Electronic Records Modules

Electronic Records Committee

Congressional Papers Roundtable

Society of American Archivists

Configuring and Using Offline Computers for Access

Jim Havron

Independent Archival Consultant

jimhavron@comcast.net

Date Published: December 15, 2016

Module#: ERCM010

Configuring and Using Offline Computers for Access

Overview and Rationale

Because situations vary in which computers are used offline, and even the definition of offline is inconsistent, this document is meant only to offer guidelines in configuring and using such machines. Most computer work is done with machines that are part of a larger network. Libraries and archives are among the few institutions that maintain offline computers for public access to records, and the use by institutional staff in any field is limited. This document notes that the term "offline" generally is used to refer to a lack of connection to the Internet or some specific network, and some of the specific situations that might be characteristic of offline computer use, including particular data storage media. It further notes that with the average computer operating system being designed for connectivity, a lack of such connection implies a need for privacy or security. The workflow section defines some issues that should be considered when configuring such computers to stand alone, unconnected, with focus primarily on public access computers and staff use computers that must connect directly with individual databases and where privacy or security are issues. Computers held in a preservation state and computers that are unconnected because they are used irregularly or for undefined purposes are also briefly discussed, as each has specific additional characteristics to be considered.

Procedural Assumptions

It is important to note that hardware and software availability, the way devices are used and communicate with media and other devices, and both risk and defense methodologies for electronic records change very quickly. It is more common than not for policies and procedures to lag behind these changes. When considering the practices in this document, be aware of this changing electronic environment. Ask questions of those who work with your technology. In short, do not assume that things have not changed since this was written.

For the purposes of this document, we can group the offline computer into four groups. The rapidly changing world of technology, the uses to which it can be put, the expectations of researchers and the types of records and documents held in an archival repository, make it impractical to cover every situation, but these four groups make acceptable starting places for adapting general concepts to specific needs. These groups are:

1. Staff or institutional use devices, computers that are designed for use offline for institutional staff use in processing, accessing, researching, and evaluating materials, as well as preparing access materials, such as catalog records or finding aids, and handling records created by the institution itself.
2. Public research devices for use by researchers physically in the archives to access electronic

materials that the institution is unwilling or unable to make available through self-service, distance research (e.g. the Internet or some network with other institutions).

3. Unassigned/multi-use computers, such as laptops used for presentations, staff computers used on an irregular basis to access specific files, or computers that may be assigned to specific tasks or individuals on a temporary basis. These computers may need to be brought online at some point, but may be unused or in very limited use most of the time.
4. Computers that hold electronic records in their original state or for preservation. These computers may be donated machines, devices with copies of records that are unprocessed or restricted, or storage servers that hold restricted records that are not generally available to researchers.

The primary assumptions when determining procedures for these computers is the criteria for access to the records, the ability to limit access, the need to deliver access, and the need to protect the machines and the records. A key part of all this is the presence of IT staff, or lack thereof. Institutional IT may have regulations relating to access and will almost certainly have a role in any configuration of computers, printers, storage, or input/output (I/O) methods. The absence of institutional IT personnel will mean that archival staff will have to determine the rules and either implement them or obtain qualified personnel to do so. For example, an institution may have firewall specifications and rules that must be applied to all firewalls, while other institutions must arrange for firewalls themselves.

Remember that, in most cases, IT staff are not archivists and do not have the same goals in handling electronic records as archivists. The rules for the larger enterprise may be at odds with those of the archives. Negotiations may be necessary, and archivists should maintain open communication with IT and regularly evaluate the tech environment to see if adjustments are needed.

The general state of institutional technology will also have an effect on the procedures used with offline computers. Within the same institution, the age and capabilities of computers and I/O, as well as the age, variety, and capabilities of software, may vary greatly. Storage is also a potential issue, with much storage used different enterprises being off site, requiring connection through a network which may make the process online by definition, depending on one's definition.

