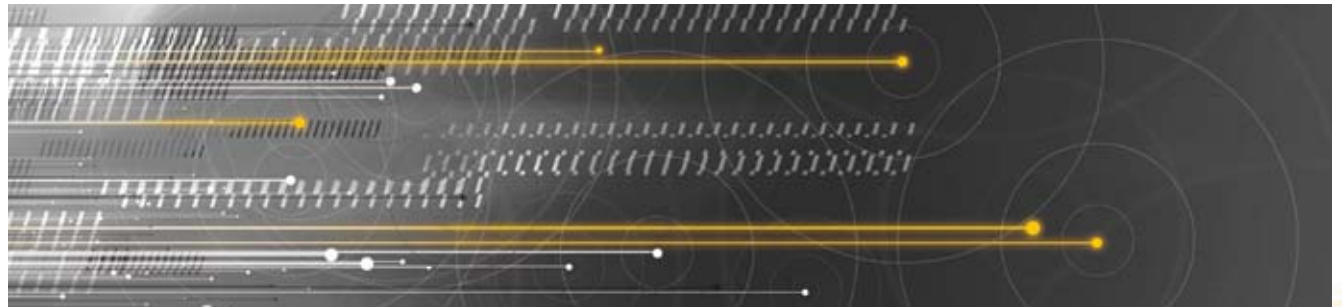


■ MPLSによるLegacy Network Migrationの課題



ソフトバンクテレコム株式会社

大矢 晃之

teruyuki.oya@tm.softbank.co.jp

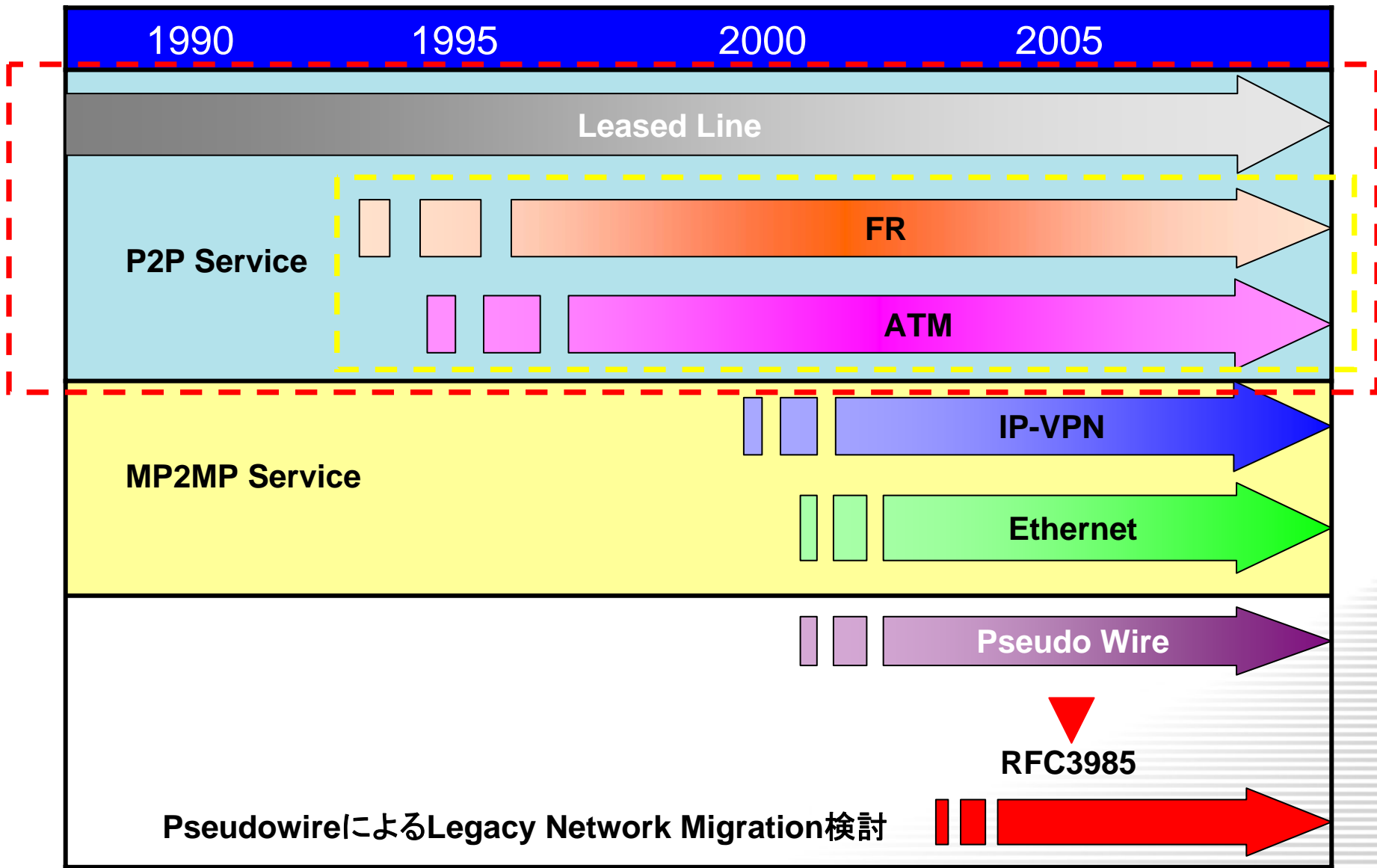
MPLS JAPAN 2007 — Oct 10, 2007

For Your Networking Universe

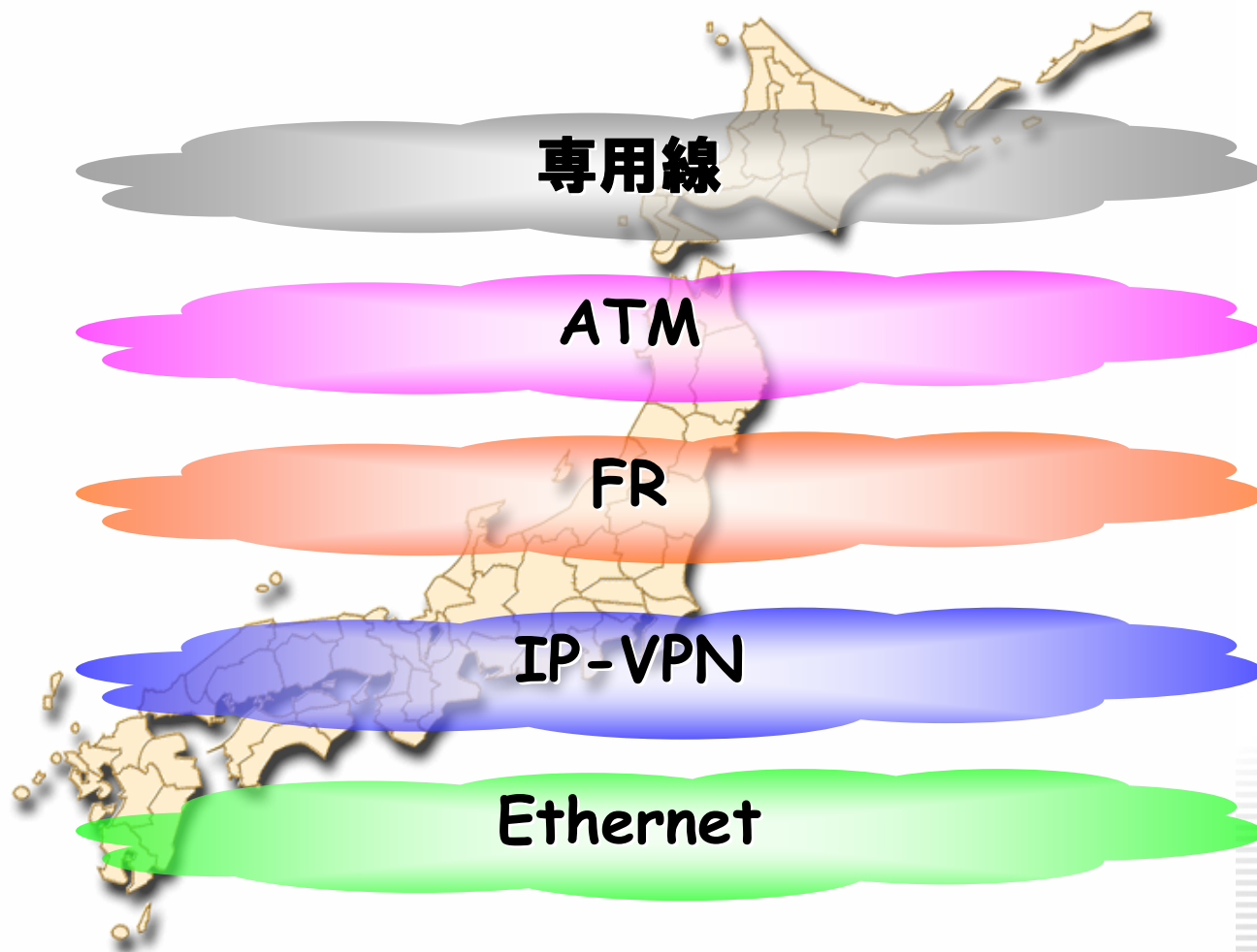
Copyrights Reserved ©SOFTBANK TELECOM

1. Legacy NW Migrationの必要性と目的
2. Requirement
3. 現状の問題点
4. Migration Scenarioの検討
5. Summary

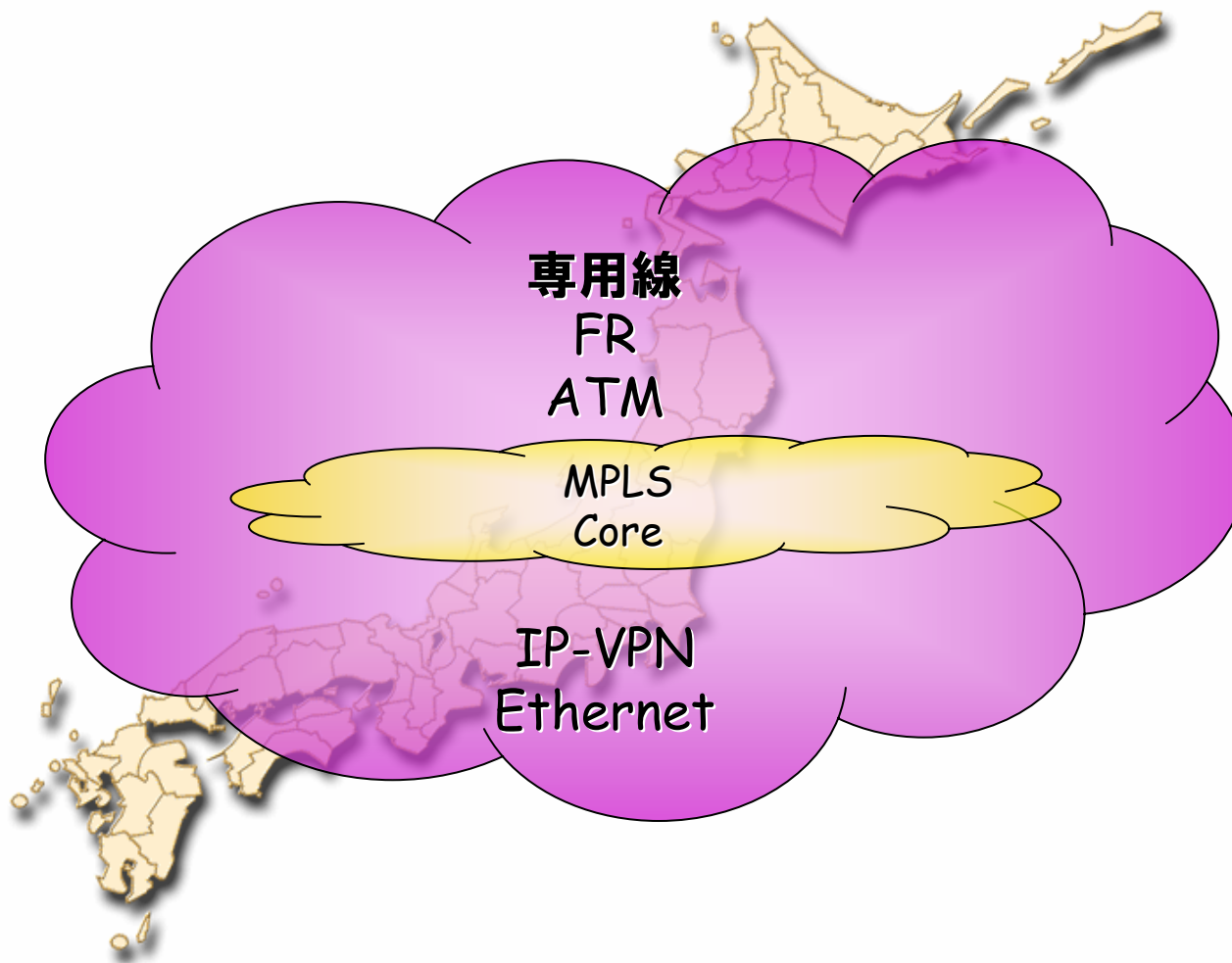
Legacy Network



サービスが生まれるたびに全国NWを構築



共通のMPLS Core上で、それぞれのサービスを論理的に分離し提供



1

設備の老朽化対応

2

新技術の導入による価値創造

3

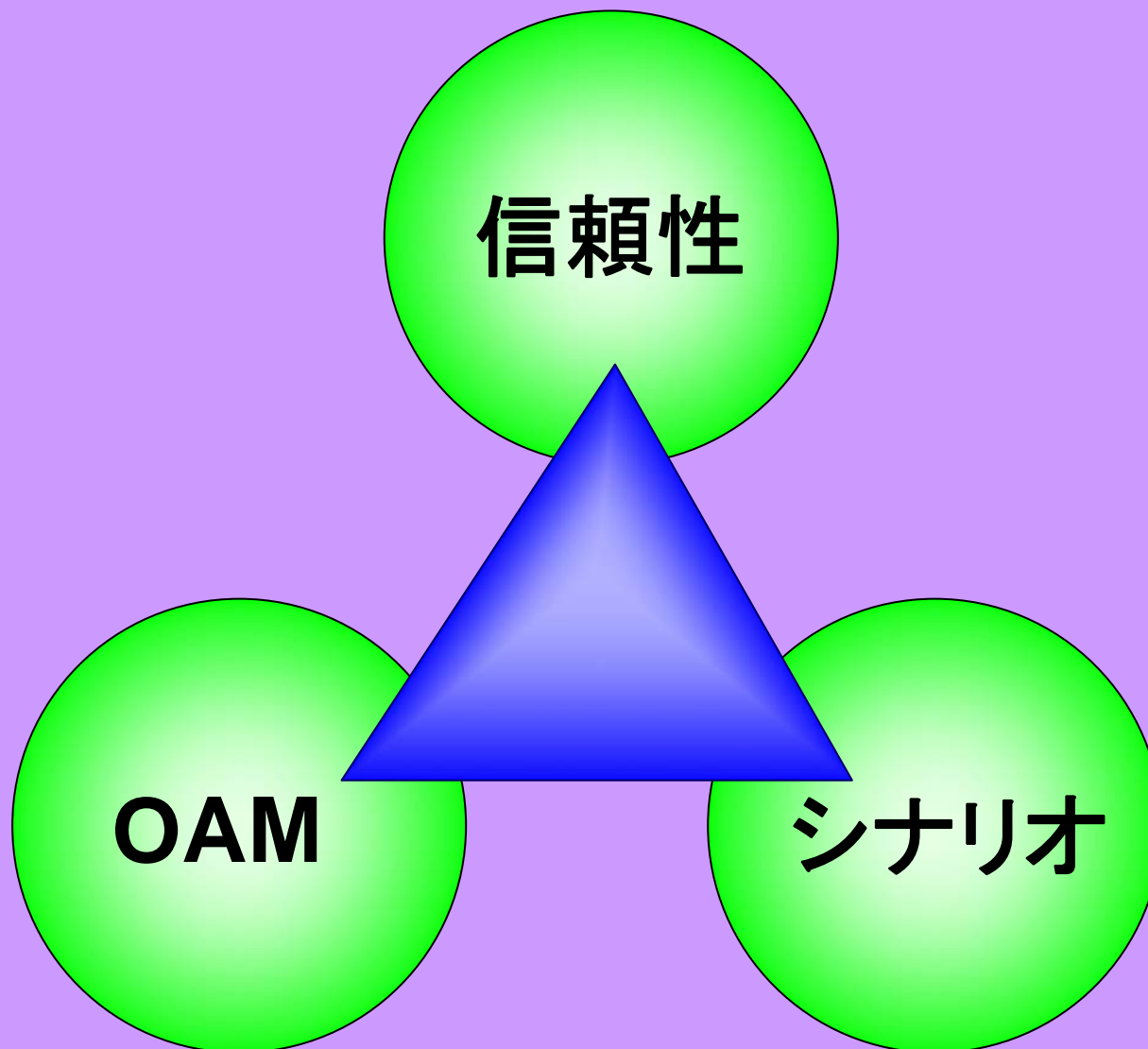
OPEX削減

4

Networkの最適化

5

オペレーションの最適化



1. Legacy NW Migrationの必要性と目的

2. Requirement

3. 現状の問題点

4. Migration Scenarioの検討

5. Summary

信頼性

HW完全冗長
共通部モジュール冗長
LineCard冗長
1+1APS/MSP
protection
sub 50ms
無瞬断upgrade

OAM

Full Rate試験
Loop
Drop/Insert
Performance Monitoring

管理

パス管理
容量管理

QOS

100% e-e帯域保証
Low Latency

信頼性

- HW完全冗長
- 共通部モジュール冗長
- LineCard冗長
- 1+1APS/MSP
- Rerouting
- 瞬断upgrade

OAM

- 豊富なOAM
- AIS/RDI
- Loopback
- CC
- Performance Monitoring

管理

- パス管理
- 容量管理
- 回線ごとのルート管理

QOS

- Service Class
(CBR/rtVBR/nrtVBR/UBR/UBR+/ABR/GFR...)
- 帯域保証(100%保証、一部保証)
- Policing/Shaping
- CLP
- PCR/SCR/MCR/MBS
- CAC
- Oversubscription

信頼性

HW完全冗長
共通部モジュール冗長
LineCard冗長
Rerouting
瞬断upgrade

OAM

LMI
Active bit/New bit
FR/ATM Interwork

管理

パス管理
容量管理
回線ごとのルート管理

QOS

帯域保証(CIR保証)
PIR/CIR
Bc/Be
DE
Oversubscription

信頼性

HW完全冗長
共通部モジュール冗長
LineCard冗長
1+1APS/MSP
protection
sub 50ms
無瞬断upgrade

OAM

Full Rate試験
Loop
Drop/Insert
Performance Monitoring
Interworking

Traffic Engineering & Management

パス管理
容量管理
回線ごとのルート管理

QOS

E-E帯域保証(100%保証、一部保証)
Low Latency
Policing/Shaping
Service Class Differentiation
Discard Priority
CAC
Oversubscription

1. Legacy NW Migrationの必要性と目的

2. Requirement

3. 現状の問題点

4. Migration Scenarioの検討

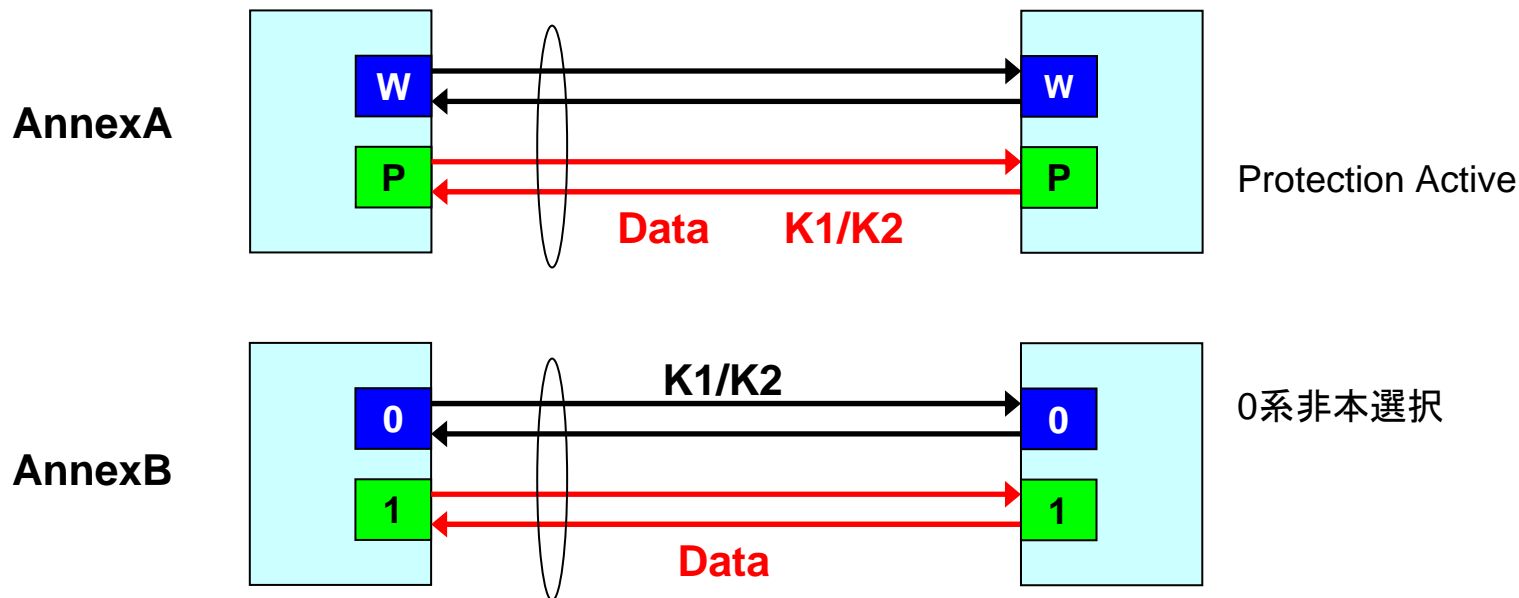
5. Summary

信頼性

Legacy ServiceをMPLS上で提供する上で、冗長機能は必須

- ✓ Hardware完全冗長
- ✓ (無)瞬断切替
- ✓ 無瞬断 Upgrade
- ✓ Line Protection
 - 1+1 APS/MSP (G.841 AnnexB)
- ✓ NW Protection
 - Local Repair(FRR) / Global Repair / Ring Protection
- ✓ Non-Stop-Routing/Signaling
 - OSPF,BGP,ISIS,LDP,RSVP...

	通称	標準	特徴
