Design of an Agile All-Photonic Network (AAPN)

Gregor v. Bochmann

School of Information Technology and Engineering (SITE)
University of Ottawa
Canada

http://www.site.uottawa.ca/~bochmann/talks/AAPN-results

uOttawa

L'Université canadienne
Canada's university

Presentation at the APOC conference

Wuhan, November 2007

Abstract

- Agile All-Photonic Networks (AAPN) is a Canadian research network (funded by NSERC and 6 industrial partners) exploring the use of very fast photonic switching for building optical networks that allow the sharing (multiplexing) of a wavelength between different information flows. The aim is to bring photonic technology close to the end-user in the residential or office environment. The talk gives an overview of the proposed overlaid star network architecture and describes new results on (a) bandwidth allocation algorithms, (b) the routing and protection of MPLS flows over an AAPN using the concept of OSPF areas, and (c) our evolving plans for building demonstration prototypes.

Overview

- Overview of the AAPN project
- Frame-by-frame bandwidth allocation
- MPLS over AAPN
- A demonstration prototype
- Conclusions

Different forms of “burst switching“

- **Question:** Can one do packet switching in the optical domain (without oeo conversion)?
- At a switching speed of 1 µs, one could switch bursts of 10 µs length (typically containing many packets)
- Traditional packet switching involves packet buffering in the switching nodes. Should one introduce optical buffers in the form of delay lines?
- The term “burst switching“ originally meant “no buffering”: in case of conflict for an output port, one of the incoming bursts would be dropped.

- **Note:** Burst switching allows to share the large optical bandwidth among several virtual connections.

AAPN

An NSERC
Research Network

The Agile All-Photonic Network

Project leader: David Plant, McGill University

Theme 1: Network architectures
Gregor v. Bochmann, University of Ottawa

Theme 2: Device technologies for transmission and switching

AAPN Professors (Theme 1 in red)

- McGill: Lawrence Chen, Mark Coats, Andrew Kirk, Lorne Mason, David Plant (Theme #2 Lead), and Richard Vickers
- U. of Ottawa: Xiaoyi Bao, Gregor Bochmann (Theme #1 Lead), Trevor Hall, and Oliver Yang
- U. of Toronto: Stewart Aitchison and Ted Sargent
- McMaster: Wei-Ping Huang
- Queens: John Cartledge (Theme #3 Lead)

- Note: Theme 2 deals with device technologies for transmission and switching

For further information see: http://www.aapn.mcgill.ca/

The AAPN research network

- Our vision: Connectivity “at the end of the street” to a dynamically reconfigurable photonic network that supports high bandwidth telecommunication services.
- Technical approach:
  - Simplified network architecture (overlaid stars)
  - Specific version of burst switching
    - Fixed burst size, coordinated switching at core node for all input ports (this requires precise synchronization between edge nodes and the core)
      - See for instance http://beethoven.site.uottawa.ca/dsrg/PublicDocuments/Publications/Hall05a.pdf
    - Burst switching with reservation per flow (virtual connection), either fixed or dynamically varying
      - See for instance http://beethoven.site.uottawa.ca/dsrg/PublicDocuments/Publications/Agus05a.pdf

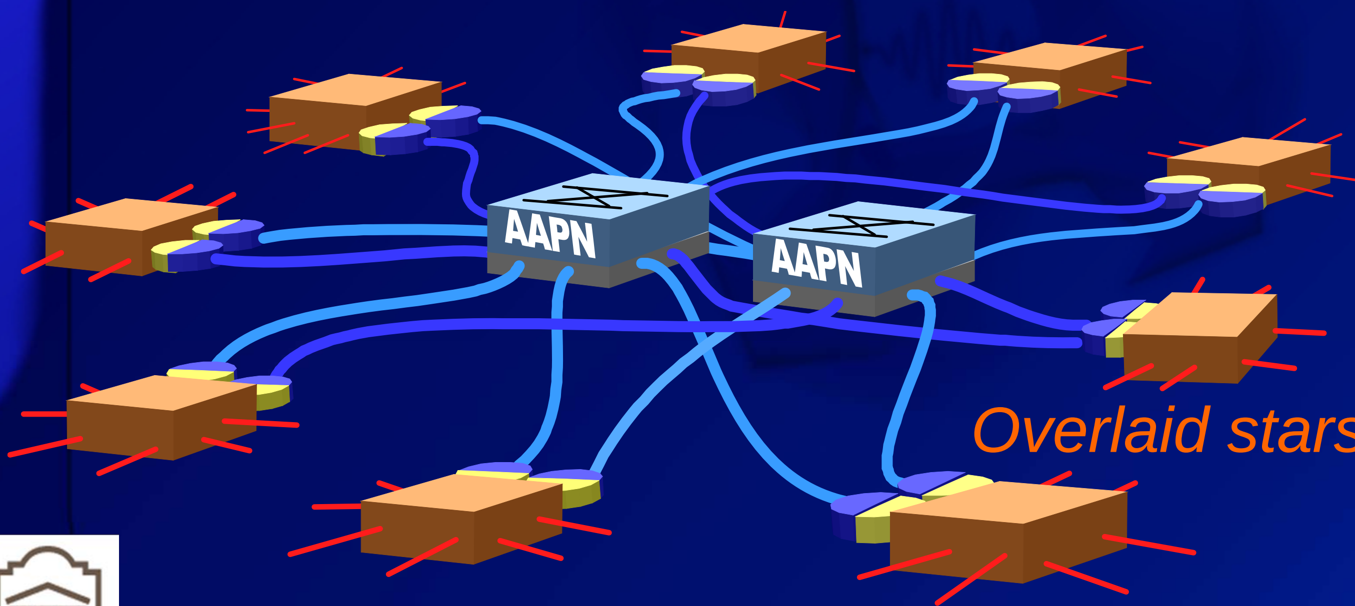
Agile All-Photonic Network

Edge node with slotted transmission (e.g. 10 Gb/s capacity per wavelength)

Fast photonic core switch (one space switch per wavelength)

