Final Report: March 1, 2008-June 30, 2009

NHPRC EMCAP Tool

Preservation of Electronic Mail Collaboration Initiative

NORTH CAROLINA DEPARTMENT OF CULTURAL RESOURCES

BEVERLY EAVES PERDUE, GOVERNOR
LINDA A. CARLISLE, SECRETARY

OFFICE OF ARCHIVES AND HISTORY
JEFFREY J. CROW, DEPUTY SECRETARY

TO: Lucy Barber, National Historical Publications and Records Commission

FROM: Kelly Eubank, Project Director, Electronic Records Archivist

SUBJECT: Electronic Mail Preservation Collaboration Initiative, Grant Number NAR RE-05701-07 (North Carolina, Kentucky, and Pennsylvania)

Progress Report: September 2008–February 2009

The Electronic Mail Preservation Collaboration Initiative aims to build a prototype tool for the collection and preservation of archival electronic mail. Staff from the states of North Carolina, Kentucky, and Pennsylvania have teamed together to leverage their knowledge and skill sets to help design and test a tool that will collect electronic mail, maintain a copy of the original e-mail and the attachment, as well as transform the e-mail itself from its native format to an XML format based on an XML schema. The team developed online training materials. The project team includes: Kelly Eubank (NC), project director, Druscilla Simpson (NC), head, IT Branch; David Minor (NC), programmer; Ed Southern (NC), head, Government Records Branch; Mark Myers (KY), electronic records archivist; Glen McAninch (KY), head, Technology Analysis and Support Branch; Linda Avetta (PA), information technology generalist administrator; Cynthia Bendroth (PA), head, Division of Records Administration and Image Services.

MAILING ADDRESS
4610 MAIL SERVICE CENTER
RALEIGH, NC 27699-4610

TELEPHONE 919-807-7280
FAX 919-733-8807

LOCATION
109 EAST JONES STREET
RALEIGH, NC 27601

Performance Objectives:

1. Continue the development of the e-mail collection and preservation software to convert e-mail from its native format into the more stable XML format and complete written documentation for the program.

Throughout March 2008, the programmer for NC focused on completing version 1.5 of the EMCAP tool within the time constraints authorized by the Chief Information Officer for the North Carolina Department of Cultural Resources (NCDCR). This work included finishing the administration console and building the functionality to import a .pst file, an additional task to the original work plan. During the course of developing the tool, the team expressed the desire that the EMCAP tool be able to ingest Microsoft Exchange personal storage table files (.pst). This resulted from the consolidation in IT services taking place in PA and NC.

Kentucky came into the grant as an already centralized Information Technology state. Pennsylvania was in the process of becoming a consolidated IT state while NC was the most decentralized with state agencies using a variety of e-mail platforms. Near the conclusion of the grant, NC, through the budgetary process, moved towards an enterprise e-mail platform of Microsoft Exchange. That being said, some agencies, such as the Secretary of State's office in NC, still use other e-mail systems such as GroupWise. With all three states moving towards using Microsoft Exchange, the EMCAP partners agreed to build tools to enable EMCAP to move from a project based initiative into a program. A key consideration for this was the ability for the tool to be able to ingest individual e-mails from other platforms such as GroupWise as well as Microsoft Exchange. For production purposes, the ideal tool for all three states would be able to synchronize directly with Microsoft Exchange and be able to ingest and export Microsoft file formats.

By the end of March, the programmer had finished the administration console piece and released it back to the group for testing. He then focused on building the .pst import function as well as fixing defects identified by the testers. During the same time, because of a different project, NC hired a temporary contractor to review and draft new e-mail guidelines and policy for e-mail use in NC and how to save and organize e-mail files. Her employment ended March 31.

In April, NC hired a temporary contractor to help test and document how to use the administrative console piece of the EMCAP tool. Mary Samouelian joined the project April 15 and began testing and documenting by developing and using test cases (see Appendix xx). Working under the direction of Ms. Eubank, the contractor diligently tested each aspect of the tool and documented her findings. For the most part, the tool performed without any major flaws with the exception of the "take account offline." When activated, the function actually wrote over e-mails from the account of the Secretary of State's office that had previously been imported and processed. Unfortunately, it was not just a matter of having the Secretary of State's staff re-copy these e-mails into their EMCAP account and then import them into the EMCAP tool for reprocessing. When we first entered into the pilot phase of the project, we asked our partners to copy e-mails into their EMCAP accounts rather than move them. This way, we would be working with copies of the e-mail rather than the originals. Training documentation was developed to demonstrate how to copy e-mails rather than move originals. The staff at the Secretary of State's office, however, mistakenly moved original e-mails into the system rather than copies. In addition, they did not retain copies of the e-mails in their accounts. As a result, access to those e-mails has been stripped. Despite this setback, according to the programmer, the e-mails still exist in the database but code needs to be written to pull those e-mails back out. We are aware of this major "bug" and it is in the queue to be fixed in the next release.

Ms. Samouelian also worked to install EMCAP v.1.5, and document the installation process. In writing the installation package, the programmer decided to focus on writing a package to install EMCAP as a first install. This was in part because previous versions of EMCAP would not be made available to the public. Only the most current version would be made available for download by the public. So, while we have documentation to install EMCAP v.1.5, we do not have documentation to upgrade from EMCAP 1.2 to 1.5. To write the detailed instructions, Ms. Samouelian installed the program a total of fourteen times; identifying defects during the process. These are available at www.emcap.ncdcr.gov (See also Addendum A).

The EMCAP team held a conference call in May to discuss the developments to date including new features, the test cases, Ms. Samouelian's experience in testing, and installing the newest version of the software in the other states. Because there was not detailed documentation on upgrading from 1.2 to version 1.5 and each state had this version in use, the programmer in NC offered to help each state install the upgrade. Each environment was slightly different and he thought it would be of greater help to be available via phone to walk through the installation. Kentucky accepted the offer, although it took some time to get the new version loaded onto the Kentucky server. Although the programmer offered to step through the installation with IT staff in Pennsylvania, PA was unable to install the upgrade because of lack of detailed documentation on how to upgrade. NC had also put in a purchase request for a license to GoToMeeting™ software. GoToMeeting™ software allows users to have virtual meetings using the Internet. This particular software allows users to share their desktop screens and would have been very helpful in this process. However, because of budget restrictions put in place at the end of the NC fiscal year, the process for purchasing became very cumbersome. All purchase orders had to be cleared through all of the normal channels as well as through the Office of State Budget and Management. As a result, the order took quite a while to complete and was not available at that time.

NC sent the test case documentation to both PA and KY. In June, the staff in KY stepped through the test cases and discovered additional issues that were addressed and fixed. Namely, KY identified issues with passwords, changing account names, and the process being unable to parse e-mails with long or malformed send to addresses.

The programmer released version 1.6 in August 2009 and with fixes to the issues identified by KY. The documentation was also updated to reflect these fixes. In addition, NC added language to the installation document to say that MS Outlook 2007 needs to be installed onto the server in order for the "Import .pst file" function to work successfully.

At the end of the project, the software has been developed and tested. We have identified issues that still need to be resolved—item counts with importing of .pst files, overwriting data in accounts with the import of a .pst file, and no de-duplication of messages and attachments. However, NCDCR continues to supports this software and NC is seeking ways to continue development of the tool and to identify continued, sustainable support.

2. Test the software to determine its scalability, efficiency, and employee compliance, using records retention schedules.

In NC, at the conclusion of Governor Michael Easley's term, the Director of the Governor's Office of Community and Citizen Services left and moved over to the Office of the State

Controller. Ms. Eubank spoke with her about her experience with EMCAP and her thoughts on its ease of use. She said that the tool was easy to use once Microsoft Outlook was loaded onto her machine and the initial bugs were worked out. She highly recommended its continued use and cautioned against returning to the previous method of transferring .pst files every six months via a DVD. In January 2009, Beverly Perdue became the first female governor in North Carolina's history. In February, Ms. Eubank contacted the office via the records analyst and the chief records officer at the Governor's office to see if the new staff might be interested in continuing its use. At the time, the staff had not been established and North Carolina was truly feeling the effects of the downturn in the economy. It took some time to establish a meeting with this group as the Legislature was in session trying to deal with the falling revenue collection. In addition, the Governor had to put forth her budget first in the budgeting process.

In late February, Ms. Eubank met with the new director and her staff as well as the IT support person to explain their responsibilities under the retention schedule, the EMCAP tool, and how it worked. The new director agreed to continue their use of EMCAP but she still needed to decide who would administer the account. This was not resolved until September 2009 as the budget crisis worsened and the Governor's office continued to experience personnel changes. The office did transfer the January to July e-mail files via a .pst file on DVD. To assist in the process, the NC team created a new account to reflect the new administration and imported the .pst file into the new account. During the process of importing the .pst file, the NC team experienced a difference in the item counts from Outlook to EMCAP. Because this was a high priority to the Governor's office, the NCDCR Chief Information Officer reassigned the programmer to work on the project to fix the discrepancies in the number count. In doing so, the programmer discovered that the open source hmail software does not like and cannot handle periods in folder names. For example, the new administration began creating folders with numbers followed by a period, e.g. 10, at the beginning so that the folders would sort properly. While this helped with sorting, hmail did not know how to process these. The NC team spoke with the office and asked them to replace the period with another character, a space, or an underscore. The office agreed to make the changes. Additionally, because the Office of Citizen and Community Services has such large e-mail accounts, going forward, the NC team decided to create a new account each year. This will help with performance and retrieval. Additionally, the scope note and finding aid for this material will include information regarding the multiple accounts for this office. In November, DCR staff met with representatives of the Governor's office to discuss usage of the tool and provide training on how to use it. We are in the process now of deploying this tool office-wide.

Kentucky

During this period, a small number of e-mails were added to EMCAP from the current and previous Commissioners of the Kentucky Department for Libraries and Archives. The current commissioner found EMCAP to be an easy and efficient means to regularly copy files to the archive in folders he has set up. The review of the previous commissioner's e-mails by the head of the State records branch continues slowly. We continued to talk to the cabinet secretary about using EMCAP and received approval of the project from staff responsible for e-mail management in the Commonwealth Office for Technology. However, we reached no agreement on transfer from the Cabinet before the end of the project.

To test scalability as much as possible EMCAP was tested (particularly the .pst import and the conversion to XML) with additional non-archival files from Glen McAninch's account. Many of these were more than ten years old which enabled us to test an earlier version of Microsoft Outlook e-mails.

Pennsylvania

Pennsylvania has had a centralized e-mail system in place for nine years using Microsoft Exchange. In the fall of 2008 Linda Avetta and Cindy Bendroth individually trained a group of EMCAP test users that consisted of the executive staff of the Pennsylvania Historical & Museum Commission (PMHC). In January, Ms. Avetta and Ms. Bendroth attended an Executive Staff Meeting to discuss their use of the product, pros and cons of the product, and any unresolved issues. Unfortunately, it was determined at that time that only two people in the group had used the tool more than a handful of times. They were reminded to use the EMCAP tool. Displayed as an additional mailbox on their Outlook Toolbar, many either forgot to use it, or found it too cumbersome to copy their e-mail. By March, it was evident by the number of e-mail messages available in EMCAP that only the state archivist was diligently and actively using it. Unfortunately, the executive director had only a few e-mail messages in the repository and they were the test e-mails that were used in the training demonstration when she was shown how to use the tool.

Also in March, PA received an update for the EMCAP server, but it was already known to have a few issues that needed to be resolved and it was also unclear whether Pennsylvania had the same install as Kentucky and North Carolina. PA observed quite a few oddities such as random folder numbering for e-mail copied to the EMCAP server that the other two states could not address. It was also unclear whether Pennsylvania received an upgrade or reinstall, and if there would be a potential negative impact with the e-mail messages already stored in EMCAP. EMCAP was written for Outlook 2007. During the pilot Pennsylvania was on Outlook 2003, and the updated version of the product had not been tested with 2003. In addition, PA requires detailed implementation documentation for all systems; not only to perform the initial install, but to provide sufficient reference for reinstallation in the event of server replacement, necessary restores, or in the event of staff turnover. As noted above, the installation could only be conducted via a coordinated effort with NC's technologist (because the GoToMeeting™ software could not be procured within the time constraints). The PA IT consolidation of PHMC's IT staff into the Governor's Office of Administration/Office of Information Technology (OA/OIT) also added a few levels of frustration due to resource scheduling. Without the upgrade, PA was not able to move forward or comment on subsequent versions of users guide documentation or test cases.

