Analysis of proposed unidirectional mode

Piers Dawe   Agilent Technologies

Eric Lynskey   UNHIOL

Pat Thaler   Agilent Technologies

Richard Brand   Nortel Networks

Ulf Jönsson   Ericsson

Orlando, March 2004

Introduction

- P802.3ah (EFM) D3.1 proposes changes in PMA, PCS or RS to allow non-receiving DTEs to transmit frames

- Changes to be optional on some port types, mandatory on others

- This "unidirectional" feature to be used to signal remote fault (RF) by frames

- At two speeds out of three, feature not compatible with existing physical layer RF signaling

Problem statement

- Incompatible physical layer logical behaviors on similar or same PMDs
- Market, operations and procurement confusion
- Two "solutions" to same "problem" (remote fault signaling)
  - But access networks are not that different
    - Must connect to non-access networks
    - OAM and remote management may be appreciated by IT departments for same reasons as by phone companies
- New "solution" less generic than current one

EFM has contradictory objectives

- "Provide a family of physical layer specifications:
  - 1000BASE-LX extended temperature range optics"

  Implies that the new or modified port type be interoperable with legacy 1000BASE-LX

  Intended to allow "new" optics to plug into "old" slots to form compliant ports

and

- "Support far-end OAM for subscriber access networks:
  - Remote Failure Indication"

Expectations set by 100M call for interest (folded into EFM)

PHY identical to current 100Mbps Std except for a new PMD

– No change to Clause 24

– Retain all state machines, 4B/5B coding etc. of 100BASE-X

- Only need to extend Clause 26, 100BASE-FX PMD, to include SMF

- 100BASE-X dual SMF is already happening, and will have applicability even outside EFM

- However, 100BASE-X SMF will be used in the public access application space

- 100BASE-X PCS is transparent to EFM OAM

– Neither "OAM in Frames" nor "OAM on Preamble" require any changes to 100BASE-X PCS'

- **We need to honor these expectations.**

  http://www.ieee802.org/3/smfx_study/index.html

  http://www.ieee802.org/3/smfx_study/public/jonsson_1_0302.pdf

On physical layer objectives

- Remember pluggable optics
- D3.1 proposes PCS behavior to depend on optics type …
  - 1000BASE-LX - mandatory /C/ ability, optional unidirectional ability
  - 1000BASE-LX10 – forbidden /C/, mandatory unidirectional ability
  - **But they are interoperable and interchangeable! Makes no sense.**
- 100BASE-LX10 plugs into same ICs as today's 100BASE-FX
  - **It's already out there**

Unidirectional is not the same as OAM

- ALL PHYs support the OAM sublayer in general (because all can transport Slow Protocol frames). But:

1. Transmitters into a shared medium can't signal remote fault (RF). Not by modified idle, nor by OAM frames, nor by preamble.

2. Transmitters into a medium with a crosstalk issue (EFM copper, 1000BASE-T) can't signal RF by frames or by preamble, but can use PHY-specific methods.

3. Other point-to-point PHYs can use any of these three methods

- **Physical layer RF signaling is the more generic method**

On OAM / remote failure indication objective

- "Support far-end OAM for subscriber access networks:
  - Remote Failure Indication"

- D3.1 does a good job of far end OAM apart from unidirectional

- It could do a better job (faster, more compatible with 802.3, working on more port types) if it used or at least coexisted with existing RF signaling methods

- Since these RF signaling methods exist before EFM, providing or supporting "Remote Failure Indication" looks like a no-op for P802.3ah

Mix and match OAM and unidirectional

- Proposed OAM transport layer to be optional
- Proposed unidirectional capability to be:
  - Sometimes optional
    - Unidirectional mode can be switched on and off
  - Sometimes mandatory
    - Unidirectional mode can be switched on and off
  - Sometimes not applicable
- As far as interoperability goes, a DTE can't tell if:
  - its link partner has unidirectional capability but no OAM,
  - or no unidirectional capability and no OAM,
  - or has OAM but won't use it (with or without unidirectional)

Cable plant now common across speeds

	Dual MMF	Dual SMF	Notes
Before EFM			
100 Mb/s	2 km	X	Choice of FDDI or OC-3 PMD. Common PCS. Needed a standard.
1 Gb/s	550 m	5 km	Stretched to 10 km, needed a standard. Just one PCS type!
10 Gb/s	varies	10 km	
After EFM (as proposed in D3.1)		*Very valuable innovation for non-access networks, but…*	
100 Mb/s	2 km	10 km	*per D3.1, "needs" 2 PCS/PMA types*
1 Gb/s	550 m	10 km	*per D3.1, "needs" 2 PCS types*
10 Gb/s	varies	10 km	At least the 2nd RS type is optional

Pluggable optics and legacy ICs

1000BASE-LX10 is interoperable with 1000BASE-LX

Port type name does not describe minor options

e.g. 100M options, presence of OAM layer

Cannot stop anyone plugging 1000BASE-LX10 GBIC into today's 1000BASE-LX slot

And they'll call it 1000BASE-LX10!

No fix to any defect of remote fault signalling in today's p2p 1000BASE-X can completely avoid having options

There are several options today around auto-negotiation and FEFI ability and use

How to do RF signaling across 802.3?

