

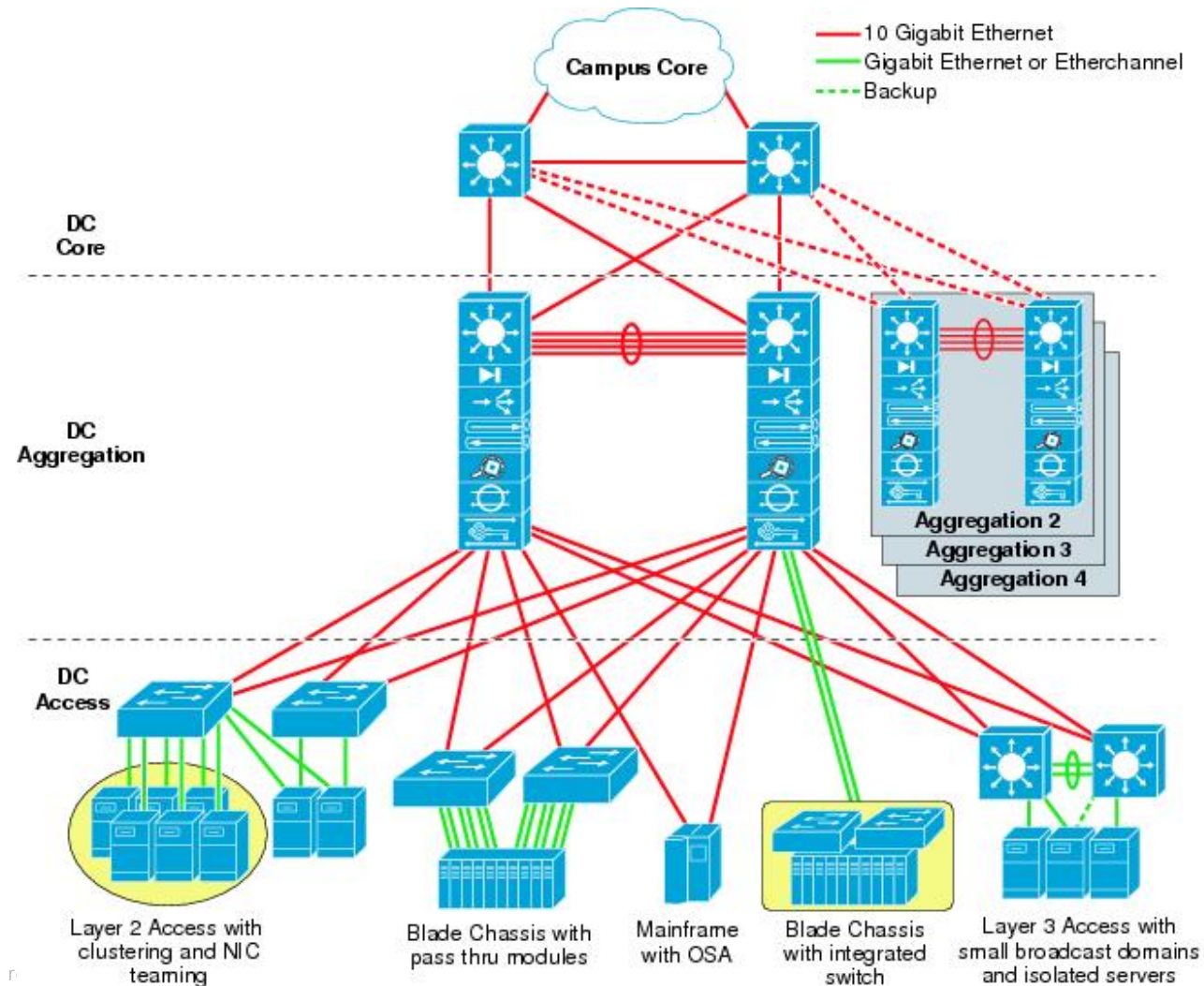
LSVR – Link State Vector Routing

Cisco Systems G.K.
Teppei Kamata
2019.11.26

本プログラムの内容

- 近年データセンターにおけるL3でのルーティングプロトコルの一つとして、Link-State Vector Routing (LSVR) がIETFにて提案されています。
- 本プログラムではなぜLSVRのようなプロトコルが提案されているのかの背景を説明した上で、LSVRの動作概要と従来のIGP/BGPによるRoutingと比較した場合のLSVRのPros/Consを解説します。

