Operational Analytics from A to Z

An integrated platform for enterprise analytics of all sizes

May 2013

A White Paper by

Dr. Barry Devlin, 9sight Consulting
barry@9sight.com

Operational analytics has become a mandatory competence for modern business. As big data moves from hype to value, as the market demands ever faster and more accurate responses to opportunities and threats, operational analytics is becoming the foundational technology for near real-time decision making and action taking.

This paper outlines the business value and technological foundations of operational analytics in a manner that is accessible to the business-savvy IT professional and the computer-aware business manager. The evolution of the concept, bridging business needs and technology capabilities, is outlined. The business drivers for immediate, detail-level decisions and actions are provided; and the broad value shown in sample implementations.

We explore the need for a well-integrated, highly cohesive and easily extensible platform to enable the level of performance required. This platform is shown to be an evolution of well-tried and tested data warehouse technologies, combined with more recent developments.

Finally, we discuss the IBM offerings for operational analytics. We focus on the capabilities of IBM InfoSphere Warehouse and associated components as a software base to meet the needs of a full-fledged implementation. And we briefly explore deployment options such as IBM PureData Systems for Operational Analytics that make such an environment available to businesses of every size and across all industries.

In short, operational analytics—from A to Z.

Contents

Sponsored by:
International Business Machines
www.ibm.com

An Asian utility company tracks real-time gas usage of over 2 million users, predicting patterns of demand and rapidly identifying leaks and pilfering. An auto parts retailer analyzes customer website behavior and sales data to closely manage inventory and immediately align marketing efforts to opportunities. A UK mobile telecommunications company scrutinizes dropped call rates and data transfer volumes, segmenting customers on real-time usage patterns to predict churn and detect fraud.

Across industries and business departments, common threads emerge. New data sources, often voluminous, are combined with traditional databases and subjected to in-depth analysis. Results are applied in near real-time in call centers and websites to influence customer behavior, take different actions or, generally, change the parameters and context of operational transactions. Detailed results are captured and fed back into the analysis. The informational-operational loop is being closed… rapidly.

Stephan Haeckel's sense and respond[1] from the 1990s has become listen and anticipate[2]. Businesses that listen to their customers and anticipate their needs consistently show higher revenue growth—as much as 60% higher—than those who don't. In fact, sensing and listening are essentially the same: capturing data as broadly and as quickly as possible and integrating it to a consistent context in a comprehensive, well-managed process. But, to anticipate goes beyond simply responding. Anticipation requires the ability to generate triggers based on potential outcomes, to predict the impact of current actions, and to model best possible outcomes. As we shall see, these are the key to understanding and benefiting from operational analytics.

The business impact is significant and the required change in behavior extreme. As a joint MIT Sloan and IBM Institute for Business Value report[3] points out: "*Taking advantage of new business models and new data, unexpected competitors are emerging and familiar customers are demanding unprecedented attention. Disruptions like these create a widening set of opportunities for players at every level. With a full range of analytics capabilities governed by an integrated analytics strategy, organizations are better positioned to widen, or narrow, the distance between themselves and competitors to their own best advantage.*"

> *Operational analytics has immediate business impact, but demands significant changes in user and IT behavior.*

Operational analytics applies to organizations of all sizes in all industries along the entire length of the supply chain. Understanding customers' behavior and optimizing their experience and profitability is equally important regardless of the size of the customer base. Retailers, telcos and utility companies can all benefit from improving their retention rates. Interactions with suppliers and partners are also subject to the type of improvements that operational analytics can bring. Let's explore how this new world emerges from prior business intelligence undertakings.

From data warehouse to operational analytics

Plus ça change, plus c'est la même chose [4]

The more things change, the more they stay the same. The field formerly known as BI has been assailed in the past twenty years by a wide variety of terminology—much of it trying to suggest novelty where little really existed. *Data warehouse* (DW) of the 1980s became *business intelligence* (BI) in the 1990s and has lately been transformed by some into *business analytics* (BA). The 1990s *operational data store* (ODS) was transfigured to *operational BI* by the 2000s and more recently has evolved into *operational analytics*. It is now difficult to find an adjective that has not been placed in front of "analytics" and declared the latest must-have for decision support. Add the ubiquitous *big data* to the mix and you have a recipe for total confusion. And yet, *plus ça change*…

In truth, technological support for business decision making has evolved over the years. The emphasis has moved from the collection and cleansing of data to the delivery of a wide variety of information to the business. From looking in the rearview mirror of yesterday's performance to tracking business operations in flight. From simple reporting of facts to highly automated decisions based on predicted situations. Figure 1[5] shows this evolution on two fronts: the behaviors driving business users and the responses required of IT providers.

