



Digital Process & Architecture: Data, Deployment & Notifications

Ross Smith, Chief Architect

Agenda

What we're here to talk about

- Major Architecture Topics
- What is the ELK Stack?
- Opportunities



1

Major Architecture Topics

Major Architecture Topics

1. Data Loading

- Data Cleaning vs. Upstream Dependencies (Addresses, DISTRIB/Multiple Product Sources)
- Re-Using Data Across Systems

2. Deployment Pipeline

- IT Controlled Process vs. DS-Controlled

3. Logging and Diagnostics

- Local Logging vs. Aggregated

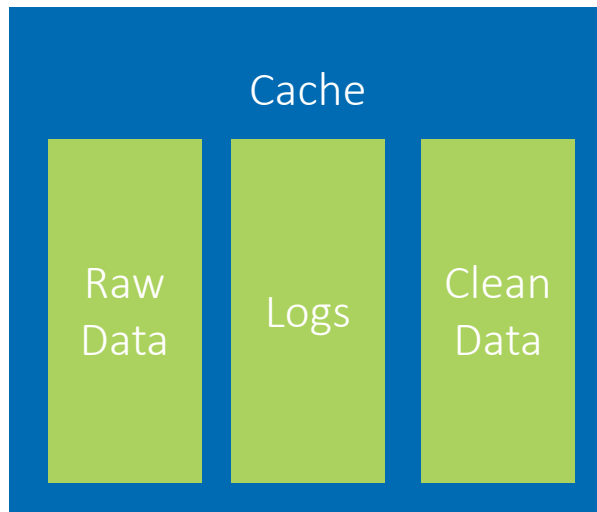
4. User Event Notifications

- In-App vs. Polling vs. Aggregated

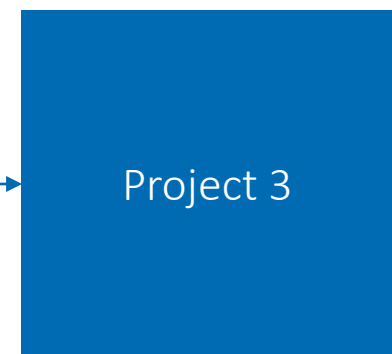
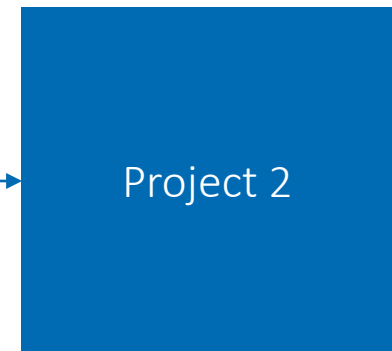
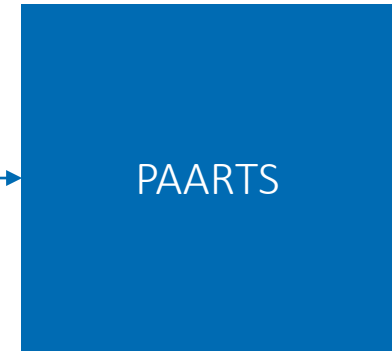
AAR IT
Hardware



AAR DS
Hardware



Proposed Caching Layer Architecture





2

What is the Elastic Stack (ELK)?






















Introduction to Elasticsearch

- An open-source, broadly-distributable, readily-scalable, enterprise-grade search engine. Accessible through an extensive and elaborate API, Elasticsearch can power extremely fast searches that support your data discovery applications.
- Goals: Reduce technical debt involved with upgrading legacy systems for modern architectures
- Allow new applications to read against data from legacy systems by leveraging their existing processes
- Elastic offers flexible models starting with Free Open Source to Multi-Tiered Enterprise Support Subscriptions.



