

IPFIX 関連の RFC を 読んでみよう

浅間 正和 @ 有限会社 銀座堂

IPFIX とは

- IP Flow Information eXport の略
- ネットワーク・デバイスが IP Traffic Flow 情報をエクスポートする為のプロトコル
- IETF IPFIX WG によって標準化が進められている
- エクスポートする情報を Template として自由に定義することができる
- Cisco 社独自プロトコルの NetFlow v9 がベースとなっている
- 主要な RFC :
 - [RFC5470] Architecture for IP Flow Information Export
 - [RFC5101] Specification of the IP Flow Information Export (IPFIX) Protocol for the Exchange of IP Traffic Flow Information
 - [RFC5102] Information Model for IP Flow Information Export
 - [RFC5472] IP Flow Information Export (IPFIX) Applicability
- ※ 他にも、匿名性を確保するための RFC6235、 双方向 Flow に対応するための RFC5103、大規模なネットワークで Flow 情報を集約するための RFC6183 等
もりだくさん
- もしかしてまだそんなに実装されていない...？

IP Traffic Flow（あるいは単に Flow）とは

- ネットワーク内の Observation Point（後述）を一定時間間隔内に通過するパケットの集合
- 特定の Flow に属するすべてのパケットは共通のプロパティの集合を持つ
- 各々のプロパティは以下の値に対してある関数を適用した結果として定義する
 1. ひとつ以上のパケット・ヘッダ・フィールド（例：宛先 IP アドレス）、トランスポート・ヘッダ・フィールド（例：宛先ポート番号）、アプリケーション・ヘッダ・フィールド（例：RTP ヘッダ・フィールド）
 2. ひとつ以上のパケットそのものの特徴（例：MPLS ラベルの数）
 3. パケットの取り扱いから得られたひとつ以上のフィールド（例：Next Hop IP アドレス、送信インターフェース）
- あるパケットが、ある Flow に定義された全てのプロパティを完全に満たしているとき、そのパケットはその Flow に属していると定義される

