



**SOLID
QUALITY
MENTORS**

Búsquedas en SharePoint 2010

Daniel A. Seara

dseara@solidq.com



MEJORAS



SQLU Summit

En general

- **Escalabilidad**– contenidos y búsquedas ya no son un problema
- **Experiencia de usuario**– navegación visual
- **Búsqueda social**– libera el valor de las personas y su experiencia
- **Relevancia** –mejoras de ranking y soporte de idiomas
- **Plataforma** –herramientas para personalizar, extender o construir aplicaciones
- **Administración y confiabilidad** – mejores “crawls”, Admón por scripts, monitoreo
- **Proces. De contenido enriquecido**– extrae metadatos de contenido empresarial

Productividad de las búsquedas

Productos para cada necesidad



Complete intranet search



High end search delivered through SharePoint

Comunes en toda la línea

- Framework de interfaz común
- Integración y búsquedas sociales
- Integración con la Plataforma SharePoint
- Para usuarios finales y administradores

- Fwk de conector común (BDC)
- Experiencia de desarrollo y APIs
- Capacidades de Admón e instalación
- Ventajas operativas (SCOM, scripting)



SQLU Summit

Conceptos de búsquedas

Search Center - Interfaz de usuarios para realizar consultas e interactuar con los resultados

Query Servers - Acepta consultas y retorna resultados

Query Federation - Retorna resultados desde índices no-SharePoint

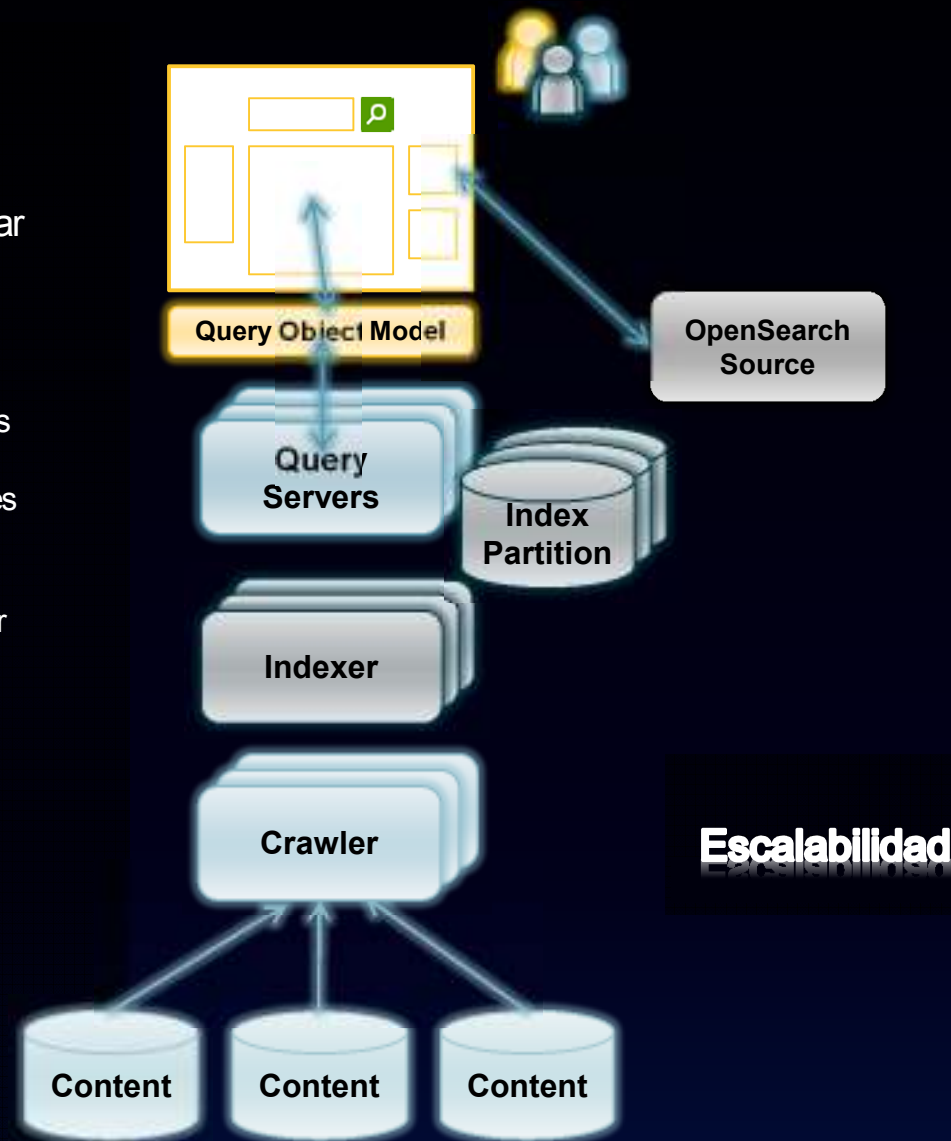
Indexing - Extrae información de los ítems para mejorar eficiencia la coincidencia

