AWS
TRANSFORMATION DAY
BENELUX

Automate best practices and operational health for your AWS resources

Julien Simon, Principal Technical Evangelist, AWS @julsimon
Maarten Mortier, CTO, Showpad @maartengm

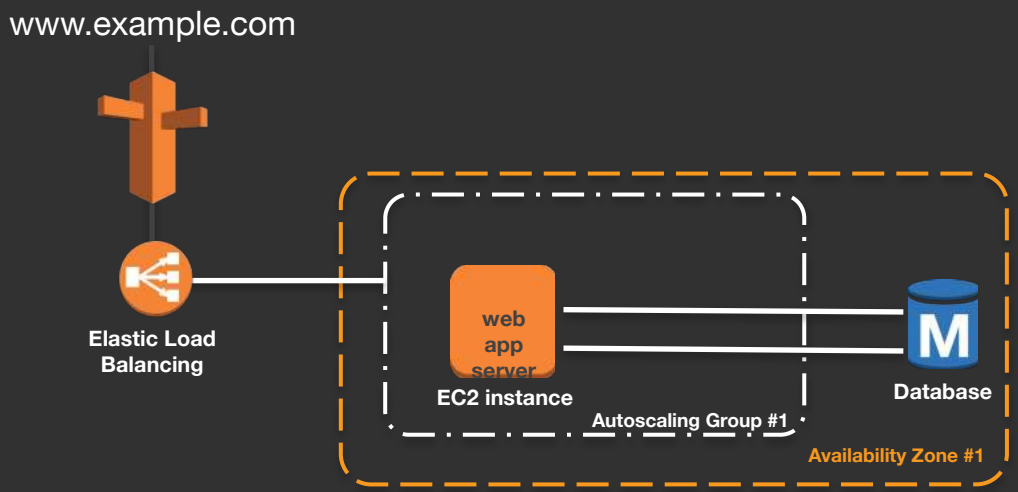
23/05/2017

amazon web services

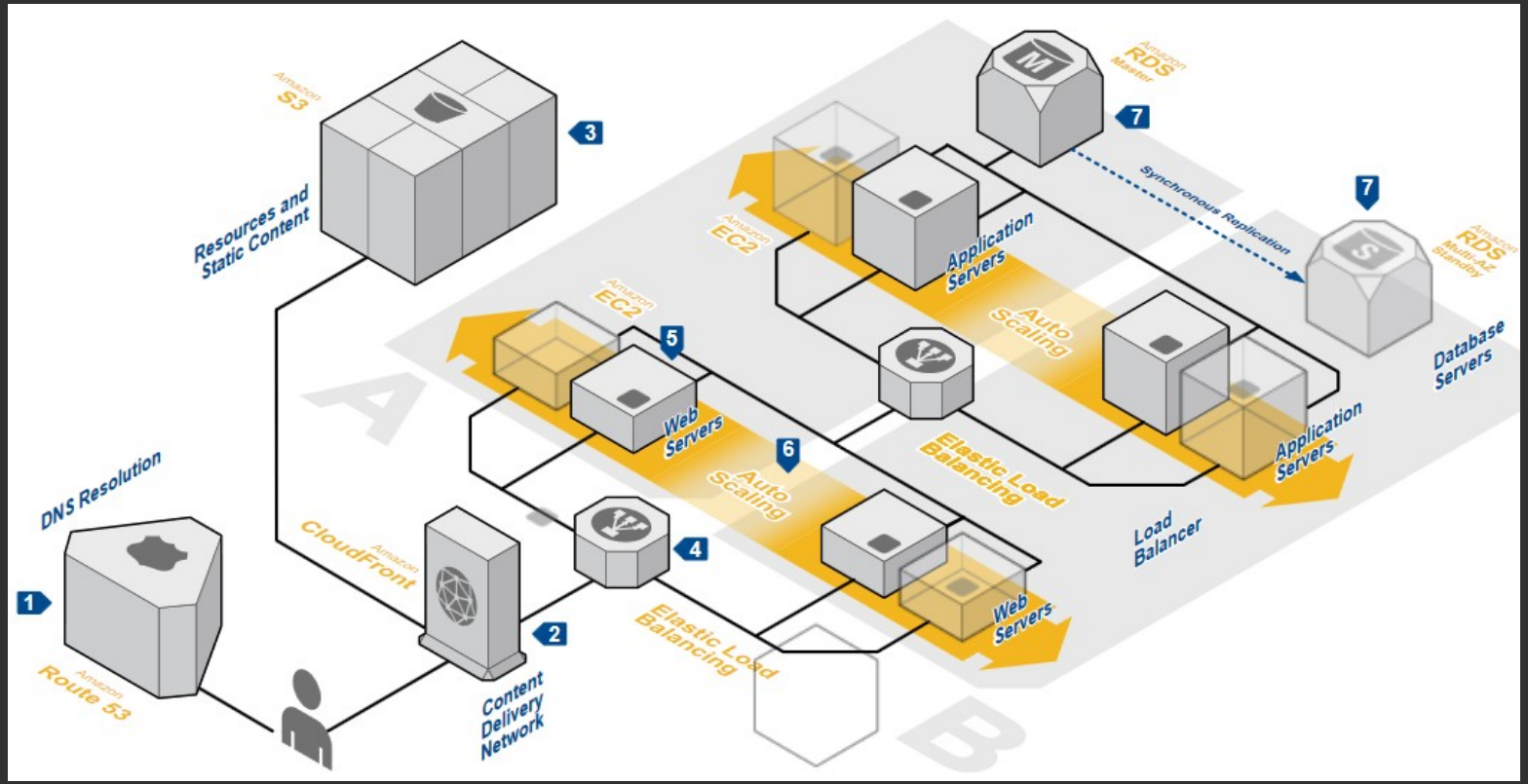
What to expect from this session

- Learn about Showpad's journey to full cloud automation
- Learn how to customize health alerts and automate remediation actions and with the Personal Health Dashboard
- Learn how to automate best practices with Trusted Advisor

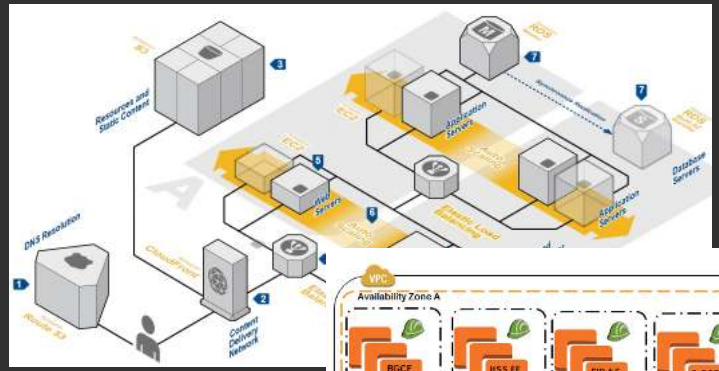
It probably started like this…

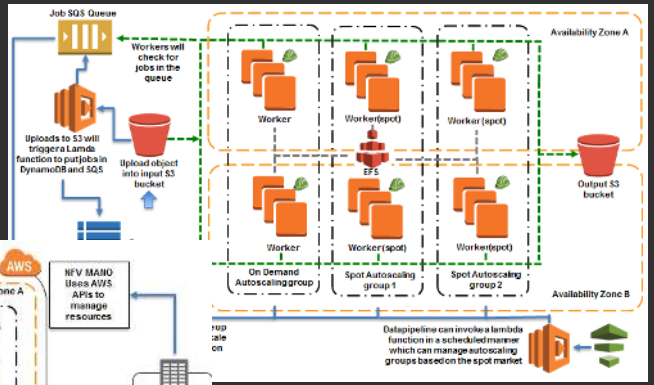


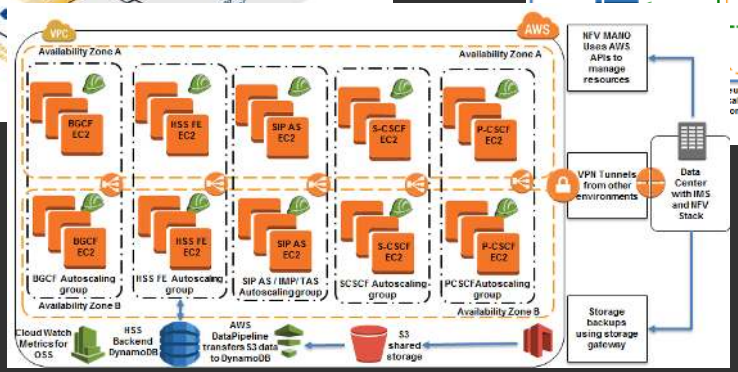
A few clicks in the console later….



