AMD

AMD's
Video Processing Engine

Harry Wentland

About Me

- Work for AMD
- I'm a display kernel guy who recently started getting involved in Video and usermode drivers
- I know little about GStreamer itself but would like to learn more
- Played around with mpv and ffmpeg

What is VPE?

Why

What is VPE?

PowerTOP 2.15 Overview Idle stats Frequency stats Device stats Tunables WakeUp

The battery reports a discharge rate of 4.55 W
The energy consumed was 91.3 J
The estimated remaining time is 13 hours, 36 minutes

<ESC> Exit | <TAB> / <Shift + TAB> Navigate |

radeontop 1.4, running on UNKNOWN_CHIP bus c3, 120 samples/sec

Graphics pipe	0.00%
Event Engine	0.00%
Vertex Grouper + Tesselator	0.00%
Texture Addresser	0.00%
Texture Cache	0.00%

~ : bash — Konsole

New Tab Split View Copy Paste Find...

```
[hwentlan@hwstxlaptop ~]$ AMDGPU_HDR_CASE=1 mpv ~/Videos/Samsung\ Wonderland\ Two\ HDR\ UHD\ 4K\ Demo.ts --vo=dmabuf-wayland --geometry=1280x720 --hwdec=vaapi --loop --no-terminal
Cannot load libcuda.so.1
[hwentlan@hwstxlaptop ~]$ AMDGPU_HDR_CASE=1 mpv ~/Videos/Samsung\ Wonderland\ Two\ HDR\ UHD\ 4K\ Demo.ts  --geometry=1280x720  --loop --no-terminal
[hwentlan@hwstxlaptop ~]$
```

PowerTOP 2.15 Overview Idle stats Frequency stats Device stats Tunables WakeUp

The battery reports a discharge rate of 19.9 W
The energy consumed was 398 J
The estimated remaining time is 3 hours, 10 minutes

<ESC> Exit | <TAB> / <Shift + TAB> Navigate |

radeontop 1.4, running on UNKNOWN_CHIP bus c3, 120 samples/sec

Graphics pipe	10.00%
Event Engine	0.00%
Vertex Grouper + Tesselator	5.00%
Texture Addresser	0.00%
Texture Cache	0.00%

Samsung Wonderland Two HDR UHD 4K Demo.ts - mpv

mpv – 4k 24Hz HDR Video
default settings

~ 20 W

ENDEAVOUR
DEFINITION: A purposeful or industrious undertaking.
Especially one that requires effort or boldness.

~ : mpv — Konsole

New Tab Split View Copy Paste Find...

```
[hwentlan@hwstxlaptop ~]$ AMDGPU_HDR_CASE=1 mpv ~/Videos/Samsung\ Wonderland\ Two\ HDR\ UHD\ 4K\ Demo.ts --vo=dmabuf-wayland --geometry=1280x720 --hwdec=vaapi --loop --no-terminal
Cannot load libcuda.so.1
[hwentlan@hwstxlaptop ~]$ AMDGPU_HDR_CASE=1 mpv ~/Videos/Samsung\ Wonderland\ Two\ HDR\ UHD\ 4K\ Demo.ts  --geometry=1280x720  --loop --no-terminal
```

PowerTOP 2.15 | Overview | Idle stats | Frequency stats | Device stats | Tunables | WakeUp

The battery reports a discharge rate of 11.5 W
The energy consumed was 231 J
The estimated remaining time is 5 hours, 50 minutes

<ESC> Exit | <TAB> / <Shift + TAB> Navigate |

radeontop 1.4, running on UNKNOWN_CHIP bus c3, 120 samples/sec

Graphics pipe	13.33%
Event Engine	0.00%
Vertex Grouper + Tesselator	10.83%
Texture Addresser	0.00%
Texture Cache	0.00%

