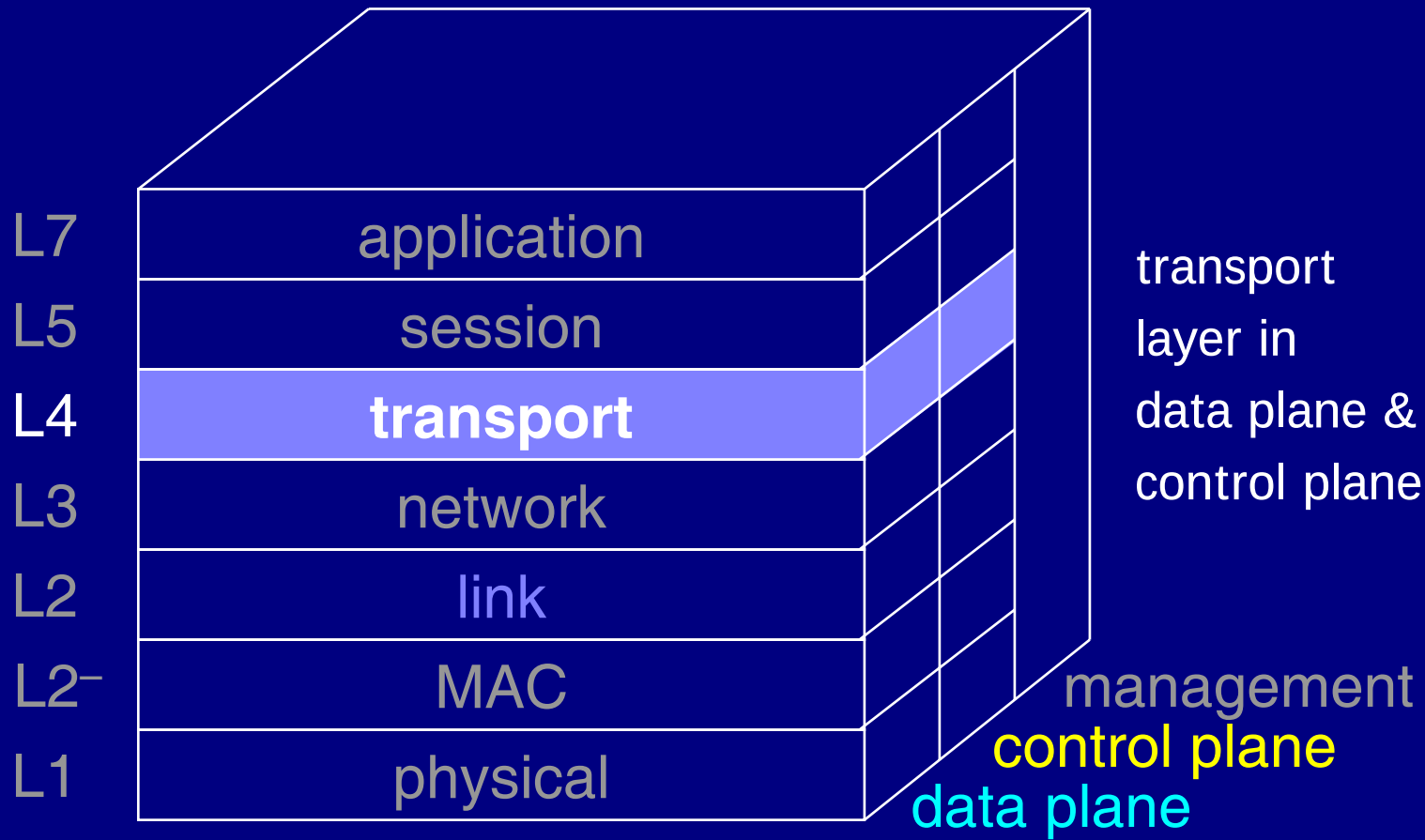
KU EECS 800
Survivable, Resilient, and DT Networking Transport Layer

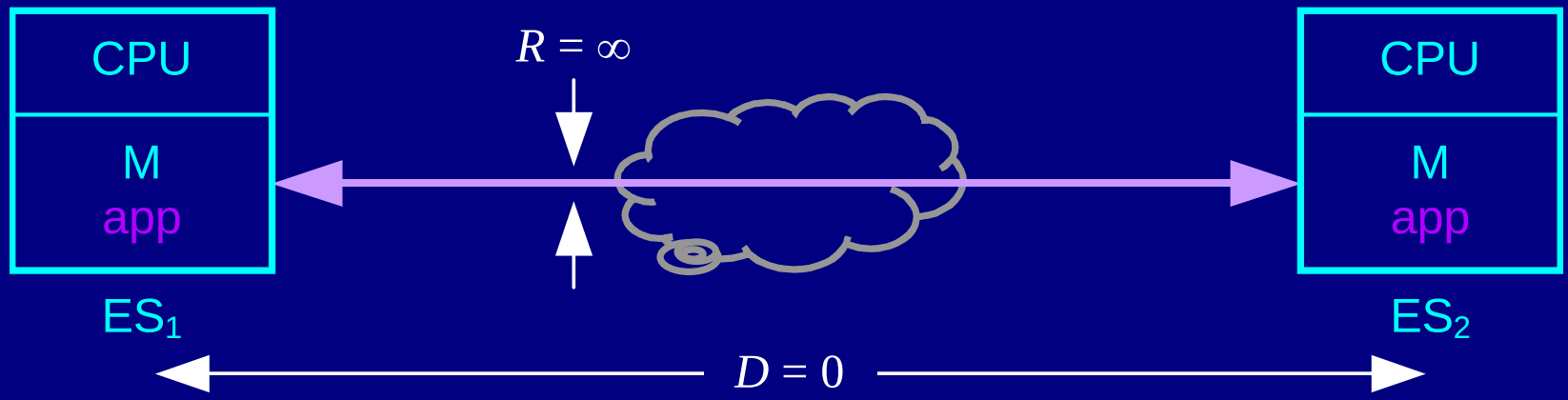
James P.G. Sterbenz
The University of Kansas Department of Electrical Engineering and Computer Science

jpgs@eecs.ku.edu
3036 Eaton
++1 785 864 8846
http://www.ittc.ku.edu/~jpgs/courses/eecs800