Pennsylvania, however, is very interested in the management of e-mail in general. The new Right-To-Know Law dramatically increased the need to properly manage e-mail and records management finally received the visibility it deserves. A group of IT and records management staff in PA, including Ms. Avetta, investigated tools the commonwealth could use to manage e-mail. Some features investigated include the categorization of e-mail, formats for storing e-mail, and the association of proper retention. Due to the economic climate (PA's budget was finally passed 5 months late in November), this project was tabled due to lack of funding for an enterprise solution. Additionally, Ms. Avetta participated with and was a voting member for the commonwealth's security domain on an enterprise-wide e-mail encryption project. Vendor demos were held early in the calendar year 2009). The need for e-mail encryption when sending sensitive content to external recipients was budgeted in the Governor's Office of Administration/Office of Information Technology last year. Ms. Avetta's participation was fortunate due to archival and records management concerns with encryption.

Ironically, the very challenge that PA had with EMCAP in the beginning of the project in seeking approval to use an open source software product (which was prohibited), has eventually led to

the OA/OIT reexamining their policy. Now staff may bring open source products before an OA/OIT board on a regular basis for possible approval. This is encouraging because it is no longer out-right prohibited but is assessed on a case-by-case basis as standard procedure.

3. Produce different types of training methodologies, including Web-based training, implementing work under development by Pennsylvania.

North Carolina

The NC team developed several different training tools on how to use the EMCAP tool. We produced several text documents with images (for the visual learners) for Microsoft 2003, Microsoft 2007, and Novell GroupWise. The documents are comprehensive in nature and include screen shots to illustrate the directions in the text. Additionally, the NC team developed video training to show users how to move their mail using MS Outlook as well as an online tutorial for users of GroupWise (http://www.records.ncdcr.gov/EmailPreservation/training.htm). The team produced the online tutorial using SoftChalk which allows users to print a certificate at the end of the successful completion of the tutorial.

Kentucky

Kentucky also produced a text document for Outlook 2003. Since KY did not have any EMCAP clients using Outlook 2007, there was no urgency for producing a 2007 training manual, since North Carolina did produce one. Apart from the EMCAP specific training, KY has continued to do e-mail management training for state and local officials throughout the grant period. Kentucky also created a PowerPoint e-mail presentation with audio attached (using Adobe flash) so that individuals can take the training online without the instructor being present.

Pennsylvania

Pennsylvania also produced a text document for Outlook 2003. Since Outlook 2007 was not available in PA during the EMCAP pilot testing, an Outlook 2007 guide was never written. Without the final version of EMCAP installed in PA, training was somewhat difficult.

4. Test ways of providing access to these XML files and the feasibility of doing so, whether through existing online catalogs, a third party vendor, or Web interfaces.

The EMCAP team was unable to address this objective. Given the different levels of IT consolidation, different levels of IT support within each of the states, and the difficulty in securing IT support because of the changing IT environments, the project team was not able to really address this objective. Changes in the political landscape, including a lawsuit against the former Governor's office for destruction of e-mail (NC), issuance of Executive Orders on E-mail Archiving and Retention (NC), and heightened awareness of roles and responsibilities do provide some measure of hope that we will be able to address this objective in the near future.

In KY's tests, some attachments saved to the file system during the conversion of e-mail messages to XML use the Transport Neutral Encapsulation file format. These files are simply containers for one or more regular file attachments and have the file name WIN.DAT. EMCAP does not extract the constituent attachments within these WIN.DAT files. While these files can be read by third party tools, some of which are free of cost, this poses an unnecessary barrier to

access. Future versions of EMCAP should extract the constituent attachments so that they can be easily listed and accessed.

Outreach

During the course of this grant, the project team conducted a number of outreach activities. We presented at state, regional, and national conferences. The team presented at the Mid-Atlantic Regional Archives Conference, the Society of North Carolina Archivists, the Society of American Archivists, and the Collaborative Electronic Records Project (CERP) end of project meeting. Halfway through the project, staff from the EMCAP project teamed with staff from the CERP project to collaborate on the development of the eXtensible Markup Language (XML) schema to parse the e-mail collected. Since the CERP team was further along in their project in February of 2008, the programmer for that project wrote a parser using Squeak based on the XML schema developed jointly. The teams built independent parsers but both worked successfully. Some of the older e-mail was problematic and could not be parsed.

In August 2008, staff from the EMCAP project teamed with staff from the CERP project to submit a workshop proposal to the Society of American Archivists titled “Achieving E-mail Account Preservation with XML”. The workshop focused on the projects, the tools developed, and the XML schema. The Society of American Archivists accepted the proposal and the workshop was offered at the annual meeting in August 2009.

The tool itself has generated tremendous interest. The Web site still receives over 300 hits a month. The workshop, picked up by SAA, has been offered four times to date—2 times at the annual meeting, at a conference sponsored by the Rocky Mountain Archives Society and the Philadelphia Museum of Art. The team has also engaged staff from the National Center for Advanced Systems and Technologies (NCAST) and from the Data Intensive Cyber Environments (DICE) group. We discussed our tool and the schema and sought their feedback. Based on those discussions, we continue to make the schema and the tool available for testing and feedback.

Lessons Learned:

This project has had a tremendous impact on all of the participating organizations. The key lesson learned is that technology can and does work. Often, the key to success is knowing the right people in the right organizations. At the beginning of this project, all three states engaged the appropriate people across government. However, as states began consolidation of Information Technology resources, many of the people who were appropriate in the beginning were not the people who would see this project through. In PA, for instance, the IT consolidation that took place there meant that the staff had to identify and engage new people who could sponsor this project. Since the project was still in the pilot phase and it involved some open-source software, PA was not able to participate as fully as they had previously committed.

The project team also learned the value of Web 2.0 tools to help foster communication and collaboration. The staff in KY provided login credentials to the NC and PA staff into their network. This allowed the project team to leverage the KY wiki tools as well as their calendaring program through MS Exchange in order to schedule meetings. The wiki tools provided space for us to collaborate on our project documents and work products.

In NC, we gained valuable skills in project and staff management. For future tool development, we will be able to build on the relationships formed and the working style developed to ensure

that projects are completed in a timely manner. The Archives staff has a much better relationship with the IT staff and has been able to develop the skill set to discuss technical work with them while also being able to explain the project in more generic terms to those who are not technical but key supporters.

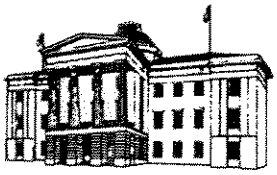
Conclusion:

The Preservation of Electronic Mail Collaboration Initiative project team is very thankful for the opportunity to work on this project provided by the National Historical Publications and Records Commission. It truly has been an exciting and exhilarating learning experience. The partners all have a better understanding of the technologies involved. Moreover, the partners have a much greater appreciation of the amount of work it takes to develop tools, work in collaboration with each other (both across state boundaries as well as the challenges faced with internal resources, priorities, and bureaucratic nuances), and push forward the deployment of an e-mail capture and preservation tool.

The EMCAP tool, developed by North Carolina and tested by Kentucky, does indeed work with different e-mail software programs. Pennsylvania tested the earlier version of the EMCAP tool but was unable to test the later product release that proved to be successful. Users in North Carolina found it relatively easy to use. Archivists, once they learned the system, found the set up of accounts and administration of them, simple and straightforward to manage. The EMCAP project will be carried forward on a limited scale in KY. The KDLA Commissioner continues to use EMCAP to store official correspondence and the previous commissioner's files are also stored in the e-mail archive. As part of the EMCAP project for test purposes we collected and are storing e-mail releases from two governor's communication offices. These are not considered archival, because they are duplicates of the official press releases on the governor's Web site. In NC, EMCAP continues to be supported both in the Governor's office and at the department level. The state did purchase a conventional e-mail archiving system but is looking to help continued development and use of EMCAP.

A major outcome of this grant in NC is that the State Archives earned a seat at the table with the IT groups in each of their states where one did not necessarily exist before. With Governor Perdue's issuance of Executive Order No. 18 (Addendum B), the electronic records archivist was named a voting member of the E-mail Archiving solution team for the selection of an e-mail archiving solution for the state of NC. It is an opportunity that would not have happened without NHPRC's funding and support of this project.

State of North Carolina



BEVERLY EAVES PERDUE
GOVERNOR

EXECUTIVE ORDER NO. 18
E-MAIL RETENTION AND ARCHIVING POLICY

WHEREAS, the North Carolina Public Records Law declares that the public records and information compiled by the agencies of North Carolina government are the property of the people; and

WHEREAS, all e-mail messages sent and received in the transaction of state business are public records; and

WHEREAS, a transparent government and the citizens' right to access public records are of paramount importance; and

WHEREAS, Governor Easley issued Executive Order Number 150, entitled E-mail Retention and Archiving, on January 9, 2009; and

WHEREAS, I have reviewed Executive Order Number 150 and determined that some of the provisions in the aforementioned order should be clarified.

NOW, THEREFORE, pursuant to the authority vested in me as Governor by the Constitution and the laws of the State of North Carolina, **IT IS ORDERED THAT:**

RESCISSION

1. Executive Order Number 150, dated January 9, 2009, is hereby rescinded.

EMPLOYEE RESPONSIBILITIES

2. Executive Branch employees shall treat all e-mail messages which they send or receive via state government e-mail accounts as public records and shall handle and maintain them in compliance with the Public Records Law and records retention schedules in the same manner as paper documents or other tangible records.

3. Employees have no expectation of privacy in their electronic correspondence, and all employees shall assume that information on the State's e-mail system is subject to public review and to review by state officials.

4. All outgoing e-mails sent from Executive Branch State e-mail accounts shall include language notifying the recipient(s) that the message is subject to the Public Records Law and may be disclosed to third parties.

5. Executive Branch employees shall not permanently delete any e-mail messages that they **send** for at least 24 hours, and shall not permanently delete any e-mail messages they **receive** for at least 24 hours except that they may immediately and permanently delete any e-mail messages they **receive** that are not clearly related to the transaction of State business, such as e-mails containing advertising materials or offensive materials. After 24 hours, Executive Branch employees shall retain or delete e-mails they have sent or received according to the retention schedules for their agency established by the Department of Cultural Resources.

6. Executive Branch employees who conduct State business via personal e-mail accounts shall ensure that all public records are retained in accordance with this Executive Order and are retained pursuant to the Public Records Law and applicable record retention schedules.

7. Executive Branch employees shall not use State e-mail accounts for political purposes, to conduct private commercial transactions or to engage in private business activities. Executive Branch employees may use State e-mail for limited family or personal communications so long as those communications do not interfere with their work.

AGENCY RESPONSIBILITIES

8. All Executive Branch agencies shall copy all e-mails sent and received by their employees on backup tapes at least once daily. The Office of Information Technology Services (ITS) will provide this backup service to all agencies for which it provides e-mail services. Each Executive Branch agency that does not use ITS e-mail services shall employ a back-up system that creates a back-up copy of the messages in all e-mail systems of the agency at least once daily. All backup tapes created after the issuance of Executive Order 150 and prior to the implementation of a single e-mail archive system will be maintained for 10 years. After implementation of an e-mail archive system, backup tapes will be maintained for such period as ITS may establish.

9. ITS will procure an e-mail archive system as soon as practicable and provide that system to all agencies for which it provides e-mail services. ITS will make this archive system available to other Executive Branch agencies as soon as practicable. E-mails shall be retained in this system for 10 years. ITS will consult with the North Carolina Department of Cultural Resources (DCR) to identify e-mails that should be preserved beyond 10 years.

