University of Glasgow

Wider, Not Faster

Colin Perkins

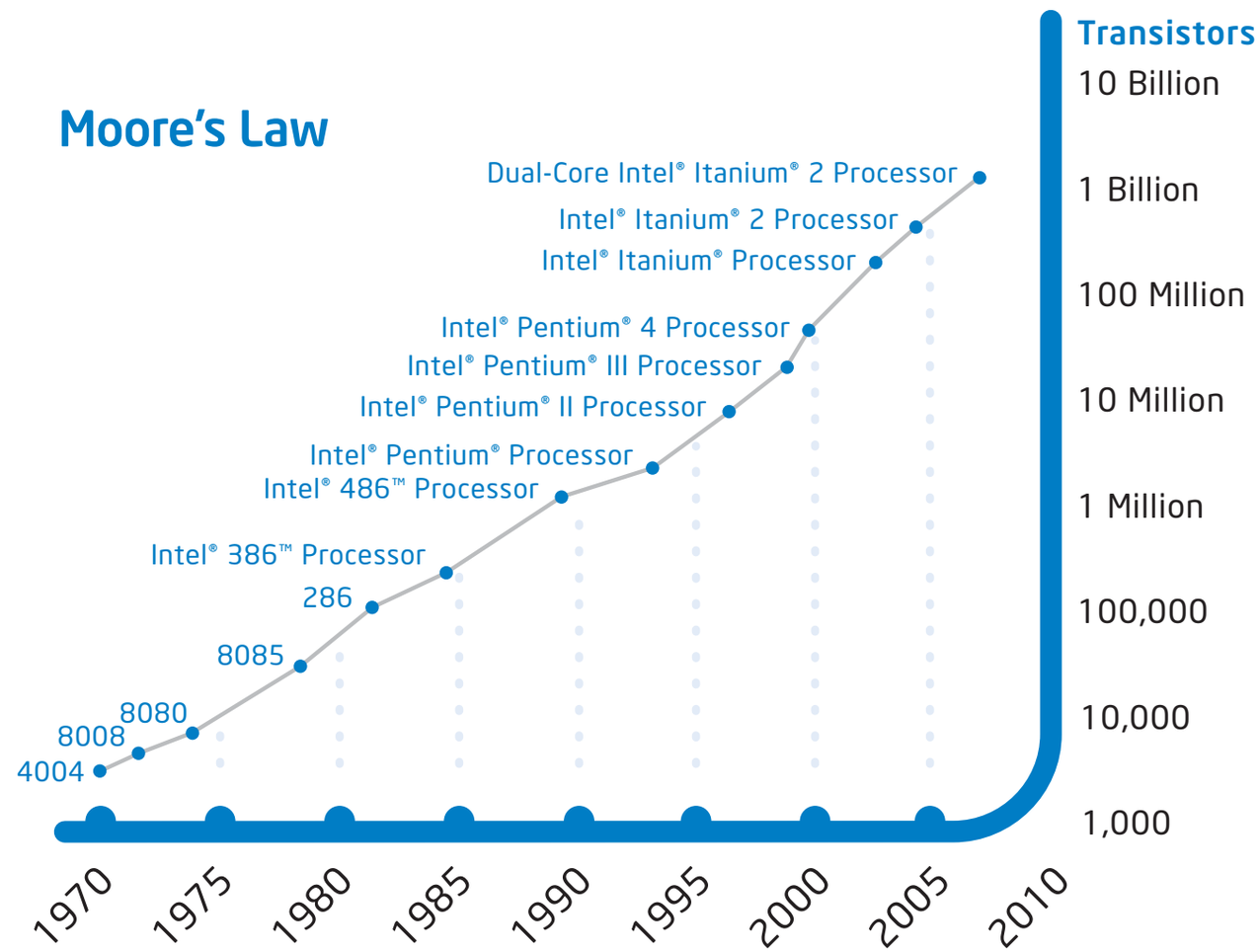
http://csperkins.org/

Talk Outline

- Trends in semiconductor devices
- Implications for networking
  - Network layer issues
  - Transport protocols
- Impact on applications

