

# ForCES Protocol

---

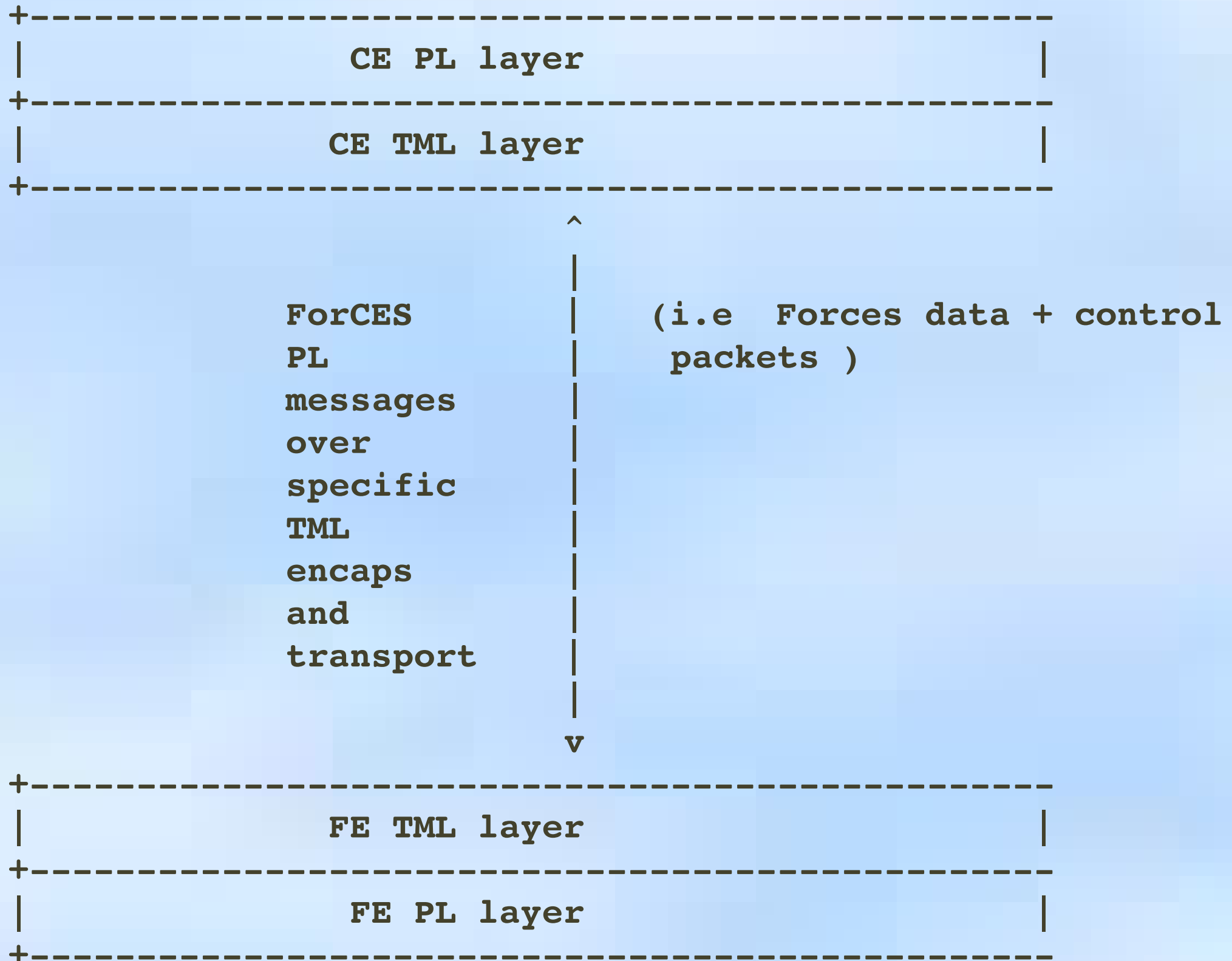
Avri Doria  
for the ForCES Protocol Team  
IETF60  
August 2004

[psg.com/~avri/forces/forces-ietf60-fpteam.pdf](http://psg.com/~avri/forces/forces-ietf60-fpteam.pdf)