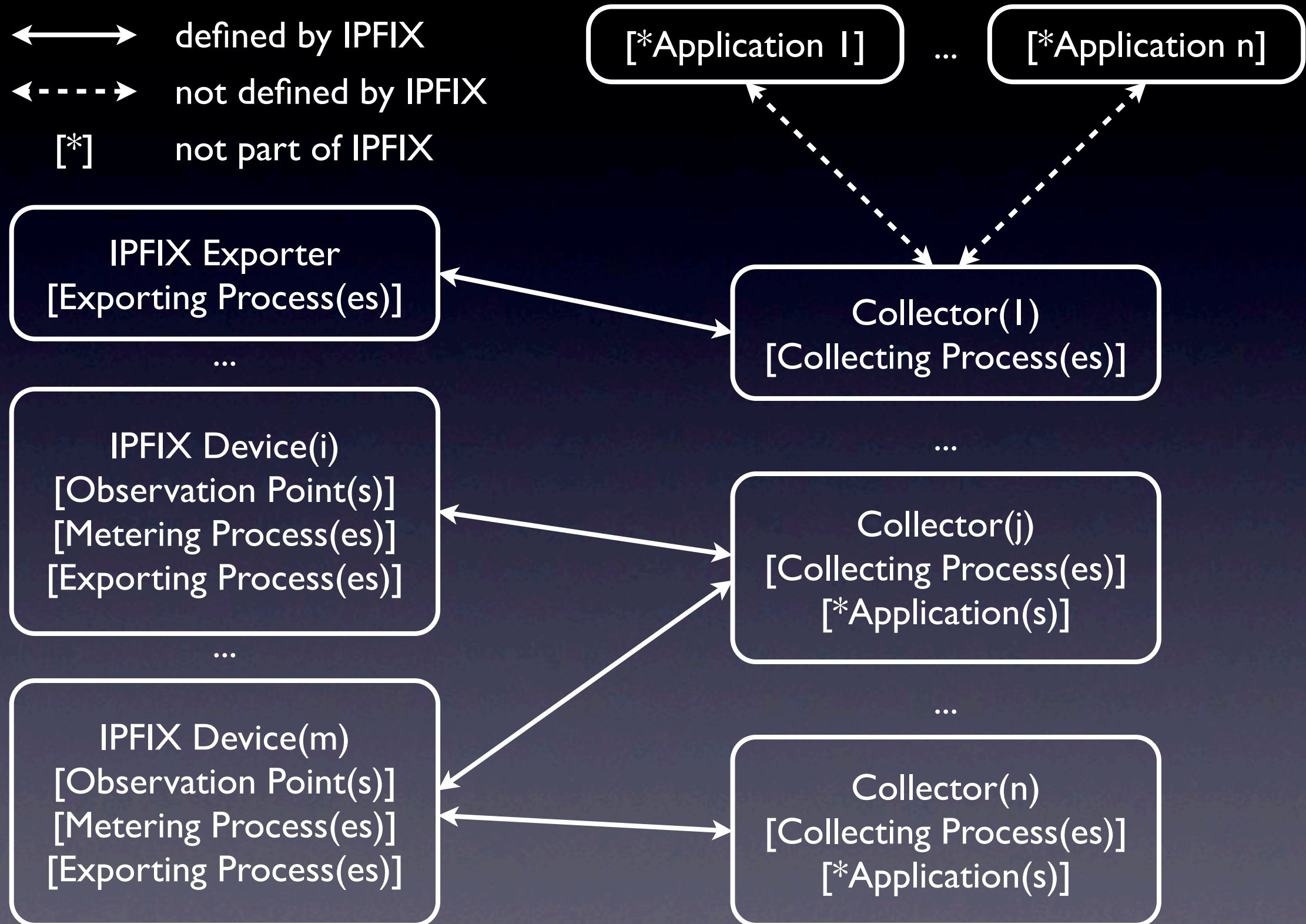
登場人物（１）

- Observation Point（観測点）
 - ネットワーク上で IP パケットを観測できる場所（例：共有メディア上にアタッチされたプローブ、ルータのあるひとつのポート、ルータのインターフェースのセット）
 - すべての Observation Point はひとつの Observation Domain に関連づけられる
 - Observation Point は他の Observation Point のスーパーセットの場合がある
- Observation Domain（観測領域）
 - Observation Point の集合
 - 各 Observation Point での Flow 情報は Observation Domain の単位で Metering Process によって集約される
- Metering Process
 - Observation Point で観測されたパケットのヘッダと特徴から Flow レコードを生成するプロセス
 - パケット・ヘッダのキャプチャ、タイムスタンプ生成、サンプリング、分類、Flow レコードの分類といった機能からなる

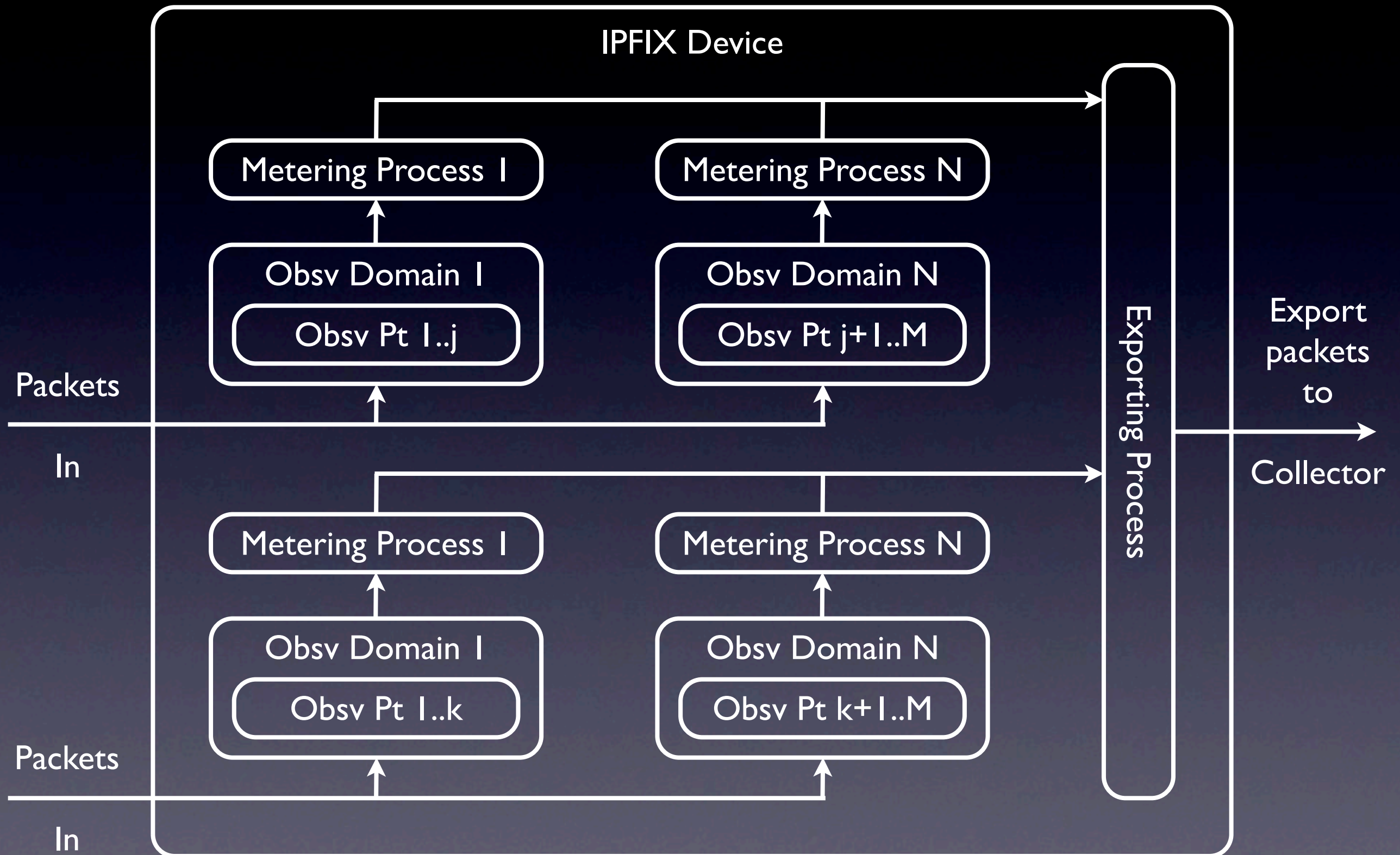
登場人物（2）

- Exporting Process
 - ひとつ以上の Metering Process によって生成された Flow レコードをひとつ以上の Collecting Process に送信するプロセス
- Exporter
 - ひとつ以上の Exporting Process を実行しているデバイス
- IPFIX Device
 - ひとつ以上の Exporting Process を実行しているデバイス
 - Exporting Process のほかに任意の数の Observation Point と Metering Process が実行されているかもしれない
- Collecting Process
 - ひとつ以上の Exporting Process から Flow レコードを受信するプロセス
 - Collecting Process は受信した Flow レコードを処理したり保存したりするかもしれない
- Collector
 - ひとつ以上の Collecting Process を実行しているデバイス

