Review of 802.1ag framework

David Elie-Dit-Cosaque, Maarten Vissers

ALCATEL

Goals

- Review ideas in 802.1ag framework (see Norm's presentation) to
  - identify unstated assumptions
  - discuss key underlying concepts
  - questions
- Introduce proposals
- Discuss the role of OAM (and relation to the network management plane)
- Consistency with ITU-OAM architecture

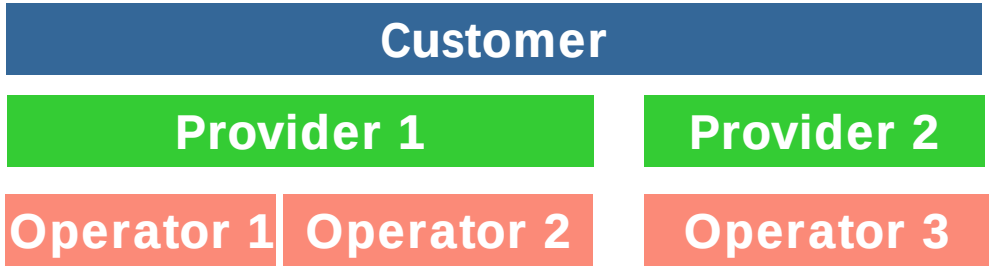
Why Ethernet OAM?

- Operators
  - They already have a detailed, resilient management plane (NMS)
  - They have link management tools to detect link failures
    - For Operators, Ethernet OAM is used to (a) detect fabric failures and (b) in cases where NMS does not exist, or where the NMS is unable to correlate link failure to a connection.
- Providers
  - Work is in progress (in ITU-T) to provide limited access to providers to the operator management plane
- Customers
  - Customers will probably not have access to the Provider/Operator management plane
  - A customer might need to verify SLA's provided to him
- Operators will need minimal Ethernet OAM. Providers will need more comprehensive Ethernet OAM for themselves and to allow customers better monitoring functionality.

Scenario requiring Ethernet OAM

> A likely scenario requiring Ethernet OAM is the following:

- An Operator owns provider bridges with redundant management plane (withstand at least 2 link failures)
- A Provider could use the services of multiple operators and has no or limited access to the Operator's management plane
- A Customer could use the services of multiple providers and has no access to the provider management plane



> Ethernet OAM is needed in cases where there is no interface between the management plane and the link monitoring tool.