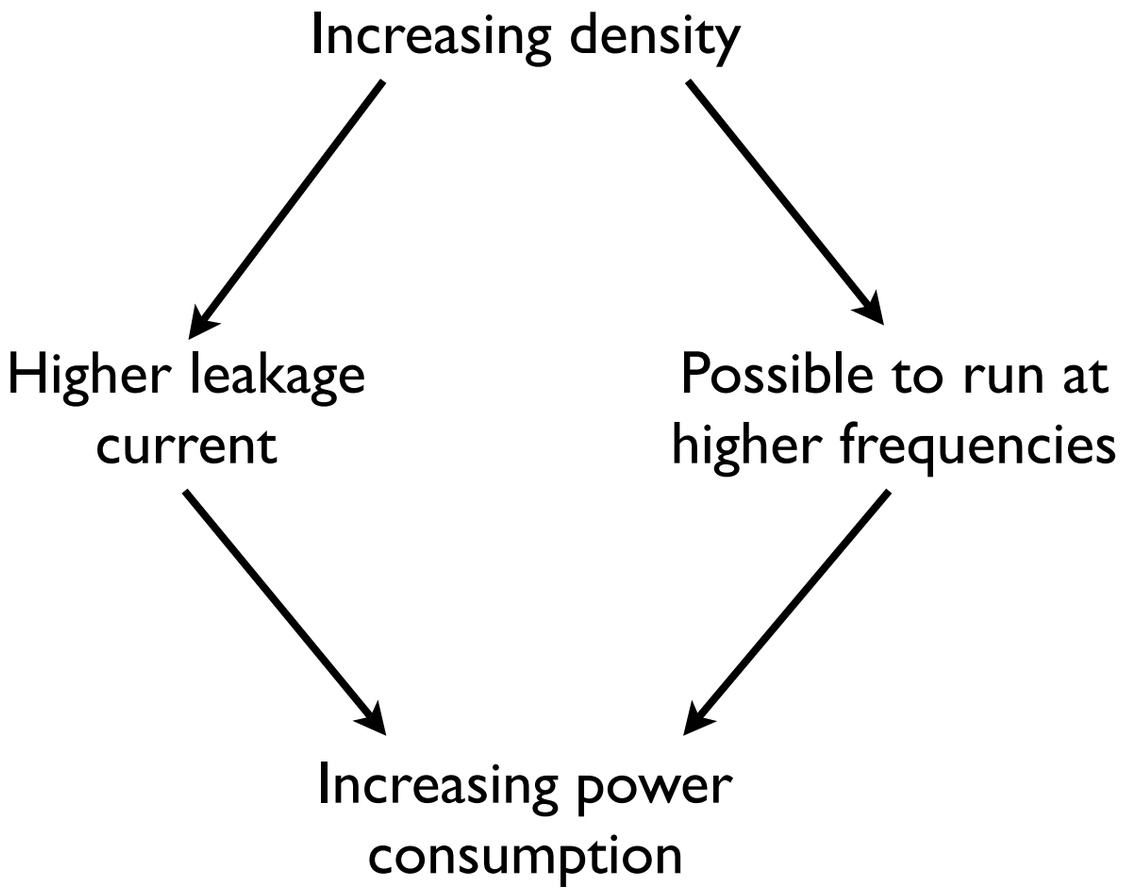
Trends in Semiconductor Devices

- Exponential growth in semiconductor density
- Corresponding growth in CPU performance
  - More logic → better algorithms
  - Bigger caches; more memory
- Router vendors benefit from the same curve
  - Larger routing tables fit in fast, on-chip memory