IPFIX リファレンス・モデル



IPFIX デバイス



Metering Process

- IPFIX デバイス内の Flow 測定にかかわるすべての Observation Point は少なくともひとつの Metering Process と関連づけられていなければならない
- Observation Point に届いたすべてのパケットはその Observation Point に関連づけられたそれぞれの Metering Process に届けられる
- 大雑把には Metering Process は Observation Point を通過するパケットを観察し、タイムスタンプを押し、選択基準(Selection Criteria)に基づき Flow に分類する
- Metering Process は Observation によって生成されたすべての Flow を管理する機能ブロックで、典型的には以下のものが含まれているかもしれない
 - Observation Domain からきた全ての Flow Records のデータベースを管理する
これには新しい Flow Records の作成、既に存在する Flow Records の更新、Flow Records 統計情報の計算、追加の Flow プロパティの伝達、Flow 特有ではないパケットの取り扱いに基づく情報（いくつかのケースでは AS 番号やルータの状態等）の追加が含まれる
 - 生成された Flow Records、観測されたパケット等、といった Metering Process 自身の統計情報の管理

Flow Expiration と Flow Export

- Flow は以下の条件のもとで期限が切れた (Expired) と考えられる
 1. Flow に属するパケットが一定期間観測されなかった場合
この“一定時間”は Metering Process において設定が可能なはずで、その値が最も小さい値である 0 秒の場合、直ちに期限が切れることを意味する
その場合ひとつのパケットをひとつの Flow としてレポートする
 2. IPFIX デバイスがリソース上の制約を受けたとき
例えば Flow Records を保持するメモリが不足した場合、本来よりも速く期限が切れるかもしれない
 3. 長い間走っている Flow は Metering Process が通常の基準があるエクスポート・ポリシーに基づき期限が失効する；これは Metering Process で設定可能
期限失効時、その Flow Record はそのまま管理されるかもしれない
- Exporting Process は期限が切れたフローをエクスポートしようと判断する
- 長い間存在している Flow に対して、Exporting Process は Flow Records を通常の基準か、あるエクスポート・ポリシーに基づきエクスポートする；この周期性やエクスポート・ポリシーは Exporting Process において設定することが可能

Selection Criteria for Packets

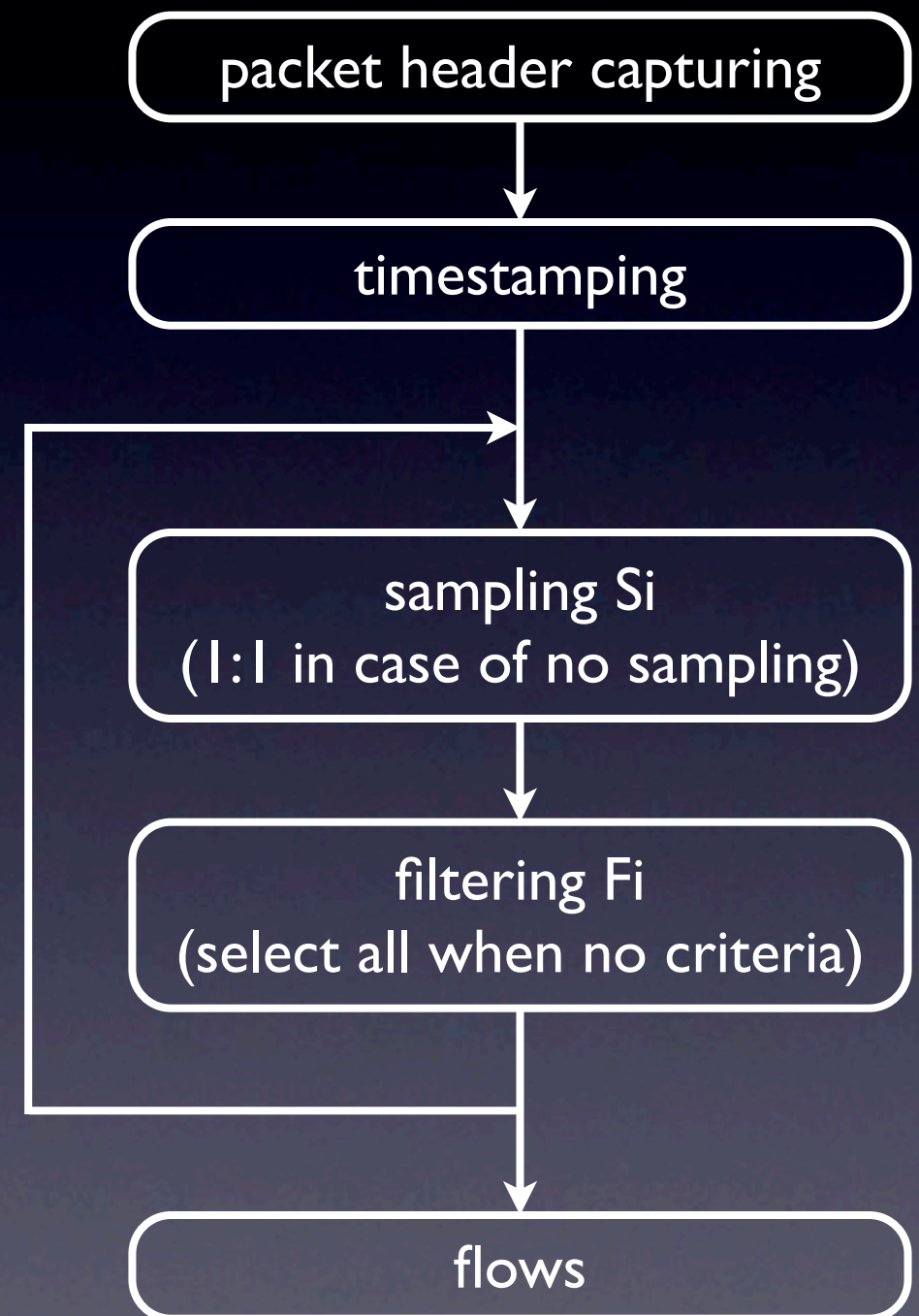
- Metering Process は Observation Point での測定で特定の packets のみが選択されるようなルールを定義するかもしれない
- それは以下の 2 つの方法のひとつか、それらを順番に組み合わせたものかもしれない

