SDN Japan 2013

September 18-20, 2013 @Ebisu Garden Hall

Segment Routing !

- Path/Flow programmability 進化のかたち

18 September 2013

Miya Kohno (mkohno@cisco.com)

Agenda

- 背景 – “Value”, “Simplification”
- Segment Routingの概要
  - 基本動作
  - コントローラとの連携
- Use case
- Segment Routingのメリット

SDN Taxonomy

- SDN

広義のSDN

- Abstraction
- Virtualization
- Cloud Resource Orchestration
- Service control
- Workflow Automation
- …

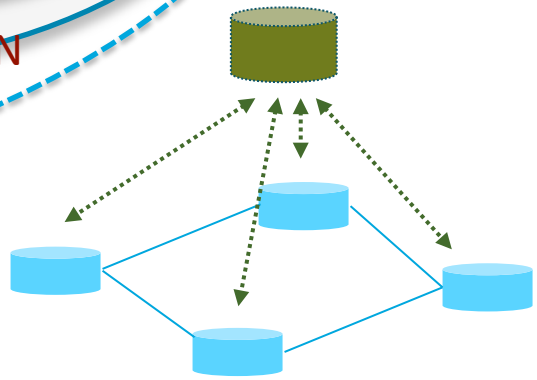
- Flow/Path programmability

狭義のSDN

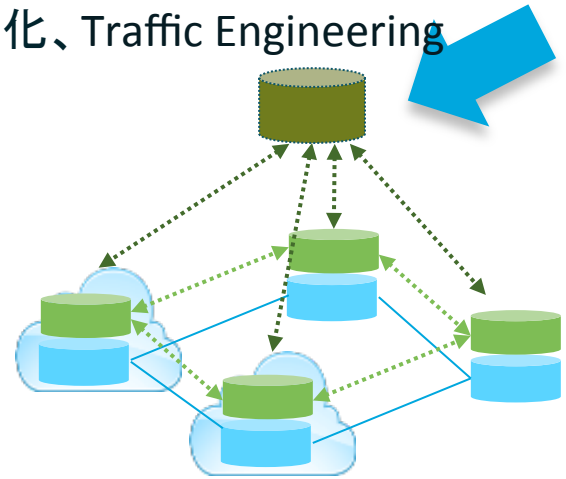
- Flow/Path Programmability

Hybrid Model (vs Controller/Agent Model)

- 自律分散コントロールプレーンは残す
  - Scaling, Robustness
  - Fault Detection, Protection, LAG..
- 集中型intelligenceが適する物は集中で
  - 可視化、分析、
  - 自動化、Traffic Engineering

