

FAST Search for SharePoint

> > *SharePoint search on steroids*

BJØRN OLAV KÅSIN
MICROSOFT ENTERPRISE SEARCH
GROUP
JAN 15TH 2010

Microsoft Enterprise Search- Productivity

PRODUCTS FOR EVERY CUSTOMER NEED



Microsoft®
SharePoint® Server 2010

Complete intranet search



Microsoft®
FAST™ Search Server 2010
for SharePoint®

*High end search delivered
through SharePoint*

Common across the product line

- Common UI Framework
- Social search features and integration
- SharePoint platform integration
- End user and site administrator enablement
- Common Connector Framework (BDC)
- APIs and developer Experience
- Admin & deployment capabilities
- Operations advantages (SCOM, scripting)

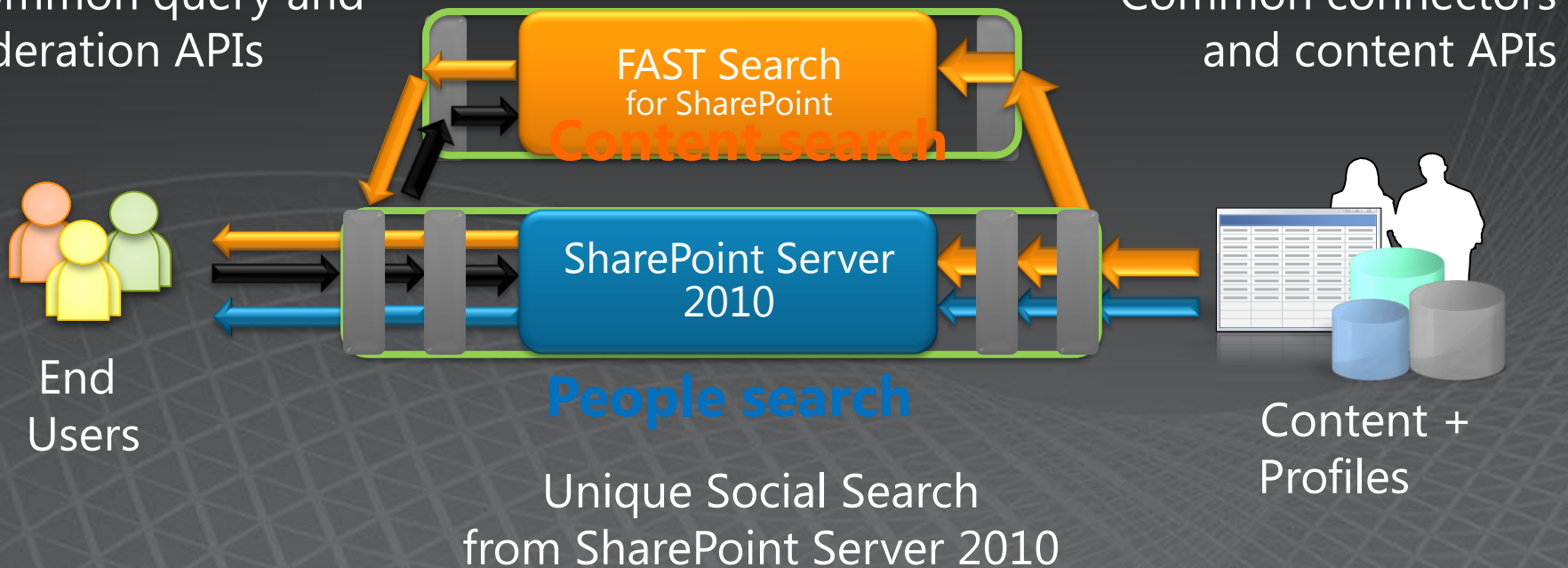
FAST Search Extends SharePoint Server

Superset of capabilities; common base functionality and platform

Best High-End Content Search
based on FAST ESP

Common query and
federation APIs

Common connectors
and content APIs



Common IT Pro and developer experience

End user

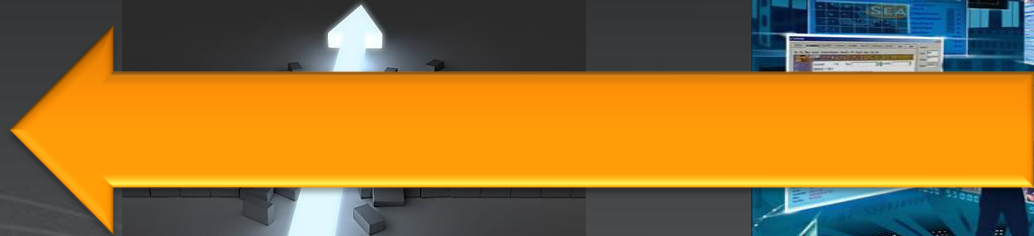
IT Professional

Developer

“Go beyond the search box”

“Eliminate Compromise”

“Do more with search”



End user

IT Professional

Developer

“Go beyond the search box”

“Eliminate Compromise”

“Do more with search”



- Find, explore, and connect (vs. “10 blue links”)
- Social search to connect people and expertise
- **Visual, conversational search**
- **Use Context to support the needs of diverse groups**
- **Precise Control of the user experience**

Go Beyond the Search Box

VISUAL, CONVERSATIONAL SEARCH

Sorting on any property

Visual Best Bets

Emergency department dashboard
medical professionals decision-making
more efficiently
> Learn more

Relevance
Relevance
freshness (Newest on top)
freshness (Oldest on top)
Date (Newest on top)
Date (Oldest on top)
Document Size (Ascending)
Document Size (Descending)
goodness (Ascending)
goodness (Descending)
Site Quality (Ascending)
Site Quality (Descending)

Refinement with
counts
on any property

Scrolling PowerPoint
Previews

Thumbnails

Similarity
Search

Result Type

Any Result Type

Microsoft PowerPoint (11)

Company

Any Company

SAP (2)

Microsoft (2)

Schneider Electric (1)

Author

more ▾

Site

Any Site

<http://www.microsoft.com>
(2)

<http://www.microsoft.com>
(2)

<http://www.microsoft.com>
(1)

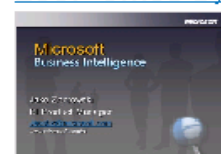
<http://www.microsoft.com>
(1)

more ▾

1-10 of 12 results



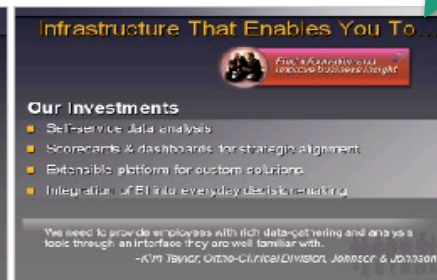
Amplify The Impact Of Your People Business Productivity Infrastructure Built On 2007 Office System



We need to provide employees with rich data-gathering interface they are well familiar with. Kim Taylor, Ortho-
Johnson Microsoft Business Intelligence An end-to-end
Scorecard Manager 2005 Performance

Authors: System Account Date: 2009-07-31T01:09:45

http://fastdemo76/Shared%20Documents/Demo%20documents/ITPro_BI_Deep_Dive_V2.pptx
Close Preview | View In Browser | Similar Results



Everyday Work with Windows Vista Best Practices Everyday Work with Windows Vista Best Practices



Understand the enhanced search capabilities, file organization improvements in Windows Vista to help you work smarter.

Authors: System Account Date: 2009-07-31T01:09:45

http://fastdemo76/Shared%20Documents/Demo%20documents/EverydayWorkWithWindowsVista_BP_E.ppt
Preview | View In Browser | Similar Results

User Context Matters

28:25-31:40



Renee Lo, *Engineer*

What should I know about implementing ERP?



Alan Brewer, *Sales*

What should I know about selling ERP consulting?

FAST Search • Search Results

All Sites People

erp

Preferences Advanced

Result Type 1-10 of 86 results

Sort by: Relevance

Any Result Typ...

Microsoft Word (79)

Microsoft Powe... (2)

Microsoft Exce... (2)

Site

Any Site

fastdemo17.fas... (84)

Author

Any Author

System Account (84)

Mike Tiano (8)

Rob Barker (5)

Don Funk (5)

show more v

Modified Date

Any Modified D...

Past 24 Hours (0)

Past Week (84)

Past Month (84)

Past Six Month... (84)

Past Year (84)

Earlier (0)

People Matches

Kari Furse
Salesperson
Sales Engagement
Management

Justin Thorp
Manager
Operations

Ari Suominen
Engineer
Engineering
Operations

View more people v

ERP Implementation Toolkit v5 is here!

ERP Web Design Toolkit now available

Download the latest ERP web design toolkit. Learn how the Long Distance Learning Engine will improve your productivity and make you a more valuable member of the team!

<http://demo/Customers/Graphic%20Design%20Institute%20-%20Distance%20Learning%20Engine%20Development%20-%20Sales%20Pitch.pptx>

Fourth Coffee - ERP Implementation - Requirements

... Considerations 6 Overview The ERP Implementation for Fourth ... Business Objectives In the ERP Implementation, Fourth Coffee ... reporting Constraints The ERP system must be fully ...

Authors: System Account Date: 10/12/2009 Size: 55KB

<http://fastdemo17.fastsearchdemos.net/Custo...fee-ERPImplementation-Requirements.docx>

View In Browser | Similar Results

Fourth Coffee - ERP Implementation - Statement of Work

FAST Search • Search Results

All Sites People

erp

Preferences Advanced

Result Type 1-10 of 86 results

Sort by: Relevance

Any Result Typ...

