unixshop | LinuxBizWorld | AbcLinuxu | 64bit | PráceAbc | dFklub | SMSinzerce | 2cars | stickfish.cz

Školení OKsystem

- Administrace SUSE Linuxu
- Administrace SUSE Linuxu
- Základy SUSE Linuxu
- Využití operačního systému Linux ve školském prostředí
- Základy SUSE Linux Enterprise Server 10

Najdi školení

Anketa

Na MTA mám povoleno TLS šifrování:

pouze pro příchozí spojení (8%)

pouze pro odchozí spojení (1%)

pro všechny spojení (21%)

šifrování úplně zakazuji (4%)

nespravuji žádný mail server (66%)

Hlasuj Celkem 457 hlasů

Komentářů: 32, poslední dnes 17:09
Navrhněte novou anketu

Zprávičky

Centrum | Napsat zprávičku

DJVU versus PDF

dnes 15:53 | Přečteno: 171x | Za vším hledej Linux | poslední úprava: dnes 15:53

Ve svém minulém příspěvku jsem se pokusil upozornit na možnost konverze postscriptových dokumentů do formátu DJVU a jen velice zběžně jej představil jako formát podobný co do funkčnosti formátu PDF. Shodou náhod kolega obdržel k připomínkování návrh direktivy o doporučených formátech pro výměnu dat, která doporučuje ODT formát pro dokumenty které nemají konečnou revizi a PDF pro dokumenty které jsou již uzavřené.

Vzhledem k tomu, že PDF dokumenty lze bez problému převést na DJVU vcelku o nic nejde, ale proč ještě převádět PDF do nějakého DJVU? Podotýkám že smyslem zápisu není pohanět formát PDF, spíše zdůraznit přednosti formátu DJVU tam, kde má PDF slabiny.

Kde kulhá PDF

I když dva dělají totéž, není to totéž. Rozdíl mezi PDF a DJVU je kromě struktury dokumentu také v typu použité komprese. PDF má podobně jako postscript lineární zápis tj. prohlížeč čte dokument a postupně vykresluje objekty specifikované pomocí souřadnic a atributů na stránku. Text je renderován podle použité znakové sady a případných dalších nastavení. Ovšem obrázky jsou vloženy v kvalitě odpovídající určení dokumentu. Využívá se např. toho, že v případě dokumentu určeného pro čtení z monitoru je zbytečné pro ně nastavovat vyšší rozlišení než 100dpi. Vnitřní struktura PDF dokumentu se může lišit také v závislosti na tom jak byl vytvořen, přesto že na první pohled vypadá pokaždé v prohlížeči stejně.

Například dokument uložený přímo do PDF např. z OpenOffice, nebo vytištěný přes CUPS-PDF bude malý, protože text není uložen jako vyrenderovaný obrázek, ale přímo jako text. Má to však jeden háček. Pokud si tento dokument otevře na svém stroji někdo, kdo nemá nainstalované písmo, jaké bylo použito při generování PDF dokumentu, tak se může stát, že místo textu uvidí jen hromadu čtverečků. Tomu lze zabránit tím, že se do PDF dokumentu vloží i font. Ovšem tím jeho velikost poněkud nabobtná.

Další možností je, text renderovat jako grafiku, to má ale háček v tom, že jej nelze dále editovat, ani prohledávat. Poslední možností je vložit celý dokument do PDF jako obrázek. Casto se to tak dělá, pokud obsahuje nějaké podpisy, razítka, atp. soubor tím ale řádně nakyne - možnost jeho fulltextového prohledávání může být zachovaná pokud do něj OCR aplikace vloží rozpoznaný text. Možnost takto zpracovat PDF nabízí většina moderních OCR aplikací - bohužel většina z nich je "only for MS Windows"

V DJVU dokumentu je každá stránka samostatným objektem, který na rozdíl od PDF rovnou pracuje s několika vrstvami. Základem je obrazová vrstva, která je přítomna vždy. Kromě ní mohou být přítomny i další textové vrstvy s metadaty, hyperlinky, skrytým textem obsahu stránky a poznámkami. Díky wavelet kompresi je však taková stránka co do objemu dat vždy menší než u PDF. Naopak čím vyšší rozlišení měl výchozí dokument, o to lepší výsledek komprese.

