

PERCEPATAN PERHITUNGAN HASIL TES SISTEM ONLINE MENGUNAKAN KOMPRESI DATABASE

Surya Sujarwo

Computer Science Department, School of Computer Science, Binus University
Jl. K.H. Syahdan No. 9, Palmerah, Jakarta Barat 11480
surya.ss@binus.edu

ABSTRACT

The article explores the influence of database compression on processing time in generating online test results through experimental study. The study uses three level compressions consists of none, row and page in table used by online test system. The experiment shows that using page compression results in faster processing time and less space compared to none and row compression, and the optimal row for compression is between 1001 and 2000 rows. In summary, to reduce space usage and processing time, it is better to use page compression in user answer table of online test system and to group the participant in schedule so that users' answer will be between 1001 and 2000 rows.

Keywords: database compression, processing time, online test result

ABSTRAK

Artikel ini membahas pengaruh penerapan kompresi database terhadap waktu pemrosesan dalam menghitung dan menghasilkan hasil tes online melalui metode eksperimen. Penelitian ini menggunakan tiga level kompresi yaitu kompresi none, row dan page pada tabel yang digunakan oleh sistem tes online. Hasil penelitian menunjukkan bahwa menggunakan kompresi page menghasilkan waktu proses perhitungan yang lebih cepat dan tempat yang digunakan lebih sedikit dibandingkan dengan kompresi none dan row, dan jumlah baris yang optimal untuk kompresi yaitu antara 1001 sampai 2000 baris. Kesimpulannya, untuk mengurangi penggunaan tempat dan mempercepat waktu perhitungan lebih baik menggunakan kompresi page pada tabel jawaban yang digunakan pada sistem tes online dan mengelompokkan peserta tes pada jadwal sedemikian rupa sehingga jawaban peserta berada di antara 1001 sampai 2000 baris.

Kata kunci: kompresi database, waktu pemrosesan, hasil tes online

PENDAHULUAN

Pada aplikasi tes online diperlukan tempat untuk penyimpanan data berupa jadwal tes, kelompok soal tes, soal tes, jawaban soal tes, peserta tes, dan jawaban peserta. Tempat untuk penyimpanan data bisa berupa database atau *file*. Untuk pengaksesan data yang cepat dalam membaca dan menulis data tes kita dapat menggunakan database server yang menyimpan data secara terstruktur menggunakan struktur data B-Tree yang dapat mengakses data melalui *key* yang diberikan secara cepat (McGehee, 2008).

Selain memerlukan penyimpanan data, aplikasi tes online memerlukan aplikasi yang bisa diakses melalui komputer-komputer pada waktu bersamaan, yang bisa menggunakan arsitektur client-server atau menggunakan web. Pada umumnya sistem tes online menggunakan aplikasi web yang bisa diakses melalui aplikasi browser pada komputer yang terhubung melalui jaringan internet atau intranet. Untuk membuat aplikasi web yang cepat kita bisa menggunakan pemrograman web yang tersedia secara luas saat ini seperti asp.net (Shepherd, 2010). Pada asp.net tersedia framework untuk mengakses database secara mudah menggunakan Entity Framework (Lerman, 2010) yang menyediakan semua pengaksesan database pada code pemrograman.

Proses yang ada dalam aplikasi tes online salah satunya adalah menghitung hasil tes yang didapat dari menghitung jumlah jawaban peserta yang sesuai dengan jawaban soal tes yang ada sesuai dengan jadwal tes yang ada. Dalam proses perhitungan ini yang menjadi proses yang paling lama adalah pengambilan data dari database ke dalam aplikasi. Untuk meningkatkan waktu pengambilan data kita bisa menggunakan kompresi (McGehee, 2008) pada penyimpanan data yang dilakukan oleh database server. Pengaruh kompresi ini diterapkan pada index yang ada pada tabel dengan menggunakan teknik untuk mempercepat pengaksesan data melalui index (Bhuiyan & Hoque, 2009).

Penelitian ini menggunakan metode *experimental* (Freitas, 2009) (Dodig-Crnkovic, 2002) untuk mengukur seberapa pengaruh kompresi dalam mempercepat proses perhitungan hasil tes online.

