

# CAPE: Configurable & Automated Production Engine

August 28, 2017



**INTERMAP.COM** 

# **The Automation Challenge**

One of the most common challenges our customers face is the difficulty in integrating geospatial data and new technology with existing data and technologies. The complexity comes from managing large volumes of data, a plethora of tools, and the highly-detailed process flows that are necessary for producing results. These challenges include:

- · Performing complex analysis, extraction, merging, or creation of data sets
- · Automating workflows to produce repeatable, high throughput, accurate results that interact with the user
- Integrating with existing tools to combine the best available technologies
- · Accelerating creation of consistent results, simulations, and optimizations

Intermap was one of first in the industry to recognize the need for a scalable infrastructure capable of modernizing workflows and solving these complex problems.

### Introducing CAPE

To meet these demands, Intermap developed a Configurable & Automated Production Engine (CAPE). The goal of CAPE was to allow Intermap and its customers to solve large scale Automated workflows bridge the gap between users, applications, and information, problems that were too complex, difficult, manually intensive, or susceptible to process failure.



accelerating an operator's ability to create meaningful results

#### **CAPE Key Benefits:**

- Perform high volume, complex analysis, and production that would otherwise be impractical.
- · Get results quicker by accelerating complex workflows through automation
- Maximize the value of existing tools, processes, and datasets by connecting them together in a single environment
- Save thousands of man hours in labour by automating tasks, decisions and complex calculations
- Increase quality and reliability by tracking, managing, and eliminating mistakes and bottlenecks
- · Scale your business to petabytes of data and hundreds of processing nodes

The initial release of CAPE was an immediate success, providing automated production and analysis of global datasets for multiple simultaneous customers. CAPE has successfully completed hundreds of projects around the world in geospatial, telecom, insurance, advertising, and government applications.

In one project alone, through automation, CAPE has eliminated over 100,000 + man-hours of labour, processed over 20PB of data, and performed complex calculations of tens of thousands of CPU hours.

# **CAPE Technical Description**

CAPE is a software infrastructure that can be configured to run multiple distinct production workflows. A CAPE production system combines manual and automated steps to produce a product or a result. CAPE provides the following key features:

- Definition of complex workflows and processes which automate tasks and perform complex logic (like loops and conditional statements) to answer key business questions
- · Execution of 3rd party applications and Web Services
- Customizable interface for operators that allows them to interactively perform quality control (QC) steps, provide input, and monitor status
- · Monitors metrics on performance of processes, operators, jobs, and resources
- Horizontally and vertically scalable platform capable of distributing jobs and processes on hundreds of nodes and CPUs

CAPE's architecture allows multiple projects to be set up within a single production system that uses different inputs and operators. This allows one system to be run by multiple users in an organization. The CAPE architecture is very flexible as it allows full parallelization of resources. Entities may be used in multiple jobs, and processing is automatically scaled to the available hardware depending on the workflow. Processing can be configured for parallelization through various rules in a workflow. CAPE is composed of five main components which are shown in the following diagram and described in the sections that follow.

CAPE



Illustration of the main components of the CAPE Architecture, including the Process Manager, CAPE Database, Load Balancer, Web Server, and the Processing Nodes

#### **CAPE** Database

The CAPE Database stores configuration information for CAPE and all workflows it is configured to run. It also stores information about specific projects that will be, are being, or have been run within a production system. Specifically, data is stored for:

- System configuration and generalization
  - Information that is required for the system and subject to change depending on the process
    Resources, users, configured processes
- Process management
  - Configuration information for each process including entity types, job types, job commands
  - Information for each instance of a process including specific job information type, command, status, and priority
- Load balancing
  - Information on processing servers: availability, address, memory, CPU, I/O, and load
  - Information on jobs: Estimated memory, CPU, I/O required, type of job, etc.
- Metrics generation
  - Complete history of all job execution and entity manipulation is stored, allowing metrics to be derived

### Load Balancer

Load balancing is implemented in CAPE to ensure jobs are distributed evenly across available processing servers. The load balancer determines which jobs can be run based on CPU, memory, I/O, and other requirements. In a typical configuration, there are more available jobs than processing servers, so a backlog of jobs is formed and the load balancer sends those jobs to the servers when computing resources become available. CAPE can prioritize specific jobs, but if all priorities are the same, then the system processes the jobs in the order generated.

The load balancer also monitors jobs for completion. The status of each processing node is monitored, and if a server goes down, the jobs previously assigned to it will be reassigned. Workflows and jobs can be cancelled, and the load balancer will ensure all running processes corresponding to those jobs are killed.



An Example of a CAPE workflow used in automated data generation. The green boxes represent jobs, blue boxes represent data (entities), and orange boxes represent decisions.

### **Process Manager**

The Process Manager node monitors workflows and performs the logic contained in the workflow. It monitors the status of each job and the steps in the workflows. It also performs logical decisions that add jobs, configure jobs, link jobs, and determines job execution order. The Process Manager can also be triggered on things like changes to data, rework, or user interaction before determining which step / action to take.