Hardware and Software Requirements

Hardware and software for computers designated solely for offline use can be problematic. To begin with, one must consider the task that the equipment is to accomplish. We will assumed that the specific software needed for the job, apart from considerations of whether the computer is online or offline, will be handled based upon the specifics of that use. For instance, if the computer is being used for cataloging or producing finding aids, the appropriate software will be acquired. There are a wide variety of applications for accessing different types of databases, viewing or editing images, and handling audio/video. Although the online/offline capabilities of the computer may make a difference in the choice of these applications, they generally depend more on the specific task that they are designed to accomplish than on network access. For this reason they will not be discussed here.

Four primary considerations must be made in deciding what configuration one wishes to use with the offline computer. They will be pretty much the same whether one is dealing with a staff use computer, public researcher computer, or multi/occasional use computer. (Data preservation computers have their own special needs and will be discussed later.)

1. Access levels – Access to records should be set to the minimum level allowed only. One might think of this as "need to know." Records should only be accessible to those authorized, and software and hardware should be used to assure such restrictions are in place. This may mean virtual machines, data filters, firewalls, or physical separation of record storage from the machine. Hardware and software for multi-factor authentication may be required.
2. Input – Records may be stored on all manner of media. Optical media, such as CDs and DVDs are still fairly common, but it is increasingly the case that records are on hard drives, often external drives, or even remote servers. Storage media must be considered when configuring offline devices to allow access to the necessary records, and one must decide whether this means allowing some degree of online contact, in which case network security must be configured to segregate the computer as much as possible to protect the limited nature of access. Firewalls, encryption, and malware scanning software will be considerations here.
3. Output – One of the biggest dilemmas in having the controlled access desired in an offline computer is determining how the machine will provide output for the work done. This may be easier to control on computers used by staff if the output is going to storage with controlled access, but there may still need to be special software or hardware to accomplish the task. In the case of both staff use and public use computers, special care must be taken with allowing connections through USB or other ports, possibly requiring limiting output storage hardware. Users today are used to being able to transfer records through online networks and cloud storage, so policies will need to address the alternatives offered. When possible, a dedicated printer connected directly to the computer is preferred to a network device.
4. Software and firmware maintenance – Several issues must be considered when configuring a computer for use solely offline. Not the least of these is the rapidly increasing movement of

computing to “the cloud,” not just for storage, but as a computing platform. Many software suites now function online, in some cases requiring frequent online connection and others actually doing the work entirely online. This is becoming increasingly problematic, even for computers that generally work as part of online networks. Even offline computers must update software and firmware, requiring some sort of access to the updates.

In addition:

1. Plans for backup of data must be made, which may require special hardware and software.
2. The degree to which these factors are a concern will, of course, be reflected in the degree to which the computer is meant to actually be offline.
3. A computer that is meant to isolate records as much as possible from unwarranted access will have backup security needs that may be fairly stringent.
4. A computer that is offline in some respects, but in such a way as to restrict access to a networks as opposed to always prohibited may have different security needs when it comes to backup.
5. For security reasons, records will need to be backed up in another location, so in all cases some access will be required. The speed and efficiency of backup may be affected by the computer’s state as online or offline.
6. Obsolescence of individual machines could be an issue, not only in maintenance of older machines for preservation, but because it is increasingly difficult to totally disconnect from networks on a permanent basis, meaning new machines can be problematic.

Workflow

To reiterate, the assumption is that we are setting up parameters for using computers in an offline state, in at least a large part to control access to records and maintain control over them according to professional standards. Because of the great variety of electronic record formats, types of hardware and software, rapidly changing nature of technology available to creators, users, and preservers of such records, and a myriad of other reasons, specific recommendations are not possible here. Additionally, if one is looking for the best way to use a computer that is generally connected to a network but which one wishes to occasionally use offline, the place to look for guidelines and practices is under the category of online computers. This is not to say that some of these practices might not be of worth in such situations, but as will later be shown, special steps must be taken to secure access in these cases.

