The Bridge Queue Placement Problem

Norman Finn

Cisco Systems

Version 2

IEEE Std. 802.1Q-2006

Subclause 8.6 The Forwarding Process

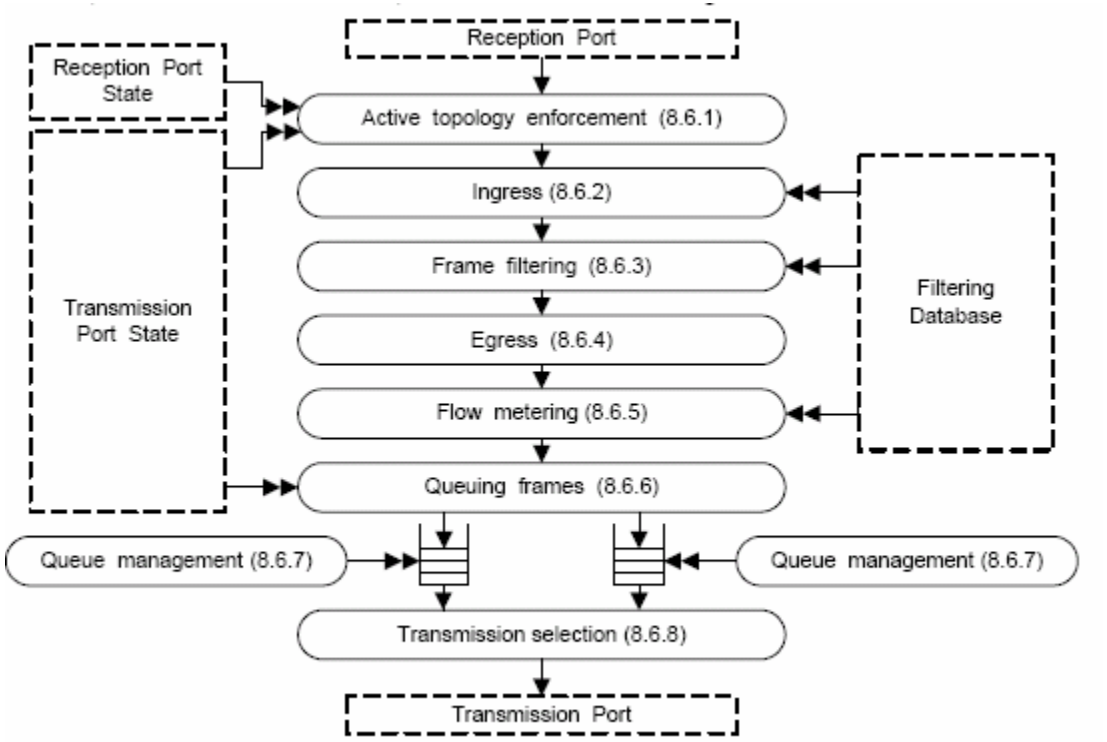


Figure 8-9—Forwarding Process functions

IEEE Std. 802.1ag-2007

Subclause 22 CFM in systems

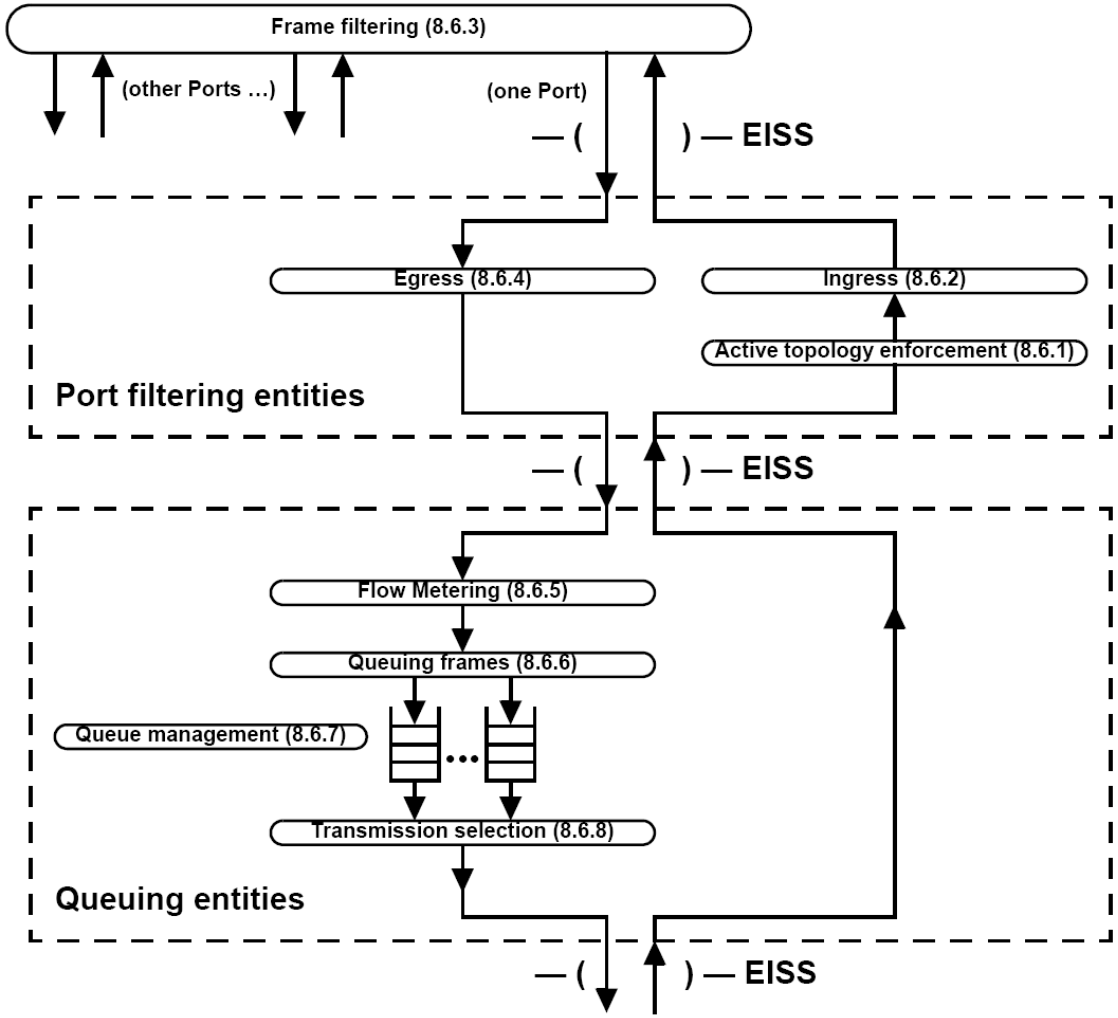


Figure 22-2—Alternate view of Forwarding process

IEEE Std. 802.1ag-2007

Subclause 22 CFM in systems

Port filtering entities are separated from queuing entities, and both are placed in the “pants leg”, instead of in the Forwarding Process

