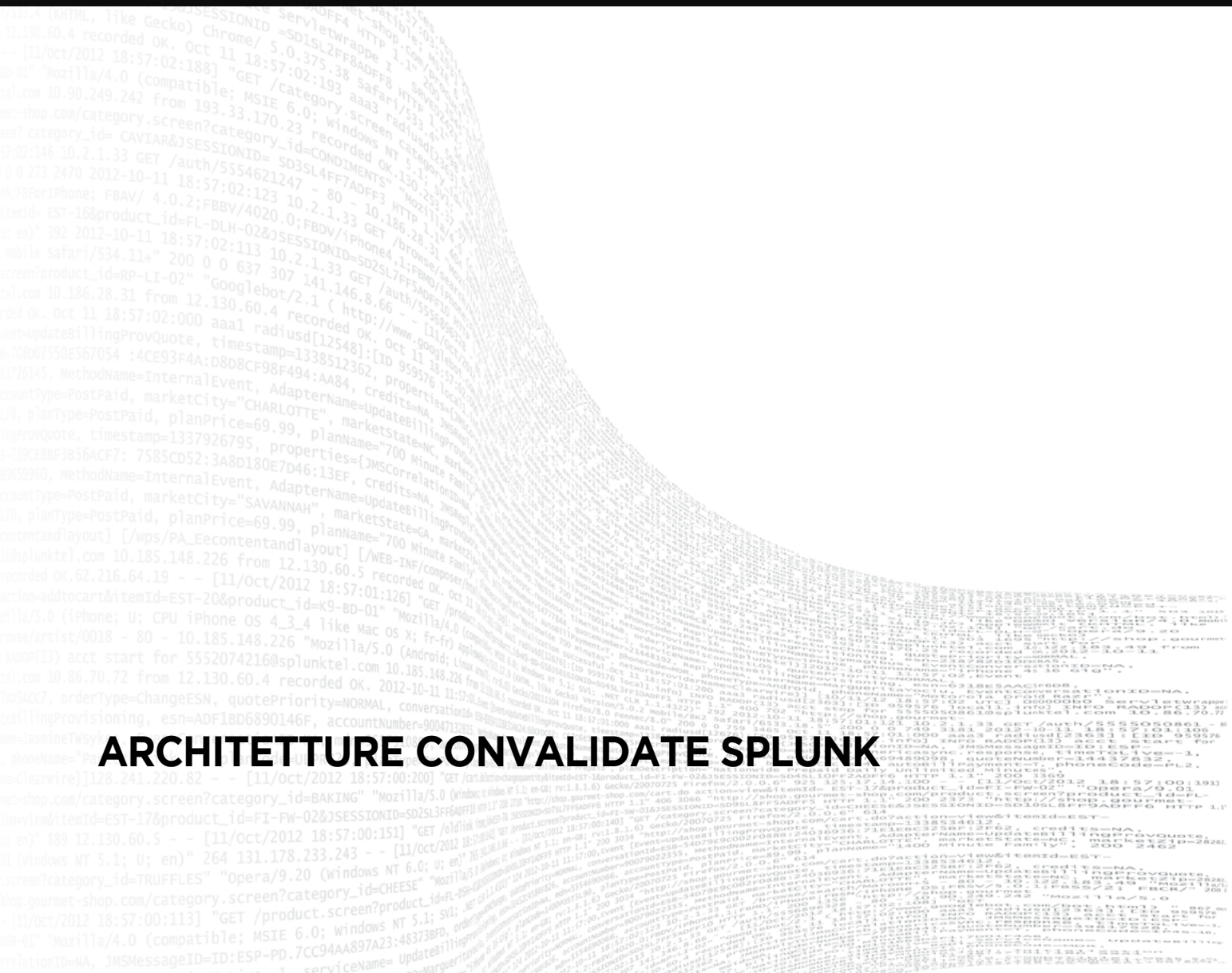
splunk>

WHITE PAPER

ARCHITETTURE CONVALIDATE SPLUNK



Sommario

Introduzione

Le Architetture convalidate Splunk (Splunk Validated Architectures – SVA) sono architetture di riferimento consolidate per un deployment stabile, efficiente e ripetibile di Splunk. Molti attuali clienti di Splunk hanno sperimentato una rapida adozione ed espansione, da cui sono emerse alcune sfide inerenti all'aumento delle dimensioni. Al contempo, i nuovi clienti di Splunk sono sempre più alla ricerca di linee guida e architetture certificate che assicurino loro la solidità delle basi del loro deployment iniziale. Le SVA sono state sviluppate per aiutare i clienti ad affrontare queste crescenti esigenze.

Le SVA aiutano sia i clienti già acquisiti sia i nuovi clienti di Splunk a realizzare un ambiente nel quale la manutenzione e la risoluzione degli errori sono semplificate. Le SVA sono pensate per aiutare l'utilizzatore a ottenere i migliori risultati possibili riducendo al contempo al minimo i costi di proprietà (cost of ownership). Inoltre l'intera struttura di Splunk sarà fondata su un'architettura che consente di scalare il deployment con l'evolvere delle esigenze nel tempo.

Le SVA offrono opzioni di topologia che tengono conto di un ampio spettro di necessità organizzative, in modo che sia possibile comprendere e trovare facilmente una topologia adatta alle proprie esigenze. Il processo di selezione delle Architetture convalidate Splunk aiuta a trovare l'abbinamento tra gli specifici requisiti e la topologia che meglio si adatta alle esigenze dell'organizzazione. Per coloro che si approcciano a Splunk per la prima volta, per il deployment iniziale consigliamo di implementare un'Architettura convalidata. Per i clienti che hanno già esperienza, consigliamo di esplorare la possibilità di allinearsi alla topologia di un'Architettura convalidata. A meno che non si abbiano requisiti unici che richiedono la realizzazione un'architettura personalizzata, molto probabilmente un'Architettura convalidata soddisferà i requisiti mantenendo un buon rapporto costi-benefici.

Questo whitepaper presenta una panoramica generale delle SVA. In questo whitepaper sono presentate le risorse necessarie per portare a termine il processo di selezione della SVA, tra cui il questionario dei requisiti, i diagrammi della topologia del deployment, i principi di progettazione, e le linee guida generali.

Per assistenza nell'implementazione di un'Architettura convalidata Splunk, contattare i servizi professionali di Splunk (https://www.splunk.com/en_us/support-and-services/splunk-services.html).

Struttura del documento

Le SVA si ripartiscono in tre grandi aree tematiche:

1. Topologie di indicizzazione e ricerca
2. Componenti dell'architettura per la raccolta dei dati
3. Principi e best practice di progettazione

L'indicizzazione e la ricerca riguardano i livelli dell'architettura che mettono a disposizione le capacità primarie di indicizzazione e ricerca di un deployment Splunk. La sezione relativa ai componenti per la raccolta dei dati guida il lettore nella scelta del corretto meccanismo di raccolta dei dati in funzione dei propri requisiti.

I principi e le best practice di progettazione si applicano all'architettura nel suo insieme e aiutano a fare la scelta giusta al momento di determinare i dettagli del deployment.

Motivi per usare le Architetture convalidate Splunk

L'implementazione di un'Architettura convalidata consente di progettare e distribuire Splunk con maggiore sicurezza. Le SVA aiutano a risolvere alcune delle difficoltà più comuni che devono affrontare le organizzazioni, tra cui:

Prestazioni

- Le organizzazioni desiderano vedere un miglioramento delle prestazioni e nella stabilità.

Complessità

- Talvolta le organizzazioni rimangono impantanate nelle difficoltà di deployment personalizzati, specialmente quando hanno vissuto una crescita troppo rapida od organica. In questi casi, è possibile nell'ambiente sia stata introdotta un'inutile complessità. Questa complessità può diventare una barriera reale nel momento in cui si tenta di scalare le dimensioni.

Efficienza

- Per ricavare i massimi benefici dal deployment Splunk, le organizzazioni devono migliorare l'efficienza dell'operatività e accelerare il time to value.

Costo

- Le organizzazioni sono alla ricerca di un modo per ridurre il costo totale di proprietà (TCO) soddisfacendo al contempo tutte le proprie esigenze.

Agilità

- Le organizzazioni hanno necessità di adattarsi parallelamente alla crescita e all'aumento delle dimensioni.

Manutenzione

- Spesso si rende necessaria l'ottimizzazione dell'ambiente per ridurre l'onere di manutenzione.

Scalabilità

- Le organizzazioni devono essere in grado di scalare in modo efficiente e trasparente.

Verifica

- Le parti interessate all'interno dell'organizzazione vogliono la certezza che il deployment Splunk sia realizzato sulla base delle best practice.

Colonne portanti delle Architetture convalidate Splunk

Le Architetture convalidate Splunk sono realizzate sulle seguenti colonne portanti. Per maggiori informazioni su questi aspetti si rimanda alla successiva Appendice "A".

DISPONIBILITÀ	PRESTAZIONI	SCALABILITÀ	SICUREZZA	GESTIBILITÀ
Il sistema è **sempre operativo** ed è in grado di ripristinarsi a seguito di guasti e interruzioni, programmati o meno.	Il sistema è in grado di **mantenere un livello di servizio ottimale** in diversi pattern di utilizzo.	Il sistema è progettato per scalate a tutti i livelli, consentendo di **gestire in modo efficace carichi di lavoro crescenti**.	Il sistema è **progettato per proteggere i dati, le configurazioni, e gli asset** continuando al contempo a generare valore.	Il sistema è **gestibile a livello centralizzato e a tutti i livelli.**

Questi pilastri sostengono direttamente il servizio **Gestione e supporto della piattaforma** nel modello di Splunk basato sui centri di eccellenza.

Cosa offrono le Architetture convalidate Splunk

Facciamo osservare che le SVA non comprendono le tecnologie di deployment né il dimensionamento del deployment. Il motivo alla base di questo è:

- nel contesto delle SVA, le tecnologie di deployment, come ad esempio i sistemi operativi e l'hardware dei server, sono considerate scelte di implementazione. Differenti clienti arriveranno a scelte differenti, non è quindi possibile generalizzare.
- Il dimensionamento del deployment esige una valutazione del volume di inserimento di dati, dei tipi di dati, dei volumi delle ricerche, nonché dei casi d'uso delle ricerche, che tendono a essere molto specifici per cliente e in generale non hanno alcuna rilevanza ai fini dell'architettura fondamentale stessa del deployment. Gli attuali strumenti di dimensionamento possono aiutare con questo processo una volta che è stata creata l'architettura del deployment. Splunk Storage Sizing (https://splunk-sizing.appspot.com/) è uno degli strumenti disponibili.

Le SVA offrono:	**Le SVA non offrono:**
✅ Opzioni di deployment in cluster e non in cluster. ✅ Diagrammi dell'architettura di riferimento. ✅ Linee guida per aiutare a scegliere la giusta architettura ✅ Raccomandazioni specifiche per livello. ✅ Best practice realizzare il deployment Splunk	❌ Scelte di implementazione (sistema operativo, server fisico, virtuale o cloud ecc.). ❌ Dimensionamento del deployment. ❌ Un'approvazione prescrittiva dell'architettura. Nota: le SVA offrono raccomandazioni e linee guida con l'ausilio delle quali prendere la decisione giusta per la propria organizzazione. ❌ Un suggerimento di topologia per ogni possibile scenario di deployment. In alcuni casi, la presenza di fattori unici può richiedere lo sviluppo di un'architettura personalizzata. Gli esperti di Splunk sono a disposizione per assistere con le eventuali soluzioni personalizzate necessarie. I clienti possono contattare il proprio Splunk Account Team. Chi non è ancora cliente, può contattarci qui (https://www.splunk.com/en_us/talk-to-sales.html).

Ruoli e responsabilità

Le Architetture convalidate Splunk sono molto importanti per le considerazioni che decisori e amministratori sono chiamati a fare. Gli architetti d'impresa, i consulenti, gli amministratori Splunk e i provider di servizi gestiti devono essere tutti coinvolti nel processo di selezione della SVA. Segue una descrizione di ciascuno di questi ruoli:

Ruolo	**Descrizione**
Architetti d'impresa	Si occupano di progettare il deployment Splunk per soddisfare le esigenze dell'impresa.
Consulenti	Si occupano di fornire servizi per l'architettura, la progettazione e l'implementazione di Splunk.
Tecnici Splunk	Si occupano di gestire il ciclo di vita di Splunk.
Fornitori di servizi gestiti	Soggetti che distribuiscono ed eseguono Splunk sotto forma di servizio per i clienti.

Presentazione generale del processo di selezione di Architetture convalidate

Il processo di selezione delle Architetture convalidate Splunk aiuta a individuare l'architettura più semplice e più snella che soddisfa tutte le esigenze dell'organizzazione.

Passaggi del processo di selezione	Obiettivi	Considerazioni
Passaggio 1: definire i requisiti per: a) indicizzazione e ricerca b) meccanismo/i di raccolta dei dati	*Definire i requisiti.*	• I decisori, le parti interessate e gli amministratori devono collaborare per identificare e definire i requisiti dell'organizzazione. • Se è già presente un deployment, è possibile valutare l'architettura corrente per comprendere cosa comporterebbe spostarsi su un modello convalidato. *Per un questionario di supporto alla determinazione dei requisiti, vedere il successivo passaggio 1.*
Passaggio 2: scegliere una topologia per: a) indicizzazione e ricerca b) ogni meccanismo di raccolta dei dati	*Scegliere una topologia che soddisfa i requisiti individuati.*	• Si dovrà scegliere la topologia che meglio soddisfa i requisiti. • Occorre tenere il tutto semplice e coerente con la SVA, così da apprezzare la facilità del percorso verso la scalabilità. *Per diagrammi e descrizioni delle opzioni di topologia, vedere il successivo passaggio 2.*
Passaggio 3: applicazione di principi e best practice di progettazione	*Definire una priorità per i principi di progettazione ed esaminare le best practice di implementazione specifiche per livello.*	• Ogni principio di progettazione consolida una o più colonne portanti delle Architetture convalidate Splunk. • Occorre attribuire una priorità ai principi di progettazione in funzione delle esigenze della propria organizzazione. • Le raccomandazioni specifiche per livello indirizzano l'implementazione della topologia. *Per una scomposizione dei principi di progettazione, vedere il successivo passaggio 3.*

Passaggio 1a: definizione dei requisiti di indicizzazione e ricerca

Per selezionare la corretta topologia del deployment, occorre esaminare in modo approfondito i propri requisiti. Una volta definiti i propri requisiti, sarà possibile scegliere il modo più semplice e più efficace dal punto di vista dei costi per distribuire Splunk. Più avanti è presentato un questionario che aiuta a determinare i requisiti principali per i livelli di indicizzazione e ricerca del deployment.

Il questionario dei requisiti si concentra sulle aree che avranno un impatto diretto sulla topologia del deployment. Consigliamo quindi vivamente di registrare le risposte alle domande che seguono prima di scegliere una topologia nel passaggio successivo.

Aspetti da tenere in considerazione

Esame dei casi d'uso

In sede di determinazione dei propri requisiti, occorre pensare ai casi d'uso previsti per l'infrastruttura Splunk. Ad esempio, la topologia per il caso d'uso di un DevOps di dipartimento spesso è più semplice rispetto a un caso d'uso mission-criticial (ma non sempre). Occorre considerate approfonditamente i casi d'uso che coinvolgono:

- Ricerca
- Disponibilità
- Requisiti di conformità (aspetto particolarmente importante se serve una fedeltà dei dati al 100% e disponibilità continua)
- Altri scenari di casi d'uso specifici per la propria organizzazione

A seconda dei propri scenari di casi d'uso, è possibile che il deployment debba offrire altre caratteristiche architetturali.

Pensare al futuro

Per determinare i propri requisiti occorre pensare alle esigenze immediate, ma occorre anche considerare la crescita e la scalabilità future. Scalare il deployment può comportare un costo, l'inserimento di altro personale, o l'esigenza di altre risorse che potrebbe essere utile iniziare a programmare sin da subito.

Categorie di topologia

Segue una guida delle categorie di topologie di SVA. Queste categorie vengono utilizzate nel questionario che segue. Riferimenti a queste categorie si trovano anche nei successivi passaggi del processo di selezione della SVA.

Categorie del livello di indicizzazione

Codice categoria	Spiegazione
S	La categoria "S" indica l'indexer di un deployment Splunk con un solo server
D	La categoria "D" indica l'esigenza di livello di indexer distribuiti con almeno 2 indexer
C	La categoria "C" indica l'esigenza di un livello di indexer in cluster (è necessaria la replica dei dati)
M	La categoria "M" indica l'esigenza di un livello di indexer in cluster multisito

Categorie del livello di ricerca

Codice categoria	Spiegazione
1	La categoria "1" indica che una search head singola potrebbe soddisfare i requisiti
2	La categoria "2" indica più sono necessarie più search head per soddisfare il requisito
3	La categoria "3" indica che è necessari un cluster di search head per soddisfare il requisito
4	La categoria "4" indica che per soddisfare il requisito è necessario un cluster di search head distribuito su più siti (un SHC "esteso")
+10	La categoria "+10" indica che per supportare l'app Enterprise Security è necessaria una search head dedicata (in cluster). Aggiungere 10 alla categoria della topologia del livello di ricerca e leggere attentamente la descrizione della topologia per gli specifici requisiti per questa app.