Samsung Wonderland Two HDR UHD 4K Demo.ts - mpv

mpv – 4k 24Hz HDR Video
--hwdec=vaapi

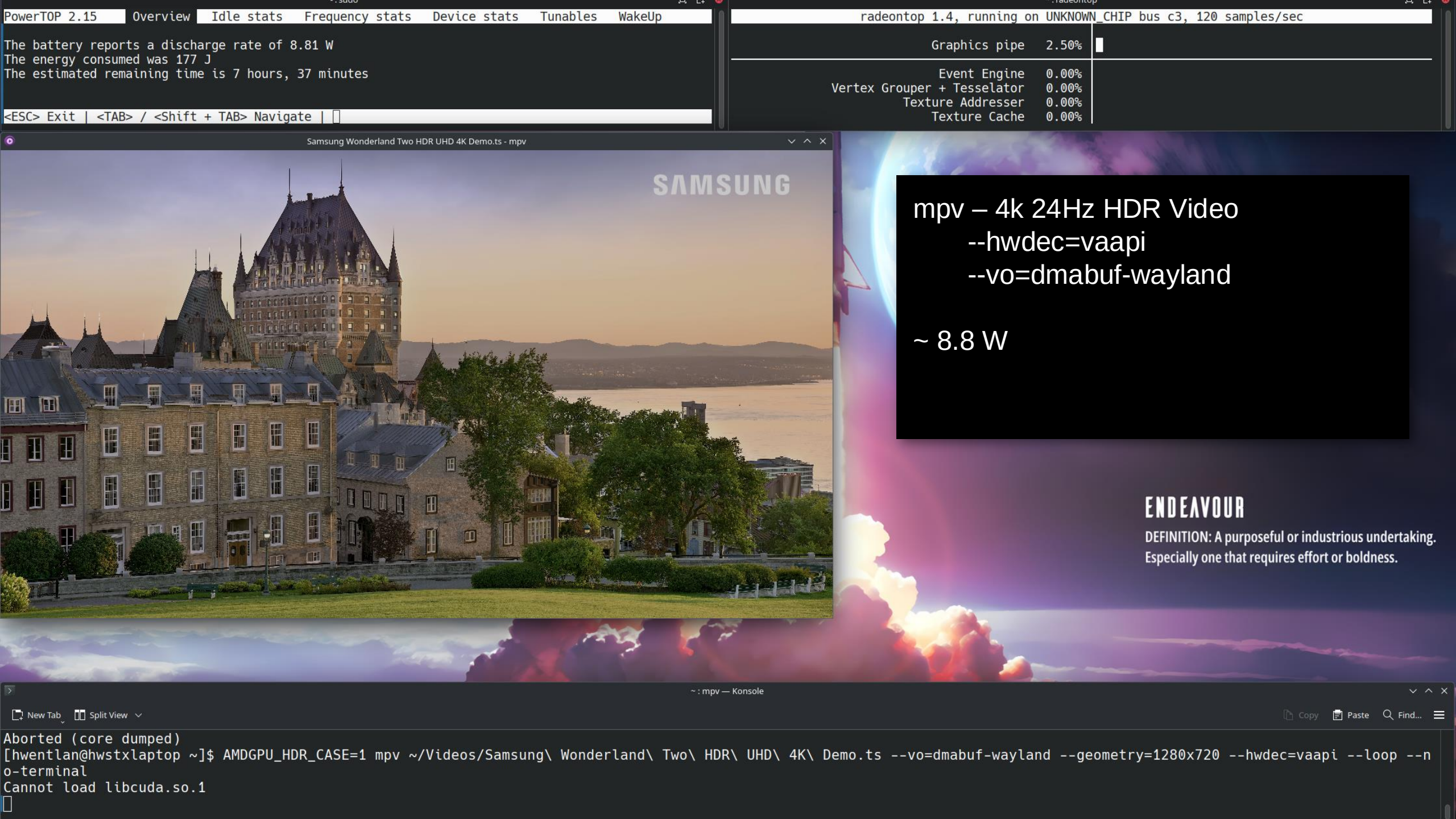
~ 11.5 W

ENDEAVOUR
DEFINITION: A purposeful or industrious undertaking.
Especially one that requires effort or boldness.

~ : mpv — Konsole

New Tab | Split View | Copy | Paste | Find...

```
o-terminal --vf=scale_vaapi=format=rgba
Cannot load libcuda.so.1
[hwentlan@hwstxlaptop ~]$ source ~/git/setup-vpe.source
[hwentlan@hwstxlaptop ~]$ AMDGPU_HDR_CASE=1 mpv ~/Videos/Samsung\ Wonderland\ Two\ HDR\ UHD\ 4K\ Demo.ts --vo=gpu --geometry=1280x720 --hwdec=vaapi --loop --no-terminal
```


PowerTOP 2.15 Overview Idle stats Frequency stats Device stats Tunables WakeUp
The battery reports a discharge rate of 8.81 W
The energy consumed was 177 J
The estimated remaining time is 7 hours, 37 minutes
<ESC> Exit | <TAB> / <Shift + TAB> Navigate |
radeontop 1.4, running on UNKNOWN_CHIP bus c3, 120 samples/sec
Graphics pipe 2.50%
Event Engine 0.00%
Vertex Grouper + Tesselator 0.00%
Texture Addresser 0.00%
Texture Cache 0.00%
Samsung Wonderland Two HDR UHD 4K Demo.ts - mpv
SAMSUNG
mpv – 4k 24Hz HDR Video
--hwdec=vaapi
--vo=dmabuf-wayland
~ 8.8 W
ENDEAVOUR
DEFINITION: A purposeful or industrious undertaking.
Especially one that requires effort or boldness.
~ : mpv — Konsole
New Tab
Split View
Copy
Paste
Find...
Aborted (core dumped)
[hwentlan@hwstxlaptop ~]$ AMDGPU_HDR_CASE=1 mpv ~/Videos/Samsung\ Wonderland\ Two\ HDR\ UHD\ 4K\ Demo.ts --vo=dmabuf-wayland --geometry=1280x720 --hwdec=vaapi --loop --no-terminal
Cannot load libcuda.so.1

PowerTOP 2.15 Overview Idle stats Frequency stats Device stats Tunables WakeUp

The battery reports a discharge rate of 9.50 W
The energy consumed was 191 J
The estimated remaining time is 7 hours, 4 minutes

<ESC> Exit | <TAB> / <Shift + TAB> Navigate |

radeontop 1.4, running on UNKNOWN_CHIP bus c3, 120 samples/sec

Graphics pipe	0.00%
Event Engine	0.00%
Vertex Grouper + Tesselator	0.00%
Texture Addresser	0.00%
Texture Cache	0.00%

Samsung Wonderland Two HDR UHD 4K Demo.ts - mpv

mpv – 4k 24Hz HDR Video
--hwdec=vaapi
--vo=dmabuf-wayland
--vf=scale_vaapi=format=rgba

(using VPE)

~ 9.5 W

Especially one that requires effort or boldness.