Microsoft Word (79)

Microsoft Powe... (2)

Microsoft Exce... (2)

Site

Any Site

fastdemo17.fas... (84)

Author

Any Author

System Account (84)

Mike Tiano (8)

Rob Barker (5)

Don Funk (5)

show more v

Modified Date

Any Modified D...

Past 24 Hours (0)

Past Week (84)

Past Month (84)

Past Six Month... (84)

Past Year (84)

Earlier (0)

People Matches

Shmuel Yair
Salesperson
Sales Engagement
Management

Kari Furse
Salesperson
Sales Engagement
Management

Justin Thorp
Manager
Operations

View more people v

NEW SALES INCENTIVE ON STRATEGY WORKSHOPS

Strategy Workshops for ERP Project Sales

Come to the Strategy Workshops for ERP Project Sales and learn to deliver compelling Sales Pitches. This Cold Call Sales Guide is just an example of what you will learn.

<http://demo/Customers/Proseware,%20Inc.%20-%20Cold%20Call%20Sales%20Guide%20-%20Sales%20Pitch.pptx>

Fourth Coffee - ERP Implementation - Sales Pitch

... Consulting Fourth Coffee ERP Implementation IT Implementation ... team will implement SAP as an ERP system. The Contoso team ... Proposed deliverables: ERP Implementation documentation ...

Authors: System Account Date: 10/12/2009 Size: 500KB

<http://fastdemo17.fastsearchdemos.net/Custo...fee-ERPImplementation-SalesPitch.pptx>

Preview | View In Browser | Similar Results

Fourth Coffee - ERP Implementation - RACI

Contoso Consulting - RACI Matrix: Fourth Coffee, ERP Implementation Responsible Accountable Consulted Informed Discovery Brian Groh Mike Tiano ...

Authors: System Account Date: 10/12/2009 Size: 422KB

<http://fastdemo17.fastsearchdemos.net/Custo...urthCoffee-ERPImplementation-RACI.pptx>

Preview | View In Browser | Similar Results

Fourth Coffee - ERP Implementation - Request for Proposal

... prospective Vendors for the ERP Implementation. Contoso team will implement SAP as an ERP system... The scope of ... present. DRAFT DRAFT ERP Implementation Page 1 of 3 ...

Authors: System Account Date: 10/12/2009 Size: 107KB

Go Beyond the Search Box

SHAPING THE USER EXPERIENCE

- Site admin/Search admin control
 - Visual Best Bets
 - Promote/Demote documents and sites
 - UI extensibility (web parts, ..)
 - Relevancy profiles and parameters
 - User Context parameter & admin
- End User Control
 - Sorting, Ranking, and Navigation
 - Admin-enabled controls
- Linguistics and term control
 - Keywords, phrases, synonyms, spellcheck
 - Multilingual searching control
 - Lists for metadata extraction



Go Beyond the Search Box

DEEP REFINERS WITH EXACT COUNTS

Refine Results:

Any Date

Between 2009-05-14T... (33)

Between 2009-05-14T... (30)

Before 2009-05-14T1... (30)

2009-05-14T12:55:15... (29)

Any Result Type

Adobe PDF (115)

Microsoft PowerPoin... (7)

Any Author

System Account (45)

Cappgemini - 11 rue ... (16)

10-K Wizard Technol... (8)

Andrew Frank (4)

more v

Any Site

www.accenture.com (26)

www.cappgemini.com (13)

www.10kwizard.com (8)

www.bearingpoint.co... (7)

more v

Any Company

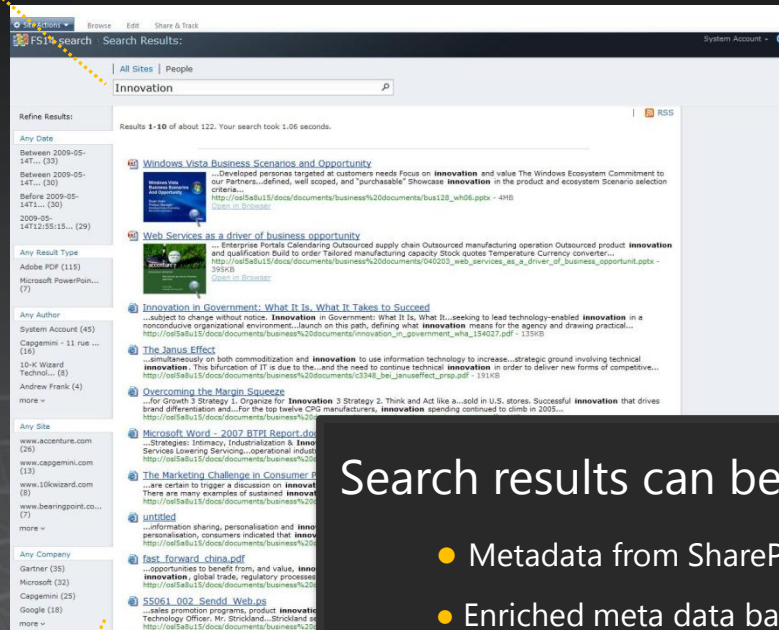
Gartner (35)

Microsoft (32)

Cappgemini (25)

Google (18)

more v



Search results can be refined based on:

- Metadata from SharePoint (such as Author)
- Enriched meta data based on entity extraction (such as Company)

Custom refiners can be made by:

- Enabling new Managed Properties to be used as refiners
- Configuring the Refinement Web Part to display the new refiner

Go Beyond the Search Box

BROADER, BETTER LANGUAGE COVERAGE

- Can search in any language
(including klingon!);
- 84 languages detected to allow
language-specific handling
- Lemmatization improves recall
(*'better' includes 'good'*)
- Phrase search includes stopwords
(*"a room with a view"*)
- Only nouns and adjectives are
expanded (higher precision)
(*'book' -> 'books', not 'booked'*)

Afrikaans	Hausa	Pashto, Pushto
Albanian	Hebrew	Persian
Arabic	Hindi	Polish
Armenian	Hungarian	Portuguese
Azerbaijani	Icelandic	Punjabi
Basque	Indonesian	Rhaeto-Romance
Bengali, Bangla	Irish	Romanian
Bosnian	Italian	Russian
Breton	Japanese	Sami (Northern)
Bulgarian	Kannada	Serbian
Catalan	Kazakh	Slovak
Chinese-S	Kirghiz	Slovenian
Chinese-T	Korean	Sorbian
Croatian,	Kurdish	Spanish
Czech	Latin	Swahili
Danish	Latvian, Lettish	Swedish
Dutch	Letzeburgesch	Tagalog
English	Lithuanian	Tamil
Estonian	Macedonian	Telugu
Faroese	Malay	Thai
Finnish	Malayalam	Turkish
French	Maltese	Ukrainian
Galician	Maori	Urdu
Georgian	Marathi	Uzbek
German	Mongolian	Vietnamese
Greek	Norwegian	Welsh
Greenlandic	Norwegian-B	Yiddish
Gujarati	Norwegian-N	Zulu

Go Beyond the Search Box

END USER PERSPECTIVE



Microsoft®
SharePoint Server 2010

Great OOB search experience

- Excellent OOB relevance
- Federated results
- PC, phone, or browser access
- Metadata based refiners
- Related searches
- View in Browser
- Phonetic & nickname matching
- Relevance and refiners
- Integrated presence
- Click-through drives relevance
- Colleague & Expertise suggestions
- Query suggestions



Microsoft®
FAST™ Search Server 2010
for SharePoint

Most visual and dynamic
user experience

- Same capabilities, *plus*:
- Broader, better language support
 - Deeper query syntax

- Same capabilities, *plus*:
- Deep refiners with counts
 - Similarity search
 - Sorting on any property

Same capabilities

Same capabilities

- Thumbnails & Scrolling Previews
- Visual Best Bets

- User context from user profile
- Multiple relevance profiles

One-stop search center
to *find* answers fast

Refinement to *explore*
information quickly

Social search to *connect*
with people and expertise

Search gets better with use

Visual cues for rapid
recognition of information

Contextual to meet the
needs of diverse groups

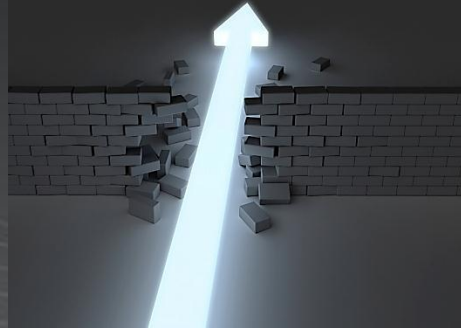
End user

“Go beyond the search box”



IT Professional

“Eliminate Compromise”