The Traditional Data Centre



DC Networkの発展に関する著名な文献



Amin Vahdat

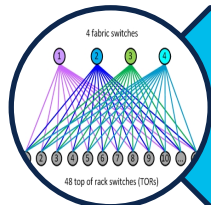
“Scale-Out Networking in the Data Centre”
IEEE Micro 30, 4 (July 2010), 29-41



James Hamilton, “Data Centre Networks Are in My Way”



“Jupiter Rising: A Decade of Clos Topologies and Centralised Control in Google’s Datacentre Network”



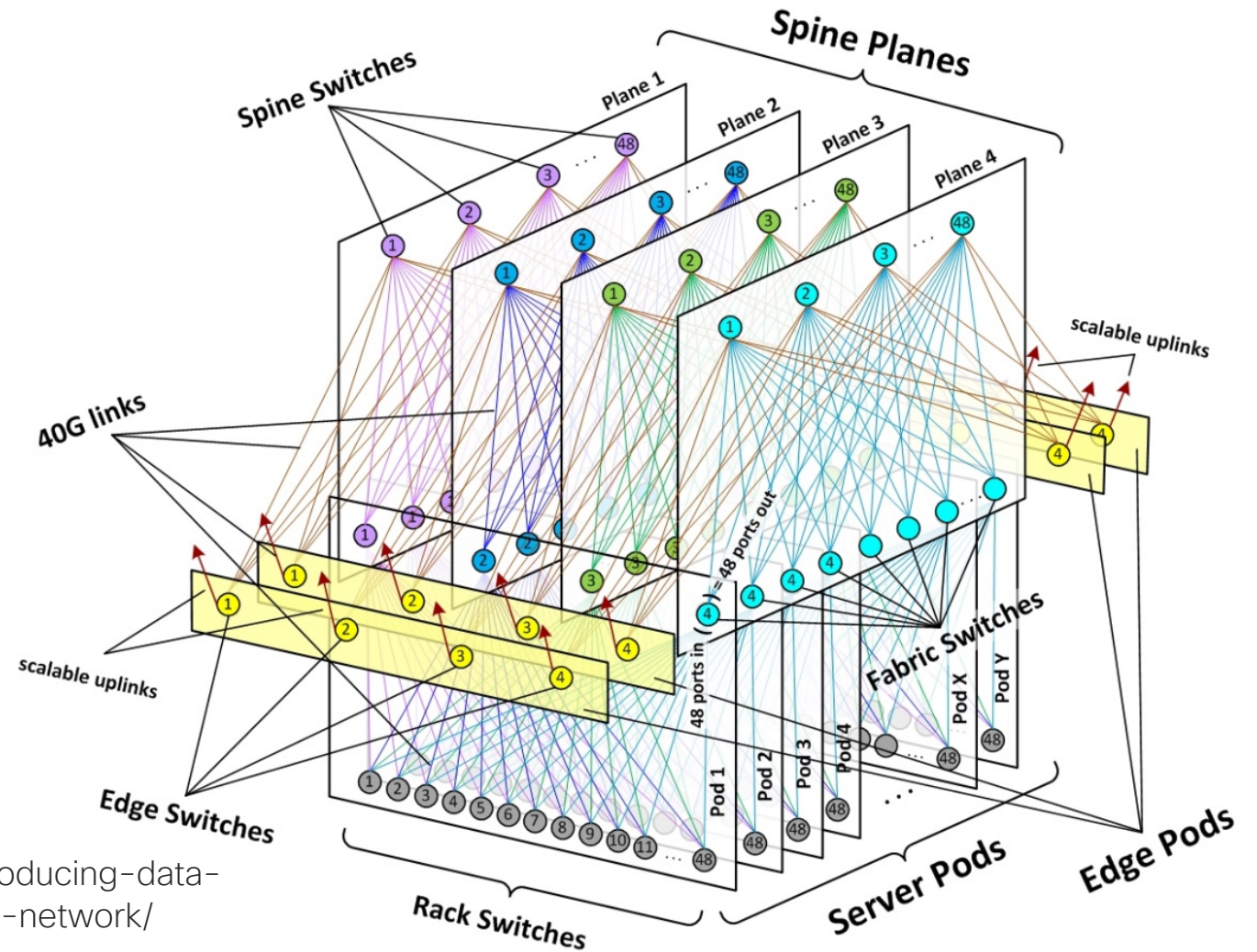
“Introducing data centre fabric, the next-generation Facebook data centre network”



