



## IX - GRUPO DE ESTUDO DE OPERAÇÃO DE SISTEMAS ELÉTRICOS - GOP

### AQUISIÇÃO DE DADOS DE USINA HIDROELÉTRICA E INTEGRAÇÃO COM SISTEMAS DE BUSINESS INTELLIGENCE

**LÊNIO OLIVEIRA PRADO JÚNIOR (\*) (1); RICARDO EMERSON JULIO (1); EDSON BORTONI (1); GUILHERME SOUSA BASTOS (1); MARCELO OLÍMPIO DA FONSECA (2); PEDRO PAULO DE CASTRO VIANA (2) UNIFEI (1); JIRAU ENERGIA (2)**

#### RESUMO

A proposta deste trabalho é aplicar IT (*Information Technology*) aos processos de OT (*Operational Technology*) na usina hidrelétrica Jirau, implementando uma arquitetura de recuperação de dados brutos armazenados no PI System e sistemas auxiliares, processamento através de técnicas de *Big Data* e ETL (*Extract, Load and Transform*), e armazenamento para uso em análises e relatórios. O conhecimento obtido da análise dos dados auxilia no planejamento estratégico, visando atender às restrições do sistema usando o melhor número possível de unidades geradoras, reduzindo ao máximo o consumo energético dos serviços auxiliares, resolvendo o problema do comprometimento da unidade (*Unit Commitment - UC*).

#### PALAVRAS-CHAVE

*Business Intelligence*, Usina Hidrelétrica de Jirau, PI System, *Big Data*, ETL.

#### 1.0 - INTRODUÇÃO

A Usina Hidrelétrica (UHE) Jirau está situada no estado de Rondônia, no local denominado Ilha do Padre, a 120 quilômetros ao longo do Rio Madeira, na cidade de Porto Velho. Com a capacidade instalada de 3.750 MW, possui 50 turbinas do tipo Bulbo. A UHE Jirau é uma usina à fio d'água e a operação de seu reservatório é realizada conforme curva guia estabelecida na Resolução ANA nº 269/2009. Nos períodos de cheia, a vazão afluyente não utilizada para a geração de energia elétrica deve ser liberada através das comportas do vertedouro, ou seja, o recurso hídrico deve ser aproveitado da melhor maneira possível para geração energética, pois a água não pode ser armazenada além do nível máximo permitido.

Uma característica relevante do rio Madeira refere-se à elevada quantidade de sedimentos e troncos transportados em seu curso. Apesar da existência de infraestrutura para o descarregamento de troncos para a jusante do barramento, constituída por linhas de *log booms* que direcionam os troncos para um vertedouro específico, um volume significativo de troncos acumula na tomada d'água das casas de força da usina. Este acúmulo de troncos e sedimentos resulta em restrições de potência decorrentes da obstrução das grades de proteção das tomadas d'água das unidades geradoras. Quando o nível de sujeira alcança o limite máximo, são necessárias atuações para solucionar o problema. O procedimento de limpeza da sujeira é um processo recorrente, porém é necessário o desligamento da unidade geradora para sua realização, o que acarreta perdas financeiras enquanto o equipamento se encontra desligado.

O elevado número de unidades geradoras representa um grande desafio operacional, uma vez que é desejável manter o máximo rendimento no processo de geração de energia com o menor custo possível. Adicionalmente, é necessário manter as unidades geradoras e os equipamentos auxiliares funcionais, para que seja possível aproveitar ao máximo o recurso hídrico, enquanto este se encontra disponível.

A tomada de decisões operacionais em um cenário com as características mencionadas se torna complexo, e muitas vezes depende da experiência dos operadores para determinar o modo de operação das unidades geradoras e identificar possíveis pontos de atenção, o que aumenta o grau de complexidade operacional da usina.

Dado o panorama exposto, é imprescindível viabilizar métodos que auxiliem operadores e gerentes na tomada de decisões, além de disponibilizar informações para que se tenha uma visão sobre as decisões tomadas e qual o impacto de tais decisões no que se refere à produção de energia, estado operacional das unidades geradoras e também pontos para possíveis melhorias (3).

Visando atender estas demandas, este trabalho apresenta as técnicas, ferramentas e procedimentos desenvolvidos a fim de prover um ambiente que forneça aos operadores da usina uma melhor visualização de informações para auxiliar na tomada de decisões e verificar o status operacional da usina.

## 2.0 MATERIAIS E MÉTODOS

Existem na Usina Hidrelétrica Jirau inúmeros sistemas de monitoramento, gestão e acompanhamento das atividades, sendo que cada um destes sistemas auxilia os operadores em determinada etapa do processo da operação. O volume de dados gerados por sensores, atuadores, alarmes e inúmeras outras fontes para obter informações de nível de reservatório, estados dos equipamentos, temperatura, nível de óleo, status das unidades geradoras, sistemas de resfriamento, etc., é muito grande. Procedimentos existentes na usina dependem diretamente de informações obtidas de várias fontes para a tomada de decisões, sejam elas técnicas, gerenciais ou comerciais.

