GSA | Praktische Hinweise

Google Search Appliance

Handbuch für Bereitstellungsszenarien

Mai 2014

Handbuch für Bereitstellungsszenarien

In diesem Dokument werden Szenarien für die Bereitstellung einer Google Search Appliance (GSA) beschrieben.

Über dieses Dokument

Die Empfehlungen und Informationen in diesem Dokument sind das Ergebnis unserer praktischen Arbeit vor Ort mit verschiedenen Kunden und in verschiedenen IT-Umgebungen. Wir danken unseren Kunden nd Partnern dafür, dass sie uns ihre Erfahrungen und Erkenntnisse zugänglich gemacht haben.

Inhalt	In diesem Leitfaden werden erweiterte GSA-Konfigurationen beschrieben, die für eine Architektur bei Einbindung weiterer Inhaltsquellen in die GSA erforderlich sein können.
Primäre Zielgruppe	GSA-Administratoren ohne Vorkenntnisse über das Produkt, erfahrene GSA-Administratoren und Functional Analysts für die GSA
IT-Umgebung	Umgebung, in der die GSA für die öffentliche Suche auf Internet- und Intranetwebsites und in freigegebenen Dateien konfiguriert ist
Bereitstellungsphasen	Erstkonfiguration der GSA und Einbindung weiterer Inhaltsquellen in die GSA
Weitere Ressourcen	• Learngsa.com bietet verschiedene Schulungs- und Informationsmaterialien für die GSA. • In der GSA-Produktdokumentation finden Sie umfassende Informationen über die GSA. • Über das Google for Work Support Portal haben Sie Zugriff auf den Google-Support • GSA – Notes from the Field ("GSA – Tipps aus der Praxis") bietet Hilfe bei der Entwicklung und Bereitstellung von Unternehmens-Suchlösungen, die auf der Google Search Appliance (GSA) basieren.

Inhalt

Kapitel 1: Einfache Suche auf einer öffentlichen Website

Überblick über das Szenario

Acme Inc. ist ein großer multinationaler Produzent von Unterhaltungselektronik mit einer umfangreichen externen Webpräsenz. Die Webinhalte des Unternehmens umfassen allgemeine Unternehmensinformationen sowie spezielle Marketingmaterialien für die einzelnen Produktbereiche. Darüber betreibt das Unternehmen einige Supportforen für die eigenen Produkte. Im Anwendungsbeispiel für dieses Szenario möchte das Unternehmen die Google Search Appliance einsetzen, um ein Suchfeld für die allgemeine Suche auf der gesamten Website sowie gesonderte Suchfelder für die einzelnen Produktbereiche bereitzustellen. Alle externen Webinhalte von Acme Inc. sind öffentlich; es gibt keine Zugriffsbeschränkungen für bestimmte Nutzer und/oder Gruppen.

Anforderungen

- Sämtliche im Web verfügbaren öffentlichen Inhalte indexieren
- Ein Feld für die allgemeine Suche bereitstellen, das Suchergebnisse für den gesamten indexierten Websiteinhalt sowie für die einzelnen Produktbereiche zurückgibt
- Gesonderte Suchfelder bereitstellen, deren Ergebnisse sich nur auf einen bestimmten Produktbereich beziehen
- Das Suchfeld und die Ergebnisseite an die Vorgaben für das Unternehmensbranding von Acme Inc. anpassen
- Das Supportforum von Acme Inc. stündlich indexieren, da sich die Inhalte rasch ändern und/oder neue Inhalte hinzukommen
- In Spitzenzeiten 20 Suchanfragen pro Sekunde verarbeiten und bei GSA-Problemen/-Ausfällen hohe Verfügbarkeit bieten
- Inhalte aus verschiedenen Sprachen in den Suchergebnissen nicht vermischen

Annahmen

- Es gibt für die Webinhalte von Acme Inc. jeweils eigene Seiten für jede Sprache.
- Es sind Webpropertys vorhanden, auf denen Suchfelder platziert werden.

Wichtige Aspekte

- Ausreichende Kapazität sicherstellen, um in Spitzenzeiten 20 Suchanfragen pro Sekunde verarbeiten zu können
- Entscheiden, ob Ergebnisse direkt über die GSA oder über eine spezielle Darstellungsschicht einer Webanwendung dargestellt werden sollen
- Anhand von Berichten oder Analysedaten einschätzen, wie Nutzer mit den Suchergebnissen interagieren

Empfohlene Vorgehensweise

Die von Google empfohlene Vorgehensweise für die Implementierung einer einfachen Suche auf einer öffentlichen Website umfasst folgende Bereiche:

- Bereitstellungsarchitektur
- Konfiguration von Crawling und Indexierung
- Frontend-Konfiguration
- Administrative Tätigkeiten

Bereitstellungsarchitektur

Um die allgemeine Last zu bewältigen und zusätzlich Failover-Kapazität bereitzustellen, setzt Acme Inc. insgesamt drei GSAs in einer Produktionskonfiguration ein. Zwei der drei Geräte werden im Sinne der Planung ausreichender Kapazitäten in einer Aktiv/Aktiv-Konfiguration verwendet. Die dritte GSA wird als Hot-Backup für die Ausfallsicherung verwendet.

Acme Inc. konfiguriert die drei GSAs so, dass sie gespiegelt werden. Dabei ist eine GSA das Hauptgerät. Auf ihr sollten alle Konfigurationsänderungen durchgeführt werden. Um eine hochverfügbare Aktiv/Aktiv-Konfiguration zu erhalten, wird der GSA ein Load Balancer vorgeschaltet. Der Load Balancer erfüllt die beiden folgenden Funktionen:

- Er verteilt den Suchanfragen-Traffic gleichmäßig auf die beiden aktiv/aktiv konfigurierten GSAs.
- Er sendet einen Ping an die beiden aktiv/aktiv konfigurierten GSAs. Erhält er von einem der aktiven Geräte keine Antwort, schaltet er von diesem Gerät auf die Hot-Backup-Einheit um.

Da die GSA in einer bestehenden Webanwendung bereitgestellt wird, empfiehlt Google, Suchanfragen und Antworten der GSA über eine spezielle Darstellungsschicht einer Webanwendung zu verarbeiten. In diesem Fall wird die GSA als Dienst verwendet. Dabei reicht die Weboberfläche Suchanfragen weiter und parst dann die von der GSA zurückgegebenen Suchergebnisse im XML-Format gemäß den Marketing- und Brandingrichtlinien für die Seitenformatierung. Da die Suchseite nicht direkt auf der GSA ausgegeben wird, sollte die GSA nicht im öffentlichen Netzwerk sichtbar sein, hinter einer Firewall im Netzwerk von Acme Inc. platziert werden. Dabei sollte die Umfeldsicherheit über Netzwerkfirewalls hergestellt werden.

Konfiguration von Crawling und Indexierung

Acme Inc. konfiguriert jeweils eine Sammlung für jede Sprache des Webauftritts. So kann der Parameter `site` verwendet werden, um die Anfragen nach Sprache zu sortieren – je nachdem, auf welcher Seite der Nutzer die Suche gestartet hat. Da für jeden Produktbereich eine auf die eigenen Dokumente beschränkte Suche gewünscht wird, konfiguriert Acme Inc. außerdem Sammlungen für die einzelnen Produktbereiche.

Acme Inc. konfiguriert Start-URLs für Top-Level-Seiten. Für Inhalte, die sich häufig ändern, kann die Crawling-Rate entsprechend angepasst werden, um sicherzustellen, dass die Inhalte mindestens einmal täglich gecrawlt werden. Um das Crawling noch gezielter zu steuern, können die Admin API oder ein Webfeed verwendet werden, um sicherzustellen, dass bestimmte Seiten mehrmals am Tag in die Crawling-Warteschlange aufgenommen werden.

Frontend-Konfiguration

Jedes Suchfeld auf der Webpräsenz von Acme Inc. ist mit einer Reihe von Suchanfrageparametern verknüpft. Diese Parameter werden zusammen mit der Suchanfrage an die GSA weitergereicht. Sie sorgen dafür, dass genau die passenden Ergebnisse auf der Suchergebnisseite angezeigt werden.

Ein Beispiel: Ein Suchfeld auf einer englischen Produktseite sollte sowohl die Sammlungsparameter für die Sprache Englisch als auch für den jeweiligen Produktbereich weitergeben. Die Dokumenttypdefinition der Ergebnisdatei sollte geprüft werden, um zu ermitteln, in welchen XML-Elementen die GSA Informationen zurückgibt. Diese Elemente sollten vom Frontend geparst und dann auf der Seite entsprechend angezeigt werden.

Administrative Tätigkeiten

Acme Inc. erstellt mithilfe der Funktion "Erweiterte Suchberichte" Berichte, die aufzeigen, wonach Nutzer gesucht und worauf sie in den Suchergebnisseiten geklickt haben. Die Berichte sollten häufig erstellt und analysiert werden, da sie ein guter Anhaltspunkt für die allgemeine Zufriedenheit mit der Suchfunktion sind.

Alternative Vorgehensweise

Auch die Suche direkt über die GSA – ohne vorgeschaltete Weboberfläche – ist vorstellbar. Acme Inc. könnte dabei das Stylesheet für ein Frontend entsprechend anpassen. Eine umfassende Anpassung ist hier allerdings schwieriger. Andererseits ist für diese Vorgehensweise eventuell weniger Entwicklungsaufwand nötig und die neuen, vorkonfigurierten Frontend-Funktionen, die auf der GSA verfügbar sind, können so leichter genutzt werden.

