**UNICORN VYSOKÁ ŠKOLA S.R.O.**

BAKALÁŘSKÁ PRÁCE

**2023 Valerij Šlovikov**

**UNICORN VYSOKÁ ŠKOLA S.R.O.**

**Softwarový vývoj**



**BAKALÁŘSKÁ PRÁCE**

Využití GPU ve webových prohlížečích

**Autor BP:** Valerij Šlovikov

**Vedoucí BP:** Ing. Michal Gregor

**Čestné prohlášení**

Prohlašuji, že jsem svou bakalářskou práci na téma Využití GPU ve webových prohlížečích vypracoval/a samostatně pod vedením vedoucího bakalářské práce a s použitím výhradně odborné literatury a dalších informačních zdrojů, které jsou v práci všechny citovány a jsou také uvedeny v seznamu použitých zdrojů.

Jako autor/ka této bakalářské práce dále prohlašuji, že v souvislosti s jejím vytvořením jsem neporušil/a autorská práva třetích osob a jsem si plně vědom/a následků porušení ustanovení § 11 a následujících autorského zákona č. 121/2000 Sb.

Dále prohlašuji, že odevzdaná tištěná verze bakalářské práce je shodná s verzí, která byla odevzdána elektronicky.

V……………………. dne ……….. …….……………………………

Valerij Šlovikov

**Poděkování**

Děkuji vedoucímu bakalářské práce Ing. Michalu Gregorovi za účinnou metodickou, pedagogickou a odbornou pomoc a další cenné rady při zpracování mé bakalářské práce. Dále bych chtěl poděkovat za jeho trpělivost a vstřícný přístup během konzultací.



# Využití GPU ve webových prohlížečích

# GPU Utilization in Web Browsers

# Abstrakt

Tato bakalářská práce se zabývá možnostmi a využitím 3D grafiky na webu, především s ohledem na technologie WebGL, WebGL2 a WebGPU. Cílem je porovnat jejich funkční rozdíly, zhodnotit podporu různými prohlížeči, bezpečnostní aspekty a demonstrovat reálné aplikace včetně ukázkových projektů. Součástí práce je také průzkum knihoven, jako jsou Three.js a Babylon.js, se zaměřením na jednoduchost a efektivitu práce. Zvláštní pozornost je věnována parallax mappingu a jeho využití ve spojení se šroubovicí, jakož i potenciálu WebGPU pro akceleraci AI výpočtů přímo v prohlížeči.

Praktická část se zaměřuje na tvorbu několika ukázkových aplikací, včetně vykreslení jednoduchých tvarů, „Hello World“ příkladů a pokročilejších scén využívajících parallax mapping. Dále je popsána implementace canvas komponenty v rámci frameworku uu5 a provádí se analýza výkonu jednotlivých řešení. Výstupem je soubor poznatků a doporučení, jak efektivně navrhovat a vyvíjet 3D aplikace na webu.

*Klíčová slova:* 3D grafika, WebGL, WebGL2, WebGPU, parallax mapping, Three.js, Babylon.js, AI, výkon, webové prohlížeče

This bachelor’s thesis focuses on the possibilities and usage of 3D graphics on the web, primarily examining WebGL, WebGL2, and WebGPU. The goal is to compare their functional differences, evaluate browser support, explore security considerations, and demonstrate real-world applications via sample projects. The research also explores various libraries such as Three.js and Babylon.js, focusing on ease of use and development efficiency. Special attention is given to parallax mapping, its integration with a helix model, and the potential of WebGPU for accelerating AI computations directly in the browser.

The practical part centers around creating several sample applications, including rendering simple shapes, “Hello World” demos, and more advanced scenes featuring parallax mapping. Additionally, the implementation of a canvas component within the uu5 framework is described, accompanied by a performance analysis of the different approaches. The outcome is a set of insights and recommendations for effectively designing and developing 3D applications on the web.

Keywords: 3D graphics, WebGL, WebGL2, WebGPU, parallax mapping, Three.js, Babylon.js, AI, performance, web browsers

**Obsah**

[Využití GPU ve webových prohlížečích 6](#_Toc202298726)

[GPU Utilization in Web Browsers 6](#_Toc202298727)

[Abstrakt 7](#_Toc202298728)

[1. Úvod do 3D Grafiky ve webových aplikacích 10](#_Toc202298729)

[1.1. Proč je 3D grafika na webu důležitá 10](#_Toc202298730)

[1.2. Popis WebGL, WebGL2 a WebGPU 10](#_Toc202298731)

[1.3. Porovnání WebGL/WebGL2 a WebGPU 13](#_Toc202298732)

[1.4. Aplikace 3D grafiky na webu 13](#_Toc202298733)

[1.5. Výhody a omezení 15](#_Toc202298734)

[1.6. Stručná historie a vývoj 16](#_Toc202298735)

[2. Porovnání WebGL, WebGL2 a WebGPU 19](#_Toc202298736)

[2.1. Funkční rozdíly 19](#_Toc202298737)

[2.2. Podpora prohlížečů 20](#_Toc202298738)

[2.3. Bezpečnostní aspekty 20](#_Toc202298739)

[3. Knihovny pro práci s 3D grafikou 20](#_Toc202298740)

[3.1. Three.js, Babylon.js, atd. 21](#_Toc202298741)

[3.2. Porovnání jednoduchosti práce s těmito knihovnami 21](#_Toc202298742)

[4. Parallax Mapping a Šroubovice 22](#_Toc202298743)

[4.1. Využití v 3D grafice 22](#_Toc202298744)

[5. Akcelerace AI pomocí WebGPU 23](#_Toc202298745)

[5.1. Základy strojového učení na grafických procesorech 23](#_Toc202298746)

[5.2. WebGPU a jeho výhody pro AI 23](#_Toc202298747)

# Úvod do 3D Grafiky ve webových aplikacích

3D grafika na webu otevřela nové možnosti pro vývojáře, designéry i koncové uživatele, poskytující bohaté a interaktivní zážitky přímo v internetovém prohlížeči. Již nemusíme čekat na vzdálenou budoucnost, kdy 3D modely, animace, a dokonce i celé virtuální světy budou dostupné na dosah kliknutí. Tyto technologie už jsou tady a postupně se stávají standardní součástí moderních webových aplikací.

