FlexE2.0で夢の1.6Tbps IF

Nokia Solutions&Networks
鹿志村康生
yasuo.kashimura@nokia.com

Flex Ethernet 2.0

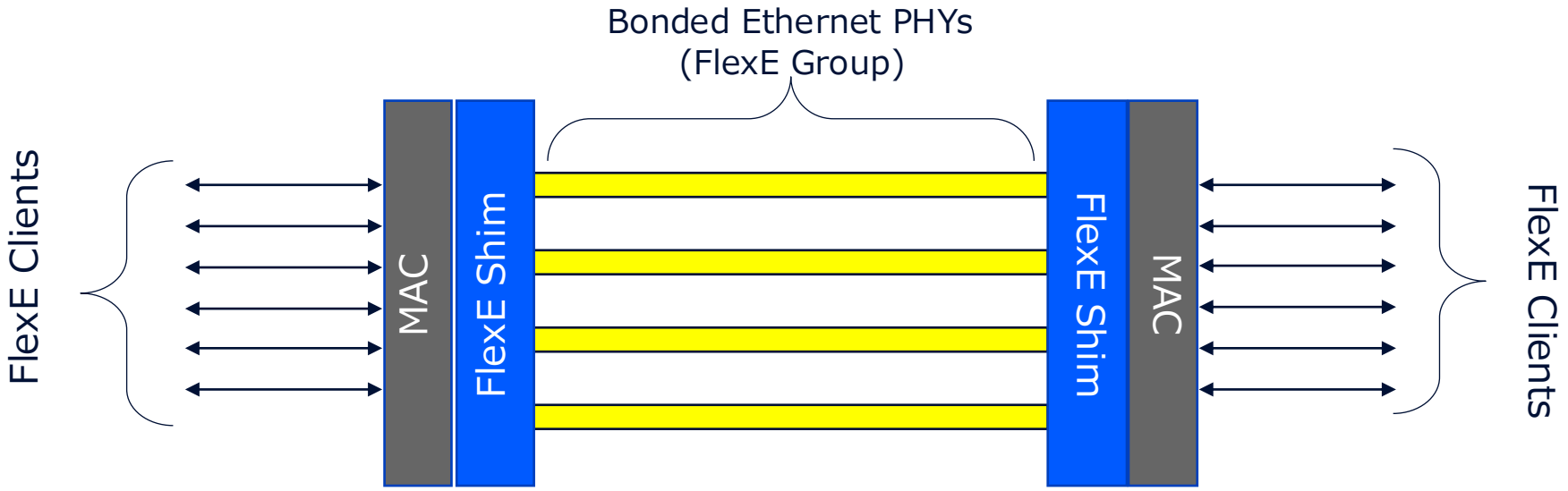
OIFで策定、2018年にFlexE 2.0版をpublish

複数のEthernetリンクをBonding

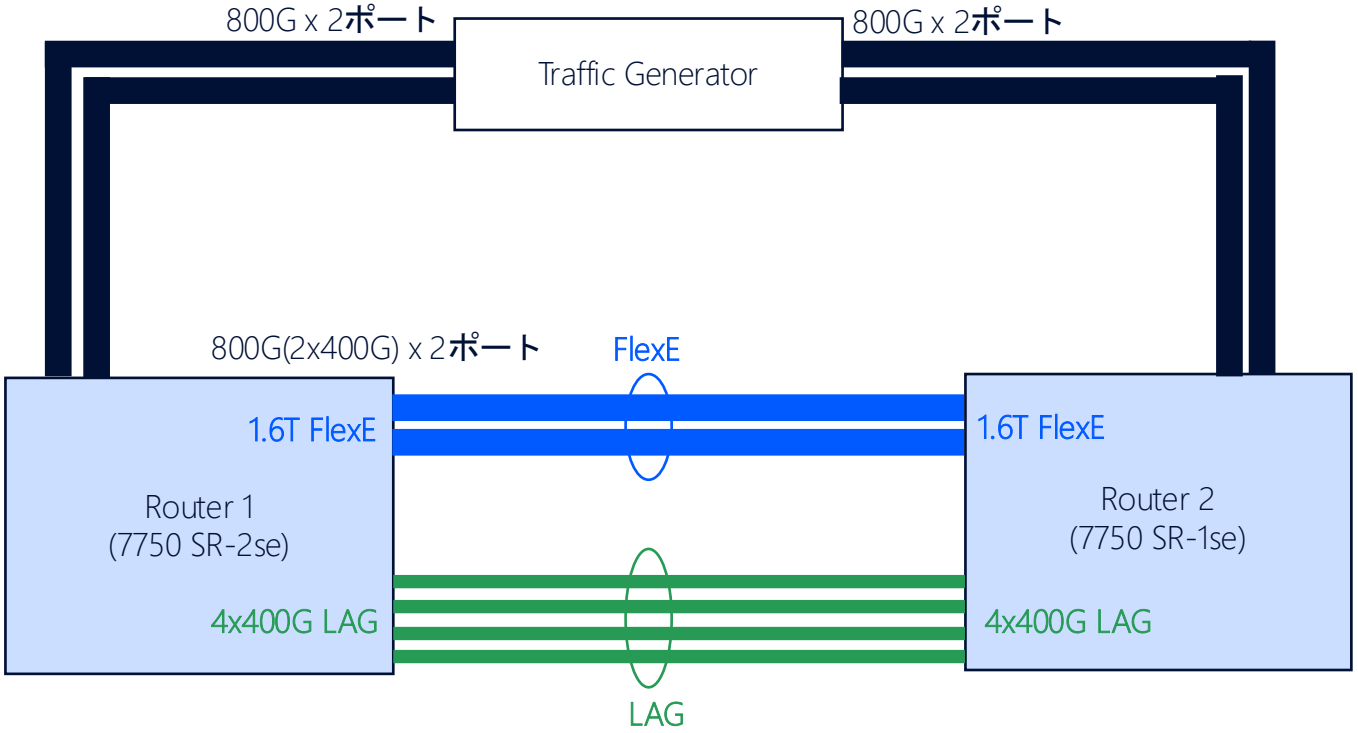
FlexE 2.0で200G/400Gをサポート

クライアント: 10G, 40G, n x 25G

Flowレベルではなく Bitstreamレベルで Trafficを分散、Elephant Flowも分散可能でより効率的なリンク帯域利用を可能とする。



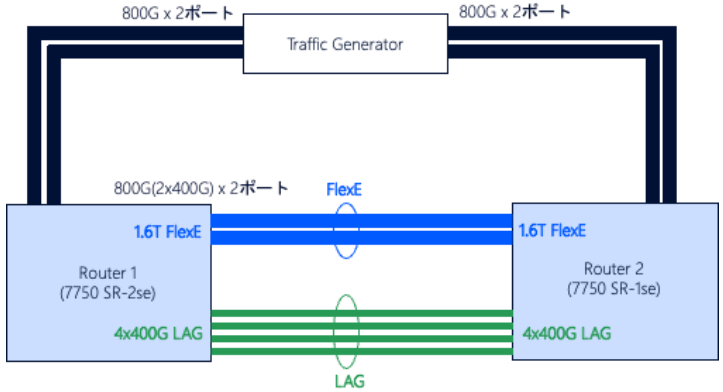
トポロジー



トポロジー

7750 SR-2se

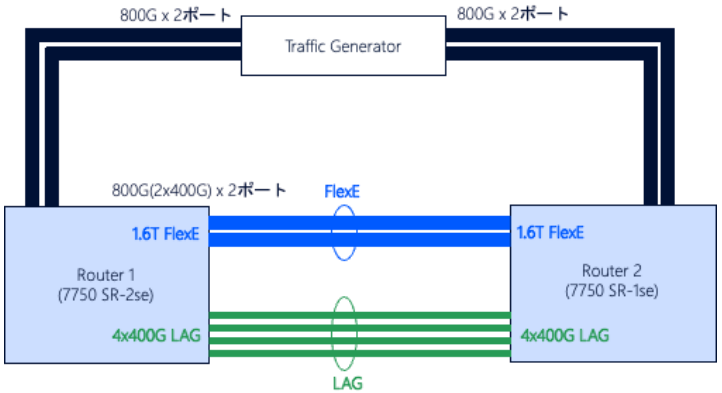
```
===============================================================================
Ports on Slot 1
===============================================================================
Port          Admin Link Port    Cfg  Oper LAG/ Port Port Port   C/QS/S/XFP/
Id            State      State   MTU  MTU  Bndl Mode Encp Type   MDIMDX
-------------------------------------------------------------------------------
1/1/c1        Up         Link Up                          conn   Reserved (10*
1/1/c1/1      Up         Link Up                          flxmbr
1/1/c1/2      Up         Link Up                          flxmbr
1/1/c2        Up         Link Up                          conn   Reserved (10*
1/1/c2/1      Up         Link Up                          flxmbr
1/1/c2/2      Up         Link Up                          flxmbr
1/1/c3        Down       Down                             conn   400G-ZR-Amp
1/1/c4        Down       Down                             conn
1/1/c5        Down       Down                             conn
1/1/c6        Down       Down                             conn
1/1/c7        Down       Down                             conn
1/1/c8        Up         Link Up                          conn   400G-ZR-Amp
1/1/c8/1      Up    Yes  Up       9212 9212    1 hybr dotq cdgige
1/1/c9        Up         Link Up                          conn   400G-ZR-Amp
1/1/c9/1      Up    Yes  Up       9212 9212    1 hybr dotq cdgige
1/1/c10       Down       Down                             conn
1/1/c11       Down       Down                             conn
1/1/c12       Up         Link Up                          conn   400G-ZR-Amp
1/1/c12/1     Up    Yes  Up       9212 9212    1 hybr dotq cdgige
1/1/c13       Up         Link Up                          conn   400G-ZR-Amp
1/1/c13/1     Up    Yes  Up       9212 9212    1 hybr dotq cdgige
```



