Simplified Sizing : Introducing New Splunk Sizing Calculator

Jeff Champagne – Principal Architect
Mustafa Ahamed – Principal Architect

October 2018 | Version 1.0

Forward-Looking Statements

During the course of this presentation, we may make forward-looking statements regarding future events or the expected performance of the company. We caution you that such statements reflect our current expectations and estimates based on factors currently known to us and that actual events or results could differ materially. For important factors that may cause actual results to differ from those contained in our forward-looking statements, please review our filings with the SEC.

The forward-looking statements made in this presentation are being made as of the time and date of its live presentation. If reviewed after its live presentation, this presentation may not contain current or accurate information. We do not assume any obligation to update any forward-looking statements we may make. In addition, any information about our roadmap outlines our general product direction and is subject to change at any time without notice. It is for informational purposes only and shall not be incorporated into any contract or other commitment. Splunk undertakes no obligation either to develop the features or functionality described or to include any such feature or functionality in a future release.

Is This Session For Me?

What will I learn?

- Are you a Splunk Admin or Architect?
- Have you wondered if you can get more out of your existing hardware?
- Do you want to learn more about how Splunk does Benchmarking?

Agenda

- Introductions
- Primer on Sizing & Benchmarking
- Splunk Enterprise Sizing Calculator
- ES Sizing calculator
- Next Steps
- Q&A

10s, 10s, 10s Across the Board!

Rate Our Session Please

Jeff Champagne

Principal Architect, Core Platform

- Member of Global Field Architecture team
- Leads Voice of the Customer program
- Member of the Splunk Architecture Council
- Background in enterprise architecture & financial services/trading systems
- Former customer, joined Splunk in 2014

New York City

Mustafa Ahamed

Director, Platform Architecture

- Member of Global Field Architecture team
  - Focused on APAC
- Led Splunk Enterprise Product Mgmt for 6 years
- Launched features like SH/Index Clustering and Pipeline Parallelization
- Joined Splunk in 2011