Achten Sie bei dieser Vorgehensweise darauf, dass im GSA-Index keine gesicherten Inhalte als "Öffentlich" markiert sind, da Nutzer direkt auf der GSA Suchanfragen ausführen können. Ein Reverseproxy kann verwendet werden, um den Zugriff auf die GSA einzuschränken, indem erlaubte URL-Muster auf die weiße Liste gesetzt werden.

Die Anzahl der auf der GSA definierten Sammlungen sollte sich unter 200 bewegen. Um dies zu erreichen, können Sie, statt separate Sammlungen für die Produktbereiche anzulegen, auch einen Metadatenparameter für jeden der Produktbereiche verwenden. Der Parameter wird dann zusammen mit den Inhalten indexiert. Er wird als Filter auf Suchanfragen für einen bestimmten Produktbereich angewendet und bewirkt, dass die GSA nur Inhalte für den betreffenden Produktbereich abruft.

Überblick über die Projektaufgaben

In der folgenden Tabelle sind die Projektaufgaben und Aktivitäten für die Implementierung einer einfachen Suche auf einer öffentlichen Website aufgelistet.

Aufgabe	Aktivität
Bereitstellungsarchitektur planen	• GSAs im Rack montieren und verkabeln • Geräte konfigurieren und Spiegelung einrichten • Load Balancer vor den GSAs konfigurieren • Umfeldsicherheit für die GSAs einrichten
Crawling und Indexierung konfigurieren	• Start-URLs für das Crawlen von Inhalten einrichten • Sammlungen für Sprachen und Produkteinheiten konfigurieren • Inhalte ermitteln, die sich häufig ändern, und sicherstellen, dass diese ein- oder mehrmals täglich indexiert werden
Frontend konfigurieren	• Vorhandene Weboberfläche für das Hinzufügen von Suchfeldern aktivieren • XML-Antwort von der GSA parsen und Ergebnisse gemäß den Unternehmensrichtlinien für Benutzeroberflächen darstellen

Langfristige Optimierungen

- Suche und Funktionen anhand der Berichte über Suchmuster der Nutzer optimieren
- Inhalte für KeyMatches ermitteln
- Komplexere Synonymlisten aktivieren

Dynamische Navigation basierend auf Metadaten und Attributen aktivieren

Kapitel 2: Einfache interne Suche

Überblick über das Szenario

Acme Inc. hat eine große interne Webpräsenz, die sich über verschiedene Teile der Welt erstreckt. Im Anwendungsbeispiel für dieses Szenario möchte das Unternehmen die Suche für alle internen Websites und -seiten an einem Ort zusammenführen, sodass die Mitarbeiter beim Recherchieren von Informationen nicht extra verschiedene Websites aufrufen müssen. Obwohl alle Nutzer Zugriff auf das Intranet von Acme Inc. haben, können nicht alle von ihnen die Informationen auf den verschiedenen Websites in der Unternehmensdomain aufrufen. So ist beispielsweise der Zugriff auf Informationen aus der Personalabteilung bei der Suche wünschenswert. Folglich ist die gesicherte Bereitstellung personenbezogener Daten eine wichtige Anforderung.

Anforderungen

- Folgende Inhalte indexieren:
    - Dateifreigaben im Unternehmen
    - Interne Webseiten
    - Informationen aus der Personalabteilung
- Ein Feld für die allgemeine Suche bereitstellen, das eine Ergebnisseite für den gesamten indexierten Inhalt und für die einzelnen Produktbereiche zurückgibt
- Gesonderte Suchfelder bereitstellen, deren Ergebnisse sich nur auf einen bestimmten Produktbereich beziehen
- Das Suchfeld und die Ergebnisseite gemäß den Vorgaben für das Unternehmensbranding von Acme Inc. anpassen
- Suchergebnisse für gesicherte Inhalte nur für Nutzer anzeigen, die zum Ansehen der Inhalte berechtigt sind
- Failoverfunktionen bei Problemen/Ausfällen einer GSA bieten

Annahme

Es ist ein Mechanismus zur Authentifizierung von Nutzern vorhanden.

Wichtige Aspekte

- Entscheiden, ob Ergebnisse direkt über die GSA oder über eine spezielle Weboberfläche ausgegeben werden sollen
- Entscheiden, ob die Sicherheit mithilfe der GSA oder über die der GSA vorgeschaltete Anwendung verwaltet werden soll
- Entscheiden, wie der Google Search Appliance-Connector für Dateisysteme konfiguriert werden soll, um Inhalte aus freigegebenen Dateien zu indexieren.
- Anhand von Berichten oder Analysedaten einschätzen, wie Nutzer mit den Suchergebnissen interagieren

Empfohlene Vorgehensweise

Die von Google empfohlene Vorgehensweise für die Implementierung einer einfachen internen Suche umfasst folgende Bereiche:

- Bereitstellungsarchitektur
- Konfiguration von Crawling und Indexierung
- Konfiguration der sicheren Suche
- Frontend-Konfiguration
- Administrative Tätigkeiten

Bereitstellungsarchitektur

Um auch Failoverfunktionen zu berücksichtigen, werden bei Acme Inc. insgesamt zwei GSAs in einer Produktionskonfiguration eingesetzt. Die beiden GSAs werden in einer Aktiv/Passiv-Konfiguration verwendet. Eine GSA dient dabei als Hauptgerät und die andere als Hot-Backup bei Ausfällen. Acme Inc. konfiguriert beide GSAs so, dass sie gespiegelt werden. Eine GSA ist dabei das Hauptgerät. Auf ihr sollten alle Änderungen an der Konfiguration durchgeführt werden. Um eine Aktiv/Passiv-Konfiguration zu erhalten, wird ein Load Balancer vor die GSAs geschaltet. Die Rolle des Load Balancers sieht wie folgt aus: Er sendet einen Ping an die aktive GSA. Erhält er von dem aktiven Gerät keine Antwort, schaltet er von diesem Gerät auf das Hot-Backup-Gerät um.

Da die GSA intern bereitgestellt wird, wird empfohlen, die Ergebnisse direkt über die GSA auszugeben und mithilfe des GSA-Stylesheets anzupassen. In diesem Fall kann Acme Inc. das Stylesheet mit dem Layout-Assistenten ändern, einem XSLT-Assistenten auf der GSA, mit dem rasch bestimmte Funktionen zur Anzeige hinzugefügt werden können. Sind weitere Änderungen erwünscht, kann auch das Stylesheet entsprechend manuell geändert werden. Beachten Sie, dass das ESO keine benutzerdefinierten XSLT-Änderungen unterstützt.

Ein Reverseproxy wird in der Architektur benötigt, wenn Query Fidelity (d. h. die korrekte Abbildung des Umgebungszustands in den Suchergebnissen) erforderlich ist. Dadurch wird sichergestellt, dass Suchanfragenparameter nicht manipuliert werden oder nicht ad-hoc an die GSA übermittelt werden können. Wenn gesicherte Inhalte im Index als "Öffentlich" markiert sind und Sicherheit durch eine Anwendungsschicht auf der Basis von Metadaten realisiert wird, sollte den GSAs ein Reverseproxy vorgeschaltet und sollten Suchanfragen gefiltert werden, damit niemand direkt auf die GSA zugreifen und eigene Suchanfragen senden kann. Dies ist erforderlich, um sicherzustellen, dass Nutzer die URL nicht so manipulieren, dass sie mit Metadaten versehene Inhalte sehen, die für sie eigentlich nicht sichtbar sein sollten oder aus einer Sammlung stammen, für die sie keine Zugriffsrechte haben.

Der Google Search Appliance-Connector für Dateisysteme, der zum Indexieren von Inhalten aus Dateifreigaben verwendet wird, sollte auf einem externen Server in einer Produktionsumgebung gehostet werden. Der Connector läuft in einer JVM und wird mit integriertem Tomcat bereitgestellt.

Konfiguration von Crawling und Indexierung

Acme Inc. konfiguriert Start-URLs für Top-Level-Seiten. Um Inhalte nach den einzelnen Abteilungen von Acme Inc. zu unterscheiden, kann für jede der Abteilungen eine Sammlung angelegt werden.

Zum Indexieren von Dateifreigaben sollte der Google Search Appliance-Connector für Dateisysteme verwendet werden. Der Connector unterstützt Folgendes:

- Early-Binding-Autorisierung (ACLs)
- Szenarien, in denen Datumsinformationen über den letzten Zugriff auf durchsuchte Dateien und Verzeichnisse gepflegt werden müssen
- Bei der Freigabe handelt es sich um eine nicht über HTTP zugängliche Windows-DFS-Domain-Root-Freigabe.

Konfiguration der sicheren Suche

Acme kann zur Sicherung von Inhalten eine der folgenden Strategien verwenden, die sich je nachdem unterscheidet, ob eine Autorisierung erforderlich ist:

- Nur Authentifizierung erforderlich, keine Autorisierung
- Authentifizierung und Autorisierung erforderlich

Nur Authentifizierung erforderlich

Wenn eine Authentifizierung erforderlich ist, aber keine Autorisierung:

- Crawlen Sie Inhalte mit einem Admin-Konto und markieren Sie sie als "Öffentlich".
- Platzieren Sie die gecrawlten Inhalte in einer Sammlung.
- Die Anwendungsebene über der GSA übernimmt die Authentifizierung. Sobald ein Nutzer für eine Seite mit einem Suchfeld authentifiziert wurde, wird eine Suche in der Sammlung durchgeführt, in der die Inhalte platziert waren.
- Bei dieser Strategie werden öffentliche und gesicherte Inhalte vermischt. Wenn Sie möchten, dass bestimmten Nutzern keine gesicherten Inhalte angezeigt werden, verwenden Sie einen Reverseproxy vor der GSA. So kann sichergestellt werden, dass nur zulässige Suchanfragen an die GSA gesendet werden. Der Reverseproxy sorgt dafür, dass nicht authentifizierte Nutzer nicht direkt auf die GSA zugreifen können, wo Nutzer ihre eigenen Suchparameter erstellen und auf Suchanfragen durchführen können.

