Modelling of Optical Interconnects for Parallel Processing

G.Russell, J.F. Snowdon, T.Lim

(Department of Physics, Heriot-Watt University)

P.M. Dew, I.Gourlay

(Informatics Research Institute, University of Leeds)

Heriot-Watt University
Department of Physics

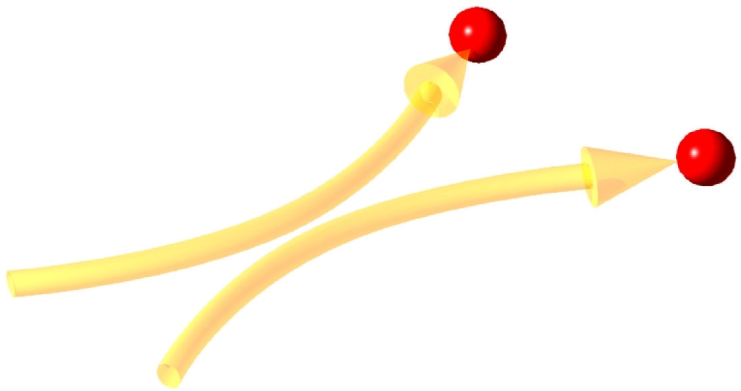
Structure of the Presentation

- Why Optics?
- History
- BSP modelling
- Examples
- Conclusions
- Further Work

Heriot-Watt University
Department of Physics

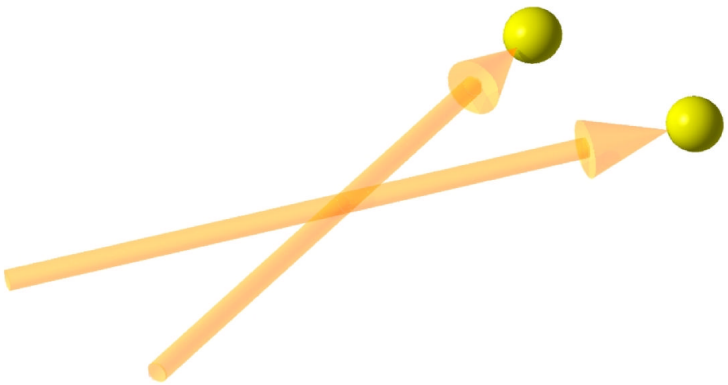
Why Optics?

Electrons



Since electrons carry mass and charge they interact strongly (Coulomb Interaction). Ideally suited for switching.

Photons



Photons do not carry mass or charge and are non-interacting in free space. They are ideally suited to interconnection.

Why Optics?

- As the length or density of electrical interconnection increases they suffer from increased wire resistance, residual wire capacitance, fringing fields and inter-wire cross-talk.
- Optical interconnection does not suffer from any of these problems due to its non-interacting nature: however this makes optics inefficient when trying to implement any type of gate.
- Optoelectronics attempts to make the best of both worlds by using electronics for switching and optics for communication.

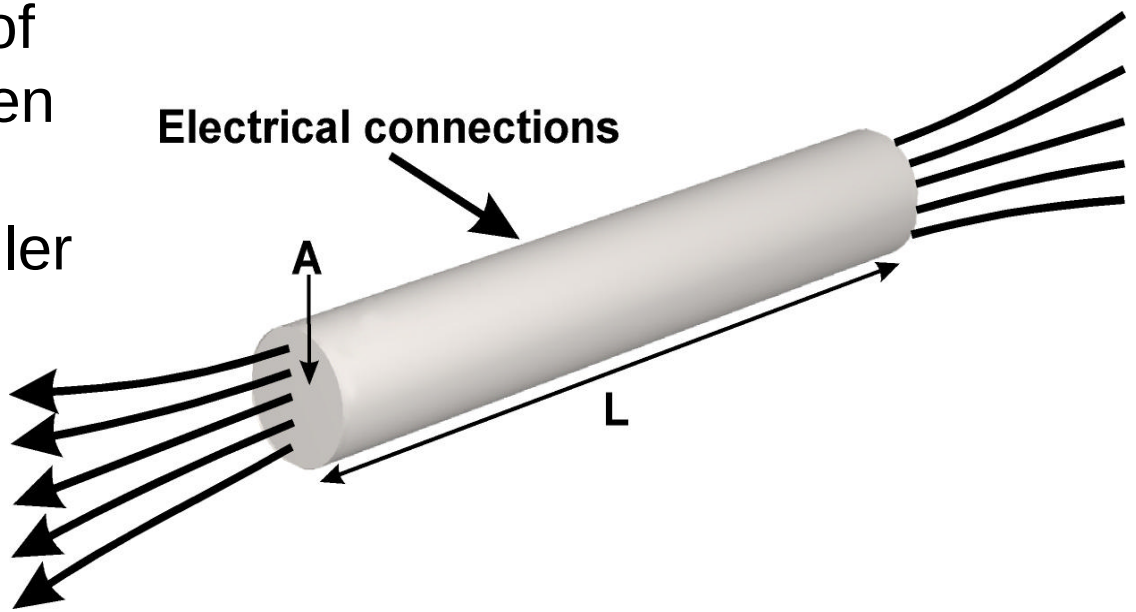
Heriot-Watt University
Department of Physics

Bandwidth Limitations

The maximum bandwidth of electronic systems has been estimated by Burton Smith (Tera Corp.) and David Miller (Bell Labs) as:

$$B_{max} = 500.\frac{A}{L^2}\ THz$$

Aspect ratio



Consider chip connections for a 10x10mm chip (100mm^2):

- Edge connections: 400 with ~100µm diameter lines A=3mm^2
- 2D Solder Bump array: 2,000 with ~15µm diameter pads A=3mm^2

Thus for a 10cm electrical connection across a board

B_{max} ~150GHz

SIA Roadmap

I/O figures for high performance ASIC systems taken from Semiconductor Industry Association.

