White Paper

INFORMATION ARCHIVING WITH EMC SOURCEONE

Something for Everyone

Abstract

This White Paper discusses compliance and legal requirements that Russian and Polish organizations must consider when storing their email and unstructured data; risks and problems associated with that data; and how a strong Information Archiving / Management program, enabled with EMC's SourceOne, can help organizations meet those requirements.

November 2012

EMC²

Part Number h11245.

Table of Contents

Information Archiving With EMC SourceOne

Something for Everyone

By Chris Dale, e-Disclosure Information Project[1]
James D. Shook, Esq., EMC Corporation[2]

Executive Summary

While archiving has been a useful tool for the IT department for many years, teams in compliance, human resources and legal departments are beginning to learn of the benefits delivered by a sound archiving initiative. These go beyond mere cost-saving and defensibility – there are positive business benefits as well. This White Paper will explore some of the best practices, challenges and benefits from deploying an archive.

This paper has been developed specifically to help businesses which are based in Russia or Poland, or which transact business in those countries. Many of the principles, however, are applicable everywhere and for companies of all sizes. An appendix includes notes about some of the relevant regulations.

[1] http://edisclosureinformation.co.uk/edisclosureproject.htm
[2] http://www.emc.com/products/family/emc-sourceone-e-discovery-family.htm

The Information Explosion Continues

The amount of electronic data created and stored continues to grow at a staggering rate. The amount of information created and stored more than doubles every two years. [3] In 2011, the world generated and consumed 1.8 zettabytes of information, with businesses responsible for about 80% of all information.[4] By 2020, organizations will be dealing with 50 times more information than they handled in 2009.[5]

Much of this information is in the form of email. In fact, email is arguably the most important repository for corporate information, containing "in excess of 80 percent of the business-critical content for an organization."[6] With the vast majority of this information never existing physically (i.e. printed to paper), it is more important than ever to properly manage information in its native electronic format.

How much of this information is useful will vary from company to company. Companies must keep certain material for statutory reasons or because of regulatory requirements; some of it may be required for actual, pending or anticipated litigation; some of it should be treated as a valuable asset of the business. The value of all of this information comes with a downside -- there are costs and risks involved with merely storing information, especially if it is held beyond its useful life. In the past, these costs have mainly been operational in nature – the cost of storage, backup, etc and of the staff to manage it. Today, many organizations should worry about the compliance risks tied to maintaining data. Although the unit cost of storing information has dramatically decreased[7], the risks and costs associated with compliance and legal issues have significantly increased.

Compliance Risk

The days of simply collecting and storing information, without planning how to manage, secure, access and delete that content, are gone. Maintaining compliance with applicable - and sometimes conflicting - laws cannot be done without organizations thinking about their compliance framework, creating policies to meet those requirements and deploying appropriate technology and tools to enable those policies. Consider the following:

- Corporations must be mindful of their employees' activities throughout the globe, which can require them to provide oversight and internal audits. The failure to undertake this work can be disastrous. Avon Corporation has now spent $280 million on its internal investigation into allegations that it violated the (US) Foreign Corrupt Practices Act with bribes originating in France, Brazil and possibly China.[8] The initial allegations came from a whistle blower, not from the company's auditors or investigators, further complicating matters. The four-year-old investigation is still incomplete, and one can only speculate on a final fine / settlement amount.

[3] "Extracting Value From Chaos", IDC Digital Universe (Sponsored by EMC), June 2011.
[4] Id.
[5] Id.
[6] "Proper Email Governance For the Protection Of Your Business & Brand", A Frost & Sullivan White Paper.
[7] From 2005 to 2011, the cost of acquiring, storing, managing, etc. a gigabyte of information has fallen to 1/6 of the original cost. IDC, Extracting Value From Chaos.
[8] www.fcpaprofessor.com/category/avon. Last visited September 20, 2012.