# Team members

- Ligang Dong, Zhejiang Gongshang University
- Avri Doria, ETRI
- Ram Gopal, Nokia
- Robert Haas, IBM
- Jamal Hadi Salim, Znyx
- Hormuzd M Khosravi, Intel
- Weiming Wang, Zhejiang Gongshang University

# Framework



# Protocol Phases

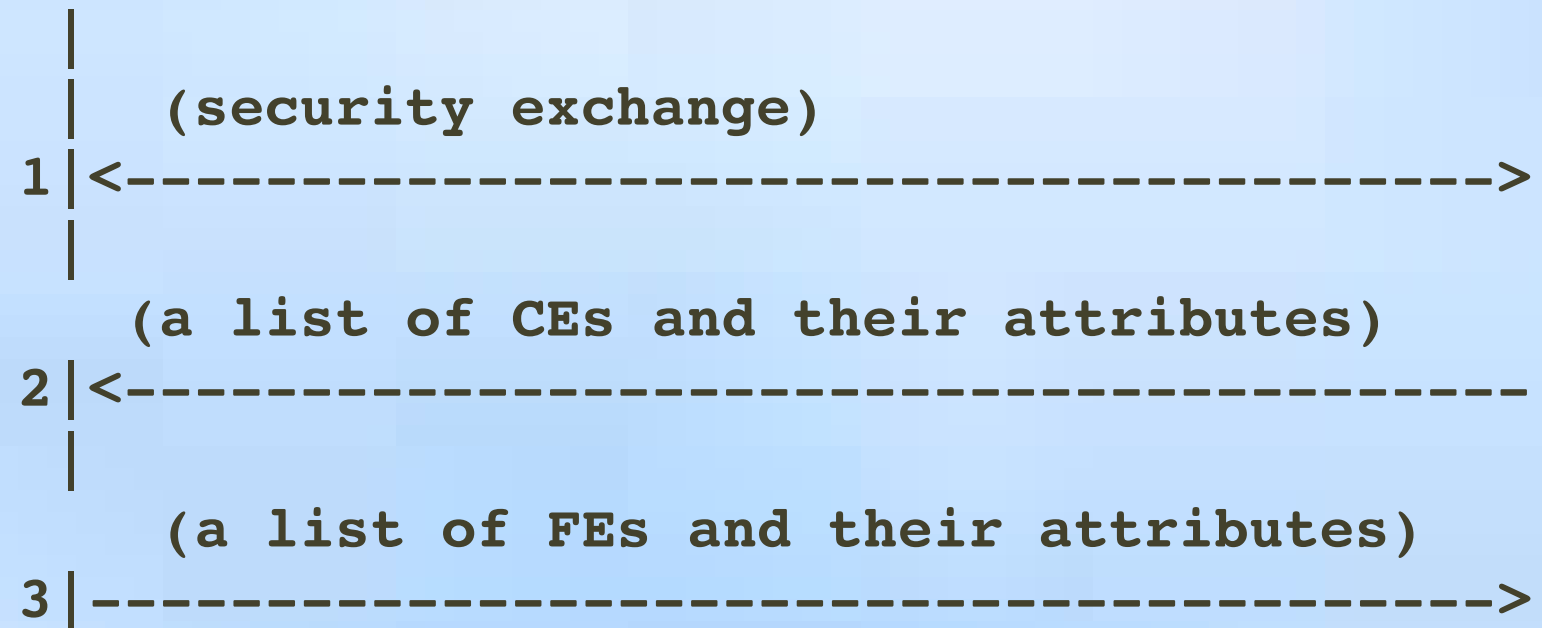
- Pre Association
  - CEM <--> FEM
  - FE <--> FEM
  - CE <--> CEM
  - Not part of protocol
- Post Association
  - FE <--> CE