Year	1999	2002	2005	2008	2011	2014
Process size (nm)	180	130	100	70	50	35
Chip size (mm^2)	450	509	622	713	817	937
On-chip clock (GHz)	1.2	1.6	2.0	2.5	3.0	3.6
I/O Bus speed (MHz)†	480	885	1035	1285	1540	1800
I/O Pads††	368	464	584	736	927	1167
Off-chip data rate (Gbs^{-1})	**177**	**410**	**604**	**946**	**1428**	**2100**

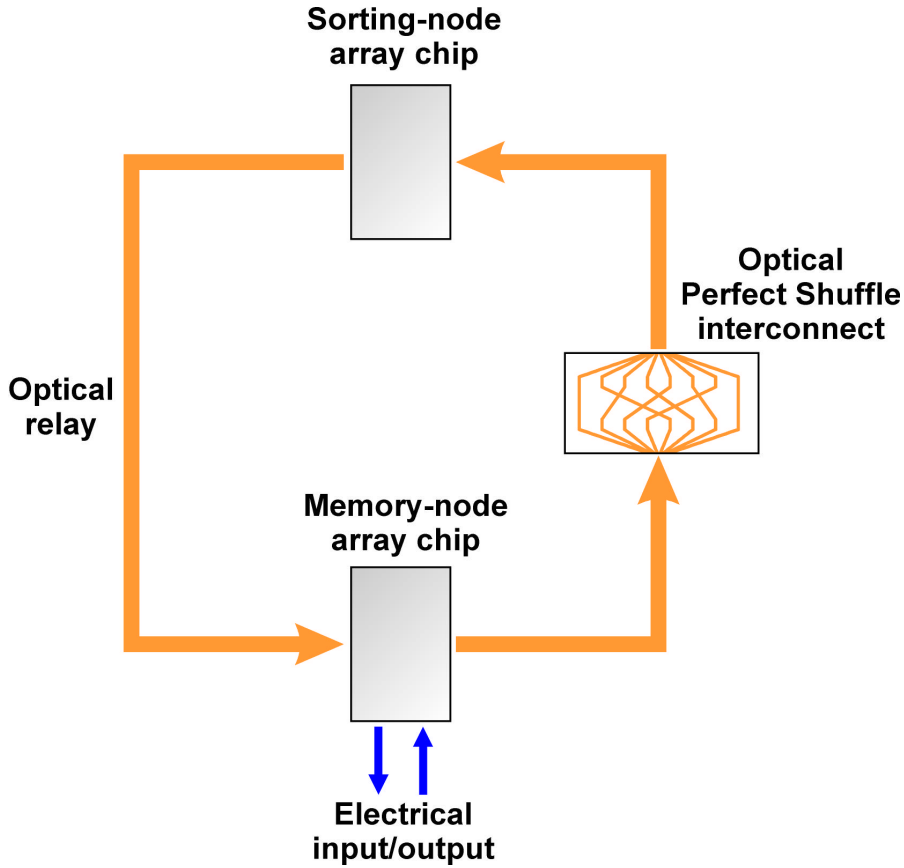
† Chip to board (off-chip) speed (high performance for peripheral buses.

†† Chip to package pads (peripheral).

Heriot-Watt University
Department of Physics

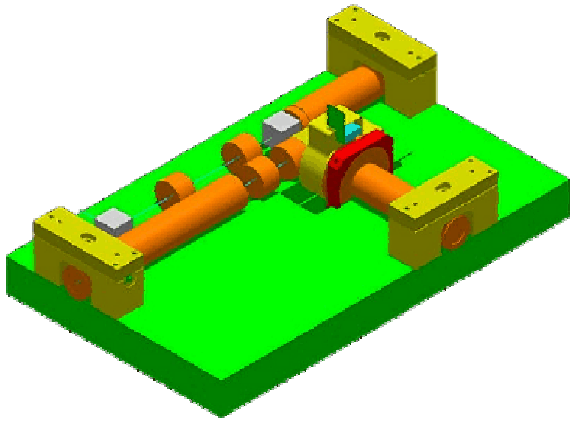
Origins : Heriot-Watt sorting demonstrator

