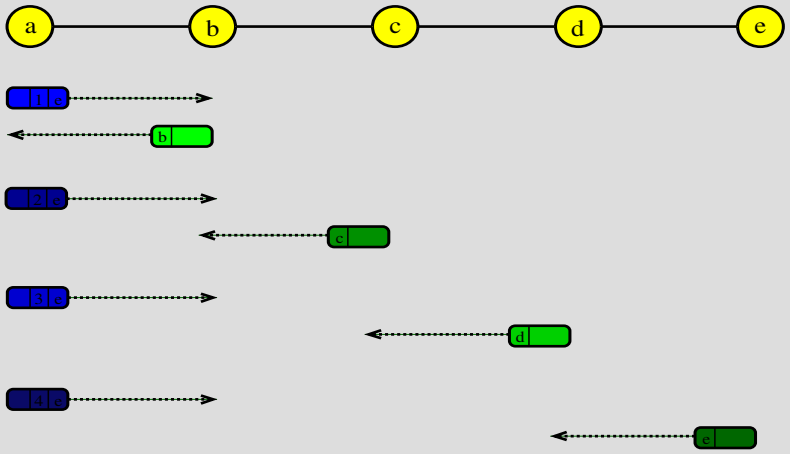
# Optimising Path Discovery

## Doubletree and Paris Traceroute

Tony McGregor

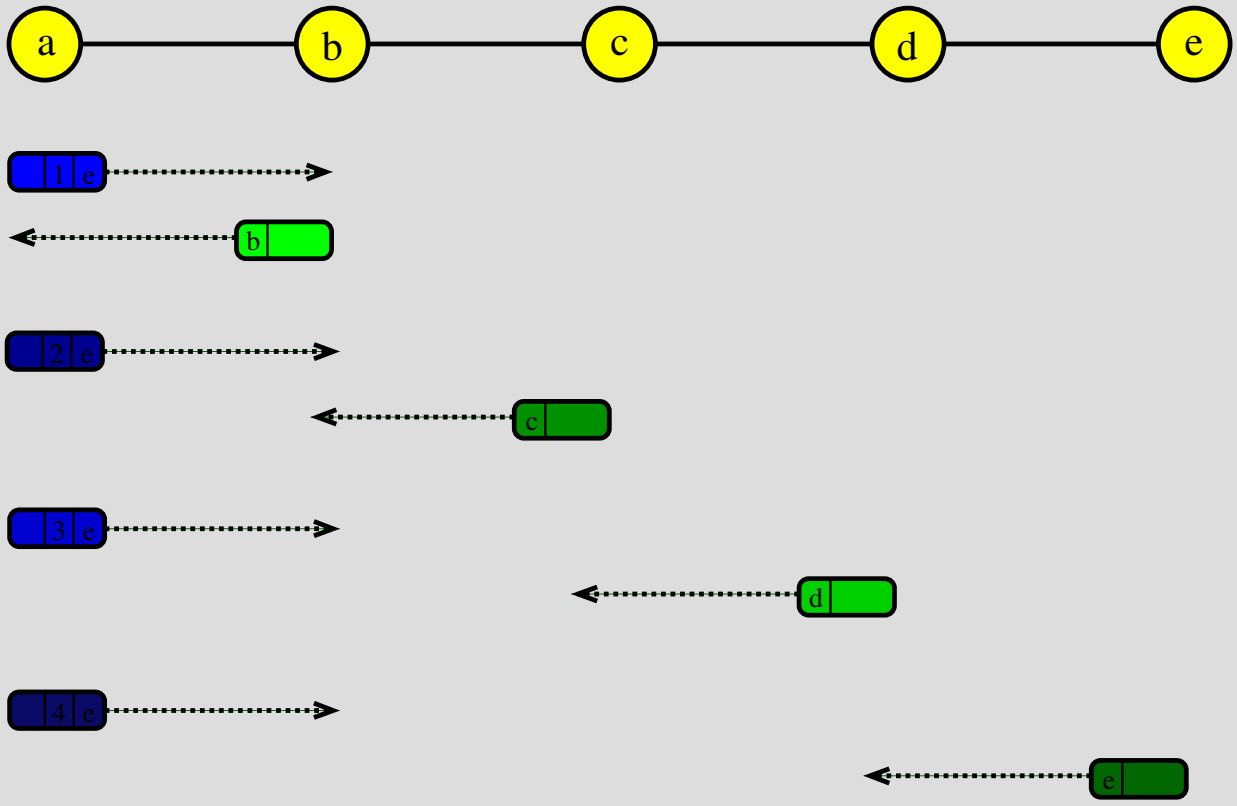
RIPE NCC Visiting Researcher
tony.mcgregor@ripe.net

The University of Waikato
tonym@cs.waikato.ac.nz

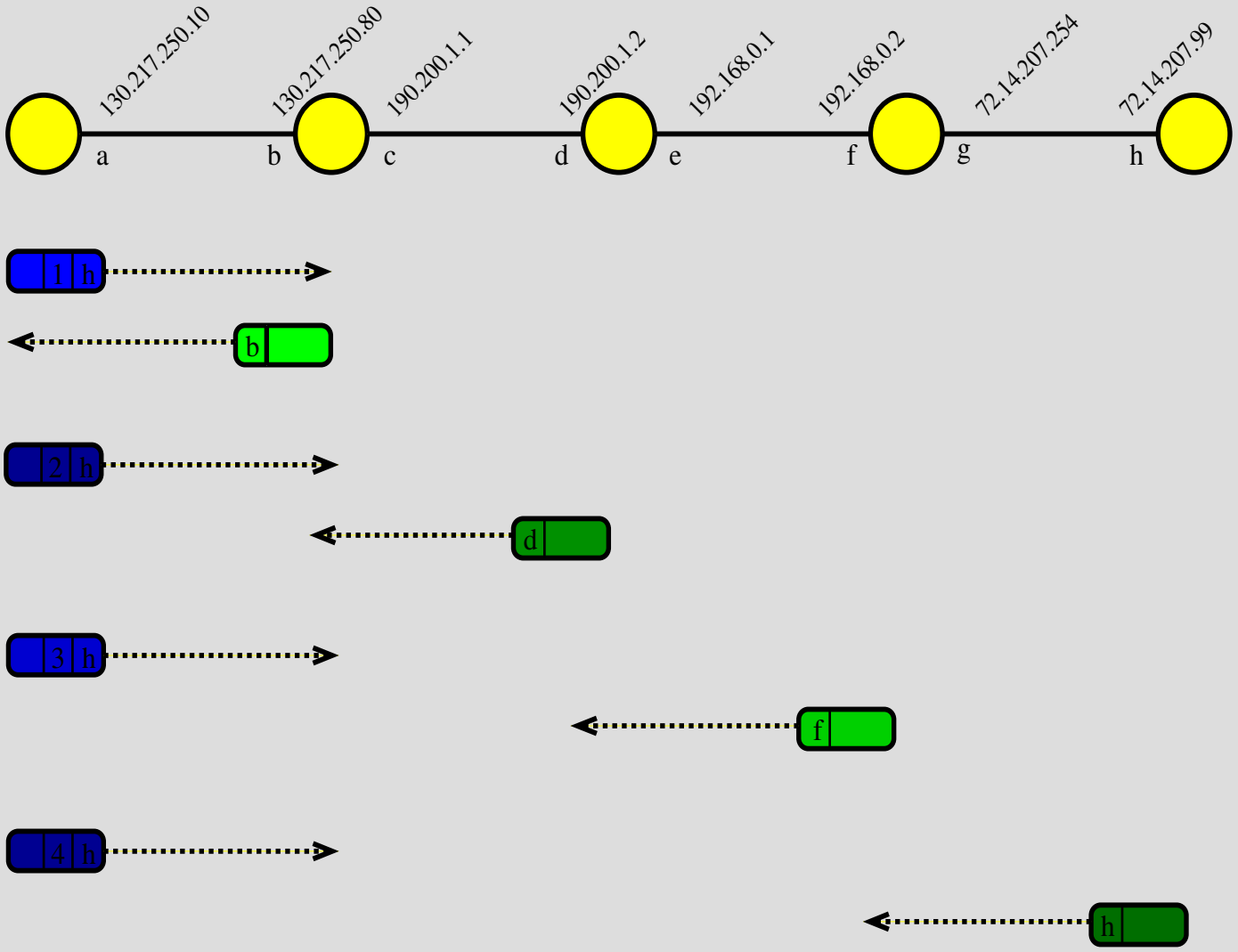


# About Finding The Path between Two Monitors

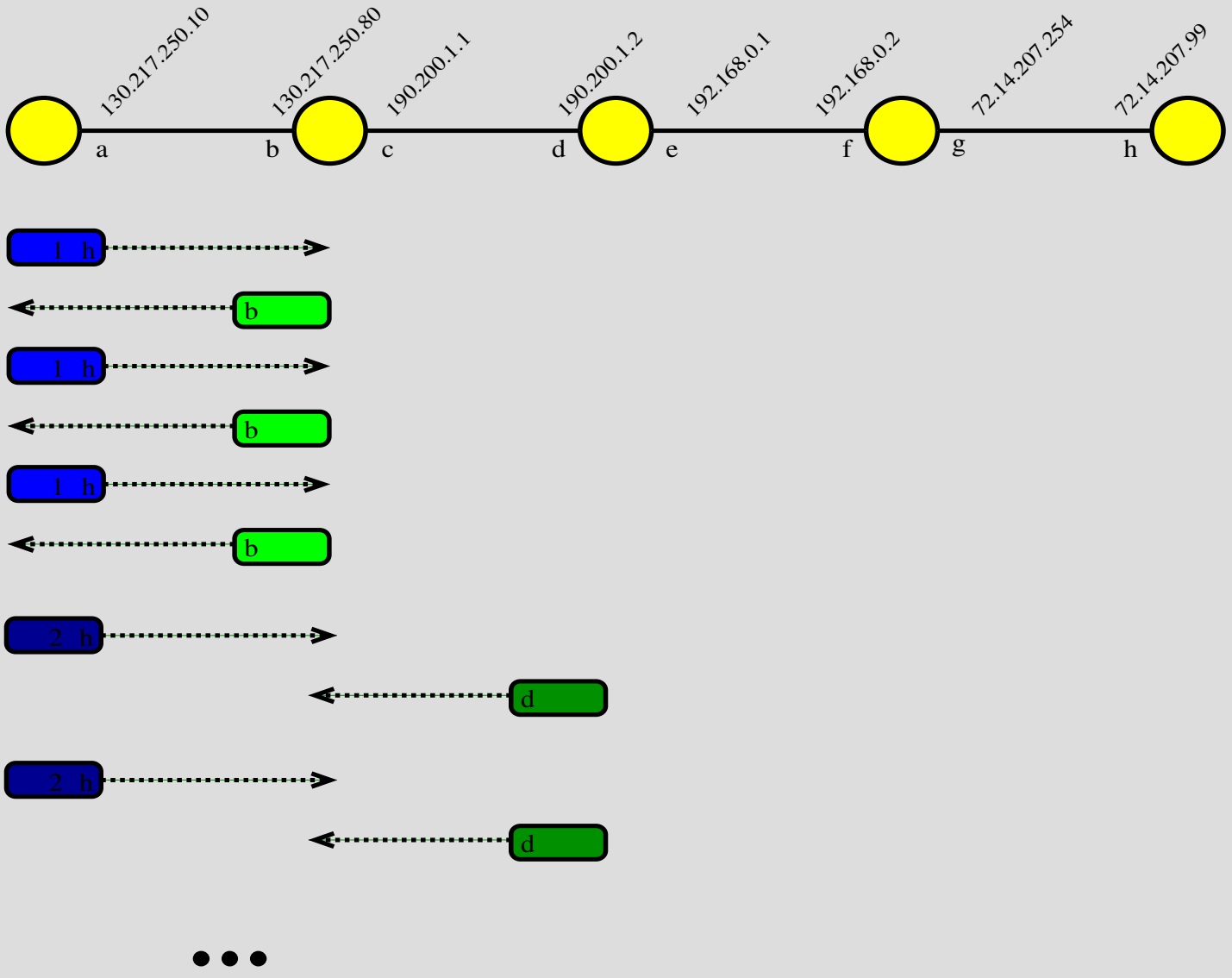
- Traceroute



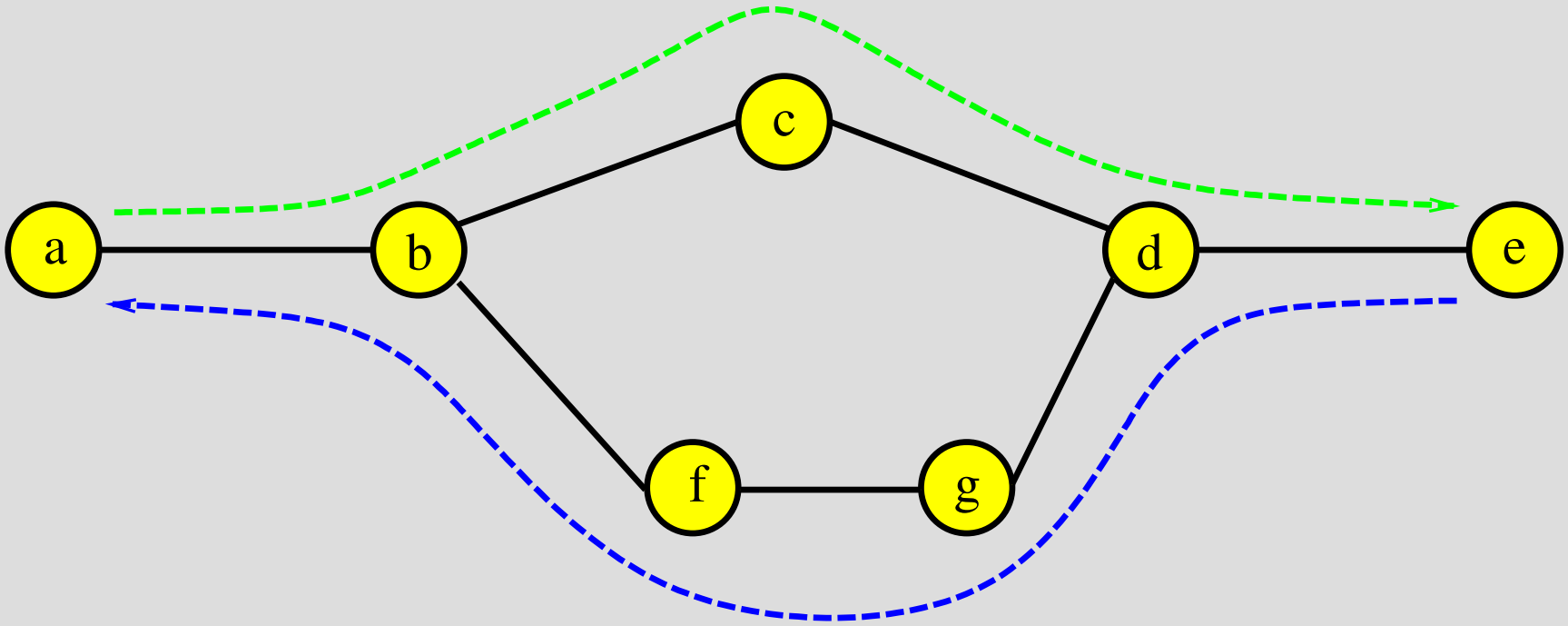
# Traceroute
## (the real story)



