Pneuron: The Pursuit of Painless Agility

by Robin Bloor
Chief Analyst

"Instead of taking the data to one place and piling it up in a heap to be processed, why not move the processing to the various places where the data naturally resides or is gathered?"

What Are We Doing Wrong?

It has long been believed that the best way to harness corporate data is to pour a great quantity of it into a large, powerful and expensive database, and then bombard it with queries to serve a wide variety of BI and analytic applications. If this mighty data hub starts to wilt under the pressure, then you buy more powerful hardware, or build data marts to offload some of the work, or do both. An inevitable side effect to this game plan is the birth of a data integration project to establish and manage the passage of raw data from multiple sources into the burgeoning data mountain – and this data integration project must grow to maturity before the much anticipated business benefits can be mined.

Most organizations have pursued a data strategy like this ever since the idea of a data warehouse took root, and long after the beguiling vision of a single data warehouse was confounded by reality. Indeed the adherence to this approach has shaped a major part of the IT industry in respect of how computer hardware and computer software have evolved. As a consequence, the assumption that data must be centralized before it can be properly exploited has become self-perpetuating, to some degree.

The Cost of Centralizing Data

Data centralization projects are often big-ticket projects. They are expensive in respect of computer hardware, the networking capacity required, the big database software involved, the integration software that's needed, and the team of consultants or developers who see the project through. They are expensive because of the time it takes for the business benefits to arise. They are expensive because maintenance costs are high, and because, very often, such projects beget child projects several years down the line. Whether it's risk management or customer churn, data volumes grow and new sources of data march into the data landscape, and the need arises to pile up yet another data mountain, bigger and slightly different than the last one.

Corporations repeat this expensive and cumbersome operation of centralizing data because there seems to be no alternative, but in fact there is. Most business people who have a little familiarity with computers will realize, if they think about it, that programs are small and data is big. Even very complex programs are usually only megabytes in size, whereas databases are measured in gigabytes, terabytes or even petabytes. And while programs do sometimes change, they do not change very frequently. But data changes all the time and new data is always being added.

So instead of taking the data to one place and piling it up in a heap to be processed, why not move the processing to the various places where the data naturally resides or is gathered? Why not stop moving the data, and instead, move the processing?

The payoff for doing this is not just that you no longer need to spend so much time and effort planning and maintaining data flows. Queries that data analysts and business analysts may wish to ask of the data may actually execute faster. A good example of where this makes a big difference is what is called a "restrictive query." Consider the situation where an analyst would like a list of Asian women between the ages of twenty and thirty with red hair who have bought a particular style of

shoes from one of the company's retail outlets in California. A centralized data warehouse will, most likely, have a vast fact table with maybe billions of rows that has joined together customers and purchases they made. To answer this query the whole fact table will need to be scanned.

But if the data had never been centralized, this workload could be shared across several databases. And if, as might be the case, there are no red haired Asian women in that age range in the customer database, the query will be resolved with very little resource usage. This will also be the case even if there are quite a few such customers.

The Pneuron Alternative

Pneuron is a distributed solutions platform that has been built to minimize the flow of data by moving the analysis, application and processing to the data. The platform has four primary components: a Design Studio for building applications; a UI environment that provides a rich visualization capability for performing interactive "what if" analyses, processing results and integration with BI tools, legacy applications, mobile endpoints, etc.; the Cortex; and the Pneurons. It can be used to build any type of business application, but it is particularly appropriate for big data analytics or where sources of a solution are distributed and diverse, and it can reduce costs dramatically.