Opto-electronic interface

- *Provisions sub-multiples of a wavelength*
- *Large number of edge nodes*



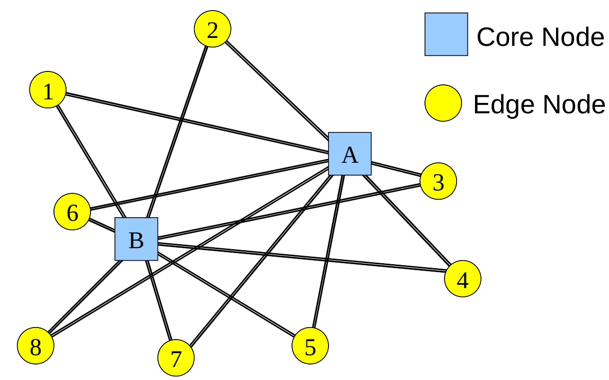
Overlaid stars architecture

Starting Assumptions

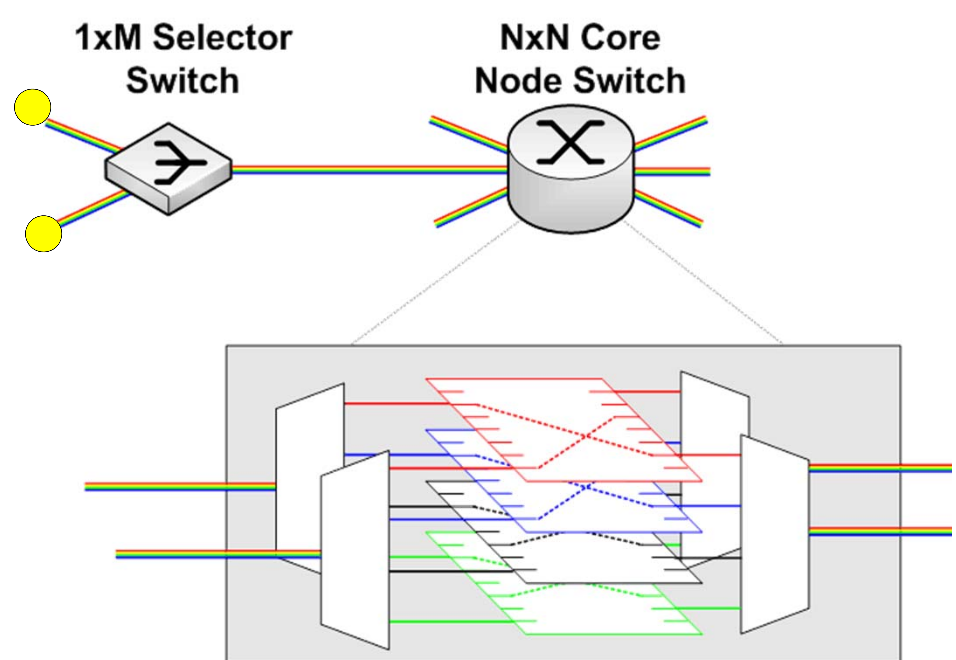
- Avoid difficult technologies such as
  - Wavelength conversion
  - Optical memory
  - Optical packet header recognition and replacement
- Current state of the art for data rates, channel spacing, and optical bandwidth (e.g. 10 Gbps)
- Simplified topology based on overlaid stars
- Large number of simple edge nodes (e.g. 1000)
- Fixed transmission slot length (e.g. 10 μsec)
- No distinction between long-haul and metro networks
- This requires
  - Fast optical space switching (<1 μsec)
  - Fast compensation of transmission impairments (<1 μsec)

AAPN Architecture

- Overlaid stars



- Port sharing is required to allow a core node to support large numbers of edge nodes
  - A *selector* may therefore be used between edge and core nodes
- A wavelength stack of bufferless transparent photonic switches is placed at the core nodes
  - a set of space switches, one switch for each wavelength



Deployment aspects - Questions

- Long-haul or Metro ?
  - connectivity "at the end of the street"; to a server farm
  - AANP as a backbone network ?
- High capacity (many wavelengths) or low capacity (single or few wavelengths) ?
- Multiple core nodes ?
  - For reliability
  - For load sharing
- Transmission infrastructure ?
  - Using dedicated fibers
  - Using wavelength channels provided by ROADM network

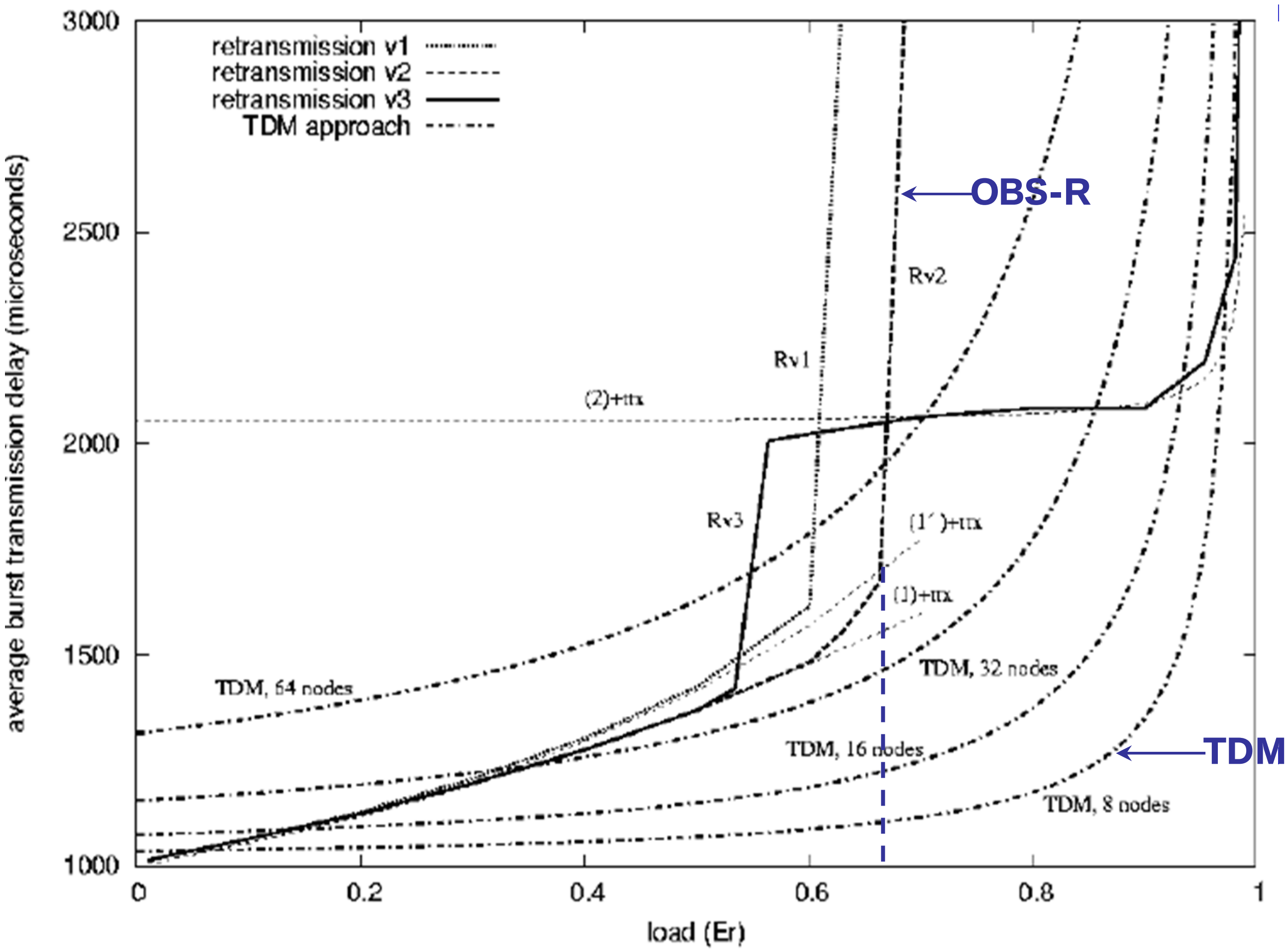
Overview

- Overview of the AAPN project
- Frame-by-frame bandwidth allocation
  (work by my PhD student Cheng Peng)
- MPLS over AAPN
- A demonstration prototype
- Conclusions

Comparing Burst-Mode Schemes

- Long-haul AAPNs: long propagation delays for signalling
- Two modes of slot transmission:
  - With reservation (long signalling delay)
  - Without reservation, as proposed for "Burst Switching" (loss probability due to collisions)
- Collaboration with Anna Agusti-Torra (Barcelona)
  - New method: Burst switching with retransmission (to avoid losses)
  - Comparison with TDM (see next slide)
- Method to avoid long signaling delays: TDM
  - Allocate unused time slots; these free slots can be used without signaling delays (they were allocated in advance)

TDM vs. OBS



What kind of technologies should be employed in the AAPN, TDM or OBS?

