Active Probing for Available Bandwidth Estimation

Sridhar Machiraju

UC Berkeley OASIS Retreat, Jan 2005

Joint work with D.Veitch, F.Baccelli, A.Nucci, J.Bolot

Outline

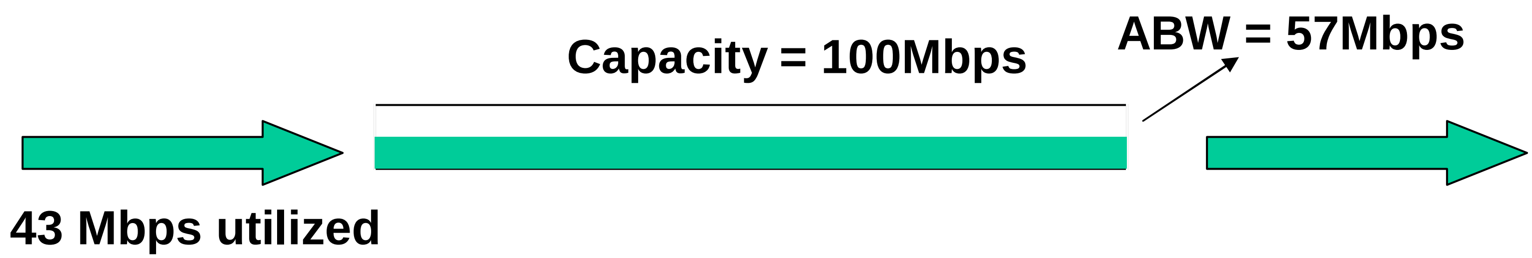
- Motivation
- Packet pair problem setup
- Describing the system
- Solving with i.i.d. assumption
- In practice
- Conclusions and future work

Estimating Available Bandwidth

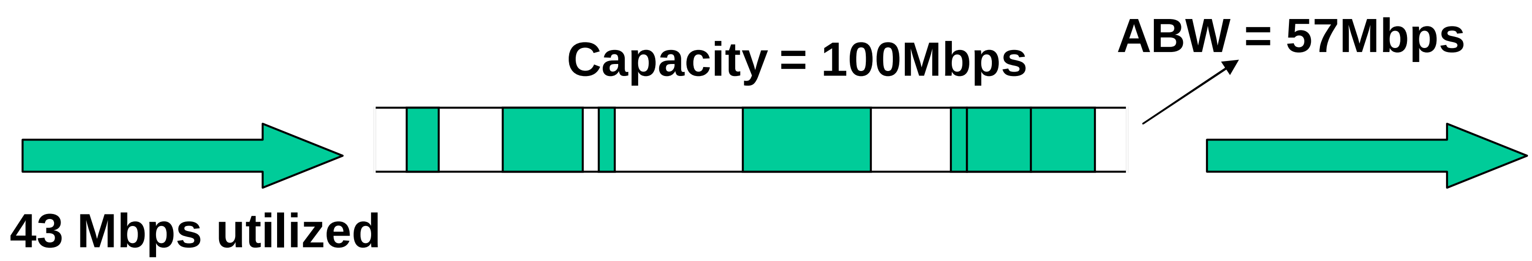
- Why
  - Path selection, SLA verification, network debugging, congestion control mechanisms
- How
  - Estimate cross-traffic rate and subtract from known capacity
    - Use **packet pairs**
  - Estimate available bandwidth directly
    - Use packet trains

Available Bandwidth

- Fluid Definition - spare capacity on a link

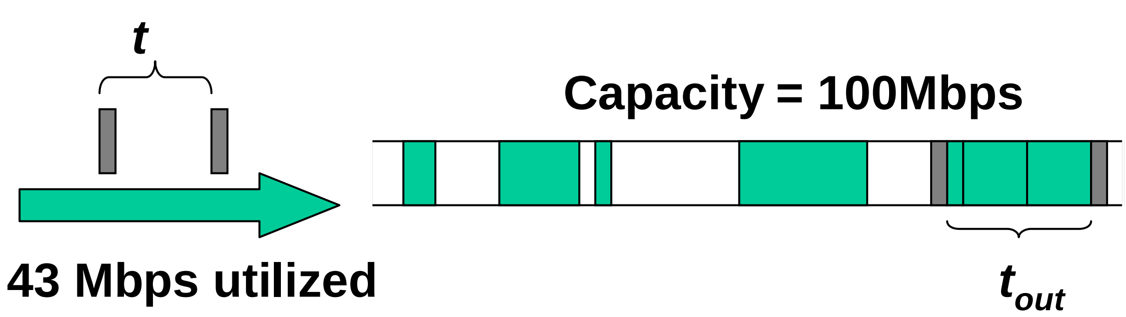


- In reality, we have a discrete system
  - Avail. b/w averaged over some time scale



Packet Pair Methods

- Assume single FIFO queue of known capacity C
  - Queuing at other hops on path negligible
- Send packet pair separated by t
- Output separation t_{out} depends on cross-traffic that arrived in t



Problem Statement

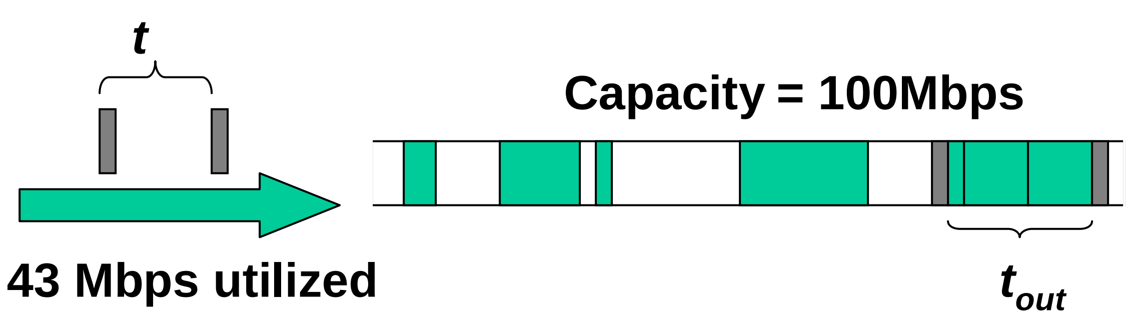
- What information about the cross-traffic do packet pair delays expose?
  - Consider multi-hop case
  - How to use packet pair in practice
- Might not apply in practice if
  - Non-FIFO queuing
  - Link layer multi-path

