

MPLS JAPAN 2008
MPLS/PWE QoSとATM QoS

Tetsuya Innami
SoftBank Mobile Corp.
<tinnami@bb.softbank.co.jp>

October 29th, 2008

ATMをMPLS Pseudowireで!?

■ Motivation。私の場合。

- ✓ いま使ってるATM S/WはそろそろEoLなりそう。ATM自体はまだまだ消えそうもない……。むしろ増えてる?!
- ✓ 普通に考えたら新規のATM S/W買って、入れ替えたり増設すればいいんだけど。。。。
 - ATM S/Wをこの先いつまで使うかと聞かれると、つらい。
 - ATM専用機器だとなかなかお金出してもらえない……。高いし……。
- ✓ そもそも回線の容量が足りてない?
 - ってことは、ATM S/Wの台数も増やしてかなきゃいけない?!
 - Fibre・回線もATM用に取ってこなくちゃいけない?!

■ Routerでなんとかやってみる?!

- ✓ RouterにATMカードをいれてPWE。
 - ATMカード以外、Routerは別用途にも使える。
- ✓ インフラの回線まで含めて集約することで、スケールメリットがでてくれば、コストも安くなるはず。



しかしQoSで大きな落とし穴がありました……。

まずは、基本的なことの復習。



ところでATMのQoSってどんなだったっけ？！

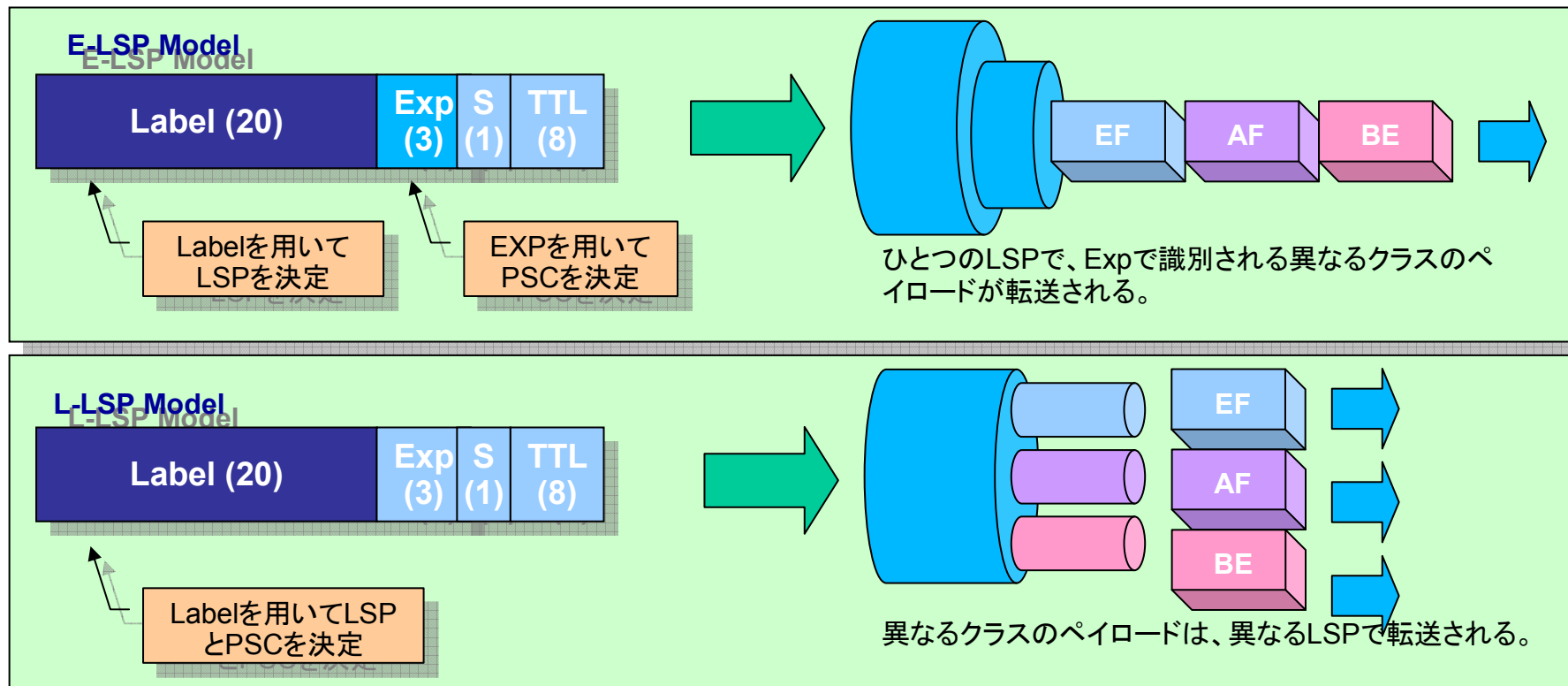
- ATMでは、いろいろなアプリケーション・ペイロードの性質(帯域・バースト性・遅延・etc)に合わせた、サービス品質を提供するメカニズムが用意されています。

