

Technická dokumentace - specifikace požadovaného plnění

Výpočetní uzly z projektu *PhD Infra ZČU*

1. Základní informace a požadavky na výpočetní uzly

1.1 **Obecné požadavky**

- 1.1.1 **Předmětem** plnění veřejné zakázky je kompletní řešení, sestávající se z dodávky, instalace a zprovoznění **uzlů (serverů) výpočetního clusteru** v konfiguraci podle specifikace uvedené dále.
- 1.1.2 Řešení musí obsahovat výpočetní uzly následující specifikace:
 - 1.1.2.1 **4** identické uzly v lokalitě ZČU
(*Identickým uzlem* se rozumí osazení zcela shodných komponent do každého uzlu, včetně použitých pamětí)
- 1.1.3 Součástí plnění je také dodávka a instalace potřebné síťové infrastruktury zahrnující kabeláž.
- 1.1.4 Součástí předmětu plnění je rovněž poskytnutí záruky za jakost a řádnou funkčnost dodaného plnění, včetně technické podpory (dále také jen „záruka“) na dobu alespoň **36 měsíců** s reakcí následující pracovní den a opravou v místě instalace do 14 dnů od nahlášení.
- 1.1.5 **Instalací zprovozněním** se rozumí instalace hardware do stávajících rack skříní, zapojení všech síťových rozhraní, zapojení do elektrické sítě, nastavení BIOSu a BMC dle pokynů zadavatele a spuštění hardware a ověření bezchybného chodu všech komponent.

1.2 **Detailní technické požadavky na uzly clusteru - každý výpočetní uzel musí splňovat tyto podmínky:**

- 1.2.1 Provedení do standardního 19" racku, velikost uzlu musí být maximálně 2U. Servery budou umístěny do stávajících racků, racky nejsou součástí zakázky.
- 1.2.2 Sdílení komponent mezi více uzly clusteru není povoleno.
- 1.2.3 Uzel (výpočetní jednotka se samostatnou pamětí, chipsetem, procesory, diskem, atd.) musí mít procesory se sdílenou pamětí v architektuře x86_64 a celkem minimálně 128 fyzických jader.
- 1.2.4 Minimální požadovaný výkon uzlu měřený nástrojem SPECfp2017 ve variantě FP, rate, baseline je 1180 bodů. Zároveň výkon v tomto benchmarku, přepočtený na jedno jádro CPU, tj. výkon celého uzlu vydělený počtem fyzických jader v uzlu (počítají se pouze fyzická jádra, nikoli technologie hyperthreading), musí být alespoň 9.
- 1.2.5 Operační paměť uzlu musí být alespoň 1.5TB DDR5 ECC. Rychlost pamětí nesmí být horší než rychlost paměti použité ve SPEC benchmarku v bodu 1.2.4. Na všech použitých paměťových kanálech musí být stejná skladba DIMMů, všechny paměťové kanály musí být obsazené.
- 1.2.6 Každý uzel clusteru musí mít přístup k lokálnímu NVMe prostoru, určenému pro operační systém a lokální dočasná data, vše realizováno alespoň dvěma totožnými SSD NVMe disky s celkovou kapacitou alespoň 15TB. Součet rychlostí lineárního čtení/zápisu všech disků musí být alespoň 12/6 GB/s, součet I/O výkonu všech disků musí být alespoň 2000k/300k IOPS pro náhodné čtení/zápis. Součet výdrže disků musí být alespoň 1.6 PB zápisů (1600TBW) na každý TB celkové kapacity všech disků, případná reklamace

disku nesmí být zamítnuta kvůli jeho opotřebení zápisy, pokud dodavatel neprokáže, že byl překročen alikvotní podíl z celkové výdrže odpovídající podílu jednotlivého disku na celkové kapacitě všech disků.

- 1.2.7 Uzel musí mít alespoň jedno 10G ethernet rozhraní typu SFP+ umožňující 10G připojení do switche Cisco Catalist řady 9500, který je součástí stávající infrastruktury. Ethernetová rozhraní musí podporovat UDP TNL offload v linuxovém ovladači (příznaky NETIF_F_GSO_UDP_TUNNEL a NETIF_F_GSO_UDP_TUNNEL_CSUM) a bootování přes PXE.
- 1.2.8 Uzel musí obsahovat minimálně dva GPU akcelerátory s podporou CUDA API a tenzorových jader. Každá karta musí být v provedení pro rackové servery a splňovat následující parametry: teoretický výkon minimálně 90 TFLOPS v FP32 a paměť alespoň 48 GB.
- 1.2.9 Uzel musí mít redundantní napájení, zdroje musí mít certifikaci 80 PLUS Titanium.
- 1.2.10 Nabízený procesor i další komponenty (motherboard, BIOS) musí podporovat virtualizaci, včetně virtualizace I/O (např. VT-d v terminologii firmy Intel, AMD-Vi v terminologii firmy AMD).
- 1.2.11 Výkon SPEC benchmarku požadovaný v bodech 1.2.4 lze prokázat předložením oficiálního výsledku z webu www.spec.org dosaženého na ekvivalentním stroji (typ a počet procesorů totožný s dodanými procesory; počet a frekvence paměťových DIMMů a organizace paměti totožné s dodanou pamětí, velikost paměti se může lišit) nebo výsledkem spuštění benchmarku na uzlu nakonfigurovaném dle uvedené technické specifikace. V případě pochybností ohledně deklarovaného SPEC skóre, je dodavatel povinen na vyzvání přeměřit výkon dodaného HW samostatným během SPEC benchmarku.
- 1.2.12 Všechny disky musí podporovat vyčtení provozních stavů a statistik pomocí standardu SMART, být měnitelné za chodu (hot-swap) a přístupné zepředu.
- 1.2.13 Uzel clusteru musí umožňovat přístup ke konzoli (klávesnice + monitor) jak lokálně tak po síti (síťový KVM nebo BMC).
- 1.2.14 Uzel musí podporovat bootování z externího zařízení a to jak lokálně (boot z USB – CD-ROM, flash disk, harddisk), tak po síti (síťový KVM nebo BMC).
- 1.2.15 Základní deska musí umožňovat změnu pořadí bootovacích zařízení.
- 1.2.16 Základní deska musí obsahovat management controller (BMC) kompatibilní se specifikací IPMI 2.0 nebo vyšší. BMC musí umět monitorovat minimálně funkčnost ventilátorů a zdroje, aktuální spotřebu, teplotu CPU a základní desky; dále musí BMC poskytovat základní vzdálený power management (vypnout, zapnout, reset). Požadujeme možnost změny bootovacího zařízení vzdáleně pomocí BMC nebo KVM.
- 1.2.17 Grafické rozhraní KVM (BMC nebo externího) musí fungovat ve webovém prohlížeči s podporou HTML5.
- 1.2.18 Vzdálený přístup k BMC musí být realizovatelný buď po sdíleném síťovém kabelu nebo samostatným kabelem. V případě použití samostatného kabelu pro BMC je dodavatel povinen dodat i switche s minimálním počtem portů odpovídajícím počtu uzlů plus 3 a rychlostí portů alespoň 1Gb/s.

