

# Thor

## Hromovládce datových serverů

Po uvedení vysoce úspěšného storage serveru X4500 („Thumper“) přichází Sun Microsystems s jeho inovovaným následovníkem, storage serverem X4540 s kódovým označením Thor.



Stejně jako bájný Thor ze severského panteonu, i tento Thor bude mít svého „nohsleda“, a to rozšiřující JBOD modul J4500, nebo též „Loki“. X4540, stejně jako jeho předchůdce X4500, nabízí mohutnou výpočetní sílu zkombinovanou s masivní diskovou kapacitou v podobě 48 disků o velikosti až 1TB. To vše v pouhých čtyřech RU výšky. Po hardwarové stránce disponuje X4540 dvěma čtyřjádrovými procesory AMD Opteron 2356 s frekvencí 2.3GHz. Oproti svému předchůdci nabízí rozšiřitelnost paměti až na 64GB při použití 4GB DIMM modulů. Asi nejdůležitější vylepšení se však skrývá v IO subsystému. Thor nyní používá Nvidia čipset, který má pro IO vyčleněno devět PCIExpress x8 linek. Šest z nich vede na šest SATA řadičů LSI, ke kterým jsou připojeny disky – osm na každý řadič. Zbývající tři PCIExpress linky jsou vyvedeny na externí sloty, kde jsou využitelné pro různé rozšiřující karty – zejména pro SAS adaptér a připojení externích JBOD modulů. Server X4540 tak vykazuje oproti svému předchůdci X4500 téměř trojnásobný nárůst propustnosti!

Management serveru je, podobně jako u ostatních opteronových strojů z produkce Sun Microsystems, zajištěn ILOM rozhraním, které je realizováno vestavěným servisním procesorem. Toto rozhraní je přístupné pomocí seriové linky nebo ethernetového portu. Mezi poskytované funkce patří zejména možnost sledování vnitřních senzorů stroje, vzdálené vypínání a zapínání serveru nezávisle na OS, a také možnost vytvořit vzdálenou konzoli, včetně připojení vzdálené DVD mechaniky pro pohodlnou instalaci OS z dohledové konzole.

### Loki, aneb když 48TB nestačí

Nevýhodou předchozí generace storage serveru byla nemožnost kapacitního rozšíření, neboť neexistoval vhodný JBOD (just-a-bunch-of-disks) modul. Nyní odpadá i tato nevýhoda. Chystá se totiž uvedení tří JBOD zařízení pod souhrnným označením J4000. Jednotlivá zařízení se samozřejmě liší kapacitou, použitými typy disků a podporovanými hosty. Nejmenší z těchto modulů je J4200. Toto zařízení nabízí ve 2RU výšce pozice pro 12 disků SAS (Serial-

attached- SCSI) nebo SATA 3,5" a také 2 IO moduly. IO modul si lze představit jako SAS expandér (analogie např. k FC switchi), ke kterému jsou připojeny jednotlivé disky a také SAS porty, pomocí kterých se zařízení připojí k serveru, případně pro kaskádové připojení dalších JBOD trayů – J4200 umožní kaskádově propojit až 4 moduly, tedy celkem získat až 48 disků. IO modul tedy neobsahuje žádný RAIDový řadič, ani vyrovnávací paměť. Jeho úkolem je pouze zviditelnit instalované pevné disky na příslušných host portech modulu.

O něco větší kapacitu nabídne zařízení J4400, které ve 4RU výšky pojme 24 disků, opět s možností volby SAS nebo SATA protokolu. Dva IO moduly, stejně jako v předchozím případě, nabízejí konektivitu přes 6 portů SAS. Z modulů J4400 půjde kaskádovým propojením sestavit úložiště až s 96 pevnými disky, později se plánuje až 192 SAS nebo SATA disků v jednom úložišti. Tak obrovskou kapacitu však nemusí nutně využívat pouze jeden připojený server. Jak J4200, tak J4400 pochopitelně podporují připojení více hostů, a to jak napřímo přes SAS porty, tak pomocí externího SAS expandéru. Z hlediska hostu se připojení realizuje pomocí SAS adaptéru, kde jsou k dispozici osmiportové SAS HBA, včetně podpory různých typů RAIDu. JBOD zařízení J4200 a J4400 jsou univerzálně použitelná a levné storage platformy, připojitelná k široké škále serverů. Pro kapacitní rozšíření storage serveru X4540 je však primárně určen poslední ze zástupců řady J4000 – JBOD modul J4500, alias Loki.

Tento modul nabízí největší osaditelnou kapacitu z celé produktové řady – 48 disků. Podporována je však pouze SATA technologie 3,5", stejně jako u storage serveru Thor. V modulu Loki lze použít SATA disky o kapacitách 500GB, 750GB nebo 1TB. Ve spojení s možností zkaskádovat až 2 Loki moduly (později až 10), můžeme dostat neuvěřitelnou kapacitu až 144TB. S přibývajícím kapacitou a počtem disků ve storage zařízení ale vyvstává důležitý problém – a tím je...

