Characterizing the Transport Behaviour of the Short Message Service

Earl Oliver
David R. Cheriton School of Computer Science
University of Waterloo

MobiSys 2010
June 17, 2010

Takeaways

Takeaways

- Studied the channel characteristics of SMS from the perspective of mobile devices sending bursts of messages (bulk senders)

Takeaways

- Studied the channel characteristics of SMS from the perspective of mobile devices sending bursts of messages (bulk senders)
- Design and implement an efficient and reliable SMS-based data transport protocol

SMS in mobile systems

- Prolific use of SMS

SMS in mobile systems

- Prolific use of SMS
  - ubiquity

SMS in mobile systems

- Prolific use of SMS
  - ubiquity
  - reliability

SMS in mobile systems

- Prolific use of SMS
  - ubiquity
  - reliability
  - low-cost*

SMS in mobile systems

- Prolific use of SMS
  - ubiquity
  - reliability
  - low-cost*
  - convenient APIs

SMS in mobile systems

- Prolific use of SMS
  - ubiquity
  - reliability
  - low-cost*
  - convenient APIs
  - endpoint addressability

Existing use of SMS

- Typically constrained to single messages

Existing use of SMS

- Typically constrained to single messages
- Stop-and-wait protocol used to send larger amounts of data

Existing use of SMS

- Typically constrained to single messages
- Stop-and-wait protocol used to send larger amounts of data
- Easy to implement

Existing use of SMS

- Typically constrained to single messages
- Stop-and-wait protocol used to send larger amounts of data
- Easy to implement
- Network communication often a secondary concern

Alternatives to SMS

Alternatives to SMS

- Cellular data services

Alternatives to SMS

- Cellular data services
  - low-latency, high data rate

Alternatives to SMS

- Cellular data services
  - low-latency, high data rate
    - end-points behind NAT

Alternatives to SMS

- Cellular data services
  - low-latency, high data rate
    - end-points behind NAT
  - sparsely deployed in developing regions

Alternatives to SMS

- Cellular data services
  - low-latency, high data rate
    - end-points behind NAT
    - sparsely deployed in developing regions
- MMS, EMS

Alternatives to SMS

- Cellular data services
  - low-latency, high data rate
    - end-points behind NAT
      - sparsely deployed in developing regions
- MMS, EMS
  - can transfer large amounts of data

Alternatives to SMS

- Cellular data services
  - low-latency, high data rate
    - end-points behind NAT
      - sparsely deployed in developing regions
- MMS, EMS
  - can transfer large amounts of data
    - poorly supported and not universally available

Goal

- Develop a portable library to exchange *large* mounts of data over SMS

Goal

- Develop a portable library to exchange *larg*e mounts of data over SMS
- How do we e*fficiently* maximize the use of the SMS channel?

Goal

- Develop a portable library to exchange *larg*e mounts of data over SMS
- How do we e*fficiently* maximize the use of the SMS channel?
  - Minimize message overhead

Goal

- Develop a portable library to exchange *larg*e mounts of data over SMS
- How do we e*fficiently* maximize the use of the SMS channel?
  - Minimize message overhead
  - Maximize data throughput

Outline

- Channel characterization
- Transport protocol design
- Evaluation

Previous work

Service center

Source

Destination

- Zerfos (IMC 2006)

Previous work

Delay and loss rate

Service center

Source

Destination

- Zerfos (IMC 2006)

Our work

Service center

Source

Destination

- We examine channel properties from the **perspective of mobile devices** using the service as ***mass message senders***

Our work

Transmission time, delay, loss rate, message reordering

Source

Black box

Transmission time, delay, loss rate, message reordering

Destination

- We examine channel properties from the **perspective of mobile devices** using the service as ***mass message senders***

Sources of variability

Base station and controller

Service center

Base station and controller

Source

Destination

Sources of variability

Base station and controller

Service center

Base station and controller

Dedicated channel request

Source

Destination

Sources of variability

Base station and controller

Service center

Base station and controller

Dedicated channel request

Source

Destination

Sources of variability

Service center

Base station and controller

Source

Destination

Sources of variability

Base station and controller

Service center

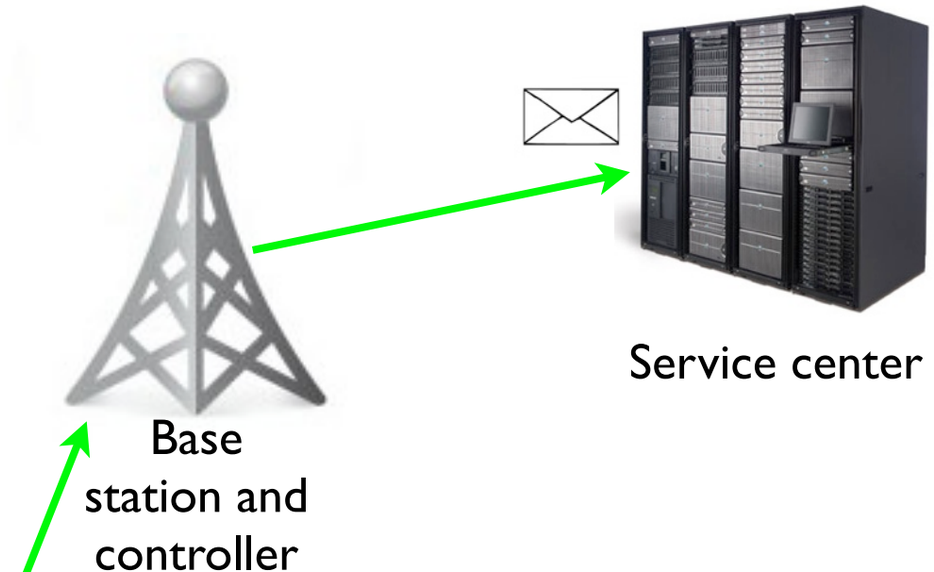
Base station and controller

Channel allocation

Source

Destination

Sources of variability



Message
transmit

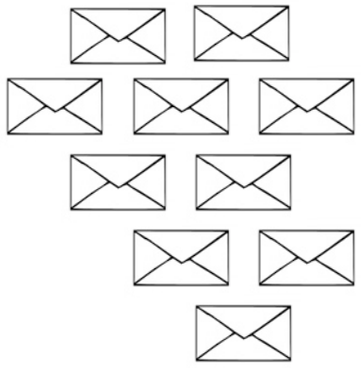
Source

Destination

Sources of variability

Base station and controller

Service center



Base station and controller

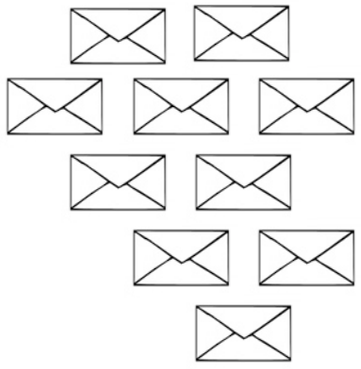
Source

Destination

Sources of variability

Base station and controller

Service center



Base station and controller

Source

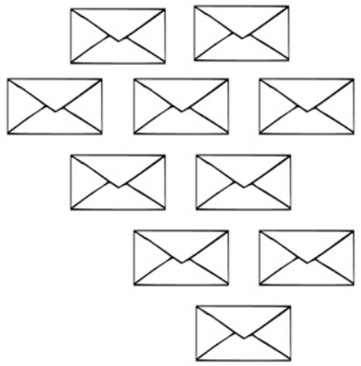
Transmit time

Destination

Sources of variability

Base station and controller

Service center



Base station and controller

Page and dedicated channel request

Source

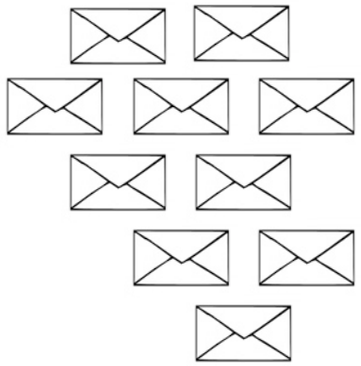
Transmit time

Destination

Sources of variability

Base station and controller

Service center



Base station and controller

Page and dedicated channel request

Source

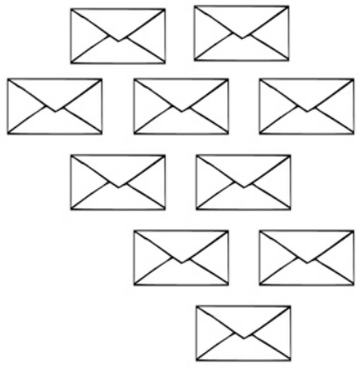
Destination

Transmit time

Sources of variability

Base station and controller

Service center



Base station and controller

Channel allocation

Source

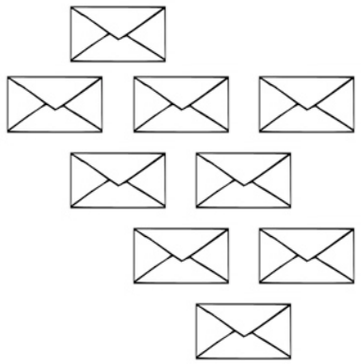
Transmit time

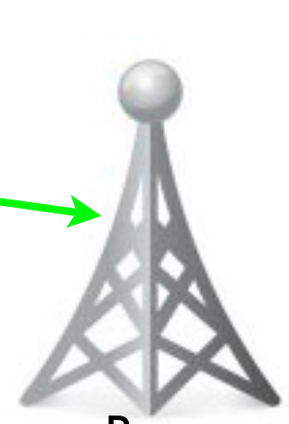
Destination

Sources of variability

Base station and controller

Service center





Base station and controller

Source

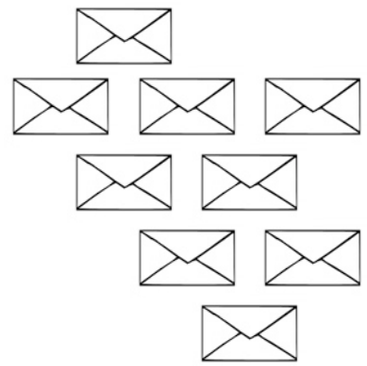
Transmit time

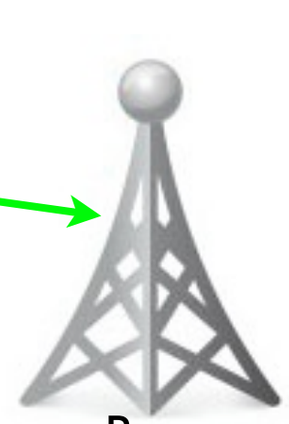
Destination

Sources of variability

Base station and controller

Service center





Base station and controller

Source

Destination

Transmit time

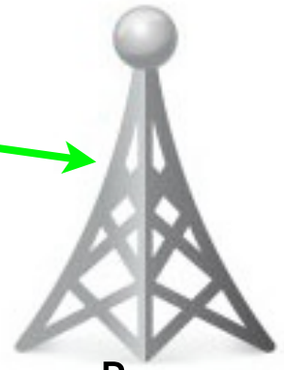
Delay

Sources of variability

Base station and controller

Service center

Loss rate, reordering rate



Base station and controller

Source

Destination

Transmit time

Delay

Experiment

Channel characteristics

- Transmission time
- Delay
- Loss rate
- Reordering rate

X

Experimental parameters

- Intra-burst time
- Transmission index
- Network interface
- Time-of-day

Experiment

Intra-burst time

Transmission index

Network interface

Time-of-day

Experiment

Intra-burst time

Transmission index

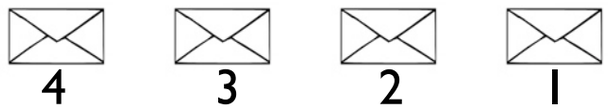
Network interface

Time-of-day

Experiment

Intra-burst time

Transmission index

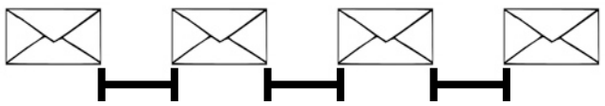


Network interface

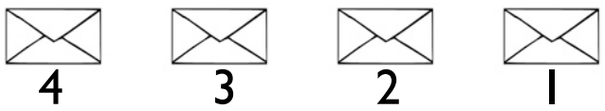
Time-of-day

Experiment

Intra-burst time



Transmission index

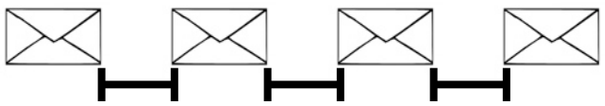


Network interface

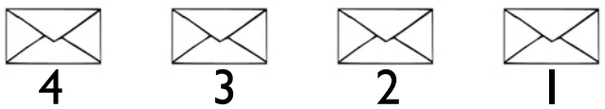
Time-of-day

Experiment

Intra-burst time



Transmission index

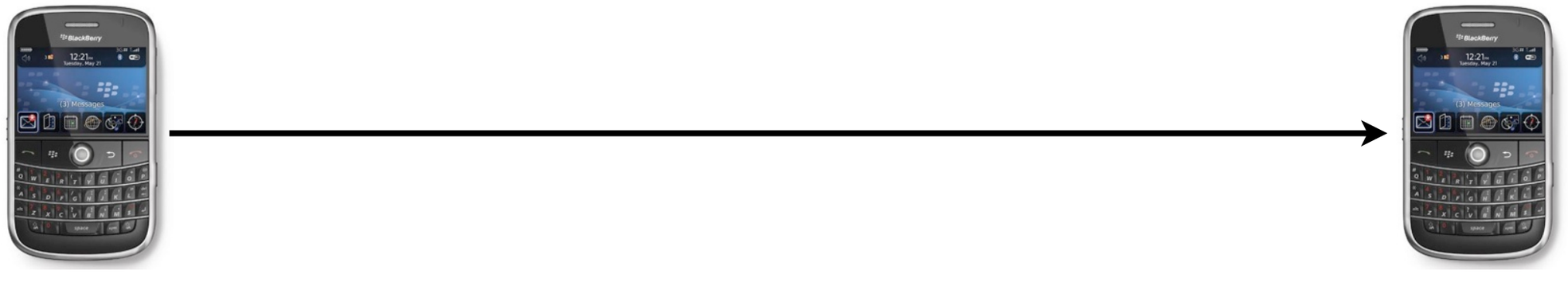


Network interface

Time-of-day

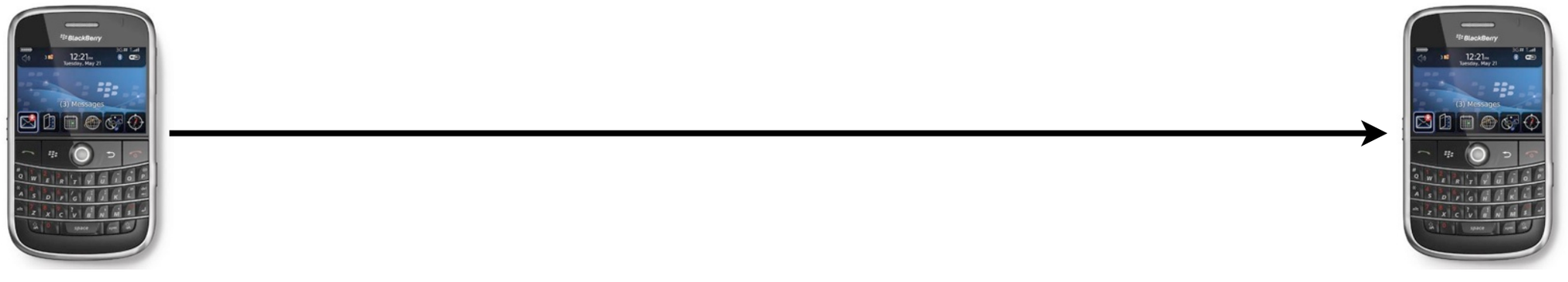
Methodology

Methodology



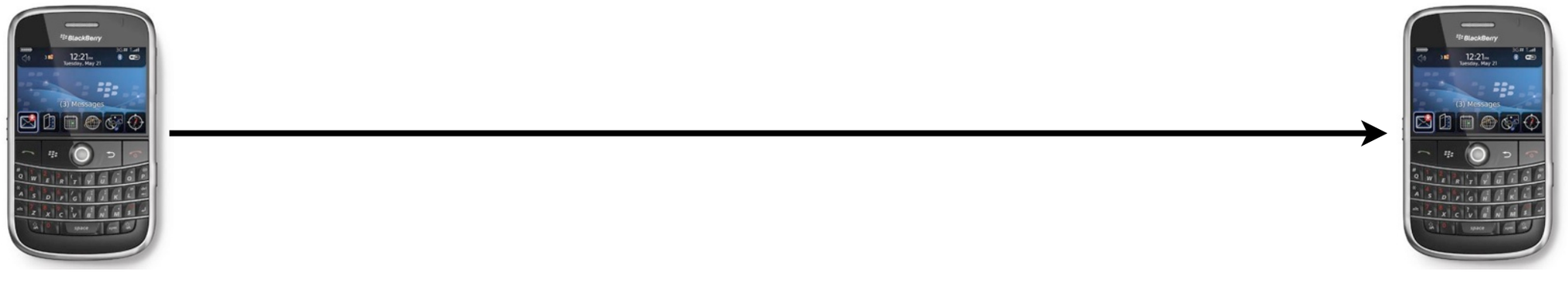
- Transfer repeated bursts of messages between pairs of stationary BlackBerrys and USB tethered Nokia cell phones