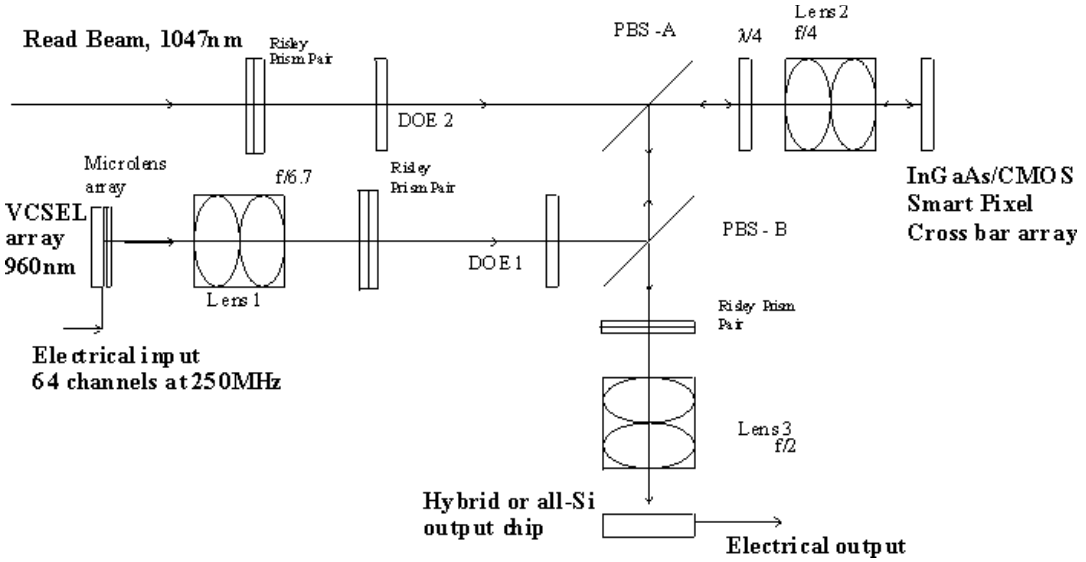
Batcher's bitonic sort



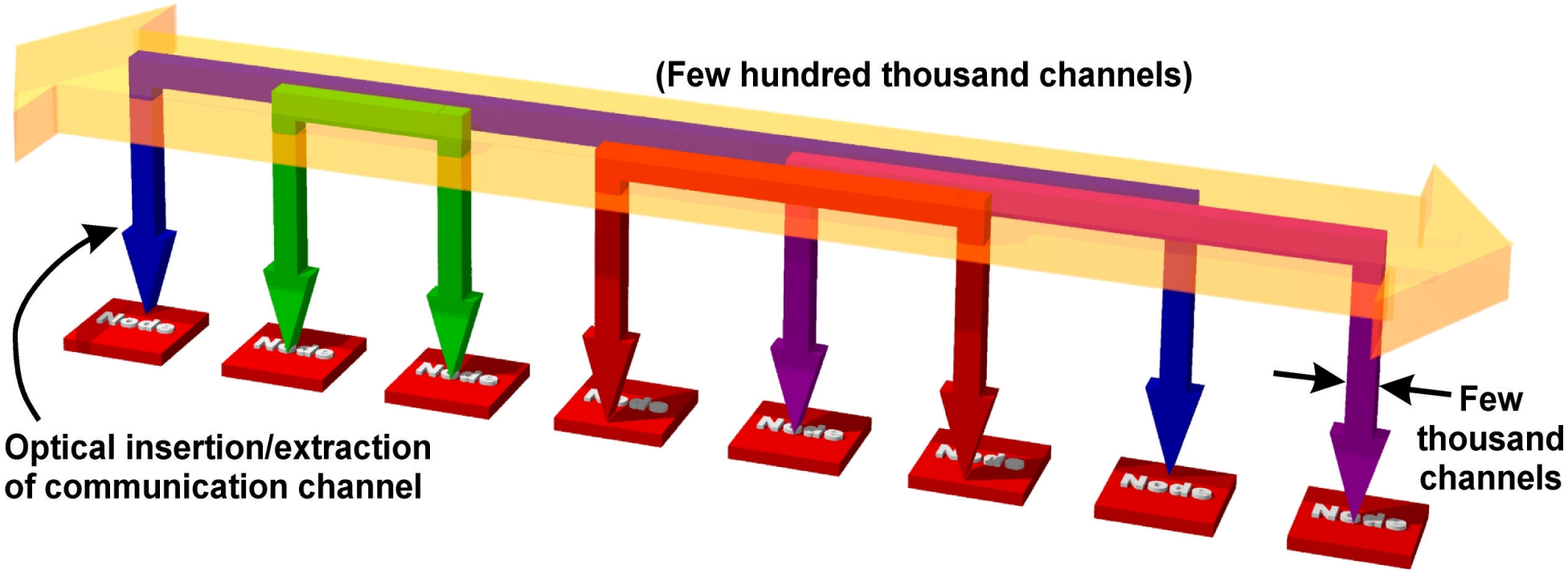
Origins : Heriot-Watt SPOEC demonstrator

Smart-Pixel Opto-Electronic Connects



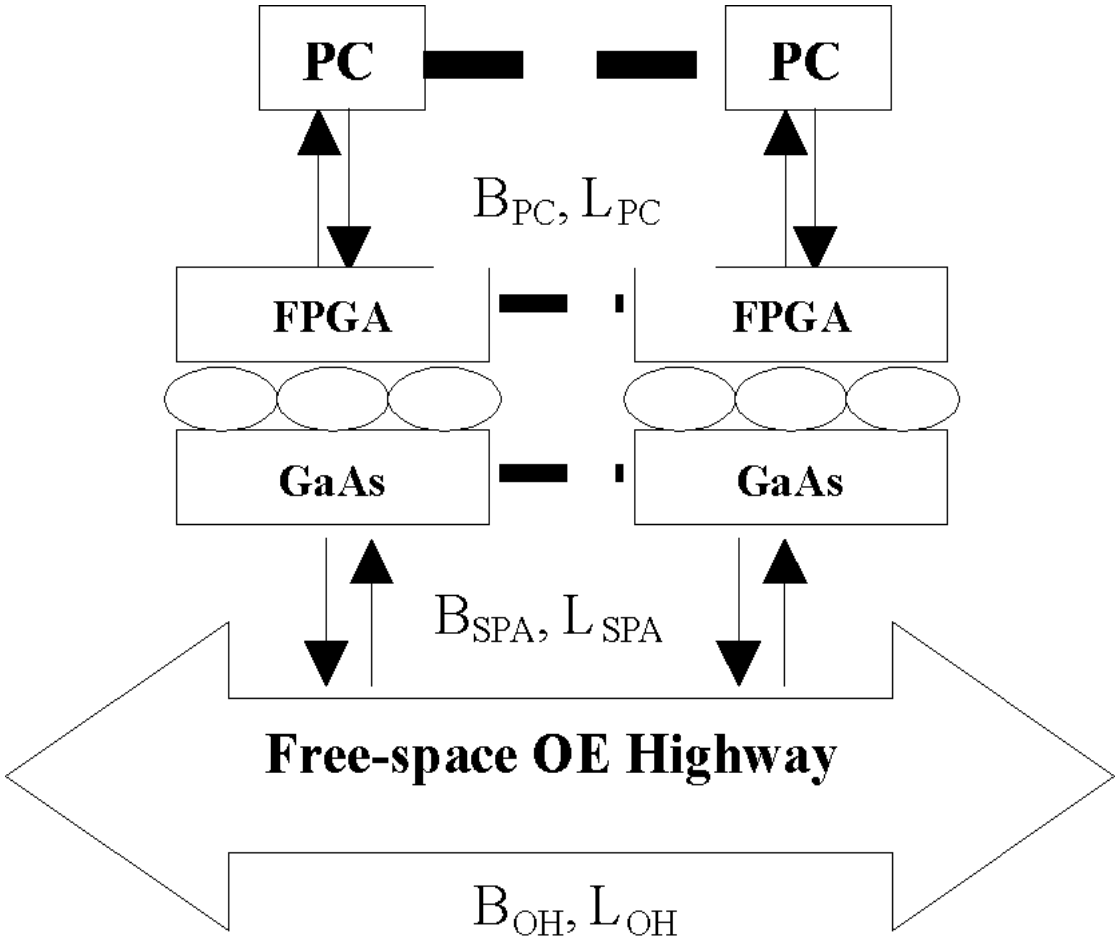


Optical Highways



The concept of optical highways is to provide a general purpose multiprocessor harness with several thousands of channels passed node to node via an OFPGA interface.

Opto-Electronic System Architecture



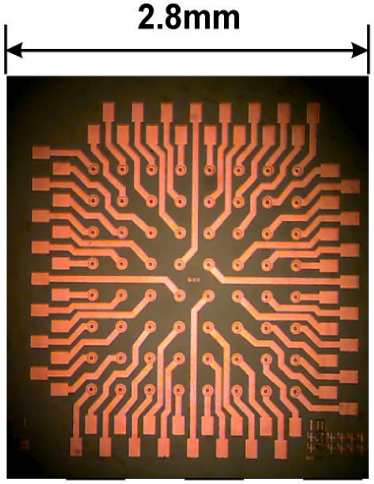
Heriot-Watt University
Department of Physics