And before you know it…





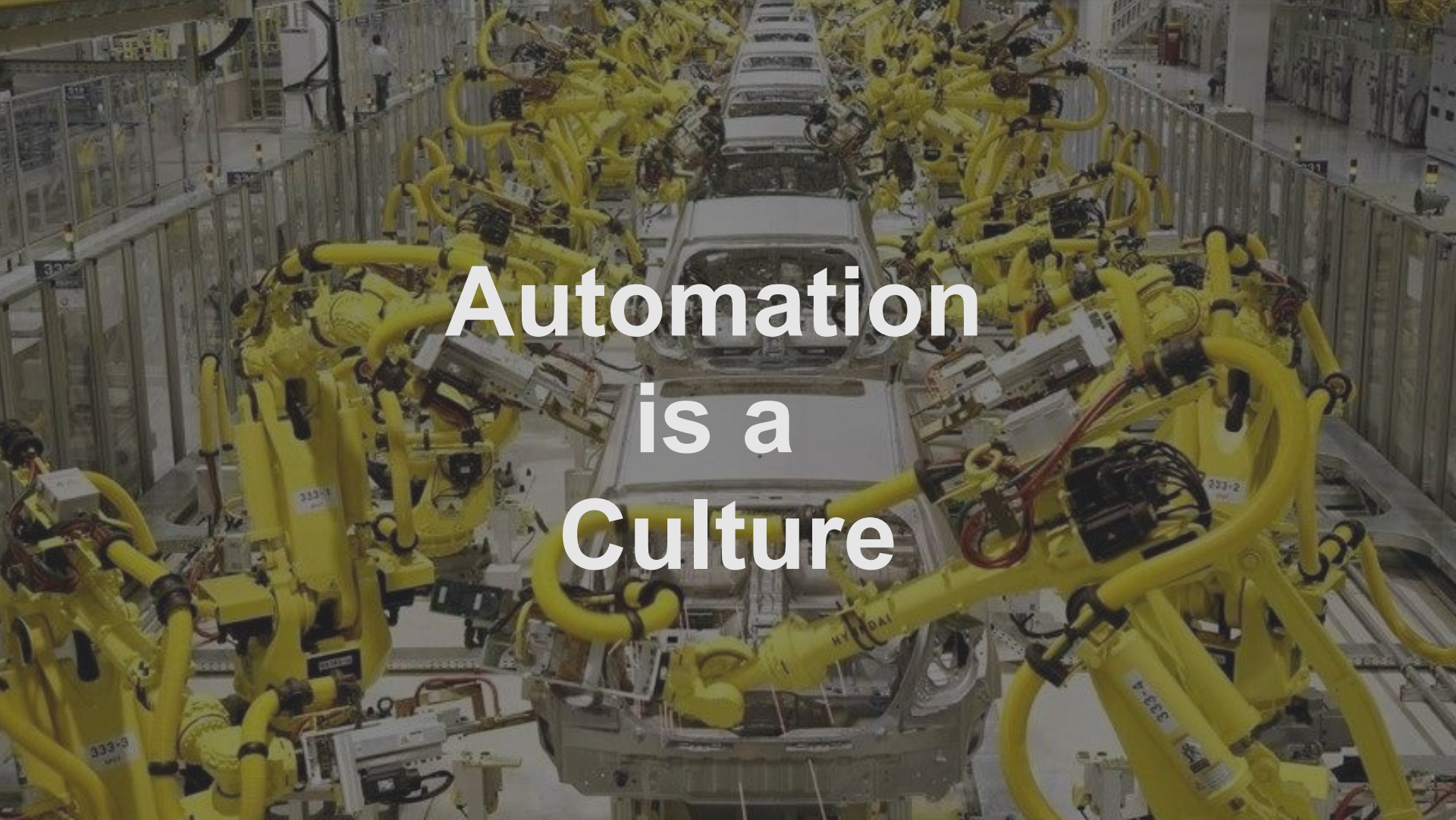


Sales Enablement **SaaS Platform**

Content Distribution and Content Usage Analytics

250 employees worldwide (SF, Portland, Ghent, London)
€25m ARR

Migrating to the AWS cloud


Automation
is a
Culture

HOW LONG CAN YOU WORK ON MAKING A ROUTINE TASK MORE EFFICIENT BEFORE YOU'RE SPENDING MORE TIME THAN YOU SAVE?

(ACROSS FIVE YEARS)

HOW MUCH TIME YOU SHAVE OFF	HOW OFTEN YOU DO THE TASK: 50/DAY	5/DAY	DAILY	WEEKLY	MONTHLY	YEARLY
1 SECOND	1 DAY	2 HOURS	30 MINUTES	4 MINUTES	1 MINUTE	5 SECONDS
5 SECONDS	5 DAYS	12 HOURS	2 HOURS	21 MINUTES	5 MINUTES	25 SECONDS
30 SECONDS	4 WEEKS	3 DAYS	12 HOURS	2 HOURS	30 MINUTES	2 MINUTES
1 MINUTE	8 WEEKS	6 DAYS	1 DAY	4 HOURS	1 HOUR	5 MINUTES
5 MINUTES	9 MONTHS	4 WEEKS	6 DAYS	21 HOURS	5 HOURS	25 MINUTES
30 MINUTES		6 MONTHS	5 WEEKS	5 DAYS	1 DAY	2 HOURS
1 HOUR		10 MONTHS	2 MONTHS	10 DAYS	2 DAYS	5 HOURS
6 HOURS				2 MONTHS	2 WEEKS	1 DAY
1 DAY					8 WEEKS	5 DAYS

Key Areas of Automation

Change Management
Incident Management
Success Management

Change Management

Managing Infrastructure Changes
Managing Infrastructure Needs (Autoscale)
Managing Code Changes (Test Automation)
Detecting Changes, CVE, ..