ということによくあるCLOSの例

The Data Centre Fabric

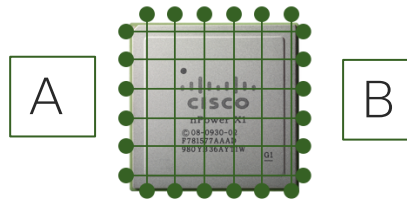
- “hundreds of thousands of servers”
- “multi-petabit bisection bandwidth”



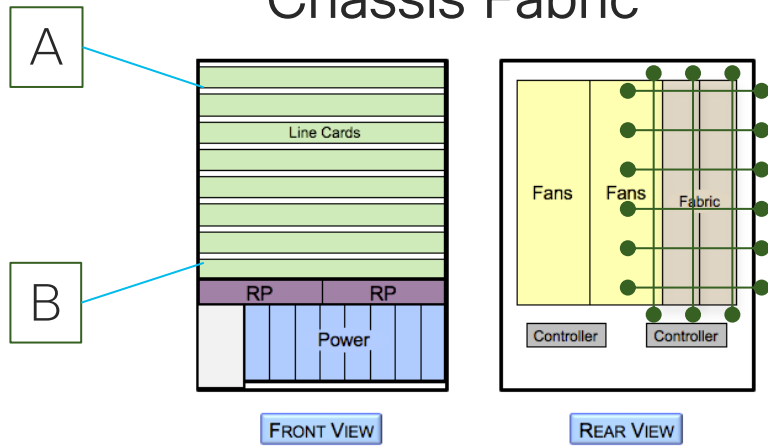
<https://code.facebook.com/posts/360346274145943/introducing-data-center-fabric-the-next-generation-facebook-data-center-network/>

