

世界の巨大インターネット企業の最新トレンド: ALL-L3DCネットワークのトレンド

2022年1月26日

Ruijie Networks Japan

趙 漢卿

・ 趙 漢卿 (ちょう かんきょう) ・ Ruijie Networks CO.,LTD

- 2017年入社
- 技術開発部 セールスエンジニア

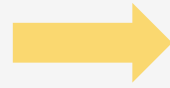


- 所在地：中国 福建省
- 設立：2000年
- 売上：1,814億円 (2021年度)
- 従業員数：7,000+ (R&Dエンジニア：3,500+)
- 事業内容：ネットワークスイッチ、ワイヤレス、ルータ、セキュリティなど
世界50ヶ国+で展開中。

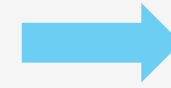
中国のインターネット企業における対応の推移

Ruijie

スイッチ仮想化
(スタッキング)
CSS/VSS/IRF/VSU



ポート仮想化
M-LAG/vPC



仮想化なし
(疑似ALL-L3)
De-stacking



導入済み

Tencent

導入済み

Baidu 百度

導入済み



導入済み

DiDi

導入済み



導入済み

美团网
meituan.com

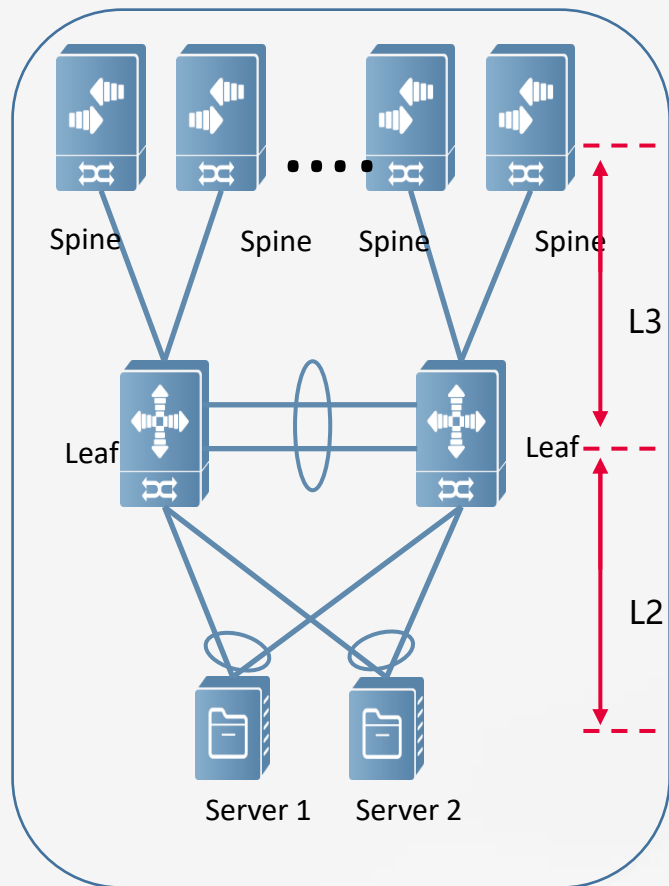
導入済み

TikTok

導入済み

導入実績：
De-stackingを利用される
ToRスイッチは10万台以上。

運用実績：
高い安定性、De-stackingに関する故障0件。
3 | www.ruijienetworks.com



MLAG(Multichassis Link Aggregation Group)

メリット

- 原則ファームウェアバージョンに差があっても組める。
- 冗長プロセスのコンバージェンスまでの時間が短い。
- それぞれのSWのCPUなどリソースをフル活用可能。

デメリット

- 設定が複雑。MLAGを組むポート間はコントロールのカップリングの為に情報の同期、状態確認などが必須。
- MLAG構成では配下セグメントはL2になる。同一ブロードキャストドメインであるが故の課題は抱える。
- MLAGも未だにメーカー依存の仕様を持つことが多い。インターオペラビリティに注意が必要。

MLAGの課題おさらい

- ✓構築する為の設定が複雑である
- ✓MLAGを組むスイッチ間同期に問題が発生すると障害に
- ✓L2通信であるが故の課題を抱える
- ✓メーカー依存仕様であることが多い

