IBM Cloud for Education Applications Lab

Evolution and Development

Written by:
Andy Rindos
Ravi Patil
Xianqing Yu
Young Oh
James Jushchuk
Blain Dillard
Anand Singh
Arunava Majumdar

IBM **Cloud**

IBM

Early concept

Nearly two decades ago, the term “cloud computing” was only just beginning to attract attention from industry and government. But North Carolina State University had already begun work on a radical solution that would replace their existing on-campus server labs, which up to that point had required significant upkeep and investment. They called their new development the "Virtual Computing Laboratory" (VCL).

This early cloud management software stack enabled faculty, students, and staff to request — anytime, anywhere — virtual machines, entire dedicated servers, and even integrated research clusters of servers. All of these available systems operated inside their central data centers (see Figures 1 and 2). While it initially specialized in delivery of Infrastructure as a Service (IaaS), the VCL was capable of supporting any NIST-defined or other recognized cloud service paradigms (PaaS, SaaS, XaaS). The solution was agnostic to the operating system, hypervisor, and hardware requested. And the physical infrastructure was distributed across multiple data centers within the state of North Carolina, creating a seamless (location-agnostic) reservoir of accessible resources. Within a few years, NC State had built a multi-thousand (physical) server cloud that provided access to users from the University of NC System (state universities), the NC Community College System, and select NC K–12 school districts.

At the same time, the VCL-based cloud paradigm was being explored and adopted by many other universities in the region. These included the UNC System, Duke, Clemson, the Virginia Commonwealth schools, the University System of Maryland, the Tennessee Board of Regents, historically black colleges and universities (HBCUs) participating in the Technology Transfer Program (TTP), and others along the US east coast. It was also adopted by select campuses in the California State University System and other US schools, as well as institutions in Canada, Mexico, Europe, and South America.

IBM was involved in developing the solution from the beginning, and these efforts began many years before IBM released its initial Cloudburst and SmartCloud offerings. At its inception, the NC State VCL-based cloud was running, for the most part, on the very first IBM BladeCenter Servers. Other cloud instances in Canada incorporated IBM Power servers. The first release of IBM BladeCenter Servers at NC State used the IBM-developed open source deployment engine xCAT (Extreme Cloud Administration Toolkit) to stand up its bare metal and virtual servers and clusters. And through VCL, the NC State cloud was able to loan its hardware resources to its huge and separate HPC server farm, that was managed by IBM Spectrum Symphony LSF software — especially at night, when traditional cloud resource requests waned.

Open source contribution

The IBM WebSphere organization contributed by assisting VCL in becoming an open source Apache Software Foundation incubator project in 2008. It has since graduated to a “top level” project in 2012, indicating that it had successfully acquired a self-sustaining contributor community. In 2009, IBM Global Education Industry General Manager, Mike King, announced the launch of the “IBM Cloud Academy.” (The name has since been adopted by another IBM program.) This initiative, including member universities from around the globe, established an IBM/university ecosystem dedicated to working together to adopt and improve the newly emerging cloud computing configuration. It would go on to launch joint research and development projects, create shared reusable assets, pursue joint funding proposals from state and federal programs, and host workshops and academic conferences — including the annual International IBM Cloud Academy Conference (ICACON see below).