HashiCorp

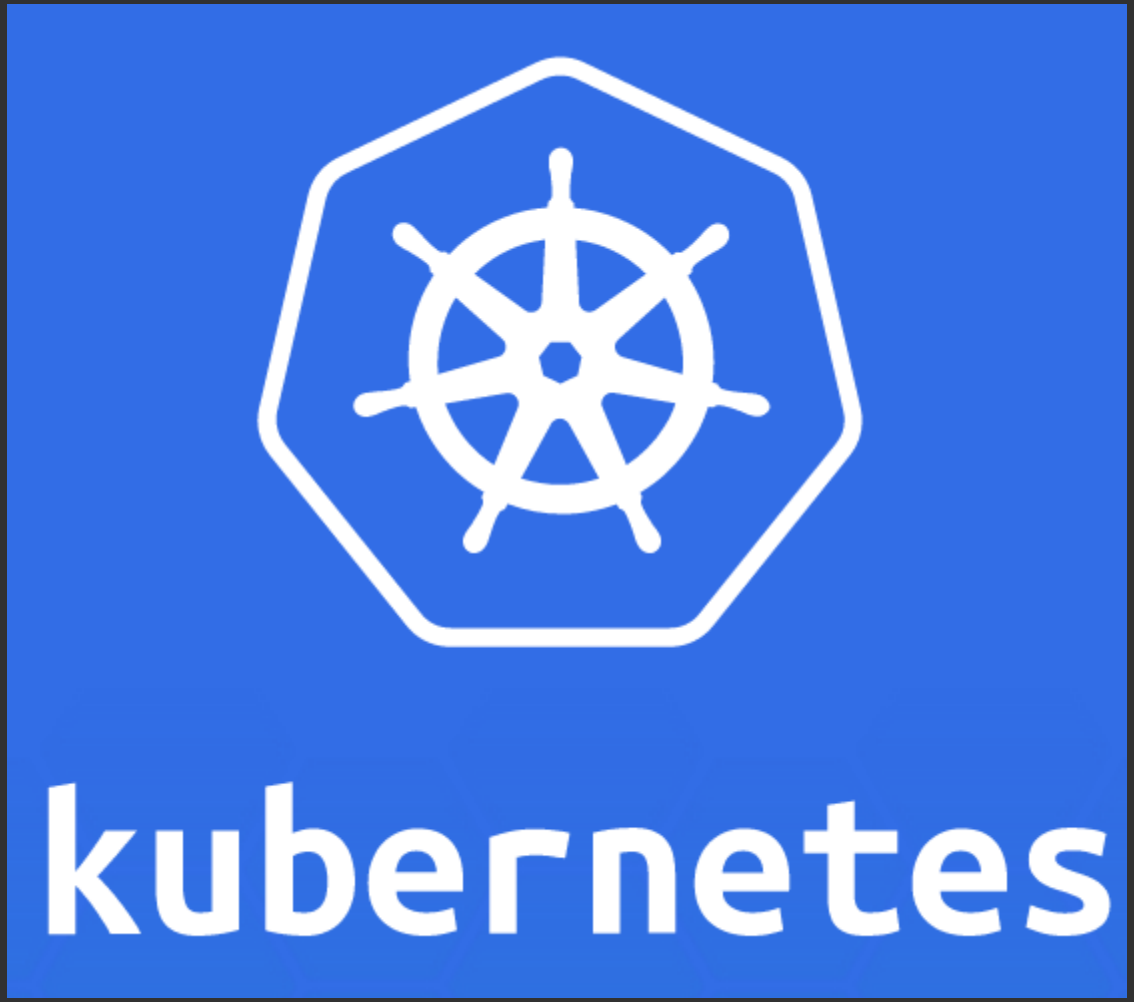
Terraform


CloudFormation

```
module "vpc" {
  source = "../../../modules/aws/vpc"
  name   = "${var.name}-vpc"
  cidr   = "${lookup(module.vpc_info.vpc_cidr, var.environment)}"
}

module "public_subnet" {
  source    = "../../../modules/aws/public_subnet"
  name      = "${var.name}-public"
  vpc_id    = "${module.vpc.vpc_id}"
  cidrs     = "${lookup(module.vpc_info.public_subnets, var.environment)}"
  azs       = "${module.vpc_info.azs}"
  ipv6_cidr = "${cidrsubnet(module.vpc.ipv6_cidr,4,0)}"
}

module "nat" {
  source            = "../../../modules/aws/nat"
  name              = "${var.name}-nat"
  azs               = "${module.vpc_info.azs}"
  public_subnet_ids = "${module.public_subnet.subnet_ids}"
}

module "ephemeral_subnets" {
  source          = "../../../modules/aws/private_subnet"
  name            = "${var.name}-ephemeral"
  vpc_id          = "${module.vpc.vpc_id}"
  cidrs           = "${lookup(module.vpc_info.ephemeral_subnets, var.environment)}"
  azs             = "${module.vpc_info.azs}"
  nat_gateway_ids = "${module.nat.nat_gateway_ids}"
  ipv6_cidr       = "${cidrsubnet(module.vpc.ipv6_cidr,4,1)}"
}

module "private_subnet" {
  source          = "../../../modules/aws/private_subnet"
  name            = "${var.name}-private"
  vpc_id          = "${module.vpc.vpc_id}"
  cidrs           = "${lookup(module.vpc_info.private_subnets, var.environment)}"
  azs             = "${module.vpc_info.azs}"
  nat_gateway_ids = "${module.nat.nat_gateway_ids}"
  ipv6_cidr       = "${cidrsubnet(module.vpc.ipv6_cidr,4,2)}"
}
```

```
+ module.geoip_microservice.aws_lambda_function.geoip_lambda
    arn:                                "<computed>"
    description:                        "Query the approximate geographical location of an IP"
    environment.#:                      "1"
    environment.0.variables.%:          "2"
    environment.0.variables.environment: "staging"
    environment.0.variables.region:     "eu-central-1"
    filename:                           "/home/jenkins/terraform/terraform/providers/aws/staging/eu_central_1/.t
on.zip"
    function_name:                      "geoIP"
    handler:                            "handler.geoIP"
    invoke_arn:                         "<computed>"
    last_modified:                      "<computed>"
    memory_size:                        "128"
    publish:                            "false"
    qualified_arn:                      "<computed>"
    role:                               "${aws_iam_role.geoip_lambda.arn}"
    runtime:                            "nodejs6.10"
    source_code_hash:                   "g6bGERZQXMaz6EKPYq7fdQUctCANOSTvQlH7ocEEt2c="
    timeout:                            "3"
    version:                            "<computed>"

+ module.geoip_microservice.aws_lambda_permission.allow_api_gateway
    action:        "lambda:InvokeFunction"
    function_name: "geoIP"
    principal:     "apigateway.amazonaws.com"
    source_arn:    "arn:aws:execute-api:${var.region}:${module.aws_account.account_id}:${aws_api_gateway_rest_api
s_api_gateway_resource.geoip_gateway.path}"
    statement_id:  "AllowExecutionFromApiGateway"


Plan: 22 to add, 1 to change, 1 to destroy.
```


Infrastructure changes are CI’d


kubernetes

Autoscaling Kubernetes

	Nodes	Pods
Horizontal	# of nodes	# of pods
Vertical	resources for a node	resources for a pod

```
scale_up_group:
  members: [kubernetes_node_host]
   # This defines a scale group whose members may be scaled up, incrementing by 1.
   # The scale worflow is called when the following criteria are met
   # The Hyperkube process total CPU will be more than 3 for a total of 10 seconds.
   # No more than 6 hosts will be allowed.
  policies:
    auto_scale_up:
      type: scale_policy_type
      properties:
        policy_operates_on_group: true
        scale_limit: 6
        scale_direction: '<'
        scale_threshold: 8
        service_selector: .*kubernetes_node_host.*.process.hyperkube.cpu.percent
        cooldown_time: 60
      triggers:
        execute_scale_workflow:
          type: cloudify.policies.triggers.execute_workflow
          parameters:
            workflow: scale
            workflow_parameters:
              delta: 1
              scalable_entity_name: kubernetes_node
              scale_compute: true
```

```
Total Pod Memory Allocated: 2048MB
============================
Used Pod Memory: 192
Pod memory Percent: 9%
============================
Used Pod Memory: 357
Pod memory Percent: 17%
============================
Average Pod Memory: 8
```

Lessons learned

Pre-scaling
Cache Priming
Spot Instances
Scheduled Scaling


docker
REGISTRY
TM & © 2015 Docker, Inc.