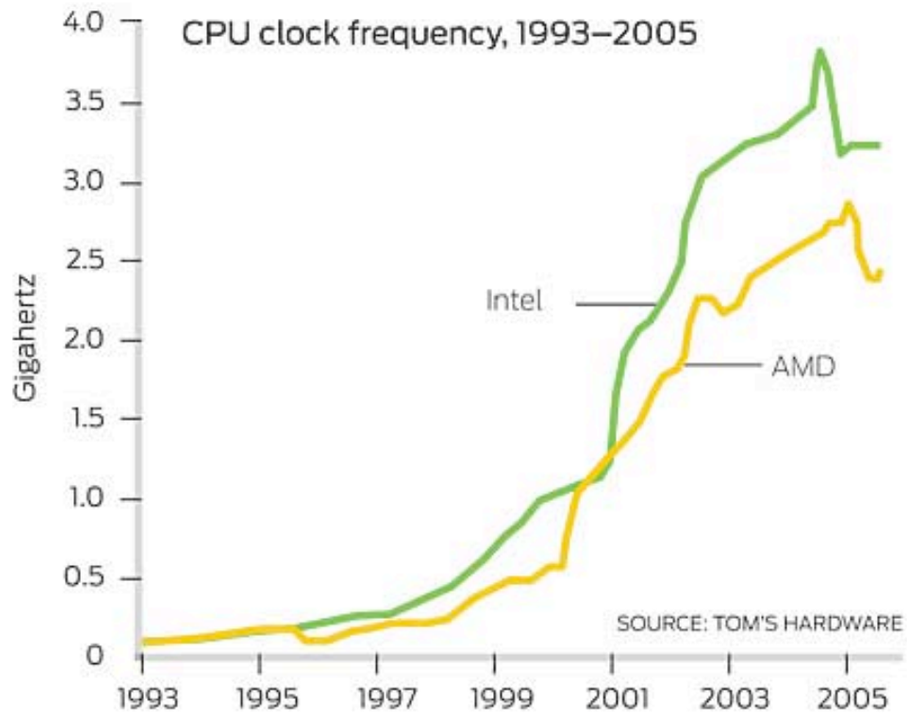


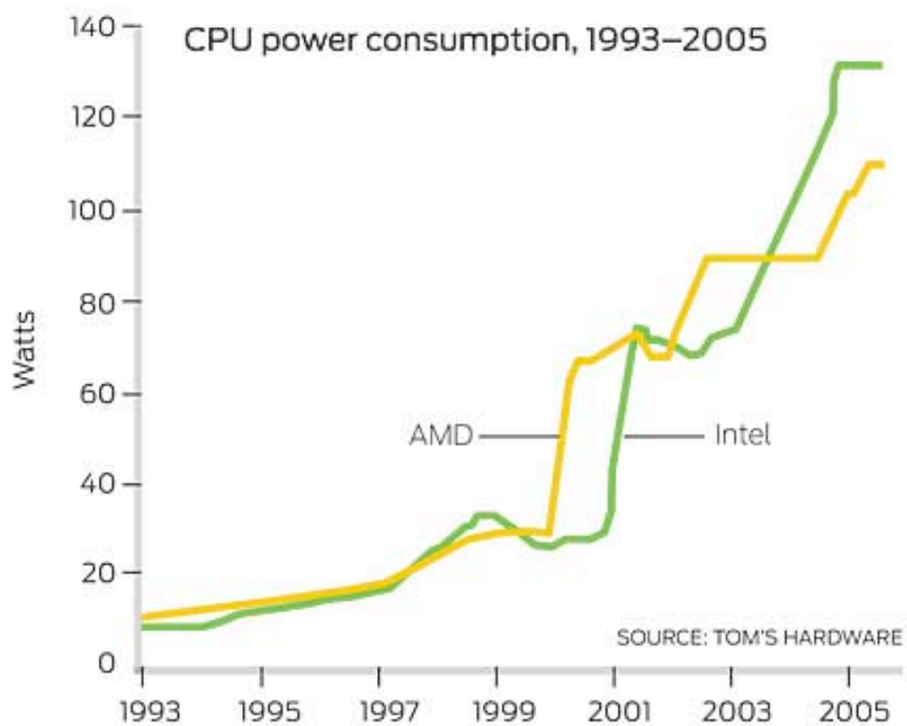
Source: Intel Tera-Scale Computing Research Overview

