FORCE10™

Higher Speed Ethernet

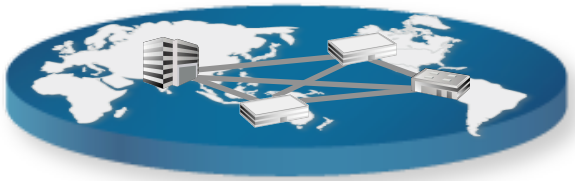
Agenda

- 100 GbE Technology Requirements
- IEEE P802.3ba Standards Update

What is Driving the Need for 40 GbE and 100 GbE?

- Research network trends
  - 10 GbE WANs
  - Cluster and grid computing
- Data center trends
  - Lots of GbE and 10 GbE servers
  - Cluster computing
    - High-end servers pushing almost 8 Gbps
    - Storage flows > 1 Gbps
- ISP / IX trends
  - High bandwidth applications
  - 10 GbE peering
- *Something* has to aggregate all those 10 GbE links
  - LAG is an interim solution

Government/Research



Data Center

Service Provider

Higher Speeds Drive Density – Everyone Benefits!

- Even if you don't need 100 GbE, you still benefit
- 100 GbE technology will drive 10 GbE port *density up* and *cost down*
  - Just as 10 GbE did for GbE
- Assuming routers have the switching capacity we can support these line-rate combinations on a single line card
  - 1 x 100 GbE port
  - 10 x 10 GbE ports
  - 100 x 1 GbE ports
  - And even more oversubscribed port density…

Higher Speeds Drive Density – Peering Benefits!

- This is great for peering!
- 100 GbE IX peering fabrics
- 100 GbE WAN PHY (cheaper than SONET)
- Lower equipment cost with higher densities for peering
  - Cheaper 10 GbE ports
  - Really cheap GbE ports

Higher Speeds Drive Switch/ Router Requirements

- New capacity and density drive architectural requirements needed to support 100 GbE
- Massive hardware and software scalability
  - >200 Gbps/slot fabric capacity for any reasonable 100 GbE port density (local switching capacity is useless here!)
  - Support for several thousand interfaces
  - Multi-processor, distributed architectures
- Really fast packet processing at line-rate
  - 100 GbE is ~149Mpps or 1 packet every 6.7ns

  (10 GbE is *only* ~14.9Mpps or 1 packet every 67ns)

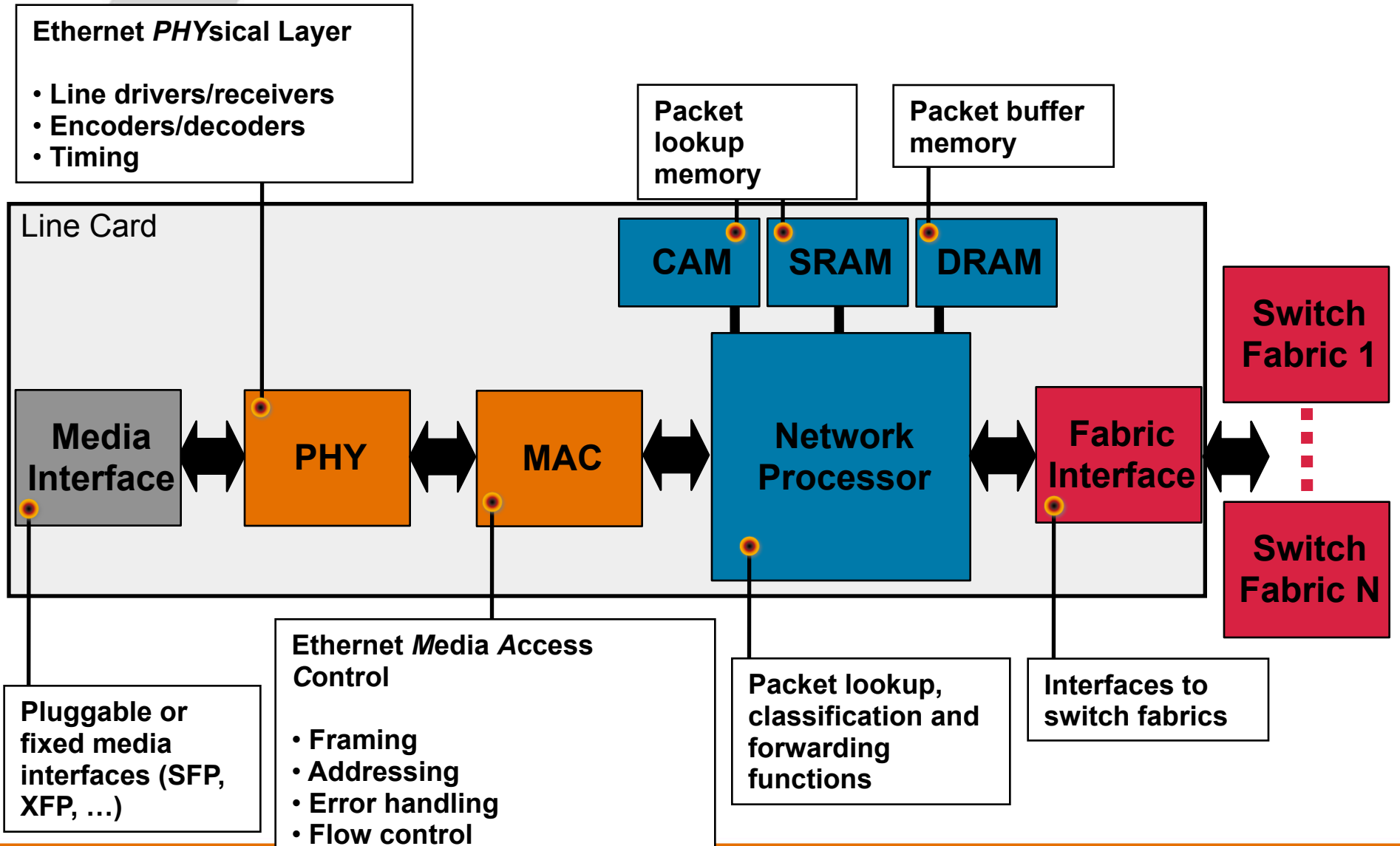
Higher Speeds Drive Switch/ Router Requirements

- Complete system resiliency
  - Hitless forwarding at Tbps speeds
  - Redundant hardware
  - HA software, hitless upgrades
  - DoS protection and system security
- Chassis design issues
  - N+1 switching fabric
  - Channel signaling for higher internal speeds
  - Clean power routing architecture
  - Reduced EMI
    - Conducted through power
    - Radiated into air
  - Cabling interfaces

Anatomy of a 100 Gbps Solution: Slot Capacity

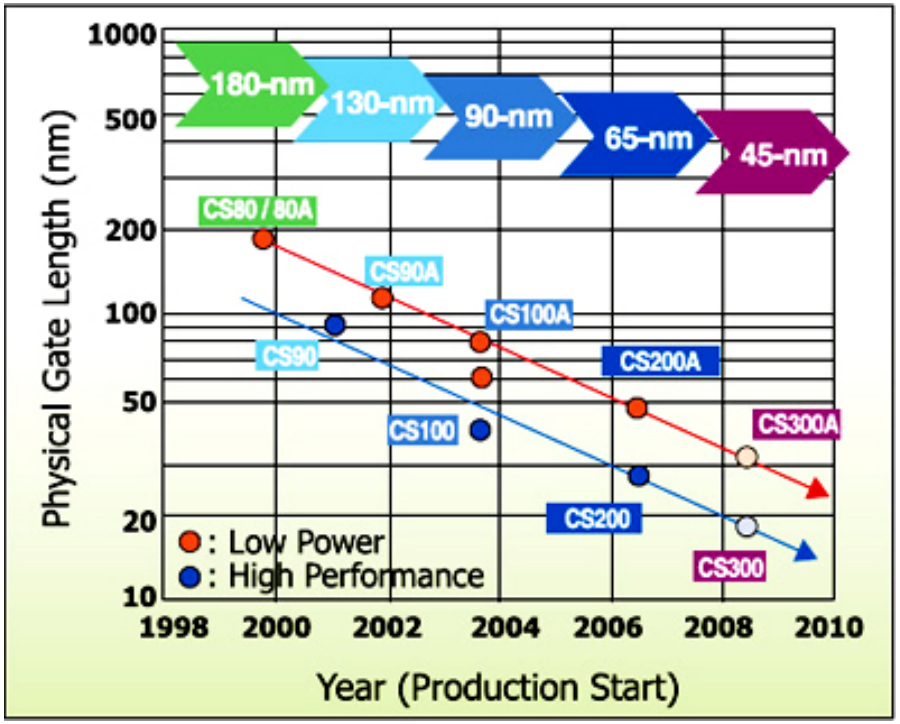
Year System Introduced	Full Duplex Raw Slot Capacity
2000	40 Gbps
2004	60 Gbps
2006 – 2007 (in design now)	120 Gbps
2009	500 Gbps Required For Reasonable 100 GbE Port Density