Most of the work in providing safe and effective access to records through offline computers is in configuring them for the specific needs for which they will be used. The majority of the rest lies in training and communication.

General Points to Consider for Standalone Computer Policies

1. Whenever possible, machines, and users, with different levels of access, should be separated.
2. Whenever possible, records with differing levels of sensitivity or authorized access, should be separated.

3. In the event that physical separation (such as limiting access to a specific computer only, or requiring use of a specific computer to access a single storage device) is not possible, use roadblocks and compartmentalization. Put as many hurdles as possible between the entities with different access needs as possible, such as firewalls or access codes, and use virtual machines or drive partitions to separate the resources for the task into separate compartments. (Most users of computers with Windows operating systems will see what appears as two separate hard drives on their computer out of the box. There is, in fact, a single drive that has been logically partitioned into two separate areas, which the computer sees and treats as separate drives.)
4. Use credential protocols, such as user name/password, a token, or a card, to control access to different "compartments" or equipment. Multi-factor access protocols, using more than one of these methods of authenticating the user's identity and right to have access may be desired. The level of credentialing is likely going to be dependent on the sensitivity of access. A single username/password combination might suffice to access a machine used to produce inventories and finding aids, but a much higher state of security should be necessary to gain access to sealed records of constituent casework with PII or information covered under HIPAA.
5. Access may be dependent on a person's position or role in the organization, rather than tied to an individual. In such cases, credentials may need to be shared, although it would be preferred for each individual to have separate credentials. Whenever an individual leaves a position or role with access restrictions, credentials for that role should be changed.
6. There may be a need to frequently change storage media to control access. External hard drives may be required for different collections, or the records used may need to be brought to the computer through temporary storage, such as flash media. While the use of optical media such as DVDs and CDs is becoming obsolete, it is still prevalent in many archives and libraries, and provides a viable alternative for keeping records separated until accessed.
7. If a computer is used with restricted access, active memory should be cleared immediately after use.
8. Change passwords or other credentials often.
9. Consult with your IT department, or whoever provides IT support periodically to be sure they understand your processes and have made no changes in them without your knowledge. IT staff in large institutions often change things to respond to a certain need, but the changes can influence some processes without knowing. (The author of this module once found that firewall rules had been changed to keep audio and video files from being accessed across the network boundaries of a college campus, at least on university owned machines. Our archives provided online access to many of our oral history collections, so, needless to say, this was a problem.) We cannot expect IT to understand our needs unless we make them clear.
10. Avoid generic login accounts that work with the same credentials on more than one staff-use machine. If possible, all staff should have their own login. There will likely be some other method of credentialed access to research machines.
11. Discuss machine use regulations regularly with staff.
12. Consider blocking USB access to computers used by staff. This may need to be weighed carefully against staff morale, in the event personnel use the computer to listen to music from mobile

devices while working. This is still a significant security vulnerability, just the same.

13. Some simple log should be produced to show who is using what machine and for what. This information will allow administrators to track the last user of a machine that develops a problem and ask questions. It will also allow administrators to task staff with simple maintenance, anti-malware scans, and to be sure that machines do not sit idle for too long without being updated or scanned.
14. Be sure that standalone computers are included in your emergency preparedness and disaster recovery plans, along with any special software or records stored on them.
15. Develop a policies and procedure guide specific to your machines and collections, and update it whenever there are changes.

