Autonomy
IDOL Reference Architecture

Reference architecture for HP Autonomy’s IDOL Knowledge Management solution
# Table of contents

Introduction .............................................................................................................................................................................. 3
Overview: HP Autonomy’s IDOL Knowledge Management solution ................................................................. 4
Pre-deployment considerations ........................................................................................................................................ 5
Scenarios ................................................................................................................................................................................... 6
Scenario 1 – Single server ...................................................................................................................................................... 7
Scenario 1 – Single server sample bill of materials ........................................................................................................ 8
Scenario 2 – Enterprise environment .................................................................................................................................. 8
Scenario 2 – Enterprise environment sample bill of materials .................................................................................... 10
Conclusion ............................................................................................................................................................................... 11
About HP Autonomy ............................................................................................................................................................ 12
### Introduction

The Autonomy Intelligent Data Operating Layer (IDOL) platform allows you to automate the way you analyze and understand unstructured data, which makes up over 90% of enterprise data. HP Autonomy's IDOL Knowledge Management solution provides industry-leading search and analysis of data, with integrated retrieval and interaction capabilities. To deal with the massive amounts of content in modern electronic data volumes, the solution can easily scale to meet the needs of your enterprise.

This white paper provides an overview of the sizing and architecture required by the IDOL Knowledge Management solution to process, store, and present data using the system's industry-leading platform.

The sizing and architectures presented in this document are based on the HP Autonomy's IDOL technology and the HP ProLiant DL Gen8 server platform. The configurations have been developed by the HP Autonomy technical team to provide an enterprise-spanning IDOL solution that efficiently balances performance, storage, and cost.

**Among the basic functionalities provided by the IDOL platform are:**

- Automatic Query Guidance (AQG)
- Hyperlinking
- Highlighting
- Summarization
- Spell check and thesaurus
- IAS security (document / user)
- Keyword and exact phrase search
- Boolean and proximity search
- Parametric search
- Metadata search
- Conceptual search

**Additional IDOL Knowledge Management solution features include:**

- HP Autonomy Intelligent Universal Search user interface
- User personalization—e-mail and alerts
- Agent training
- Expertise profiling

This white paper can assist in the rapid design and deployment of the IDOL software on HP servers for data volumes and user bases of various sizes. Two deployment scenarios are explored here: a limited-in-scope single server IDOL deployment and a highly scalable enterprise IDOL deployment.
Overview: HP Autonomy’s IDOL Knowledge Management solution

HP Autonomy’s IDOL Knowledge Management solution is a massively scalable platform that provides insight into the volumes of unstructured data that can reside across your enterprise. The solution is designed to deal with hundreds of types of data and data repositories in a highly integrated manner, allowing you to effectively perform a wide range of activities such as discovering knowledge articles or documents, analyzing large volumes of data for a particular purpose such as eDiscovery, or locating content experts.

IDOL solutions deliver components essential to processing, understanding, retrieving, and interacting with data. These components can be grouped into the following basic categories:

- **IDOL Connector Framework Services** – The direct interface between IDOL and the various data repositories, CFS connects to over 400 repositories, providing a common interface for IDOL and GUIs to interact with data, abstracting repository-specific protocols and APIs.

- **IDOL server** – The underlying engine and data index of all HP Autonomy solutions, the IDOL server provides over 300 search functions and maintains an understanding of all unstructured content extracted from documents. This category also includes component services necessary for the distribution capabilities of the IDOL Server platform.

- **Graphical User Interfaces (GUIs)** – The web server that serves the end-user interface with the IDOL server, this component provides the end user with easy access to advanced IDOL functions, advanced mapped security, and simple-to-use interfacing with data, wherever the data may reside. This information is then presented in intuitive interfaces, increasing the productivity of end-users.
Below is a system diagram of an IDOL solution, detailing the communication that occurs between the component services. Each graphic can potentially represent a single software service running on a server, or a multitude of servers or databases performing the same task. For example, the Enterprise IDOL deployment below may have more than one server running several IDOL Connectors.

Pre-deployment considerations
There are a number of key factors you should consider before designing and deploying an Autonomy IDOL solution. The following subsections describe the design decisions required in creating a baseline configuration for each component.

Operating system
Autonomy IDOL solutions support many operating systems, including the Windows family of operating systems, as well as many flavors of Linux and Unix. Commonly deployed OS platforms include Windows Server 2008 and Red Hat Enterprise Linux 5.x/6.x. It is recommended that all servers used in a solution utilize the same OS for ease of management.

