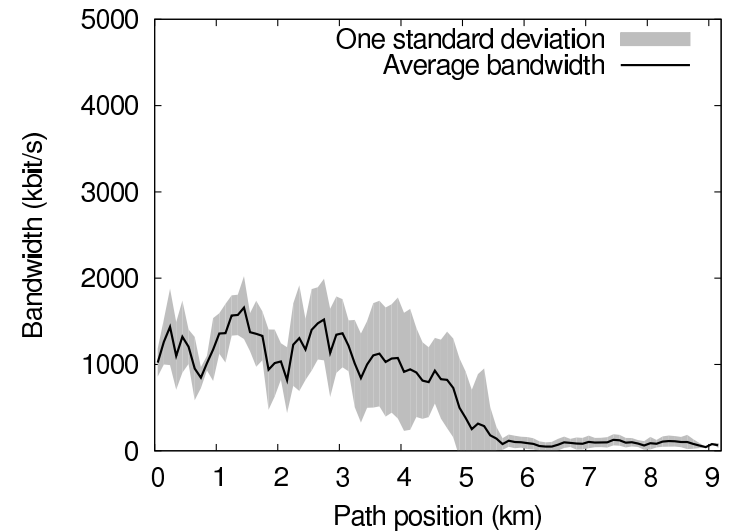
- Very simple 2D plot, relating only to a set of very specific image pairs
- 100% deterministic and repeatable, no point in expressing errors
- definition ahead of time: boolean condition that defines “match” (adopted from an independent study that developed a good comparison method)



Emulation example

Documenting the repeatability of bandwidth measurement on a typical commuter path

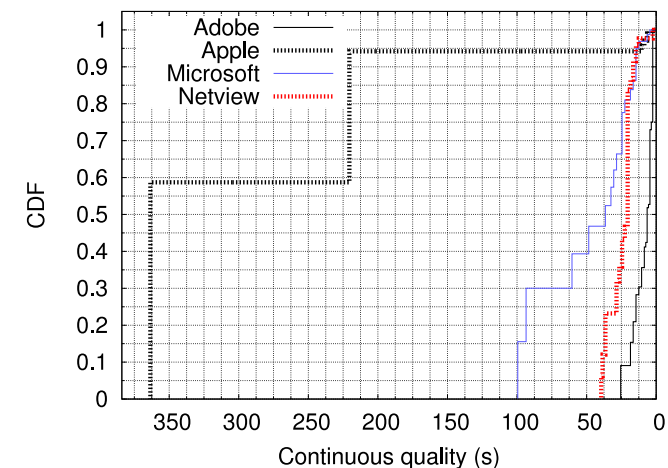
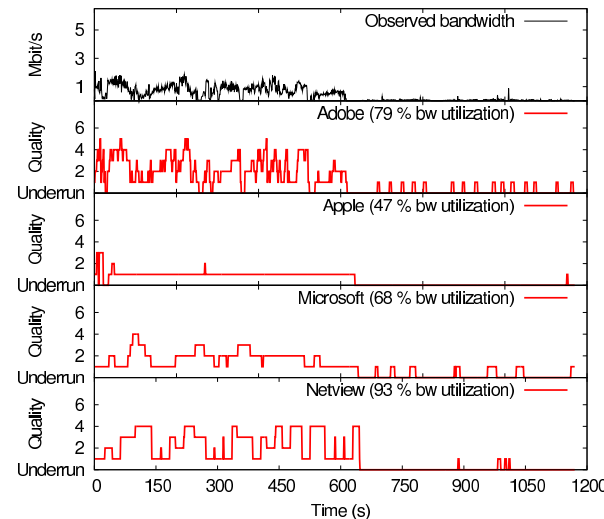
- Map shows the subway route from Stovner to Oslo S
- graph shows the measured bandwidth by distance from Stovner
- figure does not represent any specific measurement run, the measurements have been collected, and the graph shows both the average bandwidth and 1 standard deviation
- not only anecdotal but valid for predictions



Emulation example

Comparing the bandwidth efficiency and stability of several HTTP Adaptive Streaming methods

- Upper graph shows quality development over time
- by itself, it has only anecdotal value
- Lower graph shows CDF of quality changes
- Apple HLS is most stable (a desirable property)
- but upper graph exposes that the price for this is very low quality

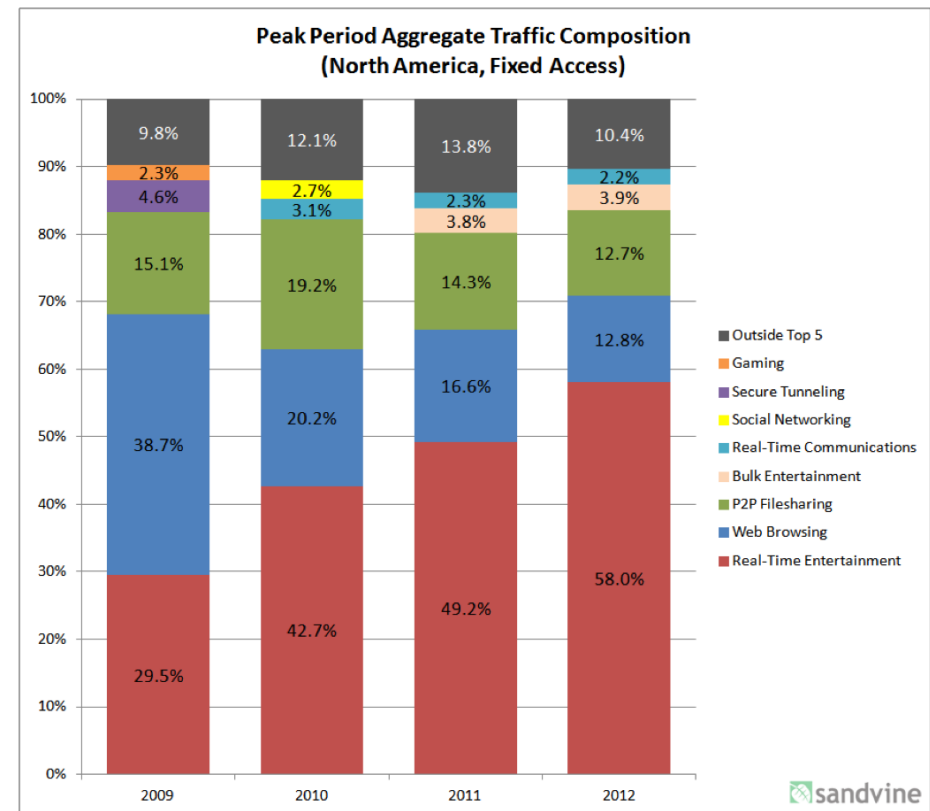


Measurement example

Development of traffic shares
over time

A graph using percentages to
express the share of
application types on the
Internet

