RH RQ

Bringing great research ideas into open source communities

John Goodhue

Power surge: the push for sustainability in high-performance computing and AI workloads

+

Scaling the PEAKS of sustainability with insights from Kepler and machine learning

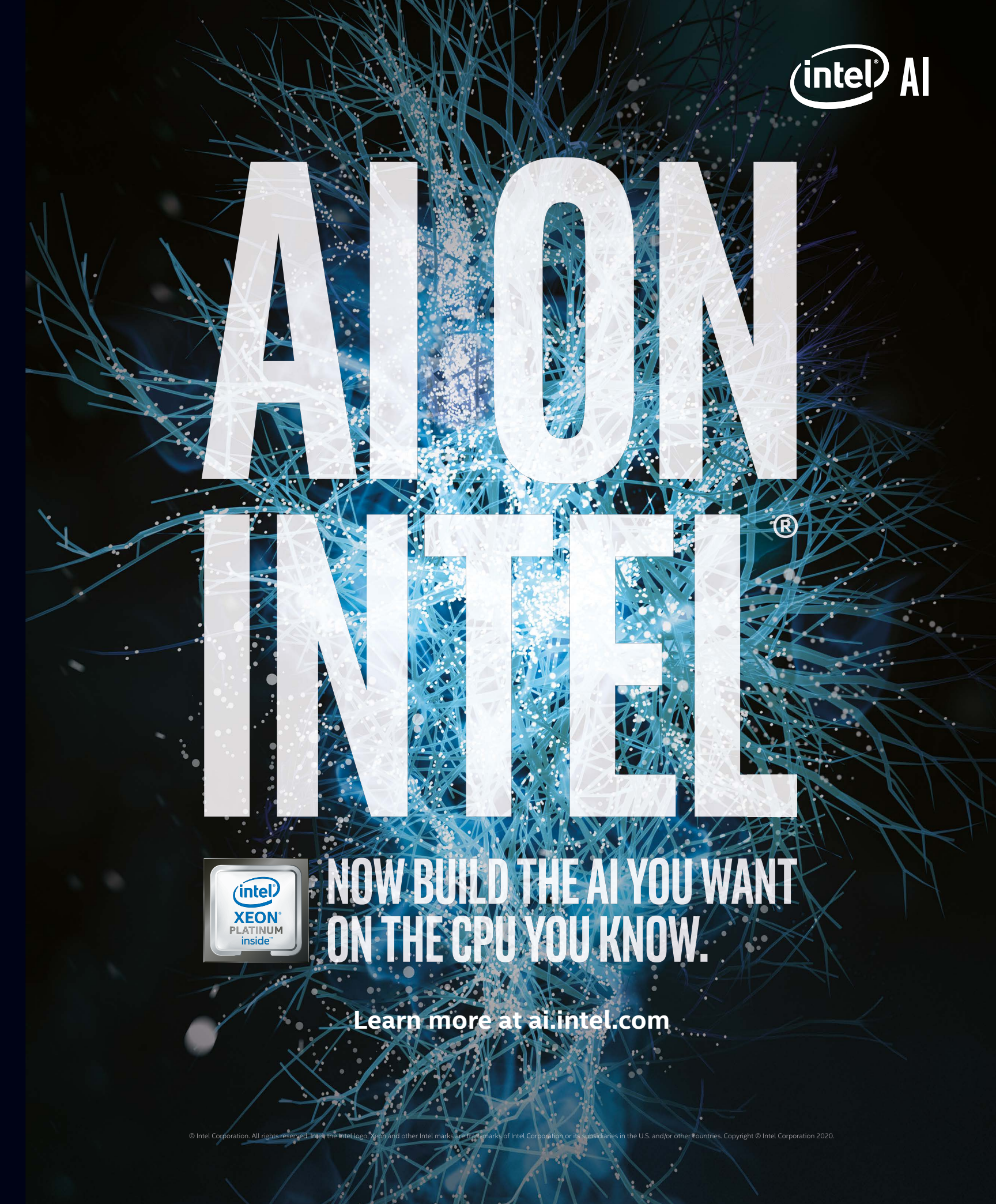
The Open Education Project is ready to scale

Open source authentication exposed

COLUMN

The Mass Open Cloud: enabling AI

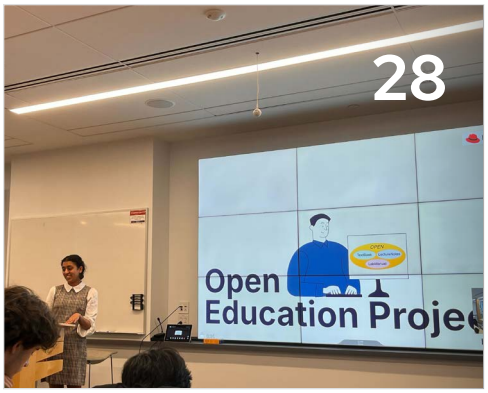
Red Hat Research Quarterly
Volume 6:2 | November 2024 | ISSN 2691-5278


intel AI
AI ON INTEL®
intel XEON PLATINUM inside™
NOW BUILD THE AI YOU WANT ON THE CPU YOU KNOW.
Learn more at ai.intel.com
© Intel Corporation. All rights reserved. Intel, the Intel logo, Xeon and other Intel marks are trademarks of Intel Corporation or its subsidiaries in the U.S. and/or other countries. Copyright © Intel Corporation 2020.

RESEARCH QUARTERLY

VOLUME 6:2

Table of Contents



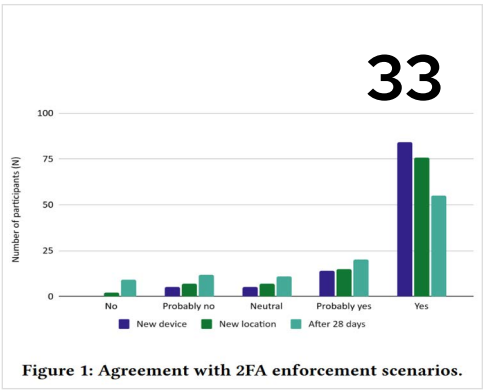


Figure 1: Agreement with 2FA enforcement scenarios.

Departments

Features

Red Hat

RESEARCH QUARTERLY

VOLUME 6:2

From the director

About the Author
Heidi Picher Dempsey is the US Research Director for Red Hat. She seeks out and cultivates research and open source projects with academic and commercial partners in operating systems, hybrid clouds, performance optimization, networking, security, AI, and operations.

From particles to prototypes: what we learn from managing open clouds

For those active in the early years of cloud computing, the challenges of open AI systems may feel strangely familiar. Do large-scale research collaborations have a lesson for today's AI developers and engineers? We think so.

by Heidi Picher Dempsey

With the proliferation of cloud computing in the early 2000s, IT organizations faced a new challenge: how to manage services in multiple different infrastructures efficiently, especially when cloud applications could span multiple services, which made isolating the root cause of a problem very difficult. Proactive monitoring, alerting, and response for a single application might require analyzing statistics from a large regional database service, a national high-speed backbone, some key regional and local caching infrastructure, and the local servers and laptops at multiple offices or campuses that actually initiated and provided the end user's window into the application. Each application had a different set of indicators used to determine whether it was operating correctly, so IT organizations customized their operations tools to correlate and communicate the state of those indicators to operators, whose goal it was to find and fix problems before users reported them.

In effect, IT organizations were building models of how healthy applications behaved and distributing that knowledge over multiple organizations and infrastructures so that it could be used to solve user problems. If this sounds like an AI domain, you're right. But at a time when people had just made it through the Year 2000 issue focused on using dates with four digits instead of two in all these services, there were no practical options for applying ML techniques to observability and many obstacles to even collecting the necessary data to apply those techniques. Instead, engineers focused on developing home-grown tools that could help sort through the data flooding in from diverse sources, looking for connections to solve individual problems.

Fortunately, the research community had seen into this future already, with the challenges of mass data storage, processing, and transfers that came with beginning construction on the Large Hadron Collider (LHC) in 1998. The LHC project enabled a large international community of over 10,000 very smart, enthusiastic researchers and developers in 100 countries to combine forces to build a distributed data collection and analysis tool for LHC experiments that spanned many infrastructures we would later come to call clouds. Emblematic of the large-scale collaboration needed for this challenge, the LHC tunnel itself belongs to no one country. It spans the France-Switzerland border near Geneva. The Worldwide LHC Computing Grid services launched in 2003, and fast development, test, deployment, and operations cycles for these services became the standard long before the term DevOps became popular in industry.

Critically, the LHC services required open source software for a large-scale distributed collaboration like this to work. Open source software, in turn, hastened the transition of these research-driven approaches for observing and managing large-scale data and distributed systems to some few enterprises that had started building their own large distributed infrastructures to manage services instead of science experiments. Google and Amazon began presenting at more academic conferences and publishing papers about cloud management, but they did not release their tools as open software. Splunk, which was founded in 2003 in San Francisco to provide a web-style interface to data collected and integrated in a central database, was an early market leader. Although Splunk was not open sourced, other tools like Kibana, Grafana, and Prometheus were evolving and began to be released widely in the 2010s.

The Large Hadron Collider tunnel, built by the European Organization for Nuclear Research (CERN)

Still, there was no agreed-upon definition of standard APIs and conventions for handling telemetry data consistently no matter what tool was being used to analyze it, so the barriers to large-scale collaboration between different organizations with applications that spanned multiple commercial services still existed. Open Telemetry, which finally began to be defined in 2019 with seed funding and technical committees in the Cloud Native Computing Foundation (CNCF), provided an increasingly popular way to avoid vendor lock-in for telemetry data and simplify software development to manage that data with a smaller set of APIs and conventions.

So why the history lesson? Because AI accelerated the pace of cloud development, but the problem of collecting, managing, and analyzing telemetry to support ever more complex services in the cloud still remains. The challenge of keeping not just the software that analyzes telemetry, but the data itself—along with the results and recommendations that AI service management software generates—visible to those who build, manage, and use cloud services from multiple providers is an even more challenging active research area than what we've faced before. Even just defining what constitutes a truly open AI system is controversial. The Open Source Initiative has been working on a definition of open source AI for years—as this issue goes to print, we're looking at an October 2024 announcement.

The ways in which infrastructure, data, and AI systems from different owners can be combined or federated into large-scale services have continued to proliferate. AI developers, network engineers, application developers, and university researchers are currently working together in organizations like the AI Alliance and Horizon Europe on prototypes that are meant to keep telemetry open and manageable by those who use the cloud, as well as those who create and run the commercial services they depend on.

