
GERÊNCIA DE INFORMAÇÃO DE PROCESSOS INDUSTRIAIS: UM ESTUDO DE CASO NA PRODUÇÃO DE PETRÓLEO E GÁS

Alessandro J. de Souza*
ajdsouza@dca.ufrn.br

Wany Leydiane S. de Andrade*
wany@dca.ufrn.br

Gustavo Bezerra P. Leitão*
gugarn@hotmail.com

Clauber G. Bezerra*
clauber@dca.ufrn.br

Rafael H. Feijó*
rafaelheider@yahoo.com.br

Luiz Affonso Guedes*
affonso@dca.ufrn.br

*DCA-UFRN
Caixa Postal 6101
CEP 13081-970 - Natal-RN

ABSTRACT

With the coming of intelligent field equipments, a great variety of data coming of those equipments can become available for other applications. However, now the great challenge in industrial automation is to transform that great volume of data in useful information. The production information management system (on the Web), presented in this work, it allows data visualization through Internet/Intranet, aiming at to facilitate the complex analysis of stored data in supervisory systems. The use of Web technologies, guarantees to system the advantage of the portability, economy with software licenses and maintenance, since in the days today most of the computers possesses a Browser. In that way, the information management system is constituted of an environment capable to answer to engineers' of processes requests for analyses, through the integration and unification of processes data.

KEYWORDS: PIMS, MES, J2EE, SCADA, OPC, DDE, Database.

RESUMO

Com o advento dos equipamentos de campo inteligentes, uma grande variedade de dados proveniente desses equipamentos pode ser disponibilizada para outras aplicações. Porém, o grande desafio atual na automação in-

dustrial é transformar esse grande volume de dados em informação útil. O sistema de gerência de informação da produção (via Web) apresentado neste trabalho permitindo a visualização de dados via Internet/Intranet, objetivando facilitar a complexa análise dos dados armazenados nos sistemas supervisórios. O uso de tecnologias Web, garante ao sistema a vantagem da portabilidade, economia com licenças de software e manutenção, já que nos dias de hoje quase todos os computadores possuem navegadores (Browser). Dessa forma, o sistema de gerência da informação se constitui de um ambiente capaz de responder aos pedidos de análise dos engenheiros de processos, através da integração e unificação dos dados dos processos.

KEYWORDS: PIMS, MES, J2EE, SCADA, OPC, DDE, Banco de Dados.

1 INTRODUÇÃO

Para melhor representar uma arquitetura de um sistema de automação, pode-se dividi-lo em níveis, conforme mostra a figura 1. No primeiro nível, encontram-se os sensores e atuadores, como sensores de nível, pressão, temperatura, de fins de curso, válvulas, inversores de frequência, etc. No nível seguinte, encontram-se os controladores lógicos programáveis (CLP) e os Supervisory Control and Data Acquisition (SCADA) (Daneels and Salter, 1999), que se comunicam com os CLPs. Até o iní-

cia da década de 1990, os sistemas de controle formavam ilhas de automação, onde cada sistema controlava o seu parque sem possibilitar a integração das informações. Com a disponibilização dos dados da produção, desde o chão-de-fábrica até o produto final, pode-se subir mais um nível na pirâmide da automação e, neste caso, referir-se aos sistemas de gerência de informação dos processos que são englobados com o termo geral de Enterprise Production Systems (EPS), onde estão incluídos os Plant Information Management System (PIMS) e os Manufacturing Execution Systems (MES). No nível mais alto da pirâmide, figura 1 encontram-se os sistemas corporativos de gestão da planta, Enterprise Resource Planning (ERP), responsáveis pela transformação desses dados em informações de negócio. A integração negócio-manufatura é um processo chave para as indústrias e requer troca de informações entre os processos de negócio e os sistemas de manufatura de modo a criar e manter vantagens competitivas no mercado.



Figura 1: Pirâmide de Automação.

