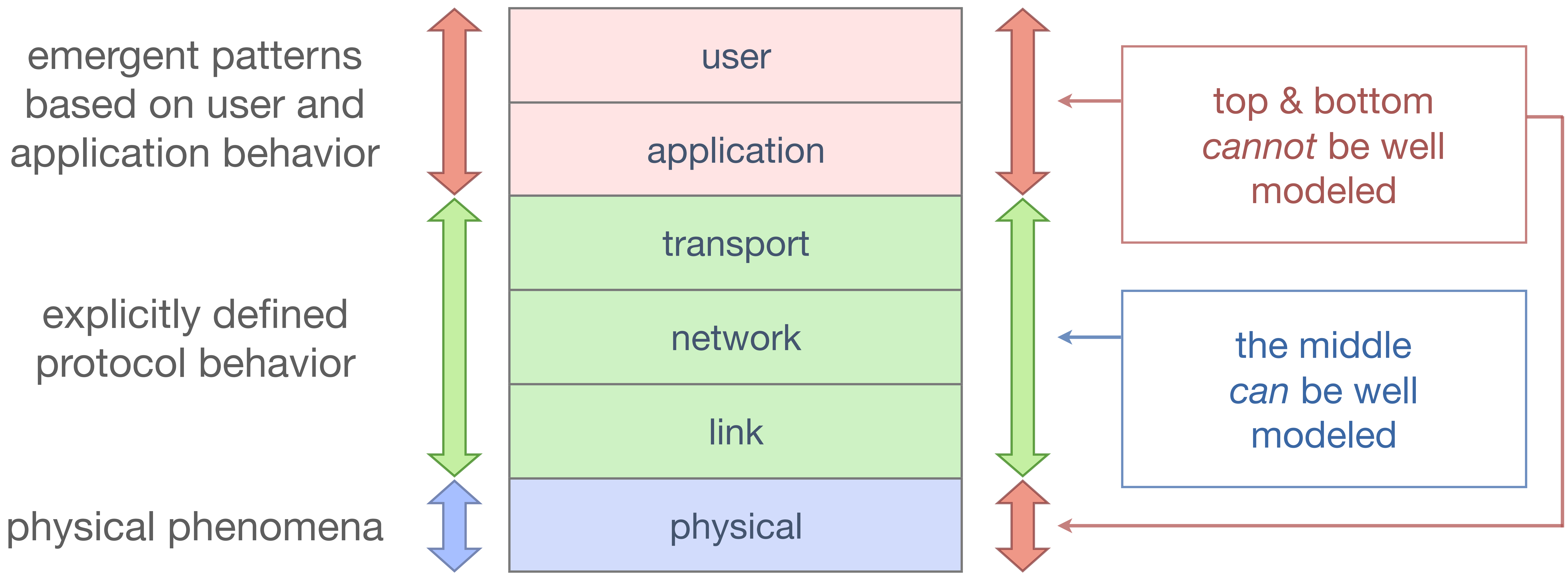
Towards Realistic Models of Wireless Workload

Stefan Karpinski, Elizabeth Belding, Kevin Almeroth
University of California, Santa Barbara

The Realism Deficit

The protocol stack in a (wireless) network simulation:



Experimental test-beds side-step modeling the bottom and middle, but not the top

Catch-22

The problem with modeling traffic behavior...

- someone needs to decide what characteristics of traffic are important
- but if you know that, then you have your model already

Previous approaches have subjectively decided which characteristics are important

- makes it impossible to objectively compare different models
  - each fits own criteria optimally and others' criteria poorly

Our approach cuts through this Gordian knot:

- models are realistic if they induce performance characteristics similar to real traffic
  - no reliance on subjective choice of arbitrary statistical criteria
  - requires simulation to measure realism — we cannot just look traces