# Pre Association Phase

<-----F1 ref pt----->

FE Manager

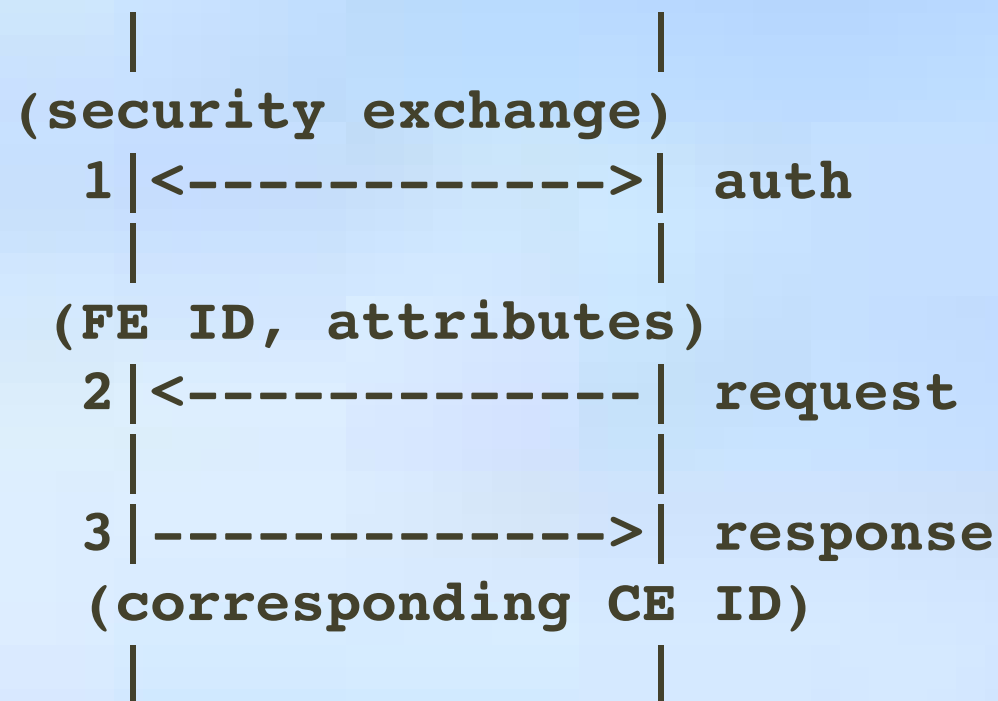
CE Manager



<----Ff ref pt---->

FE Manager

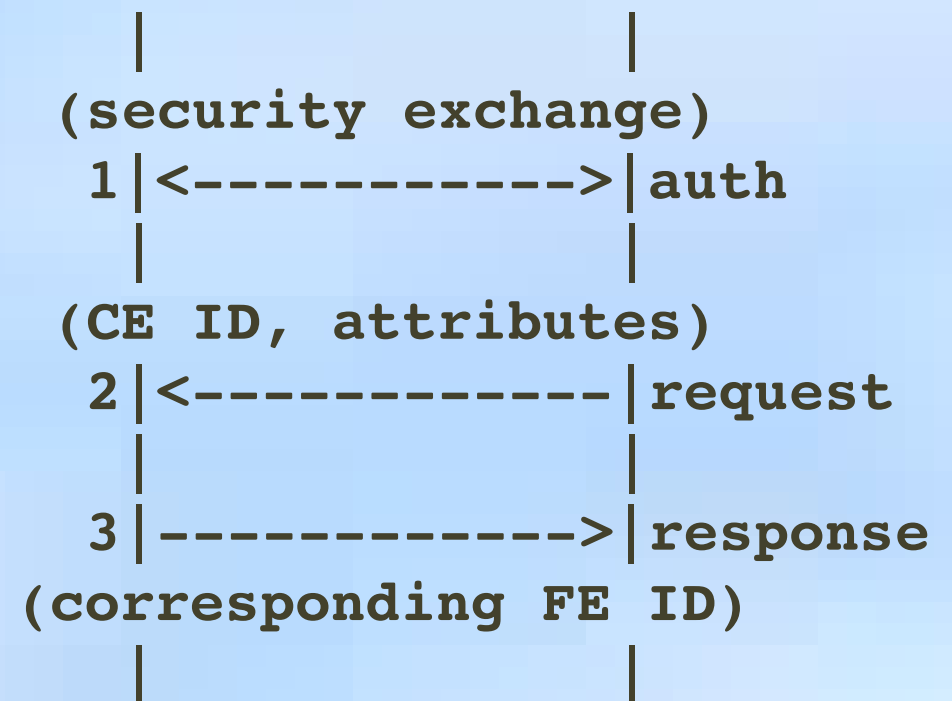
FE



<--Fc ref pt----->

CE Manager

CE



# Post Association Phase



# PL – TML Separation

- PL – ForCES Protocol Layer
  - Provides IETF defined protocol services to FE
- TML – Transport Mapping Layer
  - Resolves underlying transport issues