O nível mais alto da pirâmide, a integração das informações de chão-de-fábrica aos dados corporativos possibilita a otimização da gerência e a integração de todo processo de manufatura refletindo, positivamente, nos aspectos financeiros. Dentre os vários benefícios existentes na interação negócio-manufatura, podemos destacar: a disponibilidade para comprometimento; redução do tempo de produção e a implantação e otimização da cadeia de suprimento. É no nível de gerência de negócio que não podemos deixar de ressaltar a importância da Internet, possibilitando, por exemplo, aos clientes de uma determinada indústria consultar o status de seu pedido de fornecimento a qualquer instante, obtendo esta informação em tempo de execução.

1.1 Objetivos do Trabalho

Baseado nas necessidades de gerenciar e disponibilizar os dados gerados no chão-de-fábrica é proposto através deste trabalho uma estratégia para gerência de dados históricos e on-line de plantas industriais através da Internet, sendo esta de caráter genérico, ou seja, reúne os módulos considerados básicos (aquisição de dados,

compactação e armazenamento e visualização) para a disponibilização de dados de um sistema de automação industrial através da Internet e sua posterior validação.

A estrutura proposta permite ao usuário remoto o acesso aos dados do sistema de automação industrial. Nesta estratégia são abordados três contextos distintos, o sistema SCADA, o servidor de aplicações e o cliente; seus módulos e como se dá à intercomunicação entre eles. Como características principais à estratégia propõe a independência de plataforma tanto do lado cliente quanto do lado servidor; independência dos dispositivos e protocolos de comunicação que estão operando no sistema de automação industrial; sistema de controle de sessão de usuários remotos; configurabilidade do sistema pelo cliente remoto, além de basear-se na utilização de software livre.

Para validar a estratégia proposta o sistema de gerência esta sendo implantado como estudo de caso na produção de petróleo e gás em uma das unidades da PETROBRAS UN-RNCE, Brasil.

2 GERÊNCIA DE INFORMAÇÃO DE PROCESSOS

Basicamente, a gerência da manufatura é feita através de um dos sistemas EPS que possuem um repositório, onde são concentradas todas as informações relevantes das células de produção, diretamente ligadas aos sistemas de supervisão e controle. Esses sistemas coletam as informações dos sistemas de supervisão, CLPs, Sistema Digital de Controle Distribuído (SDCDs) e sistemas legados e os armazenam em uma base de dados em tempo real. Tal base tem a característica de possuir grande capacidade de compactação e alta velocidade de resposta às consultas em sua base histórica. Devido a este fato, torna-se capaz de armazenar um grande volume de dados com recursos mínimos, quando comparada às soluções convencionais. Basicamente, a informação contida na base de dados consiste de uma lista de registros temporais formada por uma ou mais tabelas, as quais possuem tipicamente os seguintes campos: Time stamp; Identificador do tag; Valor; e Qualidade do dado;

A aquisição de dados consiste em uma das tarefas mais difíceis na implantação de tecnologias de conexão com os sistemas que compõem as células de produção, pois por mais modernos e organizados que sejam esses sistemas, sempre apresentam uma grande heterogeneidade. Os sistemas de gerência da manufatura já dispõem de uma grande variedade de drivers de comunicação cobrindo a maioria dos sistemas existentes e englobando as mais novas tecnologias de troca de informação, tais como o OLE for Process Control - OPC (Iwanitz and Lange, 2001).

Outra tecnologia de comunicação ainda bastante utilizada é o Dynamic Data Exchange - DDE (Petzold, 1998). O mecanismo DDE é um protocolo baseado no conceito de mensagem e que permite a troca de informações entre aplicações Windows.

3 SISTEMA DE INFORMAÇÃO PROPOSTO

O sistema de gerência de informação via Web, proposto neste trabalho, tem como objetivo possibilitar que usuários obtenham os dados da célula de produção que estão nos sistemas SCADA. Após a captura desses o sistema faz um tratamento de compactação e os armazena em uma base de dados possibilitando a consulta a dados históricos. Além da consulta a dados históricos, o sistema possibilita a visualização de dados online do processo. Todo o sistema pode ser configurado e parametrizado remotamente, através de um simples browser via protocolo HyperText Transfer Protocol - HTTP. A figura 2 ilustra a arquitetura do sistema.