Clock Rate and Power Consumption



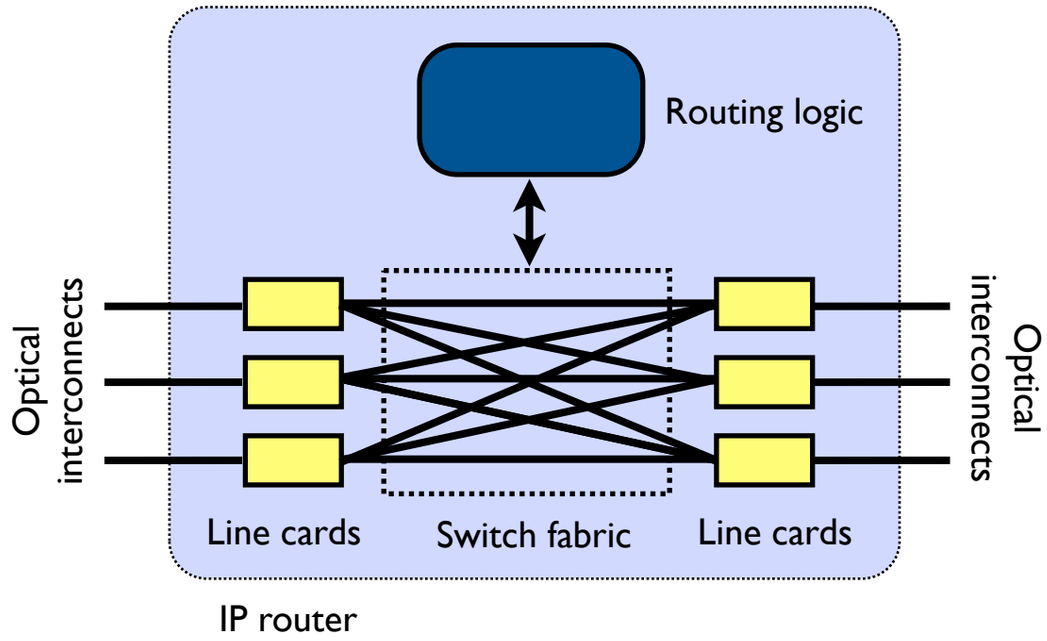
Semiconductor devices are no longer getting faster





So what? We're network researchers...

- Consider router architecture
  - Interconnect and switching fabric optical
  - Routing and forwarding logic uses semiconductor devices

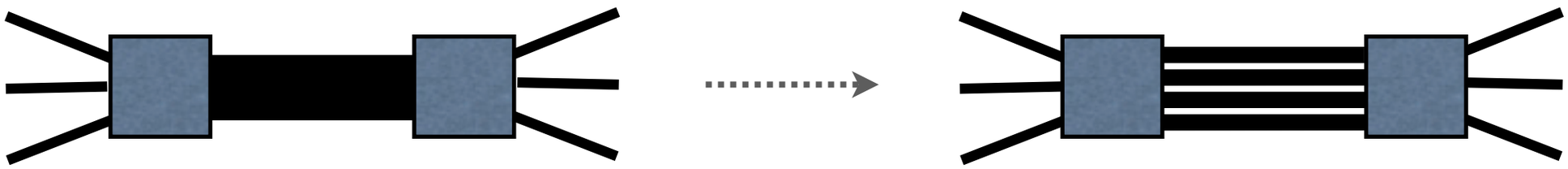


- Limited by power consumption:
  - Performance of routing logic
  - Core network links run at 40Gbps, whereas CPUs are power limited at 3-4GHz...
  - All-optical routers not (yet?) practical

