A Reconfigurable Bi-Directional Optical Data Interconnect

G. A. Russell, K. Symington and J. F. Snowdon

Department of Physics, Heriot-Watt University, Edinburgh

http://www.optical-computing.co.uk

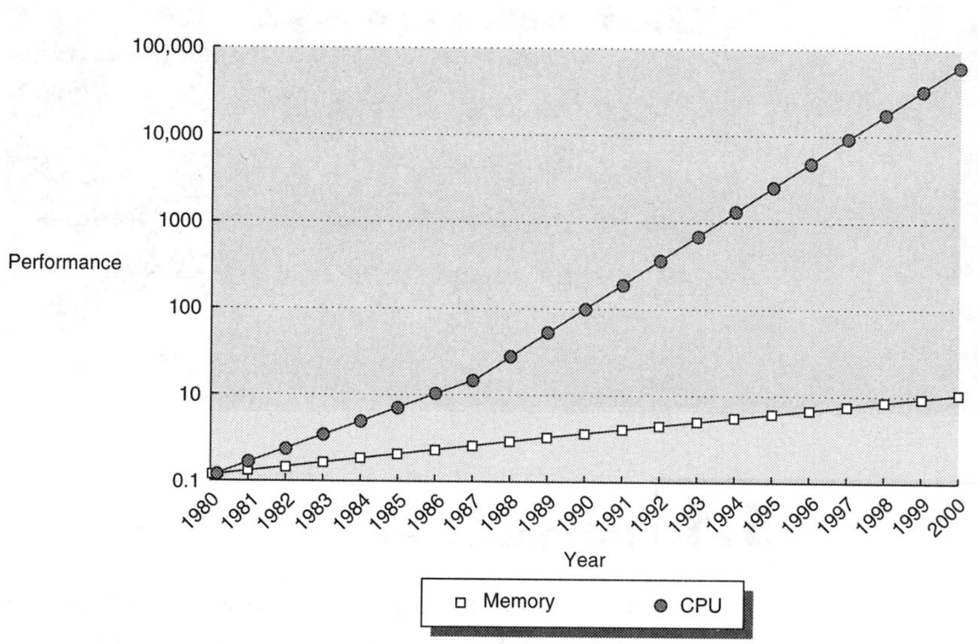
The Problem with Computers

•Processor performance increasing faster than memory bus speed

•Processor performance limited by data bottleneck

•Large computational clusters require highly connected topologies with long distance hops.

•Simulation and analysis shows optical interconnect must be embedded deeply in the system or electronic bottlenecks dominate.



HERIOT WATT UNIVERSITY *Department of Physics*

The Solution - The Optical Highway

Why Optics?

•Inherent parallelism,

•Higher temporal and spatial bandwidths,

•Higher interconnection densities,

•Less signal cross-talk,

•Immunity from electromagnetic interference and ground loops,

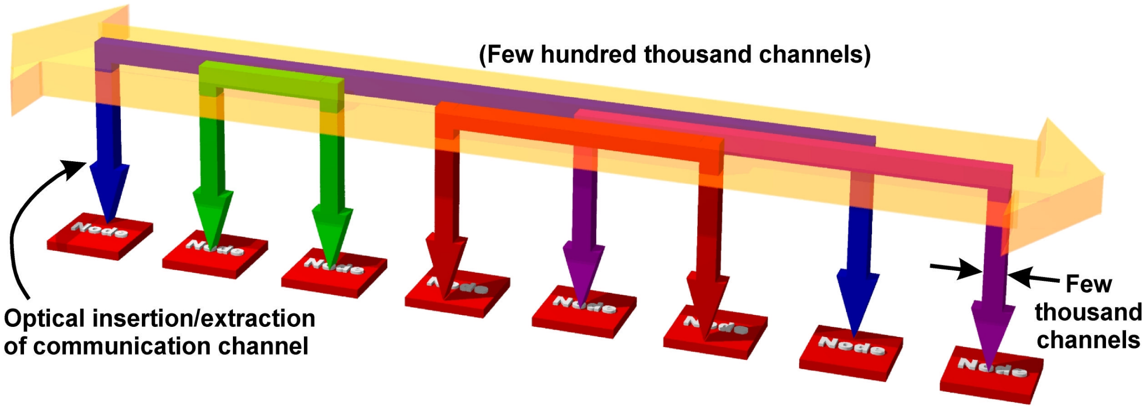
•Freedom from quasi-planar constraints,

•Lower signal and clock skew,

•Lower power dissipation,

•Larger number of fan-ins fan-outs,

•Potential for reconfigurable interconnects.