# Traceroute
## (the really real story)



# Traceroute
## (more reality)

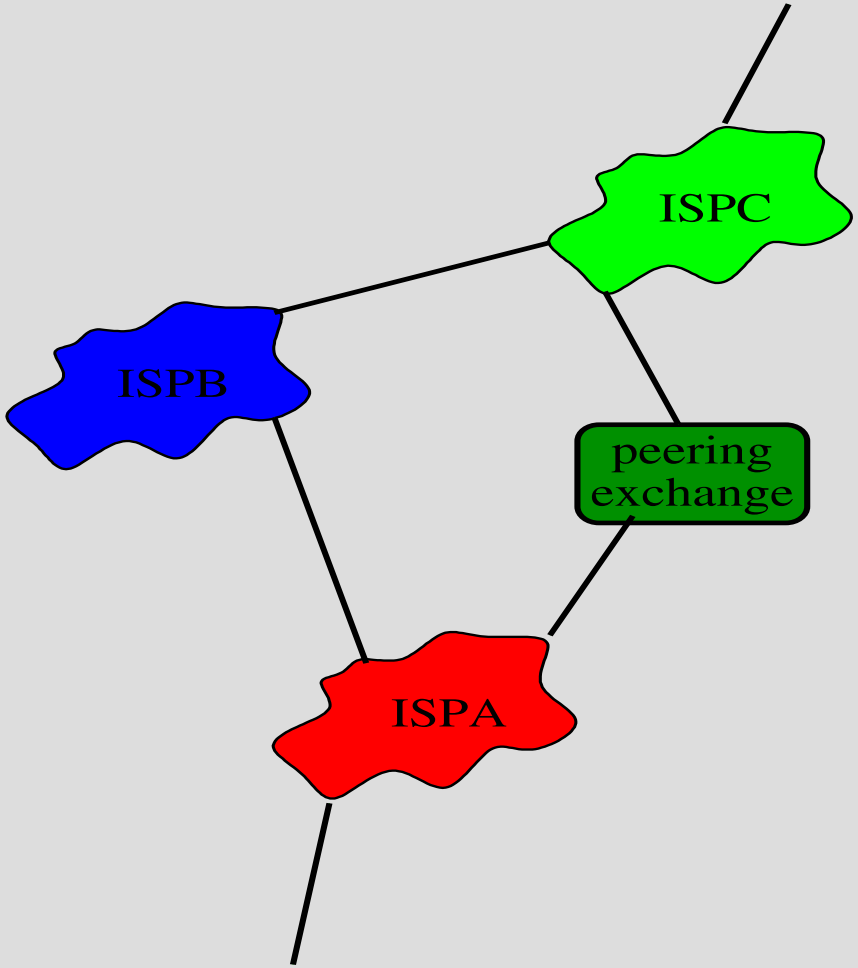


# Traceroute
## (But wait ... there's more)

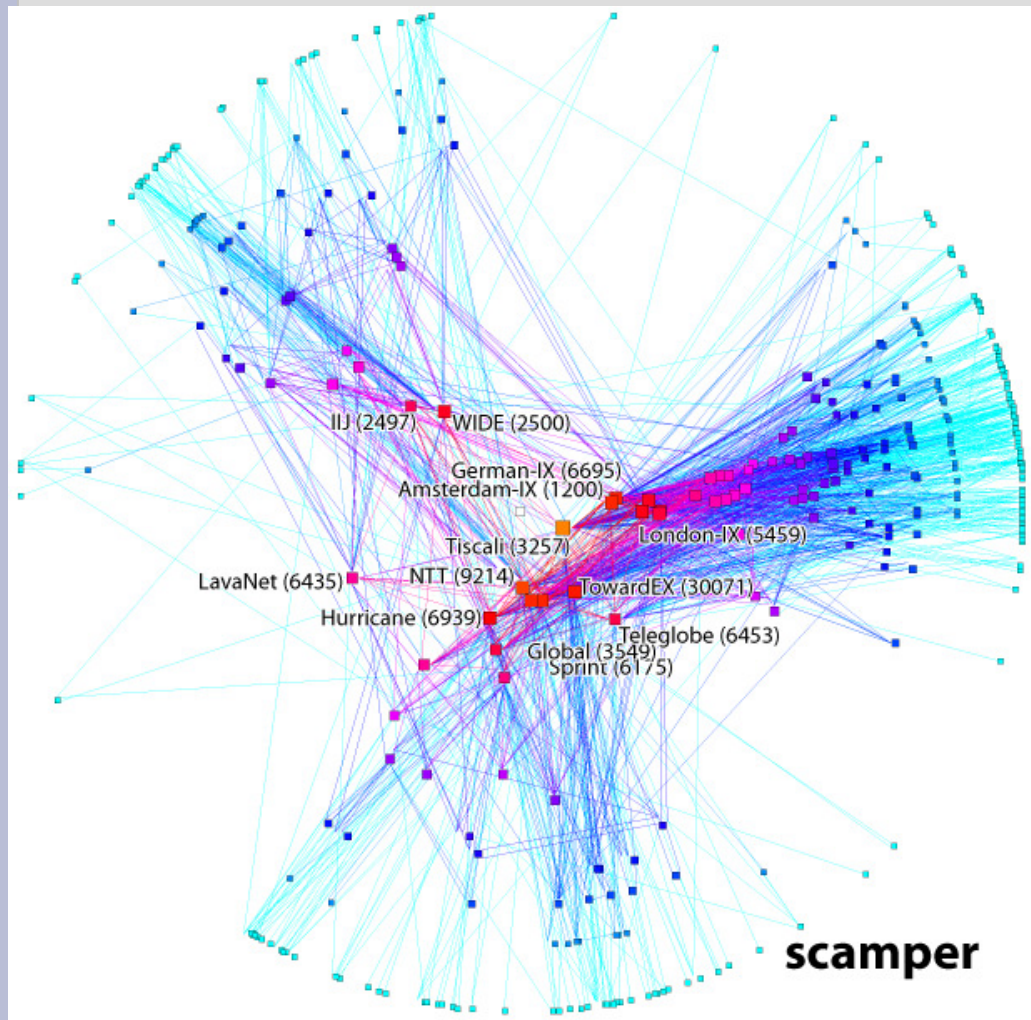
- RFC1918 addresses
- ICMP disabled
- ICMP filters
- False routes
- Load balancing
- Interface coalition
- ISPs hide topology
  - TTL 0
- Tunnels
  - MPLS
  - IPv6

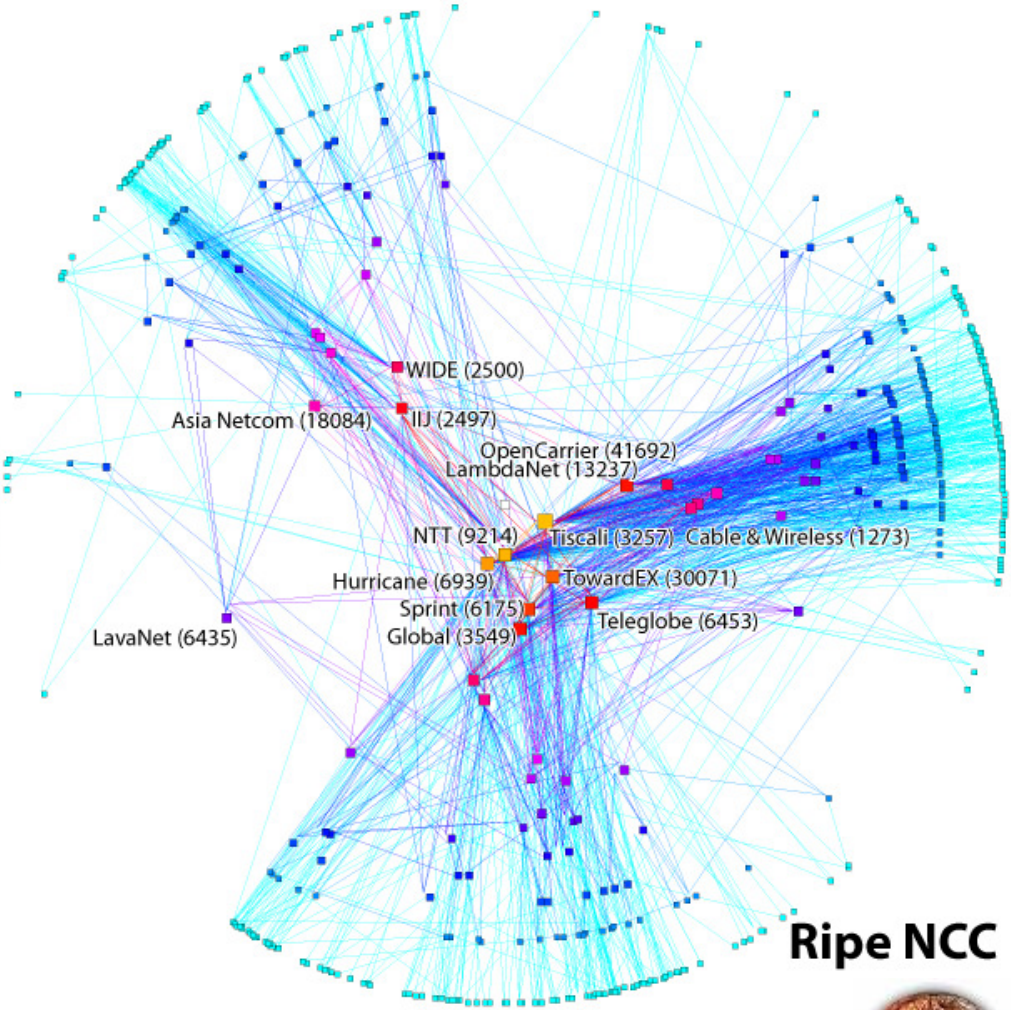
# Why Not Use Routing Data?

- It's different!
  - More detailed
  - Route changes
  - Load balancing
  - Aggregation
  - Policy
  - etc



