AWS Data & Analytics

Turn Real-Time Insight into Action

aws marketplace

Overview

Turn Real-Time Insight into Action

In today's big data digital world, your organization produces large volumes of data with great velocity. Generating value from this data and guiding decision making require quick capture, analysis and action. Without strategies to turn data into insights, the data loses its value and insights become irrelevant. Real-time data integration and analytics tools play a crucial role in harnessing your data so you can enable business and IT stakeholders to make evidence-based decisions.

Table of Contents

Table of Contents

Turn Real-Time Insight into Action

In this e-book, we explore the possibilities of turning real-time insight into action on AWS through two use cases: implementing IT operational analytics through packaged solutions on AWS and enabling custom-made application to create business actions by incorporating a real-time data pipeline. For each use case, we outline the specific steps you can take to implement AWS and solutions in AWS Marketplace when managing your organization's big data.

Improving Decision Making with Data Insights

Your IT initiatives to implement and maintain analytics applications in the cloud can transform your organization's access to data. These applications enable your organization to access and refresh data in real time, making the most of sophisticated technology tools.

According to Robert Kim, director of software and cloud solutions from Presidio, an IT solutions provider focused on digital infrastructure, cloud and security solutions, *"You can't use traditional analytics and application platforms if you want to do anything real time. The reason why businesses are using cloud-native application and stateless application delivery is because you have to get away from the limitations of latency."* When it comes to real-time access, cloud solutions are key.

Tips from the Pros

Real-World Advice from Experts that Live Data Everyday

Ronald van Loon, Social Media Influencer

Adversitement delivers data management solutions and customer insights to optimize the customer experience.

Businesses have to make real-time actions based on streaming (IoT, devices, people, business) data to continuously refine the customer experience, improve sales and reduce costs. For example, cross-channel sales require live actions if businesses want to retain customers and obtain that sale. If the data cannot be processed in real time and action is taken days later, the customer has already moved on to a competitor.

Log data is a challenge for DevOps and IT departments, which have to pinpoint small but potentially detrimental issues in bulk streams of log data. Cognitive solutions, using machine learning algorithms, matches human domain knowledge with log data, along with open source repositories, discussion forums, and social thread. It makes a data reservoir of insights that hold the solutions for issues faced by IT and DevOps daily.

Ronald van Loon, Director, Adversitement. @Ronald_vanLoon

Use Case #1: Implementing an IT Operational Analytics Solution

For IT organizations today, it's essential that your IT operations and management teams adopt tools that allow them to gain visibility and optimize operations. As businesses become more digital, collecting and analyzing large volumes of data across IT and business sources can lend great insights into IT operations. However, collecting and analyzing this data in a way that lends value to the business poses a significant challenge. Searching IT operations data is a powerful mechanism when operators know what they are looking for. However, as data volumes grow, search becomes less useful. So finding the right operational analytics solution is vital to gaining visibility into IT operations and making effective use of large quantities of data.

With a robust IT operational analytics solution, you can accelerate application delivery and monitor and troubleshoot in real time, while helping to ensure security and compliance. By integrating and analyzing your IT operational data, you can gain insight into IT operations, manage IT infrastructure events and processes and take immediate action through automated processes and workflows.

An IT operational analytics solution enables you to take timely action, gain visibility, and ensure continuous insight. Operational analytics solutions allow you to leverage your organization's existing IT systems, giving you analytics-driven processes and workflows that align with your goals. Users can visualize information from enterprise systems and data sources to gain intelligence on key IT events and processes. In real time, you can compare and analyze events from multiple data sources to gain valuable knowledge. Advanced analytics can predict threats and opportunities and prescribe the next best actions.

Analyzing IT Operational Analytics with AWS

Amazon Web Services offers analytical processing tools that enable you to gain insights from your IT operational data. These tools can help you identify patterns and trends in your data, gain insight into anomalies before they happen and improve decision making across the organization.

Amazon Kinesis enables you to collect, process and analyze real-time, streaming data. The solution offers capabilities to cost-effectively process streaming data at scale with the ability to choose the tools that suit the requirements of your application. With Amazon Kinesis, you can ingest real-time data such as video, audio, application logs, website clickstreams and Internet of things (IoT) telemetry data for machine learning, analytics and other applications.

Further enabling analytical processing is AWS Lambda, a solution that lets you run code without provisioning or managing servers. You pay only for the compute time you consume. You can run code for any type of application or backend service, with no administration. You just upload your code and Lambda takes care of everything required to run and scale your code with high availability.

Implementing IT Operational Analytics with AWS Seller Solutions

AWS and its partners offer multiple solutions to implement IT operational analytics. AWS solutions include Amazon S3, Amazon Kinesis, Amazon Redshift, AWS Lambda and Amazon DynamoDB. Partner solutions include Splunk Cloud, NetApp ONTAP Cloud and Palo Alto Networks Add-on for Splunk. Below is an overview of these solutions and the technical requirements and aspects of implementation.