Developer

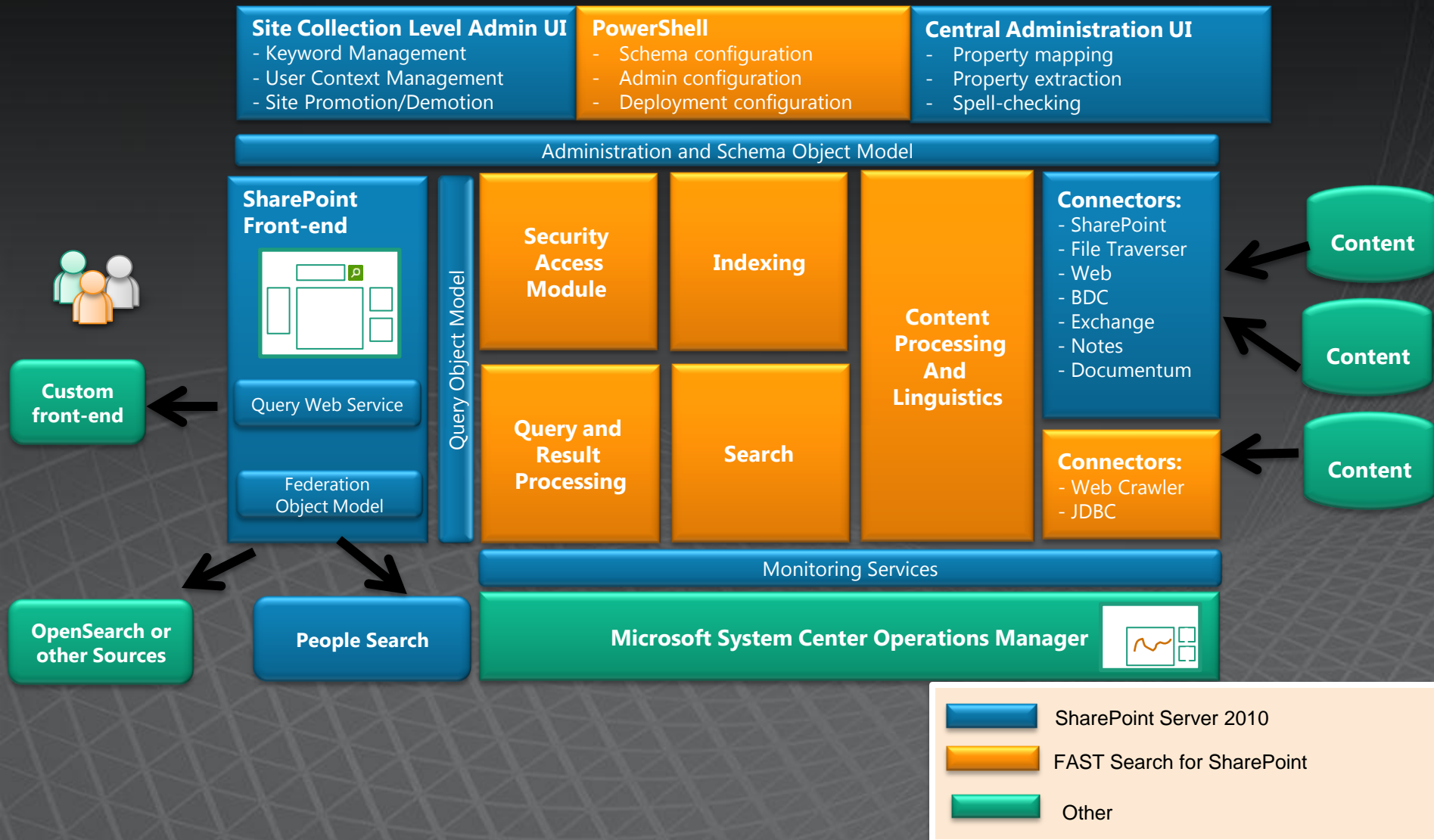
“Do more with search”



- **Scale-out and Performance to handle any business need**
- Easy, enterprise-class deployment and manageability
- Secure, broad connectivity
- **Advanced Content Processing out-of-the-box**
- **Easy to configure powerful user experiences**

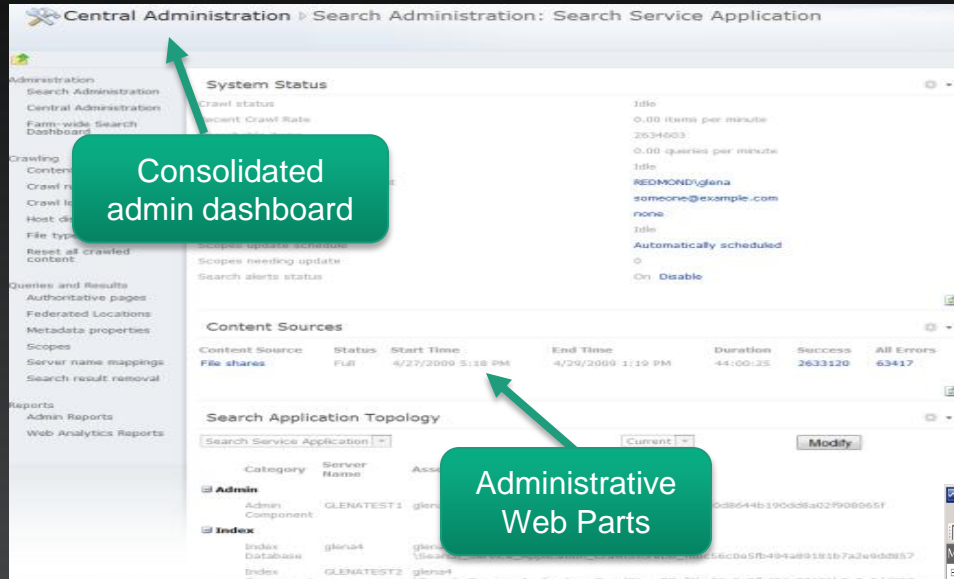
FAST Search for SharePoint

Summary of architectural elements



Eliminate Compromise

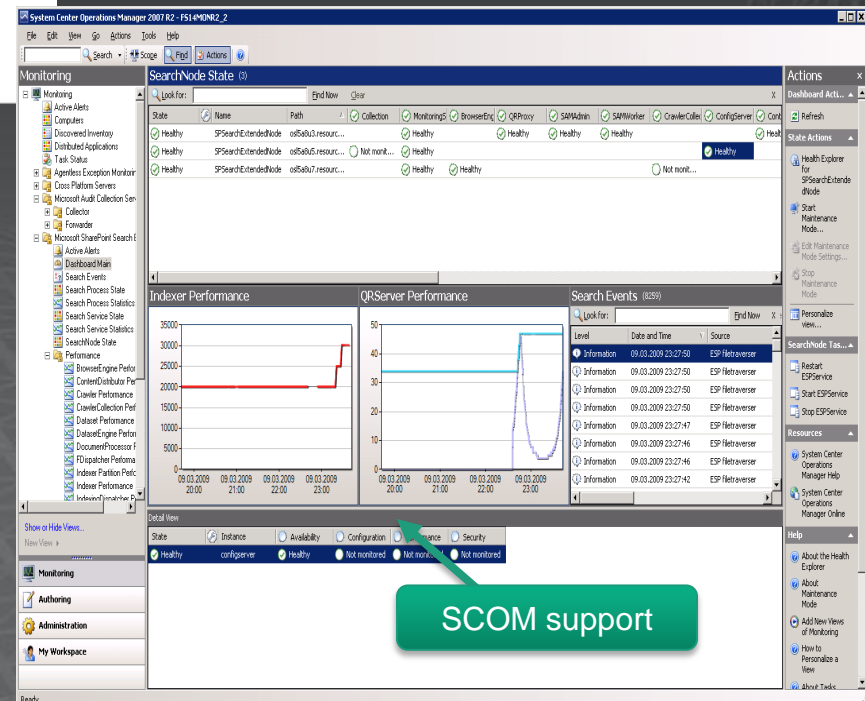
ENTERPRISE-CLASS MANAGEABILITY



PowerShell and
Virtualization support

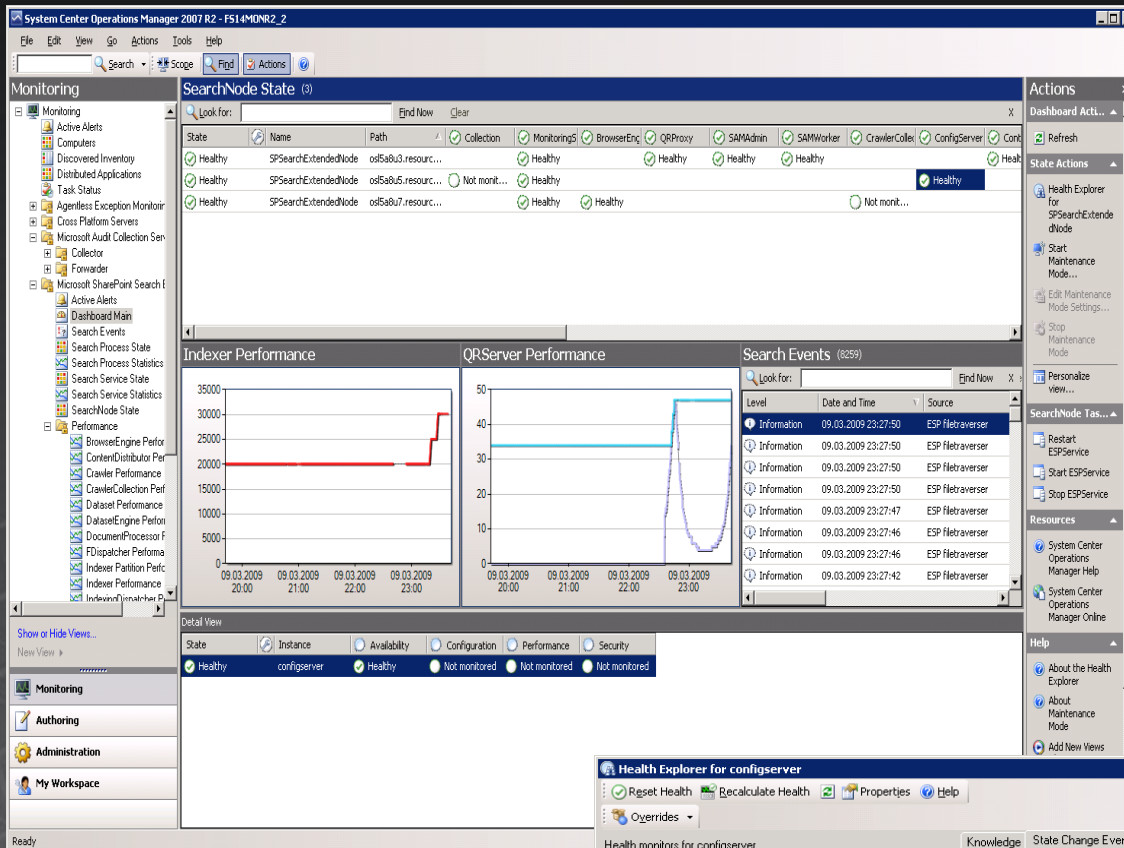
Add-FASTSearchResource -Path
dictionaries\wholewords1.xml