	Traditionalist	Compromise	Radical
	Physical layer RF	Both, where applicable	Depending on port type: OAM layer RF or physical layer RF
Ad-vantages	In line with 802.3 Compatibility, interoperability, parts sourcing Allows for fast protection switching Low editorial workload and low schedule risk Satisfies "extended 1000BASE-LX" objective	Compatibility, interoperability, parts sourcing Extensible to other signals than RF Satisfies "extended 1000BASE-LX" objective	Extensible to other signals than RF
Disad-vantages	Doesn't immediately fix any bug in clause 37	More options Does not all have to be done now, in EFM	Inconsistent across port types Market confusion Lack of compatibility, interoperability, parts sourcing Likely to be flouted More options Slower protection switching?
What would be done	Simplify cl.66 to address 1000BASE-PX-D only	Modify cl.66 per comments to D3.1 Leave any difficult state-diagram changes to revision PAR	Continue arguing about which ports behave which way, debate more port names, leave the real world and the standard diverged

When/if we decide to proceed with changes to RF signaling in 802.3…

	Don't change from 802.3	Change later	Change "now" (in 802.3ah) (whatever the change is)
Ad-vantag es	In line with 802.3 Compatibility, interoperability, parts sourcing Market simplicity	Opportunity to have thought-through consistent strategy May be extensible to other signals than RF	Any benefits take effect soon
Disad-vantag es	Believed weak for 1000BASE-X	Delay in "fixing" 1000BASE-X	Need to get it right, this week Stuck with the consequences Still have to agree which port types, naming, options, this week
What would be done	Simplify cl.66 to address 1000BASE-PX-D only	Simplify cl.66 to address 1000BASE-PX-D only Leave any difficult state-diagram changes to revision PAR	Agree an RF strategy for 802.3 Agree detail of how it works, show interoperability and compatibility (or not) with cl. 24, 36, 37, 46 Agree which ports behave which way Agree any more port names

Conclusions from study

We cannot use OAM layer remote fault indication as a generic mechanism: too many PHY types can't use it

This job is better done, with minimum latency, by existing PHY mechanisms

No need to introduce new options and incompatibilities to 10G and 100M to fix a bug in clause 37 (Gigabit)

May wish to allow both OAM and physical RF signaling for complex networks

Layering: specifications of PCS/PMA/RS and PMD (e.g. wavelength) should be decoupled

Want to be able to use today's PCSs with today's being-standardized optics

What SHOULD we specify?

- Legacy-friendly
- Try to be consistent across speeds
- Support interoperable and interchangeable PMDs
- *Other…?*

Proposal 1

- Allow capability for unidirectional transmission of frames:

  – Mandatory in 1000BASE-PX-D

  – Forbidden or discouraged in 1000BASE-PX-U

  – An option for other 1000BASE-X

    • **Not tied to PMD type such as wavelength**

  – An option for 100BASE-X and 10G

    • **Not tied to PMD type such as wavelength**

Alternative proposal

- We don't need to introduce new options and incompatibilities to 10G and 100M to fix a bug in clause 37.

- Fix clause 37!
  - This would have to be an option, and preferably interoperable with today's 1000BASE-X whether option is on or off.
  - May be better done in the revision PAR than in EFM

- Restrict unidirectional transmission of frames to its necessary place in 1000BASE-PX-D

Backup material follows

- Detailed tables of RF by port type
- State transition diagrams for RF signaling
- Compatibility with RPR?
- More detailed stuff
- History of unidirectional in EFM drafts
- Layering

Detail of proposal 1 by speed

		Physical layer RF signaling			Unidirectional capability			Compatibility	
		No RF	Optional	Mandatory	No	Optional	Mandatory		
802.3	100BASE-X	Opt	Opt		No				
	1000BASE-X	?	?	?	No				Buggy?
	10G			Yes	No				
D3.1	100BASE-X	or	or			or	or	No	Depending on PMD
	1000BASE-X	or	?	?	?	or	or	No	Depending on PMD
	10G			Yes		Opt		Yes	
Proposal 1	100BASE-X	Opt	Opt			Opt		Yes	
	1000BASE-X	?	Opt			Opt		Yes	
	10G			Yes		Opt		Yes	
Proposal 2	All	As 802.3 except PX-D			No except PX-D		PX-D only		

- Per 802.3ah D3.0 and D3.1:
- EFM's new OAM sublayer is to be optional
- Among other messages, it is to contain an optional remote fault signal to be transported by slow protocol frames
- Mandatory changes proposed to state diagrams or mandatory behavior of physical layers:
- 100BASE-X PMA and PCS
- 1000BASE-X PCS
- 10G RS
- "Required" for a sublayer that's optional!?

What we learnt in January   2 of X

- These changes to depend on optics type …
  - 1000BASE-LX - mandatory /C/ ability, optional unidirectional ability
  - 1000BASE-LX10 – forbidden /C/, mandatory unidirectional ability
  - **But they are interoperable and interchangeable! Makes no sense.**

- Further proposed changes would defeat the 100BASE-X far end fault indication (FEFI)

- All this is called "Unidirectional ability"

- The OAM sublayer can work without unidirectional ability and is itself optional

What we learnt in January   3 of X

- per D3.0 57.2.12,
  - "When a link is operating in Unidirectional OAM mode, the OAM sublayer ensures that only Information OAMPDUs with the Link Fault critical link event indication set and no Information TLVs are sent once per second across the link."

- and per 57.2.10.1,
  - "Link fault" means "The PHY has determined a fault has occurred in the receive direction of the local DTE.
  - Note: The definition of the specific faults comprising the ... Link Fault flags is implementation specific and beyond the scope of this standard."

