

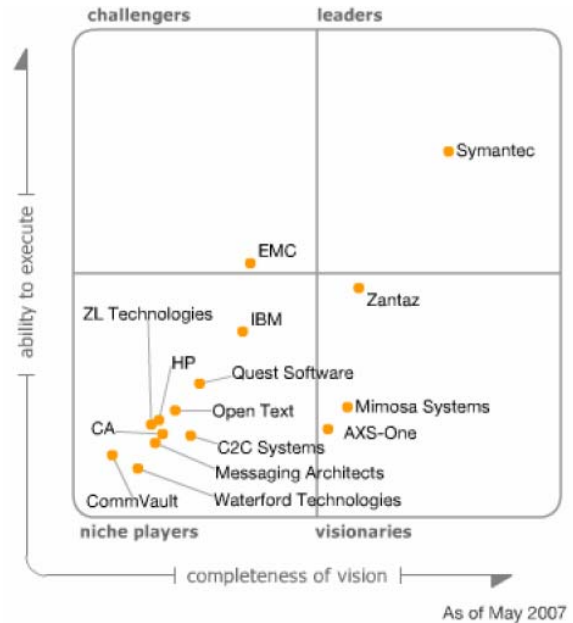
# 資料歸檔與郵件歸檔比較說明

## 1.0 全球資料歸檔市場現況

根據 Gartner 2007.05 的全球郵件歸檔調查報告，目前主要進行資料歸檔廠商有 EMC、IBM、Symantec、HP、CA、AXS-ONE。而在郵件歸檔全球市場軟體授權產值將由 2006 年 2.7 億美元提升到 2011 年 10 億美元。

調查報告中亦指出，郵件歸檔產品優先評估的是能完整歸檔企業郵件（All Email Message），並提供有效的搜尋調閱（Searchable archive）與歸檔管理，符合法規遵循、企業資安規範與稽核作業。

同時產品的效能必需能夠滿足企業端的需求，但目前就 Gartner 所觀察到的狀況，將不同資料來源統一集中歸檔管理是未來儲存廠商努力的願景，但現階段由於資料型態的差異與不同，包括技術整合的困難，要完成集中式的管理，短期是無法被實現的，因此報告中建議可以先針對電子郵件進行妥善的歸檔，後續再根據儲存廠商提供的解決方案進行異質資料整合。



## 2.0 資料歸檔解決方案應用趨勢 - 以 EMC、IBM 為例

摘錄 Gartner 對於 EMC 與 IBM 的郵件歸檔解決方案的觀察，建議尚需注意項目如下，

廠商	Gartner 觀察
EMC	<p>Cautions :</p> <ul style="list-style-type: none"> <li>● Companies with large numbers of mailboxes and complex compliance requirements should wait for EMC to deliver its planned large-company e-mail archiving solution or expect to migrate in the future.</li> <li>● Companies requiring robust support for <b>double-byte characters</b> should test thoroughly to see whether current functionality is sufficient.</li> </ul>
IBM	<p>Cautions :</p>

	<ul style="list-style-type: none"> <li>● Multiple e-mail archiving products, since the acquisition of FileNet, are confusing to sales personnel and prospects, and <b>will require careful monitoring of road maps.</b></li> <li>● Although preloaded systems offer some ease-of-deployment support, deployment and management is <b>still complex and requires knowledge of DB2. Compliance supervision</b> will require integration with a <b>third-party solution.</b></li> </ul>
--	--

參考資料歸檔廠商的產品藍圖規劃（如下表），未來亦將採取獨立執行歸檔作業趨勢

EMC	IBM
<ul style="list-style-type: none"> <li>● <b>Archiving for Mail and Messaging</b></li> <li>● Enterprise Application Archiving</li> <li>● File System Archiving</li> <li>● Integrated Content Archiving</li> <li>● SharePoint Archiving</li> </ul>	<ul style="list-style-type: none"> <li>● <b>DB2 Content Manager Solution for Email Archive</b> <ul style="list-style-type: none"> <li>■ e-mail capture</li> <li>■ easy and reliable access to archived e-mail</li> <li>■ archive and retrieval activity logs</li> <li>■ flexible and open content management</li> </ul> </li> </ul>

