

RANCANG BANGUN APLIKASI PENGAMBILAN BERITA SECARA OTOMATIS MENGGUNAKAN *CONTENT SYNDICATION* BERBASIS XML DENGAN PLATFORM MICROSOFT .NET

Diana Purwitasari, Febriliyan Samopa, Ade Afrian

Jurusan Teknik Informatika,

Fakultas Teknologi Informasi, Institut Teknologi Sepuluh Nopember

Kampus ITS, Jl. Raya ITS, Sukolilo – Surabaya 60111, Tel. + 62 31 5939214, Fax. + 62 31 5913804

Email: diana@its-sby.edu , iyan@its-sby.edu,

ABSTRAK

Banyaknya kebutuhan akan informasi di internet menyebabkan penyedia jasa situs berita untuk memberikan berita yang selalu yang terbaru. Salah satu alternatif solusi adalah dengan melakukan *content syndication*. *Content syndication* adalah proses dimana suatu isi berita dikirimkan atau disediakan, biasanya dengan biaya tertentu, dari penyedia berita, biasanya disebut *originators*, ke pasar yang membutuhkan atau *subscribers*. *RSS (Rich Site Summary)* adalah format yang secara umum digunakan untuk melakukannya. *RSS* pada dasarnya adalah suatu file yang berada di suatu situs, yang menyediakan informasi tentang isi dari situs tersebut. File tersebut biasa disebut sebagai *RSS Feeds* dan dapat di ambil dan diolah untuk mendapatkan informasi tentang isi situs tersebut.

Dibuat sebuah aplikasi untuk pengambilan situs berita secara otomatis menggunakan *content syndication* yang memerlukan aplikasi pada proses background untuk mengambil *RSS Feeds* secara berkala pada komputer yang berfungsi sebagai server. Server yang mengambil berita dari situs penyedia terdiri dari aplikasi yang mengatur konfigurasi berita tersebut, dan sebuah *windows service* untuk mengambil *RSS feeds* kemudian mengolahnya secara otomatis. Sedangkan aplikasi untuk membaca berita dari *RSS server* terdapat pada client berupa sebuah komponen *plug-in*.

Uji coba pertama dilakukan dengan menguji keberhasilan aplikasi dalam mengatur konfigurasi skema, atribut tabel, dan pengaturan kategori situs penyedia *RSS*. Sedangkan ujicoba kedua dilakukan dengan melakukan perbandingan hasil pencarian berita yang didapat dari program dengan berita dari situs lain yang tidak menerapkan *content syndication*. Dari hasil pengujian diketahui bahwa aplikasi dengan *content syndication* mampu melakukan pencarian berita dan memberikan hasil yang lebih baik.

Kata Kunci: *Content Syndication, RSS, Windows Service, Band Object.*

1. PENDAHULUAN

Content syndication adalah proses dimana suatu isi/informasi/berita dikirimkan atau disediakan, biasanya dengan biaya tertentu, dari penyedia berita yang biasanya disebut *originator* ke sejumlah besar distributor atau *subscriber*¹.

Model *syndication* telah digunakan selama beberapa dekade dalam industri berita dan pertelevisian. Pada tahun-tahun terakhir ini, model bisnis ini juga dengan sukses telah digunakan di internet. *Content syndication* secara *online* melibatkan koleksi informasi dari banyak sumber daya, pengepakan informasi tersebut dalam format yang standar, dan menyediakannya untuk para distributor yang mencari berita untuk penggunaan dalam platform media semacam website, situs WAP, dan outlet berbasis SMS.

Keberadaan pelaku *content syndication* untuk memperkenalkan skalabilitas dalam pasar. Oleh

karena itu pelakunya menjadi *middle-tier* diantara pembuat berita di satu sisi dan penerbit di sisi yang lain. Organisasi yang lebih suka mengadopsi solusi *content syndication* adalah organisasi yang mendistribusikan informasi ke jaringan yang luas. Secara umum, pelaksanaan secara on-line sama dengan secara off-line. Perbedaan yang mendasar berhubungan dengan penyekalaan dan kelangkaan. Pelaku sindikasi offline biasanya mempunyai jalur distribusi yang tetap dan terbatas seperti koran, radio, dan TV. Internet merubah hal ini dan mendistribusikan informasi dengan cara yang jauh lebih luas. Dengan *content syndication*, informasi apapun dapat berada dimana saja, karena sudah tidak ada lagi jarak antara pembuatan dan pendistribusian.

2. RDF SITE SUMMARY(RSS)

RSS adalah singkatan dari "*Really Simple Syndication*", "*Rich Site Summary*" atau "*RDF Site*

Summary"². RSS adalah format yang secara umum digunakan untuk melakukan *content syndication*. RSS pada dasarnya adalah suatu file yang berada di suatu situs, yang menyediakan informasi tentang isi situs tersebut. File tersebut biasa di sebut sebagai "RSS Feeds" dan dapat di ambil dan diolah untuk mengambil informasi tentang isi situs tersebut.