How to Implement Splunk Cloud on AWS

Splunk Cloud Overview

With Splunk Cloud in the AWS cloud, you gain the flexibility of the AWS infrastructure to tailor your Splunk deployment, and you can modify your deployment on demand as your needs change. When you build on AWS, you no longer need to consider lead times waiting for hardware to change or to scale your Splunk Cloud deployment.

When you implement Splunk Cloud, your machine data becomes accessible and usable. You can search, monitor and analyze data from any source to gain valuable intelligence and insights across your entire organization. With a full range of powerful search, analysis and visualization capabilities with pre-packaged content for use cases, you can discover and share insights.

Splunk Cloud Features and Benefits

With Splunk Cloud on AWS, you can manage, investigate, analyze and visualize your IT organization's operational data with efficiency and security. It enables:

Data collection and indexing regardless of format or location. The structure and schema are applied only at search time, so you can analyze the data without limitation.

Data search and investigation using the Splunk Search Processing Language (SPL). Splunk automatically normalizes your data formats and provides more than 140 commands to perform statistical searches, calculate metrics and look for specific conditions within a rolling time window.

Data correlation by identifying relationships between events or activities. Splunk Cloud correlates data based on time, location or custom search results. Use the Transaction command to identify related events such as a transaction or session and investigate failed transactions. Let the Event Pattern Detection command automatically find common or rare patterns in your data. You can also let users discover and share insights with the point-and-click Pivot interface.

Visualization and reporting to make sense of volumes of data. Different user types can use custom dashboards, and you can generate custom reports in PDF format or export results for broader use.

Data integrity through security and administration tools. Splunk Cloud software provides secure data handling, access controls, auditability, assurance of data integrity and integration with enterprise single sign-on solutions.

How Splunk Cloud Works

To deploy Splunk Cloud on AWS, you can implement the Quick Start and set up the following:

- An Amazon Virtual Private Cloud (VPC) configured across two or three Availability Zones. The Quick Start provisions one public subnet in each Availability Zone.
- Two Elastic Load Balancing (ELB) load balancers: one to load-balance HTTP web traffic to the search head instances and the other to load-balance HTTP event traffic destined for the Splunk HTTP Event Collector (HEC) across all indexer instances.
- An AWS Identify and Access Management (IAM) user with fine-grained permissions for access to AWS services necessary for the initial deployment process.
- Appropriate security groups for each instance or function to restrict access to only necessary protocols and ports.

- In the public subnets, EC2 instances for Splunk Enterprise, including the following:
  - Splunk indexer cluster with the number of indexers you specify (3-10), distributed across the number of Availability Zones you specify. The Splunk receiver (splunktcp) and Splunk HEC are enabled across all indexers.
  - Splunk search heads, either stand-alone or in a cluster, based on your input during deployment. In the latter case, the search heads are distributed across the number of Availability Zones you specify.
  - Splunk license server and indexer cluster master, co-located.
  - Splunk search head deployer, where applicable.
  - (Optional) User-provided Splunk apps and add-ons, loaded and pre-installed across indexers and search heads, based on your input.