O sistema é composto por três módulos independentes, conforme mostra figura 3: módulo de comunicação, módulo de compactação e módulo de visualização. O módulo de comunicação opera como interface de comunicação entre o sistema e o supervisor, através do qual é feita a seleção dos dados que serão requisitados e seus parâmetros. Em seguida, o módulo de compactação utiliza um algoritmo para compactar os dados, os quais são então enviados via rede e armazenados no banco de dados relacional para serem manipulados pelo módulo de visualização.

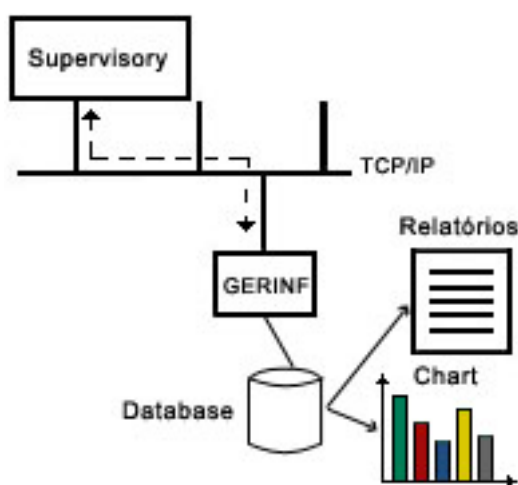


Figura 2: Arquitetura do Sistema

3.1 Módulo de Comunicação

Através do sistema SCADA é possível ligar ou desligar bombas, abrir ou fechar válvulas. Sistema SCADA é um software destinado a promover a in-

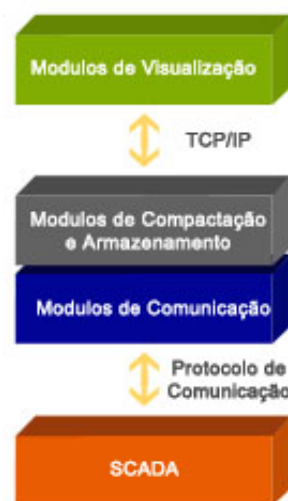


Figura 3: Módulos do Sistema

terface homem/máquina, que proporciona uma supervisão plena de um processo através de telas devidamente configuradas, representando o processo, onde estas podem ser animadas em função das informações recebidas ou enviadas pelo CLP, controlador, etc. Atualmente existem vários programas supervisórios, tais como INTOUCH, ELIPSE, AIMEX, FIX-32, VIEW, CIMPLIST, entre outros.

O módulo de comunicação é responsável pela coleta dos dados de produção para o seu posterior armazenamento em uma base de dados relacional, que servirá de consulta para o módulo de visualização. Os dados podem ser lidos ciclicamente pelo módulo de comunicação ou enviados por iniciativa do servidor de dados quando ocorre algum evento, como a alteração do valor de uma variável. Essa coleta de dados pode ser realizada a partir dos sistemas de supervisão ou diretamente a partir dos CLPs, utilizando alguma tecnologia de comunicação, como o OPC, DDE ou Drivers nativos dos equipamentos. Espera-se que um software de gerência possua interfaces para a maioria dos sistemas de importância comercial.

3.2 Módulo de Compactação e Armazenamento

A compressão de dados tem como principal objetivo eliminar informações redundantes, para tornar menor o espaço de armazenamento das informações e diminuir custos com armazenamento e transmissão de dados.

Foram desenvolvidos diversos algoritmos para realizar a compressão de dados. Deseja-se que o algoritmo de compressão de dados tenha uma taxa de compressão alta, não elimine informações relevantes, e permita que os dados reconstituídos sejam o mais próximo possível

dos dados originais.

Existem vários tipos de algoritmos de compressão de dados, por isso deve-se escolher o que melhor satisfaz a cada tipo de aplicação.

