

# OPTIMIZACIJA SKLADIŠTA PODATAKA KORIŠTENJEM ORACLE OLAP 11g



Josip Pojatina  
Agrokor ICT  
Ulica grada Vukovara 284/c  
10000 Zagreb  
Tel: +385 1 2372 065  
Fax: +385 1 2365 774  
Mob: +385 99 2271 418  
E-mail: [josip.pojatina@agrokor.hr](mailto:josip.pojatina@agrokor.hr)  
Web: [www.agrokor.hr](http://www.agrokor.hr)

## SAŽETAK:

Oracle OLAP (ex Express OLAP) je prva multidimenzionalna baza podataka. Nakon akvizicije, Oracle je počeo s ugradnjom Express OLAP engine u Oracle bazu (9i rel. 2). Svaka nova verzija je nudila sve veću integraciju relacijske i multidimenzijske baze. Oracle OLAP je jedini OLAP server koji je implementiran u samom kernelu baze, a ne kao odvojeni server. To mu daje brojne prednosti pred konkurentima počevši od jednostavnijeg backup-a, security-a i osvježavanja OLAP kocki, jednostavnijeg ETL-a, pristupanja OLAP-u korištenjem SQL-a, pa sve do Cube MV-ova i njihovo korištenje za query rewrite čime se mogu zamijeniti deseci postojećih MV-ova. Sve to uz moćan OLAP DML programski jezik, stavlja Oracle OLAP daleko ispred konkurentskih proizvoda.

## 1. UVOD

Prosječna veličina skladišta podataka posljednjih 10-ak godina bilježi strmoglav rast. Dok su do prije samo nekoliko godina skladišta podataka veličine nekoliko terabajta smatrana velikim, danas takvu veličinu skladišta podataka ostvaruju čak i manje tvrtke, dok su skladišta podataka veličine petabajta posve uobičajena u velikim kompanijama.

Takav vrtoglav rast skladišta podataka se može objasniti s jedne strane tehnološkim napretkom tehnologije koja omogućuje upravljanje tako velikom količinom podataka, dok s druge strane stoji jasno izražena potreba tvrtki da se što više podataka smjesti u skladišta podataka. Samo podaci koji se u praksi koriste imaju stvarnu vrijednost i donose prednosti onima koji ih znaju iskoristiti kroz jedan od vidova poslovne inteligencije počevši s klasičnim složenim izvještavanjem, preko OLAP-a, pa sve do rudarenja podacima (data mining-a) i statističke analize podataka.

Usprkos povećanju brzine procesora, sve većim količinama memorije, te sve bržim uređajima za skladištenje podataka (storage sustavi), jasno je da problem držanja većih skladišta podataka pod kontrolom ipak zahtjeva pored stručnog osoblja i korištenje raznih tehnologija koje omogućuju optimiziranje skladišta podataka s ciljem zadovoljavanja ograničenja koje diktira konkretna situacija. Primjeri takvih ograničenja su vremenski rok za punjenje skladišta podataka tijekom (najčešće) noćne obrade, potom vrijeme odziva (response time) na razne upite od strane korisnika prilikom pokretanja izvještaja i sl.

Tema optimiranja skladišta podataka je iznimno široka (napisani su deseci knjiga s tom temom) i obuhvaća područje optimiranja i ispravnog konfiguriranja sistema pohrane podataka (storage), operacijskog sustava, mreže, pa sve do optimizacije baze podataka (instance), ETL procesa, i u konačnici dizajna skladišta podataka. Oracle nudi cijeli arsenal različitih tehnologija za optimiranje skladišta podataka. Jedan od najjačih oružja, jedinstven isključivo za korisnike Oracle baze podataka, je mogućnost korištenja Oracle OLAP 11g servera za dobivanje spektakularnih performansi i vremena odziva. Upravo je korištenje Oracle OLAP opcija u svrhu optimiranja skladišta podataka tema ovog referata.

## 2. OLAP – ŠTO JE TO?

Kako bi se moglo pokazati na koji način Oracle OLAP 11g može pomoći prilikom optimiranja skladišta podataka, potrebno je najprije definirati što je OLAP, na koji način se on uklapa u sustav

poslovne inteligencije te koje su mu glavne karakteristike. OLAP je akronim za „Online analytical processing“.

Sam termin OLAP je relativno nedavno skovao Dr. E.F. Codd, poznat računalni znanstvenik i to 1993., iako su se multidimenzionalne baze pojavile nekoliko desetljeća ranije. Iako termin nije najbolje odabran (budući da u akronimu nema ni traga pojma multidimenzijski što i je suština svakog OLAP-a, dok ističe elemente koji nisu toliko bitni – npr. analytical processing ili čak elemente koji nisu točni – npr. online, budući da ne treba biti online da bi se radilo s OLAP-om), te usprkos tomu što su predloženi i drugi nazivi (Nigel Pendse je primjerice ponudio akronim FASMI – fast analysis of shared multidimensional information, koji daleko bolje opisuje suštinu onoga što se danas podrazumijeva pod pojmom OLAP), danas je OLAP uobičajen, prepoznatljiv i opće prihvaćen pojam.

Osnovno obilježje OLAP-a je njegova multidimenzionalnost. Pored toga, sukladno FASMI akronimu, ostala obilježja OLAP-a su:

- Fast - odnosi se na vrijeme odziva sistema na korisnikove upite
- Analysis - opisuje prvenstvenu namjenu OLAP-a a to je analiza podataka
- Shared - podrazumijeva višekorisnički pristup OLAP kockama, kao i zaštitu podataka
- Multidimensional – ovo je suština OLAP-a koji pohranjuje podatke u višedimenzionalne strukture podataka
- Information - svrha OLAP-a je dostavljanje informacija korisnicima

