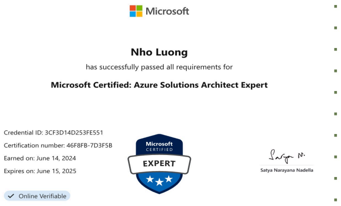
ELK STACK

Author: Nho Luong
Skill: DevOps Engineer Lead



Microsoft

Nho Luong

has successfully passed all requirements for

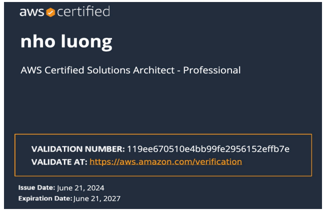
Microsoft Certified: Azure Solutions Architect Expert

Credential ID: 3CF3D14D253FE551
Certification number: 46F8FB-7D3F5B
Earned on: June 14, 2024
Expires on: June 15, 2025

✓ Online Verifiable

Satya Narayana Nadella

Microsoft
CERTIFIED
EXPERT



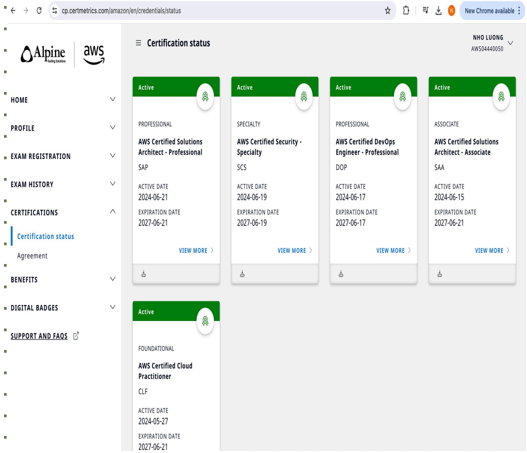
aws certified

nho luong

AWS Certified Solutions Architect - Professional

VALIDATION NUMBER: 119ee670510e4bb99fe2956152effb7e
VALIDATE AT: https://aws.amazon.com/verification

Issue Date: June 21, 2024
Expiration Date: June 21, 2027



cp.certmetrics.com/amazon/en/credentials/status

Certification status

NHO LUONG
AWS04440050

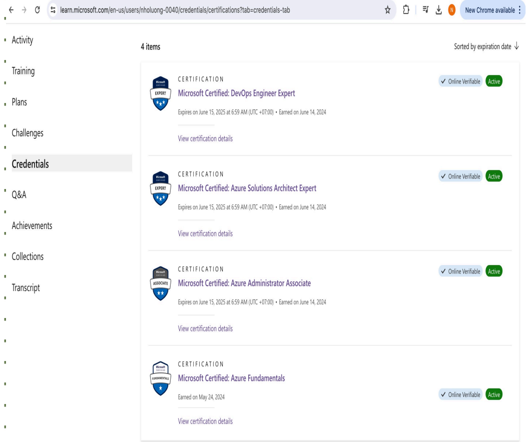
Active
PROFESSIONAL
AWS Certified Solutions Architect - Professional
SAP
ACTIVE DATE
2024-06-21
EXPIRATION DATE
2027-06-21
VIEW MORE

Active
SPECIALTY
AWS Certified Security - Specialty
SCS
ACTIVE DATE
2024-06-19
EXPIRATION DATE
2027-06-19
VIEW MORE

Active
PROFESSIONAL
AWS Certified DevOps Engineer - Professional
DOP
ACTIVE DATE
2024-06-17
EXPIRATION DATE
2027-06-17
VIEW MORE

Active
ASSOCIATE
AWS Certified Solutions Architect - Associate
SAA
ACTIVE DATE
2024-06-15
EXPIRATION DATE
2027-06-21
VIEW MORE

Active
FOUNDATIONAL
AWS Certified Cloud Practitioner
CLF
ACTIVE DATE
2024-05-27
EXPIRATION DATE
2027-06-21



learn.microsoft.com/en-us/users/nholuong-0040/credentials/certifications?tab=credentials-tab

4 items

Sorted by expiration date ↓

CERTIFICATION
Microsoft Certified: DevOps Engineer Expert
Expires on June 15, 2025 at 6:59 AM (UTC +07:00) • Earned on June 14, 2024
View certification details
✓ Online Verifiable Active

