IIJ SA-W1 and some functions.

Masanobu SAITOH

msaitoh@netbsd.org

1. Introduction
2. Ethernet Bridge Framework
3. netipsec+opencrypto
4. Other additions and changes

This document will be put somewhere after finishing NetBSD developer summit.

Bugfixes and improvements

- I wrote as follows in my mail:

> Except the following functions, almost all fixes and enhancements “for Marvell SoC" have been merged to -current:
>
> - L2 cache support
> - bug fixes, add functions and enhancement for mvicesa (crypto accelerator)

- For other than Marvell SoC?

Not merged yet.

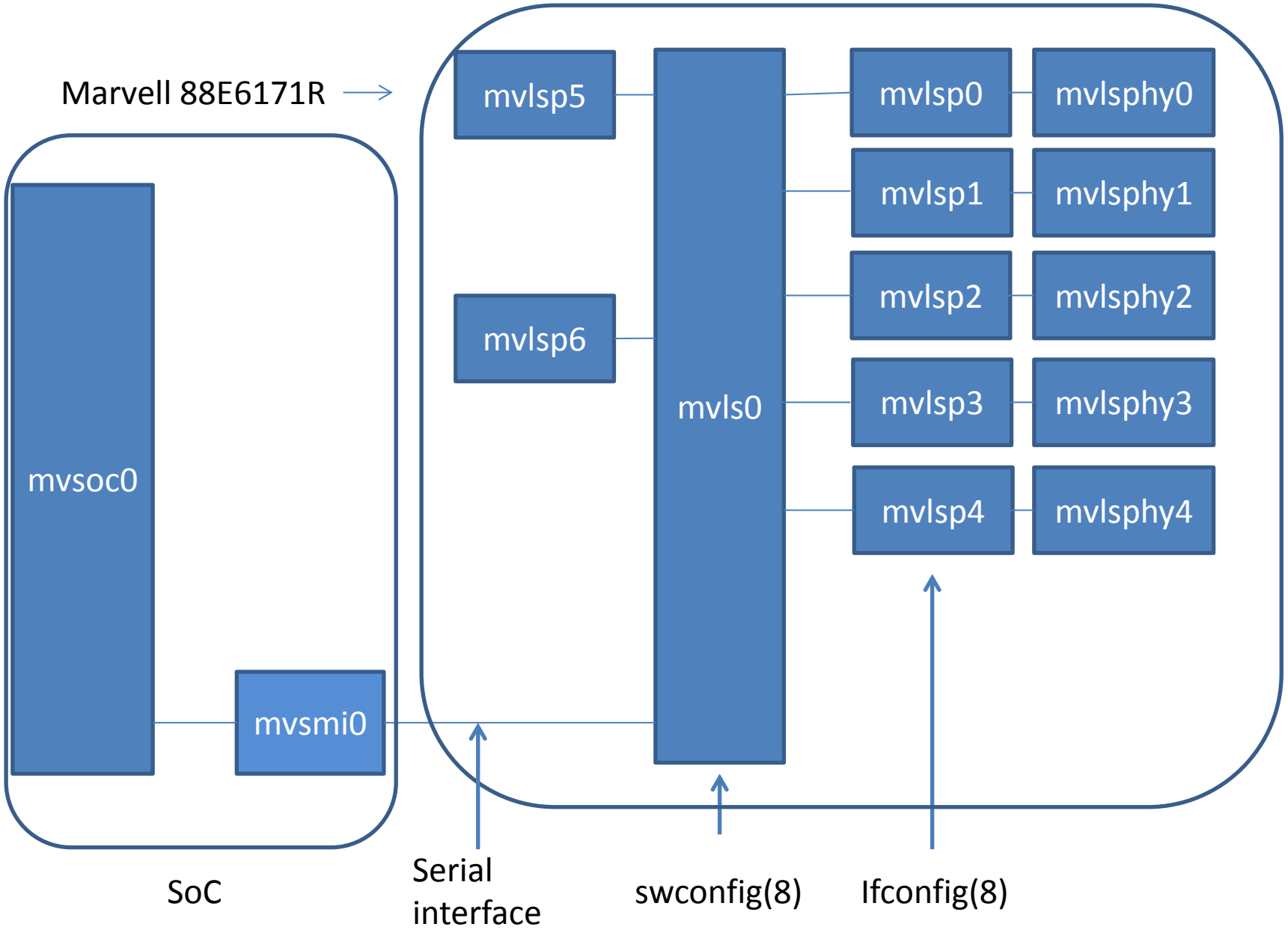
- New functions
  - Framework for Ethernet Switch
- Important fixes
  - netipsec+opencrypto
- A lot of not serious bugfixes and improvements