など



これらを解決することができるのがDe-Stackingです

De-stackingって何？

同一MAC/IPを持つ2台のアクティブスイッチで構成
+
偽装LACP
+
ホストルートアドバタイズメント

De-stacking概要1 / 疑似ALL-L3通信を実現

✓ Active-Active gateway

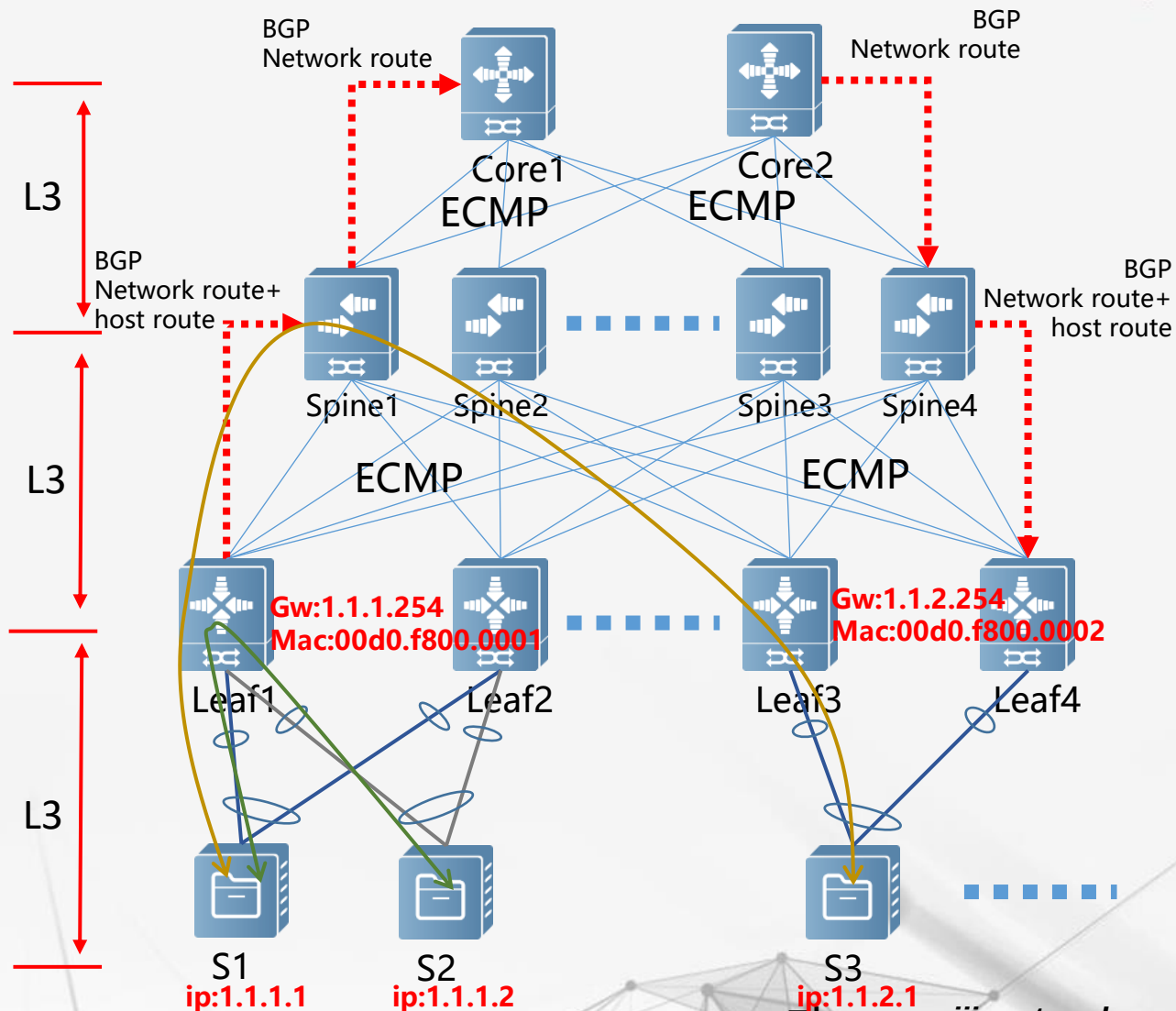
- Leaf1とLeaf2のSWは同IPとMAC。
- どちらもアクティブ。

✓ 偽装LACP

- Leaf1とLeaf2のSystem IDを同じに、Device IDとPort IDは異なるように設定する。
→サーバーからは、Leaf1とLeaf2は同じ一つのスイッチとして見える。

