

Infrastruktura pro zpracování záznamů přednášek

Eva Hladká, Miloš Liška

eva@fi.muni.cz, xliska@fi.muni.cz



Fakulta informatiky
Masarykova univerzita v Brně

SCO2005, Brno

Záznamy přednášek

- Regulérní studijní materiál
- Pravidelné přednášky i jednorázové akce
- Stále větší objem obého

Záznamy přednášek

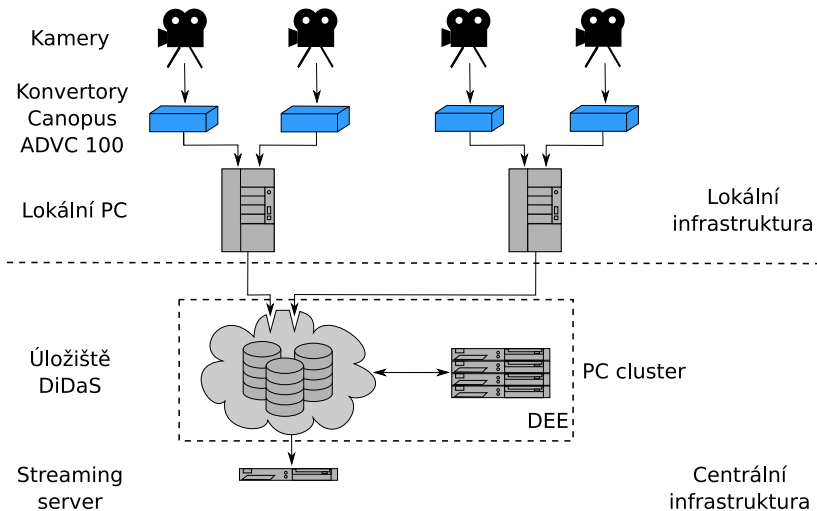
- Kde zaznamenávat a kde zpracovávat
- Záznam v posluchárně při přednášce
 - Přirozené prostředí
 - Neklade dodatečné nároky na pedagoga
- Zpracování centrální (poměr cena a výkon)

Požadavky na celý systém

- Bezobslužnost
- Škálovatelnost
- Robustnost

Fáze záznamu a zpracování přednášky

- 1 Akvizice obrazu a zvuku
 - 2 Transport dat ke zpracování
 - 3 Kódování dat a jejich úprava
 - 4 Transport dat na streaming server
 - 5 Prezentace dat
- Fáze 1 nutně lokální
 - Fáze 2 lokální/centrální
 - Fáze 3–5 centrální



Vybavení poslucháren

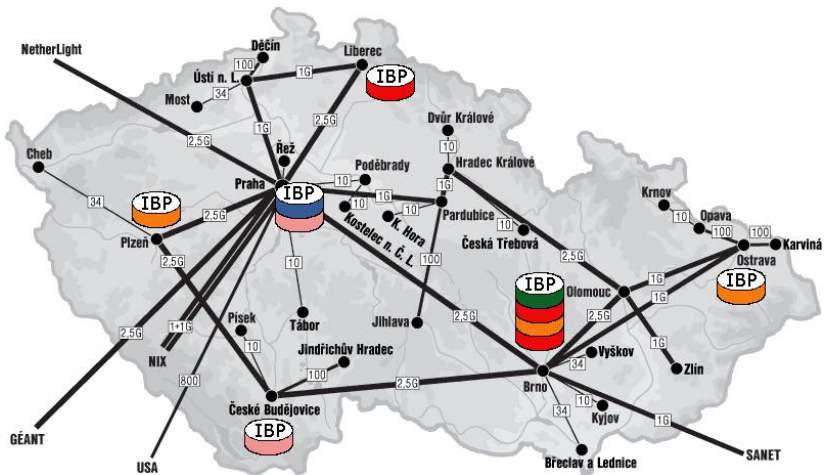
- Jedna nebo více kamer
- Zvukový systém
- Lokální počítač pořizující záznam
 - Záznam ve formátu rawDV, vhodný pro další (i vícenásobné) zpracování
 - Datový tok 25 Mb/s
 - 2 h přednáška = 24 GB dat

Síťová infrastruktura

- Nosný prvek infrastruktury
- Doporučená šířka pásma 1 Gb/s
- Všechny stroje ve vybudovaném systému jsou připojeny linkami s kapacitou 1 Gb/s

Datové úložiště

- Projekt DiDaS
- Distribuované datové úložiště
- Celková kapacita 16 TB
- 8 serverů rozmístěných po celé republice, připojených k páteřní síti CESNET2
- Ukládání primárních dat a dočasných dat při zpracování
- Přístup k datům prostřednictvím metadat



Kódování záznamů 1/2

- PC cluster v rámci infrastruktury Metacentra
- 5 uzlů
 - 2× P4 3 GHz CPU, 2 GB RAM
- Zpracování v systému Distributed Encoding Environment (DEE)

Kódování záznamů 2/2

Používané formáty záznamů:

- DivX; 512×384 ; datový tok 512 kb/s
- RealMedia 9 SureStream; 384×288 ; datové toky 768 kb/s, 512 kb/s resp. 256 kb/s
- mp3

Výkon systému:

- Záznam přednášky lze zpracovat za 90 minut
- Pro srovnání cca. 4 hodiny na samostatném PC

Streaming server

- Úložiště pro zpracované přednášky
 - Aktuální kapacita až 4 TB
 - Data jsou zálohovaná
- Poskytuje záznamy uživatelům
 - Webové rozhraní
 - HelixServer
- Statistiky o využívání záznamů

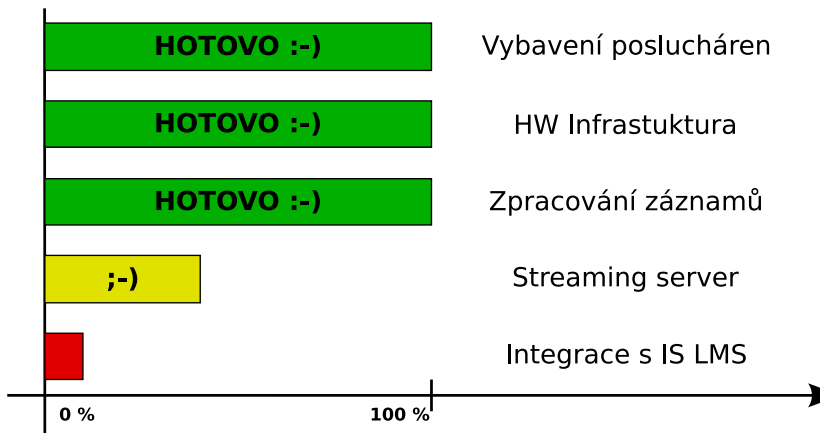
Komunikace IS ↔ streaming server

- Výměna textových zpráv (protokol http)
 - Formát zprávy:
`semestr:1433:kod:datum:url:velikost:poznamka`
- Autorizace vyučujícím
- Autentizace při přístupu
- Zpětná vazba (diskuse)

Dosažené výsledky

- Kompletně zprovozněný HW
- Prakticky hotová SW část
- Produkční stav

- Bezobslužnost
- Jednoduché použití
- Efektivnost
- Vysoká škálovatelnost
 - V současné době 62 h (FI) + 16 h (PrF) zaznamenaných a zpracovaných přednášek týdně
 - Předpoklad až 224 h přednášek týdně



Další vývoj

- High-definition (HD) video
- Možnost využití nových formátů
- Automatické indexování videa založené na rozpoznávání řeči

Děkuji za pozornost.

Otázky?