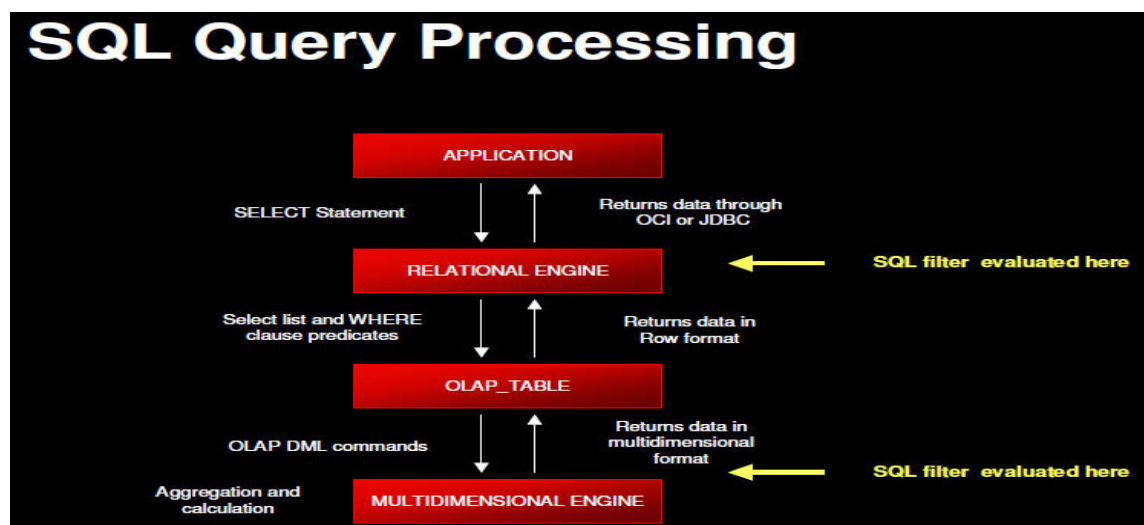
OLAP je zbog svoje multidimenzionalnosti, mogućnosti izrade složenih kalkulacija, mogućnost interaktivnog dobivanja neograničenog broja različitih izvještaja/analiza iz postojećeg seta podataka kao i jednostavnošću interpretacije rezultata izvještaja/analiza, zauzeo centralno mjesto svakog rješenja za poslovnu inteligenciju uzimajući najbolje osobine klasičnih izvještaja na bazi skladišta podataka s jedne te statističke analiza i rudarenja podataka (dana mining-a) s druge strane.

### 3. POVIJEST ORACLE OLAP-A

Povijest Oracle OLAP-a počinje početkom 1970 i to pod nazivom Express. Express je tako postao prva, a ujedno i jedna od najpopularnijih multidimenzionalna baza na tržištu. Mnoge aplikacije pisane u Express-u pred nekoliko desetljeća rade još i danas.

Oracle je akvizicijom kupio Express 1995 s ciljem da ga integrira u Oracle bazu. Prvi rezultat integracije je vidljiv na prelazu 2002-2003 godinu, kad je Oracle izdao Oracle 9i bazu podataka s integriranim OLAP serverom pohranjenim kao BLOB unutar baze.

S izlaskom 10g baze (2004-2006 godina), Oracle je dodao nove mogućnosti OLAP serveru poput kompresije i sl. S izlaskom 10g release 2 baze, Oracle je po prvi put omogućio pristup OLAP-u putem običnog SQL-a (kroz OLAP\_TABLE funkciju), integrirajući time relacijski i multidimenzijski dio Oracle baze. Na slici 1 je prikazana arhitektura pristupanja OLAP podacima iz SQL-a.



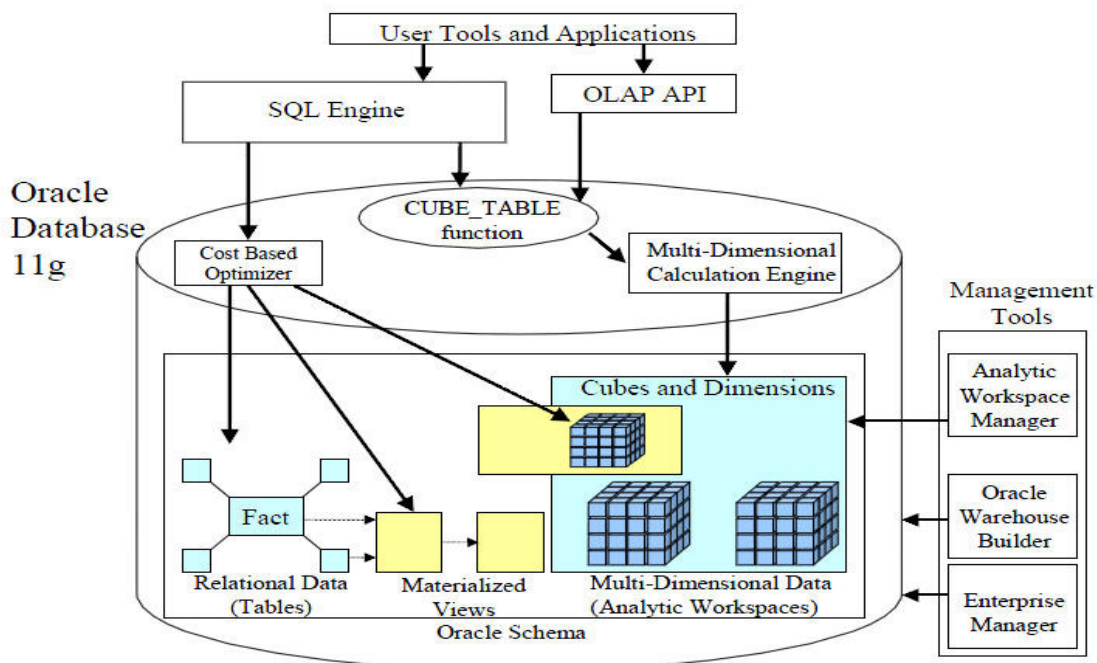
Slika 1: Prikaz pristupanja multidimenzijskim (OLAP) podacima direktno iz SQL-a

Ovo je jedna od kapitalnih funkcionalnosti koja je jedinstvena isključivo za Oracle bazu i bit će detaljnije objašnjena kasnije (uz live demo).

S pojavom Oracle 11g baze (2007-2009), Oracle je otišao još jedan korak dalje u integraciji OLAP 11g servera s relacijskim dijelom baze. Oracle je naime proširio logiku materijaliziranih view-ova, koji sad tretiraju OLAP kocke kao materijalizirane view-ove (sumarne tablice). CBO (Cost base optimizer) je sad sposoban raditi query rewrite s OLAP kockama na isti način kao što je to dosad radio s klasičnim sumarnim tablicama (materijaliziranim view-ovima).

Oracle OLAP sad automatski izlaže dimenzije i kocke kao view-ovi unutar star scheme. Ovime čak i klasične aplikacije mogu putem običnog SQL-a pristupiti OLAP podacima i raditi složene kalkulacije nad multidimenzijskim podacima. S posljednjom verzijom Oracle baze (11g), postupak potpune integracije Oracle relacijskog i OLAP dijela baze je 15 godina nakon akvizicije Express-a završen. Na slici 2 je prikazana OLAP arhitektura u 11g verziji baze.