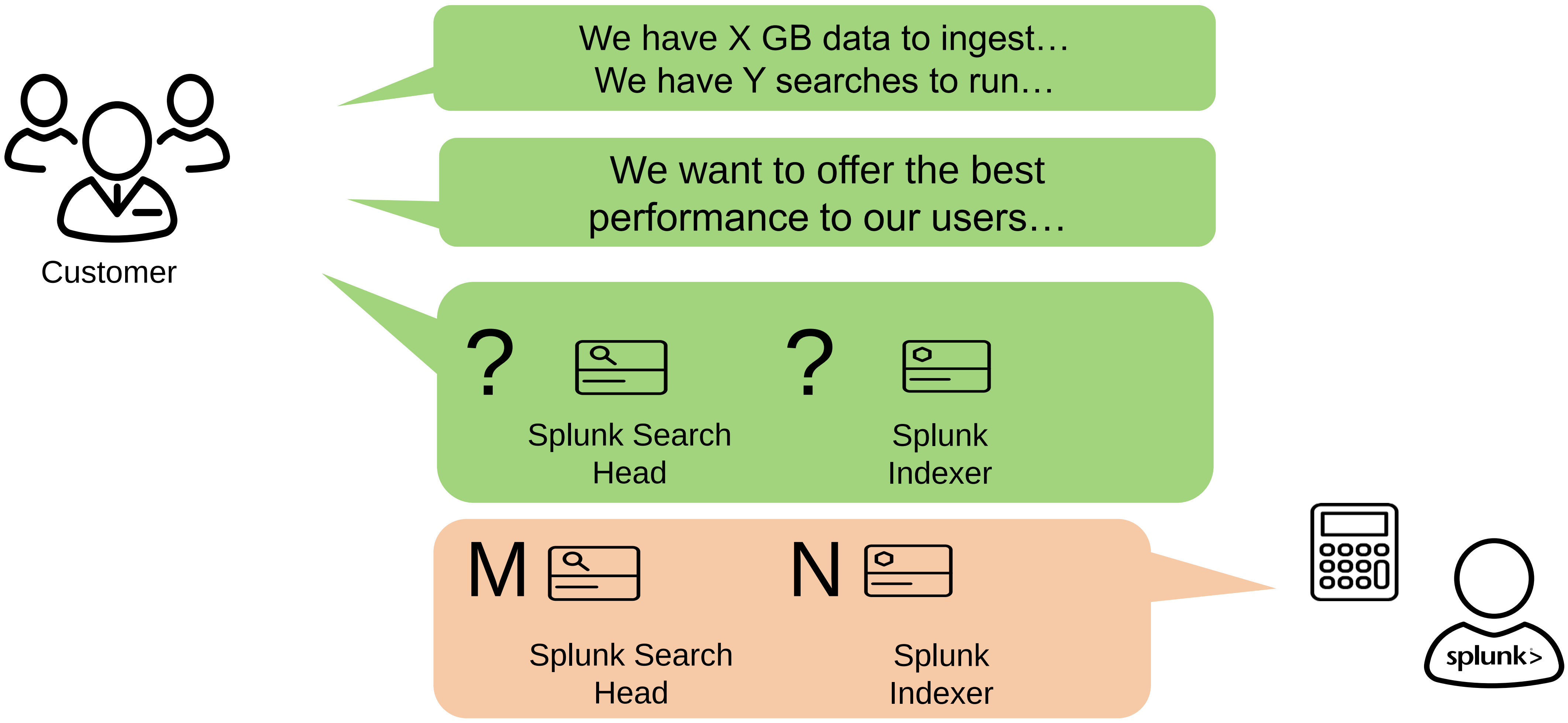
Chennai

Benchmarks & Sizing

Making Sense of the Madness

Introduction to Sizing

Finding the right hardware fit for your workload



Customer

We have X GB data to ingest…
We have Y searches to run…

We want to offer the best performance to our users…

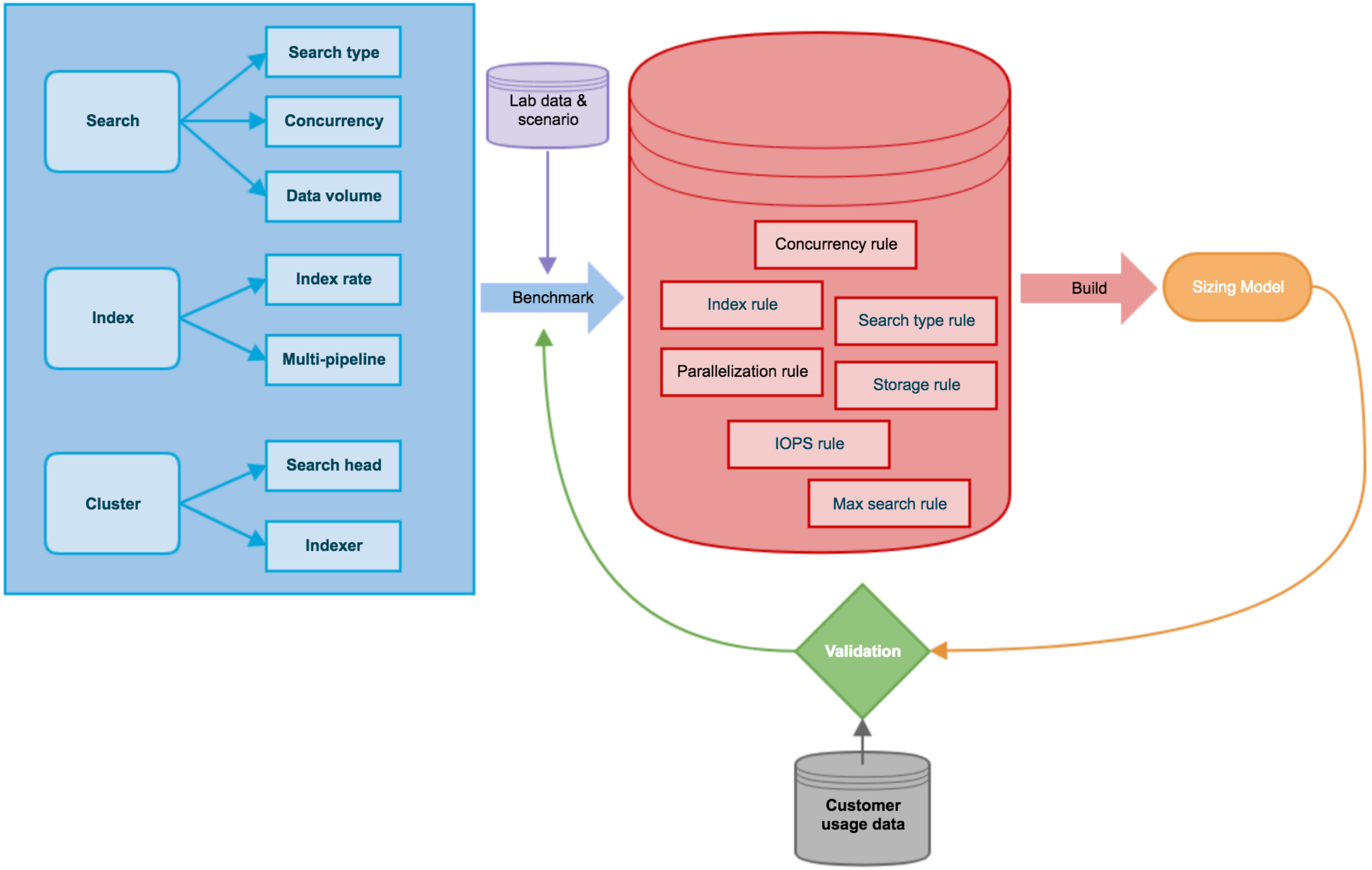
? Splunk Search Head

? Splunk Indexer

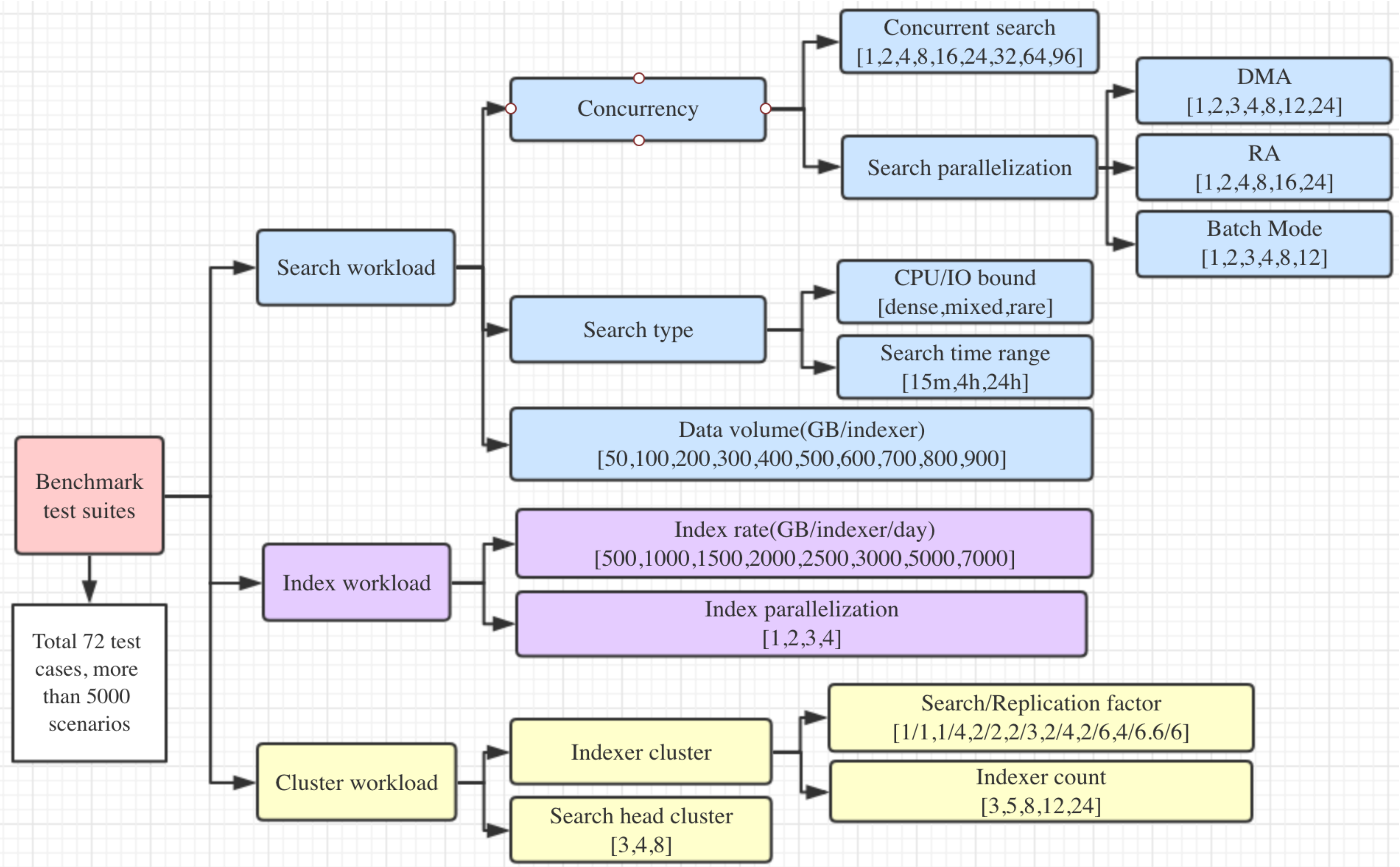
M Splunk Search Head

N Splunk Indexer

The Science Behind Benchmark Testing



Benchmark Test Suites



Benchmark test suites

Total 72 test cases, more than 5000 scenarios

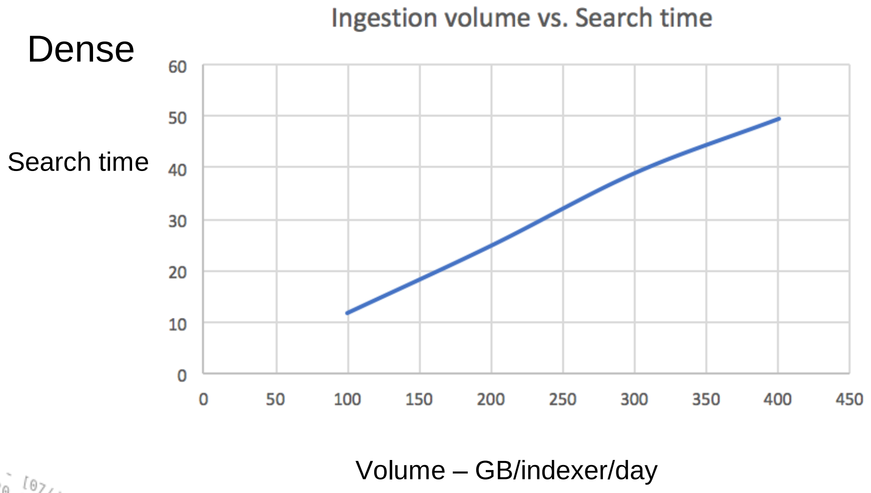
- Search workload
  - Concurrency
    - Concurrent search [1,2,4,8,16,24,32,64,96]
    - Search parallelization
      - DMA [1,2,3,4,8,12,24]
      - RA [1,2,4,8,16,24]