トポロジー

7750 SR-2se

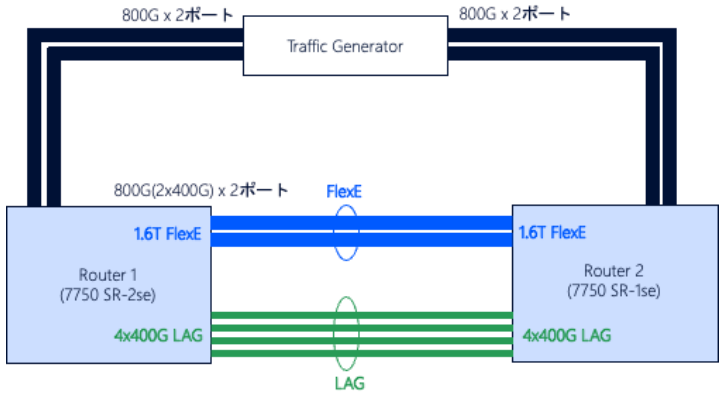
```
===============================================================================
MDA 1/1 Flex Group 11
===============================================================================
Description             : (Not Specified)
Admin State             : Up
Oper State              : Up
Group Number            : 0x00001
Hold Time Up            : 0 seconds
Hold Time Down          : 0 seconds
Number of members       : 4
Client Configured       : 1/1/f11/1
Max Bandwidth           : 1600G
Config Alarms           : overhead-frame-not-locked
                          deskew-failed
                          local-phy-fault
                          remote-phy-fault
                          bad-group-number
                          bad-phy-number
                          bad-phy-instance-number
                          flexe-map-mismatch
                          client-calendar-error
                          calendar-in-use-error

Alarm Status            : None

Reported Alarms         : None
===============================================================================

===============================================================================
Members
===============================================================================
Port                    : 1/1/c1/1
PHY Number              : 1
Config Alarms           : None

PHY Inst Config Alarms  : None

Alarm Status            : None

Reported Alarms         : None
-------------------------------------------------------------------------------
Port                    : 1/1/c1/2
PHY Number              : 2
Config Alarms           : None

PHY Inst Config Alarms  : None
```