What we learnt in January

- Unidirectional ability …
- … is to be an option on some legacy port types.
- … is not compatible with Auto-Negotiation or half-duplex mode.
- … is not compatible with loop-timed PMD systems: 1000BASE-T, future 10GBASE-T, or EFM copper.
- … cannot be used in the upstream direction of 1000BASE-PX.

What we learnt in January   5 of X

- When a 10GBASE-{anything but T} port's receive side is down;
  - If in unidirectional able mode, it is to transmit link fault OAMPDUs embedded in RF (because receive side is down), which conveys the same information as RF.
  - If not in unidirectional able mode, RF is transmitted.

- For 100BASE-X;
  - If in unidirectional able mode, it is to transmit link fault OAMPDUs embedded in idles.
  - If not in unidirectional able mode, FEFI or idle (not AN for fiber optical 100BASE-X) is transmitted.

- For 1000BASE-X but not from ONU to OLT of 1000BASE-PX;
  - If in unidirectional able mode, it is to transmit link fault OAMPDUs embedded in idles.
  - If not in unidirectional able mode, idles or /C/, indicating remote fault, are transmitted.

Thus:

For 10G, unidirectional duplicates an existing PHY function for signalling remote fault that works

For 100BASE-LX, unidirectional doesn't duplicate, but disables an existing optional PHY function for signalling remote fault, and adds an optional OAM layer RF function

For 1000BASE-X, unidirectional disables a questionable PHY function for signalling remote fault and adds an optional OAM layer function

"Issue is in Cl.37 auto-negotiation logic: doesn't transmit remote fault until link is back up (in both directions)"? Not sure this is true: looks like plain vanilla /C/ will signal RF

(Unidirectional ability is not compatible with auto-negotiation)

Signal remote fault in PHY or OAM sublayer?

- Option 1   PHY layer
  - Standardized for most PHYs
  - Mandatory for 100BASE-FX, *Auto-negotiation* mandatory but *suspect for 1000BASE-X*, mandatory and clean for 10G
  - *Hardware oriented – fast*

- Option 2   OAM layer
  - New, optional
  - Can be defined for most PHYs, but not all because of collisions and crosstalk issues
  - *No good for 2BASE-TL, 10PASS-TS, little use for 1000BASE-T,*
  - *Software oriented – slower*

- Option 3   Both
  - Could be useful for e.g. multi-hop links
  - Has to be by options to allow graceful evolution

Compatibility of remote fault signaling with Gigabit Ethernet physical layers

PMD type	PHY layer remote fault signal possible (per Cl. 36, 37, 40)?	OAM sublayer remote fault signal possible?
1000BASE-T	Yes	MASTER to SLAVE, not vice versa[1]
1000BASE-SX	Yes[2]	Yes
1000BASE-LX	Yes[2]	Yes
1000BASE-LX10	Yes[2]	Yes
1000BASE-BX10	Yes[2]	Yes
1000BASE-PX-D	Yes[2, 3]	Yes[3]
1000BASE-PX-U	No[3]	No[3]

Notes

1 1000BASE-T uses loop timing to control crosstalk – not allowed to transmit frames if it can't hear. Analogous to PX-U – DTE at center could signal RF to periphery using OAM sublayer, but DTE at periphery can't tell the center.

2 Cl. 36 and 37 offer a PHY level remote fault mechanism within Auto Negotiation. Is it buggy?

3 1000BASE-PX-U (customer's end) is not allowed to transmit at all if it can't hear, else it might talk over another customer. So customer's end can't report problems to central office, even if central office can report problems to customers

Compatibility of remote fault signaling with 2B/10P, Fast and 10G Ethernet physical layers

PMD type	PHY layer remote fault signal possible (per Cl. 24? And others)?	OAM sublayer remote fault signal possible?
2BASE-TL	Yes	No[1]
10PASS-TS	Yes	No[1]
100BASE-Tn?	Yes?	Yes
100BASE-FX	Yes[2]	Yes
100BASE-LX10	Yes[2]	Yes
100BASE-BX10	Yes[2]	Yes
10GBASE-R	Yes	Yes
10GBASE-W	Yes	Yes?
10GBASE-LX4	Yes	Yes

Conclusion: Ethernet PHYs mostly have physical layer RF signal (so does SONET)

Notes

1   DSL based PHYs have own remote fault signaling, can't allow unidirectional transmission of frames for crosstalk reasons?

2   100BASE-X remote fault signal is called Far End Fault Indication

Options for remote fault signaling with 2B/10P, Fast and 10G Ethernet physical layers

PMD type	802.3, DSL standards, 1000BASE-LH spec		D3.1		This proposal		Notes
	PHY capability	Operation	PHY capability	Operation	PHY capability	Operation	
2BASE-TL	*? 1*		*PHY*	*PHY, plus optionally by OAM*	*PHY only?1*	*PHY only?1*	
10PASS-TS	*? 1*		*PHY*	*PHY, plus optionally by OAM*	*PHY only?1*	*PHY only?1*	
100BASE-Tn?	*PHY? AN?*		*PHY? AN?*	*PHY, plus optionally by OAM*		*PHY, plus optionally by OAM*	
100BASE-FX	*None or PHY (FEFI)*	*None or FEFI*	*None, FEFI or unidir*	*None, FEFI or unidir*	*None, FEFI, unidir or both*	*None, FEFI, by OAM, or both*	*Be more*
100BASE-LX10, BX10			*Optional FEFI, mandatory unidir*	*PHY (FEFI) or Auto-Negotiation or by OAM*			
1000BASE-T	*PHY and AN*	*PHY and AN*	*PHY and AN, optional unidir?*	*?, optionally by OAM*			
1000BASE-SX, LX	*PHY Auto-Negotiation*	*PHY AN or none*	*PHY AN, optional unidir*	*None, PHY AN or by OAM*	*PHY AN, optional unidir*	*None, PHY AN or by OAM*	*consistent across*
1000BASE-LX10	*?*	*?*	*PHY AN, mandatory unidir*	*None, or by OAM*			
1000BASE-PX-D			*Unidirectional*	*Unidirectional*	*Unidirectional*	*Unidirectional*	
1000BASE-PX-U			*PHY AN*	*Recommend none*	*PHY AN?*	*Recommend none*	
10GBASE-R/W/LX4	*PHY (RF)*	*PHY (RF)*	*PHY (RF), optional unidir*	*PHY (RF), optionally by OAM also*	*PHY (RF), optional unidir*	*PHY (RF), optionally by OAM also*	*p2p speeds!*
10GBASE-T	*"No unidirectional capability" 2*						