Anatomy of a 100 Gbps Solution: Line Card Components



Anatomy of a 100 Gbps Solution: ASIC Selection

- Typical 10 Gb/s MAC today
  - 64-bit wide
  - 156.25MHz
  - 90nm or 130nm silicon
- 90nm process geometry vs 130nm
  - 100% more gates
  - 25% better performance
  - ½ power consumption
- 65nm and 45nm process geometry
  - Expected to be available by 2010
  - Will give us more design options



Anatomy of a 100 Gbps Solution: Memory Selection

- Memory technology is usually the last thought, but these are critical components
- Advanced Content-Addressable Memory (CAM)
  - Less power per search
  - Need 4 times more performance assuming 2 x 100 GbE ports on each network processor
  - Enhanced flexible table management schemes
- Memories
  - Use DRAMs with SERDES interfaces when performance allows to conserve cost
  - Quad Data Rate III SRAMs for speed (PS and Xbox gaming industry is driving this technology)
- Force10 works with JEDEC (Joint Electron Device Engineering Council) to advance serial memory technology
  - JEDEC is an international semiconductor engineering standardization body

Feasible Technology for 2010: Media Interfaces

- Increasing port densities and higher speed interfaces need new media interfaces
- Force10 participates in MSAs to design new pluggable media
  - QSFP: quad SFP (four interfaces in one SFP)
  - SFP+: smaller 10 GbE media
- New optics are needed for higher speed Ethernet
  - SONET/SDH OC-768 optics available today but cost close to 30 x 10 GbE optics

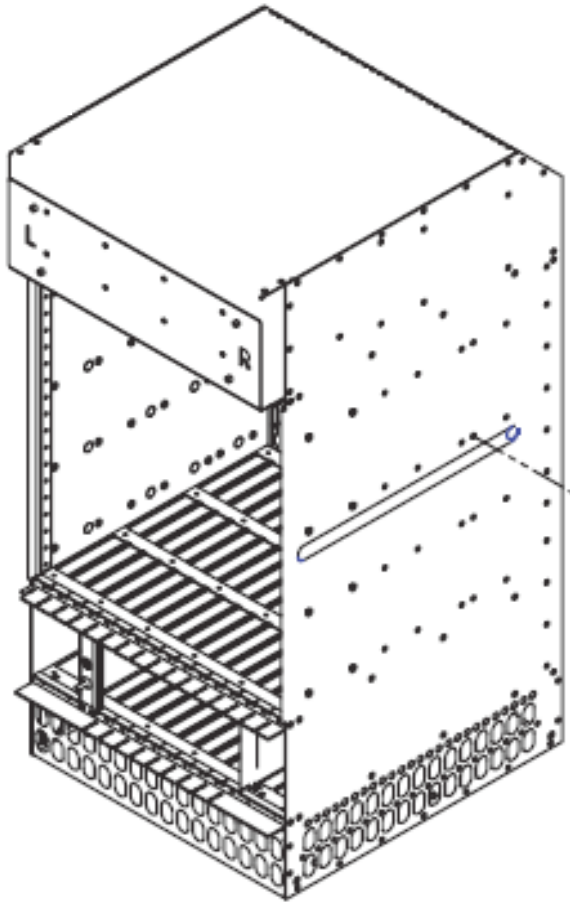

SFP


XFP


QSFP

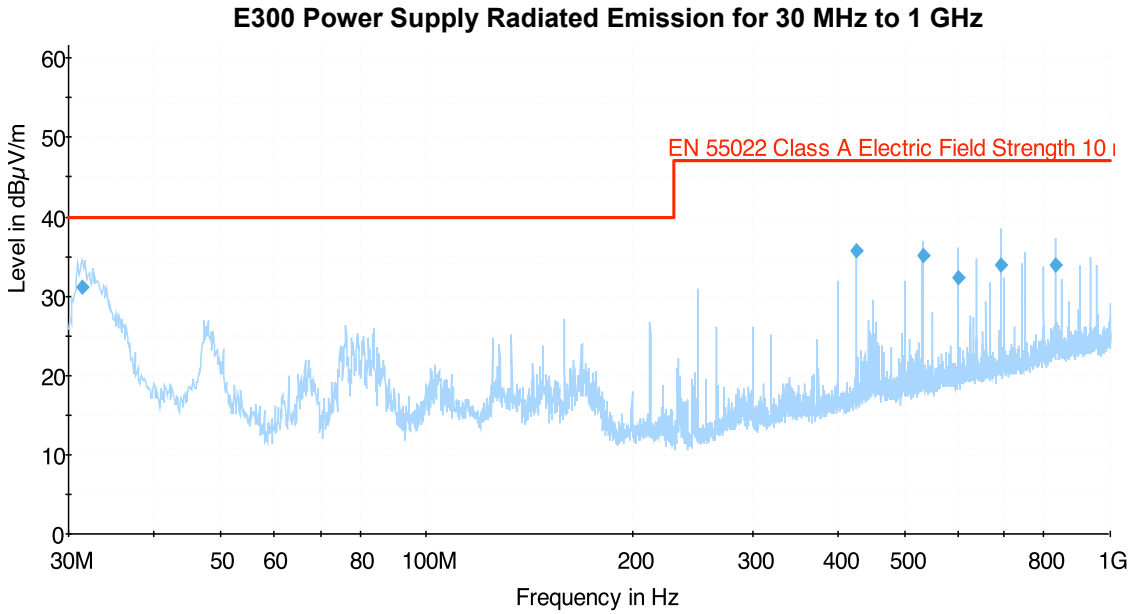
Anatomy of a 100 Gbps Solution: Chassis Design

- Chassis design issues to consider
  - Backplane and channel signaling for higher internal speeds
  - Lower system BER
  - Connectors
  - N+1 switch fabric
  - Reduced EMI
  - Clean power routing architecture
  - Thermal and cooling
  - Cable management
- All design aspects must also meet local regulatory standards



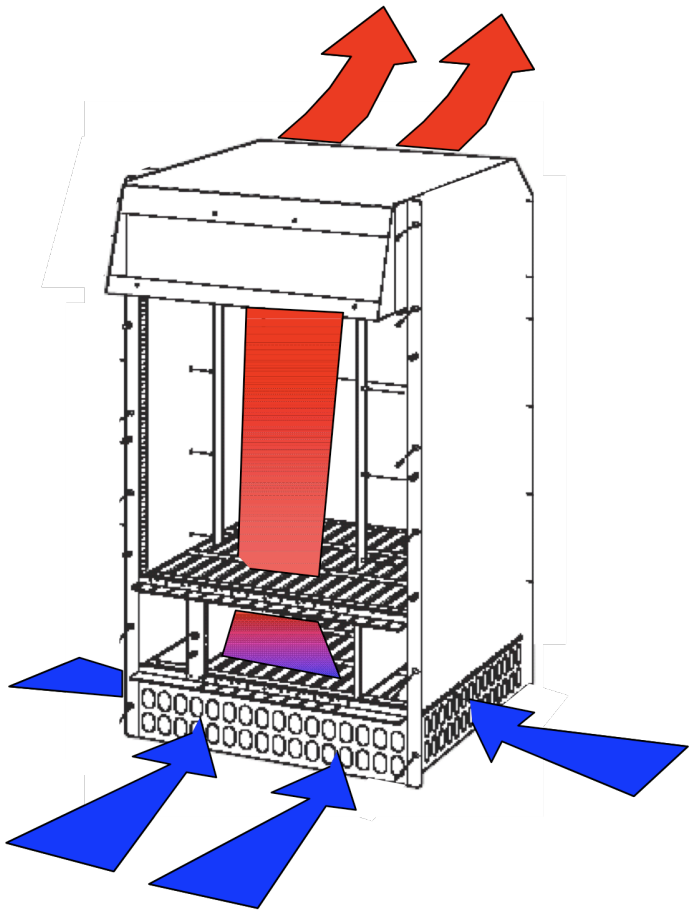
Anatomy of a 100 Gbps Solution: Designing for EMI Compatibility

- EMI is Electro-magnetic Interference, ie noise
- Too much noise interfere can cause bit errors, so we need to minimize the EMI
- Concerned about two types of EMI
  - Conducted EMI through power
  - Radiated EMI through air



Anatomy of a 100 Gbps Solution: Thermal and Cooling

- System is already designed for cooling >400W/slot
- Front to back filtered airflow for carrier deployments
- 1+1 cooling redundancy
- Active health monitoring for fan status reporting



Agenda

- 100 GbE Technology Requirements
- IEEE P802.3ba Standards Update

The Push for Standards: Interplay Between the OIF & IEEE

- OIF defines multi-source agreements within the Telecom Industry
  - Optics and Electronic Dispersion Compensation (EDC) for signal integrity
  - SERDES definition
  - Channel models and simulation tools
  - … the components inside the box
- IEEE 802 covers LAN/MAN Ethernet
  - 802.1 and 802.3 define Ethernet over copper cables, fiber cables, and backplanes
  - 802.3 leverages efforts from OIF
  - …the signaling and protocols on the wire

Per IEEE-SA Standards Board Operations Manual, January 2005

At lectures, symposia, seminars, or educational courses, an individual presenting information on IEEE standards shall make it clear that his or her views should be considered the personal views of that individual rather than the formal position, explanation, or interpretation of the IEEE.