Umístění switche se předpokládá ve stejném racku jako instalované uzly.

1.3 Požadavky na síťové komponenty

- 1.3.1 Součástí nabídky musí být tato propojovací kabeláž pro připojení všech uzlů do síťových switchů:

- 1.3.1.1 10G SFP+ ethernet kabely pro běžný provoz. Kalkulujte délku 20m. Podle konkrétní situace na sále a polohy v racku je následně žádoucí dodat kabely na míru. Kabely musí být kompatibilní se switchi Cisco Catalyst řady 9500.
- 1.3.1.2 Je-li použita varianta samostatného kabelu pro IPMI (viz 1.2.18) tak i 1G kabely pro BMC. Kabely kalkulujte na míru, při instalaci je žádoucí, aby nevznikaly zbytečné kabelové smyčky.

2. Další požadavky:

- 2.1 Veškerá zařízení by mělo být možno koupit a používat bez jakéhokoliv přídavného software (externího, nepočítá se firmware, jehož trvalá licence je součástí dodávky zařízení). Pokud je programové vybavení nutnou součástí nabízeného plnění (například SW pro vzdálenou správu), musí být jasně specifikovány důvody a cena za oprávnění k užívání takového software musí být zahrnuta do nabídkové ceny (oprávnění bude uděleno na dobu neurčitou, resp. na dobu trvání majetkových práv autora).
- 2.2 Všechny komponenty, které jsou touto technickou specifikací požadovány, musí být použitelné v prostředí operačního systému Linux (zejména, ale nikoliv výhradně 64bit Debian), tj. musí být podporovány distribučním nebo originálním jádrem nebo s využitím externích ovladačů dostupných ve zdrojovém kódu.
- 2.3 V nabídce musí být uvedena celková maximální spotřeba sestavy (maximální spotřeba odpovídá spotřebě při plném zatížení všech komponent).
- 2.4 V nabídce musí být uvedena velikost celé instalace v počtu U ve standardním 19" racku.
- 2.5 Všechna zařízení instalovaná v racku musí mít výstup teplého vzduchu směřovaný dozadu racku.
- 2.6 Dodavatel musí nakonfigurovat EFI/BIOS všech serverů jednotně dle požadavku zadavatele tak, aby stroj bylo možné nabootovat z PXE a následně z pevného disku.
- 2.7 U zařízení umožňujících vzdálenou správu předá dodavatel zadavateli seznam s umístěním těchto zařízení v racku (pozice v racku, MAC adresa) a provede nastavení BMC dle pokynů zadavatele (jednotné heslo, IP adresa z dodaného rozsahu, VLAN...) tak, aby bylo možné stroje ovládat na dálku.
- 2.8 Vybraný dodavatel je povinen na výzvu zadavatele prokázat původ dodaného plnění.
- 2.9 Veškeré vyvázání kabeláže bude provedeno způsobem umožňujícím opakované beznástrojové rozebrání a znovupoužití.
- 2.10 Veškeré kabely musí být opatřeny na obou koncích shodným, v rámci instalace jednoznačným, označením, aby bylo možné identifikovat odkud kam vedou bez nutnosti procházet celou kabelovou trasu.
- 2.11 Instalace do volných pozic ve stávajících stojanech (rackách) dle domluvy s odběratelem v místě instalace.
- 2.12 Pro instalaci je k dispozici rack Rital celkové výšky 42U a vnější šířky 80 cm a hloubky 95cm. Proudění chladného vzduchu je ve směru z přední do zadní části racku. Napájení je redundantní, rack je v zadní části vybaven dvěma vertikálními PDU 3x16A/230V (na každé fázi je 8xC13).
- 2.13 Komponenty, u kterých jsou požadovány redundantní napájecí zdroje musí být schopny trvalého běhu jen na jednu větev a celá sestava nesmí překročit proudový limit ani při zapnutí po neřízeném výpadku napájení.
- 2.14 Nabídka musí obsahovat seznam komponent s jasným označením typů dílů, tak aby bylo možné z otevřených zdrojů ověřit, že splňují požadované technické parametry.

Alternativně je možné požadované parametry doložit pomocí datasheetů od výrobce.