Julia for Infrastructure

Ajay Mendez
ajay@kinant.com

Agenda

- Julia for Startups
  - Our journey and why Julia made sense

- Julia for Infrastructure
  - How we used Julia to build a *data ubiquity platform*

Our Journey

Datastore for Backups and Archives

- Compression at scale
- Scale horizontally using commodity nodes

Data Ubiquity Platform

- Checkpoint, move and share data
- Like git for data

Data Governance Platform

- Is sensitive data being copied and shared?
- How much can be saved by eliminating redundancies?
- Is it easy to find all relevant data sets for an analytics job?
- Are you sure all copies of data marked for deletion are removed?

Fail Fast and Fail Early

Idea!

- Compression at scale
- Find redundancies in petabytes
- Prototype in C

"No matter the programming language chosen, a professional developer will write an average 10 lines of code a day."
-- Fred Brooks, The Mythical Man Month

"The only way to get software written faster is to use a more succinct language"
-- Paul Graham, Succinctness is Power

kinant

Build Something Useful Fast!

December ‘16
Datastore for backup and archive. C Prototype.

January ‘17
Port to Julia

March ‘17
Pivot 1: Data Ubiquity Platform

May ‘17
Pivot 2: Data Governance Platform

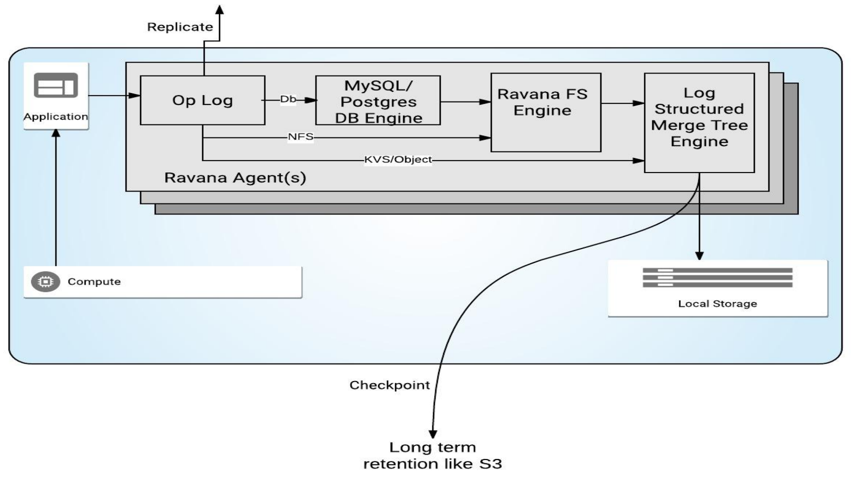
November ‘17
Pivot 1 complete had we used C/C++

Dec Jan Feb Mar Apr May Jun Jul Aug Oct Nov Dec

Building A Data Ubiquity Platform

Ravana.jl

- Like Git for large data
- Persistent cache
- S3 for long term retention
- Replicated oplog for availability
- Fast checkpoint, clone and restart



Ravana Architecture

What Worked

- Debugging - REPL to the rescue
- Rapid prototyping - the prototype is the product
- Forget about on disk formats!
- Building distributed systems with remotecall()
- Increase throughput and responsiveness with tasks

Challenges

Challenge	Work Around
Lack of threads	Use @threadcall() judiciously. Not an elegant solution. Multiplexing **m** tasks on **n** threads is the way to go.
Buffer bloat	Use ring buffer
Hard to ship binaries	Use PKG3?
Language instability/compatibility	Waiting for 1.0

Summary

Julia is great for infrastructure projects!

Julia is a competitive advantage for startups!

contactus@kinant.com
ajay@kinant.com