



Mechanisms for Optimal LAG/ECMP Component Link Utilization in Networks

draft-ietf-opsawg-large-flow-load-balancing-04

IETF 87

Ram Krishnan, Sanjay Khanna (Brocade Communications)

Lucy Yong (Huawei)

Anoop Ghanwani (Dell)

Ning So (Tata Communications)

Bhumip Khasnabish (ZTE)



CURRENT STATUS

- **Key changes after WG adoption**
 - Differences between LAG/ECMP
 - Exemplary Information Model

LAG vs ECMP

- Elaborate on the issue of local vs global optimization
- Scheme works well for LAGs and “local” ECMP
 - Links are between the same set of switches
- “Non local ECMP” not as straightforward
 - See draft for example

EXEMPLARY INFORMATION MODEL (1)

- Configuration parameters for Flow Re-balancing
 - Observation interval for large flow
 - Minimum bandwidth threshold of large flow during observation interval
 - Imbalance threshold between component links for load-balancing
 - Rebalancing interval
- System Configuration and Identification Parameters
 - IP address
 - LAG ID
 - Component link ID

EXEMPLARY INFORMATION MODEL (2)

- Large flow recognition and decision to rebalance in router
 - Need configuration parameters as described above
- Large flow recognition and decision to rebalance in external management station
 - Information for Alternative Placement of Large Flows -- Component link ID, can leverage IPFIX for flow identification
 - Information for Redistribution of Small Flows -- LAG ID, Component link ID, % of traffic in each Component link

EXEMPLARY INFORMATION MODEL (3)

- Large flow recognition is being done on a router, but the decision to rebalance is being made in an external management station
 - Flow identification for export -- can leverage IPFIX
- **Monitoring and statistics**
 - Link Utilization
 - Number of times rebalancing was done.
 - Time since the last rebalancing event.



NEXT STEPS

- Ready for WG last call