Transport Layer
Layer/Plane Cube Model

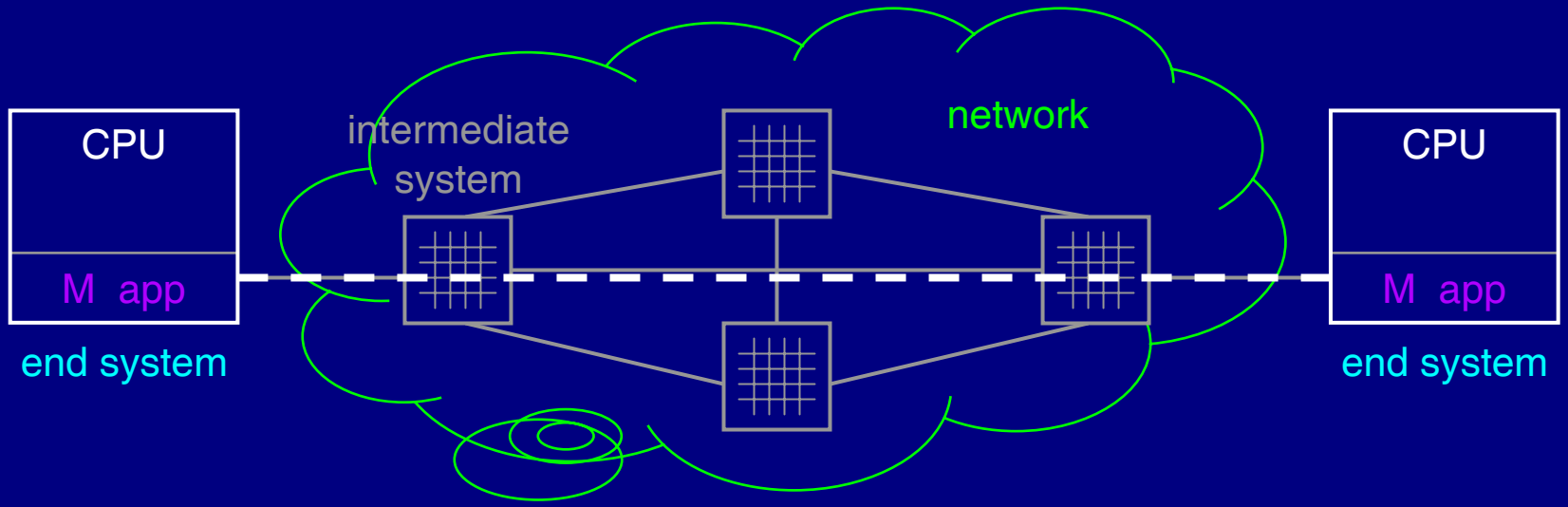


Transport Layer
Ideal Network Model and Motivation



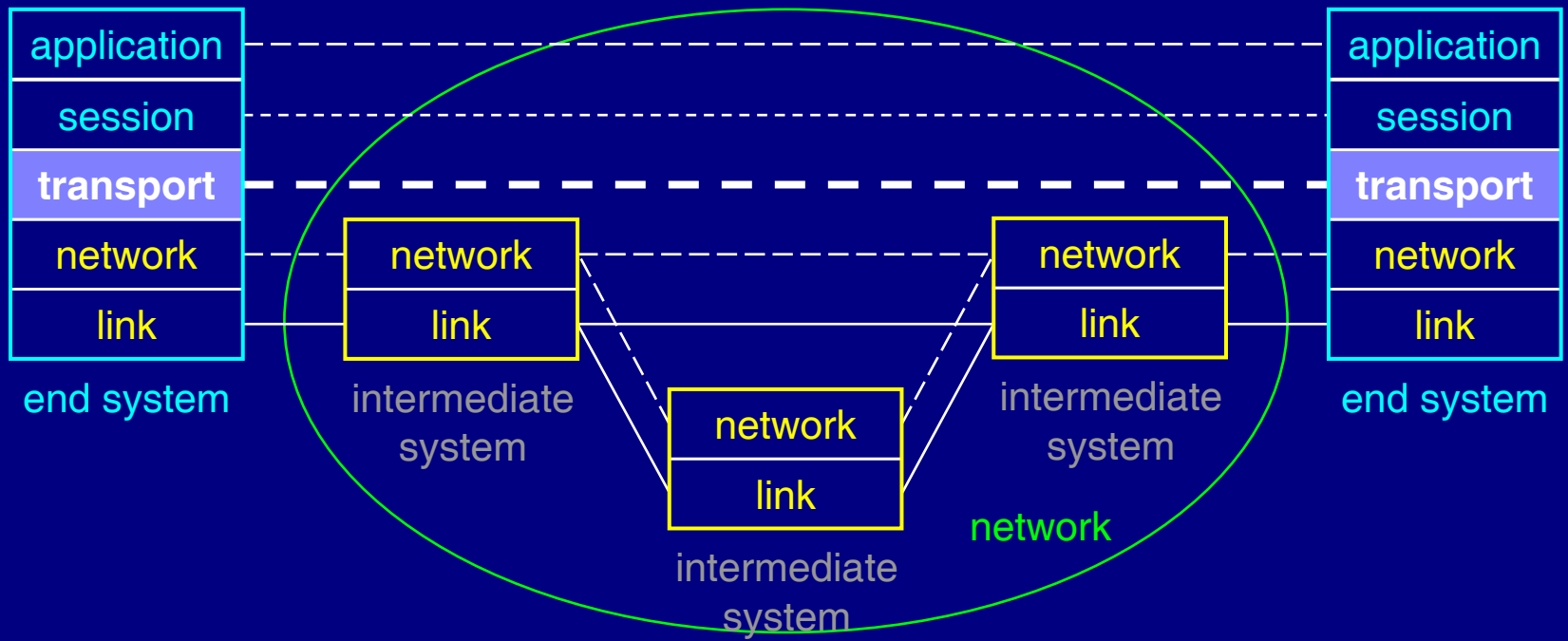
- Ideal network model
  - zero end-to-end delay
  - unlimited end-to-end bandwidth
  - perfect end-to-end transmission (no errors)
- Real networks not ideal
  - need an end-to-end protocol

Transport Layer
Transport Association Definition



- *Transport association* between *end systems*
  - *transport connection* in the case of connection state
  - may be point-to-point or multipoint
  - recall: the network is translucent: delay, errors, loss