## Proč je 3D grafika na webu důležitá

V době, kdy uživatelé očekávají stále atraktivnější a dynamičtější obsah, představuje 3D grafika výrazné odlišení od běžných 2D webů. Ať už se jedná o prezentační web pro architektonický projekt, online konfigurátor produktu, nebo interaktivní vzdělávací nástroj, 3D modely a animace vtahují uživatele mnohem hlouběji do obsahu. Tím se zvyšuje míra zapojení i délka stráveného času na webu, což může výrazně přispět k úspěchu projektu.

Dříve byla 3D grafika doménou specializovaného softwaru a hardwaru. Uživatelé potřebovali výkonné stroje a programy, jako je Blender, 3ds Max, Maya, případně herní enginy typu Unity či Unreal Engine. Nasazení 3D prvků na webu se často řešilo přes pluginy jako Flash nebo Unity Web Player, což bylo komplikované z pohledu bezpečnosti i uživatelského komfortu. Vstupem technologií jako WebGL, WebGL2 a WebGPU do mainstreamových prohlížečů se 3D grafika stává daleko přístupnější a demokratizovanější, a to jak pro vývojáře, tak pro uživatele.

## Popis WebGL, WebGL2 a WebGPU

**WebGL (Web Graphics Library)** je API vycházející z OpenGL ES 2.0, které umožňuje vykreslování 2D i 3D grafiky přímo v prohlížeči bez nutnosti instalovat jakékoli externí pluginy. Je výsledkem snah skupiny Khronos Group a poprvé bylo oficiálně vydáno v roce 2011 (The Khronos Group, 2011). Od té doby si WebGL našlo cestu do všech hlavních prohlížečů včetně Google Chrome, Mozilla Firefox, Safari a Microsoft Edge (Can I use, 2024).

* **Fungování:**  
  WebGL využívá kódy zvané shadery (vertex a fragment shadery), které běží na grafické kartě (GPU). Programátor píše tyto shadery v jazyce GLSL (OpenGL Shading Language) a skrze JavaScript s nimi komunikuje. Na nízkoúrovňové úrovni tak dochází ke komunikaci mezi prohlížečem a grafickým hardwarem uživatele.
* **Hlavní rysy:**
  + **Vykreslování bez pluginů** – stačí aktuální prohlížeč.
  + **Podpora GPU** – umožňuje akcelerované vykreslování i složitých scén, her a interaktivních projektů.
  + **Omezení** – některé pokročilejší funkce moderních GPU (jako 3D textury nebo některé formy pokročilých shaderů) nebyly v původním WebGL dostupné.
* **Využití v praxi:**
  + Interaktivní grafy, vzdělávací simulace (např. 3D anatomie, fyzikální modely).
  + Základní 3D hry a vizualizace CAD modelů přímo v prohlížeči.

S vydáním **WebGL2** se podpora 3D grafiky na webu dostala o krok dál. WebGL2 vychází z **OpenGL ES 3.0,** čímž rozšiřuje původní WebGL o celou řadu funkcí, které byly dříve dostupné pouze na nativních platformách (The Khronos Group, 2017).

* **Co nového přináší:**
  + **Pokročilé vykreslovací techniky**: Například transform feedback, který umožňuje zachycovat výsledky z vertex shaderu a dále je zpracovávat.
  + **Instanced rendering**: Možnost vykreslovat opakované instance stejného objektu (např. tisíce stromů v lese) efektivněji.
  + **3D textury a rozšířená správa bufferů**: Poskytuje více možností, jak textury ukládat a pracovat s nimi.
  + **Vylepšené shadery**: Podpora rozšířeného GLSL, které umožňuje použít komplexnější výpočty a funkce přímo na GPU.
* **Kompatibilita:**
  + V současnosti mají hlavní prohlížeče (Chrome, Firefox, Edge) (Can I use, 2024) již poměrně stabilní podporu WebGL2, Safari stále postupně dohání.
  + Na mobilních zařízeních může být situace o něco komplikovanější, některá starší zařízení neimplementují OpenGL ES 3.0 dostatečně stabilně.
* **Přínosy v praxi:**
  + **Lepší výkon** u složitých 3D scén (např. masivní scény s mnoha objekty).
  + **Realistické efekty** jako volumetrické světlo, pokročilé stínování či post-processing.
  + **Efektivnější práce s daty** díky vylepšené správě bufferů a paměti.

Nejnovějším přírůstkem v oblasti webové 3D grafiky a GPU akcelerace je **WebGPU**. Stojí za ním opět konsorcium Khronos Group ve spolupráci s dalšími klíčovými hráči. WebGPU se inspiruje moderními grafickými API, jako jsou **Vulkan, Metal** (Apple) a **Direct3D 12** (Microsoft), a přináší na web nižší úroveň přístupu k GPU a další výhody, které WebGL/WebGL2 neposkytují (MDN Web Docs, 2024).

* **Klíčové vlastnosti:**
  + **Větší kontrola nad hardwarem**: WebGPU je nízkoúrovňové API, které dává vývojářům jemnější řízení nad tím, jak se grafika (a výpočty) na GPU provádí.
  + **Vyšší výkon**: Díky nízkoúrovňové správě paměti, front vykreslování (rendering queues) a asynchronním výpočtům lze dosáhnout lepšího využití GPU.
  + **Podpora výpočtů (compute shaders)**: Kromě grafiky lze WebGPU využívat i pro obecné paralelní výpočty, což je ideální pro úlohy strojového učení (ML) nebo vědeckých výpočtů.
  + **Flexibilita**: Architektonicky vychází z designu „moderních“ API (Vulkan, Metal, Direct3D 12), proto je nastavení pipeline a resource managementu pro zkušené GPU programátory výrazně flexibilnější než WebGL.
* **Stav podpory v prohlížečích:**
  + WebGPU je stále relativní novinka. V Chrome a ve Firefoxu existuje experimentální podpora (často schovaná za příznaky „flags“).
  + Safari a Edge také postupně zavádějí podporu, ale musí se počítat s tím, že není tak stabilní jako u WebGL2.
  + Vývojový ekosystém okolo WebGPU se teprve rodí, nicméně existují už knihovny (např. wgpu, Dawn od Googlu, WebGPU adaptace v Babylon.js) usnadňující práci.
* **Příklady využití:**
  + **Kombinace 3D a AI**: Například rendering scény a zároveň strojové učení pro herní logiku v reálném čase.
  + **Pokročilé grafické efekty**: Realistické stíny, GI (global illumination), fluidní simulace, hair & fur simulace atd.
  + **Výpočetně náročné úlohy**: WebGPU se neomezuje jen na grafiku – lze ho použít pro parallel computing, zpracování velkých datasetů přímo na GPU, generování terénu apod.