Notes

0   Needs checking!

1   DSL based PHYs have own remote fault signaling?, can't allow unidirectional transmission of frames for crosstalk reasons?

2   Per 10GBASE-T SG

What is remote fault signaling for?

- It stops the MAC forwarding data frames
  - Saves wasting upstream bandwidth and misleading customers

- Other uses outside of 802.3
  - Informs router tables?
  - Triggers protection switching?
    - If Ethernet to carry voice and video and compete with SONET "50 ms switching", should be fast
    - Other?

- In Ethernet (and SONET), it's done at the physical layer

- RPR doesn't use RF, uses probe packets instead

- Fibre Channel uses ordered sets?

Compatibility with RPR?

Do D3.1's proposals for unidirectional operation have implications for 802.17 RPR?

- 802.17 has its own 1G and 10G RSs and modifies the PHYs
- 801.17 does not generate RF indications, and ignores them (except possibly for PHY_LINK_STATUS.indication)
- 802.17 does not use autonegotiation and assumes full-duplex operation
- Conclusion: At least for 1G/10G p2p links: legacy 802.3, D3.1 proposal, and this presentation's proposal, are all compatible with, or replaced by, 802.17
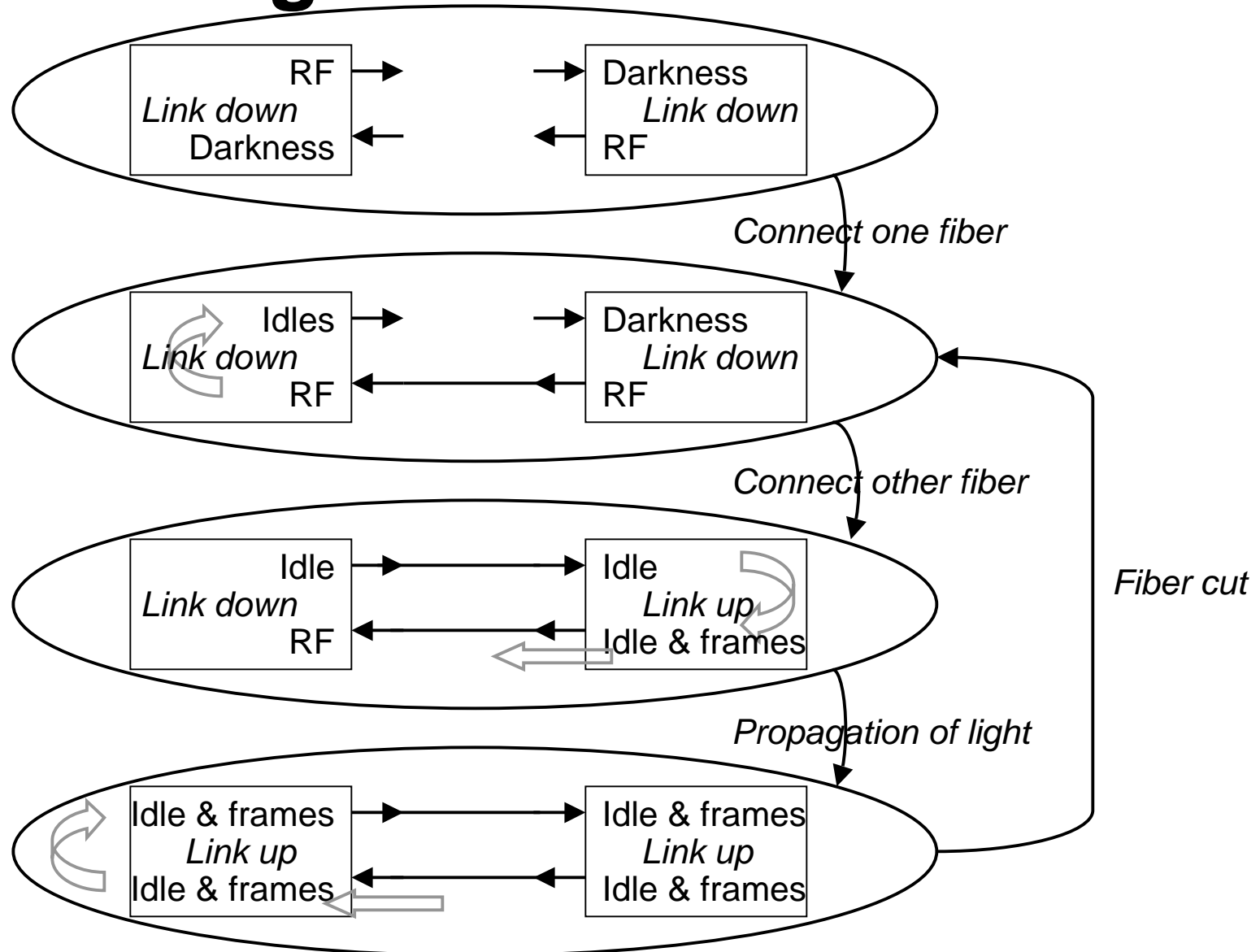
Not an issue