Outline

- Motivation
- Packet pair problem setup
- Describing the system
- Solving with i.i.d. assumption
- In practice
- Conclusions and future work

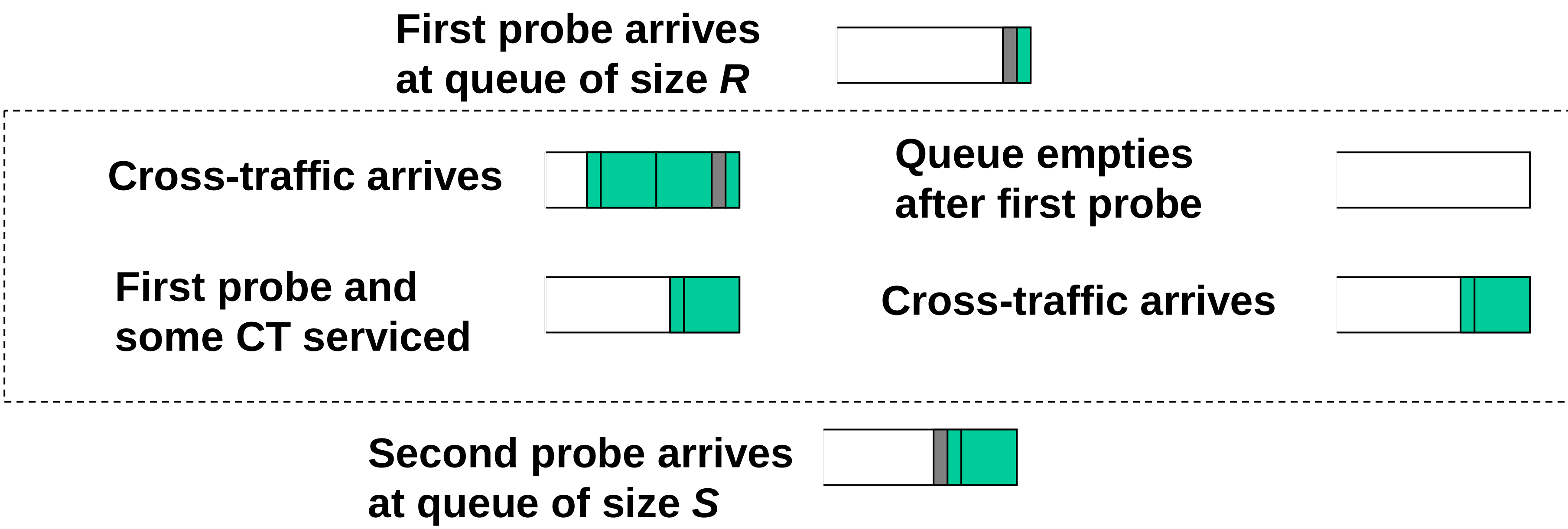
Small Probing Period *t*

- Assumption is that queue is busy between packets of pair
- Packet pair will not work for arbitrarily large *t*



How Large Can τ Be?

- Same packet pair delays
  - Different amounts of intervening cross-traffic



- τ cannot be more than transmit time of first probe!

Tradeoffs

- Small enough *t* not possible/applicable in multi-hop path
  - Under-utilized link of lower capacity before bottleneck
  - Queue sizes of other hops comparable in magnitude to *t*
- Larger *t* can be used only if some assumption made about cross-traffic!
  - Independent increments, long range dependent

Problem Setup

- ***t***-spaced packet pairs of size ***sz***
- Consecutive delays **R** and **S** of probes
  - Delay is same as encountered queue size
  - Remove need for clock synchronization later
- **A**(**T**) is amount (service time) of cross-traffic arriving at (single) bottleneck in time **T**
- What can we learn about the probability laws governing **A**(**T**)?

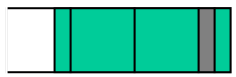
Outline

- Motivation
- Packet pair problem setup
- Describing the system
- Solving with i.i.d. assumption
- In practice
- Conclusions and future work

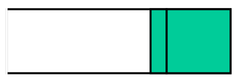
Busy vs. Idle Queue

First probe arrives at queue of size *R*

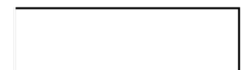
Cross-traffic arrives



First probe and some CT serviced



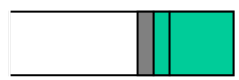
Queue empties after first probe



Cross-traffic arrives

queue is always busy

Second probe arrives at queue of size *S*



queue is idle at least once

$$S = R + sz + A(t) - t$$

$$S = \sup_{s \text{ in } [0,t]} A(t-s) - (t-s)$$

System equations

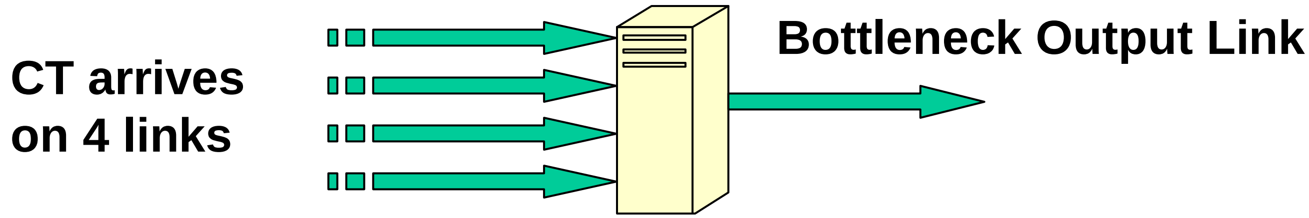
- Second delay, ***S*** is linearly related to first delay, ***R*** and amount of cross-traffic **OR**
- ***S*** is related only to amount of cross-traffic