AI telemetry solutions can help manage the data flood to find connections and make better recommendations to users and service owners, but they need a feedback loop that is open and accessible to give collaborators a chance to understand and influence decisions that affect them and their data. Red Hat and our research partners are actively collaborating on early prototypes for this work—as one example, expect to hear more about the Co-ops project, a novel framework for collaborative development of training AI models at scale that moves beyond the limitations of traditional federated learning, in the coming months. We're in the process of building multiple large clusters to support AI in this and other work in the Mass Open Cloud (MOC) that requires importing data from distributed sources, training and updating models, then distributing results worldwide. We're also excited to see results from the SEMLA project (Securing Enterprises via Machine-Learning-based Automation), which is looking at ways to integrate LLMs into system development and network configuration.

After many years of encouraging collaboration among diverse groups of humans to build federated cloud and data management systems, we're now bringing AI systems to the stakeholders' table. Will we build a new kind of open DevOps between human developers and AI operators for telemetry? We don't know yet, but we have some very smart humans and AI systems already colliding in a new kind of accelerator to find out.

> You can't manage what you can't collect!

IN THE QUARTERLY

AI systems are table stakes in another major challenge: sustainable computing. This issue of RHRQ features an interview with John Goodhue, the director of the Massachusetts Green High-Performance Computing Center (MGHPCC), the datacenter supporting projects in the MOC Alliance. John spoke with Red Hat Principal Software Engineer Parul Singh, a leader in both open telemetry and sustainability projects, and Boston University postdoc Han Dong, a researcher focused on using AI/ML to dynamically balance performance and energy efficiency. They suggest that we are just at the beginning of making plans for sustainable energy use and understanding the complex variables involved in reducing the impact of computing on the environment, from climate change to power grids. You can't manage what you can't collect! Parul and Han also give us a peek into PEAKS—the Power Efficiency Aware Kubernetes Scheduler. PEAKS uses insights from Kepler, an open source tool for collecting resource utilization metrics, and machine learning algorithms to dynamically tune the Linux kernel in a way that optimizes energy efficiency while still meeting specified performance requirements.

Meanwhile, we continue to scale the capabilities of the MOC for AI in part because of the research and educational opportunities it can support. Democratizing access to educational opportunities through open source technology is something we've long been active in at Red Hat Research. Danni Shi gives us an update on the Open Education Project (OPE), headed by PI Jonathan Appavoo. OPE has moved to the MOC Alliance's New England Research Cloud (NERC) and successfully supported classes for hundreds of students at Boston University. Danni describes the functional enhancements Red Hat engineers and BU faculty and students achieved in the past year to make OPE ready to host courses from universities around the world at an affordable price—and potential new users have already started to reach out.

As discussions about the possibilities and perils of AI take center stage in research and industry, these engineers and educators are already demonstrating how AI and large-scale systems are making a real-world impact, and what we need to measure to make sure that they continue to do that in many different collaborations.

MOC ALLIANCE

MAKING THE CLOUD LESS, WELL, CLOUDY

The Mass Open Cloud Alliance (MOC Alliance) is a collaboration of industry, the open-source community, and research IT staff and system researchers from academic institutions across the Northeast that is creating a production cloud for researchers. Of course, a collaboration is only as good as its collaborators.

Follow the MOC Alliance as they create the world's first open cloud.

@mass-open-cloud

www.massopen.cloud

contact@massopen.cloud

About the Author
Chris Tate
is a principal software engineer on the Red Hat Research team and a lead software engineer for logging, metrics, alerts, and AI/ML smart data research in the New England Research Cloud environment.

About the Author
Thorsten Schwesig
is a principal software engineer on the Red Hat Research team and part of MOC Alliance leadership for Red Hat. Thor is enthusiastic about collaborating with others to automate tasks and improve workflows.

Observability cluster added to the MOC Alliance's New England Research Cloud

Updates to NERC infrastructure enable fine-grained resource permissions for observability data.

by Chris Tate and Thorsten Schwesig

Observability data provides essential insights for optimizing performance, troubleshooting, and using resources sustainably. For users of the New England Research Cloud (NERC), part of the Mass Open Cloud (MOC) Alliance, this data also provides critical information for innovative research projects. Until recently, access to this data was restricted for most users.

A STANDALONE CLUSTER

NERC container infrastructure is based on OpenShift and includes several clusters (e.g., an infra cluster, prod cluster, and test cluster) operated within a VPN. Access to these clusters is therefore limited. This restriction especially affects observability data, such as metrics, logs, and traces. As the amount of observability data continues to grow, it becomes increasingly useful for research and teaching, independent of the applications, models, and data that generate it.

Initially, the observability data and systems in NERC, such as Thanos, Prometheus, Grafana, and Loki, ran on the infra cluster, which put higher demands on this cluster, which in turn can affect its operation in extreme cases. To enable access to observability data outside the VPN—and to relieve the infra cluster and separate tasks—we developed and implemented the idea of a standalone observability cluster.

Since March 2024, the NERC Observability Cluster has been running in its base version and has already successfully met several requirements. The cluster captures and stores metrics and logs with an increased retention rate and is accessible outside the VPN, which makes it much easier for researchers and educators to use. Additionally, we have made static dashboards for NERC data available in Grafana, providing a first basic visualization of the collected data to support analysis and monitoring, along with the ability to develop new dashboards.

CONTROLLING DATA ACCESS

With the NERC Observability Cluster in place, our next step was implementing fine-grained access control. With multiple research projects and classes hosted on NERC, maintaining data privacy compliance is essential. We needed to ensure that specific user groups, such as admins, researchers, professors, students, and apps (via API access), can access the data they need, and only the data they need.

Our primary challenges were ensuring seamless integration and maintaining high security standards. We accomplished this in May 2024 by introducing a new keycloak-permissions-operator to both operatorhub.io and Red Hat OpenShift to automate a previously missing feature of the Red Hat build of the Keycloak Operator. It introduces an advanced authorization feature of Keycloak and makes it easy to configure user, group, and application access to resources. We configure Keycloak for resource definitions, scopes, and permissions and set up a secure proxy to validate access tokens. We initially built these resources for the AI for Cloud Ops project team to give them access to certain metrics only on the prod OpenShift cluster. However, this operator was very reusable for other customers and projects as well.

The next step was to deploy a reverse-proxy (prom-keycloak-proxy) with fine-grained resource permissions authentication and authorization between applications on NERC and Red Hat Advanced Cluster Management (ACM) observability metrics. We've also shared this work with the ACM Observability team, which has features for fine-grained access to metrics on its roadmap.

FUTURE ENHANCEMENTS

In the next phase of this project, we will develop and implement mechanisms for data anonymization to ensure both privacy and usability of the data for research. We also plan to implement traces and develop interactive, dynamic dashboards that allow personalized and detailed data analysis. Additionally, the retention rate will be further optimized to support long-term analyses.

In time, we aim to introduce a proactive alerting and optimization system that captures event-based logs and provides targeted recommendations and optimizations. Additionally, we will continuously optimize the cluster's scalability and performance. We plan to promote use of the cluster by more research projects and institutions and integrate additional observability systems and data sources for a more comprehensive analysis of system performance.

The NERC Observability Cluster represents a significant improvement in the accessibility and usability of observability data for research and education. With ongoing development, it will meet growing demands and provide a solid foundation for innovative research projects. The key ideas and tools we've used can also be applied to other kinds of data that require fine-grained access control.

Keep up with our work on NERC Observability on GitHub.

The NERC Observability Cluster represents a significant improvement in the accessibility and usability of observability data for research and education.


Be bold. Be boundless.
Be a Baskin Engineer.
UC SANTA CRUZ
Baskin
Engineering
engineering.ucsc.edu

Publication highlights

Red Hat Research collaborates with universities and government agencies to produce peer-reviewed publications that bring open source contributions along with them. These research artifacts illustrate the value that open industry-academia collaborations hold not just for participants, but for technological advancement across the field of computer engineering. This is a sampling of recent papers and conference presentations; papers marked with a (🏅) were awarded special recognition. To see more visit the publications page of the Red Hat Research website (research.redhat.com/publications).

AI AND MACHINE LEARNING

"**Advancing cloud sustainability: a versatile framework for container power model training**," Sunyanan Choochotkaew (IBM Research), Chen Wang (IBM Research), Huamin Chen (Red Hat), Tatsuhiro Chiba (IBM Research), Marcelo Amaral (IBM Research), Eun Kyung Lee (IBM Research), and Tamar Eilam (IBM Research). In (2023) *Proceedings, IEEE Computer Society's 31st International Symposium on Modeling, Analysis, and Simulation of Computer and Telecommunications Systems, MASCOTS* (Stony Brook, NY), pp. 1-4.

🏅"**AutoAnnotate: reinforcement learning based code annotation for high level synthesis**," H. Shahzad (Boston University), Ahmed Sanaullah (Red Hat), Sanjay Arora (Red Hat), Ulrich Drepper (Red Hat), and Martin Herbordt (Boston University). In (2024) *25th International Symposium on Quality Electronic Design (ISQED)* (San Francisco, CA), pp. 1-9.

🏅"**AV4EV: open source modular autonomous electric vehicle platform for making mobility research accessible**," Zhijie Qiao (University of Pennsylvania, Autoware Foundation), Mingyan Zhou (University of Pennsylvania), Zhijun Zhuang (University of Pennsylvania), Tejas Agarwal (University of Pennsylvania, Autoware Foundation), Felix Jahncke (University of Pennsylvania, Technical University of Munich), Po-Jen Wang (Autoware Foundation), Jason Friedman (University of Pennsylvania, Autoware Foundation), Hongyi Lai (University of Pennsylvania, Autoware Foundation), Divyanshu Sahu (University of Pennsylvania), Tomáš Nagy (University of Pennsylvania, Czech Technical University), Martin Endler (University of Pennsylvania, Czech Technical University), Jason Schlessman (Red Hat, Autoware Foundation), and Rahul Mangharam (University of Pennsylvania, Autoware Foundation). In (2024) *IEEE Intelligent Vehicles Symposium (IV)* (Jeju Island, South Korea), pp. 2942-47.

"**Further optimizations and analysis of Smith-Waterman with vector extensions**," Reza Sajjadinasab (Boston University), Hamed Rastaghi (Boston University), Hafsah Shahzad (Boston University), Sanjay Arora (Red Hat), Ulrich Drepper (Red Hat), Martin Herbord (Boston University). In (2024) *IEEE International Parallel and Distributed Processing Symposium Workshops* (San Francisco, CA), pp. 561-70.