Recent Developments

- Defined architectures and nomenclature (100GBASE-ER4, etc)
- Adopted baseline proposals for all objectives
- Finished Draft 1.0
- On schedule: the 40 GbE and 100 GbE standards will be delivered together in June 2010
- Crystal ball says there is already demand for other PMDs outside the scope of 802.3ba (100 GbE serial, etc)
  - Standard defines a flexible architecture that enables many implementations as technology changes
  - Expect MSAs

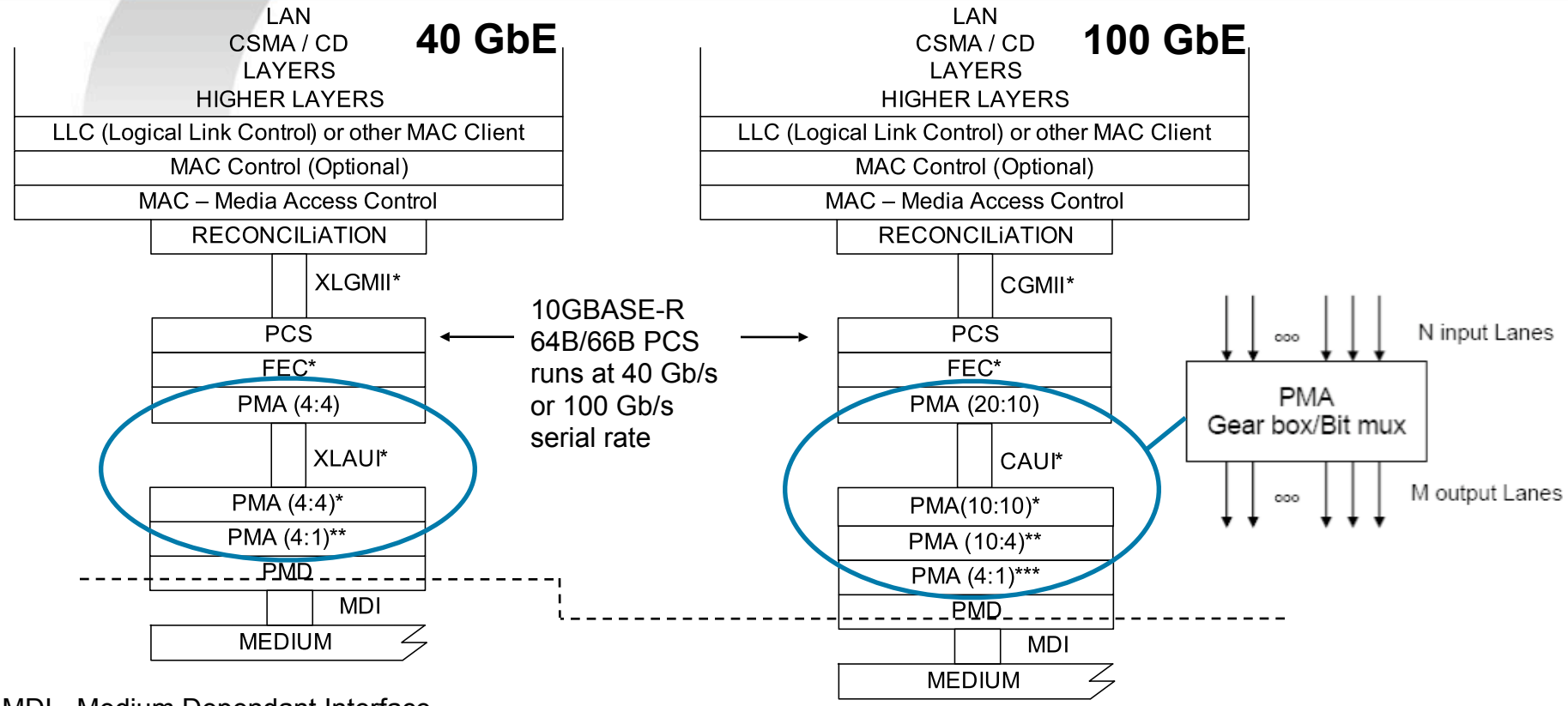
Draft 1.0 is Here!

- Significant milestone, captures all objectives but not technically complete yet (292 pages long)
- Draft 2.0 will technically complete for WG ballot
- Gives people an idea of the technology and solutions for them to scope out implementation
- Technical specifications could still be modified and clarified as people start to implement
  - Looks good on paper but stuff will come up
  - Crossing Is and dotting Ts
- Careful balance between marketing timing and standards process/technical compliance

Summary of Reach Objectives and Physical Layer Specifications

Reach	40 GbE	100 GbE	Solution
1m Backplane	40GBASE-KR4	✘	4 x 10 Gb/s (reuse 10GBASE-KR)
10m Copper Cable	40GBASE-CR4	100GBASE-CR10	n x 10 Gb/s (reuse 10GBASE-KR)
100m OM3 MMF	40GBASE-SR4	100GBASE-SR10	n x 10 Gb/s
10km SMF	40GBASE-LR4	100GBASE-LR4	4 x 10 Gb/s and 4 x 25 Gb/s
40km SMF	✘	100GBASE-ER4	4 x 25 Gb/s

Overview of Architecture



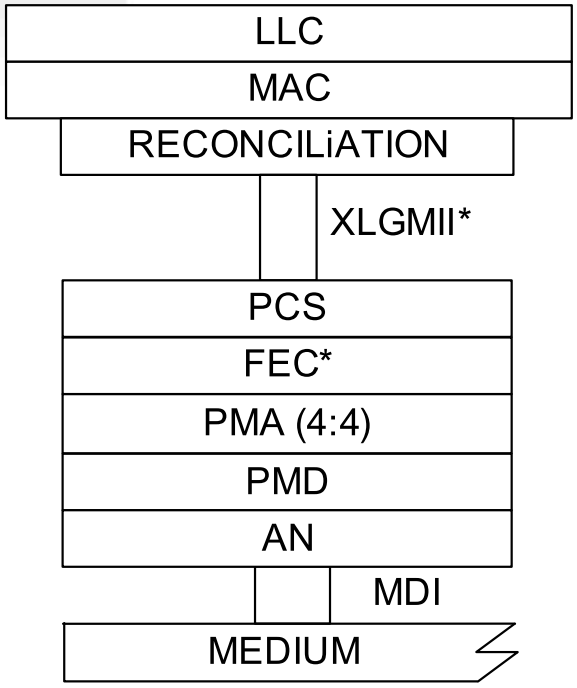
MDI - Medium Dependant Interface
PCS - Physical Coding Sublayer
PHY - Physical Layer Device
PMA - Physical Medium Attachment
PMD - Physical Medium Dependent
CGMII - 100 Gigabit Media Independent Interface
XLGMII - 40 Gigabit Media Independent Interface
CAUI - 100 Gigabit Attachment Unit Interface
XLAUI - 40 Gigabit Attachment Unit Interface

- Consistent with previous Ethernet rates, extension to 40 Gb/s and 100 Gb/s data rates
- Same frame format
- No changes to the MAC
- New interface definitions

"Provide appropriate support for OTN"

- No WIS (WAN Interface Sublayer) like we have for 10 GbE
- Define transparent mapping of 40 GbE into existing ODU3
  - Transcoding to be specified by the ITU-T SG15
  - Coordination between ITU-T SG15 and IEEE on control block types
- Define new ODU4 tier for 100 GbE
- Link fault signaling for 802.3ba Ethernet over OTN is feasible