Methodology



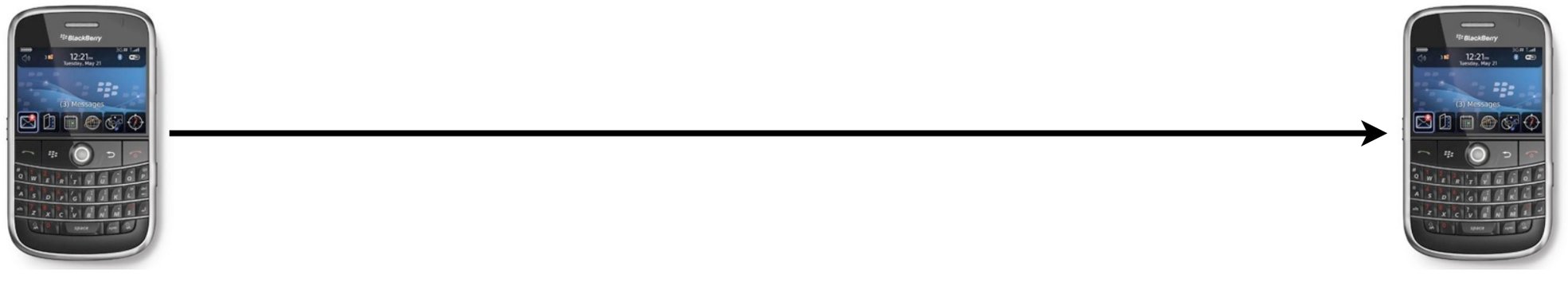
- Transfer repeated bursts of messages between pairs of stationary BlackBerrys and USB tethered Nokia cell phones
  - Major Canadian GSM provider (Rogers)

Methodology



- Transfer repeated bursts of messages between pairs of stationary BlackBerrys and USB tethered Nokia cell phones
  - Major Canadian GSM provider (Rogers)
  - 80,000 messages

Methodology



- Transfer repeated bursts of messages between pairs of stationary BlackBerrys and USB tethered Nokia cell phones
  - Major Canadian GSM provider (Rogers)
  - 80,000 messages
- See paper for details of variable isolation and experiment methodology

Summary of key results

Finding	Design impact
Transmission time	

Summary of key results

Finding	Design impact
Transmission time	
Transmission time is independent of intra-burst time, index, time-of-day.	Protocol does not need to regulate message transmission.

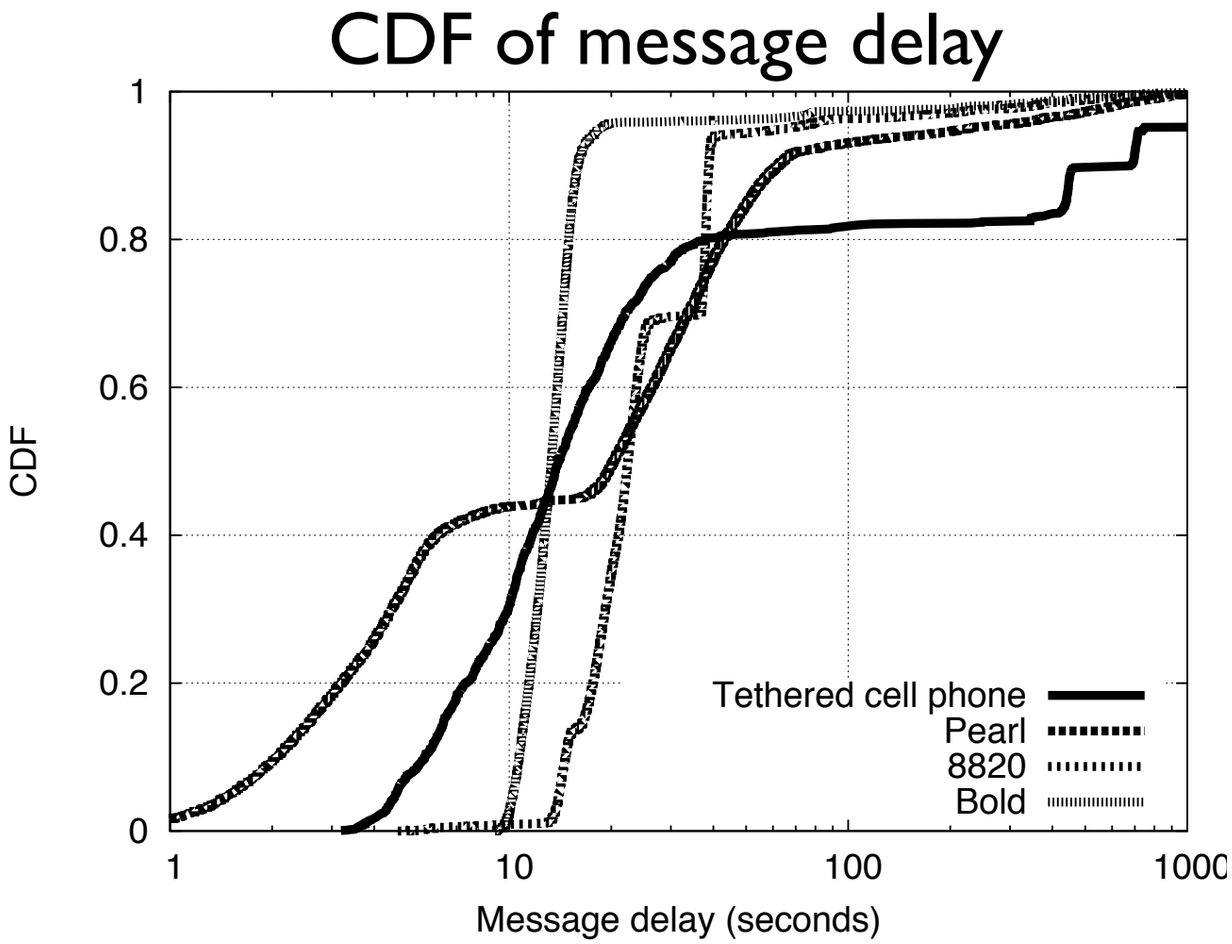
Summary of key results

Finding	Design impact
Transmission time	
Transmission time is independent of intra-burst time, index, time-of-day.	Protocol does not need to regulate message transmission.
Message loss	
Loss rate independent of experiment parameters. (0 - 4%)	Messages are more likely to be highly delayed and reordered than lost.

Finding	Design impact
Message reordering rate	
Messages are reordered at a mean rate of 3.4%.	The protocol must be tolerant of message reordering.

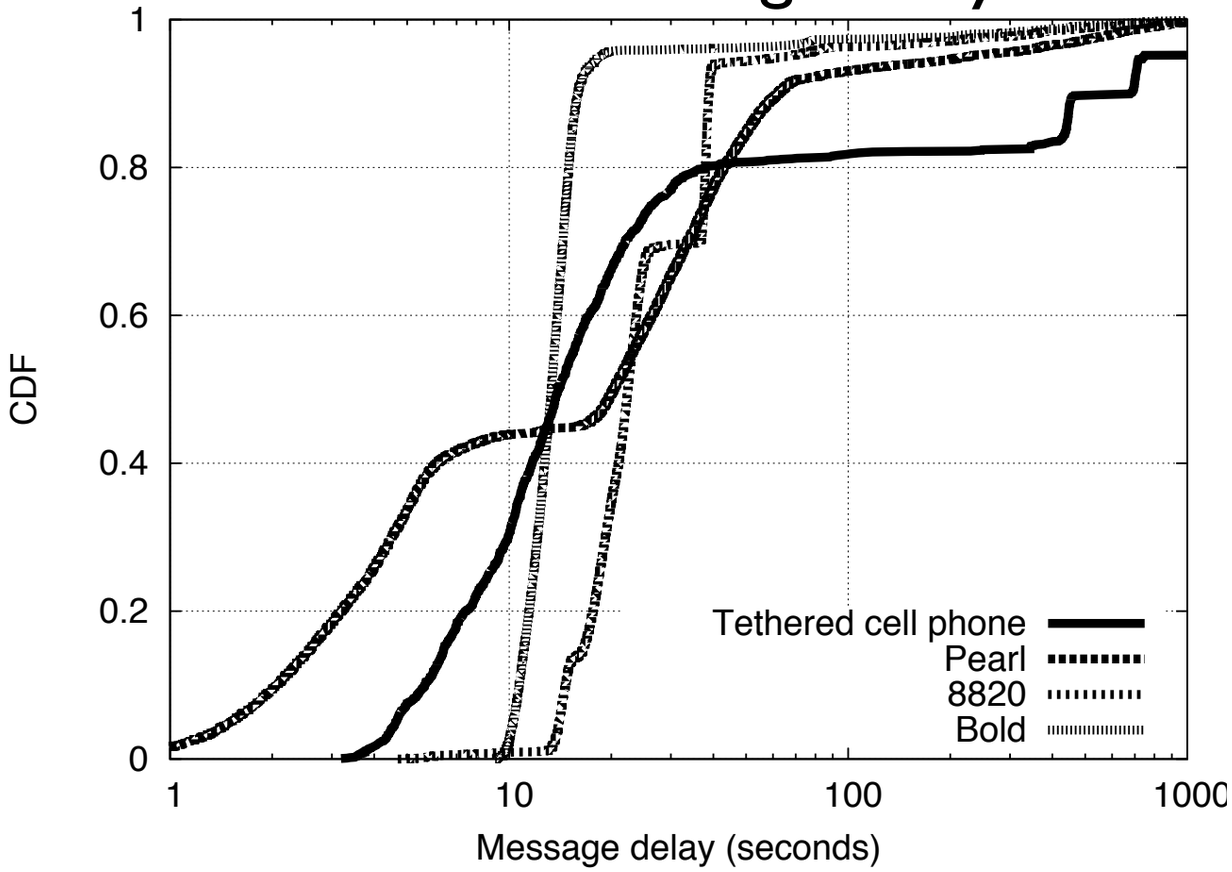
Finding	Design impact
Message delay	
Delay is independent of intra-burst size.	Protocol does not need to regulate message transmission.

Finding	Design impact
Message delay	
Delay is independent of intra-burst size.	Protocol does not need to regulate message transmission.
Delay is highly correlated with the network interface and time-of-day.	The protocol must tolerate highly variable delay.

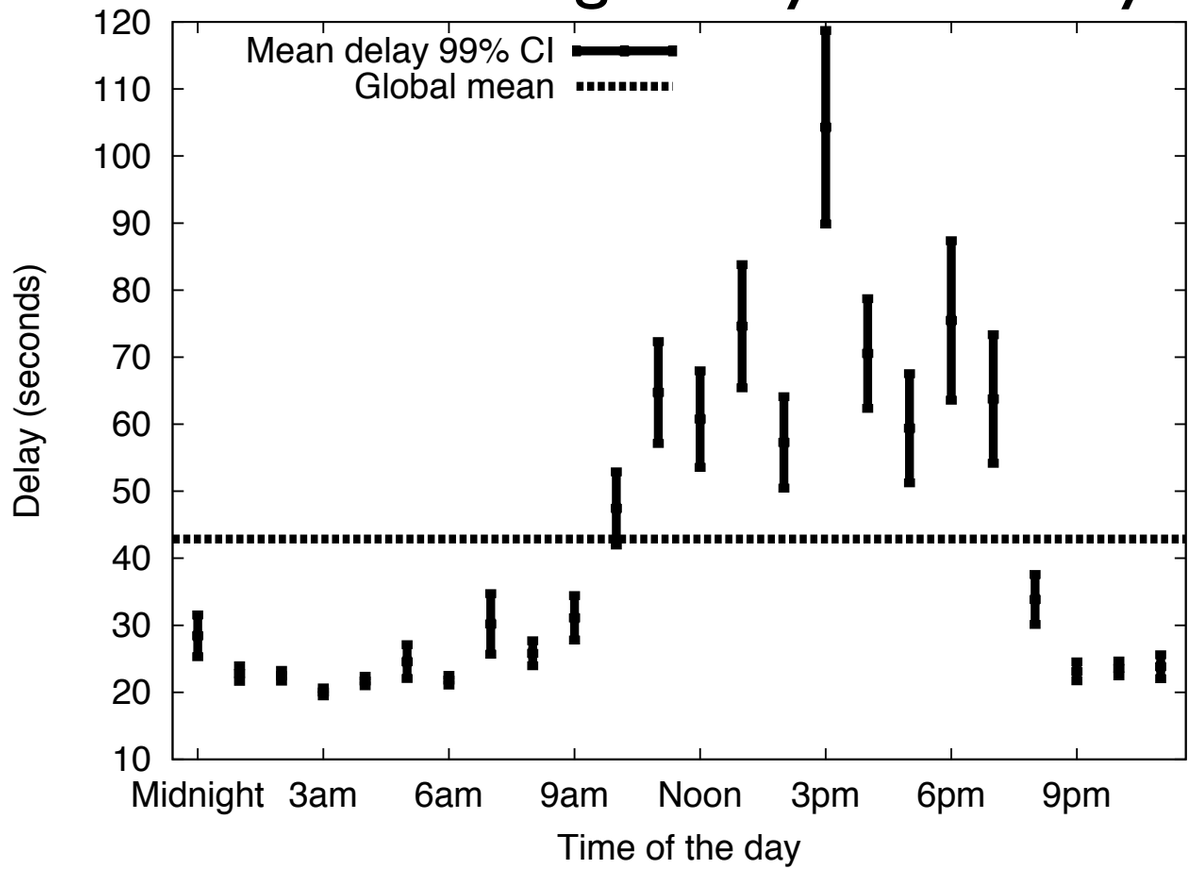


Delay is highly correlated with the network interface and time-of-day.	The protocol must tolerate highly variable delay.