10. DCR shall provide Executive Branch employees with mandatory online training for managing e-mail as public records.

11. DCR shall conduct random audits of State agencies in the Executive Branch to ensure that employees are in compliance with the records retention and disposition schedules.

12. Executive Branch agencies not subject to this Order, the Legislative Branch and the Judicial Branch, are encouraged and invited to participate in this Executive Order.

DURATION

13. This Executive Order shall be effective immediately and shall remain in effect until rescinded.

IN WITNESS WHEREOF, I have hereunto signed my name and affixed the Great Seal of the State of North Carolina at the Capitol in the City of Raleigh, this the seventh day of July in the year of our Lord two thousand and nine, and of the Independence of the United States of America the two hundred and thirty-fourth.

Beverly Eaves Perdue
Governor

ATTEST:

Elaine F. Marshall
Secretary of State

EMCAP Pilot

Version 1.5 Installation Guide

North Carolina Department of Cultural Resources

Government Records Branch

E-mail Capture and Preservation Tool Grant

June 2009

Contents

Purpose for the Guideline 3

Brief Overview of EMCAP 3

Requirements Prior to Installation 3

EMCAP Installation Instructions 5

EMCAP Post Installation Steps 19

Purpose for the Guideline

This guideline provides detailed instructions for installing EMCAP, the E-mail Capture and Preservation software. For more information about how this software functions and its design considerations as well as the project's background, please see our web site at http://www.records.ncdcr.gov/emailpreservation

Brief Overview of EMCAP

EMCAP runs on the Windows Server platform and requires Microsoft SQL Server. EMCAP is at its core an IMAP accessible e-mail store where the store is implemented using a set of XML document, one for each account. Each message is stored as real XML in the actual XML document. The body of each message and its attachments can also be stored within the XML document, but are often stored as external files referenced by the XML document. In addition, a set of industry standard mbox files are kept for each account, one per folder.

Since the amount of work required to convert e-mail messages in their native *Internet Message Format* to XML can be large, the conversion is achieved through a batch process that can be scheduled to run nightly.

Requirements Prior to Installation

Prior to the installation of EMCAP, the server must be configured properly. This will ensure the appropriate environment exists to undertake the installations and then run the installed programs effectively and efficiently. It will be necessary to ask your IT department to set up the proper environment.

Operating System	Recommended: Windows Server 2003 Service Pack 2 (32-bit x86) (http://www.microsoft.com/downloads/details.aspx?FamilyId=95AC1610-C232-4644-B828-C55EEC605D55&displaylang=en) When you install a new service it's a good rule of thumb to install any critical Windows updates to make sure that the operating system is up to date. This is something you would get from Microsoft. In addition, make sure that the server has some type of RAID (Redundant Array of Independent Disks) storage. Note: Windows XP is an operating system that is not quite up to the task of running a server service which you can rely on like EMCAP. It is recommended that this OS not be used.
Prerequisite Software	• Microsoft SQL Server or Express, 2005 or later. The compact version is not recommended. (http://www.microsoft.com/Sqlserver/2005/en/us/express.aspx) • Microsoft SQL Server Management Studio (Express or full version). The full version is recommended.

	• Microsoft .NET Framework 2.0 (http://www.microsoft.com/downloads/details.aspx?FamilyID=0856EACB-4362-4B0D-8EDD-AAB15C5E04F5&displaylang=en). • Microsoft Core XML Services (MSXML) 6.0 Service Pack 1 (http://www.microsoft.com/DOWNLOADS/details.aspx?FamilyID=d21c292c-368b-4ce1-9dab-3e9827b70604&displaylang=en) • IIS (Internet Information Services) 5.1 or better. Using Add or Remove Programs, select the Add/Remove Windows Components icon. Make sure that the Internet Information Services (IIS) is installed. The typical installation settings are fine.
Internet Requirements	Microsoft Internet Explorer 6.0 SP1 or later (prerequisite for .NET Framework)
RAM	At least 2 GB of system RAM, 4 GB preferred
Hard Disk Space	At least 20 GB of free disk space. Note: The amount of disk space required depends on: • the number of users for whom you will be collecting archival e-mail messages, • the number of years that you want to store this mail at the collection point, • the average number of messages each user will be archiving, • the average size of each message being archived. Typically this will be 200-500 MB per user, per year.
Processor	2.0 GHz Pentium 4 processor or better.
Domain Name	In order for an e-mail client to connect to the archival server, the archival server must be reachable via the World Wide Web or your local area network by its DNS (Domain Name System). The archival server's Internet Protocol (IP) address must be registered with your local DNS servers or with the set of public DNS servers that are used with the World Wide Web.

EMCAP Installation Instructions

Step 1: Download the EMCAP Version 1.5 .zip file from the EMCAP website and extract the files to a corresponding folder on the computer's desktop. Double click setup to start the installation.

Step 2: The Welcome to the Emcap Installer Setup Wizard appears. Click **Next**.

Step 3: On the next screen, you will be asked to confirm the folder where you want to install the Emcap Installer. Make note of this location. You can keep the default or click the browse button to select a new location. At the bottom of the screen, choose either "Everyone" if you want the Emcap installer for anyone who uses the computer or "Just Me" if you're installing Emcap for yourself. Click **Next**.

Step 4: Click **Next** to start the installation of the Emcap Installer.

Step 5: The next screen will confirm success of the installation. Click **Close** at the bottom of the screen to exit this wizard.

Step 6: The MainForm of the Emcap Installer will appear. You will need to follow the order of Emcap Installer for a successful installation of version 1.5. To begin, select the **1. Install hMailServer** button.

Note: You may see a message indicating that the hMailServer Setup has detected another email server. This is most likely due to the fact that the Microsoft SMTP server was installed as part of IIS, therefore this message can be safely ignored. Click **OK**.

It is strongly recommended that the EMCAP software does not use the underlying hMailServer for its SMTP service (it is very likely that your institution already has such a server as part of its normal e-mail service). Although the underlying hMailServer's SMTP service could be used, the best practice is to disable this service to eliminate the chance that any of your customers will configure their e-mail clients in such a way as to enable the sending of mail from their archival account.

Step 7: The Welcome to the hMailServer Setup Wizard will appear. This screen verifies that you are about to install hMailServer 5.0-B 326 on your computer. Click **Next**.

Step 8: The next screen will ask whether or not you accept the hMailServer license terms. Select "I accept the agreement" and then click **Next**.

Step 9: On the Select Destination Location page (Figure 1), you can keep the default location or select **Browse** for another location to install the hMailServer. Make note of this location. Click **Next**.

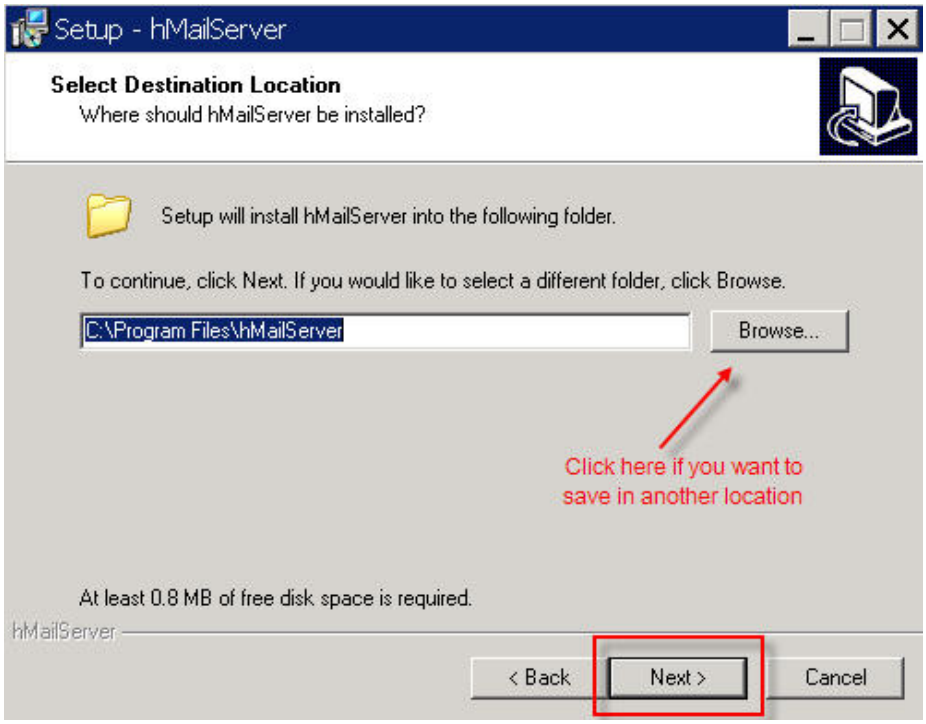
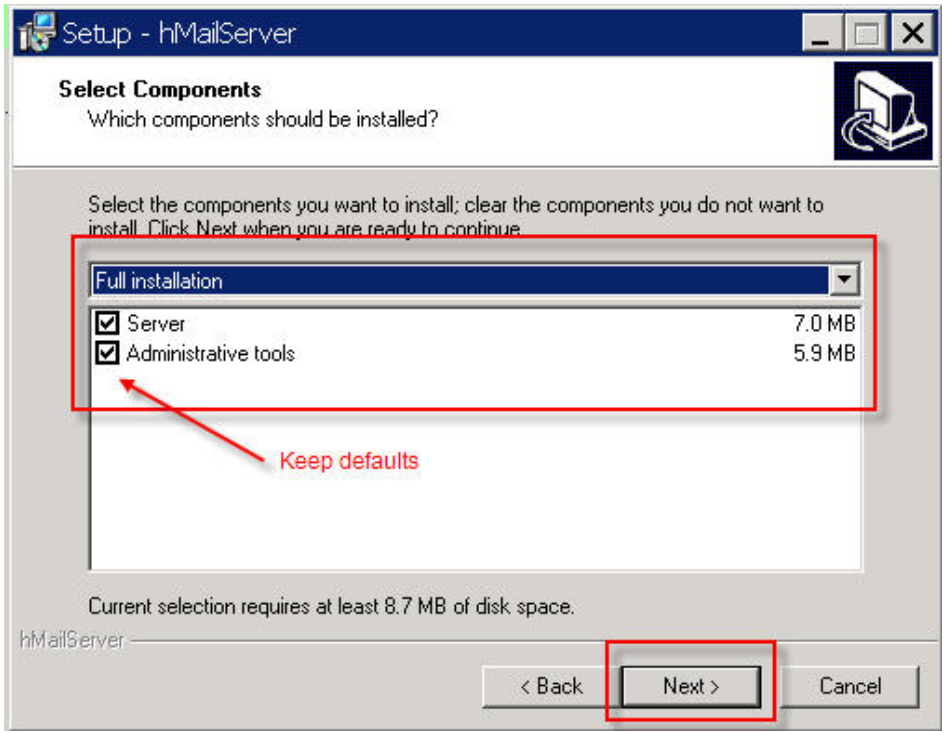


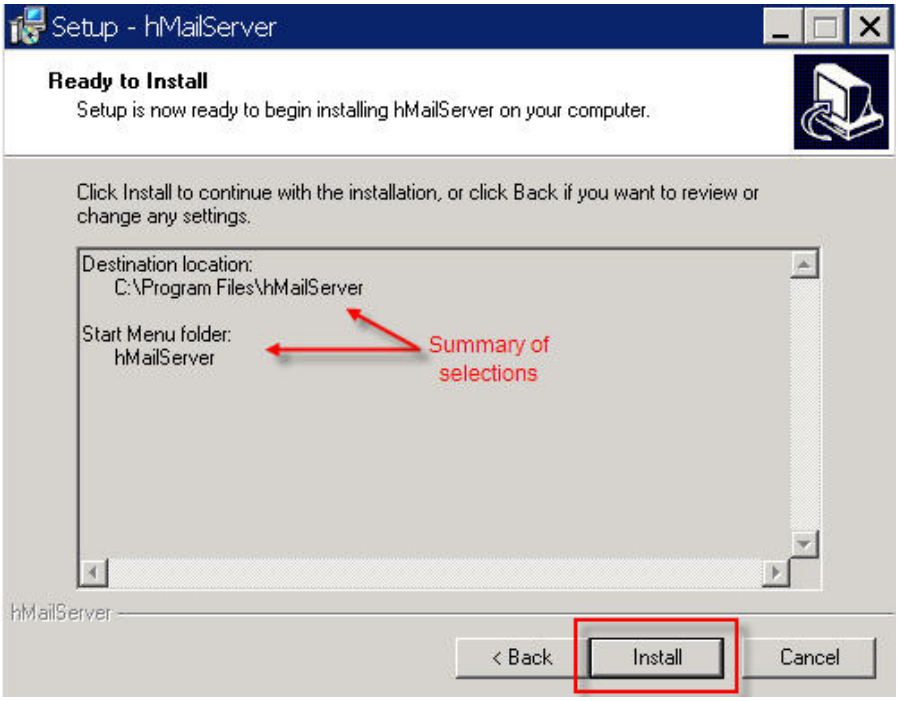
Figure 1.