SDH	ITU-T AnnexA	<ul style="list-style-type: none"> ITU-T G.841 Section. 7.1 (1998/10) (旧)ITU-T G.783 Annex A (1997/04) <p>MSP(Multiplex Section Protection)</p>	<p>勧告仕様のには、以下の選択が可能だが、ベンダ、機器によって実装は異なる。</p> <ul style="list-style-type: none"> 1+1/1:N Bi-directional/Uni-directional Revertive/Non-Revertive <p>K1/K2 byte W系: “0001”、P系: “0000”</p>
	TTC AnnexB NTT仕様	<ul style="list-style-type: none"> ITU-T G.841 Annex B (1998/10) TTC JT-G.783 (2001/04 第3版) (旧)ITU-T G.783 Annex B (1997/04) <p>MSP(Multiplex Section Protection)</p>	<p>1+1、Bi-directional、Non-Revのみ。 Working/Protectionという物理実装位置による機能の役割はなく、0系/1系という呼称で、状態によって機能の役割が変わる。 (例:「1系現用選択強制切替中」)</p> <p>K1/K2 byte 0系: “0001”、1系: “0002” ITU-Tとは切替優先度が異なる。</p>
SONET	SONET ANSI	<ul style="list-style-type: none"> ANSI T1.105.01 (2000) [Telcordia GR-253-core (旧Bellcore)] <p>APS(Automatic Protection Switching)</p>	<ul style="list-style-type: none"> 1+1/1:N Bi-directional/Uni-directional Revertive/Non-Revertive <p>K1/K2 byte W系: “0001”、P系: “0000” ITU-Tとは切替優先度が異なる。 K2 Byte 6-8 ビットの定義がある</p>



主な違い

	AnnexA	AnnexB
セクション名	W系: “0001”、P系: “0000”	0系: “0001”、1系: “0002”
セクション切替に使うK1/K2Byte	常時Protection系の信号	非本選択系の信号
Lockout動作	W系で固定、という意味。P系運用中にLOCKOUTを行うと、W系選択に切り替わる。	現状で固定、という意味。Lockout信号はなく、リクエストの送信を凍結する。

AnnexA Format

	1	2	3	4	5	6	7	8
K1Byte	リクエスト				Protectionを使用するCH			
K2Byte	Protection使用中のCH				0	セクションの警報		

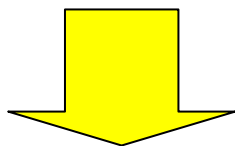
AnnexB Format

	1	2	3	4	5	6	7	8
K1Byte	リクエスト				リクエスト要求元セクション			
K2Byte	本選択中のセクション				0	セクションの警報		

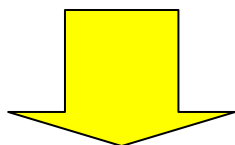
切替優先順位

	AnnexA		AnnexB	
優先順位	要求内容	bit表示	要求内容	bit表示
	ロックアウト	1111	—	
	強制切替え(FS)	1110	強制切替え(FS)	1110
	信号不良(SF)	1101	—	
	—		信号不良(SF)	1100
	信号劣化(SD)	1011	—	
	—		信号劣化(SD)	1010
	手動切替え(MS)	1000	—	
	復旧待ち(WTR)	0110	復旧待ち(WTR)	0110
	確認応答(RR)	0010	確認応答(RR)	0010
	切戻し不要(DNR)	0001	—	
	要求なし(NR)	0000	要求なし(NR)	0000

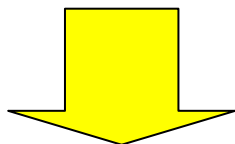
- ✓ 歴史的な経緯より、日本国内ではAPS/MSPはAnnexB(NTT仕様)が圧倒的に多い。
- ✓ 海外ではAnnexA or SONETしか使わない。
- ✓ AnnexAとAnnexBでは仕様に大きな違いがあり、相互接続は不可。



- ✓ ATM装置のマイグレーションは必要。
- ✓ セクション2重化のため、およびマイグレーションシナリオ上、AnnexBは必須。

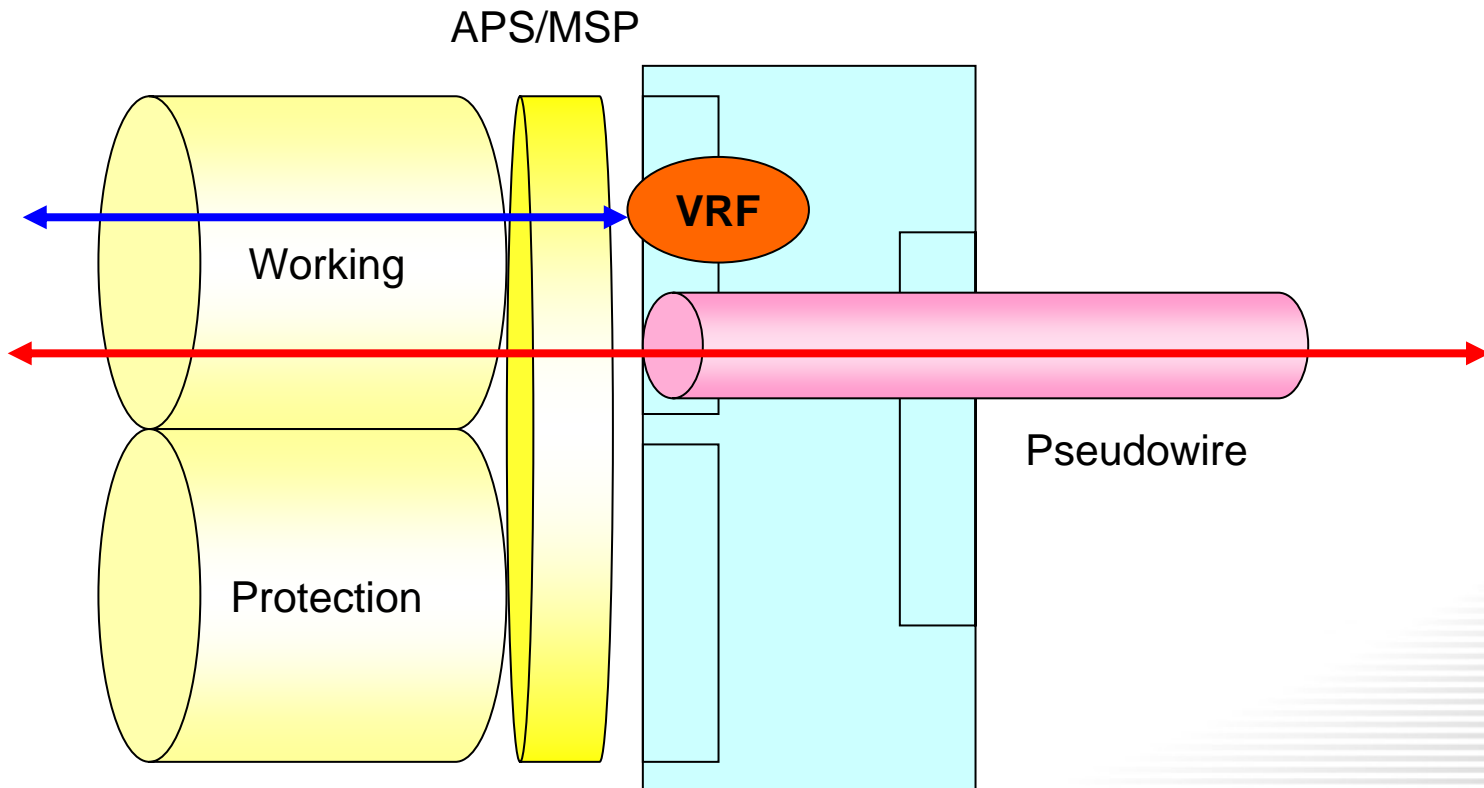


- ✓ 国内メーカーはATMは開発したがない。特にマルチサービスエッジ系はほとんどない。
- ✓ 海外メーカーは日本独自仕様(AnnexB)を開発したがない。



マイグレーションのための新装置がない

さらに、ただAnnexB対応するだけではなく、
APS/MSP portとVRF、PWE3の連携が必要

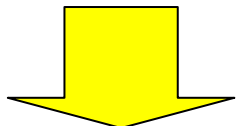


Legacy のイメージ

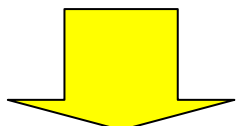
- ✓ Legacy ServiceはHW冗長がある
- ✓ Legacy Serviceは安定している