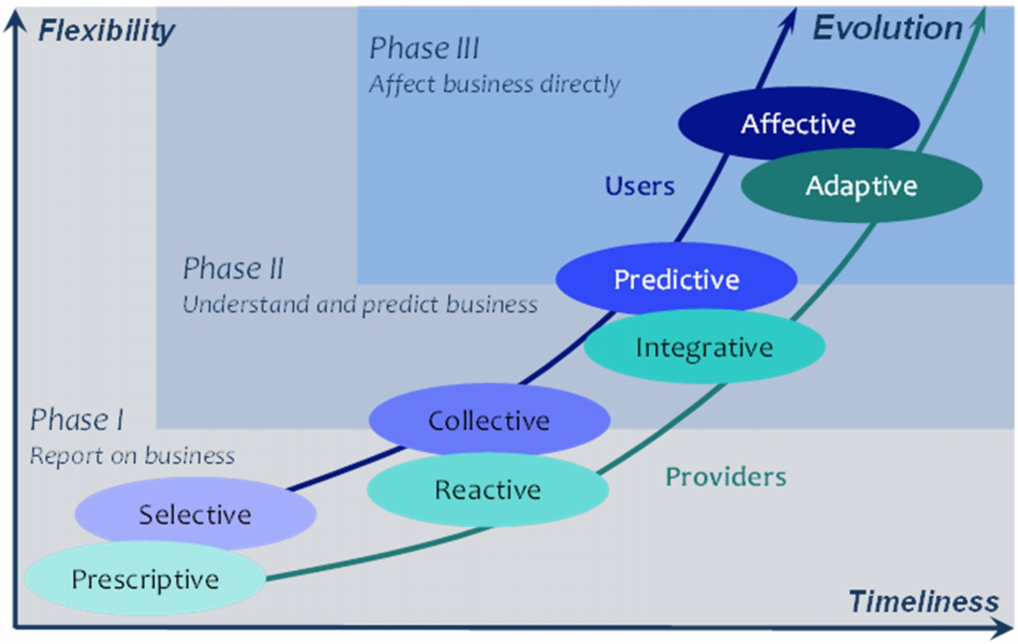

Figure 1: *The evolution of decision making support*

As this evolution proceeds apace, business demands increasing flexibility in what can be done with the data and increasing timeliness in its provision. In Phase I, largely fixed reports are generated perhaps on a weekly schedule from data that IT deem appropriate and furnish in advance. Such reporting is entirely backward looking, describing selected aspects of business performance. Today, few businesses remain in this phase because of its now limited return on investment; most have already moved to Phase II.

This second phase is characterized by an increasing awareness of the breadth of information available collectively across the wider business and an emerging ability to use information to predict future outcomes. In this phase, IT is highly focused on integrating data from the multiple sources of operational data throughout the company. This is the traditional BI environment, supported by a data warehouse infrastructure, described in a companion white paper *"Integrating Analytics into the Operational Fabric of Your Business"*[6], pages 5-6. The majority of businesses today are at Phase II in their journey towards full decision making support and leaders are beginning to make the transition to Phase III.

Phase III marks a major step change in decision making support for most organizations. On the business side, the need moves from largely *ad hoc*, reactive and management driven to a process view, allowing the outcome of predictive analysis to be applied directly, and often in real time, to the business operations. This is the essence of the behavior called operational analytics. In this stage, IT must become highly adaptive in order to anticipate emerging business needs for information. Such a change requires a shift in thinking from separate operational and informational systems to a combined operational-informational environment. This is where the action is today. This is where return on investment for leading businesses is now to be found. The following two sections will first, define operational analytics and second, introduce the architectural platform required to support it.

The move from "simply" predicting customer behavior to truly affecting it requires a highly adaptive, yet process-oriented, IT environment

Operational analytics—detailed, immediate and high value

> *"Three things remain as the basis for competition: efficient execution, smart decision making and the ability to wring ... value from business processes—all of which can be gained through efficient use of analytics."* [7]

The joint MIT Sloan and IBM Institute for Business Value report referenced earlier defines ***analytics*** as *"the use of data and related insights developed through applied analytics disciplines (for example, statistical, contextual, quantitative, predictive, cognitive and other models) to drive fact-based planning, decisions, execution, management, measurement and learning"*. The word *operational* refers to the day-to-day actions required to run the business—the online transaction processing (OLTP) systems that record and manage the detailed, real-time activities between the business, its customers, suppliers, etc. This is in contrast to informational systems where data is analyzed and reported upon.

> *Every operational transaction has an associated decision; every action is preceded by a decision.*

Operational actions demand real-time decisions. Sometimes the answer is so obvious that we don't even see the question. An online retailer receives an order for an in-stock jacket from a signed-in customer; without question, the order is accepted. But the implicit question—*what should we do with this order?*—is much clearer if the item is out of stock, if we have a higher margin jacket available that the customer might like, or a matching pair of trousers when we know the customer recently searched for chinos. There is a subtle progression here from reacting, through assuming to predicting based on the best current information. The decision may be obvious but, sometimes it is worth asking: is a better outcome possible if we explored different decisions and thus took a different action?