代表的なATM QoS Class

CBR (Constant Bit Rate)	固定ビットレート	連続的な一定セル・ペイロードの転送に向いています。送信者が要求する帯域・遅延の等の品質が、ATM網によって保証される最も優先度の高いQoS Classです。
VBR (Variable Bit Rate)	可変ビットレート	セル・ペイロードが変動するような転送に向いています。ATM網が保障するセル・レートとバースト幅の最大値を指定できます。さらに音声データのように遅延に弱いアプリケーション用のrt-VBR、それ以外アプリケーションに使われるnrt-VBRに分類されます。
UBR (Unspecified Bit Rate)	未指定ビットレート	ベストエフォート型の転送に向いています。帯域は保障されませんが、利用可能な空き帯域が割り当てられます。

MPLSにおけるQoS

- ATMと異なり、MPLSの転送メカニズム自体にはQoSを担保する仕組みがありません。
- IP転送と同様に、Label Header内のExperimental Use FieldをコードポイントとしたDiffServ ModelのQoSを使うことが一般的です(E-LSP: EXP-Inferred-PSC LSP)。
- 異なるQoS Classのペイロードを、Class毎に確立したLSPに割り当てる方式も存在します(L-LSP: Label-Only-Inferred-PSC LSP)。



ちょっとややこしいこと。



私がPseudowireにのせたいATM Traffic

■ Pseudowire

- ✓ VPで数千～数万本。
- ✓ 各VPには十数本のVCが含まれる。
- ✓ わりと頻繁にVPC単位の新設・繋ぎ換えの作業が発生します。
- ✓ 運用的な理由により、Pseudowireの数は極力抑えなければならない。

