EXALEAD'S CLOUDVIEW

A New Challenger in Enterprise Information Access and Search

Overflight

Exalead, an information access company, has differentiated itself from the hundreds of companies licensing search, content processing, and text analysis. The company recently unveiled its CloudView product line. Does Exalead have the infrastructure and engineering to deliver next-generation information access across multiple platforms? The quick answer, based on ArnoldIT.com's research is, "Yes." This white paper discusses the company's strategic technology advantages and identifies potential uses of Exalead's technology.

Company Snapshot

Exalead S.A. is privately held. The company's headquarters are in Paris, France. The firm's U.S. office is located in San Francisco, California. The firm's United Kingdom office is located in Glasgow, Scotland.

Selected customers include:

- ARF (USA): Exalead CloudView - Improved knowledge delivery service for marketing decision makers
- Advantage West Midlands (UK): Exalead CloudView - internal search facility for file system data, Business Objects data, email and email archive (Zantaz).
- Coface Services (France): Exalead CloudView - extranet encompassing 100 million records
- Fish4 (UK): Exalead CloudView - Matching candidates' résumés to job openings
- Friendster (USA) – Social network search application
- GEFCO (France): Exalead CloudView – Logistics "track and trace" application
- IBM PriceWaterhouseCoopers (France) : Exalead CloudView and Exalead Desktop - enterprise search
- Skyrock: Exalead CloudView – Business intelligence system to access to more than 500 million Web logs
- Taya IT (Arabic): Exalead CloudView - Arabic language search portal
- ViaMichelin (France): Exalead CloudView- Points of Interest search and localization

Exalead serves more than 180 customers worldwide. The company has more than 80 employees in offices in France, the United Kingdom, and the United States. The company's annual turnover is in the US$20 million range. The firm is at or near break even. Growth has averaged 100 percent per year for the last three years.

Technical Foundation: Modern, Agile, and Scalable

In the early 1990s, Digital Equipment Corp., makers of the ground breaking mini-computer and the revolutionary DEC 10, had devised a multi-core processing central processing unit or CPU. The Alpha chip bundled a number of innovations into one poker-chip sized piece of silicon. DEC decided to demonstrate the power of 64 bits, a flexible memory architecture, support for multi-threading, and speed.

Digital Equipment gathered together the top engineers and computer scientists and put them to work. Innovation after innovation flowed from DEC's research labs into AltaVista.com. The system provided online translation, an index to content in newsgroups, advanced search support of Boolean operators, and relevance ranking that worked.

By 1997, AltaVista.com--the name of the DEC search system--had become the number one search engine. Engineers and research scientists realized that large-scale tasks such as indexing tens of millions of pages and delivering queries to hundreds of thousands of users.

The AltaVista.com service had a Palo Alto, California nerve center. But by 1998, DEC had become Compaq Computer. Compaq Computer, in turn, was purchased by Hewlett-Packard. HP had little interest in making proprietary CPUs or running a Web search system.

Many of the brightest minds from the AltaVista.com project left the company. A number of engineers gravitated to Google. François Bourdoncle, trained at L'École Polytechnique, returned to Paris and began work in 2000 on the information access system that today is Exalead.

The knowledge gained from AltaVista.com's successes and failures had a profound influence on Dr. Bourdoncle's venture. François Bourdoncle is a pioneer of the search engine software market. He has an extensive background in engineering and research and development and played a leading role on the LiveTopics project for AltaVista while at L'École des Mines de Paris. He was also a senior researcher with Digital Paris Research Laboratory and Digital Systems Research Center in Palo Alto, California.

The advantage is that Exalead is one of the most up-to-date systems in terms of its architecture. In the opinion of the Beyond Search team, Exalead is nose-and-nose with the likes of Google.

Four Building Blocks

Exalead approached search from an engineering base. In an interview with Dr. Bourdoncle, he identified four foundation stones for the Exalead system, which is called CloudView:

1) Commodity hardware
2) Smart software
3) Snap in scaling
4) Adaptable architecture

Let's look at each of these building blocks.

Commodity Hardware

Anyone can contact a Hewlett Packard or Sun Microsystems' reseller and purchase high-speed, exotic hardware. These companies charge a premium for their engineered products and their 24x7 repair service. The combination of a brand name and an engineer on call was the only way to buy servers, routers, and other plumbing for a mission-critical computing system.

