
L7|ESP SDK Documentation

Release 3.3

L7 Informatics

Sep 08, 2026

APPLICATION

1	Overview	1
1.1	Getting Started	1
1.2	Project Structure	4
1.3	Administration	7
1.4	Introduction	8
1.5	Content	9
1.6	Extensions	11
1.7	Testing	24
1.8	Windows Setup	34
1.9	Overview	36
1.10	Create Workflow	38
1.11	Create Test	43
1.12	Run Tests	47
1.13	Seed Content	48
1.14	Summary	49
1.15	Container Infrastructure	49
1.16	Content Installation	101
1.17	Hardware Recommendations	103
1.18	Database Configuration	104
1.19	System Tuning	121
1.20	HA Cluster Setup	124
1.21	Database User and Grants	125
1.22	Unix Application User	125
1.23	Application Startup on Boot	126
1.24	L7/ESP Revision History	127
1.25	Security Considerations	127
1.26	Backup and Disaster Recovery	131
1.27	Questions/Feedback	134

OVERVIEW

The L7|ESP SDK is a collection of software tools used for building content within L7 Informatics' Enterprise Science Platform. With the server development kit (SDK), you can do things like:

- Create a locally-running copy of L7|ESP for testing out the software.
- Develop *Workflows*, *Protocols*, *SampleTypes*, and other content for L7|ESP using programmatic definitions.
- Develop integration tests for verifying content functionality.

Having a set of tools for developing with the L7|ESP platform is paramount to leveraging what L7|ESP has to offer. This set of tools allows developers to create L7|ESP content programmatically and outside of the L7|ESP API, which drastically speeds up the development and testing cycles for customizing an L7|ESP installation.

Application Run L7|ESP locally and review the structure of a L7|ESP SDK project.

Development Discover how L7|ESP can be configured and extended to meet any process need.

Tutorial Learn how to develop and test content within L7|ESP using the L7|ESP SDK.

Deployment Guidance for running the L7|ESP platform successfully in production.

1.1 Getting Started

You will first need to install a container runtime as the L7|ESP software is distributed as an OCI-compliant container image.

1.1.1 Container Runtime

Docker Desktop or equivalent (e.g. Rancher Desktop)



A tool designed to make it easier to run container image(s) so you can focus on development, testing and deploying content to L7|ESP.

The L7|ESP SDK is delivered as a container image, and contains the following tools:

L7|ESP - Enterprise Science Platform



The entire L7|ESP server software suite is at your fingertips. Create local or remote environments for development in just a few minutes.



Git - Version Control System

Software version control system used collaboration and change tracking between developers who are working on the same L7|ESP SDK project in parallel.



Conda - Pipeline Python Environment

Python distribution and language-agnostic package manager used when writing custom pipeline scripts and integrations throughout development.



Ansible - Configuration Management Tool

IT automation tool used to prevent configuration drift when installing the same content or customizations to different L7|ESP environments.

1.1.2 Project Directory

An L7|ESP SDK project is a directory that contains all managed content and configuration for a given L7|ESP site and all its **DTAP** environments.

If you received a deployment bundle, you can use `tar` to extract the project directory:

```
$ tar xf deployment.tar.gz
```

To create a new L7|ESP SDK project, you can use `cruft` to copy a template repository:

```
$ pipx run cruft create git@bitbucket.org:lab7io/customerrepo_template.git
```

To clone an existing L7|ESP SDK project, you can use `git` to clone the repository:

```
$ git clone git@bitbucket.org:lab7io/acme.git
```

Note: The latter commands require access to L7 Informatics' Bitbucket account.

Throughout this documentation, we'll refer to an `acme` project in examples. ACME Corp is a fictitious project used during product testing and, therefore, does not represent a real entity.

1.1.3 Start Application

Typically, the server, the web UI, and the Python client will be included as a part of the L7|ESP SDK image.

Docker Compose

Most L7|ESP SDK projects use Docker's Compose plugin to manage container life-cycle. Compose is a tool for managing container configuration in a single file called `docker-compose.yml`.

Confirm Version

It's highly recommended that you ensure you are using the latest L7|ESP SDK image at any time, so you have the latest L7|ESP patches (security or otherwise).

You can confirm this by checking the first line in `Dockerfile`:

```
$ docker compose config | grep image
image: l7esp/server:3.0.0-sdk.N
```

Start

Once you've installed all prerequisites and requirements for using the L7|ESP SDK, you can start a local instance of the application using Docker Compose:

```
$ docker compose up --detach
```

This task will take a few minutes while it pulls the L7|ESP SDK container image and creates a running L7|ESP SDK container for you to develop in. If you run into any issues, please contact [L7 Informatics' Support](#).

After provisioning, you'll have a running L7|ESP instance that you can develop content for locally. For more information on configuring your application and creating content definitions, see the [Development](#) section of the documentation. For an in-depth example of creating content using the L7|ESP SDK, see the [Tutorial](#) section.

You may also check the container logs to see if there are any errors:

```
$ docker compose logs --follow
```

Note: Press `Ctrl+C` to exit the container logs.

Shell

After starting the project, you'll need to enter the L7|ESP SDK container's shell to do development and testing. All necessary tools are installed inside the container to simplify the process of getting started with the L7|ESP SDK.

To get a shell into the L7|ESP SDK container image, you can do the following:

```
$ # get a shell into the container
$ docker compose exec server bash
```

Note: At this point, you should have a Bash prompt inside the container, instead of your host machine.

Installation

Now, to install L7|ESP SDK project inside the L7|ESP SDK container running on your machine, run:

```
$ make install
```

Virtual Environments

To perform certain tasks, you may need to switch Python virtual environments. For example, when doing pipeline script development, it is often useful to enable the Python virtual environment containing the L7|ESP Python client:

```
$ source ~/data/extensions/client/bin/activate
```

Note: Once you're finished using a virtualenv, you can deactivate the environment by running the `deactivate` command.

For more information on the `virtualenv` Python program, see the [documentation](#) section.

1.2 Project Structure

The sections below detail an overview of the structure and usage patterns of the L7|ESP SDK. For information about the L7|ESP Python client, please see the [documentation](#) section. For an in-depth example of creating content using the L7|ESP SDK, see the [Tutorial](#) section.

At a high level, the L7|ESP SDK was designed to speed up and simplify the process of building content for the L7|ESP platform. It provides a set of tools that enable rapid content development and testing, and also provides a contract by which content developers, internal and external to L7 Informatics, can communicate ideas about content.

1.2.1 Files and Directories

To start understanding and using the L7|ESP SDK, let's first go over the structure of the repository and important components:

- `Makefile` - A file used primarily for administration of L7|ESP. With the `Makefile`, you can start, stop, and reset the application. You can also use it to run testing commands.
- `docker-compose.yml` - A Docker Compose configuration file that creates a containerized development environment. This file will use Docker Compose for provisioning a development system to mirror the production environment as closely as possible.
- `roles/` - Custom Ansible roles for project deployments. General purpose roles will be defined in `app/esp-content/roles`, but roles specific to each project's deployment can be included here. Also, this folder contains development (`dev.yml`) and production (`prod.yml`) playbooks to manage development/deployment environments.
- `conf/` - All L7|ESP configuration files and data sources (`lab7.conf`).
- `content/` - All content related to this project installation, including all importable configuration files and scripts.
- `extensions/` - All extension points installed in the software. These points are reserved for custom expressions and custom API endpoints.

- `patches/` – Any patches that need applied to this project installation. This can patch the core L7|ESP codebase or any configuration files.
- `tests/` – All client-side and server-side integration tests specific to a project’s workflows and configuration.

1.2.2 Organizing Content and Tests

Now that we’ve gone over some of the components at a high level, let’s dive into some detail about the internal structure of some of these subdirectories. Here’s a high-level layout of how the L7|ESP SDK directory structure might look for an L7|ESP SDK with a couple of L7|ESP *Workflows* and *Protocols* already defined:

```

sdk/
├── README.md
├── Makefile
├── docker-compose.yml
├── requirements.txt
├── pyproject.toml
├── docs/
├── conf/
│   └── lab7.conf
├── content/
│   ├── admin/
│   │   └── Users.yml
│   ├── workflows/
│   │   └── QC-Workflow.yml
│   ├── protocols/
│   │   ├── QC-Quantification.yml
│   │   └── QC-Report.yml
│   ├── pipelines/
│   │   └── QC-Report-Pipeline.yml
│   ├── tasks/
│   │   └── generate_qc_report.py
│   └── inventory/
│       ├── Sample-Types.yml
│       └── Container-Types.yml
├── extensions/
│   ├── server/
│   │   ├── expressions.py
│   │   ├── invokables.py
│   │   └── requirements.txt
│   └── client/
│       └── expressions.js
├── patches/
│   ├── conf/
│   │   └── logs.patch
│   └── esp/
│       └── sso.patch
├── roles/
│   └── container.yml
└── tests/
    ├── __init__.py
    ├── conftest.py
    ├── resources/
    │   └── EXP001.txt
    ├── test_status.py
    └── test_content.py

```

We’ll go over details for specific files later in this documentation.

Note: How you organize your `tests` and `content` directories is up to you, but the remainder of the repository must use the above structure (the L7|ESP SDK assumes a specific file structure).

1.2.3 Notes on Specific Files

Let's go over some of the more important files in more detail. We've already covered the purpose of the `Makefile` and `docker-compose.yml` in the root directory of the repository, so let's continue with others.

Configuration

- `pyproject.toml` - A configuration file used for specifying high-level metadata about the Python project. In the `[tool.pytest.ini_options]` section, you can change the default options used by `pytest`. Usually, this file won't need to be changed, and will only need updates if you wish to adjust how certain Python tools function. Here's an example `pyproject.toml` file:

```
[tool.pytest.ini_options]
addopts = "--log-level=INFO -v -s -p no:warnings"
testpaths = [
    tests
]
```

- `conf/lab7.conf` - A file containing all L7|ESP configuration needed to run the application. For more information on this file, please see documentation for the L7|ESP platform.

Ansible Roles

- `roles/container.yml` - An Ansible playbook used for configuring the container environment with common content used for seeding the application in both development and production environments. Here's an example container playbook:

```
---
- hosts: localhost
  connection: local
  vars_prompt:
    - name: "l7esp_password"
      prompt: "Password for L7|ESP superuser: `admin@localhost`?"
      default: password
      private: yes
  roles:
    - role: l7esp_sdk
      vars:
        env: dev # This will allow database access from outside.
        db_archive: true # Whether to archive the database during installs.
      seed:
        - 'roles/seed/content.yaml' # List of seed files (from `roles/packages.
        ↪yaml`)
      tasks: [] # Any additional pre/post tasks you may have (optional)
```

Information about the `seed` section of this config is included in the *Content* section later in the documentation.

Testing

- `tests/conftest.py`, `tests/__init__.py` - Configuration files used for testing content within the L7|ESP SDK. These files generally won't need to be altered throughout L7|ESP SDK use.

1.3 Administration

1.3.1 Docker Compose + Ansible

The L7|ESP SDK relies heavily on Docker Compose for managing the container associated with running the application. For context, see the *Getting Started* section of the documentation.

1.3.2 Running the Application

As was alluded to in the documentation, you can start the application at any time using:

```
~$ docker-compose up --detach
```

For development, all content, configuration, and L7|ESP-related packages are mounted into the container. Therefore, whatever changes you make to the repo outside of the container will be reflected in the running L7|ESP instance.

The first time you run `docker-compose up --detach`, Docker Compose will try to pull the container image and subsequently create a container with all of the roles you specify in your container playbook (`roles/container.yml`).

After the first `docker-compose up --detach`, Docker Compose will only try to start the container, and won't try to pull the image again. If you need to re-pull the container image for some reason, you can use:

```
~$ docker-compose pull
~$ docker-compose up --detach
```

These commands will re-pull the L7|ESP SDK container image, and recreate the container if required.

After you've done all of your development and you want to stop the container, you can run:

```
~$ docker-compose down
```

This command will stop all containers without removing any data volumes. If you've halted the container, you can bring the application back up using:

```
~$ docker-compose up --detach
```

If you want to completely blow away the container **and** data volumes so you can start from scratch, you can run:

```
~$ docker-compose down --volumes
~$ docker-compose up --detach
```

Project repositories are designed to quickly get developers up and running, so there's generally no negative consequences for tearing-down a container (don't be afraid!).