Methodology

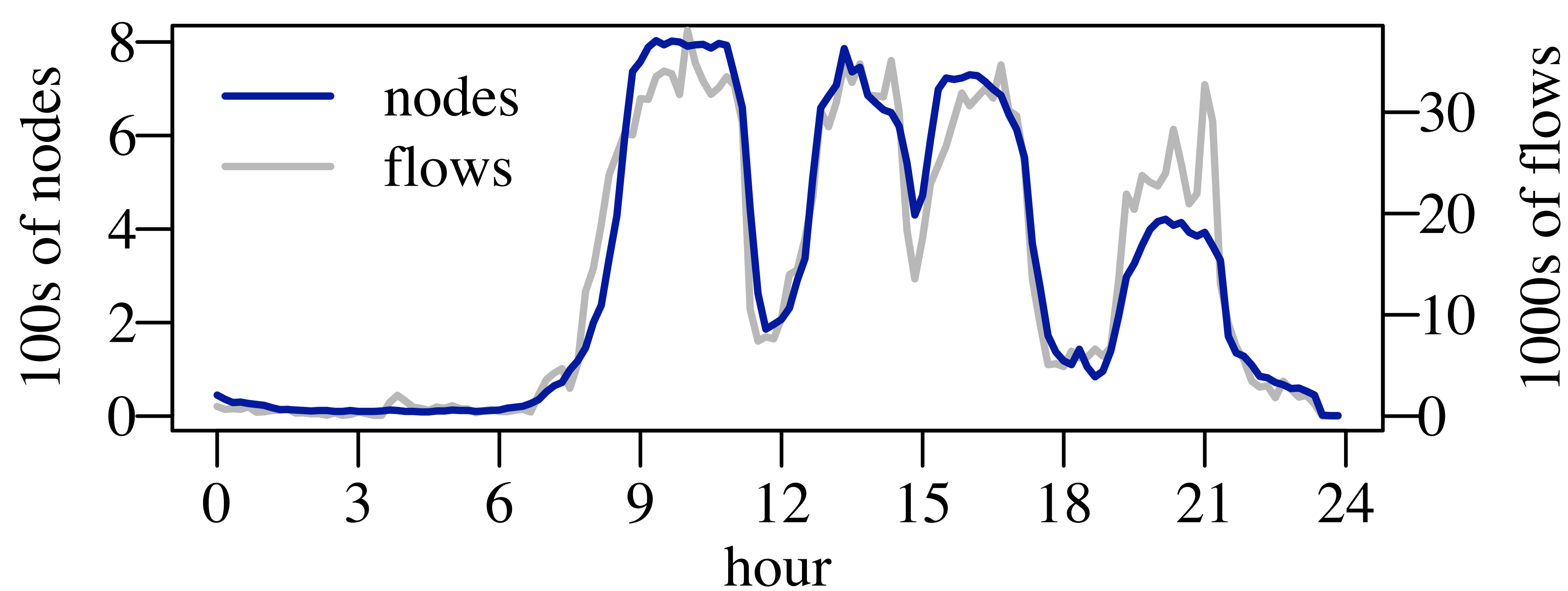
Differential analysis with respect to performance metrics:

- simulate using original trace traffic *vs.* synthetic traffic
  - preserve as many features as the synthetic model will allow
- compare induced values of important performance metrics
  - if they differ drastically, then the model is unrealistic
  - if they are consistently close, then the model is realistic

Methodology

For the comparison, we use a 24-hour trace from the 60th IETF meeting

- large, heavily utilized 802.11g network
  - 18 access points
  - 2082 wireless users
  - 2.1 million flows, 58 million packets, 52 billion bytes



This trace exhibits very broad variety of behaviors

- traffic model accuracy will not depend on any particular network condition

Deconstructing Traffic

First, extract application-level behavior:

- break traffic into flows — defined by protocol, src+dst IPs+ports
- for each flow: sequence of transmissions with data size and timestamp

Behavior consists of three (mostly) independent levels:

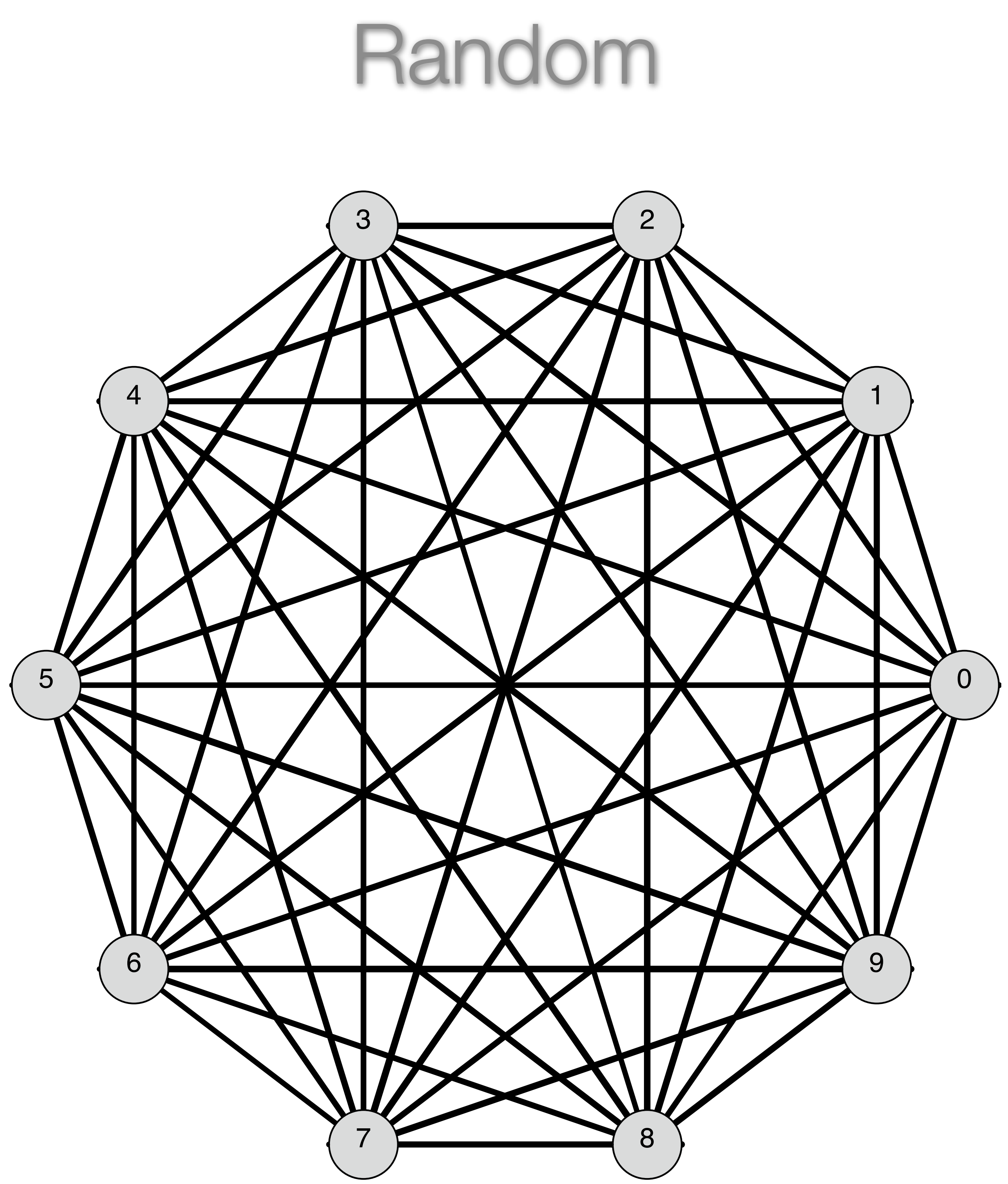
- **flow topology:** how flows are mapped between end-points (nodes) in network
- **flow behavior:** flow duration, total data & packets transmitted
- **packet behavior:** individual sizes of packets & intervals between packet transmissions