Komprese grafických objektů

Zatím co DJVU používá ke kompresi obrázků primárně vlastní wavelet kompresi IW44, Wavelet kompresi JPEG-2000 začalo do PDF Adobe implementovat až od Acrobata verze 6, ovšem kvůli zachování kompatibility se staršími prohlížeči se stejně používají u PDF dokumentů klasické komprese na

Aleš Kapica - kenyho_stesky

Rozepsané zápisy

- Distribuce - cesta do pekel?
- Až Linus zaklepe na nebeskou bránu...
- Bože to je opruz.. proč nejsem chytřejší..
- Připojení stroje s OS Windows do Linuxu přes Sambu II.

Aktuální zápisy

- DJVU versus PDF
- Použití DJVU pro skladování elektronické dokumentace
- Použití DJVU pro skladování elektronické dokumentace
- Jaký typ linuxového souborového systému upřednostňujete?
- Co takhle rovnou ospamovat blázna?

Archív ?

- červenec 2006 (3)
- červen 2006 (4)
- květen 2006 (1)
- duben 2006 (4)
- březen 2006 (7)
- únor 2006 (2)

Navigace ?

- kenyho_stesky, hlavní strana
- kenyho_stesky, stručný souhrn
- RSS kanál
- Všechny blogy
- Všechny blogy, stručný souhrn

Nastavení ?

- Uprav zápis
- Smaž zápis
- Vlož anketu
- Vlož nový zápis
- Nastavit blog
- Přejmenovat blog
- Upravit kategorie
- Upravit odkazy

dnes 15:44 | Zajímavý software

Feather Linux je live distribuce založená na Knoppixu, která se vejde i na flash disk o kapacitě 128MB (přehled software). Na Linux.com je představována jako distribuce s dobrou komunitou, ale nejistou budoucností.

Luboš Doležel | (Komentářů: 0)

dnes 15:32 | Komunita

JRuby je plně kompatibilní interpreter Ruby napsaný v Javě. LinuxJournal přináší rozhovor s vývojáři tohoto software - Charlesem Nutterem a Thomasem Enebo. Baví se mj. o přínosu JRuby pro komunitu Ruby nebo o kvalitách Javy.

Luboš Doležel | (Komentářů: 0)

dnes 15:23 | Zajímavý článek

Článek na TechTarget učí uživatele OpenOffice.org Writeru využívat všech funkcí, které tento software nabízí při práci se seznamy. Podrobně se rozebírá například číslování nebo odsazování.

Luboš Doležel | (Komentářů: 0)

dnes 11:32 | Zajímavý článek

Rádi byste přistupovali ke svému počítači odkudkoliv z internetu, ale brání vám v tom dynamicky přidělovaná IP adresa od vašeho providera? Jak se s tím vypořádat radí článek na serveru tips.linux.com.

Jiří "Geo" Lužnický | (Komentářů: 3)

dnes 11:22 | Nová verze

Vyšlo Gnome 2.15.4, další z verzí na cestě k 2.16.0. Seznam změn je poměrně obsáhlý. Je však třeba upozornit na velkou změnu v gnome-vfs (některé funkce se přesunuly do libbonobo) a problémy se spouštěním některých programů. Podrobnější seznam změn a návod jak řešit výše uvedené problémy najdete v upozornění na Gnomedesktop.org.

Jiří "Geo" Lužnický | (Komentářů: 1)

dnes 11:15 | Zajímavý článek

Server Softpedia.com přinesl recenzi známého VoIP programu Ekiga. Ze špatných vlastností zmiňuje existenci chyb v programu (kterých je sice málo, ale jsou), neexistující podporu USB telefonů a kiosk módu. Na druhou stranu chválí, že program podporuje jak SIP, tak H.323. Dále pak kvalitní adresář a dobře fungující dockování do panelu.

Jiří "Geo" Lužnický | (Komentářů: 1)

včera 19:02 | IT novinky

Po dlouhé době testování je venku nová verze Unicode standardu. Mezi novinkami je např. pět nových písem (na své si přijdou především semitisté s féničtinou a sumersko-babylónským klínopisem), mnoho jednotlivých znaků, ale též vylepšené algoritmy a konzervativnější politika přijímání změn. Více na stránkách Unicode Konsorcia.

