

Key Differentiators

What sets Ideal Analytics apart from traditional BI tools

Ideal-Analytics is a suite of software tools to glean information and therefore knowledge, from raw data. Self-service, real-time, on-demand ad-hoc analysis and high performance exploration functionality with plug-ability, scalability & security, available in both SaaS and on-premise model

1 On-Demand Self-Serving Analytics

Ideal Analytics incorporates an on-demand analytics methodology. We have moved away from traditional data-warehousing concept of transforming and loading transactional data from OLTP (or other data sources) to OLAP, and querying the data-marts with proprietary languages like MDX. Instead, we just load (without transforming) the transactional data into columnar storage and create an interactive and complete visualization over the aggregation of the transactional data with the help of standard SQL queries. This approach has the following advantages.

- ✓ We do not need any analysis storage (like OLAP store) and are not dependent on specific technologies like MDX. Instead we create interactive visualizations with the help of SQL queries, but without compromising performance as compared to OLAP store, by using columnar stores.
- ✓ There is no design phase for analytics views, as opposed to traditional data-warehousing tools, which need customized design models (created by tool consultants) to create the views on. In Ideal Analytics, the users themselves (administrative as well as end users) can configure the data-sources and interactive views are created off-the-shelf.
- ✓ Another advantage of this approach is micro batching for frequent incremental update, which is practically impossible for traditional systems with OLAP storage for data warehousing.

2 Large Data Handling

Most of the existing BI tools in the market has the limitations on volume of data and analysing those in real-time. Ideal Analytics wins the race at ease in this aspect and can analyse millions of records on-demand with its unique data loading optimization technique.

The large volume of data handling capability has been one of the key parameters from the design phase itself and it has been taken care of at various layers, by various means; that we explain next.

2.1 Data Storage Layer: Warehouse

From storage perspective, the major concerns, with respect to large data, is two-fold, firstly the query performance and secondly the storage space. We wanted to have a data warehouse, which will provide lightning performance and also storage advantage, and the solution to this was columnar database. Columnar databases provide 100-1000 times faster responses on select queries, and also offers excellent data compression capabilities. We have evaluated quite a number of columnar databases available in the market on the following criteria:

- ✓ Query Performance on millions of synthetically generated data
- ✓ Capability of handling number of concurrent queries
- ✓ Actual storage space requirement and the percentage of compression achieved
- ✓ The loading mechanism and the time taken for loading millions of synthetically generated data
- ✓ SQL interface provided by the database – this is important from ad-hoc analysis perspective, because query needs to be generated dynamically depending on the analysis requirement

With the above criteria, we evaluated Infobright, Infinydb, Exasol, Sybase IQ, MonetDB, Vertica and Oracle Exadata. The result went in favour of Infobright, and we chose them as our default db provider. **Although we are not tightly coupled with any vendor and can connect to other databases if required in specific cases, very easily.**



2.2 Data Processing Layer: Middleware

In the data processing layer, we have our proprietary “Intelligent Caching Window” mechanism, which practically makes the performance independent of data volume. We mix the techniques of intelligently formed SQL queries and data buffering by which we are able to apply the business logic to whole data but send a subset of that to the client.

2.3 Data Visualization Layer: RIA Frontend

Handling large data in web-based frontend applications pose two challenges:

1. Large volume of data pushed into the browser memory causes the browser process either to hang or crash
2. From visualization perspective, viewing lots of data points on the screen makes the view clumsy and practically incomprehensible

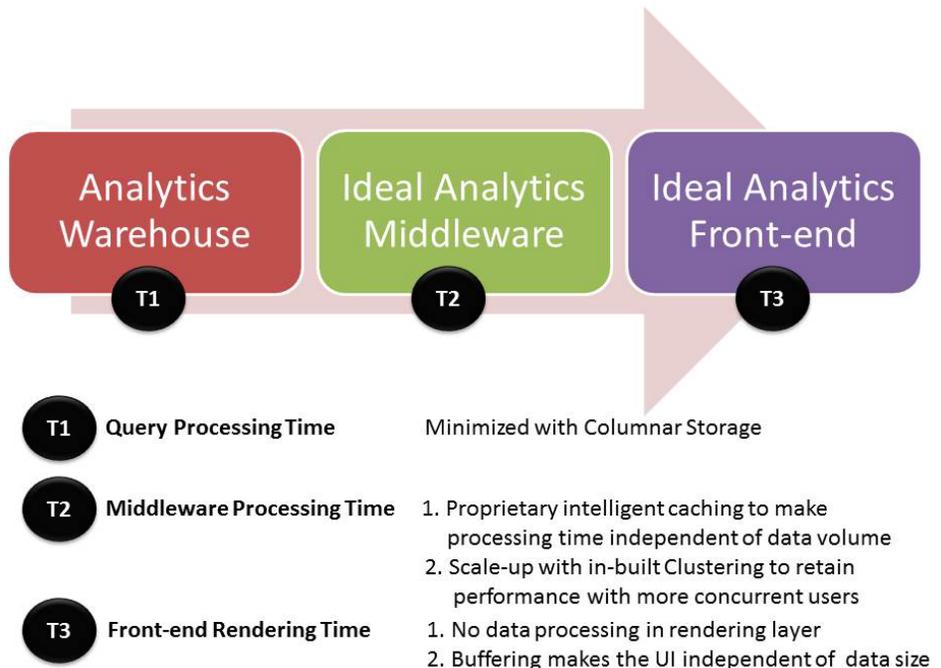
For these two limitations, most of the web-based BI products available in the market either poses a restriction in the data volume, or shows a serious amount of degradation in their performance, when they are made to work with large volume of data. In the frontend of Ideal-Analytics, we have solved the two problems in the following manner:

1. The RIA is tailor made to play with the middleware’s “Intelligent Caching Window” mechanism so that the browser is never over loaded with terabytes of data.
2. The visualization has also been made specially suitable so that users can select the amount of data they want to see, at a time on the screen, and for viewing more they will scroll smoothly. The scroll data, most of the time comes from the buffer window instead of being queried from the db, thus providing a smooth user experience even with very large volume of data.