以上資料顯示，不論是 EMC 或 IBM 等儲存設備廠商的歸檔解決方案，均存在大量郵件調閱的技術困難，以及 Double Bytes 編碼支援性問題，以 IBM 來說不僅有上述問題外，還需與現有系統 DB2 整合，資料檢索調閱尚需評估其他解決方案（3<sup>rd</sup> solution）配合進行，才能符合。

### 3.0 郵件歸檔市場競爭趨勢與分析

目前全球市場趨勢郵件歸檔廠商區分為三大類，有企業訊息管理(Enterprise Message Management)、企業內容管理(Enterprise Content Management)、儲存設備(Storage)，Openfind MailBase 分別結合三種類型廠商技術優勢，提供完整的郵件歸檔功能。郵件歸檔主要解決「管理」問題，同時又要兼具「安全」與「效能」，根據這三點，分別說明各類廠商主要技術與應用。

郵件訊息管理廠商，以自行開發的工具執行歸檔功能，查詢功能較為陽春（例如無法提供全文或多欄位檢索），需倚賴 3<sup>rd</sup> Parter 整合，同時若無搭配工具軟體，歸檔郵件需先還原至檔案伺服器後，始可進行查詢。

儲存設備廠商，主要功能還是著重在空間有效管理，儲存是附加裝置，倚賴前端系統提供郵件資料來源。簡單來說，就是將空間做最有效的利用，此領域已發展相當多的儲存技術（例如 DAS/NAS/SAN）。並以使用率或時間去做區分，發展資訊生命週期管理。

內容管理廠商，主要為垃圾郵件(anti-spam)或防火牆(Firewall)廠商，以 Gateway 的技術應用於郵件的管理，包含郵件的內容攔截與行為的偵測。主要問題在於多數廠商無法跳脫 Gateway 環境，產生有內對內信件歸檔問題。同時，此類廠商多發展 UTM 架構，但企業郵件使用量與應用多元化，技術與效能是否能配合企業的成長而跟進，也是要審慎評估。

【表一】郵件歸檔市場競爭簡易比較表（Openfind 整理，版權所有）

類別	屬性	主要缺點
EMM	郵件系統相關開發商。(例如 Microsoft, IBM,等)	<ul style="list-style-type: none"> <li>● 多數無法過濾重複信件</li> <li>● 雙字元辨識 (Double bytes)</li> <li>● 缺少郵件權限調閱管理功能</li> <li>● 無法直接查詢歷史郵件</li> <li>● 採用 SQL 資料庫，附檔檢索效能有限</li> <li>● 管理功能仍需倚賴進行合作夥伴技術支援</li> </ul>
ECM	內容管理廠商，多為 UTM 產品。(例如一般國產廠商)	<ul style="list-style-type: none"> <li>● 多數無法歸檔內部互寄郵件</li> <li>● 缺少郵件權限調閱管理功能</li> <li>● 無法直接查詢歷史郵件</li> <li>● 採用 SQL 資料庫，附檔檢索效能有限</li> </ul>

		<ul style="list-style-type: none"> <li>● 多數為低硬體成本，後續維護風險高</li> <li>● 多功能整合設備（UTM）運作效能待評估</li> <li>● 多數為簡易郵件記錄器（Log）功能，無歸檔與調閱管理功能</li> </ul>
Storage	提供資料備份 (Backup) 與還原 (Recovery) 解決方案，（例如 CA, EMC, HP, Symantec）	<ul style="list-style-type: none"> <li>● 多數無法過濾重複信件</li> <li>● 採用 SQL 資料庫，附檔檢索效能有限</li> <li>● 缺少郵件權限調閱管理功能</li> <li>● 雙字元辨識（Double bytes）等問題</li> <li>● 在地化服務支援度需評估</li> <li>● 多數已安裝代理程式進行，需變動伺服器設定</li> <li>● 需先進行系統或郵件還原後始可查詢</li> </ul>