RSS dapat didiskripsikan sebagai aplikasi metadata XML sederhana berdasar RDF yang

berdasar pada konsep *channel*. Dalam hal ini *channel* adalah file XML yang berada di server dan mengandung elemen semacam judul, deskripsi, dan *link*. File XML ini kemudian diperoleh dari pengarang melalui penerbit yang biasanya merubahnya menjadi format HTML, WML atau sejenisnya dan menampilkannya dalam website mereka. Versi yang dipergunakan disini versi 2.0.

```

<?xml version="1.0" ?>
<rss version="2.0">
<channel>
<title>New York Times: International</title>
<link>http://www.nytimes.com/pages/world/index.html</link>
<description>Headlines, abstracts and links for the latest New York Times articles, for Radio
UserLand.</description>
<language>en-us</language>
<copyright>Copyright 2003 The New York Times Company</copyright>
<lastBuildDate>Thu, 10 Jul 2003 04:15:46 GMT</lastBuildDate>
<category domain="http://www.weblogs.com/rssUpdates/changes.xml">rssUpdates</category>
<docs>http://backend.userland.com/rss</docs>
<generator>UserLand Frontier v9.1b3</generator>
<webMaster>webmaster@userland.com</webMaster>
</channel>
</rss>

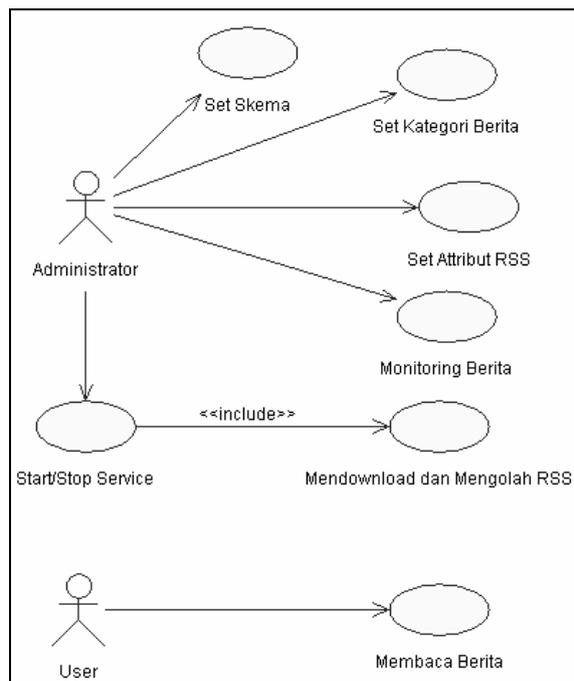
```

Gambar 1 Contoh File RSS 2.0

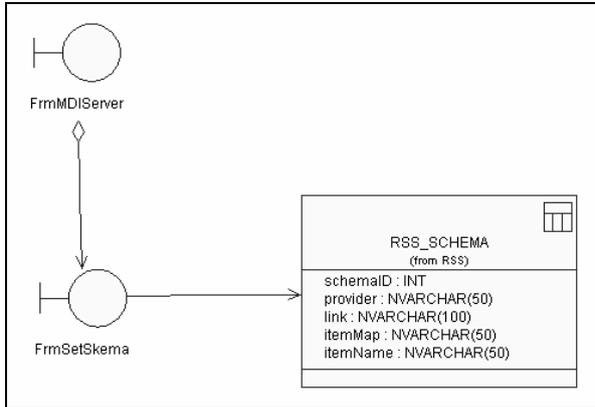
3. PERANCANGAN PROSES UNTUK CONTENT SYNDICATION

Windows service berfungsi untuk menjalankan aplikasi sebagai proses pada *background*. *Windows service* telah dieksekusi sebelum pengguna melakukan log/masuk ke sistem (jika *windows service* tersebut diatur untuk mulai pada proses *boot-up*). Selain itu juga dapat digunakan untuk mengakses fungsi-fungsi yang ada di server secara *remote*. *Windows service* juga diatur sedemikian hingga pengguna dapat melakukan *start* secara manual.

Server RSS untuk mengambil berita dari berbagai situs secara otomatis akan dijadikan sebagai *windows service* yang mampu melakukan pendeteksian apabila terdapat berita baru pada situs-situs penyedia berita. Sedangkan pada komputer pengguna akan dipasang komponen *plug-in* di Internet Explorer untuk melihat berita terbaru dari server RSS.