## Porovnání WebGL/WebGL2 a WebGPU

| **Vlastnost** | **WebGL** | **WebGL2** | **WebGPU** |
| --- | --- | --- | --- |
| **Založeno na** | OpenGL ES 2.0 | OpenGL ES 3.0 | Vulkan, Metal, Direct3D 12 |
| **Pokročilé funkce** | Omezené | Lepší správa textur, instanced rendering | Nízká úroveň, compute shadery, fronty, paměťové bariéry |
| **Snadnost použití** | Relativně snadná | Stále dosti podobné WebGL | Vyšší složitost – nutné detailní nastavení pipeline |
| **Výkon** | Střední | Vyšší než WebGL | Nejvyšší (efektivnější správa GPU) |
| **Podpora v prohlížečích** | Široká (Chrome, Firefox, Safari, Edge) | Dnes již široká, ale může chybět v některých starších verzích | Postupně přibývá, dosud experimentální v některých prohlížečích |

Z tabulky je zřetelné, že **WebGPU** staví na úplně jiné úrovni přístupu k GPU než WebGL/WebGL2. Zatímco WebGL2 rozšiřuje stávající řešení založené na OpenGL ES, WebGPU vychází z moderních API, která jsou navržena s ohledem na efektivitu a flexibilitu v rámci současného grafického hardware.

Pro **běžné webové 3D aplikace** a pro rychlé prototypování je zřejmě stále nejrozumnější využití **WebGL2** (nebo knihoven nad ním), jelikož dosud nabízí nejširší podporu a spolehlivost napříč prohlížeči.  
Naopak **WebGPU** ocení ti, kdo potřebují skutečně využít potenciál GPU naplno – ať už pro fotorealistické renderování, masivní výpočty, nebo integraci s ML (např. neural nety běžící přímo na GPU v prohlížeči) (Chickerur, 2024).

## Aplikace 3D grafiky na webu

3D grafika na webu není jen zajímavý trend – už dnes se naplno využívá v mnoha oblastech a s vývojem nových technologií se její uplatnění neustále rozšiřuje. Níže uvádím několik hlavních oblastí, kde se 3D řešení uplatňují:

* **Interaktivní vizualizace**
  + **Vzdělávání**: V různých online kurzech, výukových platformách nebo virtuálních laboratořích najdeme 3D modely molekul, anatomických částí těla či historických artefaktů. Studenti si mohou detailně prohlédnout objekty ze všech stran nebo s nimi dokonce interagovat.
  + **Průmyslové využití**: Zobrazování 3D modelů strojů a zařízení v inženýrských či výrobních firmách. Díky tomu lze urychlit proces návrhu, údržby nebo školení zaměstnanců, kteří se s modelem mohou seznámit předem.
  + **Prezentační a marketingové účely**: 3D grafické prvky dokážou efektivně zaujmout a představit produkt či koncept atraktivním způsobem (např. architektonické návrhy).
* **3D e-commerce**
  + **Online prohlídky produktů**: Mnohé e-shopy nabízejí 3D modely nábytku, elektroniky nebo třeba bot, takže zákazníci si mohou zboží „osahat“ ze všech stran.
  + **Konfigurátory**: Auta, kola, domácí spotřebiče nebo i oblečení je možné konfigurovat online (vybrat barvu, materiál či doplňky) a ihned vidět výsledek v reálném 3D zobrazení.
  + **Zvýšení důvěry zákazníků**: 3D náhled pomáhá zákazníkům lépe pochopit, co kupují, což může snižovat počet vratek a zvyšovat spokojenost s nákupem.
* **Hry**
  + **Webové hry s 3D prostředím**: Díky stále rychlejším GPU a pokročilým API jako WebGL2 se webové hry už přibližují kvalitou tradičním desktopovým hrám.
  + **Multiplayer**: Propojení 3D enginů a technologií jako WebSockets nebo WebRTC umožňuje hrát v reálném čase s ostatními hráči přímo v prohlížeči.
  + **Experimenty a prototypy**: Web je ideální platforma pro rychlé sdílení ukázek či dem – nemusí se instalovat žádná aplikace, stačí odkaz.
* **Virtuální a rozšířená realita (VR/AR)**
  + **WebXR**: Standard umožňující propojit 3D grafiku s VR a AR brýlemi, takže je možné vytvořit například virtuální prohlídky, simulace či hry přímo v prohlížeči.
  + **Rozšířená realita**: S chytrým telefonem lze v prohlížeči zobrazit virtuální objekty přes kameru v reálném prostředí (např. náhled, jak by vypadal nábytek v obývacím pokoji).
  + **Integrace s gesty a ovladači**: Uživatelé se mohou v 3D prostředí volně pohybovat a interagovat s objekty pomocí speciálních ovladačů.
* **Umělecké projekty a simulace**
  + **3D galerie a generativní umění**: Umělci mohou vystavovat svá díla ve virtuálních prostorách nebo vytvářet interaktivní vizuální instalace.
  + **Hudební vizualizace**: 3D animace reagující na hudební podněty (spektrální analýza, basové linky, rytmy) můžou být atraktivním doplňkem koncertních streamů či DJ setů.
  + **Fyzikální a vědecké simulace**: Například simulace proudění tekutin, modely vesmíru, molekulární dynamiky atd. Přímo v prohlížeči se dají demonstrovat komplexní jevy, a to i s možností interakce (změna parametrů za běhu).

## Výhody a omezení

Ačkoli je 3D grafika na webu mocná a slibná, nelze opomenout některé limity, které s sebou přináší. Níže jsou shrnuty hlavní klady a zápory, na které by měl vývojář pamatovat.