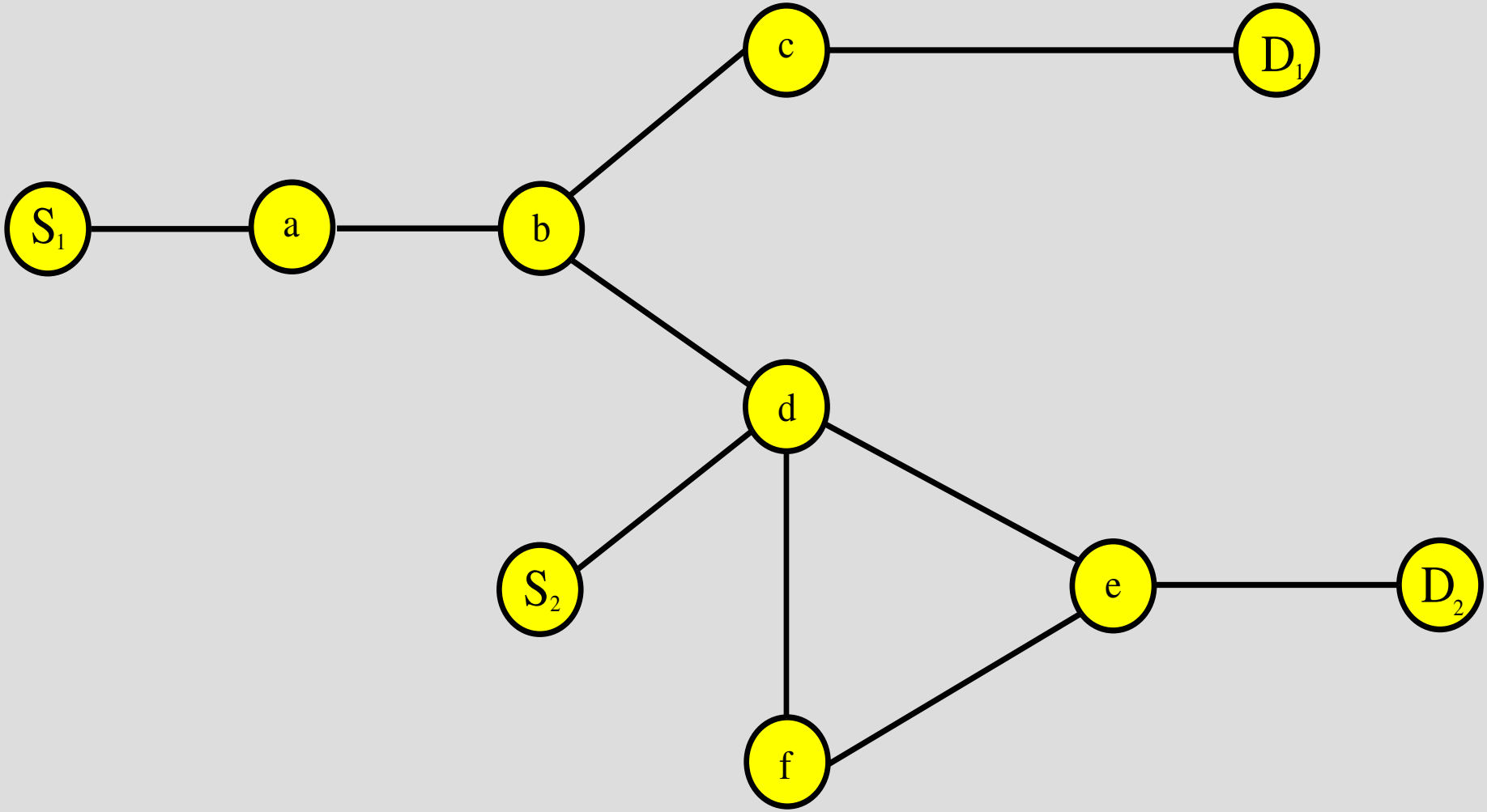
# Scamper vs BGP



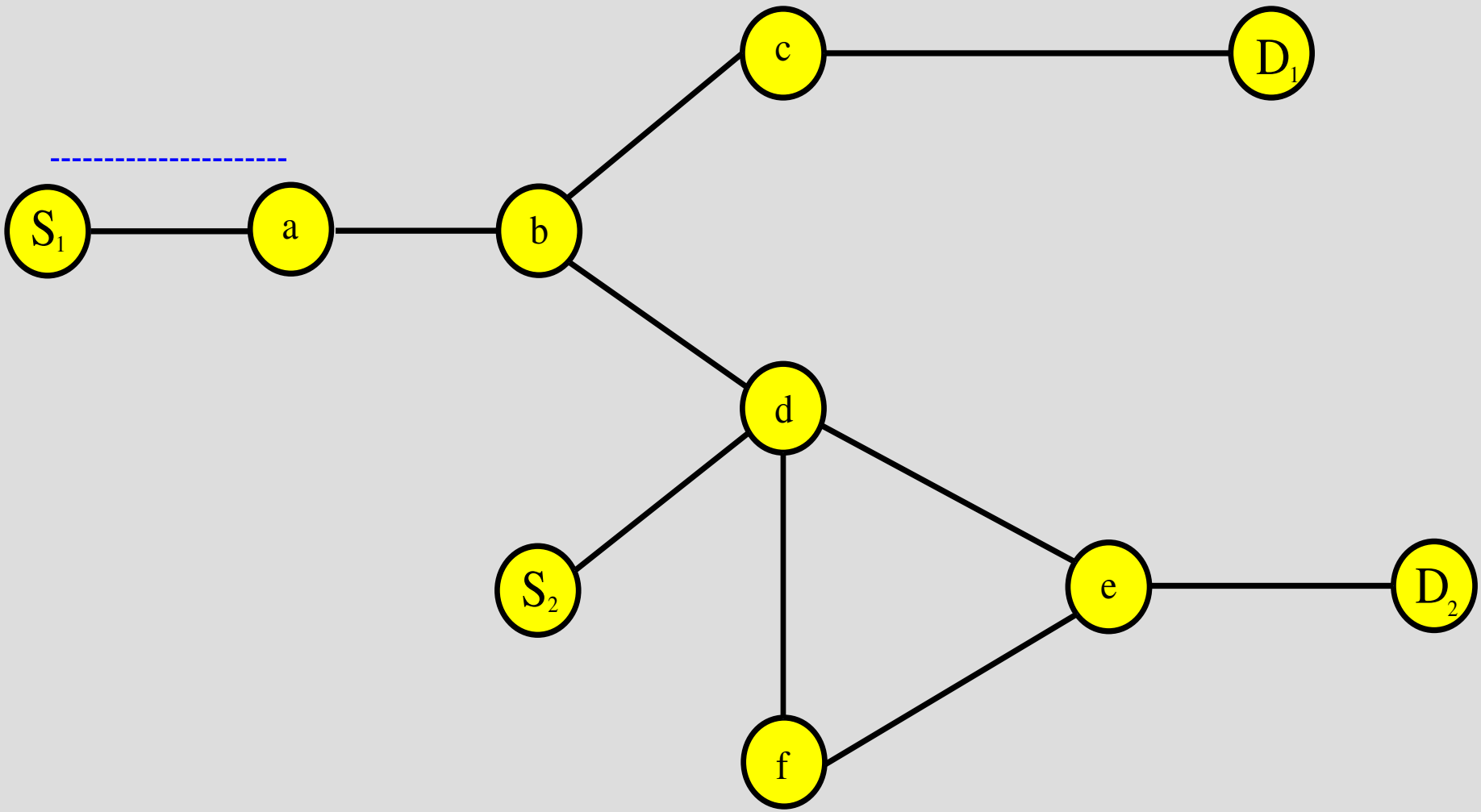

caida

# Wasted Probes
## (doubletree)

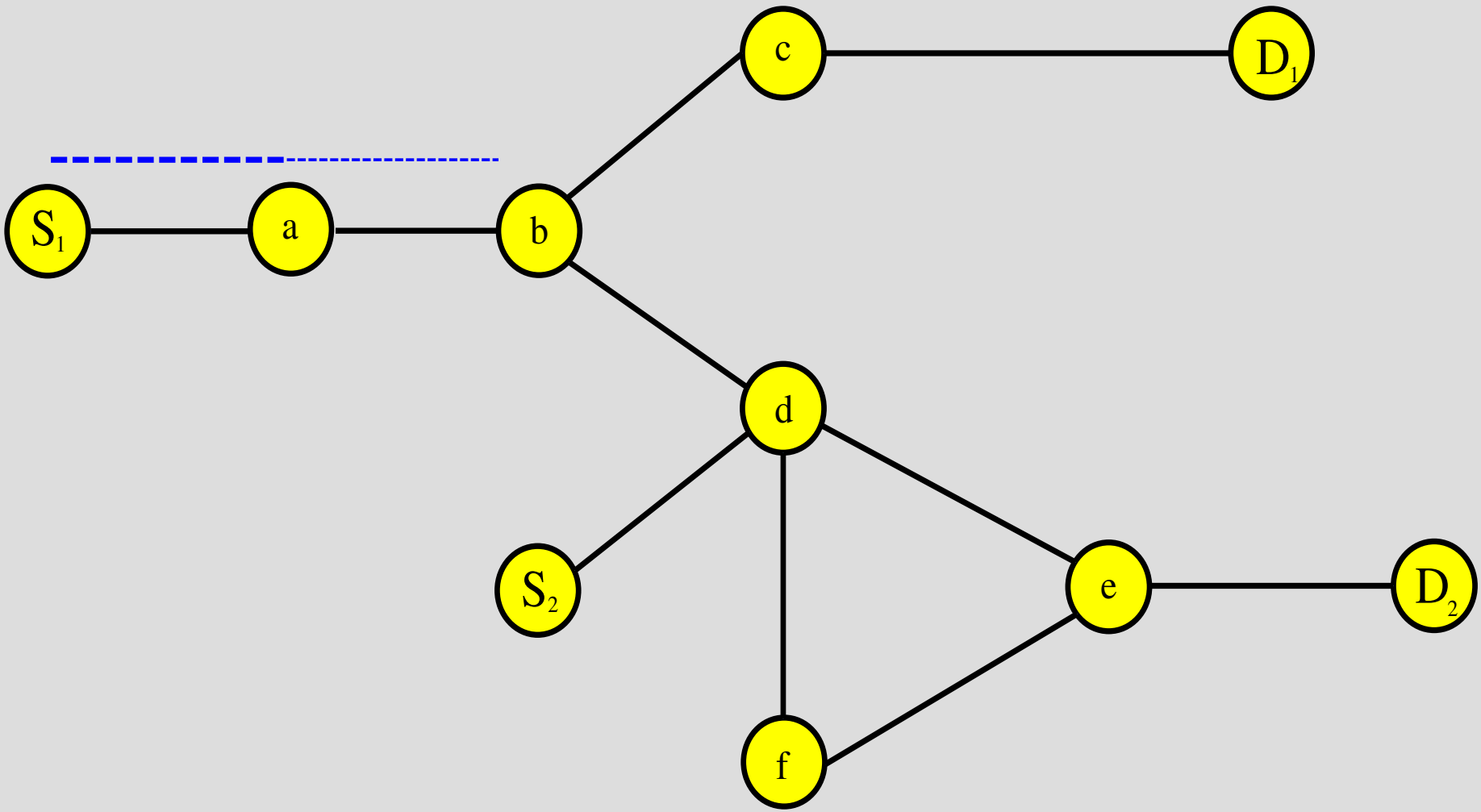



Donnet B et al. “Efficient Algorithms for Large-Scale Topology Discovery” Sigmetrics 05

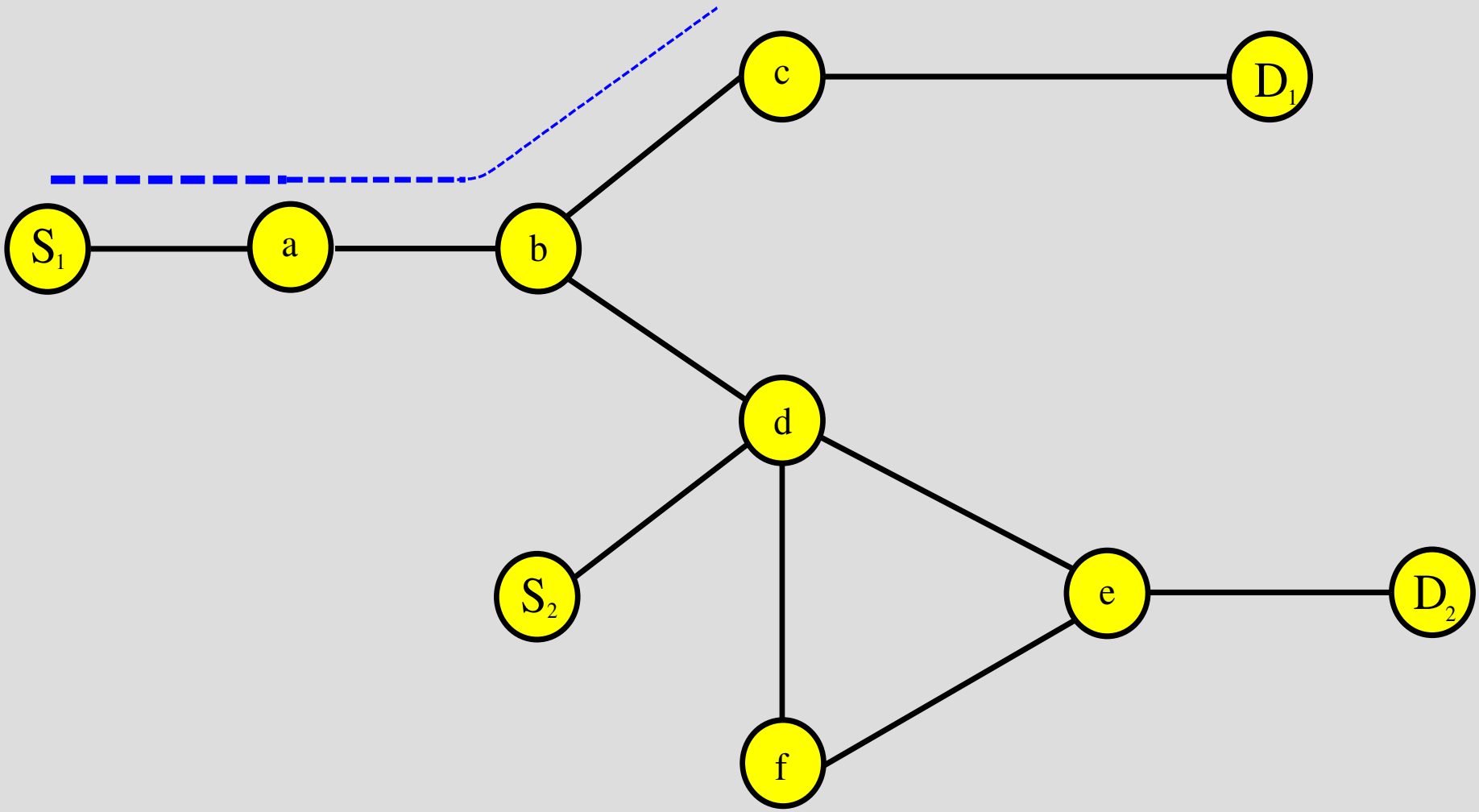
# Wasted Probes
## (doubletree)



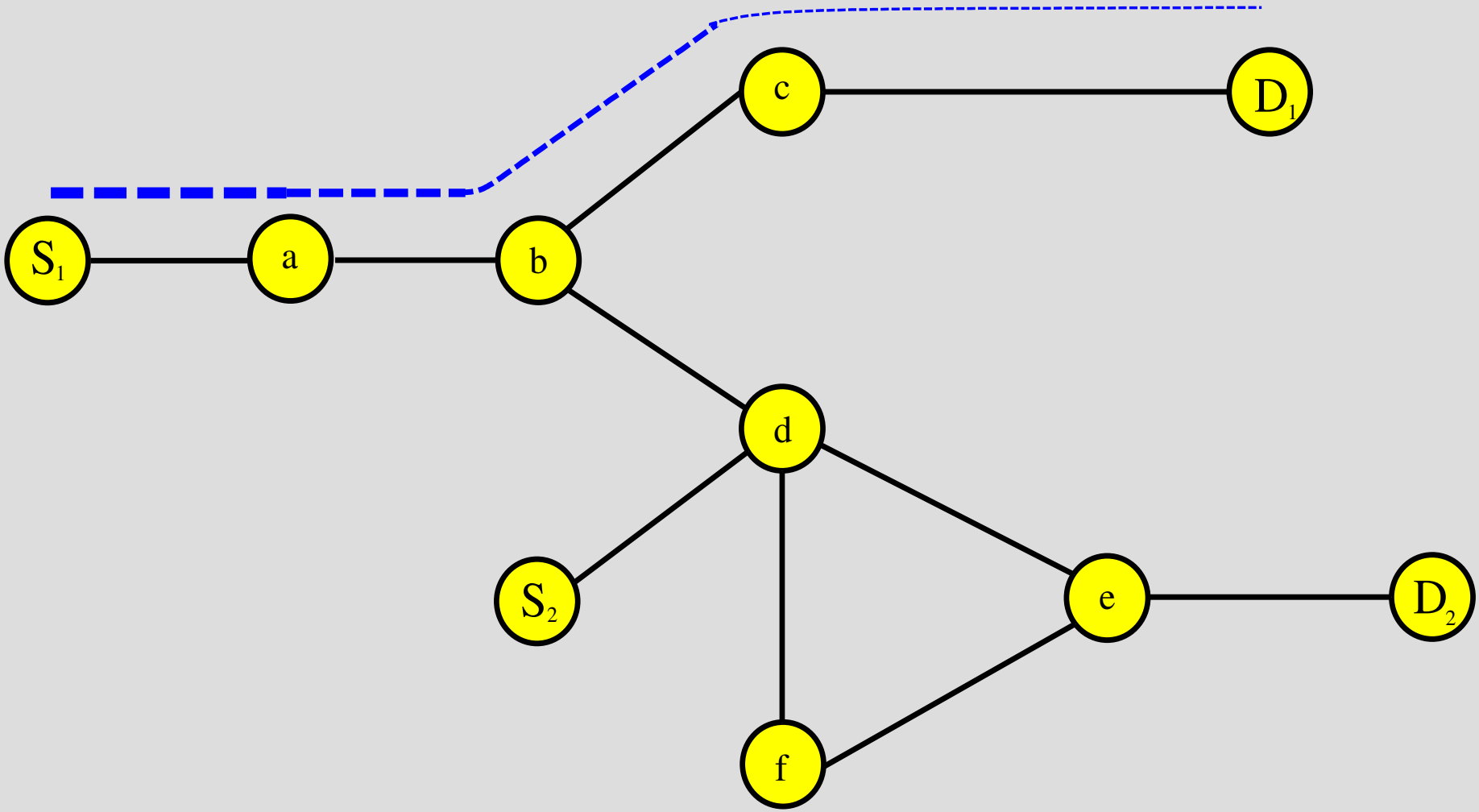
# Wasted Probes
## (doubletree)



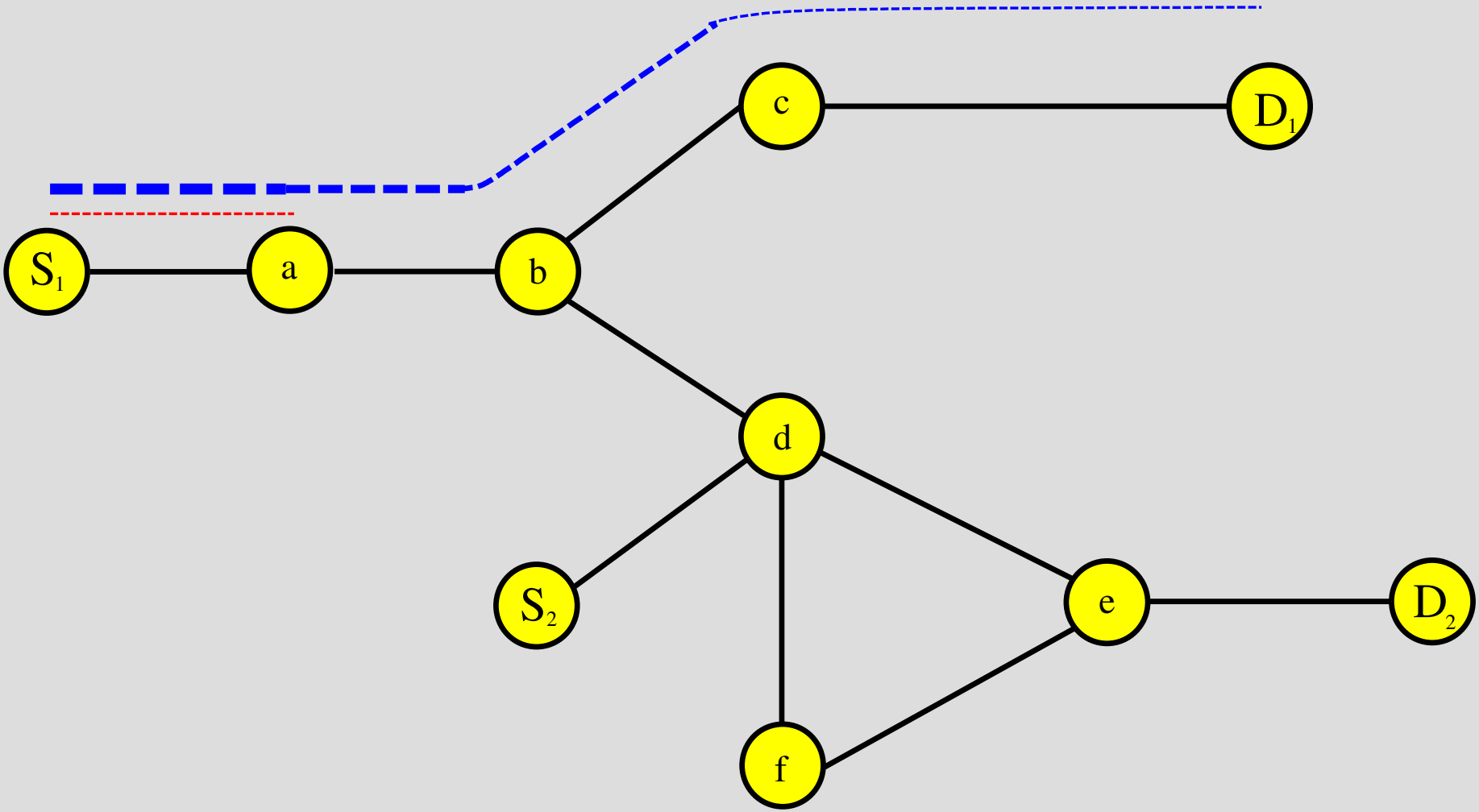
# Wasted Probes
## (doubletree)