Todo aparato de hardware comumente utilizado em usinas e indústrias é conhecido como OT (*Operational Technology*) e é utilizado para controlar, monitorar e alterar os dispositivos físicos. Sistemas SCADA são exemplos de OT, e são utilizados para obter dados em tempo real, disponibilizando interfaces gráficas, gerando alarmes e alertas, com a finalidade de auxiliar o operador na tomada de decisões. Os PLCs (*Programmable Logic Controllers*) são dispositivos que realizam a interface entre os sensores, coletando e processando dados, e tomando decisões baseadas em parâmetros definidos, disponibilizando informações de saída para outros sistemas.

Usualmente, OT é utilizada por funcionários específicos com alto grau de conhecimento e treinamento. É comum que OT não recebam atualizações ou modificações por longos períodos, e que possuam sistemas operacionais específicos e especializados no processo em que são inseridos. Por outro lado, IT (*Information Technology*) são ferramentas computacionais, incluindo hardware e software, que são utilizadas pelos funcionários como ferramenta de suporte indispensável às operações de negócio. Estas ferramentas comumente recebem atualizações e melhorias com maior frequência.

Neste trabalho, técnicas de IT são aplicadas aos processos de OT na usina hidrelétrica Jirau, com o intuito de efetuar *Big Data Analytics* através de um ambiente tecnológico de ponta com as mais recentes técnicas de comunicação e processamento de dados (1, 2). Dentre as ferramentas e tecnologias utilizadas para alcançar o objetivo estão:

- **PI System:** é um portfólio integrado de software para coletar, armazenar, visualizar, analisar e compartilhar dados operacionais com usuários dentro e fora da empresa. Compreende, entre outras ferramentas complementares, um banco de dados temporal, onde as informações são armazenadas utilizando-se *Tags* que possibilitam acrescentar carimbos de data e hora. Trata-se de uma poderosa ferramenta que possibilita organizar de maneira simples, embora extremamente eficaz, os ativos e produtos existentes na organização, exibindo-os em formato de árvore, em uma estrutura organizacional condizente com a realidade da empresa.
- **Python:** é uma linguagem de programação de propósito geral, robusta e eficiente, muito utilizada no desenvolvimento de aplicativos, serviços e aplicações web, scripts e principalmente para análise de dados e *machine learning*. Para possibilitar a execução eficiente das transformações necessárias nos dados, e também integrar todas as ferramentas utilizadas neste trabalho, é imprescindível utilizar uma linguagem aprimorada, de fácil aprendizado e que possui amplo suporte e desenvolvimento de novas funcionalidades, de modo a permitir que novas técnicas e novos aplicativos sejam integrados com as atualizações que forem sendo desenvolvidas.
- **Apache Cassandra:** um banco de dados *NoSQL*, onde as informações não precisam ser armazenadas em forma de tabelas, contendo sempre a mesma estrutura, como obrigatoriamente acontece em um banco de dados relacional. As principais características deste banco de dados são o fato dele ser distribuído, gratuito, de código aberto, escalável e de alta disponibilidade, sem comprometimento do desempenho. As informações a serem armazenadas podem ter formatos aleatórios, de modo que diversos tipos de estruturas de dados podem ser armazenados, dependendo da necessidade da aplicação. Tais características foram o motivo da escolha desse banco de dados para uso no projeto de Jirau, uma vez que a sua tolerância a falhas e o armazenamento e processamento eficaz de grandes volumes de dados é propícia para análise de dados de missão crítica.
- **Apache Spark:** é uma ferramenta para análise unificada de dados, processamento de *Big Data* e *machine learning*, capaz de realizar o processamento de forma paralela e distribuída. O Apache Spark, assim como o Apache Hadoop estão entre as ferramentas mais utilizadas para processamento de *Big Data* e *Data Science*, sendo que o Spark efetua o processamento de informações em memória, sendo assim, cerca de 100 vezes mais rápido que o Hadoop. O Hadoop oferece um sistema de arquivos distribuído, eficaz quando é necessário uso intenso de leitura e gravação em disco. Existem diversos módulos integrados para *streaming* de dados, uso de SQL e processamento de gráficos existentes para o Spark, possibilitando que novas fontes de dados sejam acessadas e manipuladas de maneira simples.

- Power BI: é um conjunto de ferramentas de visualização da Microsoft, que se conecta a uma grande variedade de fonte de dados com a finalidade de criar indicadores, *dashboards* e disponibilizar informações de forma gráfica e intuitiva, auxiliando na exibição de uma maneira visualmente envolvente e interativa. As ferramentas oferecem elementos gráficos que permitem apresentar os dados de maneira atrativa, com customizações, filtros e seleção de parâmetros que podem ser utilizados por quem irá visualizar as informações.
- Aplicação Web: foi desenvolvida uma ferramenta web que a usina utilizará para definir quais informações serão utilizadas no processamento de BI. Trata-se de um conjunto de telas, funcionalidades e serviços desenvolvidos para permitir a configuração e execução de todo processo descrito neste documento.