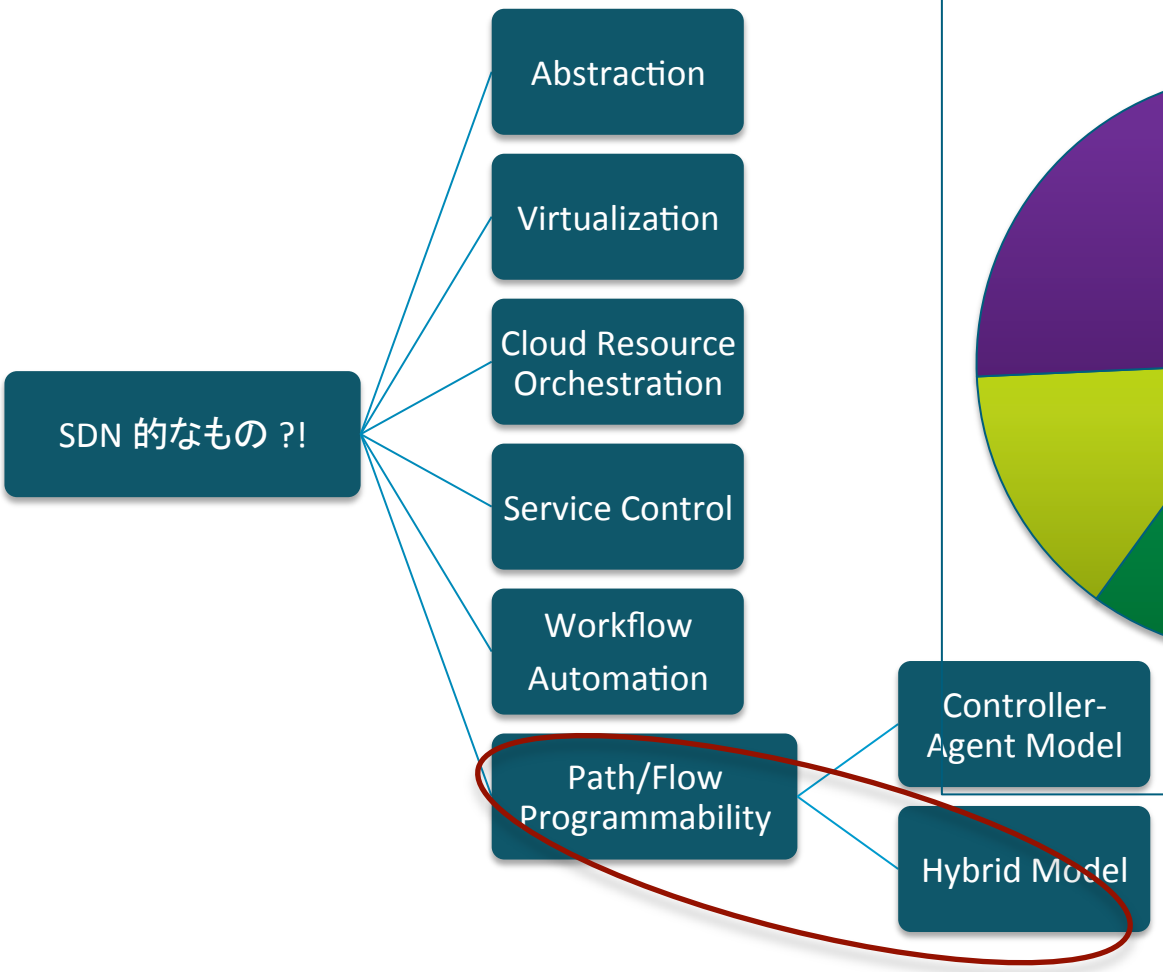


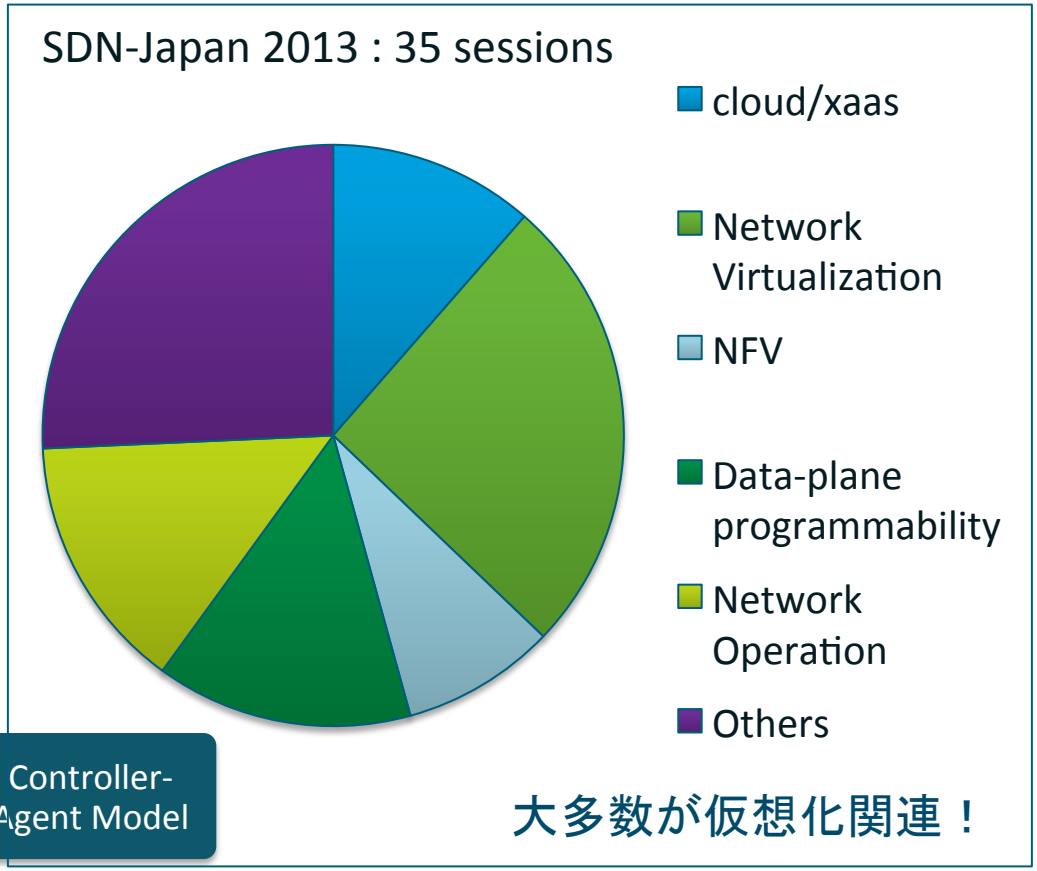
Controller/Agent Model



Hybrid Model

SDN Taxonomy ?!

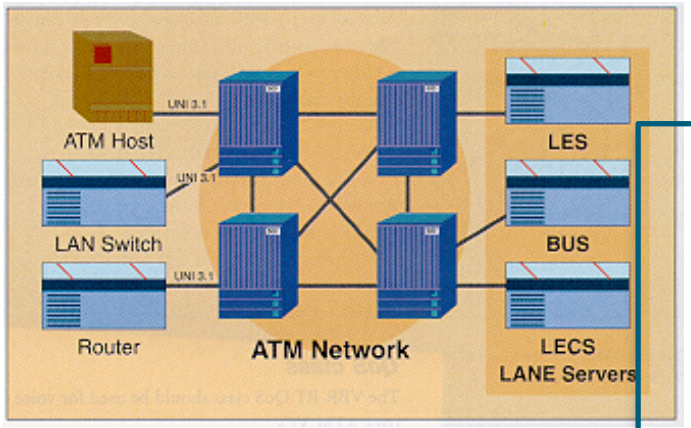




- “Virtualization Japan”に改名する?（冗談です）
- 「インフラのしくみそのものよりも、その上でどういう価値を出すか」が重要

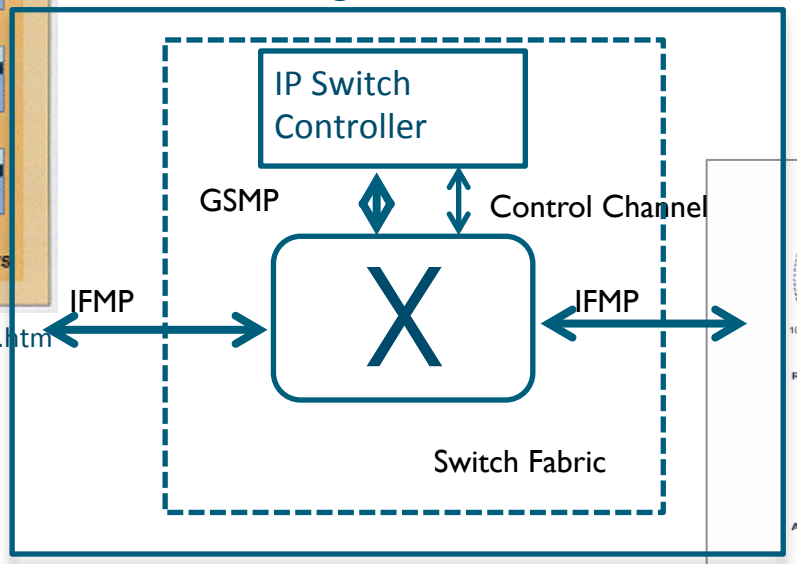
Flow/Path Programmability – lessons learnt

- LAN Emulation (LANE)



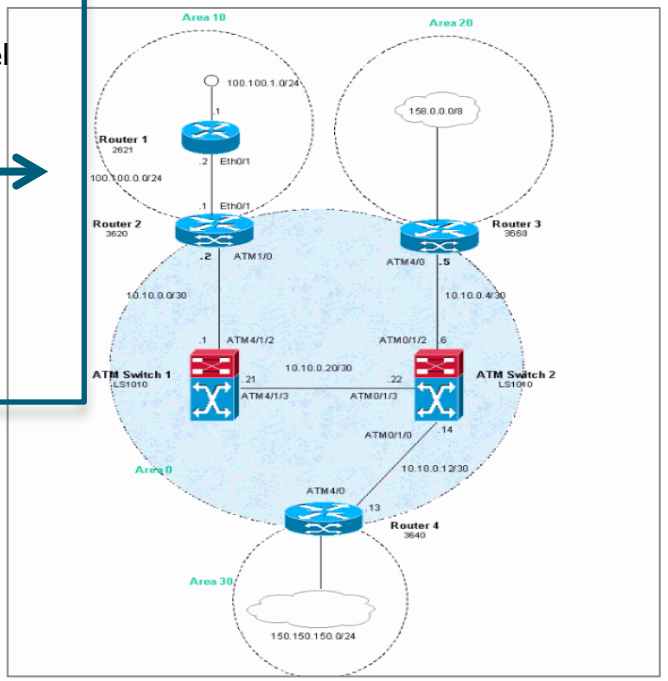
http://home.mira.net/~marcop/ATM_applications.htm

- IP Switching



http://www.cse.wustl.edu/~jain/cis788-97/ftp/ip_switching/index.htm を参考に作成

- MPLS - IP+ATM/VC merge



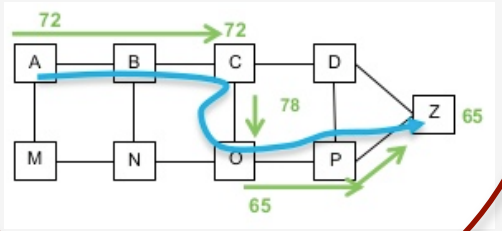
http://www.cisco.com/en/US/tech/tk436/tk798/technologies_configuration_example09186a00801c2d73.shtml