* **Výhody**
  + **Dostupnost a jednoduchá distribuce**: Každý, kdo má moderní prohlížeč, může 3D aplikaci používat ihned, bez nutnosti stahovat a instalovat další software. To výrazně usnadňuje šíření a sdílení – stačí poslat odkaz.
  + **Okamžité aktualizace**: Jakmile je aplikace (nebo hra) hostovaná na serveru, uživatel vždy získává nejnovější verzi. Odpadá tak starost, že někdo používá zastaralou instalaci.
  + **Multiplatformnost**: Web funguje na různých operačních systémech (Windows, macOS, Linux, Android, iOS) i zařízeních (PC, notebook, tablet, mobil). Pokud je projekt navržen správně, uživatelé ho mohou spustit kdekoli.
  + **Komunita a ekosystém**: Existuje široká paleta knihoven (Three.js, Babylon.js, PlayCanvas apod.), návodů a open-source projektů, což zjednodušuje vývoj.
* **Omezení**
  + **Výkon a kompatibilita**: Výsledná rychlost a kvalita vizualizace závisí na konkrétní kombinaci prohlížeče, zařízení a GPU. Na starších telefonech nebo počítačích nemusí 3D scény běžet plynule, případně se nemusí vůbec načíst.
  + **Komplexita vývoje**: I když existují knihovny, 3D programování obecně vyžaduje znalost matematiky (vektory, matice), shaderů a optimalizačních technik. Pro kvalitní výsledky tak potřebuješ čas a zkušenosti.
  + **Bezpečnostní omezení**: Prohlížeče běží v tzv. sandboxu, aby chránily uživatele i systém. To může některá nízkoúrovňová volání blokovat nebo omezovat (např. přímý přístup k GPU paměti je silně regulován).
  + **Potřeba fallback řešení**: Pokud chceš zajistit podporu i pro uživatele se staršími prohlížeči, musíš myslet na alternativy (např. 2D verzi nebo statické obrázky), pokud prohlížeč nepodporuje WebGL2 či WebGPU.

## Stručná historie a vývoj

Historie 3D grafiky na webu je poměrně mladá ve srovnání s tradičním (desktopovým) 3D renderingem. Přesto se během posledních dvou dekád odehrál výrazný posun od experimentálních pluginů k robustním standardům integrovaným přímo v prohlížečích. V této kapitole projdeme hlavní milníky, které umožnily, aby se 3D obsah stal běžnou součástí webových stránek.

* **Počátky (VRML, Flash, Unity Web Player)**
  + **VRML (Virtual Reality Modeling Language)**: Vznikl už v polovině 90. let a byl jedním z prvních pokusů o přenesení 3D grafiky do prostředí webu. Umožňoval popis 3D objektů a scén pomocí textových souborů s příponou \*.wrl. Přestože VRML dokázal do jisté míry zobrazit 3D modely v prohlížeči, nikdy nezískal masivní popularitu. Důvodem byla nízká podpora v mainstreamových prohlížečích a nutnost instalace speciálních pluginů, což komplikovalo uživatelský zážitek.
  + **Flash**: Ačkoli se Flash spojoval primárně s 2D animacemi a interaktivními prvky, existovaly i pokusy o využití 3D (např. papervision3D). Opět se ale jednalo o proprietární technologii vyžadující instalaci pluginu a s limitovaným přístupem k GPU akceleraci.
  + **Unity Web Player**: V době, kdy Unity začalo nabírat na popularitě jako herní engine, nabídl plugin Unity Web Player, který umožňoval spouštět 3D hry a aplikace v prohlížeči. I zde ale platilo, že se jednalo o uzavřenou technologii vyžadující zvláštní instalaci – a s rozšiřováním mobilních zařízení, kde pluginy často nefungovaly, se stávalo toto řešení stále méně životaschopné.
* **Nástup WebGL (2011)**
  + V roce 2011 byla oficiálně vydána první stabilní verze WebGL 1.0, která vycházela z OpenGL ES 2.0. Tento krok představoval zásadní zlom, protože WebGL se stal standardem přímo podporovaným prohlížeči (nejprve Chrome a Firefox, později Safari a další).
  + **Hlavní přínosy WebGL 1.0**:
    - Odstranění nutnosti používat pluginy – uživatelé mohli spustit 3D obsah „z krabice“, pokud měli aktuální prohlížeč.
    - Založení na existující technologii (OpenGL ES 2.0) zajistilo relativně rychlý růst a možnost využívat znalosti z herního průmyslu.
  + Po vydání WebGL 1.0 začaly vznikat první populární knihovny jako Three.js, které vývojářům usnadnily práci s tímto poněkud nízkoúrovňovým API. Během krátké doby se objevila řada experimentálních dem a her, což dokládalo potenciál 3D grafiky na webu.
* **WebGL2 (2017)**
  + S rostoucími nároky na kvalitu 3D zobrazení a výkon vznikla potřeba posunout WebGL dál. Výsledkem byl standard **WebGL2**, publikovaný konsorciem Kronos Group v roce 2017.
  + **Novinky ve WebGL2**:
    - **Podpora OpenGL ES 3.0** – objevily se funkce jako 3D textury, *transform feedback*, *instanced rendering*, jež usnadnily vykreslování komplexnějších scén a efektů.
    - **Vylepšené shadery a uniformy** – umožňovaly pokročilejší grafické efekty přímo na GPU.
    - **Vyšší stabilita a kompatibilita** – mnozí vývojáři ocenili lepší funkce pro práci s buffery a texturami, což vedlo k efektivnějšímu využití GPU.
  + Přijetí WebGL2 bylo postupné, protože ne všechny prohlížeče (zejména mobilní) ho implementovaly současně a beze zbytku. Postupem času se však WebGL2 stal běžně dostupným v Chrome, Firefoxu či nových verzích Edge (Chromium), zatímco Safari podporu postupně vylepšuje.
* **Příchod WebGPU (2022+)**
  + Nejmladší kapitolu v tomto vývoji představuje WebGPU, které se začalo objevovat v experimentální podobě v roce 2020 a do stabilnější fáze se dostává zhruba od roku 2022.
  + **Inspirace moderními API**: WebGPU se opírá o principy a architektury Vulkanu (Khronos), Metalu (Apple) a Direct3D 12 (Microsoft), což jsou nízkoúrovňová API zaměřená na vysoký výkon a detailní kontrolu nad hardwarem.
  + **Rozdíl oproti WebGL/WebGL2**: Zatímco WebGL/WebGL2 jsou postaveny nad OpenGL ES, WebGPU jde cestou přímého, nižšího přístupu k GPU. To poskytuje vyšší efektivitu a flexibilitu při vykreslování, ale také zvyšuje komplexitu implementace.
  + **Možnosti využití**:
    - **Pokročilé grafické efekty** (fotorealistické osvětlení, ray tracing v budoucnu, fluidní simulace).
    - **Strojové učení** (computational shadery dovolují masivní paralelní výpočty, což pomáhá akcelerovat trénování nebo inferenci neuronových sítí).
  + **Současný stav podpory**: WebGPU je v mnoha prohlížečích k dispozici jako experimentální funkce (např. Chrome Canary, Firefox Nightly), je ale jen otázkou času, kdy se stane plnohodnotným standardem se stabilní podporou napříč platformami.