- More individuals are availing themselves of the EU Data Privacy Act's right to require organizations to provide information that it has retained on that individual, which can be disruptive and expensive. (Similar regulations exist in Russian and Poland). Proper compliance with such demands can consume resources which might be better used elsewhere; failure to comply can be even more expensive. A French company was fined €10,000 (US$12,281) in July 2012 for being either unable or unwilling to provide an employee with data requested under France's data privacy law to help him with a road accident claim.[9]

- Companies must be wary of being hauled into faraway courts, especially in the United States where litigation requirements often conflict with data privacy principles in Europe and Asia. In one case, a German company was ordered to produce responsive, third-party personal data to a US Court, even though the data was stored in Germany and the company argued that the production would violate German law.[10] The court found held that such blocking statutes "do not deprive an American court of the power to order a party ... to produce evidence even though the act of production may violate that statute."

- In these difficult economic times, some elements of the "fraud triangle" – motive, rationalization and opportunity – can be enhanced. Employees suffering through financial difficulties or layoffs can find more motives and rationalization in illegal activities. The mixture of policies and technology which comes from an archiving strategy brings enhanced audit and investigation capabilities; these protect the organization by removing the third side of the triangle – opportunity.

How Archiving With SourceOne Enhances Compliance and Operations

It is often assumed that the term "archive" implies a place to keep larger volumes of old material for long periods. In fact, information management policies are frequently used to enable and enforce significantly shorter retention periods for electronic content.[11]

An "Information Archiving / Management" initiative, including email management as a key component, can significantly assist organizations in meeting their compliance requirements while simultaneously delivering critical operational benefits. Information Management can be defined as the structure on which an organization establishes, enables and enforces a policy related to the retention and classification of information, particularly email, but also applying to other common repositories such as fileshares and Sharepoint.

[9] http://www.pcworld.com/article/258964/french_company_fined_and836410000_for_failing_to_share_gps_tracking_data_with_employee.html
[10] AccessData Corp. v. ALSTE Tech. GMBH, 2010 WL 318477 (D. Utah Jan. 21, 2010)
[11] See "Developing Effective Email Management Policies", James D. Shook.

Some of the key benefits derived from implementing an archive include[12]:

- Complying with regulatory obligations to retain – and delete - data in an appropriate and secure manner;
- Responding quickly and efficiently to legal requests for data in a defensible manner;
- Having ready access to data to support internal audit, investigation and other compliance needs; and
- Operational benefits including self-service access to archived content, reduced storage requirements and improved system performance.

The following sections take a closer look at how EMC's SourceOne family delivers these benefits for the compliance and operational challenges commonly seen by Polish companies.

Complying With Regulatory Obligations To Retain Data

Even with an archive relieving many of the operational issues for data retention, information cannot be held indefinitely. The Russian Federal Law "On Personal Data" of 2006 and amended in 2011 ("Russian OPD"), the Polish Act on Personal Data Protection ("Polish PDP") and the 1995 EU Data Protection Directive [13] all require that data should only be held while it is current and useful.[14] Timely deletion is specifically required by the Polish PDP (Article 26), implied in requirements of other countries and supported elsewhere as a best practice:

> *Data Controllers should retain Protected Data only as long as necessary to satisfy legal or business needs. While a legal action is pending or remains reasonably anticipated, Data Controllers should preserve relevant information, including relevant Protected Data, with appropriate data safeguards.*[15]

In addition, repositories that contain protected information must be secure. The Russian OPD recommends encryption and other technical and organizational measures to prevent unauthorized access to personal data. Likewise, the Polish PDP's security requirements require:

> *The controller shall be obliged to implement technical and organisational measures to protect the personal data being processed, appropriate to the risks and category of data being protected, and in particular to protect data against their unauthorised disclosure, takeover by an unauthorised person, processing with the violation of the Act, any change, loss, damage or destruction.*[16]

---

[12] See also "Qualifying the Costs and Benefits of Archiving Your Email and Other Electronic Content", Osterman Research, October 2011.
[13] Directive 95/46/EC on the protection of individuals with regard to the processing of personal data and on the free movement of such data
[14] In some cases, a request may be required to force deletion of the data, but inaccurate data or improper processing may result in damages.
[15] The Sedona Conference International Principles on Discovery/Disclosure/Data Protection, Guideline 6.
[16] Polish PDP, Article 36

To meet this type of requirement, it may be necessary to eliminate local caches of messages (i.e. PSTs/NSFs). These troves of information, often created by users precisely in order to subvert data retention policies, cannot be properly managed and pose security risks. An unmanaged PST file can contain years of email messages, which might in turn represent personal and sensitive personal information on employees, forecasts, business plans and strategies, trade secrets, etc. Such information may be potentially discoverable in litigation or to a regulator, quite apart from any breach of data protection and privacy laws.

