A Method for Data Integration in Enterprises Using Web Service

Alieh AliAbadi
MSc Student; Department of Computer and Information Technology;
Eshragh Institute of Higher Education;
aliyeh_aliabadi@yahoo.com

Mohsen Mohammadi
PhD in Information Technology; Assistant Professor;
Computer Department; Esfarayen University of Ttehcnology;
Corresponding Author mohsen@esfarayen.ac.ir

Iranian Journal of Information Processing and Management

Received: 15. Apr. 2017 **Accepted: 22, Jul. 2017**

Iranian Research Institute for Information Science and Technology (IranDoc)
ISSN 2251-8223
eISSN 2251-8231
Indexed by SCOPUS, ISC, & LISTA
Vol. 33 | No. 4 | pp. 1637-1658
Summer 2018

Abstract: Increasing the speed and reducing the use of resources in data integration process have always been the goal of developers and researchers in the process of data integration. The purpose of this study is to provide a solution using metadata as well as web browsing to speed up the process, so as to improve resources such as memory. The proposed solution is implemented using the three-layer architecture approach, which includes business logic layers, software layer and data access layer. After implementing the proposed strategy, it was tested on 5000 database records in a case study (Shahsavand Tea Company). The solution presented in addition to a comparison with several similar cases has also been surveyed by experts. The results show that the proposed solution has been able to increase the data transfer speed and improve the use of memory resources in the given data volume. Also, according to the answers given to the questionnaire by experts, it was found that user-friendly software design has been able to facilitate the use of the tool for users.

Keywords: Web Service, Data Integration, Database, ETL Process, Meta data

ارائهٔ راهکاری جهت تجمیع داده‌ها در سازمان‌ها با استفاده از وب‌سرویس

عالیه علی‌آبادی
دانشجوی کارشناسی ارشد؛ مهندسی کامپیوتر؛ مؤسسهٔ آموزش عالی غیرانتفاعی اشراق؛
aliyeh_aliabadi@yahoo.com

محسن محمدی
دکتری فناوری اطلاعات؛ استادیار؛ گروه کامپیوتر مجتمع آموزش عالی فنی-مهندسی اسفراین؛ پدیدآور رابط mohsen@esfarayen.ac.ir

پژوهش‌نامهٔ پردازش و مدیریت اطلاعات

دریافت: ۱۳۹۶/۰۱/۲۶ **پذیرش:** ۱۳۹۶/۰۴/۳۱ مقاله برای اصلاح به مدت ۱ روز نزد پدیدآوران بوده است.

فصلنامه | علمی پژوهشی
پژوهشگاه علوم و فناوری اطلاعات ایران (ایرانداک)
شاپا (چاپی) ۸۲۲۳-۲۲۵۱
شاپا (الکترونیکی) ۸۲۳۱-۲۲۵۱
نمایه در LISTA، ISC، و SCOPUS
jipm.irandoc.ac.ir
دورهٔ ۳۳ | شمارهٔ ٤ | ص‌ص ۱۶۳۷-۱۶۵۸
تابستان ۱۳۹۷

چکیده: بهبود سرعت و کاهش استفاده از منابع در فرایند تجمیع داده‌ها همواره از اهداف توسعه‌گران و محققان در زمینهٔ فرایند تجمیع داده‌ها بوده است. هدف از این پژوهش، ارائهٔ راهکاری با استفاده از متادیتا و همچنین، وب‌سرویس، جهت افزایش سرعت فرایند است، به‌طوری که بهبود منابعی مانند حافظه را نیز به همراه داشته باشد. راهکار پیشنهادی با استفاده از روش معماری سه‌لایه پیاده‌سازی شده است و شامل لایه‌های منطق کسب‌وکار، لایهٔ نرم‌افزار، و لایهٔ دستیابی به داده‌هاست. پس از پیاده‌سازی راهکار پیشنهادی، بر روی ۵۰۰۰ رکورد پایگاه داده در مطالعهٔ موردی (شرکت چای شاهسوند) تست شد. برای راهکار ارائه‌شده علاوه بر مقایسه با چندین مورد مشابه، از افراد خبره نیز نظرسنجی شده است. نتایج نشان می‌دهد که راهکار پیشنهادی توانسته در حجم داده‌هایی مشخص سرعت انتقال داده‌ها را افزایش و میزان استفاده از منابع حافظه را بهبود دهد. همچنین، طبق ارزیابی پاسخ‌های داده‌شده به پرسشنامه مشخص شد که طراحی کاربرپسندِ نرم‌افزار توانسته است سهولت استفاده از ابزار را برای کاربران در پی داشته باشد.

کلیدواژه‌ها: وب‌سرویس، تجمیع داده، پایگاه داده، فرایند ETL، متادیتا

۱. مقدمه

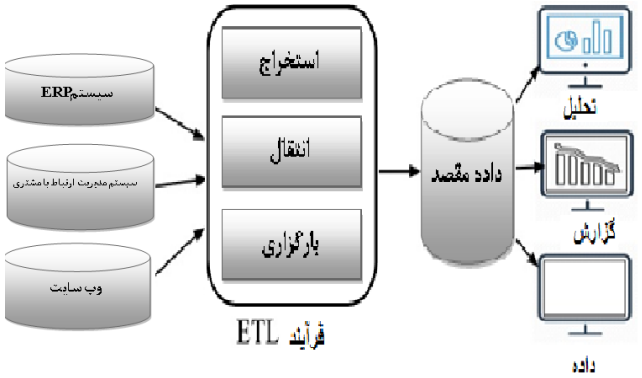
امروزه کلید ماندگاری در دنیای کسب‌وکار قدرت تحلیل، برنامه‌ریزی، عکس‌العمل مناسب و تغییر شرایط کسب‌وکار در اسرع وقت و با حداکثر سرعت ممکن است. بنابراین، وجود یک پایگاه داده که داده‌ها را از منابع اطلاعاتی مختلف سازمان جمع‌آوری می‌کند، ضروری به نظر می‌رسد. در واقع، این پایگاه داده مخزن اصلی کلیهٔ داده‌های حال و گذشتهٔ یک سازمان است که همیشه جهت انجام عملیات گزارش‌گیری و آنالیز در دسترس مدیران است .از این طریق مدیران می‌توانند به تمامی داده‌های سازمان از گذشته تاکنون جهت تجزیه و تحلیل و تهیهٔ گزارش‌های مختلف دسترسی داشته باشند (Kadir 2013).

زمانی‌که سازمان‌ها به اهمیت داده‌ها در سیستم‌های اطلاعاتی کسب‌وکار خود پی بردند، اهمیت تجمیع داده‌ها بیشتر احساس شد. آن‌ها متوجه شدند اطلاعاتی که از این داده‌ها به‌دست می‌آید، می‌تواند یک مزیت رقابتی مهم محسوب شود که به بقای سازمان در محیط رقابتی کسب‌وکار کمک کند. دلایل مختلفی را می‌توان برشمرد که اهمیت تجمیع داده‌ها را روشن‌تر می‌سازد. دو مورد از مهم‌ترین این دلایل را می‌توان این‌گونه بیان کرد که اولاً با توجه به مجموعه‌ای از سیستم‌های اطلاعاتی، وجود منابع یکپارچه از داده‌های این سیستم‌ها می‌تواند دسترسی راحت‌تر به داده‌ها را از طریق تنها یک نقطهٔ دسترسی فراهم آورد، دوم این‌که با توجه به داده‌های خاصی که هر بار مورد نیاز است، این داده‌ها از سیستم‌های اطلاعاتی مختلف، اما مکمل هم ترکیب می‌شوند. بدین‌صورت که ابتدا استخراج داده‌ها از منابع مختلف سیستم‌های اطلاعاتی صورت می‌گیرد و سپس، داده‌ها با هم ترکیب می‌شوند. نمایش داده‌ها به شیوه‌ای است که همانند یک واحد کل به نظر می‌رسند. به این ترتیب، اطلاعات به‌صورت جامع‌تر در اختیار کاربر قرار می‌گیرد که به جهت تصمیم‌گیری‌های استراتژیک مناسب‌تر هستند (Salem, Boussaïd and Darmont 2013).