Beachten Sie, dass mit einem Reverseproxy eine weitere Komponente zur Architektur hinzukommt. Weitere Informationen finden Sie in Kapitel 7: Reverseproxy für Umfeldsicherheit und aus anderen Gründen implementieren.

Authentifizierung und Autorisierung erforderlich

Wenn sowohl eine Authentifizierung als auch eine Autorisierung erforderlich sind:

- Crawlen Sie Inhalte mit dem Konto eines Nutzers, der Zugriff hat, und markieren Sie sie nicht als "Öffentlich".
- Nutzer müssen möglicherweise ihre Anmeldedaten eingeben, wenn sie eine Suche ausführen. Die Suchergebnisse werden mittels HEAD-Anfragen in den Dienst-Endpunkten autorisiert.

- Legen Sie fest, wie Sie die Einbindung in die verfügbaren Authentifizierungsmechanismus realisieren möchten. Es gibt folgende Möglichkeiten:
    - Kerberos, weitere Informationen siehe das Kerberos-Szenario in Kapitel 6
    - Integrierte Windows-Authentifizierung (NTLM) mithilfe der SAML Bridge
    - LDAP oder einfache Eingabeaufforderung für Nutzername/Passwort durch die GSA
    - Cookie-Umsetzung für die Einbindung in die Formularauthentifizierung und Rückgabe eines bestätigten Nutzernamens an die GSA

Frontend-Konfiguration

Jedes Suchfeld auf der Webpräsenz ist mit einer Reihe von Suchanfrageparametern verknüpft. Diese Parameter werden zusammen mit der Suchanfrage an die GSA weitergereicht. Sie sorgen dafür, dass genau die passenden Ergebnisse auf der Suchergebnisseite angezeigt werden.

So gibt ein Suchfeld auf der Seite der Personalabteilung beispielsweise die Sammlungsparameter für diese Abteilung weiter. Google empfiehlt, dass Acme Inc. die Ergebnisseite mithilfe des Layout-Assistenten anpasst. So können bestimmte Funktionen aktiviert oder deaktiviert werden. Ein weiterer Vorteil bei der Verwendung des XSLT-Assistenten ist die höhere Wahrscheinlichkeit, dass zukünftige Versionen des XSLT damit kompatibel sind.

Administrative Tätigkeiten

Acme Inc. erstellt mithilfe der Funktion "Erweiterte Suchberichte" Berichte, die aufzeigen, wonach Nutzer gesucht und worauf sie in den Suchergebnisseiten geklickt haben. Die Berichte sollten häufig erstellt und analysiert werden, da sie ein guter Anhaltspunkt für die allgemeine Zufriedenheit mit der Suchfunktion sind.

Alternative Vorgehensweisen

Sie können für die sichere Suche "Nur Authentifizierung erforderlich" anstelle einer Anwendung, die sich vor der GSA befindet und die Authentifizierung durchführt, auch die Funktion "Umfeldsicherheit" auf der GSA verwenden. Sie stellt sicher, dass die GSA ohne Nutzerauthentifizierung keine Ergebnisse bereitstellt. Wenn die Umfeldsicherheit aktiviert ist, muss die GSA Nutzer erst mit einem der konfigurierten Mechanismen authentifizieren, bevor Suchergebnisse bereitgestellt werden. Treten Fehler bei der Authentifizierung auf, stellt die GSA keine Ergebnisse bereit, auch keine öffentlichen Ergebnisse.

Verwenden Sie Richtlinien-ACLs für Inhalte, auf die nur authentifizierte Nutzer Zugriff haben. Bei diesem Ansatz kann eine Gruppe "Jeder" verwendet werden, um den Zugriff auf diese Inhalte zu steuern. Die Gruppe "Jeder" muss hierfür zum Zeitpunkt der Authentifizierung aufgelöst werden.

Überblick über die Projektaufgaben

Die folgenden Tabelle zeigt die Projektaufgaben und Aktivitäten für die Implementierung einer einfachen internen Suche.

Aufgabe	Aktivität
Bereitstellungsarchitektur planen	• GSAs im Rack montieren und verkabeln • GSAs konfigurieren und Spiegelung einrichten • Load Balancer vor den GSAs konfigurieren • Umfeldsicherheit für die GSAs einrichten • Server beschaffen, auf dem der Dateisystem-Connector gehostet werden soll
Crawling und Indexierung konfigurieren	• Start-URLs für das Crawlen von Inhalten einrichten • Sammlungen für die einzelnen Abteilungen konfigurieren • Dateisystem-Connector installieren und konfigurieren • Sicherheitsmechanismen ermitteln und Crawler-Zugriff konfigurieren
Frontend konfigurieren	• Vorhandene Weboberfläche für das Hinzufügen von Suchfeldern aktivieren • XSLT-Änderungen für die einzelnen Frontends mithilfe des Assistenten konfigurieren

Langfristige Optimierungen

- Suche und Funktionen anhand der Berichte über Suchmuster der Nutzer optimieren
- Inhalte für KeyMatches ermitteln
- Komplexere Synonymlisten aktivieren
- Entitätserkennung aktivieren, um Dokumente mithilfe von textbasierten Wörterbüchern, Begriffen oder regulären Ausdrücken automatisch mit Metadaten anzureichern
- Dynamische Navigation anhand von Metadaten und Attributen aktivieren
- Expertensuche in Büro- und/oder Abteilungseinträgen aktivieren
- Bereiche ermitteln, für die OneBoxes nützlich sein können

Kapitel 3: Interne Suche im Intranet, im Dateisystem und in SharePoint

Überblick über das Szenario

Acme Inc. stellt verschiedene Datenkorpora bereit, die auf unterschiedlichen Servern im Netzwerk des Unternehmens zugänglich sind. Diese Datenspeicher können über verschiedene Datenverwaltungsanwendungen, z. B. SharePoint, sowie als gesicherte Dateifreigaben aufgerufen werden.

Für die Informationsrecherche müssen die Mitarbeiter also verschiedene Anwendungen öffnen, was mühsam und zeitaufwendig ist. Und nicht nur das: Die Produktivitätseinbußen durch wiederholte Informationsrecherche in voneinander getrennten Systemen und durch die bestehenden, ineffizienten Suchtools machen sich mittlerweile im Betriebsergebnis bemerkbar.

Anforderungen

- Folgende Inhalte gesichert indexieren:
    - Gesicherte Dateifreigaben
    - Daten aus dem SharePoint-Portal, die für das Hosting interner Websites verwendet werden
- Suchergebnisse für gesicherte Inhalte nur für Nutzer anzeigen, die zum Ansehen der Inhalte berechtigt sind
- Standardbenutzeroberfläche für den Datenzugriff erstellen
- Benutzerdefinierte Oberflächen für interne und externe Nutzer erstellen
- Eingerichtetes System muss messbaren geschäftlichen Nutzen bringen

Annahmen

- In SharePoint sind mehr als 500.000 Dokumente vorhanden.
- Eine automatische Lösung für die Analyse ist wünschenswert.

Wichtige Aspekte

- Entscheiden, ob der integrierte oder der externe Google Search Appliance-Connector für SharePoint verwendet werden soll
- Entscheiden, ob webfähige SMB-Dateifreigaben gecrawlt oder der Dateisystem-Connector verwendet werden soll
- Entscheiden, ob Ergebnisse direkt über die GSA oder über eine spezielle Darstellungsschicht einer Webanwendung ausgegeben werden sollen
- Entscheiden, ob die Sicherheit mithilfe der GSA oder über eine vorgeschaltete Anwendung verwaltet werden soll
- Entscheiden, ob eine stille Authentifizierung erforderlich ist, bei der Nutzer nicht mehrmals von der GSA aufgefordert werden, ihre Anmeldedaten einzugeben

Empfohlene Vorgehensweise

Die von Google empfohlene Vorgehensweise für die Implementierung einer internen Suche im Intranet, im Dateisystem und in SharePoint umfasst folgende Bereiche:

- Nutzwertanalyse
- Bereitstellungsarchitektur
- Konfiguration von Crawling und Indexierung
- Konfiguration von Serve-Time-Authentifizierung und -Autorisierung

Nutzwertanalyse

Um den geschäftlichen Nutzen der eingerichteten Suchlösung zu beurteilen, führt Acme Inc. eine kurze Studie durch, um die Zeit zu erfassen, die auf den bestehenden Plattformen für die Suche nötig ist. Zur Erfassung dieser Informationen sollten automatisierte Tools verwendet werden. Wenn Analysetools vorhanden sind, sollten diese verwendet werden, Informationen zur Nutzung oder Dauer der Suche auf den aktuellen Systemen zu erfassen. Sind keine Analysetools vorhanden, wäre es für Acme Inc. sinnvoll, ein Analyseprogramm zu implementieren, mit dem die Effektivität der Suche künftig automatisch evaluiert werden kann.

Nach Abschluss der GSA-Einrichtung führt Acme Inc. eine Evaluation durch, um die Effektivität der neuen Lösung zu messen. Um die richtigen Messwerte zu erkennen, wird das Unternehmen ähnliche Anwendungsbeispiele zum Vergleich heranziehen, die schon vor Beginn der Bereitstellung evaluiert wurden.

