



Národní digitální knihovna

Úskalí udržitelnosti a rozvoje v praxi

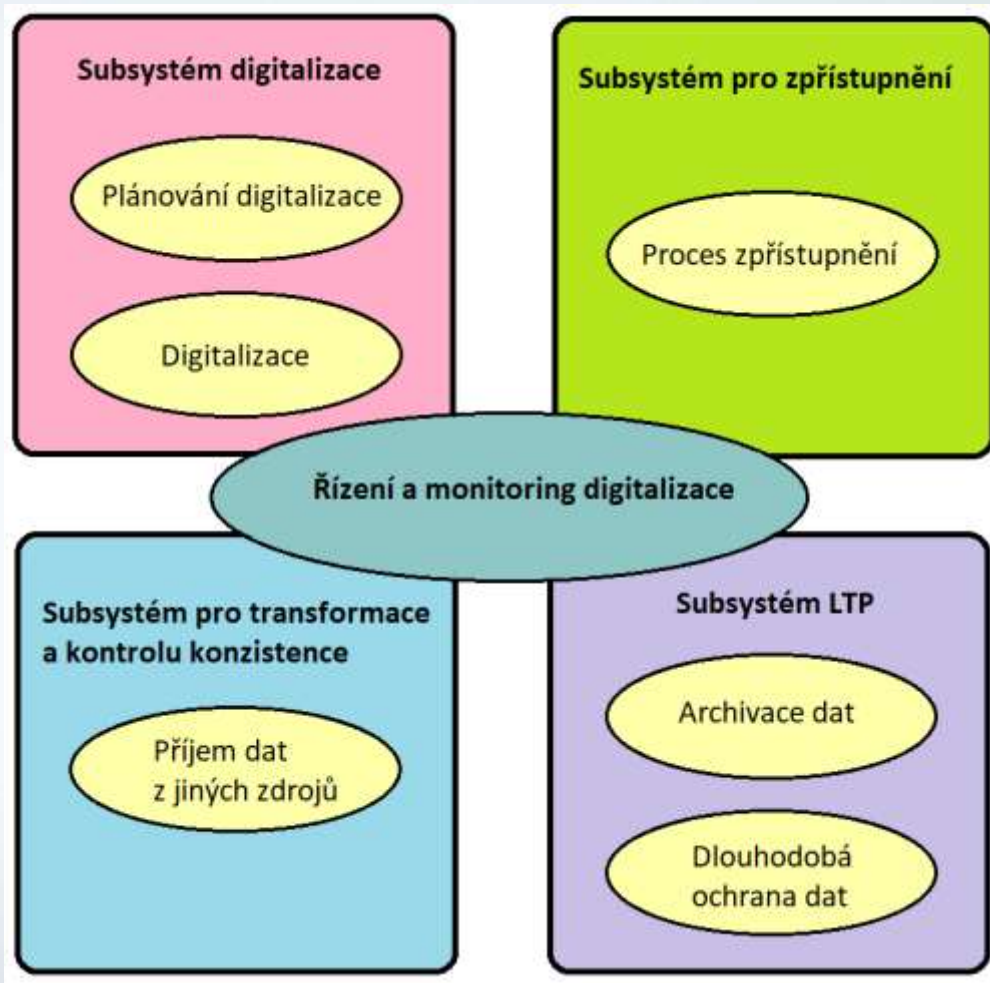
Petr Kukač

Konference CDA, 8. 11. 2018

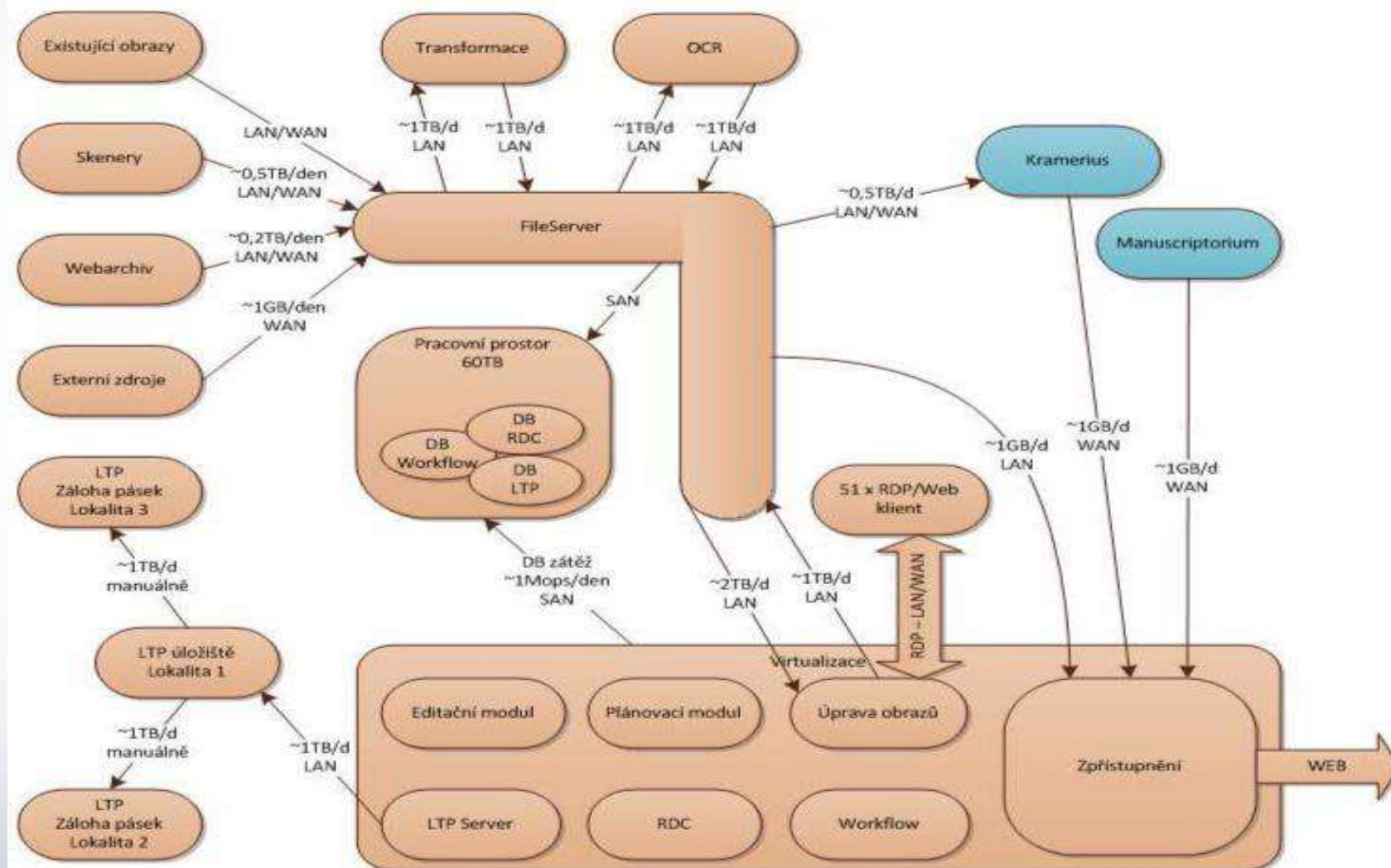
Poslání, úkol, význam

- ❖ **Digitalizovat moderní bohemikální produkci:** ve finále až 310 milionů stran
- ❖ **Zachránit obsah z ohrožených fondů,** ale i formu
- ❖ **Dát vzniknout webarchívu,** provádět systematické sklizně
- ❖ **Zpřístupnit fondy digitální formou,** částečně namísto tradiční
- ❖ **Bezpečně uložit digitální objekty na desítky let**
- ❖ **Bezpečně uložit digitální objekty jiných institucí**

Procesní model systému



Logické členění



HW vybavení

- **14x blade servery** (SQL, virtualizace, digitalizace)
- **3x rack servery** (file server cluster, zálohování)
- **124 TB diskových úložišť**
- **archivační appliance** IBM Information Archive
- **zálohovací páskové knihovny**
- **SAN infrastruktura**
- **51 pracovních stanic**
- **UPS, klimatizace, ...**

- **1 centrální datové (výpočetní) centrum v Praze**
- **2 digitalizační pracoviště (Praha a Brno)**

Skenery

<i>Typ</i>	<i>NK ČR Praha</i>	<i>MZK Brno</i>
4DigitalBooks DL 3003	1	1
4DigitalBooks DL Mini	1	1
4DigitalBooks Scan2Page	1	0
i2S SupraScanTM Quartz A0 HD	0	1
Treventus-ScanRobot® 2.0 MDS	2	2
Canon DR-X 10C	1	1
Plustec OpticBook A300	3	3
i2S SupraScanTM Quartz A1 HD	1	0
4Digital Books ScanVpage 600 Jumbo	1	0