CDF of message delay

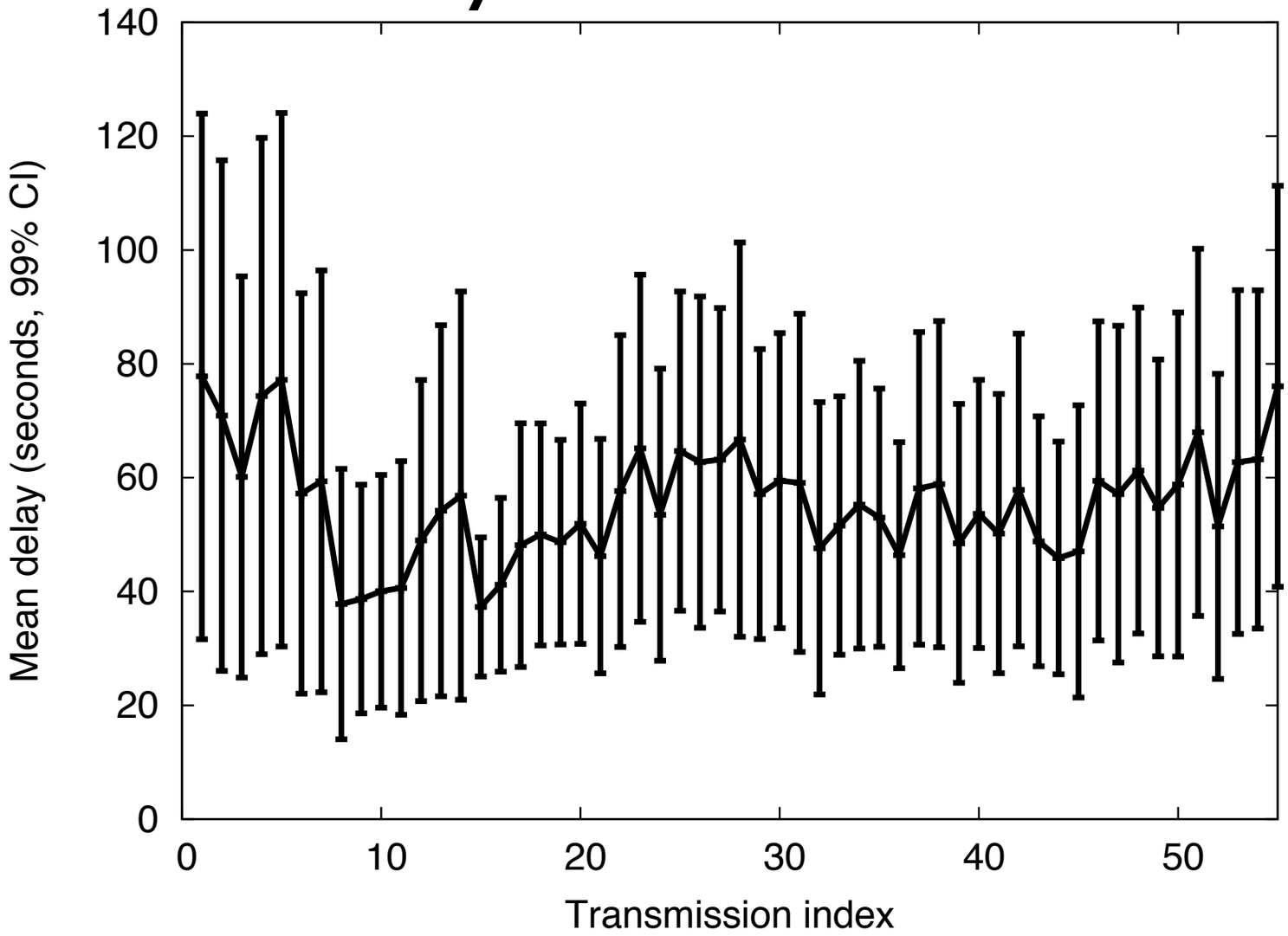


Mean message delay over a day



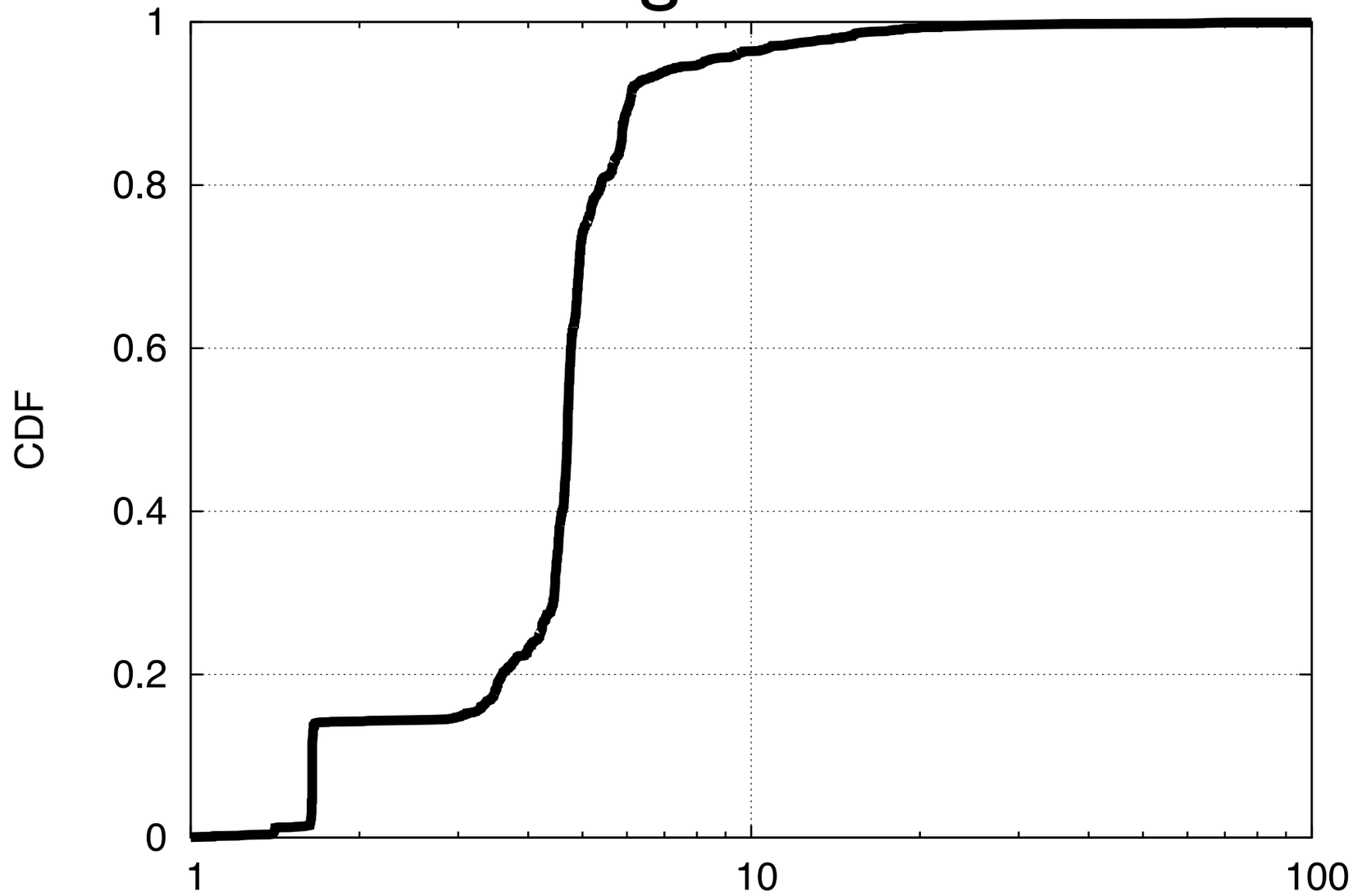
Delay is highly correlated with the network interface and time-of-day.	The protocol must tolerate highly variable delay.

Mean delay vs. transmission index



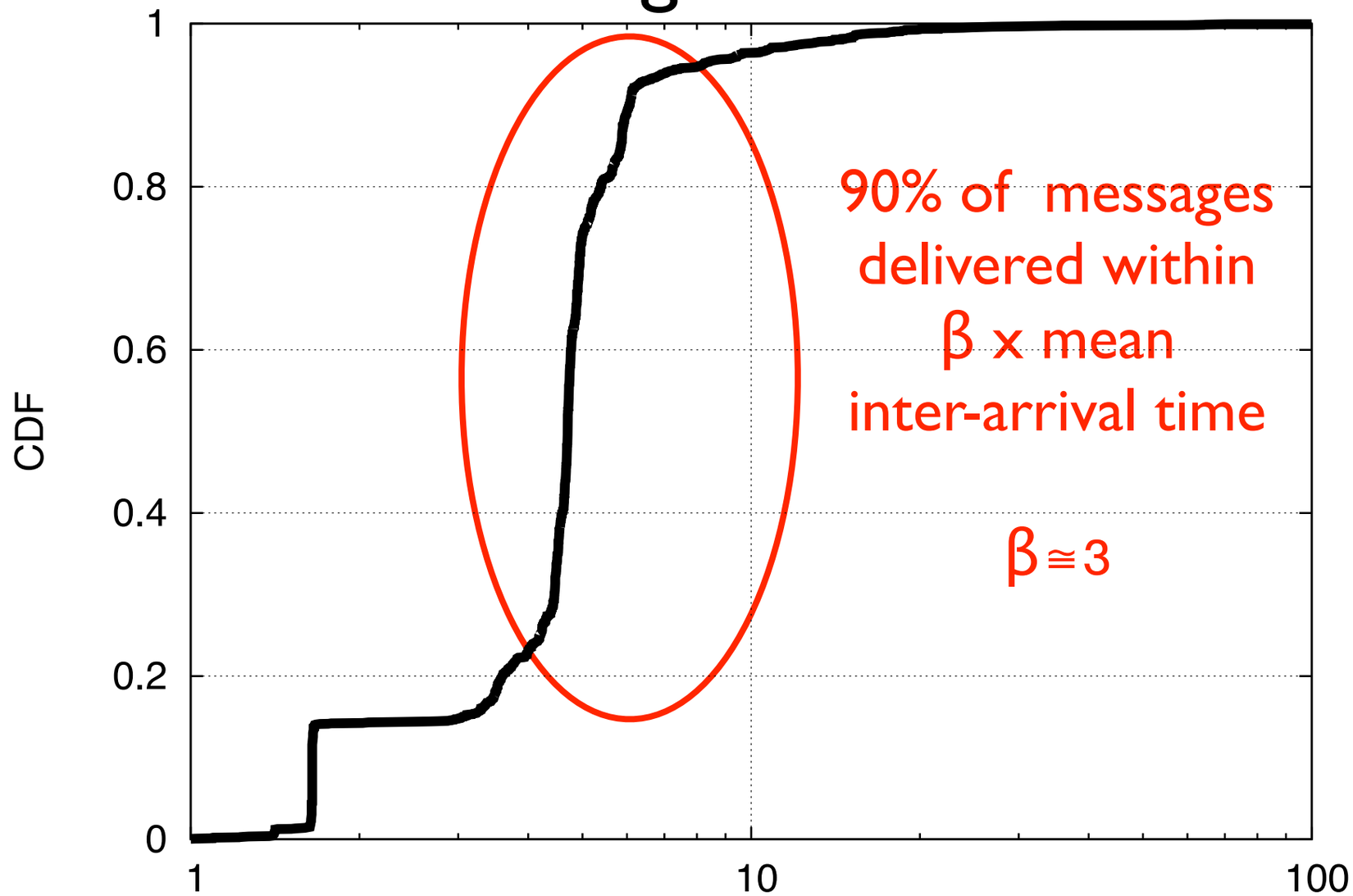
Delay is independent of transmission index.	The protocol may be agnostic to the quantity of messages transmitted.

CDF of message inter-arrival time



Despite high delays and reordering, messages are received at a steady rate in batches.	The protocol may assume a relatively constant message arrival rate.

CDF of message inter-arrival time



Despite high delays and reordering, messages are received at a steady rate in batches.	The protocol may assume a relatively constant message arrival rate.

- Preliminary experimentation

- Preliminary experimentation
  - Bi-directional SMS communication

- Preliminary experimentation
  - Bi-directional SMS communication
  - Increases transmission time and delay

- Preliminary experimentation
  - Bi-directional SMS communication
  - Increases transmission time and delay
  - Message reordering (~45%)

- Preliminary experimentation
  - Bi-directional SMS communication
  - Increases transmission time and delay
  - Message reordering (~45%)
  - Losses are frequent (> 20%)

Outline

- Channel characterization
- Transport protocol design
- Evaluation

Transport Protocol Design

- Goals

Transport Protocol Design

- Goals
  - Minimize message overhead

Transport Protocol Design

- Goals
  - Minimize message overhead
  - Maximize throughput

Transport Protocol Design

- Goals
  - Minimize message overhead
  - Maximize throughput
- Flow control and error control
  - Simplified version of NETBLT: SMS-TP

SMS-TP

SMS-TP

Fragment 1

SMS-TP

Fragment 1

Selective ACK

SMS-TP

Fragment 1

Selective ACK

Waits and avoids bidirectional communication

SMS-TP

Fragment 1

Selective ACK

Waits and avoids bidirectional communication

Unfettered message transmission

SMS-TP

Fragment 1

Selective ACK

Unfettered message transmission

Waits and avoids bidirectional communication

Single selective ACK tolerates message reordering

Final ACK

SMS-TP

Fragment 1

Selective ACK

When to send ACK?