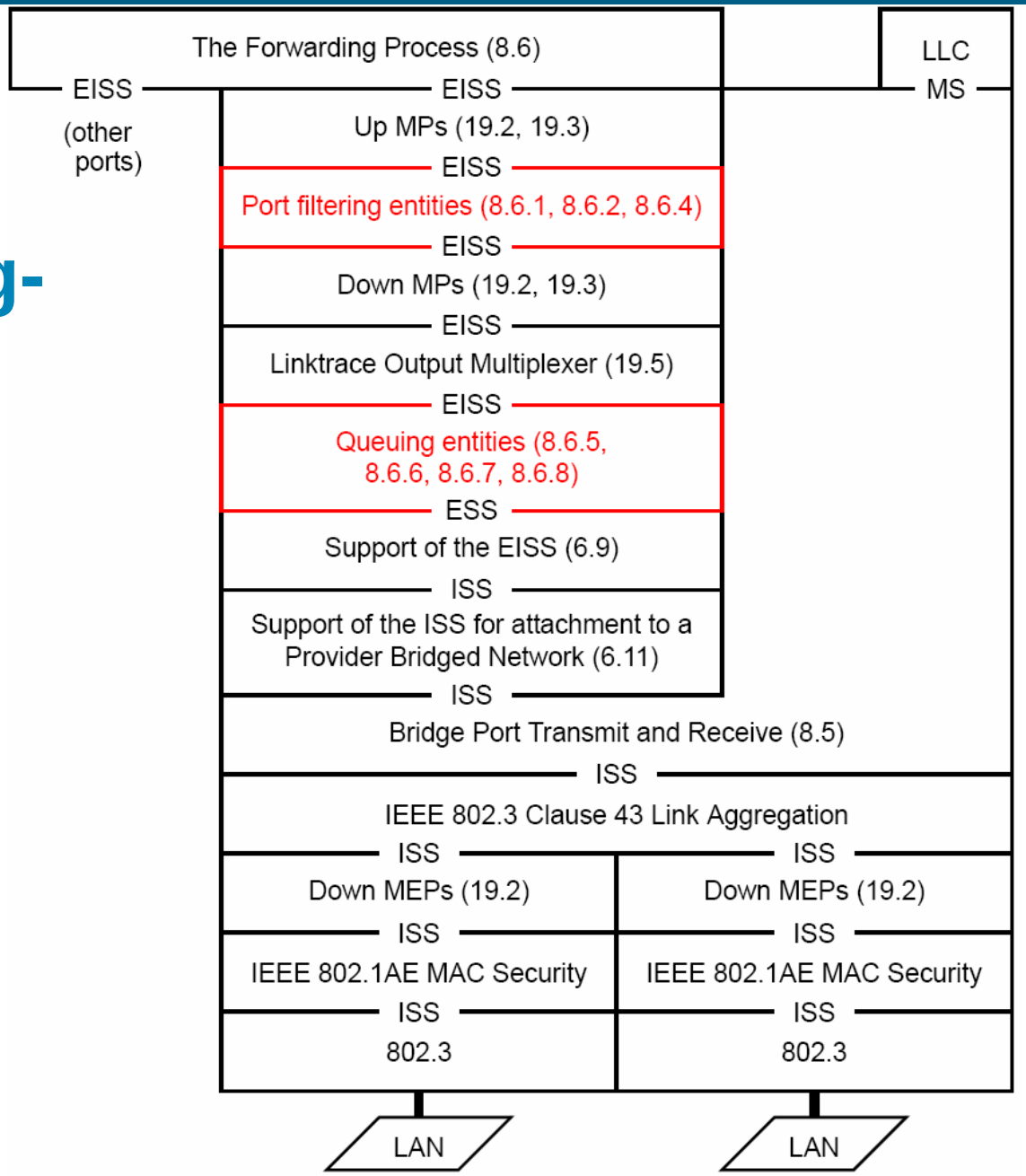


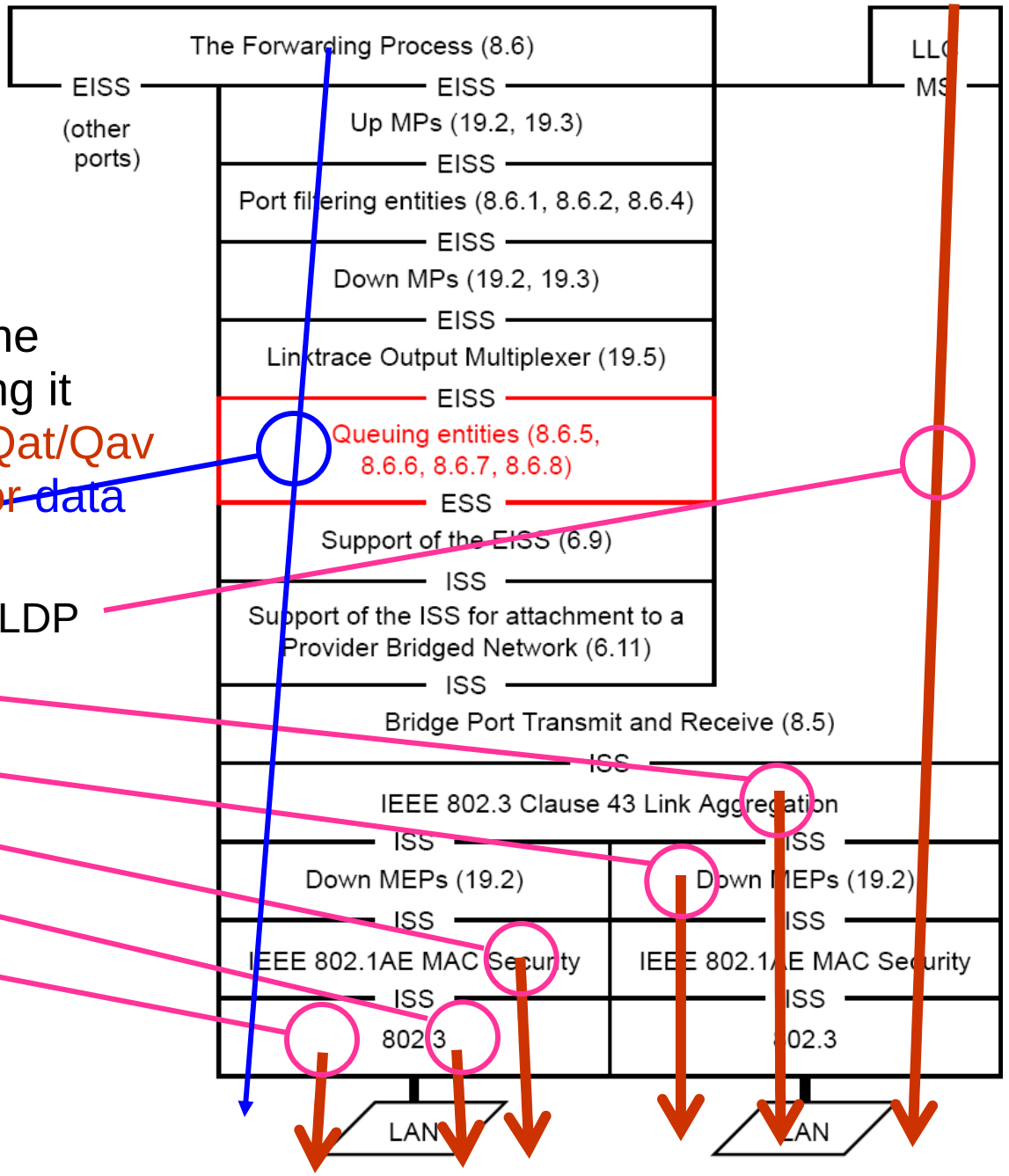
Figure 22-8—Maintenance Point placement relative to other standards

Two Projects

- P802.1Qat specifies a protocol to reserve bandwidth.
- P802.1Qav specifies transmission selection algorithms for the queues to support bandwidth-reserved data streams.
- P802.1au specifies a protocol to generate congestion notifications.
- P802.1Qaz specifies transmission selection algorithms for the queues to support congestion-controlled data streams.