Figure 2 Relational and Multidimensional Data in Oracle Database 11g



Slika 2: Prikaz OLAP arhitekture u Oracle 11g verziji baze

Danas je Oracle OLAP 11g jedini OLAP server u potpunosti integriran u samu jezgru baze podataka. Sve prednosti ovakve arhitekture donose slijedeća poglavlja. Iako je povijest Oracle OLAP-a u ovom odlomku prikazana prilično površno, i ovakav prikaz je dovoljan da se shvati da je u ovom referatu riječ ne samo o prvom OLAP serveru, već i o jednom od najduže živućih softverskih proizvoda uopće, koji se nadograđivao kroz protekla desetljeća. Na slici 1 i 2 je prikazana arhitektura Oracle OLAP-a i njeno mjesto u Oracle bazi.

#### 4. OLAP - TIPIČNA PODRUČJA PRIMJENE

Područje primjene OLAP tehnologije je svugdje gdje se traži napredna analiza (time-series analysis, non-additive calculation, financial calculations, statistical forecasting, modelling...) ogromne količine podataka. Primjeri primjene OLAP tehnologije su:

- Odjel prodaje i marketinga (proračun profitabilnosti proizvoda, profiliranje kupaca, analiza distribucije, performance/efikasnost prodaje, praćenje prodaje, analiza promocija, predviđanje prodaje...)

- Odjel financija (financijsko modeliranje, upravljanje i praćenje novčanih tokova, buĝetiranje, konsolidacija...)
- Odjel plana i analize (analiza investicija/akvizicija/reorganizacija, dugoročno planiranje, alokacija resursa, planiranje kapaciteta, planiranje ljudskih resursa...)
- Rukovodstvo kompanije (analiza trendova, praćenje ključnih pokazatelja, analiza konkurentnosti...)

Oracle OLAP daje mogućnost napredne analize podataka (korištenjem value set-ova, dimenzija, hijerarhija, relacija i sl.) koji se nalaze u skladištu podataka, kao i kreiranje naprednih kalkulacije nad podacima. Kako se Oracle OLAP instalira kao dio Oracle baze podataka, tako je i njegova najčešća primjena upravo u okruženjima u kojima se i skladište podataka također nalazi u Oracle bazi. Drugi (samostojeći) OLAP serveri (Oracle/Hyperion Essbase, Cognos, Microsoft Analysis Server...) se uglavnom koriste u heterogenim okruženjima koja trebaju podržati niz različitih baza podataka i tehnologija. Kako je Oracle-ov udio na tržištu skladišta podataka veći od 33% s snažnom tendencijom rasta, Oracle OLAP opcija će ipak za veliki dio korisnika biti itekako interesantna.

## **5. OPTIMIZACIJA SKLADIŠTA PODATAKA KORIŠTENJEM ORACLE OLAP 11G**

Premda većina pod pojmom optimizacija skladišta podataka podrazumijeva smanjivanje vremena odziva (response time) prilikom pokretanja izvještaja i raznih SQL upita prema skladištu podataka, pojam optimizacije skladišta podataka je bitno veći. U slijedećim poglavljima će biti opisani razni aspekti optimizacije skladišta podataka kao posljedica korištenja Oracle OLAP 11g.

### **5.1. Izrada sigurnosnih kopija multidimenzijskih podataka**

Prvo što bi svakom administratoru koji se brine za izradu sigurnosnih kopija podataka unutar kompanije (backup administrator) trebalo biti na pameti je na koji način napraviti sigurnosnu kopiju (backup) podataka.

Kako je dosad u više navrata napomenuto, Oracle OLAP je jedini OLAP server koji „živi“ unutar baze podataka. Svi ostali analitički OLAP serveri koji se nalaze na tržištu su odvojeni od baze podataka kao samostalni/samostojeći serveri.

To u stvarnosti znači da dok ostali OLAP serveri zahtijevaju izradu posebnih procedura za pohranu sigurnosnih kopija (backup) podataka, pohrana Oracle OLAP podataka se odvija istovremeno s pohranom svih ostalih podataka koji se nalaze u Oracle bazi korištenjem standardnog Oracle-ovog alata za tu namjenu (RMAN).

Zbog izvrsnih performansi kojom RMAN izrađuje sigurnosne kopije podataka (paralelizam) te kompresije podataka, optimizira se vrijeme trajanja izrade sigurnosnih kopija podataka kao i prostor na disku. Pritom ne treba zaboraviti da se takvim centraliziranim pristupom čitava procedura za izradu sigurnosnih kopija podataka značajno pojednostavnjuje, ona postaje robusnija, a troškovi održavanja se smanjuju.

### **5.2. Zaštita OLAP podataka**

Druga posljedica objedinjavanja relacijske i multidimenzijske baze unutar jednog proizvoda je svakako postavljanje zaštite podataka. Kako je Oracle OLAP 11g dio Oracle baze podataka, postavljanje sigurnosti pristupa OLAP podacima se ne razlikuju standardnom postupku postavljanja sigurnosti pristupa svim ostalim relacijskim podacima unutar baze podataka.

Kompletna zaštita podataka se postavlja na jednom mjestu a to je Oracle baza podataka, što znači da nema dupliciranja sigurnosnih postavki. To značajno, napose u većim kompanijama, pojednostavnjuje postupak administriranja sigurnosnim postavkama, dok je cjelokupan sustav sigurnosti robusniji.

Optimizacija se sastoji u vremenu potrebnom za upravljanjem i održavanjem sustava sigurnosti, kao i vremenu potrebnom za usvajanjem novih znanja administratora OLAP servera. Dok u slučaju Oracle OLAP 11g, administrator baze podataka koristi stečeno znanje da bi postavio/održavao na jednom mjestu jedinstven sustav sigurnosti pristupa i OLAP i relacijskim podacima, OLAP

administratori samostojećih OLAP servera moraju najprije savladati načine postavljanja sigurnosti na tim serverima, da bi potom najčešće duplicirali prava pristupa OLAP podacima iz skladišta podataka.

Time je sustav sigurnosti ranjiviji (moguće su dvije točke upada u sustav umjesto jedne kao u slučaju Oracle OLAP-a), dok je održavanje i upravljanje takvim sustavom zbog dupliciranog sustava složenije, vremenski zahtjevnije i podložno greškama.