Transport Layer
End-to-End Protocol



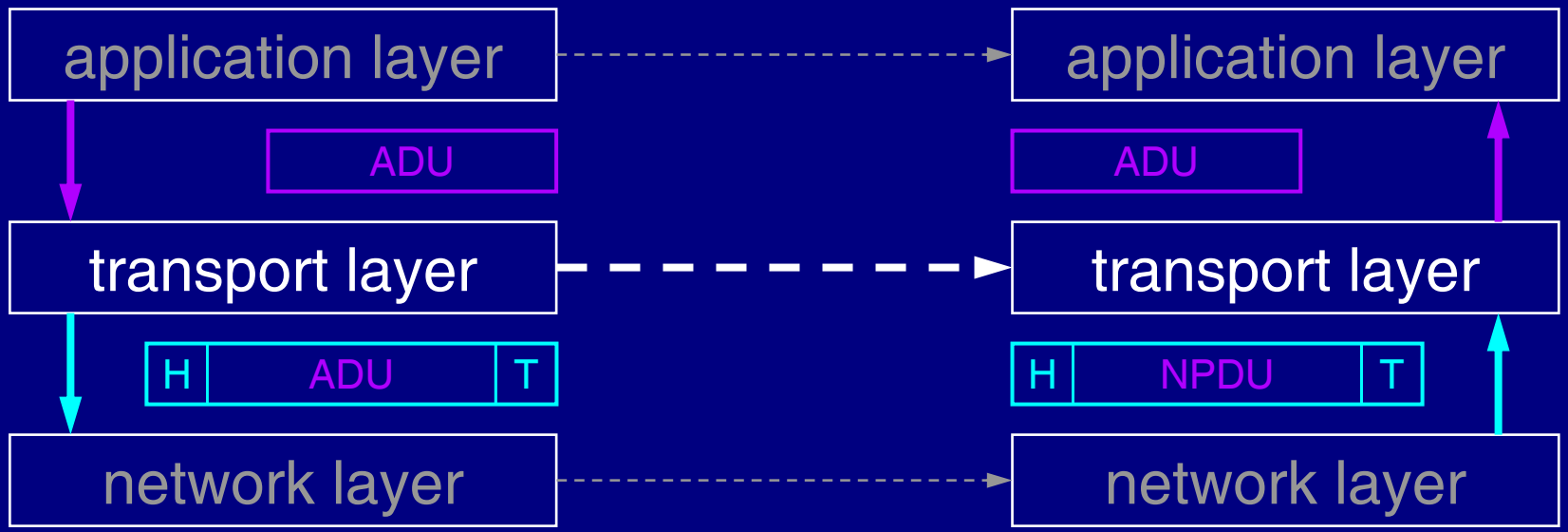
- Transport protocol
  - is responsible for end-to-end transfer of a TPDU

Transport Layer
Service and Interfaces

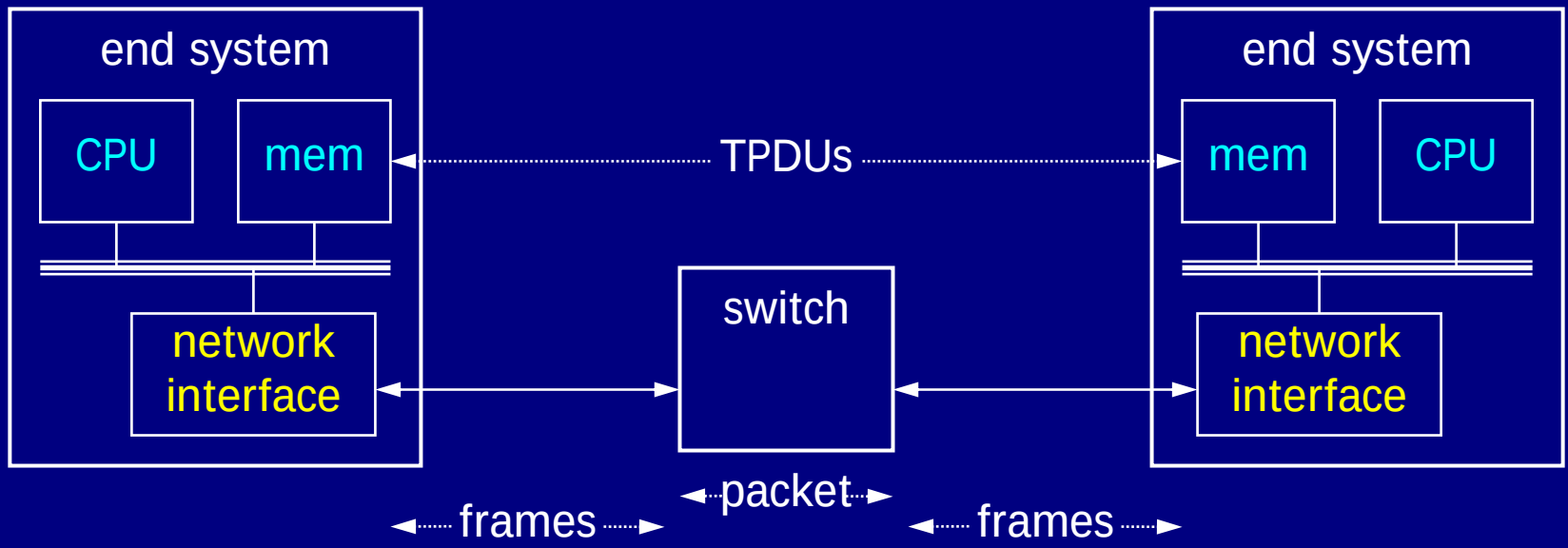
- Transport layer is E2E analog of HBH link layer
  - link layer (L2) transfers packets HBH
  - network layer (L3) determines the path of hops
- Transport layer (L4) service to application layer (L7)
  - transfer PDU E2E (end-to-end)
    - sender: encapsulate ADU into TPDU and transmit
    - receiver: receive TPDU and decapsulate into ADU
  - multiplexing to application on end system
  - optional reliability: error checking/correction/retransmission
    - recall end-to-end arguments:
      E2E reliability with HBH error control only for performance
    - may be done by application (A2A: application-to-application)
  - flow control and assistance for network congestion control

Transport Layer
Service and Interfaces

- *Transport PDU* encapsulates *application data unit*
  - ADU – application data unit
  - TPDU – transport protocol data unit
    - TPDU = header + ADU + optional trailer (protocol dependent)