Os algoritmos se classificam em algoritmos com e sem perda de informação. Os algoritmos sem perda de informação permitem que os dados reconstituídos sejam idênticos aos dados originais, mas proporcionam uma baixa taxa de compressão. Os algoritmos com perda de informação permitem apenas uma estimativa dos dados originais, mas tem uma alta taxa de compressão.

Entre os principais algoritmos de compressão estão:

Os algoritmos de codificação de repetição (sem perda de informação): basicamente uma palavra muito freqüente, ou uma repetição contínua de um determinado caracter, é substituído por uma codificação que ocupa menos espaço. É mais utilizado em compressões de documentos de texto. Como exemplo temos a codificação run-length;

Algoritmos de amostragem a intervalos fixos (com perda de informação): são salvas amostras periódicas dos dados, porém para aumentar a taxa de compressão o período de amostragem deve ser aumentado, causando assim perda de informações intermediárias relevantes. Esta abordagem é bastante utilizada em PIMS (Plant Information Management System);

Algoritmos de amostragem variável (com perda de informação): são mais eficientes; é um aprimoramento do algoritmo anterior porque salva amostras apenas quando há mudanças significativas dos dados, mantém a qualidade dos dados, além de permitir uma alta taxa de compressão.

A seguir será detalhado o funcionamento de dois algoritmos de amostragem variável: o Swinging Doors e Boxcar/Backslope.

3.2.1 Algoritmo Boxcar/Backslope

O algoritmo Boxcar/Backslope (Misra et al., 1998) contém dois critérios que são consultados para decidir se um dado será armazenado ou não. Os critérios são o Boxcar e o Backslope. Esse algoritmo contém ainda três parâmetros que são definidos de acordo com a variável cujas informações serão comprimidas: o desvio de compressão, o tempo máximo de compressão e a qualidade do dado.

Sempre que o tempo transcorrido entre o último valor armazenado e o último valor recebido for maior que o tempo máximo de compressão, o último valor recebido será armazenado.

Se a qualidade do último dado recebido mudar, o valor

anterior ao último recebido será armazenado.

A cada novo dado que é recebido, são verificamos os dois critérios. Se ambos os critérios forem violados, então o dado será armazenado no Banco de Dados. Violar o critério Boxcar significa que o último valor recebido apresentou um desvio maior que o limite do desvio de compressão, em relação a reta paralela ao eixo do tempo, e que passa pelo último valor armazenado. Violar o critério Backslope significa o último valor recebido apresentou um desvio maior que o limite do desvio de compressão, em relação a reta definida com os dois últimos valores armazenados.

Se apenas um dos critérios for violado, nenhum dado será armazenado.

A figura 4 ilustra o funcionamento deste algoritmo, onde, a bola preta representa um dado que foi armazenado, e a bola branca representa um dado que não foi armazenado. A bola amarela representa um dado que violou os dois critérios do algoritmo, e portando será armazenado no Banco de Dados.

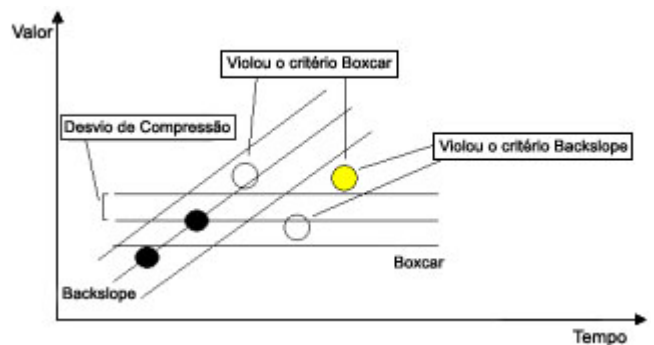


Figura 4: Algoritmo Boxcar/Backslope

3.2.2 Algoritmo Swinging Doors

O algoritmo Swinging Doors contém três parâmetros que são definidos de acordo com a variável cujas informações serão comprimidas. São eles: tempo mínimo de compressão, tempo máximo de compressão e desvio de compressão. Quanto maior for o tempo máximo de compressão e o desvio de compressão, maior será a taxa de compressão.

