

APM X-Gene SoC Performance Monitoring Unit (PMU)

X-Gene SoC PMU consists of various independent system device PMUs such as L3 cache(s), I/O bridge(s), memory controller bridge(s) and memory controller(s). These PMU devices are loosely architected to follow the same model as the PMU for ARM cores. The PMUs share the same top level interrupt and status CSR region.

PMU (perf) driver

The xgene-pmu driver registers several perf PMU drivers. Each of the perf driver provides description of its available events and configuration options in sysfs, see `/sys/devices/<BcX/iobX/mcbX/mcX>/`.

The "format" directory describes format of the config (event ID), config1 (agent ID) fields of the perf_event_attr structure. The "events" directory provides configuration templates for all supported event types that can be used with perf tool. For example, "Bc0/bank-fifo-full" is an equivalent of "Bc0/config=0x0b".

Most of the SoC PMU has a specific list of agent ID used for monitoring performance of a specific datapath. For example, agents of a L3 cache can be a specific CPU or an I/O bridge. Each PMU has a set of 2 registers capable of masking the agents from which the request come from. If the bit with the bit number corresponding to the agent is set, the event is counted only if it is caused by a request from that agent. Each agent ID bit is inversely mapped to a corresponding bit in "config1" field. By default, the event will be counted for all agent requests (config1 = 0x0). For all the supported agents of each PMU, please refer to APM X-Gene User Manual.

Each perf driver also provides a "cpumask" sysfs attribute, which contains a single CPU ID of the processor which will be used to handle all the PMU events.

Example for perf tool use:

```
/ # perf list | grep -e l3c -e iob -e mcb -e mc
    l3c0/ackq-full/                                [Kernel PMU event]
<...>
    mcb1/mcb-csw-stall/                            [Kernel PMU event]

/ # perf stat -a -e l3c0/read-miss/,mcb1/csw-write-request/ sleep 1

/ # perf stat -a -e l3c0/read-miss,config1=0xfffffffffffffffe/ sleep 1
```

The driver does not support sampling, therefore "perf record" will not work. Per-task (without "-a") perf sessions are not supported.