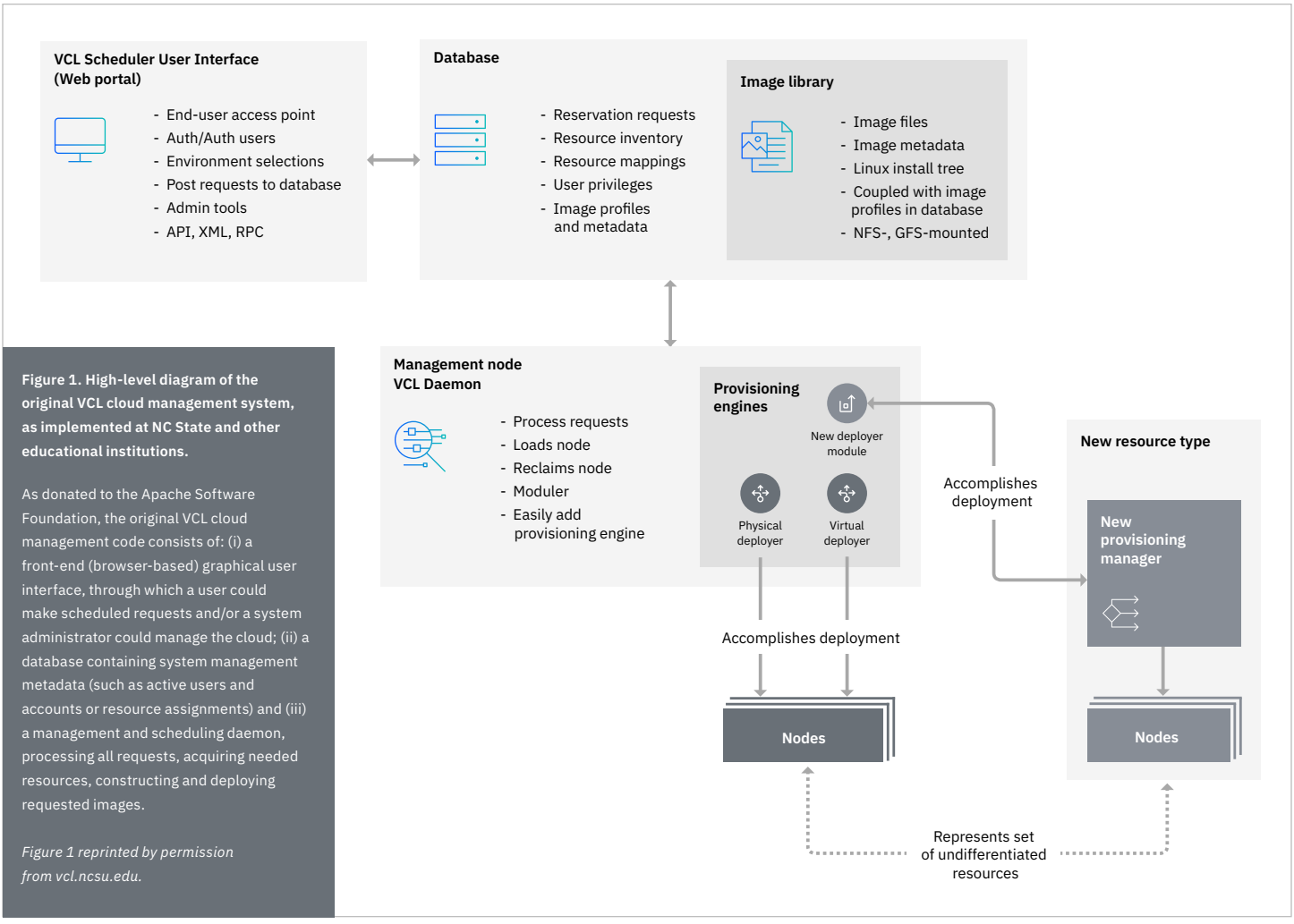

Figure 1. High-level diagram of the original VCL cloud management system, as implemented at NC State and other educational institutions.

As donated to the Apache Software Foundation, the original VCL cloud management code consists of: (i) a front-end (browser-based) graphical user interface, through which a user could make scheduled requests and/or a system administrator could manage the cloud; (ii) a database containing system management metadata (such as active users and accounts or resource assignments) and (iii) a management and scheduling daemon, processing all requests, acquiring needed resources, constructing and deploying requested images.

Figure 1 reprinted by permission from vcl.ncsu.edu.

