

Big Data: estejamos atentos a essa nova realidade

Por Antonio Pedro Timoszczuk*

É inegável que o avanço tecnológico que viabilizou as redes de comunicação e o aumento da capacidade dos computadores trouxe a possibilidade de interconectar não só as empresas, mas também as pessoas e o exemplo mais conhecido é a internet. No âmbito das empresas, foram abertas possibilidades para aumentar a eficiência dos processos com a implantação de sistemas de controle e supervisão, os quais possibilitam o acompanhamento em tempo real da operação. Na área administrativa a substituição do papel pelo registro eletrônico sem sombra de dúvida permitiu novas abordagens para análise dos negócios e agilizou a tomada de decisão.

Toda essa evolução tem como pano de fundo o processamento de dados e as informações que norteiam a razão de ser de qualquer sistema de computação. Em especial, nos dias atuais, a possibilidade de acesso e a quantidade de informações geradas são enormes. Apenas para dar uma ideia do que isso significa, tomemos como exemplo a atividade na internet e teremos números realmente impressionantes. A cada minuto são enviados mais de 200 milhões de e-mails, o Google recebe mais de 2 milhões de consultas, 100 mil tweets são enviados, 571 web sites são criados e mais de 680 mil conteúdos são compartilhados no Facebook. Tudo isso representa uma quantidade imensa de dados e informações gerada a cada dia,

mês e ano, que são armazenados e podem se tornar um problema ao invés de uma solução.

Considerando um paralelo, o setor elétrico não é diferente, com os sistemas de supervisão e controle que geram “toneladas” de dados e o processamento das informações dos consumidores pelas empresas de distribuição. Recentemente, com a expectativa da implantação efetiva do conceito de smart grid e da medição inteligente, devem ser gerados dados em uma quantidade que excede tudo o que foi feito até hoje, desencadeando um verdadeiro tsunami de dados. O que fazer então para estar preparado para lidar com tal volume de dados?

A capacidade de processamento de dados e informações tem evoluído ao longo do tempo. Inicialmente, com o conceito de bancos de dados locais até os modernos repositórios de dados e informações ou data warehouses, armazenando e permitindo a recuperação de informações estruturadas. Posteriormente, a denominada fusão de dados introduziu o conceito de cruzamento de dados e com isso agregou maior valor às informações produzidas. Passou-se então para a era da inteligência em negócios ou business intelligence, conhecido pela sigla BI, onde as informações são tratadas e geram indicadores para os negócios. Recentemente, uma expressão ganhou destaque no setor

de processamento de dados: o Big Data, representando um verdadeiro hype.

Mas o que significa o ‘Big Data’ afinal? Todos os comentários a respeito destacam que o Big Data tem um potencial imenso e que revolucionará os negócios na era digital, proporcionando novas visões de cenários, melhorando a tomada de decisões, gerando lucro em áreas antes não imaginadas. Entretanto, se formos buscar uma definição formal para o seu significado, provavelmente iremos encontrar tantas definições quanto o número de pessoas para as quais façamos essa pergunta. A primeira ideia liga a expressão big data à grande quantidade de dados e informações. Mas o que representa uma grande quantidade hoje pode não sê-lo amanhã e, para isso, basta se lembrar dos números na internet que citei como exemplos anteriormente. Muitas organizações importantes, dentre elas fornecedoras de software e equipamentos, além de consultorias que atuam no tema, apresentam diversas definições que vão desde a quantidade de dados processada, sua diversidade e complexidade, até a definição de ambientes integrados de conhecimento. Entretanto, acredito que os 3 vês, muito citados na literatura, captam a maioria das definições existentes para o big data e são eles: o Volume de dados, a sua Variedade, que representa a complexidade dos dados e informações, e por último, a Velocidade

com que são tratados. Essas características foram viabilizadas especialmente por meio de avanços tecnológicos na capacidade de processamento e de armazenamento, o que permite o acesso grandes volumes dados a uma velocidade nunca antes pensada. Sob essa ótica, surgem dúvidas sobre qual a diferença entre o BI tradicional e o Big Data, além da velocidade.

Em adição às características que citei, pessoalmente prefiro pensar no big data sob um outro ponto de vista, que é a forma como os dados ou informações são tratados. No BI tradicional, geralmente, os dados e informações estão muito bem estruturados e regras de relacionamento e filtros são aplicados de forma a produzir resultados, na forma de indicadores, que sejam significativos para a análise desejada. Busca-se a causalidade de algum fato ou tendência. No mundo do Big Data, por outro lado, busca-se a correlação dos dados e informações e deixa-se com que os dados “falem” por si só, ou seja, busca-se o que e não o porquê. Pode parecer estranho, mas ao buscar o relacionamento entre os dados e informações, geralmente surgem conclusões inesperadas e visões para um problema nunca antes imaginadas podem ser vislumbradas. E quando buscam-se correlações, a validade e o valor dessas correlações devem contar com o elemento humano, que é quem vai dizer se o que a máquina encontrou tem utilidade ou não.

Essa abordagem já resultou em modelos sofisticados de predição de ocorrência de surtos de gripe, até empresas especializadas em indicar o melhor momento de adquirir uma passagem aérea. Peço ao leitor que note que estamos falando de validade e valor da informação e introduzimos um termo importante que é a predição ou o olhar futuro. Em suma, o Big Data visa tratar volumes consideráveis de informação, sem a preocupação com a sua estruturação e a uma velocidade maior do que os sistemas tradicionais de BI.

E o que isso tudo tem a ver com o

sistema elétrico? Tudo, diria eu. O setor elétrico vem sendo pressionado por uma redução nos seus custos por um lado e, pelo outro, por resultados que a empresa deve proporcionar para garantir o retorno do investimento e proporcionar a sua expansão. Este é um desafio considerável onde as análises tradicionais encontram-se no seu limite, especialmente pelo considerável aumento do volume de informações produzidas pelos sistemas de controle e supervisão, sem contar as informações administrativas e financeiras, tudo isso com pressão por respostas em tempo recorde. Diversas empresas já têm os olhares voltados para esta nova estratégia que é o Big Data. O Operador Nacional do Sistema Elétrico (ONS) anunciou recentemente que está executando um teste piloto aplicando o Big Data para avaliar as condições do sistema de energia brasileiro.

Portanto, considerar a estratégia de Big Data pode ser a saída para encontrar caminhos que levem as empresas a soluções que produzam os tão almejados resultados de eficiência e economia. Lidar com o cenário complexo das tecnologias do smart grid e tirar o máximo proveito comercial, ou, da mesma forma, estar preparado para a realidade do gerenciamento de novas fontes de energia que vêm sendo integradas ao sistema e acomodar os prosumers – que são os consumidores e produtores de energia ao mesmo tempo – são desafios que podem encontrar apoio na estratégia do Big Data. Estejamos, então, atentos e atualizados quanto às possibilidades do Big Data.

**Antonio Pedro Timoszczyk é engenheiro eletricista, com ênfase em eletrônica, mestre e doutor em Engenharia de Sistemas Eletrônicos. Atualmente, é gerente de P&D e Inovação na Fundação Ezute (nova denominação da Fundação Atech), pesquisador colaborador da Universidade de São Paulo (USP) e diretor do instituto Efort.*