

**INTERLAN**  
Computer Networks

IBM - Globálna expertíza a technológie

José Manuel Barroso:  
"... **kultúra** je srdcom  
európskej **integrácie**"



# ***Slovakiana***

## ***Stratégia budovania infraštruktúry dátové sklady a archívy***

Národná konferencia o digitalizácii  
16.-17. júna 2009

TP Hotel Družba Senec



IBM: Prestaňme hovoriť, začnime konať

16.6.2009 | © 2009 IBM Corporation

## Infraštruktúra pre Slovakianu – Zadanie

- Infraštruktúra pre systém pre dlhodobú archiváciu, organizáciu a sprístupnenie digitálneho kultúrneho obsahu
  - Systém dlhodobej archivácie digitálneho obsahu (**Digitálny archív**)
  - Systém pre správu digitálneho obsahu (**Národný agregátor**)
    - Nástroj pre popis objektov (**Katalogizačný modul**)
    - **Systém pre správu národných autorít**
    - Systém pre správu zbierok
  - Systém na sprístupnenie a prezentáciu obsahu (**Portál kultúry**)
    - **Integračné rozhranie** na Ústredný portál verejnej správy (ÚPVS)
- Technické požiadavky
  - Dostupnosť
  - Spoľahlivosť
  - Integrácia
  - Certifikácia a štandardizácia