A Figura 1 apresenta o diagrama do *framework* idealizado e implementado.

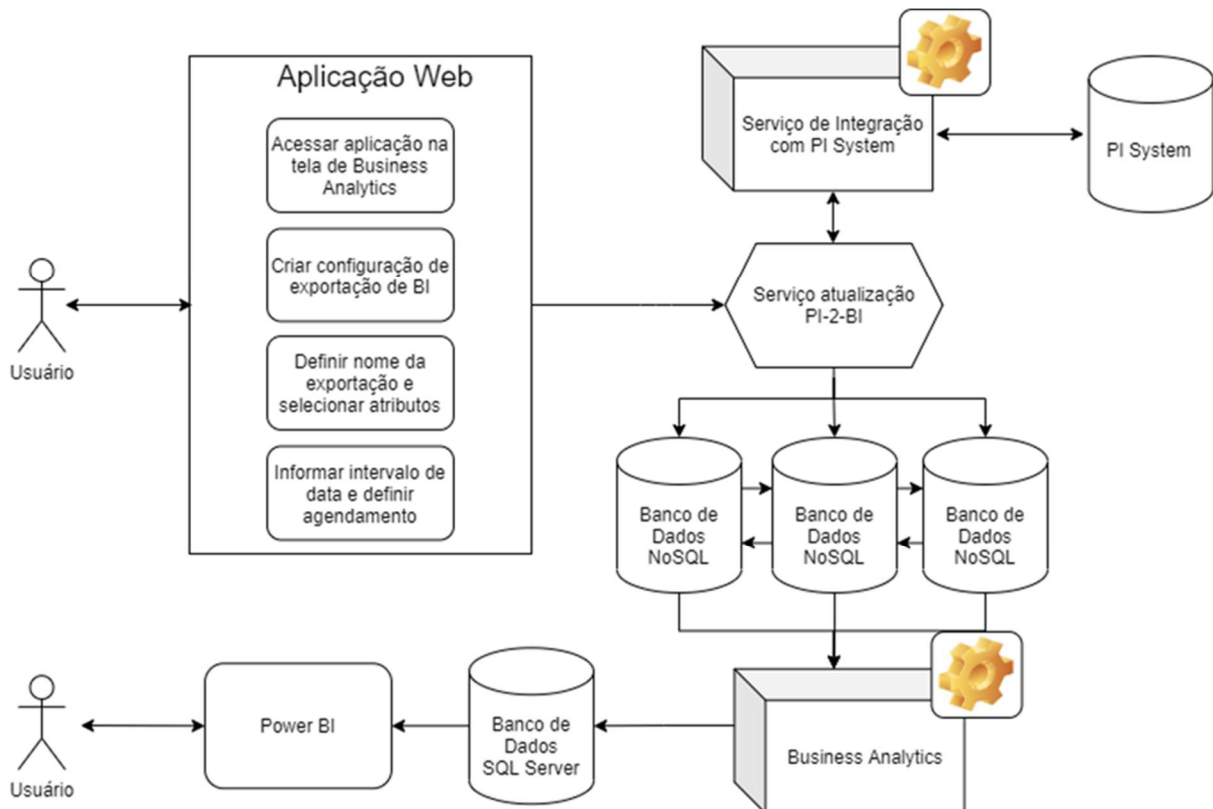


FIGURA 1: Diagrama do *framework* desenvolvido.

O PI System é utilizado para armazenar de forma temporal os dados de OT que são coletados pelos PLCs. Esta é a principal fonte de dados utilizada neste projeto. Adicionalmente, são efetuadas integrações com outros sistemas utilizados pela usina, como o sistema SAU (Sistema de Acompanhamento de Usina), para coletar dados específicos de manutenção e informações que apenas são registradas neste sistema.

Para viabilizar a coleta no PI System de informações a serem processadas pela ferramenta de BI proposta, foram desenvolvidos módulos no sistema que possibilitam selecionar os atributos desejados para análise e que exportam essas informações para o banco NoSQL.

O integrador de dados desenvolvido permite, através de uma interface web, selecionar a base de dados a ser pesquisada, os atributos a que se deseja ter acesso, o intervalo de datas para obtenção dos dados e o intervalo de execução da tarefa de coleta. Este integrador fornece meios de acessar, de forma rápida, os atributos relacionados aos equipamentos cujos dados estão armazenados no PI, possibilitando assim ter acesso a dados históricos necessários para as análises. Ao salvar a configuração de exportação, é criado um agendamento para que os dados sejam exportados.

Um módulo desenvolvido no projeto monitora os agendamentos de exportação de informações do PI para execução de BI, e então efetua a coleta dos dados previamente definidos pelo utilizador, armazenando essas informações para execução do processamento no banco de dados NoSQL Cassandra.

