Carol Mitchell
Don’t Store Images In A Transactional Database

Gwen Shapira
Oracle & Hadoop For Fun & Profit

Kellyn Pot’Vin
Training Days Director Perspective

Edward Roske
Essbase From 1992 To 2012

Member Focus - Vincent Giasolli
OBIEE Tips & Techniques - Dan Hotka
Sharon combines her technology background with her Regis degree to help empower her own community. How will you be influential?

The School of Computer & Information Sciences in Regis University’s College for Professional Studies offers four undergraduate and six graduate degrees that immerse students in the latest technologies and software. Students benefit from our state-of-the-art learning environments, experienced faculty and strong industry relationships.

Busy adults choose Regis University for the academic rigor and flexibility.

> Regionally accredited and nationally ranked
> Accelerated 5- and 8-week classes
> Online and campus-based learning
> Multiple start dates throughout the year

Sharon Joan Whitehair
M.S. in Database Technologies
Learn more at:
CPS.Regis.edu/beinfluential

Focused on Delivering the right Solution for the right Project

Arisant focuses on understanding business requirements to ensure the most effective implementation of the right solution. Our strategy is aimed at providing an honest and expert brand of consulting services for both the Private and Public Sectors. Contact us today to discuss any of the following services, and more.

- Consulting Services
  - Architecture
  - Analysis
  - Design
  - Implementation
  - Project Management
- Staffing/Support
- Managed Services
- Managed Hosting Services
- Oracle Software Sales
- Oracle Hardware Sales

Arisant Business Solutions through Oracle Technology

Oracle Platinum Partner

Managed Services
Business Intelligence
Oracle Licensing
Identity & Access Management

Managed Hosting Services
Oracle Cloud 11

303-330-4065
sales@arisant.com
http://arisant.com
contents

features

6 Don’t Store Images In A Transactional Database
   by Carol Mitchell
   Content Management Solutions

10 Integrating Oracle & Hadoop
   by Gwen Shapira
   When The Relational Database
   Is Not Enough

14 RMOUG 2013
   by Kellyn Pot’Vin
   The Conference Director Perspective

20 Essbase
   by Edward Roske
   1992 To 2012

monthly features

4  Letter From Our President
19  Stan Yellott Scholarship Fund
29  RMOUG Board of Directors
30  RMOUG Calendar of Events
31  May 2013 QEW

On the Cover:
   Jeff Stephenson is an application developer who has been working with Oracle databases and PL/SQL since 2005. He is currently working with a great group of people in the Information Systems department at Colorado State University. Away from work, Jeff enjoys exploring and photographing the beautiful expanses of the west with his wife and two daughters. This image of Longs Peak was taken on a sunrise hike last summer up to Dream Lake in Rocky Mountain National Park. This fourteener is a waymark in northern Colorado, majestically towering above the Front Range.
From The President

Tim Gorman

...and another RMOUG Training Days conference enters the history books!

Based on documentation unearthed recently, we found that the conference last month was our 24th consecutive annual event, with the first having occurred in January 1990 at the old US West Training Center off Wadsworth Blvd in Lakewood CO. I have the agenda from the 2nd conference, held in January 1991, and we will post that agenda to our website.

So, that means that next year’s conference, Training Days 2014 at the Colorado Convention Center from Wednesday through Friday, 03-05 February 2014, will mark our Silver Jubilee, our 25th conference! Mark your calendars now, as it will be better and more useful than ever!

But before looking ahead to next year, I hope you’ll join me in thanking the volunteers and friends who made our just-completed Training Days 2013 such a huge success. Later in this newsletter, we’ll be listing all of the volunteers who served in the following capacities...

Organizing Committee

Starting in July 2012, these are the people who actively met, planned, and worked weekly over the eight months leading up to the conference, adapting to changes and making the decisions. Led by RMOUG’s Training Days director Kellyn Pot’vin, here are the organizers of the conference...

• Kellyn Pot’vin
• Peggy King
• John Jeunnette
• Pat Van Buskirk
• Tim Gorman

Each member met at least once weekly, took responsibility for different major parts of the conference, and worked with TeamYCC on the details of the conference. Kellyn in particular was deeply involved in every aspect of the conference, right up to the final minute of the final day.

Abstract Selection Committee

I led this committee, and starting in September 2012, a team of 55 volunteers from all around the world used the web-based abstract-review software provided by John Jeunnette and Prairie Systems to read, review, and score over 300 abstracts submitted by speakers from all over the world. By the end of October, all 300+ abstracts had been reviewed and scored on a scale of 0 (lowest) to 5 (highest) by several different technical reviewers, ensuring that each abstract was carefully vetted to provide the best possible content for the conference attendee.

And then, on a dark and stormy night before Halloween at the Goodson Recreation Center in Centennial, about a dozen of the abstract selection committee met to select an agenda for the conference, based on the reviewed and scored abstracts.

Here is the method we used...

• Based on the total number of abstracts received, determine how many of the 156 time-slots will be occupied by presentations from each of the six topic tracks...
• Leaving aside one entire time period on the first day for “vendor presentations,” which include the “Women In Technology” forum and special new lab for the high-school students from Pine Creek HS in Colorado Springs
  ◦ We ended up with a target total of 143 presentation time-slots to fill
• Based on the proportion of 300+ abstracts submitted for each of the six tracks, we allocated a total number of the 143 time-slots for each track
  ◦ Ranked on review scorings, we chose the “top N” abstracts for each track
    ◦ Limiting each author to no more than two abstracts, we weeded out presentations from authors with more than two abstracts and substituted with abstracts that did not score in the “top N”
  • RMOUG’s goal is to accept as many presenters as possible, balancing the opportunity to speak with quality and innovation
  ◦ We assigned the two biggest tracks (i.e. DBA and App Dev) to four rooms apiece across all time periods.
  ◦ The remaining five rooms were allocated for the remaining four tracks
• We then laid out the chosen abstracts onto twelve 4’x8’ work tables to visually emulate the “Schedule At A Glance” available on the RMOUG website during the conference
  ◦ And then we worked to ensure that certain presenters spoke on certain days to fit their travel schedules, that a presenter wasn’t presenting twice in the same time period, etc.

...and we accomplished all of that in less than four intense hours! Some years, this selection and scheduling process took 6-8 hours, but we’re evidently getting better at it.

After TeamYCC contacted the selected speakers to confirm that they would present, we were able to post the initial agenda before the Thanksgiving holiday. All in all a massive task from start to finish, but with so many volunteers providing their time and expertise, it was just one of many massive tasks comprising this conference.

Room Ambassadors

Pat Van Buskirk and Peggy King led the effort to recruit, train, and manage the schedules of the room ambassadors at the conference. RMOUG has some volunteers who have been ambassadors at Training Days conferences going back into the 1990s, and internationally we have a reputation for the nicest and most experienced
volunteers. Although the job may seem easy, greeting attendees to the rooms, distributing and collecting evaluation forms, there is a deadly serious aspect to the job. In 1996, the RMOUG Training Days conference was shut down by the fire marshal for having too many attendees in a room. We were able to react quickly enough and convince the fire marshal to reverse his decision, but for a few minutes the situation was quite dicey. Room ambassadors ensure that we don’t have a repeat of that catastrophe, and nobody does it better than our volunteers!

**Registration Volunteers**

Pat and Peggy once again led the effort to recruit, train, and schedule our smiling volunteers at the registration booths. Our own “directors of first impressions” are one of the reasons that the conference is so popular, as they check people in, provide them materials, and answer questions.

**Speaker Airport Shuttle**

Starting around 10 years ago, RMOUG realized that the distance from DIA to the convention center represented a $60+ taxi fare each way for our many out-of-town speakers. While many of these out-of-town guests were traveling on company expense accounts, a fair number are self-employed or traveling on their own nickel, so we felt that it would be pleasant to welcome our guests at the airport and drive them to their local accommodations, whether a hotel or a friend’s house. Along the way, volunteer drivers have the time and opportunity to chat with our guest speakers, get to know them personally, and share with them some of the things to do and see around Denver. All in all, it is ultimately pleasant but is often a tedious, time-consuming, and frustrating job, as flight delays turn simple pickups into lengthy airport stakeouts, not to mention mounting parking fees. This year, the conference had a total of 65 out-of-town speakers. Of that total, 37 asked for rides from the airport to the conference. And of those 37 requested rides, we were able to provide rides for 36, leaving only one arriving speaker to shell out for the taxi fare or Super Shuttle. Returning speakers to the airport after the conference is always more difficult, and we were able to accommodate only a small number of the returning speakers.

But the friendliness and willingness to help exemplifies RMOUG in the minds of our out-of-town visitors, and there are quite a large number of highly-respected Oracle professionals around the world who have wonderful memories of Denver.

Good news: the new railway from DIA to Union Station is expected to be completed in 2015, so there is a good chance that next year’s conference will be the last hurrah for the fabled RMOUG airport shuttle.

So, please join me in a huge “thank you” for all of these wonderful volunteers, your colleagues and co-workers.

Also, please join me in thanking TeamYCC, who have helped RMOUG produce this conference over the past 16 years, since 1997. Led by Veronica Snow, and joined by colleagues Kathleen McCasland and Lori Lorusso, the wonderful women of TeamYCC fly out of their Wilmington NC home and help provide the conference with its professional look and feel.

Most of all, congratulations to the incomparable Kellyn Pot’vin, whose inspiring ideas and bold vision of what a world-class conference should comprise, made Training Days 2013 into the benchmark for the rest of the global Oracle community to aim for. Thank you, Kellyn!

And Now, About Training Days 2014...

As I mentioned earlier, next year’s conference will be held once again at the Colorado Convention Center from Wednesday through Friday, 03-05 February 2014. Besides the fact that it will be the 25th annual conference, there are two major things to note about this conference...

**Wednesday through Friday? Not Tuesday through Thursday, or Monday through Wednesday? What is RMOUG up to?**

We are not choosing different sets of dates to confuse anybody. Rather, these were the best dates offered to RMOUG by the Colorado Convention Center and the city of Denver.

So yes, next year the half-day University Sessions and the OTN Labs and other special pre-conference events will be held on Wednesday 03-February 2014, and the main 2-day Training Days conference will be held on Thursday-Friday 04-05 February 2014.

At least we won’t be stomping all over St Valentine’s Day next year, eh?

**How the city of Denver views the RMOUG Training Days conference, and what that means for us**

A few years ago, the city of Denver renovated the Colorado Convention Center and doubled its size. Since then, the city has instructed the management company for the convention center that priority will be given to events which generate more than 200 rooms per night at nearby hotels. In order to maximize the economic impact of the convention center, the city only wants space at the convention center used for events which bring visitors into Denver to spend money on hotels, meals, and entertainment. At present, we believe that RMOUG is generating about 80-100 rooms per night, and we are working to substantiate that as fact.

So, about two years ago, RMOUG received an email from the management company describing this policy and encouraging RMOUG to see this policy as an opportunity to “explore other event venues in the Denver metro area” for our Training Days conference. What this means for RMOUG is that, starting last year, we are no longer permitted to book the Colorado Convention Center 18 months in advance, as we have been accustomed to doing since 1999. Instead, if the venue still has space 12 months prior to our event, then they will book us. That is what happened for the just-completed Training Days 2013, and that is what happened with the upcoming Training Days 2014. And that is why both events were not on the same sets of days that the prior dozen conferences were on. We’ve gotten lucky the past two years to get in at all.

We cannot assume that Training Days 2015 will be at the Colorado Convention Center.

