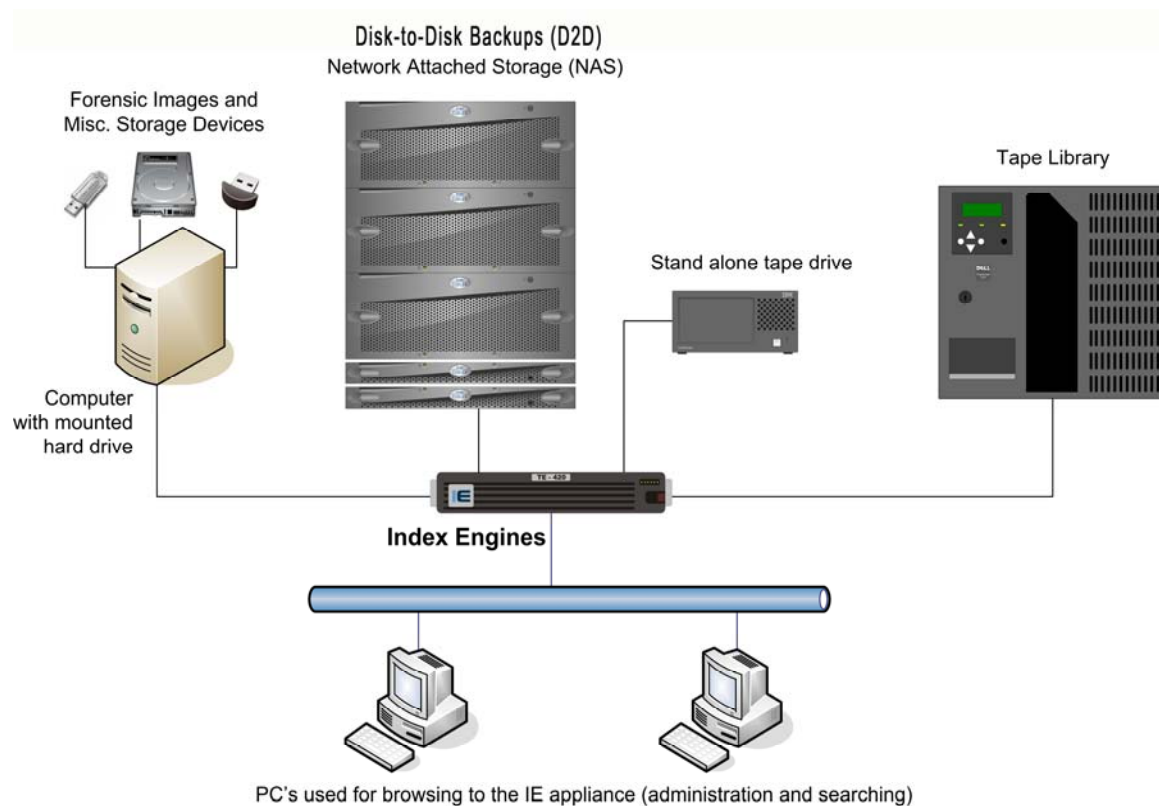


Patented Platform Explained

Index Engines - Performance for the Enterprise

The Index Engines platform indexes unstructured and email data unprecedented speeds and with extreme scalability. Architected to keep up with backup speeds and high speed networks, Index Engines makes discovery of hundreds of TBs of enterprise data achievable. The intelligent engine can process multiple indexing streams in parallel, at speeds of 1TB/hour using a single node. The Index Engines platform is designed with maximum flexibility allowing the platform to index content from any source including LAN, Tape, VTL, and D2D systems. In addition to the complex unstructured and semi-structured formats like Exchange and Lotus Notes, the Index Engines platform understands and can directly process NDMP input and most backup formats. All of these file types and containers can be indexed at unprecedented speeds on an economical platform. How Index Engines achieves these capabilities is explained below.

Typical Configuration



Enterprise Class Indexing

Index Engines purpose-built indexing operating system was designed from the ground up to meet the demands of the enterprise. Unlike traditional project-based indexing products, Index Engines has delivered enterprise class processing speeds, scalability and access to proprietary environments.

All components of the platform, including the database, word scraping, and query engine have been developed by Index Engines in order to meet Enterprise speed and scalability requirements. Traditional indexing products that use third party and open source components, typically designed for Internet-class indexing, fail when faced with terabytes of data using a single node.

The success of the Index Engines architecture has been proved based on the efficient processing of enterprise data. Full content and metadata indexing is performed at 1TB/hour/node. The resulting index footprint is only 4 to 8% of the original data. Traditional indexers cannot claim the same speed and efficiency since they have not invested in an architecture that scales to meet the challenges of the enterprise.

Sequential Processing

As data streams to Index Engines it is scanned and processed sequentially in order to maintain high speed throughput. Sequential processing of streaming data is a core component of Index Engines intellectual property and thus a unique advantage of the platform. Traditional indexing requires random disk access in order to process data. Random access to disk is slow because of unavoidable hardware limitations.

Without sequential processing of streaming data enterprise-class indexing cannot occur within reasonable timeframe and with a realistic number of processing nodes. Traditional vendors will claim to meet these requirements; however they will be forced to utilize dozens of processing nodes just to keep up. 1TB/hour per node is not a speed metric that random disk access can achieve.

Leveraging Purpose Built Protocols

Data can only be processed as quickly as it can be scanned. Index Engines has taken advantage of higher speed protocols for access to enterprise data. NDMP was developed as a high speed protocol to efficiently dump bulk data from a NAS device over the network. NDMP is significantly faster than CIFS or NFS for accessing bulk data over the network. Using NDMP, data can stream at speeds of up to 100 MB/second per GigE link. Multiple streams can be configured to process 1TB/hour. Traditional network protocols, such as NFS and CIFS, are supported by Index Engines but more simultaneous streams are required for fast processing speeds. Index Engines platform supports sustained indexing using NFS/CIFS of 800GB/hour/ node, 80% of the NDMP indexing performance.

Index Engines knowledge of proprietary backup formats and the proprietary sequential processing technology make it possible to directly process Tapes, VTLs, and D2D subsystems directly at these high speeds. With the knowledge of backup formats these subsystems can be read and managed directly to create detailed knowledge of backup content.

Architected for the Enterprise

Beyond the speed of processing, a true enterprise-class indexing platform needs to be scalable and adaptable to complex IT environments. In addition to unprecedented processing speeds, Index Engines has focused on architecting an efficient platform to store the resulting index. Traditional indexes can reach sizes of 20 to 100% of the original data source, turning indexing activity into a storage challenge. Index Engines index footprint is 4 to 8% of the original data source. This allows for extreme scalability, supporting up to 1 billion files/emails per node.

Gaining access to proprietary environments is the final key to enterprise indexing. Email is locked away in proprietary databases making it difficult to access. Using standard tools to gain access to this data (MAPI for Exchange as an example) is slow and cumbersome. Index Engines employs its sequential processing technology to these databases allowing for significantly faster processing of the content.

Choosing Index Engines

Enterprise-class indexing cannot be achieved using traditional approaches. Index Engines has invested significant time and resources in building a true enterprise class indexing platform. This

system has been purpose built to deliver detailed knowledge of enterprise content to support key business initiatives. Index Engines is the right approach for managing all enterprise data quickly and efficiently.

Contact Index Engines for more details at:

Web: www.indexengines.com

Email: info@indexengines.com

Phone: 732-817-1060

Address: 960 Holmdel Rd., Holmdel, NJ 07733