Ethernet switch framework

Ethernet switch framework

- Designed and implemented by Hikaru Abe.
  - Almost all commits for mvgbe(4) by me were originally worked by him (not me).
- Control switch function using swconfig(8)
- Control media setting using ifconfig(8)

device tree on SA-W1



Ethernet switch drivers on SA-W1(dmesg)

mvsmi0 at mvsoc0 unit 0 offset 0x72004-0x72007: Serial Management Interface
mvls0 at mvsmi0 addr 0-31 gpio 11 irq 107 single-chip rev 2: Marvell Gigabit Ethernet Switch
mvlsp0 at mvls0 port 0: Marvell Gigabit Ethernet Switch External Port
mvlsphy0 at mvlsp0 phy 0: Marvell 88E6171 Gigabit Switch PHY, rev. 0
mvlsphy0: 10baseT, 10baseT-FDX, 100baseTX, 100baseTX-FDX, 1000baseT-FDX, auto
mvlsp1 at mvls0 port 1: Marvell Gigabit Ethernet Switch External Port
mvlsphy1 at mvlsp1 phy 1: Marvell 88E6171 Gigabit Switch PHY, rev. 0
mvlsphy1: 10baseT, 10baseT-FDX, 100baseTX, 100baseTX-FDX, 1000baseT-FDX, auto
mvlsp2 at mvls0 port 2: Marvell Gigabit Ethernet Switch External Port
mvlsphy2 at mvlsp2 phy 2: Marvell 88E6171 Gigabit Switch PHY, rev. 0
mvlsphy2: 10baseT, 10baseT-FDX, 100baseTX, 100baseTX-FDX, 1000baseT-FDX, auto
mvlsp3 at mvls0 port 3: Marvell Gigabit Ethernet Switch External Port
mvlsphy3 at mvlsp3 phy 3: Marvell 88E6171 Gigabit Switch PHY, rev. 0
mvlsphy3: 10baseT, 10baseT-FDX, 100baseTX, 100baseTX-FDX, 1000baseT-FDX, auto
mvlsp4 at mvls0 port 4: Marvell Gigabit Ethernet Switch External Port
mvlsphy4 at mvlsp4 phy 4: Marvell 88E6171 Gigabit Switch PHY, rev. 0
mvlsphy4: 10baseT, 10baseT-FDX, 100baseTX, 100baseTX-FDX, 1000baseT-FDX, auto
mvlsp5 at mvls0 port 5: Marvell Gigabit Ethernet Switch Internal Port
mvlsp6 at mvls0 port 6: Marvell Gigabit Ethernet Switch Internal Port

swconfig(8)

```
# swconfig
usage:  swconfig <dev> group <groupid> [member '<port>...']
        swconfig <dev> -group <groupid>

        swconfig <dev> portfdb <fdbid> [member '<port>...']
        swconfig <dev> -portfdb <fdbid>

        swconfig <dev> vlan <vlanid> [member '<port>[<(u)ntag,(t)ag>]...']
        swconfig <dev> -vlan <vlanid>
        swconfig <dev> defaultvlan <port> <vlanid>

        swconfig <dev> mirror-rx <dstport> '<srcport>...'
        swconfig <dev> -mirror-rx
        swconfig <dev> mirror-tx <dstport> '<srcport>...'
        swconfig <dev> -mirror-tx

        swconfig <dev> nolearning|-nolearning <port>
        swconfig <dev> notagged|-notagged <port>
        swconfig <dev> nountagged|-nountagged <port>

        swconfig <dev> flushfdb <fdbid>
        swconfig <dev> showfdb <fdbid>
```