SMS-TP

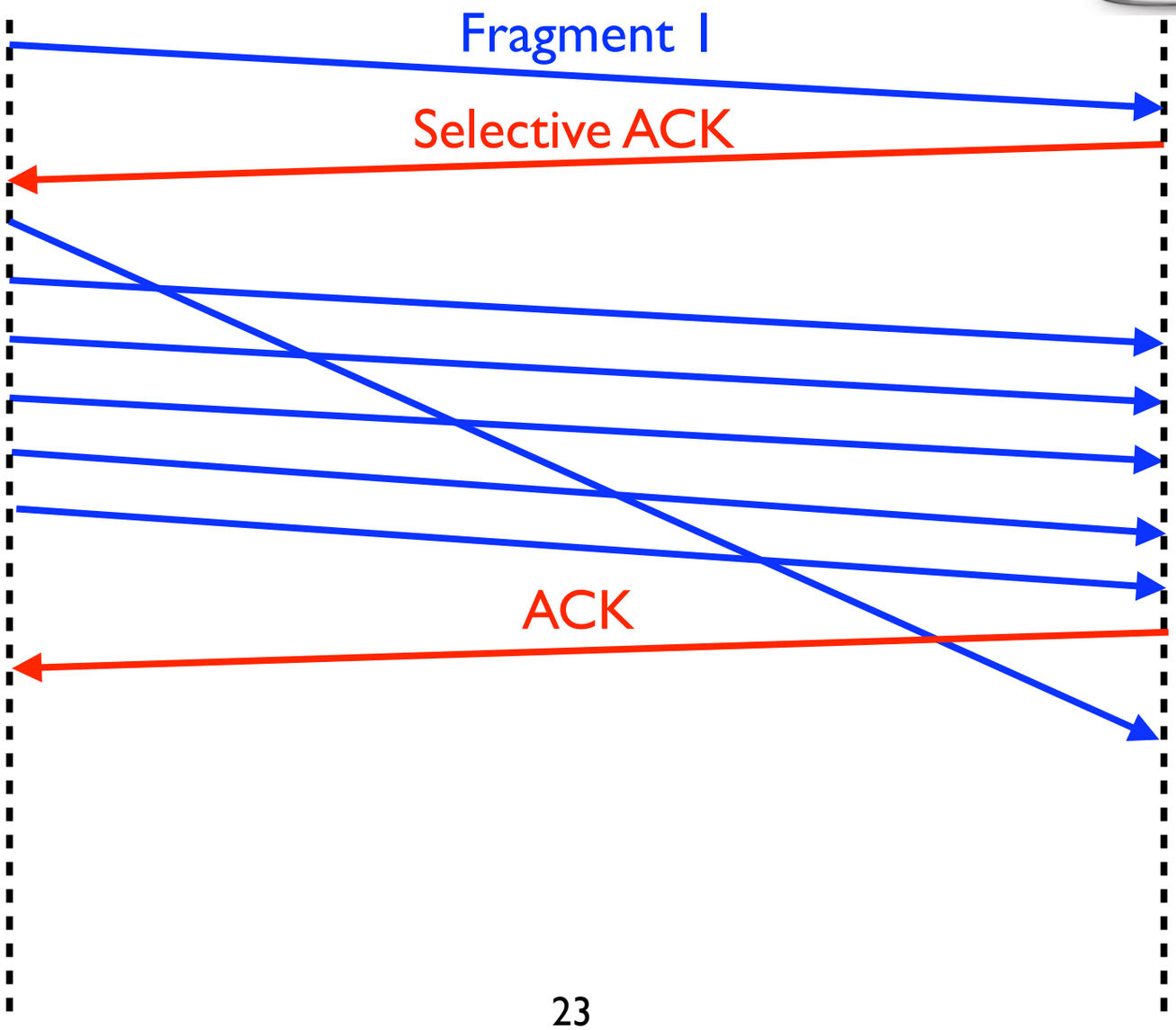
Fragment 1

Selective ACK

ACK

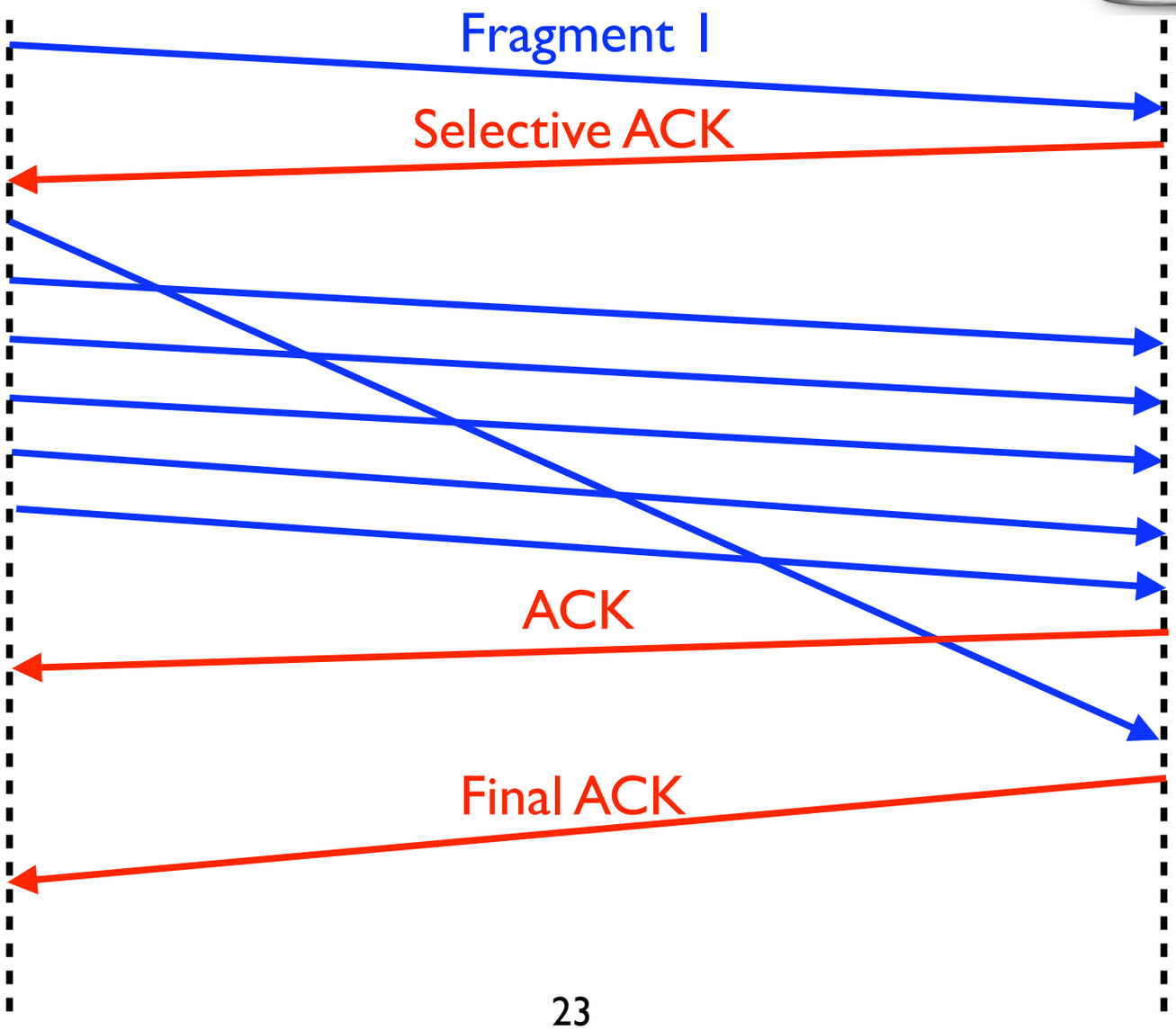
When to send ACK?

SMS-TP



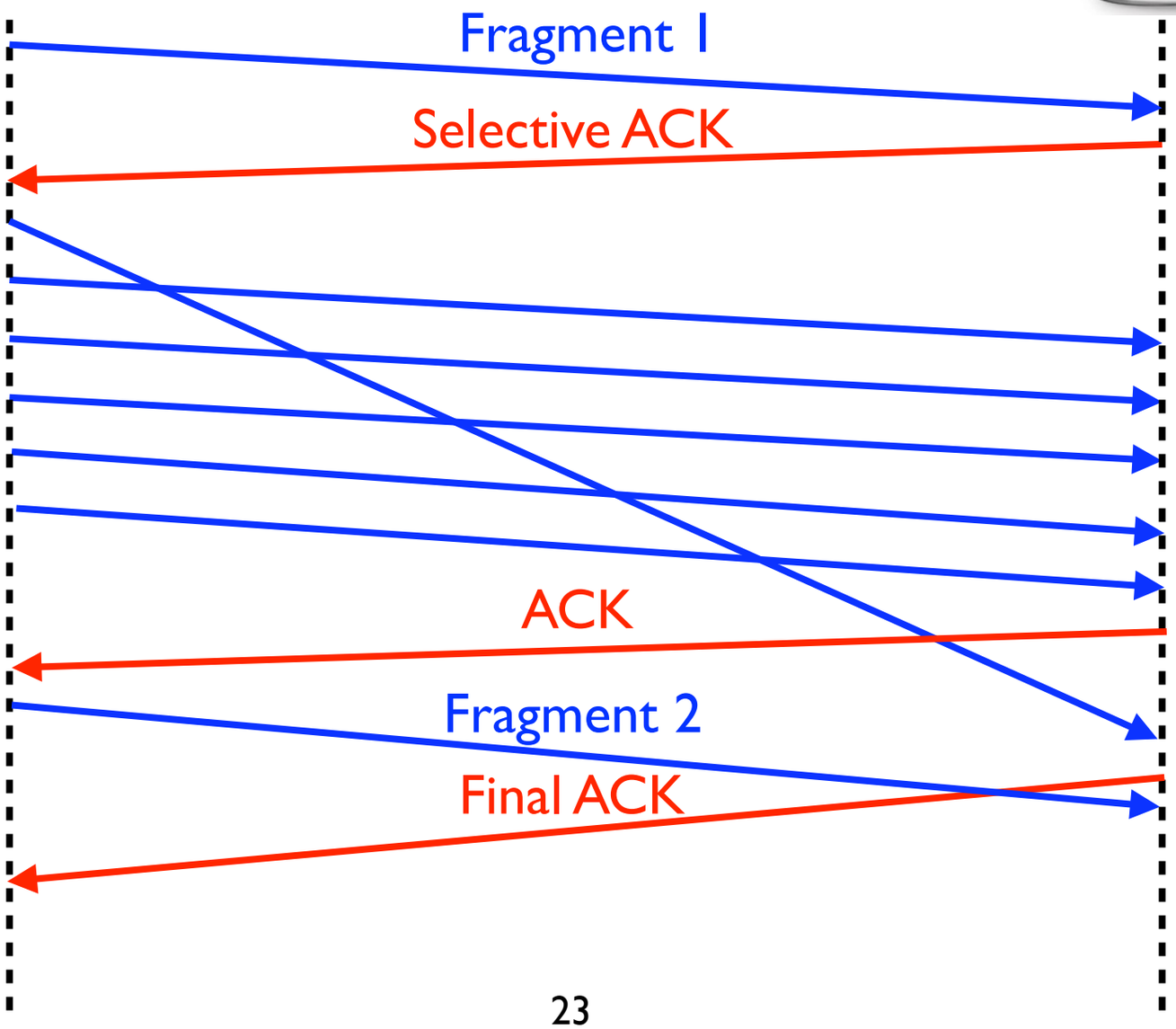
When to send ACK?

SMS-TP



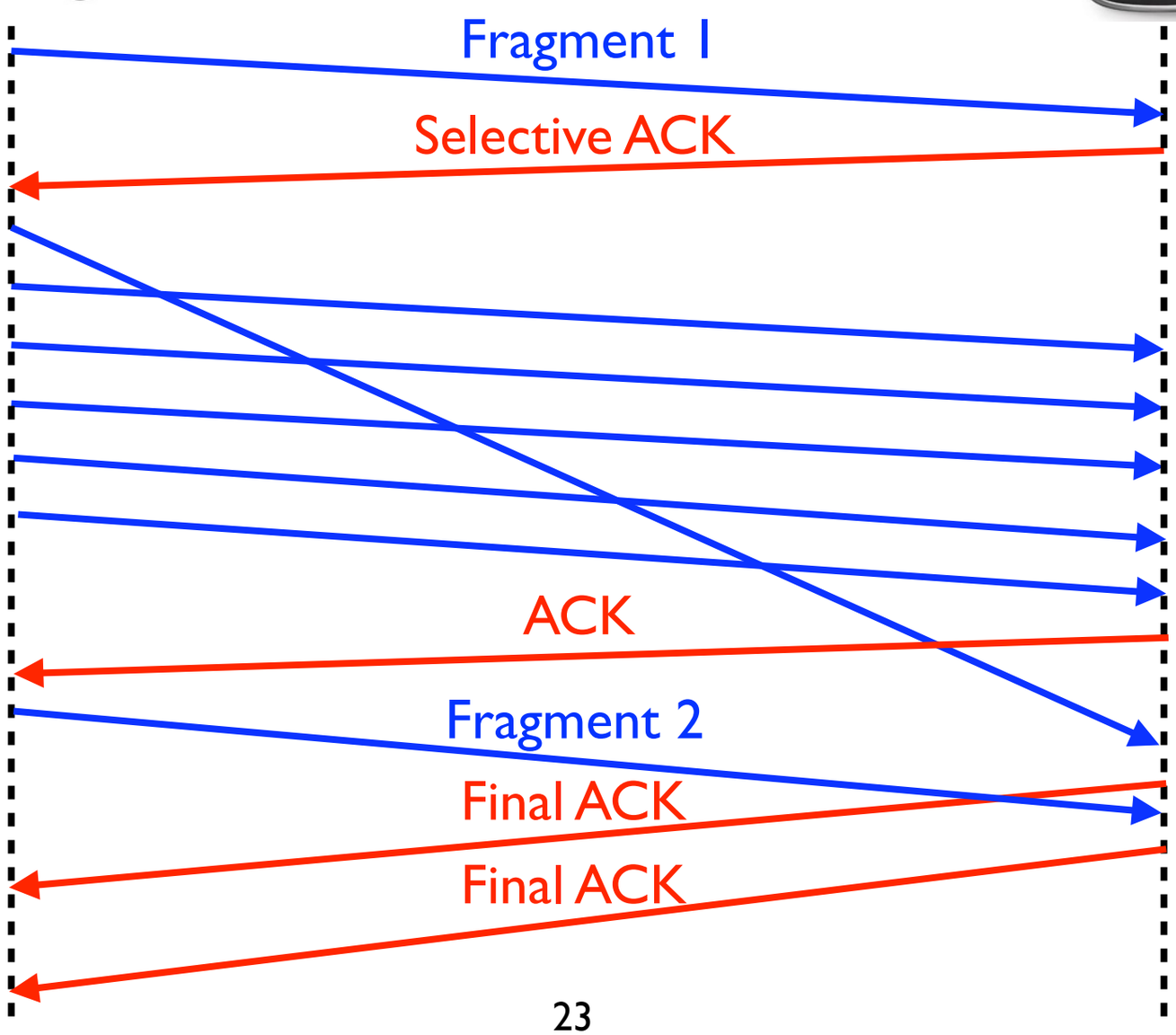
When to send ACK?

SMS-TP



When to send ACK?

SMS-TP



When to send ACK?

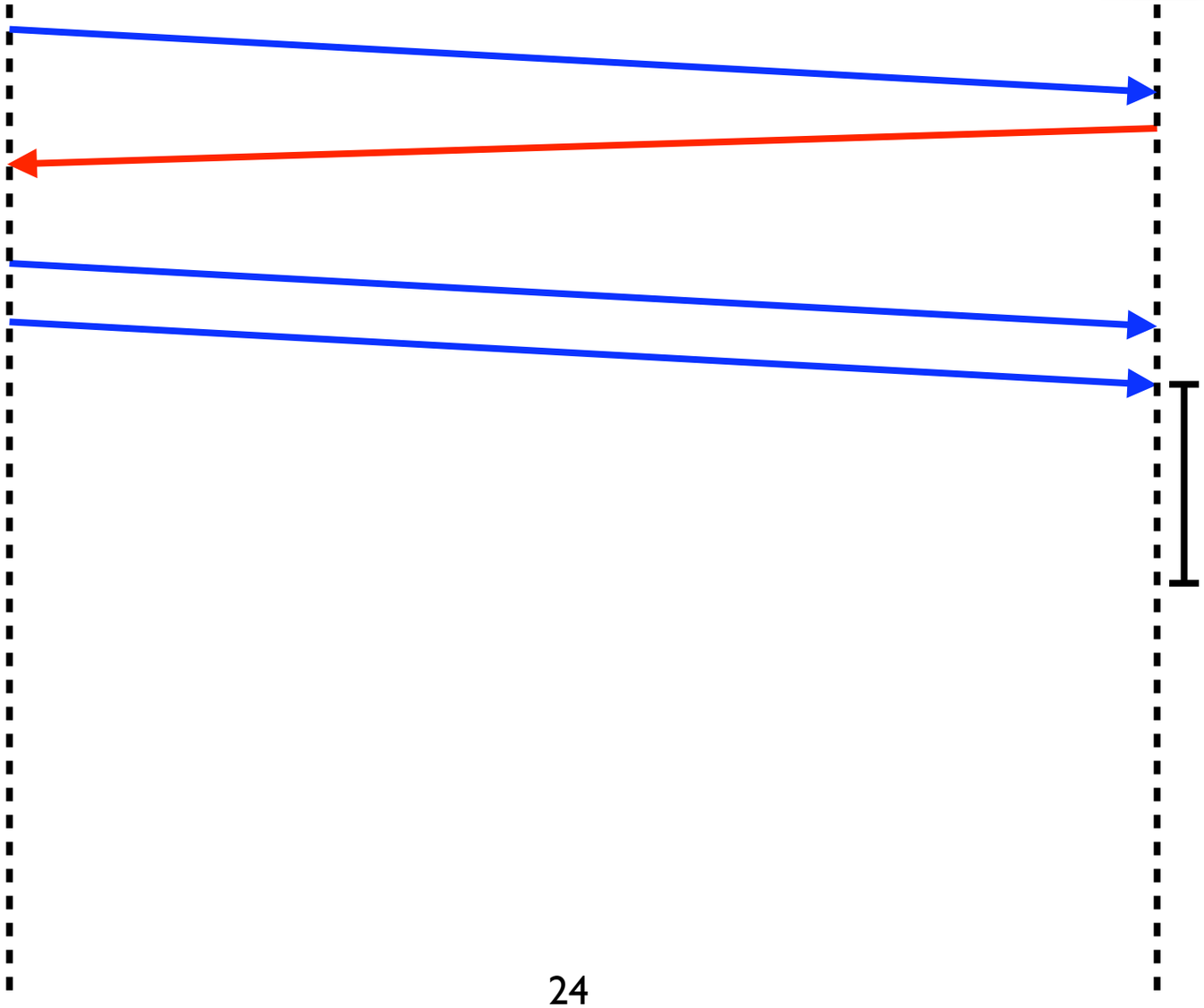
SMS-TP (receiver)

SMS-TP (receiver)

90% of messages are
delivered within
β x mean inter-arrival time

SMS-TP (receiver)