Introduction to Elasticsearch

17 systems in ranking, December 2017

Rank			DBMS	Database Model	Score		
Dec 2017	Nov 2017	Dec 2016			Dec 2017	Nov 2017	Dec 2016
1.	1.	1.	Elasticsearch 	Search engine	119.78	+0.37	+16.51
2.	2.	2.	Solr	Search engine	66.30	-2.86	-2.70
3.	3.	3.	Splunk	Search engine	63.79	-1.08	+8.87
4.	4.	4.	MarkLogic	Multi-model 	11.15	-0.40	+0.82
5.	5.	5.	Sphinx	Search engine	6.03	+0.15	-1.00
6.	6.	 8.	Microsoft Azure Search	Search engine	4.11	+0.24	+2.08
7.	7.	 9.	Algolia	Search engine	3.06	+0.28	+1.44
8.	 9.	 7.	Amazon CloudSearch	Search engine	2.73	+0.37	+0.58
9.	 8.	 6.	Google Search Appliance	Search engine	2.73	-0.03	+0.16
10.	 11.	10.	Xapian	Search engine	0.61	+0.04	+0.11
11.	 10.	 12.	CrateDB	Multi-model 	0.59	-0.07	+0.38
12.	12.	 14.	SearchBlox	Search engine	0.23	-0.01	+0.13
13.	 15.	 15.	DBSight	Search engine	0.00	-0.02	-0.01
14.	 13.	 16.	Exorbyte	Search engine	0.00	-0.06	±0.00
14.	 16.	 11.	Indica	Search engine	0.00	±0.00	-0.36
14.			Manticore Search	Search engine	0.00		
14.	14.	 16.	searchxml	Multi-model 	0.00	-0.04	±0.00

Source: <http://db-engines.com/en/ranking/search+engine>

The Elastic Stack



By combining the massively popular Elasticsearch, Logstash, and Kibana, Elastic has created an end-to-end stack that delivers actionable insights in real time from almost any type of structured and unstructured data source.

The Elastic Stack

Elasticsearch

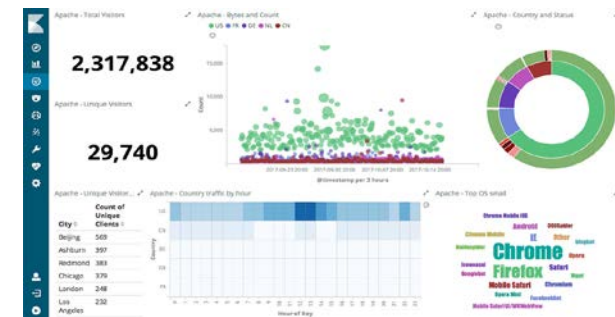
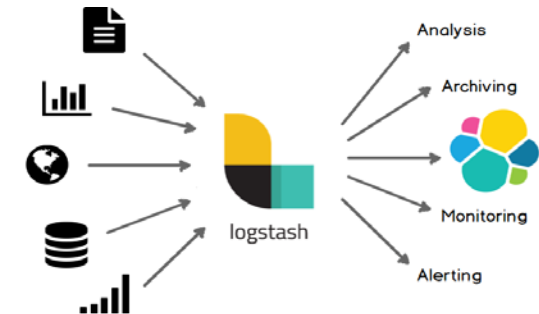
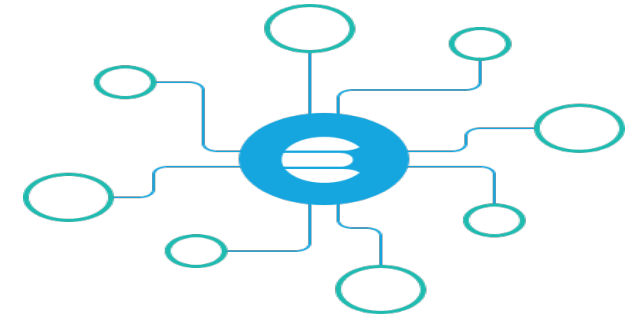
- Horizontally scalable database
- Stores structured or unstructured data
- Freetext search of billions of records in seconds

Logstash

- Smart transactional monitoring
- Loads log files and deltas into Elasticsearch
- Handles high number of high volume, varied data sources (“Beats”)

Kibana

- Light configuration-driven and WYSIWYG dashboard and reporting console
- GUI to monitor Elasticsearch data in real time





