

White Paper

Realizing eDiscovery Benefits with EMC SourceOne eDiscovery-Kazeon

By Katey Wood, Analyst

June, 2011

This ESG White Paper was commissioned by EMC and is distributed under license from ESG.

Contents

Introduction	3
eDiscovery Spending: No End in Sight	3
In-house Factors with the EMC eDiscovery Benefits Calculator	3
Litigation Profile and Investigative Approach.....	4
Data Retrieval Considerations	4
Comparing Current eDiscovery Costs	5
Determining Costs: Visible and Invisible	5
Pricing Options and Transparency.....	5
eDiscovery Technology Investment: The Value Proposition	6
“Risk Creep:” The Price of Poor eDiscovery.....	7
Organizational Costs and Inefficiencies.....	7
Additional Risk Factors	7
The Bigger Truth	8

All trademark names are property of their respective companies. Information contained in this publication has been obtained by sources The Enterprise Strategy Group (ESG) considers to be reliable but is not warranted by ESG. This publication may contain opinions of ESG, which are subject to change from time to time. This publication is copyrighted by The Enterprise Strategy Group, Inc. Any reproduction or redistribution of this publication, in whole or in part, whether in hard-copy format, electronically, or otherwise to persons not authorized to receive it, without the express consent of the Enterprise Strategy Group, Inc., is in violation of U.S. copyright law and will be subject to an action for civil damages and, if applicable, criminal prosecution. Should you have any questions, please contact ESG Client Relations at (508) 482-0188.

Introduction

Determining eDiscovery costs and requirements can be difficult for enterprise customers given that the work has traditionally been outsourced to third parties, sometimes with minimal or inconsistent tracking and transparency. Understanding the internal eDiscovery process, points of inefficiency, and the less visible costs of poor process and preparation is integral to weighing the potential benefits of further investment. This white paper uses the EMC eDiscovery Benefits Calculator¹ as a basis for evaluating and comparing eDiscovery costs.

eDiscovery Spending: No End in Sight

eDiscovery costs have grown substantially since the 2006 changes to the Federal Rules of Civil Procedure made clear that all electronically-stored information is discoverable in US courts. According to a 2010 Duke University survey of litigation costs at major companies, the average discovery costs per case ranged from \$621,880 to \$2,993,567 from 2006-2008, topping \$2,354,868 to \$9,759,900 per case for companies at the high end of the spectrum.² Moreover, growth in eDiscovery expenses shows no sign of flagging: in 2010, leading global law firm Fulbright & Jaworski's 7th Annual Litigation Trends Survey Report indicated that electronic discovery was the top litigation area targeted for increased spending with more than 40% of the largest companies planning to increase electronic discovery spending in the next year.

These costs are due to escalating data volumes—the Economist estimated worldwide data growth from 150 exabytes (a billion gigabytes) to 1200 exabytes from 2005-2010³—along with the effort required to extract this data from disparate enterprise systems and endpoint devices, handle it defensibly, and, finally, pay attorneys to examine it. Inefficiencies in the process are common—most enterprise companies retain data as standard policy regardless of its utility (or risk) to the business. This irrelevant data is still subject to discovery, often at high cost. Duke estimated that in major cases going to trial in 2008, the ratio of pages discovered to pages entered as exhibits at trial was as high as 1000/1. On average, 4,980,441 pages of documents were produced, but only 4,772 exhibit pages were marked.

Increasing efficiency in the litigation response process is essential to curbing ballooning discovery costs. At a macro level, it's easy to make the case for investing in technology solutions; eDiscovery tools have the potential to better automate and focus large-scale data collection, speeding and refining the scope of litigation response while reducing downstream costs for service providers and attorneys working with overly-broad swaths of custodial data.

It can, however, be difficult to determine current expenses and requirements in order to make an intelligent comparison. Because eDiscovery is traditionally outsourced, it may not be tracked consistently at an internal level or it may be billed back to the client with few line items for cost specified. Moreover, there are a number of critical risk factors involved in performing eDiscovery poorly, although these may only be realized once data has been mishandled and results in case setbacks, financial sanctions, and unfavorable exposure. This white paper will discuss factors involved in determining eDiscovery spending as well as the benefits and considerations involved in in-sourcing eDiscovery with a technology investment in the context of the EMC eDiscovery Benefits Calculator.