Questionario 1: definizione dei propri requisiti per i livelli di indicizzazione e ricerca

♦ *Vedere la legenda sopra per una spiegazione dei codici delle categorie di topologia. Se la risposta è "sì" a più di una domanda, usare il codice categoria della topologia corrispondente alla domanda con il numero più alto.*

#	Domanda	Considerazioni	Impatto sulla topologia	Categoria della topologia livello indexer ♦	categoria della topologia livello ricerca ♦
1	I volumi giornalieri attesi di inserimento di dati sono inferiori a ~300GB/giorno?	Valutare la crescita a breve termine dei volumi di inserimento giornalieri (~6-12 mesi)	Candidato per un deployment con un server singolo, a seconda delle risposte alle domande sulla disponibilità	S	1
2	Serve un'elevata disponibilità per la raccolta dei dati/indicizzazione?	Se non si prevede di utilizzare Splunk per monitorare i casi d'uso che richiedono un inserimento continuo di dati, potrebbe essere accettabile una temporanea interruzione del flusso di dati in ingresso, presumendo che i dati di log non vadano persi.	Richiede un deployment distribuito per supportare l'inserimento continuo	D	1
3	Presumendo una search head	Se il caso d'uso è rappresentato dal	Richiede un indexer in cluster con un	C	1

#	Domanda	Considerazioni	Impatto sulla topologia	Categoria della topologia livello indexer ♦	categoria della topologia livello ricerca ♦
	disponibile per eseguire una ricerca: i dati devono essere completamente ricercabili in ogni momento, non è possibile sopportare alcun impatto sulla completezza dei risultati della ricerca?	calcolo di metriche di prestazioni e dal monitoraggio generico dell'uso per mezzo di funzioni aggregate, ad esempio, il guasto di un singolo indexer potrebbe non influenzare in modo significativo il calcolo di statistiche su un grande numero di eventi. Se il caso d'uso è invece il controllo della sicurezza e l'individuazione delle minacce, molto probabilmente non sono accettabili punti ciechi nei risultati della ricerca	fattore di replica pari almeno a due (2). Nota: mentre un fattore di replica pari a 2 offre una minima protezione contro possibili errori dei nodi di un singolo indexer, il fattore di replica consigliato (e predefinito) è 3.		
4	Si gestiscono più datacenter e serve il ripristino automatico dell'ambiente Splunk in caso di guasto del datacenter?	I requisiti di ripristino di emergenza possono imporre il funzionamento continuo di due strutture (attiva/attiva) oppure prescrivere obiettivi di RTO/RPO per il ripristino di emergenza manuale	Il funzionamento continuo richiede un clustering di indexer multisito e almeno due search head attive per garantire il failover sia al livello di inserimento di dati/indicizzazione sia al livello di ricerca.	M	2
5	Ipotizzando un inserimento di dati continuo e senza perdite, è richiesta un'elevata disponibilità per il livello di ricerca a	Se Splunk viene usato per il monitoraggio continuo in tempo quasi reale, probabilmente	Richiede search head ridondanti, eventualmente il clustering delle search head	D/C/M	3

#	Domanda	Considerazioni	Impatto sulla topologia	Categoria della topologia livello indexer ♦	categoria della topologia livello ricerca ♦
	contatto con l'utente?	non sono tollerabili interruzioni nel livello di ricerca. La situazione potrebbe invece essere diversa per altri casi d'uso.			
6	Serve supporto per un grande numero di utenti concomitanti e/o un significativo carico di lavoro per ricerche pianificate?	I requisiti per più di ~50 utenti/ricerche concomitanti impongono solitamente di scalare in orizzontale il livello di ricerca	Può richiedere una topologia che utilizza un cluster di search head nel livello di ricerca	D/C/M	3
7	In un ambiente con più datacenter, occorre che gli artefatti degli utenti (ricerche, dashboard e altri oggetti knowledge) siano sincronizzati tra i siti?	Da questo dipende se gli utenti avranno un'esperienza aggiornata e coerente in caso di guasto di un sito.	Richiede un cluster di search head "esteso" su più siti, con opportuna configurazione. **Importante:** se da una parte un SHC esteso può migliorare la disponibilità della ricerca per gli utenti durante la completa indisponibilità di un sito, non è possibile garantire che tutti gli artefatti vengano sempre replicati in entrambi i siti. Questo aspetto può avere ripercussioni su specifiche applicazioni che fanno affidamento su artefatti coerenti e aggiornati, come ad es. l'app Splunk Enterprise Security. Il clustering di search head da solo non può offrire una soluzione completa di disaster recovery. Gli altri benefici di	M	4

#	Domanda	Considerazioni	Impatto sulla topologia	Categoria della topologia livello indexer ♦	categoria della topologia livello ricerca ♦
			un SHC vengono mantenuti.		
8	Vi è l'intenzione di distribuire l'app Splunk Enterprise Security (ES)?	Assicurarsi di aver letto e compreso le specifiche limitazioni a cui è soggetta l'app Splunk Enterprise Security, documentate con ciascuna topologia.	ES necessita di un ambiente con search head dedicato (in modalità standalone o in cluster).	D/C/M	+10
9	È presente un ambiente geograficamente distribuito sottoposto a normative di conservazione dei dati?	La normativa di alcuni paesi non consente ai dati generati nel paese di lasciare i sistemi che si trovano in quel paese	Alcune regolamentazioni non consentono il deployment di un livello di indicizzazione centrale per Splunk e richiedono lo sviluppo di un'architettura personalizzata in collaborazione tra Splunk/Partner e il cliente che valuta approfonditamente i dettagli di tale deployment. In altre parole, non esiste una SVA per soddisfare questo requisito.	Personalizzata	Personalizzata
10	Sono previsti criteri di sicurezza molto restrittivi che impediscono di posizionare nella stessa sede fonti specifiche dei dati di log su server/indexer condivisi?	È possibile che, in funzione delle policy aziendali, i dati di log altamente sensibili non possano trovarsi nella stessa sede di altri set di dati a minore rischio nello stesso sistema fisico/nella stessa zona di rete.	Servono più ambienti di indicizzazione indipendenti, eventualmente con un livello di ricerca ibrido e condiviso. Questo aspetto esula dal campo di applicazione delle SVA e richiede lo sviluppo di un'architettura personalizzata.	Personalizzata	Personalizzata

Come determinare il codice della categoria della topologia

Sulla base delle risposte al questionario dei requisiti che precede, si otterrà un indicatore combinato della categoria topologica che consente di individuare la migliore topologia per le proprie esigenze. Seguono le istruzioni e alcuni esempi.

Istruzioni

1. Trascrivere le domande a cui si è risposto con un "sì".
2. Se la risposta è stata "sì" a più di una domanda, seguire la raccomandazione di topologia corrispondente alla domanda con il numero più alto. Se sono presenti più opzioni di topologia (ad esempio, "D/C/M"), guardare alle precedenti domande per determinare quale è l'opzione più adatta al proprio caso.
3. Il codice della categoria topologica inizia con la lettera che rappresenta il livello indexer (ad esempio, "C" o "M"). Questa lettera è seguita dal numero che rappresenta il livello di ricerca (ad esempio, "1" o "13").

Esempio #1

Ipotizziamo che la risposta sia "sì" alle domande #3, #5 e #8. Si otterrà la categoria topologica "C13", la quale indica la necessità di un livello di indicizzazione in cluster con due cluster di search head.

Esempio #2

Ipotizziamo ora che la risposta sia stata "sì" solo alla domanda #1. Si arriverà alla categoria topologica "S1", che indica come topologia ideale un deployment Splunk con un unico server.

Passaggio 2a: scelta di una topologia per l'indicizzazione e la ricerca

Solitamente si distinguono le topologie tra deployment non in cluster e deployment in cluster. I deployment non in cluster richiedono il numero minimo di componenti distinti e presentano eccellenti caratteristiche di scalabilità. Occorre tenere a mente che anche se i deployment non in cluster presentano caratteristiche limitate di disponibilità e ripristino di emergenza, questa opzione di deployment può essere comunque una buona scelta per la propria organizzazione.

Ricorda: lo scopo primario del processo di selezione della SVA è consentire di creare quello che serve senza introdurre componenti non necessari.

Nota

Rimane comunque possibile implementare una topologia che offre benefici in più rispetto alle esigenze immediate, ma occorre tenere a mente che essa probabilmente comporterà anche costi non necessari. Inoltre, l'introduzione di una maggiore complessità è solitamente controproducente per l'efficienza operativa.

Nota importante sui diagrammi di topologia

Le icone nei diagrammi di topologia rappresentano **ruoli funzionali Splunk** e non presuppongono una infrastruttura dedicata per eseguirli. Vedere l'Appendice per indicazioni su come i ruoli Splunk possono trovarsi nella stessa infrastruttura/sullo stesso server.

Utilizzo del codice della categoria topologica

Prima di selezionare un'opzione di topologia, è altamente consigliato di compilare il questionario dei requisiti per determinare il codice della categoria topologica. In caso contrario, tornare indietro e completare il passaggio precedente. Una volta ottenuto il codice della categoria topologica, sarà possibile identificare l'opzione di deployment che meglio si adatta ai requisiti indicati.

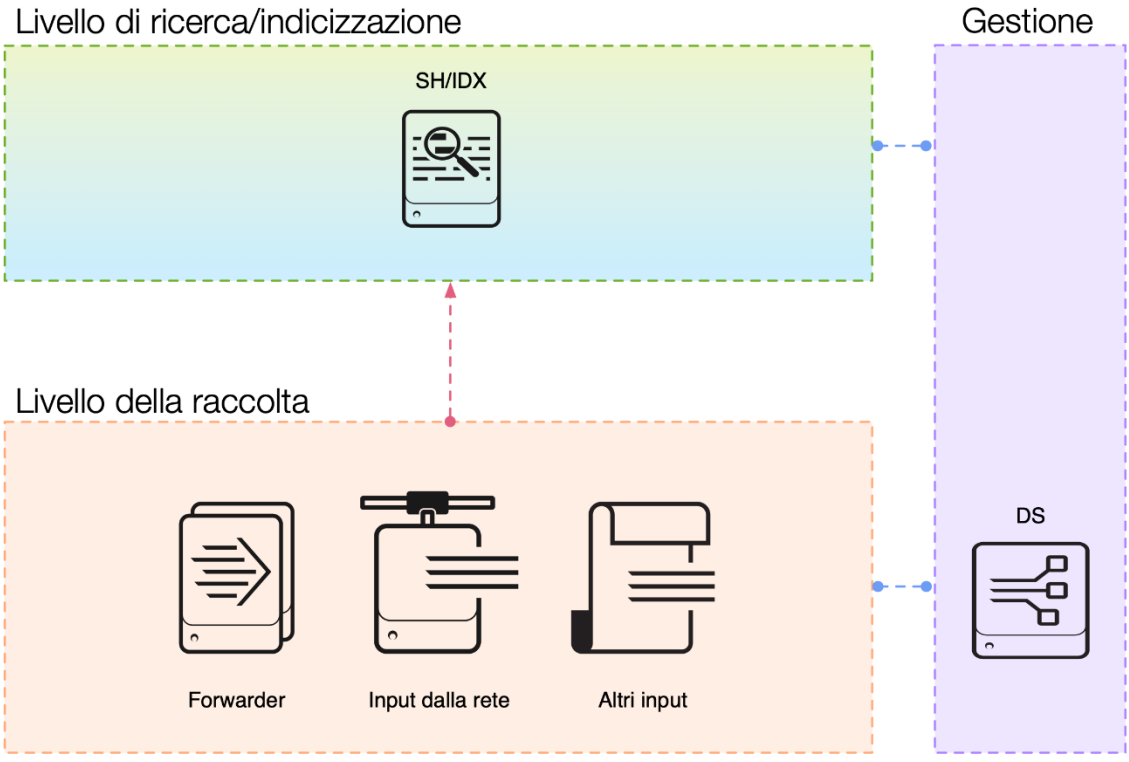
Opzioni di deployment non in cluster

Seguono le seguenti opzioni di topologia:

Tipo di deployment	Codice/i della categoria topologica
Deployment a server singolo	S1
Deployment distribuito non in cluster	D1 / D11

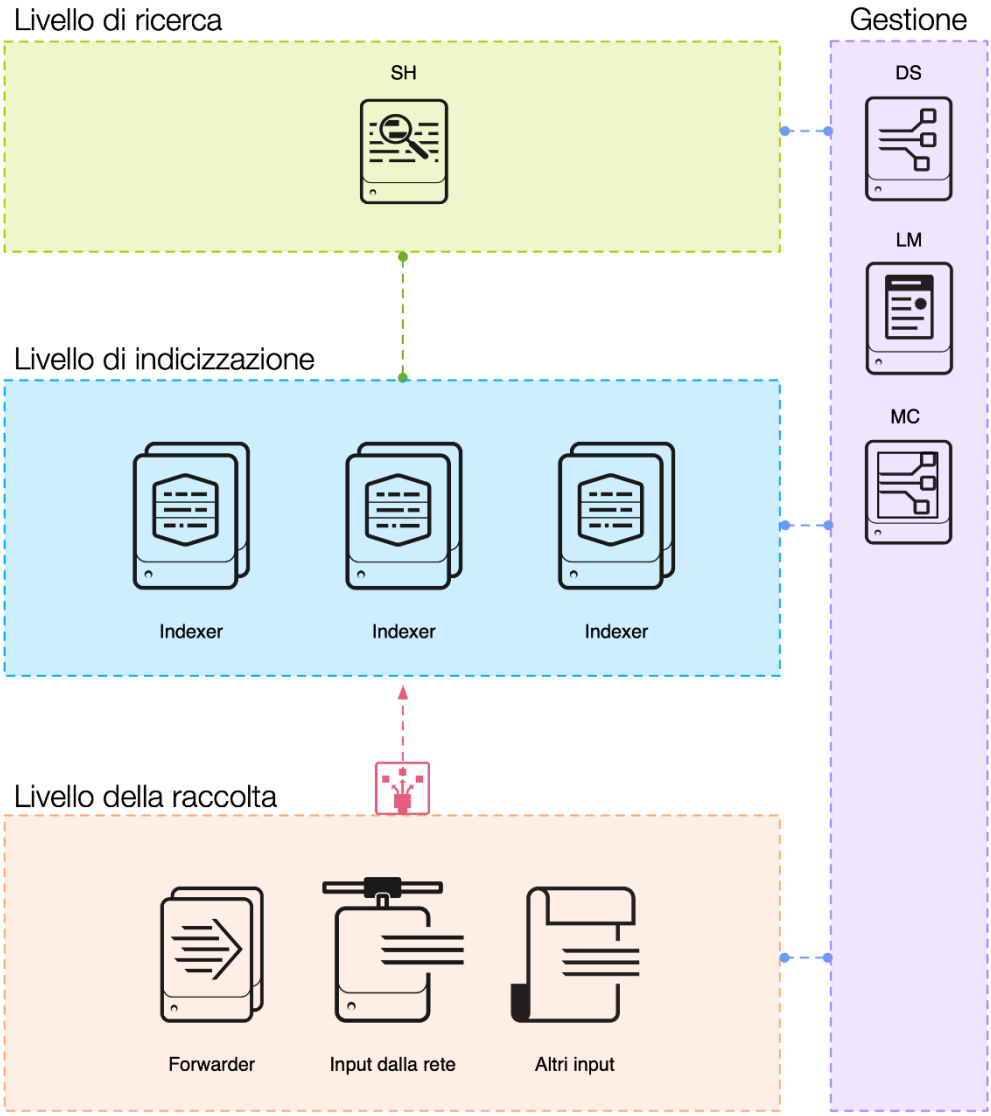
Per un'illustrazione dei componenti della topologia, si rimanda alla successiva Appendice "B".

Deployment a server singolo (S1)



Descrizione del deployment a server singolo (S1)	Limitazioni
Questa topologia del deployment offre una soluzione a costi contenuti se l'ambiente soddisfa tutti i seguenti criteri: a) non è necessario offrire elevata disponibilità né ripristino di emergenza automatico per il deployment Splunk, b) il volume giornaliero di inserimento di dati è inferiore a 300GB/giorno, e c) il numero di utenti è limitato e i casi d'uso di ricerca non sono critici. Questa topologia trova impiego solitamente per i casi d'uso di minore rilevanza e non essenziali per l'impresa (spesso di natura dipartimentale). Tra i casi d'uso opportuni vi sono ambienti di test per l'onboarding, piccoli casi d'uso DevOps, ambienti di test e integrazione di applicazioni, e scenari simili. Tra i principali benefici di questa topologia vi sono la facilità di gestione, le buone prestazioni di ricerca per volumi di dati contenuti, e un TCO fisso.	• No elevata disponibilità di ricerca/indicizzazione • Scalabilità limitata dalla capacità hardware (percorso di migrazione diretta verso deployment distribuito)

Deployment distribuito non in cluster (D1 / D11)



Descrizione del deployment distribuito non in cluster (D1 / D11)	Limitazioni
È necessario spostarsi verso una topologia distribuita in una delle seguenti situazioni: a) il volume giornaliero di dati da inviare a Splunk è maggiore della capacità di un deployment con un unico server, o b) si desidera/occorre garantire l'elevata disponibilità della funzione di inserimento di dati. Il deployment di più indexer indipendenti consente di scalare la capacità di indicizzazione in modo lineare e di incrementare in modo implicito la disponibilità per l'inserimento di dati. Il TCO aumenta in modo prevedibile e lineare, man mano che si aggiungono i nodi indexer. L'introduzione (consigliata) del componente della console di monitoraggio (MC) consente di tenere sotto osservazione lo stato di salute e la capacità del deployment distribuito. Inoltre la MC mette a disposizione un sistema di allarme centralizzato che consente di ricevere notifiche nel caso in cui il deployment si trovi in una condizione non ottimale. La/le search head dovranno essere configurate manualmente con l'elenco delle destinazioni di ricerca disponibili ogni volta che si aggiunge un nuovo indexer. **Nota per i clienti ES:** se il codice della categoria è D1 (si intende quindi distribuire l'app Splunk Enterprise Security), per distribuire l'app è necessaria un'**unica** search head **dedicata** (non raffigurata nel diagramma della topologia). Il livello della raccolta deve essere configurato con l'elenco degli indexer target (tramite un deployment server) ogni volta che viene aggiunto un nuovo indexer. Questa topologia del deployment può scalare in modo lineare fino a oltre 1000 nodi indexer e può quindi supportare volumi di inserimento di dati e ricerca estremamente elevati. Grazie all'esecuzione di ricerche parallele su più indexer (map/reduce), è possibile mantenere le prestazioni di ricerca anche in set di dati di grandi dimensioni. Anche se non specificamente separato come topologia distinta, si può usare un cluster di search head per aumentare la capacità di ricerca al livello di ricerca (vedere il livello di ricerca nella topologia C3/C13).	• No elevata disponibilità per il livello di ricerca • Elevata disponibilità limitata per il livello di indicizzazione, l'errore di un nodo può generare risultati della ricerca incompleti per le ricerche storiche

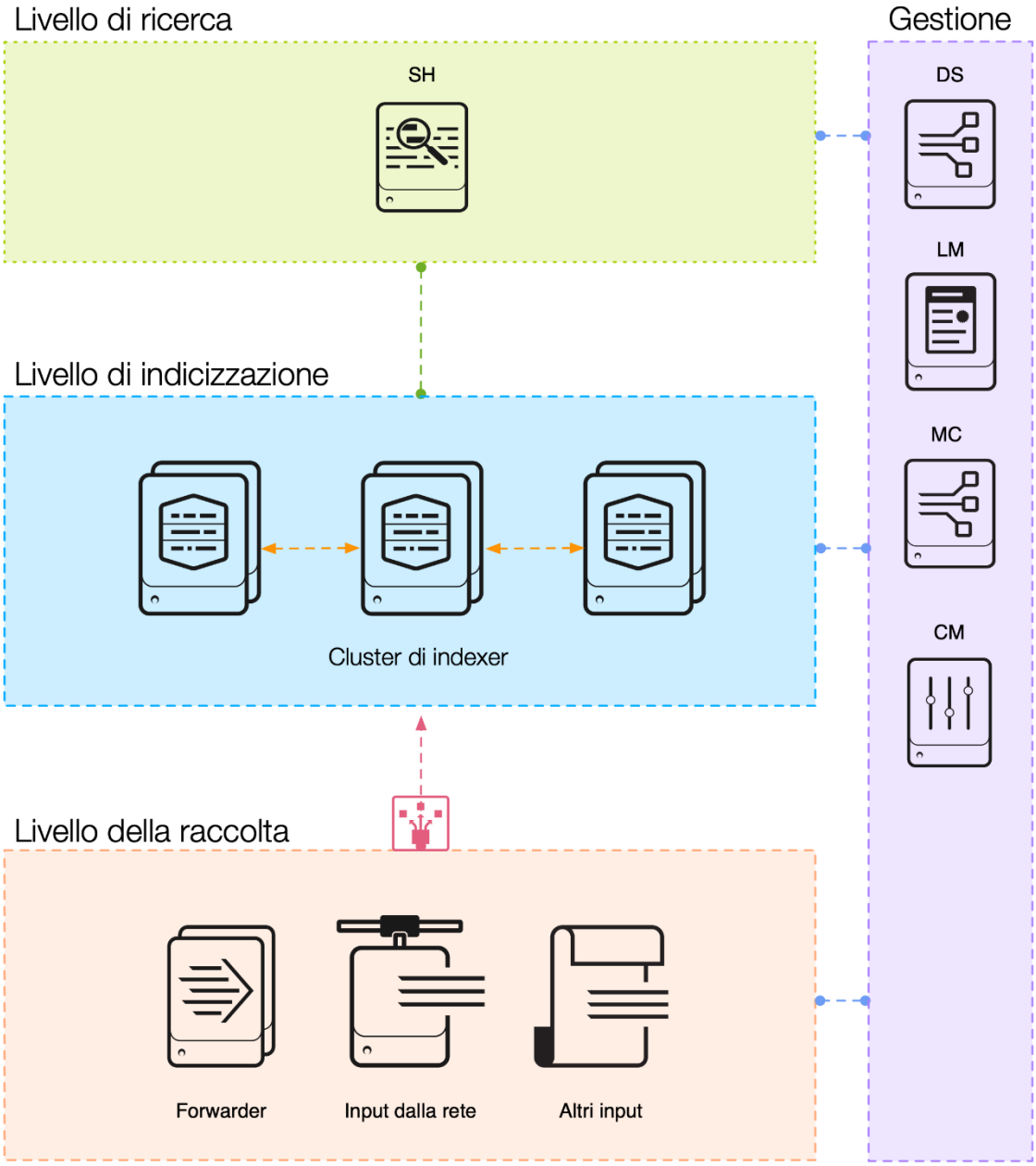
Opzioni di deployment in cluster

Più avanti sono presentate le seguenti opzioni di topologia:

Tipo di deployment	Codice/i della categoria topologica
Deployment distribuito in cluster - un solo sito	C1 / C11
Deployment distribuito in cluster + SHC - un solo sito	C3 / C13
Deployment distribuito in cluster - multisito	M2 / M12
Deployment distribuito in cluster + SHC - multisito	M3 / M13
Deployment distribuito in cluster + SHC - multisito	M4 / M14

Per un'illustrazione dei componenti della topologia, si rimanda alla successiva Appendice "B".

