Monitoring & Logging in High Scale Environments

Kobi Biton

Challenges

- 1000+ Machines and growing...
- 70+ Services with complex dependencies
- 3 Data Center
- High Profile Customers
- Scale

SO WHY?

The Business side...

Every X Minute of downtime

- Direct Revenue Lost
- Impacts Reputation
- Customers lose confidence

Few more Good Reasons

- Capacity Planning
- Identify Bottlenecks Across Our Services
- Shape Long Term Production Strategies
- Evaluate Products Performance

Our Goals

- Monitoring Systems As Prevention Tools
- Low MTTR times when Prevention was not possible

MTTR

Some Numbers...

- We process ~ 1.2M System & Application Metrics Per Minute
- Up to 2500 Log Metrics Per Sec
- ~ 1000 Machines a.k.a Producers
- 3 Data Centers
- Metrics Can be added by all users

Principles we believe In

- Collect your metrics once
- Use a middle tier with message brokers
- If It moves measure It
- Create Dashboards (Its so easy with graphitus)
- Push your metrics rather then pull
- Technology will break so monitor It!
- Use Open Source but consider the home grown Development effort

Our Dear Friends

Nagios®

Graphite

collectd
The system statistics collection daemon

RabbitMQ™
Messaging that just works

Commercial

- KeyNote
- New Relic
- Boundary

Architecture Components
Producers & Shippers Tier

- System Metrics – collectd to RMQ
- Application Metric – Coda Hale's yammer metrics library
- Event BUS – JAVA/Python/CLI to Logstash
- System, Application Logs – Logstash to RMQ
- Application Logs – Log4j to Graylog

Architecture Components
Middle (Transport) Tier

- Message Broker – RabbitMQ 3.0.1
- Fanout Exchanges (Why?) - Logs/Metrics
- RabbitMQ Sharding – RMQ server per type
- Broker failover - DNS Based