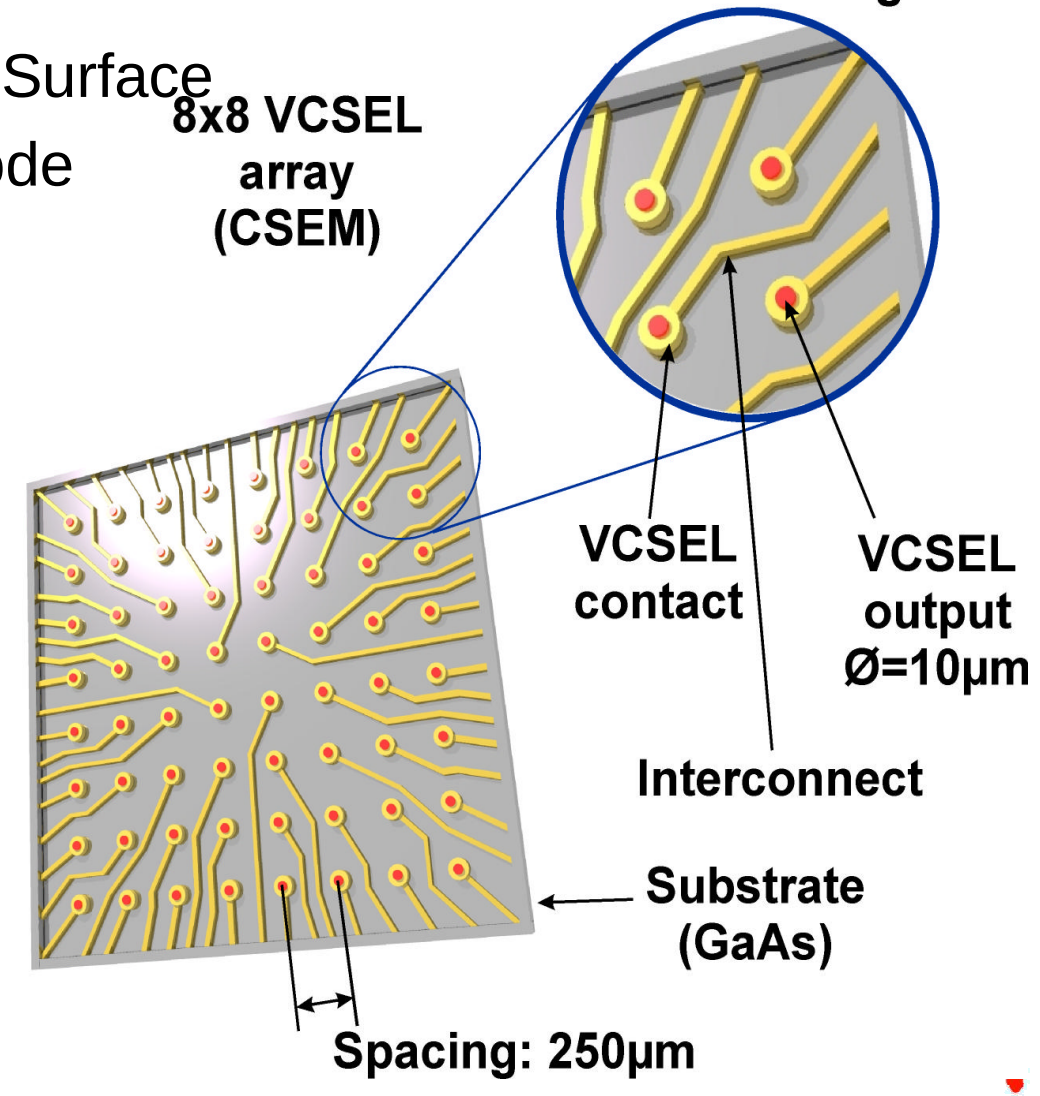
The VCSEL Array

The VCSEL (Vertical Cavity Surface Emitting Laser) is a laser diode that emits from the surface of the substrate.

Typical speeds are >1GHz.