"XVO: generalized visual odometry via cross-modal self-training," Lei Lai (Boston University), Zhongkai Shangguan (Boston University), Jimuyang Zhang (Boston University), and Eshed Ohn-Bar (Boston University). In (2023) *IEEE/CVF International Conference on Computer Vision (ICCV)* (Paris, France), pp. 10060-71.

CLOUD COMPUTING AND EDGE

"Can OS Specialization give new life to old carbon in the cloud?" Han Dong (Boston University), Sanjay Arora (Red Hat), Orran Krieger (Boston University), Jonathan Appavoo (Boston University). In (2024) *Proceedings of the 17th ACM International Conference on Systems and Storage (SYSTOR '24)* (Virtual, Israel) pp. 83-90.

"CCO: Cloud Cost Optimizer," Adi Yehoshua (Red Hat), Ilya Kolchinsky (Red Hat), and Assaf Schuster (Technion, Israel). In (2023) *Proceedings of the 16th ACM International Conference on Systems and Storage (SYSTOR '23)* (Haifa, Israel), p.137.

"Enabling cost-benefit analysis of data sync protocols," Novak Boskov (Boston University), Ari Trachtenberg (Boston University), and David Starobinski (Boston University). In (2023) *Computer* 56:10, pp. 62-71.

"Experiences and lessons learned from PHYSICS: a framework for cloud development with FaaS," George Kousiouris (Harokopio University of Athens, Greece), Marta Patiño (Universidad Politécnica de Madrid, Spain), Carlo Sánchez (Atos, Spain), and Luis Tomás Bolivar (Red Hat). In (2024) *Euro-Par 2023: Parallel Processing Workshops* (Limassol, Cyprus), pp. 219-23.

EMERGING AND SPECIALIZED HARDWARE

"Effortless locality on data systems using relational fabric," Tarikul Islam Papon (Boston University), Ju Hyong Mun (Boston University), Konstantion Karatsenidis (Boston University), Shahin Roozkhosh (Boston University), Denis Hoornaert (Technical University of Munich), Ahmed Sanaullah (Red Hat), Ulrich Drepper (Red Hat), Renato Mancuso (Boston University), and Manos

Athanassoulis (Boston University). In (2024) *IEEE Transactions on Knowledge and Data Engineering*, pp. 1-12.

"**Improved models for policy-agent learning of compiler directives in HLS learning to drive anywhere**," Robert Munafo (Boston University), Hafsah Shazad (Boston University), Ahmed Sanauallah (Red Hat), Sanjay Arora (Red Hat) Ulrich Drepper (Red Hat), and Martin Herbordt (Boston University). In (2023) *IEEE High Performance Extreme Computing Conference (HPEC)* (Virtual).

"**Performance evaluation of VirtIO device drivers for host-FPGA PCIe communication**," Sahan Bandara (Boston University), Ahmed Sanaullah (Red Hat), Zaid Tahir (Boston University), Ulrich Drepper (Red Hat), Martin Herbordt (Boston University). In (2024) *IEEE International Parallel and Distributed Processing Symposium Workshops (IPDPSW 2024)* (San Francisco, CA), pp. 169-176.

"**Quantifying the gap between open source and vendor FPGA place and route tools**," Shachi Khadilkar (University of Massachusetts, Lowell), Ahmed Sanaullah (Red Hat), and Martin Margala (University of Louisiana, Lafayette). In (2023) *IEEE High Performance Extreme Computing Conference (HPEC)* (Virtual).

SECURITY, PRIVACY, AND CRYPTOGRAPHY

"**The adoption rate of JavaCard features by certified products and open source projects**," Lukas Zaoral (Red Hat), Antonin Dufka (Masaryk University, Czechia), and Petr Svenda (Masaryk University, Czechia). In (2024) *22nd International Conference on Smart Card Research and Advanced Applications* (Amsterdam, Netherlands), pp. 169-89.

"**Everlasting ROBOT: the Marvin attack**," Alicja Kario (Red Hat). In (2024) *Computer Security – ESORICS 2023* (The Hague, Netherlands).

"**Fingerprint forgery training: easy to learn, hard to perform**," Agata Kruzikova (Masaryk University, Czechia) and Vashek Matyas (Masaryk University, Czechia). In (2023) *Proceedings of the 18th International Conference on Availability, Reliability and Security (ARES '23)* (Benevento, Italy), pp. 1-7.

"**SREP: out-of-band sync of transaction pools for large-scale blockchains**," Novak Boskov (Boston University), Ari Trachtenberg (Boston University), and David Starobinski (Boston University). In (2023) *IEEE International Conference on Blockchain and Cryptocurrency (ICBC)* (Dubai).

"**TPMScan: a wide-scale study of security-relevant properties of TPM 2.0 chips,**" Petr Svenda (Masaryk University, Czechia), Antonin Dufka (Masaryk University, Czechia), Milan Broz (Masaryk University, Czechia), Roman Lacko (Masaryk University, Czechia), Tomas Jaros (Masaryk University, Czechia), Daniel Zatovic (Red Hat), and Josef Pospisil (National Cyber and Information Security Agency, Czechia). In (2024) *IACR Transactions on Cryptographic Hardware and Embedded Systems* 2, pp. 714-34.

"**Two-factor authentication time: how time-efficiency and time-satisfaction are associated with perceived security and satisfaction**," Agata Kruzikova (Masaryk University, Czechia), Michal Muzik (Masaryk University, Czechia), Lenka Knapova (Masaryk University, Czechia), Lenka Dedkova (Masaryk University, Czechia), David Smahel (Masaryk University, Czechia), and Vashek Matyas (Masaryk University, Czechia). In (2024) *Computers & Security* 138.

"**Uncovering CWE-CVE-CPE relations with threat knowledge graphs**," Zhenpeng Shi (Boston University), Nikolay Matyunin (Honda Research Institute, Germany), Kalman Graffi (Technische Hochschule Bingen, Germany), David Starobinski (Boston University). In (2024) *Cryptography and Security* 27: 1, pp. 1-26.

"**Understanding similarities and differences between software composition analysis tools**," Pranet Sharma (Boston University), Zhenpeng Shi (Boston University), Sevval Simsek (Boston University), David Starobinski (Boston University), and David Sastre Medina (Red Hat). Forthcoming in *IEEE Security and Privacy*.

"**What Johnny thinks about using two-factor authentication on GitHub: a survey among open source developers**," Agata Kruzikova (Masaryk University, Czechia), Jakub Suchanek (Masaryk University, Czechia), Milan Broz (Masaryk University, Czechia), Martin Ukrop (Red Hat), and Vashek Matyas (Masaryk University, Czechia). In (2024) *Proceedings of the 19th International Conference on Availability, Reliability and Security (ARES '24)* (Vienna, Austria), Association for Computing Machinery, pp. 1–11.

Says who?

/Keep your options open
redhat.com/options

Red Hat

A AWS

B Azure

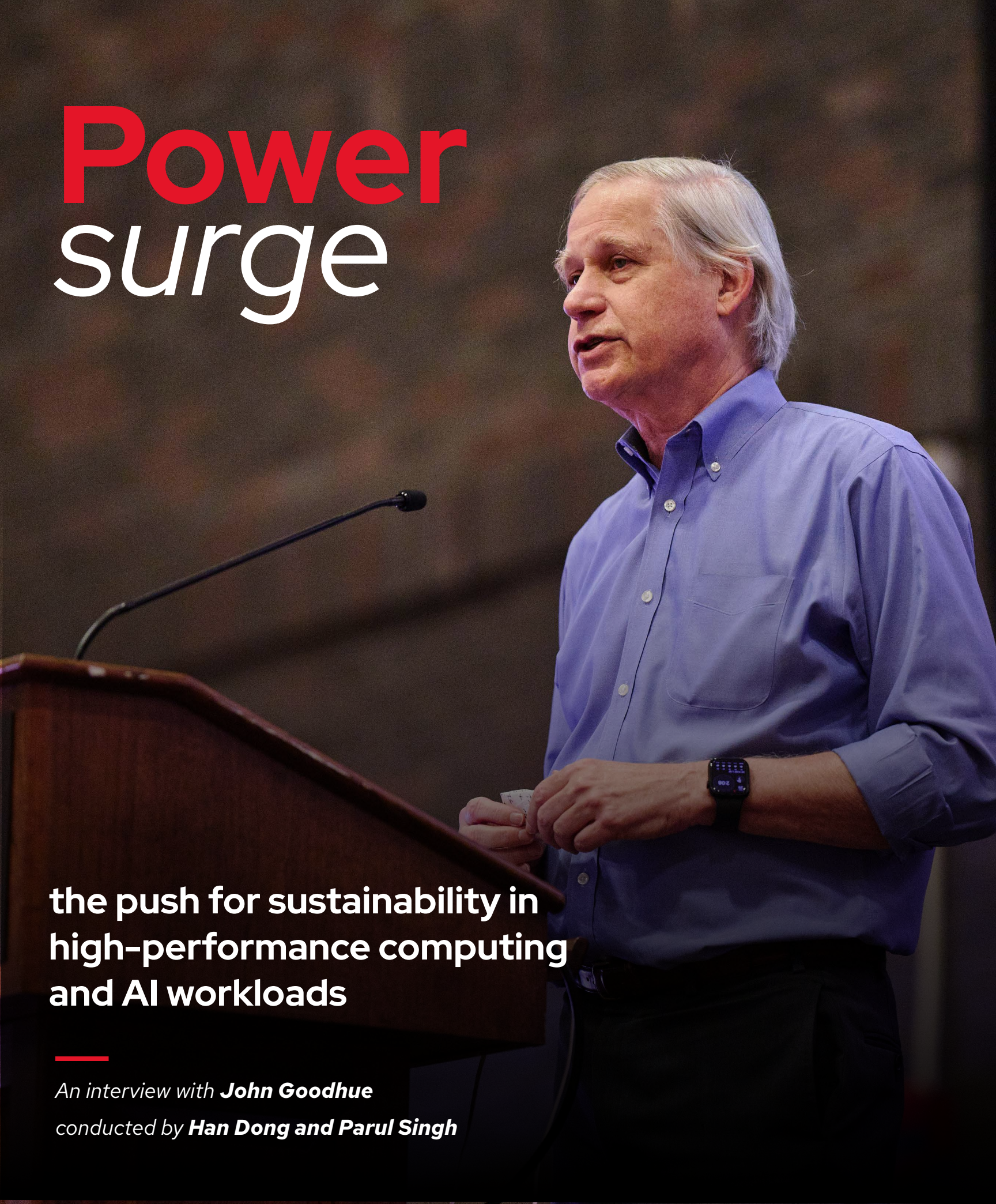
C Google Cloud

D All of the above

/Keep your options open
redhat.com/options

Red Hat

Power *surge*



the push for sustainability in high-performance computing and AI workloads

An interview with ***John Goodhue***

conducted by ***Han Dong and Parul Singh***

RESEARCH QUARTERLY

VOLUME 6:2

Red Hat

Interview

John Goodhue has perspective. He was there at the birth of the internet and the development of the BBN Butterfly supercomputer, and now he's a leader in one of the toughest challenges of the current age of technology—sustainable computing. Comparisons abound: one report says carbon emissions from cloud computing equal or exceed emissions from all commercial flights combined. Another suggested that by 2027 AI workloads could be using as much energy as a country the size of the Netherlands. That leaves computing and research communities with a conundrum: how do we solve real-world global challenges without worsening climate change or devastating power grids?