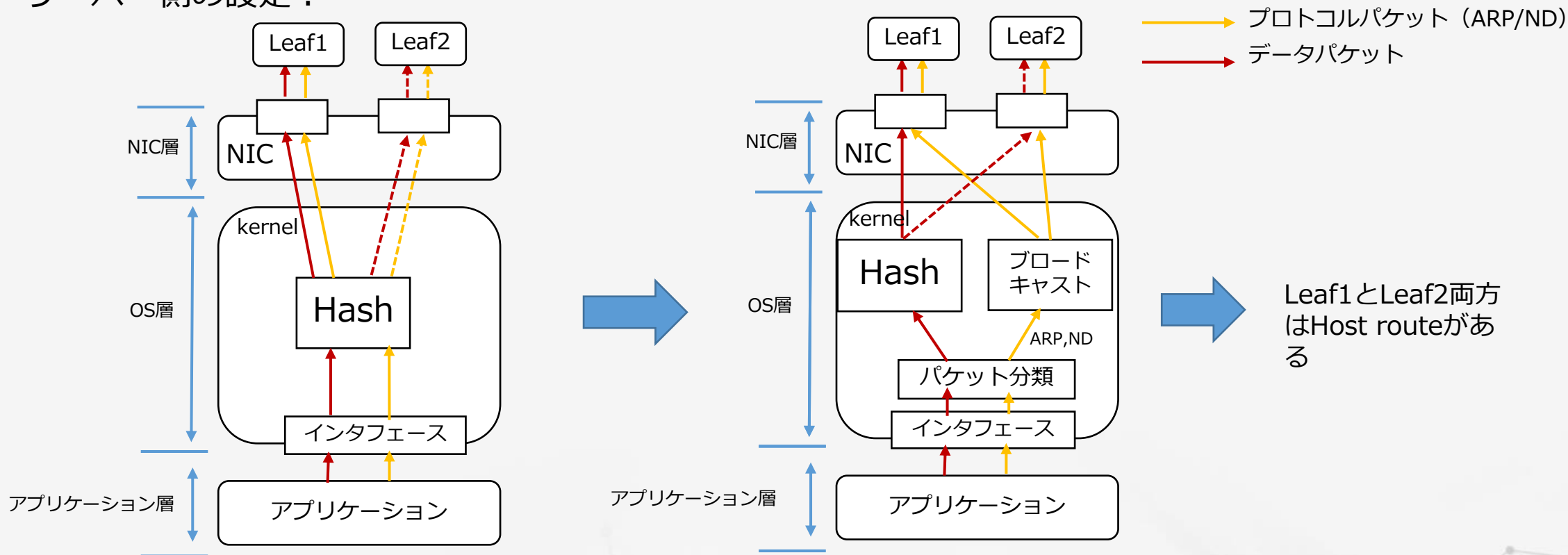
✓ ホストルートアドバタイズメント(通知) (※主にBGPでの実績が多い)

- Leaf1/2のL2ブロードキャストを無効化。
- 代わりにプロキシARPを有効化。
※全てのARP requestを代理応答する。
- スイッチが学習したARP情報をホストルートに変換する。
- 該当ホストルートはBGPなどのルーティングプロトコルによりネットワークに通知され共有される。



De-stacking概要2 / サーバ上で必要な設定 *Ruijie*

サーバー側の設定：



サーバーNICのBonding modeをmode 4に設定するだけ (LACPモード)