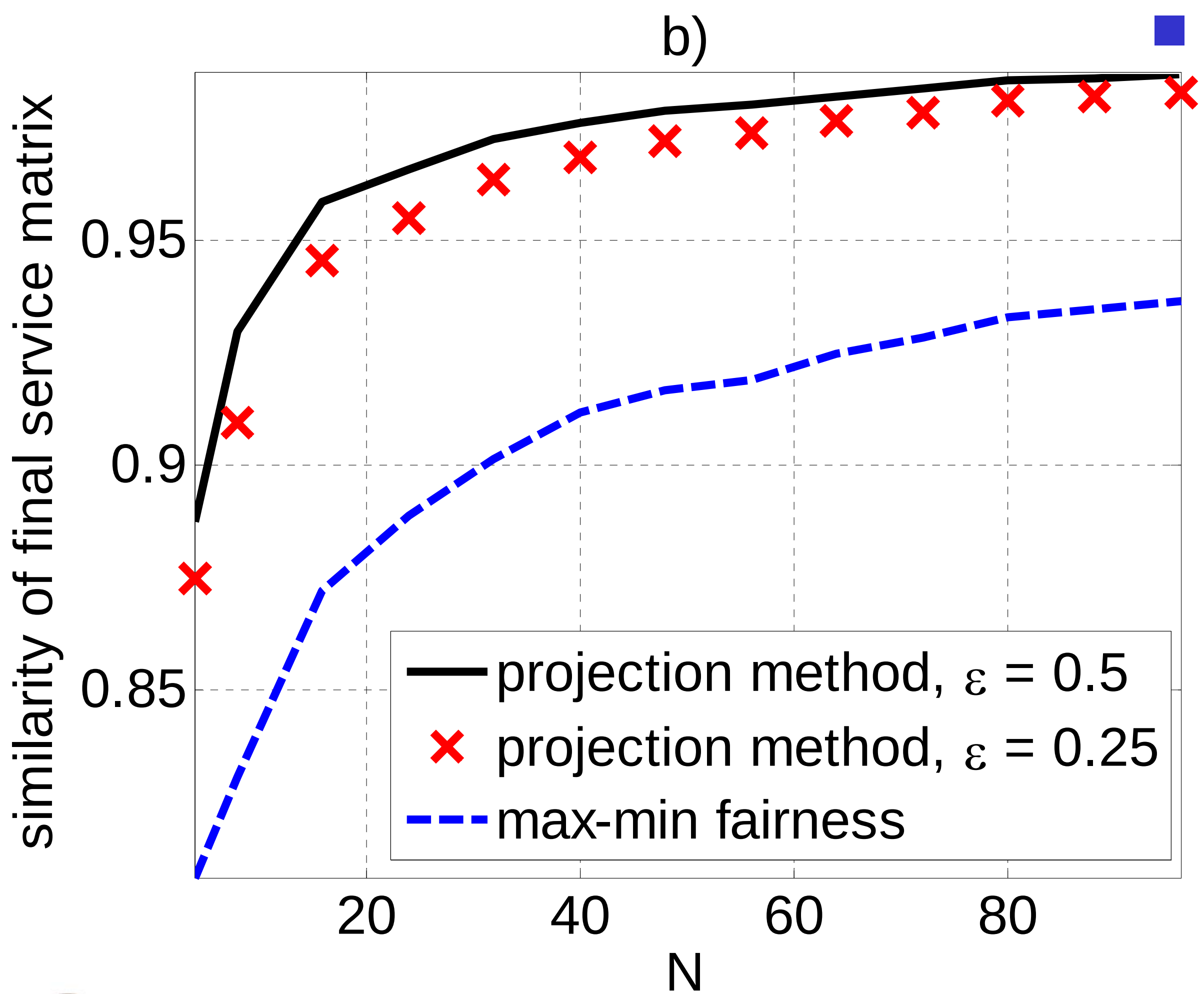
- The delay of OBS w/ retransmission (OBS-R) degrades sharply when the load is beyond 0.6 but is negligible at lower load.
- The delay of TDM maintains better delay performance at the high load compared with OBS-R.

TDM shows a better performance than OBS-R especially at the high load.

Birkhoff - von Neumann Approach

- The BvN decomposition approach calculates the timeslot schedules for a frame from the traffic demands between all node pairs.
- Two steps:
  - Constructing a service matrix from a traffic matrix
  - Decomposing the service matrix into switch permutations. (problem has $O(N^{4.5})$ complexity)
- The main challenges of BvN Decomposition are:
  - **How to construct a service matrix that closely reflects the traffic demand for all source-destination pairs?**
  - **How to find a heuristic decomposition algorithm with low complexity that allows a practical implementation?**