- no absolute values, only percentages
- color as well as order allows easy recognition of types, as well as appearance of new types

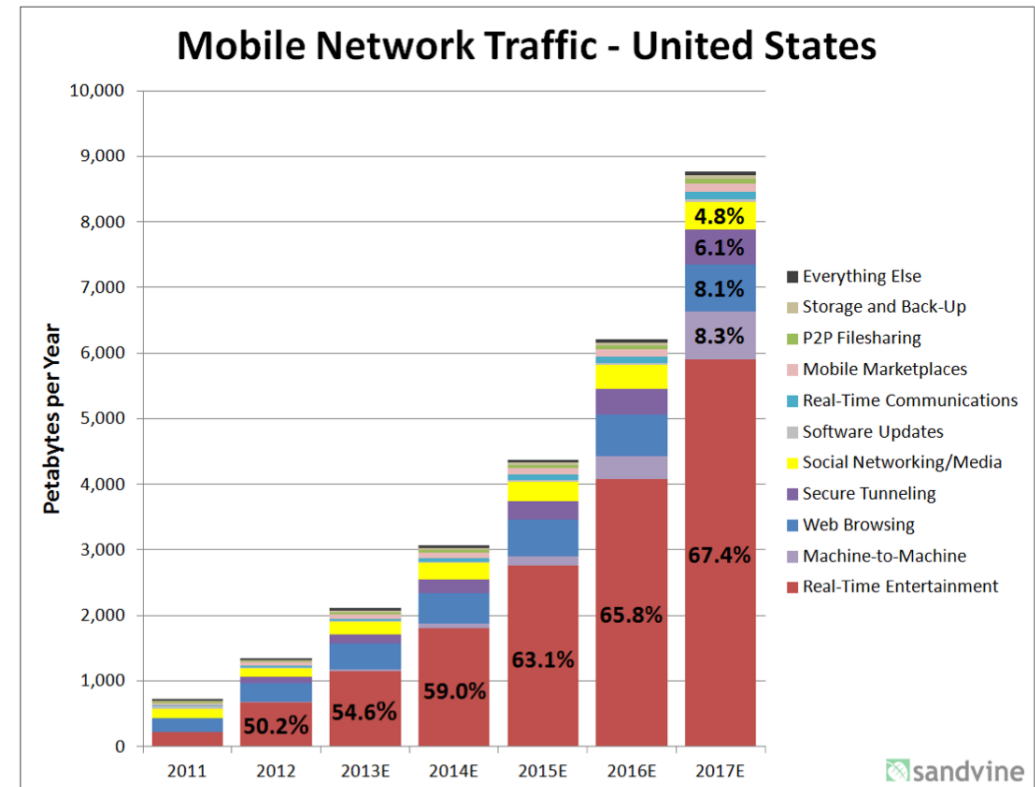


Measurement example

Development of absolute mobile traffic over time

A graph using absolute value to communicate the rapid growth of mobile traffic

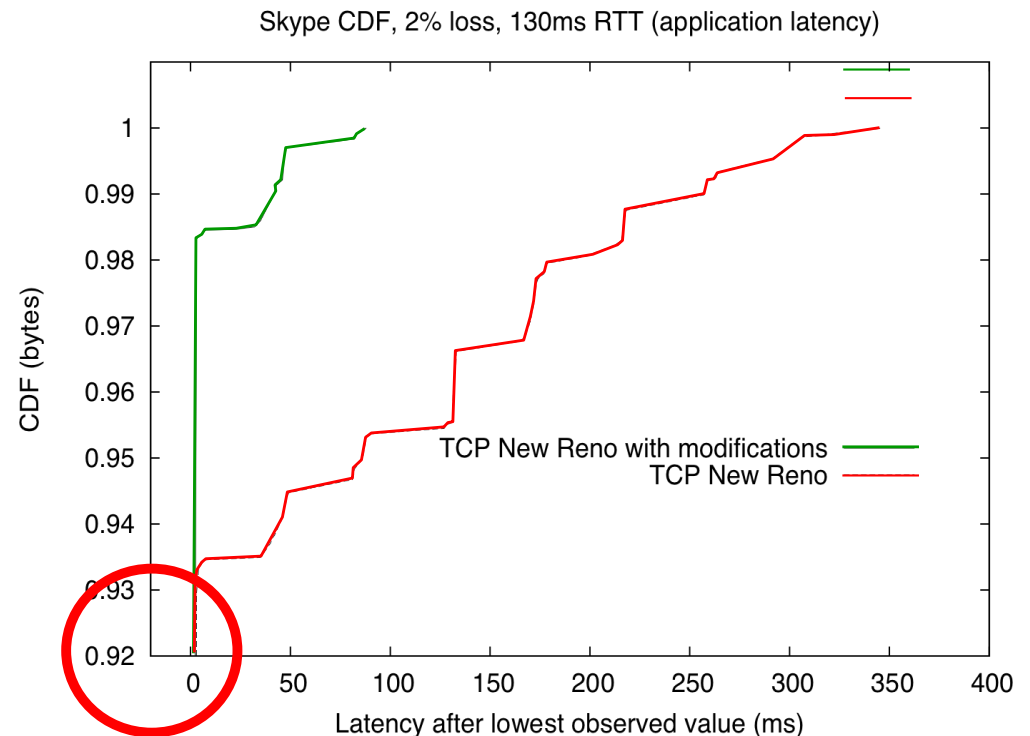
- percentages provided as text in graph
- color as well as order allows easy recognition of types, as well as appearance of new types
- note "E" for estimates



Measurement example

Hypothesis: “Thin-stream” modifications to Linux’s implementation of TCP New Reno reduce latency.

