



Legion Performance Analytics

Strategies and Tactics

May 2022

Legion Performance Analytics

Introduction

- logs, metrics, traces
- latency, stability, satisfaction
- for local, distributed & cloud native applications



Legion Performance Analytics Strategies

Table of Contents

- Whole stack solution
- record cheap, read maybe
- All the data, one protocol
- From inception to live
- Progress & roadmap

Legion Performance Analytics Strategies

Whole stack solution

- off the shelf components are not good enough
- instrumentation
 - low overhead (~ 40 ns / event)
 - generic and flexible format (like protobuf with references)

Legion Performance Analytics Strategies

Whole stack solution

- database
 - scalable in writing
 - low cost when unused
 - bursty reads
 - write like a data lake, read like a data warehouse

Legion Performance Analytics Strategies

Whole stack solution

- user interface
 - flame charts with billions of entries
 - graphs with (at least) thousands of nodes
 - tight integrations with time series and lists
 - web based
 - mashup of rad telemetry + prometheus + kibana + grafana

Legion Performance Analytics Strategies

Table of Contents

- Whole stack solution
- **record cheap, read maybe**
- All the data, one protocol
- From inception to live
- Progress & roadmap

Legion Performance Analytics Strategies

Record cheap, Read maybe

- low overhead instrumentation
 - thousands of events per frame
 - recording is serializing with heterogenous queue
 - platform-specific memory layout
 - batching
 - fast compression using lz4

Legion Performance Analytics Strategies

Record cheap, Read maybe

- cheap ingestion
 - Event block payload in S3 without decompression
 - MySQL: metadata about processes, streams and blocks

Legion Performance Analytics Strategies

Record cheap, Read maybe

- pay for what you read
 - ETL on demand
 - decompression of structured event blocks
 - parse events to build trees and graphs
 - write in parquet on S3 with lambda
 - query using AWS Athena & datafusion

Legion Performance Analytics Strategies

Table of Contents

- Whole stack solution
- record cheap, read maybe
- **All the data, one protocol**
- From inception to live
- Progress & roadmap

Legion Performance Analytics Strategies

All the data, one protocol

- Structured events
 - time series are not general enough
- Stream definition contains memory layout of events
- Instrumented apps are free to upload any event in any stream
 - analytics relies on tagged streams
 - analytics expect and process specific event types
- Forward & backward compatibility

Legion Performance Analytics Strategies

All the data, one protocol

- Custom binary protocol could be extended
 - Crash dump
 - Images
 - Video

Legion Performance Analytics Strategies

Table of Contents

- Whole stack solution
- record cheap, read maybe
- All the data, one protocol
- **From inception to live**
- Progress & roadmap

Legion Performance Analytics Strategies

From inception to live

- Development
 - High event density
 - low constant costs
- Live
 - High scalability
 - Fast adaptability
 - configure output verbosity of instrumented app
 - to-the-minute live data

Legion Performance Analytics Strategies

Table of Contents

- Whole stack solution
- record cheap, read maybe
- All the data, one protocol
- From inception to live
- **Progress & roadmap**

Legion Performance Analytics Strategies

Progress

- Instrumentation libraries: Rust, Unreal
- Ingestion in the cloud
 - Rust on k8s (gRPC + http)
 - MySQL Aurora Serverless + S3
- Analytics/ETL
 - Rust on k8s (gRPC-web)
 - Cache on S3
- UI: Svelte, Typescript, Canvas

Legion Performance Analytics Strategies

June Priorities

- Regulations: GDPR, Pipedata, bill 64
- UI improvements (logs, l10n/i18n, timeline, metrics)
- Unreal module
- Lakehouse: just-in-time parquet generation + query engine