Após efetuado o armazenamento, o processamento pelo Apache Spark é acionado para recuperar as informações do banco Cassandra e executar os métodos de Extração, Transformação e Carregamento (ETL) dos dados. Este é o ponto principal do *framework* proposto, pois é nesta etapa que os dados brutos são tratados, transformados e agrupados, de modo a se tornarem informações utilizáveis de maneira mais simples pelas ferramentas de visualização, no caso o Power BI. Também são executadas nesta etapa, cálculos e tratamentos nos dados, substituindo valores incorretos e executando validações.

A implementação do processamento dos dados utilizando o Apache Spark foi criado de modo que seja simples adicionar novos tratamentos de ETL, ou seja, caso seja necessário um método diferente, é necessário apenas criar um novo modelo de manipulação dos dados utilizando a linguagem Python, cadastrar na aplicação através da interface web e utilizar nas configurações de exportação de BI. Essa arquitetura desenvolvida torna o sistema mais robusto e adaptável às necessidades da usina, possibilitando melhorias e correções sem a necessidade de modificações no restante da aplicação.

Atualmente na aplicação foram desenvolvidos dois módulos; o primeiro transforma os atributos que são exportados em formato de linha para colunas, de acordo com a data e a unidade geradora relacionada, resultando em uma elevada redução na quantidade de registros a serem manipulados, facilitando a visualização de dados relacionados a um determinado ativo e consequentemente o tempo de processamento. O segundo módulo realiza a agregação dos dados de hora em hora, utilizando a média como função, a fim de prover uma visualização sumarizada das informações. Este módulo é útil quando a visualização das informações sem agregação se torne inviável ou pouco significativa.

Por fim, as informações resultantes, já processadas, são salvas em um banco de dados relacional. As informações deste banco são consumidas diretamente pelo Power BI, que tem acesso nativo a este tipo de ambiente. A escolha desse esquema utilizando um banco relacional deve-se ao fato de que, na usina, foi escolhido este banco para permitir o acesso dos usuários às informações de BI. Também por opção da usina, foi utilizado o Power BI como ferramenta de análise e criação de dashboards, porém outras ferramentas para essa finalidade podem ser usadas utilizando a plataforma desenvolvida.

Os usuários do sistema, interagem com o *framework* através da aplicação web, onde são configurados os parâmetros de exportação com os atributos existentes no PI System. Outra forma de interação dos usuários é através do Power BI. No Power BI, as informações armazenadas são utilizadas para a criação de indicadores de desempenho KPIs (*Key Performance Indicator*) e relatórios que resumizam e apresentam os dados de modo interativo, com capacidades de filtragem e seleção de parâmetros de customização.

### 3.0 APLICAÇÃO WEB

Nesta seção são apresentadas algumas telas da aplicação web desenvolvida no projeto, demonstrando como são criadas as configurações de exportação de informações a serem utilizadas no processamento de BI.

Na Figura 2 é apresentada a tela de cadastro de configuração de exportação para BI. As informações dos elementos e atributos apresentados na tela são obtidas do PI System, através do integrador desenvolvido, e o usuário pode selecionar quais são os atributos necessários para realização de sua análise.