'Optical Highway' carrying hundreds of thousands of channels linking non-local nodes with thousands of channels

Technologies - Optical Components

Polarisation Beam Splitters (PBS)

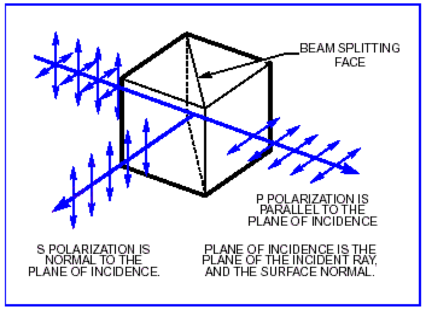
Used to route data streams into and out of Optical Highway.

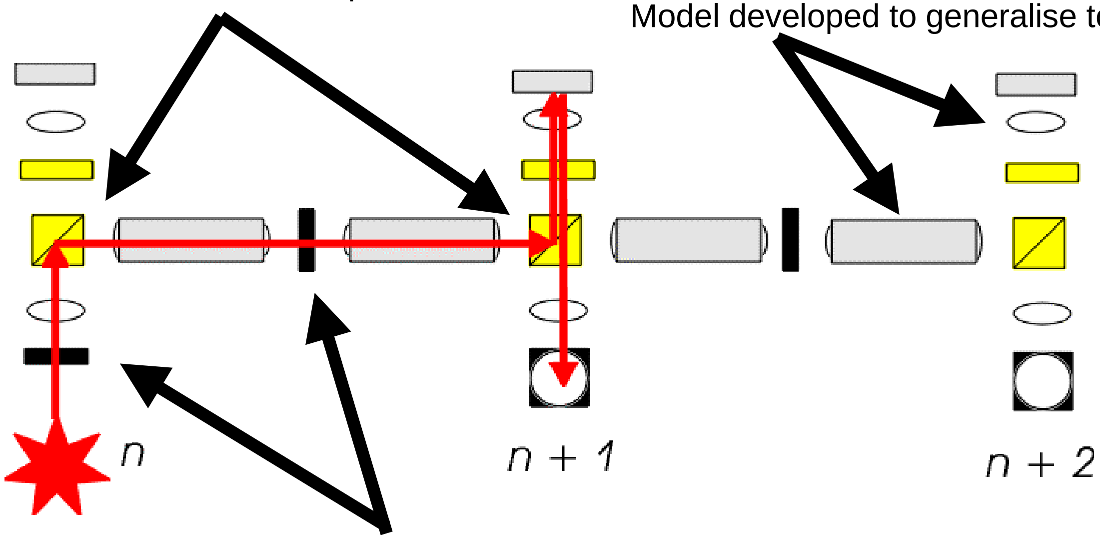
Patterned Birefringent Plates are used to selectively flip polarisations of the beams to be routed out of the Optical Highway.

Lens

Combination of bulk and micro lenses can be used.

Simple modelling and ray-tracing used to insure aberration limited spot size small enough to eliminate cross-talk between channels. Model developed to generalise to multiple relay stages.





$$A_L = \frac{kf}{(f/\#)^2}$$

$$A_T = \frac{A_L \phi}{(f - A_L)}$$

Liquid Crystal Plates

Liquid Crystal Plates can be used as reconfigurable Birefingent elements. Although too slow for packet switching they are fast enough to reconfigure the topology for fault tolerance and algorithmic reasons.

LC plates used in experiment removed from Steroscopic glasses designed for 3D computer gaming.

