Run Xenomai as ACRN hypervisor Real-Time VM OS

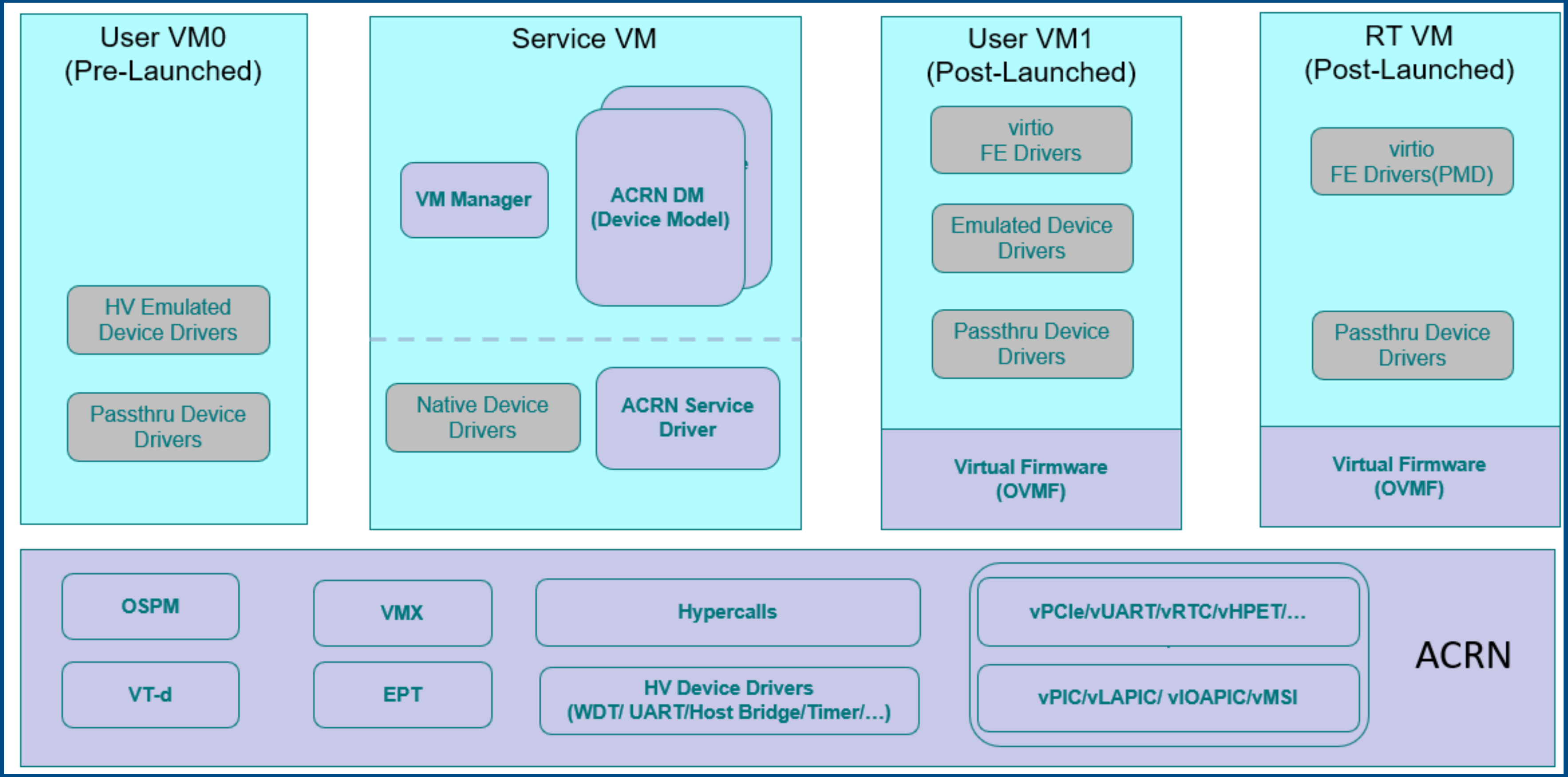
ACRN Introduction

ACRN™ is a flexible, open-source, lightweight hypervisor - intended to enable consolidation of heterogeneous workloads, and to streamline IoT edge development.

- A Linux Foundation Project
- Launched in March 2018
- Version 1.0 released in May 2019

License: BSD-3-Clause. The ACRN hypervisor and ACRN Device Model software are provided under the permissive BSD-3-Clause license, which allows "redistribution and use in source and binary forms, with or without modification" together with the intact copyright notice and disclaimers noted in the license.

