

## Výpočetní cluster pro porozumění jazyku (LUCC)

Jedná se o dodávku čtyř serverů pro projekt, zabývající se porozumění přirozenému jazyku.

Dva servery jsou určeny na zpracování neuronových sítí pracovištěm vyvinutými nástroji. Pro účely specifikace platformy uvažujeme zadavatelem již vlastněné GPU karty Nvidia RTX s referenčním nebo pasivním chlazením a bez omezení použití v datových centrech – tedy profesionální karty, určené na výpočty.

Dva další servery jsou určeny k vytvoření LustreFS svazku se ZFS filesystemem na OSS serverech. MDS server má zadavatel k dispozici. Uvedená dvojice serverů bude sloužit jako OSS servery, postavené na NVMe technologii.

Instalaci software provede zadavatel. Součástí dodávky je návrh a kompletace dodávaných strojů, jejich dodání a zajištění požadovaných záručních podmínek.

Součástí dodávky nejsou rackové skříně ani jiné, v zadávací dokumentaci neuvedené komponenty.

**Nepřekročitelný cenový limit: 1.832.231,- Kc bez DPH**

Je požadováno dodání následujících systémů:

### I. Dva servery pro zpracování neuronových sítí

- Provedení, určené pro montáž do skříně Rack, dodání včetně lyžin.
- Prostorové nároky: max 4U
- 2x CPU s podporou HT, 20 jader (celkem tedy 40 fyzických / 80 logických jader). Frekvence CPU minimálně 2,2GHz, technologie Deep Learning Boost, podpora pro virtualizaci, instrukční sada AVX-512, výrobní proces procesoru 14nm nebo lepší.
- Požadovaná architektura CPU je x86\_64.
- výkonnost procesoru alespoň 7.8 v SPECint2017, sloupec Base dle spec.org: <https://www.spec.org/cpu2017/results/cint2017.html>
- výkonnost serveru (2 CPU) alespoň 110 v SPECint\_rate2017, sloupec Base dle spec.org: <https://www.spec.org/cpu2017/results/rint2017.html>
- 8x GPU akcelerator. Jsou požadovány GPU akcelerátory minimálně s čipem Nvidia Turing. Konkrétní výrobce GPU karet není určen, tento čip je vyžadován vzhledem k nutnosti zachovat homogenní výpočetní prostředí a kompatibilitu s existujícím programovým vybavením. Kritickým parametrem je podpora CUDA 10.0.
- Parametry pro GPU: paměť 16GB GDDR6, 256-bitové paměťové rozhraní, Nvidia NVLink, 3072 CUDA jader, 384 Tensor jader, 48 RT jader 12, referenční chlazení, 12 nm technologie, EULA nesmí omezovat použití karet v datových centrech!
- 2x hot-swap SSD, minimální kapacita každého z nich 240GB, DWPD minimálně 3,6 po dobu 5 let
- Paměť specifikace minimálně DDR4 ECC Registered, 2666MHz. Servery musí mít osazeno minimálně 384GB RAM
- Paměťové moduly mohou být maximálně dual rank.
- Modul vzdálené správy přes internet (zařízení musí umožňovat KVM-over-LAN, pro tuto funkci musí mít vlastní síťový konektor).
- 2x rozhraní Ethernet 10Gbps SFP+
- Alespoň 1x 1Gbps Ethernet port (nebo zpětně kompatibilní 10G-BaseT)
- 1x 3m SFP+ 10Gbps kabel
- Je požadováno redundantní napájení minimálně N+1 (výpadek jednoho zdroje nezpůsobí výpadek serveru), zdroje hot-swap (vyměnitelné za běhu systému), certifikace zdrojů 80 PLUS Platinum.