## Základy architektury GPU a vykreslovací pipeline

Moderní grafické procesory jsou masivně paralelní zařízení navržená pro SIMT (single‑instruction‑multiple‑thread) výpočty. Každé jádro (Streaming Multiprocessor u NVIDIE / Compute Unit u AMD) obsahuje stovky až tisíce ALU a sdílené rychlé cache, které umožňují zpracovat tisíce vláken současně (The Khronos® Vulkan Working Group, 2025).

Ve vykreslovací pipeline dnes najdeme fixní kroky (tesselátor, rasterizér) i programovatelné fáze, ve kterých běží shadery psané v jazycích GLSL, WGSL nebo SPIR‑V. WebGL 1/2 vystavuje pouze vertex a fragment shader; WebGPU umožňuje navíc compute, fragment, vertex a brzy i mesh/task shadery (The Khronos Group, 2017). Tato programovatelnost dovoluje provádět nejen klasické vykreslování trojúhelníků, ale i obecné GPGPU výpočty (např. konvoluční jádra pro ML) (W3C, 2025).

### Propustnost vs. latence

GPU dosahují vysoké propustnosti, protože paralelizují tisíce operací, ale jejich single‑thread latence je vyšší. Pro reálný výkon je proto klíčové minimalizovat přepínání stavů, upravovat data dávkově a u WebGPU se vyhnout synchronizačním bariérám, pokud to není nezbytné (Bernhardsson, 2024).

## Integrace GPU v prohlížečích a bezpečnostní model

Každý prohlížeč vkládá grafickou část do vlastního *GPU procesu*, aby oddělil potenciálně nebezpečné shadery od renderovacího enginu. Před spuštěním shaderu proběhne **kompilace a validace** – pro WebGPU dokonce dvoufázová (WGSL → SPIR‑V‑like IR → nativní ISA) (Beaufort, 2023).

**Sandboxing:** Grafické API nemá přímý přístup k paměti GPU; přidělování bufferů zprostředkuje prohlížeč. Tím se omezuje riziko *memory‑leak* i post‑Spectre útoků založených na časování. Firefox implementuje tzv. *robust buffer access* chránící před přetečením indexu (Mozilla Security Team, 2024).

### Prohlížečová podpora v roce 2025

* **WebGL 2.0** – podporováno všemi hlavními desktopovými prohlížeči ≥ 2018.
* **WebGPU** – od **Chrome 113 (04/2023)** ve výchozím stavu; Firefox Nightly a Safari Technology Preview je zatím nutné zapnout ručně v sekci *experimental features*, ale podle oficiálního plánu vývoje by mělo být stabilní vydání dostupné do Q4 2025 (W3C News, 2024).

## Souhrn teoretických poznatků

1. WebGPU přináší **explicitní model správy zdrojů** inspirovaný Vulkanem a umožňuje až 3–4× rychlejší vykreslování komplexních scén oproti WebGL 2  (Chickerur, 2024).
2. Compute shadery otevírají dveře k běhu ML modelů lokálně v prohlížeči (např. *Transformers.js* inference ~3× rychlejší) (WebGPU Experts, 2024).
3. Správné využití **pipeline state objects** a **bind‑skupin** je klíčem k minimálnímu CPU overheadu (Bernhardsson, 2024).

# Porovnání WebGL, WebGL2 a WebGPU

Vývoj 3D grafiky na webu je úzce propojen s technologiemi WebGL, WebGL2 a WebGPU. Každá z nich má jiné možnosti, výkonové charakteristiky, a také různý stupeň podpory mezi prohlížeči. Následující podkapitoly se věnují funkčním rozdílům, podpoře jednotlivých prohlížečů a bezpečnostním hlediskům.

## Funkční rozdíly

* **WebGL (OpenGL ES 2.0 pro web)**
  + Základy 3D vykreslování: Poskytuje nízkoúrovňové rozhraní, kde programátor řeší buffery, shadery (vertex/fragment) a transformace.
  + Nižší nároky na znalosti GPU pipeline: V porovnání s novějšími API je WebGL „jednodušší“ a vhodný pro základní 3D scény.
  + Omezená podpora některých pokročilých funkcí: Např. geometry shadery, transform feedback nebo 3D textury zde nebyly nativně dostupné.
* **WebGL2 (OpenGL ES 3.0 pro web)**
  + Rozšířené funkce z OpenGL ES 3.0: Například 3D textury, transform feedback (umožňuje sbírat data z vertex shaderů), instanced rendering (efektivní vykreslování více kopií stejného objektu).
  + Lepší správa paměti a shaderů: Větší flexibilita při práci s různými formáty textur, buffery a větší kontrola nad GPU zdroji.
  + Zpětná kompatibilita: Většina kódu z WebGL 1.0 lze ve WebGL2 používat, ale k dispozici je zároveň více vylepšení a optimalizací.
* **WebGPU (inspirace Vulkan, Metal, Direct3D 12)**
  + Nízkoúrovňový přístup: Dává vývojáři mnohem větší kontrolu nad vykreslovacím řetězcem (pipeline). Je tedy možné optimalizovat výkon na úrovni, která v WebGL/WebGL2 nebyla dostupná.
  + Compute shadery: Umožňují provádět obecné výpočty na GPU, takže WebGPU je ideální nejen pro grafiku, ale i pro strojové učení nebo vědecké simulace.
  + Vyšší výkon a flexibilita: WebGPU dosahuje až 3,5násobně rychlejšího výkonu pro výpočetně náročné úlohy (Chickerur, 2024).
  + Vyšší složitost vývoje: Na rozdíl od WebGL/WebGL2 je nutné spravovat více zdrojů ručně (komandové fronty, paměťové bariéry apod.), což vyžaduje hlubší znalosti grafické pipeline.