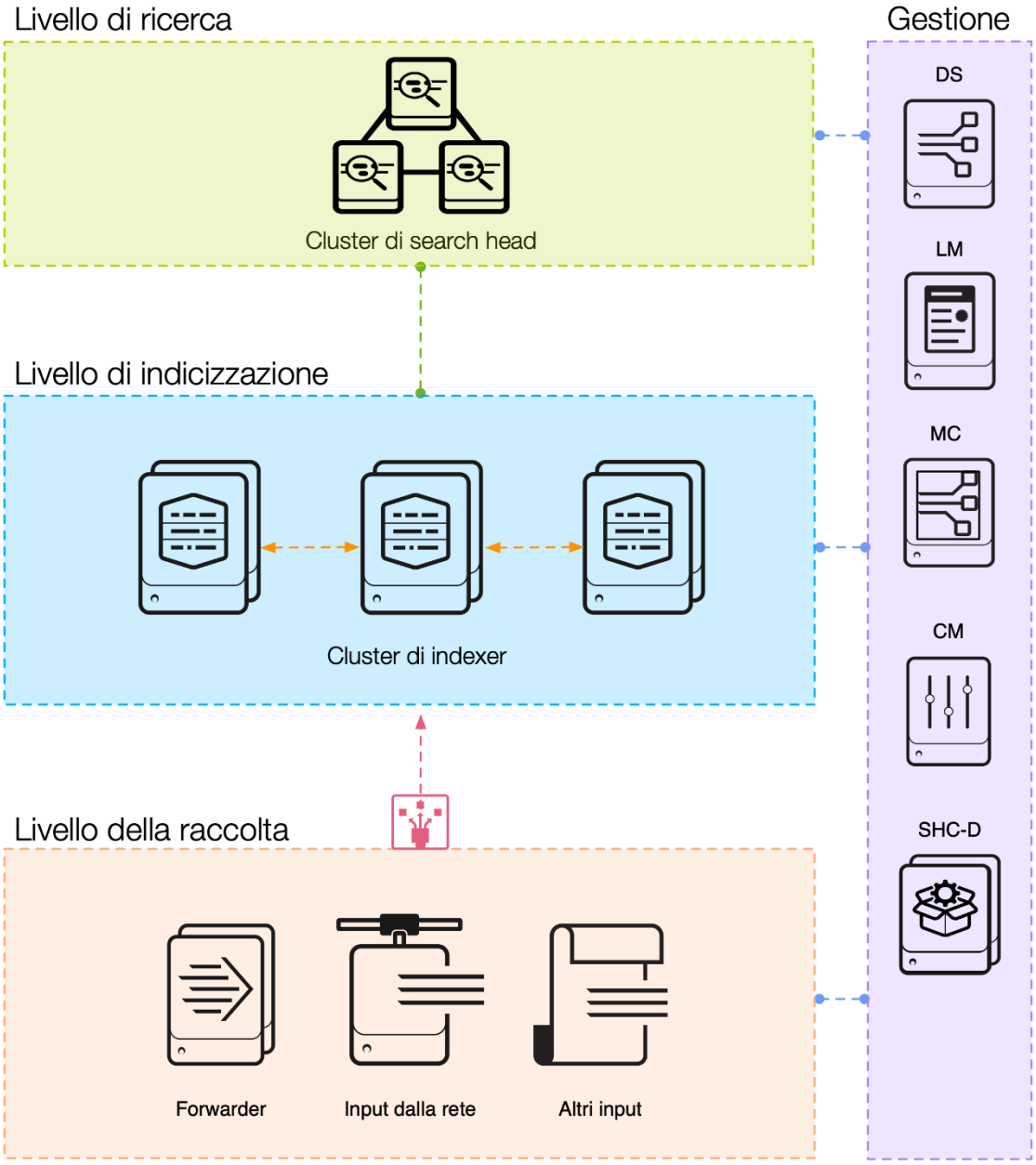
Deployment distribuito in cluster - un solo sito (C1 / C11)



Descrizione del deployment distribuito in cluster - un solo sito (C1 / C11)	Limitazioni
Questa topologia introduce il clustering di indexer in abbinamento a criteri di replica dei dati opportunamente configurati. Questo offre un'elevata disponibilità dei dati in caso di errore del nodo peer di un indexer. Occorre tuttavia ricordare che questo vale solo per l'indicizzazione e non mette al riparo da malfunzionamenti della search head. **Nota per i clienti ES:** se il codice della categoria è C11 (si intende quindi distribuire l'app Splunk Enterprise Security), per distribuire l'app è necessaria un'**unica** search head **dedicata** (non raffigurata nel diagramma della topologia). Questa topologia richiede un ulteriore componente Splunk detto master cluster (CM). Il CM si occupa del coordinamento e dell'applicazione dei criteri di replica dei dati configurati. Il CM	• No elevata disponibilità per il livello di ricerca • Il numero totale di bucket unici nel cluster di indexer è limitato a 5MM (V6.6+), 15MM bucket • No capacità di disaster recovery automatica in caso di indisponibilità del datacenter

Descrizione del deployment distribuito in cluster - un solo sito (C1 / C11)	Limitazioni
serve anche come fonte autorevole per i peer del cluster disponibili (indexer). La configurazione delle search head è semplificata perché la configurazione avviene sul CM anziché sui singoli peer di ricerca. È possibile configurare il livello di inoltro per individuare gli indexer disponibili attraverso il CM. Questo semplifica la gestione del livello di inoltro. Occorre ricordare che la replica dei dati all'interno del cluster avviene in modalità non deterministica. Non è possibile controllare dove vengono conservate le copie richieste di ciascun evento. Inoltre, anche se la scalabilità è lineare, vi sono limitazioni relativamente alla dimensione totale del cluster (~50PB di dati ricercabili in condizioni ideali). Si consiglia di prevedere la console di monitoraggio (MC) per tenere sotto osservazione la condizione di salute dell'ambiente Splunk.	

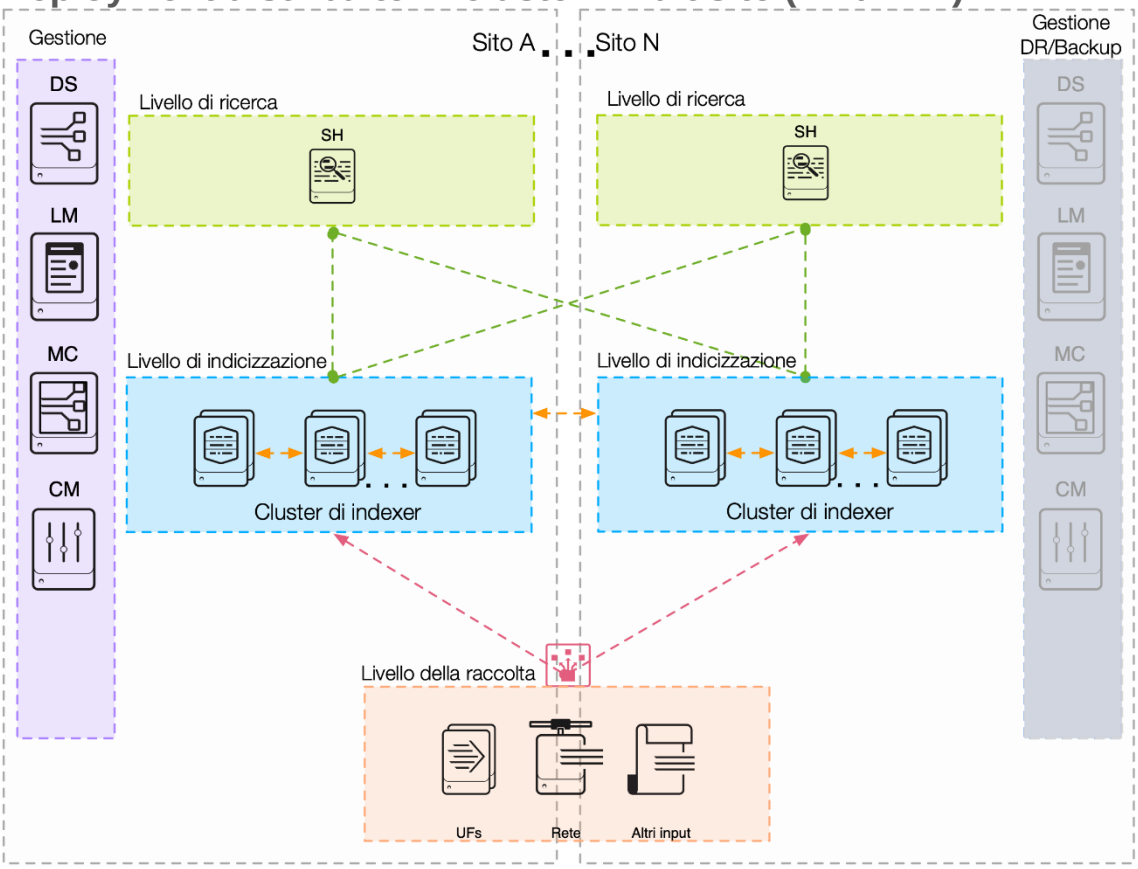
Deployment distribuito in cluster + SHC - un solo sito (C3 / C13)



Descrizione di deployment distribuito in cluster + SHC - un solo sito (C3 / C13)	Limitazioni
Questa topologia aggiunge scalabilità orizzontale ed elimina il single point of failure (singolo punto di vulnerabilità) dal livello di ricerca. Per implementare un SHC servono almeno tre search head. Per gestire la configurazione SHC, per ogni SHC serve un ulteriore componente Splunk chiamato deployer del cluster di search head. Questo componente è necessario per distribuire i cambiamenti necessari ai file di configurazione nel cluster. Il deployer del cluster di search head non ha requisiti di elevata disponibilità (non ha un ruolo runtime). L'SHC mette a disposizione il meccanismo per aumentare la capacità di ricerca disponibile oltre quella che una singola search head può offrire. Inoltre, l'SHC consente la distribuzione del carico	• No capacità di disaster recovery in caso di indisponibilità del datacenter • ES richiede SH/SHC dedicati • La gestione di un deployment ES su SHC è supportata, ma impegnativa (comporta PS)

Descrizione di deployment distribuito in cluster + SHC - un solo sito (C3 / C13)	Limitazioni
di lavoro delle ricerche pianificate all'interno del cluster. L'SHC offre anche un failover utente ottimale in caso di indisponibilità della search head. Davanti ai membri dell'SHC è necessario un bilanciatore di carico di rete che supporta le sessioni "sticky" per garantire il corretto bilanciamento del carico degli utenti nel cluster. **Nota per i clienti ES:** se il codice della categoria è C13 (si intende quindi distribuire l'app Splunk Enterprise Security), per distribuire l'app è necessario un cluster di search head **dedicato** (non raffigurato nel diagramma della topologia). Il livello di ricerca può contenere search head in cluster e non in cluster a seconda della capacità e delle esigenze dell'organizzazione (anche questo non è raffigurato nel diagramma della topologia).	• L'SHC non può avere oltre 100 nodi

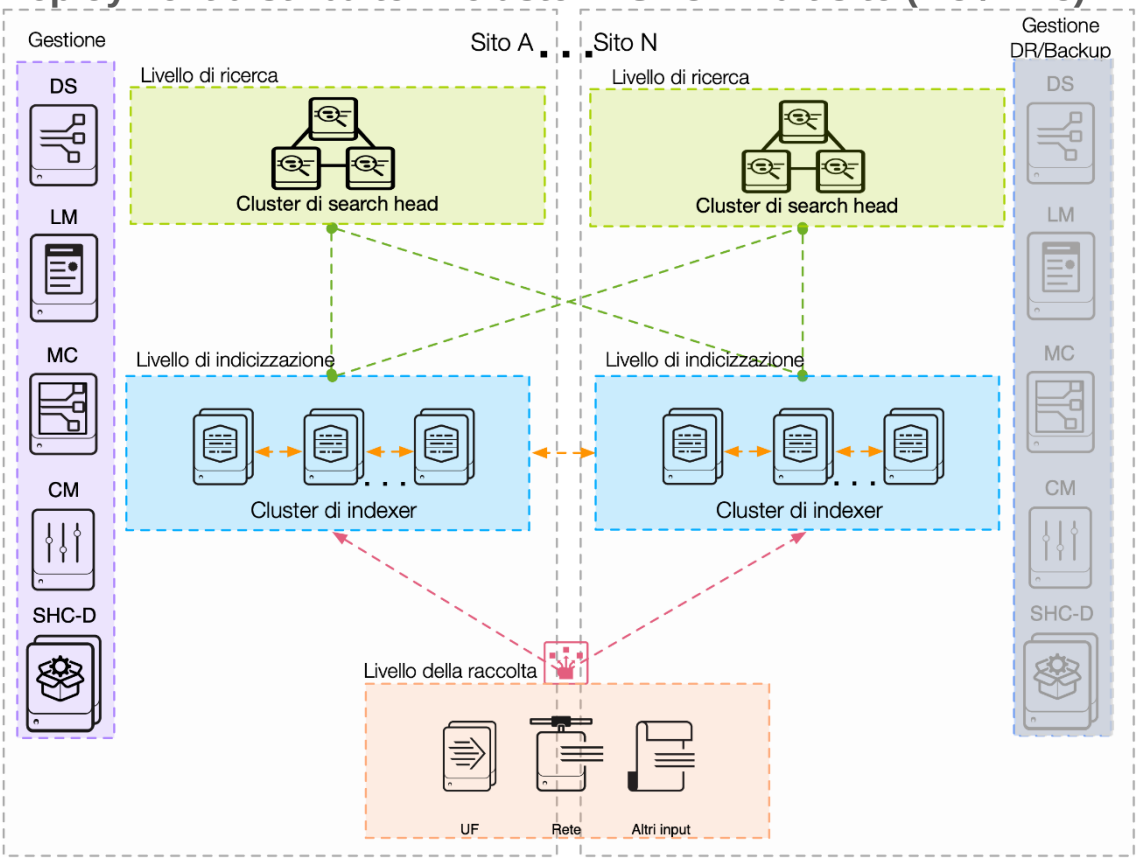
Deployment distribuito in cluster - multisito (M2 / M12)



Descrizione di deployment distribuito in cluster - multisito (M2 / M12)	Limitazioni
Per offrire un ripristino di emergenza quasi automatico in caso di evento catastrofico (ad es. indisponibilità del datacenter), il clustering multisito è l'architettura del deployment da preferire. Un cluster multisito in salute richiede una latenza di rete accettabile tra i siti, come specificato nella Documentazione Splunk.	• Nessuna condivisione della capacità delle search head disponibili e nessuna replica degli artefatti di ricerca **tra i siti** • Un malfunzionamento delle funzioni di gestione deve

Descrizione di deployment distribuito in cluster - multisito (M2 / M12)	Limitazioni
Questa topologia consente di replicare in modo **deterministico** i dati su due o più gruppi di peer del cluster di indexer. Sarà possibile configurare la replica del sito e il fattore di ricerca. Questo fattore di replica del sito consente di specificare dove vengono inviate le repliche, e garantisce che i dati vengano distribuiti su più sedi. Viene comunque gestito da un unico nodo primario del cluster, di cui occorre eseguire il failover sul sito di DR in caso di emergenza. Il clustering multisito offre la ridondanza dei dati in sedi distribuite e fisicamente separate, con la possibilità di una distribuzione separata geograficamente. Gli utenti possono eseguire automaticamente il failover sul sito di DR per assicurare la disponibilità. Questa topologia non offre però un meccanismo per sincronizzare automaticamente tra i siti la configurazione del livello di ricerca e gli artefatti runtime. La capacità disponibile delle destinazioni di ricerca (indexer) tra i siti può essere utilizzata per l'esecuzione delle ricerche in un modello attivo/attivo. Se possibile, si può configurare la site-affinity per assicurare che gli utenti che hanno effettuato l'accesso alla search head di uno specifico sito effettuino ricerche solo negli indexer locali. **Nota per i clienti ES:** se il codice della categoria è M12 (si intende distribuire l'app Splunk Enterprise Security), per distribuire l'app è necessaria un'**unica** search head **dedicata** (non raffigurata nel diagramma della topologia). Per la search head ES, il failover comporta l'impostazione di una search head "shadow" nel sito di failover che viene attivata e utilizzata solo in caso di disaster recovery. Si consiglia di rivolgersi ai servizi professionali di Splunk per progettare e implementare un meccanismo di failover del sito per il deployment di Enterprise Security.	essere gestito fuori da Splunk in caso di guasto del sito • La latenza tra i siti per la replica degli indici deve essere compresa nei limiti consigliati

