

انبارهای داده (Data Warehousing)

پیاده‌سازی یک نمونه با Oracle Warehouse Builder 11g

تألیف و ترجمه:
مهندس رضا نخلبند



**NAGHOOS
PUBLICATION**

سرشناسه : نخلبند، رضا، ۱۳۶۶ -
 عنوان و نام پدیدآور : انبارهای داده = Data Warehousing : پیاده‌سازی یک نمونه با
 مشخصات نشر : تهران: انتشارات ناقوس، ۱۳۹۰.
 مشخصات ظاهری : ۲۰۸ ص: مصور، نمودار.
 شابک : ۹۷۸-۹۶۴-۳۷۷-۴۹۵-۰ : بها: ۸۰۰۰۰ ریال به همراه DVD
 وضعیت فهرست‌نویسی : فیا
 یادداشت : واژه‌نامه.
 موضوع : ذخیره‌سازی داده‌ها
 موضوع : اوراکل (برنامه کامپیوتر)
 رده بندی کنگره : QA۷۶/۹/۳۳ ۱۳۹۰
 رده بندی دیویی : ۰۰۵/۷۲
 شماره کتابشناسی ملی : ۲۳۵۷۰۸۱



NAGHOOS
PUBLICATION

www.naghoospress.ir

انتشارات ناقوس

چاپ اول

کلیه حقوق برای مؤلف محفوظ
 است. تکثیر تمامی یا قسمتی از
 این اثر به صورت حروفچینی یا
 چاپ مجدد، چاپ الکترونیک،
 فتوکپی و انواع دیگر چاپ ممنوع
 است و پیگرد قانونی دارد.

نام کتاب : Data Warehousing = انبارهای داده :
 (پیاده‌سازی یک نمونه با Oracle Warehouse Builder 11g)
 ناشر : انتشارات ناقوس
 مؤلف : رضا نخلبند
 چاپ اول : ۱۳۹۰
 تیراژ : ۱۰۰۰ جلد
 لیتوگرافی : علامه
 چاپ : علامه
 صحافی : علامه
 طراح جلد : حمید قاسمی
 صفحه‌آرایی : مریم رضائی
 قیمت به همراه DVD : ۸۰۰۰۰ ریال
 شابک : ۹۷۸-۹۶۴-۳۷۷-۴۹۵-۰
 ISBN : 978-964-377-495-0

مراکز پخش:

۱- انتشارات ناقوس: میدان انقلاب، خیابان کارگر جنوبی، خیابان روانمهر، کوچه دولتشاهی، پلاک ۲

تلفن و فاکس: ۶۶۲۰۸۵۲۰ - ۶۶۲۰۱۷۹۸ - ۶۶۲۱۷۰۷۲ - ۶۶۲۶۶۷۹۳

۲- کتابفروشی الیاس: خیابان انقلاب، نبش فروردین تلفن: ۶۶۴۰۵۰۸۲

فهرست مطالب

۱۳	مقدمه
۱۵	فصل اول: انبارهای داده
۱۵	۱-۱ سیستم‌های کامپیوتری و پیدایش انبار داده
۱۶	۱-۱-۱ پیدایش انبار داده
۱۶	۱-۱-۲ انبار داده چیست؟
۱۸	۱-۱-۳ اهداف یک انبار داده
۱۹	۱-۱-۴ طراحی یک پایگاه داده متفاوت
۲۰	۱-۱-۵ استخراج انبار داده از سیستم‌های عملیاتی
۲۱	۲-۱ معرفی برخی از مفاهیم و مؤلفه‌های انبار داده
۲۱	۲-۲-۱ مرکز داده‌ها
۲۲	۲-۲-۲ انبار داده یا مرکز داده؟
۲۲	۲-۲-۳ مدل‌سازی ابعادی
۲۵	۲-۲-۴ جدول Fact
۲۵	۲-۲-۵ جدول بُعد
۲۶	۲-۲-۶ کلیدهای مخزن
۲۷	۳-۱ معماری یک انبار داده
۲۸	۳-۲-۱ معماری های سرور
۲۸	۳-۲-۱-۱ تک پردازشگر و تک دیسکی
۲۹	۳-۲-۱-۲ تک پردازشگر و چند دیسکی
۳۰	۳-۲-۱-۳ چند پردازشگر و چند دیسکی
۳۱	۳-۲-۱-۴ سرورهای دسته‌ای
۳۲	۳-۲-۱-۵ پردازشگرهای موازی در سطح وسیع
۳۲	۳-۲-۱-۶ معماری پایگاه داده اوراکل
۳۴	۳-۲-۲ ترکیبات سخت‌افزار
۳۴	۳-۲-۲-۱ حافظه
۳۴	۳-۲-۲-۲ پردازشگرها
۳۵	۳-۲-۲-۳ ذخیره‌سازی
۳۵	۳-۲-۳ پارتیشن‌بندی داده‌ها
۳۶	۳-۲-۳-۱ چگونه داده‌ها را پارتیشن‌بندی کنیم؟

