Federal Virtual Training Environment (FedVTE)

A decade of continued cost-saving benefits

Marie Baker
April Galyardt
Dominic Ross

May 2017

Cyber Workforce Development

http://www.sei.cmu.edu

This material is based upon work funded and supported by the Department of Homeland Security under Contract No. FA8702-15-D-0002 with Carnegie Mellon University for the operation of the Software Engineering Institute, a federally funded research and development center sponsored by the United States Department of Defense.

DM17-0320

A Decade of Federal Virtual Training Environment

The Federal Virtual Training Environment (FedVTE) is an online, on-demand training system containing cybersecurity and certification prep courses, at no cost to federal, state, and local government employees. The goal of FedVTE is to help support the need to cultivate a skilled cyber workforce that is prepared to protect and defend in cyberspace. While the platform has gone through several evolutions and transitions, a virtual training platform has been available to the federal cyber workforce for over ten years.

FedVTE has been a valuable asset to the federal government by enabling agencies to realize savings in training costs, as well as by utilizing videos specifically purposed for bandwidth and storage efficiency. This report provides an overview of the history of FedVTE, examines a year of FedVTE user activity, followed by detailing the unique video compression technique and its cost-effectiveness. The report concludes by looking at potential strategies the FedVTE program could consider to further service cyber workforce knowledge and skill building.

FedVTE background history and evolution

The U.S. Department of Defense *(DoD) Directive 8570.01 Information Assurance Training, Certification, and Workforce Management,* was one of the first efforts to combat the lack of skilled operators to defend and protect in cyberspace. It was designed to provide the basis for an enterprise-wide solution to train, qualify, and manage the DoD Information Assurance (IA) workforce. The directive called for immediate action for training all employees with access to a DoD system, at a baseline requirement.

The manual, 8570.01-M, last updated in 2012, provides guidance for adhering to the directive. This includes categorizing all IA positions within four categories, and assigning a level or specialty within each category. All individuals holding an IA position must obtain an approved industry certification from within their category, to be deemed qualified for their position.

To obtain an industry certification, each individual mandated by this directive must pass a certification exam. As such, training was needed to prepare for exams. Even the most experienced technical operator would need a refresher on concepts that are covered in the broad scope of certification exam objectives. While agencies were required to adhere to the 8570 directive, it wasn't funded. Each agency would need to budget for training and exam costs for each employee, and allot time away from operations to attend training. Because of how financing would impact individual agencies, and the inability for operators to attend training, it was realized that an enterprise-wide training solution would be most efficient for the enterprise-wide directive. DoD's Virtual Training Environment (VTE) was spawned.

The Software Engineering Institute (SEI), a Federally Funded Research and Development Center (FFRDC) at Carnegie Mellon University, developed and originally hosted VTE from 2007-2012. During this time, the Defense Information Systems Agency (DISA) funded both content development and platform support. When VTE was ready and required to be transitioned away from the SEI in 2012 the benefits the Federal government as a whole, and not just the DoD, were receiving from this training program

were a major consideration. The Department of Homeland Security (DHS), agreed to take on hosting and managing the platform, rebranding it the Federal Virtual Training Environment (FedVTE). Upon transition away from the SEI, the platform was originally hosted by the State Department for DHS, and since May 2015, done so by USA Learning.

Brief historical usage numbers during evolution and transition period

SEI VTE, 2008

- 16,730 active learners
- 120,000 hours of training delivered

State Dept. FedVTE, Sep 2012 – Nov 2014

- Over 153,000 trained
- 62,380 Course completions

The current instantiation of FedVTE serves over 160,000 registered users, and hosts over 60 cybersecurity-related courses. While the 8570 directive was officially replaced by the DoD 8140 *Cyberspace Workforce Management Policy Update* in August 2015, the manual for implementing the new directive has not been released. Until the 8140 manual is finalized, the 8140 directive will assume the 8570 manual, 8570.01-M.

The 8140 directive is expected to be more versatile and comprehensive. It is driven by the NICE Cybersecurity Workforce Framework, which is an initiative to establish consistency when defining job roles within cyber operations through 32 cyber specialty areas categorized into seven high-level areas of operations. Each of the 32 specialty areas are further detailed through related tasks, knowledge, skills, and abilities (KSAs) specific to that job role. The 8140 offers more granularity for categorizing the DoD cyber workforce, enables a clearer picture of workforce strengths and gap areas, and helps guide training plans to specific task goals.

The NICE Cybersecurity Workforce Framework has been adopted and leveraged by several organizations and agencies for identifying their own workforce, and training needs. For instance, the DoD Cyberspace Workforce Framework (DCWF) uses the structure and guidance of the NICE framework, and incorporates mission essential tasks and KSAs from the Joint Cyberspace Training and Certification Standards (JCT&CS) to identify a gamut of DoD cyber workforce roles. The Office of Personnel Management (OPM) assigned codes corresponding to each DCWF category, specialty area, and task. The codes are used to uniformly catalog the workforce within areas of work and specialty, within training and certification tracking databases.

FedVTE continues to support the 8570 directive through its evolution to the 8140. The courses within FedVTE are organized according to the NICE Framework categories and specialty areas. This helps guide users and employers to content based on operational needs.

Besides offering courses to help prepare individuals to sit for a certification exam, courses can be completed to fulfill Continuing Education (CE) requirements to maintain a certification. For instance,

CompTIA, a certification agency with four certifications listed in the 8570, maintains a list of FedVTE courses that are pre-approved for CE credits and can be applied to keep one of their certifications in good standing. The certification prep courses are only about a third of the courses available in FedVTE. The remaining content covers a broad spectrum of cybersecurity-related topics for novice through advanced users.

Completion certificates for each course are available, allowing users to provide record of training credit hours to certification agencies, or employers. The Army Training and Certification Tracking System (ATCTS) is just one example of an agency that leverages FedVTE completion reports to update user training records.

Data is not available to positively assert the primary reason users access FedVTE training. It's unknown whether it's required by directive or employer, or driven by personal development, for instance. However, analyzing course progress reports can provide some insight.

FedVTE course progress and completion data from the 2016 calendar year was analyzed to determine if it could reveal trends and needs of the cyber workforce utilizing the system. The data included any registered user who used FedVTE within the year, accessing any course. A high-level summary of the system and data:

- 139,355 registered users by the end of 2016
- 67,704 new user registrations within the 2016 calendar year
- 61,825 users recorded activity within the 2016 calendar year data
- 72,468 course completions earned in 2016
- 72,468 course completions equate to over 1 million hours of training

Users by department

There were 62,484 distinct accounts active in 2016, but only 61,825 distinct email addresses. The approximately 750 users (1.1%) who may have multiple accounts, most likely transferred departments during the year, causing their account duplication in the data. For this analysis, we will focus on the activity of distinct user accounts.

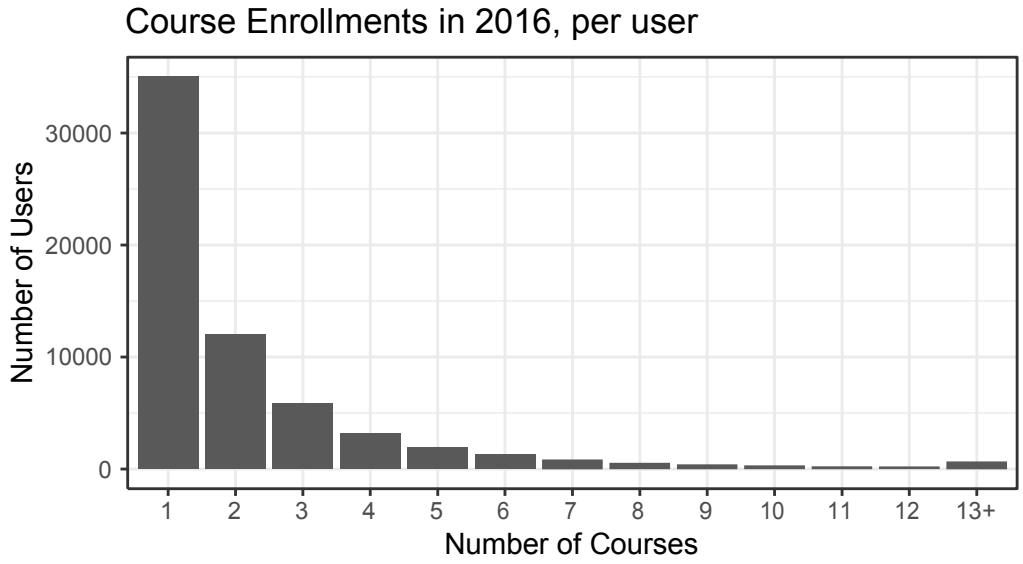
Users come from 62 different departments. Department options for users to select from when registering an account in FedVTE, are ordered here by number of active users in 2016:

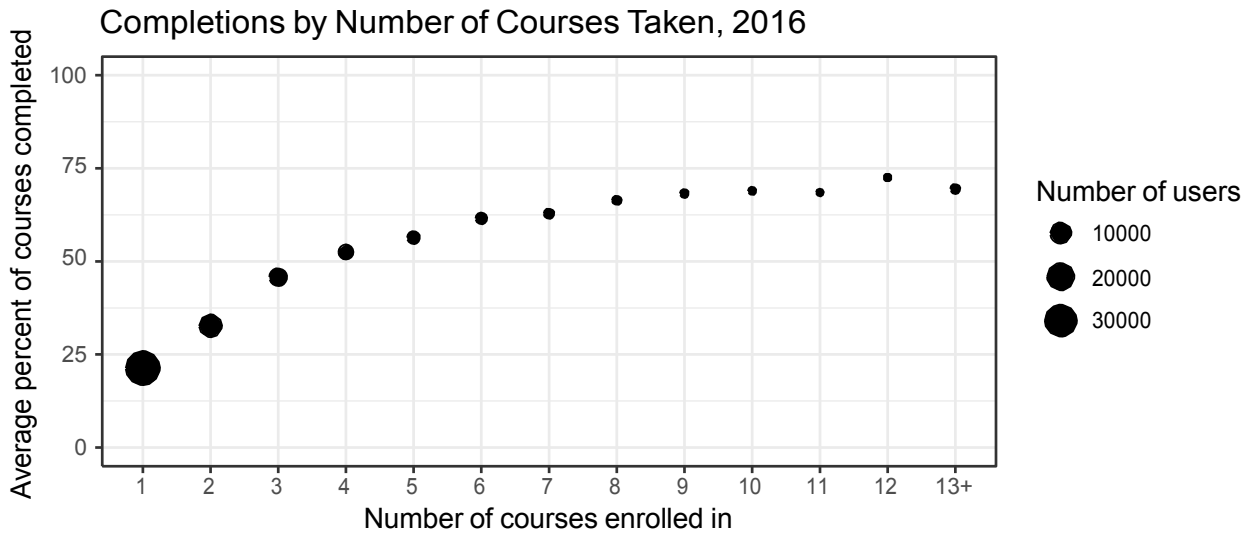
Department	Number of Users
Army	11,031
Air Force	10,160
Dept of Defense	10,155
VETERAN (not Govt. Employee or Contractor)	5,526
Navy (Military)	4,490
Dept of the Navy (Civilian)	2,990
Dept of Homeland Security	2,629
Marine Corps	2,465
State & Local Governments	2,055
Navy	1,638

Dept of the Navy	1,432
Other Federal Depts, Agencies, Commissions, Boards, or Councils	1,048
None Recorded	784
Dept of Veterans Affairs	628
Dept of Treasury	546
Dept of Justice	474
Dept of Transportation	439
Social Security Administration	434
Dept of State	414
Dept of Health and Human Services	357
Dept of Energy	286
Tennessee Valley Authority	282
Dept of Commerce	251
US Special Operations Command	169
Federal Small Independent Agencies	167
US Postal Service	156
Dept of Agriculture	140
National Aeronautics & Space Administration	131
Dept of Labor	124
Dept of Education	109
Federal Reserve System	106
Smithsonian Institution	101
US Courts	95
Dept of the Interior	89
General Services Administration	88
Central Intelligence Agency	63
Federal Commissions, Boards, Councils	61
Environmental Protection Agency	43
Peace Corps	34
Veterans Administration	32
Armed Forces Retirement Home	30
Tribal Governments	29
Dept of Housing and Urban Development	24
National Archives and Records Administration	20
Library of Congress	18
Pension Benefit Guaranty Corporation	17
Office of Management & Budget	15
Massachusetts Institute of Technology	14
National Credit Union Administration	14
Consumer Financial Protection Bureau (CFPB)	13
National Science Foundation	13
US House of Representatives	12
US Senate	12
Small Business Administration	11
Territorial Governments	8
Government Printing Office	3
Air Force Intelligence Agency	2
Massachusetts Institute of Technology, Lincoln Labs	2
Pretrial Services Agency (PSA)	2
Advisory Council on Historic Preservation	1
National Academy of Sciences	1
Other Agencies or Commissions	1

For further analysis, departments with less than 100 users along with "None Recorded" will be combined into the "Other Federal Departments, Agencies, Commissions, Boards, or Councils" group.

Course enrollments and completions per user





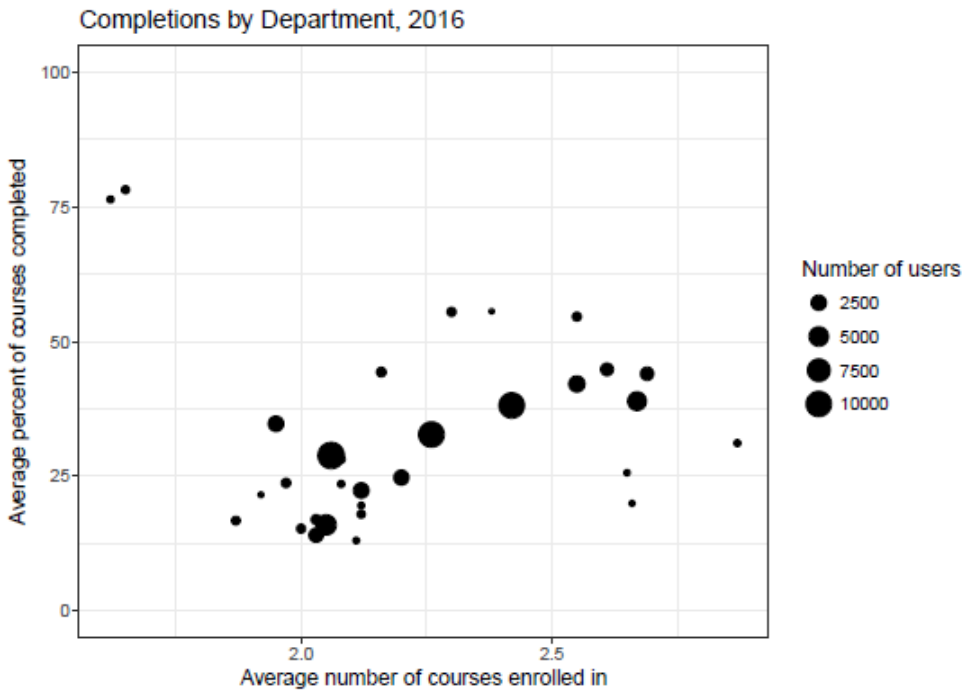
- Of the 61,825 users, 56% were enrolled in one course, 19.2% were enrolled in two courses, while only 2.3% were enrolled in 10 or more.
- Of the 35,012 people who enrolled in a single course during 2016, 78.7% did not complete that course.
- Of the 27,468 people who enrolled in more than one course during 2016, 63.1% completed at least one course.
- In general, we see that users who enrolled in more courses also had a higher rate of completion

Course enrollments and completions aggregated by department (And email group)

Department	Number of Users	Avg courses	Avg completions
Army	11,030	2.06	0.82
Air Force	10,160	2.26	1.08
Dept of Defense	10,155	2.42	1.22
VETERAN (not Govt. Employee or Co	5,526	2.05	0.60
Navy (Military)	4,490	2.67	1.50
Dept of the Navy (Civilian)	2,990	2.55	1.44
Dept of Homeland Security	2,626	2.12	0.76
Other Federal Depts, Agencies, Co	2,599	1.95	0.80
Marine Corps	2,465	2.20	0.84
State & Local Governments	2,055	2.03	0.43
Navy	1,638	2.69	1.50
Dept of the Navy	1,432	2.61	1.49
Dept of Veterans Affairs	628	2.03	0.65
Dept of Treasury	546	2.16	1.21
Dept of Justice	474	1.97	0.71
Dept of Transportation	439	2.30	1.38
Social Security Administration	434	2.55	1.64
Dept of State	414	2.00	0.53
Dept of Health and Human Services	357	1.87	0.46
Dept of Energy	286	2.12	0.65
Tennessee Valley Authority	282	1.65	1.15
Dept of Commerce	251	2.08	0.73
US Special Operations Command	169	2.08	0.80
Federal Small Independent Agencie	167	1.62	1.09
US Postal Service	156	2.87	1.51
Dept of Agriculture	140	2.12	0.54
National Aeronautics & Space Admi	131	2.11	0.53
Dept of Labor	124	2.65	0.87
Dept of Education	109	2.66	0.90
Federal Reserve System	106	1.92	0.58
Smithsonian Institution	101	2.38	1.52

Same thing aggregated by email address.