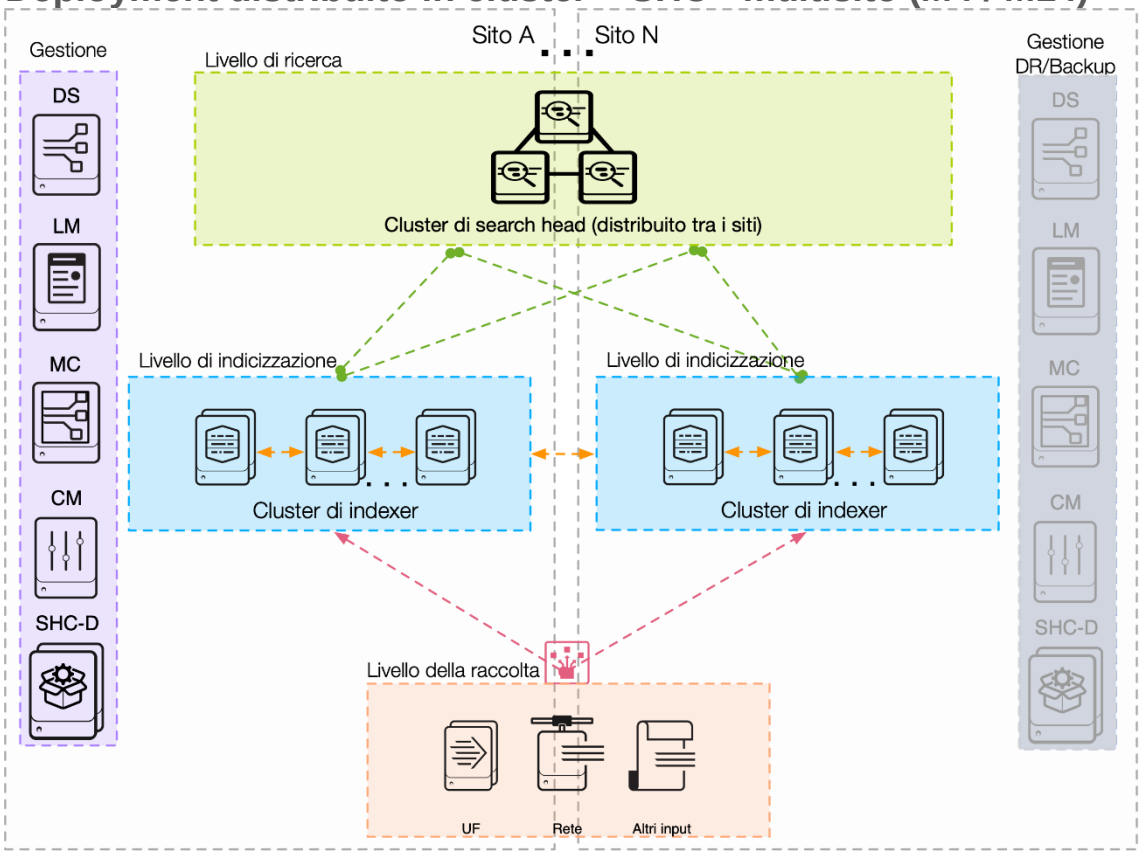
Deployment distribuito in cluster + SHC - multisito (M3 / M13)



Descrizione del deployment distribuito in cluster + SHC - multisito (M3 / M13)	Limitazioni
Questa topologia aggiunge scalabilità orizzontale ed elimina il single point of failure (singolo punto di vulnerabilità) dal livello di ricerca in ciascun sito. Per implementare un SHC servono almeno tre search head (per sito). Per gestire la configurazione SHC, per ogni SHC serve un ulteriore componente Splunk chiamato deployer del cluster di search head. Questo componente è necessario per distribuire i cambiamenti necessari ai file di configurazione nel cluster. Il deployer del cluster di search head non ha requisiti di elevata disponibilità (non ha un ruolo runtime). L'SHC offre i seguenti benefici: a) maggiore capacità di ricerca disponibile, oltre quella che una singola search head può offrire, b) distribuzione del carico di lavoro delle ricerche pianificate nel cluster, e c) failover utente ottimale in caso di indisponibilità di una search head. Davanti ai membri dell'SHC in ogni sito è necessario un bilanciatore di carico di rete che supporta le sessioni "sticky" per garantire il corretto bilanciamento del carico degli utenti nel cluster. **Nota per i clienti ES:** se il codice della categoria è M13 (si intende distribuire l'app Splunk Enterprise Security), per distribuire l'app è richiesto un unico cluster di search head dedicato *contenuto all'interno di un sito* (non esplicitamente raffigurato nel diagramma della topologia). Per poter ripristinare un ambiente ES SH da un guasto, è possibile usare tecnologie di terze parti per eseguire un	• Nessuna replica degli artefatti di ricerca tra i siti, gli SHC operano in modalità standalone • La latenza tra i siti per la replica degli indici deve essere compresa nei limiti documentati • L'SHC non può avere oltre 100 nodi

Descrizione del deployment distribuito in cluster + SHC - multisito (M3 / M13)	Limitazioni
failover delle istanze delle search head, oppure si può prevedere il provisioning di un "warm standby" di ES SH mantenuto sincronizzato con l'ambiente ES primario. Si consiglia vivamente di rivolgersi ai servizi professionali di Splunk in caso di deployment di ES in un ambiente con esigenze di elevata disponibilità/disaster recovery.	

Deployment distribuito in cluster + SHC - multisito (M4 / M14)



Descrizione del deployment distribuito in cluster + SHC - multisito (M4 / M14)	Limitazioni
Si tratta dell'architettura convalidata più complessa, pensata per i deployment che hanno rigidi requisiti di elevata disponibilità e ripristino di emergenza. Consigliamo vivamente di coinvolgere i servizi professionali di Splunk per un corretto deployment. Se correttamente distribuita, questa topologia offre un funzionamento continuo dell'infrastruttura Splunk per la raccolta dei dati, l'indicizzazione, e la ricerca. Questa topologia comporta l'implementazione di un cluster "esteso" di search head distribuito su uno o più siti. Offre un failover ottimale per gli utenti in caso di errore di un nodo di ricerca o di un datacenter. Gli artefatti di ricerca e gli oggetti knowledge a runtime sono replicati nell'SHC. Serve un'attenta configurazione per assicurarsi che la replica avvenga nei diversi siti, perché di per sé l'SHC non rileva il sito (la replica degli artefatti non è deterministica).	• La latenza di rete tra i siti deve essere entro limiti documentati • Il failover dell'SHC può richiedere dei passaggi manuali solo se sopravvive una minoranza dei membri del cluster

Descrizione del deployment distribuito in cluster + SHC - multisito (M4 / M14)	Limitazioni
È possibile configurare la site-affinity per accertarsi che il collegamento WAN tra i siti venga utilizzato solo nei casi in cui non è possibile soddisfare la ricerca in locale. Davanti ai membri dell'SHC è necessario un bilanciatore di carico di rete che supporta le sessioni "sticky" per garantire il corretto bilanciamento del carico degli utenti nel cluster. **Nota per i clienti ES:** se il codice della categoria è M14 (si intende distribuire l'app Splunk Enterprise Security), per distribuire l'app è richiesto un unico cluster di search head dedicato *contenuto all'interno di un sito* (non esplicitamente raffigurato nel diagramma della topologia). ES richiede che sia disponibile un set coerente di artefatti runtime, aspetto che non può essere garantito in un SHC esteso in caso di indisponibilità di un sito. Per poter ripristinare un ambiente ES SH da un guasto, è possibile usare tecnologie di terze parti per eseguire un failover delle istanze delle search head, oppure si può prevedere il provisioning di un "warm standby" di ES SH mantenuto sincronizzato con l'ambiente ES primario. Si consiglia vivamente di rivolgersi ai servizi professionali di Splunk in caso di deployment di ES in un ambiente con esigenze di elevata disponibilità/disaster recovery.	

Passaggio 1b: definizione dei requisiti per la raccolta dei dati

Il livello della raccolta dei dati è un componente fondamentale di un deployment Splunk. Esso consente a qualunque dispositivo nell'ambiente di inoltrare dati al livello di indicizzazione ai fini dell'elaborazione, rendendoli così disponibili per le ricerche in Splunk. Il fattore più importante è assicurare che l'inoltro e l'indicizzazione avvengano nel modo più efficiente e affidabile, trattandosi di un aspetto critico per il successo e le prestazioni del deployment Splunk.

Considerare i seguenti aspetti dell'architettura del livello di raccolta dei dati:

- L'origine dei dati. Provengono da file di log, fonti Syslog, input di rete, strumenti di log degli eventi del sistema operativo, applicazioni, message bus o da altre fonti?
- Requisiti di latenza dell'inserimento di dati e capacità
- Distribuzione ideale degli eventi tra gli indexer del livello di indicizzazione
- Tolleranza agli errori e ripristino automatico (elevata disponibilità)
- Requisiti di sicurezza e titolarità dei dati

Questa sezione delle SVA si concentra sui metodi comuni di raccolta dei dati. Questa sezione presenta anche l'architettura e le best practice per ogni metodo di raccolta dei dati, richiamando i possibili aspetti da considerare al momento di scegliere l'implementazione.

Considerazioni importanti sull'architettura e motivo della loro importanza

Data la centralità del ruolo del livello di raccolta dei dati, è importante capire le considerazioni più rilevanti che entrano in gioco nella progettazione dell'architettura.

Anche se in base alle proprie esigenze alcune di queste considerazioni possono non essere pertinenti, quelle riportate in grassetto nella tabella che segue descrivono aspetti fondamentali pertinenti per qualunque ambiente.

Considerazione	Perché è importante?
I dati sono inseriti correttamente (timestamp, interruzioni di riga, troncamento)	L'importanza di una distribuzione ideale degli eventi tra gli indexer non può essere sovrastimata. Il livello di indicizzazione lavora con la massima efficienza se tutti gli indexer disponibili vengono utilizzati in modo analogo. Questo vale sia per l'inserimento di dati sia per le prestazioni di ricerca. Un unico indexer che gestisce un volume significativamente superiore di inserimento di dati rispetto agli altri indexer può avere ripercussioni negative sui tempi di risposta delle ricerche. Per gli indexer che hanno capacità limitata di archiviazione sul disco locale, una distribuzione degli eventi non omogenea può anche rendere i dati vecchi prima di soddisfare i criteri di conservazione dei dati configurati.
I dati vengono distribuiti in modo ottimale tra i diversi indexer disponibili	Se i dati non vengono inseriti correttamente a causa di un'errata configurazione dei timestamp e delle interruzioni di riga, la ricerca all'interno di questi dati diventa molto difficile, perché i confini tra gli eventi devono essere forzati al momento della ricerca. L'estrazione non corretta del timestamp, o l'assenza del timestamp, può causare un'indesiderata assegnazione implicita del timestamp. Questo può confondere gli utenti e rendere molto più difficile del necessario sfruttare i dati.
Tutti i dati raggiungono il livello di indicizzazione in modo affidabile e senza perdite	Ogni dato di log raccolto per gli scopi di un'analisi affidabile deve essere completo e valido, in modo che le ricerche eseguite su tali dati forniscano risultati validi e precisi.
Tutti i dati raggiungono il livello di indicizzazione con una latenza minima	I ritardi nell'inserimento di dati fanno lievitare il tempo intercorso tra un evento potenzialmente critico che si verifica e la possibilità di cercarlo e di rispondere ad esso. Una latenza minima di inserimento è spesso cruciale per i casi d'uso di monitoraggio che prevedono allarmi indirizzati al personale o interventi automatizzati.
I dati sono protetti durante il transito	Se i dati sono sensibili o devono essere protetti durante la trasmissione su reti non sicure, può essere necessario ricorrere alla crittografia dei dati per prevenire l'intercettazione non autorizzata da parte di terzi. In generale consigliamo di attivare SSL su tutte le connessioni tra componenti Splunk.
L'uso di risorse di rete è contenuto al minimo	L'impatto sulle risorse di rete della raccolta di dati di log deve essere contenuto al minimo per non influenzare il traffico di rete critico per l'azienda. Per le reti basate su linee a noleggio, ridurre l'utilizzo della rete contribuisce anche a contenere il TCO totale del deployment.
Autenticazione/ autorizzazione delle fonti di dati	Per prevenire che fonti di dati malevole producano effetti sull'ambiente di indicizzazione, è utile considerare l'autenticazione/autorizzazione del collegamento. Questo aspetto può essere risolto con controlli di rete, oppure applicando meccanismi a livello di applicazione (ad es. SSL/TLS).

Data l'estrema importanza per il deployment, le indicazioni contenute in questo documento si concentrano sulle architetture che supportano una distribuzione degli eventi. Quando un ambiente Splunk non presenta le prestazioni di ricerca attese, quasi sempre la causa è il mancato rispetto dei requisiti minimi di prestazioni dell'archiviazione e/o una distribuzione non omogenea degli eventi, che limita la possibilità di sfruttare la parallelizzazione della ricerca.

Ora che sono chiare le considerazioni più importanti riguardanti l'architettura, è giunto il momento di scoprire quali specifici requisiti di raccolta dei dati occorre soddisfare.

Questionario 2: definizione dei requisiti per la raccolta dei dati

Rispondendo alle seguenti domande si otterrà l'elenco completo di componenti per la raccolta dei dati che servono per il proprio deployment. È possibile usare i codici nella colonna più a destra per trovare maggiori dettagli su ciascuno dei componenti, più avanti nel testo.

#	Domanda	Considerazioni	Impatto sulla topologia	Componenti rilevanti per la raccolta dei dati
1	Occorre monitorare file locali o eseguire script di raccolta dei dati sugli endpoint?	Si tratta di un requisito primario per quasi tutti gli scenari di deployment Splunk.	Occorre installare lo universal forwarder sugli endpoint e gestire la sua configurazione a livello centrale.	UF
2	Occorre raccogliere i dati di log inviati attraverso syslog da dispositivi sui quali non è possibile installare software (apparecchi, switch di rete, ecc.)?	Syslog è un protocollo di rete diffuso che trova spesso impiego per dispositivi realizzati appositamente che non consentono l'installazione di software personalizzato.	Occorre avere un'infrastruttura di server syslog che serva come punto di raccolta.	SYSLOG HEC
3	Occorre supportare la raccolta di dati di log da applicazioni che effettuano il log su API anziché scrivere su dischi locali?	La scrittura in file di log sugli endpoint impone di mettere a disposizione spazio su disco e di gestire questi file di log (rotazione, cancellazione, ecc.). Alcuni clienti desiderano abbandonare questo modello ed effettuare il log direttamente in Splunk utilizzando le librerie di logging disponibili.	Occorre usare Raccolta eventi HTTP di Splunk (HEC) o altre tecnologie che svolgano funzione di collettore dei log.	HEC
4	Occorre raccogliere dati da un provider di dati di eventi in streaming?	Molte aziende hanno adottato un modello di eventi ad hub, in cui una piattaforma di dati in streaming (come AWS Kinesis o Kafka) serve come trasporto dei messaggi tra i produttori dei dati di log e i consumatori.	Occorre integrare il provider di dati in streaming e Splunk.	KAFKA KINESIS HEC
5	Sono presenti criteri di sicurezza non modificabili che non consentono ai produttori di log di stabilire una connessione TCP direttamente con il livello di indicizzazione?	Talvolta le topologie di rete consistono di più zone di rete con regole firewall restrittive tra loro, che possono rendere impossibile abilitare in modo generalizzato il flusso del traffico tra le zone traffico sulle porte	Occorre un livello di inoltro intermedio che consenta il flusso del traffico tra le zone di rete.	IF

		Splunk. Configurare e mantenere le regole dei firewall per le singole fonti/singoli indirizzi IP target sarebbe eccessivamente oneroso.		
6	Occorre raccogliere dati di log utilizzando programmi, ad es., chiamando REST API o interrogando database?	Splunk offre diversi input modulari che consentono l'esecuzione di script verso API per una grande varietà di casi d'uso di inserimento di dati, tra cui DBX per la raccolta di dati da database relazionali.	Il livello di raccolta dei dati richiederà uno o più nodi di raccolta dei dati (DCN) implementati con uno Splunk Heavy Forwarder.	DCN
7	Occorre inoltrare i dati (o un loro sotto insieme) verso altri sistemi oltre e in aggiunta a Splunk?	Alcuni casi d'uso richiedono che i dati indicizzati in Splunk vengano inoltrati anche a un altro sistema. Spesso i dati inoltrati consistono di un sottoinsieme dei dati originari, oppure di dati che sono stati modificati prima di essere inoltrati.	A seconda del caso d'uso specifico, per supportare l'instradamento e il filtro basati su eventi è possibile che sia necessario un livello di inoltro intermedio realizzato con un Heavy Forwarder. In alternativa si possono inoltrare i dati dopo l'indicizzazione usando il comando cefout presente nella App Splunk per CEF.	HF
8	Sono presenti siti remoti con limitazioni della larghezza di banda e occorre filtrare significativamente i dati prima di inviarli sulla rete?	Per poter filtrare i dati prima di trasmetterli occorre un (heavy) forwarder che esegua il parsing. La larghezza di banda in uscita della rete utilizzata da un HWF è circa 5 volte quella di uno UF, pertanto ha senso filtrare unicamente se viene eliminato un numero significativo di eventi (regola indicativa: >50% dei dati originari). Idealmente, si dovrebbe regolare la granularità del log per ottenere la riduzione necessaria del suo volume.	Se non è possibile ridurre il volume del log alla fonte, occorre un HF intermedio nel sito remoto che esegua il parsing dei dati originari e filtri gli eventi sulla base di una configurazione.	IF HF
9	Occorre mascherare/offuscare i dati sensibili prima di trasmetterli su una rete pubblica per l'indicizzazione?	Talvolta proteggere il traffico del forwarder con SSL non è sufficiente per proteggere i dati sensibili in transito su reti pubbliche e si rende necessario mascherare	Se non è possibile mascherare i dati nell'applicazione che li produce, occorre un HF intermedio nel sito che esegua il parsing dei dati originari e applichi	IF HF

		alcune parti degli eventi prima della trasmissione (SSN, dati CC, ecc.). Idealmente, i dati dovrebbero essere mascherati nell'applicazione che produce i dati di log.	le regole di masking richieste sulla base di una configurazione, prima di inviare i dati agli indexer.	
10	Occorre registrare metriche usando statsd o collectd?	Statsd e collectd sono tecnologie diffuse utilizzate per raccogliere metriche da sistemi host e applicazioni.	Splunk supporta specifici tipi di indici e metodi di raccolta per consentire di alimentare questi indici utilizzando UF, HF o HEC.	METRICHE
11	Serve che uno dei componenti per la raccolta dei dati abbia un'elevata disponibilità?	Solitamente non riguarda gli endpoint, la disponibilità può essere invece rilevante per altri componenti per la raccolta dei dati, come forwarder intermedi o nodi di raccolta dei dati.	Occorre valutare in che modo un guasto influenzerà la disponibilità di ciascun componente, e come affrontarla.	HA

Passaggio 2b: selezione dei componenti per la raccolta dei dati

Dopo aver completato il questionario, si avrà a disposizione un elenco dei componenti per la raccolta dei dati necessari per soddisfare i requisiti del deployment. Questa sezione presenta in maggiore dettaglio ogni componente dell'architettura di raccolta dei dati. Prima di procedere, alcune indicazioni di natura generale.