۳۶	۲-۳-۳-۱ پارتیشن‌بندی محدوده
۳۸	۳-۳-۳-۱ پارتیشن‌بندی درهم
۳۸	۴-۳-۳-۱ پارتیشن‌بندی لیستی
۴۰	۵-۳-۳-۱ پارتیشن‌بندی مرکب

فصل دوم: معرفی و نصب پایگاه داده و OWB

۴۲	۱-۲ معرفی Warehouse Builder
۴۴	۱-۱-۲ ابزار OWB چیست؟
۴۵	۲-۱-۲ اجزاء OWB و معماری آن
۴۶	۲-۲ نصب پایگاه داده اوراکل 11g
۴۹	۱-۲-۲ تنظیمات Listener
۵۲	۳-۲ ساخت پایگاه داده
۶۱	۱-۳-۲ Workspaces تنظیمات مخزن و
۶۶	۴-۲ معرفی محیط Design Center
۶۷	۱-۴-۲ اجزاء محیط Design Center

فصل سوم: تعریف و وارد کردن ساختار داده منابع

۷۱	۱-۲ منابع شرکت فرضی ما
۷۲	۲-۲ تعریف و وارد کردن متادیتای منبع
۷۳	۳-۲ ساخت یک مازول پایگاه داده اوراکل
۷۶	۴-۲ ساخت یک مازول پایگاه داده SQL Server
۷۷	۵-۲ ساخت اتصال پایگاه داده SQL Server
۸۱	۶-۲ تنظیمات اوراکل برای اتصال به SQL Server
۸۱	۱-۶-۳ ساخت فایل پیکربندی سرویس ناممکن
۸۲	۲-۶-۳ ویرایش فایل‌های listener.ora و tnsnames.ora
۸۵	۷-۲ ساخت یک مازول ODBC برای SQL Server در Warehouse Builder
۸۸	۸-۲ وارد کردن متادیتای منبع از یک پایگاه داده
۹۲	۹-۲ تعریف کردن دستی متادیتای منبع توسط Data Object Editor
۹۵	۱۰-۲ وارد کردن متادیتای منبع از فایل‌ها

فصل چهارم: طراحی ساختار مقصد

۱۰۴	۱-۴ طراحی انبار داده در OWB
۱۰۴	۱-۱-۴ ساخت حساب کاربری مقصد
۱۰۶	۲-۱-۴ ساخت مازول مقصد
۱۰۹	۲-۴ ساخت ابعاد و مکعب در OWB
۱۱۰	۱-۲-۴ ایجاد بُعد زمان
۱۱۵	۲-۲-۴ ایجاد بُعد محصول
۱۱۹	۳-۲-۴ ایجاد بُعد فروشگاه