RMOUG is indeed exploring other venues in the city, and around the city. We are also considering ways to push ourselves larger, so that the city of Denver once again wants our business. The board of directors of RMOUG would very much like to hear which approach you prefer -- moving to another venue or staying at the convention center by increasing the number of room-nights associated with the conference.

As winter yields to spring, the cycle will be starting all over again. Please become a more-active member of this wonderful organization by volunteering to be an “at-large” member of the board of directors. An online application form can be found at “http://www.rmou.org/about/current-board-of-directors/board-application/”.

We look forward to working with you!
Don’t Store Images In A Transactional Database

by Carol Mitchell

Users are much more efficient when they have direct access to documents related to their Oracle EBS transactions, entities, etc. This has been proven by many companies for different business purposes. So the electronic file attachment functionality in Oracle E-Business Suite applications is great! Or is it?

Calling all DBA’s! Do you have megabytes, gigabytes, or even terabytes of unstructured data stored within your Oracle EBS transactional database due to users attaching electronic files using the attachment functionality? The users need this data in order to have the right information at the right time. Typically you will see users storing invoices, vendor contracts, packing slips, inventory documents, GL journal entry supporting documentation, HR employee information, just to mention a few, as this helps them become more efficient.

We all know that processing power is very cheap these days, and disk space is even cheaper. But throwing hardware at the problem of storing too much unstructured content in a structured database is not the right long-term solution.

This article will cover the benefits of using file stores managed by external Content Management software with integration to the content instead of storing the unstructured data in the Oracle E-Business Suite transactional table.

Why Not Store Content in a Database?

Most DBA’s agree that any object larger than one megabyte should only be referenced in a database pointing to a file store. Of course, everyone has differing opinions on the subject.

But when it comes to the Oracle E-Business Suite applications and what they are being used for, storing large files in the database results in performance issues, fragmentation, and longer backups. From an overall company viewpoint the downsides are worse -- storing content within a database results in difficulty to securely share across organizations, lack of document versioning capabilities, inability to control the retention of the data, inability to use the data to automate workflows, and other company business uses.

Content Management Solutions

Use of content management applications for securing and sharing unstructured content is considered a best practice solution, especially with the growth of paper and electronic files being generated every day. Per AIIM’s (The Global Community of Information Professionals) President, John Mancini in “The Big Data Balancing Act” white paper, 2013. “As I think about this question of harnessing our information assets and optimizing them to our competitive advantage, it has never been more important to balance the needs of technology with the needs of the business.”

Content management applications store the content efficiently and securely in a file store with only the search (index) data stored within a database. With content management, paper and electronic files can be captured and stored with security capabilities at the document level if needed.

Content can be searched based on any “metadata” (property, index) field defined for the type of document. Taxonomy structures can be created to store the content in a logical way for navigation, discovery, and find-ability. It is critical for the success of businesses.

And the next step beyond simple content management is Enterprise Content Management (ECM) that provides much more functionality than just “storing the data.” Many ECM applications include records management with email management to meet compliance and e-Discovery requirements, Business Process Management (BPM or Workflow) to support automation of business processes with associated content, web content management, social content management, analytics, and other creative uses for the data.

Oracle EBS Integration with Content Management

Storing your unstructured content in a content management application can still allow the users to quickly access their content from within the Oracle E-Business Suite applications by storing direct links within the applications to the content in the file store. A lot of vendor companies provide integration functionality for Oracle EBS applications by modifying the Oracle form or web page to support manually linking the document, or by screen scraping and doing a search, or even by storing the URL link in the attachment table. Different solutions offer different amounts of control, functionality, and required development efforts.

The use of a content management application and an integration solution solves the concerns over storing content in the Oracle EBS database. It also addresses the issue of document versioning control, and allows secure sharing of content across the organization for multiple purposes. Among the integration solutions out there, questions still remain:

1. Is there a way to automate the link creation?
2. What if there is an exception? We need to ensure compliance.
3. How can we do this integration without custom development?
4. If we do custom development, who will support it?
5. How can content be accessed across applications?
6. What about the metadata differences due to transaction changes?
7. What about the security of documents available from Oracle EBS?
8. How can documents already stored in the database be pulled out and put into a content management application with appropriate metadata, but still allow the users access?
9. Do the users have to print their documents, scan them and then have them go into the content management application rather than attach the files electronically?
10. What happens when we need to do an upgrade to Oracle or the document management application?

These are some of the questions that are asked when this approach to Oracle EBS related document storage is proposed. There is no simple answer to these questions. Some companies will accept different combinations of these issues, but not others.

As companies have growing needs to manage their content, and they see the benefits of having their content available at the right time and the right place, the above-mentioned issues need to be addressed.

One option available to address these issues, which is a solution designed exclusively for the Oracle E-Business Suite, is a CMMC product called ‘DocSavi.’ While this is NOT the only solution available, the below information is presented as it may be a beneficial approach for your specific company needs.

**DocSavi™ (Patent Pending)**

CM Mitchell Consulting (CMMC) is a Denver based Oracle/IBM/Kofax software and professional services company helping Oracle clients since 1992. To provide a solution for the above issues, CMMC has developed a product which provides you with all of the functionality that you need to ‘clean up’ the database by moving the content to ensure it is stored in an efficient, secure, content management application going forward.

DocSavi™ is a configurable, integrated and fully supported product, not a development tool. It allows content to be mapped from your transactional database to your content management application through a friendly web-based (Web Logic, Webshere, JBoss) Java user interface.

This interface allows the system to be configured, monitored, and any errors to be identified and easily resolved.

If a document is scanned or an electronic file is ‘saved’ to a content management application that is mapped to an Oracle EBS table within DocSavi, the document location is sent through Oracle’s advanced queuing to create a record within the standard Oracle attachment table containing a direct connection to the image within the content management application. And unlike other integration solutions, this integration is done automatically through a web service with full exception monitoring, reporting and correction capabilities.

By utilizing the attachment functionality inherent in the Oracle E-Business Suite applications, no modifications need to be made to the software, only configuration. With so many users wanting to use this capability for iExpenses, iSupplier, AP, and HR, the solution uses the standard Approval Management Engine configuration to address any approval functionality needed. And this functionality can be used across all of the applications to support having the right data at the right time.

In order to ensure upgrades do not affect the functionality of the solution, the DocSavi engine is stored in a separate schema on the Oracle database.

To ensure the security of the documents, validation of the user’s access rights to the documents is controlled by the content management application.

And, to ensure that the metadata within the content management application is accurate, a synchronization process is available to update the metadata if the corresponding field within Oracle EBS is updated.

Using this DocSavi integration product also allows the users to continue to ‘save’ their electronic files while working within their Oracle E-Business Suite transactions or entity records, and an automated web service will upload this content back into the content management system, along with the search (index) criteria for storage. And once the location is sent onto the Oracle tables for retrieval purposes, the file is removed from the transactional database. This functionality can be used to do a one-time migration of documents from the Oracle database to a content management solution.

**Process Flow**

Using content management along with the Oracle E-Business Suite applications allows paper, electronic documents, email, etc.
to be organized utilizing metadata (tags, indexes) to fully support searches within a single document type with a certain property value, or across multiple document types, or across multiple content repositories. Content can be secured based on business roles and access needs, and versioning control can be recorded to help the business meet regulatory compliance requirements.

**DocSavi Product Differentiators**

Some of the major differentiators of the DocSavi solution:

- Configured using a Web based UI versus development
- Does not modify the Oracle form or web page
- Direct access to the document versus a search
- Supports all tables within the Oracle EBS suite of applications
- Uses standard Oracle attachment functionality for access from within Oracle EBS
- Can access documents across applications
- Supports upload of documents stored electronically in the Oracle attachment table into FileNet P8, leaving a URL link behind for access
- Uses Oracle standard Approval Management Workflow setup and rules versus needing development efforts to define approval processes thru-out the applications
- Synchronizes the metadata between the two systems when fields are changed within Oracle to ensure metadata consistency
- Uses FileNet P8 security for access regardless of Oracle responsibility
- Fully supported product
- Reduced cost to deploy and maintain

DocSavi is an Oracle Validated Fusion Middleware Integration for the Oracle E-Business Suite. As described by Oracle, “partners with validated integrations are able to provide customers with standards-based product integrations, tested and validated by Oracle. Customers benefit from improved risk management and smoother upgrade capability, leading to a lower cost of ownership and greater overall satisfaction.” DocSavi is also an IBM verified solution.

**Conclusion**

The biggest benefit of storing unstructured documents in an Enterprise Content Management application rather than a transactional table is that typically it is just that – for the enterprise. Not only can an ECM application provide you with a secure location to store needed documents for Oracle E-Business suite processing, but it can be used across the entire organization to store other critical company documents. This type of application can allow companies to organize all of their unstructured data so that it becomes useful data that can help make the company more efficient, more compliant, assist in decision making, and best of all, make e-discovery a simple process.

But, alas, this is a topic for another day . . . .

Carol Mitchell, CEO
CM Mitchell Consulting Corporation
www.cm-mitchell.com
303-526-2796

Carol has over 20 years of experience in the information systems industry. Her career spans the areas of project management and implementation of integrated product and software solutions, Oracle E-Business Suite applications, and Enterprise Content Management software.

Carol is a Certified Information Professional, a Certified Document Imaging Architect and a Certified Information Systems Auditor. She has been an IBM Information Champion for the last four years. Carol has served as a chapter president for the American Business Women’s Association, the Information System Auditor’s Association, the Oracle Application Users Group, and the Rocky Mountain Chapter of the Association of Image and Information Management (AIIM). Carol actively participates in the various users groups conducting speaking and training sessions within the information management industry.

Carol Mitchell is president and founder of CM Mitchell Consulting Corporation located in Littleton, CO. CM Mitchell Consulting is an advanced business partner and authorized reseller of IBM FileNet P8 Suite and IBM Datacap Capture. CMMC is an Oracle Gold partner and reseller of Oracle WebCenter. CMMC is also a partner with Kofax, AnyDoc and Fujitsu for capture and scanning products. The company provides professional services, software integrations and document management life-cycle solutions.

---

**Become A Member For These Outstanding Benefits**

- Quarterly Education Workshops
- Special Annual Training Days Rates
- Database Labs
- List Server Subscriptions
- SQL>Update Magazine

www.rmooug.org/member.htm
I have been working with Oracle’s BI tools for years. I am quite the Discoverer expert (a free tool now from Oracle Corp...OBIEE ...standard edition) but this tool is now on the back burner and I hardly even field questions on it anymore.

OBIEE is the replacement product in Oracle’s BI offering. This tool originated from Oracle’s purchase of Siebel and Oracle has enhanced it, put it in with weblogic and SOA, and spruced up the end user piece.

I do training on both the admin/repository side and the end user reports/dashboards side.

This series of articles will show you some of the newer features of the tool. This article will discuss how to get started, or, come up with a good development environment where all these pieces are already installed, etc.

Remember OTN? Oracle Technical Network? This site is your friend. Oracle has produced a rather large VMBox image that has all of the BI tools, including BI Publisher, already installed and working.

The Oracle BI Sample is available for download from from OTN at

The Sample v107 is Oracle BI 11.1.1.5 and the Sample v207 is the Oracle BI 11.1.1.6. This download is a complete working OBIEE environment for VMBox:

This download is HUGE: 26GB.

I have purchased quite the PC to run this image. It takes a lot of horsepower.

Toshiba P750, Quad Core processors running at 2.20 GHz.
  •  64 Bit Windows 7 Professional
  •  8gb RAM, 500GB (fast seek time) Hard Drive

Bring the image up in the VMBox environment...open VMBox...then open the Sample v207 environment.