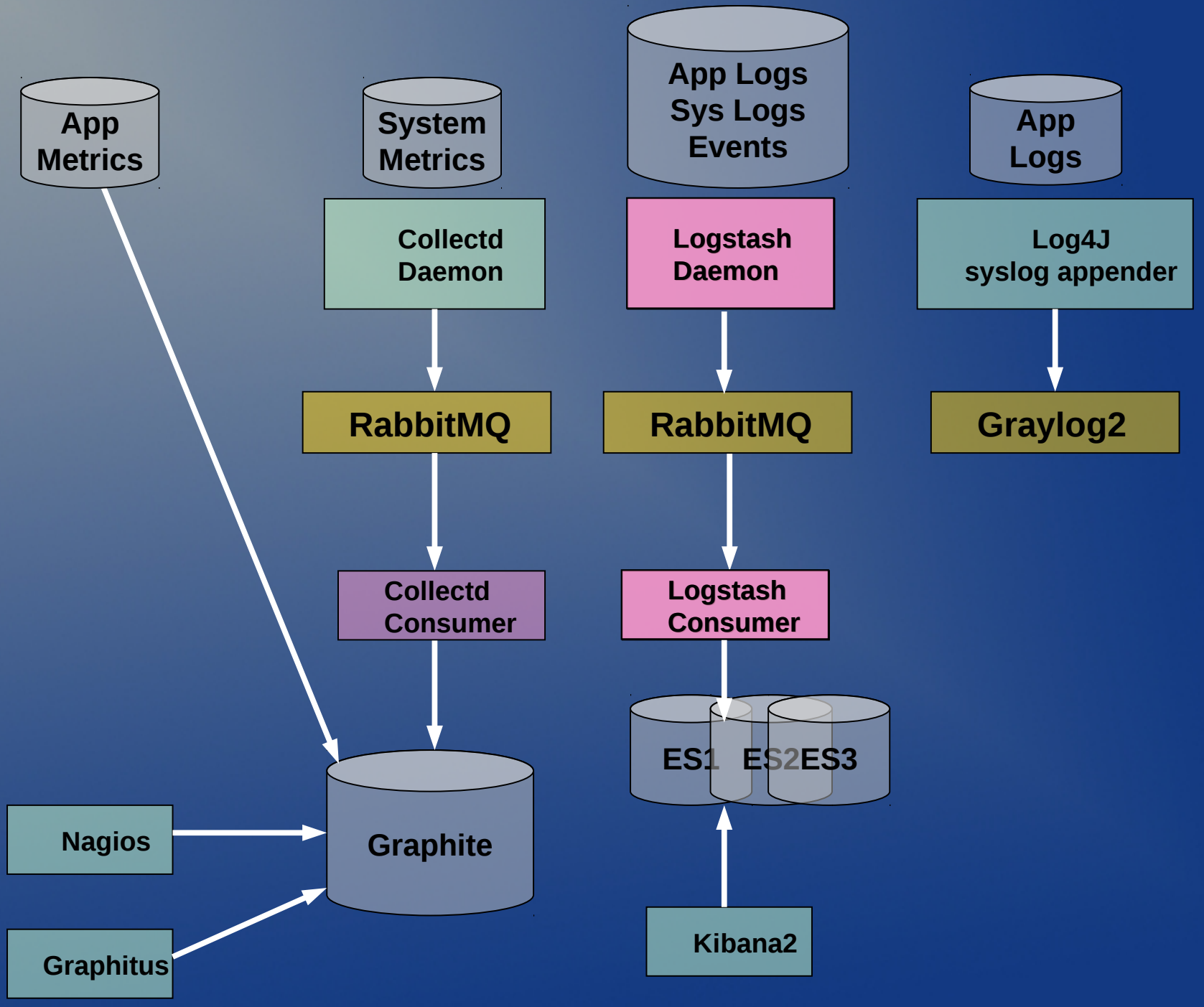
Architecture Components
Consumers Tier

- System Metrics – Collectd Consumes to Graphite
- Logs – Logstash Consumes to Elasticsearch Consumer on every ES Cluster Node
- App Metrics – Shifting from Direct Line Feeder To Consume
- Queues and Exchanges are Dynamically created (will not survive broker restarts)