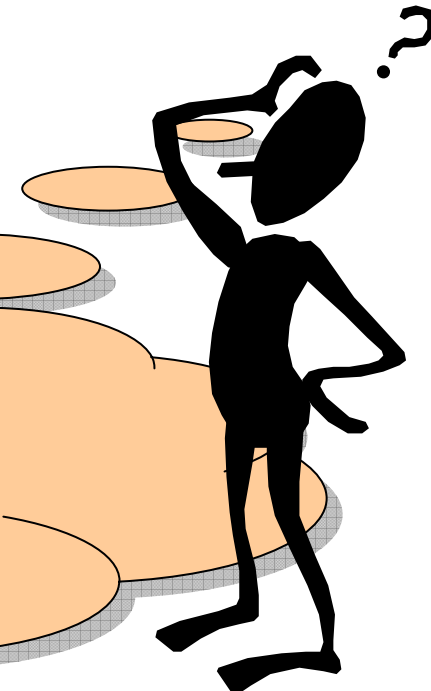
■ QoS Class

- ✓ 1つのVPに含まれるいくつかのVCには、異なるQoS Classのものが含まれてます。
- ✓ 一部帯域を保証する必要があるようなClassがあります。

■ MPLS PSN

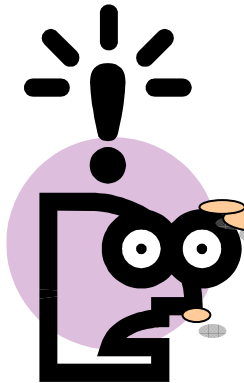
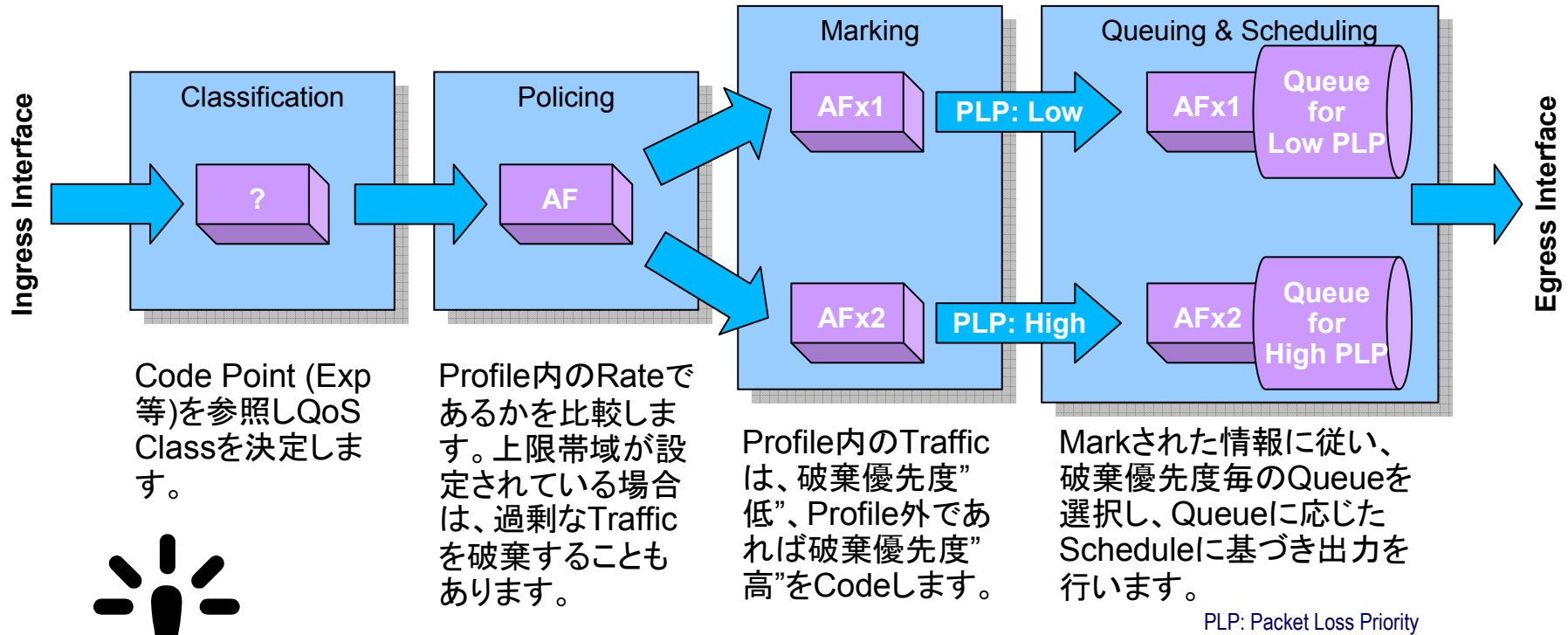
- ✓ ATM PW以外にも多くのTrafficが流れてます。
- ✓ もちろんいろいろなQoS Classが混ざってます。

って言うか、ルーターでやるってことは普通の優先制御でなんとかするってことってだよな?!
がっつり、帯域確保なんて無理っぽいんだけど...



帯域確保が必要なQoS Classの実現

- 特定のClassに属するTrafficの一部の帯域を保証するためには、input rateに応じたMarkingを行いqueuing/schedulingに反映させることで実現します。



ATMのCLPは消えてるかもなので、Oversubscribeしてるときとか気をつけよう。

この動作をHop毎に繰り返します。

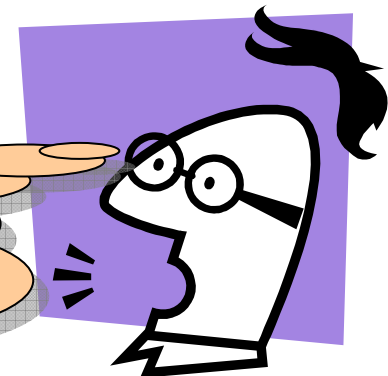
Code Pointの問題

■ E-LSP Modelで使えるCode Point(Exp)は3bitsだけ。

- ✓ $2^3=8$ 通りしかありません。
- ✓ 各Classで破棄優先度を区別する必要があると4 Classしか使用できない…。例えば、

Class Of Service	Packet Loss Priority	Exp
Network Control	Low	7
	High	6
Expedited Forwarding	Low	5
	High	4
Assured Forwarding	Low	3
	High	2
Best Effort	Low	1
	High	0