Note: After the application has been started and provisioned initially, other make targets, like `reset`, `reload`, and `clear`, can be used for managing updates throughout L7|ESP SDK development.

1.3.3 Resetting the Database

There are currently two ways to remove existing content in the application. The first (and fastest) of these two methods is:

```
make clear
```

When you run `make clear`, the L7|ESP SDK will use the L7|ESP Python client to `drop` all existing models in the application. This drop won't completely wipe the database, and all internal history for each of the models will still be in the system. `make clear` is primarily used for quickly clearing the system between tests. For a more comprehensive reset of everything in the database, you can use:

```
make reset
```

This command will take down the application, completely remove and re-migrate the database, and then bring the application back up. Just like with `make clear`, all content will be removed from the application. Although this way of resetting the database is slower than `make clear`, it's also more comprehensive in resetting the application to its default state.

1.3.4 Reloading the Application

If the application goes down or you need to restart it for some reason, you can reload it using:

```
~$ make reload
```

This command will stop and restart all services related to the L7|ESP application. If the application is not working for some reason, this command should be attempted before contacting technical support.

1.3.5 Importing Content

After the database resets, all of the initial content configured in the repository will be removed. To re-import content, you can use:

```
~$ make import
```

This command will reload any content defined as part of your development configuration in *roles/packages.yml* or *roles/dev.yml*.

1.4 Introduction

There are several types of development you can do with the L7|ESP SDK:

Content Developing and seeding configuration for content (i.e. *Workflows*, *Protocols*, *SampleTypes*, etc.).

Extensions Developing custom expressions and custom endpoints to use within content.

Testing Testing your workflows and protocols to create automated test suites.

Windows Setup Windows development environment setup instructions using WSL2

1.5 Content

While developing with the L7IESP SDK, all content is defined within the `content/` folder of the L7IESP SDK. That folder can be organized however you like, but a directory structure similar to the following is recommended:

```
content/
├── admin/
│   └── Users.yml
├── workflows/
│   └── QC-Workflow.yml
├── protocols/
│   ├── QC-Quantification.yml
│   └── QC-Report.yml
├── pipelines/
│   └── QC-Report-Pipeline.yml
├── tasks/
│   └── generate_qc_report.py
└── inventory/
    ├── Sample-Types.yml
    └── Container-Types.yml
```

In the example above, note that directories are used to separate different types of content files. Generally, most of the content added to the `content` directory will be in a YAML format that the L7IESP Python client can import into the application. For more information about the various types of YAML config files that can be imported into the application, please see the L7IESP Python client documentation.

1.5.1 Referencing Resources

In addition to configuration files, resources referenced in the application (i.e. scripts or artifacts used during pipeline execution) should be referenced using a special path identifier that works in both development and production. Below are useful paths for referencing custom or stock content:

- `$LAB7DATA` - The path to the application data folder, which contains the application database, pipeline artifacts, and all custom content defined in the repository.
- `$LAB7DATA/content` - Path to the `content` directory from the L7IESP SDK. Any files you create in the `content` directory will be automatically mapped to this production directory on install/update.
- `$LAB7DATA/common` - Path to common (not project-specific) content resources. Any stock content not part of a specific project repository may be referenced via this path.

In addition to these special paths, there are several default ParamGroups you can use when running *Pipelines*:

- `param('apps', 'python')` - A reference to the version of Python installed with the application. This version of Python contains any requirements you specify in the `requirements.txt` file within the L7IESP SDK, the L7IESP Integrations Python module, and also the L7IESP Python client. It's generally recommended to use this param group instead of using *python* before calling scripts.
- `param('apps', 'integration')` - A reference to the L7IESP Integrations Python entrypoint, which is installed alongside the application for instrument integrations and other types of support.
- `param('apps', 'client')` - A reference to the L7IESP Python client entrypoint, which might be useful for some types of activities (i.e. ingests).

For example, if we have a script in our content directory that references another file in a separate directory, we can define our *Pipeline* to reference those two files, as follows:

```
# content folder with script and resource
content/
├── pipelines/
│   └── my-custom-pipeline.yml
└── tasks/
    ├── my_custom_pipeline.py
    └── pipeline_config.json
```

```
# contents of content/pipeline/my-custom-pipeline.yml
```

My Pipeline:

```
tasks:
  - My Pipeline Script:
      cmd: "param('apps', 'python') $LAB7DATA/content/tasks/my_custom_pipeline.py -
↳c $LAB7DATA/content/tasks/pipeline_config.json"

  - My Integration Script:
      cmd: "param('apps', 'integration') instrumentsupport illumina checkindexes --
↳worksheet '{{sample_sheet_uuid}}'"

  - My Ingest Script:
      cmd: "param('apps', 'client') ingest my_ingest {{infile}}"
```

1.5.2 Loading Content

After developing and testing all content that needs to be in your production instance, you need to set up configuration to specify which content is loaded into L7|ESP by default. There are two components to this:

1. Create **seed** files that define which configuration to import and what types of objects they represent in the system. A **seed** file is a standard format that is used to import configuration via the L7|ESP Python client. For information about the structure of these files, see the L7|ESP Python client documentation.
2. Update your container playbook to import those **seed** files (`roles/container.yml`).

For a simple example, let's consider a `content/` folder with a *Workflow*, two *Protocols*, and a *Pipeline* that we want to go into production. We can define a **seed** file at `roles/qc.yml`, like so (you can name the seed file however you like):

```
# Contents of: roles/qc.yml

- model: Workflow
  data: $LAB7DATA/content/workflows/QC-Workflow.yml

- model: Protocol
  data: $LAB7DATA/content/protocols/QC-Quantification.yml

- model: Protocol
  data: $LAB7DATA/content/protocols/QC-Report.yml

- model: Pipeline
  data: $LAB7DATA/content/pipelines/QC-Report-Pipeline.yml
```

In the example above, `$LAB7DATA` is an environment variable referencing the L7|ESP SDK root directory. It can be used both in development and production deployments to reference the same location.

Once we've created our **seed** file, we can update our container playbook, `roles/container.yml`, to reference that seed file for the content that L7|ESP will load into the system by default:

```
---
seed:
  - '{{ sdk }}/roles/qc.yml'
```

In the file above, `{{ sdk }}` is an Ansible variable referencing the L7|ESP SDK root directory. It can be used both in development and production deployments to reference the same location.

Once you've completed all of this configuration, you can test it out using:

```
~$ make import
```

This command will use Ansible to import all of the content, which is the exact process that is used during install/update in production to import content.

1.5.3 Biobuilds

BioBuilds is a curated collection of open-source bioinformatics tools, pre-built for Linux on both Intel x86_64 and IBM POWER8 systems as well as Mac OS X.

Here we add *samtools* to the list of Biobuilds tools that will be included on install:

```
# contents of roles/container.yml
---
biobuilds:
  - bwa
  - samtools
```

1.6 Extensions

The `extensions` folder in the L7|ESP SDK allows developers to create custom expressions and endpoints in the application. Here's an example of the structure of that directory:

```
├── extensions/
│   ├── server/
│   │   ├── expressions.py
│   │   ├── invokables.py
│   │   └── requirements.txt
│   └── client/
│       └── expressions.js
```

L7|ESP offers a number of extension mechanisms for cases when the out-of-the-box functionality is insufficient to meet a particular need, both for server-side and client-side functionality. On the client-side, the `invokables.js`, `renderers.js`, and `fas-overrides.js` files provide entry to augment and customize the user experience. On the server-side are custom **Expressions**, **Invokables**, **Transition Strategies**, **Data Source Loaders**, and **API Callbacks**.

1.6.1 Server Side Extensions

L7|ESP provides a number of server-side extension points detailed below, in order from “most-used” to “least-used”.

Custom Expression Functions

Custom server-side expressions available in L7|ESP **Worksheets** can be created in the `extensions/server/expressions.py` file.

Users may contribute new **Expression** functions to L7|ESP. **Expressions** must be callable or subclasses of `lab7.expression.APICall`. They are contributed using the `lab7.extensions.expression` decorator.

@expression Arguments:

- `name` - the name of the **Expression**. If not specified, the callable object is examined for `name` or `__name__` attributes, in order. If found, the value is used for the name, otherwise an error is raised.
- `context` - the **Expression** evaluation context to add to. Currently, only `Lab7ExpressionContext.ALL` is supported.

Expressions may accept any arguments and return any value, although in most contexts, the return value will ultimately be coerced to a string. For example, contributing a quantile function might look like:

```
from esp.extensions import expression

@expression
def quantile(array, quant):
    import numpy as np
    return np.percentile([float(x) for x in array], float(quant*100))
```

After registering the extension with L7|ESP (see “Registering server-side-extensions”), the `quantile` function will be available in all contexts.

Here are some additional examples of using the `@expression` decorator:

```
@expression
def myfunc(a, a):
    return int(a) + int(b)
```

Now `myfunc` is accessible from, e.g., LIMS default values:

```
{{ myfunc(1, 2) }}
```

To override the name by which the **Expression** is registered, supply the `name` argument to the decorator.

```
@expression(name='add2')
def myfunc(a, a):
    return int(a) + int(b)
```

Now the expression is:

```
{{ add2(1, 2) }}
```

By default, **Expressions** are set to available in every evaluation context. To limit to a specific context, supply the `context` argument.

```

from lab7.extensions.expressions import Lab7ExpressionContext

@expression(name='add2', context=Lab7ExpressionContext.ALL)
def myfunc(a, a):
    return int(a) + int(b)

```

The **Expression** is available in every context. If no context is provided, the context is set to `Lab7ExpressionContext.ALL`.

Note: Currently, the only context used in L7|ESP is `Lab7ExpressionContext.ALL`, but support for context-specific functions will be added in a future release of L7|ESP.

In advanced **Expression** scenarios, users may need access to the database. In these cases, **Expressions** should be objects that extend from `lab7.expressions.APICall`. Objects that extend from `APICall` have access to several useful attributes, including

- `self.ctx` - the **Expression** evaluation context. See “Using Expressions in L7|ESP” for information on what is available in various **Expression** evaluation contexts.
- `self.session` - the SQL Alchemy database session
- `self.agent` - The active L7|ESP user
- `self.cache` - A dictionary shared across **Expressions** and **Expression** evaluations for a single evaluation event. For instance, when evaluating a LIMS worksheet, the cache dictionary is created at the beginning of evaluation and retained until the entire worksheet is finished. This allows **Expression** authors to lookup a value 1x/in-bulk, and cache the results. For database-backed **Expressions**, it is highly encouraged to use the cache to help ensure sample sheet saving stays fast. (In one example at L7, a sample sheet save with 100 samples went from 30+ seconds down to 3 seconds by adding caching to one **Expression**).

The following example creates a “service_types” **Expression** function that looks up the list of service types from the DB, but only if the list has not already been fetched before in this evaluation context.

```

@expression(name="service_types")
class ServiceTypes(APICall):
    """Return a list of service types

    Args:
        names_only (bool): If true (the default), the returned list is only the_
        ↪service type names. If false, the
        ↪returned list is the full service type dictionary.
    """

    def api_func(self, names_only=True):
        key = '__service_type_list__{}'.format(names_only)
        if key not in self.cache:
            service_types = query_service_types(params=['name'], agent=self.agent,
            ↪session=self.session)
            if names_only:
                self.cache[key] = [x['name'] for x in service_types]
            else:
                self.cache[key] = service_types
        return self.cache[key]

```

Custom Queries

Custom queries can be contributed to the backend by defining the query in a yaml file and placing the yaml file in `/opt/l7esp/data/content/queries`. Queries added this way are available through API call from `/api/v2/queries/{query_file_name}`. The YAML file MUST have the suffix `.yaml` and the filename must not have any spaces. The structure of the YAML file is as follows:

```
name: {query name, spaces ok}
description: |+
    QUERY DESCRIPTION
query: SQL
parameters:
  param1:
    description: param1 description
    type: list # list of values, such as what might be used in an IN clause.
  singleton_param:
    description: param 2 description. Param 2 is a single value.
```

For instance, a simple query to fetch project names might be:

```
name: Project Name List
description: |+
    Fetch the name for all projects in L7|ESP, even ``archived`` projects.
query: |+
    SELECT name from resource where cls=540
parameters:
```

If the file was placed in `/opt/l7esp/data/content/queries/project_name_list.yaml`, it would then be possible to call `/api/v2/queries/project_name_list`. The API call returns a JSON object with various keys, including name (the value of name in the yaml file) and one of either results or error depending on if the query ran successfully. If not, error will contain the error message. If successful, results will contain the results as a list of JSON objects. For instance, `/api/v2/queries/project_name_list` from above might return:

```
{
  "name": "Project Name List",
  "results": [
    {"name": "project 1"},
    {"name": "project 2"}
  ]
}
```

Custom Endpoints

L7|ESP allows users to contribute new backend endpoints to the application without modifying the core server. These endpoints are called Invokables and are contributed using the `lab7.extensions.invokable` decorator.