- Bootování operačního systému: konfigurovatelné pořadí zařízení, podpora bootování přes USB nebo DVD, a podpora bootování ze vzdáleného iso obrazu prostřednictvím vzdáleného managementu (Baseboard Management Controller (BMC), prostřednictvím Kernel-based Virtual Machine (KVM) po LAN).
- Je vyžadováno vzdálené ovládání vypnutí/zapnutí/reset, vzdálená sériová konzole (serial-over-lan) a konzole KVM - vše dostupné přes LAN nástroji pro operační systém Linux. Funkcionalita vypnutí/zapnutí/reset musí být dostupná nástroji na příkazové řádce použitelnými ve skriptu.
- Všechny jednotky/servery musí být schopny plnohodnotného provozu v operačním systému (OS) Linux. OS Linux je v současné době používán na všech výpočetních kapacitách instalovaných na pracovišti zadavatele. Jedná se o rozšířený a bezplatný OS. Použití jiného OS by znamenalo velmi významné zvýšení nákladů na instalaci a správu výpočetních kapacit. Zadavatel vyvíjí vlastní software pro výzkumné účely, který je závislý na operačním systému Linux jakožto jediné kompatibilní platformě. Používané distribuce jsou Ubuntu 18.04 v 64-bitové verzi.
- Záruční doba na server: minimálně 3 roky, next-business-day, on-site
- Záruční doba na GPU karty: minimálně 3 roky, mail-in

## II. Dva OSS servery pro LustreFS

- Provedení, určené pro montáž do skříně Rack, dodání včetně lyžin.
- Prostorové nároky: 1U na řešení
- 1x CPU s podporou HT/CMT, 8 jader s podporou HT/CMT (16 vláken)
- Požadovaná architektura CPU je x86\_64.
- výkonnost procesoru alespoň 7.53 v SPECint2017, sloupec Base dle spec.org:  
<https://www.spec.org/cpu2017/results/cint2017.html>
- výkonnost serveru (1 CPU) alespoň 55 v SPECint\_rate2017, sloupec Base dle spec.org: <https://www.spec.org/cpu2017/results/rint2017.html>
- 2x hot-swap SSD, kapacita každého z nich alespoň 240GB, DWPD minimálně 3.6 po dobu 5 let
- PCIe x16, gen 4
- podpora minimálně pro 10 NVMe
- osazeno 8x hot-swap minimálně 3.84TB NVMe, TLC, DWPD minimálně 1.0 po dobu 5 let
- Paměť specifikace minimálně DDR4 ECC Registered, 3200MHz. Servery musí mít osazeno alespoň 256GB RAM
- Paměťové moduly mohou být maximálně dual rank. Všechny kanály musí být obsazeny. Paměťová kapacita serveru musí být rozšiřitelná s použitím identických paměťových modulů až na 768GB.
- Modul vzdálené správy přes internet (zařízení musí umožňovat KVM-over-LAN, pro tuto funkci musí mít vlastní síťový konektor).
- Alespoň 1x rozhraní 1Gbps Ethernet
- 2x rozhraní 25GbE SFP28
- 1x 2m 10GE SFP28-SFP28 kabel
- Je požadováno redundantní napájení minimálně N+1 (výpadek jednoho zdroje nezpůsobí výpadek serveru), zdroje hot-swap (vyměnitelné za běhu systému), certifikace zdrojů 80 PLUS Platinum nebo vyšší.
- Bootování operačního systému: konfigurovatelné pořadí zařízení, podpora bootování přes USB nebo DVD, a podpora bootování ze vzdáleného iso obrazu prostřednictvím vzdáleného managementu (Baseboard Management Controller (BMC), prostřednictvím Kernel-based Virtual Machine (KVM) po LAN).
- Je vyžadováno vzdálené ovládání vypnutí/zapnutí/reset, vzdálená sériová konzole (serial-over-lan) a konzole KVM - vše dostupné přes LAN nástroji pro operační systém Linux. Funkcionalita vypnutí/zapnutí/reset musí být dostupná nástroji na

## Příloha č. 1 – Technická specifikace

- příkazové řádce použitelnými ve skriptu.
- Všechny jednotky/servery musí být schopny plnohodnotného provozu v operačním systému (OS) Linux. OS Linux je v současné době používán na všech výpočetních kapacitách instalovaných na pracovišti zadavatele. Jedná se o rozšířený a bezplatný OS. Použití jiného OS by znamenalo velmi významné zvýšení nákladů na instalaci a správu výpočetních kapacit. Zadavatel vyvíjí vlastní software pro výzkumné účely, který je závislý na operačním systému Linux jakožto jediné kompatibilní platformě. Používané distribuce jsou Ubuntu 16.04 v 64-bitové verzi.
  - Záruční doba minimálně 3 roky, next-business-day, on-site

