Nouveau/NVK Update

Faith Ekstrand

XDC 2024

About me

- Faith Ekstrand

  - @gfxstrand@mastodon.gamedev.place

- Been around freedesktop.org since 2013

  - First commit: wayland/31511d0e, Jan 11, 2013

- At Intel from June 2014 to December 2022

  - NIR, Intel (ANV) Vulkan driver, SPIR-V → NIR, ISL, other Intel bits

- At Collabora since January 2022

  - Work across the upstream Linux graphics stack, wherever needed

  - Currently the lead developer / maintainer of NVK

Nouveau kernel

- Mostly bug fixes this year

- GSP firmware support is merged

- GSP firmware is now shipping in linux-firmware
  and distros

  Well, almost....

Red Hat announced Nova

```
Subject: Nova and staging Rust abstractions
Date: Wed, 20 Mar 2024 18:58:43 +0100    [thread overview]
Message-ID: <Zfsj0_tb-0-tNrJy@cassiopeiae> (raw)

Hi all,

In this mail I briefly want to announce the Nova project and subsequently talk about the first
efforts taken in order to upstream required Rust abstractions:

We just started to work on Nova, a Rust-based GSP-only driver for Nvidia GPUs. Nova, in the long
term, is intended to serve as the successor of Nouveau for GSP-firmware-based GPUs.

With Nova we see the chance to significantly decrease the complexity of the driver compared to
Nouveau for mainly two reasons. First, Nouveau's historic architecture, especially around nvif/nvkm,
is rather complicated and inflexible and requires major rework to solve certain problems (such as
locking hierarchy in VMM / MMU code for VM_BIND currently being solved with a workaround) and
second, with a GSP-only driver there is no need to maintain compatibility with pre-GSP code.
```

What is Nova?

- From scratch NVIDIA kernel driver

- Written in Rust
  - Sharing some Rust abstractions with Asahi
  - Also adding new Rust abstractions

- GSP-only (so Turing+)

- Still in very early development stages

5

Since XDC 2023...

We're Vulkan 1.3 conformant

Software in the Public Interest, Inc. 2024-02-28 Vulkan_1_3	🔗770
NVIDIA GeForce RTX 4090 D	CTS Version: 1.3.7.3
NVIDIA GeForce RTX 4090	CPU: 13th Gen Intel(R) Core(TM) i7-13620H
NVIDIA GeForce RTX 4080	OS: Linux 6.8.0-rc1-drm-misc-fixes
NVIDIA GeForce RTX 4070 Ti SUPER	
NVIDIA GeForce RTX 4070 Ti	
NVIDIA GeForce RTX 4070 SUPER	
NVIDIA GeForce RTX 4070	
NVIDIA GeForce RTX 4060 Ti	

Last year, I submitted Vulkan 1.0 conformance *during* XDC

We implemented a few features...

- shaderStorageImageMultisample
- shaderFloat64
- shaderInt64
- shaderInt16
- shaderResourceResidency
- sparseResidency2Samples
- sparseResidency4Samples
- sparseResidency8Samples
- sparseResidencyAliased
- sparseResidencyImage2D
- sparseResidencyImage3D
- variableMultisampleRate
- KHR_8bit_storage
- KHR_16bit_storage
- KHR_calibrated_timestamps
- KHR_compute_shader_derivatives
- KHR_dynamic_rendering_local_read
- KHR_fragment_shader_barycentric
- KHR_incremental_present
- KHR_index_type_uint8
- KHR_line_rasterization
- KHR_load_store_op_none
- KHR_maintenance5
- KHR_maintenance6
- KHR_maintenance7
- KHR_pipeline_executable_properties
- KHR_pipeline_library
- KHR_present_id
- KHR_present_wait
- KHR_shader_atomic_int64
- KHR_shader_expect_assume

- KHR_shader_float_controls
- KHR_shader_float_controls2
- KHR_shader_float16_int8
- KHR_shader_integer_dot_product
- KHR_shader_maximal_reconvergence
- KHR_shader_quad_control
- KHR_shader_relaxed_extended_instruction
- KHR_shader_subgroup_extended_types
- KHR_shader_subgroup_rotate
- KHR_shader_subgroup_uniform_control_flow
- KHR_shader_terminate_invocation
- KHR_synchronization2
- KHR_vertex_attribute_divisor
- KHR_vulkan_memory_model
- KHR_workgroup_memory_explicit_layout
- KHR_zero_initialize_workgroup_memory
- EXT_attachment_feedback_loop_layout
- EXT_calibrated_timestamps
- EXT_conservative_rasterization
- EXT_color_write_enable
- EXT_depth_bias_control
- EXT_depth_clamp_control
- EXT_depth_range_unrestricted
- EXT_descriptor_buffer
- EXT_device_generated_commands
- EXT_display_control
- EXT_image_drm_format_modifier
- EXT_dynamic_rendering_unused_attachments
- EXT_graphics_pipeline_library
- EXT_host_image_copy

