



TECHNICKÁ SPECIFIKACE POŽADOVANÉHO PLNĚNÍ

Rozšíření clusteru datových služeb se bude skládat z následujících hlavních komponent:

1. **Diskový subsystém s čistou kapacitou alespoň 2000 TB**
2. K tomuto subsystému **nejméně dva front-end servery** pro jeho zpřístupnění protokolem CIFS a NFSv4 v režimu active-active z clusterového souborového systému (konfigurace SW pro export protokolem CIFS a NFSv4 není součástí dodávky).
3. Tento systém bude dodaný bez serverových stojanů (racků). Musí se vejít nejvýše do dvou standardních 42U racků zadavatele (Rittal TS-IT).

Hlavní komponenty budou propojeny prostřednictvím SAS technologie. SAS infrastruktura bude součástí dodávky. Dodané zařízení musí být plně funkční jako celek a splňovat všechny dále uvedené požadavky.

Požadavky zadavatele na jednotlivé části:

1. Architektura:

- 1.1 Servery a diskové subsystémy musí být samostatné jednotky propojené technologií SAS.
- 1.2 Propojení serverů a polí musí být realizované redundantně, propojení musí být realizováno tak, aby výpadek jakékoliv propojovací cesty neznamenal výpadek celého systému.
- 1.3 Veškeré kabely nezbytné pro zajištění propojení SAS infrastruktury musí být součástí dodávky.
- 1.4 Součástí dodávky musí být veškeré nezbytné kabely pro zajištění Ethernet konektivity serverů i polí. V případě serverů musí ke každému serveru být dodán jeden 10Gb optický kabel pro rozhraní 10GBASE-LR, jeden kabel pro 1Gb Ethernet rozhraní a jeden kabel pro management rozhraní. V případě polí musí být dodán nejmeně jeden kabel pro každý kontroler pole pro 1Gb Ethernet rozhraní. Maximální očekávaná délka Ethernet kabelů je 20 metrů, po dohodě se zadavatelem budou dodány kabely na míru. Ke každému serveru musí být navíc dodán SFP+ modul do serveru a SFP+ modul do switchu, modul do switchu musí být stoprocentně kompatibilní s HP A5820X-24XG-SFP+ JC102A, moduly musí být typu LR.

2. Front-end servery musí splňovat tyto podmínky:

- 2.1 Nejmeně dva front-end servery, které disková pole s kapacitou 2000 TB zpřístupní protokolem CIFS a NFSv4 jako jeden nebo několik filesystémů. Veškeré požadavky na servery včetně výkonostních musí být splněny všemi (oběma) front-end servery.
- 2.2 Každý server musí mít minimálně dvanáct procesorových jader se sdílenou pamětí v architektuře x86_64. Výkon celého serveru ve SpecINT2006 rate, baseline musí být alespoň 450 bodů.
- 2.3 Každý server musí mít alespoň 128 GB RAM ECC.
- 2.4 Každý server musí být osazen dvěma systémovými disky s kapacitou alespoň 120 GB každý, rychlost otáčení ploten minimálně 10000RPM případně SSD srovnatelných parametrů. V případě SSD je požadováno DWPD alespoň 2.
- 2.5 Každý server musí mít aspoň jednu jednoportovou 10Gb Ethernet kartu s optickým rozhraním 10GBASE-LR a zároveň alespoň jedno 1Gb Ethernet rozhraní s možností PXE bootu.

- 2.6 Každý server musí mít SAS rozhraní pro připojení do SAS infrastruktury o celkové rychlosti alespoň 32Gb/s.
- 2.7 Servery musí mít duální napájení. Zdroje i disky musí být vyměnitelné za chodu.
- 2.8 Server musí umožňovat centralizovaný přístup ke konzoli (klávesnice + monitor) a zároveň musí podporovat bootování z externího zařízení. Externím zařízením se rozumí jak lokální (KVM switch, boot z USB – CD-ROM, flash disk, harddisk), tak síťové (síťový KVM nebo BMC, boot z virtuálního média).
- 2.9 Servery musí umožňovat změnu pořadí bootovacích zařízení.
- 2.10 Servery musí obsahovat management controller (BMC) kompatibilní se specifikací IPMI 2.0 nebo vyšší. BMC musí umět monitorovat minimálně funkčnost ventilátorů, 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.
- 2.11 Funkcionalita IPMI musí být přístupná z příkazové řádky běžící na vzdáleném linuxovém systému připojeném k BMC přes LAN.