RHRQ asked Han Dong and Parul Singh, engineers working on open source projects related to energy efficiency, to talk with John about his work as director of the Massachusetts Green High-Performance Computing Center (MGHPCC), a joint venture of Boston University, Harvard, MIT, Northeastern, and the University of Massachusetts system. The MGHPCC provides computing and storage resources for over 20,000 faculty and student researchers and educators, including the MOC Alliance-supported New England Research Cloud (NERC) and the Red Hat Collaboratory at Boston University, and works to maximize energy efficiency while minimizing environmental impacts. They offer us a deep dive into the factors that bear on sustainability, from renewable energy sources and hardware choices to scheduling and tuning policies. The latter two, by the way, are the subject of Han and Parul's article on the PEAKS project, also in this issue. —Shaun Strohmer, Ed.

About the interviewer
Han Dong
is a postdoc in the Computer Science department at Boston University and a contributor to several projects at the Red Hat Collaboratory at Boston University.

About the interviewer
Parul Singh
is a principal software engineer and open source upstream contributor in the Emerging Technologies group of the Red Hat Office of the CTO.

Parul Singh: Tell us about your background in hardware. How did that lead you to eventually running the MGHPCC?

John Goodhue: Coming out of school, I ended up spending equal time on networks and computing. It was an exciting time for networking technology, with the technologies that became the foundations for the Internet just coming into play. At the same time, the company I was working at (BBN) was working on tiny networks that fit in a cabinet, interconnecting a large number of microprocessors to form a single computer system. That work led to the BBN butterfly, the second or third generation of the thing that we now call a cluster.

I ended up at the MGHPCC after leaving a startup in the high-performance computing business. It's obviously a very different thing: you can think of it as level minus one in the OSI (Open Systems Interconnection) stack. At the time we started to build it in 2010, the world was just beginning to realize that energy efficiency in datacenters might matter. Before then, even the idea of turning your air conditioners off in the winter was kind of a "why bother." That's how far the dark ages people were. We drove pretty hard on energy efficiency, and our timing was kind of lucky that way.

The Holyoke Dam as seen from South Hadley, MA, during the "freshet," or spring thaw. Photo by Simtropolitan, CC BY-SA 3.0

Parul Singh: From your perspective, what got people to start thinking about sustainability in computing?

John Goodhue: There was a turning point in the thinking in the industry around 2005, exacerbated by the recession and also influenced by growing awareness of the need to pay attention to climate change. An early challenge was finding ways to work with our architects and engineers to take steps that seemed radical at the time but are fairly commonplace. Though even today, developers of datacenters are under great pressure to get the thing built and into operation so they can start to get return on investment. Being thoughtful about energy efficiency and environmental footprint often takes a back seat as a result.

Going forward, we are beginning to see people think about what's going to happen when three things converge. First, you're seeing a dramatic increase in the amount of power consumed by datacenters, which is estimated to be as much as 3% of total worldwide energy demand, and may double in the coming years. Second, you're seeing large segments of the rest of the economy pivoting to electricity. The transportation sector and electric cars are a good example, but it's happening everywhere. Third, there's the introduction of renewables, which will make supply more intermittent.

The MGHPCC is lucky—maybe there's a combination of skill and luck—that we settled ourselves in a city (Holyoke, MA) whose municipal electric company delivers 100% green energy to us through hydroelectric power. But they're also forward-looking and interested in creative ways to make sure that the energy supply we receive has as light a footprint on the grid as possible. If you look out maybe 10 years, you may see, for example, battery storage as a factor. Our load is steady 24/7, but maybe we can lighten that load during peak hours and then make it heavy during non-peak hours in a managed way that potentially both saves money and offsets the impact of renewables. Tracking those types of change is a challenge that all datacenter operators face.

Han Dong: One thing we don't really know in the systems field is how power purchase agreements work with datacenters. Can you talk about that?

John Goodhue: The energy markets are a fascinating thing. They were developed when deregulation happened and transmission, generation, and delivery were split into three different things. The MGHPCC is dealing with the municipal electric company, which was grandfathered into being able to do all three. You don't get a discount for consuming more—or less, for that matter. They currently encourage us to have a load that's steady with the ability to drop or decrease if there's advantage to doing so.

If you look at the transmission rate, it's based on monthly peak consumption. There's an advantage to shaving peaks

that accrues to our supplier. We need to have a piece of the agreement that incentivizes us to do that. If you look at something called the forward capacity charge, that's based on your annual peak usage. That leads to a double incentive during one month of the year, but nobody knows what month that is until the end of the year. If you go to the ISO New England website, there is endless material on how the markets work. It's remarkable that the grid and energy markets work as well as they do given the complexity of the market and the engineering that supports them.

Han Dong: You mentioned the MGHPCC is 100% renewable. What does that mean?

John Goodhue: I like to say there are three ways of viewing it. First, where do the electrons come from? A quantum physicist would say who knows—everywhere. The more practical view is the electrical engineering view. There is a dam on the Connecticut River that generates way more energy than what we consume. It is connected to the same power substation that our primary power feed is connected to. From an electrical engineering point of view, you can't get any more tightly coupled than that to a renewable source. And there's the added benefit that we're not relying on transmission lines that have other environmental impacts. It's all right there within a half mile. The third view is the market view, where renewable energy like hydro or solar energy carries a premium. So we also pay the price premium for renewable energy. At both an electrical engineering level as well as the market level, we are 100% carbon free.

Han Dong: Is it reasonable to consider renewables free energy?

John Goodhue: It is absolutely not free. There are capital costs and maintenance costs. The only difference is that you're not burning fuel from Texas or Pennsylvania or the Middle East. The upfront costs can be higher. The good news is that solar is getting to a point where the lifetime cost of delivering a kilowatt hour of renewable energy continues to decline, but "free" is not a description you would attach to any energy source.

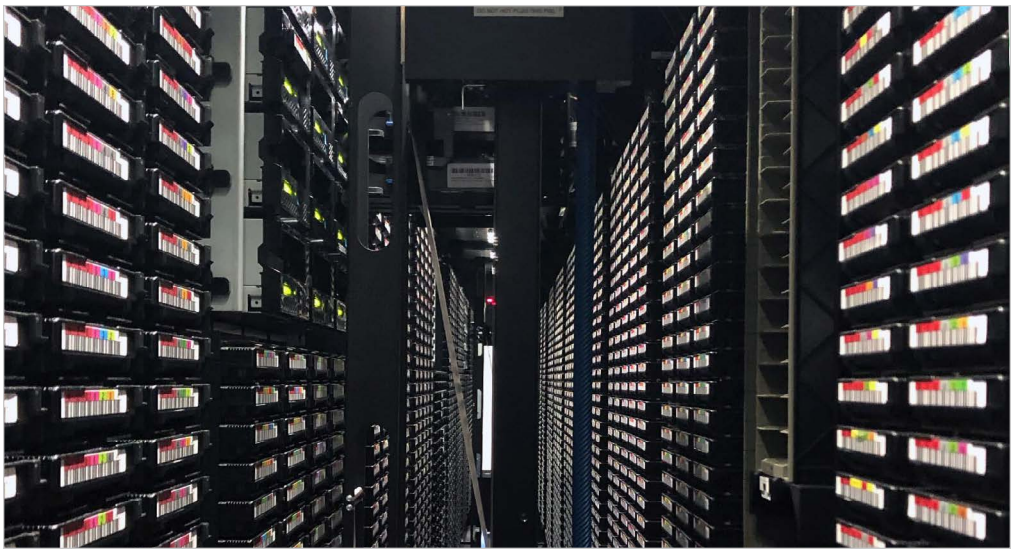
Han Dong: When you say upfront costs, are you talking about the embodied carbon that's attached to some of these sources that we don't really think about?

John Goodhue: There are things that emit carbon and there's things that don't, and the money the MGHPCC pays for electricity generation goes to places that don't emit carbon.

Han Dong: At an NSF Workshop last month, one of the issues discussed was embodied carbon in terms of datacenter servers. The standard longevity, at least for hyperscalers, is that these servers last at most about six years and then they replace them. There's a bigger push towards prolonging the life of some of this older hardware. How do we do that?

John Goodhue: That happens at a few different levels. First, lifetimes started to increase around 2005. Before then, CPU clock rates were steadily increasing from about the late 1980s through the early 2000s, and the performance

Developers of datacenters are under great pressure to get the thing built and into operation. Being thoughtful about energy efficiency and environmental footprint often takes a back seat as a result.



Racks of the Northeast Storage Exchange, created to address the escalating need for research data storage, housed at the MGHPCC

improvements in every three-year cycle were significant enough to justify replacing the server.

That equation changed as clock rates hit the wall at about three megahertz. As a result, lifetimes are now closer to five and six years—at least that's what we've been seeing. GPUs are an exception, as architectural changes introduce compelling improvements every two years or so. But even there you can keep the older ones around, to support less performance-sensitive workloads for them, and the motherboard, at least, has to change.

Han Dong: What are some examples of the architectural changes that are driving GPU performance?

John Goodhue: Many of them have to do with AI workloads, which can use fixed-point arithmetic, and algorithm improvements have made it possible to use much lower predictions. It's a complete U-turn compared to the drive for better floating-point performance needed for the scientific simulation applications.

Parul Singh: When I was working at Boston University, I learned that the MGHPCC datacenter is cooled naturally. Could you say more about how that's done, using water and the cold temperatures in the New England area?

John Goodhue: I'll unpack that a little bit. The cooling system for the datacenter really operates in three stages. We circulate cold water—not extremely cold, about 65 degrees Fahrenheit— into the computer room and run it through heat exchangers that either remove heat from the air that's ejected by the servers or remove heat through water that circulates right next to the chips. The second stage transfers heat to a different water loop that circulates through a set of cooling towers. The cooling tower uses two kinds of processes to remove heat from the water that circulates through them. One is evaporation—evaporation is a cooling process, as we learned in elementary school—and the other is, say in January, it's just cold outside so the water cools down.