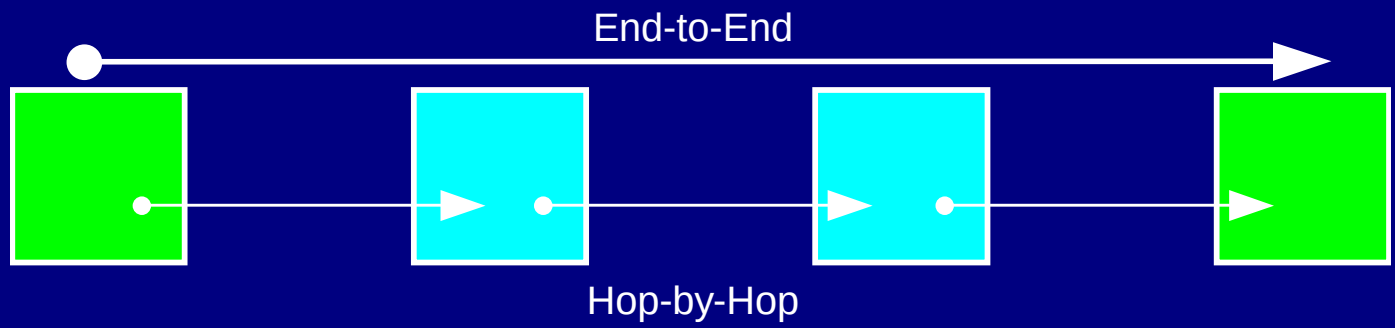


Transport Layer
Functional Placement



- Transport layer functionality only in end system
  - end system: host software or network interface (NIC)

Transport Layer
Hop-by-Hop vs. End-to-End Semantics



- Hop-by-hop functions
  - link layer protocols
  - link compression and FEC
  - network forwarding
  - link / subnet error control
  - embedded protocols (e.g. protocol boosters)
- End-to-end functions
  - transport protocols
  - source routing
  - end-to-end encryption
  - session protocols
  - application protocols

Transport Layer
End-to-End Argument

- Hop-by-hop function composition ≠ end-to-end
- Examples
  - HBH encryption has data in clear inside network nodes
  - HBH link error control doesn't cover network layer errors

End-to-End Argument T-3

Functions required by communicating applications can be correctly and completely implemented only with the knowledge and help of the applications themselves. Providing these functions as features within the network itself is not possible.

Transport Layer
Hop-by-Hop Performance Enhancement

- E2E functions replicated HBH if performance benefit
- Example
  - per link error control in high bandwidth-×-delay networks reduce E2E control loop when error

Hop-by-Hop Performance Enhancement Corollary T-3A

It is beneficial to duplicate an end-to-end function hop-by-hop if the result is an overall (end-to-end) improvement in performance.

Resilience Strategy
End-to-End Resilience & Disruption Tolerance

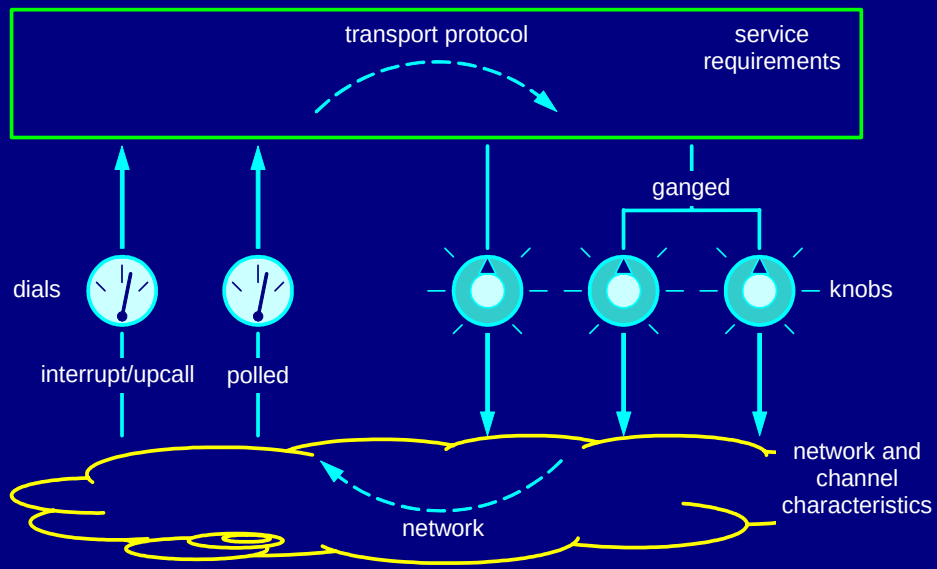
- Introduction to survivability and disruption tolerance
- Resilience strategy
  - network survivability and resilience
    1. maintain survivable connectivity when possible
    2. survivable communication even when not connected
    3. resilient network mechanisms
    4. technologies to enhance survivability
  - end-to-end resilience and disruption tolerance
  - disruption-tolerant user-controlled adaptive applications
- Summary

E2E Survivability and DT
Internet Transport Protocols

- Current transport protocols assume
  - strongly connected stable end-to-end paths
  - symmetric path
  - reliable medium
- TCP and SCTP
  - combined feedback error+flow+congestion control
    - reliable ACK stream required for self-clocking
  - unable to discriminate channel loss from congestion
    - congestion indicated by loss (switch drops)
    - channel loss results in throttling source
      - *wrong* response in uncongested network

