Scalability and Accuracy in a Large-Scale Network Emulator

Nov. 12, 2003

Byung-Gon Chun

Contents

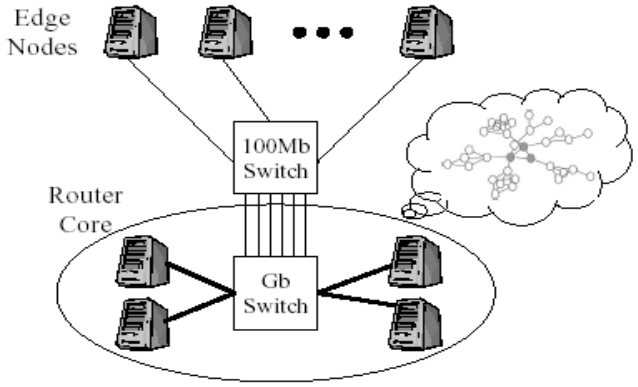
Motivation

- Real-world experiments
  - It is hard to deploy and administer research software at distributed sites
  - Results are not reproducible
- Simulation
  - It misses important real system interactions
  - Tools do not support direct execution of software
- Emulation
  - Unmodified applications
  - Almost reproducible results
  - Important system interactions

Contribution

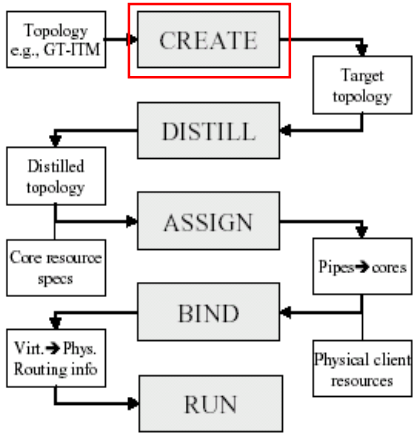
- *Large-scale* network emulator that can emulate *network dynamics* in a *topology*
- Techniques to trade increased *scalability* for reduced *accuracy*
- Evaluation of a range of systems

Architecture



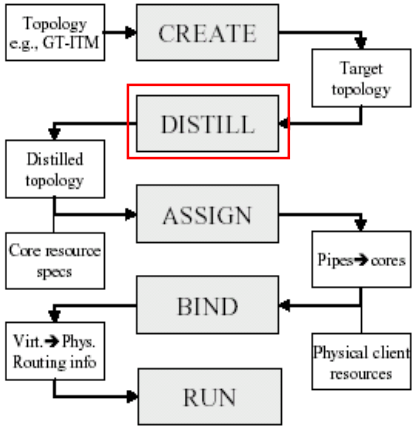
EN
VN
CR
Pipe

ModelNet Phases



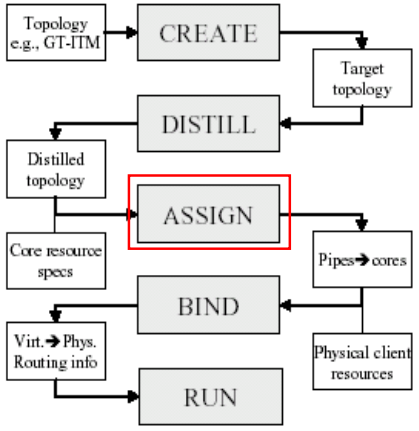
- Generate a network topology
- Source: Internet traces, BGP dumps, Synthetic topology
- Target: GML

ModelNet Phases



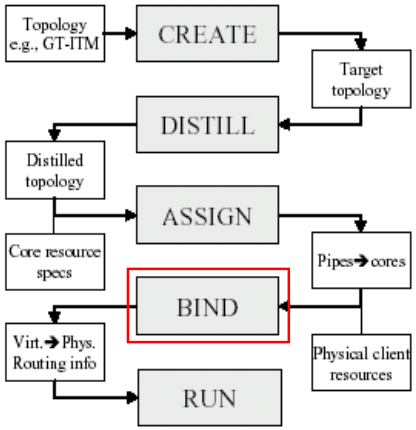
- Source: GML
- Target: Pipe topology
- Optionally, simplify the network

ModelNet Phases



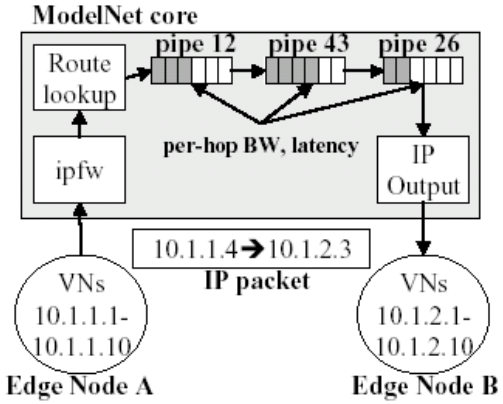
- Map components of the distilled topology to core nodes
- Minimize the bandwidth demands
- Greedy algorithm

ModelNet Phases



- Assign VNs to ENs
- Multiplex VNs
- Bind an EN to a CR
- Generate configuration scripts
- Install pipes and routing tables to CRs
- Set up IP addr for VNs

Core Router



- Routing table
  - Memory vs Computation Tradeoff
- Descriptors referencing buffered packets
- Schedule descriptors on pipes using a heap
- Admission control of packets injected
- Multihop and multi-core

Discussion

- Resource isolation
  - UDP flows
- Routing protocol
  - Assumes shortest path routing and instantaneous repair of node and link failures
  - Routing protocol dynamics (update propagation)
  - Flexibility to inject routing algorithms
    - *Policy-based* routing