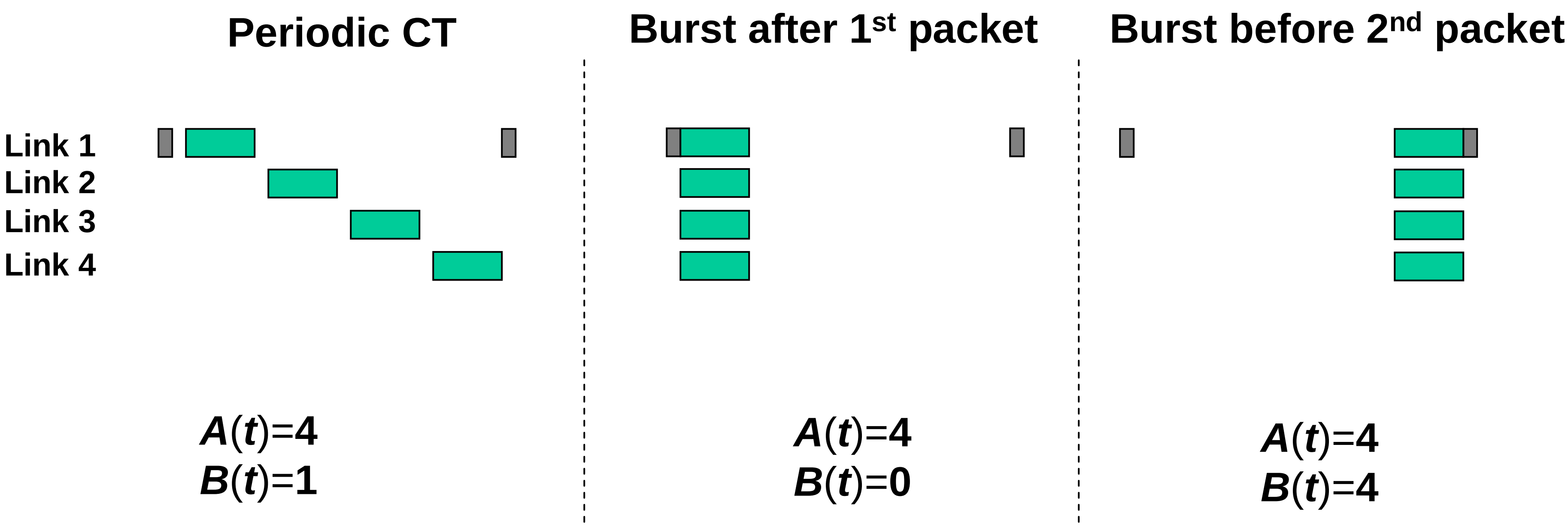
$$S = \max(R + sz + A(t) - t, B(t))$$

$$A(t) = \text{Amount of CT in time } t$$
$$B(t) = \sup_{s \text{ in } [0,t]} A(t-s) - (t-s)$$

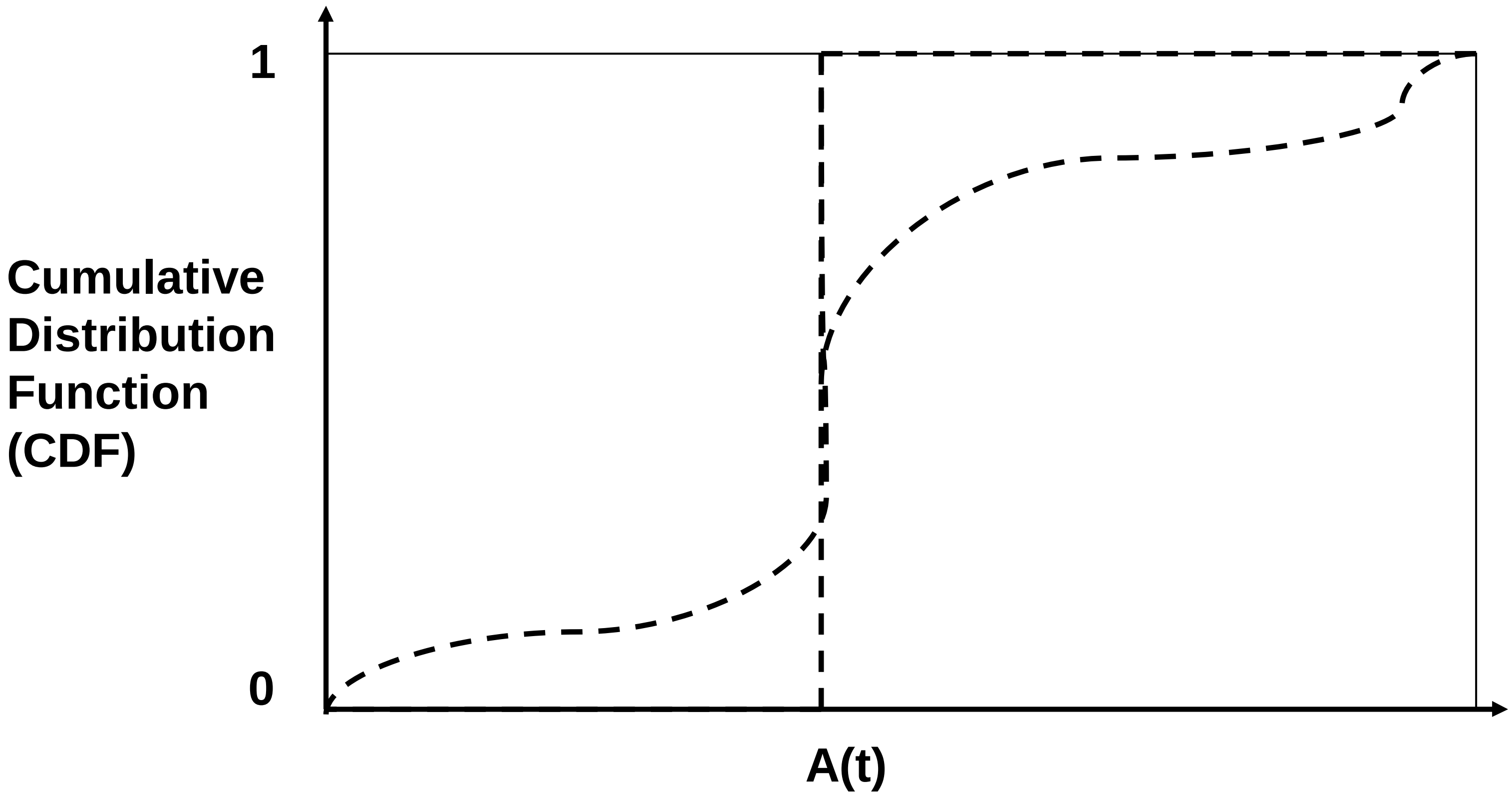
CT measured in time units of service time at bottleneck