実際は他のTrafficも含めてもう少し細かい分け方をしたくなるので、少し節約しないとExpだけのCodingでは足らなくなってしまう。。。



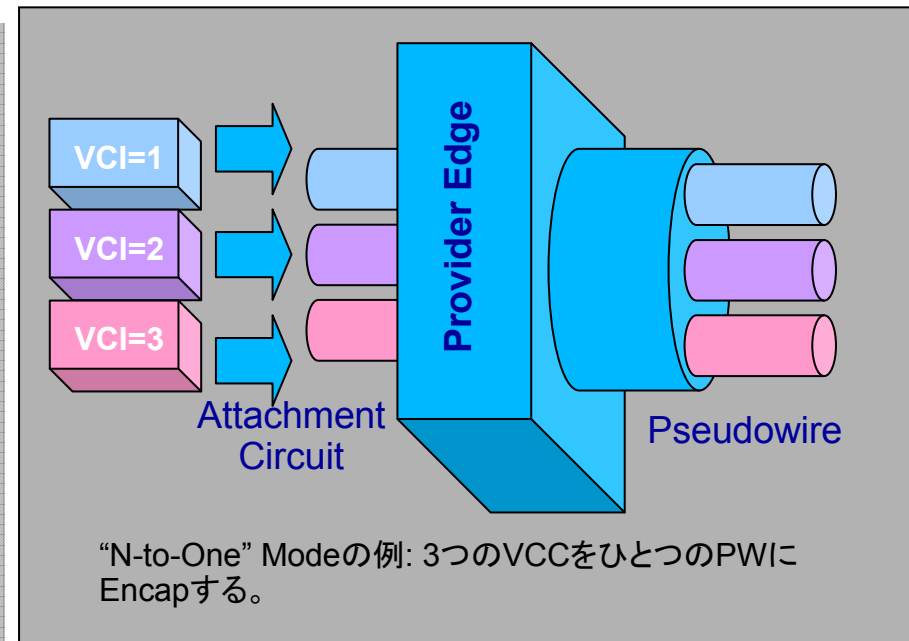
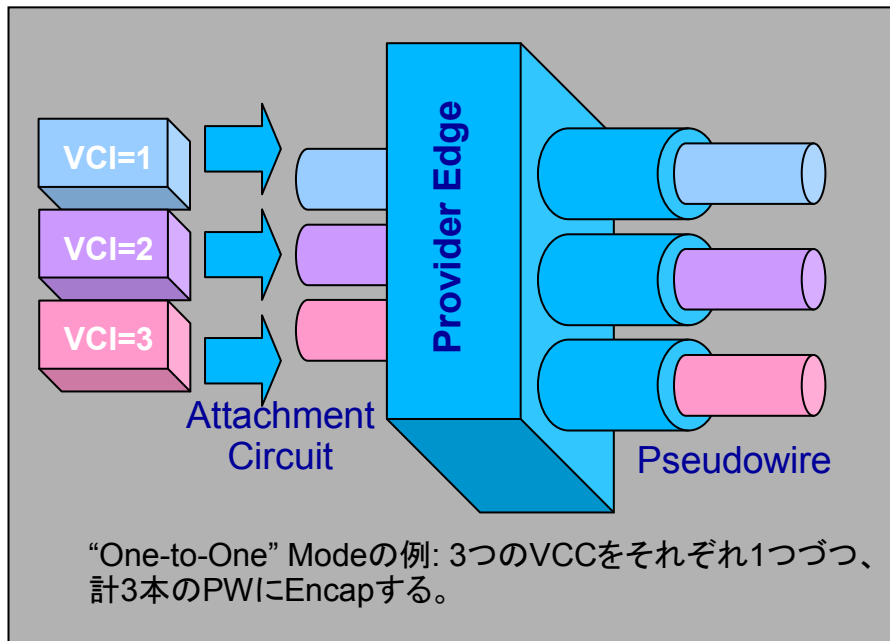
実際のインプリ方針。