There should be an Admin, End User, and another account.
I usually pick the Admin login. ALL passwords are ‘Admin123’. IF these accounts do not exist, email me for the setup and start scripts. The Deployment Guide has instructions for starting the various services.

I have been using the ‘weblogic’ user id and ‘Admin123’ passwords to work the examples and illustrations also found under the documentation tab/tutorial tabs where the Sample v207 image was downloaded.

I found the Deployment Guide very helpful with startup tasks and accounts/passwords. I have been working with the User Guide for the OBIEE product. I have found many of the tutorials helpful as well.

Once you have the environment running...open a web browser (I prefer Google Chrome for Oracle things...). follow these steps:
  •  Do an IPCfg command from a DOS prompt on the OBIEE computer. This is the IP address used next
  •  http://<ip address from above step>/7001/analytics
    o login: weblogic
    o password: Admin123

This information should get you started. The next two articles will illustrate building reports, reports with prompts then how to add these articles to dashboards...a cool environment for running your OBIEE queries and reports.

Let me know how I can help you with your OBIEE implementation.

Dan

Dan Hotka is a Training Specialist and an Oracle ACE Director who has over 35 years in the computer industry, over 29 years of experience with Oracle products. His experience with the Oracle RDBMS dates back to the Oracle V4.0 days. Dan enjoys sharing his knowledge of the Oracle RDBMS. Dan is well-published with 12 Oracle books and well over 200 printed articles. He is frequently published in Oracle trade journals, regularly blogs, and speaks at Oracle conferences and user groups around the world.

Dan Hotka - Author/Instructor/Oracle ACE Director
www.DanHotka.com
Dan@DanHotka.com
DanHotka.Blogspot.com
515 771-3935
Modern data warehouses face challenges that are not easily resolved by traditional relational databases:

1. Storing and processing unstructured data – images, video, PDFs.
2. Storing and processing large bodies of text, especially when the requirement includes natural language processing. The text can be email, contracts, j
3. Storing and processing large amounts of semi-structured data – XML, JSON and log files.

Traditional RDBMS data warehouses have limitations that make them less than effective for those new data types that are now required by the organization. They can be stored and processed in a relational database – but there is little benefit to doing so, and the cost can be high.

There are also some types of jobs where a traditional data warehouse is not the optimal solution – ad-hoc analytical queries can be a bad fit for the existing schema and cause performance issues that will wreak havoc in the carefully scheduled batch workloads running in the data warehouse.

Advanced machine-learning algorithms are typically not implemented in SQL, which means that the developers will need to read massive amounts of data from the data warehouse and process it in the application.

Cleanup and transformation of data from different sources require large amounts of processing. While ETL jobs often run in the data warehouse, many organizations prefer to move these jobs to a different system that is more suitable to the task, rather than invest in larger data warehouse servers and the associated database license costs.

What is Hadoop?

Hadoop is a platform designed to use inexpensive and unreliable hardware to build a massively parallel and highly scalable data-processing cluster.

It is designed to be a cost-effective and scalable way to store and process large amounts of unstructured and semi-structured data. This type of data processing poses some technical challenges – Clearly large amounts of cheap storage are required to store large amounts of data, but large disks are not enough. High-throughput access to the data is required to allow timely processing of the data. In traditional storage systems throughput did not keep up with the increases in storage space. On top of the storage system, large number of processors is required to process the data, and applications that are capable of utilizing large number of processors in parallel. And of course we also want the system to integrate cleanly with our existing infrastructure and have nice selection of BI and data analysis tools.

Hadoop was designed to address all these challenges based on two simple principles:
1. Bring code to data
2. Share nothing

The first principle is familiar to all DBAs – forever we plead the developers to avoid reading large amounts of data for processing in their Java apps. When the amounts of data are large this is slow and can clog the network. Instead, package all the processing in a single SQL statement or PL/SQL procedure and perform all the processing where the data is. Hadoop developers submit jobs to the cluster that stores the data.

The second principle is due to the difficulty of concurrent processing. When data is shared between multiple jobs, there has to be locking and coordination in place. This makes development more difficult, imposes overheads and reduces performance. Hadoop jobs are split into multiple tasks; each task processes small part of the data and ignores the rest. No data is shared between tasks.

Hadoop is made of two components: HDFS, a distributed and replicated file system, and Map-Reduce, an API that simplifies distributed data processing. This minimalistic design – just a file system and job scheduler allows Hadoop to complement the relational database. Instead of creating a schema and loading structured data into it, Hadoop lets you load any file onto HDFS and to use Map-Reduce to process and structure the data later.

HDFS provides redundant and reliable storage for massive amounts of data, based on local disks of all nodes in the Hadoop cluster. File sizes are typically very large, and to reflect that, Hadoop’s default block size is 64MB (compare with Oracle’s 8KB?). Sustained high throughput is given priority over low latency, to maximize throughput of large scans. HDFS provides reliability by copying each block to at least three different servers in the cluster. The replication doesn’t just provide failover in case a node crashes; it also allows multiple jobs to process the same data in parallel on different servers.

Map-reduce is a method to distribute processing jobs across the servers. Jobs are split into small, mostly independent tasks. Each task is responsible for processing data in one block, and whenever possible it will run on a server that stores that block locally. The design maximizes parallelism by eliminating locks and latches.

As the name suggests, map-reduce has two phases: map and reduce. Map tasks filter and modify the data. This is analogous to
the “where” portion of a query and to non-aggregating functions applied to the data. The reduce phase applies the data aggregation: group by and aggregating functions such as sum and average.

Since map-reduce jobs are limited to filtering and aggregating, more complex analytical queries do not translate well to map-reduce and are therefore difficult to implement in Hadoop. However, since Map-Reduce jobs are just Java code and can import existing Java libraries, Map-Reduce offers flexibility that does not exist in pure SQL. For example, OCR libraries can be used to allow Map-Reduce jobs to extract data from image files.

Don’t go imagining that Hadoop is the ultimate silver bullet! Hadoop has several well-known limitations – At this point in time there are not many enterprise products for Hadoop and this means that building solutions on top of Hadoop is likely to require more resources than the building the same solution on top of Oracle. Hadoop is best suited for batch processing, and low-latency real-time analysis is still in beta stages. In addition Hadoop can be tricky to debug, tune and troubleshoot – it is not as clearly documented and instrumented as Oracle database is.

One of the design goals of Hadoop is to allow loading data first and structuring it later. One of the main ways to provide structure for the semi-structured files stored in Hadoop is by using Hive - a data warehouse system on top of Hadoop that allows defining a schema for files in HDFS and then querying the data in the tables using a SQL-like language which is translated automatically into Map-Reduce jobs. This greatly simplifies the task of data processing in Hadoop and means that learning Java and writing Map-Reduce jobs from scratch is not mandatory.

## Common Enterprise Use Cases

### ETL for semi-structured data

This is probably the most common use-case for Hadoop that I run into.

There is a system somewhere that is generating log files or XMLs, and these files include data that needs to get into the data warehouse. Of course, in order to extract the data and process it so it will fit into the star schema, some processing is required. It can be done on the relational database, but why waste the most expensive CPU cycles in the enterprise on text processing?

Instead, our customers use Flume to load log data into Hadoop, write Hadoop Map-Reduce jobs or use Hive to process the data and then use one of the techniques we’ll discuss later to load the data into a relational database.

### ETL for structured data

This looks like a strange use-case at first. The data is already structured, why bother with Hadoop? But anyone who dealt with ETL jobs knows that one person’s structured data is another’s messy data source. OLTP data requires a lot of aggregation and transformation to fit into a data warehouse. Traditionally we did much of this work in the data warehouse itself – but data warehouses can be expensive and difficult to scale. As the amounts of data grow, so does the amount of time it takes run the ETL jobs.

Instead, load the data into Hadoop and add enough servers to make sure the job is split into small manageable chunks and finishes within the SLA.

### Historical reports

Normally, searching for data within a relational database is easy enough. The only problem is when the data is not in the relational database. Since relational databases typically use expensive storage and since performance typically degrades as the amounts of data grow, the amount of data stored in a relational database is typically limited.

Many online stores have been in business since the late 1990s. But how many of those businesses have detailed purchasing information for nearly 20 years of data? Marketing departments have learned not to ask for too-detailed information from the past, because it is unlikely the data was stored.

But with Hadoop using cheap storage, there is no reason not to store all the data your organization may need. And Hadoop clusters are not tape storage – they can process the data. So marketing departments can look at long-term trends and IT departments no longer throw away potentially valuable data.

### Needle in Hay Stack

Network equipment can generate huge amounts of logs. No one would dream of loading all the data into a relational database. But sometimes this data has to be searched.

For example when it is determined that there was a security breach, many terabytes of log file data has to be scanned for few lines indicating the source of the breach.

The solution is simple – just run grep. And Hadoop was built to run “grep” fast and in parallel on huge amounts of data. It is trivial to implement too.

### Search unstructured data

I work for Pythian, a remote DBA company. Naturally, many of our customers feel a bit anxious when they let people they haven’t
even met into their most critical databases. One of the ways Pythian deals with this problem is by continuously recording the screen of the VM that our DBAs use to connect to customer environments. We give our customers access to those videos and can replay them to check what the DBAs were doing. However, watching hours of DBAs just doing their job can be boring. We want to be able to find the interesting moments – catch the DBA just as he’s running “drop table” or when he solves a difficult issue.

To provide this capability, Pythian streams the recordings to our Hadoop cluster. There we run OCR in parallel on each screen-shot captured in the recording and isolate the text on the screen and the DBA keystrokes. Other jobs mine the data for suspicious key words (“drop table”), patterns (credit card numbers) and root access. This system lets our customers verify their databases are safe without excessive work on our or their part.

Connecting the Big Dot

Hopefully by now you see how Hadoop can complement a traditional data warehouse and how you can use Hadoop to process more data types or just more data.

But clearly for all this wonderful stuff to work in practice, we need a way to get data from a relational database into Hadoop, and more important – from Hadoop into a relational database.

Let’s look into techniques for doing just that.

Sqoop

Sqoop is an Apache project designed to transfer data between Hadoop and relational databases.

Its main modules are sqoop-import, which transfers data from a relational database into Hadoop and sqoop-export, which transfers data from Hadoop to relational databases.

Both modules work in similar manner. Sqoop starts multiple map-reduce jobs which connect to the database and run queries in parallel. Sqoop-import runs select statements and writes the results on HDFS, sqoop-export reads data from HDFS and runs DML statements to insert the data into database tables.

The main argument to sqoop-import is --table, which allows the developer to specify which database table to import into Hadoop. Sqoop-import also allows specifying exactly which columns to import and a where clause to limit the rows imported.

For example:

Sqoop import --connect jdbc:oracle:thin:@//dbserver:1521/masterdb
--username hr --table emp
--where "start_date > '01-01-2012'"

Will import into Hadoop data about employees who started work at the beginning of last year.

By default sqoop-import runs 4 map processes to read the data from the database, but the number of map processes can be modified in a parameter. To split the data between the map processes, sqoop-import looks for the primary key, checks the minimum and maximum values and divides that range between the map processes.

Suppose our emp table has primary key of (id), the range of ids in our table to 1 to 12 and we are running two map processes.

In this case one map process will run the query:

```
SELECT * FROM emp WHERE start_date > '01-01-2012'
and id >= 1 and id < 6
```

and the other map process will run:

```
SELECT * FROM emp WHERE start_date > '01-01-2012'
and id >= 6 and id < 13
```

You can specify a column other than primary key to split the work between map processes:

```
Sqoop import --query SELECT a.*, b.* FROM a
JOIN b on (a.id == b.id) WHERE a.type='PRIME' and
$CONDITIONS' --split-by a.id --target-dir /user/
foo/joinresults
```

In this example we decided to run 16 map processes and split the work by customer_id, even though the primary key is shop_id.