CERTIFICATION
Microsoft Certified: Azure Solutions Architect Expert
Expires on June 15, 2025 at 6:59 AM (UTC +07:00) • Earned on June 14, 2024
View certification details
✓ Online Verifiable Active

CERTIFICATION
Microsoft Certified: Azure Administrator Associate
Expires on June 15, 2025 at 6:59 AM (UTC +07:00) • Earned on June 14, 2024
View certification details
✓ Online Verifiable Active

CERTIFICATION
Microsoft Certified: Azure Fundamentals
Earned on May 24, 2024
View certification details
✓ Online Verifiable Active

What is elasticsearch.

Schema-less

Distributed

REST-ful, Document-oriented, and speaks JSON

For searching and analytics

and more…

Author: Nho Luong
Skill: DevOps Engineer Lead

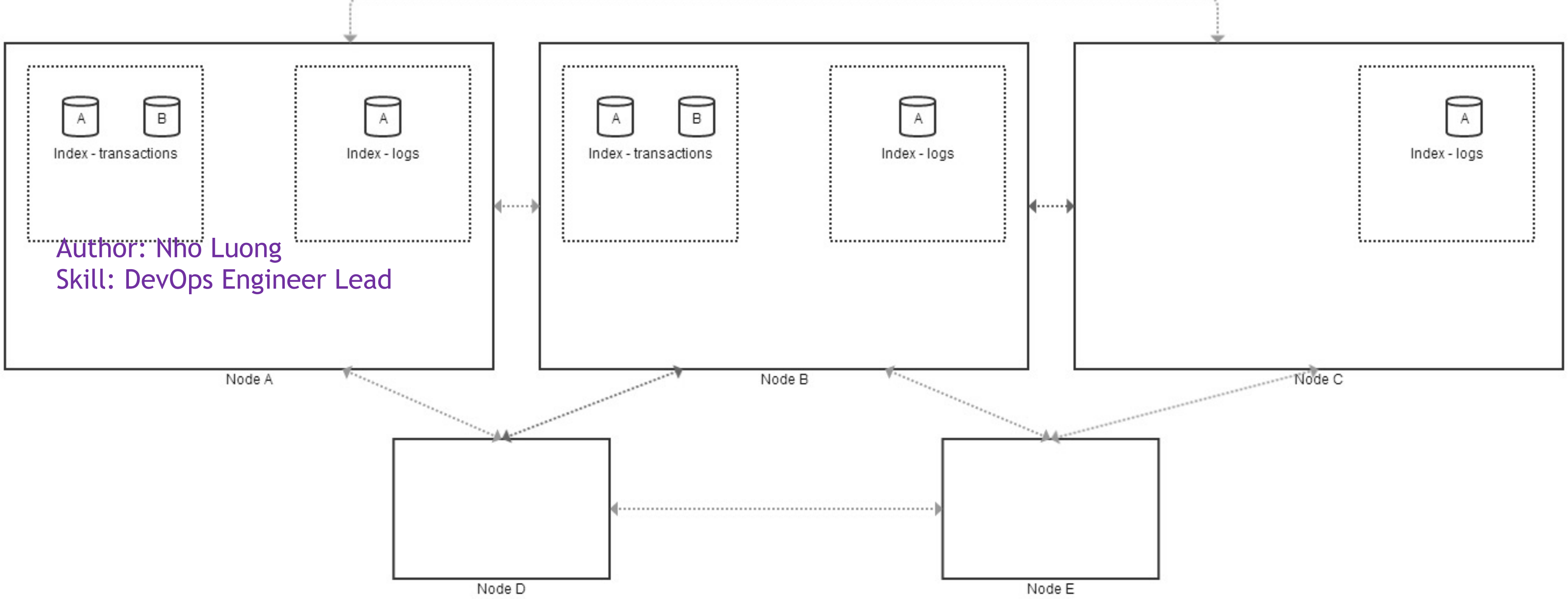
Architecture

Built on top of Apache Lucene

Runs on the JVM

Distributed in nature - cluster can have data, master or load balancing nodes

Highly available and fault tolerant



Legend

⛁ Shard

⬚ Index

□ Node (instance of ElasticSearch process)

Author: Nho Luong
Skill: DevOps Engineer Lead

ElasticSearch - document oriented

- and schema-less

Movies example...