Operational analytics is *the process of developing optimal or realistic* ***recommendations for real-time, operational decisions*** *based on insights derived through the application of statistical models and analysis against existing and/or simulated future data,* ***and applying these recommendations in real-time interactions****.*

Why operational analytics is important today

> *"Analytics themselves don't constitute a strategy, but using them to optimize a distinctive business capability certainly constitutes a strategy."* [8]

The two quotes above are from Kaplan and Harris' 2007 book *"Competing on Analytics: The New Science of Winning"*. They were and remain true, but operational analytics is now a hot topic because:

1. Business operations provide ever larger quantities of data for statistical analysis through adaptive websites and increasingly automated and digitized data collection from process sensors
2. Social media is providing growing volumes and varieties of information for analysis
3. As business change continues to accelerate, competition for business increases
4. Data storage and processing continue to increase in power and decrease in cost, making operational analytics a financially viable approach for ever smaller businesses
5. Making many small, low-value decisions better can often make a bigger contribution to the bottom-line than a few, high value ones; and the risk of failure is more widely spread

Enterprise decision management expert, James Taylor, points out[9] that operational data volumes are large enough to provide statistically significant results and the outcomes of decisions taken can be seen and tracked over relatively short timeframes. Operational analytics thus offer a perfect platform to begin to apply the technological advances in predictive analytics and test their validity.

Business use cases—a brief overview

Leading-edge companies are applying operational analytics across the full gamut of business processes along the full length of the value chain. *"Integrating Analytics into the Operational Fabric of Your Business"*[6] provides a functionally-oriented view, emphasizing marketing, customer retention and fraud. Here, we take an industry-oriented view.

Financial services were early adopters of data warehousing solutions to enable a single view of the customer across their many and varied systems. With a move to near real-time availability of transaction and customer interaction data, this basic integration infrastructure has been extended to move fraud detection to new levels. One large financial institution in North America has been using operational analytics to move beyond discovering individual suspicious transactions to identifying patterns of fraudulent behaviors that may affect multiple customers over dispersed locations and acting quickly to prevent loss. Furthermore, such tracking of usage patterns can be further extended to anticipate customers' needs and improve customer service.

Retailers were amongst the first to adopt analytics, back when it was called data mining. Since cash registers became powerful enough to store purchase records, market basket analysis, especially in conjunction with loyalty cards, has been the norm since the early 1990s. But analyzing sales patterns is only part of the story; much real value emerges from integrating this knowledge with the supply chain—from the stores back to the warehouses and often as far as the suppliers—to ensure there are enough of the right products at the optimum price on the shelves exactly when customers want them. The requirement is for a system equally comfortable with operational queries and deep analyses, as seen with GS Retail, a leading Korean retail chain, which has combined operational store management functionality and marketing analysis in a single operational analytics environment.

Cell / mobile phones transformed the telecommunications industry in the mid-1990s as phones became linked to individuals rather than locations, a much larger and more cost conscious customer set and as widespread market liberalization drove fierce competition among providers. Smart phones have continued the trend, with mobile subscriptions exceeding six billion in 2012[10]. In this industry, understanding and anticipating customer attrition and taking action to manage it are survival issues. Operational analytics sits in the middle of marketing, service centers and operations as the conduit of information about customer behaviors to proactively manage customer churn in the call center and run marketing campaigns that attract or retain customers. Hutchison 3G, a leading mobile provider in the UK under the brand "3", use behavioral segmentation—based on analysis of call and mobile broadband usage—rather than traditional demographic segmentation to identify trends and patterns in user behavior before they leave in order to identify likely churners and intervene proactively. The system is also used to detect patterns indicating fraudulent use.

> *Operational analytic systems deployed in one business area can be extended to or reused in others.*

This is but a taster of what companies are doing. The benefits are well-defined and substantial, both in cost savings and revenue generation. And as we've seen, once installed, these systems typically enable other equally significant gains in different business areas than those initially targeted.

An integrated technological platform

"What technology is really about is better ways to evolve." [11]

Traditional BI, described as phases I and II in figure 1, has long been delivered via a strictly layered data architecture centered around an integrated, consistent and reliable repository—an enterprise data warehouse (EDW)—of historical information with specialized data marts for decision support[12]. The emergence of operational analytics and "big data" necessitates a revamp of that old architecture and the introduction of an integrated information platform shown in figure 2 and described in more depth in *"The Big Data Zoo—Taming the Beasts"*[13]. This platform for all information types consists of a number of database and analytic technologies, each optimized for a particular type of processing and access, called *pillars* and named for the business role they support.