Twistlock™

Twistlock

- Dashboard
- Configure
  - ACCESS
  - DEFENDERS
  - SYSTEM
- Defend
  - ACCESS
  - RUNTIME
  - TRUST
- Monitor
  - ACCESS
  - RUNTIME
  - TRUST

About Enterprise 1.5.35

CONTAINERS | HOSTS | IMAGES | REGISTRY

SCAN NOW | CSV

Search images

HIDE CHART

REGISTRY	REPOSITORY	TAG	HOSTS	VULNERABILITIES	COMPLIANCE	STATUS	RUNNING	EXPLORE
docker.io	twistlock/private	defender_test	eran-pc	0	0	✓	true	
docker.io	twistlock/private	console_test	eran-pc	0	0	✓	true	
docker.io	examples/myproject	1.0	eran-pc	7	0	⚠	true	
docker.io	examples/myproject	1.0.1	eran-pc	0	0	✓	true	

docker.io/examples/myproject:1.0.1

ID: e64edfe60d2a9c6370c01030c517e9cd1ed2fe49f324fdde881de4170977b1df

OS distribution: Alpine Linux v3.4

VULNERABILITIES | COMPLIANCE | PROCESS INFO | PACKAGE INFO | HOSTS

ID	TYPE	SEVERITY	DESCRIPTION

There is no data to show

Incident Management

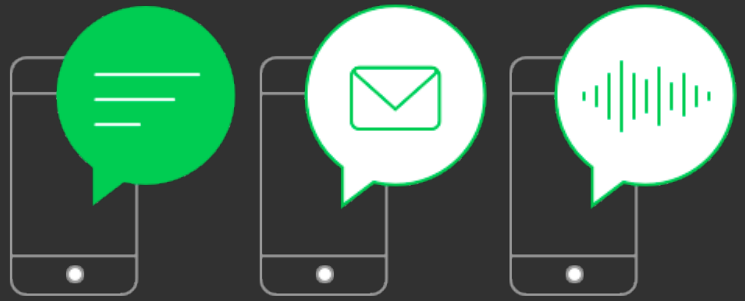

WORK METRICS
THROUGHPUT
SUCCESS
ERROR
PERFORMANCE
RESOURCE METRICS
UTILIZATION
SATURATION
ERROR
AVAILABILITY
EVENTS
CODE CHANGES
ALERTS
SCALING EVENTS
ETC

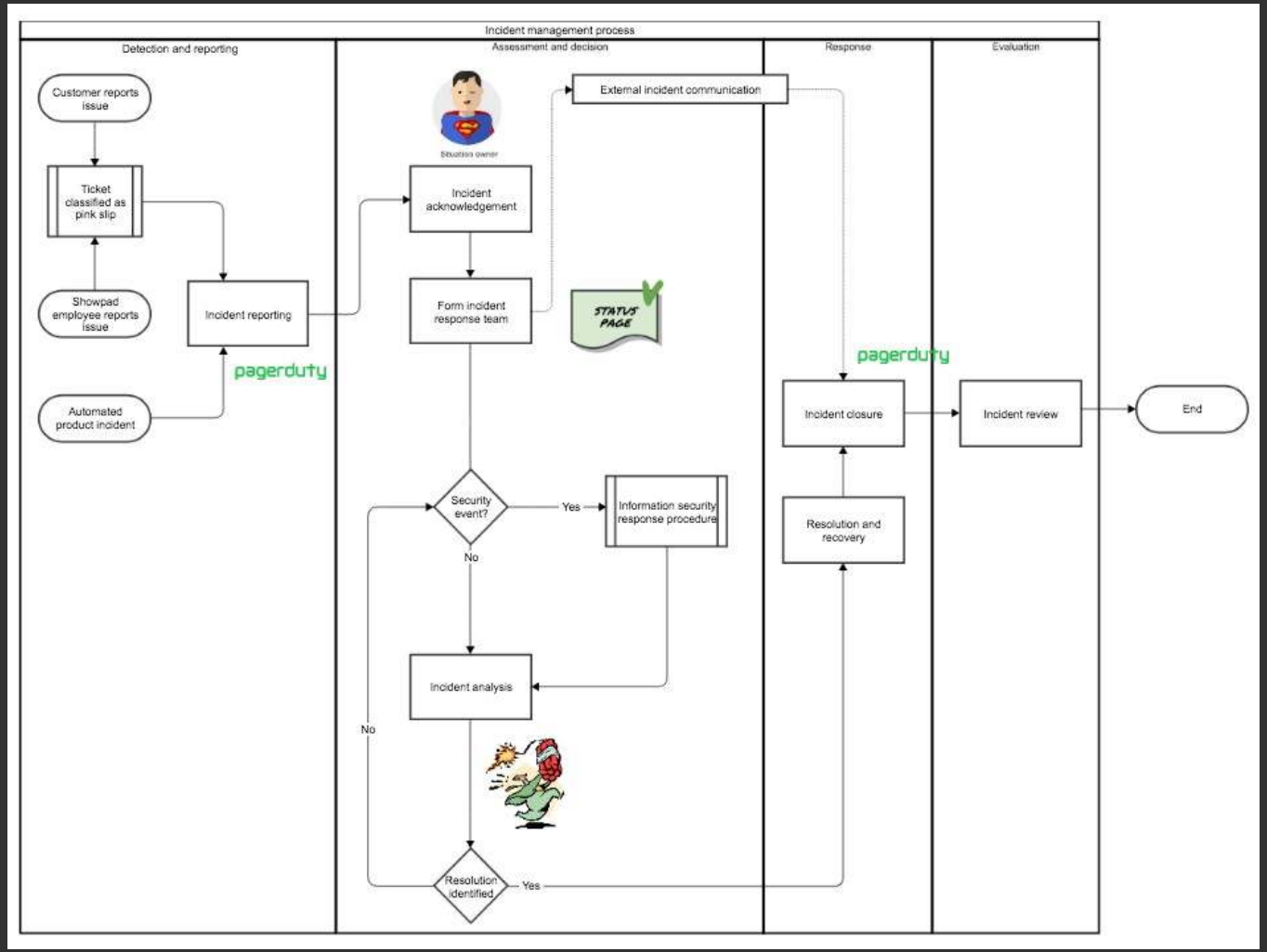
DATADOG


0

pingdom

pd




Incident management process
Detection and reporting
Assessment and decision
Response
Evaluation
Customer reports issue
Ticket classified as pink slip
Showpad employee reports issue
Automated product incident
Incident reporting
pagerduty
Incident acknowledgement
External incident communication
Form incident response team
STATUS PAGE
Security event?
Yes
No
Information security response procedure
Incident analysis
No
Resolution identified
Yes
pagerduty
Incident closure
Resolution and recovery
Incident review
End

Lessons learned

Enum & Prioritize Channels
Low and High Watermarks
Involve CEO on High Severity Channels

Lessons learned

Don’t overautomate customer communication

Success Management

Dashboard-First Development

HTTP Caching Prod 1d

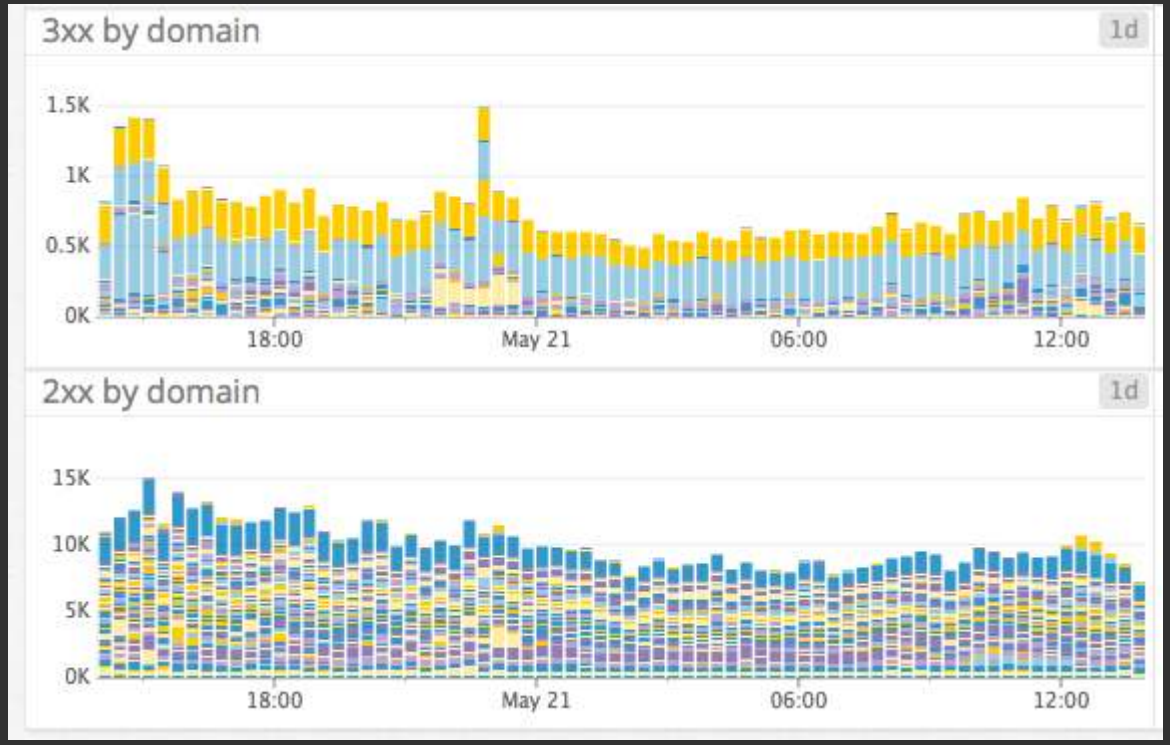
76.90%

Http caching prod 2d

67.37%

APU Processing time average 1d

0.516s



Each Feature gets a Dashboard

Micro-service using Amazon API Gateway to push any metrics document out and have historical info