۱۲۲	۴-۲-۴ ایجاد یک مکعب
۱۲۳	۲-۴ مروری بر ابزار Data Object Editor
۱۲۵	۴-۲-۱ اجزاء محیط Data Object Editor
۱۲۹	فصل پنجم: معرفی و پیاده‌سازی فرآیند ETL
۱۲۹	۱-۵ فرآیند ETL چیست؟
۱۳۰	۱-۱-۵ معرفی ناحیه Staging
۱۳۰	۲-۵ عملیات نگاشت و عملگرهای آن در OWB
۱۳۱	۵-۲-۲ Mapping Editor پنجره‌های
۱۳۲	۵-۲-۲ Canvas چیدمان صفحه
۱۳۴	۵-۲-۳ عملگرهای OWB
۱۳۴	۵-۲-۴ عملگرهای مبدأ و مقصد
۱۳۶	۵-۲-۴ عملگرهای جریان داده
۱۳۷	۵-۲-۳ عملگرهای قبل/بعد از عمل پردازش
۱۳۸	۲-۵ طراحی و ساخت یک نگاشت ETL
۱۳۹	۵-۲-۱ ساخت جدول ناحیه Staging به وسیله Data Object Editor
۱۴۲	۵-۲-۲ طراحی و ساخت نگاشت ناحیه Staging
۱۴۳	۵-۲-۳ افزودن جداول منبع
۱۴۵	۵-۲-۲ افزودن جدول مقصد
۱۴۶	۵-۲-۳ اتصال منبع به مقصد
۱۵۶	۴-۵ طراحی و ساخت نگاشت STORE
۱۵۶	۴-۱ افزودن جداول منبع و مقصد
۱۵۷	۴-۲ افزودن عملگرهای تبدیل
۱۶۱	۴-۲ ایجاد یک عملگر جستجوی کلید
۱۶۲	۴-۳-۱ ساخت یک جدول خارجی
۱۶۳	۴-۳-۲ ساخت و بارگذاری یک جدول جستجو
۱۶۵	۴-۴ استفاده از عملگر جستجوی کلید در نگاشت STORE
۱۶۶	۴-۴-۱ افزودن یک عملگر تبدیل SUBSTR
۱۶۷	۴-۴-۲ افزودن یک عملگر CONSTANT
۱۶۸	۴-۴-۳ افزودن یک تبدیل TO_NUMBER
۱۶۹	۴-۴-۴ افزودن یک عملگر Key Lookup
۱۷۱	۵-۵ طراحی و ساخت سایر نگاشت‌های انبار داده
۱۷۲	۵-۱-۵ طراحی و ساخت نگاشت PRODUCT
۱۷۳	۵-۲-۵ طراحی و ساخت نگاشت مکعب SALES
۱۷۴	۵-۲-۵ صفات بُعد در مکعب
۱۷۶	۵-۲-۵ صفات مقداری و سایر صفات در مکعب
۱۷۶	۵-۲-۳ نگاشت صفات و مقادیر به مکعب

۱۸۱	فصل ششم: پیاده‌سازی فیزیکی انبار داده
۱۸۱	۱-۶ عملیات اعتبارسنجی
۱۸۴	۲-۶ عملیات تولید کد
۱۸۴	۱-۲-۶ Design Center در تولید کد
۱۸۷	۲-۲-۶ Data Object Editor در تولید کد
۱۸۸	۲-۲-۶ Mapping Editor در تولید کد
۱۸۹	۳-۶ عملیات گسترش فیزیکی
۱۸۹	۱-۳-۶ Control Center Service چیست؟
۱۹۰	۲-۲-۶ گسترش فیزیکی در Editor
۱۹۳	۲-۲-۶ Control Center Manager ابزار
۱۹۸	۴-۶ عملیات اجرای کد
۲۰۱	پیوست ۱: THE POS TRANSACTIONAL SOURCE DATABASE
۲۰۳	پیوست ۲: THE WEBSITE ORDER MANAGEMENT DATABASE
۲۰۵	لیست علائم و اختصارات
۲۰۷	منابع

سپاسگزاری

رسول خدا فرمودند:

« مَنْ عَلَّمَنِي حَرْفًا فَقَدْ سَيَّرَنِي عَبْدًا »

شکر و سپاس یزدان پاک را که در پناه مساعدت‌های بی‌دریغش، این امکان دست داد تا این کتاب را به رشته تحریر درآورم. اهل قلم خوب می‌دانند که پایان دادن یک اثر، منوط به همکاری بسیاری از افراد است که جا دارد از آنها تشکر و قدردانی گردد. لذا لازم می‌دانم تا از پدر و مادر بزرگوایم که باعث دلگرمی در ارائه این پروژه شدند، قدردانی نمایم. همچنین از استاد گرامی، جناب آقای مهندس گداز و تمامی عزیزانی که در تهیه این کتاب یاریم نمودند، سپاسگزاری می‌کنم.