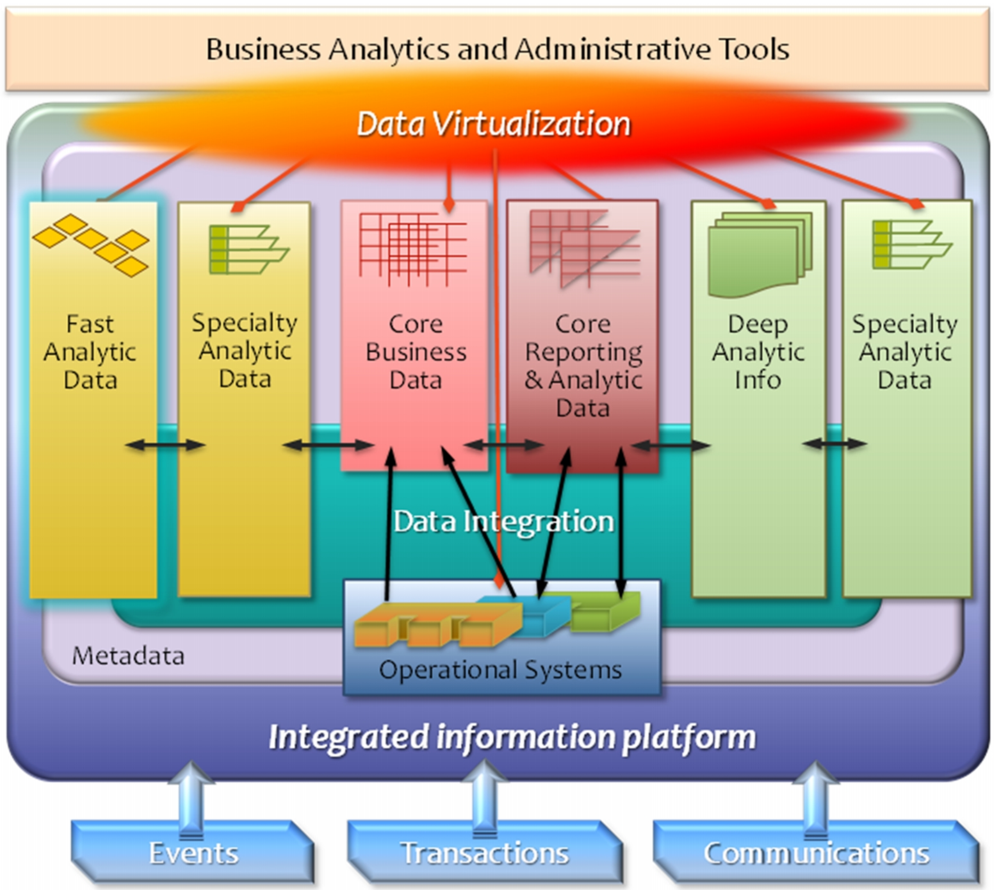

Figure 2: *The integrated information platform*

The first and central ***core business data*** pillar is the consistent, quality-assured data found in EDW and MDM systems. Traditional relational databases, such as IBM DB2, are the base technology. Application-specific reporting and decision support data often stored in EDWs today are explicitly excluded. ***Core reporting and analytic data***, the second pillar, covers these latter data types. In terms of technology, this pillar is also ideally a relational database. Data warehouse platforms such as IBM InfoSphere Warehouse, IBM Smart Analytics System and the new IBM PureData System for Operational Analytics, play strongly here. Business needs requiring higher query performance may demand an analytical database system built on massively parallel processing (MPP), columnar databases or other specialized technologies, such the new IBM PureData System for Analytics (powered by Netezza Technology).

Deep analytic information requires highly flexible, large scale processing such as the statistical analysis and text mining often performed in the Hadoop environment. ***Fast analytic data*** requires such high-speed analytic processing that it must be done on data in-flight, such as with IBM InfoSphere Streams, for example. This data is often generated from multiple sources that need to be continuously analyzed and aggregated with near-zero latency for real-time alerting and decision-making. At the intersection of speed and flexibility, we have ***specialty analytic data***, using specialized processing such as NoSQL, XML, graph and other databases and data stores. It appears twice in the platform because it applies to both the machine-generated data generated by events on the left and human-sourced information from communications on the right.

We thus see that the central pillar of the platform are closely aligned to classic BI and the traditional data warehouse architecture, with the important difference that in data marts used for reporting and analysis, data can—and often should—be fed directly from the operational systems to reduce the latency inherent in the old three-layer approach. Operational analytics is also focused in the same part of the platform, but with a clear understanding that other pillars can be involved as required and that the contextual integration provided across the pillars is a vital aspect to ensure consistency—both semantic and temporal—of the data used and the maintenance of the closed-loop process that operational analytics demands.