ACRN Architecture



Since ACRN v3.0

ACRN 3.2

- Enabling New Generation Intel® Processors
- Hypervisor-Managed Processor Performance Policy Controls
- New Debianization Solution for ACRN
- Service VM Upgraded to use Ubuntu 22.04

- Passthrough PMU (performance monitor unit) to user VM only in debug builds
- Added tarfile member sanitization to Python tarfile package extractall() calls
- Run executables with absolute paths in board inspector

- Refined shutdown & reset sequence
- Hypervisor Real Time Clock (RTC)

- Redesigned ACRN Configuration and more ACRN Configuration Improvements
- Improved Board Inspector Collection and Reporting
- Sample Application with Two Post-Launched VMs
- Multiple-Displays Support for VMs
- Improved TSC Frequency Reporting

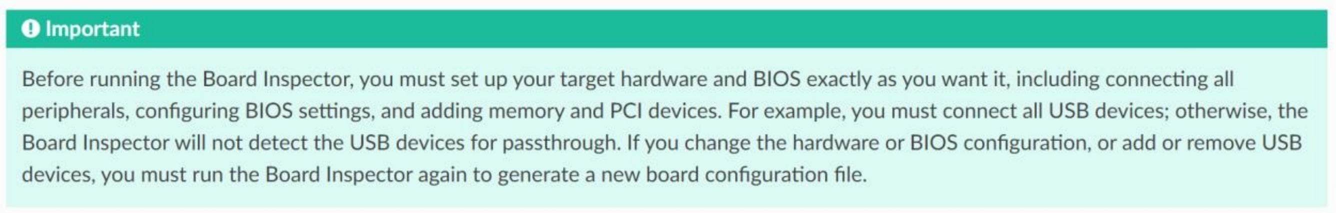
- Mitigation for Return Stack Buffer Underflow security vulnerability
- ACRN shell commands added for real-time performance profiling

ACRN 3.0

Improvement of Board Inspector Collection and Reporting

Steps

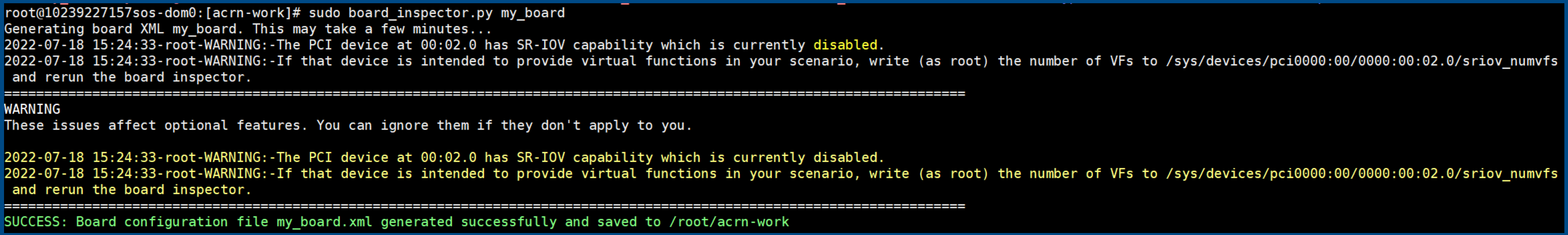
- Set Up the Target Hardware
- Install OS on the Target Board
- Configure Target BIOS Settings
- Generate a Board Configuration File



Important

Before running the Board Inspector, you must set up your target hardware and BIOS exactly as you want it, including connecting all peripherals, configuring BIOS settings, and adding memory and PCI devices. For example, you must connect all USB devices; otherwise, the Board Inspector will not detect the USB devices for passthrough. If you change the hardware or BIOS configuration, or add or remove USB devices, you must run the Board Inspector again to generate a new board configuration file.

Example output for generating board configuration file



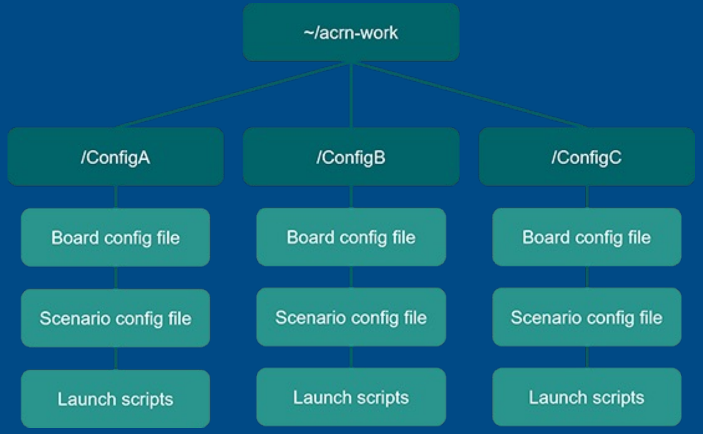
```
root@10239227157sos-dom0:[acrn-work]# sudo board_inspector.py my_board
Generating board XML my_board. This may take a few minutes...
2022-07-18 15:24:33-root-WARNING:-The PCI device at 00:02.0 has SR-IOV capability which is currently disabled.
2022-07-18 15:24:33-root-WARNING:-If that device is intended to provide virtual functions in your scenario, write (as root) the number of VFs to /sys/devices/pci0000:00/0000:00:02.0/sriov_numvfs
 and rerun the board inspector.
========================================================================================================================
WARNING
These issues affect optional features. You can ignore them if they don't apply to you.

2022-07-18 15:24:33-root-WARNING:-The PCI device at 00:02.0 has SR-IOV capability which is currently disabled.
2022-07-18 15:24:33-root-WARNING:-If that device is intended to provide virtual functions in your scenario, write (as root) the number of VFs to /sys/devices/pci0000:00/0000:00:02.0/sriov_numvfs
 and rerun the board inspector.
========================================================================================================================
SUCCESS: Board configuration file my_board.xml generated successfully and saved to /root/acrn-work
```