انجام عکس‌العمل مناسب و سریع مستلزم این است که داده‌ها با حداکثر سرعت ممکن در اختیار قرار گیرند. ارائهٔ یک راهکار بهینه جهت تجمیع داده‌ها به‌شدت وابسته به انتخاب فرایندی است که تحت عنوان استخراج، تبدیل و بارگذاری[1] (ETL) شناخته

1. Extract, Transform, Load (ETL)

می‌شود. همان‌طور که در شکل ۱، نشان داده شده، این فرایند شامل استخراج داده‌ها از مبدا، تبدیل آن‌ها به نوع مناسب، و بارگذاری آن‌ها در مقصد نهایی است. فرایندهای ETL از نظر سرعت و عملکرد یکسان نیستند. امروزه، اهمیت بهبود این فرایندها در تصمیم‌گیری‌های کسب‌وکار رو به افزایش است. می‌توان گفت که ETL مؤثرتر، منجر به تصمیم‌گیری مؤثرتر و بازدهی بهتر می‌شود (Ali 2014). در این میان، سرعت همواره یکی از معیارهای مهم در تجمیع داده‌ها بوده و همواره بهبود سرعت در تجمیع داده‌ها یکی از چالش‌های مهم توسعه‌گران بوده است (Leigh and Lau 2016; Agrawal 2008; Greene 2016).



شکل ۱. فرایند تجمیع داده‌ها (Alandi 2016)

هدف از این مقاله ارائهٔ راهکاری جهت تجمیع داده است. با توجه به اهمیت سرعت و همچنین، سعی در کاهش هزینه‌ها، سعی بر آن شد تا راهکار ارائه‌شده بتواند با استفاده از وب‌سرویس جهت انجام عملیات ETL و همچنین، استفاده از متادیتا جهت تشخیص داده‌هایی که باید منتقل شوند، سرعت را افزایش داده و به زمان واقعی نزدیک‌تر کرد.

۲. پیشینهٔ تحقیق

تجمیع داده‌ها فرایندی نسبتاً پیچیده است. هیچ روش فراگیری برای تجمیع داده‌ها وجود ندارد و بسیاری از تکنیک‌های مورد استفادهٔ متخصصان حوزهٔ فناوری اطلاعات هنوز در مرحلهٔ رشد و تکامل است. بعضی از روش‌های تجمیع داده‌ها برای یک سازمان، بهتر از سایر روش‌ها عمل می‌کنند که این موضوع کاملاً به نیازهای آن سازمان بستگی

خواهد داشت (Madhavan 2007). به‌عنوان اولین دستاوردها، تحقیقات در زمینهٔ تجمیع داده‌ها در سیستم‌های چند پایگاه داده[1] در سال ۱۹۸۰ انجام شد (Florea, Diaconita and Bologa 2016). این اولین پایه و اساس تحقیقات بر روی تجمیع داده‌ها بود. تحقیقات بر روی میانجی‌ها (De Faria Cordeiro, Campos and Borges 2015)، سیستم‌های همتابه‌همتا[2] (Abiteboul, Nguyen and Ruberg 2006) و تجمیع بر اساس فناوری وب‌سرویس‌ها (Oracle Data Integrator 2016) ادامه پیدا کرد.

«لندرز و روزنبرگ» کاهش استفاده از منابعی مانند حافظه را یکی از اهداف مهم در تجمیع داده‌ها می‌دانند و ترکیب دو یا چند عملیات و تبدیل آن‌ها به یک فرایند را جهت افزایش بهره‌وری و کاهش استفاده از حافظه پیشنهاد می‌کنند (Landers and Rosenberg 1986). بعضی از نویسندگان استفاده از یک رابط کاربری مناسب در ابزارهای ETL را مناسب می‌دانند (Majumdar, Saha and Wang 2016). نتایج تحقیقات آن‌ها نشان می‌دهد که یک رابط کاربری گرافیکی انتقال داده بین منابع را تسهیل می‌کند و این رابط با پیچیده‌تر شدن نیازهای سیستم مانند نیاز به شناسایی منبع و شناسایی تغییرات مفید خواهد بود. با توجه به آنچه بیان شد ارائهٔ راهکاری که زمان و میزان استفاده از منابعی مانند حافظه را بهبود و انجام فرایند تجمیع را برای کاربران تسهیل بخشد، مطلوب خواهد بود.

«زارترز» با توجه به اهمیت فرایند ETL در تجمیع داده‌ها، روشی را برای افزایش عملکرد ETL پیشنهاد داده است (Zuters 2011). روش ارائه‌شده در این پژوهش بر مبنای فشرده‌سازی داده‌ها در مبدأ و رفع فشرده‌سازی از روی داده‌ها در مقصد عمل می‌کند. در این روش داده‌ها با استفاده از یک نرم‌افزار فشرده می‌شوند و به این ترتیب، حجم داده‌هایی که باید منتقل شود، کاهش می‌یابد.

«پریت داندا و ویشال گورل» نیز روش‌های مختلف بهبود عملیات ETL در تجمیع داده‌ها را بررسی کرده‌اند. نتایج تحقیق آن‌ها نشان می‌دهد که بهبود در فرایند درخواست[3] به‌عنوان ستون فقرات فرایندهای ETL، می‌تواند این فرایندها را تا حد زیادی بهبود دهد. آن‌ها به این منظور استفاده از حافظهٔ میانی[4] در فرایند درخواست را به‌عنوان راهکاری جهت افزایش سرعت ETL مطرح می‌کنند (Preet Dhanda 2016).

ابزارهای ETL زیادی تاکنون ارائه شده است که هر یک دارای مزایا و معایبی است.

1. Multi Database 2. Peer to peer 3. query process 4. cache

این ابزارها با رفع نیاز به کدنویسی می‌توانند در وقت و زمان صرفه‌جویی نمایند. از مهم‌ترین این ابزارها می‌توان به IBM Datastage، Pentaho Kettle، Microsoft SSIS، Pervasive Data Integrator و Informatica Power center اشاره نمود.

در جدول ۱، ابزارهای مختلف ETL مقایسه شده‌اند. مقایسهٔ ابزارها در این بخش با توجه به سادگی استفاده از این روش‌ها، قابلیت استفادهٔ مجدد، نزدیکی به زمان واقعی، اعتبارسنجی داده‌ها، تجاری یا منبع باز بودن آن‌ها، رابط کاربری و پیچیدگی انتقال صورت می‌گیرد. سادگی استفاده جهت درک و فراگیری راحت‌تر یک ابزار، رابط کاربری گرافیکی[1] مناسب و کاربرپسند، و قابلیت استفادهٔ مجدد از سرویس‌ها سه معیار مهم در انتخاب یک ابزار ETL مناسب هستند (Adeptia 2015; Alandi 2016). همان‌طور که قبلاً اشاره شد، سرعت در فرایند ETL و نزدیک شدن هرچه بیشتر به زمان واقعی یکی از مهم‌ترین ویژگی‌های ETL است. این ویژگی به‌عنوان یک معیار مهم در مقایسهٔ ابزارهای ETL در جدول ۱، آورده شده است. از جملهٔ معیارهای مهم دیگر یک ابزار ETL می‌توان به اعتبارسنجی داده‌ها و تجاری (C) یا منبع باز بودن (O) ابزارها اشاره کرد (Vilakkumadathil 2016).