~ : mpv — Konsole

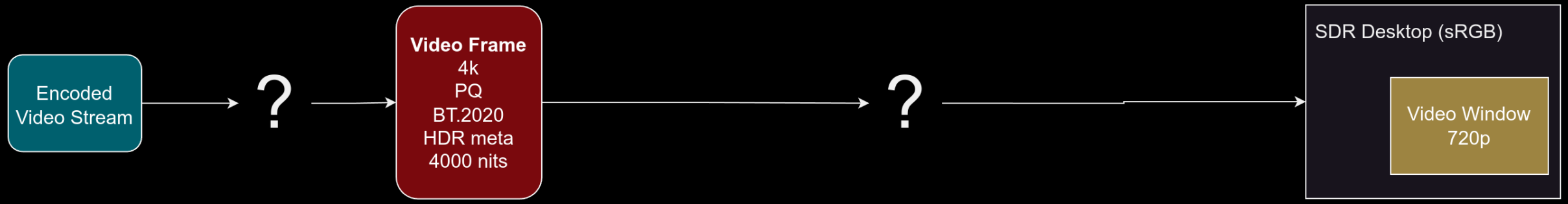
New Tab Split View Copy Paste Find...

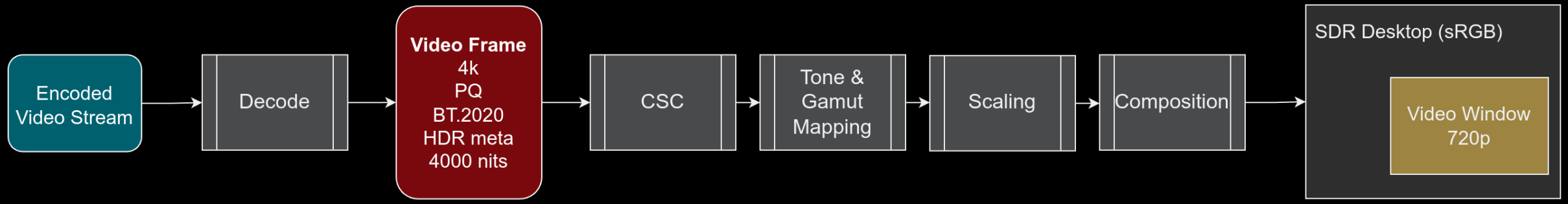
```
Aborted (core dumped)
[hwentlan@hwstxlaptop ~]$ AMDGPU_HDR_CASE=1 mpv ~/Videos/Samsung\ Wonderland\ Two\ HDR\ UHD\ 4K\ Demo.ts --vo=dmabuf-wayland --geometry=1280x720 --hwdec=vaapi --loop --no-terminal --vf=scale_vaapi=format=rgba
Cannot load libcuda.so.1
```

What did we see

- HW Decode isn't enabled by default and saves lots of power
- --vo=dmabuf-wayland gives us better power consumption, but can't handle HDR (yet)
- --vf=scale_vaapi video filter engages VPE
  - Maybe there's a better way?
- VPE can handle HDR conversions
- VPE vs GPU saves power

How the video ends up on the screen...





PowerTOP 2.15 Overview Idle stats Frequency stats Device stats Tunables WakeUp

The battery reports a discharge rate of 19.9 W
The energy consumed was 398 J
The estimated remaining time is 3 hours, 10 minutes

<ESC> Exit | <TAB> / <Shift + TAB> Navigate |

radeontop 1.4, running on UNKNOWN_CHIP bus c3, 120 samples/sec

Graphics pipe	10.00%
Event Engine	0.00%
Vertex Grouper + Tesselator	5.00%
Texture Addresser	0.00%
Texture Cache	0.00%

Samsung Wonderland Two HDR UHD 4K Demo.ts - mpv

mpv – 4k 24Hz HDR Video
default settings

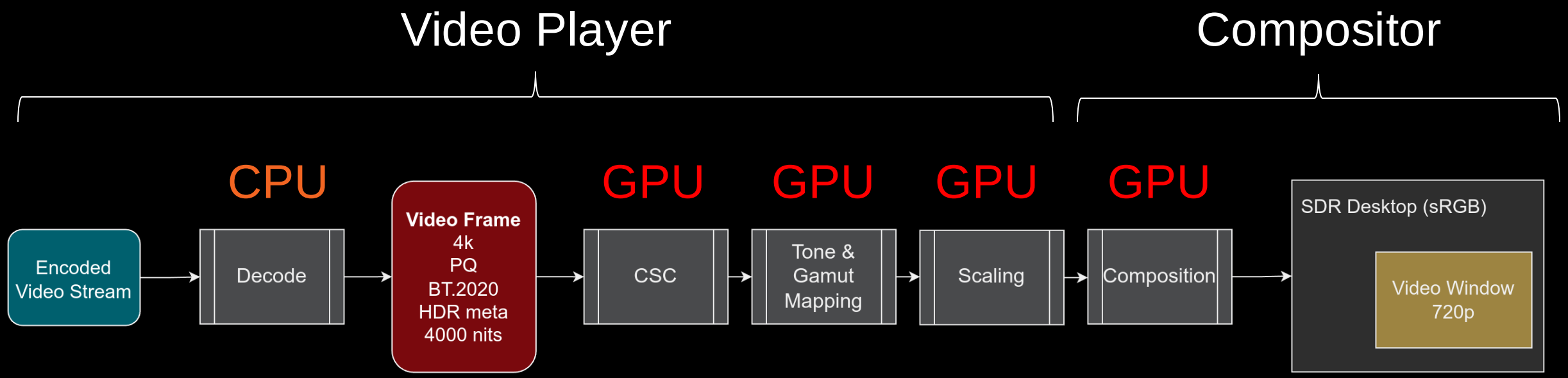
~ 20 W

ENDEAVOUR
DEFINITION: A purposeful or industrious undertaking.
Especially one that requires effort or boldness.

