Mellanox watchdog drivers

for x86 based system switches

This driver provides watchdog functionality for various Mellanox Ethernet and Infiniband switch systems.

Mellanox watchdog device is implemented in a programmable logic device.

There are 2 types of HW watchdog implementations.

Type 1:
: Actual HW timeout can be defined as a power of 2 msec. e.g. timeout 20 sec will be rounded up to 32768 msec. The maximum timeout period is 32 sec (32768 msec.), Get time-left isn't supported

Type 2:
: Actual HW timeout is defined in sec. and it's the same as a user-defined timeout. Maximum timeout is 255 sec. Get time-left is supported.

Type 3:
: Same as Type 2 with extended maximum timeout period. Maximum timeout is 65535 sec.

Type 1 HW watchdog implementation exist in old systems and all new systems have type 2 HW watchdog. Two types of HW implementation have also different register map.

Type 3 HW watchdog implementation can exist on all Mellanox systems with new programmer logic device. It's differentiated by WD capability bit. Old systems still have only one main watchdog.

Mellanox system can have 2 watchdogs: main and auxiliary. Main and auxiliary watchdog devices can be enabled together on the same system. There are several actions that can be defined in the watchdog: system reset, start fans on full speed and increase register counter. The last 2 actions are performed without a system reset. Actions without reset are provided for auxiliary watchdog device, which is optional. Watchdog can be started during a probe, in this case it will be pinged by watchdog core before watchdog device will be opened by user space application. Watchdog can be initialised in nowayout way, i.e. oncse started it can't be stopped.

This mlx-wdt driver supports both HW watchdog implementations.

Watchdog driver is probed from the common mlx_platform driver. Mlx_platform driver provides an appropriate set of registers for Mellanox watchdog device, identity name (mlx-wdt-main or mlx-wdt-aux), initial timeout, performed action in expiration and configuration flags. watchdog configuration flags: nowayout and start_at_boot, hw watchdog version - type1 or type2. The driver checks during initialization if the previous system reset was done by the watchdog. If yes, it makes a notification about this event.

Access to HW registers is performed through a generic regmap interface. Programmable logic device registers have little-endian order.