Link Aggregation and its Applications

Asante
47709 Fremont Blvd
Fremont, CA 94538
USA

SALES
408-435-8388

TECHNICAL SUPPORT
408-435-8388: Worldwide

www.asante.com/support
support@asante.com
Overview

The feature of Link aggregation or Trunking exists in many of the Ethernet switches these days. It is a method of combining similar physical links into one logical link or ‘fat pipe’ with increased bandwidth and reliability. Link Aggregation also provides load balancing where network traffic across the logical link is distributed over the physical link members so there is less possibility that a single link is oversubscribed.

Link Aggregation has the important benefits below:

- Better connection availability in case of single link failure
- Better utilization of physical links to their capacity due to load balancing
- Upgrade to higher performance links can be unnecessary since the ‘fat pipe’ functions as a high performance link itself.

Feature descriptions

Link Aggregation or Trunking is a feature on an Ethernet switch to group multiple physical links into a trunk or ‘fat pipe’. The number of physical links that can be in a trunk is usually limited. The common number is 8. One requirement is that each link in the same trunk must have the same physical characteristics. They must be either all Gigabit(GE) or Fast Ethernet(FE) links and they should be operating in full duplex mode.

Each trunk of a switch can be connected to a trunk of another switch. Connecting 2 trunks means connecting a link member of one trunk to a link member of the other trunk. It is not necessary to connect all trunk members to all trunk members of the other trunk. Therefore it is not necessary for the 2 trunks to have the same number of link members. However in practice, two connecting trunks will have the same number of link members. It is necessary that link members of both trunks have the same link characteristics (GE or FE) and configuration(auto-negotiation and other link setup).
Referring to the diagram above, 3 PC’s, one server and 2 switches are in the network. Switch S1 and S2 have Trunks (Link Aggregation Groups) configured, with each trunk having 4 link members. When a frame is received in S1 and determined to be forwarded to S2, one of the links (1, 2, 3, and 4) is selected to carry the frame to S2. The selection is based on a configured hash algorithm in S1. Normally, the hash is calculated based on Ethernet source and destination addresses. Hash can also be derived based on IP addresses or TCP/UDP ports on some high-end switches. Frames received by S2 that are to be forwarded to S1 will go through the same process. Since links in the trunk must carry traffic in both directions at the same time, it is required that they operate in full duplex mode.
Examples of Link Aggregation

Trunking technology can be used to improve the performance and management of a workgroup network in a number of ways. This section presents some likely scenarios for setup of Link Aggregation in a network backbone.

There scenarios below are common in a network where Link Aggregation is deployed:

- Switch-to-switch
- Server-to-switch
- Server-to-server

Switch-to-Switch

In this scenario, multiple workstations are joined by one aggregated link group (trunk) between switches S1 and S2. By aggregating multiple links, higher speed connections can be achieved without hardware upgrade.

In diagram above, two switches (S1 and S2) are connected using four GE (1000 Mb/s) links. If one link fails between these two switches, the other links in the link aggregation group take over the traffic and the connection is maintained. In the mean time, the traffic between the 2 work groups is distributed over the 4 GE links.

This configuration reduces the number of ports of the switch available for connection to external devices. Therefore there is a trade-off between port usage and additional capacity for a given switch pair.
Switch-to-Server

Most server platforms can saturate a single 100 Mb/s link with many of the applications available today. Thus, link capacity becomes the limiting factor for overall system performance. Consequently, aggregating these links to improve access to applications and file servers is a likely implementation scenario. The diagram above illustrates a connection between a server and a switch, and shows a trunk with four physical links. The switch is connected to nine clients, and to an Ethernet hub that is connected to three additional clients.

Under this configuration, the server could be another type of equipment, such as a router, or high-performance workstation that also supports Link Aggregation.

In Diagram above, one server is connected to a switch using four 100 Mb/s links. In this application, link aggregation is used to improve performance for the link-constrained server. By aggregating multiple links, better performance is achieved without requiring a hardware upgrade to either the server or the switch. Aggregation on the server side can
generally be achieved through software changes in the device driver for the LAN adapter card that has multiple interfaces.

**Server-to-Server**

Upgrading server-to-server links is another likely scenario. Applications such as data warehousing and data distribution demand the frequent movement of large amounts of data between servers, and providing its unimpeded movement is important. Also Link Aggregation connection may be useful for multi-processing or server redundancy applications where high performance is needed to maintain real-time server data coherence. The diagram above illustrates how trunks can be employed between 3 servers to achieve this goal.
The diagram above illustrates a configuration that utilizes 2 trunking capable switches, an internet router, a server and a bunch of workstations. In this configuration, two ports are grouped together using Link Aggregation to provide up to 2 Gbps* full duplex communication between 2 trunking switches. Also a 2 port trunk is also configured between a switch and the server, providing a 2 Gbps full duplex communication path. A router is connected using a 10/100 port, providing access to the wide area network.

