

CAPA

COMPUTAÇÃO DE ALTO DESEMPENHO



Santos Dumont, do LNCC: upgrade fará com que ele se torne o mais rápido do país

Expansão do supercomputador Santos Dumont e aquisição de novas máquinas por Inpe, Unicamp e consórcio de universidades paulistas ampliarão o poder de processamento de dados no país

Domingos Zaporolli

O Brasil projeta avanços significativos em sua infraestrutura de supercomputadores, com impacto sobre os recursos de computação de alto desempenho (HPC) disponíveis para a pesquisa científica. As principais iniciativas são do Laboratório Nacional de Computação Científica (LNCC), em Petrópolis, na região serrana do Rio de Janeiro, do Instituto Nacional de Pesquisas Espaciais (Inpe), em Cachoeira Paulista, no interior paulista, duas instituições vinculadas ao Ministério da Ciência, Tecnologia e Inovação (MCTI), e do Centro para Computação em Engenharia e Ciências da Universidade Estadual de Campinas (CCES-Unicamp).

Um supercomputador consiste em milhares de pequenos computadores chamados nós. Cada nó é equipado com sua própria memória e vários processadores. Atuando em conjunto, eles são capazes de realizar cálculos extremamente complexos em elevada velocidade, na casa dos quintilhões (10^{18}) de operações por segundo. Essas supermáquinas são uma poderosa ferramenta de pesquisa e também dão suporte a setores da economia como defesa, energia e saúde. Equipamentos mais potentes, explicam especialistas, possibilitam que cientistas resolvam problemas complexos em menor tempo.

Esses equipamentos podem apoiar estudos em diferentes áreas, como simulações da formação do Universo para compreender sua evolução, modelagem molecular na busca por novos fármacos e previsão do tempo e ocorrência de eventos climáticos extremos. Eles também são usados para ampliação de fronteiras de exploração de óleo e gás, pesquisa de novos materiais, desenvolvimento de projetos aeroespaciais e elaboração do potencial energético de fontes renováveis de energia.

“Com a incorporação da inteligência artificial nas atividades econômicas e sociais, a necessidade de supercomputadores capazes de lidar de forma eficiente com grande volume de dados e oferecer respostas rápidas às demandas será cada vez maior”, avalia o cientista da computação Paulo José da Silva e Silva, do Instituto de Matemática, Estatística e Computação Científica (Imecc) da Unicamp.

O LNCC programa para este ano a expansão do poder computacional do Santos Dumont, máquina dedicada exclusivamente à investigação acadêmica, que deverá passar a ocupar, com o upgrade, o posto de mais potente do país. Hoje, o SDumont, como é mais conhecido, tem capacidade computacional de 5,1 petaflops. Com a atualização, chegará a algo entre 22 e 25 petaflops – flops é um acrônimo em inglês para operações de ponto flutuante ou cálculos por segundo.

Cada petaflops equivale à capacidade de processar 1 quatrilhão (10^{15}) de operações matemáticas por segundo. O SDumont processa 5,1 quatrilhões de operações matemáticas em apenas 1 segundo. Se a mesma operação fosse realizada em computadores pessoais modernos, seriam necessários 24,5 mil aparelhos trabalhando em conjunto (ver *infográfico na página 16*).

Hoje, o maior supercomputador da América Latina é o Pégaso, da Petrobras, adquirido em 2022 da empresa francesa Atos por R\$ 300 milhões. Ele entrou em operação em dezembro daquele ano com poder para processar 21 petaflops. O equipamento figura na 45ª posição no ranking independente Top500, divulgado em novembro de 2023, que lista as máquinas mais potentes do mundo.

A 62ª edição do levantamento, lançado pela primeira vez em 1993, foi elaborada por pesquisadores da Universidade do Tennessee, do Laboratório Nacional Lawrence Berkeley e do National Energy Research Scientific Computer Center (NERSC), todos nos Estados Unidos.