Decorator arguments:

- Name - Name of the **Invokable**. This will appear in the API as `/api/invoke/<name>`. If name is not supplied, the callable will be examined in turn for a “name” or “__name__” attribute and use the attribute value if found. If no attribute is found, a value error will be raised.
- Session - The value of session is the name of the callable argument that should receive the SQLAlchemy session. If no session is required, the argument may be omitted.
- Files - The value of files is the name of the callable argument that should receive the uploaded files data structure. (For details on the value, see <https://www.tornadoweb.org/en/stable/guide/structure.html>).

The decorated **Invokable** objects must be callable. The first argument of an **Invokable** MUST be “agent”, which receives the authenticated user making the request. Note that **Invokable** endpoints may only be requested by authenticated users. Aside from session and files argument (see arguments above and examples below), all other arguments are specific to the **Invokable** and supplied by requesting agents by posting `{"kwargs": {"arg1": "value1", ...}}`. The return value of the callable must be JSON-serializable, the request should be made with the HTTP header `Content-Type: application/json` and requestors should expect an returned content type of `application/json`.

Note: In a future release of L7|ESP, **Invokables** will be reworked so the `kwargs` key may be omitted so JSON object keys are mapped directly to function arguments. i.e. `{"kwargs": {"arg1": "value1", ...}}` would simplify to just `{"arg1": "value1", ...}`. Support for `kwargs` will be maintained, however, for backwards compatibility.

If you need an **Invokable** for non-JSON returns, contact your L7 field application scientist.

Custom invokables (endpoints) available through the application can also be created in the `extensions/server/invokables.py` file. Here’s an example of creating a custom endpoint to return an ‘ok’ status on GET request:

```
# contents of extensions/server/expressions.py

# endpoint definitions
class Ping(object):
    """
    Ping server and return OK response.
    """

    def __call__(self, agent, *args, **kwargs):
        return {'status': 'ok'}

# export (boilerplate)
INVOKABLES = {
    'ping': Ping,
}
```

Once this custom invokable has been defined, you can access the `/api/invoke/ping` URL via authenticated request:

```
>>> from esp import base
>>> base.SESSION.get('/api/invoke/ping')
{'status': 'ok'}
```

More Examples:

```
from lab7.extensions import invokable
@invokable
def custom_api(agent, a, b):
    return {'result': int(a)+int(b)}
```

The above function can now be “invoked” by issuing a request like `POST /api/invoke/custom_api`

With a content body of `{"kwargs": {"a": 1, "b": 2}}`

And will return an `application/json` response of `{ 'result': 3 }`

Supply the name argument to the **Invokable** decorator to override the name:

```
@invokable(name='add2')
def custom_api(agent, a, b):
    return {'result': int(a)+int(b)}
```

Now the API call is to POST /api/invoke/add2

The first argument to **Invokable** must be `agent`. The **Invokable** machinery supplies the `agent` as the resolved **User** object corresponding to the authenticated user who issued the request.

Supply the `session` and `files` argument to the decorator to signal that your **Invokable** should receive the request session and uploaded files, respectively.

```
@invokable(session='sessionarg', files='filesarg')
def myapi(agent, a, b, sessionarg, filesarg):
    pass
```

Now the multipart POST API call to /api/invoke/myapi will pass the arguments `a` and `b` from the `kwargs`, the `agent` as the standard first argument, the request session to the `sessionarg` argument, and the request files (Tornado's data structure) to the `filesarg` argument of `myapi`.

Custom Transition Strategies

Users may contribute their own **Transition Strategies** to L7|ESP. **Transition strategies** are used when routing samples from one workflow to the the next. A custom **Transition Strategy** might be used to send a notification when samples route to a particular workflow; or to adjust the routing so the sample going into the next workflow is a relative of the sample ending the previous workflow (for instance, routing a tissue sample for processing when a fluid sample fails).

Transition strategies are contributed using the `lab7.extensions.transition_strategy` decorator.

Arguments:

- Name - Name of the strategy as used in, e.g., **Workflow Chain** configurations. If not specified, the strategy object is examined for `name` and `__name__` attributes in-turn. If found, the value is used, otherwise, an error is raised.
- Virtual - if `True`, the transition is “virtual”. The default is `False`. Currently, the only difference between virtual and non-virtual transitions in L7|ESP is how they are rendered in **Workflow Chain** views. Virtual transitions render as dashed orange curves. Non-virtual transitions are blue and, typically, straight. Non-virtual transitions are used to inform the initial graph layout; virtual transitions are rendered after the graph structure is defined.

Strategies must be callable and must accept the following two parameters:

1. `sample_uuids` - The list of `sample_uuids` being transitioned. The list only contains UUIDs that pass the transition rule.
2. `context` - an instance of `WFCTransitionContext` which simplifies passing various useful data to the **Transition Strategies**, such as the incoming workflow instance (`from_experiment` property), the outgoing workflow instance (`to_experiment` property), the active transition, the active SQLAlchemy session, etc.

Strategies must return a list of `sample_uuids` that will be put into the downstream `to_experiment`.

The following example **Transition Strategy** transitions the parents of the samples in the final workflow sample set to the next workflow. An example use-case: sequencing libraries are pooled into a “library pool”. The library pool is sequenced. After sequencing, the libraries need to be pushed to bioinformatics, but the sequencing workflow has the library pool. You could use `submit_parents` to move the pool parents, the libraries, into bioinformatics if the sequencing passes QC.

```
@transition_strategy
def submit_parents(sample_uuids, context):
```

(continues on next page)

(continued from previous page)

```

import lab7.sample.api as sapi
parents = []
found = set()
for uuid in sample_uuids:
    deps = sapi.get_sample_dependencies(
        uuid, parents=True, children=False, uuids_only=True,
        agent=context.agent, session=context.session)
    for parent in deps['parents'][::-1]:
        if parent not in found:
            found.add(parent)
            parents.append(parent)
return parents

```

Custom Protocol Actions

You can write your own protocol actions like other server side extensions. You will need to use the `@protocol_action` decorator, provide a name, and any params that it needs. For example:

```

@protocol_action(
    name="Sync Item Field",
    desc="Copies a LIMS value into an inventory item field."
    parameters=[
        ExtensionParameter(
            name="prot_field", type="string", label="Protocol Source Field",
            help="The ID of the protocol field to copy from"
        ),
        ExtensionParameter(
            name="item_field", type="string", label="Item Target Field",
            help="The ID of the protocol field to copy from"
        )
    ]
)
def sync_field_value_with_items(rows, context):
    """
    Copies a LIMS value into an inventory item field. The target field must be a
    ↪field on the item entity (eg: initial_qty, lot_id, status, expiration_timestamp)
    The user will need to specify both the source (the protocol column name) and
    ↪target item field.
    """
    from lab7.resource.models import ResourceActionType
    source_id = context.action["parameters"]["prot_field"]
    target_id = context.action["parameters"]["item_field"]

    msg_template = (
        "Updated item field `{target_id}` from `{}` to `{}` " +
        "via protocol action `{context.action_name}` " +
        "of protocol `{context.protocol.name}` " +
        "in sheet `{context.sheet.name}` ({context.sheet.uuid}) "
    )

    actions = []

    for row in rows:
        try:
            old_value = getattr(row.sample, target_id)

```

(continues on next page)

(continued from previous page)

```
new_value = row.get_value(source_id)

setattr(row.sample, target_id, new_value)
actions.append((
    msg_template.format(old_value, new_value),
    ResourceActionType.property_change(target_id, old_value, new_value)
))
except ValueError:
    pass

return actions
```

In the above example, we define a function and apply the `@protocol_action` decorator such that the function will be called when a protocol executes the ‘Sync Item Field’ action.

The `@protocol_action` decorator accepts the following parameters:

- `name` - The name of the action.
- `desc` - A description of the action.
- `parameters` - A list of `ExtensionParameter` objects.

Since our example copies a value from a protocol field to an item field, we pass in two `ExtensionParameter` objects - one for the protocol field name (`source`) and another for the item field name (`target`).

An `ExtensionParameter` object accepts the following parameters:

- `name` - The name of the parameter.
- `type_` - The type of the parameter (one of: `string`, `sample_type`, `sample_id`).
- `label` - The label for the parameter.
- `default` - A default value for the parameter (optional).
- `required` (`bool`) - Whether the parameter is required or not (default: `True`).
- `help` - A help text for the parameter.

The function will receive a list of rows (`rows`) and a context object.

The context object contains the following properties:

- `action` - The action object, containing a `parameters` dict where the keys are the parameter names and the values are the parameter values.
- `action_name` - The name of the action.
- `action_def` - The action definition object.
- `agent` - The agent making the request.
- `condition` - The condition object.
- `executed_actions` - A list of executed actions.
- `final_protocol_state` - The final protocol state.
- `final_row_state` - The final row state.
- `initial_protocol_state` - The initial protocol state.
- `initial_row_state` - The initial row state.
- `newly_approved_rows` - A list of newly approved rows.

- `newly_completed_rows` - A list of newly completed rows.
- `newly_failed_rows` - A list of newly failed rows.
- `newly_unapproved_rows` - A list of newly unapproved rows.
- `newly_uncompleted_rows` - A list of newly uncompleted rows.
- `newly_unfailed_rows` - A list of newly unfailed rows.
- `protocol` - The protocol object.
- `protocol_ended_active` - The protocol ended active.
- `protocol_started_active` - The protocol started active.
- `session` - The current database session
- `sheet` - The sheet object.
- `sheet_env` - The sheet environment object.
- `tab_idx` - The tab index.

API Callbacks

API Callbacks are a powerful mechanism for augmenting the behavior of any existing L7|ESP endpoint.

Note: **API Callbacks** provide a lot of power, but can also easily contribute performance problems into a running L7|ESP instance if care is not taken. Use with caution. At L7, **API Callbacks** are used in only a handful of implementations where customers asked for functionality and the following criteria were met:

- a. There was no other mechanism to accommodate the request in a manner that would meet the user needs
 - b. The request was specific enough to one customer that adding it as a core feature would not benefit any other customers and would add additional complexity to the codebase and UI.
-

API Callbacks are contributing using the `lab7.extensions.api_callback` decorator

Callback arguments:

- **callback (callable function):** The callable object to register. The call must accept four arguments.
 - `agent (User)`: The resolved esp User object for the requesting user
 - **request_details (RequestDetails):** A RequestDetails object with information about the request including
 - * **endpoint (string):** The endpoint being called (e.g. `/api/samples`) For GraphQL endpoints, the `api_callback` machinery automatically adds the `operationName` to the endpoint, so the endpoint will be `/graphql/<operationName>` For instance: `/graphql/AddItemToContainer`. Note also that GraphQL endpoints are always POSTed to, even for data retrieval.
 - * `when (string)`: Either "before" or "after"
 - * `method (string)`: The request method (POST, PUT, GET, or DELETE)
 - * **status (string|int):** Request status. If when is "before", it will be an empty string. If when is "after" it is the integer status code of the request.
 - `data (dict)`: All request parameters, including query and body parameters.
 - `session (Session)`: Database session, which is optional and may be None

- `name (string)`: The name of the callback. If no name is provided, the Callable is checked in order for a `__name__` or `name` property and, if found, this property is used. Otherwise, an error is raised.
- `endpoints (string|list[string])`: An API endpoint this callback is valid for. The string or list of strings are treated as exact matches except that `'*'` is a valid wildcard.
- `methods (string|list[string])`: The HTTP method the callback is valid for. The special string `'*'` (default) indicates the callback is valid for all methods. A list of methods may be supplied such as `method=['PUT', 'POST']`.
- `when (string|list[string])`: When the callback is valid. May be `'before'`, `'after'`, the list `['before', 'after']`, or `'*'` (shortcut for `['before', 'after']`)
- `stop_on_failure (Boolean)`: If we should stop processing and rollback on a failure

Examples:

```
from lab7.extensions import api_callback

@api_callback
def global_callback(agent, session, request_details, data):
    with open('mylog', 'a') as log:
        log.write("Called {} ({})\n".format(request_details.endpoint, request_details.
        ↳when))

@api_callback(name='specific_callback', endpoints='/api/samples', methods='POST',
        ↳when='before'):
def specific_callback(agent, session, request_details, data):
    if 'name' not in data:
        raise ValueError('This ESP installation does not support auto-sample-name-
        ↳generation')
```

The first example above would append to a file `mylog` before and after every API call in L7|ESP.