Service matrix construction

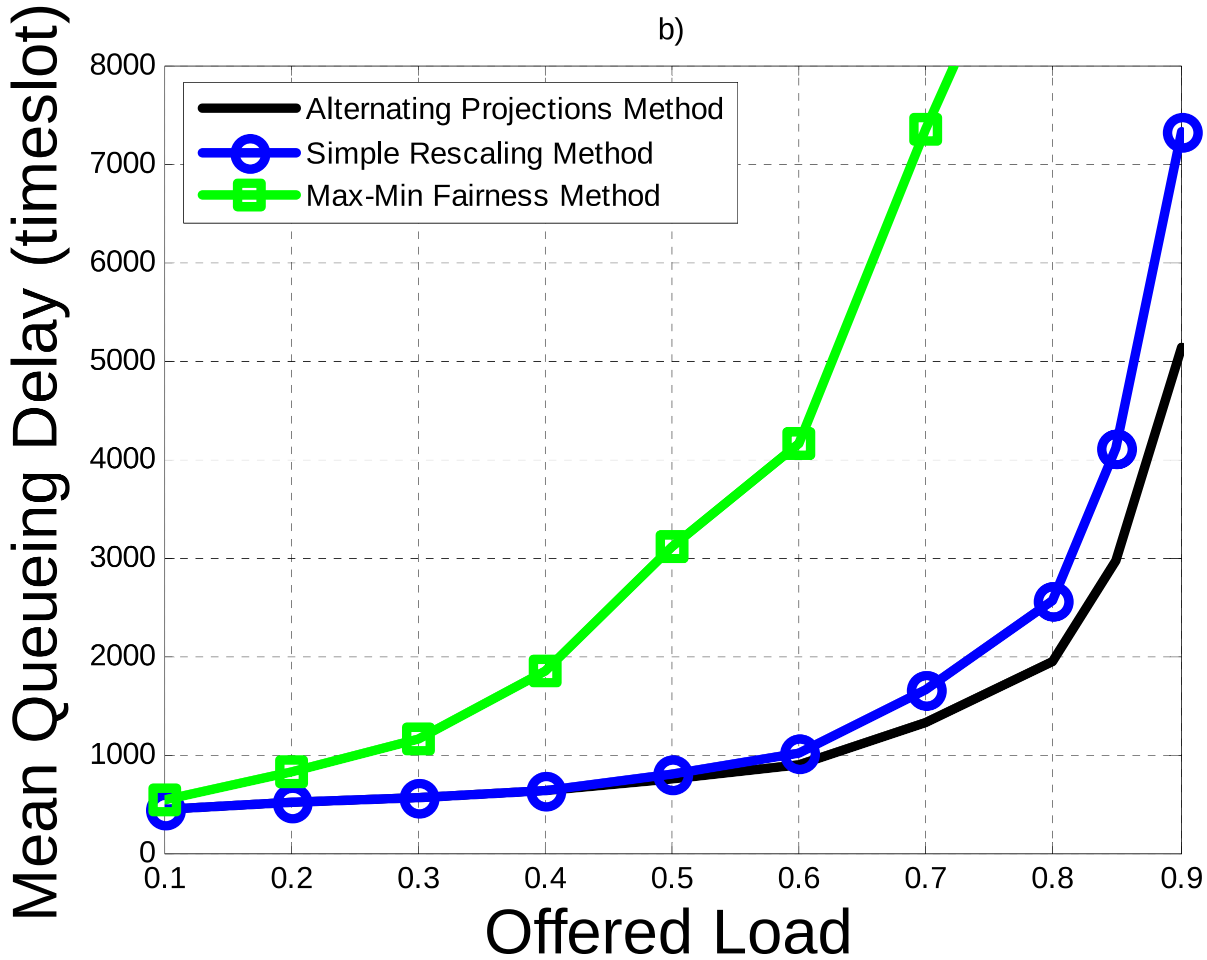


- New algorithm: Alternating Projections Method

Similarity Comparison

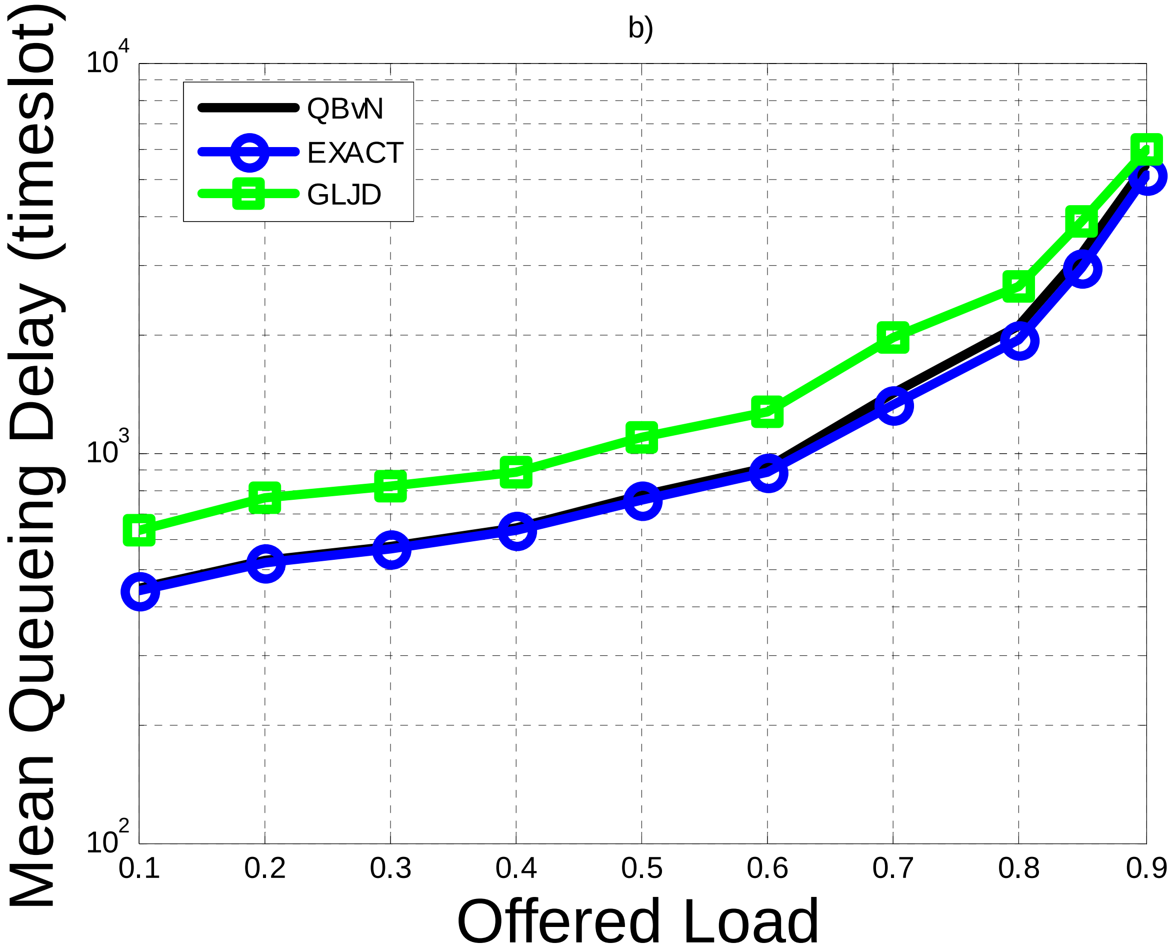
- Compared with Max-min fairness method [7]
- The service matrices obtained with this Alternating Projection method have very high measures of similarity to the original traffic matrix, with an average similarity greater than 95% for N>=32.

Service matrix construction: queuing delay



- Delay performance
  - Long-haul scenario, N=16, 1000km
  - Tested under self-similar traffic
  - Compared with
    - Max-min fairness method [7]
    - Simple rescaling method [8]
- **Conclusion**: performs better than the max-min fairness method or the simple rescaling method.