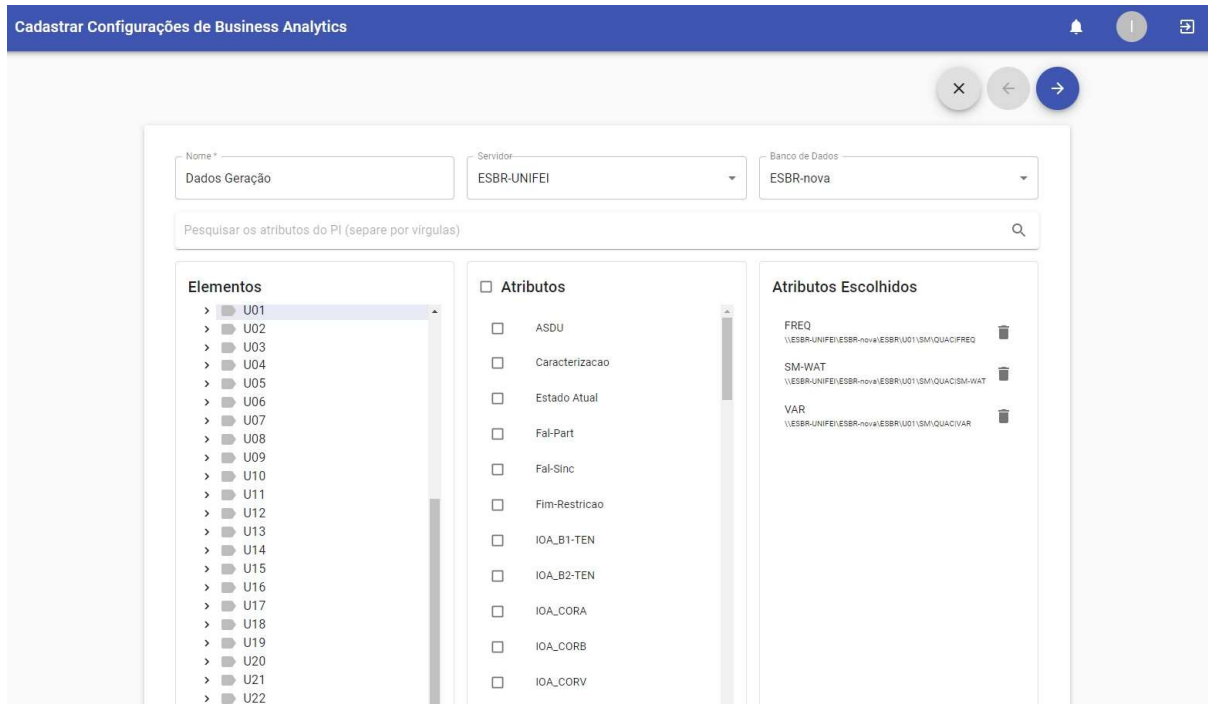


FIGURA 2: Cadastro de configuração de *Business Analytics*.

Na Figura 3 são listados os módulos ETL desenvolvidos no projeto até o momento. Para cada módulo apresentado na imagem, há um programa desenvolvido na linguagem Python que efetua, utilizando o Apache Spark, o ETL dos dados. Adicionar novos módulos é bem simples, de modo que à medida que novas demandas de tratamento de informações forem efetuadas, um novo programa em Python será desenvolvido e então cadastrado na tela mostrada na Figura 3, sem a necessidade de modificações em outras partes da plataforma.

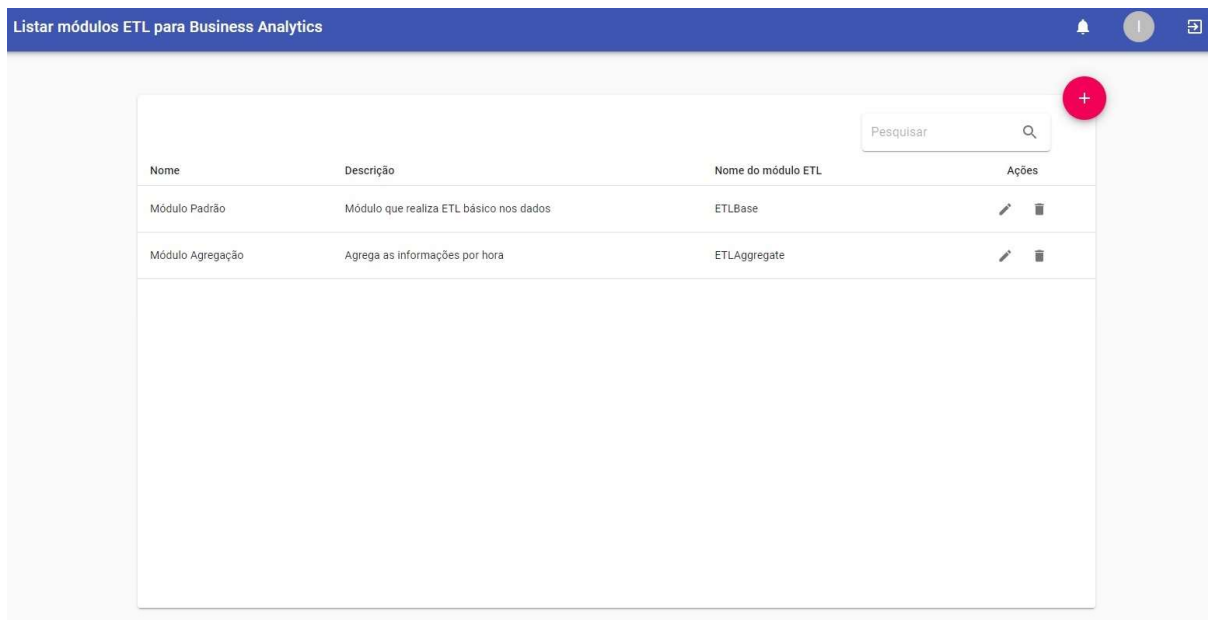


FIGURA 3: Listagem dos módulos ETL para *Business Analytics* desenvolvidos no projeto.

#### 4.0 DASHBOARDS

Nesta seção são apresentados os *dashboards* criados utilizando o Power BI, utilizando as informações após o processamento de ETL.

A Figura 4 exibe um *dashboard* que apresenta o valor de geração máximo possível detalhado por unidade geradora, considerando a queda bruta disponível.