In some cases, however, organizations actually need the ability to retain specific information for long periods, and in a compliant manner. For example, the Polish Accounting Act and the Telecom Act can require content - agreements or other critical information – to be retained for five years or even longer. Some organizations may want to retain certain data for longer periods, matching the duration of statutes of limitation relevant to their region or industry sector. A long retention period is generally permitted under most data privacy regimes if there is an appropriate business purpose. If there is a specific, compliance requirement to keep the data, then it is ipso facto important and cannot be entrusted to a quota-driven email server or even a PST file that can become corrupted or lost at the mercy of a single employee.

EMC's SourceOne Family can help organizations to meet all of these challenges. Retention policies can easily be matched to specific user groups, such as legal, finance, HR, etc., with retention periods automatically assigned to messages based upon this hierarchy. Retention is applied automatically, so that when a message or file reaches its retention period – say one year – it is deleted without end-user intervention. Organizations then can be certain that their policies are being followed and enforced.

Retention of longer-term, targeted content is also supported. With email, employees can use the user-directed archive to separate long-term business records from personal and other sensitive information and retain those messages for whatever longer term period is appropriate. Because these UDAs are contained within the archive, and not in an individual employee's corporate mailbox or PST, they are under stronger company control and not subject to being lost as a result of an employee laptop theft or departure. Similarly, files in file shares and SharePoint that meet pre-defined requirements can also be retained for a longer term.

SourceOne also provides the operational capabilities that eliminate the need for unsecure PSTs or NSFs. Advanced deduplication and the use of archival storage cut costs so that employees can have a virtually limitless mailbox, limited by organizational policy and not by size. The SourceOne Offline Access feature insures that these archived messages remain available, in a secure manner, to employees even when they are not on the network. Existing PST/NSF files can even be absorbed into the archive during the implementation process.

Responding To Legal Requests For Data

Email messages and files must often be located and preserved to meet specific requirements such as regulatory inquiries, investigations or litigation. For example, these capabilities can

be necessary to properly respond to individuals who demand information, under the Russian ODP or Polish PDP, about their own personal data being held by an organization.

Companies transacting business in the UK or the US, or which have dealings with companies located in those jurisdictions, also face the possibility of litigation in their courts. The legal system in those countries includes a process called discovery or disclosure, where both parties must provide information, including electronically stored information, which is relevant to the dispute – even if it is harmful to their own case. The process of identifying, preserving, collecting, analyzing and reviewing relevant electronic information – "e-Discovery" in the US or "e-Disclosure" in the UK – is often difficult and expensive for companies to comply. The failure or inability to comply with the obligations accurately and quickly can result in severe penalties from the Court. Quite apart from the risk of such sanctions, cases and regulatory positions cannot be prosecuted or defended properly if the organization fails to retain information required to establish its position.

However, in situations when data is required to be held by law, it is usually inappropriate to simply retain all of the content held by the organization as a safety measure. Retaining just the data which one might reasonably be expected to produce, particularly with email content, is the best way to maintain compliance with the Russian OPD and Polish PPD and is also a best practice:

> *Preservation or discovery of Protected Data should be limited in scope to that which is relevant and necessary to support any party's claim or defense in order to minimize conflicts of law and impact on the Data Subject.*[17]

This can become even more critical if the data must be transferred, which transfers are restricted by both the Russian OPD and Polish PPD, subject to certain exceptions. An established process which takes account of foreseeable demands, along with supporting tools, can make it simple and easy to find data when required. Ad hoc processes tend to waste time, require more staff and result in the loss of business resources.

EMC's SourceOne Family provides the capabilities to meet these requirements – and more. Discovery Manager allows precise searching of the SourceOne archive, using keywords, date ranges, custodians, etc. to insure that only relevant content is located. This targeted content can then be saved separately, insuring that regular, untargeted content continues to be retained only for the appropriate retention period. Once collected, relevant information can be further culled, reviewed and tagged, insuring that only targeted collections are retained and, when necessary, that appropriate data owners can be notified, perhaps because their consent is required.