- *Cumulative Distribution Function* (CDF) provides the percentage of measurement points up to a given X value
- useful if number of samples not identical
- useful if number of samples is quite large



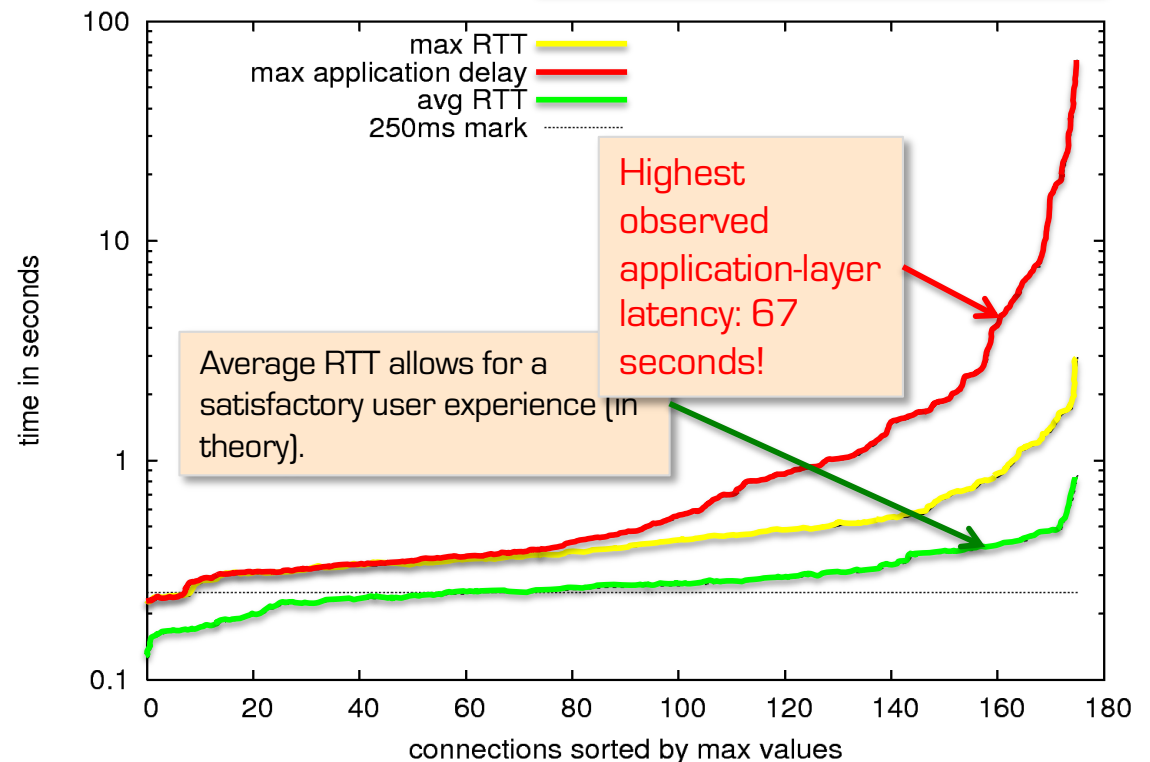
starts far away from 0

Measurement example

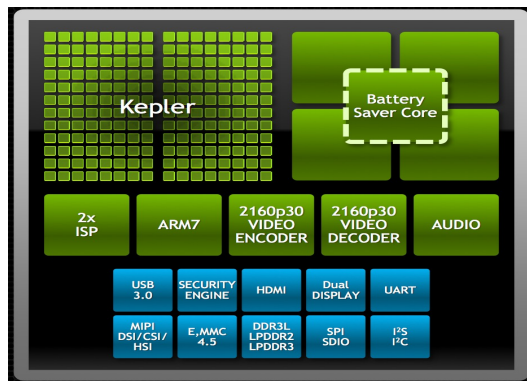
Application-layer behaviour of a popular MMORPG estimated from a single server-sided measurement probe



- simple 2D presentation, both dimensions observed
- *data sorting* can provide more information than *histograms* or *cumulative distribution functions*

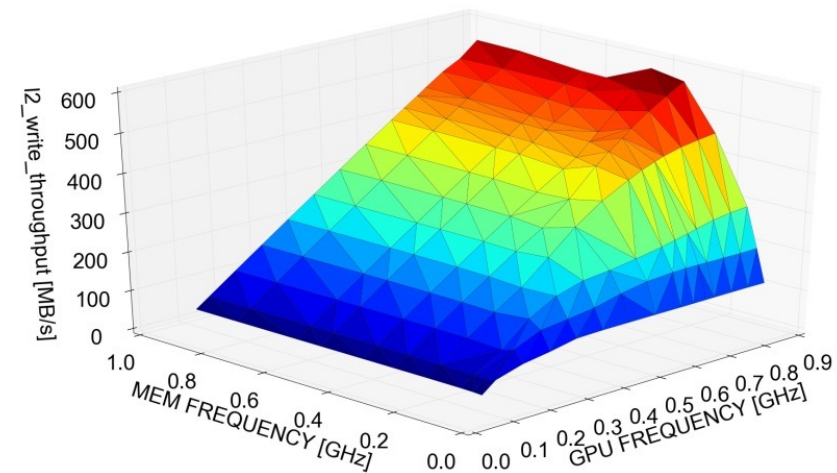


Measurement example



NVidia Tegra K1

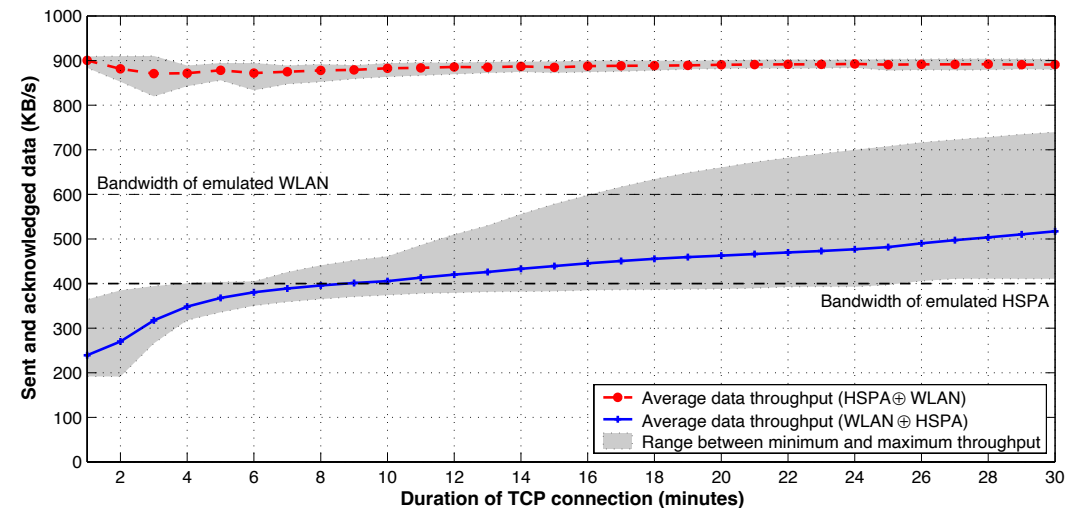
impact of frequency on throughput



- note: *4 dimensions* in the presentation
- additional dimension can be used to add information or to add expressiveness to one or more of the dimensions