METODE

Penelitian ini menggunakan ASP.NET 4 sebagai aplikasi server dan Entity Framework 4 dalam pengaksesan database, dan SQL Server 2008 sebagai database server. Metode penelitian yang digunakan adalah metode *experimental*. Data yang digunakan dalam *experimental* adalah data operasional ujian online dari tanggal 13 November 2010 sampai 18 Mei 2013 sebanyak 5.941.236 baris data yang menyimpan jawaban mahasiswa. Jadwal tes yang akan dihitung hasilnya sebanyak 1412 jadwal. Untuk setiap jadwal yang ada akan dihitung waktu prosesnya untuk database yang menggunakan tabel yang tidak terkompresi, tabel yang terkompresi menggunakan *row* dan tabel yang terkompresi menggunakan *page*.

Tabel yang digunakan untuk menyimpan data jawaban mahasiswa memiliki struktur pada Tabel 1. Index yang digunakan pada tabel antara lain kolom "Id" menggunakan *clustered index ascending* dan kolom "Note" menggunakan *non-unique, non-clustered index ascending*.

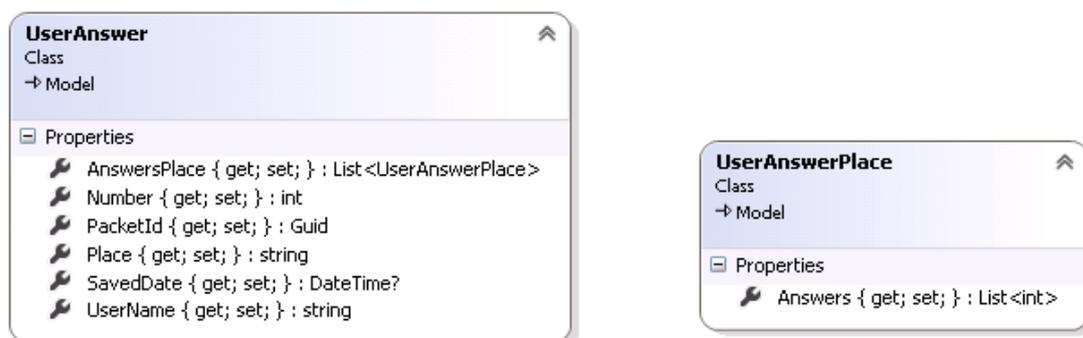
Tabel 1 Struktur Tabel Penyimpanan Jawaban Mahasiswa

Nama Kolom	Tipe data	Ukuran
Id	uniqueidentifier	
Type	varchar	150
Name	nvarchar	256

Note	nvarchar	256
SaveDate	datetime	
UserId	uniqueidentifier	
Object	varbinary	MAX

Jawaban mahasiswa disimpan pada kolom *Object* di tabel jawaban menggunakan format *xml* dari class *UserAnswer* dan *UserAnswerPlace* yang diubah dalam bentuk *byte*. *Properties* pada class *UserAnswer* dan *UserAnswerPlace* dapat dilihat pada Gambar 1. Contoh bentuk *xml* dari class *UserAnswer* dan *UserAnswerPlace* sebagai berikut:

```
<UserAnswer xmlns="http://schemas.datacontract.org/2004/07/Business.Model"
xmlns:i="http://www.w3.org/2001/XMLSchema-instance">
  <AnswersPlace>
    <UserAnswerPlace>
      <Answers xmlns:a="http://schemas.microsoft.com/2003/10/Serialization/Arrays">
        <a:int>1</a:int>
      </Answers>
    </UserAnswerPlace>
  </AnswersPlace>
  <Number>25</Number>
  <PacketId>8d5fcddd-420a-42bf-b1cf-e49902947286</PacketId>
  <Place i:nil="true"/>
  <SavedDate>2010-11-15T09:44:59.7607311+07:00</SavedDate>
  <UserName>1401084912</UserName>
</UserAnswer>
```



Gambar 1 Struktur class *UserAnswer* dan *UserAnswerPlace*

Besar data jawaban mahasiswa yang disimpan didapat dari jumlah *byte* yang diperoleh dari perubahan *xml* ke dalam bentuk *byte*. Jumlah *byte* yang digunakan pada jawaban mahasiswa dapat dilihat pada Tabel 2.