Index Partition - Subconjuntos de índices generales

Crawling - Recorre URLs para obtener ítems para el catálogo

Connectors - Saben como procesar distintos contenidos

Content Sources - Orígenes de contenidos



EN DETALLE



SQLU Summit

Interfaz de usuario

- Refinamiento de búsquedas
 - Para propiedades comunes
 - Fácil de personalizar
- Centro de búsquedas
 - Alcances, web parts, "best bets", mejores respuestas, búsq. Avanzada
 - Federación de búsquedas con soporte OpenSearch
- Experiencia centralizada
 - Metadata, taxonomía y "social tags" basados en refinamiento de resultados
 - "Quiso decir" mejorado
 - "Pre y post" sugerencias
 - Vínculo "Ver en el navegador" (para muchos documentos Office)
 - Sintaxis de búsqueda mejorada



Interfaz de usuario

- **Ranking de relevancia**
 - Novedades: URL coincidentes, social tags, resultados de clics previos, coincidencia de frase implícita, extracción de metadata, etc.
 - Mejoras de resúmenes y “expresiones ruido”
- **Multilingües**
 - Detección Automática de muchos documentos
 - Manejo de palabras compuestas “Innovationszyklen” and “innovation”, “zyklen”
 - Ranking en colecciones multilingües
- **Nuevos clientes**
 - Navegadores Smartphone
 - integración en búsquedas de Windows 7

Nueva sintaxis de consultas

- Operadores booleanos y propiedades
 - ("SharePoint Search" OR "Bing Search") AND (title:"keyword syntax" OR title:"query syntax")
- Comodines de búsquedas
 - Micro* author:bill*
- Mejoras en operadores aplicables a propiedades
 - =, >, <, <=, >=
- Se pueden crear rangos de búsquedas

Búsqueda social

- Mejor experiencia
 - Minería de correos para obtener perfiles, intereses, colegas
- Búsquedas estilo “Libreta de direcciones”
 - Coincidencia fonética de nombres
 - Coincidencia de seudónimos
- Relevancia
- Metadata, mejores coincidencias, contenido recién publicado



Arquitectura escalable

- Metas del motor de 2010
 - Latencias de menos de un segundo a gran escala
 - Índices frescos
 - Mejor resolución con mayor disponibilidad
- Filosofía básica
 - Componentización del sistema
 - Remoción de cuellos de botella por escalado

¿Cómo escala 2007?

“Todo el índice”



Query

“Cuello”

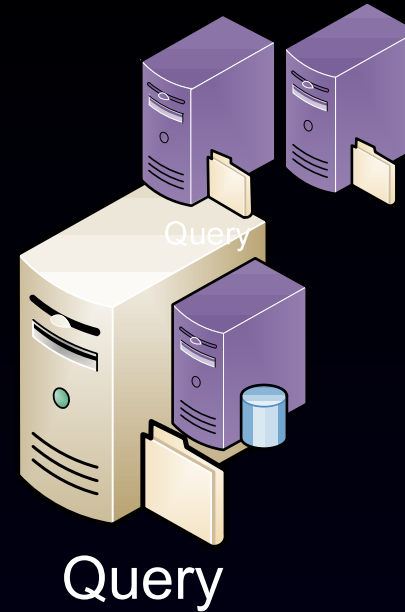
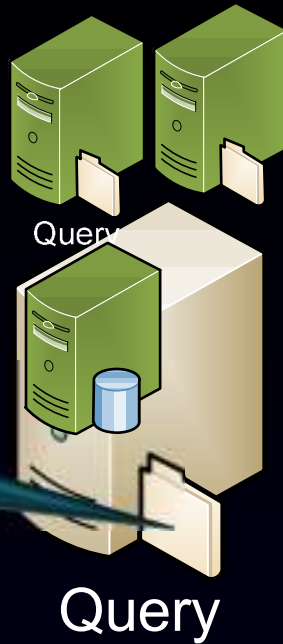
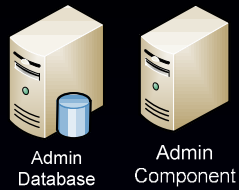