今までのMPLS

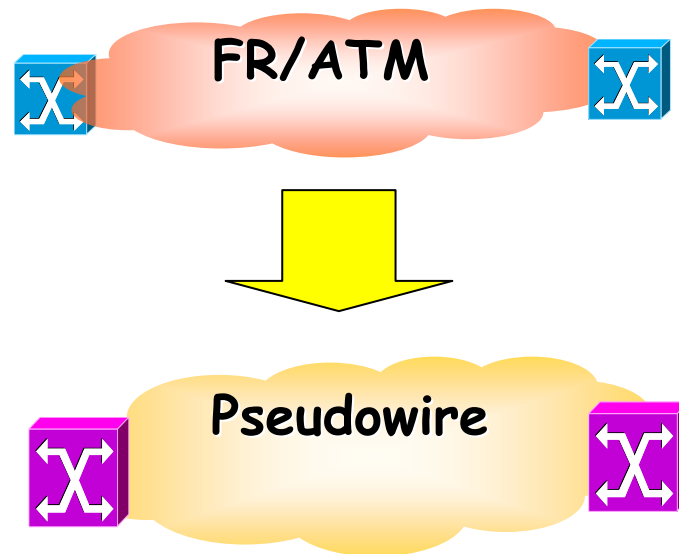
- ✓ なかなか冗長機能を満足する
Multi Service Edgeが登場しない
- ✓ Line Card冗長がまだできない



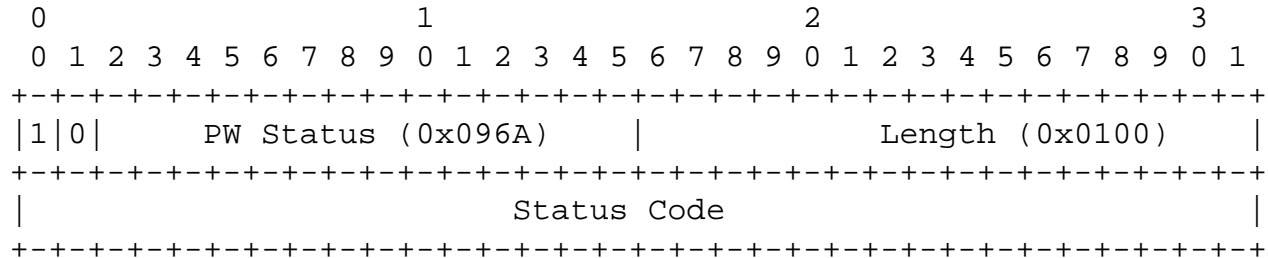
がんばって冗長機能を実装したとして、
装置単体の信頼性を全面的に信用して大丈夫？



装置冗長も考えましょう！



PW Status TLV

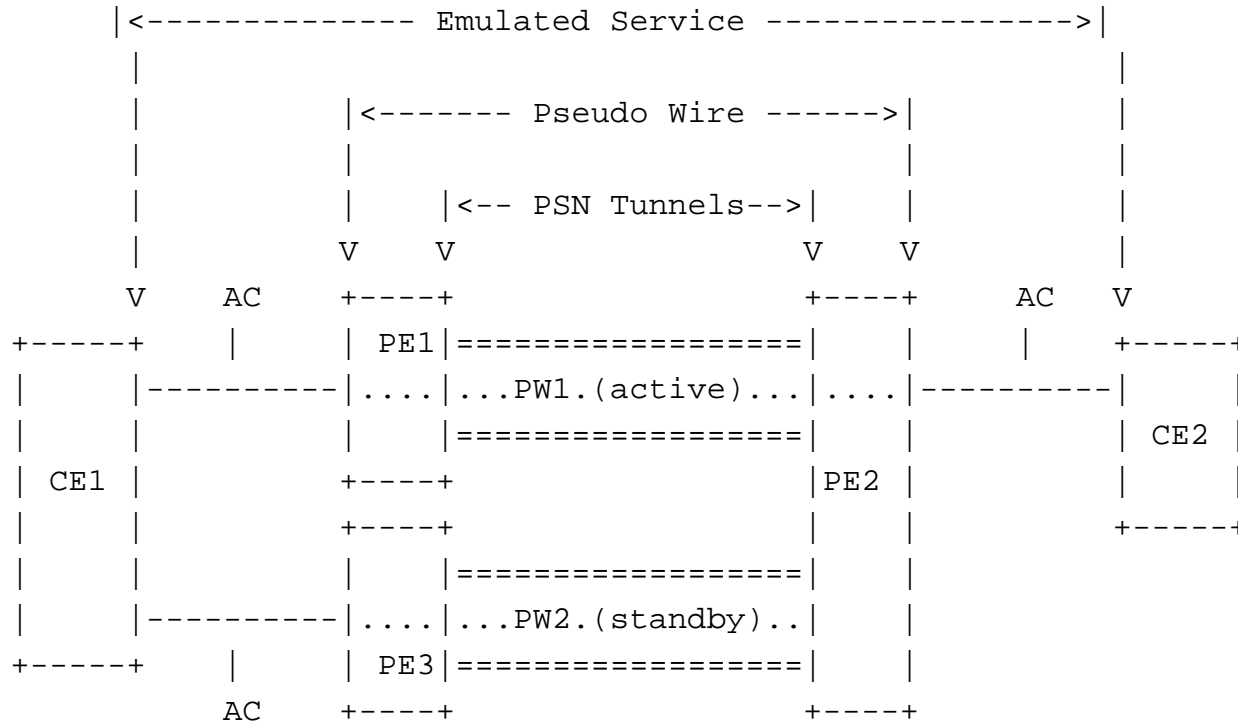


Status Code

Bit Mask	Description
0x00000000	- Pseudowire forwarding (clear all failures)
0x00000001	- Pseudowire Not Forwarding
0x00000002	- Local Attachment Circuit (ingress) Receive Fault
0x00000004	- Local Attachment Circuit (egress) Transmit Fault
0x00000008	- Local PSN-facing PW (ingress) Receive Fault
0x00000010	- Local PSN-facing PW (egress) Transmit Fault

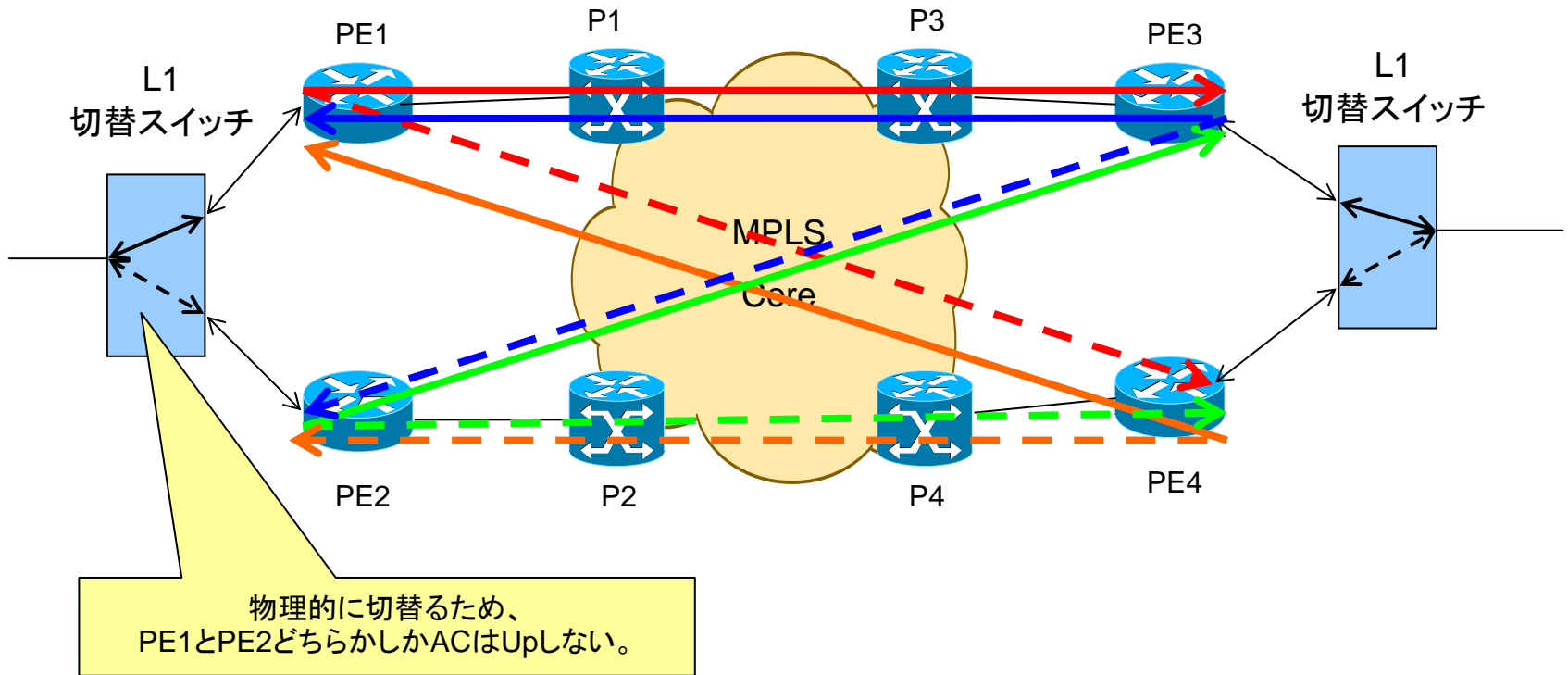
draft-muley-dutta-pwe3-redundancy-bit-01.txt

Bit Mask	Description
0x00000020	- When the bit is set, it represents "PW forwarding standby". When the bit is cleared, it represents "PW forwarding active".



draft-muley-pwe3-redundancy-01

I-D Status: Expired (Expiration Date: August 2007)