There are two ways of moving the heat from the computer center loop to the cooling tower loop. Most of the year, we just use a heat exchanger because the water from the cooling tower loop is cold enough. There are times during the year in Massachusetts when it's too hot outside to allow us to cool the water enough using just heat exchangers. In that case, we use chillers, which are just refrigeration units that move heat from one place to another. The chillers are the biggest energy hogs in the building, so we operate to minimize the use of that resource. As I mentioned earlier, most datacenters didn't bother to minimize the use of chillers until the early 2000s, when datacenters started to get more energy conscious.

Han Dong: None of this information you're talking about is exposed to a user. Is there any value in getting this information to an end user? If there's a way to let someone know before they deploy their job that the cooling costs are likely to be high at a given time, is it worth incentivizing them to think, "Maybe I shouldn't run my heavy workload today because of the potential cooling costs?"

John Goodhue: I'll start with an analogy. I've spent a lot of time building networks, and one of the key reasons the internet works at all is layering, right? When I send a packet with a source and a destination address in a checksum using IPV4, I don't need to know anything about what happened next except whether the packet was dropped. Even that layer doesn't tell me how to figure it out. There's the layer above that, which works on end-to-end reliability when it's needed, or maybe it's tolerant of dropping packets on the way up the stack in the session that we're running. If every layer had to know what was happening in every other layer, you could never change it and it would frequently break.

The analogy isn't perfect but it holds. There are signals the grid can give our energy supplier that our energy supplier can pass through to us about peak demand, for example, and backing off when a multi-peak happens. It's feasible to pass those signals through, but you would need to react quickly. There is some promising research at BU and other universities looking at how to do that while minimizing impact on performance.

It is also possible to react to a signal at a lower level in the stack, by lowering the CPU clock rate. This has the advantage of being invisible to the application software, other than a decrease in performance.

Han Dong: The operating system has mechanisms to do exactly what you're talking about. It just doesn't have the policies there right now.

Let's address the question everyone's concerned about. From your perspective as director of a high-performance computing facility, what are the biggest challenges presented by generative AI and LLMs? For example, in terms of workload, how much of the work in MGHPCC consists of things like generative AI and large language models? I'm assuming it's increasing, right?

John Goodhue: It is increasing. If I use GPU installations as a proxy, it's growing substantially. The reason I can't say exactly is there's been a wave of installations with people who put in orders last year but supply chain problems kept them from actually getting their stuff until the last three months or so, so the systems are just beginning to come up.

Han Dong: Are a lot of them doing training jobs right now or are they also setting up inferencing servers on the MGHPCC?

John Goodhue: That's a good question. I suspect that there tends to be more training, compared to industrial workflows, but I don't know what the mix is.

Han Dong: Are you thinking about how the MGHPCC might adapt? Do you have to change how you think about the cooling and layout of the datacenter?

John Goodhue: The MGHPCC has been able to take this change in stride, in part due to early design decisions, and in part because we started supporting systems with high power density in 2016, well before the AI wave appeared.

Not much has changed with the AI boom, with one exception, which is the amount of power that gets consumed per square foot. Ten years ago, high-end enterprise workloads were six kilowatts per rack (KPR), and research computing workloads were probably more like 12 KPR. GPU-heavy research computing workloads can now be as much as 60 or 70 KPR, going up to 100 KPR. So we are seeing the amount of computing resources per square foot increase by factors of two and three. How we distribute power hasn't changed our cooling much at all.

The technology is evolving rapidly, and it's essential that research computing and datacenters do not fall behind their for-profit counterparts.

More broadly, the Coalition for Academic Scientific Computation (CASC) has organized several working groups around the topics of AI, energy efficiency, and building and maintaining datacenters like the MGHPCC. The technology is evolving rapidly, and it's essential that research computing and datacenters do not fall behind their for-profit counterparts.

Parul Singh: Thanks for your time, John. This has been quite interesting and very helpful.

THERE ARE MANY UNIVERSITIES IN MASSACHUSETTS, BUT ONLY ONE FLAGSHIP *FOR* MASSACHUSETTS.

As the Commonwealth's flagship public research university, UMass Amherst is committed to pursuing progress for our great state in computer science, technology, engineering, and more. Learn why we've soared to #26 in *U.S. News & World Report* rankings of top-tier public universities and find your degree.

Learn more at **umass.edu**

University *of* Massachusetts Amherst

RESEARCH QUARTERLY

VOLUME 6:2

Red Hat

Feature

Scaling the PEAKS of sustainability with insights from Kepler and machine learning

A proposed Kubernetes scheduler plugin aims to introduce energy efficiency as a factor in dynamic scheduling while still meeting performance requirements.

by Han Dong and Parul Singh

About the author
Han Dong
is a postdoc in the Computer Science department at Boston University and a contributor to several projects at the Red Hat Collaboratory at Boston University.

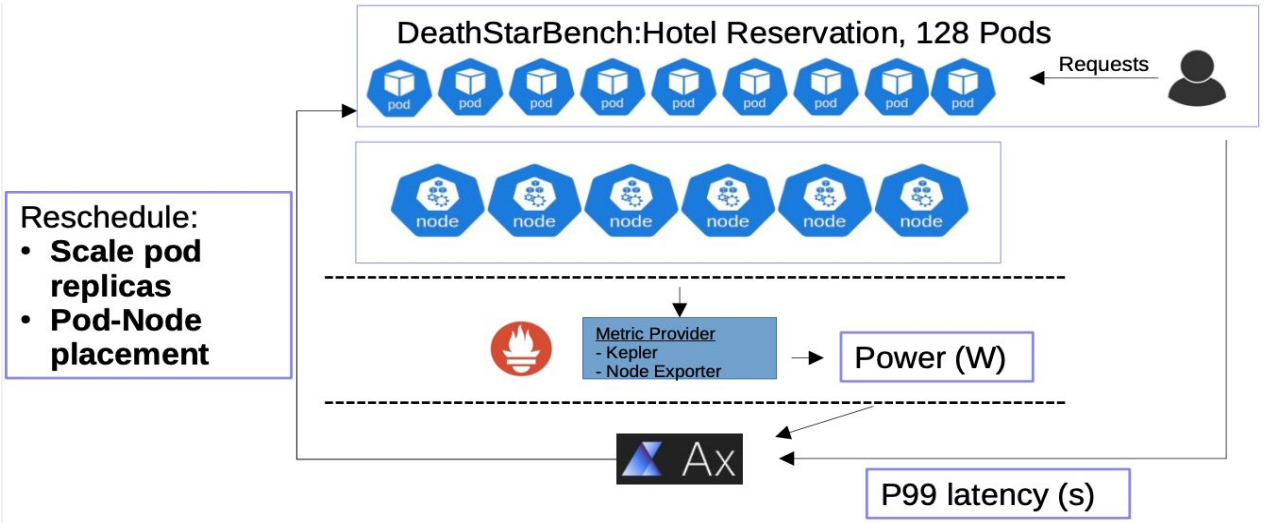
About the author
Parul Singh
is a principal software engineer and open source upstream contributor in the Emerging Technologies group of the Red Hat Office of the CTO.

Businesses in many sectors are setting aggressive sustainability goals, from transitioning to renewable energy sources to reducing existing consumption. Nowhere is the pressure to meet these goals more urgent than in the technology sector, where datacenters running AI workloads consume rapidly increasing amounts of electricity. Making those goals achievable requires greater visibility into energy usage and an intelligent way to adapt dynamically in response to changing requirements.

That's the idea behind the Power Efficiency Aware Kubernetes Scheduler (PEAKS), a collaborative project involving engineers and researchers from Boston University, Red Hat, and IBM. PEAKS is a scheduler that can factor in sustainability goals such as power utilization. It uses observability metrics to schedule tasks on a Kubernetes or OpenShift cluster more optimally. For example, an AI training algorithm is a long-running job that can go over days to train a model, or it could even be a continuous process to keep retraining the model whenever new data is found. PEAKS can find the nodes best suited for a particular workload to meet the application performance goals while reducing overall energy consumption.

PEAKS BACKGROUND

PEAKS uses insights from projects based in both industrial and academic research. The first is Kepler (Kubernetes-based Efficient Power Level Exporter), an open source project founded by Red Hat's Office of the CTO, with early contributions from IBM Research and Intel. Kepler offers a way to estimate power consumption at the process, container, and Kubernetes pod levels. The second set of insights come from a constellation of projects based at the Red Hat Collaboratory at Boston University that apply machine learning to Linux kernel configurations to optimize performance and energy tradeoff,



A motivational demonstration of applying dynamic tuning to pod scheduling in Kubernetes. We use the Bayesian optimization library provided by Ax.dev to automatically scale the number of pod replicas as well as the node the pods will be deployed on. The workload is the Hotel Reservation benchmark from DeathStarBench. Over time, this online learning approach seeks to improve both the P99 latency and power consumption to support user requests.

such as Automatic Configuration of Complex Hardware and Discovering Opportunities for Optimizing OpenShift Energy Consumption.

Kepler

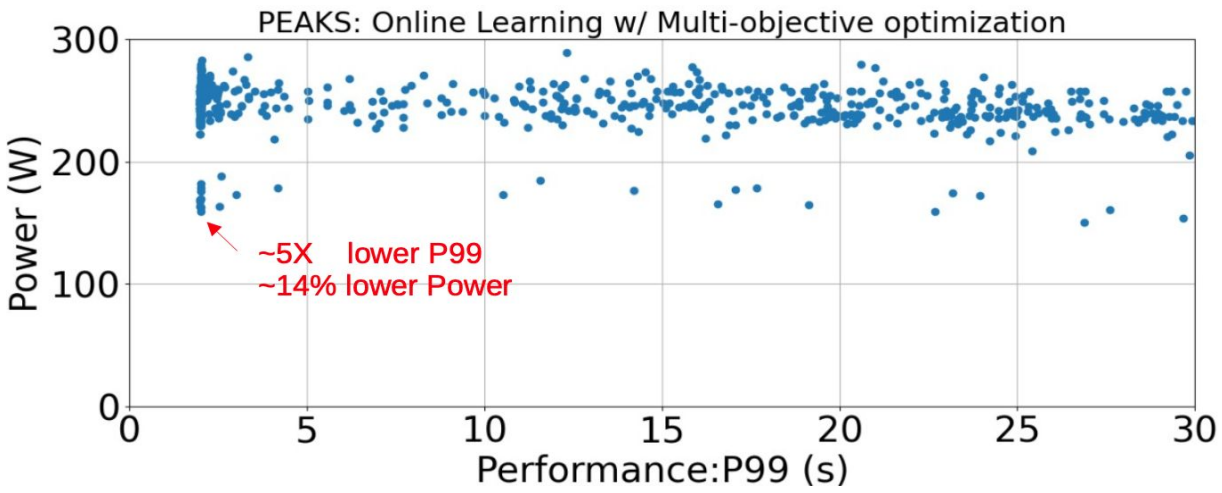
Kepler utilizes a BPF program integrated into the kernel's pathway to extract process-related resource utilization metrics. Kepler also collects real-time power consumption metrics from the node components by using various APIs or by using regression-based trained power models when no real-time power metrics are available in the system.