"It can be used to build any type of business application, but it is particularly appropriate for big data analytics or where sources of a solution are distributed and diverse, and it can reduce costs dramatically."
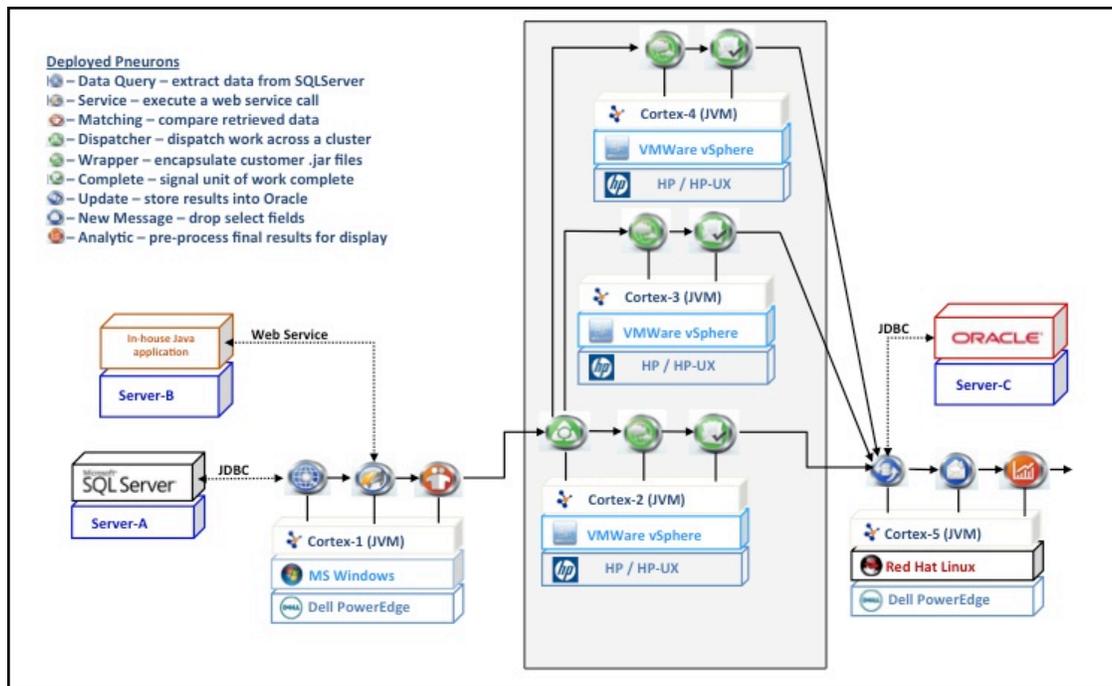

Figure 1. A View of a Sample Pneuron Deployment

Figure 1 provides an example of how Pneuron works when deployed. Applications are built visually – without coding – by configuring Pneurons. These are relatively small executable bundles of functionality, which, when connected together, function like a network and execute the whole application, report or operating model. Some have APIs built in and can query data in files, documents or databases, or they can interact directly with applications. Others can carry out statistical calculations, analysis, predictive modeling, mobile integration, matching and even wrapping of

legacy applications. Currently there are about 50 or so general purpose Pneurons that can be selected to build applications.

The Cortex orchestrates the network of Pneurons by managing the allocation of processing resources and handling all inter-Pneuron communications. As workload increases or decreases, the Cortex dynamically adds or subtracts capacity to sustain network performance and an optimal use of processing infrastructure. As a rule it maintains a very minimal footprint, acting more as a dynamic performance, availability and resiliency facilitator for the Pneurons.

Users also have the option to create a cluster of Cortices to cover the possibility of server failure or volume spikes, so if one node fails, the cluster will dynamically and elastically re-form around still active nodes and ensure no work is lost during the transition. The cluster distributes workload across multiple servers and can be deployed across servers that also run other applications. In addition, it elastically adds and removes Pneurons based on workload, the availability of processing capacity, and multiple other user-defined parameters.

"The Pneuron platform behaves as if it were a single application running in one place, but under the covers it is intelligently distributed."

A convenience of centralization is that it simplifies the implementation of global rules, such as data access rules, encryption rules and so on. The Cortex achieves the same goal by orchestrating and distributing such rules, using the appropriate Pneurons, to the points where they apply. The Pneuron platform behaves as if it were a single application running in one place, but under the covers it is intelligently distributed, end to end. Essentially, physical and logical deployment is presented in the same layer, through the visual capabilities of the Design Studio.

It is also very light in its use of hardware and network resources. It logically locates Pneurons to where computer resources are available and chooses those locations so that network traffic, especially data traffic, will be minimized. Typically new users of the platform discover that they do not need to buy any new hardware to implement it. It is unlikely to be disruptive to the performance of the applications or the data sources it connects to.

The Bottom Line

The Pneuron platform has a unique and agile architecture that is both technology and deployment agnostic, and leverages a non-invasive integration approach to protect existing investments. While it can be used to build any type of business application from business intelligence to accelerating conversions, mergers and acquisitions, it is particularly appropriate for big data analytics applications (risk management, fraud, network security, customer management, etc.) and systems with widely distributed data. It offers a completely new and innovative approach. And it is far more economic, far swifter in development and far easier to maintain. It is truly a game changing capability.

We advise businesses that have always followed the data centralization route to investigate this technology, especially if they are weary of throwing good money after bad.

Company: Pneuron	Products: Pneuron
www.pneuron.com	Versatile development and execution platform