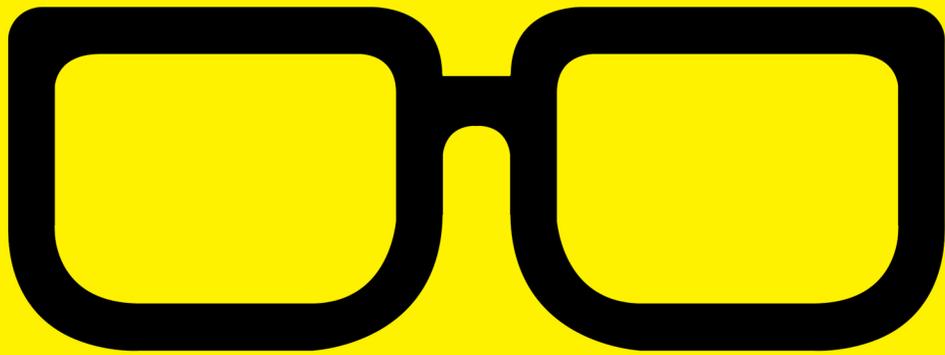


SPONSORED BY
botfactory.io
from DivvyCloud

GEEK GUIDE



BotFactory: Automating the End of Cloud Sprawl

Table of Contents

About the Sponsor	4
Introduction	5
What Is BotFactory?	6
The Classic Entrepreneur’s Opportunity—Solve Your Own Problem First	7
Not a Single Tool, but a Platform	8
Taming Your Cloud with Bots	12
Active Policy Enforcement Bot Example.....	16
Case Studies	19
General Electric	19
Software AG Government Solutions	21
Discovery Communications	21
Getting Started at BotFactory.io	22
Documentation	24
Resources	25

JOHN S. TONELLO is the Director of IT and Communications Manager for NYSERNet Inc., New York’s regional optical networking company, serving the state’s colleges, universities and research centers. He’s been a Linux user and enthusiast since building his first Slackware system from diskette more than 20 years ago.

GEEK GUIDES:

Mission-critical information for the most technical people on the planet.

Copyright Statement

© 2016 *Linux Journal*. All rights reserved.

This site/publication contains materials that have been created, developed or commissioned by, and published with the permission of, *Linux Journal* (the “Materials”), and this site and any such Materials are protected by international copyright and trademark laws.

THE MATERIALS ARE PROVIDED “AS IS” WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE, TITLE AND NON-INFRINGEMENT. The Materials are subject to change without notice and do not represent a commitment on the part of *Linux Journal* or its Web site sponsors. In no event shall *Linux Journal* or its sponsors be held liable for technical or editorial errors or omissions contained in the Materials, including without limitation, for any direct, indirect, incidental, special, exemplary or consequential damages whatsoever resulting from the use of any information contained in the Materials.

No part of the Materials (including but not limited to the text, images, audio and/or video) may be copied, reproduced, republished, uploaded, posted, transmitted or distributed in any way, in whole or in part, except as permitted under Sections 107 & 108 of the 1976 United States Copyright Act, without the express written consent of the publisher. One copy may be downloaded for your personal, noncommercial use on a single computer. In connection with such use, you may not modify or obscure any copyright or other proprietary notice.

The Materials may contain trademarks, services marks and logos that are the property of third parties. You are not permitted to use these trademarks, services marks or logos without prior written consent of such third parties.

Linux Journal and the *Linux Journal* logo are registered in the US Patent & Trademark Office. All other product or service names are the property of their respective owners. If you have any questions about these terms, or if you would like information about licensing materials from *Linux Journal*, please contact us via e-mail at info@linuxjournal.com.

About the Sponsor

DivvyCloud

DivvyCloud software enables organizations to achieve their cloud computing goals by simplifying and automating compliance and optimization of public and private cloud infrastructure. Using DivvyCloud, customers can leverage programmatic Bots to identify and remediate security, cost and scale problems in real time. DivvyCloud was founded by seasoned technologists who understand first hand what is necessary to succeed in today's fast-changing, multi-cloud world.

Definitions:

Bot·Fac·to·ry /bät'fakt(ə)rē/ noun

1. Simple point and click tool for users to configure, build and deploy Bots.
Also see BOT.

Bot /bät/ noun

1. Programmatic automation to analyze and correct cloud infrastructure per user-defined rules and actions.

BotFactory: Automating the End of Cloud Sprawl

JOHN S. TONELLO

Introduction

It's no secret that the rapid growth and shift to the cloud has given IT organizations concern from almost the very start. The early promise of Software-as-a-Service—and all the other as-a-service offerings—has quickly given way to concern about security, fragmentation and vendor lock-in. If you're like most, you've developed a healthy fear of black-box deployments, cloud sprawl and the gut feeling that things are more than a little out of control.

Yes, the savings, speed and scale of cloud-based services make them impossible to ignore. In fact, for many, it's getting harder and harder to justify on-premises infrastructure. Its cost and overhead simply will turn whatever it is you do into something uncompetitive.

Fortunately, many vendors feel your pain and have come forward with tools for monitoring and managing it all. The most promising of these are vendor-agnostic and extensible, enabling you to mix legacy systems with cloud systems seamlessly, all while integrating with the tools you

BotFactory is an “if-this-then-that” automation platform that enables you to anticipate and fix problems with your public and private cloud infrastructure in real time—not just learn about problems that grow into massive to-do lists.

already know and love.

This is the promise of DivvyCloud and its compelling BotFactory automation platform. Rather than providing just tools, the engineers at DivvyCloud have developed a framework for managing near and far-flung infrastructure. The company’s approach is designed to be used by anyone trying to manage cloud resources, from a single IT guy or gal trying to run a handful of Amazon Web Services (AWS) accounts for a small development shop to a huge enterprise like General Electric.

With DivvyCloud’s data-harvesting capabilities coupled with the automation offered by BotFactory, security teams can keep data and infrastructure secure; DevOps teams can do their deployments more easily without running afoul of infrastructure policies, and operations teams can automate the mundane and deliver the value of cloud by controlling costs and employing best practices.