Heuristic decomposition algorithm



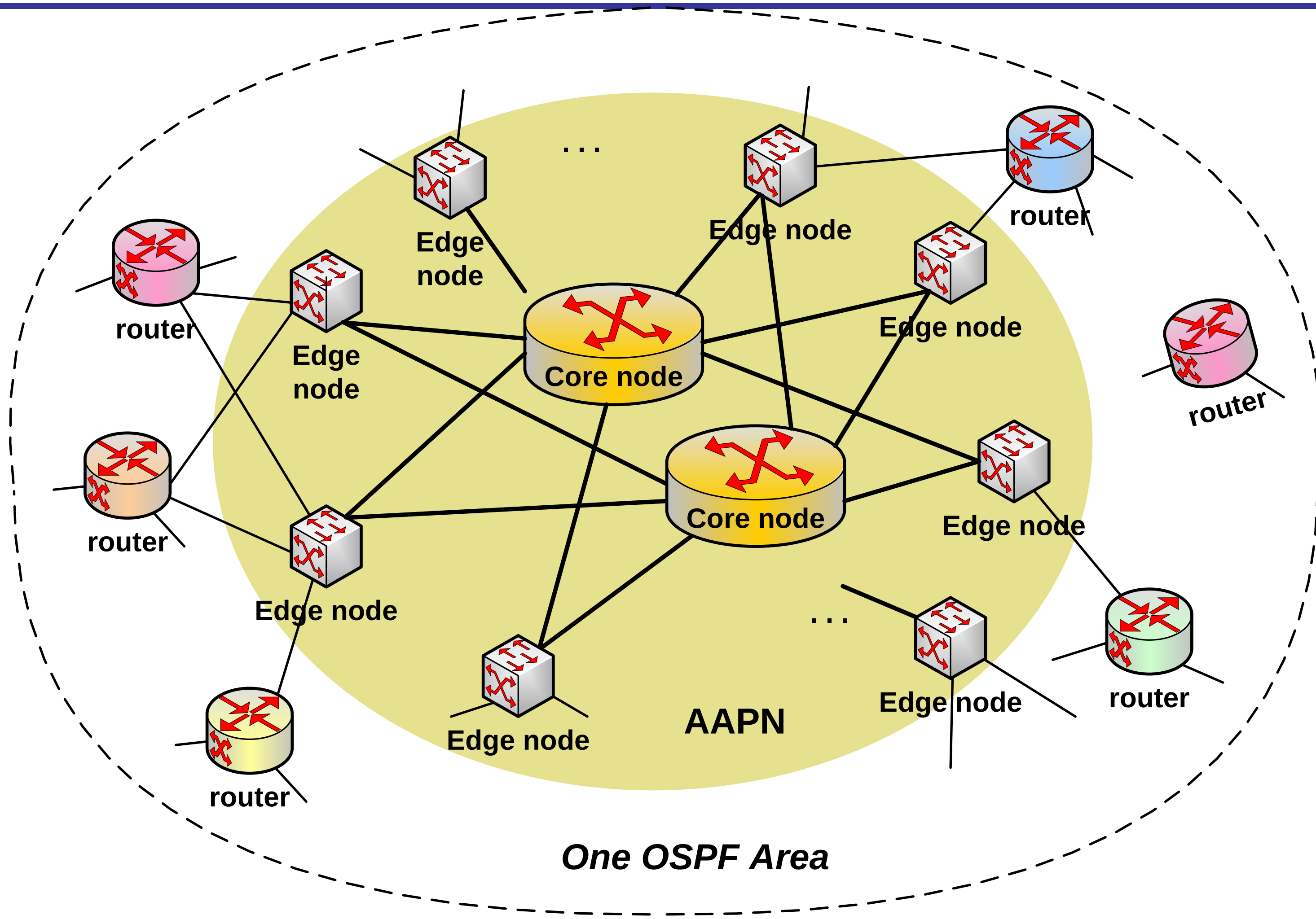
- **New algorithm:** Quick BvN (QBvN)
  - Long-haul scenario, N=16, 1000km
  - Tested under self-similar traffic
  - Compared with
    - **Benchmark**: Exact BvN
    - Greedy Low Jitter Decomposition (GLJD) [11]
- **Conclusions:**
  - excellent performance (close to optimum)
  - Low complexity O(NF)

Overview

- Overview of the AAPN project
- Frame-by-frame bandwidth allocation
- MPLS over AAPN

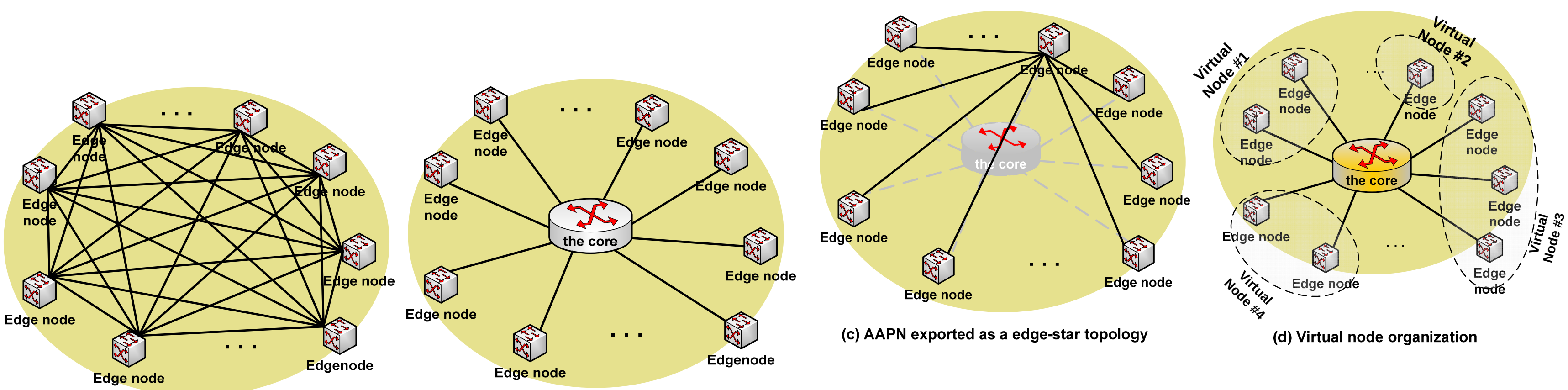
  (work by my PhD student Peng He)
- A demonstration prototype
- Conclusions

IP/MPLS routing over an AAPN



Solving the scalability problem

- Applying OSPF in a straightforward manner over an AAPN:
  - AAPN with N edge nodes corresponds to N x N links in the routing table (1 000 000 links in case of 1000 edge nodes)
- Different approaches to introduce abstraction (we speak about "virtual routers")



(a) AAPN exported as a full-meshed topology

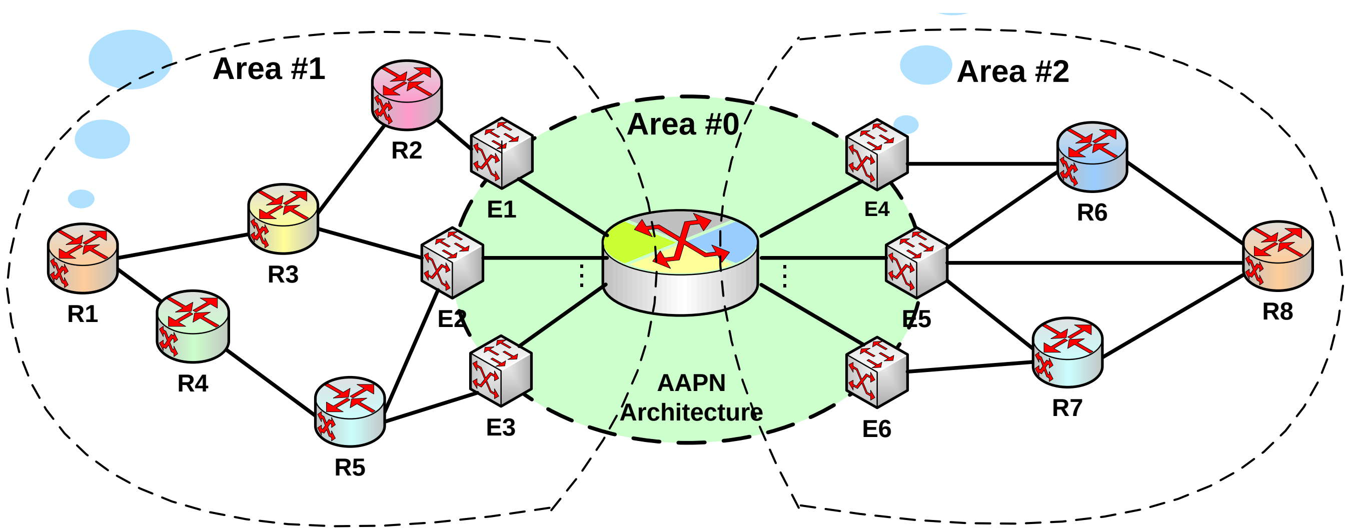
(b) AAPN exported as a core-star topology

