

SUPERPOČÍTAČ
DEVANA

VÝKON PRE VAŠE INOVÁCIE
NÁRODNÉ KOMPETENČNÉ CENTRUM PRE HPC

800 TFlop/s
celkový výkon

140 univerzálnych
CPU uzlov

8 akcelerovaných
GPU uzlov

32 NVIDIA A100
40GB akcelerátorov

296 Intel Xeon
procesorov

9472 CPU
jadier

38 TB
RAM

1 PB
dátové úložisko

100 Gb/s
Infiniband sieť

130 kW
príkon



LUCIA DEMOVIČOVÁ

NÁRODNÉ KOMPETENČNÉ CENTRUM PRE HPC SLOVENKO

„Technológie ako strojové učenie, umelá inteligencia, spracovanie veľkých objemov dát či pokročilé simulácie dnes poháňajú inovácie prakticky vo všetkých odvetviach – a superpočítače ako Devana sú motorom mnohých z týchto aplikácií.“

Nepremeškajme na Slovensku šancu využiť možnosti, ktoré HPC (high performance computing) ponúka. Devana je prístupná aj slovenským firmám pri vývoji inovatívnych postupov a modelov.

Naša ponuka však obsahuje oveľa viac ako len prístup k samotnej infraštruktúre. V Národnom kompetenčnom centre ponúkame bezplatné možnosti získať know-how, nové kompetencie a spoluprácu na reálnych projektoch s našim tímom odborníkov.

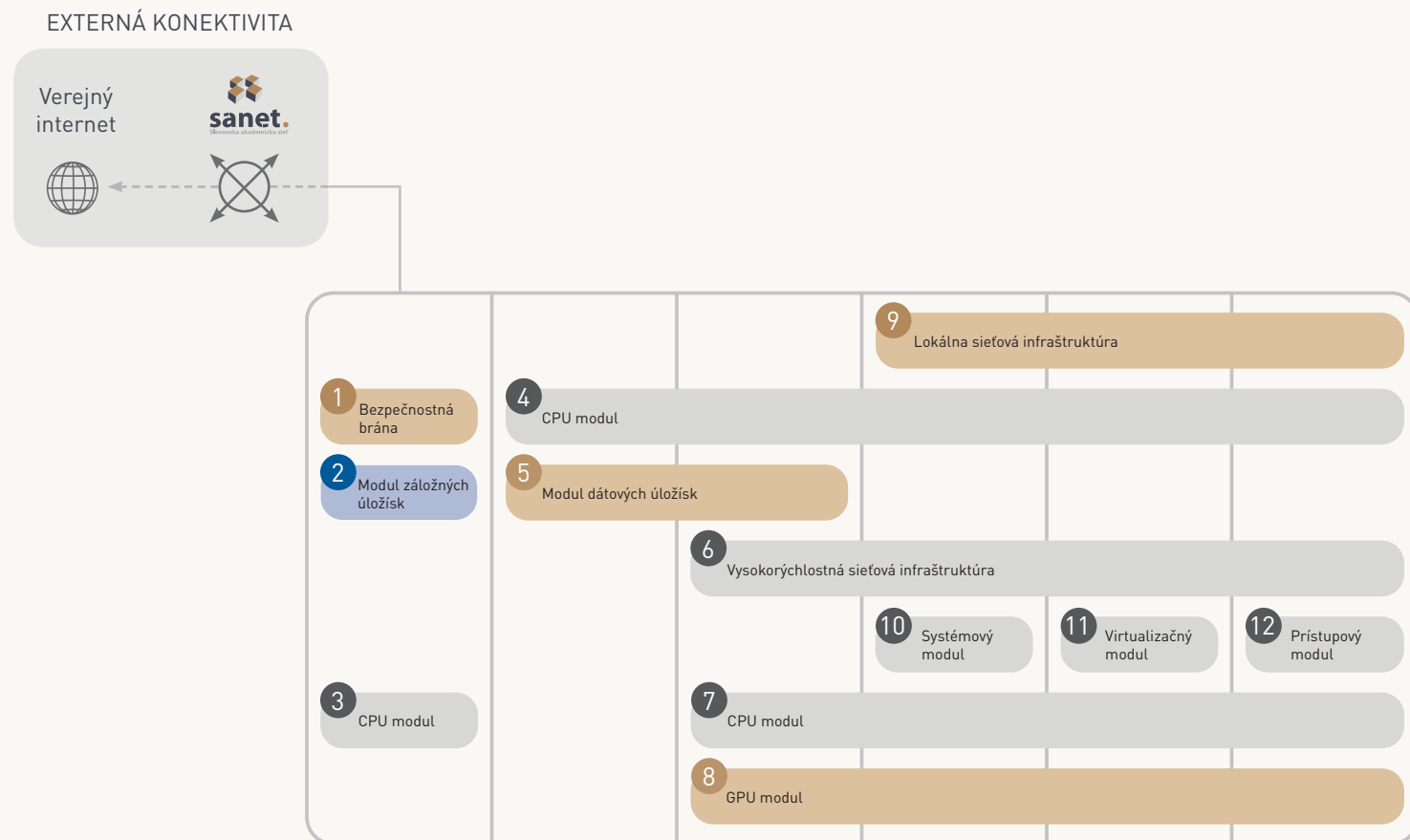
Našou misiou je pomôcť vyrásť slovenskému HPC ekosystému, podporovať nasadenie pokročilých digitálnych riešení a zároveň byť bránou k európskej HPC komunite.“