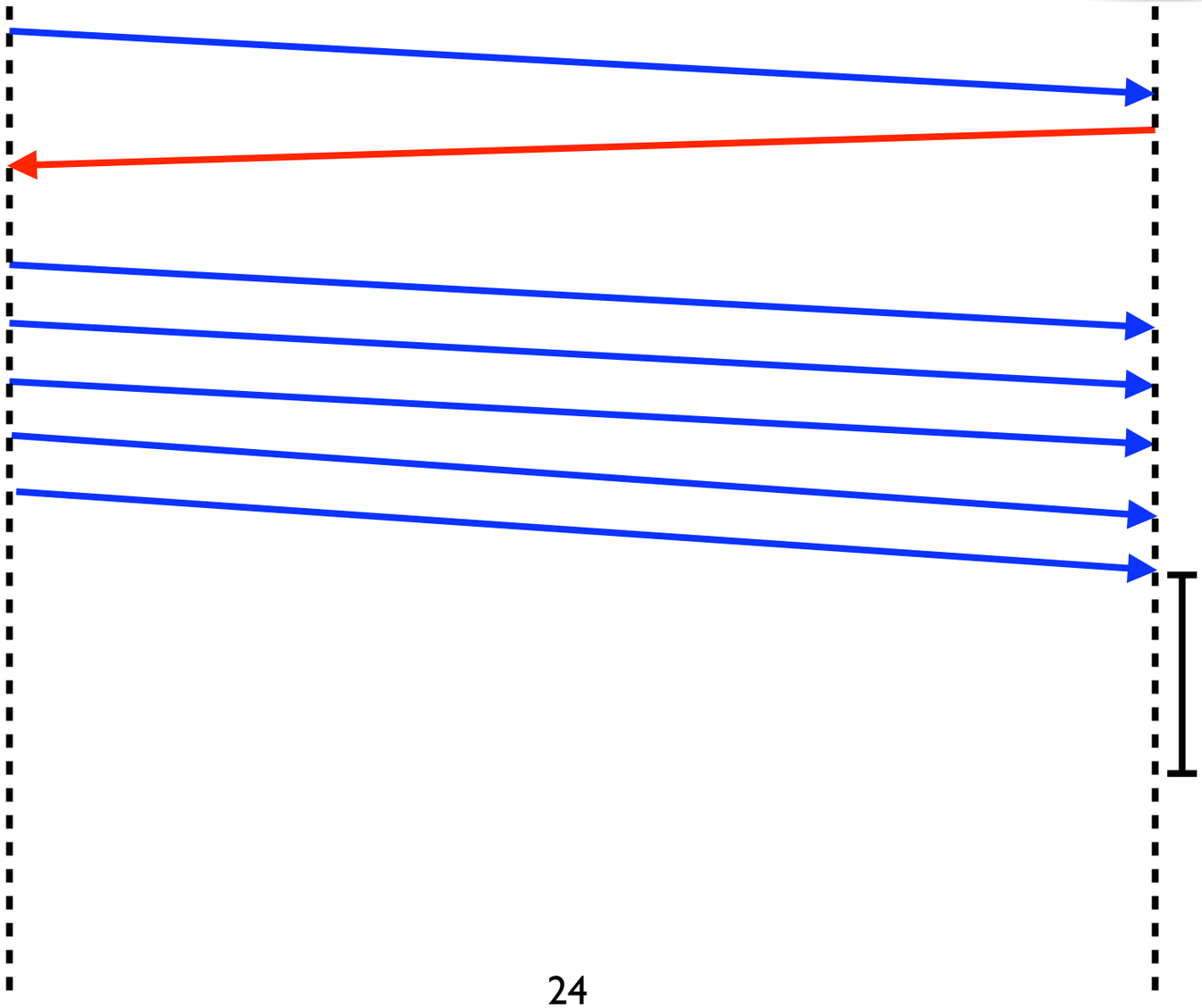
90% of messages are
delivered within
β x mean inter-arrival time



EWA of
inter-arrival
time

SMS-TP (receiver)

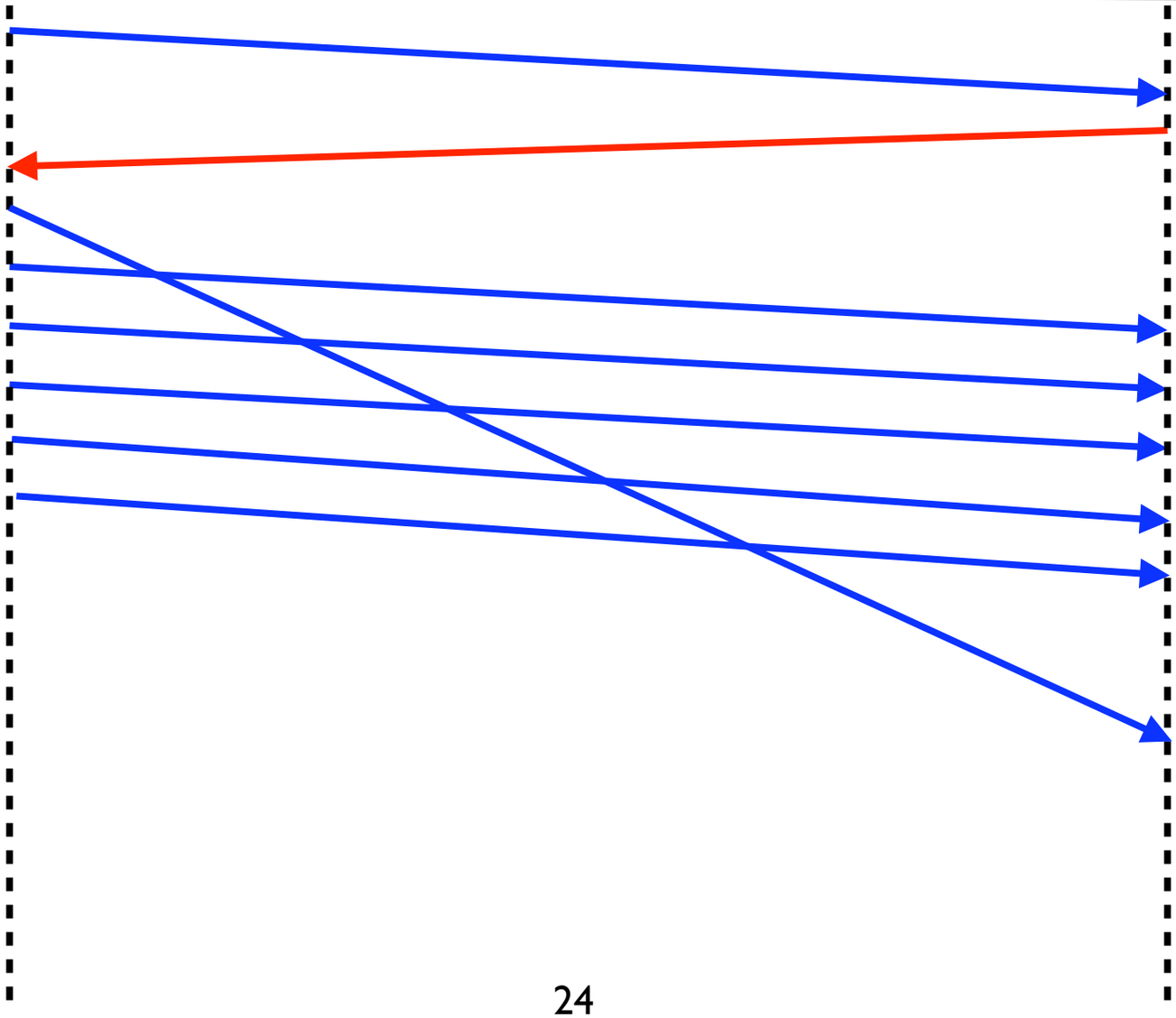
90% of messages are
delivered within
β x mean inter-arrival time



EWA of
inter-arrival
time

SMS-TP (receiver)

90% of messages are delivered within β x mean inter-arrival time

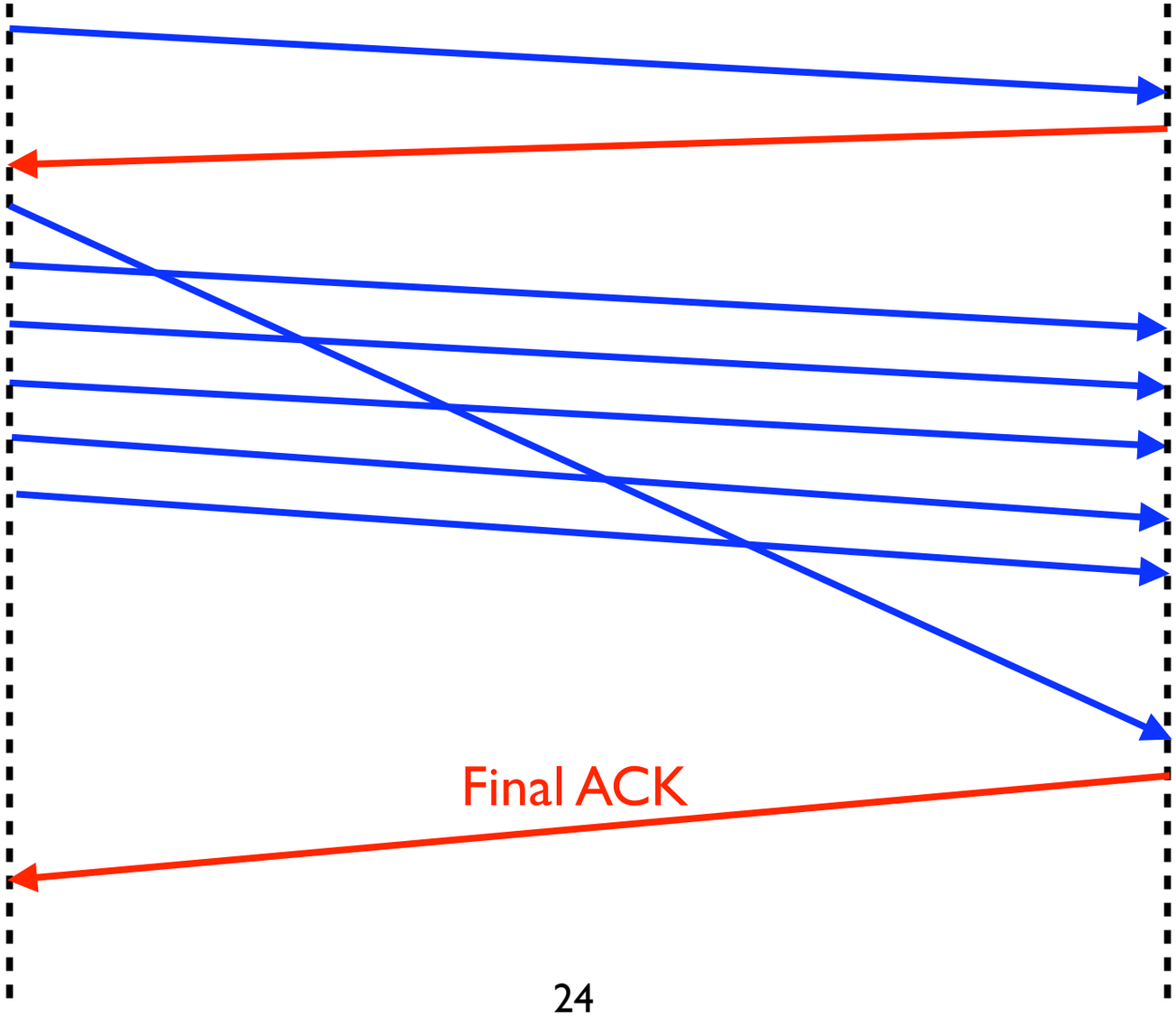


EWA of inter-arrival time

Message arrives within ACK timer

SMS-TP (receiver)

90% of messages are delivered within β x mean inter-arrival time

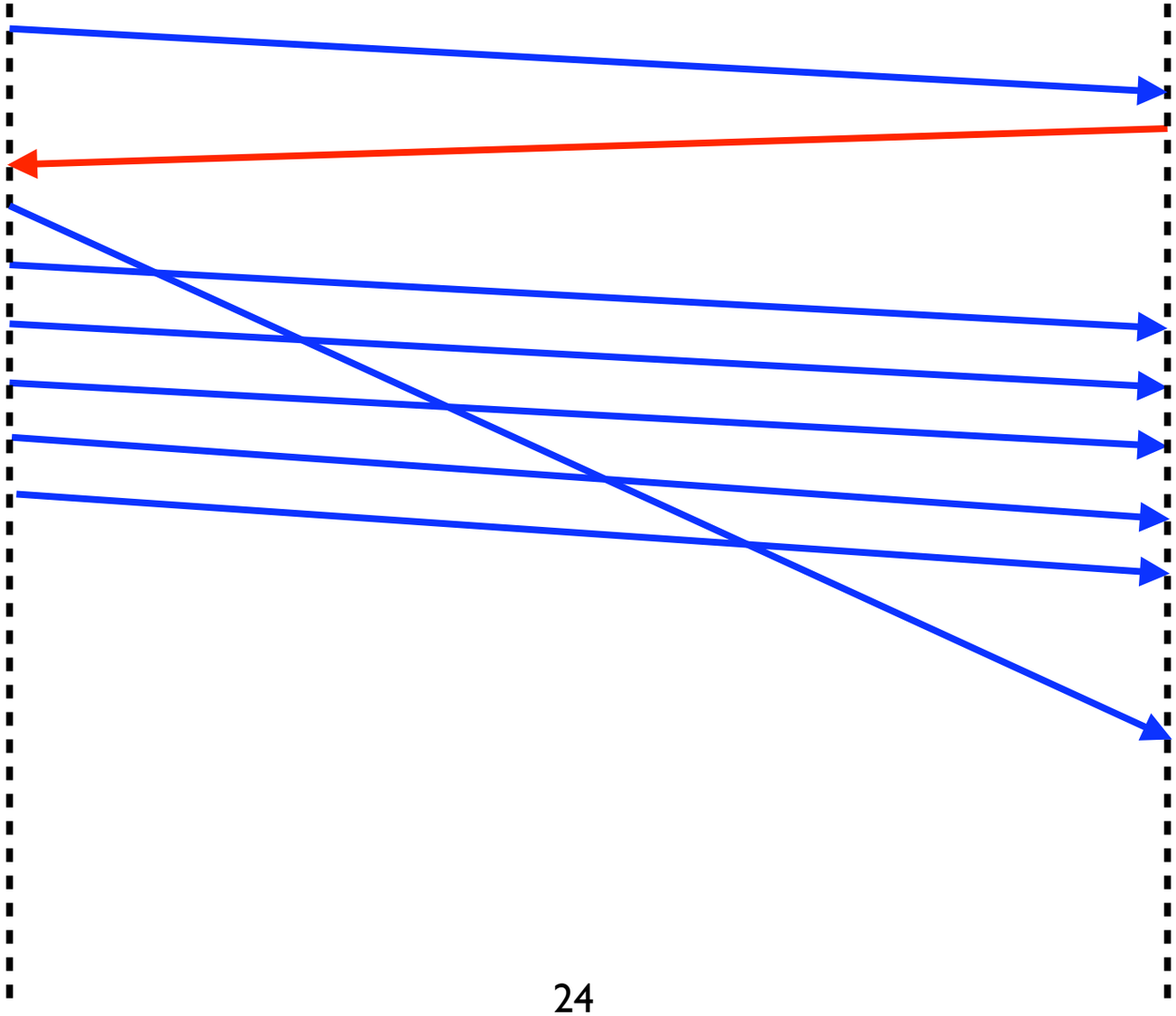


EWA of inter-arrival time

Message arrives within ACK timer

SMS-TP (receiver)

90% of messages are delivered within β x mean inter-arrival time



EWA of inter-arrival time

Message arrives within ACK timer

Has the data been transferred?

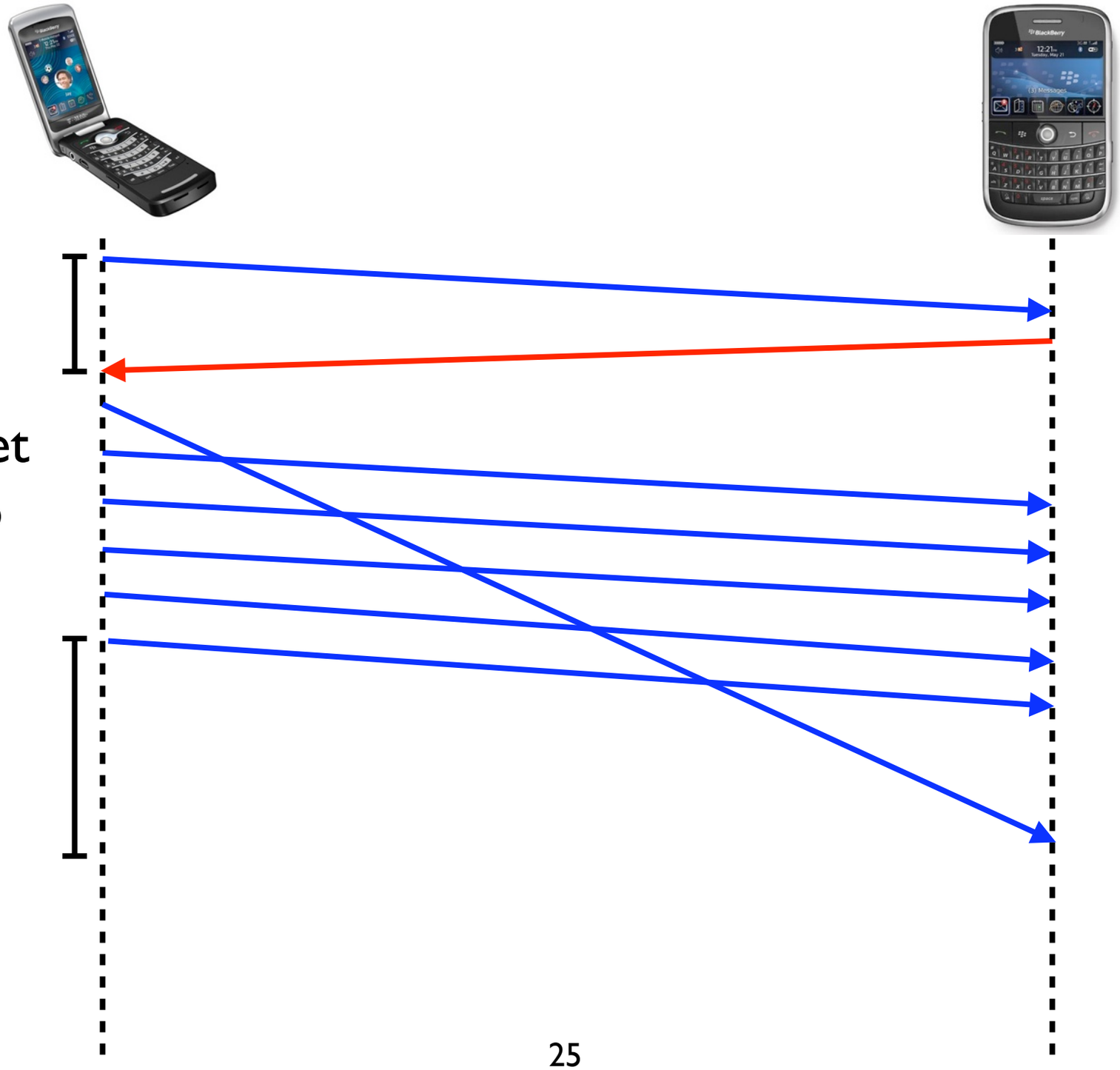
Retransmit?