[17] The Sedona Conference, supra, Guideline 3.

Ready Access To Data Supporting Internal Compliance Requirements

Many companies are also proactively monitoring their information for audit, security and even as a precaution against severe laws such as the Anti-Bribery Act in the UK and the Foreign Corrupt Practices Act in the US and other countries. These laws are broad and far-reaching, requiring only a very tenuous relationship with the legislating country to trigger its impact. Proactive positioning requires organizations to review certain information – usually email – created by individuals that have been identified as high risk or whose actions have brought them to the attention of security or audit processes.

EMC's SourceOne Family provides robust capabilities to meet these requirements. Emails can be journaled (captured at the point of being sent or received), archived and indexed for immediate searching. Messages containing certain keywords or meeting other criteria can be quickly and securely reviewed by appropriate personnel, with all actions being logged to further insure compliance with internal and external regulations.

Operational Benefits

Although compliance concerns are important, providing operational benefits remains an important goal for most companies looking to implement an archive. Reducing backup windows, cutting storage costs and simplifying email support issues are all common organizational goals – and if they are not then they should be.

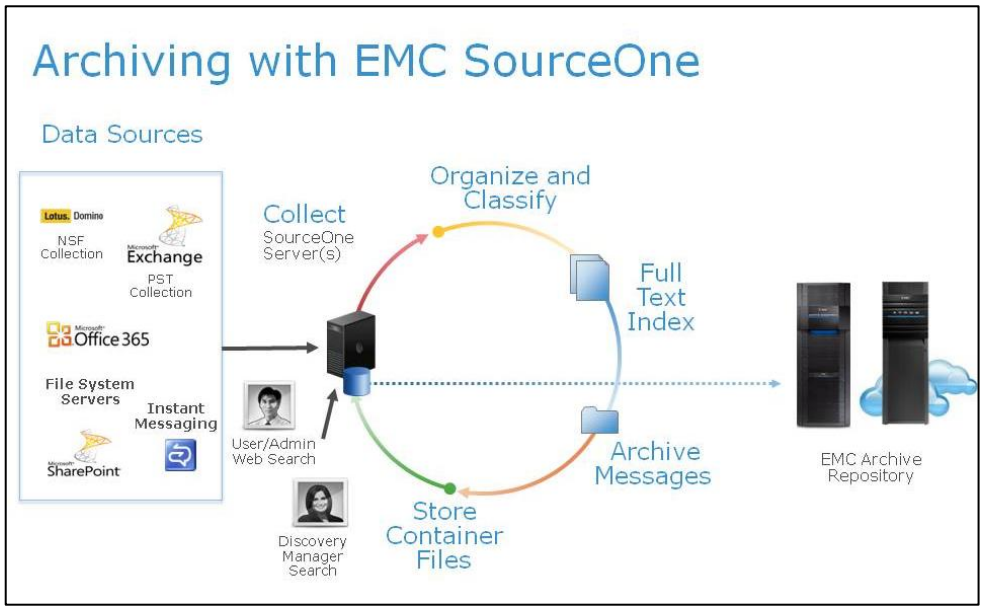
Email Archiving

Today, many organizations maintain significant operational benefits through the implementation of an email archive. In a standard corporate email environment without an archive, all email messages are collected and maintained by the corporate email server(s), which are centralized and maintained by the IT Department. In a typical default mode, all messages are stored on the server, which quickly grows in size, resulting in (a) operational problems for the server; (b) difficulty in creating a daily backup of the server's content for business continuity purposes; and (c) continuing expense in purchasing new storage, typically requiring a higher performance, more expensive tier of storage to meet the requirements of the email system. These costs, large as they are, can seem relatively insignificant when compared with the burden of searching through the servers on an ad hoc basis for discovery purposes, often repeatedly over the same volumes.