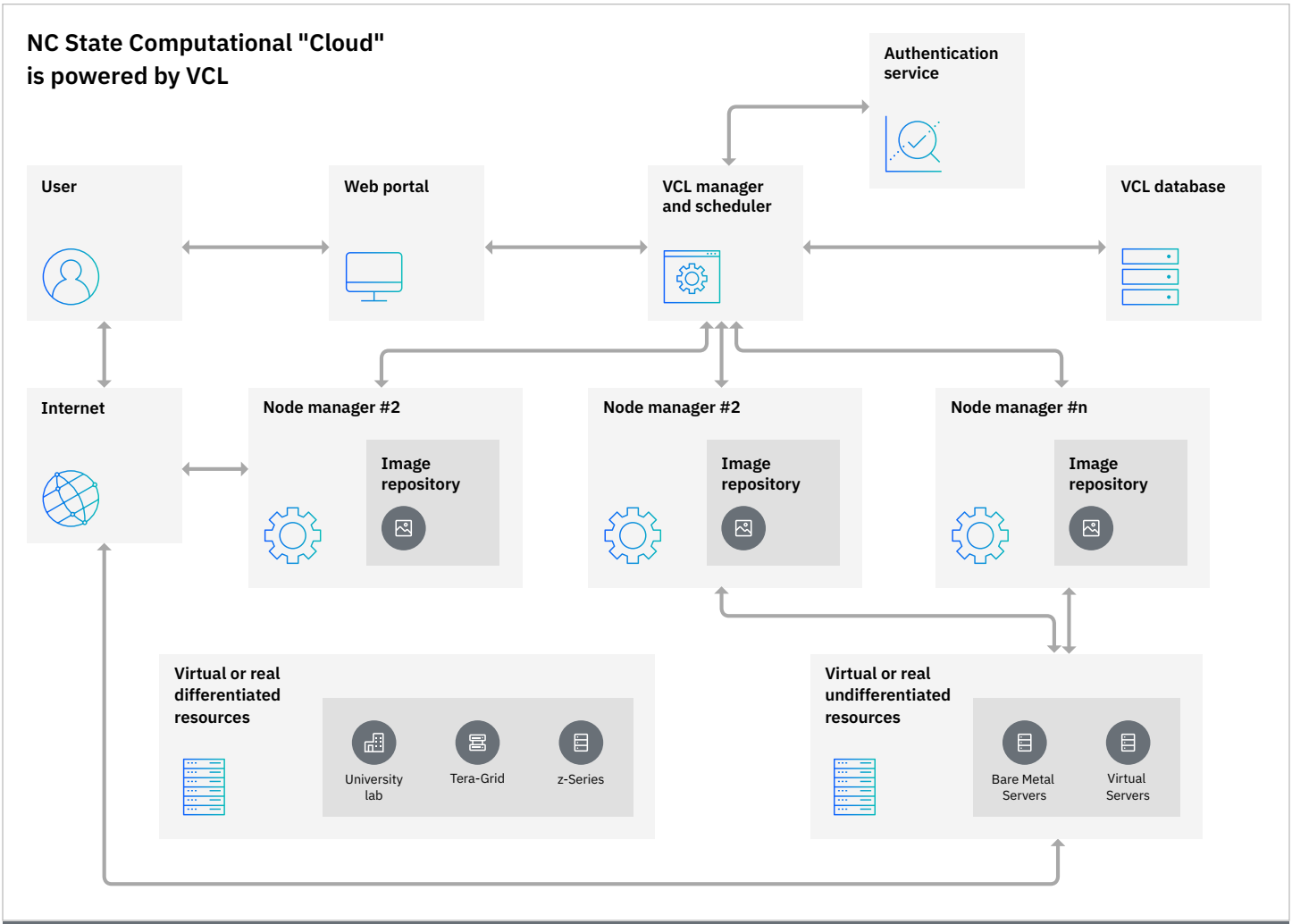

Figure 2. High-level diagram of the original VCL cloud management system.

As developed by NC State University and others, scalability and performance are maximized by defining multiple, local node managers and replicated image libraries (with deployable images) for some specified set of resources, based on number, locality, or function. Each node manager is controlled by the (high availability) central VCL manager/scheduler. Based on the request (examples: for a virtual machine {VM} using a particular hypervisor, for a dedicated server using a particular OS, or for resources in another cloud), the VCL can utilize a specific device driver that executes the deployment of the specified resources. New device drivers can be easily added as needed, including device drivers for containers/Kubernetes, IBM Cloud Paks, or others.

Figure 2 reprinted by permission from and vcl.apache.org.

Given their lower-energy consumption profile and cost-of-ownership value proposition, IBM BladeCenter servers were continually added to the growing NC State super-cloud. They were also often deployed by the many campuses who were adopting Apache VCL. And IBM integrated its various cloud solutions as they emerged, developing and donating device drivers to the Apache open source project. These new device drivers provided additional capacity and brought needed services to the academic community.