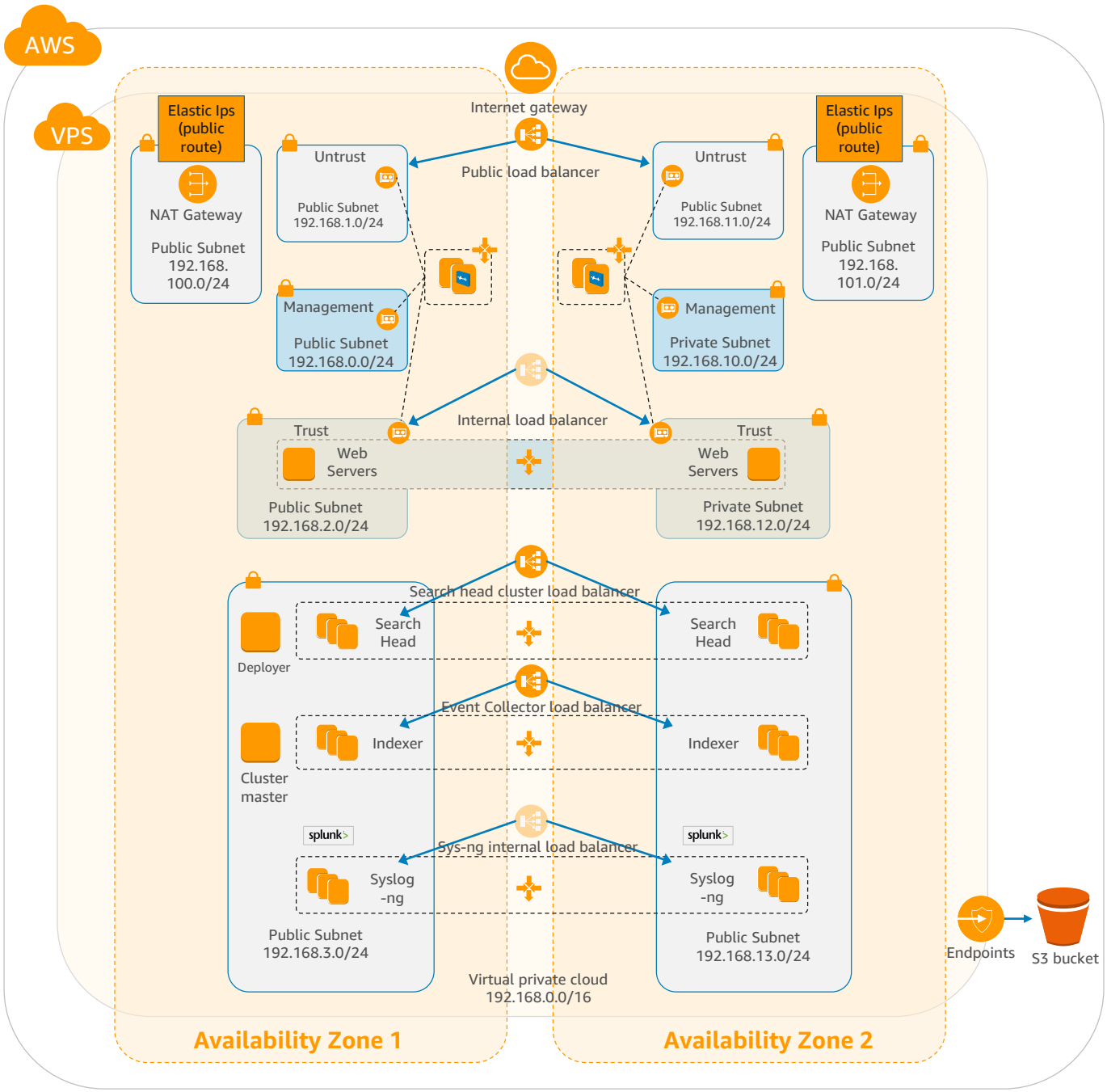
There are two deployment options:

- Deploy Splunk into a new Amazon VPC (end-to-end deployment). This option builds a new AWS environment consisting of the VPC, subnet, security groups and other infrastructure components and then deploys Splunk into this new VPC.
- Deploy Splunk into an existing Amazon VPC. This option provisions Splunk in your existing AWS infrastructure.

Architecture

Deploying the Quick Start Guide for a new VPC with default parameters builds the following Splunk Cloud environment in the AWS cloud:

Figure 1: Splunk Cloud on AWS



Tips from the Pros

Real-World Advice from Experts that Live Data Everyday

Splunk

Splunk helps businesses maximize machine data, turning it into answers. Organizations use market-leading Splunk solutions with machine learning to solve their toughest IT, IoT and security challenges.

As more of your critical workloads move to AWS, you need to gain critical security, operational and cost management insights across your entire AWS and hybrid environment. Splunk provides that kind of visibility with Splunk Enterprise, Splunk Cloud on the AWS Marketplace, or Splunk Insights for AWS Cloud Monitoring.

In addition, the Splunk App for AWS, included in Splunk Insights for AWS Cloud Monitoring, provides users with dashboards, visualizations and alerts right out of the box. The Splunk App for AWS offers a rich set of pre-built dashboards and reports to analyze and visualize data from numerous AWS services—including AWS CloudTrail, AWS Config, AWS Config Rules, Amazon Inspector, Amazon RDS, Amazon CloudWatch, Amazon VPC Flow Logs, Amazon S3, Amazon EC2, Amazon CloudFront, Amazon EBS, Amazon ELB and AWS Billing—all from a single, free app.

With Splunk, customers can integrate multiple data sources into a searchable repository to easily troubleshoot and monitor their environment. Splunk captures complex interactions between the infrastructure and application components, making it easy to perform root cause analysis.

Splunk Enterprise is perfect for deploying on AWS. It's self-contained and can be easily deployed on any EC2 instance. Splunk Enterprise also scales horizontally, making it ideal for an AWS deployment.

Tyillere Hansen, Sr. Manager, Global Strategic Alliances Marketing, Splunk

Leveraging AWS Marketplace to Make Sense of Firewall Data

Palo Alto Networks Add-on for Splunk Overview

The Palo Alto Networks Add-on for Splunk gives visibility into application traffic, helping your security team by enforcing policy-based threat control and prevention and by identifying the root cause of issues. Both Palo Alto Networks and Splunk provide technologies that help protect your workloads from cyber-attacks and provide visibility, analytics and reporting across cloud, on-premises and hybrid environments. Splunk and Palo Alto Networks both operate under the understanding that the ability to detect cyber threats is an absolute necessity, but the value of this is limited if you don't possess the ability to intervene.