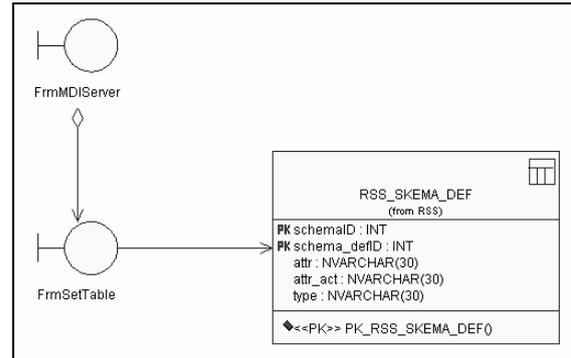
Gambar 2 Diagram Use Case Aplikasi Pengambilan Berita Secara Otomatis Menggunakan Content Syndication



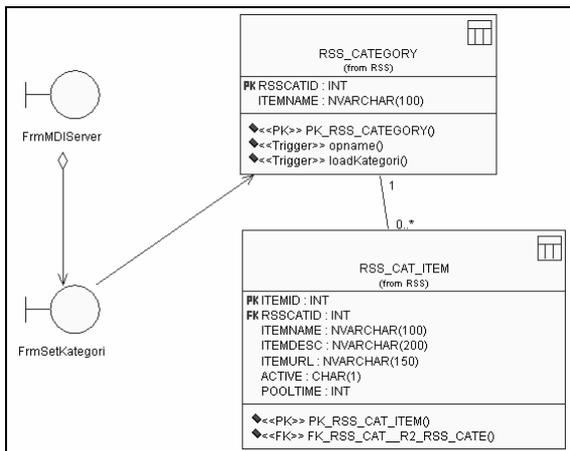
Gambar 3 Diagram Class Proses Set Skema XML dari RSS

Pada gambar 3 RSS Server dengan aplikasi administrator dilakukan pengaturan untuk menyiapkan skema XML dari RSS yang akan di *download*. Setiap *provider* menerapkan versi RSS yang berbeda, sehingga perlu disimpan setiap skema RSS. Hal ini juga ditujukan apabila RSS mengalami perkembangan ke versi selanjutnya.

ditujukan agar tetap dapat melakukan pengambilan berita dan pengolahan RSS walaupun *provider* memiliki versi RSS yang berbeda agar sistem tetap dapat berfungsi untuk versi RSS selanjutnya.



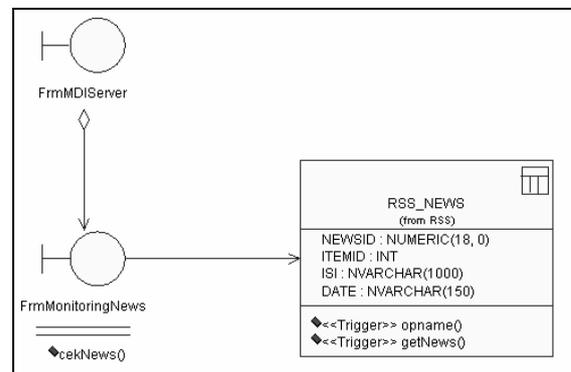
Gambar 5 Diagram Class Proses Set Atribut RSS



Gambar 4 Diagram Class Proses Set Kategori Berita

Kemudian di gambar 4 terjadi modifikasi kategori berita sesuai dengan kebutuhan serta pengaturan alamat situs penyedia berita untuk masing-masing kategori.

Selain itu administrator dapat memilih atribut elemen-elemen pada skema RSS untuk disimpan pada database pada gambar 5. Misal situs A memakai RSS versi 0.9 sehingga elemen-elemen yang perlu disimpan adalah *title*, *link*, *description*. Sedangkan situs B memakai RSS versi 2.0 sehingga elemen yang perlu disimpan adalah *title*, *link*, *description*, *pubDate*, *Author* dan *Guid*. Penyimpanan atribut