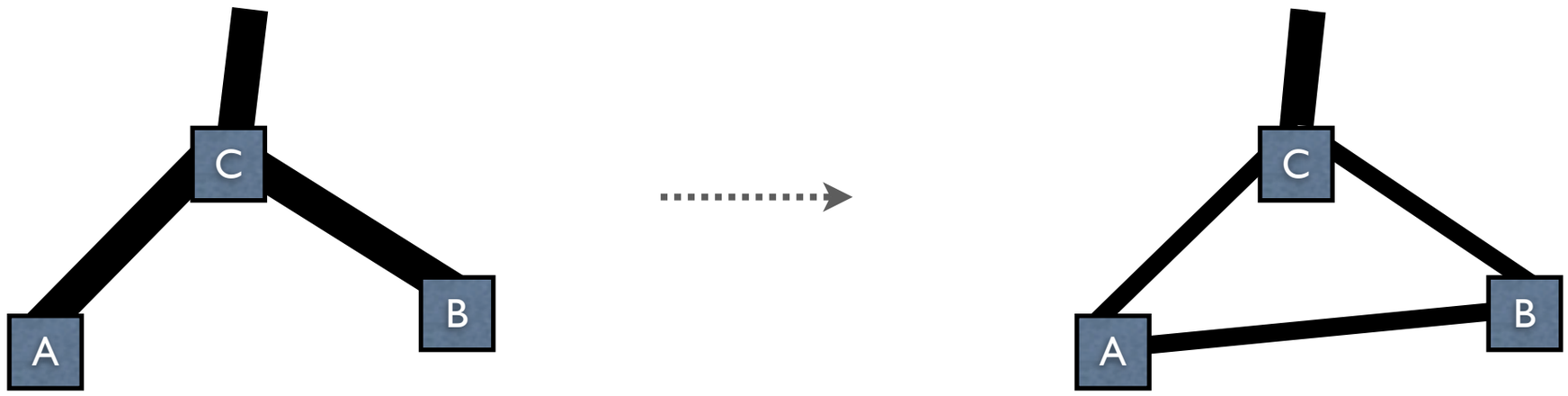
Wider, Not Faster

- CPU designers reducing clock speed and introducing parallelism (“multicore”)
- Same will *inevitably* happen in networking
  - Benefit due to reduced power consumption will outweigh complexity of using several slow links in place of a single fast link

Implication: Richer Network Connectivity

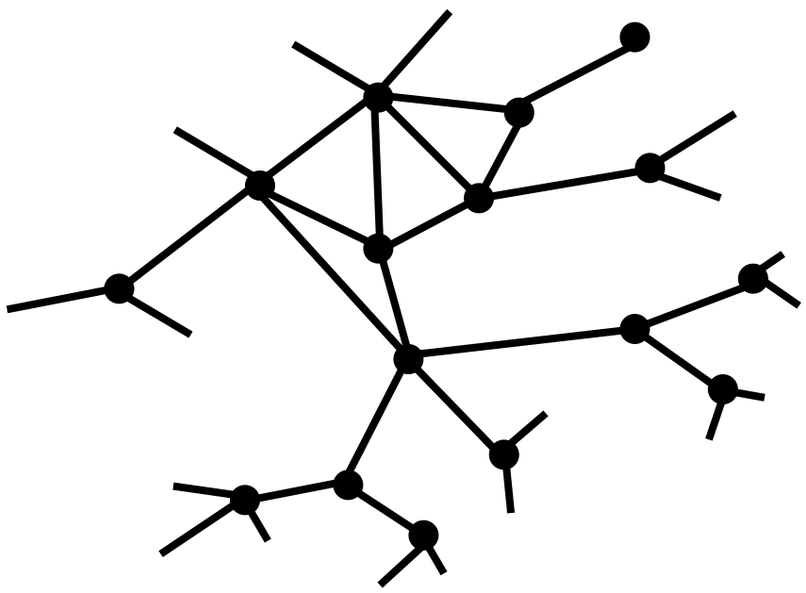


Easy to handle; only local impact on routing



Hard to handle; globally visible in topology

Problem: Routing Table Growth



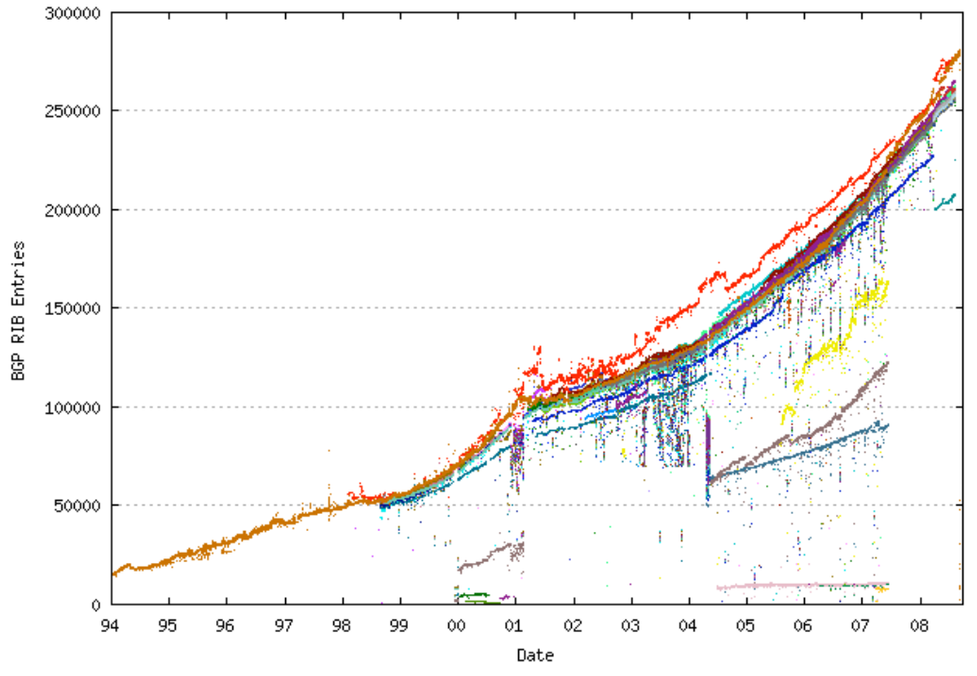
- Hierarchical aggregation of routing prefixes requires a mesh-like core, with tree-like edge networks
- Trend is making the edges more mesh-like → aggregation will fail, since hierarchy violated