The second example is only called POSTing to the `/api/samples` endpoint and is called just before the standard L7|ESP handler. In this example, the callback is used to add a customer-specific constraint that sample POST always supplies a name. This would have the effect that samples could **only** be added by API call since the UI for sample creation relies on samples having names.

Note that callback registration is a two-step process. The first step is to declare and register the callback for invocation, as above. The second step is to add the name of the callback to the system configuration. This allows extension modules to make callbacks available for use while still requiring explicit configuration to enable them for a given installation. This is accomplished by adding an `api_callbacks` section to the system configuration:

```
api_callbacks:
  before:
    - global_callback
    - specific_callback
  after:
    - global_callback
```

Note that it is an error to add a callback to a section that is not supported by the callback registration. For instance, adding `specific_callback` to the `after` block is an error since `specific_callback` declares `when` as `"before"`. This check ensures callbacks are used in the manner intended by callback authors.

Data Source Loaders

One of L7|ESP's configuration extensions is **Data Source Loaders**. **Data Source Loaders** offer a mechanism for flexibly loading relatively static data into the server memory space. Examples include assay configurations, Illumina index ID to index sequencing mappings, and more. Although many uses of **Data Source Loaders** have been supplanted by the newer database-backed dynamic configuration, they are still useful for any data source that can not be mapped to JSON, or which would be undesirable to map to JSON. For instance, ontology files in `*.obo` format, which can be read by such programs as `obonet` into a `networkx` graph. Out of the box, **Data Source Loaders** support YAML, JSON, CSV, and TSV file formats, but users can contribute their own **Data Source Loader** to augment, or even replace the out-of-box functionality. For instance, if a YAML or JSON file should be parsed into a specific Python object instead of generic `dict` and `list`. **Data Source Loader** logic follows the “chain of command” pattern, wherein a list of **Data Source Loaders** is traversed and the first **Data Source Loader** that can properly parse the data is used. Built-in **Data Source Loaders** can therefore be overridden just by inserting a custom **Data Source Loader** ahead of the built-in in the chain. **Data Source Loaders** are contributed using the `lab7.extensions.datasource_loader` decorator.

`datasource_loader` arguments:

- `where(callable)` - “where” in the chain of command to place this **Data Source Loader**. This is a callable that receives the **Data Source Loader** and positions it in the list. `lab7.extensions.datasource_loaders` provides the following out-of-box locators:
 - `beginning` - Position the **Data Source Loader** as the first **Data Source Loader** in the list. If two or more **Data Source Loaders** all specify `beginning`, they will be inserted in the beginning in LIFO order.
 - `end` - Position the **Data Source Loader** as the last **Data Source Loader** in the list. If two or more **Data Source Loaders** all specify `end`, they will be appended to the end in LIFO order.
 - `after(prior)` - Position the **Data Source Loader** after **Data Source Loader** given by `prior` where `prior` is the class of the other **Data Source Loader**.

The **Data Source Loader** is expected to be an object conforming to the `DataSourceLoader` protocol. The `DataSourceLoader` protocol consists of two methods:

- `can_load(datasource)` - returns `True` if the **Data Source Loader** is capable of handling `datasource`, where `datasource` is the configuration for the data source, including the `url` to load and, optionally, the data type.
- `load(name, datasource)` - returns the loaded object structure. `Name` is the name of the data source in the configuration.

Data source loading then works by searching the list for a **Data Source Loader** that can load the current data source and, when found, loading it. Custom **Data Source Loaders** are checked before out of the box **Data Source Loaders**, allowing for overriding default loading behavior for out-of-the-box formats.

Note that the `url` key of `datasource` is guaranteed to exist and to be an already stripped string value when calls to `can_load` or `load` are made.

Registering server-side extensions

Server side extensions are added to L7|ESP by means of Python entry points.

- Users new to Python entry points should review the official [Entry Points Specification](#).
- For a less formal example of using entry points, see [Python Entry Points Explained](#).

L7|ESP's extension entry point is `l7.esp.extensions`. To add your extensions to L7|ESP, create a Python package to hold the extensions (but see below for SDK users), make sure all modules with extensions are loaded when the package is loaded, and add the appropriate `entry_point` configuration to the `setup.py` file. For example, if the extensions are in an `acme` Python package with a structure:

```
acme
├── acme
│   ├── __init__.py
│   ├── api_callbacks.py
│   ├── dataloaders.py
│   ├── expressions.py
│   ├── invokables.py
│   └── transition_strategies.py
└── setup.py
```

Then the `__init__.py` file would have:

```
import acme.api_callbacks
import acme.dataloaders
import acme.expressions
import acme.invokables
import acme.transition_strategies
```

And the `setup.py` file entry point configuration would look like:

```
setup(
    name='acme',
    ...,
    entry_points={
        'l7.esp.extensions': [
            'acme=acme'
        ]
    })
```

Users taking advantage of the standard SDK layout and tooling for developing their L7|ESP instance can place extensions in the `extensions/server` directory. Files in this directory are auto-packaged into the `l7ext` python package at build time. Within the `extensions/server` directory, the auto-packaging ensures `expressions.py`, `invokables.py`, `transition_strategies.py`, `api_callbacks.py`, and `dataloaders.py` are auto-loaded when the `l7ext` package is loaded, so users can add to those files without worrying about anything else.

1.6.2 User-interface extensions

The user interface supports the following extensions: `invokables.js`, `renderers.js`, and `fas-overrides.js`, in descending order of usage likelihood.

`invokables.js`

`invokables.js` should be placed in `Lab7_ESP/current/lib/www/static/js`. It is loaded in LIMS, Reports, and Applet apps. Therefore, any functions, objects, or variables placed in this file are available to protocol event handlers (`onrender`, `onchange`) as well as to reports and applets. Other than loading symbols from this file, L7|ESP does not do anything special with this file.

renderers.js

renderers.js should be placed in Lab7_ESP/current/lib/www/static/js. It is loaded by the LIMS app. Unlike invocables.js, L7|ESP will examine renderers.js for specific symbols and process them if found.

manipulateGridOptions

manipulateGridOptions allows users to customize the behavior of LIMS grids. It is called with the argument gridOptions and should return a grid options object (usually by modifying the passed-in object, then returning it). The list of options is extensive, but the most common use is to register additional column rendering and/or editing components. For instance:

```
manipulateGridOptions: function(gridOpts) {
  // Any custom renderers MUST be added to gridOpts.components in order for transpose_
  // to work.
  // Other custom logic can be added, but default is no-op.
  var renderers = this.getRenderers()
  gridOpts.components = {
    ...gridOpts.components,
    ...renderers
  }
  return gridOpts
},
```

manipulateColumnDefs

manipulateColumnDefs receives a columnDefs argument and is expected to return the same. columnDefs is a list of column definition dictionaries. Using this function, users can alter the behavior of specific columns in a worksheet. For example - overriding the component used to render or edit a particular column value. For instance:

```
manipulateColumnDefs: function(columnDefs) {
  var renderers = this.getRenderers()
  console.log(columnDefs)
  columnDefs.forEach(x => {
    if (!x.meta) {
      return
    }
    let meta = null
    try {
      meta = JSON.parse(x.meta)
    } catch(err) {
      meta = x.meta
    }
    if (meta && meta.augment && meta.augment.renderer && renderers[meta.augment.
    //renderer]) {
      x.cellRendererSelector = function() {
        return {component: meta.augment.renderer}
      }
      x.transposeRendererName = meta.augment.renderer
    }
  })
  return columnDefs
}
```

The above block examines the `meta` property of each column definition to see if the path `meta.argument.renderer` matches a known custom renderers. If so, it uses the custom renderer in place of the standard renderer for that column.

Note: The functions in `renderers.js` are called for protocol for every workflow in the system. Operations in this file should therefore be lightweight to avoid performance penalties.

fas-overrides.js

`fas-overrides.js` is a file loaded in global scope on every page. As with `api_callback` on the server-side, this file should only be used in circumstances where other techniques fail to adequately address user-needs and the functionality required is so specific to a single use-case that others would not benefit from the functionality as a more general feature.

1.7 Testing

1.7.1 Running Tests

As you develop content within a project repository, it's generally good practice to do test-driven development. Test creation will be covered in later sections, but for running tests, you can use two different commands:

```
~$ make test
```

This command is the first option, or you can directly invoke `pytest` for running the tests.

```
~$ pytest
```

Since `pytest` is used for running the test suite, you can use it to run specific tests or test classes in your repository. For example:

```
~$ # A single test file
~$ pytest tests/test_qc_workflow.py

~$ # All tests in a class
~$ pytest tests/test_qc_workflow.py::TestQCWorkflow

~$ # A specific test method in a class
~$ pytest tests/test_qc_workflow.py::TestQCWorkflow::test_check_qc_values
```

In addition to running tests locally, you can also run tests on a different server. To specify a different L7|ESP instance on the command line, use the `--host` and `--port` options:

```
~$ pytest --host test.l7informatics.com --port 8005
```

Note: For running tests on a public URL, you may need to connect to the application via SSL. To do so, you'll need to use the `--ssl` option and set your port to 443. For example:

```
~$ # connecting to https://test.l7informatics.com
~$ pytest --host test.l7informatics.com --port 443 --ssl
```

Here are other options useful during test execution:

```

~$ pytest -h
usage: pytest [options] [file_or_dir] [file_or_dir] [...]

positional arguments:
  file_or_dir

...

custom options:
-N, --no-import          Skip IMPORT definitions when running tests.
-C, --teardown           Teardown content after running tests.
-P, --port               Port for accessing L7|ESP.
-S, --ssl                Use SSL when connecting to L7|ESP.
-H, --host               Host for accessing L7|ESP.
-U, --email              Email for admin user.
-X, --password           Password for admin user.

```

1.7.2 Creating Tests

There are generally two types of tests you can write to verify functionality as you develop content within L7|ESP:

- **Functional Tests** - Simple tests that can be run outside of the context of testing *Workflows* and related content. Examples of this type of testing include testing connections, integration points, server-side extensions, etc.
- **Content Tests** - Tests driven by configuration that test the functionality of *Workflows* and other types of content. These types of tests are generally written using classes and L7|ESP SDK tools for automating much of the content testing.

Functional Tests

As a gentle introduction, let's write a simple test to verify that L7|ESP is connected and accepting requests. We'll need to create a test file in the `tests` directory named `test_status.py`. The contents of that file should look something like:

```

# -*- coding: utf-8 -*-

import pytest

def test_status():
    import esp
    running = esp.status()
    assert running, 'Could not connect to L7|ESP via python client!'

```

After creating tests, you can run them using `pytest` (this example uses the test defined above):

```

~$ pytest tests/test_status.py b
===== test session starts.
↳ =====
platform darwin -- Python 3.7.1, pytest-3.10.0, py-1.7.0, pluggy-0.8.0 -- /usr/local/
↳ opt/python/bin/python3.7
cachedir: .pytest_cache
collected 1 item

tests/test_status.py::test_status Connection established!