*The integrated information platform contains **all** the information generated and used by the enterprise.*

Closed-loop operational analytics

"We have learned that… the past will be a poor guide to the future and that we shall forever be dealing with unanticipated events." [14]

In theory, traditional BI should be a closed-loop process—data is copied from the operational environment into and analyzed in BI; this drives decisions, which influence the operational environment and the outcome can be compared with what was expected. In practice, in strategic and tactical BI, the first three steps proceed very slowly and the final loop-closing step is difficult to perform due to the wider scope of the decisions being taken. Businesses that implemented operational BI in the past have also struggled to close this loop, but for technological reasons. Operational analytics demands that the loop be closed. The actual loop required today is also somewhat more complex. Fortunately, recent technology advances support this demand.

Figure 3 details this closed-loop process for operational analytics against the integrated information platform (simplified for clarity).

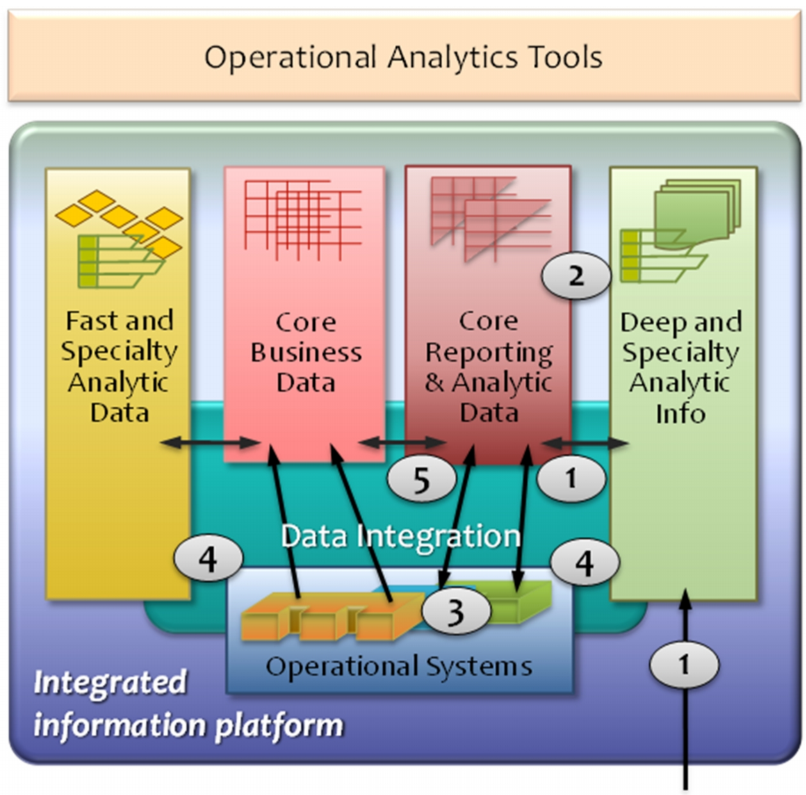

Figure 3: *Closed-loop operational analytics*

1. Gather information for analysis: Information is collected from a variety of sources—existing warehouse data, operational systems, social media, and so on.
2. In-depth analysis: Perform statistical analysis on the gathered data to discover correlations between behaviors and actions, building models. Depending on the types of data and volumes involved this is performed in a combination of relational databases and Hadoop.

3. Model application: One or more models are applied to the operational environment. In some cases, they are applied to all interactions; in others, real-time analysis determines if / when they are activated.

4. Real-time analysis: Analyze behaviors and actions in real-time (for example, web clicks) and apply models as appropriate in real-time as operational activity is occurring.

5. Feedback to in-depth analysis: Note results and feed back into the in-depth analysis stage for further modeling.

From an IT perspective, steps (1), (2) and (5) have very different processing characteristics than (3) and (4), especially in organizations with large data volumes and high velocity decision needs. The former involve reading and number-crunching of potentially large volumes of data with relatively less demanding constraints on the time taken. The latter require the exact opposite—fast response time for reading and writing relatively small data volumes. We now explore these characteristics using IBM InfoSphere Warehouse and related IBM products.

Operational analytics is a process that requires combining operational and informational processing characteristics.

(1) Gather information for analysis

Data integration has long been at the heart of BI. Traditionally, the focus has been on reconciling and consolidating data from disparate operational systems—classic ETL (extract, transform and load). While this aspect remains vital, operational analytics requires that it be supplemented by two further features: (1) speed and timeliness of data acquisition and (2) sources beyond traditional operational systems, including "big data" and sources within the warehouse itself.

