

**Vymezení požadavků na výpočetní systém**  
**Příloha č. 1 Zadávací dokumentace veřejné zakázky**  
**„Výpočetní systém pro strojové učení a umělou inteligenci pro IT4Innovations“**

## **Obsah**

Vymezení požadavků na výpočetní systém Příloha č. 1 Zadávací dokumentace veřejné zakázky „Výpočetní systém pro strojové učení a umělou inteligenci pro IT4Innovations“ .....	1
1 Záměr zadavatele .....	1
2 Předmět zakázky.....	2
3 Legenda .....	2
4 Výpočetní systém .....	2
4.1 Požadavky – Hardware .....	2
4.2 Požadavky – Software .....	3
4.3 Požadavky – Instalace.....	4
5 Obecné požadavky .....	4
5.1 Dodávka .....	4
5.2 Dokumentace .....	5
5.3 Prohlášení o shodě .....	5
5.4 Likvidace odpadů.....	5
6 Infrastruktura zadavatele .....	5
6.1 Datové centrum.....	5
6.2 Rack pro instalaci výpočetního systému .....	5

### **1 Záměr zadavatele**

Záměrem zadavatele je modernizace a posílení kapacit výpočetní infrastruktury národního superpočítačového centra IT4Innovations (dále jen IT4Innovations), které zadavatel provozuje.

V rámci veřejné zakázky „Výpočetní systém pro strojové učení a umělou inteligenci pro IT4Innovations“ bude realizován nový specializovaný výpočetní systém určený pro provádění výpočetních úloh v oblastech strojového učení (machine learning) a umělé inteligence (artificial intelligence). Pořízovaný výpočetní systém je dále v dokumentu označován jako „výpočetní systém“ či jen „systém“.

Výpočetní systém bude provozován národním superpočítačovým centrem IT4Innovations a bude využíván k poskytování služeb v oblasti náročných výpočtů (High Performance Computing).

Výpočetní systém bude umístěn a provozován v datovém centru zadavatele, které se nachází v budově IT4Innovations, v areálu Vysoké školy báňské – Technické univerzity Ostrava.

## 2 Předmět zakázky

Předmětem veřejné zakázky „Výpočetní systém pro strojové učení a umělou inteligenci pro IT4Innovations“ je dodávka specializovaného výpočetního systému určeného pro provádění výpočetních úloh v oblastech strojového učení a umělé inteligence využívající technologii GPU, tj. dodávka hardware a software, instalace do datového centra zadavatele a poskytování servisních služeb.

## 3 Legenda

V následujícím textu jsou uváděny následující značky:

**SPEC\_číslo** označuje pro snazší identifikaci jednotlivé požadavky zadavatele veřejné zakázky.

**SPEC\_číslo (I)** označuje požadavek zadavatele veřejné zakázky na informaci, kterou dodavatel musí uvést v nabídce.

## 4 Výpočetní systém

### 4.1 Požadavky – Hardware

- SPEC\_1 Výpočetní systém musí sestávat z právě jednoho výpočetního serveru osazeného výpočetními GPU kartami, které jsou propojeny vysokorychlostní komunikační sítí.
- SPEC\_2 Výpočetní systém musí poskytovat agregovaný výpočetní výkon minimálně 2 Pflop/s v režimu FP16 (half precision floating point).
- SPEC\_3 Výpočetní systém musí poskytovat vysoce propustnou operační paměť (HBM) GPU karet o celkové velikosti minimálně 512 GB.
- SPEC\_4 Výpočetní server výpočetního systému musí splňovat následující požadavky:
- Architektura x86-64
  - Dva procesory na výpočetní server
  - Minimálně 24 fyzických CPU jader na procesor serveru
  - Teoretický výpočetní výkon procesoru serveru minimálně 2Tflop/s v režimu dvojité přesnosti (double precision floating point)
  - Kapacita operační paměti RAM serveru minimálně 1.5 TB
  - Dva lokální disky technologie NVMe SSD, každý o kapacitě minimálně 960GB (pro operační systém)
  - Lokální disky technologie NVMe SSD o celkové kapacitě minimálně 30TB (pro data)
  - Minimálně 8 portů Infiniband 100 Gb/s (pro připojení do výpočetní sítě zadavatele)
  - Minimálně 2 porty Ethernet 25/10 Gb/s (pro připojení do LAN sítě zadavatele)
- SPEC\_5 Výpočetní systém musí obsahovat minimálně 16 GPU karet určených pro výpočty, všechny GPU karty musí být shodného typu.
- SPEC\_6 Každá GPU karta musí splňovat následující požadavky:
- Podpora technologie CUDA
  - Teoretický výpočetní výkon minimálně 125Tflop/s v režimu FP16 (half precision floating point)
  - GPU paměť minimálně 32 GB HBM2
  - Propustnost GPU paměti minimálně 900 GB/s