3. Diskový subsystém musí splňovat tyto podmínky:

- 3.1 Diskový subsystém se bude skládat z alespoň dvou nezávislých diskových polí. Všechna disková pole musí být identická (vč. počtu disků, typu, počtu a konfigurace RAID skupin).
- 3.2 Kapacita polí bude složena z klasických rotačních disků (SAS, SATA nebo FC).
- 3.3 Celková kapacita (součet velikostí blokových zařízení exportovaných z diskových polí na server) musí být minimálně 2000 TB. Zabezpečení rotačních disků musí být pomocí RAID 6 v konfiguraci 16+2 nebo lepší, např. 15+2, nebo pomocí ekvivalentní technologie se stejnou úrovní zabezpečení (počet paritních disků). RAID skupin může být v poli více, na front-end serverech mohou být softwarově spojeny do jediného blokového zařízení pomocí softwarového RAID 0 nebo spojením za sebe. Všechny RAID skupiny musí být nakonfigurovány stejně a musí být realizovány pomocí externího kontroleru. SW RAID ani RAID realizovaný na HBA kartě na front-end serveru není přípustný. Všechny RAID skupiny musí být připojitelné na všechny front-end servery. RAID musí být nakonfigurován tak, aby rebuild neběžel více jak 48 hodin (během plného provozu, je přípustná degradace výkonu).
- 3.4 Do kapacity 2000 TB nejsou počítány paritní ani hot-spare disky.
- 3.5 Každé diskové pole musí obsahovat nejméně 2 hot-spare rotační disky na každých 30 disků, přidělitelné k libovolnému RAIDu z rotačních disků.
- 3.6 Každé diskové pole musí obsahovat minimálně dva diskové řadiče funkční v režimu active-active.
- 3.7 Je požadována plná redundance komponent diskových polí, včetně řadičů, zdrojů napájení, ventilátorů a SAS řadičů (v diskových serverech i polích, v serverech lze využít dualportových SAS karet).
- 3.8 Minimální velikost cache každého z řadičů 8 GB. Celý diskový systém musí mít alespoň 32 GB využitelné cache při současném splnění bodu 3.7, tj. je-li bod 3.7 realizován cache mirroringem, musí mít systém alespoň 64 GB cache.
- 3.9 Zabezpečení cache hardwarových RAID řadičů při výpadku proudu a poruše jednoho z řadičů (flash RAM nebo kapacitory).
- 3.10 Každý řadič každého pole musí mít alespoň jedno SAS propojení do každého front-end serveru
- 3.11 Firmware diskových polí musí být možné aktualizovat bez výpadku pole.
- 3.12 Zdroje, ventilátory, disky a řadiče musí být typu hot-swap (musí být vyměnitelné bez přerušení provozu pole).
- 3.13 Vzdálený management a monitoring serverů i diskových polí, varování o poruchách disků a řadičů pomocí SNMP zpráv. Vzdálený management musí být plně použitelný z Linuxu a musí být realizován, jak pomocí CLI, tak pomocí webového prohlížeče.
- 3.14 Disková pole musí být redundantně připojena ke každému front-end serveru s využitím dodávaných SAS propojů.

4. Ostatní

- 4.1 Součástí nabídky musí být spotřeba zařízení v jednotlivých serverových stojanech a celková maximální spotřeba sestavy (maximální spotřeba odpovídá spotřebě při plném zatížení všech komponent, tedy všech front-end serverů a diskových polí).
- 4.2 Všechny komponenty, které jsou touto technickou specifikací požadovány, i front-end servery a všechny jeho komponenty musí být použitelné v prostředí operačního systému Linux (zejména, ale nikoliv

výhradně Debian, CentOS), tj. musí být podporovány distribučním nebo originálním jádrem nebo s využitím externích ovladačů dostupných ve zdrojovém kódu.

- 4.3 Diskové pole a front-end servery musí být ve formátu vhodném pro montáž do standardního 19" racku zadavatele (Rittal TS-IT), výška racku je 42U.
- 4.4 Zatížení jednoho racku musí být nejvýše 1500 kg. Maximální celkový příkon zařízení v jednom racku musí být nejvýše 18 kW.

5. Měření výkonu

Součástí nabídky budou výkonnostní testy dle následujícího popisu:

- 5.1 Sestava s kapacitou 2000TB musí poskytovat celkovou průchodnost alespoň 3000000 kB/s při sekvenčním čtení 16 velkých souborů ze všech uzlů a 2000000 kB/s při sekvenčním zápisu 16 velkých souborů ze všech uzlů, všemi uzly se rozumí uzly dle sekce 2.
- 5.2 Rychlost úložiště bude měřena na všech front-endech dodaných v konfiguraci dle sekce 2. Rychlost pro sestavu bude měřena nad dvěma nebo více XFS svazky lokálně připojenými na dodané front-end servery (vždy jeden svazek připojený k jednomu front-end serveru).
- 5.3 Konfiguraci operačního systému pro účel akceptačních testů a jejich předvedení zajistí dodavatel.
- 5.4 Ověření výkonu bude prováděno pomocí iozone -t 16 -Mce -s1000g -r256k -i0 -i1 -+m nodes. Kde nodes popisuje jednotlivé front-endy v souladu s dokumentací nástroje iozone. Podstatné pro průchodnost jsou údaje „Children see throughput for 16 initial writers” (pro zápis) a „Children see throughput for 16 readers” (pro čtení).
- 5.5 Dodavatel v akceptačních testech demonstruje požadovanou průchodnost úložiště na dodané sestavě nakonfigurované dle výše uvedené technické specifikace. Je-li bod 3.8 splněn pomocí cache mirroring, musí být v průběhu testů cache v tomto režimu zapnutá.
- 5.6 Pro rychlost úložiště je pro zadavatele podstatná rychlost naměřená programem iozone (verze 3.x z <http://www.iozone.org>). **Výstup programu iozone je nutné přiložit k nabídce.**