



Informační systém MU

architektura a implementace v rozsáhlém prostředí velké organizace

Michal Brandejs, Miroslav Křipač
Masarykova univerzita

DATAKON 2007



Témata přednášky

- Motivace
- Architektura systému
- Distribuce strukturovaných dat
- Distribuce nestrukturovaných dat



Masarykova univerzita

- **Nejvýběrovější univerzita v ČR***
(40 000 uchazečů, 60 000 přihlášek)
- **9 fakult, řada fyzických lokací**
(spektrum oborů: humanitní, přírodovědné, umělecké, technické)
- **Největší instituce na Moravě**
- **40 237 studentů, 5 816 zaměstnanců**
- **Integrované prostředí**
(jednotná legislativa, přijímací řízení, plně přístupné studium)
- **Rozsáhlé kurzy**
(až tisíc studentů navštěvujících jeden předmět)



Informační systém MU

- Jednotný webový systém:
<https://is.muni.cz/>
- Elektronická podpora administrativy školy
(studijní agenda, kompletní elektr. podpora výuky, komunikace...)
- „ERP“ a „DSS“ univerzity
(žádné indexy, on-line přihlášky, automatické hodnocení, ...)
- Striktní přímý on-line přístup k datům
- Další přidaná hodnota
(elektronické Obchodní centrum, „antiplagiátorský systém“, ...)



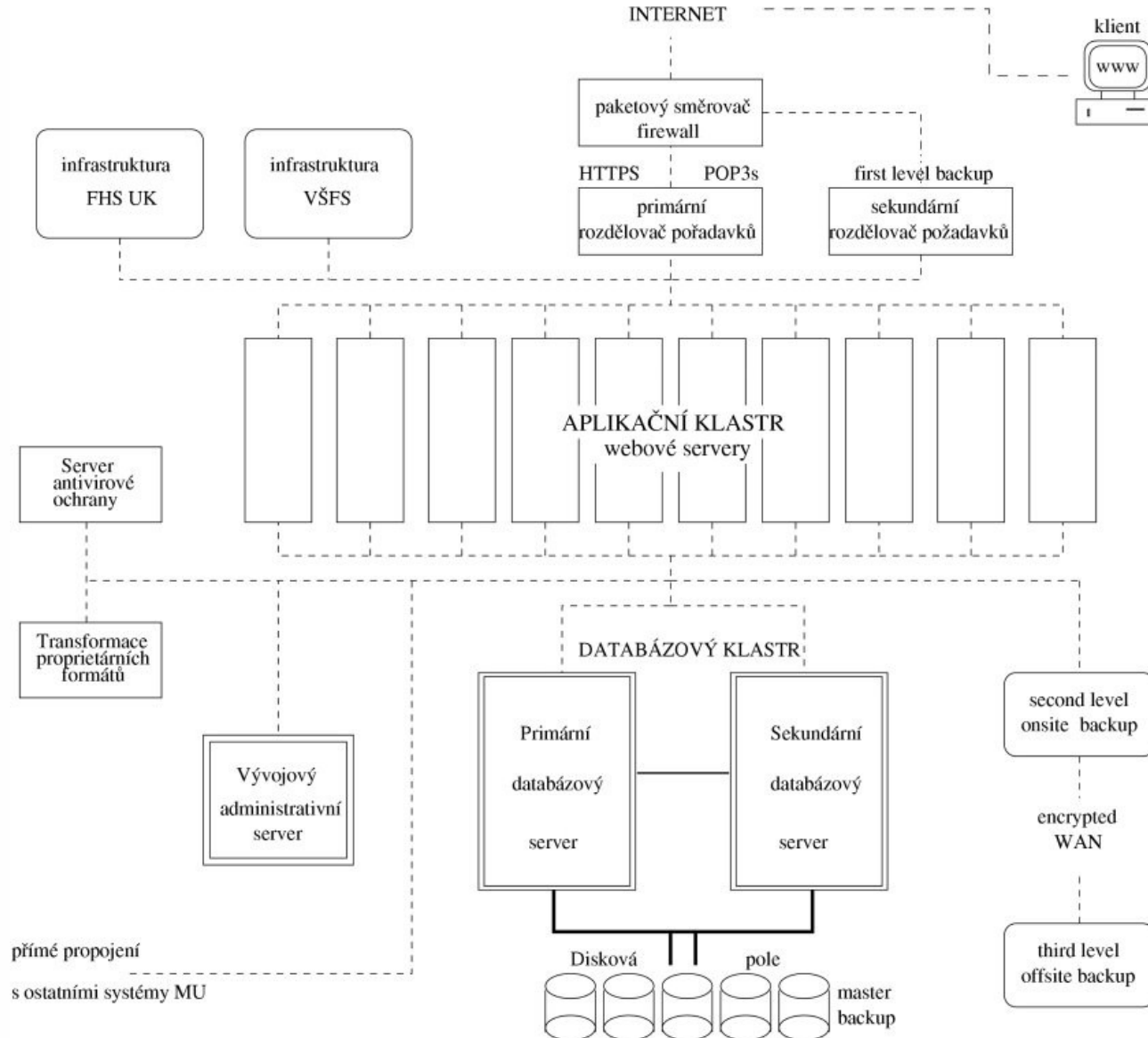
Informační systém MU

- Asi 45 000 aktivních uživatelů
(celkem 140 000)
- Až 27 000 uživatelů každý den
- 2 mil. komplexních operací za den
(špičkový výkon až 500 000 operací za hodinu)
- Tisíce uživatelů ve stejný okamžik
- 2s na vyčerpání kapacity semináře
- Nonstop provoz
(„noční uživatelé“, zahraniční přístup)