1. Sampling Functions, S_i

連続して入ってくる packets のうち測定のために選択するかどうかを判断するための関数（例：毎度 100 番目の受信 packets をサンプルする）

2. Filter Functions, F_i

packet・ヘッダの値や packet 処理で得られた値、packet そのもののプロパティで定義される関数を満たす packet のみを選択する関数



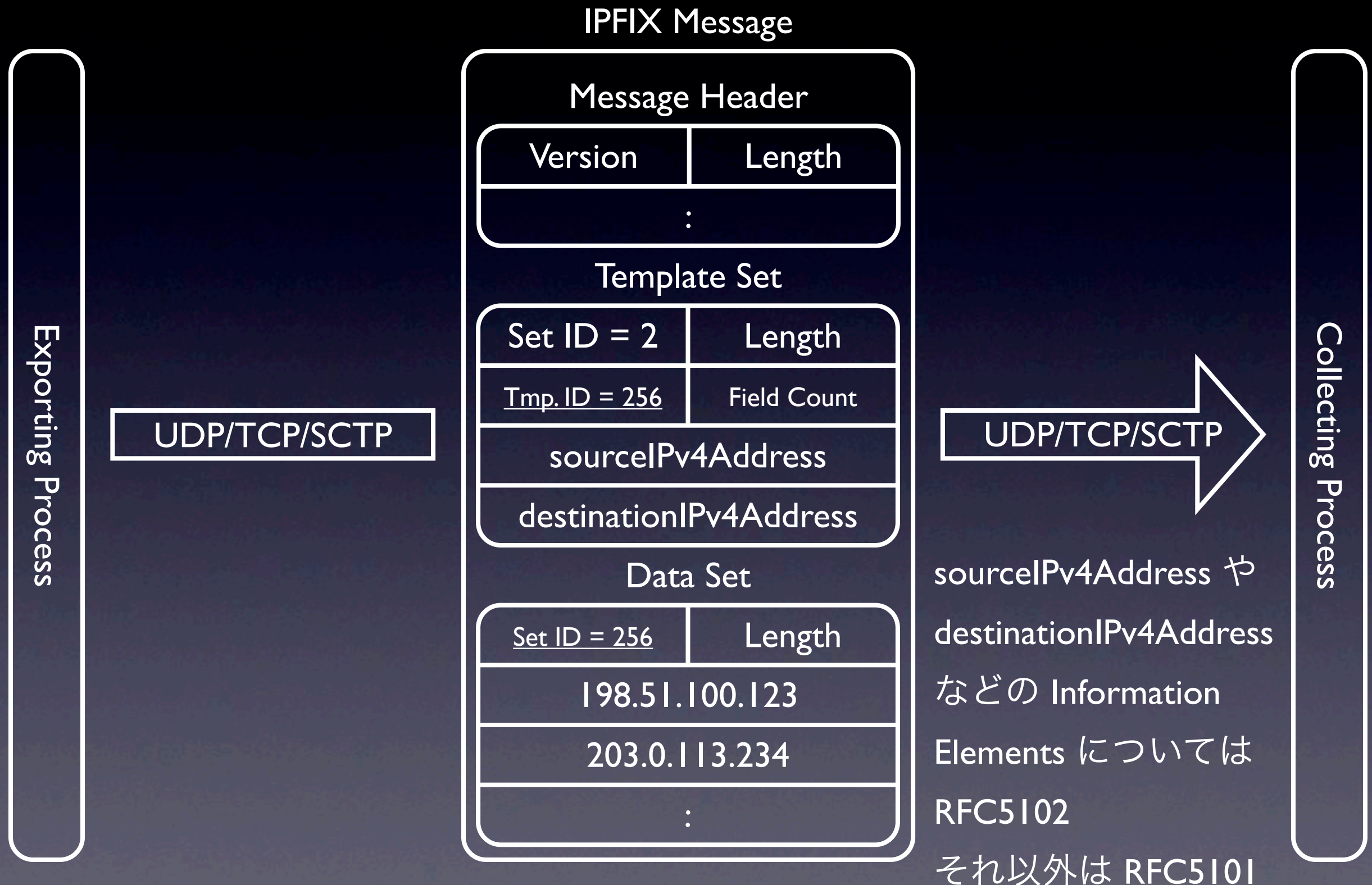
Observation Point と Observation Domain

- Flow Record は、それがどの Observation Point で測定されたのかがわかることで良く分析することができるので、IPFIX デバイスはその情報を Collectors に送信することが推奨される
- Observation Point がひとつしかない場合や Observation Point のある場所に関する情報が関係ないような場合は、Metering Process は Observation Point に関する情報を Flow Records に追加しないかもしれない
- Observation Domain は IPFIX デバイスのなかで単一の Observation Points グループを識別するための論理的なブロックを表す
- 全ての {Observation Point, Metering Process} の組は単一の Observation Domain に属する
- IPFIX デバイスは複数の Observation Domain を持つことができる
- すべての Observation Domain は IPFIX デバイスのコンテキストのなかで一意的な ID (Observation Domain ID) を持っていなければならない