Broadcast domain

Forwarding DB

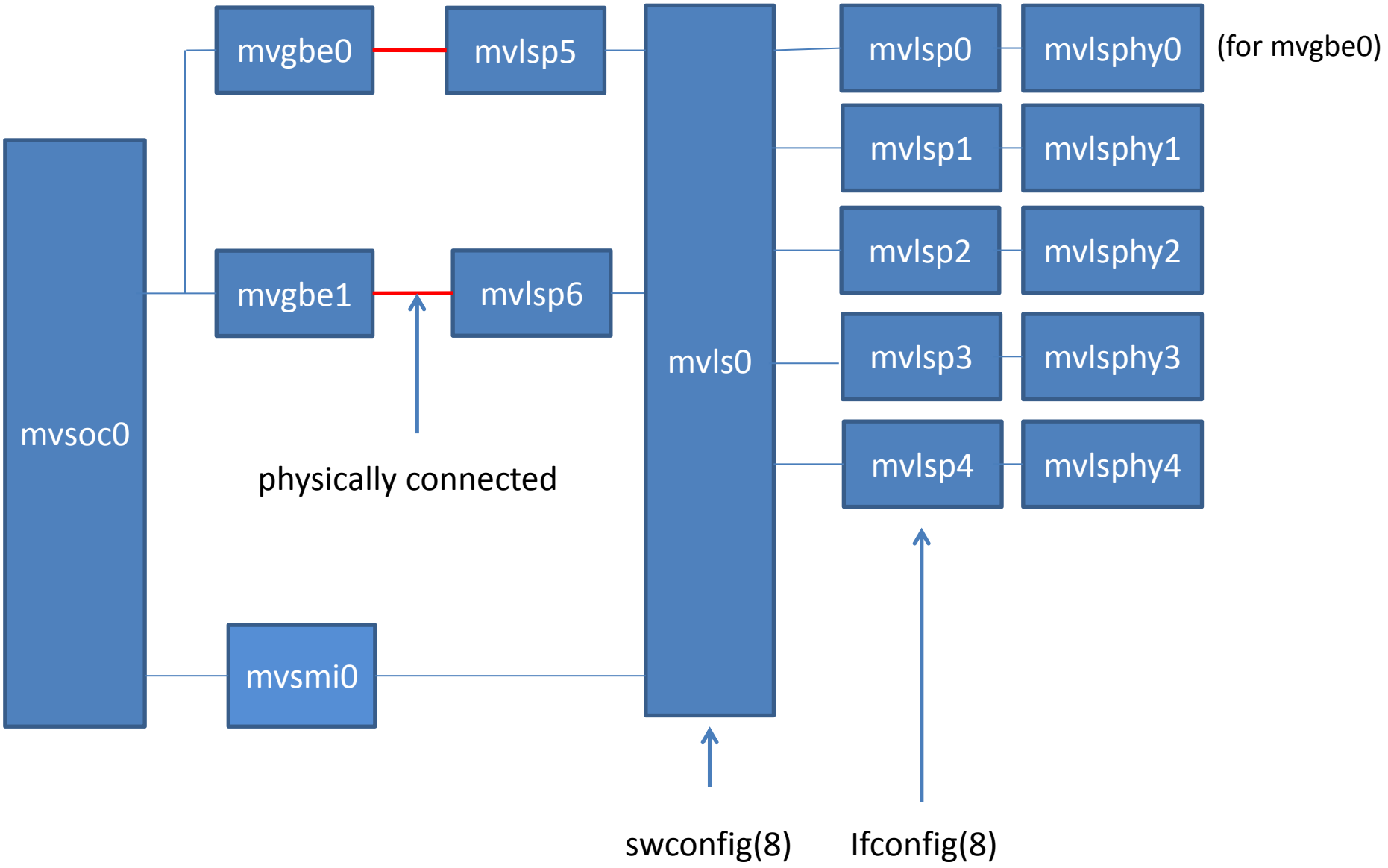
vlan

mirroring

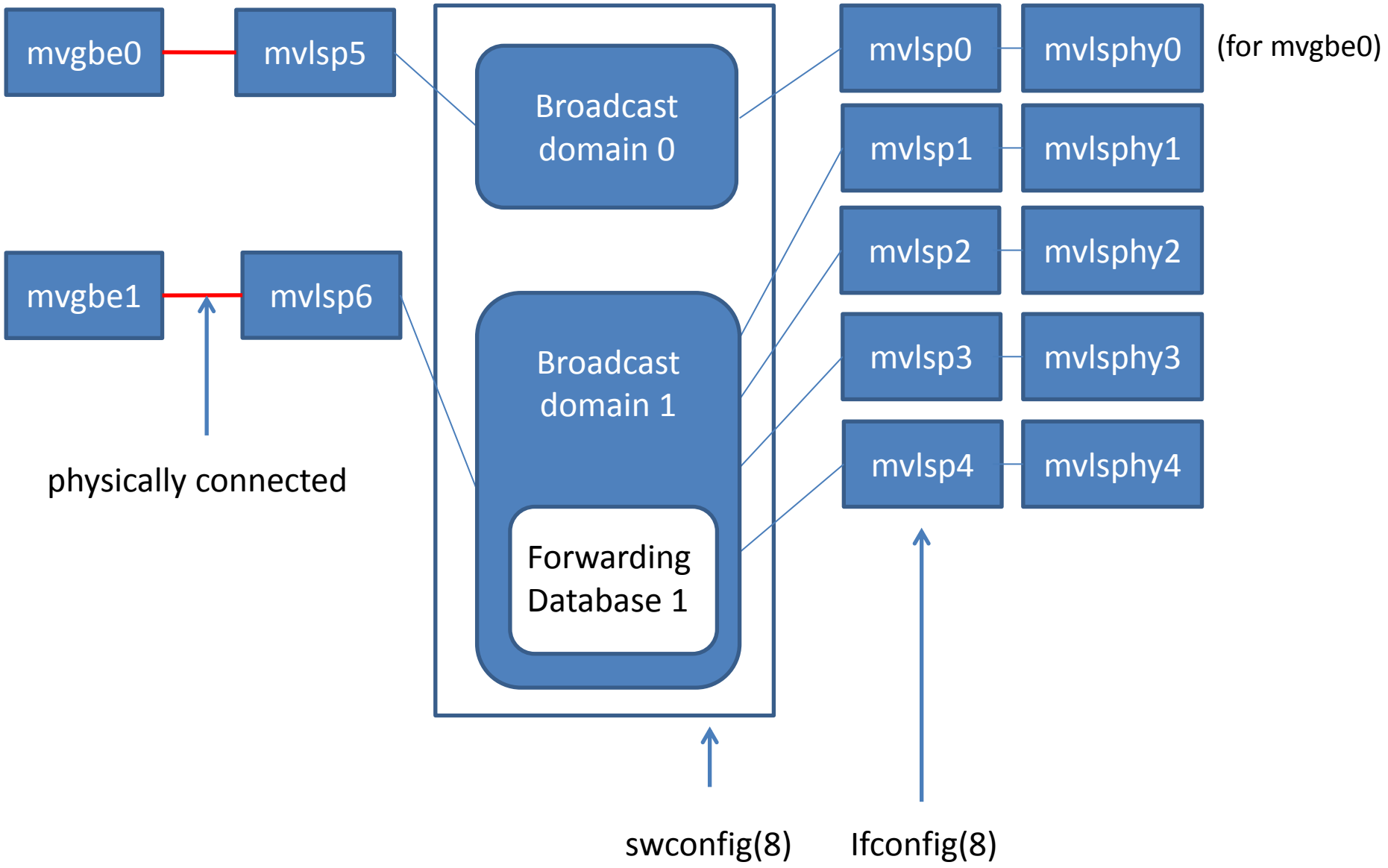
control

admin

Relation between mvgbe and switch



Relation between mvgbe and switch



Settings

```
# cat ifconfig.mvls0
!swconfig $int group 0 member '0 5' nolearning 0 nolearning 5
!swconfig $int group 1 member '1 2 3 4 6'
!swconfig $int portfdb 1 member '1 2 3 4 6’
Up
```

