Conntrack-tools: connection tracking userspace tools

Pablo Neira Ayuso
<pablo@netfilter.org> Netfilter Project

conntrack-tools:status

- Two programs (blame me the very similar for the name!):
  - conntrack: command line interface, it is a replacement of /proc/net/nf_conntrack.
  - conntrackd: daemon to handle state synchronization (to enable highly available stateful firewalls).
- Last release: 0.9.7. 31th May, 2008.
- Next release: next week :)
  - New features and improvements
  - Tons of bugfixes
  - New documentation: the conntrack-tools HOW-TO (http://dune.lsi.us.es/~pablo/conntrack-tools.html)

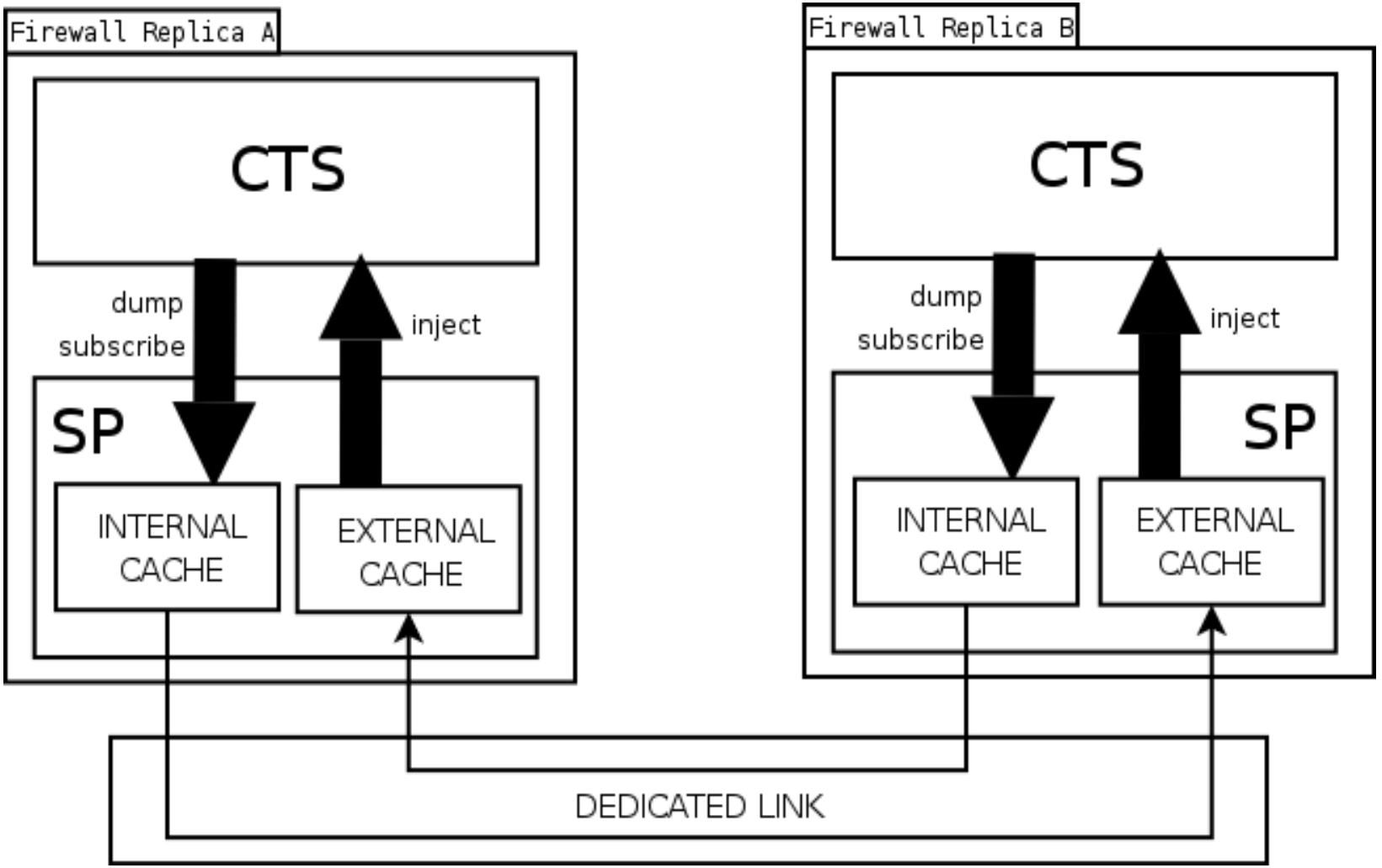
conntrack CLI

- Similar syntax than iptables (so it's not that user friendly). Once we have nftables, it would be nice to use a similar syntax.
- Features:
  - List the current state table (in /proc-compatible and XML)
  - Filtering options: No need to use grep.
  - Flexible updates and deletion.
  - Event listening.

conntrackd: the daemon

- The daemon can be used as:
  - a simple flow-based statistics collector.
  - state table synchronizer: it propagates asynchronously the state-changes between stateful firewall replicas to achieve high availability.
    - Currently only Primary-Backup supported, but the architecture allows multiprimary.
    - TLV-based message format: around 76-100 bytes per message.
    - Message batching.