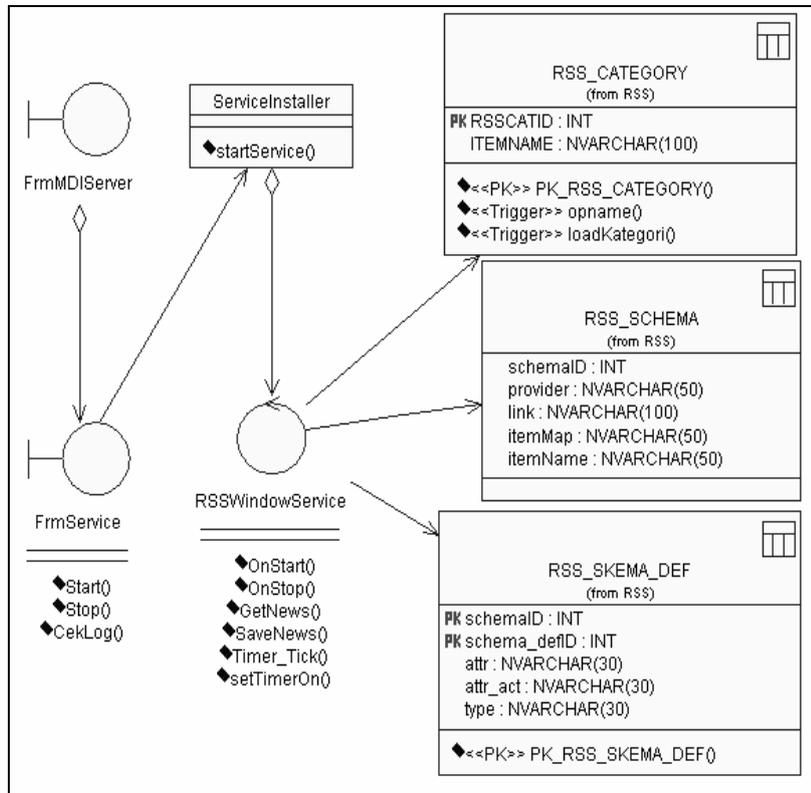


Gambar 6 Diagram Class Proses Memantau Berita

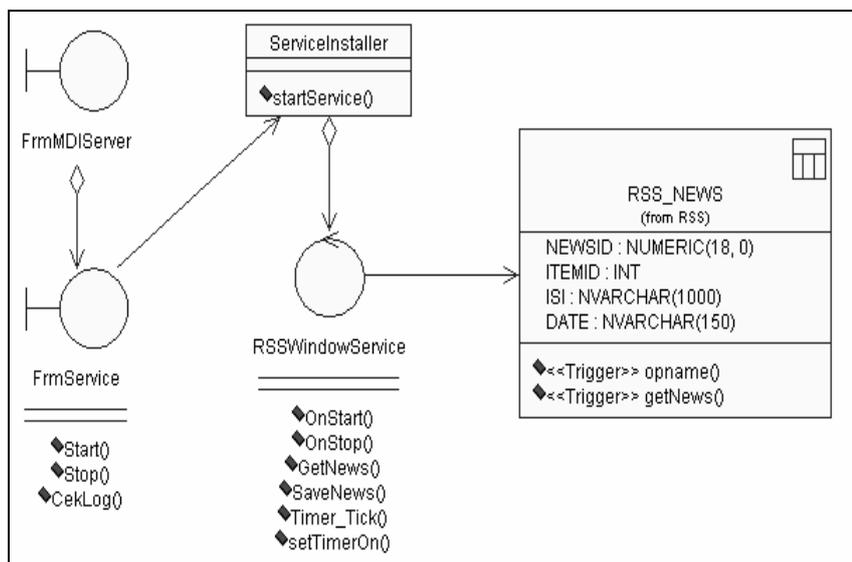
Pada gambar 6 terjadi proses untuk memantau hasil *download* dan pengolahan RSS feeds. Administrator dapat menampilkan daftar berita yang sudah pernah di *download* dan menyimpan data – data berita tersebut.

Ditunjukkan pada gambar 7 bahwa proses pengambilan berita dilakukan dengan tujuan agar aplikasi dapat melakukan *download* RSS feeds secara terus-menerus sesuai kategori yang telah ditentukan.

Ketika suatu RSS feeds telah berhasil di-*download*, maka pada gambar 8 RSS feeds itu akan ditampung kedalam suatu dokumen yang telah disiapkan skemanya. Sehingga setiap elemen atribut RSS dapat diambil dan disimpan kedalam database.



Gambar 7 Diagram Class Proses Pengambilan Berita

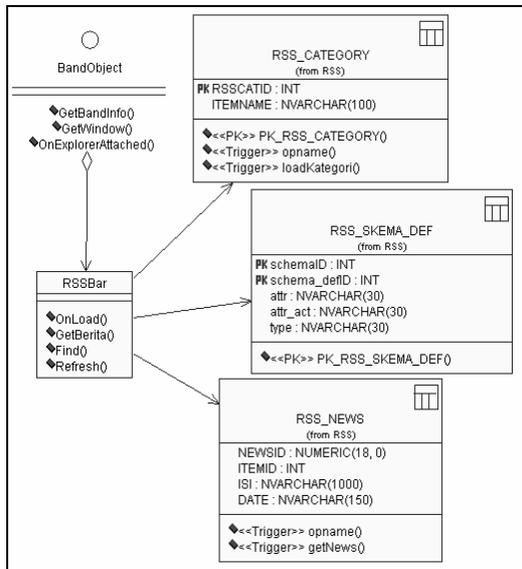


Gambar 8 Diagram Class Proses Pengolahan RSS

Pengguna yang ingin membaca hasil RSS feeds yang telah diambil dapat melakukan pembacaan hasil dengan terlebih dahulu menjalankan aplikasi Internet Explorer. Lalu memilih untuk menjalankan *explorer bar* pada bagian *RSSBar*. Lalu *Band Object* akan melakukan inialisasi untuk menampung *RSSBar* yang telah didefinisikan, sehingga menampilkan *explorer bar* RSS. Proses pengambilan berita dimulai dengan pengguna melakukan permintaan berita melalui, *class RSSBar* tersebut, lalu *class* akan melakukan proses pembacaan berita sesuai skema

RSS-nya. Setelah itu *class RSSBar* akan mengambil berita dari database.