O Brasil tem nove supercomputadores na relação – o melhor desempenho desde 1993, quando o Top500 começou a ser elaborado – e ocupa a 11ª colocação entre as nações com os maiores parques



de supercomputadores (ver *infográfico na página 21*). A Petrobras é dona de seis dessas máquinas. As outras três pertencem à empresa de softwares MBZ – os equipamentos NOBZ1 e A16A, ambos com arquitetura Lenovo – e à companhia de softwares e pesquisa SiDi – o supercomputador Iara, com tecnologia Nvidia, voltado a investigações na área de inteligência artificial (ver *infográfico na página 17*). Por uma questão técnica, o Santos Dumont não foi avaliado. O ranking só considera equipamentos com arquitetura computacional homogênea, ou seja, dotados de nós computacionais com a mesma configuração. Não é o caso da máquina do LNCC, que tem quatro arquiteturas distintas trabalhando em conjunto.

O líder do ranking é o Frontier, primeiro supercomputador na casa do exaflops, com capacidade de processar 1 quintilhão de operações por

O técnico Ruy Marvulle Bueno configura o supercomputador Coaraci, da Unicamp, para disponibilizá-lo aos usuários

O UNIVERSO DOS FLOPS

Unidade de medida equivale ao número de cálculos matemáticos que um computador faz por segundo

- Mega**flops = 1 milhão (10^6) de operações por segundo
- Giga**flops = 1 bilhão (10^9) de operações por segundo
- Tera**flops = 1 trilhão (10^{12}) de operações por segundo
- Peta**flops = 1 quatrilhão (10^{15}) de operações por segundo
- Exa**flops = 1 quintilhão (10^{18}) de operações por segundo

FONTE WILLIAN WOLF (UNICAMP)



Pégaso, da Petrobras: em operação desde 2015, é o maior supercomputador da América Latina e o 45º mais rápido do mundo

segundo, do Laboratório Nacional de Oak Ridge, nos Estados Unidos (ver reportagem na página 20).

A Petrobras será a responsável pelo investimento orçado em cerca de R\$ 100 milhões para o upgrade do SDumont. “Nossa expectativa é de que a instalação das atualizações seja concluída entre julho e agosto”, diz o matemático Wagner Vieira Léo, coordenador de Tecnologia da Informação e Comunicação do LNCC.

MARCO NA PRODUÇÃO CIENTÍFICA

O SDumont entrou em operação em 2015 e foi a primeira estrutura de computação do país com capacidade em petaflops – na época, 1,1. Em 2019, foi reconfigurado para 5,1 petaflops com apoio do consórcio petrolífero Libra, liderado pela Petrobras. Ele atua como nó central (Tier-0) do Sistema Nacional de Processamento de Alto Desempenho

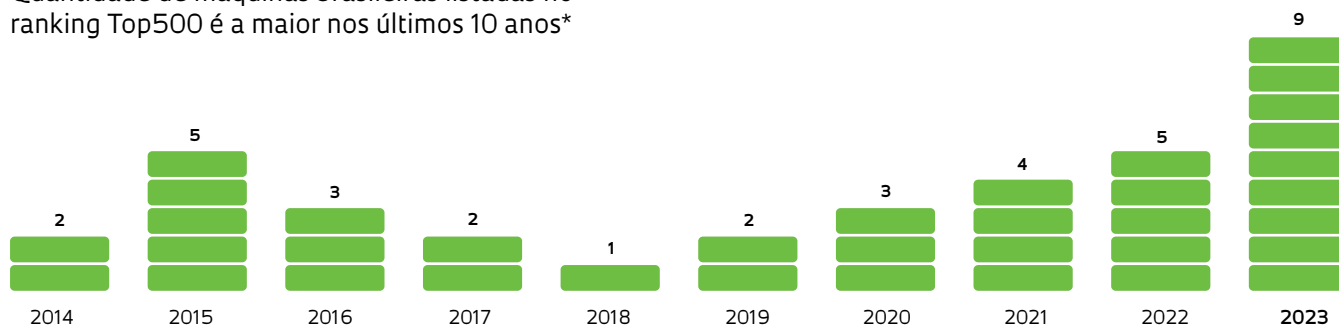
(Sinapad), composto por nove unidades de HPC vinculadas ao MCTI.