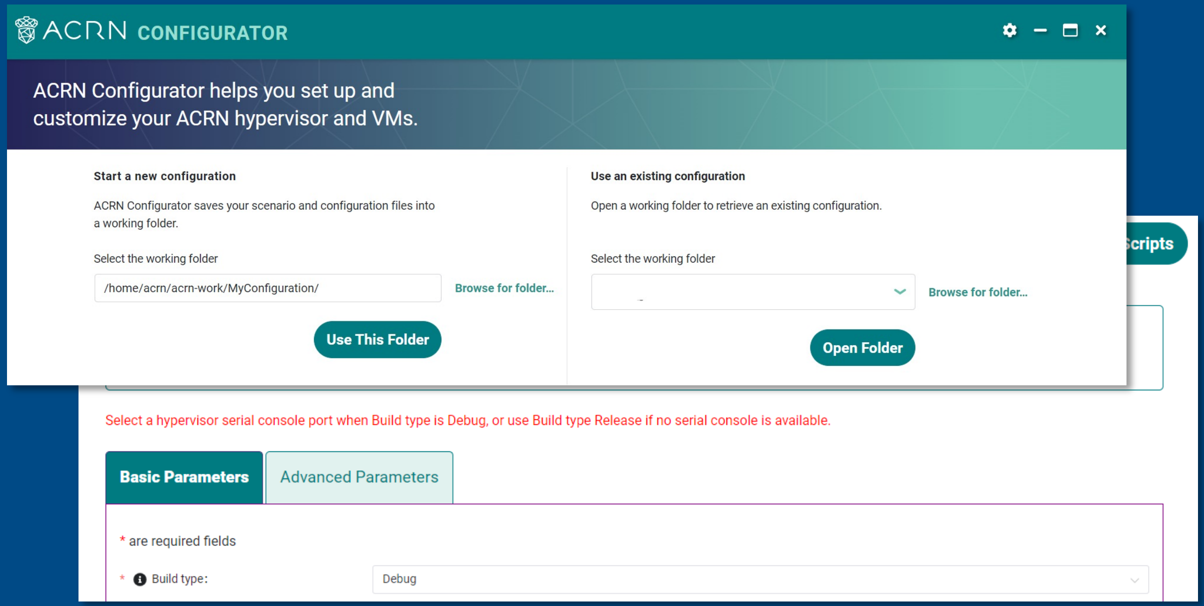
New ACRN Configurator

Generate Scenario Configuration File and Launch Script

ACRN Configurator

- Reads board information from the board configuration file generated by the Board Inspector Tool
- Helps you configure a scenario of hypervisor and VM settings
- Generates a scenario configuration file that stores the configured settings in XML format
- Generates a launch script for each post-launched User VM





ACRN Configurator helps you set up and customize your ACRN hypervisor and VMs.

Start a new configuration

ACRN Configurator saves your scenario and configuration files into a working folder.

Select the working folder

/home/acrn/acrn-work/MyConfiguration/

Use an existing configuration

Open a working folder to retrieve an existing configuration.

Select the working folder

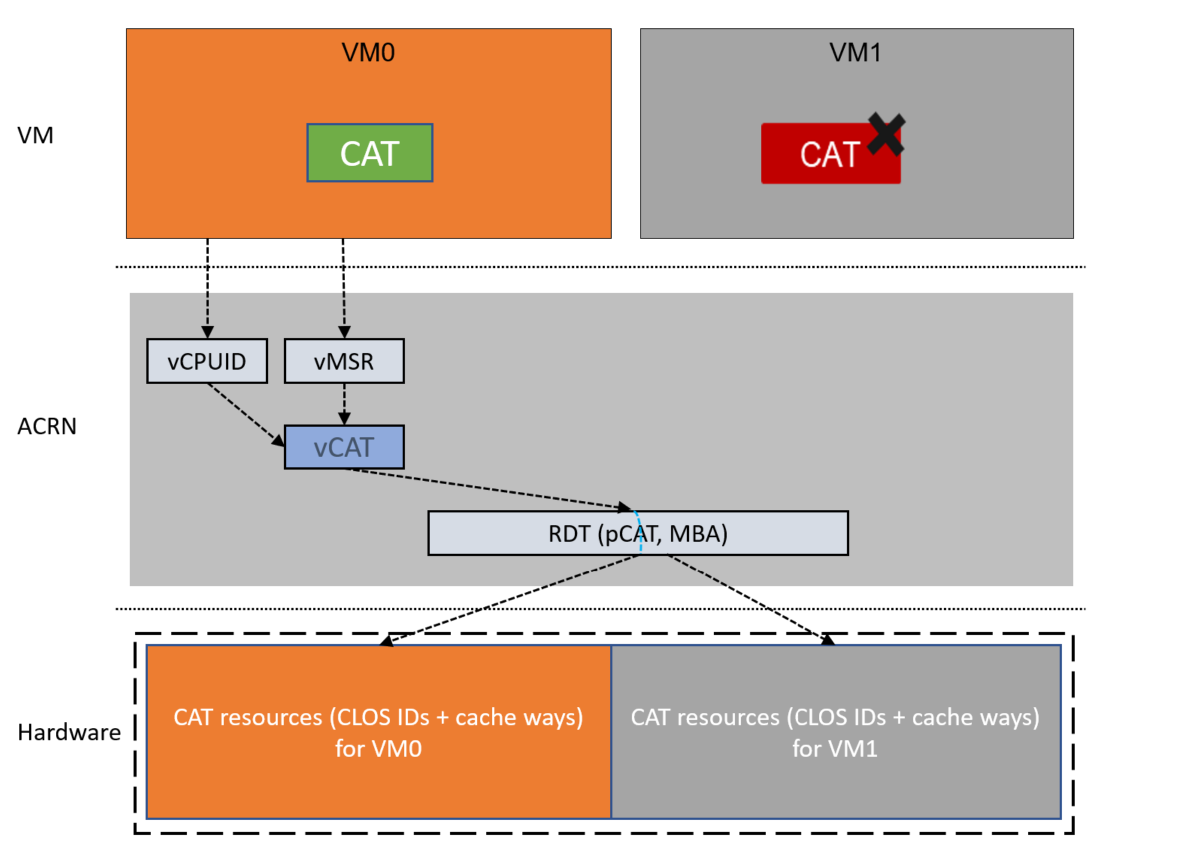
Select a hypervisor serial console port when Build type is Debug, or use Build type Release if no serial console is available.