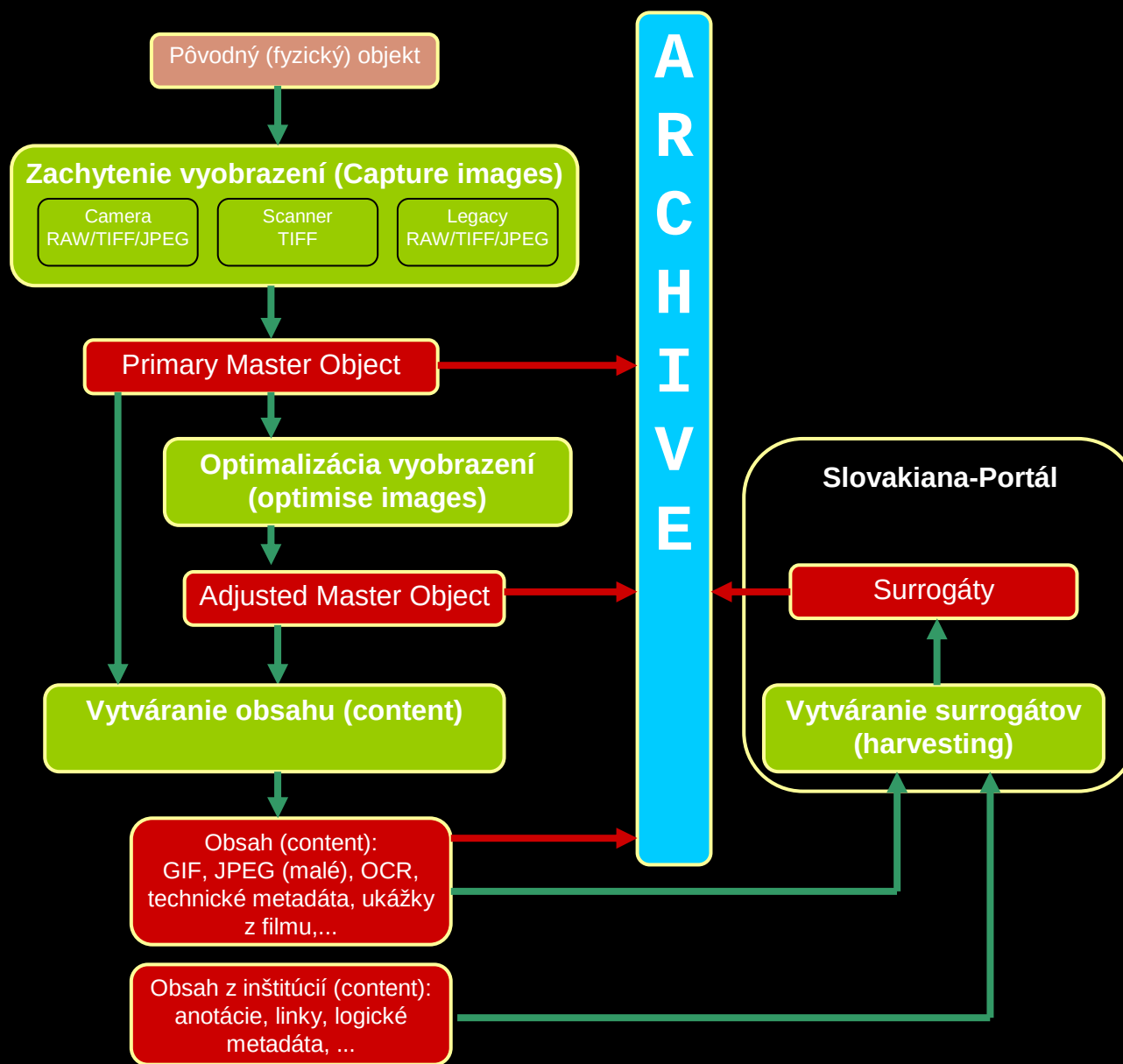


## Infraštruktúra pre Slovakianu – Ohraničenie

- Funkčné požiadavky Europeany (národný agregátor, štandardy)
- 3 lokality, vzájomné zálohovanie, Disaster recovery
- Centrálné úložisko obsahu (thumbnails) pre inštitúcie
- Content management
- Harvestovanie do Slovakiana-Portál
- Dlhodobá archivácia
- Škálovateľnosť