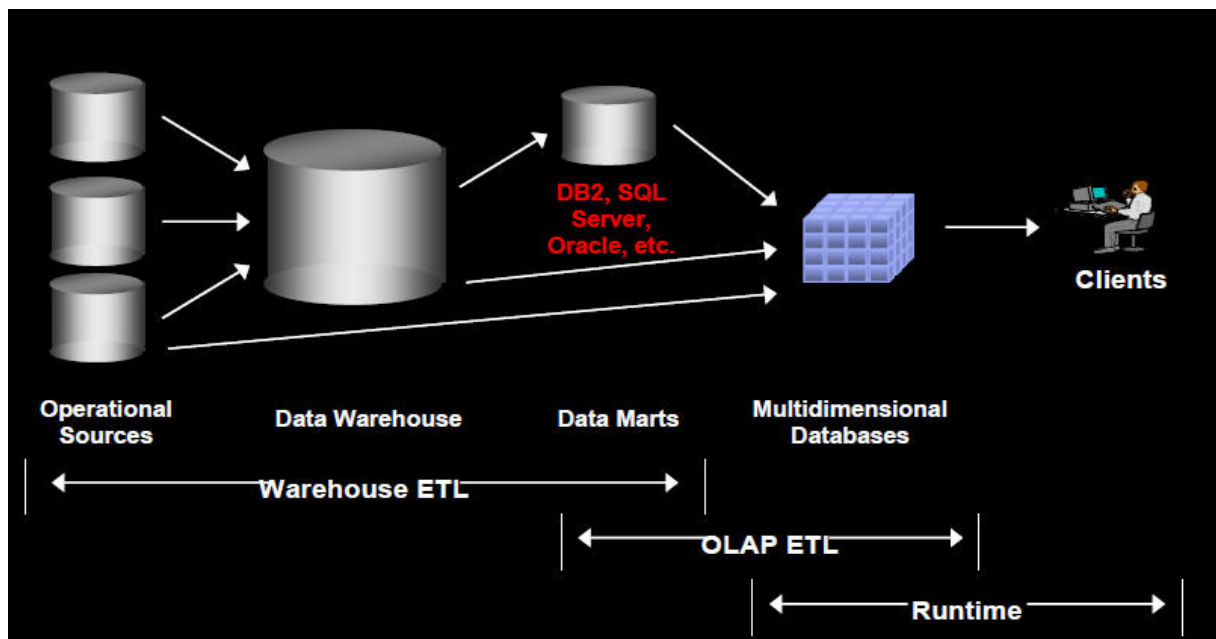
### 5.3. Pojednostavnjivanje procedure ETL-a i punjenja OLAP kocki

Prilikom izgradnje skladišta podataka, vremenski je najzahtjevniji postupak izrade procedura za punjenje skladišta podataka svježim podacima (najčešće iz relacijskog OLTP izvora podataka).

Ukoliko se koristi samostojeći OLAP server, tada je potrebno izraditi i separatnu proceduru (OLAP ETL) koja će, nakon što je skladište podataka osvježeno s novim podacima, napuniti i OLAP kocke sa svježim podacima.

U slučaju Oracle OLAP 11g, postupak osvježavanja je iznimno jednostavan i svodi se na osvježavanje kocki pozivom standardnog Oracle package-a (dbms\_cube.refresh\_mview).

Optimizacija skladišta podataka se u ovom slučaju sastoji od pojednostavnjivanja osvježavanja OLAP kocki budući da je cijeli postupak osvježavanja i skladišta podataka i OLAP kocki jedan jedinstveni proces (a ne dvije separatne procedure kao u slučaju samostojećih OLAP servera), dok je brzina osvježavanja bitno veća jer nema seljenja podataka sa skladišta podataka na separatan OLAP server, dok sama OLAP tehnologija omogućuje puno efikasnije osvježavanje podataka. Na slici 3 je prikazan klasičan ETL koji se sastoji od dva odvojena procesa: jedan ETL za DW a drugi za OLAP.



Slika 3: Klasičan ETL u kojem je osvježavanje OLAP kocki odvojeno od punjenja DW-a

### 5.4. Smanjenje hardverskih zahtjeva

Kao još jedna posljedica korištenja Oracle OLAP 11g je svakako i optimizacija u pogledu hardverskih zahtjeva potrebnih za rad sustava poslovne inteligencije.

Kako je Oracle OLAP server sastavni dio Oracle baze, nema potrebe za kupovinom dodatnog servera što je nužnost za samostojeće OLAP servere. Pored toga korištenjem Oracle OLAP-a je moguće puno efikasnije podesiti i sustav za pohranu podataka (storage) nego što je to slučaj sa samostojećim serverima.

Kako cijena servera i sustava za skladištenje podataka pogotovo u srednjim i velikim tvrtkama nije nezanemariva, optimizacija se u ovom slučaju postiže efikasnijom upotrebom postojećih resursa (servera i skladištenja podataka).



## 5.5. Mogućnost pristupa OLAP podacima direktno iz SQL alata

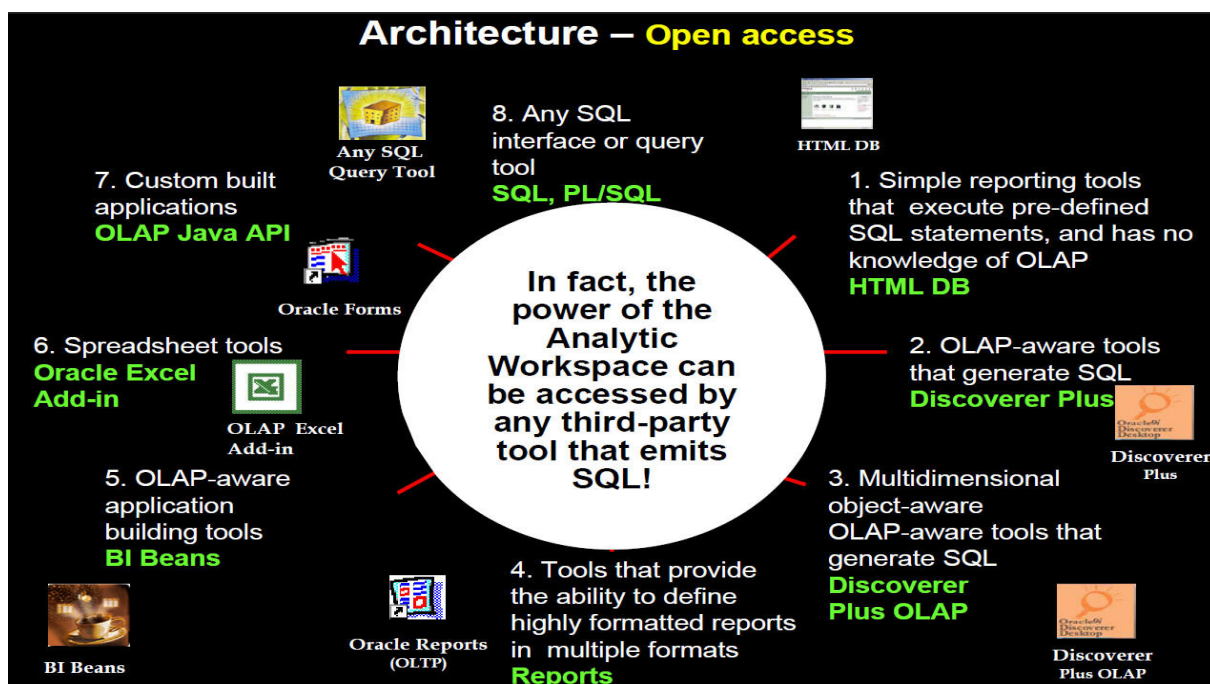
Ovo je zasigurno jedna od najvećih prednosti Oracle OLAP 11g pred svim ostalim konkurentima na tržištu danas.