- index a document (PUT and POST)
- check for existence of a document
- retrieve fields
- delete
- update
- update with optimistic concurrency
- update partial
- upsert

Author: Nho Luong
Skill: DevOps Engineer Lead

ElasticSearch - Distributed

- Start a node - open Marvel
  - Data is allocated within the node
- Start another node
  - Highlight data being redistributed to the new node
- Discovery mechanisms - multicast vs. unicast
- Master election, sharding

Author: Nho Luong
Skill: DevOps Engineer Lead

ElasticSearch - RESTful

Get index stats - number of shards (partitions of data), replicas, state and size

Get cluster health - overall health status, number of shards and nodes

Get cluster state - metrics of all indices, settings and mappings of all indices, some metrics, info on all shards in all indices

ElasticSearch - Concepts

Index - highest level bucket to store documents, indicates some physical storage

Type

Relational DB	⇒ Databases	⇒ Tables	⇒ Rows	⇒ Columns
Elasticsearch	⇒ Indices	⇒ Types	⇒ Documents	⇒ Fields

Mapping - the definition of a type (think schema) and how ElasticSearch should analyze, parse and store the fields of this type

Analysis:

first, tokenizing a block of text into individual *terms* suitable for use in an inverted index, then normalizing these terms into a standard form to improve their "searchability" or *recall*.

Author: Nho Luong
Skill: DevOps Engineer Lead

ElasticSearch - search and analytics

Search

- Structured search - working with exact values, between date ranges, numbers, enumerated strings, etc…
- Full-text search - natural language and other text, relevance is usually concern here instead of exact matches

ElasticSearch - search in depth

Analyzes all documents and keeps an inverted index data structure for fast matching

Inverted index example

Document: “The quick brown fox jumped over the lazy dog.”

```
Term    | Doc 1 |
--------------------------
The     |   x   |
quick   |   x   |
brown   |   x   |
fox     |   x   |
jumped  |   x   |
over    |   x   |
the     |   x   |
lazy    |   x   |
dog     |   x   |
```

Author: Nho Luong
Skill: DevOps Engineer Lead

Inverted index example

Document: “The quick brown fox jumped over the lazy dog.”
Document: “Quick, the fox, was lazy.”

```
Term    | Doc 1 | Doc 2 |
--------------------------
The     |   x   |       |
quick   |   x   |       |
brown   |   x   |       |
fox     |   x   |   x   |
jumped  |   x   |       |
over    |   x   |       |
the     |   x   |   x   |
lazy    |   x   |   x   |
dog     |   x   |       |
Quick   |       |   x   |
was     |       |   x   |
```

Author: Nho Luong
Skill: DevOps Engineer Lead

ElasticSearch - Query DSL

Simple search
Compound search
Query vs. Filters
Range filter
Aggregations
Significant terms (‘the uncommonly common’)

Author: Nho Luong
Skill: DevOps Engineer Lead

ElasticSearch - search examples

NFL data - fuzzy description, more like this
NFL data - bool query
NFL data - all IND offense
NFL data - aggregations - average down and distance, 2nd half yard to go