In-house Factors with the EMC eDiscovery Benefits Calculator

There are many factors involved in determining eDiscovery budgets and response methods. Each company and individual case has its own unique requirements, which vary in the financial impact and potential benefits garnered by making investments in the process. Because eDiscovery has traditionally been outsourced to law firms and service providers, enterprise litigants may be understandably unprepared or reluctant to manage all of its functions in-house—and may even be uncertain of their own costs, requirements, and processes.

¹ EMC offers an iPad/iPhone/Android Phone version of the "EMC eDiscovery Benefits Calculator" and a more detailed version that is only available through EMC sales.

² Source: Duke University, *Litigation Cost Survey of Major Companies*, 2010.

³ "Data Deluge," *the Economist*, 2010.

For the purposes of this writing, EMC eDiscovery Benefits Calculator will serve as a point of comparison for potential cost savings in enterprise eDiscovery. EMC's Benefits Calculator helps identify the cost differential between a fully-outsourced model, and discovery with the use of EMC SourceOne eDiscovery-Kazeon data management functions for early case assessment.

It's worth noting in a broader context that EMC SourceOne eDiscovery-Kazeon is available in multiple deployment options for the marketplace: it can be used transactionally through service providers, purchased for in-house deployment, or even purchased internally and operated by outsourced staff as a "managed service," depending on the requirements of the customer. This is important, as most enterprise customers, while they would like to increase efficiency and transparency in the process, may lack the immediate appetite or resources to initiate a new internal cross-functional business process around eDiscovery. Thus, staffing costs can be factored into the model through a combination of variable hourly rates and time spent on various tasks. These can be manipulated to reflect either internal or outsourced operation of the tool. Customers committing in-house staff to operating tools should potentially factor in additional salary, training, and even potential talent acquisition costs.

Litigation Profile and Investigative Approach

With the many options available, the initial step in considering eDiscovery methods from a legal standpoint is recognizing that every organization has a "litigation profile" based on the typical volume and types of investigations it experiences. This affects the approach it takes to eDiscovery and whether an in-house investment is justified.

Some baseline factors for considering an internal eDiscovery strategy include:

- What is the typical/anticipated volume of lawsuits, regulatory investigations, and internal investigations?
- What types of cases are common (labor dispute, IP dispute, class action) and what are the stakes in litigating?
- How many investigations involve eDiscovery overall?
- Are cases risky to the extent that outsourcing is preferable? Do they require specialized data handling?
- Could cases be resolved sooner if they were investigated in-house?

Organizations may prefer to handle day-to-day internal investigations in-house or may want to investigate low-impact lawsuits themselves at an early stage in order to determine case merits and potentially settle. EMC SourceOne eDiscovery-Kazeon offers early, pre-collection access to many data locations and types to allow this, plus functionality through first-pass review for routine investigations by in-house staff. The EMC eDiscovery Benefits Calculator incorporates optional predictive estimates of how many cases could potentially be resolved early on through the use of its product to illustrate the value of *avoiding* litigation, when possible.

Conversely, using service providers lets enterprise litigants outsource the risk surrounding eDiscovery self-collection to a third party, shielding the litigant from potential claims of process indefensibility. Criminal disputes, regulatory investigations, or cases involving specialized data handling (as in international collections) are typically outsourced to a service provider for better defensibility and outside counsel for review, even with an existing technology investment and established in-house process.

Data Retrieval Considerations

From an IT standpoint, there are fundamental logistical considerations for determining the costs involved in litigation response and the benefits of technology investment:

- What is the average volume of data on any given case, and overall annually?
- How many custodians are typically involved?
- What are the most frequent data locations and types required?
- How accessible is data in terms of cost and time to retrieve it?

These factors are more easily available or estimated for enterprise litigants and can be factored in as variables in the EMC eDiscovery Benefits Calculator for considering the benefits of using EMC SourceOne eDiscovery-Kazeon. Configurable variables are available for case sizes and frequency, typical custodian counts, volume of data, and different data locations and types supported by EMC, plus the time taken for collection and production tasks through typical methods. Part of the value proposition is EMC's over-the-network approach that centralizes and simplifies the process of identifying and collecting data from supported locations and types across numerous custodians.

The EMC eDiscovery Benefits Calculator also incorporates configurable per-GB costs for various tasks: collection and preservation, processing and analysis, review, and production. This obviously allows customers to compare costs from traditional methods to expenses using EMC SourceOne eDiscovery-Kazeon. It can also illustrate downstream costs potentially avoided by using EMC's more targeted data collection, or the benefits of avoiding litigation entirely by investigating in-house and making more informed strategic case decisions earlier in the overall discovery process.

