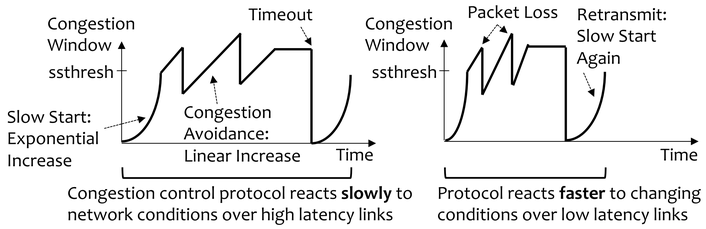
# Cascaded TCP: Big Throughput for Big Data Applications in Distributed HPC

Umar Kalim*, Mark Gardner*+, Eric Brown+, Wu-chun Feng*
Department of Computer Science*, Office of IT+, Virginia Tech,
{umar, mkg, brownej, wfeng}@vt.edu

synergy.cs.vt.edu

## Motivation

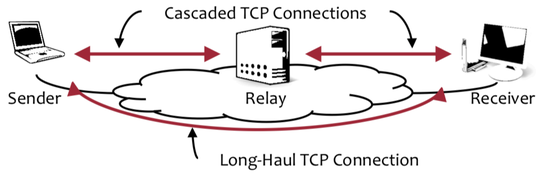
- Vanilla TCP implementations are typically unable to saturate links over high latency and high capacity paths ➔ poor utilization
- Cause: congestion-control feedback is coupled with acknowledgements:
  - Larger the latency, the longer it takes to increase the window size
  - Higher the capacity, the longer it takes to saturate the link




- Bulk data transfer solutions cannot be applied as such to streaming apps
- Streaming Big Data application operating over WAN:


## Hypotheses

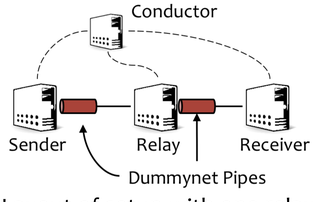
- Throughput improves when layer-4 relay(s) are used to convert a long-haul TCP connection to a **cascade of connections**
  - Particularly for long-lived TCP connections over WAN – long haul




- Improvement in performance justifies costs of layer-4 processing at relays

## Approach & Experimental Setup

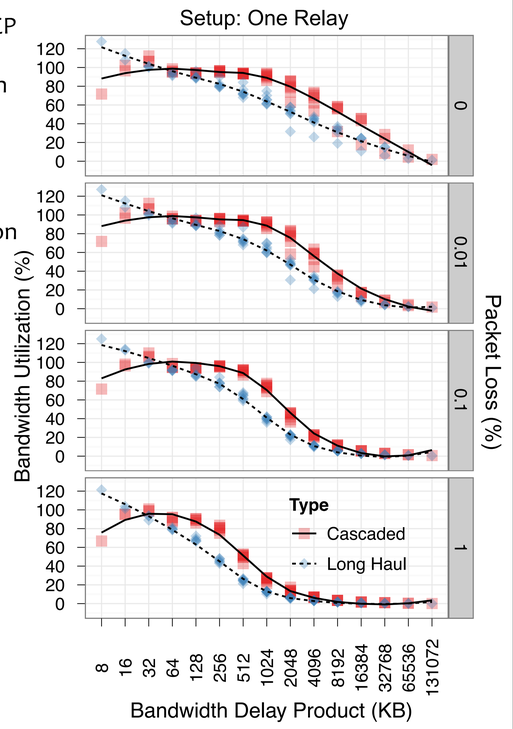
- Goal: Characterize performance of Cascaded TCP vs. Long-Haul TCP for Big Data in an emulated environment
- Configuration:
  - FreeBSD v9.0
  - TCP New Reno
  - Dummynet ( ~ 1000Hz)
  - netcat
  - iperf
- A conducting node sets up the receiver, relay(s) and sender, and maintains timestamps