Architecture Components
Backend Tier

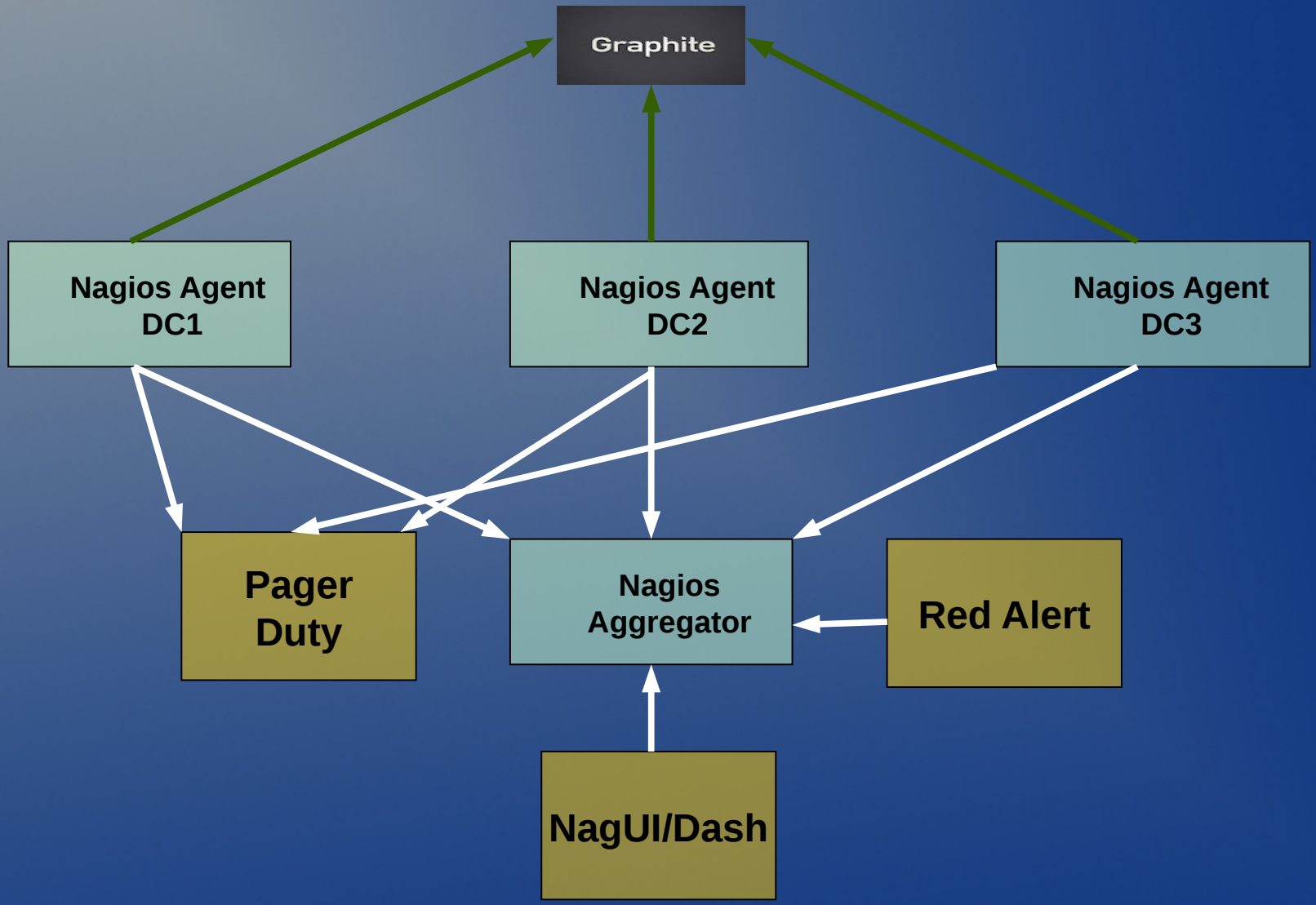
- Graphite Servers
- Nagios , Using Perl plug-in to get the metric results
- Elasticsearch , Kibana2 dashboard for searches
- Graylog2, Mainly for Alert Streams
- Graphitus, Graphite Dashboards

