Connect. Communicate. Collaborate

GN2-JRA3 efforts in the control plane area

ビクター ライス , Victor Reijs, HEAnet
victor.reijs@heanet.ie

September 11th, 2006

GLIF, Tokyo

Connect. Communicate. Collaborate

Outline

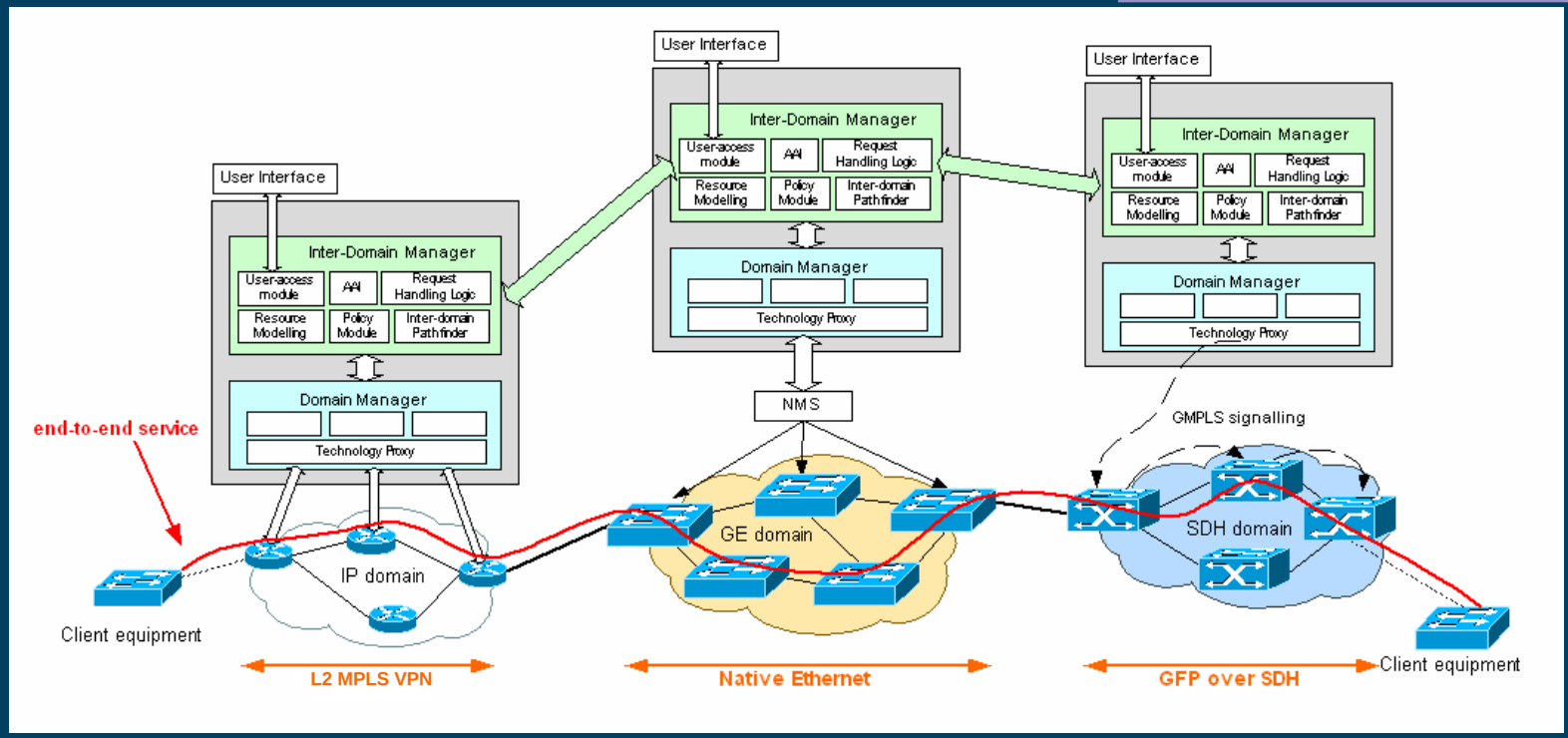
- GN2-JRA3 activity…
- IDM multi-domain issues…
- Stitching framework for multi technology domains…
- Multi-domain pathfinder…
- Questions…
- Appendix…

GN2-JRA3 activity

- A ‘Joint Research Activity’ investigating the piloting of ‘Bandwidth on Demand’ services to the NREN community
- The environment:
  - Multi administrative domains
  - Multiple technology domains, e.g.
    - GFP over SDH, L2 MPLS VLL, Native Ethernet
  - Multi aggregated (e.g. GMPLS/UCLP) and provisioning (e.g. human/NMS/BLUEnet/DRAC) domains
  - Requirements for:
    - end-to-end non-contended deterministic capacity
    - a standardized interface for service requests by clients
    - service level indication to clients
    - advance reservation (scheduled)

GN2-JRA3 architecture

Connect. Communicate. Collaborate



- Inter-Domain Manager (IDM) and Domain Manager (DM)
- Each domain participating in BoD service provisioning needs to operate an IDM and honor the IDM-DM and IDM-IDM defined interfaces
- GN2-JRA3 will provide:
  - The IDM module
  - Reference implementation for several DMs (human NOC, Ethernet-based, etc.)

Outline

- GN2-JRA3 activity…
- IDM multi administrative domain issues…
- Stitching framework of multi technology domains…
- Multi-domain pathfinder…
- Questions…
- Appendix…

IDM multi-domain issues

- Domain independence on resource usage policies and technological choices
- A service and network abstraction schema to describe implementation over very different networks
  - a schema which allows to clearly specify which type of service is requested
  - a network abstraction which allows inter-domain information exchange independently of the underlying technologies
- Stitching framework for multi technology domains…
- Multi-domain pathfinding procedure…
- In-advance reservation (scheduling)
- Monitoring of Layer 1/2 technology domains
- Authentication and authorization

Connect. Communicate. Collaborate

Outline

- GN2-JRA3 activity…
- IDM multi-domain issues…
- Stitching framework for multi technology domains…
- Multi-domain pathfinder…
- Questions…
- Appendix…

Stitching framework for multi technology domains

Connect. Communicate. Collaborate

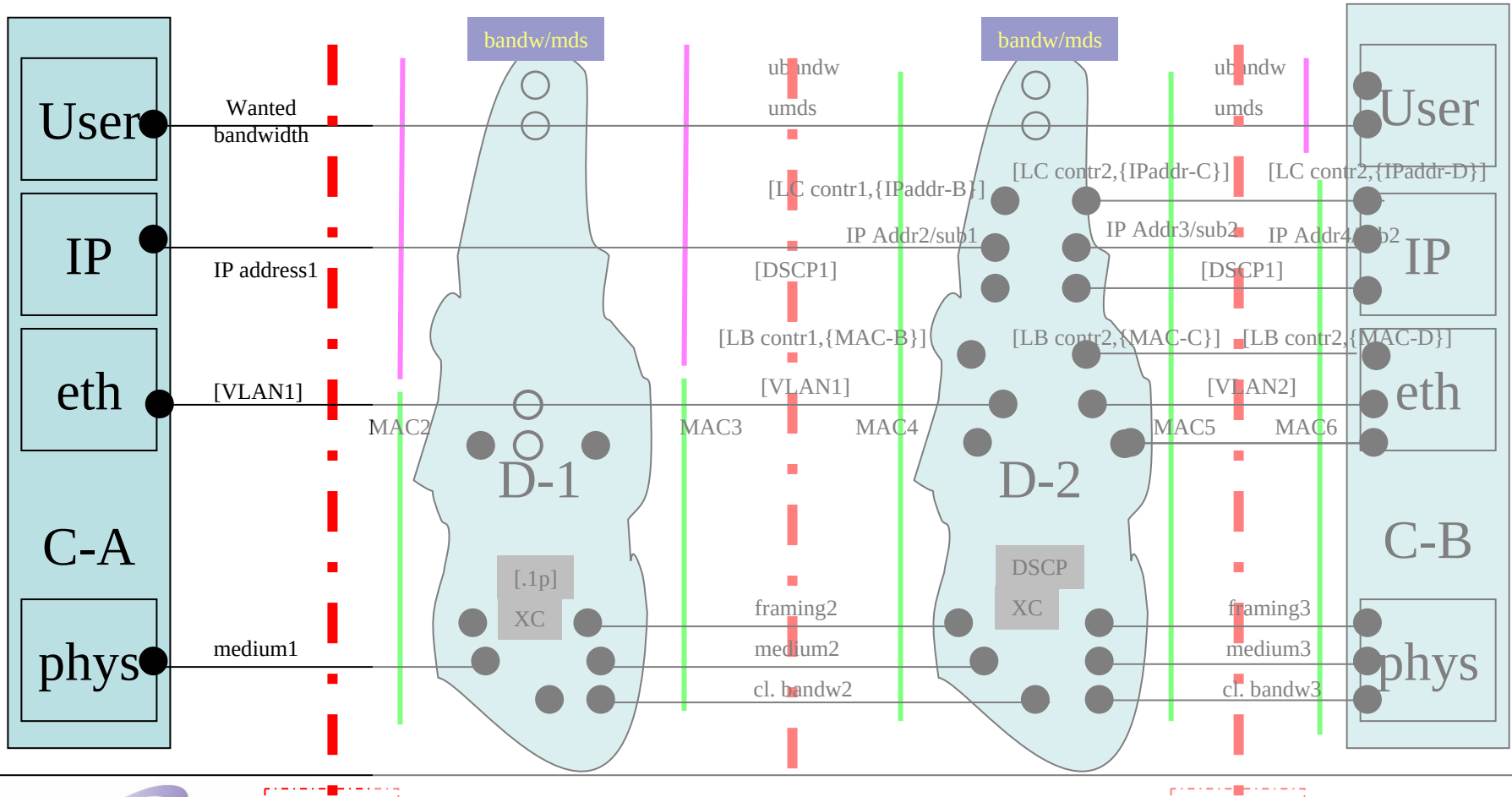
- Stitching is making sure that one can make connections over different technology
- We do stitching (adaptation layer) in networking already for years:

  host – ethernet – IP – ethernet - host
- Different layers
- Different peering points
- This presentation only tries to make the steps more explicit

Layers in stitching

- Each layer (not necessarily OSI layers) has parameters:
  - Layer D+: User (human/application): e.g. IP Port number, wanted bandwidth, wanted max. datagram size, latency, reservation, etc.
  - Layer C (~routing):
    IP layer: e.g. IP address and subnet, DSCP
  - Layer B (~switching):
    L2 MPLS VLL: e.g. VC_ID#
    Ethernet: e.g. VLAN#, MAC address
    SDH/SONET: e.g. VCG#, VC-3/VC-4
  - Layer A (~physical):
    interface: e.g. framing, medium (lambda), clock speed

Ethernet – IP

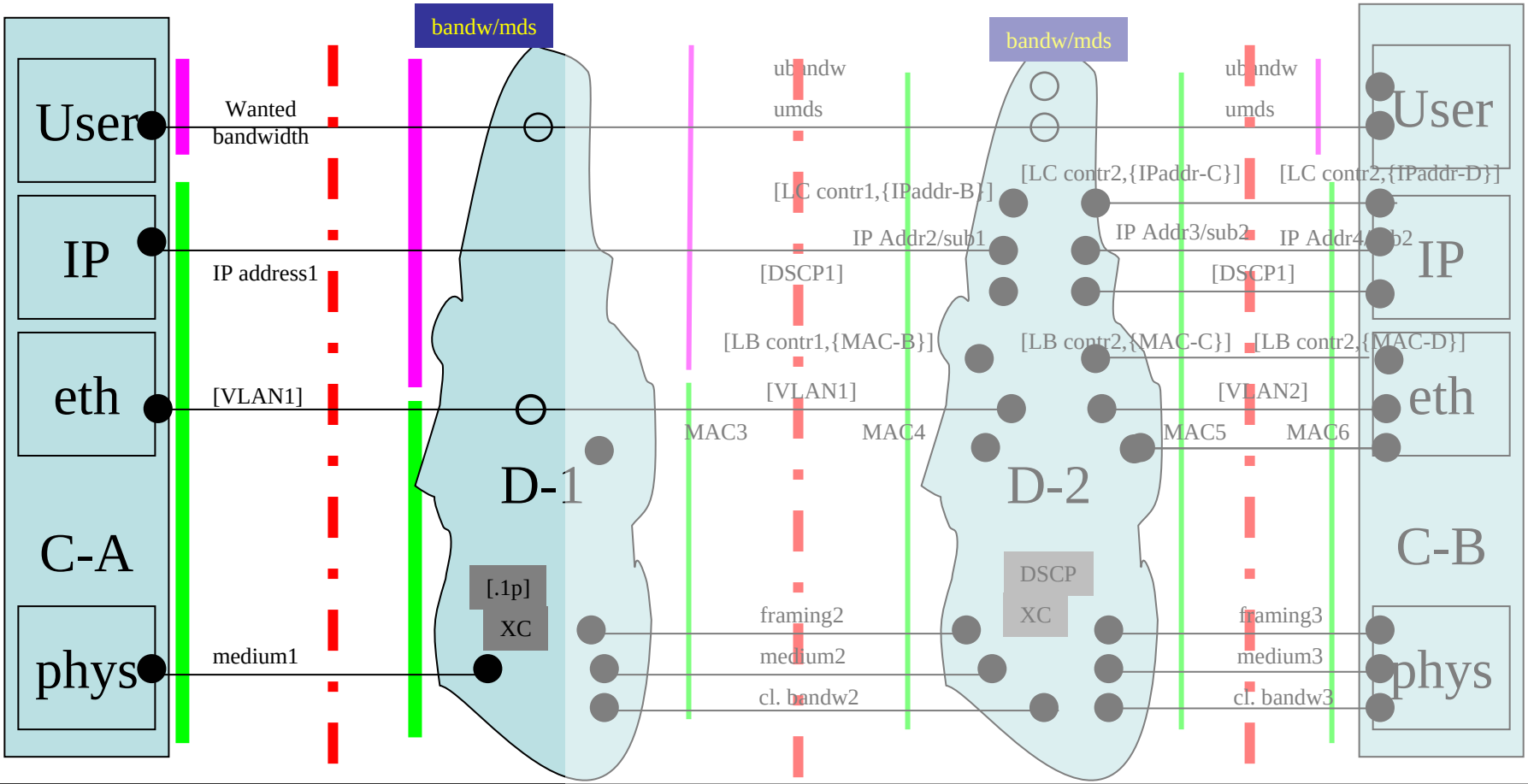

User
IP
eth
C-A
phys
Wanted bandwidth
IP address1
[VLAN1]
medium1
bandw/mds
bandw/mds
ub ndw
umds
ub ndw
umds
[LC contr1,{IPaddr-B}]
IP Addr2/sub1
[DSCP1]
[LB contr1,{MAC-B}]
[VLAN1]
[LC contr2,{IPaddr-C}]
IP Addr3/sub2
[DSCP1]
[LB contr2,{MAC-C}]
[VLAN2]
[LC contr2,{IPaddr-D}]
IP Addr4/sub2
[LB contr2,{MAC-D}]
MAC2
MAC3
MAC4
MAC5
MAC6
D-1
D-2
[.1p]
XC
DSCP
XC
framing2
medium2
cl. bandw2
framing3
medium3
cl. bandw3
User
IP
eth
C-B
phys
CC1
CC2
CC3

Communication channel between domains (1/3)

Connect. Communicate. Collaborate

- There are two parts to the comm. channel:
  - Configuration part
    Handled by e.g. DM, CLI, provisioning system
    These are parameters that are directly being configured in the technology domain.
  - Stitching part
    Handled by e.g. inter-domain routing protocol, IDM, phone, e-mail
    The stitching part is for parameters that are not directly configured

Ethernet – IP


User
IP
eth
C-A
phys
Wanted bandwidth
IP address1
[VLAN1]
medium1
bandw/mds
D-1
[.1p]
XC
ub. bandw
umds
[LC contr1,{IPaddr-B}]
IP Addr2/sub1
[DSCP1]
[LB contr1,{MAC-B}]
[VLAN1]
MAC3
framing2
medium2
cl. bandw2
bandw/mds
D-2
DSCP
XC
MAC4
ub. bandw
umds
[LC contr2,{IPaddr-C}]
IP Addr3/sub2
[DSCP1]
[LB contr2,{MAC-C}]
[VLAN2]
MAC5
framing3
medium3
cl. bandw3
[LC contr2,{IPaddr-D}]
IP Addr4/sub2
[LB contr2,{MAC-D}]
MAC6
User
IP
eth
C-B
phys
CC1
CC2
CC3

Communication channel between domains (2/3)

Connect. Communicate. Collaborate

- In principle there is only one comm. channel between the each domain (e.g. D-1) incl the client domain (e.g. C-A)
- The communication can be between different BoD system parts; IDM, DM, humans or technology/provisioning domain
- Also it must be defined how to derive essential technology domain specific parameters related to switching (grey background) and performance (blue background).