ATM Traffic Encapsulation Modes

■ RFC4717

- ✓ “One-to-One” Mode
 - 1つのコネクション(VCC or VPC)を1つのPWに
- ✓ “N-to-One” Mode
 - N本($N \geq 1$)のコネクション(VCC or VPC)を1つのPWに



Modeの選択

- どの方法を選ぶかを決めるとき、気をつけること。今回は、
 1. VC単位で異なるQoS Classが存在してます。
 2. VP単位で頻繁にProvisioningが発生する。
 3. Pseudowireの数は極力少ない方がいい。
- 1.の要件を満たすために、
 - ✓ Attachment Circuit毎に、VCIを参照して適当なExp値をMarkingする。
- 2.と3.を満たすために、
 - ✓ Attachment CircuitのVPI毎に1つのPWを設定する。



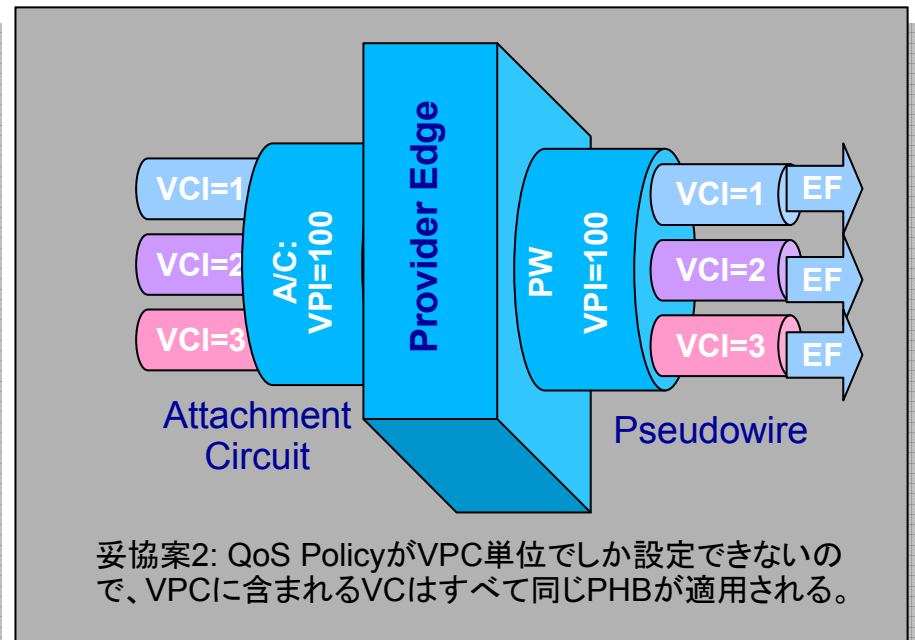
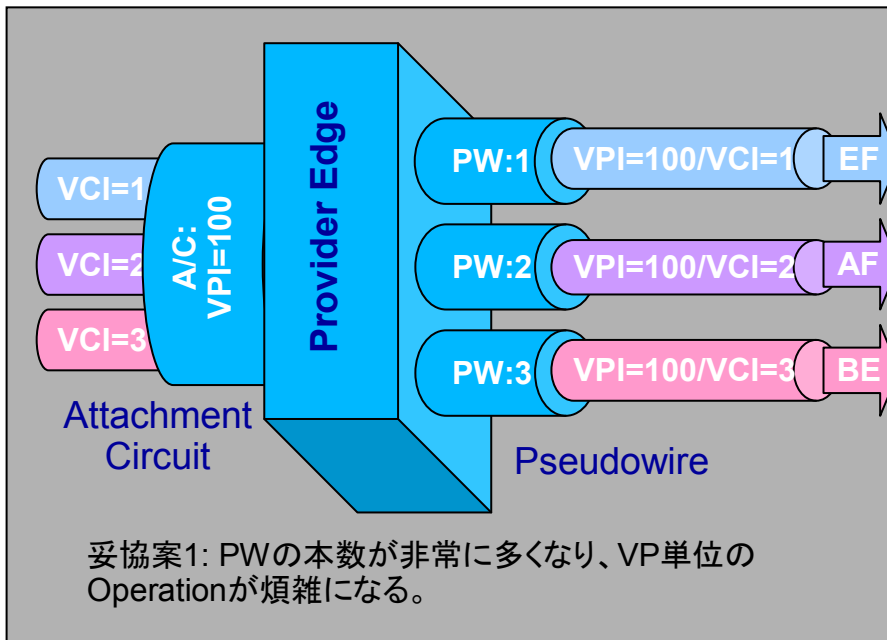
妥協案・・・