$A(t)$ and $B(t)$


CT arrives
on 4 links
Bottleneck Output Link

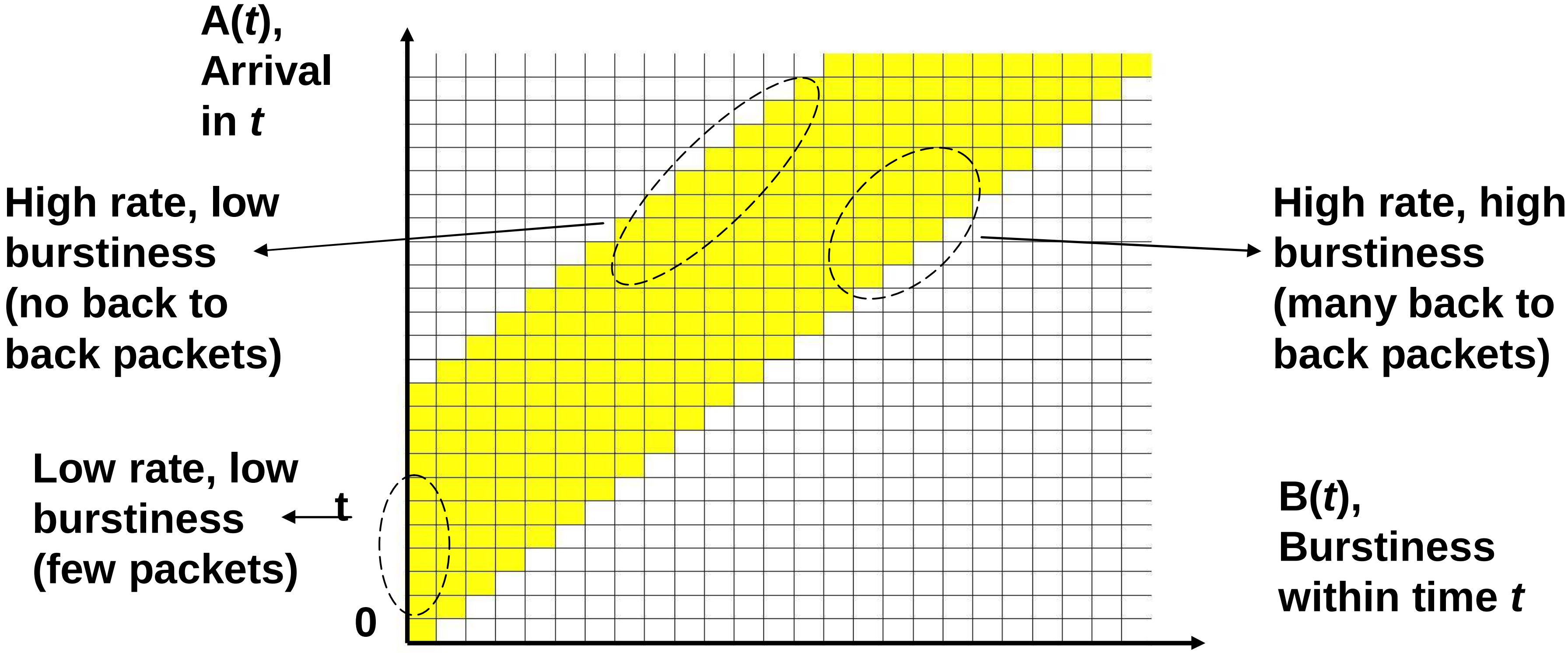

Periodic CT
Burst after 1st packet
Burst before 2nd packet
Link 1
Link 2
Link 3
Link 4
A(t)=4
B(t)=1
A(t)=4
B(t)=0
A(t)=4
B(t)=4

CDF of *A*(*t*)



- Cumulative distribution of *A*(*t*)
  - Arrival process in *t* time units
  - Step function if similar amounts of cross-traffic arrives every time period of size *t*

Joint Prob. Distribution of B(t),*A(t)*



- $A(t) > B(t) > A(t) - t$ (Density non-zero **only** in strip of size t)
  - $B(t)$ depends on the capacity of link
- Given $A(t)$, smaller $B(t)$ implies $A(t)$ amount of traffic is well-spread out within t time units