Communication channel between domains (3/3)

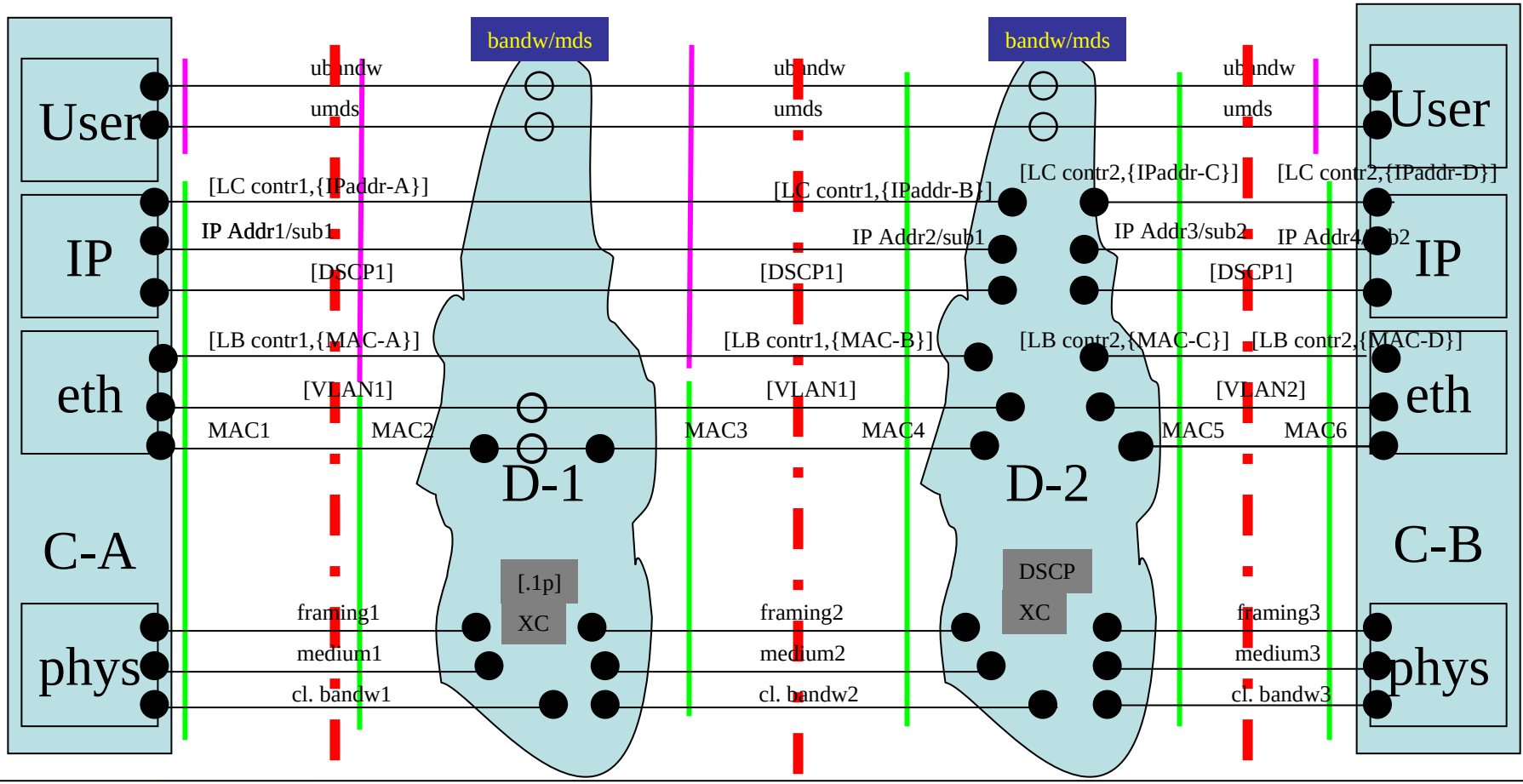
Connect. Communicate. Collaborate

- The client domain can be a network, but also a host.
- Parameters related to Layer A/B/C ‘control’ parameters:
  - Layer A system:
    - Layer A ‘control’ parameters are e.g. speed and duplex negotiations
  - Layer B/C host:
    - normally one IP address and one MAC address (but can have multiple ones and could use autoconfig)
    - normally static routing (but can be other)
  - Layer B/C network:
    - a range of addresses (MAC and/or IP): {MAC-A} or {IPaddr-A}
    - a routing/control protocol (like static, OSPF, spanning tree, etc.)

How parameters are treated

- All parameters start and stop at a certain domain (black filled dot: ●): peering points
- Some parameters are used within the domain (black open dot: ○) and go to other domains: *learning* points
- These *learning* points can result in additional parameters (grey background) specific to a technology/provisioning domain (minimal it is making a connection (XC/LSP))
- Some parameters are not used in a specific domain, but need to be transported by IDM protocol (straight line through domain) to next domain
- Some parameters can have end-to-end significance (e.g. possibly VLAN#, etc.)
- Some parameters are important for end-to-end performance, like the maximum datagram size and the actual data bandwidth (blue background).

Ethernet – IP


bandw/mds
bandw/mds
User
IP
eth
C-A
phys
User
IP
eth
C-B
phys
ub.andw
umds
[LC contr1,{IPaddr-A}]
IP Addr1/sub1
[DSCP1]
[LB contr1,{MAC-A}]
[VLAN1]
MAC1
MAC2
framing1
medium1
cl. bandw1
ub.andw
umds
[LC contr1,{IPaddr-B}]
IP Addr2/sub1
[DSCP1]
[LB contr1,{MAC-B}]
[VLAN1]
MAC3
MAC4
framing2
medium2
cl. bandw2
ub.andw
umds
[LC contr2,{IPaddr-C}]
IP Addr3/sub2
[DSCP1]
[LB contr2,{MAC-C}]
[VLAN2]
MAC5
MAC6
framing3
medium3
cl. bandw3
[LC contr2,{IPaddr-D}]
IP Addr4/sub2
[LB contr2,{MAC-D}]
D-1
D-2
[.1p]
XC
DSCP
XC
CC1
CC2
CC3

Other dependencies

- Some things are not essential for this framework: as they are just as solve-able as when using manual procedures
- When automating the following need to be incorporated:
  - an (collapsed) abstract notation (XML schema)
  - constraint analysis (technology agnostic IDM)
  - path finding issues
  - (advance) reservations
  - IDM-IDM and IDM-DM protocol
  - implementation issues
  - incompatible technologies (LAN-WAN PHY 10 GE)
  - etc.
- These things will be defined elsewhere in GN2-JRA3, but stitching will certainly provide input to these.