Problem:

- Many frames bypass the queuing entities, making it impossible for P802.1Qat/Qav to guarantee latency for data streams:
  - 802.1 xSTP, MxRP, LLDP
  - 802.3 LACP
  - 802.1 CFM
  - 802.1 Security
  - 802.3 OAM
  - 802.1 Pause

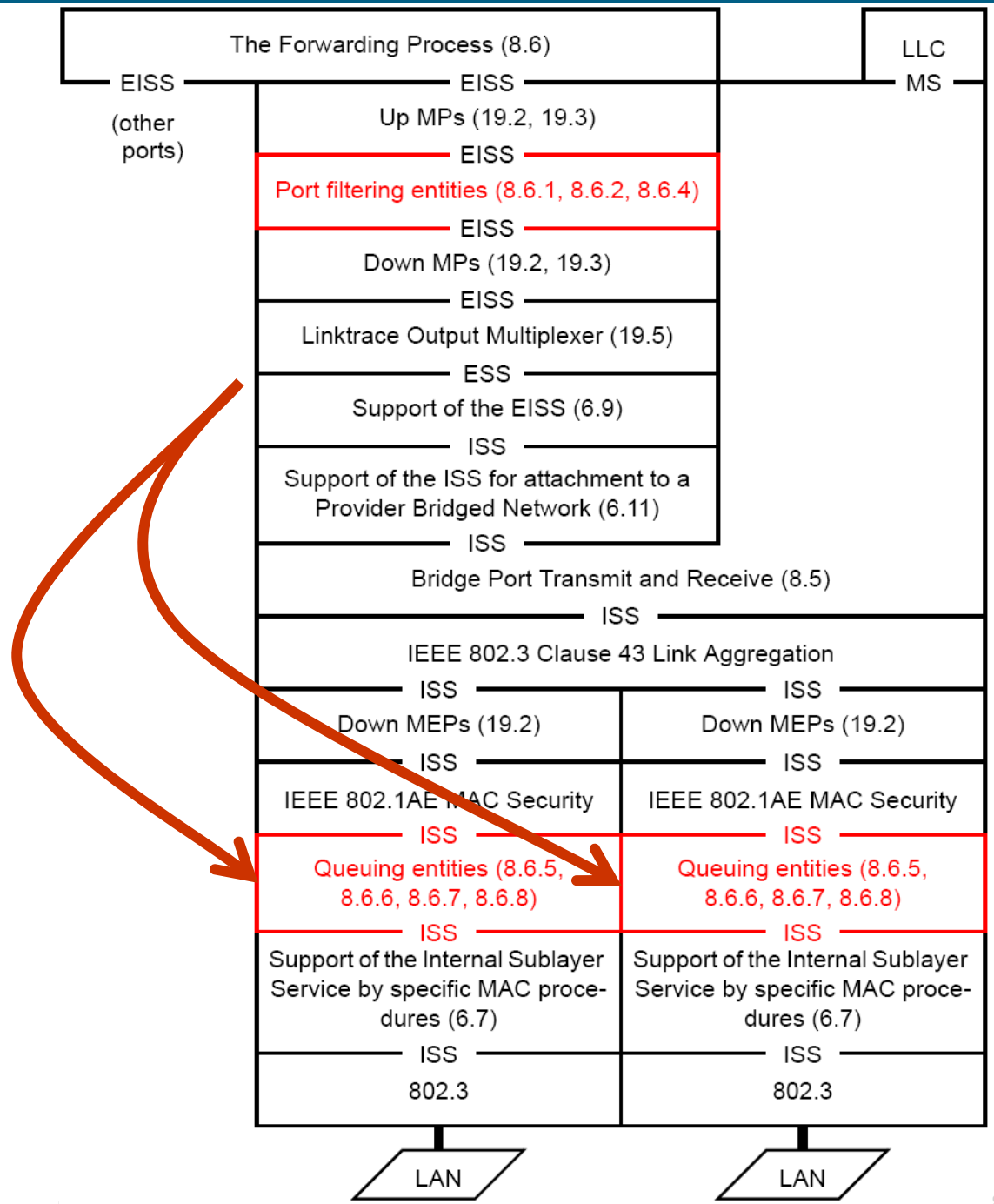


Current queue position kills P802.1Qat/Qav

- Frames can be introduced into the stack a numerous places around and/or below the priority queues.
- Given this situation, it is not wise to leave the difficult task of ensuring the AVB TG's latency guarantees as an exercise to the reader, if we want the guarantees to be met.