40GBASE-KR4: 1m Backplane

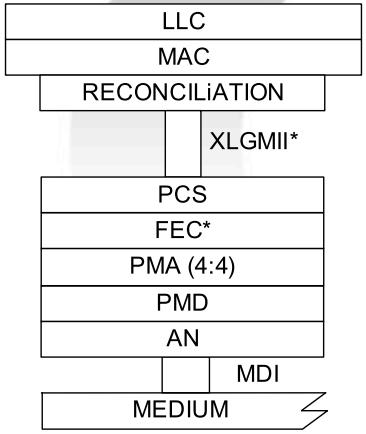


40GBASE-KR4
CSMA / CD Layers

* - Optional

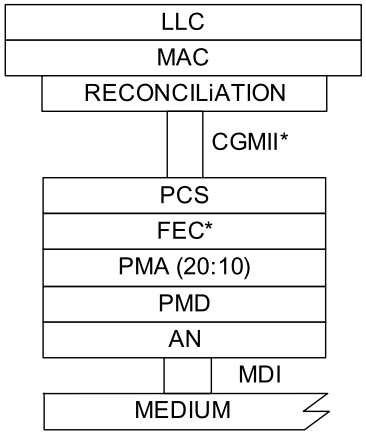
- 1 Gigabit to 40 Gigabit Ethernet solution for backplanes
- Reuses 10GBASE-KR
  - 4 x 10 Gb/s
  - Same Tx / Rx electrical characteristics and testing
- Auto-negotiation
  - Speed
  - Capabilities (FEC, pause ability)
- Influence of EEE (Energy Efficient Ethernet) pending

40GBASE-CR4 and 100GBASE-CR10: 10m Copper Cable



40GBASE-CR4
CSMA / CD Layers

* - Optional



100GBASE-CR10
CSMA / CD Layers

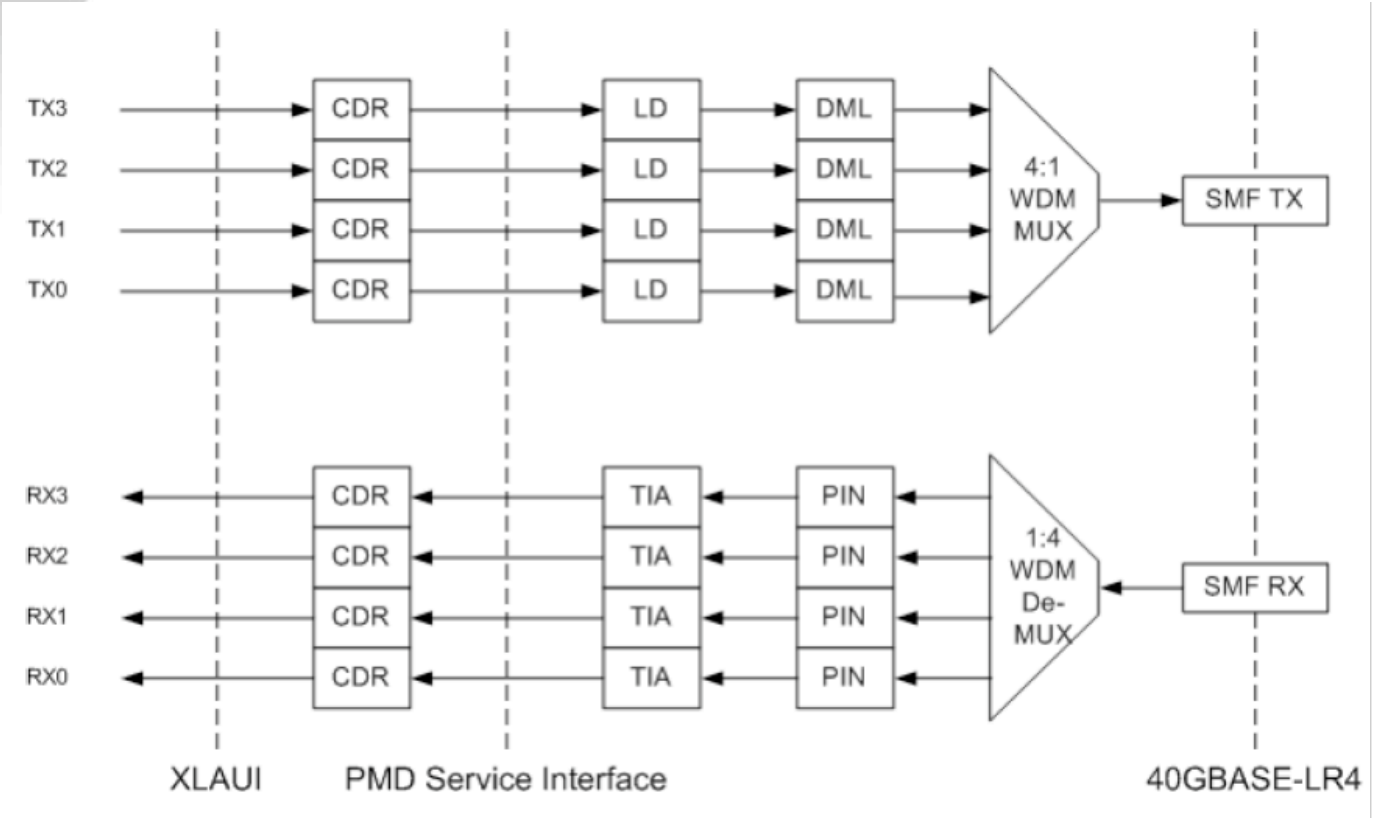
* - Optional

- Reuses 10GBASE-KR
  - 40GBASE-CR4: 4 x 10 Gb/s
  - 100GBASE-CR10: 10 x 10 Gb/s
- Cable parameters based on 10GBASE-CX4
- Auto-negotiation
  - Speed
  - Capabilities (FEC, pause ability)
- 4 x MDI
  - QSFP: enables common form factor for fiber and copper
- 10 x MDI
  - SFF-8092: enables common form factor for fiber and copper

40GBASE-SR4 and 100GBASE-SR10: 100m OM3 MMF

- 40GBASE-SR4
  - 4 Tx / 4 Rx parallel lanes over 4 + 4 OM3 parallel fibers connected to a high density SFF
- 100GBASE-SR10
  - 10 Tx / 4 Rx parallel lanes over 10 + 10 OM3 parallel fibers connected to a high density SFF
- Interest in going beyond 100m
  - How far can the adopted proposal really go?

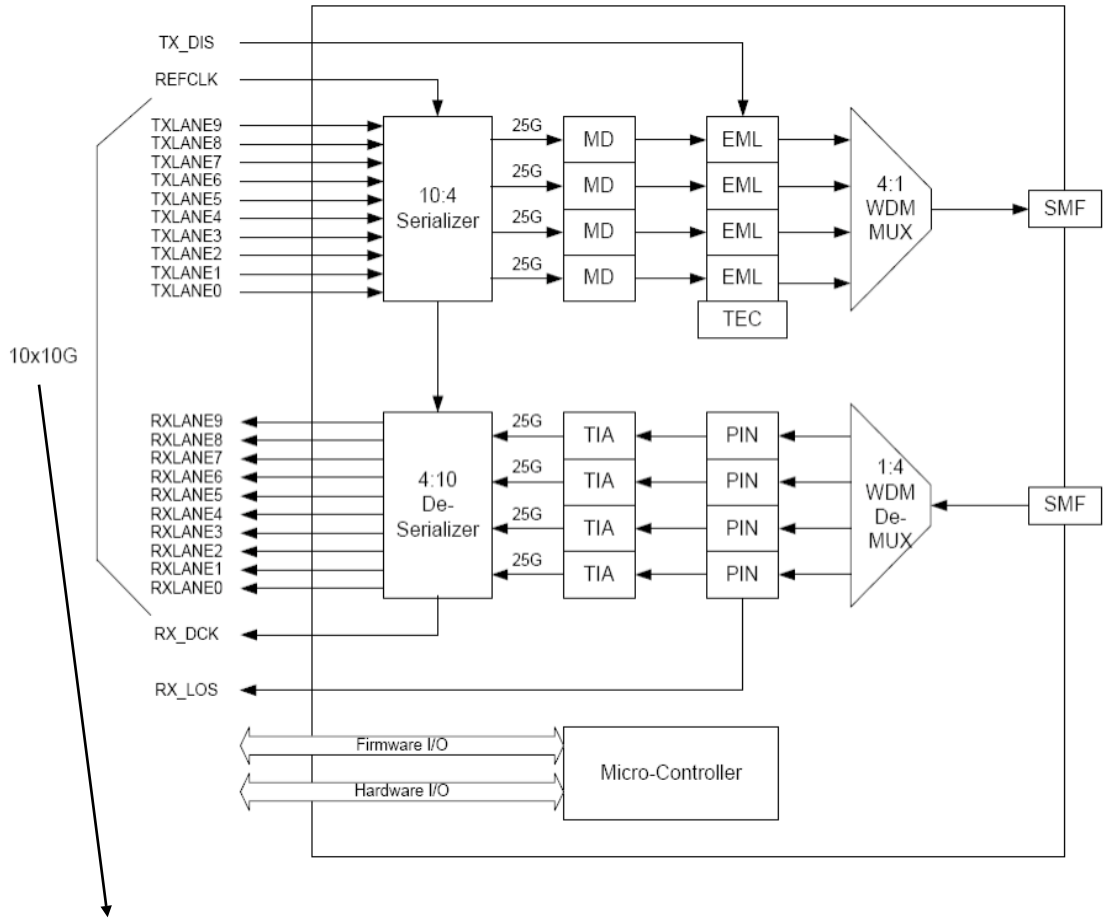
40GBASE-LR4: 10km SMF



- Uses ITU G.694.2 CWDM grid for LAN applications (CWDM) specification
- Wavelengths 1270nm, 1290nm, 1310nm and 1330nm