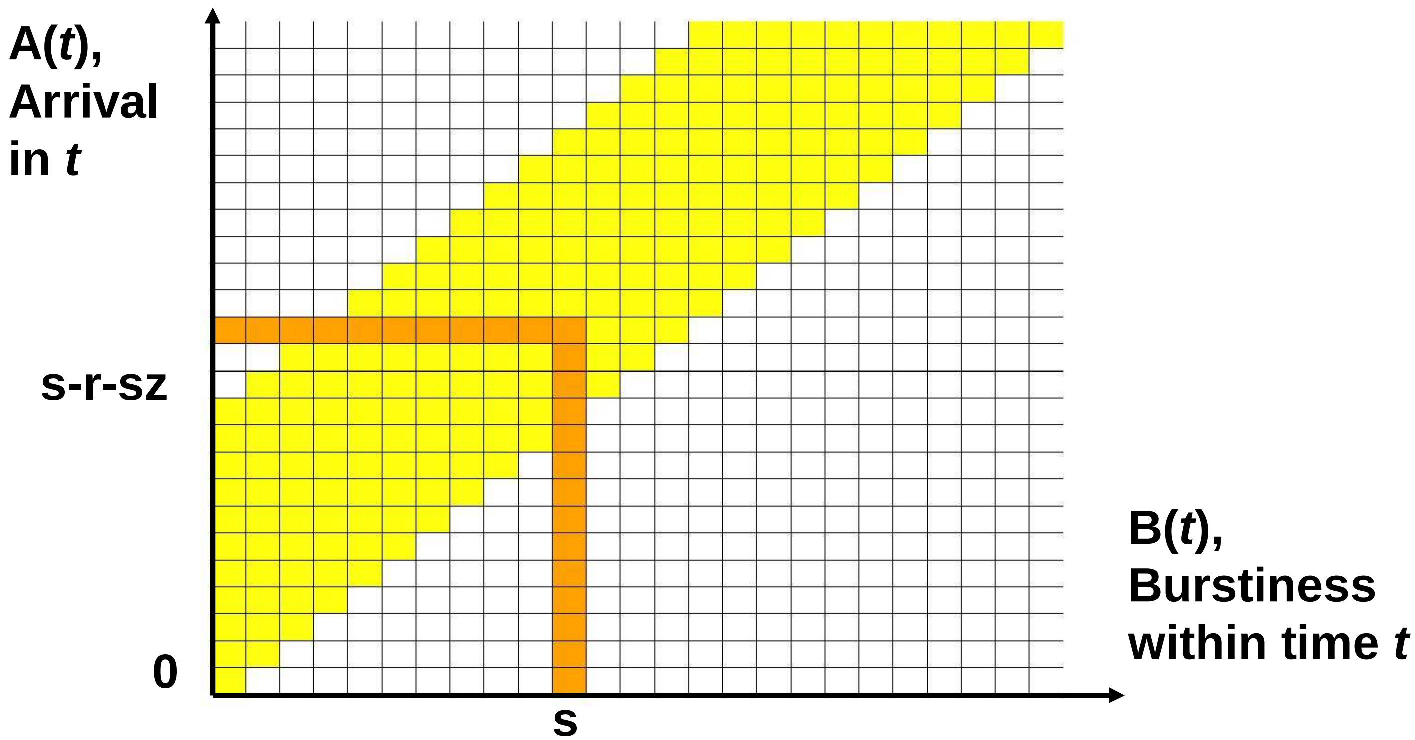
Outline

- Motivation
- Packet pair problem setup
- Describing the system
- Solving with i.i.d. assumption
- In practice
- Conclusions and future work

Solving System Equations

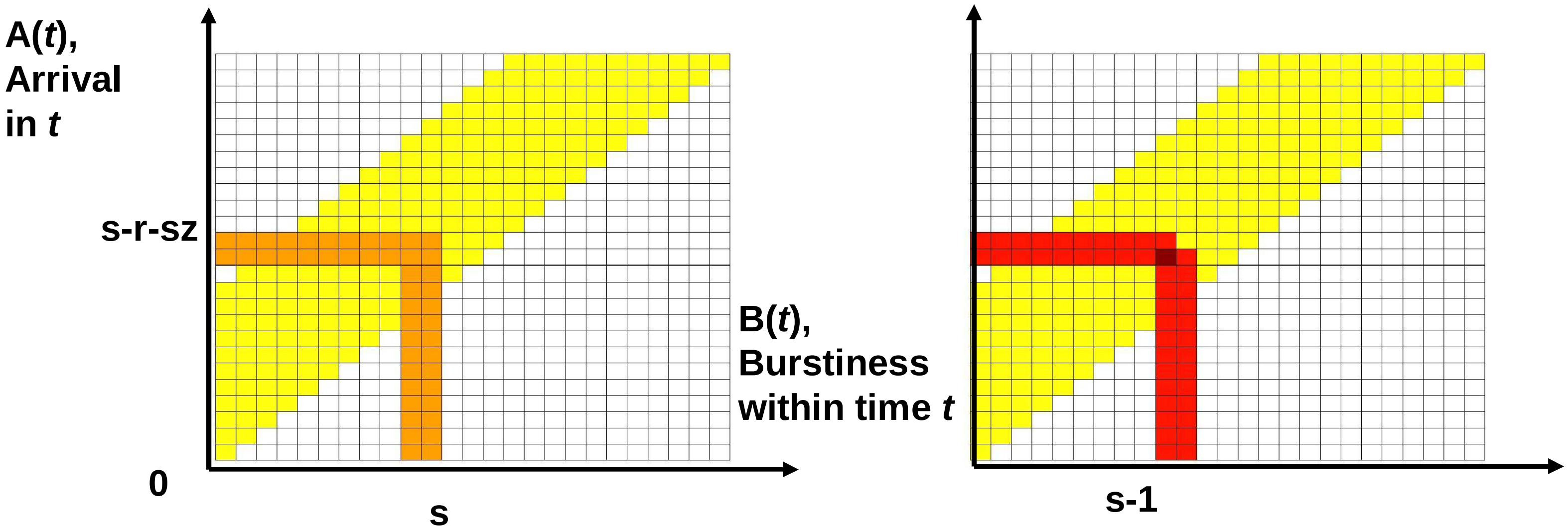
- Assumptions on CT needed for arbitrary ***t***
- Our assumption on CT
  - **A**(***t***) is i.i.d. in consecutive time periods of size ***t***
  - Traces from OC-3 (155Mbps link) at real router show that dependence between consecutive **A**(***t***) is 0.16 to 0.18
- Given delays ***R*** and ***S*** of many packet pairs
  - Use conditional probabilities $f_r(s)$=P(**S**=s|**R**=r)

Packet Pair Delays in (**B**,**A**) Space



- Consecutive delays (**r**,**s**) occur because cross-traffic had (**B**,**A**) anywhere on a right angle
- $f_r(s)$=P(**S**=s|**R**=r) is sum of (joint) probabilities along right angle