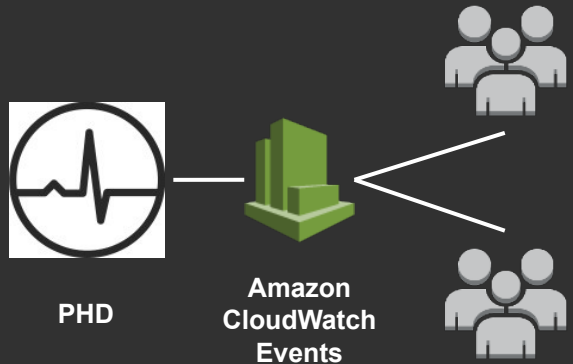
Automation applied to an efficient operation will magnify the efficiency. Automation applied to an inefficient operation will magnify the inefficiency.

Bill Gates

AWS Health and Personal Health Dashboard

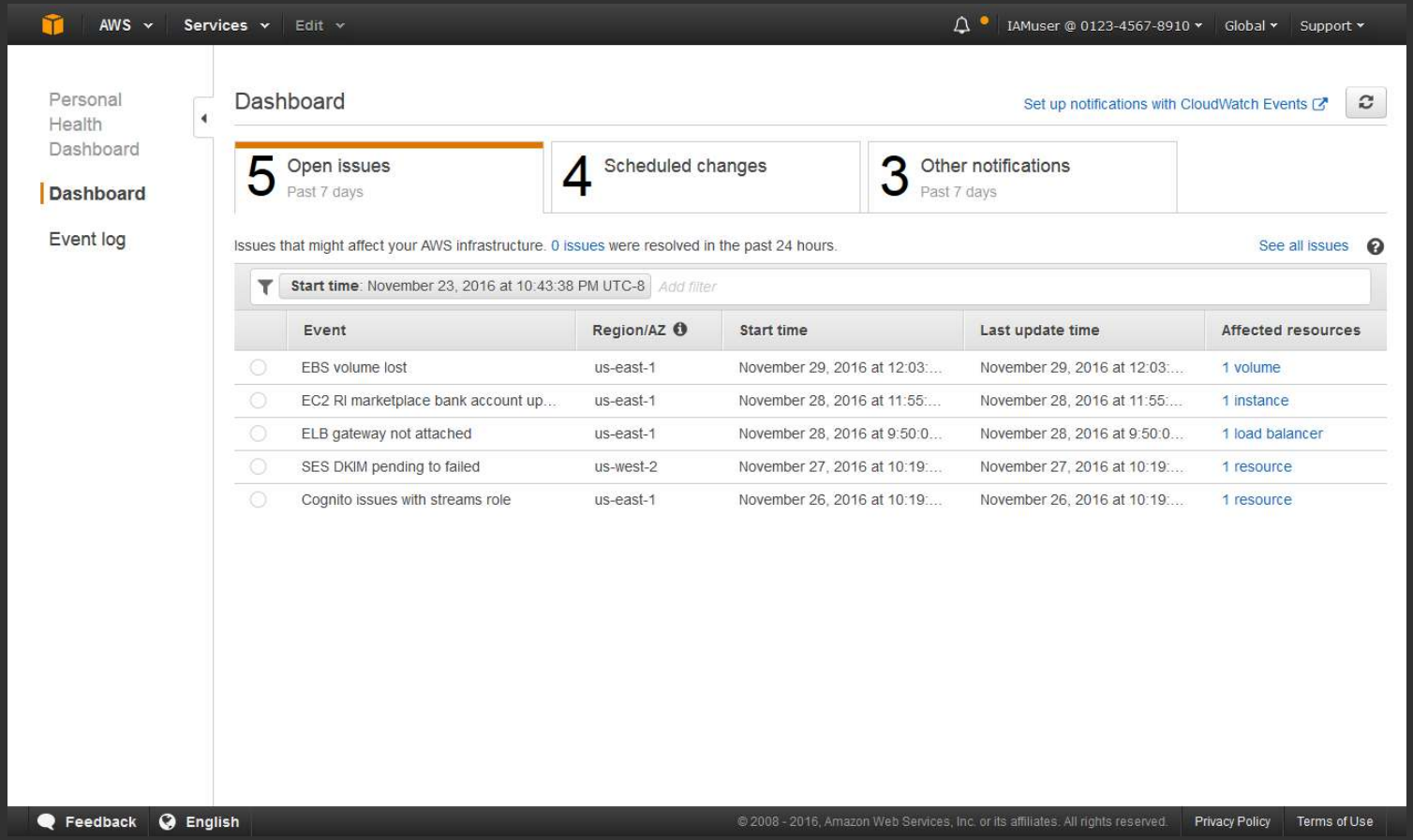
Visibility and transparency into your resources

Remediation guidance and knowledge articles



Custom notifications and automated actions

AWS Personal Health Dashboard



AWS Personal Health Dashboard

- Available for all AWS customershttps://phd.aws.amazon.com/
- Notifications and actions through CloudWatch Events
- Persistent notifications through new navigation bar alerts
- Hundreds of operational, billing, security event types

AWS Services and resources you use

Health Service

Personal Health Dashboard

API

- describe-events
- describe-event-details
- describe-affected-entities
- …

Push notifications through CloudWatch Events

- Set Rules to extract events of interest
- Set Targets for rules (Amazon SNS, Amazon SQS, AWS Lambda, Amazon Kinesis)