An email archive works by offloading the content from the email server to a less expensive tier of storage, applying and enforcing corporate policies over the content, and making it readily accessible for search. At a high level, this work is done in five steps:

1. Collecting emails from the email server on a pre-established schedule;
2. Organizing and classifying the messages based upon company established criteria;
3. Generating a full-text index for search and eDiscovery purposes, enabling search by keyword, date range, sender or recipient, etc. ;
4. Bundling these messages together to create an "archive" of messages; and

5. Moving the containers with the archived messages to less expensive, archival storage.



The EMC SourceOne Family includes an email management module which delivers best-in-class features and performance in these areas, too. Its scalable architecture has already been proven to meet the demands of huge organizations. The flexibility of a master-worker architecture enables the system to scale upon demand, including on-demand deployment of VMWare (virtualized) environments. Its integration with EMC's best-of-breed EMC storage solutions – such as Data Domain, Atmos and Centera -- insures that storage requirements are also integrated into and met by a complete EMC solution.

File share and SharePoint Archiving

Organizations are seeing rapid growth in their file share and SharePoint environments, resulting in many of the issues previously seen with email. To meet these challenges, the *EMC SourceOne Family* also offers modules for file share and SharePoint archiving. Content meeting certain criteria – such as age - can be archived, retained for a pre-set period and indexed for immediate search requirements.

Conclusions

The problems raised by growing data volumes will not go away or get smaller. Volumes are increasing, and so is the need to access that information. Even non-litigious companies are open to attack by regulators, local, EU-wide and US, whose power and scope increases year by year.

The reasons for addressing the issue are not merely defensive and not merely for cost reduction, though these provide motive enough. There are more positive business reasons for setting up an archive and its surrounding policies, including better use of what is left. After all, information is a key business asset.

Technology is the servant to policies which are in part universal, in part local to the jurisdiction and in part specific to an industry sector. Dealing with it needs a human element – someone to take responsibility not just for infrastructure of boxes and wires and not just for security, but for business purposes.

One issue which companies face is that the people who are responsible for the infrastructure are not generally the same people as the users, who are in turn distinct from those who face the demands of litigation or of regulatory intervention. A yet further group of people have the task – or perhaps no-one does – of maximizing the business value of the assets which lie in an organization's data stores.

This divergence of both responsibilities and control brings problems of communication, of specification and of budgets. The gap between the corporate players, and their external lawyers, can be bridged by the involvement of a consultative partner who understands the business needs as well as the software and boxes which are available to address the problems and help uncover the asset value of the ever-increasing volumes of data which companies have.

Appendix

This is a brief summary of some of the important compliance and legal requirements that apply to many organizations in Russia and Eastern Europe:

Russian Federal Law On Personal Data (2006/2007/2011)

Russian law regulates the processing of personal data by "data operators." Data cannot be processed without prior written notification of the regulating agency, the Roscomnadzor, unless an exemption applies to the processing. There are ten specific exemptions to the notification requirements, including: (1) data processing under employment law; (2) data is received in connection with a contract; (3) data relates to a public association or religious organization; (4) the data was publicly available from the data subject; (5) the data consists of only basic information – surname, first name and patronymic. "Sensitive personal data" – including racial or ethnic origin, political opinions and religious beliefs and other commonly understood "sensitive" categories – is entitled to additional protections.

Data may be processed upon consent of the data subject or upon other exemptions including judicial purposes; for the purpose of an agreement with the data subject; statistical or scientific purposes; to protect the data subject; and other reasons. In the case of consent from the subject, written consent is not required, but the burden of showing consent is on the data operator. Thus, obtaining written consent - -which may be electronic – is generally advisable. Individuals have the right to obtain information being held about them via a free subject access request, although there are exceptions which make the provision not as broad as the similar provision in the EU. Penalties, including damages, are provided for violations.

Polish Act On Personal Data Protection (1997/2004)

Poland first implemented the Polish Act on Personal Data Protection ("Polish PDP") in August 1997, and amended the Act in 2004. Similar to other laws enacting the European Union's Data Protection Act, the Polish PDP provides extensive rights to data subjects.

Only "personal data" and "sensitive personal data" are actually protected by the Polish PDP. But those definitions cover a very broad segment of information and thus, as a practical matter, all email can be treated as potentially subject to its terms.[18] In addition, almost any activity, including mere storage, is deemed to be "processing" and is subject to the requirements of the Act. (Article 7). Personal data can be collected only for "specified and legitimate purposes", without any further processing, and should not be maintained longer than to meet those purposes (Article 26). Proper security methods must be implemented to protect the data (Article 31). The Act also provides extensive rights for data subjects to understand what data is being held about them, including the ability to update and (in some cases) delete the information upon request (Article 32).

