# *A Network of Time-Division Multiplexed Wiring for FPGAs*

**Rosemary Francis**

Simon Moore

Robert Mullins

UNIVERSITY OF CAMBRIDGE

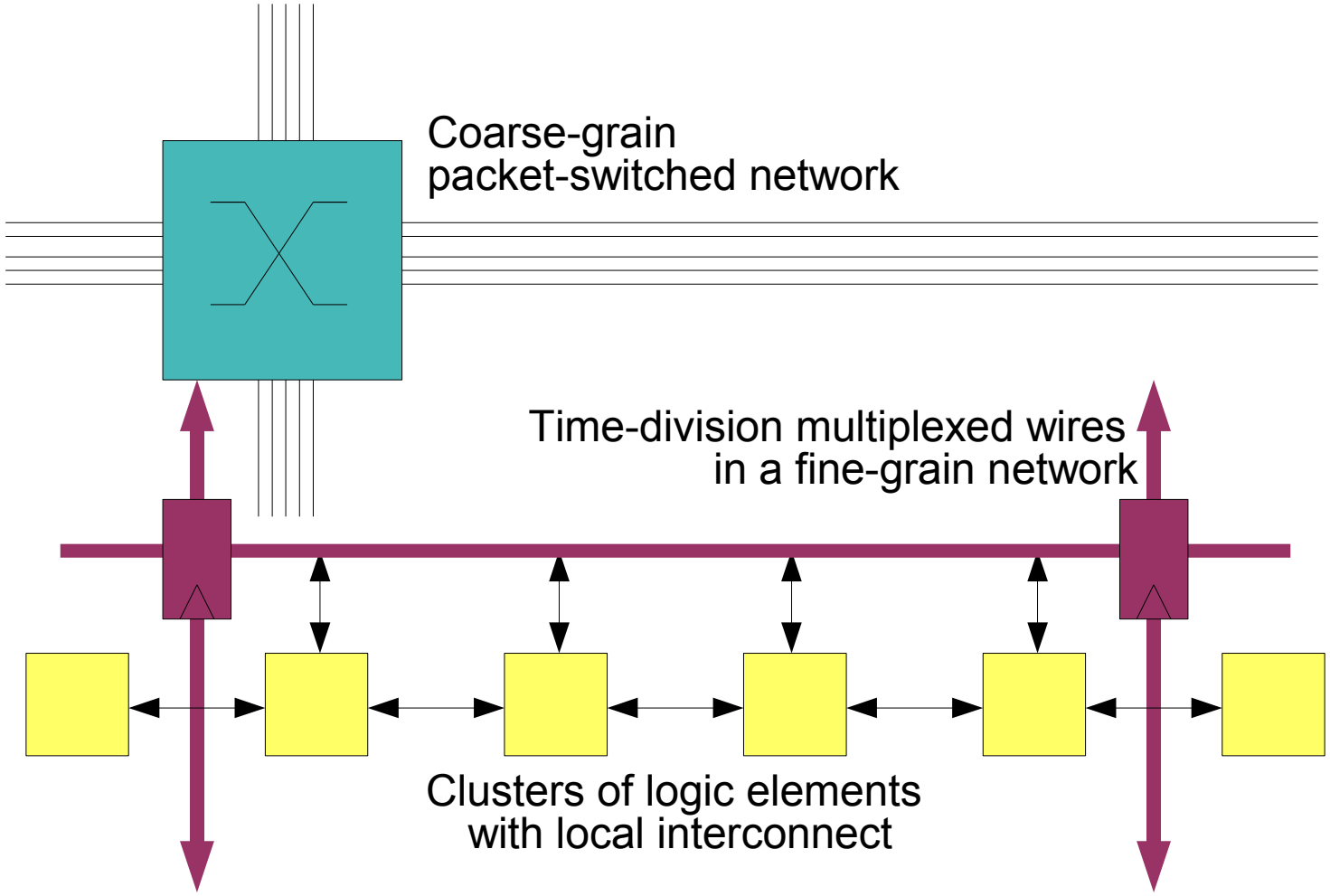
# Motivation

- FPGAs are now home to complex Systems on-Chip
- ...but still optimised for single-core designs
- FPGA global wiring is simple in comparison with ASIC Networks-on-Chip
- Networks for FPGAs use lots of logic
- Hard blocks are limited by the soft IP blocks