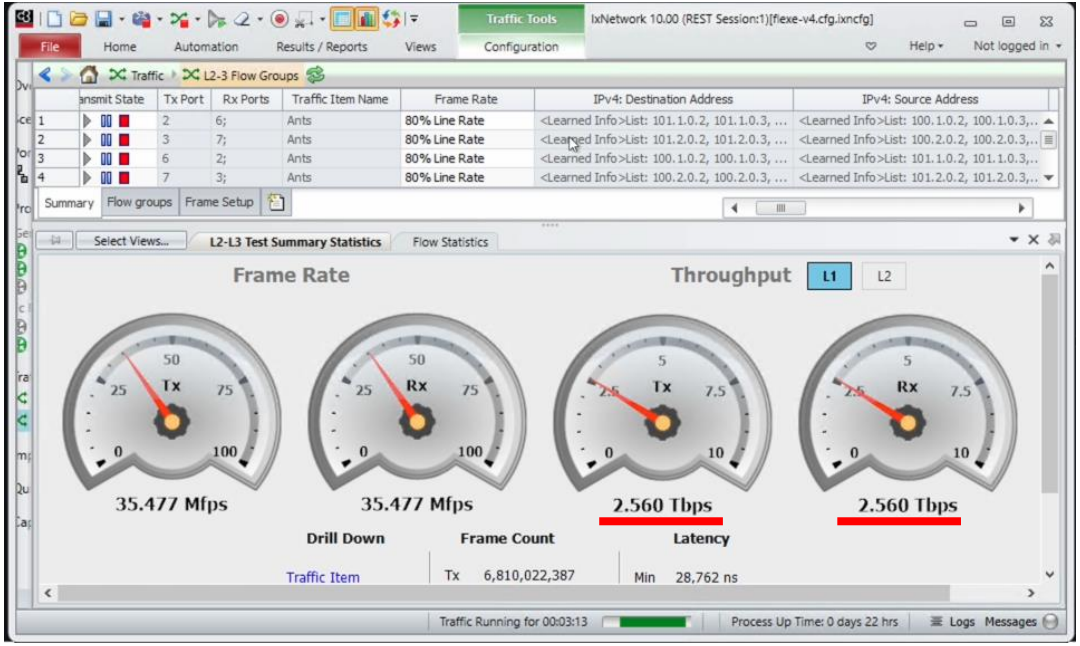
トポロジー

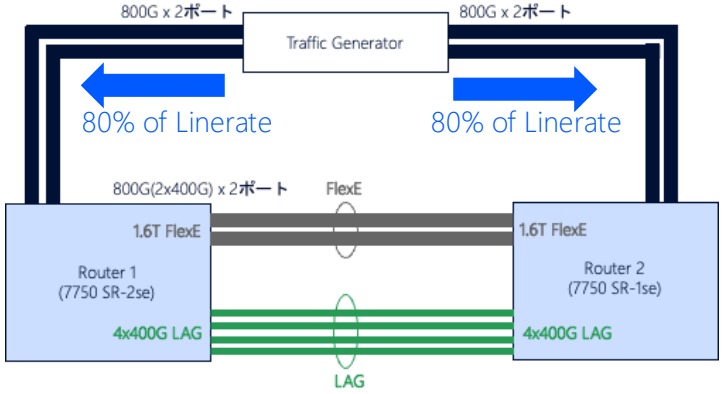
7750 SR-2se

```
===============================================================================
Ethernet Interface
===============================================================================
Description        : Flex-E Client
Interface          : 1/1/f11/1                   Oper Speed        : 1600 Gbps
FP Number          : 1                           MAC Chip Number   : 1
Link-level         : Ethernet                    Config Speed      : N/A
Admin State        : up                          Oper Duplex       : full
Oper State         : up
Config Duplex      : N/A
Physical Link      : Yes                         MTU               : 9212
Single Fiber Mode  : No                          Min Frame Length  : 64 Bytes
IfIndex            : 1610895713                  Hold time up      : 0 seconds
Last State Change  : 04/25/2025 06:57:07         Hold time down    : 0 seconds
Hold Time Down Rmng: 0 cs                        Hold Time Up Rmng: 0 cs
Last Cleared Time  : N/A
Phys State Chng Cnt: 1
RS-FEC Config Mode : None
RS-FEC Oper Mode   : None

Configured Mode    : hybrid                      Encap Type        : 802.1q
Dot1Q Ethertype    : 0x8100                      QinQ Ethertype    : 0x8100
PBB Ethertype      : 0x88e7
Ing. Pool % Rate   : 100                         Egr. Pool % Rate  : 100
Ing. Acc. Wt.      : 50                          Egr. Acc. Wt.     : 50
Ing. Net. Wt.      : 50                          Egr. Net. Wt.     : 50
Net. Egr. Queue Pol: default
Egr. Sched. Pol    : n/a
HwAggShaper Sch Pol: n/a
Monitor Port Sched : Disabled
Monitor Agg Q Stats: Disabled
Monitor Oper Group : none
Monitor HwAggShap *: Disabled
Auto-negotiate     : N/A                         MDI/MDX           : N/A
Oper Phy-tx-clock  : not-applicable
Accounting Policy  : None                        Collect-stats     : Disabled
Acct Plcy Eth Phys : None                        Collect Eth Phys  : Disabled
Egress Rate        : Default                     Ingress Rate      : Default
Oper Egress Rate   : Unrestricted
Load-balance-algo  : Default                     LACP Tunnel       : Disabled
Access Bandwidth   : Not-Applicable              Booking Factor    : 100
Access Available BW: 0
Access Booked BW   : 0
Sflow              : Disabled
Discard Rx Pause   : Disabled
Tx Pause Frames    : Enabled
```