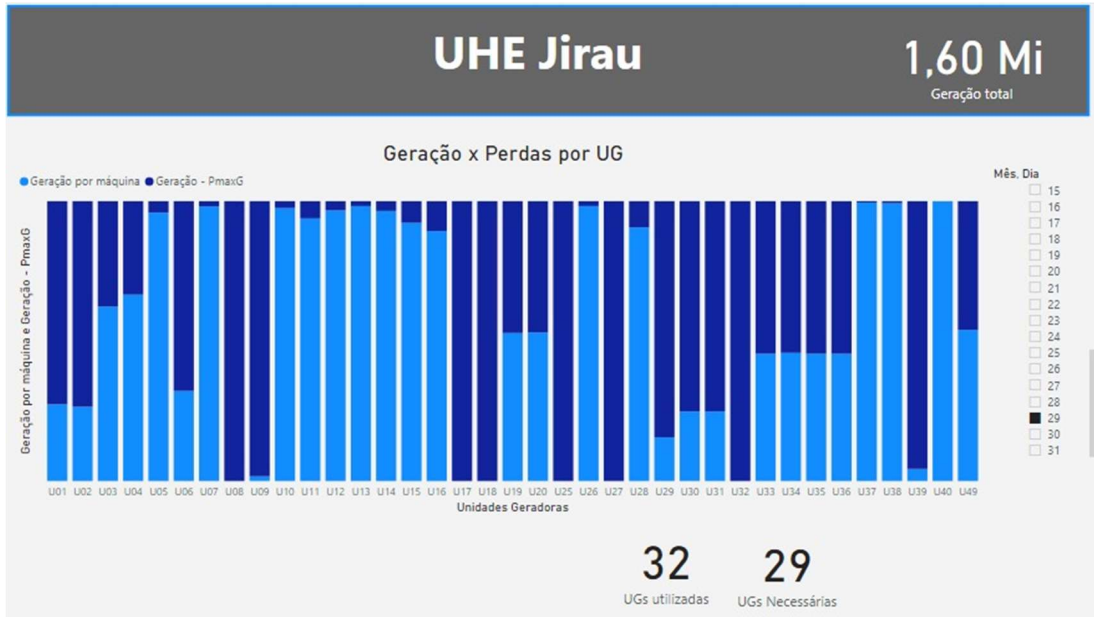


FIGURA 4: Geração máxima possível considerando a queda bruta

No *dashboard* apresentado na Figura 5, é exibida a geração efetiva realizada na usina em relação a geração máxima possível de ser efetuada. O *dashboard* é interativo, e possibilita a seleção do mês que se deseja visualizar, entre outros filtros e seleções.

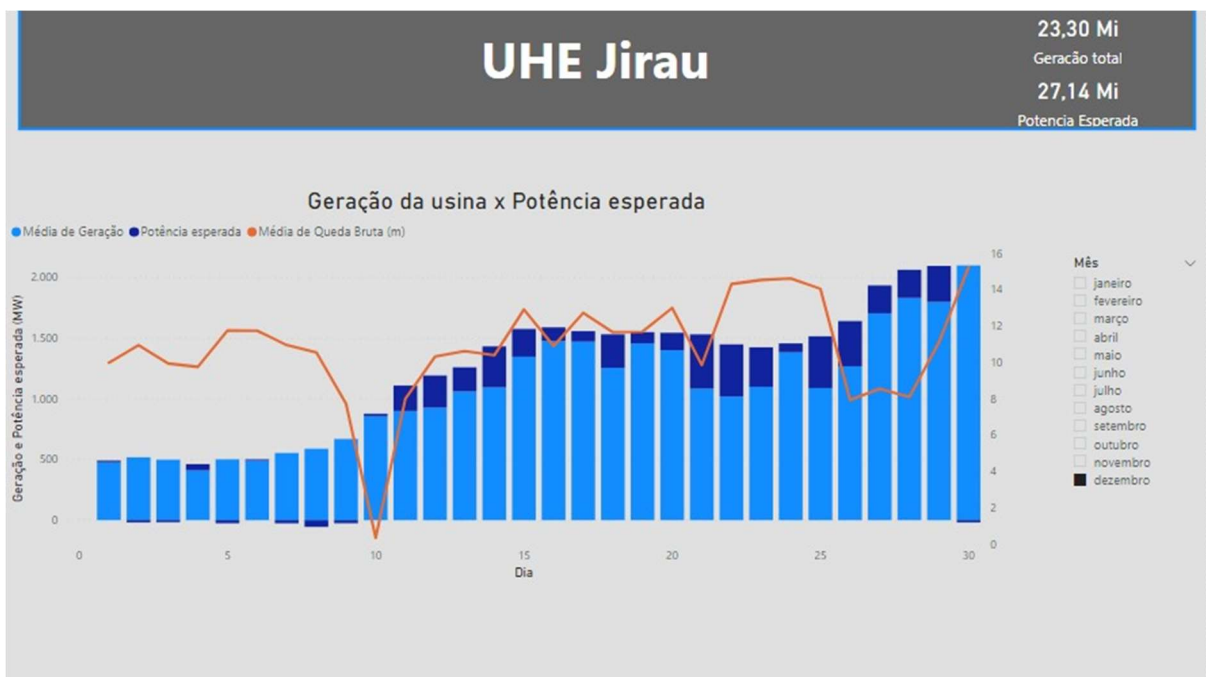


FIGURA 5: Geração efetiva da usina em relação à potência esperada.

