splunk>	**Splunk 4.3**	
	Event and log analysis tool	
	Systeemvereisten: Windows/Unix/Linux	
	Ontwikkeld door:	Splunk
Gratis/commerciële licentie	**Contactpersoon:**	Bob.lannoy@smals.be

Functionaliteiten

Splunk positioneert zich aan de hand van de term “Operational intelligence”. Het is een pakket dat allerlei (log)informatie verzamelt en doorzoekbaar maakt voor analyses. Het soort informatie dat je kan analyseren is heel breed: besturingsysteemlogs en metrieken, webserver logs, logs van netwerktoestellen of eigen logs met eigen data.
De belangrijkste functionaliteiten zijn:

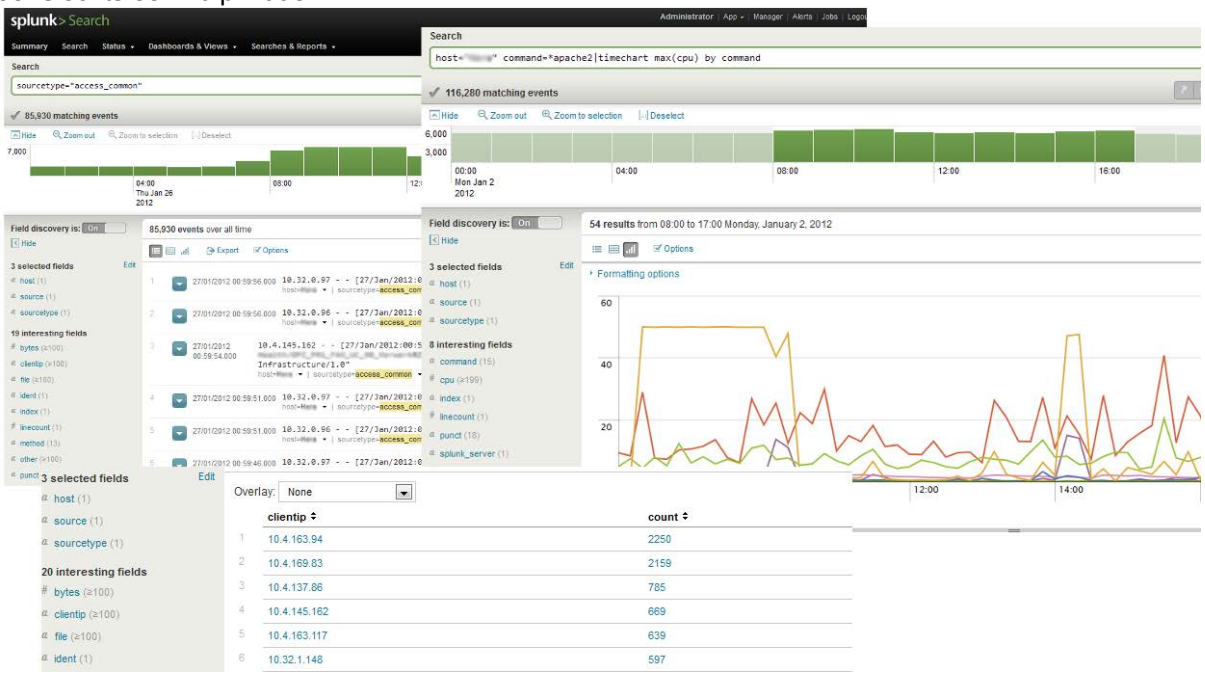
- Indexeren van ruwe event-informatie uit diverse bronnen: logs en informatie van besturingssystemen, applicatieservers, netwerkinformatie, ...
- Met behulp van *forwarders* kan je info decentraal verzamelen en naar een centraal systeem sturen
- Web interface om de data te analyseren
- Detectie van velden in de loglijnen en mogelijkheid om eigen velden met regular expressions af te bakenen
- Zoekbalk dient als commandobalk om zowel eenvoudige als complexe operaties uit te voeren
- Tijdslijn laat toe om in te zoomen op specifieke momenten
- Full text search op de geïndexeerde informatie
- Het gebruik van velden en combinaties ervan om specifieke informatie op te halen. Een voorbeeld is het extraheren van informatie uit de URLs van webserver logs om het aantal bezoekers per pagina op te halen.
- Visualisatie van gegevens (verloop in tijd van bepaalde waarden en datareeksen)
- Aanspreken van externe bronnen voor verrijking van gegevens, zoals het opzoeken van IP-adressen, volledige namen van gebruikers op basis van de user-id die in de log voorkomt, ...
- Bijkomende *Apps* die het makkelijker maken om event/loginformatie van specifieke systemen beter te verwerken zoals herkennen van velden, specifieke visualisaties, ...
- Alerting mogelijkheden (enkel in betalende versie)

