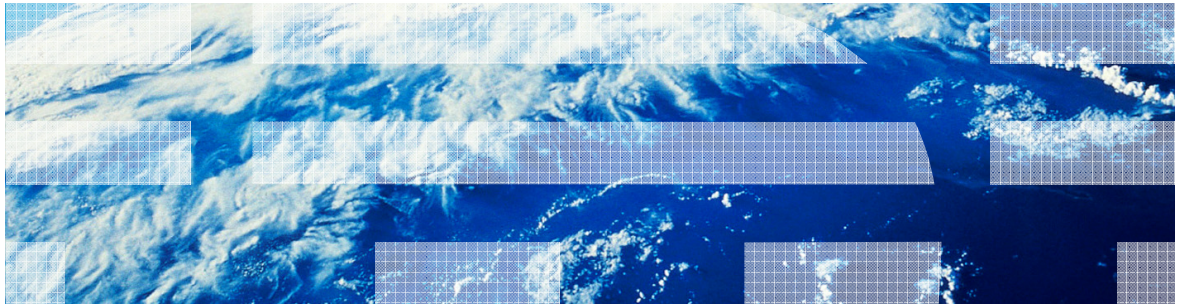
Searching Object Stores



Searching

- What do users search for?
  - Documents
  - Folders
  - Custom objects stored in a P8 system
- Search methods
  - Simple searches
  - Search templates
  - Stored searches
- Search results
  - Each search method returns results based on configured criteria
- Two types of searches
  - Property-based searches
  - Content-based retrieval

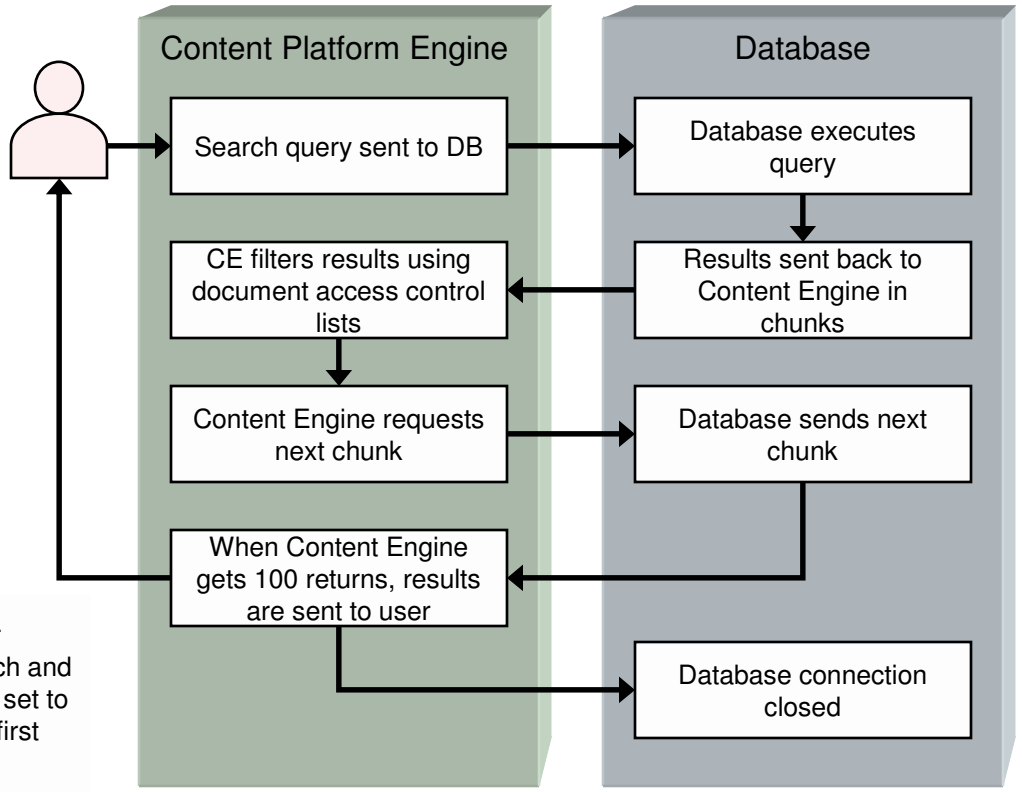
What is property-based searching?