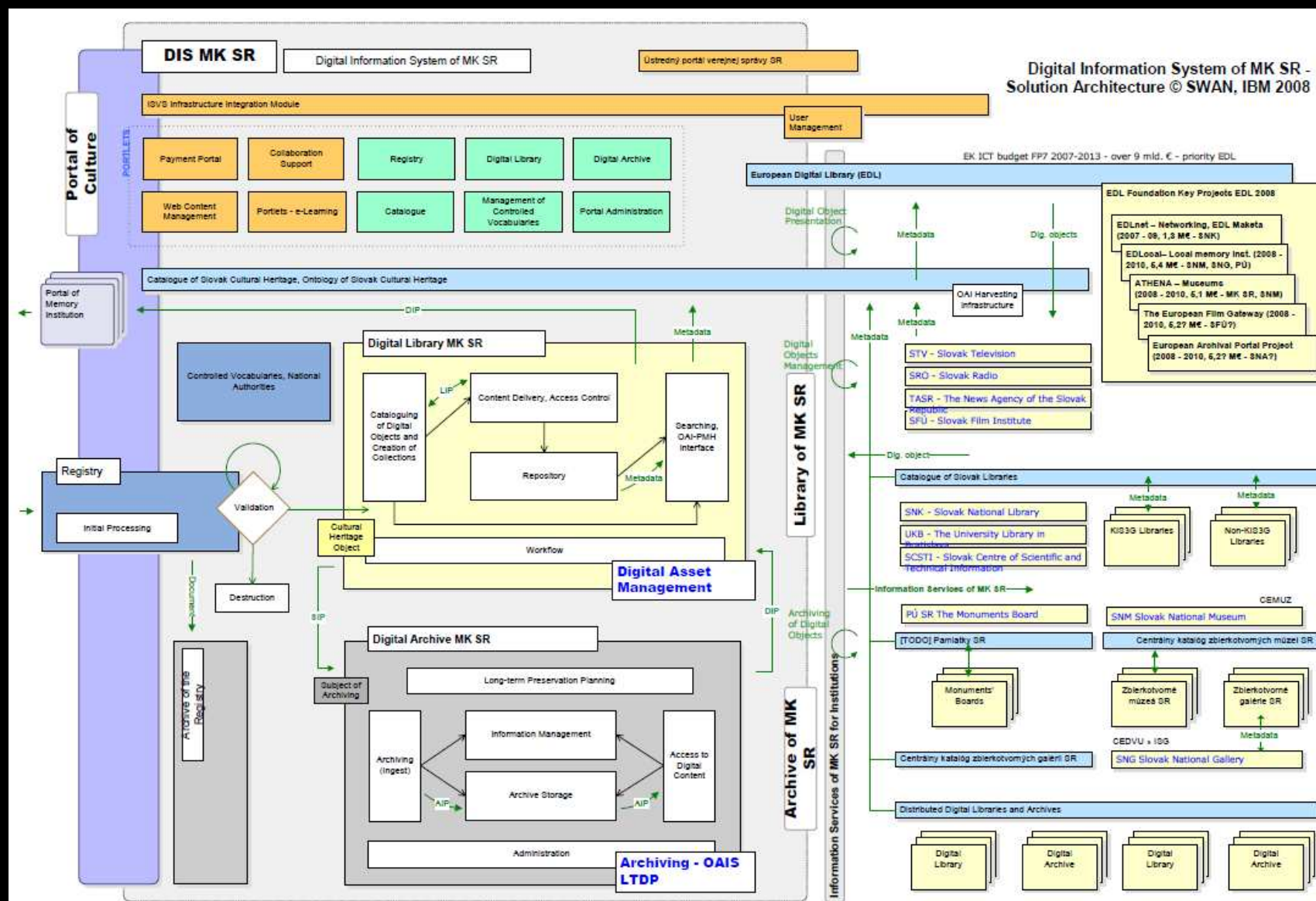


# Všeobecný proces digitalizácie





# Architektúra systému Slovakiana



## Funkcia dátových úložísk v systéme...

- Dátové úložiská sú kritickou časťou celého systému sú jeho srdcom
- Všetky digitálne objekty budú uložené v dátových skladoch
- Dáta budú zaevidované, oindexované a systematicky uložené do databáz a zložiek podľa vopred určeného systému
- Dáta budú virtuálne uložené na rôznych úložných systémoch podľa veľkosti, častosti prezerania, rýchlosti prístupu, ako aj podľa katalógov a budú chránené proti hackerom
- Pre užívateľa budú dáta prístupné ihneď a nepretržite cez internetový prehliadač, budú zdieľané a ľahko dostupné



# Princípy a komponenty riešenia

- Spracovanie masívneho množstva dát  
Terrabyty až Petabyty počas desiatok rokov
- Infraštruktúra musí byť robustná a redundantná a spĺňať náročné záťažové kritériá, zároveň automatizovaná a Green
- Celý systém musí byť postavený podľa otvorených štandardov IT a využívať komponenty Open source ako aj overené postupy a štandardy
- Systém musí byť prístupný v režime 7x24x365, flexibilný, dopĺňateľný a otvorený pre najnovšie technológie





# Komponenty a ich funkcia

- Overené technológie IBM na poli serverov ako aj úložných pamäťových systémov a softvéru
- Na úrovni rýchlej dostupnosti dát využijeme diskové polia IBM DS 5000, kvôli vysokému výkonu, dostatočnej kapacite, spoľahlivosti a rozširovateľnosti budú tvoriť tzv. Rýchlu cache pamäť pre Páskový automatický archív
- Virtualizáciu zabezpečí SAN Volume Controller = zdieľanie dát medzi užívateľmi, efektívne ukladanie dát a ich zrkadlenie medzi lokalitami
- Automatizované robotické páskové knižnice IBM na uloženie všetkých dát do digitálneho archívu a pre všetky formáty dát
- TSM – automatické ukladanie dát





## Referencie

- Prvá stavebná sporiteľňa = 1. projekt digitalizácie všetkých dokumentov na Slovensku, scan zmlúv, elektronický obeh dokumentov, ich manažment, automatické spracovanie, a následné uloženie do digitálneho archívu, prístup mobilných pracovníkov do archívu cez Web rozhranie, z akéhokoľvek miesta 24 hod. denne
- Spracovanie 2 Mio zmlúv so všetkými súvisiacimi dokladmi, obrazmi, podpisovými vzormi, na ďalších 30 rokov
- Podpora vedenia PSS počas celého trvania projektu 24 mesiacov

## Referencie SLK

- SLSP centrálné úložisko všetkých dát bankového systému, DWH, CRM , digit. archív 500 TB, DR
- VÚB centrálné úložisko dát, DWH, digit. Archív, virtualizácia, DR
- ALLIANZ SP centrálné úložisko dát, register. Digit archív, SAP pre 6 EU krajín, DR
- Sociálna poisťovňa úložisko dát a register poistencov
- T-mobile úložisko dát, DWH, digit. Páskový archív
- SPP centrálné úložisko dát, DWH pre SAP, DR
- SK telekom digitálny archív,
- Tatrabanka centr. úložisko v digit. Archíve, pásk. knižnice
- USSTEEL centrálné úložisko dát, manažment a dig. Archív a ďalší

## Referencie zahraničné

- Národná knižnica Holandského kráľovstva
- Ermitáž Ruská federácia, <http://www.hermitagemuseum.org>
- Národné kultúrne dedičstvo Egypta  
<http://www.eternalegypt.org>
- Digitalizácia archívu filmov Warner Brothers
- Archív CNN
- Európsky Hydrometeorologický ústav, ústavy krajín EÚ
- Walt Disney Studios archív
- Digitalizácia pamiatok vo Vatikáne – Pieta
- RTL tv CASE study
- C.E.R.N. Švajčiarsko a ďalšie