Points to Consider for Standalone Computer Configuration

1. The first step in determining the process of configuring your offline computer arrangement is to determine what the needs are in regards to both access, and denial of access. A part of this is to determine the risk of losing control of accessibility. In the not unlikely event that more than one access arrangement is needed, the archivist must determine the need for additional machines, establishment of virtual machines, or acceptance of a modification of access.
2. Computers that are configured for offline use only, sometimes called stand-alones or "air gap" devices, should be considered most often when truly sensitive data is being handled by the device. The handling of such information will be decided based on the level of that sensitivity. Sensitivity of data, of course, can be determined by numerous factors, from HIPAA-protected data to records that are restricted by donor agreement for use only with donor permission. The archivist must determine the level of effort and resources that should be expended to prevent unauthorized connection with the records.
   Consider the following:

   A. Records that are stored on a stand-alone computer will be on the hard drive in some form even after they have been deleted. Additional steps can be taken to make it more difficult to recover deleted records, and casual access to the computer may not reveal the remnants of the data to an unauthorized user, but there are simple software packages that can, recover, decode and copy many deleted files. If you use a computer offline to protect sensitive records, do not assume they are safe when the computer is not in use or the records have not been thoroughly removed from storage.
   B. If the data itself is very sensitive, one should consider encryption while in storage.
   C. Records used on an offline computer but stored on external storage media, will exist in memory on the computer, potentially after the external media is connected.
   D. The longer a computer is used without a network connection for updating software, the more vulnerabilities will exist in the software. Design a process by which software can be updated periodically.
   E. If the computer is dedicated to a specific task that only requires certain software, consider disabling or even removing unneeded software. For example, if the computer is

only used to view records in document or PDF formats, there is no need for active Microsoft PowerPoint on the machine.

F. If the computer is not truly a stand-alone device, but is rather a computer that has inactive connections or is not connected because it is a virtual machine, additional care must be taken to determine that access is limited to authorized personnel only.

3. There may be collections of records that have different levels of access allowed. Administrators should determine the level of access. For example, a database of constituent correspondence may include personally identifiable information that should be protected, but also data that is appropriate for researcher access. It might even be appropriate for the archivist to view restricted information to answer specific questions that are acceptable research queries, but might not be answerable without viewing restricted records as well. In such cases, filters or access restrictions should be in place and online connections should not be used for the open records so that the restricted records will not be put at risk.
4. The most common single way that electronic information is acquired without authorization is through physical transport. Customers, staff, researchers, thieves and spies regularly remove data on portable storage, either intentionally or accidently. Solid state flash memory, such as memory cards and USB thumb drives transport data out of an enterprise while at the same time often leaving malware behind. In addition to these storage devices, media players, phones, and other mobile devices are on the list of culprits. Because we are speaking of online/offline in terms of connection to other devices, this list comprises means of connection, effectively allowing communication almost as much as if the computer is online. Control of storage devices must be part of securing access, both to protect electronic records and the computers.

A good way to begin is to list the vulnerabilities to the electronic records and compare it to both limits to access required and the amount of access desired. Remember that this examination must be made in context of the environment of resources available. It would also be good to remember that security and assurance of data integrity are often at odds with convenience or expectations of accessibility.

Public Access Computers Specifics

1. Public research computers that are not intended to be connected to some network should have external connections disabled when not in use.
2. Wireless and Ethernet should be at least disabled by an administrator in such a way that only an administrator can re-enable them. This should also be considered for any other input connection, including card readers and USB ports.
3. Discuss with IT professionals about <u>disconnecting</u> v. <u>disabling</u>; not necessarily physically disconnecting, but removing the ability of the machine to reconnect without reconfiguring as opposed to turning off the access port. Consider disconnecting wireless, Ethernet, HDMI, card readers, USB, or other ports, as opposed to simply disabling them. This will take considerably more time on the part of IT personnel, but may be necessary in places where computers are not always within sight of staff and where there are a high volume of researchers. For example, a religious archives has a computer that is used to access copies of news articles stored in a

database and available for viewing but not for free digital copies. The computer is in a small recess in the research area and is not within sight of the staff. Researchers managed to reactivate disabled USB ports, and make copies of the articles, along with some images that were available for research but not for distribution.