```

(continues on next page)

(continued from previous page)

```
PASSED
```

```
===== 1 passed in 0.68 seconds_
↪=====
```

If L7|ESP was not running and the tests failed, you'd see the following:

```
~$ pytest tests/test_status.py b
===== test session starts_
↪=====
platform darwin -- Python 3.7.1, pytest-3.10.0, py-1.7.0, pluggy-0.8.0 -- /usr/local/
↪opt/python/bin/python3.7
cachedir: .pytest_cache
collected 1 item

tests/test_status.py::test_status FAILED

===== 1 failed in 0.68 seconds_
↪=====
```

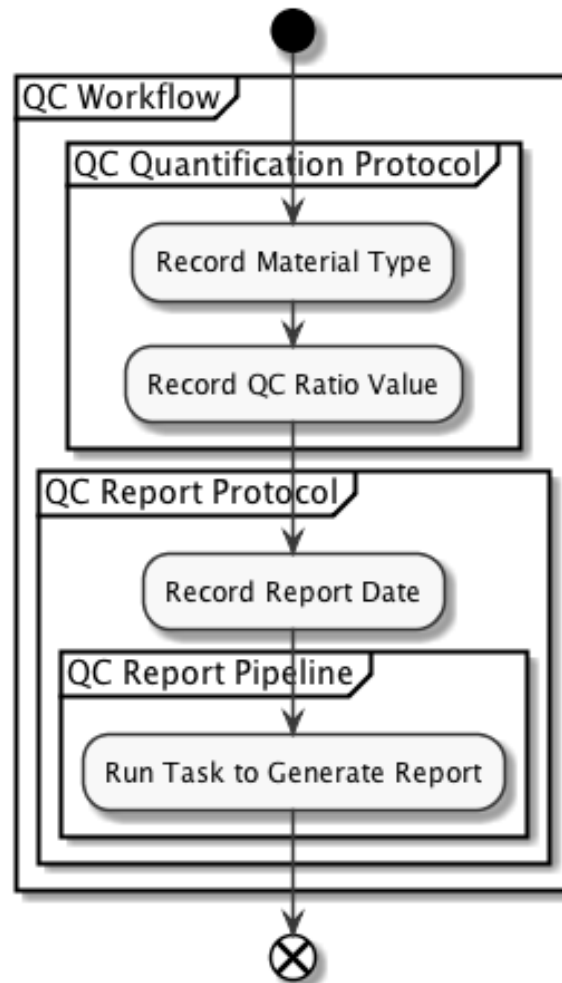
Now that we've created a test for some specific functionality in the application, let's talk about how to create tests for custom content defined in the L7|ESP SDK.

Content Tests

Along with functional testing, we can leverage L7|ESP SDK tools to help integration testing for content defined in the L7|ESP SDK. By now, you're familiar with the config-style format for defining content, so let's talk about defining tests to verify the functionality of a simple QC Workflow. For this example, our *Workflow* will have two parts:

1. A *Protocol* to capture QC metadata about a *Sample*.
2. A *Protocol* to generate a *Report* using a *Pipeline*.

Here's a schematic detailing the components of that *Workflow* at a high level:



Now that we know what type of content we’re going to create, let’s make some content files. First, we need to create our *Workflow*:

```
# Contents of: content/workflows/QC-Workflow.yml

QC Measurement Workflow:
desc: Workflow to measure QC values and generate a QC report.

tags:
- quickstart
- qc

protocols:
- QC Quantification:
  protocol: standard
  variables:
    - Type:
      rule: dropdown
      dropdown:
        - 'DNA'
        - 'RNA'
    - Ratio:
```

(continues on next page)

(continued from previous page)

```

        rule: numeric

- QC Report:
  protocol: pipeline
  pipeline: QC Report
  variables:
    - material:
      rule: text
      value: "{{ column_value('Type', 'QC Quantification') }}"
      visible: false
    - ratio:
      rule: numeric
      value: "{{ column_value('Ratio', 'QC Quantification') }}"
      visible: false
    - Report Date:
      rule: date

```

And we also need to create the *Pipeline* that our *Workflow* references:

```

# Contents of: content/pipelines/QC-Report.yml

QC Report:
  report:
    name: Result Report
    elements:
      - - depends:
          - {file: Result Report, tasknumber: 1}
          type: raw_file
      - []
    tasks:
      - Generate QC Report:
          desc: Task to generate QC report from workflow metadata
          cmd: |+
              # Simple Task that determines if the specified 'ratio' is in the proper_
↪range

              RNA_MIN=1.8
              RNA_MAX=2.1

              DNA_MIN=1.7
              DNA_MAX=2.0

              TYPE="{{ material }}"
              RATIO="{{ ratio }}"

              if [ $TYPE = "RNA" ]; then
                  PASS=`echo "$RNA_MIN <= $RATIO && $RATIO <= $RNA_MAX" | bc`;
              elif [ $TYPE = "DNA" ]; then
                  PASS=`echo "$DNA_MIN <= $RATIO && $RATIO <= $DNA_MAX" | bc`;
              fi

              echo $PASS

              if [[ $PASS = 1 ]]; then
                  echo "<b>Your sample <font color='green'>Passed</font> and contains pure
↪$TYPE</b>" >> result.html
              else

```

(continues on next page)

(continued from previous page)

```

        echo "<b>Your sample <font color='red'>Failed</font> and is NOT pure
↪$TYPE</b>" >> result.html
    fi
    files:
    - Result Report:
      file_type: html
      filename_template: "{{ 'result.html' }}"

```

Now that we have our content defined, we can use L7|ESP SDK tools to write a simple test that will import these files and verify that they exist in the system. This test also enables us to verify that there are no syntax errors in our *Workflow* definitions. Our test file for doing so can look like the following, `tests/test_qc_workflow.py`:

```

# -*- coding: utf-8 -*-

# imports
import os
import unittest

from . import CONFIG, RESOURCES, CONTENT
from esp.testing import ModelLoaderMixin

# tests
class TestQCWorkflows(ModelLoaderMixin, unittest.TestCase):
    IMPORT = dict(
        Workflow=[
            os.path.join(CONTENT, 'workflows', 'QC-Workflow.yml'),
        ],
        Pipeline=[
            os.path.join(CONTENT, 'pipelines', 'QC-Report-Pipeline.yml'),
        ]
    )

```

Running this test will result in the following output:

```

~$ pytest tests/test_qc_workflow.py

===== test session starts _
↪=====
platform darwin -- Python 3.7.1, pytest-4.1.0, py-1.7.0, pluggy-0.8.0 -- /usr/local/
↪opt/python/bin/python3.7
cachedir: .pytest_cache
collected 1 item

tests/test_qc_workflow.py::TestQCWorkflows::test__content
INFO:root:Clearing existing content from database.
INFO:root:Creating Task: Generate QC Report
INFO:root:Creating Pipeline: QC Report Pipeline
INFO:root:Creating PipelineReport: QC Report
INFO:root:Creating Protocol: QC Quantification
INFO:root:Creating Protocol: QC Report
INFO:root:Creating Workflow: QC Workflow
Successfully imported config data.
PASSED

===== 1 passed in 3.47 seconds _
↪=====

```

Along with simply testing *Workflow* definitions, we can also test that the *Workflow* works properly when used in an *Experiment*. To do this test, we can use the same format as above, but add a `DATA` attribute for what *Projects/Experiments* need to be created. For this example, we'll need to create a *Project* and two *Experiments* for testing out the *Workflow*. Here are config files for our two *Experiments*:

```
# Contents of tests/resources/QC-Test-1.yml

My QC Experiment 1:
  submit: True
  project: QC Project
  workflow: QC Workflow

  samples:
    - ESP001
    - ESP002

  protocols:
    - QC Quantification:
      data:
        ESP001:
          Type: DNA
          Ratio: 1.8
        ESP002:
          Type: RNA
          Ratio: 1.9
```

```
# Contents of tests/resources/QC-Test-2.yml

My QC Experiment 2:
  submit: True
  project: QC Project
  workflow: QC Workflow

  samples:
    - ESP003
    - ESP004

  protocols:
    - QC Quantification:
      data:
        ESP003:
          Type: DNA
          Ratio: 1.7
        ESP004:
          Type: RNA
          Ratio: 2.0
```

After creating these two configs, we can update our tests to run them after *Workflows* are created. Here is an updated version of our original test file:

```
# -*- coding: utf-8 -*-

# imports
import os
import unittest

from . import CONFIG, RESOURCES, CONTENT
```

(continues on next page)

(continued from previous page)

```

from esp.testing import ModelLoaderMixin

# tests
class TestQCWorkflows(ModelLoaderMixin, unittest.TestCase):
    IMPORT = dict(
        Workflow=[
            os.path.join(CONTENT, 'workflows', 'QC-Workflow.yml'),
        ],
        Pipeline=[
            os.path.join(CONTENT, 'pipelines', 'QC-Report-Pipeline.yml'),
        ]
    )
    DATA = dict(
        Project=[
            {'name': 'QC Project'}
        ],
        Experiment=[
            os.path.join(RESOURCES, 'QC-Test-1.yml'),
            os.path.join(RESOURCES, 'QC-Test-2.yml'),
        ]
    )

```

Let's go over the components of this test class in more detail:

- `ModelLoaderMixin` - A test mixin used for managing content imports and setup/teardown during testing. For testing out content in this way, you'll want to include both `esp.testing.ModelLoaderMixin` and `unittest.TestCase` in your tests.
- `IMPORT` - A class variable that is a dictionary of mappings from model type to models that should be imported as a part of the tests. For information on the types of models you can create configuration for, see documentation for the Python client.
- `DATA` - A class variable that is similar to `IMPORT`, but is reserved specifically for creating *Experiments* and *Projects* for testing out content.

Under the hood, any content defined within the `IMPORT` and `DATA` blocks of the tests will be using the Python client to create the content by calling `Model.create()` for a specific model. For example, if your tests look like:

```

class TestQCWorkflows(ModelLoaderMixin, unittest.TestCase):
    IMPORT = dict(
        SampleType=[
            os.path.join(RESOURCES, 'inventory', 'Sample-Types.yml'),
            {
                'name': 'My Sample Type',
                'desc': 'Description for My Sample Type'
            }
        ]
    )

```

The configuration will be created with this equivalent code:

```

from esp.models import SampleType
SampleType.create(
    os.path.join(RESOURCES, 'inventory', 'Sample-Types.yml')
)
SampleType.create(
    name='My Sample Type',

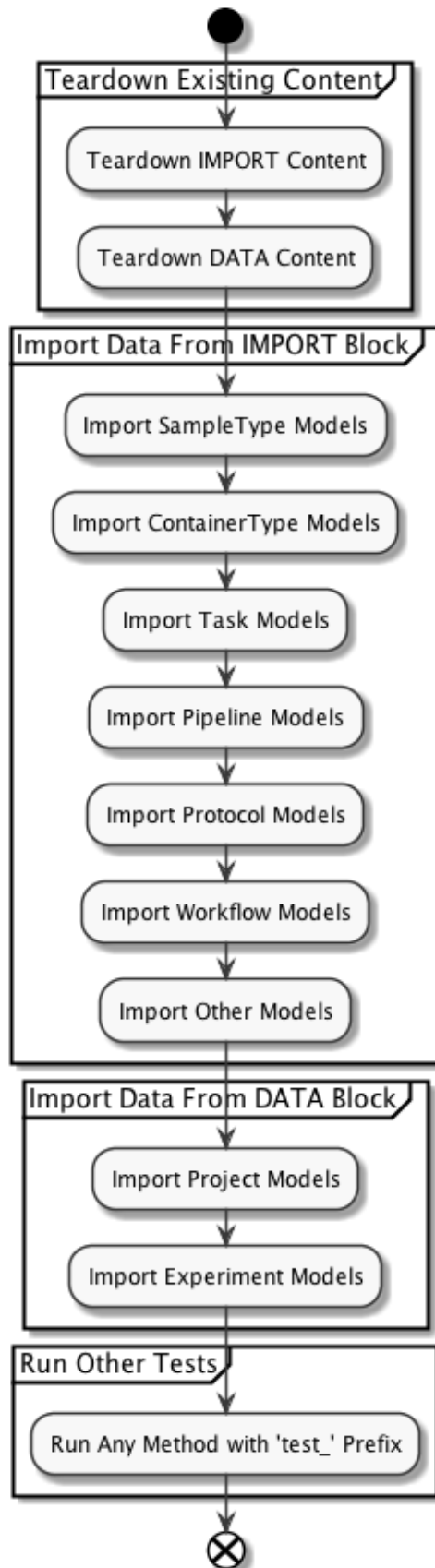
```

(continues on next page)

(continued from previous page)

```
desc='Description for My Sample Type'  
)
```

For additional context, the following flow diagram describes testing steps using the format described above (i.e. any test using a `ModelLoaderMixin`):



1.7.3 Using Configuration

In addition to explicitly defining what configuration you want to include in your testing, you can also import default configuration, specified in **seed** files within the repository (see the *Content* section for more information on those seed files).