Innovative,
but just emulation is NOT good !!
既存機能のエミュレーションではだめ。

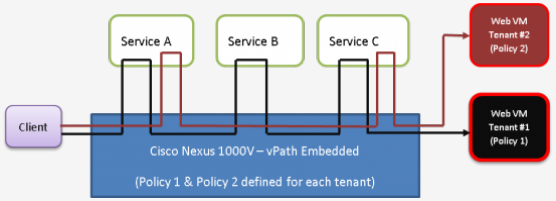
Flow/Path Programmability 進化のかたち

- Flow/Path それ自体が価値を提供することが求められる
- 「Applicationの要請にどれだけ答えられるか」

Segment Routing



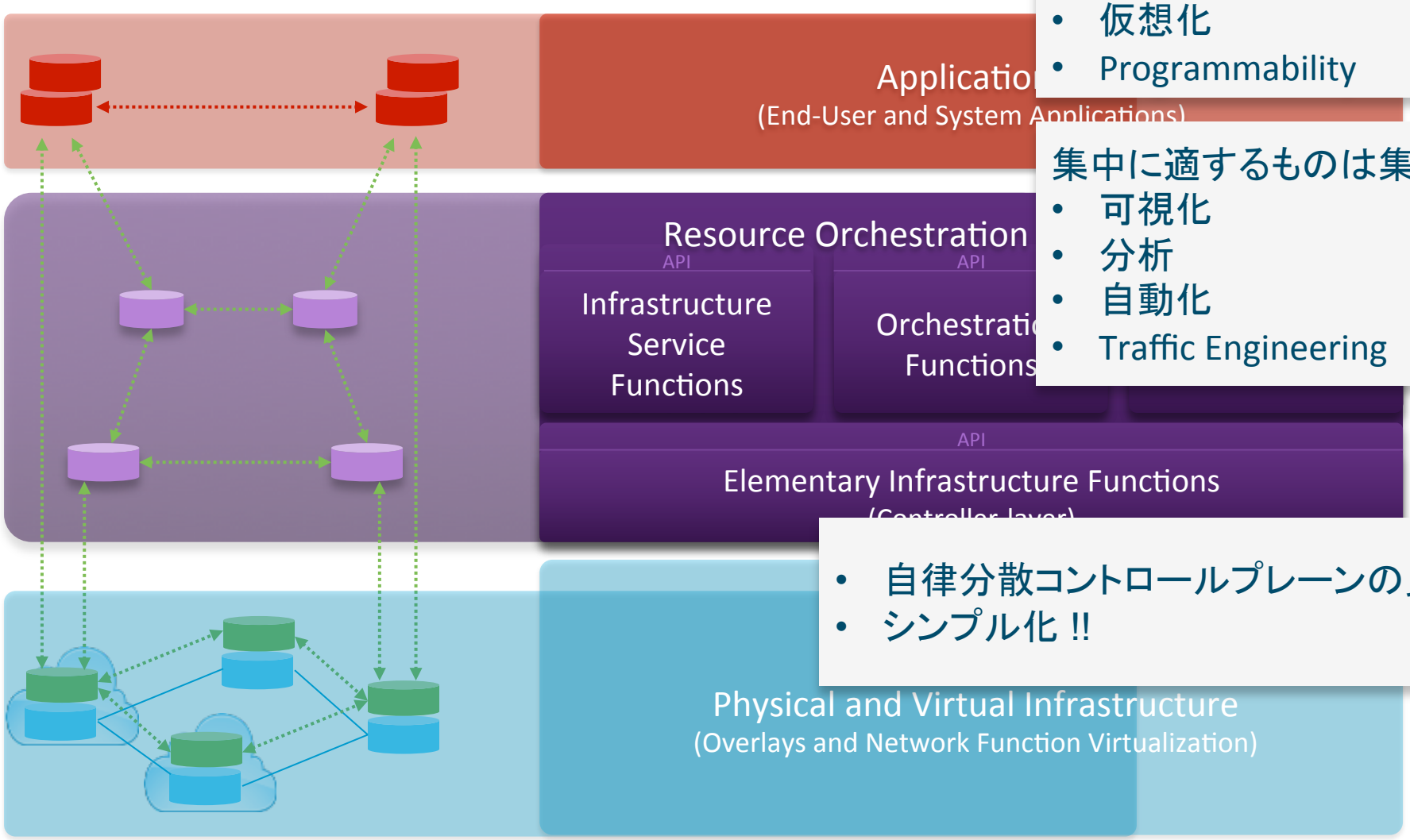
Network Service Chaining



Policy-aware forwarding

仮想化も重要！、ハードウェアも重要 !!

Flow/Path Programmability 進化のかたち
- Segment Routing 開発の背景



- 抽象化
- 仮想化
- Programmability

集中に適するものは集中

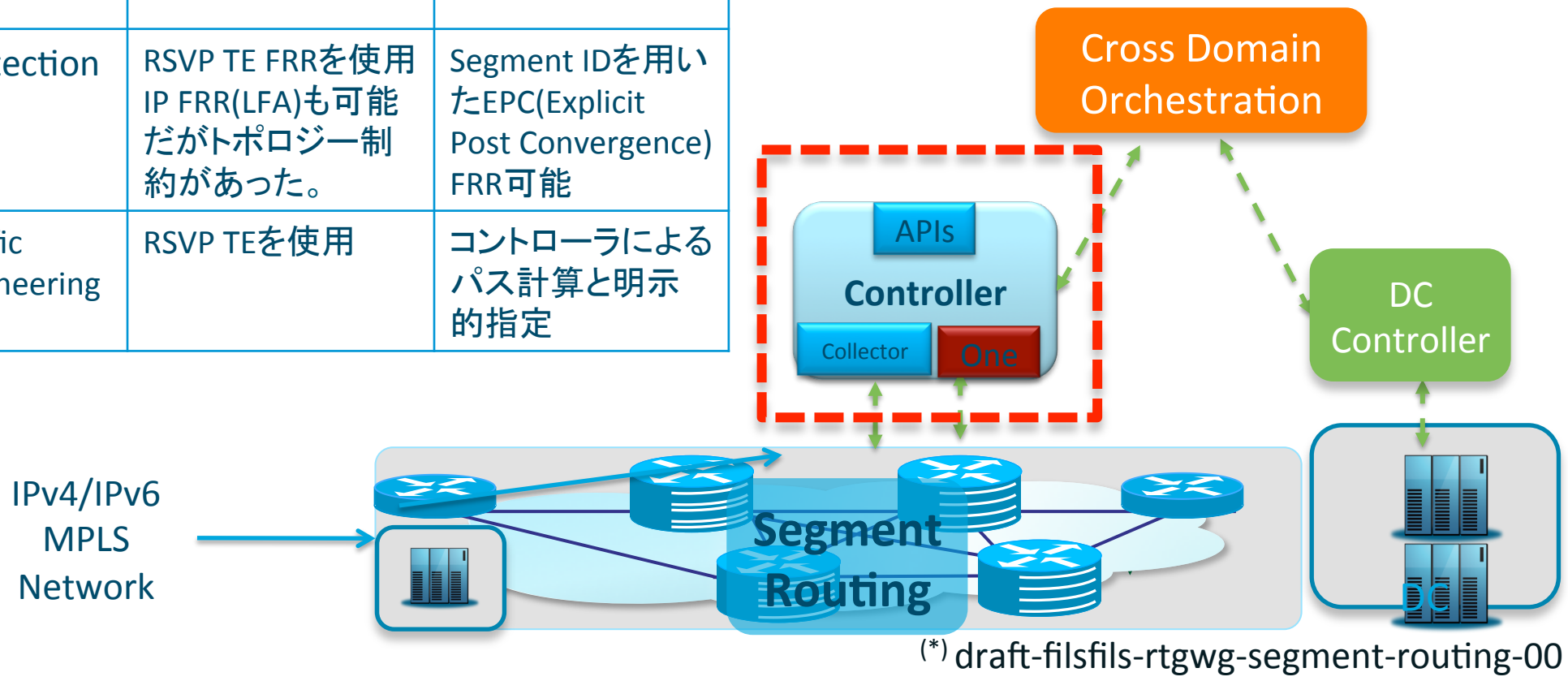
- 可視化
- 分析
- 自動化
- Traffic Engineering

- 自律分散コントロールプレーンの見直し
- シンプル化 !!

Segment Routing(*)によるNWのシンプル化

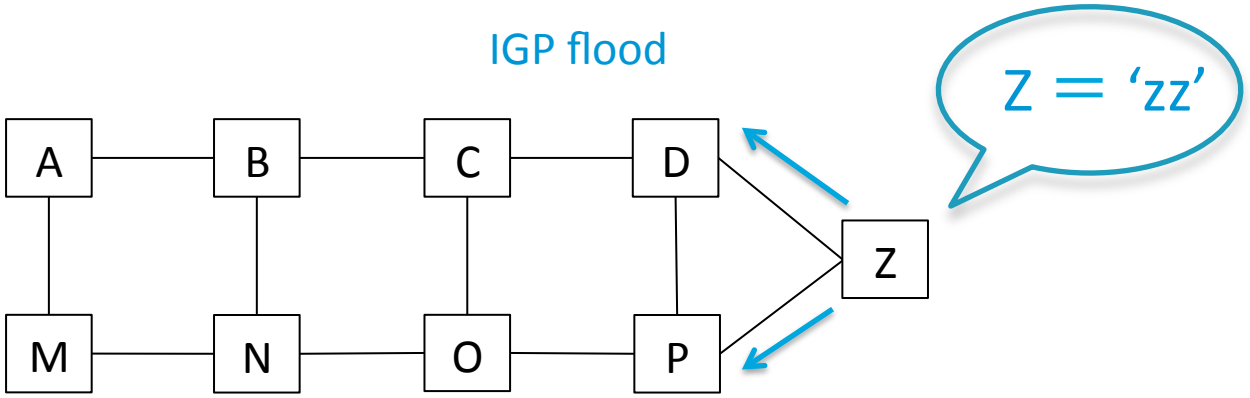
	MPLS	Segment Routing
転送ヘッダ	LDPやRSVPにより Labelを配布	IGPにより Segment IDを配布
Protection	RSVP TE FRRを使用 IP FRR(LFA)も可能 だがトポロジー制 約があった。	Segment IDを用い たEPC(Explicit Post Convergence) FRR可能
Traffic Engineering	RSVP TEを使用	コントローラによる パス計算と明示 的指定