Petr Tomášek | (Komentářů: 9)

včera 11:52 | IT novinky

Virtualizace pomocí Xenu byla dlouho použitelná jen k běhu open source systémů. Lidem z XenSource se však překvapivě podařilo získat podporu ze strany Microsoftu, a tak bude možno použít Xen k paralelnímu běhu Linuxu a Windows na jednom fyzickém systému.

Luboš Doležel | (Komentářů: 14)

Red Hat | Reklama

Pracujte pro Red Hat jako administrátor GNU/Linuxu!

včera 11:41 | Zajímavý článek

Lukáš Hasík si při pokusu o zvládnutí kvízu o historii NetBeans uvědomil, že toho zas tak moc neví. Proto ve svém blogu uveřejnil zajímavé odkazy, kde se můžete o historii tohoto vývojového prostředí něco dozvědět.

Jiří "Geo" Lužnický | (Komentářů: 4)

včera 11:34 | Zajímavý software

Vyšla první verze programu Dates. Tento malý a jednoduchý kalendář, který používá pro ukládání dat Evolution Data Server, je primárně určen pro kapesní počítače, jako je Nokia 770. Jeho GUI je ale napsané i v čistém GTK+, takže by neměl být problém program používat i na jiných zařízeních.

bázi LZW a JPEG.

Práce s poznámkami a vloženým textem

Vezměte prosím v úvahu, že toto téma zpracovávám z hlediska open source a Linuxu zvlášť, tudíž komerční aplikace pro jiné systémy zcela opomíjím. Ba troufám si říct, že v Linuxu lze kupříkladu pro práci s DJVU nalézt nástroje lepší. Pokud jde ale o možnost práce s poznámkami a textem ve vytvořeném PDF, tak mi to spíš připomínalo procházku slzavým údolím.

Už, už jsem se chystal naříkat, že žádná z volně dostupných aplikací pro linux neumožnuje u PDF práci s poznámkami a pak jsem ji našel! Multivalent ale má několik pih na kráse. Je napsán v Javě, tudíž vyžaduje nainstalovanou javu od Sunu a.. ..je pomalý jako všechny aplikace v javě. Prakticky totéž ovšem lze vytknout i čtečce djvu dokumentů Java DjVu Viewer. Výhodou je ale fakt, že je lze použít všude tam kde se dá java od Sunu nainstalovat. Druhou pihou na kráse je že nepracuje s unicode, takže na diakritiku s klidem zapomente, i když se domnívám že to bude spíš problém poznámek v PDF.

U DJVU dokumentů není problém k editaci textových údajů a poznámek použít nástroje které jsou součástí DjVuLibre, co bohužel prozatím schází je nějaký grafický nástroj, který by usnadnil orientaci v textové vrstvě a zjednodušil vkládání hypertextových odkazů a poznámek, ale třeba se toho někdo chytí jako zajímavého a užitečného projektu.

Doporučené stránky

Pro vás, co máte hlubší zájem o srovnání možností formátu DJVU mám dva odkazy na práce Adolfa Knolla, obé ovšem v jazyce anglickém:
Efektivnost wavelet komprese
Testování nových obrazových formátů

Hodnocení

Stav: 2,62 Počet hlasů: 8

0 1 2 3

Nástroje: Tisk

Komentáře

Nástroje: Sleduj (0) ?, Tisk

Vložit další komentář

dnes 16:31 Kníže Ignor | blog: stoupa
Re: DJVU versus PDF
Odpovědět | Admin | Link | Blokovat | Sbalit

Je to sice v příspěvku obsažené, ale chtěl bych zdůraznit, že djvu opravdu nelze považovat za náhradu pdf. PDF toho umí mnohem víc, je daleko univerzálnější. DjVu si sice vede lépe v naskenovaných dokumentech, ale pro jiné věci se už zas tak moc nehodí.