Measurement example

Linux TCP's ability to recover from out-of-order delivery of packets

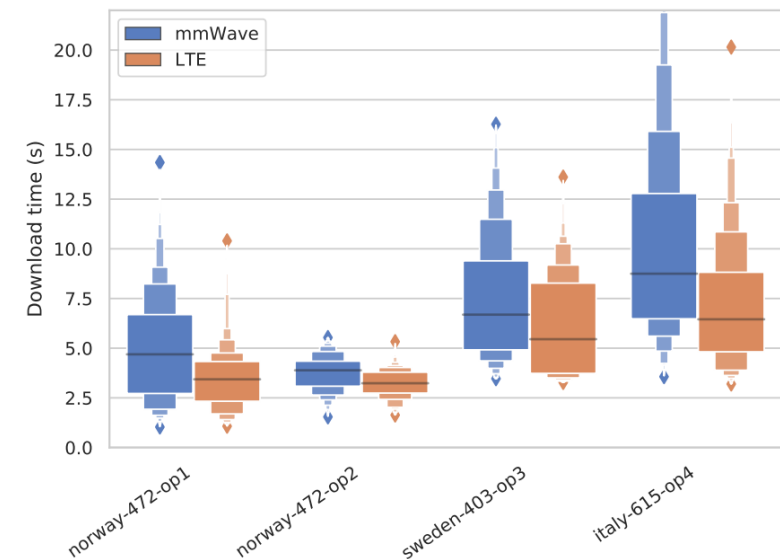
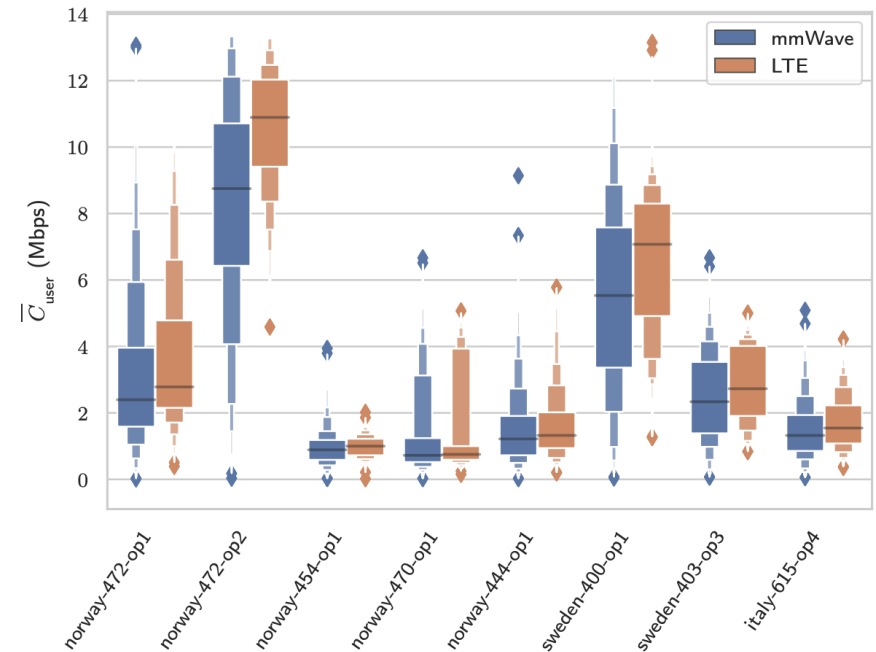


- 2D graph studying values for the average bandwidth of very long-lived TCP flows whose packets are alternately sent over 2 very different paths
- details of short-term TCP behaviour are completely hidden
- smoothness achieved by averaging
- *shaded areas illustrate uncertainty* (range from min to max average throughput)

Measurement example

Investigation of the current operator SPEPs (used in 4G LTE networks) under mmWave-like dynamics using MONROE testbed

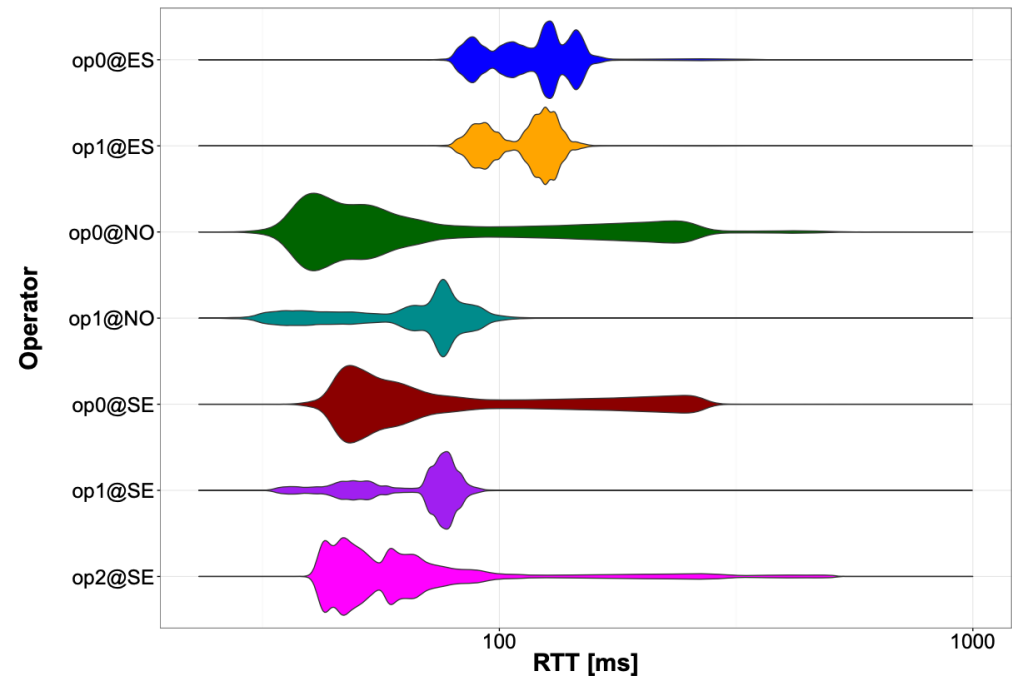
- Different colors highlight the different settings
- Boxplots display the first two letter values (the median and quartiles); letter-value plots display further letter values therefore more suitable for large datasets.