PE1 → PE3 Pri

PE1 → PE4 backup

PE2 → PE3 Pri

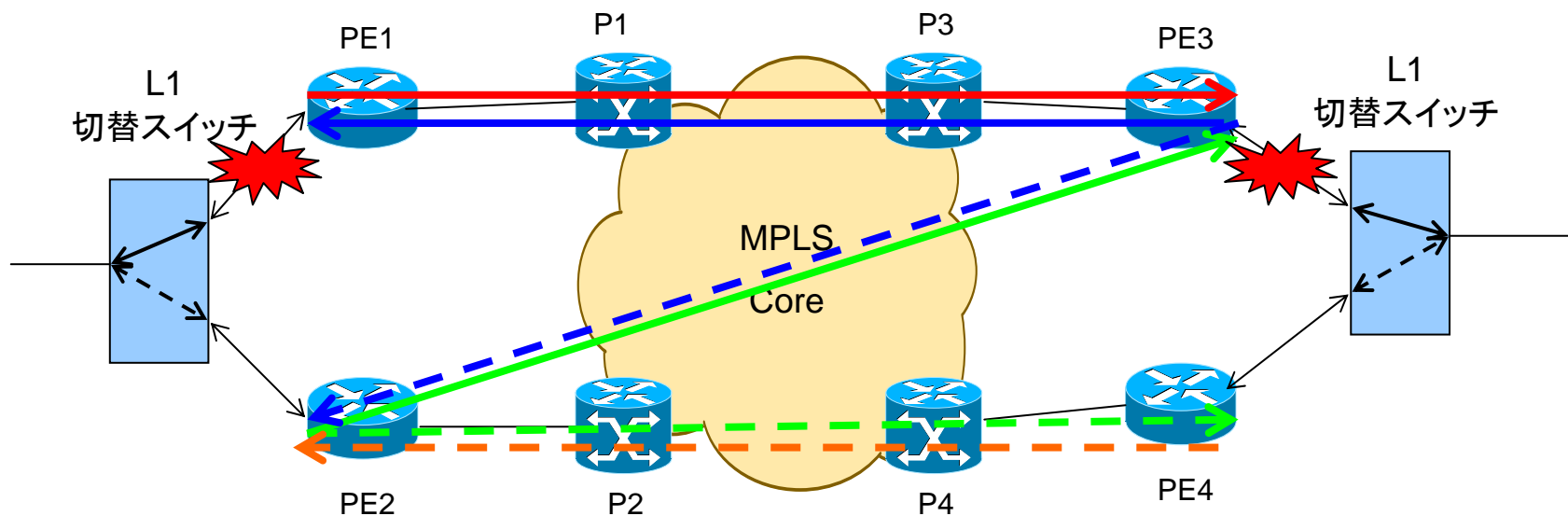
PE2 → PE4 backup

PE3 → PE1 Pri

PE3 → PE4 backup

PE4 → PE1 Pri

PE4 → PE2 backup



PE1 → PE3 Pri

PE1 → PE4 backup

PE2 → PE3 Pri

PE2 → PE4 backup

PE3 → PE1 Pri

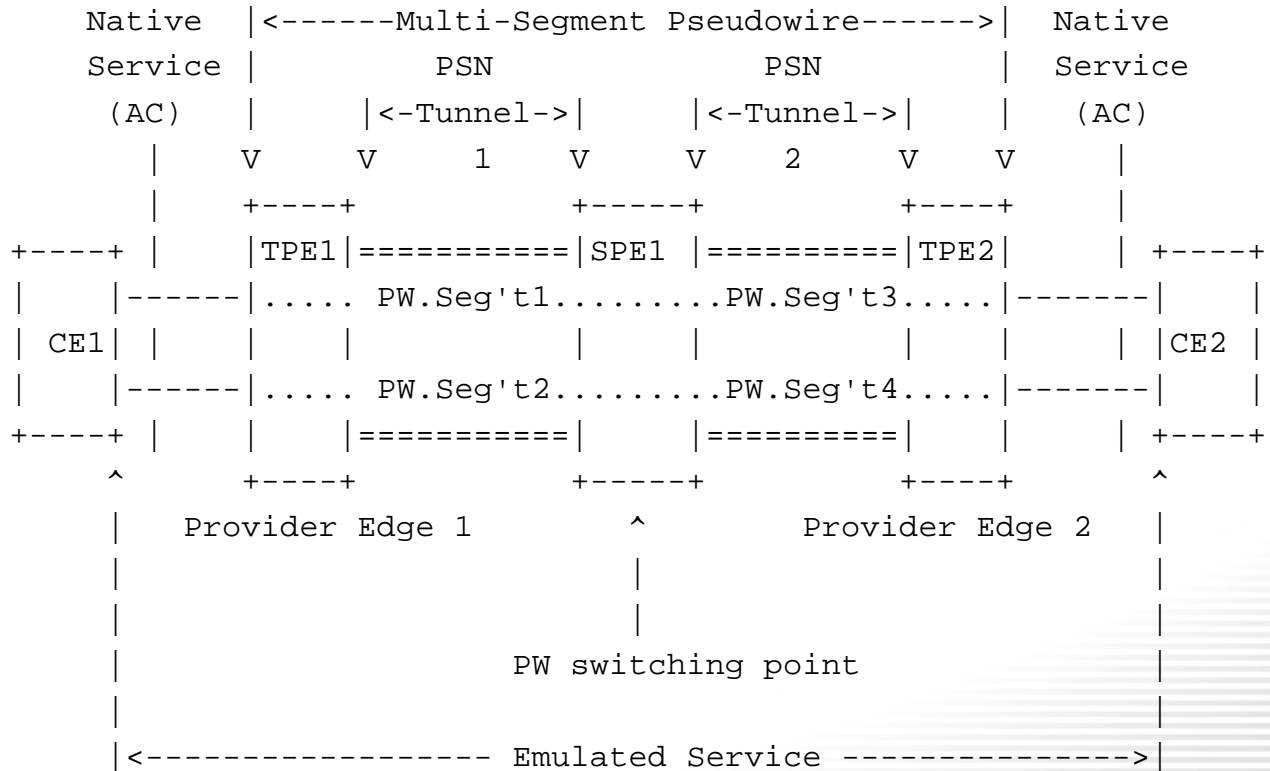
PE3 → PE4 backup

PE4 → PE1 Pri

PE4 → PE2 backup

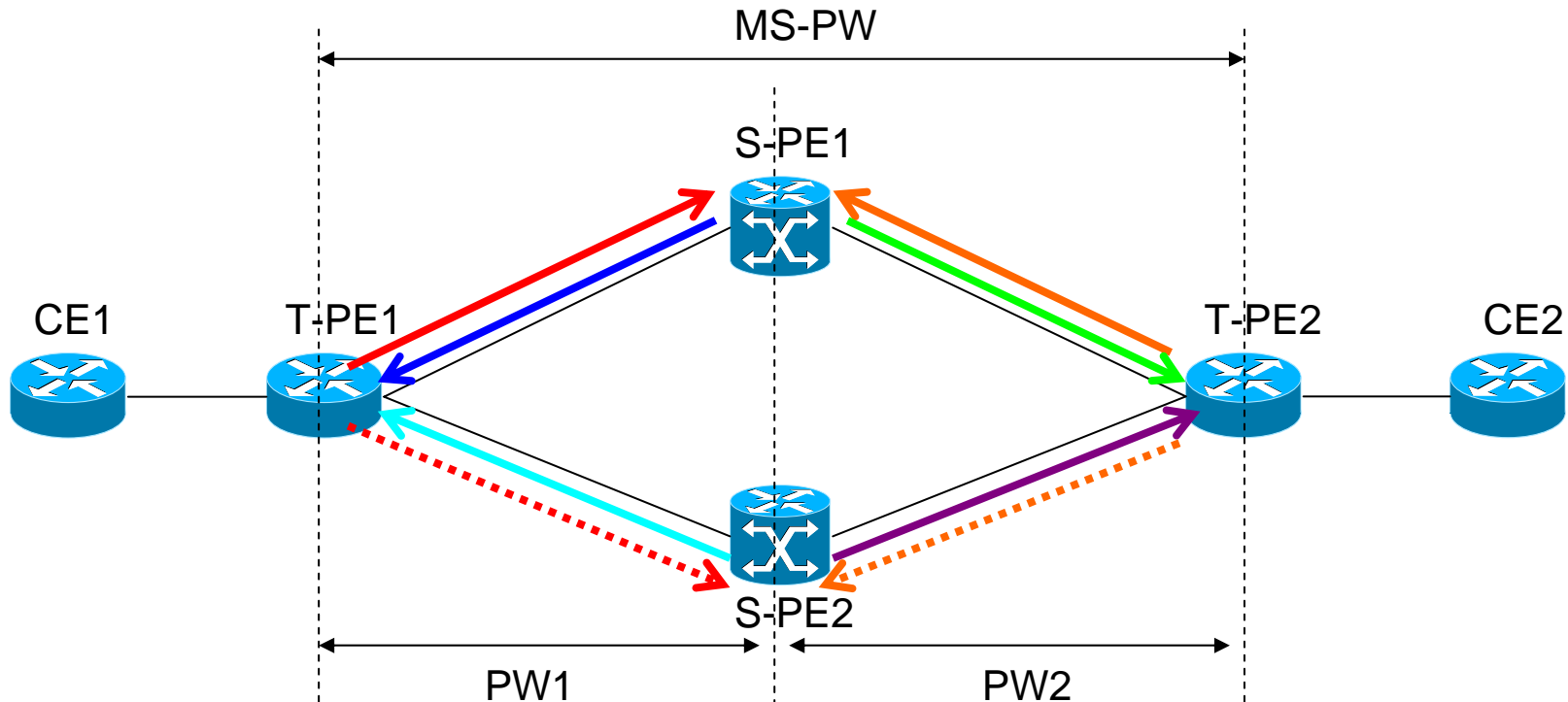
PWE3のMPLS NWが拡大していくと、
 Domainの分離やTunnel LSPのScalabilityを目的として、
 Multi Segment Pseudowireが必要になってくる。

draft-ietf-pwe3-ms-pw-arch-03



MS-PW & PW Redundancy

MS-PWではPW Redundancyが必須。



Supports Extension to VCCV (MH-VCCV)

- T-PE1 → S-PE1 Pri
- T-PE1 → S-PE2 Backup
- T-PE2 → S-PE1 Pri
- T-PE2 → S-PE2 Backup

- S-PE1 → T-PE1 Pri
- S-PE2 → T-PE1 Pri
- S-PE1 → T-PE2 Pri
- S-PE2 → T-PE2 Pri

Traffic Engineering/Management & QOS

	ATM	Pseudowire
Signaling	PNNI1.0 AutoRoute(Cisco),etc	T-LDP RSVP-TE
CAC	Per VP/VC to every Interface	Per TE Tunnel to every Interface Per VP/VC to TE Tunnel
TE	Based on routing cost Explicit Path Per VP/VC	Explicit Path Per Tunnel
Protection Rerouting	No Protection Rerouting per each VP/VC Down Time : 数100ms~数10s (depend on # of connections)	Local repair (FRR) Global repair(Path Protection) Down Time : 数10~数100ms

ATM QOSとEXP Mappingの例

ATM			Pseudowire	
Service Class	Discard Priority	優先度	EXP	Queue
CBR	CLP=0		7	1
	CLP=1		6	
rtVBR	CLP=0		5	2
	CLP=1		4	
nrtVBR	CLP=0		3	3
	CLP=1		2	
UBR (UBR+)	CLP=0		1	4
	CLP=1		0	

実際の運用考えると、EXPを全部使うわけにはいかない

>>> EXP MappingとQueueアサインのポリシーが重要。

特に同一Class内のDiscard Priorityの実装。

EXPの拡張はありえないのか？ DSCP Like？

- ✓ Per PWでCACLしたい。
 - Tunnel帯域 >> Σ PW帯域だと、無駄が多くなる。
 - こまめにTunnel帯域増が必要になる。

- ✓ Tunnelごとに收容しているPWと合計帯域、帯域使用率が見たい。
 - 外部DBでの管理は可能だが、
装置側で簡単に見れるコマンド、MIBがあるとうれしい。

- ✓ PWごとに乗せるTunnelを選択したい。Tunnel Selection

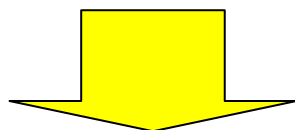
- ✓ EXP 8Classの制限が何とかならないか

- ✓ PWを提供するMPLS NWの拡張を考慮して、
Inter-Area TEは必須。(with Protection)

OAM

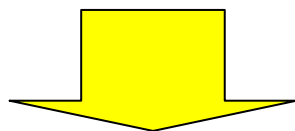
専用線:

伝送装置には回線レベルのFull Rate試験機能がある
(Drop/Insert 内部試験機、DSU Loop2)



ATM専用線:

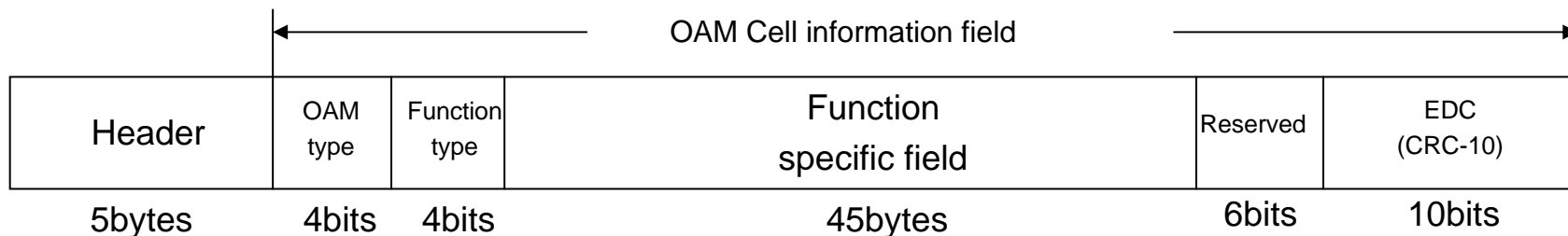
ATM-SWにはVP、VCLレベルのFull Rate試験機能がない



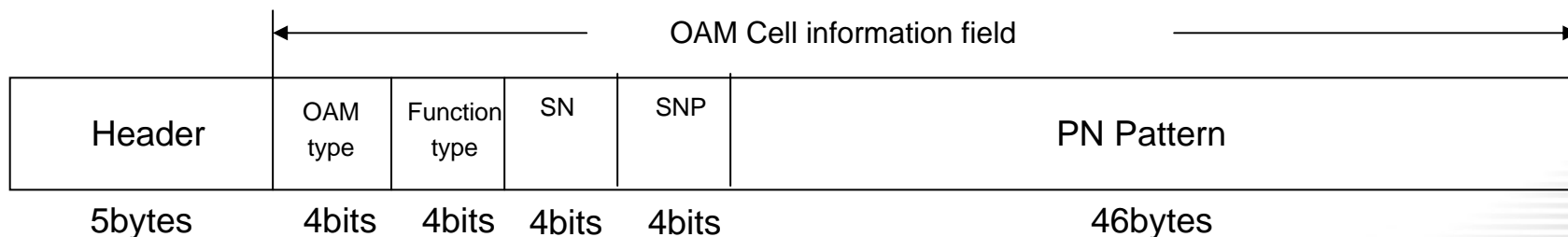
伝送装置のような、Drop/Insert 内部試験機を
ATM-SWに実装しよう！

これを使って、開通時や切り分け時の回線試験をしよう！
相互接続回線でも使いましょう！

標準OAM Cell Format



導通試験 Cell Format (NTT仕様?)



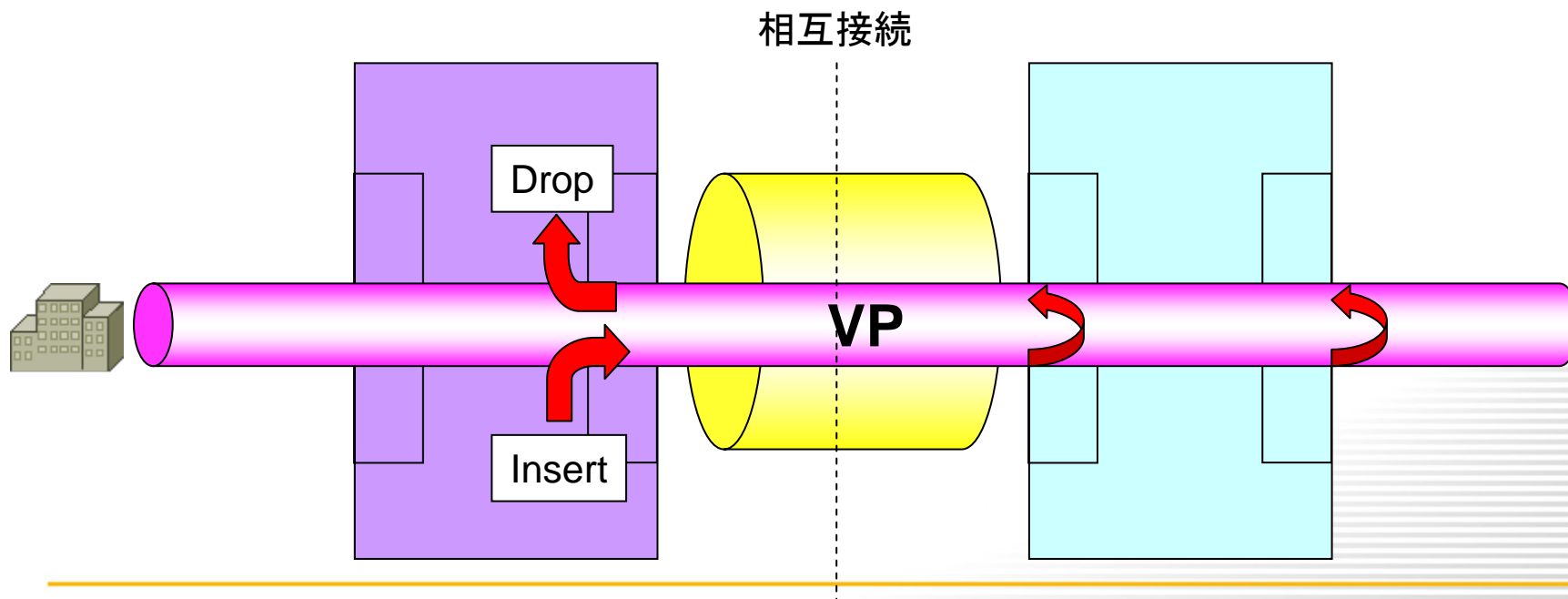
OAM Type=0011、Function Type=0000 : 標準にはない。