http://grouper.ieee.org/groups/802/3/ba/public/sep08/cole_01_0908.pdf

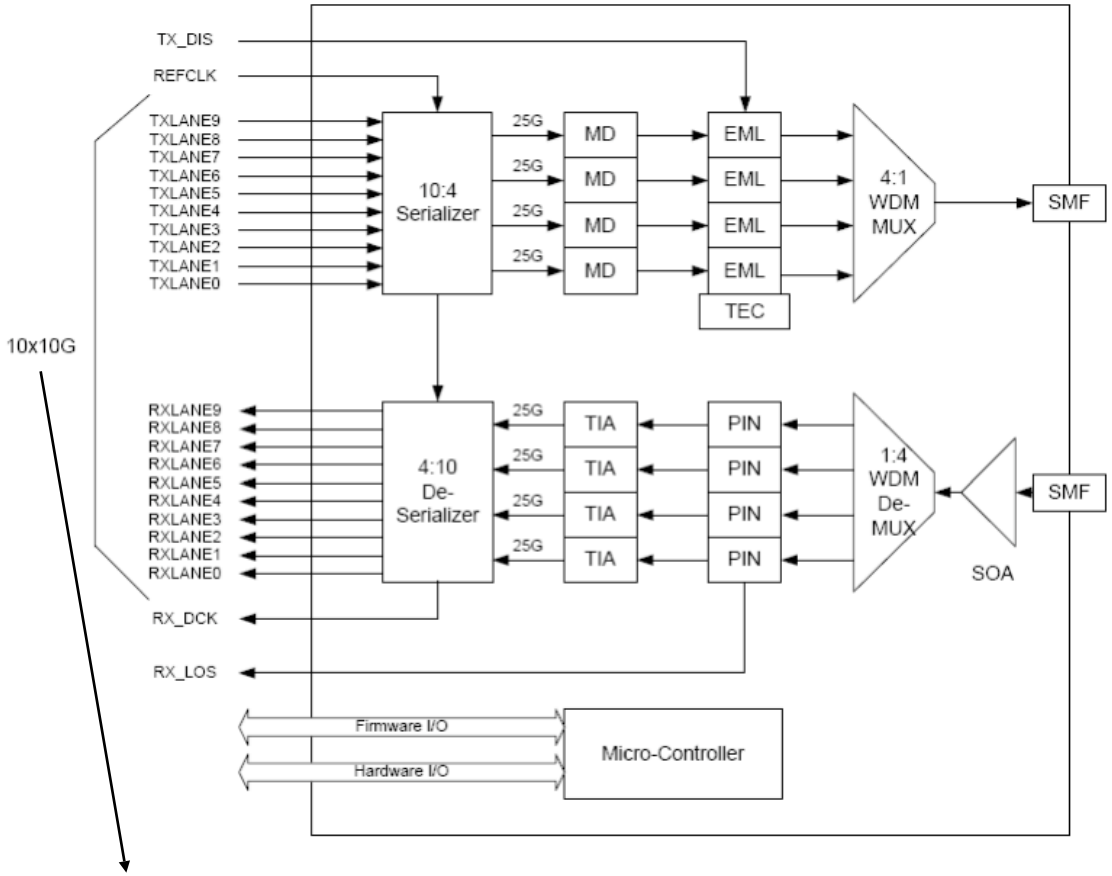
100GBASE-LR4: 10km SMF



- Uses ITU G.694.1 widely spaced DWDM grid for LAN applications (LAN WDM) specification
- Wavelengths 1295nm, 1300nm, 1305nm, and 1310nm
- Same grid as 40km