مقدمه

در دنیای امروزه که دنیای تکنولوژی و ارتباطات است، خواسته یا ناخواسته با داده‌ها و اطلاعات، سر و کار داریم. شاید در حالت عادی این داده‌ها و اطلاعات برای شما بی‌اهمیت باشند، ولی زمانی که نیاز به یک تصمیم‌گیری دارید، ناچارید به دنبال جمع‌آوری این داده‌ها و اطلاعات بروید و آنها را در کنار هم به‌چینید و تحلیل و بررسی کنید تا بهترین تصمیم‌گیری را داشته باشید. حال هرچه این تصمیم‌گیری بزرگتر باشد، به داده‌ها و اطلاعات بیشتر و وسیع‌تری نیاز دارید، تا آنجا که به تصمیم‌گیری‌های استراتژیک سازمانی و تجاری برسید. واضح است که در چنین حالتی، به سادگی و با ابزارهای عادی نمی‌توانید چنین تصمیم‌گیری‌های بزرگ و مهمی را انجام دهید. اینجاست که نقش سیستم‌های پشتیبان تصمیم^۱ و ابزارهای داده^۲ پررنگ می‌شوند و به کمک شما می‌آیند.

در این کتاب سعی کرده‌ام که شما را با مبحث ابزارهای داده آشنا کنم و مراحل طراحی و پیاده‌سازی آن را بطور مفصل توضیح دهم. بیشتر مباحث این کتاب به پیاده‌سازی آن اختصاص داده شده که برای پیاده‌سازی، از پایگاه داده اوراکل^۳ و ابزاری که برای طراحی و مدیریت ابزارهای داده در اختیار ما قرار می‌دهد، استفاده شده است.

دلیل انتخاب پایگاه داده اوراکل، این بوده که این پایگاه داده از قدرتمندترین پایگاه‌های داده موجود تا این زمان است و توانایی بالایی در پردازش داده‌های با حجم بالا دارد و هر ساله نسخه‌های جدیدتری از آن با امکانات پیشرفته‌تر نسبت به قبل را در اختیار ما قرار می‌دهد. در حال حاضر جدیدترین نسخه آن نسخه 11g می‌باشد که در این پروژه از آن استفاده شده است.

اوراکل ابزاری به نام Warehouse Builder دارد که از نسخه 9i به بعد آن را در پکیج خود قرار داده است و از دیدگاه من این ابزار، یکی از بهترین ابزارها برای طراحی و مدیریت ابزار داده است و در این پروژه از نسخه 11g که جدیدترین نسخه آن در حال حاضر می‌باشد، استفاده شده است.

با توجه به اینکه برای پیاده‌سازی ابزار داده، نیاز به یک پایگاه داده مرتبط با این مبحث داریم، از یک پایگاه داده که مربوط به یک فروشگاه فرضی با بیش از ده هزار تراکنش است، استفاده کرده‌ام که در CD ضمیمه این کتاب اسکرپت‌های مربوط به ساخت و بارگذاری داده‌های آن، قرار داده شده است.

¹ DSS

² Data Warehousing

³ Oracle

در بخش‌هایی از این کتاب از مطالب کتاب‌ها و مقالاتی استفاده شده است که در قسمت منابع، معرفی شده‌اند. امیدوارم که مطالب این کتاب برای شما مفید واقع شود و از این طریق گام کوچکی در جهت پیشرفت علمی خوانندگان عزیز برداشته باشم.

با توجه به اینکه هیچ انسانی کامل و عاری از خطا نیست و من هم از این قاعده مستثنی نیستم، از خوانندگان عزیز تقاضا دارم که نظرات و پیشنهادات خود را در جهت هرچه بهتر شدن کیفیت محتوای این کتاب از طریق ایمیل اینجانب، مطرح سازند. همچنین در صورت وجود هرگونه ابهام یا سؤال در مورد مباحث این کتاب، با من در ارتباط باشید. اینجانب تمام سعی خود را در جهت پاسخگویی به شما عزیزان خواهم کرد.

Email : reza_nakhlband@yahoo.com

reza.nakhlband@gmail.com

Yahoo ID : reza_nakhlband

با آرزوی موفقیت برای تمامی رهروان علم و دانش

با تشکر - رضا نخلبند

تابستان ۹۰

www.ketab.ir