An IP-over-OPS network

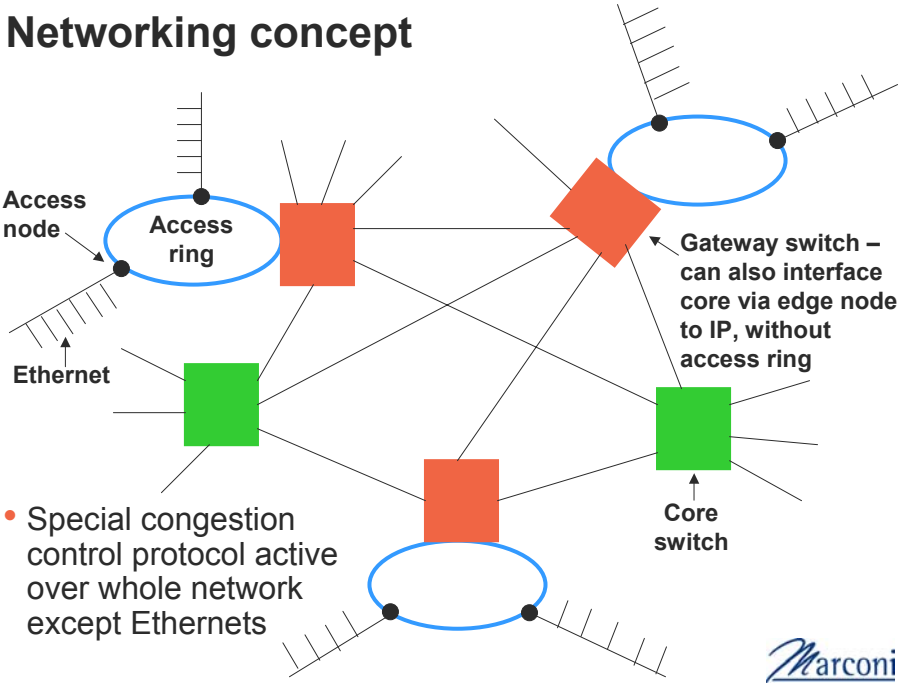
David K. Hunter & Derek R. McAuley,
Marconi Labs, Cambridge UK

With acknowledgements to:
Sven Östring, Andrew Moore & Rolf Neugebauer of Marconi Labs, and staff at Centre for Photonics Communications Systems Research, University of Cambridge

Rationale of SOAPS project (Smoothed Optical ATD Packet Switching)

- 3-5 year timescale
- Electronics used for:
  - Buffering packets and forwarding them from access & IP to the core (and vice-versa)
  - Header processing
  - Control and packet scheduling within switches/routers
- Optics used for:
  - Switching and minimal buffering in core
- Don't try to reproduce present-day electronic IP routers in optics
- Probable cost benefits through:
  - Reduction in e/o and o/e conversion
  - Hardware simplification
- Routing protocols left for later study.

Networking concept



- Special congestion control protocol active over whole network except Ethernets

Marconi

Overview: distinguishing features of SOAPS

- Smoothing and congestion control at the edge of the network.
- Reduced core optical buffering, due to edge processing.
- Fixed-length “slot” structure treated as a byte stream.
- Routing of slots, each encoded using Coarse WDM (CWDM).
- Electronics for edge switching and buffering.
- Optics for core switching and buffering.
- Electronics for control and header processing throughout the network.
- Programme of work designed to show operation with real traffic scenarios and real protocols.

Marconi

Optical buffering in SOAPS

- Many superimposed CBR flows of varying bitrates and random phase on each link (as above)
- Bernoulli traffic yields a pessimistic estimate of buffer depth [Roberts TCOM Feb 91, Karol TCOM Dec 87]
  - Even for 900 CBR flows
- 30 slot buffer with Bernoulli traffic:
  - 10^{-10} PLP for load of 0.7
  - $2x10^{-7}$ PLP for load of 0.8 [Hluchyj JSAC Dec 88]

Marconi

Implementation of optical buffering

- 50ns slot
  - 250 bytes at 40Gb/s (including say 10ns guard band)
  - 1 + 2 + … + 30 slots = 465 slots = 4650m fibre delay-line
- 5km fibre = drum of fibre:
  - Inner diameter 60mm
  - Outer diameter 100mm
  - Thickness 13.5mm
  - Easily temperature controlled
- Proposed core node OPS architecture uses 244 slots for 4x4 switch
- Wavelength cannot be used to select delay (using e.g. AWG or Fibre Bragg Grating) since each slot is itself multi-wavelength