As an example, one donated device driver enabled VCL installations to request additional resources from the public IBM SmartCloud. This established one of the first examples of a true hybrid cloud, providing faculty and students access to combine university private and IBM

public resources. Such hybridization efforts led ultimately to a multi-university Internet2 NET+ sponsored initiative. This initiative sought to establish a community of schools that could request cloud-based resources from one another, including from IBM data centers, via the Internet2 high-speed backbone interconnecting universities and gigapops across the US. It has additionally provided connectivity into the high-speed academic backbones of other countries.

IBM also created VCL-deployable images of its most popular software products, that were available for free to everyone through the IBM Academic Initiative. This provided students and faculty with a variety of software assets that could be integrated into classroom instruction or research agendas. This practice is continued today, as IBM makes all of its Watson, SPSS, Hyperledger Blockchain, IoT, Security, and other software services and technologies available through the IBM Cloud for Education.

And the regular IBM Cloud Academy member meetings and workshops, notably the annual International IBM Cloud Academy Conference (ICACON), brought IBM and universities together on a regular basis, as they worked together on common technological and other challenges.

Feature and function relevance to academic needs

The popularity of Apache VCL within the academic community can be attributed to the many specialized features and functions specifically designed to address the needs of education and research (and campus business processes) that NC State incorporated into the original solution. These include:

- Block reservations, enabling a professor to schedule predefined VMs (or servers) to be ready and deployed for every student, just-in-time, every time a given class meets, throughout an entire semester or year.
- Flexible policy administration, that enables faculty or IT staff to precisely define what resources can be accessed by a professor or student, based on what classes they have enrolled in or what software/hardware resources they have purchased.
- Federated identity management, enabling faculty and students to access their cloud accounts using intranet IDs (typically email addresses) and passwords that the university has issued.
- Automatic license management, guaranteeing that a university does not exceed the number of simultaneous licenses that it has purchased for a given piece of software, as well as detailed reports of such license (and resource) usage over time, that can be used to optimize purchasing decisions.
- Faculty empowerment (given proper approval) to create application images as needed (without central IT intervention), that can then be permanently stored and for example deployed (as multiple copies) to an entire class of students.
- Through the creation of new device drivers, easy integration of any existing or newly arising services, platforms, hypervisors/containers, operating systems, or hardware (including GPUs) that might be required for teaching, research, business or other activities.
- Strong security features that have been included from the beginning, for example, two-factor sign-on authentication.
- Easy customization, enabling the access portal and other solution webpages to be personalized with the logo or more of a given institution.
- High performance features (including image pre-loading and local image library replication) that minimize image deployment and response times, typically to a few seconds or less.

Most importantly, as a result of the Apache VCL open-source community, the available specialized features and functions have been expanded over the decades so that the solution addresses a significant portion of the needs and wants of academia today. It has been designed and continuously improved by universities for universities. Furthermore, the solution has demonstrated its stability and effectiveness over the decades, as it continues to support numerous campus production clouds without major incident. It has easily scaled to thousands of servers without detrimental performance impacts through the design principles developed by NC State and others. And it naturally supports a hybrid cloud infrastructure that is consistent with IBM Cloud philosophy as espoused today.

IBM Cloud for Education Applications Lab

As related above, a large number of universities adopted an on-premise (private) Apache VCL-based cloud. Many of these are still in operation today. However, back in the early 2000s, there were some universities that wanted to adopt such a cloud paradigm, but did not have the necessary infrastructure, support personnel resources, or budget to properly deploy and maintain such a campus private cloud. Others who adopted it later transitioned to a public cloud solution for business, economic, or other reasons.

Recently, a number of these latter universities (many original members of the former IBM Cloud Academy) approached IBM seeking the specific cloud capabilities provided by this Apache VCL-based solution. However, they were interested in a hosted solution, that did not involve new investment in their on-premise infrastructure or personnel. (Many in fact outright requested Apache VCL running within an IBM data center.) But they were also interested in the Watson/AI, IoT, blockchain, and other services that the traditional IBM Cloud offers. Some were even interested in the possibility of augmenting their on-campus resources with the on-demand services and hardware that were occasionally needed, typically during times of peak demand. Such universities therefore sought a hybrid cloud paradigm (as was explored during our Internet2 NET+ project mentioned above).