If we move all of the queuing and BCN generation down to a lower position in the diagram, P802.1Qat/Qav can guarantee latency.

But, that makes it impossible for P802.1au/Qaz to do Backward Congestion Notification.



Low queue position kills P802.1au/Qaz

- The shim is below the EISS, so has no way of knowing on what VLAN to send the BCN in response to a frame belonging to a flow that needs to be controlled.

What is essential

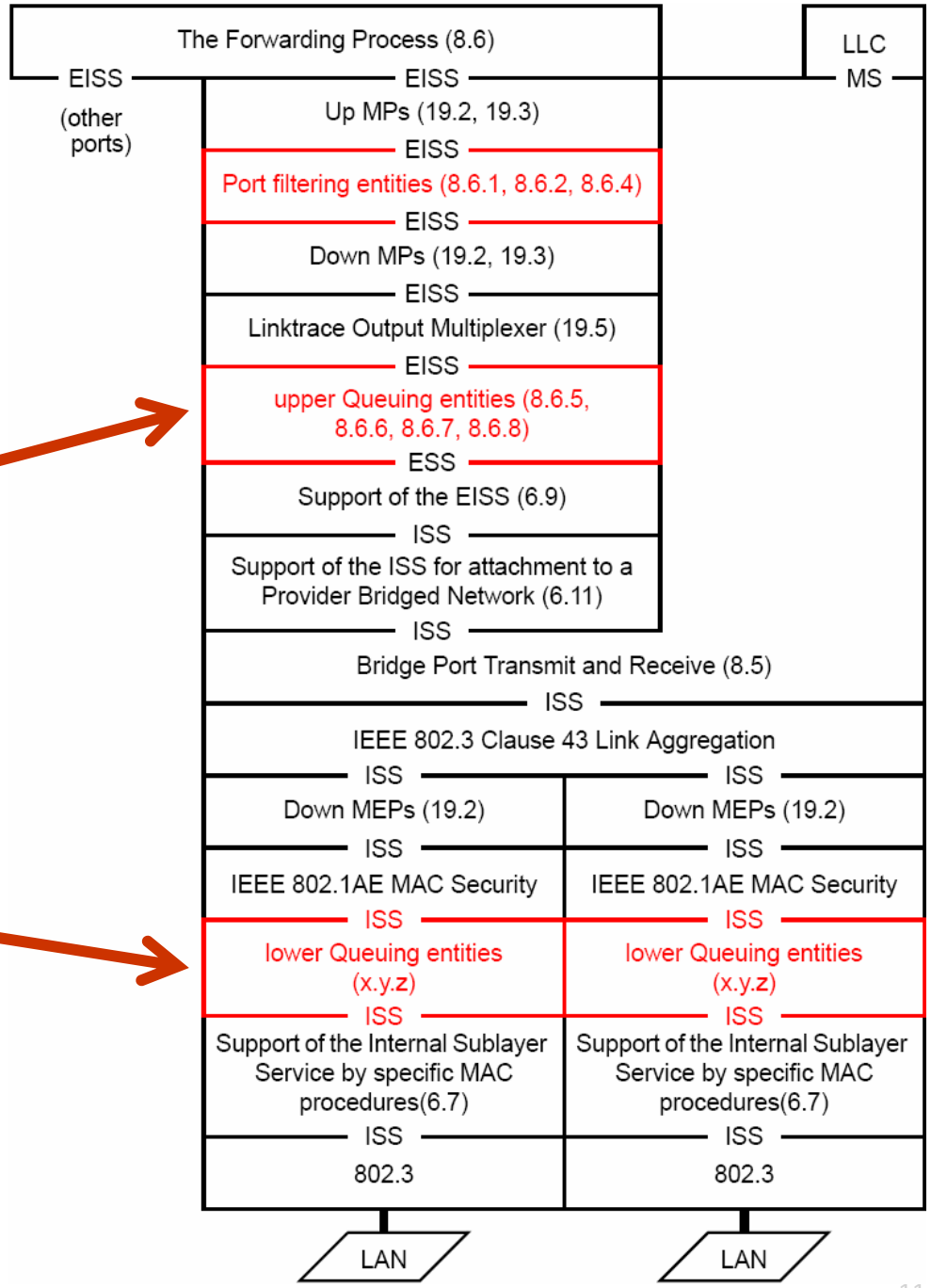
- Any frame emitted by a time-sensitive queue (at least, the P802.1Qat/Qav reserved stream queues) must be emitted by the “6.7 Support of the Internal Sublayer Service by specific MAC procedures”, at the bottom of the stack, in preference to any frame generated by any other source above it in the stack.