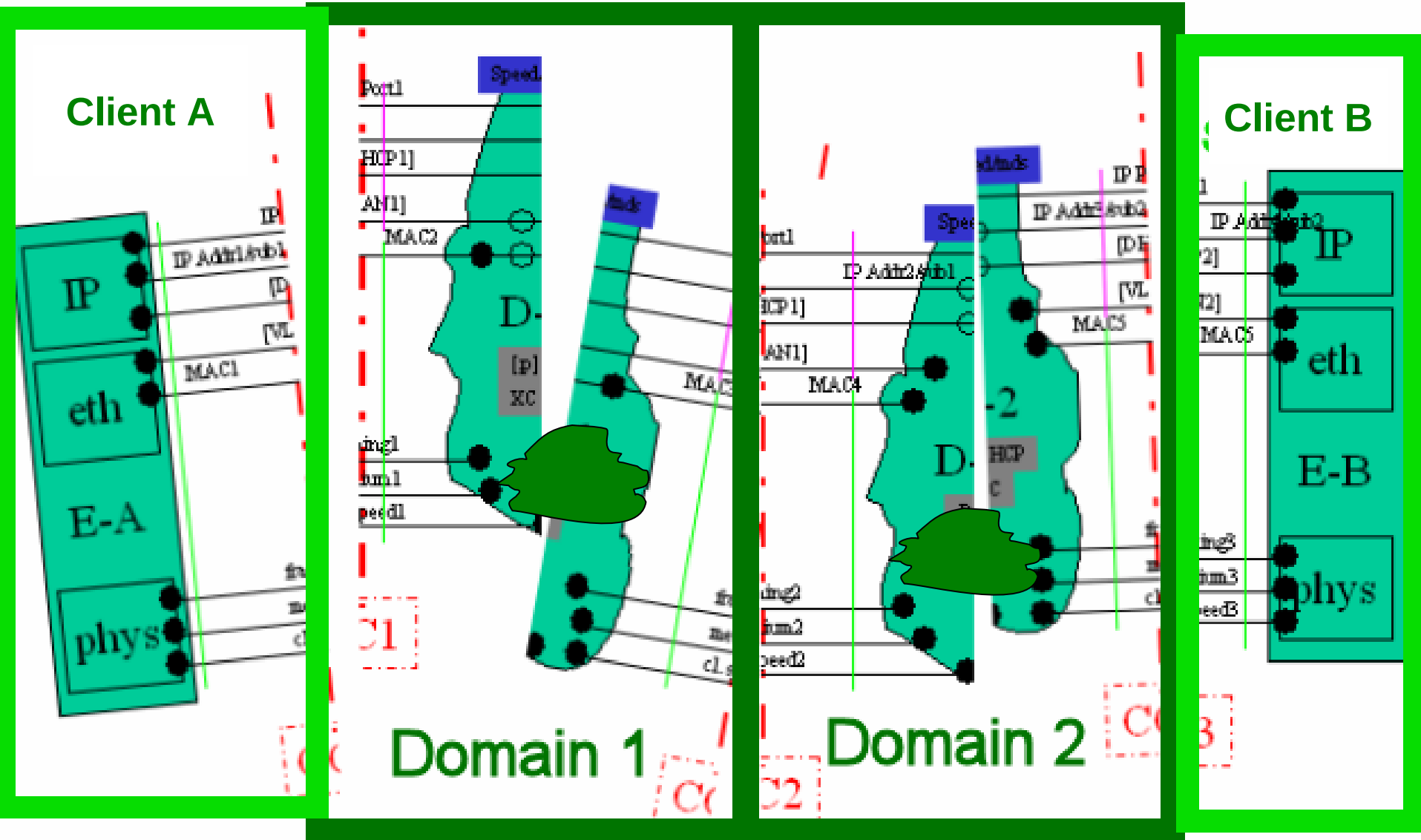
Some scenarios covered by GN2-JRA3

Connect. Communicate. Collaborate

- SDH-SDH (scenario 5a)
- SDH-SDH-SDH (scenario 5b)
- Ethernet – L2 MPLS VLL (scenario 6)
- L2 MPLS VLL – L2 MPLS VLL (scenario 8a: no VC_ID# stitching)
- L2 MPLS VLL – L2 MPLS VLL (scenario 8a: VC_ID# stitching)
- Ethernet –PIP (scenario 9)
- Ethernet –Ethernet (scenario 14a: .1Q between domains)
- Ethernet –Ethernet (scenario 14b: .1ad between domains)
- L2 MPLS VLL – OOO (scenario 15)
- A jigsaw puzzle…

Stitching jigsaw puzzle

Connect. Communicate. Collaborate



Connect. Communicate. Collaborate

Outline

- GN2-JRA3 activity…
- IDM multi-domain issues…
- Stitching framework for multi technology domains…
- Multi-domain pathfinder…
- Questions…
- Appendix…

Multi-domain pathfinder (1/3)

The main task of the Pathfinder sub-module in the IDM is to provide a list of inter-domain paths able to fulfill a BoD request

The main components contribute to the BoD request realisation:

- the BoD request
- between domains announcing interdomain links and their (collapsed) topologies. It allows each domain to apply a policy in link announcing and to build its view of the BoD domains network.
- a “pruning” algorithm which, for each BoD request, simplifies the abstract network topology removing non relevant components. As an example, removing all links with capacity lower than requested.
- a pathfinding (routing) engine based on a constraint based shortest path algorithm
- a reservation module that will check the actual provisioning

*figure adopted from DRAGON project material

Multi-domain pathfinder (2/3)

- The pathfinder operates on collapsed abstract topology view of each domain. Current assumption is that each domain is collapsed to a point like structure with only interdomain links. The full topology is known only internally.
- Collapse stages:

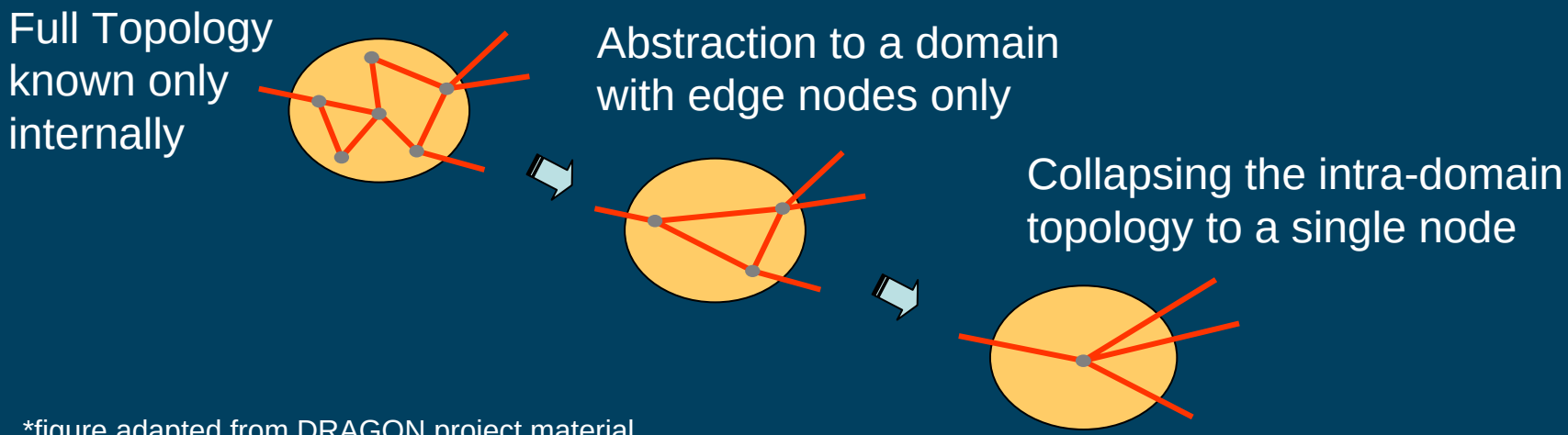


*figure adapted from DRAGON project material

- The pathfinding protocol used is OSPF with Traffic Engineering standard extension (Quagga), enriched by GN2-JRA3 specific constraints.

Multi domain pathfinder (3/3)

- For the IDM Phase 1, the Quagga OSPFv2 routing daemon implementation with custom defined Opaque LSAs is used as the pathfinding protocol engine and signaling sub-module
  - As the Quagga OSPFv2 daemon is a SPF (Shortest Path First) engine and not a constraints-based SPF engine, the Pathfinder module is required to perform additional CSPF computations
- Based on TE information for the advertised topology, the Pathfinder sub-module applies a constraint-based algorithm to create a list of paths to be handed back to the Reservation module
  - Each path in the list represents an inter-domain route over a set of interconnected domains, and includes the ingress and egress interface in each transit domain

Connect. Communicate. Collaborate

Outline

- GN2-JRA3 activity…
- IDM multi-domain issues…
- Stitching framework for multi technology domains…
- Multi-domain pathfinder…
- Questions…
- Appendix…

Time for questions?

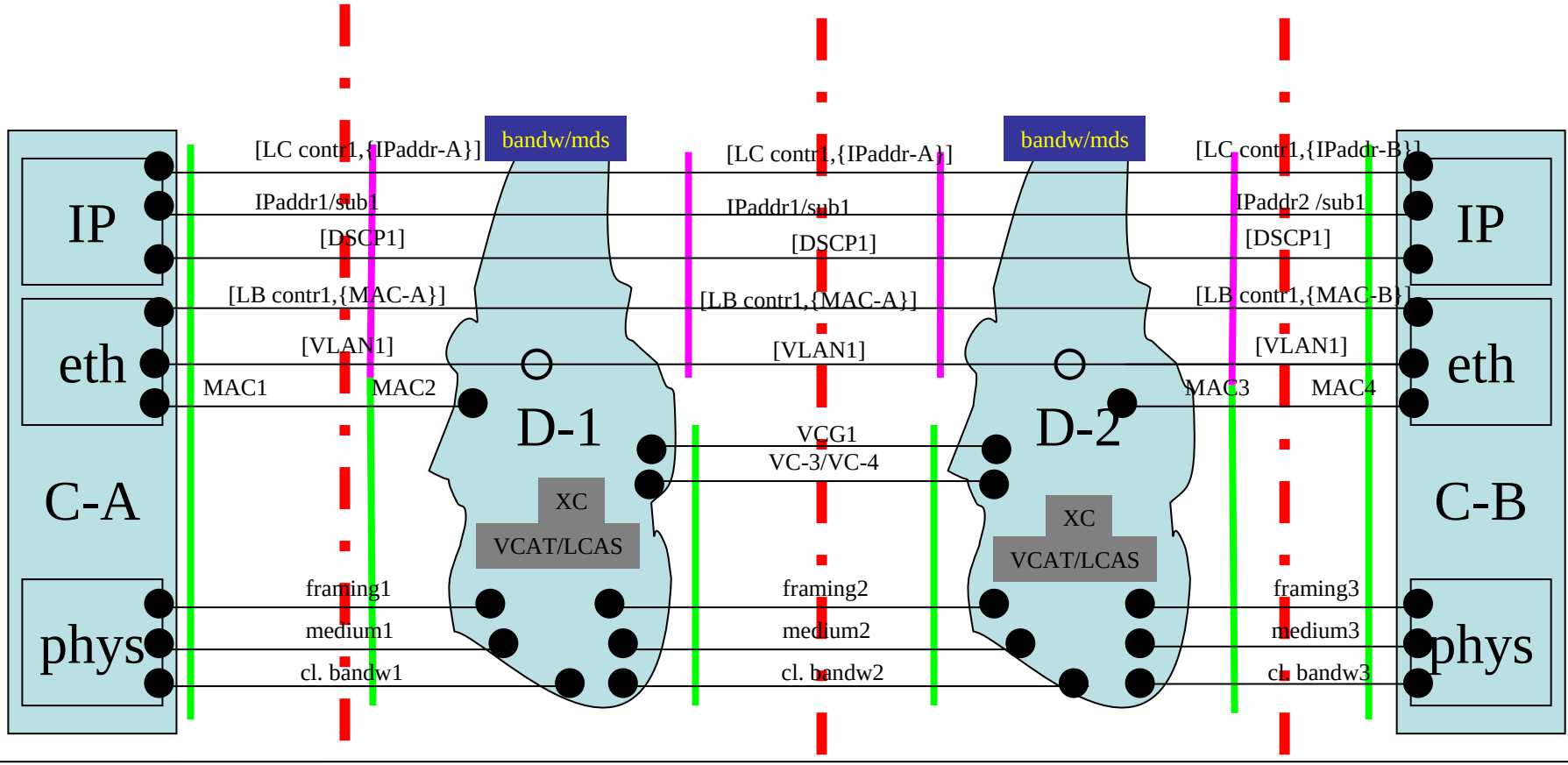
- How to evaluate technology constraints in a technology agnostic way at the IDM level?
- How to get optimum between pruning due to constraints and checking at reservation?
- Hoe much information can be collapsed abstract topology hold?
- Piloting the stitching template and its scenarios out on the Louisiana-Brno link