O *dashboard* mostrado na Figura 6 estratifica a informação de geração por unidade geradora. A representação leva em consideração a geração efetuada apenas pelas unidades que estavam sendo utilizadas. Também é apresentando o valor máximo de geração possível de cada unidade geradora.

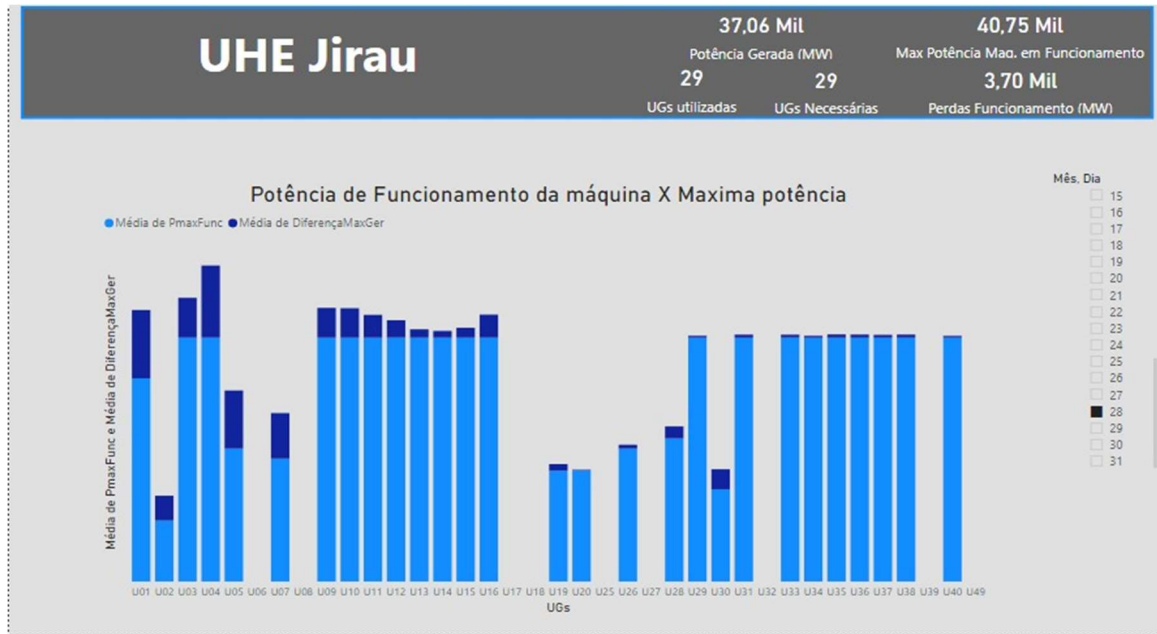


FIGURA 6: Geração possível x geração executada por unidade geradora.

## 5.0 CONCLUSÃO

Devido ao grande número de unidades geradoras, e consequentemente a quantidade de dados gerados por toda a estrutura associada à operação de usinas hidrelétricas, neste caso, especificamente a usina de Jirau, a elaboração de técnicas e ferramentas que possibilitem agregar valor aos dados e viabilize meios de visualizar e entender as informações contidas nesses dados se torna extremamente importante.

Com o intuito de viabilizar métodos para aprimorar a visualização e análise dos dados, e de acordo com os conceitos da indústria 4.0, efetuar a integração entre OT e IT de modo a aperfeiçoar a maneira como a operação é realizada, através de técnicas de processamento de dados e tecnologias de comunicação se tornam tarefas cada vez mais necessárias, principalmente em um cenário tão específico e desafiador como é o caso da usina hidrelétrica de Jirau, que devido à complexidade da operação, precisa de ferramentas que disponibilizem uma visão em tempo real do status de equipamentos e da operação como um todo.

A fim de agregar nos dados, e possibilitar uma melhor visão computacional do status da operação, este trabalho contribui para aperfeiçoar o processo de extração de informações utilizando um *framework* paralelo e distribuído, para integrar e transformar os dados e torná-lo disponível para operadores, supervisores e gerentes, para que estes possam operar da melhor forma possível a usina, aproveitando ao máximo os recursos hídricos disponíveis com o menor custo possível.

Com o conhecimento extraído, é possível auxiliar o planejamento da hidrelétrica ao longo de um horizonte de tempo de modo a atender a todas as restrições do sistema, tais como utilização do número adequado de unidades geradoras para atender a demanda esperada e redução no consumo energético dos serviços auxiliares, resolvendo o problema do comprometimento da unidade (*Unit Commitment - UC*).

Foi apresentado a arquitetura desenvolvida, composta por um conjunto completo de softwares e tecnologias integradas de forma a agregar valor aos dados e entregar resultados relevantes para a usina, através de cubos e visões de dados. Neste documento foram apresentados alguns *dashboards* simples, que utilizam dados após o processamento de BI, apenas com o intuito de demonstrar o funcionamento da plataforma.

