

tiagokatcipis

software engineer

contact

+(55)(48)991060132

tiagokatcipis@gmail.com

GitHub

LinkedIn

Blog

languages skill

English: advanced
(reading, writing);
intermediate (speaking)

programming languages

C, C++, Go, Python,
Lua, Nash, Bash,
Javascript

protocols

DBus, RTMP, HTTP,
SIP, RTP, AMQP

cloud

Kubernetes, AWS,
Azure, Docker,
CoreOS, Terraform

documentation

Gtk-Doc, Doxygen,
Docbook, LDoc,
LaTeX, Docco

automation

Ansible, Make

dev environment

Vagrant, Docker
Compose

monitoring

Prometheus, Grafana,
StatsD, Sysdig

education

2006–2011 **Bachelor of Computer Science**

UFSC

experience

2018–2019 **Neoway**

Software Engineer - Lambda Team

Florianópolis, Santa Catarina

As a software engineer on the lambda team I was responsible for researching and developing new captcha solvers (audio and image). Solving audio captchas involved studying and applying the following algorithms and tools:

- Noise removal using sox
- Audio segmentation by silence using sox
- Extracting features (MFCC) from segmented audio
- Normalize extracted features with rescaling
- Train SVM model
- Recognize characters using trained SVM model
- Label image captchas using TensorFlow

Main accomplishments:

- Development of tool to measure assertiveness of captcha solving services
- Solved audio captcha with 80% of assertiveness
- Solved image captcha with 40% assertiveness using TensorFlow
- Tooling to automate dataset and trained models upload at Azure BlobFS

2017–2018

Neoway

Florianópolis, Santa Catarina

Software Engineer - Data Platform Team

As a software developer on the data platform team I started to be responsible for a more broad context involving multiple teams, building not only services but also tooling to be used by those teams.

Besides the technical challenges of providing a platform I also performed code reviews and coaching across multiple teams. There was a lot of challenges like:

- Migrating the entire infrastructure of multiple teams from AWS to Azure
- Defining network topology and security measures
- Creating backups for VMs with different operational systems at Azure
- Specify a new DSL to develop web scrappers
- Scale the ingestion of huge CSV files (terabytes per file)
- Define solutions for infrastructure building that crossed multiple teams boundaries

Main accomplishments:

- Fully automated infrastructure provisioning at Azure
- Easy to build parallel version of the entire infrastructure (including networking and security)
- Development and testing of a library to automate infrastructure building
- Specified and developed generic backup strategy at Azure
- Entire infrastructure access through a bastion host as a security measure
- New service that allowed arbitrarily huge files to be ingested easily (scaling horizontally)

2015–2017

Neoway

Software Engineer - Data Capture Team

Florianópolis, Santa Catarina

Technical lead of the team responsible for capturing data from the web and publishing it to the entire company, which involved solving hard problems as:

- Scraping the web (Scrapy and Selenium)
- Parsing data from multiple formats (HTML, PDF, SWF)
- Maintaining all downloaded raw data at S3
- Defining and documenting good protocols for proper service integration
- Developing services on different languages, like Go and Python
- Migrating the whole data pipeline to a new service oriented architecture
- Coach the team on good test techniques for the new architecture
- Developed new services to provide good proxies and captcha breaking

Besides the technical challenges I also helped the team to apply some development practices, like TDD (Test Driven Development), code review and continuous integration. Together we built a DevOps culture to enable infrastructure as code on our team, we were responsible for the whole solution, from the development to testing and deployment (including monitoring the production system).

Main accomplishments:

- Technical leadership and coaching for a team of 6 people
- Built a new crawling framework, from scratch to production
- Fully automated development environment with docker compose
- Fully automated dashboard and alert creation on Sysdig Cloud
- Cluster orchestration at AWS/Azure using Kubernetes, CoreOS and Docker
- Migrating all infrastructure from Codero to AWS
- Later, migrated the infrastructure from AWS to Azure
- Active participation on building the tools to automate infrastructure on Azure
- Aided on the migration of more than 100 crawlers to Kubernetes
- Actively participated on the screening and interviews of new candidates for the team
- Coached new members
- Implemented a real time monitoring for the data production pipeline using StatsD and Sysdig
- Gave talks inside the company and on events like TDC (The Developers Conference) and GopherCon

2012–2015 **Dígitro**
Lead Software Engineer

Florianópolis, Santa Catarina

Worked on the development of a VoIP phone with touchscreen in a very strict environment (a Blackfin DSP processor with no MMU and 64MB of RAM). The project was built from scratch in C and used DBus to integrate different processes. This project involved learning considerably about SIP and VoIP state machines and also memory issues because of the lack of an MMU, like external memory fragmentation. The development environment was rather complex and involved cross compilation with very specific tools. This build of this development environment was done using Vagrant and Ansible.

Some refined functions on the VoIP phone required integration with our PBX that used a rather complex proprietary protocol. To solve the integration problem we developed a PBX REST service in Lua that integrated with the current legacy protocols exposing them in a more web friendly way. This service was implemented in Lua.

My last project at Dígitro was a solution to web audio playback with very specific audio effects (like silence removal, change in pitch) that had to be developed using HTML5. It was the rewrite of an old system that worked using Flash. Some of the components of the old project, like Gstreamer plugins, were reused, but the service itself had to be rewritten using REST and HTTP to do the media streaming (and control). This was developed in Javascript (NodeJS) integrating with the C code for streaming using IPC (Inter Process Communication). The original audio files were encoded in IMA ADPCM and another proprietary CODEC and were exported through SFTP. Part of the problem was also integrating with this legacy infrastructure and transcode it to a more web friendly format.