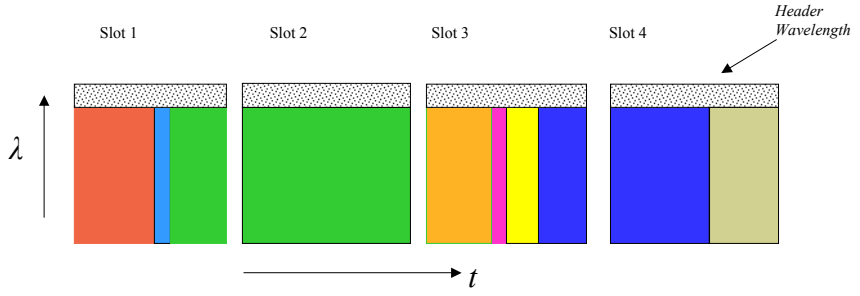
Marconi

Packet header format

- Header on separate wavelength
- Encoded in standard GigE frames
  - Use off-the-shelf components
- Header fields:
  - Label – determines path
  - Data length – allows empty or partially full slots
  - Sequence – facilitates detection of out-of-order slots e.g. after a fault
  - Source edge node – required when reassembling datagrams
  - Congestion indication – ECN information (see later)
  - OAM – relevant functionality of SDH OAM

Marconi

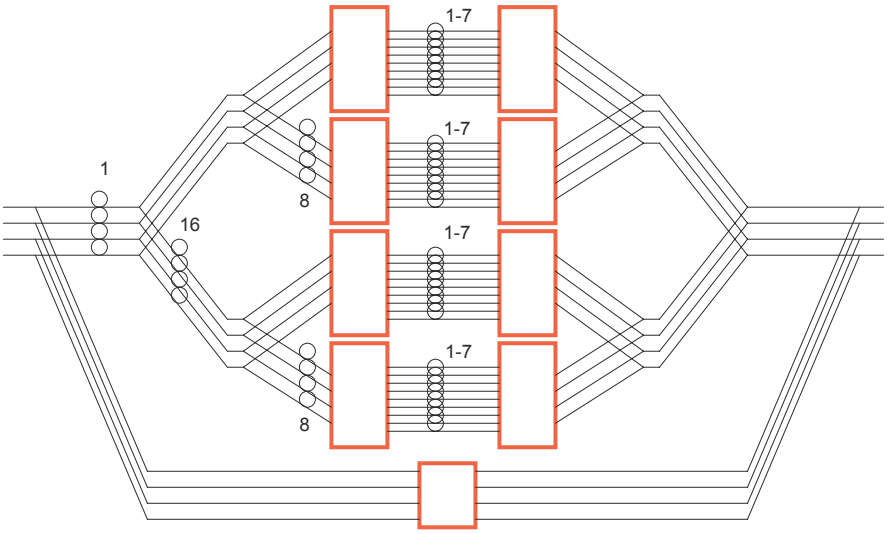
CWDM slot format



- Coarse WDM – aim for uncooled components
- Slot payload treated as a bitstream
- Low bitrate header on separate wavelength
  - Low cost o/e and e/o components in core

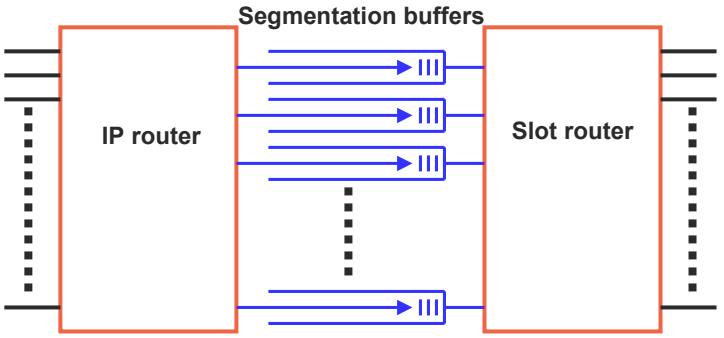
Marconi

Core router design



Marconi

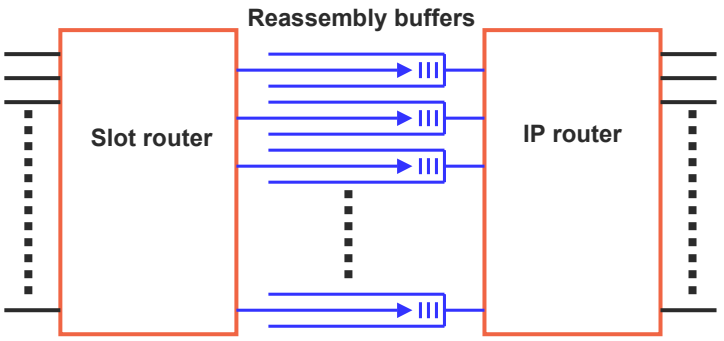
Edge router design – edge to core



- IP router forwards according to destination edge node
- Data is groomed according to destination edge node
- Slot router forwards according to exit link derived from destination edge node number in header