Department	Number of Users	Avg courses per user	Avg completions per user
mil	40,882	2.34	1.17
gov	8,602	2.09	0.79
other	12,996	2.10	0.67



For the majority of departments, the same pattern holds at the group level that was observed at the individual level. Individuals who enroll in more courses tend to have a higher rate of completion.

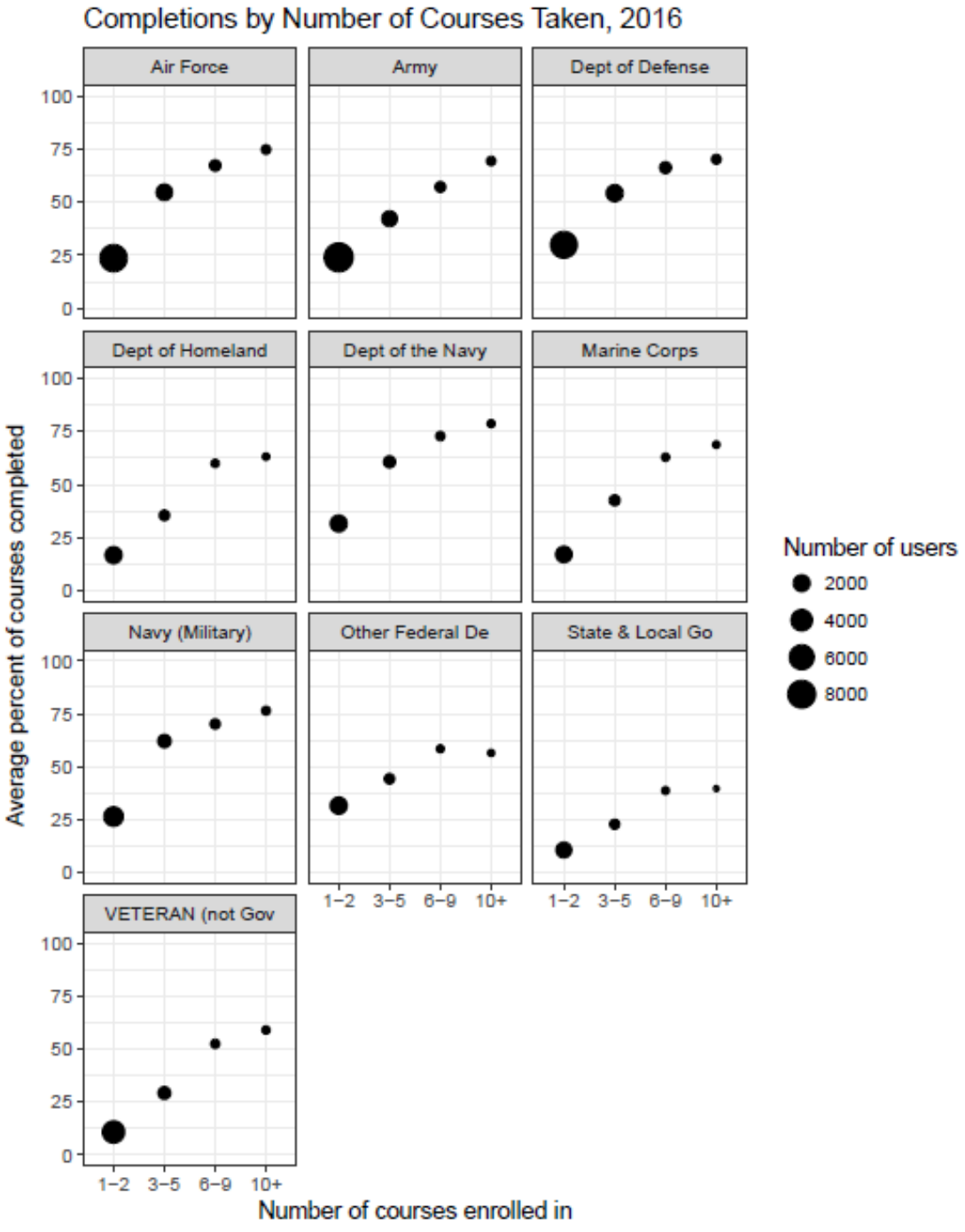
However, there are 2 departments where the average number of course enrollments is less than 2 per person, but the average completion rate is higher than 75%. *State Agency* is used to anonymize actual agency name.

Department	Number of Users	Avg courses	Avg completions	Avg perc. complete
State Agency	282	1.65	1.15	78.2
Federal Small Independent Agencies	167	1.62	1.09	76.4

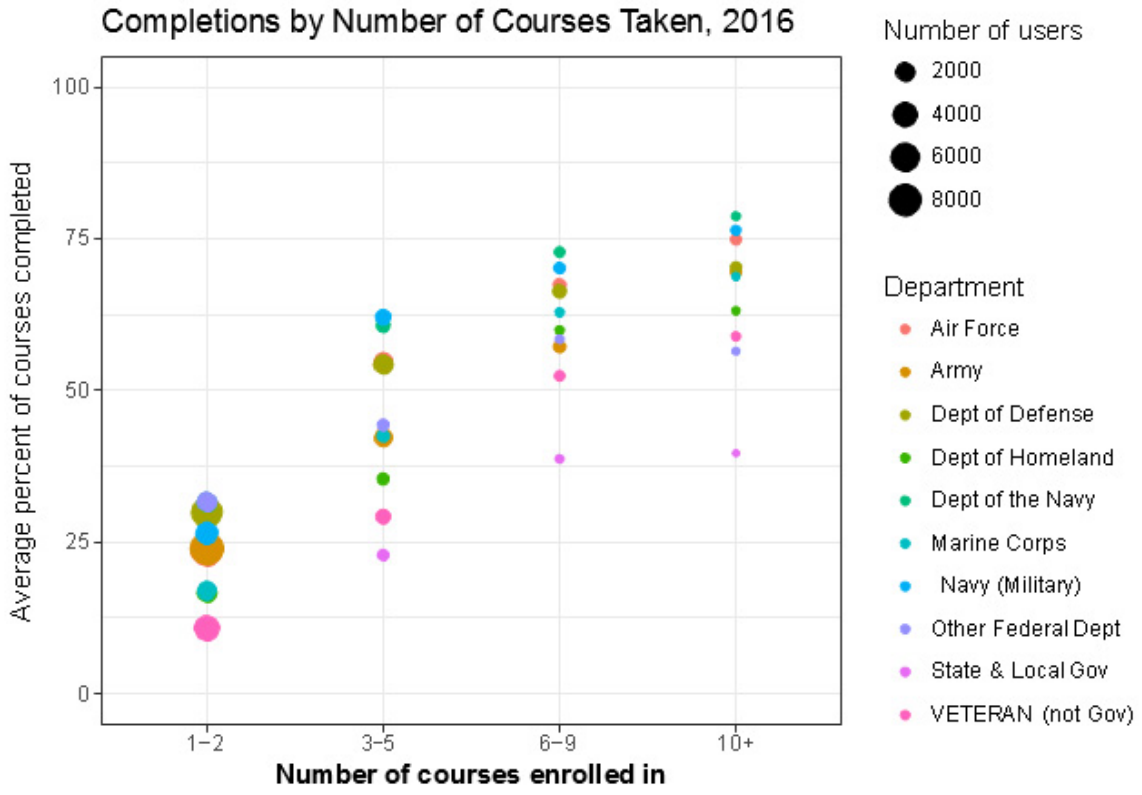
This high completion rate is largely being driven by one course for each department. Many users in these 2 departments took one course and completed it (so for those users the % of courses completed is 100%). It is reasonable to suspect that these agencies require their employees to take this specific course.