Development of 25 Gb/s electrical signaling will drive 2nd generation

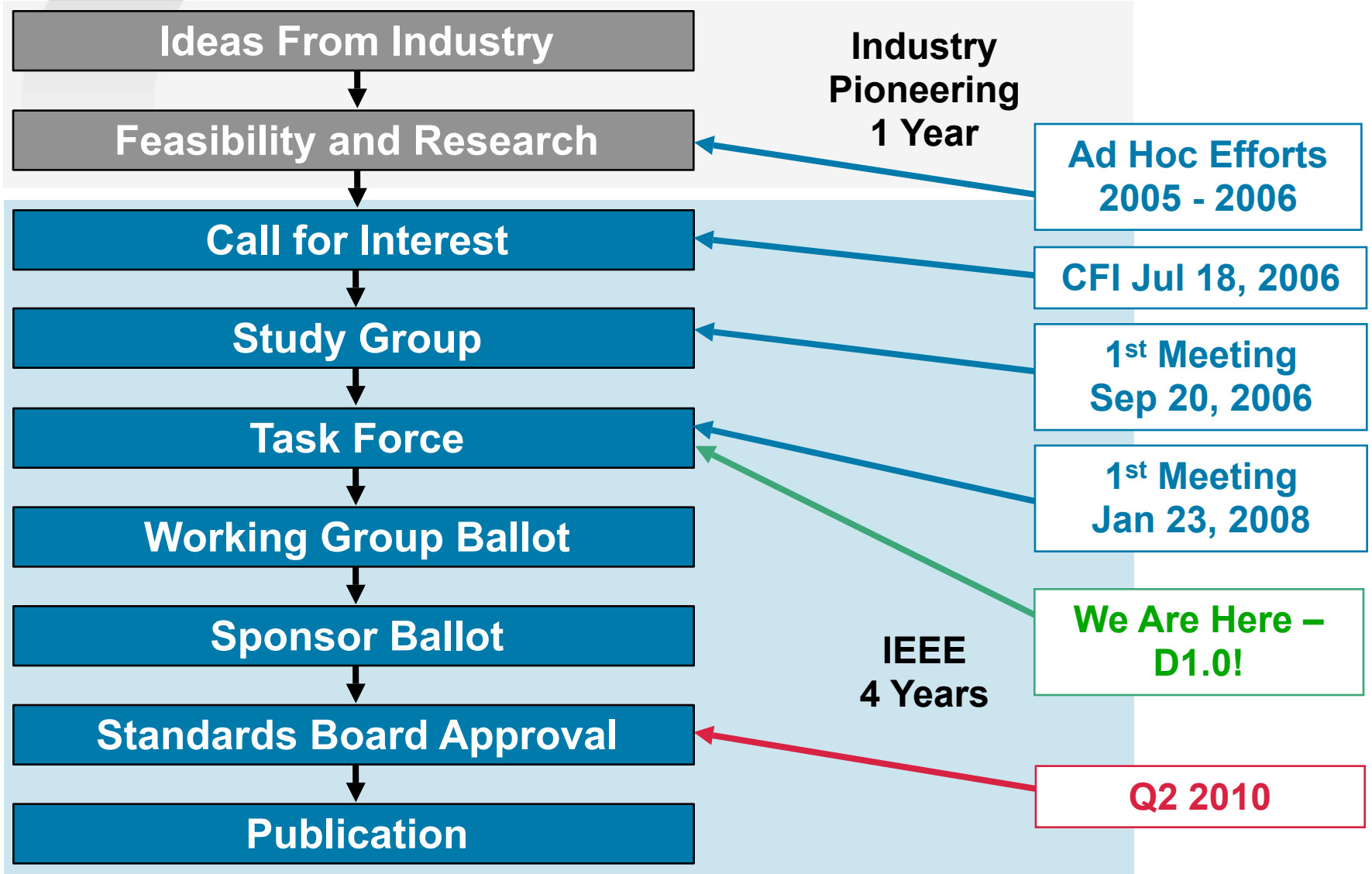
100GBASE-ER4: 40km SMF



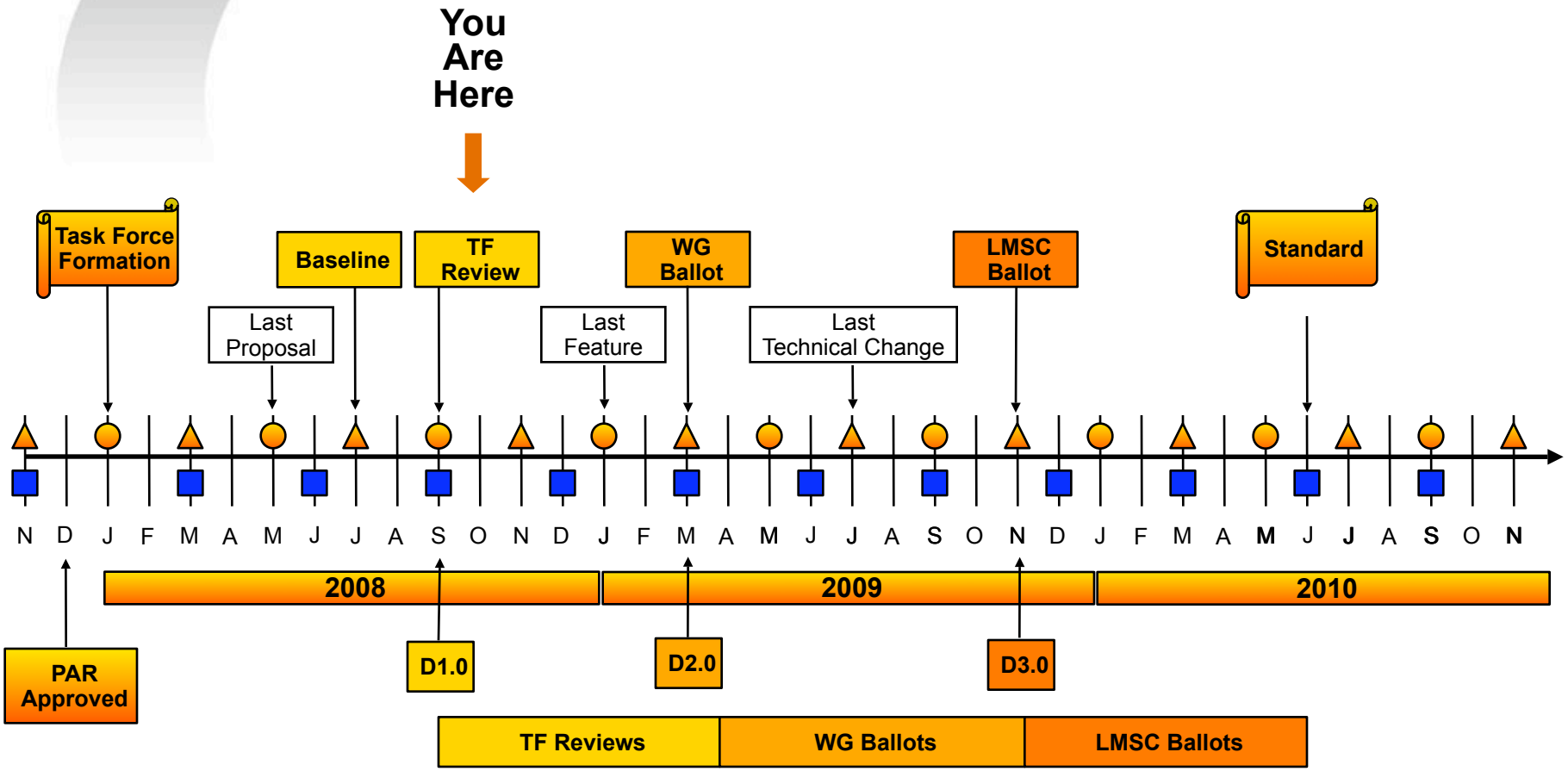
- Uses ITU G.694.1 widely spaced DWDM grid for LAN applications (LAN WDM) specification
- Wavelengths 1295nm, 1300nm, 1305nm, and 1310nm
- Same grid as 10km

Development of 25 Gb/s electrical signaling will drive 2nd generation

Where are we now after ~3 years?



IEEE P802.3ba Task Force Timeline – March 2008 Plenary



Future Meetings for 2008 and 2009

- September 2008 Interim
  - September 15 – 19, Seoul
  - Review Draft 0.9
  - Resolve 40 GbE 10km SMF objective
- After September Interim
  - Generate Draft 1.0
  - Begin Task Force review
- November 2008 Plenary
  - November 9 – 14, Dallas
  - Draft 1.0 comment resolution, work towards Draft 2.0
- January 2009 Interim
  - January 12 – 16, New Orleans
- March 2009 Plenary
  - March 8 – 13, Vancouver

More Information is Here

- http://grouper.ieee.org/groups/802/3/ba/

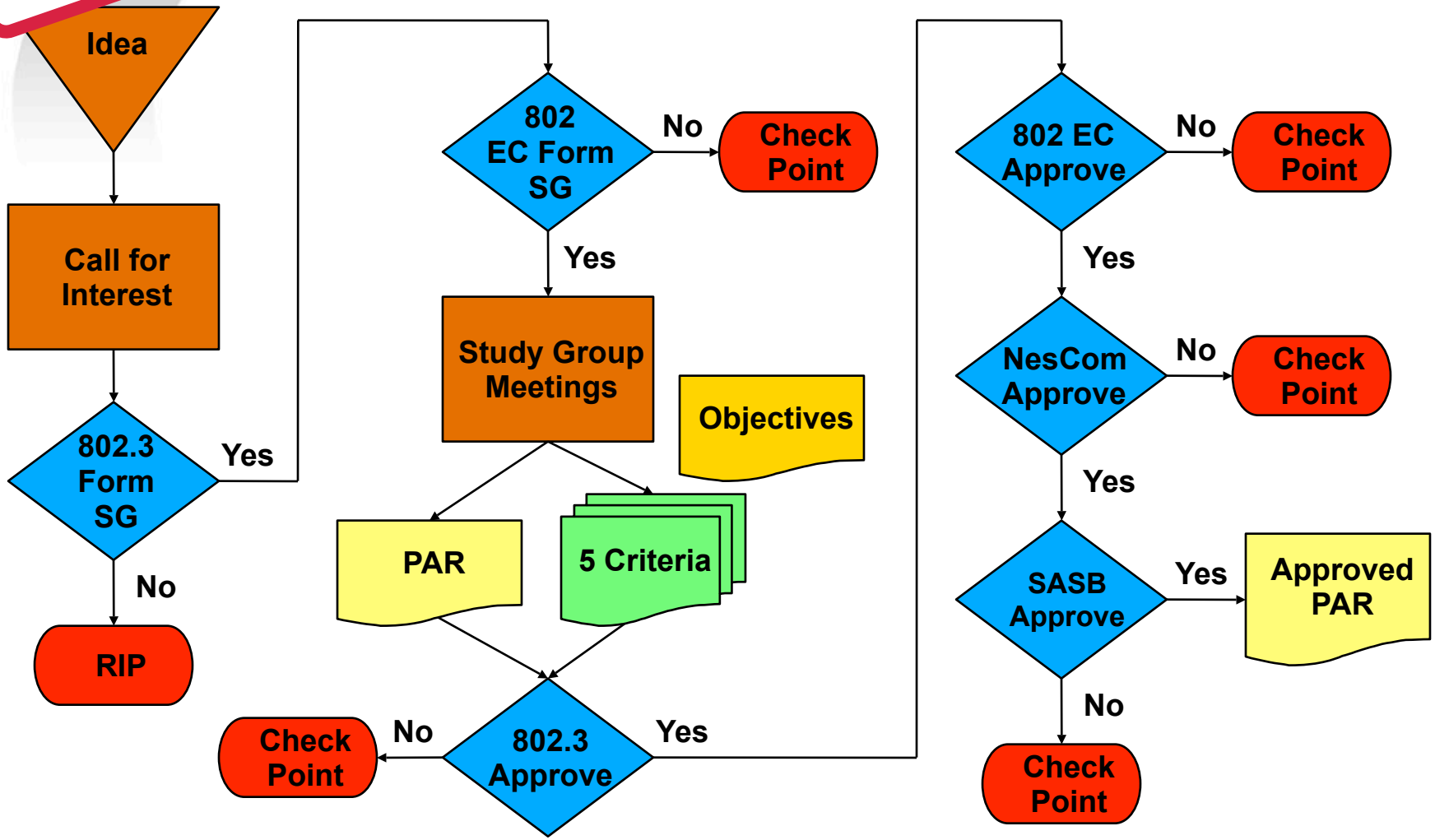
803.3ba Nomenclature Suffix Summary

Speed	Medium		Coding Scheme	Lanes	
	Copper	Optical		Copper	Optical
40G = 40Gb/s 100G = 100Gb/s	K = Backplane C = Cable Assembly	S = Short Reach (100m) L = Long Reach (10km) E = Extended Long Reach (40km)	R = 64B/ 66B Block Coding	n = 4 or 10	n = Number of Lanes or Wavelengths
				n=1 is not required as serial is implied	