~ : mpv — Konsole

New Tab Split View Copy Paste Find...

```
[hwentlan@hwstxlaptop ~]$ AMDGPU_HDR_CASE=1 mpv ~/Videos/Samsung\ Wonderland\ Two\ HDR\ UHD\ 4K\ Demo.ts --vo=dmabuf-wayland --geometry=1280x720 --hwdec=vaapi --loop --no-terminal
Cannot load libcuda.so.1
[hwentlan@hwstxlaptop ~]$ AMDGPU_HDR_CASE=1 mpv ~/Videos/Samsung\ Wonderland\ Two\ HDR\ UHD\ 4K\ Demo.ts  --geometry=1280x720  --loop --no-terminal
```



PowerTOP 2.15 | Overview | Idle stats | Frequency stats | Device stats | Tunables | WakeUp

The battery reports a discharge rate of 11.5 W
The energy consumed was 231 J
The estimated remaining time is 5 hours, 50 minutes

<ESC> Exit | <TAB> / <Shift + TAB> Navigate |

radeontop 1.4, running on UNKNOWN_CHIP bus c3, 120 samples/sec

Graphics pipe	13.33%
Event Engine	0.00%
Vertex Grouper + Tesselator	10.83%
Texture Addresser	0.00%
Texture Cache	0.00%

Samsung Wonderland Two HDR UHD 4K Demo.ts - mpv

mpv – 4k 24Hz HDR Video
--hwdec=vaapi

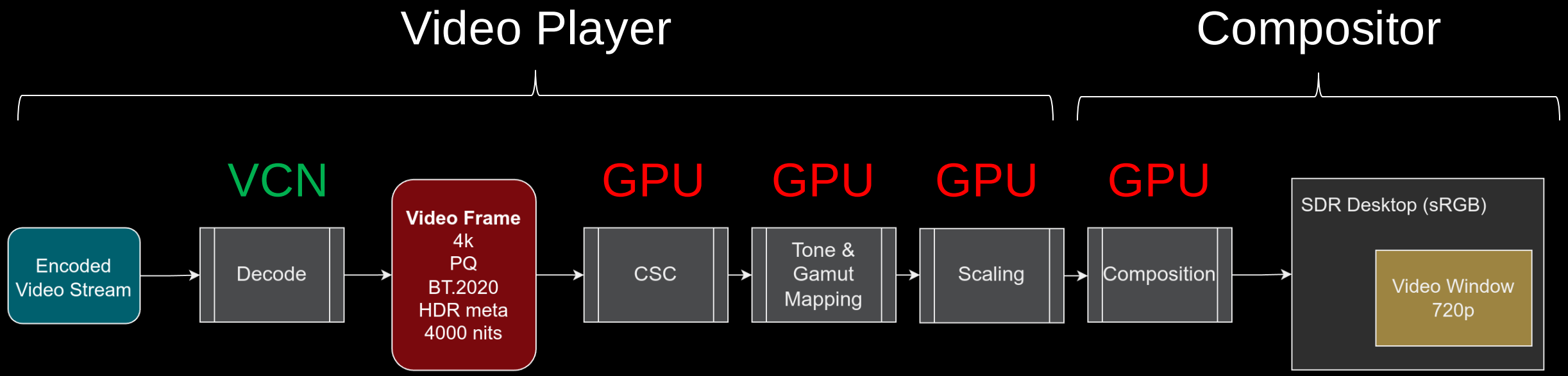
~ 11.5 W

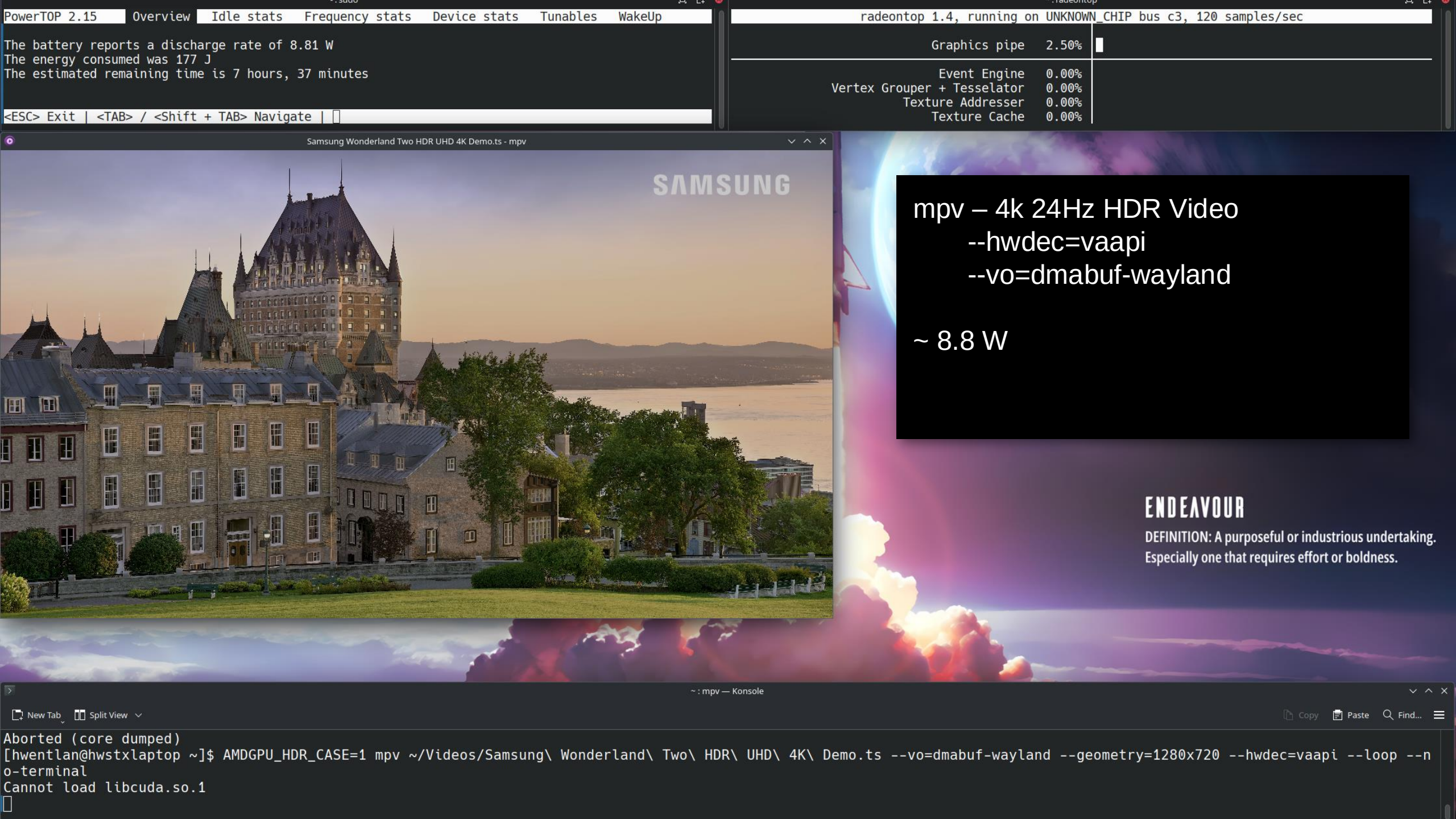
ENDEAVOUR
DEFINITION: A purposeful or industrious undertaking.
Especially one that requires effort or boldness.