HERIOT WATT UNIVERSITY Department of Physics

Modelling the Optical Highway

Optics Modelling

There are a number of limits on the possible interconnectivity of an optical highway due to the optics.

- optical power
- aberration
- device densities

For the purposes of this modelling these limits are assumed to be completely independent and that the minimum of the three is the final limit.

Model interfaces with electronic modelling of PC buses and algorithmic model of real problems

$$p_{power} = \frac{2}{\ln(\xi)} \cdot \ln\left(\frac{NEP\sqrt{\frac{8 \cdot B_{optical}}{C_w}}}{P_{VCSEL}} \right) + 1$$

$$p_{ab} = \frac{\pi\phi^2}{h_{\max}\left(A_T^2 + \frac{\lambda^2}{16} \right) + s_{VCSEL}^2}$$

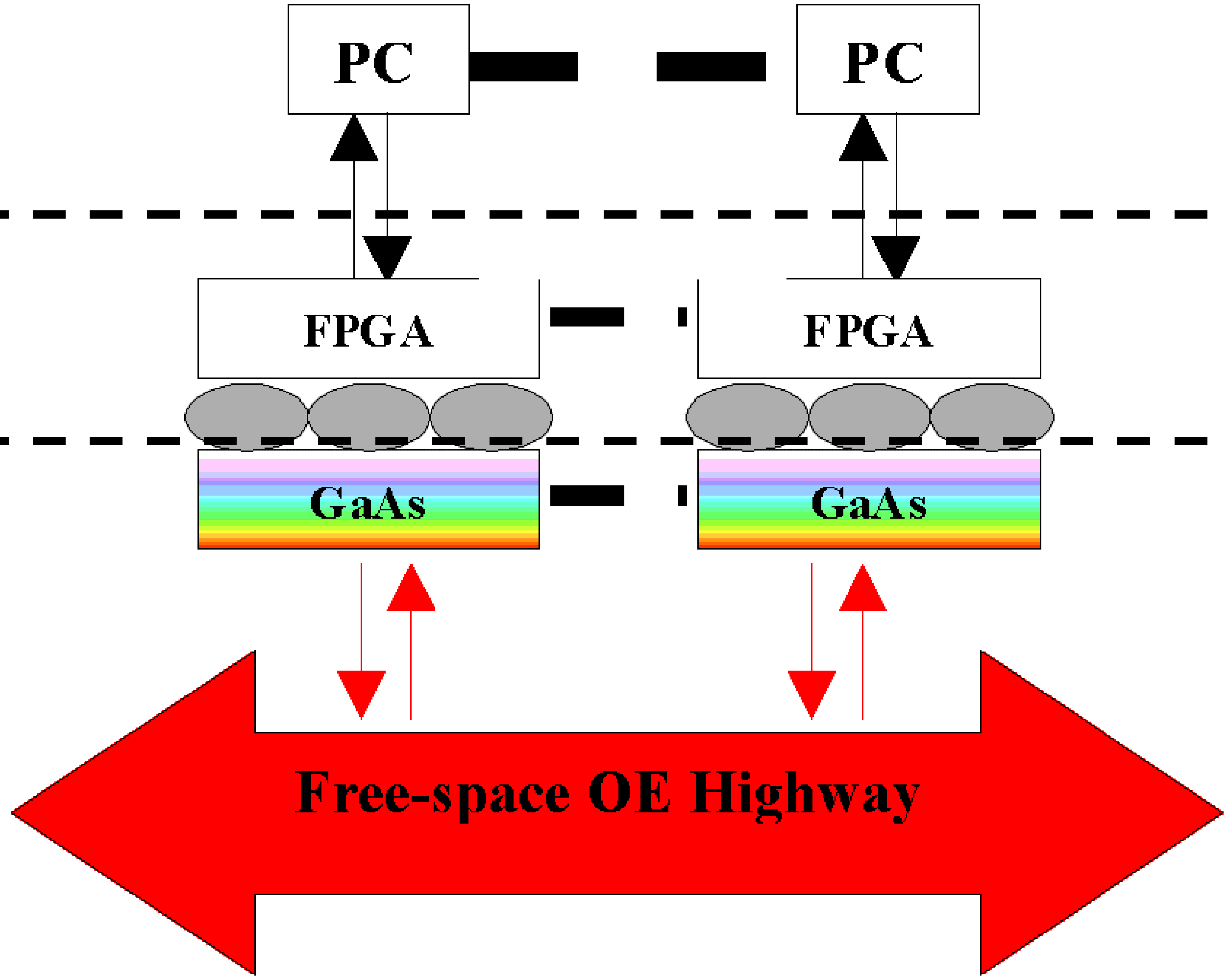
$$p_{device} = \frac{N_d}{2 \cdot C_w}$$