SMS-TP (sender)

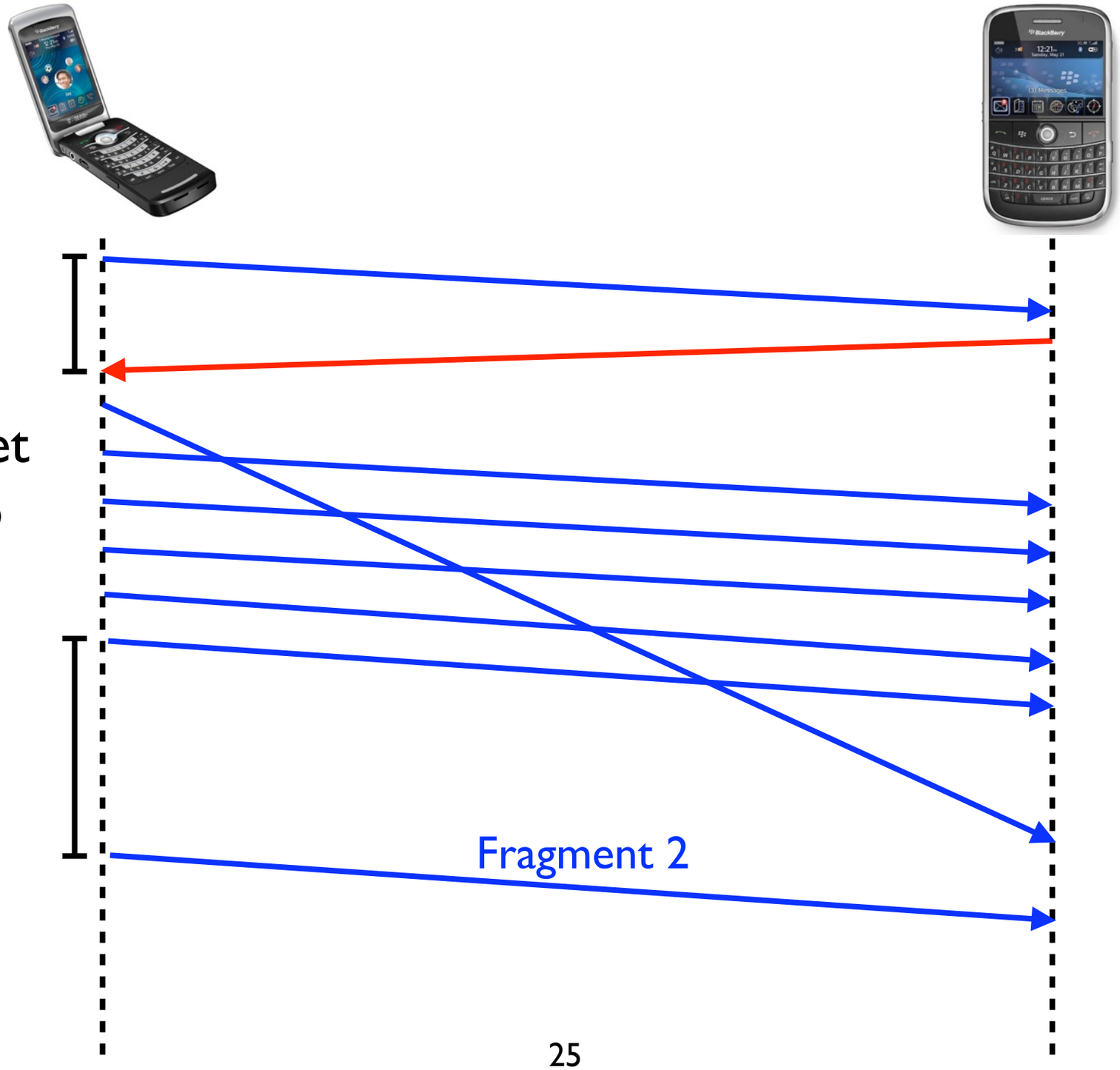
SMS-TP (sender)

Setup delay
(1 RTT)

SMS-TP (sender)



SMS-TP (sender)



SMS-TP (sender)

Setup delay (1 RTT)

ACK timer set to γ x setup delay

ACK timer expires

Fragment 2

Final ACK

Detects duplicate fragment, retransmit final ACK

Outline

- Channel characterization
- Transport protocol design
- Evaluation

Evaluation

- Implementation

Evaluation

- Implementation
  - CLDC compliant Java library

Evaluation

- Implementation
  - CLDC compliant Java library
  - Free for download

Evaluation

- Implementation
  - CLDC compliant Java library
  - Free for download
- Evaluation

Evaluation

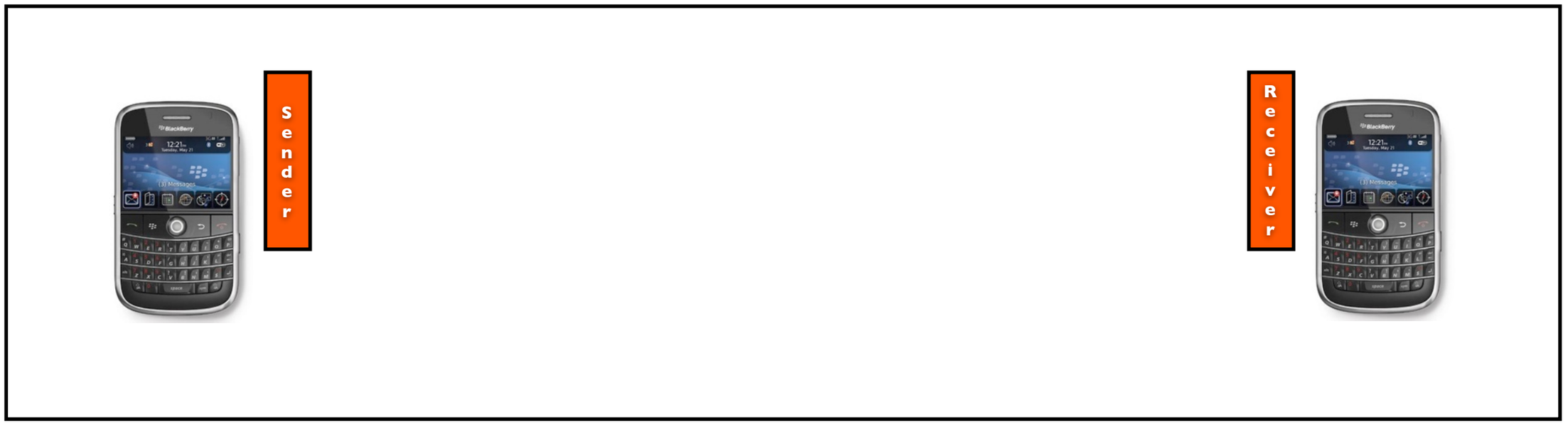
- Implementation
  - CLDC compliant Java library
  - Free for download
- Evaluation
  - Field trial

Evaluation

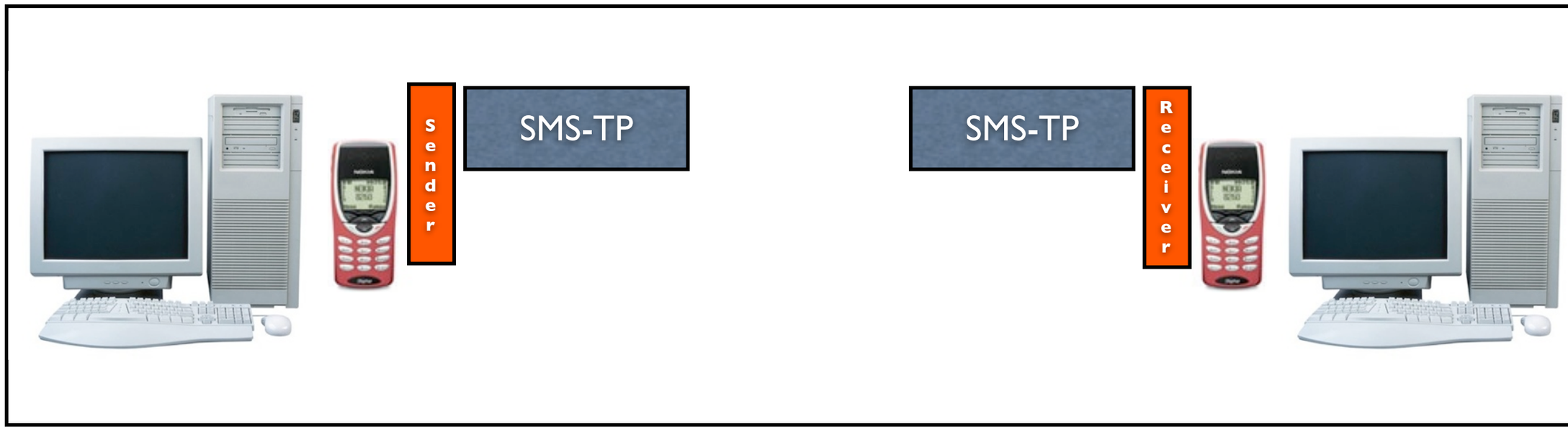
- Implementation
  - CLDC compliant Java library
  - Free for download
- Evaluation
  - Field trial
  - Simulation