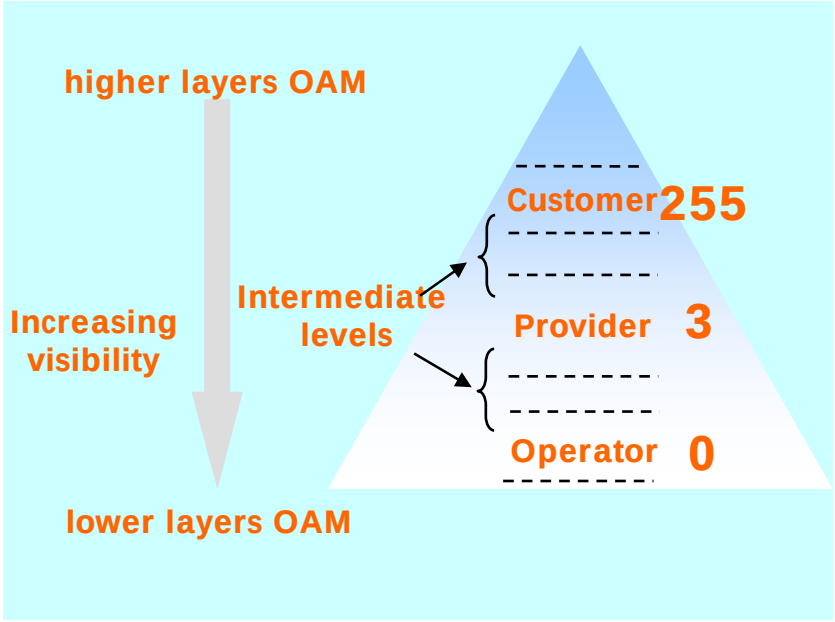
ALCATEL

Requirements for Ethernet OAM

- Requirements for Ethernet OAM are different for the operator and for the customer
- Ethernet OAM must be able to:
  - monitor the health of links (because providers and customers might not have access to the management layer)
  - check connectivity of ports
  - detect fabric failures
  - provide the building blocks for error localization tools
  - give appropriate scope to customers, providers and operators (hierarchical layering of OAM)
  - avoid security breaches (be “leak-proof”)
  - easy to configure (Plug-and-Play operation)
  - optional: perform basic performance monitoring (delay, jitter, frame loss) end to end.

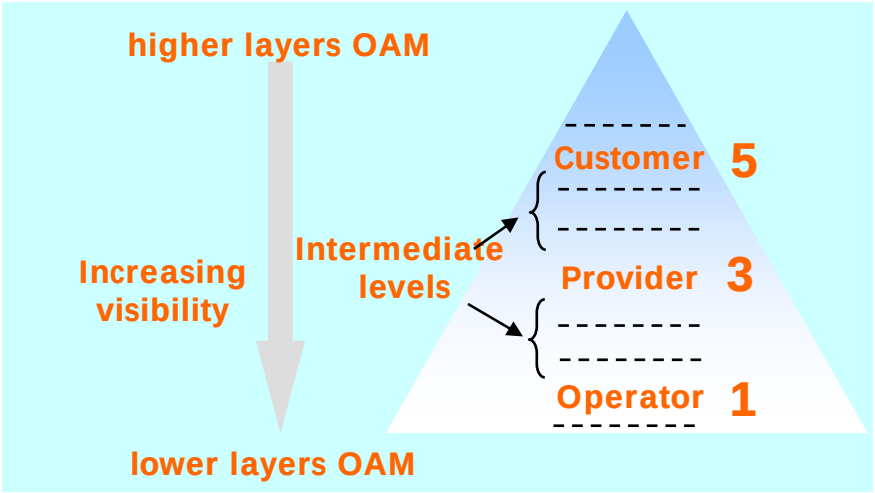
OAM Levels - Proposal for level to layer mapping

- The Operator gets ”full visibility”. Customer has “limited visibility”
- Mapping of layers to levels is confusing.
- Level is a numerical value.
- Proposal for levels with fixed numbers: higher layer (customer) is associated with a higher level number.
- There can be up to 256 layers, and hence 256 levels.



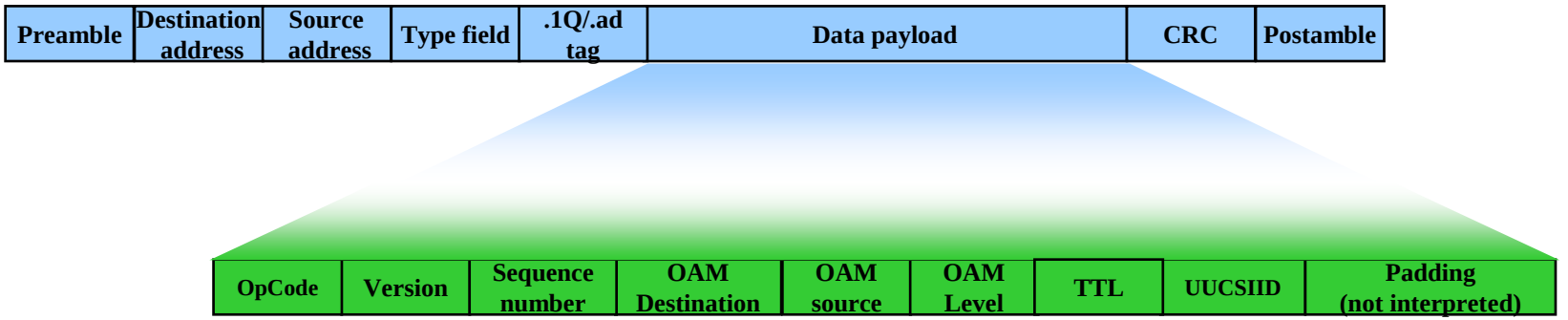
- Eg. Customer can be at Level 255, and Operator can be at Level 0.
- Different layers can also have relative levels - discussed later.
- For fixed levels, the LP’s have to be provisioned manually.