# CASE Study : RTL Television Kolín, Germany



## RTL Group

Najväčší Broadcaster v Nemecku

36 TV, 33 Radio Staníc v 11 krajinách

1.182 zamestnancov / \$1.761 Mio.  
Obratu

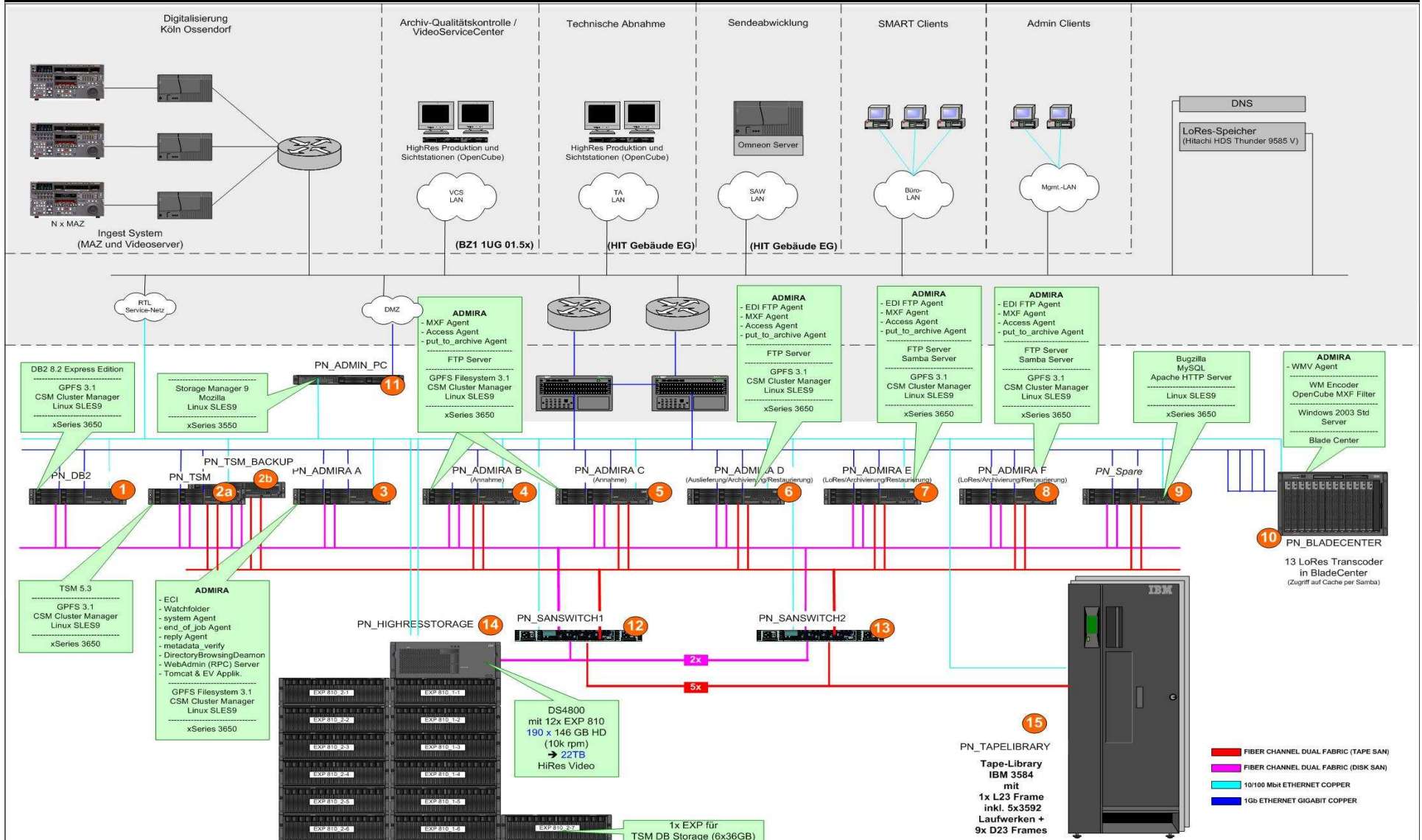
Globálny Leader v Content Manažmente

300 Programami v 40 krajinách

8.500 Hodín Programu za rok



# Design : rovnaké komponenty ako pre MKSR



Ďakujem za pozornosť

## RTL požiadavky a výzvy projektu

- Strategický Projekt na digitalizáciu Audio a Video Archivov
- vymeniť starý archív za nový a zachovať všetky dáta
- Obdobie od vzniku až po rok 2013
- 147.000 Hodín Video Contentu
- Znížiť náklady
- Vysoké požiadavky na bezpečnosť dát
- Zlepšiť dostupnosť dát na 7 x 24 hodín
- Nároky na kapacitu > 50 PB

## Riešenie IBM

### Global Business Services

- Koncept a Implementácia Contentbank
- Essence a dátový Management postavený na IBM ADMIRA
- IT-Technológie:
  - IBM TS3500 Pásková knižnica s Enterprise Tape mechan. IBM TS1120
  - IBM Cartridges
  - IBM DS4800 Storage Diskové polia
  - IBM System x Server
- Software:
  - IBM Tivoli Storage Management, DB2
  - IBM General Parallel File System



## Starý RTL Program-Archív ( Video pásky)



# Nový RTL Digitálny Archív - IBM pásková knižnica



18/12/2006



# IBM TS3500 pohľad dovnútra knižnice