- RSVP，LDPは不要
- ネットワーク内のステートが大幅に減少
- ApplicationやOrchestrationと容易に連携



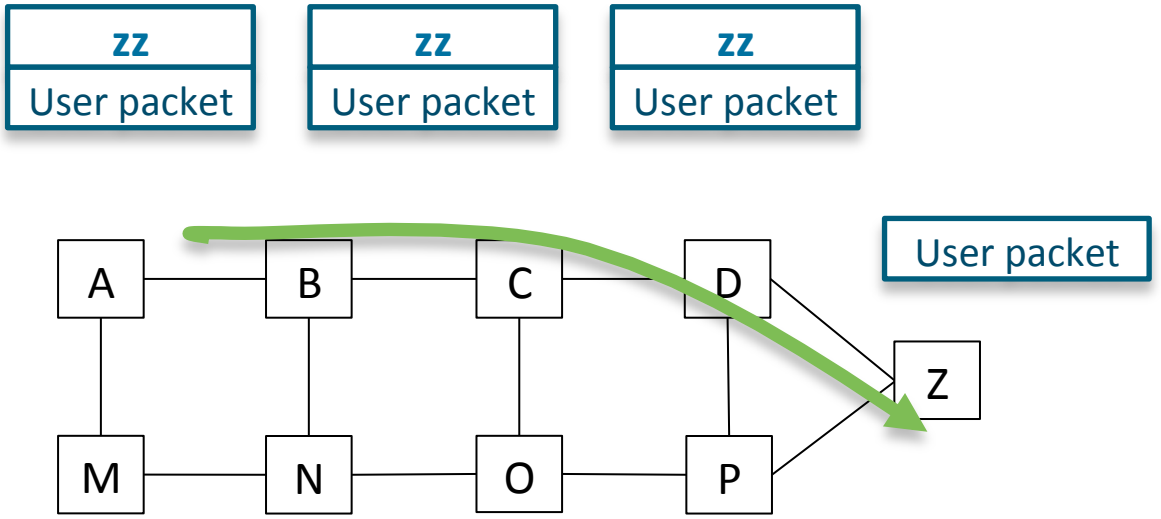
(*) draft-filsfils-rtgwg-segment-routing-00

Segment Routing – 基本動作



- 各ノードは自分のSegment IDを、IGPにより広報する
  - OSPFの場合、opaque LSAを使用 (draft-psenak-ospf-segment-routing-extensions)
  - 2種類のSegment ID
    - Node Segment ID (Node-SID) : ノードへのbest pathを表す
    - Adjacency Segment ID (Ajd-SID) : インタフェースへのone hop path

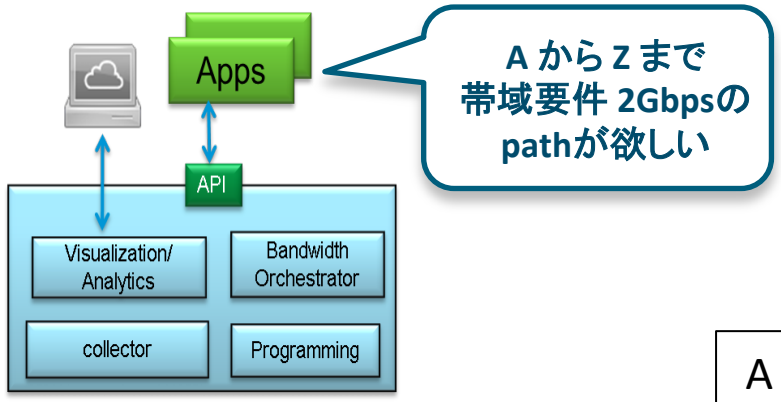
Segment Routing – 基本動作



- ノードZを宛先とするパケットには、zzをつけて転送する

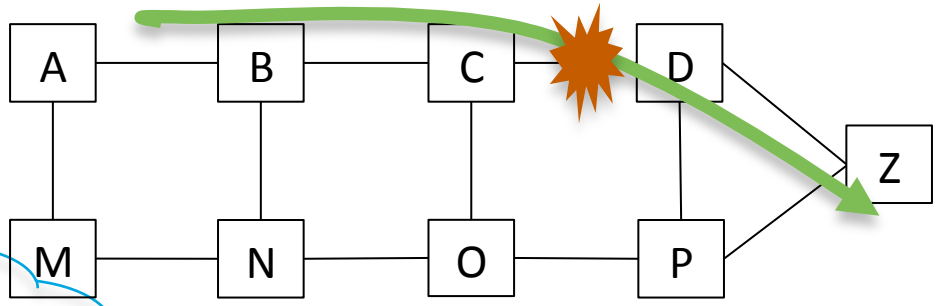
(Segment ID=global significant <-> c.f. MPLS Label = local significant)

