GAPP
SYSTEM

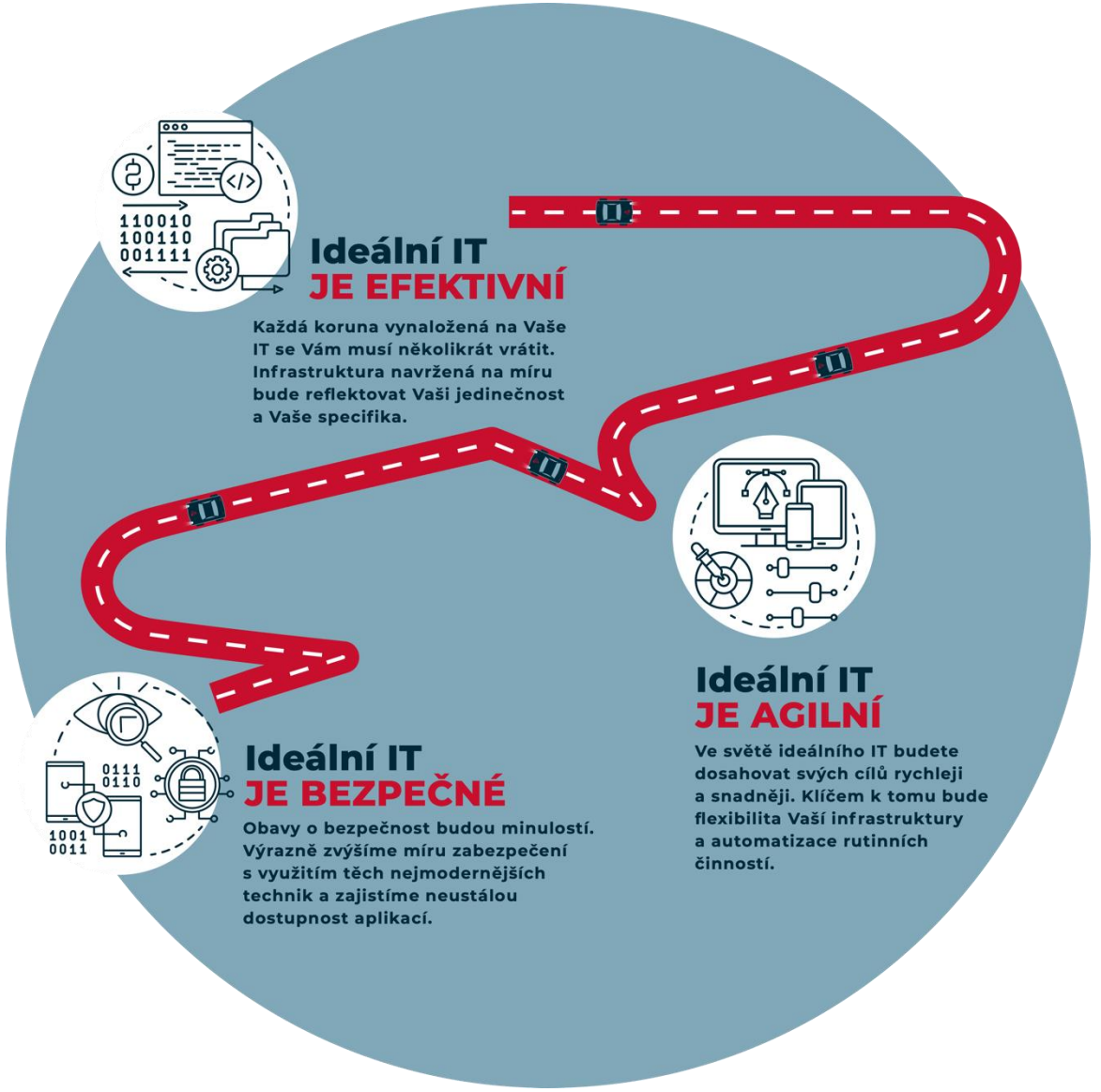
Provoz AI/ML workloadů v datových centrech

David Gottvald

Co my v GAPPu vlastně děláme?