Architektura systému

SCHÉMA ARCHITEKTURY JÁDRA IS MU





Klientská část

- **Pouze běžný webový prohlížeč HTTPs**
(Mozilla Firefox, MSIE, Safari, Opera, Nokia Web Browser...)
- **Nízké nároky na základní funkčnost**
(pokročilé technologie jen volitelně: JS, AJAX, cookies...)
- **Nízké náklady na údržbu aplikací**
- **Tiskový subsystém s přímým přístupem**
(lokální tisk, podpora různých protokolů, distribuovaná správa)
- **Autentizace jménem a heslem**
(vylepšená HTTP autentizace, podpora odhlašování)
- **Podpora pro WAP přístup**



Vstupní část

- **Dvě připojení do sítě CESNET**
(v blízkosti brněnského uzlu, redundantní spojení)
- **Dva nezávislé směrovače**
(automatická zastupitelnost, sdílené s Fakultou informatiky)
- **Interní síť ve dvou fyzických lokacích**
- **Dvě nezávislé brány**
(automatická zastupitelnost, sdílené s Fakultou informatiky)
- **Rozdělování požadavků**
(distribuce zátěže na úrovni TCP spojení)
- **OEM x86-64, Linux Fedora Core**



Aplikační část

- **Aplikační cluster ve dvou lokacích**
(20 x OEM x86-64 dual-core, Fedora Core Linux)
- **Vlastní aplikační prostředí**
(Apache HTTP Server, Perl, mod_perl, 1400 aplikací)
- **Vlastní systém distribuované cache**
(slouží pro „read-mostly“ data, invalidace)
- **Plná zastupitelnost**
(automatická detekce a obnovení v případě výpadku uzlu)
- **Lineární horizontální škalovatelnost výkonu**
(data nejsou sdílena na aplikační úrovni s výjimkou cache)



Podpůrná infrastruktura

- **Zálohovací systémy**

(on-site i off-site backup, automatické, ruční kontrola)

- **Antivirová ochrana**

(ochrana souborů uložených v systému)

- **Převod proprietárních formátů**

(automatická konverze DOC, ODT, XLS, PPT, PDF, TXT)

- **Scanování a rozpoznávání textu**

(písemné testy vč. hodnocení, archivace závěrečných prací)

- **Napojení na externí systémy**

(ekonomický IS, vyhodnocení matematických úloh, ...)



Datová část

- Nestrukturovaná data
 - vlastní distribuované datové úložiště

- Vysoce strukturovaná data
 - relační databázový systém



Nestrukturovaná data

- Nejsou interpretována systémem
(dokumenty, multimediální data, obrazy CD/DVD, ... až TB dat)
- Nepodléhají vnitřním změnám
(změna znamená smazání a vložení celé nové verze)
- Převažuje čtení nad zápisem
(zápis typicky není paralelní, netransakční)
- **IS MU: Distribuované datové úložiště**
(vlastní řešení, distribuce dat na diskové kapacity apl. clusteru)
- Zajištění konzistence a rozložení zátěže
- Velmi efektivní využití a navýšení zdrojů
(až 4 disky/1U uzel -> využití existující infrastruktury)



Strukturovaná data

- Základní údaje spravované systémem
(osoby, kurzy, zkušební termíny, známky, ...)
- Masivně paralelní on-line přístup pro zápis
(změny se projeví ihned, zásadní důsledky pro čtení ostatních)
- Malý objem dat x velká výpočetní složitost
(nestačí data jen načíst/uložit, je nutné je on-line interpretovat)
- Je vhodné i tato data **distribuovat** na DB úrovni?