To address this demand, the IBM Cloud university relations team therefore created the "IBM Cloud for Education Applications Lab" (see Figure 3). With this solution, Apache VCL is designed to manage dedicated networking, storage, and servers within various IBM data centers located around the world, specifically set aside for the academic community. Since Apache VCL is location-agnostic, it can manage a geographically distributed set of resources. This enables faculty and students to access cloud services from any IBM data center that is physically closest to their campus. It also addresses in-country or regional data residency requirements.

The team is also integrating requested services via IBM Cloud, enabling a user to seamlessly request, say, IBM Cloud Pak for Data from within the Apache VCL services menu. Additionally (as a result of supporting various university certification programs), the solution offers pre-loaded images of many academically popular IBM software applications. For example, downloadable images of IBM Security QRadar, with increased storage and security specifically designed for cybersecurity hands-on classes, are available. IBM Internet of Things and IBM Watson services are also extremely popular for university research projects involving sensor-based streaming data that then needs to be intensively analyzed in the cloud.

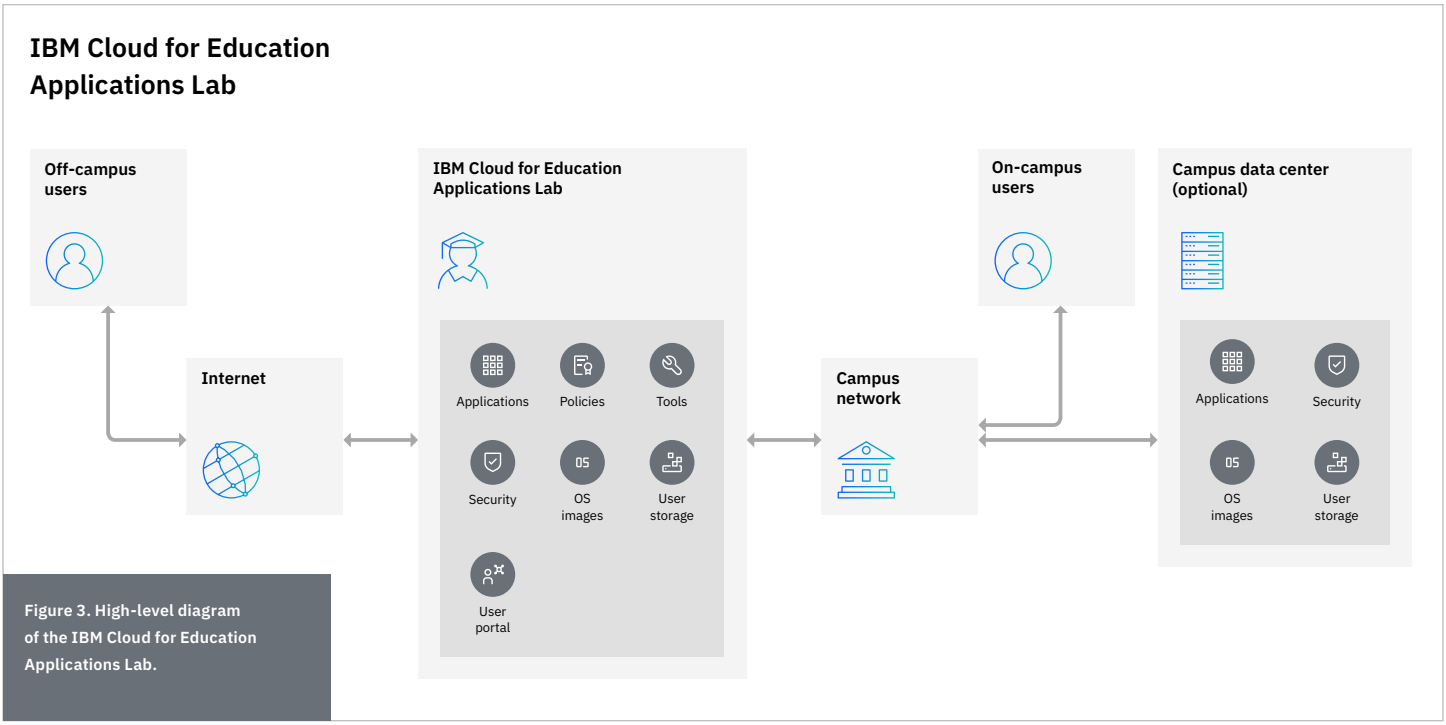

Figure 3. High-level diagram of the IBM Cloud for Education Applications Lab.