Measurement example

Investigation of latency for 4G LTE networks using MONROE testbed

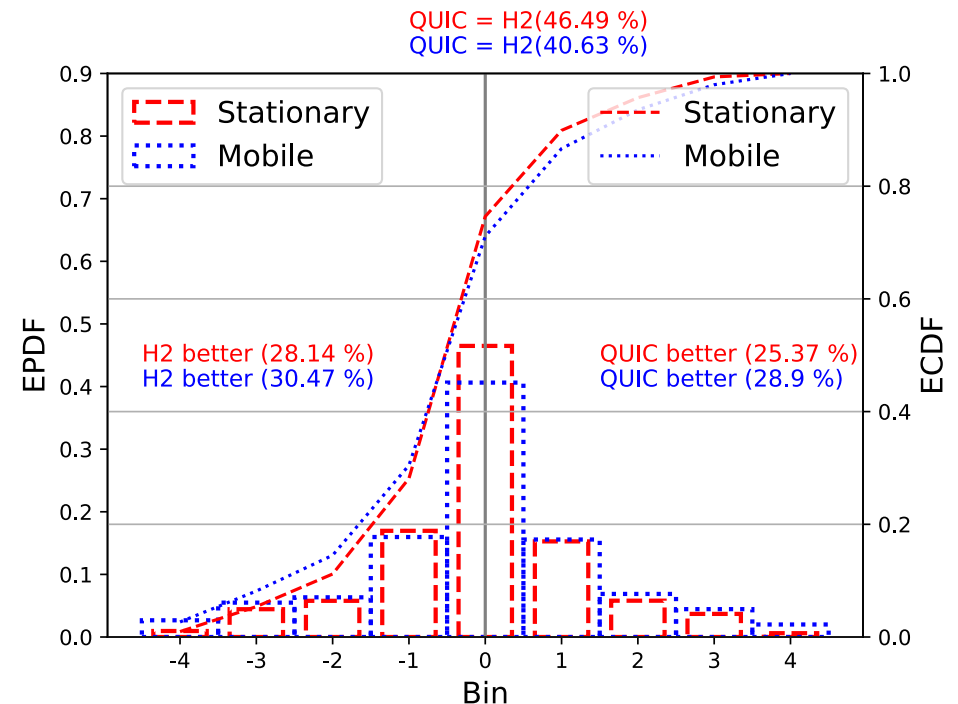
- Different colors highlight the different settings
- Violin plots are similar to box plots, except that they also show the probability density of the data at different values.



Measurement example

Performance comparison of HTTP2 and QUIC over 4G networks under stationary and mobile scenarios

- Color together with pattern is used to differentiate different settings
- Empirical PDF is presented together with empirical CDF
- The bins are used to illustrate the relative difference levels



Measurement example

The impact of COVID-19 on the web performance over mobile networks

- 2D matrix form together with heatmap
- The heatmap together with values in it provide more information
- highly aggregated data
- concept of certainty (e.g. confidence intervals) gets lost

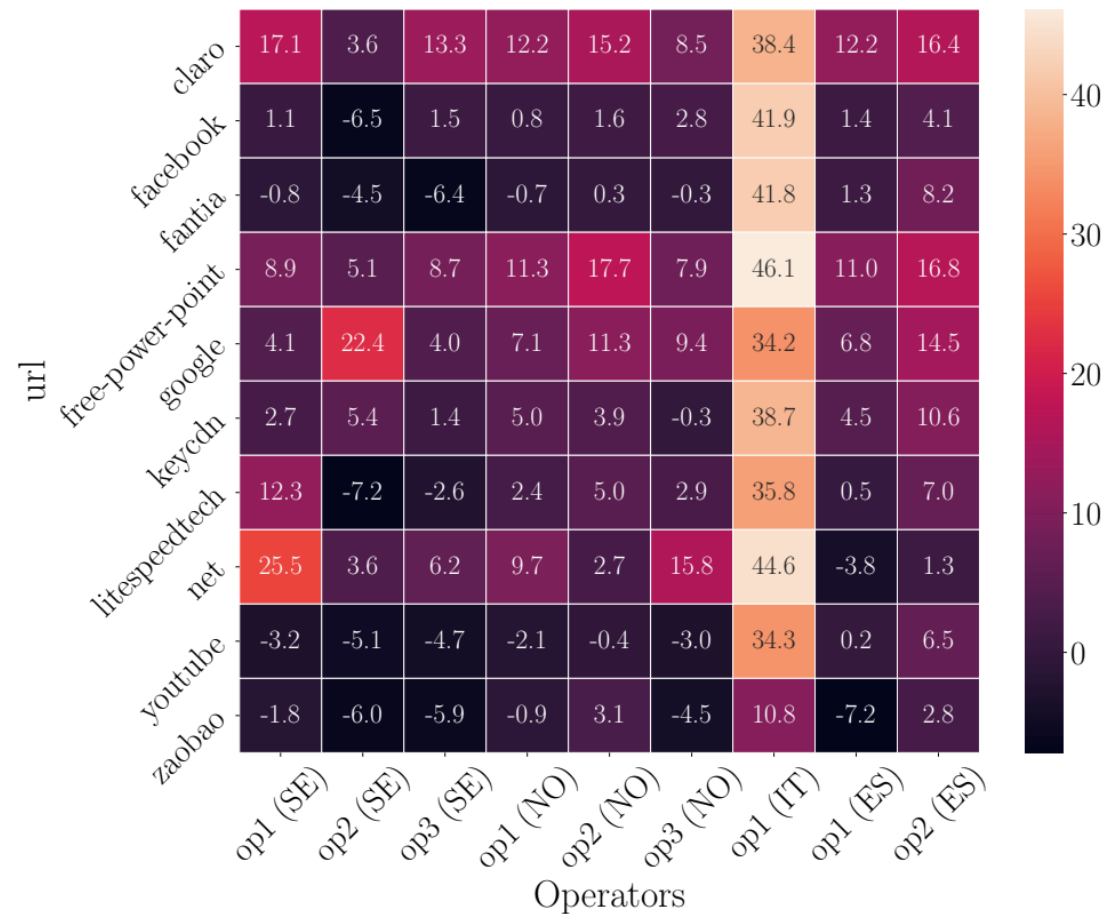
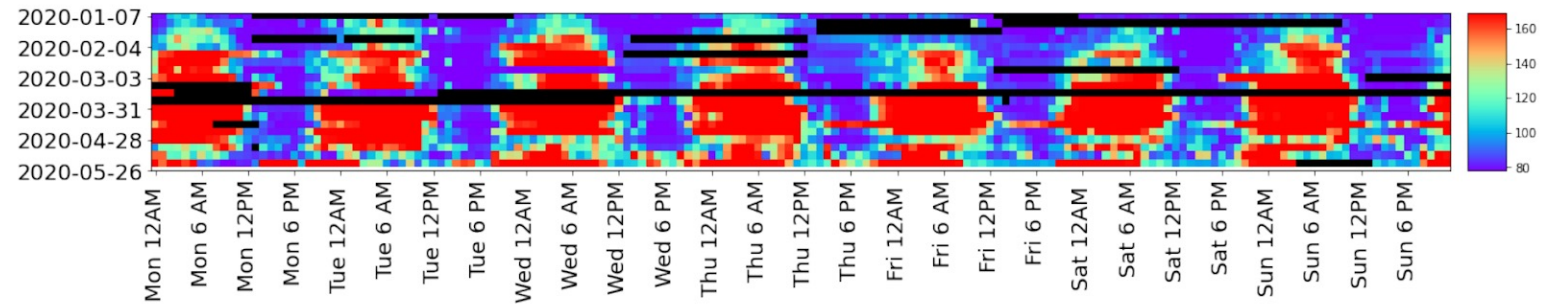