Step 10: On the next screen, select the components you wish to install. The Server and Administrative tools are default selections. Keep these defaults. Click **Next**.



Step 11: On the Select Start Menu Folder page, use the default Start Menu Folder “hMailServer”. At the bottom of the screen, keep “Don’t create a Start Menu folder” unchecked. Click **Next**.

Step 12: The Ready to Install page appears. This essentially is a review of your selections. If you are satisfied with these selections, click the **Install** button.



Step 13: Prior to the finalizing the installation, you will be prompted to enter a hMailServer Security password. Make note of this password. Re-enter the password to confirm and click **Next**.

Step 14: On the Completing the hMailServer Setup Wizard, uncheck “Run hMailServer Administrator” and then click **Finish**.

Step 15: A Database Setup window will appear (Figure 2), which will walk you through the steps of setting up the database. Click **OK**.

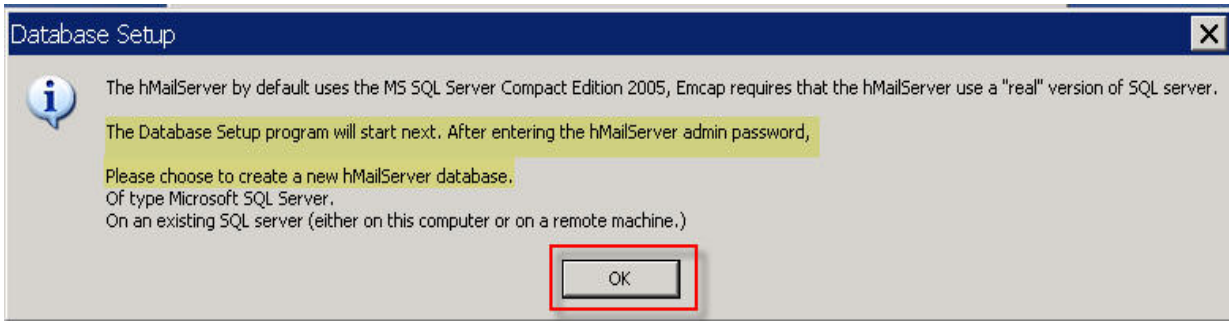
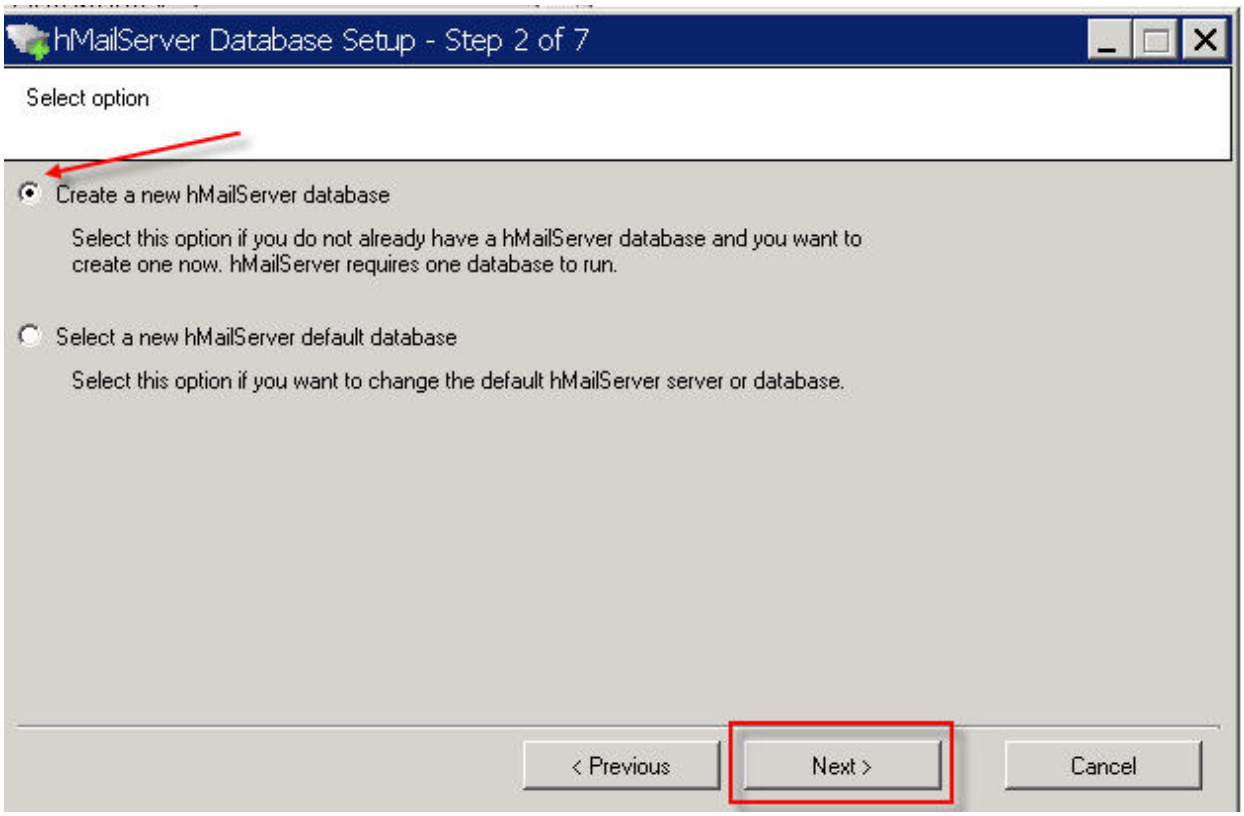


Figure 2.

Step 16: Enter the password you created in Step 13 and click **OK**.

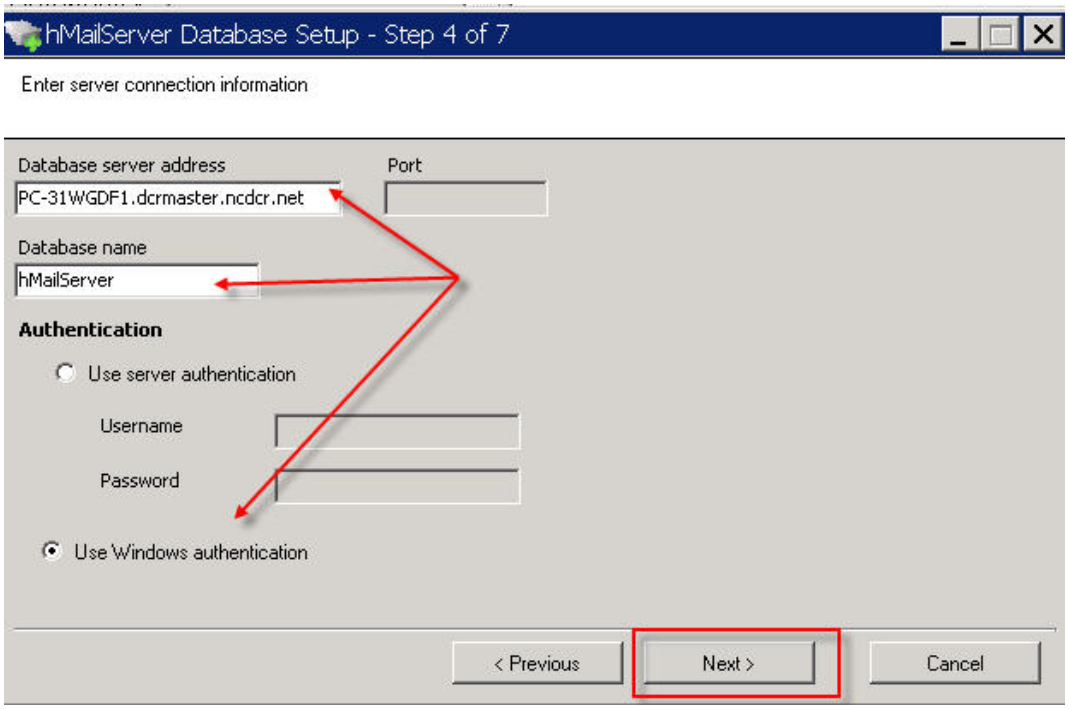
Step 17: The hMailServer Database Setup wizard will appear. Using this wizard, you can create a new hMailServer database and select a new default database. Click **Next**.

Step 18: On the next page, select “Create a new hMailServer database”, and then click **Next**.



Step 19: On the next page, keep the default “Microsoft SQL Server” as the database server type. Click **Next**.

Step 20: On the next page, type in the full DNS name in the Database server address field. For Database name use “hMailServer” and then check “Use Windows authentication”. Click **Next**.



Step 21: On the next screen (Figure 3) you will need to select the hMailServer dependency from the drop down list. Select the first in the list that begins with “SQL Server”. Click **Next**.

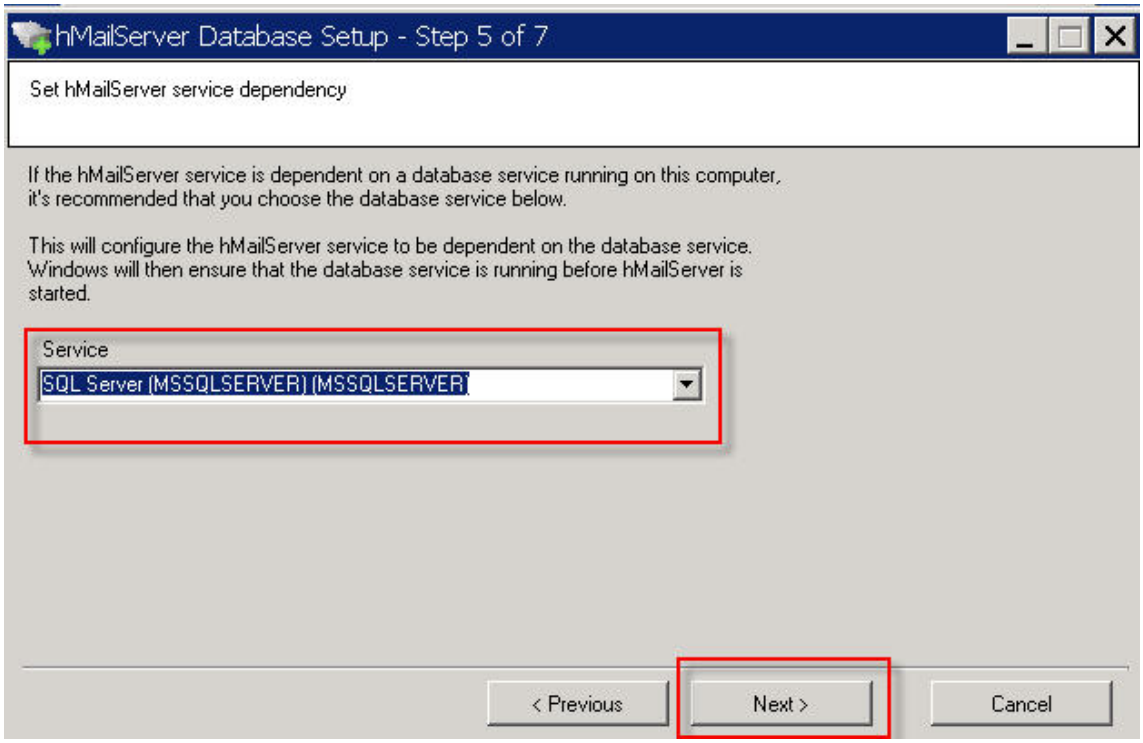


Figure 3.