Bezpečnost
Cloud
AI

#idealniIT



ARCHITEKTURA INFRASTRUKTURY PRO AI

GAPP SYSTEM

Příklad obecné architektury - onPrem

Web frontend cluster

- Servery pro interakci s uživateli

Vysoce výkonný LLM Cluster

- Servery pro řešení zadaných úloh

Vykonné RAG appliance

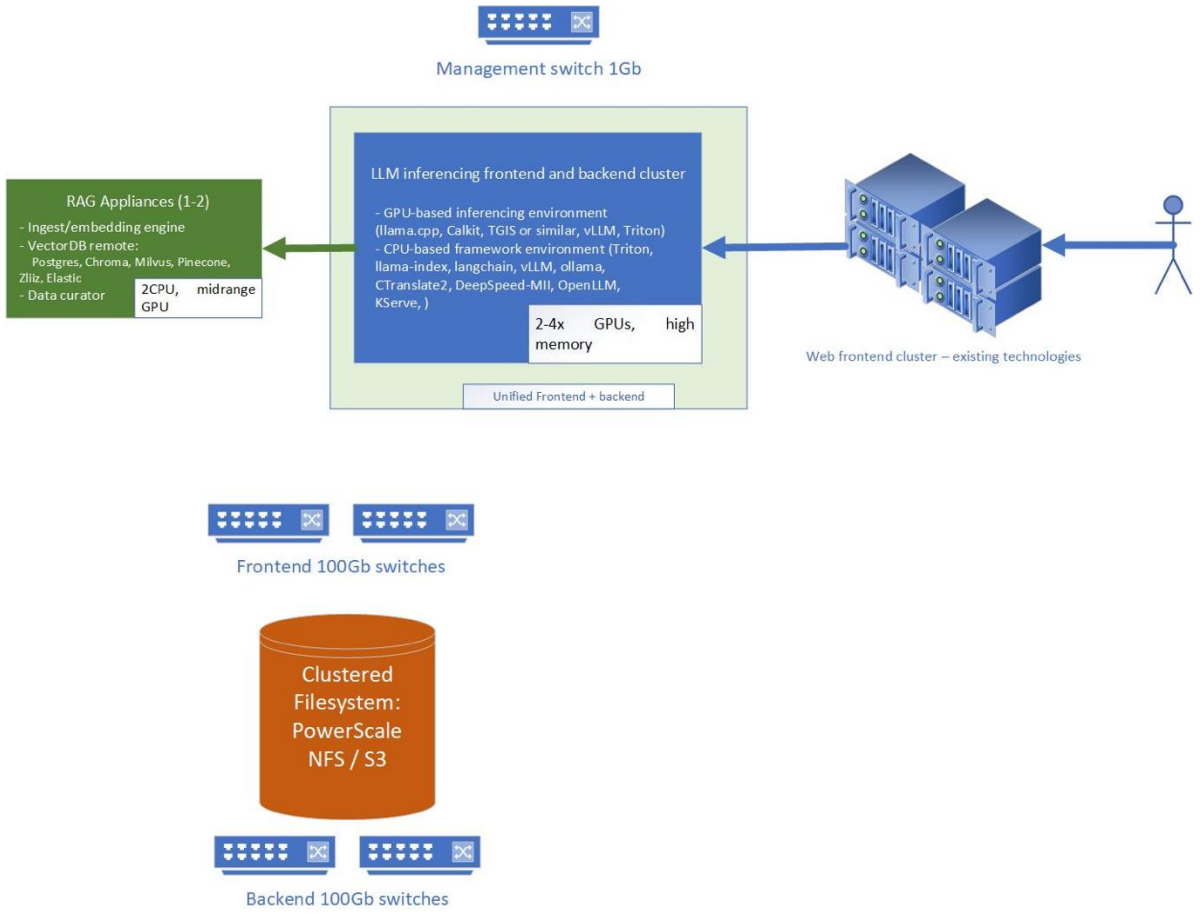
- Retrieval-Augmented Generation Appliance pro vyhledávání informací a generaci textu slouží pro zpřesňování výstupů velkých LLM modelů