* are required fields

* Build type: Debug

vCAT

ACRN vCAT exposes some number of virtual CAT resources to VMs and then transparently map them to the assigned physical CAT resources in the ACRN hypervisor; VM can take advantage of vCAT to prioritize and partition virtual cache ways for its own tasks.


VM
VM0
CAT
VM1
CAT
ACRN
vCPUID
vMSR
vCAT
RDT (pCAT, MBA)
Hardware
CAT resources (CLOS IDs + cache ways) for VM0
CAT resources (CLOS IDs + cache ways) for VM1

ACRN shell commands added for real-time performance profiling

VM Exit Profiling

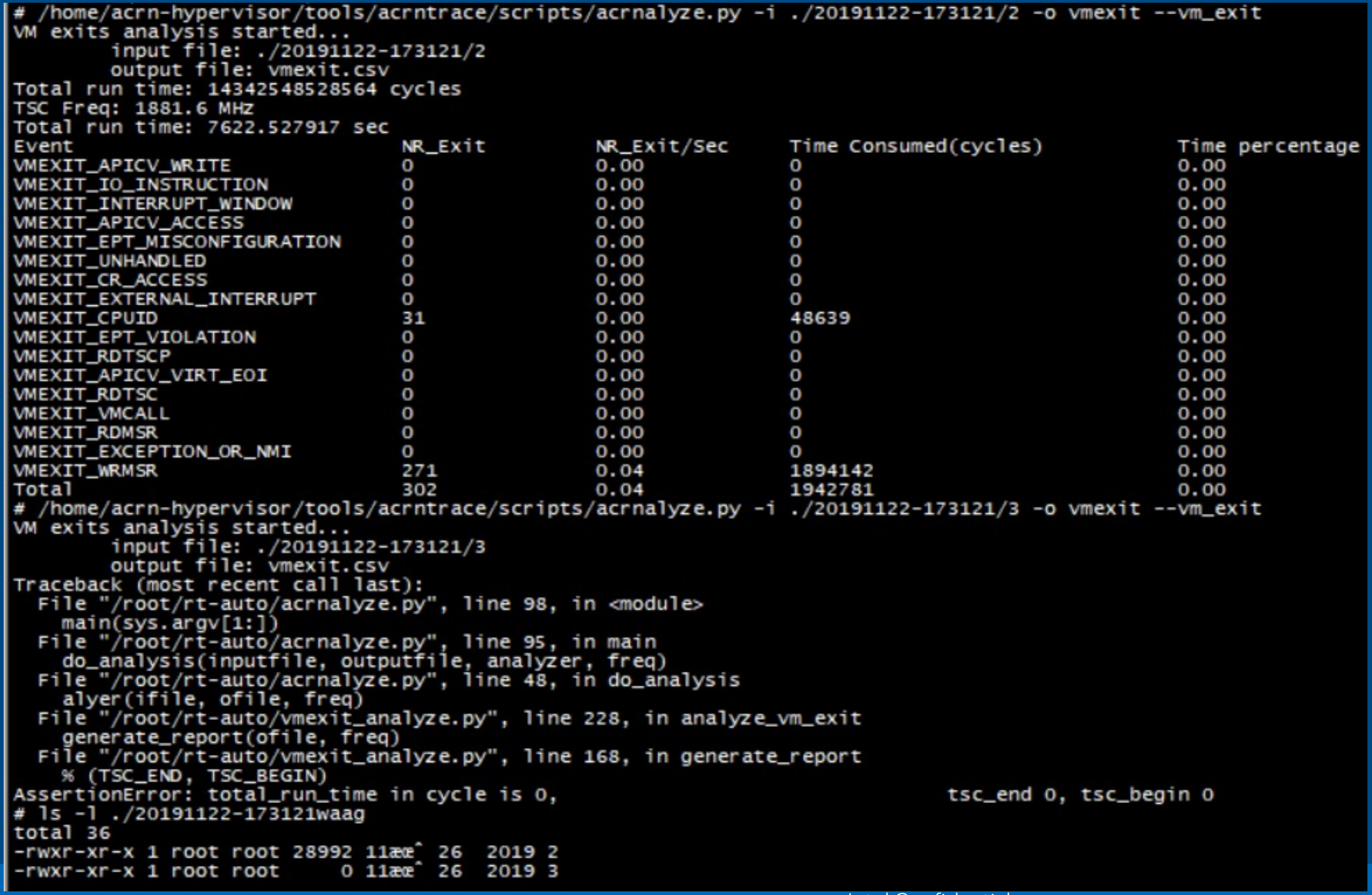
A VM exit marks the point at which a transition is made between the VM currently running and the VMM (hypervisor) which must exercise system control for a particular reason. It's a key source of performance degradation in a virtualized systems. We supported 2 ways of VM exit profiling in ACRN