The Palo Alto Networks App leverages the data visibility provided by the Palo Alto Networks security platform with Splunk's investigation and visualization capabilities to deliver security reporting and analysis. The app enables security analysts, administrators and architects to correlate application and user activities across network and security infrastructures from a real-time and historical perspective.

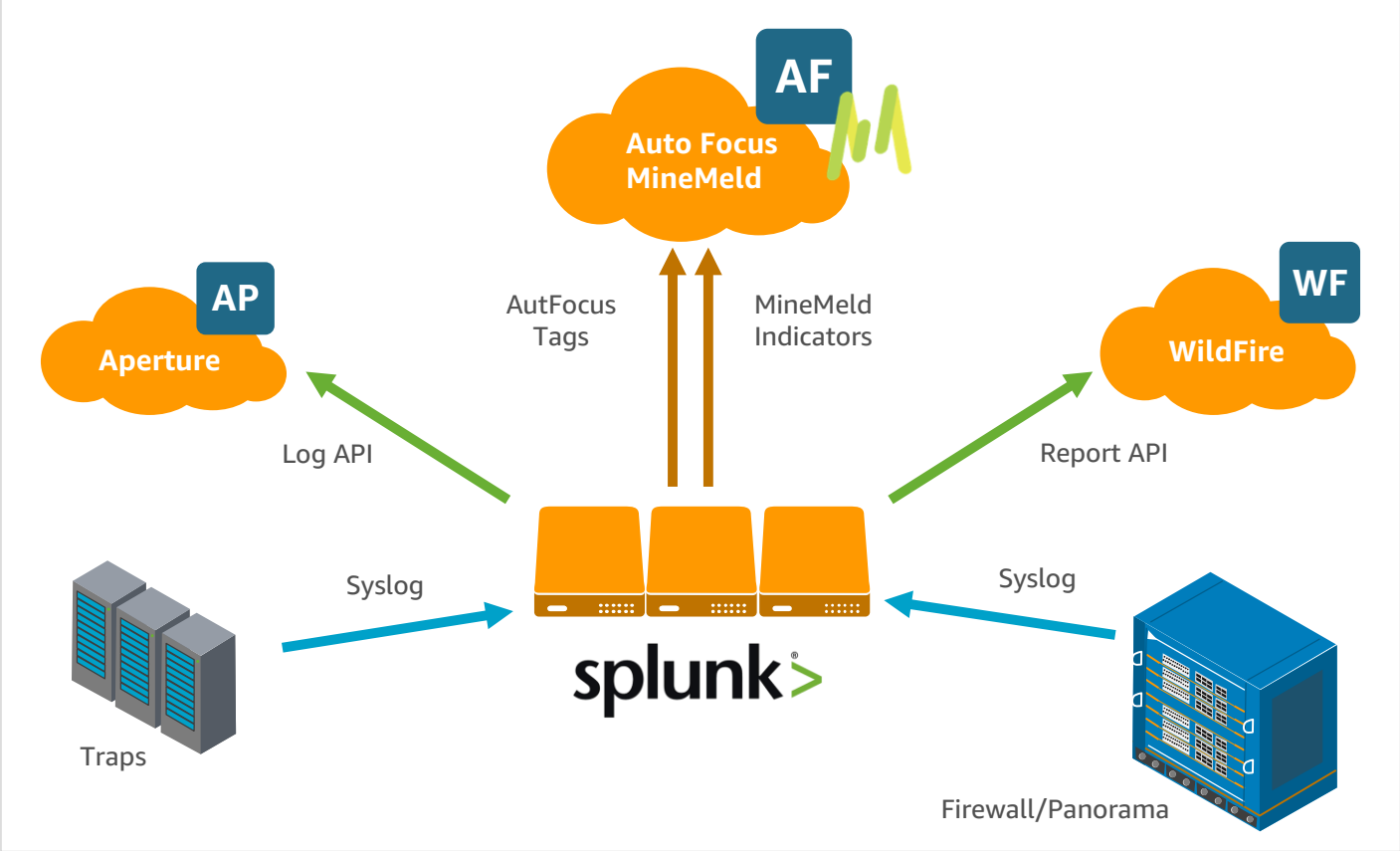
How Palo Alto Networks Add-on for Splunk Works

When you implement Palo Alto Networks App or Add-on, you can increase visibility into network traffic and malicious activity, both in the network and on the endpoint. When you combine this visibility with Splunk, you can make correlations and perform analytics across disparate data streams. These correlations can be with different kinds of Palo Alto Networks data, for example, correlating WildFire reports against traffic logs to detect infected hosts or correlating firewall logs with endpoint logs. The real power of Splunk, however, is finding correlations and performing analytics across multiple sources of data and multiple vendors, for example, correlating firewall logs with webserver logs or advanced endpoint security logs with Windows event logs.