Once all the data are collected, Kepler can calculate the energy consumed by each process. It does this by dividing the power used by a given resource based on the ratio of the process and system resource utilization. Then Kepler aggregates the per-process energy consumption into totals for containers and Kubernetes pods. Data collected and estimated for the containers are then stored as time-series data in Prometheus.

Determining per-container energy consumption in a Kubernetes or OpenShift environment is a more difficult problem than it might seem at first glance. The way we proportion energy consumption is based not just on the computation for a given node. Many open source projects that report energy consumption give you only the dynamic energy consumption or the power you've consumed by running a computation. Kepler is following the greenhouse protocol, where you also must account for idle energy: the energy used for the maintenance of the nodes, such as cooling or just keeping those nodes alive. We do this by finding the idle energy of the whole node and distributing that evenly over all the containers. This is not strictly accurate, but it is a consistent basis using the best-case scenario.

Kepler also addresses the challenge of calculating energy consumption in a virtualized environment. It's relatively easy to calculate the energy consumption of processes or containers running on bare metal, where there's no abstraction. But in a virtualized and cloud environment such as AWS, you don't know how many VMs are running on a host. That data may not even be available to you because there could be multiple isolated tenants running their VMs on a particular physical machine.

In the multitenant scenario, cloud providers don't generally provide a solution. You must trust that AWS or Google report your total energy consumption not at the process level, but at the tenant level. In other



The result of the experiment: every dot in the figure is a single deployment configuration of the pods. The online learning approach was able to find a solution that is able to improve on the default behavior (proportional scaling) by a factor of 5 for performance while saving 14% in power.

words, they should report the total consumption of the server or servers hosting your processes, or the total consumption of a cluster that provides all your virtual nodes. There are certain hypervisors that can provide this data from their host machines, but most do not. If you're running your workloads on the public cloud, most of the time you won't have this information.

Kepler tackles this problem by using models for various CPU architectures. We created these models by running benchmarks on bare metal machines with various CPU architectures. Kepler uses these models to estimate the energy consumption of a virtual machine on a specific host based on the benchmarks for machines with the same CPU architecture. (For greater detail, read "Exploring Kepler's potentials: unveiling cloud application power consumption" on the Cloud Native Computing Foundation blog.)

Dynamic tuning

Policies in the Linux kernel can enable it to run in a more energy-efficient manner. With literally thousands of tunable parameters, there are many options to push the energy efficiency of Linux, especially in the way it runs workloads. By replacing some of the default Linux hardware policies with policies that are specifically tuned for energy efficiency, we can save a lot of energy. However, we must also ensure that the policy balances energy efficiency with performance for running applications. Service Level Objectives (SLOs) specify these required performance bounds for a system. For example, an SLO could be set so that 99% of incoming requests to the kernel are satisfied within one millisecond, which means we can exploit variations in this one-millisecond range to maximize energy savings. As these requests come in, a properly tuned system can respond to them in a way that satisfies the one-millisecond SLO while also saving energy.

To accomplish this, we use a machine learning algorithm—in this case, Bayesian optimization. The algorithm learns about the behavior of the application that's running and automatically tunes relevant kernel parameters to move towards that energy efficiency space. In testing with an assumed SLO of 99% response latency below 500 µs, we achieved energy savings of up to 50% compared to untuned default Linux. Even at a more stringent SLO of 99% latency below 200 µs, Bayesian optimization adapted to this performance requirement with up to 30% energy savings.

PEAKS

PEAKS is applying the same type of tuning techniques to

We take inspiration from previous kernel tuning work to decompose scheduling problems into a set of machine learning problems.

scheduling. There are two main areas in scheduling to consider: pod placement and autoscaling.

Pod placement involves selecting the physical node where a particular pod will run. Not all applications require the latest and greatest hardware. In some cases, it might be even better to utilize older hardware due to its existing embodied carbon costs (the amount of greenhouse gas emissions associated with the upstream stages of a product's life). Different generations of hardware (e.g., CPUs, GPUs) can have different Thermal Design Points (TDPs) due to their processor node technologies and other architectural characteristics. This is an interesting opportunity for tuning to maintain performance SLOs while slowly migrating functionality to older-generation hardware with drastically lower TDPs than the most recent versions.

Autoscaling involves determining the number of necessary pods to run as application needs change over time. The total number of pods can scale dynamically to meet SLOs as the workload changes. We can also leverage this information to improve energy efficiency. For example, if the load on datacenter servers follows a diurnal pattern, we may take advantage of this pattern to reduce energy consumption during periods of light load while still meeting performance goals.

To explore tuning for these scheduling aspects, we take inspiration from the previous kernel tuning work to decompose the above scheduling problems into a set of machine learning problems, using popular techniques including Bayesian Optimization, Bandit Optimization, Gradient Descent, and Simulated Annealing.

FUTURE MILESTONES AND CHALLENGES

One of the biggest challenges we have had in the Kepler project is validating its metrics. This is a general problem in AI when using models: how can you stand by the accuracy of your estimations? That's why we are running various benchmarks to validate the metrics on Kepler. If we know the metrics are off by a certain acceptable amount, we report that. One option we are exploring is using the power distribution unit (PDU) data from the Massachusetts Green High Performance Computing Center (MGHPCC) to verify whether the values measured by Kepler are comparable to measured power usage at the rack level.

In terms of kernel tuning, the next level to explore is how we can further tune the energy of jobs based on a varying abundance of green energy over time or location, or on the basis of cooling requirements. For example, we are investigating how kernel tuning could reduce carbon emissions from a refrigeration cooling system while also considering SLOs. Running jobs at night, when cooling needs are lower than during the day, might be an option for applications such as database backups, which must run daily but allow flexible scheduling.

You can learn more about PEAKS as a proposed Kubernetes scheduler plugin in the GitHub Kubernetes Enhancement Proposals (KEP) repository and in Han Dong's PEAKS presentation at DevConf.US.

UMass Lowell is proud to collaborate with Red Hat, a Select Preferred Partner, and celebrate more than a decade of working together on research, philanthropy and building the next generation of Red Hat's workforce.

UMASS LOWELL

Red Hat

RESEARCH QUARTERLY

VOLUME 6:2

Feature

About the Author

Danni Shi is a senior software engineer at Red Hat, leading the development effort for the OPE project. She is dedicated to advancing open source education and improving the accessibility of technology for all.

The Open Education Project is ready to scale

Enhancements to the pioneering platform for open source education have made it more reliable, easier to use, and much more affordable for new users.

by Danni Shi

What would it mean to open source education? For starters, we'd need a way for educators to create and publish their own high-quality open source materials—lectures, presentations, textbooks, and lab manuals—so they aren't locked into proprietary texts or software. We'd also want a way to deploy these materials in a live and interactive manner, at university scale, that makes it easy to collaborate and share content. To really make it open, we'd need to maximize accessibility, so a student can engage from anywhere simply by opening a web browser.

That's the aim of the Open Education Project (OPE), which launched in 2022 when it first received support from the Red Hat Collaboratory Research Incubation Award program. In the May 2023 issue of RHRQ ("Open source education: from philosophy to reality"), I wrote about our progress with OPE and future milestones. Since then, we've run multiple successful Boston University computer science courses on OPE and made several improvements. We're now ready to host courses from other departments and other universities from around the world.

OPE ON NERC

One of the biggest accomplishments of the past year is moving OPE from AWS to the New England Research Cloud (NERC). That gives us three useful benefits. First, it reduces the cost per student. NERC provides student access at cost, which is a fraction of the cost of AWS. In 2022, courses running on OpenShift AI on AWS cost roughly $150 per student, even after efforts to minimize expenses. After moving to NERC, the final cost is roughly $18 per student. This change makes OPE a potential solution for a much wider range of schools, including colleges and K-12 schools in under-resourced regions.

Second, we have more flexibility, because we are using our dedicated OpenShift AI cluster for classes. This means we can manage the cluster to support as many classes as we wish while ensuring that we have the scalability and load balancing for large class sizes. Third, we have more customized monitoring and management, which helps efficiently manage our resources. For example, we can monitor

Danni Shi presents OPE at the 2024 MOC Alliance Workshop.

when students' notebooks have been idling for a long time and shut them down after a specified period of time, keeping in mind that students in a machine learning class may need a longer window to run training processes. We also have a monitor that garbage-collects notebooks when students have a wrong image or wrong container size, so we can ensure students are using the right container size to start their notebooks.

The monitoring cluster also helps us find anything abnormal and better debug any issues, and we can capitalize on NERCs Observability Cluster to monitor resource usage and better understand student resource usage patterns. This also makes billing cleaner. Until the move to NERC, we could not differentiate between the usage in different classes maintained by a university. Moving forward, we can bill for individual class usage, so departments running less resource-intensive courses won't be charged the same rate as, say, a computer science department running several machine learning courses.

One significant challenge we faced was maintaining privacy for student notebooks. The problem was that each class had its own OpenShift namespace, mapped from a ColdFront project. Previously, when students launched the Jupyter Workbench via the Data Science project in their designated namespace, they could access all notebook instances within that namespace. This level of access allowed them to view, log in, and stop others' notebooks, posing a significant privacy concern. To address this issue, we enabled the **`rhods-notebooks`** namespace and directed students to use the Jupyter title that belongs to the **`rhods-notebooks`** namespace. This way, each student has access exclusively to their individual notebooks. This solution led to classes being able to efficiently share resources within the

We're now ready to host courses from other departments and other universities from around the world.

Intern Meera Malhotra presents her work on OPE at Boston University.

`rhods-notebooks` namespace—which is also important when accounting resource usage for each class.

FUNCTIONAL ENHANCEMENTS

Stability and reliability

One of our goals over the past year was to develop an automated test framework for OPE content that could potentially become part of the supported Red Hat OpenShift AI platform. We now have test automation that includes a GitHub CI/CD workflow, so when we build the containers, we use that workflow to test them. We also developed a GitHub CI/CD workflow that triggers when changes are committed to the container source code. This workflow verifies the functionality of the container image, including package versions, Jupyter features, and the user interface. This ensures that changes to the source code do not introduce errors, maintaining the build's integrity and reliability.