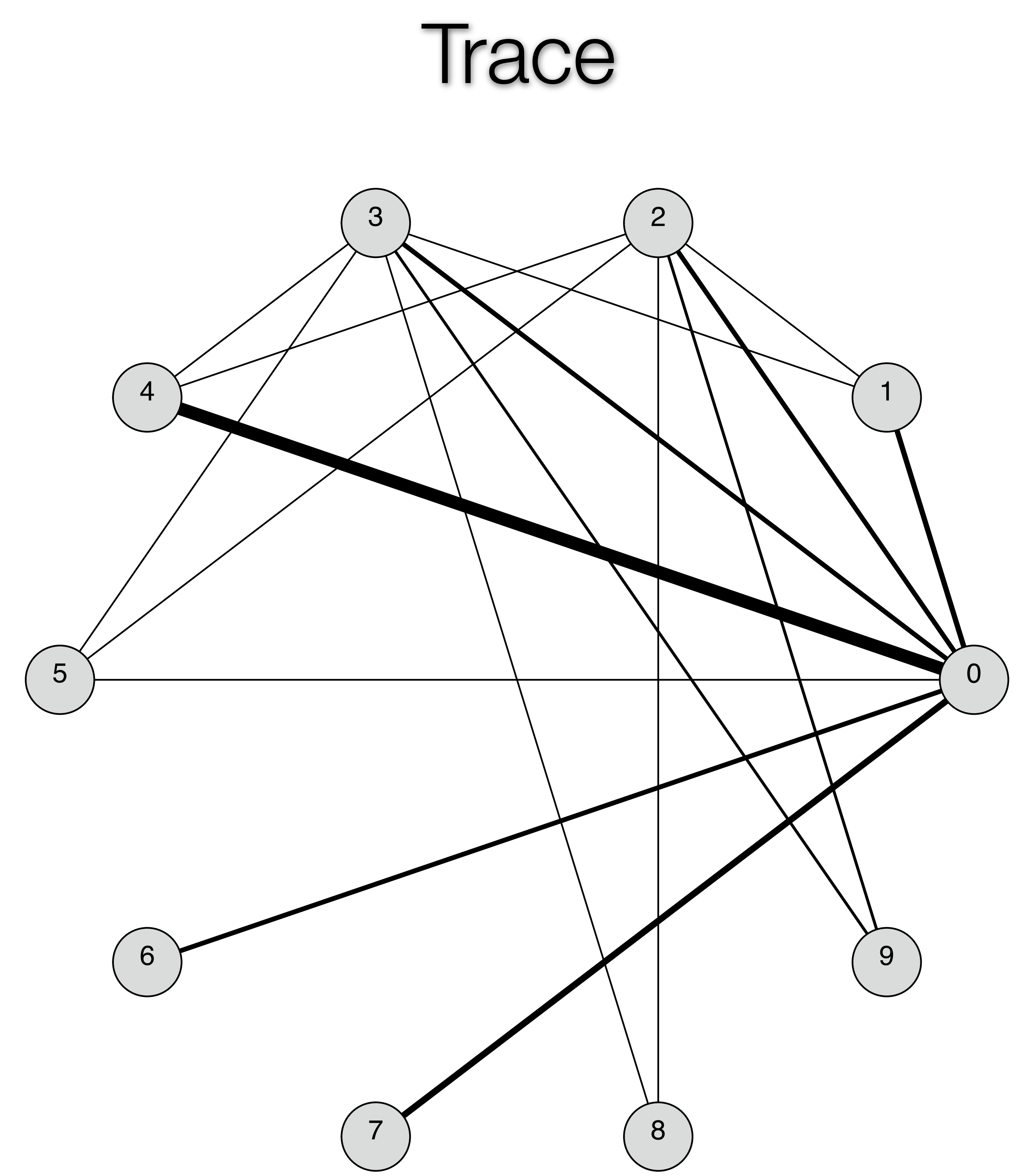
For example, CBR traffic has **uniform** packet behavior

- but most experimental evaluations also have
  - **uniform** flow behavior & **uniform** or **random** end-point topology