To import a seed file in your testing, you can update your test class to look like the following (keeping with the example used above):

```
class TestQCWorkflows(ModelLoaderMixin, unittest.TestCase):
    SEED = os.path.join(ROLES, 'demos', 'content.yml')
    DATA = dict(
        Project=[
            {'name': 'QC Project'}
        ],
        Experiment=[
            os.path.join(RESOURCES, 'QC001.yml'),
            os.path.join(RESOURCES, 'QC002.yml'),
        ]
    )
```

1.8 Windows Setup

1.8.1 Installation Prerequisites

1. Linux on Windows with WSL
 - <https://docs.microsoft.com/en-us/windows/wsl/install>
2. Ubuntu v20.04-LTS (download and install from Microsoft store app)
 - <https://www.microsoft.com/store/productId/9MTTCL66CPXJ>
3. Windows Terminal (download from Microsoft store app)
 - <https://www.microsoft.com/store/productId/9N0DX20HK701>
4. Docker Desktop
 - Install Desktop Docker with WSL2 backend
 - <https://docs.docker.com/desktop/windows/wsl/>
5. VS Code (Remote-wsl and python extensions are helpful)
 - Install VS Code for your Windows version: <https://code.visualstudio.com/>
 - Then from within the Extensions install the following extensions:
 - remote-wsl (<https://marketplace.visualstudio.com/items?itemName=ms-vscode-remote.remote-wsl>)
 - remote-containers (<https://marketplace.visualstudio.com/items?itemName=ms-vscode-remote.remote-containers>)
 - python (<https://marketplace.visualstudio.com/items?itemName=ms-python.python>)

1.8.2 Setup Instructions

1. Run (or Setup for first time) Ubuntu in Windows.

Note: Windows may not be ready to install a Linux environment & will inform you to activate Windows Subsystem for Linux first. If you get such an error, see (<https://docs.microsoft.com/en-us/windows/wsl/install>) for additional details (and follow these instructions):

- Press Win+R
- Type: cmd.exe
- To gain elevated admin privileges, press: Ctrl+Shift+Enter
- At the command line, type: `cd C:\WINDOWS\system32`
- Then: `wsl --install -d Ubuntu-20.04`

2. Next follow the prompts to install the actual Ubuntu system.

3. Launch Ubuntu and set your username and password

- Close Ubuntu window

4. Launch windows terminal:

- In the Windows Terminal settings:
 - In “Startup” settings: Set the Default System to “ubuntu 20.04” (and save changes)
 - Under “Profiles”, change the starting directory for “Ubuntu 20.04” to:


```
* \\wsl$\Ubuntu-20.04\home\<username>
```

Note: Replace “<username>” with the username you created during the Ubuntu Install.

5. Update/install the following packages in your windows terminal linked to Ubuntu

```
sudo apt update -y && sudo apt upgrade -y
sudo apt install -y python3-pip python3-venv nano
```

6. After Ubuntu is set up, add a WSL configuration file in the /etc folder of your Ubuntu instance so it mounts the drive with metadata and the SSH works (do these with windows terminal)

```
sudo nano /etc/wsl.conf
# Type in this and save the file:
[automount]
options = "metadata"
```

7. Add a WSL configuration file in the Windows user directory (change to your username), to limit memory usage (by default WSL will consume 50% or 8GB of memory or whichever amount is smaller)

- Add Global memory limits for the Windows Environment via Powershell to your Windows home directory (e.g. C:\Users\<username>\.wslconfig):

```
sudo nano /mnt/c/Users/<user>/.wslconfig
# Type in this and save the file:
[wsl2]
processors=2 # Limits WSL2 VM use to two virtual processors
```

(continues on next page)

(continued from previous page)

```
memory=6GB # Limits WSL2 VM memory to 6GB
```

- Capability to add memory limits on the installed distribution (Ubuntu) Environment. → are not necessary to perform, but is referenced here **for** awareness:
- Access the ``/etc`` directory by entering:

```
cd /etc
# Verify you have a wsl.conf file in the /etc/ directory contents & then edit:
ls -la | grep wsl.conf
nano wsl.conf
```

Note: To see which options you can place into the configuration file, view this article: - <https://docs.microsoft.com/en-us/windows/wsl/wsl-config#configure-global-options-with-wslconfig>

1. Run Docker Desktop, enable integration with WSL2 Ubuntu & link your docker to Bitbucket.

- In Docker Desktop you will need to add the Ubuntu distribution to your WSL2 settings for WSL integration:
 - Go into “Resources” and select sub-menu “WSL Integration”: Enable integration with additional distros...
 - Turn on switch for: “Ubuntu-20.04” (first make sure “Enable integration with my default WSL distro is checked”)

If ‘Ubuntu-20.04’ switch is not present, trouble shoot using the following:

- In Powershell, type in the following command to check which version of Ubuntu you are running:

```
wsl --list -verbose
# Version should be 2 for all names.
NAME                STATE      VERSION
Ubuntu-20.04        Running    2
docker-desktop       Running    2
docker-desktop-data  Running    2

# If Ubuntu-18.04 is in Version 1, enter the following command:
wsl --set-default-version 2
# Check again which version of Ubuntu you are running.
```

If you are running version 2, Refresh Docker Desktop and the Ubuntu-20.04 switch should be visible.

1.9 Overview

Below, you can find code-heavy examples of how to use the L7|ESP SDK. These examples are not meant to be comprehensive, but rather give users a feel for how to best leverage tools in the L7|ESP SDK.

It is highly recommended that readers of this document complete the L7|ESP Quick Start tutorial through the UI and Python client (see core and client L7|ESP documentation) before completing this tutorial via the L7|ESP SDK. This tutorial mirrors building out that content, so users should come in with an understanding of how content definitions work within the context of L7|ESP.

In addition, this documentation refers heavily to the L7|ESP Python client. Please see that documentation for information on usage patterns.

1.9.1 Tasks

To demonstrate building out content for a specific lab, we'll build a simple end-to-end *Workflow*. Upon completion of this example, we'll have used L7|ESP to do the following:

- Create a simple *Protocol* that tracks values associated with *Samples*.
- Create a simple *Pipeline* that generates a *Report* for each *Sample*.
- Create a *Workflow* that allows the *Protocol* and *Pipeline* to be used together for processing *Samples*.
- Organize content in the L7|ESP SDK so that it can be referenced during development.
- Create tests for evaluating functionality of that *Workflow*.
- Update the L7|ESP SDK configuration to automatically load that content on install.

In summary, at the end of this tutorial, we will have a fully functional laboratory management system with one *Workflow* already set up!

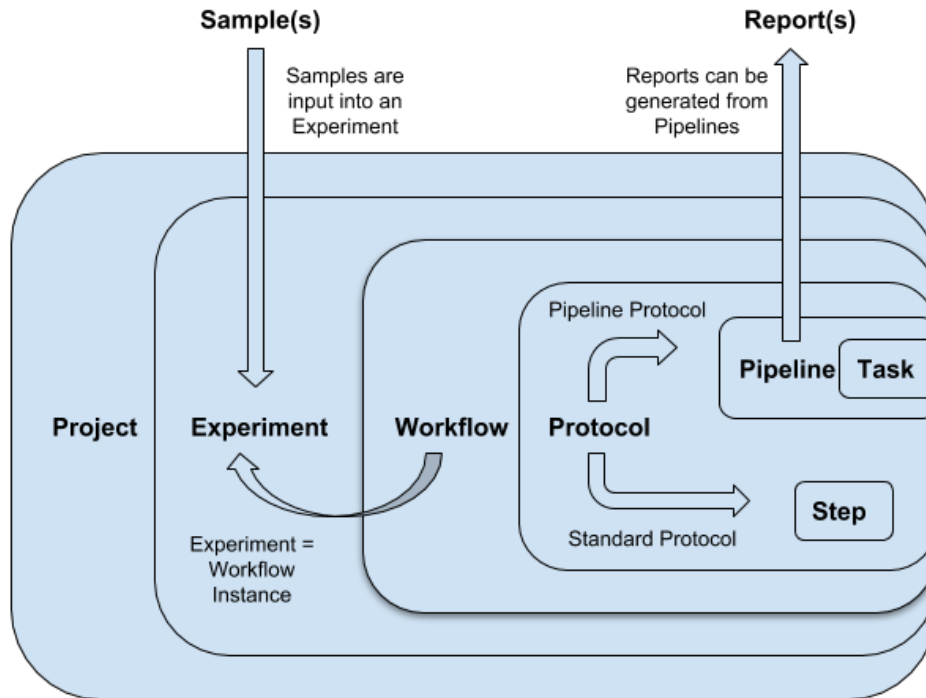
Note: L7|ESP uses many common terms for elements within the platform. Throughout this manual, terms with specific meanings in L7|ESP will always be capitalized, while lowercase versions of the same words are used in the general sense. For example, a laboratory protocol can be captured in an L7|ESP *Protocol*, and then grouped into an L7|ESP *Workflow*.

1.9.2 Terminology

While this list is not entirely comprehensive, it covers basic terminology for the scope of this introductory demo.

- **User:** Any person who uses L7|ESP
- **File:** A reference to a file registered as a Resource and its associated metadata
- **Sample:** A physical sample registered within L7|ESP. *Note that this term has a general purpose, and can be used to represent a raw sample, an aliquot, etc.*
- **Step:** A procedural step defined within an L7|ESP *Protocol*
- **Task:** A computational step in an L7|ESP *Pipeline*
- **Protocol:** A sequence of steps performed on a *Sample*. Note that this guide has the user create both standard and pipeline *Protocols*. Pipeline *Protocols* allow users to run *Pipelines* as part of a larger *Sample*-processing *Workflow*.
- **Pipeline:** An ordered collection of one or more L7|ESP *Tasks*
- **Report:** A collection of visual elements that display results from an L7|ESP *Pipeline*
- **Workflow:** An ordered collection of L7|ESP *Protocols*
- **Experiment:** An instance of a particular L7|ESP *Workflow* with associated *Samples* and data
- **Project:** A grouping mechanism for L7|ESP *Experiments*, *Pipelines*, *Samples*, *Reports*, etc.

A high-level diagram that displays the general relationship between these terms can be viewed below:



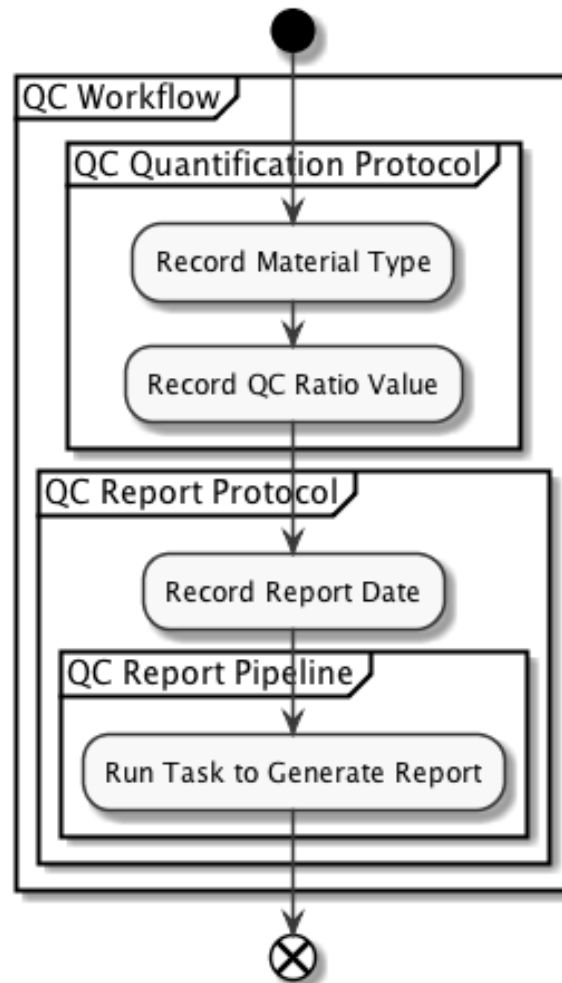
1.10 Create Workflow

For this tutorial, the user will build a simple QC *Workflow* that records a QC value and calculates a Pass/Fail result as a *Report*. By this point, you've already seen this *Workflow* in other tutorials, so we'll focus on how to create the content in the context of the L7|ESP SDK.