SPOZNAJTE
DEVANU

DEVELOPMENT
AND ENABLING

SCHÉMA architektúry



1 Bezpečnostná brána

- kontrola prístupu
- prevencia nepovoleného prístupu
- detekcia nepovoleného prístupu

2 Modul záložných úložísk

- zálohovanie systémových dát
- zálohovanie používateľských dát

3 CPU modul

- akcelerované HPC/HPDA

5 Modul dátových úložísk

- časť pre trvalé dáta používateľov/projektov
- časť pre dočasné dáta používateľov/úloh

6 Vysokorychlostná sieťová infraštruktúra

- sieť na prenos a zdieľanie dát
- sieť na medziuzlovú komunikáciu

8 GPU modul

- akcelerované HPC/HPDA/AI úlohy

9 Lokálna sieťová infraštruktúra

- manažmentová sieť
- viacúčelová servisná sieť

10 Systémový modul

- Správa HPC klastra
- Správa úloh a prístupu k zdrojom

11 Virtualizačný modul

- Používateľský portál
- Ostatné používateľské služby
- Ostatné služby na riadenie procesov HPC klastra

12 Prístupový modul

- používateľský prístup
- základné používateľské pracovné prostredie

Architektúra DEVANY

Univerzálny modul pre HPC aplikácie

Výkonovou jednotkou univerzálneho výpočtového modulu je server SD630 v2 vo formáte 0,5 U. Štyri takéto servery sú osadené v jednej 2U jednotke DA240 (Obr. 1). Táto platforma je optimalizovaná na maximalizáciu výkonu na jednotku plochy/objemu data-centra pri zachovaní jednoduchosti prevádzky a údržby. Samotná konfigurácia výpočtového uzla obsahuje 2 úsporné no výkonné procesory Intel Icelake 6338 32C/2.0GHz/205W, 256GB operačnej pamäte a 100Gb HDR Infiniband adaptér na pripojenie do výpočtovej siete (Obr. 2). Čo sa týka diskov výpočtové uzly sú rozdelené do dvoch kategórií podľa výkonu a kapacity lokálneho úložiska. V oboch prípadoch sa ale jedná o technológiu SSD NVMe pripojenú cez PCI express zbernicu.