Indexer
Query
+ Query

“Cuello”

“Punto de fallo”

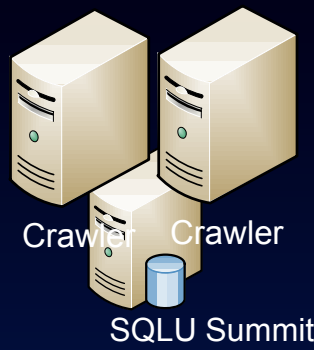
En 2010



- Multiple Indexers
- Stateless Crawlers
- Crawl Distribution
- Query Mirroring
- Query Partitioning
- Multiple Property DBs
- Admin Database + Admin Component

“The whole index”

“Bottleneck”



Distribución de contenido

- “Crawling”
 - Balanceador de carga distribuye entre “crawlers” las bases de datos
 - Cada “Crawler” analiza el contenido de la base de datos que tiene asignada
 - Configuración predeterminada puede modificarse por roles de distribución
 - Ej. Agregar un nuevo conector
- Consultas
 - Baja latencia si todos los índices tienen el mismo tamaño
 - Distribución por hash del documentId
 - Las particiones de crawl indexan y propagan a los servidores de consultas

Otras mejoras del motor

- Soporte para expresiones regulares en reglas de Crawl
- Soporte nativo para repositorios sensibles a mayúsculas
- Habilidad de priorizar Content Sources para distribuir los recursos de crawl
- Nueva 'Política de Crawl' para definir como se tratan los errores
- Menor tiempo de detención de indexación por Backups