- SPEC\_7 GPU karty výpočetního systému musí být vzájemně propojeny vysokorychlostní GPU sítí.
- GPU síť musí umožňovat současnou komunikaci libovolných GPU karet mezi sebou.
  - Každá GPU karta musí poskytovat celkovou teoretickou propustnost 150GB/s připojení do GPU sítě v každém směru, celkově 300GB/s v obou směrech.
  - Topologie GPU sítě musí být plně neblokující.
- SPEC\_8 Součástí dodávky musí být aktivní optické kabely (s příslušnými transceivery na obou koncích), případně sady modulů a optického kabelu, pro připojení do sítí zadavatele:
- 2 ks aktivní optické kabely QSFP+ FDR délky 30m - pro připojení do Infiniband výpočetní sítě zadavatele, Mellanox SX6025 FDR kompatibilní
  - 2 ks aktivní optické kabely SFP+ 10GbE délky 25m - pro připojení do Ethernet LAN sítě zadavatele, Cisco N7K-M224XP-23L kompatibilní
- SPEC\_9 (I) Dodavatel musí v nabídce uvést detailní konfiguraci hardware výpočetního systému – výpočetního serveru, GPU karet a GPU sítě.

## 4.2 Požadavky – Software

- SPEC\_10 Součástí dodávky musí být veškeré potřebné programové vybavení a licence potřebné pro legální použití výpočetního systému v plném rozsahu zadavatelem a uživateli služeb zadavatele. Programovým vybavením systémů se rozumí veškeré softwarové vybavení - firmware, ovladače hardware, operační systém, knihovny, softwarové stacky, aplikační software, management software, atd. výpočetního systému.
- SPEC\_11 Výpočetní server výpočetního systému musí používat operační systém typu Linux, 64-bitový, nativně používající instalační balíky typu RPM nebo DEB.
- SPEC\_12 Použitý operační systém musí mít dobrou dostupnost aktualizací, zejména těch, které jsou určeny ke zvýšení bezpečnosti a spolehlivosti.
- SPEC\_13 Výpočetní systém musí podporovat běh následujících open-source frameworků pro strojové učení a umělou inteligenci na GPU kartách systému:
- TensorFlow
  - torch
  - PyTorch
  - MXNet
  - theano
  - caffe, caffe2
- SPEC\_14 Výpočetní systém musí umožňovat běh kontejnerů (tj. použití virtualizace na úrovni operačního systému) typu Docker. Kontejnery musí podporovat GPU karty.
- SPEC\_15 Výpočetní systém musí poskytovat transparentní přístup do GPU pamětí GPU karet mezi GPU kartami prostřednictvím GPU sítě.
- SPEC\_16 (I) Dodavatel musí v nabídce uvést název operačního systému a popis software výpočetního systému.

### **4.3 Požadavky – Instalace**

- SPEC\_17 Dodavatel musí výpočetní systém (hardware a software) nainstalovat a zprovoznit v infrastruktuře zadavatele – v datovém centru IT4Innovations (tzv. instalace do datového centra zadavatele).
- SPEC\_18 Instalaci do datového centra zadavatele se rozumí veškeré dodávky a činnosti, jejichž výsledkem bude zprovoznění výpočetního systému v prostorách a infrastruktuře datového centra IT4Innovations.
- SPEC\_19 Instalace do datového centra zadavatele musí být provedena dle platné legislativy a předpisů a dle požadavků a doporučení výrobců zařízení.
- SPEC\_20 Dodavatel je povinen respektovat infrastrukturu zadavatele. Infrastruktura zadavatele je popsána v kapitole 6 Infrastruktura zadavatele.

#### **4.3.1 Požadavky - umístění**

- SPEC\_21 Všechna zařízení výpočetního systému budou dodavatelem nainstalována do racku zadavatele umístěného na datovém sálu zadavatele.