Indicazioni generali sull'architettura di inoltro

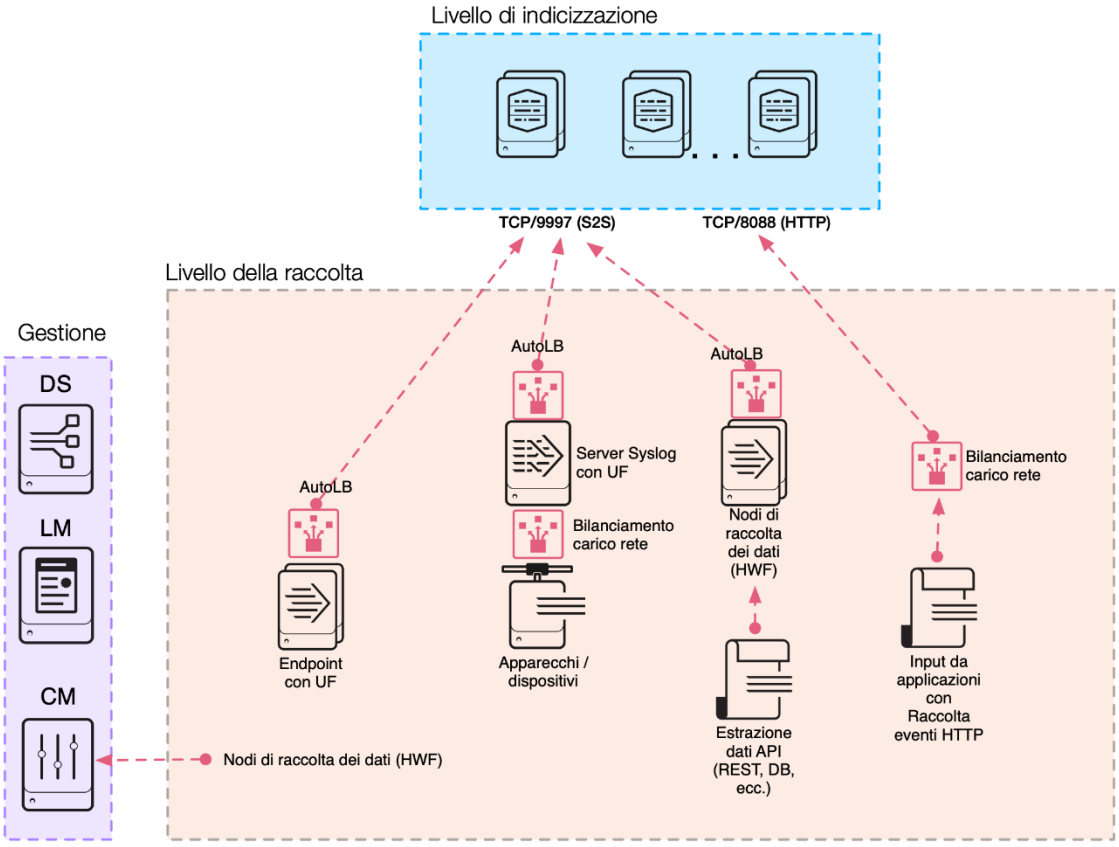
Idealmente, il livello di raccolta dei dati è per quanto possibile "flat", ossia le fonti di dati vengono raccolte in locale da uno universal forwarder e inoltrate direttamente al livello di indicizzazione. Si tratta di una best practice perché garantisce una latenza minima di inserimento dei dati (tempo necessario affinché siano disponibili per la ricerca) e consente una corretta distribuzione degli eventi tra gli indexer disponibili. Attenendosi a questa best practice si semplificano la gestione e l'operatività. Spesso vediamo che i clienti realizzano un livello di inoltro intermedio. In generale, occorre evitare questa situazione a meno che non sia l'unico modo per soddisfare i requisiti. Ala luce del possibile impatto dei forwarder intermedi, in questo documento è presente un'apposita sezione dedicata ad approfondire questo argomento.

Vi sono endpoint che non consentono l'installazione dello universal forwarder (in altre parole, dispositivi di rete, apparecchi) e che effettuano il log con il protocollo syslog. Un'architettura separata secondo le best practice per la raccolta di queste fonti di dati è illustrata nella sezione intitolata Raccolta di dati da Syslog.

Per le fonti di dati la cui raccolta deve avvenire per mezzo di programmi (API, accesso a database), si consiglia di implementare un nodo di raccolta dei dati (DCN) basato su un'installazione completa di Splunk. Questo è detto anche heavy forwarder. Non si consiglia di eseguire questo genere di input al livello delle search head in ambienti diversi da quello di sviluppo.

Il seguente diagramma mostra una generica architettura di raccolta dei dati che risponde a queste linee guida.

Presentazione generale della topologia di raccolta dei dati



Il diagramma che precede mostra il Deployment Server (DS) nel livello amministrazione, utilizzato per gestire le configurazioni sui componenti per la raccolta dei dati. È indicato anche il License Master (LM) perché i nodi di raccolta dei dati hanno necessità di accedere al LM per abilitare le funzionalità di Splunk Enterprise. Il master cluster (CM), se disponibile, può essere usato dai forwarder per l'indexer discovery, eliminando la necessità di gestire gli indexer disponibili nella configurazione di output dei forwarder.

Nel diagramma che precede, AutoLB rappresenta il meccanismo automatico di bilanciamento del carico integrato in Splunk. Questo meccanismo serve a garantire una corretta distribuzione degli eventi per i dati inviati utilizzando il protocollo proprietario di Splunk S2S (porta di default 9997). Nota: l'utilizzo di un bilanciatore del carico di rete esterno per il traffico S2S non è attualmente supportato né consigliato.

Per bilanciare il traffico proveniente dalle fonti di dati che comunicano con un protocollo industriale standard (come HTTP o syslog), si utilizza un bilanciatore del carico di rete che assicura una distribuzione uniforme del carico e degli eventi tra gli indexer nel livello di indicizzazione.

(UF) Universal Forwarder

Lo universal forwarder (UF) rappresenta la scelta migliore se vi è l'esigenza raccogliere grandi set di dati dai sistemi presenti nell'ambiente. Si tratta di un meccanismo di raccolta dei dati appositamente realizzato, con requisiti di risorse ridotti al minimo. Lo UF dovrebbe essere la scelta predefinita per raccogliere e inoltrare i dati di log. Lo UF offre:

- Funzione di checkpoint/riavvio per una raccolta dei dati senza perdite.
- Protocollo efficiente che riduce l'utilizzo della larghezza di banda della rete.
- Possibilità di limitazione.

- Bilanciamento del carico integrato tra gli indexer disponibili.
- Crittografia di rete opzionale con SSL/TLS.
- Compressione dei dati (da usare solo senza SSL/TLS).
- Più metodi di input (file, log degli eventi di Windows, input di rete, input da script).
- Limitate capacità di filtro degli eventi (solo log di eventi di Windows).
- Supporto a pipeline parallele di inserimento per aumentare la capacità/ridurre la latenza.

Con poche eccezioni per dati ben strutturati (json, csv, tsv), lo UF non esegue il parsing delle fonti di dati in eventi, pertanto non può eseguire alcuna azione che richieda una comprensione del formato del log. Inoltre è dotato di una versione ridotta di Python, che lo rende incompatibile con qualunque app di input modulare che necessiti per funzionare dell'installazione completa di Splunk.

È normale che vengano distribuiti numerosi UF (da 100 a 10.000) sugli endpoint e sui server in un ambiente Splunk e che vengano gestiti centralmente, attraverso un server del deployment Splunk, oppure uno strumento di gestione delle configurazioni di terze parti (come ad es. Puppet o Chef).

(HF) Heavy Forwarder

Lo heavyweight forwarder (HWF) è un deployment completo di Splunk Enterprise configurato per operare come forwarder con indicizzazione disabilitata. Un HWF non svolge solitamente altri ruoli. La differenza principale tra uno UF e un HWF è che il HWF contiene la pipeline di parsing completa, eseguendo le stesse funzioni di un indexer, senza però scrivere e indicizzare realmente gli eventi su disco. Questo consente al HWF di comprendere i singoli eventi e di rispondere ad essi, ad esempio per mascherare i dati oppure per applicare un filtro ed eseguire l'instradamento sulla base dei dati dell'evento. Poiché si tratta di un'installazione completa di Splunk Enterprise, può ospitare input modulari che necessitano di un'installazione completa di Python per funzionare correttamente per la raccolta dei dati oppure funzionare come endpoint per la Raccolta eventi HTTP di Splunk (HEC). Il HWF svolge le seguenti funzioni:

- Esegue il parsing in eventi.
- Filtra e instrada sulla base dei dati dei singoli eventi.
- Ha un peso sulle risorse maggiore rispetto allo UF.
- Ha un peso maggiore sulla larghezza di banda della rete rispetto allo UF (~5x).
- Ha un'interfaccia grafica per l'amministrazione.

Di norma, i HWF non vengono installati sugli endpoint ai fini della raccolta dei dati. Al contrario vengono usati su sistemi standalone per implementare nodi di raccolta dei dati (DCN) o livelli di inoltro intermedi. **Usare un HWF solo quando non è possibile soddisfare i requisiti di raccolta dei dati da altri sistemi con uno UF.**

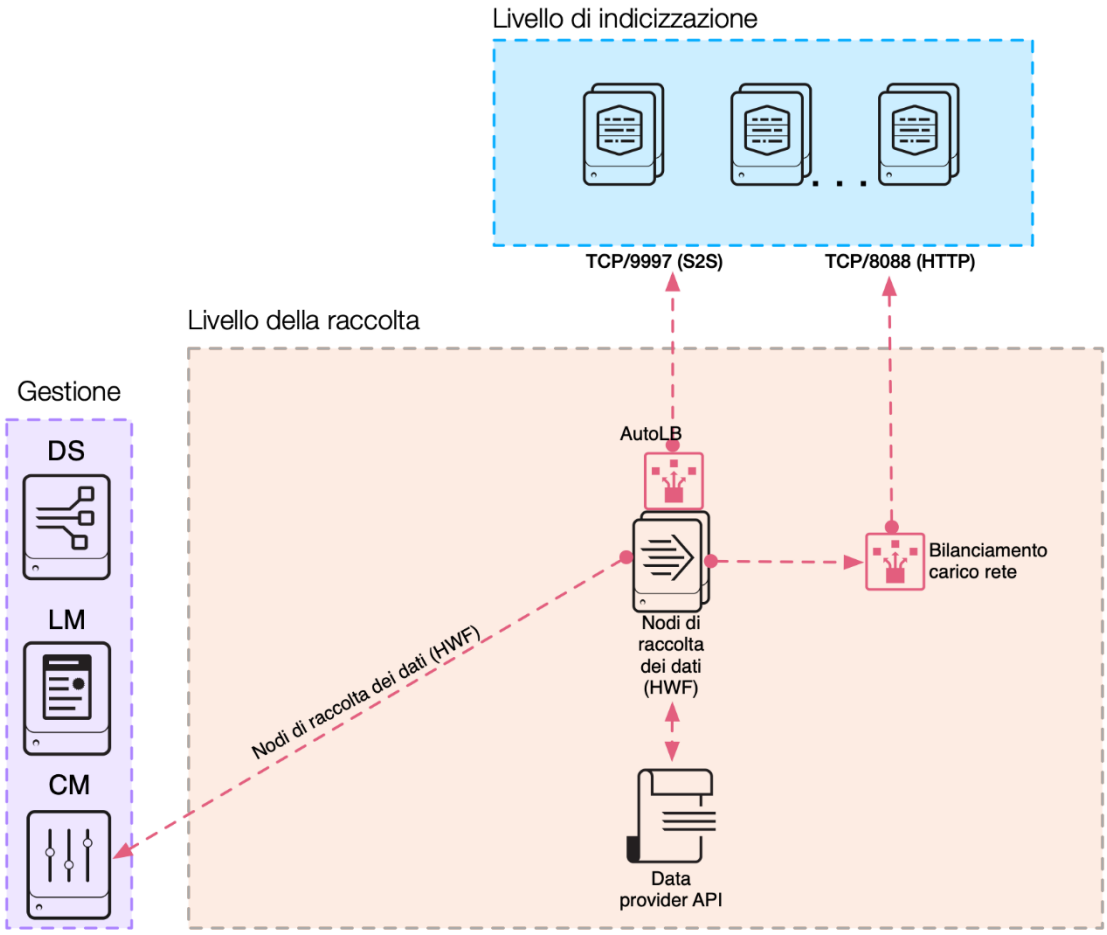
Esempi di requisiti di questo genere sono:

- Lettura dei dati da un RDBMS ai fini del loro inserimento in Splunk (input da database).
- Raccolta dei dati da sistemi raggiungibili tramite API (servizi cloud, monitoraggio VMWare, sistemi proprietari, ecc.).
- Messa a disposizione di un livello dedicato che ospiti il servizio di raccolta eventi HTTP.
- Implementazione di un livello di inoltro intermedio che richiede un parsing forwarder per l'instradamento/il filtro/il mascheramento.

(DCN) Heavy Forwarder come nodo di raccolta dei dati

Alcune fonti di dati richiedono una raccolta per mezzo di qualche forma di API. API di questo tipo possono essere REST, servizi web, JMS e/o JDBC come meccanismo di interrogazione. Splunk e sviluppatori terzi mettono a disposizione una vasta gamma di applicazioni che consentono a queste API di interagire. Il caso più frequente è che queste applicazioni vengono implementate utilizzando il framework di input modulare di Splunk, che per funzionare correttamente richiede un'installazione completa del software Splunk. La best practice per questo caso d'uso è implementare uno o più server che funzionando come heavy forwarder configurati per operare come nodo di raccolta dei dati (DCN).

Topologia dei nodi di raccolta dei dati

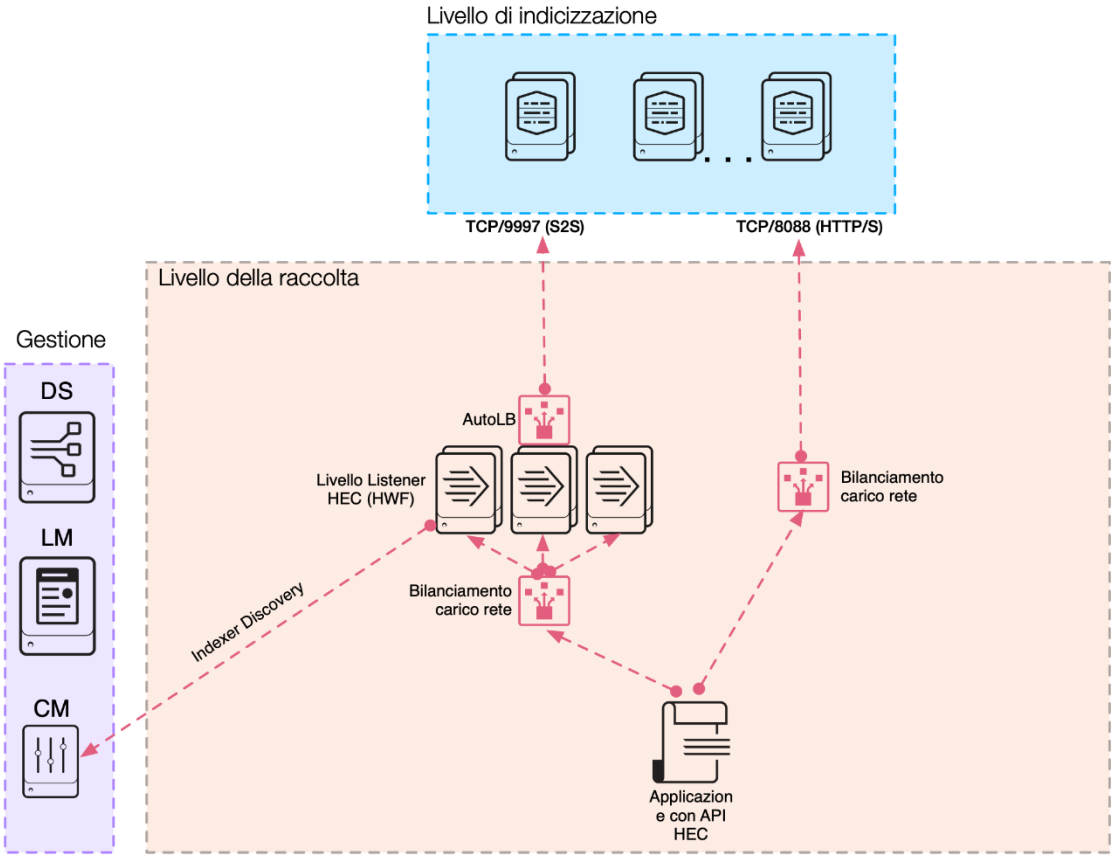


Raccolta eventi HTTP (HEC)

HEC mette a disposizione un servizio listener che accetta connessioni HTTP/S lato server, e API lato client, che consentono alle applicazioni di pubblicare i dati utili di log direttamente nel livello di indicizzazione o su un livello dedicato di ricezione HEC composto da uno o più heavy forwarder. HEC offre due endpoint che supportano la trasmissione dei dati in formato grezzo e in formato JSON. Con JSON è possibile inserire anche altri metadati nel payload dell'evento, che possono contribire a una maggiore flessibilità nella ricerca all'interno dei dati in un momento successivo.

Il diagramma che segue illustra le due opzioni di deployment per HEC:

Scelte di topologia HEC



Il livello di amministrazione contiene il license master (richiesto da HF) e il deployment server per gestire gli input HTTP sui componenti in ascolto. Nota: se il livello di indicizzazione si trova in cluster e riceve il traffico HEC direttamente, la configurazione HEC è gestita attraverso il master cluster anziché il deployment server.

La decisione su quale topologia del deployment scegliere dipende in gran parte dalle esigenze specifiche. Un livello di listener HEC dedicato introduce un altro componente architetturale nel deployment. Il lato positivo è che può essere scalato in modo indipendente e offre un livello di isolamento dal livello di indicizzazione dal punto di vista della gestione. Inoltre, poiché il livello HEC dedicato richiede un HF, esso eseguirà il parsing di tutto il traffico in entrata, sgravando parte del carico di lavoro dall'indexer.