What Is BotFactory?

BotFactory is an “if-this-then-that” automation platform that enables you to anticipate and fix problems with your

public and private cloud infrastructure in real time—not just learn about problems that grow into massive to-do lists. It leverages DivvyCloud’s underlying real-time data harvesting and automation platform.

The DivvyCloud founders come from Electronic Arts, makers of massive multiplayer online games, where speed and scale matter. Let’s face it, gamers aren’t a patient lot willing to cut you slack when their avatars get cut down at critical moments just because a bunch of your VMs failed. As EA engineers, they had to manage 5,000 servers in five countries serving millions of paying customers. Theirs was a highly competitive market, which put pressure on them to move quickly.

They soon learned this was nearly impossible with existing tools and staff, so they struck out to make it easier to track resources across projects, manage security groups and automate the life-cycle management of cloud resources. Instead of taking hours to provision resources, they wanted something that would enable them to create new instances, manage existing ones, monitor costly cloud resources and not wait all day for it to happen.

The Classic Entrepreneur’s Opportunity—Solve Your Own Problem First

Their answer was to build the DivvyCloud platform. The company’s next step is to put this capability into the hands of everyone with BotFactory. With DivvyCloud, they found that the data harvested across multiple platforms could become the basis for a dynamic set of extensible Bots to manage it all. Fed by near-real-time

information, they reasoned, BotFactory Bots can automate everything from adding new DNS entries for newly spawned VMs to triggering alerts for security holes they didn't want.

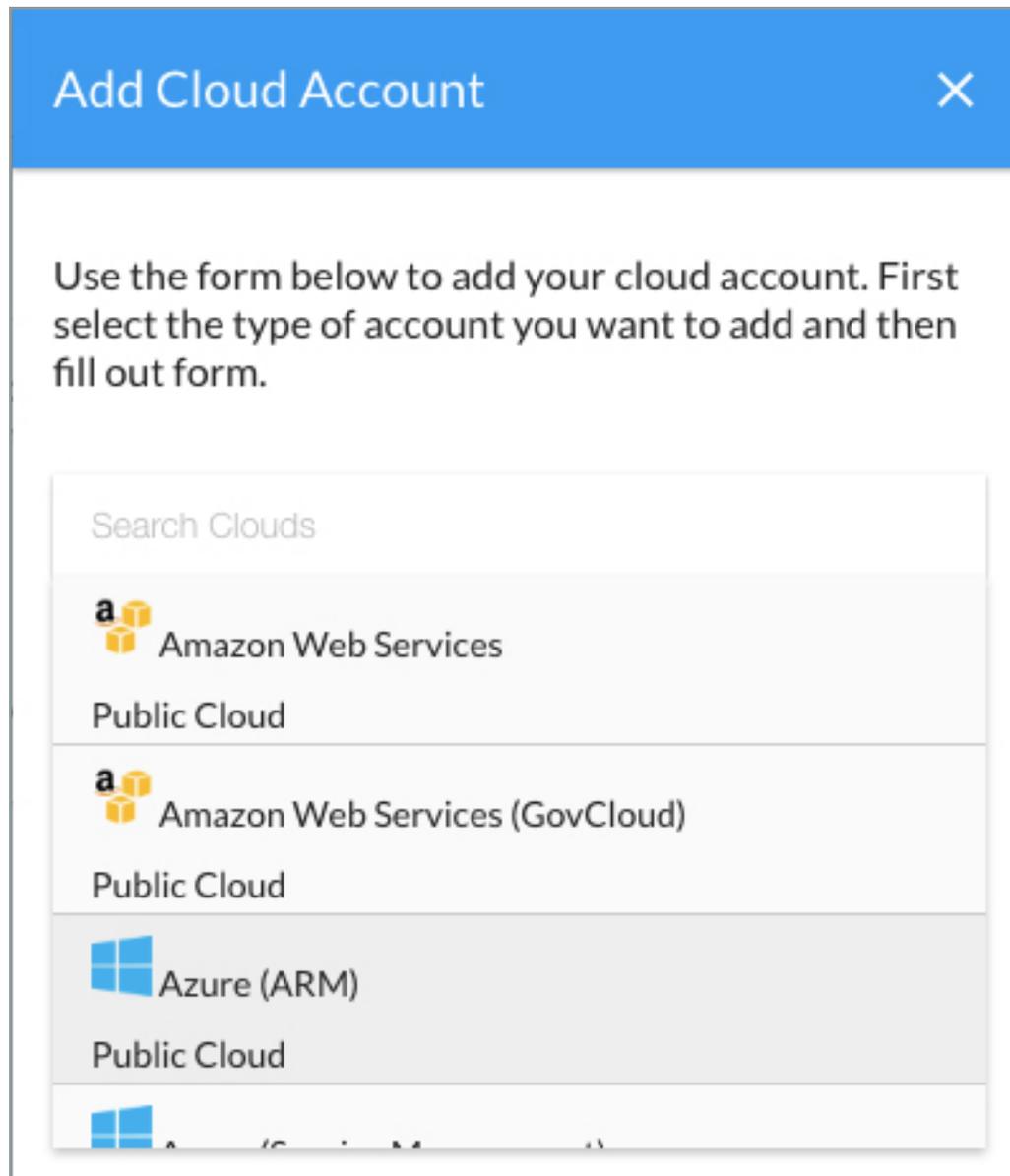
They also wanted it to work across most public and private clouds—and to work across them all at the same time. They designed DivvyCloud and BotFactory to connect easily to Amazon Web Services, Google Compute Engine, IBM SoftLayer, Microsoft Azure, OpenStack, Rackspace, VMware and other clouds, giving you a centralized way to stem cloud sprawl quickly and stop wasting valuable resources, time—and money.