Até dezembro de 2023, o SDumont já havia sido utilizado em 430 projetos de pesquisa, que proporcionaram 1.044 artigos científicos, 43 livros ou capítulos de livros, 398 dissertações de mestrado ou teses de doutorado e 11 patentes. Hoje, ele tem mais de 2 mil usuários ativos. “O Santos Dumont é um marco na produção científica brasileira. Antes dele, o pesquisador precisava buscar parcerias no exterior para realizar seus trabalhos”, afirma Léo.

Entre as contribuições recentes do equipamento, está o sequenciamento de 19 genomas do coronavírus, tarefa realizada em 48 horas por pesquisadores do LNCC e das universidades federais de Minas Gerais (UFMG) e do Rio de Janeiro (UFRJ) em março de 2020, um mês depois dos primeiros casos confirmados de Covid-19 no Brasil.

FORTE EVOLUÇÃO

Quantidade de máquinas brasileiras listadas no ranking Top500 é a maior nos últimos 10 anos*



* O SANTOS DUMONT NÃO FOI AVALIADO PELO RANKING TOP500

FONTE 62ª EDIÇÃO DO RANKING TOP500

POR DENTRO DE UM SUPERCOMPUTADOR

Dotado de milhares ou até milhões de núcleos de processamento de alto desempenho, o equipamento pode pesar entre 20 e 40 toneladas

Alta velocidade de processamento e grande capacidade de memória são as duas principais características dos supercomputadores. Eles são formados por um conjunto de máquinas que atuam de forma integrada e pesam em média entre 20 e 40 toneladas. Os equipamentos são distribuídos em racks enfileirados e ocupam uma sala específica refrigerada. O Frontier, mais potente supercomputador do mundo, conta com 74 racks que ocupam uma sala de 680 metros quadrados.

O Santos Dumont, supercomputador do Laboratório Nacional de Computação Científica (LNCC), fica na cidade serrana de Petrópolis (RJ), que tem clima ameno, e ainda

assim demanda três geradores de 750 quilovolts-ampère (kVA) cada um, sendo dois em funcionamento e um para substituição, em caso de necessidade. É necessário 1.500 kVA para manter o sistema funcionando e a refrigeração contínua.

A velocidade de processamento de um supercomputador é estipulada conforme a rapidez que ele processa operações (flops). Os aparelhos mais potentes da atualidade processam operações na faixa de dezenas ou centenas de petaflops – cada petaflops equivale a 1 quatrilhão (10^{15}) de operações matemáticas por segundo – ou, na dimensão de exaflops, 1 quintilhão (10^{18}) de operações matemáticas por segundo.

Para realizar essa tarefa, essas supermáquinas somam milhares ou milhões de núcleos de processamento de alto desempenho, entre unidades de processamento central (CPU) e unidades de processamento gráfico (GPU). Com capacidade computacional de 5,1 petaflops, o Santos Dumont tem 36.472 núcleos de CPU e 1.134 nós computacionais.

A memória volátil ou RAM (*random access memory*) de um notebook comum varia entre 2 e 32 gigabytes (GB). Em um supercomputador, ela é medida em terabyte (TB, que equivale a mil GB). O Pégaso, da Petrobras, com 21 petaflops de capacidade de processamento, tem uma memória RAM de 678 TB.

DIFERENÇA ABISMAL

Compare a diferença do poder computacional dos supercomputadores e de um notebook comum



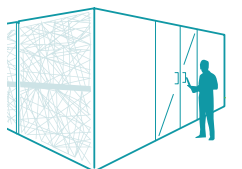
Notebook
(208 gigaflops)