Strukturovaná data

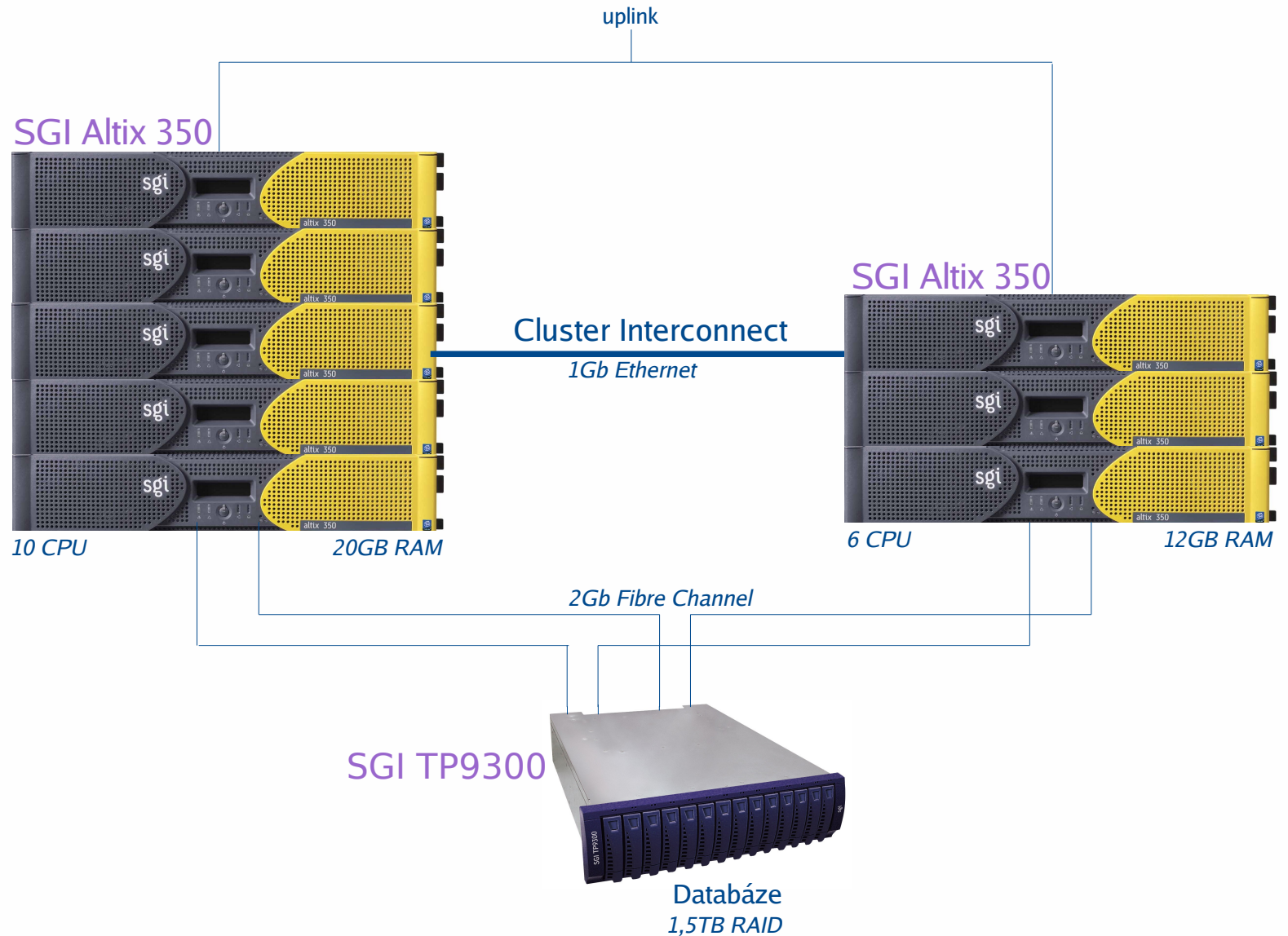
- Vlastnosti distribuovaného systému
 - relační databázový systém ve formě clusteru
 - cluster se sdíleným diskovým prostorem
 - jedna databáze spravovaná více uzly v síti (jedna data)
 - + izolace hardwarové a částečně softwarové chyby
 - + automatické zotavení z chyb, vyšší dostupnost
 - + možnost snadnější distribuce zátěže, škálovatelnost
 - + architektura založená na komoditním hardware
 - + možnost postupného růstu nákladů
 - + vyšší zabezpečení provozu během údržby
 - režie softwarového zajištění horizontální škálovatelnosti
 - náklady na softwarové licence
 - složitost realizace, chybovost, propagace globálních chyb



Strukturovaná data v IS MU

- **Problém:** režie zajištění konzistence při zátěži
- Jak ponechat zmiňované výhody clusteru?
- **Řešení:** architektura SGI NUMAflex (Altix 450)
 - vertikální škálovatelnost = sdílená paměť z pohledu DBMS
 - + velmi vysoký výkon i propustnost na HW úrovni (*52 jader*)
 - + vysoká stabilita („jednoduchý“ DBMS server) (*10 min výpadku*)
 - + možnost postupného růstu i dočasného navýšení zdrojů
 - + násobné umístění všech komponent pro zajištění dostupnosti
 - + neproprietární řešení IA64 – cena blíže komoditním HW (*60,- Kč*)
 - + bez nutnosti dalších licencí
 - rekonfigurace může znamenat výpadek celého systému
 - méně obvyklá architektura

Distribuovaná varianta IS MU





Nedistribučovaná varianta IS MU

SGI Altix 450



52 CPU core

104GB RAM

SGI TP9300



Databáze
1,5TB RAID



Dotazy, komentáře, ...

kripac@fi.muni.cz