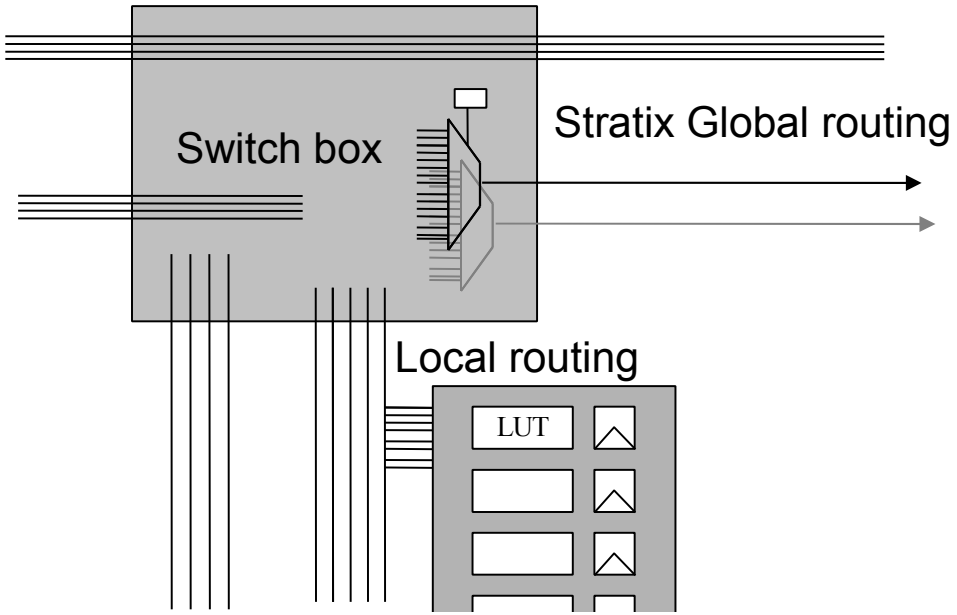
# Goals

- Use TDM components for effective soft NoC implementation
- Funnel data to high-speed hard blocks
  - Hard NoC
  - Multipliers
  - Block RAM
- Determine optimum TDM architecture
  - What are the costs?
  - Is it possible to design for global and local routing?