  - Possibility of multiple TML layer definitions
- FEM/CEM responsible for configuring FE/TML- CE/TML connection

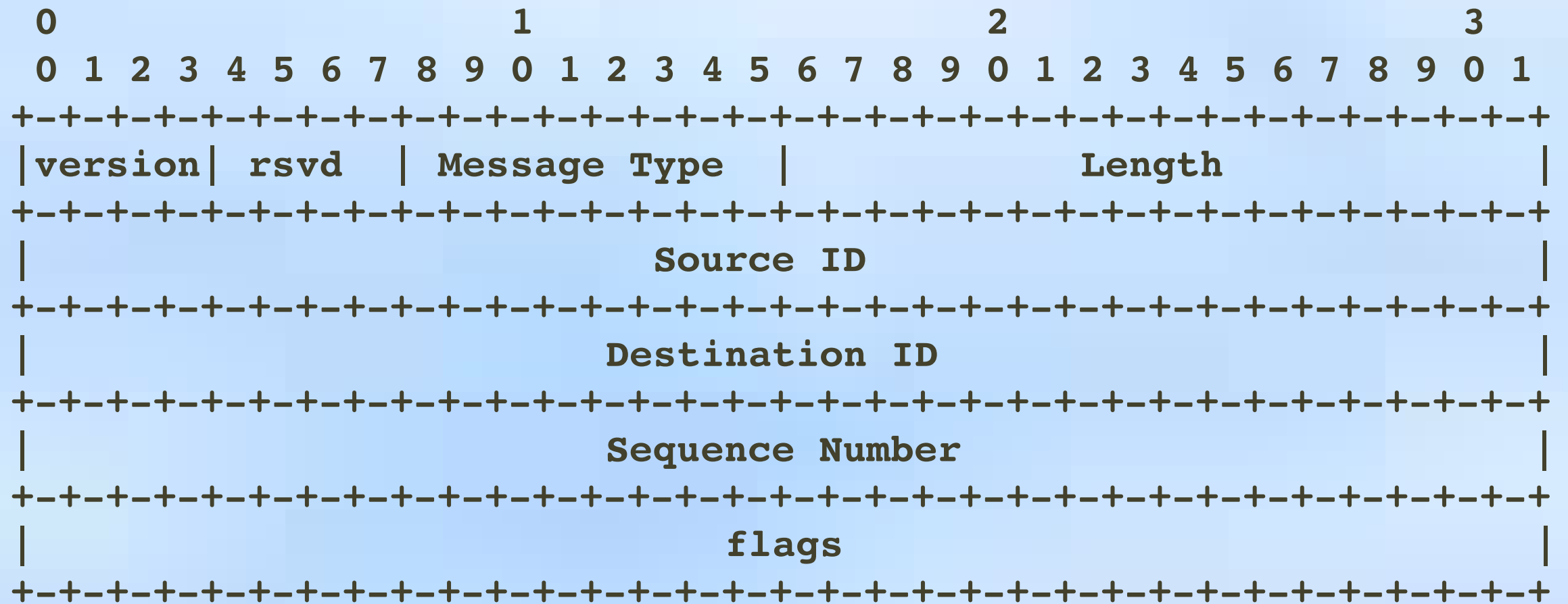
# TML Requirements

- Reliability
- Security
  - Endpoint authentication
  - Message Authentication
  - Confidentiality
- Congestion control
- Address mapping
  - Unicast
  - Multicast
  - Broadcast
- Timeliness
- High availability support
- Encapsulations
- Prioritization