#### **4.3.2 Požadavky - napájení**

- SPEC\_22 Zařízení výpočetního systému budou napájena z PDU v racku zadavatele určeného pro instalaci. Součástí dodávky jsou potřebné napájecí kabely. V případě, že řešení PDU v racku zadavatele nevyhovuje nabízenému řešení, dodavatel jako součást dodávky zajistí vyhovující řešení PDU.

#### **4.3.3 Požadavky - chlazení**

- SPEC\_23 Chlazení tepla vyprodukovaného zařízeními výpočetního systému bude zajišťovat infrastruktura zadavatele.

#### **4.3.4 Požadavky - transport**

- SPEC\_24 Při transportu a instalaci zařízení nesmí být překročena únosnost podlahy přepravní trasy. Únosnost podlahy datového sálu a jeho přístupové chodby (místnosti 219 a 223) je 25kN/m<sup>2</sup>.  
Únosnost podlahy přístupových prostor (místnosti 217 a 218) je 5kN/m<sup>2</sup>. Tento úsek musí být pro transport materiálu dodavatelem dočasně ošetřen instalací ochranných desek pro rozklad zatížení (kupříkladu položením překližkových desek), tak aby výsledné zatížení podlahy v průběhu fyzické dodávky nepřekročilo 5kN/m<sup>2</sup> a předešlo se tak poškození podlahy.

#### **4.3.5 Požadavky - sítě**

Žádné požadavky. Připojení výpočetního systému do sítí zadavatele provede zadavatel za použití dodavatelem dodaných kabelů (viz SPEC\_8).

## **5 Obecné požadavky**

### **5.1 Dodávka**

- SPEC\_25 Součástí dodávky musí být doprava, instalace, zprovoznění systému a předvedení funkčnosti systému.

## 5.2 Dokumentace

SPEC\_26 Součástí dodávky musí být poskytnutí dokumentace dodávaného hardware a software (manuály) v anglickém jazyce.

## 5.3 Prohlášení o shodě

SPEC\_27 Ke všem dodaným zařízením musí být doloženo prohlášení o shodě.

## 5.4 Likvidace odpadů

SPEC\_28 Součástí dodávky musí být likvidace veškerých odpadů vzniklých realizací dodávky. Zadavatel nebude uchovávat obaly, obalový materiál.

# 6 Infrastruktura zadavatele

Výpočetní systém bude instalován a provozován v datovém sálu datového centra zadavatele v budově IT4Innovations. Budova IT4Innovations se nachází v areálu kolejí Vysoké školy báňské – Technické univerzity Ostrava, na adrese Studentská 6231/1B, 708 00 Ostrava-Poruba.

## 6.1 Datové centrum

Datové centrum zajišťuje redundantní zdroj zálohovaného napájení (dvě nezávislé napájecí větve, redundance 1+1) rozvodné soustavy 400/230V;3+N+PE; 50Hz; TN-S.

Chlazení pro datový sál poskytuje pět samostatných chladících kapalinových okruhů - dva okruhy teplé vody teploty (cca 30°C) a tři okruhy studené vody (cca 11,5°C), je použita směs 35% propylenglykol a 65% voda.

Datový sál o výměře 511m<sup>2</sup> poskytuje světlou výšku mezi systémovou podlahou a stropním podhledem 2.9m.

## 6.2 Rack pro instalaci výpočetního systému

Pro účely instalace výpočetního systému je k dispozici jeden standardní 19“ EIA-310 rack 42U (rack unit) umístěný na datovém sálu zadavatele. Využitelná hloubka racku je cca 90cm.

Rack je vybaven dvěma jednofázovými PDU (Power Distribution Unit) EATON ePBZ31, každé na jiném okruhu a jiné fázi. PDU má vstupní parametry 240V / 32A. PDU má 20ks C13 zásuvek (10A) a 4ks C19 zásuvek (16A) (ve dvou stejných skupinách po 10ks C13, 2ks C19). PDU je vybaveno dvěma jističi 16A typu C. PDU jsou v racku umístěna ve svislé poloze v zadní části po stranách racku.

Rack je chlazen systémem chlazení v zadních dveřích racku připojeným na okruh studené vody, chladící výkon je dimenzován pro maximální tepelný výkon 25 kW/rack.