  (Better still, 802.3ah OAM should be enqueued, as well, but that’s another story.)

- P802.1aq BCN must be generated at the upper position.

Solution 1: We could define two places for queues.

- Queues can go here.
- Queues can go here



How is this a "solution"?

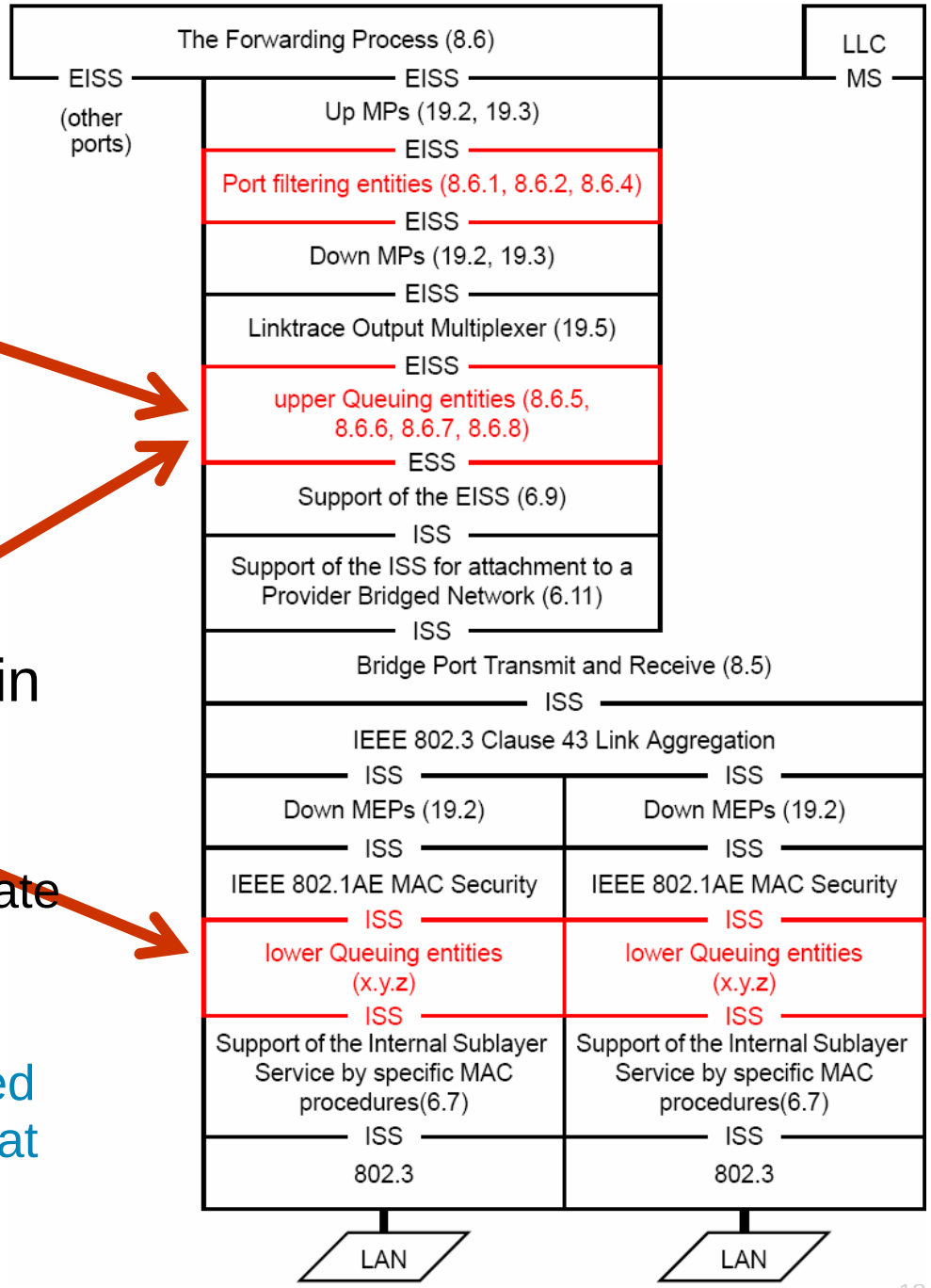
- P802.1au/Qaz queues must go here.

  BCN generation works properly in this position.

- P802.1Qat/Qav rate shaper queues could go in either place.

  But data frames introduced into the stream below the rate shaper queues will affect worst-cast latency.

  So, at least those introduced frames must be enqueued at the lower position.