Bereitstellungsarchitektur

Acme Inc. wird den externen SharePoint-Connector verwenden, da mehr als 500.000 SharePoint-Dokumente vorhanden sind. Wenn Dateifreigaben webfähig gemacht werden können, können sie direkt von der GSA gecrawlt werden.

Die Ergebnisse werden direkt über die GSA ausgegeben, indem benutzerdefinierte Frontends für verschiedene Datenspeicher verwendet werden. Für die Suche in SharePoint wird das Suchfeld für SharePoint eingerichtet und verwendet.

Für die Indexierung von Dateifreigaben bietet sich der Google Search Appliance-Connector für Dateisysteme an. Der Connector eignet sich zum Beispiel für folgende Szenarien:

- Early-Binding-Autorisierung (ACLs)
- Szenarien, in denen Datumsinformationen über den letzten Zugriff auf durchsuchte Dateien und Verzeichnisse gepflegt werden müssen
- Bei der Freigabe handelt es sich um eine nicht über HTTP zugängliche Windows-DFS-Domain-Root-Freigabe.

Konfiguration von Crawling und Indexierung

Acme Inc. konfiguriert Crawling und Indexierung für die folgenden Arten von Inhaltsquellen:

- SharePoint: Um SharePoint-Inhalte zu indexieren, installiert und konfiguriert Acme den SharePoint-
  Connector auf einem separaten Server. Außerdem werden auf jedem SharePoint-Web-Frontend in der Serverfarm Google-Webdienste für SharePoint installiert. Als Connector-Konfigurationsoption werden ACLs in die GSA eingelesen. Da zur Indexierung von SharePoint-Inhalten der SharePoint-
  Connector verwendet wird, ist der Active Directory-Gruppen-Connector erforderlich, um die Active Directory-Gruppenmitgliedschaften eines Nutzers während des Servings aufzulösen. Diese werden für die Inhaltsautorisierung und für die Zuordnung von Active Directory-Nutzern/-Gruppen
  zu lokalen SharePoint-Gruppen benötigt.

- Dateifreigaben: Um Dateifreigaben zu indexieren, konfiguriert Acme webfähige Dateifreigaben auf der Seite **Inhaltsquellen > Web-Crawling > Start- und Sperr-URLs** (vor Version 7.2: **Crawl und Index > URLs crawlen** in der GSA-Admin-Konsole.

Konfiguration von Serve-Time-Authentifizierung und -Autorisierung

Acme Inc. verwendet Kerberos als bevorzugten Authentifizierungsmechanismus zwischen der GSA und dem Inhaltsserver. Hierfür werden folgende Aufgaben ausgeführt:

- Active Directory-Dienstkonto für die GSA erstellen
- Kerberos auf der GSA konfigurieren
- Den SharePoint-Connector konfigurieren, damit dieser Inhaltsfeeds einschließlich ACLs-Feeds an die GSA übermittelt
- Active Directory-Gruppen-Connector für die Cache-Speicherung von Active Directory-Objekten in einer Datenbank konfigurieren, damit Gruppen während des Servings schnell aufgelöst werden können
- SharePoint-Connector-Instanz als Authentifizierungsmechanismus konfigurieren, um Gruppen für die Sitzung eines Nutzers aufzulösen, die für die Filterung von Inhalten im Index anhand der ACL verwendet werden sollen.

Da Inhaltsfeeds einschließlich ACLs von SharePoint an die GSA übermittelt werden, werden Inhalte im Index der GSA mithilfe der ACLs autorisiert, die beim Inhaltscrawling eingelesen wurden.

Alternative Vorgehensweisen

- GSA-Connector für Dateisysteme verwenden, um Inhalte von Dateifreigaben zu indexieren
    - Der Vorteil bei dieser Vorgehensweise besteht darin, dass ACLs zusammen mit den Inhalten eingelesen werden, was eine Early-Binding-Autorisierungsentscheidung ermöglicht, die eine bessere Leistung bietet.
    - Damit dies funktioniert, müssen die richtigen Gruppen für den Nutzer dann bei der Authentifizierung aufgelöst werden, da für die Early-Binding-ACL-Autorisierungsentscheidung Gruppen benötigt werden. Zu diesem Zweck könnte der Active Directory-Gruppen-Connector verwendet werden, der auch für SharePoint erforderlich ist.

Überblick über die Projektaufgaben

Die folgende Tabelle zeigt die Projektaufgaben und Aktivitäten für die Implementierung einer internen Suche im Intranet, im Dateisystem und in SharePoint.

Aufgabe	Aktivität
Bereitstellungsarchitektur planen	• SharePoint/Active Directory-Gruppen-Connector-Server konfigurieren
Crawling und Indexierung konfigurieren	• SharePoint-Connector konfigurieren • Active Directory-Gruppen-Connector konfigurieren • Dateifreigaben-Orte unter "Crawl und Index" in der GSA-Admin-Konsole konfigurieren
Frontend konfigurieren	• Frontends für verschiedene Datenspeicher anpassen
Serve-Time-Authentifizierung/-Autorisierung konfigurieren	• GSA für Kerberos aktivieren • Connectorbasierte Autorisierung konfigurieren • Mechanismus für die Gruppenauflösung konfigurieren ○ Wenn der SharePoint-Connector verwendet wird, ist dies mit hoher Wahrscheinlichkeit eine auf dem SharePoint-Connector basierende Authentifizierung, die ausschließlich für die Gruppenauflösung konfiguriert wird.

Langfristige Optimierungen

Implementieren Sie das Google-Suchfeld für SharePoint, um eine Suche direkt in SharePoint zu ermöglichen.

Kapitel 4: Über Feeds indexieren

Überblick über das Szenario

Acme Inc. besitzt eigene Ladengeschäfte, die unter zwei verschiedenen Marken betrieben werden. Im Anwendungsbeispiel für dieses Szenario möchte das Unternehmen eine Suche implementieren, bei der die Mitarbeiter der Ladengeschäfte die Produktdatenbank durchsuchen können. Die Produkte werden aktuell in einer Datenbank gespeichert; bestimmte Daten befinden sich dabei in Unternehmensanwendungen. Der Crawler kann die Produktseiten nicht direkt indexieren. Es gibt ein Web-Frontend, das Produktinformationen anzeigt, wenn die Produktnummer in der URL bereitgestellt wird.

Anforderungen

- Inhalte über Produkte indexieren
- Suche innerhalb bestimmter Einzelhandelsmarken bereitstellen
- Eine parametrische Navigation im linken Bereich ermöglichen nach:
    - Preis
    - Kategorie
    - Zeitpunkt des Produkteingangs

Annahmen

- Es gibt ein Web-Frontend, das Produktseiten anzeigen kann. Dieses Web-Frontend enthält nicht alle für die Indexierung der Produkte erforderlichen Metadaten.
- Die Produkte sind nicht gesichert, sondern für beliebige Nutzer sichtbar.

Wichtige Aspekte

- Entscheiden, ob der GSA-Connector für Datenbanken verwendet werden soll, um die Produktdatensätze auf die GSA zu übertragen
- Entscheiden, ob ein Inhaltsfeed oder ein Webfeed verwendet werden soll, um die Produktdatensätze auf die GSA zu übertragen
- Metadaten für die Indexierung zusammen mit den Inhalten definieren, die für die dynamische Navigation und erweiterte Suchfunktionen verwendet werden

Empfohlene Vorgehensweise

Die von Google empfohlene Vorgehensweise für die Indexierung über Feeds umfasst folgende Bereiche:

- Bereitstellungsarchitektur
- Konfiguration von Crawling und Indexierung
- Metadaten im Fokus
- Frontend-Konfiguration

Bereitstellungsarchitektur

Da die Datensätze zusammen mit den erforderlichen Metadaten nicht nur mithilfe von Datenbankabfragen erstellt werden können, wird der Datenbank-Connector nicht für die

Inhaltsindexierung verwendet. Acme Inc. verwendet stattdessen einen benutzerdefinierten Feed. Ein benutzerdefinierter Feed ist eine Anwendung, die XML-Dateien erstellt, die die Datensätze für die Indexierung auf der GSA enthalten. Der wichtigste Schritt, um die XML-Dateien mit den Datensätzen auf die GSA zu übertragen, ist eine POST-Aktion auf der Feeds-Protokollschnittstelle auf der GSA.

Zusätzlich zur GSA ist ein weiterer Server (Windows oder Linux) erforderlich, auf dem die Feeds-Anwendung gehostet wird. Diese Anwendung führt einige Abschnitte der Programmlogik aus, um einen Datensatz zu erstellen und Datensätze auf die GSA zu übertragen.

Konfiguration von Crawling und Indexierung

Es wird empfohlen, Inhaltsfeeds zu verwenden, um die Produktdatensätze auf die GSA zu übertragen. Auf diese Weise können Inhalte für die Indexierung individuell angepasst werden. Bei dieser Vorgehensweise wird auch die Cachefunktion der GSA genutzt, bei der benutzerdefinierte Produktseiten im GSA-Index zwischengespeichert werden. So können Mitarbeiter in den Ladengeschäften einfach die Version aus dem Cache aufrufen – und den umständlichen Weg über das bestehende Produkt-Web-Frontend vermeiden.

Die Feeds-Anwendung muss so konzipiert sein, dass sie für jede Produktzeile in der Datenbank das erforderliche HTML und die zugehörigen Metadaten erstellen und in die GSA einlesen kann. Es wird ein Mechanismus benötigt, der alle gelöschten, geänderten und hinzugefügten Datensätze nachverfolgt.