まずはLAGでTraffic流してみる

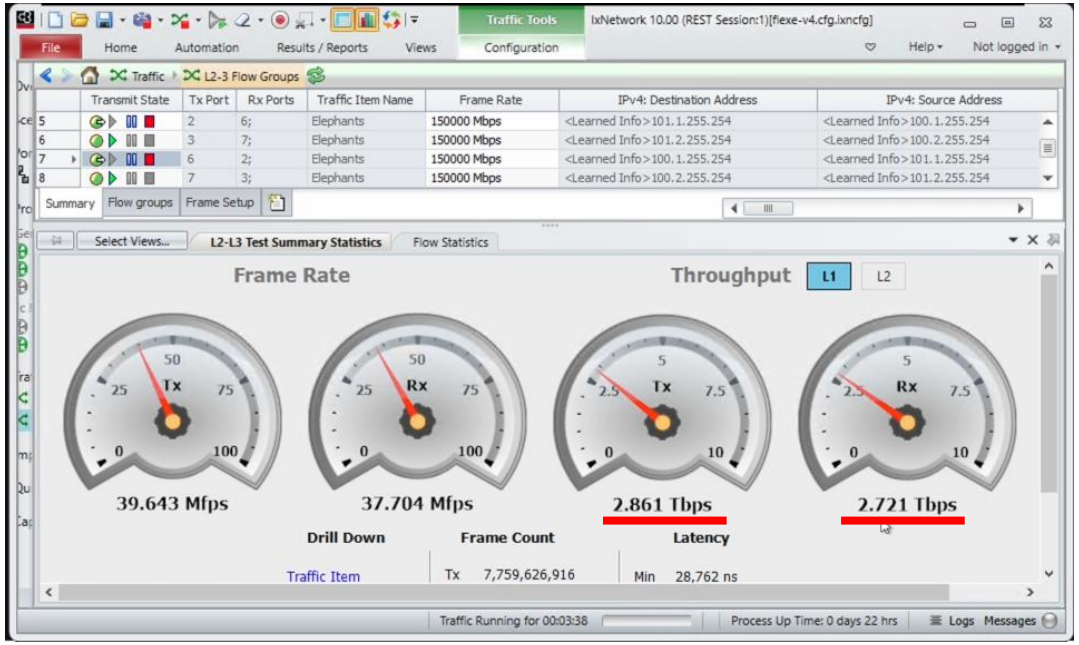


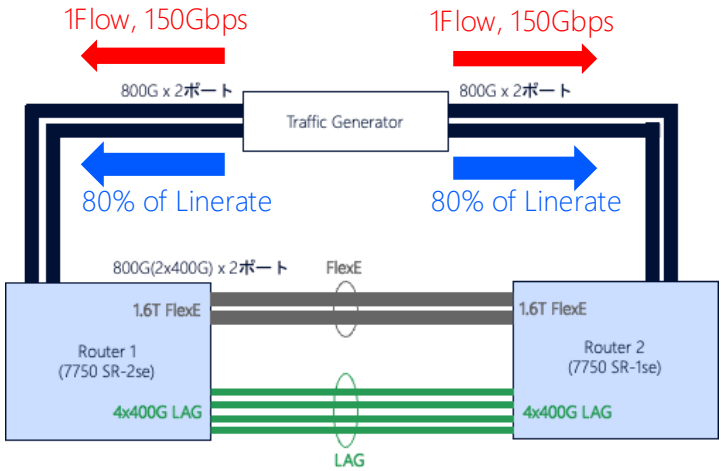


Base Traffic:
800Gbps x 2port x 80% x 2(**双方向**) = 2.56Tbps

400G LAG 1port**あたり**320Gbps(**片方向**)