- EXT_image_sliced_view_of_3d
- EXT_legacy_vertex_attributes
- EXT_load_store_op_none
- EXT_map_memory_placed
- EXT_memory_budget
- EXT_multi_draw
- EXT_nested_command_buffer
- EXT_pipeline_creation_cache_control
- EXT_pipeline_creation_feedback
- EXT_pipeline_robustness
- EXT_post_depth_coverage
- EXT_primitive_topology_list_restart
- EXT_primitives_generated_query
- EXT_queue_family_foreign
- EXT_scalar_block_layout
- EXT_shader_image_atomic_int64
- EXT_shader_module_identifier
- EXT_shader_object
- EXT_shader_replicated_composites
- EXT_shader_subgroup_ballot
- EXT_shader_subgroup_vote
- EXT_subgroup_size_control
- EXT_swapchain_maintenance1
- EXT_texel_buffer_alignment
- GOOGLE_decorate_string
- GOOGLE_hlsl_functionality1
- GOOGLE_user_type
- NV_compute_shader_derivatives
- NV_shader_sm_builtins
- VALVE_mutable_descriptor_type

Including everything needed by DXVK

nvk: Tracker issue for DXVK support

⊖ Closed Issue created 1 year ago by **Echo J.**

This will serve as a meta tracker issue for anything required to get DXVK running. This isn't for DXVK bugs, just features.

The list below will provide a missing list of features for three key DXVK versions and various feature levels (I'll also link to issues/MRs if possible)

List of missing (and required) features for DXVK:

DXVK v1.5.1 (the last Vulkan 1.0 release)

FL11_0 (Feature Level 11):

And everything Zink needs for GL 4.6

nvk: Tracker issue for Zink support

⊖ Closed | 🏷 Issue created 1 year ago by **Faith Ekstrand**

This will serve as a meta tracker issue for anything required to get Zink running. This isn't for zink bugs, just features.

List of required features: https://docs.mesa3d.org/drivers/zink.html

Issues or MRs (if there is no issue) for various Zink feature versions:

OpenGL 2.1:

- ☑ VK_KHR_timeline_semaphore (nouveau/mesa!136 (closed))
- ☑ VK_EXT_provoking_vertex (nouveau/mesa!127 (merged))

And most of what VKD3D wants...

We're still missing a few things:

- Ray tracing

- Fragment shader interlock

- ~~Fragment shading rate~~

- Mesh shaders (in progress)

The new compiler supports Maxwell+

nvk/nak: Use NAK on Maxwell+

Merged **Faith Ekstrand** requested to merge ⌥ `gfxstrand/mesa:nak/sm50` 📋 into `main` 2 months ago

Overview 0   Commits 36   Pipelines 3   Changes 27

This MR has a bunch more Maxwell fixes for NAK and fixes all but one known compiler issue with the Vulkan CTS. The one remaining issue is a precision issue with 64 to 16-bit float conversions that also exists with codegen. With this, we can switch to NAK by default for Maxwell.

👍 0   👎 0

✅ **Merge request pipeline #1235185 passed**
Merge request pipeline passed for `36311964` 2 months ago ❓

Lots of games work now!

Horizon: Zero Dawn

Cyberpunk 2077

GPU   0%   0°C
                    0^{MHz}
CPU   21%   4499^{MHz}
VRAM   0.0^{GiB}
RAM   4.7^{GiB}
VULKAN   26^{FPS}   38.0^{ms}
NVIDIA GeForce RTX 2080
Mesa 24.3.0-devel (git-b4acc3fc42)
Frametime   min: 31.0ms, max: 41.6ms

Use the mouse wheel or press X to accelerate.