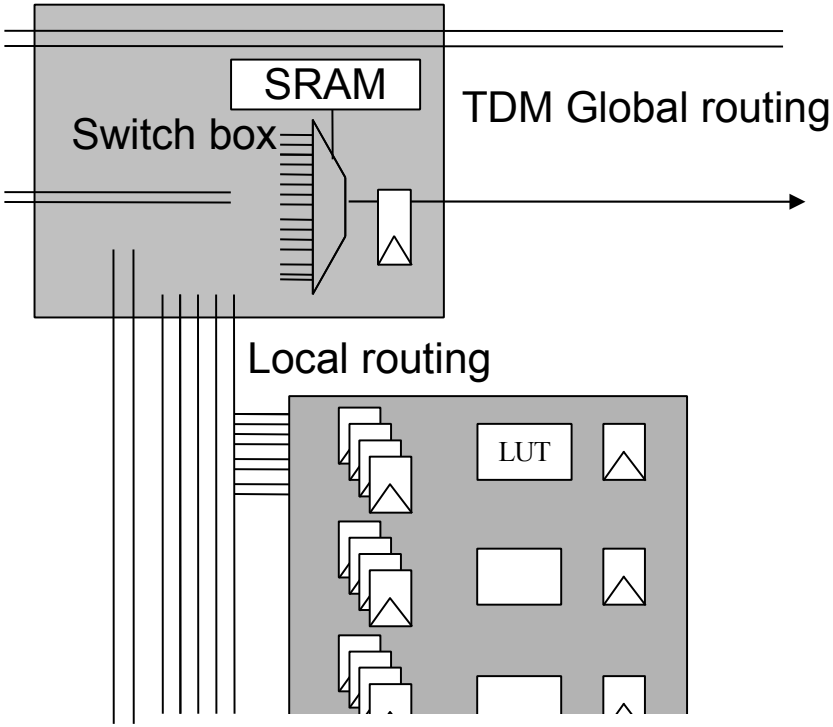
# Hierarchy of interconnect



# Architecture: Stratix vs TDM

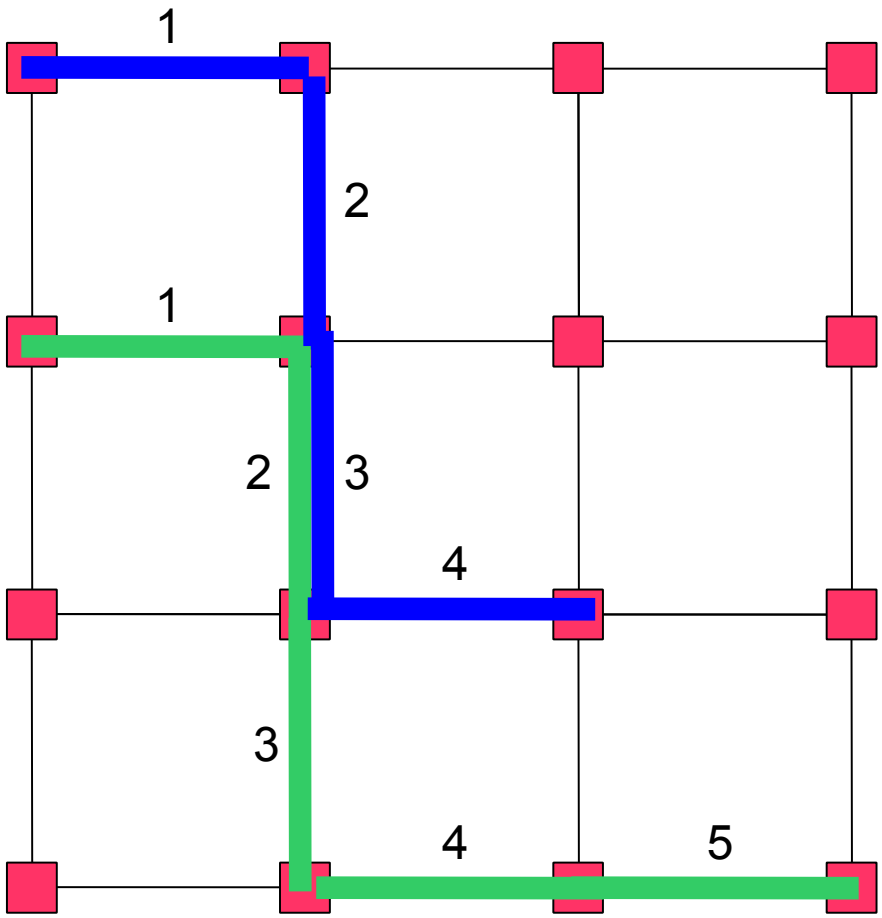



Cluster of logic elements



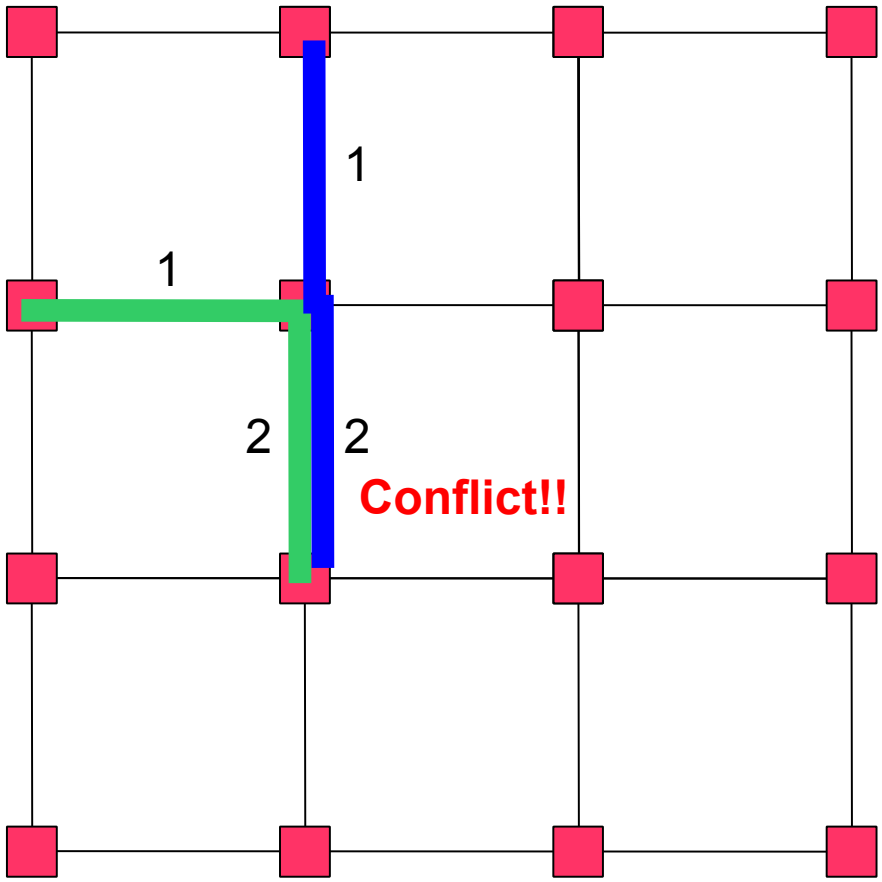

Cluster of logic elements
with latched inputs

# Wire Sharing



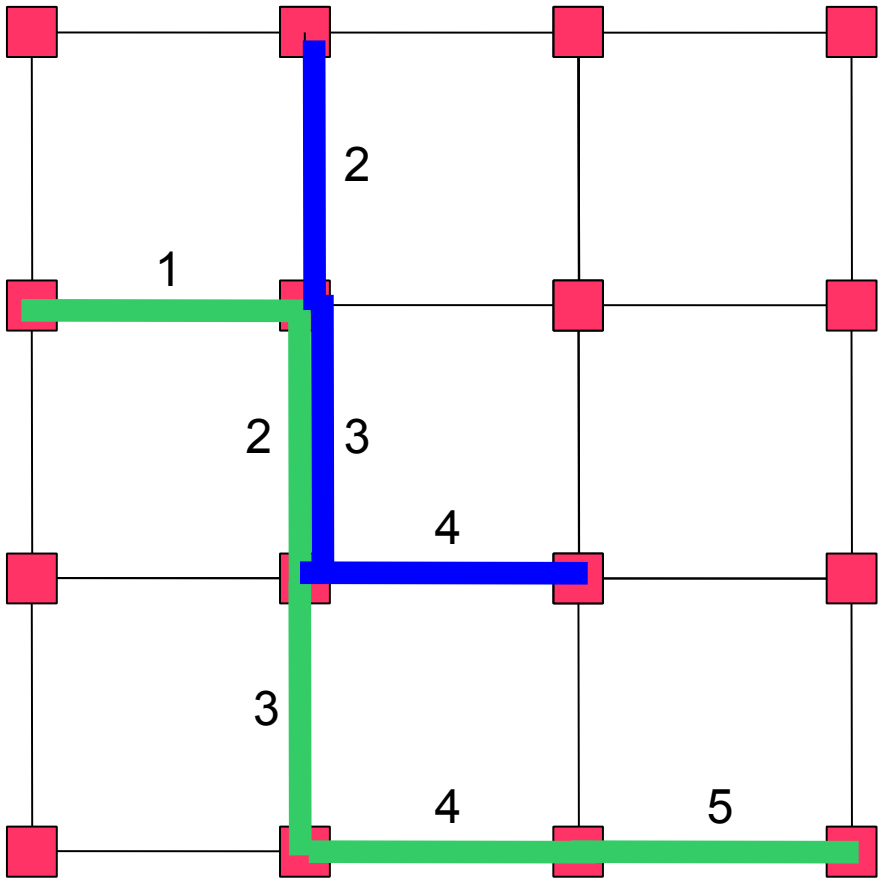

- Many wires can be shared without a problem

# Wire Sharing




- Many wires can be shared without a problem
- Other configurations require a more intelligent approach

# Wire Sharing




- Many wires can be shared without a problem
- Other configurations require a more intelligent approach
- Signals can be delayed to allow more efficient wire use without rerouting