■ 妥協案その1

- ✓ “N-to-One (N=1)” Modeの Encapsulation、PWはVCCにする。
- ✓ Exp Markingはちゃんとできるけど、PWの数がVPCのときの十数倍になる。
 - 運用の人に嫌われそう。
 - Provisioningも大変・・・。

■ 妥協案その2

- ✓ “N-to-One (N=1)” Modeの Encapsulation、PWはVPCにする。
- ✓ PWの数は増えないけど、VCI毎の Markingが出来ない。
 - 一番優先度の高いClassのMarkingをすべてのVCに行うことになるので、帯域の無駄使い・・・。全体のTrafficに余裕があるうちはいいけど。。。



どうしてこんなことに・・・。



今時のRouterのしくみ

■ IngressのModuleで、

- ✓ Egress Module/Interface/Labelの決定
- ✓ Classification, Policing
- ✓ Marking
- ✓ Encapsulation
- ✓ Switch Fabricへの転送

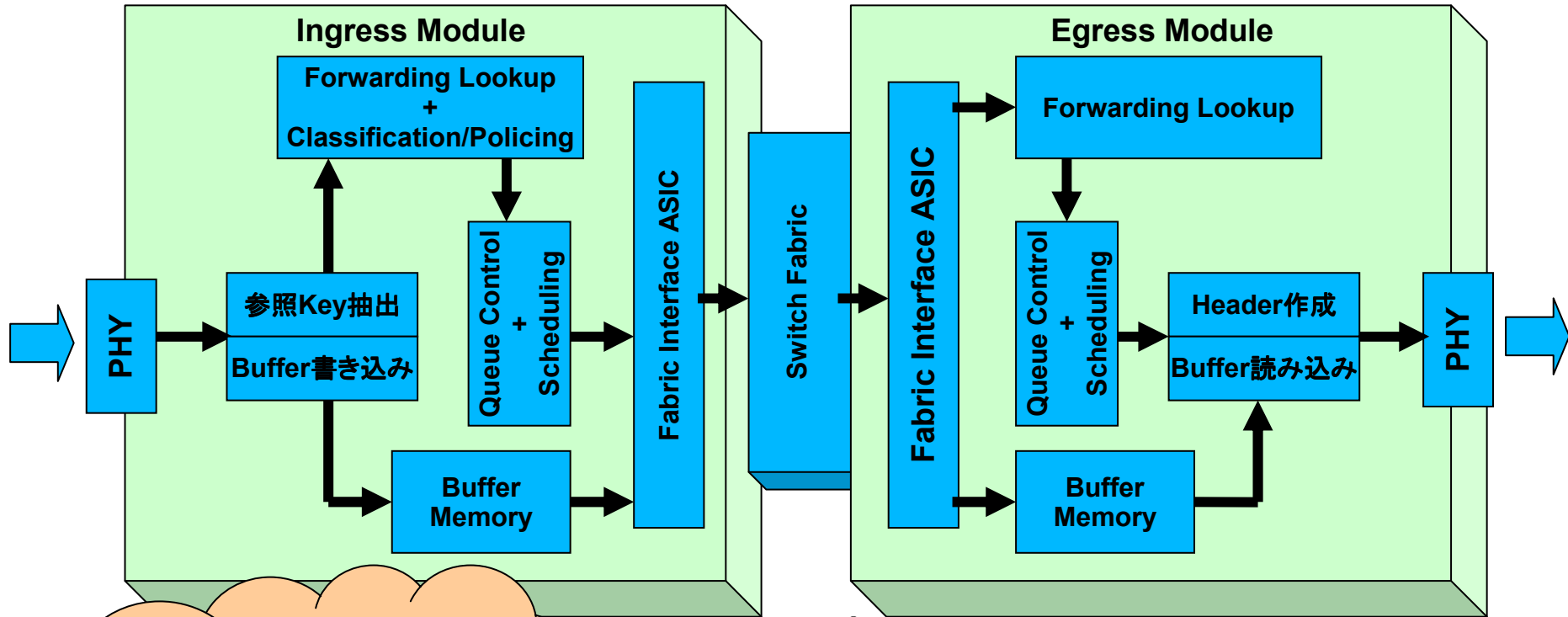
という、一連の流れをやってしまいます。

- ✓ 再帰的にLookupしてClassification/Markingを行ったり、転送先のLookupとClassificationを完全に別々に行うのは不得意・・・。
- ✓ EgressのModuleでは、Ingressで決定されたRuleに従った転送を実施するだけ？！