CRCがないため、ATM-SWで違反セルとして廃棄される。

特に相互接続においては、帯域・品質保証や切り分け・正常性確認のため、VPやVC単位のFull Rate試験機能が必要。

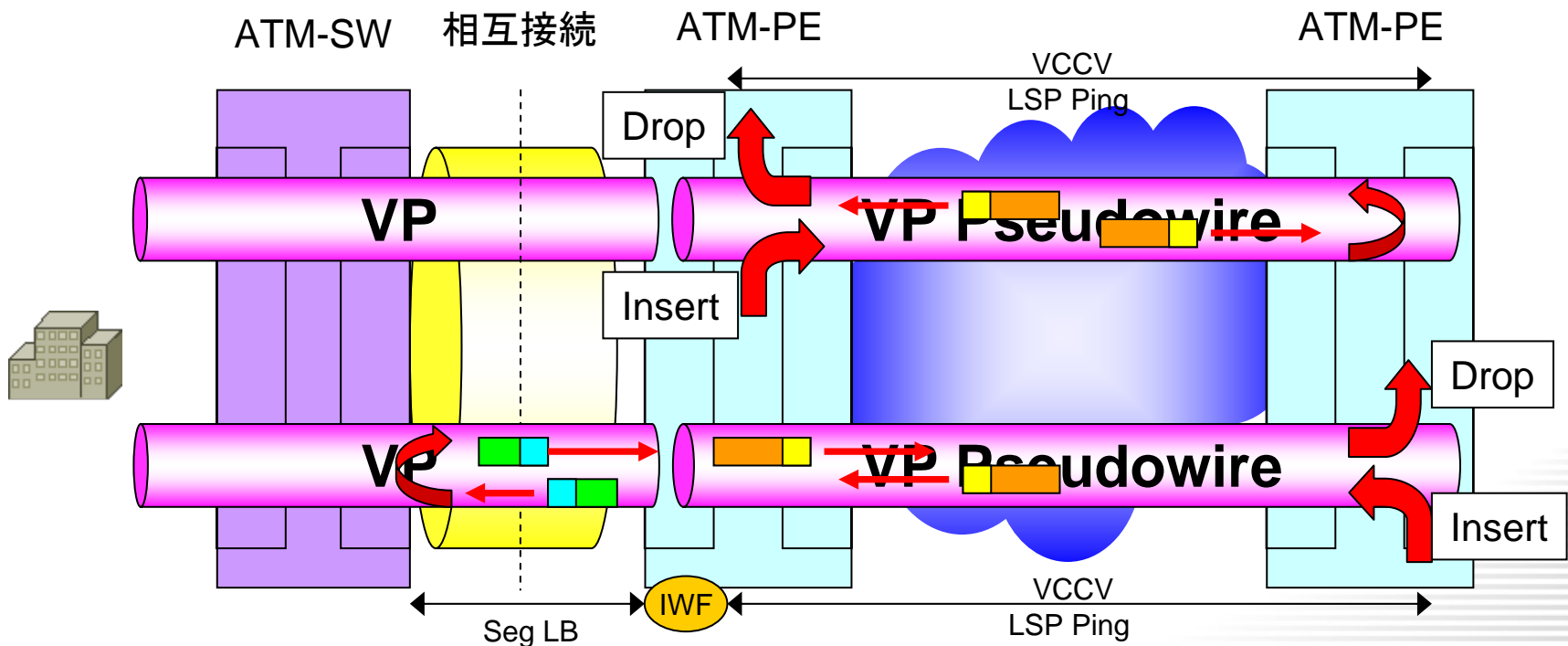
特殊なセルを使用してもテストできるように、HeaderのVPI/VCIのみを見て、単純にLoopbackできる機能がほしい。

しかし、標準セル以外は廃棄するように作りこまれたATMチップが多く、対応可能なベンダーは少ない。



VCCV(Virtual Circuit Connectivity Verification)により、PWのConnectivity Checkが可能。
従来の標準ATMにはなかった機能。

今はPingしかないが、Full Rateでのスループット・品質測定ができるまで拡張されるとうれしい。
また、Legacy NWとの相互接続を考えると、Native ATM OAMへの対応も必要。

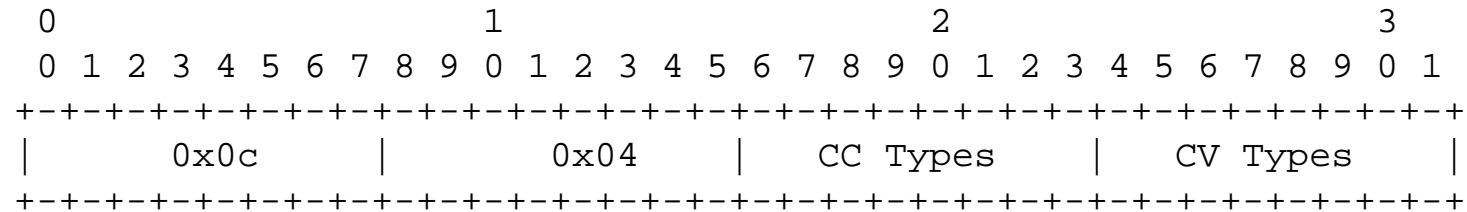


VCCV-OAM interworkなんてできたらおもしろい？

The VCCV parameter ID is defined as follows in [RFC4446]:

Parameter ID	Length	Description
0x0c	4	VCCV

The format of the VCCV parameter field is as follows:



MPLS Control Channel (CC) Types:

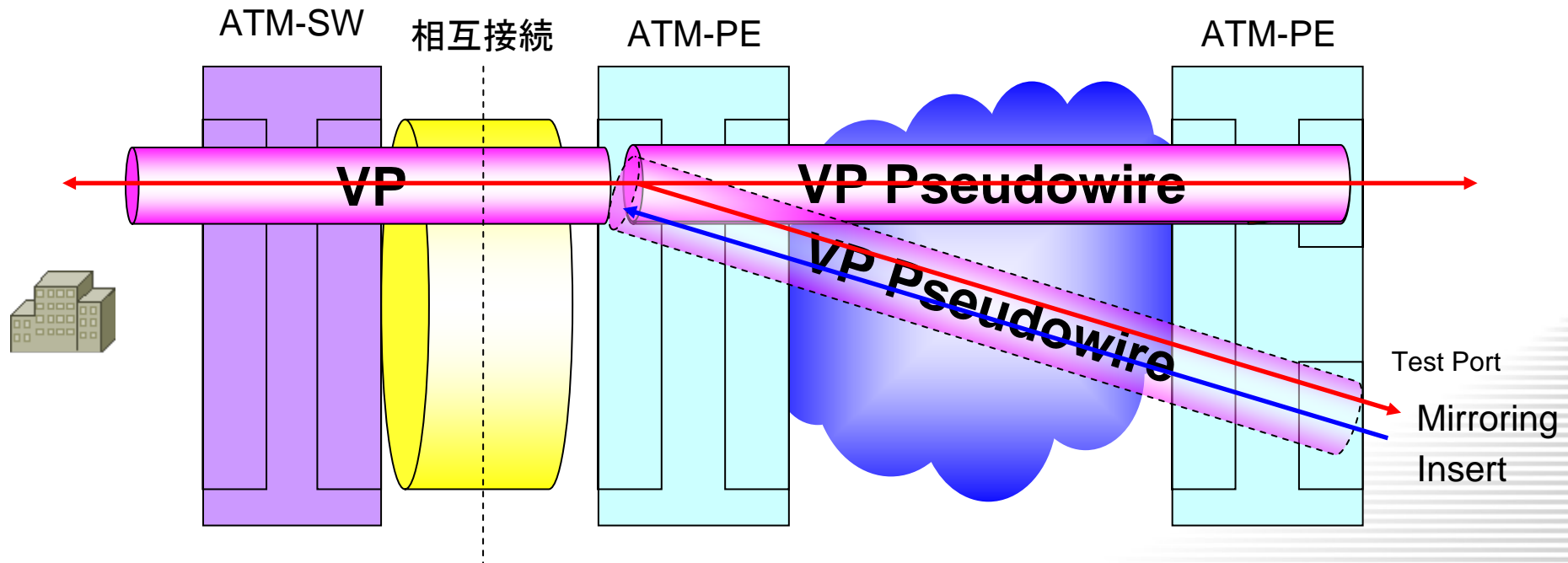
- | Bit (Value) | Description |
|--------------|--|
| Bit 0 (0x01) | Type 1: PWE3 Control Word with 0001b as first nibble (PW-ACH, see [RFC4385]) |
| Bit 1 (0x02) | Type 2: MPLS Router Alert Label |
| Bit 2 (0x04) | Type 3: MPLS PW Label with TTL == 1 |
| Bit 3 (0x08) | Reserved |
| Bit 4 (0x10) | Reserved |
| Bit 5 (0x20) | Reserved |
| Bit 6 (0x40) | Reserved |
| Bit 7 (0x80) | Reserved |

MPLS Connectivity Verification (CV) Types:

Bit (Value)	Description
Bit 0 (0x01)	- ICMP Ping
Bit 1 (0x02)	- LSP Ping
Bit 2 (0x04)	- Reserved
Bit 3 (0x08)	- Reserved
Bit 4 (0x10)	- Reserved
Bit 5 (0x20)	- Reserved
Bit 6 (0x40)	- Reserved
Bit 7 (0x80)	- Reserved

これだけ「Reserved」していれば、当然拡張されていくんだよね？

P2MP PWのような機能があると、外部試験機を接続しての試験など
高度な試験・解析を実施しやすい。
ユーザconnectionは残したまま、内部試験機と同等のことができる。



VPLレベル

F4 segment

F4 end-end

VCLレベル

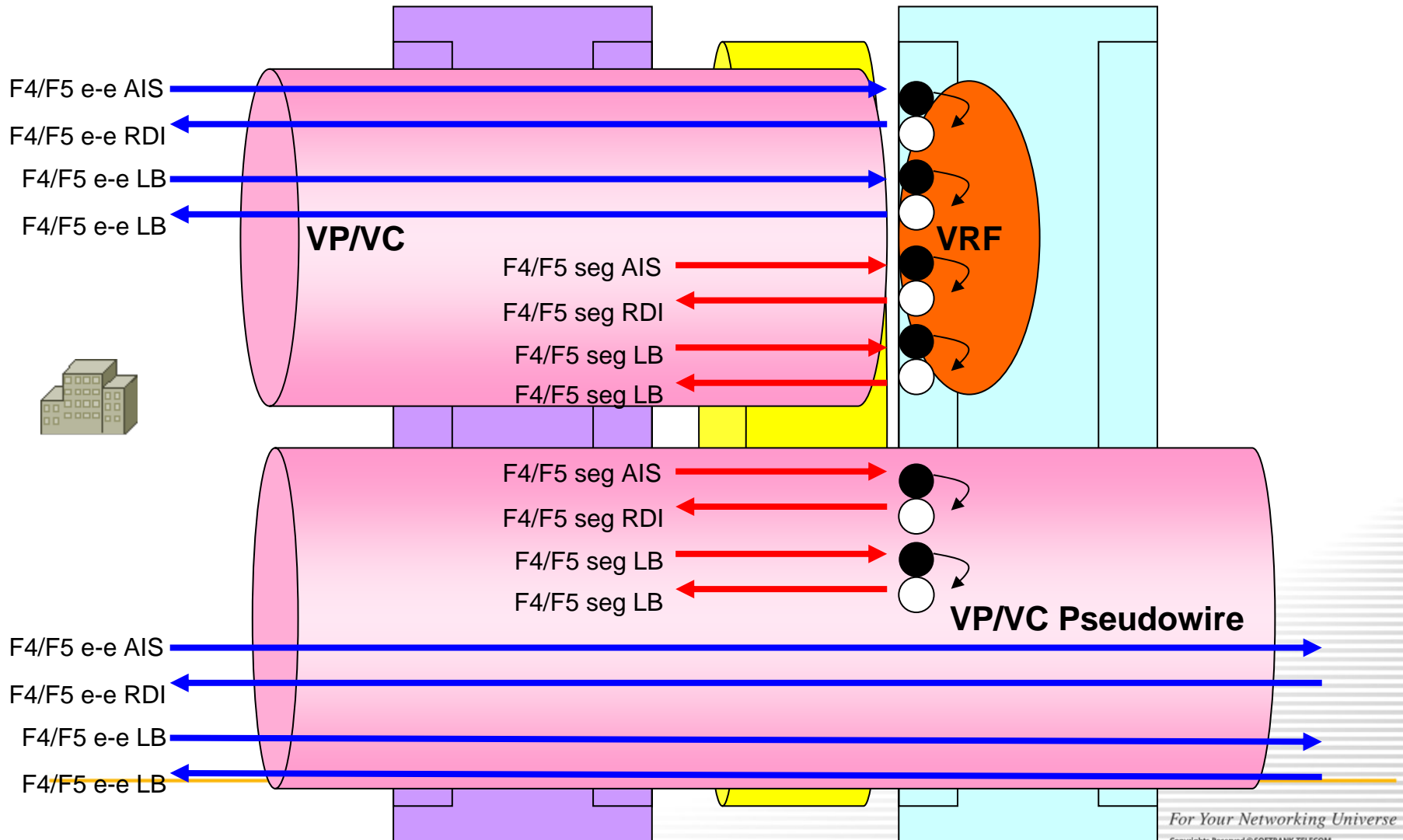
F5 segment

F5 end-end

OAM type	Code	Function type	Code
Fault Management	0001	AIS	0000
		RDI (FERF)	0001
		OAM Cell Loopback	1000
		Continuity Check	0100
Performance Management	0010	Forward Monitoring	0000
		Backward Reporting	0001
		Monitoring and Reporting	0010
Activation Deactivation	1000	Performance Monitoring	0000
		Continuity Check	0001

MPLS Edgeに期待するATM OAM動作

VRFではsegもe-eも全てterminate、PWE3ではsegはterminate、e-eはスルー。



～RFC4717より

VCCの場合

The PEs SHOULD be able to pass the following OAM cells **transparently**:

- F5 Alarm Indication Signal (AIS) (segment and end-to-end)
- F5 Remote Defect Indicator (RDI) (segment and end-to-end)
- F5 loopback (segment and end-to-end)
- Resource Management
- Performance Management
- Continuity Check
- Security

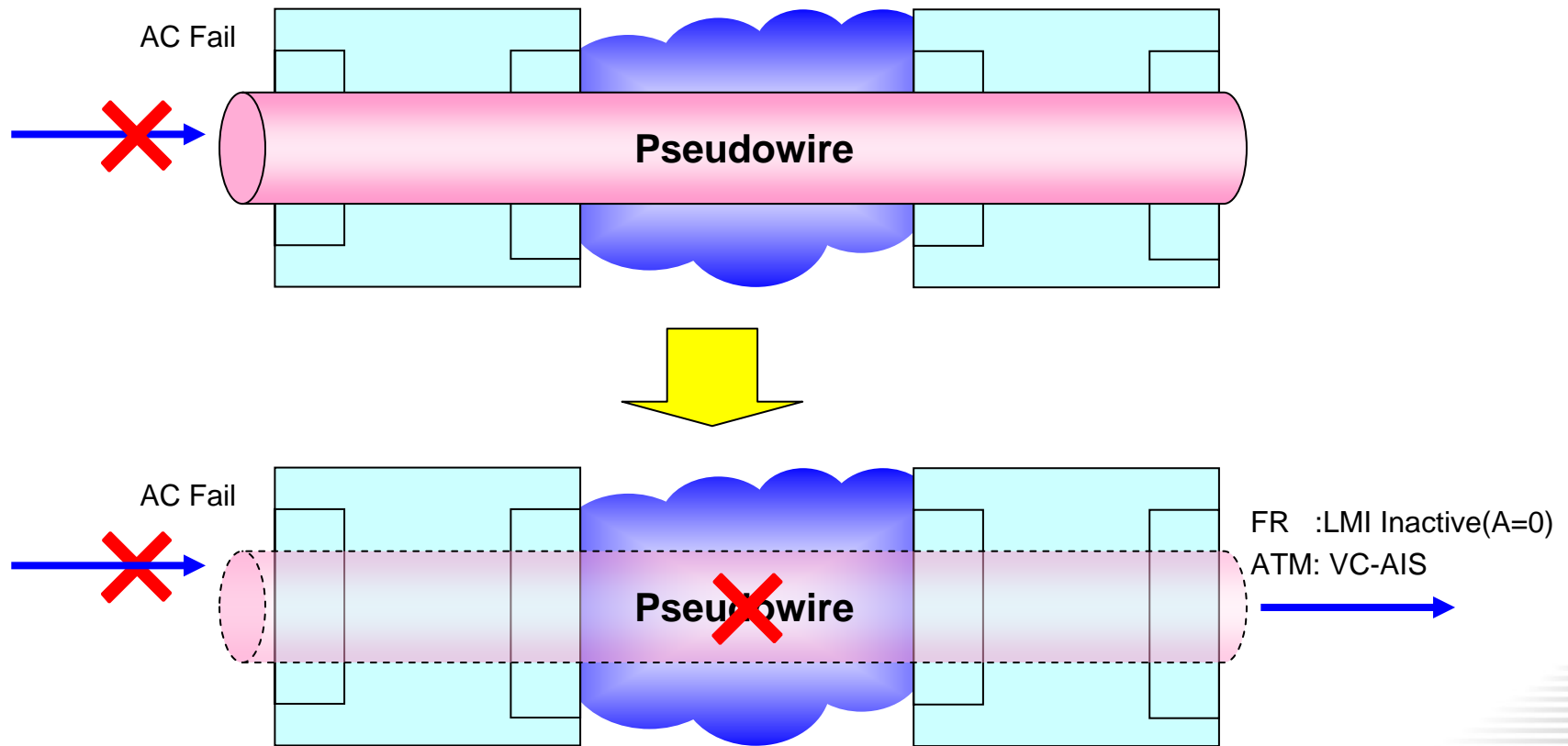
However, if configured to be an administrative segment boundary, the PE SHOULD terminate and process F5 segment OAM cells.