جدول ۱. مقایسهٔ ابزارهای ETL

ابزار	سادگی استفاده	نزدیکی به زمان واقعی	قابلیت استفادهٔ مجدد	اعتبارسنجی داده‌ها	ارائهٔ پیش‌نمایش	رابط کاربری کاربرپسند	تجاری یا منبع باز
IBM Datastage		√		√	√		C
Pentaho Kettle	√			√	√		O/C
Microsoft SSIS				√			C
Pervasive Data Integrator				√	√		C
Informatica Powercenter		√		√	√	√	C

همان‌طور که در جدول ۱، نشان داده شده، Pentaho از آن دسته از ارائه‌دهندگانی است که نسخه‌ای از نرم‌افزار خود را آزاد و به‌صورت منبع باز در اختیار کاربران خود قرار داده است. البته، این ارائه‌دهندگان عموماً برای کاربرانی که از آن‌ها در محیط واقعی

1. graphic user interface

و در محدوده‌ای وسیع‌تر استفاده نمایند، نسخهٔ تجاری با قابلیت‌های ویژه دارند. همچنین، با توجه به اطلاعات جدول می‌توان مشاهده کرد که IBM Datastage و Informatica Powercenter روش‌هایی هستند که به سمت نزدیک شدن به زمان واقعی حرکت کرده‌اند و در پرداختن به دیگر شاخص‌ها نیز به‌طور یکسان عمل کرده‌اند. شاخص‌هایی مانند اعتبارسنجی داده‌ها و پیچیدگی انتقال که در واقع، به همان مرحلهٔ انتقال در فرایند ETL اشاره دارد، در تمام روش‌ها مورد توجه قرار گرفته‌اند. این امر می‌تواند نمایانگر این موضوع باشد که این شاخص‌ها در بحث تجمیع داده‌ها به شاخص‌های پایه‌ای تبدیل شده‌اند که باید در تمام پژوهش‌های آینده در زمینهٔ تجمیع داده‌ها مورد توجه قرار گیرند. قابلیت استفادهٔ مجدد را در هیچ‌کدام از ابزارهای آورده‌شده در جدول شاهد نیستیم. همچنین، سادگی استفاده و ارائهٔ رابط کاربری کاربرپسند از آن دسته از شاخص‌هایی است که کمتر به آن پرداخته شده است و در کنار دیگر شاخص‌ها تقریباً نادیده گرفته شده‌اند.

۳. راهکار پیشنهادی

راهکار ارائه‌شده با استفاده از وب‌سرویس جهت انجام عملیات ETL، دارای دو فاز اصلی است. فاز اول مربوط به طراحی پایگاه داده و فاز دوم در ارتباط با طراحی سرویس‌ها جهت انجام عملیات استخراج، تبدیل و بارگذاری داده‌هاست. همان‌طور که در شکل ۲، نشان داده شده، هر فاز دارای مراحلی است. مراحل فاز اول بر اساس روش‌هایی که تاکنون وجود داشته، شامل ایجاد جداول متادیتا، ایجاد نماها و ایجاد تریگر[1]هاست. مراحل فاز دوم که هدف اصلی و نوآوری این تحقیق است، در ارتباط با استفاده از وب‌سرویس‌ها جهت عملیات استخراج، تبدیل و بارگذاری داده‌هاست. جزئیات هر یک از فازهای اصلی راهکار پیشنهادی در قسمت ذیل تشریح شده است.

۳-۱. فاز طراحی پایگاه داده

گام اول: ابتدا در هر پایگاه داده‌ای که می‌خواهیم در تجمیع داده‌ها شرکت کند، چهار جدول مربوط به ذخیره‌سازی ابرداده‌ها باید ایجاد شود.

- ◇ جدول tblDataTransferSuccessType

در این جدول اطلاعات مربوط به وضعیت انتقال ذخیره شده است.

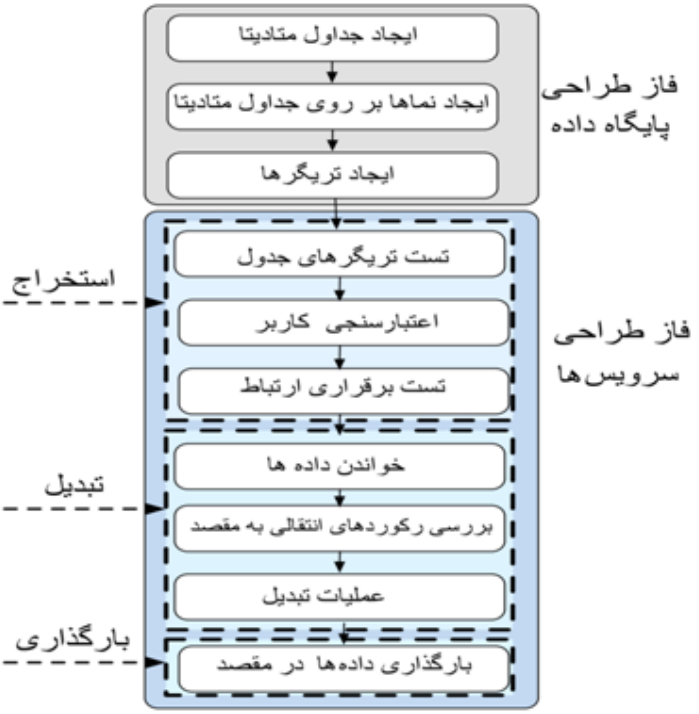
1. trigger

◇ جدول tblDataTransferTransactionType

در این جدول اطلاعات مرتبط با نوع عملیات انجام‌گرفته ذخیره شده است.

◇ جدول TransferAttempt

این جدول، جدول ذخیره‌سازی مرتبط با اطلاعات رکوردهایی است که در تجمیع شرکت کرده‌اند.



شکل ۲. مراحل روش پیشنهادی

◇ جدول tblDataTransferAttemptDetail

این جدول، جدول ذخیره‌سازی داده‌های مرتبط با جدول قبلی است.

◇ جدول tblDataTransferTransaction

در این جدول، اطلاعات مرتبط با رکوردهای اضافه‌شده به‌روزرسانی یا حذف شده و ذخیره می‌گردد.

گام دوم: در این گام تعدادی نما[1] جهت نمایش داده‌ها در رابط کاربری در هر پایگاه داده ایجاد می‌شود. لیست این نماها در ادامه آورده شده است.

- ◇ vwDataTransferSuccessTypeMajor
- ◇ vwDataTransferSuccessTypeMinor
- ◇ vwDataTransferTransactionTypeMajor
- ◇ vwDataTransferTransactionTypeMinor

گام سوم: اجرای سندی[2] که تریگر مربوط به درج، حذف و به‌روزرسانی داده‌ها را ایجاد می‌کند. این تریگرها جداول مربوط به ابرداده‌ها را تغذیه می‌کنند. به این صورت که با درج هر رکورد در یک جدول اطلاعاتی مانند زمان درج، کاربر درج‌کننده، جدول مقصد و نوع فرم ثبت‌شده توسط تریگر درج در جدول tblDataTransferTransaction ذخیره می‌شوند.

پس از فاز طراحی پایگاه داده، به فاز طراحی سرویس‌ها می‌رسیم. در این فاز، چهار گام اول مربوط به فرایند استخراج هستند. سپس، دو گام در فرایند تبدیل و در نهایت، گام آخر مربوط به فرایند بارگذاری داده‌هاست.

۳-۲. فاز طراحی سرویس‌ها

گام اول: این گام شامل تست تریگرها و جداول پایگاه داده است. وب‌سرویس قبل از اقدام به تجمیع داده‌ها، ابتدا در پایگاه دادهٔ مبدأ شروع به بررسی جداول و تریگرهای شرکت‌کننده در تجمیع می‌کند و در صورت گذر موفقیت‌آمیز از این مرحله، وارد گام بعد می‌شود.

گام دوم: در این گام، وب‌سرویس بر اساس آدرس پایگاه دادهٔ مقصد که در لایهٔ نرم‌افزار مشخص می‌شود، تست ارتباط با مقصد را انجام می‌دهد و در صورت موفقیت‌آمیز بودن این گام وارد گام بعد می‌شود.

گام سوم: پس از تست برقراری ارتباط با پایگاه دادهٔ مقصد، پایگاه داده بر اساس نام کاربری و کلمهٔ عبور که در لایهٔ نرم‌افزار وارد شده، اعتبارسنجی ارتباط با مقصد را انجام

1. view
2. script

می‌دهد و در صورت نادرست بودن، پیغام مناسب را نمایش می‌دهد.

گام چهارم: در این گام که آخرین گام از مرحلهٔ استخراج است، داده‌ها از منابع استخراج می‌شوند و پس از آن وارد فرایند تبدیل می‌گردند.