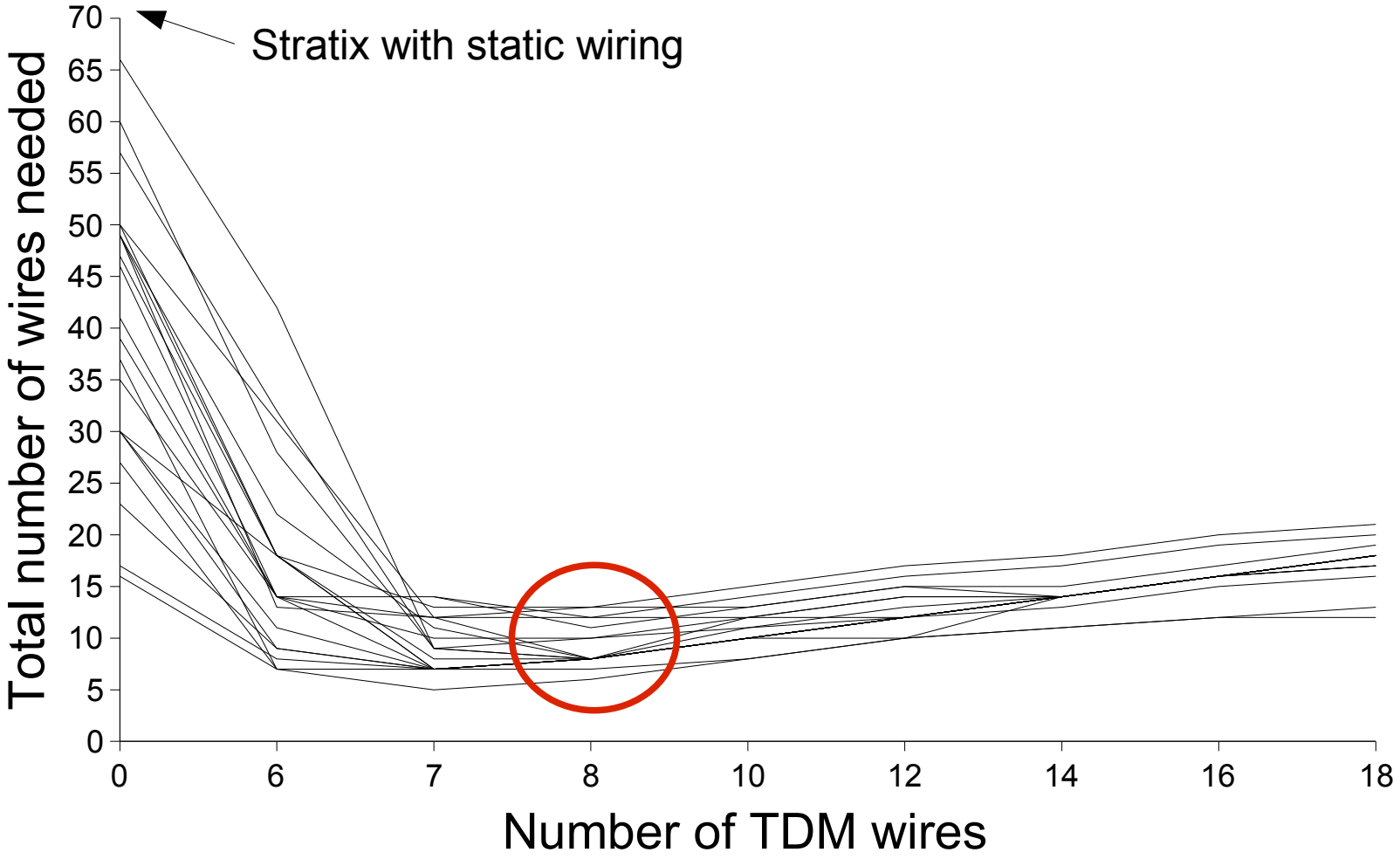
# Our Scheduler

- Our scheduler
  - maps benchmarks from a Stratix FPGA to a TDM FPGA
  - resolved TDM conflicts after place and route
- Benchmarks
  - IP cores taken from the Altera University Suite
- Aim
  - To reduce the amount of wiring as far as possible using TDM wiring with realistic characteristics

# Parameter selection (1 of 3)

- Assume infinite time slots to reduce wiring
  - Determine minimum number of TDM wires