      - Batch Mode [1,2,3,4,8,12]
  - Search type
    - CPU/IO bound [dense,mixed,rare]
    - Search time range [15m,4h,24h]
  - Data volume(GB/indexer) [50,100,200,300,400,500,600,700,800,900]
- Index workload
  - Index rate(GB/indexer/day) [500,1000,1500,2000,2500,3000,5000,7000]
  - Index parallelization [1,2,3,4]
- Cluster workload
  - Indexer cluster
    - Search/Replication factor [1/1,1/4,2/2,2/3,2/4,2/6,4/6.6/6]
    - Indexer count [3,5,8,12,24]
  - Search head cluster [3,4,8]

Search Workload – Search Type

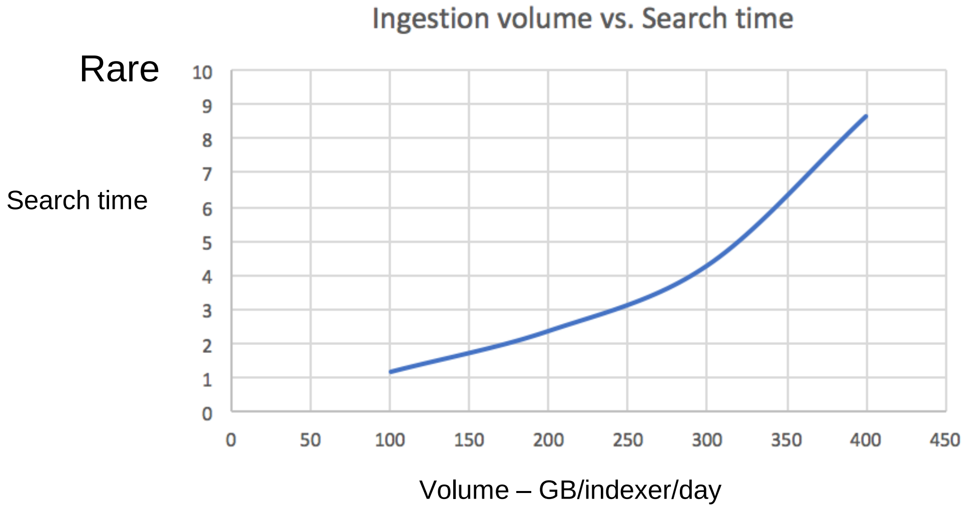
- Log
  - Syslog
- Search
  - Dense – CPU bound
  - Mixed – between dense and rare
  - Rare – IO bound
- Search Time Range (Based on Cloud Perf Benchmarks CloudFY19)
  - 15m – short search
  - 4h – moderate search
  - 24h – long search