بررسی رکوردهای انتقالی به مقصد: این گام شامل تشخیص رکوردهایی است که از پایگاه دادهٔ مبدأ به مقصد انتقال داده شده است، اما اکنون در پایگاه دادهٔ مقصد وجود ندارد. در این بخش وب‌سرویس بر مبنای داده‌های موجود در جداول TransferAttempt و TransferAttemptDetail، و نوع هر رکورد که آیا قبلاً انتقال موفقیت‌آمیز داشته یا نه، در صورت وجود رکوردهایی که در مبدأ وجود دارند و در مقصد وجود ندارند، پیغامی را به کاربر نرم‌افزار نمایش می‌دهد. کاربر در صورت لزوم می‌تواند مجدداً انتقال را انجام دهد. همچنین، در این گام رکوردهایی که تاکنون در تجمیع شرکت نکرده‌اند، مشخص می‌شوند. در جدول tblDataTransferTransaction، وب‌سرویس به‌دنبال رکوردهایی می‌گردد که در انتقال شرکت نکرده‌اند و در صفحهٔ نمایش نرم‌افزار کاربری آن‌ها را به کاربر نمایش می‌دهد.

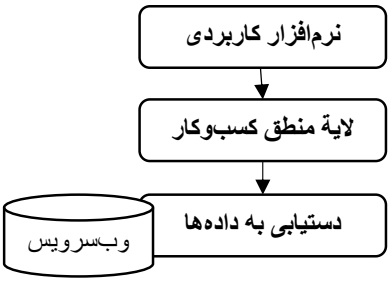
عملیات تبدیل: در این گام، بر اساس فیلدهای کلید خارجی مرتبط با رکوردی که باید منتقل شود، عملیات انتقال صورت می‌پذیرد و در صورت نیاز پیغام خطای مناسب به کاربر نمایش داده می‌شود. در توضیح بیشتر این گام می‌توان گفت که فرایند تبدیل داده‌ها در روش پیشنهادی شامل محاسبه و بررسی فیلد کلید برای جلوگیری از تکراری بودن با استفاده از یک الگوریتم آماری و در برخی فیلدها شامل تبدیل نوع آن‌ها به یک نوع دیگر و یا تفکیک یک فیلد به دو فیلد است.

بارگذاری داده‌ها در مقصد: در این گام که می‌توان آن را همان گام بارگذاری فرایند ETL دانست، داده‌ها بر اساس شمارهٔ کلید مرتبط با نوع فرم‌ها در پایگاه دادهٔ مقصد درج می‌گردند. در صورت برخورد با هرگونه خطا از جمله قطع ارتباط با مقصد، اطلاعات مناسبی در جداول TransferAttempt و TransferAttemptDetail درج می‌گردد. برای مثال، اگر عملیات انتقال یک رکورد موفقیت‌آمیز نباشد، یک رکورد در جدول TransferAttempt درج می‌شود و نوع فیلد IsSuccess آن برابر با مقدار false مقداردهی می‌شود. به این ترتیب، سرویس در انتقال‌های بعدی متوجه رکورد منتقل‌نشده می‌شود و مجدداً سعی بر ارسال آن می‌کند.

۳-۳. پیاده‌سازی راهکار پیشنهادی

در این قسمت ابتدا به تشریح نحوهٔ پیاده‌سازی راهکار پیشنهادی می‌پردازیم و سپس، نحوهٔ اجرای آن در مطالعهٔ موردی توضیح داده خواهد شد.

راهکار پیشنهادی با استفاده از روش معماری سه‌لایه پیاده‌سازی شده است که امروزه در طراحی سیستم‌های نرم‌افزاری به کار می‌رود. در این معماری، سیستم نرم‌افزاری به چند زیرسیستم تقسیم می‌شود و قسمت‌های مستقل سیستم به‌صورت لایه‌های جداگانه طراحی می‌گردند. این لایه‌ها علاوه بر این‌که هر کدام وظیفهٔ مستقل خود را دارند، با هم نیز در ارتباط بوده و به لایه‌های دیگر سرویس‌دهی می‌کنند. از آنجا که در این نوع معماری، تمامی لایه‌ها مستقل از یکدیگر عمل می‌کنند، در صورت نیاز می‌توان با سرعت بالا و هزینهٔ پایین بر روی هر لایه تغییراتی اعمال کرد؛ بدون آن‌که نیاز به تغییر در سایر لایه‌ها وجود داشته باشد. در نتیجه، هرگونه تغییر و گسترش برنامه آسان‌تر خواهد شد. لایه‌ها عبارت‌اند از لایه‌های دستیابی به داده‌ها که در اینجا همان وب‌سرویس است و عملیات ذخیره، بازیابی، حذف و یا به‌روزرسانی داده‌ها از پایگاه داده را انجام می‌دهد. لایه‌های بعدی عبارت‌اند از: منطق کسب‌وکار برای تبدیل داده‌ها جهت نمایش در لایهٔ کاربری نرم‌افزار. (شکل ۳) (Li and Wu 2009). منبع کدهای وب‌سرویس و تریگرهای نرم‌افزار به‌طور کامل در پیوست‌های ۲ ضمیمه شده است.



شکل ۳. معماری سه‌لایهٔ ابزار پیشنهادی

مطالعهٔ موردی

مطالعهٔ موردی پژوهش بر روی «شرکت شاهسوند» انجام شده است. این شرکت در سال ۱۳۷۲ به‌منظور تأمین بخشی از نیاز بازار محصولات غذایی کشور تأسیس گردید و

هم‌اکنون با تولید ۱۷ گروه از محصولات غذایی یکی از تولیدکننده‌ها در صنعت غذاست. این شرکت، دارای ۱۷ شعبه و ۱۲ زیرشعبه در سراسر ایران است. بین سال‌های ۱۳۷۷ تا ۱۳۸۳ که سیستم‌ها به‌صورت جزیره‌ای بود، برای تجمیع داده‌ها به‌صورت دستی عمل می‌کرد. به این صورت که داده‌ها از فایل‌هایی مانند فایل‌های متنی به‌صورت دستی در فایل مقصد تجمیع می‌شدند. پس از آن تحقیق و بررسی بر روی راهکارهای مختلف تجمیع داده‌ها انجام شد. بعد از آن برای مدتی از ابزار تجمیع داده‌ها در SQL Serve (SSIS[1]) جهت تجمیع اطلاعات ۱۷ شعبه استفاده شد. اما نیاز به نیروی متخصص جهت کار با این ابزار و بررسی اشکالات به‌وجودآمده در فرایند تجمیع باعث شد که به استفاده از روشی روی بیاورند که علاوه بر پاسخگو بودن به نیازهایی مانند تجمیع داده‌ها با سرعت مناسب، با داشتن رابط کاربری مناسب و کاربرپسند، برای تمام کارکنان قابل درک و استفاده باشد.