# Infinite number of time slots

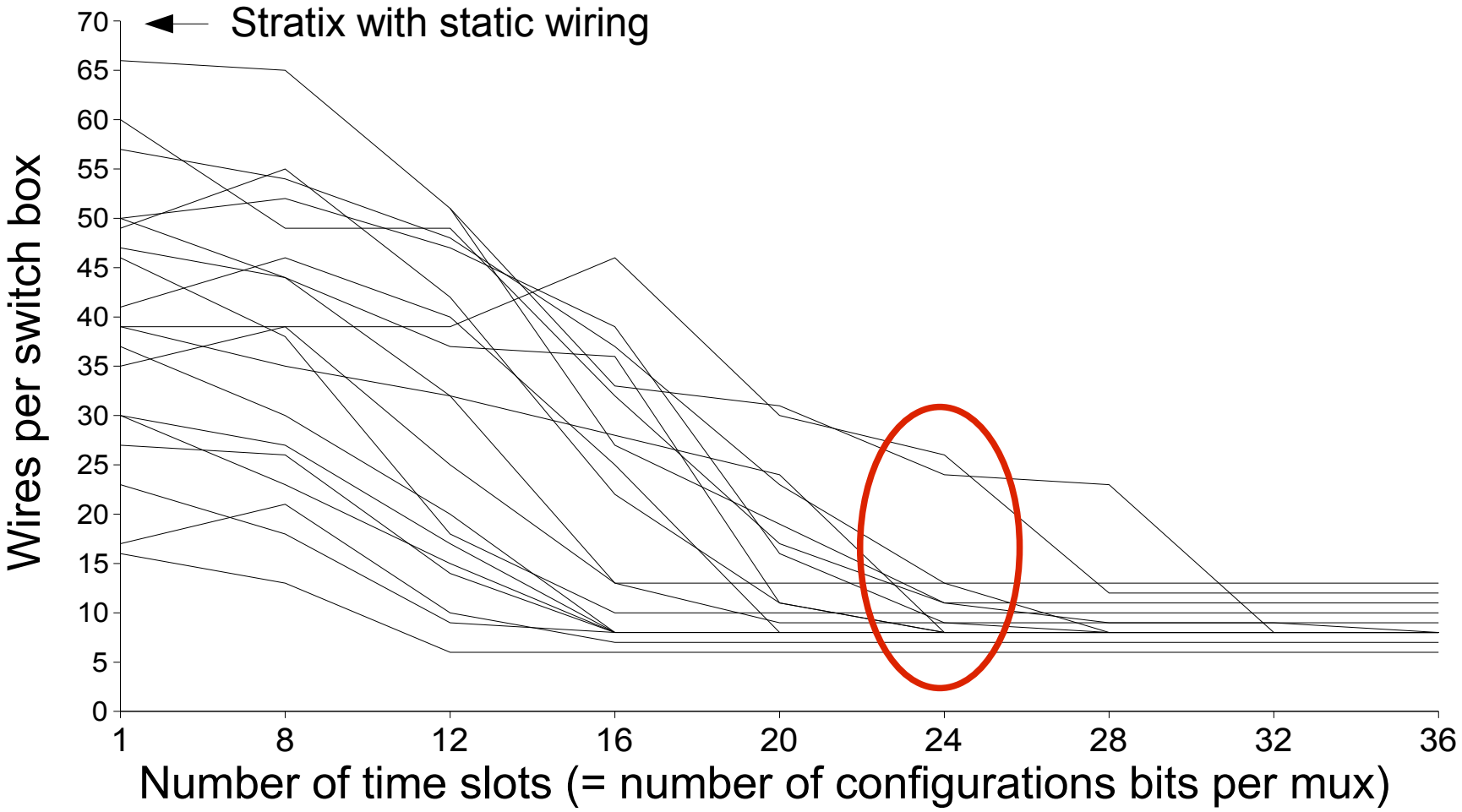



UNIVERSITY OF CAMBRIDGE

# Parameter selection (2 of 3)

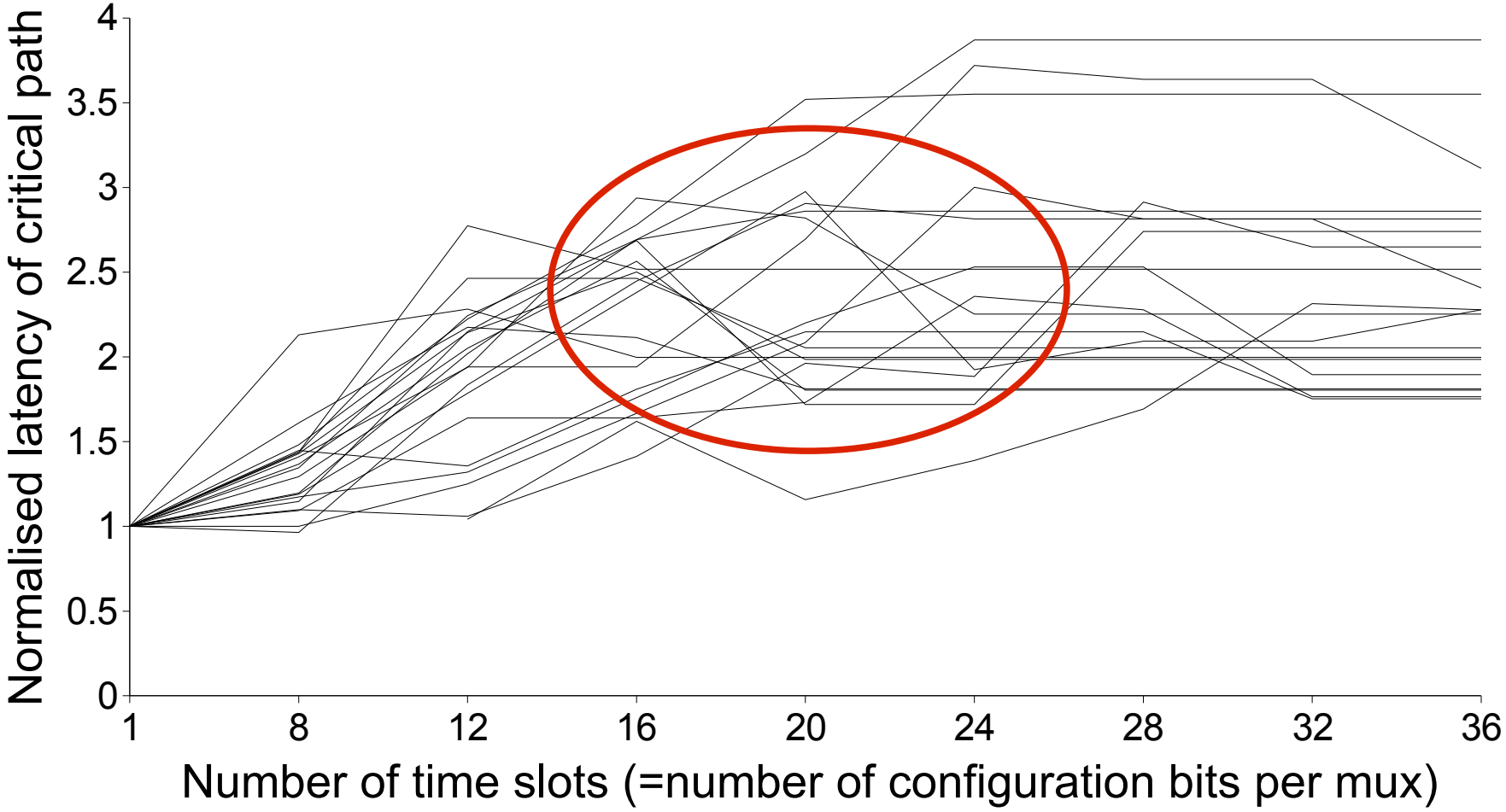
- Assume infinite time slots to reduce wiring
  - Determine minimum number of TDM wires
- Vary number of time slots
  - Determine optimum number of time slots
  - Investigate the effect this has on latency