The Palo Alto Networks VM-Series next-generation firewall complements AWS security groups and web application firewalls by classifying and controlling application traffic on AWS based on the application identity, and then applying threat prevention policies to block known and unknown cyber-threats. Splunk Enterprise provides security visibility by capturing and analyzing logs from the Palo Alto firewall using the Palo Alto Networks App for Splunk.

Splunk can collect data from all Palo Alto Networks products, each providing visibility and control. Splunk, the Palo Alto Networks App for Splunk and AWS work together to protect your workloads from cyberattacks and provide visibility, analytics and reporting across cloud, on-premises and hybrid environments.

Figure 3: Palo Alto Networks Add-on for Splunk on AWS



This "Security and analytics environment on AWS" Quick Start builds an enterprise-class security and analytics environment on the Amazon Web Services (AWS) Cloud, using the Palo Alto Networks VM-Series next-generation firewall, Splunk Enterprise, and the Palo Alto Networks App for Splunk, along with complementary services from AWS

Leveraging AWS Marketplace to Make Sense of File Storage Data

NetApp ONTAP Cloud Overview

NetApp ONTAP Cloud is a data management solution that provides protection, visibility and control for your cloud-based workloads. With NetApp ONTAP Cloud, you get a universal storage platform that addresses your cloud data needs. You can employ this common data management platform in the cloud and on-premises, allowing users to benefit from a single data fabric without having to retrain IT staff on new data management methods.

How Splunk and NetApp Work Together

Splunk helps you gain visibility into your NetApp storage, so you can monitor the health of your entire NetApp Data ONTAP environment. Splunk provides a unified, single pane-of-glass view into all systems throughout the enterprise so you can identify sub-optimal performance and pinpoint the exact source of performance degradations. With a higher-level overview of all the NetApp systems across your enterprise, your IT infrastructure team can drill down to a particular incident and easily isolate the source of degradation at any point in time.

You can observe the most important performance metrics like latency, as well as log data from your NetApp storage systems. Also, you can combine NetApp's storage log monitoring with Splunk software's alerting capability to be notified of an important issue the moment it arises. With Splunk and NetApp, analyzing storage performance trends in real time does not require you to deploy multiple monitoring solutions. With these insights, you can spot issues like abnormal latency of a particular volume and compare it with other metrics of interest.

The Splunk App for NetApp Data ONTAP provides real-time comprehensive visibility, flexible operational analytics and cross-tier view into your NetApp storage systems. With Splunk and NetApp you can quickly isolate your storage issues, optimize your storage performance and proactively plan your NetApp systems capacity.

For more information, visit our Quick Start Guide on NetApp ONTAP on AWS.

Use Case #2: Building a Real-Time Data Pipeline for Analytics

In today's IT environments, operations data is generated continuously by a steadily increasing number of diverse data sources. Real-time analytics have become essential for IT organizations looking to streamline operations. Whether you are gathering insights from Internet of things data, analyzing trends and patterns in event logs or monitoring network traffic, real-time data brings tremendous value.

Developing a process for capturing, storing and processing this data can enable real-time IT automation. However, traditional data processing infrastructures – especially those that support applications – were not designed for the large quantities of data generated by our mobile, streaming and online world. Making sense of volumes of data and gaining visibility into IT operations requires building a real-time data pipeline.

By building a real-time data pipeline for analytics, you can process and move data between different compute and storage services, whether in the cloud or on-premises, at specified intervals. This enables you to regularly access your data where it is stored, transform and process it at scale and generate insights. In addition to packaged solutions that consume your real-time systems data, you can build a custom application to provide real-time data analytics.

Tips from the Pros

Real-World Advice from Experts that Live Data Everyday

Slalom Consulting

Slalom Consulting provides program management, business process improvement, software development for companies from the Fortune 500 to emerging companies. The company also provides specialized solutions including enterprise messaging and collaboration, business mobility, custom relationship management (CRM) and enterprise resource planning (ERP).

It is very easy to move your workloads to the cloud and be focused on the operations of technology, and lose sight of the user experience. Whether that user experience is an internal user, or whether it is an external user, I tell all of our clients to really think carefully about providing a positive user experience. A good example of this is ensuring the performance they experience with reporting and analytics is meeting their needs and expectations. This is great place to focus to help make sure your project is successful.

It is also critical that roles and responsibilities are well-defined and that those are well defined from a cloud perspective. It is important to work closely and early with your stakeholders such as corporate security teams to avoid misunderstandings and project delays.