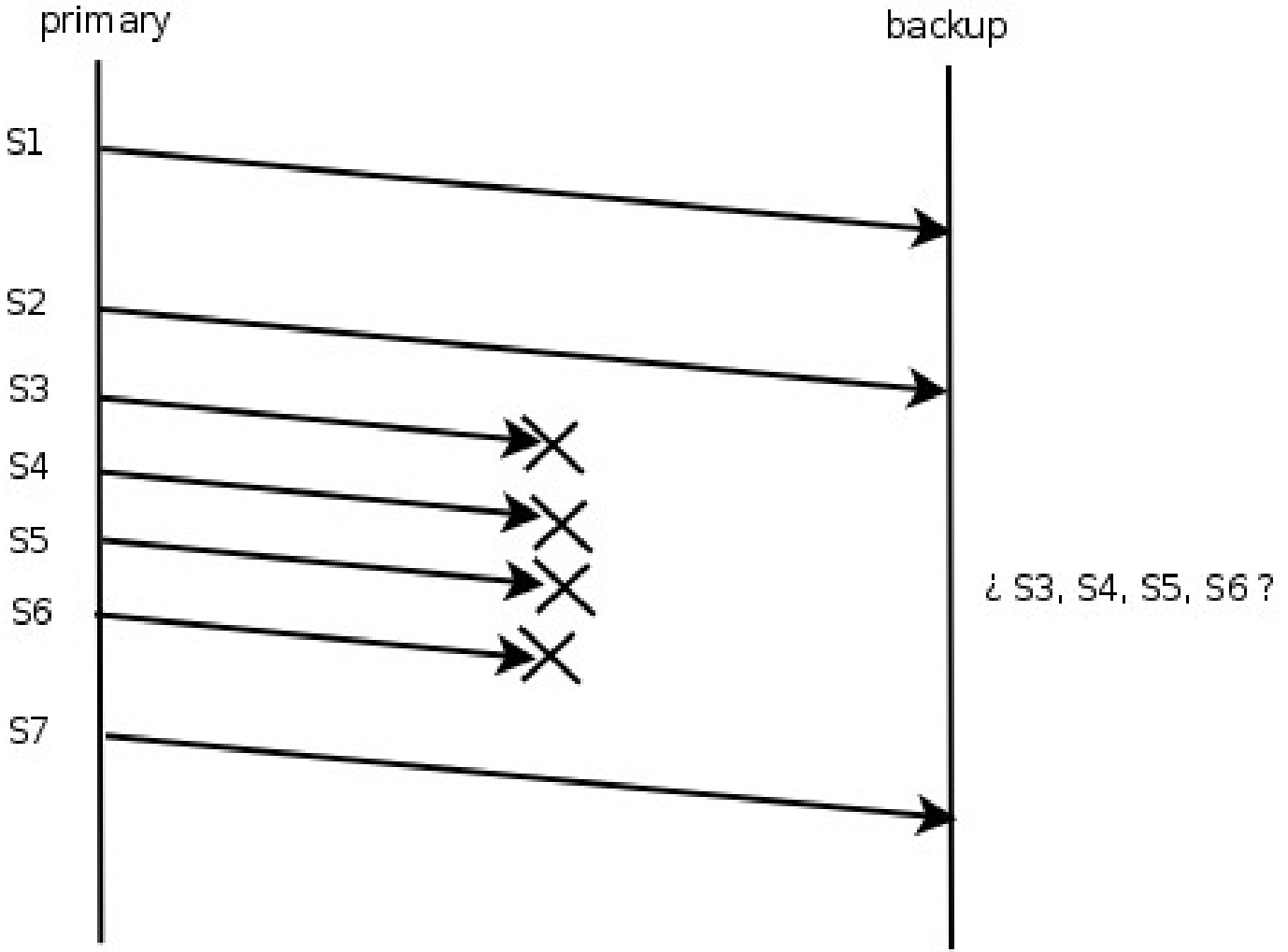
conntrackd: design



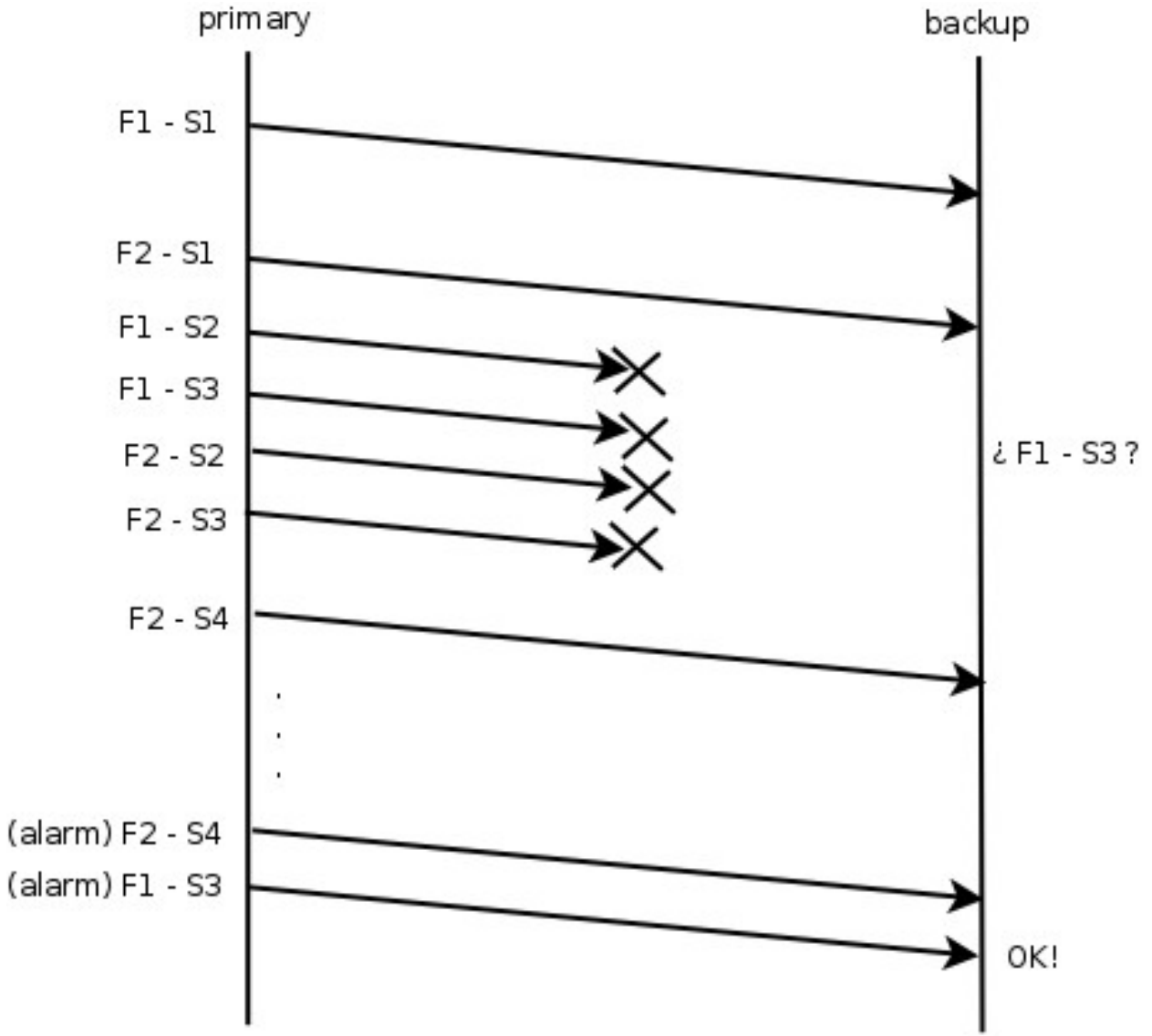
Replication protocols

- Asynchronous replication based on multicast. Three approaches – as for now:
  - **NOTRACK**: like pfsync, a best effort protocol, no sequence tracking at all.
  - **ALARM**: Every N seconds a state message is sent (spamming but resolve well unconsistencies between nodes).
  - **FT-FW:** Reliable protocol (with sequence tracking). Still experimental but a lot of progress on the way.