۳-٤. اجرای روش پیشنهادی

۳-٤-۱. طراحی و پیاده‌سازی پایگاه دادهٔ شعب

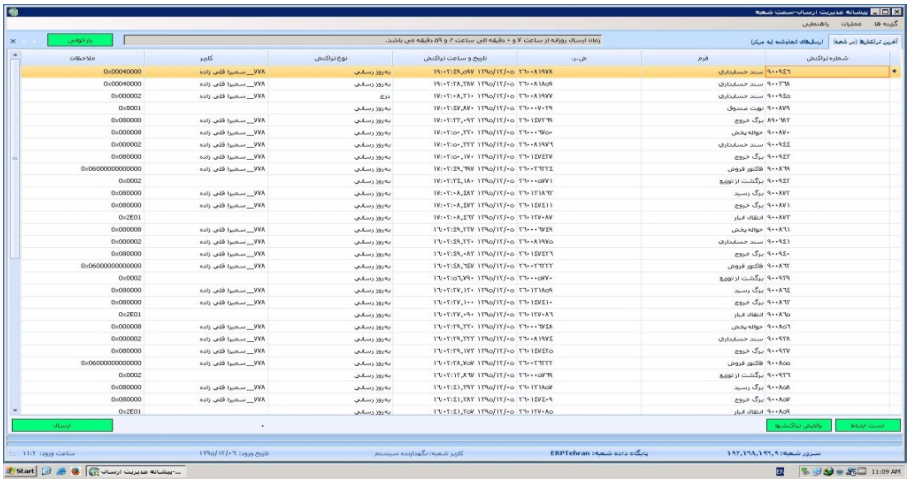
ابتدا جداول متادیتا و نماهایی را که برای فرایند تجمیع لازم است، بر روی پایگاه دادهٔ مورد نظر ایجاد و پیاده‌سازی می‌شوند. پس از آن طراحی و پیاده‌سازی منبع دادهٔ مقصد صورت می‌گیرد، به‌طوری که نحوهٔ پیاده‌سازی جدول داده‌ها در مقصد همانند پیاده‌سازی جدول داده‌ها در مقصد است که می‌تواند شامل جداول مربوط به فروش، انبارداری، منابع انسانی، مالی و ... باشد. بهتر است منبع دادهٔ مقصد در یک سرور با فضای هارددیسک کافی و با سرعت پردازندهٔ مناسب نصب شود. مرحلهٔ بعد در ارتباط با استقرار وب‌سرویس است. پس از انتشار سرویس‌ها، آن‌ها را باید جهت استفاده بر روی یک سرویس اطلاعاتی اینترنت (IIS[2]) مستقر کرد. بدیهی است که سرویس‌ها جهت انجام فرایند تجمیع باید از آدرس منبع دادهٔ مقصد باخبر باشند. پس از طراحی و استقرار پایگاه دادهٔ شعب در مطالعهٔ موردی، نرم‌افزار ارائه‌شده تجمیع و بر روی سیستم شعبه‌ها نصب شده و مانند هر نرم‌افزار تحت ویندوز دیگری به آسانی قابل نصب خواهد بود. نکتهٔ مهم تنظیمات لازم در فایل تنظیمات نرم‌افزار است، به‌طوری که وب‌سرویس فایل مربوطه را بتواند به منبع داده در شعبهٔ خود متصل نماید. این ارتباط می‌تواند از طریق اینترنت و یا اینترانتِ شرکت باشد.

1. SQL Server Integration Service
2. Internet Information Service

در نهایت، پس از تست سخت‌افزاری موفقیت‌آمیز آدرس وب‌سرویس، می‌توان نرم‌افزار را اجرا کرد. شکل ۴، صفحهٔ ورود به نرم‌افزار تجمیع داده‌ها را نشان می‌دهد. همان‌طور که در شکل نشان داده شده، پس از اجرای نرم‌افزار توسط کاربر، ابتدا اتصال به وب‌سرویس توسط نرم‌افزار بررسی می‌شود. در صورت برقراری ارتباط، صفحهٔ ورود ظاهر می‌شود.

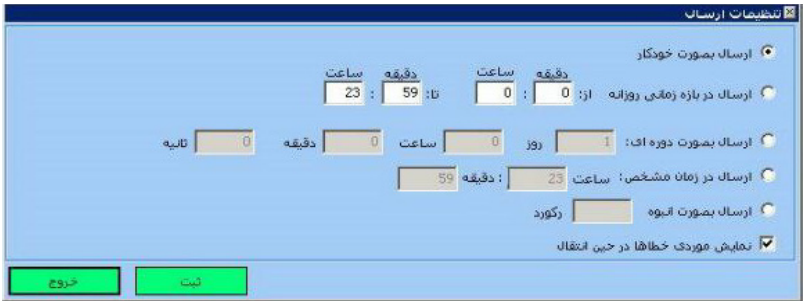
شکل٤. صفحهٔ ورود به نرم‌افزار تجمیع داده‌ها

شکل ۵، صفحهٔ اصلی نرم‌افزار را نشان می‌دهد. پس از ورود موفقیت‌آمیز کاربر به نرم‌افزار صفحهٔ اصلی نرم‌افزار نمایش داده می‌شود. همهٔ عملیات کاربر، از طریق منوهای این صفحه انجام می‌شود. در ابتدای ورود در صفحهٔ اصلی تمام رکوردهایی که باید تجمیع شوند اما هنوز در فرایند تجمیع شرکت نکرده‌اند، به کاربر نمایش داده می‌شود. در صورتی که رکورد منتقل‌نشده‌ای وجود نداشته باشد، پیام No data to display نمایش داده می‌شود. همچنین، با کلیک بر روی دکمهٔ ارسال در سمت راست پایین صفحه، وب‌سرویس شروع به استخراج، تبدیل و بارگذاری داده‌ها در مقصد می‌کند. با کلیک بر روی تست ارتباط می‌توان از اتصال نرم‌افزار به پایگاه دادهٔ مقصد از طریق وب‌سرویس اطمینان حاصل نمود. مادامی که نرم‌افزار در حال اجرا باشد، نرم‌افزار هر رکوردی را که در پایگاه دادهٔ مقصد درج می‌شود بدون وقفه برای مقصد ارسال می‌کند.



شکل ۵. صفحهٔ اصلی نرم‌افزار

یکی دیگر از قابلیت‌های مهم در این نرم‌افزار امکان ایجاد برنامهٔ زمان‌بندی برای فرایند تجمیع داده‌هاست. همان‌گونه که در شکل ۶، نمایش داده شده، می‌توان یک زمان‌بندی جدید برای انجام عملیات تجمیع ایجاد نمود. به این ترتیب، می‌توان در صورت نیاز، زمان‌بندی را به‌گونه‌ای تنظیم نمود که به‌عنوان مثال، در پایان هر روزِ کاری تجمیع داده‌ها به‌صورت خودکار انجام شود.



شکل ٦. نمایی از صفحهٔ ایجاد دوره برای تجمیع داده‌ها

۳-٤-۲. ارزیابی راهکار پیشنهادی

در جدول ۲، مقایسهٔ ابزار پیشنهادی به همراه سه ابزار مهم دیگر آورده شده است. ابزارها بر مبنای معیارهای سرعت و استفاده از حافظه مورد بررسی قرار گرفته‌اند. این

بررسی بر اساس انتقال ۵۰۰۰ رکورد از منبع دادهٔ مبدأ به منبع دادهٔ مقصد است. وب‌سرویس بر روی ویندوز سرور ۲۰۱۲ نصب شده است و ابزارهای تجمیع بر روی ویندوز ۸ و بر روی یک سیستم اجرا شده‌اند تا از نظر سرعت پردازنده و مقدار حافظه در شرایط یکسان تست شوند. مشخصات هر یک از ابزارهای ETL در پیوست ۴، آورده شده است.

جدول ۲. ارزیابی ابزار پیشنهادی

ابزار ETL	سرعت (دقیقه)	استفاده از حافظه (مگا بایت)
IBM Data Stage	۱۷	۸۹۳
Oracle Integration Tools	۱۳	۸۳۰
SQL Server Integration Service	۱۷	۸۰۴
راهکار پیشنهادی	۱۵	۷۹۹

با توجه به این جدول می‌توان گفت که راهکار پیشنهادی توانسته است نسبت به IBM و SSIS زمان را کاهش دهد. همچنین، راهکار پیشنهادی توانسته مقدار استفاده از حافظه را در زمان اجرا نسبت به دیگر روش ها کاهش دهد. جهت روشن‌ترشدن نتایج ارزیابی، در زمان تست هر یک از ابزارها با دیتاست مورد نظر، تصاویری ثبت شده که در پیوست شماره ۳، آورده شده است. برای بررسی عملکرد نرم‌افزار از نظر معیار سهولت استفاده از نظر کاربران و ارزیابی رضایت آنان پرسشنامه‌ای تهیه شده که به پیوست آورده شده است. این پرسشنامه از Kibugu (2016) اقتباس شده است. این پرسشنامه توسط ۰۲ نفر که ۸ نفر از آن‌ها در حوزهٔ کار با نرم‌افزار متخصص بوده‌اند، تکمیل شده است. در جدول ۳، می‌توان نتایج این بررسی را مشاهده نمود. همان‌طور که در جدول ۲، نشان داده شده، پنج سؤال اصلی از ۰۲ افراد خبره بر مبنای مقیاس «لیکرت» نظرسنجی شده است؛ به‌طوری که میانگین اکثر آیتم‌ها بین ۳ تا ۴ بر مبنای مقیاس ۴ امتیازی است. بنابراین، می‌توان نتیجه گرفت که رابط کاربری ارائه‌شده توانسته است استفاده از نرم‌افزار را برای کاربران «شرکت شاهسوند» سهولت بخشد. جهت ارزیابی سطح اطمینان داده‌ای به‌دست‌آمده جدول ۲، ضریب «کرونباخ» آلفا نیز توسط نرم‌افزار SPSS محاسبه شده است. ضریب «کرونباخ» آلفا برای داده‌های فوق حدود ۰/۸۲۷ است که نشان‌دهندهٔ قابلیت اطمینان بالا می‌باشد.