D'altra parte, l'hosting del listener HEC direttamente sull'indexer garantisce probabilmente una migliore distribuzione degli eventi nel livello di indicizzazione, perché HTTP è un protocollo ben compreso da tutti i bilanciatori del carico di rete e opportuni criteri di bilanciamento del carico aiutano ad assicurare che gli indexer meno impegnati vengano serviti per primi.

Nell'ottica di realizzare l'architettura più semplice possibile che soddisfi i propri requisiti, consigliamo di considerare l'hosting del listener HEC sugli indexer, presumendo che si disponga di sufficiente capacità di sistema per farlo. Questa decisione può essere facilmente invertita in un momento successivo qualora sorga l'esigenza, semplicemente con il deployment di un livello HF opportunamente dimensionato e configurato, e modificando la configurazione di LB per utilizzare gli indirizzi IP del HF anziché dell'indexer. Questa modifica dovrebbe risultare trasparente per le applicazioni client.

Nota: se serve la conferma indexer dei dati inviati tramite HEC, è consigliato un livello di listener HEC dedicato per contenere al minimo la duplicazione dei messaggi causata dai riavvii a rotazione degli indexer.

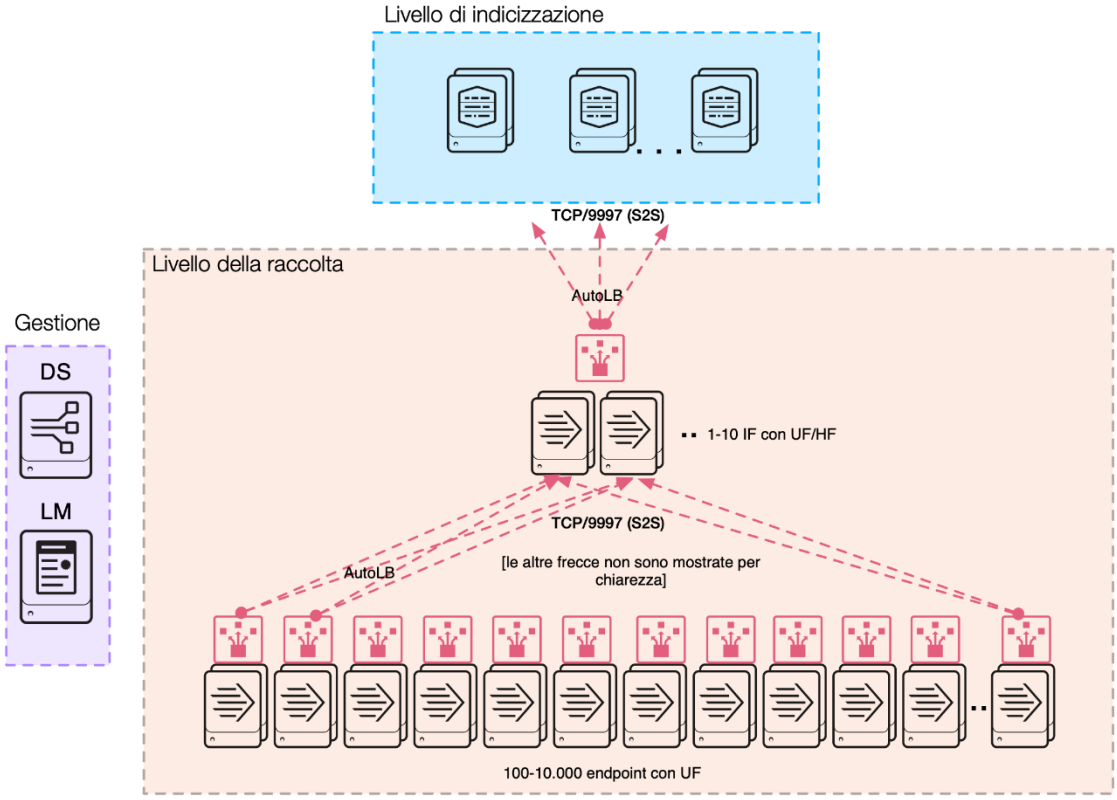
Nota: questa architettura del deployment HEC mette a disposizione il trasporto per alcuni degli altri componenti per la raccolta dei dati presentati più avanti, in particolare per la raccolta dei dati Syslog e di metriche.

(IF) Livello di inoltro intermedio

In alcune situazioni si rendono necessari dei forwarder intermedi per l'inoltro dei dati. I forwarder intermedi ricevono i flussi di log dagli endpoint e li inoltrano al livello indexer. I forwarder intermedi introducono delle difficoltà architetturali che esigono un'attenta progettazione per evitare ripercussioni negative sull'intero ambiente Splunk. L'aspetto più importante è che i forwarder intermedi concentrano i collegamenti da un numero compreso tra 100 e 10.000 forwarder terminali, inoltrando i dati agli indexer usando un numero di collegamenti notevolmente inferiore. Questo può influire sulla distribuzione dei dati nel livello di indicizzazione, perché solo un sottoinsieme di indexer riceve il traffico in un dato momento. Tuttavia, questi effetti negativi possono essere contenuti prevedendo correttamente il dimensionamento e la configurazione.

Il diagramma che segue illustra bene questa difficoltà:

Topologia inoltro intermedio



In uno scenario con un solo forwarder intermedio, tutti gli endpoint si collegano a questo unico forwarder (potenzialmente anche migliaia), e il forwarder intermedio si collega a turno a un solo indexer per volta. Non si tratta di uno scenario ottimale perché è probabile che si verifichino le seguenti conseguenze:

- Un grande flusso di dati da molti endpoint viene fatto passare attraverso un imbuto che esaurisce le risorse di sistema e di rete.
- Target di failover limitati per gli endpoint in caso di guasto del forwarder intermedio (il rischio dovuto al guasto è inversamente proporzionale al numero di forwarder intermedi).
- Numero limitato di indexer serviti in un dato momento. Le ricerche su brevi periodi di tempo non beneficeranno della parallelizzazione come potrebbero.

Inoltre, i forwarder intermedi aggiungono un ulteriore livello architetturale al deployment, che può complicare la gestione e la risoluzione dei problemi e aggiungere latenza al percorso di inserimento di dati. È consigliabile provare a evitare l'uso di livelli di inoltro intermedi, a meno

che non sia l'unica opzione per soddisfare i propri requisiti. Si può considerare l'uso di un livello intermedio se si hanno:

- dati sensibili che occorre offuscare/rimuovere prima di inviare agli indexer attraverso la rete. Un esempio è quando occorre usare una rete pubblica.
- Rigidi criteri di sicurezza non consentono un collegamento diretto tra gli endpoint e gli indexer, come ad esempio reti multizona o indexer basati su cloud.
- I limiti di larghezza di banda tra gli endpoint e gli indexer impongono di filtrare un sottoinsieme significativo di eventi.
- È richiesto l'instradamento a target dinamici in funzione degli eventi.

Considerare le esigenze di dimensionamento e configurazione per ogni livello di inoltro intermedio, per accertarsi della disponibilità di questo livello, prevedere una capacità di elaborazione sufficiente a gestire tutto il traffico, e supportare una buona distribuzione degli eventi tra gli indexer. Il livello IF presenta i seguenti requisiti:

- Numero sufficiente di pipeline per l'elaborazione dei dati in generale.
- Infrastruttura IF ridondante.
- Configurazione di bilanciamento del carico di Splunk opportunamente calibrata. Ad esempio, autoLBVolume, EVENT_BREAKER, EVENT_BREAKER_ENABLE, possibilmente forceTimeBasedAutoLB se necessario.

Le indicazioni generali suggeriscono di avere un numero di pipeline di elaborazione dei forwarder intermedi pari a due volte il numero di indexer nel livello di indicizzazione.

Nota: una pipeline di elaborazione non corrisponde a un server IF fisico. Disponibilità di sufficienti risorse di sistema. Sono ad esempio disponibili, CPU core, memoria e larghezza di banda sulle schede NIC, un unico IF può essere configurato con più pipeline di elaborazione.

Se serve un livello IF (vedere questionario), prevedere come scelta predefinita l'uso di UF per il livello, perché offrono una capacità superiore con un consumo inferiore di risorse, sia per il sistema sia per la rete. Usare un HF se le capacità dello UF non soddisfano i requisiti.

(SYSLOG) Raccolta di dati da Syslog

Il protocollo Syslog offre una fonte diffusa di dati di log in ambiente aziendale. I livelli più scalabili e affidabili di raccolta dei dati contengono un componente di inserimento syslog. Esistono molti modi per inserire in Splunk i dati di syslog. Considerare i seguenti metodi:

- **Universal forwarder (UF)/heavy forwarder (HF)**: Usare uno Splunk UF o HF per monitorare (inserire) file scritti da un server Syslog (ad es. `rsyslog` o `syslog-ng`).
- **Syslog Agent verso HEC**: Usare un syslog agent in grado di emettere l'output verso Splunk HEC. (esistono moduli di terze parti per rsyslog e syslog-ng in grado di indirizzare l'output verso HEC). .
- **Input diretto da TCP/UDP**: Splunk è in grado di ascoltare una porta TCP o UDP (la porta di default è UDP 514) e acquisire le fonti da qui (**non** consigliato in produzione).

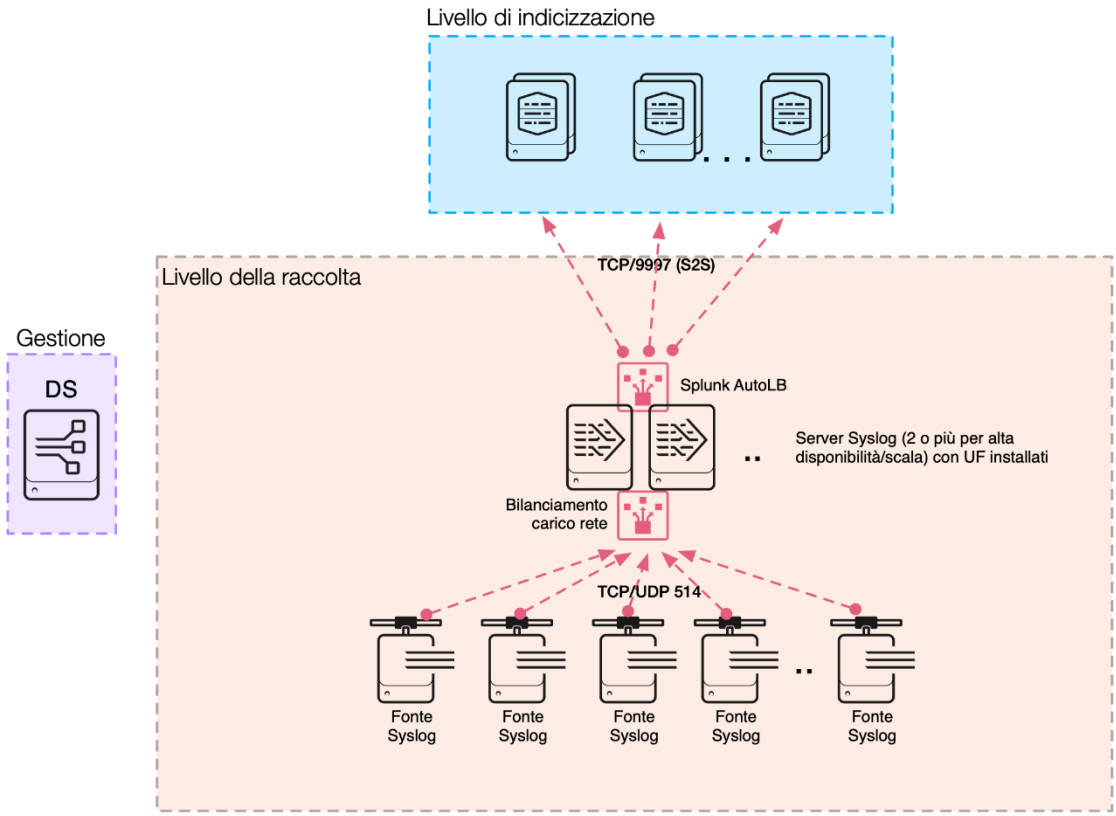
Syslog (monitoraggio file insieme a un SCD)

Splunk può usare il monitoraggio, con `input.conf`, su uno UF/HF per elaborare e inserire le fonti syslog che vengono scritte su disco su un endpoint, utilizzando un syslog collection daemon (SCD). Soluzioni molto comuni, `rsyslog`, `syslog-ng` e anche Fastvue offrono soluzioni commerciali e gratuite scalabili e semplici da integrare e gestire sia in ambienti a bassi volumi sia in ambienti distribuiti su grande scala.

Per saperne di più su come configurare i monitor, vedere Monitora file e cartelle in *Caricamento dei dati.*

Questa architettura supporta un corretto onboarding dei dati nello stesso modo di uno universal forwarder su qualunque altro endpoint. Si può configurare il SCD per individuare molteplici tipi differenti di log e scrivere gli eventi di log su opportuni file e opportune cartelle da cui un forward Splunk possa prelevarli. Scrivendo gli eventi su disco, si aggiunge anche un livello di persistenza al flusso di log di syslog, elemento in grado di limitare l'esposizione alla perdita di dati per i messaggi trasmessi utilizzando un protocollo UDP non affidabile per il trasporto.

Topologia della raccolta di dati da Syslog con UF



Il diagramma mostra le fonti syslog che inviano dati con TCP o UDP sulla porta 514 verso un pool di server Syslog con bilanciamento di carico. Più server garantiscono un'elevata disponibilità per il livello della raccolta e possono prevenire la perdita di dati durante le operazioni di manutenzione. Ogni server Syslog è configurato per applicare al flusso syslog delle regole che producono la scrittura di eventi syslog su file/cartelle dedicati per ciascuna fonte (eventi firewall, syslog sistema operativo, switch di rete, IPS, ecc.). Lo UF installato per ogni server monitora i file e inoltra i dati al livello di indicizzazione affinché vengano elaborati nell'apposito indice. Splunk AutoLB serve per distribuire i dati in modo omogeneo tra gli indexer disponibili.

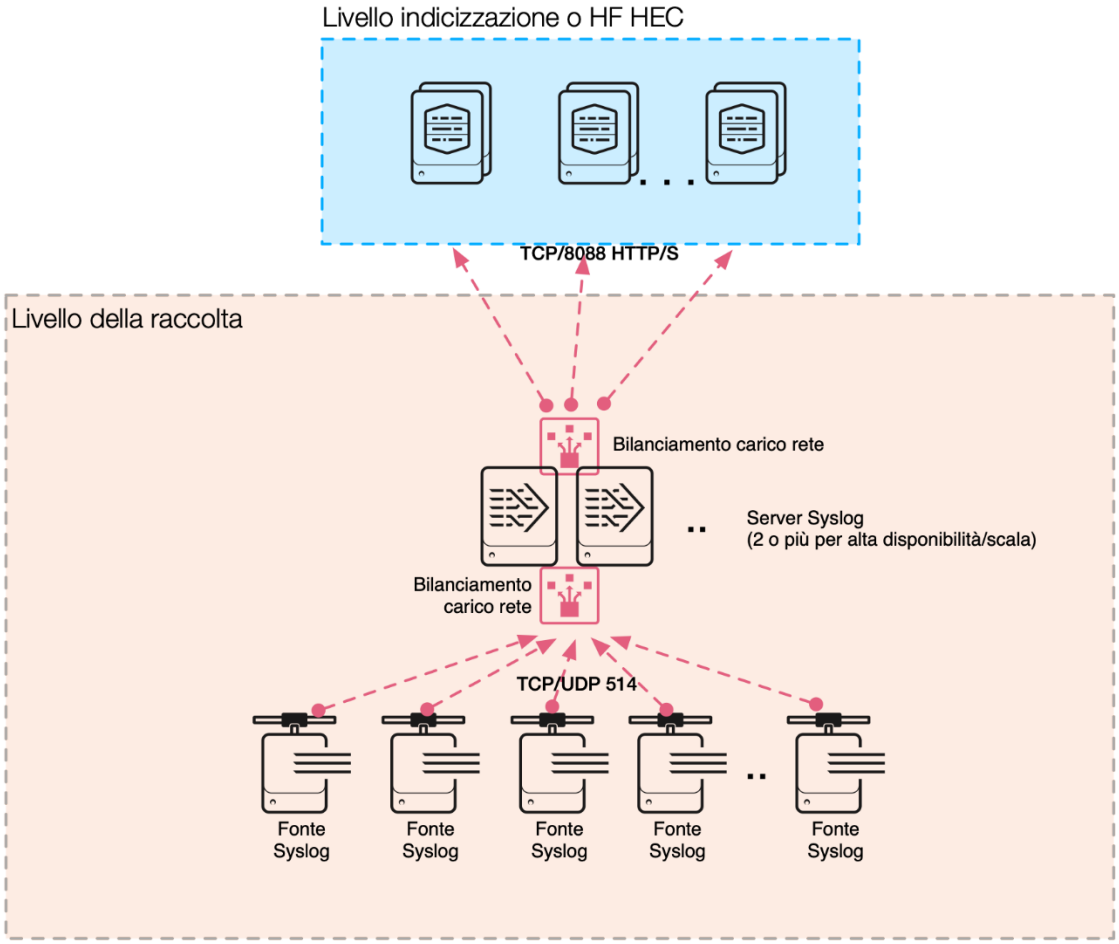
Il deployment server mostrato nel livello di amministrazione può essere usato per gestire a livello centrale la configurazione UF

Syslog Agent verso HEC

Alla luce della crescente diffusione di HEC, aumenta il numero di deployment che utilizzano il proprio deployment di HEC per l'inserimento dei dati syslog. Per maggior informazioni, vedere il post sui blog di Splunk Syslog-ng and HEC: Scalable Aggregated Data Collection in Splunk.

Il diagramma che segue mostra le fonti di invio dei dati sulla porta 514 usando un bilanciatore del carico di rete verso una server farm syslog. Vengono applicati opportuni criteri di syslog con una destinazione syslog personalizzata, uno script python che utilizza HEC API, e gli eventi sono inviati a un listener HEC, anche con un bilanciatore del carico di rete per l'indicizzazione:

Topologia della raccolta di dati da Syslog con HEC



Un beneficio di questa topologia è la possibilità di eliminare l'esigenza di distribuire e configurare UF/HF. Il bilanciatore di carico HTTP serve i listener HEC sull'indexer (o un livello di listener HEC dedicato) per assicurarsi che i dati vengano distribuiti in modo uniforme tra i diversi HEC. Configurazione del bilanciatore del carico almeno con i criteri "Collegamenti minori".