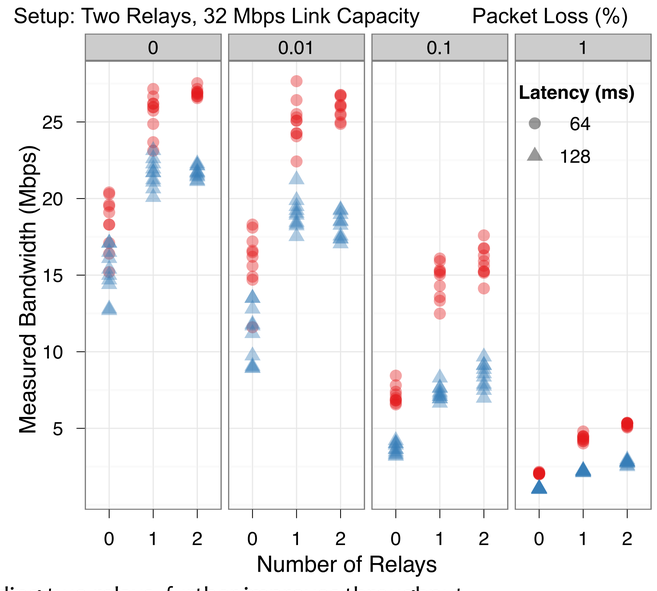



Layout of setup with one relay

## Results

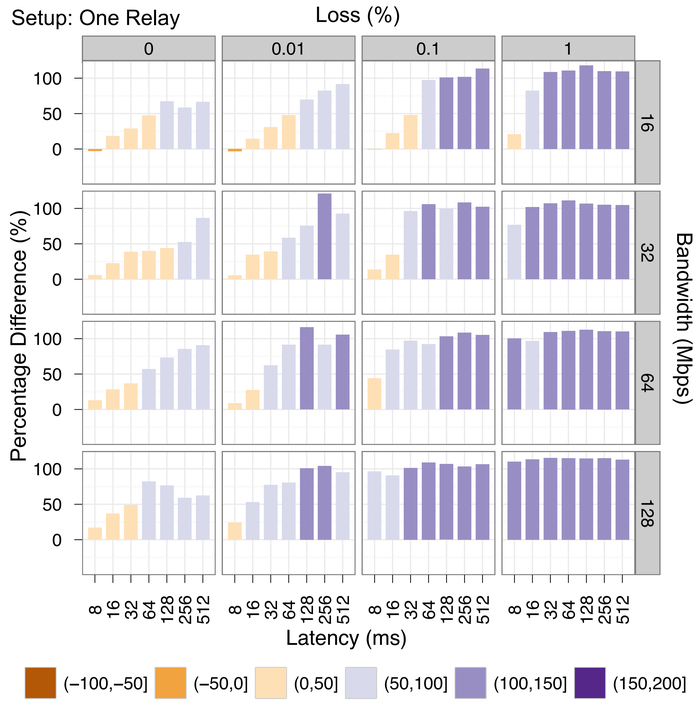
- Increased throughput observed with Cascaded TCP
- Cross-over point is between a bandwidth-delay product of 32KB and 64KB
- As bandwidth-delay products increase, utilization is accentuated before we see diminishing returns
- Cascaded TCP ameliorates impact of losses; as losses increase, the cross-over point – where Cascaded TCP performs better – moves towards smaller bandwidth delay products (i.e., 64KB to 32KB)
- Owing to relay setup time, for small bandwidth-delay products, long-haul TCP results in better utilization than Cascaded TCP




- Adding two relays, further improves throughput
- As high as 90% utilization is observed when the E2E latency is 64ms and loss is 0% (Note: protocol overheads limit peak to ~94%)

- Even when utilization is low (at high bandwidth-delay products), Cascaded TCP shows 100% improvement in throughput




Long-Haul TCP vs. Cascaded TCP: Percentage Differences

## Related Work


- H. Y. Pucha and C. Hu, "Slot: Shortened Loop Internet Transport using Overlay Networks," Purdue University, TR-ECE-5-12, 2005.
- A. Cohen, S. Rangarajan, and H. Slye. "On the Performance of TCP Splicing for URL-Aware Redirection", In USENIX Symposium on Internet Technologies and Systems - (USITS), 1999.


## Future Work

- How many relays are needed for optimal performance?
- Is the overhead of relay setup within tolerable limits?
- What are the implications of for TCP semantics?
- Are results also applicable to other congestion-control algorithms?
- Use the Cascaded TCP framework for NRAO/VLA data transfer

## Acknowledgements

Research was funded in part by Virginia Tech and Juniper Networks