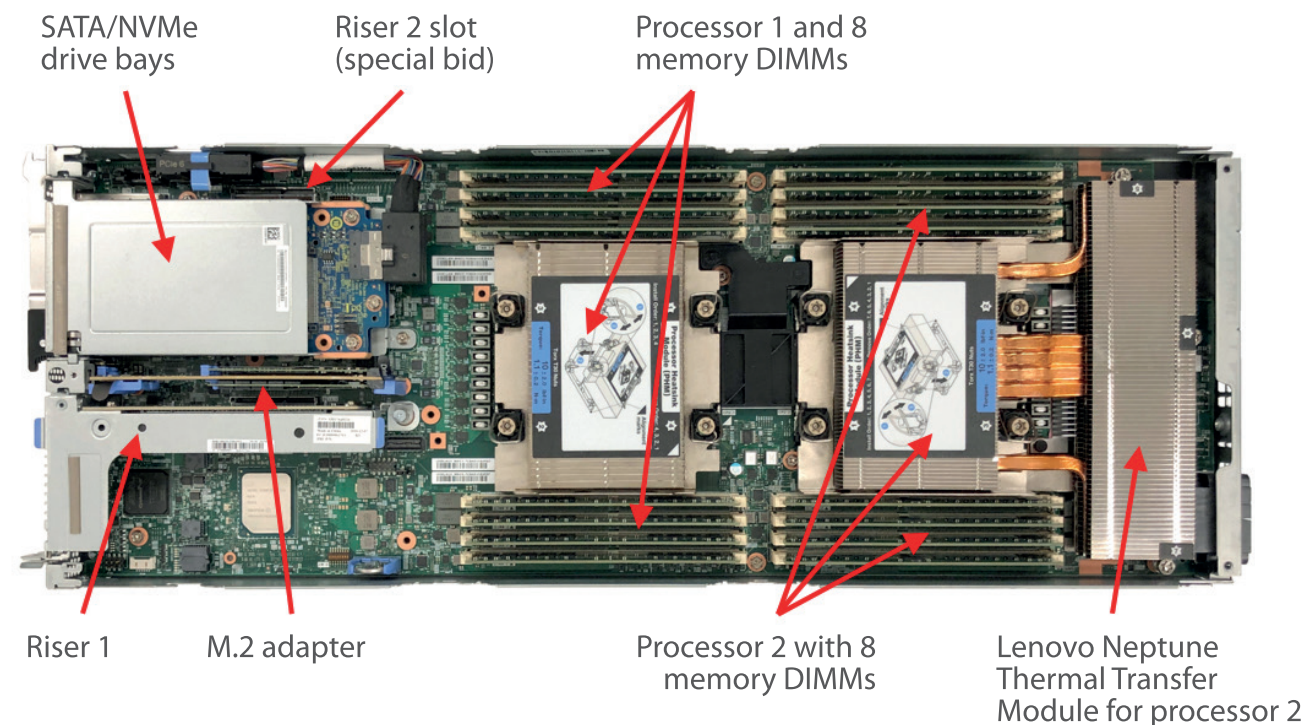


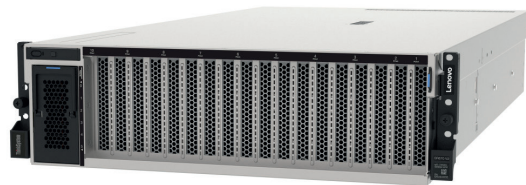
OBRÁZOK 1

Štyri univerzálne výpočtové uzly
Lenovo ThinkSystem SD630 v2
v jednotke DA240.

OBRÁZOK 2

Pohľad na vnútorné rozloženie
systémov výpočtového uzla
SD630 v2.



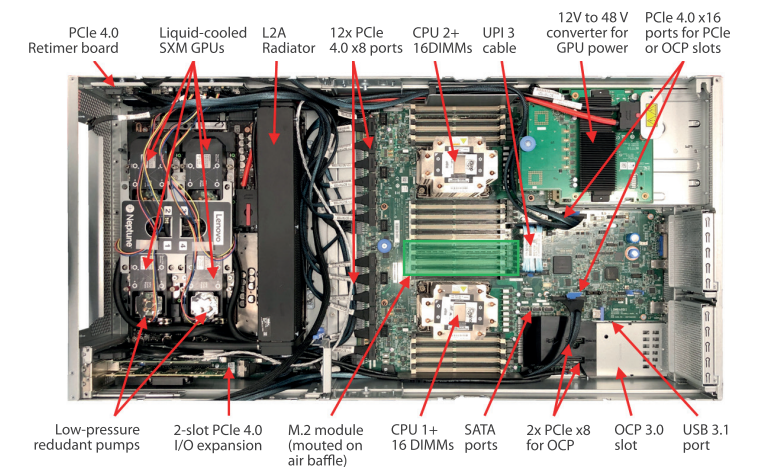
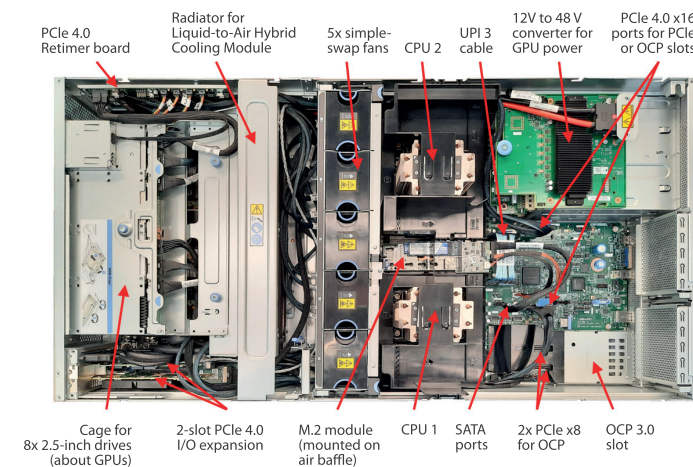


