Rearranging Harvard.edu for the Future
Cloud and DevOps Program Revitalizes HPAC Sites to Optimize for Best Practices

When the Cloud & DevOps program at Harvard University Information Technology (HUIT) launched in January 2015, it brought on a support portfolio of some of Harvard’s most prominent websites — including www.harvard.edu and a number of other cloud-based sites maintained by Harvard Public Affairs and Communications (HPAC). These sites were first deployed to Amazon Web Services (AWS) in July 2014 by a special-projects team from HUIT Infrastructure that successfully migrated from Rackspace while also re-platforming for the scalability, resiliency, and disaster recovery offered by AWS. In the meantime, a number of new or improved AWS features had made parts of the HPAC infrastructure outdated or in need of optimization. While the HPAC sites had enjoyed many cloud benefits — including automated response to capacity demands — managing code deployment in particular was problematic. With this in mind, the Cloud & DevOps program partnered with HPAC in June 2015 to build a more stable cloud deployment framework in tandem with a redesign of www.harvard.edu (dubbed “HPAC 2.0”).

The Opportunity
After assuming support for HPAC websites (“HPAC 1.0”), the Cloud & DevOps team worked together to produce a holistic plan for improving the portfolio. First, they examined service tickets opened by HPAC in the preceding year and found that the majority of issues were in infrastructure (S3 bucket configuration, AMI build/spinup, logging), code deployment (orchestration, code sync), or Jenkins job workflow (S3/database sync, Drupal module configuration). The team then conducted a deep dive into the existing HPAC environments for insight into these categories and to identify opportunities for architecture improvement; in doing so, they found that most issues could be resolved by re-architecting the code deployment process. The team also found that a number of new AWS features, as well as matured best practices, could be used to provide added security and stability to the deployed cloud infrastructure. What’s more, the timing couldn’t be better; the scheduled www.harvard.edu relaunch provided a prime opportunity for the team to implement their holistic plan and pave the way for improvements to other HPAC sites.

The Solution
Work began in early June 2015 to build new staging and production environments for HPAC 2.0, including (but not limited to) these new AWS features and HUIT cloud best practices:

- **AWS Account:** Direct billing and conformance with HUIT account standards
- **AWS VPC:** Implementation of Direct Connect, designation of Harvard-recognized IP ranges, and VPN lockdown for STG
- **AWS EC2:** AMI build process re-architected to contain most application components on launch, resulting in quicker deployment to auto-scaling groups
- **AWS CodeDeploy/Lambda:** Custom code deployment scripts that polled S3 for changes were replaced with an AWS automated process; S3 layout was streamlined and sharing of S3 between environments was eliminated
- **AWS CDN:** Process configuring CloudFront to restrict access to the Drupal admin module was replaced with a more secure Apache solution; a DDoS service was placed in front of CloudFront to mitigate denial-of-service attacks
- **AWS Memcache:** Improvements made to prevent clashes between Drupal’s internal cache and AWS memcache
- **Workflow:** Jenkins jobs used to sync code between environments were rewritten to reflect infrastructure changes; added new jobs to provide status of site URLs as well as cache clearing and CloudFront invalidation

What’s Next?
With a planned go-live Aug. 21, 2015, HPAC and the Cloud & DevOps program have built a comprehensive transition plan using traditional release management principles modified for the cloud. The plan is being executed using a master release schedule supported by a validated cloud readiness checklist, detailed deployment and decommissioning plans, and a comprehensive operational runbook. Issues are being managed jointly in Basecamp, with release tracking using a standard change ticket in ServiceNow. By Aug. 21, we look forward to reporting an issue-free release, as well as a reduction in support requests over time. In addition, the infrastructure implementation handbook created for HPAC 2.0 will serve as a valuable model to build the 2.0 versions of the rest of HPAC’s websites — as well as provide a means for HPAC to manage these migrations on its own.

Find other case studies and learn more about the program at [cloud.huit.harvard.edu](http://cloud.huit.harvard.edu) August 2015