We also needed to develop tests for the cluster's functionality and scalability. For example, if we launch 300 Juypter notebooks at one time, how much latency will there be? In our previous ROSA cluster in AWS, when a large group of students launched all their notebooks at the same time, it could take nearly 20 minutes for each notebook to start as the AWS cluster scaled up. We wrote scalability tests to confirm that we have no latency issues starting large numbers of notebooks on the new cluster.

To make container builds faster and more reliable, we also reduced the container image size by using lightweight base images and multistage builds. Associate Software Engineer Isaiah Stapleton successfully squashed the image, meaning that an image potentially larger than 10 gigabytes—an excessive volume for an ordinary laptop—is reduced to less than two gigabytes per base image. Isaiah also developed a customized image-build process, allowing users to rebuild an existing OPE image with custom add-ons, such as modifying unique identifiers (UID), group identifiers (GID), and group settings within the container image.

Usability tools

The ability for users to create textbooks and interactive content is an essential element of OPE. To make it easier for authors to create static content or presentations, we created command-line tools to integrate OPE features. Adding a repository, creating a new OPE project, and building and publishing books can all be accomplished from the command line. The tool can be used inside an OPE container, which means that authors have a consistent lab environment while constructing textbooks. The command-line tool was developed in summer 2023 by BU intern Ke Li, and summer 2024 intern Meera Malhotra is working to improve it. Meera's contributions include enhancing the OPE tools usage documentation, simplifying the installation process, and adding tooling to the container environment for ease in managing dependencies for installations.

Meera is also working on a new OPE textbook, Beyond the Classroom (red.ht/beyond-the-classroom), with the goal of bridging the knowledge gap between education and writing code in the workforce. Meera gathered input from experienced Red Hat engineers to help identify areas where students need more instruction to transition successfully to the workforce. The textbook contains chapters on adjusting to writing code outside of the

classroom, learning how to understand regex, reading man pages, using the command line, and Git fundamentals. The book also includes interviews with engineers who share their insights with junior programmers. For example, security expert Lily Sturmann gives advice about how to write more secure code with clean coding standards, and Principal Software Engineer Sally O'Malley discusses the benefits of learning the command line and on-the-job learning. The interviews give not only technical advice but advice on adjusting to the workforce and combating impostor syndrome.

Autograder

BU student Ross Mikulskis, with mentoring from Isaiah, worked on an OPE Autograder, an application that receives homework submissions via a POST request, runs tests, and sends the results back to the client. The autograder integrates via API with the Gradescope platform, which manages student submissions, runs the autograder at scale, and distributes the results to both student and teacher. By automating the grading process we are able to provide immediate and consistent feedback to students, while enhancing the ability for these classes to scale.

OPE IN THE REAL WORLD

OPE infrastructure on NERC was used for three classes at Boston University in the Spring 2024 semester: Computer Systems (CS 210), with more than 320 students; Introduction to Operating Systems (EC 440), with more than 70 students, and Tools for Data Science (CS 506), with more than 220 students. Positive student feedback was that OPE was straightforward and easy to use with no onboarding difficulties.

Isaiah Stapleton and Danni Shi (left) work with Red Hat summer interns.

We have been approached by others interested in OPE, including Thomas McKenna from BU Wheelock College of Education and Human Development, the founder and creator of Phenomena for NGSS, an educational website designed to support teachers in learning more about phenomena-based instruction. In a recent OPE meeting, Thomas shared his Phenomena website and educational use cases and expressed his desire to use OPE to improve the teaching experience. Chris Simmons from the Massachusetts Green High-Performance Computing Center (MGHPCC) utilizes the Jupyter platform to train researchers and found that OPE aligns with several aspects of his work. A recent visit from students from Beijing, China, also sparked interest in collaboration with computer science faculty at Tsinghua University. We welcome their interest and encourage others to join us.

Building OPE is a highly collaborative learning process. Working with students and professors generates new and often innovative ideas on how to improve the books and platform. Students like Meera, who first encountered OPE in the CS 210 class, are also very active in creating the OPE platform. They bring valuable first hand feedback that helps us identify which parts of OPE need improvement and which things we want to maintain. This continuous feedback from both professors and students ensures that OPE effectively meets educational needs.

The OPE project is open to all. See the Open Education Project page for more details and links to repositories to help you start your own textbooks and classes.

MUNI FI

Masaryk University
Faculty of Informatics

Your **research,**

projects

and **education** partner.

FI.MUNI.CZ

RESEARCH QUARTERLY

VOLUME 6:2

Red Hat

Feature

Open source authentication exposed: how open source developers perceive user authentication

Ensuring security in open source software starts before a line of code is written. What role should communities and developers play?

by Agáta Kružíková

Open source projects are used in commercial products by many companies, from Microsoft and Google to Red Hat. The developers behind these projects and their user accounts are the first element in the supply chain, making their security crucial for maintaining trust and a good reputation. However, these developers are often independent, and essentially anyone can start contributing to open source projects. Since these independent developers are not bound by any company IT security policies, they may not use secure authentication, which means their accounts may be at risk of being misused or stolen. The impact of a compromised developer account can be substantial, given the large number of project users.

Supply chain attacks are a concern for proprietary software as well. A notable example is the 2020 sophisticated attack on SolarWinds, which highlighted the vulnerability of such systems. In early 2021, Microsoft disclosed a critical vulnerability in its Exchange Server to selected security partners, but there were signs that this information was later leaked without authorization. These incidents have seriously undermined trust in the contributors involved in the software development process, given their widespread impact on customers. Several high-profile incidents in open source projects have shown that weak authentication can lead to account misuse or theft. For example, the ctx project, available through the Python Package Index (PyPI), was compromised and replaced with malicious software following an account takeover. Similarly, security researchers were able to regain control of a widely used package in the npm (JavaScript package manager) repository by reclaiming an expired domain. These are not isolated incidents, but part of a broader pattern of security issues in open source ecosystems.

Since there may be no formal security policy for a project and user accounts, it is important to consider users' perceptions of proper account security and authentication mechanisms. Users are more likely to adopt secure behaviors if they

About the Author
Agáta Kružíková earned her PhD at the Centre for Research on Cryptography and Security at Masaryk University in Czechia. Her research focused on authentication from the usable security point of view, particularly user perception of two-factor authentication.

Research booth advertisement on DevConf.CZ

perceive them as meaningful and usable. To better understand user behavior in open source projects, we conducted two user studies. We focused on a specific aspect of upstream development: user authentication. For this purpose, we chose GitHub as the platform for open source projects, given that it is one of the most widely used development platforms.

WHAT 2FA METHODS ON GITHUB ARE USED, AND HOW ARE THEY PERCEIVED

The goal of the first study was to map the usage of authentication methods on GitHub and to understand how these methods are perceived by two groups: people who do not use the methods and those who are current users. The study was conducted online in November 2020 among Red Hat employees who volunteered to participate based on an invitation email sent to the mailing list. We collected data from 83 participants. We discovered that most users primarily use two-factor authentication (2FA) and consider usernames and passwords to be the most user-friendly method. In terms of security, hardware and software tokens were seen as the most secure options. However, the use of a third-party service for fallback authentication (via Facebook, which is no longer offered by GitHub) was generally not favored. You can find more information in the RHRQ article "User authentication for open source developers: what do they use?" (Aug. 2021) and in the conference paper "Authentication of IT professionals in the wild: a survey"(*Security Protocols* XXVIII, Oct. 2023).

WHY WE FOCUSED ONLY ON AUTHENTICATION

While discussing the results from the first study during an academic workshop and DevConf talk, some attendees questioned why we did not consider commit signing. While we believe that commit signing is also an important security measure, it serves a different purpose than user authentication. Our goal was to understand whether developers recognize that different security measures address different aspects of security, emphasizing the importance of authentication.

GITHUB ANNOUNCEMENT

However, in May 2022, GitHub announced plans to enforce two-factor authentication for contributors by the end of 2023 (and later extended this deadline to 2024 for some user accounts). They also reported that "only approximately 16.5% of active GitHub users and 6.44% of npm users use one or more forms of 2FA." This finding contrasts with our first study, where we discovered that 81% of participants used 2FA. However, our sample consisted only of people contributing to public projects. Also, our participants were employees of an open source-friendly company (in the first study) or attendees of an open source conference (in the second study). This suggests that we did not investigate the representative sample of GitHub users.

PERCEPTION OF 2FA

In our second, follow-up study, we examined user perceptions of GitHub's planned 2FA enforcement. Additionally, we investigated the role of authentication in the context of other security measures that can be implemented to enhance project security.

To improve the response rate in the second study and leverage the opportunity for personal contact, which may be more engaging, we continued our investigation at the 2023 Red Hat-sponsored DevConf.CZ in Brno, Czechia, with a research booth from Masaryk University. DevConf.CZ participants were invited to take part in the survey and were rewarded with merchandise provided by Red Hat. Data for this study were collected in June 2023, during the period when GitHub's 2FA enforcement was in progress.

To gain a better understanding of our participants, we asked them to provide details about a public project of their choice to which they contribute. We collected only publicly available data about these projects, such as the number of commits or stars. Based on this data, we observed a wide range of projects—from very small ones where the participant was the sole contributor to some significantly larger projects.

Overall, we collected data from 110 open source developers on GitHub. Our participants shared some common characteristics typical of this sample (e.g., around 80% male), but compared to the Stack Overflow Developer Survey, our sample was predominantly from the EMEA region and had higher education levels than the general developer population. Participants also rated themselves as generally knowledgeable and experienced in information technology security. Additionally, our participants held varying levels of access rights to the reported projects. Notably, 19% of participants had higher access rights than they actually used. Regularly reviewing and updating access rights for all project members could help in removing unnecessary permissions.

WHAT CHANGED?

Let's compare the results between the first and second studies. Some authentication methods offered by GitHub changed between the studies. For example, in the first study, fallback methods included recovery codes, SMS codes, and Facebook, while in the second study, only recovery codes were available as fallback options. In both studies, the primary second-factor methods included SMS, software tokens, and hardware tokens. However, in the second study, the software token could be activated either as an authentication application (as in the first study) or as GitHub Mobile via push notification, which was a new feature available in the second study only. Despite these changes, software tokens were positively perceived in both studies in terms of usability and security, and they remained the most commonly used methods.