4. If ports must be left open, use USB port blockers. These devices fit into the USB ports like little plugs and are not locks, even though they come with Keys. The keys are just tools that make it easy to remove them from the port, a task that can also be accomplished with tweezers or two small kitchen knives. USB locks are available, but pricey and, while more of a hindrance to access than the port blocker, they are not infallible. Still, both of these items will hinder someone trying to access the computer ports.
5. Determine policies and procedures for computer output. How will the researcher acquire the product of his or her work? While traditional note taking is still an option, it increasingly fails to meet the needs or desires of the researchers, particularly with the large amount of data in records and the skill sets of today's researchers. Some options to consider:
    A. Printers are still handy, although researchers increasingly desire electronic formats for their research outcome. A printer hardwired into the actual computer would be the best option. Barring that, a network computer that is separated from the network by a firewall, (as is the research standalone computer) is acceptable. It may also be possible for staff to copy the records and print them from a staff network computer.
    B. If there are no use restrictions on any records accessed by the machine, and USB ports are enabled or can be enabled, the files can be copied directly to portable media, such as a <u>clean</u> flash drive, by the researcher. The clean part means that the archives is assured of the safety of storage device. USB drives with a fairly large amount of storage can be purchased in bulk (which can mean as few as 10 for those with tight budgets) and sold to researchers for an amount that might be less than the cost of paper copies. Much less, if the files are large in number.
    C. Optical disks, generally CDs or DVDs, can also be used if the software and hardware are available. The best situation would be to have the drives disabled or write-blocked on the machine so that data can only be written to the optical disk (write blocking keeps data from being written to the drive on the machine), or to use an external drive operated by staff. As with printers, the files could be transferred on known safe media to a staff machine where the disk can be burned. As with the USB storage mentioned above, the institution should provide the disks.
    D. As with printers, the files could be transferred on known safe media to a staff machine where it can be emailed or otherwise provided to the researcher.
    E. In some cases, research will be done by a staff member with access to records that the average researcher does not have. Staff must be particularly careful in such cases to only copy appropriate data.
6. Keep in mind that the "computer virus" (malware) has existed long before the Internet was commonly in use. Floppy disks and, later, diskettes often spread malware from device to device. Today, the Internet of Things, the vast amount of devices that communicate with

each other in the background of our activities, has provided access to networks and devices that were believed to be defended. Malware scripts are very small and usually well hidden. As of mid-2016, security specialists generally agree that malware lives undetected on a device for up to 256 days, unless it reveals itself or proper security discovers it. Any device that a researcher inserts into a computer could spread malware or download data if one is not alert.

7. Methods of supplying the results of research will vary with institution, the records, and the technological arrangement. It will be a "people issue" as much as anything. Just as it is often difficult for researchers to understand why they are restricted to pencil and paper in the manuscript room or why all records have not been digitized, it is very difficult for them to understand the precautions taken with standalone computers. It may help to consider ways that these devices can be shown to be different from other access devices, used for providing extra layers of security, not as the old computers that have not yet been updated. Rhetoric aside, perception of the value of the computer configuration as offline may be a useful tool in obtaining compliance from researchers.

Unassigned/Multi-Use Computers Specifics

Some institutions have computers that are not in regular use. These may be laptop computers used to run presentation software, field work devices used to inventory or take notes away from the institution, or mobile devices used for audio/visual work. Regardless, these devices should receive the same treatment as those already listed, with some additions.

1. Determine if the devices in question are to always be used in an offline state. We tend to consider these machines to be auxiliary devices, and in many cases the need for network connection will vary.
2. Whether the device is used online or offline, care must be taken to provide regular maintenance, including charging, malware scans, and updates.
3. If the computer will be used offline, develop a checklist for the user to help check the system and be sure all network connections are properly disabled.
4. If records that are not meant to be publicly accessible are to be used with mobile devices, encrypt and password protect the data and the device.
5. Use physical locks on mobile devices. The records on the computer can be stolen or corrupted if the device itself is taken or unauthorized access is attempted.
6. Just as it is a good idea to have a checklist for system configuration that assures offline status, one should make note of the settings when the device is acquired and then restore those settings. Apart from courtesy, this reduces the chance that someone will try to adjust settings on the computer and inadvertently change something that may be missed the next time the computer is adjusted for offline use.

