Resilient P2P Multicast from the Ground Up

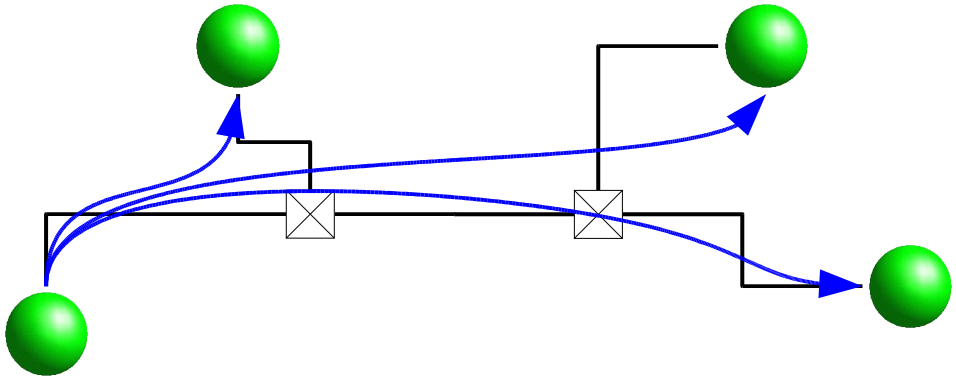
Stefan Birrer & Fabián E. Bustamante
Department of Computer Science
Northwestern University
www.aqualab.cs.northwestern.edu

The Need for Group Communication

- The need for group communication
  - Online gaming (e.g. www.station.sony.com)
  - Video conferencing (e.g. Access Grid)
  - Bulk data dissemination (e.g. BitTorrent)

The Need for Group Communication

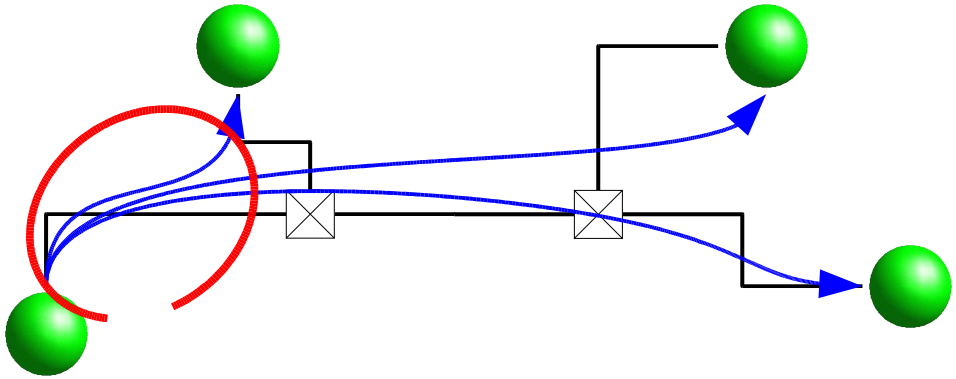
Unicast



- The need for group communication
  - Online gaming (e.g. www.station.sony.com)
  - Video conferencing (e.g. Access Grid)
  - Bulk data dissemination (e.g. BitTorrent)

The Need for Group Communication

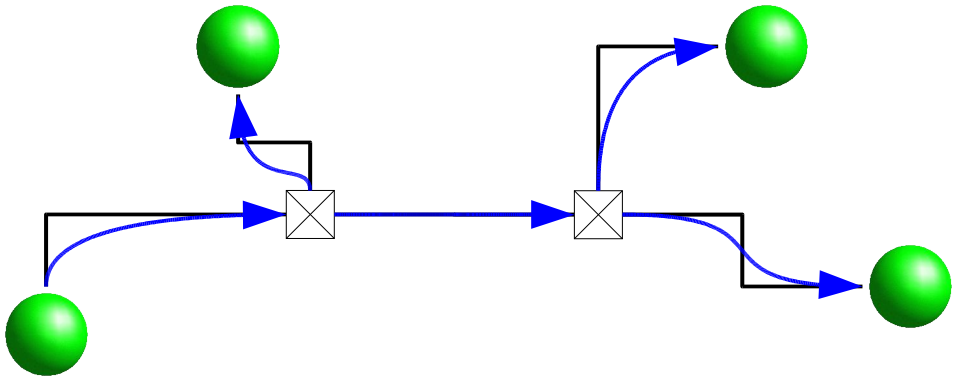
Unicast



- The need for group communication
  - Online gaming (e.g. www.station.sony.com)
  - Video conferencing (e.g. Access Grid)
  - Bulk data dissemination (e.g. BitTorrent)

IP Multicast as one Solution

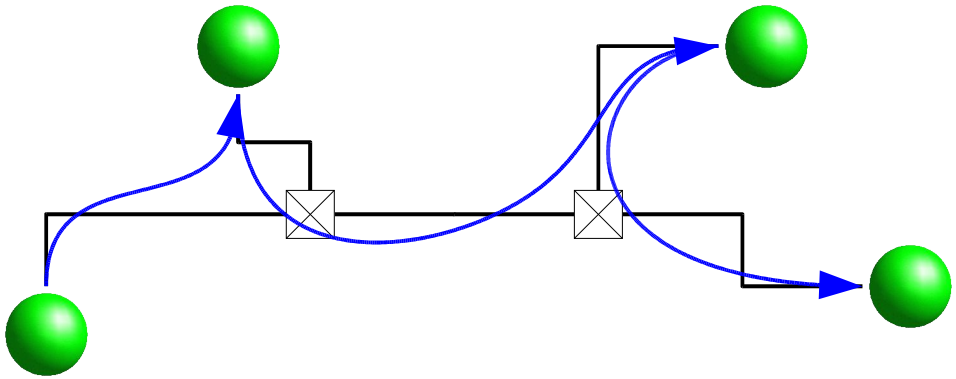
IP Multicast



- Router replicate messages
- Efficient group communication

End System Multicast

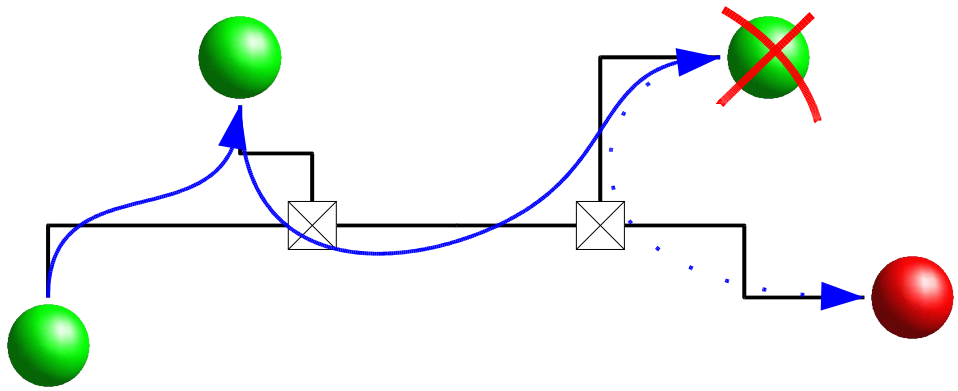
ESM



- But, deployment issues with IP Multicast
  - Security, scalability, ...
- Application-layer or end-system multicast

The Problem with Transiency

ESM



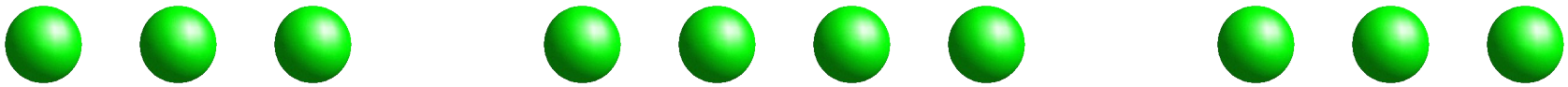
- Median Session Uptime, a good indicator
  - 1 hour to 1 minute [Bustamante03,Gummadi03]

Nemo - Resilient Overlay Multicast

Achieve high delivery ratio w/o paying extra - in latency, duplicates, control traffic