During these projects I always practiced and advocated TDD (even on embedded systems in C) and helped people on the team to write automated tests for their code. I also coached two new members in the team.

2010–2012 **Dígitro**
Software Engineer

Florianópolis, Santa Catarina

I started working on a solution to web audio playback with very specific audio effects (like silence removal, change in pitch) that had to be developed using Flash. To solve that problem I worked with two different open source C++ projects that did reverse engineering of the RTMP protocol to develop our own Flash Media Server. I worked directly with the integration of the server playback logic with Gstreamer and the plugins that enabled the desired effects on playback.

The next project was a solution to biometric identification. This has been solved by building a service around a third party C library that built and scored voice models. I developed a REST service in Lua that binded to the C code that built the voice models and used MongoDB to store the voice models and perform searches on the database.

Then I developed a service to perform searches for specific words in audio files. The service would replace a old one so it had to implement the same textual/proprietary stateful protocol of the old one. The service itself was implemented in Lua with bindings to C code that did the actual search of words.

2008–2010 **Dígitro**
Trainee

Florianópolis, Santa Catarina

I started helping in the development of an cross platform (Windows and Linux) audio streaming library for a VoIP softphone, aiming at porting the current application that was Windows only to Linux. The library was C code being cross compiled to Windows using mingw.

After that I got involved developing a prototype of a voice biometrics system. It involved building a small web server in Python that binded to third party libraries that did the build and scored voice models and a cross platform Python application that provided a user friendly GUI to the web server.

2007-2008 **Cyclops / LAPIX**
Trainee

Florianópolis, Santa Catarina

Worked on adding new features on the system responsible to integrate medical equipment to the DICOM system. I developed a cross platform application that generated a customized GUI based on a XML description of a form, making it easier to people working at hospitals to manipulate the forms and save them back at XML. Like a domain specific XML editor.

This involved learning C++ and XML parsing, together with developing cross platform GUI applications, on this case using WxWidgets. The code has been tested using CppUnit.

I also automated the setup of the development environment using Python.

open source projects

2017-now	mdtoc A very simple table of contents generator for markdown.	https://github.com/madlambda/mdtoc
2016-2018	nash Nash is a shell language focused on simplicity and having a nicer syntax than traditional shells and support to containers. It also strives to be safer than traditional shells.	https://github.com/NeowayLabs/nash
2016-2018	klb klb is used to automate infrastructure creation on AWS and Azure. I got involved on designing the support for Azure since this was the tool used to migrate Neoway production infrastructure from AWS to Azure.	https://github.com/NeowayLabs/klb
2013	CppUTest CppUTest is a C /C++ based unit xUnit test framework for unit testing and for test-driving code. In this project I worked both on improving the documentation and at adding new native types to the mock framework (which involved some refactoring).	http://cpputest.github.io
2012	GStreamer GStreamer is a library for constructing graphs of media-handling components. I contributed with a plugin named <i>removesilence</i> and some documentation for the GstCheck documentation.	http://www.gstreamer.net
2010-2011	Pattern detection on H.264 This is my Bachelor's Thesis and it consists of a prototype of a H.264 CODEC that uses OpenCV and H.264 internal algorithms to do pattern detection and object tracking integrated on the encoding process. Metadata generated on the encoding process is integrated on the video bit-stream on conformance with the standard.	https://github.com/katcipis/h264.pattern.detection
2010-2011	LuaSofia Lua binding for the Sofia-SIP library. Contributed to the project from the start, helping to make decisions about the design of the software and documenting it.	https://github.com/ppizarro/luasofia
2010	GPS tracking system System designed to provide the location of a device at the receive of a position request using SMS.	https://github.com/katcipis/gps.tracking
2010	LuaNotify Lua library that implements a simple Pub/Sub system inspired on glib GSignal API.	https://github.com/katcipis/luanotify

presentations

2018

Object Orientation in Go

The Developers Conference

For people that come from a background on Java or other classic object oriented languages (like C++) there is also some discussion on if Go is actually object oriented.

In this presentation I try to present Go as a language that is more object oriented than these classic languages, at least according to the original foundations of object orientation.

Presentation source can be found [here](#).

2016

Building Resilient Services in Go

GopherCon Brazil

Resilience is not about never failing, but how do you recover from it. How can you prevent your services from locking down or exhausting all its resources ? How to perform graceful service degradation ? Can this kind of behaviour be tested properly ?

On Go we have some new features, like Contexts, that helps us to model timeouts and cancellation properly.

They can be combined with other useful features as select and channels to model timeouts and resource pools, which can be essential to provide proper service degradation instead of total failure of the system.

On this talk I try to answer this questions using new features available on Go 1.7, direct from production ready software.

Presentation source can be found [here](#).

2016

Real Life Kubernetes

The Developers Conference

On this presentation we will give a short introduction on Kubernetes and show the experience of learning and using Kubernetes on production for two very distinct systems.

The first one is a data acquisition system, involving running multiple instances of different crawlers, storage services, captcha breaking services, message brokers (like RabbitMQ) and database integration outside the cluster.

The second one is a web application, involving network analysis using graphs with the ultimate goal of fraud prevention. The application is strongly bounded with the microservices architecture and the twelve factor app.

Graph and document databases, cache layers, a message broker and a distributed filesystem are some of the technologies surrounding the application ecosystem.

Presentation source can be found [here](#).