Example: 100GBASE-ER4

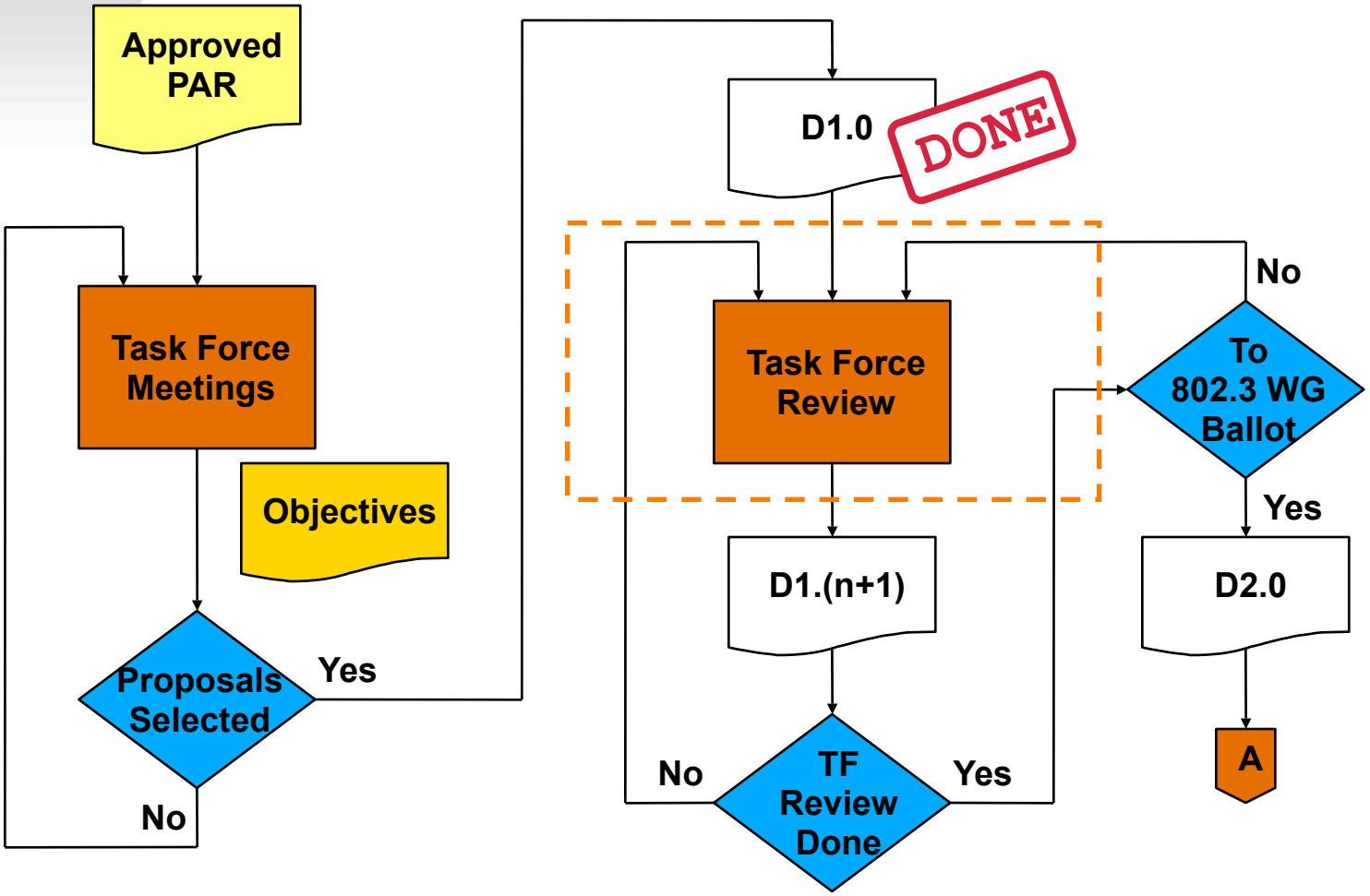
Overview of IEEE 802.3 Standards Process (1/5)- Study Group Phase

DONE

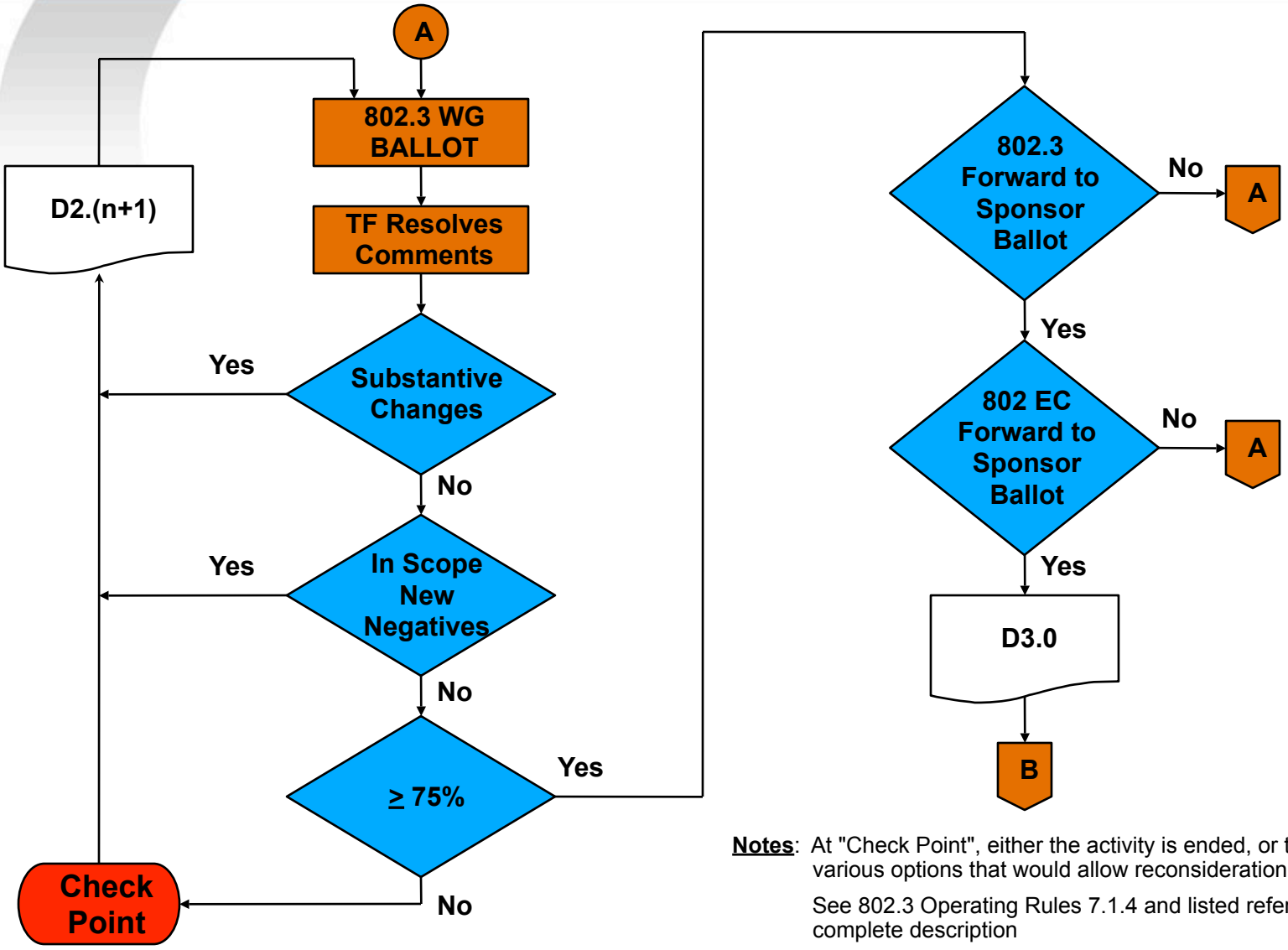


Note: At "Check Point", either the activity is ended, or there may be various options that would allow reconsideration of the approval.