It is important to note that certain data types and data repositories are native to certain operating systems, and can only be properly processed and interfaced with the specific OS. For example, Microsoft Outlook PSTs and Microsoft SharePoint utilize APIs native to the Windows family of operating systems. Full details of supported OS platforms can be found in the corresponding IDOL installation and administration documentation.
Storage
When considering the number of disks a server requires, there are several factors to weigh:

• **I/O performance** – IDOL Index disk specifications indicated in each server section are for optimal performance. Substandard write speeds may impact the performance of initial index creation and synchronization of new content. Substandard read speeds may impact performance of operations on the index. Ideally, each IDOL engine should have access to its own disk or partition, independently capable of 120MB/s bandwidth even when all the engines are accessed simultaneously. Because IDOL generally performs tasks sequentially, IOPs are much less a consideration than bandwidth and vary by block size.

• **Disk redundancy** – HP Autonomy recommends no RAID, RAID 0, 1 or 10 depending upon the customer’s hardware redundancy level support requirements.

• **Storage capacity** – The size of the disks utilized for the Autonomy IDOL servers determines the capacity of each hardware unit. Due to bandwidth bottlenecks of conventional HDDs, it is recommended that the size of the partitions utilized for the IDOL services should not exceed 300GB.

Network
To properly interface with data repositories, Autonomy IDOL solutions require high-speed network access to the data repositories. In practice, this is accomplished by placing connectors physically “close” to the original data sources.

For communications between different IDOL components, such as IDOL distribution components to IDOL engines, a high-speed network allows index files to be passed quickly and for large result sets to be collated expediently. It is recommended that 1GbE Ethernet or higher be used.

Scenarios
Depending on the specifications of the solution, an IDOL Knowledge Management solution deployment can range from a single hardware unit to an environment consisting of multiple IDOL servers, a number of application servers, and a quantity of connector servers.

This reference architecture is designed around the main sizing considerations for the two example scenarios—single server and enterprise environment—and provides insights on how the solutions can scale. Sample architectures are provided as well as the hardware specifications needed to build out the solution.

The IDOL platform can easily scale to meet any system specifications. The metrics below detail some of the considerations taken into account for each scenario and in general for IDOL deployments.

**Initial data volumes:**
• Total number of data repositories
• Total number of documents across the repositories
• Total data volume (size on disk)

**Data update rates:**
• Regularity of rescan period
• Length of rescan windows
• Average number of documents updated/created per period

**System capacity:**
• Total number of users
• Expected number of concurrent users
• Hours of usage per day
• Mapped Security
• High Availability
Scenario 1 – Single server

For small or departmental deployments serving a small number of documents and a moderate number of users, it is most cost-effective to deploy the entire solution on a single server.

For example, if you have the following metrics:

Initial data volume:
- 2 local data repositories (for example a website or file share)
- Up to 20M documents across the repositories (based on industry average file size)\(^1\)
- Up to 2TB of total data volume (original document size on disk)

Data update rates:
- Rescan daily
- Rescan for 8 hours
- Up to 200,000 documents updated/created per period (network and repository dependent)

System capacity:
- 100 users
- Up to 10 concurrent users
- 8 hours of usage per day
- Mapped security
- High availability on disk only

In a single-server IDOL Search deployment, all services (see Figure 2) reside upon a single hardware unit. Resources on the server are devoted to specific services, segregating the functionality to the extent possible. High availability here is assumed to be done via RAID on the local disk configuration.

Predominantly, communication is internal to the server, with only communications to and from the appliance being to the data repositories and end users (via the application GUI).

Required specification:
- HP DL380p G8
- Intel Xeon processors (8 cores) or equivalent
- 32+ GB of RAM
- One (1) 146 GB partition for OS
- One (1) 146 GB partition for Connectors
- One (1) 146 GB partition for GUI Applications
- One (1) 146 GB partition for IDOL Components Distribution Services
- Four (4) 146 GB partitions for IDOL Engines
  - Each partition should have dedicated SAS drives or equivalent, when all drives/LUNs are simultaneously accessed, each should be capable of 120MB/s bandwidth and 180-200 IOPS (bandwidth being the more important metric)
- Note: All partitions should be in RAID1 configuration for disk redundancy (8 partitions created from 16 physical drives)
- Note: HP Autonomy does not recommend that this server be virtualized

---

\(^1\) A document is composed of a single object containing metadata and one or more pages of text as well as a single record from a structured source. For example a single Microsoft Word document, PDF document, or HTML page and a record from a single or multi-table database containing one or more values extracted as one unique row.
Scenario 1 – Single server sample bill of materials