Exemplary Behavior: Flow Topology

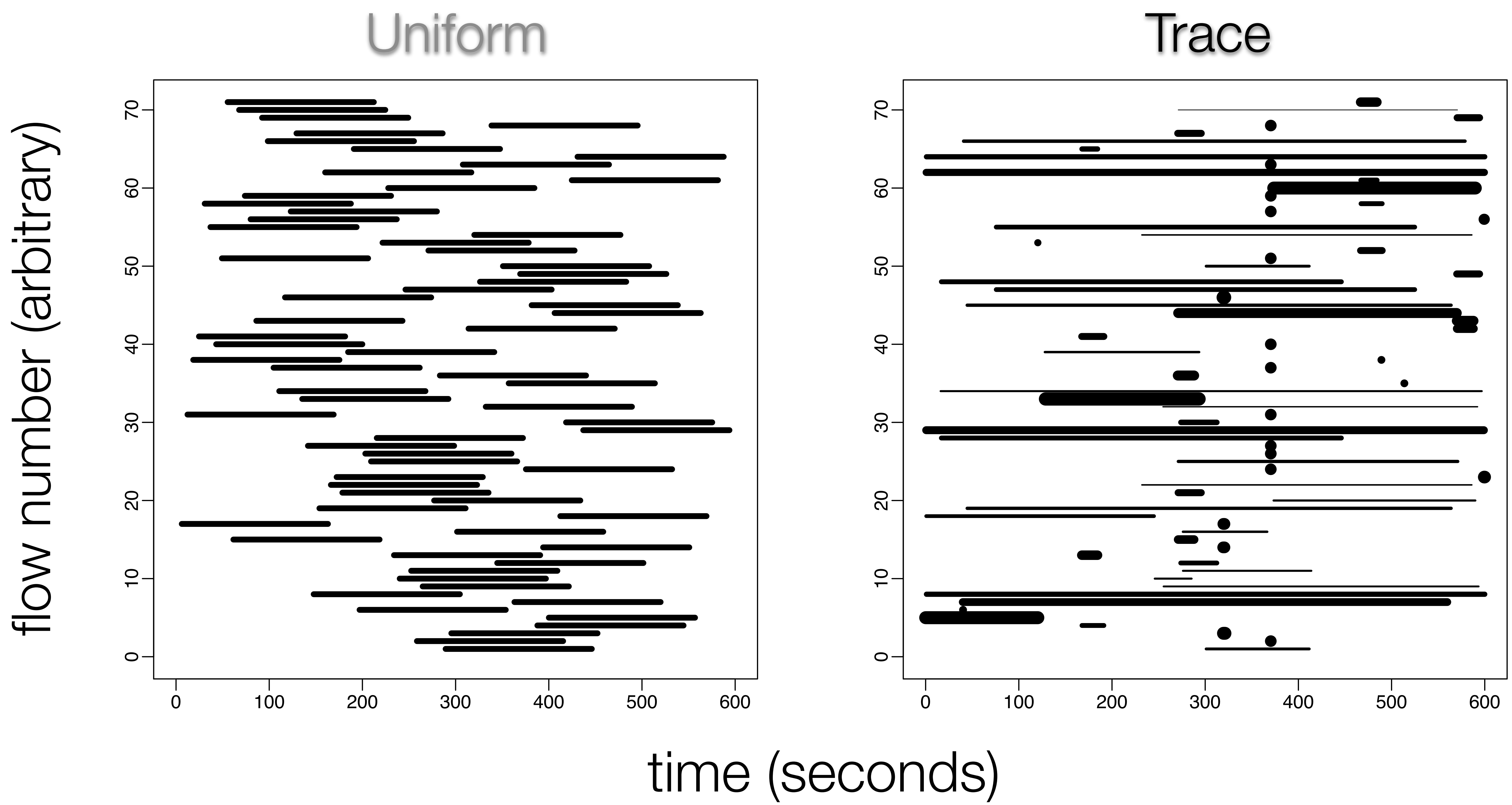
Example scenario: 9 nodes, 756 flows





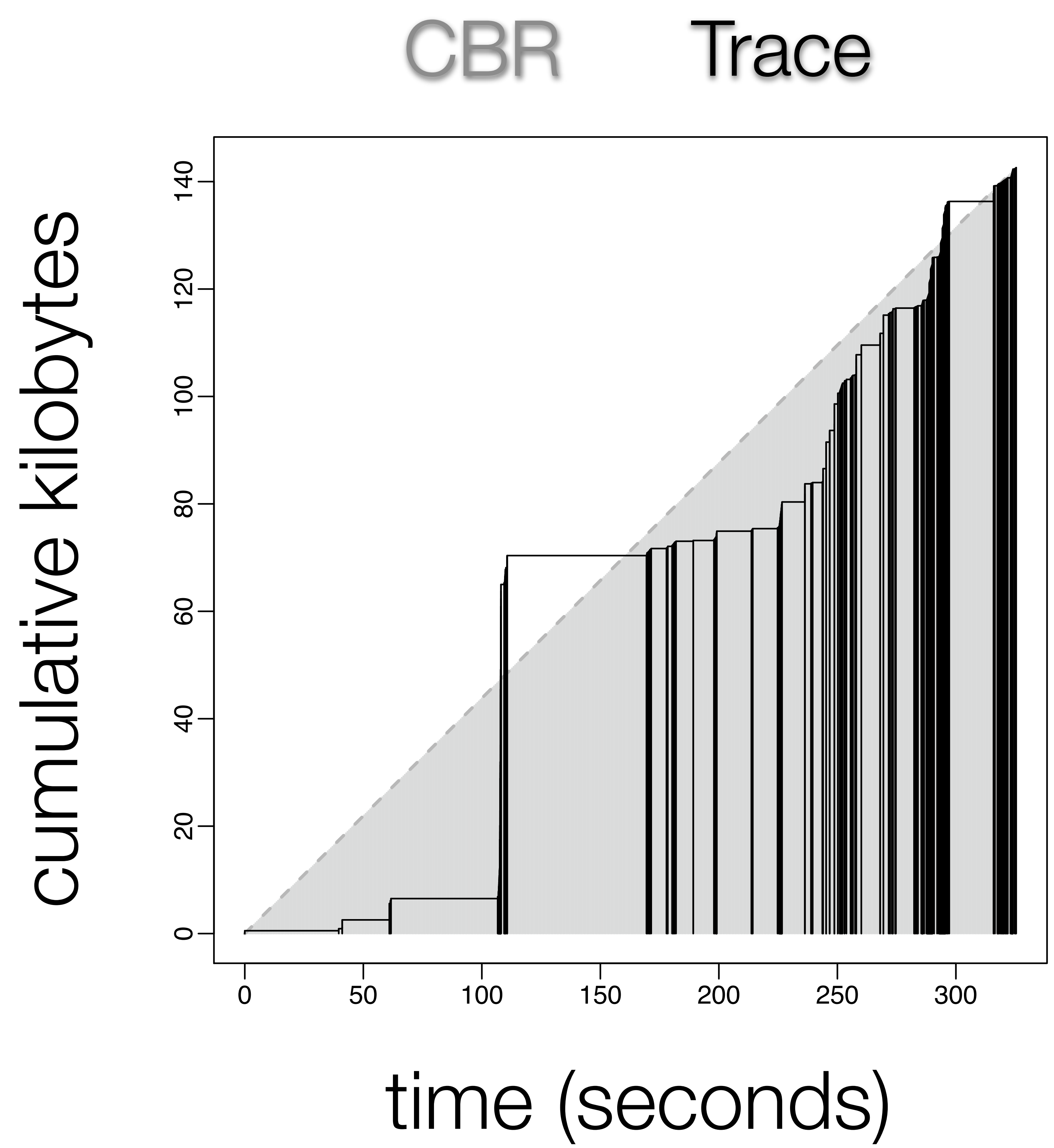
Exemplary Behavior: Flow Behavior

Example scenario: 71 flows in 10 minutes



Exemplary Behavior: Packet Behavior

Example flow: 512 packets, 143 kilobytes, over 325.43 seconds



Traffic Models

Commonly used experimental models:

- **RandomUniformCBR**
- **UniformCBR**

Models changing only packet behavior:

- **SampleTime**
- **SampleSize**
- **SampleTimeSize**

Models changing only flow (end-point) topology:

- **ShuffleEndPoints**
- **RandomEndPoint**
- **SampleEndPoint**
- **SampleEndPointsJoint**