OAM Levels



- Rules:
  - every LP at higher layer is also a MP for the lower layer
  - OAM frames at a given level *n* remain at that level
  - OAM frames are either processed or discarded by same level LP/MP:
    - discarded when the frame originates from outside the OAM domain
    - processed when the frame originates from inside the OAM domain
- higher layer OAM frames (e.g. customer) are relayed transparently by lower layer OAM MP/LP (e.g. operator).
- lower layer OAM frames are terminated before a higher layer OAM LP or MP.
- a node can have multiple levels.

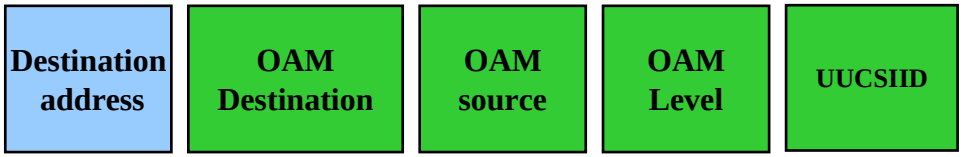
OAM Frame format



- **Destination MAC@** : Multicast MAC @ (for Traceroute, Connectivity check) or unicast (for Loopback)
- **Opcode:** specify the operation being performed (e.g. Ping, Connectivity check, Traceroute, loop detection,error message ...)
- **Version:** specify the version of the OAM protocol
- **Sequence number:** might be useful to detect out of order OAM frame (useful for loop detection)
- **OAM destination:** the MP or LP that is being tested
- **OAM source:** the MP or LP that is sending the OAM frame
- **OAM level:** the OAM level to which this frame is applicable
- **TTL:** Field used to prevent OAM frame from looping indefinitely. Used also in Ping and Traceroute
- **UUCSIID:** Globally unique Service VLAN identifier
- **Padding:** allow OAM frames to have variable sizes thus enabling MTU discovery.

Continuity Check / MP Connectivity Check?

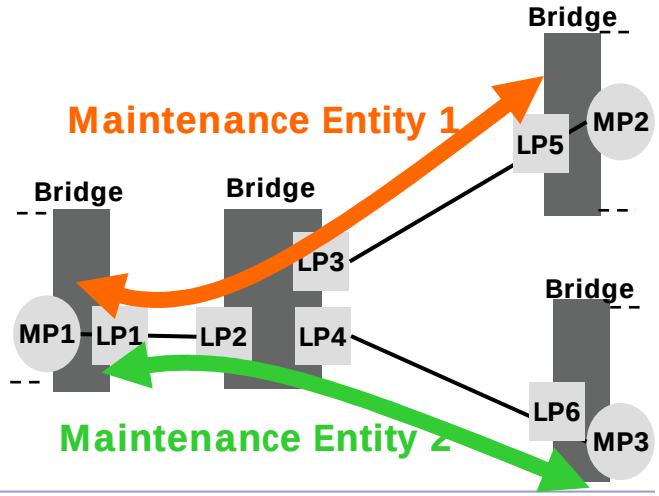
- **Goals:** a) MPs periodically check the status of a Maintenance Entity (ME)
  b) allow MPs to discover other MPs across an OAM domain
  c) Detect merged VLAN conditions
  d) Discover information for alarm suppression
- It only indicates that there is a problem, but does not indicate where
- **Minimal set of used fields :**



- **Database :**
  - in LP: level information, ...
  - in MP: builds the CC database (or MP database)

	MP MAC @	OAM Level
MP 1	MAC @ 1	12
MP 2	MAC @ 2	2

  - **Basic Description :**
    CC frames are unidirectional. Multicast by MPs to other MPs in their level periodically . Each Each MP receiving CC frame adds it to its reachable MP database.
    The database lists the MAC address and level of reachable MPs



ALCATEL

Proposal for CC traffic priority

- How do we treat CC traffic? Is it high priority traffic or is it like ping best effort?
- Proposal: CC traffic has highest priority.
- Pros:
  - The purpose of CC is to detect that the link/fabric is functioning.
  - One CC for all classes of service. Else, need a CC for every class of service.
- Cons:
  - Problems in the data path for data traffic may not be caught.