Elephant Flow(150Gbps)を追加



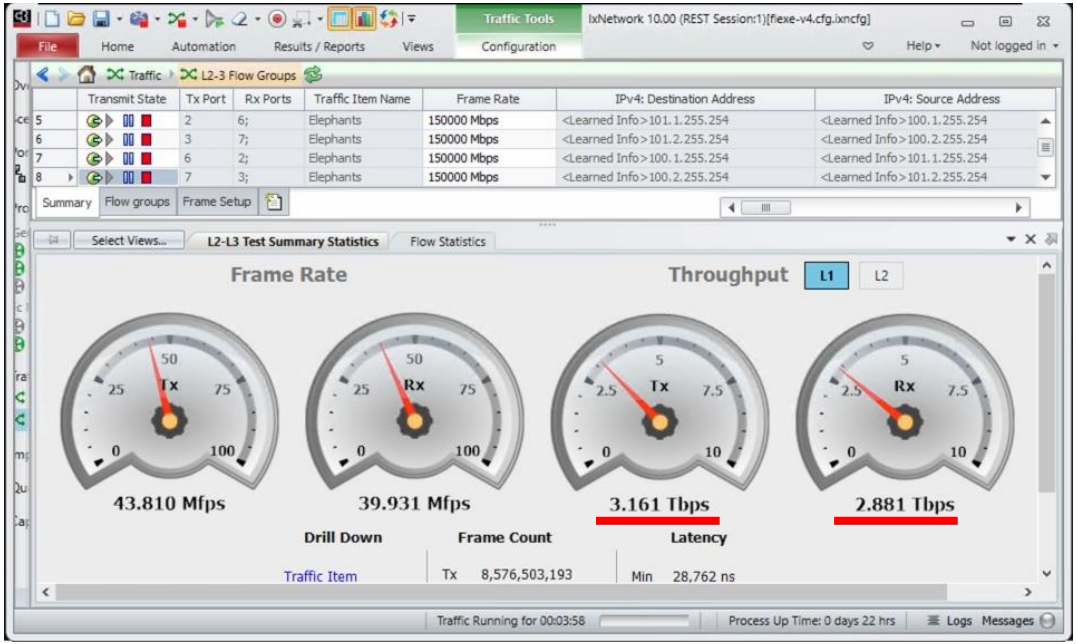


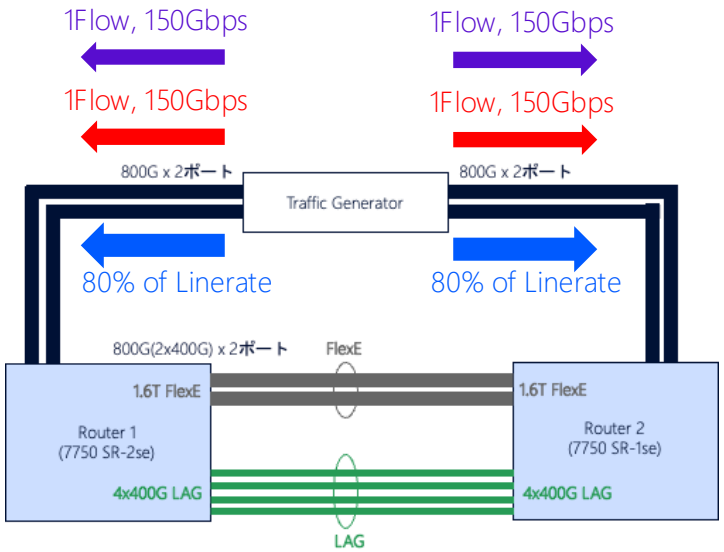
Base Traffic:
800Gbps x 2port x 80% x 2(**双方向**) = 2.56Tbps