Tabel 2 Jumlah *Byte* yang Digunakan pada Jawaban Mahasiswa

Jumlah byte	Jumlah	Jumlah byte	Jumlah	Jumlah byte	Jumlah
436	7	487	184	536	6
437	85	488	1792	537	141
438	1484	489	1384	538	346
439	3128	490	6087	539	105

464	8	491	35836	540	2264
465	4	492	69202	541	14907
466	46	493	17	542	26976
467	455	494	13949	543	1
468	4539	495	44690	544	5
469	11185	496	521	545	60
477	1	497	6015	546	623
478	1	498	79590	547	6250
479	34	499	976871	548	80539
480	112	500	2449832	549	669687
482	1	508	234	550	1428771
483	1	509	656		
484	21	510	630		
485	28	511	1922		
486	2	535	1		

HASIL DAN PEMBAHASAN

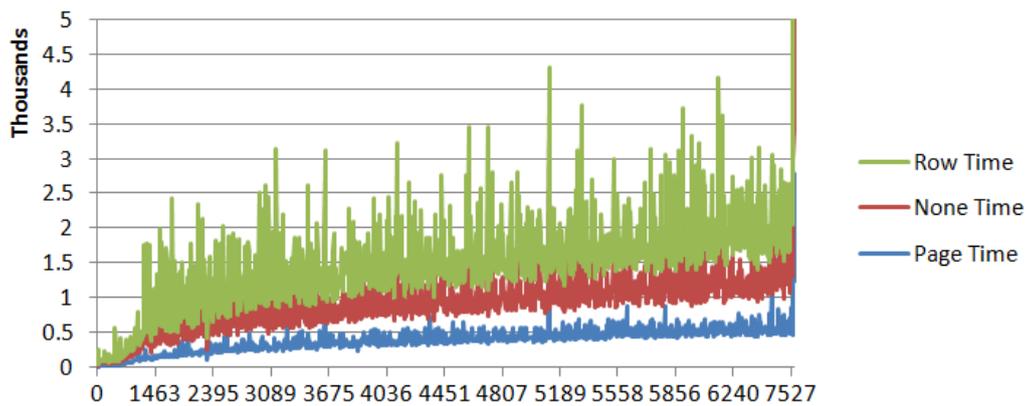
Besar *space* yang digunakan pada tabel jawaban mahasiswa menggunakan tiga macam kompresi dapat dilihat pada Tabel 3, dengan keterangan setiap ukuran dihitung dalam ukuran kilobyte (KB), *reserved* adalah ukuran *file* yang digunakan untuk menyimpan data pada tabel, data adalah ukuran untuk menyimpan data tanpa index, *indexes* adalah ukuran untuk menyimpan index dan *unused* adalah ukuran yang tersedia pada *file* yang tidak digunakan.

Tabel 3 Ukuran *Space* yang Digunakan Sesuai Jenis Kompresi

Tipe kompresi	Ukuran Reserved(KB)	Ukuran Data(KB)	Ukuran Indexed(KB)	Ukuran Unused(KB)
<i>None</i>	6.268.392	5.528.824	738.008	1.560
<i>Row</i>	6.253.264	5.528.896	722.800	1.568
<i>Page</i>	3.528.288	2.815.784	712.120	384

Tabel 3 menunjukkan *space* yang terpakai jika menggunakan kompresi *page* adalah sebesar 56.29% dari *space* yang terpakai jika tidak menggunakan kompresi (*none*), dan *space* yang terpakai jika menggunakan kompresi *row* adalah sebesar 99.76% dari *space* yang terpakai jika tidak menggunakan kompresi. Hal ini terjadi karena menggunakan kompresi *page* data dikompresi dengan menggabungkan baris-baris yang ada menjadi satu sedangkan yang kompresi *row* data dikompresi setiap barisnya. Data pada baris ukurannya kecil sehingga jika dikompresi tidak akan terlalu berkurang bahkan bisa lebih besar, sedangkan jika dikompresi dengan menggabungkan data pada baris-baris yang ada maka ukurannya akan berkurang banyak.