## Efektivní správa velkého počtu disků v JBOD zařízení

Storage server Thor se se svými širokými možnostmi využití neomezuje jen na podporu jednoho operačního systému. Na zařízení může běžet jak Solaris, tak Linux. Solaris však nabízí velkou výhodu v podobě

souborového systému ZFS (Zettabyte File System), který se velmi dobře uplatní při správě a efektivní konfiguraci tak vysokého počtu disků. Největší přidaná hodnota ZFS ve spojení s technologií Thor a Loki je zejména ve snadné administraci díky integraci s volume managerem, a vysokém stupni bezpečnosti, ideálním pro použití s levnými SATA disky. Koncept ZFS spočívá v tom, že pevné disky se sdruží do tzv. poolů, kde mají definován svůj RAID layout, a z tohoto poolu se následně alokují jednotlivé filesystemy, které již jsou chráněny daným typem raidu – k dispozici je mirror, RAIDZ a RAIDZ2, což jsou obdoby RAID5 a RAID6, které zaručují odolnost vůči výpadku jednoho, resp. dvou disků. Další stupeň zabezpečení pak představuje použití kontrolních součtů, kdy se ke každému zapsanému bloku vygeneruje kontrolní suma a uloží v bloku nadřazeném. Kromě toho jsou v ZFS všechny zápisové operace tzv. Copy-on-Write a navíc v podobě transakcí, které se buď dokončí celé, nebo se neprovedou vůbec. Tak se zajistí, že na discích je za všech okolností konzistentní stav a nepřijde se o data.

## Technologie pro rozsáhlá úložiště s nízkými náklady

K čemu se tedy dá storage server typu X4540 použít? Způsobů použití je mnoho, o čemž vypovídá i vysoký počet prodaných kusů předchůdce Thora – serveru X4500. V zásadě lze rozlišit dva hlavní způsoby využití tohoto zařízení. První způsob ze storage serveru vytvoří vysoce propustný NAS server s širokými možnostmi přístupu a flexibilní konfigurací. V tomto případě lze s výhodou využít vlastnosti operačního systému Solaris, který v sobě integruje jak NFS/CIFS server, tak iSCSI target komponentu. Můžeme tedy vytvořit volumy na ZFS storage poolech s příslušným stupněm zabezpečení, a tyto volumy zviditelnit po iSCSI jako blokovaná zařízení pro širokou škálu serverů. To vše bez nutnosti pořízení licencí software třetích stran a s plnou podporou Sunu. S využitím iSCSI můžeme X4540 použít také jako komponentu v storage řešení hierarchického typu, které obsahuje různé typy úložišť pro příslušné typy dat a zároveň poskytuje funkcionalitu automatické migrace mezi vrstvami podle předem definovaných politik. Takové řešení lze postavit např. se sunovským souborovým systémem SAM-FS (Storage Archive Management). Díky vysoké propustnosti a bohatě sizovanému IO najde X4540 uplatnění

i jako media server pro D2D2T zálohy – podporované zálohovací softwary jsou v tomto případě EBS, Veritas Netbackup nebo BakBone NetVault.

X4540 však kromě kapacity nabízí i vysoký výpočetní výkon v podobě dvou čtyřjádrových procesorů a nezanedbatelného množství paměti. Praxe již ukázala, že zařízení tohoto typu lze výhodně využít v řadě úloh, kdy server obsahuje i jistý stupeň aplikační logiky. Příkladem je zejména Data Warehouse appliance, vzniklá ve spolupráci s firmou Greenplum, kdy je výpočetní modul Thoru schopen v rámci zpracování dotazu projít značné množství dat. Z jiného soudku je pak využití X4540 ve stále rostoucí oblasti HPC (High Performance Computing). Zde vznikají ohromná množství dat jako produkty výpočetních operací nad množstvím dat vstupních. Oba typy těchto dat je třeba uložit tak, aby k nim byl rychlý a zejména paralelní přístup z velkého počtu výpočetních uzlů (výpočetní grid). V současnosti se pro tuto úlohu často prosazuje souborový systém Lustre, což je objektový souborový systém, který pro ukládání vlastních dat používá velké množství tzv. OSS (Object Storage Server), které se vyznačují velkou kapacitou a jistým výpočetním výkonem. Jako OSS pak mohou velmi dobře posloužit právě storage servery X4540.

V současné době vzniká mnoho dalších projektů využívajících předností storage serverů Thumper a Thor. Jmenovat můžeme např. aplikace v oblasti VTL – jako deduplikační zařízení pro Dilligent ProtecTier, či přímo jako VTL appliance v sunovském produktu VTL Value. X4540 umí dokonce s využitím Luminex Channel Gateway fungovat i jako mainframe storage přes FICON protokol. S přibývajícím výkonem a možnostmi kapacitního rozšíření se úměrně zvětšují i počty různých aplikací využívající toto zařízení. O oblíbě tohoto konceptu u zákazníků zajisté svědčí strmě vzrůstající počty prodaných kusů a samozřejmě také velká pozornost věnovaná X4540 ve vývojových laboratořích Sunu. To vše dává záruku vývoje a podpory tohoto úspěšného produktu i do budoucna.

Více informací naleznete na [www.sun.com/servers/x64/x4540/](http://www.sun.com/servers/x64/x4540/)