Kako je u točki 3 (Povijest Oracle OLAP-a) navedeno, Oracle je u 10g verziji baze prvi put omogućio pristupanje OLAP multidimenzijskim podacima direktno iz SQL-a. To u praksi znači da će prednosti OLAP tehnologije osjetiti sve aplikacije koje mogu pristupiti Oracle bazi podataka i pokrenuti SQL upit nad njom poput Oracle BI Publisher-a, Oracle Reports Builder, Cognos, SAP Business Objects, Cristal Reports, TOAD for Oracle, Oracle APEX (besplatni alat), pa čak i najobičniji SQL\*Plus.

Time je i povrat investicije za kupovinu Oracle OLAP opcije zbog takve širine primjene brži u odnosu na samostojeće OLAP servere. Isto tako, treba naglasiti da za korištenje OLAP servera nije potrebna kupovina skupih specijaliziranih BI alata, što u konačnici također smanjuje ukupnu investiciju u OLAP.

Simba Technologies je surađujući s Oracle-om razvila Simba MDX Provider for Oracle OLAP, čime je omogućeno korištenje alata poput Microsoft Excel Pivot tablice za pristupanje Oracle OLAP-u te izradu interaktivnih izvještaja (Simba MDX for Oracle OLAP je tema drugog referata istog autora na ovogodišnjem HROUG-u).

Ovu mogućnost, kao posljedica integracije OLAP servera u jezgru Oracle baze, konkurentski OLAP serveri neće moći postići niti u daljnjoj budućnosti. Na slici 4 je prikazana mogućnost pristupa i korištenja Oracle OLAP-a iz raznih (a ne samo specijalističkih) alata.



Slika 4: pristupanje OLAP podacima korištenjem raznih alata

## 5.6. Korištenje kompleksnih kalkulacija kao posljedica direktnog pristupa OLAP podacima iz SQL-a

Kako je u poglavlju 4 (Tipična područja primjene OLAP-a) navedeno, jedno od glavnih prednosti OLAP-a općenito je i bogatstvo i kompleksnost kalkulacija koje je OLAP server sposoban izvesti nad podacima spremljenim u složenim multidimenzionalnim strukturama.

Kako je Oracle omogućio pristup dimenzijama i OLAP kockama direktno iz SQL-a, te kako su izračunate mjere (calculated measures) sastavni dio OLAP kocki, implicitno proizlazi da su sve SQL aplikacije dobile mogućnost korištenja složenih mjera koje ne postoje u relacijskom dijelu baze. Time je ionako bogat Oracle-ov SQL značajno proširen.

## 5.7. Pojednostavnjivanje SQL upita kao posljedica direktnog pristupa OLAP podacima iz SQL-a

Iako je Oracle-ov SQL jezik vodeći u industriji, zbog njegove primarne namjene za rad s relacijskim podacima, nije se pokazao pogodnim za izradu složenih analitičkih upita nad podacima.

Tijekom posljednjih 10-ak godina, Oracle se trudio popraviti situaciju u tom pogledu uvodeći brojne analitičke funkcije (model, lag, lead i sl.).

Iako su analitičke funkcije veoma moćne (napisane su knjige koje se bave analitičkih funkcija), ipak su upiti koji su ih koristili postajali kompleksni i teško razumljivi. Pored toga, i nadalje je ostao veliki jaz između mogućnost kalkulacija OLAP-a i SQL-a.

S druge strane, kako je OLAP tehnologija primarno namijenjena za analizu složenih podataka, upiti koji koriste OLAP tehnologiju su bitno jednostavniji i razumljiviji u odnosu na analogne SQL upite.

Pored toga, kako su podaci spremeni u OLAP-u optimizirani za složene ad-hoc upite, dohvaćanje multidimenzijskih podataka je bitno brže u odnosu na analogne SQL upite.

Time su, koristeći pristup OLAP-u direktno iz SQL-a (ili iz bilo kojeg alata koji se može spojiti s Oracle bazom), dobivene dvije koristi: upiti koji koriste OLAP su bitno brži, a njihova sintaksa je značajno jednostavnija. Kako je na slici 5 prikazano, korištenjem OLAP-a se bitno smanjuje kompleksnost upita.

<pre>SELECT [breakout columns],        sales,        sales_prior_year        sales_ytd,        sales_ytd_prior_year FROM sales_cube_view WHERE [star join]</pre>	<pre>WITH sales_dense AS   (SELECT [breakout columns]    sales,    SUM(sales) over(PARTITION BY [breakout columns]    ORDER BY [time column] ASC range BETWEEN unbounded    preceding AND CURRENT ROW) AS sales_ytd  FROM   (SELECT [breakout columns]    a.sales   FROM    (SELECT [breakout columns]     SUM(f.sales) sales    FROM [table list]    WHERE [star join and other filters]    GROUP BY [breakout columns])   a PARTITION BY(breakout columns)   RIGHT OUTER JOIN    (-- need list of all time periods    SELECT DISTINCT [time columns]    FROM time_dim    b ON([join on relevant time level]))   ) - Continued...</pre>
--	--

Slika 5: Pojednostavnjivanje SQL upita korištenjem OLAP-a

## 5.8. Cube organized materialized views

Oracle je tijekom godina sve čvršće integrirao Express multidimenzijsku bazu u svoju vodeću relacijsko/objektnu bazu podataka.