208 bilhões de cálculos matemáticos por segundo



Coaraci
(801 teraflops)
801 trilhões de cálculos por segundo

3,8 mil vezes mais rápido do que o notebook



Santos Dumont
(5,1 petaflops)
5,1 quatrilhões de cálculos por segundo

24,5 mil vezes mais rápido do que o notebook



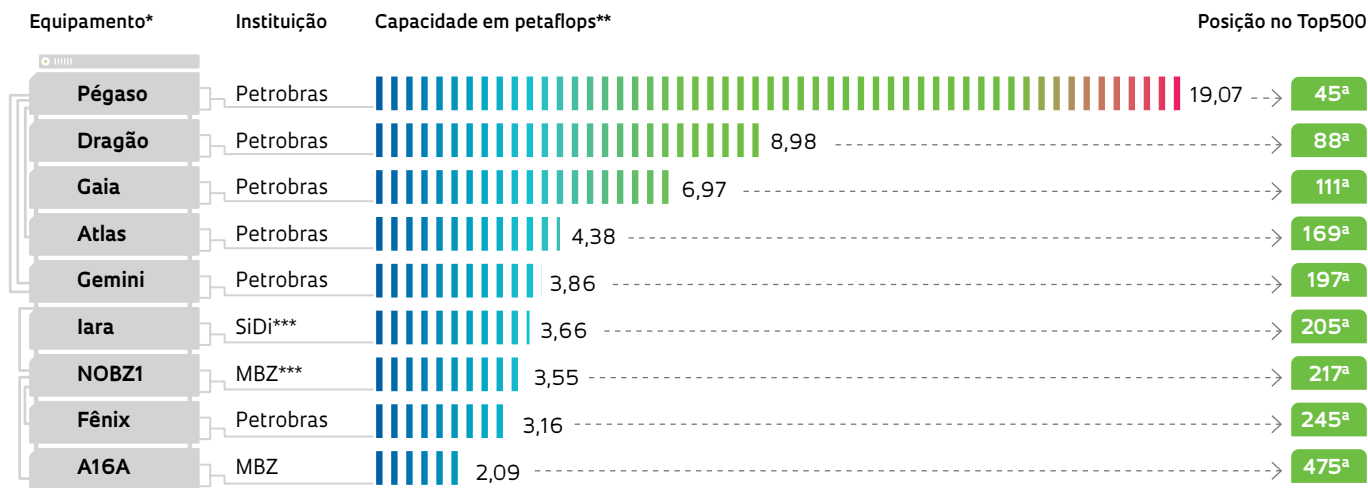
Frontier
(1,19 exaflops)
1,19 quintilhão de cálculos por segundo

O Santos Dumont também auxiliou cientistas da Universidade Federal de Pernambuco (UFPE) e da Fundação Oswaldo Cruz em Pernambuco (Fiocruz Pernambuco) no desenvolvimento de uma candidata à vacina contra o vírus zika. Ainda na área médica, o equipamento permitiu que equipes do Instituto do Coração (InCor) do Hospital das Clínicas da Faculdade de Medicina da Universidade de São Paulo (USP) e do Laboratório de Modelagem em Hemodinâmica (HeMoLab) do LNCC criassem modelos computacionais que simulassem o escoamento sanguíneo em artérias coronárias para o diagnóstico de risco de infarto do miocárdio.

“Hoje, a computação de alto desempenho está presente em quase todas as áreas da pesquisa científica. Até mesmo em humanidades se utiliza a análise avançada de dados no apoio à definição de políticas públicas”, exemplifica Silva, do Imecc-Unicamp, que também é o coordenador

SUPERMÁQUINAS BRASILEIRAS

Noves dos 500 supercomputadores mais potentes do mundo estão instalados no país



* O SANTOS DUMONT NÃO FOI AVALIADO PELO RANKING TOP500 ** A CAPACIDADE AFERIDA PELO TOP500 PODE VARIAR EM RELAÇÃO AO VALOR DIVULGADO PELAS INSTITUIÇÕES *** EMPRESAS DE SOFTWARE E PESQUISA

FONTE 62ª EDIÇÃO DO RANKING TOP500

do Centro Nacional de Processamento de Alto Desempenho em São Paulo (Cenapad-SP), uma das estruturas HPC que compõem o Sinapad.

O principal equipamento do centro paulista é o supercomputador Lovelace, adquirido em 2021 com recursos da Financiadora de Estudos e Projetos (Finep) e expandido em 2023 com o apoio da FAPESP. Nomeada em homenagem à matemática inglesa pioneira da computação, a máquina fabricada pela Dell Technologies apresenta em sua configuração atual capacidade de processamento de 388 teraflops. Cada teraflops proporciona 1 trilhão de operações por segundo. O centro também é aparelhado com o Tyr, que utiliza o sistema IBM Power 750, com capacidade de 37 teraflops. No total, os equipamentos do centro disponibilizam 425 teraflops de processamento.