- via acrntrace
- via VM Exit command in HV console

ACRN shell commands were added to sample vmexit data per virtual CPU to facilitate real-time performance profiling

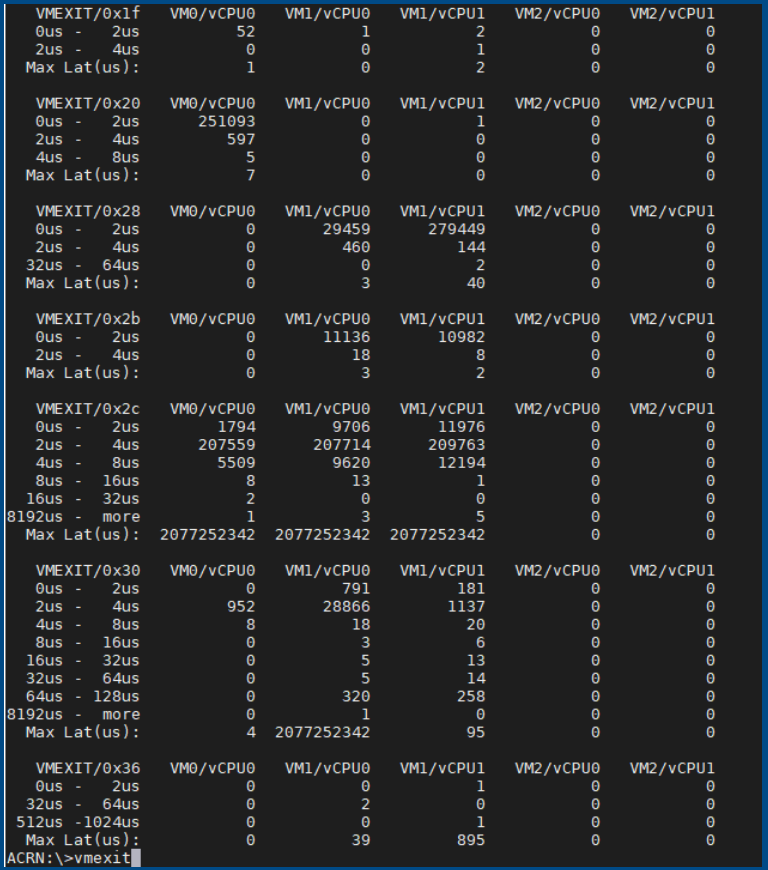
- vmexit enable | disable : enabled by default
- vmexit clear : clears current vmexit buffer
- vmexit [vm_id] : outputs vmexit reason code and latency count information per vCPU for a VM ID (or for all VM IDs if none is specified).

VMEXIT data profiling I



```
# /home/acrn-hypervisor/tools/acrntrace/scripts/acrnalyze.py -i ./20191122-173121/2 -o vmexit --vm_exit
VM exits analysis started...
        input file: ./20191122-173121/2
        output file: vmexit.csv
Total run time: 14342548528564 cycles
TSC Freq: 1881.6 MHz
Total run time: 7622.527917 sec
Event                           NR_Exit         NR_Exit/Sec     Time Consumed(cycles)           Time percentage
VMEXIT_APICV_WRITE              0               0.00            0                               0.00
VMEXIT_IO_INSTRUCTION           0               0.00            0                               0.00
VMEXIT_INTERRUPT_WINDOW         0               0.00            0                               0.00
VMEXIT_APICV_ACCESS             0               0.00            0                               0.00
VMEXIT_EPT_MISCONFIGURATION     0               0.00            0                               0.00
VMEXIT_UNHANDLED                0               0.00            0                               0.00
VMEXIT_CR_ACCESS                0               0.00            0                               0.00
VMEXIT_EXTERNAL_INTERRUPT       0               0.00            0                               0.00
VMEXIT_CPUID                    31              0.00            48639                           0.00
VMEXIT_EPT_VIOLATION            0               0.00            0                               0.00
VMEXIT_RDTSCP                   0               0.00            0                               0.00
VMEXIT_APICV_VIRT_EOI           0               0.00            0                               0.00
VMEXIT_RDTSC                    0               0.00            0                               0.00
VMEXIT_VMCALL                   0               0.00            0                               0.00
VMEXIT_RDMSR                    0               0.00            0                               0.00
VMEXIT_EXCEPTION_OR_NMI         0               0.00            0                               0.00
VMEXIT_WRMSR                    271             0.04            1894142                         0.00
Total                           302             0.04            1942781                         0.00
# /home/acrn-hypervisor/tools/acrntrace/scripts/acrnalyze.py -i ./20191122-173121/3 -o vmexit --vm_exit
VM exits analysis started...
        input file: ./20191122-173121/3
        output file: vmexit.csv
Traceback (most recent call last):
  File "/root/rt-auto/acrnalyze.py", line 98, in <module>
    main(sys.argv[1:])
  File "/root/rt-auto/acrnalyze.py", line 95, in main
    do_analysis(inputfile, outputfile, analyzer, freq)
  File "/root/rt-auto/acrnalyze.py", line 48, in do_analysis
    alyer(ifile, ofile, freq)
  File "/root/rt-auto/vmexit_analyze.py", line 228, in analyze_vm_exit
    generate_report(ofile, freq)
  File "/root/rt-auto/vmexit_analyze.py", line 168, in generate_report
    % (TSC_END, TSC_BEGIN)
AssertionError: total_run_time in cycle is 0,                               tsc_end 0, tsc_begin 0
# ls -l ./20191122-173121waag
total 36
-rwxr-xr-x 1 root root 28992 11æœˆ 26  2019 2
-rwxr-xr-x 1 root root     0 11æœˆ 26  2019 3
```