Fabrics Come In All Shapes and Sizes

NPU Fabric

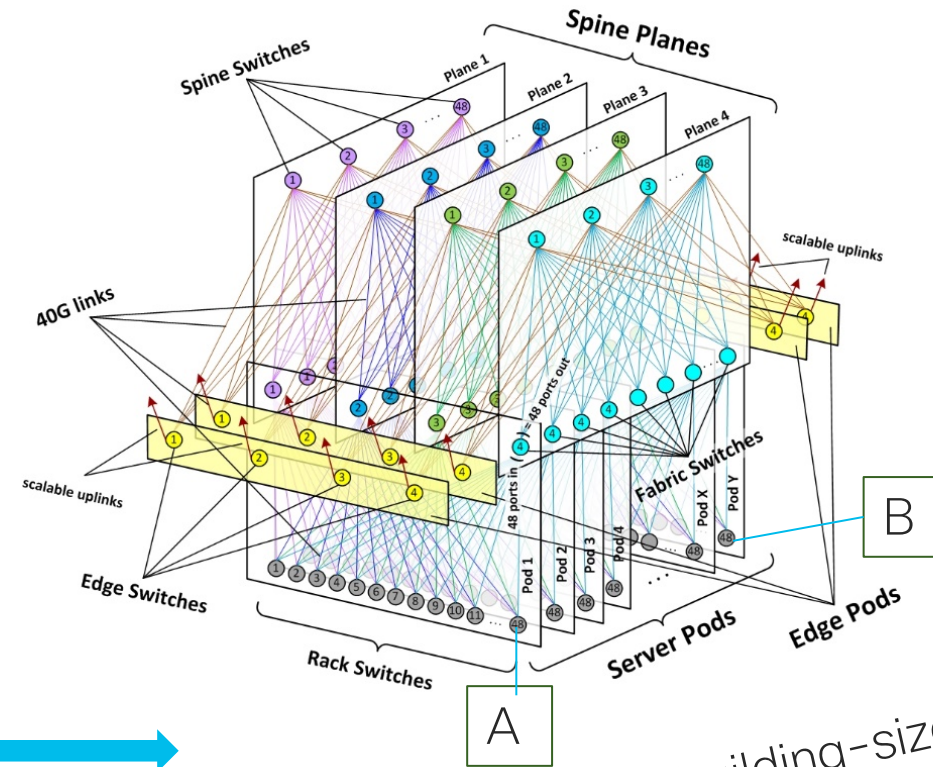


Chassis Fabric



← Internal

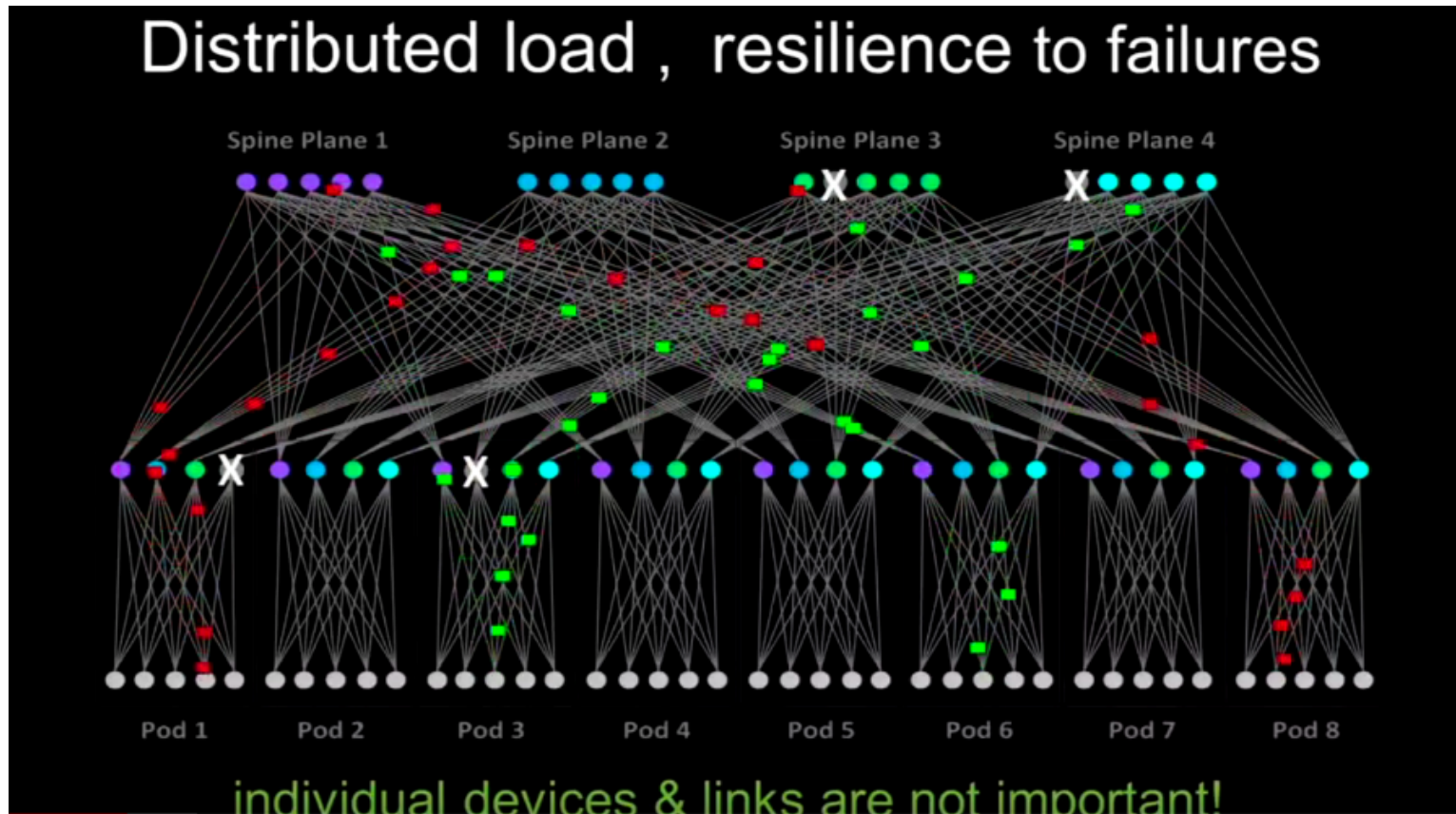
Data Centre Fabric



→ External

“building-sized router”