## 4.0 Openfind MailBase 結合郵件與搜尋核心技術

Openfind MailBase 是目前市場上唯一具備高速「搜尋引擎」技術的郵件歸檔產品，郵件採用標準資料格式（RFC822 & MIME）提供明確基本欄位屬性，例如寄（收）件人、郵件主旨等，很清楚辨識郵件溝通的歷程，可以從帳號、時間，關鍵字詞、附件等檢索條件，快速利用搜尋引擎找到相關郵件獲得所需要的資訊。範圍包括各類 Microsoft Office、Adobe PDF、RFF、ZIP、TAR 等各種附檔格式。



功能支援「附檔辨識」與「加密附檔」檢索，協助稽核人員找出可疑郵件，同時再調閱權限，不僅止功能設定，可搭配群組結構 (LDAP/AD)，進行調閱權限設定。例如 A 與 B 員工，因權限不同，在相同關鍵字查詢條件下，會查到不同比數的信件結果。

**【表二】** 郵件歸檔相關產品歸檔功能比較

產品品牌	Openfind	其他國產品牌	其他國外品牌
搜尋引擎	○	×	×
調閱權限	○	×	×
郵件追蹤	○	×	×
離線查詢	○	×	×

## 5.0 Openfind MailBase 與資料歸檔產品整合應用案例

MailBase 於 2005 年日本上市以來，由於日本個人情報保護法的推動下，陸續與儲存設備廠商合作，提供客戶符合法遵循的郵件歸檔解決方案。在國際化或整合彈性已累積相當豐富，相關合作記錄如下：

合作廠商	內容
EMC	<ul style="list-style-type: none"> <li>➤ MailBase 與 EMC Japan 聯袂推出 Mail Archive Solution</li> <li>➤ 說明： 為了提供日本企業客戶更完整的 <a href="#">Mail Archive Solution</a>，CyberSolutions 於 2007 年 3 月與 <a href="#">EMC Japan</a> 策略聯盟，共同整合了 <a href="#">MailBase</a> 軟體 與 <a href="#">EMC Centera™</a> 儲存設備，為企業提供高效能、高品質並符合法規要求之最佳解決方案。</li> </ul>
EMC	<ul style="list-style-type: none"> <li>➤ 大規模メール監査システムとコンプライアンス・ストレージによる内部統制</li> <li>➤ 說明：與 EMC 合作推出符合「個人情報保護法」的郵件歸檔解決方案。</li> </ul>
EMC	<ul style="list-style-type: none"> <li>➤ NEC ネットエスアイ in EMC Forum 2007</li> <li>➤ 說明：MailBase 日本經銷商 NEC 與 EMC 舉辦座談會，提供客戶進行郵件歸檔導入建議。</li> </ul>
EMC	<ul style="list-style-type: none"> <li>➤ 『まだ間に合う！2008 年メール統制元年に備えよ！』 ～(株)ツムラにおける"CyberSecure シリーズ導入までの変遷と次の一手"』</li> <li>➤ 說明：日本經銷商 Cybersolution 與 EMC 合作郵件歸檔研討會。</li> </ul>
IBM	<ul style="list-style-type: none"> <li>➤ Lotus Notes/Domino 対応のコンプライアンスソリューション「メールアーカイブおまかせパック」</li> <li>➤ 說明：與 IBM 合作提供 Notes 郵件歸檔解決方案</li> </ul>
IBM	<ul style="list-style-type: none"> <li>➤ PFU ブース in LotusDay 2007</li> <li>➤ 說明：日本經銷商 PFU 於 IBM Lotus Day 提公郵件歸檔解決方案</li> </ul>
IBM	<ul style="list-style-type: none"> <li>➤ 『企業を守るメール監査とデータ管理』セミナー</li> <li>➤ 說明：日本經銷商 Cybersolutions 與 IBM 一起合作推動 Notes 個人情報保護法解決方案</li> </ul>