Not a Single Tool, but a Platform

At its core, DivvyCloud harvests data from your systems no matter where they are, and BotFactory gives you a straightforward browser-based interface to identify non-compliant resources and fix them. Connecting your cloud resources is straightforward too. For vCenter 5.5 or 6.0, you add the IP address of your cluster and enter the credentials. For AWS, Azure and Google, you create service accounts for BotFactory to use. For OpenStack, you can add up to six core services, including Nova, Cinder and Quantum/Neutron.

Once connected, DivvyCloud immediately and continuously collects data from your cloud providers' API endpoints, not with agents, so the data stream is robust and not limited by host OS configurations. It then continuously monitors your environments for changes. New resources, changed resources, or metadata

or deleted resources create hook points, which can be used as starting points for automation and compliance scanning with BotFactory.



Add Cloud Account [X]

Use the form below to add your cloud account. First select the type of account you want to add and then fill out form.

Search Clouds

-  Amazon Web Services
Public Cloud
-  Amazon Web Services (GovCloud)
Public Cloud
-  Azure (ARM)
Public Cloud
-  Azure (Government)

FIGURE 1. Easily add your public and private cloud accounts to BotFactory by entering your credentials.

Since DivvyCloud can scan all your systems at once, you get insight on all your instances, volumes, networks, load balancers, SSL certificates, users and other resources from a single pane of glass, in a matter of minutes.

When you add multiple clouds, the system normalizes the data from your disparate sources and pours it into a database. Since it writes only changes, the database isn't clobbered, resulting in fast analysis. The cloud resource data is stored in a MySQL database while historical data (or in the enterprise platform, custom data feeds) is stored in ElasticSearch.

Hooks are at the core of the system, signaling when, say, new resources are created. Built-in plugins, or ones you build yourself via BotFactory, automate various actions, and the architecture, written in Python, is highly extensible. You can dig into almost any data point provided via API and customize it.

Other cloud-monitoring solutions, including evident.io, a security and compliance automation tool, and CloudHealth, a cloud analytics platform, similarly offer insight into your resources, but DivvyCloud and the BotFactory framework move quickly from monitoring to remediation. The combination gives you the ability to move beyond just knowing what's happening to actually taking action—and doing it quickly, automatically and consistently.

The fact is, there's a lot more going on inside your cloud infrastructure than inside a typical data center. In a cloud environment, there's continuous deployment and continuous activity. The rate of change is so high that monitors alone

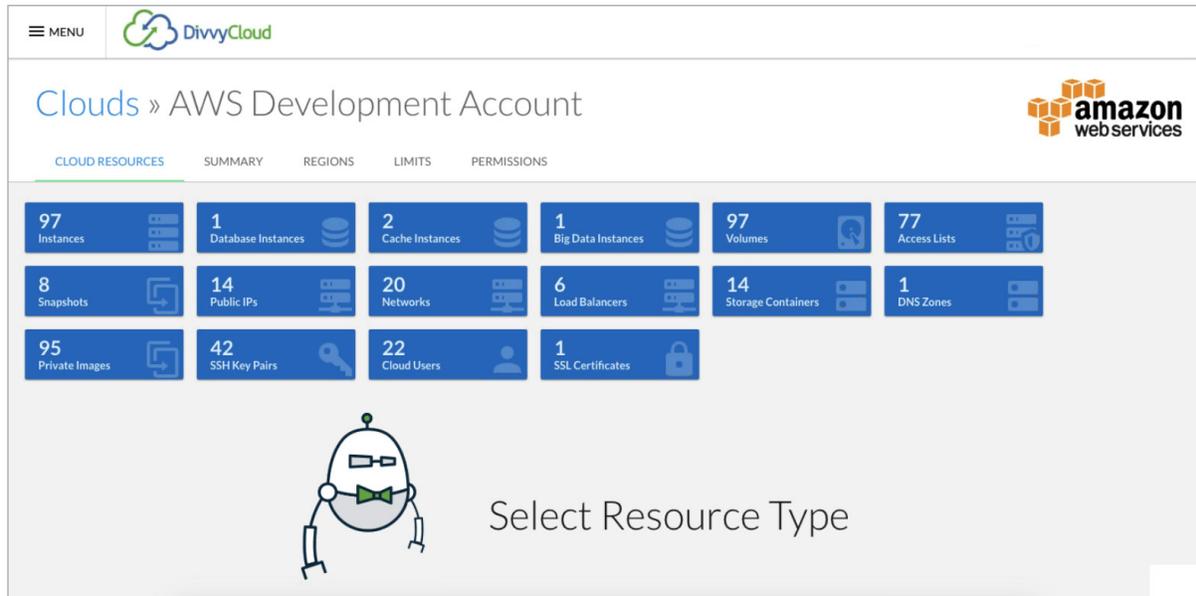


FIGURE 2. Create Bots by selecting resource types across all your cloud accounts and applying automation tasks.

often fall short, because by the time you get around to fixing problems the monitors uncover, the information is out of date.

BotFactory seeks to change that by using DivvyCloud's near-real time data to help you eliminate as much as 90% of the time you spend on mundane tasks, such as cleaning up a rapidly changing environment or setting up "safe" instances that DevOps teams can use and deploy on their own—without jeopardizing your environment or otherwise running off the rails.

The Bots themselves are pieces of automation code directed at your cloud infrastructure resources. Each one contains some metadata about who created it, the scope of analysis and resources you want it to look at, and conditions to act upon.