Step 22: On the next screen, click **Next** to perform the scheduled tasks. Select **Close** to exit the wizard.

Step 23: Back on the main form of the Emcap Installer, select the **2. Modify hMailServer** button.

Step 24: On the Modify hMailServer window (Figure 4), check “Use Windows Integrated Security”, the hMailServer Administrator Password, and the select **Make Modifications**. A success pop-up window will appear. Click **OK**.

Figure 4.

Step 25: Back on the main form of the Emcap Installer, select the **3. Create Emcap Database** button.

Step 26: The Create/Update EMCAP Database window will appear (Figure 5). From the drop down box in the Install database in folder field, select the appropriate folder. Keep all other defaults on the window and select **Create Database.** You must choose a location to store the SQL files.

The database can be installed on a remote server; however the remote server must have a folder named either C:\SQLData or D:\SQLData. If neither of these folders is the desired final location, the database can be moved to another location on the local server via the SQL Management Studio by detaching the database, moving the data and log files and then reattaching. The database can be moved to another SQL server similarly, but the connection strings must be updated to reflect the new location. For information on how to update the connection strings see the last section of this document entitled “Troubleshooting."

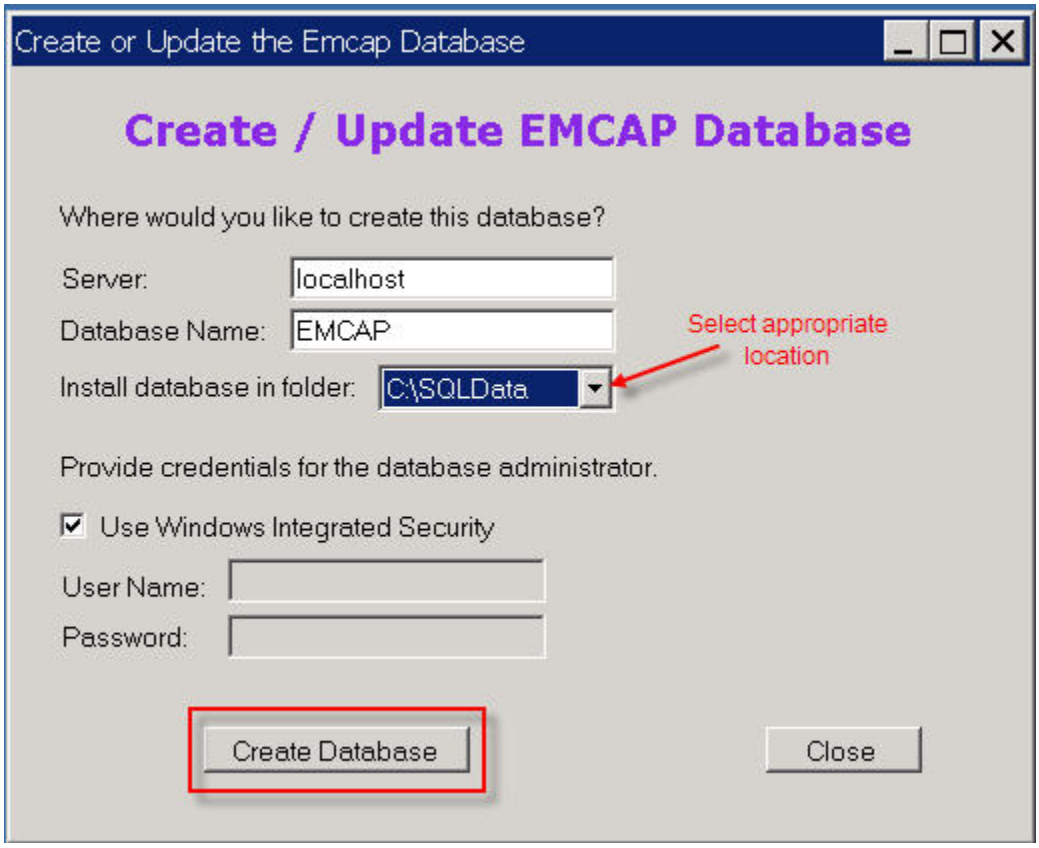


Figure 5.

Step 27: A pop-up will appear verifying that the EMCAP database has been created will appear. Click **OK**.

Step 28: A success pop-up will appear verifying that the EMCAP database has been updated. Click **OK**. Click the **Close** button on the Create /Update EMCAP Database window.

Step 29: Back on the main form of the Emcap Installer, select the **4. Choose SMTP Settings** button.

Step 30: The Send Mail Settings window will appear (Figure 6). Enter the name or IP address of the server to which SMTP traffic will be routed. The Account Name and Account Password fields can be left blank if the SMTP server does not require a user account for authentication. In the Sender's Address field, type in the e-mail address of the primary contact for EMCAP. The Recipient's address will automatically be populated; however you can change this if necessary. Click **Save** to save the information entered.

Figure 6.

Step 31: Back on the main form of the Emcap Installer, select the **5. Install Emcap Software** button.

Step 32: The Install Emcap Software window will appear (Figure 7). You can keep all defaults on this; however you will need to type in the hMailServer Administrator Password you previously created. Select **Install** at the bottom of the window.

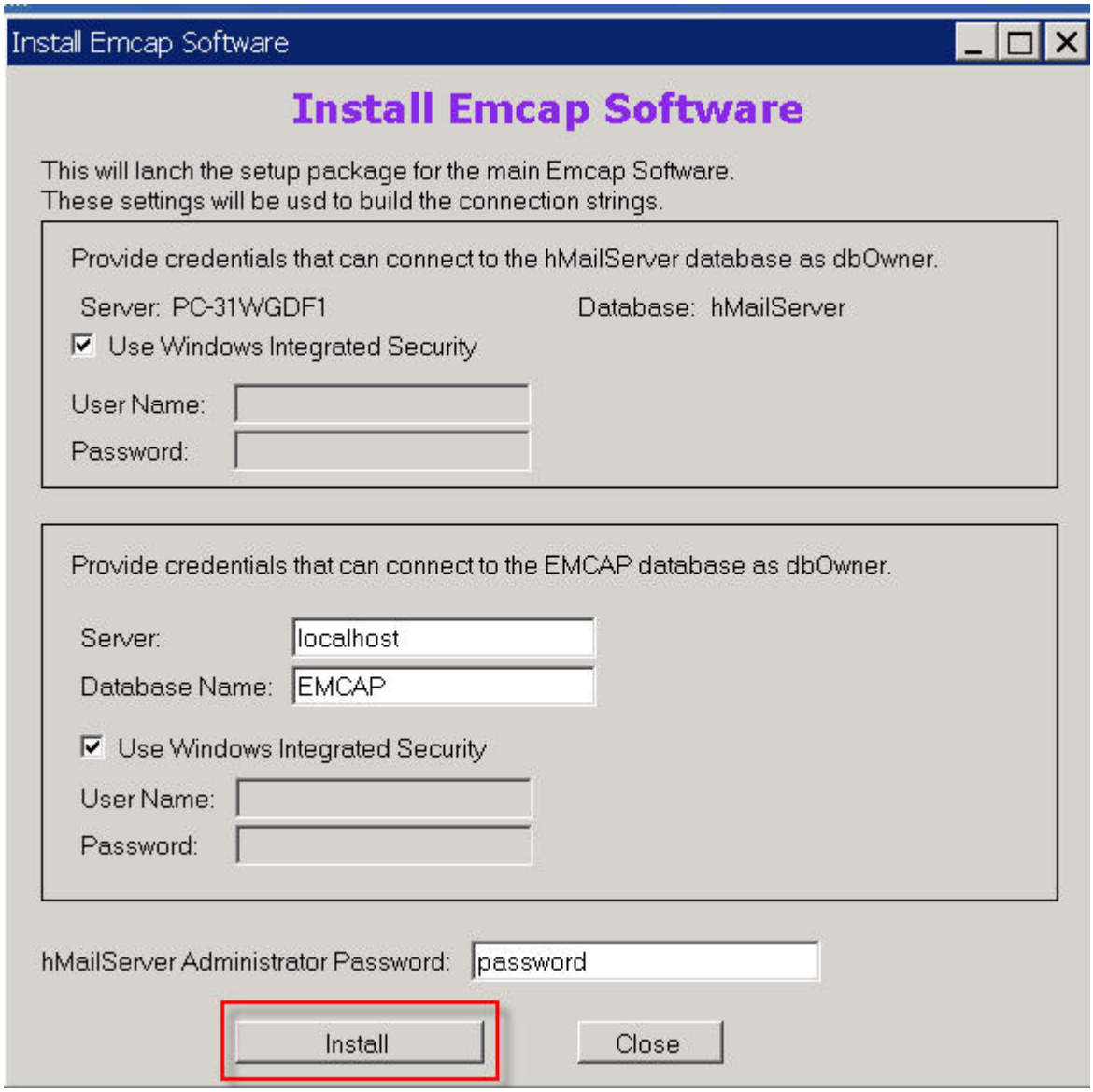


Figure 7.

Step 33: The Emcap Setup Wizard will appear, which will guide you through the steps required to install Emcap on your computer. Click **Next**.

Step 34: On the Select Installation Folder, you can keep the default location or select **Browse** for another location to install Emcap. Make note of this location. At the bottom of the screen, choose either "Everyone" if you want Emcap installed for anyone who uses the computer or "Just Me" if you're installing Emcap for yourself. Click **Next**.

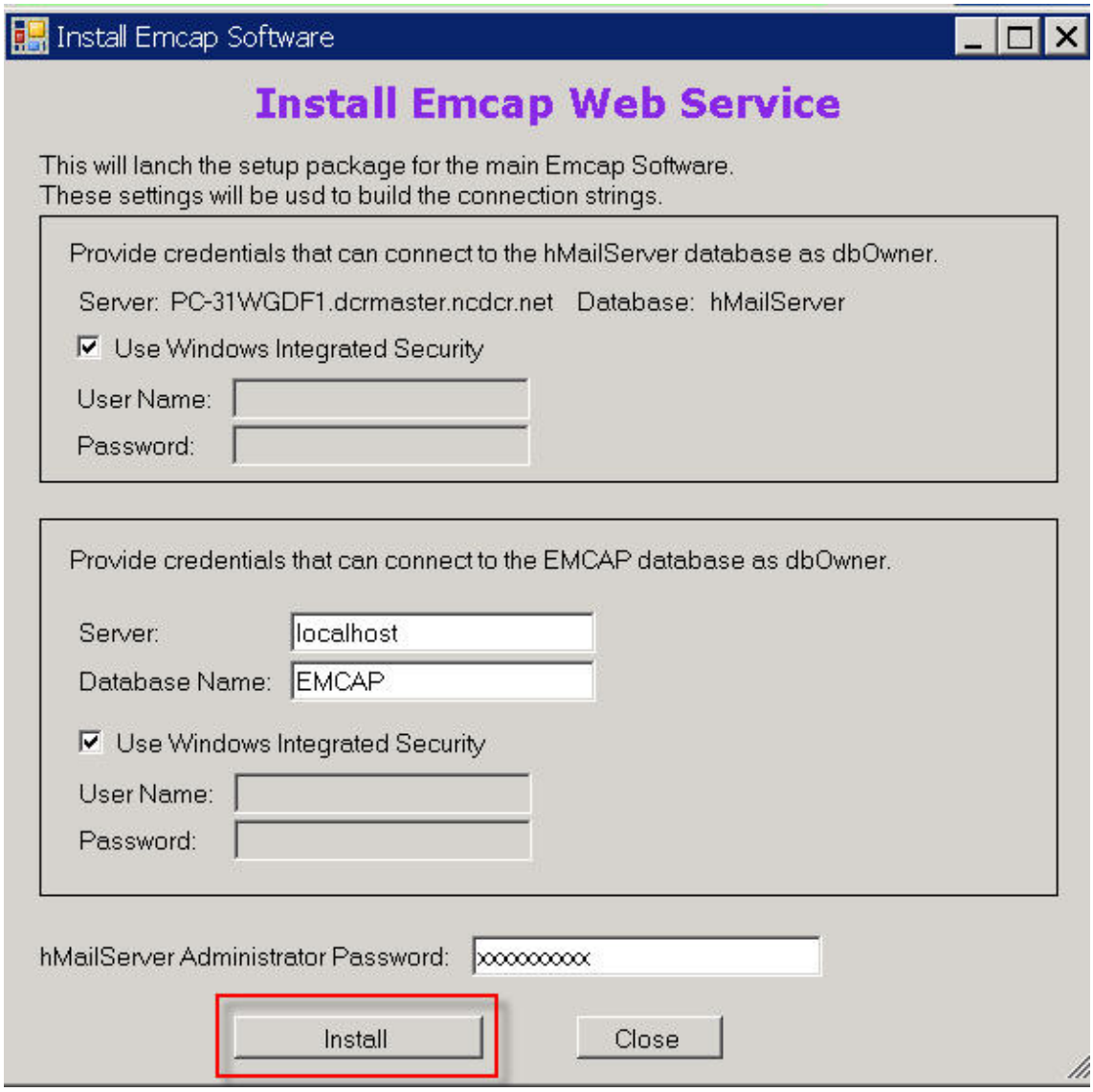
Step 35: Click **Next** to start the installation.