VMEXIT data profiling II

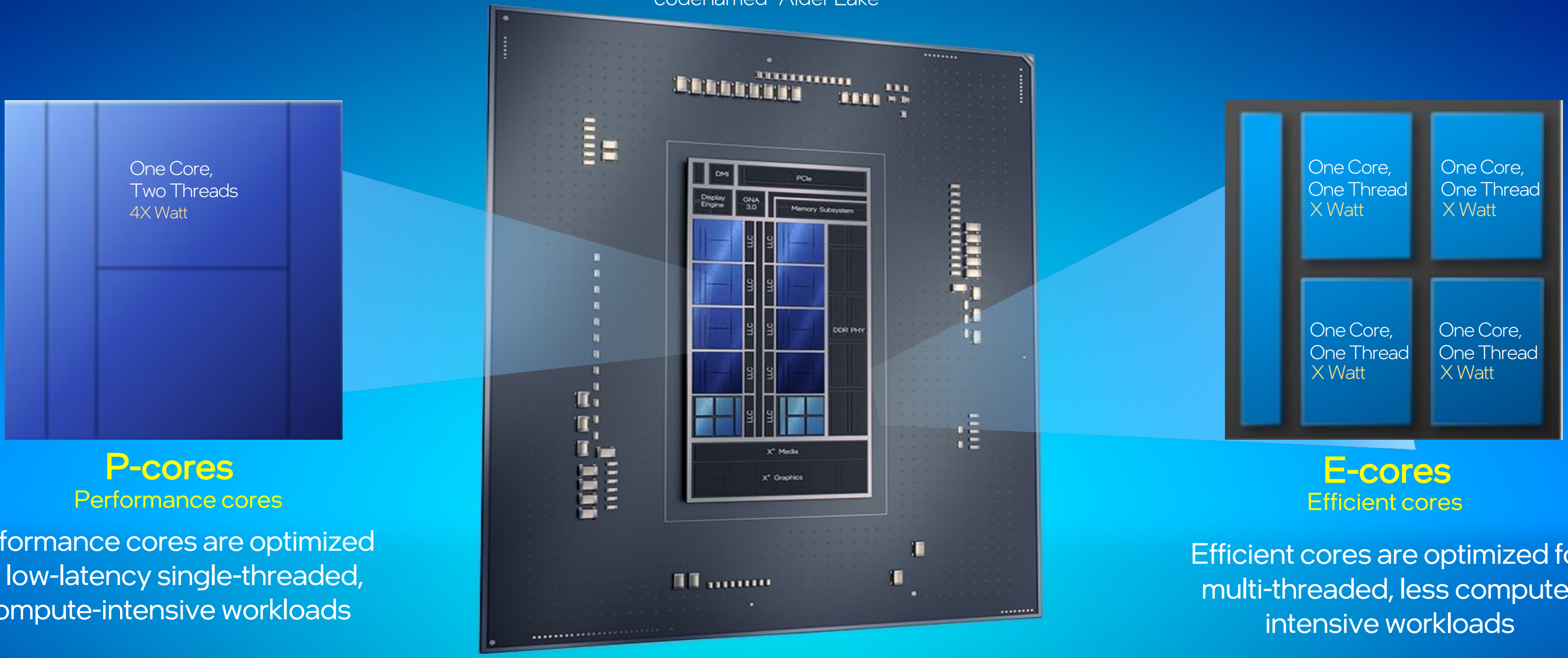


```
   VMEXIT/0x1f   VM0/vCPU0   VM1/vCPU0   VM1/vCPU1   VM2/vCPU0   VM2/vCPU1
   0us -   2us          52           1           2           0           0
   2us -   4us           0           0           1           0           0
  Max Lat(us):           1           0           2           0           0

   VMEXIT/0x20   VM0/vCPU0   VM1/vCPU0   VM1/vCPU1   VM2/vCPU0   VM2/vCPU1
   0us -   2us      251093           0           1           0           0
   2us -   4us         597           0           0           0           0
   4us -   8us           5           0           0           0           0
  Max Lat(us):           7           0           0           0           0

   VMEXIT/0x28   VM0/vCPU0   VM1/vCPU0   VM1/vCPU1   VM2/vCPU0   VM2/vCPU1
   0us -   2us           0       29459      279449           0           0
   2us -   4us           0         460         144           0           0
  32us -  64us           0           0           2           0           0
  Max Lat(us):           0           3          40           0           0

   VMEXIT/0x2b   VM0/vCPU0   VM1/vCPU0   VM1/vCPU1   VM2/vCPU0   VM2/vCPU1
   0us -   2us           0       11136       10982           0           0
   2us -   4us           0          18           8           0           0
  Max Lat(us):           0           3           2           0           0

   VMEXIT/0x2c   VM0/vCPU0   VM1/vCPU0   VM1/vCPU1   VM2/vCPU0   VM2/vCPU1
   0us -   2us        1794        9706       11976           0           0
   2us -   4us      207559      207714      209763           0           0
   4us -   8us        5509        9620       12194           0           0
   8us -  16us           8          13           1           0           0
  16us -  32us           2           0           0           0           0
8192us -  more           1           3           5           0           0
  Max Lat(us):  2077252342  2077252342  2077252342           0           0

   VMEXIT/0x30   VM0/vCPU0   VM1/vCPU0   VM1/vCPU1   VM2/vCPU0   VM2/vCPU1
   0us -   2us           0         791         181           0           0
   2us -   4us         952       28866        1137           0           0
   4us -   8us           8          18          20           0           0
   8us -  16us           0           3           6           0           0
  16us -  32us           0           5          13           0           0
  32us -  64us           0           5          14           0           0
  64us - 128us           0         320         258           0           0
8192us -  more           0           1           0           0           0
  Max Lat(us):           4  2077252342          95           0           0

   VMEXIT/0x36   VM0/vCPU0   VM1/vCPU0   VM1/vCPU1   VM2/vCPU0   VM2/vCPU1
   0us -   2us           0           0           1           0           0
  32us -  64us           0           2           0           0           0
 512us -1024us           0           0           1           0           0
  Max Lat(us):           0          39         895           0           0
ACRN:\>vmexit
```