Proposal: Introduce a new message similar to CC but with an additional Timestamp field

- If OAM traffic has the same priority as data traffic, then:
- Proposal: create a new CC like message with timestamp
- CC with timestamp will indicate a one way delay.

Pros:

- Could report the health and status of a Maintenance Entity (e.g. the route between two MPs) at no extra messaging
- Could be used to trigger a fault isolation tool (e.g. Ping)

Cons:

- One additional field

Multicast Traceroute (is really LP Connectivity Check)

- **Goal:** Allow MP's to discover intermediate LPs/MPs inside an OAM domain
  Is it Really Needed? Currently, it is not very useful with absolute encoding[see later slides]
- **Minimal set of used fields:**
  - **Multicast** MAC @ used in the destination address to perform flooding at a given OAM level
  - OAM Frames from higher layers are not processed by the OAM entity within and simply flooded on the ports => better performance

- **Database**
  - in LP: none
  - in MP: Builds the Traceroute database (or LP database)

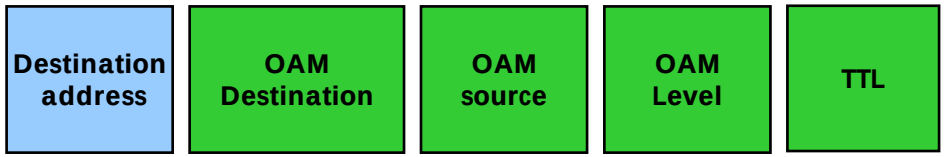
	LP MAC @	OAM Level
LP 1	MAC @ 1	12
LP 2	MAC @ 2	2

- **Basic Description:**
  A Traceroute frame is sent by MPs. Traceroute relies on multicast to flood the OAM frame to each LP and MP in a given OAM domain. Each LP or MP receiving a Traceroute frame will respond with a Traceroute reply frame addressed to the **OAM source** in the received OAM frame. The LP/MP puts its address as the new **OAM source**. After receiving the Traceroute reply, the MP adds the **OAM source** (in the received frame) in its Traceroute database.

Both CC and Multicast Traceroute set the required environment for the Ping

Loopback message (is really Ping)

- **Goal:** Check the connectivity between an MP and a LP or MP. Also used to locate errors (usually more than one Ping needed)
- **Minimal set of used fields:**



- **Database**
  - in LP: none
  - in MP: none

- **Basic Description:**

  A MP first does a lookup of its Traceroute or Connectivity check database respectively in order to get the MAC @ of its target (MP or LP). The originating MP builds a ping message unicast to the destination MP or LP, with a **TTL** set to its max value. The ping is forwarded to the destination LP or MP and the TTL is decremented at each hop. The receiving MP or LP sends a Ping Reply**.** If there is an error condition, we need to return an error reply that indicates the error condition.

Problem with multicast trace route - How does a LP know to which level it must respond to?

- A LP will see multiple CC's corresponding to different levels.
- How does it know to which level it must respond to?
- There are two approaches
  - Absolute level encoding- LP is configured with a certain level.
  - Relative level encoding (Stack approach) - LP discovers its level.
- If Absolute level --> LP already knows which level it is and therefore what is the benefit of MP "discovering" LPs ?
- If Relative level --> then Traceroute is useful.
- These approaches have not yet been discussed in 802.1ag, but have been discussed in depth in ITU-T.
- The choice of these approaches affects the configuration of MP/LP's.

Relative or Absolute OAM levels ?