In-house or third-party monitoring and event management systems

AWS service integrations

Resource and service-level insights into health


Amazon EC2


Amazon EBS


Amazon RDS


Elastic Load Balancing


Amazon ElastiCache


AWS Certificate Manager


Amazon VPC


Amazon SES


AWS Direct Connect


Amazon Cognito


Amazon Elasticsearch Service


AWS CloudTrail

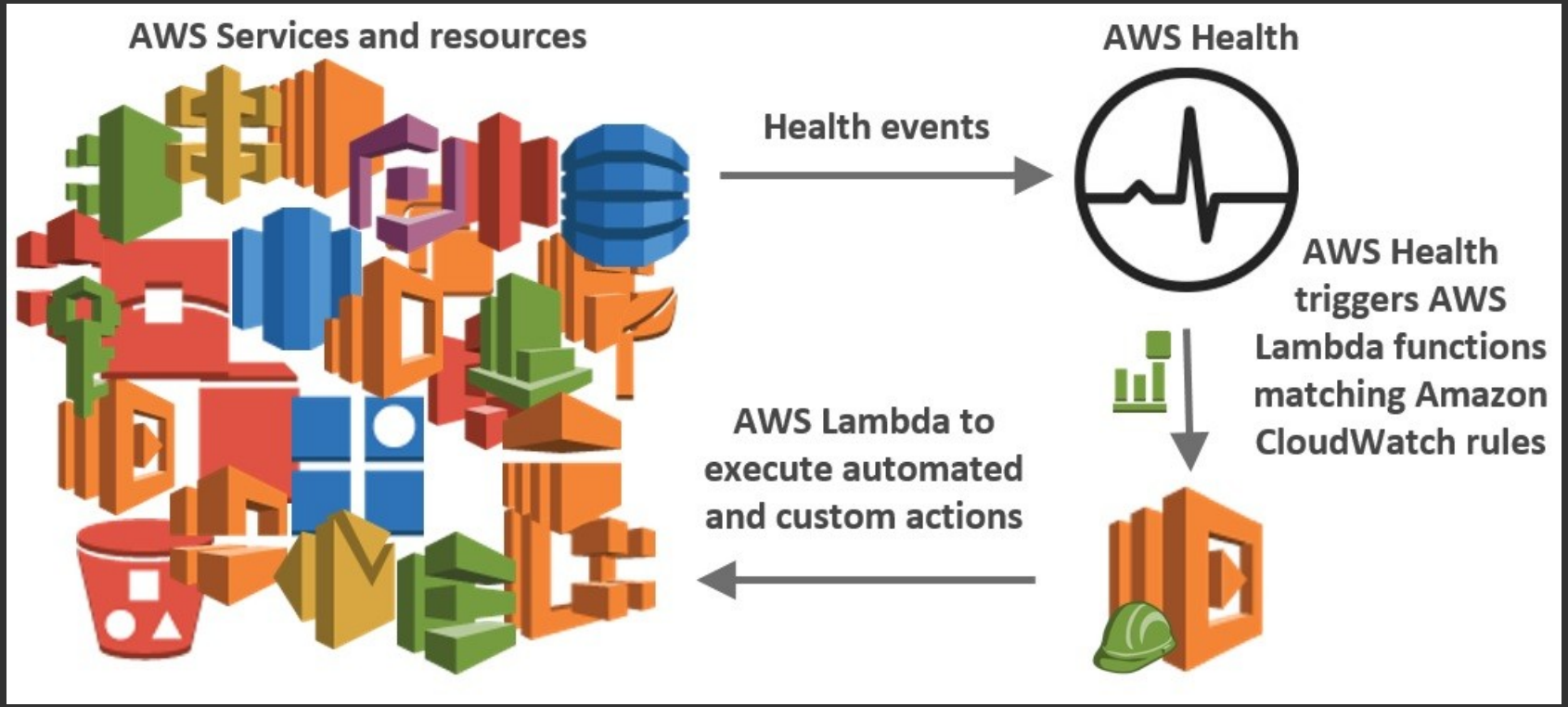
Service-level insights into health

All AWS services

Aws-health-tools

- Open Source and community driven on Github https://github.com/aws/aws-health-tools
- Customized alerts in response to AWS Health events
- Automated actions in response to AWS Health events

How does it work?



Examples

SMS Notifier – Send custom text or SMS notifications.

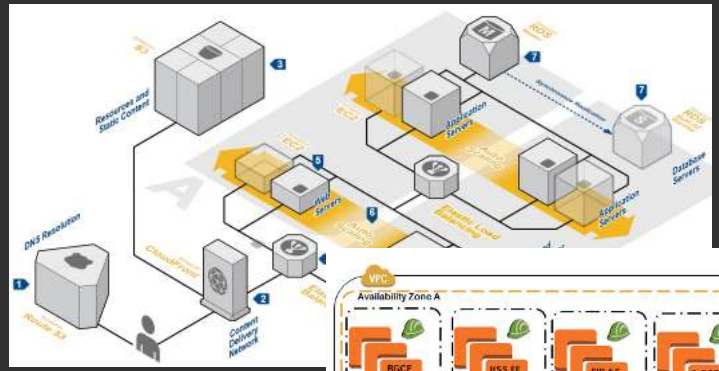
SNS Topic Publisher – Publish to an SNS topic.

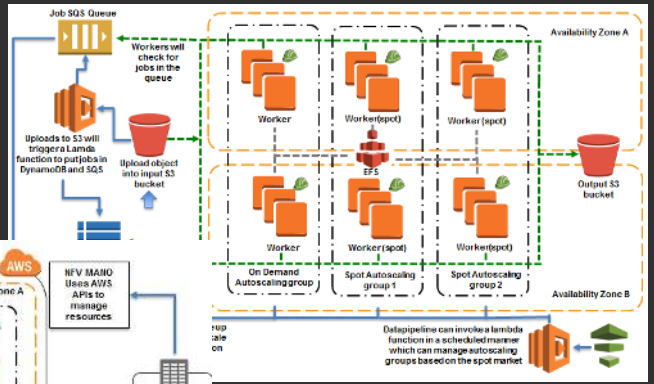
Slack Notifier – Post to a Slack channel.

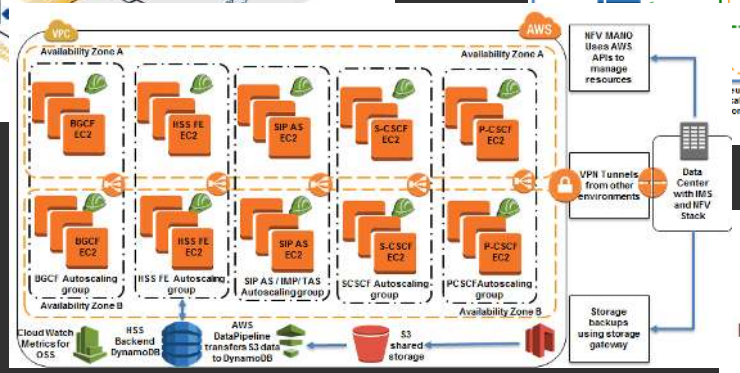
Instance Store Degraded Drive – Stop or terminate an instance that has a degraded instance store drive.

Disable AWS CodePipeline Stage Transition – Stop deployment when an issue arises.

Health automation: check. Time to optimize!







So what is Trusted Advisor?

AWS Trusted Advisor provides best practices (or checks) in four categories: **cost optimization, security, fault tolerance**, and **performance** improvement.

Red action strongly recommended
Yellow investigation recommended
Green no problem detected

AWS Trusted Advisor

Over **50 million** recommendations provided to AWS customers resulted in **$500m+** in cost savings for users of Trusted Advisor

Let’s look at an example

Low Utilization Amazon EC2 Instances 

Warns when EC2 instances appear to be underused.

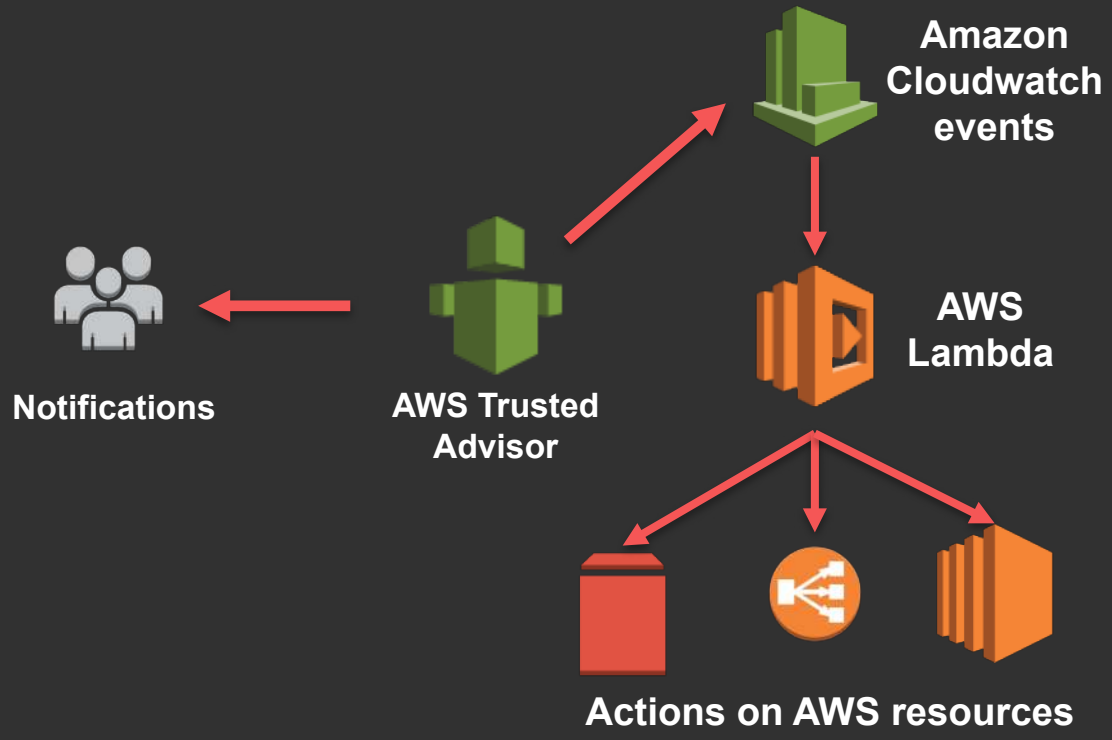
Alert Criteria

Yellow An instance had 10% or less daily average CPU utilization and 5 MB or less network I/O on at least 4 of the previous 14 days.