It is critical to remember that the split-by column should be either indexed or a partition key, otherwise our example will result in 16 parallel full table scans on the shops table, which is likely to cause severe performance issues.

Sqoop-import can be even more flexible and allow you to import the results of any query:

```
sqoop import --query 'SELECT a.*, b.* FROM a
JOIN b on (a.id == b.id) WHERE a.type='PRIME' and
$CONDITIONS' --split-by a.id --target-dir /user/
foo/joinresults
```

In this case $CONDITIONS is a place-holder for the where-clause that sqoop-import adds for splitting work between the map processes. Since the query is free-form it is our responsibility to figure out the split column and the right place to add the condition in the query.

Sqoop-export is not quite as flexible as sqoop-import and it is important to be aware of its limitations. Sqoop-export works by turning each input line into an insert statement and commits 1000 statements. This means that failure of sqoop-export job will leave the database in an unknown state with partial data inserted. This can cause re-tries of the export job to fail due to collisions or to leave duplicate data - depending on the constraints on the table.

By default sqoop-export fails if any insert statement fails. This can become a problem if there are unique indexes and existing data in the target table. Sqoop-export can be configured to re-try unique-constrain violations as updates, effectively merging the data into the target table.

In general sqoop-export is not as flexible and configurable as sqoop-import. Developers can specify source files on HDFS, target tables, number of map processes that perform the inserts and the update behavior:

```
sqoop export
--connect jdbc:oracle:thin:@//dbserver:1521/masterdb
--table bar
--export-dir /results/bar_data
```

Due to those limitations, sqoop-export is rarely used for databases other than MySQL (sqoop-export has special “direct mode” for MySQL which overcomes many of these issues).

Fuse-DFS

Oracle highly recommends using external tables to load data into a data warehouse (http://www.oracle.com/technetwork/database/bi-datawarehousing/twp-dw-best-practices-for-implem-192694.pdf) and with good reasons:

- External tables allow transparent parallelization of the data access
- They allow you to avoid staging tables. Instead you use
SQL to read data from the source files, transform it and place it in target tables, all in one command.

- External tables allow parallel direct path writes, farther improving performance

So when I had to load large amounts of data from Hadoop into an Oracle Data Warehouse, naturally I looked at ways I could use external table. With parallel reads and writes, direct path writes and one-step ETL - it is guaranteed to beat Sqoop-export performance, without having to worry about partial data load.

If Oracle could access files on HDFS, this would be no problem at all. As you’ll see soon, it can do just that using Oracle’s Hadoop Connectors. But those connectors have significant license fees, which some of my customers were unwilling to invest. By default, Oracle external tables can only access files on the server’s file system. Which means we need to mount HDFS as a posix-like file system. This is done with Fuse-DFS, which is relatively easy to install and configure:

```
sudo yum install hadoop-0.20-fuse
hadoop-fuse-dfs dfs://<name_node_hostname>:<namenode_port> <mount_point>
```

Now that every user can see the HDFS files, it is easy enough to create an external table to read them. The catch is that files on Hadoop are not always plain text, and you may need to add a preprocessor to read them. You can read a detailed description of how to do that on the Cloudera blog: http://blog.cloudera.com/blog/2012/09/exploring-compression-for-hadoop-one-dbas-story/

The important benefit of using fuse-dfs and external tables is that it allows you to use standard Oracle tools and leverage all your hard-won experience as an Oracle tuning expert to squeeze every last bit of performance out of the data load process. Sqoop-export does not give you the flexibility to do this.

**Oracle Loader for Hadoop**

Oracle Loader for Hadoop is a high-performance utility used to load data from Hadoop into Oracle database.

Oracle Loader for Hadoop runs as a Map-Reduce job on the Hadoop cluster, shifting the processing work to Hadoop and reducing load on the Oracle database server.

The Map-Reduce job partitions the data, sorts it and converts the data into Oracle Database file formats before loading the data into Oracle database using direct write path. All this pre-processing is done on the Hadoop cluster, close to the data origins and where processing power is easier to scale. Oracle database only has to place prepared data blocks into data files.

Loading pre-sorted data into Oracle tables means that index creation will be faster and require less IO and CPU. Compression, especially HCC will also be faster, take less CPU and result in higher compression ratios than when compressing unsorted data.

In version 2 of the connector, Oracle added support for Avro file type, for Hadoop compression, and for loading data from Hive tables.

If part of your ETL process includes frequent data loads from Hadoop to Oracle, the performance benefits of using Oracle Loader for Hadoop are difficult to ignore. The main drawback is that it is not open-source and requires a license to run.

**Oracle SQL Connector for Hadoop**

In previous version, this connector was called Oracle Direct Connector for HDFS, and provided a pre-processor for creating an external table from files residing in HDFS. This connector was benchmarked by Oracle and shown to be about 5 times faster than using Fuse-DFS for the external tables.

In version 2, the connector was re-written with new interface and is now more powerful and easier to use. It runs as Map-Reduce job, and automatically creates the external table using either data in the Hive data dictionary or by assuming that all columns in a delimited text file are varchar2 type. Just like Oracle Loader for Hadoop, the SQL Connector also supports Avro file types and compressed data.

Once the external table exists, it can be used for ETL the same way any external table can and the connector transparently handles parallelism.

**Closing notes and tips for the aspiring Hadooper**

- It is important to remember that while Hadoop offers exciting new possibilities, Oracle database is a powerful and well-understood platform. I always advise customers to first make sure they are using Oracle correctly before venturing out to new platforms. Are you using external tables for your ETL? Efficient direct path writes? Is your data partitioned correctly? Are you using partition-wise joins and star transformations? Moving an ETL process to Hadoop is far more challenging than making sure the existing process is optimal, so start with data warehouse 101 tuning.

- At the time of writing, Hadoop is still best optimized for batch processing and jobs. Real-time ability to query Hadoop is still in beta and even simple Map-Reduce jobs take few seconds to process. If real-time analytics is part of the requirements, I’d wait before adopting Hadoop.

- As always before embarking on a new data warehouse project, make sure you have clear requirements, goals and deliverables. Make sure Hadoop’s capabilities and limitations make sense in the context of those requirements and goals. It is easy to get excited about adopting new technology while losing the overall picture.

- Have realistic expectations - Hadoop is relatively new technology. It is not as mature as Oracle and can be much more challenging to deploy, tune and troubleshoot. Projects will take longer than you are used to and that unexpected snags and bugs will show up.

- When things go wrong, don’t forget the basics. Hadoop is different from relational databases in many aspects - but many skills and tasks apply: If things are slow, find the bottleneck. Use operating system tools to find out what is slow - storage, network, CPUs? Expect performance problems as data and workload grows. Plan to grow the cluster accordingly.

- Make sure you are using the right tool for the job - structured data, real-time reports, BI integration, frequent updates and OLTP-like workloads belong in Oracle data warehouse. Unstructured and semi-structured data, large bodies of text and data whose structure can change frequently without notice belong in Hadoop.

- Because Hadoop cluster can be created from any combination of servers, there is no excuse not to have a “toy cluster” to try new ideas. Perhaps your first Hadoop cluster is a server from QA that no one uses and two laptops - its enough to get started and explore whether Hadoop can add any value to your data warehouse, and no one will stop the project due to high costs.

- It’s a new world out there, have fun exploring.

Gwen Shapira is a senior Oracle DBA at The Pythian Group and a member of the OakTable network. She contributes to the Oracle user community by speaking in Oracle conferences and writing in her blog. She also volunteers at the North California Oracle User Group (NoCOUG) where she is chair of the conference program.

---

**SQL>UPDATE** • Spring 2013 13
With RMOUG Training Days 2013 behind us, I now get to look back at my first year as a conference director. The planning and work, both daunting and challenging, focused much of my time for the last eight months. The conference already is a phenomenally successful event, so I knew I needed to make sure not to touch anything that was working well and only enhance what was already there. Much of the planning for abstract selection, scheduling, volunteers, etc. is carefully coordinated and understood by many at RMOUG. Outside of these main duties as Training Days Conference director, I felt there were only three new challenges that I needed to take on:

1. Introduce the Guidebook Mobile Application. I’d utilized it at Denver’s SQLPass conference and was incredibly impressed with what was by far, the best mobile event application I’d ever seen.
2. Rework the Volunteer/Speaker Reception. Attendance had declined and it needed just a little revamping to get it going again.
3. Introduce a WIT session and see if we could get an RMOUG WIT program going.

The first challenge seemed to be an impossible challenge at first. The Guidebook app, for the version and features we wanted, was going to cost a good amount of money. I wasn’t sure we were going to be able to accomplish this, as I was already starting to find ways to cut cost due to normal cost increases for food and such. I was bound and determined we try to offer the conference for the same incredible price we had the last couple of years and I didn’t want the mobile app to impact it.

I was very thankful when Arisant stepped up and offered to sponsor the Guidebook mobile application and we were on our way. The guidebook for the mobile application must be built, which does take considerable man-hours. I don’t know what I would have done without Team YCC’s Lori Lorusso stepping in and helping us out with this feat. Anyone who downloaded and utilized the app, (over 80% of our attendees) were offered a great interface to view the schedule, speaker and keynote information, to create their own schedule and access to social media.

With the Mobile Application out of the way, we were onto another pet project of mine, the Speaker/Volunteer Reception. This had historically been a successful event, with only a few complaints, so I knew I just wanted to tweak it a little. We decided to move to the larger venue and after inspecting a number of places through-out downtown Denver, we chose the Rock Bottom Brewery, where we now, not only had a private bar area, but also TV’s, pool tables and an extended menu. Vikki Lira assisted in making this event memorable by joining the ACE dinner, (which is almost impossible to plan during such a short event...) to ours, creating the new VIP Reception. Needless to say, the event was a great success, a larger turnout then we’ve ever had in the past and I received very positive feedback from our guests.

Two down and one to go! With all the actual planning for the session selections made, knowing we had great speakers, great content and great attendees in the way of our Rocky Mountain Oracle User group, I was now onto my last challenge. I had long wondered why we experienced a less than similar ratio of women in attendance at the Training Days conference as I knew were in the workplace. I was then startled upon presenting at the Denver SQLPass conference and impressed by the female attendance of at least 40%. Fellow presenter, Jen McCown was there as well and she is responsible for a number of the WIT, (Women in Technology) events for the MSSQL side. She shared with me her experience on what made a successful program and I took this to heart.

There were three rules that I needed to follow very closely:

1. Do not make it a “cake social” or “book club” kind of event. If you wouldn’t offer it to men in attendance, don’t offer it to the women. If you do, you will lose the interest of the true women technologists.
2. Do not be afraid to discuss all areas that are challenges for women in the workplace and keeping young women from entering database career fields. At the same time, if anyone starts to “male-bash”, this must be quelled immediately. Thank the attendee for their opinion, but explain that this is not the right forum for it.
3. Have men involved - at least one on a panel and invite men to attend. Their contribution and their points of view aren’t just valuable, they learn as well from the sessions.