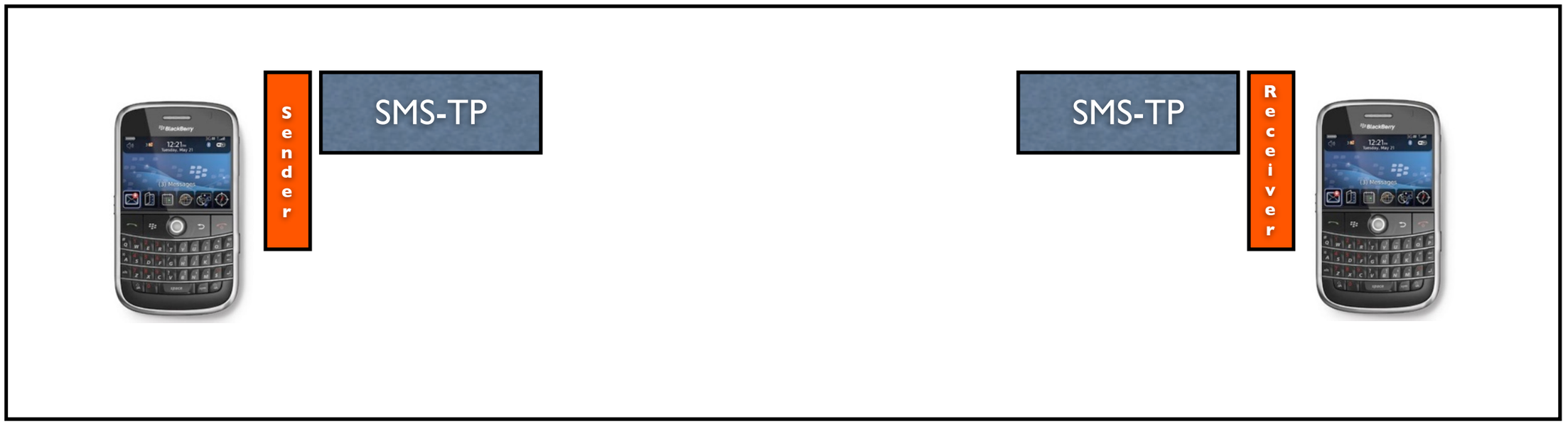
Field trial

Field trial

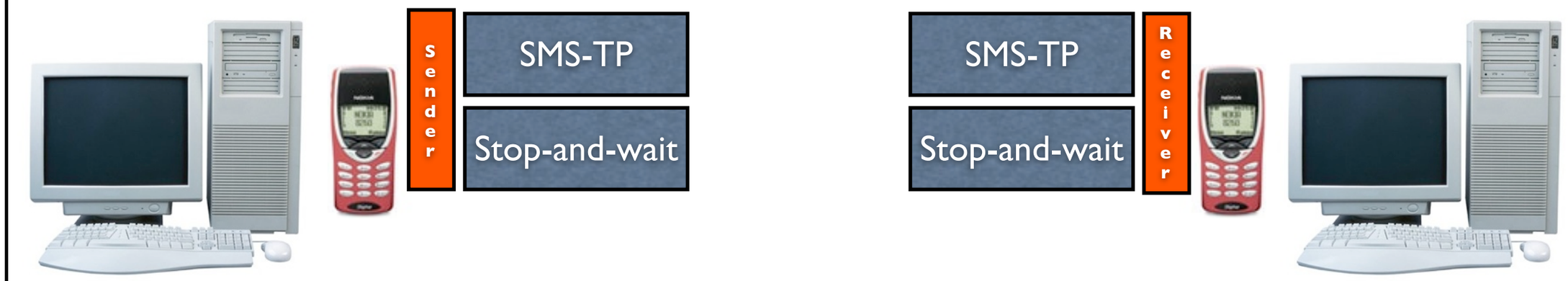


Field trial



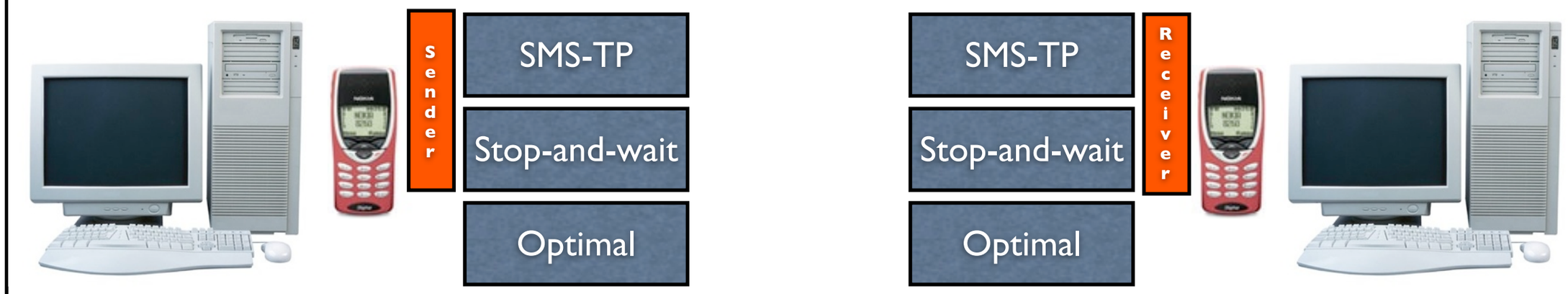


Field trial


Sender
SMS-TP
Stop-and-wait
SMS-TP
Stop-and-wait
Receiver

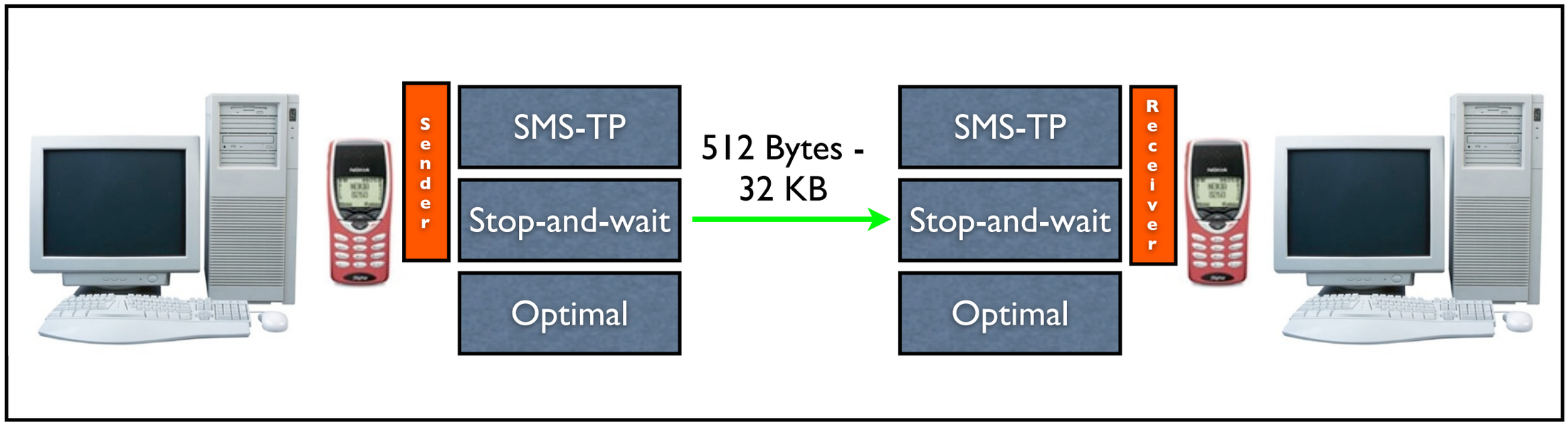

Sender
SMS-TP
Stop-and-wait
SMS-TP
Stop-and-wait
Receiver

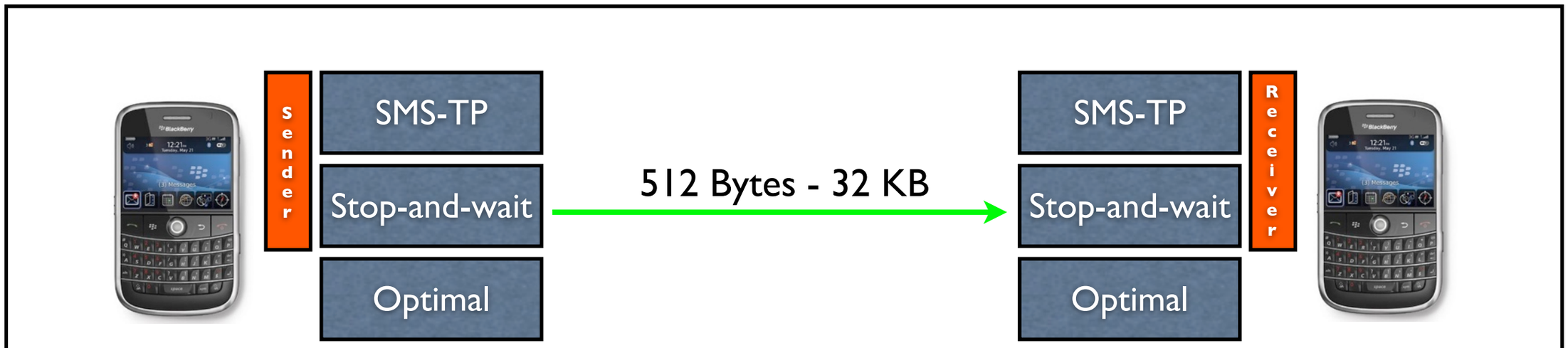
Field trial


Sender
SMS-TP
Stop-and-wait
Optimal
SMS-TP
Stop-and-wait
Optimal
Receiver


Sender
SMS-TP
Stop-and-wait
Optimal
SMS-TP
Stop-and-wait
Optimal
Receiver

Field trial


Sender
SMS-TP
Stop-and-wait
Optimal
512 Bytes - 32 KB
SMS-TP
Stop-and-wait
Optimal
Receiver


Sender
SMS-TP
Stop-and-wait
Optimal
512 Bytes - 32 KB
SMS-TP
Stop-and-wait
Optimal
Receiver

Outperforms existing by a factor of two due to consolidation of ACKs

8.5% off calculated optimal

Pipelined transmission results in:

545% increase in performance

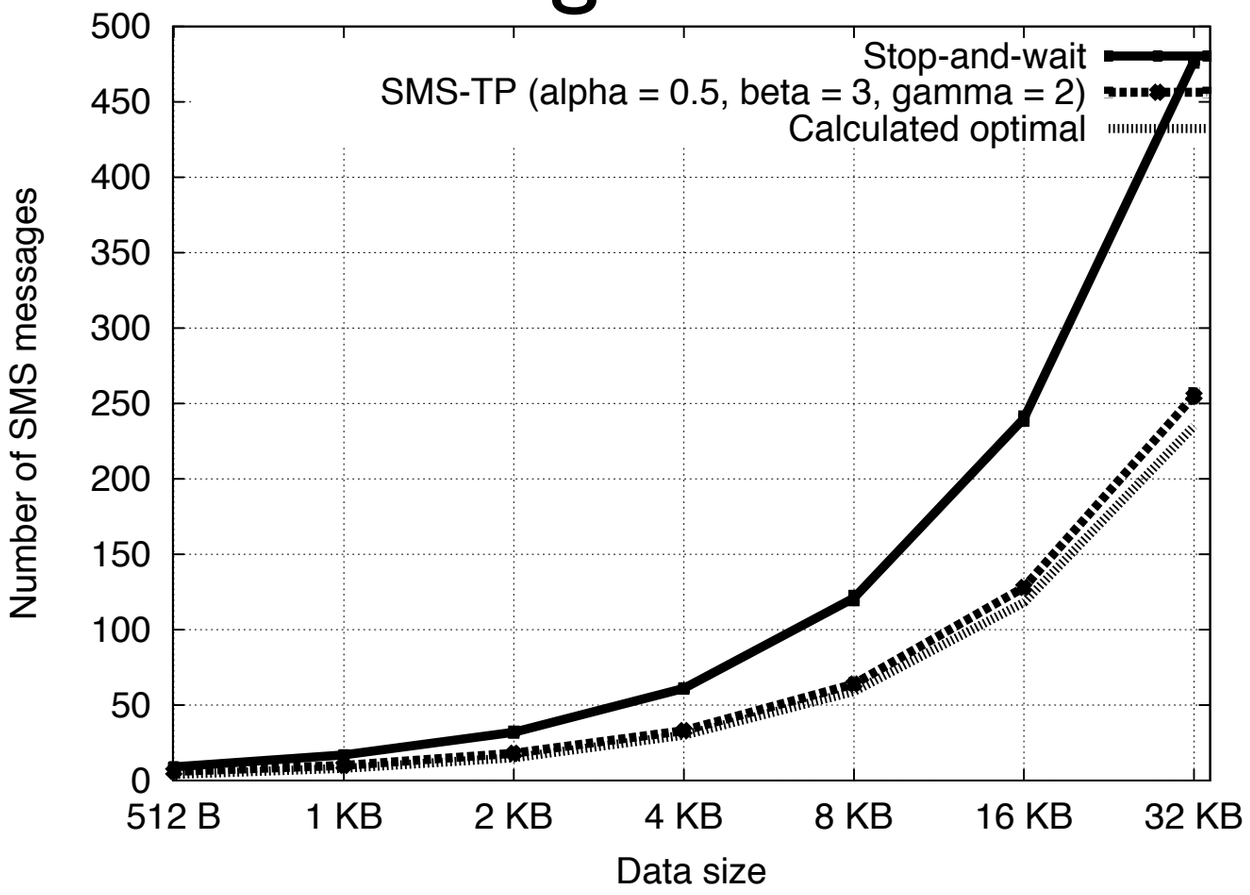
Outperforms existing by a factor of two due to consolidation of ACKs

8.5% off calculated optimal

Pipelined transmission results in:

545% increase in performance

Message overhead



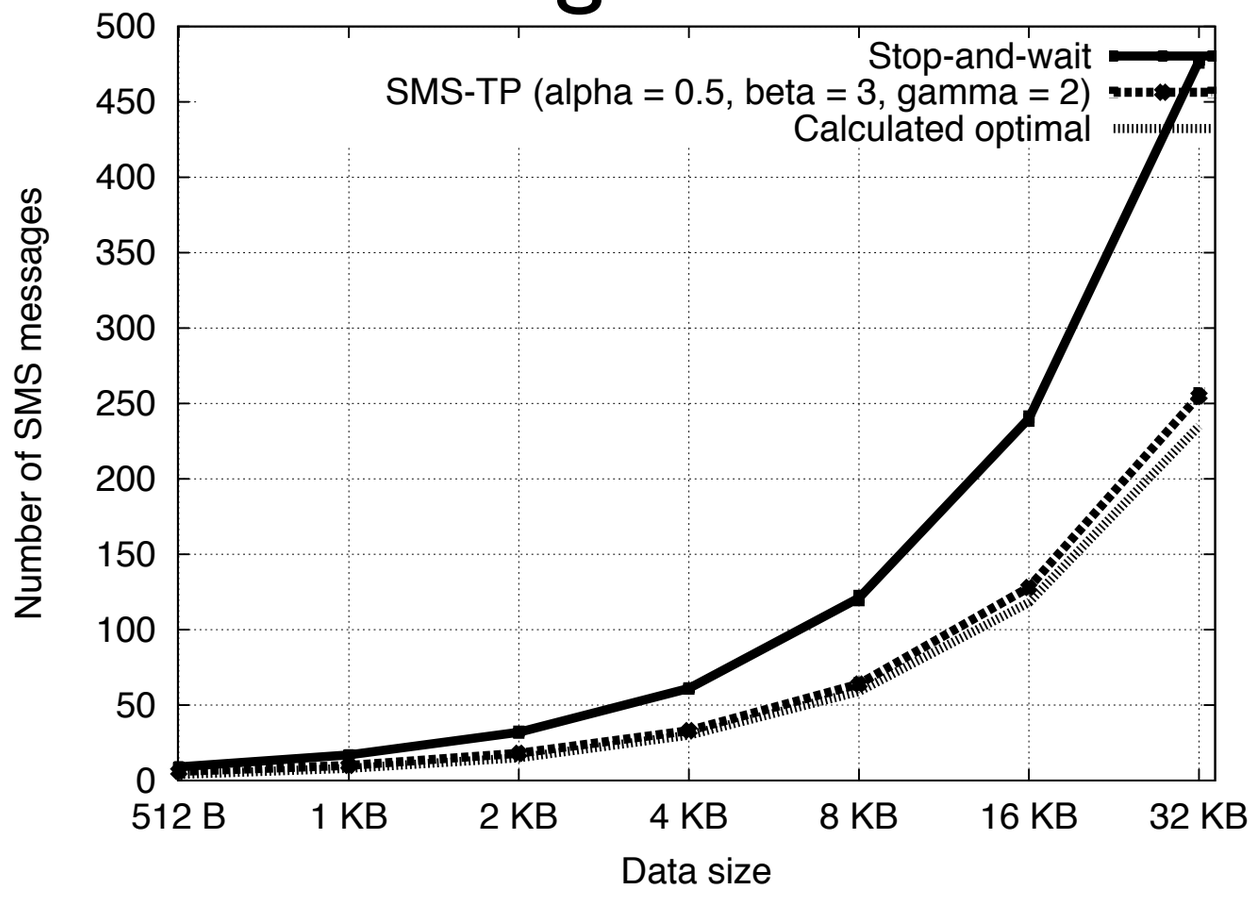
Outperforms existing by a factor of two due to consolidation of ACKs

8.5% off calculated optimal

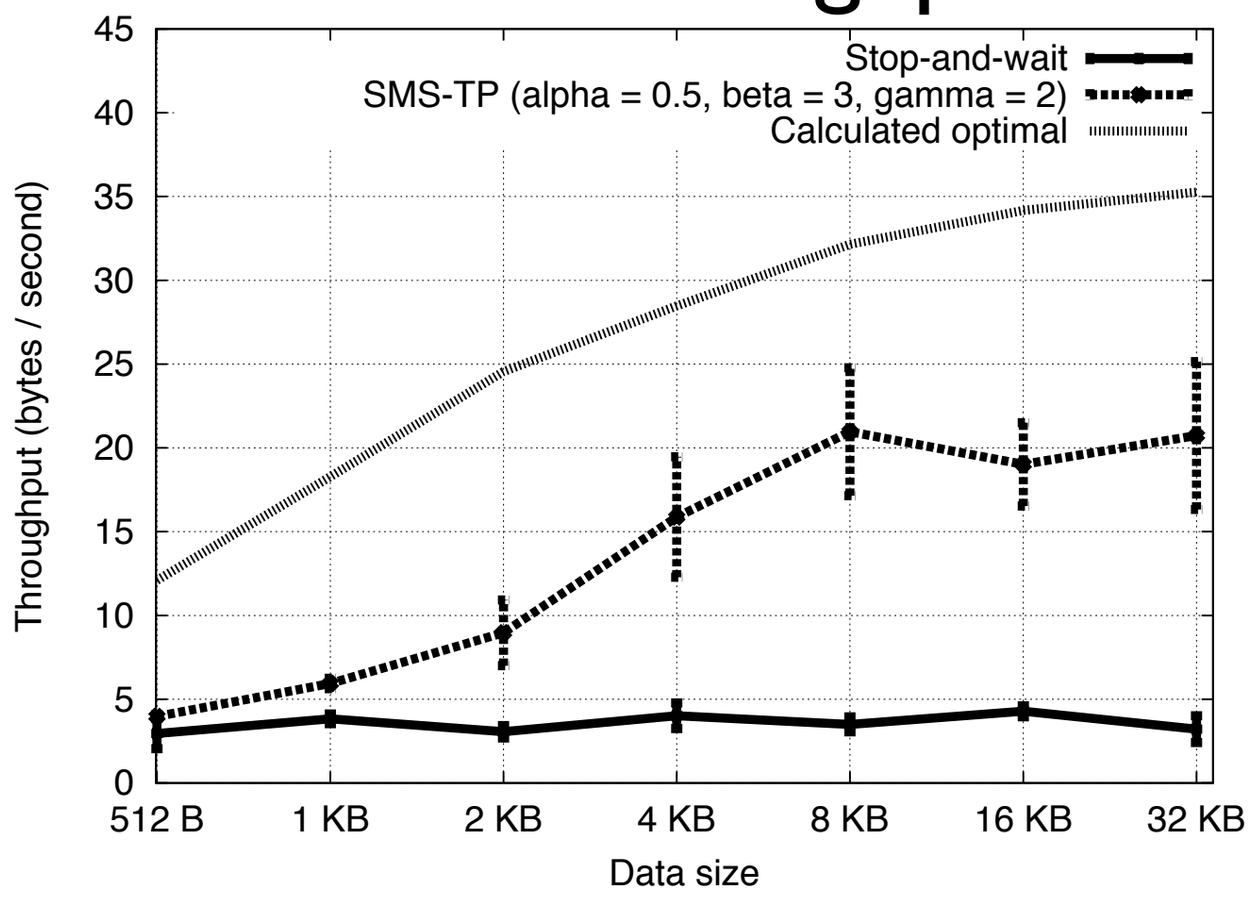
Pipelined transmission results in:

545% increase in performance

Message overhead



Data throughput



Simulation

Simulation

- Study may not reflect conditions of a developing region (Zerfos'06)

Simulation

- Study may not reflect conditions of a developing region (Zerfos'06)
  - Loss rate

Simulation

- Study may not reflect conditions of a developing region (Zerfos'06)
  - Loss rate
  - Delay

Simulation

- Study may not reflect conditions of a developing region (Zerfos'06)
  - Loss rate
  - Delay