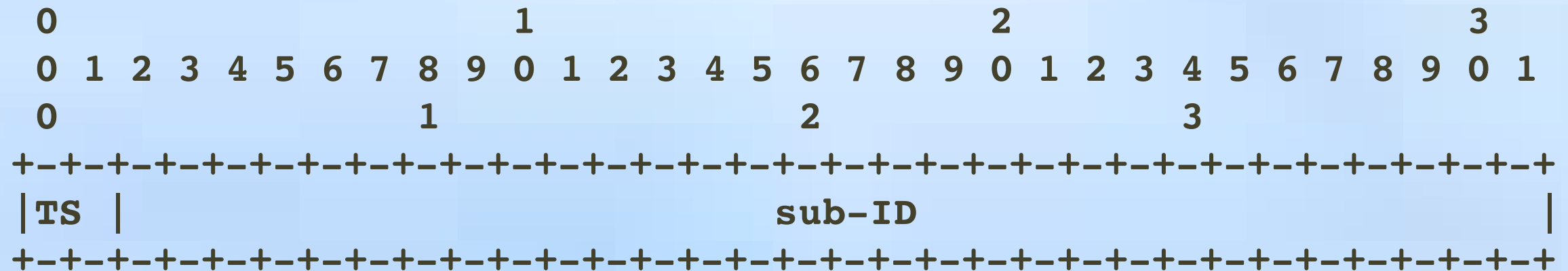
# Common Header



# Header Fields

- Version – 2 bit major, 2 bit minor
- Command/Message Type
- Source/Dest ID
- Sequence – Unique to the PDU
- Length – Header + Message in DWORDS
- Flags: Still need determination of flags that belong in header
  - ACK – only one defined so far
  - Atomicity ?
  - Batching ?
  - Throttle flag ?

# Source/Dest ID field



TS	Corresponding ID range	Assignment
--	-----	-----
0b00	0x00000000 to 0x3FFFFFFF	FE IDs ( $2^{30}$ )
0b01	0x40000000 to 0x7FFFFFFF	CE IDs ( $2^{30}$ )
0b10	0x80000000 to 0xBFFFFFFF	reserved
0b11	0xC0000000 to 0xFFFFFFFF	multicast Ids ( $2^{30} - 16$ )
0b11	0xFFFFFFFF0 to 0xFFFFFFF0	reserved
0b11	0xFFFFFFF0	all CEs broadcast
0b11	0xFFFFFFF1	all FEs broadcast
0b11	0xFFFFFFF2	all FEs and Ces (NE) broadcast

# Header Issues

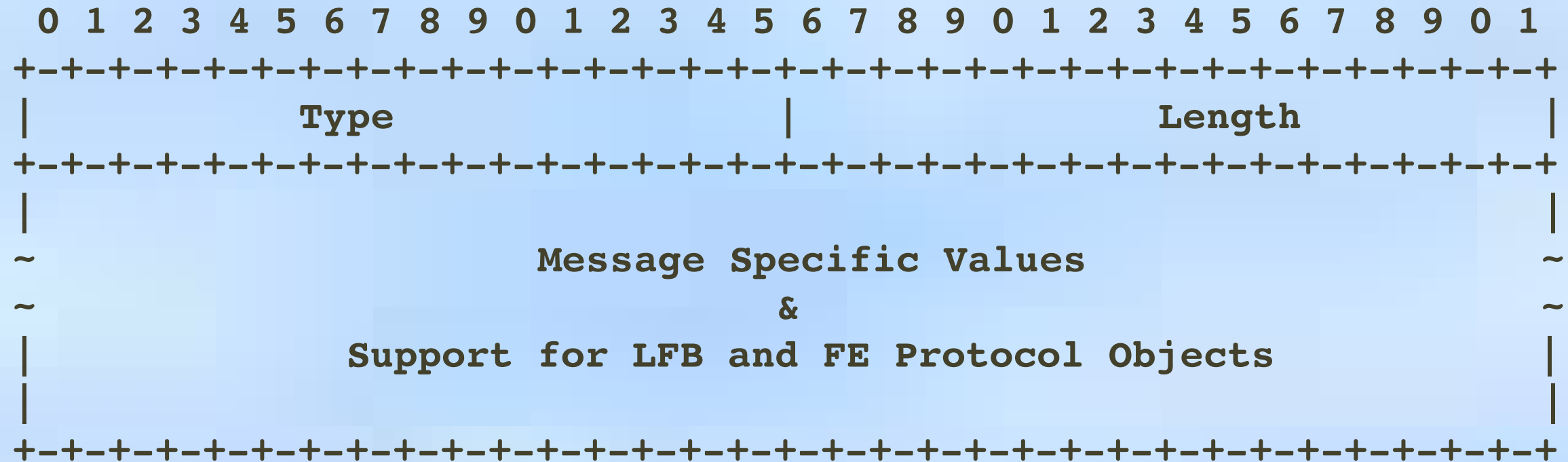
- Need to discuss how PL IDs map to the underlying TML level addresses
  - IPv4 – unicast, multicast
  - IPv6 – link local, global, ipv4 mapped, unicast, multicast
  - VPN IDs
  - Ethernet
- Do flags for Atomicity, Batching etc.. belong in the header or specific messages such as the configuration message