OBRÁZOK 3

Akcelerovaný výpočtový uzol
Lenovo ThinkSystem SR670 v2
s kapacitou až 8 PCIe GPGPU
kariet.

Akcelerovaný modul pre AI/ML

Akcelerovaný modul nového superpočítača pozostáva zo serverov typu SR670 v2. S možnosťou osadenia až ôsmimi akcelerátormi typu GPGPU má samotný server výšku 3U (Obr. 3). Kôli binárnej a inštrukčnej kompatibilitate majú akcelerované uzly rovnaký typ a počet procesorov ako univerzálne uzly. Takisto aj v prípade operačnej pamäte a lokálnych diskov je ich konfigurácia zhodná (Obr. 4). V porovnaní s univerzálnymi uzlami je posilnené pripojenie do výpočtovej siete na dvojnásobok, teda na 200Gb/s. Výkonovou časťou uzla pre aplikácie strojového učenia a umelej inteligencie sú štyri GPGPU akceľátory Nvidia A100 40GB vo formáte SXM, prepojené medzi sebou uniformnou rýchlosťou prostredníctvom NVlink technológie. Každá karta so svojimi 6912 CUDA a 432 tenzorovými jadrami poskytuje slušný výpočtový výkon aj pre tie najnáročnejšie aplikácie z oblasti AI/ML.



OBRÁZOK 4

Pohľad na vnútorné rozloženie systémov
výpočtového uzla SD670 v2 v dvoch úrovniach.

Architektúra DEVANY

Lenovo

Lenovo

Lenovo

Lenovo

Lenovo

Výpočtová sieť

Interná komunikačná sieť superpočítača určená na medziuzlovú komunikáciu a prístup k dátam na zdieľaných úložiskách je HDR infiniband (implicitné použitie technológie RDMA) s priepustnosťou 100 Gb/s a latenciou maximálne 10 mikrosekúnd. Posilnené dvojnásobné priepustnosti sú na uzly akcelerovaného modulu a servery poskytujúce prístup k dátam na zdieľaných úložiskách. Topológia tejto siete je stromová – neblokujúca sieť (t.j. blokovací pomer je 1:1). Skladá sa z dvoch vrstiev prepínačov, L1 vrstva prepínačov (tzv. “leaf” prepínače) je určená na pripojenie koncových bodov, L2 vrstva prepínačov (tzv. “spine” prepínače) prepája jednotlivé L1 prepínače navzájom.



Zdieľané dátové úložiská

Modul dátového úložiska obsahuje dva nezávislé fyzické klaster, klaster HOME a klaster SCRATCH. Prvý je navrhnutý ako bezpečné úložisko súborov určene primárne na ukladanie trvalých dát používateľov. Je to komplexné hardvérové riešenie použiteľné na nasadenie súborového systému distribuovateľného prostredníctvom NFS protokolu (predpokladá sa použitie verzie NFSv4.0 alebo vyššej). Dva obslužné NFS servery je možné nakonfigurovať v active-active alebo active-passive móde. SCRATCH je výkonné úložisko súborov určené pre krátkodobé skladovanie dočasných dát generovaných počas behu úloh. Finálna konfigurácia zahŕňa dva metadata servery a štyri dátové servery (V prípade Aurela boli dáta a metadata distribuované rovnomerne cez všetky dátové servery). Na úrovni súborového systému je použité „open source“ riešenie BeeGFS, umožňujúce v prípade potreby bezproblémové navýšenie šírky pásma pridaním ďalších dátových serverov.