Výkonné All-flash úložiště

- Výkonné a škálovatelné úložiště pro uložení dat

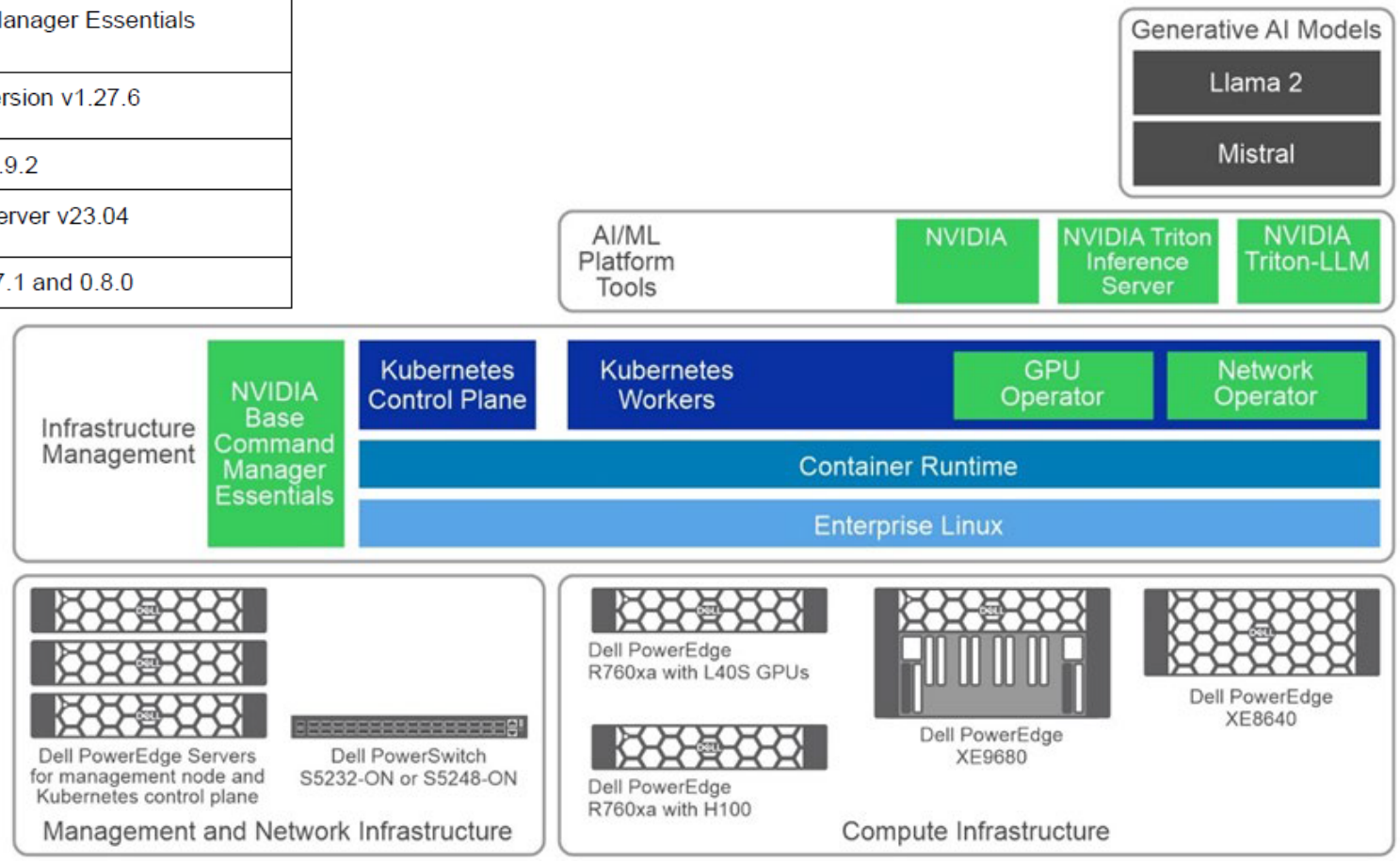
Vysoce propustné switche

- Switche, které slouží pro interní komunikaci celé AI/ML infrastruktury



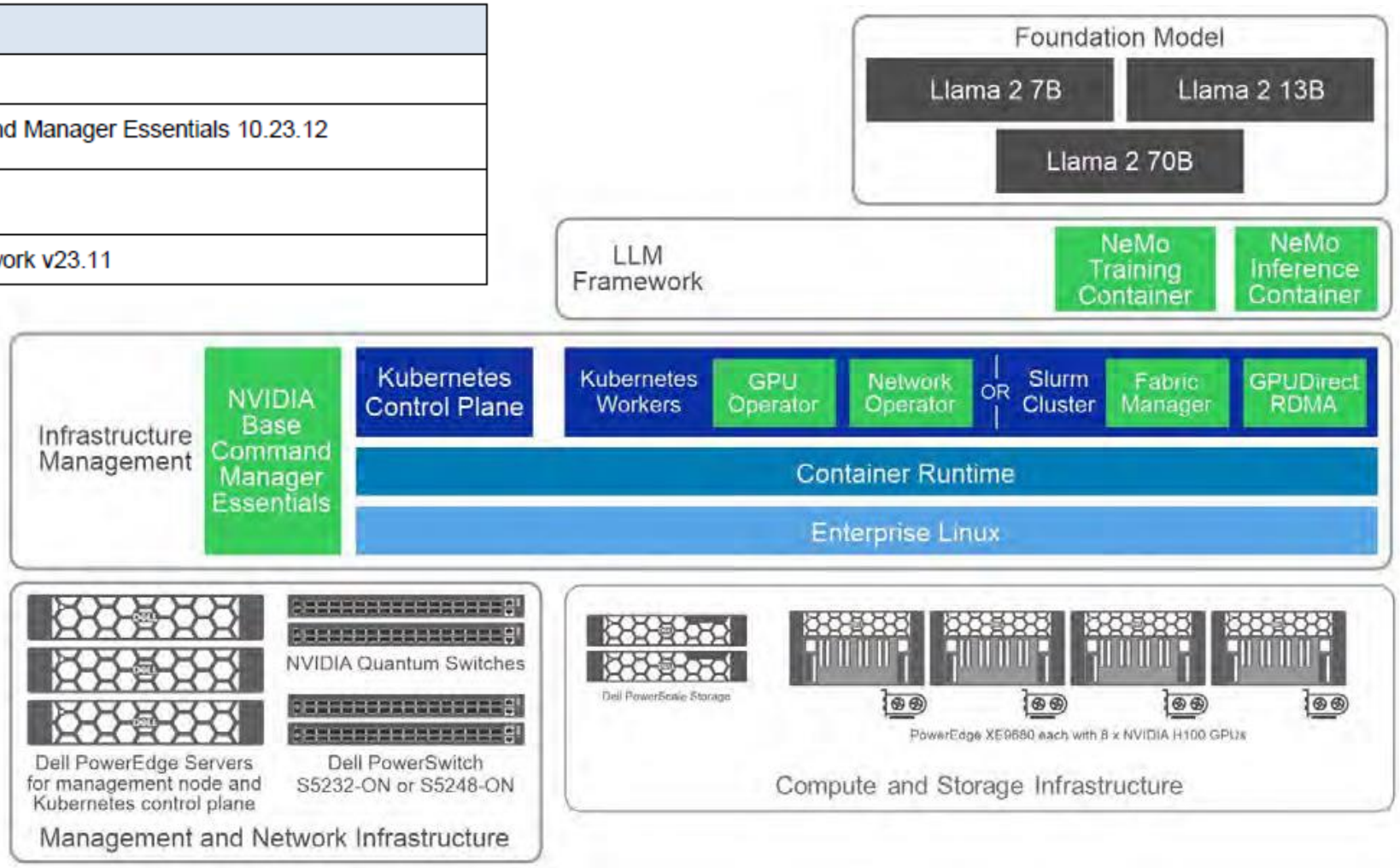
Příklad architektury LLM clusteru - Inference

Component	Details
Operating system	Ubuntu 22.04.1 LTS
Cluster management	NVIDIA Base Command Manager Essentials 10.23.12
Kubernetes	Upstream Kubernetes - Version v1.27.6
GPU operator	NVIDIA GPU operator v22.9.2
Inference server	NVIDIA Triton Inference Server v23.04
Inference Engine	NVIDIA TensorRT-LLM 0.7.1 and 0.8.0