Nemo - Resilient Overlay Multicast

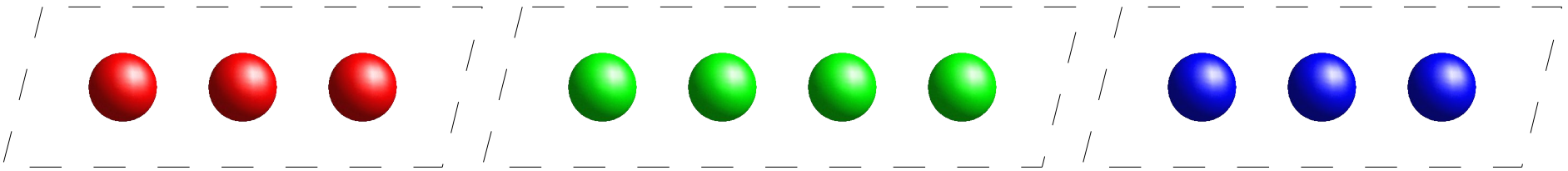
Achieve high delivery ratio w/o paying extra - in latency, duplicates, control traffic



Nemo - Resilient Overlay Multicast

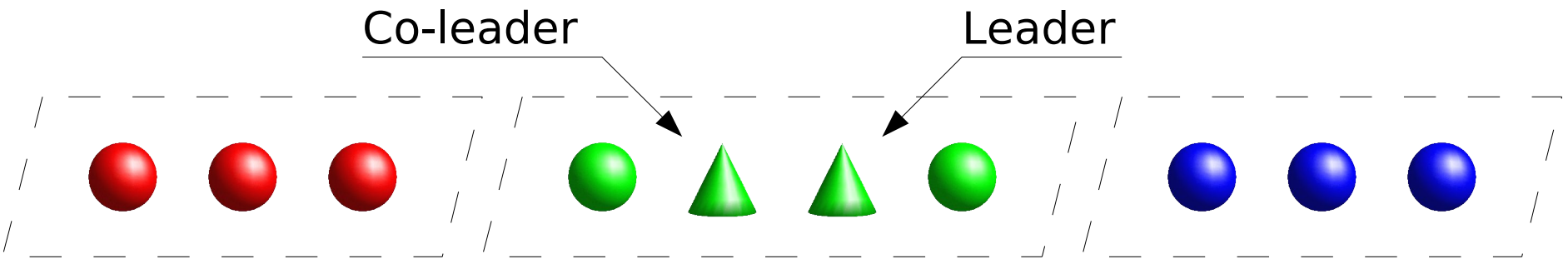
Achieve high delivery ratio w/o paying extra - in latency, duplicates, control traffic

Cluster based on proximity



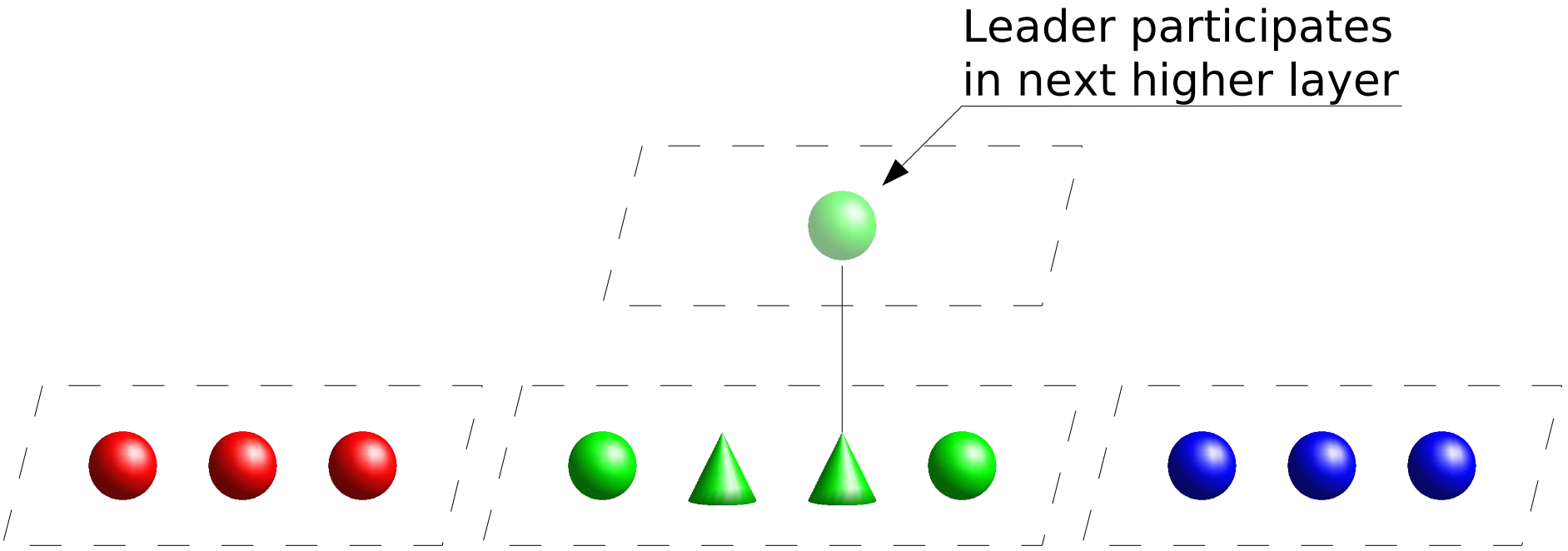
Nemo - Resilient Overlay Multicast

Achieve high delivery ratio w/o paying extra - in latency, duplicates, control traffic



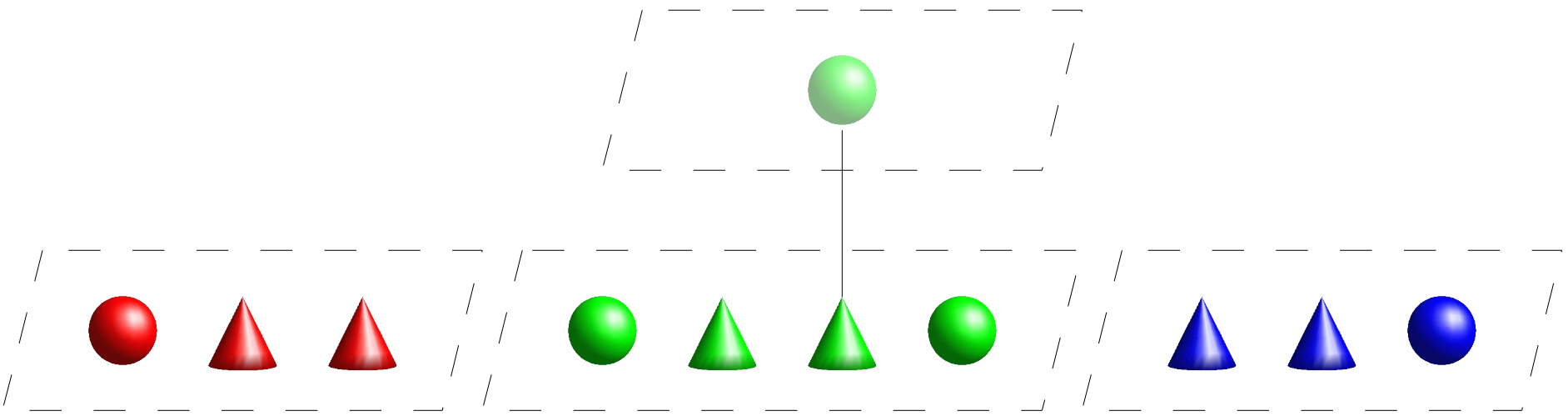
Nemo - Resilient Overlay Multicast

Achieve high delivery ratio w/o paying extra - in latency, duplicates, control traffic