Search Workload – Volume

- Dense
  - Dense searches (CPU bound) show a linear relationship with volume
- Rare
  - Rare searches (IO bound) show 2 kinds of linear relationship with volume
  - Memory can cache volume, low IOPS, memory not enough to cache volume, IOPS grows fast
- Mixed
  - mixed_search_time = dense_search_time * dense_ratio + rare_search_time * rare_ratio



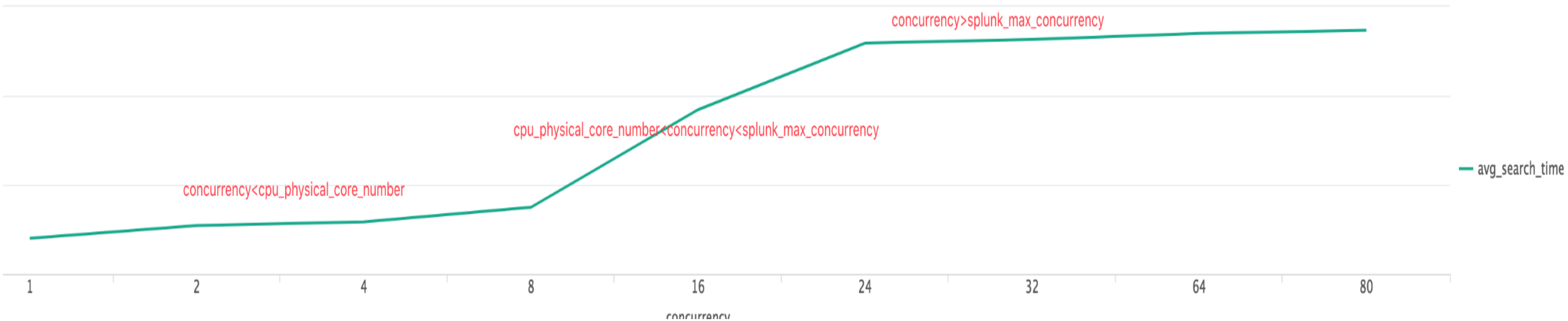
Test env: aws i3.4xlarge env



Test env: aws i3.4xlarge env

Search Workload - Concurrency

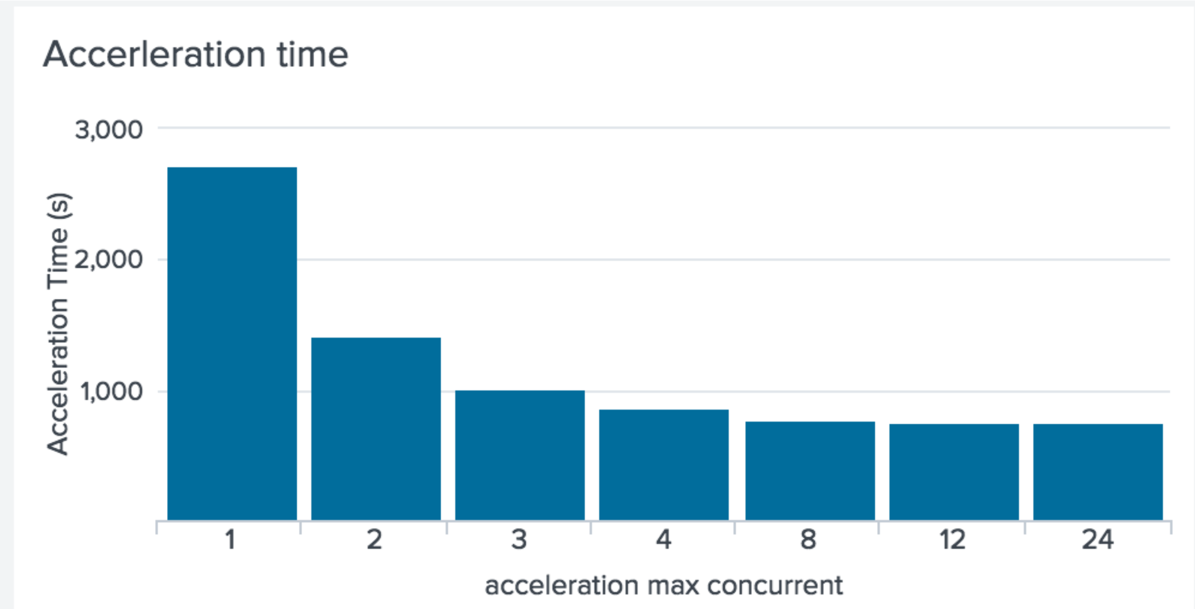
- Concurrency(Search concurrency) vs. CPU core
  - Stable stage. Search concurrency less than CPU core number.
  - Pressure stage. Search concurrency between CPU core number and max search concurrency of Splunk.
  - Max stage. Search concurrency above max search concurrency of Splunk.