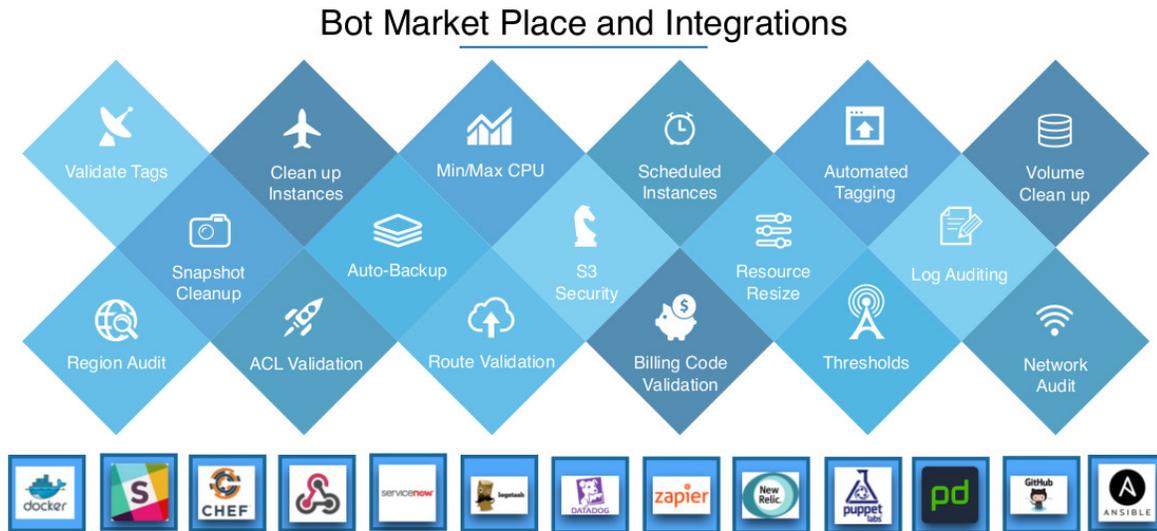


FIGURE 3. BotFactory integrates with a broad array of tools and platforms, and lets you choose from a wide variety of predefined tasks.

Bots can not only report back to you about your resources, they also can share the information with other tools you trust. DivvyCloud and BotFactory can integrate with Ansible, Puppet, Chef, Splunk, ServiceNow and a broad array of tools and services you already know. You can use Bots to automate tasks via those tools and move beyond simple orchestration and log analysis.

Taming Your Cloud with Bots

Once you've attached your cloud resources to DivvyCloud, you can begin to deploy Bots and start automating quickly. If you're familiar with how Splunk gives you rapid insight and turns one-dimensional logs into three-dimensional analysis, you'll be comfortable inside BotFactory.

The templates are widely varied and available on GitHub, so you can pull down what you need from the stock bots or modify them for your custom needs—all within BotFactory.

The interface lets you define the rules and policies you and your organization want to enforce with step-by-step questions and templates. You start by being asked what kind of problem you're trying to solve—security, optimization, best practices, curation or other miscellaneous tasks.

You then set up conditions that trigger actions. BotFactory's "if-this-then-that" approach makes it easy to decide what actions to take, including adding the resource to a report, generating an alert, creating an audit trail log or even terminating a rogue resource.

Select the types of resources you want to manage—volumes, snapshots and database instances, for example—and your Bots will scan all of your disparate cloud resources at once and give you a unified view of the issues.

Next, choose from any of more than 100 Bot templates to begin your analysis. For example, audit security rules and find systems with SSH ports open to the world, find load balancers with no nodes behind them, or identify unattached (and, therefore, wasted) volumes. The templates are widely varied and available

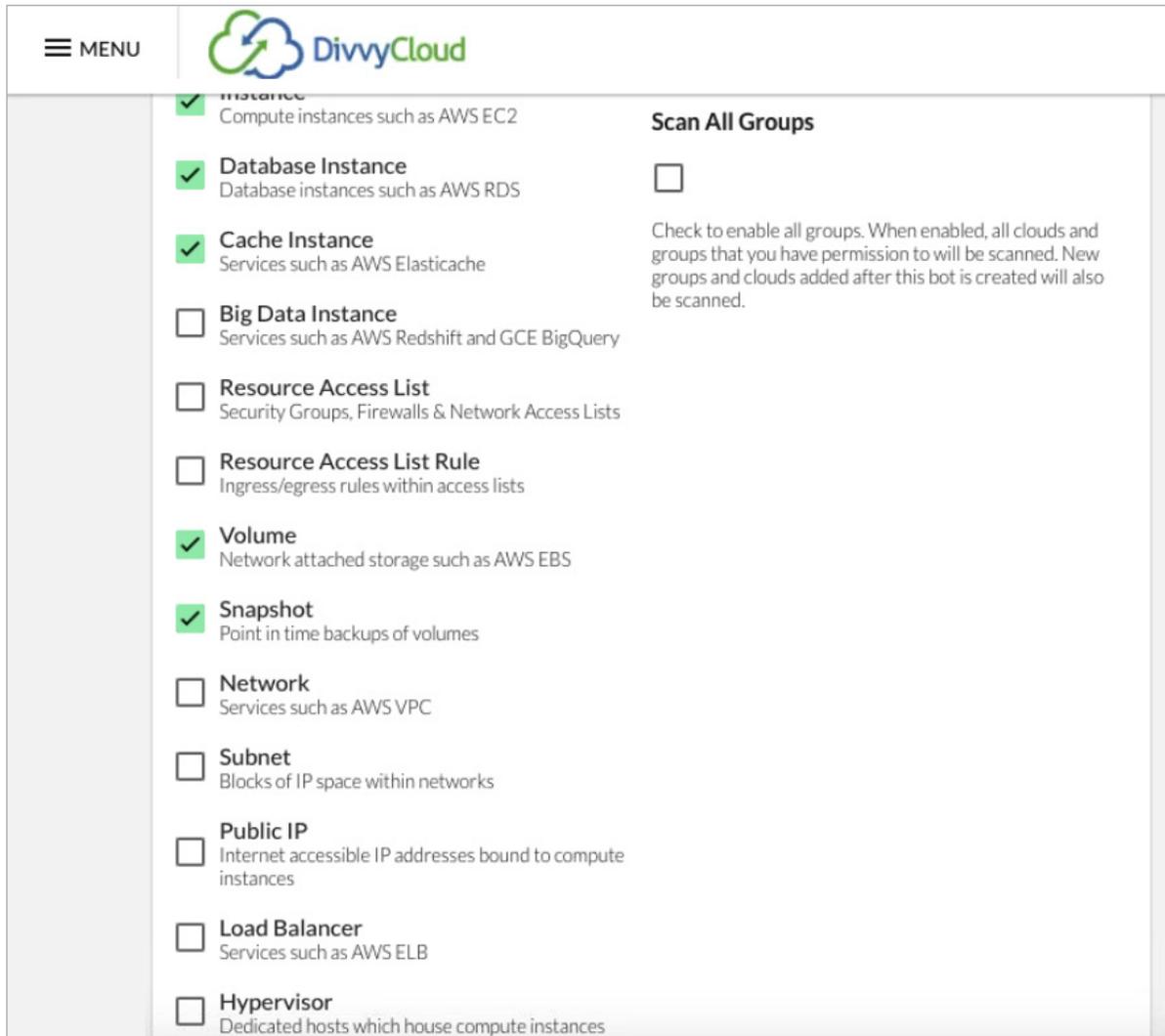
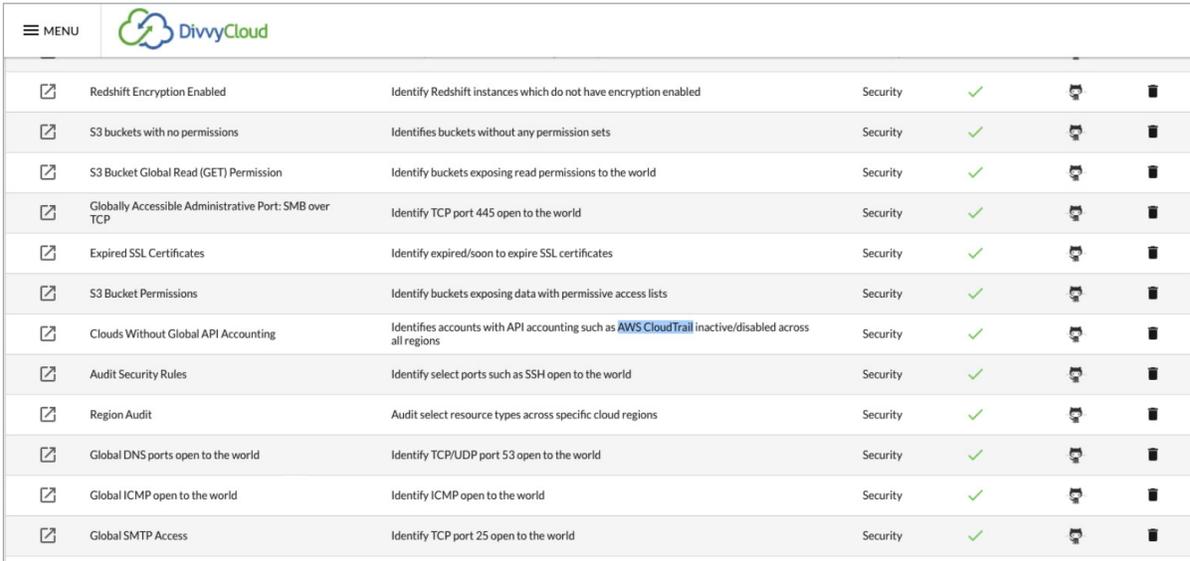