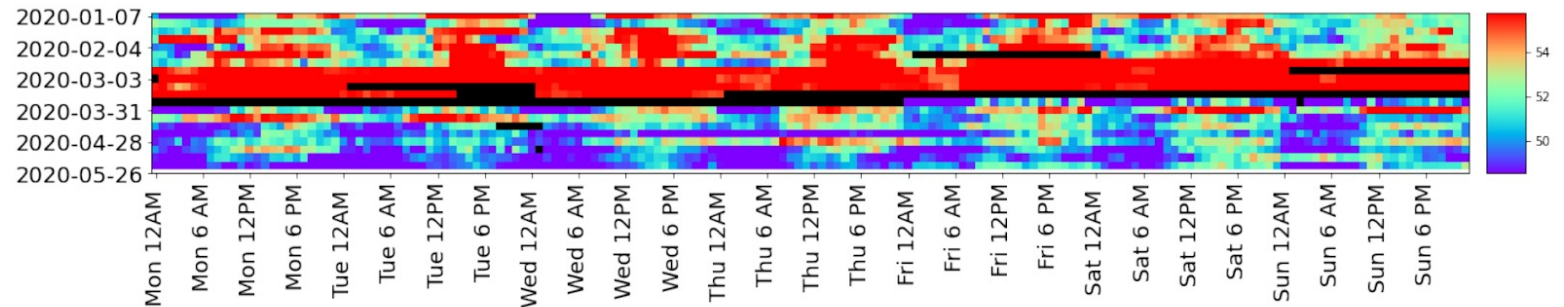
Figure 1: Change in PLT(%) from February to March.

Measurement example

The impact of
COVID-19 on
the
performance
of mobile
networks



(a) RTT [ms] for op1 (IT).

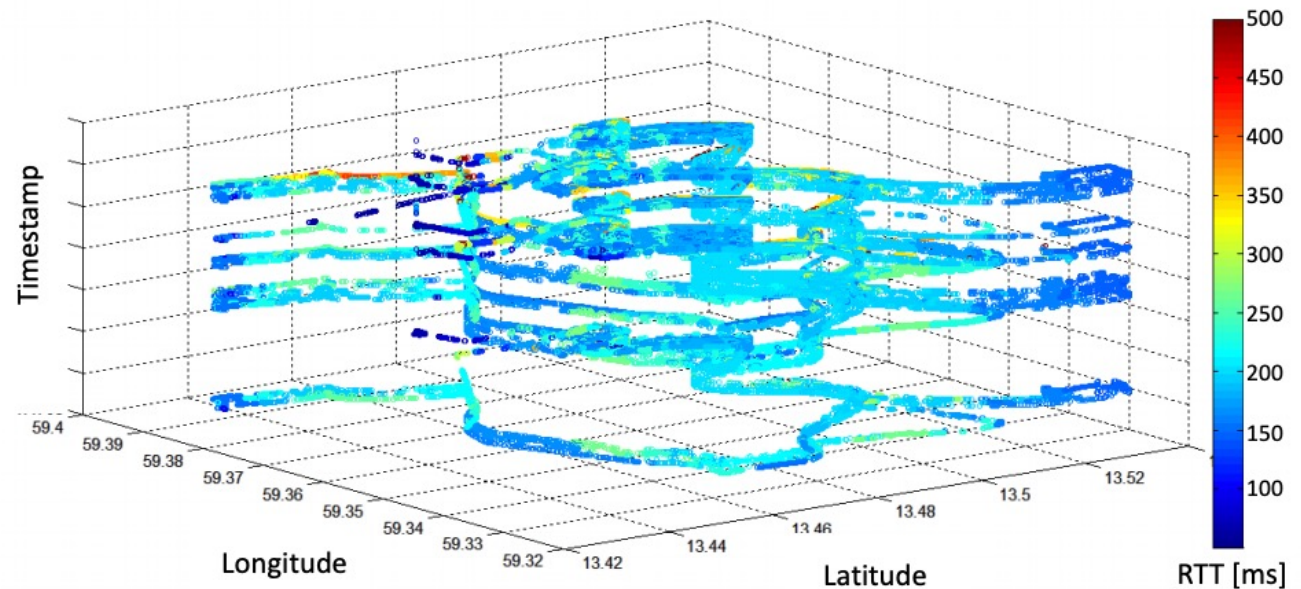


(c) RTT [ms] for op1 (SE).