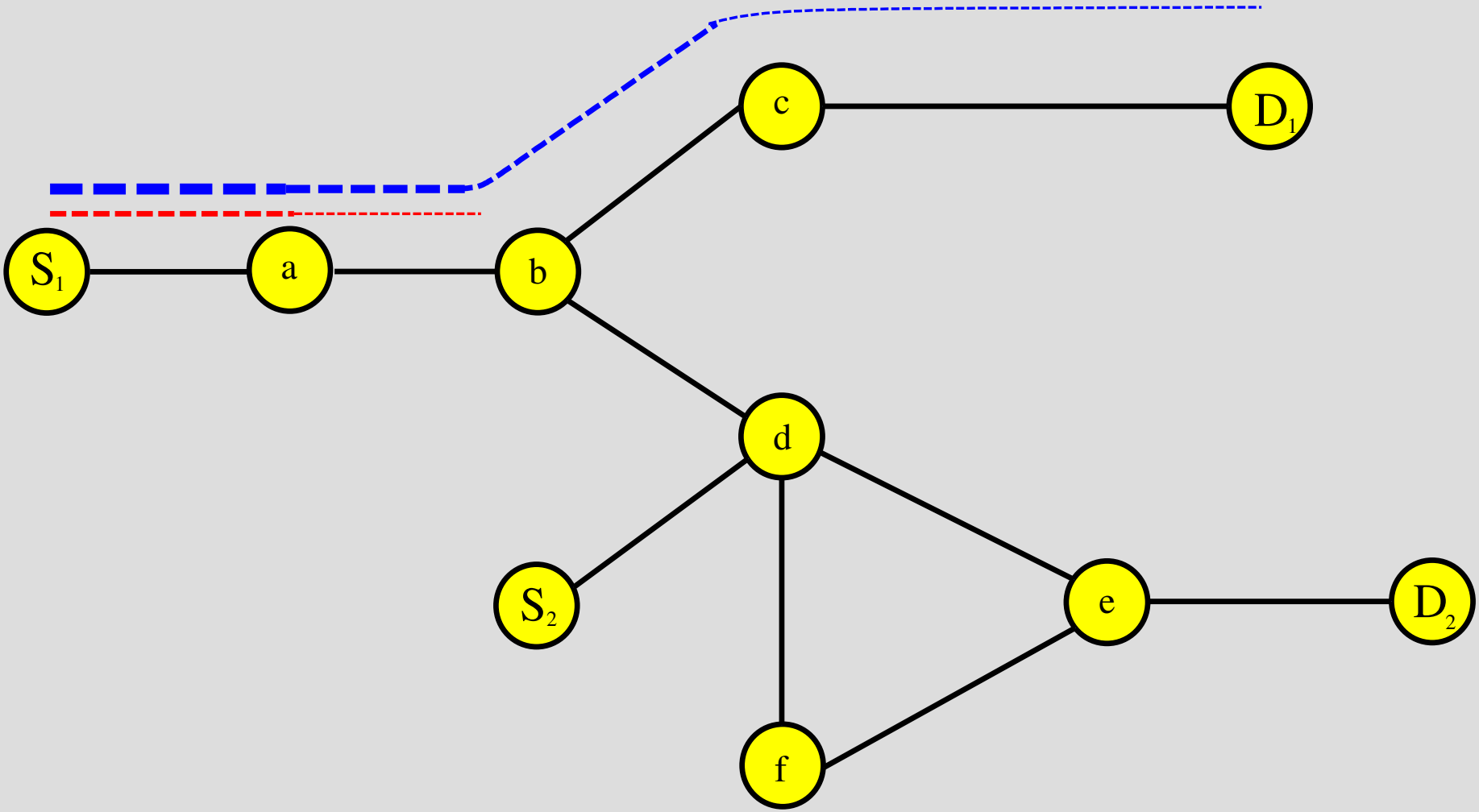
# Wasted Probes
## (doubletree)



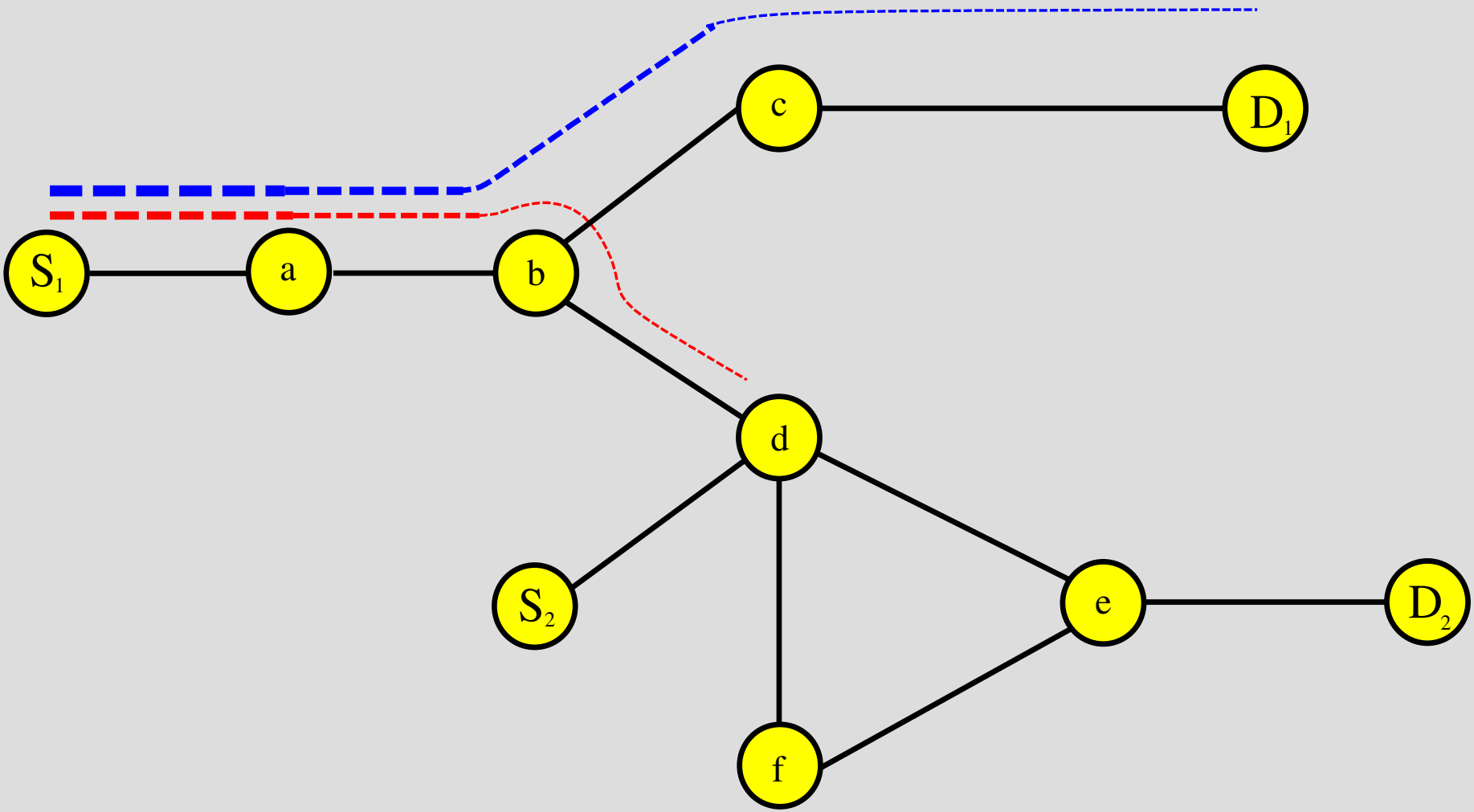
# Wasted Probes
## (doubletree)



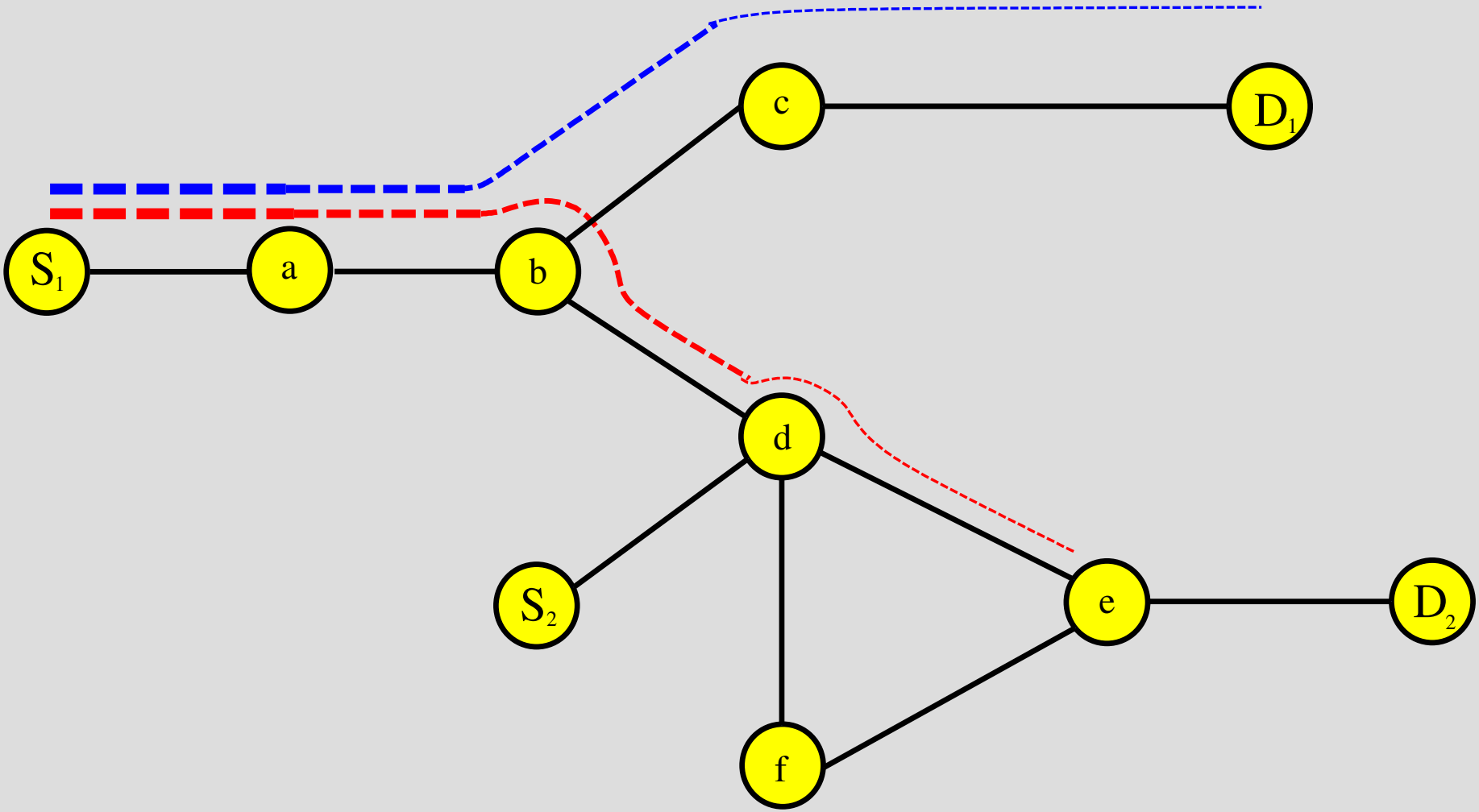
# Wasted Probes
## (doubletree)



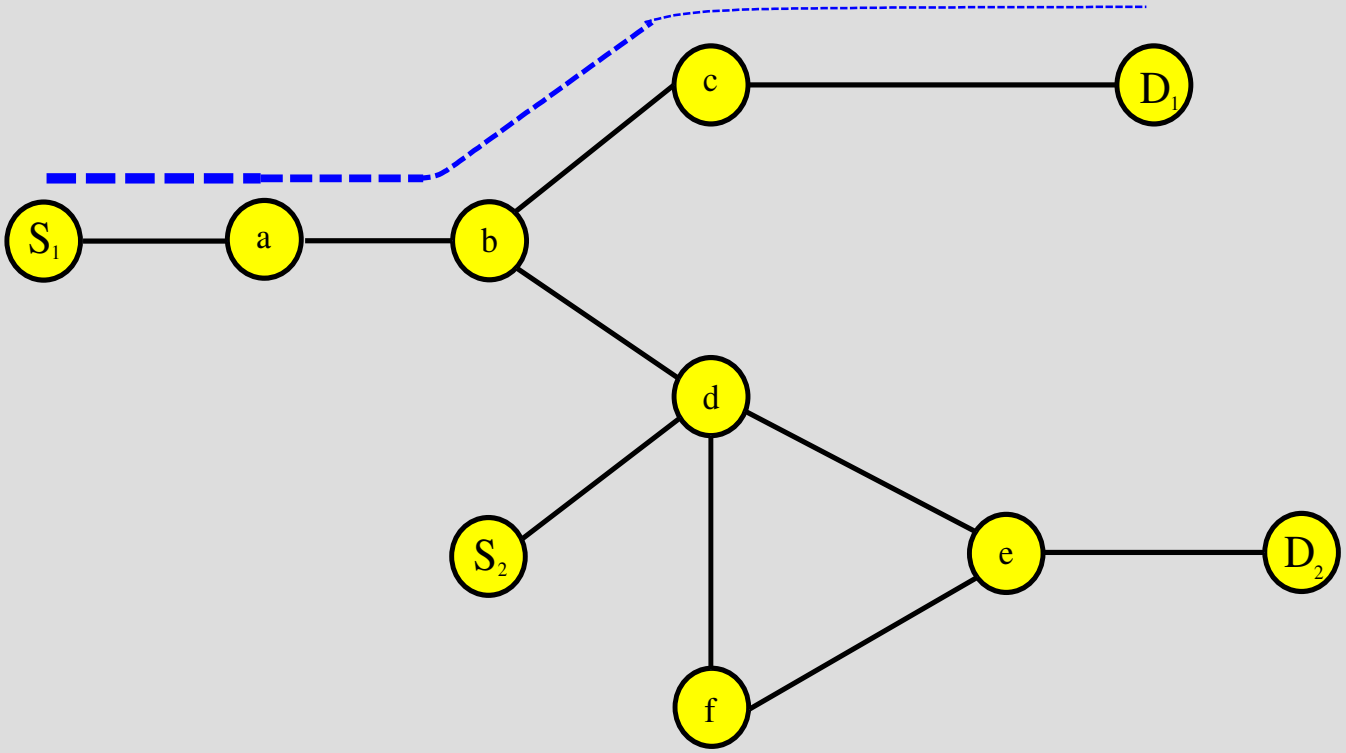
# Wasted Probes
## (doubletree)



