

Link Aggregation Control Protocol Scenarios

Mick Seaman

This note describes a number of scenarios to illustrate the effect of LACP use in connecting point to point links between two systems.

Terminology

Throughout “connection” means a single physical link connected to one port on each system.

“Initial connection” is used to refer to connections between two ports, neither having participated in LACP previously, or which have been returned to their initial state.

The term “restored connection” refers to connection between two ports on the two systems that were previously connected, and have not been connected to any other systems or ports in the meantime.

When a link is described as being “brought up”, that means that the link is made available for the use of high layer protocols through an aggregate port.

When a port or system is said to have a “unique key” it means that it knows that a given port will never be aggregated with another port because it is the only port on the system with that “key”. The scenarios described assume that that system will indicate the uniqueness by not setting the Aggregate flag in the protocol.

Assumptions

The 802.3 hardware provides proper signaling of link enabled/disabled conditions as detected by presence or loss of light etc.

The aggregate port used for a multiple link aggregate corresponds to the lowest numbered port in that aggregate. If this restriction is not imposed some of the disruption scenarios described here will not occur, on the other hand the user is likely to experience the greatest difficulty in understanding how his systems have come to be in their current states.

1. Initial Connection, Single Link, Both Systems LACP Capable, set Automatic

The link will be brought up with no delay. The user will see no differences between this scenario and one in which LACP is not present.

2. Initial Connection, Single Link, One System LACP Capable, set Desirable

The link will be brought up without delay, and data transfer will proceed as if LACP were not present. The LACP capable system will send LACP messages, first at the fast periodic rate

(once per second) and then at the slow rate (say once per minute).

3. Initial Connection, Single Link, Both Systems LACP Capable, one port (at least) set Desirable, one port (at least) has a Unique Key

The link will be brought up at both ends without delay as if LACP was not present. Then as each system receives LACP messages from the other the link will be brought down and then up again¹.

4. Restored Connection, Single Link, Both Systems LACP Capable, one port (at least) set Desirable, one port (at least) has a Unique Key

After a delay not exceeding two seconds (the time taken for both systems to send LACP messages), the link is brought up at both ends.

5. Initial Connection, Single or Multiple Links, Both Systems LACP Capable, one port (at least) set Desirable

The link will be brought up at both ends without delay as if LACP was not present. Then as each system receives LACP messages from the other the link will be brought down. It will stay down for a brief period (3 to 5 seconds as set by the wait while timer) to allow other connections to join the same aggregate. At the end of this period from the time the last such connection is made, and following a further delay of not more than two seconds, all the links will useable as part of a single aggregate.

¹ A minor change to the selection machine specification would avoid the need to bring the link down and up again. However since this effect will only occur on the very first time the two ports/systems are connected, this is a very minor improvement. The argument for the currently specified behavior is that the aggregate port is guaranteed to go down and come up again each time the identity of the remote partner changes. The specified behavior stops link aggregation from obscuring a physical wiring change that would be detected with an ordinary physical link. Of course the special case of transitioning between no perceived identity and a known partner could be special cased, as could any transition where individual rather than aggregate links are involved. If suppression of the “glitch” described is really wanted, that would be a better solution than running a “hold down” timer to stop the port coming up until information exchange has taken place.

6. Remove Connection from a Multiple Link Aggregate

There is no user disruption at the aggregate port level² unless the physical connection is the last physical connection to be removed. In that case the aggregate port will be brought down. Otherwise frames are distributed to the remaining links.

7. Restore Connection to a Multiple Link Aggregate

After a brief period (3 to 5 seconds as set by the wait while timer) to allow other connections to join the same aggregate, the link is added to the aggregate and following a further delay of no more than two seconds the link is available to carry new or redistributed conversations as part of the aggregate.

If other connections remained in the aggregate throughout the period that this connection was removed and restored, there will have been no disruption at the aggregate port level (a few frames in flight might have been lost).

If all the connections are removed and restored the aggregate port will have been brought down and then up again, however the same aggregate port will be used and there will only be one "down up" transition.

8. Initial Connection added to a Multiple Link Aggregate

If this is an initial connection for both ends of the link, the link will be brought up without delay, then as each system receives LACP messages it will be brought down, and after a brief period added to the existing aggregate.

If the connection added is not the lowest numbered port for either system, there will be no disruption at the user aggregate port level. If it is the existing aggregate port will go down and the new lower numbered aggregate port will be brought up a few seconds later.

9. Reused Connection added to a different Multiple Link Aggregate

After a brief waiting period the connection will be added to the new aggregate. There will be no disruption to the new aggregate in either system unless the added link has the lowest numbered port in that system.

If the port in either system was previously the lowest numbered port in a different aggregate and this is the first time that port has been reused, that aggregate will be brought down

² There may be disruption at a later time if the removed connection is reused in a different aggregate. See below.