- Property-based searching
  - Is searching for items based on the value of assigned properties.
  - It applies to documents, folders, and custom objects.
  - Properties and target values can be specified in the search criteria.
  - Criteria can be combined to narrow search results.
  - It can use custom- or system-defined properties.
- Examples of property-based searches:
  - Find all documents that were added by a certain user after a certain date.
  - Find all folders that have the word “sales” in the title.

Content Engine and the database

- All document metadata is stored in a database
  - The Content Platform Engine generates database queries whenever documents are accessed
  - Examples: Performing a search, opening a folder
- Many Content Engine changes affect the database
  - Adding a custom property
    - This action adds one column to the document class database table.
  - Adding a stored search or search template
  - Adding a new single index

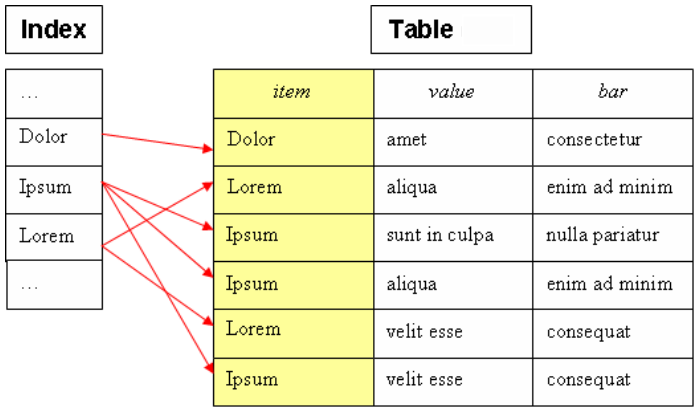
How a user request can affect the database



Example: A user performs a search and has preferences set to display only the first 100 documents.

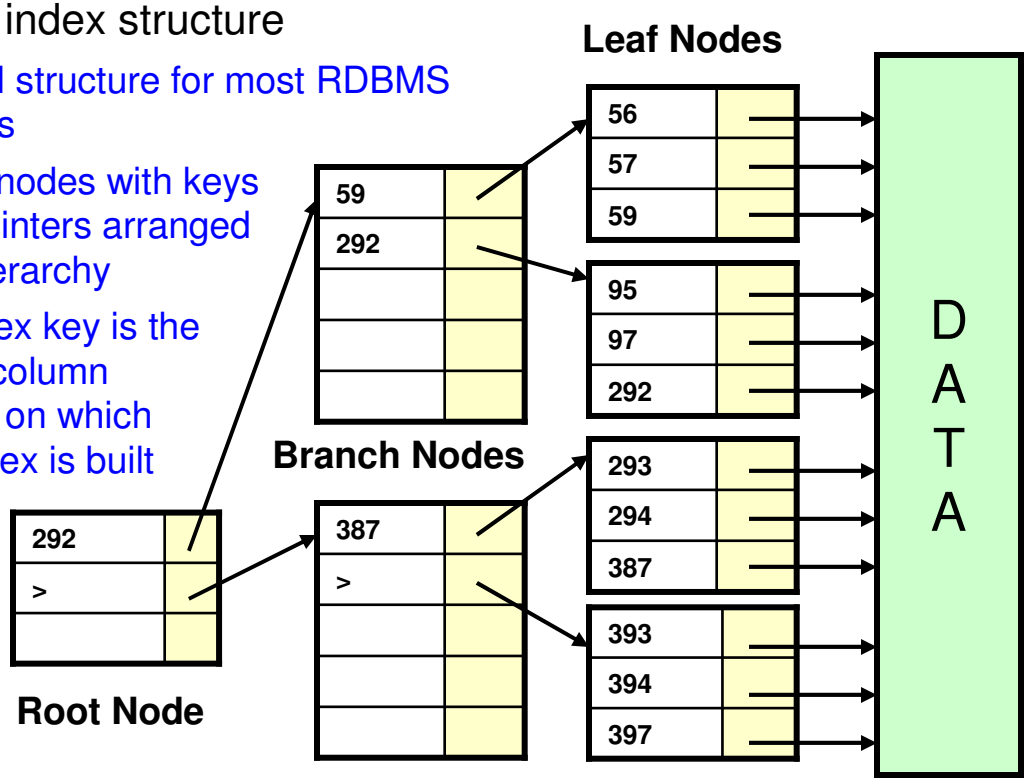
What are property indexes?