### **Processing Node**

CAPE Processing Nodes perform the heavy lifting by running jobs as directed by the Process Manager. Processing Nodes can be added or removed as required by the system. Each Processing Node:

- · Executes jobs locally that are arbitrary, and may include batch files, executables, and web services
- · Monitors the job for any abnormal conditions, and restarts if necessary
- Manages local resources to maximize job performance
- · Captures metrics of running jobs, and reports status to the database

#### Web Service

CAPE's Web Service allows external operators to control and monitor the CAPE system. The Web Service provides two main interfaces: a Web Management User Interface for operators and a web service interface for integrating with applications.

The Web Management User Interface is accessible through a standard web browser by an operator. It allows the operator to use the CAPE system to:

- Set up and monitor projects through customizable easy-to-use pages
- Monitor system performance and status of production systems, bottlenecks, or results
- Notifies when jobs are ready or when actions are needed
- Interact with results or make decisions about the next course of action

The Web Services interface allows 3rd party applications to connect with CAPE.

- · Start new jobs or production systems
- Change jobs, configure jobs (may require customization)
- · Monitor the status of a production system
- · Get the results of a workflow



Example of CAPE Web Management User Interface. In this example, an operator is configuring the "Universal Projector" project for generation of a new data set

# **Case Study: Lloyds of London**

## The Problem

Lloyd's of London, a billion-dollar insurance firm, was unable to underwrite flood insurance effectively because it could not assign a price to the risk at specific property locations. Its existing process involved many tools and datasets which were of poor quality and inefficient. Many different groups within its business had different ways of solving the problem of determining risk, which created confusion and sub-optimal underwriting. Lloyd's came to Intermap for a streamlined system to combine its processes and analyses into a single configurable engine that could generate the answers they need with the push of a button.

# The Solution

Intermap provided the geospatial data, measurement, and modeling expertise for the complex calculations. In addition, Intermap provided custom automation to rapidly calculate flood risk information to Lloyd's. Intermap combined all data required for underwriting into a single access point. This included:

- High resolution bare-earth elevations
- Precise water locations
- · Measurements of height above water and other flood baselines
- · Measurements of distance to water
- Accurate flood models

To deliver a solution to underwriters, Intermap's insurance software (InsitePro) combined all information into a single risk score, using an automated geospatial analytic that activates with a single button push. Underwriters are now able to quickly and easily assess flood risk and apply pricing to specific properties based on dependable and high quality data.

### The Outcome

Lloyd's underwriters are able to get highly accurate risk scores in just a few seconds. The results are both repeatable and reliable and have allowed Lloyd's to competitively expand its business during its migration from government flood insurance to private flood insurance in the United States.



A one-button Risk Score for underwriting flood insurance at a very specific property. Underwriters can use this output to quickly and accurately assess insurance risk.



# **Case Study: New South Wales**

### The Problem

In 2012, the Land and Property Information division of the government of New South Wales in Australia realized it had a problem needing a creative solution. Beginning several years earlier, it began a state-wide data collection program using airborne optical imagery platforms to collect high-density stereo imagery for over 600,000km2 of the state. In addition to a complete image dataset, it wished to turn the stereo data into a seamless, highly accurate bare earth Digital Terrain Model (DTM), but didn't have the manpower, tools, or time to manipulate the vast quantities of data they had collected into the desired output. With the clock already ticking toward the desired program completion date, they needed to look externally to find a partner to produce the desired products, with the required accuracy specifications, in the limited time available.

### The Solution

Intermap was able to respond to the call for help with a unique and effective solution capitalizing on its proprietary technology, skilled workforce, and extensive experience with large data production projects. The key to Intermap's solution was utilizing its CAPE system to shepherd the large volumes of raw data from New South Wales through a complex workflow involving

many different tools. The process of Digital Elevation Model (DEM) generation, and the management steps required to create the desired outputs in an efficient, inexpensive, and timely manner, were executed using mostly automated processes enabled by CAPE. A proprietary DEM Fusion process ensured a consistent and seamless product. A final DEM editing process was completed by a skilled team of experts using ISO 9001:2008 processes to create the highly accurate DTM product sought by the client.

### The Outcome

The initial pilot project awarded to Intermap allowed New South Wales to evaluate Intermap's CAPE system against other vendors using different methodologies. The quality, consistency, and capability for lowcost and high-volume processing enabled by CAPE has led to 6 consecutive years of solesource contract awards, helping the client stay on track to complete its massive statewide program by 2018.



New South Wales DSM (above) and final DTM (below). The DTM was created utilizing a combination of automation and manual steps in a CAPE workflow. The approach produced a low-cost, high-quality product that was on schedule and under budget.



# About Intermap www.intermap.com

Intermap Technologies is a global geospatial solutions leader focused on improving the ways in which people, businesses, and governments use location-based information. Intermap provides products, services, and solutions that help businesses and governments solve their geospatial challenges.

Copyright 2017. Intermap Technologies<sup>®</sup>. All rights reserved.