Band Object diperkenalkan mulai Shell versi 4.71 dan menyediakan kemampuan sama seperti window biasa. *Band Object* adalah COM dalam sebuah *container*. Untuk *explorer bar* bentuk dari *container*-nya adalah Internet Explorer.



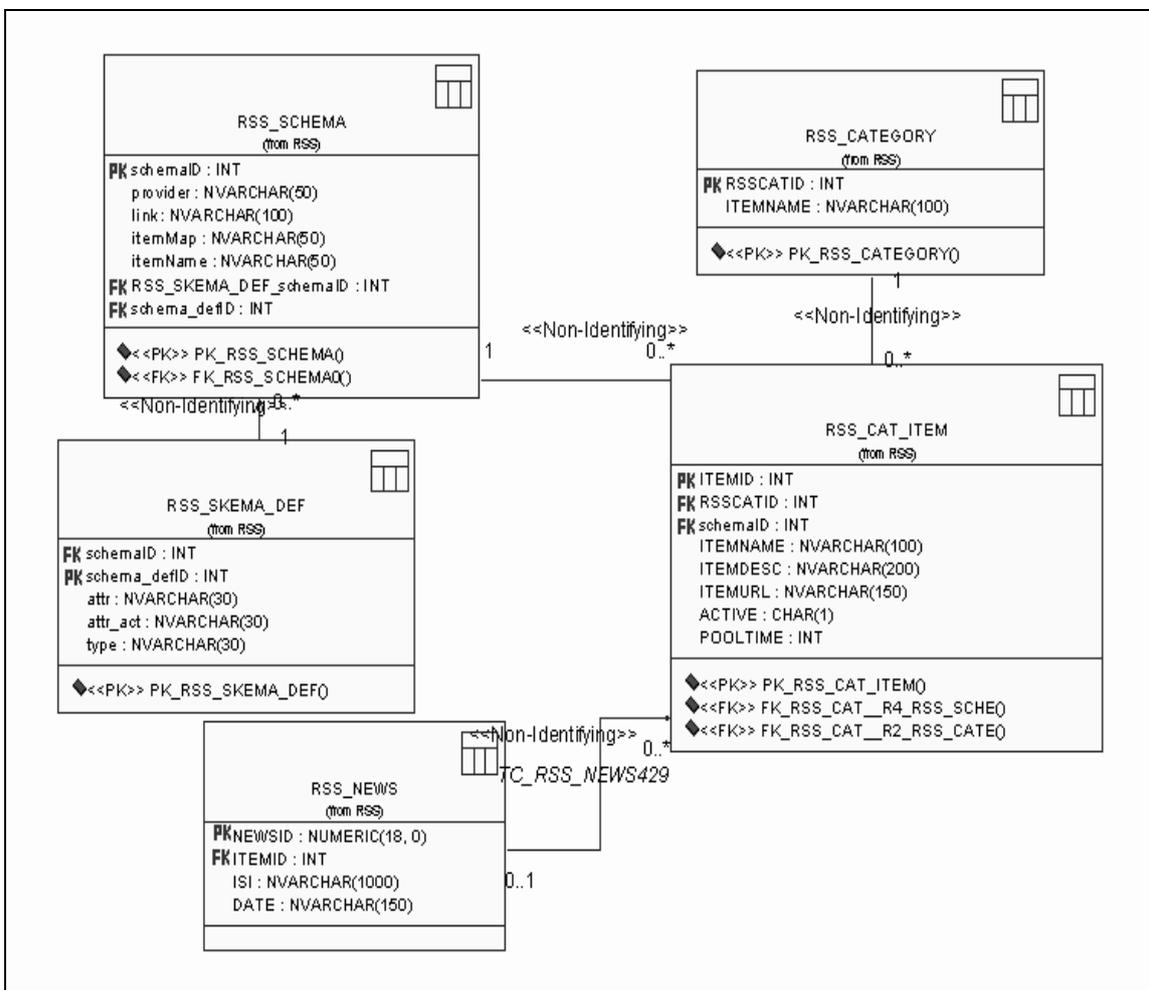
Gambar 9 Diagram Class Proses Pengambilan Berita

Kemudian dilakukan perancangan basis data dengan mengidentifikasi semua tabel yang

diperlukan untuk menyimpan data-data yang diperlukan. Selain itu juga diidentifikasi semua relasi antar tabel.

Obyek-obyek pada data model diagram dapat dijelaskan sebagai berikut:

- RSS_SCHEMA untuk masing-masing provider berita RSS yang didapatkan dengan melaksanakan skema dari contoh RSS milik situs penyedia berita.
- RSS_SKEMA_DEF menyimpan data definisi dari skema.
- RSS_CATEGORY menyimpan data tentang kategori yang disiapkan untuk mengelompokkan berita yang akan dikumpulkan.
- RSS_CAT_ITEM menyimpan data tentang lokasi situs pada masing-masing kategori.
- RSS_NEWS menyimpan *content* RSS yang telah di *download*.