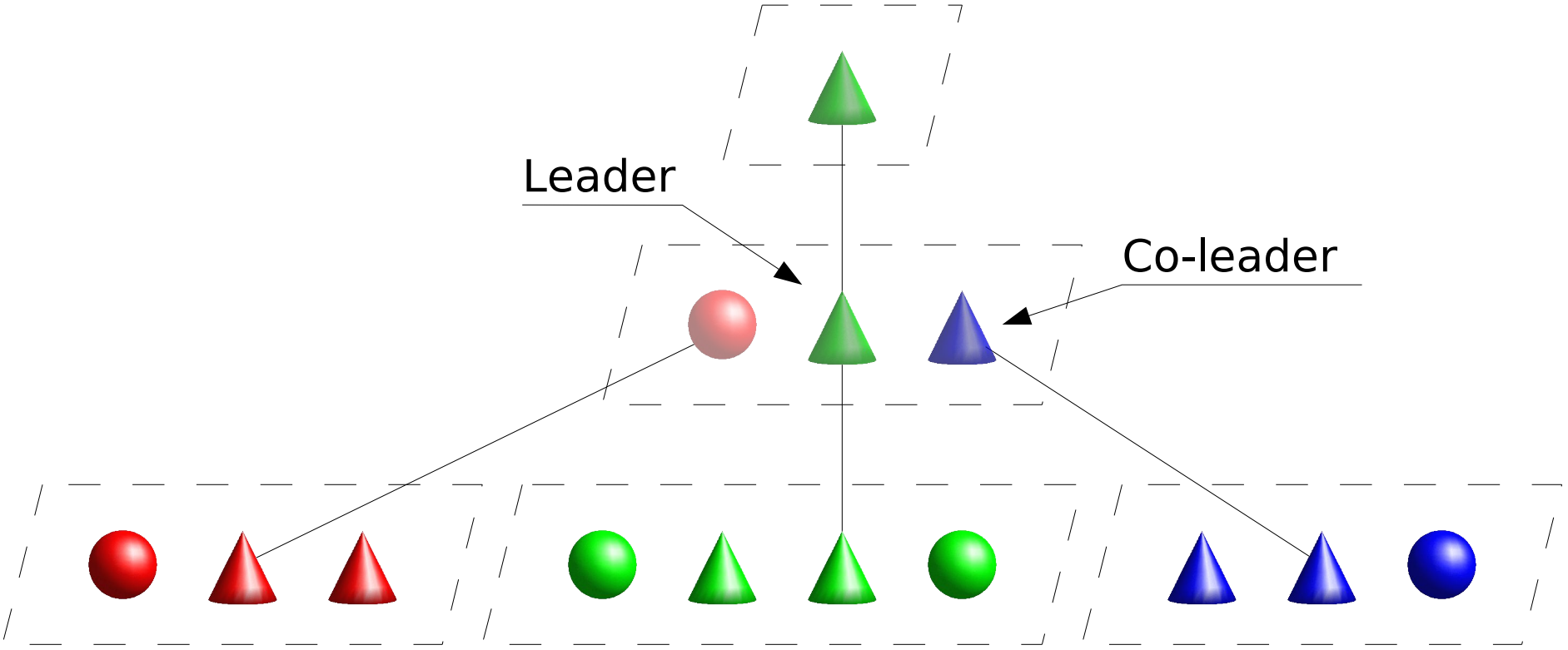
Nemo - Resilient Overlay Multicast

Achieve high delivery ratio w/o paying extra - in latency, duplicates, control traffic

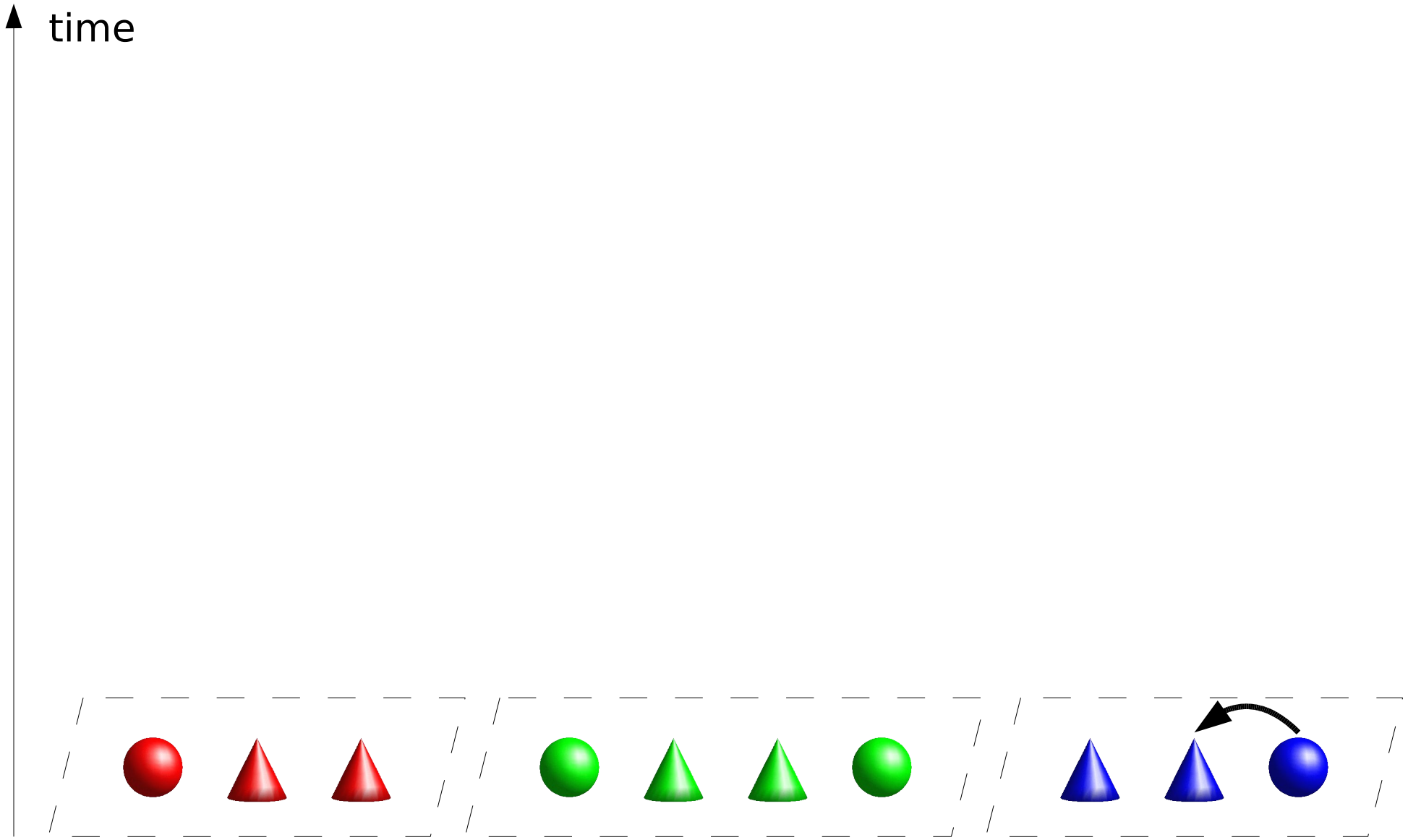


Nemo - Resilient Overlay Multicast

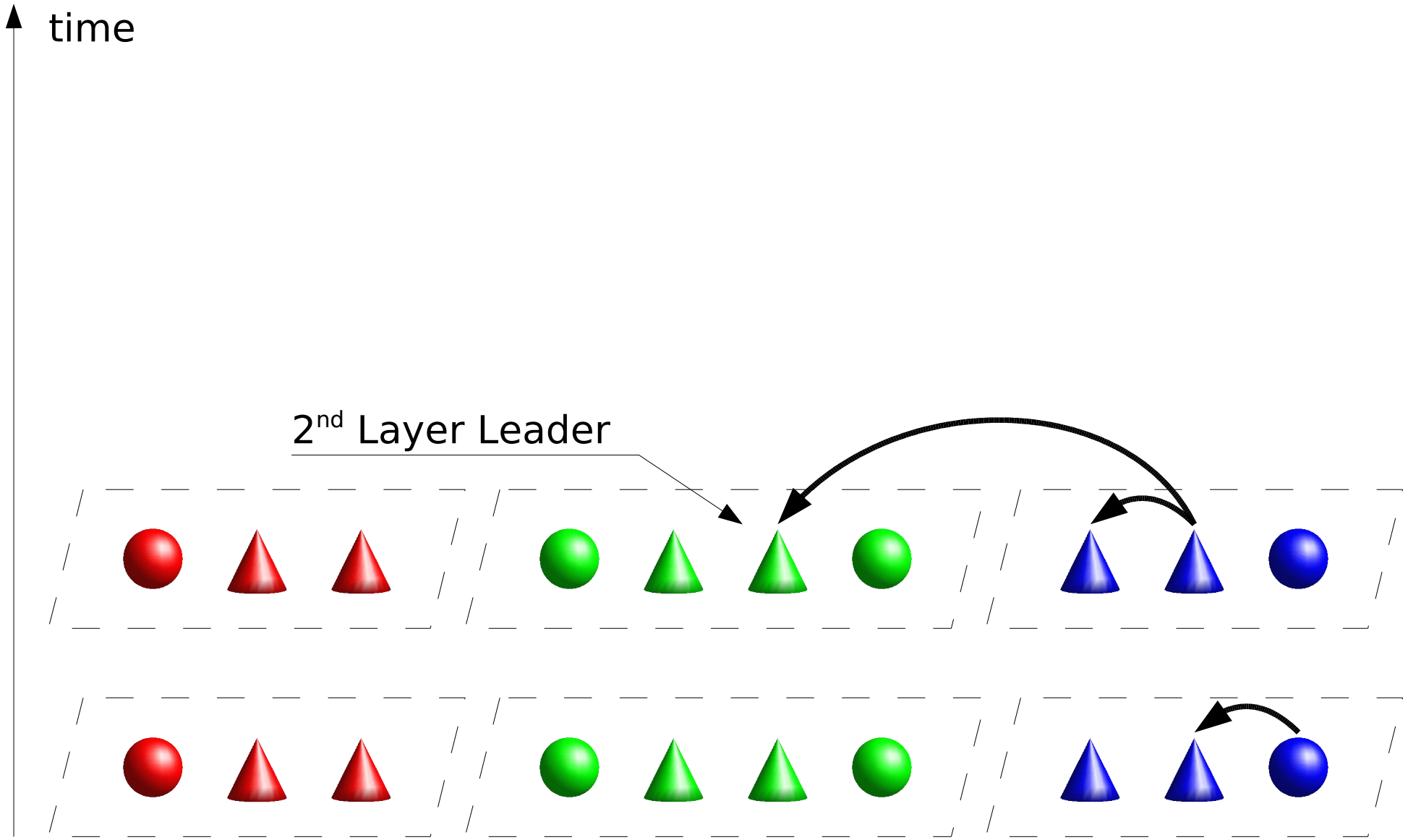
Achieve high delivery ratio w/o paying extra - in latency, duplicates, control traffic