Recommended Action

Consider stopping or terminating instances that have low utilization.

Building Automation



Automation Setup – safety first!

Tag resources subject to TA optimization actions.

Create an IAM policy and role for the Lambda function to use.

Setup up a Cloudwatch event rule to trigger the Lambda function.

Setup the Lambda function to take actions recommended by Trusted Advisor.

Cost Optimization

3 5 0

$1,059.76

Potential monthly savings

Filter by tag

Tag Key | Tag Value | Apply filter | Reset

View: Investigation recommended

Cost Optimization Checks

Low Utilization Amazon EC2 Instances — *Refreshed: 4 minutes ago*

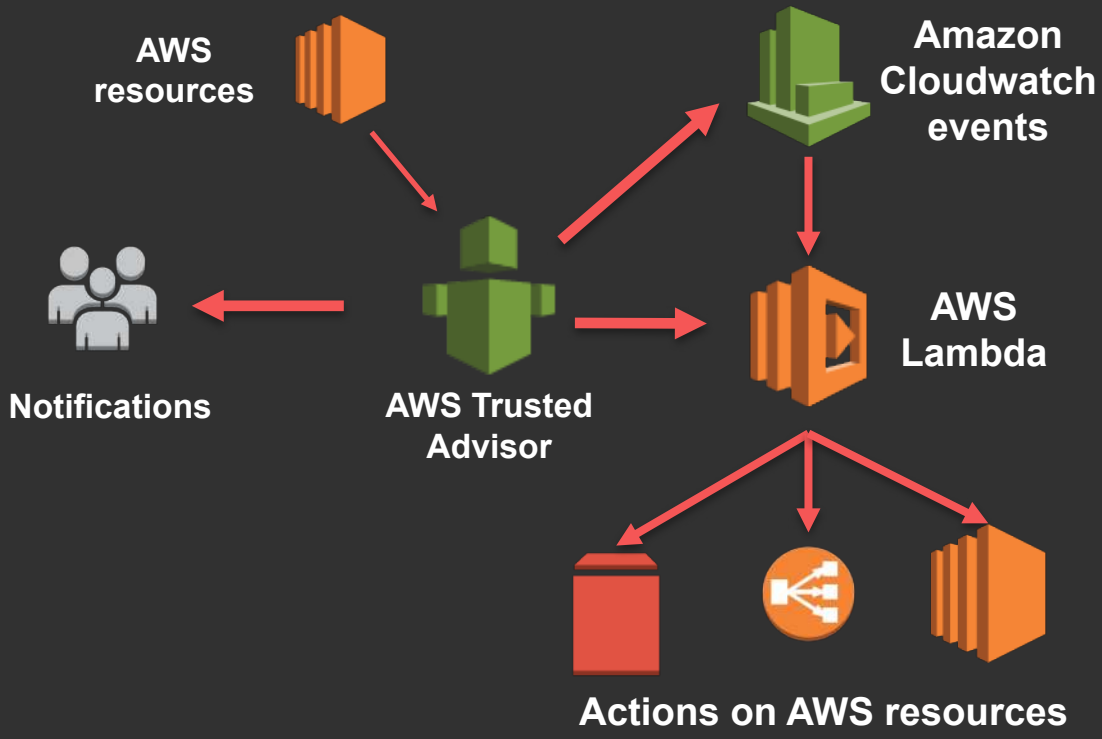
Checks the Amazon Elastic Compute Cloud (Amazon EC2) instances that were running at any time during the last 14 days and alerts you if the daily CPU utilization was 10% or less and network I/O was 5 MB or less on 4 or more days.

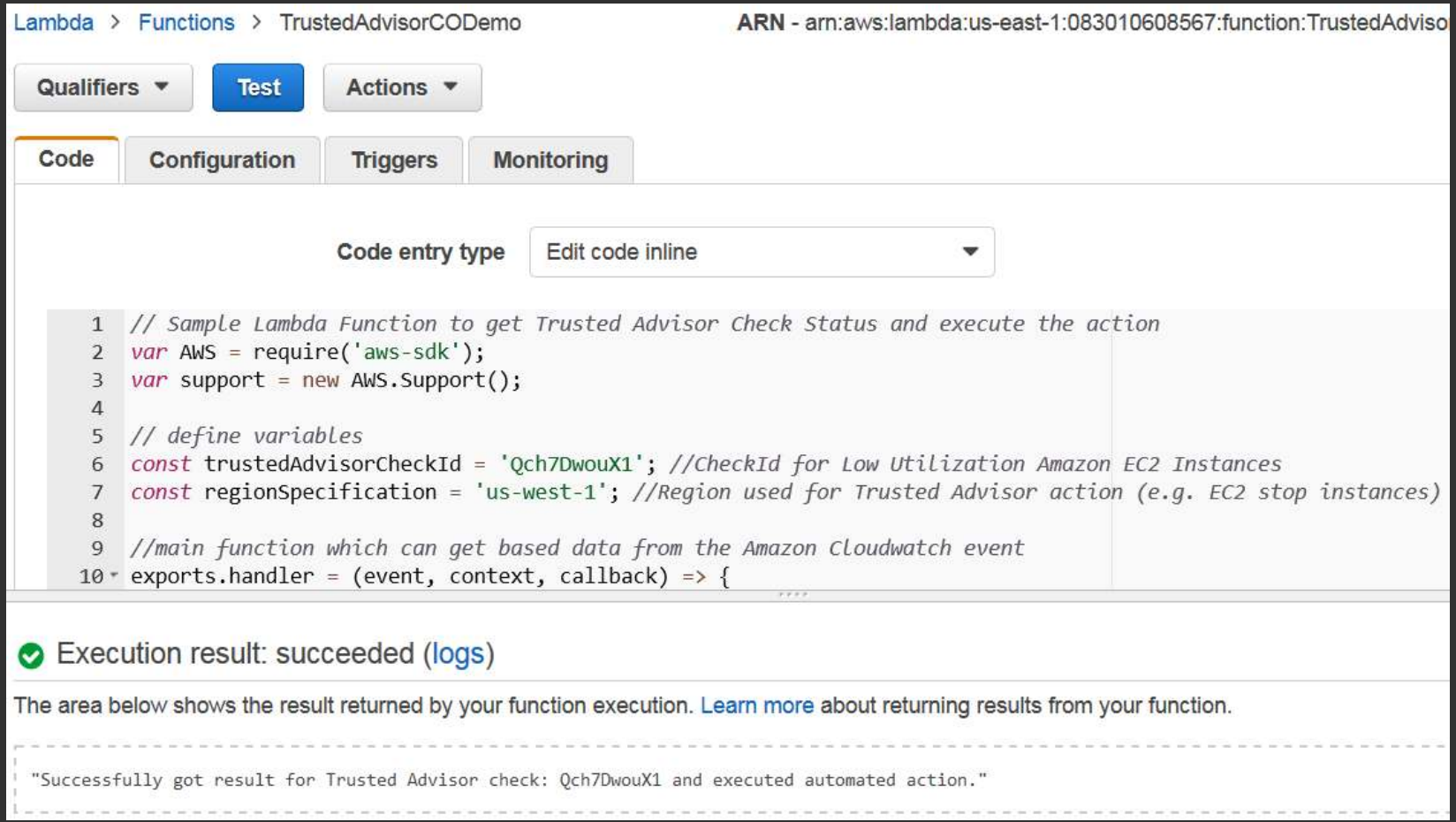
14 of 17 Amazon EC2 instances have low average daily utilization. Monthly savings of up to $519.12 might be available by minimizing underutilized instances.