The advent of sharply reduced costs for CPUs, memory, and hard drives led to a new computing paradigm. Low cost servers made it possible to create redundant systems. Lower storage costs eliminated the need for unreliable and expensive tape back up systems. Redundant servers and storage devices could deliver reliability by allowing a device to fail and automatically switching to a back up system. The savings in hardware and other devices can be significant, often 25% of the cost of the exotic, brand name systems. As a result, designing a system for 99.999996 uptime with lower cost hardware delivers speed, flexibility, and performance. Instead of waiting weeks for a replacement device, a server can be swapped in minutes.

The Exalead system makes use of such advanced engineering principles as redundancy, data replication, and software configurable devices. The result is that Exalead reduces the cost of processing large amounts of data. When additional capacity is needed, commodity hardware eliminates the bottleneck created when a brand name vendor has to manufacture a replacement part.

Like Google and a handful of other companies in the information access business, Exalead has developed systems expressly for performance, reliability, scalability, and lower operating costs. Even with the sharply reduced prices now offered by brand name vendors, Exalead delivers at a significantly lower hardware cost.

Smart Software

Exalead's scientists and engineers are careful to explain that the code is "smart". The idea is that Exalead's system does the work that once a subject matter expert or technician had to perform. Dr. Bourdoncle does not use the phrase "artificial intelligence" nor does he talk about "automatic learning systems." The idea is that modules of Exalead's system adapt to the information flowing through the system and to the actions users take. "The combination," he says, "allows Exalead to deliver actionable information without the time delay imposed by manual tuning and knowledge base editing."

The secret at Exalead is mathematics. L'École Polytechnique is well known for its strong commitment to mathematics, algorithmic thinking, and number theory. The influence of those mental models has left a clear imprint on the Exalead system. If Henri Poincaré or

André-Mariue Ampère were to visit Exalead, both of these Polytechnique graduates would recognize the Exalead approach.

The smart software is implemented within the rigorous environment of software engineering best practices. Exalead's engineers are young, but these men and women document their code, follow the engineering practices Dr. Bourdoncle refined at DEC's research labs in France and America, and discipline of professional programmers. Specific engineering best practices include:

- Organized error handling
- Well specified features
- Automated quality assurance procedures.

Run-and-gun code is all too common at some content processing companies. Exalead is an engineering firm focused on information access and solutions.

Snap In Scaling

The technical literature and even popular media are awash in stories about cloud computing and green data centers. Microsoft is in the midst of a wrenching and expensive transition from its traditional approach to data centers and systems architecture. The August 2008 technical article for an ACM publication makes it clear that Microsoft has figured out that AltaVista.com's architecture makes sense in today's data center. Exalead is more advanced than Google when it comes to enterprise architecture. Microsoft, as of October 2008, is several years behind Google and lags Exalead in enterprise and cloud computing.

Exalead was untouched by the Google Web indexing and advertising business. In fact, Exalead implemented the innovations documented in "Towards a Next Generation Data Center Architecture: Scalability and Commoditization" when it set up operations in Paris in 2000.

Other innovations related to scaling include:

1) Enhanced file read speeds to accelerate query processing
2) Built in connectors for common file systems and file types, including Microsoft Exchange and Lotus Notes
3) Fast content processing which adapts to content volume and size of documents
4) On-the-fly index updating to ensure that search results are fresh.

Adaptable Architecture

Exalead's system can run in a range of configurations to meet licensees' specific requirements. CloudView is also available as a hosted service. The Exalead system can run on the customer's hardware and as a standalone system or integrated with other enterprise software and systems.

Google builds technology for its Web indexing and advertising business. The enterprise initiative is an afterthought. A separate engineering team has the task of creating a version of Google's Web search that is suitable for enterprise use in the Google Search Appliance. Microsoft's multi-billion dollar rebuilding of its global data centers must support downloads, license authentications, online applications, and other tasks related to moving desktop applications built for on premises installation and use to the cloud. Neither of these massive organizations is as focused as Exalead on building the next-generation, best-of-breed information access platform.,

Exalead has invented, refined, and deployed what it calls "adaptable architecture." Dr. Bourdoncle said:

> It is okay to build an architecture to do one thing like Google. It is not so good to build an architecture that has to do what desktop applications did on a designated PC but in a completely different cloud environment. We did not do what Google or Microsoft did. Instead, we focused on engineering a system that can process large amounts of data on a small number of machines. When you want to move those machines to the cloud or move the cloud to your office, no problem. We can also integrate the CloudView architecture into a heterogeneous computing environment. That's no problem for us either. The result is that with Exalead we have what we call the 'adaptable architecture'.