O algoritmo Swinging Doors cria uma área de cobertura no formato de um paralelogramo e com largura igual ao dobro do desvio de compressão, este paralelogramo se estende desde o último valor armazenado até o último valor recebido. Se qualquer um dos valores recebidos entre o último armazenado e o último recebido ficar fora da área de cobertura, então o valor anterior ao último recebido será armazenado no Banco de Dados. O último valor recebido é sempre armazenado se o tempo transcorrido desde o último valor armazenado for maior

que o tempo máximo de compressão. O funcionamento deste algoritmo é explicado nas figuras 5a e 5b.

A bola preta representa um dado que foi armazenado; A bola branca representa um dado que não foi armazenado; e a bola amarela representa o último dado recebido. A figura 5a apresenta um paralelogramo formado pelo último valor armazenado e o último valor recebido. Nenhum dos valores intermediários ficaram fora do paralelogramo, portanto nada será armazenado.

A figura 5b mostra outra situação, onde um dos valores intermediários ficou fora da área do paralelogramo. Logo, o valor anterior ao último valor recebido (em amarelo), será armazenado no Banco de Dados.

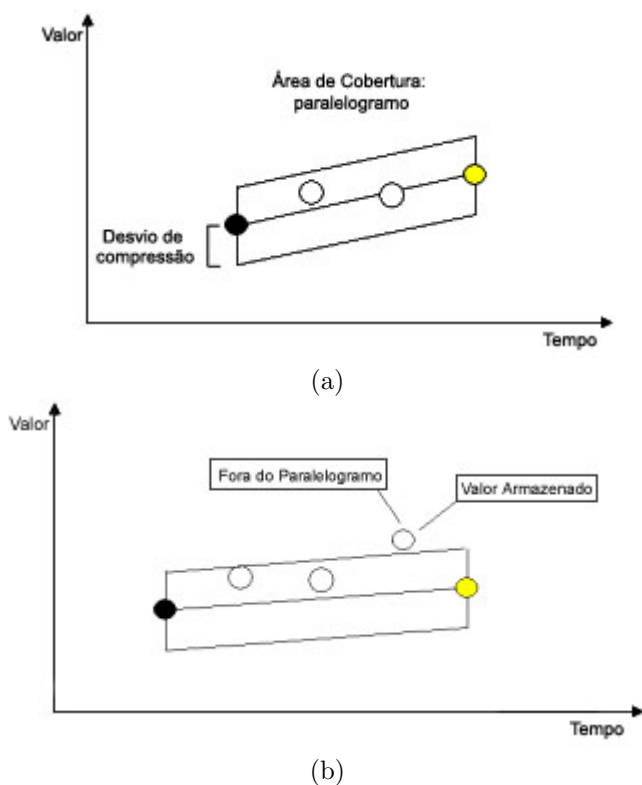


Figura 5: Algoritmo de compactação de dados Swinging Doors.

3.3 Módulo de Visualização

Sistemas tradicionais de gerência de processo têm sido instalados como sistemas desktop com limitadas interações com os ambientes de escritório da corporação, o que leva à necessidade dos usuários destes ambientes fazerem uso de outros aplicativos para obter dados do processo de produção. Outro fator negativo é que os atuais sistemas voltados para o pessoal de processo, dificultando o entendimento e operação por parte do pessoal dos setores administrativos.

Após analisar as inúmeras aplicações para resolver

esse problema de conectividade entre o chão-de-fábrica, gerência de processos e administração, tornou-se claro que a solução estaria no desenvolvimento de um sistema em ambiente Web. O navegador Web (browser) se tornou para o usuário final, o ambiente onde ele faz tudo o que precisa na empresa, desde a troca de e-mails até a visualização de relatórios, gráficos e imagens. Neste contexto, o sistema de gerência via Web possibilitará que cada setor da corporação tenha sua visão segundo as necessidades de informações desejadas, a partir do login do usuário no sistema.