Static/Preservation Computers Specifics

Computers that are meant to hold data in a secure state for preservation are generally kept offline and have controlled access. They tend to be servers, computers designed to hold data until needed and that are made to run for longer periods of time than a desktop. Because they are generally not accessed on a regular basis, the primary requirements are to have a full backup of the data, and both the primary and backup be tested on a regular basis for physical and logical integrity. Checksums are nice, but in many cases not adequate. A checksum can tell you that there is something wrong with the 32G video file on the server, but not where the corruption is or what form it takes. If the data is sensitive enough, a full bit-to-bit image of the original disk should be made.

If the computer is the original device that held the records, and it is intended to preserve it as exactly as possible, the computer should be write protected. The mere action of powering up a computer will write data to the disk, let alone the act of connecting outside communication, even looking at the system. Write protection can be applied and a bit-to-bit image can be made of the original disk. If the device contains very valuable information, it might be worth making two disk images, one as a preservation backup, and one that can be examined by archivists. Even the drive being examined should be write protected when used. Malware scans may be done to this drive, as well as analysis of the data storage patterns. It might be worth noting that malware that is not in motion, rather lying dormant on a computer, may not be detected by initial scans. Later scans using updated signature files in the scanning software could detect malware that was not detected in previous scans.

All preservation devices should be kept in a stable environment and examined periodically. It would be wise to keep track of the availability of parts for the devices in the event that replacement parts are needed. This is true of software as well. It does little good to preserve data that is machine readable, but not have a machine to read it.

Summary

Most enterprises do not use offline, or unconnected, computers except for special circumstances in which special care must be taken to control access. With data defacement and simple opportunities to attack a system being major motivating factors in cyber-attacks today, libraries and archives are vulnerable. They have used offline computers for years for numerous reasons, but the primary reason for deliberately doing so today is to remove avenues of attack for threats to data or access control. The variables are too great to write hard and fast rules for every institution, but by considering the concepts of controlling access to make data accessible only to those with the right to access it, to keep systems up to date without putting them at risk through network connection, and by devising reasonable steps to access data and disseminate it in a secure manner, offline computers can be quite useful in preserving records and providing controlled access in this day of being connected.