3

Opportunities

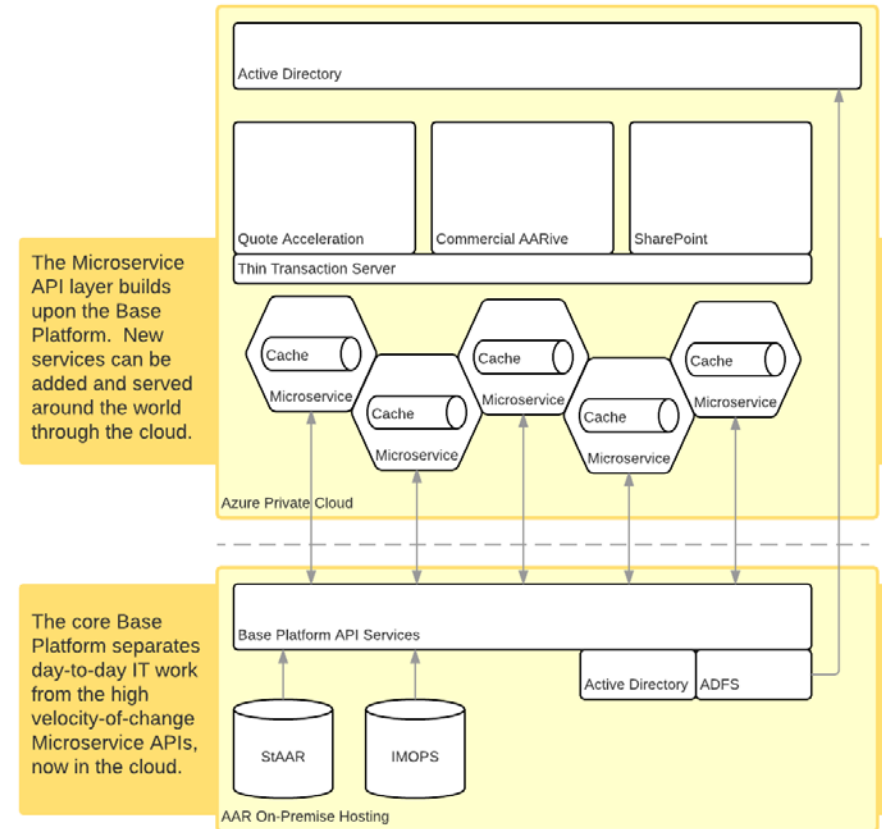
Platform API Services Architecture

Recommended Future State Architecture

- Scalable AAR applications in the cloud, with data supplied by microservices
- Microservices intelligently cache and provide data from IMOPS, preserving the system of authority
- Support microservices with on-premise Platform API Service layer closely tied to IMOPS DB

This is the PITSS architecture proposed in June 2016.

Elasticsearch enables and simplifies this future.



APIs and Analytics

Read-Only APIs

- Reduce the query load on IMOPS
- Exposes default Elasticsearch query APIs, reducing the work in custom development
- Scalable to serve wide variety of use cases with minimal change in tech footprint0

System Analytics

- Monitor uptime and health of both on-premise AAR systems and cloud-based applications
- Designed to integrate with cloud-based infrastructure health checks and management

Proof of Concept



Proof of Concept Estimate

ID	POC Task Description	PITSS Hours	PITSS Cost
1	Finalize project design	20	\$2,800.00
2	Stand up Oracle DB with partial IMOPS schema with journal tables	32	\$4,480.00
3	Write simple CRUD web service against DB tables	16	\$2,240.00
4	Set up Gatling to fire calls against CRUD API	16	\$2,240.00
5	Stand up ELK stack in AWS EC2 or Azure	8	\$1,120.00
6	Write a simple Elasticsearch provider in Java based on journal tables	32	\$4,480.00
7	Write a reconciliation API to test for inconsistency and latency across DB and ELK stack	32	\$4,480.00
8	Write process documentation, ELK API integration recommendations, benchmark results, and implementation recommendations for AAR	40	\$5,600.00
9	Meetings (estimated 10% time)	20	\$2,800.00
	TOTAL	216	\$30,240.00