～RFC4717より

If the ingress PE cannot support the generation of OAM cells, it MAY notify the egress PE using a pseudowire-specific maintenance mechanism such as the **PW status message** defined in [RFC4447].

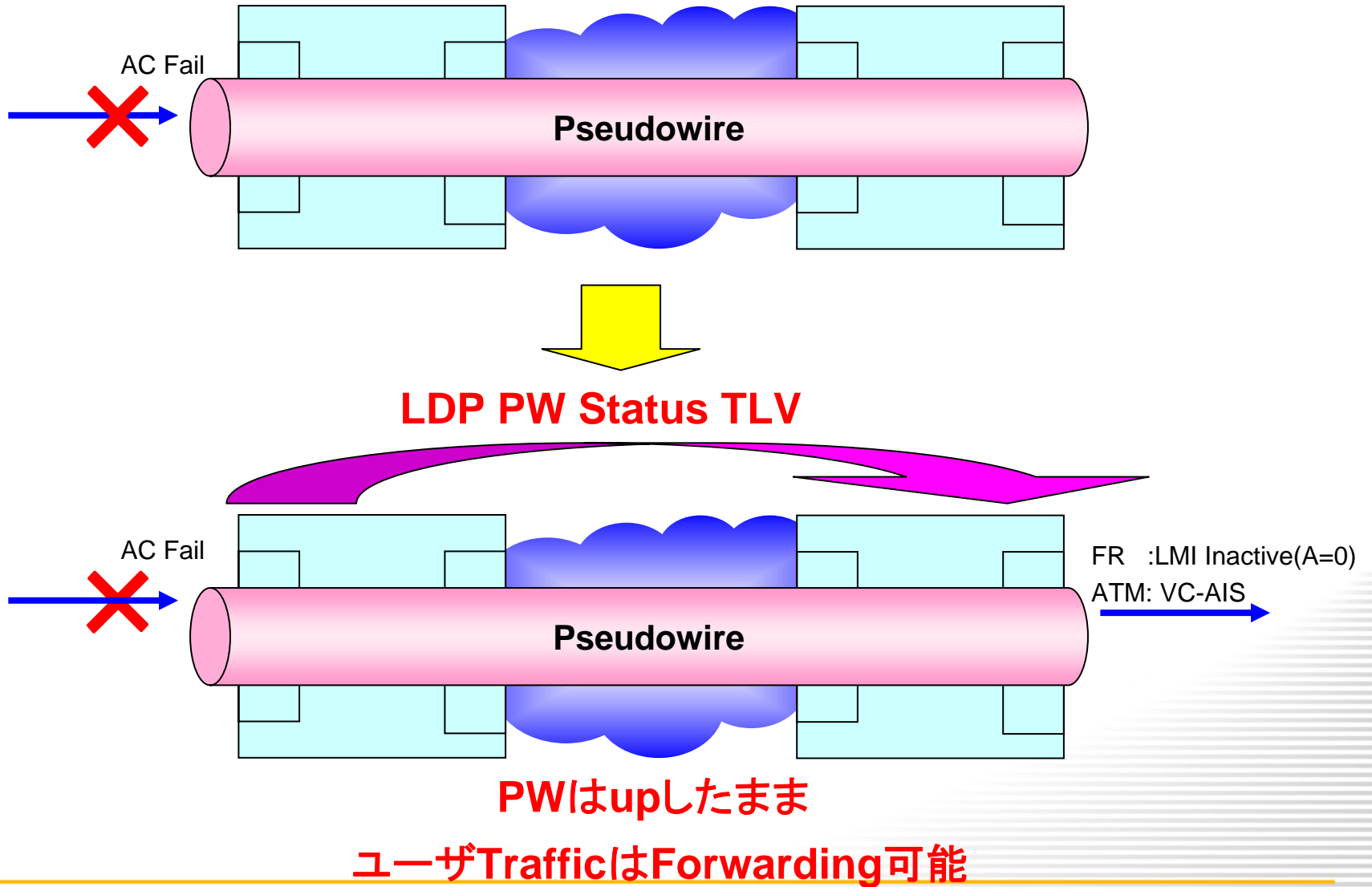
Alternatively, for example, the ingress PE MAY **withdraw the pseudowire (PW label) label** associated with the service.

Upon receiving such a notification, the egress PE SHOULD generate the appropriate F4 AIS (for VPC) or F5 AIS (for VCC).



Label withdrawでPWをDownさせることで通知

復旧時、AC upするまで数10秒間PWがupしない



- ✓ 同一Port上にVRF(terminate end)とPWE3の混在
- ✓ Segmentとend-end OAM cellの区別
(Segment Boundary)
- ✓ 既存ATMとのInterworkを考慮したOAM Mapping

draft-ietf-pwe3-oam-msg-map-05

I-D Status: Expired (Expiration Date: September 2007)