Nástroje pre vás

- Jupyter Notebooks – interaktívne prostredie pre váš HPC development
- Open OnDemand – zjednodušený prístup cez webové rozhranie a prehľadné pracovné prostredie aj pre menej skúsených používateľov
- RStudio server – porozumejte svojim dátam vďaka efektívnej analýze
- Ansys – pokročilé simulácie fyzikálnych javov a vlastností pre lepšie produkty a kvalitnejší dizajn
- Singularity – platforma pre kontajnerizáciu
- Technická podpora a helpdesk v režime „next business day“

Príležitosti

- Získajte bezplatný prístup pre výskumné a vývojové projekty v režime open science.
- Škálujte riešenia – využite testovací prístup a vyskúšajte rôzne prístupy.
- Optimalizujte a spolupracujte na vašich projektoch s našimi odborníkmi.
- Aj produkčné výpočty môžete realizovať v rámci schém štátnej pomoci – Európske centrá digitálnych inovácií Hopero a HealthHub.
- Dohodnite si komerčný prístup s možnosťou dynamickej alebo statickej alokácie.
- V prípade potreby vám pomôžeme získať prístup k európskym pre-exascale superpočítačom EuroHPC JU.

Geodeticca

V rámci našej spolupráce sme mali možnosť využiť výpočtový výkon superpočítača Devana, ktorý nám poskytol neoceniteľnú pomoc pri zdokonaľovaní našej technológie automatického generovania topologicky čistých objektov z leteckých a satelitných snímok Zeme. Vďaka vysokému výkonu Devany sme boli schopní v krátkom čase dosiahnuť významné zlepšenie v presnosti našich algoritmov. Kľúčovou časťou našej spolupráce bolo tréningovanie umelej inteligencie na detekciu, klasifikáciu a následnú vektorizáciu objektov, pričom nám Devana umožnila v rýchlom čase nájsť optimálne parametre pre náš algoritmus. Táto spolupráca nám otvorila nové možnosti pre ďalší výskum a rozvoj v oblasti geopriestorových dát, čo v konečnom dôsledku prispelo k rozšíreniu našich služieb a zvýšeniu spokojnosti našich klientov.

Skymove

Vďaka výnimočným vlastnostiam nového superpočítača sme mali možnosť otestovať výpočtové procesy pre klasifikáciu priestorových údajov a dokázali sme identifikovať procesy, ktoré povedú k efektívnejšej práci a lepším výsledkom po dobu najbližších rokov. Vidíme veľký potenciál aj pri úlohách, ktoré sú spojené s umelou inteligenciou.

Nettle

Superpočítač Devana využívame pri tréningovaní jazykových modelov, ktoré nám pomáhajú s extrakciou kľúčových entít v texte. To nám pomáha vytvárať chatbot a voicebot aplikácie s novou úrovňou user experience.

Shark.Aero

HPC systém sme využili na CFD (výpočtová dynamika tekutín) simulácie rovnomerného prúdenia v 3D okolo čiastočného modelu lietadla pri rôznych rýchlostiach a uhloch nábehu. Vďaka HPC sme boli schopní spustiť viacero simulácií súčasne. Prístup k vysokovýkonnej infraštruktúre nám umožnil pracovať na väčšej sieti, čo bolo pre nás veľkým prínosom nakoľko náš lokálny počítač mal obmedzenú pamäť.

Codium

Akcelerovaný modul HPC systému Devana nám umožnil otestovať viaceré prístupy riešenia prevencie patologického hrania online hazardných hier. Výkonné GPGPU akcelerátory sa uplatnili pri tréningovaní a fine-tuningu zložitých AI modelov.

HPC tím



Michal
PITOŇÁK



Lukáš
DEMOVIČ



Marián
GALL



Filip
HOLKA



Miloslav
VALČO



Michal
KADÚC



Halyna
HYRYAVETS



EURÓPSKA ÚNIA
Európsky fond regionálneho rozvoja
OP Integrovaná infraštruktúra 2014 – 2020



MINISTERSTVO
INVESTÍCIÍ, REGIONÁLNEHO ROZVOJA
A INFORMATIZÁCIE
SLOVENSKEJ REPUBLIKY



MINISTERSTVO
DOPRAVY
SLOVENSKEJ REPUBLIKY



NATIONAL
SUPERCOMPUTING
CENTRE

Projekt Národné kompetenčné centrum pre vysokovýkonné počítanie
je spolufinancovaný Európskou úniou.
Kód projektu v systéme ITMS2014+: 311071AKF2