

SocketTimeout

Socket Timeouts

TCP Socket Timeouts are caused when a TCP socket times out talking to the far end.

Socket timeouts can occur when attempting to connect to a remote server, or during communication, especially long-lived ones.

They can be caused by any connectivity problem on the network, such as:

- A network partition preventing the two machines from communicating.
- The remote machine crashing. This cannot easily be distinguished from a network partitioning.
- The remote machine is a VM and it has been destroyed. Again, to the calling application, this looks like a connectivity problem.
- A change in the firewall settings of one of the machines preventing communication.
- The settings are wrong and the client is trying to talk to the wrong machine, one that is not on the network. That could be an error in Hadoop configuration files, or an entry in the DNS tables or the `/etc/hosts` file.
- If its over a long-haul network (i.e. out of cluster), it may be a transient failure due to the network playing up.
- If you have some proxy settings, they may be wrong.
- If using a client of an object store such as the Amazon S3 and [OpenStack](#) clients, socket timeouts may be caused by remote-throttling of client requests: your program is making too many PUT/DELETE requests and is being deliberately blocked by the far end. This is most likely to happen when creating many small files, or performing bulk deletes (e.g. deleting a directory with many child entries). It can also arise from a transient failure of the long-haul link.

Comparing this exception to the [ConnectionRefused](#) error, the latter indicates there is a server at the far end, but no program running on it can receive inbound connections on the chosen port. A Socket Timeout usually means that there is something there, but it or the network are not working right

Identifying and Fixing Socket Timeouts

The root cause of a Socket Timeout is a connectivity failure between the machines, so try the usual process

1. Check the settings: is this the machine you really wanted to talk to?
2. From the machine that is raising the exception, can you resolve the hostname.
3. Is that resolved hostname the correct one?
4. Can you ping the remote host?
5. Is the target machine running the relevant Hadoop processes?
6. Can you telnet to the target host and port?
7. Can you telnet to the target host and port from any other machine?
8. On the target machine, can you telnet to the port using localhost as the hostname. If this works but external network connections time out, it's usually a firewall issue.
9. If it is a remote object store: is the address correct? Does it go away when you repeat the operation? Does it only happen on bulk operations? If the latter, it's probably due to throttling at the far end.

Remember: These are [your network configuration problems](#) . Only you can fix them.