Connect. Communicate. Collaborate

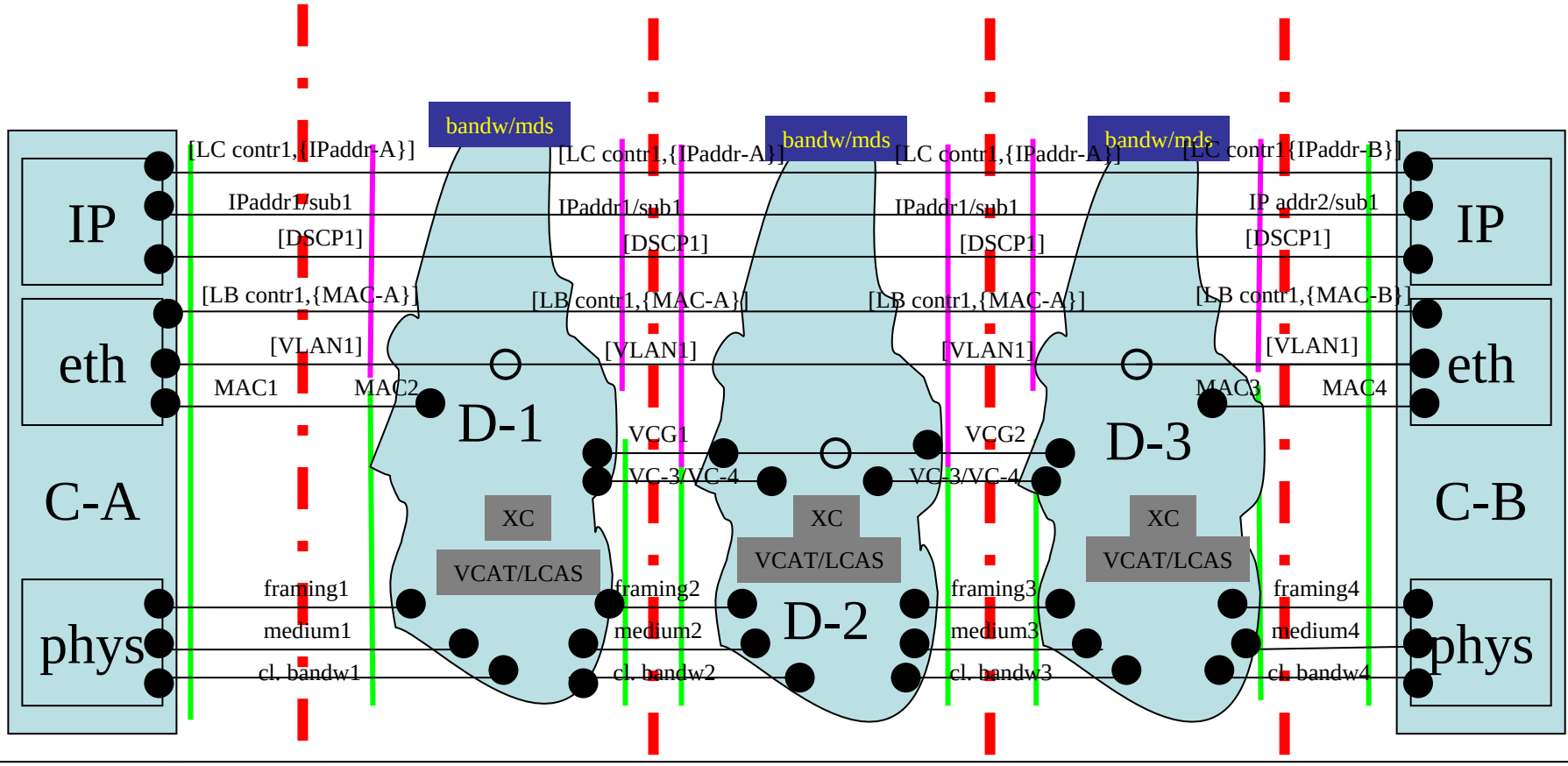
Outline

- GN2-JRA3 activity…
- IDM multi-domain issues…
- Stitching framework for multi technology domains…
- Multi-domain pathfinder…
- Questions…
- Appendix…

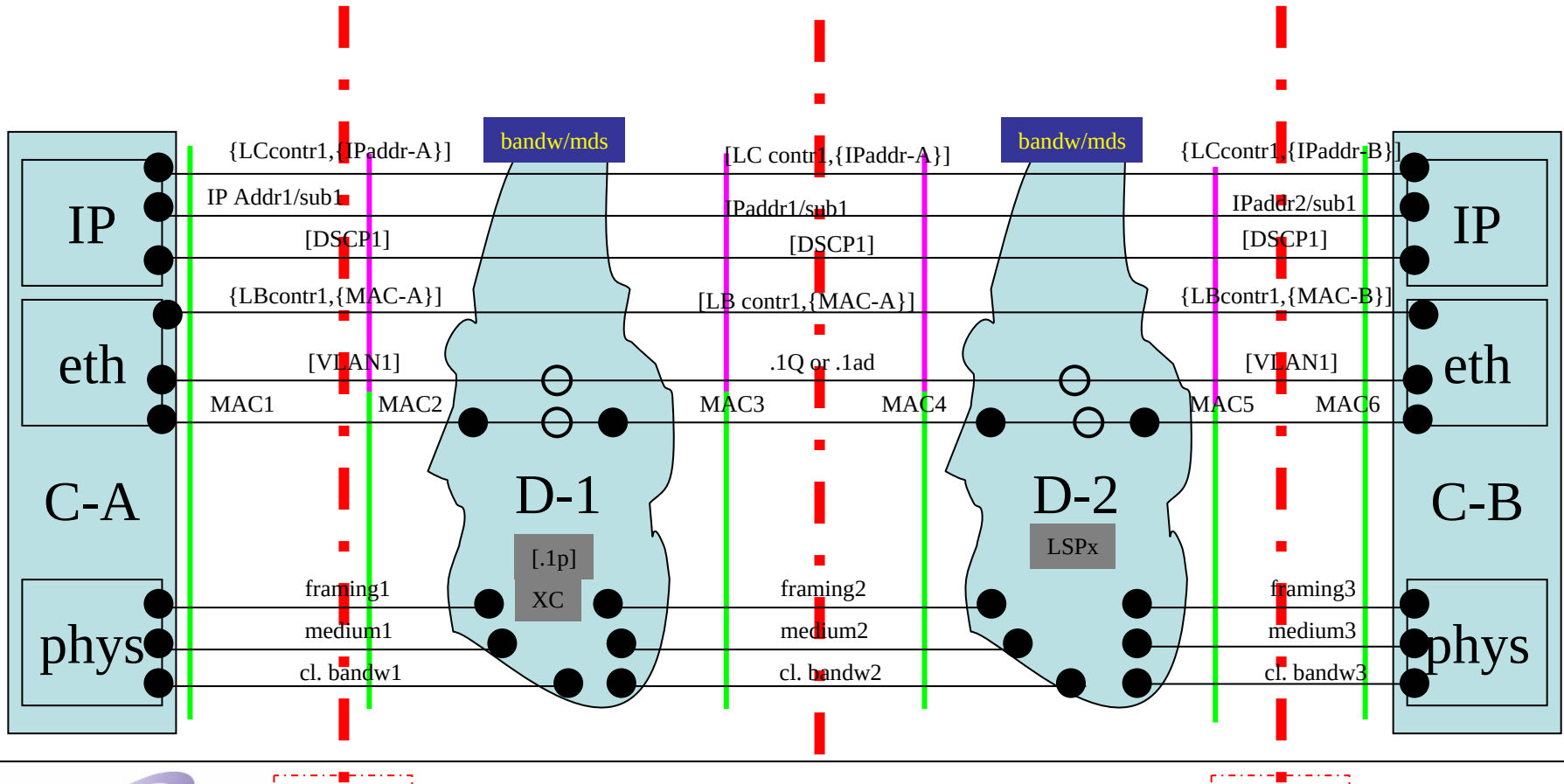
SDH – SDH (sc5a)