SW vybavení

- ✓ **Framework SAFE**
- ✓ **Digitalizační workflow**
- ✓ **Logica Scanning Application**
- ✓ **Transformační modul**
- ✓ **Editační modul**
- ✓ **ScanTailor**
- ✓ **ABBYY Recognition Server**
- ✓ **Long Term Preservation**
- ✓ **Řídící a dohledové centrum**

Kramerius

Výkony

- ❖ Celkový počet digitalizovaných stran do XII/2014:
26 milionů
- ❖ Počet stran digitalizovaných za 1 den na 1 pracovišti:
13 200
- ❖ Objem dat uložených v LTP do XII/2014:
623 TB
- ... dnes ...*
- ❖ Počet stran digitalizovaných k 2. 11. 2018:
51 658 068
- ❖ Průměrný denní výkon v počtu stran na 1 pracovišti:
9 000

Nové požadavky

- Vývoj standardního HW – uplatňuje se při nutnosti výměny, částečně odcloněno virtualizací
- Vývoj operačních systémů, SQL serverů, JAVA, apod. – ukončení podpory, změny v licencování, portování aplikací, nutné updaty
- Vývoj „konfekčních“ aplikací – ABBYY Recognition server, GIMP, Adobe – nové verze, licenční modely
- Ukončení podpory specifických aplikací – u „open source“ sice nižší riziko, ale když k tomu dojde, není nástupce; u komerčního je obvykle jasný navazující produkt a garance migrace, musí však být vysoutěžený a zakoupený
- Vývoj knižních skenerů umožňuje skenovat ve vyšší kvalitě, ale také v nových typech a verzích souborů

Nové požadavky

ještě pokračování ...

- Roste celkový objem uložených dat: nutno rozšiřovat kapacity pro LTP i provozních úložišť
- Roste počet titulů k zpřístupnění: nutno hledat výkonnější systém, Kramerius v současné verzi je na hranici možností
- Změny ve výběru předloh: fyzický stav se zhoršuje (od destrukce přes automaty k ručnímu skenování) – s tím souvisí pracnost a celkový výkon
- Posun v typu předloh: mono x perio, s tím souvisí velikost dat jedné dávky, pracnost na přípravě i dalších pracovištích
- Větší důraz na kvalitu: dodatečné opravy jsou náročnější než kvalita na „první dobrou“

Nové požadavky

další pokračování ...

- Vývoj souborových formátů – TIFF5 -> TIFF6; souvisí spíše s LTP
- Vývoj norem, specifikací, legislativy – GDPR, Metamorphose, FADGI (1, 2, ..., 4) ...
- Vývoj metatadových specifikací (2016, 2018) – změny hlavně u periodik, snahy o zavedení nových typů dat
- Urychlení konverzí dat ze starších digitalizací, jejich zpřístupnění a dlouhodobé uložení
- Příjem a zpracování e-bornů: definice standardů, formáty pro uložení i zpřístupnění, nástroje pro validaci a zobrazení, nové postupy,...
- Digitalizace dalších fondů jiných než papírové: audio, video, obsah optických nosičů, disket, ...

Co s tím?

- Maximální zainteresovanost vedení knihovny
- Jasná strategie, pevné zakotvení v dlouhodobé koncepci instituce
- Zvýšení povědomí vlastních zaměstnanců o významu a prospěšnosti digitalizace pro celou knihovnu
- Alokace finančních prostředků
 - na běžný provoz
 - na investice do obnovy
 - na rozvoj
 - **NA LIDI**

Nestačí? Co ještě?

- Propagace významu digitalizace u veřejnosti
 - Pro záchranu a uchování
 - Pro dosažitelnost
- Zapojení do mezistátních aktivit, meziresortních projektů, ...
- Celkové zvýšení atraktivnosti digitální knihovny pro čtenáře

Společenská poptávka

Významnost pro zřizovatele

Priorita pro stát

Spolehlivé a garantované financování na léta dopředu

Děkuji za pozornost

Petr Kukač
Národní knihovna ČR
Odbor digitalizace
petr.kukac@nkp.cz