Metadaten im Fokus

Metadaten sind im Umgang mit Produktinhalten besonders wichtig. Metadaten helfen Endnutzern dabei, gezielter zu suchen und Inhalte präzise anhand bestimmter Kategorien einzugrenzen. Acme Inc. kann bestimmte Metadaten ermitteln, die für Überschriften in der dynamischen Navigation verwendet werden können. Nutzer können dann mit nur einem Mausklick gezielt verschiedene Metadatenkategorien auswählen. Andere Metadatenwerte können in Suchanfragen verwendet werden, um die Anfrage auf bestimmte Inhalte einzugrenzen.

Frontend-Konfiguration

Wenn Inhaltsfeeds zum Einlesen von Inhalten in den Index der GSA verwendet werden, bietet das u. a. folgenden Vorteil: Auf der GSA wird eine gecachte Version des benutzerdefinierten Inhalts gespeichert, der für den Feed erstellt wurde. Diese Version kann dann auf dem Frontend angezeigt werden.

Ein Beispiel hierfür wäre die Indexierung von druckfreundlichen Seiten, die ausgedruckt und als Datenblätter ausgegeben werden können. Möchte ein Nutzer z. B. eine Produktüberblicksseite ansehen, so kann er den gecachten Inhalt auf der GSA aufrufen. Möchte der Nutzer eine ausführlichere Ansicht, kann er auf den Link klicken. Er wird dann zum Web-Frontend weitergeleitet, wo die detaillierte Produktseite für das betreffende Element angezeigt wird. Acme Inc. muss das GSA-XSLT-Stylesheet ändern, damit die Produkte und die zugehörigen Metadaten entsprechend angezeigt werden.

Acme Inc. passt auch das Frontend so an, dass erweiterte Suchfunktionen für Nutzer auf der Basis von Sammlungen und definierten Metadaten verfügbar sind. Nutzer könnten z. B. Metadaten per Auswahl in einem Drop-down-Menü oder über ein Optionsfeld als Suchbegriffe an die Suchanfrage anhängen und so das Produktkorpus entsprechend eingrenzen.

Alternative Vorgehensweisen

- Wenn alle Inhalte und Metadaten aus Datenbankabfragen abgeleitet werden können, können Sie den Datenbank-Connector für das Einlesen der Daten aller Produkte verwenden.
- Wenn die Frontend-Anwendung für die Produktanzeige in der Lage ist, alle zur Indexierung erforderlichen Informationen anzuzeigen, sollten Sie einen Webfeed für die Indexierung aller Produkte verwenden.
- Anstatt den Prozess zum Einlesen von Metadaten während der Indexierung zu definieren, können Sie alternativ auch mittels Wörterbüchern oder XML-Strukturen mit regulären Ausdrücken Entitätserkennungsregeln konfigurieren, um die Dokumente während der Indexierung automatisch mit Entitäten zu taggen.

Überblick über die Projektaufgaben

Die folgende Tabelle enthält die erforderlichen Projektaufgaben und Aktivitäten für die Indexierung von Inhalten über Feeds.

Aufgabe	Aktivität
Bereitstellungsarchitektur planen	• GSAs im Rack montieren und verkabeln • Server bereitstellen, auf dem die Feedanwendung gehostet wird
Crawling und Indexierung konfigurieren	• Verfolgungs-URLs für eingelesene Inhalte konfigurieren • Separate Sammlungen für einzelne Marken konfigurieren • Logik für die Erstellung der Feedinhalte entwerfen • Feedanwendung entwerfen, die XML-Datensätze schreibt und an die GSA weitergibt
Frontend konfigurieren	• Dynamische Navigation aktivieren • XSLT so ändern, dass Datensätze zusammen mit den gewünschten Metadaten angezeigt werden • Seite oder Seitenbereich für die erweiterte Suche erstellen, auf der Suchanfragen entsprechend den gewünschten Metadaten eingeschränkt werden können

Kapitel 5: Cookie-Umsetzung mit stiller Authentifizierung

Überblick über das Szenario

Im Anwendungsbeispiel für dieses Szenario möchte Acme Inc. sein SiteMinder-SSO-System und drei verschiedene Inhaltsquellen mithilfe der folgenden Mechanismen in die GSA integrieren:

- URL-ACLs
- Connectorbasierte Autorisierung
- Öffentliche Inhalte

Die bevorzugte Lösung aus der Sicht des Unternehmens ist die stille Authentifizierung der Nutzer, nachdem diese sich im Hauptportal des Unternehmens angemeldet haben.

Anforderungen

- Folgende Inhalte indexieren:
  - Livelink
  - Webbasierte Personenverzeichnis-Anwendung
  - Lotus Connections
- Ein Feld für die allgemeine Suche bereitstellen, das eine Suchergebnisseite mit den relevantesten Links in allen indexierten Inhalten zurückgibt
- Suchergebnisse für gesicherte Inhalte nur den Nutzern anzeigen, die zum Ansehen der Inhalte berechtigt sind
- Eine nahtlose stille Authentifizierung bereitstellen, bei der Nutzer suchen können, nachdem sie sich zunächst im Hauptportal angemeldet haben
- Inhalte in Lotus Connections werden auf der Basis von programmdefinierten Lotus Connections-Gruppen sowie LDAP-Gruppen abgesichert.

Annahmen

- Es ist ein SiteMinder-SSO vorhanden, mit dem Nutzer für die gesicherten Inhaltsquellen – Livelink und Connections – authentifiziert werden.
- Alle Inhaltsquellen verwenden dieselbe Identität.
- Der bestehende GSA-Connector für Livelink dient zur Einbindung der GSA in Livelink.
- Connections-Inhalte werden in die GSA eingelesen.
- ACLs können zusammen mit Connections-Inhalten eingelesen werden, um die Vorteile der URL-ACL-Funktion der GSA zu nutzen.

Wichtige Aspekte

- Bestätigen Sie die Annahme, dass alle Inhaltsquellen dieselbe Identität verwenden.
- Stellen Sie fest, ob die Inhaltsquellen native Gruppen nutzen oder Gruppen, die mit Active Directory/LDAP synchronisiert werden.
- Bestätigen Sie die Annahme, dass Lotus Connections die URL-ACL-Funktion der GSA verwenden kann, um ACL-Informationen zusammen mit dem Inhaltsfeed einzulesen.

Empfohlene Vorgehensweise

Die von Google empfohlene Vorgehensweise für die Implementierung einer Cookie-Umsetzung, die eine stille Authentifizierung ermöglicht, umfasst folgende Bereiche:

- Überblick über die Architektur
- Authentifizierung
- Autorisierung

Überblick über die Architektur

Die folgende Tabelle enthält die Inhaltsquellen, die Acme Inc. in die GSA integrieren möchte.

Inhaltsquelle	Integrationsmethode
Lotus Connections	• Feed, der Connections-Inhalte und ACLs für jedes einzelne Dokument enthält • Programmeigene Connections-Gruppen sowie LDAP-Gruppen werden zusammen mit dem Inhalt eingelesen. • Inhalte werden mithilfe der SeedList-Funktion von Lotus Connections synchron gehalten.
Open Text Livelink	• Inhalte werden durchsucht und über den Livelink-Connector in die GSA eingelesen. • Der Connector hält die Inhalte synchron. • ACLs werden nicht zusammen mit den Inhalten eingelesen. • Die connectorbasierte Autorisierung wird verwendet. Dokumente werden im Batchverfahren autorisiert.
Webbasiertes Personenverzeichnis	• Inhalte sind im Web verfügbar und werden direkt von der GSA gecrawlt.

Authentifizierung

Gemäß den Autorisierungsmechanismen von Acme Inc. benötigen beide gesicherten Inhaltsquellen eine bestätigte Identität mit einem Nutzernamen, um Inhalte für Nutzer während des Servings zu autorisieren. Der Livelink-Connector benötigt einen bestätigten Nutzernamen, um die connectorbasierte Autorisierung auszuführen. Ein gültiger Nutzername muss zusammen mit den zugehörigen Gruppen für die GSA bereitgestellt werden, damit Inhalte im Index mit ACLs autorisiert werden können.

Da der Zugriff auf beide Inhaltsquellen über das SiteMinder-SSO von Acme Inc. erfolgt, wird für die Nutzer während ihrer Sitzung ein Cookie erstellt, mit dessen Hilfe sie nach der Anmeldung im Hauptportal auf die Quellen zugreifen und eine Suche durchführen können. Es wird eine formularbasierte Authentifizierungsregel der universellen Anmeldung eingerichtet, um innerhalb des Authentifizierungsprozesses eine durch SiteMinder geschützte Beispiel-URL abzurufen. Wenn ein Nutzer bereits im Portal angemeldet und ein SiteMinder-Cookie vorhanden ist, wird der Nutzer autorisiert und muss keine Anmeldedaten eingeben.

Autorisierung

Für die Autorisierungsmechanismen beider Inhaltsquellen sind Anmeldedaten erforderlich:

- Livelink erfordert eine bestätigte Identität mit einem Nutzernamen.
- Connections erfordert einen Nutzernamen und Gruppen.

Daher muss die Beispiel-URL-Seite, die durch das SiteMinder-SSO geschützt ist, folgende Informationen über denjenigen Nutzer zurückgeben, dem der SSO-Cookie für die Authentifizierung zugeordnet ist:

- Einen Nutzernamen
- Eine Liste der mit dem Nutzer verknüpften Gruppen

Dieser Vorgang wird auch als "Cookie-Cracking" bezeichnet.

Hierfür wird eine JSP- oder ASP.NET-Prozedur erstellt, die nach Prüfung der Authentizität des SSO-Cookies eine HTTP-Antwort an die GSA zurückgibt. Die Antwort hat den OK-Statuscode "200" und enthält einen bestätigten Nutzernamen und eine Liste der zugehörigen Gruppen im Header ("X-Username" und "X-Groups"). Beachten Sie, dass programmeigene Connections-Gruppen zusammen mit den Nutzer-LDAP-Gruppen zurückgegeben werden müssen, damit Connections-ACLs im Index unterstützt werden.