IP
eth
C-A
phys
[LC contr1,{IPaddr-A}]
IPaddr1/sub1
[DSCP1]
[LB contr1,{MAC-A}]
[VLAN1]
MAC1
MAC2
framing1
medium1
cl. bandw1
bandw/mds
D-1
XC
VCAT/LCAS
[LC contr1,{IPaddr-A}]
IPaddr1/sub1
[DSCP1]
[LB contr1,{MAC-A}]
[VLAN1]
VCG1
VC-3/VC-4
framing2
medium2
cl. bandw2
bandw/mds
D-2
XC
VCAT/LCAS
[LC contr1,{IPaddr-B}]
IPaddr2 /sub1
[DSCP1]
[LB contr1,{MAC-B}]
[VLAN1]
MAC3
MAC4
framing3
medium3
cl. bandw3
IP
eth
C-B
phys
CC1
CC2
CC3

SDH – SDH – SDH (sc5b)



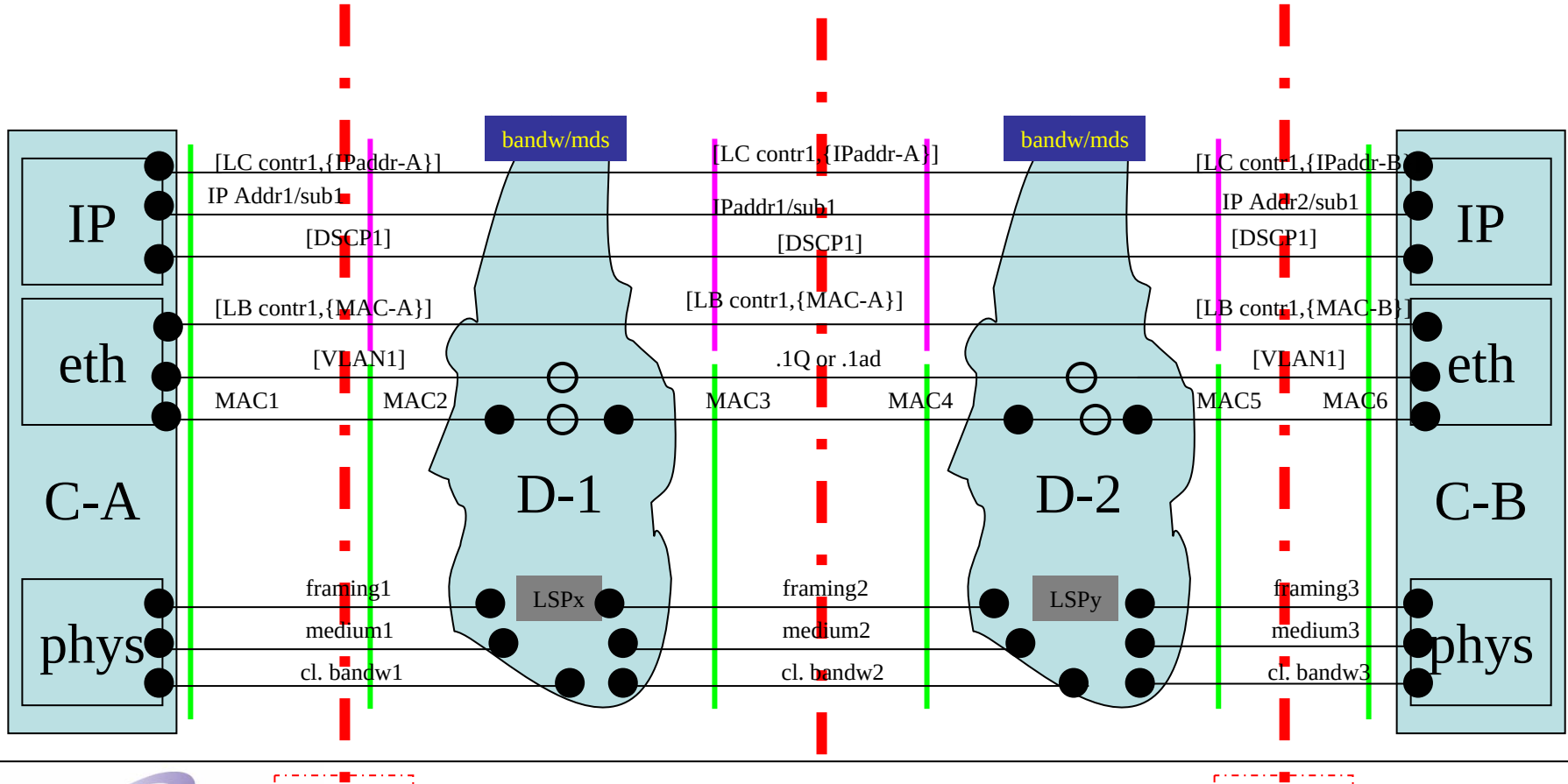
Ethernet – L2 MPLS VLL (sc6)



L2 MPLS VLL – L2 MPLS VLL (sc8a)

(no VC_ID# stitching)