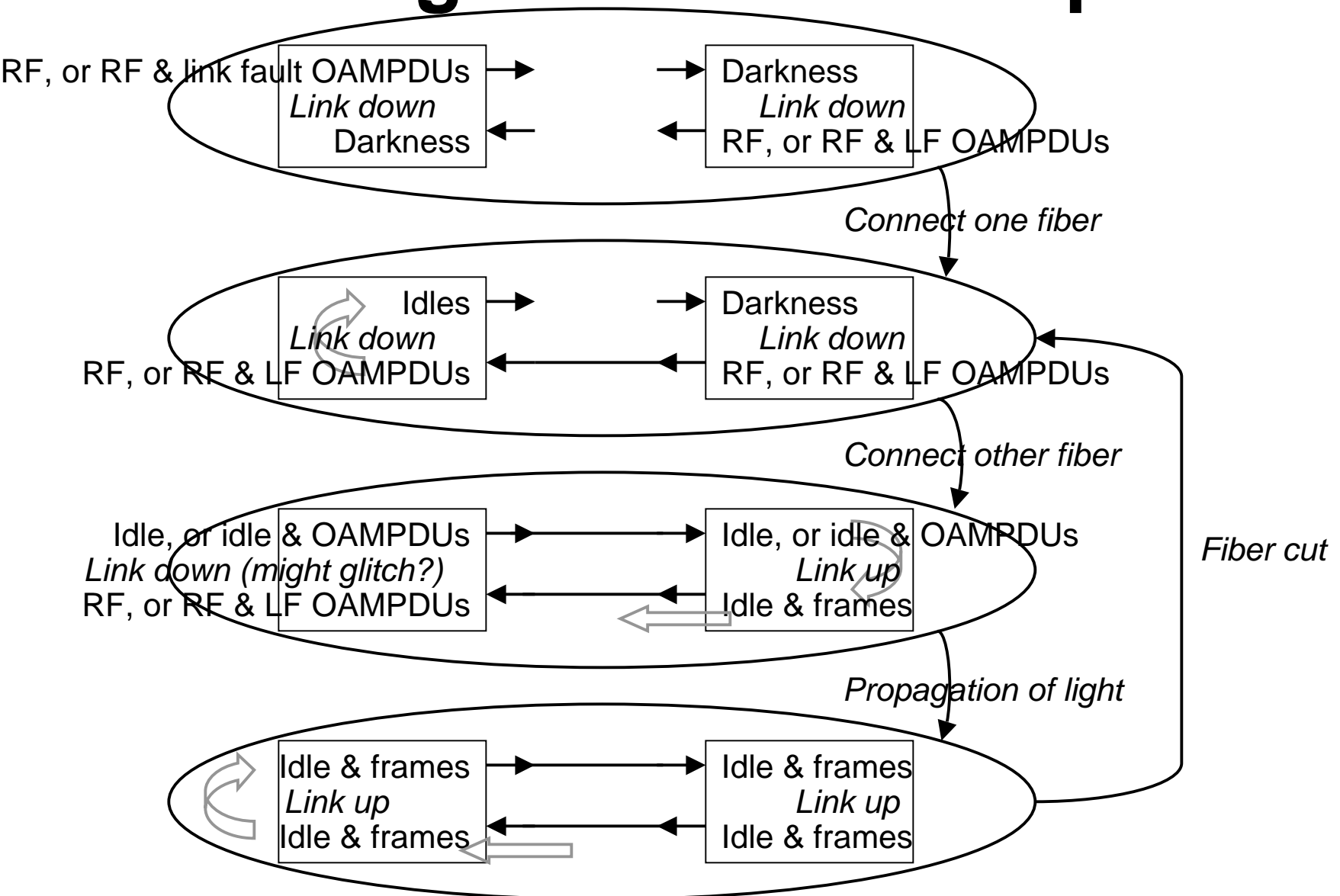# Options in existing optical port types

- 100BASE-FX
  - FEFI
- 1000BASE-X optical
  - AN?
  - RF
  - Extensions of AN?
- 10GBASE-R
  - None, at present
- Notice these options do not spawn new port types or names
  - Options strongly disliked by standards bodies because of feature creep – but may be needed for backwards compatibility
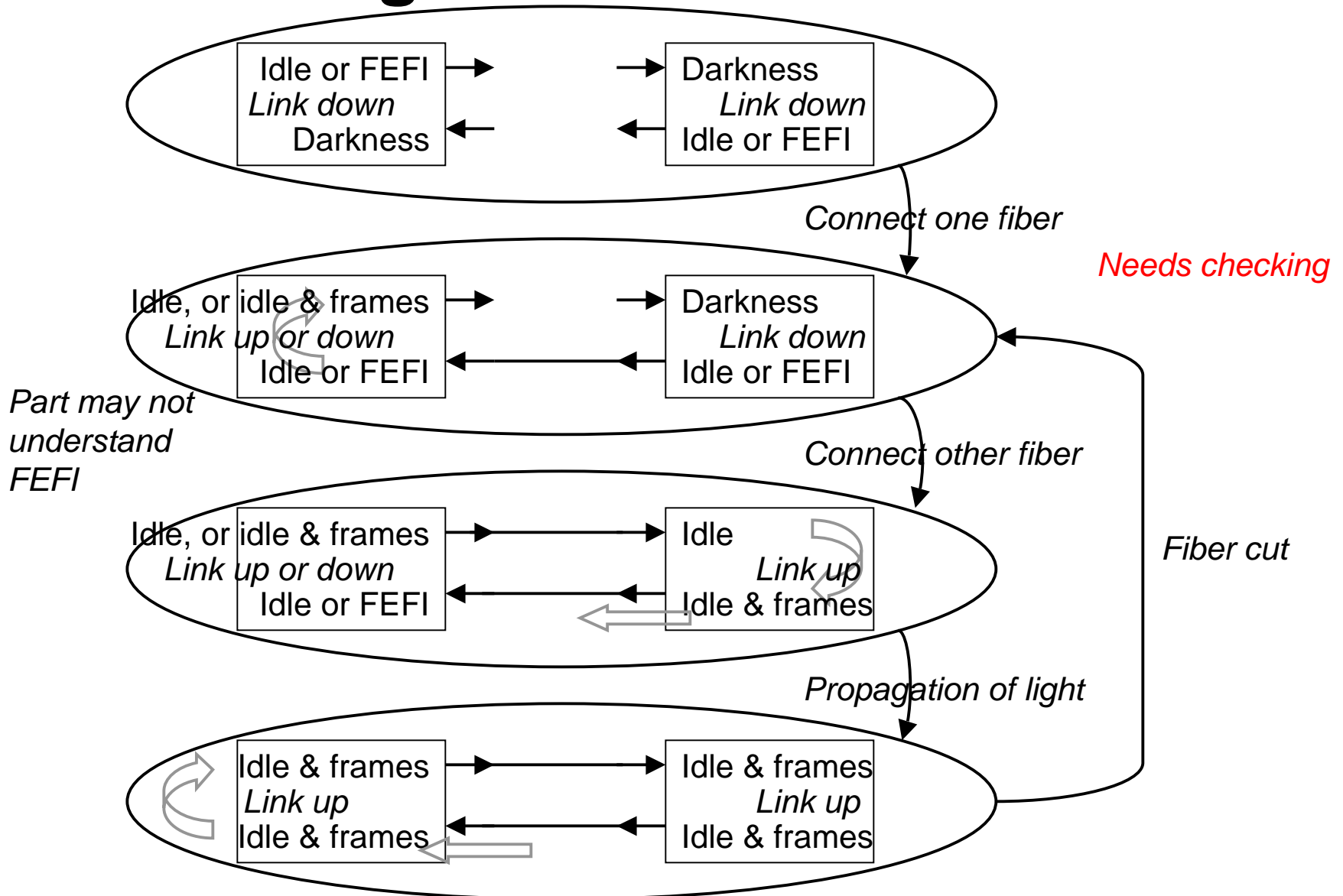 - Some options known by project e.g. 802.3ad

State diagram: remote fault in 10G

State diagram: RF in 10G per D3.1

State diagram: FEFI in 100BASE-FX

Idle or FEFI
Link down
Darkness

Darkness
Link down
Idle or FEFI

Connect one fiber

Needs checking

Idle, or idle & frames
Link up or down
Idle or FEFI

Darkness
Link down
Idle or FEFI

Part may not understand FEFI

Connect other fiber

Idle, or idle & frames
Link up or down