S verzijom 11g Oracle baze, korisnici su dobili još jednu jedinstvenu funkcionalnost - izlaganje OLAP kocki kao da se radi o sumarnim tablicama (materijaliziranim view-ovima) relacijskom dijelu Oracle baze.

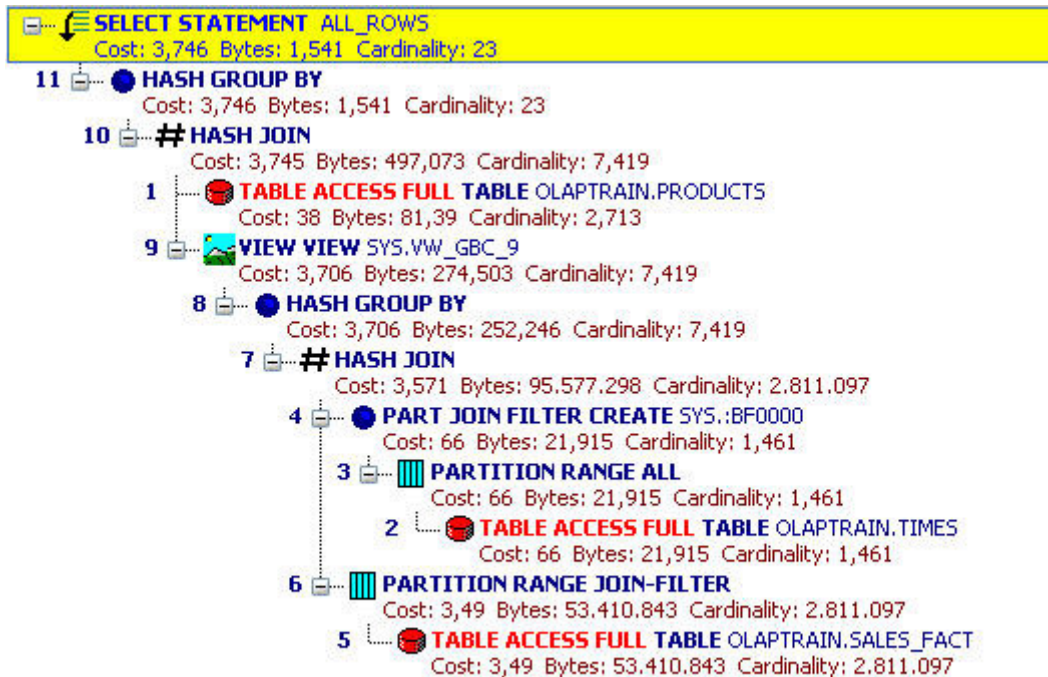
Ovo je možda, uz izlaganje multidimenzijskih objekata klasičnom SQL-u, vjerojatno najvažnija funkcionalnost Oracle OLAP-a, koja značajno proširuje njegovo područje primjene na dio koji je donedavno bio rezerviran isključivo za klasične sumarne tablice (materijalizirane view-ove).

Oracle CBO (cost base optimizer) od 11g verzije baze prilikom generiranja planova izvršavanja SQL upita, uzima u obzir i postojanje OLAP kocki kao mogući izvor sumarnih podataka.

Ukoliko CBO prilikom generiranja plana izvršavanja za neki SQL upit procijeni da mu je „jeftinije“ (u smislu potrošnje resursa, brzine i efikasnosti) uzeti neke podatke iz OLAP kocki, on će to i napraviti. U daljnjem tekstu bit će navedene glavne prednosti ove nove funkcionalnosti.

### 5.8.1. Query rewrite s OLAP kockama

Sumarna tablica bez mogućnosti query rewrite-a ne bi imala nikakvog smisla. Stoga je Oracle morao riješiti cijeli niz problema kako bi query rewrite radio na isti način kao i klasičan query rewrite s relacijskim sumarnim tablicama. Korisnik, osim velike brzine dobivanja traženih podataka, nije uopće svjestan da je u pozadini Oracle napravio query rewrite koristeći podatke iz OLAP kocke, jer se cijeli proces query rewrite-a odigrava za korisnika potpuno transparentno. Na slici 6 je prikazan klasičan plan izvršavanja SQL upita, dok je na slici 7 prikazan plan izvršavanja istog SQL upita korištenjem sumarnih podataka iz OLAP kocke (cube query rewrite).



Slika 6: Plan izvršavanja bez query rewrite-a



Slika 7: Plan izvršavanja korištenjem podataka iz OLAP kocke (cube query rewrite)

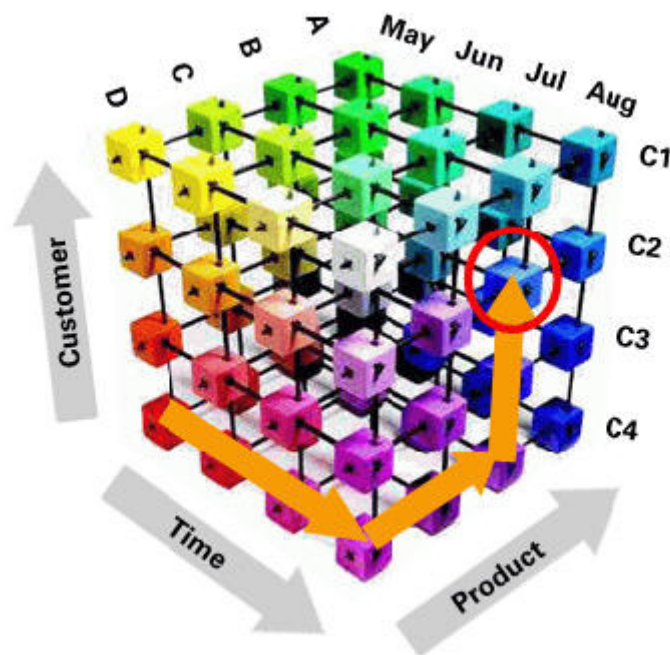
### 5.8.2. Veliki broj kombinacija u kojima dolazi do query rewrite-a

Kombinacija tehnologije query rewrite-a i OLAP-a omogućilo je Oracle-u da značajno poveća broj situacija u kojima dolazi do query rewrite-a.