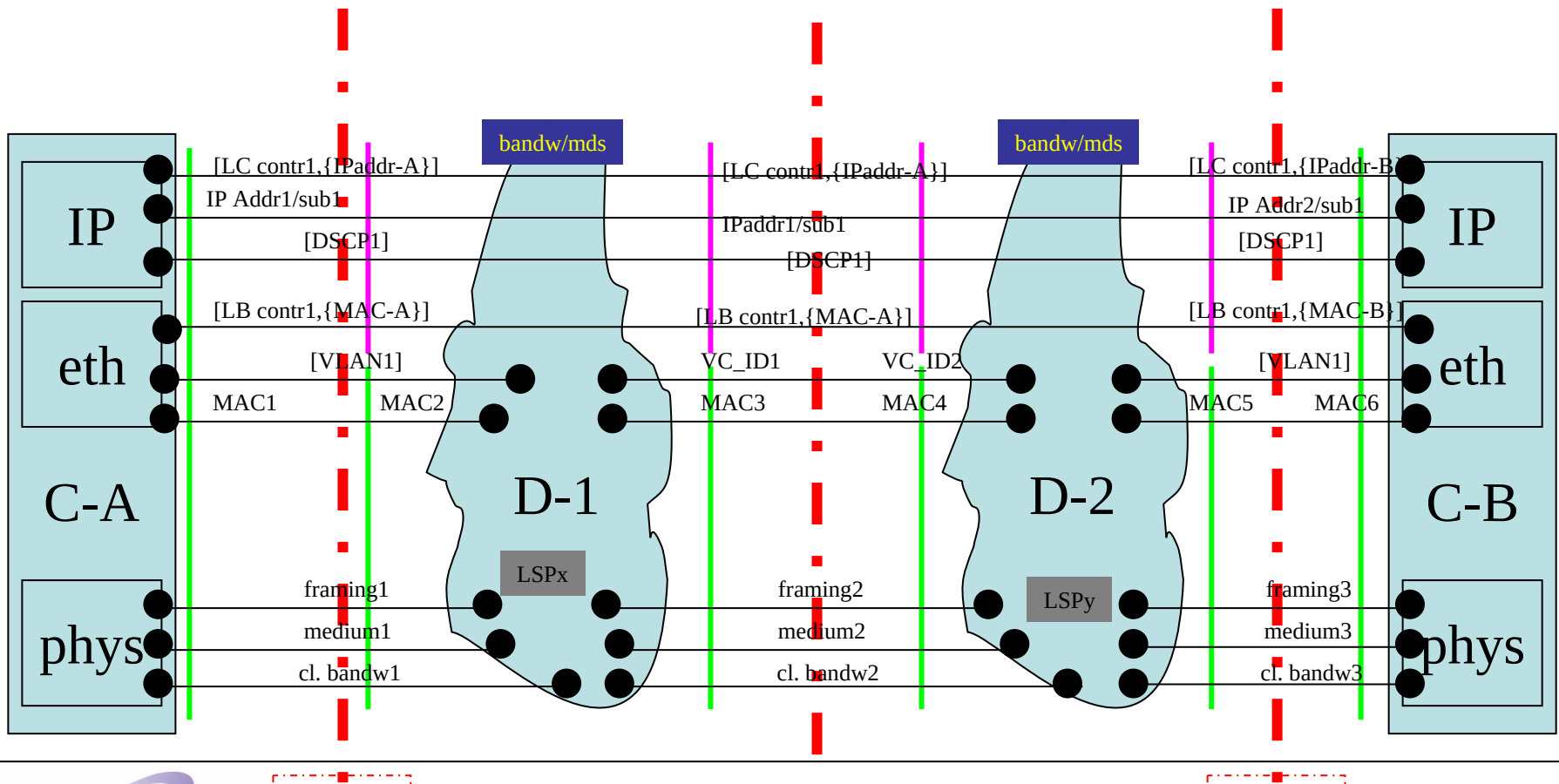
Connect. Communicate. Collaborate



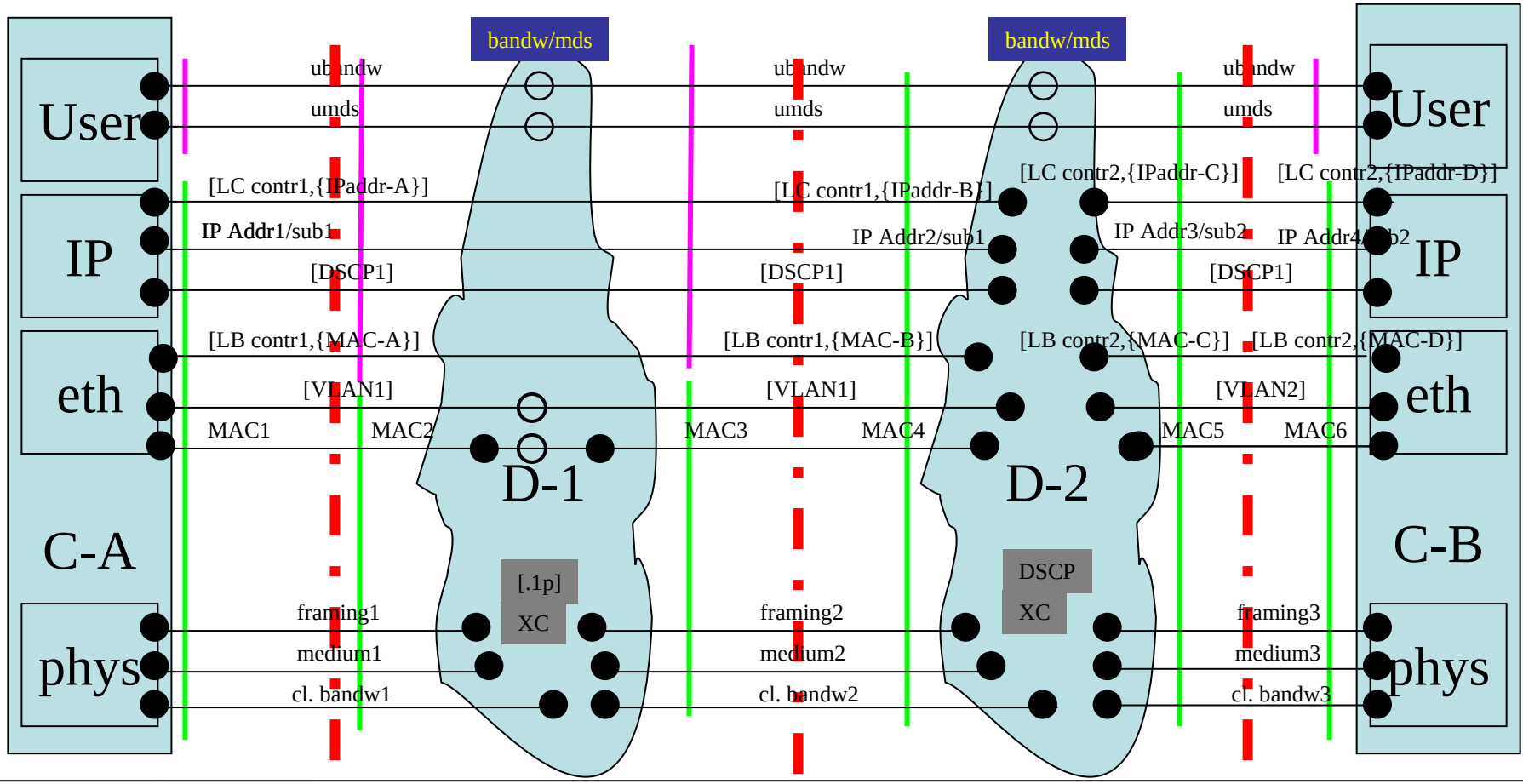
L2 MPLS VLL – L2 MPLS VLL (sc8b)

(VC_ID# stitching)

Connect. Communicate. Collaborate



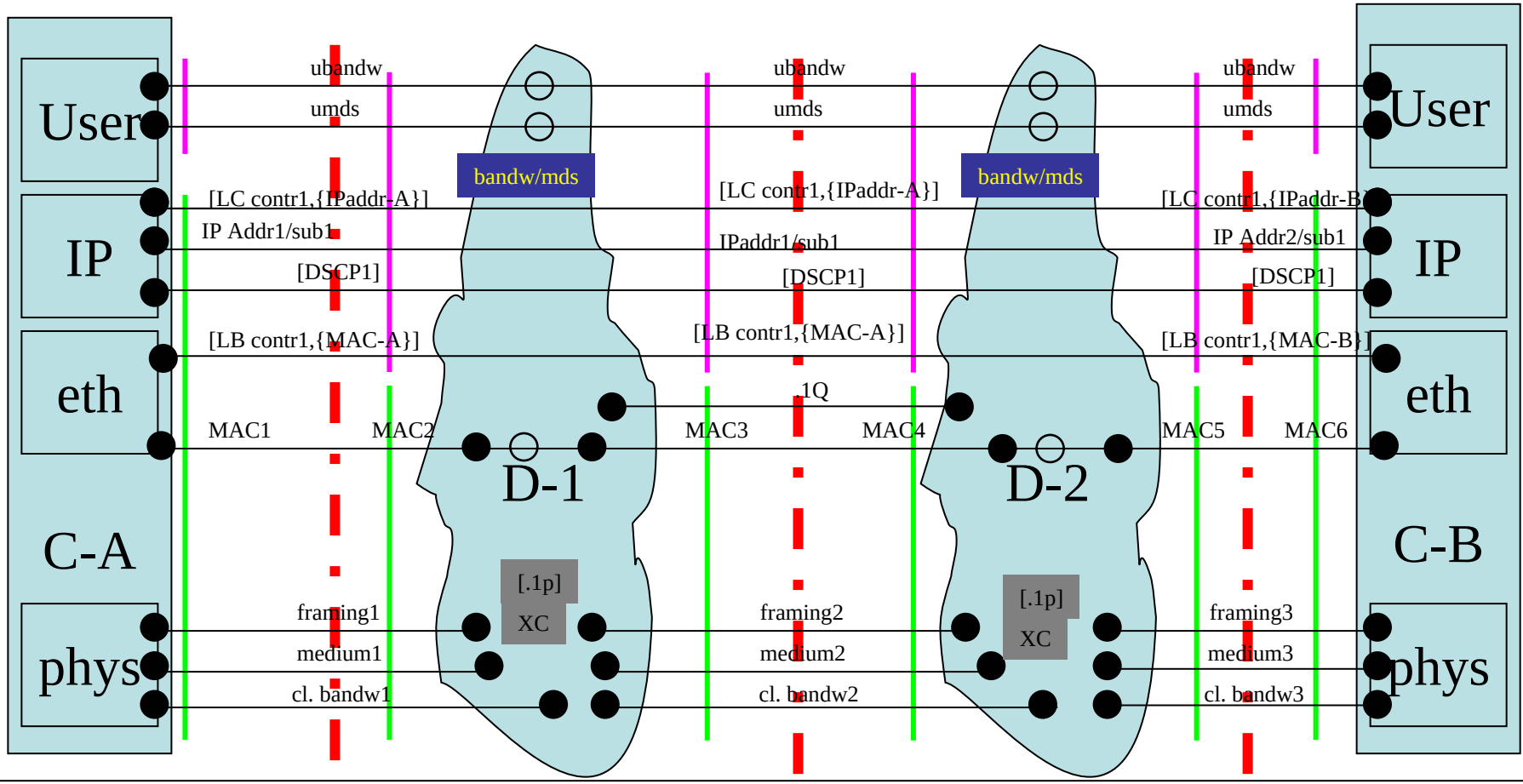
Ethernet – PIP (sc9)