Introducing Intel® Hybrid Technology

Available today on select 12th Gen Intel® Core™ processors

12th Gen Intel® Core™ processor
codenamed "Alder Lake"



P-cores

Performance cores

Performance cores are optimized for low-latency single-threaded, compute-intensive workloads

P-cores provide heavy lifting power for **single & limited threading performance**

E-cores

Efficient cores

Efficient cores are optimized for multi-threaded, less compute-intensive workloads

E-cores provide **multi-threaded throughput and power efficiency**

Flexible Performance and Throughput

Introducing Intel® Hybrid Technology for the Edge

New Intel® Core™ processors that deliver flexible performance to meet the growing demands of future edge solutions

A new generation of Intel Core processors have arrived, which deliver a breakthrough in innovation, featuring our new performance hybrid architecture. These new processors can deliver increased flexibility for your future edge services and solutions.

By integrating two all-new core microarchitectures into a single processor, this advanced technology can help increase compute efficiency, sustainability, and support the deployment of intelligent edge services.

What does this mean for future edge solutions?

- Dynamic, controllable and scalable performance
- Enhanced power efficiency to meet deployment requirements and sustainability
- Fine-tune processing controls to optimize edge workloads and services

P-cores
Performance cores

High performance cores to support demanding performance and throughput

E-cores
Efficient cores

Additional cores for enhanced multi-threaded performance and scaling

What's Available Today?

- Find the latest 12th Gen Intel Core processors with Intel Hybrid Technology (link)
- Access Intel Hybrid Technology enabled SDKs and frameworks
- Learn more about Intel Thread Director, which provides feedback to the OS for optimal processing decisions: Video overview, Learn more
- Intel Hybrid Technology is supported today on Windows 11 operating system (link), limited support today on Windows 10 LTSC and Linux

Looking Ahead To Future Edge Usages

- Future support on Windows 11 LTSC and Linux for Intel Thread Director
- Future Intel Hybrid Technology software enhancements for edge services include:
  - OS scheduling to automatically optimize assigning tasks to cores
  - Assign priority tasks/workloads to specific cores or core types
- Edge use case enablement resources and proof-of-concept support