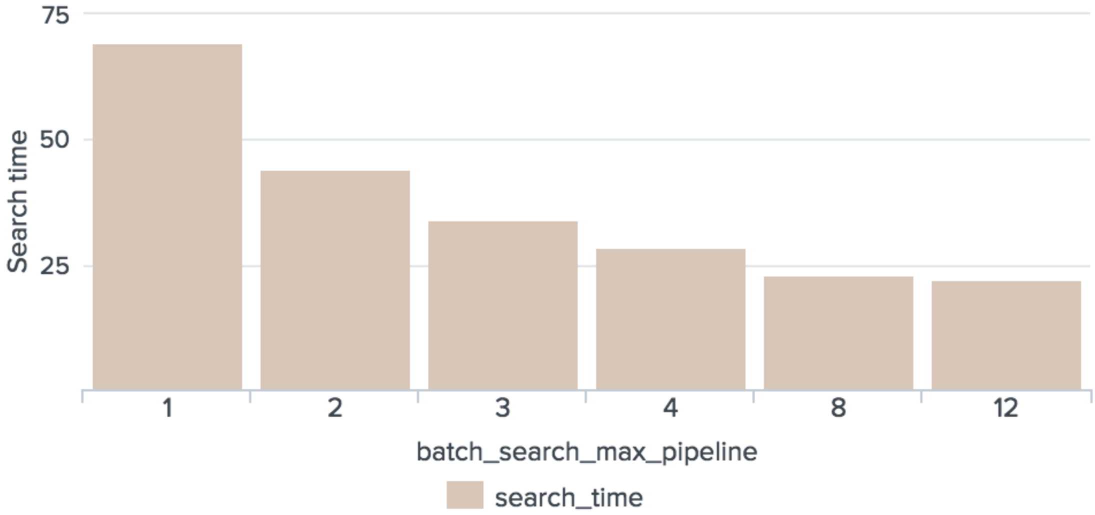
Test result of aws i3.4xlarge
1 search head, 3 indexers as a culster
Dense workload

Search Workload - Parallelization

- Parallelization(the option in Splunk conf file) vs. CPU core
  - Parallelization < Available CPU core. Acceleration is obvious and diminishing.
  - Parallelization >= Available CPU core. Acceleration is not obvious.



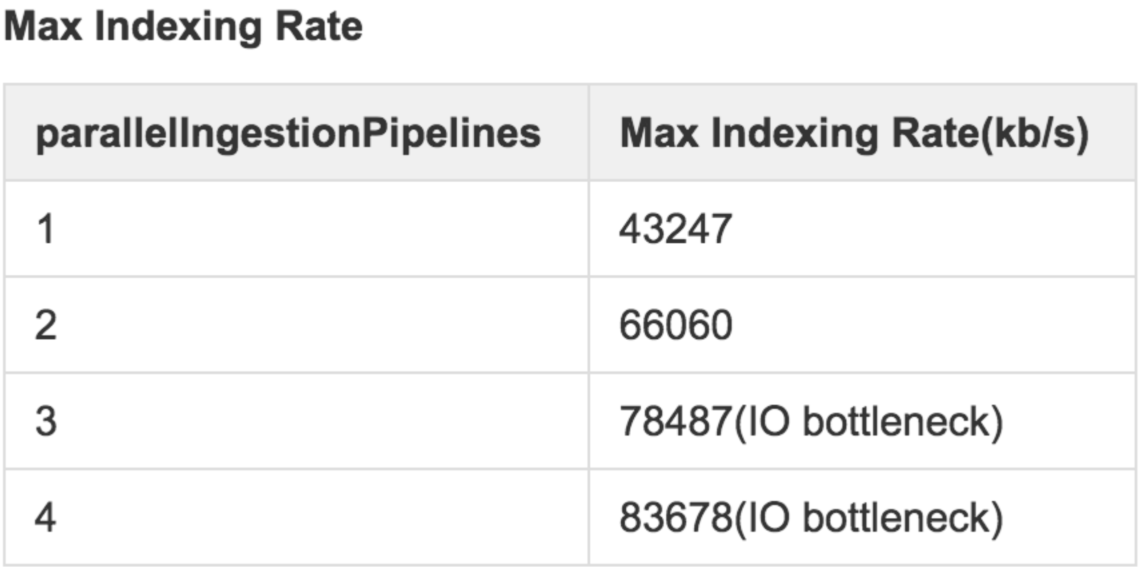
- Test result of aws i3.4xlarge env
- DMA max concurrent test



- Test result of aws i3.4xlarge env
- Batch search max pipeline test

Indexing Workload

- Basic (result is from benchmark test in aws m5.xlarge env)
  - CPU utilization has a linear correlation with event volume, more events more CPU usage.
  - 500G/indexer/day will lead to around 1 additional CPU core.
- Parallelization (result is from benchmark test in aws m5.xlarge env)
  - Adding 1 more parallel ingestion pipeline will increase the max indexing rate increases around 1.53 times, lead to around 3-4 additional CPU cores and about 200-300 IOPS utilization.

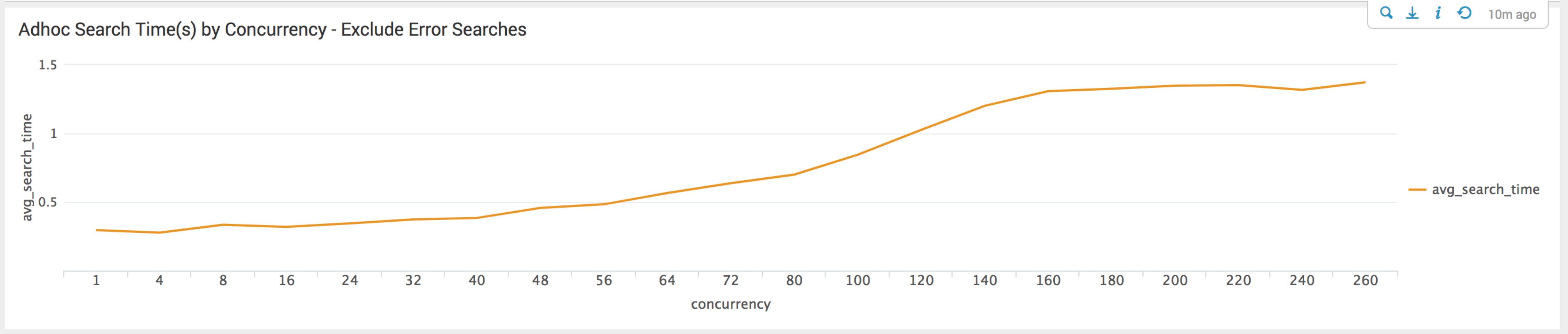


Max Indexing Rate

parallelIngestionPipelines	Max Indexing Rate(kb/s)
1	43247
2	66060
3	78487(IO bottleneck)
4	83678(IO bottleneck)