A utilização de um sistema Web trará as seguintes vantagens: facilidade e baixo custo de manutenção e atualização dos sistemas, pois para efetuar a manutenção do sistema será necessário alterar apenas arquivos HTML, Applets Java e Scripts que se encontram no servidor; possibilidade de execução de aplicações em máquinas sem grande poder de processamento e sem a necessidade de constantes atualizações de hardware e software; o sistema não fica condicionado à instalação em nenhuma máquina cliente, pois todos os computadores têm browser previamente instalados.

O sistema foi desenvolvido baseado na tecnologia Java, mais especificamente a plataforma Java 2 Enterprise Edition (J2EE) (Singh et al., 2002). Esta tecnologia consiste em um conjunto de especificações coordenadas e um guia de práticas que permitem o desenvolvimento, instalação, execução e gerenciamento de aplicações n-camadas no servidor.

O principal objetivo da utilização da Plataforma J2EE foi explorar suas capacidades para suporte a aplicações corporativas tais como: robustez, estabilidade, segurança e performance. Tais capacidades foram de extrema importância para nosso projeto, haja vista, a excessiva quantidade de requisições recebidas e volume de informações tratadas em cada requisição.

O desenvolvimento do sistema seguiu um padrão de projeto, Design Patterns, bastante difundido no desenvolvimento de aplicações Web, chamado Model-View-Controller - MVC (Tichy, 1997), possibilitando um maior controle sobre as requisições vindas das máquinas clientes para a aplicação. Esse padrão foi implementado através do Framework de aplicação Struts da Apache Software Foundation (Husted et al., 2004).

O módulo de Visualização consiste em um ambiente genérico tanto para o suporte aos demais módulos (comunicação e compactação) do sistema, quanto possibilitar que usuários possam construir dinamicamente suas consultas.

O suporte dado pelo módulo de visualização ao módulo de comunicação corresponde à parametrização necessária para que este possa buscar as informações

nos sistemas SCADA. Dentre os parâmetros necessários é possível escolher o tipo de protocolo de comunicação a ser utilizado, tipo de supervisor e variáveis (tags) que devem ser monitoradas em cada sistema SCADA. Outros parâmetros como unidade de valor, qualidade do dado, também, podem ser especificados no módulo de visualização. Todos os parâmetros configurados são enviados ao módulo de comunicação para que esse possa buscar os dados e ter a melhor desempenho possível.

O suporte fornecido ao módulo de compactação e armazenamento consiste, também, em fornecer parâmetros necessários ao funcionamento deste, de acordo com o algoritmo de compactação escolhido.

A geração de consultas dinâmicas, ponto principal do módulo de visualização, consiste na busca de informações em um banco de dados relacional gerado pelo módulo de compactação e armazenamento. As consultas dinâmicas são geradas a partir do acesso aos metadados do banco, que é proporcionado através da interface DatabaseMetaData do Java 2. Esta interface possibilita a montagem de consultas baseada nos metadados do banco de dados. Desta forma, o sistema não se refere apenas a consultas pré-existentes (criadas por desenvolvedores), mas sim na construção de consultas pelo próprio usuário.

Além de visualizar consultas históricas o módulo de visualização, também, possibilita a visão on-line dos dados do processo.

Desta forma, através do módulo de visualização é possível chegarmos como as informações do processo a todos os setores da corporação de forma rápida, independente de plataforma e com custo de software e manutenção bastante baixos.

4 ESTUDO DE CASO NA INDÚSTRIA DE PETRÓLEO E GÁS

Como foi citado na introdução desse artigo, o objetivo do sistema de gerência de informação é proporcionar aos usuários dos diversos setores da corporação uma visualização dos dados históricos e on-line da planta possibilitando análises e auditorias para tomada de decisões. Nesse contexto, o sistema de gerência está sendo avaliado como estudo de casos em uma das plantas de produção de Petróleo e Gás da PETROBRAS UN-RNCE.