The Ultimate In Network Availability

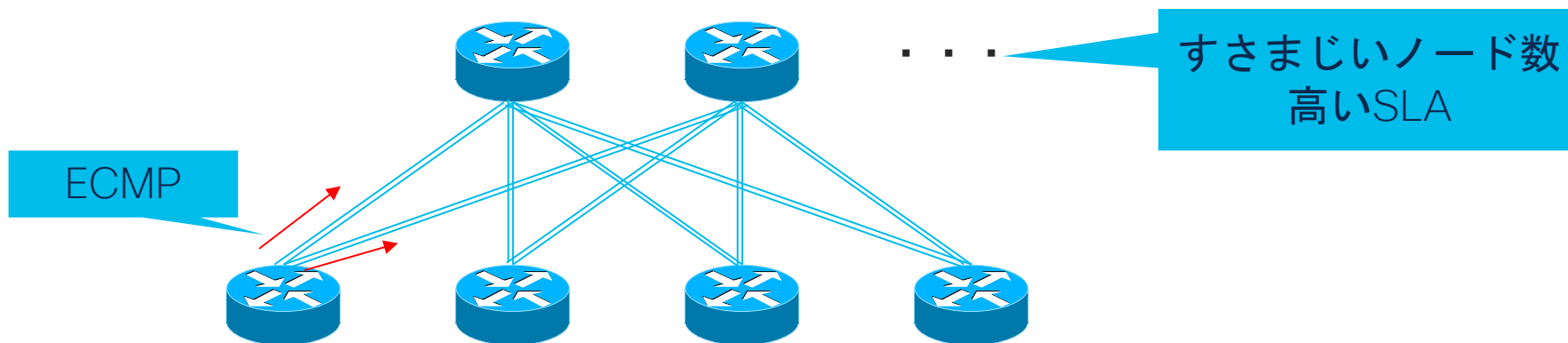


- Independent control planes – loosely coupled, significantly distributed
- Many Paths Between Servers
- Small Blast Radius
- Can Sustain Multiple Failures

<https://www.youtube.com/watch?v=kcl3fGEait0>

こんな状況に合わせたRouting Protocol?

- Link State Vector Routing & Routing In Fat TreeがIETF100で提案
- このセッションではLSVRを説明します



LSVRの特徴

- 一言でいうと、BGPでLink state情報を用いたRoutingを行う
- OSPFやISISと同じようにSPF計算を用いたRoutingが可能
- BGPで経路交換を行うのでMassive Scaleと呼ばれるCLOS NWにおいても安定して利用可能
- BGP LSと同じAF 16388に新しいSAFIを(これから定義)を用いて、Link State情報を広報する。(そのためBGP LSと同じフォーマット)

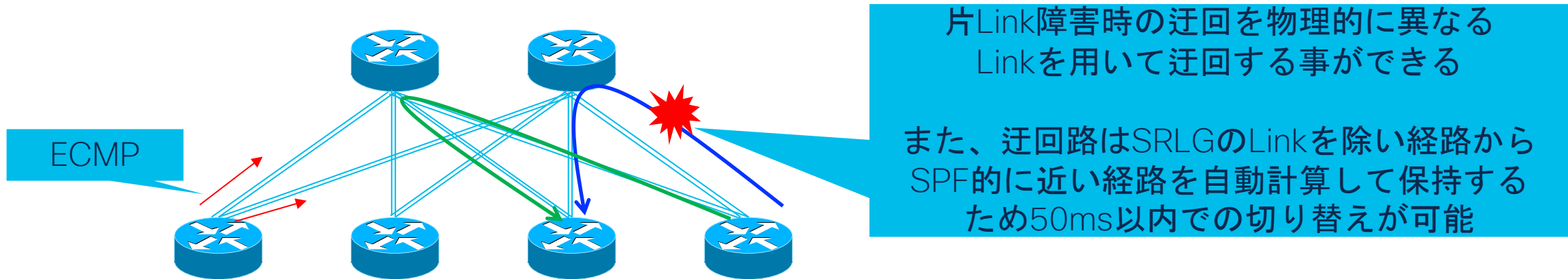
Motivation for LSVR(1)