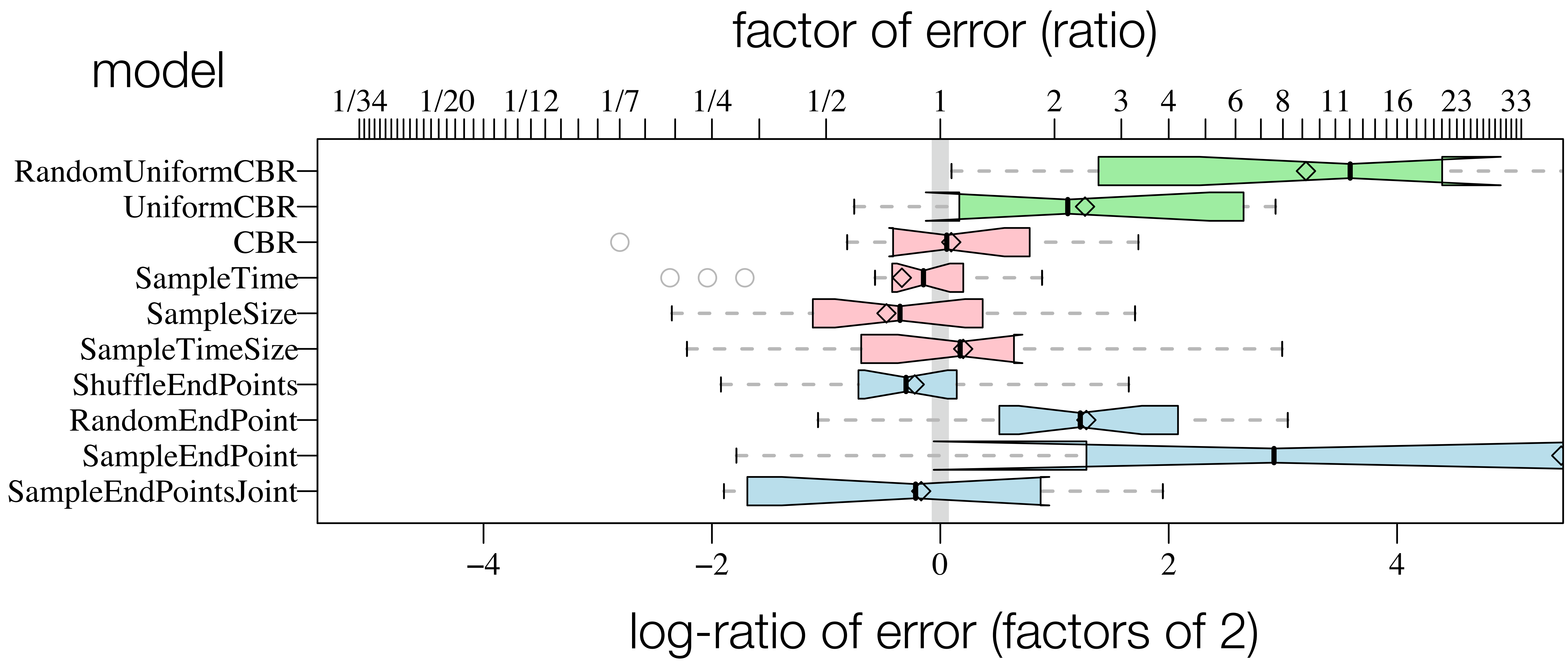
Performance Metrics

A selection of important wireless metrics

- Application Layer
  - **end-to-end delay**
  - **received throughput**
- Network Layer
  - **AODV control overhead** (RREQ/RREP/RERR)
- Link Layer (Medium Access)
  - **packet retransmission rate**
    - link retransmissions per application data unit initiated

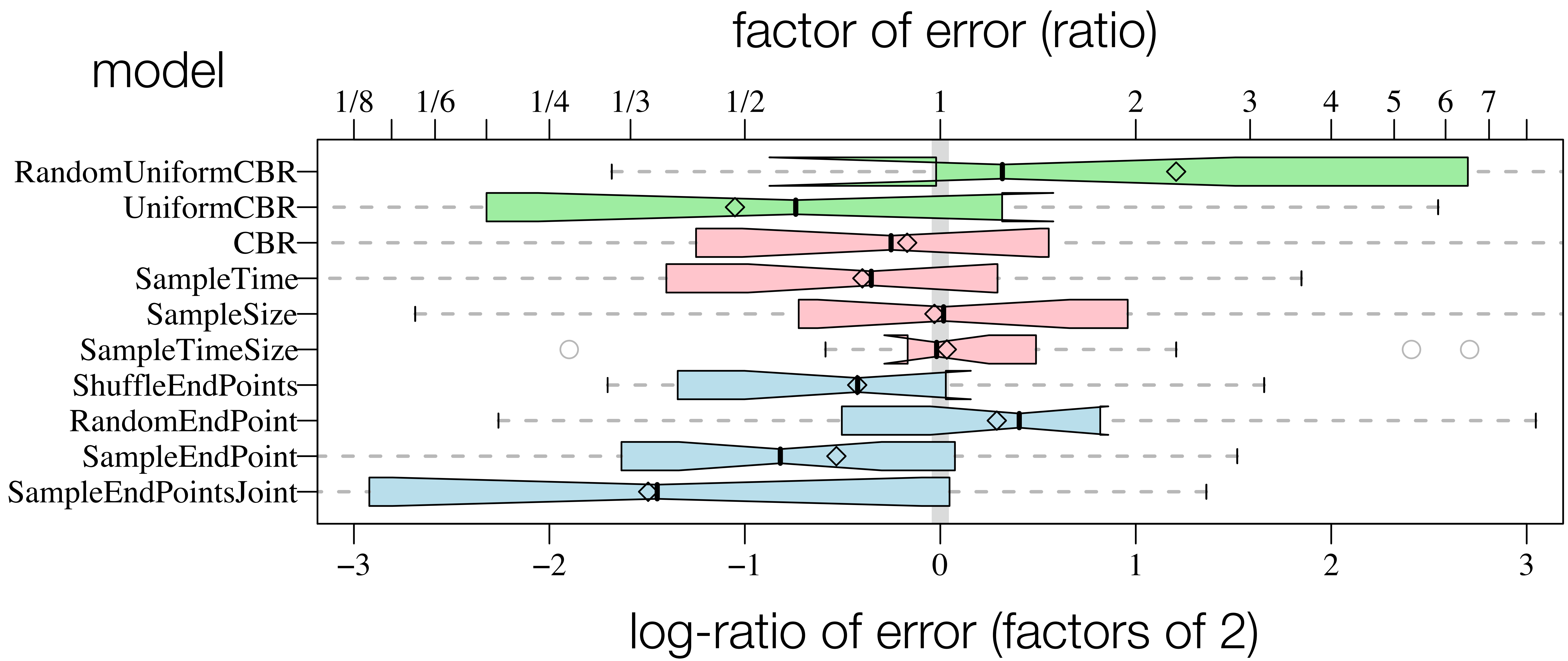
Results: Received Application Throughput