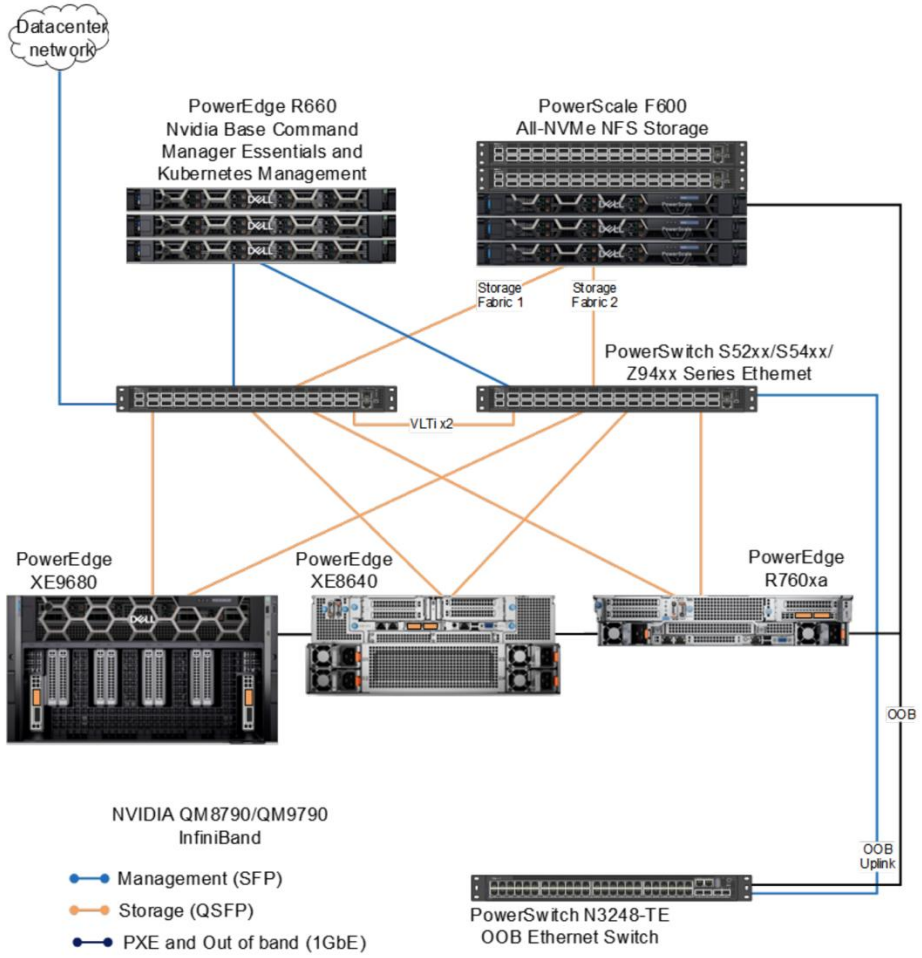


Příklad architektury LLM clusteru - Trénování

Component	Details
Operating system	Ubuntu 22.04.1 LTS
Cluster management	NVIDIA Base Command Manager Essentials 10.23.12
Slurm cluster	Slurm 23.02.4
AI framework	NVIDIA NeMo Framework v23.11



Jak může vypadat infrastruktura pro LLM



Management servery

- Dell PowerEdge R660 (NVIDIA Base Command Manager Essentials, Kubernetes control plane)

LLM servery (Worker nody)

- Dell PowerEdge R760XA (LLM model)
- Dell PowerEdge XE8640 (LLM model)
- Dell PowerEdge XE9680 (LLM model)