400G LAG 1port**あたり**320Gbps(**片方向**)
+ 150Gbps**が1リンクに片寄り**Drop**発生**

(320Gbps * 3 + 400Gbps) * 2 = 2.72 Tbps

さらにもう1本
Elephant Flow(150Gbps)を追加



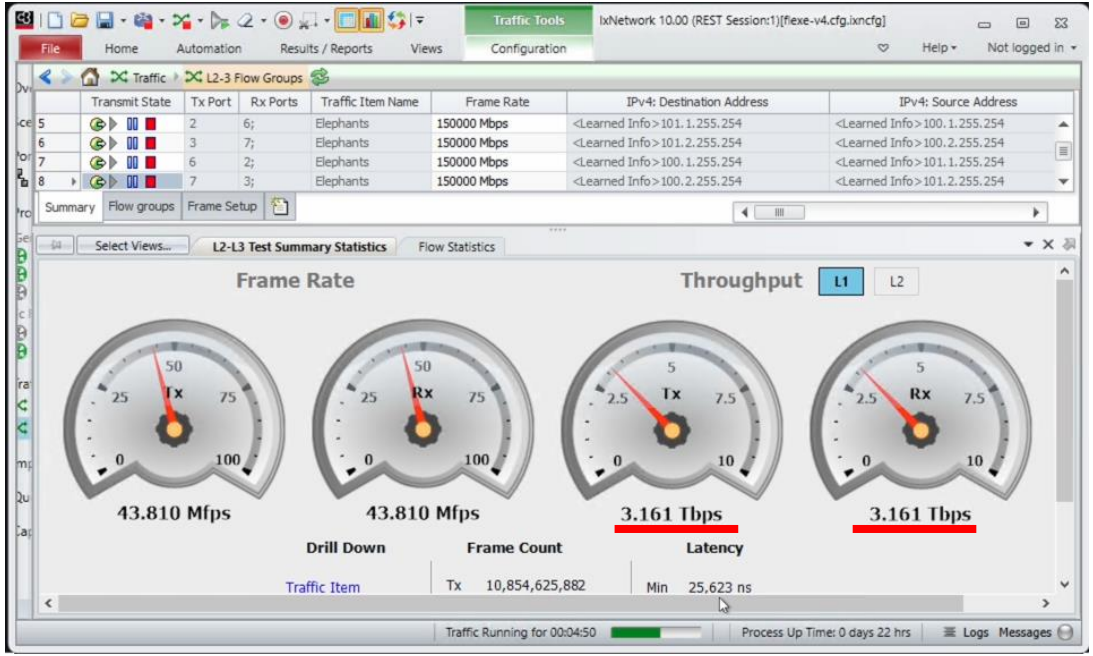


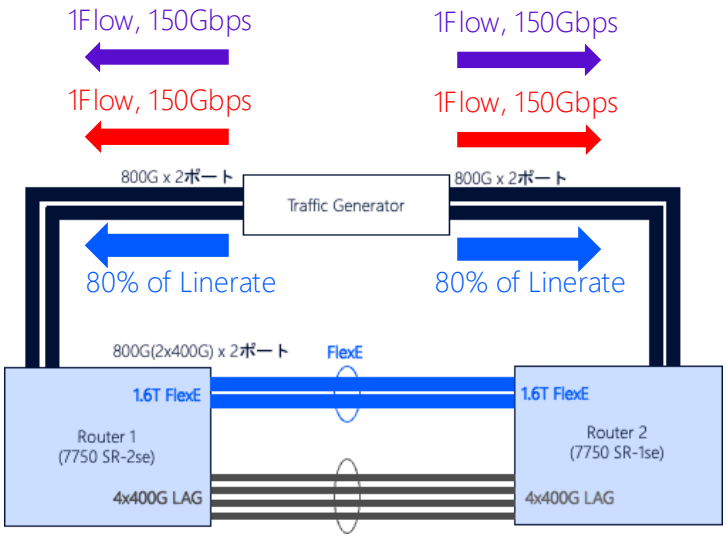
Base Traffic:
800Gbps x 2port x 80% x 2(**双方向**) = 2.56Tbps

400G LAG 1port**あたり**320Gbps(**片方向**)
+ 150Gbps**が**1**リンクに片寄り**Drop**発生**
+ 150Gbps**が別のリンクに片寄り**Drop**発生**