- 複数の Observation Domain を持つ場合一意的な ID は Exporting 機能にパラメータとして渡されなければならない

Exporting Process と Collecting Process

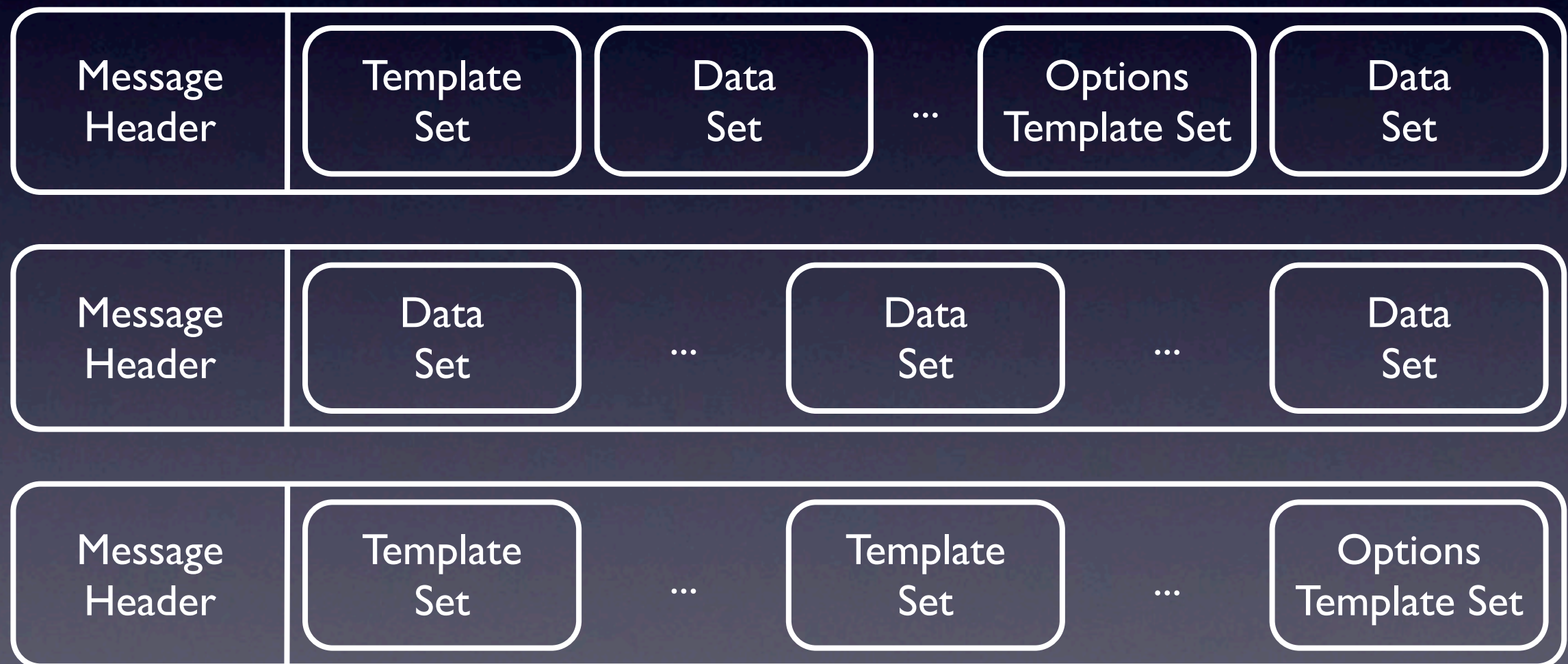
- Exporting Process は IPFIX プロトコルを用いてひとつ以上の IPFIX Collectors にデータを送信する機能ブロック
- Exporting Process は、一方で Metering Process から Flow Records を取得する境界となり、他方で Collector の Collecting Process とやり取りをする
- もしかしたら Selection Criteria for Packets で述べたように特定の Flow Records だけがエクスポートされるような追加のルールが Observation Domain で定義されているかもしれない
- Collecting Process は Flow Record の Template ID を Flow Record の Information Elements を解釈する際に用いる
- そのためには IPFIX Exporter は IPFIX Collector がすべての Flow Record の Template ID を知っていることを保証しなければならない
- また Flow Records を解釈するために IPFIX Collector は Metering Process で各々の Flow に用いられた関数 $F()$ も知っている必要がある
- Collecting Process は異なる <Exporting Process, Observation Domain> の組からの IPFIX メッセージを識別しそれらを保存しなければならない