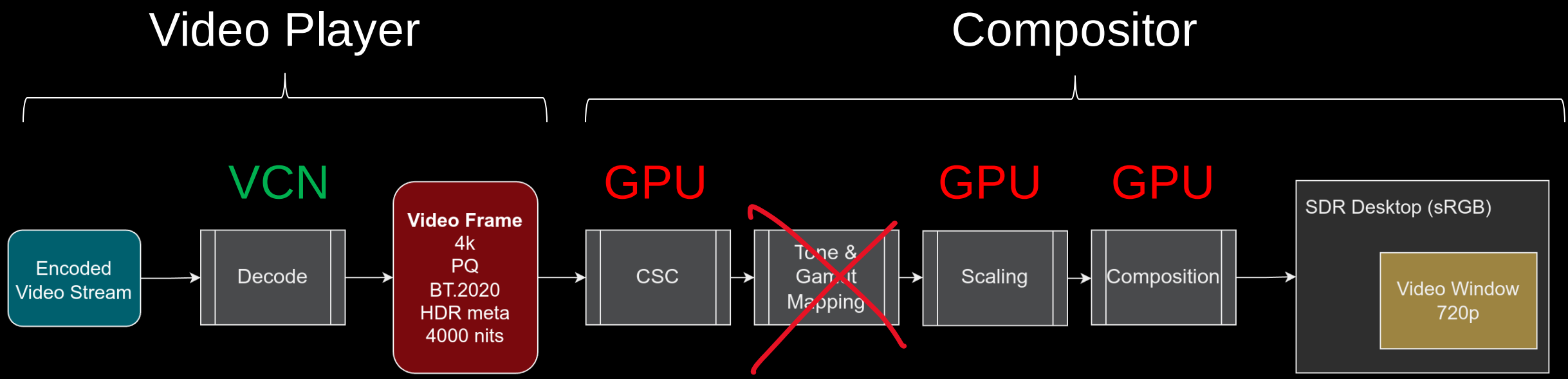
~ : mpv — Konsole

New Tab | Split View | Copy | Paste | Find...

```
o-terminal --vf=scale_vaapi=format=rgba
Cannot load libcuda.so.1
[hwentlan@hwstxlaptop ~]$ source ~/git/setup-vpe.source
[hwentlan@hwstxlaptop ~]$ AMDGPU_HDR_CASE=1 mpv ~/Videos/Samsung\ Wonderland\ Two\ HDR\ UHD\ 4K\ Demo.ts --vo=gpu --geometry=1280x720 --hwdec=vaapi --loop --no-terminal
```




PowerTOP 2.15 Overview Idle stats Frequency stats Device stats Tunables WakeUp
The battery reports a discharge rate of 8.81 W
The energy consumed was 177 J
The estimated remaining time is 7 hours, 37 minutes
<ESC> Exit | <TAB> / <Shift + TAB> Navigate |
radeontop 1.4, running on UNKNOWN_CHIP bus c3, 120 samples/sec
Graphics pipe 2.50%
Event Engine 0.00%
Vertex Grouper + Tesselator 0.00%
Texture Addresser 0.00%
Texture Cache 0.00%
Samsung Wonderland Two HDR UHD 4K Demo.ts - mpv
SAMSUNG
mpv – 4k 24Hz HDR Video
--hwdec=vaapi
--vo=dmabuf-wayland
~ 8.8 W
ENDEAVOUR
DEFINITION: A purposeful or industrious undertaking.
Especially one that requires effort or boldness.
~ : mpv — Konsole
New Tab
Split View
Copy
Paste
Find...
Aborted (core dumped)
[hwentlan@hwstxlaptop ~]$ AMDGPU_HDR_CASE=1 mpv ~/Videos/Samsung\ Wonderland\ Two\ HDR\ UHD\ 4K\ Demo.ts --vo=dmabuf-wayland --geometry=1280x720 --hwdec=vaapi --loop --no-terminal
Cannot load libcuda.so.1



PowerTOP 2.15 Overview Idle stats Frequency stats Device stats Tunables WakeUp

The battery reports a discharge rate of 9.50 W
The energy consumed was 191 J
The estimated remaining time is 7 hours, 4 minutes