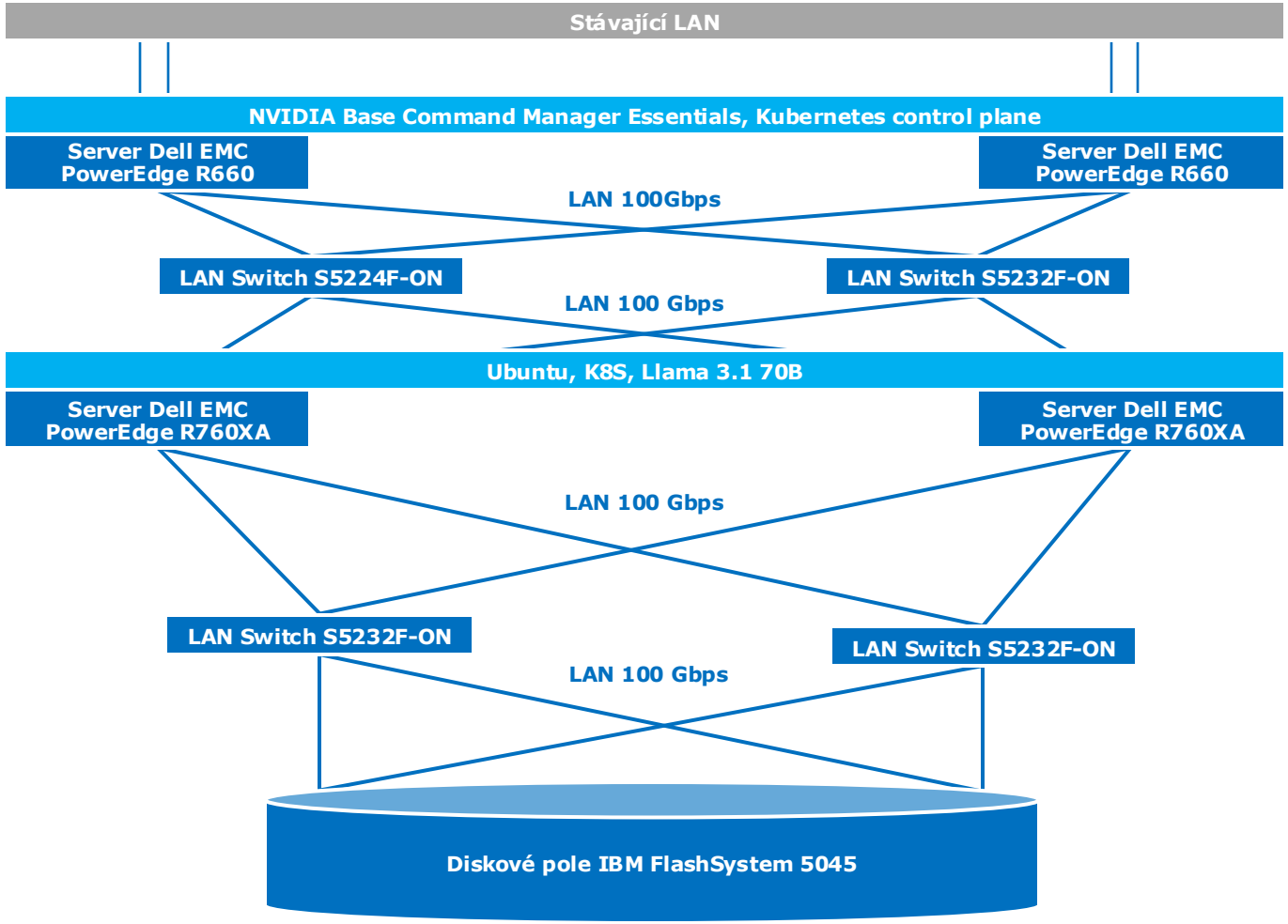
Výkonné datové úložiště

- Dell PowerScale F210 (data, databáze apod.)
- Dell PowerScale F710 (data, databáze apod.)

LAN infrastruktura pro interní komunikaci

- Dell PowerSwitch S52xx
- Dell PowerSwitch S54xx
- Dell PowerSwitch Z94xx

Příklad: Infrastruktura pro model Llama 3.1 70B



- **2x Server pro management a WEB UI:** Dell EMC R660 (2x Intel Xeon Gold 5418Y, 512 GB RAM, 4x 3,84 TB SSD, 4 x 10/25 GbE, 2 x 100 GbE, 5Y NBD)
- **LAN infrastruktura pro FrontEnd:** 2x Dell EMC S5232F-ON Switch, 32x 100GbE QSFP28
- **2x Server pro LLM model Llama 3.1 70B:** Dell EMC R760XA (2x Intel Xeon Gold 6542Y, 2x GPU NVIDIA H100 NVL, 1024 GB RAM, 2x 960 GB SSD, 4 x 100 GbE, 5Y NBD)
- **Datové úložiště:** 3 nody Dell EMC PowerScale F710 včetně LAN switchů pro BackEnd 2x S5232F-ON
- **Výkon:** 192 aktivních session, 1000 - 2000 uživatelů, TTFT < 2s
- **SW:** Ubuntu/Ubuntu Pro s K8s nebo Rancher, MLOPs nástroje Kubeflow a MLflow

MANAGEMENT SERVERY

GAPP
SYSTEM

Management server Dell PowerEdge R660

Kompaktní vysoce výkonný server s velmi širokým využitím

- Standardizace smíšených zátěží
- Databázové aplikace a analytika
- Infrastruktura pro virtualizaci

