Suricata - Task #3695

research: libhwloc for better autoconfiguration

05/11/2020 12:13 PM - Victor Julien

Status: Assigned
Priority: Normal
Assignee: Shivani Bhardwaj
Category:
Target version: 7.0beta1
Effort:
Difficulty:

Description

https://www.open-mpi.org/projects/hwloc/

hwloc-ls gives us a nice view into the system. What the NUMA nodes are, which devices are connected to each node. Also what the cpu id's are for the nodes.

Example output:

```bash
$ hwloc-ls
Machine (63GB total)

NUMANode L#0 (P#0 31GB)
  Package L#0 + L3 L#0 (30MB)
    L2 L#0 (256KB) + L1d L#0 (32KB) + L1i L#0 (32KB) + Core L#0
      PU L#0 (P#0)
      PU L#1 (P#24)
    ...
    L2 L#11 (256KB) + L1d L#11 (32KB) + L1i L#11 (32KB) + Core L#11
      PU L#22 (P#11)
      PU L#23 (P#35)
  HostBridge L#0
    PCIBridge
      PCI 1000:0086
      Block(Disk) L#0 "sda"
    PCIBridge
      PCI 19ee:4000
      Net L#1 "ens1np0"
      Net L#2 "ens1np1"
    PCIBridge
      PCI 8086:1d6b
      PCI 8086:1502
      Net L#3 "eno1"
    PCIBridge
      PCI 8086:10d3
      Net L#4 "enp1s0"
    PCIBridge
      PCI 10de:128b
      GPU L#5 "renderD128"
      GPU L#6 "controlD64"
      GPU L#7 "card0"
    PCI 8086:2826
NUMANode L#1 (P#1 31GB)
  Package L#1 + L3 L#1 (30MB)
    L2 L#12 (256KB) + L1d L#12 (32KB) + L1i L#12 (32KB) + Core L#12
      PU L#24 (P#12)
      PU L#25 (P#36)
    ...
    L2 L#23 (256KB) + L1d L#23 (32KB) + L1i L#23 (32KB) + Core L#23
```
There are 4 NICs in this machine: 2 Dual port Netronome cards (ens3np* on NUMA node 1, ens1np* on node 0. Built-in NICs enp1s0 and eno1 also on node 0).

We could use this info in properly setting up CPU affinity for Suricata.

I'm assuming that libhwloc exposes this info in way that Suricata would use it.

Goals:
- review hwloc availability and versions for our 'tier 1' and 'tier 2' supported OS', distros.
- create a PoC where configure detects and enables libhwloc and prints the NUMA node for the interface Suricata intends to use (single iface is ok for the PoC)
- determine if the lib is suitable for the autoconfig goal

Bigger picture:
- idea is to allow an option to suri like --numa-from-nic (name TBD) that would take the numa node for the nic, then set cpu affinity and thread counts to only use that numa node.
- in multi-nic capture, setup threads incl affinity according to numa config
- if possible, detect and warn on misconfiguration by numactl (e.g. nic is on numa node 0, threads are forced on node 1)
- simplify manual configuration. E.g. instead of cpu: [ 0, 2, 4, 6, 8, 16, 18, 20, 22 ] something like numa: [ 0 ]

History
#1 - 05/12/2020 08:58 AM - Victor Julien
- Status changed from New to Assigned
- Assignee set to Shivani Bhardwaj
- Target version set to 6.0.0beta1

#2 - 06/01/2020 02:29 PM - Victor Julien
- Target version changed from 6.0.0beta1 to 7.0beta1