Manage Reliably
*WITH A CONSOLIDATED ADMIN
DASHBOARD, ADMINISTRATIVE
WEB PARTS, BUILT-IN USER AND
SYSTEM MONITORING, AND SCOM
SUPPORT*

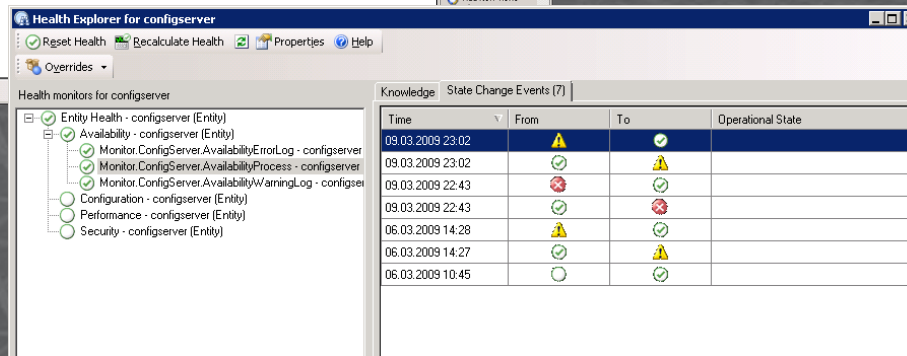


Monitoring and Alerting

GRANULAR STATE, PERFORMANCE AND ALERTING OVERVIEW



- Dedicated Management Pack for System Center Operations Manager
- Support of standard Windows monitoring services
- Detailed monitoring of state and performance at node level
- Extensive alerting options



Management

54:40-56:00

- Central Admin UI for core configurations
 - Managed Properties
 - Entity extraction
 - Spell check
- Powershell support for scripted administration
 - Schema
 - Rank
 - Crawling
- Site Admin UI for search management
 - Keyword management
- Extensible search analytics reporting

Eliminate Compromise

EASY TO CONFIGURE POWERFUL USER EXPERIENCES

Empower end users to personalize result sets with sorting

Visually highlight content that is a 'best bet' for a given query

Create more intuitive navigation by converting keywords into refiners

Search Action Links
Search Results Action Links

Add Managed Property | Crawled Property Categories | Refresh

Property Name ↑	Type	Sort Property	Query Property	Refiner Property	Mappings to Crawled Properties
DocKeywords	Text	No	Yes	No	keywords(Text), Office:5(Text), 5(Text)
docrank	Integer	Yes	Yes	No	
DocSubject	Text	No	Yes	No	Office:3(Text), ows_EMail(Text), ows_EM
	Integer	Y	Yes	No	
	Text	No	No	No	
	Text	No	Yes	No	
	Datetime	No	No	No	ows_End_x0020_Date(Datetime)
FileExtension	Text	No	Yes	No	FileExtension(Text), ows_File_x0020_Typ ext(Text)

Enable end-user control of Refinement, sorting, languages

Control property extraction and use in refinement, sorting, and ranking

Easy admin of user context, best bets, visual best bets, & promotion/demotion

Results Per Page: 10

Limit Characters In Summary: 185

Limit Characters In URL: 90

Refinement controls:

- Date: Newest
- Companies: Ascending
- documentsignature: Ascending

Search results:

- product management
- productivity
- user experience
- UT Design
- write
- companies
- documentsignature

Simplified search administration

Tune relevance in FS2010

● **Tuning of keyword driven rank**

High level and easiest way of tuning relevance is with usage of the SharePoint administration UI

- Promote/demote sites & documents
- Use search keywords
 - Best Bets or Visual Best Bets
 - Used with synonyms (hits in synonyms do not give any rank)

Tune relevance in FS2010, cont...

● **Tuning of static rank**

- Added at index time into different placeholders
- Not possible to tune them at query time
- Very effective at query time since there are no rules or calculation, only retrieval and summary of the values
- Default the FS2010 containers are:

urldepthrank	Used to boost shorter URLs
docrank	Rank boost based on the number of and relative importance of links pointing to this item
siterank	Rank -number of and relative importance of links pointing to the items on this site
hwboost	FS14 placeholder for generic usage of static rank points, default value = 0

Tune relevance in FS2010, cont...

- **Tuning of static rank, example:**

Modify the impact of the static rank managed properties:

```
PS C:\> $rp = Get-FASTSearchMetadataRankProfile -Name default
PS C:\> $np = New-FASTSearchMetadataRankProfile -Name URLboost -Template $rp
PS C:\> $qc = $np.GetQualityComponents()
PS C:\> $qc | where-object -filterscript { if ( $_.ManagedProperty
Reference.Name -eq "urldepthrank" ) { $_.Weight=200; $_.Update() } }
```

- **other, examples:**

- **Modify which properties to use**
- **Use custom managed properties for static rank boost**
- **.....**

Eliminate Compromise

IT PRO PERSPECTIVE



Microsoft®
SharePoint Server 2010

Enterprise class search system
integrated with SharePoint

- Enterprise Scale-out (to 100M docs)
- Full Fault Tolerance
- Native 64 bit; Hyper-V support

- Wizard - driven installation
- Consolidated search dashboard

- PowerShell support
- SCOM support
- Full search reporting

- Full set of connectors OOB
- Easy to add new sources via BDC
- Strong security



Microsoft®
FAST™ Search Server 2010
for SharePoint

Highly capable **and**
easy to manage

Same capabilities, *plus*:

- Extreme scale-out

Same capabilities

Same capabilities

Same capabilities, *plus*:

- Enterprise Web Crawler

- Content Processing pipeline
- Metadata Extraction

- Easy setup of User Context, Visual Best Bets, Promotion
- Give end users control over sorting, ranking, navigation

