

The face of Business Intelligence has changed dramatically since the advent of Web 2.0 technologies and the proliferation of the NOSQL movement. Some web sites now generate Gigabytes of data every day. Some have grown into the Terabyte and even the Petabyte territories. Having to deal with such enormous amounts of information has led to some radical rethinking in the way data is stored, mined and used. Hence the advent of virtualized storage infrastructures like Hadoop, BigTable and Cassandra which are essentially algorithm driven virtualization environments for storing these oceans of information. This now falls under a realm that is called “Big Data”.

So, how does this relate to Business Intelligence? While companies were in possession of such vast amounts of data, there was always a need for turning this data into useful information and useful information into actionable intelligence. The traditional mechanisms that were in place for gaining such intelligence were Transactional Data Warehouses, Master Data Hubs and Business Intelligence Tools to slice and dice information contained in these data silos.

One of the major challenges has always been the scattered and distributed nature of information within organizations. This was further compounded by M&As where additional such repositories were inherited. Maintaining Data Integrity was always an issue. The advent of Data Cleansing tools like Trillium and Hyperion only helped to a certain extent. While the goal was to gain meaningful intelligence out of this data, all that could be obtained was information. The tried and tested methodology was to run BI tools like Cognos, Business Objects or Oracle BI against these data stores and obtain reports. However, with today’s real-time and agile internet, these reports are becoming obsolete before they even hit the reviewer’s desk.

In came real-time analytics. The future of BI—which I am calling BI 2.0—will be increasingly driven by real-time analytics acting upon non-traditional data stores such as these virtualized environments. The BI landscape has undergone a sea change. In this article we will talk about the various ways in which BI is being perceived and handled in the current IT landscape.

If Traditional BI = Data Warehouse + Reporting Tools

Then BI 2.0 = (Past + Present) Data/Analytics + Future Business Analytics + Reporting Tools.

I have used this simple equation by way of demonstration: BI 2.0 is all about studying and analyzing Past, Present and Future trends within Transactional data (as opposed to Master Data) using archived and real-time data feeds and then turning this data into knowledge for stronger decision-making.

Given today’s requisite business agility, the true and trusted mechanisms of the DW-driven mode of acquiring BI is now considered almost prehistoric. This is not to say that these technologies are obsolete; they certainly are not. They do have their uses. But, this model is

being increasingly supplanted by more agile and real-time mechanisms like Business Activity Monitoring (BAM) and Complex Event Processing (CEP).

BAM is typically tied to real-time monitoring and intelligence gathering with respect to web services. A BAM tool is an important component of a Service Oriented Architecture.

While CEP (Complex Event Processing) is similar to BAM, it is more of a services agnostic technology. It is sort of a real-time message aggregator that can process messages from various different sources including web services. One example is a transaction in which attached to order detail there is included the related stock price information for the companies that produce the ordered products. In this case, this stock detail feed that would normally come from the likes of a Bloomberg or Reuters can be combined with the Order Data and be displayed.

CEP is another emerging trend for real-time analytics. To use an analogy, what mashups does for a Web 2.0 presentation tier, CEP attempts to do at the backend messaging layer by being able to process messages from disparate sources for richer, deeper analysis.

It's Real-Time or It's of No Real Use:

When it comes to Traditional BI tools, the data retrieved from them, in today's parlance, is considered outdated. Intelligence is gleaned after the fact—after significant time has elapsed between the actual activity and the time when it is analyzed in either an individual or aggregated fashion. In this Internet Age, even the lapse of a day could render that data obsolete.

BAM and CEP are real-time tools; data is analyzed as the transactions are being executed. It is this optimally effective decision-making in the moment that makes BI 2.0 so attractive to new age businesses.

BI Technology Trends: The Past

We need not go into the details of how BI functions were being performed or are still being approached using DW-driven technology. If Business needed insights into key trends then they would follow this well beaten path:

- Set up a de-normalized DW or DataMart
- ETL information from transactional data stores into DW
- Use reporting tools like Cognos to generate the required MIS reports.