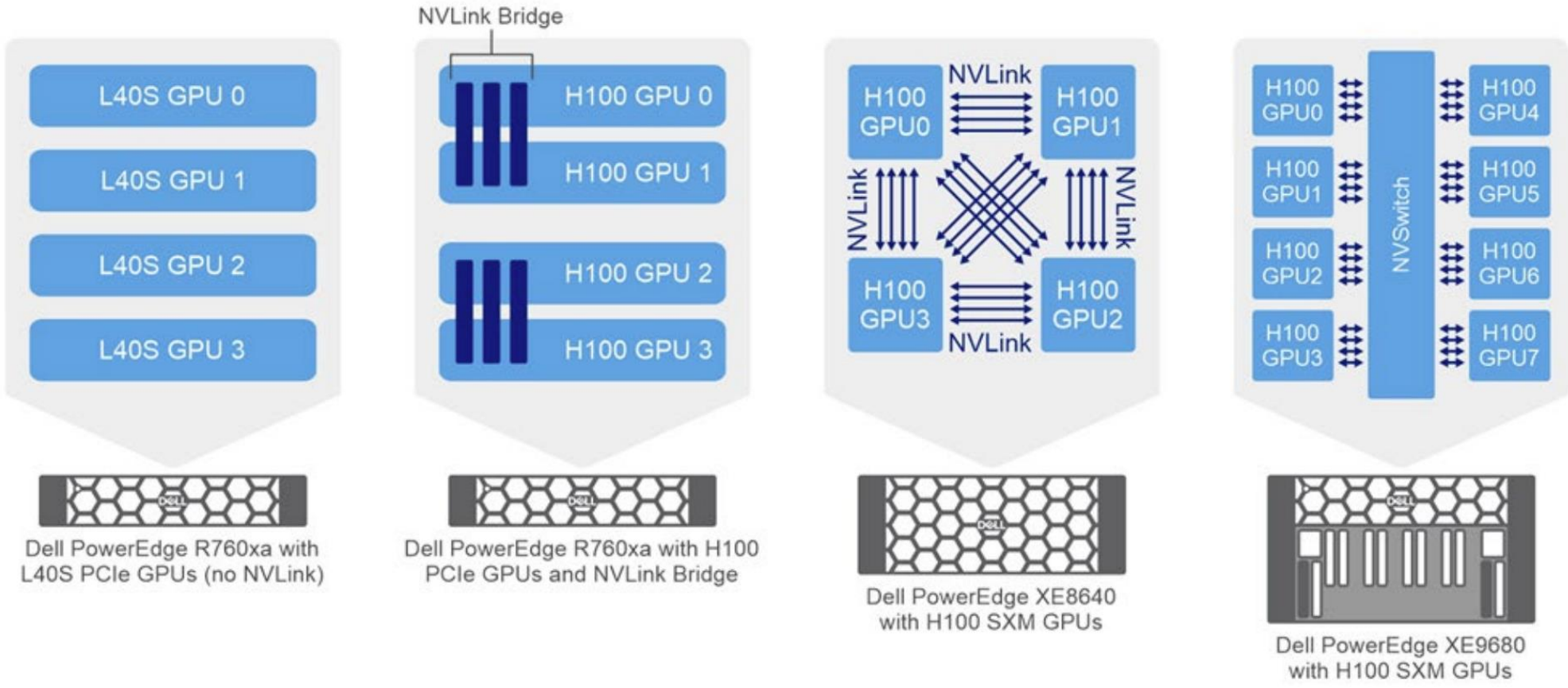
1U dvou-socketový server

- 1 – 2 Intel XEON Scalable nebo Max CPU 4. a 5. generace, 56 resp. 64 jader
- 32 DDR5 DIMM slotů, podporuje až 8 TB RAM (4800 MT/s a 5600 MT/s)
- Až 10 x 2.5“, SAS/SATA/NVMe HDD/SSD
- Až 16 x EDSFF E3.S Gen5 NVMe SSD

WORKER (LLM)
SERVERY

GAPP
SYSTEM

Servery pro Worker nody



LLM server Dell PowerEdge R760XA

Vysoce výkonný server optimalizovaný pro využití GPU

- Trénování a inferencing AI-ML/DL
- Pokročilá analytika
- Infrastruktura pro virtualizaci desktopů

2U dvou-socketový server

- 1 – 2 Intel XEON Scalable nebo Max CPU 4. a 5. generace, 56 resp. 64 jader
- 32 DDR5 DIMM slotů, podporuje až 8 TB RAM (4800 MT/s a 5600 MT/s)
- Až 4 x 400 W DW PCIe x16 GPU
- Až 12 x 75 W SW PCIe x8 GPU
- Až 8 x 2.5", SAS/SATA/NVMe
- Až 6 x 2.5-inch NVMe
- Až 6 x EDSFF E3.S Gen5 NVMe SSD

LLM server Dell PowerEdge XE8640

Specializovaný server optimalizovaný pro AI, HPC a aplikace náročné na výkon

- Trénování a inferencing AI-ML/DL
- Modelování a simulace v rámci HPC
- Velmi náročné aplikace z pohledu výkonu

4U dvou-socketový server

- 1 – 2 Intel XEON Scalable nebo Max CPU 4. a 5. generace, 56 resp. 64 jader
- 32 DDR5 DIMM slotů, podporuje až 4 TB RAM (4800 MT/s a 5600 MT/s)
- Až 4 x NVIDIA HGX H100 80GB 700W SXM5 GPU, plně propojené prostřednictvím NVIDIA NVLink
- Až 8 x 2.5", SAS/SATA/NVMe
- Až 8 x EDSFF E3.S Gen5 NVMe SSD

LLM server Dell PowerEdge XE9680

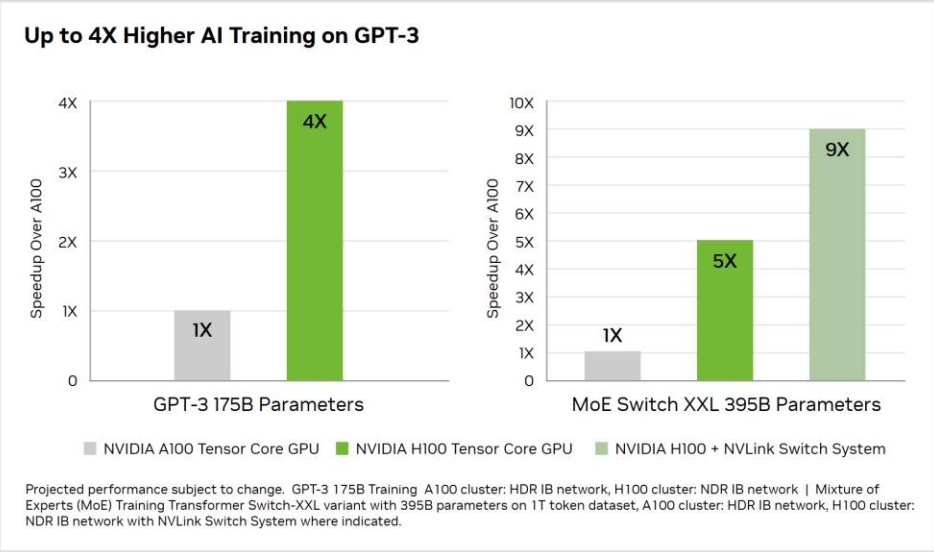
8-cestný GPU server s extrémním výkonem pro AI-ML/DL a HPC