IBM InfoSphere Warehouse V10 addresses the first of these points with a new continuous data ingest feature that allows loading of ***DB2 for Linux, UNIX and Windows (LUW)*** tables from pipes or files, avoiding staging tables and allowing other applications to use data while the table loads. By eliminating or reducing the downtime associated with traditional scripted or batch ETL approaches, the required data is made available earlier to users or automated processes, enabling faster analytics and tactical decision making.

Operational analytics demands information from a wider variety of sources and in a broader set of structures than traditional BI. ***InfoSphere Information Server*** and ***DataStage*** provide a comprehensive set of functionality for information gathering and provision from all common sources of relational and non-relational data, both in batch mode and in real-time, in many cases. In addition, the in-built InfoSphere Warehouse Replication features simplify the movement and copying of data within the DB2 environment between Core Business and Reporting/Analytic Data areas.

Altogether, this is a comprehensive set of functionality for all aspects of information gathering.

(2) In-depth analysis

In this phase of operational analytics, the main emphasis is on extensive statistical investigation and more traditional querying of a wide range of relevant data. Finding unexpected correlations across hundreds or even thousands of attributes of behaviors, characteristics and actions is the goal. With so many attributes and often millions of records, this phase should not be expected to be real-time. However, timeliness is still a requirement, especially after the initial cleansing, preparation and exploration phase. Past experiences of overnight (or longer) data mining runs and frequent exports / imports between different platforms are no longer acceptable for ongoing analysis.

Balancing the need for data scale against timeliness is a key design concern in operational analytics.

Scale and timeliness are achieved in a number of different ways, depending on the characteristics of the data used and the urgency of the analytic need. Initial preparation and exploration often benefits from use of a Hadoop-based platform, shown as Deep and Specialty Analysis in figure 3. In this area, ***InfoSphere BigInsights*** provides a platform for managing and using large quantities of multi-structured data. When such data is well-understood, optimized and ready for production use, a portion of it is often moved into the relational environment, where IBM offers three base platforms:

1. ***IBM DB2 Advanced Enterprise Edition***, a Linux, Unix and Windows offering, which is the principal focus of this paper
2. ***IBM DB2 Analytics Accelerator***, a hardware/software offering on System z, using Netezza technology closely integrated with and accessed transparently via DB2 on z/OS, described elsewhere[6]
3. ***IBM PureData System for Analytics***, a system based entirely on Netezza technology, optimized for large-scale, focused strategic and predictive analytics

In the operational analytics sphere, the requirement is for a relational platform capable of (i) the in-depth analytics phase—when these needs are less extreme—and (ii) the later real-time analytics phase, as well as (iii) providing the platform for Core Business Data.

DB2 for LUW supports these three needs through its data partitioning functionality, based on a massively scalable, shared-nothing architecture, and providing high performance query processing against relational as well as native XML data. Version 10 offers adaptive compression, which automatically determines the optimal compression between row- and page-based approaches, potentially reducing storage costs and improving performance, especially for large I/O-bound warehouse applications and query workloads. Storage costs are further reduced by Multi-Temperature Data Management features which enables the DBA to categorize data by temperature. Typical categories are *hot* (frequent access, short response time), *warm* (less frequent access, longer response times), *cold* (infrequent access, never updated) and *dormant* (archived), stored on different speed / cost hardware from solid-state disk through to traditional spinning disks, and even magnetic tape (managed via Optim). This categorization also allows the DB2 Optimizer to prioritize access to hot data.

InfoSphere Warehouse provides embedded data mining, modeling and scoring capabilities, including standard data-mining model algorithms such as clustering, associations, classification and prediction (additional algorithms may be imported in the industry-standard Predictive Model Markup Language format) in support of the type of in-depth, predictive analysis characteristic of this phase of operational analytics. In addition, text analytic functionality allows the direct use of freeform information such as call center notes, customer feedback and so on as part of the overall analysis. The high performance characteristics mean that this in-depth analytics can be performed on the same platform and data as the later real-time analysis phase.

(3) Model application

Model application—the process by which interactions between the business and its customers are changed on the fly to improve business performance—is clearly an operational process. In a traditional BI approach, we would say it occurs outside the BI system and move quickly along. After all, operational systems are outside the scope of business intelligence. However, the term *operational analytics* alone implies that we must deal with both the operational and the analytical. There are four levels of sophistication:

The extent of integration between operational and analytical is driven by the level of automation needed by the business.

1. ***Influencing a manual process:*** in a call center, a representative's interaction with a customer is directed by the model developed in previous analysis of the customer's total lifetime value, simi-

lar metrics and demographically-based predictions of what response or offer should be made. The call center application is loosely connected to the analytics system through pre-loaded model data with no real-time data. The agent retains significant power to direct the interaction.