FIGURE 4. The “if-this-then-that” nature of the Bots allows you to monitor and manage a variety of resources across all your cloud platforms with a few clicks.

on GitHub, so you can pull down what you need from the stock Bots or modify them for your custom needs—all within BotFactory.

BotFactory comes with more than 50 Bots that are



Checkbox	Template Name	Description	Category	Status	Bot Icon	Trash Icon
<input type="checkbox"/>	Redshift Encryption Enabled	Identify Redshift instances which do not have encryption enabled	Security	✓		
<input type="checkbox"/>	S3 buckets with no permissions	Identifies buckets without any permission sets	Security	✓		
<input type="checkbox"/>	S3 Bucket Global Read (GET) Permission	Identify buckets exposing read permissions to the world	Security	✓		
<input type="checkbox"/>	Globally Accessible Administrative Port: SMB over TCP	Identify TCP port 445 open to the world	Security	✓		
<input type="checkbox"/>	Expired SSL Certificates	Identify expired/soon to expire SSL certificates	Security	✓		
<input type="checkbox"/>	S3 Bucket Permissions	Identify buckets exposing data with permissive access lists	Security	✓		
<input type="checkbox"/>	Clouds Without Global API Accounting	Identifies accounts with API accounting such as AWS CloudTrail inactive/disabled across all regions	Security	✓		
<input type="checkbox"/>	Audit Security Rules	Identify select ports such as SSH open to the world	Security	✓		
<input type="checkbox"/>	Region Audit	Audit select resource types across specific cloud regions	Security	✓		
<input type="checkbox"/>	Global DNS ports open to the world	Identify TCP/UDP port 53 open to the world	Security	✓		
<input type="checkbox"/>	Global ICMP open to the world	Identify ICMP open to the world	Security	✓		
<input type="checkbox"/>	Global SMTP Access	Identify TCP port 25 open to the world	Security	✓		

FIGURE 5. Built-in templates enable you to analyze real-time data from your cloud resources. BotFactory connects to GitHub to make it easy to pull in and modify what you need.

focused on giving you immediate return on your investment. For example, you can quickly look for servers that are under-utilized or enforce operational best practices across every system you manage.