Disabling/Enable USB ports temporarily in MS Windows

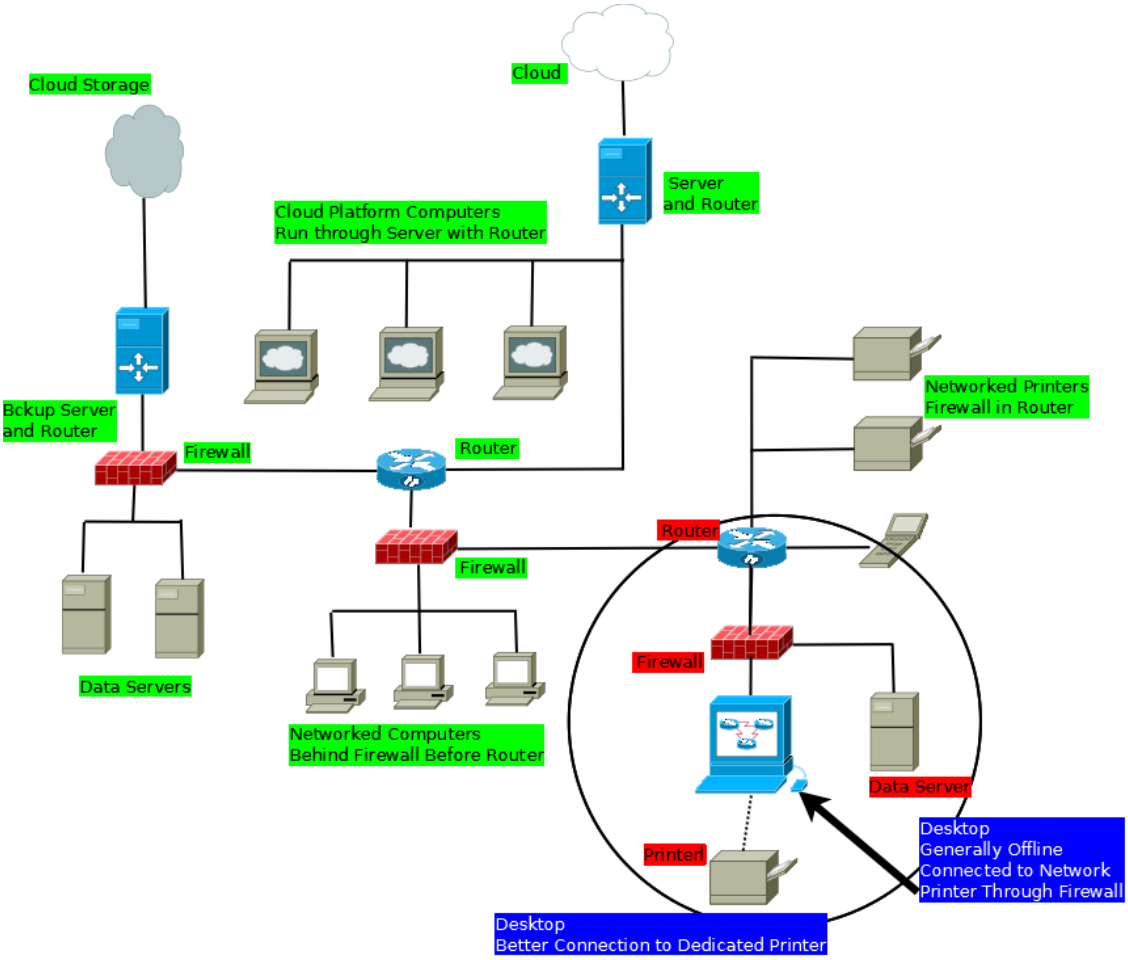
There are several scripts that will do this in the registry. They depend a bit on the version of Windows.

You can generally disable the ports from the device manager in your control panel. Again, the path to this depends on the version of Windows. It is generally under “system” and or “Hardware”. These can be found in some of the configurations after Vista under the “maintenance” tab. If you have your control panel set to the classic view, it is on the opening screen.

In Windows 8 and 10 it pops up quickly in the search category off the Start Menu.

If you expand the Device Manager, (some older configurations have hardware to be opened first,) you should be able to find the Universal Serial Bus controllers (aka, USB). Expand that and you will find the USB Root Hub. There will be one for each active USB connection on your motherboard, whether there is a port or not (there usually is or it would not be physically activated.) If you right click on each of the ports, you will have an option to disable or enable, depending on the state of the port at the time. If it offers the choice you wish, disable for example, click it. **Do not click disconnect!** Repeat for each of the Hubs. Undo your actions by coming back to the USB Root Hubs and repeating the process.

Example of an offline computer (in circle, lower right) **that is connected to a network in a limited manner.** There is an extra firewall in between the computer and the network router. This provides some security. If the computer used a dedicated computer connected directly to the computer, (dotted line) instead, it would reduce network connection. Notice this computer is also connected to a server for data. The server is also behind the firewall. If the configuration was changed, the computer could be disconnected from the network, assuming others didn't need the same data server.