ElasticSearch - search examples

NFL 2013 data - get touchdowns by quarter

NFL 2013 data - get significant terms in description by teams

Author: Nho Luong
Skill: DevOps Engineer Lead

What is LogStash

Data import/export tool for time series and log data

Design inspired by Unix utilities which pipe in/out to each other

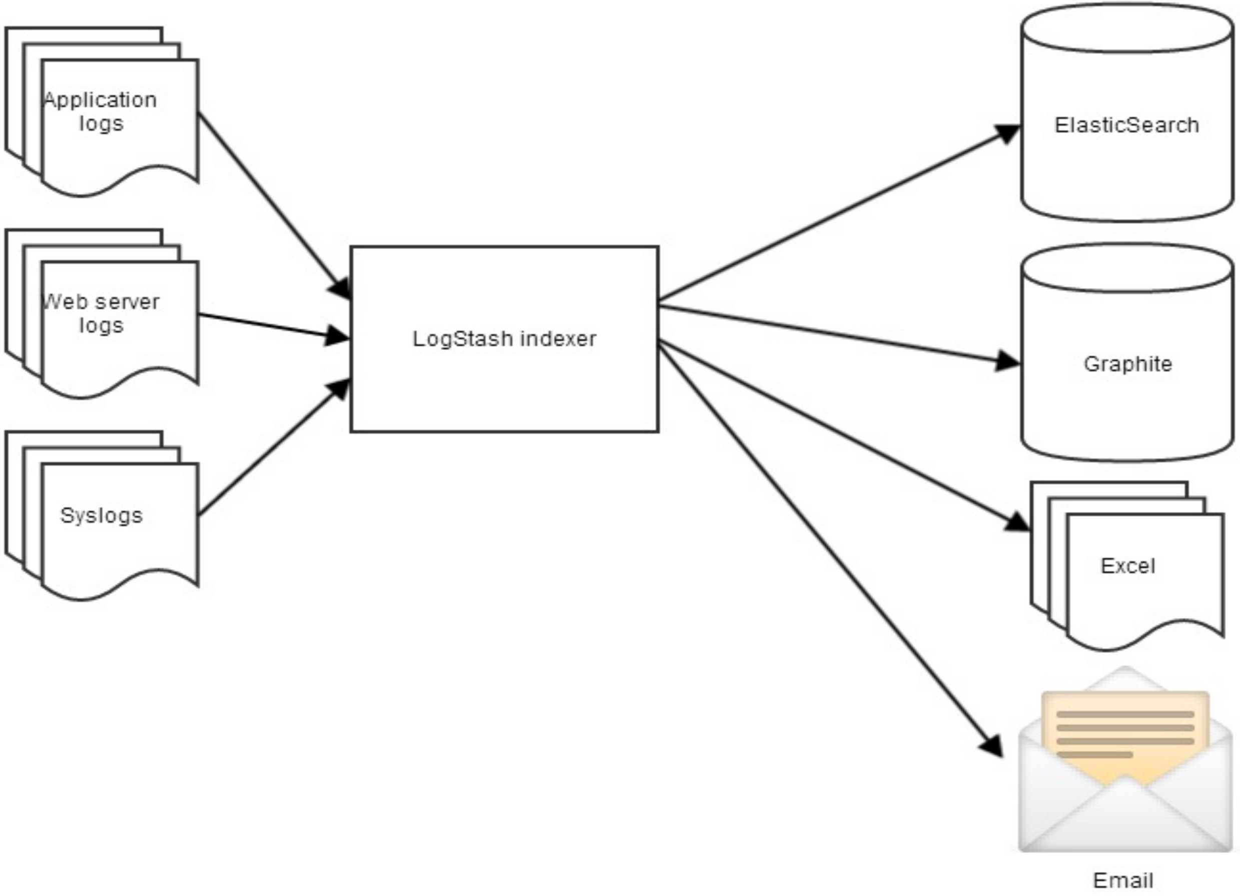
What problem does it solve?

How to parse and analyze log data from many sources?