- Absolute OAM level encoding uses an integer value to indicate a specific level.
  - MPs and LPs only process the OAM frames with a level field equal to their configured level
- Relative OAM level encoding also uses a integer value to indicate if a OAM packet must be processed at the current level
  - the originating MP always insert “Level 0” in the OAM Level field
  - The OAM Level field is incremented when the OAM frame hits an MP source
  - The OAM Level field is decremented when the OAM frame hits an MP sink.
  - If MP source: then it inserts OAM frames, and it increments the level of frames that enter.
  - If MP sink: then it extracts OAM frames, and it decrements the level of frames that enter.
  - Already discussed at ITU-T

Absolute OAM encoding

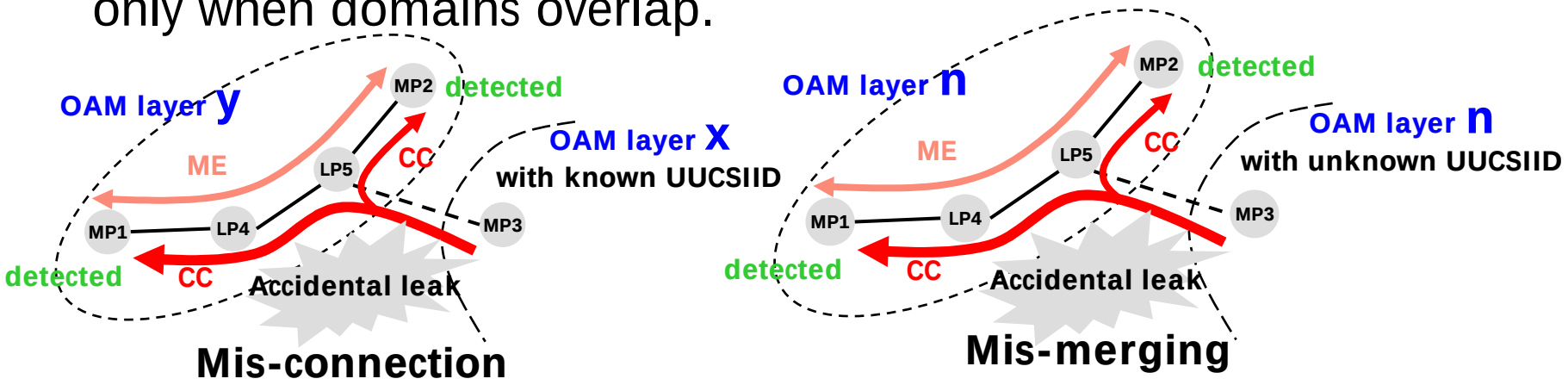
- Advantages
  - Any packet contains explicitly its level ID
  - Possible use of Multicast address (for CC and TR) to relieve the bridge from processing OAM frames at every layer
  - Does not require new hardware
- Drawbacks
  - Every MP/LP must be configured with its level

Relative OAM encoding

- Goal: Define a “Minimal configuration MP”
- Advantages
  - MPs need not be configured with their Level ID.
  - MPs and LPs don’t need to know their absolute level for correct operation
- Drawbacks
  - By directly looking at an OAM packet, it is not possible to tell which level it originated from
  - LPs are not capable of discovering their absolute level
  - The burden to select which frame to terminate or to pass through is on every bridge OAM entity (Shim layer or Bridge brain).

OAM misconfiguration scenarios

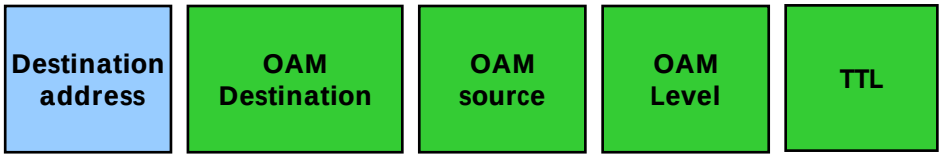
- Mismerging: get OAM frames from different service VLANs. Thus, receive CCs (with wrong UUCSIID) that are not part of the ME - due to corrupt bridge filtering database?
  - To detect mismerging, need to probe a branch that the originating MP does not know about.
  - Tools: Multicast Ping, Multicast Traceroute (same as multicast ping with unnecessary messaging)
- Misconnecting: Get valid CCs with wrong MPs. This is possible only when domains overlap.