Marconi

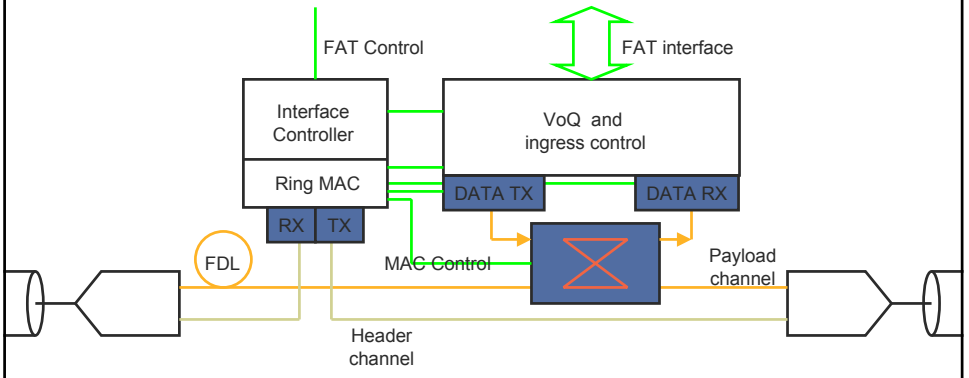
Edge router design – core to edge



- Slot router forwards according to source edge node
- Slots are groomed according to source edge node
- IP router forwards as normal

Marconi

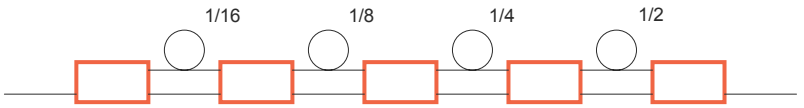
Ring access node for demonstrator



- Photonic 2x2 switch does not save on o/e and e/o costs
- May as well use electronic switch
- Photonic switch will provide valuable information

Marconi

Synchronisation

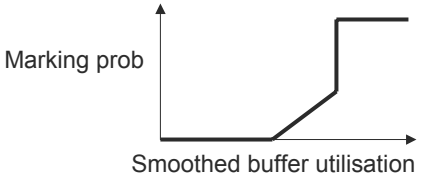


- Implemented using standard techniques:
- Coarse synchronisation
  - Variable switched delay-lines (as above)
  - Only necessary to any extent in the core
- Necessary to detect start of each slot and compute required delay
- Need arises due to:
  - Drift and wander in transmission
  - Optical packet switch operates synchronously

Marconi

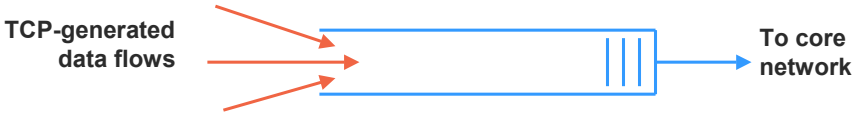
Slot marking schemes for the core

- Mark slots to indicate degree of congestion and allow TCP sources to adjust their rates
- Marks are fed back to source
- Kelly's virtual buffer
  - Model virtual buffer which is smaller than real one
  - Subject it to same traffic
  - Mark slots when virtual buffer overflows
- RED-type marking
  - Probability a packet is marked depends on buffer utilisation



Marconi

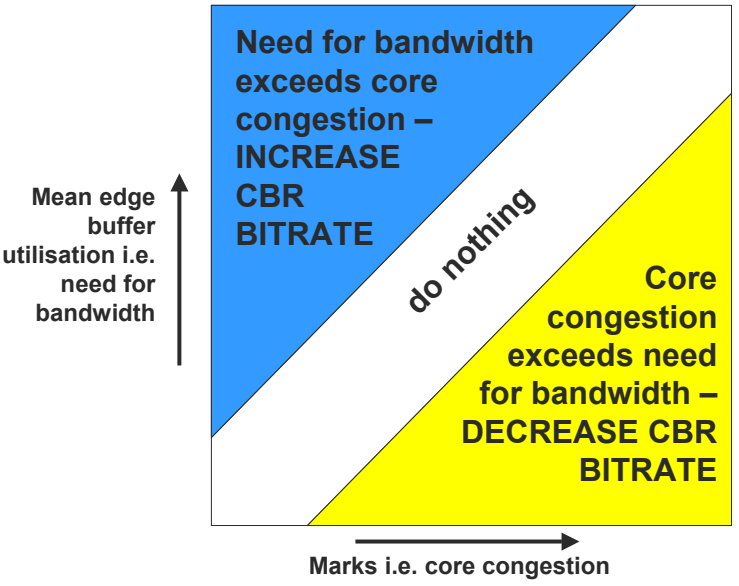
Edge segmentation and smoothing buffer



- High utilisation = need for more core capacity
- Low utilisation = capacity requirements satisfied
- Datagrams are dropped instead of marked (adapt one of previous schemes)
- Slots leave the buffer as a CBR (Constant Bit Rate) stream
- Rate of CBR stream is periodically adjusted in response to edge buffer utilisation and core marks…

Marconi

Co-ordination of core and edge congestion



Marconi

Further work

- Objective: show validity of basic concept with real traffic and real protocols.
- Physical layer and switch design:
  - Refine switch architectures by simulation.
  - Refine existing devices for use in this application.
  - Investigate problems of physical implementation (noise, crosstalk, loss).
  - Investigate viability of different optical switching technologies, for this application.
- Protocols and the control plane:
  - Develop, simulate and refine protocols.
  - Investigate routing protocols suitable for edge and core parts of the network.
- Evaluation:
  - Build a test-bed implementation of an optical packet switch plus a ring network, to carry real Internet traffic.

Marconi