Idle or FEFI

Idle
Link up
Idle & frames

Fiber cut

Propagation of light

Idle & frames
Link up
Idle & frames

Idle & frames
Link up
Idle & frames

State diagram: 100BASE-LX10 in D3.1: example of FEFI but not unidirectional aware at one end, unidirectional but not FEFI aware at other
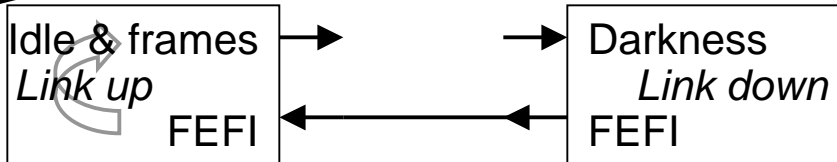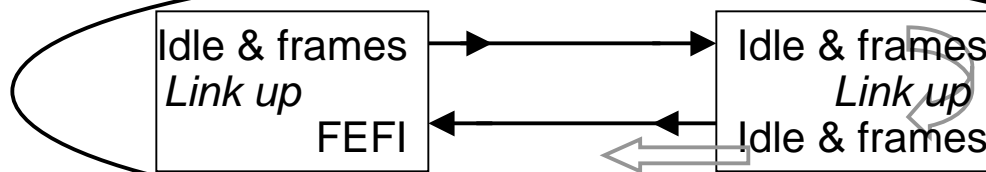
Starts in non-unidirectional mode

Idle *Link down* Darkness	Darkness *Link down* FEFI

Connect one fiber

Needs checking

Part does understand FEFI

Idle & frames *Link up* FEFI	Darkness *Link down* FEFI

Connect other fiber

Fiber cut

Idle & frames *Link up* FEFI	Idle & frames *Link up* Idle & frames

Propagation of light

Idle & frames *Link up* Idle & frames	Idle & frames *Link up* Idle & frames

Option 3 if we believe "access is different"

- Think of two "buckets"