(c) AAPN exported as a edge-star topology

(d) Virtual node organization

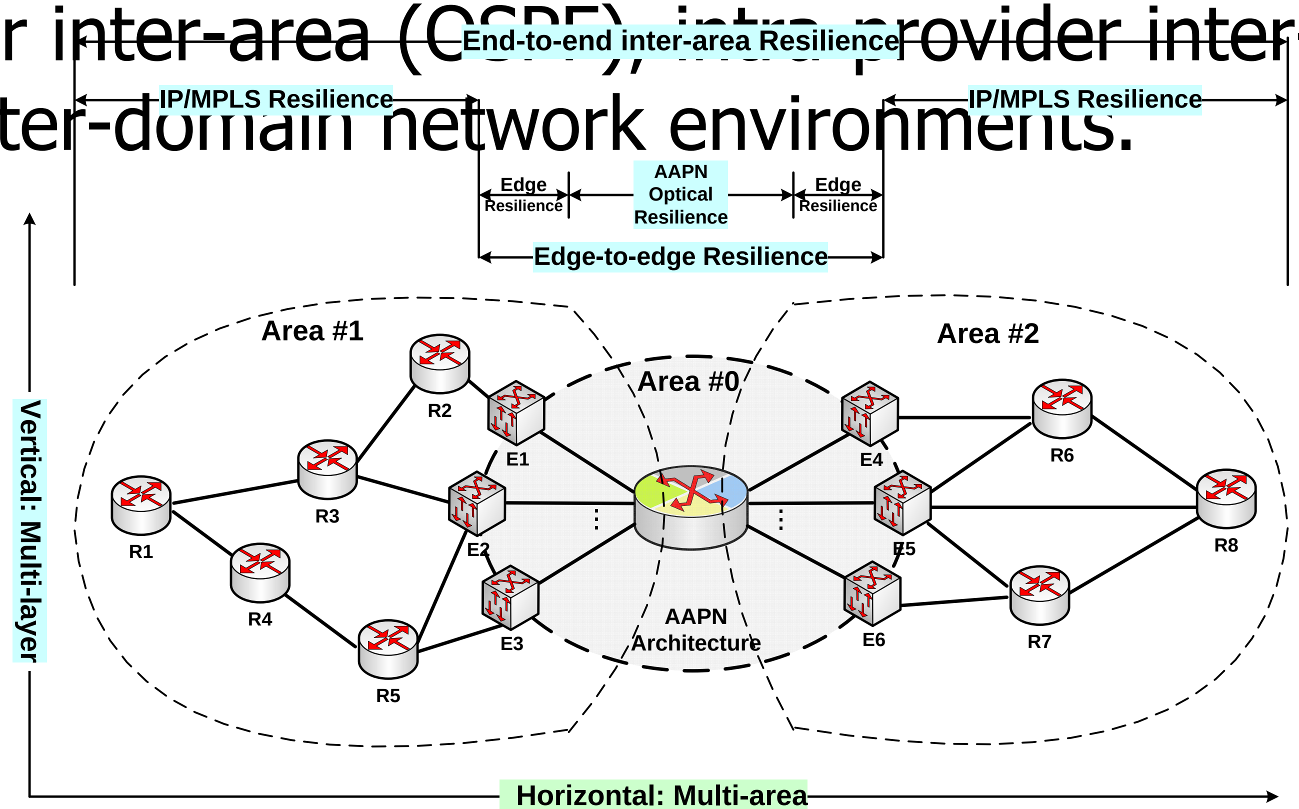
Result: Optimal OSPF inter-area routes

- OSPF can support multiple routing areas with only local routing information
- Finding optimal inter-area routes is an open research problem
- This can be solved when the OSPF areas are interconnected by an AAPN: MPLS over AAPN Traffic Engineering Framework



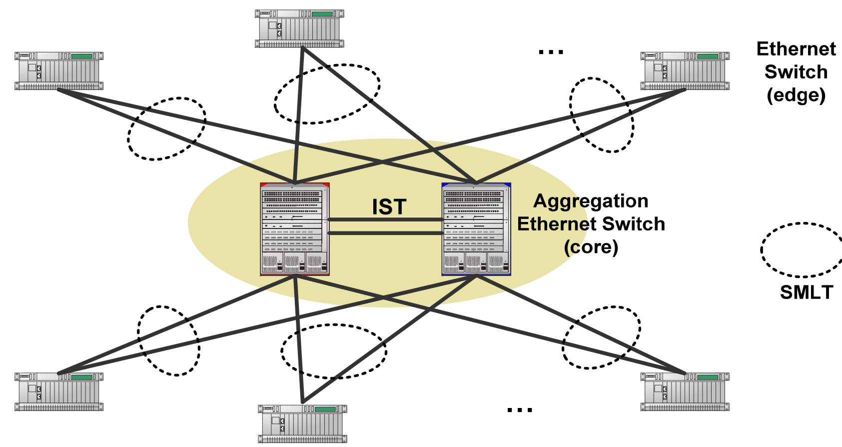
Multi-layer resilience of MPLS flows over AAPN

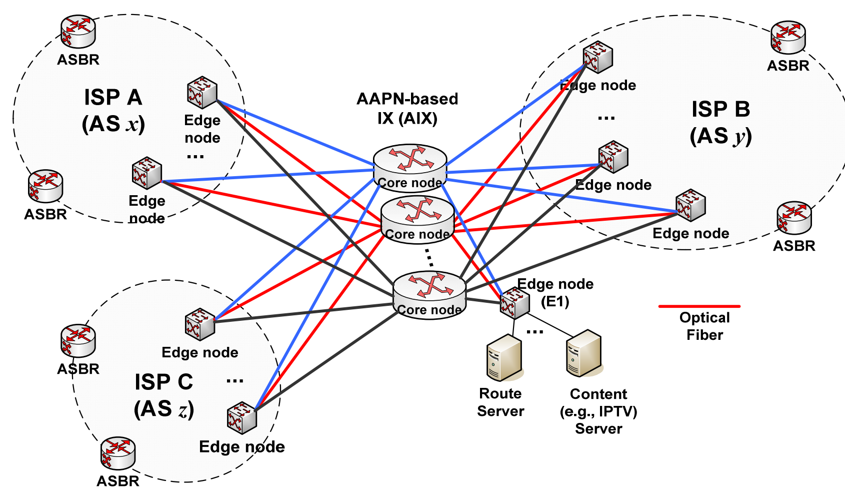
- Using our traffic engineering framework as a starting point, we are working on multi-layer resilience of MPLS flows over AAPN, especially for inter-area (OSPF), intra-provider inter-AS and inter-domain network environments.