Adam Hood, Practice Director of Data Engineering and Data Science Solutions, Slalom Consulting

Building a Real-Time Data Pipeline with AWS Partner Solutions

To process large amounts of real-time or streaming data you need a data pipeline. The most significant challenges associated with creating a data processing pipeline are minimizing latency and achieving a near real-time processing rate. With the right tools, however, these challenges can be overcome.

A data pipeline architecture consists of a number of layers:

- Data ingestion
- Data collection
- Data processing
- Data storage
- Data query
- Data visualization

AWS and its partners offer solutions to help with each aspect of building a real-time data pipeline. AWS solutions include Amazon EC2, Amazon S3 and Amazon Redshift. Seller solutions like Matillion ETL, Attunity and SoftNAS can help with various layers of your data pipeline. Specifically, AWS seller TIBCO Spotfire provides data visualization capabilities for your pipeline. Below is an overview of the solution and the technical requirements and aspects of implementation.

Download a full tutorial on How-to implement an Intelligent Analytical System integrating AWS Services and Solutions in AWS Marketplace.

TIBCO Spotfire Cloud Overview

TIBCO Spotfire Cloud is a SaaS analytics solution designed for data exploration. You can explore, visualize and create real time dashboards for Amazon Redshift, RDS, Microsoft Excel, SQL Server, Oracle and more. Visualization cam be embedded into web pages and custom applications.

TIBCO Spotfire Case Study

Challenge

Data growth presents a real problem when you are trying to make use of raw data. A food sales and marketing firm in the Southeastern U.S. needed to wrangle and cleanse unusable data to determine how to support its customers. *"My team helps brands and our sales team understand trends and opportunities at particular retailers. We're trying to determine the price point that gets the greatest lift for a product,"* said a representative from the company.

Analysts were spending time and resources cleansing data and multiple divisions were doing some of the same tasks. They needed to become more efficient at deriving insights and getting actionable information, rather than simple reports.

Solution

TIBCO provided them with the capability to visualize the data with interactive graphics that could rapidly refresh and show live data and analytics across 80 categories. *"TIBCO set itself apart with interactive graphics. You click on a graphic, and two or three others show you which data is affected in real-time."*

Results

The company's business has grown since implementing TIBCO Spotfire analytics. Spotfire played a key role in meeting new client needs and helping them succeed with new promotions. *"For our division, the Spotfire system absolutely paid for itself within the first year."*

TIBCO Spotfire Cloud Features and Benefits

TIBCO Spotfire Cloud enables you to rapidly create data visualizations, including drill-down, for Amazon Relational Database Service (RDS) and Redshift. You can connect to on-premises data sources through VPC, VPN or AWS Direct Connect (requires additional services from AWS). The solution enables you to share visualizations, dashboards and analysis securely with anyone. Spotfire's collaboration features facilitate discussions with peers, advisors and customers.

Spotfire connects to your AWS data, allowing you to build dashboards from a variety of data sources. You can then visualize predictive and statistical analytics with built-in functions or write your own R code using the integrated TIBCO Enterprise Runtime for R engine.

Spotfire for AWS is launched within your Amazon AWS account, so you can add drivers, users, configure changes and more. End-to-end security can be configured in accordance with the AWS shared security model and uses native AWS access management tools. Everything from the back-end data connections to individual user entitlements can be controlled and monitored.

How TIBCO Spotfire Cloud Works

With TIBCO Spotfire Cloud for AWS on a single Amazon Machine Image (AMI), you get a multi-user platform with a variety of services preinstalled and pre-configured. It offers:

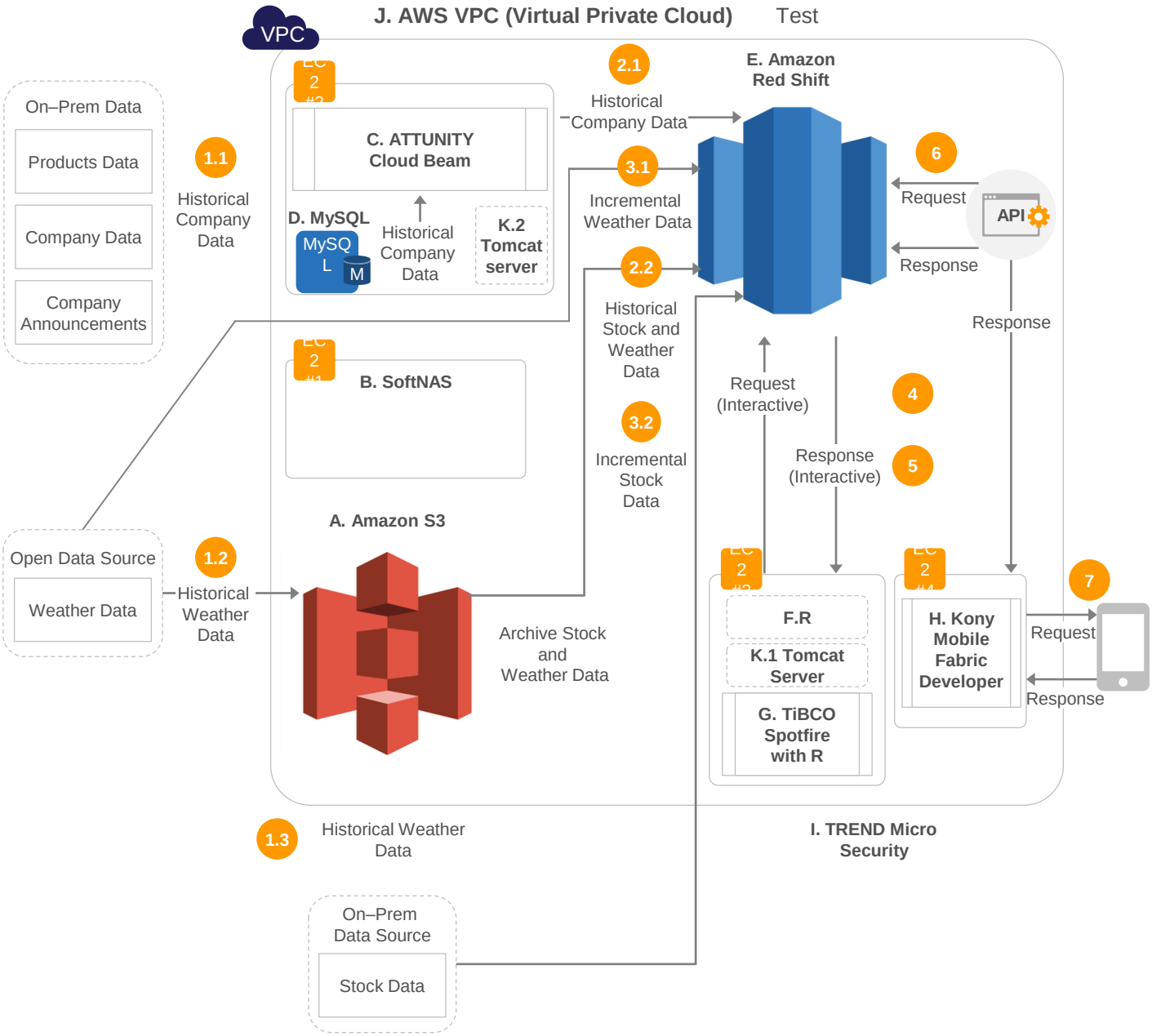
- Connections to dozens of data sources, in the cloud and on-premises.
- Data exploration through best-of-breed, interactive visual analytics.
- Dashboard and guided analytics authoring and sharing.

Spotfire enables you to access data in two primary ways: in-memory and in-database. With in-memory, you load all or subsets of data into the Spotfire in-memory data engine and Spotfire performs all calculations. When using this approach, you have access to an extensive set of transformations and calculations that are built into Spotfire.

For ad-hoc analysis with in-database, you leave the data in the underlying database. Spotfire issues dynamic queries with every user interaction, bringing back only a small set of aggregated results required to refresh the visualizations. This method is scalable to large data sizes, beyond what can fit in memory and there are no upfront load times.

Scaling TIBCO Spotfire Cloud for AWS

High–Level Architecture & Data Flow Design



When you need to scale, Spotfire offers two options. You can scale up, making more system resources available to an existing Spotfire instance. Alternatively, you can scale out, adding new instances to an existing Spotfire deployment. You should consider scaling your instance when any of the following occur:

- Slow response to user interactions.
- Slow data loading from information links.
- Low on-system memory (RAM).
- Windows paging to virtual memory.

With a variety of Amazon EC2 instance types, you can select the most cost-effective instance type according to the required system capacity and expected compute tasks. You can scale up or down with these steps:

- Stop the Spotfire instance you want to re-scale.
- Select the desired instance type and size (larger or smaller).
- Restart the instance.

Scaling out is a more advanced setup, in which multiple EC2 instances containing Spotfire services can be either load-balanced or service-optimized.

Tips from the Pros

Real-World Advice from Experts that Live Data Everyday

TIBCO Software

TIBCO Spotfire is a comprehensive platform for visual analytics that helps businesses prepare, clean and connect their data. Then they can visualize the data and apply advanced predictive capabilities, as well as scaling easily.

I think the most appealing thing with TIBCO Cloud(tm) Spotfire is that it allows IT to focus on getting the data in order and less on having to help business users get answers. You can focus on providing a foundation of data and governance in place instead of spending your time helping individual end users. Spotfire is truly a self-service visual analytics application that you can deploy very broadly.

Nowadays, where data lakes are more popular, IT organizations are spending less time upfront preparing the data. They want to load data into the data lake and refine it as they go. Spotfire can fill a gap here since it allows IT to give the business users value right away because it includes self-service data preparation capabilities that business users can use on their own.