**IEEE 802.3ad Link Aggregation standard**

The Link Aggregation (LAG) is made a standard by IEEE referred as 802.3ad. The standard describes the Link Aggregation as follows: “Link Aggregation allows one or more links to be aggregated together to form a Link Aggregation Group, such that a MAC client can treat the Link Aggregation Group as if it were a single link”. The standard also defined LACP, Link Aggregation Control Protocol, for networking devices that intend to support LAG. Any 2 devices supporting LACP can be connected together with multi-links which can form a Link Aggregation group via LACP. Below is a brief summary of 802.3ad LAG:
• LAG is performed above the MAC
• LAG assumes all links are:
  – full duplex
  – point to point
  – same data rate
• Provides graceful recovery from link failures
• Traffic is distributed packet by packet
• All packets associated with a given “conversation” are transmitted on the same link to prevent mis-ordering
• Does not change packet format
  – No added headers or sequence numbers
  – Type/Length interpretation unchanged
• Does not require added buffers
  – No fragmentation or reassembly
• Does not re-order or mis-order packets
• Does not add significant latency
• Does not increase the bandwidth for a single conversation
• Achieves high utilization only when carrying multiple simultaneous conversations

Frame Distributor

The Frame Distributor is the part of the standard that specifies how a link is selected to carry a packet.
This standard does not mandate any particular distribution algorithm(s); however, any distribution algorithm shall ensure that, when frames are received by a Frame Collector, the algorithm shall not cause
a) Mis-ordering of frames that are part of any given conversation, or
b) Duplication of frames.
The above requirement to maintain frame ordering is met by ensuring that all frames that compose a given conversation are transmitted on a single link in the order that they are generated by the MAC Client; hence, this requirement does not involve the addition (or modification) of any information to the MAC frame, nor any buffering or processing on the part of the corresponding Frame Collector in order to re-order frame.

Asante Link Aggregation Support

Asante management switches support LACP and 802.3ad LAG. To configure link aggregation on Ethernet interfaces, one must configure the Ethernet interface, create the LAG (Link Aggregation Group), and add the Ethernet interface as a member link in the LAG. The example below is based on Asante’s IC3648 L2+ management switch.

Configuring an Ethernet Physical Interface
To configure a member link, one must know the configuration of the ports of the LAG of the other switch or server. The port configuration of the port members of the LAGs must be the same. This will insure that the operating speed and duplex will be the same once the ports of the LAGs are connected. Therefore ports of each LAGs must be either configured to be auto-negotiate enabled or disabled. If disabled, the port speed must be configured to be the same and duplex must be configured to be FULL. To configure an LAG on the switch, perform the following steps:

1. Specify a Fast Ethernet or Gigabit Ethernet interface for which you want to create a member link.

   \[\text{switch}(\text{config})\#\text{interface g1}\]

2. Specify the speed and the duplex mode for the Ethernet interface.

   \[\text{switch}(\text{Interface g1})\#\text{speed 100fd}\]

3. To configure additional member links, repeat steps 1 to 2.

   NOTE: All of the member links that you configure must have the same physical layer characteristics, such as speed and duplex mode.

Configure the LAG

There are 6 LAGs (1-6) defined in IC3648. LAG 6 is reserved for Gigabit LAG. To configure a LAG and add member links, perform the following steps:

1. Specify port to add.

   \[\text{switch}(\text{config})\#\text{interface 9}\]

2. Add port to LAG 4.

   \[\text{switch}(\text{Interface 9})\#\text{addport 4}\]

802.3ad Link Aggregation Configuration Example

The following example displays configuration of LACP for IC39480 and a MAC server running OS X.

LACP setup for MacOS 10.4:

1. Log in to the server as an administrative user.

2. Open System Preferences.

3. Click Network.

4. Choose Network Port Configurations from the Show pop-up menu.

5. Click New.

6. Choose Link Aggregate from the Port pop-up menu.

   Note: You'll only see this option if you have two or more Ethernet interfaces on your system.

7. Type the name of the link aggregate in the Name field.

8. Select the ports to aggregate from the list.

9. Click OK.

10. Click Apply Now.

By default the system gives the link aggregate the interface name bond<num>, where <num> is a number indicating precedence. For example, the first link aggregate will be named bond0, the second bond1, and the third bond2.

The interface name bond<num> assigned by the system is different from the name you give to the link aggregate port configuration. The interface name is for use in the command line, while the port configuration name is for use in the Network pane of System Preferences.

**LACP setup procedure for IC39480**

Use the following command lines to setup LAG 2 that has port 1 and 2 as its members. LAG 2 is connected to the MAC server:

```
switch (config)# interface 1
switch (Interface 1)# addport 2
switch (Interface 1)# exit
switch (config)# interface 2
switch (Interface 2)# addport 2
switch (Interface 2)# exit
```
Summary

Today, corporations are feeling the impact as more users and powerful applications threaten to overwhelm network performance. Gigabit Ethernet has been the work horse to provide relief to networks overtaxed by bandwidth-hungry applications and an increased number of users. Now high performance server’s widespread deployment has triggered a need for a corresponding increase in speed in network backbones and to corporate servers.

The Asante implementation of Link Aggregation provides a dramatic increase in network performance. With Asante switches, users have the ability to establish a scaleable ‘fat pipe’ to carry higher an aggregated data rates than any single Gigabit Ethernet or Fast Ethernet link can accommodate, while retaining the capability to seamlessly and safely migrate to ultra high-speed Ethernet in the future. The other benefit is the increased resiliency of the network connection. By using the automatic configuration protocol LACP the Asante switch can provide redundancy with automatic switching to the standby link in case the active link fails.