Utilizing Intel® Hybrid Technology

OS Scheduler

Rely on OS scheduler
to assign tasks to cores

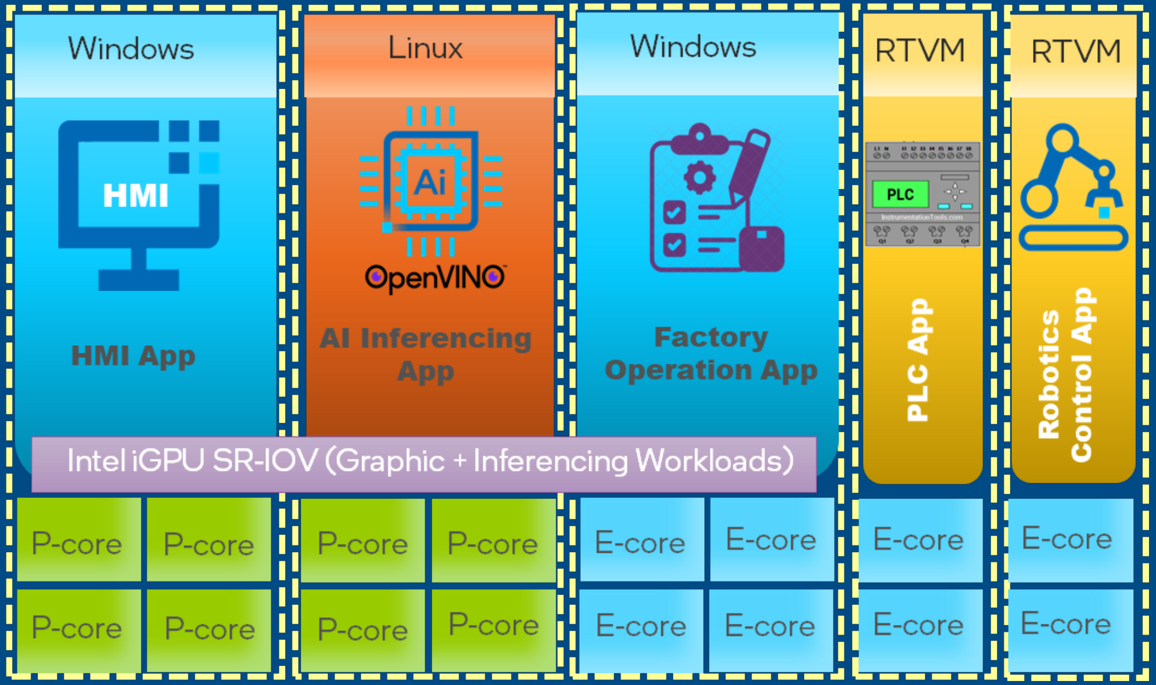
Core Affinity

Pin workloads to
specific cores

Power Throttling

Automatic core
assignment
for power efficiency

Industrial Workload Consolidation Use-Cases



Windows
HMI
HMI App

Linux
Ai
OpenVINO
AI Inferencing App

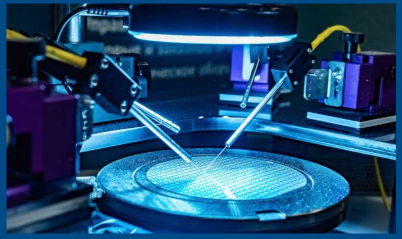
Windows
Factory Operation App

Intel iGPU SR-IOV (Graphic + Inferencing Workloads)

RTVM
PLC
PLC App

RTVM
Robotics Control App

P-core	P-core	P-core	P-core	E-core	E-core	E-core	E-core
P-core	P-core	P-core	P-core	E-core	E-core	E-core	E-core



Workloads consolidation unites various OSes onto a single platform, reducing purpose-built hardware and infrastructure, increase compute asset utilization to achieve optimum cost.

Windows based applications can be allocated based on compute intensity requirement.

AI inferencing with AVX2 enabled advantages on P-cores to provide high inferencing performance.

Real-time apps taking advantages of TCC-enabled-CPU such as on ADL-S to provide optimized workloads performance.

ADL-S: 12th Gen Intel® Core™ processors

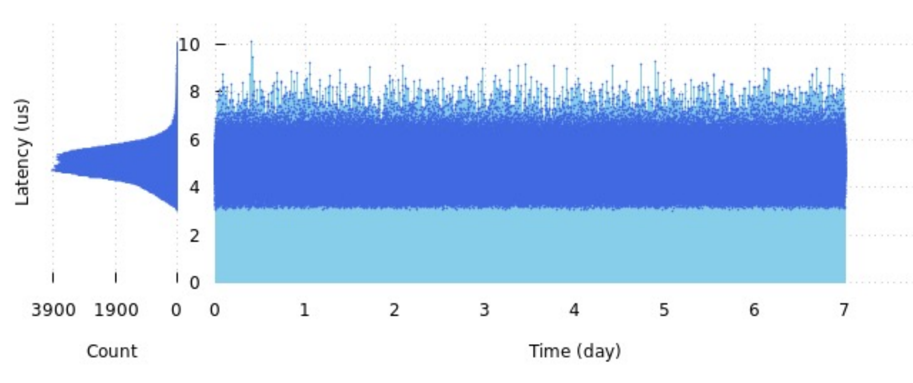
ADL-S RTVM + WaaG Benchmark

Industrial Workload Consolidation scenario benchmarking with ACRN

- P-cores allocation:
  - **SOS** - 1 core
  - **Windows** - 2 cores, 3D graphic workload will use iGPU acceleration
- E-cores allocation:
  - **RTVM** – 2 cores for Xenomai RTVM

Configuration	
Hardware	12th Gen Intel(R) Core(TM) i9-12900E (8P + 8E)
BIOS Config	TCC mode on, Hyperthreading enable, Turbo On, SRIOV enabled
VM config	WaaG (P-cores) 1 RTVM (1 Xenomai) (E-Cores)
Test Contents	WaaG: Passmark RTVM Xenomai: latency + stress-ng

WaaG Test selection	Passmark benchmark percents
CPU	50%
RAM	50%
2d Gfx	50%
Disk	50%
Sound	50%
Network	50%



intel®