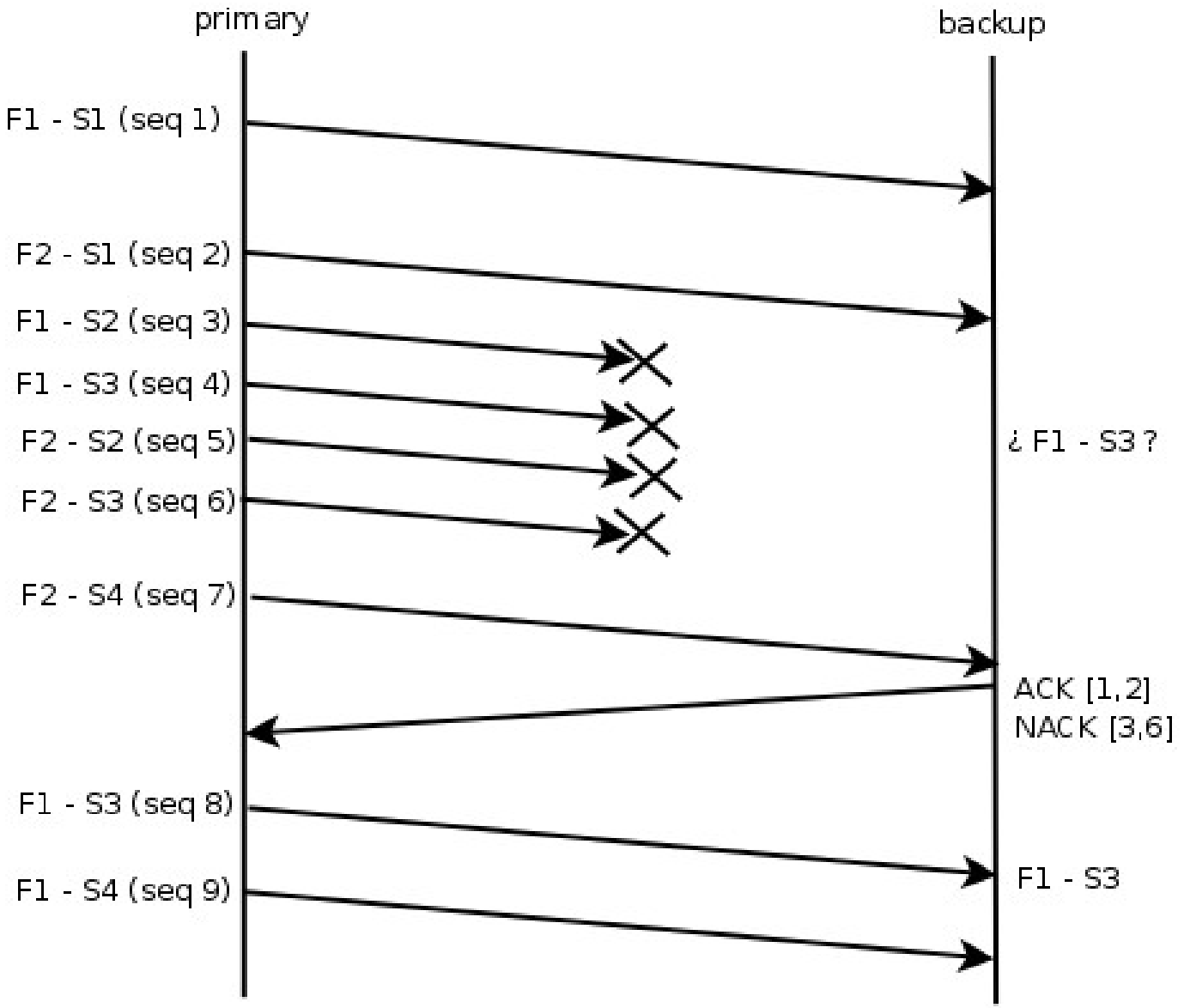
NOTRACK



ALARM


primary
backup
F1 - S1
F2 - S1
F1 - S2
F1 - S3
F2 - S2
F2 - S3
¿ F1 - S3 ?
F2 - S4
(alarm) F2 - S4
(alarm) F1 - S3
OK!

FT-FW


primary
backup
F1 - S1 (seq 1)
F2 - S1 (seq 2)
F1 - S2 (seq 3)
F1 - S3 (seq 4)
F2 - S2 (seq 5)
F2 - S3 (seq 6)
F2 - S4 (seq 7)
¿ F1 - S3 ?
ACK [1,2]
NACK [3,6]
F1 - S3 (seq 8)
F1 - S4 (seq 9)
F1 - S3

Changes: from previous NFWS

- What was pending and it has been done this year:
  - Support for netfilter kernel-space filtering based on BSF since 2.6.25.
    - You may only replicate only TCP Established state to reduce the CPU consumption (a normal HTTP connection requires 6 messages).
  - No need to disable TCP window tracking since 2.6.22.
  - IPv6 support (needs more testing)
  - Natively support for related conntracks
  - NAT sequence adjustment
  - Improved netlink overrun handling
  - Documentation :-)

TODO

- What needs to be done:
  - Redundant dedicated link
  - **Multiprimary support (hash-based load sharing)**, some target similar to CLUSTERIP with changes.
    - Logic: If a packet arrives.
      - Should I handle this? hash(src) % localnum
    - Using a hub:
      - Use the same virtual MAC for all nodes: VRRP MAC.
    - Using a switch:
      - Use multicast link layer (RFC1122 violation), use arptables.
  - TIPC