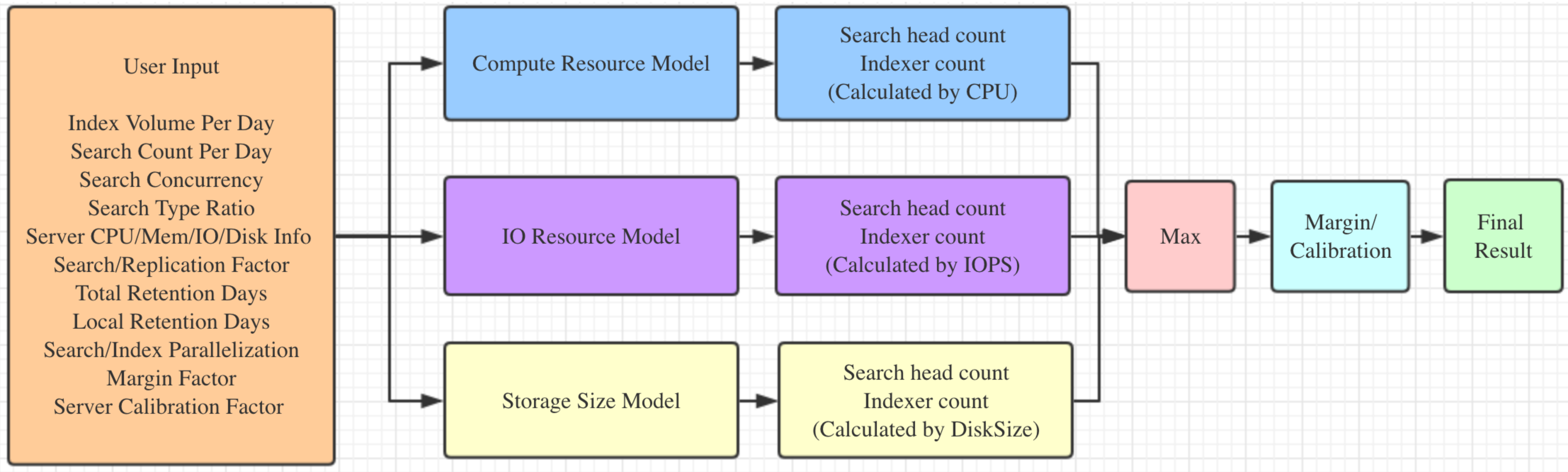
Cluster Workload

- Indexer cluster workload(Test different search/replication factor）
  - Disk IO write Ops and network throughput nearly doubled when double the search/replication factor.
  - Search time is almost the same when double the search/replicate factor.
- Search head cluster workload(Run max capacity searches on each search head in a cluster)
  - Almost the same trend with single search head.
  - The max concurrency of search is sum(max_search_each_sh).



- Test result of aws m5x.large env
- 3 search heads as a search head cluster
- 6 indexers as a indexer cluster
- Mixed workload

Sizing Model Introduction



User Input

Index Volume Per Day
Search Count Per Day
Search Concurrency
Search Type Ratio
Server CPU/Mem/IO/Disk Info
Search/Replication Factor
Total Retention Days
Local Retention Days
Search/Index Parallelization
Margin Factor
Server Calibration Factor

Compute Resource Model

Search head count
Indexer count
(Calculated by CPU)

IO Resource Model

Search head count
Indexer count
(Calculated by IOPS)

Storage Size Model

Search head count
Indexer count
(Calculated by DiskSize)