E2E Survivability and DT

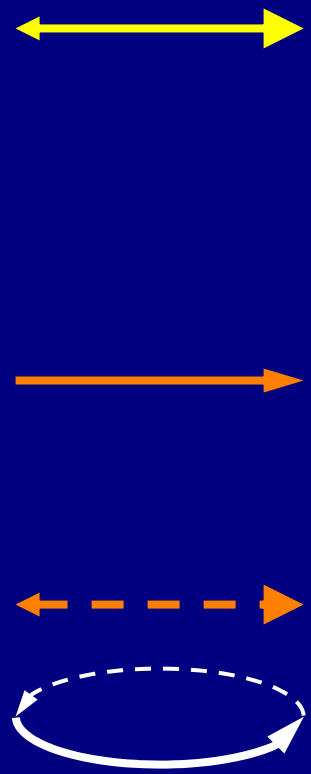
Knobs and Dials



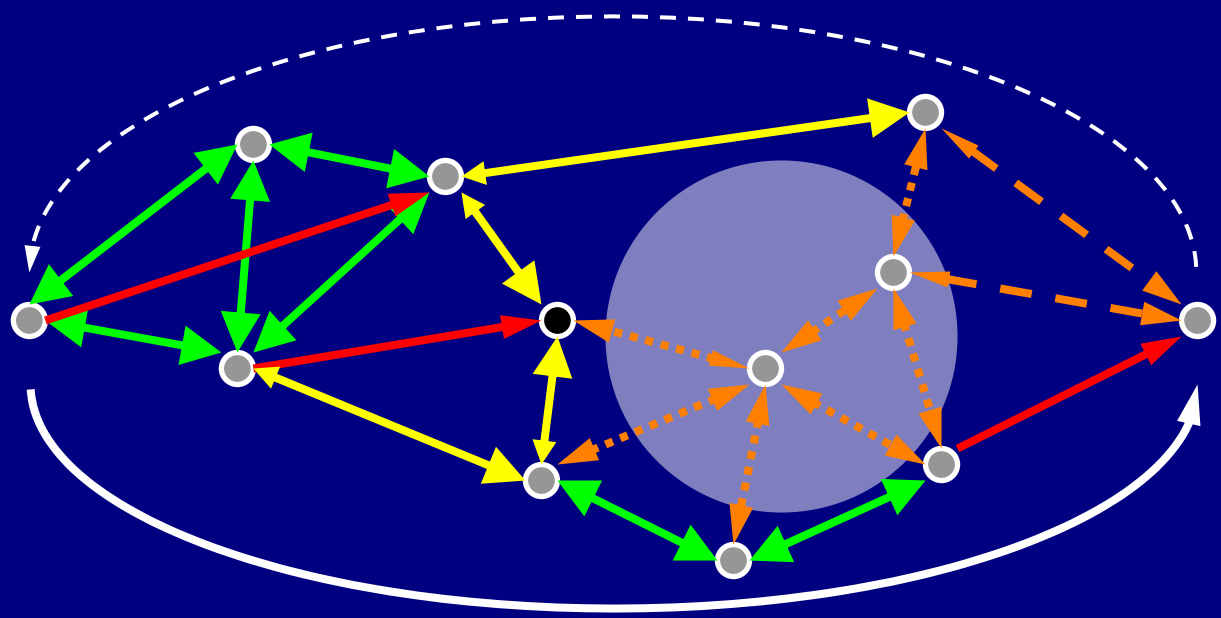
- Transport protocols should interact with network
  - *dials* feed back topology and path characteristics
  - *knobs* influence lower layer parameters and behaviour
  - example: error control based on loss characteristics

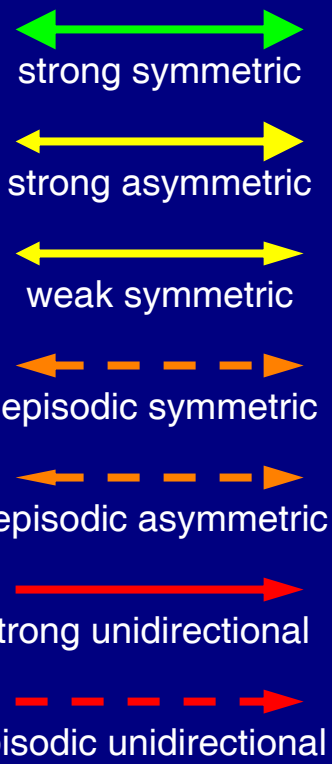
E2E Survivability and DT
Asymmetric Paths

- Asymmetric channels result from
  - asymmetric transmission power
    - intentional (LPD) or available power
  - antenna characteristics and directionality
  - terrain and location
- Unidirectional channels result from
  - asymmetric transmission power
  - radio silence
- Path connectivity may be episodic
- Asymmetric and unidirectional E2E
  - concatenation of channels
  - forward and reverse may follow different paths



E2E Survivability and DT
Asymmetric End-to-End Paths

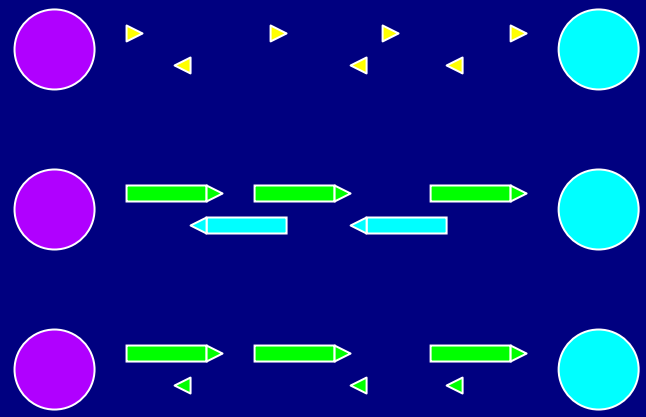




- Asymmetric end-to-end path challenges
  - how to find best paths through network
  - how to characterise entire path

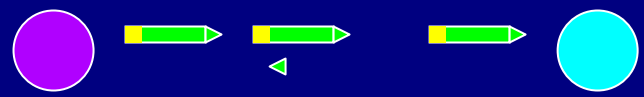
E2E Survivability and DT
Bidirectional Paths

- Bidirectional path *required* for
  - pairwise synchronisation
    - signalling messages
  - bidirectional data communication
    - application issue
  - closed-loop feedback control
    - ACKs for *reliable* data transfer
    - *even* if data transfer unidirectional



E2E Survivability and DT
Open Loop Control

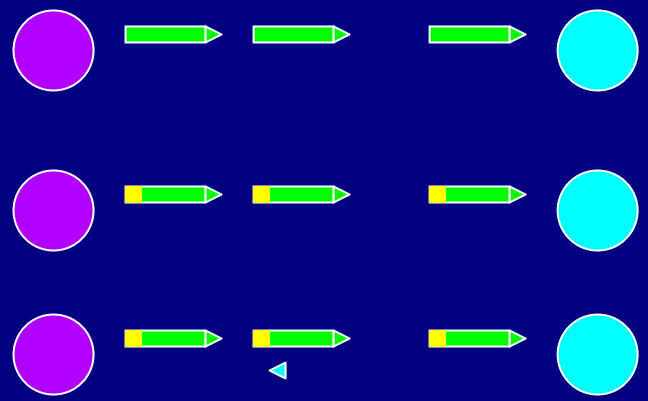
- Survivability with asymmetric channels needs:
  - open-loop control
  - with feedback only when necessary
- Open-loop rate control
  - congestion feedback from network only when necessary