Gambar 10 Diagram Class Proses Pengambilan Berita

Secara garis besar antar muka untuk aplikasi server mempunyai lima fasilitas utama :

1. Fasilitas untuk pengaturan skema dengan fungsi membuat skema provider RSS.
2. Fasilitas untuk pengaturan atribut elemen RSS yang akan diambil.
3. Fasilitas pengaturan kategori berita yang mempunyai fungsi untuk menambah, mengubah, dan menghapus kategori. Fasilitas ini juga berfungsi untuk menentukan alamat situs dan skema untuk masing-masing kategori.
4. Fasilitas melakukan pengawasan terhadap berita-berita yang pernah di *download*.
5. Fasilitas untuk mengaktifkan atau menonaktifkan service pengambilan berita dari *provider* berita.

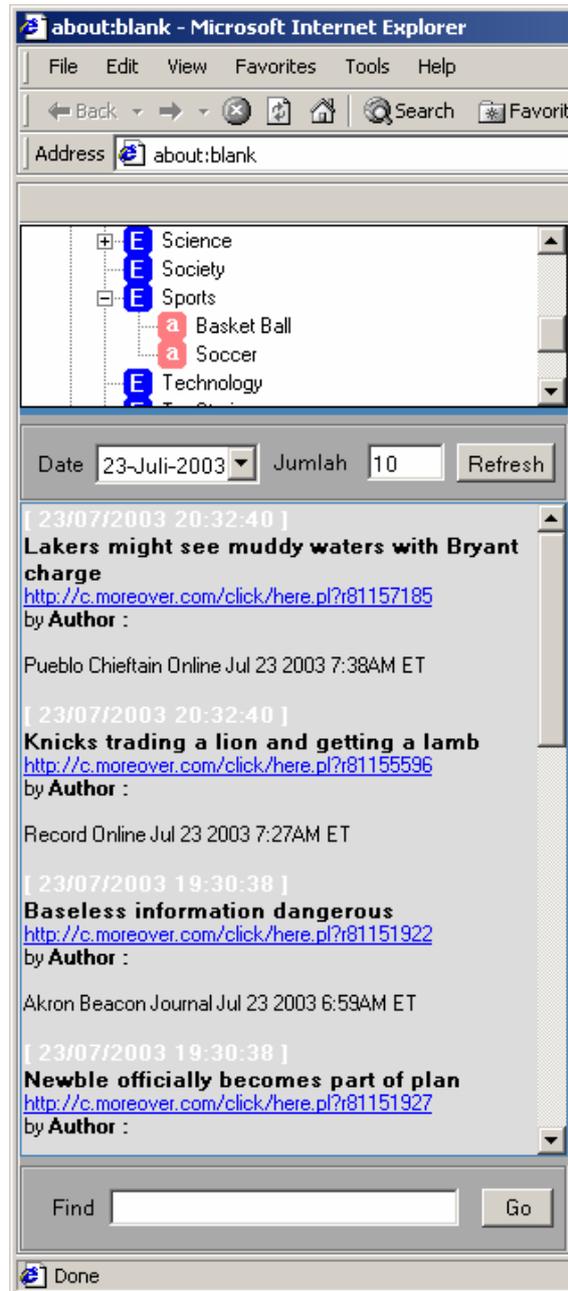
Sedangkan aplikasi pengguna secara garis besar akan dirancang agar mempunyai fasilitas-fasilitas sebagai berikut:

1. Fasilitas untuk menampilkan kategori-kategori berita yang ada di server RSS.
2. Fasilitas komponen tombol untuk melakukan *refresh* berita dari server RSS.
3. Fasilitas filter inputan seperti tanggal dan jumlah berita yang ditampilkan.
4. Menampilkan daftar berita sesuai filter tanggal, kategori, dan jumlah tertentu.
5. Fasilitas pencarian apabila pengguna ingin mencari suatu berita dengan *keyword* yang dimasukkan.

BAND OBJECT

Proses pembuatan komponen *content syndication* menggunakan *class Band Object* berisi *interface* untuk memanggil dan memanfaatkan *COM DeskBand Object* yang dimiliki *Microsoft Windows*. Ada beberapa tipe *Desk Band* yaitu *Communication Band* adalah komponen yang berada di bagian bawah dari *instance IE4+* (Internet Explorer 4+). Sedangkan *Information Band* berada di sebelah kiri dari *instance IE4+*, seperti *bar search* untuk mencari suatu file atau folder di dalam suatu directory. Serta *Toolbar Band* yang berada di bagian *toolbar* dari Internet Explorer. Untuk melihatnya IE4+ dijalankan kemudian klik "View" menu item dan pilih "Explorer Bar" menu item. Separuh atas separator adalah *Information Band*, sedangkan bagian bawah menu item adalah *Communication Band*.