- All PHYs that use dual fibers or data Cat-n type twisted pair in first bucket
  - Use existing RS/PCS/PMA with its RF
  - Avoid worst of interoperability and market confusion
  - Consider optional compatible RF in frames

- Access-only PHYs, that use single fiber or telephony cable, in second bucket
  - Should strive for consistent Ethernet-like behavior, even in bucket 2

History of unidirectional in EFM drafts

- D0.9 Cl.57 says some physical layer devices "should support", others "may"
- D1.0 Modifies Cl.24 and 36 for all PHYs (not just EFM). Nobody notices?
- D1.1 Cl.57 says some physical layer devices "support", others "may"
- D1.2 Cl.57: "OAM functionality is mandatory for some, optional for all other"
- D1.3 Above text removed again: OAM is optional. P2MP called out specifically in Cl.57
- D1.414   Changes to Cl.46 for all 10G ports. Nobody notices?
- D1.9, D2.0   Cl.57 EFM copper doesn't support RF in frames
  - Comments point out that retroactive change is bad
- D2.1 Changes to 24, 36, 46 removed: Cl.66 modified RS/PCS/PMA added: some PHYs "require" (editor's words). Optics clauses continue to point at Cl. 24, 36.
- D2.2, D3.0
  - Comments objecting to mandatory and incompatible nature of changes
- D3.1
  - More comments objecting to mandatory and incompatible nature of changes

Summary: we've been slow in picking up these problems!

With hindsight, EFM copper's lack of support, as well as 1000BASE-T, 10GBASE-T, on top of backwards compatibility issues, defeated the stated main attraction of RF in frames
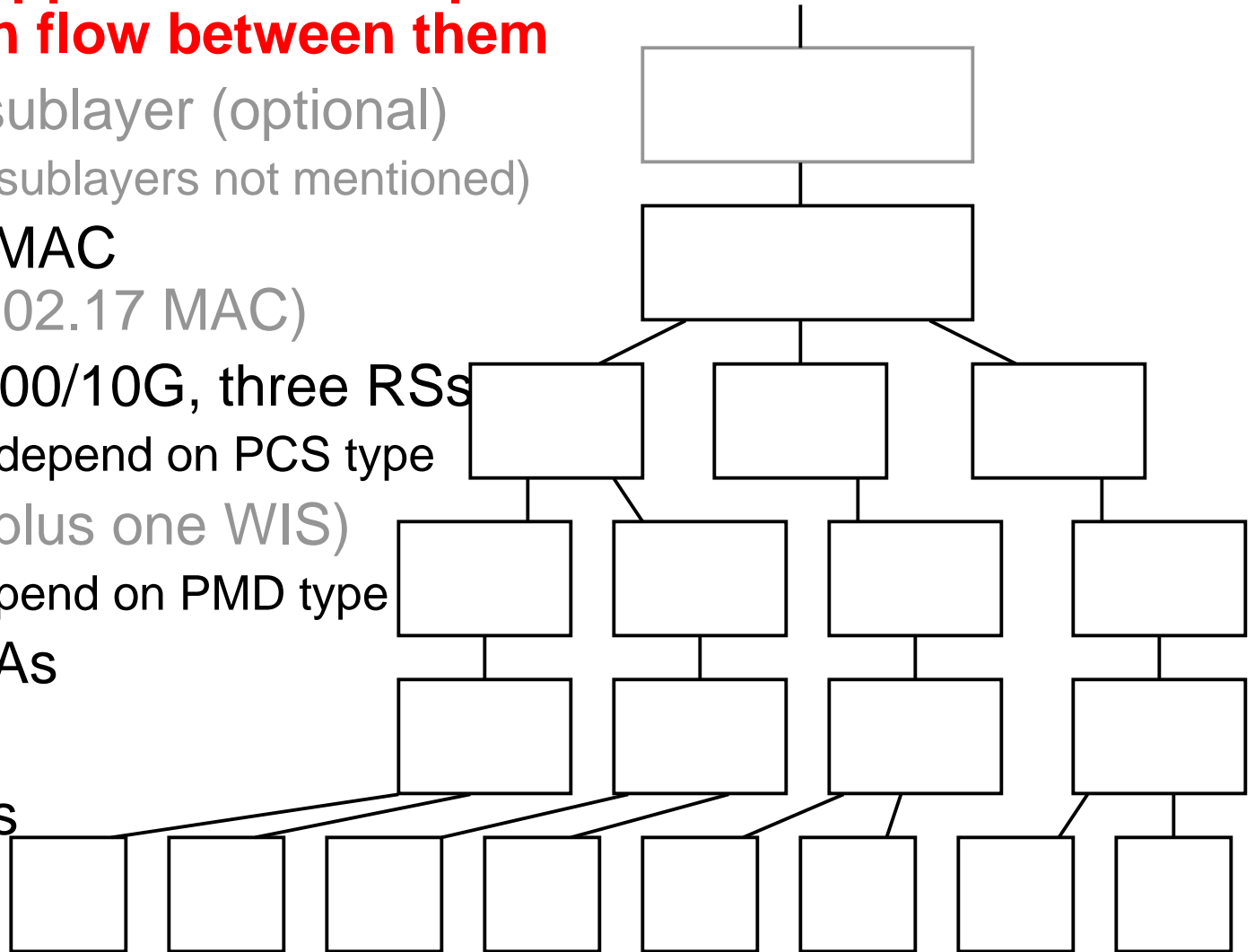
Layering

Layers are supposed to be separate with minimum information flow between them

- One OAM sublayer (optional)
  (many other sublayers not mentioned)