I started by emailing who I thought would be excellent panelists for the session. I was thrilled with the exchange of emails between Gwen Shapira, Alex Gorbachev, Karen Morton, Maria Colgan and Debra Lilley. Many of them had thought about this subject and had incredible value to offer to the session. Gwen had to pull out of the conference due to work demands, then Alex, just a couple days before the conference was alarmed to find out that there had been some confusion with his travel visa, so he was unable to...
attend. I immediately thought of another friend and colleague that was another excellent choice and so Dominic Delmolino stepped into Alex’s shoes and was a solid, masculine voice in our incredibly impressive panel of women! The attendees to this session, both male and female, did a fantastic job of contributing, offering their own insights and experience to the panel’s view. Post the session, I was blown over by the amount of attendees that came to me to talk further about WIT, their ideas and their hopes for the WIT program. I’ve since reached out to our members, planning what will be some evening sessions and a QEW, (quarterly education workshop) for RMOUG’s very own WIT program.

Outside of my duties at conference director, serving on committees, picking up folks from the airport and all the other chaos, I made sure to only allow one of my sessions to be put into the RMOUG session schedule- “EM CLI, the Enterprise Manager Command Line Interface”. On the day before, due to a last minute cancellation, I had them pencil in my “Sherlock Holmes for the DBA”. Both went extremely well, considering that I was just a *bit* distracted…😊

Some of the conference highlights from my perspective?

- Having folks like Frits Hoogland, Pete Scott, Debra Lilley and Martin Widlake coming from Europe to present and Andrey Goryunov attending, coming all the way from Australia.
- Carlos Sierra approaching me via email and wanting to be part of RMOUG, then he and Frits jumping in with a couple sessions to fill in for last minute cancellations due to storms and illness.
- Maria Colgan approaching me at KSCOPE and asking to do more at RMOUG and so she is now part of OTN’s workshops.
- Jeff Smith approaching me about having a special session involving the Pine Creek Elementary Students and so the Oracle Vendor Showcase became an Oracle Careers Session.
- The RMOUG speaker gift blankets. I hope everyone gets as much use out of them as I hope they will.
- Having Mogens Norgaard as our keynote speaker.
- Having my company Enkitec put so much into this conference and so many of my coworkers submit abstracts in support.
- The WIT session, not just my wonderful co-panelists, but even more so when I received feedback and realized how successful we’d been!
- Receiving the IOUG Volunteer of the Year award for 2013.
- Everyone who sent me wonderful emails after the conference, letting me know how much they enjoyed Training Days, the WIT session, the mobile app or to just tell me how much they enjoyed one of my technical sessions.

I want to thank all those that not only sent me thanks for the wonderful conference, but also offered me feedback so this next year we can improve even more. I truly can’t believe how much we offer, between great speakers, great content, people and venue at the price we do. We couldn’t do all of it without the continued support of our speakers, exhibitors, sponsors, board and most of all, our members!

Congratulations to Kellyn Pot’Vin as recipient of the 2013 Volunteer of the Year - A much deserved award!!
The more detailed breakdowns of Measures (Sales, COGS, Margin, etc.) are necessary to understand the overall performance. Second, notice that the values in the second report are identical to those in the first report. Likewise, Budgeted Profit is 129,380 in both reports.

Essbase is now referred to as a database (Oracle's relational database), so Essbase is now an OLAP database. Some of them (such as the ability to write-back data) were more interesting than others. While some other databases at the time met one or more of the qualifications, the only OLAP database to meet all ten was Arbor Software's Essbase. (Remember that Arbor is the company that commissioned the study.)

It's interesting to note that while Essbase was a database (as Scenario and Measures both are), there are no restrictions on which dimensions can be down the side or across the columns (if you wanted to see a nice trend of negative) in every month. To represent this in Excel, we use the time met one or more of the qualifications, the only OLAP database was Arbor Software's Essbase. (Remember that Arbor is the company that commissioned the study.)

In Oracle's world, there is only one database to meet all ten was Arbor Software's Essbase. (Remember that Arbor is the company that commissioned the study.)

We've now introduced a third dimension. Most people call this dimension Category, Ledger, or Version. It is Essbase traditionally referred to this dimension as the Category dimension. However, in Oracle's database, this is called the Scenario dimension. Some people like to call the dimension that contains Actual, Budget, Forecast, etc. to pivot our market dimension down to our columns so that we can compare profitability across different regions? To do this, we'd either have to have a series of linked spreadsheet formulas or hire a temporary employee to print out all the spreadsheets and type them in again with the markets now in the columns. Within Excel, it's problematic to represent more than one dimension over to the rows, our report will now look like this:

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Measures</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Sales</td>
<td>129,380</td>
</tr>
<tr>
<td>B</td>
<td>COGS</td>
<td>45,678</td>
</tr>
<tr>
<td>C</td>
<td>Profit</td>
<td>83,702</td>
</tr>
</tbody>
</table>

In the bottom-right corner, you'll see our familiar actual profit and loss statement:

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Categories</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>January</td>
<td>Sales</td>
<td>129,380</td>
</tr>
<tr>
<td>February</td>
<td>COGS</td>
<td>45,678</td>
</tr>
<tr>
<td>March</td>
<td>Profit</td>
<td>83,702</td>
</tr>
</tbody>
</table>

For the first few years, everyone called Essbase (and its competitors like Cognos, Business Objects, and Express) either an MDDB or OLAP database. The problem was that this was very difficult to explain to a casual user. Since casual users (CEOs, CFOs, etc.) are the ones who tend to sign checks at most companies, this produced a marketing problem of the highest order. What is the actual purpose of these OLAP databases? One of the overarching beliefs was that OLAP/MDDB databases "helped users make decisions and then provide them with the information they needed to support those decisions." Since 1992 when Essbase was released, other terms have been bandied about at various times including EIS (either SQL server or market) or we could hire a temporary employee to print out all the spreadsheets and type them in again with the markets now in the columns.
space. They should be since they created the term in the first place.

Beer, Cream Soda, Fruit Soda, and Diet Drinks. If we ever wanted

Income Statements tend to have accounts, time periods, scenarios,

needs to manage the performance of her company. Nowadays, this

is two-fold. First, financial minds tend to understand Essbase

some have recently started using the phrase “Business Analytics”
to refer to both EPM and BI, so you may sometimes hear this one

Hyperion made) was been seen as a financial tool. The reason for

the IT organization often felt left out and turned to inferior prod-

organizations, departments, companies and years on them. Since

marketing. Since Essbase was so easy to explain to end users in

relational databases do a poor job at multi-dimensional data,

EPM should be able to handle it whether the original number is

Sarbanes-Oxley can put you in jail, and EPM can help keep you

these days thanks in no small part to Sarbanes-Oxley bringing

compliance and management of data to the forefront. Simply put,

EPM (Enterprise Performance Management) is meant to

as scorecarding, planning, and financial consolidation. If there is

instance, the Product dimension has five children: Colas, Root

member of an Essbase dimension (like “100-10”). An “alias” is the

queries. A “member name” is the short, computery name for the

a well funded marketing campaign by IBM in the late

While counting with generations is pretty straight-forward,

Note that there are members that don’t have any children.

What if you want to refer to all the members into which May

Qtr2 is the parent of May. Year is the parent of Qtr2.

As level-1. Since the level-0 members of the Year dimension are

Level-0 members are sometimes also referred to as “leaves,”

In the picture above, May is childless. We refer to childless mem-

members as being “level-0”. If you ever want all of the bottom, child-

years on them. Since

It is also worth pointing out that

as scorecarding, planning, and financial consolidation. If there is

A “dimension” defines different categories for your data. A

descendants of Year would include all four quarters and all twelve

If Colas, Root Beer, and the other rug rats are all the children

What if you want to refer to all the members into which May

If you want to count down the hierarchy, use generations

To truly succeed in the world of Essbase, there are some

Just as the parents of the level-0 members are level-1 mem-

Historically, Essbase (and pretty much every other product

To refer to those members on a report without hard coding them,

we need both levels and generations? Well, in some dimensions

we could say “give us all the children of Product.”

While counting with generations is pretty straight-forward,

other rug rats are all the children

we could say “give us all the children of Product.”

we could say “give us all the children of Product.”

and it definitely was nice to have both options available to me. The

Parents of level-1 members are level-2 members. Their

Parents of level-1 members are level-2 members. Their

Parents of level-1 members are level-2 members. Their

Parents of level-1 members are level-2 members. Their

Parents of level-1 members are level-2 members. Their

Parents of level-1 members are level-2 members. Their

Parents of level-1 members are level-2 members. Their

Parents of level-1 members are level-2 members. Their

Parents of level-1 members are level-2 members. Their

Parents of level-1 members are level-2 members. Their

Parents of level-1 members are level-2 members. Their

Parents of level-1 members are level-2 members. Their

Parents of level-1 members are level-2 members. Their

Parents of level-1 members are level-2 members. Their

Parents of level-1 members are level-2 members. Their

Parents of level-1 members are level-2 members. Their

Parents of level-1 members are level-2 members. Their

Parents of level-1 members are level-2 members. Their

Parents of level-1 members are level-2 members. Their

Parents of level-1 members are level-2 members. Their

Parents of level-1 members are level-2 members. Their

Parents of level-1 members are level-2 members. Their

Parents of level-1 members are level-2 members. Their

Parents of level-1 members are level-2 members. Their

Parents of level-1 members are level-2 members. Their

Parents of level-1 members are level-2 members. Their

Parents of level-1 members are level-2 members. Their

Parents of level-1 members are level-2 members. Their

Parents of level-1 members are level-2 members. Their

Parents of level-1 members are level-2 members. Their

Parents of level-1 members are level-2 members. Their

Parents of level-1 members are level-2 members. Their

Parents of level-1 members are level-2 members. Their

Parents of level-1 members are level-2 members. Their

Parents of level-1 members are level-2 members. Their

Parents of level-1 members are level-2 members. Their

Parents of level-1 members are level-2 members. Their

Parents of level-1 members are level-2 members. Their

Parents of level-1 members are level-2 members. Their

Parents of level-1 members are level-2 members. Their

Parents of level-1 members are level-2 members. Their

Parents of level-1 members are level-2 members. Their

Parents of level-1 members are level-2 members. Their

Parents of level-1 members are level-2 members. Their

Parents of level-1 members are level-2 members. Their

Parents of level-1 members are level-2 members. Their

Parents of level-1 members are level-2 members. Their

Parents of level-1 members are level-2 members. Their

Parents of level-1 members are level-2 members. Their

Parents of level-1 members are level-2 members. Their

Parents of level-1 members are level-2 members. Their

Parents of level-1 members are level-2 members. Their

Parents of level-1 members are level-2 members. Their

Parents of level-1 members are level-2 members. Their

Parents of level-1 members are level-2 members. Their

Parents of level-1 members are level-2 members. Their

Parents of level-1 members are level-2 members. Their

Parents of level-1 members are level-2 members. Their

Parents of level-1 members are level-2 members. Their

Parents of level-1 members are level-2 members. Their
Stan Yellott Scholarship Fund
Eligibility Requirements

- Registered or the intent to register at an accredited post secondary educational institution in the United States
- Minimum GPA 2.5 from a current official transcript
- Currently enrolled in a computer science/information technology class and/or intends to enroll in a computer science the following term

Award notification will be given within 45 days following application deadlines. Upon acceptance of the scholarship, additional information may be required.

For Details, Visit the RMOUG Website
www.rmoug.org

RMOUG Scholarship Mission

To provide educational opportunities to members of the organization about the information technology industry in general, and in particular the technology of Oracle Corporation to include databases, storage, networking and application development, specifically the products and services of the Oracle Corporation.

To collect, post and distribute information about Oracle technologies and other related technologies to members.

To provide members with the ability to help their peers maximize their knowledge and skills working with products in information technology and in particular Oracle products.

To provide a consolidated channel of communication, conveying needs, concerns, and suggestions, for members of the organization, to Oracle Corporation and other vendor corporations involved with Oracle related technology.