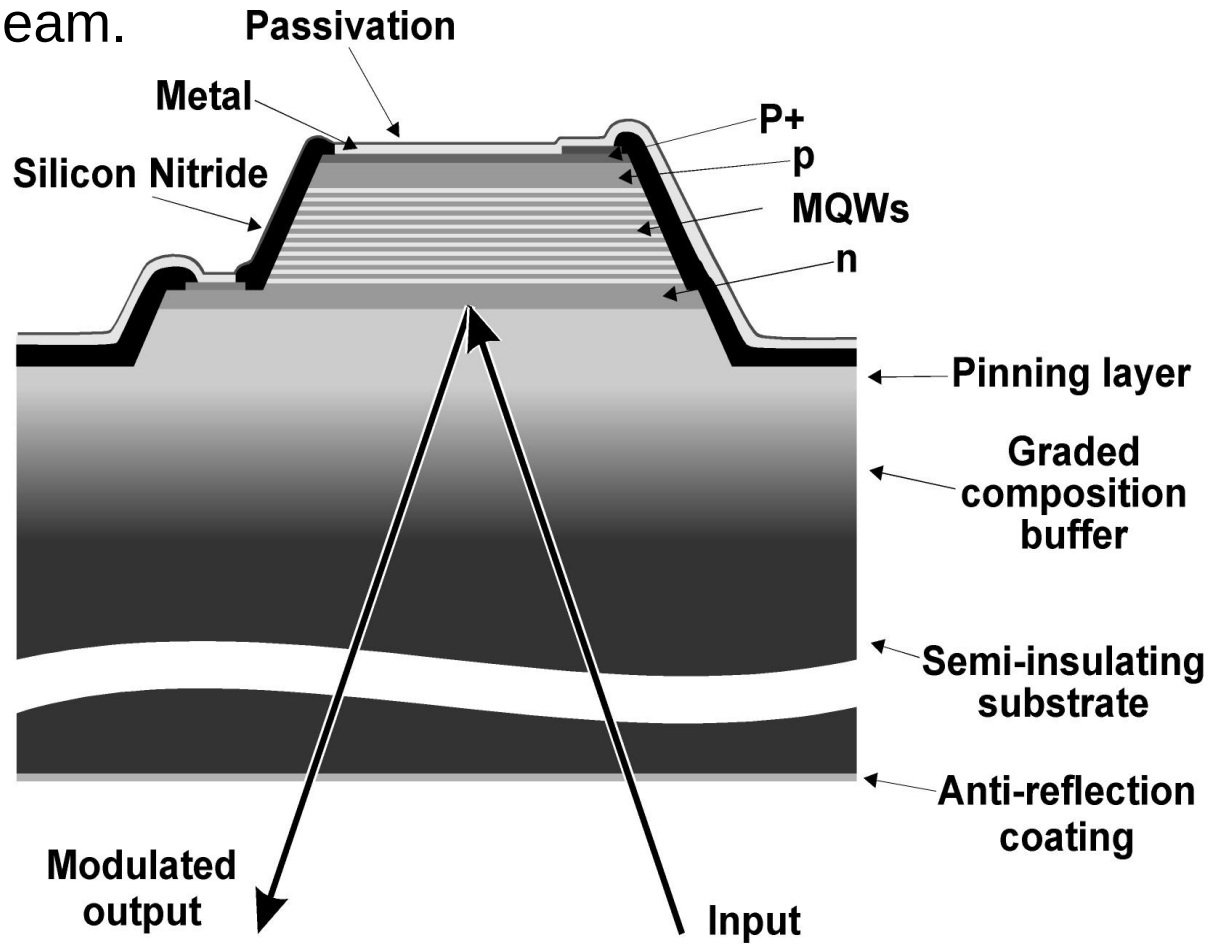
VCSEL array photograph
2.8mm



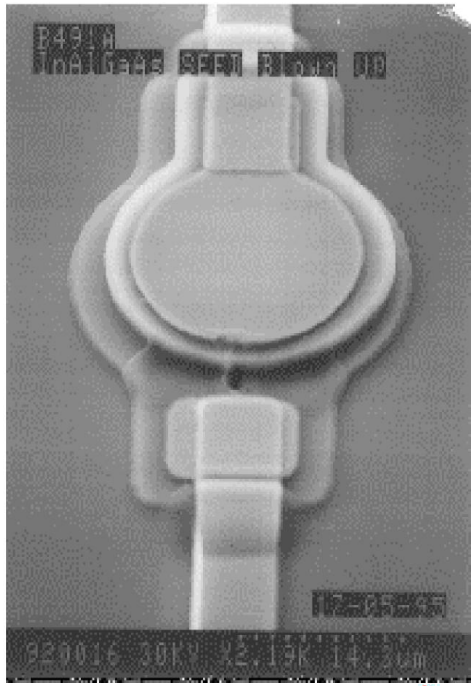


MQW Modulators

The MQW array does not produce its own output but modulates an input beam.



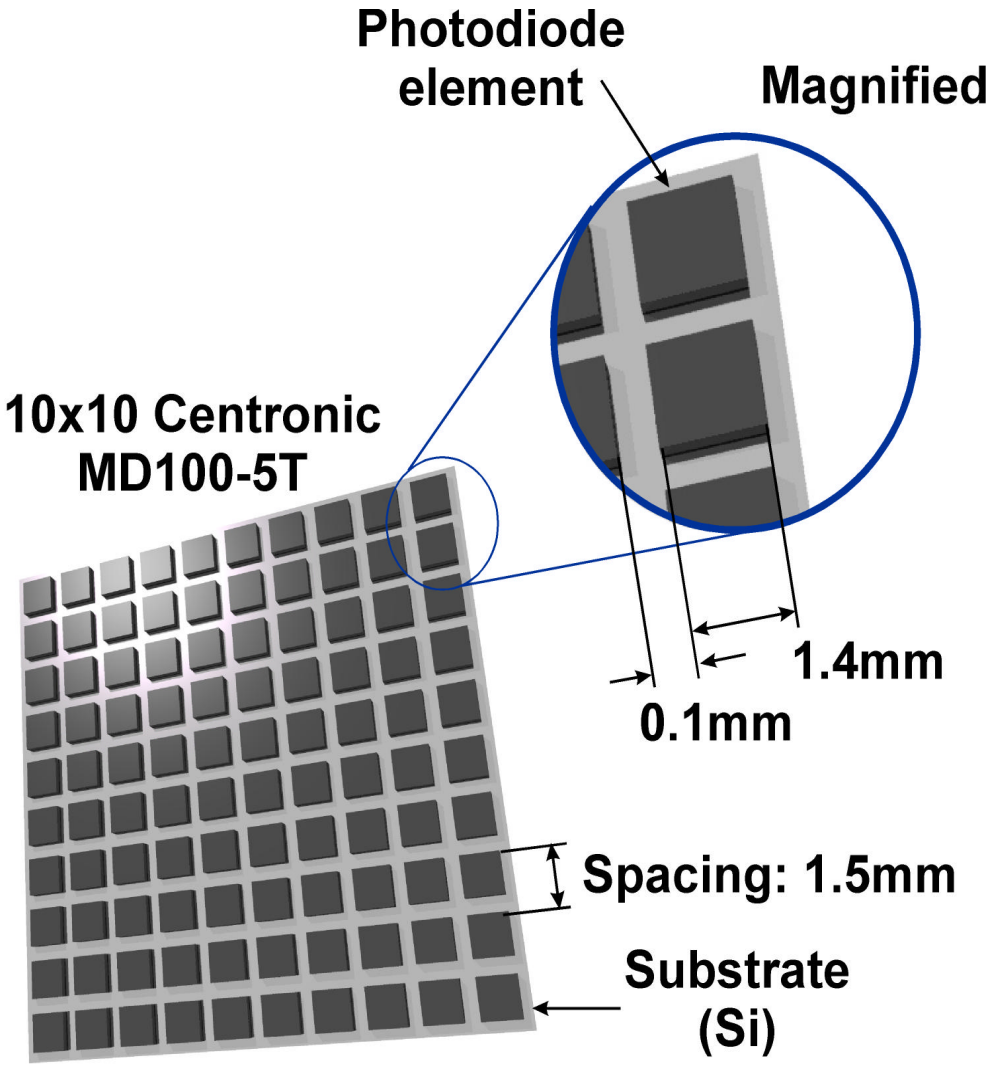
MQW Photograph



Detector Arrays

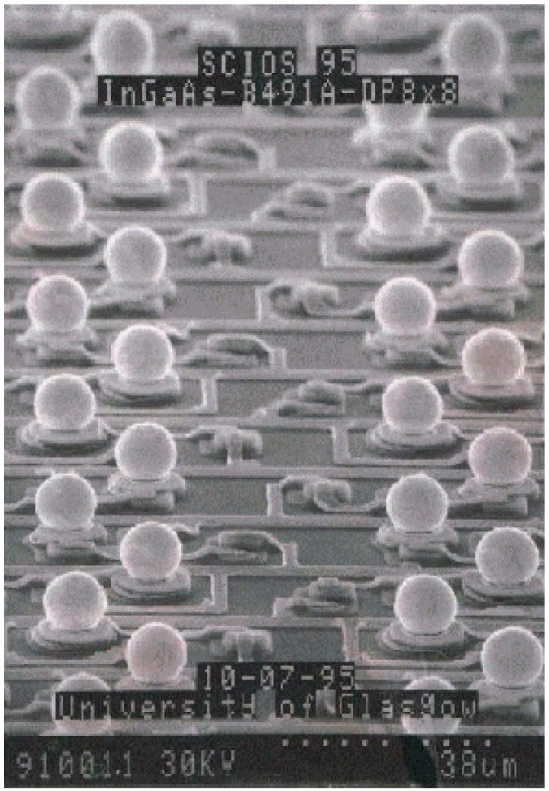
Photodetectors act as input devices and are currently available in a wide range off-the-shelf.