While these numbers are useful and illustrative, it's important to remember that per-GB charges are not always obvious or readily available to enterprise litigants, whether they've been outsourcing the process or not. In the next section, we'll consider the less clear-cut costs of eDiscovery.

Comparing Current eDiscovery Costs

Determining Costs: Visible and Invisible

Quantifying and comparing the comprehensive costs/benefits of eDiscovery expenses and risk factors is more complicated. Enterprise litigants may have difficulty even accurately determining and itemizing current "spend" on eDiscovery for a number of reasons:

- Much of eDiscovery today is outsourced and specific charges may not be broken out as line items.
- Pricing options for both software and services are myriad and complex.
- Internal and external metrics around data volumes, timeframes, and man-hours are not tracked consistently.
- "Soft" organizational costs for spiky staffing demands, operational disruptions, and poor process management are difficult to quantify.
- Less apparent risk factors in performing eDiscovery poorly are often grossly underestimated, or may only be realized once they result in sanctions or irreversible consequences for a case.

The following sections will consider more complex or less obvious cost factors in eDiscovery not necessarily reflected in the EMC eDiscovery Benefits Calculator.

Pricing Options and Transparency

eDiscovery has traditionally been outsourced to service providers and law firms with budgetary expenses concentrated in a few broad areas: service provider fees, technology and hosting fees, and attorney document review fees.

In practice, there is little uniformity in the market around pricing. Actual services performed by outsourcers may or may not be broken out into individual line items and expenses depending on pricing, provider, and customer demand. Law firms may even bill service providers and technology used in a case to clients without specifying their costs, prompting questions around incentives they have to reduce them. A lack of transparency in services rendered and pricing can make it difficult to determine what is being charged and how or to draw an "apples-to-apples" comparison to alternate methods. Even the most transparent pricing options can be difficult to parse and compare in practice.

This is not due to a lack of options. For example, service providers might charge a combination of per-GB, per-document, or per-page fees varying by data type for technical functions like processing, culling, and production; hourly rates for forensic collection and litigation support staffing; and per-custodian rates for services like mailbox password-cracking and legal hold notification. Attorney review charges may be by the hour, by the page or document, at a flat rate, or some combination depending on how review is structured among different teams and review stages. Technology can be charged by the click, by the seat, by data volume processed, by processing compute resources used, by data volume culled and exported to a review tool, or at a monthly or annual per-GB rate for ongoing hosting.

This is only the first layer of pricing complexity. Add to this the fact that some providers have caps on volume that require an additional premium once a maximum is reached, or may price processing or collections at an “all-in” rate for one day of work, but require another full day of charges for iterative collections if the litigant must go “back to the well” to collect more data. Processing charges may vary before and after data is expanded from container files like PSTs since volume increases dramatically once embedded objects and attachments are extracted from decompressed files. Culling charges may be based either on the amount of data culled out (since it saves money on reviewing it) or the amount of data left over and promoted to full review. Similarly, more advanced analytics in review tools can incur additional charges.

Understandably, total costs in these models can be wildly unpredictable. Attempts to curb spending can even exacerbate the problem. For example, under-collecting at an early stage can result in more costly iterative collections downstream if something crucial is missed or custodians and date ranges expand significantly later. Similarly, culling seemingly irrelevant data prior to review can eliminate critical evidentiary context for attorneys examining documents downstream, resulting in longer review times or additional collections to clarify. In these instances, cost-savings gained by clear-cutting data, under-investing in knowledgeable staff, and underestimating the importance of ongoing collaboration between legal and IT discovery teams can actually *increase* the overall costs of discovery—particularly if they necessitate additional work for service providers and external law firms.

eDiscovery Technology Investment: The Value Proposition

Anecdotally, total eDiscovery charges can range anywhere from hundreds to thousands of dollars per gigabyte overall. Crunching the numbers can be difficult; various factors involved in the particular case, budgeting, organizational priorities, and IT systems contribute to choosing the most cost-efficient providers and tools for a given investigation.

The higher-level factors in controlling costs are not mysterious. Overall costs are limited by:

1. Getting the right data (and limiting irrelevant data) from the earliest stage possible.
2. Minimizing staffing costs and improving project collaboration between cross-functional teams.
3. Understanding key case metrics for volume and data contents to refine project scope and timeframes.
4. Limiting (or eliminating) the overall duration of a case.

eDiscovery solutions like EMC SourceOne eDiscovery-Kazeon can address several inefficiencies in this process with technical features and automation:

- Pre-collection “in-place” data identification and analysis allows litigation support teams to investigate the metadata and full-text contents of data prior to collecting and processing it. This potentially lets in-house counsel make a try vs. settle determination early in the process, eliminating further downstream costs of pursuing litigation.
- Advanced analytics, which when used in targeted collection and downstream filtering can reduce the amount of irrelevant data collected, preserved, and eventually reviewed. Often, solutions allow iterative, automated collection as investigation criteria are refined to lighten the burden on IT in responding to litigation requests.