E2E Survivability and DT
Open Loop Error Control

- Open-loop error control: FEC
  - unreliable transfer
    - optional per link FEC
  - quasi-reliable transfer
    - FEC for probabilistic reliability
  - reliable transfer
    - requires bi-directional path
    - infrequent adaptive selective ACKs
    - distinct from:
      - flow control (E2E)
      - congestion control

  note: SCTP does none of this



E2E Survivability and DT
End-to-End Transport Mechanisms

- Flow control
  - rate that *receiver* can accept
  - purely end-to-end
- Congestion control
  - rate that *network* can accept without congesting
  - network feedback to end systems
- Error control
  - retransmission of corrupt and lost packets
  - link and network-based error characteristics
  - application-dependent reliability requirements

E2E Survivability and DT
Explicit Loss/Congestion/Delay Discrimination

- Absence of expected packet or ACK arrival
  - three *distinct* and unrelated causes:

1. Congestion: packet dropped in network
   - congestion control: queue overflow (tail drop)
   - congestion avoidance: intentional packet drop
2. Corruption: packet lost or delivered corrupted
   - channel error causing bit errors
3. Delay: packet arrival later than expected
   - store-and-forward delays in disruption tolerant network
   - long path
     - speed-of-light delay in delay-tolerant network
     - very long path around disruption

E2E Survivability and DT
Discrimination and Explicit Notification

- Discrimination and proper response essential:
  - congestion ⇒ back off
  - corruption ⇒ retransmit
  - delay ⇒ wait or retransmit via lower delay path
- Explicit notification
  - ECN: explicit congestion notification
  - ELN: explicit loss notification (due to corruption)
  - ELN *cannot* be determined from ECN (and vice versa)
    - packet that first causes congestion may then be corrupted

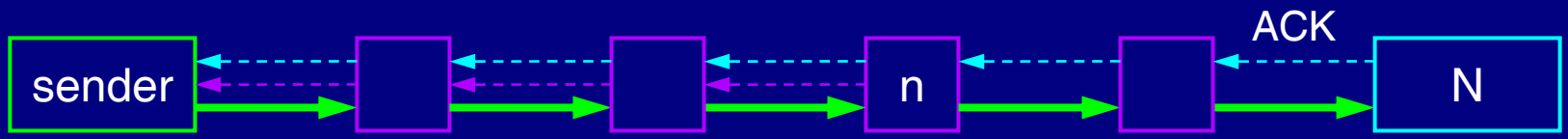
E2E Survivability and DT
Error Control Mechanisms

- Mechanisms
  - observation that error has occurred (detection)
  - notification of error
  - decision on what response to take
  - action to correct error
- Taxonomy of mechanisms
  - each mechanism may have different implementations...
    ...(e.g. E2E vs. HBH)...
    ...but they are frequently related
- ETEN: explicit transport error notification

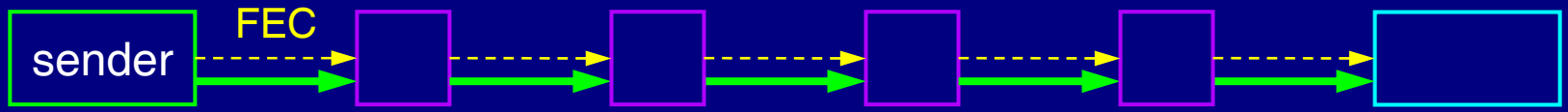
ETEN Taxonomy
Determinism and Granularity

- Determinism
  - deterministic (take action based on specific corruptions)
  - probabilistic (e.g. throttle source x% of the time)
- Granularity
  - PETEN: per packet response
  - CETEN: cumulative error rate response
- Control feedback ...
- Control locus ...
- Control band ...
- Control direction ...

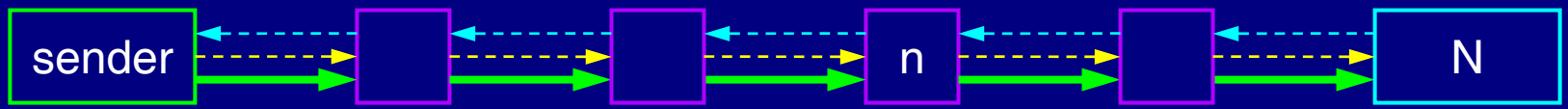
ETEN Taxonomy
Control Feedback



- Closed loop feedback from notifier nodes {n,N}
  - (N)ACK from end system or switch (e.g. congestion drop)

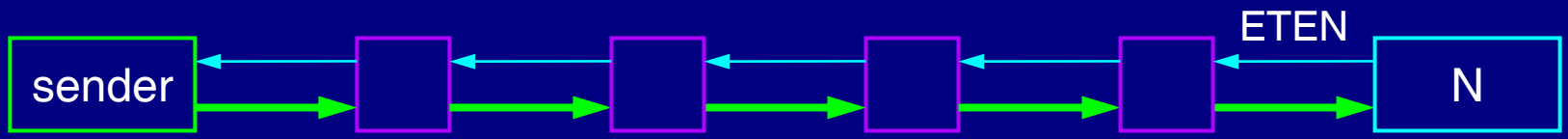


- Open loop
  - Unreliable or FEC for statistical reliability, rate control

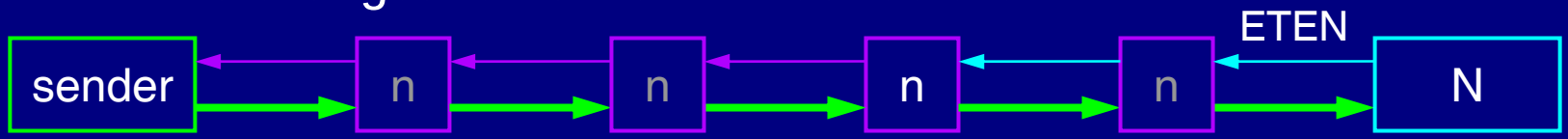


- Hybrid open + closed loop
  - FEC for statistical reliability + (N)ACKs as needed