## 6.0 結論

電子郵件是主要的溝通工具以及具法律效用的正式文件，對組織而言就是一種知識資產，和其它重要的營運資產一樣具有價值，都必須持續給予妥善保護。企業實際應用的資料型態概分為三類：

- **結構化資料**：是指有欄位、有項目的資料，例如：試算表，財務資料、人事資料等。
- **半結構化資料**：由文字組成，卻無規則可循或格式不固定，在以往是很難加以處理的，例如：Email、Office files、html 等。
- **非結構化資料**：例如照片、影像、語音、視訊、掃描文件等。

如果以檔案大小來評估，結構化資料在歸檔整理的效率是最高的，由於資料量小所需的儲存空間也不高，而非結構化資料在歸檔的處理是費時也需要更大的儲存成本。過去的資料庫都只能針對結構化的資料進行處理，而半結構化或非結構化的資訊都無法納入其中，所以造成系統使用上的瓶頸。

### 建議優先導入郵件歸檔管理

電子郵件的優勢在於不像其他的非結構性的資料不易整理萃取。以一般放在資料匣的 Office 檔案進行比較，如果沒有特別加以註記或設定，會不清楚這篇文章的作者是誰，文章主旨或機密等級為何？利用郵件來進行資料歸檔的好處在於郵件採用標準資料格式（例如 RFC 822 & MIME）提供明確基本欄位屬性，例如寄（收）件人、郵件主旨等，很清楚辨識郵件溝通的歷程，可以從時間，關鍵字詞、副本等檢索條件，快速利用搜尋引擎（Search）找到相關郵件獲得所需要的內容或做稽核舉證。因此搜尋適合用來作為調閱郵件工具的優勢就在於可以從半結構、非結構化的資料中，萃取相同的關鍵字詞，將資料分析集合，產生接近處理結構化資料的查詢效能。

### 保留彈性因應未來技術整合

根據 Gartner 的報告，目前儲存設備廠商仍存在系統效能、資料管理與搜尋技術問題，但是為未來儲存設備廠商努力完成的目標，現階段不論從管理面、使用效益、儲存成本與國際儲存廠商（如 IBM、EMC）所提供的解決方案的評估，建議應分別針對不同資料型態，獨立分開進行歸檔管理較為適合。由於 MailBase 採用標準郵件格式歸檔，整合搜尋引擎技術，輔以大型儲存設備廠商的合作經驗，未來可以配

合更完整的資料歸檔與稽核管理的計畫。

## 7.0 參考資料

1. EMC , Archiving for Mail and Messaging  
<http://www.emc.com/solutions/samples/backup-recovery-archiving/archiving-mail-messaging.htm>
2. IBM , Data Archiving Techniques and Considerations for the DB2 Family  
<http://www-1.ibm.com/support/docview.wss?rs=64&uid=swg27004347>
3. IBM , DB2 Content Manager Solution for Email Archive  
[http://www-306.ibm.com/software/data/cm/solution\\_emailarchive.html](http://www-306.ibm.com/software/data/cm/solution_emailarchive.html)
4. IBM 進軍資料歸檔市場  
<http://news.networkmagazine.com.tw/software-application/sftwnews/2008/07/22/1256/>
5. Gartner 2007 Email active-archiving Magic Quadrant



# Openfind™

網擎資訊軟體股份有限公司

地 址：台北市 100 南昌路二段 222 號 4 樓

電 話：02-2369-7575

傳 真：02-2364-8738

網 址：<http://www.openfind.com>

E-mail：[pm@openfind.com](mailto:pm@openfind.com)