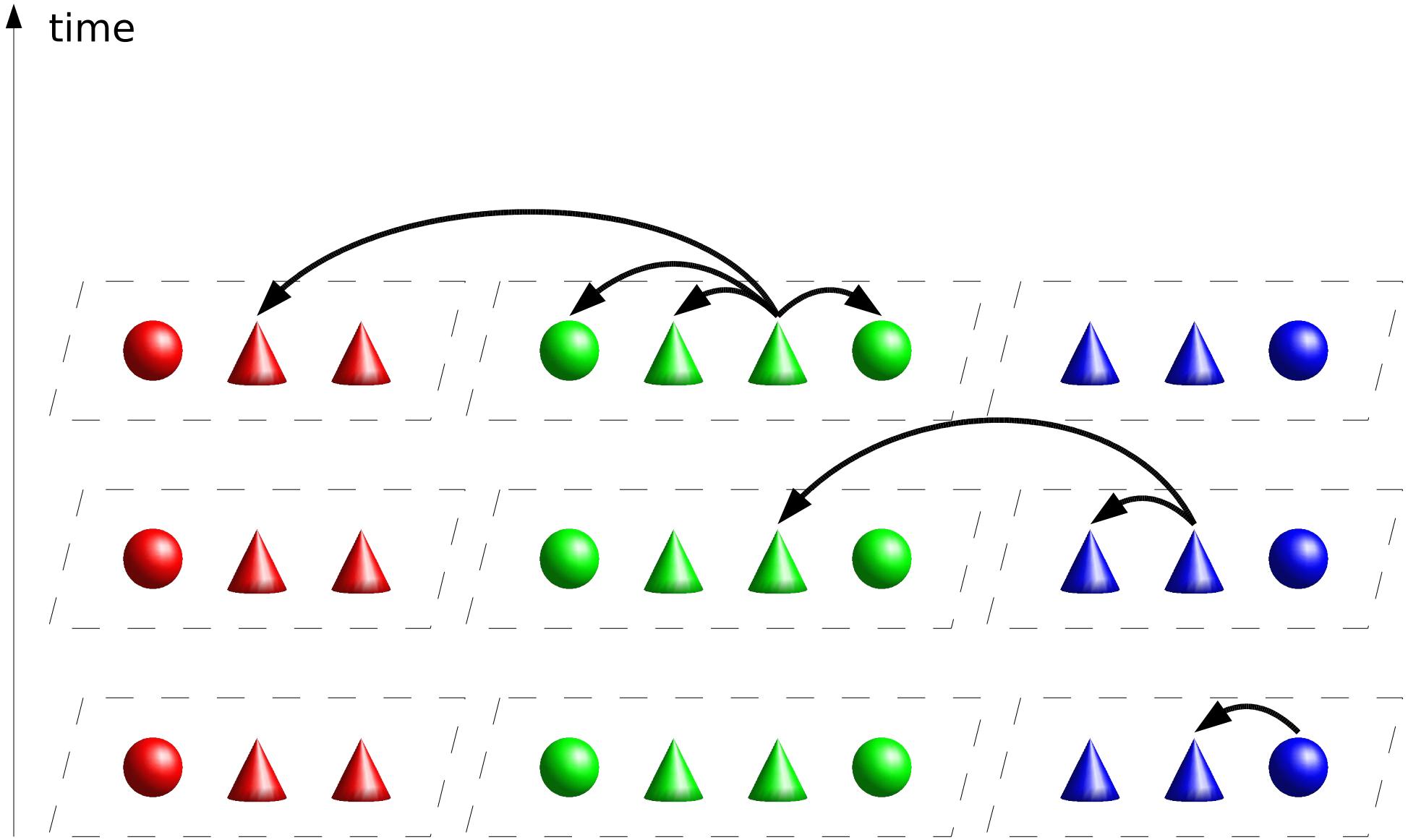
Nemo's Data Forwarding



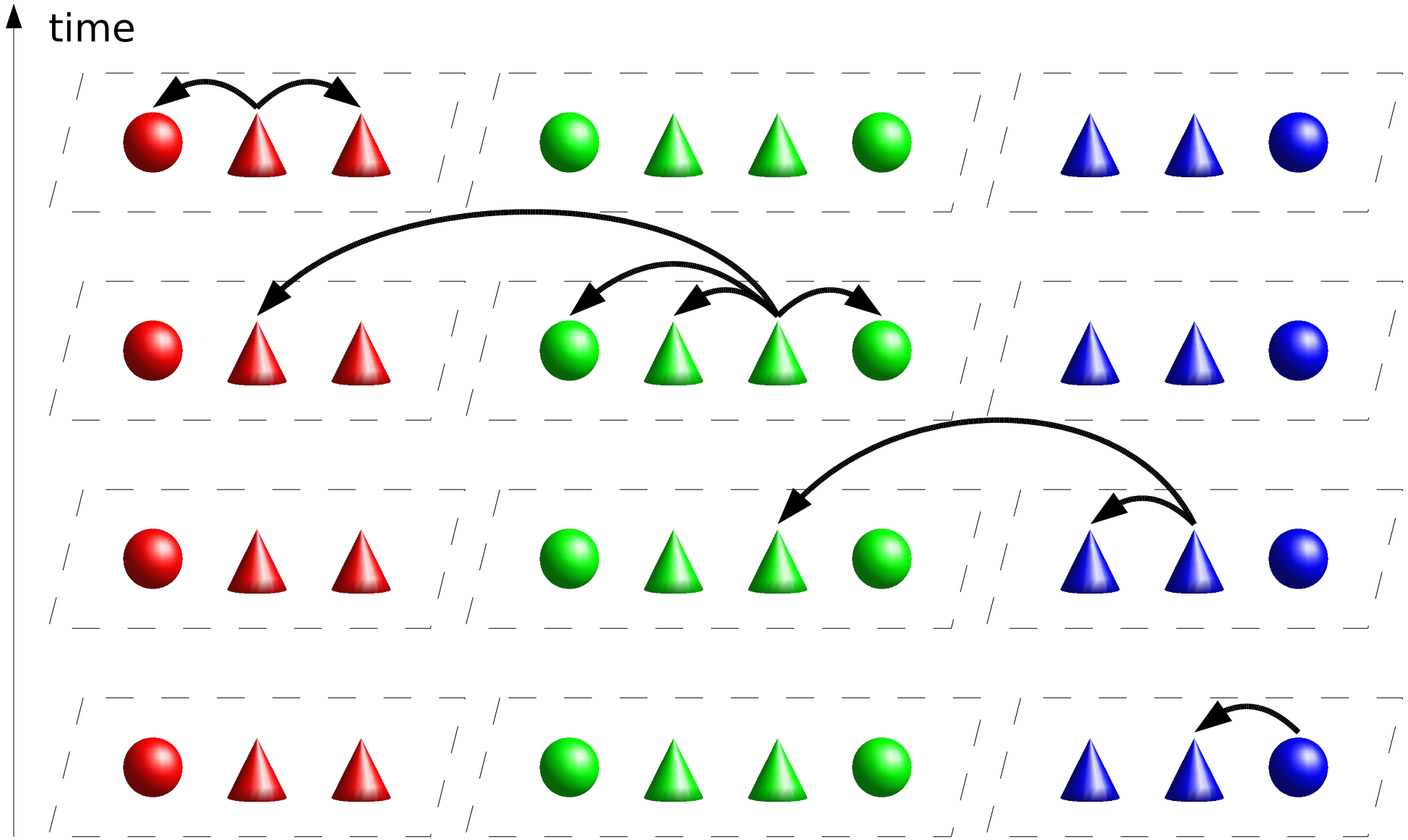
Nemo's Data Forwarding