So this was the tried and trusted method and was considered business as usual. Traditional BI is very batch-oriented in nature and not real-time. However, while real-time mechanisms may look very compelling in terms of their value proposition and ROI, it may not be feasible to implement these quickly as they are dependent on a services-based infrastructure.

Apart from this, certain types of trends may indeed require you to maintain and analyze historical data (such as stock prices) so Traditional BI will continue to be used. Traditional BI itself is not becoming obsolete; rather Traditional BI and BAM/CEP each play a specific role in attaining the true value of BI 2.0 and will, therefore, co-exist even within the same organization.

BI Technology Trends: The Present

As the competitive landscape becomes more intense, the need for real-time business intelligence is almost De Rigueur in certain industries, in order to gain business agility and competitive advantage. With the advent of technologies such as BAM and CEP, it is possible to perform real-time and even future *what if* business analytics.

BAM and CEP fall under a new category called Event Driven Architecture (EDA). These tools permit BI in both synchronous and asynchronous modes.

Business Activity Monitoring (BAM) is associated mainly with Web Services. It is a dashboard driven toolset that is comprised of various design time components and a runtime component. It is used in conjunction with other runtime components of a typical SOA implementation, such as ESBs and BPEL engines.

The usage of a BAM product entails two steps at a minimum. The first step is to design your BAM strategy in terms of what services you want to monitor, the parameters you want to consider, the mechanisms you will institute for monitoring (which is usually through a dashboard), and the actual design of the dashboard itself.

Once your strategy is in place, the next step, of course, is to implement it. Any BAM offering provides a very rich set of design time tools to implement rich color-coded dashboards which can later be deployed into the BAM runtime environment.

CEP (Complex Event Processing) is similar to BAM with the difference that CEP is not tied to web services alone and is message agnostic. You can aggregate disparate messages from myriad sources and use them as part of one CEP application. CEP is almost always event driven.

Possibly unbeknownst to you, you are experiencing the benefits of real-time analytics every time you are on the web. One very common example of this is targeted ads that are delivered to you based on the searches you conduct on sites like Google. Let's say you spend a ½ hour searching for flooring material and installation services. Subsequent activity like checking email or shopping cart activity will cause flooring related ads to show up at designated slots on your browser. All this is happening in real-time; your search characteristics are being analyzed and based on the quality and frequency, ads of vendors who have contracted or bought ad slots for these types of services will be delivered to you.

The online shopping segment is also benefitting a lot from these advanced BI paradigms. The likes of Amazon, eBay and Zappos are rejoicing at the arrival of BAM and CEP. In such high

traffic sites where the hits are measured in millions per day, attaining the level of agility and edge over the competition is virtually impossible without real-time analytics. Traditional BI strategies fall woefully short and are not equipped to deal with these types of business operations.

Take a closer look at a typical real-time business intelligence use case scenario—the shopping cart.

BI Technology Trends: The Future

There are several game changing trends that are in vogue in the BI field that promise to change the face of the industry. While BAM and CEP are the pre-cursors, they are fueling other paradigms that were not possible before.

Here's an example that is taken from advertising industry pundits whose central focus is always on "The Consumer": Smart Billboards that will glean biometric information based on facial features of anyone walking by. If *the consumer* turns out to be a young 20-something year old male, the backend system would beam back relevant information such as ads for an iPod, cell phone, energy drink, sports car or anything that fits this particular demographic.

Just 5 years ago this idea would have seemed outlandish, but not anymore. With real-time BI delivery mechanisms at the foundation, scenarios such as this are becoming reality. With a little ingenuity and BI 2.0, the possibilities are seemingly endless.