The diagram below shows a simplified block diagram for the CloudView system.
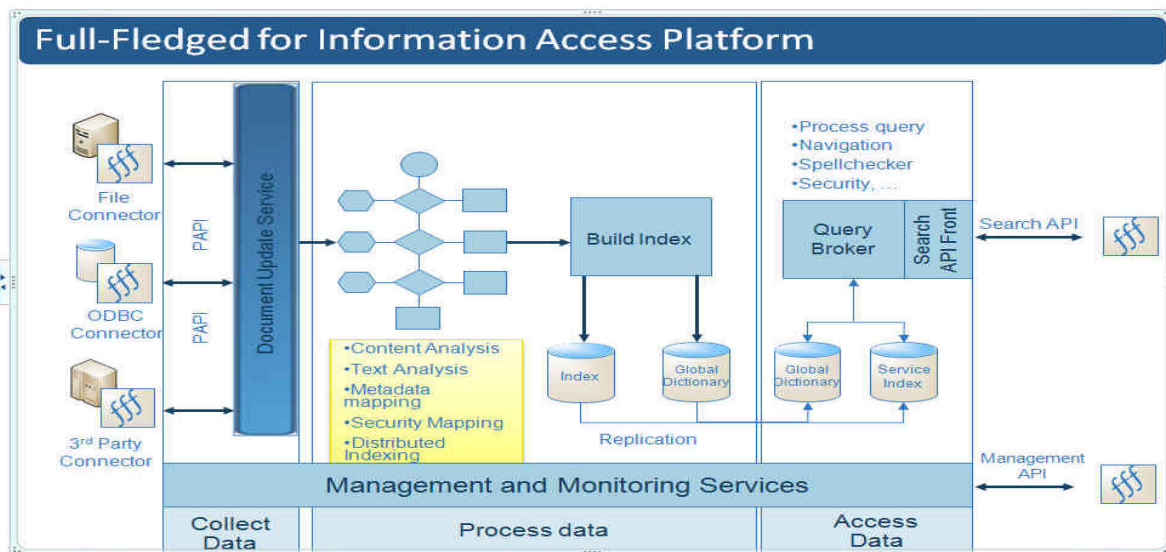
The advantages of the Exalead approach focus on a low total cost of ownership (TCO) and quicker time to value for the licensee. Plus, the approach includes:

1) Reduced costs because commodity servers perform functions ordinarily requiring expensive, specialized equipment.

2) Snap in scaling which delivers two benefits. First, less staff time is required to deploy, configure, and integrate into the existing infrastructure and the hardware itself is a commodity. The price differential between name brand servers and commodity servers can range from a savings of 25 percent to as much as 75 percent per device.
3) Smoother pipelines for data in and data out. The result is that latency is reduced without having to pay for additional caching (an expensive approach), dedicated hardware to work around congestion associated with database access, and large numbers of machines that cannot be fully utilized due to flawed parallelization and threading.

Product Options

Exalead CloudView is available under three editions: OEM, Search and 360. It is a software platform that automatically collects, structures and contextualizes high volumes of unstructured and structured content scattered across the enterprise information cloud, helping customers either improve information search, discovery and management or build extended business applications.

Exalead is reshaping the digital content landscape with a platform that uses advanced semantic technologies to bring structure, meaning and accessibility to previously unused or under-utilized data in the disparate, heterogeneous enterprise information cloud. The system collects data from virtually any source, in any format, and transforms it into structured, pervasive, contextualized building blocks of business information that can be directly searched and queried, or used as the foundation for a new breed of lean, innovative, next-generation applications with built-in, self-service reporting capabilities.

Representative Customer Uses of Exalead's System

The Exalead difference in hardware and software architecture translates to a number of useful, needed solutions for its customers. Let's look at three.

Data Management

With the volume of data in an organization doubling every six months, data management is emerging as a significant challenge. If one views the Exalead architecture as a data management solution, an organization can use the Exalead servers and software on premises or as a cloud service to capture, index, and serve structured and unstructured data to users as well as other applications. In short, Exalead's approach to architecture makes it possible to replace a traditional Codd database (Oracle RDBMS, IBM DB2, or Microsoft SQL Server). The Exalead system can output reports, business data in subsets, or as "answers" to questions. Companies like Aster Data and InfoBright have already moved ahead of traditional database solutions. But these companies, despite their next generation data management tools, lag

behind Exalead. Exalead offers both the hardware architecture and the software for next generation data management.

Business Intelligence

Exalead's system provides real-time monitoring, alerts integrated into work flows, and a real time junction operation combine to keep a licensee's data fresh, available, and in the bloodstream of the organization. A leading logistic management company in the United Kingdom coordinates its global operations using CloudView. The company's employees in more than a dozen countries can determine the status of a customer's order and access related information such as the manifest and accounting information for each shipment. This system provides integration of text and formatted records with analytics. Latency between information being known to the system and that information's availability to a decision maker or another system process is less than one second.