Industrial strength
scale-out and performance

Easy deployment

Enterprise-class
manageability

Secure,
broad connectivity

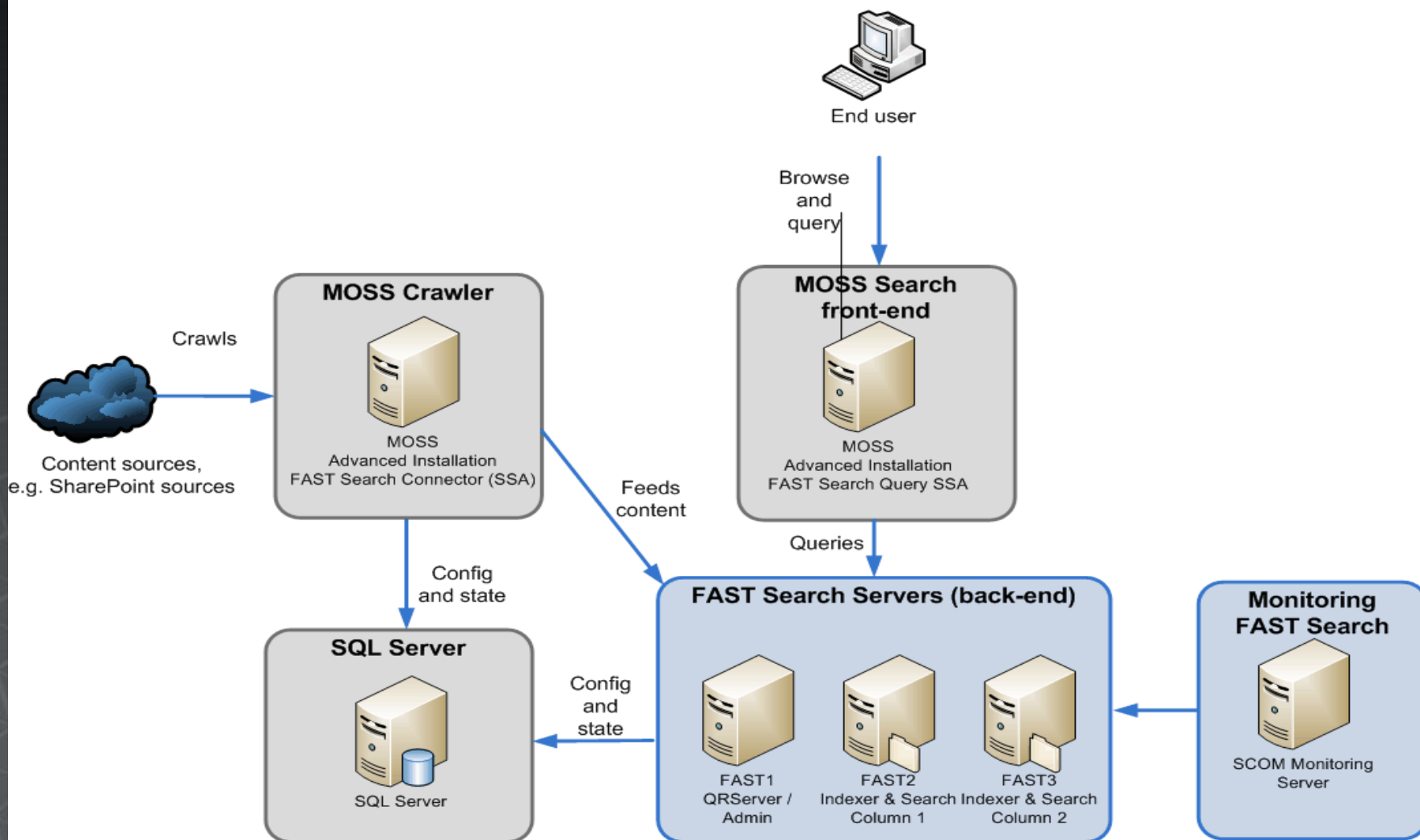
Advanced content
processing OOB

Easy to configure high-end
user experiences

Overview – Deployment

OVERALL TOPOLOGY

FAST Search multiple server deployment

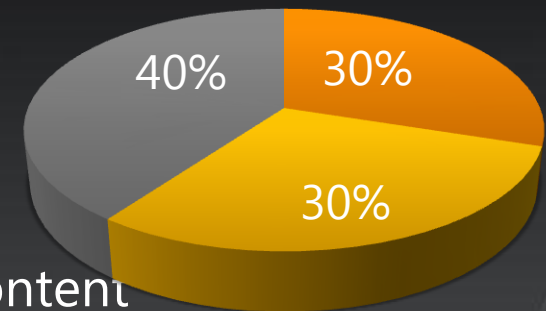


FS2010 backend

SIZING GUIDELINES

Sizing scenario

- SharePoint list items
- SharePoint attachments
- File shares



- “Fully” 64-bit
 - Maximum 2^{31} unique terms
 - In practice no process memory limits
- Limitations for realistic usage
 - Disclaimer: All numbers dependent on content
 - Assuming single node 8 core server (or better)
 - Recovery: Resetindex processing ~1M/hour
 - Feed rates: ~25 dps (or 2M/day)
 - Practical sizing: 15M – 40M documents per node
- Multi-node configurations by creating a deployment.xml
 - Provided to admin node at install time
 - Start admin node before any other nodes
 - All other nodes will get their configuration from the admin node

FS2010 backend

SIZING GUIDELINES - CONTINUED

- Search cluster
 - Find suitable number of columns
 - Optional: Add a dedicated search row
 - Consistent search latency
 - Redundancy
 - Optional: Add a backup indexer row
 - Redundancy
- "Admin" server
 - configserver, samadmin, spelltuner, sprel etc
 - Dedicated webanalyzer node(s) for larger installations
- Document processing
 - Sprinkle document processors across other nodes (will run as low priority)
 - contentdistributor and indexingdispatcher will normally not need dedicated nodes
 - Use multiple contentdistributors for redundancy
 - Add contentdistributors to keep below ~1:30 ratio with document processors

```
<host name="osl2u521.europe.corp.microsoft.com">
  <admin />
  <indexing-dispatcher />
  <content-distributor />
  <webanalyzer server="true" link-processing="true"/>
  <searchengine row="0" column="0" />
  <document-processor processes="14" />
</host>

<host name="osl2u522.europe.corp.microsoft.com">
  <query />
  <searchengine row="0" column="1" />
  <document-processor processes="12" />
</host>

<host name="osl2u523.europe.corp.microsoft.com">
  <query />
  <searchengine row="0" column="2" />
  <document-processor processes="12" />
</host>

<host name="osl2u524.europe.corp.microsoft.com">
  <content-distributor />
  <indexing-dispatcher />
  <searchengine row="0" column="3" />
  <document-processor processes="16" />
</host>
```


FS2010 backend

SIZING GUIDELINES - CONTINUED

- Powerful servers allow for denser configuration
 - More CPU cores
 - More memory
 - Better disks
- Index and search are the heaviest disk users
- Document processing consumes most of the CPU capacity
- “Typical” usage under load for most important components

Component	CPU	Memory	Disk	Notes
Indexer	4-8 cores	4-8 GB	+++++++++	
Search	2-4 cores	4-8 GB	+++++	
Document processor	1 core	1 GB	0	Default 4x DP
Webanalyzer	1 core ++	2 GB ++	++	Configurable in deployment.xml
Spelltuner	1 core	2 GB	+	Nightly runs

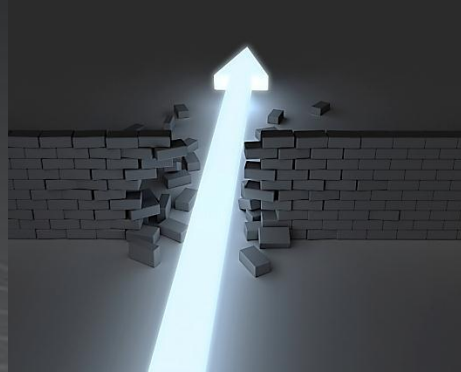
End user

“Go beyond the search box”



IT Professional

“Eliminate Compromise”



Developer

“Do more with search”

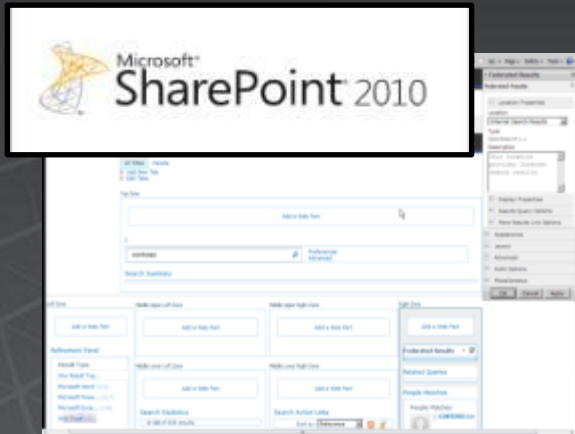


- Work across a range of configuration and development
- Familiar, integrated development tools and framework
- **Richer capabilities enable search-driven applications**

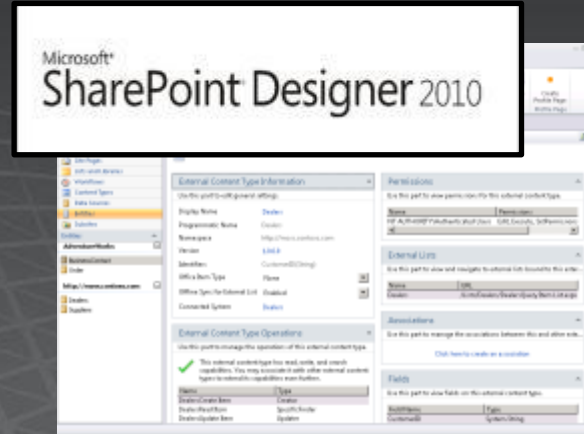
Do More With Search

ACROSS A SPECTRUM OF CUSTOMIZATION AND DEVELOPMENT

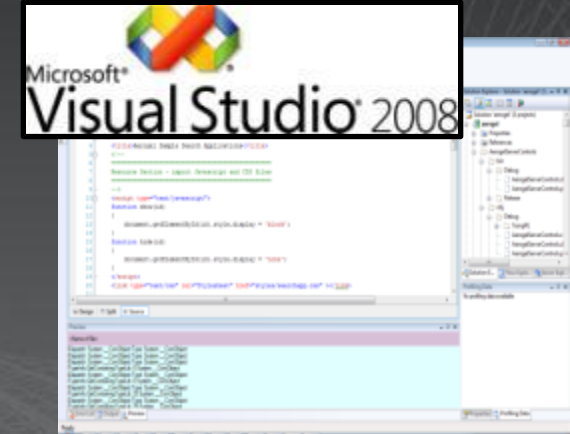
Configure



Extend



Create



Example Applications



Configure

Intranet Search
People Search
Site Search

Extend

Research Portal
Case Management
Voice of the Customer
Parts Search
.....

Create

IP Portfolio mgmt
BI/Search
Collaboration Kiosk
Compliance
Intel/Surveillance
Drug Discovery
Authoring co-pilot
....