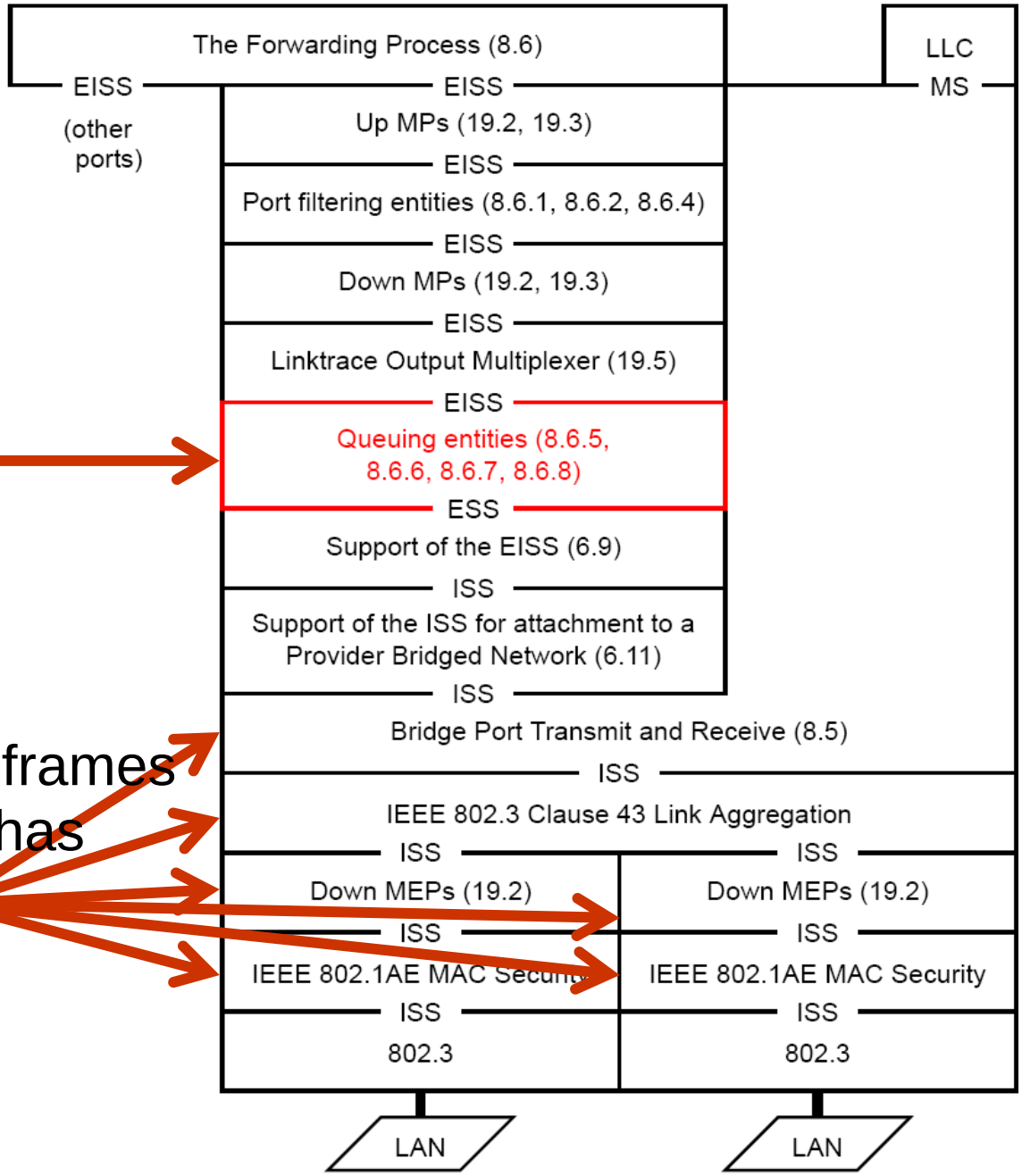
Solution 1: meets the essentials

- If the lower queues are drained by user_priority, but the frames emitted by time-controlled queues are:
  - All placed in the same queue;
  - That queue is drained ahead of any other queue; and
  - No frames at the time-controlled priority levels are inserted below the time-controlled queues;

  Then latency guarantees can be met.
- For example, if user_priorities 4 and 5 go to two separate rate-controlled queues at the upper position, then at the lower position:
  - Frames at all other priorities (from all sources) are placed into some number of other queues.
  - The priority 4 and 5 queue is drained ahead of all other queues, so is almost always empty.

Solution 2: We could define selection rules.

- Queues go here.
- Every “Y” at which frames can be introduced has selection rules.