جدول ۳. نتایج نظرسنجی از افراد خبره

خبره	زمینه	Q1	Q2	Q3	Q4	Q5	میانگین
Expert 1	طراح نرم‌افزار	۳	۳	۴	۴	۳	۳/۴
Expert 2	طراح نرم‌افزار	۴	۳	٤	۴	۴	۳/۸
Expert 3	طراح نرم‌افزار	۳	۴	۳	۴	۳	۳/۴
Expert 4	طراح نرم‌افزار	۳	۴	۳	۴	۳	۳/۴
Expert 5	طراح نرم‌افزار	۴	۴	۳	۴	۴	۳/۸
Expert 6	طراح نرم‌افزار	۳	۳	۴	۴	۴	۳/۶
Expert 7	تحلیلگر سیستم	۳	۴	۳	۳	۴	۳/۴
Expert 8	طراح نرم‌افزار	۳	۴	۴	۳	۴	۳/۶
Expert 9	طراح نرم‌افزار	۱	۳	۳	۳	۲	۲/۲
Expert 10	تحلیلگر سیستم	۱	۳	۳	۳	۲	۲/۲
Expert 11	تحلیلگر سیستم	۳	۴	۴	۴	۳	۳/۴
Expert 12	کاربر پایگاه داده	۲	۴	۴	۳	۴	۳/۲
Expert 13	کاربر پایگاه داده	۴	۳	۳	۳	۴	۳/۶
Expert 14	کاربر پایگاه داده	۴	۳	۳	۳	۲	۳
Expert 15	طراح پایگاه داده	۳	۴	۴	۳	۲	۳/۲
Expert 16	طراح پایگاه داده	۱	۲	۲	۳	۳	۲/۲
Expert 17	طراح پایگاه داده	۴	۴	۴	۳	۳	۳/۶
Expert 18	طراح پایگاه داده	۴	۴	۴	۳	۴	۳/۶
Expert 19	طراح پایگاه داده	۳	۴	۴	۴	۳	۳/۴
Expert 20	طراح پایگاه داده	۳	۳	۳	۴	۴	۳/۶

٤. نتیجه‌گیری

این پژوهش به‌دنبال ارائهٔ راهکاری جهت تجمیع داده است. با توجه به اهمیت سرعت در دنیای امروز و همچنین، سعی در کاهش هزینه‌ها سعی بر آن شد تا راهکاری ارائه شود که بتواند سرعت را افزایش و استفاده از منابع را کاهش دهد. بنابراین، در راهکار ارائه‌شده سعی بر آن شد که با استفاده از وب‌سرویس جهت انجام عملیات ETL و همچنین، استفاده از متادیتا جهت تشخیص داده‌هایی که باید منتقل شوند، سرعت را افزایش داده و به زمان واقعی نزدیک‌تر کرد. همچنین، با توجه به اهمیت سهولت استفاده از ابزار سعی بر آن شد که رابط گرافیکی کاربرپسندی ارائه شود تا بتواند سهولت استفاده از ابزار ارائه‌شده را برای کاربران در پی داشته باشد. راهکار پیشنهادی به‌صورت گام‌به‌گام شرح داده شد. در ادامه، نحوهٔ پیاده‌سازی ابزار ETL و همچنین، راه‌اندازی ابزار جهت تجمیع

داده‌ها در شعب مختلف بیان شد. طبق ارزیابی که انجام شد، مشخص شد که راهکار پیشنهادی توانسته در حجم داده‌هایی مشخص سرعت انتقال را افزایش و میزان استفاده از منابع را بهبود بخشد. همچنین، طبق ارزیابی پاسخ‌های داده‌شده به پرسشنامه مشخص شد که طراحی کاربرپسند نرم‌افزار توانسته است سهولت استفاده از ابزار را برای کاربران در پی داشته باشد.

با وجود تحقیقات و تلاش‌های بسیار زیادی که انجام شده، می‌توان گفت که هنوز هم می‌توان سرعت را در فرایند ETL افزایش داد و با توجه به نیاز به دسترسی به داده‌های جامع و باطراوت نزدیک‌تر شدن به زمان واقعی توصیه می‌شود. خودکارسازی یکسری فرایندها می‌تواند انجام شود؛ به‌خصوص ایجاد جدول‌ها و نماها و تریگرها که در حال حاضر به‌صورت دستی ایجاد می‌شوند. با وجود تحقیقات فراوان در خصوص تجمیع داده‌ها در سراسر دنیا، این موضوع همچنان در کشور ما مغفول است و تحقیقات بیشتر در این حوزه جهت دستیابی به یک راهکار و اجرای آن در شرکت‌ها توصیه می‌شود.

فهرست منابع

Abiteboul, S., O. Benjelloun, and T. Milo. 2002. *Web services and data integration*. IEEE Confernace, Singapore 2002, WISE publishing.

Abiteboul, S., B. Nguyen, and G. Ruberg, 2006. Building an active content warehouse. *Processing and managing complex data for decision support* 1: 63-72.

Adepetia, E. T. L. 2015. Compare data integration vendors. https://adeptia.com/products/etl_vendor_comparison.html (accessed Feb. 20, 2015).

Agrawal, D. 2008. *The reality of real-time business intelligence*. in International Workshop on Business Intelligence for the Real-Time Enterprise. Springer Confernace, Berlin, Heidelberg. 2008 Aug 24.

Alandi, A. 2016. Technical Comprehensive Survey of ETL Tools. *International Journal of Applied Engineering Research* 11 (2): 2557-2559.

Ali, F. S. E. 2014. A Survey of Real-Time Data Warehouse and ETL. *International Scientific Journal of Management Information Systems* 9 (3): 3-9.

Arenas, M. 2003. The hyperion project: from data integration to data coordination. *ACM Sigmod Record* 32 (3): 53-58.

Carey, M. J. 1995. Towards heterogeneous multimedia information systems: The Garlic approach. in Research Issues in Data Engineering, 1995: Distributed Object Management, Proceedings. RIDE-DOM'95. Fifth International Workshop on IEEE. Taiwan.

De Faria Cordeiro, K., M. L. M. Campos, and M. R. Borges. 2015. aDApTA: Adaptive approach to information integration in dynamic environments. *Journal of Computers in Industry* 71: 88-102.

Florea, A. M. I., V. Diaconita, and R. Bologa. 2016. Data integration approaches using ETL. *Database Systems Journal* 6 (3): 19-27.

Gour1, V., S. S. Sarangdvot, and S. tanwar. 2010. Improve Performance of Extract, Transform and Load (ETL) in Data Warehouse. *International Journal on Computer Science and Engineering* (IJCSE) 1:

786-789.

Greene, L. A., 2016. Generating and implementing data integration job execution design recommendations. Google Patents. https://patents.google.com/patent/US9501377B2/en, (access 20 Jun 2017).

Halevy, A., A. Rajaraman, and J. Ordille. 2006. Data integration: the teenage years. in Proceedings of the 32nd international conference on Very large data bases. VLDB Endowment. Seoul, Korea

Jain, T., S. Rajasree, and S. Saluja. 2012. Refreshing datawarehouse in near real-time. *International Journal of Computer Applications* 46 (18): 24-29.

Kadir, R. A. L. A. 2013. Impact of DBM Best practices on Community Rehabilitation Projects. *Journal of the planner and development* 27: 89-113.