ALCATEL

Proposal: Multicast Loopback message (is really Multicast Ping)

- **Goal:** Used to locate mismerge errors.
- **Minimal set of used fields:**

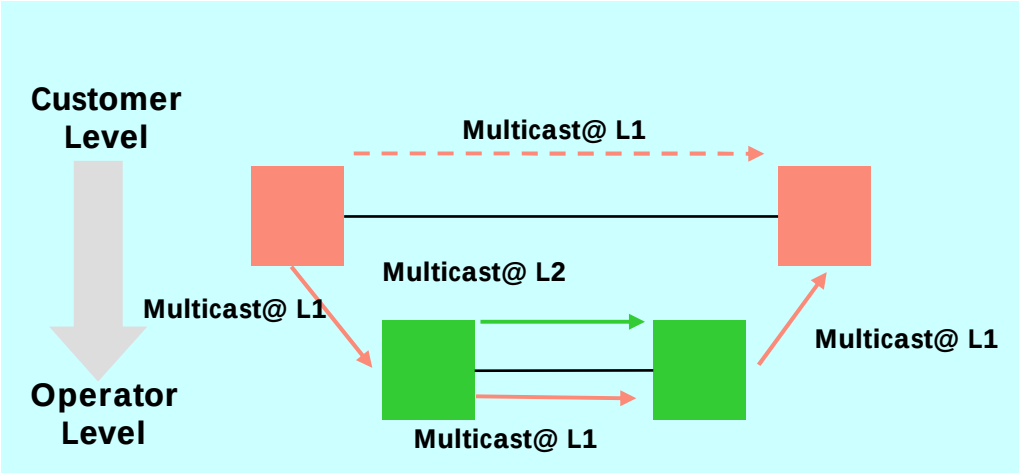


- **Database**
  - in LP: none
  - in MP: none
  - OAM destination is known -- else it becomes Traceroute.
  - Destination address - is the multicast address at this level.
- **Basic Description:**

  A MP first does a lookup of its Traceroute or Connectivity check database respectively in order to get the MAC @ of its target (MP or LP). The originating MP builds a ping message multicast to the destination MP or LP, with a **TTL** set to its max value. The ping is multicast within that level and the destination will receive it. The TTL is decremented at each hop. The receiving MP or LP sends a Ping Reply**.**
- Compared to unicast ping, the destination and other nodes receive it as well.