# Wasted Probes
## (doubletree)



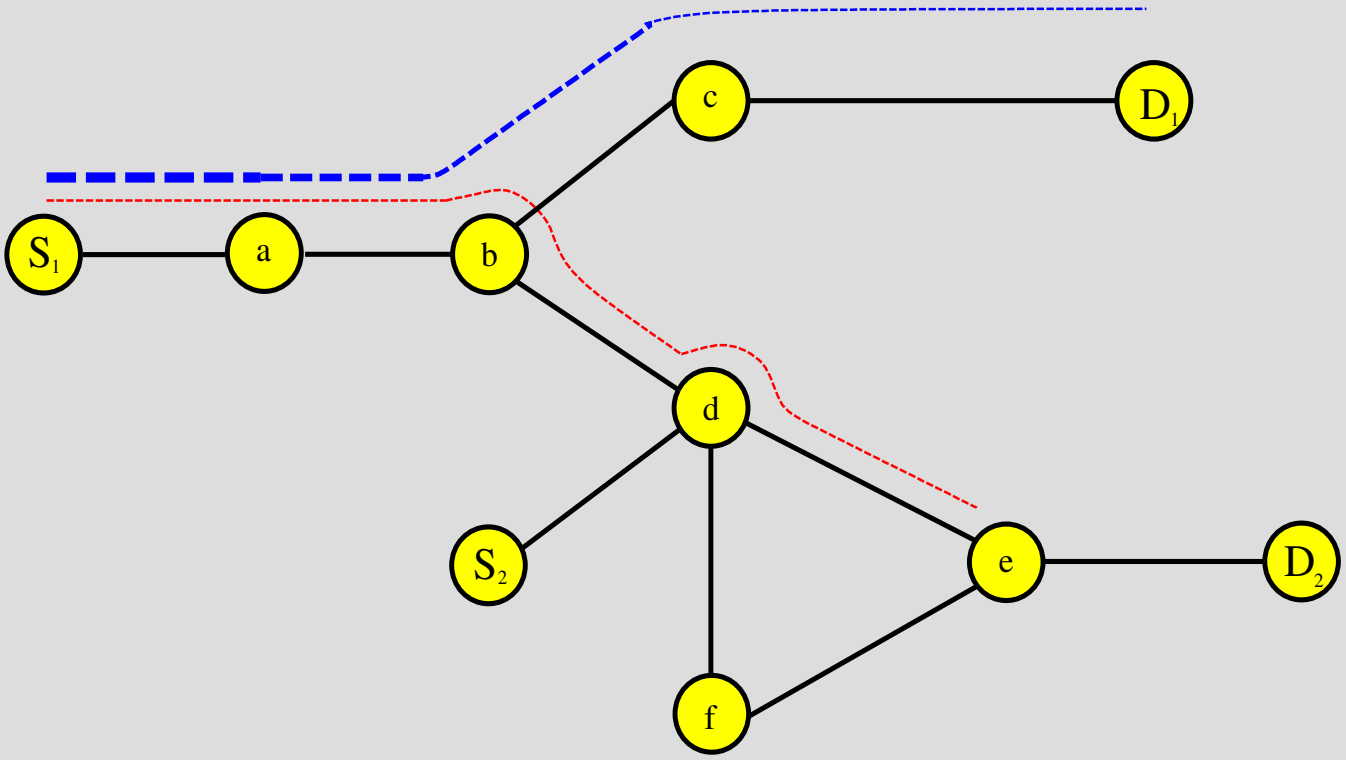
# Intra-monitor Redundancy




Stop set

S1, a, b, c, D1

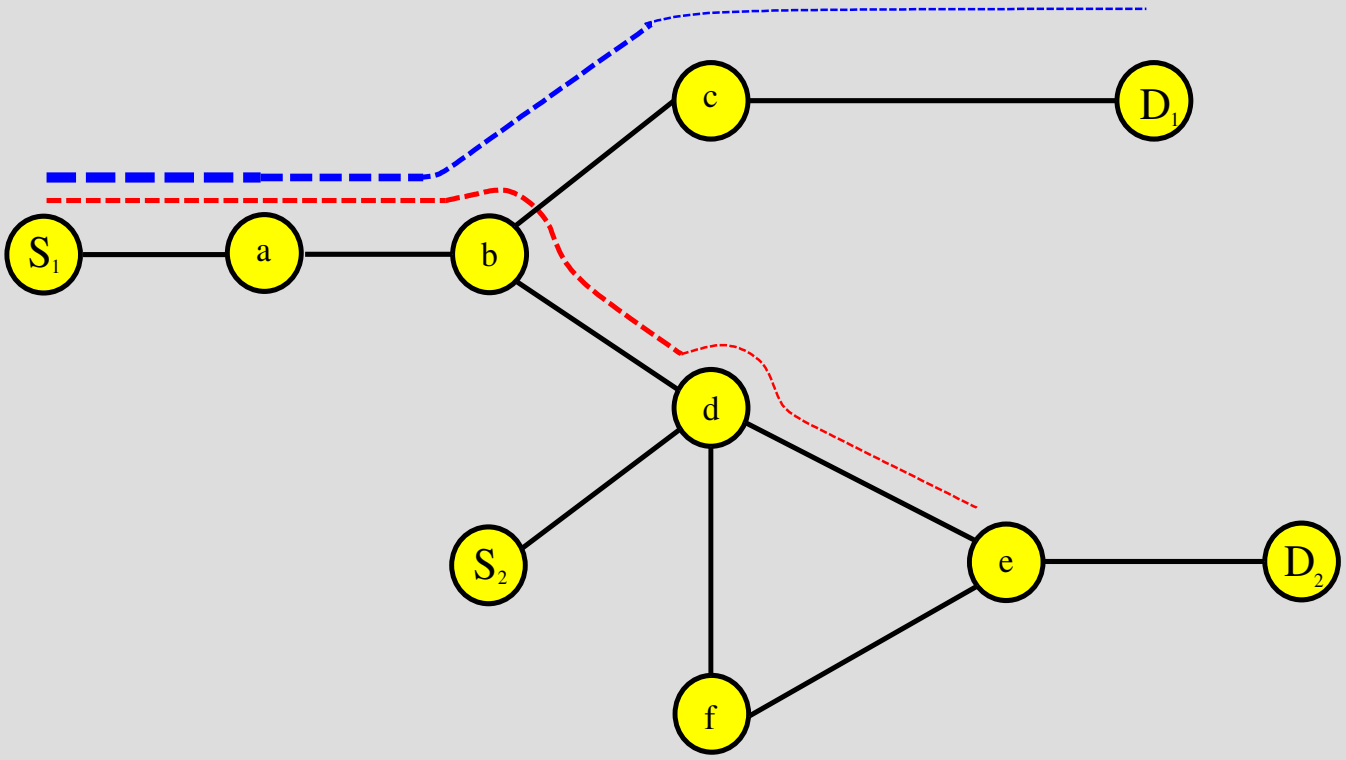
# Intra-monitor Redundancy




Stop set

S1, a, b, c, D1, e

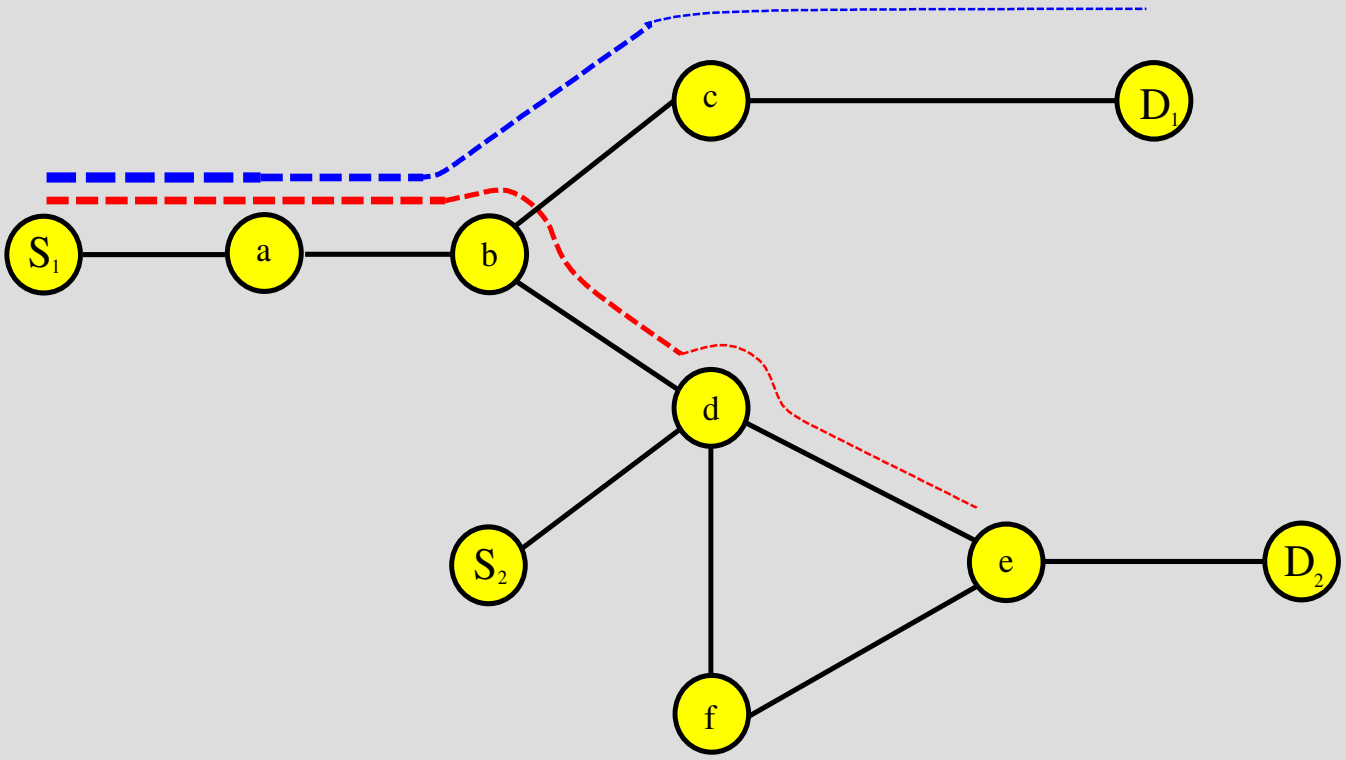
# Intra-monitor Redundancy




Stop set

S1, a, b, c, D1, e, d

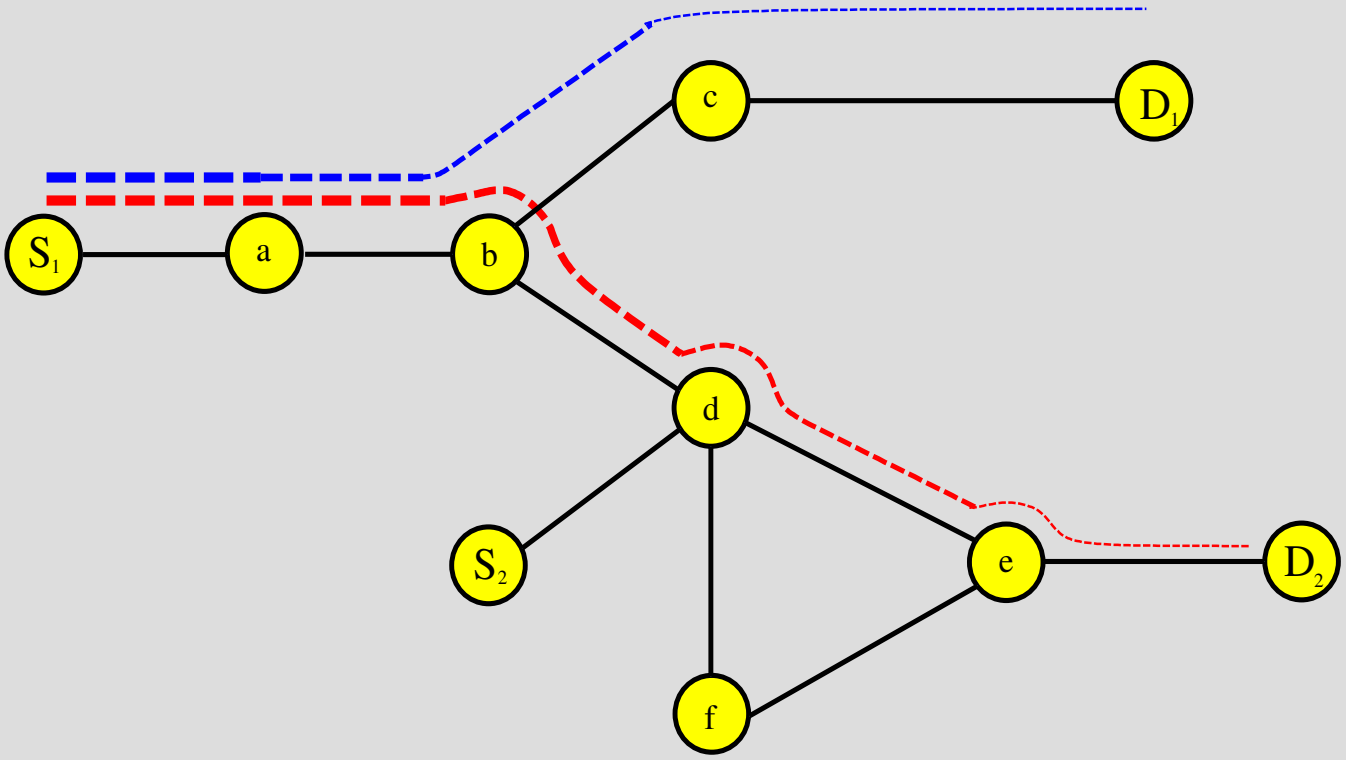