Ethernet – Ethernet (sc14a)

(.1Q)

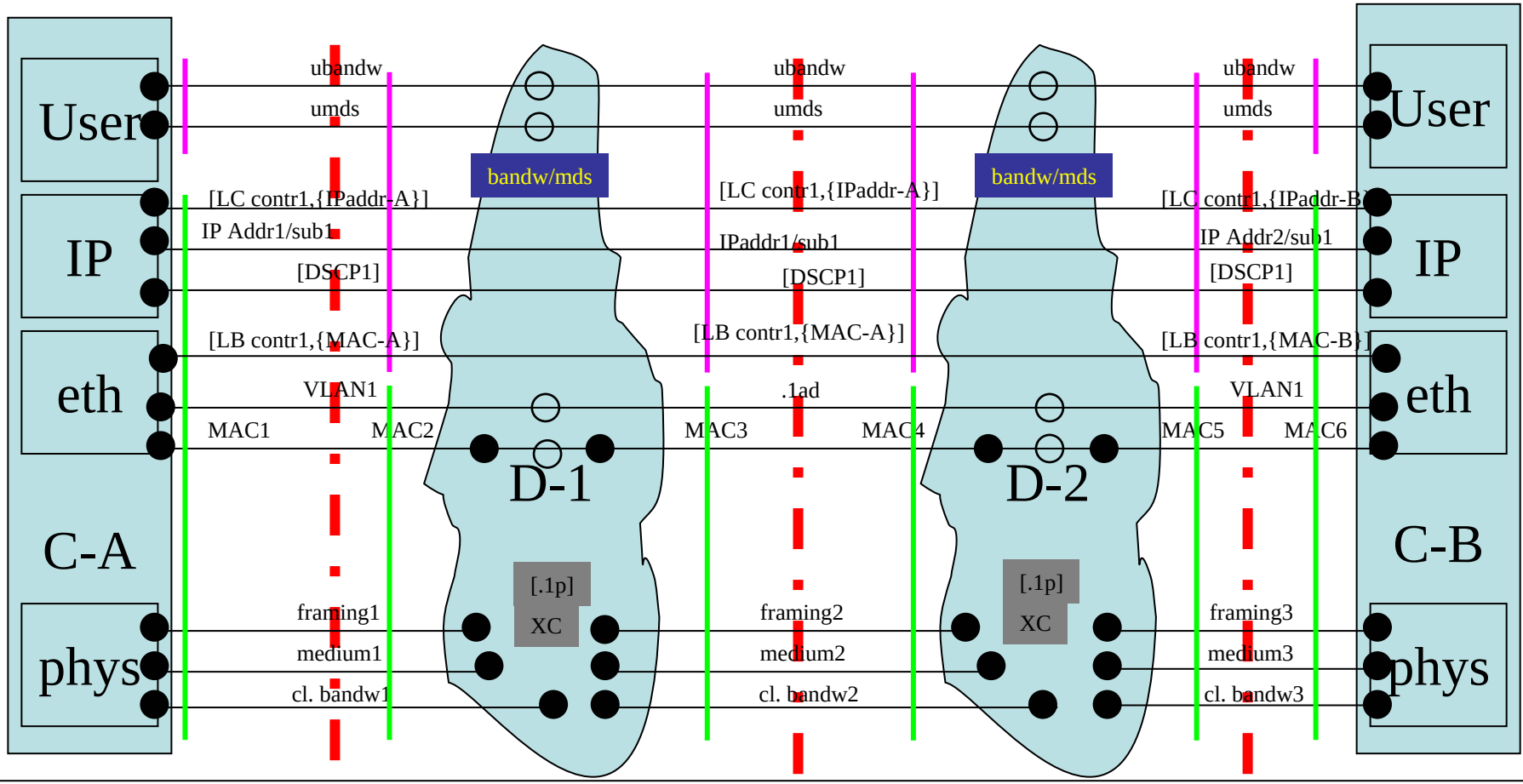
Connect. Communicate. Collaborate



Ethernet – Ethernet (sc14b)

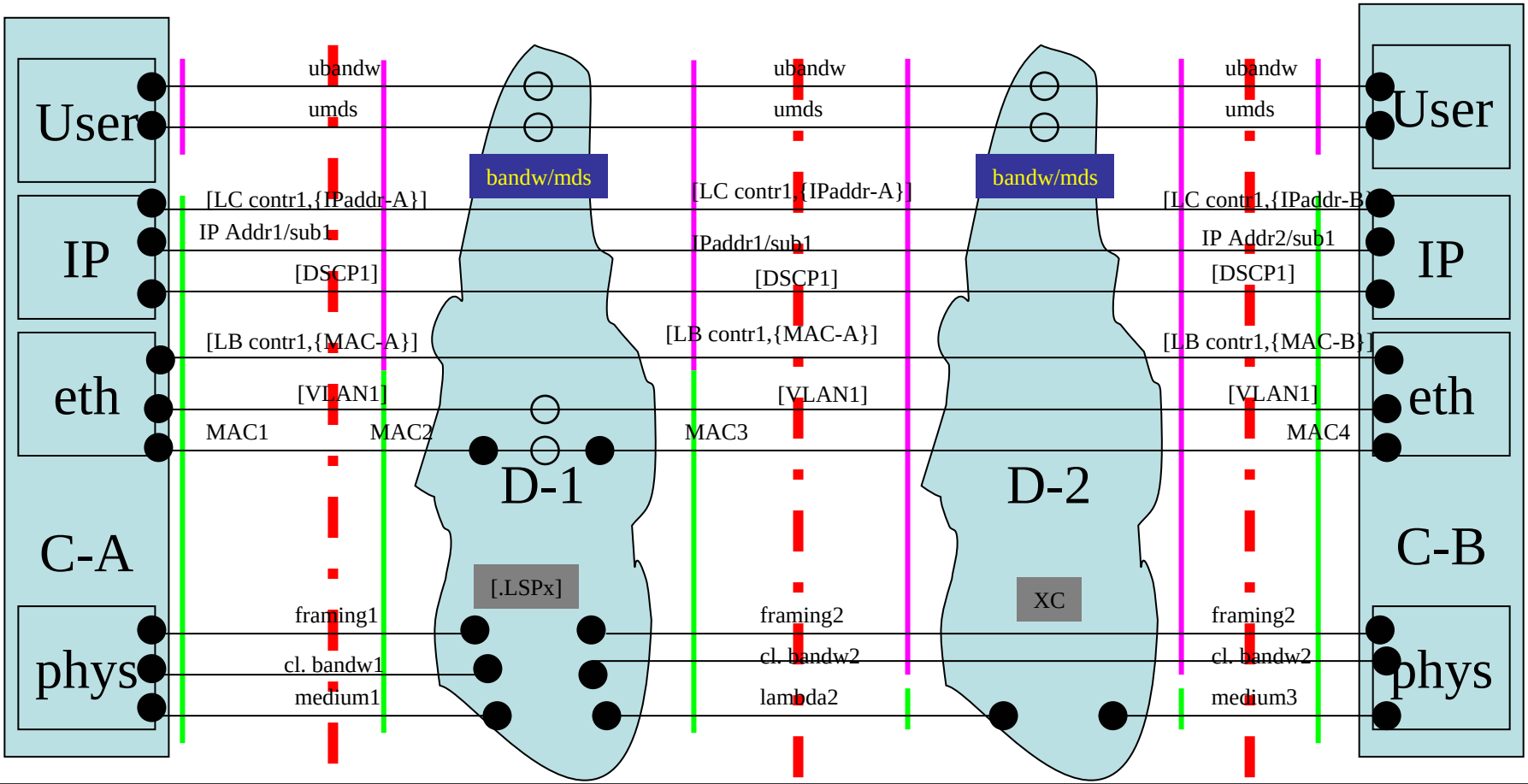
(.1ad)

Connect. Communicate. Collaborate



L2 MPLS VLL – OOO (sc15)

Connect. Communicate. Collaborate



Things to elaborate

- Multiplexing/demultiplexing VLANs to different destination
- Do VLAN# need to be network unique (or only interface unique)?
- How are MAC addresses handled in ethernet networks (each interface unique MAC address or can there be duplications).
- More then two domains
- Incorporate findings into/from path finding (paths from pathfinder, stitching feasibility check, etc.)

Connect. Communicate. Collaborate