Desde 1994, quando foi constituído o Cenapad-SP, até o final de 2023, as pesquisas feitas no centro geraram em torno de 4,4 mil publicações acadêmicas, 294 teses de doutorado e 331 dissertações de mestrado. Hoje são 221 projetos ativos

e 737 usuários de todo o país. “A infraestrutura tem 100% de ocupação e ainda há uma grande demanda represada na fila de espera”, relata o pesquisador da Unicamp.

Léo, do LNCC, e Silva compartilham a visão de que, após a atualização do SDumont, a principal prioridade da computação de alto desempenho no país voltada à pesquisa acadêmica é o upgrade da infraestrutura das demais unidades do Sinapad, uma vez que todas apresentam demanda de uso maior do que a capacidade de processamento (ver box na página 18).

Em janeiro deste ano, um primeiro reforço tornou-se realidade. O Coaraci, “mãe do dia” em tupi, novo supercomputador do CCES-Unicamp, passou por testes em dezembro de 2023 e foi disponibilizado para a comunidade acadêmica no mês seguinte. É uma máquina Dell de 801 teraflops, adquirida com apoio do Centro de Engenharia e Ciências Computacionais (CECC), um dos Centros de Pesquisa, Inovação e Difusão (Cepid) financiados pela FAPESP.

5,7 milhões de vezes mais rápido do que o notebook

“É o computador mais potente dentro de uma universidade brasileira e receberá pesquisas nas mais diversas áreas de interesse científico. Nossa avaliação preliminar é de que o Coaraci terá uma ocupação intensa, 24 horas por dia, sete dias por semana”, diz o engenheiro mecânico William Wolf, assessor técnico do CCES. O equipamento foi instalado no Centro de Computação John David Rogers, abrigado no Instituto de Física Gleb Wataghin da Unicamp.

Para Wolf, que é coordenador do Laboratório de Ciências Aeronáuticas da Unicamp, o Coaraci irá viabilizar trabalhos que eram difíceis de ser realizados por equipes brasileiras. As pesquisas aeroespaciais feitas na Unicamp, por exemplo, rodam em supercomputadores nos Estados Unidos e na França, um processo que leva em média seis meses por conta das longas filas e da burocracia com a preparação e a revisão de projetos internacionais. Além disso, esses projetos precisam entrar em uma fila de espera global e muitas vezes os resultados precisam ser compartilhados com grupos estrangeiros antes de ser processados.

“Esses dados são extremamente difíceis de obter, e nosso maior interesse é tratá-los. Em outros casos, os dados são sensíveis mesmo, pois se trata de inovação tecnológica”, explica Wolf. “Com o Coaraci, vamos fazer as pesquisas localmente, de forma rápida e segura, sem a necessidade de compartilhar dados que, em muitos casos, podem ser estratégicos para o desenvolvimento tecnológico no Brasil.”

O Inpe, instituição que também forma o Sinapad, dará início ao processo de licitação que levará à compra de um novo supercomputador capaz de aprimorar a previsão climática no país. O investimento programado é de R\$ 200 milhões em recursos do Fundo Nacional de Desenvolvimento Científico e Tecnológico (FNDCT). O valor será empregado em equipamentos e infraestrutura física nas instalações do instituto em Cachoeira Paulista (SP) voltados a dar o adequado suporte à máquina.

“O investimento será executado em etapas, iniciando-se este ano e com conclusão prevista para 2027”, detalha o doutorando em engenharia e tecnologia espaciais Ivan Márcio Barbosa, coordenador de Infraestrutura de Dados e Supercomputação do Inpe.

A primeira fase tem orçamento previsto de R\$ 47,5 milhões para a aquisição de equipamentos que vão gerar poder computacional de aproximadamente 2 petaflops. Ao final do processo em 2027, o sistema de alto desempenho do Inpe deverá somar 8 petaflops – capacidade cerca de 15 vezes superior à da supermáquina atual do instituto, o Tupã, adquirido em 2010, que oferece 550 teraflops (ver Pesquisa FAPESP nº 306).