Do More with Search



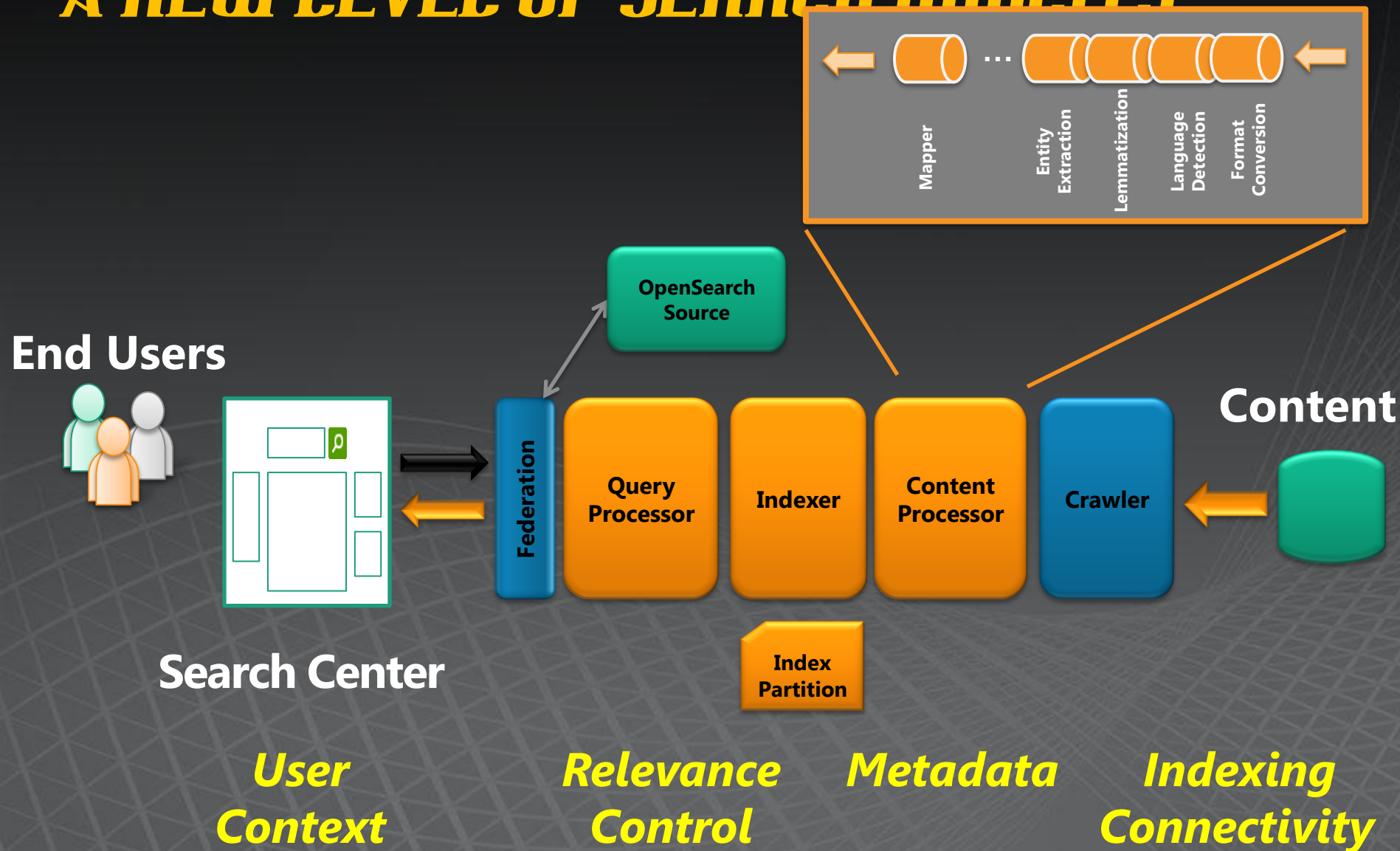
Common Framework, APIs, Development Tools

Richer Capabilities
enable search-driven
applications

Advanced Content Processing

A NEW LEVEL OF SEARCH QUALITY

Content
Processing
Pipeline



Advanced Content Processing



Improving Insight and Decision
Making Using M **PRODUCT (Custom)**
Intelligence and **SQL Server 2008**

Rafal Lukawiecki
Strategic Consultant, Project Botticelli Ltd
rafal@projectbotticelli.com

CONCEPT (Custom)

Business Intelligence
BI - Improving Business Insight

"A broad category of applications and technologies for gathering, storing, analyzing, sharing and providing access to data to help enterprise users make better

COMPANY (OOTB)

— Gartner



Site Actions ▾ Browse Page Publish

Search C ▸ Search Results

All Sites People

windows server

Preferences Advanced

10 of 99 results Sort by: Relevance

Managing Windows Server 2008 R2
... to be familiar with **Windows Server** 2008 R2 management ... entitled Managing **Windows Server** 2008 R2 white paper ... Streamlined management in **Windows Server** 2008 R2 Overview ...
Date: 8/13/2009 Size: 1MB Rank 1011837
http://janhsdemo2/Shared Documents/WS08R2_Management_Technical_Overview-FINAL.pptx
Preview | View In Browser | Similar Results

Windows Server 2008 R2: A Technical Overview
Windows Server 2008 R2: A Technical ... 3 year cadence of **Windows Server** releases Each major ... update (e.g., **Windows Server** 2003, followed by WS ...
Date: 5/9/2009 Size: 4MB Rank 14612
<http://janhsdemo2/Shared Documents/Windows Server 2008 R2 - A Technical Overview.pptx>
Preview | View In Browser | Similar Results

Improving Insight and Decision Making Using Microsoft Business Intelligence and SQL Server 2008
... Intelligence and **SQL Server** 2008 Rafal ... Microsoft, **Windows**, **Windows Vista** ... Platform Suite: **SQL Server**, Excel, Sharepoint ... Performance Point **Server** Automate KPI ... Microsoft, **Windows**, **Windows Vista** ...
Date: 4/23/2009 Size: 3MB Rank 12075
<http://janhsdemo2/Shared Documents/1 Improv ... siness Intelligence and SQL Server 2008.pptx>
Preview | View In Browser | Similar Results

Advanced Windows PowerShell Scripting.pptx
... partners Exchange **Server** • **Windows** Diagnostics • **Server** ... pick up your copy of **Windows Server** 2008 R2 RC from the ... Learn More about **Windows Server** 2008 R2: <http://janhsdemo2/Shared Documents/Advanced Windows PowerShell Scripting.pptx>
Date: 5/13/2009 Size: 1MB Rank 12074
Preview | View In Browser | Similar Results

Windows PowerShell: Tips from the Expert
... PowerShell and **Windows Server** 2008 R2 **Windows** ... pick up your copy of **Windows Server** 2008 R2 RC from the ... Learn More about **Windows Server** 2008 R2: <http://janhsdemo2/Shared Documents/Windows PowerShell Tips from the Expert.pptx>
Date: 5/13/2009 Size: 1MB Rank 12074
Preview | View In Browser | Similar Results

Result Type
Any Result Type
Microsoft PowerPoint (49)
Adobe PDF (29)
Microsoft Word (18)
Web Page (2)

Company
Any Company
Microsoft (97)
IBM (14)
SAP (12)
Gartner (12)
show more ▾

Product
Any Product
Microsoft Office (31)
SQL Server (16)
Microsoft Exchange (11)
FAST ESP (7)
show more ▾

Concept
Any Concept
Compliance (19)
Business Intelligence (17)
Business Process (10)

Extract properties to be used for refinement

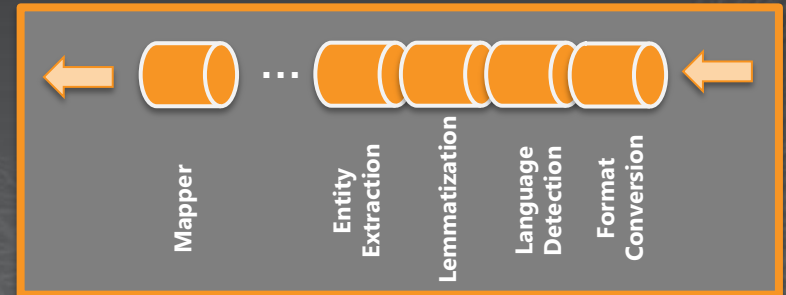
Content Pipeline Stages

Default

- **Format Conversion**
iFilters, OutSideIn
- **Language and encoding detection**
- **Lemmatizer**
Linguistics normalization
- **Tokenizer**
Word breaking
- **Entity Extraction**
Companies, locations
- **DateTimeNormalizer**
Date normalization
- **Vectorizer**
Create document vector for similarity searching
- **WebAnalyzer**
Anchor text and link cardinality analysis
- **PropertiesMapper**
Map to crawled properties
- **PropertiesReporter**
Report detected properties

Optional

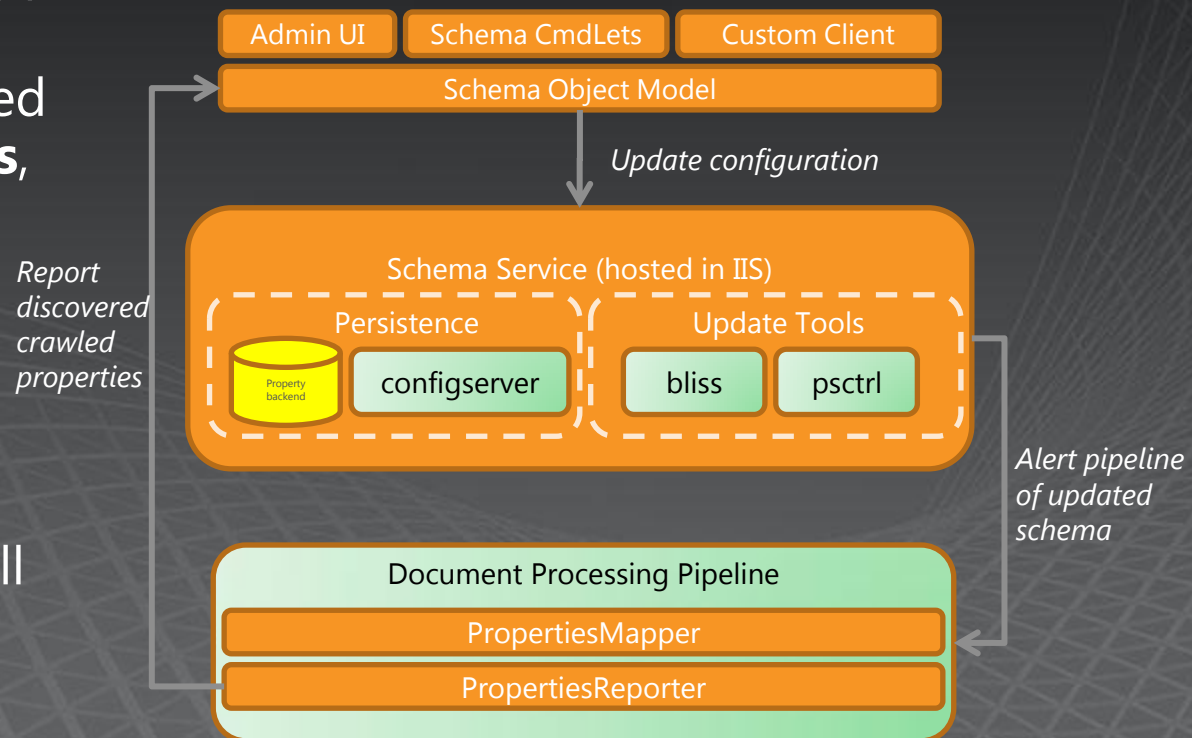
- XML Properties mapper
- Offensive Content Filter
- Verbatim (wholeword) extractor
 - Loads dictionary for custom extraction, e.g product names
- Field Collapsing
- Entity Extraction
 - Persons