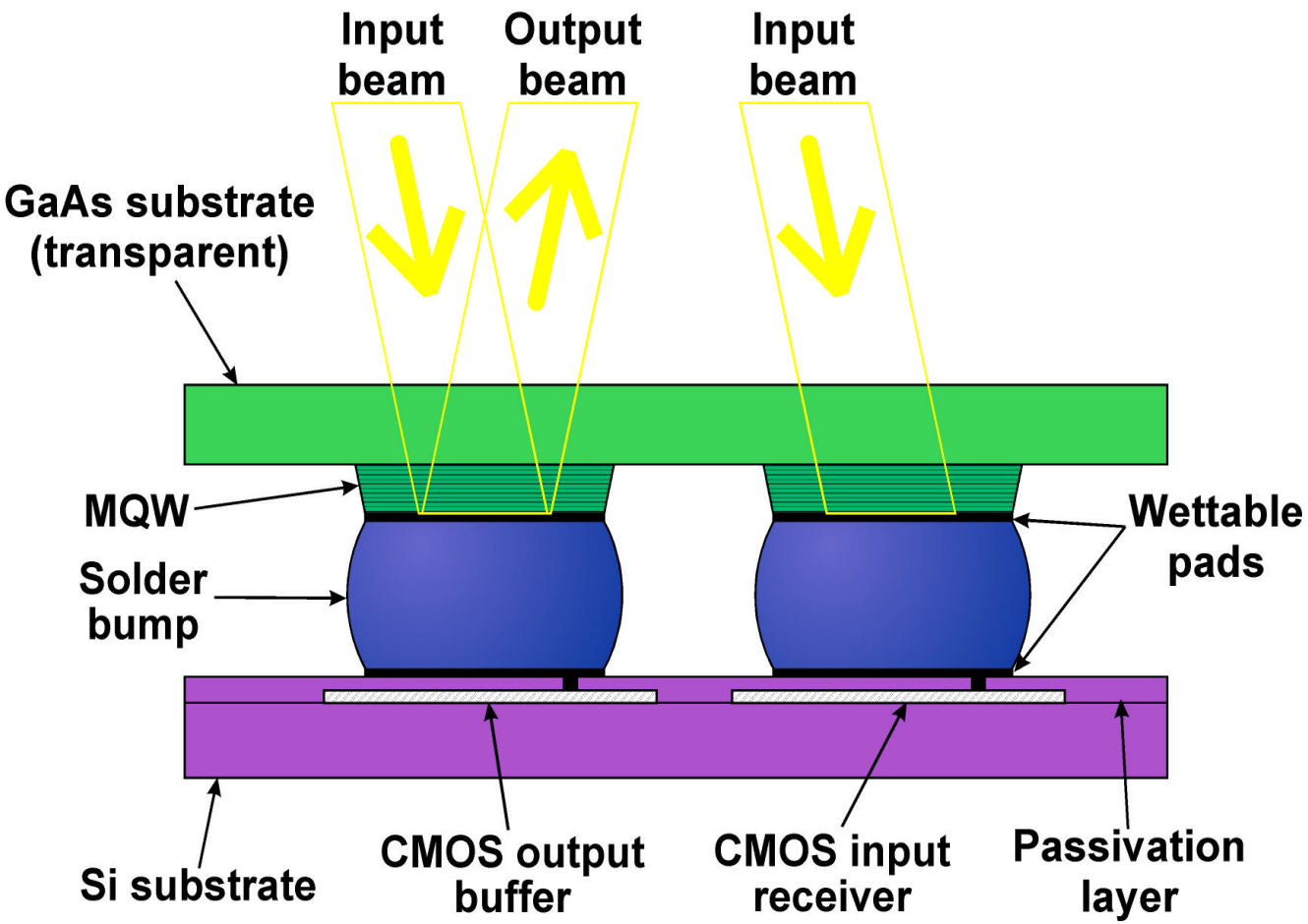
The faster they are driven the more power they require.



Flip-Chip Bonding

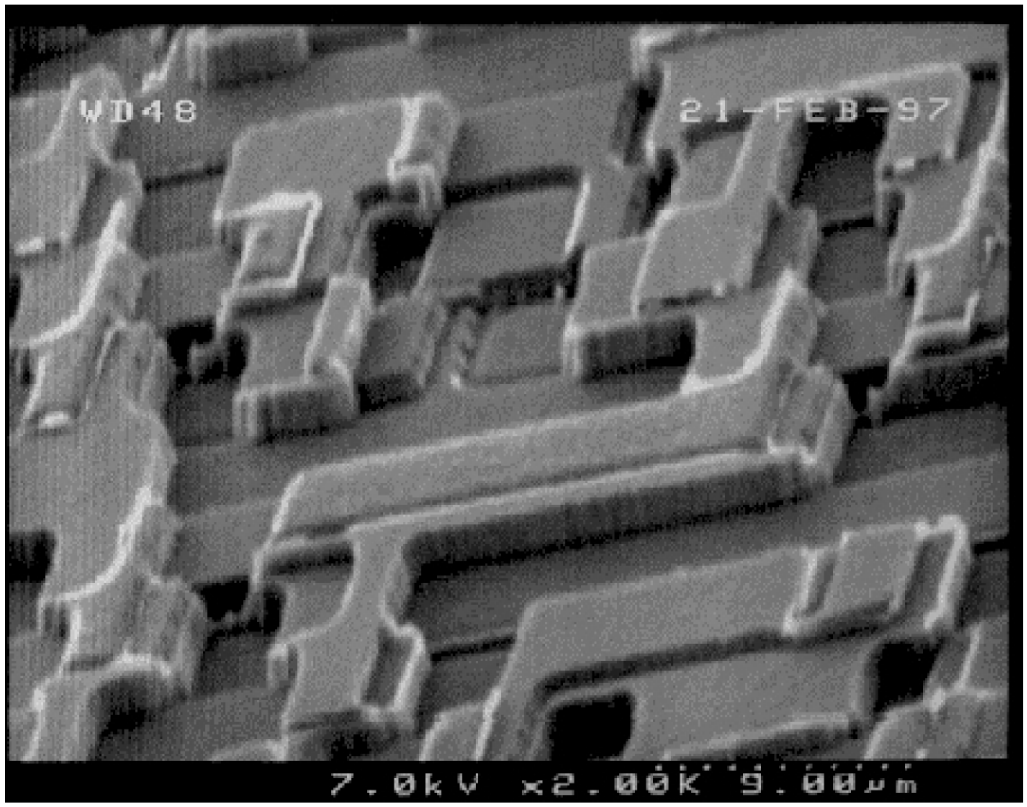
Solder bumps ready for flip-chip bonding process.