Segment Routing – Controllerによる制御



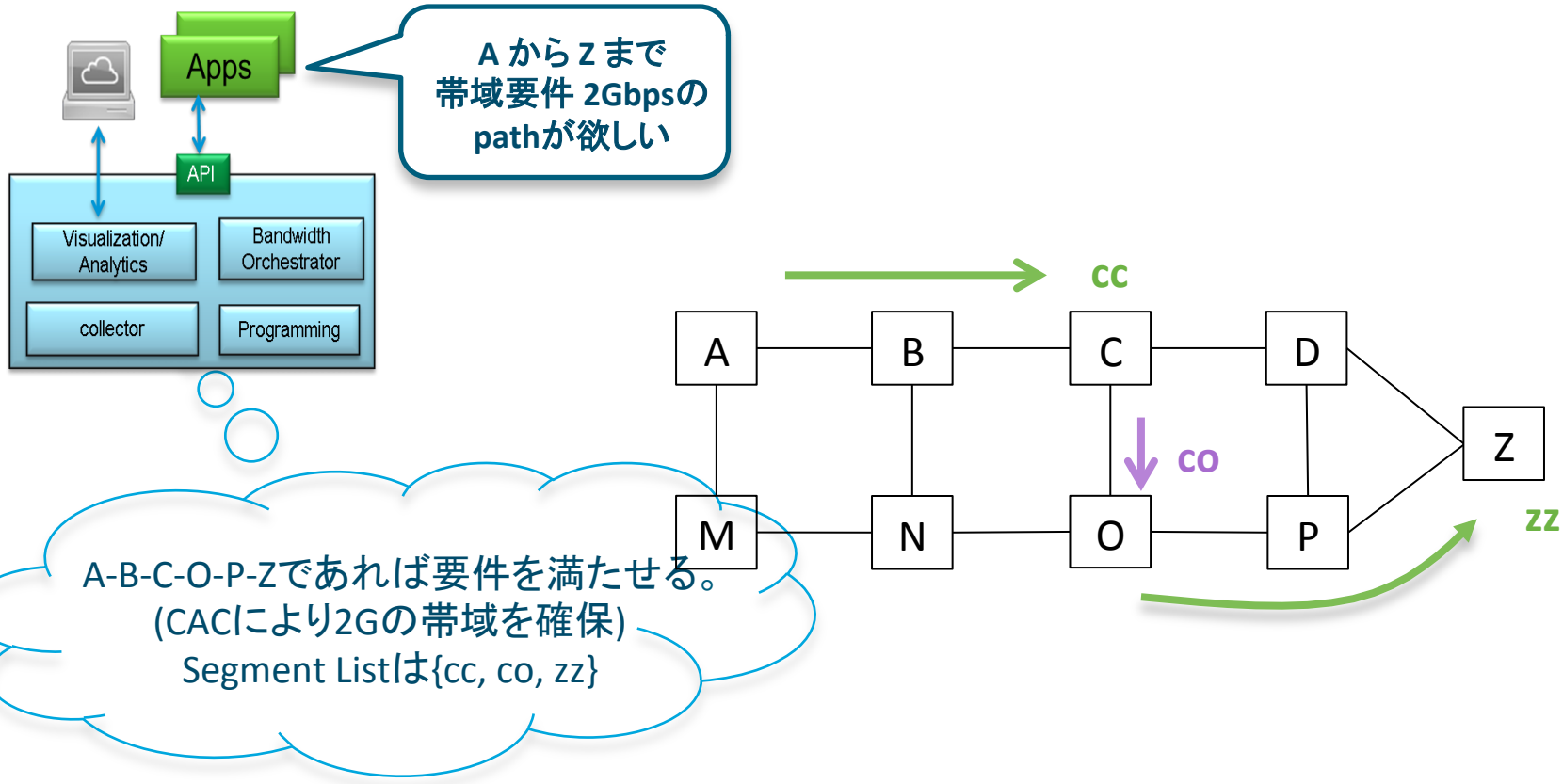
A から Z まで
帯域要件 2Gbpsの
pathが欲しい

Bandwidth = Full



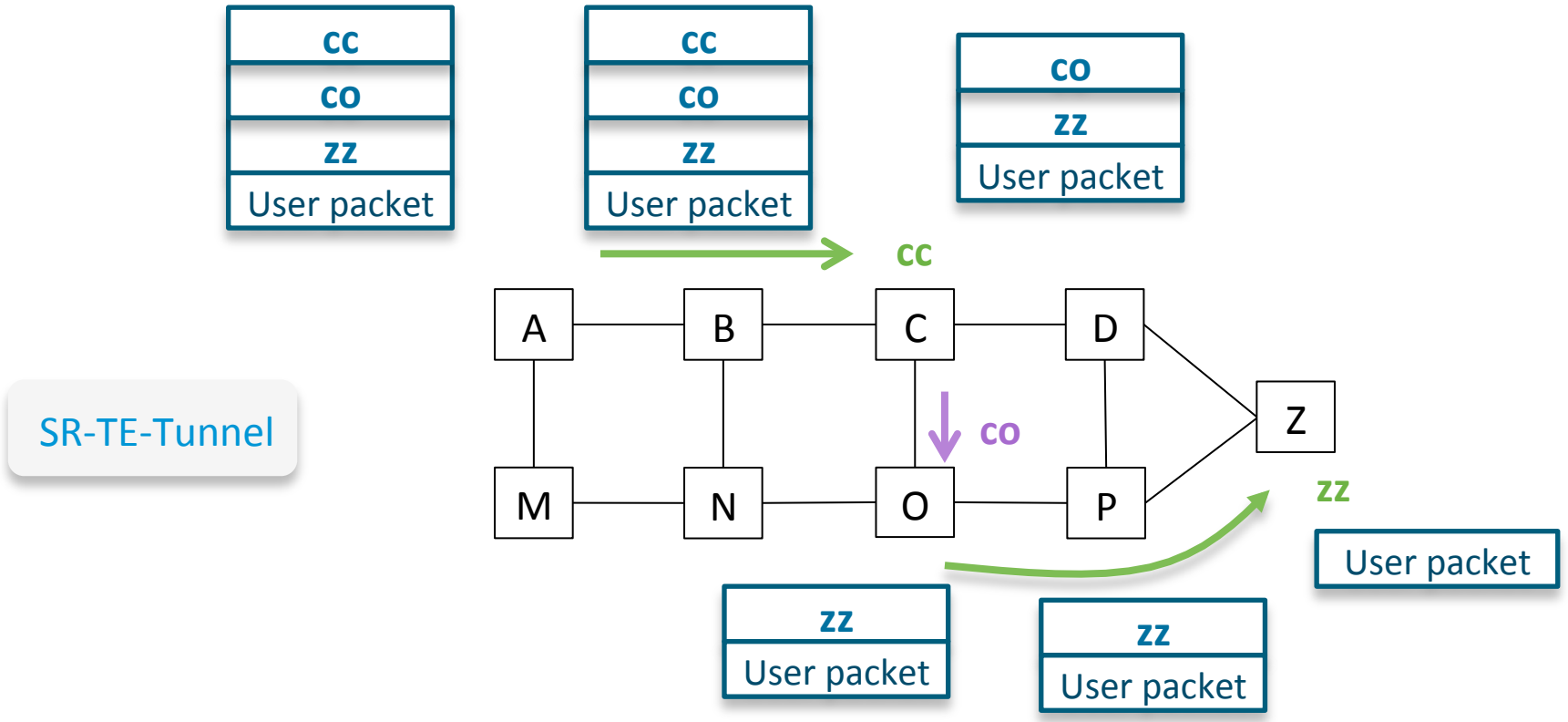
C–D間回線の帯域使用率が高いため、
SPF計算による経路では
そのSLA要件を満たせない

Segment Routing – Controllerによる制御



- Controller
  - そのSLA要件を満たすパスを発見する
  - NodeおよびAdjacency Segmentのリストをencodeする

Segment Routing - Source Routing

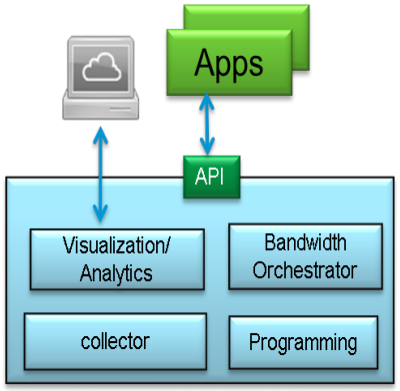


- SourceにてHeader Stackを生成し、パケットを送出する

  Pathは、Segment idのStackとして表現される。(転送ヘッダ == Label, IPv6 data planeも可)

- 中継ノードはヘッダーに応じてforwarding、RSVP stateを持つ必要がない

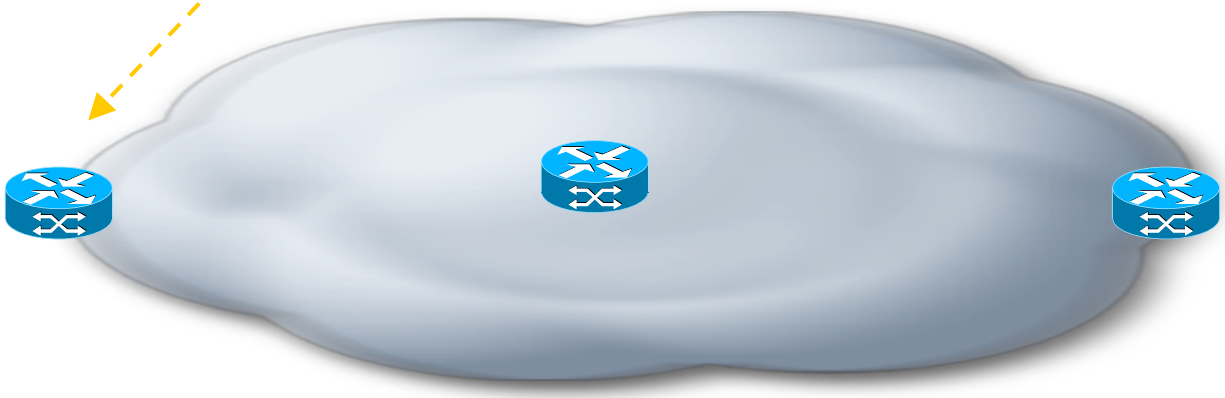
Controller : PCEP Capability Negotiation