ETEN Taxonomy
Control Locus



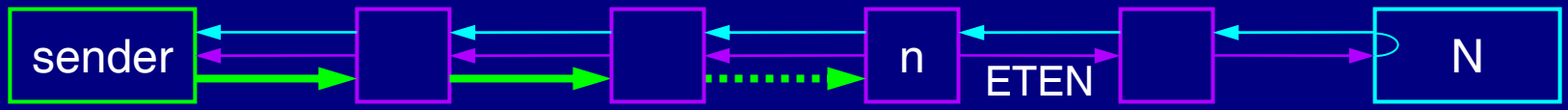
- End-to-end
  - no changes to network infrastructure



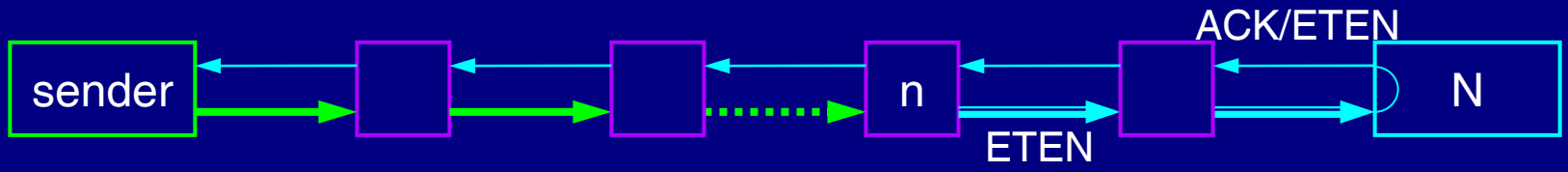
- Some hop-by-hop
  - deployed as needed (e.g. optional n at wireless access links)

- All hop-by-hop (N indicates *required* notifier functionality)
  - deployment challenges: impractical for Internet as a whole

ETEN Taxonomy
Control Band

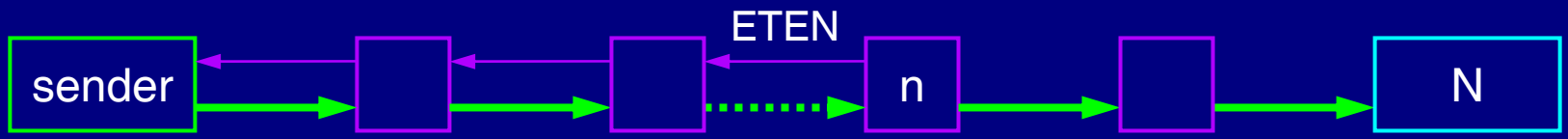


- Out-of-band
  - ETEN signalling messages
    - may be forward or backward

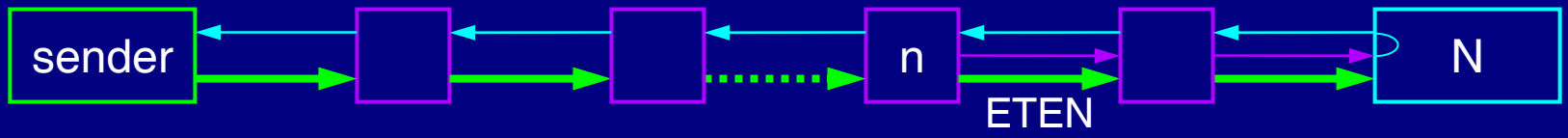


- In-band
  - ETEN information carried by packets; options:
    - corrupted packets not dropped but marked (violates RFC 1812)
    - carried by subsequent packets in flow (header field or option)

ETEN Taxonomy
Control Direction



- Backward
  - ETEN messages returned by switches
    - out-of-band unless inserted in reverse flow
    - requires HBH ETEN



- Forward
  - ETEN forwarded to receiver; turned back to sender
    - may be in- or out-of band
    - longer delay in response for interactive over high bw-×-delay

ETEN Simulation
TCP with Oracle ETEN

- TCP Oracle ETEN
  - instantaneous notification of corruption-based loss
  - perfect per packet loss discrimination
  - perfect response: retransmit rather than throttle
    - if loss within RTO
  - upper bound of possible benefit for bulk traffic
    - does *not* cause immediate retransmission:
      would matter for transactions but not for bulk transfer
- Simulation
  - ns-2 **`FullTcpSack`**
  - 536B segments, 30 min. runs