Standard Bots are designed to fix common cloud problems, such as the following:

- Finding common security vulnerabilities in cloud infrastructure and fixing them automatically.
- Enforcing customer-specific organizational policies, such as tagging or regional data governance.
- Identifying and terminating wasted resources, such as

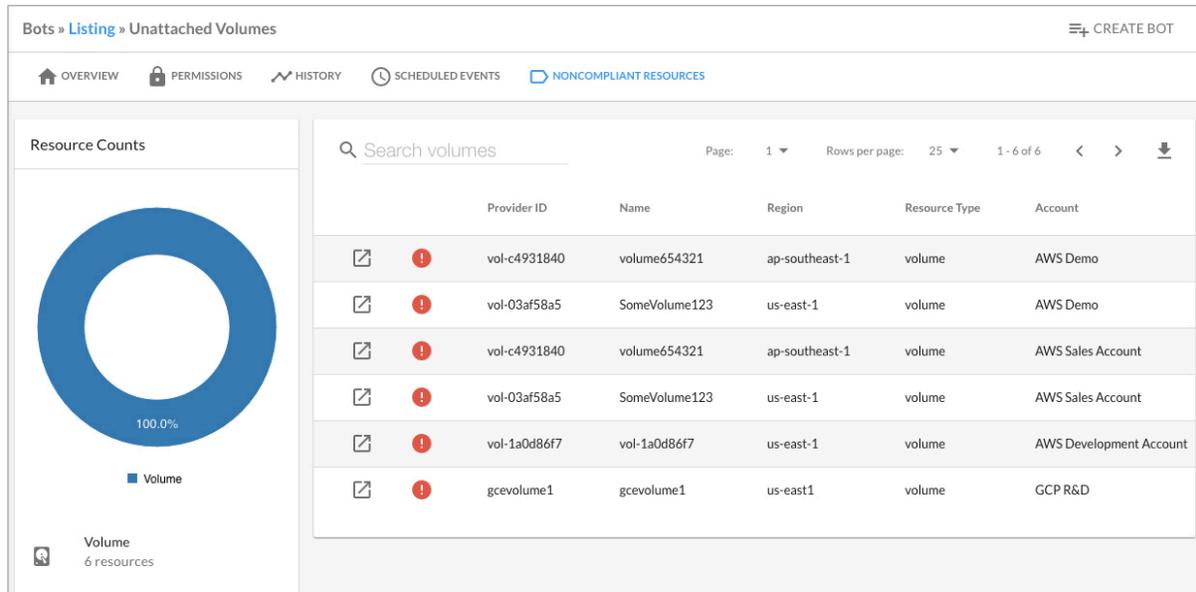


FIGURE 6. Results of Bot activity show you non-compliant resources and enable you to act quickly to fix them.

automatically turning off test and development resources outside business hours.

Active Policy Enforcement Bot Example

By taking advantage of if-this-then-that scenarios, you can enforce policies automatically. This is the sort of self-healing many shops pine for—that is, the ability to set up rules for resources that provide flexibility while programmatically solving problems before they occur.

In the following Bot example, an admin deploys a policy to prevent new server instances from being provisioned in unauthorized or unapproved regions of the country or world. This might be to maintain rapid application response times or avoid higher-cost instances. Such a policy could

help avoid serious security problems or reduce wasted resources, or both.

This particular Bot contains information on whitelisted regions (in this case, Amazon's us-east-1 and us-west-2 data centers) and notifies the admin via email that a resource was deployed in an unauthorized place. The Bot creates a

```
15  "filters": [  
16    {  
17      "negate": true,  
18      "config": {  
19        "regions_list": [  
20          "us-east-1",  
21          "us-west-2"  
22        ]  
23      },  
24      "name": "divvy.filter.resource_in_regions"  
25    }  
26  ],  
27  "groups": [],  
28  "schedule": {  
29    "time_of_day": {  
30      "_type": "TimeOfDay",  
31      "tz": "UTC",  
32      "second": 0,  
33      "minute": 0,  
34      "hour": 3  
35    },  
36    "_type": "Daily"  
37  },  
38  "ondemand_enabled": true,  
39  "actions": [  
40    {  
41      "run_when_result_is": true,  
42      "config": {},  
43      "name": "divvy.action.mark_non_compliant"  
44    },  
45    {  
46      "run_when_result_is": true,  
47      "config": {  
48        "message_body": "Resources located in Blacklisted Regions are non-compliant.",  
49        "sender_address": "info@divvycloud.com",  
50        "hours": 24,  
51        "recipient_list": [  
52          "info@divvycloud.com"  
53        ]  
54      }  
55    }  
56  ]  
57 }
```

FIGURE 7. Bots can be tailored to meet your needs—in this case, allowing instances to be created only in whitelisted regions: AWS's us-east-1 and us-west-2.

hook point based on the “creation” event and takes action. In this case, the action is to send an email message, but it could just as easily be set to schedule the instance for termination and removal.

You can begin to see how Bots can give you quite a bit of control over your environment. For example, you could set up a Bot (or series of Bots) in BotFactory to analyze data harvested across all your cloud resources to see which are being underutilized. Take development VMs. Say your team uses the instances only between 6am and 8pm weekdays. The rest of the time, they’re idle—but still costing you money.

If you have 500 active development instances, the scheduled instance Bot can shut them down automatically from 8pm to 6am as well as Saturday and Sunday. If you’re paying \$.12 per hour for each instance, that Bot would save you about \$6,000 each week—more than \$300,000 a year.

In this same example, these 500 instances—spread across AWS and Azure, say—can be managed centrally. There’s no need to log in and manage any of those VMs separately or set up separate scripts customized to each environment.

Perhaps idle development servers aren’t your biggest problem, but over-sized instances are. We’ve all built VMs with the best of intentions (and guesswork) and spun up servers based on what we think utilization will be. If you’re like most, you add extra resources for good measure, because you’d rather have too much than too little.

Still, there’s significant cost if even 25% of your instances are over-sized. To rein it in, you could set loose a BotFactory Bot to downsize any instance running at less

than 20% of the CPU capacity you've commissioned. In cloud-vendor terms, saving CPUs will save you money.

If you had 500 instances on AWS—125 of which are running at 20% of CPU or less—your optimization Bot could downsize them from C4.4xlarge to C4.xlarge automatically. At a difference of \$.72 per hour, resizing those instances for one week would save you more than \$15,000, or more than three-quarters of a million dollars during the course of a year.

Case Studies

General Electric These sorts of potential numbers got the attention of Thomas Martin, General Electric's application modernization leader, who saw DivvyCloud as a way to support the company's vision to aggregate data across all its cloud vendors. GE also wanted something that could work retroactively and for future deployments—and across all of its business units.

Already a large consumer of cloud services, GE needed a better way to keep track of the resources it deploys to support 10,000 applications worldwide, in both production and non-production environments.