→ Can we route without aggregation?

(Compact routing)

→ Can we aggregate better?

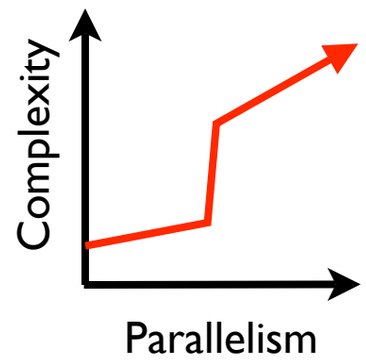
(Multiple prefixes per AS, aggregating per upstream connection; locator-identifier split)



Source: Geoff Huston/APNIC

Opportunity: Additional Forwarding Cycles?

- Introducing link parallelism will greatly increase routing complexity
  - Step-change in table size and churn
  - Re-think of the network architecture

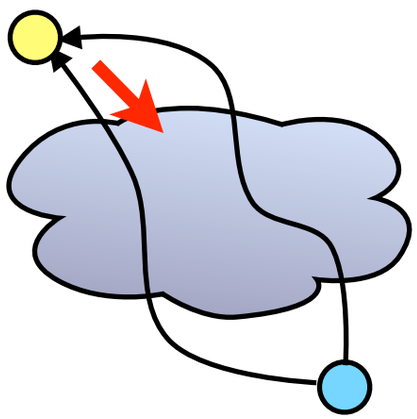


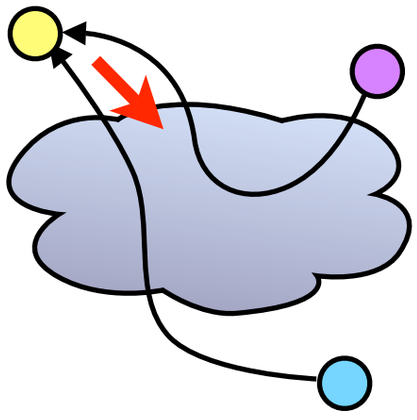
- Further increases less problematic
  - Network will be settled in new equilibrium
  - What is the trade-off between power to drive a link and to make forwarding decisions?
    - Can we introduce more parallelism than needed, while keeping line card processor performance, to give more cycles per packet while still saving power?
    - May allow very interesting network architectures...

Implications for Transport: Multi-path

- To make effective use of capacity, transport protocols must use multiple paths
  - How to see multiple paths?
    - Is multi-homing sufficient? Or is visibility into the core (source routing) needed?
  - Multi-path congestion control
    - Multi-path TCP
    - Layered and multi-description coding for real-time media → what transport? how to adapt? usability?
    - Effects on traffic engineering and load balancing
  - Likely solve mobility as a side-effect

Implications for Transport: Reliability





- Multi-path transports can build on reliable multicast protocols
  - Responses from multiple paths vs. responses from multiple participants; scaling techniques, avoiding implosion
  - Massive re-ordering at transport layer
  - Application level framing
  - De-coupling transport from path choice
- Managing delivery on a single path insufficient for reliability

Implications for Applications

- Concurrency is going to be an issue for the network, as well as within the end-system
  - APIs will evolve, expose multi-path behaviour
  - Berkeley Sockets API likely insufficient

- What will be the typical approach for writing networked applications?
  - Should be coding in Erlang rather than C/Java?

Conclusions

- Trends in semiconductor device behaviour → increased parallelism in network
- Expect this to cause a rethink in network and transport protocol design
- How to evolve network and applications?