Simulation

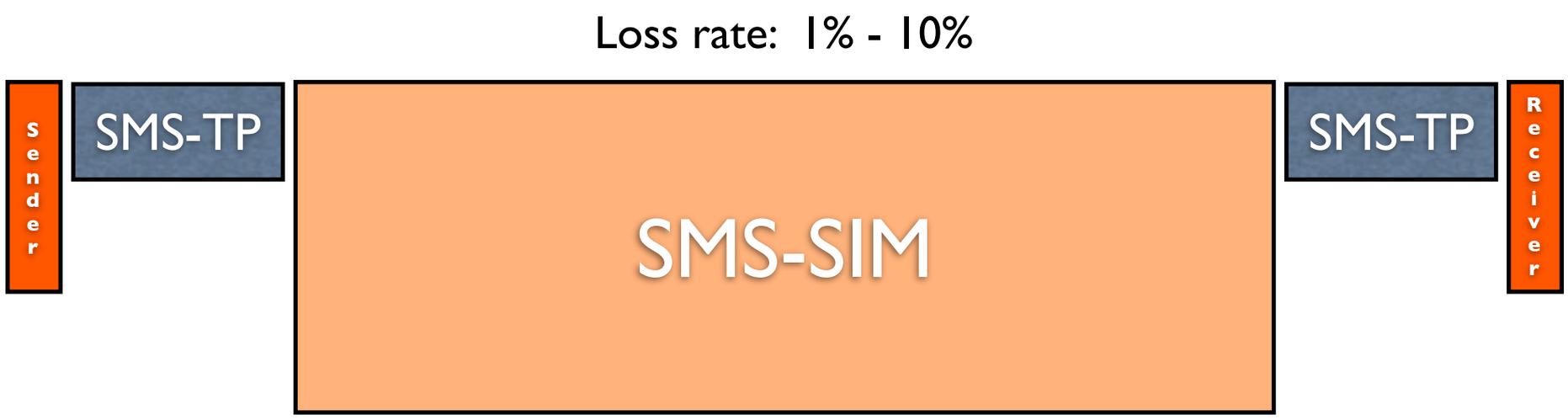
- Study may not reflect conditions of a developing region (Zerfos'06)
  - Loss rate
  - Delay

Simulation

- Study may not reflect conditions of a developing region (Zerfos'06)
  - Loss rate
  - Delay

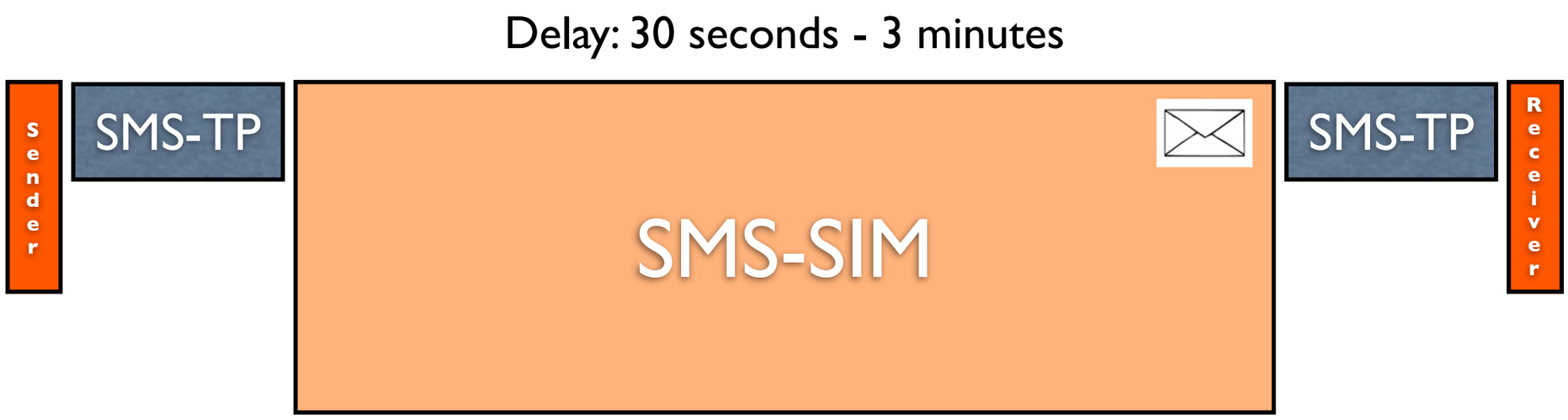
Simulation

- Study may not reflect conditions of a developing region (Zerfos'06)
  - Loss rate
  - Delay



Simulation

- Study may not reflect conditions of a developing region (Zerfos'06)
  - Loss rate
  - Delay



Delay and loss have a negligible impact on message overhead
- except under high loss and high delay situations

S-

A loss rate increase has statistically insignificant impact on throughput

6 fold increase in delay, 50% decrease in throughput

50% better than stop-and-wait in a well provisioned network

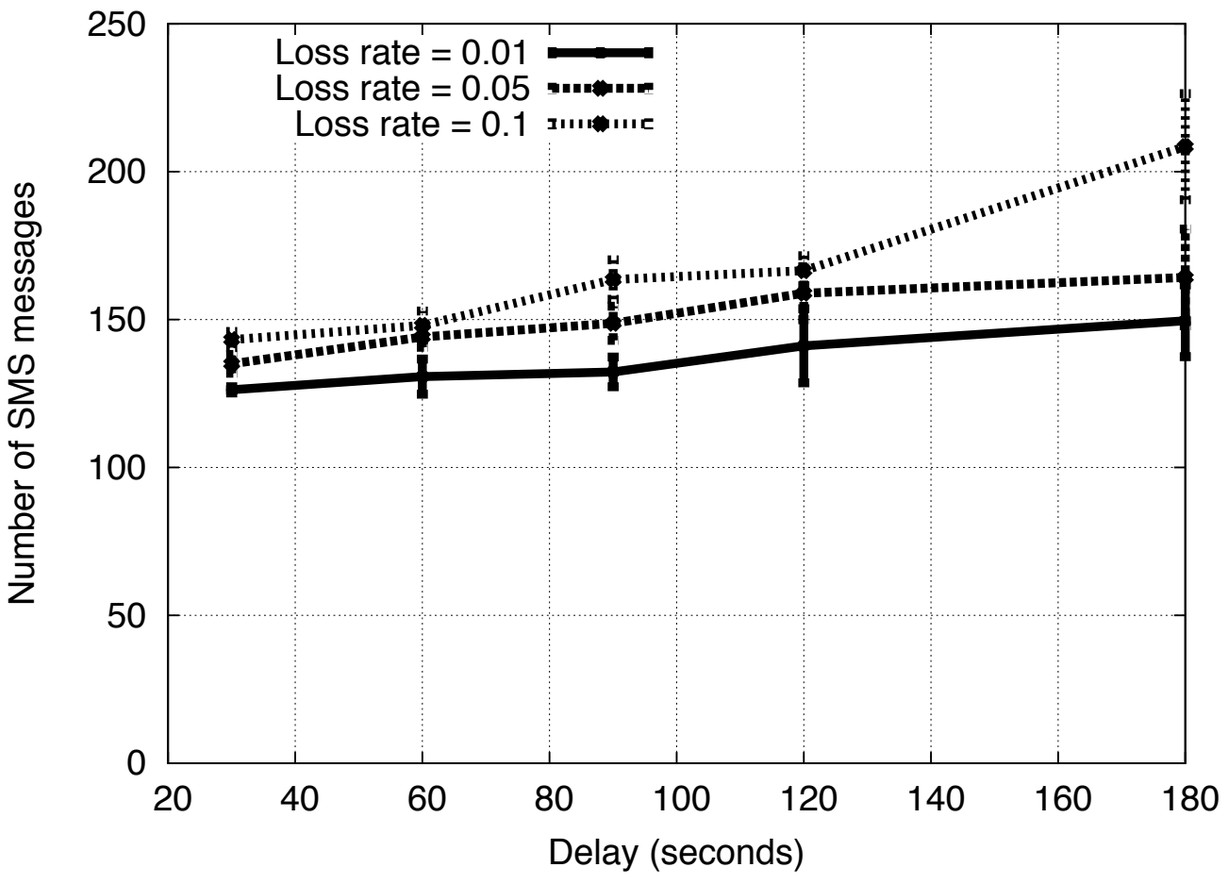
Delay and loss have a negligible impact on message overhead
- except under high loss and high delay situations

A loss rate increase has statistically insignificant impact on throughput

6 fold increase in delay, 50% decrease in throughput

50% better than stop-and-wait in a well provisioned network

Message overhead



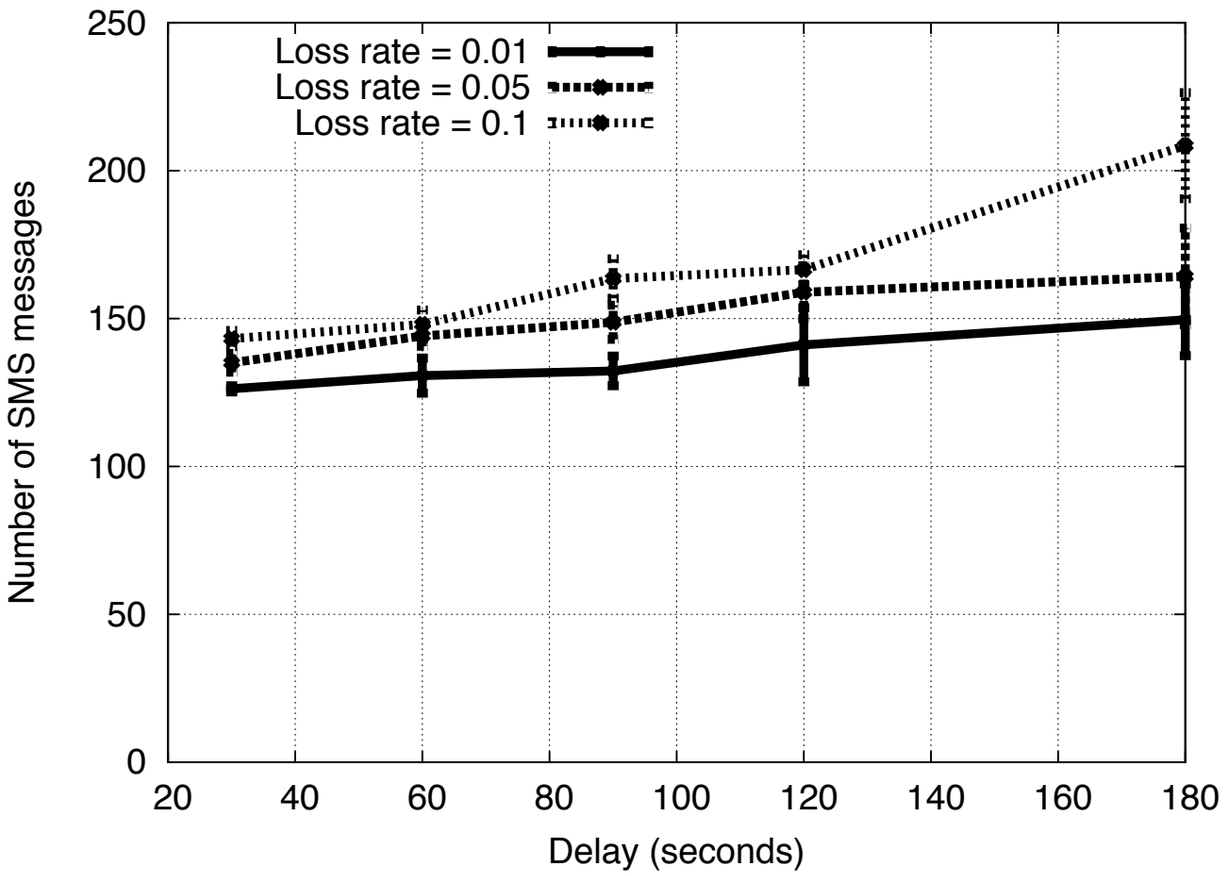
Delay and loss have a negligible impact on message overhead
- except under high loss and high delay situations

A loss rate increase has statistically insignificant impact on throughput

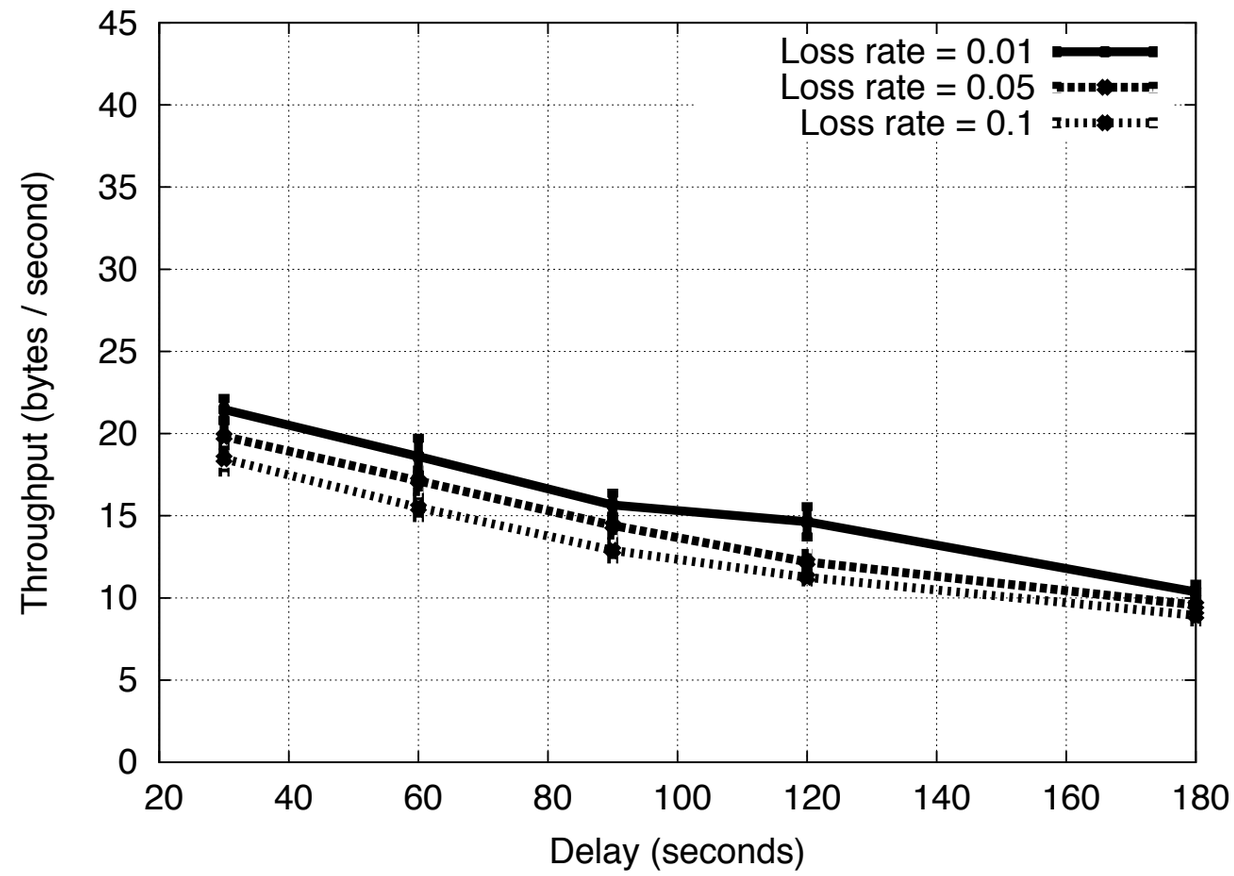
6 fold increase in delay, 50% decrease in throughput

50% better than stop-and-wait in a well provisioned network

Message overhead



Data throughput



Summary

- Studied the channel characteristics of SMS from the perspective of mobile devices sending bursts of messages

Summary

- Studied the channel characteristics of SMS from the perspective of mobile devices sending bursts of messages
- Design and implement an efficient and reliable SMS-based data transport protocol

Summary

- Studied the channel characteristics of SMS from the perspective of mobile devices sending bursts of messages
- Design and implement an efficient and reliable SMS-based data transport protocol
  - Reduces message overhead by 50%

Summary

- Studied the channel characteristics of SMS from the perspective of mobile devices sending bursts of messages
- Design and implement an efficient and reliable SMS-based data transport protocol
  - Reduces message overhead by 50%
  - Increases throughput by as much as 545%

Why not TCP?

- Significant delays
- Messages rarely lost
- Reordering is common
- Does not suffer from congestion drops

Questions?