- Database indexes
  - Stored in the object store database.
  - Can greatly reduce the time required for a query on the indexed property (from several minutes to only seconds).
  - Can be created by the database administrator based on database statistics.
- Efficiency considerations
  - The number of indexes must be balanced for efficiency.
  - Some indexes might be dropped for faster bulk uploads.



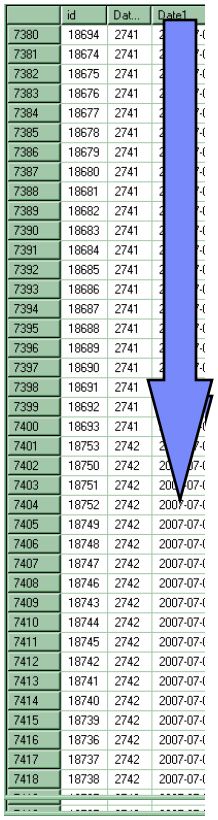
How does a property index work?

- B+ tree index structure
  - Typical structure for most RDBMS indexes
  - Set of nodes with keys and pointers arranged in a hierarchy
  - An index key is the set of column values on which the index is built



Locating data without an index

- Example
  - Locate all purchase orders for product ID 18677
  - A full *table scan* is required. All rows in the table must be searched.
    - Each row is examined to determine whether it satisfies the statement's WHERE clause.
- Full table scans
  - Are *sometimes* more efficient than index range scans.
    - Database optimizer determines that the query will access most of the blocks in the table although an index exists.
  - Use larger I/O calls.



	id	Dat...	Date1
7380	18694	2741	
7381	18674	2741	
7382	18675	2741	
7383	18676	2741	
7384	18677	2741	
7385	18678	2741	
7386	18679	2741	
7387	18680	2741	
7388	18681	2741	
7389	18682	2741	
7390	18683	2741	
7391	18684	2741	
7392	18685	2741	
7393	18686	2741	
7394	18687	2741	
7395	18688	2741	
7396	18689	2741	
7397	18690	2741	
7398	18691	2741	
7399	18692	2741	
7400	18693	2741	
7401	18753	2742	
7402	18750	2742	
7403	18751	2742	
7404	18752	2742	
7405	18749	2742	
7406	18748	2742	
7407	18747	2742	
7408	18746	2742	
7409	18743	2742	
7410	18744	2742	
7411	18745	2742	
7412	18742	2742	
7413	18741	2742	
7414	18740	2742	
7415	18739	2742	
7416	18736	2742	
7417	18737	2742	
7418	18738	2742	

Indexing properties

- Database index
  - You can create a database index for any property definition (except a system-owned property).
  - Known as a *single index*.
    - Example: Employees always search for case numbers, so you need to index the CaseNumber property.
  - Do not confuse with full-text indexing used for content-based retrieval.
- Single indexes
  - Creation of a single index for a property increases the time required to add documents to an object store.
  - Use single indexes sparingly and only for search criteria.
  - Best used on unique-valued properties.
  - Discuss single indexes with your database administrator.