The importance of BI as a whole has come into sharp focus due to the increasing adoption of EBI (Enterprise Business Intelligence) tools. With that said, maturity levels being attained in the ERP Software and DW domains, Top Management and executives at the "C" level are demanding more real-time or near real-time intelligence due to the changing competitive landscape.

One case in point is the Advertising Industry where trends have to be tracked on a daily basis. This is due to game changing technology and ad delivery systems brought about by the likes of Google and Yahoo. Now "Joe's Discount Mart" has the same advertising power in his hands as a Target or a K-Mart. The playing field has been leveled as never before. People are increasingly moving away from Print media due to various reasons, chief among them being,

- Time lag between ad printing and delivery (magazines to hit stores) vs Instant ad delivery when you open your browser
- No real-time analytics with print ads, you really don't know how many people looked at your ad vs a browser click tells you that someone has not only looked at your ad but has evinced interest in it by clicking on the ad link.
- Ad may be outdated in a few days' time vs no such risk with online ad delivery, ads can be refreshed every minute if necessary.

This one small example illustrates the need for BI (real-time or historical) in projecting and forecasting business and user trends.

Another major push for the current BI trends is MDM. Master Data Management (MDM) in conjunction with Data Warehousing for current and historical Transactional Data along with real-time tools like BAM and CEP are the cornerstones for efficient and effective Enterprise Business Intelligence. Without these critical components, EBI will not exist and will not be effective.

All the options we have talked about so far have dealt with BI needs driven by traditional modes of MDM implementations. There is another whole new realm that has to deal with Big Data management.

Before we delve deeper into Big Data, an analogy is on order. What Cryogenics is to Rocket Technology, MapReduce and Hadoop is to Big Data. The technologies used to enable Big Data, driven primarily by algorithms like MapReduce and their implementations like Hadoop, are similar to applying cryogenics to data. It is the same data but being presented in an entirely different and condensed medium.

While the types of BI needs required of Big Data Installations may not conform to the traditional definition of BI, which for the most part has to deal with MIS reports with very precise financial forecasts and related information, this can certainly form a part of or type of need with reference to BDM (Big Data Management).

BDM is absolutely necessary for sites that have reached the petabyte stage or even sites with greater than say 500TB and whose data storage needs are expanding exponentially. The well known ones are of course the usual suspects like Facebook, Yahoo, Google etc. While even very small sites run by under 10 member teams can generate terabytes of information in a very short time. While they do not have the wherewithal to invest oodles of money in building their own Algorithm (like MapReduce) driven storage virtualization platform, one that is hosted in the cloud like Amazon S3 would certainly be of great interest to them.

The future of Business Intelligence will feature these two terms in increasing frequency: “Real Time” and “Big Data”. The trend will be toward gaining real-time Business Intelligence by acting upon data stored in virtualized Big Data environments.

While traditional DW-DH-BI implementations will remain relevant, we will see an increased trend towards the former.

We'll take a closer look at MDM and BDM soon in separate Perspectives.

Real-time Business Intelligence - Use Case

Let us now look at a common use case for real-time business intelligence—the Shopping Cart. Shopping on the web is one of the most common online transactional activities today, and it is increasing year over year. For high traffic sites like eBay and Amazon, business cannot be made truly agile without the use of real-time analytics. Traditional BI just will not suffice.

So let us consider the following example:

Acme Mart is a medium-sized discount store with brick and mortar stores throughout the country.

It also sells merchandise through its high traffic website acmediscountstore.com.