- 2D information (date vs time of the day) together with heatmap
- Diurnal pattern comparison
- Visual patterns are visible, but details are hidden

Measurement example

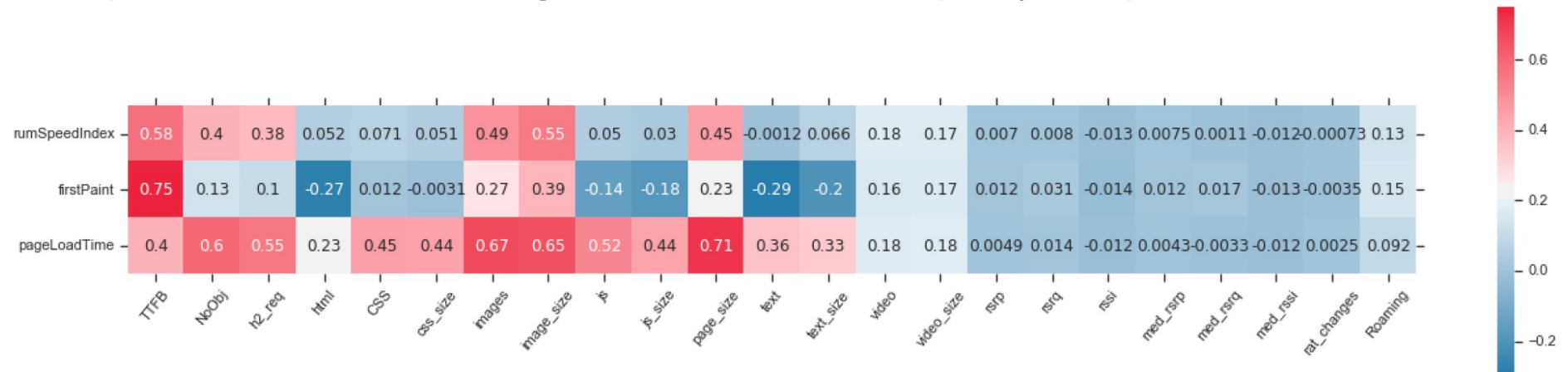
Investigation of latency over multiple trips under mobility for mobile networks



- 3D information together with heatmap
- Multiple laps are shown using the Y-axis offset based on relative timestamps to visually show the different trips
- Potentially have the map to visualize the location as well

Measurement example (with some ML flavour)

The impact of different browsing features on the web quality of experience

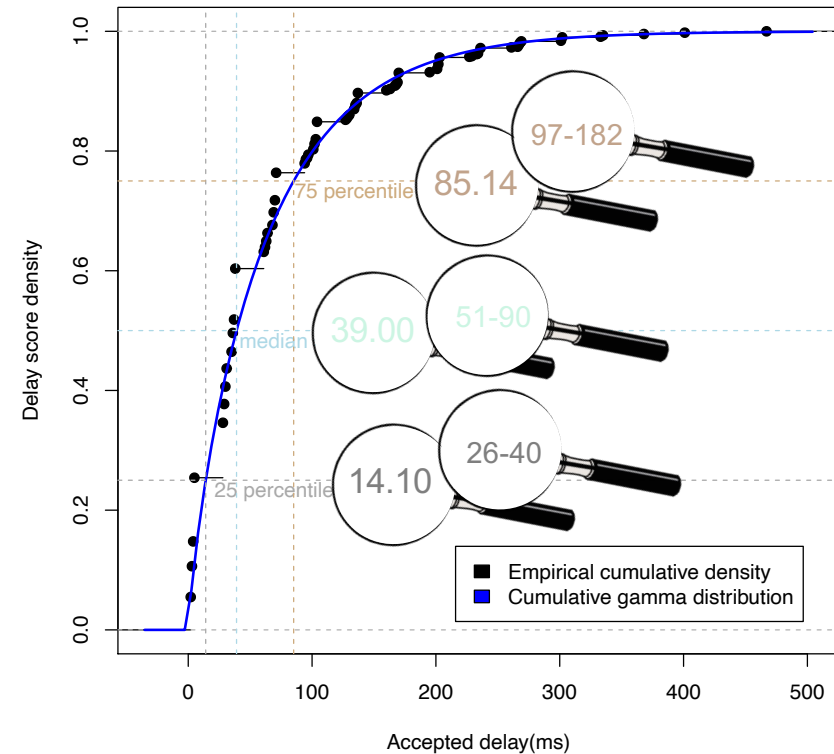


- Spearman's correlation matrix
- Heatmap has negative and positive values

User study example

Hypothesis: users can detect that they are experiencing hand-eye latency below 100ms

- *Cumulative Distribution Function* (CDF) provides the percentage of measurement points up to a given X value (using dots in this case)
- *matched with a function* (here cumulative gamma distribution)
 - better describe the distribution and validate generality
 - create simulations

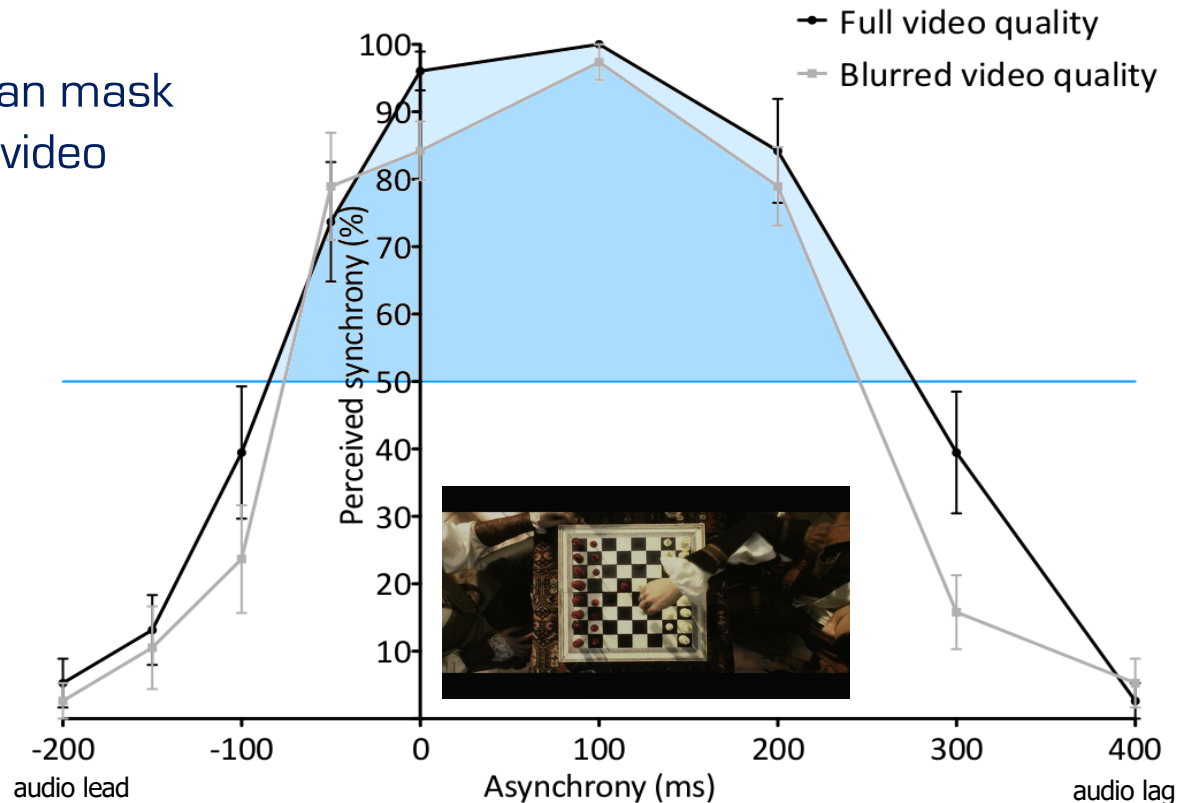


User study example

Hypothesis: poor video quality can mask asynchrony between audio and video streams