Im Folgenden wird ein vollständiger und erfolgreicher Authentifizierungs- und Autorisierungsablauf beschrieben:

1. Der Nutzer meldet sich im Intranetportal des Unternehmens über eine SiteMinder-Anmeldeseite an und ein SiteMinder-Cookie wird in seiner Sitzung erstellt.

2. Der Nutzer führt eine Suche auf der GSA durch, dabei werden Cookies mit dem entsprechenden Geltungsbereich weitergegeben.

3. Die GSA ruft eine Beispiel-URL ab. Dabei werden die auf diesen Abruf zugeschnittenen Cookies weitergegeben. Das Cookie dient dazu, die Authentifizierung der durch SiteMinder geschützten Seite zu bestätigen.

4. Die Seite gibt den Code "200" zusammen mit dem bestätigten Nutzernamen und den Gruppen zurück, die mit dem Nutzer verknüpft sind, dessen Sitzungscookie an die Seite weitergegeben wurde.

5. Die GSA betrachtet den Abruf als erfolgreich und verknüpft den Nutzernamen und die Gruppen mit der Anmeldedatengruppe für bestätigte Identitäten.

6. Der Nutzername und die Gruppen werden verwendet, um weitere Inhalte auf der GSA zu autorisieren. Der Nutzername wird an die connectorbasierte Autorisierung für Livelink übergeben. Nutzername und Gruppen werden übergeben, um damit Autorisierungsprüfungen für ACLs vorzunehmen, die mit Inhalten im Index verknüpft sind.

Alternative Vorgehensweise

Führen Sie für alle Inhalte Autorisierungsprüfungen per HEAD-Anfrage aus. Hierfür sind keine Gruppen und kein Nutzername erforderlich, die bzw. der mit der bestätigten Identität verknüpft sind.

Überblick über die Projektaufgaben

In der folgenden Tabelle sind die Projektaufgaben und Aktivitäten für die Implementierung des Cookie-Crackings aufgeführt, das eine stille Authentifizierung ermöglicht.

Aufgabe	Aktivität
Bereitstellungsarchitektur planen	• Anwendung entwickeln, die mithilfe eines Cookies den Nutzernamen und die zugehörigen Gruppen eines Nutzers zurückgibt • Diese Anwendung auf einem Web-/Anwendungsserver bereitstellen, der durch SiteMinder geschützt ist; für Apache kann ein SiteMinder-Plug-in für die Integration in das SSO verwendet werden
Crawling und Indexierung konfigurieren	• Connectors für die Indexierung der Inhalte konfigurieren • Crawler für die Indexierung der öffentlichen Inhalte konfigurieren
Cookiebasierte Authentifizierung mit einer Beispiel-URL unter "Authentifizierungsmechanismen der universellen Anmeldung" konfigurieren	• Wenn dies konfiguriert ist, ruft die GSA die Beispiel-URL auf, wodurch die Seite geöffnet wird. Die Cookie-Cracker-Seite gibt dann den Nutzernamen und die Gruppen zurück, die mit der bestätigten Identität verknüpft sind, die die Suche ausführt.

Langfristige Optimierungen

- Prüfen Sie die Architektur, wenn neue Inhaltsquellen hinzukommen, um zu ermitteln, ob die bestehende Architektur geändert werden muss, um die neuen Inhalte zu integrieren.

- Falls später ACLs für Livelink verfügbar werden, passen Sie den Cookie-Cracker so an, dass er auch Livelink-Gruppen zurückgibt. Wenn es Konflikte bei Gruppennamen gibt, können Sie einen Cookie-Cracker in einer separaten Anmeldedatengruppe (Namespace) verwenden.

Kapitel 6: Stille Authentifizierung – Integration in NTLM und SAML Bridge

Überblick über das Szenario

Acme Inc. verwendet NTLM mit integrierter Windows-Authentifizierung (IWA) mit einem Active Directory-Backend. Im Anwendungsbeispiel für dieses Szenario bevorzugt das Unternehmen eine nahtlose, stille Authentifizierung für seine Nutzer, bei der diese nach ihrer Anmeldung in der Windows-Domain Suchen ausführen können und mit dem Internet Explorer im Internet navigieren.

Anforderungen

- NTLM-geschützte Inhalte indexieren und sicher ausgeben
- Ein Feld für die allgemeine Suche bereitstellen, das eine Suchergebnisseite mit den relevantesten Links in den indexierten Inhalten zurückgibt
- Suchergebnisse für gesicherte Inhalte nur den Nutzern anzeigen, die zum Ansehen der Inhalte berechtigt sind
- Eine nahtlose stille Authentifizierung bereitstellen, bei der Nutzer Suchen ausführen können, nachdem sie sich in der Windows-Domain angemeldet haben

Annahmen

- Der IIS-Server, auf dem die Inhalte gehostet werden, akzeptiert HEAD-Anfragen.
- Alle gecrawlten Inhalte befinden sich unter derselben Windows-Domain, in der sich Nutzer anmelden.

Wichtige Aspekte

- Bestätigen Sie die Annahme, dass alle Inhaltsquellen dieselbe Windows-Domain verwenden, in der sich Nutzer anmelden.
- Stellen Sie sicher, dass alle Server in der Bereitstellungsarchitektur mit demselben Zeitserver synchronisiert werden.
- Stellen Sie sicher, dass ein Zertifikat in den IIS verfügbar ist, das verwendet werden kann, um SAML-Post-Binding-Anfragen zu signieren.
- Um die Kommunikation mit der SAML Bridge über HTTPS zu ermöglichen, müssen Sie die GSA mit dem Stammzertifikat der SAML Bridge konfigurieren.

Empfohlene Vorgehensweise

Die von Google empfohlene Vorgehensweise für die Implementierung der stillen Authentifizierung durch die Integration mit NTLM und SAML Bridge umfasst folgende Bereiche:

- Authentifizierung
- Autorisierung
- Ablauf der Authentifizierung und Autorisierung mit der SAML Bridge

Authentifizierung

Acme Inc. verwendet die SAML Bridge, um Nutzer mit Active Directory zu authentifizieren und so die Vorteile der stillen Authentifizierung zu nutzen, die über die integrierte Windows-Authentifizierung (IWA) möglich ist. Hierfür muss der Domaincontroller, auf dem Active Directory ausgeführt wird, folgende Anforderungen erfüllen:

- Die Windows 2003-Kerberos-Erweiterung muss verfügbar sein, da Kerberos für die Authentifizierung zwischen SAML Bridge und Inhaltsserver verwendet wird.
- Die Funktionsebene der Domain muss auf Windows Server 2003 festgelegt sein.
- Active Directory muss so konfiguriert werden, dass die SAML Bridge berechtigt ist, delegierte Anmeldedaten des Nutzers zu verwenden, um Inhalte auf dem Inhaltsserver aufzurufen.

Um die SAML Bridge zur Verwendung durch die GSA beim Ausführen der Authentifizierung zu konfigurieren, verwenden Sie die Seite **Suche > Sichere Suche > Authentifizierungsmechanismen der universellen Anmeldung** (vor Version 7.2: **Serving > Authentifizierungsmechanismen der universellen Anmeldung > SAML**) in der Admin-Konsole.

Da SAML-Post-Binding ab SAML Bridge Version 2.8 als Option verfügbar ist, wird empfohlen, Post-Binding für die Authentifizierung mit der SAML Bridge zu verwenden. Eine Schritt-für-Schritt-Anleitung zur Konfiguration der SAML Bridge für das Post-Binding finden Sie in folgendem Wiki: http://code.google.com/p/google-saml-bridge-for-windows/wiki/SAMLBridge28features

Autorisierung

Da eine Autorisierung mit NTLM erforderlich ist, wird die SAML Bridge auch für die Autorisierung verwendet. In diesem Fall muss die SAML Bridge als Autorisierungsanbieter auf der Seite **Suche > Sichere Suche > Zugriffssteuerung** (vor Version 7.2: **Serving > Zugriffssteuerung**) in der Admin-Konsole konfiguriert werden. Die GSA delegiert dann Autorisierungsprüfungen für einzelne Dokumente an die SAML Bridge.
Die SAML Bridge antwortet entsprechend mit PERMIT oder DENY.

Ablauf der Authentifizierung und Autorisierung mit der SAML Bridge

1. Ein Nutzer stellt eine Suchanfrage nach gesicherten Inhalten.
2. Die Authentifizierungs-SPI der GSA wird verwendet, um die Delegierung an die SAML Bridge zwecks Authentifizierung vorzunehmen. NTLM, das im Browser des Nutzers konfiguriert ist, dient zur Authentifizierung des Nutzers.
3. Nachdem der Nutzer authentifiziert wurde, ermittelt die GSA die relevantesten Ergebnisse für den Nutzer. Wenn die Ergebnisse gesicherte Dokumente enthalten, verwendet die GSA die Autorisierungs-SPI, um die Autorisierungsprüfungen für diese Dokumente an die SAML Bridge zu delegieren.
4. Die SAML Bridge erhält ein Kerberos-Ticket, das auf diesen Nutzer ausgestellt ist, und tritt unter der Identität des Nutzers auf dem Inhaltsserver auf.
5. Die SAML Bridge sendet eine PERMIT- oder DENY-Nachricht zurück an die GSA und die GSA zeigt die Ergebnisse, auf die ein Nutzer zugreifen darf, auf einer Suchergebnisseite an.