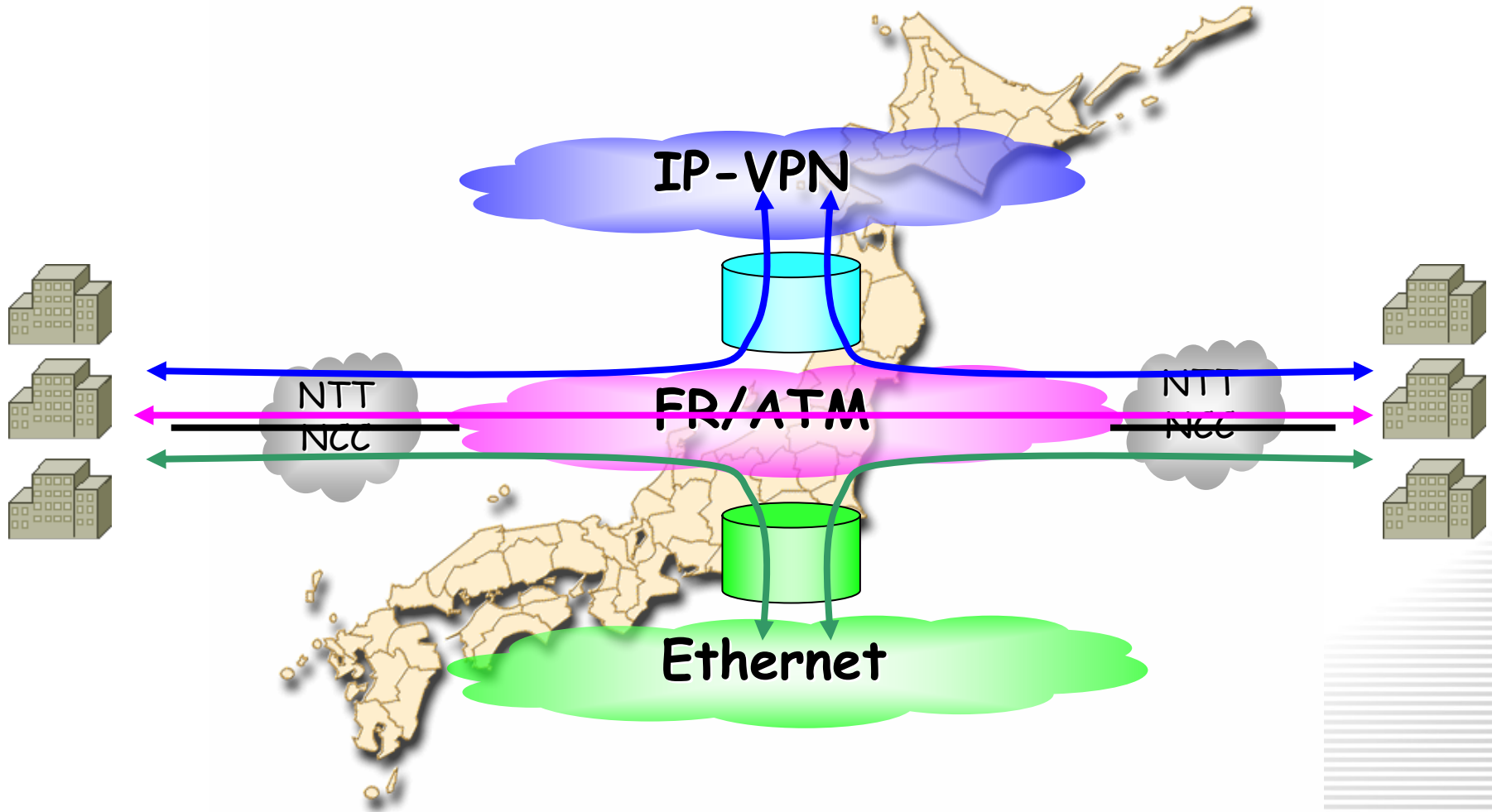
1. Legacy NW Migrationの必要性と目的

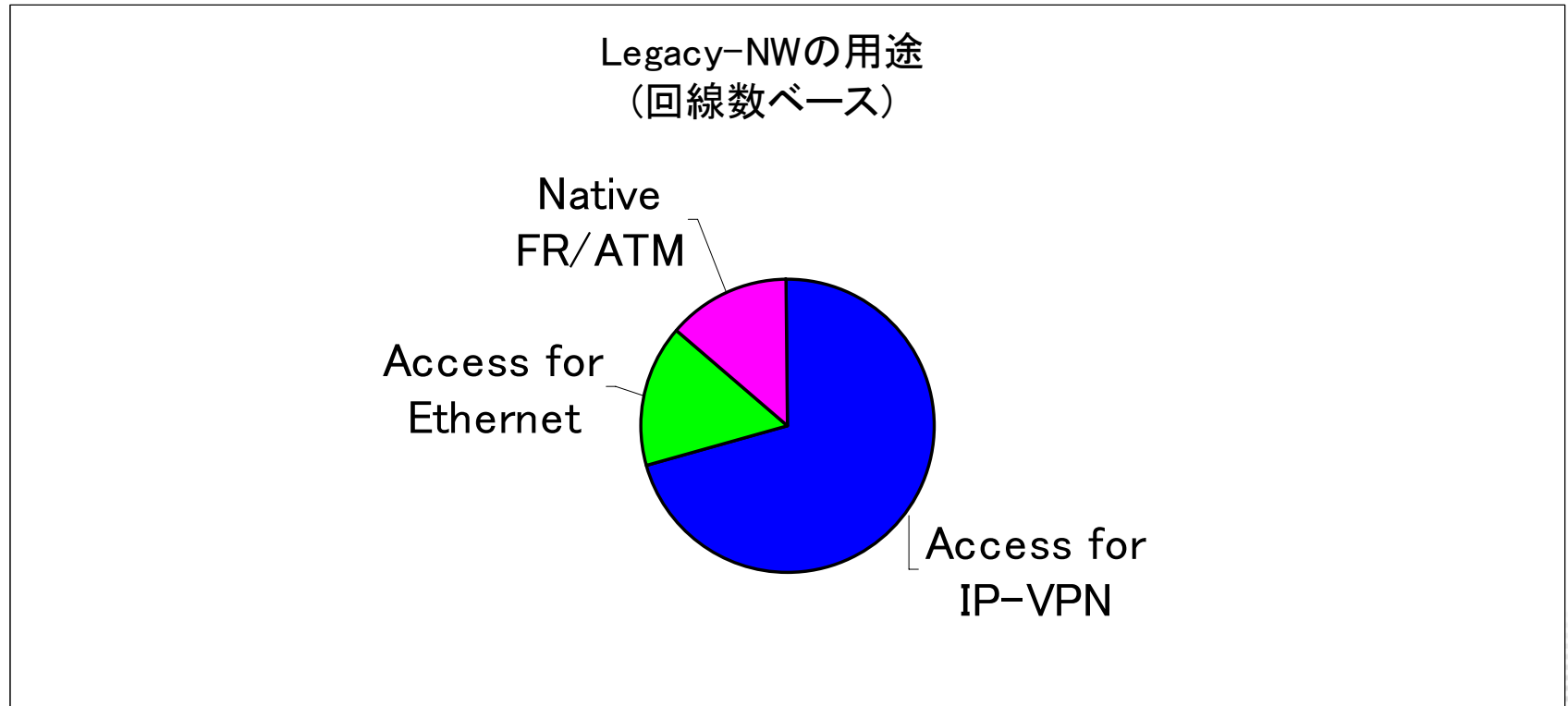
2. Requirement

3. 現状の問題点

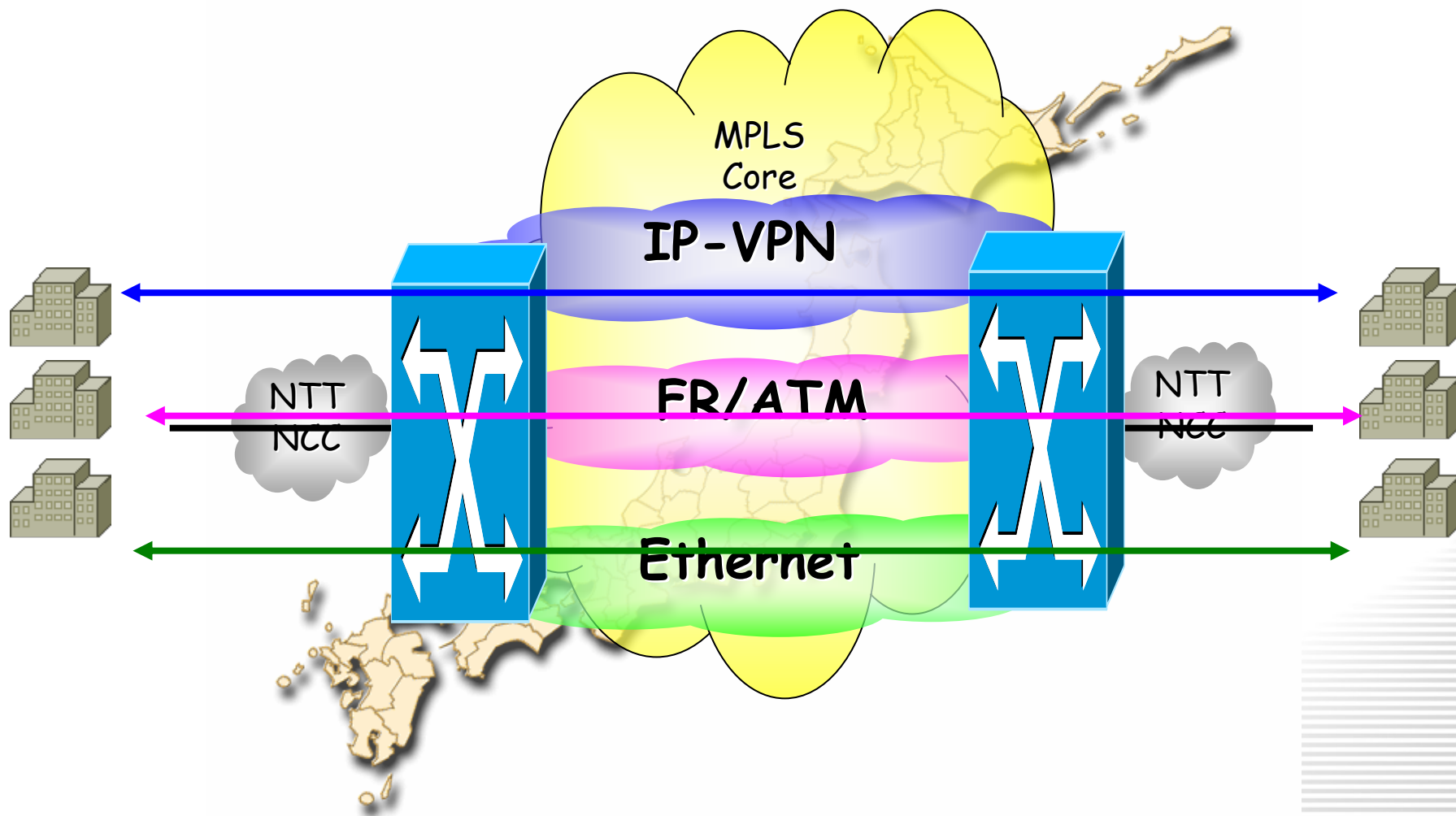
4. Migration Scenarioの検討

5. Summary

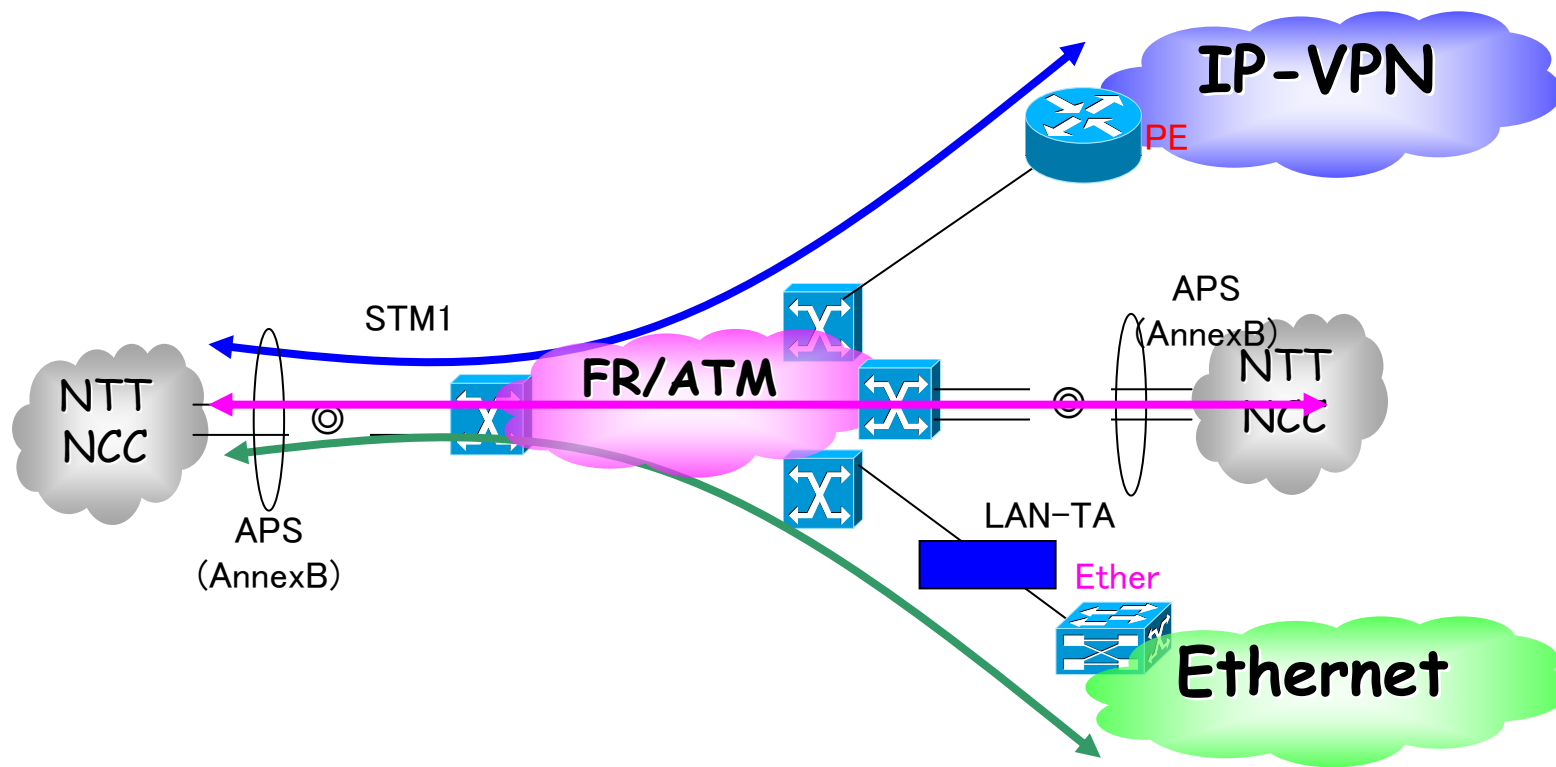




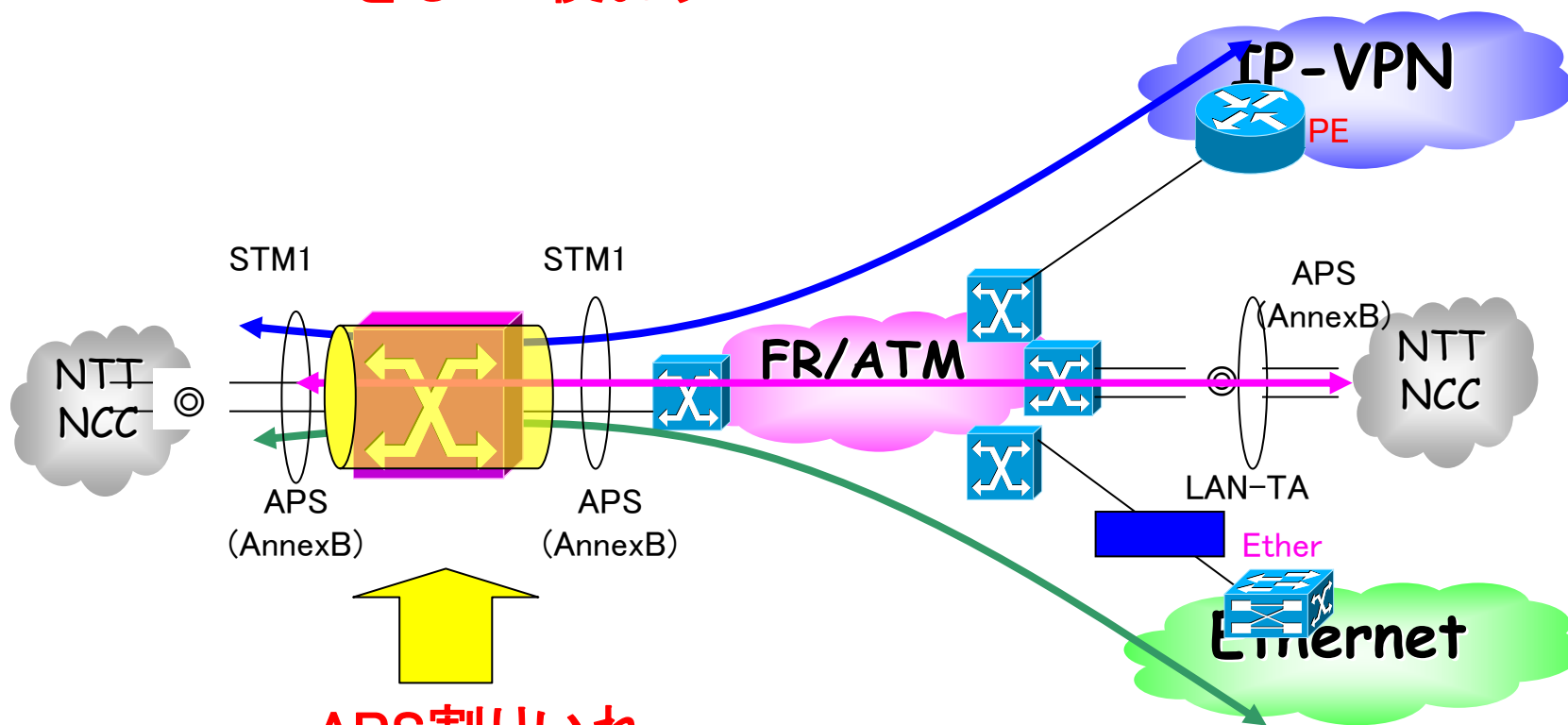
こんなことをやってみたくなる



Migration Scenario step0: 現状



Port Modeをもっと使おう！



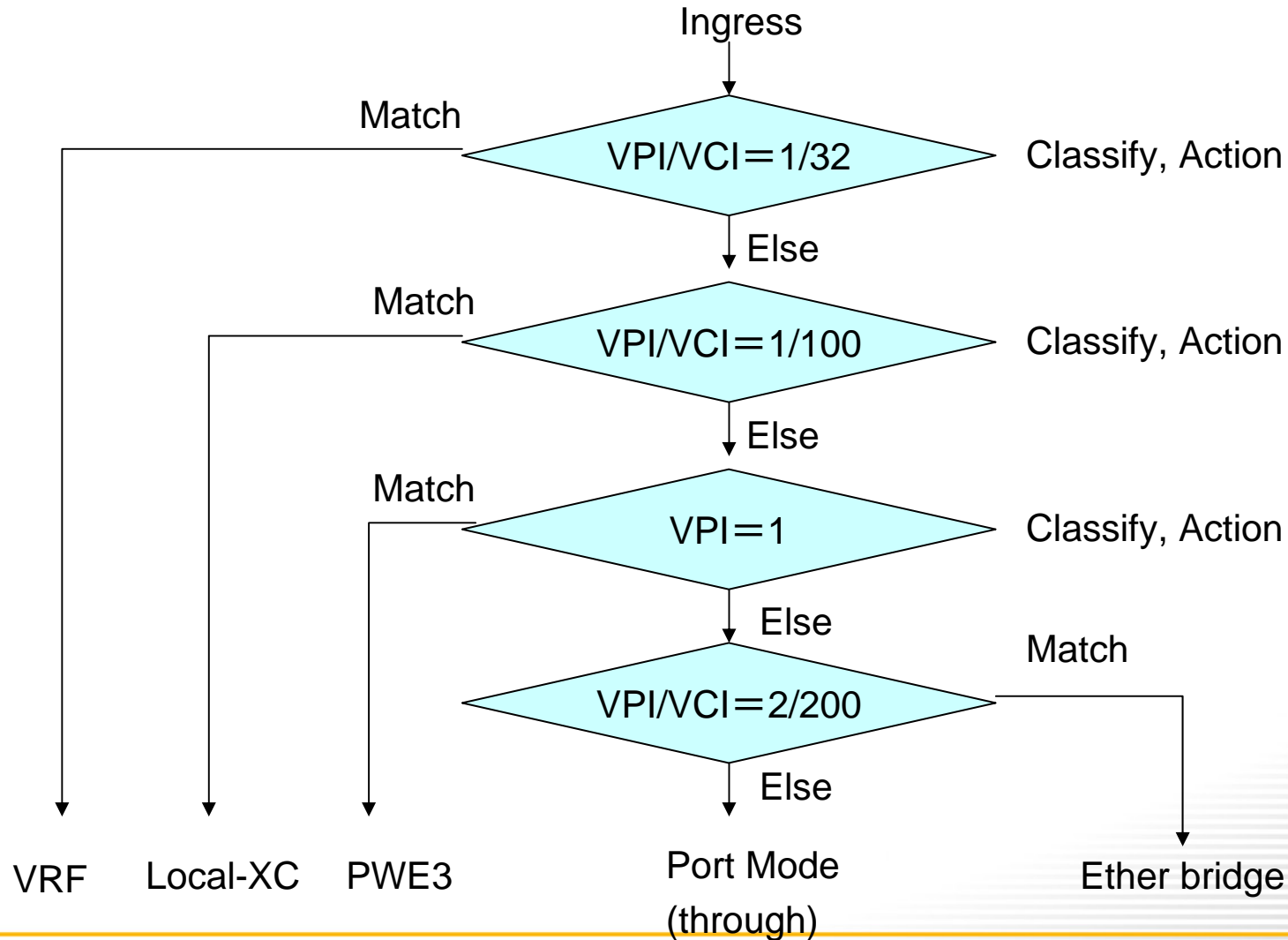
APS割りいれ

Port Mode PWE3(Connetctionパススルー)