"We needed to be able to look across all our providers and be able to scale, deploy policy and enable our developers the freedom to develop, but at the same time be able to ensure they're staying between the bumper rails of safety", Martin said.

For its part, GE tries to fit the right cloud provider and location to the right purpose, so it takes advantage of internal, hybrid and public clouds. The company's developers work in the different vendor spaces and create the equivalent

of individual Bots, or policies, for those environments.

Good as these were, GE couldn't use these one-off Bots on other vendors' platforms or get data from across all its platforms at the same time.

GE initially deployed DivvyCloud to aggregate resource data across its cloud providers, and it began to look at ways to save money and make it all a little less risky. In particular, Martin said he wanted to make sure they were curating the environment, making it secure and using only what they needed.

His engineers began their Bot journey by building an application inventory. With 9,000 application workloads, they needed to know what cloud infrastructure was tied to which applications. They created application identifiers as tags to help figure that out, and that tagging provided a bit of metadata that fundamentally changed how GE tracked its cloud resources.

They created a Bot they dubbed the "Reaper" to scan across all instances and cloud vendors constantly to make sure all active infrastructure is, indeed, tied to a valid application ID. If not, Reaper kills it off and notifies the resource owner.

To ensure DivvyCloud could work at "GE scale", GE spun up 7,000 servers in the cloud and deployed its new DivvyCloud Bots across a variety of vendors. GE found that it could detect and take action in about 60 seconds of finding a problem.

GE initially deployed 15 core Bots and, like Lego bricks, began building on them. The company has another 15–20 coming online, including ones it's converting from its own previously created in-house Bots.

With five Bots looking for unscheduled production

instances; resizing instance CPU capacity; and shutting down aged snapshots, unused instances and unattached volumes, GE has the potential to save millions annually in cloud costs.

Software AG Government Solutions is an independent, wholly owned subsidiary of Software AG. Customers include some of the biggest agencies in the public sector. Software AG Government Solutions embraced AWS for better elasticity, cost savings and access to compute resources. The team needed a new approach to manage their AWS infrastructure to control costs, manage access and data security, and increase the speed and flexibility of self-service provisioning.

Using Bots for scheduling and security, the team saved thousands of dollars per month and maintained an improved security posture for their development and demo environments.

According to Chris Borneman, Vice President, “DivyCloud provides next-generation automation, management and insights into our cloud infrastructure. Their knowledge and expertise were invaluable.”

Discovery Communications For Discovery Communications, a \$6.5 billion global media company, the core of its business relies on a hybrid cloud infrastructure to host content. Not only is there a lot of it, but it’s accessed frequently, needs to move easily from one region of the world to another, and has to be easy to manage.

The company had both legacy systems and green-field cloud-based infrastructure it needed to manage, with an eye toward making sure it was all secure and fully utilized.

At this scale, Discovery Communications needed to automate as much of its infrastructure maintenance as possible, said Dave Duvall, Senior Vice President of Infrastructure and Support Services.

“There was a lot of discussion among our engineering and architecture teams about the long-standing metaphor of servers as pets vs. herds”, he said. “We quickly agreed that a core operational tenet would be toward an ‘automate everything’ philosophy. Logging in to our servers to troubleshoot or fix a problem would be an exceptional event and an operational failure.”

The company deployed DivvyCloud Bots so it could detect events and act on them, and used them to help manage the migration to a hybrid cloud strategy. By setting up Bots to keep a close watch on everything in real time, the company can focus its staff resources on bigger issues. “Automation allows us to leverage auto-scaling, workload grooming and other approaches to ensure our costs are managed and, most importantly, contained”, Duvall said.

Getting Started at BotFactory.io

If you want to start experimenting with your own Bots, you can set up a hosted BotFactory.io account for free at <https://botfactory.io>. Each free account comes with more than 100 Bots that will do real-time analysis on a single cloud account. You can also play around with automated remediation with any one Bot in “active mode” at any given time. It’s a good way to see how Bots can take automated action to fix problems, not just report on them.

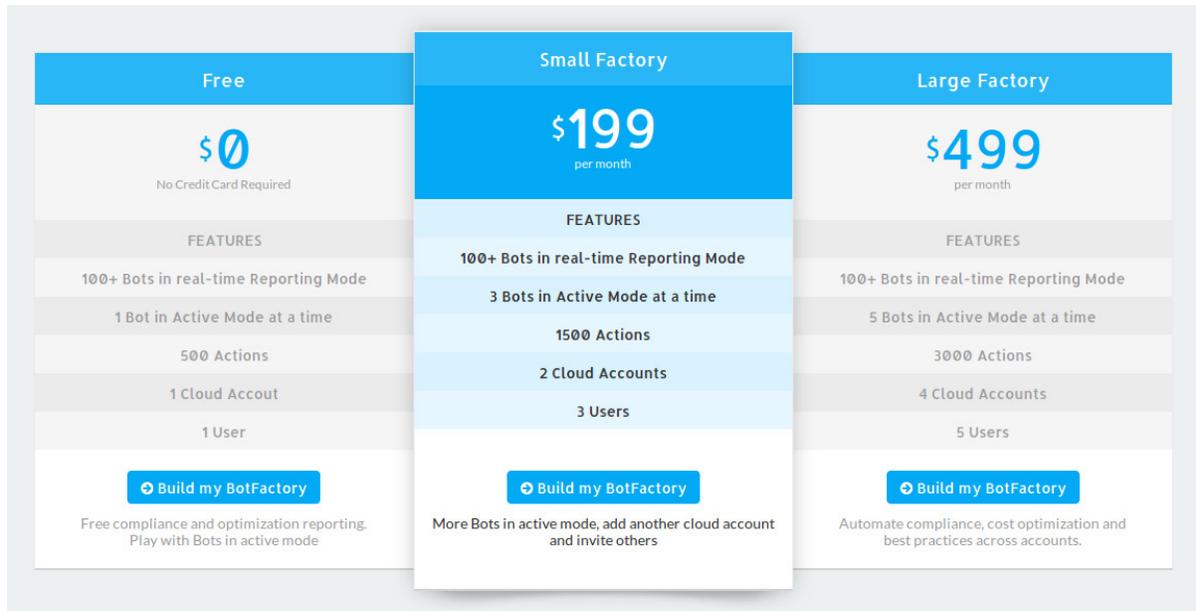


FIGURE 8. BotFactory.io’s hosted solution comes in three options, including a free version that allows you to connect one cloud platform and run one Bot at a time.