Department	Course	Enrolled	Completed	% Complete
Federal Small Independent Agencies	Cyber Security Overview for Managers	126	125	99.2
State Agency	Cloud Computing Security	105	97	92.4

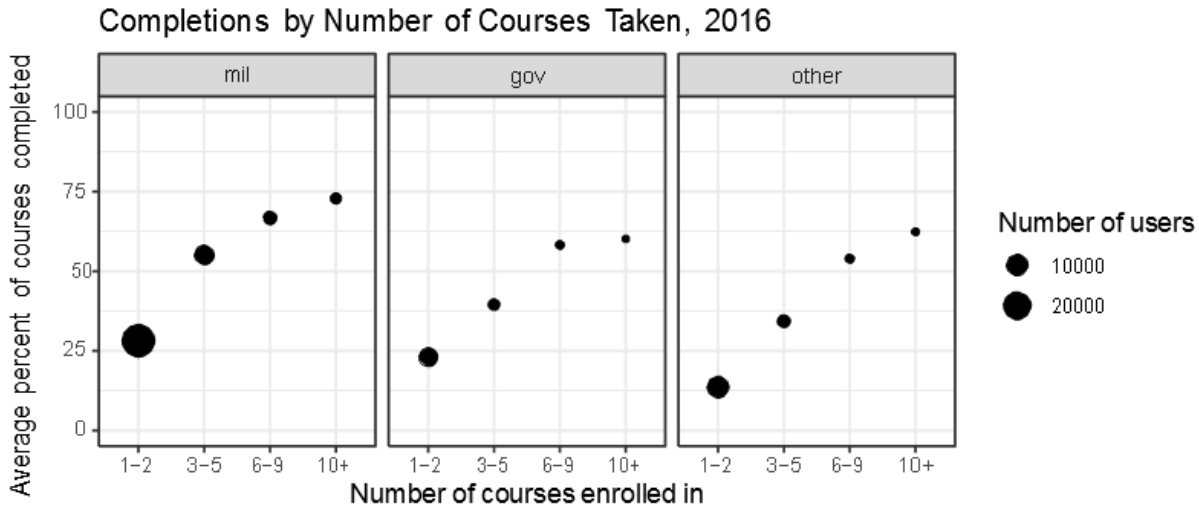
For the 10 largest departments (with over 2000 users each), following are patterns between departments. These 10 departments account for 86.6% of the active users in 2016.



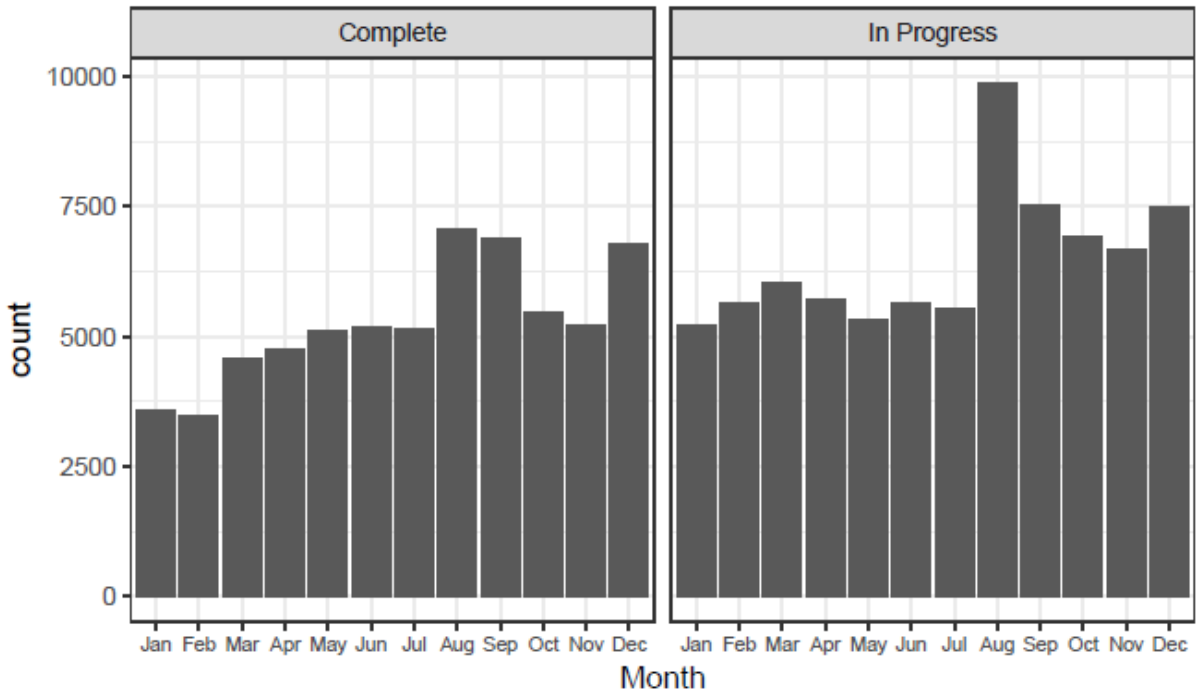
An alternate version of the previous graph:



Completions by number of courses taken, by email group



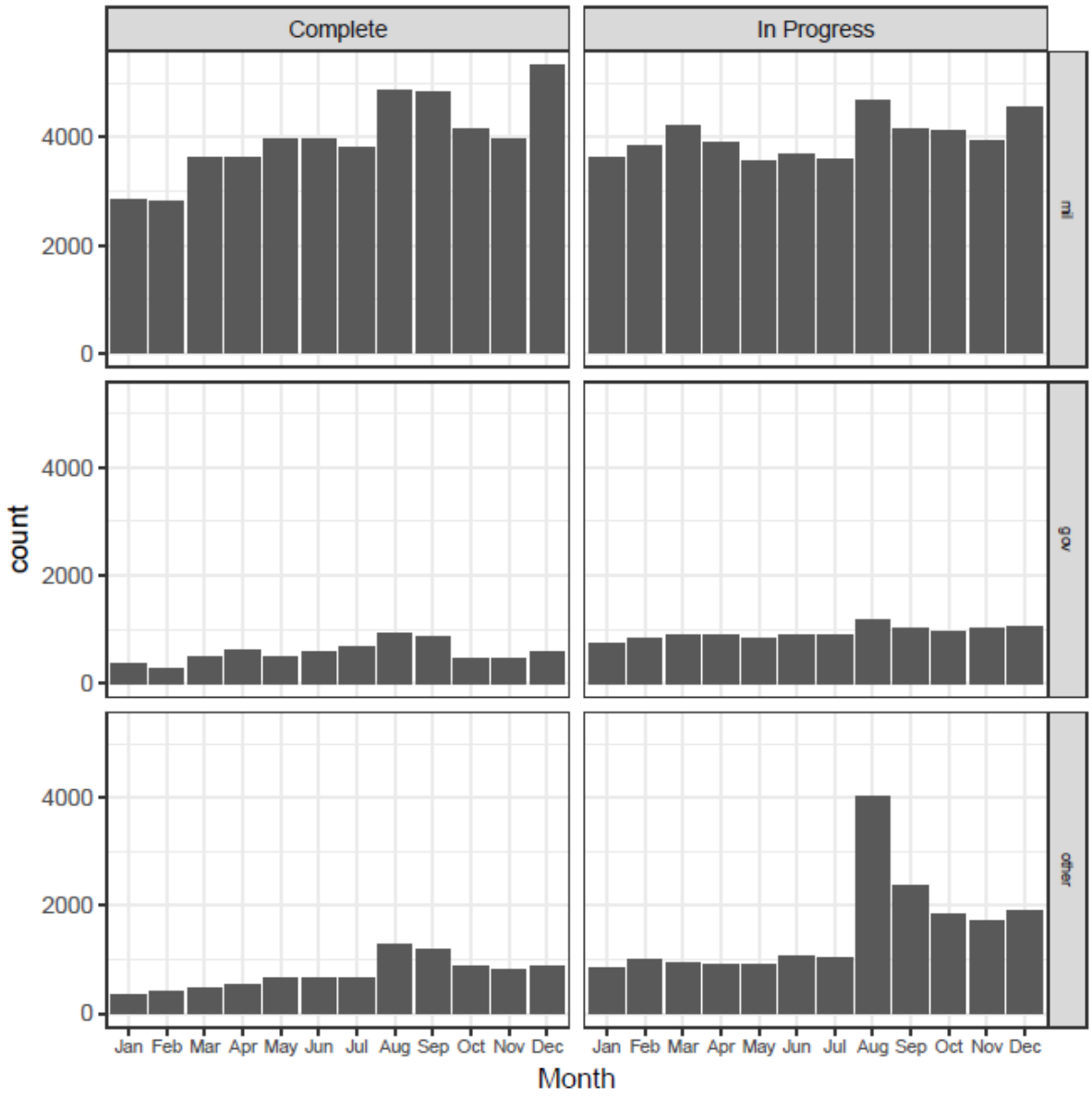
Courses completed overtime



- The data shows peak activity in August/September,
- Smaller peaks in completions in December and March.
- The August peak in completions is accompanied by an even larger peak in beginning courses.