2. ***Directing a manual process:*** similar to the previous case, but where the connection between the call center and analytic applications approaches real-time. Phase 4 of operational analytics (real-time analysis) must be invoked so that the agent has both real-time data and recommendations based upon it; the agent has more limited freedom in the interaction.
3. ***Automated offer process:*** characteristic of retail websites, offers seen by customers are generated automatically from models based on past interactions as well as the current clickstream. The analytic environment is tightly connected to the operational system and real-time analysis is performed in a closed loop approach.
4. ***Integrated model and offer process:*** real-time analysis and operational activities occur in an integrated system, allowing multiple models to run simultaneously for different customers and self-tune according to results. Real-time analysis and operations are virtually indistinguishable.

As we progress through these levels, moving from separate but connected operational and informational environments to a fully integrated, real-time, automated operational-informational environment, the demands on the database system increase dramatically. Thus, highly efficient query processing and optimization become vital. By utilizing specific optimization techniques within DB2, such as star schema optimizations, the database can deliver high performance for analytical workloads in combination with high performance point query access. The Workload Manager allows fine-grained prioritization of specific types of queries and applications. Improved Index management, including Jump Scan and Pre-fetching, as well as Predicate Evaluation Avoidance also contribute to the performance improvements.

(4) Real-time analysis

As with the previous phase, performance is critical in real-time analysis. We've already discussed this aspect at length for DB2. Another feature that is particularly applicable in real-time analysis, as well as being useful in any BI-related function (and one that has been hand-coded in many EDW developments since the time immemorial!) is Time Travel Queries, also known as Bi-temporal Tables. With this feature, DB2 takes over the management of specified timestamp columns in tables, enabling faster and easier historical and trend analytical queries. With new data being continuously added from real-time activities to the analysis, this feature allows historical tables to be more easily built and maintained, and analysis confined to specific time periods.

Both here, and in the in-depth analysis phase, the included ***IBM Cognos Business Intelligence*** function provides business users with a rich set of BI capabilities to easily access data from the entire operational analytic environment; and reporting and analysis features allow them to deliver relevant information how, when and where it is needed. In addition, customers can take advantage of IBM Cognos Dynamic Cubing functionality to deploy high performance OLAP reporting cubes. ***IBM Mashup Center*** is also included as a foundation for the analytical and visualization needs of small organizations and departmental uses.

(5) Feedback to in-depth analysis

The fifth phase in the operational analytics is relatively straight-forward technically: simply updating the data used in in-depth analysis with additional data and new results discovered in real-time. Its importance should not be underestimated, however; this step finally closes the loop between the operational and informational activities. It has long been a stumbling block in BI systems, because of mismatches in design and ownership between the two environments. The reason why it is easier here is because of the conceptual unification of these two environments in operational analytics. But, at a practical level, the widespread use of a single database environment is the main facilitator. The features and power of InfoSphere Warehouse already discussed, especially Replication, make this possible. The Advanced Security feature, in which DB2 transparently manages access to data on a row and column level within tables, also facilitates the different types of data use in such a mixed operational/informational environment.

An integrated software platform, supporting a wide range of operational and informational needs, is key for operational analytics.

Deployment options

Among the earliest vendors to recognize the importance of decision support and the extent to which it could be supported by relational database technology was IBM. They published the first architectural description[15] of a data warehouse in 1988, with implementations on DB2 and SQL/DS. Since then, IBM has created and/or acquired a broad portfolio of software and solutions that cover the entire breadth of decision support—from strategic BI to business analytics, from predictive to operational analytics, from small to mid-range businesses to the largest multinationals. This section briefly sketches out the main deployment options for the extensive functionality previously described.

As we saw earlier, the heart of any operational analytics deployment is a data warehouse environment, but one that closely integrates with some significant aspects of the operational environment where a large proportion of the core business data originates. Thus, from a software point of view, IBM's operational analytics environment is built around DB2 (on whatever platform) as well as InfoSphere Warehouse and its supporting and complementary tools. Where deeper and/or more specialized analytics is required, Netezza- or Hadoop-based technology may be added. So any company building a bespoke solution on an IBM base will start from these software components.

However, IBM also offers a number of pre-built and integrated solutions.

Announced in October 2012, IBM PureData System for Operational Analytics is the latest member of the IBM PureSystems family of expert integrated systems. It provides a pre-integrated hardware, software and expertise platform optimized for operational analytics and built upon DB2 for LUW and InfoSphere Warehouse on the IBM AIX platform. It is fully scalable to 96 Power7 cores and more than a petabyte of raw, uncompressed data. It is designed to handle over 1,000 concurrent operational queries.