BotFactory.io’s Small Factory and Large Factory offerings give you more Bots, actions and cloud accounts. With the Small Factory, which is \$199 per month, three users can put three Bots in active mode and connect two cloud accounts. The Large Factory gives you five active Bots, four cloud accounts and five users for \$499 per month.

For larger needs, DivvyCloud offers its Enterprise version, which includes BotFactory. While BotFactory.io is hosted by DivvyCloud, the enterprise platform is intended to be installed on your own on- or off-premises cloud instances. Enterprise customers also can purchase developer licenses that allow teams to code their own custom filters, actions and integrations for use within BotFactory.

Documentation

To help you get up and running, BotFactory and DivvyCloud offer good documentation, online help and built-in GitHub integration. The API documentation in particular is straightforward and clear, making it easy to set up calls for clouds, projects, resource groups, instances, networks, routes, volumes, DNS and users.

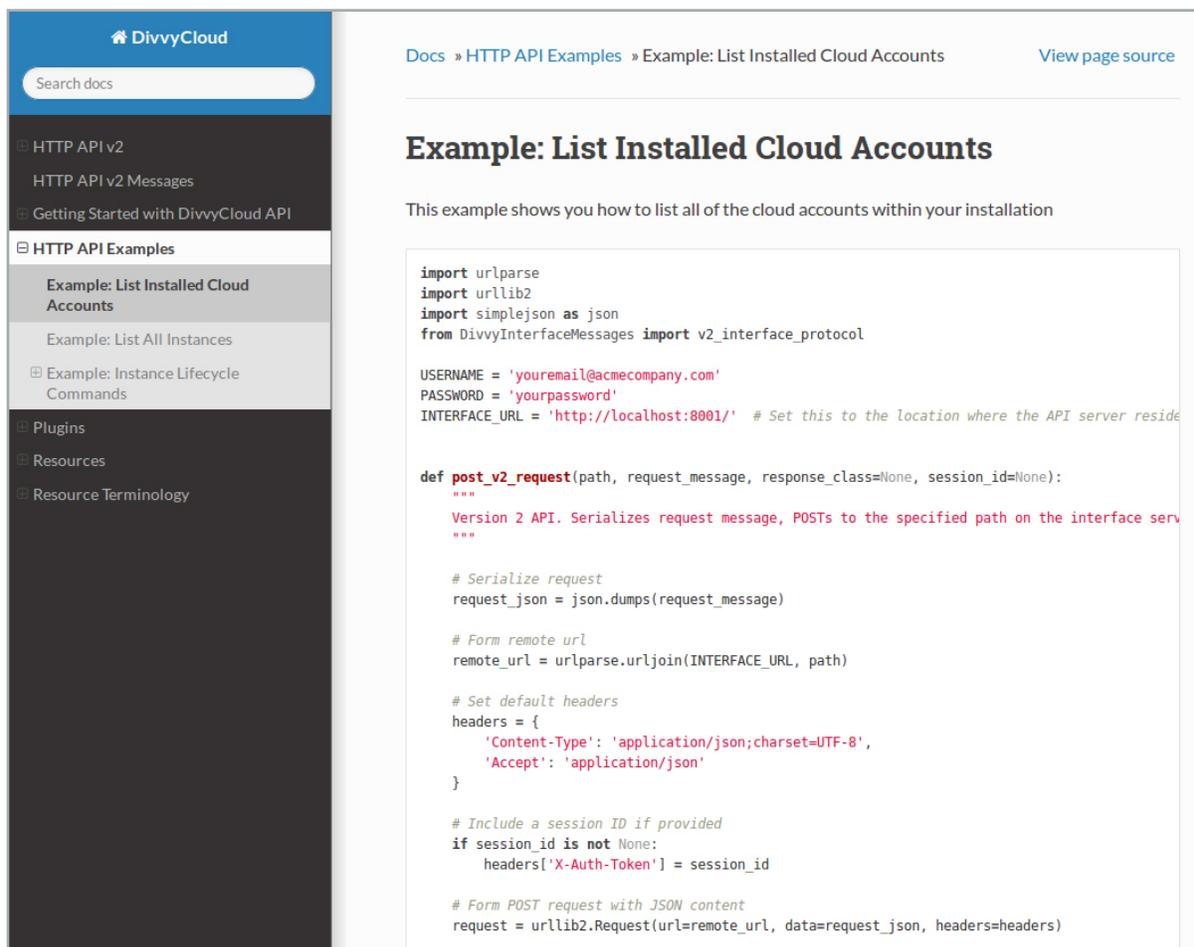


FIGURE 9. DivvyCloud’s API documentation is robust, and it’s a good way to get an idea of how BotFactory uses extensible hooks and plugins to analyze and manage your cloud infrastructure.

Specific examples provide working code and clear explanations of hook points and plugins. The API information actually provides great insight into how BotFactory and DivvyCloud automation works, and it's a good place to start exploring capabilities.

If you're eager to find a way to manage and automate your ever-growing cloud infrastructure centrally, BotFactory.io certainly can make life easier—and keep the concern at bay.■

Resources

BotFactory.io: <https://botfactory.io>

DivvyCloud: <http://divvycloud.com>

Supported Clouds:

<http://divvycloud.com/solution/supported-clouds>

Connecting to Clouds: <http://support.divvycloud.com/knowledgebase/topics/70317-clouds>

Documentation: <http://support.divvycloud.com>

API Documentation: <http://docs.divvycloud.com/api/15.16>