SR Cap TLV in PCEP
Open message

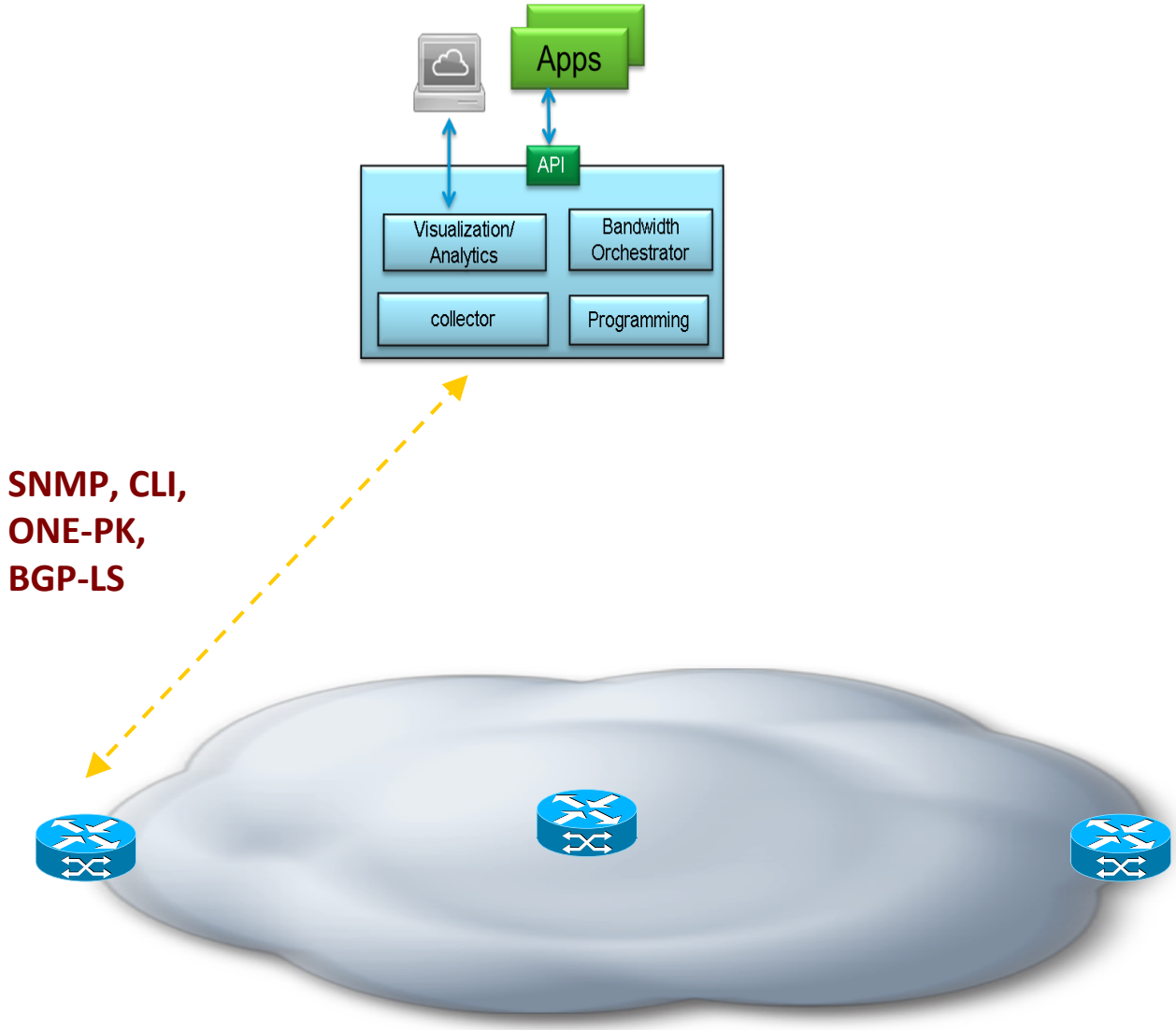
```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|           Type= TBD           |            Length=4           |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                               Flags                           |R|I|S|U|
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 1: STATEFUL-PCE-CAPABILITY TLV format



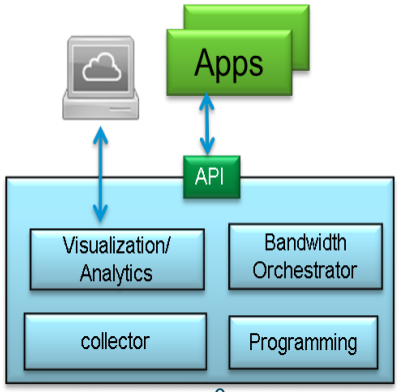
draft-sivabalan-pce-segment-routing

Controller : Topology Information Gathering



BGP-LS : draft-ietf-idr-ls-distribution

Controller : Stateful PCEP



PCCreate
PCRpt
PCUpdate

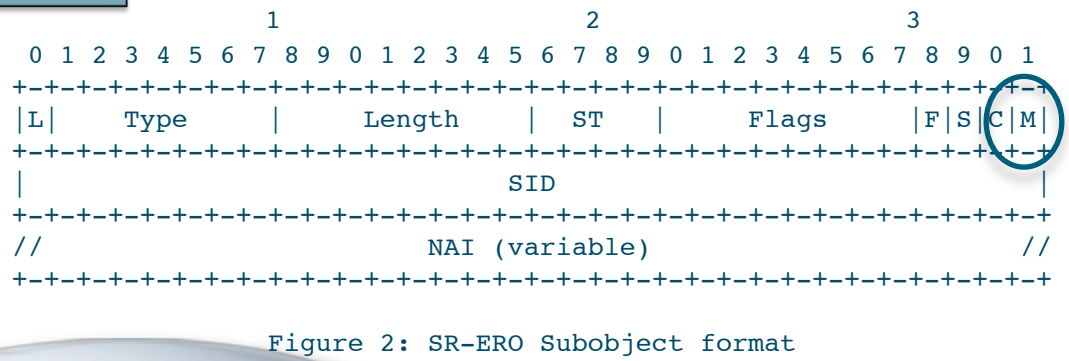


Figure 2: SR-ERO Subobject format

ST: Segment ID type
SID: Segment ID
NAI: Node or Adjacency Information (IPv4/v6)

SR-TE tunnelの設定 – PCEP extension

- Stateful PCE:
  - Constraintに応じたパス計算を行う
  - TE tunnelの生成を行う

- PCEにより計算されたSR pathは、Segment ID (SID)のstackとして表現される
  - MPLS networkの場合
    - ✓ Segment ID ＝Label(20-bit)
    - ✓ TTL および TC bits (exp bits) はPCC(*)によりセットされる

draft-sivabalan-pce-segment-routing

(*) PCC: Path Computation Client, ここではノードのこと

Segment Routing – Hybrid SDN

- SRは、自律分散コントロールプレーンとの共存が前提
  - 基本情報(Node/Adj Segment)はIGPで配布する
  - 明示的なTEが必要な部分のみ、add-on的にSR-TE pathをはる
    - 明示的Traffic Engineering
    - Controllerの障害時も、Forwardingは停止しない
  - 障害検出、OAM、protection（少なくとも一時的な）は、自律的に行う
  - EPC FRRにより、topology非依存のProtectionが可能
  - SR非対応ノード(LDP onlyなど)とも共存可能
  - Bidirectional Co-routed LSPのようなtransport的アプリケーションとも親和性が高い
  - Meta dataを運ぶことができる

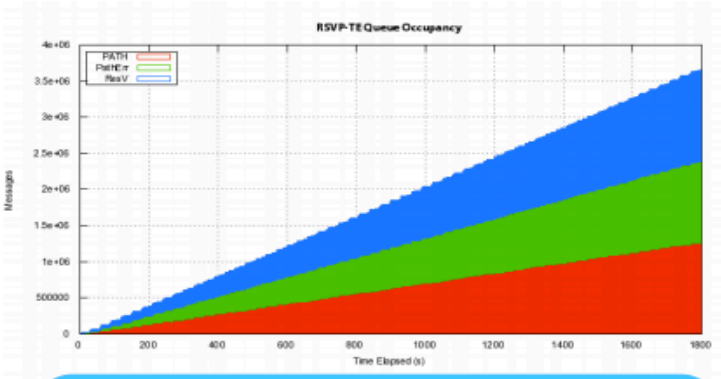
Use case #1

BT: Performance Based LSPs

Path Constraints and Technology Options.

- **Requirement for a number of types of constrained service/flow routing:**
  - Co-routing.
    - Considering SRLG/Node/Link diversity or bi-directional paths.
  - Affinity-based routing.
    - Diverging from SPT based on constraining available paths by colour/admin-group.
  - Performance-managed services.
    - Latency, available bandwidth, etc.
- Clearly, a number of these constraints can be delivered by RSVP-TE today.

- **Per-service/flow routing requires a significant increase in the number of RSVP-TE LSPs when compared to current deployments:**