RFC5101 と RFC5102



IPFIX Message Format

- IPFIX Message は Message Header とそれに続くひとつ以上の Sets によって構成される
- Sets は Data Set、Template Set、Options Template Set のいずれかの型のいずれか
- Byte Order は Network-Byte order (Big-Endian)



Message Header Format

- Version (2 octets)
 - NetFlow v9 の次ということで 10 (0x000a) を入れる
- Length (2 octets)
 - Message Header と Sets を含んだトータルの長さ (単位はオクテット)
- Export Time (4 octets)
 - Exporter から送信された時刻の 1970年1月1日 0:00 UTC から経過した秒数
- Sequence Number (4 octets)
 - PR-SCTP ストリームで送信されたすべての IPFIX Data Records の数を 2^{32} で割った余り
 - 使用するトランスポート層によって意味が変わる
- Observation Domain ID (4 octets)
 - Exporting Process に対してローカルにユニークな Observation Domain を識別する 32 bit の識別子
 - IPFIX Message 全体が特定の Observation Domain に関する情報ではない場合 (例えば Exporting Process の統計情報の場合) は 0 となる場合がある

Field Specifier Format

- E (1 bit : Enterprise bit)
 - この bit が 0 の場合は IETF によって仕様が定められる Information Element
 - この bit が 1 の場合はベンダーが独自に仕様を定める Information Element
- Information Element ident. (7 bit)
 - Information Element を識別する識別子
 - E bit が 0 の場合は RFC5102 を参照
 - E bit が 1 の場合はベンダーの仕様書を参照
- Field Length (2 octets)
 - エンコードされる Information Element の長さ (単位はオクテット)
- Enterprise Number (4 octets : Optional)
 - E bit が 1 のときのみ存在するフィールドで E bit が 0 のときは存在しない
 - IANA によって割り当てられた Enterprise Number [PEN] が埋まる
<http://www.iana.org/assignments/enterprise-numbers>

Set and Set Header Format

- Set Header
 - Set ID (2 octet)
 - Set を識別するための ID
 - 2 は Template Set のために予約
 - 3 は Options Template Set のために予約
 - 256~ は Data Set のために使われる
 - 4~255 は将来の為の予約
 - 0, 1 は歴史的な理由で使用しない
 - Length (2 octet)
 - Set 全体の長さ (単位はオクテット)
- record
 - 後述の Template Record Format、Options Template Record Format、Data Record Format を参照
- Padding
 - Set が aligned boundary から始まるよう padding が入れられる場合がある



Template Record Format

- Template Record Header
 - Template ID (2 octet)
 - すべての新しく生成される Template Record には一意な Template ID が与えられる
 - その一意性はトランスポート・セッションと Observation Domain において局所的となる
 - Template ID 0~255 は Template Sets と Options Template Sets と他の作成される Sets のために予約されている
 - Data Sets の Template ID は 256~65535 の範囲となる
 - Field Count (2 octet)
 - この Template Record に含まれるフィールドの数

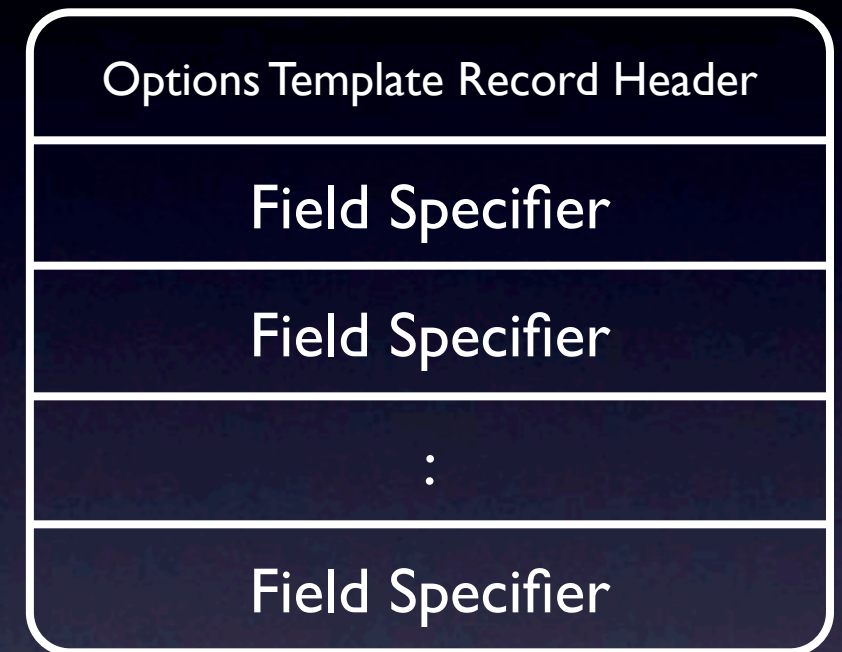
Template Record Format



Options Template Record Format

- Options Template Record Header
 - Template ID (2 octet)
 - この Options Template Record の Template ID
 - 255 より大きい値となる
 - Filed Count (2 octet)
 - この Options Template Record に含まれるすべてのフィールドの数 (Scope Filed 含む)
 - Scope Field Count (2 octet)
 - この Options Template Record に含まれる Scope Fileds の数 (必ず 0 以外となる)
- Scope
 - 対応する Data Records に含まれる Information Elements のコンテキストを指定するもの
 - 少なくとも LineCardId, TemplateId, exporterIPv4Address, exporterIPv6Address, and ingressInterface を Scope として指定できるよう実装されなければならない