Data Flow

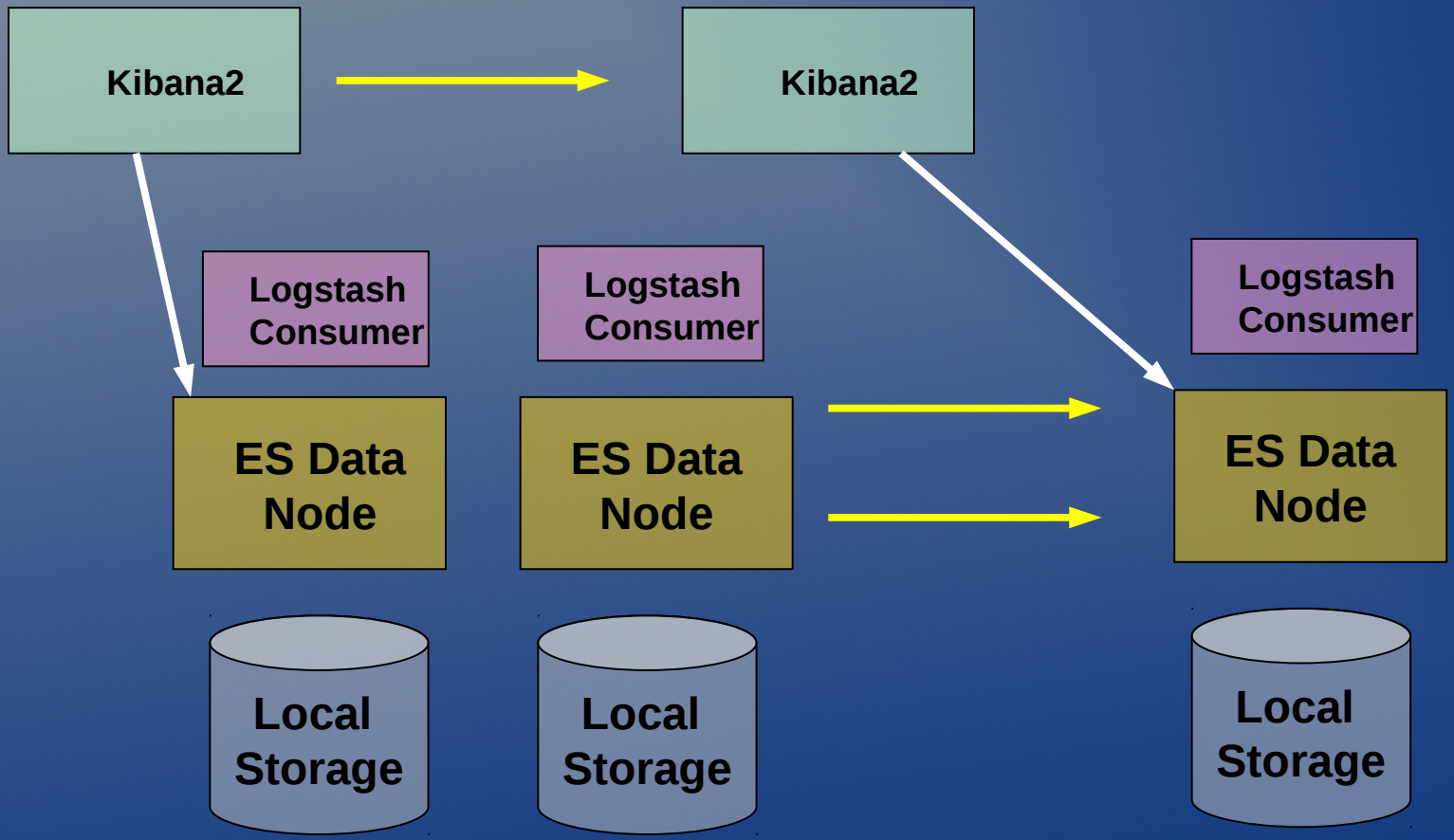

App Metrics
System Metrics
App Logs
Sys Logs
Events
App Logs
Collectd Daemon
Logstash Daemon
Log4J syslog appender
RabbitMQ
RabbitMQ
Graylog2
Collectd Consumer
Logstash Consumer
ES1 ES2ES3
Graphite
Nagios
Graphitus
Kibana2