Inicialmente podemos constatar que a integração entre os módulos do sistema (módulo de visualização Web, comunicação, compactação e armazenamento) foi bastante satisfatória.

Durante o experimento foi possível observar a operação do módulo de comunicação capturando dados de um sistema supervisor do tipo Intouch. A partir do momento

que os dados estão sendo capturados pelo módulo de comunicação, o mesmo repassava ao módulo de visualização para acesso on-line (figura 7) e ao módulo de compactação e armazenamento, que é responsável por armazenar os dados em uma base de dados PostgreSQL para geração e avaliação de dados históricos da produção (figura 6).

Para o módulo de compactação e armazenamento foi escolhido o algoritmo de amostragem variável Swinging Doors, devido à necessidade de se armazenar um volume de informações muito grande, que são coletados durante anos, este algoritmo nos garante uma alta taxa de compressão e mantém a qualidade dos dados armazenados. Para o nosso sistema, o parâmetro tempo mínimo de compressão não foi utilizado, essa adaptação foi feita para que informações relevantes não fossem perdidas.

Através do módulo de visualização via Web foi possível observar o ganho de agilidade, pois além de disponibilizar as informações da planta via Web, o módulo tem a capacidade de configurar os demais módulos do sistema, como por exemplo: cadastro de novos supervisor e suas variáveis (tags), tempo de amostragem para captura de dados, entre outros.

Atualmente, o sistema de gerência está em fase experimental. Depois de implantado a versão final, toda a produção poderá ser avaliada através dos históricos formados, sendo possível visualizar a produção de anos anteriores.

Para adquirir as informações provenientes do campo de produção, utilizamos uma interface de comunicação entre o sistema de gerência proposto e o supervisor utilizado, neste caso específico o Intouch. Isso ocorre, devido ao fato de que os supervisórios já existem e é inviável manuseá-los para gerar um banco de dados. Dessa forma, é necessário acessar no supervisor para adquirir tais informações.

Na implementação da interface de comunicação, utilizamos o protocolo DDE. O DDE (Dynamic Data Exchange) é um protocolo para troca dinâmica de informações entre aplicativos Windows. Através dele é possível a duas aplicações (um servidor e um cliente) conversarem através de mensagens padronizadas, que permitem ao cliente obter dados a partir do servidor. Cada cliente também pode requisitar dados de diversos servidores, onde para cada um deles um canal de comunicação é estabelecido. No contexto deste projeto, os servidores DDE serão definidos como sendo os softwares supervisórios, enquanto que o cliente será o software de gerência.

De acordo com esta abordagem, o software de gerência inicia uma conversação com o software supervisor, solicitando-lhe os dados relevantes ao usuário para o

gerenciamento do processo em questão. Feito isto, o supervisor começa a transmitir os dados do processo que está sendo monitorado para o software de gerência a cada intervalo de amostragem do sistema e as envia para o banco de dados temporal do sistema, para a sua recuperação quando solicitado pelo usuário.

Levando em consideração a linguagem utilizada (JAVA), é importante ressaltar que este sistema não depende de sistema operacional ser executado, ou seja, qualquer sistema operacional como Linux, Windows pode ser servidor do sistema. Se a empresa julgar necessário é possível disponibilizar estas informações na internet, isto dependerá somente da infra-estrutura da empresa, pois o sistema já possui suporte para este fim.

As telas apresentadas encontram-se com valores fictícios, tendo em vista a necessidade de manter o sigilo das informações da empresa.