Solution 2: meets the essentials

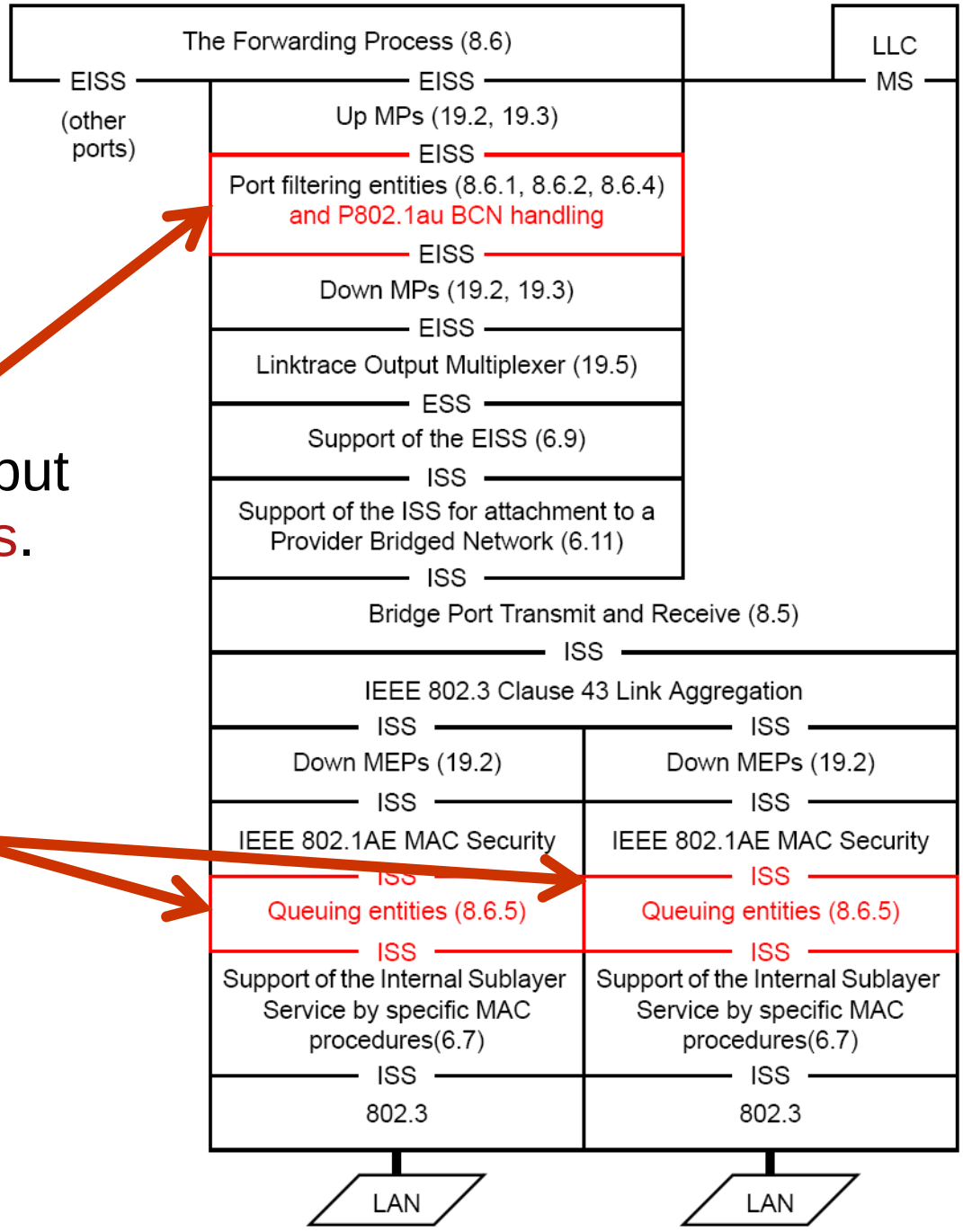
- All queues are at the upper position.
- But, at every point below the queues where frames are introduced into the stream, a rule is applied that says, “Frames at rate-controlled user_priority levels are:

  Never introduced; and

  Are always accepted in preference to frames at any other priority level.
- The problem with this formulation is that, since the (E)ISS is a service definition, there is no such idea as, “frames offered to a Y-shim’s two upper SAPs at the same time”, and hence, it is not obvious to everyone how to specify the necessary rule.

Solution 3: We could move all the queues down.

- No queues go here, but BCN generation does.
- All queues go here



Solution 3: meets the essentials

- All queues are at the lower position:

  Rate controlled queues are drained ahead of non-rate controlled queues.

  The queue configuration mechanism knows on which physical link a given reserved stream will be transmitted, so knows which link's queues to configure to carry that stream.

  But, this places burdens on P802.1Qat to include the upper-layer information needed at reservation time to determine which run-time link a stream will use.

  And, does it know, for a given P802.1au congestion controlled frame, on what physical link the frame will go out?

- The BCN generation is at the upper position, and uses management information to know whether a given frame should be sampled, and/or a BCN generated.

All these ideas have problems

- Basic problem: All queues at the top leaves undefined how you actually make this work.
- Solution 1: BCN at the top and others at the bottom is a complex solution that is unlikely to match any implementation.
- Solution 2: All queues at the top and selection rules at every Y makes very large changes to the document, and is logically unpleasing.
- Solution 3: All queues at the bottom and BCN operates via the management interface begs the question of how BCN or P802.1Qat knows on what aggregated link a frame will be output.