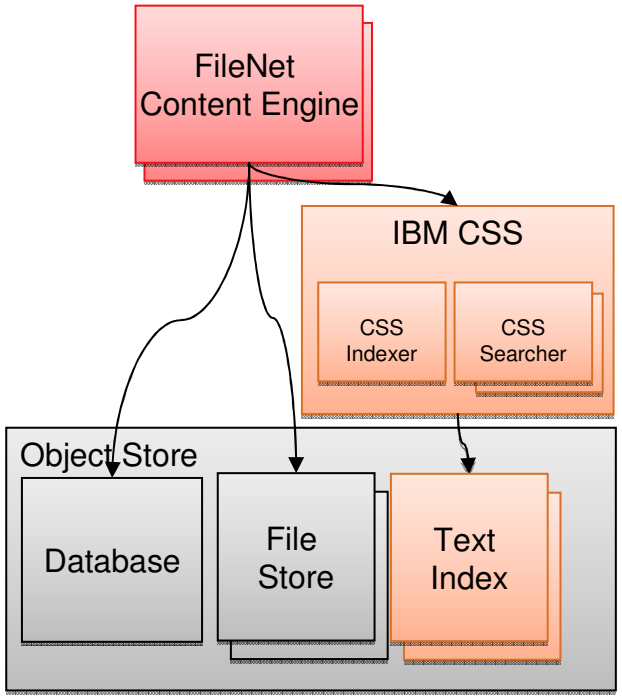
Property indexing guidelines

- Do not manually edit IBM FileNet database cells.
- Create stored searches and search templates for users.
  - Ensure that they search on indexed properties.
  - Prevent users from creating open searches.
- Create single indexes when needed.
  - Discuss any indexes with the database administrator.
  - Create single indexes for properties that are searched often.
  - Do not create single indexes for every custom property.
  - The database administrator can index system-owned properties, if needed.

What is content-based retrieval (CBR)?

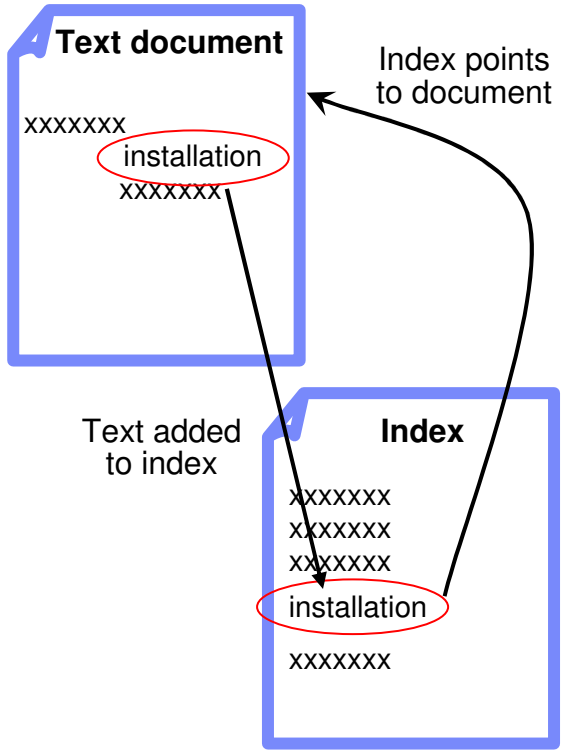
- Searches content, annotations, and properties for
  - Words, phrases
  - Words in proximity
- Supports most document types:
  - Microsoft Office documents, PDF, HTML, ASCII, and so forth
  - Can search in XML tags.
- Most Content Platform Engine search utilities support CBR.
  - Stored searches, search templates, Content Navigator, Workplace XT, and Query Builder
- Documents must have been indexed.
  - IBM Content Search Services provides full-text indexing.

Content Search Engine Architecture



What is a content index?

