Ovladače a architektura CUDA

CUDA (Compute Unified Device Architecture) je platforma pro paralelní výpočty vyvinutá společností NVIDIA. Umožňuje softwaru využívat grafický procesor (GPU) pro obecné výpočty, což je základem pro běh moderních velkých i malých jazykových modelů (LLM/SLM).

1. Vztah mezi Ovladačem a CUDA Toolkit

Je důležité rozlišovat mezi dvěma komponentami:

- **GPU Driver (Ovladač):** Základní software, který komunikuje přímo s hardwarem. Musí být nainstalován na hostitelském OS.
- **CUDA Toolkit:** Sada knihoven a nástrojů pro vývojáře. Obsahuje soubory potřebné k tomu, aby aplikace (např. PyTorch) věděla, jak instrukce na GPU posílat.

Kompatibilita (Forward/Backward)

Každá verze CUDA Toolkitu vyžaduje **minimální verzi ovladače**.

- Novější ovladač obvykle podporuje starší verze CUDA.
- Starý ovladač **nebude fungovat** s novou verzí CUDA (např. CUDA 13 vyžaduje ovladače z roku 2025/2026).

2. Tabulka verzí pro rok 2026

Verze CUDA Toolkit	Minimální verze ovladače (Windows)	Minimální verze ovladače (Linux)
CUDA 13.0	560.xx a vyšší	560.xx a vyšší
CUDA 12.6	551.xx a vyšší	550.xx a vyšší
CUDA 12.1	528.xx a vyšší	525.xx a vyšší

3. Instalace a ověření

Správný postup instalace:

```
1. Nejdříve aktualizujte **NVIDIA Driver** na nejnovější stabilní verzi
(Studio nebo Game Ready).
2. Restartujte systém.
3. Nainstalujte **CUDA Toolkit** (pokud jej vaše aplikace nevyžaduje již
"přibalený" uvnitř, jako to dělá např. instalátor Ollama).
```

Příkazy pro ověření stavu:

Chcete-li zjistit, co vidí systém, použijte v terminálu:

```
# Zobrazení verze ovladače a maximální podporované verze CUDA
nvidia-smi

# Zobrazení verze aktuálně nainstalovaného kompilátoru CUDA (Toolkit)
nvcc --version
```

> **Pozor:** Často se stává, že `nvidia-smi` ukazuje např. CUDA 12.6, ale `nvcc` ukazuje 11.8. To znamená, že váš ovladač sice moderní CUDA podporuje, ale máte nainstalované staré vývojářské nástroje.

4. Specifika pro AI aplikace (cuDNN)

Pro hluboké učení (Deep Learning) je k CUDA nutné doinstalovat ještě knihovnu **cuDNN** (CUDA Deep Neural Network library). Ta obsahuje optimalizované rutiny pro standardní operace, jako jsou konvoluce nebo aktivační funkce.

- Většina moderních frameworků (PyTorch, TensorFlow) si cuDNN stahuje automaticky v rámci svých balíčků (např. přes `pip install`).
- Pokud instalujete ručně, verze cuDNN musí přesně odpovídat verzi CUDA Toolkitu.

— **Související dokumentace:**

- Příručka pro aktualizaci ovladačů
- GPU v Dockeru
- Nastavení Python prostředí pro AI

— **Správce:** @SysAdmin_Team **Poslední revize:** Leden 2026

From:
https://www.serviceit.cz/ - **IT ENCYKLOPEDIE**

Permanent link:
https://www.serviceit.cz/doku.php?id=it:aktualizace_ovladacu_pro_ai

Last update: **2026/01/04 16:44**