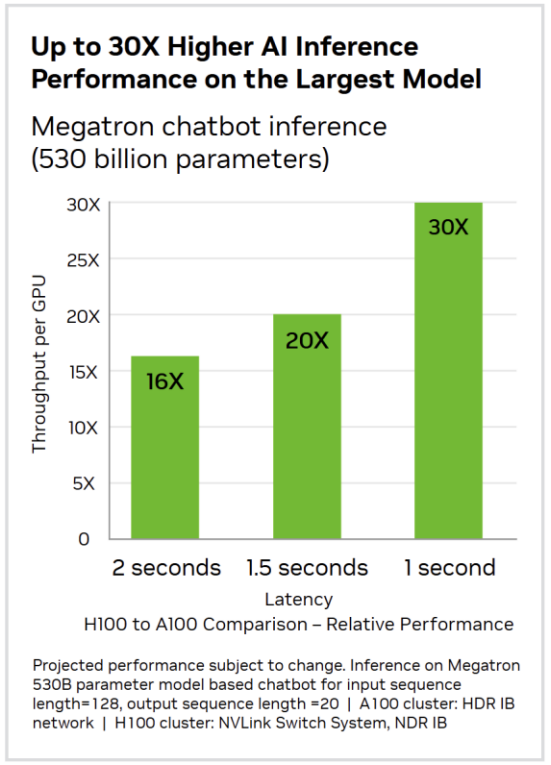
- Trénování a inferencing AI-ML/DL
- Podpora až osmi vysoce výkonných GPU
- HPC v oblastech velkých jazykových modelů, doporučovacích enginů, výpočty molekulární dynamiky a genových sekvencí

6U dvou-socketový server

- 1 – 2 Intel XEON Scalable nebo Max CPU 4. a 5. generace, 56 resp. 64 jader
- 32 DDR5 DIMM slotů, podporuje až 4 TB RAM (4800 MT/s a 5600 MT/s)
- 8 x NVIDIA HGX H100/200 s podporou NVIDIA NVLink
- 8 x AMD Instinct MI300X s podporou AMD Infinity Fabric nebo Intel Gaudi 3 s podporou RoCE
- Až 8 x 2.5", SAS/SATA/NVMe SSD
- Až 16 x E3.S NVMe direct drive

GPU NVIDIA H100





Technical Specifications

	H100 SXM	H100 NVL
FP64	34 teraFLOPS	30 teraFLOPS
FP64 Tensor Core	67 teraFLOPS	60 teraFLOPS
FP32	67 teraFLOPS	60 teraFLOPS
TF32 Tensor Core*	989 teraFLOPS	835 teraFLOPS
BFLOAT16 Tensor Core*	1,979 teraFLOPS	1,671 teraFLOPS
FP16 Tensor Core*	1,979 teraFLOPS	1,671 teraFLOPS
FP8 Tensor Core*	3,958 teraFLOPS	3,341 teraFLOPS
INT8 Tensor Core*	3,958 TOPS	3,341 TOPS
GPU Memory	80GB	94GB
GPU Memory Bandwidth	3.35TB/s	3.9TB/s
Decoders	7 NVDEC 7 JPEG	7 NVDEC 7 JPEG
Max Thermal Design Power (TDP)	Up to 700W (configurable)	350-400W (configurable)
Multi-Instance GPUs	Up to 7 MIGs @ 10GB each	Up to 7 MIGS @ 12GB each
Form Factor	SXM	PCIe dual-slot air-cooled
Interconnect	**NVIDIA NVLink™:** 900GB/s **PCIe Gen5:** 128GB/s	**NVIDIA NVLink:** 600GB/s **PCIe Gen5:** 128GB/s
Server Options	NVIDIA HGX H100 Partner and NVIDIA-Certified Systems™ with 4 or 8 GPUs NVIDIA DGX H100 with 8 GPUs	Partner and NVIDIA-Certified Systems with 1–8 GPUs
NVIDIA Enterprise	Add-on	Included

*With sparsity

VÝKONNÉ DATOVÉ ÚLOŽIŠTE

GAPP SYSTEM

All-Flash datové úložiště Dell PowerScale

Scale-Out datové úložiště s vysokou flexibilitou

- Scale-out architektura - Distribuovaná plně symetrická klastrovaná architektura s operačním systémem OneFS.
- Modulární design - PowerScale 1U nebo 2U pro montáž do racku s minimálně 3 uzly s backendovým připojením Ethernet nebo InfiniBand.
- Škálovatelnost - Cluster může škálovat až 252 uzlů se škálováním kapacity i výkonu. Jediný cluster může dodat až 186 PB RAW kapacity.
- Vysoká dostupnost - Bez jediného bodu selhání. Samoopravný design chrání před selháním disku nebo uzlu.
- Operační systém - PowerScale OneFS vytváří cluster s jediným systémem souborů a jedním globálním jmenným prostorem. Je plně žurnálovaný a distribuovaný.

Dell PowerScale F210 - ScaleOut NAS

1U Dell PowerEdge R660

- 1 x CPU Intel Xeon Silver 4410Y (2G/12C)
- 128 GB RAM
- 4x 10/25 GbE nebo 4x 40/100 GbE pro FE a BE
- 4x 2,5“ NVMe SSD (1,92(960 GB) TB, 3,84 TB, 7,68 TB a 15,36 TB)
- Min. 3 nody - Max. 252 nodů
- RAW kapacita nodu: 8 TB – 61 TB
- RAW kapacita clusteru: 23 TB - 15 PB

Dell PowerScale F710 - ScaleOut NAS

1U Dell PowerEdge R660

- 2 x CPU Intel Xeon Gold 6442Y (2,6G/24C)
- 512 GB RAM
- 2x 10/25 GbE nebo 2x 100/200 GbE pro FE a 2x 100/200 GbE pro BE
- 10x 2,5“ NVMe SSD (3,84 TB, 7,68 TB, 15,36 TB, 30,72 TB)
- Min. 3 nody - Max. 252 nodů
- RAW kapacita nodu: 38 TB – 308 TB
- RAW kapacita clusteru: 115 TB - 77 PB