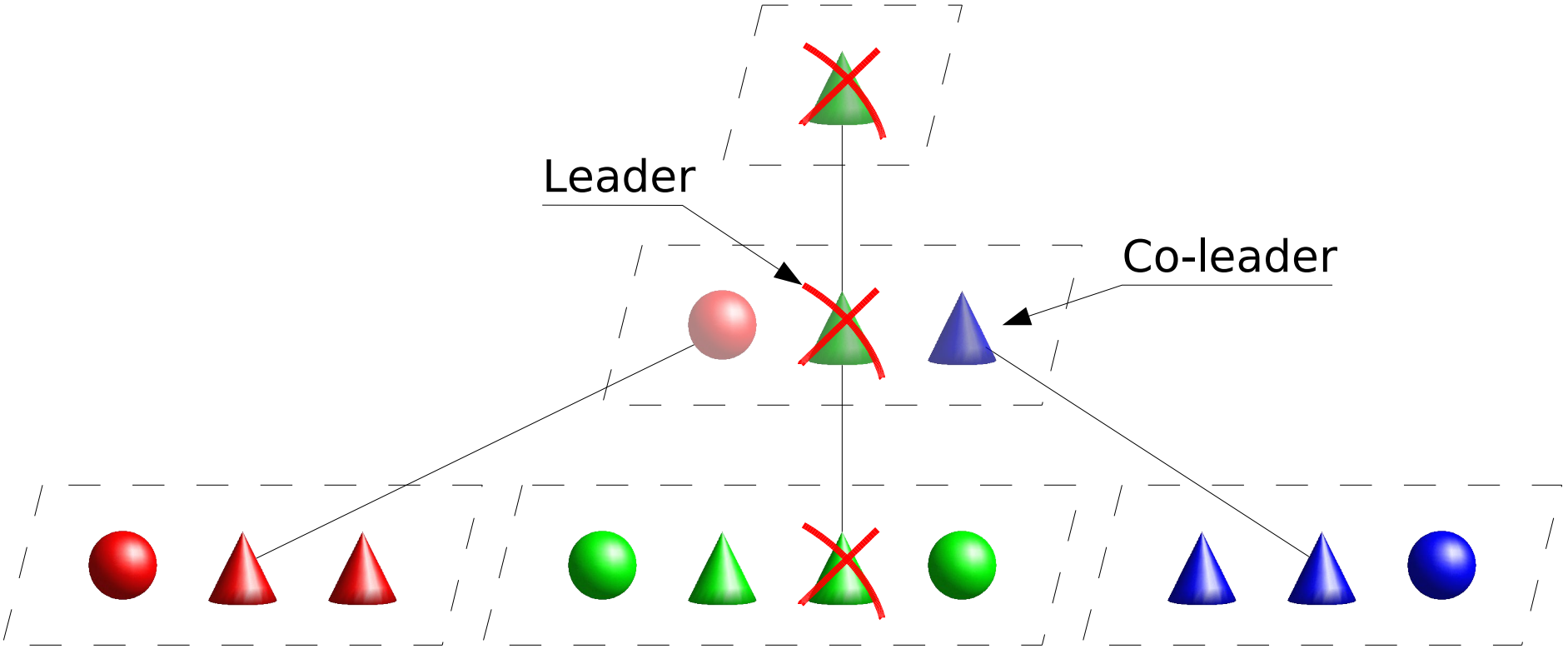
Nemo's Data Forwarding



Nemo's Data Forwarding

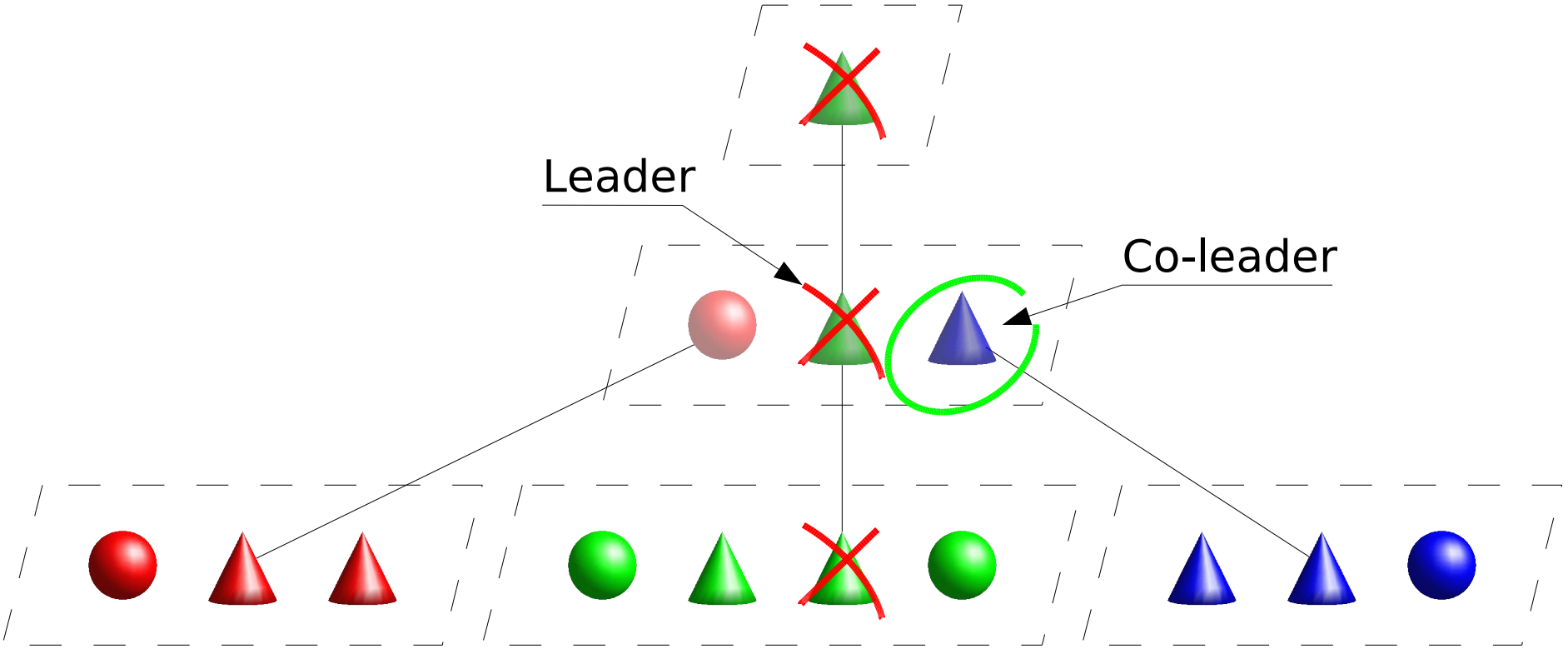


Peer Failure