To encourage members to present their information technology experiences using Oracle and other products and services.

To provide a consolidated channel of communication between members of the RMOUG and other communities in related information technology industries.

To promote educational opportunities for students of information technology through directed funding and services for educational purposes.

RMOUG is committed to supporting others in the pursuit of technical knowledge.

The Scholarship Fund started in 2001 to encourage future IT professional in their efforts to broaden their knowledge. In 2007, RMOUG voted to rename the scholarship fund to honor the memory of Stan Yellott. Stan was a long time member of RMOUG where he supported the user community by serving on the RMOUG board. Stan focused on expanding Oracle educational opportunities. Stan’s vision was to include high school and college students as the next generation of IT professionals.

Become A Member For These Outstanding Benefits

- Quarterly Education Workshops
- Special Annual Training Days Rates
- Database Labs
- List Server Subscriptions
- SQL>Update Magazine

www.rmoug.org/member.htm
While traditional Oracle relational developers have heard of Essbase, many don’t know what it really does. Is it ROLAP, some form of star schema, maybe a snowflake schema, or potentially some other s-noun schema heretofore unknown? Well, it’s actually none of those but to explain what it is will require going back about 20 years.

While a few pages isn’t enough to teach Essbase to everyone who reads this article, I did want to explain the history of Essbase for those who might want to understand why it’s good to have an option beyond the Oracle database itself.

Essbase wasn’t owned by Oracle until 2007. Prior to that, Essbase was produced by a company named Hyperion Solutions Corporation. Although Hyperion was founded in 1981, the Essbase product came along in the early 1990’s compliments of a company whose only real product was Essbase: Arbor Software. Up until 1998 when Hyperion and Arbor “merged”, the two companies were fierce competitors who were just as likely to spit on each other in waiting rooms as work together. (I’m kidding, but only slightly, because I can seem to recall at least one spitting incident.)

Arbor Software was founded in 1991 by Jim Dorrian and Bob Earle. They noticed at the time that companies were beginning to use spreadsheets not just for presentation of information but as a place to store data and business logic. Often, multiple sources of data were being consolidated together in spreadsheets and they were even seeing companies begin to release analysis to the public based on data in spreadsheets.

Jim/Bob wanted to build a database for spreadsheets. Essbase actually stands for Extended SpreadSheet database. Thanks to some creativity (and some venture capital from Hummer Winblad), they released the first version of Essbase in 1992. This original release of the product garnered three whole paragraphs of press coverage in Software Magazine on May 15, 1992. Here it is in all its “babe in the woods waiting to be eaten by bears” naïveté:

Data server “feeds” 1-2-3 or Excel
Arbor Software Corp.’s Essbase data server
Software Magazine; May 15, 1992

Following a three-year development effort, start-up Arbor Software Corp., Santa Clara, Calif., has built a data server that “feeds” popular desktop offerings, including 1-2-3 from Lotus Development Corp., and Excel from Microsoft Corp., on client machines.

“We conceded the front end to [widely installed] spreadsheets,” said James Dorrian, president and co-founder. “We built the product with two assumptions: that people knew their spreadsheets and that people knew their jobs.”

According to Marketing Vice President Michael Florio, the OS/2-based $22,000 Essbase offers users in client/server environments simultaneous access to large volumes of multidimensional spreadsheet data.

Notice that it was originally developed to run on OS/2 and its claim to fame was that it fed spreadsheets. Also, notice that you could get a server license for only $22,000 which sort of goes to show you that technology doesn’t always get cheaper over time.

The first version of the product wasn’t nearly as user friendly as it is today. Ignoring the Herculean steps required to actually build an Essbase database, retrieving data into Excel (or Lotus, at the time) required writing requests to Essbase in a language known as “Essbase Report Scripting.”

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>&lt;PAGE</td>
<td>Measures,</td>
<td>Product,</td>
<td>Market</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Profit</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Product</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Market</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>&lt;COLUMN</td>
<td>Scenario</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>&lt;CHILD</td>
<td>Scenario</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>&lt;ROW</td>
<td>Year</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>&lt;ICHILD</td>
<td>Year</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>!</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Essbase Report Scripting wasn’t very intuitive to end users, so some genius at Arbor came up with what became synonymous with Essbase: the Essbase Spreadsheet Add-in for Excel. It provided a much friendlier interface that finance and accounting people loved (since they already loved Excel):

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td>Profit</td>
<td></td>
<td>Product</td>
<td>Market</td>
</tr>
<tr>
<td>2</td>
<td>Actual</td>
<td>Budget</td>
<td></td>
<td>Variance</td>
<td>Variance %</td>
</tr>
<tr>
<td>3</td>
<td>Qtr1</td>
<td>24703</td>
<td>30580</td>
<td>-5877</td>
<td>-19.21844343</td>
</tr>
<tr>
<td>4</td>
<td>Qtr2</td>
<td>27107</td>
<td>32870</td>
<td>-5763</td>
<td>-17.5329459</td>
</tr>
<tr>
<td>5</td>
<td>Qtr3</td>
<td>27912</td>
<td>33980</td>
<td>-6068</td>
<td>-17.85756327</td>
</tr>
<tr>
<td>6</td>
<td>Qtr4</td>
<td>25800</td>
<td>31950</td>
<td>-6150</td>
<td>-19.24882629</td>
</tr>
<tr>
<td>7</td>
<td>Year</td>
<td>105522</td>
<td>129380</td>
<td>-23858</td>
<td>-18.44025352</td>
</tr>
<tr>
<td>8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Eventually, the Essbase Spreadsheet Add-In was replaced by a superior called Hyperion Smart View.

Smart View is the Office Add-in for all of the Oracle EPM System (formerly known as Hyperion) products including Essbase, Hyperion Planning, Hyperion Financial Management, Web Analysis, Financial Reporting, Strategic Finance and Enterprise. Unlike its predecessor, the Excel Add-In, Smart View works across all Office products: Excel, Word, PowerPoint, and so on.

**OLAP and Other Terms**

When Essbase was first released, no one was quite sure what it was. Was it some sort of spreadsheet on steroids? Was it a temporary employee who was really good at typing? Was it a database? If so, why didn’t it have records and fields and most importantly, why didn’t it let us IT types write SQL to access it?

Everyone was pretty sure what it wasn’t: a typical relational database. The creators originally called it a “data server.” Shortly after Essbase was created, they commissioned a study by the late Dr. E.F. Codd (the same Ph.D. who came up with the original rules for what constituted a true relational database) to determine what the heck Essbase was.

Dr. Codd was definitely impressed. He felt that this wasn’t a relational database yet it was definitely a database and a very important new type to boot. He called it an “OLAP” database to separate it from every other database up to that point.

To put it simply, all databases prior to Essbase were built for the purpose of storing transactions. The goal for these systems was to get individual records into the database as quickly as possible and to get those same records back out again as quickly as possible. A side goal was to store that data in as small a space as possible, because those were the days when hard drive space cost as much as a good mule. Summarization of these types of databases was possible, but definitely not the objective of the database design. Dr. Codd classified traditional relational databases “OLTP” (On-Line Transaction Processing).

He knew that Essbase was the first database designed purely to support analysis. Knowing that this was going to be The Next Big Thing, he created a term to describe these databases: OLAP (On-Line Analytical Processing). There were several features that Essbase offered that no previous database could handle.

**Multi-Dimensional Databases**

First of all, Essbase was a multi-dimensional database (MDDB or MDB, for short). What did the good doctor mean when he said Essbase was multi-dimensional? Simply that any of the dimensions set up in a database could be put in the rows or the columns (or applied to the whole page/report).

All databases up to this point were two-dimensional: records and fields. Essbase had no theoretical dimension limit (though there was certainly a practical limit). The Sample.Basic database that comes with Essbase has five base dimensions: Year, Measures, Product, Market, and Scenario. It also has five more “attribute” dimensions: Caffeinated, Ounces, Pkg Type, Population, and Intro Date. The ASOSamp.Basic database that comes with Essbase is a sample Aggregate Storage database with 14 dimensions. The largest database I’ve ever seen had over 100 dimensions, but I think they were just trying to show off. In general, typical Essbase block storage databases have five to ten base dimensions. By base dimension, we mean dimensions that show up all the time (like the five mentioned above in Sample.Basic).

While any relational database can be set up to give the appearance of having multiple dimensions, it takes a lot of up front work by developers. Essbase and other OLAP databases have dimensionality built-in.

**Optimized for Retrieval**

Essbase databases were also optimized for retrieval at any level of the hierarchy, even the very topmost number that might represent every dollar the company has ever made in its history. OLTP databases (relational databases) were nicely optimized for retrieval of detailed records but definitely not hierarchical information. By pre-saving summarized information, Essbase allows analysis to happen from the top down with no decrease in performance.

For OLAP databases, the hierarchy is native to the database itself. This is far different from relational databases that store the data in one table and then have one or more other tables that can be joined in to view data in a rolled-up fashion. For Essbase, the hierarchy is the database. When you change the hierarchy logic in Essbase as to how a product is grouped or a market rolls-up, you actually change where the data is stored.

Because hierarchy is inherent to OLAP databases, drill-down (sometimes known as “slicing and dicing” but never known as “making julienne data”) is inherent as well. Essbase is great at doing Ad hoc analysis because it knows that when a user double-clicks on Qtr1, she wants to see Jan, Feb, and Mar. This is because the roll-up of months to quarters is pre-defined back on the server.

Dr. Codd came up with ten rules for defining OLAP databases. Some of them (such as the ability to write-back data) were more interesting than others. While some other databases at the time met one or more of the qualifications, the only OLAP database to meet all ten was Arbor Software’s Essbase. (Remember that Arbor is the company that commissioned the study.)

It’s interesting to note that while Essbase was a database from 1992-2007, Oracle no longer refers to Essbase as a database even though it stores data. In Oracle’s world, there is only one database (Oracle’s relational database), so Essbase is now referred to as an OLAP Server. I still think it’s a database, though.

**What is a Dimension?**

Here is a really simple Profit & Loss Statement:
Four Dimensions and More

Right now, this spreadsheet is not broken down by product or market. Within Excel, it’s problematic to represent more than three dimensions (since we’ve used the rows, columns, and tabs). One way is to have a separate file for each combination of product and market:

As you can see, this is getting ridiculous. What if we want to pivot our market dimension down to our columns so that we could compare profitability across different regions? To do this, we’d either have to have a series of linked spreadsheet formulas (which would break as soon as we added or deleted a new product or market) or we could hire a temporary employee to print out all the spreadsheets and type them in again with the markets now in the columns. While the latter method is obviously error-prone, the “rekeying” method is the one used by the majority of the companies in the world that do not own Essbase or another OLAP product.

Since Market and Product are dimensions, it should be no more difficult to put them in the columns or rows than Scenario or Measures. Producing a report with markets down the side and products across the top is no more difficult than dragging-and-dropping:

In the bottom-right corner, you’ll see our familiar actual profit for the year of 105,522. At the top of the report, you’ll see that we have to specify the three dimensions in our application that are not in our rows or columns or Essbase wouldn’t know which values to display. For instance, if we didn’t specify “Profit”, Essbase wouldn’t know if we wanted Profit, Sales, Margin, or some random measure to be named later.

DSS, EIS, BI, BPM, EPM...

Now that we all agree that Essbase is multi-dimensional (and we understand multi-dimensionality at least a wee little bit), what should we call the general category of databases that includes Essbase?
For the first few years, everyone called Essbase (and its competitors like Cognos, Business Objects, and Express) either an MDDB or OLAP database. The problem was that this was very difficult to explain to a casual user. Since casual users (CEOs, COOs, CFOs, etc.) are the ones who tend to sign checks at most companies, this produced a marketing problem of the highest order. What is the actual purpose of these OLAP databases?