In both studies, a majority of participants had enabled 2FA at the time of data collection: 81% in the first study and 68% in the second study. Most participants using 2FA in the second study had had it enabled for over a year, indicating that they did not enable it solely because of GitHub's enforcement. Regarding IT security policies, a similar trend was observed in both studies: approximately one-third of participants reported having a security policy applicable to their project, one-third reported not having such a policy, and one-third were unsure about the policy. Both studies focused on real-world developers and included a sample from this population, which is valuable given that many studies have either very small samples or use computer science students instead of experienced professionals. We would like to take this opportunity to thank all of our participants once again for their time and valuable input.

2FA ENFORCEMENT PERCEPTION

Regarding awareness of the planned 2FA enforcement, approximately half of the participants were aware of it at the time of data collection. We gained very positive insights into the planned 2FA enforcement—overall, participants supported it across

19% of participants had higher access rights than they actually used.

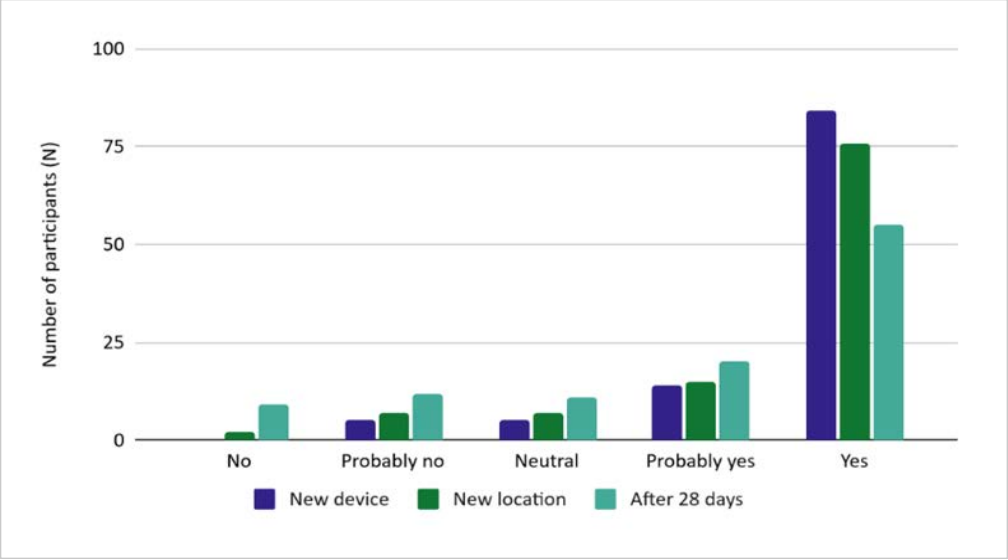


Figure 1. *Agreement with 2FA enforcement scenarios*

various scenarios and user groups. Participants were asked to consider 2FA enforcement for three scenarios: when logging from a new device, new location, or for control after 28 days. **Figure 1** shows the level of users' agreement with 2FA enforcement, separated by new devices, new locations, and after 28 days. In the case of new devices, 98 said yes or probably yes, compared to 91 for a new location and 75 for after 28 days.

When evaluating 2FA in the context of other security measures that can enhance project security, we provided participants with a list of security measures, including authentication, and asked them to evaluate them. The list was not comprehensive, as including all possible measures would have been too extensive for participants to consider. Instead, we selected a few relevant examples, including branch protection rules, signing releases, two-factor authentication, information on how to report security vulnerabilities (e.g., the SECURITY.md file), contact details for a security expert or team, signing commits, signing tags, use of release management tools (e.g., Jira), and advanced security testing (e.g., fuzzing). Participants view 2FA as an important element that should be included in a project's security policy, with a level of importance comparable to other key security measures. When comparing the project owner's use of 2FA with enforcing 2FA for all contributors, participants consider the owner's use of 2FA to be the most crucial security measure for protecting the project, and it is also the most commonly implemented. You can find more information in the academic article "What Johnny thinks about using two-factor authentication on GitHub" (Proceedings of the 19th International Conference on Availability, Reliability, and Security, Jul. 2024).

Even though only few participants reported personal or mediated (e.g., through colleagues or friends) experiences with security breaches, it is evident that most of our participants recognize the importance of securing their accounts in the context of open source and are likely aware of the consequences of account takeovers. These contributors are vital for the health and growth of the open source ecosystem. It is encouraging to see that there are user groups who value account security and support 2FA enforcement for contributors, perceiving the offered authentication methods as both usable and secure. However, there remains a concern about the users who have not voluntarily adopted 2FA on GitHub, as we identified that only a minority of users had not adopted 2FA voluntarily (and some users are strongly against the 2FA enforcement on GitHub, e.g., referring to 2FA as "nonsense" which "needs to stop,"). We leave it to future work to identify the difference between users who adopted 2FA voluntarily and involuntarily.

ACKNOWLEDGEMENTS

Thanks are due to Vashek Matyáš for supervision; all project members including Milan Brož, Martin Ukrop, and Jakub Suchánek, for their work; Red Hat for support; Cali Dolfi and Joshua Padman for consultations

THE UNIVERSAL AI SYSTEM FOR HIGHER EDUCATION AND RESEARCH

NVIDIA DGX A100

Higher education and research institutions are the pioneers of innovation, entrusted to train future academics, faculty, and researchers on emerging technologies like AI, data analytics, scientific simulation, and visualization. These technologies require powerful compute infrastructure, enabling the fastest time to scientific exploration and insights. NVIDIA® DGX™ A100 unifies all workloads with top performance, simplifies infrastructure deployment, delivers cost savings, and equips the next generation with a powerful, state-of-the art GPU infrastructure.

Learn More About **DGX** @ nvda.ws/dgx-pod
Learn More About **DGX on OpenShift** @ nvda.ws/dgx-openshift

Red Hat

RESEARCH QUARTERLY

VOLUME 6:2

Column

About the Author
Heidi Picher Dempsey
is the US Director for Red Hat Research.

Red Hat and the MOC-A: creating the open source cloud for the AI era

Our long-standing research partnership is now a strategic opportunity to advance open source AI, and everyone is invited.

by Heidi Picher Dempsey and Orran Krieger

About the Author
Orran Krieger
is the Director of Red Hat Research while on sabbatical from Boston University.

If you've followed Red Hat Research over the years, chances are you've heard of the Mass Open Cloud Alliance (MOC-A). The MOC is a production cloud service with 150 petabytes of storage, supporting containers through Red Hat OpenShift, an AI platform through Red Hat OpenShift AI, and bare metal access through Elastic Shared Infrastructure (ESI). When the MOC launched in 2014, it aimed to support research and education with a transparent alternative to opaque public commercial clouds. That goal was a perfect fit for Red Hat Research's initial focus: working with university partners to innovate in the hybrid cloud space and keep those innovations open.

The success of this relationship means we now have an unprecedented opportunity to bring new AI technologies from academic groups to the wider research community, and to grow and improve AI implementations at scale with real workloads. The MOC has become a strategic platform for Red Hat to partner with academic researchers and evolve open source software with research and educational users. Red Hat Research is working with the MOC-A to further develop this platform so researchers and engineers can improve open source AI tools and solutions, ensure that dominant AI technologies are open source, and build an ecosystem of partners and collaborators providing AI solutions and services. In addition, the MOC-A can enable access to AI infrastructure to researchers and new research startup projects that could not otherwise access open transparent AI implementations.

WHAT'S IN IT FOR YOU?

Great question, and we have a lot of answers, depending on where you operate in the technology landscape. To name a few:

Domain researchers

Domain researchers can leverage infrastructure to use in ways that won't happen with traditional public clouds. Red Hat and the MOC-A empower domain specialists to run AI workloads, support facilitation with the technology, and work with system researchers and industry partners to rapidly evolve the platform. Domain researchers also benefit from the experience and engineering collaboration of Red Hat's Office of the CTO. These experts bring the flexibility and transparency of the whole open source ecosystem to bear on projects that range from designing satellites to detecting disease. Red Hat's new

machine learning tools– InstructLab and Red Hat OpenShift AI–are also accessible to all MOC users.

Technology vendors
Technology vendors have an opportunity to reach influential audiences. Hardware providers can partner to expose new technology to a wide community of users. For example, Lenovo partnered with the MOC-A to provide 64 NVIDIA A100 GPUs under a new business model of GPU core-hour lease, and is now expanding this partnership with 192 H100 GPUs. Software developers can benefit both from building relationships with collaborating universities and driving adoption and by working with MOC-A resources, like telemetry, to optimize their solutions based on real usage.

All MOC users benefit from connections to the broader AI and cloud communities. Red Hat Research engineers are actively involved in community projects, contributing their software to the community and bringing the latest new ideas back to implement in MOC-A joint science, engineering, and start-up incubation projects. The MOC-A is an active member of the AI Alliance, supporting forums and working groups and providing technical leadership to some of the AI Alliance's most important initiatives. The MOC-A also works with multiple nonprofit organizations and government agencies to make the best of high-tech available to everyone, one project at a time.

WHAT'S HAPPENING NOW?

To galvanize our expanded focus on the MOC, in June 2024 Orran Krieger, Professor of Electrical and Computer Engineering, Boston University, and Mass Open Cloud PI, joined Red Hat Research while on sabbatical to help focus the research team–not just in Boston, but globally–on the goal of making the MOC an open, distributed platform for AI/ML workloads. With Heidi Dempsey, US Director for Red Hat Research, we are engaging industry and university partners, open source communities, and Red Hat business units to collaborate in making the MOC the platform of choice for open source AI development.

We've already begun to exploit the combined capabilities of the MOC and Red Hat OpenShift AI. One powerful example is the AI for Drug Discovery Forum, held in October to launch a new AI Alliance working group for drug discovery using new open source AI foundation models (Heidi is an interim co-lead for the working group). The forum also gave users hands-on experience running inference and using notebooks to submit various types of protein combinations to the models. A joint effort of Boston University, IBM, Red Hat, MOC-A, Cleveland Clinic, and AI Alliance, the forum demonstrates the power of interdisciplinary collaboration and open source AI models to break through research bottlenecks and ensure that solutions are ethical, accessible, and transparent.

An essential step in building this platform is building the ecosystem of contributors, users, advocates, collaborators, and partners who will benefit from it. If you are interested in partnering with Red Hat around the MOC-A, or would like more information, please contact Heidi Dempsey (hdempsey@redhat.com) or Orran Krieger (okrieger@redhat.com).

We now have an unprecedented opportunity to bring new AI technologies from academic groups to the wider research community, and to grow and improve AI implementations at scale with real workloads.

GET A LAPTOP THAT IS AI READY

AMD RYZEN™ AI TECHNOLOGY IS NOW BUILT IN

AMD
RYZEN AI

AMD RYZEN 7000 SERIES
AMD RADEON GRAPHICS

*Available on selected systems