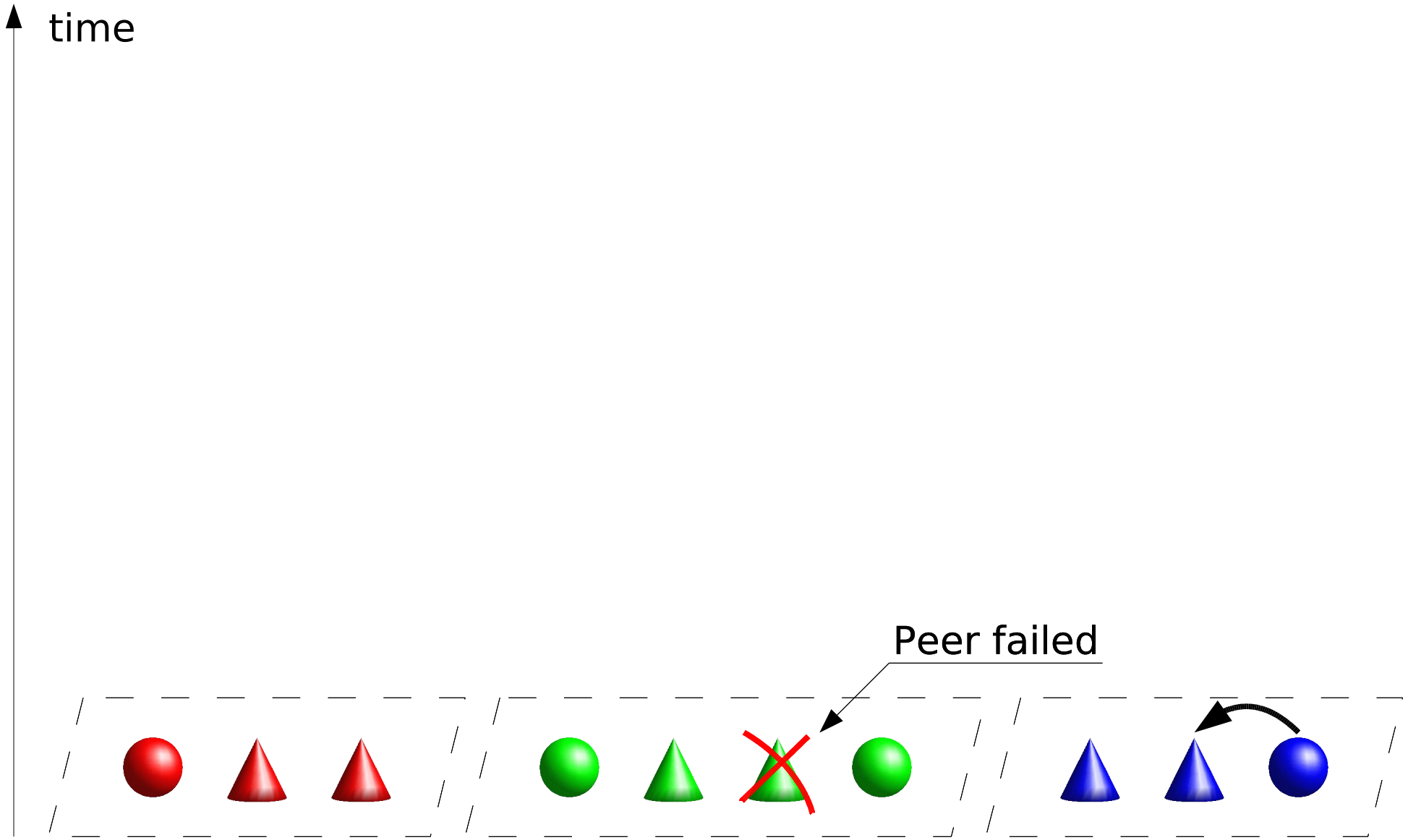


Peer Failure

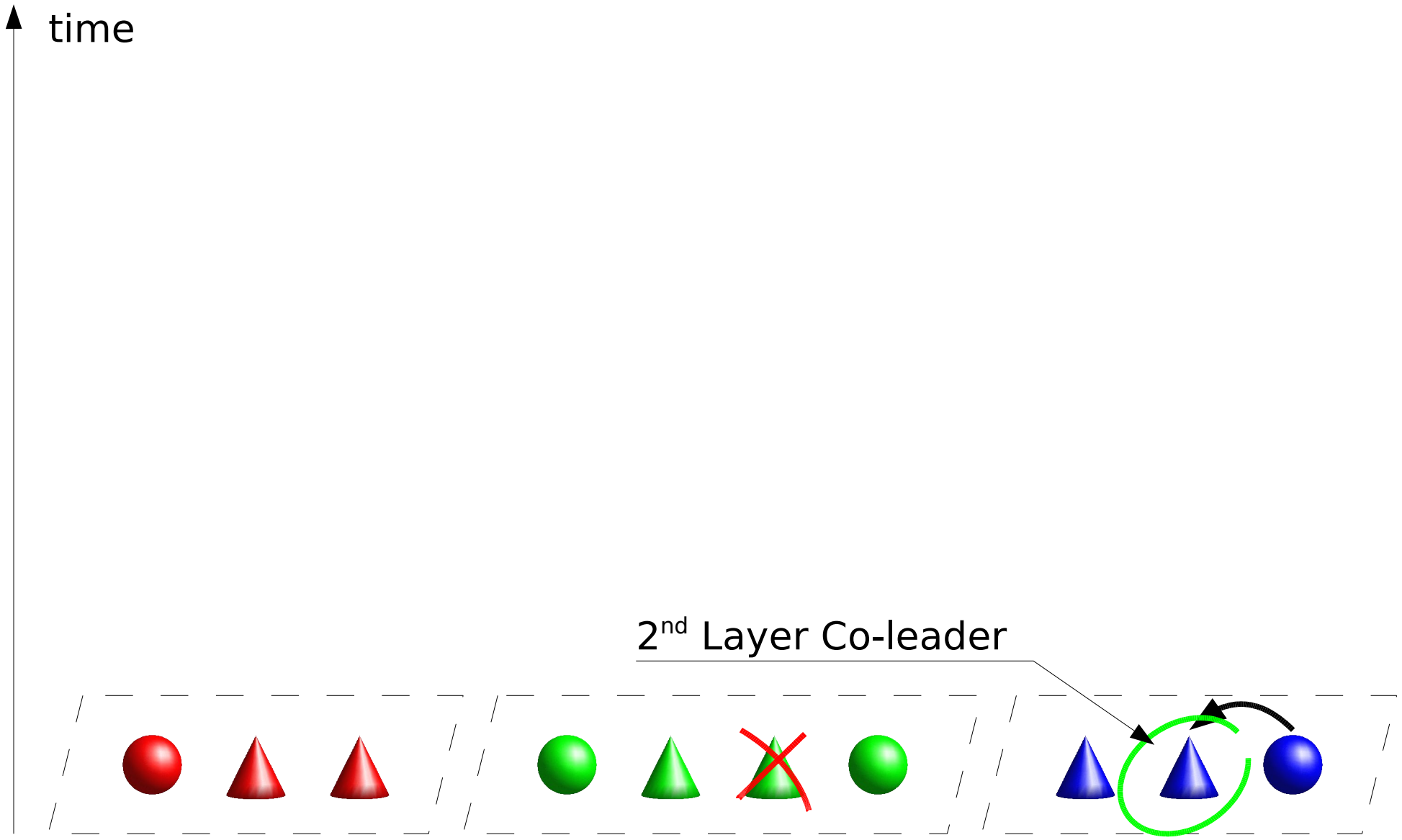
- Co-Leader shares forwarding responsibility with Leader