# Header Issues cont'd

- Sequence and subsequence numbers and their correspondence to events and replies
- Size of version field
- Seperate major/minor fields, size of fields

# Message TLVs

All Messages use the same TLV structure:



**Note: TLVs can be nested**

# Messages

- Association Messages
  - Setup, Setup Response
  - Teardown
- Query & Response Messages
- Configuration & Response Messages
- Event Notification & Response Messages
- Packet Redirect Message
- State Maintenance Message
- Heartbeat Message



# FE Protocol Object

- Logical entity in each FE responsible for control of FE protocol operations
- One and only one FE Object in an FE
- Analogous to an LFB
- Consists of the following elements:
  - FE Protocol events that can be subscribed/unsubscribed
  - FE Protocol capabilities, e.g. Protocol versions and supported FE models
  - FE Protocol attributes, e.g. Timers, failover and restart policies

# FE Object

- Also known as: FE Attributes
- One and only one instance of FE Object per FE
  - FE Events:
    - FEStatusChange  
FE Up/Down/Active/Inactive/Failover
    - FE DoS alert
    - FE capability change
  - FE attributes:
    - FE Behavior Exp. Timer
    - High Availability Mode

# Issues with Messages

- Need to synch with Model Section 5
- How should packet redirections LFBs be defined
- Is there any syntactic different between FE protocol objects and LFBs
  - I.e. Can they use the same TLV structures?
- Break out frequently used fields into common TLVs . E.g. Event Flags fields into a EF TLV

# Scalability Factors

- Batching – TBD
  - Multiple commands in one PDU
    - Issue of results and sequence numbering
- Command Pipelining – TBD
  - Separate independent parallel flows
    - i.e. given 10 Gbit pipe how do we use it?
  - Protocol should not be the bottleneck
- Uses and consequences of 2 Phase commit being discussed