Step 36: An Installation Complete window will verify that Emcap has been successfully installed. Click **Close** to exit the wizard.

Step 37: An additional pop-up window will appear verifiying the success of the installation. Click **OK**. You will see an **Admin Console** icon on your desktop.

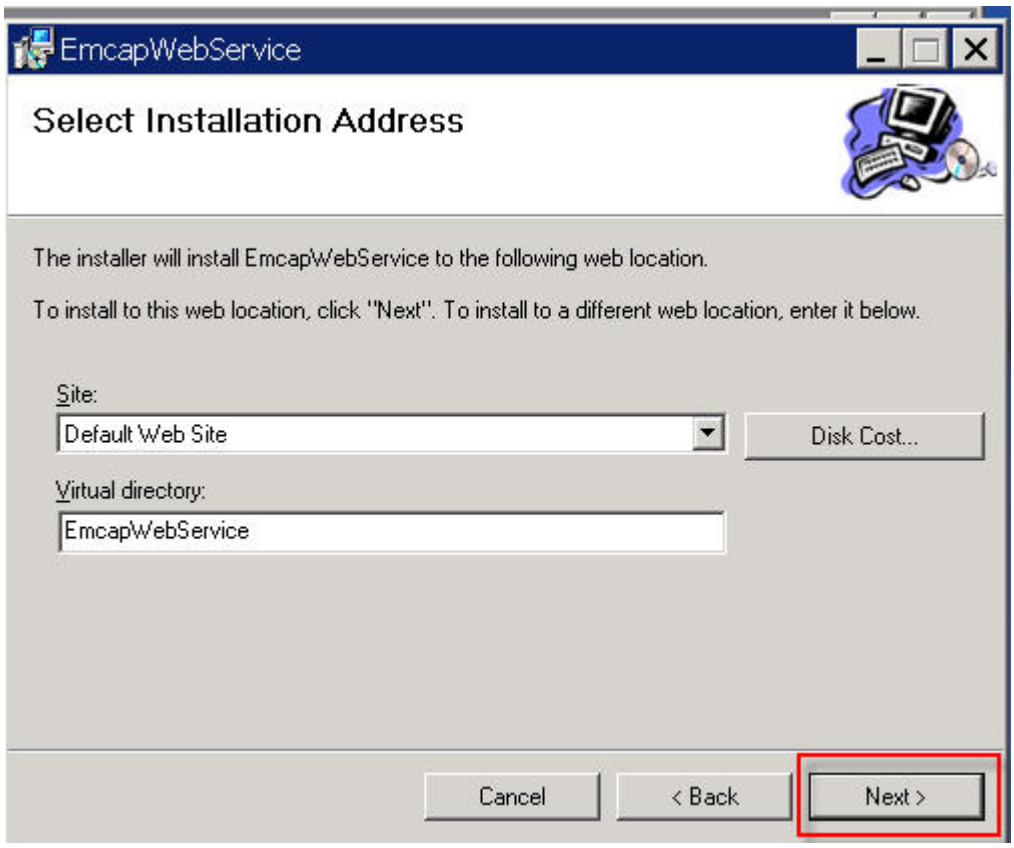
Step 38: Back on the main form of the Emcap Installer, select the **6. Install Web Service** button.

Step 39: The Install Emcap Web Service window will appear, which launches the setup package for the main Emcap software. Click **Install**.



Step 40: The EmcapWebService Setup Wizard will appear, which will guide you through the steps required to install EmcapWebService on your computer. Click **Next**.

Step 41: On the Select Installation Address screen, keep all the defaults and click **Next**.



Step 42: Click **Next** to start the installation.

Step 43: The Installation Complete screen verifies that the EmcapWebService has been successfully installed. Click **Close** to exit the wizard.

Step 44: An additional pop-up window will appear verifying the success of the installation. Click **OK**.

Step 45: Back on the main form of the Emcap Installer, select the final step **7. Install Web Site** button.

Step 46: A pop-up window will appear, verifying whether or not to install the web site on public facing server or the server you are currently using (Figure 8). Select **Yes** or **No** depending on your setup. If you decide to install the Web Site on a different server, you will need to install the Emcap Installer on that server and only perform **Step 7**.

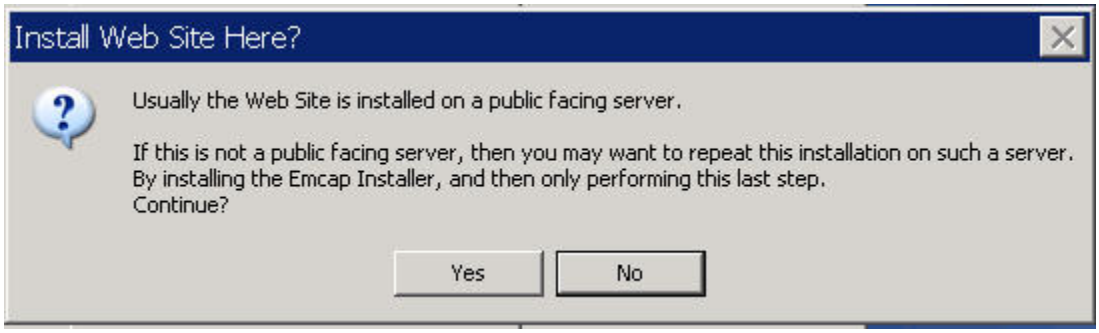
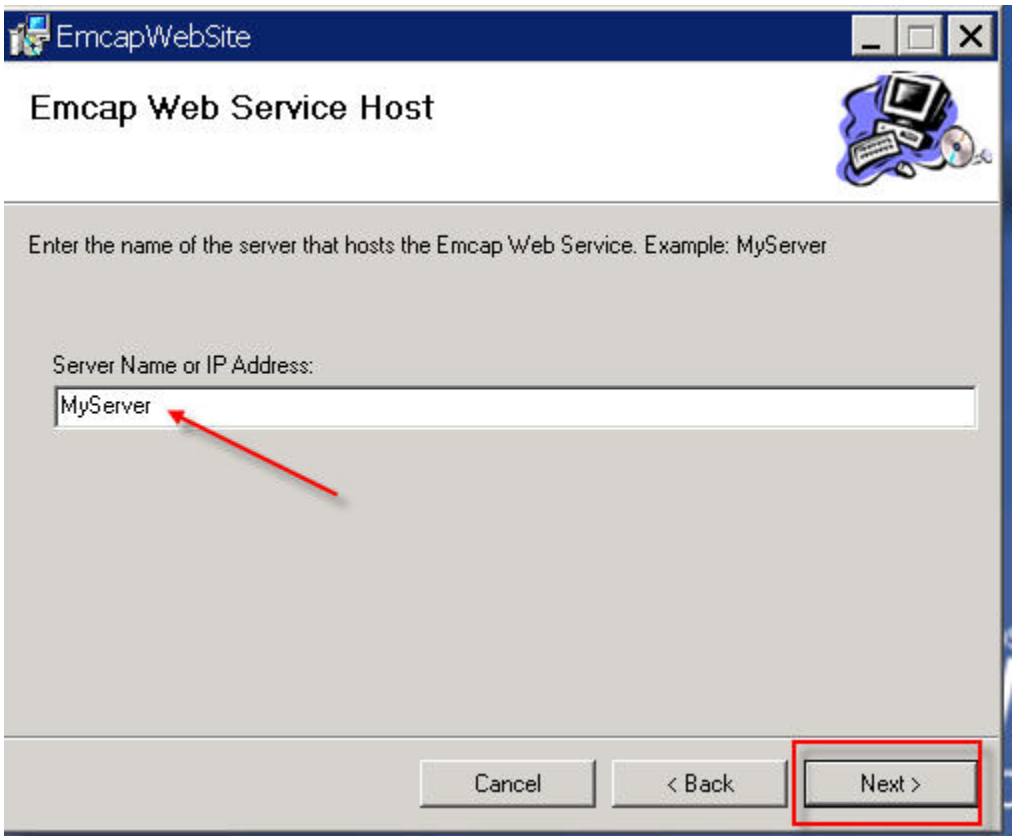


Figure 8.

Step 47: If you selected **Yes**, an EmcapWebSite Setup Wizard will appear, which will guide you through the steps required to install EmcapWebSite on your computer. Click **Next**.

Step 48: On the Select Installation Address screen, keep all the defaults and click **Next**.

Step 49: On the next screen, enter the name of the server or IP address of the server that will be hosting the Emcap Web Service. Click **Next**.



Step 50: On the next page, enter your Institution Name. This Institution Name will appear on several of the web pages. Click **Next**.

Step 51: Click **Next** to start the installation.

Step 52: The Installation Complete screen verifies that the EmcapWebSite has been successfully installed. Click **Close** to exit the wizard.

Step 53: An additional pop-up window will appear verifying the success of the installation. Click **OK**.

Step 54: You are now finished with installing Emcap. To exit the Emcap Installer main form, click the **Close** button.

EMCAP Post Installation Steps

Part 1: Configure the hMailServer

The hmailserver will work just fine “out of the box” but there are a couple of steps that can be taken to make it more secure.

Step 1: Start the hMailServer administration program.

Step 2: Select "Protocols" from the left panel, and unchecked both SMTP Server and POP3 Server, and pressed the save button. (This leaves only the IMAP protocol.)

Step 3: Select "Status" from the left panel, then stop and restart the server.

Step 4: Close the administration program.

The hMailServer supports Secure Sockets Layer (SSL) which will protect the mail sent to, and read from, the EMCAP server from anyone that may attempt to “sniff the wire.” For more information please see the documentation on the hMailServer’s web site. http://www.hmailserver.com/documentation/?page=feature_ssl

Part 2: Configure the EmcapWebService and EmcapWebSites

Until these steps are completed, the Emcap WebSite will not function correctly. All other parts of the Emcap server are ready and can be used before these steps have been completed.