  - Number of LSPs is greater than full mesh (already not recommended).
  - Scale limit of mid-point signalling during large failures.
- **Limited additional functionality is offered by having mid-point state.**
  - Generally only admission control.
  - Required in a subset of path routing scenarios.



Mid-point Overloading – Post-Mortem Model

Unbounded RSVP-TE queue growth based on inability to process PATH messages within LSP retry time – **LSPs never successfully re-signal.**

http://www.ietf.org/proceedings/87/slides/slides-87-status-0.pdf

Use case #2

DT: MPLS OAM

Segment Routing based OAM use case

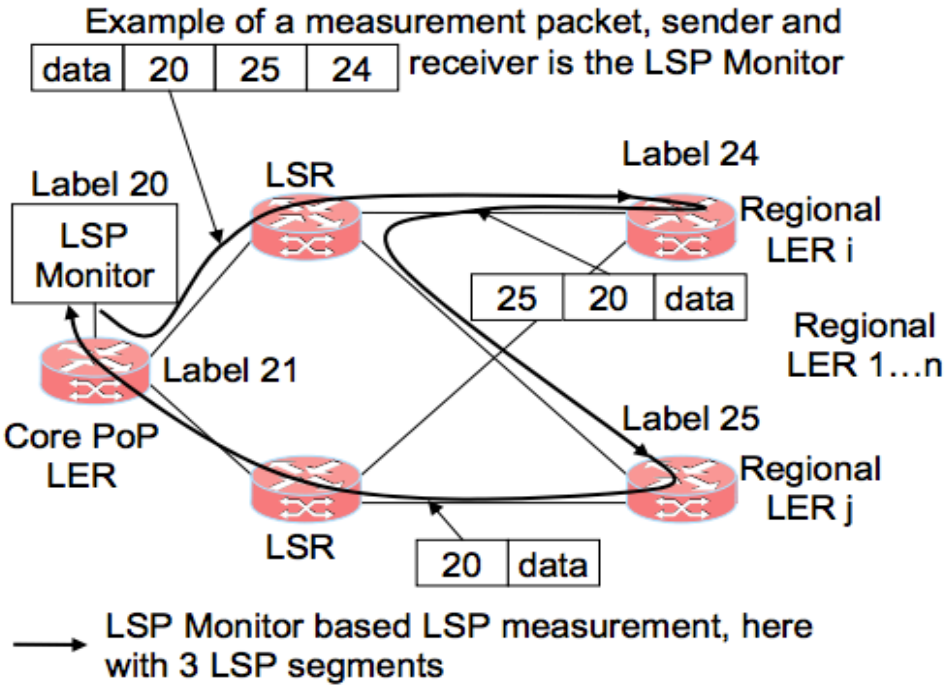
IETF 87, Berlin

Rüdiger Geib, Deutsche Telekom

Segment based Routing allows for scaleable LSP monitoring

Monitoring MPLS data plane liveliness

- source based routing allows execution of arbitrary LSP chains.
- then a ping with data plane loop can be built.
- by ISIS the LSP Monitor is aware of the network topology and its state.
- a single LSP monitor is able to adress all LSPs of a domain. A redundant design is possible if desired.
- Example to the right: the LSP monitor checks data plane liveliness between LER i and LER j. In general, by the method shown all LSPs can be monitored.



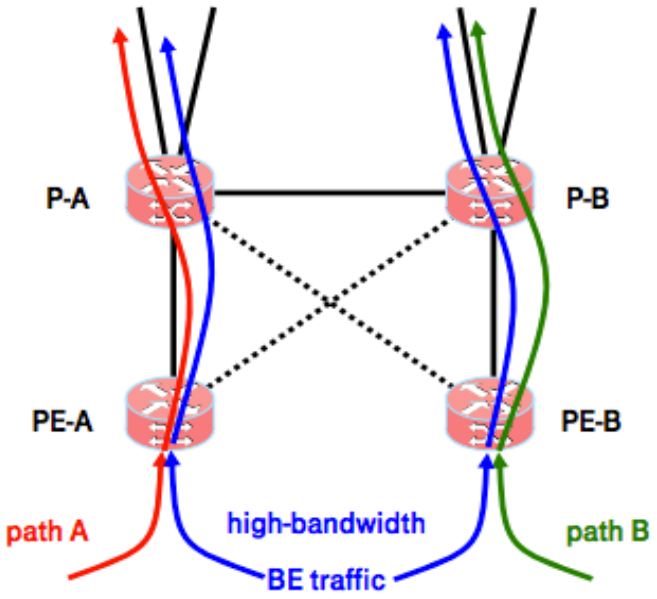
http://www.ietf.org/proceedings/87/slides/slides-87-status-1.pdf

Use case #3

DT: Disjoint Path and QoS based routing

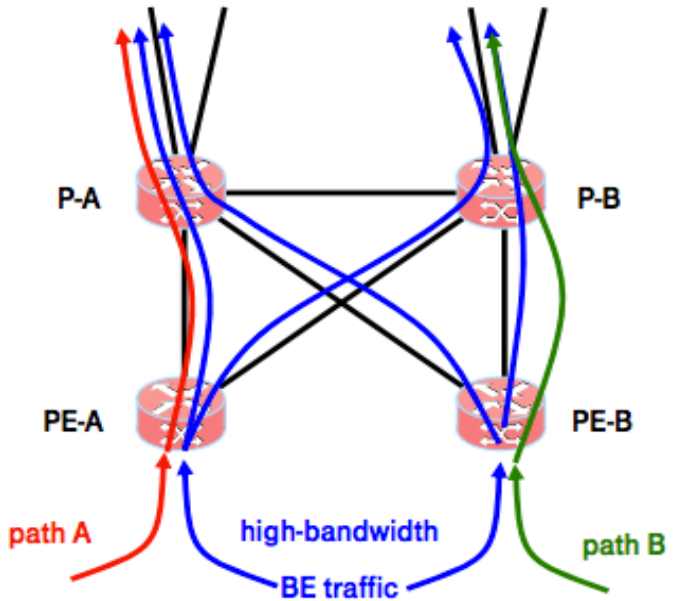
Merged network:

- Topology tailored for both disjoint paths and IP-FRR.
- Limited efficiency.



Optimized future network with SR:

- Basic topology optimized for IP-FRR and efficiency.
- Sigtran traffic constrained with A/B anycast segment to provide disjoint paths.



LIFE IS FOR SHARING.

Dr. Martin Horneffer / SR Use Cases @ DT 7/29/2013

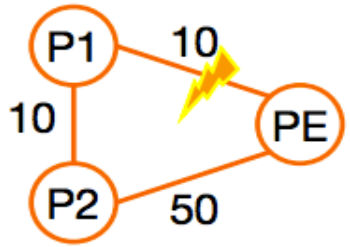
http://www.ietf.org/proceedings/87/slides/slides-87-status-2.pdf

Use case #4

Orange: FRR (Fast Reroute)

Incremental deployment in a LDP network

- As first step, Segment Routing use may be restricted to FRR backup path.
  - Keeping LDP for nominal traffic, like the way it currently is.
- If nodes are already SR capable, SR FRR can be deployed incrementally on a per PLR basis. (with incremental benefit).