$$p_{max}$$

HERIOT WATT UNIVERSITY Department of Physics

Three Level Model

- **Algorithmic (Leeds)**
  - Sorting
  - Searching
- **PC Bus and FPGA (HW)**
  - PCI
  - New Buses
- **Optical Highway (HW)**
  - Power
  - Aberration
  - Device Density
  - Topology



Extract from Parameter Table

Simulation Parameters

- Large parameter space
- One page of seven shown
- >60 parameters
- Small number of interface parameters between layers of model
- More added as model refined
- Values from past and current experimental demonstrators and commercial sources used in simulations.

1 Variable Definitions

A_l	Longitudinal spherical aberration or L.SA. Measured in meters (m).
A_t	Transverse spherical aberration or T.SA. Measured in meters (m).
B'	Optical node-to-node bandwidth in one direction. Measured in bytes per second (B/s).
B_{act}	Bandwidth transfer requested by algorithm through the I/O bus. This must be equal to or less than the available bandwidth. Measured in bytes per second (B/s).
B_{avail}	Best practically available bandwidth on I/O bus given a defined efficiency. Measured in bytes per second (B/s).
B_{IOB}	Maximum theoretical bandwidth transfer available on the I/O bus. Measured in bytes per second (B/s).
B_{mem}	Maximum bandwidth transfer available to processing system main memory. Measured in bytes per second (B/s).
$B_{optical}$	Bandwidth of data through the optical highway, node to node, down a single physical channel. It is essentially the emitter/detector raw transmission rate. Encoding must be included in this figure. Measured in bytes per second (B/s).
B_{per}	Maximum bandwidth transfer used on I/O bus by peripheral devices. This is dependent on architecture; for example some architectures use the I/O bus to access the hard-disk. This must be equal to or less than the available bandwidth. Measured in bytes per second (B/s).
B_{SPA}	Bandwidth of smart pixel array cache memory. Measured in bytes per second (B/s).
C_v	Number of virtual interconnection channels. Virtual channels are considered to carry data transparently in both directions. This is a dimensionless number.
C_p	Number of physical interconnection channels. This number represents the number of physical links that need to be established and thus there must be