The cycle time, I guess in addition to what I said, is also something critical to IT. The trend seems to be that there's the desire to delegate, delegate report writing for sure, self-service report writing, self-service dashboards, but also to some extent self-service data preparation. That's the main value to IT, that Spotfire as a platform allows you to do that in a fairly easy-to-use way.

That's why I think Spotfire fits really well with that sort of mindset; start small and grow. See what it's used for and then increase as new use cases are discovered. That's certainly how Spotfire spread even before AWS came around. The way we spread internally at companies was basically from department to department. I think that's the best advice, put a robust core in place and then scale out from there.

Tobias Lehtipalo, Sr. Director Product Management, TIBCO Software, Inc

Next Steps: AWS Marketplace

AWS and AWS Marketplace solutions together enable big data analysis to help you understand and manage your business and data, enhance decision making and take action in real time. By collecting, managing, securing and analyzing large amounts of structured, semi-structured and unstructured data, you can inform business decisions, act on key insights and guide customer interactions and business processes.

Data management solutions on the AWS Marketplace include:

Splunk Cloud

Splunk Cloud provides security and operational visibility across your AWS environment, including applications, infrastructure and AWS services such as CloudTrail, Config, VPC Flow Logs and more. Organizations of all sizes leverage Splunk visibility with AWS agility to troubleshoot applications, ensure security and compliance and monitor business-critical services in real time. Splunk Cloud makes it easy to gain end-to-end visibility across your AWS and hybrid environments. Leverage Splunk Cloud with the free Splunk App for AWS to gain critical security, operational and cost optimization insight into your AWS deployment. Whether you're managing applications, infrastructure or a security operations center in the cloud, Splunk Cloud delivers operational intelligence for a real-time understanding of your IT organization, so you can make informed decisions.

https://aws.amazon.com/marketplace/pp/B06XK299KV

NetApp ONTAP Cloud

NetApp is the leading enterprise storage software providing secure, proven NFS, CIFS and iSCSI data management for non-disruptive operation on AWS storage. With NetApp ONTAP Cloud, you get the power of ONTAP software with high availability and flexible performance and capacity options managing up to 360TB of your Elastic Block Store (EBS) and S3 capacity.

https://aws.amazon.com/marketplace/pp/B011KEZ734

Palo Alto Networks Add-on for Splunk

Palo Alto Networks is a leader in cyber-security, protecting thousands of enterprise, government and service provider networks from cyber-threats. Unlike fragmented legacy products, our security platform safely enables business operations and delivers protection based on what matters most in today's dynamic computing environments: applications, users and content.

https://aws.amazon.com/marketplace/pp/B00PJ2V04O

TIBCO Spotfire Cloud

TIBCO Spotfire Cloud is analytics SaaS application designed for data exploration. Everything you need is available in the cloud. TIBCO Software Inc. is a global leader in infrastructure and business intelligence software. Whether it's optimizing inventory, cross-selling products or averting crises, TIBCO delivers the ability to capture the right information at the right time and act on it preemptively for a competitive advantage. With a broad mix of innovative products and services, TIBCO is the strategic technology partner trusted by businesses around the world.

https://aws.amazon.com/marketplace/pp/B00PB74KYY

Matillion ETL

Matillion ETL for Amazon Redshift makes loading and transforming data on Redshift fast, easy, and affordable. The AMI takes less than five minutes to set up and delivers results much faster than traditional ETL technologies. With just a few clicks, you can load data into Redshift from dozens of sources, including S3 and RDS, multiple databases and APIs, common systems like Google Analytics, Salesforce, Netsuite and SAP, and even social media like Facebook and Twitter. Matillion ETL makes it easy to orchestrate and automate data load and transform, integrate with other systems and AWS services, leverage scripts, and much more.

https://aws.amazon.com/marketplace/pp/B010ED5YF8

Attunity

Attunity offers solutions that help you accelerate data replication, ingest, and streaming across a wide range of heterogeneous databases, data warehouses and data platforms. Used by hundreds of enterprises worldwide, Attunity solutions help you get more value from your data on the AWS Cloud easily, securely, and efficiently with zero downtime. Attunity technology is validated by AWS for Data Integration in the AWS Big Data Competency Program and for Data Migration in the AWS Migration Competency Program.

https://aws.amazon.com/marketplace/pp/B00JK3HIWQ

SoftNAS

SoftNAS Cloud Enterprise extends native AWS storage (EBS, S3) to create an enterprise-class, full-featured cloud NAS filer. Safely migrate mission-critical applications to the AWS cloud without a physical storage appliance in minutes without specialized training. For workloads with basic networking, caching, deduplication and encryption requirements, General Purpose Edition balances cost versus memory and CPU resources and is a good starting point for new users.

https://aws.amazon.com/marketplace/pp/B06Y5XGRWF