libnftables updates

NFWS 2017 Faro, Portugal

Pablo Neira Ayuso <pablo@netfilter.org>

Bootstrapping...

- High level library for nftables.
- main.c and cli.c under nftables
  - All remaining bits go to libnftables.
- Eric Leblond hacked on nft to bootstrap library in NFWS'16
  - 34 patches on top of nft 0.5
  - Available at https://github.com/regit/nftables
  - Proposed workshop:
    - 1) nft_ctx_t typedef to carry all context.
    - 2) Pass nft syntax to library call.
    - 3) Library call says OK/NOK.
  - Simple and minimal number of exposed API
    - So we move on more freely...

Cons

- nftables existing design is layered...

```
main()
  nft_run()
    nft_parse()
    nft_netlink()
      do_command()
        netlink_add_*()
```

- Netlink details are hiding in several layer on code.

Next spin...

- Revisiting library workflow:
  - Fully expose netlink socket file descriptor.
    - User have control on select()/send()/recv().
  - Rework existing layered design, split it in chunks:
    - 1) Parse nft syntax → generates list of command structure
    - 2) Command structure → generate netlink bytecode
    - 3) Send netlink bytecode to kernelspace
    - 4) Pass netlink error message and list of commands to display errors.
  - Monitor/trace mode:
    - 1) Read netlink message
      - Probably convert this to command structures?
    - 2) nft syntax output

Preparation changes

- Wraps all context information in **struct nft_ctx** or wherever this belongs to...
  - debug_level
  - void *scanner
  - struct parser_state
  - struct list_head msgs (error messages)
  - struct eval_ctx, to keep evaluation
  - struct output_ctx, to keep how things are printed
  - struct netlink_ctx, for netlink information
  - Move max_errors to struct parser_state

Preparation changes (2)

- Daemons need cache updates
  - Tear down and repopulate cache is expensive
  - Subscribe to event notification and refresh cache
- Rtnetlink is used in a number of spots
- Move all __init functions to nft_init()
  - Before doing so remove as many __init as possible.
  - Can we skip nft_init() function?

Preparation changes (3)

- Memory leaks: We have a good bunch of them
  - Run valgrind, fix them.