For this demo, our *Workflow* will have two parts:

1. A *Protocol* to capture QC metadata about a *Sample*.
2. A *Protocol* to generate a *Report* using a *Pipeline*.

Here's a schematic detailing the components of that *Protocol* at a high level:



All content developed within the L7|ESP SDK is defined within YAML config files that live in the `content` directory within the repository. For more information about the structure of the repository, see the [Project Structure](#) section of the documentation.

1.10.1 Step 1: Create a *Protocol*

Now that we know what type of content we’re going to create, let’s make some content files. First, we need to create our first *Protocol* (QC Quantification):

Using the L7|ESP SDK, we can create this configuration via YAML config file. For more information on the YAML config format or types of content that can be created with the YAML config format, please see the L7|ESP Python client documentation.

For this first *Protocol*, **create the following file**:

```
# Contents of: content/protocols/QC-Quantification.yml

QC Quantification:
  protocol: standard
  variables:
    - Type:
```

(continues on next page)

(continued from previous page)

```

    rule: dropdown
    dropdown:
      - 'DNA'
      - 'RNA'
  - Ratio:
    rule: numeric

```

In this file, we've defined an *L7IESP Protocol* that has two columns the user will input when filling out *Worksheets*: *Type* and *Ratio*. *Type* is a column of dropdown type with options *DNA* and *RNA*, and *Ratio* is a numeric column. Other types of *Protocol* columns include (also listed in the L7IESP Python client documentation):

- `string` - Column with text or string data.
- `numeric` - Column with numeric data.
- `dropdown` - Column with a dropdown selector. The options for the dropdown are specified as a list in the `dropdown` property.
- `checkbox` - Column with checkbox indicating yes/no status.
- `date` - Column with date picker for selecting calendar date.
- `attachment` - Column for uploading file attachment.
- `link` - Column with data containing external link.
- `barcode` - Column for scanning barcode data and printing various types of barcodes (QR, 1D, Mini DataMatrix, etc.).

Note: This list may not be fully comprehensive. For the full list, please see the L7IESP Python client documentation.

Along with this *Protocol*, we're going to define an *L7IESP Pipeline* and pipeline *Protocol* to include in our *Workflow*.

1.10.2 Step 2: Create a Pipeline to generate a Report:

To create a *Pipeline* for use in a *Workflow*, we need to create two things:

1. A *Pipeline* that will generate a *Report* when given *Type* and *Ratio* QC data.
2. A pipeline *Protocol* that will link the values from the *QC Quantification Protocol* to our *Pipeline*.

To start, **define the report-generating Pipeline:**

```

# Contents of: content/pipelines/QC-Report-Pipeline.yml

QC Report Pipeline:
  report:
    name: Result Report
    elements:
      - depends:
        - {file: Result Report, tasknumber: 1}
        type: raw_file
      - []
  tasks:
    - Generate QC Report:
      desc: Task to generate QC report from workflow metadata
      cmd: |+

```

(continues on next page)

(continued from previous page)

```

# Simple Task that determines if the specified 'ratio' is in the proper_
↪range

RNA_MIN=1.8
RNA_MAX=2.1

DNA_MIN=1.7
DNA_MAX=2.0

TYPE="{{ material }}"
RATIO="{{ ratio }}"

if [ $TYPE = "RNA" ]; then
    PASS=`echo "$RNA_MIN <= $RATIO && $RATIO <= $RNA_MAX" | bc`;
elif [ $TYPE = "DNA" ]; then
    PASS=`echo "$DNA_MIN <= $RATIO && $RATIO <= $DNA_MAX" | bc`;
fi

echo $PASS

if [[ $PASS = 1 ]]; then
    echo "<b>Your sample <font color='green'>Passed</font> and contains pure
↪$TYPE</b>" >> result.html
else
    echo "<b>Your sample <font color='red'>Failed</font> and is NOT pure
↪$TYPE</b>" >> result.html
fi
files:
- Result Report:
    file_type: html
    filename_template: "{{ 'result.html' }}"

```

In the block above, we're defining a *Pipeline* with a single *Task*. In that *Task*, the `cmd` block contains code that will be run as part of the *Pipeline*. This code can either be bash code or a call to an external script that will run the *Pipeline*. Also in the `cmd` block are references to `{{ material }}` and `{{ ratio }}` variables that are passed into the *Pipeline*. Next, we'll create configuration for a *Protocol* that can run this *Pipeline* as part of our *Workflow*.

1.10.3 Step 3: Create a 'pipeline' *Protocol* to record data and run the *Pipeline*:

Now, let's create our pipeline *Protocol* for running the *Pipeline* that we defined above. For this *Protocol*, we want to capture column data (Type and Ratio) from a previous *Protocol* so they can be passed as arguments (`material` and `ratio`) to the *Pipeline*. We also want to include a date column for users to specify a date when the QC data were taken.

Create the following file to capture all of this information:

```

# Contents of: content/protocols/QC-Report.yml

QC Report:
  protocol: pipeline
  pipeline: QC Report Pipeline
  variables:
    - material:
        rule: text
        value: "{{ column_value('Type', 'QC Quantification') }}"
        visible: false

```

(continues on next page)

(continued from previous page)

```

- ratio:
  rule: numeric
  value: "{{ column_value('Ratio', 'QC Quantification') }}"
  visible: false
- Report Date:
  rule: date

```

Above, we can reference our *Pipeline* using the pipeline configuration option. Additionally, for *Protocols* that are executing a *Pipeline*, `protocol:` pipeline must be explicitly specified as part of the *Protocol* definition. Finally, to carry data over from a previous *Protocol*, we're using an L7|ESP expression in the `value` field for the column. For more information on expressions and how they're used in L7|ESP, see the L7|ESP documentation.

1.10.4 Step 4: Put it all together in a *Workflow*:

Now that we our *Pipeline* and *Protocols* defined, we can put them all together in a *Workflow* that can be used in the lab.

Define the *Workflow* as follows:

```

# Contents of: content/workflows/QC-Workflow.yml

QC Workflow:
  desc: Workflow to measure QC values and generate a QC report.
  tags:
    - quickstart
    - qc

  protocols:
    - QC Quantification
    - QC Report

```

Above, since our *Protocol* configuration is already defined, we can simply reference them in the `protocols` block of the *Workflow*.

Notes

You can also define all of these items in a nested way.

For instance, if you want to define the *QC Workflow* with a single *Protocol* all in the same config file, **instead, define the Workflow this way:**

```

# workflow
QC Workflow:
  desc: Workflow to measure QC values and generate a QC report.
  tags:
    - quickstart
    - qc

  # nested protocols
  protocols:
    - QC Quantification:
      protocol: pipeline
      variables:
        - Type:
          rule: dropdown

```

(continues on next page)

(continued from previous page)

```

dropdown:
  - 'DNA'
  - 'RNA'
- Ratio:
  rule: numeric

```

1.11 Create Test

Test-driven development is highly recommended as good practice when using the L7|ESP SDK, so we'll start our tutorial with testing. As a gentle introduction, let's write a simple test to verify that L7|ESP is connected and accepting requests.

To do this task, **create a test file in the ``tests`` directory named ``test\_status.py``**. The contents of that file should look something like:

```

# -*- coding: utf-8 -*-

import pytest

def test_status():
    import esp
    running = esp.status()
    assert running, 'Could not connect to L7|ESP via Python client!'

```

Now that you've written a test, **run it using ``pytest``**:

```

~$ pytest tests/test_status.py b
===== test session starts.
↳=====
platform darwin -- Python 3.7.1, pytest-3.10.0, py-1.7.0, pluggy-0.8.0 -- /usr/local/
↳opt/python/bin/python3.7
cachedir: .pytest_cache
collected 1 item

tests/test_status.py::test_status Connection established!

PASSED

===== 1 passed in 0.68 seconds.
↳=====

```

If L7|ESP was not running and the tests failed, you'd see the following:

```

~$ pytest tests/test_status.py b
===== test session starts.
↳=====
platform darwin -- Python 3.7.1, pytest-3.10.0, py-1.7.0, pluggy-0.8.0 -- /usr/local/
↳opt/python/bin/python3.7
cachedir: .pytest_cache
collected 1 item

tests/test_status.py::test_status FAILED

===== 1 failed in 0.68 seconds.
↳=====

```

Later in this tutorial, we'll talk about different ways of structuring tests for various types of content definitions.

1.11.1 Step 1: Create test file for importing data

Now that we have our content defined, we can use L7|ESP SDK tools to write a simple test that will import these files and verify that they exist in the system. This test also enables us to verify that there are no syntax errors in our *Workflow* definitions.

Create the test file, `tests/test_qc_workflow.py`, to look like the following:

```
# -*- coding: utf-8 -*-

# imports
import os
import unittest

from . import CONFIG, RESOURCES, CONTENT
from esp.testing import ModelLoaderMixin

# tests
class TestQCWorkflows(ModelLoaderMixin, unittest.TestCase):
    IMPORT = dict(
        Workflow=[
            os.path.join(CONTENT, 'workflows', 'QC-Workflow.yml'),
        ],
        Protocol=[
            os.path.join(CONTENT, 'protocols', 'QC-Quantification.yml'),
            os.path.join(CONTENT, 'protocols', 'QC-Report.yml'),
        ],
        Pipeline=[
            os.path.join(CONTENT, 'pipelines', 'QC-Report-Pipeline.yml'),
        ]
    )
```

Run this test to get the following output:

```
~$ pytest tests/test_qc_workflow.py

===== test session starts _
↪=====
platform darwin -- Python 3.7.1, pytest-4.1.0, py-1.7.0, pluggy-0.8.0 -- /usr/local/
↪opt/python/bin/python3.7
cachedir: .pytest_cache
collected 1 item

tests/test_qc_workflow.py::TestQCWorkflows::test__content
INFO:root:Clearing existing content from database.
INFO:root:Creating Task: Generate QC Report
INFO:root:Creating Pipeline: QC Report Pipeline
INFO:root:Creating PipelineReport: QC Report
INFO:root:Creating Protocol: QC Quantification
INFO:root:Creating Protocol: QC Report
INFO:root:Creating Workflow: QC Workflow
Successfully imported config data.
PASSED

===== 1 passed in 3.47 seconds _
↪=====
```

(continues on next page)

(continued from previous page)

From this test output, we can see that all of our content was imported correctly, and references within the files were properly defined.

1.11.2 Step 2: Create *Experiment* Configs for Testing

Along with simply testing *Workflow* definitions, we can also test that the *Workflow* works properly when used in an *Experiment*. To do this, we can use the same format as above, but add a `DATA` attribute to our test class for what *Projects* and *Experiments* need to be created. For this example, we'll need to create a *Project* and two *Experiments* for testing out the *Workflow*.

Create the following config files for our two *Experiments*:

```
# Contents of tests/resources/QC-Test-1.yml
```

```
QC Experiment 1:
  submit: True
  project: QC Project
  workflow: QC Workflow

  samples:
    - ESP001
    - ESP002

  protocols:
    - QC Quantification:
      data:
        ESP001:
          Type: DNA
          Ratio: 1.8
        ESP002:
          Type: RNA
          Ratio: 1.9
```

```
# Contents of tests/resources/QC-Test-2.yml
```

```
QC Experiment 2:
  submit: True
  project: QC Project
  workflow: QC Workflow

  samples:
    - ESP003
    - ESP004

  protocols:
    - QC Quantification:
      data:
        ESP003:
          Type: DNA
          Ratio: 1.7
        ESP004:
          Type: RNA
          Ratio: 2.0
```

1.11.3 Step 3: Update Tests to Include *Experiments*

After creating these two configs, we can update our tests to run them after *Workflows* are created.