- One 802.3 MAC
  (plus One 802.17 MAC)

For 10/100/1000/10G, three RSs

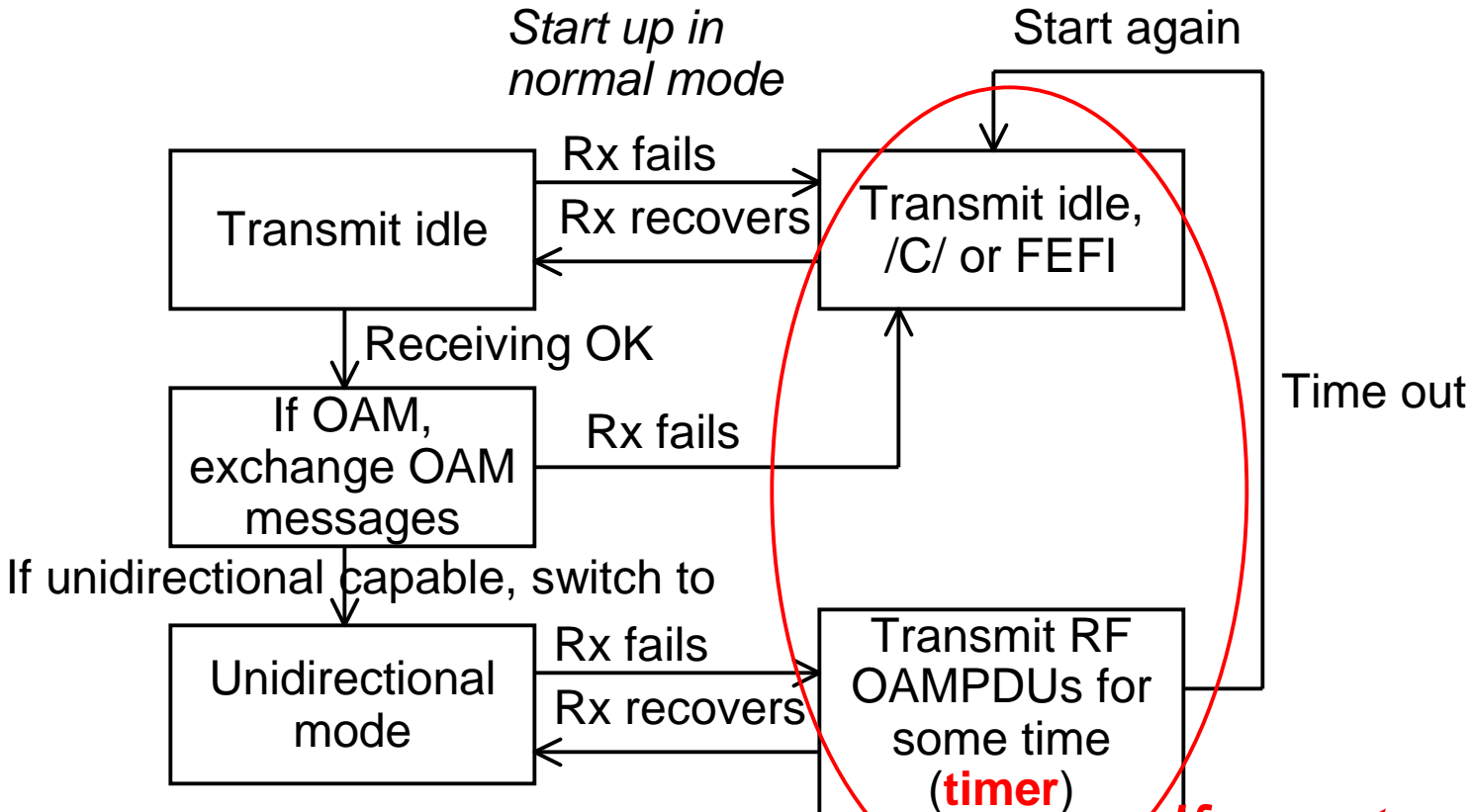- – Shouldn't depend on PCS type

Four+ PCSs (plus one WIS)

  Shouldn't depend on PMD type

- Four+? PMAs

- Many PMDs

- **Don't make it even more complicated!**

Timer needed to restart in normal mode?

Start up in normal mode

Start again

Transmit idle		Transmit idle, /C/ or FEFI

Rx fails →

← Rx recovers

↓ Receiving OK

Rx fails

If OAM, exchange OAM messages

If unidirectional capable, switch to ↓

Time out

Unidirectional mode		Transmit RF OAMPDUs for some time (**timer**)

Rx fails →

← Rx recovers

If can transmit RF OAMPDUs in RF, can we merge these states and avoid the timer?

After Rx failure, cable could be disconnected and then connected to another path, therefore need to start again from scratch

Questions 1/3

- Should EFM be very like "mainstream" (IT-oriented) Ethernet for its remote fault signaling, or different? Why?

- How should we account for dual-use PHYs (used in both "mainstream" and access networks)?

- Is "backwards" compatibility good, bad, irrelevant?

- Do we have a clear view on what remote fault signaling is for, anyway?

How to do RF signaling across 802.3?

- Traditionalist ?
  - (Physical layer RF)

- Compromise
  - (Both, where applicable)

- Radical
  - (Depending on port type: OAM layer RF or physical layer RF)

When/if to proceed with changes to RF signaling in 802.3?

- Don't change from 802.3?

- Change later?

- Change "now" (in 802.3ah) (whatever the change is)?