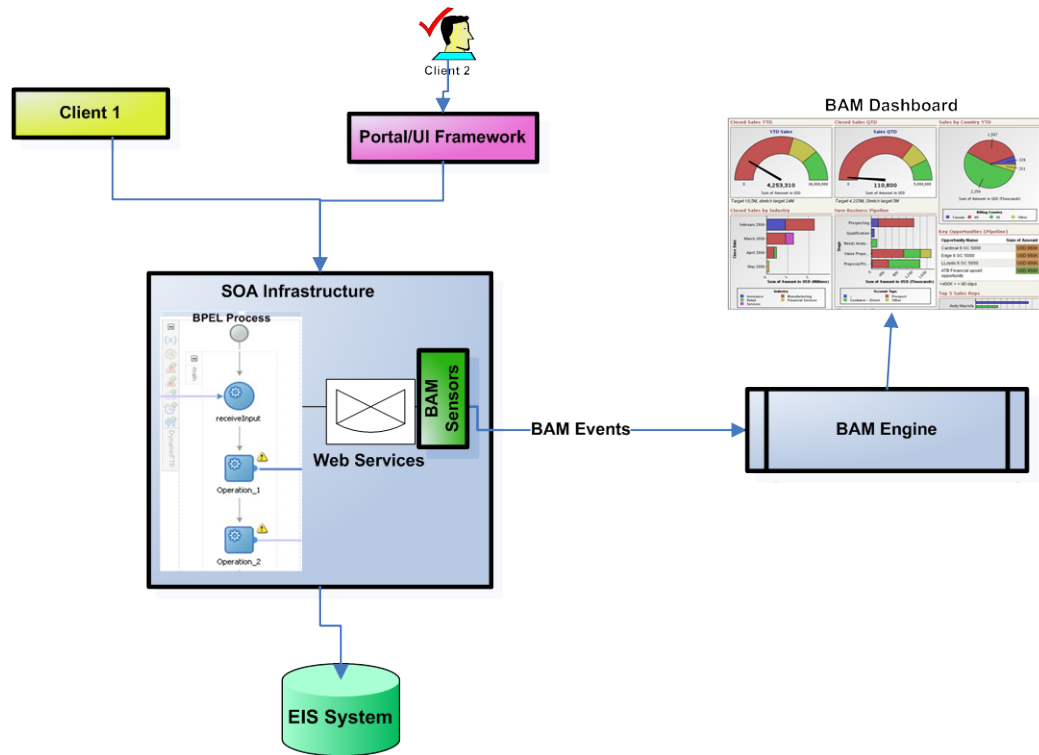
The transaction volumes both at the stores and online are quite large—to the extent of several million per day. It has a mature, established SOA infrastructure and its entire order management process is service-enabled.

Acme Mart also uses BAM (Business Activity Monitoring) to monitor services. Custom dashboards have been built to monitor sales of specific item categories and even items through sophisticated dashboards.

When a sale is completed either at the store POS or online, the information is fed to the dashboard in real-time. When it notices unusual spikes in buying patterns for a particular product type in a given area, it is able to react immediately by diverting all the trucks with that merchandise to that particular area. From a Business Agility standpoint, it scores very high and is on par with the major retailers of its kind. It has instituted these mechanisms for both the Buy and Sell sides of its business, bringing together its suppliers and manufacturers.