<ESC> Exit | <TAB> / <Shift + TAB> Navigate |

radeontop 1.4, running on UNKNOWN_CHIP bus c3, 120 samples/sec

Graphics pipe	0.00%
Event Engine	0.00%
Vertex Grouper + Tesselator	0.00%
Texture Addresser	0.00%
Texture Cache	0.00%

Samsung Wonderland Two HDR UHD 4K Demo.ts - mpv

mpv – 4k 24Hz HDR Video
--hwdec=vaapi
--vo=dmabuf-wayland
--vf=scale_vaapi=format=rgba

(using VPE)

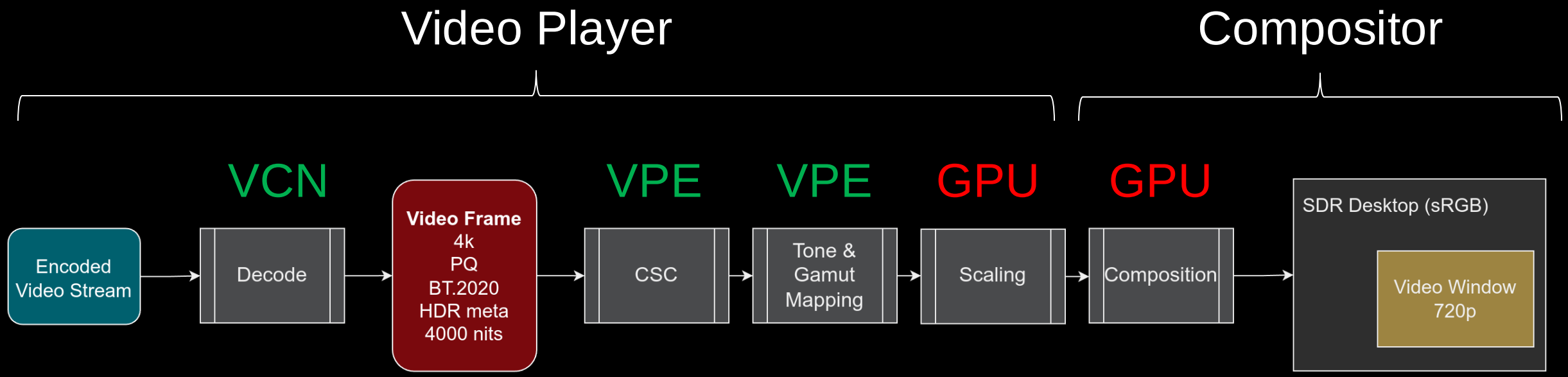
~ 9.5 W

Especially one that requires effort or boldness.

~ : mpv — Konsole

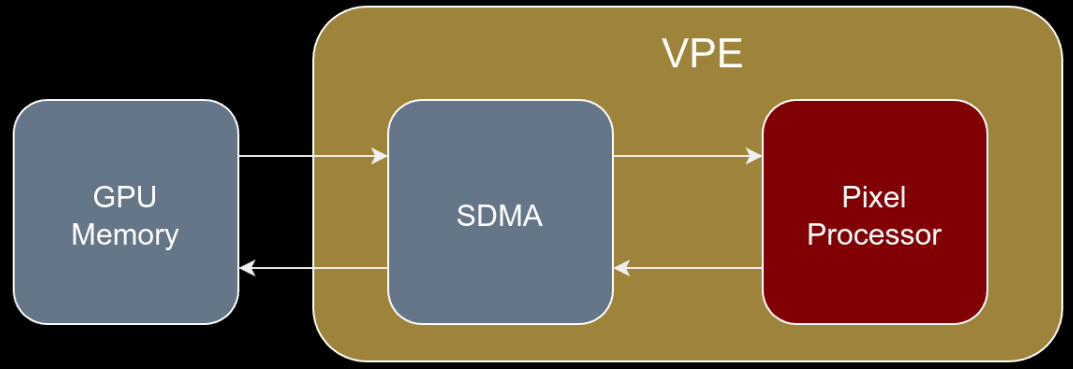
New Tab Split View Copy Paste Find...

```
Aborted (core dumped)
[hwentlan@hwstxlaptop ~]$ AMDGPU_HDR_CASE=1 mpv ~/Videos/Samsung\ Wonderland\ Two\ HDR\ UHD\ 4K\ Demo.ts --vo=dmabuf-wayland --geometry=1280x720 --hwdec=vaapi --loop --no-terminal --vf=scale_vaapi=format=rgba
Cannot load libcuda.so.1
```



What is VPE?