Kibugu, A. 2016. A Methodology for the Implementation of a Data Warehouse Using an Etl Process Model for Improved Decision Support. Doctoral dissertation, University of Nairobi.

Landers, T. and R. L. Rosenberg. 1986. *An overview of Multibase*. in Distributed systems, Vol. II: distributed data base systems. Norwood, MA, USA.: Artech House, Inc.

Leigh, J. and K. H. T. Lau. 2016. *Use of projector and selector component types for ETL map design.* Google Patents. https://patents.google.com/patent/US5283600, (access 27 Jul 2017).

Li, H. and Z. Wu. 2009. *Research on distributed architecture based on SOA*. Communication Software and Networks, 2009. ICCSN'09. International Conference on. IEEE: 670-674. *Macau, China*

Lohiya, A. S., D. M. Gojare, S. M. Karde, and A. A. Patil. 2017. Optimize ETL For Banking DDS: Data Refinement Using ETL Process For Banking Detail Data Store (DDS). *Imperial Journal of Interdisciplinary Research* 3 (3): 3-10.

Madhavan, J. 2007. Web-scale data integration: You can only afford to pay as you go. in Proceedings of CIDR. https://pdfs.semanticscholar.org/2c8e/ba60f6ad294a9c81e51dce75ace49e4c8a33.pdf, (accessed 20 Feb. 2017).

Majumdar, G., T. K. Saha, and M. Q. Wang. 2016. Data integration using automated data processing based on target metadata. Google Patents. https://patents.justia.com/inventor/michael-q-wang, (accessed 20 Oct. 2017).

Muddasir, M., N. R. Kumar, and V. Prajwal. 2016. Methods to Enhance Transformation in Near Real Time ETL. *Methods* 13730-48 :(5) .

Oracle Data Integrator. 2016. Available from: http://www.oracle.com/technetwork/middleware/dataintegrator/overview/index-0883. (acccessed Jan. 2016).

Preeti Dhanda, N. S. 2016. Extract Transform Load Data with ETL Tools. *International Journal of Advanced Research in Computer Science* 7 (1): 43-74.

Salem, R., O. Boussaïd, and J. Darmont. 2013. Active XML-based Web data integration. *Information Systems Frontiers* 15 (3): 371-398.

Vilakkumadathil, R. R. 2016. Validating code of an extract, transform and load (ETL) tool. US Patent https://patents.google.com/patent/US9244809B1/en, (accessed Feb. 25, 2017).

Ziegler, P., and K. R. Dittrich. 2004. Three Decades of Data Intecration—all Problems Solved? in Building the information Society. Springer 1: 3-12. Boston,US.

Zuters, J. 2011. Near real-time data warehousing with multi-stage trickle and flip in International Conference on Business Informatics Research. Springer Publishing. Riga, Latvia.

پیوست ۱. پرسشنامه

در این مقاله به‌دنبال ارائهٔ راهکاری جهت جمع‌آوری داده‌ها و سازماندهی اطلاعات سازمان‌ها پرداخته شده است. به‌دنبال آن هستیم تا عملکرد نرم‌افزار تجمیع داده‌ها و سهولت استفاده از آن را بررسی نماییم. خواهشمند است با توجه به نرم‌افزار تجمیع داده‌ها و با صداقت به سؤالات پرسشنامه پاسخ دهید:

بخش اول: اطلاعات فردی

۱. لطفاً نام و اطلاعات تماس خود را وارد نمایید:

نام و نام خانوادگی:

ایمیل:

لطفاً پاسخ مورد نظر خود را با ✓ مشخص نمایید.

حوزهٔ فعالیت/ تخصص شما چیست؟

☐ کارشناس IT

☐ مدیر IT

☐ کارشناس حسابداری

☐ مجری اتوماسیون اداری

☐ سایر

سابقهٔ فعالیت شما در حوزهٔ کاری خود چند سال است؟

☐ سال اول ☐ ۱-۲ سال ☐ ۳-۵ سال ☐ ۵-۱۰ سال ☐ بیشتر از ۱۰ سال

بخش اول: عملکرد نرم‌افزار

۲. آشنایی شما با نرم‌فزار تجمیع داده‌ها چقدر است؟

☐ ضعیف ☐ متوسط ☐ خوب ☐ عالی

۳. طراحی صفحات نرم‌افزار را چگونه ارزیابی می‌کنید؟

☐ ضعیف ☐ متوسط ☐ خوب ☐ عالی

۴. ناوبری نرم‌افزار را برای رفتن از یک نقطه به نقطهٔ دیگر چگونه می‌بینید؟

☐ ضعیف ☐ متوسط ☐ خوب ☐ عالی

اطلاعات ارائه‌شده در صفحات نرم‌افزار را چقدر در درک وضعیت انجام فرایند تجمیع

داده‌هـا مناسـب می‌دانیـد؟

☐ ضعیف ☐ متوسط ☐ خوب ☐ عالی

۵. تأثیر نرم‌افزار را در سهولت استفاده از ابزار تجمیع داده‌ها چگونه ارزیابی می‌کنید؟