Alternative Vorgehensweise

Wechseln Sie zu Kerberos als Authentifizierungsmechanismus für Inhaltsserver. Beim Wechsel zu Kerberos besteht die Möglichkeit, die Bereitstellung der SAML Bridge beim Konfigurieren der stillen Authentifizierung zu umgehen.

Überblick über die Projektaufgaben

In der folgenden Tabelle sind die Projektaufgaben und Aktivitäten für die Implementierung der stillen Authentifizierung und die Integration mit NTLM und SAML Bridge aufgeführt.

Aufgabe	Aktivität
Bereitstellungsarchitektur planen	• Active Directory-Konfiguration für die Installation der SAML Bridge vorbereiten • SAML Bridge auf Domaincontroller bereitstellen und Konfiguration nach Bedarf anpassen • Zertifikate für POST-Binding und Kommunikation über HTTPS konfigurieren
SAML Bridge auf der GSA konfigurieren	• Authentifizierungs-SPI so konfigurieren, dass die SAML Bridge verwendet wird • Autorisierungs-SPI so konfigurieren, dass die SAML Bridge verwendet wird

Langfristige Optimierungen

Prüfen Sie die Architektur, wenn neue Inhaltsquellen hinzukommen, um zu ermitteln, ob die bestehende Architektur geändert werden muss, um die neuen Inhalte zu integrieren.

Kapitel 7: Reverseproxy für Umfeldsicherheit und aus anderen Gründen implementieren

Überblick über das Szenario

Bei Acme Inc. gibt es streng vertrauliche Forschungs- und Planungsdokumente. In diesem Szenario möchte das Unternehmen den Zugriff auf diese Dokumente beschränken, indem alle Suchvorgänge zwingend über einen Proxy erfolgen. Der Proxy erzwingt die Authentifizierung über das System von Acme für die Einmalanmeldung (Single Sign-On – SSO), bevor er den Zugriff auf die GSA erlaubt, und beschränkt auch die Suchanfragen, die an die GSA gesendet werden können.

Anforderungen

- SSO-Anmeldung erzwingen, bevor auf die GSA zugegriffen werden kann
- Suchanfragen auf der GSA durch die Einschränkung von URL-Anfrageparametern auf eine bestimmte Sammlung beschränken

Annahmen

- In diesem Beispiel wird davon ausgegangen, dass ein Apache-Webserver verwendet werden soll. Beachten Sie, dass auch andere Webserver als Reverseproxys für die GSA verwendet werden können.
- Ein Apache-Server ist verfügbar.
- Ein Apache-Plug-in für das SSO von Acme ist verfügbar.

Wichtige Aspekte

- Wenn die GSA für sichere Suchen verwendet wird:
    - HTTPS-Datenverkehr muss über den Proxy erfolgen.
    - Aufrufe des Sicherheitsmanagers auf der GSA müssen ebenfalls über den Proxy erfolgen.
- Wird über HTTPS auf die GSA zugegriffen, muss auch der SSL-Datenverkehr über den Proxy erfolgen.
- Die GSA ist durch eine Firewall geschützt und der Zugriff ist auf den Proxyserver beschränkt.

Empfohlene Vorgehensweise

Die von Google empfohlene Vorgehensweise für die Implementierung eines Reverseproxys für die Umfeldsicherheit umfasst folgende Bereiche:

- Apache in ein SSO-System integrieren
- Anfragen an die GSA über den Proxy leiten
- Kompletten Datenverkehr über den Reverseproxy leiten

Apache in ein SSO-System integrieren

Um die Apache-Instanz mit dem SSO-System zu schützen, installiert Acme Inc. das passende Apache-SSO-Plug-in für das verwendete SSO-System. Je nachdem, ob das Plug-in eine Konfigurationsschnittstelle hat, erfolgt die Konfiguration über verschiedene Assistenten, in denen Schutzoptionen für die Anwendung ausgewählt werden können, oder indem in Apache entsprechende Ressourcenfilter für den Datenverkehr festgelegt werden.

Wenn das SSO-Plug-in konfiguriert ist, wird ein Nutzer jedes Mal, wenn der Apache-Host mit dem entsprechenden Cookie-Domainbereich aufgerufen wird, über das SSO authentifiziert. Liegt noch kein Cookie für die Sitzung des Nutzers vor, müsste dieser auf die SSO-Anmeldeseite weitergeleitet werden, damit ein Cookie erstellt wird. Anschließend darf der Nutzer auf die GSA zugreifen.

Anfragen an die GSA über den Proxy leiten

Der Mechanismus, der hierfür normalerweise verwendet wird, ist ein "VirtualHost"-Block. Die Einrichtung kann aber auch über die Hauptserverkonfiguration erfolgen. So konfigurieren Sie einen virtuellen Host für die Proxy-Verarbeitung des Datenverkehrs:

```
<VirtualHost *:80>
  ProxyRequests Off
    <Proxy *>
      Order Deny,Allow
      Deny from all
      Allow from [gsa_ip]
  </Proxy>

  ProxyPass / http://gsa32.ihrebeispielurl.de
  ProxyPassReverse / http://gsa32.ihrebeispielurl.de
</VirtualHost>
```

In Konfigurationen, bei denen die sichere Suche aktiviert ist, wird das Apache-Plug-in "mod_ssl" für die Proxyverarbeitung des HTTPS-Datenverkehrs benötigt. Darüber hinaus muss für den Apache-Server ein Zertifikat ausgestellt werden. Dieses Zertifikat muss auf der GSA installiert werden, damit die über den Proxy geleiteten Anfragen als signiert erkannt werden.

Kompletten Datenverkehr über den Reverseproxy leiten

Nachdem der Reverseproxy implementiert wurde, konfiguriert Acme Inc. eine Firewallregel, um nur Datenverkehr auf der GSA zuzulassen, der vom Apache-Host kommt. Dadurch werden alle Anfragen für den Zugriff auf die GSA zwingend über den Apache-Reverseproxy geleitet.

Alternative Vorgehensweise

Alternativ kann ein anderer Webserver für die Implementierung des Reverseproxys verwendet werden. Beispielsweise könnten Internetinformationsdienste (Internet Information Services – IIS) verwendet werden, um den Datenverkehr zu filtern.

Ab GSA-Version 6.14 kann die Funktion "Umfeldsicherheit" der GSA verwendet werden, um einen solchen Mechanismus einzurichten. Hierfür müsste ein Sicherheitsmechanismus auf der GSA konfiguriert werden, der nur die Authentifizierung übernimmt. Wenn dies entsprechend eingerichtet wurde, werden Nutzern erst dann öffentliche Ergebnisse angezeigt, nachdem die Nutzer auf der GSA authentifiziert wurden.

Überblick über die Projektaufgaben

In der folgenden Tabelle sind die Projektaufgaben und Aktivitäten zur Implementierung eines Reverseproxys für die Umfeldsicherheit aufgeführt.

Aufgabe	Aktivität
Apache-Integration in das SSO planen	• Apache-URL schützen • Apache so konfigurieren, dass das SSO-Plug-in verwendet wird, und entsprechende Ressourcenfilter einrichten, um den Datenverkehr für die SSO-geschützten Ressourcen zu filtern
Konfigurieren, dass Anfragen an die GSA über den Apache-Proxyserver geleitet werden	• Virtuellen Host für die Verarbeitung des Datenverkehrs an die GSA über den Proxy und die GSA entsprechend genauso konfigurieren • Wenn eine sichere Suche oder der Zugriff auf die GSA über HTTPS erforderlich ist, wird das Plug-in "mod_ssl" für die Proxyverarbeitung des HTTPS-Datenverkehrs benötigt.
Firewall so konfigurieren, dass ausschließlich über den Apache-Host auf die GSA zugegriffen werden kann	• Firewallregel konfigurieren, um die Umfeldsicherheit für die GSA so einzurichten, dass nur über den Apache-Proxy darauf zugegriffen werden kann

Langfristige Optimierungen

- Der Reverseproxy kann auch für andere Zwecke eingesetzt werden: saubere URLs, Firewall-Tunneling, Caching für die Leistungsverbesserung.
- Reaktionszeiten und Serving-Kapazitäten der GSA lassen sich erheblich verbessern, wenn Apache als Cache verwendet wird. So könnte beispielsweise eine Memcache-Konfiguration zum Abschnitt "VirtualHost" hinzugefügt werden:

  ```
  CacheEnable mem /
  MCacheSize 4096
  MCacheMaxObjectCount 1000
  MCacheMinObjectSize 1
  MCacheMaxObjectSize 4096
  ```

  Hierdurch werden die letzten 1.000 GSA-Antworten mit einer Größe von maximal 4 KB im Cache gespeichert.

Kapitel 8: Relevanztest

Überblick über das Szenario

Acme Inc. hat die GSA eingerichtet und folgende Inhaltsquellen integriert:

- Livelink-Inhalte
- Gecrawlte Intranetwebsite
- Personenverzeichnis-Anwendung

Im Anwendungsbeispiel für dieses Szenario möchte das Unternehmen Relevanztests durchführen, um sicherzustellen, dass die Nutzer mit den von der GSA ausgegebenen Suchergebnissen zufrieden sind, bevor die Suchlösung in der Produktionsumgebung verfügbar gemacht wird.

Anforderungen

Sicherstellen, dass die Suchergebnisse, die Nutzern angezeigt werden, für die eingegebenen Suchbegriffe relevant sind

Annahmen

- Inhalte wurden bereits in die GSA integriert und sind im Index verfügbar.
- Die Testplaner sind mit dem geschäftlichen Kontext der Inhalte im GSA-Index vertraut.