# Intra-monitor Redundancy




Stop set

S1, a, b, c, D1, e, d

# Intra-monitor Redundancy

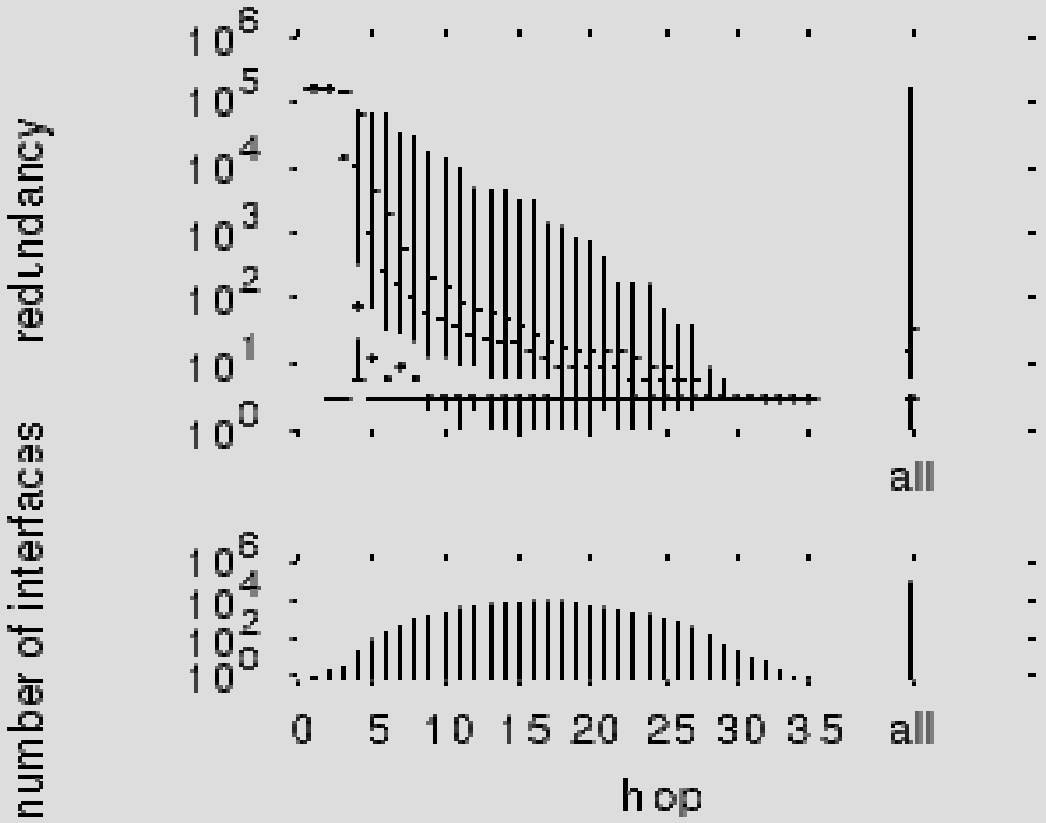


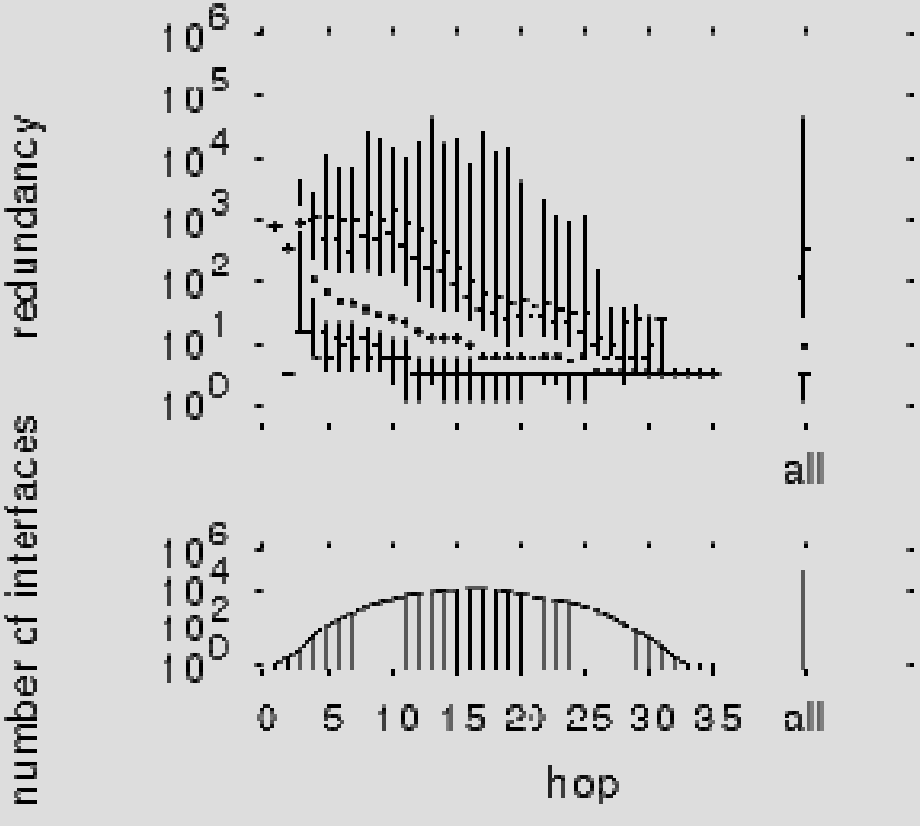

Stop set

S1, a, b, c, D1, e, d, D2

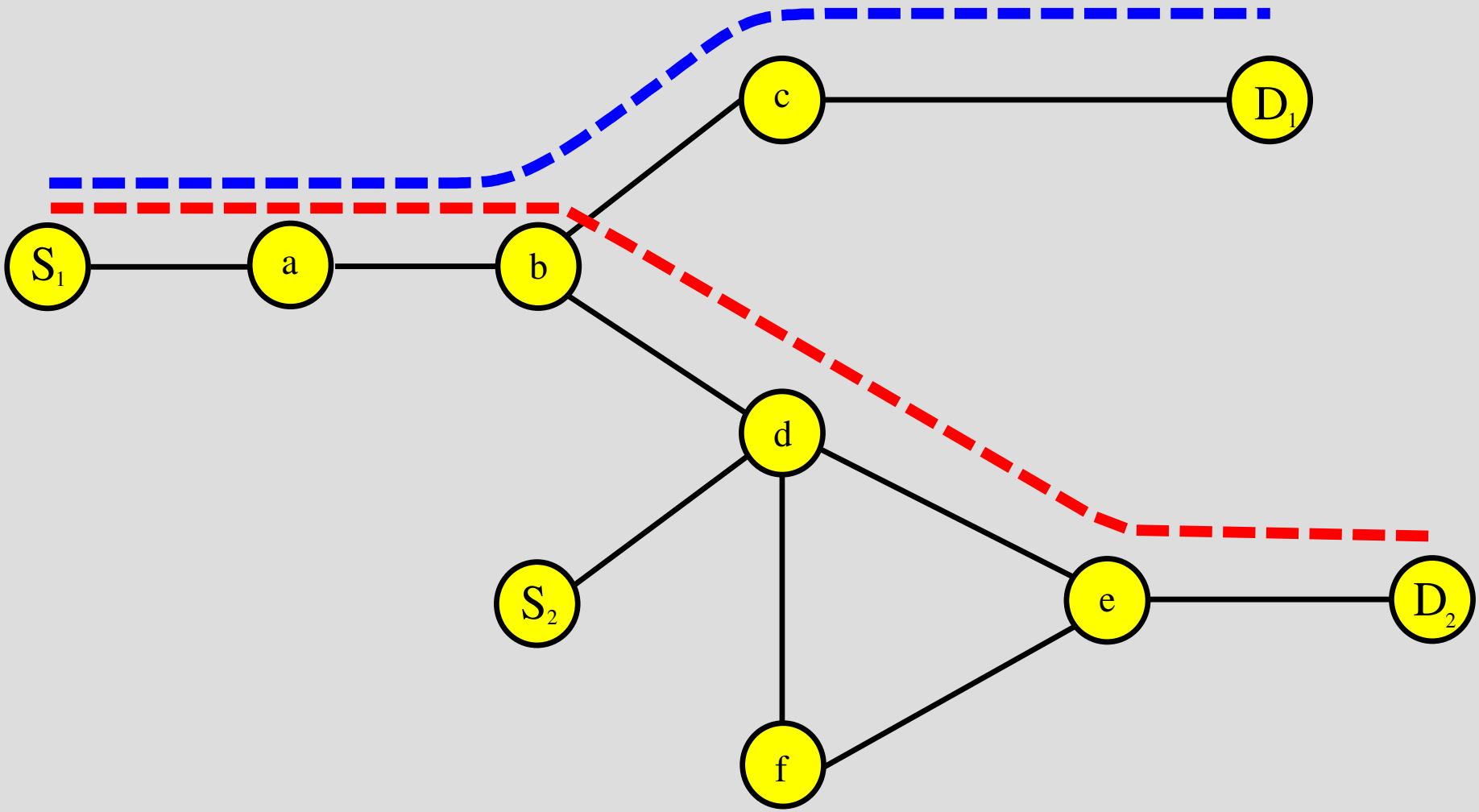
# Extent of intra-monitor redundancy

- Donnet et al. Estimate that in skitter 87% of probes were redundant.