<table>
<thead>
<tr>
<th>Qty</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>HP DL380p Gen8 25-SFF CTO Server</td>
</tr>
<tr>
<td>1</td>
<td>HP DL380p Gen8 E5-2609 FIO Kit</td>
</tr>
<tr>
<td>1</td>
<td>HP DL380p Gen8 E5-2609 Kit</td>
</tr>
<tr>
<td>2</td>
<td>HP DL380p Gen8 Intel Xeon E5-2609 (2.4GHz/4-core/80W) Processor Kit</td>
</tr>
<tr>
<td>4</td>
<td>HP 8GB 2Rx8 PC3-12800E-11 Kit</td>
</tr>
<tr>
<td>16</td>
<td>HP 146GB 6G SAS 10K 2.5in SC ENT HDD</td>
</tr>
<tr>
<td>1</td>
<td>HP Ethernet 1GbE 4P 331FLR FIO Adptr</td>
</tr>
<tr>
<td>1</td>
<td>HP 2GB FBWC for P-Series Smart Array</td>
</tr>
<tr>
<td>2</td>
<td>HP 750W CS Gold Ht Plg Pwr Supply Kit</td>
</tr>
<tr>
<td>1</td>
<td>HP Insight Control ML/DL/BL Bundle E-LTU</td>
</tr>
</tbody>
</table>

Scenario 2 – Enterprise environment

This scenario focuses on pan-enterprise deployments with large data volumes, a variety of data repositories, large user bases, or deployments that may need to expand in the future to deal with future data volume growth.

Software components are distributed across homogeneous hardware units designed to meet the specific needs of each component. These server-types are designated:

- Enterprise Connector Servers
- Enterprise IDOL Servers
- Enterprise GUI Application Servers

IDOL’s modular components provide flexibility in deployment and allow for scaling for each of the components in a linear, predictable matter. For each type of server in an enterprise environment, this document describes the scaling considerations with respect to data volumes, data updates, and system capacities as applicable to each server-type.

Enterprise Connector Servers – Scaling considerations

Connector Servers house one or more Connector Framework Services, which are used to interface with repository specific protocols and APIs, as described in the Solution Overview. Though little data is actually stored on these servers, each connector handles large amount of network and disk traffic, as the service scan for new data and process any that is found.

Assuming that a connector is continuously running, it is recommended that each connector have dedicated access to 2+ cores, 2+ GB RAM (more RAM if working with large files), and its own dedicated disk. Multiple connector instances can be collated into a physical server, in the example below, four (4) connector instances are fitted onto a single server. However, the physical server can be scaled larger or smaller based on the number of connector instances needed.

The following heuristics may be used to determine the number of connector instances required:

- One connector instance per repository-type
- A connector instance for every 200M in a repository-type
- A connector instance scans 500,000 over 24 hours (this is a conservative estimate as the figure is highly repository and network dependent).
- Replicating the system should High-Availability be required
Enterprise Connector Servers – Required specification

The specification below would support four instances of connector services.

- HP DL380p G8
- Intel Xeon processors (8 cores) or equivalent
- 8+ GB of RAM
- One (1) 146 GB partition for OS
- Four (4) 146 GB partitions for Connectors
- Note: SAN may be used in lieu of physical drives, however, note that connectors handle high volumes of transient data to disk.

Enterprise IDOL servers – Scaling considerations

The IDOL servers are distributed across multiple nodes in order to easily meet scalability demands and provide a uniform set of services for ease of management. The core component of the IDOL platform is referred to as the IDOL engine or the content engine. The IDOL engine stores and performs advanced analysis on the content extracted from documents, requiring dedicated processing power and memory to carry out these tasks. IDOL servers can be scaled first vertically by placing multiple IDOL engines on a physical server, then horizontally by adding more IDOL servers.

IDOL engine sizings can vary wildly on numerous factors including data, engine configurations, query composition, frequency and response requirements. However, for the purposes of the IDOL Knowledge Management solution, we will assume a corpus of documents from standard office productivity suites and industry average file sizes.

Autonomy would recommend additional hardware for IDOL servers be added in the event any of the following is the case with respect to the considerations listed above:

- One IDOL server for every 8TB of data in the data volume
- Replicating the system for every 250 concurrent users
- Replicating the system should high availability be required

Additionally, Autonomy would recommend one (1) additional IDOL distribution servers hardware unit be added for every eight (8) hardware units of IDOL server.

Enterprise IDOL Server – Required specification

- HP DL380p G8
- Intel Xeon processors (16 cores) or equivalent
- 128 GB of RAM
- One (1) 146 GB partition for OS
- One (1) 146 GB partition for IDOL distribution components
- Ten (10) 300 GB partitions for IDOL engines
  - Each partition should have dedicated SAS drives or equivalent, when all drives/LUNs are simultaneously accessed, each should be capable of 120MB/s bandwidth and 180-200 IOPS (bandwidth being the more important metric)
- Note: All partitions should be in RAID1 configuration for disk redundancy (12 partitions created from 24 physical drives)
- Note: HP Autonomy does not recommend that this server be virtualized

Enterprise IDOL Distribution Server – Required specification

- HP DL160 G8
- Intel Xeon processors (8 cores) or equivalent
- 16 GB of RAM
- One (1) 146 GB partition for OS
- One (1) 146 GB partition for IDOL distribution components
Enterprise GUI Application Servers – Scaling considerations

The Intelligent Universal Search and IDOL Explore Application Servers house the web servers end-users interact with. These are the only servers with direct access to the Enterprise IDOL Servers, and process the raw return into a format that is user-friendly. Additionally, these servers communicate regularly with the HP Autonomy Connector services as users retrieve the data they wish to access.