What is VPE



VPEP Formats

- Input: NV12, P010, RGB 8 / 10
- Output: RGB 8 / 10, FP16

SoC Support

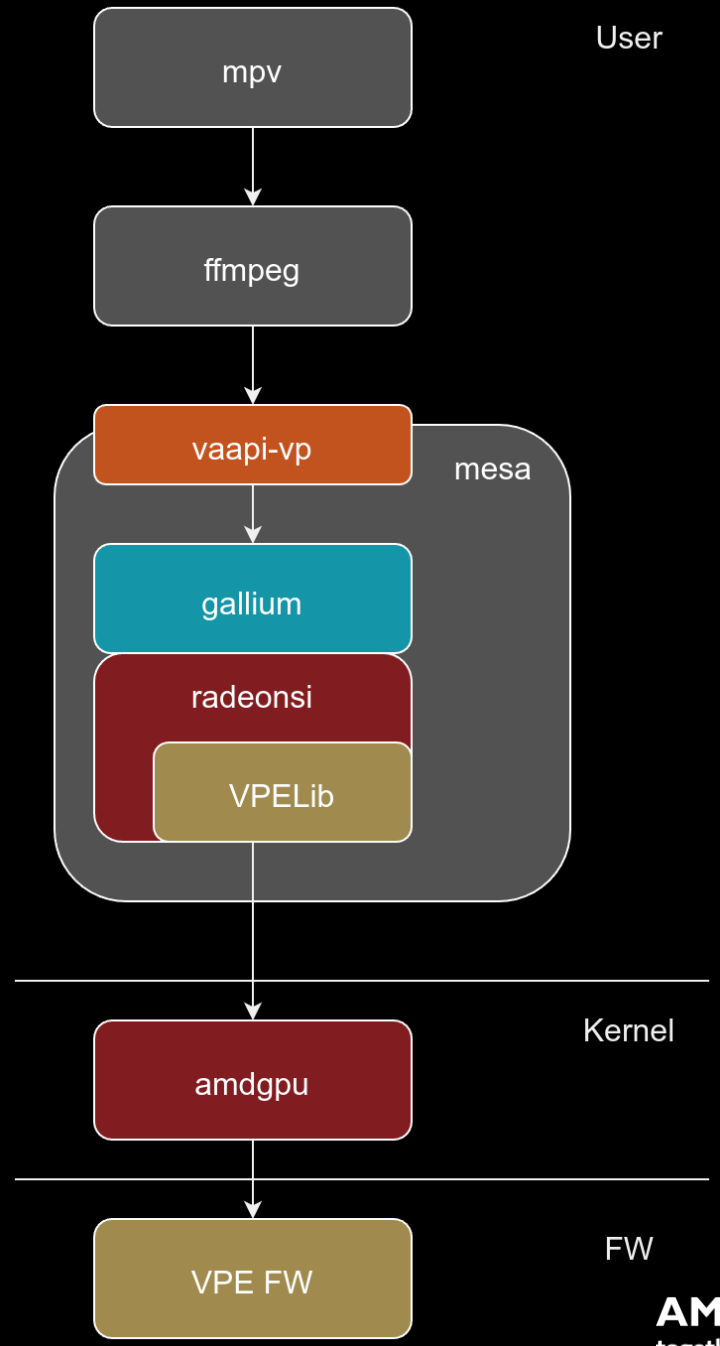
- Ryzen AI 300

VPEP Capabilities

- Tone, and Gamut-mapping via 3DLUT
- Color space conversion via 3x4 matrix
- Scaling
- Linearization
- Support for HDR and SDR curves and color spaces

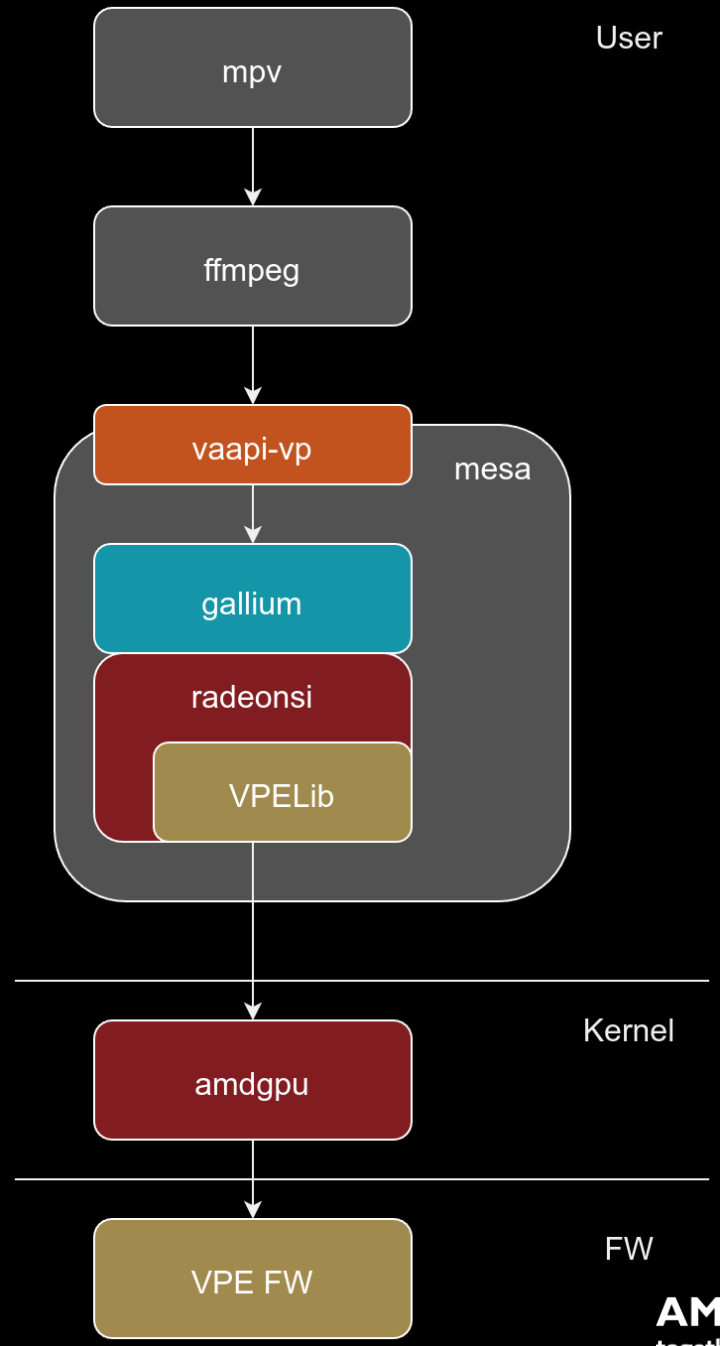
VPE in the Linux Video Stack

- Functionality exposed via VA-API Video Processing API
- src/amd/vpelib in Mesa to create commands
- VPELib is shared code with other OSes
- In mpv and ffmpeg supported via scale_vaapi video filter
- VPE's ability to construct a tone-mapping 3DLUT isn't open-sourced yet, but we'll hopefully open it soon