Multicast addresses

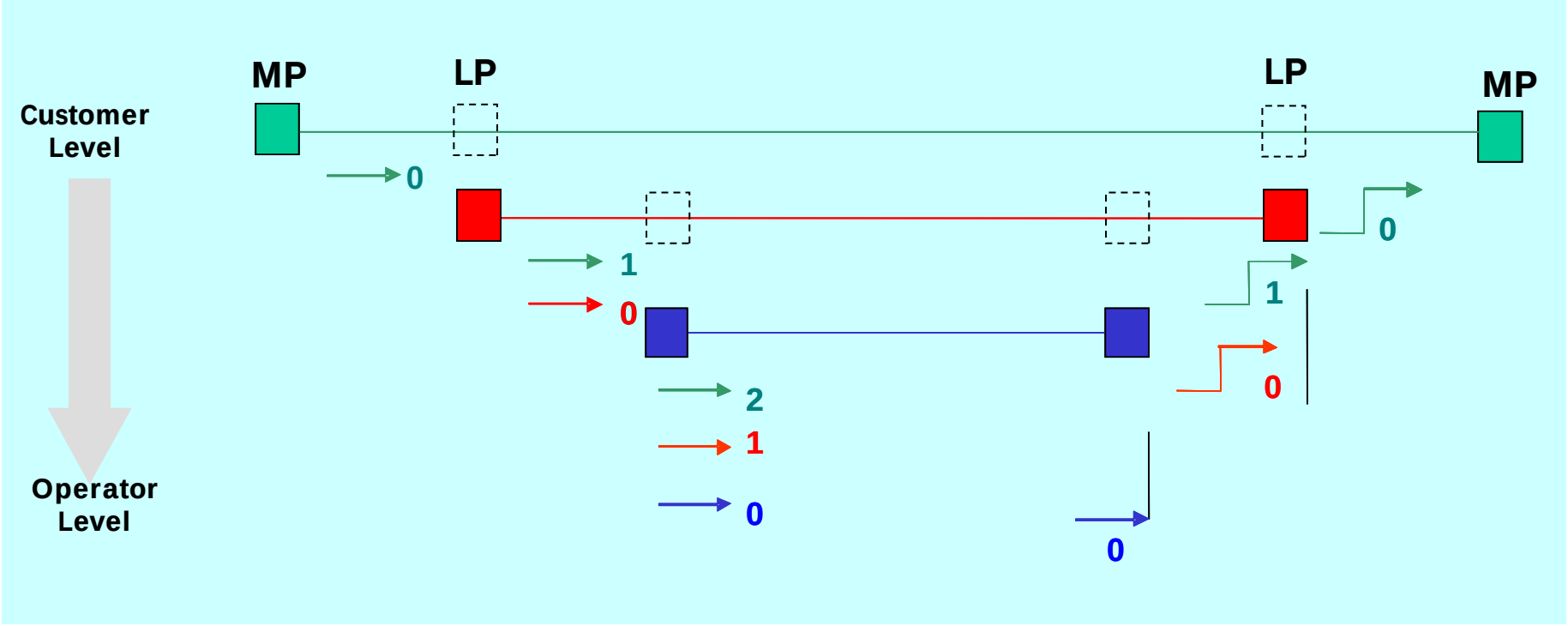
- If 1 MC address per OAM level
  - additional security
  - cheaper to process: because no need for an LP to examine the level field in the frame
- If 1 MC address for all OAM levels
  - OAM frames are multicast at all levels. Thus, every LP has to examine the level field in the frame to decide forwarding/processing (OAM shim)



ALCATEL

Relative OAM encoding

> From ITU-T



The relative level numbering is irrelevant to the absolute level numbering scheme proposed.

ALCATEL

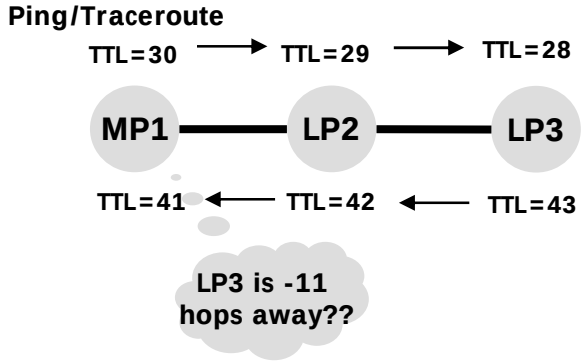
Proposal to configure LP for Absolute Encoding: Have both types of encoding in the OAM frame at same time ?

- This solves the problem of configuring the LP with its level for the case of absolute encoding. Currently needs to be done manually.
- The relative encoding indicates that the OAM frame has originated from the current layer if it is equal to 0, the absolute level encoding just indicate which level it is!
- **Thus the LPs are able to deduce their level just by looking at the OAM frames passing through them.**
- Benefits from both approaches
  - plug and play
  - only configuration needed is deciding which ports are LPs or MPs
  - absolute level "self-contained" in the OAM frame
  - increase security if a OAM frame leaks out of an OAM domain
- We propose adding a new field to the OAM frame :
  - **relative OAM level field :** 1 byte = 255 level offsets possible **(new)**

Adds only 1 byte of overhead to the OAM frame for a simpler and more secure Eth OAM protocol

Fault detection

- Problem to be solved:
  - Determine the ordering of LP's (route)
  - Location of the LP's that bound the error
  - Build a map of the network