Supercomputadores usam muita energia para alimentação elétrica e refrigeração. O Tupã, informa Barbosa, tem um custo anual de eletricidade

SÃO PAULO TERÁ NOVO CENTRO DE COMPUTAÇÃO DE ALTO DESEMPENHO

Consórcio formado por sete universidades contará com supercomputador de 5 petaflops

Em um ano, um novo supercomputador, com capacidade em torno de 5 petaflops, deverá entrar em funcionamento no país. A máquina equipará o recém-criado Centro de Supercomputação Científica do Estado de São Paulo (C3SP), uma das duas unidades contempladas com recursos de uma chamada lançada pela FAPESP e pelos ministérios da Ciência, Tecnologia e Inovação (MCTI) e das Comunicações (MCom) para apoio a centros de computação de alto desempenho.

“Essa é uma demanda antiga das instituições de pesquisa do estado, responsáveis por cerca de 60% do uso do Santos Dumont, o supercomputador do LNCC”, diz o meteorologista Pedro Leite da Silva Dias, diretor-executivo do C3SP. “Se quisermos ter um impacto maior

da pesquisa feita aqui, precisamos de máquinas competitivas.”

O C3SP é formado pelas três universidades estaduais de São Paulo (USP, Unicamp e Unesp), pelas quatro federais (UFABC, UFSCar, Unifesp e ITA) e as privadas FEI e Mauá. “Um edital para a compra do equipamento e a seleção de um centro de processamento de dados empresarial para abrigá-lo será lançado em breve. Optamos por um lugar neutro para não descaracterizar o conceito de consórcio”, explica Dias.

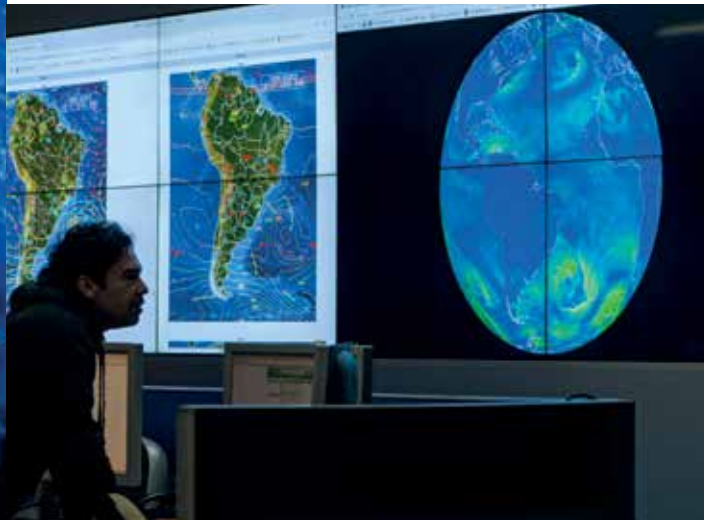
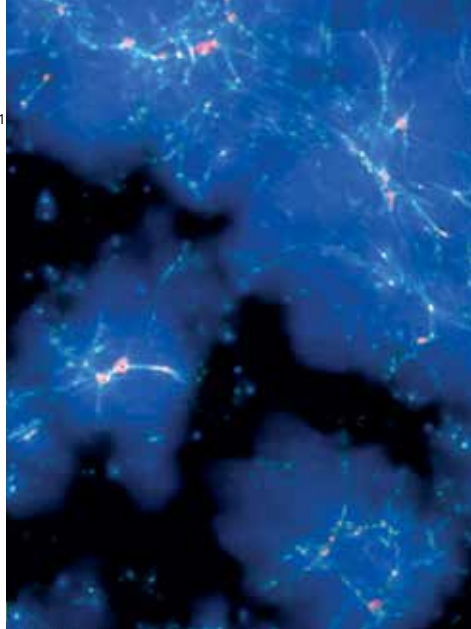
A segunda proposta contemplada foi o Centro de Computação de Alto Desempenho do Sinapad, o Sistema Nacional de Processamento de Alto Desempenho. “Nosso projeto insere-se em um contexto estratégico para a supercomputação no país, que é a revitalização do Sinapad”, afirma Antônio

Tadeu Gomes, coordenador-executivo da proposta e pesquisador do LNCC.