例えばこんなRouter

■ よくありそうなRouterの転送ロジック

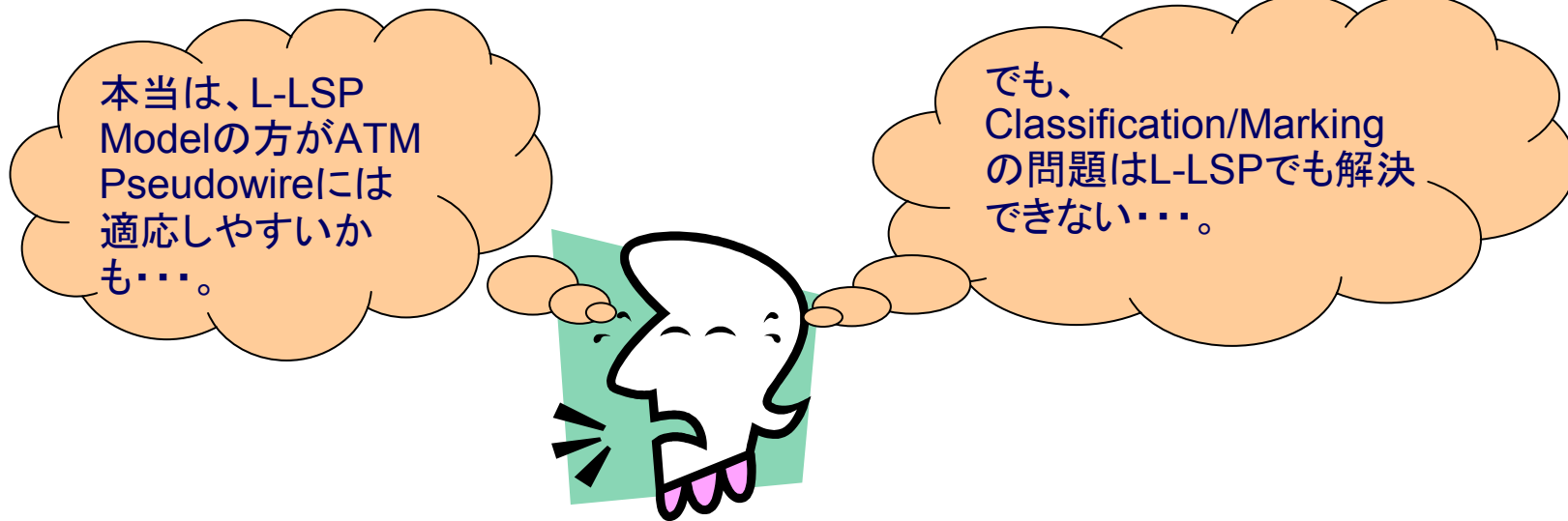


Ingress Moduleで行われるLookup/Classificationは基本的にL3(v4/v6)またはLabel参照で解決されます。従って、PW等の様に特定のMediaに依存したものは予め作成しておいたLabel TableにEntryがないとたいことが出来ないかも・・・。

Egress Moduleでは、IngressのMediaに依存した処理は基本的にできないらしい・・・

そもそもATM QoSって

- IPやMPLSのTrafficは、ばらばらと(Connectionlessで)流れてくるので、それぞれのDatagramにCode Pointを持っています。
 - ✓ 各Hopはそれら一つ一つのCode Pointを参照して振る舞いを決めます。
- ATMは、VPやVCなどの予め設定されたConnection単位にQoS的な振る舞いが予め決められています。
 - ✓ Connectionの識別ができれば、QoS的な振る舞いを決定できる。