Jedan od najvećih mana korištenja klasičnih sumarnih tablica (materijaliziranih view-ova) je limitiran broj SQL upita u kojima dolazi do query rewrite-a.

Kako OLAP kocke imaju veći broj dimenzija (u praksi najčešće od 5 do 7 dimanzija), te kako su podaci agregirani na različitim nivoima (level u OLAP terminologiji), broj kombinacija u kojima dolazi do query rewrite-a je neusporedivo veći u odnosu na klasične sumarne tablice. Time se ubrzavaju i većina drugih upita pored onih najkritičnijih za koje bi se inače kreirale sumarne tablice, što općenito ubrzava vrijeme odziva čitavog BI sustava. Slika 8 na slikovit način prikazuje način pohrane podataka u OLAP kockama.





Slika 9: Slikovit prikaz spremanja i dohvaćanja podataka unutar OLAP kocke

### 5.8.3. Smanjeno zauzeće mjesta na diskovima

Sumarne tablice (materijalizirani view-ovi) za razliku od običnih view-a zauzimaju diskovni prostor. Kako jedna OLAP kocka (kako je opisano u točki 5.8.2) može zamijeniti desetine materijaliziranih view-ova, jasno je da korištenjem OLAP materijaliziranih view-ova dovodi i do značajnog smanjenja zauzeće diskovnog prostora.

### 5.8.4. Smanjeno vrijeme i jednostavnost osvježavanja kao posljedica korištenja OLAP tehnologije

U praksi je u najvećem broju slučajeva potrebno držati sumarne tablice (materijalizirane view-ova) ažurnim, kako prilikom query rewrite-a ne bi dobili netočne podatke.

Usprkos mnogim unapređenjima, osvježavanje sumarnih tablica, pogotovo kad postoji njihova međuovisnost, je zahtjevno kako vremenski tako i u pogledu zauzeća resursa te redosljeda osvježavanja.

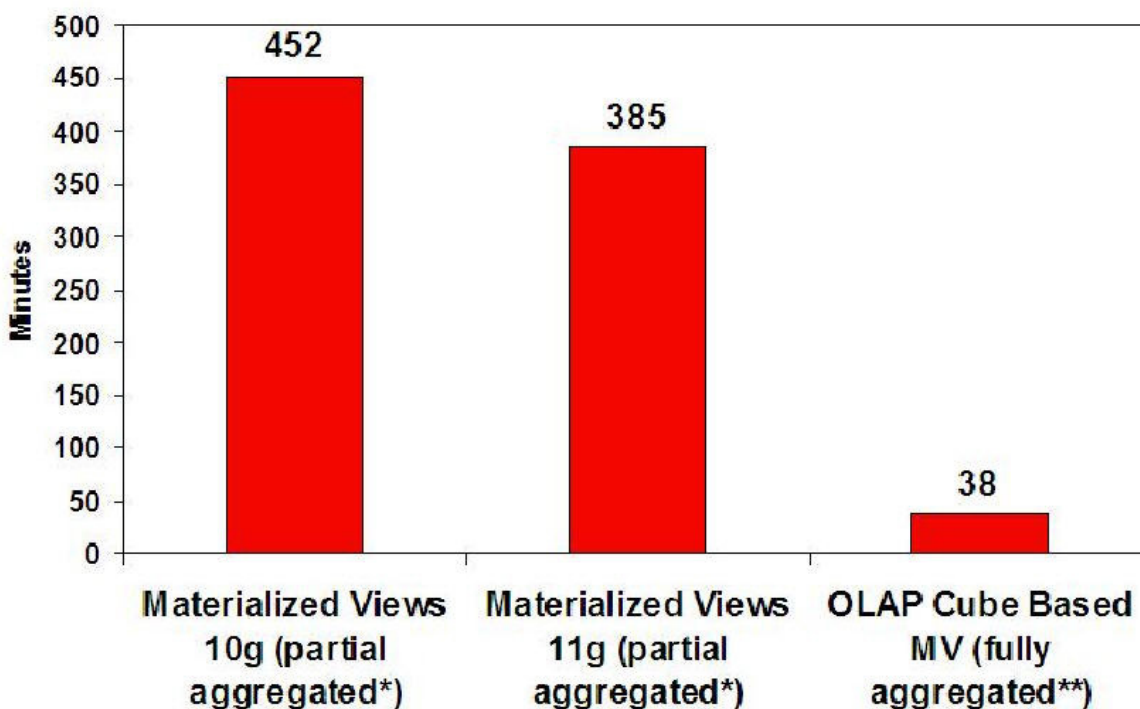
Vrijeme potrebno za osvježavanje OLAP materijaliziranih view-ova je bitno kraće dok je sam postupak osvježavanja bitno jednostavniji u odnosu na klasične relacijske sumarne tablice, jer Oracle ne mora paziti na redosljed osvježavanja u slučaju da su materijalizirani view-ovi međusobno ovisni jedan o drugom.

Na slici 10 dan je prikaz brzine osvježavanja sumarnih tablica (posebno za 10g i 11g verziju baze), kao i za OLAP materijalizirani view.

Iz slike je vidljivo da je 11g verzija baze donijela napredak u smislu poboljšavanja performansi osvježavanja klasičnih sumarnih tablica. Ipak, vrijeme potrebno za osvježavanje OLAP materijaliziranih view-ova je za red veličine manje u odnosu na klasične materijalizirane view-ove. Do tako velike razlike u performansama osvježavanja u korist OLAP materijaliziranih view-ova je došlo ne samo zbog smanjivanja broja potrebnih klasičnih sumarnih tablica, već i zbog multidimenzijske strukture podataka u OLAP kockama. Svi načini osvježavanja klasičnih sumarnih tablica su dostupni i OLAP materijaliziranim view-ovima (full, fast refresh i partition change tracking).

Iz svega navedenog može se zaključiti da je korištenje OLAP materijaliziranih view-ova itekako primamljivo i u pogledu performansi osvježavanja.

The graph below shows the results for loading and aggregating relational views and cube based materialized views for the 10g and 11g database benchmarks:



- As shown above, build times were significantly faster when using OLAP 11g. Compared to 10g and improved 11g relational materialized views, a single CPU cube built in 38 minutes (without using available parallel processing options). Other parallel tests performed the same task in only 10 to 15 minutes. By comparison, it took at least 11x longer to build partially aggregated relational materialized views on the same database server instance.