# High Availability (HA)

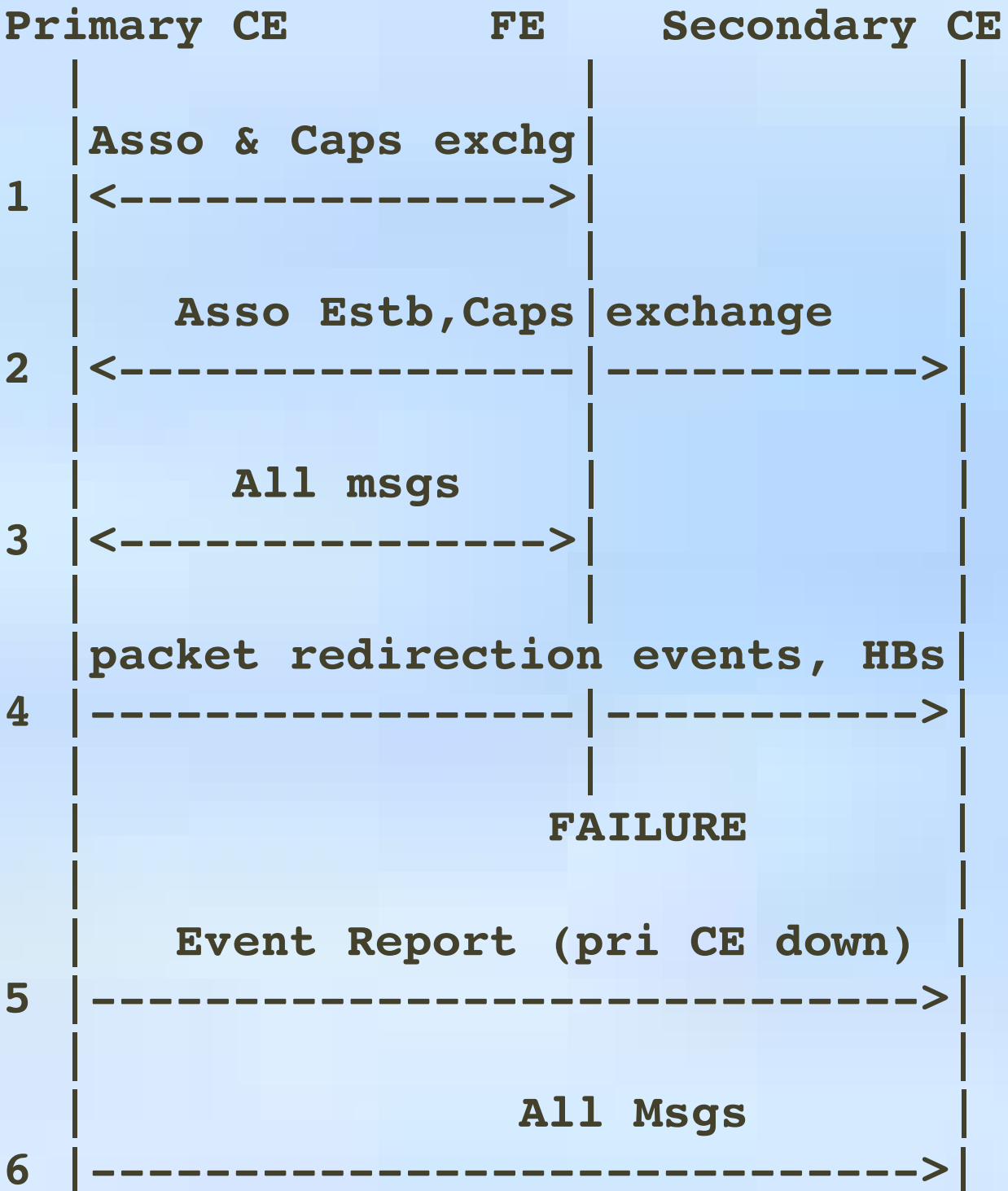
- High availability focuses on CE redundancy and not FE redundancy.
- The communication mechanism between CE $\leftrightarrow$ CE is outside the scope
- CE $\leftrightarrow$ CE coordination is expected
- FE  $\leftrightarrow$  FE coordination has not been discussed