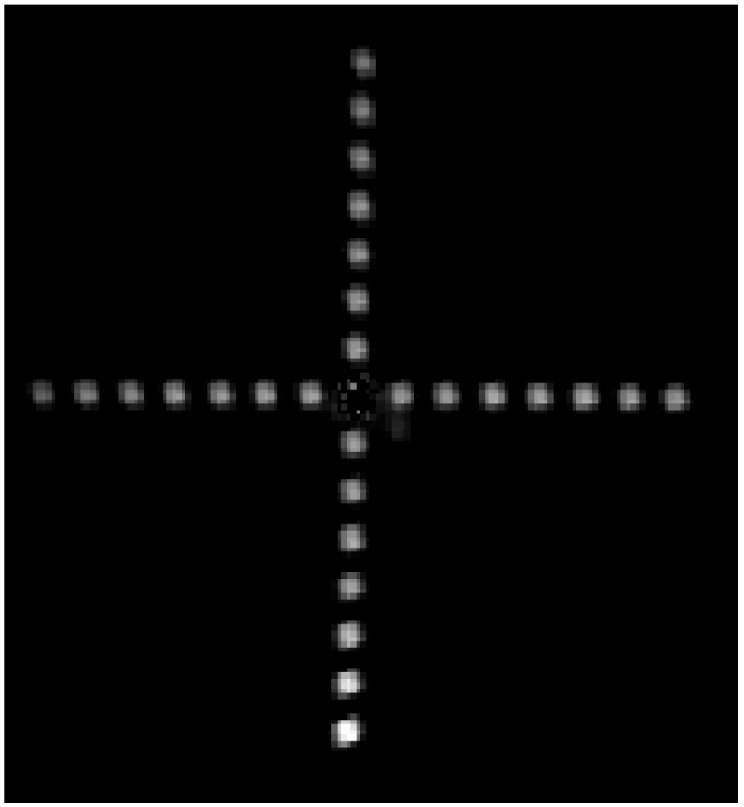


Diffractive Optic Elements (DOEs)

Sample DOE



DOE Output (Single beam input)



These elements are used as array generators and interconnection elements

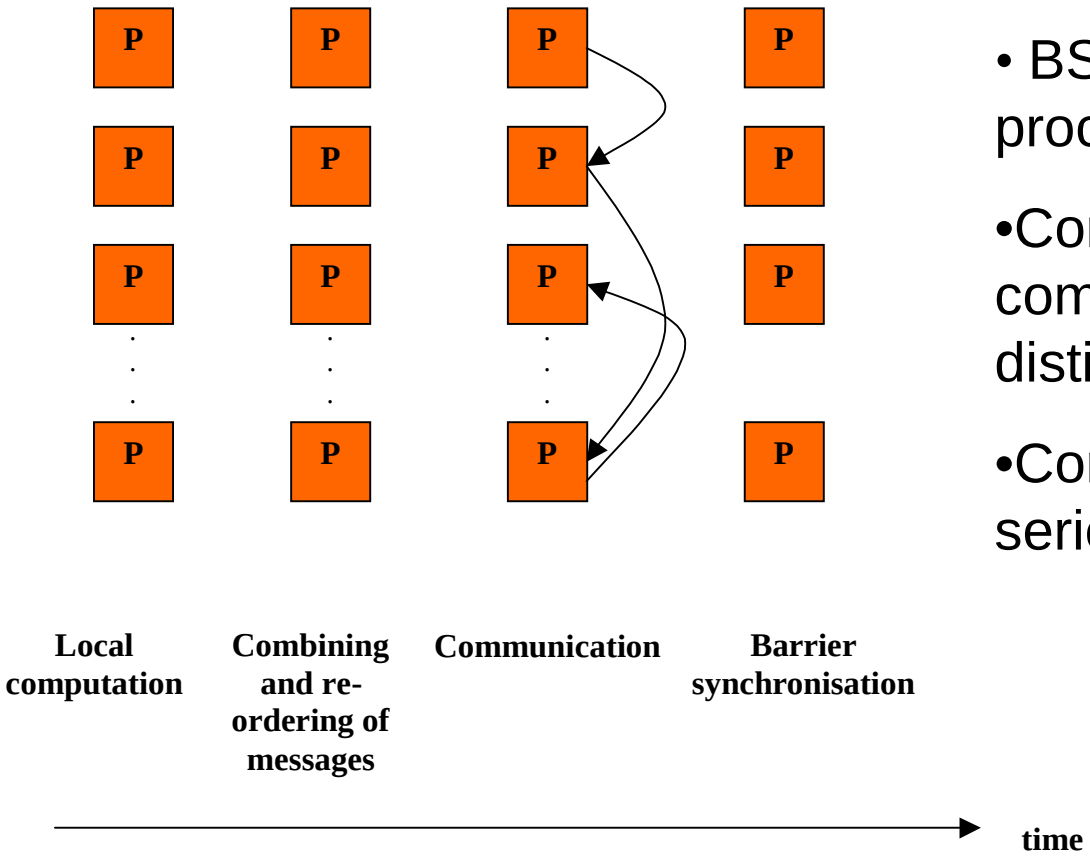
Heriot-Watt University
Department of Physics

Modelling OE systems: Why BSP?

- Communication of large messages required to exploit bandwidth?
  - BSP separates communication from computation.
  - Messages between processor pairs are combined.
- Simple cost model: few parameters.

Heriot-Watt University
Department of Physics

The BSP model



- BSP computer: processors+interconnect.
- Computation and communication occur in distinct phases.
- Computations are a series of *supersteps.*

Heriot-Watt University
Department of Physics

The BSP cost model

Cost of a superstep:

$$C = w + gh + l$$

Local computation

Communication

Barrier synchronisation

Heriot-Watt University
Department of Physics

Reducing-Sum

$$W = (\log(p) - 1) w_{SPA} + w_{PC}$$

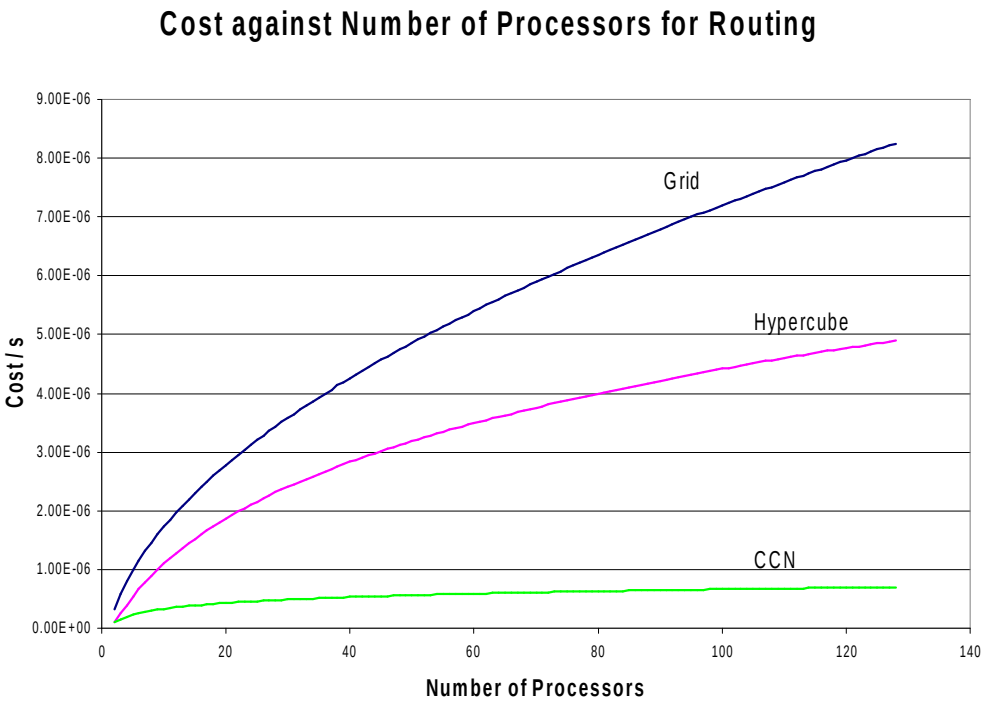
$$G = gh = 2\left(L_{PC} + L_{SPA} + R_{PC} + R_{SPA} + \frac{M_h}{B_{PC}} + \frac{M_h + M}{B_{SPA}} \right)$$

$$+ (\log_2(p) - 2)\left(L_{SPA} + R_{SPA} + \frac{M_h + M}{B_{SPA}} \right)$$

$$C = W + G$$

Heriot-Watt University
Department of Physics

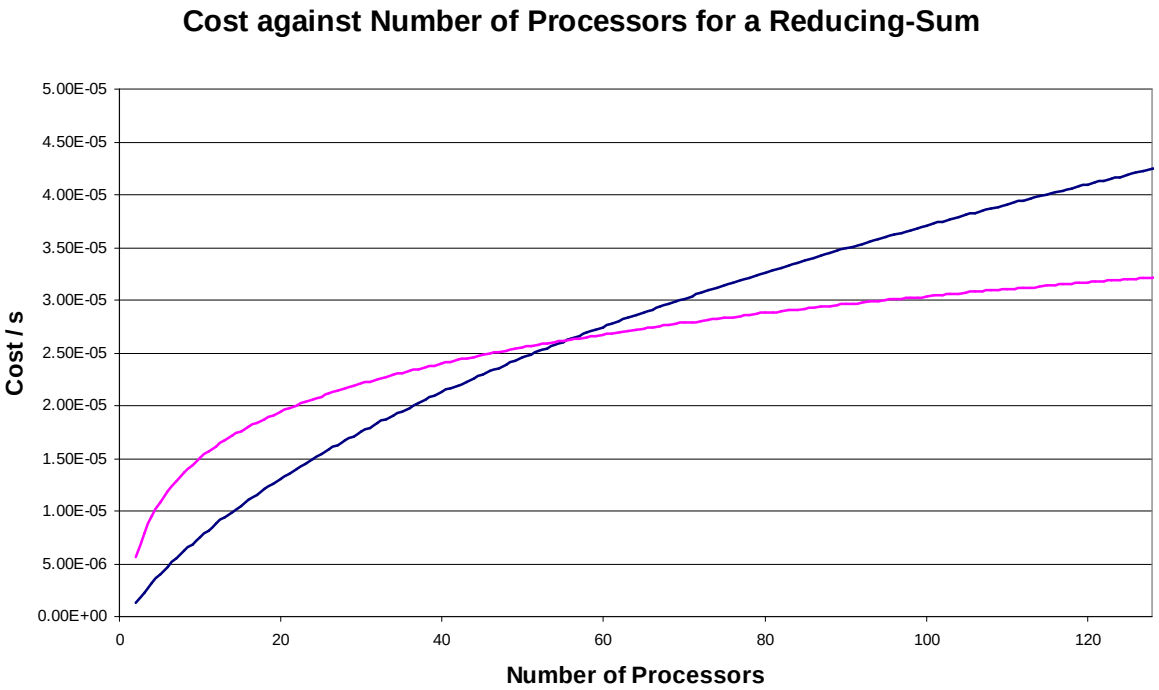
Cost of Routing



- Large electronic interconnects Pin limited - Grid or HC Topology
- Large Optical interconnects - HC+ or CNN

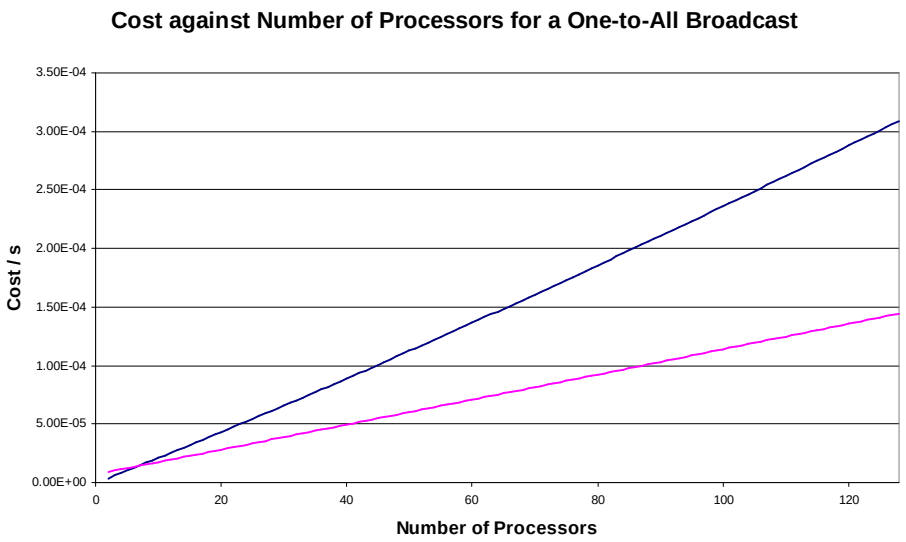
Heriot-Watt University
Department of Physics

Reducing-Sum



- Binary-tree algorithm
- Break-even 55 processors
- Most powerful cluster today >128 processors

Broadcast



- Break-even 8 processors
- Larger messages allow for added bonus from streaming and high bandwidth

Conclusions

- High connectivity reduced communication costs for small messages
- Optics provide high connectivity
- Low routing costs can speed small data and network control messaging

Heriot-Watt University
Department of Physics

Further Work

- Other operations / applications
- Design and build demonstrator system
- Topologies for high connectivity
- Generic OFPGA application

Heriot-Watt University
Department of Physics