Resolving Density in (B,A) Space



Take packet pairs such that first delay R=r
$f_r(s)=P(S=s|R=r)$

+

Take packet pairs such that first delay R=r-1
$f_{r-1}(s-1)=P(S=s-1|R=r-1)$

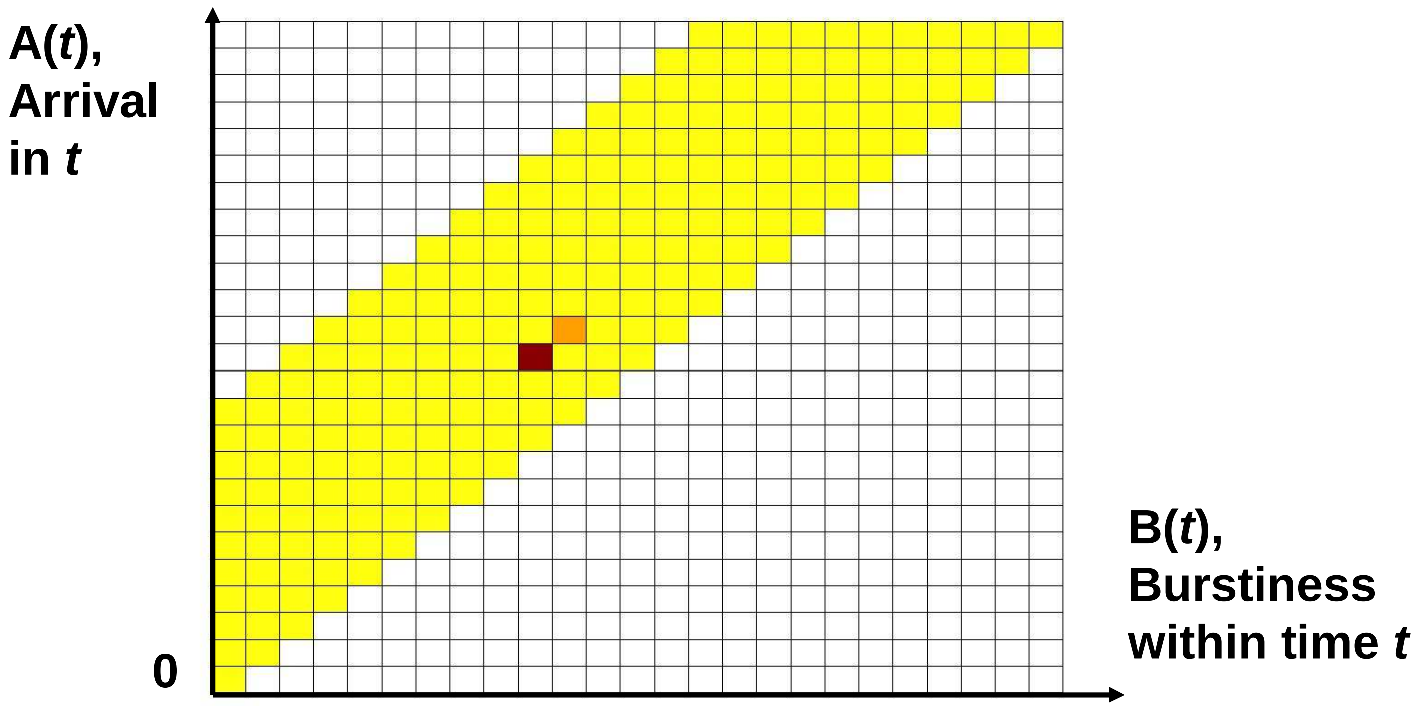
—

Take packet pairs such that first delay R=r
$f_r(s-1)=P(S=s-1|R=r)$

+

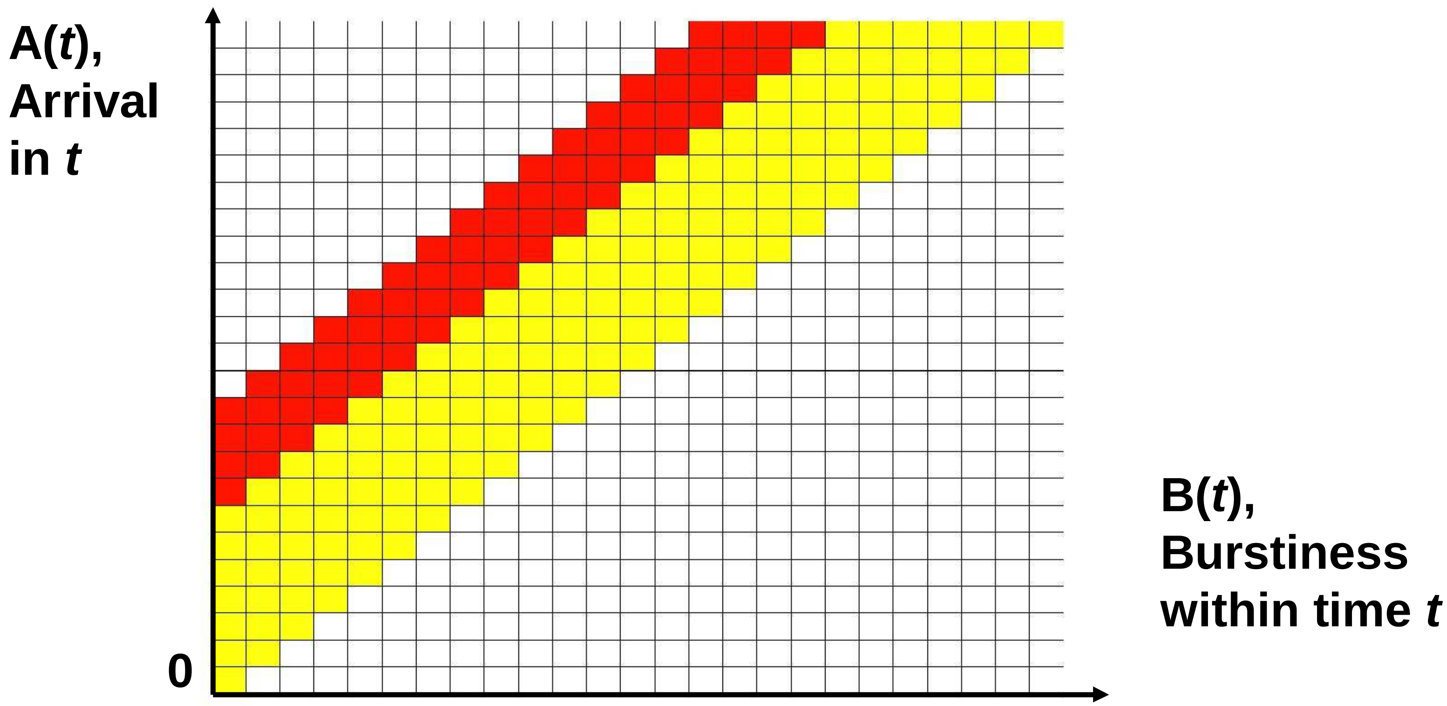
Take packet pairs such that first delay R=r+1
$f_{r+1}(s)=P(S=s|R=r+1)$