Wichtige Aspekte

- Relevanz ist schwer an konkreten Zahlen festzumachen und möglicherweise subjektiv.
- Es hat sich gezeigt, dass die vorkonfigurierten Relevanz-Algorithmen der GSA ohne Anpassungen und Optimierungen hoch relevante Ergebnisse ausgeben.

Empfohlene Vorgehensweise

Die von Google empfohlene Vorgehensweise für Relevanztests umfasst folgende Bereiche:

- Testfall vorbereiten
- Testfall ausführen
- Funktionen, die bei der Relevanzoptimierung berücksichtigt werden sollten

Testfall vorbereiten

- Ermitteln Sie verschiedene Nutzergruppen im Unternehmen, die die Suche verwenden werden.
- Ermitteln Sie anhand der Inhalte im GSA-Index einen Teil des geschäftlichen Kontexts der Art von Suchanfragen, die verschiedene Nutzer ausführen und welche Dokumente sie dabei in der Ergebnisanzeige erwarten würden.
- Stellen Sie eine Liste mit vorgegebenen Suchanfragen zusammen, die Nutzer ausführen sollen, um die Relevanz von Ergebnissen zu kommentieren. Bitten Sie die Nutzer in den Testrunden, zusätzlich zu diesen vorgegebenen Suchanfragen etwa drei eigene Suchanfragen auszuführen, um auch Kontexte zu erfassen, die bei der Vorbereitung des Testfalls nicht berücksichtigt wurden.
- Ermitteln Sie eine Reihe von Dokumenten, die Ihrer Meinung nach für eine bestimmte Suchanfrage für einen bestimmten Nutzer am relevantesten sind. Diese werden für das Scoring (die Bewertung durch die Nutzer) verwendet.
- Entwickeln Sie eine Skala, die ein Tester verwenden kann, um die Relevanz der ausgegebenen Ergebnisse zu beurteilen. Erläutern Sie die Skala und deren Funktionsweise den Nutzern, die den Test ausführen. Beispielsweise könnte eine fünfstufige Skala mit folgenden Bedeutungen verwendet werden:
    - 1: Relevanz ist hervorragend. Auf der ersten Seite werden alle besonders relevanten Ergebnisse angezeigt. Das für diese Suchanfrage im Vorfeld ermittelte Dokument wird auf der ersten Suchergebnisseite angezeigt.
    - 5: Relevanz ist sehr gering. Die Ergebnisse, die ich erwarte, werden nicht auf den ersten Suchergebnisseiten angezeigt. Das für diese Suchanfrage als relevant ermittelte Dokument wird nicht vor dem 60. Suchergebniseintrag angezeigt. Es gibt eine Inhaltsquelle, aus der gehäuft Ergebnisse angezeigt werden, wodurch alle anderen Quellen deutlich unterrepräsentiert sind.

Testfall ausführen

- Bevor Sie die Relevanz optimieren, sollten Sie eine Relevanz-Benchmark entwickeln. Bitten Sie hierfür mehrere Gruppen von Beta-Nutzern aus verschiedenen Abteilungen/Unternehmensbereichen, die die Suchlösung auch später in der Produktionsumgebung verwenden werden, die vorgegebenen Suchanfragen auszuführen.
- Bitten Sie die Nutzer, die einzelnen Ergebnisse für die ausgeführten Suchanfragen in einer Tabelle anhand der vorgegebenen Skala zu bewerten. Bitten Sie die Nutzer außerdem, allgemeine Kommentare für jede Suche anzugeben.
- Nachdem Sie mit der standardmäßigen Relevanzkonfiguration auf der GSA (keine Synonyme für die Suchanfragenerweiterung, keine Richtlinien zur Ergebnisgewichtung usw.) eine Benchmark ermittelt haben, optimieren Sie nun systematisch die Relevanzkonfiguration anhand von Nutzerfeedback und Kommentaren. Führen Sie den Test nach jedem Änderungsdurchgang erneut mit den Nutzern durch, um herauszufinden, wie sich die vorherige Änderung auf die Relevanzwahrnehmung der Nutzer ausgewirkt hat.

Funktionen, die bei der Relevanzoptimierung berücksichtigt werden sollten

In der folgenden Tabelle sind GSA-Funktionen aufgeführt, die bei der Relevanzoptimierung eine Rolle spielen.

Funktion	Kommentar
Quellengewichtung	Mithilfe der Musterübereinstimmung können Sie bestimmte Quellen höher als andere einstufen.
Quellengewichtung nach Datum, Metadatengewichtung	Hiermit können Sie Dokumente gewichten, an die bestimmte Metadaten angehängt sind.
KeyMatches	Verwenden Sie KeyMatches, um Dokumente für bestimmte Suchanfragen vorrangig anzeigen zu lassen.
Suchanfragenerweiterung	Verwenden Sie eine Richtlinie für die Suchanfragenerweiterung, um Suchanfragenbegriffe auch auf andere Begriffe auszuweiten (Synonyme).
Intelligente Auswertung	Wenn erweiterte Suchberichte aktiviert sind, verwendet die GSA die intelligente Auswertung, um Clickstreamdaten zu analysieren und bestimmte Suchergebnisse mit der Zeit aufzuwerten. Ein Beispiel: Wenn bei einer bestimmten Suchanfrage Nutzer anstatt auf das erste Ergebnis beständig auf das zweite Ergebnis auf der Seite klicken, wird das zweite Ergebnis irgendwann entsprechend aufgewertet und ganz oben auf der Seite angezeigt.
Host-Crowding/Filter	Die GSA filtert alle Kombinationen von Folgendem heraus: ● Ergebnisse aus demselben Pfad ● Ergebnisse mit identischen Titeln und Snippets
Ranking-Framework	Legen Sie eine Gewichtung nach URL fest. Beachten Sie, dass diese Lösung sehr komplex in der Verwaltung ist und nur als letztes Mittel zum Zuge kommen sollte.
Stoppwörter (ab GSA 6.10)	Verwenden Sie Stoppwörter, um zu vermeiden, dass bestimmte Begriffe in der Suchanfrage für die Ausführung der Suche verwendet werden. Verwenden Sie diese Funktion mit Bedacht, da es weitreichende Folgen haben kann, wenn Sie sie als Lösung für ein bestimmtes Problem verwenden.
Sammlungen	Inhalte können in verschiedene Sammlungen unterteilt werden, um den Dokumentkorpus zu begrenzen, der für eine Suchanfrage verfügbar ist.

Metadaten und/oder Entitäten in der dynamischen Navigation sichtbar machen, um den Nutzern eine detailliertere Suche zu ermöglichen	Anstatt die Relevanzeinstellungen der GSA anzupassen, können Sie Nutzern auch eine detailliertere Suche ermöglichen, indem Sie in der dynamischen Navigation zusätzliche Kategorien für Metadatenquellen oder Entitäten hinzufügen, die in der Entitätserkennung definiert wurden. Die dynamische Navigation ist zwar per se keine Relevanzoptimierung, kann aber die Suche für die Nutzer verbessern, da sie ihnen ermöglicht, die Ergebnisse detaillierter aufzuschlüsseln und so die gesuchten Informationen leichter zu finden.

Alternative Vorgehensweise

Sie können die Gewichtung bereits während der Indexierung anpassen. Verwenden Sie einen Inhaltsfeed und legen Sie den PageRank einzelner Dokumente fest. Mit dem PageRank-Attribut können Sie den PageRank eines Dokuments manuell festlegen. Er kann bis zu einem Wert von 99 angegeben werden, was einen sehr hohen PageRank bedeutet Der bisherige Standardwert für alle Inhaltsfeed-Dokumente betrug 96.

Überblick über die Projektaufgaben

Die folgende Tabelle zeigt die Projektaufgaben und Aktivitäten für Relevanztests.

Aufgabe	Aktivität
Test planen	• Nutzergruppen für den Test und das Relevanzfeedback ermitteln • Liste mit Suchanfragen zusammenstellen, die jeder Nutzer im Rahmen des Tests ausführen soll • Eine Reihe von relevanten Dokumenten für jede Suche und jeden Nutzer ermitteln • Relevanzskala erstellen, anhand derer die Qualität der Suchergebnisse bewertet werden soll
Test ausführen	• Relevanz-Benchmark entwickeln; dazu Nutzer die Tests ausführen lassen, bevor Änderungen an der GSA vorgenommen werden. • Nutzer anweisen, die Tests auszuführen, die Ergebnisse zu bewerten und Feedback zu geben.
Wiederholen und erneut testen.	• Die GSA-Gewichtung anhand des Feedbacks der Nutzer anpassen und Tests mit den Nutzern wiederholen. • Verfeinern Sie Ihre Einstellungen weiter und wiederholen Sie die Tests, bis Sie mit den Ergebnissen zufrieden sind.

Langfristige Optimierungen

Entwickeln Sie einen Prozess und einen Mechanismus, um Nutzerfeedback einzuholen und die ständige Verfeinerung der Suchrelevanz in der Produktionsumgebung weiterzuführen.

Zusammenfassung

Jede GSA-Bereitstellung birgt je nach Ihrer IT-Landschaft unterschiedliche Herausforderungen. Die hier beschriebenen Annahmen, Überlegungen zu wichtigen Aspekten, Vorgehensweisen, Projektaufgaben und Optimierungen sind Beispiele und nicht immer eins zu eins umsetzbar. Ihre eigene Umgebung und Zeitplanung könnte komplexer sein. Berücksichtigen Sie bei der Planung einer GSA-Bereitstellung spezielle unternehmerische und technische Anforderungen. Berücksichtigen Sie bei Ihrer Planung stets Eventualitäten.