Options Template Record Format



Data Record Format

- Field Value
 - (Template ID に応じたデータが埋められる)

Data Record Format

Field Value
Field Value
:
Field Value

いくつかの特定用途 Options Templates (1)

- Metering Process Statistics Option Template
 - Metering Process に関する統計情報をエクスポートするための Options Template
 - もし Exporter Observation Domain 上に複数の Metering Process が存在する場合は追加の Scope Field として meteringProcessId が指定されていなければならない
 - 含むべき IEs : observationDomainId (Scope)、exportedMessageTotalCount、exportedFlowTotalCount、exportedOctetTotalCount
- Metering Process Reliability Statistics Option Template
 - Metering Process の Reliability (信頼性) に関する統計情報をエクスポートするための Options Template
 - もし Exporter Observation Domain 上に複数の Metering Process が存在する場合は追加の Scope Field として meteringProcessId が指定されていなければならない
 - 含むべき IEs : exporterIPv*Address (Scope)、notSentFlowTotalCount、notSentPacketTotalCount、notSentOctetTotalCount、flowStart*、flowEnd*

いくつかの特定用途 Options Templates (2)

- Exporting Process Reliability Statistics Option Template
 - Exporting Process の Reliability (信頼性) に関する統計情報をエクスポートするための Options Template
 - 含むべき IEs : Exporting Process ID (Scope : exporterIPv4Address か exporterIPv6Address か exportingProcessId) 、 notSentFlowTotalCount、 notSentPacketTotalCount、 notSentOctetTotalCount、 time first flow dropped (flowStartMilliseconds か flowStartMicroseconds か flowStartNanoseconds か flowStartDeltaMicroseconds) 、 time last flow dropped (flowEndMilliseconds か flowEndMicroseconds か flowEndNanoseconds か flowEndDeltaMicroseconds)
- Flow Keys Option Template
 - エクスポートされる Flow 情報の Flow Key がなんであるかを指定するための Options Template
 - 含むべき IEs : templateId (Scope) 、 flowKeyIndicator

IETF-Spec. な IE を使った Template Set の例

0		15	16		31
Set ID = 2			Length = 28 octets		
Template ID = 256			Field Count = 5		
0	sourceIPv4Address = 8			Field Length = 4	
0	destinationIPv4Address = 12			Field Length = 4	
0	ipNextHopIPv4Address = 15			Field Length = 4	
0	inPacketDeltaCount = 2			Field Length = 4	
0	inOctetDeltaCount = 1			Field Length = 4	

Enterprise-Spec. IE を使った Template Set の例

0		15		16		31	
Set ID = 2				Length = 32 octets			
Template ID = 257				Field Count = 5			
0	sourceIPv4Address = 8			Field Length = 4			
0	destinationIPv4Address = 12			Field Length = 4			
1	ipNextHopIPv4Address = 15			Field Length = 4			
Enterprise number							
0	inPacketDeltaCount = 2			Field Length = 4			
0	inOctetDeltaCount = 1			Field Length = 4			

Data Set の例

0	15	16	31
Set ID = 256		Length = 44 octets	
192.0.2.12			
192.0.2.254			
192.0.2.1			
5009			
5344385			
192.0.2.27			
192.0.2.23			
192.0.2.2			
748			
388934			

IETF-Spec. IE を使った Options Temp. Set の例

0	15	16	31
Set ID = 3		Length = 24 octets	
Template ID = 259		Field Count = 3	
Scope Field Count = 1		0	lineCardId = 141
Scope 1 Field Length = 4		0	exportedPacketCount = 41
Field Length = 2		0	exportedFlowCount = 42
Field Length = 2		Padding	

IE: Identifiers

ID	Name	ID	Name
141	lineCardId	148	flowId
142	portId	145	templateId
10	ingressInterface	149	observationDomainId
14	egressInterface	138	observationPointId
143	meteringProcessId	137	commonPropertiesId
144	exportingProcessId		

IE: Metering and Exporting Process Config.

ID	Name	ID	Name
130	exporterIPv4Address	213	exportInterface
131	exporterIPv6Address	214	exportProtocolVersion
217	exporterTransportPort	215	exportTransportProtocol
211	collectorIPv4Address	216	collectorTransportPort
212	collectorIPv6Address	173	flowKeyIndicator

IE: Metering and Exporting Process Statistics

ID	Name	ID	Name
41	exportedMessageTotalCount	165	ignoredOctetTotalCount
40	exportedOctetTotalCount	166	notSentFlowTotalCount
42	exportedFlowRecordTotalCount	167	notSentPacketTotalCount
163	observedFlowTotalCount	168	notSentOctetTotalCount
164	ignoredPacketTotalCount		