  - i.e. enabling SR FRR on P1



- In the absence of SR capable node in the network, SR FRR can be deployed incrementally on a per PLR + (last) P + (first) Q basis.
  - i.e. enabling SR FRR on P1 (PLR) & SR on P2 (P) and PE (Q)
  - Note that on the Q, SR may be replaced by a T-LDP session (which is natively the case in the above example)
- More details in draft-filsfils-rtgwg-segment-routing-use-cases-01#section-6.4

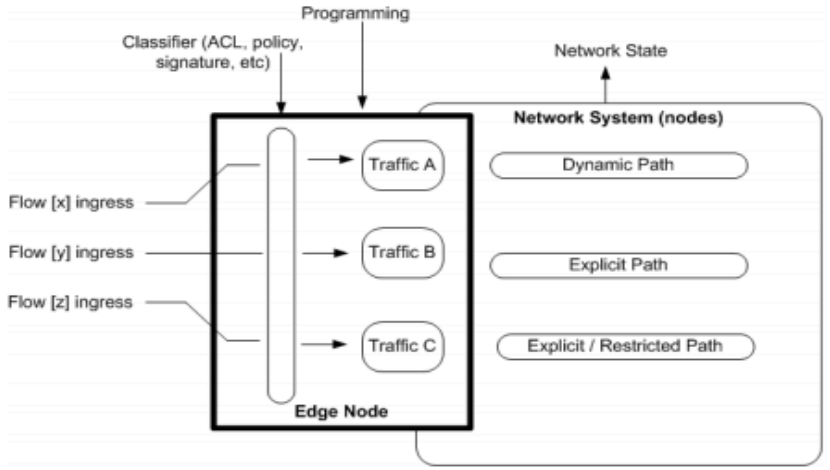
http://www.ietf.org/proceedings/87/slides/slides-87-status-3.pdf

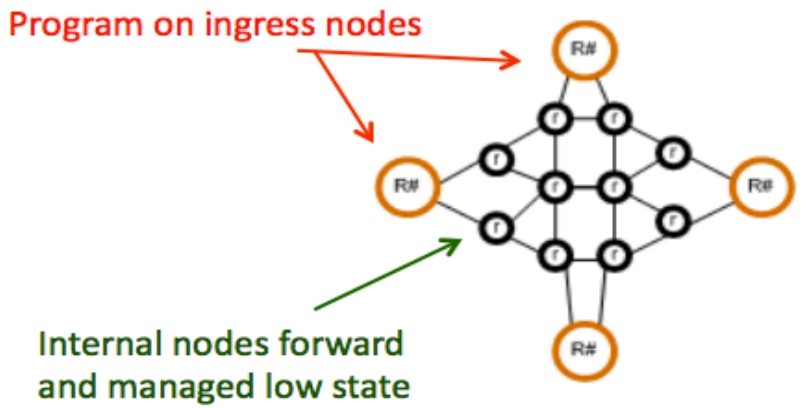
Use case #5

Rogers: Converged Multi-network operation

Programming and SDN Interaction

- Automation of the network is essential for future operation
  - Current operational modes not scalable
- SDN (path programming within this document's context) is desired, with per-flow/service network treatment
- Minimize the number of elements where programming must occur, and simplify configuration required

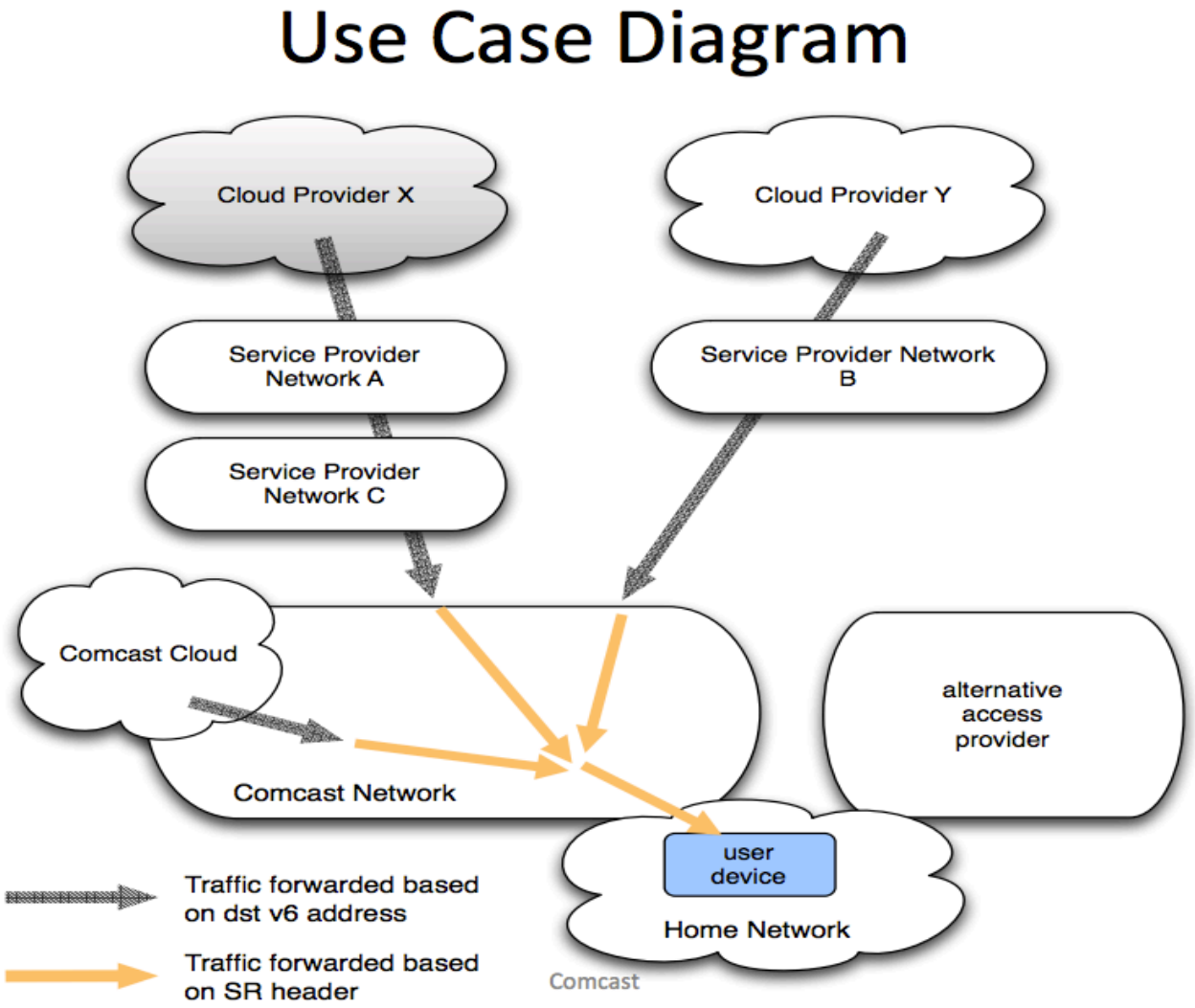




http://www.ietf.org/proceedings/87/slides/slides-87-status-4.pdf

Use case #6

Comcast : IPv6 use case



http://www.ietf.org/proceedings/87/slides/slides-87-status-11.pdf

Segment Routingのメリット

1. Simple !!!
   - 使用するControl PlaneはIGPだけ。LDP, RSVPは要らない。
     - LDP-IGP syncなどのstate syncも必要ない
2. アプリケーションの要請に柔軟に対応：ScalableなTraffic Engineering
   - Service要請に応じたパス設定、フロー設定
   - 帯域、latencyなどを加味したCSPF(Constraint SPF)
   - Stateless !
     - RSVP stateを持つ必要が無い
     - 全てのstateはヘッダ(Label Stack)にある
3. Transport的pathの運用にも有用
   - Controllerによる明示パス指定，Multi-layer PCEによるマルチレイヤ制御
   - MPLS-TP OAM, Bi-directional Co-routed LSP, Path Protection

目的に合った、
程よい、
集中と分散の配分
↓
究極のHybrid SDN

Related I-Ds

- Architecture (draft-filsfils-rtgwg-segment-routing)
- Use-Case (draft-filsfils-rtgwg-segment-routing-use-cases)
- ISIS extension for SR (draft-previdi-isis-segment-routing-extensions)
- OSPF extension for SR (draft-psenak-ospf-segment-routing-extensions)
- FRR with SR (draft-francois-sr-frr)
- PCEP extension for SR (draft-sivabalan-pce-segment-routing)
- Performance Engineered LSP using SR (draft-shakir-rtgwg-sr-performance-engineered-lsps)

http://www.segment-routing.net/

Thank you.