119 m/s
FLR 0

-X
+Z

ARG Argon Wharf

> Dock at:
MESSAGES

X4 Foundations

GPU   0%   0°C
                    0 MHz
CPU   24%   4500 MHz
VRAM   0.0 GiB
RAM   8.3 GiB
VULKAN   62 FPS   16.2 ms
NVIDIA GeForce RTX 2080
Mesa 24.3.0-devel (git-b4acc3fc42)
Frametime   min: 12.4ms, max: 57.4ms

Transport Fever 2

GPU   0%   0°C
               0 MHz
CPU   32%   4200 MHz
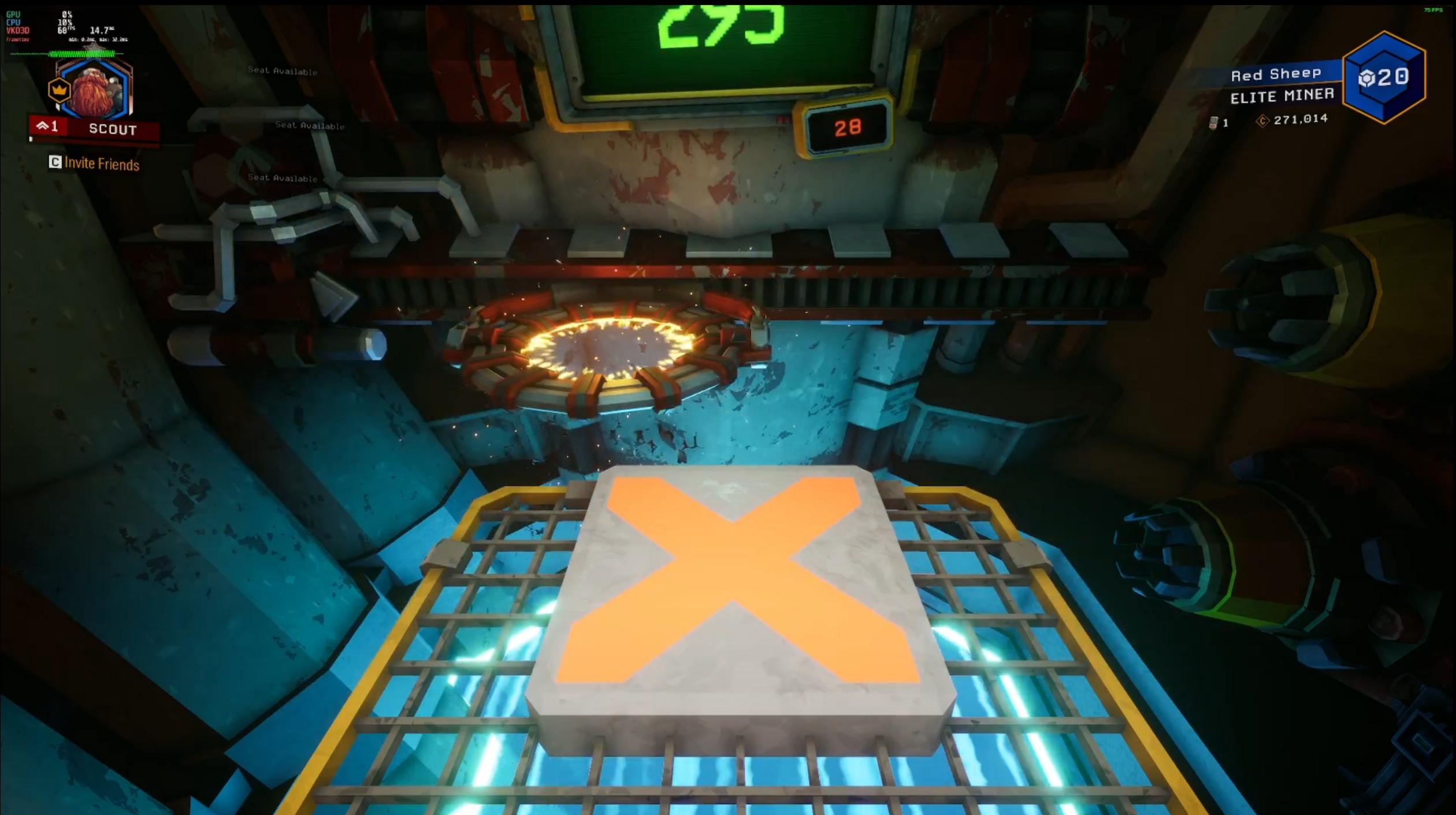VRAM   0.0 GiB
RAM   12.4 GiB
DXVK   74 FPS   13.6 ms
NVIDIA GeForce RTX 2080
Mesa 24.3.0-devel (git-b4acc3fc42)
Frametime   min: 0.4ms, max: 26.4ms