IE: IP Header Fields

ID	Name	ID	Name
60	ipVersion	193	nextHeaderIPv6
8	sourceIPv4Address	195	ipDiffServCodePoint
27	sourceIPv6Address	196	ipPrecedence
9	sourceIPv4PrefixLength	5	ipClassOfService
29	sourceIPv6PrefixLength	55	postIpClassOfService
44	sourceIPv4Prefix	31	flowLabelIPv6
170	sourceIPv6Prefix	206	isMulticast
12	destinationIPv4Address	54	fragmentIdentification
28	destinationIPv6Address	88	fragmentOffset
13	destinationIPv4PrefixLength	197	fragmentFlags
30	destinationIPv6PrefixLength	189	ipHeaderLength
45	destinationIPv4Prefix	207	ipv4IHL
169	destinationIPv6Prefix	190	totalLengthIPv4
192	ipTTL	224	ipTotalLength
4	protocolIdentifier	191	payloadLengthIPv6

IE:Transport Header Fields

ID	Name	ID	Name
7	sourceTransportPort	238	tcpWindowScale
11	destinationTransportPort	187	tcpUrgentPointer
180	udpSourcePort	188	tcpHeaderLength
181	udpDestinationPort	32	icmpTypeCodeIPv4
205	udpMessageLength	176	icmpTypeIPv4
182	tcpSourcePort	177	icmpCodeIPv4
183	tcpDestinationPort	139	icmpTypeCodeIPv6
184	tcpSequenceNumber	178	icmpTypeIPv6
185	tcpAcknowledgementNumber	179	icmpCodeIPv6
186	tcpWindowSize	33	igmpType

IE: Sub-IP Header Fields

ID	Name	ID	Name
56	sourceMacAddress	201	mplsLabelStackLength
81	postSourceMacAddress	194	mplsPayloadLength
58	vlanId	70	mplsTopLabelStackSection
59	postVlanId	71	mplsLabelStackSection2
80	destinationMacAddress	72	mplsLabelStackSection3
57	postDestinationMacAddress	73	mplsLabelStackSection4
146	wlanChannelId	74	mplsLabelStackSection5
147	wlanSSID	75	mplsLabelStackSection6
200	mplsTopLabelTTL	76	mplsLabelStackSection7
203	mplsTopLabelExp	77	mplsLabelStackSection8
237	postMplsTopLabelExp	78	mplsLabelStackSection9
202	mplsLabelStackDepth	79	mplsLabelStackSection10

IE: Derived Packet Properties

ID	Name	ID	Name
204	ipPayloadLength	18	bgpNextHopIPv4Address
15	ipNextHopIPv4Address	63	bgpNextHopIPv6Address
62	ipNextHopIPv6Address	46	mplsTopLabelType
16	bgpSourceAsNumber	47	mplsTopLabelIPv4Address
17	bgpDestinationAsNumber	140	mplsTopLabelIPv6Address
128	bgpNextAdjacentAsNumber	90	mplsVpnRouteDistinguisher
129	bgpPrevAdjacentAsNumber		

IE: Min/Max Flow Properties

ID	Name	ID	Name
25	minimumIpTotalLength	208	ipv4Options
26	maximumIpTotalLength	64	ipv6ExtensionHeaders
52	minimumTTL	6	tcpControlBits
53	maximumTTL	209	tcpOptions

IE: Flow Timestamps

ID	Name	ID	Name
150	flowStartSeconds	156	flowStartNanoseconds
151	flowEndSeconds	157	flowEndNanoseconds
152	flowStartMilliseconds	158	flowStartDeltaMicroseconds
153	flowEndMilliseconds	159	flowEndDeltaMicroseconds
154	flowStartMicroseconds	160	systemInitTimeMilliseconds
155	flowEndMicroseconds	22	flowStartSysUpTime
		21	flowEndSysUpTime

IE: Per-Flow Counters

ID	Name	ID	Name
1	octetDeltaCount	134	droppedOctetTotalCount
23	postOctetDeltaCount	135	droppedPacketTotalCount
198	octetDeltaSumOfSquares	19	postMCastPacketDeltaCount
85	octetTotalCount	20	postMCastOctetDeltaCount
171	postOctetTotalCount	174	postMCastPacketTotalCount
199	octetTotalSumOfSquares	175	postMCastOctetTotalCount
2	packetDeltaCount	218	tcpSynTotalCount
24	postPacketDeltaCount	219	tcpFinTotalCount
86	packetTotalCount	220	tcpRstTotalCount
172	postPacketTotalCount	221	tcpPshTotalCount
132	droppedOctetDeltaCount	222	tcpAckTotalCount
133	droppedPacketDeltaCount	223	tcpUrgTotalCount

IE: Miscellaneous Flow Properties / Padding

ID	Name	ID	Name
36	flowActiveTimeout	161	flowDurationMilliseconds
37	flowIdleTimeout	162	flowDurationMicroseconds
136	flowEndReason	61	flowDirection
210	paddingOctets		

IPFIX の適用性

- Accounting
 - 使用量に応じた課金
- Traffic Profiling
 - ネットワークの使われ方のトレンド解析・需要予測
- Traffic Engineering
 - ネットワーク資源の利用とトラヒック・パフォーマンスの最適化
- Network Security
 - 攻撃と侵入の検知
- QoS Monitoring
 - サービス品質をパッシブに観測
- Inter-Domain Exchange of IPFIX Data
 - 近隣プロバイダとの情報共有
- Export of Derived Metrics
 - 距離に関する情報のエクスポート