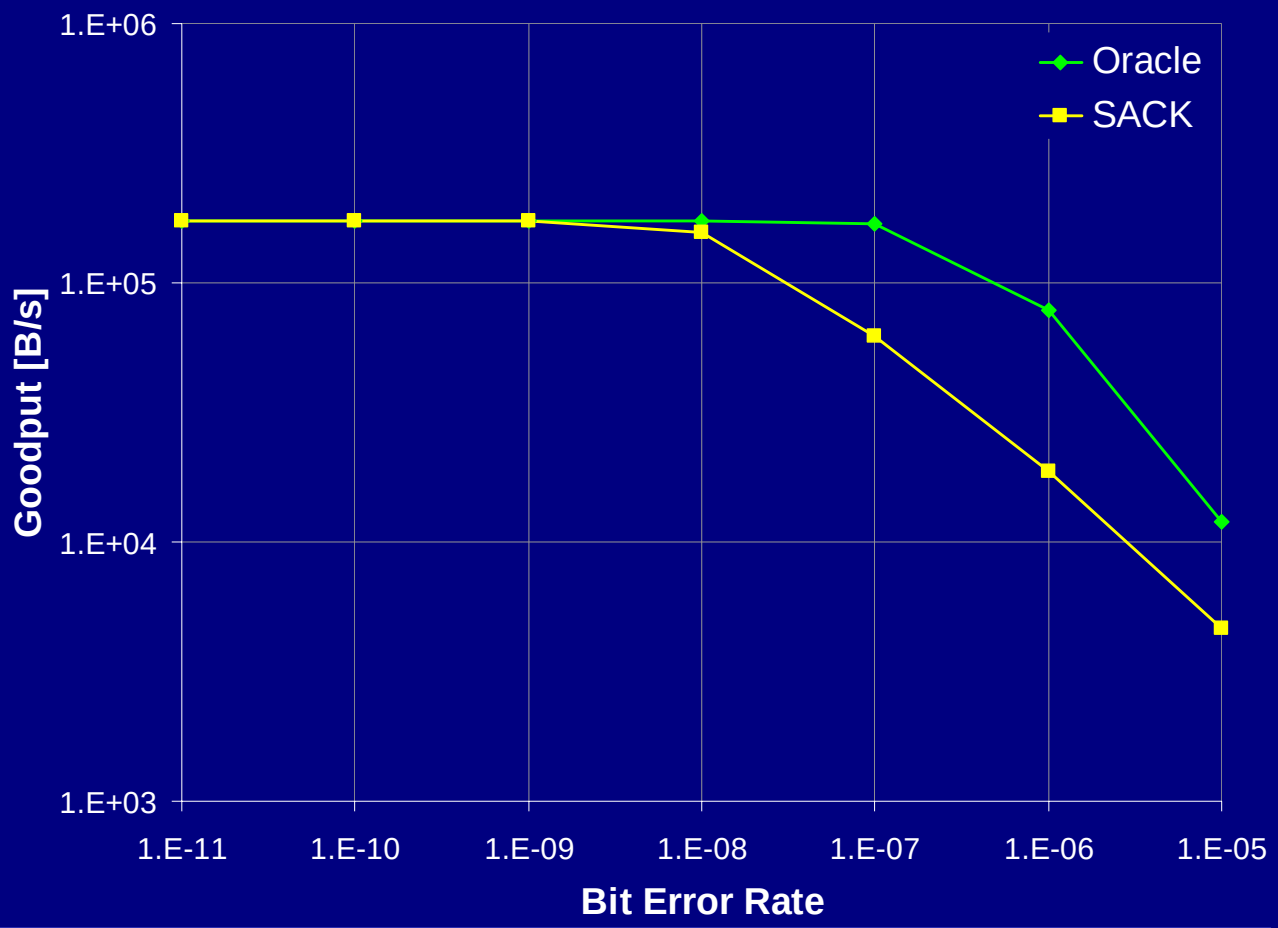
ETEN Simulation

TCP with Oracle ETEN

LTN:
250ms 1-way
1.5Mb/s

single flow

30 min. run
536B segment
2400 seg window



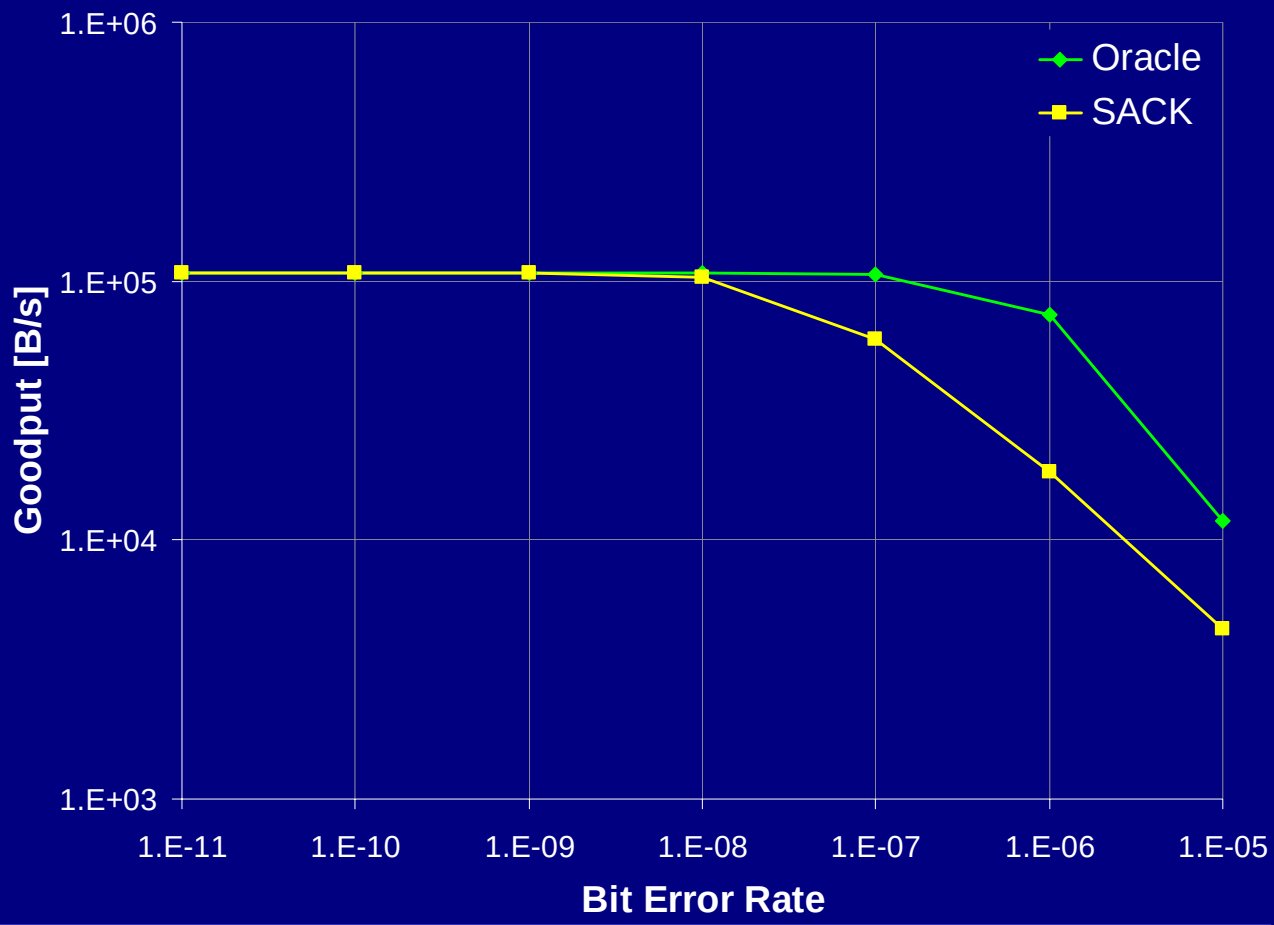
ETEN Simulation

TCP with Oracle ETEN

LTN:
250ms 1-way
1.5Mb/s

competing:
4×UDP
CBR 0.25Mb/s
on/off
exp mean 0.5s

30 min. run
536B segment
2400 seg window



Assignment for Next Class

- Read [KS+2004] and from W-DTN [JL+2005, HC+2005, TP+2005, LF+2005]
- Follow references in [KS+2004]
- Continue researching scenarios and past failures
  - e.g. NE power failure, 9/11, SBB, Hinsdale CO, *others*
- Begin to think about area of interest
- Office hours: TR 11:00–12:00 Eaton 3036
  - others *needed*?
- Makeup class Wed. 7 Aug. 09:30
  - semester schedule to be discussed Tue. 6 Aug.

End of Foils