The IBM Cloud for Education Applications Lab represents a set of dedicated resources providing the desired features and functions of Apache VCL, hybridized with the Watson/AI, IoT, Hyperledger blockchain, and security services (including Cloud Paks) provided by IBM Cloud. The resources and services are available to be distributed globally across several IBM data centers for maximum performance and adherence to local data security/privacy constraints. The Applications Lab is directly accessible through any campus network, or from off-campus via the internet, using a given institution's intranet user ID and password, and via a portal/webpage that can be customized with the institution's logo or more. It is also possible to create a hybrid solution, combining existing campus resources with those of the IBM cloud. It represents one-stop-shopping for all the cloud resources and services that a campus might need.

And while it was originally developed to address the needs of universities, the IBM Cloud for Education Applications Lab can also support the educational needs of K–12 schools and community colleges. (Both continue to be supported today by the large VCL-based cloud currently operating within the state of North Carolina.)

Because the Applications Lab only requires a browser for access, teachers and students can utilize its cloud-based resources with as little as a smartphone. Therefore, a K–12 school district would only require, say, a cart of inexpensive thin clients and internet access to take advantage of the most advanced educational applications, with more than adequate storage and computational support. These capabilities are finding popularity due to the demands created by remote learning. With its minimal requirements on school infrastructure, this solution opens the door to rural and/or disadvantaged school districts. And while internet access still remains a challenge for some areas, low bandwidth and/or Wi-Fi access is not a problem. Any computationally- and storage-intensive applications and services are executed on the servers and other resources in the cloud.

One of the important added benefits provided by the IBM Cloud university relations team and the broader IBM Cloud for Education solution has been its personalized support in designing and deploying cloud-based solutions for research, and in-classroom and virtual teaching. The team has worked with faculty and students over the decades, as part of its inherent university relations

mission. It has therefore built up a wealth of experience in the design and maintenance of classroom instruction and research infrastructure support paradigms. It offers research project design clinics, targeted workshops, and education sessions for students and faculty interested in taking full advantage of the available technologies and resources.

As part of its library of reusable assets, IBM Cloud for Education solution offers research design templates, defining the required cloud-based resources and workflows that have been deployed in support of many typical research projects. A popular template defines the required resources (computational, storage, networking) and applications and services (Watson text and video analysis, machine learning and neural networks, IoT, and more), used in analyzing streaming data from medical, environmental, control system, and other types of sensors. It addresses the typical workflows required, from edge computing and preprocessing of the massive amounts of real-time streaming data, to the AI-driven cloud-based analyses of the reduced data, ending with the plotting and presentation of the results, These reusable assets, together with personalized education, are available to all schools using the IBM Cloud for Education.

As importantly, IBM will be re-launching the updated version of the earlier “IBM Cloud Academy” partnership program, renamed “The IBM Cloud for Education Academy.”

IBM Cloud for Education Academy

When it first arrived, cloud computing offered educational institutions substantial benefits. These included economic savings, due to much lower and more predictable total costs. It also provided ubiquitous access to the most advanced technologies and services in support of research and teaching. These advances were especially beneficial to economically strapped institutions. Cloud computing additionally increased safety, by eliminating the need for students to travel to physical server labs at night. And cloud computing empowered faculty to create their own cloud-based assets, eliminating delays due to IT bottlenecks in creating them.