(320Gbps * 2 + 400Gbps * 2) * 2 = 2.88 Tbps

Routingを変更してFlexEリンクに経路を切り替える



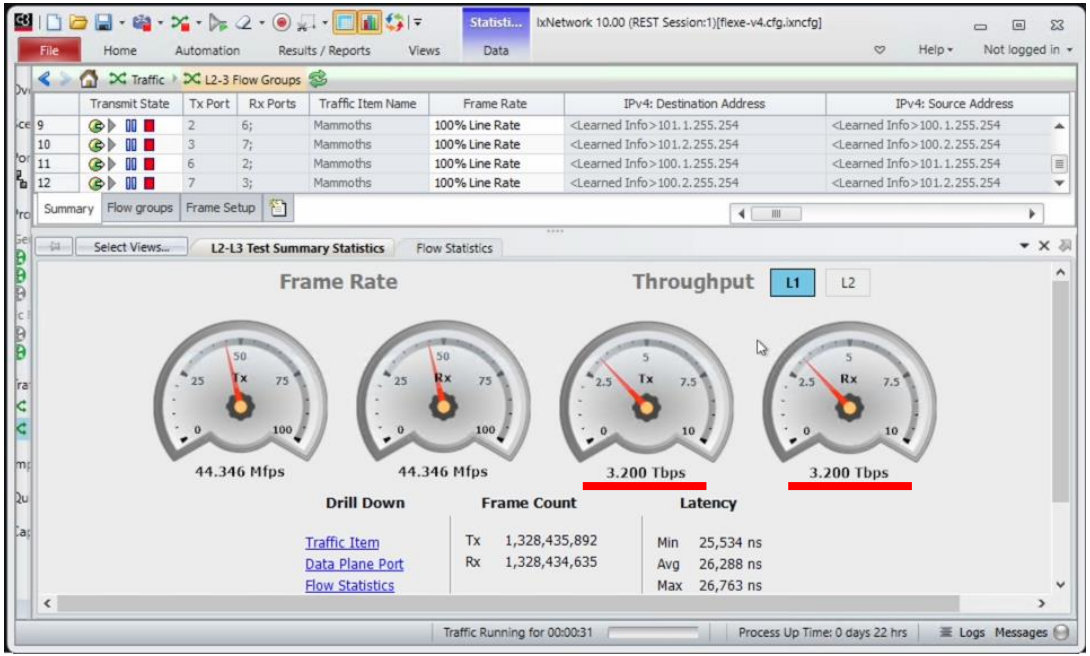


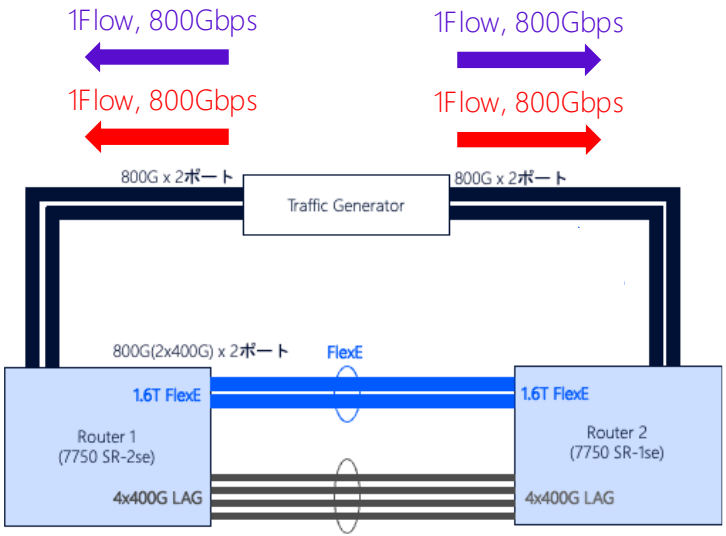
Base Traffic:
800Gbps x 2port x 80% x 2(**双方向**) = 2.56Tbps

+ 150Gbps x 2
+ 150Gbps x 2

= 3.16 Tbps, Drop**無し**

巨大なフローでLinerate流してみる

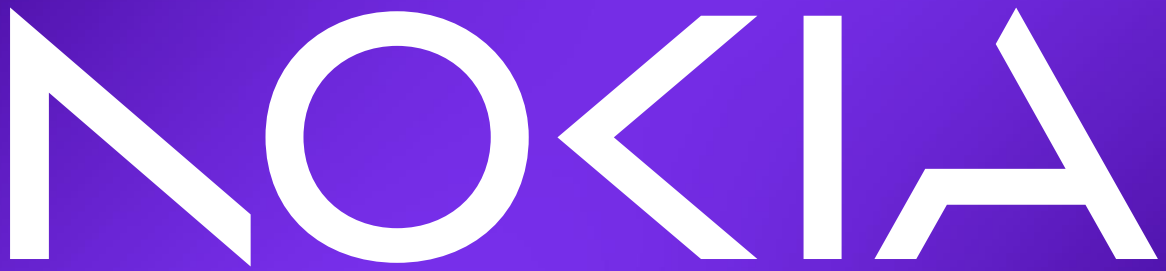




800Gbps x 4 = 3.2 Tbps, Drop**無し**

まとめ

- FlexEで偏りの無いTraffic分散が可能に。Elephant FlowもDrop無く転送可能
- FlexE2.0で400GE/800Gポートでも FlexE使用可能に。
- ZR/ZR+の長距離向けPluggableも使用可能=>WAN回線への適用可


NOKIA