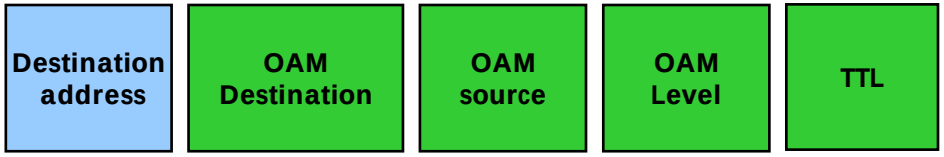
- Existing tools and why they don't work
  - Multicast traceroute - this would work if:
    - there were no branches and
    - all the flow points set the initial TTL to the same original value
  - Ping (with no error messages) - would work provided:
    - we ping all the LP/MPs in the OAM domain and
    - all the flow points sets the initial TTL to the same original value and
    - there are no branches

New tools are needed to provide Error localization
Example: unicast Traceroute?

ALCATEL

Proposal: “True” Traceroute (unicast traceroute) for Fault Detection

- **Goal:** Display the actual route between an MP and an MP or LP and the associated performance (delay , jitter, frame loss)
- **Minimal set of used fields:**
  - Traceroute uses a **Unicast** frame. The domain encoding is exclusively done in the **OAM level** field as opposed to being done in the Multicast Address.



- **Database**
  - in LP: none
  - in MP: none
- **Basic Description:**
  Traceroute reuses the Ping functionality and is described for completeness. The protocols is similar to the IP ping as follow:
  - send a ping to destination MP with TTL = 1
  - receive a **OAM Error** frame from intermediate node (e.g. TTL = 0)
  - record intermediate node address
  - increment **TTL**, send a ping to destination MP with TTL = 2
  - and so on…

For OAM error: No need for any new messages: Just an OpCode.

The end result is a route to the destination MP.

ALCATEL

Proposal: OAM needs some error/condition reporting message

- Fault between two LP's
  - some LP's will not respond to Ping messages.
  - **Need for intermediate error reporting (similar to ICMP)**
- Multiple conditions need to be reported by flow points in the network:
  - Destination Unreachable ?
  - Time to live equal to 0
  - indicates problem with the formatting of the OAM frame
  - others?…

Thus, Ping needs to have an Error Message descriptor
Ping also needs to have an additional field for Time Stamp

Clarifications needed

- Links between bridges are monitored. Link status (failure or degradation) is monitored by an autonomous system.
  - If a link failure is detected, then this condition is reported as an alarm to the NMS via SNMP. However, node failure/status requires monitoring as well.
- OAM used only to detect fabric status?
- Specify List of functionality in Ethernet OAM - [Ethernet OAM tries to solve when no access to operator management plane]
- What information is conveyed from operator to provider via the management plane?
- Role of the following functionality?
  - Next LP hop in the CC database?
  - purpose of hop count TLV?
  - MP address in CC database at the LP? (why not only level?)

Appendix: Some principles of OAM architecture

- Both Multicast address and the maintenance level encode the OAM level
- The use of multicast address is for performance reasons only
- Every bridge will have LP points at least at the lowest ME level.
- Per port - can have 4000 VLANs, and hence 4000 LP's and more than 4000 ME's
- CC messages sent every 1 second, on a periodic basis.
- Norm's Traceroute is only for discovery and connectivity check of the LP/MPs . Is this correct?
- Traceroute messages are used infrequently - triggered only after a ST reconfiguration, or change in network configuration

Appendix: Some principles of OAM architecture

- An Ethernet connection is a set of two or more ports in a stable STP topology. A spanning tree will consists of multiple connections.
- End points of these connections are MPs,intermediate points are LPs.
- Every Ethernet connection over a bridge has to be a part of a ME.
- A link may have multiple connections going over it.
- Therefore, all edge bridges in a network must be OAM capable and implement MPs.
- If there are non-OAM-capable bridges inside the domain (I.e. not at the edge!), the OAM capabilities in the domain will be preserved, maybe somewhat degraded, as long as those non-OAM capable bridges don't drop OAM frames.
- Every bridge has 2 LPs per ME/OAM domain.