## Architektonické rozdíly a výkonnostní charakteristiky

### Implicitní vs. explicitní řízení GPU

WebGL používá **implicitní stavový model** – každé volání může změnit globální kontext, což komplikuje paralelizaci a vede k vyššímu driver‑overheadu (Kenwright, 2022).   
Naopak WebGPU staví na **explicitních deskriptorech** a předem definovaných *pipeline state objects*, takže ovladač přesně ví, jak budou zdroje a stavy použity, a může optimalizovat předem (Beaufort, WebGPU Fundamentals, 2023).

### Syntetické testy v Godot enginu

Fransson & Hermansson (2024) provedli na hardwaru *Intel Core i7‑12700H + RTX 3060* sérii syntetických benchmarků v Godotu:

|  |  |  |  |
| --- | --- | --- | --- |
| **Workload** | **WebGL GPU [ms]** | **WebGPU GPU [ms]** | **Speed‑up** |
| 10 quadů | 1.64 | 0.18 | **× 9.0** |
| 1 000 quadů | 0.66 | 0.14 | **× 4.7** |
| 50 000 quadů | 2.76 | 0.50 | **× 5.5** |
| 50 000 polygonů | 150.94 | 15.75 | **× 9.6** |

*Zdroj dat: Figure 5 & 7 v [GEM 2024 paper]* (Fransson, Hermansson, &  Hu, 2024)

Autoři přisuzují zrychlení zejména:

* eliminaci **CPU‑overheadu** při validaci příkazů,
* možnosti **předkompilace shaderů** do nativního ISA

### Paměťové bariéry a synchronizace

WebGPU sice skrývá většinu interní synchronizace, ale exponuje **buffer/texture usage transitions** a umožňuje explicitně vkládat *memory barriers* v CommandEncoderu. Diskuze k návrhu (GitHub Issue #27) ukazuje, že správné rozložení bariér snižuje latenci přenosů mezi CPU↔GPU až o 50 % (kvark, 2017).

## Podpora prohlížečů

* **WebGL**: Dnes už v podstatě celoplošně podporované, běží na aktuálních verzích Chrome, Firefoxu, Safari, Edge i většině mobilních prohlížečů.
* **WebGL2**: Stabilní podpora u Chrome, Firefoxu, novějšího Edge (na jádru Chromium), situace v Safari je lepší než dřív, ale i tak může být někdy omezená. Na mnoha mobilních zařízeních funguje, ale setkáš se s různými drobnými omezeními (typ chipsetu, verze OS).
* **WebGPU**: Teprve v ranější fázi adopce. Chrome jej podporuje oficiálně (od verze 113 ve výchozím stavu, dříve za příznakem chrome://flags), ve Firefoxu a Safari je zatím převážně experimentální. Pro mobilní zařízení je to ještě složitější – řada z nich nemá nejnovější API v OS, takže je potřeba další vývoj a optimalizace. (Can I use, 2024)

## Bezpečnostní aspekty

* **Sandboxing**: Prohlížeče záměrně omezují, co kód s WebGL/WebGPU může dělat, aby chránily hardware a operační systém. V praxi to znamená, že přímý přístup do GPU paměti je filtrovaný a spravovaný prohlížečem.
* **Ověřování shaderů**: WebGL/WebGPU prochází kontrolou kódu shaderů, aby se zabránilo potenciálnímu zneužití zranitelností v ovladačích GPU.
* **Ochrana soukromí**: Přestože GPU dokáže zpracovávat velká data, prohlížeč brání přímému čtení obsahu paměti, ke kterému programátor nemá explicitní přístup. Minimalizuje se tak riziko neoprávněného získání citlivých informací.
* **Potenciální exploit (Spectre/Meltdown)**: Některé bezpečnostní problémy spojené s GPU/CPU (časování operací, spekulativní vykonávání instrukcí) se řeší i v rámci GPU sandboxu; většinou se ale jedná o hlubší úroveň než samotné WebGL/WebGPU API.

# Knihovny pro práci s 3D grafikou

Práce přímo s WebGL či WebGPU je často poměrně nízkoúrovňová – vyžaduje dobrou znalost shaderů, transformací a správy bufferů. Proto vznikly knihovny (frameworky), které vývoj značně usnadňují.

## Three.js, Babylon.js, atd.

* **Three.js**
  + Jedna z nejoblíbenějších a nejrozšířenějších knihoven pro 3D na webu.
  + Nabízí ucelený ekosystém: kamery, scény, světla, modely, materiály, animace.
  + Má rozsáhlou komunitu a tisíce pluginů, příkladů i tutoriálů.
  + Je primárně postavena na WebGL (aktuálně WebGL2), experimentální větve už chystají podporu WebGPU.
* **Babylon.js**
  + Další robustní framework s podobným zaměřením.
  + Rychle se vyvíjí a nabízí pokročilé funkce (např. PBR – Physically Based Rendering, podpora WebXR, částečně WebGPU).
  + Má skvělý nástroj Playground, kde lze interaktivně psát kód a rovnou vidět výsledek.
* **Jiné knihovny**
  + **PlayCanvas**: Herní engine s online editorem a zaměřením na výkon.
  + **A-Frame** (nad Three.js): Usnadňuje tvorbu VR/AR aplikací díky HTML-like syntaxi.
  + **Oimo.js, Cannon.js**: Fyzikální enginy pro simulaci pohybu, kolizí a dalších jevů.