From this data, it seems there may be training requirements or deadlines in March and December, the first and fourth quarter of the calendar year. The August peak is likely driven by an outreach initiative for the Hire our Heroes program. The surge of activity is in line with FedVTE records showing the average 5,000 new users per month, spiked to over 12,000 in August. New users who registered through the Hire our Heroes program accounted for over 7,000 of those.

When looking at completions and initiations over time, it is clear the August spike in enrollments was largely driven by veteran users. However, looking back a few graphs at completions by department, the veterans who only enrolled in one course had one of the lowest completion rates of all. This would indicate that there's a sizable population who were interested in the professional development resources, but didn't necessarily want to complete an entire course.



This graph shows the monthly course completion and progress data for each month during 2016, categorized by email address.

Popular courses

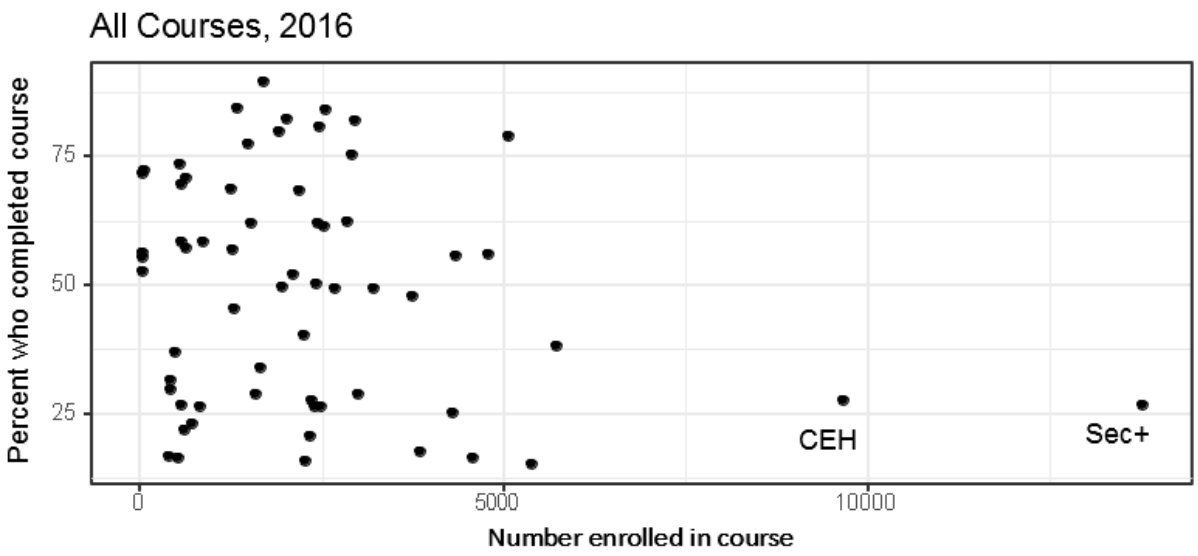
There are 62 courses in the 2016 data. The 25 most popular:

Course Name	Number of Users	Percent Completions
CompTIA Security SY0-XXX Prep	13,765	26.9
Certified Ethical Hacker vX	9,637	27.6
Advanced PCAP Analysis and Signature Development APA	5,708	38.1
ISC2 TM CISSP R Prep 2015	5,368	15.2
Cloud Computing Security	5,054	79.0
Windows Operating System Security	4,780	55.9
CompTIA A Prep 2013 220-801	4,563	16.5
CDM Module 1 : Overview	4,328	55.9
CompTIA Advanced Security Practitioner CASP CAS-002	4,294	25.2
CompTIA Network N10-005 Certification Prep	3,841	17.6
Linux Operating System Security	3,727	47.9
Wireless Network Security	3,191	49.4
Introduction to Windows Scripting	2,985	29.0
Securing Infrastructure Devices	2,931	82.2
Security and DNS	2,898	75.3
Cyber Risk Management for Technicians	2,838	62.3
DoD IA Boot Camp	2,673	49.4
CDM Module 2: Hardware Asset Management	2,538	84.3
Mobile and Device Security 2015	2,511	61.6
Advanced Windows Scripting	2,485	26.6
Securing the Network Perimeter	2,456	81.0
Cyber Security Overview for Managers	2,438	62.1
Cyber Risk Management for Managers	2,414	50.4
Certified Information Security Manager 2013 Self-Study Course	2,390	26.4
Penetration Testing	2,344	27.6

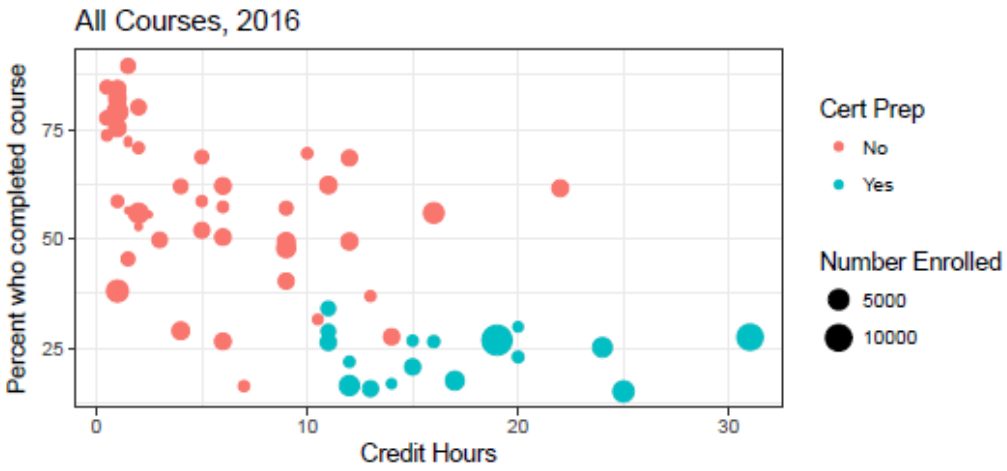
Additional Information about these 25 courses:

Four of these courses are also modules within a larger, 12-hour course. Meaning, if a user completes that entire course, they also receive completion credit for these individual modules as a standalone course; and vice versa with partial credit in the larger course for a module completed as a standalone course. The larger course is Emerging Cyber Security Threats course. The four courses above that are modules within the 12-hour course are: Cloud Computing Security, Securing Infrastructure Devices, Security and DNS, and Securing the Network Perimeter, each being one credit hour.

Seven of the most popular 25 are certification prep courses. The CompTIA Security+, which was updated from version 301 to 401 in 2016, and the Certified Ethical Hacker that was updated from version 7 to 8, have more than twice the enrolled users than any other course. However, completion of these 20+ hour long courses is relatively low, indicating they were likely used as study resources with users focusing on only topics they needed reinforced.



This graph is a good visual representation of the popularity of the two certification prep courses, Sec+ and CEH, while the percentage completing the entire courses is low.



This graph compares the training hour length between certification prep courses and non-certification prep courses, and a comparison of the percentage of users who completed either. Shorter courses that aren't certification prep, have a much higher completion percentage. This is consistent with the hypothesis that many users enrolling in the certification prep courses are using it to review or study, but not needing or required to, complete the entire course.