Perbandingan lama proses perhitungan hasil tes menggunakan kompresi yang ada bisa dilihat pada Gambar 2. Ukuran lama waktu (vertikal) menggunakan satuan *millisecond* (ms). Ukuran jumlah (horizontal) menggunakan satuan angka.



Gambar 2 Grafik perbandingan lama proses perhitungan hasil tes menggunakan kompresi *row*, *none* dan *page*

Gambar 2 menunjukkan kecenderungan semakin banyak jawaban yang diproses semakin lama waktu proses perhitungan datanya. Waktu pemhitungan yang menggunakan kompresi *row* lebih lama dibandingkan dengan tidak menggunakan kompresi, hal ini karena ukuran pada *file* yang menggunakan kompresi *row* dengan yang tidak menggunakan kompresi tidak terlalu berbeda dan untuk yang menggunakan kompresi setiap baca data selalu melakukan proses dekompresi sehingga memerlukan waktu yang lama dalam pembacaan data. Sedangkan waktu yang menggunakan kompresi *page* cenderung lebih cepat dibandingkan dengan yang tidak menggunakan kompresi, hal ini karena ukuran *filenya* berbeda sampai 43% walaupun saat membaca data pada kompresi *row* perlu melakukan proses dekompresi, namun jumlahnya jauh lebih sedikit dibanding yang menggunakan kompresi *row* dan ukuran *file* yang perlu dibaca jauh lebih sedikit dibandingkan dengan yang tidak menggunakan kompresi.

Untuk melihat persentase pengaruh waktu proses perhitungan menggunakan kompresi, data waktu di atas dikelompokkan dalam interval dan waktu prosesnya dirata-rata, kemudian di hitung persentase waktunya kemudian dirata-ratakan untuk mendapatkan rata-rata perbedaannya. Hasil perhitungan dapat dilihat pada Tabel 4.

Tabel 4 Rata-rata Perbandingan Waktu Proses Perhitungan

Interval	Page	None	Row	Page / None	None / row
0-250	29	36	36	79.18%	99.90%
251-500	56	75	101	74.70%	74.19%
501-1000	81	135	120	59.88%	111.98%
1001-2000	165	378	291	43.58%	129.68%
2001-3000	258	472	440	54.53%	107.36%
3001-5000	390	536	544	72.69%	98.51%
5001-10000	518	664	695	77.90%	95.54%
10001-100000	1,759	2,615	2,894	67.27%	90.34%
Rata-rata perbandingan waktu				66.22%	100.94%

Tabel 4 menunjukkan dengan menggunakan kompresi *page*, jika dibandingkan dengan tidak menggunakan kompresi waktu proses akan lebih cepat sekitar 66% dan menggunakan kompresi *row* tidak terlalu berbeda waktunya dibandingkan dengan tidak menggunakan kompresi.

PENUTUP

Dari hasil analisis dan pembahasan, dapat ditarik kesimpulan kompresi *page* mempunyai pengaruh yang signifikan terhadap waktu perhitungan hasil tes. Dengan data yang bagus untuk menggunakan kompresi *page* berada di antara 1001 sampai 2000 baris data.

DAFTAR PUSTAKA

- Bhuiyan, M. M., & Hoque, A. S. (2009). High performance SQL queries on compressed relational database. *Journal of Computers*, 4, 1263-1274.
- Dodig-Crnkovic, G. (2002). Scientific methods in computer science. *Conference for the Promotion of Research in IT at New Universities and at University Colleges in Sweden* (pp. 6-7). Sweden: Mälardalen University.
- Freitas, R. (2009). *Scientific Research Methods and Computer Science*. Diakses dari <http://www.map.edu.pt/mapi/2008/map-i-research-methods-workshop-2009>
- Lerman, J. (2010). *Programming Entity Framework*. Sebastopol: O'Reilly Media.
- McGehee, B. (2008). *Brad's Sure Guide to SQL Server 2008*. Cambridge: Red Gate Books.
- Shepherd, G. (2010). *Microsoft ASP.NET 4 Step by Step*. Redmond: Microsoft Press.