## 6.0 AGRADECIMENTOS

Projeto regulamentado pela Agência Nacional de Energia Elétrica - ANEEL e desenvolvido no âmbito do Programa de P&D da Energia Sustentável do Brasil S.A. (PD-06631-0011/2020).

## 7.0 REFERÊNCIAS BIBLIOGRÁFICAS

(1) AHMED, HUSSNAIN. Applying Big Data analytics for energy efficiency. (2014).

- (2) BHATTARAI, BISHNU P., ET AL. Big data analytics in smart grids: state-of-the-art, challenges, opportunities, and future directions. IET Smart Grid 2.2 (2019): 141-154.
- (3) ROSSI, BRUNO, AND STANISLAV CHREN. Smart grids data analysis: A systematic mapping study. IEEE Transactions on Industrial Informatics 16.6 (2019): 3619-3639.

## DADOS BIOGRÁFICOS



Graduado em Ciência da Computação em 2005. Mestre em Ciência e Tecnologia da Computação pela UNIFEI (2013). Atuou como analista de sistemas na Devex, empresa da área de Mineração, atualmente incorporada pela Hexagon Mining. De 2012 a 2014 atuou como especialista em sistemas e como professor no INATEL. Em 2014 atuou como professor na Unifenas. É professor efetivo no Instituto Federal do Sul de Minas Gerais. Possui licenciatura em Matemática. Doutorando como aluno regular de Engenharia Elétrica pela UNIFEI. É bolsista pesquisador no projeto de P&D na usina hidrelétrica de Jirau, e trabalha com Big Data e Machine Learning.

(2) RICARDO EMERSON JULIO  
Doutorando em Engenharia Elétrica, Mestre em Ciência e Tecnologia da Computação (2015) pela UNIFEI e graduado em Ciência da Computação pela UFLA (2007). Trabalha com redes veiculares, robótica e otimização. Possui experiência com desenvolvimento full-stack, metodologias ágeis e liderança de equipes. Já foi professor lecionando as disciplinas de Sistemas Operacionais, Projeto Integrador, Algoritmos e Desenvolvimento Web e Mobile. Engenheiro de Software com mais de 10 anos de experiência em desenvolvimento de sistemas atuando como programador, arquiteto de softwares e líder de equipe.

(3) EDSON BORTONI  
Reitor da Universidade Federal de Itajubá (UNIFEI). Prof. Titular. Engenheiro Eletricista graduado pela UNIFEI (1990), mestre em Planejamento de Sistemas Energéticos pela UNICAMP (1993), Doutor em Sistemas Elétricos de Potência pela USP (1998), Livre Docente pela USP/EESC (2012), Professor Visitante na EPFL em Lausanne, Suíça (2015) e no Politecnico di Torino, Itália (2019). Suas linhas de pesquisa se concentram em instrumentação, eficiência energética, sistemas energéticos, máquinas elétricas e sistemas de geração de energia. Recebeu os prêmios ISA Power Industry Division Achievement em 2018, Hydropower Foundation Educator of the Year em 2019, e Distinguished Lecturer pelo IEEE/PES em 2019.

(4) GUILHERME SOUSA BASTOS  
Pró-reitor de Extensão e Professor Associado da Universidade Federal de Itajubá. Graduado em Engenharia Elétrica pela UNIFEI (2001), mestrado em Engenharia Elétrica pela UNIFEI (2004), doutor em Engenharia Eletrônica e Computação pelo Instituto Tecnológico de Aeronáutica (2010), com doutorado sanduíche realizado no Australian Centre for Field Robotics (2009). Já foi coordenador do programa da Pós-Graduação em Ciência e Tecnologia da Computação (POSCOMP) e do curso de Engenharia de Controle e Automação. Tem experiência nas áreas de Engenharia Elétrica, Automação e de Computação, atuando principalmente nos seguintes temas: automação de sistemas elétricos e industriais, sistemas de decisão estocásticos, machine learning.

(5) MARCELO OLÍMPIO DA FONSECA  
Graduação em Engenharia Elétrica pela Universidade Federal de Goiás. MBA do Setor Elétrico pela Fundação Getúlio Vargas. Gerente de Operação da Jirau Energia, concessionária da UHE Jirau.

(6) PEDRO PAULO DE CASTRO VIANA  
Bacharel em Engenharia Elétrica pela Universidade Federal de Rondônia (UNIR) em 2013. Trabalhou na pesquisa de eficiência energética no campus José Ribeiro Filho-Porto Velho no ano de 2009/2010 através do Programa de Extensão Universitária (PIBEX), estagiou no Tribunal Regional do Trabalho 14ª Região RO/AC envolvendo a eficiência energética em prédios públicos, foi bolsista do programa Ciência Sem Fronteiras-Espanha, onde cursou Engenharia de Energias Renováveis, tendo apresentado o TCC com o Título, "Utilização de máquina de indução duplamente alimentada em geradores eólicos". Atualmente cursa o MBA Executivo em Administração: Negócios do Setor Elétrico.