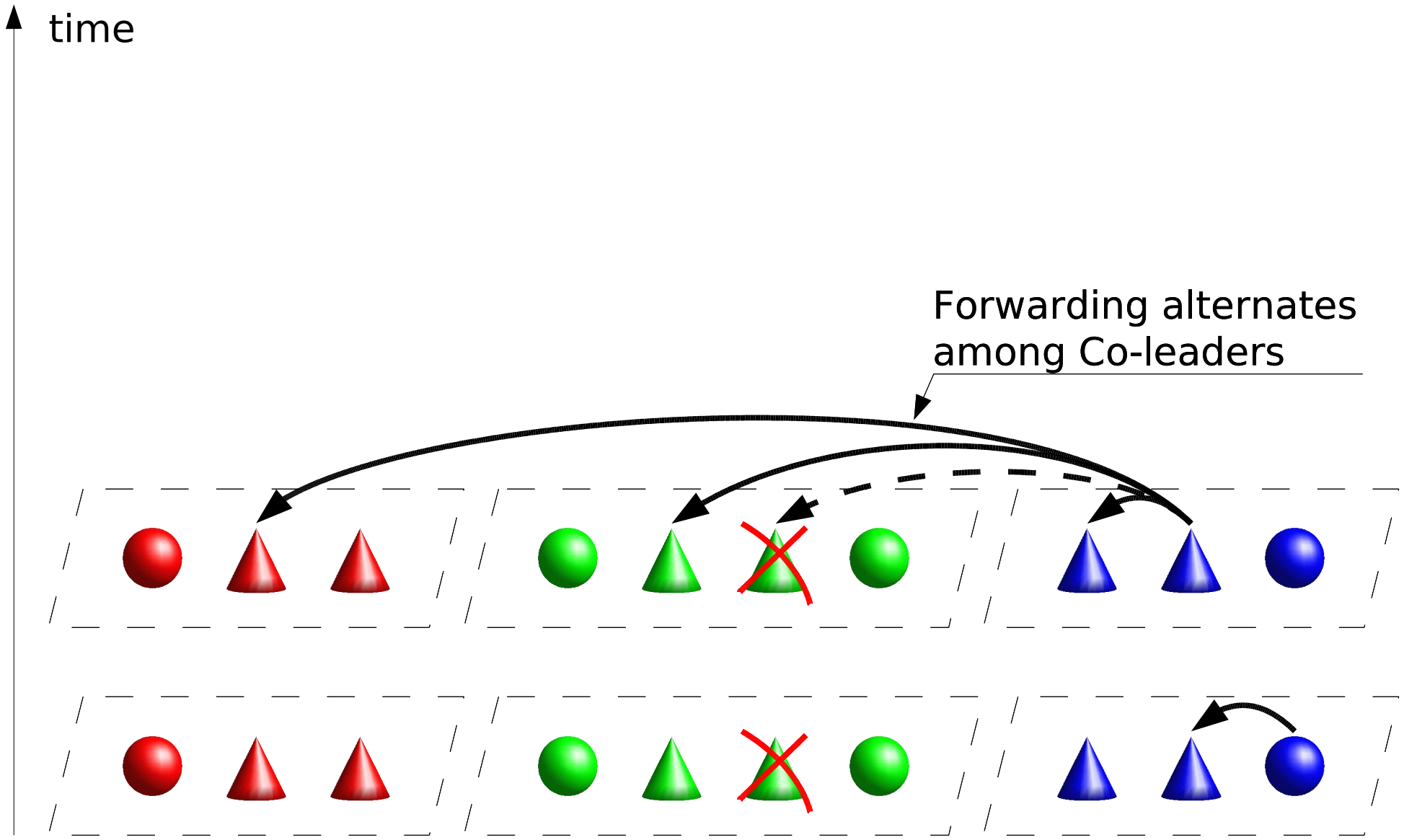
Peer Failure



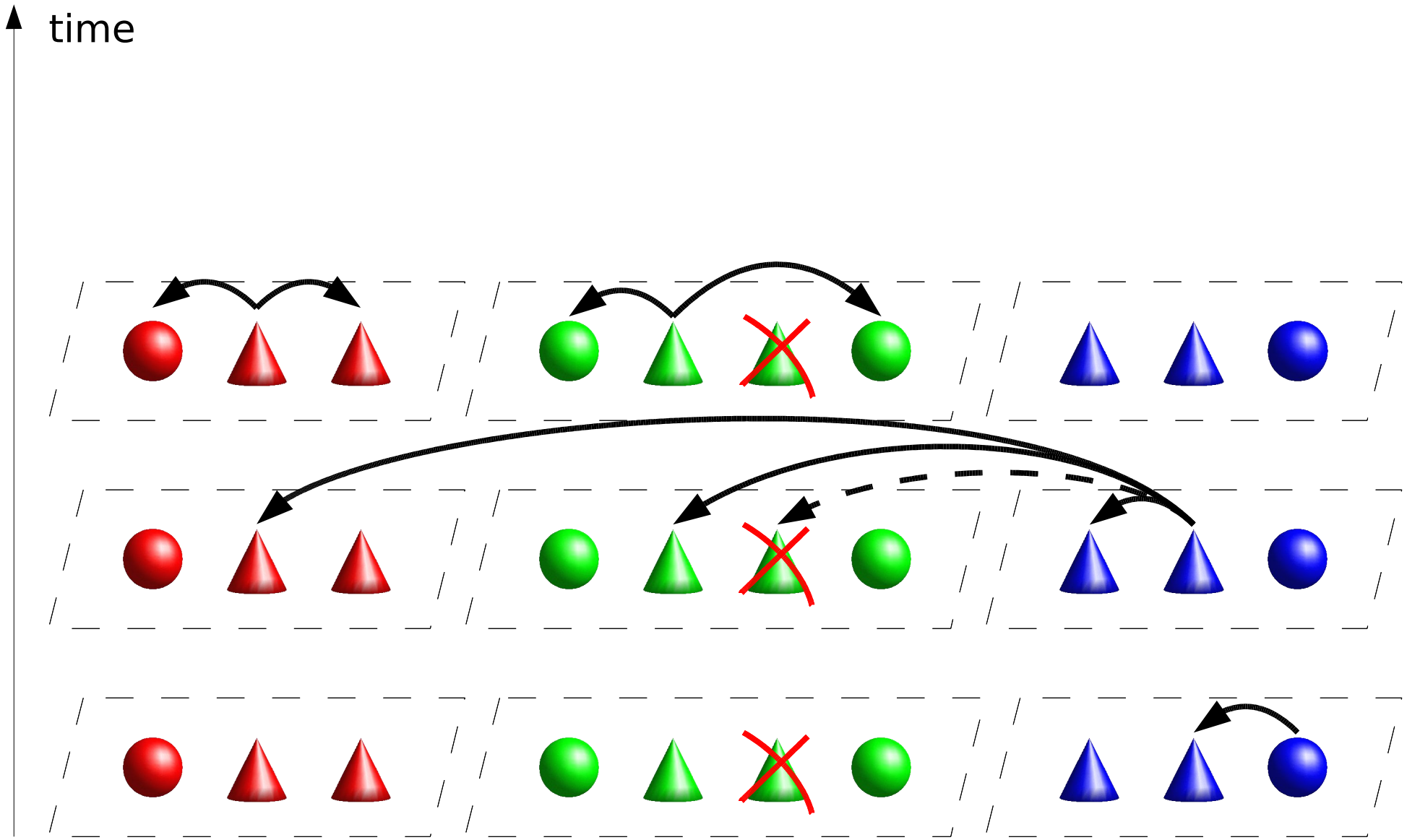
Peer Failure



Peer Failure



Peer Failure



Evaluation

- Measure effectiveness of protocol:
  Delivery ratio
- Cost of resilience:
  Latency and duplicate packets
- Methodology
  - Peers join the session in the warmup time
  - One publisher streams data
- Compare against
  - Nice [Banerjee02], Nice-PRM [Banerjee03], and Narada [Chu02]

Benefits & Costs

High Churn(MTTF 5')

512 end hosts

Best delivery ratio

Protocol	Delivery [%]	Duplicates [packets/SeqNr]
Nemo	0.998	3.16
Nice PRM(3,0.01)	0.993	12.47
Nice PRM(3,0.02)	0.994	18.20
Nice PRM(3,0.03)	0.994	24.22
Nice	0.992	7.10
Narada	0.852	0.00