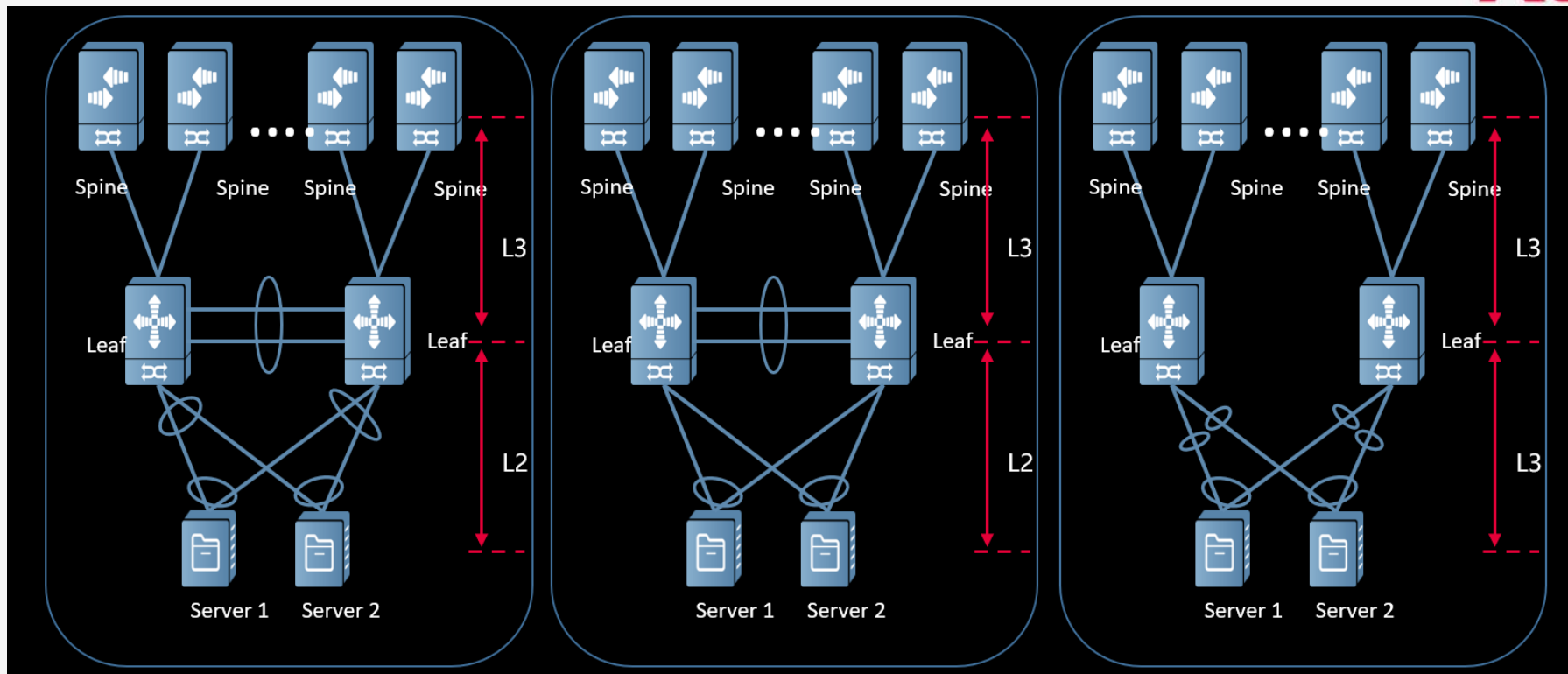
サーバー側：Linuxのコア処理変更の目的と、その挙動

- ARP (IPv4の場合) とND (IPv6の場合) パケットのみブロードキャストで両NICから送信するようになる。
- 通常のデータパケットはBondでHashの方式(通常のLACP方式)でいずれかのNICを選んで送信する。

De-stackingの特徴

メリット	
信頼性	高い。スイッチ間はデカップリングになっている為、影響なし
アップグレードリスク	低い。ファームウェアアップグレードの際にもう一台へ影響なし
設定&メンテナンス	簡単。機器間の同期、状態確認などを行わない
メーカー特定	制限なし、異なるメーカーのスイッチ同士でも組むことが可能
安定性	L2通信がないので、ブロードキャストが最大限に抑える また、De-Stacking対応スイッチは代理応答に十分耐えうる性能を持つ
デメリット	
サーバー設定	Linuxコアファイルの修正必要。 ※イメージファイルの方式でパッチインストール可能 ※ALL IP化と比較すると作業の手間やリスクは小さい
新築データセンター限定	古いサーバーでは一部だけをLinuxコアファイル変更が難しいので、新築データセンター案件向け。

ご参考：Stack/MLAG/De-Stacking比較



	Stacking	M-LAG	De-stacking
信頼性	中。管理、コントロール、データ全てカップリング。	中。プロトコルレベルのカップリング、状態、テーブルなどが同期必要。対象ポートの通信速度同期にも影響する。	高。デカップリング、お互いに影響しない。
アップグレードリスク	高。アップグレードの際にスタッキングが中断。	低。アップグレードの際にもう一台に影響しない。	低。アップグレードの際にもう一台に影響しない。
アップグレード中断時間	普通。秒～分レベル。	短い。秒レベル。	短い。秒以内。
設定&メンテナンス	複雑。管理、コントロール、データなどの情報全部設備間に同期必要。	複雑。コントロールのカップリングがあるので、情報同期、状態確認などの設定が必要。	簡単。スイッチ間同期、状態確認などの設定は不要。
サーバー設定	なし。	なし。	Linuxコア修正必要。ARP/NDをダブルNIC送信に変更。

趙 漢卿 (ちょう かんきょう)

zhaohanqing@ruijienetworks.com

THANKS

Ruijie Networks Co., Ltd.

Address: Floor 11, East Wing, Zhongyipengao Plaza, No. 29 Fuxing Road, Haidian District, Beijing, China

Post Code: 100036

Tel: (8610) 5171-5996

Fax: (8610) 5171-5872

www.ruijienetworks.com