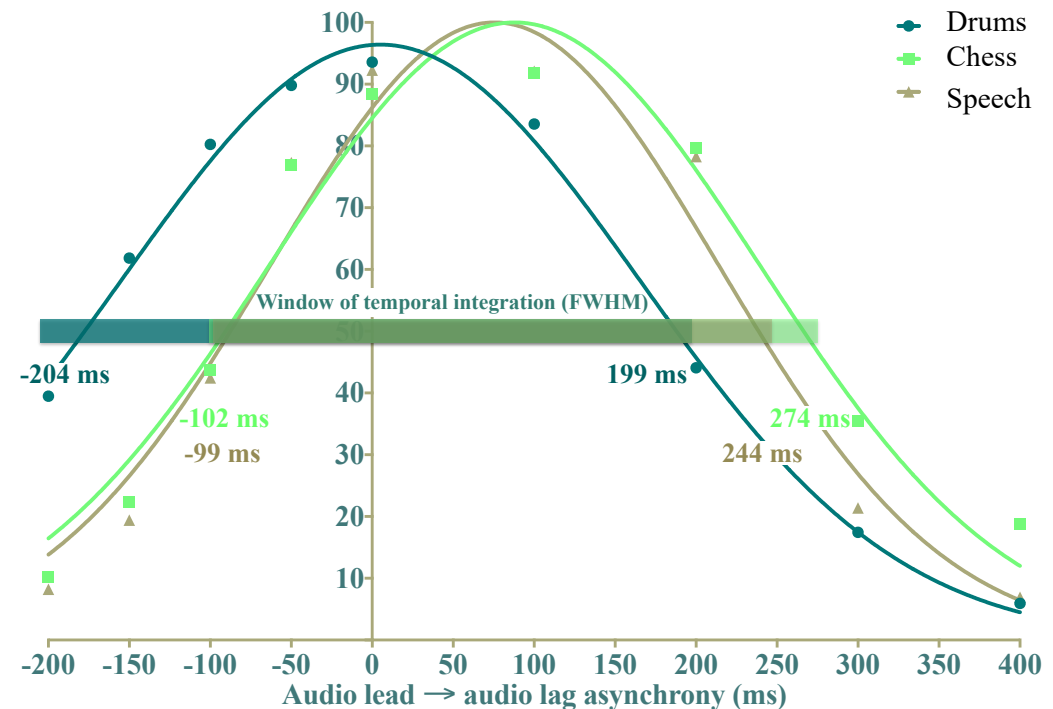
(note: proven wrong)

- note: 3 dimensions in the presentation
- sample points are plotted with error bars
- *highlight color* adds meta-information, here highlighting 50 percent of the study population
- also typical to fit a typical behavior function from samples using linear regression (shown on next slide)



User study example

Pre-study: perception of asynchrony
for different content



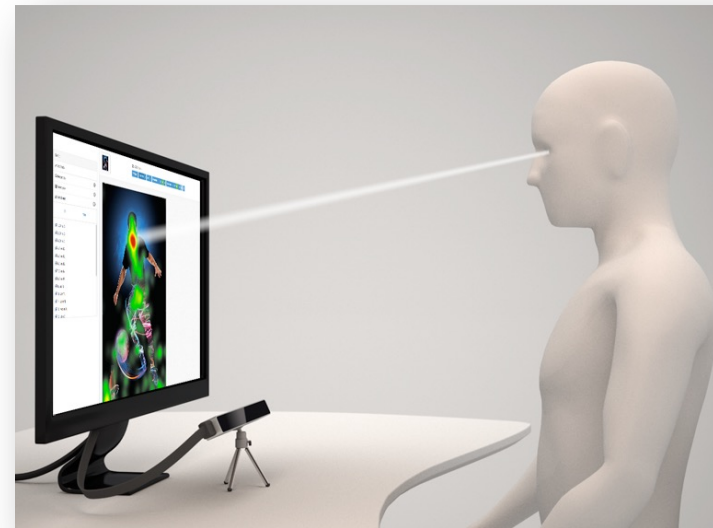
- note: 3 dimensions in the presentation
- *curves* generated from sample *using linear regression*
- horizontal bar adds meta-information, here highlighting 50 percent of the study population
- color is used to distinguish items (4th dimension) and to associate measurements with fitted curves

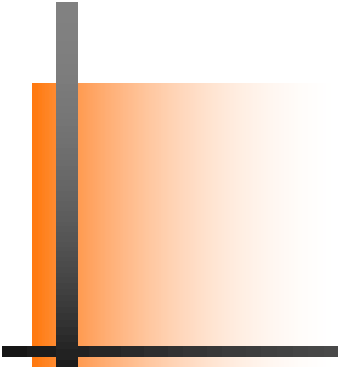
User study example

The influence of semantic relations between visual elements on human attention.

Heatmaps allow a presentation of 2D data accumulated over time.

- 2D input (axes), 1D output (color).
- Can be overlaid over base data.





Performance in distributed systems

Projects

Performance measurement projects

In this course we will give you performance analysis tasks where you will wrestle the tradeoffs, the parameters, the metrics, the methodologies, the analysis and the presentation.

We will

- introduce many of the main concepts of performance analysis
- introduce the topics that form the basis of the graded assignments
- provide example reports of good quality for you to study
- be available on email for guidance and pointers

Performance measurement projects

You must:

- Go to the literature (and the web) for details and resources to help you on the way
- Apply your own skills and judgement in the selection of metrics and methodology
- Justify your choices and try to avoid making random or biased selections
- You will face a lot of tradeoffs and difficult choices. Ask for advice. Communicate!
- This is what researchers and industry professionals are required to do in their practice