Conclusies en Aanbevelingen

Splunk is heel makkelijk op te zetten en laat toe om uitgebreide analyses te doen overheen verschillende datasets. Ondanks de uitgebreide documentatie vraagt het wel wat werk om vertrouwd te geraken met de mogelijke commando’s, het gebruik ervan en de juiste manier om iets uit te voeren.
Gezien de complexe commandline functionaliteit in de zoekbox is Splunk niet voor business users tenzij via voorgedefinieerde dashboards.
Kortom, het is alleszins een handige tool om een zicht te krijgen op de data en bepaalde gedragingen die men wenst te verklaren.

Testen en Resultaten

Na een eenvoudige installatieprocedure verwacht Splunk dat je een aantal data-inputs voorziet. Standaard kan er al heel wat data verwerkt worden zoals van Windows/Unix event logs en performantiemetrieken, IIS en Apache logs, Cisco logs en eigen logformaten. Voor een aantal van deze logs detecteert Splunk het logtype (source type) waardoor er automatisch ook een aantal standaardvelden beschikbaar zijn. In een Apache access log zal Splunk bijvoorbeeld al de diverse velden oplijsten zoals bronadres, HTTP return code, opgevraagde URL, HTTP methode, etc.
In de test werd er niet gewerkt met forwarders maar werden een aantal logs lokaal op de Splunk-machine opgeladen. Naast een test met Apache-logs hebben we ook geëxperimenteerd met een eigen custom log dewelke de CPU-belasting voor diverse processen bevatte. Na het opladen van deze eigen log kan je makkelijk velden definiëren. Dit doe je door een aantal voorbeeldwaarden in te geven waarna er automatisch een reguliere expressie gegenereerd wordt. Naast de automatische generatie van expressies is het ook mogelijk om zelf een reguliere expressie in te geven of een bestaande te wijzigen. Na het construeren van de expressie kan Splunk informatie uit de log parsen en door middel van informatievelden weergeven. Deze velden kunnen vervolgens ook gebruikt worden om informatieve plots te maken.
De Splunk commando's worden als een filter chain gemaakt. Je verfijnt de lijst met events die je wenst over te houden en daarmee kan je dan lijsten en grafieken maken. Zo kan je bijvoorbeeld snel alle HTTP-requests opsporen die fouten genereerden (*status=5* OR status=404)* en daar dan op zoek gaan naar detailinformatie.
Zoals elke tool die grote datavolumes verwerkt, wordt al snel duidelijk dat om goede analyses te kunnen uitvoeren je ook goed moet weten waar je kan zoeken en kennis moet hebben van het soort data. Splunk is hiervoor slechts een hulpmiddel.



Figuur 1: Voorbeelden van de webinterface met o.a.zoekbalk en grafiek

Gebruiksvoorwaarden

Splunk bestaat in 2 versies: Free en Enterprise. De gratis versie is beperkt tot het indexeren van 500 MB per dag.
De Enterprise versie omvat een aantal extras zoals monitoring/alerting, geschedulede PDF rapportering, load-balacing/fail-over, toegangscontrole, multi-user support, ... en specifieke Splunk Apps (Premium Apps). De website van Splunk vermeldt qua prijs dat de omzetting van een Free license in een Enterprise license voor hetzelfde volume $6000 / jaar kost.