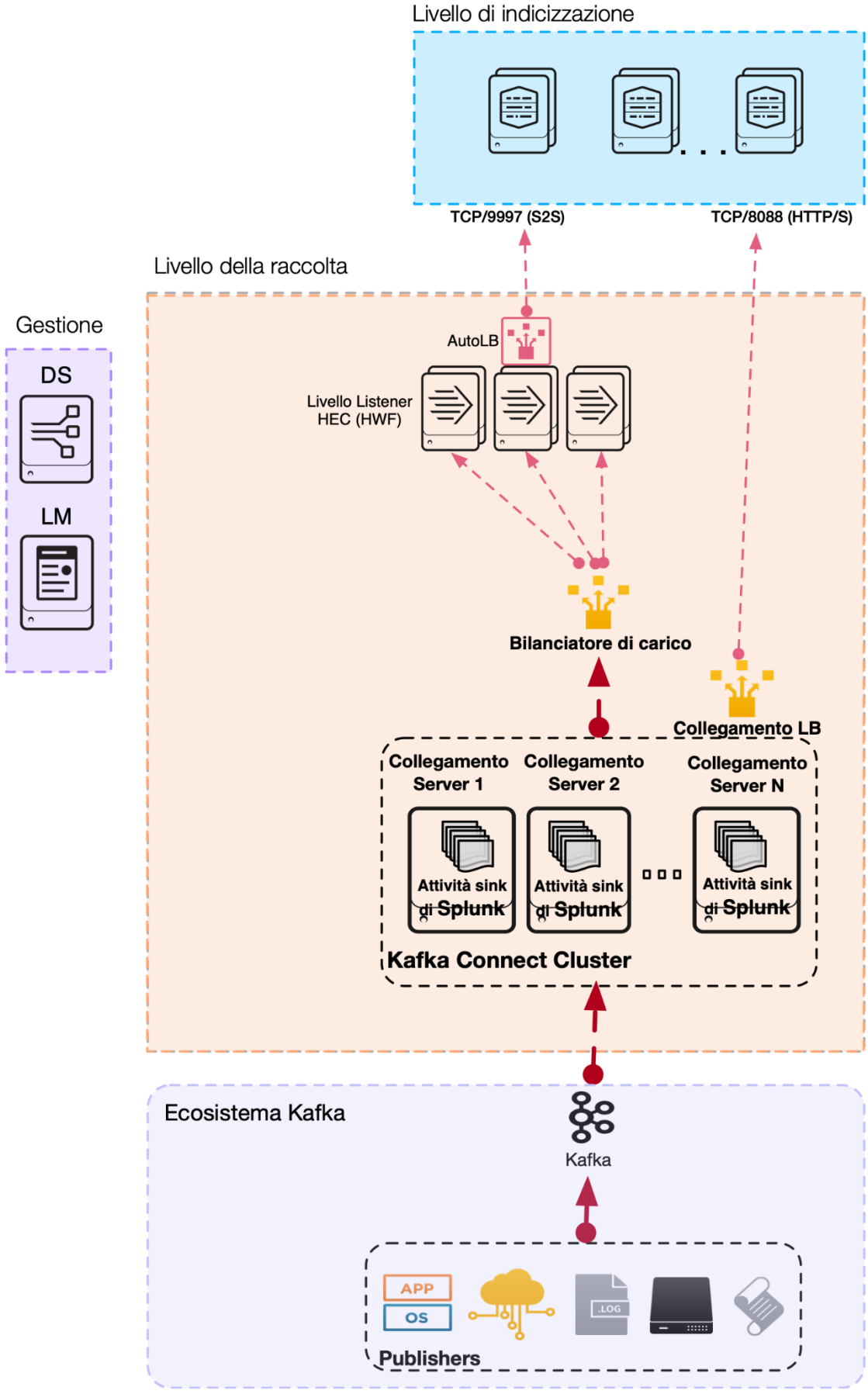
Input Splunk UDP

Splunk può utilizzare un input diretto da UDP su uno UF o HF per ricevere i dati da syslog. Per scoprire come configurare le porte TCP e UDP, vedere Ricevere dati dalle porte TCP e UDP in *Caricamento dei dati*. La possibilità di ricevere eventi sulla porta UDP 514 dipende dalla capacità dello UF/HF di essere eseguito come root. Inoltre, per evitare la possibile perdita di dati l'agente deve essere disponibile il 100% del tempo. I forwarder possono essere riavviati di frequente per l'applicazione di modifiche alla configurazione, evento che causa quasi certamente la perdita di dati. Per questi motivi**, non è considerata una best practice per un deployment di produzione**.

(KAFKA) Elaborazione di dati di log da Kafka Topics

Splunk offre un sink connector supportato per elaborare i dati da Kafka Topics, detto "Splunk Connect per Kafka". Vedere Apache Kafka Connect nel manuale di Splunk Connect per Kafka per una documentazione approfondita sul prodotto. Il pacchetto Splunk Connect per Kafka è installato in un cluster Kafka Connect opportunamente dimensionato (fuori da Splunk), dove può iscriversi ai topics configurati, e inviare gli eventi "consumati" da indicizzare utilizzando HEC:

Topologia di raccolta dei dati con Kafka e HEC

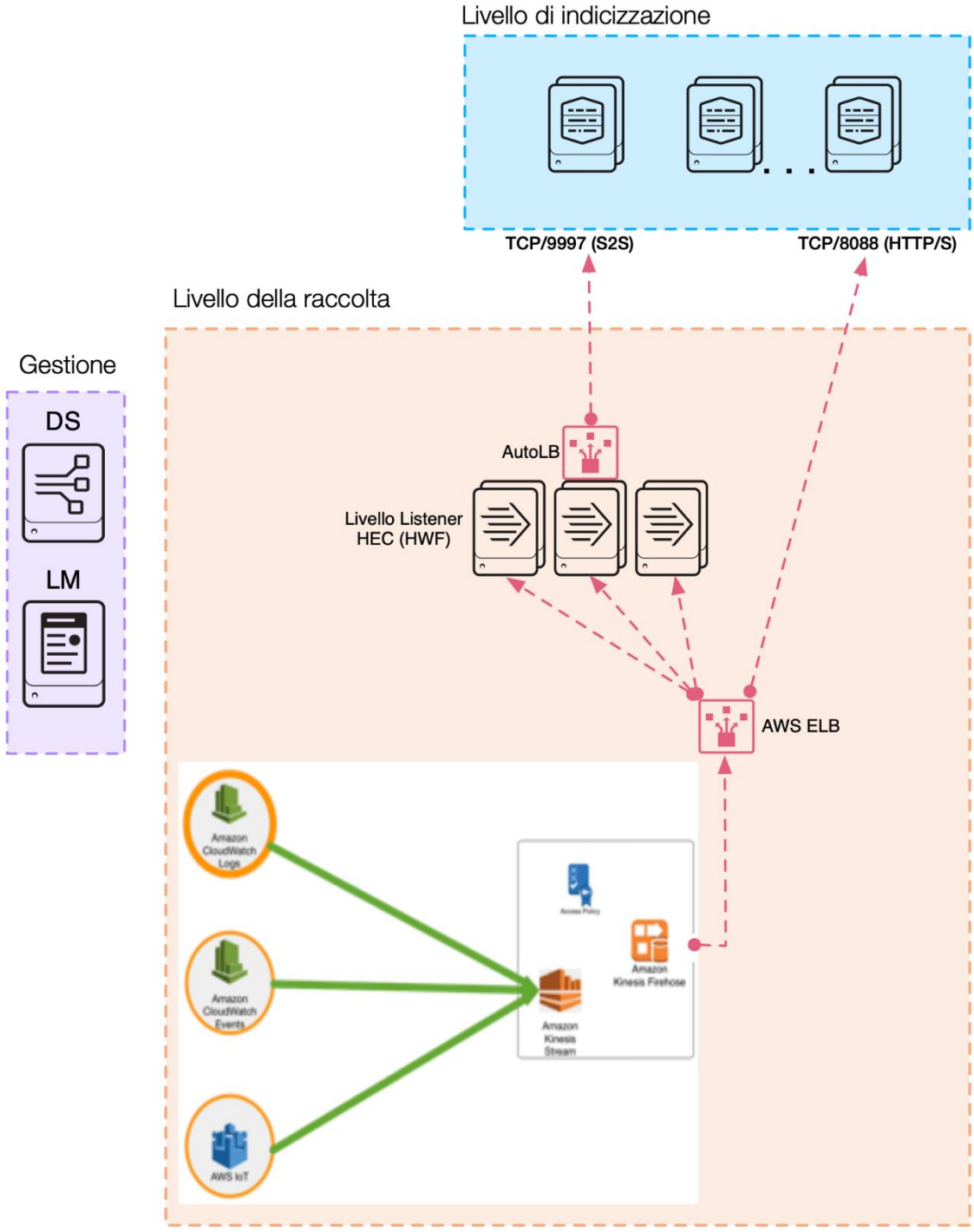


Il diagramma mostra i Kafka Publishers che inviano messaggi ai bus Kafka. Le attività che si trovano nel cluster Kafka Connect “consumano” questi messaggi tramite Splunk Connect per Kafka e inviano i dati al servizio di ascolto HEC utilizzando un bilanciatore del carico di rete. Ancora una volta, il servizio di ascolto HEC può trovarsi direttamente sull’indexer, oppure su un livello di ascolto HEC dedicato. Per maggiori dettagli si rimanda alla sezione dedicata a HEC. I componenti del livello di amministrazione sono necessari solo se viene realizzato un livello HF dedicato per l’hosting dei listener HEC.

(KINESIS) Elaborazione di dati di log da Amazon Kinesis Firehose

Splunk e Amazon hanno realizzato un’integrazione tra Kinesis e Splunk HEC che consente lo streaming di dati da AWS direttamente verso un endpoint HEC, configurabile dalla console AWS. A questo si affianca il componente aggiuntivo Splunk per Kinesis Firehose che mette a disposizione knowledge CIM-compliant per varie fonti di dati generate in AWS.

Topologia di raccolta dei dati con Amazon Kinesis



Il diagramma mostra le fonti di log AWS inviate per mezzo di un flusso Kinesis a Firehose, che (con un'opportuna configurazione), invierà i dati al servizio di ascolto HEC attraverso un AWS ELB. Ancora una volta, il servizio di ascolto HEC può trovarsi direttamente sull'indexer, oppure su un livello di ascolto HEC dedicato. Per maggiori dettagli si rimanda alla sezione dedicata a HEC.

I componenti del livello di amministrazione mostrati sono necessari solo se viene realizzato un livello HF dedicato per l'hosting dei listener HEC.

(METRICHE) Raccolta di metriche

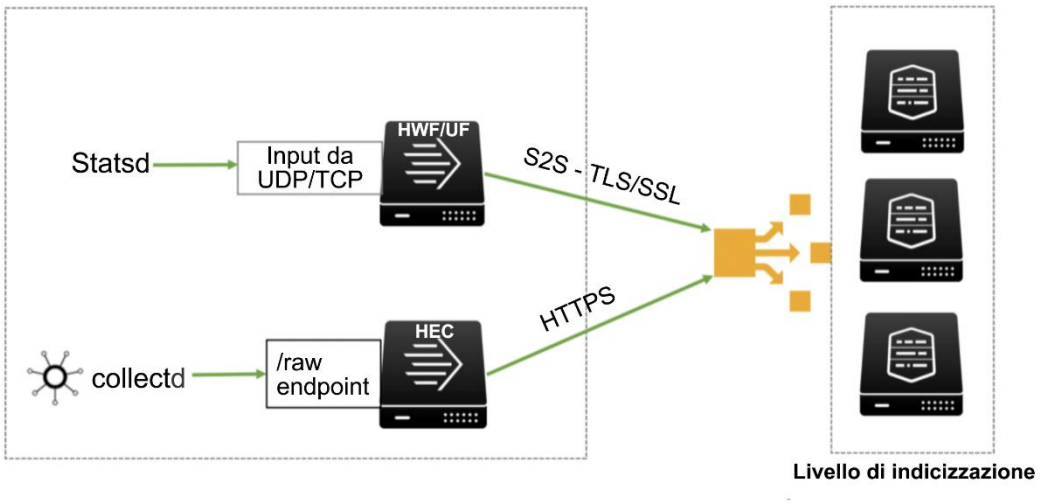
Splunk è in grado di ricevere e raccogliere da diversi software di terze parti dati sulle prestazioni dei sistemi e delle applicazioni, oppure dati metrici. Le metriche nella piattaforma Splunk utilizzano un tipo di indice speciale, ottimizzato per la memorizzazione e il recupero di dati metrici.

Sono disponibili diverse modalità per “consumare” dati metrici e il metodo di raccolta dipende dalla tecnologia impiegata. La forma più comune di raccolta di metriche assume la forma di un daemon software, ad es. **collectd**, **statsd,** oppure di un file personalizzato per i dati metrici e una configurazione valida per la fonte di dati.

Esistono principalmente due metodi per introdurre le metriche in Splunk usando agenti come **statsd** e **collectd**: o tramite un ***input TCP/UDP diretto*** o tramite **HEC**.

L’uso di **HEC** è considerato una best practice grazie alla resilienza e scalabilità degli endpoint **HEC**, e alla capacità di scalare in orizzontale in modo facile il livello della raccolta.

Topologia di raccolta dei dati metrici



Statsd supporta attualmente il trasporto via **UDP** e **TCP**, che è possibile sfruttare per ottenere un input diretto su un forwarder di Splunk, o su un indexer. Non è però considerato una best practice inviare traffico TCP/UDP direttamente al forwarder in produzione perché l’architettura non è flessibile ed è soggetta alla perdita di eventi (vedere Raccolta di dati Syslog) in conseguenza dei riavvii necessari dei forwarder di Splunk.

(HA) Considerazioni per componenti del livello di inoltro ad elevata disponibilità

Il mondo digitale condivide un concetto comune di elevata disponibilità (HA). A seconda dell’organizzazione, però, il significato può variare, tendendo ad avvicinarsi più a un ripristino di emergenza (disaster recovery – DR) piuttosto che all’elevata disponibilità. Questi due concetti, benché simili, hanno in realtà significati differenti. L’elevata disponibilità è una caratteristica di un sistema finalizzata a garantire un livello concordato di prestazioni operative, solitamente in termini di uptime, per un periodo superiore alla norma. Il disaster recovery prevede un set di criteri, strumenti e procedure che consentono il ripristino o la prosecuzione delle infrastrutture e dei sistemi tecnologici essenziali a seguito di un evento catastrofico.

Di seguito vengono presentate diverse forme di elevata disponibilità al livello intermedio/aggregato:

Livello intermedio

- Per i clienti che hanno deployment con livello intermedio/aggregato, l'elevata disponibilità dei forwarder è essenziale. A livello di applicazione, attualmente Splunk non ha un supporto nativo per l'elevata disponibilità. Esistono altre strategie per prevedere l'elevata disponibilità al livello di sistema operativo, che non sono però native in Splunk. Soluzioni comuni sono VMWare VMotion, AWS Autoscaling Groups, e Linux Clustering. Si consiglia di consultarsi con il progettista Splunk per esaminare le opzioni di progettazione a disposizione.
- Per gli ambienti con requisiti di elevata disponibilità per un livello HEC dedicato, la best practice prevede di usare un bilanciatore del carico di rete (NTLB), come NGINX, davanti a più Splunk heavy forwarder. Questo presenta il vantaggio di combinare il massimo in termini di capacità, scala e disponibilità. È disponibile un pool dedicato di istanze della Raccolta eventi HTTP il cui unico compito è ricevere e inoltrare i dati. È possibile aggiungere più istanze HEC senza dover necessariamente aggiungere altri indexer. Se l'indexer diventa un collo di bottiglia, si può aggiungere un altro indexer.
- Per gli ambienti con requisiti di elevata disponibilità per la raccolta di dati syslog, la best practice prevede di usare più server Syslog serviti da un indirizzo IP cluster (virtuale) in hosting su una soluzione di bilanciamento del carico, come HAProxy o F5, per offrire il massimo in termini di capacità, scala e disponibilità. È disponibile un pool dedicato di istanze Splunk il cui unico compito è ricevere e inoltrare i dati. È possibile aggiungere più istanze senza dover necessariamente aggiungere altri indexer. Se gli indexer diventano un collo di bottiglia, si possono altri indexer.

Livello di inoltro

- Al livello di inoltro (endpoint), l'elevata disponibilità per l'agente stesso dipende dal sistema operativo sottostante. Come minimo, occorre garantire che qualunque servizio che implementi la funzione di inoltro venga riavviato automaticamente quando si riavvia il sistema operativo host. Al di là di questo, la best practice per i forwarder prevedrebbe la configurazione e il corretto uso di AutoLB dal forwarder verso più indexer. Questo comporta anche l'uso delle conferme indexer per accertarsi che i dati arrivino al livello di indicizzazione.

Passaggio 3: applicazione di principi e best practice di progettazione

Di seguito vengono presentati i principi e le best practice di progettazione, separatamente per livello di deployment.

Livelli di deployment

I principi di progettazione delle SVA coprono tutti i seguenti livelli di deployment:

Livello	Definizione
Ricerca	• Search head
Indicizzazione	• Indexer
Raccolta	• Forwarder • Input modulari • Rete • HEC (Raccolta eventi HTTP) • ecc.
Amministrazione / Utility	• CM • DS • LM • DMS • SHC-D

Adeguare la topologia alle best practice

Per scegliere i principi e le best practice di progettazione adeguati al proprio deployment occorre tenere a mente i propri requisiti e la propria topologia. È opportuno quindi considerare le best practice solo dopo aver completato i passaggi 1 e 2 del processo di selezione delle Architetture convalidate Splunk illustrati sopra.

Best practice: raccomandazioni specifiche per livello

Di seguito vengono presentate delle raccomandazioni sui principi e sulle best practice di progettazione per ogni livello di deployment. Ogni principio di progettazione consolida una o più delle colonne portanti delle SVA: disponibilità, prestazioni, scalabilità, sicurezza, e gestibilità.