Highest enrollment courses for 10 biggest departments

Course Name	Department	Number of Users	Percent Completions
CompTIA Security+ (SY0-XX	Air Force	2,827	22.7
Certified Ethical Hacker	Air Force	1,696	29.4
(ISC)2 (TM) CISSP (R) Pre	Air Force	926	15.8
CompTIA Security+ (SY0-XX	Army	2,908	32.7
Certified Ethical Hacker	Army	1,646	24.2
Advanced PCAP Analysis an	Army	1,022	34.8
CompTIA Security+ (SY0-XX	Dept of Defense	2,266	30.6
Certified Ethical Hacker	Dept of Defense	1,834	32.2
Windows Operating System	Dept of Defense	1,207	63.5
Certified Ethical Hacker	Dept of Homeland Security	399	14.0
CompTIA Security+ (SY0-XX	Dept of Homeland Security	366	14.5
CompTIA A+ Prep 2013 220-	Dept of Homeland Security	315	13.7
CompTIA Security+ (SY0-XX	Dept of the Navy (Civilian)	732	32.1
Certified Ethical Hacker	Dept of the Navy (Civilian)	462	40.0
Windows Operating System	Dept of the Navy (Civilian)	374	62.0
CompTIA Security+ (SY0-XX	Marine Corps	715	23.5
Certified Ethical Hacker	Marine Corps	336	21.7
CompTIA Advanced Security	Marine Corps	247	22.7
CompTIA Security+ (SY0-XX	Navy (Military)	1,323	35.2
Certified Ethical Hacker	Navy (Military)	826	40.2
Advanced PCAP Analysis an	Navy (Military)	447	52.8
CompTIA Security+ (SY0-XX	Other Federal Depts, Agenc	393	22.9
Certified Ethical Hacker	Other Federal Depts, Agen	311	22.8
(ISC)2 (TM) CISSP (R) Pre	Other Federal Depts, Agen	243	11.1
Certified Ethical Hacker	State & Local Governments	295	4.1
(ISC)2 (TM) CISSP (R) Pre	State & Local Governments	232	5.2
CompTIA Security+ (SY0-XX	State & Local Governments	232	3.9
Advanced PCAP Analysis an	VETERAN (not Govt. Employ	1,227	24.7
CompTIA A+ Prep 2013 220-	VETERAN (not Govt. Employ	992	9.3
CompTIA Security+ (SY0-XX	VETERAN (not Govt. Employ	786	15.3

Again, we see the largest departments in terms of number of active users have the highest enrollment but lowest completion, in either the Security+ or CEH prep courses. The non-government employee Veteran group is the exception with a larger number of users enrolled in the 1-hour Advanced PCAP Analysis course. As this course was the first one in the list of all courses in FedVTE in 2016, and the rate of completion for the Veteran group is pretty low, it could be hypothesized that these users accessed the first course upon registration to get acquainted with the system and navigation.

Highest enrollment courses by email groups

Course Name	Number of Users	Percent Completions
CompTIA Security SY0-XXX Prep	10,639	30.0
Certified Ethical Hacker vX	6,996	32.8
Windows Operating System Security	3,830	60.1
ISC2 TM CISSP R Prep 2015	3,689	18.4
CompTIA Advanced Security Practitioner CASP CAS-002	3,574	26.7
CompTIA Security SY0-XXX Prep	1,096	12.6
Certified Ethical Hacker vX	1,093	12.1
Cloud Computing Security	995	77.2
ISC2 TM CISSP R Prep 2015	870	7.4
Advanced PCAP Analysis and Signature Development APA	804	35.2
Advanced PCAP Analysis and Signature Development APA	2,406	24.8
CompTIA Security SY0-XXX Prep	2,030	18.3
CompTIA A Prep 2013 220-801	1,833	9.1
Certified Ethical Hacker vX	1,548	14.9
CDM Module 1 : Overview	1,102	36.2

The highest course enrollment by email domain shows again the largest number of course enrollments are in certification prep courses for .mil and .gov users, with Sec+ and CEH being the top two. Completions for these lengthy courses again fairly low by both groups, indicating users partake in the topics they need.

Windows Operating System Security and Cloud Computing Security have fairly high completion rates. The Cloud Computing is a 1-hour course that was previously suggested may be required by an employer due to high completion rate by a single department. The Windows course popularity for the .mil users could indicate compliance with the 8570/8140 directive to obtain a baseline operating system course certificate.

Cost savings with FedVTE

How can this user course completion and progress data be used to estimate the potential savings in training costs FedVTE is affording to the federal government? If we look at actual costs for obtaining similar training, those numbers can be measured against course completion data from FedVTE.

Reported training cost data obtained from a Federally Funded Research and Development Center (FFRDC) from 2016, was used to determine costs to train an individual in preparation for a certification exam. Three different options for prep training were examined: traveling to attend training, subscribing to training online, and on-site instructor lead delivery.

Sending a user to a five-day industry certification prep course totaled just over $4,500. This figure includes registration for the training, hotel for 5 nights, and per diem meals for 5 days. The online subscription training included the exam fee with registration, and cost varied depending on specialty level of content and duration of access to the online material, ranging from $2,700 to $5,600. The mean of range, $4,150 will be used.

Finally, organizations may opt to have trainers come onsite to train a group of individuals in preparation for a certification exam. A 2016 invoice for onsite certification prep training for ten students including catering for five days, and onsite administering of the certification exam upon conclusion of delivery, was a little over $28,000, equating to $2,800 per individual. While the onsite delivery option is the most economical for an organization to train their users, onsite deliveries require a minimum number of students. While cheaper per individual, it's a pretty significant amount at once.

For estimating training costs and potential savings FedVTE affords the federal government, an average of the three options (4,500, 4,150, and 2,800), rounded to nearest hundred, $3,800, will be used as the cost for an individual to obtain certification exam preparatory training.

While there was a total of 72,468 course completions in FedVTE for the 2016 calendar year, not all were equivalent to a 5-day certification prep course. To help comparisons and savings estimates, only the 2016 course completions for certification prep courses will be considered. The nineteen available prep courses had a combined completion total of 13,678. If we multiply that total by the average cost previously determined for individual training, the total would be a whopping $51,976,400.

This total needs to be adjusted to account for costs associated with the FedVTE program such as hosting and administering the platform, content development, delivery, processing, and program management. These development and operational costs are shared across a couple federal agencies, namely DHS and DISA. The actual financial information for these functions aren't publically available. For the purpose of estimating, 3 million is a reasonable high-end figure for a year's worth of development, hosting and support.

All variables considered, a one-year savings in training costs FedVTE provides the federal government, would be at least 49 million dollars. This figure again only considered 19, or 1/3 of the courses available in FedVTE. It is also under 20% of all course completions for the year, and doesn't account for activity that didn't result in completion of a course. Consider, users who were preparing for a certification exam, but only referenced the topics in the prep course they needed to reinforce. This puts the estimate for money FedVTE saves the government in training costs, well over 49 million dollars in one year.

User feedback isn't available to know for sure the primary reason for utilizing training in FedVTE. Some data suggests that there are agencies that require users to complete specific courses. As illustrated previously, there were two departments that had a course where just about every user who enrolled, completed it. Having this type of feedback indicating whether users were mandated to complete training, and further, whether the employer would have to fund the training if FedVTE, weren't an option, would enable additional savings calculations with better accuracy.

Video file efficiency

Another consideration during the early development of DoD's VTE, was the difficulty some users may experience when accessing or streaming videos via the Internet due to bandwidth limitations in some deployment areas. Content streaming experience is important to consumers. Many viewers will abandon playback due to performance issues. According to the 2017 Video Streaming Perceptions Report, 20.4% of viewers will stop content viewing after a single instance of buffering.

The SEI's Cyber Workforce Development Team (CWD) has developed processes, methods and tools to produce digital video with the goal of minimizing data rate while maximizing visual quality. The fundamental technology for producing efficient video is digital compression.

Digital video files use a great deal of data. The amount of data depends on several factors including the video duration, size, quality, codec and compression settings. Digital video requires a codec. The word codec stands for compressor/decompressor. The codec converts visual images to encoded digital data.

The compression process takes the encoded digital data and converts it to a specified data rate that matches the distribution environment. Compression is both art and science and involves the manipulation of hundreds of settings. Videos must be compressed before publishing to a website or DVD.

CWD compression starts with a file that has a data rate of 100Mbits per second and ends with a file that is as low as 100Kbits per second – a reduction of 1000:1. The trick is to achieve high compression ratios while still maintaining image quality.

There are two types of compression: Lossy and Lossless. **Lossy** compression results in significantly less data than the original file. The end result is often a loss in quality in order to create the significantly smaller compressed file. With **Lossless** compression, data rate is reduced while quality is maintained. Lossless compression only removes visual data that is redundant. The result of lossless compression is a compressed file that hasn't changed much in terms of quality. Lossy compression is necessary in order to achieve the 1000:1 compression ratios that result in very low bit rates.

Bit rate is the number of bits per second needed to transmit a digital video file. A file with a higher bit rate will require more data space and take longer to move across a network. A higher bit rate often results in higher video quality. Video players, devices, and platforms vary in terms of targeted bit rate.

Current compression trends suggest an organization should compress and stream dynamic adaptive bitrate (ABR) content using APPLE HLS and MPEG-DASH in order to provide an end user the best possible viewing experience while limiting buffering interruptions. ABR standards work by detecting an end user's available bandwidth and CPU capacity in real time, and adjusting the video quality accordingly. If an end user is viewing video on a legacy mobile device with limited bandwidth, a video stream appropriate for the device and bandwidth is delivered. Alternatively, if an end user is viewing video on a modern desktop computer using an enterprise class network, a higher quality video will be streamed; if that user's bandwidth becomes limited due to heavy traffic a lower quality asset will be dynamically inserted.