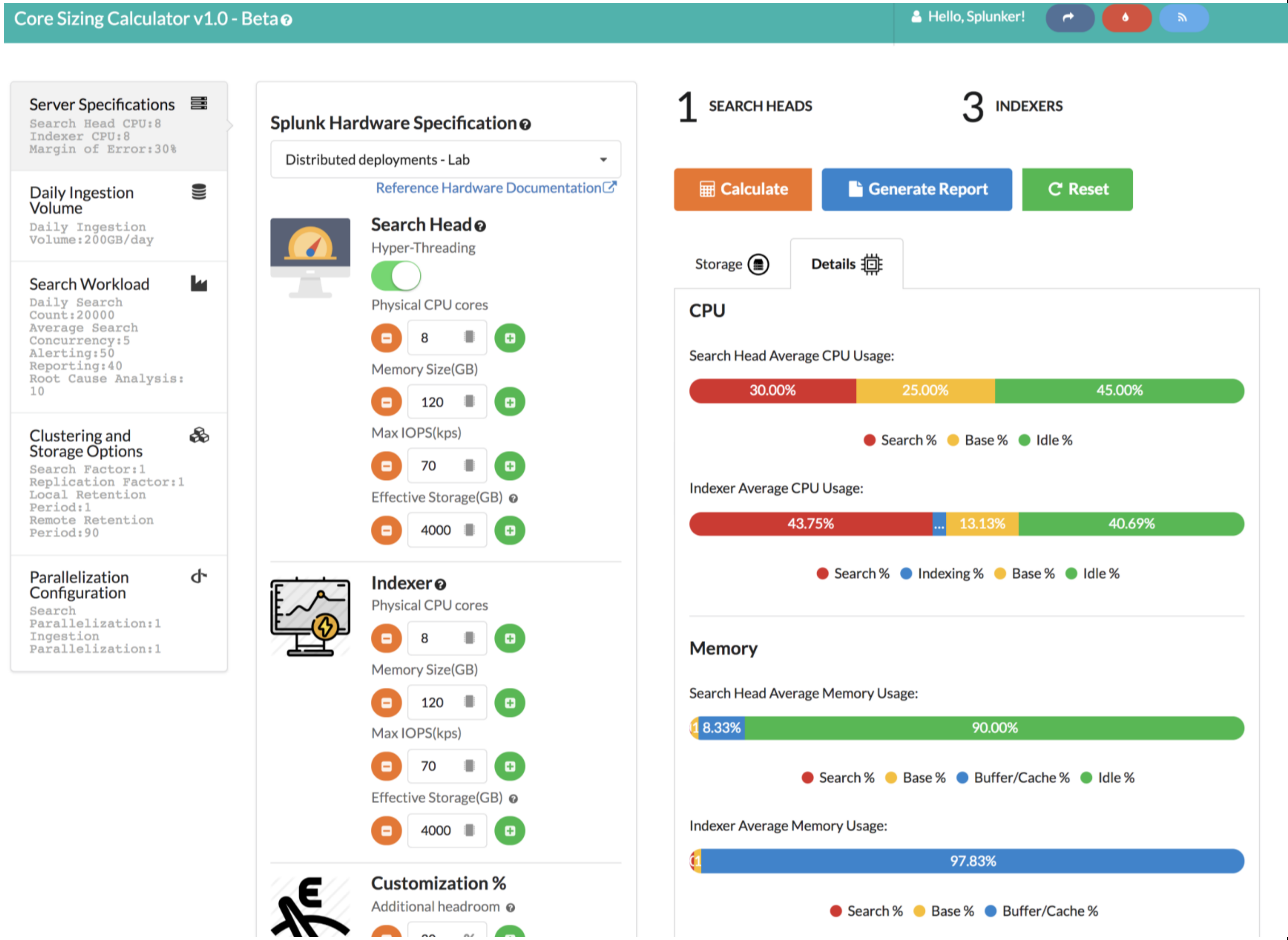
Max

Margin/
Calibration

Final
Result

Demo

Core Sizing Calculator



Core Sizing Calculator v1.0 - Beta

Hello, Splunker!

Server Specifications
Search Head CPU:8
Indexer CPU:8
Margin of Error:30%

Daily Ingestion Volume
Daily Ingestion Volume:200GB/day

Search Workload
Daily Search Count:20000
Average Search Concurrency:5
Alerting:50
Reporting:40
Root Cause Analysis:10

Clustering and Storage Options
Search Factor:1
Replication Factor:1
Local Retention Period:1
Remote Retention Period:90

Parallelization Configuration
Search Parallelization:1
Ingestion Parallelization:1

Splunk Hardware Specification

Distributed deployments - Lab

Reference Hardware Documentation

Search Head

Hyper-Threading: on

Physical CPU cores: 8

Memory Size(GB): 120

Max IOPS(kps): 70

Effective Storage(GB): 4000

Indexer

Physical CPU cores: 8

Memory Size(GB): 120

Max IOPS(kps): 70

Effective Storage(GB): 4000

Customization %

Additional headroom

1 SEARCH HEADS

3 INDEXERS

CPU

Search Head Average CPU Usage:
30.00% | 25.00% | 45.00%
● Search % ● Base % ● Idle %

Indexer Average CPU Usage:
43.75% | ... | 13.13% | 40.69%
● Search % ● Indexing % ● Base % ● Idle %

Memory

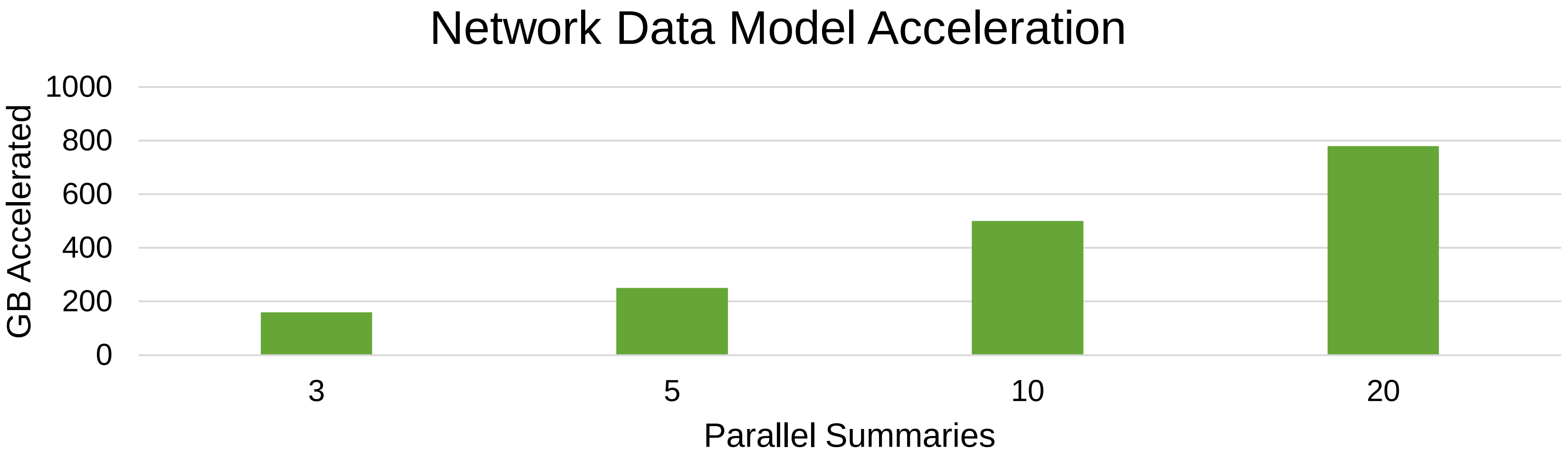
Search Head Average Memory Usage:
1 | 8.33% | 90.00%
● Search % ● Base % ● Buffer/Cache % ● Idle %