Resolving Density in (**B**,***A***) Space



- $f_r(s) + f_{r-1}(s-1) - f_r(s-1) + f_{r+1}(s)$
  - Difference in two densities adjacent along a diagonal
- Telescopic sum of differences along diagonal

Unresolvable Densities



- Width of unresolvable strip is probe size, effect of probe intrusiveness
- Joint distribution also gives us CDF of *A*(*t*)

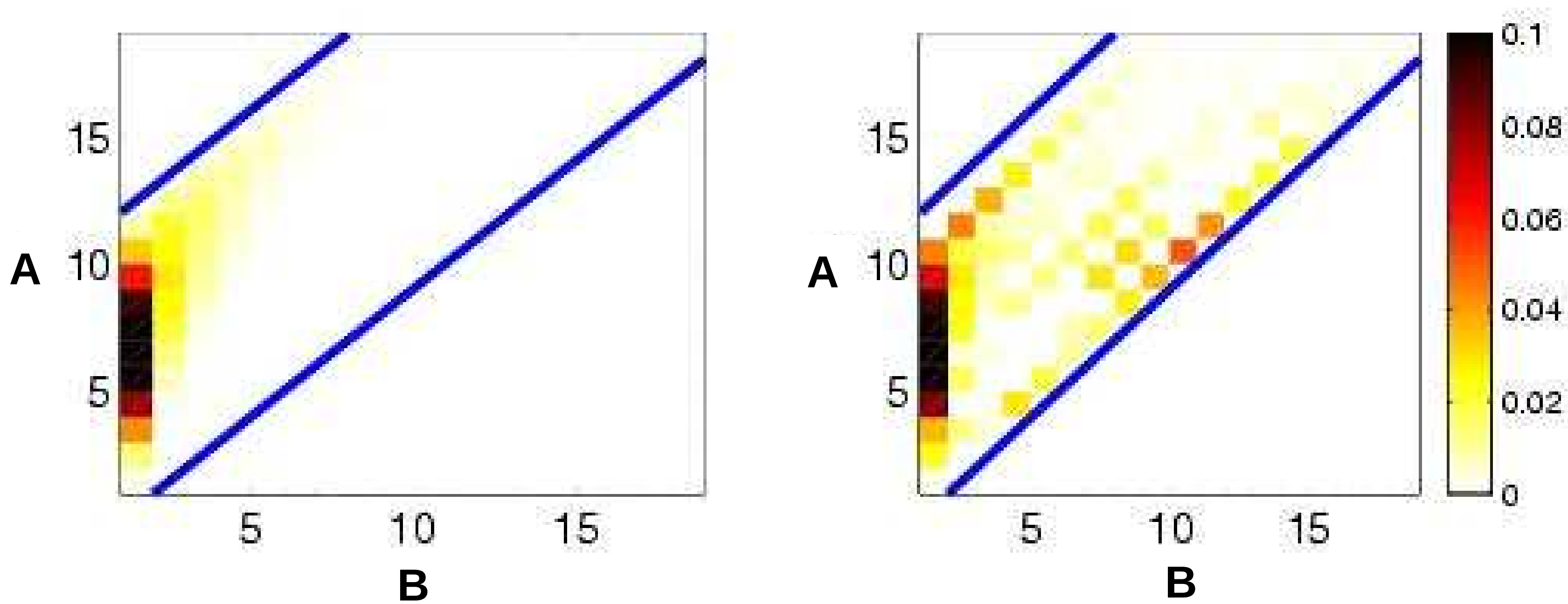
Outline

- Motivation
- Packet pair problem setup
- Describing the system
- Solving with i.i.d. assumption
- In practice
- Conclusions and future work

In Practice

- Absolute delays **R** and **S** not available
- Use minimum delay value observed and subtract from all observations
  - May not work if no probe finds all queues idle
  - Even capacity estimation may not work!
- Above limitations intrinsic to packet pair methods