Generally, ABR is a necessary solution. Creating multiple encoded assets ensures uninterrupted viewing to several different devices. Unfortunately, this approach is device and bandwidth specific while being content agnostic; it doesn't consider whether content actually benefits from variable resolutions and data rates.

Not all content needs excessive amounts of data to display in high fidelity. Sporting events contain high motion content with constant changes, while classroom training is typically only an instructor's movements. Streaming sporting events can benefit from ABR because available bandwidth can be saturated delivering a maximum image quality. When the bandwidth is diminished, the image quality dynamically decreases. However, the ability to view content is not interrupted.

A content producer may have as many as 10 variably compressed videos of the same content to support the adaptive bitrate model of delivery. Modern codecs were engineered with high-motion content in mind. Entertainment requirements differ drastically from that of training delivery, so a deep understanding of what codecs are, and how they function is important to producing efficient and effective content. Why generate multiple compressed videos when 1 will suffice?

The H.264 video codec (compressor/decompressor) technology was developed by the Moving Picture Experts Group, International Telecommunications Union, and the International Organization for Standardization/International Electrotechnical Commission Moving Picture Experts. Although the H.264 video codec can be used with multiple containers, it is most frequently wrapped with MP4. H.264 is a Lossy compression that can achieve efficient data rates in part because of motion compensation that works by analyzing video frames, and separating those frames into macro blocks. A static I-frame (Intra) is created and analyzed, then uses B-Frames (Bidirection) and P-frames (Predictive) to define future frames. When activity takes part in one macro block but not another, only the macro block that is affected will be recompressed.

Imagine a game of ping-pong: you have a static court with a single ball being directed back and forth by paddles. Our game of ping-pong only has action in a few quadrants. By referencing the static activity, motion estimates can be used to predict the action. Instead of compressing the entire frame, only a small part of the image is compressed – thus saving data. Encoding Profiles can stipulate how aggressive the codec analyzes video; tailoring a profile's parameters to your content is vital for successful compression. If you choose an overly aggressive profile without considering your content, you could see unintended image degradation.

Think of a training facility with an in-room projector: the lights in the room are dimmed so that the instructor and projector screen are both visible for the in-room participants. Yet the room is under-illuminated for the camera to accurately define all pixels, so the image appears dark. To compensate, the videographer raises input gain that brightens the overall image. But, those undefined pixels are still present and now appear as noise (tiny little dots). Although the image is more pleasant for a viewer, it is now problematic for our compression. Our compression analyzes macro blocks looking for pixel changes (activity). If we set our profile too aggressively, it can interrupt the noise as activity. The result will be an asset that is higher in data rate as well as lower in visual quality.

Compression costs include encoding, storage and delivery. ABR delivery requires the greatest cost. ABR assets are generated from a single mezzanine file. Apple recommends a 720P (1280x720) mezzanine file be encoded to 6 separate videos for HLS delivery.

1. 1280x720 3000Kbps
2. 960x540 2000Kbps
3. 768x432 1100Kbps
4. 640x360 730Kbps
5. 480x234 365Kbps
6. 416x234 145Kbps

The total data rate for all 6 formats is 7340Kbps. A video duration of one minute results in 440400Kb or 53.76MBs of data. An hour of footage using Apple's HLS compression tree results in 3.23GB of data. Local Encoding is a fixed cost requiring hardware/software and a support infrastructure. Cloud encoding costs range dramatically.

For our example, Zencoder.com offers pay-as-you-go pricing of $0.05/minute per video. Given our Apple HLS example resulting in 6 videos and 1 hour of content, the total equals 360 minutes, which equates to $18 for cloud encoding one hour of HLS content. A typical 5-day course is 30 hours x 6 formats equals 10800 minutes x $0.05 equals $540. Storing the video locally has a fixed cost for server hardware, software and infrastructure, while Cloud cost range drastically based on enterprise size and depending on vendor you may pay for storage, viewing or both.

Bandwidth fees for Cloud customers with storage costs above $1,000,000, pay $.0.005-0.008/GB. Mid-level customers with storage cost from $500,000 - $999,999, pay $0.007-0.012/GB. Retail customers pay $0.02-0.085/GB. To deliver 1 hour of content encoded to the 720P standard data rate of 3.2Mbps that equates to 1.44GB per hour of content. A mid-level CDN price of $0.06 x 1.44Gb per hour is charged $0.09 per hour viewer. A typical 5 day course contains 30 hours of content at $0.09 per hour would be $2.70 in bandwidth cost per viewer. A 5-day course that has 1000 viewers would incur a bandwidth charge of $2,700.00.

CWD's Asset Creation Collection Conversion (AC3), ecosystem allows the benefit of higher quality assets at lower data rates then industry standards.

Average Recommended Bitrate (Kbps)

Streaming Site	1280x720 (720P)	960x540 (540P)
AC3	450-500 Kbps	438 Kbps
Kaltura	3750-5000 Kbps	3750 Kbps*
Youtube	5000-7500 Kbps	3000-5000 Kbps*
Apple	3000 Kbps	2000 Kbps
BitMovin	2400 Kbps	1500 Kbps*
UStream	1200-4000 Kbps	800-1500 Kbps
Capella	2100-3100 Kbps	1500 Kbps
Netflix	2350-3000 Kbps	1800 Kbps*

*estimate based on cited data

Using AC3, CWD has the ability to deconstruct the final product so that a quality standard can be continually met without sacrificing a data rate target. AC3 targeted data for a 1280x720 (720P) asset is 500Kbps, a 1 hour video is about 220MB or .215GB. Since CWD can maintain such a low data rate there is no need to create an adaptive bitrate asset. If cloud encoding is done, a 1 one-hour asset at $0.05 per minute the cost is only $3.00, a 30-hour course cloud encoded only costs $90.00. Bandwidth savings are equally present. Using the same parameters $0.06 x .215GB equals a bandwidth cost of $0.0129 per hour viewed. A typical five-day course contains thirty hours of content at $0.0129 per hour would be $0.387 in bandwidth cost per viewer. A 5-day course with 1000 viewers would incur bandwidth charges of $387.00, resulting in savings of 86%.

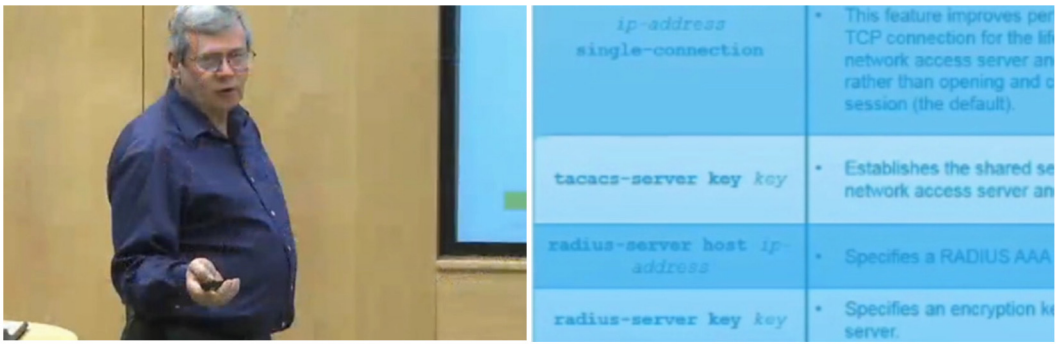
CWD consistently provides the highest quality training content at the lowest possible data rates. Using SSIMWave's Video QoE Monitor that enables comparisons using Peak Signal to Noise (PSNR) and Structured Similarity Index (SSIM), we can generate quantitative data to measure and compare data rates to video quality. SSIMplus Core is a video Quality-of-Experience (QoE) metric that measures how closely the compressed video matches the original mezzanine video, rated on a scale of 1 to 100. When compared to other compression mediums, AC3 is able to compress videos to a lowest bitrate while still maintaining a high level visual fidelity.

SSIMplus QoE scores across multiple compression types are compared below. On average, AC3's ability to transcode a 720P asset at 690 Kbps is measured as having 97.57 overall score. This metric ranks higher than the nearest of others: YouTube transcoding a 720p asset at 5000 Kbps, at a score of 97.2.