Scale Out

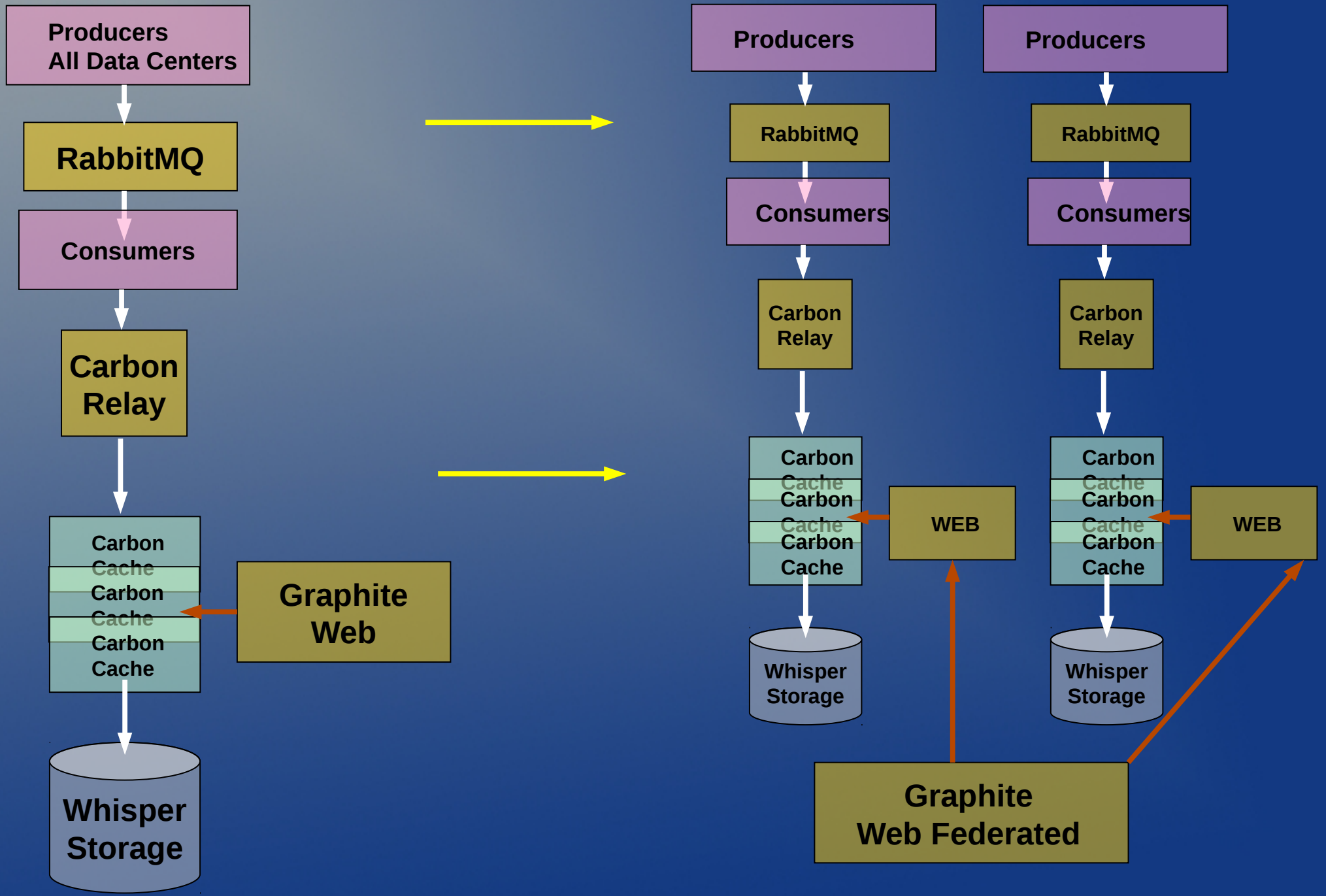
Nagios


Graphite
Nagios Agent DC1
Nagios Agent DC2
Nagios Agent DC3
Pager Duty
Nagios Aggregator
Red Alert
NagUI/Dash