- Content index
  - File that contains pointers to the character-based content in an object store
  - Like an index in the back of a book
  - Instead of page numbers, index stores references to documents.
- What is searched?
  - Content Based Retrieval searches the index file, not the actual documents



Considerations for creating searches

- Is searching easier or faster than browsing?
  - Can all users browse?
- Why are users searching?
  - What business need is being addressed?
  - Do different users have different needs?
- Where should users look?
  - Which object stores, subfolders?
- What are users looking for?
  - Type of objects and search criteria
- How should the output be presented?
  - Well-formatted search results allow users to find desired items faster.
  - Properties displayed might provide enough information so that the document does not need to be opened.

An overview of searches

- Stored search
  - Acts like a folder
    - The user clicks the icon and a list of contents is displayed.
  - Create a stored search when you have permanent criteria.
- A search template
  - Is a special kind of stored search with some property values left blank or editable
  - Can allow users to edit the criteria.
    - Property values only
- Simple searches
  - User-defined through Workplace XT > Preferences > Select Object Page > General
- All types can contain content-based search parameters.
  - Content-based retrieval is covered in another lesson.

Types of searches

Search	Reusable	Customizable at run time	Multiple Object Stores	Created with…
Stored Search	✓		✓	**Search Designer**
Search Template	✓	✓	✓	**Search Designer**
Simple Search	✓	✓		**User Preferences > Search**

– Content Engine Query Builder offers more options for administrators.
  - Accessed from Enterprise Manager.
  - Covered in another lesson.

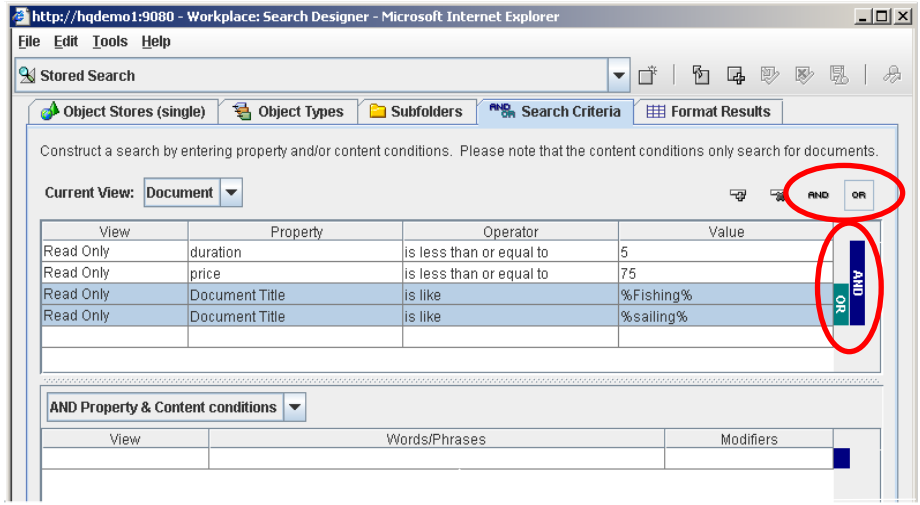
Search Designer

- Java applet for creating stored searches and search templates
  - You can test the templates immediately after creating and saving them.
  - Access from Workplace XT:
    - Tools > Advanced Tools > Search Designer

- To create a stored search, specify the following:
  - Which object stores (and optionally subfolders) you want to search
  - Whether to find documents or folders
  - The search criteria
  - How you want to display the results

Search criteria options

- Use property-operator-value format.
- Group individual criteria using AND and OR for more complex searches.



Demonstrations

- Create a stored search

What is the Query Builder?

- A search tool for Content Engine administrators
  - Accessed through Enterprise Manager
  - Use it when you need to apply actions to multiple objects that are returned by a search
  - You can perform bulk operations on the results of a search.
- Query Builder can search only one object store at a time.
  - These searches are saved as files in the file system, **not** as documents in an object store.
  - You can save them as stored searches or search templates.