Raccomandazioni per il livello di ricerca

Principi di progettazione / best practice (i propri requisiti determinano quale delle prassi seguire)		**COLONNE PORTANTI DELLE SVA**				
		DISPONIBILITÀ	**PRESTAZIONI**	**SCALABILITÀ**	**SICUREZZA**	**GESTIBILITÀ**
1	Tenere il livello di ricerca vicino (in termini di rete) al livello di indicizzazione *Ogni ritardo di rete tra il livello di ricerca e quello di indicizzazione avrà un impatto diretto sulle prestazioni di ricerca*		✓			
2	Evitare più search head indipendenti *Le search head indipendenti non consentono di condividere gli artefatti Splunk creati dagli utenti. Inoltre non scalano bene rispetto all'utilizzo delle risorse nel livello di ricerca. A meno che non vi sia la specifica esigenza di avere ambienti isolati per le search head, per scalare è disponibile un'opzione migliore.*	✓		✓	✓	✓
3	Per scalare il livello di ricerca, sfruttare il clustering di search head *Un cluster di search head replica gli artefatti degli utenti nel cluster e*	✓		✓		

	consente una pianificazione intelligente del carico di lavoro delle ricerche tra tutti i membri del cluster. Offre inoltre una soluzione a elevata disponibilità.					
4	Inoltrare tutti i log interni delle search head al livello di indicizzazione *Tutti i dati indicizzati devono essere archiviati nel solo livello di indicizzazione. Questo elimina la necessità di prevedere elevate prestazioni della capacità di archiviazione sul livello della search head, semplificando l'amministrazione. Nota: questo vale anche per qualunque altro ruolo di Splunk.*		✓			✓
5	Considerare l'uso di LDAP auth ogniqualvolta possibile *La gestione centralizzata delle identità degli utenti ai fini dell'autenticazione è una best practice aziendale generale, semplifica l'amministrazione del deployment Splunk e migliora la sicurezza.*				✓	✓
6	Assicurare un numero di core sufficienti a coprire le esigenze di ricerca concomitante *Per essere eseguita, ogni ricerca richiede un core della CPU. Se non sono disponibili core per eseguire una ricerca, questa verrà messa in coda, con conseguente ritardo per l'utente. Nota: vale anche per il livello di indicizzazione.*	✓	✓	✓		
7	Usare finestre temporali programmate per le ricerche quando possibile / uniformare il		✓	✓		

	carico delle ricerche pianificate *Spesso le ricerche pianificate vengono eseguite in momenti precisi (allo scoccare dell'ora, al minuto 5/15/30 dell'ora, a mezzanotte). Prevedere una finestra temporale all'interno della quale eseguire la ricerca può aiutare a evitare l'accumularsi di ricerche concomitanti.*					
9	Limitare il numero di diversi cluster di search head in modo da non sovraccaricare il livello di indicizzazione *Il carico di lavoro delle ricerche può essere governato solo automaticamente all'interno di un ambiente SH. I cluster di search head (SHC) indipendenti possono creare un carico di lavoro delle ricerche concorrenti maggiore di quello che è in grado di gestire il livello dell'indexer (destinazione di ricerca). Lo stesso vale per l'attenta pianificazione del numero di search head standalone.*	✓		✓		
10	Quando si realizza un cluster di search head, usare un numero dispari di nodi (3, 5, 7, ecc.) *La scelta del gestore SHC avviene con un protocollo basato sulla maggioranza. Un numero dispari di nodi garantisce che un SHC non possa mai essere diviso in un numero pari di nodi durante un guasto della rete.*	✓				✓

Raccomandazioni per il livello di indicizzazione

Principi di progettazione / best practice (i propri requisiti determinano quale delle prassi seguire)		**COLONNE FONDANTI**				
		DISPONIBILITÀ	**PRESTAZIONI**	**SCALABILITÀ**	**SICUREZZA**	**GESTIBILITÀ**
1	Consentire pipeline parallele sui server capaci per *Le funzioni di parallelizzazione consentono di sfruttare le risorse sistema disponibili che altrimenti sarebbero inutilizzate. Si osservi che le prestazioni I/O devono essere adeguate prima di abilitare le funzioni di parallelizzazione dell'inserimento.*		✓	✓		
2	Considerare l'uso di SSD per volumi HOT/WARM e di sintesi (Summaries) *Gli SSD hanno raggiunto prezzi economici ed eliminano ogni possibile limitazione di I/O che spesso è la causa di insoddisfacenti prestazioni di ricerca.*		✓			
3	Tenere il livello di indicizzazione vicino (in termini di rete) al livello di ricerca *La minore latenza possibile avrà un effetto positivo sull'esperienza dell'utente durante la ricerca.*		✓			
4	Usare la replica degli indici quando è richiesta un'elevata disponibilità dei dati storici / report *La replica degli indici assicura la presenza di più copie di ogni evento nel cluster, come*	✓				

	protezione da una possibile indisponibilità delle destinazioni di ricerca. Adeguare il numero di copie (fattore di replica) agli accordi sui livelli di servizio (SLA).					
5	Assicurare una buona qualità dell'onboarding dei dati (ad es. corretta ed esplicita definizione, per ogni fonte di dati, di interruzioni di riga, estrazione dei timestamp, TZ, e fonte, tipo di fonte, host) e stabilire un monitoraggio costante mediante la console di monitoraggio *È dimostrato che configurare esplicitamente le fonti di dati piuttosto che affidarsi alle capacità di rilevazione automatica di Splunk ha benefici significativi sulla capacità di inserimento di dati e sulla latenza di indicizzazione, specialmente in deployment con elevati volumi.*		✓	✓		✓
6	Valutare la configurazione della parallelizzazione della ricerca in modalità batch sugli indexer con potenza di elaborazione eccedente *Sfruttare le funzioni di parallelizzazione della ricerca può avere un impatto significativo sulle prestazioni di ricerca per alcuni tipi di ricerche e consente di utilizzare le risorse sistema che altrimenti sarebbero inutilizzate*		✓	✓		
7	Monitorare l'equilibrata distribuzione dei dati tra i nodi indexer (=destinazioni di ricerca).		✓	✓		✓

	Una distribuzione omogenea degli eventi/dei dati tra le destinazioni di ricerca è un fattore essenziale che contribuisce alle prestazioni di ricerca e alla corretta applicazione dei criteri di conservazione dei dati.					
8	Disabilitare la UI web sull'indexer nei deployment distribuiti/in cluster . *Non vi è motivo ragionevole per dover accedere alla WebUI direttamente sull'indexer.*		✅		✅	✅
9	Considerare i componenti aggiuntivi (add-on) tecnologici preconfigurati di Splunk per le fonti di dati conosciute *Per assicurare una buona qualità dell'onboarding dei dati da fonti ben conosciute, i componenti aggiuntivi tecnologici messi a disposizione da Splunk possono ridurre il time to value e assicurare un'implementazione ottimale, evitando di dover realizzare la propria configurazione.*		✅			✅
10	Monitorare le metriche critiche degli indexer *Splunk offre una console di monitoraggio che mette a disposizione metriche importanti sulle prestazioni del livello di indicizzazione. Esse comprendono l'utilizzo della CPU e della memoria, nonché metriche dettagliate sui componenti interni di Splunk (processi, pipeline, code, ricerche).*	✅	✅			

Raccomandazioni per il livello della raccolta

	Principi di progettazione / best practice (i propri requisiti determinano quale delle prassi seguire)	**COLONNE FONDANTI**				
		DISPONIBILITÀ	**PRESTAZIONI**	**SCALABILITÀ**	**SICUREZZA**	**GESTIBILITÀ**
1	Usare uno UF per inoltrare i dati, se possibile. L'uso dell'Heavy Forwarder dovrebbe essere limitato ai casi d'uso che lo richiedono. *AutoLB preconfigurato, capacità di riavvio, configurazione a livello centralizzato, scarsa necessità di risorse*		✓			✓
2	Usare almeno 2x pipeline di inoltro intermedie verso indexer quando si incanalano molti UF *Multiplexare un gran numero di forwarder endpoint attraverso un numero ridotto di forwarder intermedi ha impatti sull'uniforme distribuzione degli eventi tra gli indexer, che a sua volta pesa sulle prestazioni di ricerca.* ***Prevedere dei forwarder intermedi solo se assolutamente necessario.***	✓	✓			
3	Considerare di proteggere il traffico UF-IDX con SSL				✓	
4	Usare la funzione di bilanciamento del carico nativa in Splunk per distribuire i dati al livello di indicizzazione *I bilanciatori di carico di rete non sono attualmente supportati tra forwarder e indexer.*	✓		✓		

5	Usare server syslog dedicati per la raccolta di dati syslog *I server Syslog possono registrare su disco il traffico TCP/UDP in base alla fonte e consentono una configurazione del sourcetype adatta all'elaborazione con uno universal forwarder. I necessari riavvii dei forwarder non comportano una perdita di dati.*	✓				✓
6	Usare HEC per la raccolta senza agenti (anziché TCP/UDP nativo) *La Raccolta eventi HTTP (HEC) è un servizio di ascolto che consente di inviare gli eventi attraverso il protocollo HTTP[/S]. Può essere attivata direttamente sugli indexer, oppure configurata al livello di heavy forwarder; entrambi sono serviti da un bilanciatore del carico.*	✓				✓

Raccomandazioni per il livello di amministrazione / Utility

	Principi di progettazione / best practice (i propri requisiti determinano quale delle prassi seguire)	**COLONNE FONDANTI**				
		DISPONIBILITÀ	**PRESTAZIONI**	**SCALABILITÀ**	**SICUREZZA**	**GESTIBILITÀ**
1	Per gli ambienti di minori dimensioni, considerare di riunire LM, CM, SHC-D e MC su un'unica istanza *Questi ruoli server richiedono pochissime risorse e sono ottimi candidati per risiedere nella stessa posizione. Nei cluster di indexer di maggiori dimensioni, il CM*					✓

	può richiedere un server dedicato per gestire il cluster in modo efficiente.					
2	Per i deployment medio-grandi, valutare un'istanza separata di DS *Quando viene gestito un numero significativo di forwarder attraverso il Deployment Server, la richiesta di risorse aumenta fino al punto in cui per mantenere il servizio è necessario un server dedicato.*					✓
3	Per deployment estremamente grandi, valutare la presenza di più DS dietro il LB *Nota: in questo caso può essere necessario il supporto dei servizi professionali di Splunk per una corretta installazione e configurazione*	✓		✓		✓
4	Determinare se è possibile portare phoneHomeIntervalInSecs del DS oltre 60 secondi di default *Un intervallo più lungo di polling del server avrà un effetto positivo sulla scalabilità del DS*			✓		
5	Usare un DS dedicato/protetto per evitare lo sfruttamento di vulnerabilità lato client mediante la distribuzione di app *Chiunque con accesso al Deployment Server può modificare la configurazione Splunk gestita da quel DS, eventualmente anche distribuendo applicazioni malevole ai forwarder terminali. È prudente proteggere questo ruolo in modo adeguato.*				✓	
6	Usare la console di monitoraggio (MC) per monitorare lo stato di	✓	✓			✓

	salute del deployment e ricevere avvisi in caso di problemi. *La console di monitoraggio offre un set preconfigurato, specifico di Splunk, di soluzioni di monitoraggio e contiene allarmi di piattaforma estensibili che possono notificare lo stato di salute dell'ambiente.*					

Riepilogo e passaggi successivi

Questo whitepaper ha presentato un'introduzione generale sulle Architetture convalidate Splunk. Un'Architettura convalidata assicura che i requisiti dell'organizzazione siano soddisfatti nel modo più efficace del punto di vista dei costi, con la massima gestibilità e scalabilità. Le SVA offrono best practice e principi di progettazione fondati sui seguenti pilastri:

- Disponibilità
- Prestazioni
- Scalabilità
- Sicurezza
- Gestibilità

Questo whitepaper ha esaminato anche il processo di selezione delle Architetture convalidate Splunk in 3 passaggi: 1) Definizione dei requisiti, 2) Selezione della topologia, e 3) Applicazione dei principi e delle best practice di progettazione. Dopo aver presentato i molteplici benefici delle Architetture convalidate Splunk, auspichiamo che il lettore possa passare ora al processo di selezione di un'idonea topologia del deployment per la propria organizzazione.

Passaggi successivi

Cosa devo fare dopo aver scelto un'Architettura convalidata? I passaggi successivi nel percorso verso un ambiente di lavoro comprendono:

Personalizzazioni

- Considerare eventuali personalizzazioni necessarie di cui può avere bisogno la topologia scelta per soddisfare specifici requisiti.

Modello di deployment

- Decidere il modello di deployment (server fisico, virtuale, cloud).

Sistema

- Selezionare la tecnologia (server, archiviazione, sistemi operativi) secondo i requisiti del sistema Splunk.

Dimensionamento

- Raccogliere tutti i dati rilevanti che servono per dimensionare il deployment (inserimento di dati, volumi attesi di ricerche, esigenze di conservazione dei dati, repliche, ecc.) Splunk Storage Sizing (https://splunk-sizing.appspot.com/) è uno degli strumenti disponibili.

Personale

- Valutare le esigenze di personale per realizzare e gestire il deployment. È un aspetto essenziale per realizzare un centro di eccellenza Splunk.

Siamo disponibili per assistere nel processo di selezione di una Architettura convalidata e nei passaggi successivi. L'Account Team di Splunk è disponibile per eventuali domande. L'Account Team ha accesso a tutte le risorse tecniche e architetturali e sarà lieto di fornire altre informazioni.

Buono Splunking!

Appendice

Questa sezione presenta altri riferimenti usati nelle SVA.

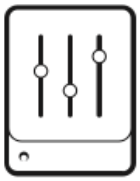
Appendice "A": illustrazione delle colonne portanti delle SVA

Pilastro	Descrizione	Obiettivi primari / Principi di progettazione
Disponibilità	La capacità di essere sempre operativo ed è in grado di riprendersi da guasti e interruzioni, programmati o meno.	1. Eliminare i single point of failure / Aggiungere ridondanza 2. Rilevare i guasti/errori programmati e non programmati 3. Tollerare i guasti programmati e non programmati, idealmente in modo automatico 4. Pianificare gli aggiornamenti graduali
Prestazioni	La capacità di utilizzare in modo efficace le risorse disponibili per mantenere un livello ottimale di servizio in condizioni di utilizzo differenti.	1. Aggiungere hardware per migliorare le prestazioni; calcolo, archiviazione, memoria. 2. Eliminare i colli di bottiglia “dal basso verso l’alto” 3. Sfruttare tutti i mezzi di elaborazione concomitante 4. Sfruttare la località (ovvero contenere al minimo la distribuzione di componenti) 5. Ottimizzare per i casi comuni (regola 80/20) 6. Evitare generalizzazioni non necessarie 7. Calcolo delle fasce orarie (pre-calcolo, calcolo nei tempi morti, calcolo condiviso/batch) 8. Certezza commerciale e precisione temporale (randomizzazione, campionamento)
Scalabilità	La capacità di assicurare che il sistema sia progettato per scalare su tutti i livelli e gestire in modo efficace maggiori carichi di lavoro.	1. Scalare verticalmente e orizzontalmente 2. Separare i componenti funzionali che devono essere scalati individualmente 3. Contenere le dipendenze tra componenti

Pilastro	Descrizione	Obiettivi primari / Principi di progettazione
		4. Progettare il prima possibile in vista della crescita futura nota 5. Introdurre una gerarchia nel progetto generale del sistema
Sicurezza	La capacità di assicurare che il sistema sia progettato per proteggere i dati e le configurazioni/gli asset pur continuando a generare valore.	1. Progettare un sistema sicuro sin dal principio 2. Usare protocolli aggiornati per tutte le comunicazioni 3. Consentire un accesso granulare e ad ampio spettro ai dati degli eventi 4. Impiegare un'autenticazione centralizzata 5. Implementare procedure di auditing 6. Ridurre le aree esposte ad attacchi o uso malevolo
Gestibilità	La capacità di assicurare che sistema sia progettato per operare a livello centralizzato ed essere gestibile a tutti i livelli.	1. Prevedere una funzione di amministrazione centralizzata 2. Gestire il ciclo di vita degli oggetti di configurazione (controllo della fonte) 3. Misurare e monitorare/profilare l'uso dell'applicazione (Splunk) 4. Misurare e monitorare la condizione di saluta del sistema

Appendice "B": componenti della topologia

Livello	Componente	Icona	Descrizione	Note
Amministrazione	Deployment Server (DS)		Il deployment server gestisce la configurazione della configurazione dei forwarder.	Deve essere distribuito in un'istanza dedicata. Può essere virtualizzato per agevolare il recupero in caso di guasto.
	License Master (LM)		Il license master è richiesto da altri componenti Splunk per abilitare le caratteristiche in licenza e tracciare il	Il ruolo del license master ha requisiti minimi in termini di capacità e disponibilità e può essere collocato insieme ad altre funzioni di amministrazione. Può

Livello	Componente	Icona	Descrizione	Note
			volume di dati inserito ogni giorno.	essere virtualizzato per agevolare il recupero in caso di guasto.
	Console di monitoraggio (MC)		La console di monitoraggio offre una dashboard per l'uso e il monitoraggio dello stato di salute dell'ambiente. Contiene inoltre diversi allarmi preconfigurati di piattaforma che è possibile personalizzare per inviare notifiche sugli aspetti operativi.	Negli ambienti in cluster, il MC può essere collocato insieme al nodo primario, in aggiunta al License Master e alla funzione di deployment server nei deployment non in cluster. Può essere virtualizzato per agevolare il recupero in caso di guasto.
	Master cluster (CM)		Il master cluster è il necessario coordinatore di tutte le attività in un cluster deployment.	Nei cluster con un gran numero di bucket di indici (elevato volume di dati/conservazione), il master cluster richiederà probabilmente un server dedicato sul quale funzionare. Può essere virtualizzato per agevolare il recupero in caso di guasto.
	Deployer del cluster di search head (SHC-D)		Il deployer del cluster di search head deve eseguire il bootstrap su un SHC e gestire la configurazione Splunk distribuita nel cluster.	L'SHC-D non è componente di runtime e ha quindi ripercussioni minime sui requisiti di sistema. Può essere collocato insieme ad altri ruoli di amministrazione. Nota: ogni SHC necessita della propria funzione di deployer SHC. Può essere virtualizzato per agevolare il recupero in caso di guasto.
Ricerca	Search head (SH)		La search head mette a disposizione degli utenti la UI per Splunk e coordina l'attività di ricerca programmata.	Le search head sono istanze Splunk dedicate nel deployment distribuito. Le search head possono essere virtualizzate per agevolare il recupero/ripristino a seguito di guasto, a condizione che dispongano di opportune

Livello	Componente	Icona	Descrizione	Note
				risorse di CPU e memoria.
	Cluster di search head (SHC)		Un cluster di search head è un insieme di almeno tre search head in cluster. Offre scalabilità orizzontale per il livello delle search head e un failover trasparente in caso di malfunzionamenti.	I cluster di search head richiedono server dedicati con specifiche idealmente identiche al sistema. I membri dei cluster di search head possono essere virtualizzati per agevolare il recupero, a condizione che dispongano di opportune risorse di CPU e memoria.
Indicizzazione	Indexer		Gli indexer sono il cuore di Splunk. Elaborano e indicizzano i dati in ingresso e servono come destinazioni di ricerca per eseguire le richieste di ricerca avviate a livello di ricerca.	Gli indexer devono essere sempre su server dedicati in deployment distribuiti o in cluster. In un deployment con un unico server, l'indexer fornisce anche le funzioni di UI per la ricerca e le funzioni di license master. Gli indexer offrono le prestazioni migliori su tradizionali server fisici o su macchine virtuali dedicate ad elevate prestazioni, se sono garantite le risorse adeguate.
Raccolta dei dati	I forwarder e gli altri componenti per la raccolta dei dati		Icona generale per qualunque componente che prende parte alla raccolta dei dati.	Comprendono gli universal forwarder e gli heavy forwarder, gli input della rete dati e altre forme di raccolta dei dati (HEC, Kafka, ecc.)