Desk Band mengimplementasikan beberapa *interface COM* seperti *IObjectWithSite* dan *IDeskBand*. Deklarasi *interface* tersebut tersedia dalam bentuk *class C++* sehingga perlu di ubah dulu ke dalam tipe Microsoft Visual Studio .Net. Untuk berinteraksi dengan pengguna dibutuhkan *usercontrol* yang mengimplementasikan *class Band Object* untuk menampilkan berita.



Gambar 11 Aplikasi Komponen RSS Client Bar pada Internet Explorer

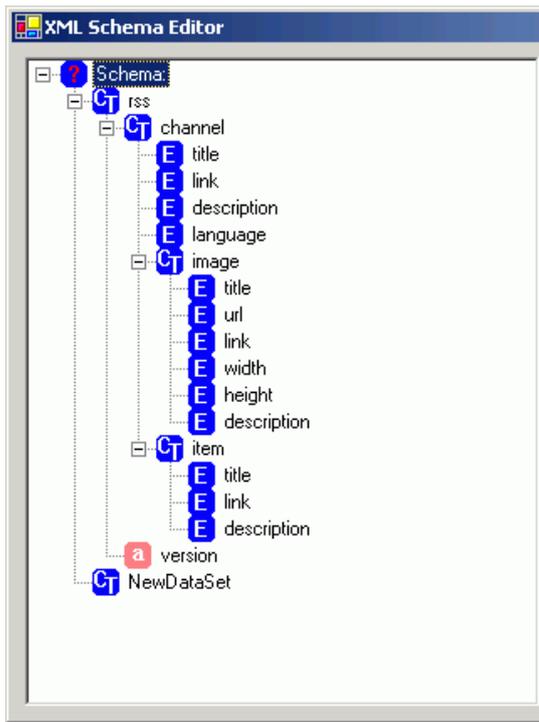
4. UJI COBA DAN EVALUASI

Uji coba dilakukan dalam jaringan lokal (intranet) menggunakan 2 buah komputer. Komputer pertama bertindak sebagai server *windows service* dan aplikasi server untuk RSS konfigurasi, sedangkan komputer kedua bertindak sebagai server basisdata. Untuk mencoba aplikasi pengguna tetap dilakukan pada komputer pertama.

Komputer yang digunakan sebagai server *windows service* memiliki spesifikasi *hardware* yang digunakan pada komputer ini adalah *Pentium III 450 Mhz* dengan memori fisik sebesar 352 Mb. Sistem operasi yang ter-install pada komputer ini adalah Microsoft Windows 2000 Professional.

Komputer yang digunakan sebagai server basisdata dengan spesifikasi hardware adalah Pentium III 450 MHz dengan memori fisik sebesar 352 mb.

Pada uji coba pengaturan skema untuk suatu provider RSS tertentu, dipilih RSS provider MoreOver.com dan provider UserLand.com, dimana situs MoreOver.com menggunakan RSS versi 0.9 sedangkan UserLand.com menggunakan RSS versi 2.0. Untuk melakukan pengaturan skema dilakukan pada form *Schema Editor* yang ada di aplikasi RSS server konfigurasi.



Gambar 12 Contoh MoreOver.xsd

Dapat dilihat pada gambar 12, skema contoh moreover.xsd. Lalu pada skema ini akan dilakukan

aktifitas *drag-drop* untuk memilih item-item mana dari skema yang akan disimpan kedalam database.

Dilakukan perbandingan suatu isi berita yang didapat dari program dengan berita dari situs lain yang tidak menerapkan *content syndication*. Dengan cara mencari suatu *keyword* di aplikasi pengguna.

1. Mencari 3 *keyword* di dua sub kategori pada aplikasi pengguna.
2. Mencari 3 *keyword* yang sama di situs portal berita.
3. Jangkauan pencarian adalah 4 hari yaitu dari tanggal 20 Juli 2003 sampai dengan 23 Juli 2003, karena aplikasi server RSS baru mulai diujicoba sejak tanggal 20 Juli 2003.

Skenario uji coba dilakukan dengan langkah-langkah sebagai berikut :

1. Memilih beberapa kata umum yang mungkin menjadi isi berita, dengan perbandingan beberapa kategori, yaitu kategori Regional sub kategori Indonesia dan kategori Sports sub kategori Soccer
2. Setiap hasil pencarian di aplikasi pengguna, akan dibandingkan dengan hasil pencarian di situs-situs berita yang tidak menerapkan *content syndication* (misalkan detik.com dan kompas.com)
3. Situs-situs yang dijadikan pembanding adalah situs-situs yang memiliki rating tertinggi di beberapa pencarian.
4. Untuk sub kategori Indonesia, *keyword* yang diujicobakan adalah Megawati, Aceh, dan Bali. Untuk kategori Soccer, *keyword* yang diujicobakan adalah Beckham, Ronaldinho dan Manchester United.