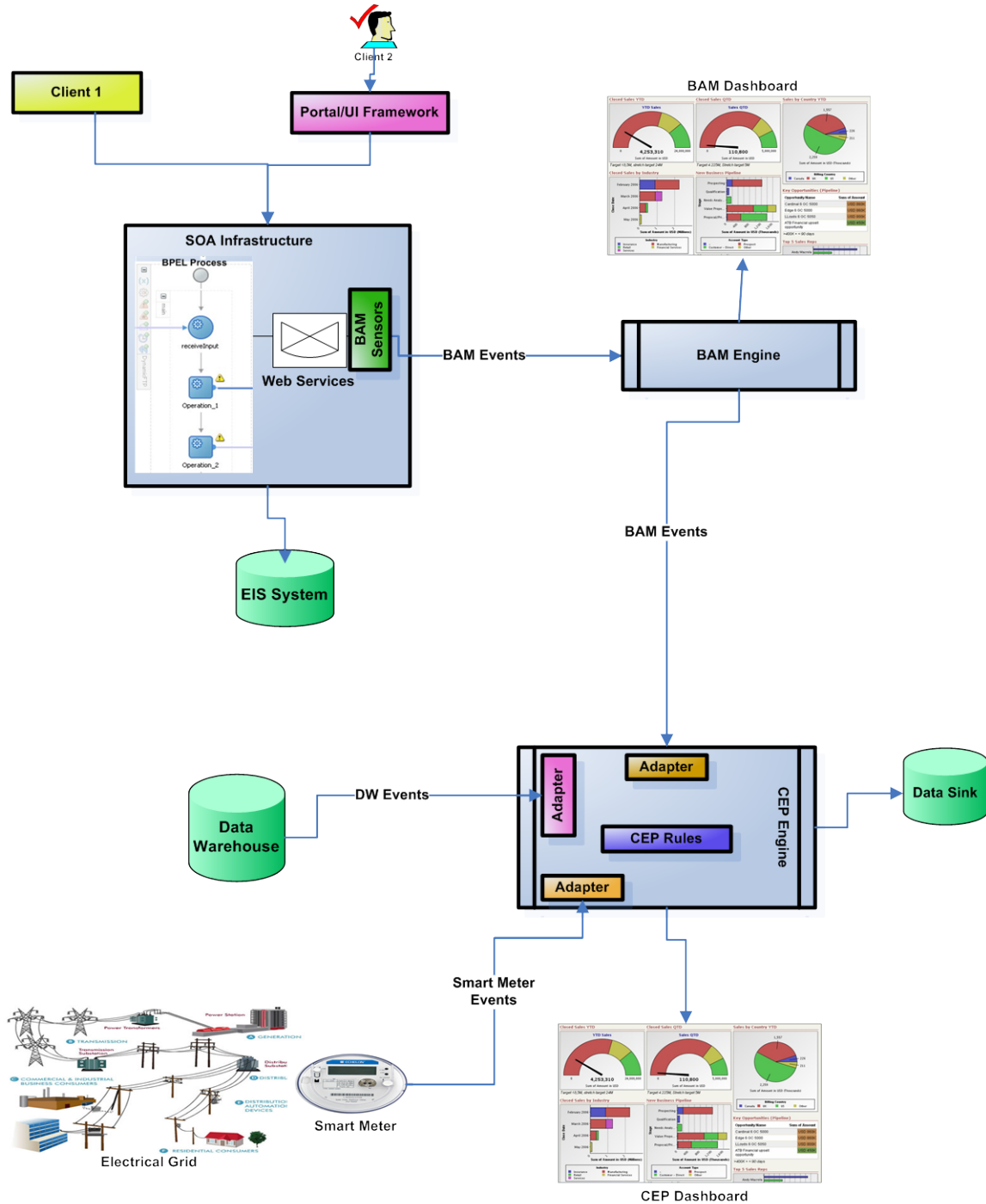
Contrast this with a similar business that uses Traditional BI mechanisms. It is quite obvious that it will not be able to match the level of agility of Acme Mart. By the time the sales data is analyzed, the opportunity to respond proactively has passed. Acme Mart is clearly more capable of reducing costs and increasing revenues via its real-time BAM-enabled BI solutions.

Typical BAM Architecture



BAM is typically tied to real-time monitoring and intelligence gathering with respect to web services. A BAM tool is an important component of a Service Oriented Architecture.

Typical CEP Architecture



While CEP (Complex Event Processing) is similar to BAM, it is more of a services agnostic technology. It is sort of a real-time message aggregator that can process messages from various different sources including web services. One example of this that can be applied to the current use case could be, along with order information, if related stock price information for

the companies that produce those products are also needed in a separate dashboard then this feed that would normally come from the likes of a Bloomberg or Reuters can be combined with this Order data and displayed.

CEP is another emerging trend for real-time analytics. To use an analogy, what mashups does for a Web 2.0 presentation tier, CEP attempts to do at the backend messaging layer by being able to process messages from disparate sources.