- Advanced analytics and visualizations can provide insight into overall data volume and contents (by keyword, custodian, metadata, or other metrics) throughout the process, offering perspective on potential costs and duration of discovery as a case progresses.

The EMC eDiscovery Benefits Calculator specifically shows the impact of accessing, analyzing, and culling data from an early stage in limiting downstream costs, since this is where EMC SourceOne eDiscovery-Kazeon provides the most value.

“Risk Creep:” The Price of Poor eDiscovery

Organizational Costs and Inefficiencies

Whether it incorporates service providers or not, creating a codified internal business process for eDiscovery is the best way to optimize efficiency, both in the short-term administration of a case and in long-term benefits for managing future eDiscovery in a repeatable, predictable manner. The costs of poor process are not always easily quantifiable, but represent real pain and inefficiency that take a significant toll on an organization:

- Unpredictable budgeting and resource-planning for litigation response, often based on hourly rates and handling large data volumes “on-demand” when it occurs.
- Staffing resources for an ad-hoc, interrupt-driven process which interferes with day-to-day operations in multiple departments.
- Lack of process and collaboration between IT and legal team, resulting in duplication of effort, requiring multiple trips “back to the well” for iterative or incomplete collections and affecting productivity and case timelines.
- Information management costs for internal data preservation and logistical problems of coordinating legal hold with records management retention and deletion policies for large amounts of individual custodians and data volumes over years of litigation.

Creating an integrated business process around eDiscovery is a nontrivial pursuit given its cross-functional tasks, mission-critical status, and tight, no-notice timeframes but ESG believes it is the only path to ensuring more optimized and predictable costs around the process.

Additional Risk Factors

While some organizations invest heavily in service providers and top tier legal representation to outsource risk and ensure case success, other litigants devote little investment of any kind to eDiscovery, performing it with limited preparation or additional resources in order to reduce immediate costs. Companies using existing tools and staff in a tactical approach may perceive little opportunity for ROI in further investment. However, risk factors around under-funded or under-managed eDiscovery have their own potential financial consequences:

- Risk of accidentally producing privileged information to opposing counsel.
- Risk of sanctions for indefensible ESI preservation and handling.
- Risk of losing a case due to mishandling or outright missing evidence.
- Risk of losing reputation from losing or mismanaging a case.

Customers making do with ad hoc eDiscovery on a shoestring should weigh these risk factors and the potential financial fallout carefully against their own process in judging overall costs, since the additional risks and expenses of eDiscovery failure possible through mismanagement are not reflected in the EMC eDiscovery Benefits Calculator. The impact of these risks varies by case, organization, and litigation profile, but once incurred, their consequences are often irrevocable.

According to a report from law firm King & Spalding’s Discovery Center published in the Duke Law Journal, both the number of cases involving eDiscovery sanctions and the number of monetary awards have risen consistently in

recent years. Courts have dismissed cases or issued default judgments against litigants with eDiscovery lapses and given adverse jury instructions, issued penalties like witness preclusion, disallowed certain defenses, or reduced burden of proof. Monetary sanctions have ranged from \$2500 to \$8.8 million dollars.⁴

The Bigger Truth

Costs for eDiscovery start out big and typically only get bigger due to the unpredictability of litigation response. Even so, total costs are still easy to underestimate in terms of existing inefficiencies, potential risk, and overall poor process. Whether an enterprise customer utilizes outside service providers or not, it can optimize its eDiscovery spending through a more standardized internal business process. This often includes incorporating eDiscovery software to better automate and track the process and facilitate more intelligent collaboration between teams.

The first step to improving eDiscovery is determining current methods and requirements, visible and “invisible” costs, and opportunities for increasing efficiency and defensibility. EMC eDiscovery Benefits Calculator offers a solid basis for understanding organizational eDiscovery costs and potential savings as a point of informed comparison, and weighing the benefits of further technology investment.

⁴ Source: “Sanctions for E-discovery Violations: By the Numbers,” Dan H. Willoughby Jr., Rose Hunter Jones, & Gregory R. Antine, *Duke Law Journal*, 2010



Enterprise Strategy Group | **Getting to the bigger truth.**