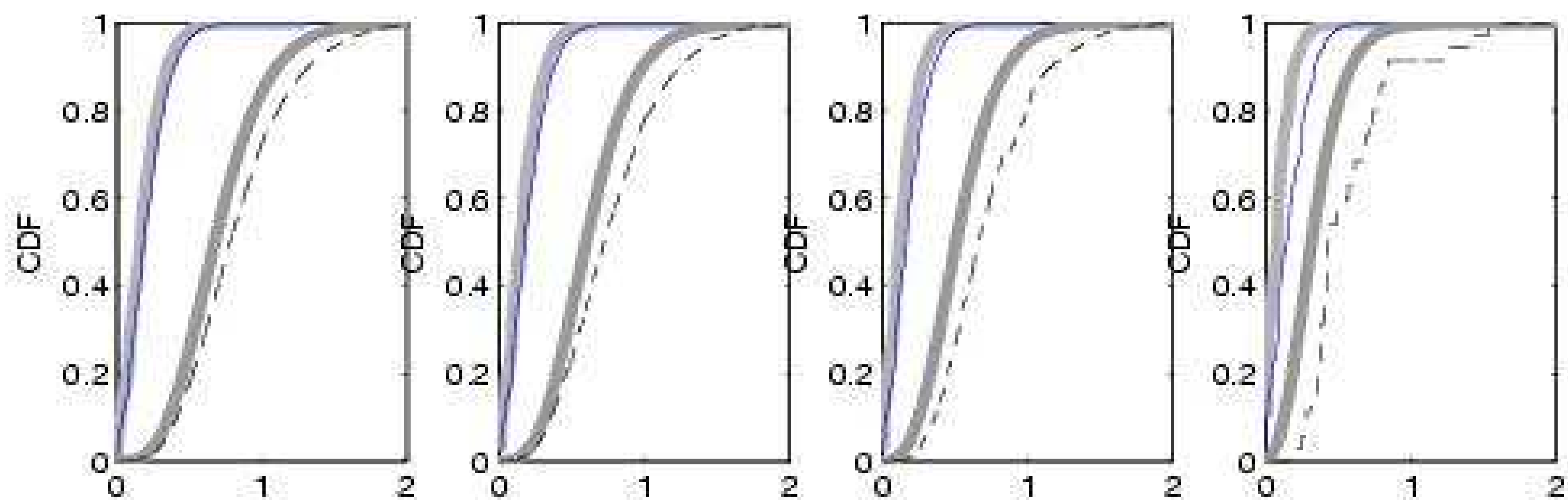
Estimation with Router Traces



- Router traces of CT (utilization – about 50%)
- *t* is 1ms; each unit is 155Mbps*0.1ms bits
- Errors due to
  - Discretization
  - *A*(*t*) not entirely i.i.d.

Estimation of *A*(*t*)

Decreasing Cross-Traffic Rate



Service time of cross-traffic arriving in 0.25, 1ms (milliseconds)

- About 10-15% error, in general
- Larger errors with lower CT rates
  - larger delay values less likely

Outline

- Motivation
- Packet pair problem setup
- Describing the system
- Solving with i.i.d. assumption
- In practice
- Conclusions and future work

Conclusions and Future Work

- Packet pair methods for (single hop) avail. b/w work either with very small ***t*** (or)
- With assumptions on CT
  - I.I.D. assumption reasonable
  - Almost complete estimation possible
  - What other assumptions possible?
- Packet pair does not saturate any link
  - Hybrid of packet pair and saturating packet train methods?

Backup Slides

A(*t*) and *B*(*t*)

Intervening cross-traffic	***A(t)*** (Arrival in *t* time units)	***B(t)*** (Arrival within *t* units)
Periodic traffic of rate ***u.C***	***u.t***	***u.t-t***
Packet burst sized ***u.t*** after first packet	***u.t*** (Size of burst)	Size of burst - *t*
Packet burst sized ***u.t*** just before 2nd pkt.	***u.t*** (Size of burst)	Size of burst

Using conditional probabilities

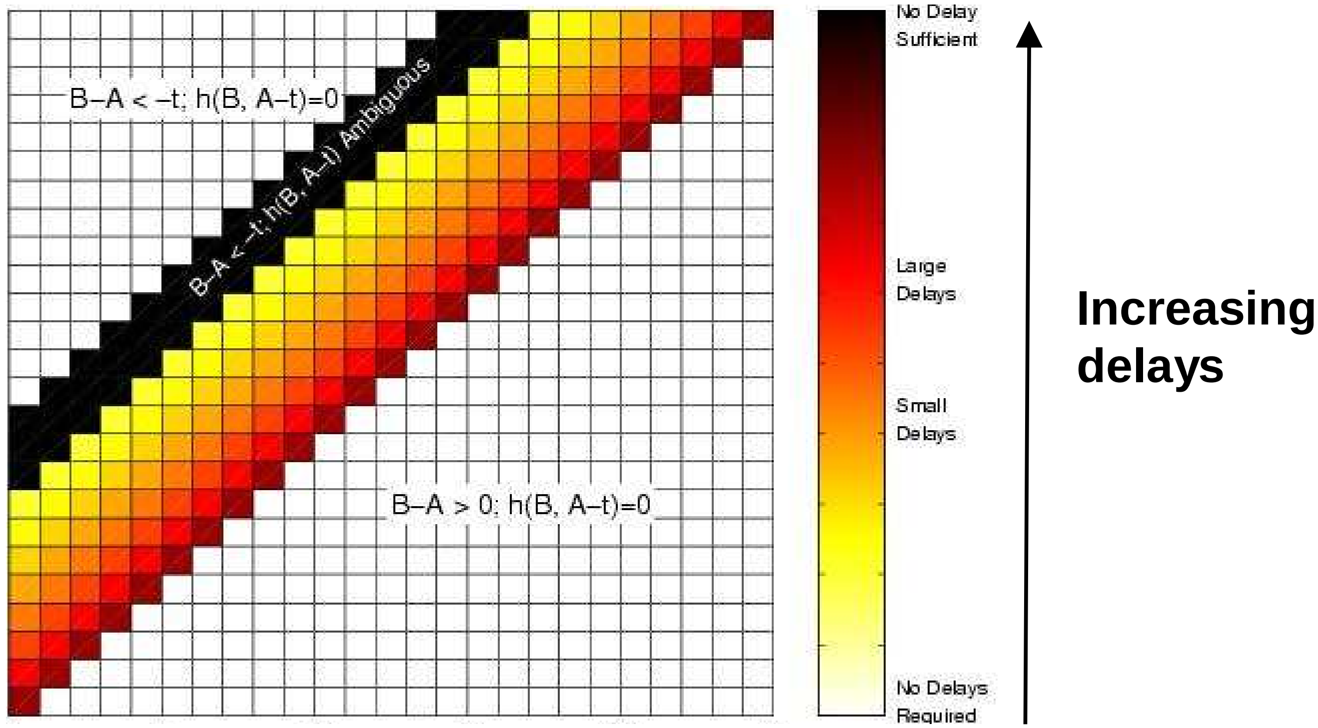
- $f_r(s)=P(\mathbf{S}=s|\mathbf{R}=r)$ is a right angle
- $F_r(s)=P(\mathbf{S}\leq s|\mathbf{R}=r)$ is a rectangle
- Horizontal bar is difference between rectangles

$$C(s-r-x, s) = F_r(s) - F_{r+1}(s)$$

- Density in the (B,A) space is the difference between two horizontal bars

$$h(s-r-x, s) = F_r(s) - F_{r+1}(s) - [F_{r-1}(s-1) - F_r(s-1)]$$

Required Delays for Estimation



Ambiguities of size ***sz*** still allow the resolution of distribution of ***A***(***t***)