Active Knowledge

For a major property management and leasing firm, Exalead provides a data management and information access platform. The Exalead system provides information to agents and provides near real time information to the company's Web site. The Exalead system manages images, textual information, and structured data. A user has a single view of the available information pertinent to a particular property or project.

Other CloudView Uses

Information Access

The Exalead system can make available from a graphical interface information from different systems, repositories, programs, and sources. The user can either submit a query or click an icon that launches a stored procedure. The user gets the needed information. The options for displaying the answer can be configured as a report, a graphic, a single line of text, or a results display with hot links to related information. Configuration options are a matter of pointing and clicking via the Exalead graphical interface or, if one prefers, by editing scripts.

Decision Support

With near real time data acquisition, the Exalead system can generate and route alerts to users. In a government agency or financial institution, timely information must reach a specific person when an event occurs. The Exalead system "watches" newly arriving data and immediately flags, formats, and pushes that information to a designated user. The Exalead system supports a range of alert options, including SMS (text messages), email, and updates to a personalized Web page displayed in a user's work flow session.

Expert Identification

The content processing and metatagging methods used by Exalead support expert finder operations. A professional working in a consulting firm needs to know who in the firm has a particular area of competence. Exalead can match a key word phrase, display a list of

suggestions, or generate a customized interface to make it possible to "discover" people with particular skills. Exalead's metatags can be used in concert with a specialized monitoring system such as Tacit Software's or Exalead's CloudView can be used as the expert finder and relationship discovery engine in an organization.

Support: Direct from Exalead Professionals

One of the issues organizations confront when adopting any enterprise solution that touches information is support. Exalead has a commitment to its licensees that is different from that offered by some vendors. A recent article in The Independent, October 8, 2008, by Rhordi Marsden reveals that Google does not return telephone calls or answers email.

Exalead does. It's a simple fact that Exalead provides each licensee with the contact information to reach an Exalead engineer for technical support. At a client's request, Exalead can train a licensee's in house team to perform most Exalead functions. If the licensee wants on site support, Exalead will place an Exalead engineer on the licensee's premises or in the licensee's data center.

Other facets of Exalead's support include:

- Customization and engineering consulting
- Training
- Integration services
- Operations support

Online Demonstrations

Exalead provides an online demonstration of the CloudView system. You can navigate to www.exalead.com and run queries against the more than eight billion Web pages the company has indexed.

ArnoldIT.com has tested the Exalead system for its Overflight real time intelligence monitoring system. You can search the full text of every Google Web log entry for the last decade using Exalead. The system automatically assigns metadata such as date information and extracts named entities. The output from a typical query appears on the following page.

The features of the ArnoldIT.com Overflight test include continuous updating. When the Beyond Search analysts compared the results from Google's own search of its Web log content, the system did not perform correct date clustering. In the figure on page 10, the entity extraction, the expanded synopsis of each document in the results list, and the optional page view [represented as a question mark for a Web log item] makes it easy to tap into Google's own, lesser known public information about its system and products.

Although Exalead is not currently looking for funding, the 2008 U.S. launch could be enhanced through a strategic partnership whose goal would be to accelerate the market penetration thanks to a better access to the market and the key executive decision makers, whose materialization could be additional funding from the partner. Partnership may then take two forms: VC or financial institution to help us gain access to key decision makers for opportunities in the U.S. market, and major strategic vendor which could complete his offer and deploy with us in a synergistic way. Selected Exalead resellers and/or integrators include Logica, Sodifrance, Acamaya, Edifixio, Adequat, Sedona, Sogeti (Capgemini),Retis, Netbureau, BNT Holland, YDS s.r.o., ODEON AST Ltd, Todayisnow, Deliverance Pte Ltd, 4D Concept, Wand, eSpirit, EntropySoft, Taya IT. The company's technology partners include DataDirect Technologies, EMC, HP, IBM, Lingway, Messaging Architects, Systran, and Vecsys.

Exalead's revenues for calendar 2008 are estimated to be about $20 million. Exalead's channel strategy is to distribute products through both direct and indirect sales. Exalead has a professional services team. Exalead offers an original equipment manufacturer version of its CloudView system designed for integration into other products and systems.

Exalead represents a next-generation information access option for organizations worldwide. The company's technology, architecture, and software move a licensee "beyond search" to make information a vital, seamless component at a compelling price point beginning at about US$50,000.

Prepared by Stuart Schram IV, Analyst, Beyond Search—January 6, 2008