Os recursos, da ordem de R\$ 50 milhões, mesmo valor destinado ao C3SP, serão usados na atualização de cinco unidades do Sinapad: o Centro Nacional de Supercomputação, na Universidade Federal do Rio Grande do Sul, o Núcleo de Computação de Alto Desempenho, na Universidade Federal do Rio de Janeiro, o Instituto Metrópole Digital, na Universidade Federal do Rio Grande do Norte, e os Centros Nacionais de Processamento de Alto Desempenho de Minas Gerais e do Ceará.

O poder computacional agregado dos cinco centros hoje é próximo de 700 teraflops. Com a revitalização, deverá crescer para algo entre 4 e 8 petaflops. “Essas unidades apoiarão a pesquisa científica em diferentes regiões do país”, destaca Gomes.

Supercomputadores são fundamentais em pesquisas que simulam a formação do Universo (ao lado), na previsão do tempo e clima (à dir.) e no desenvolvimento de projetos aeronáuticos (abaixo)



casa dos 90%, para as próximas 48 horas. Passados sete dias, fica abaixo de 50%”, diz Barbosa.

O Tupã não será desativado. “Enquanto houver peças de reposição disponíveis, a máquina será dedicada a projetos de pesquisas científicas na área de tempo e clima”, afirma o pesquisador. Atualmente essa tarefa no Inpe é realizada em um *cluster* de processamento de alto desempenho, um conjunto de máquinas menores que alcançam uma determinada capacidade de processamento. O Egeon, o *cluster* do Inpe, tem poder computacional de 200 teraflops.

A CHEGADA DE GAIA

A infraestrutura de alto desempenho voltada à pesquisa no Brasil também foi reforçada em 2023 com a aquisição por R\$ 76 milhões pela Petrobras do supercomputador Gaia, de 7,7 petaflops, montado pela Dell. O Gaia entrou em operação em agosto do ano passado e é utilizado exclusivamente pelo Centro de Pesquisas, Desenvolvimento e Inovação da Petrobras (Cenpes). Seu objetivo é desenvolver e aperfeiçoar tecnologias ligadas à geofísica, entre elas ferramentas de processamento de imagens sísmicas, que usam ondas sonoras refletidas para criar uma espécie de tomografia computadorizada da subsuperfície da Terra.

No final de 2022, a Petrobras somava uma capacidade de processamento de 63 petaflops e anunciou a meta de alcançar 80 petaflops, sem especificar a data em que planeja atingir esse objetivo. Além do Gaia, também entrou em operação em 2023 o Gemini, dotado de 3,9 petaflops. Na área de produção, os supercomputadores da petrolífera são utilizados principalmente nas atividades de processamento sísmico e engenharia de reservatórios. No primeiro caso, o objetivo é determinar a presença de petróleo e a melhor área de perfuração de poços; no segundo, a finalidade é estudar o comportamento do óleo armazenado. ■

dade de R\$ 4,8 milhões. Sua refrigeração, como na maioria dos supercomputadores, conta com um sistema misto que envolve ar condicionado e água – são 40 mil litros circulando em circuito fechado. Para receber o novo supercomputador, o Inpe implantará uma subestação de energia elétrica de 138 quilowatts de pico (kWp) e renovará sua infraestrutura elétrica, de ar condicionado e água tratada. Também implementará uma usina de geração de energia solar fotovoltaica com capacidade inicial de 300 kWp.

A renovação do sistema de previsão climática incluirá ainda uma nova modelagem de previsão numérica de tempo e clima, batizada de Monan (Modelo para Previsão dos Oceanos, Superfícies Terrestres e Atmosfera), ainda em desenvolvimento. Segundo informações preliminares divulgadas pelo Inpe, além dos métodos tradicionais baseados em solução de equações físicas, o Monan usará inteligência artificial e aprendizado de máquina para prever o início e o fim de chuvas e eventos climáticos extremos com antecedência de três dias e a previsão das tendências climáticas com uma antecipação de três meses. “Hoje, nossas previsões são para até 15 dias, mas a precisão só é alta, na