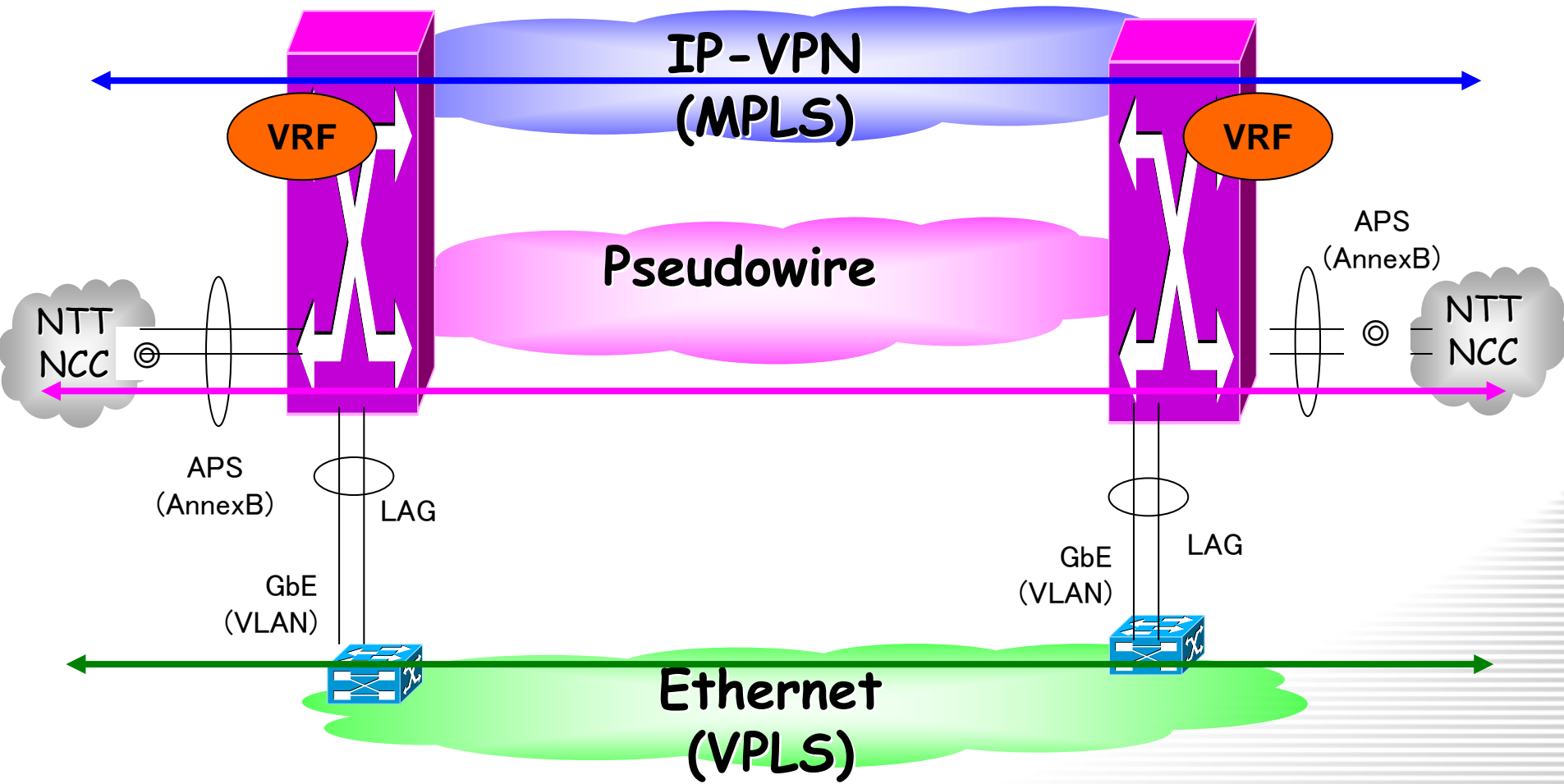
数100のConnectionは設定しないで済む

Multi Service over Port Mode

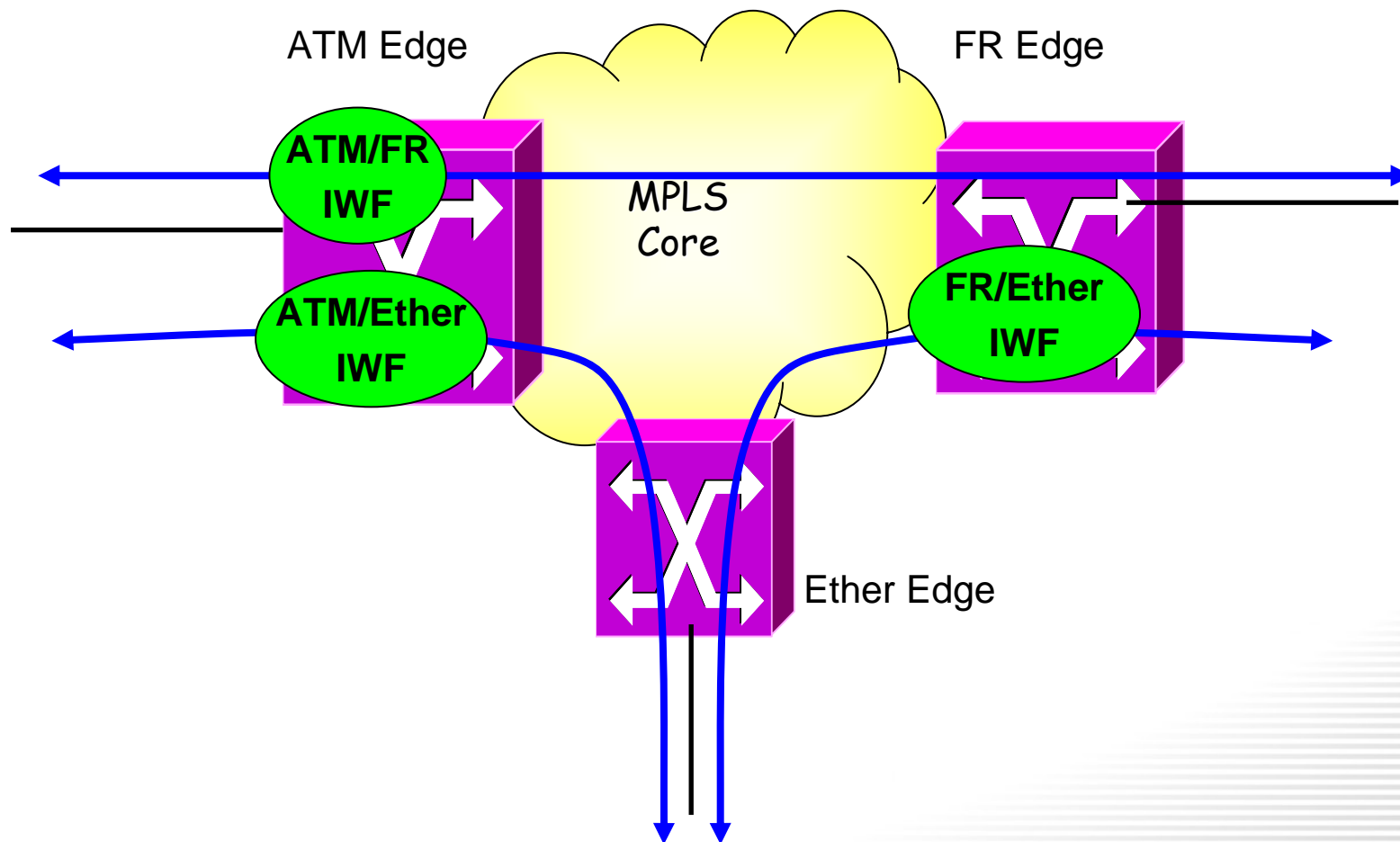
ATM-SWではできないけど、MPLS-Edgeならできる？



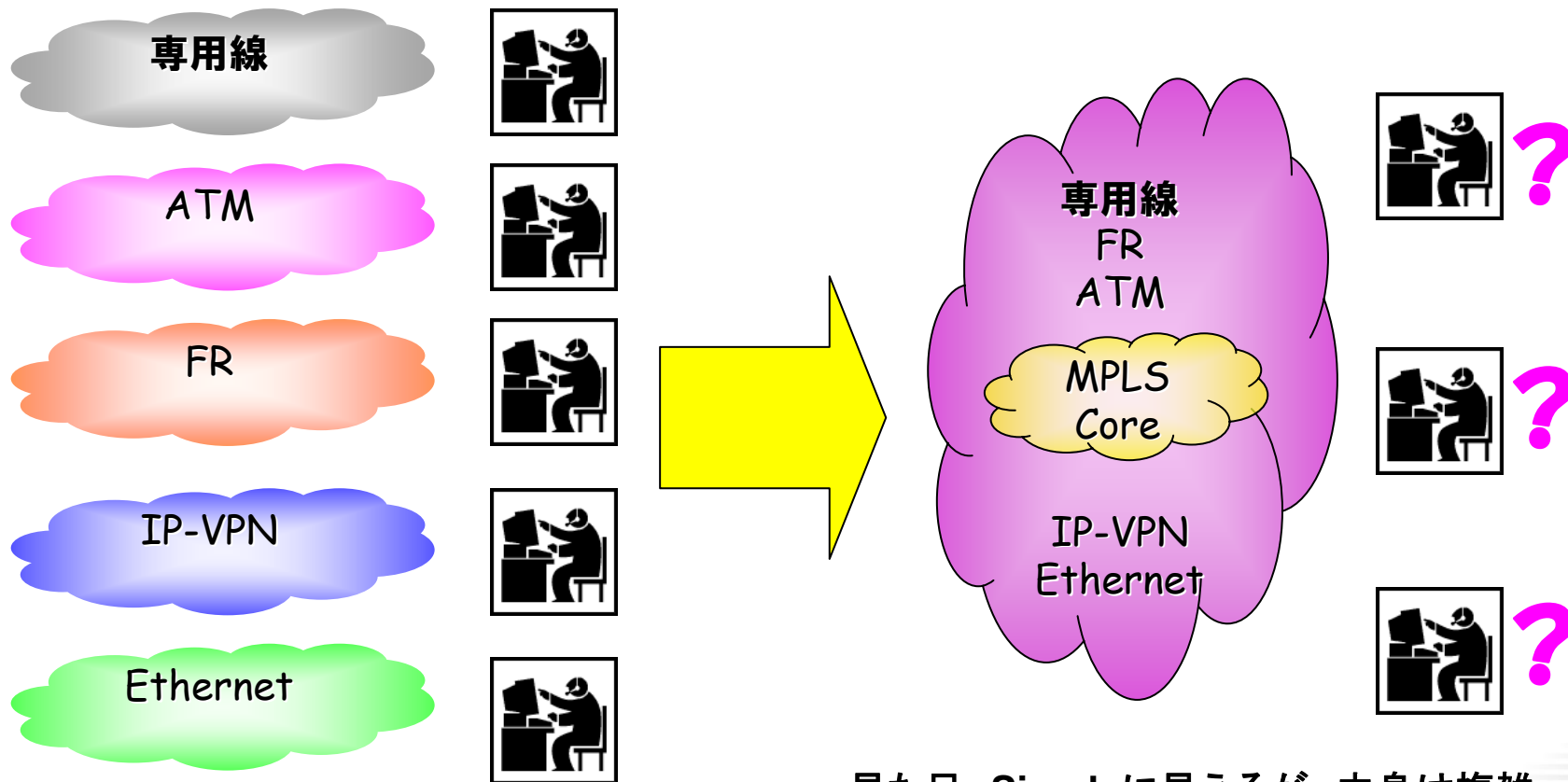
Migration Scenario :完了



LegacyとEtherの共存のためには、Interwork Functionは重要
OAMのInterworkも忘れずに！



みんな簡単に統合って言うけど
でも
Technologyの進歩だけでは
解決しきれない課題がもうひとつ



サービス毎に求められる特性や品質が違う
サービス・装置種別ごとに運用、文化が違う
きれいに運用フローできあがっている

見た目、Simpleに見えるが、中身は複雑
運用体制どうしよう？
Demarcation Pointは機能ブロック？
MPLSエンジニア足りない？
既存NWとのInterwork？

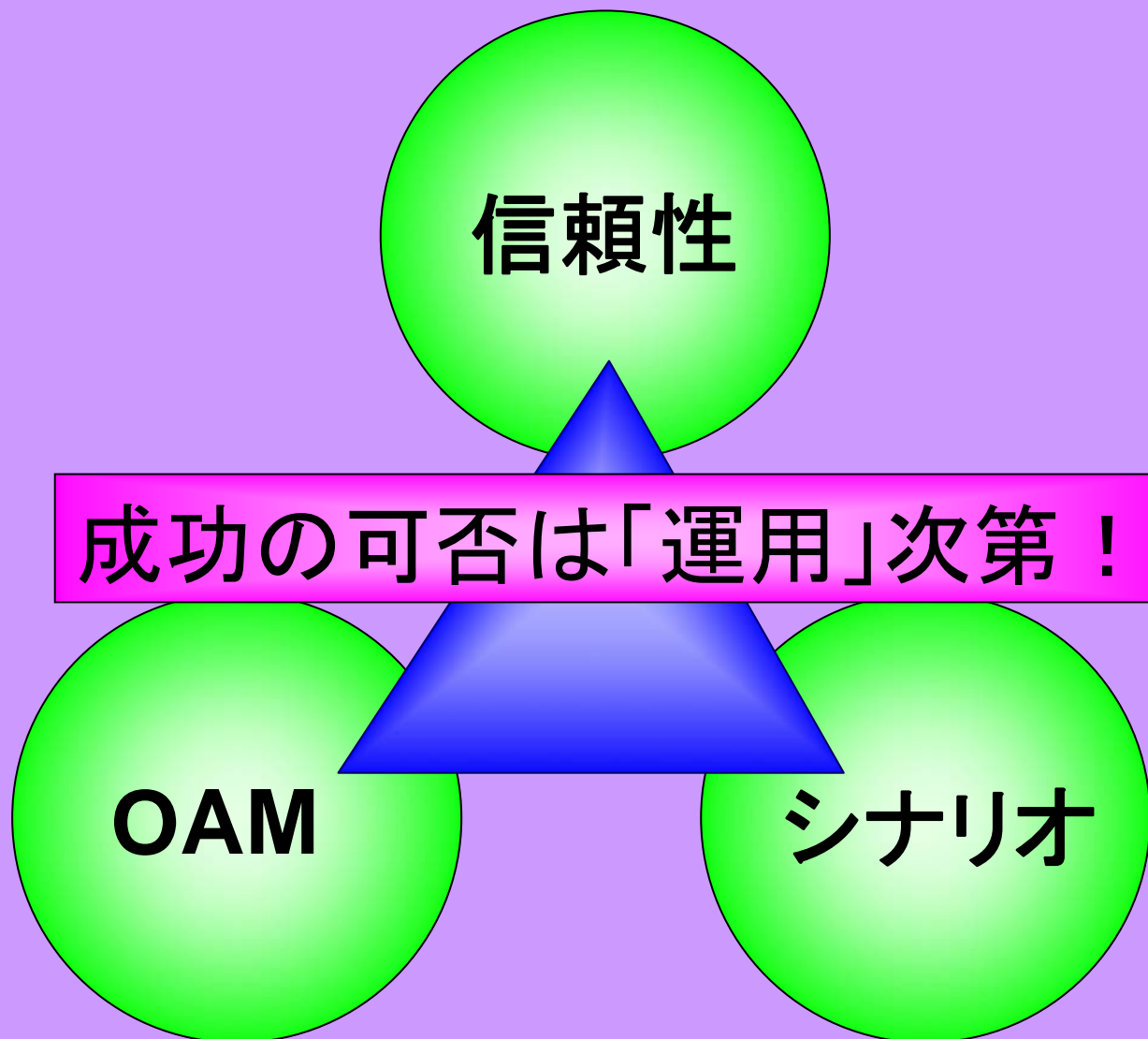
1. Legacy NW Migrationの必要性と目的

2. Requirement

3. 現状の問題点

4. Migration Scenarioの検討

5. Summary



Thank you !

ご清聴ありがとうございました！