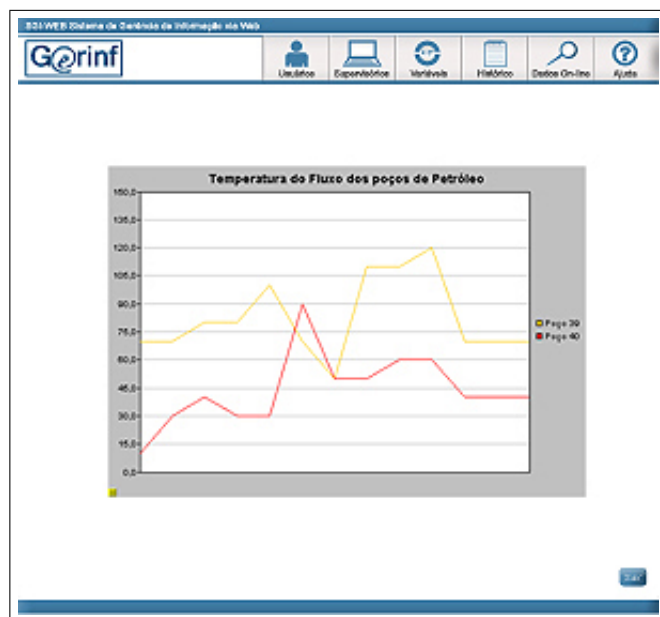


Figura 6: Dados Históricos da Planta

5 CONCLUSÕES

As tecnologias Web conquistaram grande importância dentro da área dos sistemas de automação industrial, possibilitando a expansão dos mesmos para uma rede mundial. A disponibilização de um sistema de gerência de informação dos processos industriais através da Internet vem dinamizar o fluxo da informação dentro da corporação.

Sendo assim, o presente trabalho em seu protótipo inicial mostrou possuir as características necessárias para auxiliar tomadas de decisão referentes à produção de petróleo e gás, propiciando aos usuários da empresa uma ferramenta capaz de informar quanto está sendo pro-

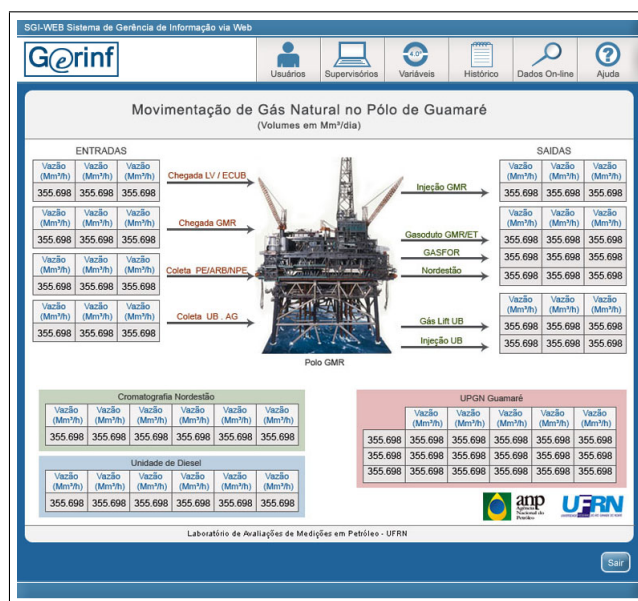


Figura 7: Dados On-line da Planta

duzido em tempo real ou fazer avaliações através de um banco de dados histórico do processo.

Atualmente se está trabalhando no aperfeiçoamento e ampliação das funcionalidades do sistema de gerência de informação, de modo que se possa obter, além de informação, o conhecimento de forma inteligente sobre os dados de operação.

REFERÊNCIAS

- Daneels, A. and Salter, W. (1999). What is scada?, *International Conference on Accelerator and Large Experimental Physics Control Systems*.
- Husted, T., Dumoulin, C., Franciscus, G. and Winterfeldt, D. (2004). *Struts in Action*, second edn, Manning Publications.
- Iwanitz, F. and Lange, J. (2001). *OLE for Process Control*, first edn, Huthing.
- Misra, M., Kumar, S., Qin, S. J. and Seemann, D. (1998). On-line data compression and errors analysis using wavelets, *AIChE Journal*.
- Petzold, C. (1998). *Programming Windows*, 5th edn, Microsoft Press.
- Singh, I., Sterans, B. and Johnson, M. (2002). *Designing Enterprise Applications with the J2EE Platform*, second edn, SUN.
- Tichy, W. F. (1997). A catalogue of general-purpose software design patterns. IEEE Computer Society.