box-and-whisker plots summarize the distributions of errors



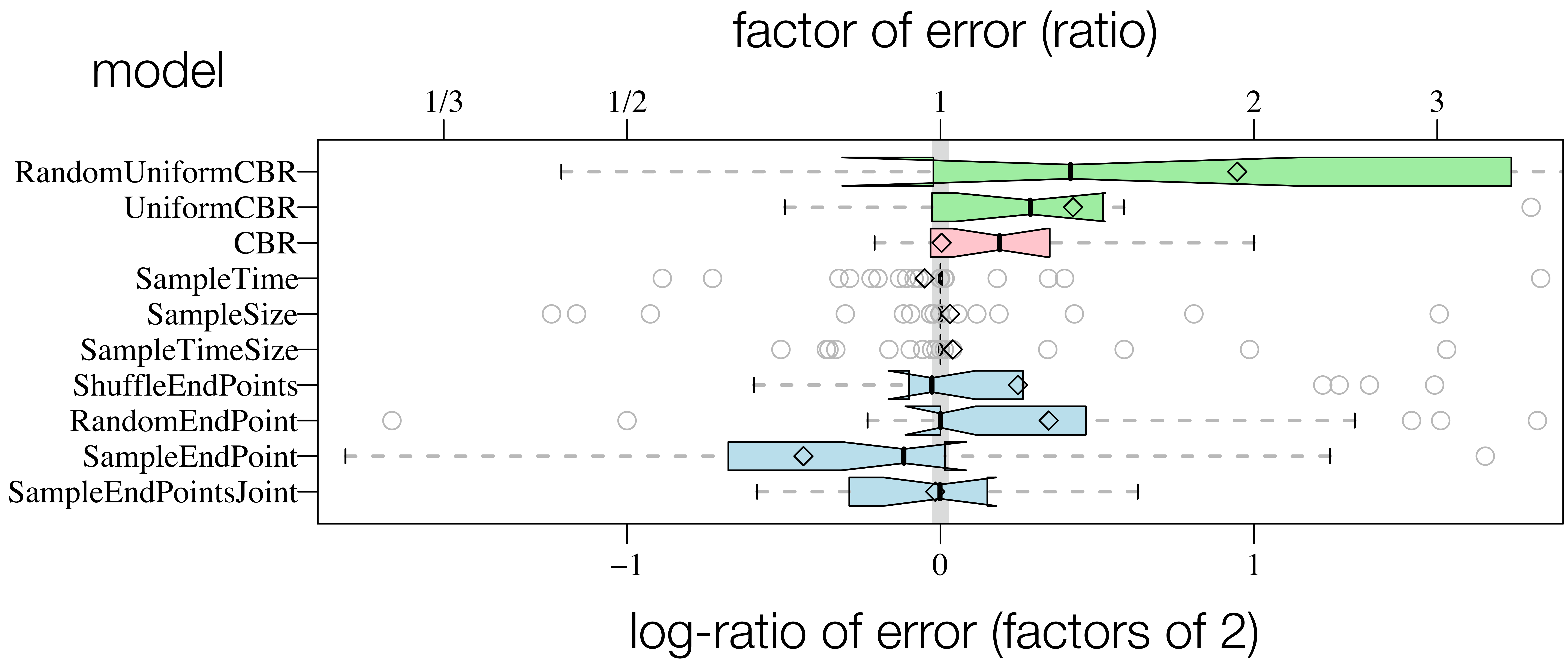
Results: End-to-End Application Delay

box-and-whisker plots summarize the distributions of errors



Results: Network Control Overhead

box-and-whisker plots summarize the distributions of errors



Conclusions

Immediate implications:

- **packet behavior does not require time-series modeling**
  - sampling time and inter-packet intervals independently is sufficiently realistic
    - this implies that flow behavior is characterized by:
      1. duration
      2. distribution of packet sizes
      3. distribution of inter-packet intervals
- **flow topology requires more complex models**
- **common traffic models do not accurately predict real-world performance**

Greater contribution:

- an objective way of evaluating traffic model realism

Questions?

Contact information:

- Stefan Karpinski 〈sgk@cs.ucsb.edu〉

  http://cs.ucsb.edu/~sgk

Research group:

- Moment Lab
  Department of Computer Science
  University of California, Santa Barbara

Please don't hesitate to email me with comments or questions about my work!