# HA cont'd

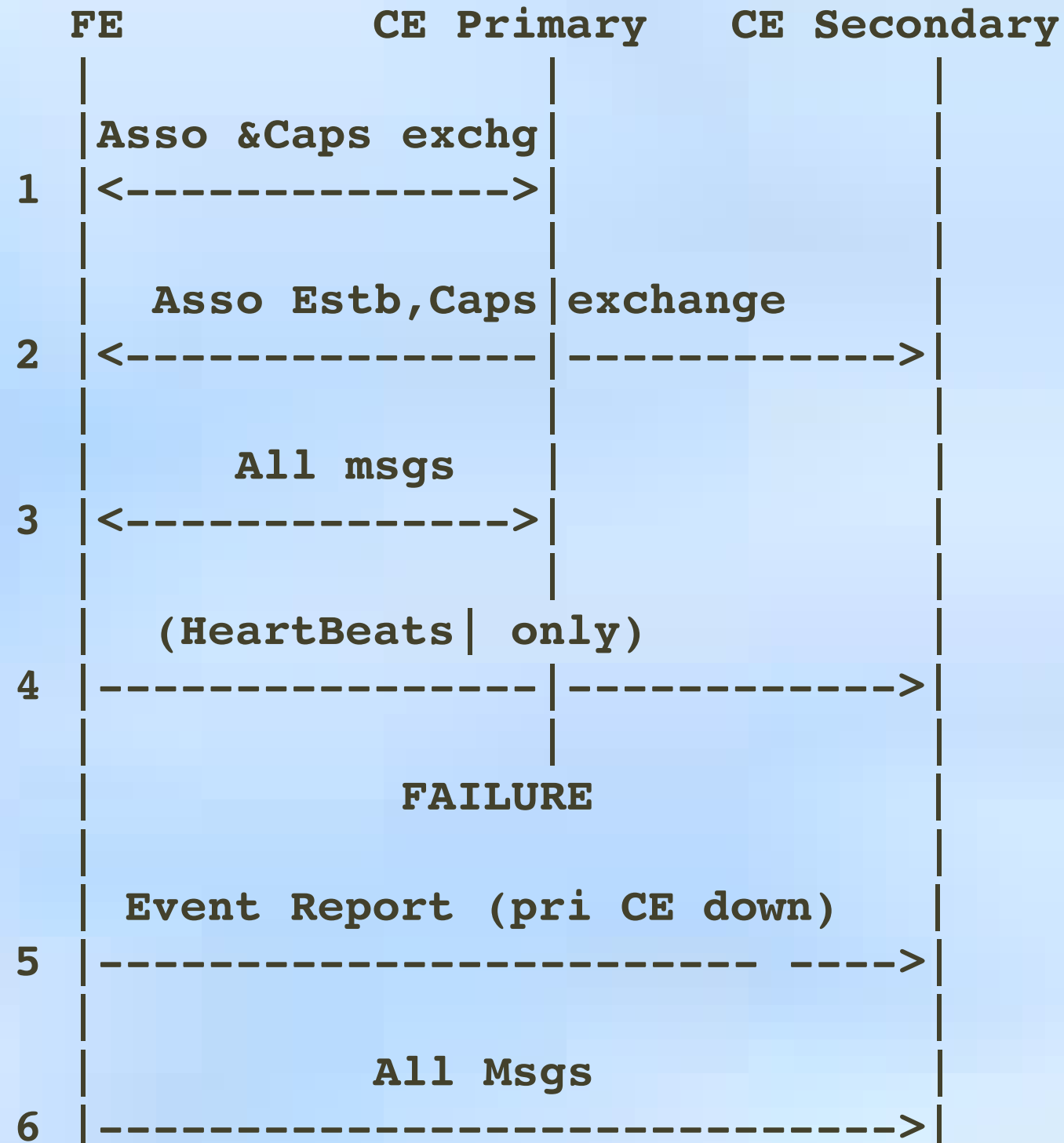
- Protocol supports CE reliability
  - One CE controlling all FEs
  - Support for secondary CEs available in configuration TLV
  - TML triggers switch over to secondary CE
  - 2 Modes
    - Report All
    - Report Primary
- Needs further discussion

# HA Modes

## CE Failover for Report All Mode



## CE Failover for Report Primary Mode



# Security

- Endpoint authentication provided in pre association phase
- ForCES PL layer does not do security.
  - Forces PL leverages underlying TML security functions
- TML Layer provides:
  - Endpoint authentication Service
  - Message Authentication Service
  - Confidentiality Service



# Security cont'd

- ForCES TML security configuration is independent of ForCES PL.
- Operator or administrator
  - has choice of any desired security level
  - Can configure appropriate policies at CE and FE endpoint.

# To Do

- Work with Model team to sync the protocol with the model
- Rework Introduction to explain much more and remove history
- Fill in all TBD sections
- Resolve open issues
- Add lots of explanatory text
- Add more examples

# Request

That the team draft:  
draft-doria-forces-protocol-01.txt  
become a WG draft

# Announcement

- 3 Implementations cooperated in forming a single NE cluster
- Interoperability / Demonstration to be held
  - At end of meeting if there is time
  - Or some other arranged time
- Plan on more extensive interoperability tests continuing in the future
  - All implementations welcome
  - Will announce on main list
  - Contact [avri@acm.org](mailto:avri@acm.org) if interested