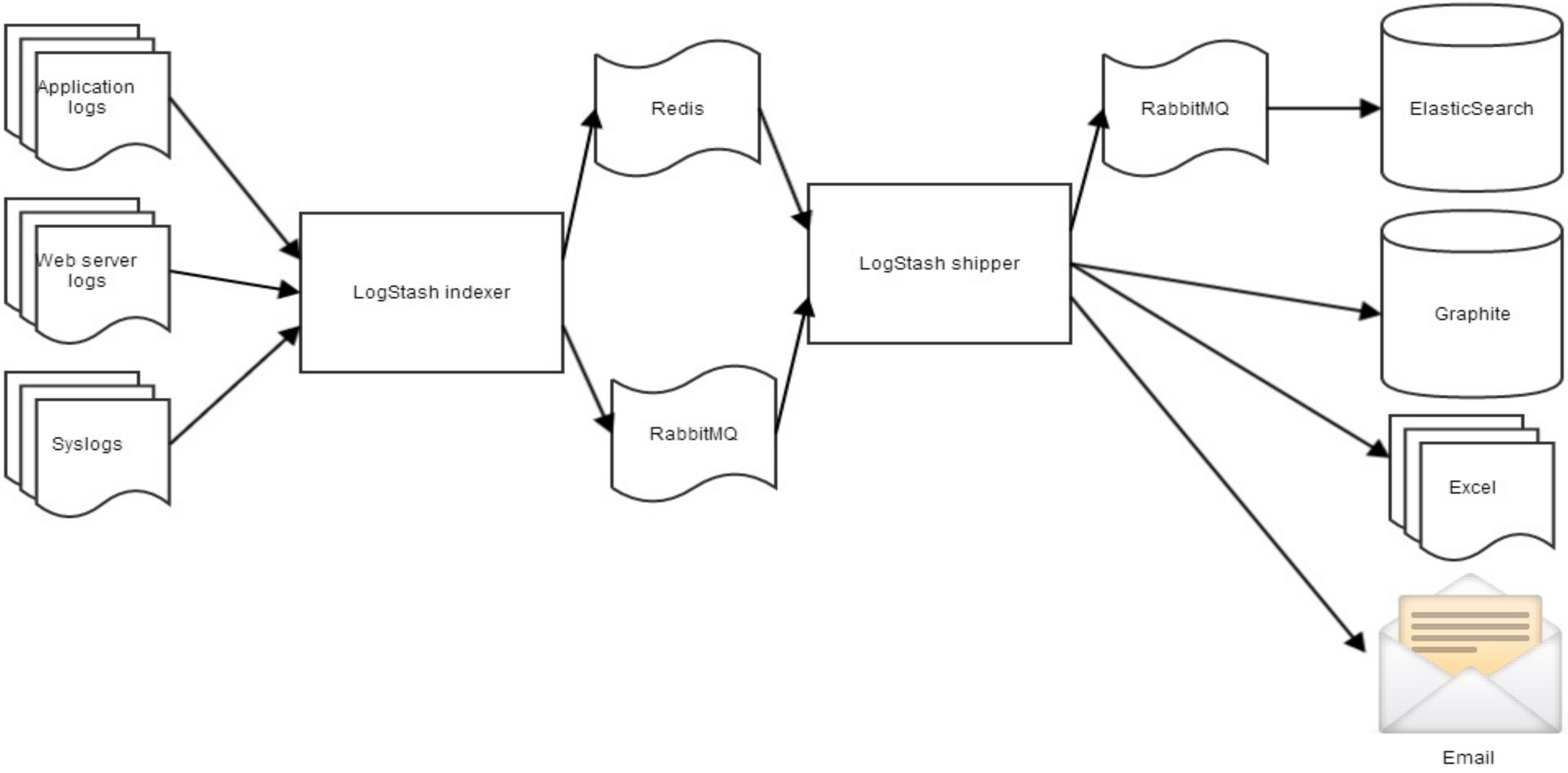
```
Mar 12 12:00:08 server2 rcd[308]: Loaded 12 packages in 'ximian-red-carpet|351' (0.01878 seconds)
[2014-05-06 08:04:00.333] [ERROR] - core - bad thing happened
[Wed Oct 11 14:32:52 2000] [error] [client 127.0.0.1] client denied by server configuration: /export/home/live/ap/htdocs/test
```

LogStash - example use cases

- Import of JSON to ES by dropping files into a folder
- Parse webserver access files across multiple servers, calculate response times and chart
- Parse application logs and send emails when an error occurs
- Stream application log data across many servers to a single log dashboard
- Drop a file into a folder to be ingested and aggregated into centralized log database



Author: Nho Luong
Skill: DevOps Engineer Lead



Author: Nho Luong
Skill: DevOps Engineer Lead

LogStash - example configuration

input
filter
output

LogStash - demo

output CPU load to CSV (load-avg.conf)

Stream tweets to ElasticSearch (twitter.conf)

Parse NodeJS server logs to ElasticSearch (mp.conf)

Author: Nho Luong
Skill: DevOps Engineer Lead

What is Kibana

Dashboard tool for data in ElasticSearch

Highly configurable/customizable, build panels with user defined charts, tables, etc...

Built on AngularJS

Author: Nho Luong
Skill: DevOps Engineer Lead

Kibana - demos

- NFL stats dashboard
- Tweets dashboard
- ElasticSearch Marvel

Thank You

Author: Nho Luong
Skill: DevOps Engineer Lead