Kibana2,ES,Logstash

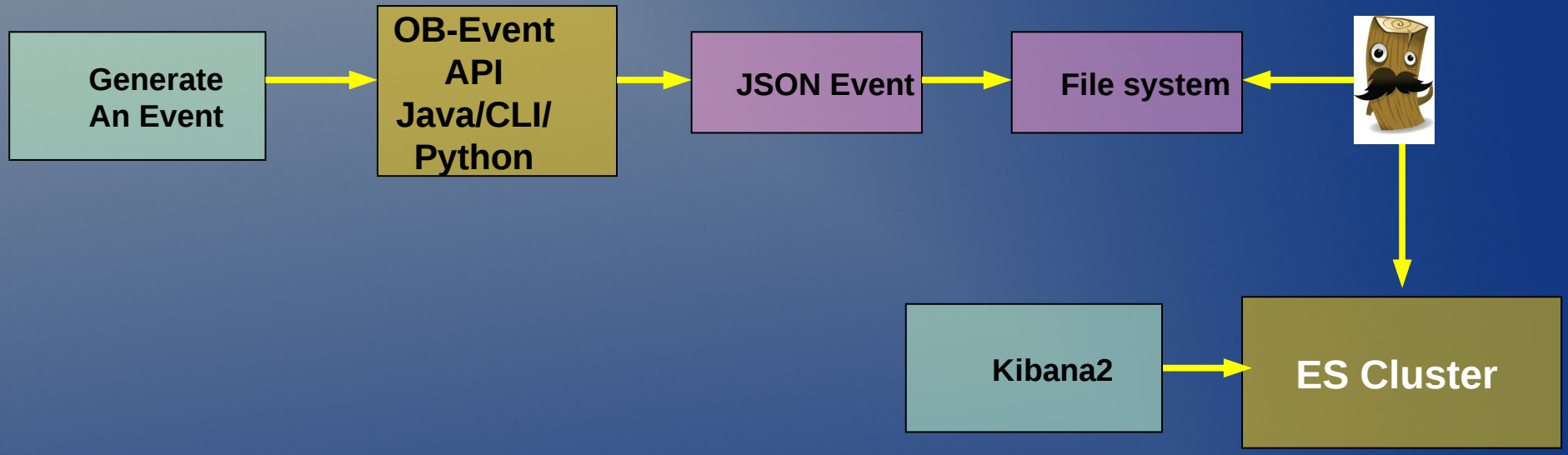


Graphite

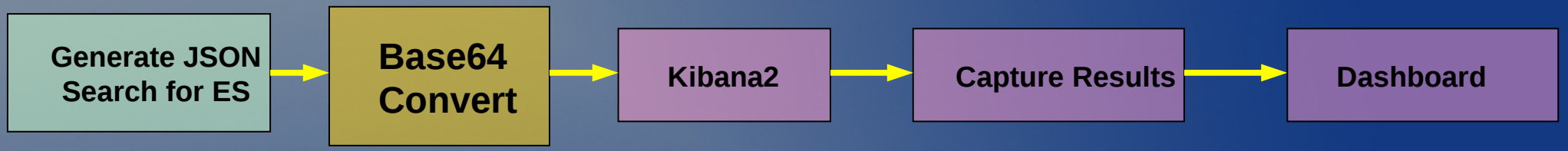

Producers
All Data Centers
RabbitMQ
Consumers
Carbon
Relay
Carbon
Cache
Carbon
Cache
Carbon
Cache
Graphite
Web
Whisper
Storage
Producers
RabbitMQ
Consumers
Carbon
Relay
Carbon
Cache
Carbon
Cache
Carbon
Cache
WEB
Whisper
Storage
Producers
RabbitMQ
Consumers
Carbon
Relay
Carbon
Cache
Carbon
Cache
Carbon
Cache
WEB
Whisper
Storage
Graphite
Web Federated

Cool Use Cases

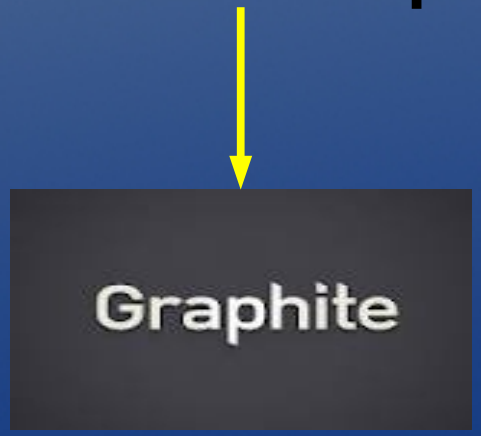
Using  As Generic Event Bus



Using  As Search API Client



Next Step...



Now Everybody wants to use it !

Self Service

- Why?
- Who uses Self Service
- Reduces “Waiting For” Scenarios

Home Grown Self Service Apps

- Red Alert – Web Frontend for Manipulating Nagios services and alerts
- Graphitus – Create your own Dashboards Using A single JSON file
- Dashanti – Multi Source Types Dashboard

https://github.com/erezmazor/graphitus

In the Oven...

- Logstash as Producer and Consumer (Why?)
- Application Metrics Via Logstash TCP Socket
- Scaling out graphite
- Kibana Search Queries to Graphite
- Graylog2 , The Future?

Logstash , ElasticSearch Tips

- Use Logstash Dynamic index names
- Number Of Shards Need to Meet Number of Nodes for even Distribution
- Regularly Rotate Old Indices
- Use Multiple Logstash .conf files with unique @type Field per file
- Minimize the number of GROK filters to avoid excessive CPU usage
- Start logstash with a per-defined HEAP limit

Thank you all !