# Determine number of time slots




UNIVERSITY OF CAMBRIDGE

# Number of time slots vs latency

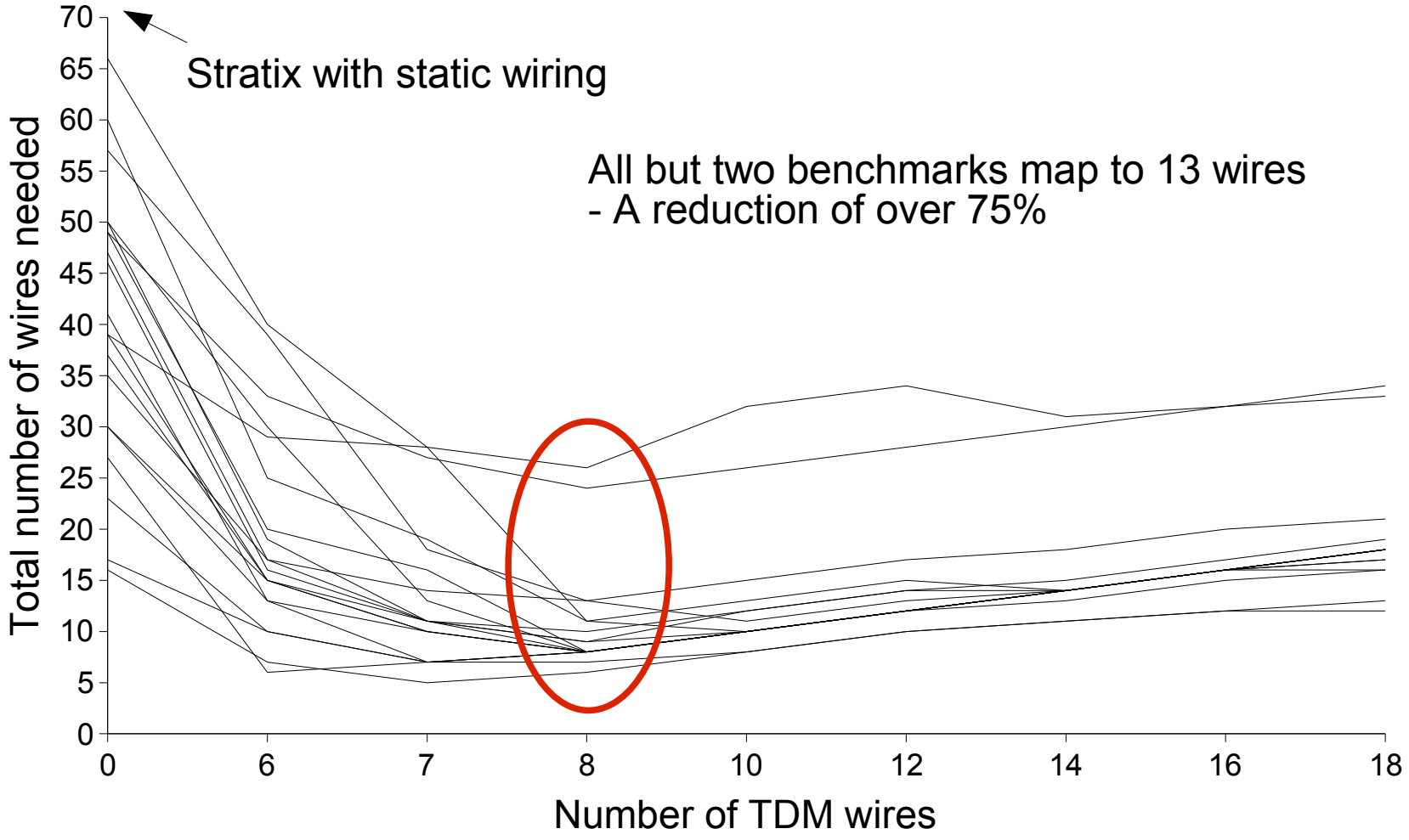



UNIVERSITY OF CAMBRIDGE

# Parameter selection (3 of 3)

- Assume infinite time slots to reduce wiring
  - Determine minimum number of TDM wires
- Vary number of time slots
  - Determine optimum number of time slots
  - Investigate the effect this has on latency
- Using optimum number of time slots
  - Re-evaluate optimum number of TDM wires

# Limited resources




UNIVERSITY OF CAMBRIDGE

# Architectural drawbacks

- Extra configuration SRAM
- High-speed interconnect clock
- Benchmarks run over three times slower
- New CAD tools needed
  - Re-routing in space as well as time
  - Optimise for TDM wiring at every stage

# Conclusions

- Using TDM wiring we can reduce the number of wires whilst increasing the data rate within channels
  - 75% less wiring * 24 time slots * 3 times slower means 2 times channel data rate
- This will allow
  - the design of effective global interconnect
  - more efficient sharing of on-chip resources
  - simplification of multi-chip designs

# Future Work

- Current scheduling algorithm gives
  - Large wire reduction, large latency penalty
- We are investigating a better compromise
  - Small wiring reduction, small latency penalties?
  - Recent new results show this is possible
- Area and power
  - Is the wiring reduction enough to justify the extra area and power costs?

# Thanks for listening...

Rosemary.Francis@cl.cam.ac.uk

UNIVERSITY OF CAMBRIDGE