The overarching belief was that OLAP/MDDB databases “helped users make decisions and then provide them the information they needed to support those decisions.” Since HUMDATPTITNTSTD makes for a lousy acronym, the term DSS was created and thus the “Decision Support Systems” term was coined.

Since 1992 when Essbase was released, other terms have been bandied about at various times including EIS (either “Executive Information Systems” or “Enterprise Information Systems” depending on whom you ask) and BI (Business Intelligence). Business Intelligence is still used fairly frequently (thanks to a well funded marketing campaign by IBM in the late 90’s), but its popularity is quickly being overtaken by EPM.

EPM (Enterprise Performance Management) is meant to include BI and expand it to also include any information a user needs to manage the performance of her company. Nowadays, this goes well beyond just a database and includes applications such as scorecarding, planning, and financial consolidation. If there is a number that needs to be manipulated, rolled-up, sliced, or diced, EPM should be able to handle it whether the original number is in an OLAP or OLTP database. It is also worth pointing out that some have recently started using the phrase “Business Analytics” to refer to both EPM and BI, so you may sometimes hear this one too.

Historically, Essbase (and pretty much every other product Hyperion made) was been seen as a financial tool. The reason for this is two-fold. First, financial minds tend to understand Essbase really well. Financial analysis is inherently multi-dimensional. Income Statements tend to have accounts, time periods, scenarios, organizations, departments, companies and years on them. Since relational databases do a poor job at multi-dimensional data, finance types started using spreadsheets. Since Essbase was a database for spreadsheets, it made it really easy to explain the value to CFOs, Controllers, VPs of Planning, and the like.

The second reason for Essbase’s traditional stereotyping as “something the bean counter uses” has to do with sales and marketing. Since Essbase was so easy to explain to end users in accounting and finance, that’s the group that the Essbase sales representatives tended to call on. The sad part about this is that the IT organization often felt left out and turned to inferior products from competing vendors because those vendors were seen as developing products that were more “IT-centric.”

As for the current market, Oracle (including the Hyperion brand) is generally accepted to be the market leader in the EPM space. They should be since they created the term in the first place in the early 21st century. EPM is quite the hot software niche these days thanks in no small part to Sarbanes-Oxley bringing compliance and management of data to the forefront. Simply put, Sarbanes-Oxley can put you in jail, and EPM can help keep you out.

### Essbase Terminology

To truly succeed in the world of Essbase, there are some handy terms to pick up. Some of them we’ve already learned.

A “dimension” defines different categories for your data. A dimension can be located on the rows, columns, or pages of your queries. A “member name” is the short, computerly name for the member of an Essbase dimension (like “100-10’). An “alias” is the longer, more descriptive name for a member (like “Cola”). All of the dimensions in a database make up the “outline.”

Here is a portion of Sample.Basic outline:

<table>
<thead>
<tr>
<th>Outline</th>
<th>Basic (Active Alias Table: Default)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Year</td>
<td>Qtr1 (+)</td>
</tr>
<tr>
<td></td>
<td>Jan (+)</td>
</tr>
<tr>
<td></td>
<td>Feb (+)</td>
</tr>
<tr>
<td></td>
<td>Mar (+)</td>
</tr>
<tr>
<td>Year</td>
<td>Qtr2 (+)</td>
</tr>
<tr>
<td></td>
<td>Apr (+)</td>
</tr>
<tr>
<td></td>
<td>May (+)</td>
</tr>
<tr>
<td>Year</td>
<td>Qtr3 (+)</td>
</tr>
<tr>
<td></td>
<td>Jun (+)</td>
</tr>
<tr>
<td></td>
<td>Jul (+)</td>
</tr>
<tr>
<td>Year</td>
<td>Qtr4 (+)</td>
</tr>
<tr>
<td></td>
<td>Aug (+)</td>
</tr>
<tr>
<td></td>
<td>Sep (+)</td>
</tr>
<tr>
<td></td>
<td>Oct (+)</td>
</tr>
<tr>
<td>Measures</td>
<td>Profit (+)</td>
</tr>
<tr>
<td></td>
<td>Income (+)</td>
</tr>
<tr>
<td></td>
<td>Total Expenses (-)</td>
</tr>
<tr>
<td>Measures</td>
<td>Sales (+)</td>
</tr>
<tr>
<td></td>
<td>Margins (+)</td>
</tr>
<tr>
<td></td>
<td>COGS (-)</td>
</tr>
<tr>
<td>Measures</td>
<td>Ratios (-)</td>
</tr>
<tr>
<td>Product</td>
<td>100 (+) (Alias: Colas)</td>
</tr>
<tr>
<td></td>
<td>200 (+) (Alias: Root Beer)</td>
</tr>
<tr>
<td></td>
<td>300 (+) (Alias: Cream Soda)</td>
</tr>
<tr>
<td></td>
<td>400 (+) (Alias: Fruit Soda)</td>
</tr>
<tr>
<td></td>
<td>500 (+) (Alias: Diet Drinks)</td>
</tr>
</tbody>
</table>

### Family Tree Relationships

The most common way to refer to members in an outline relative to each other is by using “family tree” relationships. The members directly below a member are called its children. For instance, the Product dimension has five children: Colas, Root Beer, Cream Soda, Fruit Soda, and Diet Drinks. If we ever wanted to refer to those members on a report without hard coding them, we could say “give us all the children of Product.”

The advantage to this aside from the saving in typing is that if a new product line was to be added (say, “Water”), we wouldn’t have to modify our reports. Any report designed to display the children of Product would pick up the new “Water” product and add it to the list automatically.

If Colas, Root Beer, and the other rug rats are all the children of Product, what relation is Product to its children? Assuming you didn’t fail “Birds and the Bees 101,” you’ll know that Product must be the parent of Colas, Root Beer, and the rest. In other words, the parent of any member is the one that the member rolls-up into. Qtr2 is the parent of May. Year is the parent of Qtr2.

Since Colas and Root Beer are both the children of Product, Colas and Root Beer are siblings. This is simple, but what relationship do January and May have? Well, their parents are siblings so that makes them... cousins. Correct, but “cousins” while technically correct isn’t used that often. In general, people say that January and May are at the “same level.”

What if you want to refer to all the members into which May rolls (not just the one right above)? Well, those are its ancestors which in this case would be Qtr2 and Year. Correspondingly, the descendants of Year would include all four quarters and all twelve months.

Note that there are members that don’t have any children. In the picture above, May is childless. We refer to childless members as being “level-0”. If you ever want all of the bottom, child-less members of a dimension, just ask for the level-0 members. For example, the level-0 members of the Year dimension are the months and the
Level-0 members of the Market dimension are the states.

Level-0 members are sometimes also referred to as “leaves,” because they’re at the edges of the family tree. I sometimes refer to level-0 members as “the ones who aren’t allowed to sit at the main table on Thanksgiving,” but I think I’m the only one.

All of the parents of the level-0 members are referred to as level-1. Since the level-0 members of the Year dimension are the months, then the level-1 members are the quarters. For the Market dimension, the level-1 members are the regions: East, West, South, and Central.

Just as the parents of the level-0 members are level-1 members, the parents of level-1 members are level-2 members. Their parents are level-3 members and so on up the hierarchy. There are many places in Essbase that you can specify, for example, “All the level-2 members of the Product dimension,” so remember that levels count up from the bottom of a dimension starting at 0.

If you want to count down the hierarchy, use generations instead of levels. The dimension itself is considered generation-1 (or “gen1,” for short). Its children are gen2. For the Year dimension, the gen2 members are the quarters.

Yes, the quarters are both level-2 and generation-2. Why do we need both levels and generations? Well, in some dimensions with many, many levels in the hierarchy, you’ll want to count up from the bottom or down from the top depending on which you’re closer to. We’ve seen a dimension with 17 levels in the hierarchy, and it definitely was nice to have both options available to me. The children of gen2 members are gen3 and so on down the hierarchy.

While counting with generations is pretty straightforward, levels can sometimes be a bit tricky. Look at this portion of the Measures dimension from Sample.Basic:

For this dimension, Gen1 is Measures. Gen2 is Profit and Inventory. Gen3 is Margin, Total Expenses, Opening Inventory, Additions, and Ending Inventory.

So far this is looking pretty easy, but let’s switch our focus to the levels. The level-0 members are Sales, COGS, Marketing, Payroll, Misc, Opening Inventory, Additions, and Ending Inventory. The level-1 members are Margin, Total Expenses, and Inventory. What are the level-2 members? Profit (because it’s the parent of level-1 members Margin and Total Expenses) and Measures (because it’s the parent of level-1 member Inventory).

The trickiness is that Measures is also a level-3 member because it’s the parent of Profit, a level-2 member. This means that if you ask Essbase for level-2 members, you’ll get Measures, but you’ll also get Measures if you ask for level-3 members. Notice that this counting oddity does not occur with generations.

This instance of a dimension is also known as a ragged hierarchy. Essbase handles ragged hierarchies & dimensions extremely well whereas most of its competitors, well, don’t.

Recent Changes to Essbase

One of the questions that I hear a lot is “now that Oracle owns Essbase, is it going to continue as a stand-alone product or be integrated into existing Oracle offerings?” While I don’t work for Oracle, it’s safe to say that Oracle’s path definitely appears to be to keep Essbase as an independent product.

Oracle has invested a great deal of research and development into the newest release of Essbase (11.1.2.2 as of the time of this writing) and the new features are truly impressive. From allowing text and dates in a cube to allowing attribute dimensions to vary over time to supporting 128 simultaneous CPUs during different processing steps, Oracle seems to be adding functions left and right. Instead of trying to commoditize Essbase as Hyperion seemed to be doing in its final days, Oracle is more focused on setting Essbase free to retain its lead as the number one BI/OLAP/EPM database in the world. Be assured that if you invest the time to learn Essbase, your effort will not be wasted.

For More Information

Portions of the article above were previously published in the book Look Smarter Than You Are with Essbase 11. Find the book on LuLu.com (keyword “Essbase”) if you want to continue learning more about Essbase. If you like free information (and who doesn’t in this economy), you can get all the up-to-date news in the world of Oracle EPM, Hyperion, and Essbase by visiting my Hyperion blog at http://looksmarter.blogspot.com/.

Edward Roske, Oracle ACE Director, is the co-author of the best-selling Hyperion book Look Smarter than You are with Essbase. He is CEO of interRel Consulting, a Hyperion-specific consulting firm, and blogs about Hyperion matters at LookSmarter.blogspot.com. Edward is an unimaginably frequent speaker at Oracle/Hyperion conferences and he is noted for his humorous slant to unconscionably boring information. He spends his spare time evolving into pure energy and saving busloads of nuns & orphans. He also likes puppies.
Thank you so much for all your hard work and your willingness to help out wherever there was a need. Your incredible effort was greatly appreciated by everyone at the conference. We couldn’t ask for a more dedicated group of volunteers and words can’t express how grateful we are for the great feat that is accomplished each year by our volunteers. It was a pleasure and a privilege to work with all of you at RMOUG Training Days 2013, and we look forward to seeing you again in future.