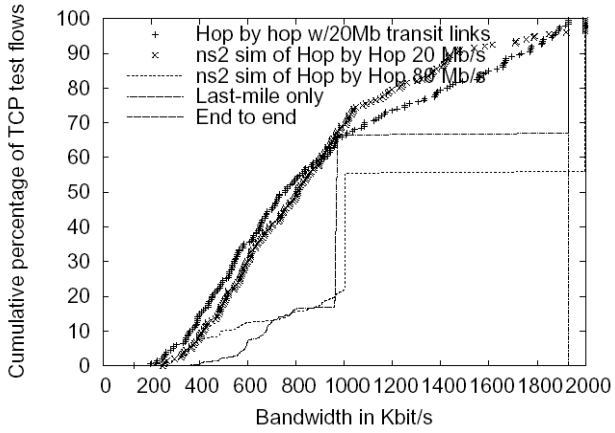
Evaluation of Implementation

- Accuracy - 1ms worst-case error per hop
- Capacity
  - 120000 packets/s with 1 hop
  - 90000 packets/s with 8 hops
- Core crossing
  - For 4 node configuration and 2 hop paths,
    0% cross-core traffic – 462.5Kpkt/s
    100% cross-core traffic – 155.8Kpkt/s
  - Depends on target topology and application communication patterns

Accuracy vs. Scalability Tradeoffs

- Continuum balancing accuracy and cost
- Distillation
  - Reduce the diameter of the network
  - Hop-by-hop
  - End-to-end: bottleneck bandwidth, latency, loss rate
  - Walk-in and walk-out(?)
  - Other approaches? – Flow-based [Narses]
  - Tradeoffs? — No wall-clock time or network size?

Accuracy vs. Scalability Tradeoffs



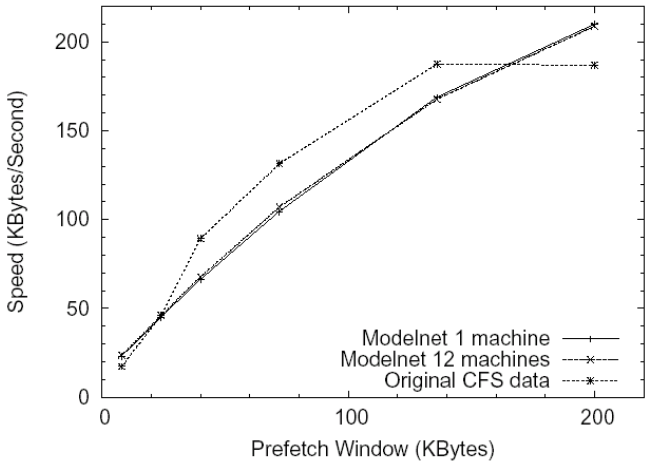
Accuracy vs. Scalability Tradeoffs

- VN multiplexing
  - Howto? Mapping of multiple VNs to a EN
  - Context switch, scheduling and resource contention
  - Concurrency model
  - Number of processes ↑ → Aggregate thp. ↓
- Network characteristics
  - Howto? Synthetic background cross traffic in CRs
  - Do not capture all packet dynamics (TCP slow start, bursty traffic)
  - Network change - recompute shortest paths
    - Fault injection like partitions, bandwidth change

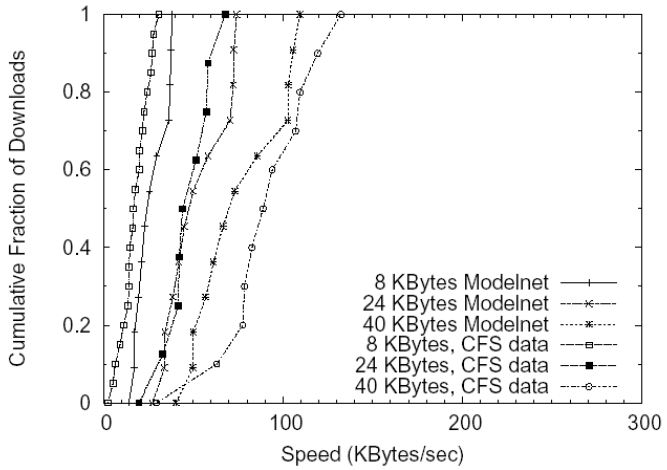
Case Studies

- 10000 node gnutella network emulation
- CFS, Replicated web services, Adaptive overlays, Ad hoc networking(?)
- CFS: reproduce the published experimental results
- Replicated web services: support for realistic Internet topologies and emulation of contention of shared pipes
- Adaptive overlays: subject systems to dynamically changing network conditions

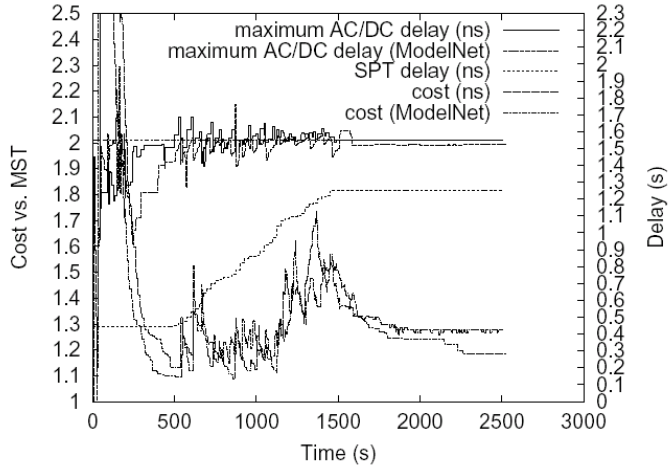
Case Studies



Case Studies



Case Studies



Discussion?