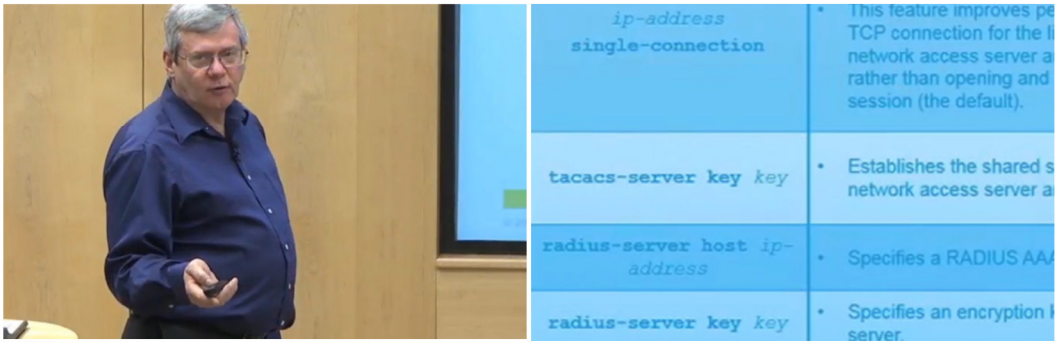
SSIMplus QoE Scores

Compression	Maximum	Minimum	Overall	Standard Deviation
AC3_540p_450Kbps	98	94	96.91	0.45
AC3_720p_466Kbps	98	94	96.33	0.59
AC3_720p_690Kbps	98	96	97.57	0.51
CBR_2pass_540p_450Kbps	94	83	91.36	1.69
VBR_2pass_540p_450Kbps	94	88	91.94	1.04
VBR_2pass_540p_900Kbps	95	90	92.68	0.77
VBR_2pass_540p_1800Kbps	95	91	93.02	0.72
YouTube_480p_358Kbps	98	92	96.39	0.88
YouTube_720p_5000Kbps	98	96	97.2	0.4
YouTube_720p_1209Kbps	98	93	96.6	0.62

Worth noting: While the tool's measure of "Excellent" quality falls in the range of 81-100, there are significant visual differences across the display of both people and text when comparing Overall scores of 93.02 and 97.57.



SSIMplus Core: 93.02 (Excellent)
Compression: VBR_2pass_540p_1800Kbps



SSIMplus Core: 97.57 (Excellent)
Compression: AC3_720p_690Kbps

CWD has been generating digital video assets in support of distributed learning, such as FedVTE for 26 years. The methodologies, processes and tools developed over three decades provide innovative cost and quality advancement. Streaming and storing bits cost money. SEI's CWD is able to provide quantitative and qualitative data that demonstrates how bit efficiency can be achieved while maintaining high-quality measurable standards.

Next steps

As detailed, FedVTE is a cost-effective resource for the federal government saving collective agencies millions in employee training costs, as well as by hosting uniquely resource-efficient video assets.

FedVTE could consider additional capabilities to further enhance its value to the cyber workforce. While different agencies may have differing training needs, the adoption and/or adaptation of the NICE Cybersecurity Workforce Framework has the cybersecurity field marching to the same drum. Job titles, and accompanying tasks and KSAs may vary, but the lexicon and methodology is universal. As such, having a learning system that tags and categorizes training assets accordingly, would aid in creating targeted training plans and assessments. As mentioned, FedVTE currently organizes the courses according to the NICE Framework categories and specialty areas. Tagging at a more granular level, the individual topical videos, to tasks and KSAs, enable users to more efficiently find content relevant to their needs.

The content within FedVTE courses is already in the form of small, easy to digest, standalone video topics. This design has allowed for ease and efficiency of course updates, and content reuse for custom course outlines. Leveraging this design further and identifying each topic's learning objective, related tasks or skills, and assessment, will create the ability for custom role or capability training and focused evaluations.

When training content is designed, the learning objectives are typically identified before the content is created. Including NICE task and KSA information as well during design, and mapping when content is added to FedVTE would result in more customizable, targeted training. A search feature where users could enter either a keyword or code to locate content related to a role, task, or KSA would support agencies' ability to uniformly train staff, maintain inventory of their human capital expertise, and identify their workforce gaps.

Another area FedVTE may want to examine, is maintaining the integrity of course completions that is earned by passing an exam. One of the challenges FedVTE is currently facing, is the publishing of the course exam answers on the Internet. This is not done purposely or desired, but the result of registered users using one of several flash card apps. These apps store and share virtual flash cards that individuals create for study and knowledge challenge purposes. While FedVTE courses have exams with large pools of questions that a random subset is pulled from to assess user's comprehension of the material covered in the course, it only takes so many users creating flash cards until the entire exam is shared. Especially when the system provides instant correct/incorrect feedback after a user answers a question. This issue is not unique to FedVTE. Many entities with exam requirements face the same answer sharing and leakage problem. There isn't a means to prevent it. While professional integrity has to be trusted to an extent, FedVTE can takes some steps to stay ahead of the question and answer sharing, or at least make it more effort on the app users' part to share an entire exam question pool.

Making the question pool smaller and updating questions periodically, will cause the virtual flash cards to be obsolete more quickly. For instance, instead of having a pool of 100 questions that 30 are randomly asked, have a pool of 50 that is refreshed after a set period of time. The number of questions and refresh intervals would depend on the popularity of a course. For instance, the two courses that are

by far the most utilized, the Security+ and CEH certification prep courses, would need more frequent question pool refresh. The same consideration is true for courses that data shows high rates of completion. While the severity of the exam questions and answers sharing problem is debatable, it is something that should be considered. Especially if there's risk to courses' good standing statuses for continuing education credits with certification agencies.

Finally, while this report was able to provide some insight into how users are accessing FedVTE content by analyzing course progress and completion data, additional data would help draw a clearer picture of the strengths of this resource, and the most effective way to mature the program. For instance, a mechanism to collect user feedback on their primary purpose for engaging in FedVTE content. Or a way for employers and training managers to convey their specific training requirements, and the driving force behind them. Whether it's an agency-specific directive, or one mandated at a higher level such as the 8570/8140. This information, as well as whether the training would be fiscally possible without FedVTE, would be beneficial to know.

Surveys would be the best instrument for collecting defined information to analyze, ensuring responses are within measurable parameters. However, there are concerns with personally identifiable information (PII) being inadvertently collected by users completing a survey within FedVTE. There are mechanisms by which users could complete a survey anonymously. For instance, a link from FedVTE to a 3rd party survey engine that doesn't require authentication and through multiple-choice questions, prevents users from revealing anything about themselves or specific employer.

Another potential feature that could help in assessing the needs of the cyber workforce, is a user forum within FedVTE. Encouraging users to provide comments and feedback on a discussion board, anonymously if desired, would not only allow FedVTE program management to have a finger on the pulse of their user base, but also enable community interactions on common issues. The discussion forum, like most online feedback mechanisms, would need monitored with posts being reviewed before published. However, this feature if utilized to its potential could see users recommending content, advising of updates to training directives, or even querying about specific topic or skill resources. All of which would be valuable information for gap analysis and program planning.

Conclusion

FedVTE is an undoubtedly valuable training resource saving federal, state, and local government agencies immeasurable costs in employee training. While more data is needed to apply an exact number to the savings, it's certainly a minimum of $50 million per year according to analysis of 2016 course completion totals versus traditional training costs. And the platform itself is unparalleled in bandwidth and storage cost-efficiency due to the video assets created by CWD's compression techniques. FedVTE's continued ability to demonstrate cost savings, as well as adapt training offerings as the cyber workforce needs mature, will help secure its value to government agencies for decades to come.

Acknowledgements

The development of this report benefited significantly from the input of CWD Technical Manager Dennis Allen, and support from colleagues Eileen Angulo, and Pamela Williams. Special thanks to John Antonucci for almost 30 years at the SEI advancing the practice of instructional video asset design and compression techniques. His expertise greatly assisted explaining in this report the innovative, cost-efficient processes employed by CWD.

References

https://static.mux.com/downloads/2017-Video-Streaming-Perceptions-Report.pdf
Average Bitrate Sources.docx

https://sidbala.com/h-264-is-magic/

https://developer.apple.com/library/content/documentation/General/Reference/HLSAuthoringSpec/Requirements.html#//apple_ref/doc/uid/TP40016596-CH2-SW1

https://zencoder.com/en/file-transcoding/pricing

http://www.streamingmedia.com/Articles/Editorial/Featured-Articles/VP9-Pricing-When-Does-Adding-VP9-Lead-to-Bandwidth-Cost-Savings-117996.aspx

https://docs.google.com/spreadsheets/d/1lKVSNGaxsrvz8e6S0JwQHDKHIqMGEKHw8-QJNA9pPJI/edit#gid=0