Tabel 1 Hasil Perbandingan Pencarian Berita

Keyword \ Source	Jumlah Item		
	aplikasi RSS	kompas.com	yahoo.com
Megawati	4	3	1
Aceh	24	3	1
Bali	40	0	7
Beckham	12	1	2
Ronaldinho	17	12	5
Manchester United	21	1	13

Dari data dapat dilihat, dalam mencari berita, hasil *content syndication* lebih baik dari situs yang tidak menerapkan *content syndication* seperti kompas.com ataupun yahoo.com.

Dilanjutkan setelah penambahan kategori, maka pengambilan berita pada *background process* hasilnya dapat dilihat sewaktu menjalankan *windows service*.

```

Start Logging : 21/07/2003 16:10:14
Berhasil Download Berita Baru : - 5 - 21/07/2003 16:12:49
Berhasil Download Berita Baru : - 5 - 21/07/2003 16:12:49
Berhasil Download Berita Baru : - 12 - 21/07/2003 16:14:16
Berhasil Download Berita Baru : - 3 - 21/07/2003 16:15:28
Berhasil Download Berita Baru : - 6 - 21/07/2003 16:16:22
Berhasil Download Berita Baru : - 6 - 21/07/2003 16:16:22
Berhasil Download Berita Baru : - 6 - 21/07/2003 16:16:22
Berhasil Download Berita Baru : - 12 - 21/07/2003 16:44:24
Berhasil Download Berita Baru : - 12 - 21/07/2003 16:44:24
Berhasil Download Berita Baru : - 12 - 21/07/2003 16:44:24
Berhasil Download Berita Baru : - 12 - 21/07/2003 16:44:24
Berhasil Download Berita Baru : - 12 - 21/07/2003 16:44:24
Berhasil Download Berita Baru : - 12 - 21/07/2003 16:44:24
Berhasil Download Berita Baru : - 12 - 21/07/2003 16:44:24
Berhasil Download Berita Baru : - 13 - 21/07/2003 17:04:29
Berhasil Download Berita Baru : - 13 - 21/07/2003 17:04:29
Berhasil Download Berita Baru : - 12 - 21/07/2003 17:09:05
Berhasil Download Berita Baru : - 10 - 21/07/2003 17:12:07
Berhasil Download Berita Baru : - 1 - 21/07/2003 17:14:42
Berhasil Download Berita Baru : - 5 - 21/07/2003 17:20:07
Berhasil Download Berita Baru : - 5 - 21/07/2003 17:20:07
Berhasil Download Berita Baru : - 5 - 21/07/2003 17:20:07
Stop Logging : 21/07/2003 17:24:26
    
```

Gambar 13 Contoh File Log pada Proses *Download*

5. KESIMPULAN

- Proses *download* berita dari situs yang menyediakan RSS feeds secara otomatis dibuat dengan *windows service* yang dapat berjalan secara *background process* di sistem operasi. Untuk men-*download* dan mengolah RSS secara otomatis dibutuhkan konfigurasi seperti skema RSS, daftar atribut elemen RSS, url *content provider*, dan lain-lain. Setiap konfigurasi disimpan di basis data. Aplikasi *windows form* digunakan sebagai *tool* untuk mengatur konfigurasi tersebut dan menyimpannya di basis data. Hasil dari aplikasi *windows service* dapat dilihat ke basis data, berupa *content* yang sudah terkategori dan berisi elemen-elemen RSS seperti *title*, *link*, *description*, dan lain-lain.
- Merancang dan membuat suatu skenario melalui *content syndication* dilakukan dengan cara menyimpan dan mengolah skema XSD dari format RSS situs *provider* tersebut. Setelah itu ditentukan kategori dan disimpan url-nya. Selanjutnya dilakukan pemilihan atribut-atribut elemen dari setiap skema, untuk menentukan format yang akan disimpan.
- Dibuat aplikasi client untuk membaca berita yang telah diolah aplikasi server. Untuk membuat komponen plug-in pada aplikasi Internet Explorer, diperlukan suatu *class interface* untuk dapat berinteraksi dengan *COM Deskband Object*. Lalu untuk tampilan

komponen dibutuhkan suatu *user control* yang akan menampilkan *content* berita.

- Dari hasil uji coba pencarian berita, hasil pencarian berita di RSS Feeds dengan suatu *keyword* mendapatkan jumlah lebih banyak dibanding dengan pencarian berita melalui portal berita biasa.

6. DAFTAR PUSTAKA

1. Hammersley, Ben, *Content Syndication with RSS*, O'Reilly, <http://www.oreilly.com/catalog/consynrss/chapter/ch04.pdf>, March 2003.
2. Belcalan, *The Journal : A brief introduction to RSS*, <http://glenn.bluemountains.net.au/mt/archives/00042.php>, May 2003
3. *Microsoft Service Developer Network Library*, Microsoft Corporation, 2003
4. Belcalan, *The Journal : A brief introduction to RSS*, <http://glenn.bluemountains.net.au/mt/archives/00042.php>, May 2003.
5. Downes, Stephen, *Content Syndication and Online Learning*, Edmonton, Alberta <http://www.downes.ca/files/rss.doc>, September 22, 2000.