The different plug-ins can be configured via UI or PowerShell

Content Processing and Schema

- Extracted document attributes reported as **Crawled Properties**
- **Crawled Properties** mapped to **Managed Properties**
- Characteristics are defined for **Managed Properties**, e.g.
 - Refiners
 - Sorting
 - Queryable
 - Type
- Definition and mapping done via UI or Powershell
 - Not via Central Admin



Content Pipeline Extensibility

- Approach

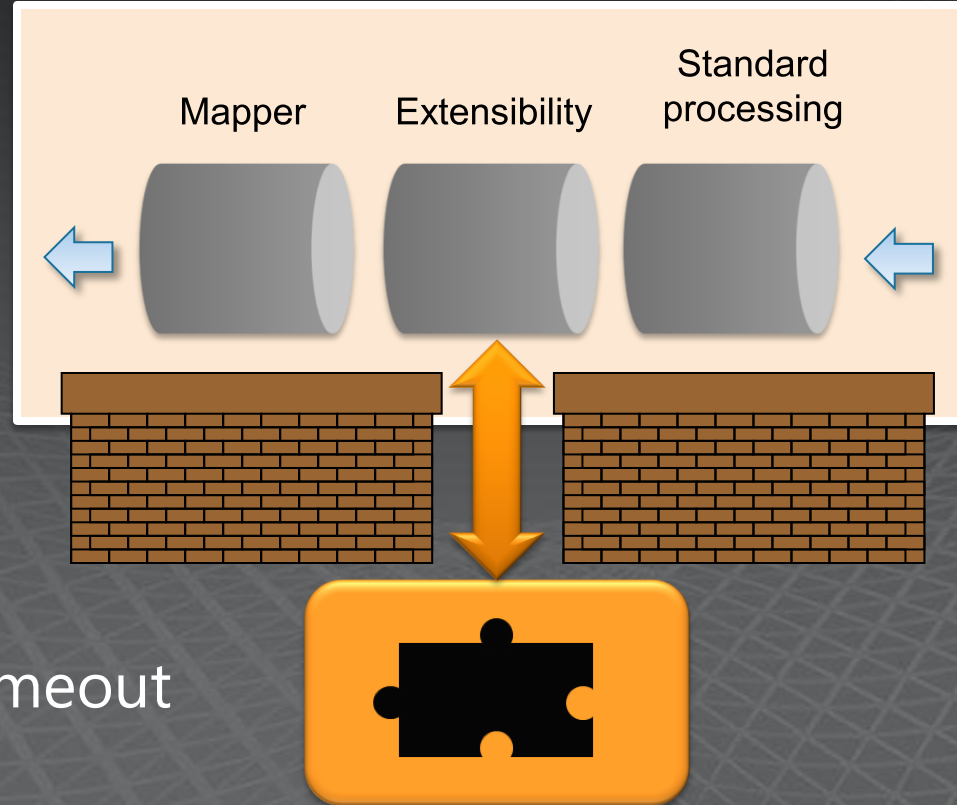
- Straightforward way to add text analysis functionality
- Flexible and supportable

- Example uses

- Sentiment analysis
- Translation
- Auto-Classification

- Mechanism

- Just before Mapper
- Runs in sandbox with timeout



Schema management – a la carte search menu

Site Actions

Give Feedback My SharePoint Benit Herstad

Central Administration Managed Properties

Crawled properties are automatically extracted from crawled content. Users can perform queries over managed properties. Use this page to create and modify managed properties and map crawled properties to managed properties. Changes to properties will take effect after documents are reprocessed or after the next full crawl.

Central Administration

Administration

- FAST Search Administration
- Central Administration

Search Service Application

Search for Managed Property Names:

Add Managed Property Crawled Properties Refresh

Property Name	Type	Sort Property	Query Property	Refiner Property	Mappings to Crawled Properties
Account	Text	No	Yes	No	SharePoint:ows_Name(Text)
anchortext	Text	No	Yes	No	
AssignedTo	Text	No	Yes	No	SharePoint:ows_AssignedTo(Text)SharePoint:ows_Assigned_x0020_To(Text)
assocqueries	Text	No	Yes	No	
Author	Text	No	Yes	Yes	Mail:6(Text)Office:#4(Text)Office:#4()OutsideIn:primary_author(Text)
body	Text	No	No	No	
charset	Text	No	Yes	No	
Companies	Text	No	Yes	Yes	SharePoint:ows_Company(Text)
companyteaser	Text	No	No	No	
concepts	Text	No	Yes	Yes	
Edit Managed Property	Text	No	Yes	No	Basic:#5(Text)SharePoint:ows_ContentType(Text)
Delete Managed Property	Text	No	Yes	No	
crawledpropertiescontent	Text	No	Yes	No	
crawledpropertynames	Text	No	Yes	Yes	
datetime	Datetime	Yes	Yes	No	
dated	Datetime	Yes	Yes	No	
datedBy	Text	No	Yes	No	
es	Text	No	Yes	Yes	
escription	Text	No	Yes	No	
ed	Text	No	Yes	No	
occlsystemid	Text	No	Yes	No	
DocComments	Text	No	Yes	No	
docdatetime	Datetime	Yes	Yes	Yes	
DocKeywords	Text	No	Yes	No	

How can you create a Managed Property?

- Map a Crawled Property to a Managed Property
- A Crawled Property is generated by default or custom content processing

Map Crawled Properties to Managed Properties. Define naming and characteristics (type, sort, query, refiner)

What can a Managed Property be used for?

- Sorting (single level, multi-level)
- Full text search (composite fields)
- Refinement
- Boost on Managed Property
- Used as part of rank profile (quality, field, managed property boost)

Richer extensibility in FAST Search for SharePoint

- Extensions to shared Web Parts
 - Core results Web Part; Search Action WebPart
- Configurable and extensible content processing
 - Custom property extractors; Pipeline configuration & extension
- Additional capabilities for custom rank models
 - Anchor Authority; Query Authority (click-through); Freshness; Quality (based on value on specific managed properties); Full-text index rank
- Extensions to Query Object Model
 - Refinement; Sorting; Find Similar; User Context; Filter (FQL, restriction on query, more "advanced scope"); Visual Best Bets; Zero result handling (resubmit flags); Spellchecking (did you mean, automatically re-write query)
- FQL (FAST Query Language)
 - Extensive relevance control per query (Xrank); Multilevel sorting and formula sorting; Near () and symbolic searching

Do More with Search

Extend OOB Web Parts – or Create New Ones

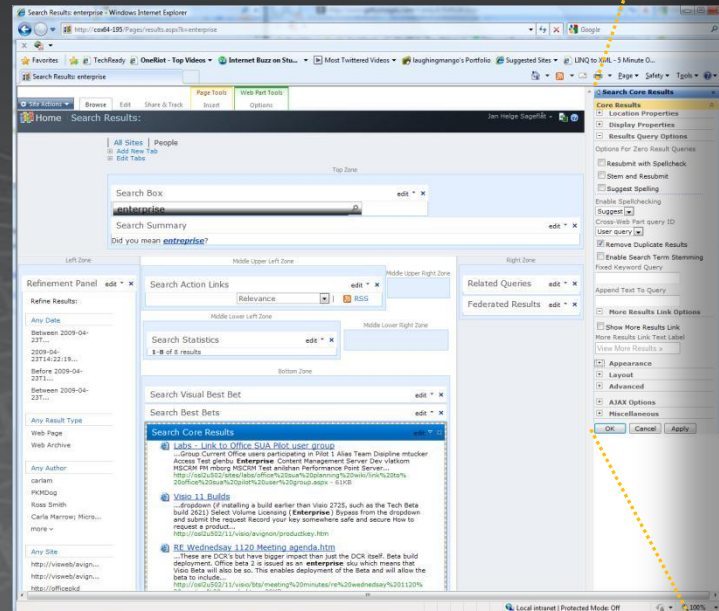
1.03:00-1.06:50

Keywords

- microsoft corporation
- view of microsoft
- microsoft office
- microsoft products
- informational purposes
- service levels
- strategic asset
- united states
- windows vista
- information technology