Internet Exchange architecture with TE

- The principle of Virtual Routers can be applied to star-like networks in other situations:
  - BGP instead of OSPF
  - Star-like Ethernet instead of AAPN
- Examples:
  - existing Internet Exchange based on Ethernet
  - Internet Exchange using an AAPN





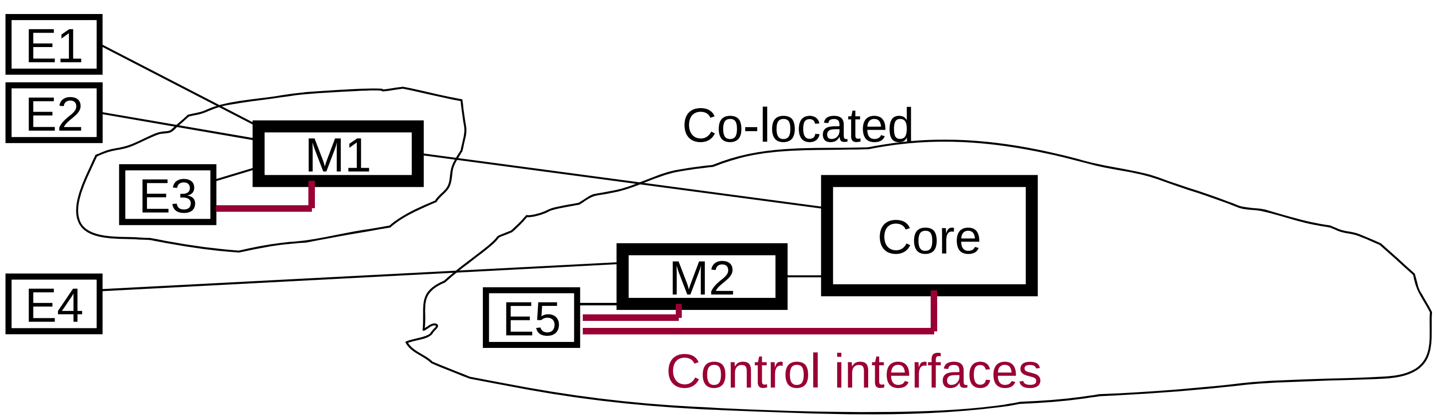
Overview

- Overview of the AAPN project
- Frame-by-frame bandwidth allocation
- MPLS over AAPN
- A demonstration prototype

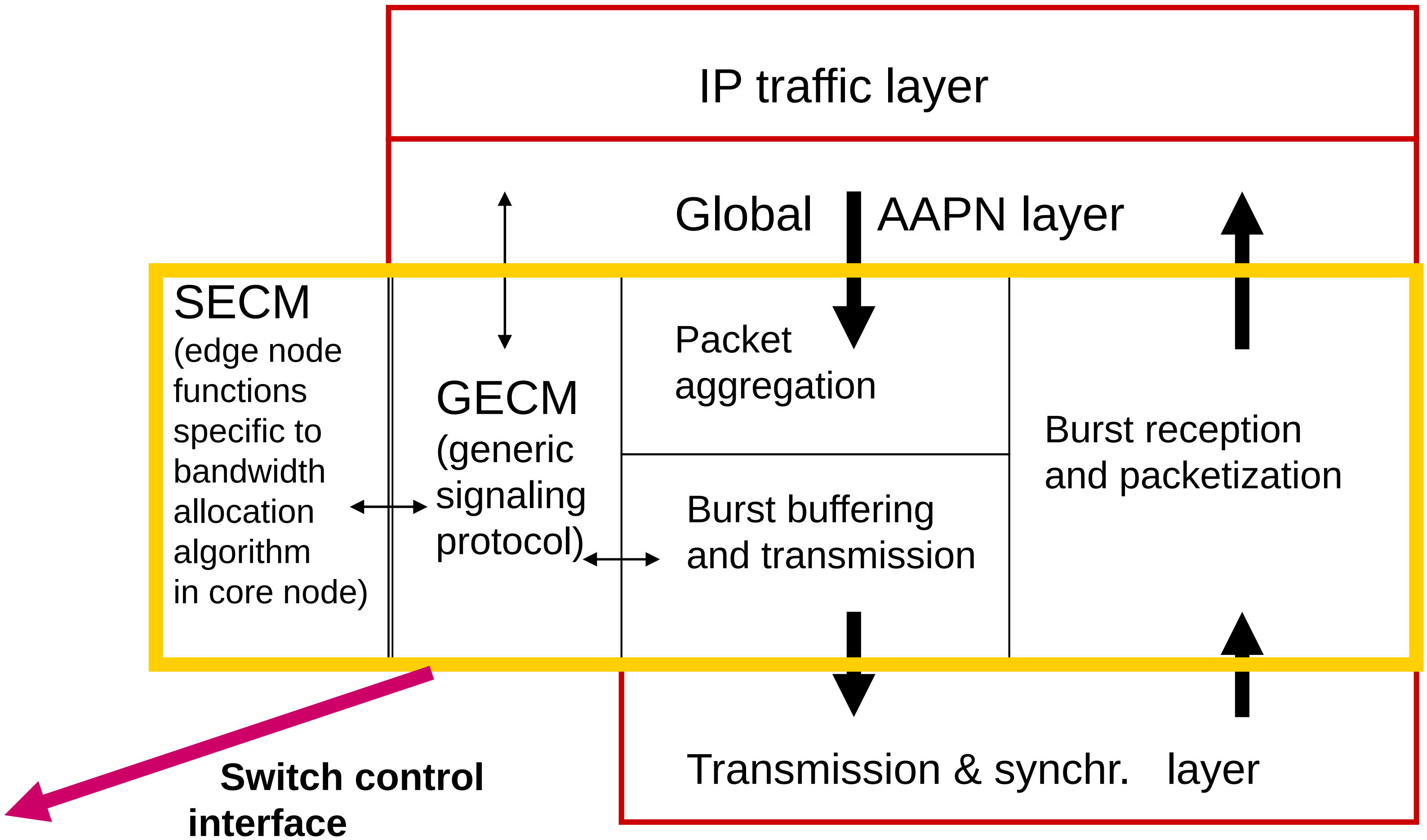
  (in collaboration with the whole Theme-1 team)
- Conclusions