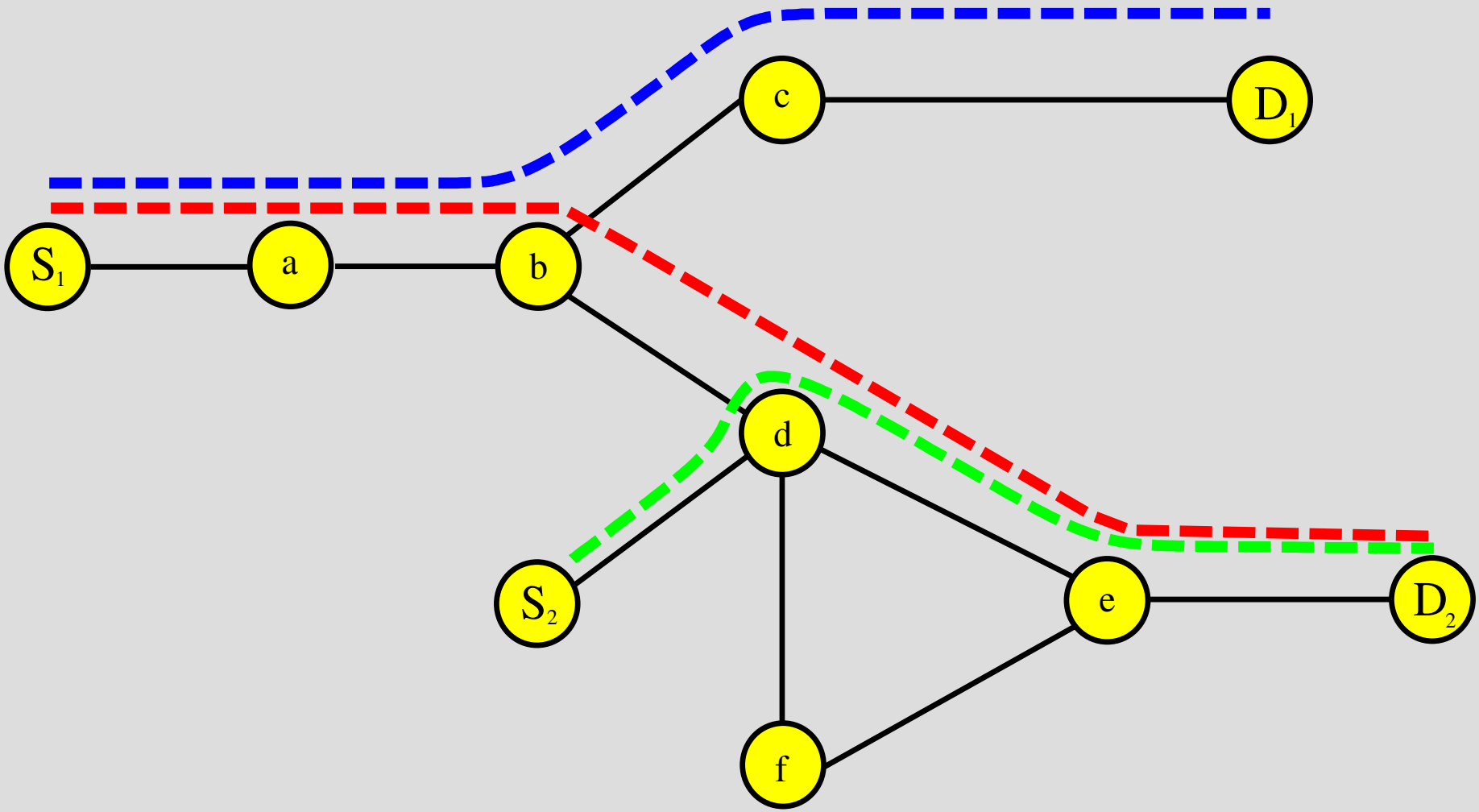


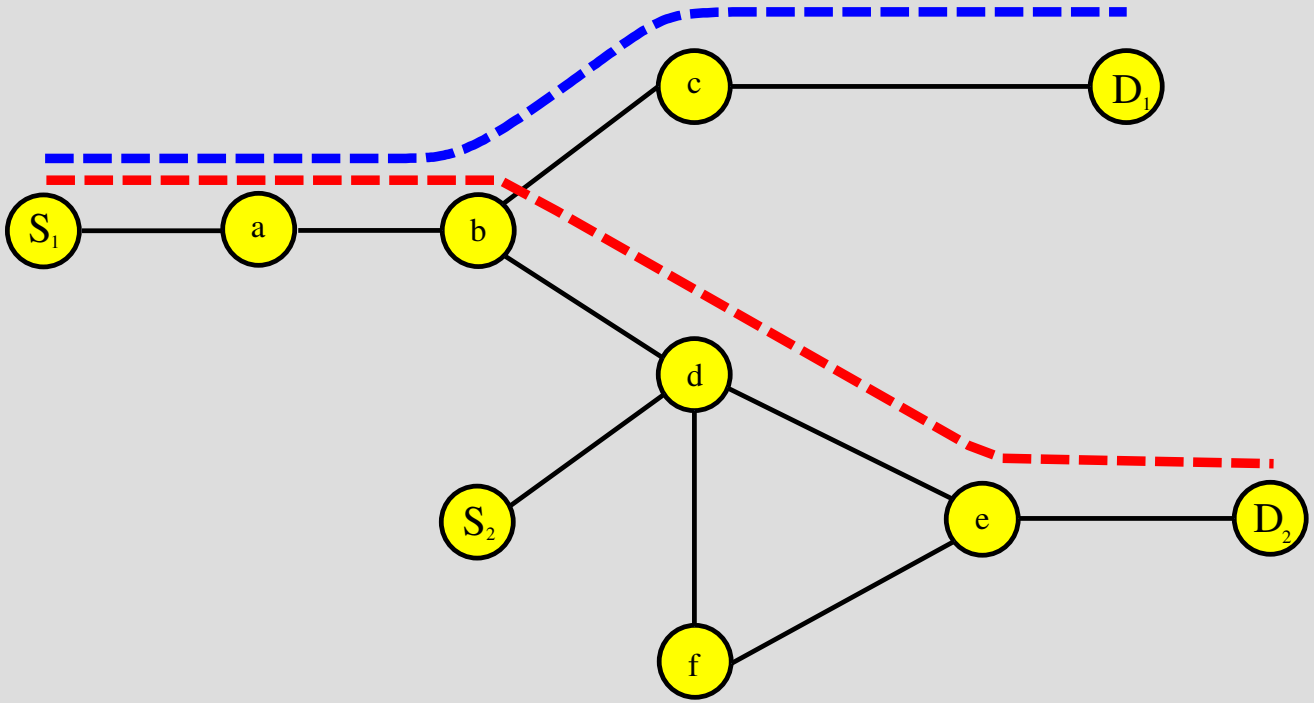
# Inter-monitor redundancy



# Inter-monitor redundancy



# Inter-monitor redundancy

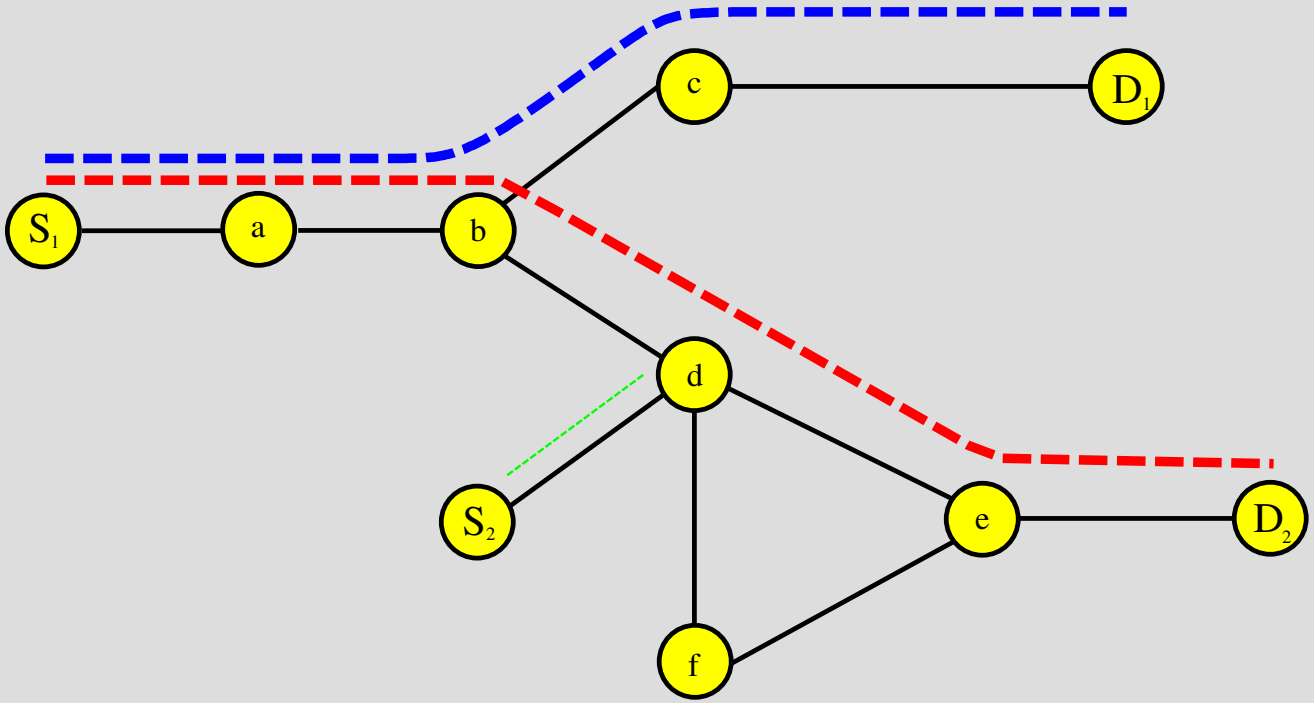



global stop set

D1: a, b, c

D2: a, b, d, e

# Inter-monitor redundancy

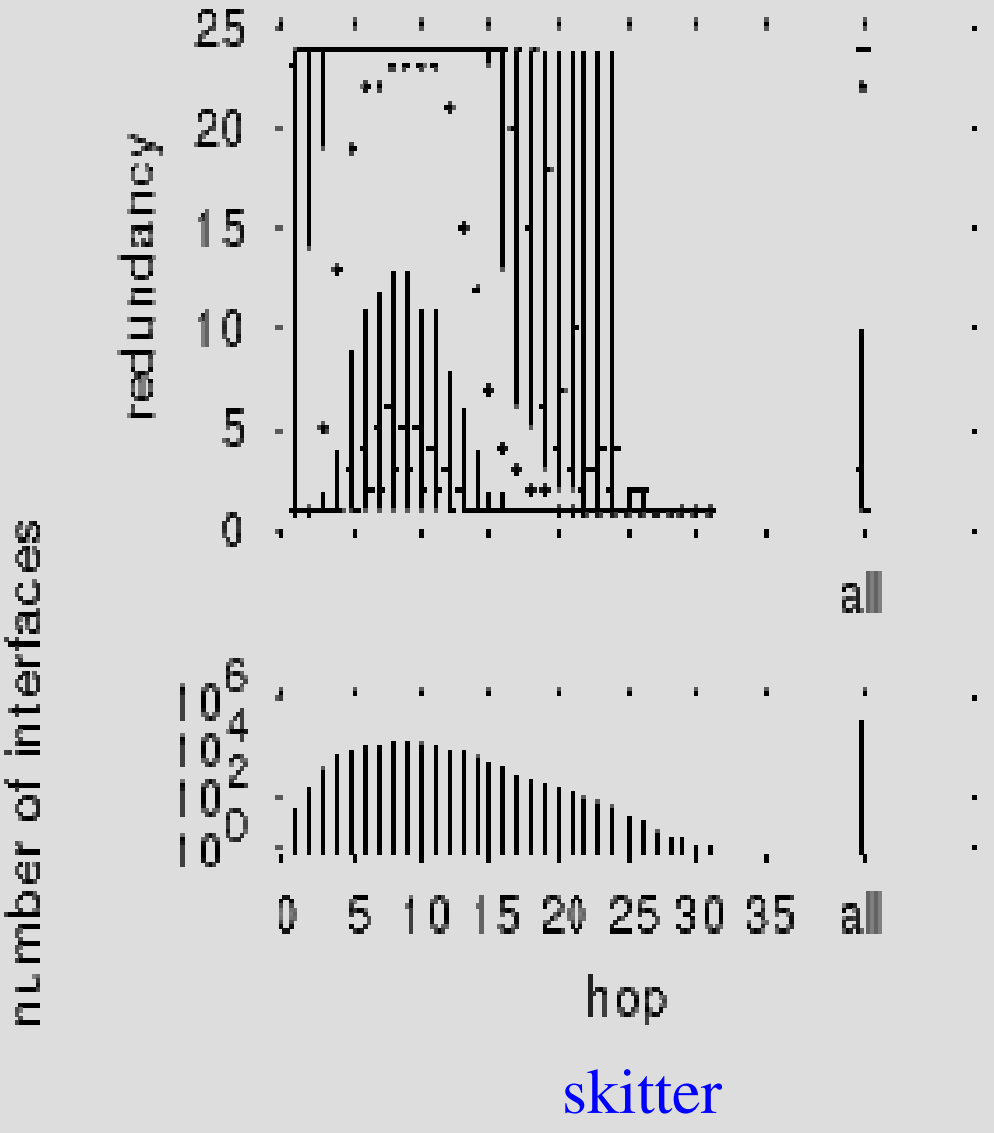



global stop set

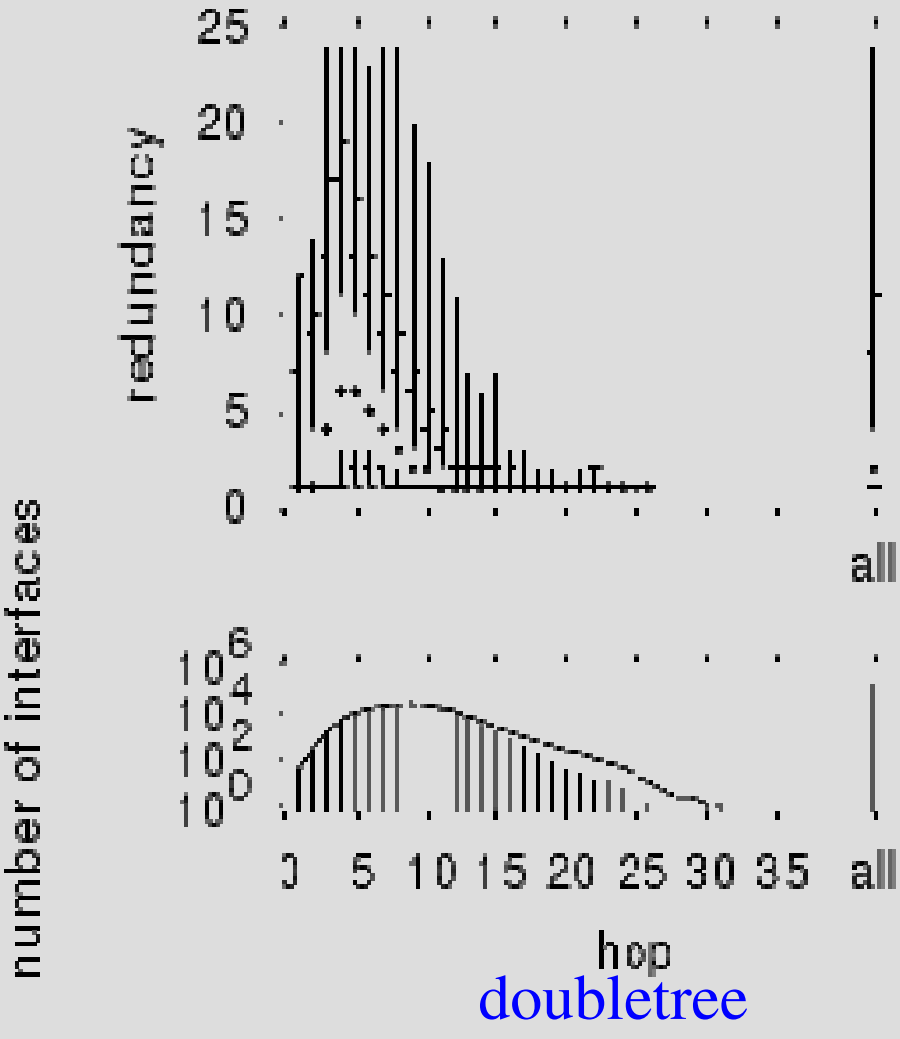
D1: a, b, c

D2: a, b, d, e

# Inter-monitor redundancy




skitter




doubletree

# Limitations

- Doubletree finds fewer nodes that repeated traceroutes
  - Path changes
  - Load balancing
- Global stop set requires communication
  - Negates savings to some extent
  - May be large if there are many destinations

# Doubletree for DAR

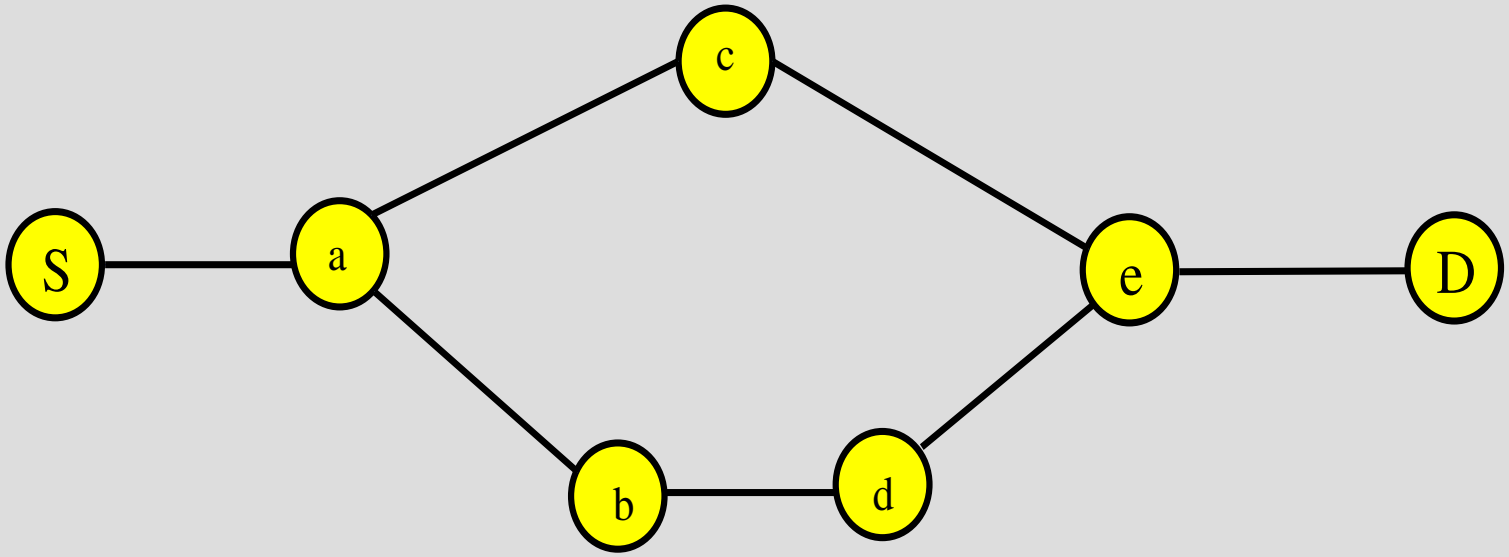
- The DAR project is investigating the design of a very large scale active measurement system
  - Design should support in the order of 100.000 probes
- Want to find path to a small number of targets from a large number of probes
  - can't just probe
  - double tree
    - A lot of redundancy so much potential for savings
    - Need to share global stop set between many probes
    - Global stop set will be smaller than for scamper

# Discovering Alternative Paths
## load balancing and Paris traceroute

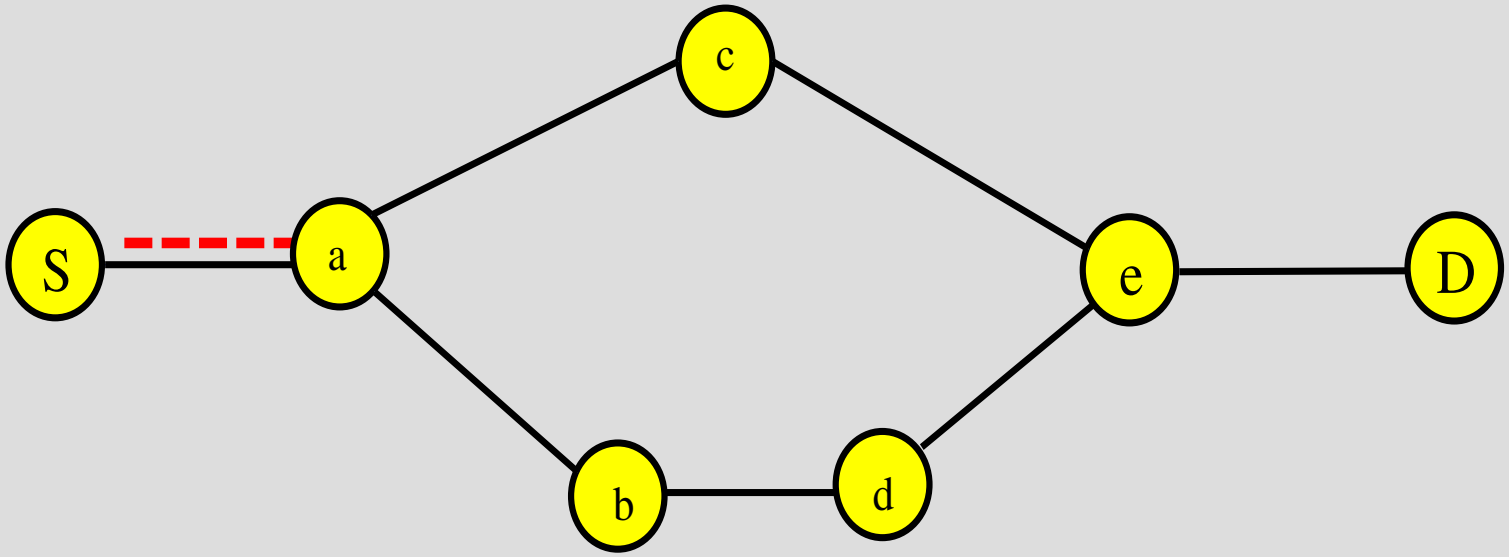
- TCP traceroute study (July 2006, 5,000 addresses)
  - 7 of 9 Tear one ISPs
  - 17 of 64 top regional ISPs
  - Over 80% of destinations appeared to have load balancers in their path

Augustin et al. “Avoiding traceroute anomalies with Paris traceroute”
Université of Pierre et Marie Curie