With Gratitude From The Training Days 2013 Conference Committee
Kellyn Pot’vin, Tim Gorman, Pat VanBuskirk, John Jeunnette and Peggy King

Recognizing Our Volunteers

Airport Shuttle Drivers

Antonio (Tree) Adams
Crystal Adams
Lisa Cocklin
Yelena Cohen
Peter Gorman
Tim Gorman
Tom Green
Prashant Khanolkar
Greg Opie
Kellyn Pot’vin
Nicholas Ratanasin
Steve Richardson

Registration & Room Ambassadors (cont’d)

Barbara Dungey
William Farrell
Carole Furnish
Mario Garcia
Peter Gorman
Tom Green
Paul Hagar
Dan Hotka
Prashant Khanolkar
Kelly Kozlowski
Rolf Kramer
Peggy Lentz
George Linehan
Elaine Magnuson
Ray Mariano
Art Marshall
Jenny McGurk
Mike McMannis
Raymond Obuch
Greg Opie
John Pack
Bill Padfield
Nicholas Ratanasin
John Reading
Steve Richardson
Wayne Stetson
Dara Tallmadge

It was the death of his wife in 2004 that changed his path. Suddenly a single father, traveling as a consultant became impossible. Businesses were tightening their belts in 2004 and Vincent decided that being closer to family would help when it came to raising his 14 year old son. Fortunately, his parents and two of his brothers lived in Hawaii. Unfortunately, the Oracle Business Suite was used by only a few government agencies. Most corporations in Hawaii are truly small business in nature, so enterprise class financial software is uncommon.

Vincent’s mother is a real estate agent in Hawaii, so Vincent decided to try Real Estate as a primary career, keeping his hand in the technology world by supplementing this decision with consulting jobs. After two years of selling houses and draining much of his savings, Vincent decided that the consulting work was the work he truly found exciting, so he decided to take the Oracle 10g Administration courses as a path to becoming an Oracle Certified Professional. The classes opened some doors and Vincent passed his Oracle Certified Associate before landing a job with National Information Consortium in downtown Honolulu for $60K a year.

Some background on the job market in Honolulu. Database management jobs start at $40K a year. Years of experience do not count towards a higher salary, only time on the job. Finding a job that paid $60K was a challenge. Certifications are very beneficial in the Hawaii marketplace. The only way to make decent money in Hawaii is to be hired directly from the mainland. Do not move there and then expect to find a good paying job.

Military consulting on the island of Oahu pays a bit more, yet the attitude and politics of the workplace in the government setting is completely different from the mainland. You are expected to do what you are hired to do, nothing more, and be on call 24/7. Vincent’s experience seemed to indicate that pushing the envelope by expressing new ideas or suggestions for improvement usually meant you were making waves in a process that already worked. The politics are very different.

Vincent had applied to the University of Hawaii upon his arrival in 2004. He had maintained his contacts through the years and made friends with the IT manager, even kayaking with him in 2005. In 2007, the University finally called and asked if Vincent was still interested in working as a database administrator. Vincent applied and managed to get a salary of $73K a year. Less than a year later, the state cut everyone’s salary 8% to balance the state budget. The Union (you are a member whether you want to be or not when accepting a state job) compensated for the cut in pay by giving all public school students every other Friday off (termed Furlough Friday), creating a secondary burden for working parents and shortening the number of days of instruction.

Vincent decided to move back to Denver for its great school system, decent salaries, normal housing prices, and to get reacquainted with his past associates and the Oracle Business Suite. Hired by Berry Petroleum in May 2011, Vincent moved back to Denver with his new family including two children ages 9 and 11. He is living in Highlands Ranch and takes light rail downtown each morning.

RMOUG is another reason for moving...
back to Denver. The opportunity to participate in training classes and the RMOUG training days each year is unparalleled. Training opportunities in Hawaii are rare and the users group there is non-existent. The Hawaii Oracle Users Group (HIOUG) did exist for a time, yet the participation was minimal.

Vincent enjoys exploring and has visited seven National Parks with his family since arriving, including Rocky Mountain, Great Sand Dunes, Petrified Forest, Grand Canyon, Arches, Mount Rushmore, and Yellowstone. The adventure of traveling for hours by car is new to his children, who have never lived off island until now.

Although the children have had a chance to play in what little snow we have had this year, they are truly looking forward to a real snow storm where they can sled downhill without having to avoid the areas where the grass is still poking out. The entire family is looking forward to learning to ski, so say a prayer we will have some snow this season.

It’s good to be back in the Rockies. If any RMOUG member has an interest in working in Hawaii, Vincent has offered to share his experiences. Vincent met his wife, Nyki, while practicing Real Estate and she remains one of the best advisors for what island life offers and the true cost of living. Nyki lived on Oahu for 25 years and raised two children there until moving to Colorado.
Reach A Targeted Oracle Audience
Advertise Now!

A full page, full color ad in RMOUG SQL>UPDATE costs as little as 70 cents per printed magazine and even less for smaller ads.

RMOUG SQL>Update Advertising Rates

<table>
<thead>
<tr>
<th>Type</th>
<th>Price</th>
</tr>
</thead>
<tbody>
<tr>
<td>Business card or 1/8 page</td>
<td>$ 50</td>
</tr>
<tr>
<td>1/4 page</td>
<td>$ 350</td>
</tr>
<tr>
<td>1/2 page</td>
<td>$ 625</td>
</tr>
<tr>
<td>Full page</td>
<td>$1,000</td>
</tr>
<tr>
<td>Inside cover</td>
<td>$1,250</td>
</tr>
<tr>
<td>Back cover</td>
<td>$1,500</td>
</tr>
</tbody>
</table>

Discounts available for RMOUG Members and Full Year Contracts

See Media Kit for deadlines and mechanical requirements.
Submit ad materials to: Pat Van Buskirk, RMOUG Newsletter Director
38101 Comanche Creek Road • Kiowa, CO 80117
303-621-7772 • E-mail: NewsletterDir@rmoug.org • www.rmoug.org

SQL>UPDATE is mailed to all RMOUG Members and distributed during Quarterly Education Workshops and Training Days each year
Meet Your Board

Tim Gorman
President & Vendors Director
E-mail: President@rmoug.org

Kelly Pot’vin
Training Days Director
E-mail: TrainingDaysDir@rmoug.org

David Fitzjarrell
Social Media Director
E-mail: TrainingDaysDir@rmoug.org

Kathy Robb
Board Member Emeritus
Arisant, LLC
E-mail: BoardME@rmoug.org

Ron Bich
Secretary & SIGS Director
E-mail: SIGSDir@rmoug.org

Thomas Green
Scholarship Director
E-mail: ScholarshipDir@rmoug.org

Carolyn Fryc
Vice President & Programs Director
E-mail: ProgramsDir@rmoug.org

Art Marshall
Web Director
E-mail: ISDir@rmoug.org

Heidi Kuhn
Administrative Assistant
Voice Mail: (303) 948-1786
Fax: (303) 933-6603
E-mail: Admin@rmoug.org

Pat Van Buskirk
Treasurer & Newsletter Director
E-mail: NewsletterDir@rmoug.org
RMOUG Events Calendar

3/21/13 Board March Board Meeting
4/6/13 DB Labs DB Labs, Regis University, 8:00 am to 1:00 pm, http://www.meetup.com/RMOUGLabs/
4/7-11/13 OAUG Collaborate (OAUG, Quest & IOUG) Conference, Colorado Convention Center, Denver, CO
4/15/13 SQL>Update Call for articles and cover photo, Summer Issue
5/17/13 QEW RMOUG Spring Quarterly Educational Workshop (QEW), Oracle/Sun office in Broomfield CO
Includes election of 2013-2014 Board of Directors
6/15/13 SQL>Update Publication, Summer Issue
8/9/13 QEW RMOUG Summer Quarterly Educational Workshop (QEW), Elitch Gardens in Denver CO
9/22-26/13 Open World Oracle Open World 2013 in San Francisco, CA (www.oracle.com/openworld)
11/14/13 QEW RMOUG Fall Quarterly Educational Workshop (QEW), Oracle DTC office in Denver, CO
11/25/13 QEW RMOUG Autumn Quarterly Educational Workshop (QEW), Oracle DTC office in Denver CO
2/5-7/14 Training Days RMOUG Training Days 2014, Colorado Convention Center in Denver, CO

Please note dates are subject to change. For the most current events calendar visit our website at www.rmoug.org.

SPONSOR A QUARTERLY EDUCATION WORKSHOP AND RECEIVE A HALF PRICE AD IN SQL>UPDATE

<table>
<thead>
<tr>
<th>Breakfast</th>
<th>Discounted Ad Rate</th>
<th>Total Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>1/4 Page</td>
<td>$350.00</td>
<td>$175.00</td>
</tr>
<tr>
<td>1/2 Page</td>
<td>$350.00</td>
<td>$312.50</td>
</tr>
<tr>
<td>Full Page</td>
<td>$350.00</td>
<td>$500.00</td>
</tr>
<tr>
<td>Inside Cover</td>
<td>$350.00</td>
<td>$625.00</td>
</tr>
<tr>
<td>Back Cover</td>
<td>$350.00</td>
<td>$750.00</td>
</tr>
</tbody>
</table>

Help RMOUG Members and Receive Recognition in An Upcoming Issue of SQL>Update

Contact Carolyn Fryc - Programs Director - 720-221-4432 - cfryc@orsportal.com

Tell Us About Yourself

Join us in sharing your Oracle experiences with other RMOUG members!

Tell us about your life, your job, or share your amusing Oracle anecdotes, tips and secrets!

Please submit all material to NewsletterDir@rmoug.org

Bring Your Horses & Guests
Franktown, CO $640,000

- 4282 Square Feet
- Fully Finished Basement
- Wet Bar, Decks
- Apartment Above Shop
- Bonus Room Above Garage
- Two Fireplaces + WB Stove
- Pasture With Loafing Shed

Entertainment Splendor In The Pines
Pat Van Buskirk • Coldwell Banker • Parker Office
(303) 243-0737 Cell (303) 841-5263 Office
www.patvanbuskirk.com pat@patvanbuskirk.com
Index To Advertisers

Arisant ................................................................. 2
DatAvail .............................................................. 18
DBAK ................................................................. 32
Quarterly Education Workshop ......................... 31
Quarterly Education Workshop Sponsorship ........... 30
Regis University .................................................. 2
Real Estate Coldwell Banker .............................. 30
RMOUG Advertising Rates ................................. 28

Reach A Targeted Oracle Audience Advertise Now!

A full page, full color ad in RMOUG SQL>UPDATE costs as little as 70 cents per printed magazine and even less for smaller ads.

Contact NewsletterDir@rmoug.org

Quarterly Education Workshops

May 17, 2013 • Oracle/Sun Office

Join us for our next Quarterly Education Workshop in May for presentations at the Oracle/Sun Office in Broomfield, CO. RMOUG hosts quarterly workshops in May, August and November of each year with the fourth and largest educational event being Training Days in February. Learn about the newest technologies, gain more insight into Oracle techniques and enjoy the camaraderie of meeting with other Oracle professionals.

If you or your organization are interested in partnering with RMOUG to host an upcoming meeting, or to submit an abstract for presentation, please contact

Carolyn Fryc, Programs Director at ProgramsDir@rmoug.org

Watch RMOUG’s Web Page for May Training Topics www.rmoug.org
Local Presence
Global Experience

Oracle Database
  Data Strategy
  Database Administration
  Data Warehouse
  Disaster Recovery
  Performance Tuning

Oracle License Management
  License Reviews
  Software Compliance

Managed Services
  Remote DBA
  Remote Applications
  Oracle Database Support
  Oracle Applications Support

Oracle Applications
  E-Business Suite R12
    Financials
    Supply Chain
    Manufacturing
    Procurement

People all over the world call us - because we fix systems that have a direct impact on business and revenue

www.dbak.com