290   300   310 NW 320   330   340   350   N

65
91
81

250ml

1

Rust

Deep Rock Galactic

GPU   0%   0°C
             0 MHz
CPU   17%   901 MHz
VRAM   0.0 GiB
RAM   5.5 GiB
VULKAN   59 FPS   16.8 ms
NVIDIA GeForce RTX 2080
Mesa 24.3.0-devel (git-b4acc3fc42)

Frametime   min: 14.6ms, max: 19.0ms

Return to Monkey Island

DXVK v2.4-73-g18ecc17e59e609e
D3D11 FL11_0
NVIDIA GeForce RTX 4090 (NVK AD102)
Driver: NVK
Version: Mesa 24.3.0-devel (git-366f63fd88)
FPS: 74.7

min: 7.2   max: 21.1
Queue submissions: 8
Queue syncs: 0
Draw calls: 4359
Dispatch calls: 0
Render passes: 53
Barriers: 122
Graphics pipelines: 741
Graphics shaders: 711
Compute pipelines: 0
Descriptor pools: 4
Descriptor sets: 96565
Vidmem heap 0: 3290 MB (13%)   2639 MB used
Sysmem heap 1:   4 MB (0%)   1 MB used
CS chunks: 80
CS syncs: 0
GPU: 96%

The Witness

TUTORIAL - VAULT

GPU
Total: $ +0   0%   0°C
  0 MHz ⚡ +0   ⏣ +0   ◉ +0.0
CPU
Result: $ 63%   900 MHz
VRAM   0.0 GiB
RAM   5.6 GiB
VULKAN   73 FPS   13.7 ms
NVIDIA GeForce RTX 2080
Mesa 24.3.0-devel (git-b4acc3fc42)
Frametime   min: 13.0ms, max: 25.8ms

ROOM TYPE   ITEMS   DECOR   LOOT

Total: $ +0 ⚡ +0   CONFIRM ALL   Show side panels

JAN 1 $ 0,000,040,000   100%   ⚡0.005   020   (•) 000   00   0   B 1 2 3 4

Pause Menu

Evil Genius 2

23

Super TuxKart

Let's talk about Zink…

NVK+Zink is OpenGL 4.6 conformant

Software in the Public Interest, Inc.	2024-05-18	OpenGL_4_6	🔗352

NVIDIA GeForce RTX 3090

NVIDIA GeForce RTX 3080 Ti

NVIDIA GeForce RTX 3080

NVIDIA GeForce RTX 3070 Ti

NVIDIA GeForce RTX 3070

NVIDIA GeForce RTX 3060 Ti

NVIDIA GeForce RTX 3060

NVIDIA GeForce RTX 3050 Ti

NVIDIA GeForce RTX 3050

NVIDIA RTX A6000

NVIDIA RTX A5500

CPU: Intel(R) Core(TM) i9-10980XE

OS: Linux 6.8.5

Display: X11 with Xwayland-23.2.6 and GNOME Shell 46.0

Important NVK features for Zink

- DRM format modifiers

  - Lets Zink use the better display path

  - NVK+Zink can run your compositor

- Shader objects and pipeline libraries

  - Lets Zink avoid stuttering

- Descriptor buffer

  - Lets Zink use its descriptor fast paths

- Multi-draw

Is Zink ready to be the primary OpenGL driver?

There are still issues…

## NVK+Zink: Discord client and Chromium slowly artifacting

○ Open 🗋 Issue created 4 months ago by **Jared Rawlings**

System information

- OS: EndeavourOS
- GPU: NVIDIA Corporation AD102 [GeForce RTX 4090] [10de:2684] (rev a1)
- Kernel version: 6.8.7-1-gfxstrand-g8a6e23f29986 # 1 SMP PREEMPT_DYNAMIC Thu, 09 May 2024 00:18:54 +0000 x86_64
- Mesa version: Mesa 24.2.0-devel (git-59babe9f)
- Xserver version (if applicable): X.Org X Server 1.21.1.13
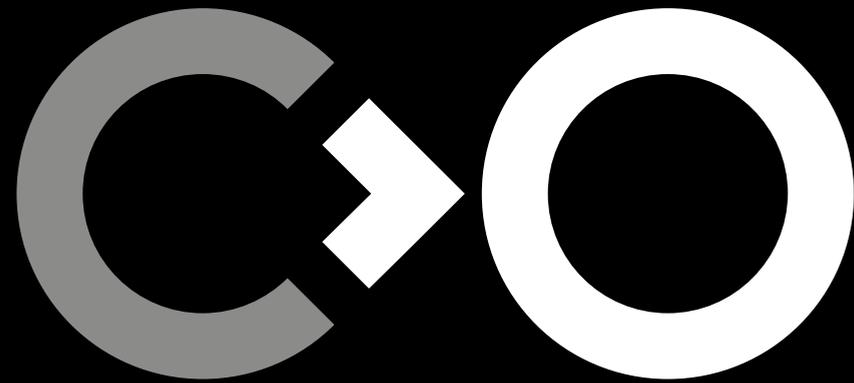- Desktop manager and compositor: KDE Plasma 6.0.4