[18] Some authorities have held that an email address – something that is contained in every email – is "personal data". Others have held that an originating IP address, usually found in the metadata of an email message, is similarly personal data.

UK Bribery Act / United States Foreign Corrupt Practices Act

Both the UK and the US have enacted extensive legislation to fight corruption – defined in the US legislation mainly as the bribery of officials for economic gain but applied by the UK Bribery Act to any commercial dealings. Only a tenuous connection to the country is required to trigger provisions of these acts, so companies in Poland with even limited US or UK involvement could be subject to their terms. Each provides substantial fines and even imprisonment for violations.

In the UK, the only defense to an alleged violation of the corporate offence of failing to prevent bribery is that the organization had in place "adequate procedures" to prevent an act of bribery being committed in connection with its business. This covers a wide range of HR, audit and other areas but as a best practice, these "adequate procedures" should include a program to monitor email and other communications for potentially improper activity to the extent that privacy laws allow this, and to take follow-up action as required.

Similarly, in the U.S., the Foreign Corrupt Practices Act (FCPA) provides stringent penalties for those making bribes or offering other inducements for business with a United States firm. Most companies with FCPA concerns will want to monitor the email and other communications of "at risk" employees and officers, again within the limits allowed by privacy laws.

Polish Accounting Act

The Polish Accounting Act provides the requirements for insuring that financial statements for Polish companies are accurate and harmonized with corresponding acts in the European Union. Among other requirements, certain data supporting the preparation of financial statements must be retained. Although there is no direct requirement to preserve or retain email, the reality is that like so many other forms of information, much of the information required to be retained is found originally or most conveniently in email messages. Some of the retention requirements include[19]:

- Payslips or their equivalent: for a period during which access to this information is required in accordance with the retirement and disability benefit regulations; for 50 years from the time that the insured stops working for a given employer; and in accordance with tax regulations, but for no less than 5 years;
- Accounting documents confirming receipts from retail sales: until the day the financial statement for a given financial year is approved, but not prior to the day that the persons to whom the assets included in retail sales were entrusted are appraised;
- Accounting documents regarding long-term investments, loans, credits, sales agreements and claims pursued in civil, criminal or tax proceedings: for 5 years from the beginning of the year that follows the financial year in which operations, transactions and proceedings are finalized, paid, settled or become outdated;
- Documents regarding warranties and complaints: for 1 year from the expiry of a warranty or the settlement of a complaint.

[19] http://europa.eu/youreurope/business/managing-business/keeping-accounts/poland/index_en.htm

Electronic Discovery and Disclosure In Litigation (Mainly US and EU)

Business disputes in the United States are often resolved in the courts through a process known as litigation. The litigation process usually includes extensive "discovery" of facts – both helpful and harmful – that the parties must disclose to one another, including relevant "electronically stored information" such as email and documents. The United Kingdom has some related processes while mainland Europe has no tradition of the discovery process.

While these requirements are limited to the United States and the United Kingdom, their impact reaches well beyond the borders of those countries. Anyone transacting business in the U.S. or UK, possibly even with a party located in those countries, may find itself subject to the jurisdiction of those courts. Organizations that are unprepared for the process of discovery often are unable to collect and produce relevant information in a timely fashion, which can result in severe penalties – both in Court in the form of sanctions, and in corporate embarrassment.

Further complicating the process is that the European Union's Data Protection and privacy requirements often conflict with the US/UK obligation to produce relevant data, particularly if the purpose is pre-trial discovery in foreign jurisdictions, further driving up the cost and risk of the process. The EU is considering a new Data Protection Regulation which will increase both the restrictions and the penalties for breach. Companies whose data, especially email, is archived, classified and, where possible, deleted in accordance with a policy of the kind described in this paper, are better able to comply with these restrictions in a timely and cost-effective manner when faced with a request or demand to send data abroad.

Records Retention

Most organizations have operationally requirements to retain certain information – "records" - for the efficient operation of a business. This information may include employee reviews and applications, financial data, communications about the creation or implementation of a contract, changes to service, internal test data, etc. Some of this information may also be subject to regulations mandating its retention for a certain period of years – i.e. tax and banking information. More frequently, this information is contained in or attached to email messages.