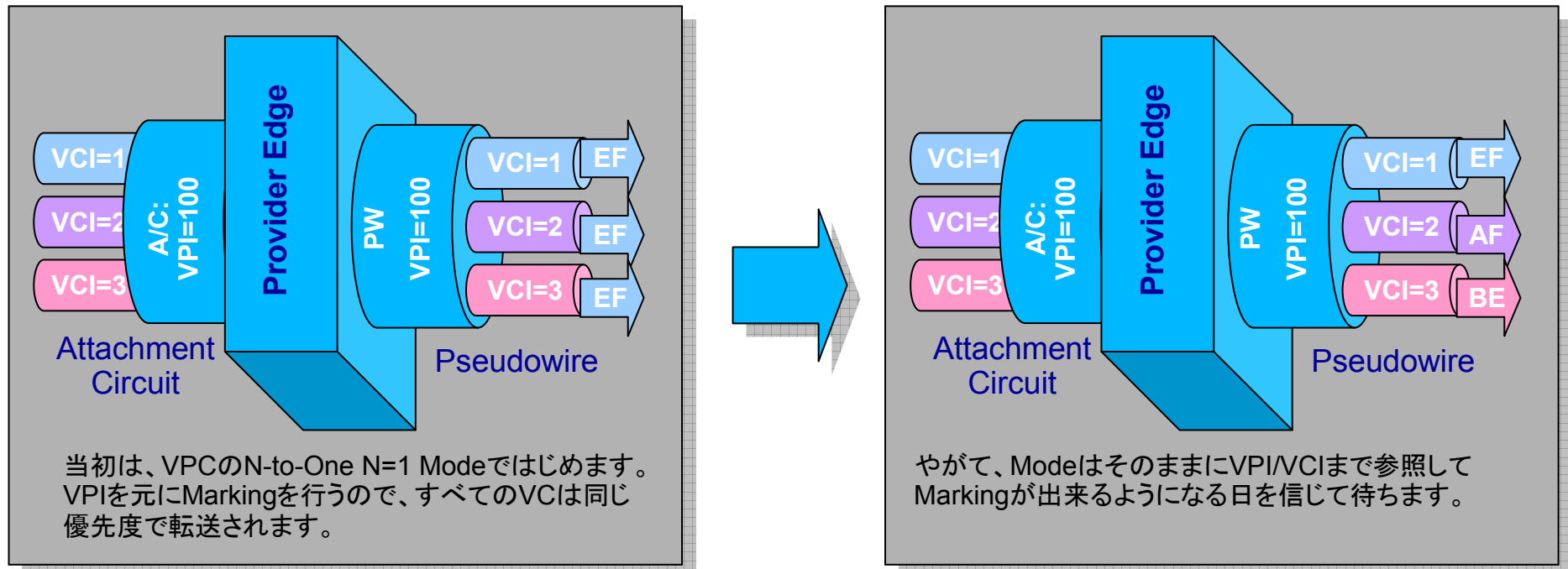


それでもやめるわけにはいかない...。



当面の妥協案と目指す最終ゴール

- まずは、
 - ✓ “妥協案2”で行きます。“N-to-One (N=1)” ModeのEncapsulation、PWはVPCではります。
 - ✓ 当面はキャパシティー的に余裕があるので、PWのTrafficは最優先として扱います。
- そのうち、
 - ✓ Ingress PEでclassificationするときにVCIまでlookupしてもらえようになる日がいつかやってくることを期待します。
 - ✓ この細工は、他の解決法と比較してもまだ実現される見込みが高いような気がしています。
 - ✓ 最初に“妥協案2”で始めておけば、移行も楽なはず。



まとめ

- MPLS Pseudowireを使ってもATM S/Wの機能を手に入れられるわけではありません。
 - ✓ 今あるATM網を完全にのせかえようとする、やがて無理なことを見付けてしまうかも・・・。
 - ✓ QoSだけでも、こんだけネタがありました。
- むしろ、MPLSやPseudowireを前提としたATMの構成に近づけていく方が自然かもしれない。
 - ✓ ATMのQoSに固執なくていい部分(UBRとか?!)だけのせてみるとか。
- まだATMを使っている・使い続けている人は多いけど、Routerにしたって、昔ほどATM関連製品を開発しているわけじゃないみたいですね。
- QoS以外にも、Convergence, OAM, Redundancyなどたくさん難問がありますよ・・・。