Extensibilidad



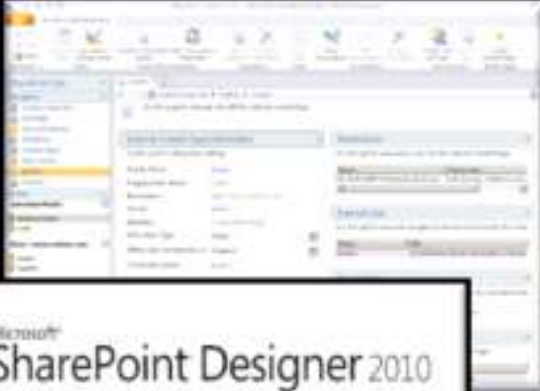
Configure

Extend

Create



Microsoft
SharePoint 2010

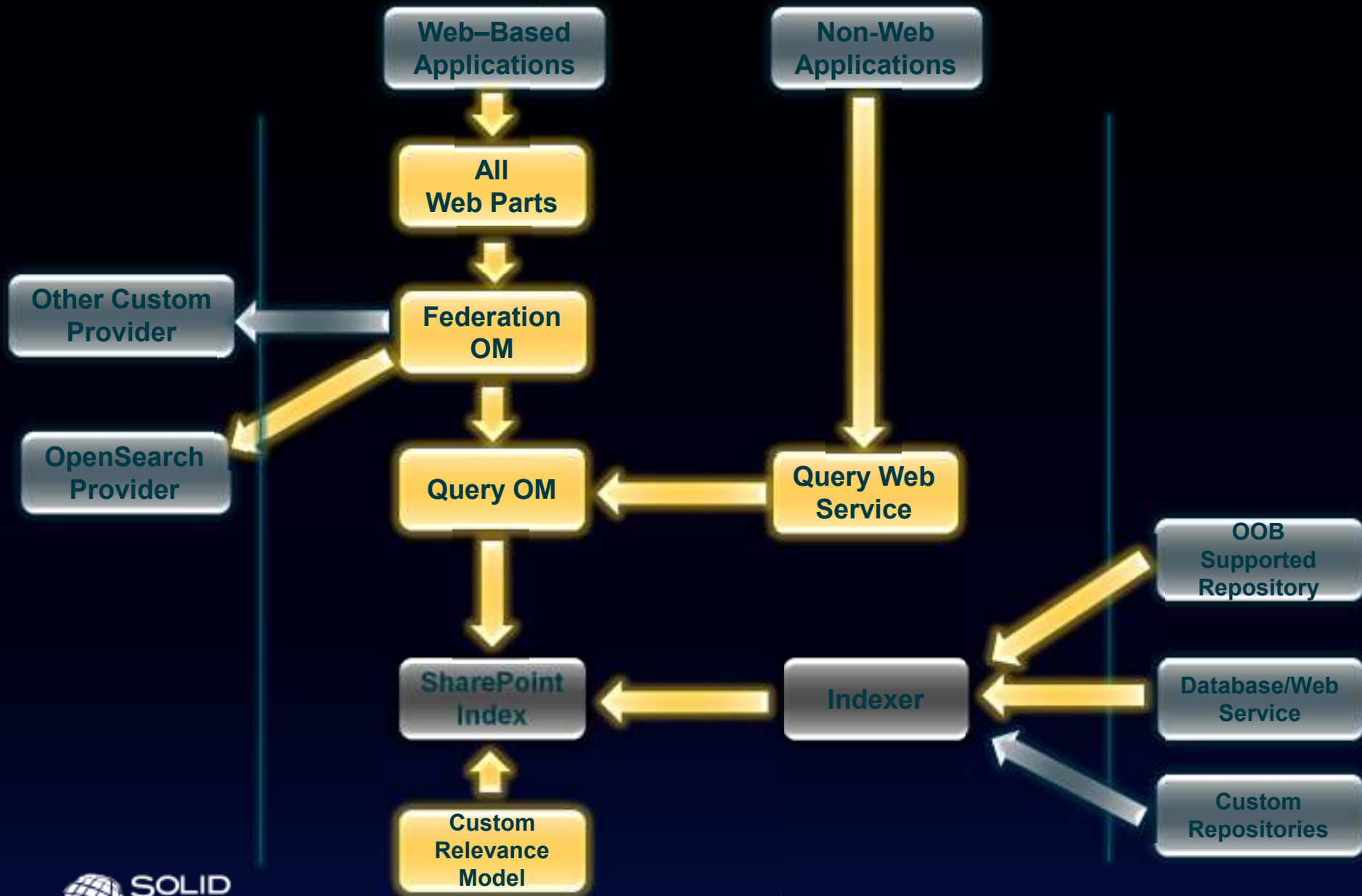


Microsoft
SharePoint Designer 2010



Microsoft
Visual Studio 2008

Arquitectura de la plataforma



Extensibilidad de la interfaz

- Cambiar propiedades del web part sin código
 - XSLT
 - config XML
 - Panel de refinamiento – control de la metadata disponible
 - Búsqueda avanzada – control de la metadata disponible
- Extender web parts del producto programáticamente
 - Todos los WP nativos son públicos
 - Extender web parts para cambiar resultados
 - Ej. Mostrar “pre visualizaciones”
- Extender el Search Center con experiencias verticales
 - Todos los web parts se comunican por una interfaz común (**SharedQueryManager**)

Framework de conectores

- Nuevas características
 - Soporta para adjuntos
 - Seguridad a nivel de Ítem
 - Crawl a través de asociación de ítems
- Mejores opciones de crawl
 - Completo (como en 2007)
 - Incremental basado en Time stamp
 - Crawl por log de cambios y cuenta de eliminaciones

Posibilidad de desarrollar conectores personalizados

Simplicidad de configuración



Tabla resumen de mejoras

Area	Mejoras o nuevas características
Relevancia	Mejores resultados <ul style="list-style-type: none">• Nuevos ingredientes, más experiencias• Mejores separadores en sílabas para más idiomas• Distancia Social, <i>tagging</i>, influencia de clics
Refinamiento	Encontrar más rápido <ul style="list-style-type: none">• Refinamientos basados en Metadata• Sugerencia de búsquedas relacionadas
Social	Conecta con personas y experiencia <ul style="list-style-type: none">• Estilo libreta de direcciones y coincidencia fonética• Análisis de casilla de correo• Comportamiento social; mejora en mySites
Plataforma	Alta escalabilidad <ul style="list-style-type: none">• Partición de índices -> latencia de menos de un segundo para 100.000.000 elementos• Soporte completo a fallos• Mejoras de conectores y relación con datos de negocio
Extensibilidad	Crear aplicaciones basadas en búsquedas <ul style="list-style-type: none">• Objetos públicos, modificables• Sintaxis de búsquedas mejoradas
TCO	Instalación y operación facilitada <ul style="list-style-type: none">• Virtualización e instalación en 64 bits• Powershell, web parts de admón, Modelo de Objetos de admón

Preguntas