ifconfig -a

```
$ ifconfig -a
mvls0: flags=41<UP,RUNNING> mtu 1500
mvlsp0: flags=41<UP,RUNNING> mtu 1500
        media: Ethernet autoselect (none)
        status: no carrier
mvlsp1: flags=41<UP,RUNNING> mtu 1500
        media: Ethernet autoselect (none)
        status: no carrier
mvlsp2: flags=41<UP,RUNNING> mtu 1500
        media: Ethernet autoselect (none)
        status: no carrier
mvlsp3: flags=41<UP,RUNNING> mtu 1500
        media: Ethernet autoselect (none)
        status: no carrier
mvlsp4: flags=41<UP,RUNNING> mtu 1500
        media: Ethernet autoselect (none)
        status: no carrier
mvlsp5: flags=41<UP,RUNNING> mtu 1500
        media: Ethernet manual (none)
mvlsp6: flags=41<UP,RUNNING> mtu 1500
        media: Ethernet manual (none)
mvgbe0:
flags=8843<UP,BROADCAST,RUNNING,SIMPLEX
,MULTICAST> mtu 1500

capabilities=3700<IP4CSUM_Rx,IP4CSUM_Tx,T
CP4CSUM_Rx,UDP4CSUM_Rx,UDP4CSUM_Tx>
        enabled=0
        ec_capabilities=1<VLAN_MTU>
        ec_enabled=0
        address: 00:e0:4d:ff:03:54
        media: Ethernet manual (none)
        status: no carrier
        inet6 fe80::2e0:4dff:feff:354%mvgbe0
prefixlen 64 scopeid 0x9
```

Future work?

- mii_attach() problem
  - Almost all Ethernet drivers use mii_attach().
  - Ethernet switch may be connected using RGMII.
- What is the best way to configure(8) Ethernet switch?

netipsec+opencrypto

problem

- How many people who are using netipsec with HW acceleration?
  - –Availability of HW acceleration chip itself.
  - –Availability of document of the HW acceleration chip.

Security Acceleration Engine

- Past:
  - Expensive
  - Difficult to get both device and document.
- Now:
  - Inexpensive if an SoC has it in it.
  - Easy to get
    - e.g. SheevaPlug

netipsec + mvcesa(4)

- Nice platform to develop netipsec with HW acceleration.
  - Cheap
  - Easy to get
- A lot of bugs:
  - in netipsec itself
  - in mvcesa(4) itself:
    - e.g. kiyohara@(the first author of the driver) said “The driver should not work ☺”

bugfixes

- Bugfixes and enhancements were done by Hiroki Suenaga and Hikaru Abe
- All functions using SA-W1 firmware work well.
- Not all functions are implemented.
  - Some functions that SA-W1 doesn't use are not implemented yet.
  - Because of the problem, an ATF test fails ☺

merging

- Those changes will be merged into –current
  - Who?
    - I can’t merge them because I’m not familiar with IPsec ☺
  - When?

Other changes

Other changes

- A lot of changes still have not been merged into –current because:
  - No time ☹
    - e.g. MEXTADD and M_EXT_RW flag problem.
  - Need discussion to merge some changes
    - See next page
  - but some of them are important

Changes which need discussion to merge

- A lot IPv6 related new functions and enhancements
- iipf, iipfnat
- npppd+pipex (OpenBSD merged this function)
- Source address selection
- Some 64bit counters (e.g. 10G network)
- netstat changes
- OpenBSD's MCLGETI
- BPF_DIRECTION_{IN,OUT}
- Routing based IPsec (ipsecif)
- L2TPv3
- mbuf utility functions
- rtadvd(8) enhancements
- ETHERCAP related changes
- EXCLUDE_1000HDX option
- Auto MDI/MDI-X
- Some DDB enhancements
- USB bugfixes
- 3G, LTE USB modem control daemon

Conclusion

- IIJ has a lot of code to improve NetBSD.
- IIJ will agree to merge those changes into NetBSD.
  - maybe not all.
- How should we merge them?
  - Some of them need discussion about the implementation before merging.
  - Japanese (communication problem)