Companies

Gartner, SAP, IBM, Oracle, Microsoft



Search Core Results

Core Results

- ☒ Location Properties
- ☒ Display Properties
- ☐ Results Query Options

Options For Zero Result Queries

- ☐ Resubmit with Spellcheck
- ☐ Stem and Resubmit
- ☐ Suggest Spelling

Enable Spellchecking

Suggest

Cross-Web Part query ID

User query

- ☒ Remove Duplicate Results
- ☐ Enable Search Term Stemming

Fixed Keyword Query

Append Text To Query

- ☐ More Results Link Options

- ☐ Show More Results Link

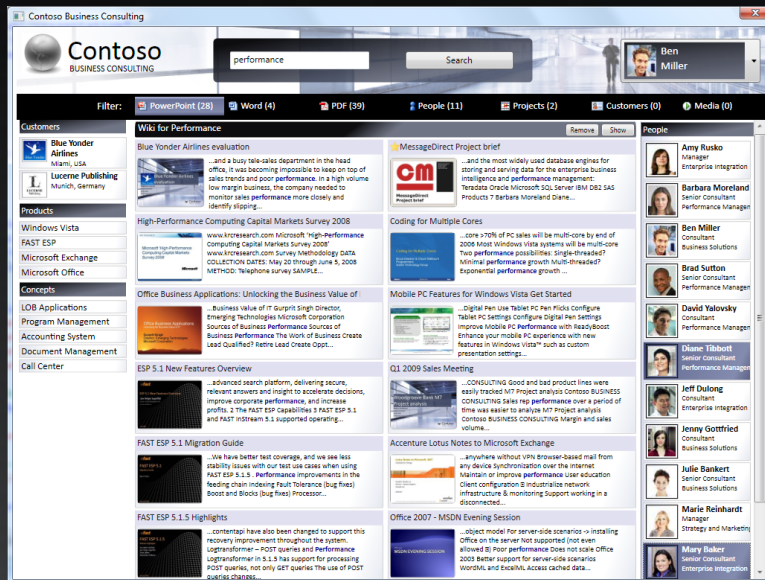
More Results Link Text Label

Appearance

- ☒ Layout
- ☒ Advanced
- ☒ AJAX Options
- ☒ Miscellaneous

Query Language Expressiveness

Soft boost using the Xrank operator



Boost all documents that contain the name



Diane
Tibbott

Boost all documents that contain the name



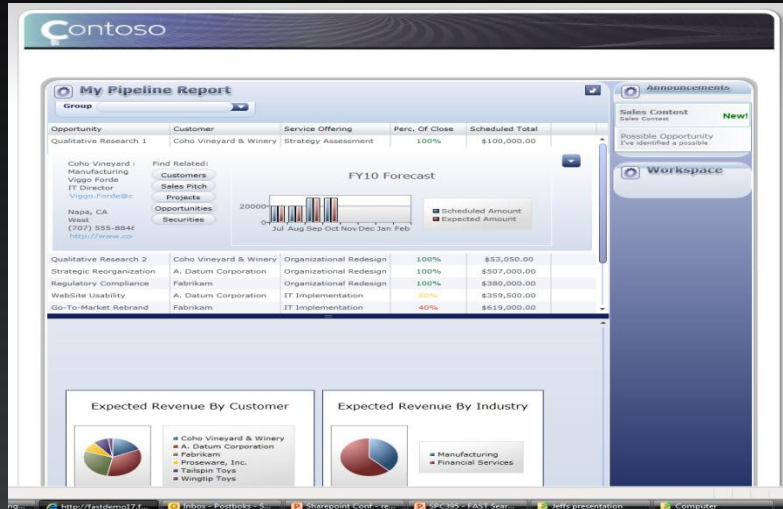
Mary
Baker

Or, in the FAST Query Language:

Single (documents containing either Mary, Diane or both gets a boost of 5000):
Xrank(string("performance"), or(person:string("diane tibbott"), person:string("mary baker")), boost=5000))

Accumulative (documents with both Mary and Diane gets a boost of 10000):
Xrank(Xrank(string("performance"), person:string("diane tibbott"), boost=5000), person:string("mary baker"), boost=5000)

XRANK - Rank power in query language



Query used behind the scenes to find related sales pitches:

```
xrank(xrank(and(and(\"salespitch\",industry:\"{industry}\"),filter(\"fileextension\":pptx)),string(\"{opportunity_name}\", mode=\"simpleany\"), boost=50000),string(\"{service_offering}\", mode=\"simpleany\"), boost=90000)
```

- Find sales pitches from same industry and only show the pptx files
- Boost pitches which includes same opportunity name
- Boost (more) the pitches covering the same service offering

Location Awareness

Geo-search using sortformula()

Content

Processing

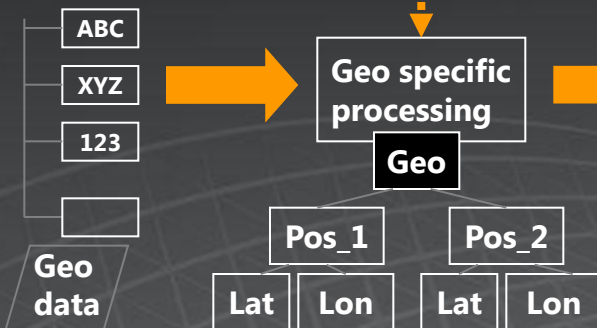
Search/Index

Query

Result

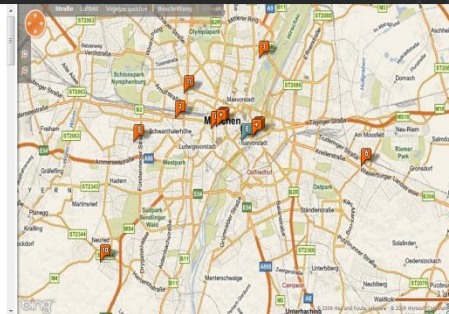
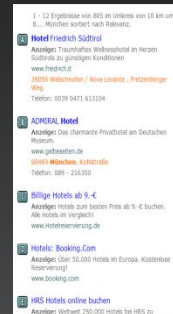


External
data



Index

Sort results by
distance from
end-user
location
(50,100)



Formula sorting allows
complete control of distance
calculations (standard or
'warped' geometries)

$$\text{Distance(from 50,100)} = \sqrt{(50 - \text{latitude})^2 + (100 - \text{longitude})^2}$$

query=hotel&sortby=+[formula:sqrt(pow(50-latitude,2)+pow(100-longitude,2))]

Do More with Search

DEVELOPER PERSPECTIVE



Microsoft
SharePoint Server 2010

Open and flexible
search platform

- Public Web Parts
- New connector framework
- Integrated with BCS
- Integrate search with BI, Workflow, Social, & Collab
- BDC tooling built into SPD
- Application tooling in VS2010



Microsoft
FAST Search Server 2010
for SharePoint

The power to build a whole new
class of search applications

- Public Web Parts with high end capabilities
- Same capabilities
- Integrated using high-end search capabilities
- Same capabilities
- Add custom property extractors
- Extend content processing
- Include external data in relevance
- Build multiple relevance profiles
- Leverage FQL to create geosearch, user-selected influencers, and much more

Customize the OOB user
experience using Web Parts

Extend connectivity using
the BDC and Federation

Combine search with other
SharePoint capabilities

Use familiar tools built for
developer productivity

Leverage advanced
content processing

Customize relevance

Use advanced query
capabilities to create powerful
applications

FAST Search for SharePoint and ESP

FAST SEARCH FOR SHAREPOINT TECHNOLOGY HERITAGE (EXAMPLES)

Many great, familiar technologies from ESP

Technology	Capability
Column/row backend architecture	Extreme scale
Content Pipeline	Metadata extraction, advanced linguistics
FAST Query Language (FQL)	"Search-driven applications"
Relevance profiles	"Contextual Search"

Many new technologies from SharePoint (or elsewhere in Microsoft)

Technology	Capability
Production, Extensible Web Parts	High-end User Experience OOB
Integrated Social Search	People & Expertise Finding, ++
Connectivity framework	Easy Connectivity with tooling in SPD
64 bit, SCOM, Hyper-V, admin console,..	Easy to deploy and manage
Office Web Apps	Thumbnails, Preview, Open in Browser

Summary

- FAST Search for SharePoint is NEW
- It combines the best of high-end search, SharePoint and Microsoft
- It extends SharePoint Server 2010
 - Superset of Capabilities
 - Highly capable and easy to manage

Microsoft TechNet

Search TechNet with Bing

Enterprise Search TechCenter

Home SharePoint Server Search Server Express FAST Search Server for SharePoint

Library 2010


Printer Friendly Version Add To Favorites Send

Enterprise Search TechCenter > FAST Search Server for SharePoint > 2010

FAST Search Server 2010 for SharePoint (Beta)

Compare search technologies

Search Technologies for SharePoint 2010 Products



Download this technical diagram to see a comparison of search capabilities in the following:

- SharePoint Foundation 2010
- Search Server 2010 Express
- Search Server 2010
- SharePoint Server 2010
- FAST Search Server 2010 for SharePoint

System Requirements for FAST Search Server 2010 for SharePoint (Beta)

You can find the system requirements for FAST Search Server 2010 for SharePoint in the downloadable [Deployment and Configuration of FAST Search Server 2010 for SharePoint \(Beta\)](#).

FAST Search Server 2010 for SharePoint Beta resources

- Planning and Architecture for FAST Search Server 2010 for SharePoint (Beta)
- Deployment and Configuration of FAST Search Server 2010 for SharePoint (Beta)
- Monitoring for FAST Search Server 2010 for SharePoint (Beta)
- FAST Search Server 2010 for SharePoint Windows PowerShell Cmdlet Overview (Beta)
- FAST Search Server 2010 for SharePoint Windows PowerShell Cmdlet Help (Beta)

Q&A

PUBLIC RESOURCES

- TechNet: <http://technet.microsoft.com/en-us/enterprisesearch/ee441234.aspx>
- MSDN: <http://msdn.microsoft.com/en-us/library/ee629339.aspx>

The background is a dark gray with a subtle, light gray wireframe grid pattern that curves across the frame. The Microsoft logo is centered in the upper half of the image.

Microsoft[®]

Your potential. Our passion.[™]

© 2009 Microsoft Corporation. All rights reserved. Microsoft, Windows, Windows Vista and other product names are or may be registered trademarks and/or trademarks in the U.S. and/or other countries. The information herein is for informational purposes only and represents the current view of Microsoft Corporation as of the date of this presentation. Because Microsoft must respond to changing market conditions, it should not be interpreted to be a commitment on the part of Microsoft, and Microsoft cannot guarantee the accuracy of any information provided after the date of this presentation. MICROSOFT MAKES NO WARRANTIES, EXPRESS, IMPLIED OR STATUTORY, AS TO THE INFORMATION IN THIS PRESENTATION.