Lots of idle tests instances in us-west-1

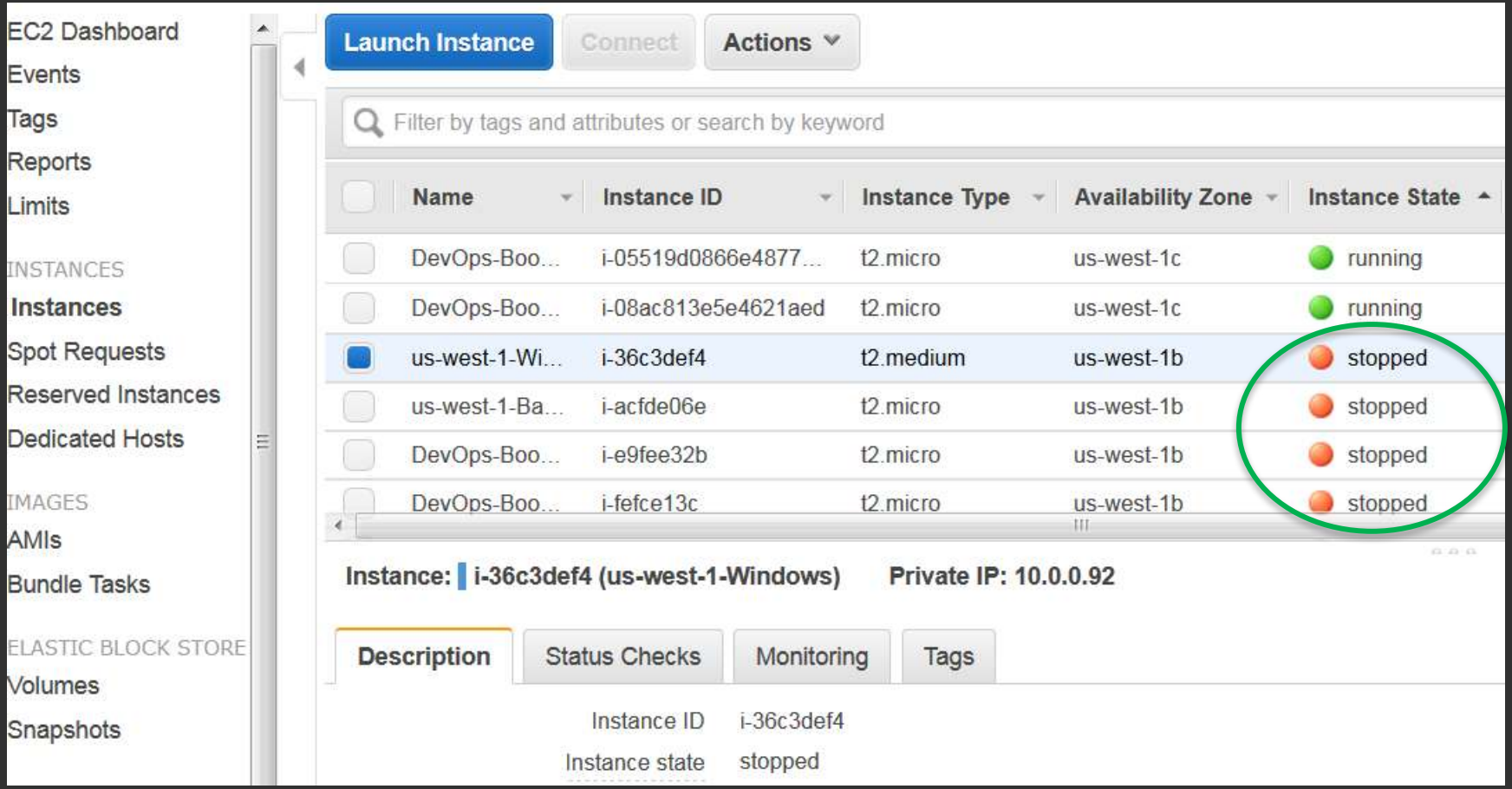
Region/AZ	Instance ID	Instance Name	Instance Type	Estimated Monthly...	CPU U
us-west-1b	i-acfde06e	us-west-1-Bastion	t2.micro	$12.24	0.0%
us-west-1b	i-36c3def4	us-west-1-Windows	t2.medium	$63.36	0.2%
us-west-1c	i-fc28833c	DevOps-Bootcamp	t2.micro	$12.24	0.4%
us-west-1c	i-102883d0	DevOps-Bootcamp2	t2.micro	$12.24	0.4%
us-west-1b	i-e9fee32b	DevOps-Bootcamp2	t2.micro	$12.24	0.4%
us-west-1b	i-fefce13c	DevOps-Bootcamp	t2.micro	$12.24	0.4%
us-east-1a	i-00e8ad544c9fbffb7	DevOps-Bootcamp	t2.micro	$9.36	0.4%

Building Automation

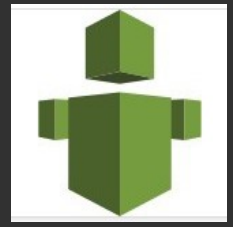




Examples available on Github https://github.com/aws/Trusted-Advisor-Tools


EC2 Dashboard
Events
Tags
Reports
Limits
INSTANCES
Instances
Spot Requests
Reserved Instances
Dedicated Hosts
IMAGES
AMIs
Bundle Tasks
ELASTIC BLOCK STORE
Volumes
Snapshots
Launch Instance
Connect
Actions
Filter by tags and attributes or search by keyword
Name
Instance ID
Instance Type
Availability Zone
Instance State
DevOps-Boo... i-05519d0866e4877... t2.micro us-west-1c running
DevOps-Boo... i-08ac813e5e4621aed t2.micro us-west-1c running
us-west-1-Wi... i-36c3def4 t2.medium us-west-1b stopped
us-west-1-Ba... i-acfde06e t2.micro us-west-1b stopped
DevOps-Boo... i-e9fee32b t2.micro us-west-1b stopped
DevOps-Boo... i-fefce13c t2.micro us-west-1b stopped
Instance: i-36c3def4 (us-west-1-Windows)
Private IP: 10.0.0.92
Description
Status Checks
Monitoring
Tags
Instance ID i-36c3def4
Instance state stopped

Trusted Advisor



Cost Optimization

- ✓ Efficient utilization of resources
- ✓ Idle resource reclamation
- ✓ Right-sizing reserved EC2 instances

Performance

- ✓ Correct configuration of services
- ✓ Utilization against AWS Service Limits
- ✓ High-utilization EC2 instances

Security

- ✓ Logging levels
- ✓ Correct configuration of users, access keys, SSL certificates, and security groups

Fault Tolerance

- ✓ Resource redundancy
- ✓ Scheduled backups
- ✓ Load balancing

Conclusion

Don’t stop at automating infrastructure provisioning

You can leverage Trusted Advisor and AWS Health to automate best practices and operational health.

These samples will get you started ☺

https://github.com/aws/aws-health-tools

https://github.com/aws/Trusted-Advisor-Tools/

AWS
TRANSFORMATION DAY
BENELUX

Thank you!

Julien Simon, Principal Technical Evangelist, AWS @julsimon
Maarten Mortier, CTO, Showpad @maartengm

amazon web services