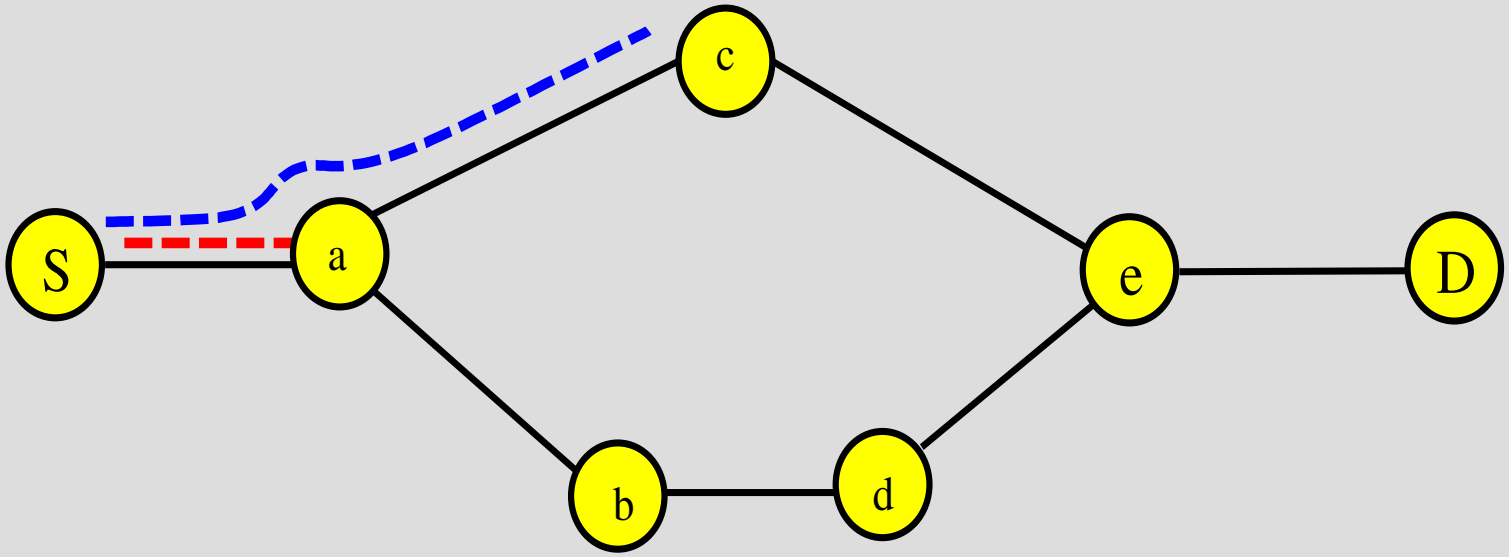
# Alternative Paths
## naïve traceroute



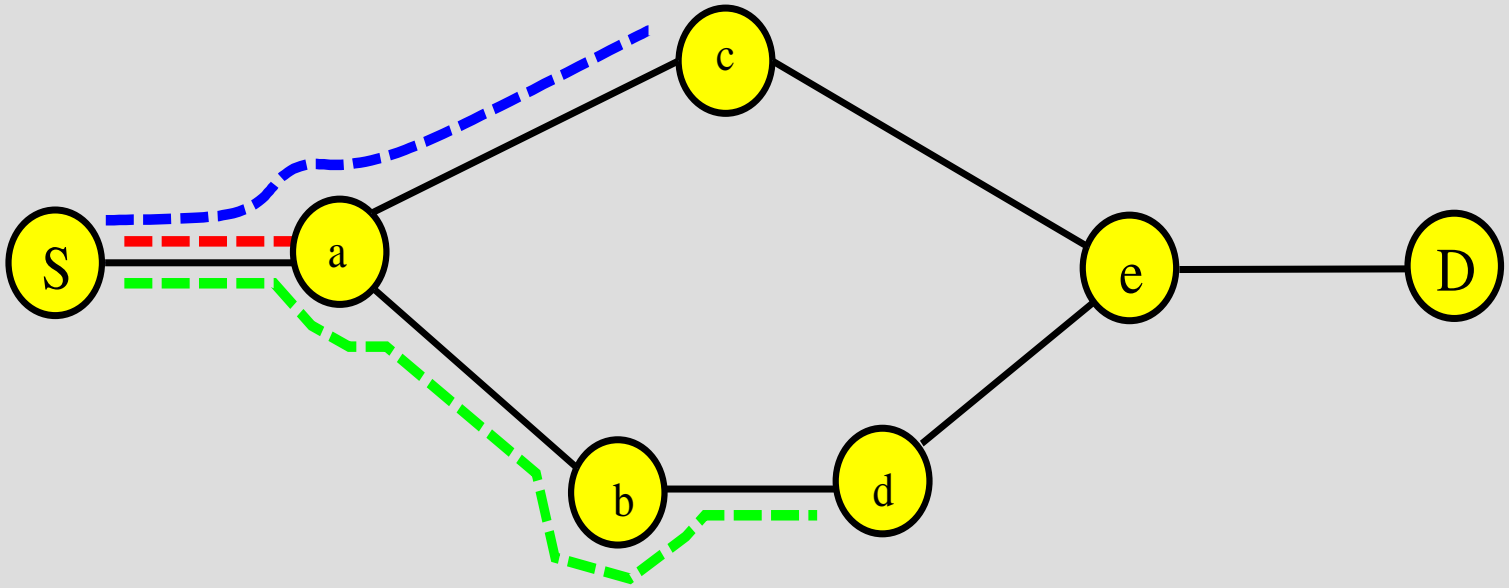
# Alternative Paths
## naïve traceroute



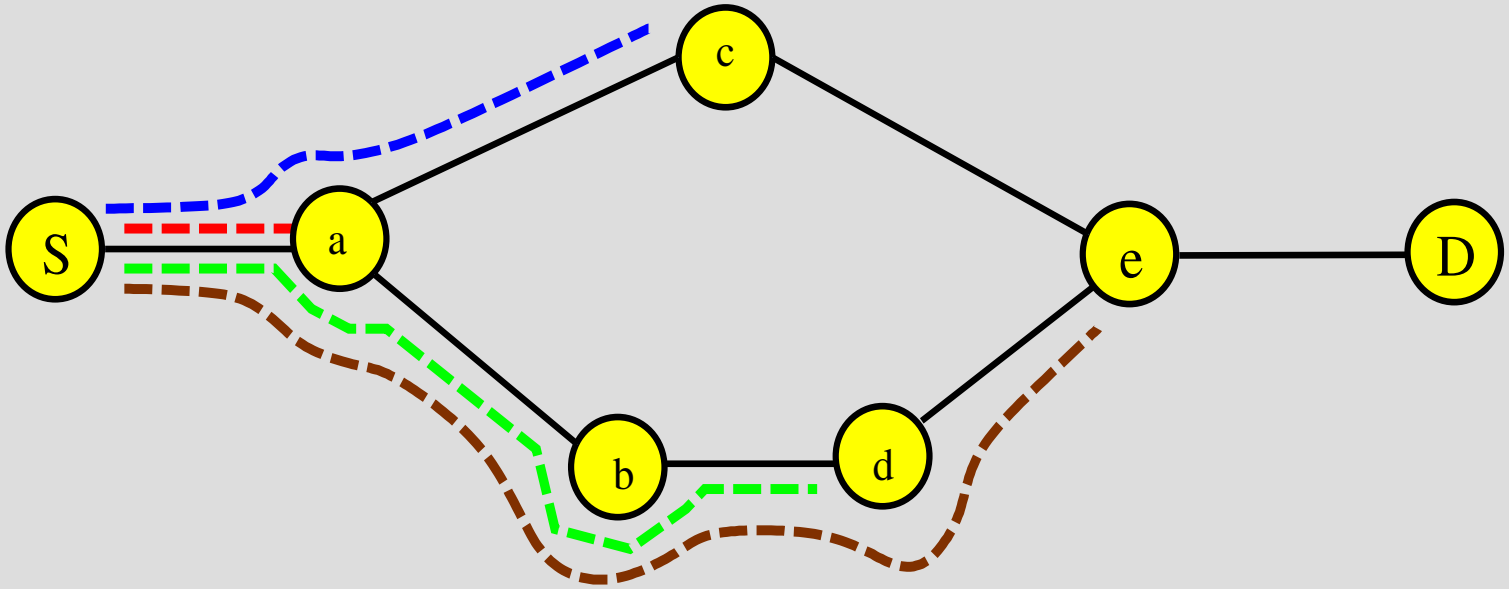
# Alternative Paths
## naïve traceroute



# Alternative Paths
## naïve traceroute

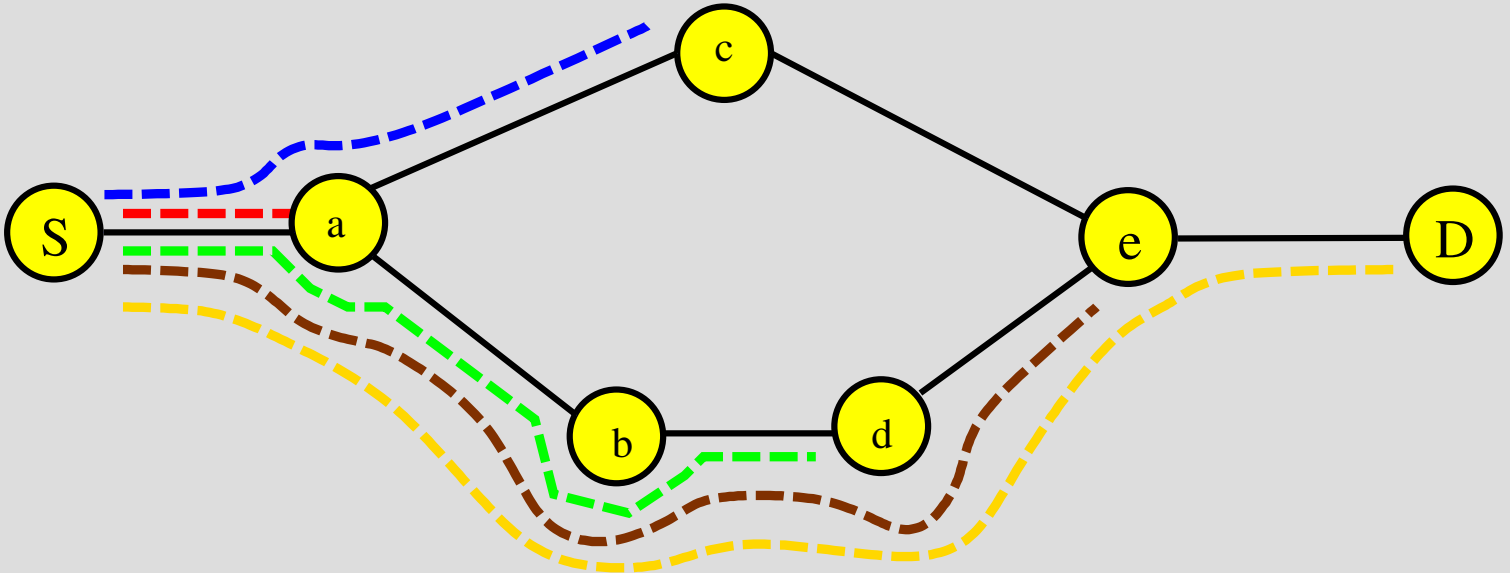


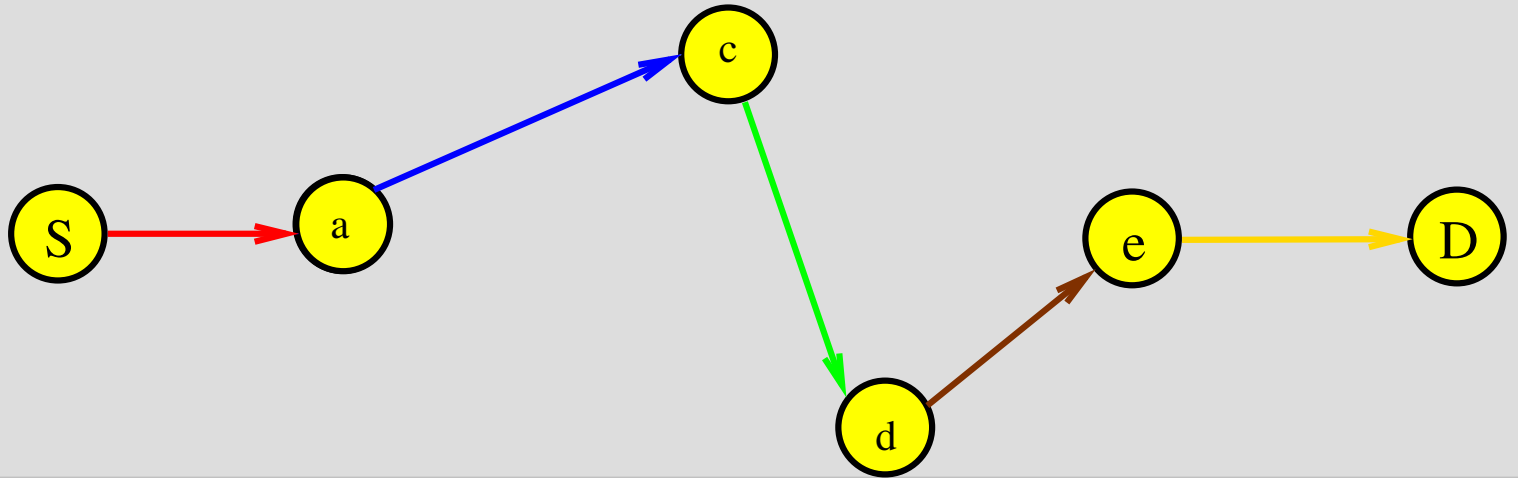
# Alternative Paths
## naïve traceroute



# Alternative Paths

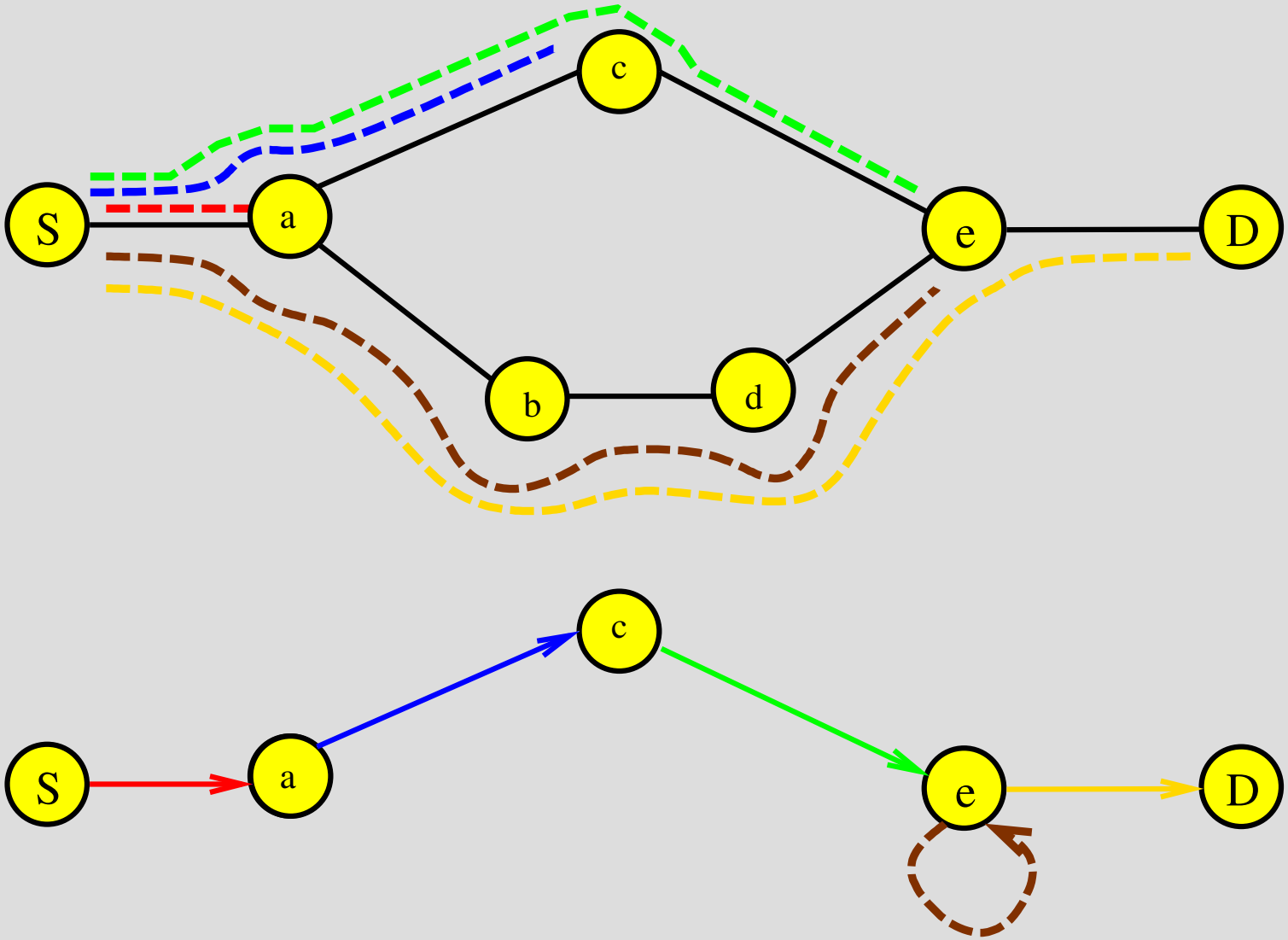
## naïve traceroute



# Alternative Paths

## naïve traceroute



# Load Balancing

- Three types of load balancers
  - Random (per-packet)
  - Destination address based
  - Flow ID based
    - Some or all of the classic 5-tuple
    - TOS
    - ICMP code and checksum
- Traceroute deliberately varies flow information
  - Match TTL expired with probe packet

IP

| Version | IHL | TOS | Total Length |
|---|---|---|---|
| Identification (+) | | Flags | Fragment Offset |
| TTL | Protocol | Header Checksum | |
| Source Address | | | |
| Destination Address | | | |
| Options and Padding | | | |

UDP

| Source Port | Destination Port (#) |
|---|---|
| Length | Checksum (#,*) |

ICMP Echo

| Type | Code | Checksum (#) |
|---|---|---|
| Identifier (*) | | Sequence Number (#,*) |

TCP

| Source Port | Destination Port |
|---|---|
| Sequence Number (*) | |

# Paris Traceroute

- Goal:
  - To discover all paths from a source to a destination
  - Doesn't address per-destination load balancing
- Create headers such that flow ID remains the same
- Still need to match ICMP TTL expired packets to probes
  - For UDP vary checksum
    - Requires manipulating the payload
  - For ICMP vary sequence number and identifier
    - So that checksum remains constant
  - For TCP vary the sequence number
    - Uses TCP half-open technique

# Results

- For a particular set of 5000 destinations from Université Pierre et Marie Curie
- Loops
  - 5.3% of traceroute paths
  - 87% caused by per flow load balancing
- Cycles
  - 0.84% of traceroute paths
  - 78% caused by per-flow load balancing
- Diamonds
  - 79% of traceroute paths
  - 64% cause by per-flow load balancing