Autonomy would recommend additional hardware be added in the event any of the following is the case with respect to the assumptions above:

- Total concurrent users accessing applications exceeds 250 users
- Replicating the system should High-Availability be required

Enterprise GUI Application Server – Required specification

- HP DL160 G8
- Intel Xeon processors (8 cores) or equivalent
- 16 GB of RAM
- One (1) 146 GB partition for OS
- One (1) 146 GB partition for GUI Applications

Scenario 2 – Enterprise environment sample bill of materials

<table>
<thead>
<tr>
<th>Enterprise Connector Server</th>
</tr>
</thead>
<tbody>
<tr>
<td>Qty</td>
</tr>
<tr>
<td>-----</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>10</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td>1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Enterprise IDOL Server</th>
</tr>
</thead>
<tbody>
<tr>
<td>Qty</td>
</tr>
<tr>
<td>-----</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td>16</td>
</tr>
<tr>
<td>24</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td>1</td>
</tr>
</tbody>
</table>
Enterprise IDOL distribution server

<table>
<thead>
<tr>
<th>Qty</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>HP DL160p Gen8 SFF CTO Server</td>
</tr>
<tr>
<td>1</td>
<td>HP DL160p Gen8 Intel Xeon E5-2650 (2.0GHz/8-core/95W) Processor Kit</td>
</tr>
<tr>
<td>2</td>
<td>HP 8GB 2Rx8 PC3-12800E-11 Kit</td>
</tr>
<tr>
<td>4</td>
<td>HP 300GB 6G SAS 10K 2.5in SC ENT HDD</td>
</tr>
<tr>
<td>1</td>
<td>HP Ethernet 1GbE 4P 331FLR FIO Adptr</td>
</tr>
<tr>
<td>1</td>
<td>HP GB FBWC for P-Series Smart Array</td>
</tr>
<tr>
<td>2</td>
<td>HP 500W Pwr Supply Kit</td>
</tr>
<tr>
<td>1</td>
<td>HP Insight Control ML/DL/BL Bundle E-LTU</td>
</tr>
</tbody>
</table>

Enterprise GUI application server

<table>
<thead>
<tr>
<th>Qty</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>HP DL160p Gen8 SFF CTO Server</td>
</tr>
<tr>
<td>1</td>
<td>HP DL160p Gen8 Intel Xeon E5-2650 (2.0GHz/8-core/95W) Processor Kit</td>
</tr>
<tr>
<td>2</td>
<td>HP 8GB 2Rx8 PC3-12800E-11 Kit</td>
</tr>
<tr>
<td>4</td>
<td>HP 300GB 6G SAS 10K 2.5in SC ENT HDD</td>
</tr>
<tr>
<td>1</td>
<td>HP Ethernet 1GbE 4P 331FLR FIO Adptr</td>
</tr>
<tr>
<td>1</td>
<td>HP GB FBWC for P-Series Smart Array</td>
</tr>
<tr>
<td>2</td>
<td>HP 500W Pwr Supply Kit</td>
</tr>
<tr>
<td>1</td>
<td>HP Insight Control ML/DL/BL Bundle E-LTU</td>
</tr>
</tbody>
</table>

Conclusion

HP Autonomy’s IDOL Knowledge Management solution allows customers to develop new insights from their unstructured data by providing a platform for industry-leading search and analysis of data, with integrated retrieval and interaction capabilities. IDOL's flexibility and scalability enables a variety of configurations and architectures, and this white paper provides several references and examples for deploying IDOL Knowledge Management solutions. Additionally, this document can guide solution architects in the rapid development and deployment in scenarios with data volumes and system capacities of various sizes.
About HP Autonomy

HP Autonomy is a global leader in software that processes human information, or unstructured data, including social media, email, video, audio, text, and web pages, etc. Autonomy’s powerful management and analytic tools for structured information together with its ability to extract meaning in real time from all forms of information, regardless of format, is a powerful tool for companies seeking to get the most out of their data. Autonomy’s product portfolio helps power companies through enterprise search analytics, business process management, and OEM operations. Autonomy also offers information governance solutions in areas such as eDiscovery, content management, and compliance, as well as marketing solutions that help companies grow revenue, such as web content management, online marketing optimization, and rich media management.

Please visit autonomy.com to find out more.