Indexer Average Memory Usage:
1 | 97.83%
● Search % ● Base % ● Buffer/Cache %

ES Sizing Calculator

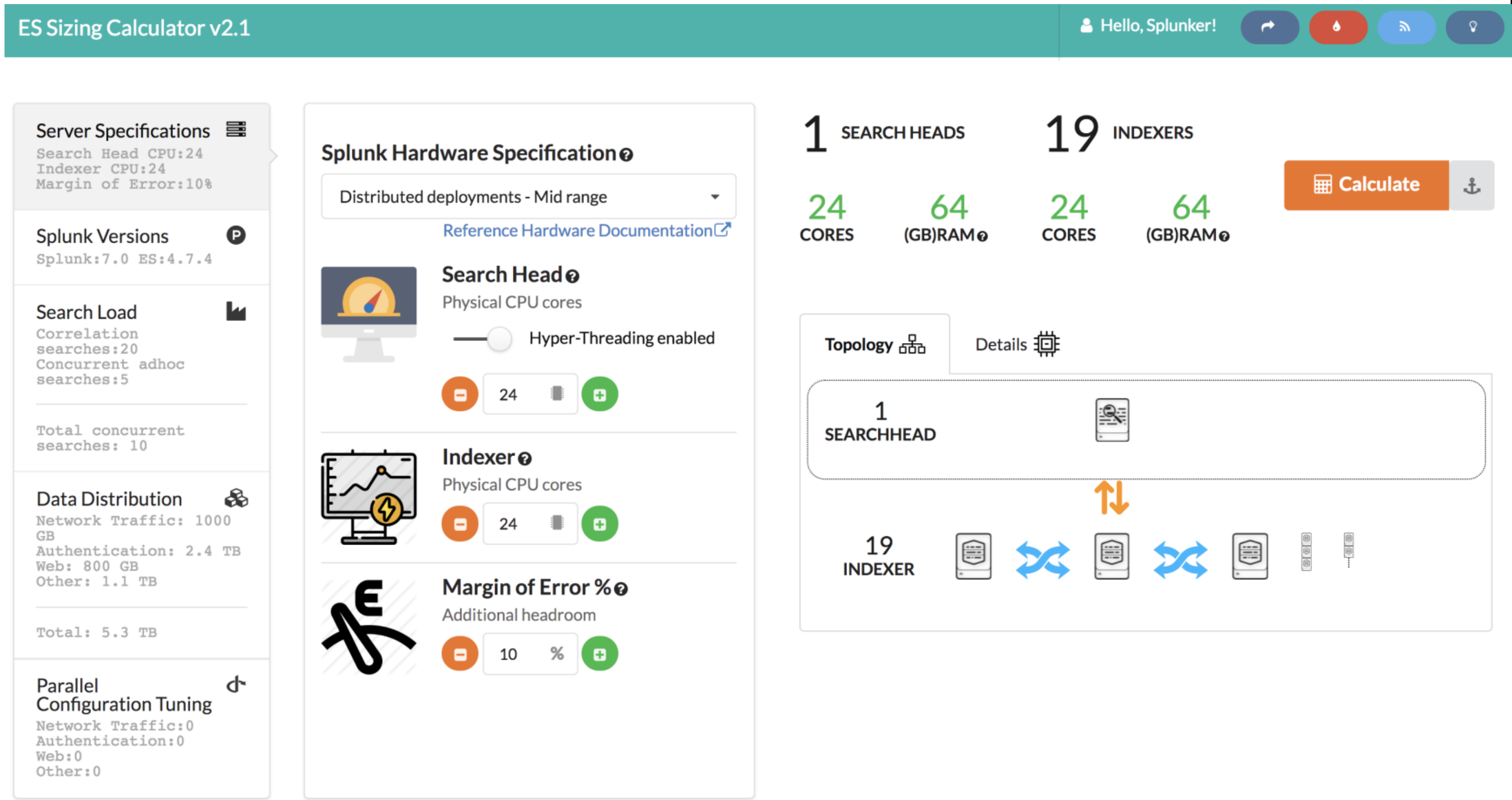
ES Sizing Challenges

- How much data can we ingest per indexer?
- How much data can we accelerate per indexer?
  - Some Data Models perform better than others
    - DM complexity
    - Cardinality of dataset
- How many searches can we run concurrently?



Demo

ES Sizing Calculator



Server Specifications
Search Head CPU:24
Indexer CPU:24
Margin of Error:10%

Splunk Versions
Splunk:7.0 ES:4.7.4

Search Load
Correlation searches:20
Concurrent adhoc searches:5

Total concurrent searches: 10

Data Distribution
Network Traffic: 1000 GB
Authentication: 2.4 TB
Web: 800 GB
Other: 1.1 TB

Total: 5.3 TB

Parallel Configuration Tuning
Network Traffic:0
Authentication:0
Web:0
Other:0

Splunk Hardware Specification
Distributed deployments - Mid range
Reference Hardware Documentation

Search Head
Physical CPU cores
Hyper-Threading enabled
24

Indexer
Physical CPU cores
24

Margin of Error %
Additional headroom
10 %

1 SEARCH HEADS
24 CORES
64 (GB)RAM

19 INDEXERS
24 CORES
64 (GB)RAM

What’s Next?

How do I get my hands on this thing?

- Live Demos @ the Customer Success Studio
  - Source=*Pavilion
- Your account teams have access to this tool now
- Public launch coming later this year
  - Individual tools at first
  - Combined Core + ES + ITSI calculator coming in 2019

Don't forget to rate this session
in the .conf18 mobile app