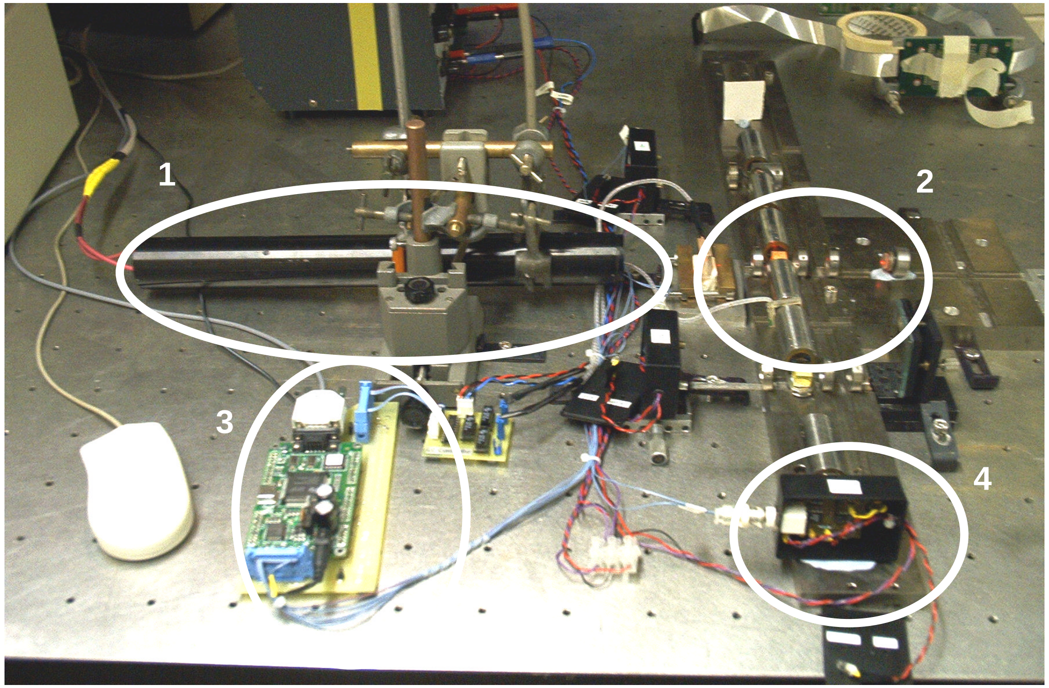
HERIOT WATT UNIVERSITY Department of Physics

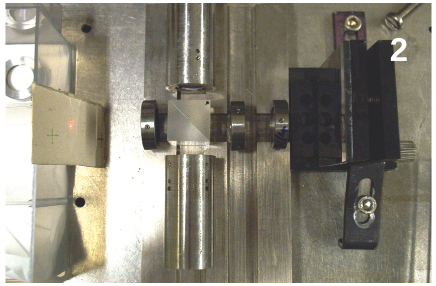
Experimental Realisation

1. Laser: HeNe for alignment, array of individual diodes for testing. Final system would use VCSEL or modulator array

2. PBS and Quarter-Wave Plates used to control the beam path by switching polarisation of beams.

Liquid Crystal Plates used to alter topology





3. DSP (Digital Signal Processor) based control board. All switching and detector control goes through DSP. Extra channel reserved for laser modulation

4. Photodetectors interfaced to PC via DSP board. Biased to trigger at correct levels.

WebCams used to measure beam quality and other optical parameters.

HERIOT WATT UNIVERSITY Department of Physics

Conclusions

•Optical interconnect can provide high bandwidth, highly scaleable communications for massively parallel computers.

•Latency penalties from optics to electronics domain changes are of set by smaller routing costs of the more direct connections available optically.

•High bandwidth required for communications bounded problems and to service the data requirements of faster processing.

•Commodity components are becoming available at the required specifications.

Research Areas and Further Work

•Large, reliable arrays of optoelectronic components (>128x128).

•Integration with silicon technologies (bonding techniques or silicon devices).

•Efficient thermal engineering.

•Simulation of real traffic patterns on interconnect.

•Topology of interconnect and components.

Acknowledgements

- **EPSRC (OSI) funded AMOS project**
- **Department of Physics, Heriot-Watt University.**
  - A.C. Walker, T. Lim, B. Layet, J.J. Casswell
- **School of Informatics, Leeds University**
  - P. Dew, I. Gourlay

References

•J.L. Hennessy and D.A. Patterson, *Computer Architecture: A Quantitative Approach*, Palo Alto, California, Morgan Kaufmann Publishers, 1996.

•I. Gourlay, P. Dew, J. F. Snowdon, G. Russell , T. Lim, B. Layet, *Supporting Highly Parallel Computing With A High Bandwidth Optical Interconnect*, AMOS project paper, Submitted Journal of Parallel and Distributed Computing, June 2001.

•G.A. Russell, T.Lim, J.F. Snowdon. Modelling and Simulation of Cache Access to Utilise a High Bandwidth Optical Processor-Memory Bus, Proc. 2002 International Topical Meeting on Optics in Computing, Taiwan, April 2002, Pages 198-200.

•http://www.optical-computing.co.uk