Describe the issue

There are still issues…
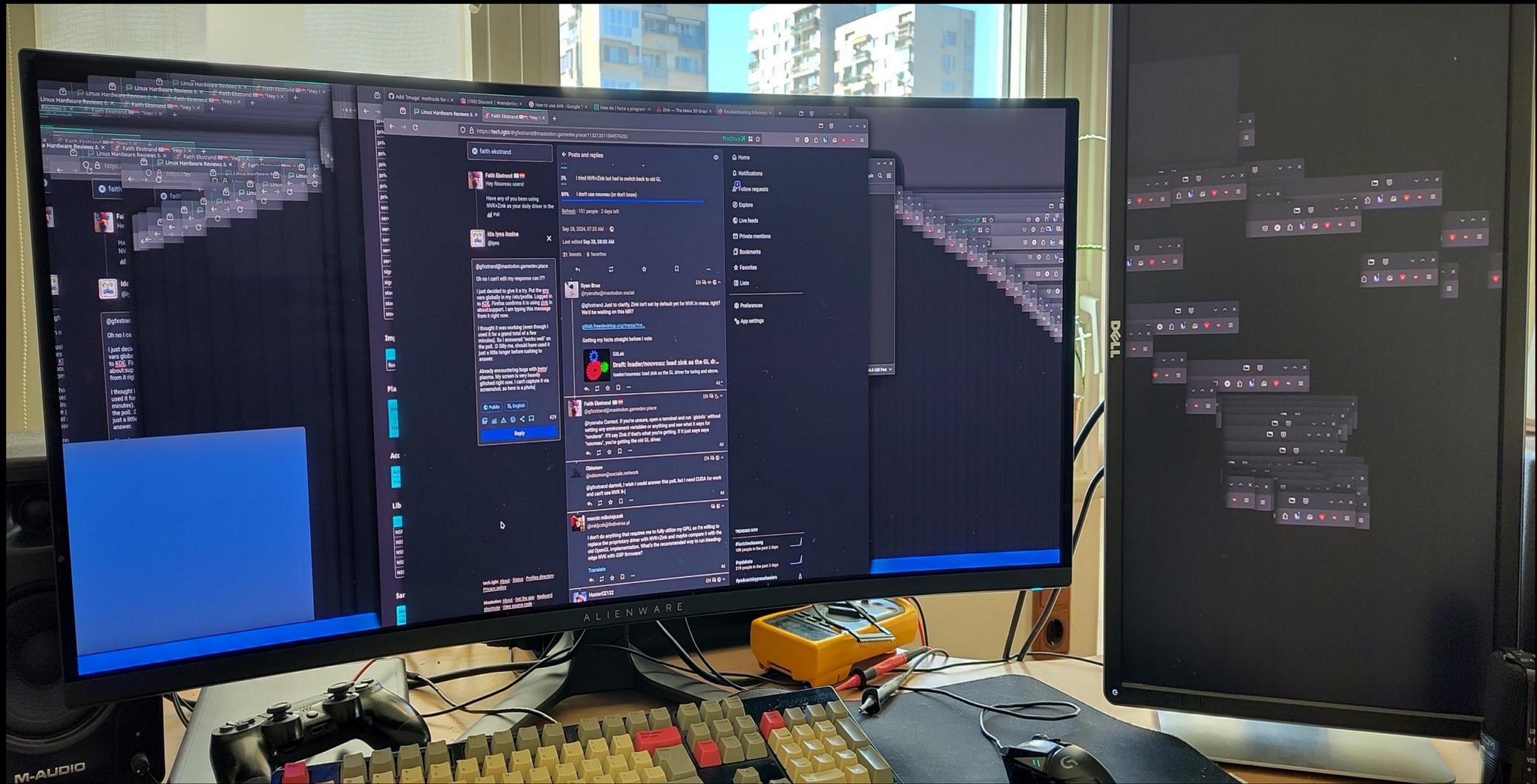
Is Zink ready to be the primary OpenGL driver?

Red Hat is going to try with Fedora 41

Thank you!