3 Performance

Ideal Analytics provides lightning-fast performance with the help of:

1. Columnar storage for analytics data for quick data retrieval
2. Unique data loading optimization technique to minimize the data traffic
3. In-built query optimization engine with intelligent caching mechanism



4 Data Load and View Update Strategy

Ideal Analytics has the capability to load data to its column storage both on-demand (pull) and automatically by receiving the data (push).

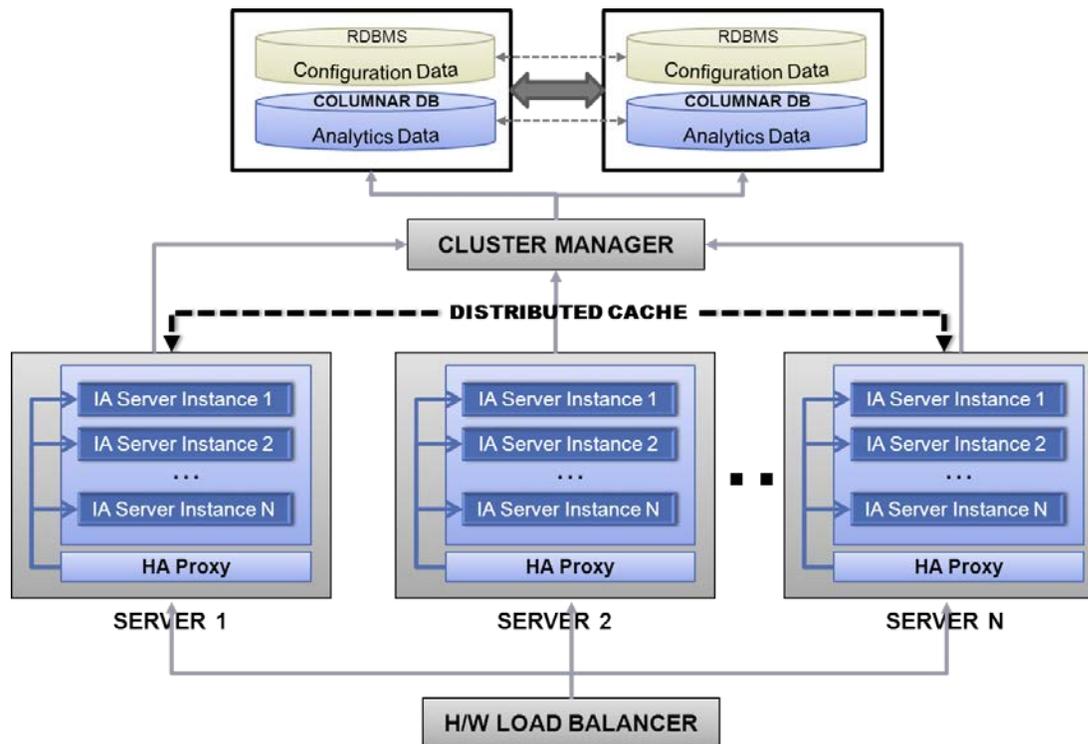
- ✓ Pull strategy: This methodology is primarily used to reload the complete data-set. In case, the backing data-source allows installation of custom agents (e.g. a database agent installed in the backing RDBMS, which has the capability of providing incremental data-set), this strategy will cater for on-demand incremental data load.
- ✓ Push strategy: In case, the backing data-store allows for installation of real-time synchronizing agents, ideal Analytics can listen for updates in the datasets and update the data-sets on a real-time basis.

The Ideal Analytics view has the capability of updating the view automatically, in case the underlying data-set is updated (as mentioned above). The view update frequency can be configured while configuring the dataset, according to the nature of the dataset.

5 Enterprise Scalability

Ideal Analytics servers are linearly scalable. Ideal Analytics instances are mostly stateless, backed by centralized and replicable instances of RDBMS storage and column storage and fronted by High availability proxy server and hardware load balancers. The only state-full nature of the server is the authentication state, i.e. login sessions, which are by default distributed cache enabled. So, a new Ideal Analytics server instance can be added to the existing cluster seamlessly, to scale up the performance.

Following is the deployment architecture:



idealanalytics

Analytics On-Demand

www.ideal-analytics.com

Self-service, real-time, on-demand ad-hoc analysis and high performance exploration functionality with plug-ability, scalability & security, available in both SaaS and on-premise model



contact@ideal-analytics.com

>> Contact Us

Office in India

202 SDF Building
Sector V, Salt Lake City
Kolkata - 700091
Tel: **+91 33 2357 6414/15**

Office in France

14 rue Séguier
75006 Paris - France
Tel: **+33 01 53 05 93 75**
Fax: **+33 (0)1 42 66 34 24**