- SPF DomainにいるすべてのBGP SpeakerがLink State情報をもたせることができる
 - ECMP
 - IP Fast Reroute(Loop-Free Alternate)による50ms Protection
 - Shared Risk Link Group(SRLG)を考慮したRouting
 - 上記を含めたOSPFでできることがBGPでもできる

※上記は必ずしもBGPでは不可能なことができるようになったという意味ではありません

Motivation for LSVR(1)

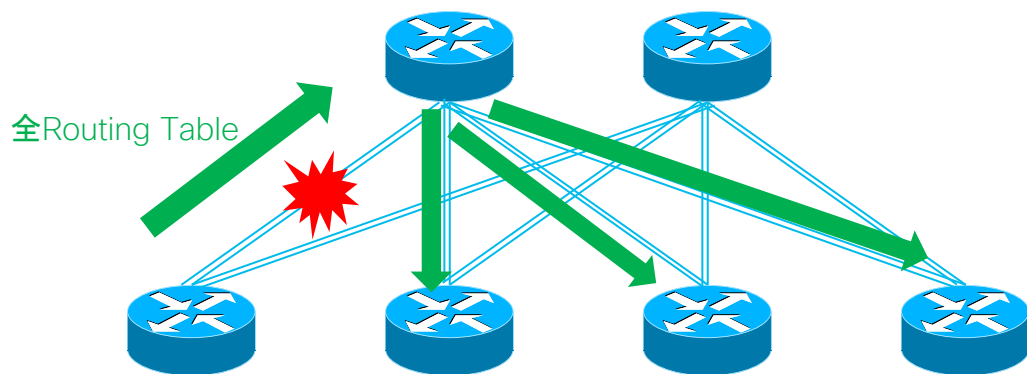
- ECMP
- IP Fast Reroute(Loop-Free Alternate)による50ms Protection
- Shared Risk Link Group(SRLG)を考慮したRouting



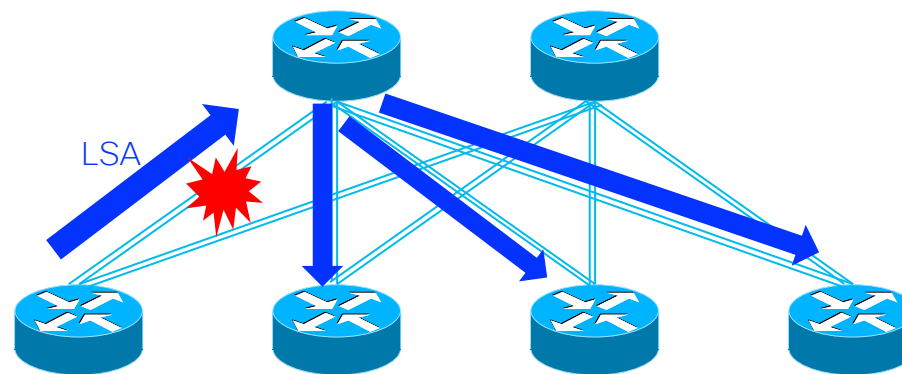
Motivation for LSVR(2)

- BGPでSPF計算を行うことによりNLRIを削減できる可能性
 - 通常BGPではDistance-vector型の経路計算を行うので、Single Link failureが多数のPrefixにImpactを与え、NLRIの広告やWithdrawが多数発生しうる
 - BGP SPFでは障害が起きたLink NLRIのみがWithdrawされる

通常BGP(Distance-vector)



BGP SPF



LSVRをIGPと比較したときのメリット

- IGPをそのままDCで使う場合の一番の懸念事項はLSA/LSPのFlooding
 - Closトポロジでは多くの重複したLSPが存在してしまう
 - FloodingにはTCPにあるReceive windowが存在しない
(Flooding自体がReliableであるために用いているので当然なのですが)
しかしReceiverのQueueがいっぱいになっている可能性は否定できない
- しかし、上記に対するIGPのImprovementも複数提案されている
 - Reduction of redundant flooding – ISIS/OSPFのExtension
<https://datatracker.ietf.org/doc/draft-ietf-lsr-dynamic-flooding>
 - Elimination of flooding to leaf nodes
<https://datatracker.ietf.org/doc/draft-ietf-lsr-isis-spine-leaf-ext>
- 上記に対する考察などを会場でお話します