Slika 10: Usporedni prikaz brzine osvježavanja klasičnih materijaliziranih view-ova za 10g i 11g verziju baze u odnosu na OLAP materijalizirane view-ove

## 6. OLAP DML programski jezik

OLAP DML programski jezik je od samih početaka Express OLAP servera bio glasovit posvojim mogućnostima i performansama, koje su se tijekom desetljeća razvoja neprestano povećavale.

U zadnjoj verziji Oracle dokumentacije vidljivo je da Oracle OLAP DML reference knjiga ima otprilike 100 stranica više (1598 stranica) od glasovite Oracle SQL Language reference dokumentacije (1508 stranica), iz čega i površni poznavatelji Oracle OLAP-a naslućuju da se radi o iznimno moćnom programskom jeziku.

Naziv programskog jezika ("OLAP DML") je pomalo zbunjujući jer upućuje na to da se radi o jeziku prvenstveno namijenjenom za manipuliranje podacima (DML u svijetu relacijskih baza podataka je akronim za data manipulation language).

U stvarnosti OLAP DML može izvršavati DML (insert, update, delete), DDL (create, alter, drop), DCL (grant, revoke) i SELECT naredbe.

OLAP DML je specijalno dizajniran programski jezik za izvođenje analitika na multidimenzionalnom setu podataka.

OLAP DML ne samo da manipulira podacima, već za razliku od SQL-a i sprema podatke unutar samog jezika (varijable, valueset-ovi...).

Glavni razlog korištenja OLAP DML-a su njegove fantastične mogućnosti koje je jedino koristeći DML programski jezik moguće do kraja iskoristiti.

Primjeri upotrebe su izrada kompleksnih izvještaja koje nije moguće dobiti niti jednim grafičkim alatom, izrade kompleksnih programa, izrade složenih kalkulacija te korištenja dimenzija/tipova podataka koje nije moguće korištenjem AW (Analytic Workspace) Manager-a kao grafičkog alata za administraciju Oracle OLAP Analytical Workspace-a.

OLAP DML izvrsno nadopunjuje ionako bogate osnovne mogućnost OLAP 11g servera s naprednim funkcionalnostima.

## 7. ZAKLJUČAK

Oracle je sa najnovijim OLAP 11g serverom, tijesno integriranim u jezgru baze, natjerao sve one koji dosad nisu razmišljali o toj opciji, da još jednom preispitaju svoju odluku da je i nadalje ne koriste.

Oracle OLAP 11g će posebno biti interesantan onima koji već imaju skladište podataka (DW) u Oracle bazi, a to je prema Gartnerovim podacima najveći postotak korisnika (više od 33%).

Područje iskoristivosti Oracle OLAP 11g je daleko veće od konkurentskih proizvoda koji su odvojeni od baze kao zasebni, samostojeći (stand-alone) OLAP serveri, jer pored ad-hoc OLAP upita koristeći skupe specijalističke BI alate, omogućuje i pristup OLAP serveru iz bilo kojeg alata koji se može spojiti na Oracle bazu (od SQL\*Plus-a, TOAD-a, SQL Developer-a, pa sve do Oracle Reports-a, Cognos-a, SAP Business Objects-a, Oracle BI Publisher-a, OBIEE itd).

Time Oracle OLAP ne samo da je neusporedivo fleksibilniji od drugih rješenja koja ne nude integraciju SQL-a i OLAP, već je područje eksploatiranja OLAP tehnologije znatno veće, što doprinosi ukupno boljem iskorištenju podataka i bržem povratu investicije.

Uz pojednostavnjivanje i ubrzavanje procesa izrade sigurnosnih kopija (backup-a), centraliziranom zaštitom podataka na jednom mjestu (u samoj bazi), značajnom povećanju brzine osvježavanja OLAP kocki u odnosu na brzinu osvježavanja klasičnih materijaliziranih view-ova, te pojednostavnjivanjem ETL-a (ne zahtjeva se zaseban ETL za punjenje skladišta podataka i zaseban ETL za punjenje OLAP kocki), te uz Cube materijalizirane view-ove jedinstvene samo za Oracle OLAP, ne treba čuditi da je interes za OLAP serverom u značajnom porastu.

Oracle je nakon akvizicije Express-a prije 15-ak godina jasno pokazao smjer kojim će nastaviti razvoj Oracle OLAP-a, a to je put njegove daljnje integracije s Oracle bazom. Ako se svemu tome pridoda i rudarenje podacima (data mining), koji je također implementiran u samoj jezgri Oracle baze, te OLAP DML iznimno moćan programski jezik Oracle OLAP 11g servera, treba samo malo mašte da bi se sve tehnologije poslovne inteligencije kombinirale u jedan hibridni sustav bez premca.

Uz sve to, ne treba zaboraviti da je Oracle OLAP (nekadašnji Express) jedan od najduže živućih softverskih proizvoda uopće, koji se kroz posljednjih 40 godina života unapređivao da bi danas postao to što jest.

Uz značajne financijske prednosti pred drugim rješenjima i arhitekturama, sigurno je da ćemo sve češće gledati rješenja bazirana na Oracle OLAP tehnologiji. Jedno od takvih rješenja je Simba MDX provider for Oracle OLAP, koji (uz standardna Oracle-ova rješenja poput OBIEE i Discoverer plus OLAP...) omogućuje direktno povezivanje Microsoft Excel Pivot tablice s Oracle OLAP 11g serverom.

Upravo to je tema druge prezentacije istog autora na ovogodišnjem HROUG-u pod nazivom: Poslovno izvještavanje korištenjem Excel Pivot tablice i Oracle OLAP-a 11g.

Za nadati se da će i ozbiljna hrvatska poduzeća prepoznati iznimnu korist koje im ovakvo jedno rješenje može donijeti u svim područjima u kojima OLAP briljira počevši od upravljanja logistikom, financijskim izvještavanjem, analizom kvalitete proizvoda pa sve do poslovnog izvještavanja i analizom prodaje i marketinga.

U današnjim okolnostima od iznimne je važnosti imati mogućnost efikasnog upravljanja vlastitim resursima putem izrade interaktivnih složenih izvještaja. Brzina kojom OLAP dohvaća podatke je impresivna (u literaturi se često uspoređuje s brzinom misli), te omogućuje svim dijelovima kompanije izrade tisuća složenih izvještaja tijekom radnog vremena. Na taj je način čak i u velikim tvrtkama, bez obzira na eksploziju količine podataka koju treba obraditi, moguće zadržati kontrolu nad svim segmentima poslovanja.