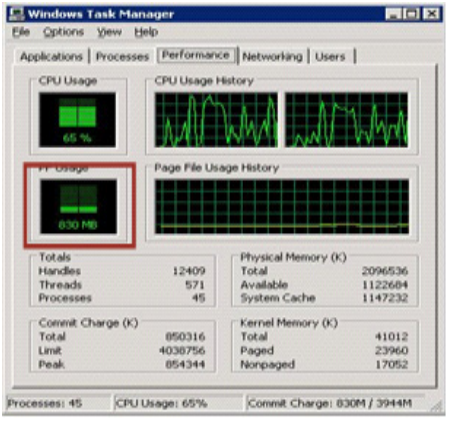
☐ ضعیف ☐ متوسط ☐ خوب ☐ عالی

پیوست ۲. تریگرهای استفاده‌شده در تجمیع

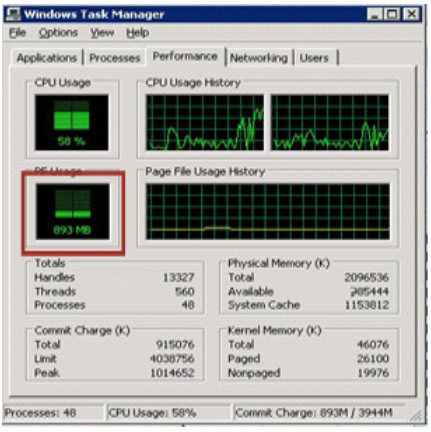
```
----------Delete Trigger---------- //
ALTER trigger [dbo].[trg_tblAccountDocument_delete] on [dbo].[tblAccountDocument] for delete
as
*/
insert into tblDataTransferTransaction (FormTypeID, FormID, TransactionDate, FarsiTransactionDate,
TypeID)
(Select 303, x.ID, GETDATE(), dbo.fnAtashGetFarsiDate (GETDATE()), 3 from deleted x)
/*
set rowcount
set nocount on
set transaction isolation level read uncommitted --uncommitted
update tblDataTransferTransaction --with (rowlock)
set TransactionDate=GETDATE(), FarsiTransactionDate=dbo.
fnAtashGetFarsiDate(GETDATE()),Notes=Notes ,TypeID=3
where exists(Select x.ID from deleted x where x.ID=tblDataTransferTransaction.FormID)
and FormTypeID=303

insert into tblDataTransferTransaction (FormTypeID, FormID, TransactionDate,
FarsiTransactionDate, Notes, TypeID)
Select distinct 303, x.ID, GETDATE(), dbo.fnAtashGetFarsiDate(GETDATE()), Null, 3)
from deleted x
where not exists(Select y.ID from tblDataTransferTransaction y with (readcommitted,
(rowlock, readpast) where y.FormTypeID=303 and y.FormID=x.ID)
Insert Trigger----------//
ALTER trigger [dbo].[trg_tblAccountDocument_insert] on [dbo].[tblAccountDocument] for insert
as
*/
insert into tblDataTransferTransaction (FormTypeID, FormID, TransactionDate, FarsiTransactionDate,
TypeID)
(Select 303, x.ID, GETDATE(), dbo.fnAtashGetFarsiDate(GETDATE()), 1 from inserted x)
/*
set rowcount 0
set nocount on
set transaction isolation level read uncommitted --uncommitted
update tblDataTransferTransaction --with (rowlock)
set TransactionDate=GETDATE(), FarsiTransactionDate=dbo.
fnAtashGetFarsiDate(GETDATE()),Notes=Notes ,TypeID=1
where exists(Select x.ID from inserted x where x.ID=tblDataTransferTransaction.FormID)
and FormTypeID=303
insert into tblDataTransferTransaction (FormTypeID, FormID, TransactionDate,
FarsiTransactionDate, Notes, TypeID)
Select distinct 303, x.ID, GETDATE(), dbo.fnAtashGetFarsiDate(GETDATE()), Null, 1)
from inserted x
where not exists(Select y.ID from tblDataTransferTransaction y with (readcommitted,
```

```
(rowlock, readpast) where y.FormTypeID=303 and y.FormID=x.ID)
Update Trigger----------//
ALTER trigger [dbo].[trg_tblAccountDocument_update] on [dbo].[tblAccountDocument] for update
as
*/
insert into tblDataTransferTransaction (FormTypeID, FormID, TransactionDate, FarsiTransactionDate,
TypeID)
(Select 303, x.ID, GETDATE(), dbo.fnAtashGetFarsiDate(GETDATE()), 2 from inserted x)
/*
set rowcount 0
set nocount on
set transaction isolation level read uncommitted --uncommitted
update tblDataTransferTransaction --with (rowlock)
set TransactionDate=GETDATE(), FarsiTransactionDate=dbo.
fnAtashGetFarsiDate(GETDATE()),Notes=convert(Varchar(1000),COLUMNS_UPDATED(),1)
,TypeID=2
where exists(Select x.ID from inserted x where x.ID=tblDataTransferTransaction.FormID)
and FormTypeID=303
insert into tblDataTransferTransaction (FormTypeID, FormID, TransactionDate,
FarsiTransactionDate, Notes, TypeID)
Select distinct 303, x.ID, GETDATE(), dbo.fnAtashGetFarsiDate(GETDATE()),)
convert(Varchar(1000),COLUMNS_UPDATED(),1), 2
from inserted x
 where not exists(Select y.ID from tblDataTransferTransaction y with (readcommitted, rowlock, readpast)
                                  where y.FormTypeID=303 and y.FormID=x.ID)
```

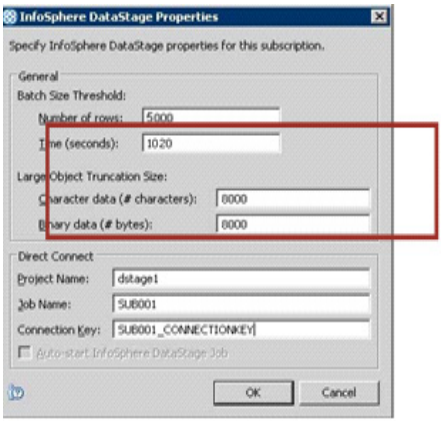
پیوست ۳. تصاویر ثبت‌شده در هنگام تست ابزارهای مختلف جهت ارزیابی



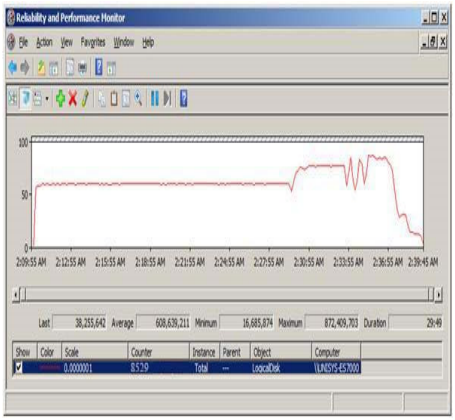
نمایی از مصرف حافظه در ابزار Oracle



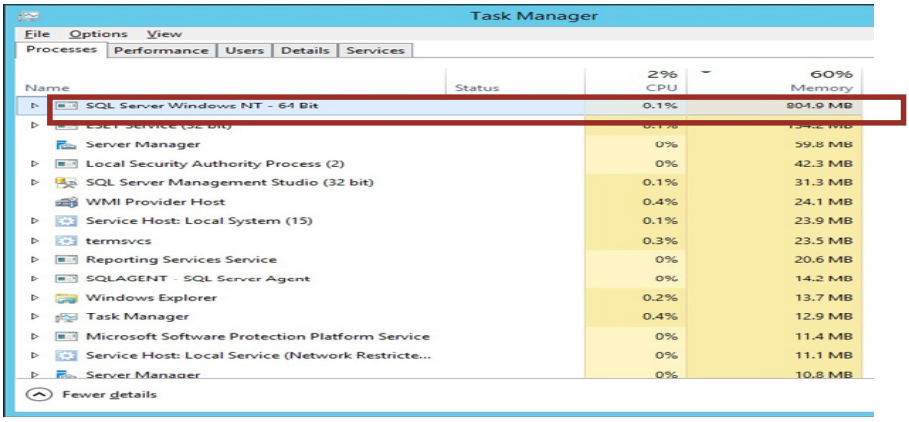
نمایی از مصرف حافظه در ابزار IBM Data Storage



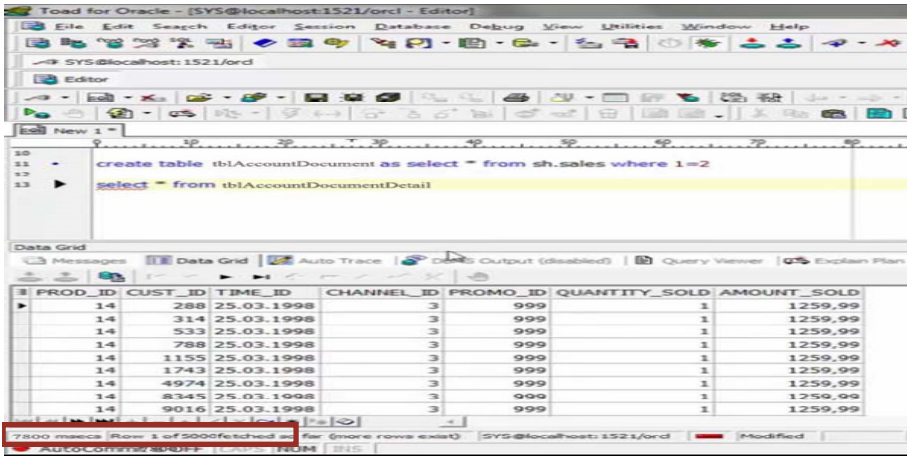
نمایی از زمان استفاده‌شده در ابزار IBM Data Storage



نمایی از زمان استفاده‌شده در ابزار SSI



نمایی از مصرف حافظه در ابزار SSIS



نمایی از زمان استفاده‌شده در ابزار Oracle

پیوست ٤. مشخصات ابزارهای مقایسه‌شده با راهکار پیشنهادی

ویژگی	ابزار
پشتیبانی از real-time web services و messaging systems	IBM Data Stage
پشتیبانی ETL با استفاده از enterprise connectivity و پردازش موازی	
پشتیبانی از real-time data	Oracle Integration Tools
پشتیبانی ETL	
پشتیبانی از data integration و workflow applications	SQL Server Integration Service
پشتیبانی ETL	

عالیه علی‌آبادی

دانشجوی کارشناسی ارشد مهندسی کامپیوتر (نرم‌افزار) است.
طراحی و پیاده‌سازی پایگاه داده و مدیریت پایگاه داده از جمله علایق پژوهشی وی است.

محسن محمدی

متولد سال ۱۳۵۴ دارای مدرک تحصیلی دکتری در رشتهٔ فناوری اطلاعات است. ایشان هم‌اکنون استادیار گروه کامپیوتر مجتمع آموزش عالی فنی مهندسی اسفراین است.
مدل‌سازی و طراحی سیستم‌های اطلاعاتی، مدیریت فرایندهای کسب‌وکار و فرایندکاوی از جمله علایق پژوهشی وی است.