Wide-Area Results

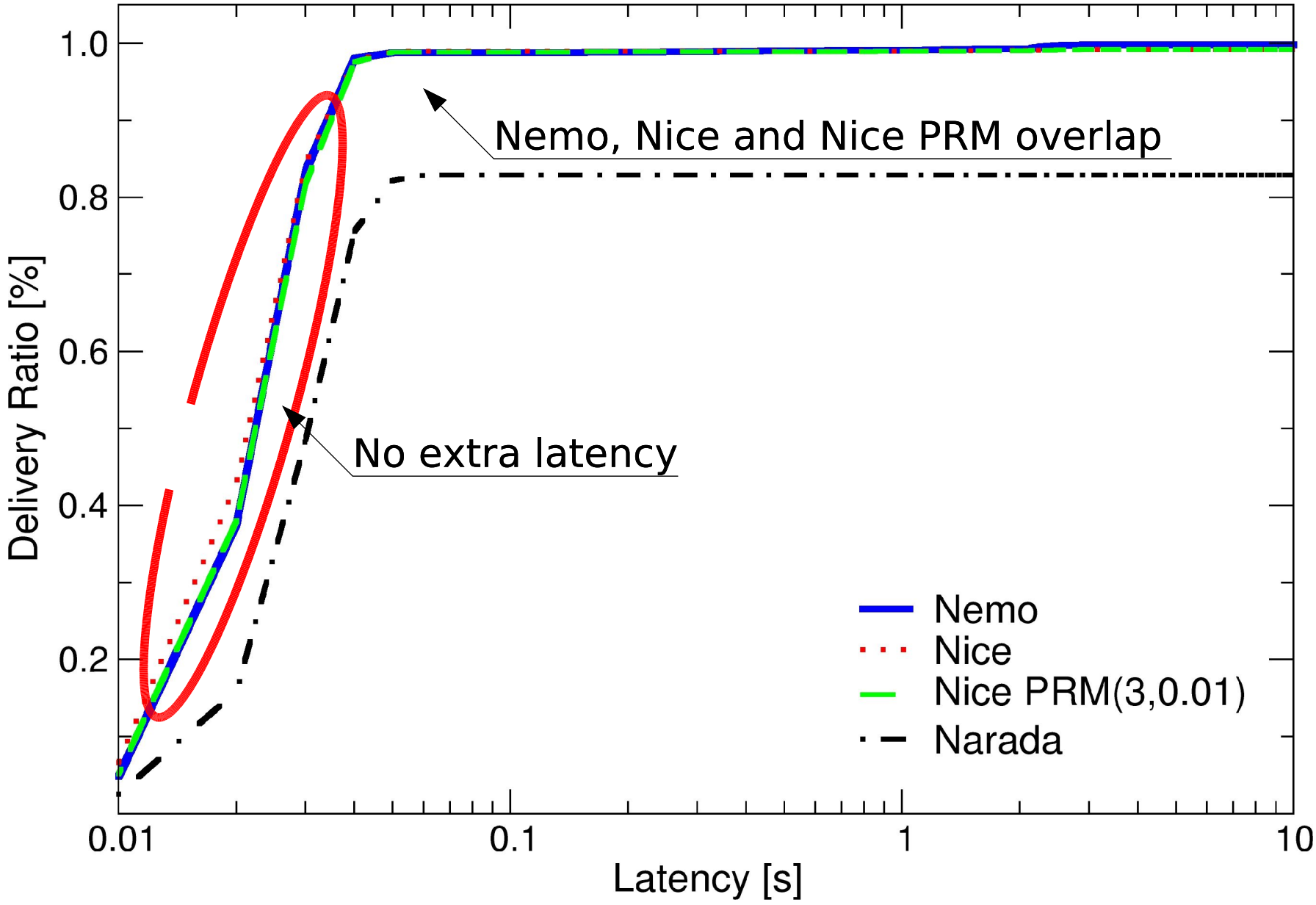
High Churn(MTTF 5')

~72 end hosts

Best delivery ratio

Protocol	Delivery [%]	Duplicates [packets/SeqNr]
Nemo	0.979	1.27
Nice PRM(3,0.02)	0.953	2.02
Nice	0.939	1.06

Benefit & Cost



Conclusions

- Multicast for efficient group communication
  - Transiency can get in the way
- **Co-leaders offer a simple yet effective solution**
  - **Improve resilience**
  - **Spread the load**
- Nemo – Resilient overlay multicast
  - 14.6% higher delivery ratio than Narada
  - 50%-85% less Duplicates than Nice & Nice PRM
  - Comparable end-to-end latency

Nemo: Resilient Overlay Multicast

?

Benefit & Cost

Low Churn(MTTF 60')

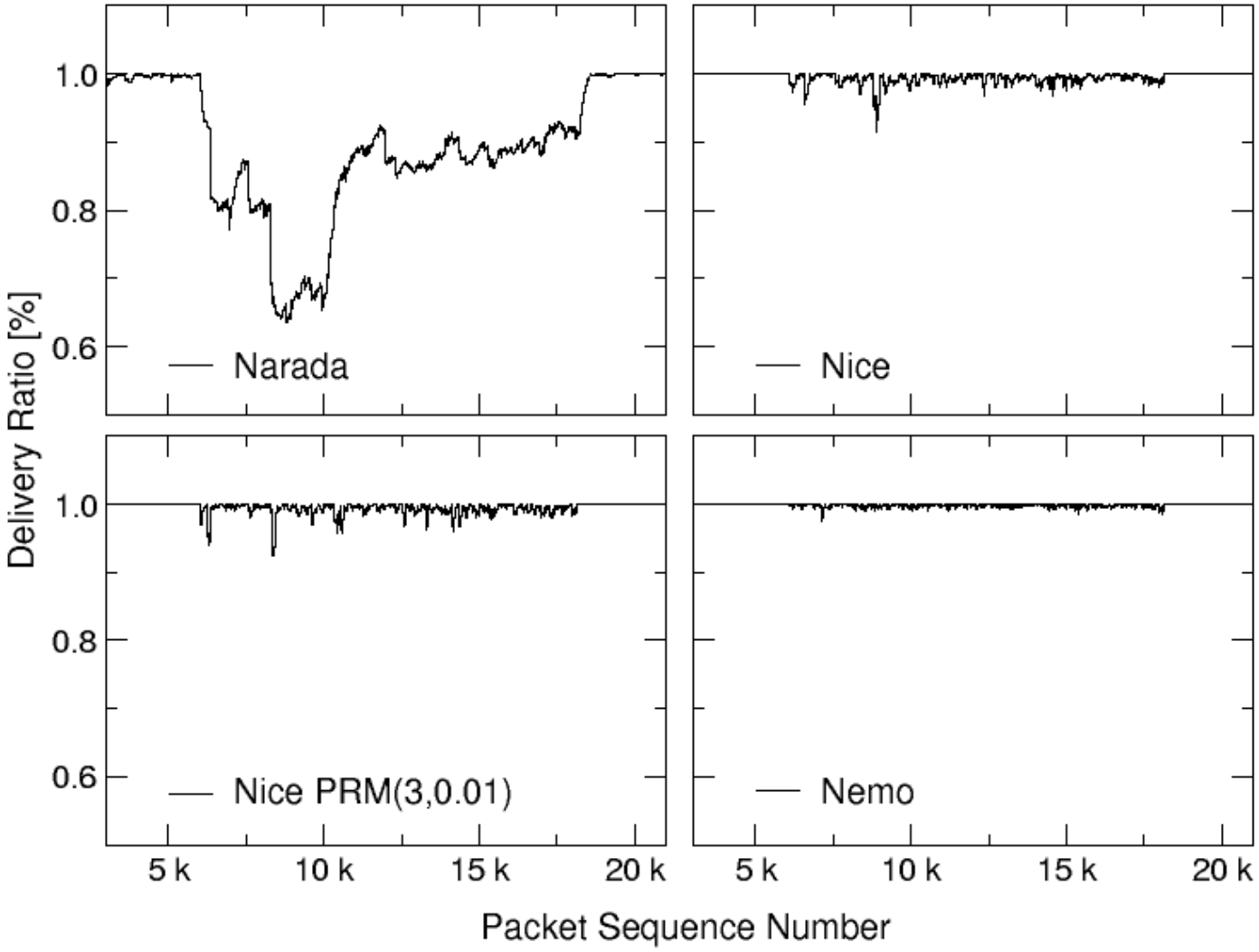
512 end hosts

Best delivery ratio

Protocol	Delivery [%]	Duplicates [packets/SeqNr]
Nemo	1.000	0.34
Nice PRM(3,0.01)	0.999	6.42
Nice PRM(3,0.02)	0.999	12.00
Nice PRM(3,0.03)	0.999	16.74
Nice	0.999	1.29
Narada	0.950	0.00

Delivery Ratio under Churn

High Churn, 512 End Hosts



Related Work

- Overlay multicast
  - Nice (Banerjee02)
  - ESM (Chu00, ...), Yoid (Francis00), ALMI (Pendarakis01), ...
- Resilient multicast
  - A lot of work on resilient IP Multicast
  - PRM - Probabilistic Resilient Multicast for Overlay (Banerjee03)
- Content Dissemination
  - Bullet (Kostic03)
  - SplitStream (Castro03)
  - BitTorrent (Cohen03)