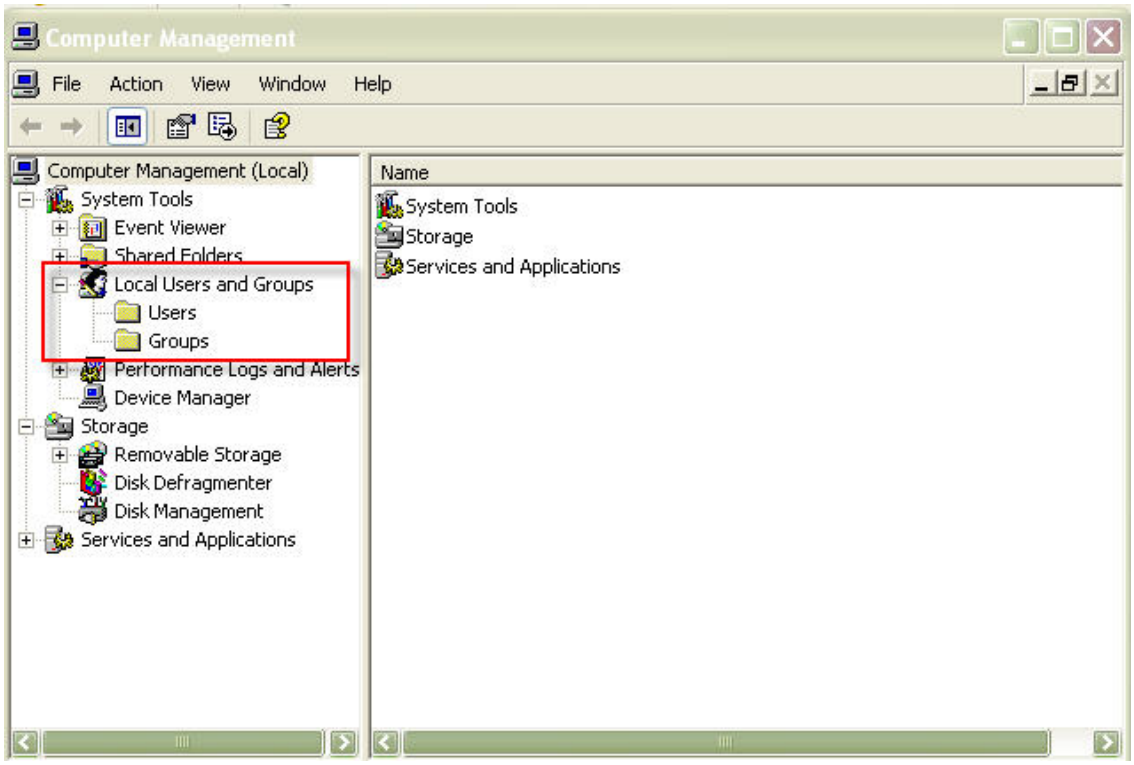
These instructions assume that both the web service and the web site are on the same computer. If this is not the case then you must make sure that the web site is configured to use an account that can be authenticated by the web service. If the database is on a different server than the web service then additional care must be

taken to make sure that the web service uses an account that can be authenticated by the database. If these servers share a common Active Directory domain, then you should probably use domain accounts instead of local accounts.

Step 1: Create a new local user account on the server named EmcapWebUser. This will be the user account that the User Support Web Site will use when logging on to the Web Service.

To do this, from the desktop, select the **Start** button → **Control Panel** → **Administrative Tools** → **Computer Management.**

The Computer Management window will appear. Expand **Local Users and Groups,** and highlight the **User** folder.



From the toolbar, select **Action** → **New User**.

A New User window will appear (Figure 9). Type "EmcapWebUser" in both the *User Name* and *Full Name* fields. Enter a description if needed. Type in a password and confirm it in the next field. **Uncheck** "User must change password at next logon" and then **check** both "User cannot change password" and "Password never expires". Click **Create** to save the user account.

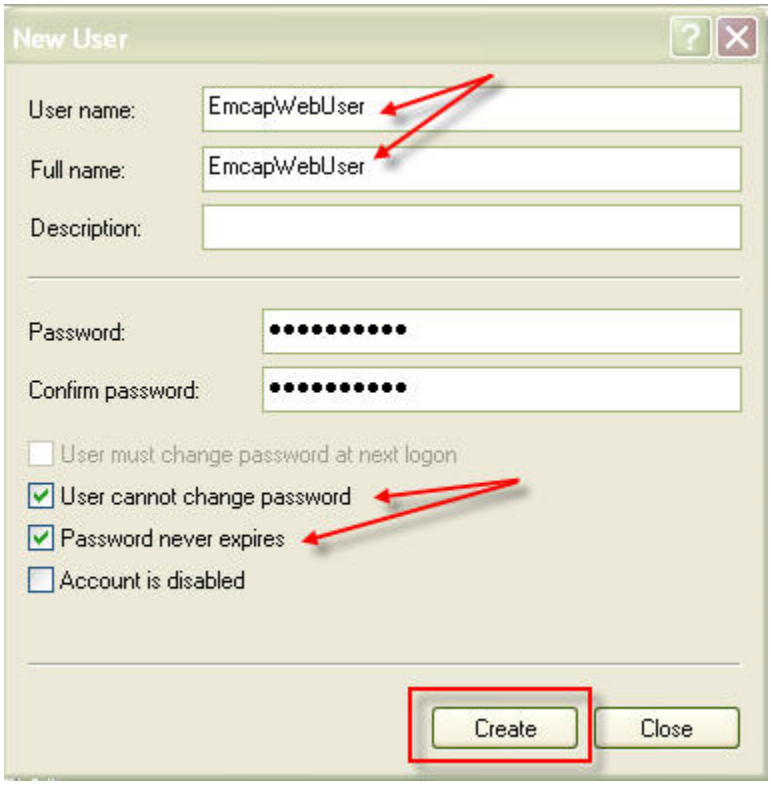


Figure 9.

Note: if the User Support Web Site is installed on a "front-end" server, then this user should be created on both servers. If both servers are part of the same domain, then this should be a domain user account.

Step 2: The New User window will be cleared so that another new account can be created. Create a new local user account on the server named EmcapUser. This will be the user account that the Web Service will use when logging on to the SQL database.

Type "EmcapUser" in both the User Name and Full Name fields. Enter a description if needed. Type in a password and confirm it in the next field. **Uncheck** "User must change password at next logon" and then **check** both "User cannot change password" and "Password never expires". Click **Create** to save the user account. Click **Close** to exit this window.

Exit out of the Computer Management window.

Step 3: To give the EmcapUser the ability to access the temporary asp.net folder, from the desktop, select the **Start** button → **Run.** In the Open field, type *cmd* and click **OK** to bring up a DOS window. At the DOS command prompt, **type** in the following command:

C:\Windows\Microsoft.net\Framework\v2.0.50727\aspnet_regiis –ga EmcapUser, and then press **Enter**.

If successful, you will receive the following message:

"Start granting emcapuser access to the IIS metabase and other directories used by ASP.NET. Finished granting emcapuser access to the IIS metabase and other directories used by ASP.NET".

Step 4: To give the EmcapWebUser the ability to access the temporary asp.net folder, at the DOS command prompt **type** in the following command:

C:\Windows\Microsoft.net\Framework\v2.0.50727\aspnet_regiis -ga EmcapWebUser, and then press **Enter**.

If successful, you will receive the following message:

"Start granting EmcapWebUser access to the IIS metabase and other directories used by ASP.NET. Finished granting EmcapWebUser access to the IIS metabase and other directories used by ASP.NET".

Note that this should be performed on the server on which the User Support Web Site was installed.

Step 5: To give the Network Service the ability to access the temporary asp.net folder, at the DOS command prompt **type** in the following command:

C:\Windows\Microsoft.net\Framework\v2.0.50727\aspnet_regiis –ga NetworkService, and then press **Enter**.

If successful, you will receive the following message:

"Start granting NetworkService access to the IIS metabase and other directories used by ASP.NET. Finished granting NetworkService access to the IIS metabase and other directories used by ASP.NET".

Note that this should be performed on both the main server and the server on which the User Support Web Site was installed if they are not one and the same

Close the DOS window.

Step 6: In SQL Server Management Studio (Figure 10), create a new login for the new user by opening the **Security** folder → **Logins** folder. Right click on Logins and select **New Login**.

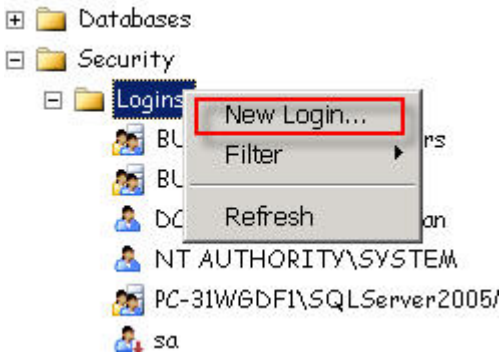


Figure 10.

A Login – New window will appear. Type in "<<server name>>\EmcapUser" in the Login Name field, and set the user's default database to be EMCAP.

On the left panel, select **User Mapping**. This will bring up the user mapping page (Figure 11). On this page:

- Check "Map" for *EMCAP* on the top screen.
- Check db_owner on the bottom screen.
- Check "Map" for *hMailServer* on the top screen.
- Check public on the bottom of the screen.

Click **OK** to save.