## Porovnání jednoduchosti práce s těmito knihovnami

* **Three.js**
  + Velmi přehledná a „tradiční“ struktura: vytvoříš scénu, kameru, renderer a poté přidáváš objekty.
  + Dokumentace je rozsáhlá, ale může být pro začátečníka trochu chaotická. Komunita je ovšem obrovská, tutoriály najdeš všude.
  + Pro jednodušší projekty je ideální – stačí pár řádků kódu a zobrazíš první 3D objekt.
  + Nevýhoda: některé pokročilejší efekty (fyzika, post-processing) je potřeba řešit buď doplňkovými knihovnami, nebo custom shaderem. (Cabello, 2025)
* **Babylon.js**
  + Má bohatou podporu WebXR a VR/AR. Přináší i editor uzpůsobený pro vytváření scén.
  + Pro úplné nováčky může být Babylon.js trochu robustnější, ale je skvěle zdokumentovaný a Playground usnadňuje experimentování.
  + Jde o „vše v jednom řešení“, do kterého se dají integrovat i pluginy pro fyziku, animace, materiály atd.
  + Disponuje docela aktivní komunitou, i když menší než Three.js. (Babylon.js Team, 2025)
* **Herní enginy (PlayCanvas, Unity WebGL build)**
  + Poskytují profesionální nástroje pro tvorbu 3D her, včetně editoru a asset pipeline.
  + V některých případech je ale obtížnější propojit tyto enginy s vlastní webovou aplikací (např. Single Page App), oproti integraci menších 3D knihoven.

Obecně platí, že Three.js a Babylon.js jsou skvělé na interaktivní 3D scény, prezentace produktů či menší hry. Pokud je cílem komplexní hra nebo rozšířené VR/AR projekty, lze sáhnout po Babylon.js (kvůli WebXR) nebo některém z herních enginů. V konečném důsledku záleží na preferencích, požadovaných funkcích a osobních zkušenostech vývojáře.

# Parallax Mapping a Šroubovice

**Parallax Mapping** (někdy též Parallax Occlusion Mapping – POM) je technika určená k vylepšení vnímání nerovností povrchu materiálu (např. kamene, cihly, dřeva) bez nutnosti používat extrémně hustou geometrii. Využívá tzv. height map (výškovou mapu), která definuje, jak hluboké jsou detaily textury. Při renderování se pak tyto „hloubky“ počítají ve shaderu a simulují efekt nerovností.

* Princip: Místo jednoduchého texturování se do fragment shaderu přidává výpočet průchodu paprsku texturou (stejně jako by byl povrch fyzicky nerovný).
* Výhoda: Relativně malý náklad na GPU oproti reálnému modelování všech detailů v geometrii.
* Nevýhoda: Při extrémních úhlech pohledu může dojít k vizuálním artefaktům (např. „odseknutí“ textury).

**Šroubovice** (Helix) je pak konkrétní 3D útvar, který lze snadno naprogramovat pomocí parametrických rovnic. Zajímavý je, když se k němu aplikuje parallax mapping, jelikož křivky a točivé plochy pěkně demonstrují, jak detailně může vyhlížet povrch.

## Využití v 3D grafice

* **Parallax mapping**
  + Herní enginy: Pro vytváření realistických povrchů, např. kamenná podlaha, nerovné cihly na zdi.
  + Architektonické vizualizace: Při pohledu na fasádu či stěny dokáže navodit dojem detailních struktur bez obrovského množství polygonů.
  + WebGL aplikace: Parallax mapping je dobrá ukázka pokročilejších shaderových technik, které lze provozovat v reálném čase i v prohlížeči (PlayCanvas, 2025).
* **Šroubovice**
  + Vizualizace v biologii (DNA helix), fyzice (závitovky, vlnovody), strojírenství (závity na šroubech).
  + Designové efekty: Může jít čistě o estetický prvek, který demonstruje prostorové křivky a materiály.

# Akcelerace AI pomocí WebGPU

S nástupem WebGPU se objevuje příležitost nejen pro lepší grafiku, ale i pro výpočty strojového učení přímo v prohlížeči. Pro některé aplikace (lokální zpracování dat, soukromí, offline režim) je GPU akcelerace na straně klienta velmi užitečná.

## Základy strojového učení na grafických procesorech

* **Paralelní výpočty**: GPU disponuje mnoha jádry (ve srovnání s CPU), která mohou zpracovávat tisíce vláken současně. To je ideální pro výpočetně náročné algoritmy, jako je trénování neuronových sítí či masivní maticové operace (např. v convolutional neural networks).
* **Vyšší propustnost**: Pro ML (Machine Learning) úlohy je důležitá rychlá propustnost při práci s velkými maticemi, což GPU dokáže lépe než CPU, které sice může mít rychlejší jádra, ale jich je výrazně méně.
* **Knihovny pro ML**: Existují projekty jako TensorFlow.js, ONNX.js nebo WebDNN, které už využívají GPU akceleraci (dříve jen WebGL, postupně i WebGPU).

## WebGPU a jeho výhody pro AI

* **Compute Shadery**: WebGPU umožňuje psát kód, který se spouští přímo na GPU, a nemusí jít o klasické vykreslování trojúhelníků. Díky tomu lze implementovat např. kernel konvoluce pro neurony, redukce, lineární algebra a další ML operace.
* **Menší overhead**: WebGPU lépe pracuje s frontami příkazů a paměťovými zdroji, takže je možné dosáhnout vyššího výkonu než při využití WebGL, který byl původně navržen výhradně pro rendering.
* **Soukromí a lokální výpočty**: Uživatel nemusí posílat citlivá data na server; model lze nahrát do prohlížeče a inference (vyhodnocení) probíhá lokálně. To se hodí např. u aplikací pro zpracování obrazu či zvuku v reálném čase (detekce gest, rozpoznávání hlasu).
* **Předzvěst budoucích standardů**: Jakmile se WebGPU ustálí, je pravděpodobné, že prohlížečové knihovny ML přejdou na WebGPU jako hlavní backend, což by mělo přinést výrazné výkonnostní skoky.

# Metodologie experimentů a testovací prostředí

Praktická část této bakalářské práce je založena na systematickém porovnání výkonu a funkčnosti technologií WebGL, WebGL2 a WebGPU prostřednictvím série kontrolovaných experimentů. Cílem je poskytnout objektivní data pro hodnocení vhodnosti jednotlivých technologií pro různé typy 3D webových aplikací.