Building an AAPN Control Platform

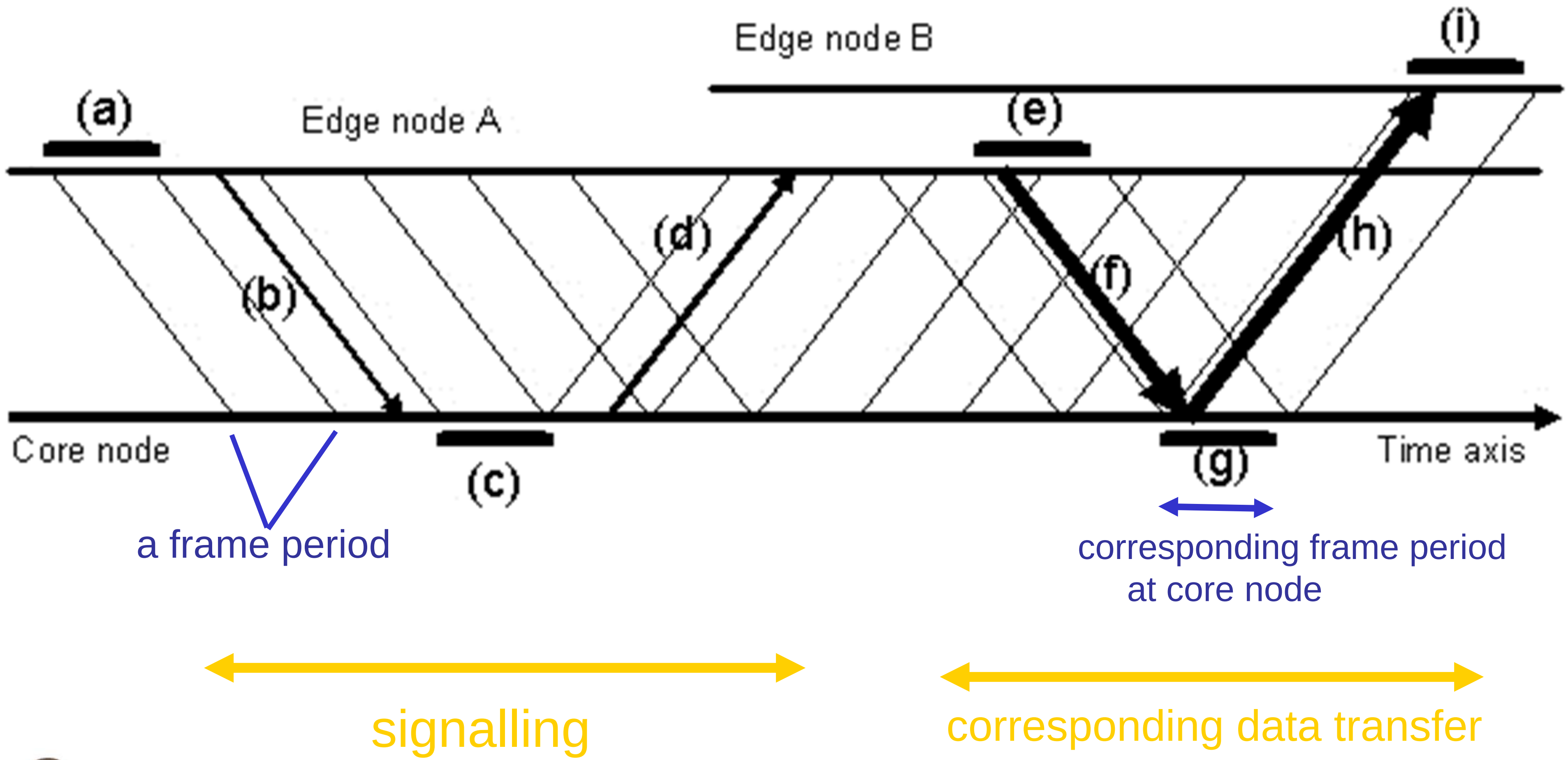
- Realizes algorithms and protocols for controlling an AAPN
  - Easily adapting to the control interfaces of different core switches
  - Easily integrating various control algorithms and protocols developed by different AAPN Theme-1 researchers
  - Initial version may be running at slower speed (using standard PCs with optical Ethernet cards)
  - Control information is exchanged through normal data blocks (no separate control channel)



Node architecture



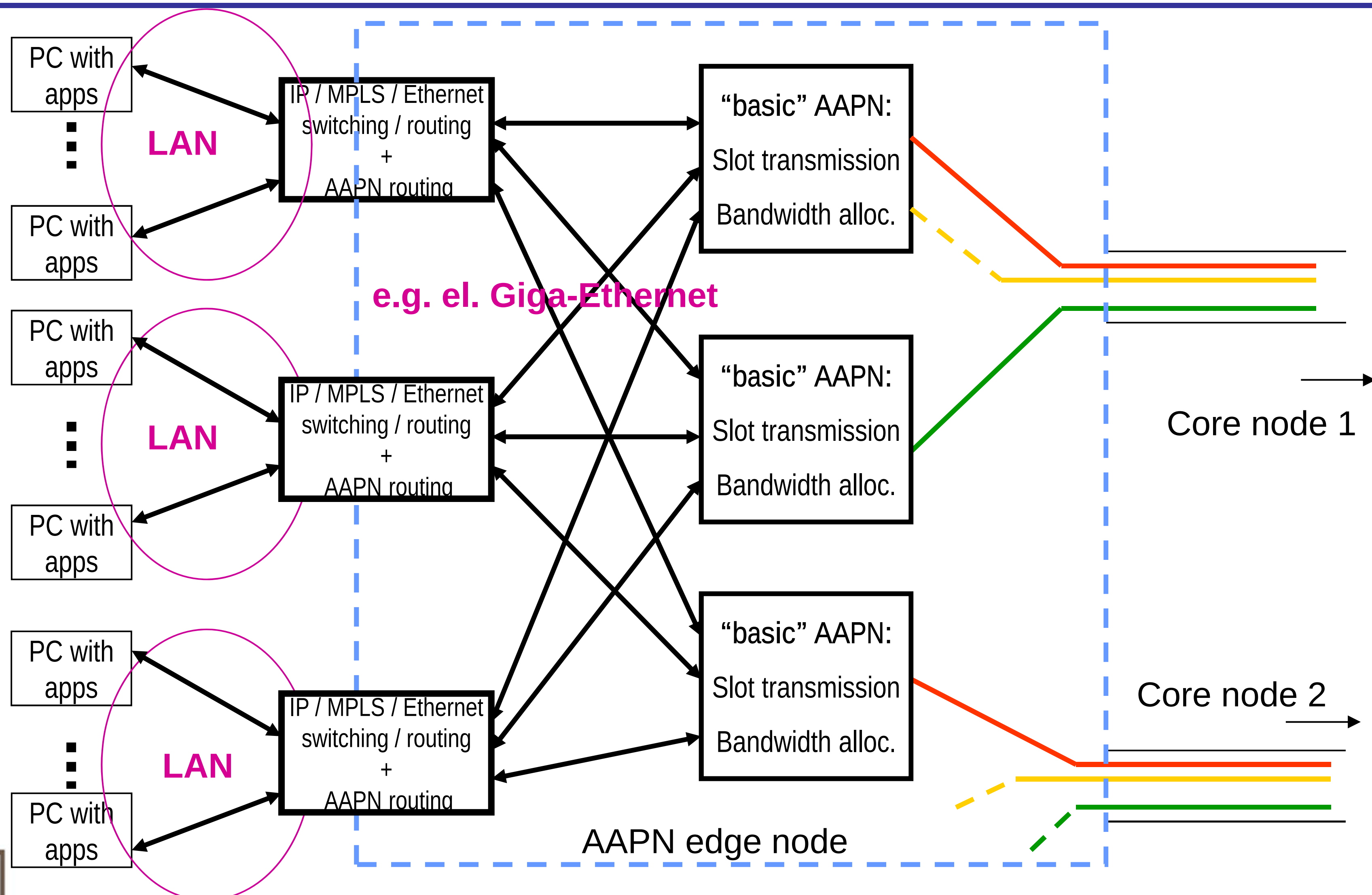
Example timing diagram



- using essentially the same control software

- Slow prototype
  - Control Platform runs on PCs, optical Ethernet transmission, MEM switch (slot time = 1 second)
    - Completed in January (used to test control platform software)
- Intermediate prototype
  - Most of Control Platform runs on PCs, optical transmission at 1 Gbps controlled by FPGAs, nano-seconds optical switch, slot time = 0.2 milliseconds
    - Expected to be completed in July
  - Improvements and extensions
    - Add basic MPLS transmission in the IP Traffic Layer to allow experimentation with real applications
    - MPLS routing and traffic engineering (TE)
    - Experiment with bandwidth allocation algorithms proposed by AAPN researchers
    - Other AAPN architectures (e.g. concentrators)

Component-based edge node architecture



Conclusions

- The AAPN is a fascinating research project
- Premise: very fast switching and simple network architecture (overlaid stars)
- Theme-1 Outcomes:
  - Proposal for slotted burst switching with bandwidth reservation (TDM, various allocation schemes)
  - Internet-integration : AAPN as an OSPF area
- Theme-2 Outcomes:
  - Various technologies for very fast switching, amplification, RRR, etc.
- Demonstration prototypes