Dell PowerScale F910 - ScaleOut NAS

2U Dell PowerEdge R760

- 2 x CPU Intel Xeon Gold 6442Y (2,6G/24C)
- 512 GB RAM
- 2x 10/25 GbE nebo 2x 100 GbE pro FE a 2x 100GbE pro BE
- 24x 2,5“ NVMe SSD (3,84 TB, 7,68 TB, 15,36 TB, 30,72 TB)
- Min. 3 nody - Max. 252 nodů
- RAW kapacita nodu: 92 TB – 737 TB
- RAW kapacita clusteru: 276 TB - 186 PB

LAN INFRASTRUKTURA PRO INTERNÍ KOMUNIKACI

GAPP SYSTEM

Dell PowerSwitch S5232F-ON

1U L3 switch – 32x 100GbE QSFP28 portů

- Škálovatelný L2 a L3 ethernetový switch s QoS, ACL a kompletní sadou standardizovaných funkcí IPv4 a IPv6 včetně OSPF, BGP a PBR.
- Výkon přepínače: 2 Tbps (4 Tbps full-duplex).
- Podpora L2 multipath prostřednictvím Virtual Link Trunking (VLT) a podpora Routed VLT
- VXLAN pro přemostění a směrování nevirtualizovaných a virtualizovaných overlay sítí.
- Konvergovaná síťová podpora pro Data Center Bridging s prioritním řízením toku (802.1Qbb), ETS (802.1Qaz), DCBx a iSCSI TLV.
- RDMA over Converged Network (RoCE) protokol pro akceleraci komunikace mezi aplikacemi a úložištěm.
- OS10 s podporou Precision Time Protocol (PTP, IEEE 1588v2).

Dell PowerSwitch S5448F-ON

1U L3 switch – 48x 100GbE SFP56-DD, 8x 400GbE QSFP56-DD

- Multi-rate 100GbE porty podporující 10/25/50/100GbE. Multi-rate 400GbE porty podporující 10/25/40/50/100/200/400GbE
- Škálovatelný L2 a L3 ethernetový switch s QoS, ACL a kompletní sadou standardizovaných funkcí IPv4 a IPv6 včetně OSPF, BGP a PBR.
- Výkon přepínače: 8 Tbps (16 Tbps full-duplex).
- Podpora L2 multipath prostřednictvím Virtual Link Trunking (VLT) a podpora Routed VLT
- VXLAN pro přemostění a směrování nevirtualizovaných a virtualizovaných overlay sítí.
- Konvergovaná síťová podpora pro Data Center Bridging s prioritním řízením toku (802.1Qbb), ETS (802.1Qaz), DCBx a iSCSI TLV.
- RDMA over Converged Network (RoCE) protokol pro akceleraci komunikace mezi aplikacemi a úložištěm.
- OS10 s podporou Precision Time Protocol (PTP, IEEE 1588v2).

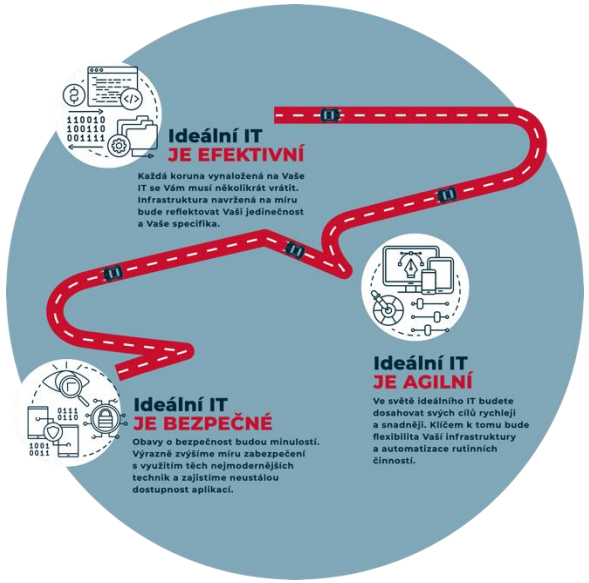
Dell PowerSwitch S9432F-ON

1U L3 switch – 32x 400GbE QSFP56-DD

- Multi-rate 400GbE porty podporující 10/25/40/50/100/200/400GbE
- Škálovatelný L2 a L3 ethernetový switch s QoS, ACL a kompletní sadou standardizovaných funkcí IPv4 a IPv6 včetně OSPF, BGP a PBR.
- Výkon přepínače: 12,8 Tbps (25,6 Tbps full-duplex).
- Podpora L2 multipath prostřednictvím Virtual Link Trunking (VLT) a podpora Routed VLT
- VXLAN pro přemostění a směrování nevirtualizovaných a virtualizovaných overlay sítí.
- Konvergovaná síťová podpora pro Data Center Bridging s prioritním řízením toku (802.1Qbb), ETS (802.1Qaz), DCBx a iSCSI TLV.
- RDMA over Converged Network (RoCE) protokol pro akceleraci komunikace mezi aplikacemi a úložištěm.
- OS10 s podporou Precision Time Protocol (PTP, IEEE 1588v2).

Proč AI řešit s GAPP System

- Máme znalosti a zkušenosti
- Jsme spolehlivý partner na trhu IT
- Dokážeme dodat ucelené řešení
- Máme k dispozici tým odborníků (Návrh, implementace a podpora)
- Máme 100% úspěšnost projektů u našich zákazníků



GAPP
SYSTEM

Děkuji za pozornost

David Gottvald
david.gottvald@gapp.cz
+420 724 954 105