VPE in the Linux Video Stack

- Radeonsi will create a buffer and call VPELib to fill it, passing
  - Src/dest rect, format, etc
- VPELib will create commands for the VPE FW and put all of that into the buffer
- Radeonsi will call an amdgpu IOCTL to pass the buffer down to HW
- Amdgpu will write buffer into the ring buffer
- VPE FW will execute from the ring buffer



Mpv Video Playback Power

- Video: 4k 24 fps HDR video[1]
  - Fullscreen video playback with mpv – GFX

```
mpv --hwdec=vaapi --hwdec-codecs=all --geometry=100% \
 --vo=gpu --no-osc --gpu-dumb-mode=yes
```

  - Fullscreen video playback with mpv - VPE

```
mpv --hwdec=vaapi --hwdec-codecs=all \
 --vf=scale_vaapi=format=rgba:w=1920:h=1080 \
 --vo=dmabuf-wayland --no-osc --gpu-dumb-mode=yes
```

- SoC Power[3] without VPE: 2.39 W
- SoC Power[3] with VPE: 1.94 W
- Savings[3]: 0.45 W

- **System**:
  - **SoC**: AMD Ryzen AI 9 HX 370 w/ Radeon 890M
  - **Panel**: 1080p60 SDR PSR1

SW:

  - **OS**: Manjaro sway[2]
  - **Kernel**: amd-staging-drm-next
  - **FW**: AMD's internal tree (should be very close to linux-firmware)
  - **Mesa**: main from July with one custom patch to support HDR to SDR conversion
  - **Compositor**: Weston

1) We generated a long, looped version of our video since mpv --loop was caching the video and skewing our results
2) We don't recommend Manjaro sway as it installs laptop-mode-tools which disables some C-state optimizations. TLP has the same issue. We removed laptop-mode-tools for our tests.
3) SoC Power measured via AMD-internal HW and tools

What kills power?
(during video playback)

Relative Power Consumption

CPU GPU Fixed-Function

Frequent GPU Wakes

- Powering the GPU up and down takes time
- The GPU burns power when it's powered, even if it doesn't have work
- Batching work helps avoid frequent GPU wakes

Memory Access

- Memory bandwidth is costly
- Pipeline operations to avoid hitting the memory bus

Thoughts

HW Decode

- Why does mpv not use HW decode by default?
- Why does totem not use HW decode by default?
- Why do browsers not use HW decode by default?

- How to enable HW decode out of the box?

VAAPI Video Processing

- What really uses VAAPI Video Processing?
- When?
- Why?

- The usage I could find seems to be limited
- Why is that?
- It's been a thing for over 10 years, with HW support by Intel

- Can we help drive wider adoption?
- Or is there a better API to do what we want?

- Video Players, Browsers, game streaming, etc.

Memory Copies

- As a newcomer to the video stack it's not easy to tell when operations will hit memory
- Memory access is costly
- Multimedia frameworks are very modular
- There seem to be a myrid of efforts around zero copy

- Do we have good tools to show memory access and/or pipeline chains?
- Do we need new or better tools?

- Zero-copy everywhere all the time

Composition

- In-progress Wayland color management and color representation APIs allow compositor to own all transformations from raw video buffer to composited framebufer
- Compositor can offload work to DRM/KMS planes with in-progress color pipeline API

- Most composition does not require blending
- Could we offload composition from GPU to VPE?

- EGL -> dmabuf -> vaapi-vp

- Could provide similar power benefits as DRM/KMS planes

Summary

Summary

VPE = video processing engine
- Incl., CSC, scaling, tone-mapping

VA-API Video Processing API

Savings[1]: 0.45 W

HW Decode by default should be a priority

VA-API VP API is barely used by anyone

Power Impact:
- CPU > GPU > Fixed Function
- GPU wakes
- Memory Bandwidth

1) SoC Power measured via AMD-internal HW and tools

Acknowledgements

- Solomon Chiu
- Peyton Lee
- Alan Liu
- Leo Li
- George Zhang


AMD

Copyright and disclaimer

- The information presented in this document is for informational purposes only and may contain technical inaccuracies, omissions, and typographical errors. The information contained herein is subject to change and may be rendered inaccurate releases, for many reasons, including but not limited to product and roadmap changes, component and motherboard version changes, new model and/or product differences between differing manufacturers, software changes, BIOS flashes, firmware upgrades, or the like. Any computer system has risks of security vulnerabilities that cannot be completely prevented or mitigated. AMD assumes no obligation to update or otherwise correct or revise this information. However, AMD reserves the right to revise this information and to make changes from time to time to the content hereof without obligation of AMD to notify any person of such revisions or changes.

- THIS INFORMATION IS PROVIDED 'AS IS." AMD MAKES NO REPRESENTATIONS OR WARRANTIES WITH RESPECT TO THE CONTENTS HEREOF AND ASSUMES NO RESPONSIBILITY FOR ANY INACCURACIES, ERRORS, OR OMISSIONS THAT MAY APPEAR IN THIS INFORMATION. AMD SPECIFICALLY DISCLAIMS ANY IMPLIED WARRANTIES OF NON-INFRINGEMENT, MERCHANTABILITY, OR FITNESS FOR ANY PARTICULAR PURPOSE. IN NO EVENT WILL AMD BE LIABLE TO ANY PERSON FOR ANY RELIANCE, DIRECT, INDIRECT, SPECIAL, OR OTHER CONSEQUENTIAL DAMAGES ARISING FROM THE USE OF ANY INFORMATION