Universities have always been among the first users to test new technologies and many universities immediately embraced the emerging cloud paradigm. To support these early adopters (as indicated above), IBM launched what is now the "IBM Cloud for Education Academy," with the goal of creating an ecosystem of university partners, working together to:

- Facilitate the adoption of cloud computing by academic institutions — for example, by removing barriers to entry by developing and sharing best practices for campus-wide rollouts
- Support collaborative projects to define and develop new cloud features/functions and technological/architectural improvements, for better scalability, performance, security, and requirements coverage — with one important vehicle being the Apache VCL open-source project.
- Develop reusable assets that could be shared across the community, including an image repository of popular open source and IBM software (the latter made available through the IBM Academic Initiative), as well as unused cloud cycles and resources that could be "borrowed" by member universities (a goal of the community's Internet2 NET+ project).
- Pursue joint funding opportunities from local, state, and federal grant programs to support basic and applied research into cloud technology and delivered applications.
- Establish proofs-of-concept around target communities (for example, to showcase cloud-based, K–12 STEM outreach programs for inner-city schools) and to improve existing infrastructure.
- Sponsor workshops, educational events and academic conferences (such as ICACON — see Figure 4) to share information on joint activities and progress, new research and technological advances, opportunities for resources and funding, and other activities — all to maintain a cooperative partnership working together to improve education and research through cloud computing technology.

Figure 4. Announcement for the International IBM Cloud Academy Conference (ICACON)

The International IBM Cloud Academy Conference (ICACON) was established by the IBM Cloud Academy to "provide an exciting opportunity for educators and researchers to share ideas and experiences in cloud computing." Open to both K–12 and higher ed, ICACON has featured experience reports and research papers on all aspects of academic loud computing. Selected papers were published in special issues of select refereed academic journals. The first ICACON was held in Research Triangle Park NC in 2012, with subsequent conferences hosted in member university cities around the world.

Given the benefits derived from the original “IBM Cloud Academy” discussed above, IBM has updated this university partner’s program with the “IBM Cloud for Education Academy.” The IBM Cloud university relations team is supporting an ecosystem of participating universities, fostering activities involving the integration of cloud computing (and the services it delivers) into curricula and research. It has been demonstrating new ways in which the technologies can be applied, while assisting in the development of new techniques, paradigms, and technological advancements. To this end, the program continues to provide the needed technical and other support of the various collaborative activities listed above. This includes hosting education workshops and classes, hackathons, project design sessions, and more.

One program that IBM is seeking to establish at participating universities is the “University Delivery Services” (UDS) program. First launched in North Carolina with IBM Global Business Services, the program works with universities to establish educational programs around key technical skills required by prospective IBM interns. These interns might then be recruited to support IBM, their university, or other services engagements.

UDS works with universities to define appropriate ways to integrate the skills training into the university curricula or activities. It has historically relied on special topics and special project courses that exist at the undergraduate and graduate levels at all schools. As such, they do not need to displace existing courses. UDS has also utilized special workshops or other educational events — all to produce an on-campus cadre of students possessing the skills required to integrate the latest technologies into all types of businesses.

The program also seeks to develop a participating ecosystem of IBM customers and business partners, who can take advantage of the services program and seek to recruit the involved students that might accompany the project deliverables. The hope is that they will also bring additional investment and participation in all IBM Cloud for Education Academy activities.

Getting started

The IBM Cloud for Education Applications Lab is currently being used by private and state/provincial universities in the US, Canada, and South America. It is available as a subscription service, for which interested university representatives may obtain free trial accounts and references upon request.

The IBM Cloud for Education university relations team regularly works with universities across the world, helping faculty define the cloud-based resources and services that they need for teaching and research purposes. They provide workshops, hackathons, targeted education sessions, and other events and services intended to assist universities in making the most effective use of IBM Cloud for Education delivered assets. The university relations team has significant experience in working with universities to roll out new academic degree programs and concentrations — especially in areas around Big Data, IoT, and quantum computing. They also sponsor project design sessions intended to help researchers optimally define the cloud-based research platforms they need. The university relations team also works directly with IT staff to help develop the integration plans (including hybridization with on-campus resources), deployable images (including VMs and containers), specialized device drivers (for example, bursting out to other public clouds), and other required cloud-based specialized features and functions.

To get started, visit our catalog page.

IBM Hybrid Cloud
IBM Corporation
New Orchard Road
Armonk, NY 10504

Produced in the United States of America
November 2021