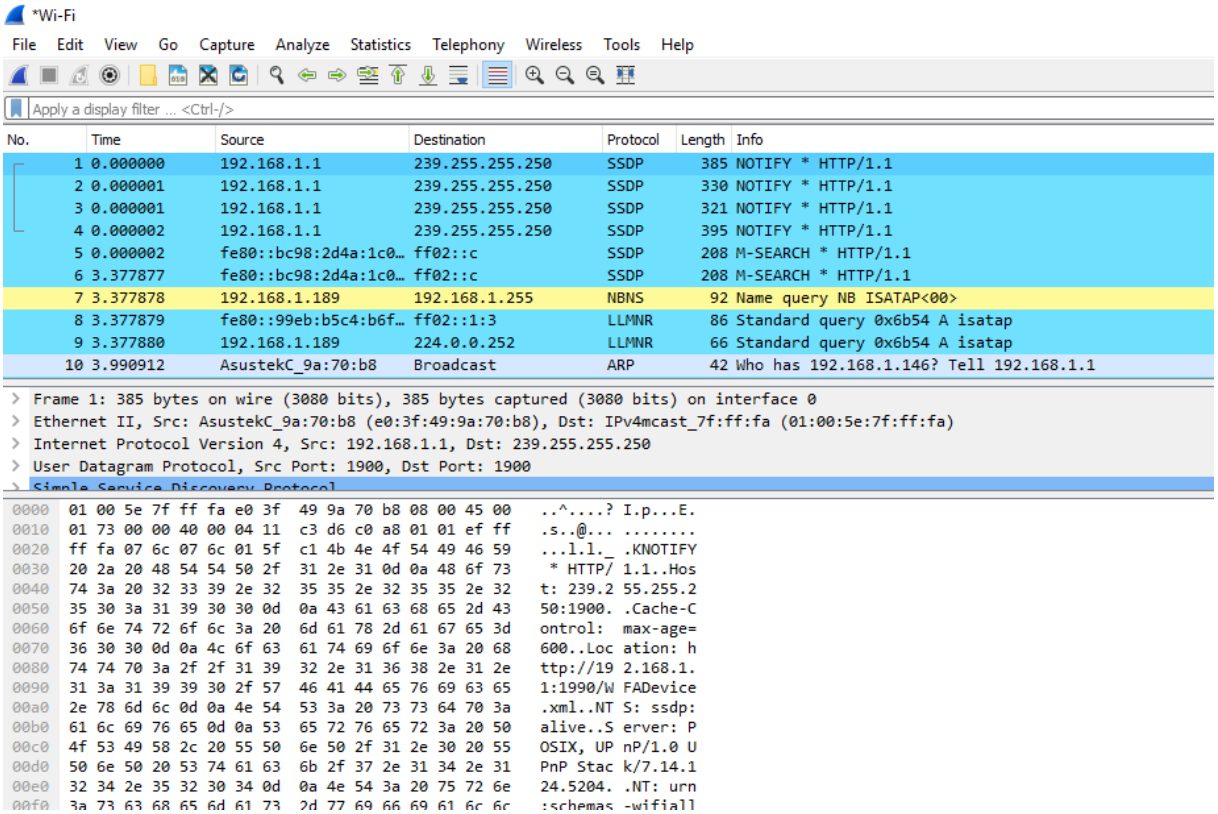


How easy is it to slip unauthorized data into an “offline” computer?

A public library history room has four computers dedicated to genealogical research. They are offline. Using a very small flash drive that is no bigger than a penny, it was possible to release a script that turned on the wireless. The flash drive was placed by casually placing the right hand behind the computer as if exploring it and checking out the situation.

The device on the right is a 32G flash drive. It was easily slipped into an unguarded USB port.

Software was then run that was able to pick up the packet information of wireless devices in the immediate area, including the computer.



Screenshot of display of packet information captured by
Wireshark software. There was much more data,
but this is a part that shows fairly standard IP info and poses no risk.

Because a backdoor route was used to access the machine as an administrator, there was no problem downloading software that allowed remote control of the computer. The penetration tester was not especially skilled, and so the process took about 10 minutes. At the end of that time, the tester could control the computer from his laptop, download software, save up to 32G of data on the small drive (they come with more storage for the same space drive), and upload anything he might have had on the drive. He admits that the flaw in the administrator access control was lucky, but it is a common vulnerability in machines of this configuration.