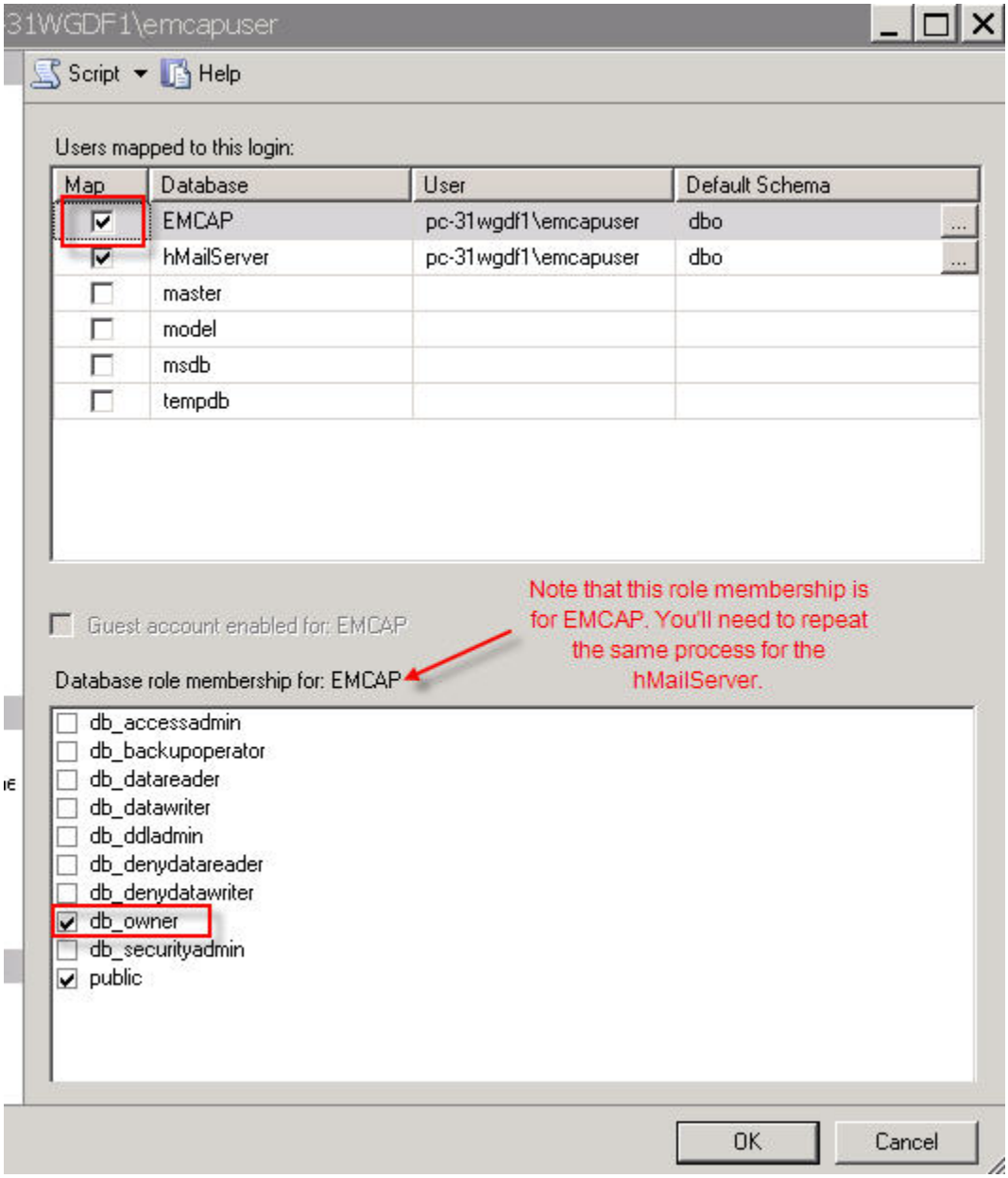


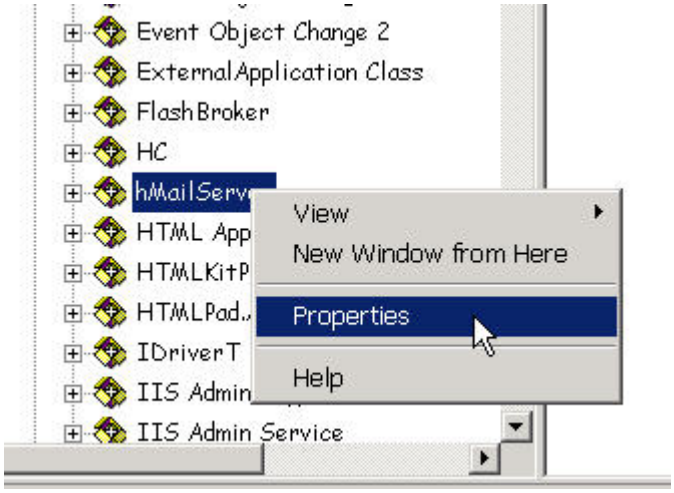
Figure 11.

Step 7: Next you will need to give the Network User permission to launch the hMailServer COM api. If the IIS Application Pool that the EmcapWebService uses runs under an identity other than Network Server, then change the following directions to use that user instead of the Network Service account.

From the desktop, select the **Start** button → **Run**. At the command line type: *dcomcnfg* and click **OK**. This will bring up the Component Services console.

Expand **Components Services** → **Computers** → **My Computer** → **DCOM Config**.

Find hMailServer. Right click on the icon and select **Properties**.



This will bring up the hMailServer Properties window (Figure 12).

Select the Security tab, and enter the following information

- For Launch and Activation Permissions, select “Customize” and press the **Edit** button. A Launch Permission window will appear. Select the **Add** button to add the Network Services user, and allow it Local Launch and Local Activation permissions. Click **OK**.

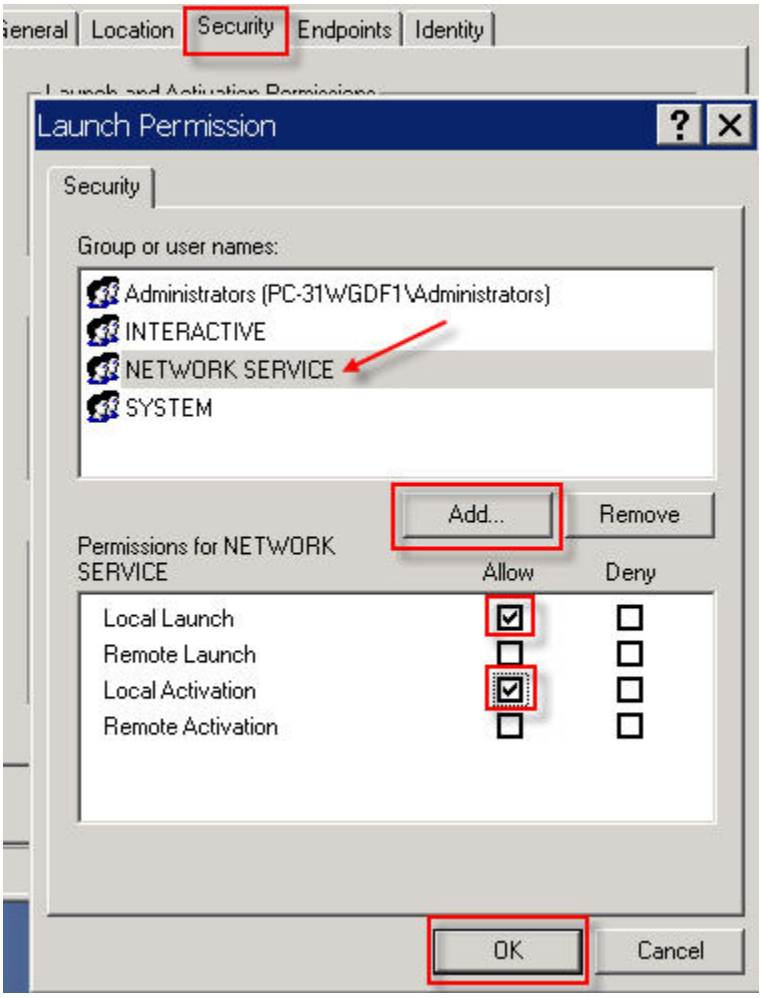


Figure 12.

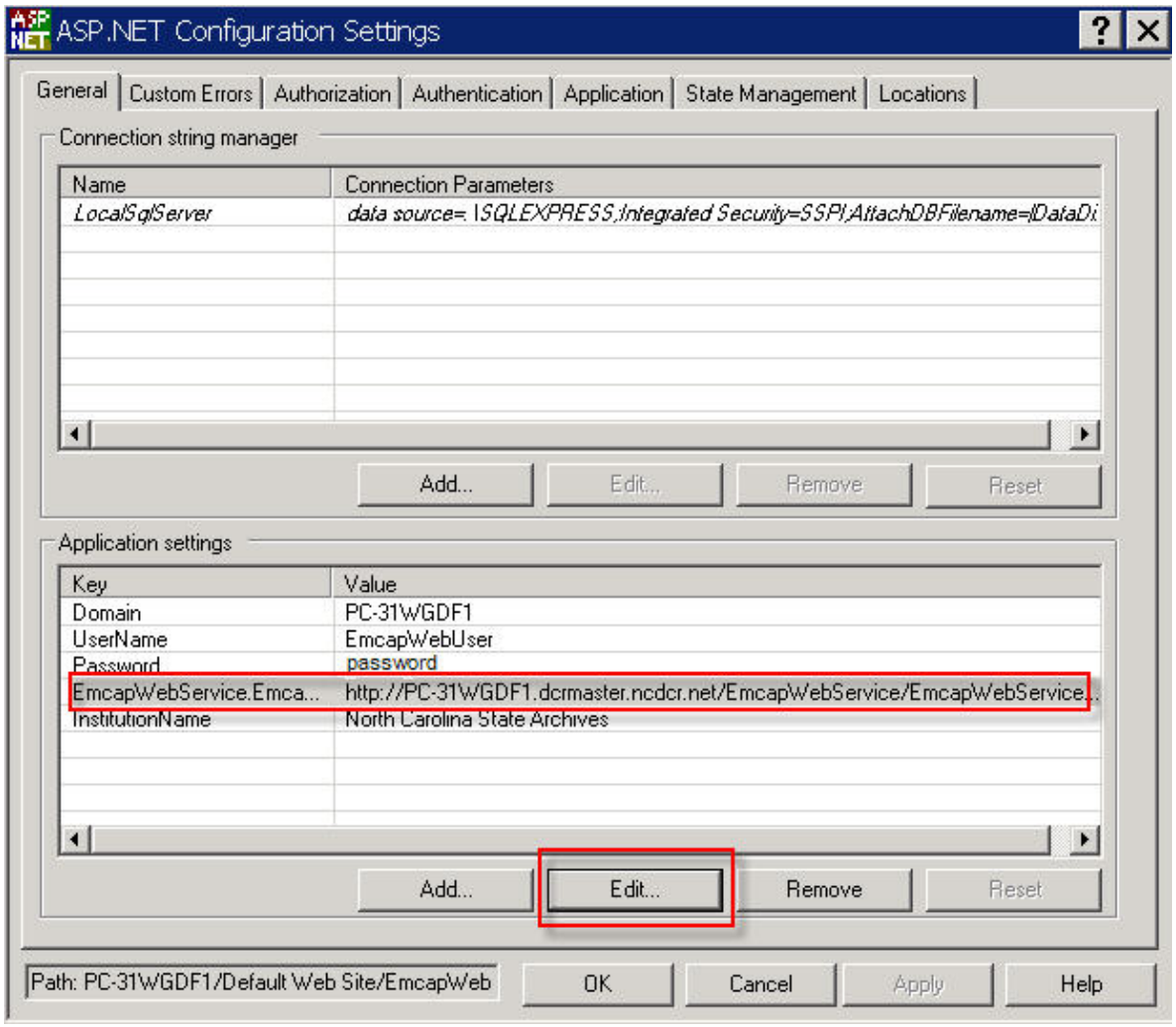
- For Access Permissions, select "Customize" and press the **Edit** button. An Access Permission window will appear. Select the **Add** button to add the Network Services user. Give it Local Access only. Click **OK**.

- Configuration Permissions will be defaulted to "Customize". Keep this default. Click **Apply** and then **OK**. Exit the Component Services widow.

Step 8: Open the Internet Information Services (IIS) and select **Web Sites** folder → **Default Web Site** → **EmcapWebSite**. Right click on the icon and select Properties.

On the ASP.NET tab, select the **Edit Configuration** button. On the General tab, update the domain and password settings to match those for the EmcapWebUser. Click **Apply**, and then **OK**. Click **OK** to exit the EmcapWebSite Properties window.

If you are experiencing problems with this step, you can troubleshoot by making sure that the Value for the EmcapWebServices can be copied and pasted into a web browser. Highlight the appropriate key and then select the **Edit** button.



From the Edit/Application Settings window, copy and paste the value into a web browser. You should see a list of operations supported by EmcapWebServ. If you do not see this list, you will need to review previous steps to determine the issue.

Step 9: While still in IIS, select EmcapWebService. Right click on the icon and select Properties.

On the **Directory Security** tab, select the **Edit** button for Anonymous access and authentication control. In the Authenticated Methods window, **uncheck** “Anonymous access” at the top of the screen and select “Integrated Windows authentication” at the bottom of the screen. Click **OK**.

On the **ASP.NET** tab, select the **Edit Configuration** button. On the Application tab, at the bottom of the screen in the Identity Settings panel, check “Local impersonation”. Type “EmcapUser” in the User Name field, and then the password that you assigned in Step 1. Click **OK**. Click **OK** to exit the EmcapWebService Properties window.

On the **General** tab, verify that the database connection strings in the Connection string manager panel have a value very similar to:

Data Source=localhost;Initial Catalog=EMCAP;Integrated Security=True;Application Name=EMCAP for the EmcapConnection string.

AND

Data Source=localhost;Initial Catalog=hMailServer;Integrated Security=True;Application Name=EMCAP for the hMailServer connection string

Note that Data Source in each case is replaced with the name of the database server or its IP Address.

Exit both windows and then exit IIS.

As a final step, open the following files in notepad and verify the database connection strings as described above:

- C:\Program Files\DCR\Emcap\ EmcapCore.exe.config
- C:\Program Files\DCR\Emcap\ AdminConsole.exe.config

Close out notepad.

Step 10: To test the Emcap Web Site; from IIS select the EmcapWebsite from the left panel. Right-click on the test.aspx page from the right panel and select **Browse**.

An Emcap Test page will appear. Click the **Go** button. After a brief pause, “Success” should be displayed at the top left of the screen.

To test the EmcapWebService; from IIS select the EmcapWebService from the left panel. Right-click on the EmcapWebService.asmx page from the right panel and select **Browse**. This should produce a page listing the available operations. From that list choose Test Connection. When **invoked**, it should return true, if not then the web service is not authenticating with the database server.