The PureData System for Operational Analytics is complemented by the IBM PureData System for Analytics (based on Netezza technology) and optimized for predictive analytics; and the IBM PureData System for Transactions (based on DB2 for LUW and optimized for purely operational workloads). They all form part of the PureSystems family of expert integrated systems. The PureData System for Operational Analytics has evolved in a technological sense from the ***IBM Smart Analytics System***, which consists for now of offerings on the system x and system z platforms. The former comprises the IBM Smart Analytics System 5600 and 5710, both aimed at the Small and Medium Enterprise sector. The latter consists of the IBM zEnterprise Analytics System 9700 and 9710 and the ***IBM DB2 Analytics Accelerator*** described in *"Integrating Analytics into the Operational Fabric of Your Business"*[6].

Despite the evolving naming, all these solutions spring from a single thought: to provide pre-configured, easily implemented systems that are increasingly optimized for specific types of data usage common in modern business environments.

Conclusions

"It is our future that lays down the laws of our today." [16]

Operational analytics is the foundation of real-time, in-the-moment business that drives business value and customer satisfaction in a fast moving market. It is a harbinger of a brave, new world where the old barriers between operational action taking and informational decision making disappear. Where a real world change—whether opportunity or threat—is investigated and evaluated, a reaction decided and action taken, and the impact assessed in one continuous, closed loop. Where this loop is closed as fast or as often as dictated only by business need.

The choice of an integrated, full function platform for operational analytics is a key step in attaining immediate business value. Towards the new combined operational / informational world. Three factors are fundamental to that choice:

1. A core database environment capable of both deep analytics and rapid response to operational needs, sufficient for the majority of day-to-day work
2. An environment that support and encourages the use of specialized stores and tools for work beyond the daily norm
3. A fully realized data integration and management environment that supports the widest variety of data movement, transformation and access needs

To design, build and deploy such a complex environment with best-of-breed components as a bespoke development is probably beyond the capacity of most IT shops in all but the largest of enterprises. Fortunately, the IBM software base of InfoSphere Warehouse and associated products offers a fully functional and highly cohesive solution. And with the type of expert integrated system exemplified by the IBM PureData System for Operational Analytics, deployment is within reach for all those ready to make the leap.

Dr. Barry Devlin is among the foremost authorities on business insight and one of the founders of data warehousing, having published the first architectural paper on the topic in 1988. With over 30 years of IT experience, including 20 years with IBM as a Distinguished Engineer, he is a widely respected analyst, consultant, lecturer and author of the seminal book, "Data Warehouse—from Architecture to Implementation" and numerous White Papers.

Barry is founder and principal of 9sight Consulting. He specializes in the human, organizational and IT implications of deep business insight solutions that combine operational, informational and collaborative environments. A regular contributor to BeyeNETWORK, Focus, SmartDataCollective and TDWI, Barry is based in Cape Town, South Africa and operates worldwide.

Brand and product names mentioned in this paper are the trademarks or registered trademarks of IBM.

---

[1] Haeckel, S.H., "*The Adaptive Enterprise; Creating and Leading Sense-and-Respond Organizations*", Harvard Business Press, (1999)

[2] "*Outperforming in a data-rich, hyper-connected world*", IBM Center for Applied Insights, YTE03002-USEN-01, (March, 2012), http://ibm.co/QZJ8XL

[3] Kiron, D., Shockley, R., Kruschwitz, N., Finch, G. and Haydock, M., "*Analytics: The Widening Divide*", (2011), http://mitsmr.com/DataReport11

[4] Jean-Baptiste Alphonse Karr in the January 1849 issue of his journal "*Les Guêpes*" ("The Wasps").

[5] Based on personal discussions with Mike Kehoe, IBM in 2007

[6] Devlin, B., "*Integrating Analytics into the Operational Fabric of Your Business*", http://bit.ly/IDAA-WP

[7] Davenport T. H. and Harris J. G., "*Competing on Analytics: The New Science of Winning*", Harvard Business School Press, (2007)

[8] Ibid.

[9] Taylor, J., "*Where to Begin with Predictive Analytics*", http://bit.ly/yr333L , 1 September 2011, accessed 8 February 2012

[10] "*The State of Broadband 2012*", Broadband Commission, International Telecommunication Union, (2012), http://www.broadbandcommission.org/Documents/bb-annualreport2012.pdf

[11] Kelly, K., "*How technology evolves*", TED Talk, (February, 2005)

[12] Devlin, B., "*Data warehouse—From Architecture to Implementation*", Addison-Wesley, (1997)

[13] Devlin, B., "*The Big Data Zoo—Taming the Beasts: The need for an integrated platform for enterprise information*", October 2012, http://bit.ly/Big_Data_Zoo

[14] Handy, C., in "*Learning Organizations*" ed. Chawla, S. and Renesch, J., Productivity Press, (1995)

[15] Devlin, B. A. and Murphy, P. T., "*An architecture for a business and information system*", IBM Systems Journal, Volume 27, Number 1, Page 60 (1988) http://bit.ly/EBIS1988

[16] Nietzsche, F., "*Human, All Too Human*", Cambridge University Press, (1878)