## Specifikace testovacího prostředí

**Hardware konfigurace**

Všechny experimenty byly provedeny na standardizovaném hardwaru s následující konfigurací:

* **Procesor:** 12th Gen Intel(R) Core(TM) i7-12700H @ 2.30 GHz
* **Operační paměť:** 64 GB RAM
* **Grafická karta:** NVIDIA GeForce RTX 4090 (mobilní verze)
* **Operační systém:** Microsoft Windows 11

Tato konfigurace představuje high-end hardware současnosti, což umožňuje otestovat plný potenciál jednotlivých technologií bez hardwarových limitací. Pro kontextualizaci výsledků jsou doplňkově provedeny testy na omezeném hardwaru prostřednictvím virtualizovaného prostředí s omezenými zdroji (4 GB RAM, 2 CPU cores, integrovaná grafika).

**Software prostředí**

**Testované prohlížeče:**

* **Mozilla Firefox** (nejnovější stabilní verze v době testování)
* **Google Chrome** (nejnovější stabilní verze v době testování)

**Vývojové nástroje:**

* Visual Studio Code s rozšířeními pro WebGL/WebGPU development
* Git pro verzování a dostupnost zdrojového kódu
* Local HTTP server pro eliminaci CORS omezení

Všechny testy jsou prováděny v čistém prostředí s restartovaným prohlížečem a vymazanou cache pro eliminaci externích faktorů ovlivňujących výkon.

## Definice testovacích scénářů

**Komplexita scén**

Testovací scénáře jsou navrženy pro postupné zvyšování komplexity od základních primitiv po pokročilé renderovací techniky:

1. **Baseline testy:** Prázdná scéna a základní RGB trojúhelník pro měření overhead jednotlivých API
2. **Objektová komplexita:** Scény s 100, 1 000 a 10 000 objektů pro testování škálovatelnosti
3. **Shaderová komplexita:** Pokročilé techniky jako parallax mapping a procedurální animace
4. **Interaktivní scény:** Testování responsivity při uživatelské interakci

Standardní rozlišení pro všechny testy je **1920×1080 pixelů** (Full HD), které představuje nejčastěji používané rozlišení v současnosti.

**Typ testovaných objektů**

Základní testovací objekty jsou **krychlové geometrie** (cube), kde každý objekt obsahuje 12 trojúhelníků (6 stěn × 2 trojúhelníky). Scéna s 10 000 objektů tedy obsahuje přibližně **120 000 trojúhelníků**, což představuje realistickou zátěž pro střední až pokročilé 3D webové aplikace.

## Měřené metriky a nástroje

### Performance metriky

**Frames Per Second (FPS):**

* Měřeno pomocí knihovny **stats.js** integrované přímo do aplikace
* Vzorkování každý frame po dobu 2-3 minut pro stabilní průměr
* Zaznamenávání minimum, maximum, průměr a medián hodnot

**Memory Usage:**

* **Heap Memory:** Měřeno prostřednictvím Chrome DevTools Memory tab
* **GPU Memory:** Monitoring prostřednictvím browser performance API
* **Memory leaks detection:** Porovnání stavu paměti před a po testu

**Loading Performance:**

* **Asset loading time:** Doba načítání 3D modelů, textur a shaderů
* **Shader compilation time:** Měření doby kompilace vertex a fragment shaderů
* **Time to first frame:** Doba od inicializace do prvního vykresleného snímku

**Nástroje pro měření**

**Browser-integrated nástroje:**

* **Chrome DevTools Performance tab** pro detailní analýzu rendering pipeline
* **Firefox Developer Tools** pro cross-browser validaci výsledků
* **WebGL/WebGPU Inspector extensions** pro GPU-specific metriky

**Custom monitoring:**

* Rozšířená **stats.js** implementace s exportem dat
* Automatizované logování metrik do JSON formátu
* Real-time monitoring s možností exportu pro statistickou analýzu

## Statistická metodologie

**Validita a reprodukovatelnost**

**Počet opakování:** Každý test je opakován **10krát** pro zajištění statistické významnosti a eliminaci náhodných výkyvů.

**Statistické vyhodnocení:**

* **Medián** jako primární metrika (odolnost proti outliers)
* **Průměr** pro srovnání s průmyslovými standardy
* **Směrodatná odchylka** pro měření konzistence výkonu

**Threshold pro validitu:** Pokud směrodatná odchylka překročí 15% průměrné hodnoty, test je zopakován s analýzou příčin variance.

**Eliminace externích faktorů**

**Před každým testem:**

* Restart prohlížeče a vymazání cache
* Ukončení nepotřebných background procesů
* Standardizace system load (idle CPU < 5%)
* Kontrola dostupné RAM a GPU paměti

**Během testování:**

* Minimalizace systémových interrupts
* Monitoring system resources pro detekci anomálií
* Automatické zaznamenávání environment conditions

**Warm-up období a JIT optimalizace**

Moderní JavaScript enginy implementují Just-In-Time (JIT) kompilaci, která může ovlivnit výkonnostní metriky zejména v prvních sekundách běhu aplikace. Proto jsou všechna měření zaznamenávána od samého začátku testu, ale s možností dodatečné analýzy prvních 10-20 sekund pro identifikaci JIT efektů. V případě zjištění výrazných optimalizačních efektů budou tyto období vyloučeny ze základních statistik s odpovídajícím zdůvodněním.

## Dokumentace a reprodukovatelnost

1. **Zdrojový kód**

Veškerý zdrojový kód implementovaných experimentů je **veřejně dostupný** prostřednictvím Git repository. V textu práce jsou uvedeny klíčové fragmenty kódu (shadery, kritické algoritmy), zatímco kompletní implementace včetně benchmarking nástrojů je referencována s možností nezávislé verifikace výsledků.

1. **Výstupní data**

**Formát výsledků:**

* **Raw data** v JSON formátu pro možnost reanalýzy
* **Statistické souhrny** v tabulkové formě
* **Vizualizace** prostřednictvím grafů (SVG formát)
* **Screenshots** reprezentativních scén (PNG formát)

1. **Baseline pro srovnání**

Pro korektní interpretaci výsledků jsou implementovány **baseline testy** měřící overhead jednotlivých API bez významné grafické zátěže. Tyto testy využívají minimální scény (prázdná scéna nebo jednotlivý trojúhelník) pro izolaci čistého API performance od komplexity 3D obsahu.

Tato metodologie zajišťuje **objektivní**, **reprodukovatelné** a **statisticky validní** výsledky, které mohou sloužit jako základ pro praktická doporučení při výběru vhodné technologie pro konkrétní typy 3D webových aplikací.