Joe se rozhodně nechce vzdát svých banánů a přenechat je mafiánovi!

dnes 16:38 Aleš Kapica | blog: kenyho_stesky | Ostrava
Re: DJVU versus PDF
Odpovědět | Admin | Link | Výše | Blokovat | Sbalit

Když už tak napiš které. Mě teda napadá především příprava dokumentů k tisku atp. Jak už jsem napsal DjVu je vhodné především k elektronické konzervaci dokumentů, tak aby bylo možno s nimi dále pracovat (ve smyslu studovat je).

dnes 17:09 Kníže Ignor | blog: stoupa
Re: DJVU versus PDF
Odpovědět | Admin | Link | Výše | Blokovat | Sbalit

No jaký má smysl do djvu převádět pdfka, které obsahují především text (třeba matematický text)? Že se mi někdy zmenší velikost je celkem dost nezajímavé oproti nevýhodám, které znamená převedení na obrázek (ve kterém lze díky přítomnosti textové vrstvy vyhledávat ☺). Je to podobné jako převádět vektorovou grafiku do bitmapy.

Joe se rozhodně nechce vzdát svých banánů a přenechat je mafiánovi!

dnes 17:58 Aleš Kapica | blog: kenyho_stesky | Ostrava
Re: DJVU versus PDF
Odpovědět | Admin | Link | Výše | Blokovat | Sbalit

Třeba takový..

```
user@stroj:~/djvutesty$ ls -alh epilog*
-rw-r--r-- 1 user users 157K 2006-07-19 17:48 epilogue.dj
-rw-r--r-- 1 user users 2,6M 2006-07-19 17:47 epilogue.ps
-rw-r--r-- 1 user users 282K 2006-07-19 17:47 epilogue.pd
```

Do DJVU převedeno přes `djvudigital` včetně textové vrsty rovnou z postscriptu (Zkusil jsem namátkou první matematickou věc na kterou jsem narazil)

dnes 18:01 Aleš Kapica | blog: kenyho_stesky | Ostrava
Re: DJVU versus PDF
Odpovědět | Admin | Link | Výše | Blokovat | Sbalit

Jiří "Geo" Lužnický | (Komentářů: 3)

Centrum | Napsat | Starší

A že nekecám, tak tady máš vytažených prvních pár řádek rovnou z toho djvu souboru:

> *The association between mathematics and graphics is ancient. Indeed, as the scholar of Greek mathematics T. L. Heath has observed, the Greek root of the word 'graphics' seems to mean in places 'to prove'. But the association is far*

Založit nové vlákno

Pracovní nabídky

- Security Architect (Any/15.2.1.4.1n)
- Programátor C/C++/Delphi
- Technical Specialist (Middleware) (Any/2.5.3.1n)
- Školitel - Unix
- Systémový architekt
- GNU/Quality Assurrance Engineer
- Programátor Java
- .NET Developer
- Application Support Analyst 2 (Billing) (Any/4.1.1.)
- Junior Unix Specialist

Rozcestník

AbcLinuxu

- Abíčko: červenec 2006
- Airline Tycoon Deluxe

LinuxBizWorld

- OSS Alliance připravuje unikátní projekt
- Red Hat podporuje OVAL formát pro bezpečnostní chyby

Chip

- Questia chce vybudovat kvalitní on-line knihovnu
- Odinstalování Service Packu 2

Svět Hardware

- Intel plánuje levný "Conroe": Core 2 Duo E4300
- BWU-100A: Blu-ray vypalovačka za 750 dolarů

Portál AbcLinuxu

- RSS a PDA
- Tým kolem Abíčka
- Jak se stát autorem
- Jak pomoci vylepšit Abíčko
- Jak pomoci s propagací Abíčka
- Hlášení chyb a námětů
- Vzkazy správcům (12)
- Inzerce

Doporučujeme

- Přidej sidebar
- Přidej mezi oblíbené
- Nastav jako domácí stránku
- linux.cz
- unixshop.cz
- pravednes.cz

Placené odkazy

- autoweb.cz
- tricko-tricka.com
- Bělení zubů White Pearl

ISSN 1214-1267 Powered by Stin

© 1999-2006 Stickfish s.r.o. Podmínky použití | Osobní údaje
Všechna práva vyhrazena.