Overview of IEEE 802.3 Standards Process (2/5) - Task Force Comment Phase



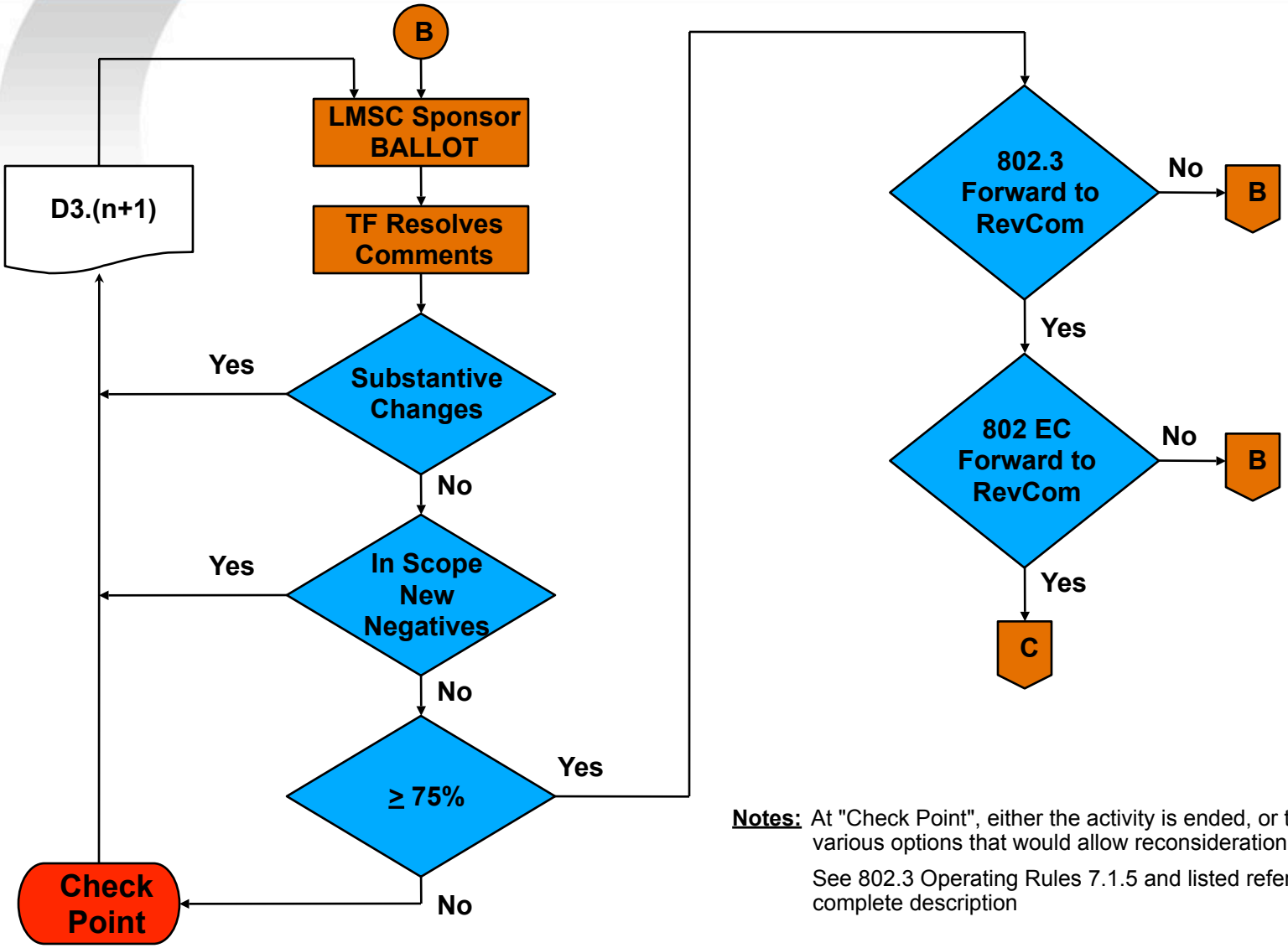
Overview of IEEE 802.3 Standards Process (3/5) - Working Group Ballot Phase



Notes: At "Check Point", either the activity is ended, or there may be various options that would allow reconsideration of the approval.

See 802.3 Operating Rules 7.1.4 and listed references for complete description

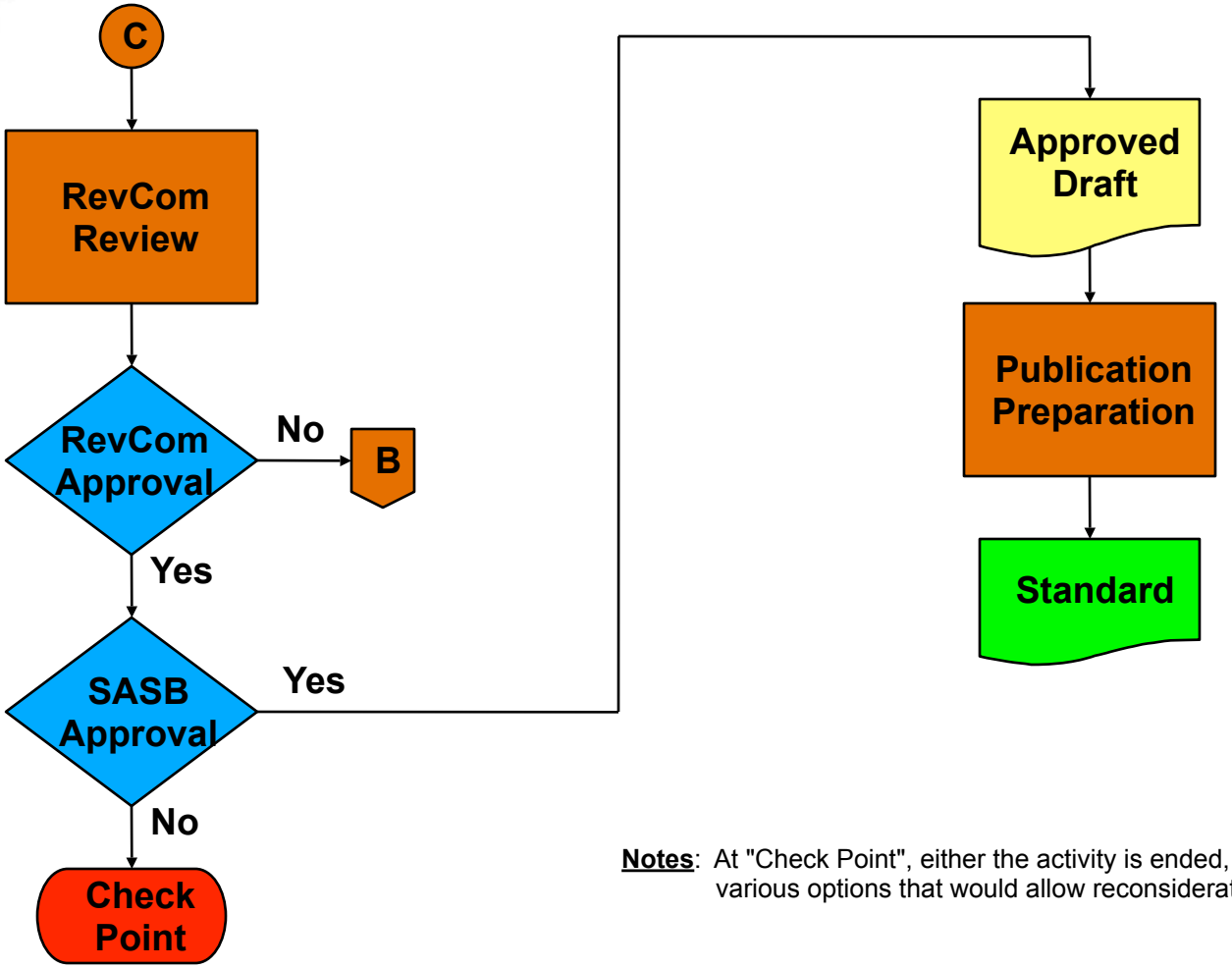
Overview of IEEE 802.3 Standards Process (4/5)- Sponsor Ballot Phase



Notes: At "Check Point", either the activity is ended, or there may be various options that would allow reconsideration of the approval.

See 802.3 Operating Rules 7.1.5 and listed references for complete description

Overview of IEEE 802.3 Standards Process (5/5) - Final Approvals / Standard Release



Notes: At "Check Point", either the activity is ended, or there may be various options that would allow reconsideration of the approval.

Thank You