Update the original test file to look as follows:

```
# -*- coding: utf-8 -*-

# imports
import os
import unittest

from . import CONFIG, RESOURCES, CONTENT
from esp.testing import ModelLoaderMixin

# tests
class TestQCWorkflows(ModelLoaderMixin, unittest.TestCase):
    IMPORT = dict(
        Workflow=[
            os.path.join(CONTENT, 'workflows', 'QC-Workflow.yml'),
        ],
        Protocol=[
            os.path.join(CONTENT, 'protocols', 'QC-Quantification.yml'),
            os.path.join(CONTENT, 'protocols', 'QC-Report.yml'),
        ],
        Pipeline=[
            os.path.join(CONTENT, 'pipelines', 'QC-Report-Pipeline.yml'),
        ]
    )
    DATA = dict(
        Project=[
            {'name': 'QC Project'}
        ],
        Experiment=[
            os.path.join(RESOURCES, 'QC-Test-1.yml'),
            os.path.join(RESOURCES, 'QC-Test-2.yml'),
        ]
    )
)
```

Finally, run the updated test to get the following output:

```
~$ pytest tests/test_qc_workflow.py

===== test session starts =====
↪=====
platform darwin -- Python 3.7.1, pytest-4.1.0, py-1.7.0, pluggy-0.8.0 -- /usr/local/
↪opt/python/bin/python3.7
cachedir: .pytest_cache
collected 1 item

tests/test_demos/test_quickstart.py::TestQuickstart::test__content
INFO:root:Clearing existing content from database.
INFO:root:Creating Task: Generate QC Report
INFO:root:Creating Pipeline: QC Report Pipeline
INFO:root:Creating PipelineReport: QC Report
INFO:root:Creating Protocol: QC Quantification
INFO:root:Creating Protocol: QC Report
INFO:root:Creating Workflow: QC Workflow
```

(continues on next page)

(continued from previous page)

```

INFO:root:Creating Project: QC Project
INFO:root:Creating Sample: ESP001
INFO:root:Creating Sample: ESP002
INFO:root:Creating Experiment: QC Experiment 1
INFO:root:Submitting Experiment: QC Experiment 1
INFO:root:Creating SampleSheet: QC Experiment 1
INFO:root:Creating Sample: ESP003
INFO:root:Creating Sample: ESP004
INFO:root:Creating Experiment: QC Experiment 2
INFO:root:Submitting Experiment: QC Experiment 2
INFO:root:Creating SampleSheet: QC Experiment 2
Successfully imported config data.
PASSED

===== 1 passed in 8.54 seconds_
↪=====

```

From this test output, we can see that all of our testing data was defined correctly (along with our content).

1.12 Run Tests

1.12.1 Step 1: Run All Integration Tests

Now, putting it all together, **run all of the tests we've defined** in the L7|ESP SDK using the test make target:

```
~$ make test
```

This command will clear existing related content from the database, and run both the `test_status.py` and `test_qc_workflow.py` tests.

1.12.2 Step 2: Clear the Database

By default, artifacts created during tests are not removed after a test run so that developers can visit the application UI and see any resulting content. When tests are re-run, the test suite will automatically tear down content from a previous run and then proceed with executing tests. Usually, you won't need to clear the database between running tests, but if you want to run everything from a clean slate, you can use the commands below to clear the database outside of a test run. There are two ways to do this, both outlined in the [Administration](#) section of the documentation:

```

~$ # keep the application running and use
~$ # the python client to run tests
~$ make clear

~$ # completely wipe the database and reload
~$ # the application
~$ make reset

```

1.13 Seed Content

Finally, now that we've created and tested a *Workflow* that's ready to be deployed in our production L7|ESP instance, we can update our production configuration to import the *Workflow*. There are two components to this process:

1. Add configuration to a new or existing **seed** file that can be imported on install/update.
2. If necessary, add a reference to the **seed** file in our production configuration file, `roles/packages.yml`.

1.13.1 Step 1: Create Seed File to Load Content

As a first step, we'll need to create a **seed** file to reference our configuration. In the L7|ESP SDK, a **seed** file is a file used by the client to import configuration for specific models. In the file, for each item to import into the system, you specify configuration and a model that the configuration describes.

Create this example seed file for importing our *Workflow*, *Protocols*, and *Pipeline* defined above:

```
# Contents of: roles/workflows.yml

- model: Workflow
  data: $LAB7DATA/content/workflows/QC-Workflow.yml

- model: Protocol
  data: $LAB7DATA/content/protocols/QC-Quantification.yml

- model: Protocol
  data: $LAB7DATA/content/protocols/QC-Report.yml

- model: Pipeline
  data: $LAB7DATA/content/pipelines/QC-Report-Pipeline.yml
```

In the file above, `$LAB7DATA` is an environment variable referencing the L7|ESP SDK root directory. It can be used both in development and production deployments to reference the same location.

1.13.2 Step 2: Update Deployment Config to Use Seed Data

Once we've created our **seed** file, **update the shared configuration file**, `roles/packages.yml`, to reference that seed file for the content L7|ESP will load into the system by default:

```
---

biobuilds:
  - bwa

seed:
  - '{{ sdk }}/roles/workflows.yml'
```

In the file above, `{{ sdk }}` is an Ansible variable referencing the L7|ESP SDK root directory. It can be used both in development and production deployments to reference the same location.

1.13.3 Step 3: Test Seeding Content Locally

Before deploying changes, we'll want to test that the seed files and changes to production configuration are properly imported by the software.

To do so, **use the `import` make target**:

```
~$ make import
```

Since the same commands are used to seed the content in development as in production, this practice is sufficient for ensuring that content will be imported the next time you update your production instance.

1.14 Summary

That's it! In summary, by going through this tutorial, you now know how to:

1. Create content in the L7|ESP SDK to meet various lab workflow and analysis needs.
2. Organize content in the L7|ESP SDK so it can be referenced during development.
3. Update L7|ESP SDK configuration to automatically load content on install.
4. Write tests for *Workflow* definition integrity and functionality.
5. Run tests to verify functionality.

1.15 Container Infrastructure

By adopting a container model, the L7|ESP platform aims to remain cloud-agnostic and simple to deploy.

Containers solve a number of problems ranging from immutable deployments to self-healing infrastructure.

This also means that there are many solutions for running and managing containers in production to choose from.

1.15.1 Kubernetes

Kubernetes, also known as K8s, is an open-source system for automating deployment, scaling, and management of containerized applications.

The L7|ESP platform can be deploy to any Kubernetes cluster using the [Helm](#) package manager and the provided [Chart](#).



Amazon Elastic Kubernetes Service

Requirements

To install the L7|ESP Helm chart on Amazon Elastic Kubernetes Service (Amazon EKS), you must have:

- An AWS account to deploy to
- Access keys for the AWS account (see [Managing access keys for IAM users](#))

The `kubectl` and `aws` command line tools will be use respectively to programmatically access the Kubernetes cluster and the AWS account. To install the CLI tools, follow the link below and choose the correct operating sytem:

- [Kubernetes CLI](#)
- [AWS CLI](#)

Infrastructure as Code

CloudFormation Template

The latest CloudFormation template is as follows:

CloudFormation template

```
AWSTemplateFormatVersion: '2010-09-09'
Description: EKS cluster using a VPC with two public subnets

Parameters:
  EKSClusterName:
    Type: String
    Default: ""
    Description: The desired name of your AWS EKS Cluster.

  EKSVersion:
    Type: String
    Default: 1.22
    AllowedValues:
      - 1.16
      - 1.17
      - 1.18
      - 1.21
      - 1.22
    Description: The desired version of your AWS EKS Cluster.

  EKSNodeGroupName:
    Type: String
    Default: "NodeGroup01"
    Description: The desired name of your AWS EKS Node Group.

  NodeAutoScalingGroupDesiredCapacity:
    Type: Number
    Default: 2
    Description: Number of desired Worker Node.
    # MinValue: 1
```

(continues on next page)

(continued from previous page)

```

# MaxValue: 5

NodeAutoScalingGroupMinSize:
  Type: Number
  Default: 1
  Description: Minimum size of Node Group ASG.

NodeAutoScalingGroupMaxSize:
  Type: Number
  Default: 5
  Description: Maximum size of Node Group ASG. Set to at least 1 greater than
↳NodeAutoScalingGroupDesiredCapacity.

EKSWorkerNodeInstanceType:
  Type: String
  Default: t2.large
  AllowedValues: [t2.nano, t2.micro, t2.small, t2.medium, t2.large, t2.xlarge, t2.
↳2xlarge,
  t3.nano, t3.micro, t3.small, t3.medium, t3.large, t3.xlarge, t3.2xlarge,
  m4.large, m4.xlarge, m4.2xlarge, m4.4xlarge, m4.10xlarge,
  m5.large, m5.xlarge, m5.2xlarge, m5.4xlarge,
  c5.large, c5.xlarge, c5.2xlarge, c5.4xlarge, c5.9xlarge,
  g3.8xlarge, r5.large, r5.xlarge, r5.2xlarge, r5.4xlarge, r3.12xlarge,
  i3.xlarge, i3.2xlarge, i3.4xlarge, i3.8xlarge,
  d2.xlarge, d2.2xlarge, d2.4xlarge, d2.8xlarge]
  ConstraintDescription: Must be a valid EC2 instance type
  Description: EC2 instance type for the node instances.

EKSIAMRoleName:
  Type: String
  Default: ""
  Description: The name of the IAM role for the EKS service to assume.

EKSPKeyPair:
  Type: "AWS::EC2::KeyPair::KeyName"
  Default: "devopskey"
  Description: The name of Key Pair to establish connection with Worker Node.

VpcBlock:
  Type: String
  Default: 10.0.0.0/16
  Description: The CIDR range for the VPC. This should be a valid private (RFC
↳1918) CIDR range.
  AllowedPattern: (\d{1,3})\.\d{1,3})\.\d{1,3})\.\d{1,3})/(\d{1,2})
  ConstraintDescription: must be a valid IP CIDR range of the form x.x.x.x/x.

PublicSubnet01Block:
  Type: String
  Default: 10.0.0.0/24
  Description: CidrBlock for public subnet 01 within the VPC.
  AllowedPattern: (\d{1,3})\.\d{1,3})\.\d{1,3})\.\d{1,3})/(\d{1,2})
  ConstraintDescription: must be a valid IP CIDR range of the form x.x.x.x/x.

PublicSubnet02Block:
  Type: String
  Default: 10.0.1.0/24
  Description: CidrBlock for public subnet 02 within the VPC.

```

(continues on next page)

(continued from previous page)

```

AllowedPattern: (\d{1,3})\.\d{1,3})\.\d{1,3})\.\d{1,3})/(\d{1,2})
ConstraintDescription: must be a valid IP CIDR range of the form x.x.x.x/x.

AvailabilityZonePublicSubnet01:
  Type: CommaDelimitedList<AWS::EC2::AvailabilityZone::Name>
  Default: us-east-1a
  Description: Availability Zone for the Public Subnet 01.

AvailabilityZonePublicSubnet02:
  Type: CommaDelimitedList<AWS::EC2::AvailabilityZone::Name>
  Default: us-east-1b
  Description: Availability Zone for the Public Subnet 02.
Metadata:
  AWS::CloudFormation::Interface:
    ParameterGroups:
    -
      Label:
        default: "Worker Network Configuration"
      Parameters:
        - VpcBlock
        - PublicSubnet01Block
        - AvailabilityZonePublicSubnet01
        - PublicSubnet02Block
        - AvailabilityZonePublicSubnet02
    -
      Label:
        default: "EKS Cluster Information"
      Parameters:
        - EKSClusterName
        - EKSVersion
        - EKSNodeGroupName
        - NodeAutoScalingGroupDesiredCapacity
        - EKSWorkerNodeInstanceType
        - EKSIAMRoleName
        - EKSKeyPair
        - NodeAutoScalingGroupMinSize
        - NodeAutoScalingGroupMaxSize

Mappings:
  ServicePrincipals:
    aws-cn:
      ec2: ec2.amazonaws.com.cn
    aws-us-gov:
      ec2: ec2.amazonaws.com
    aws:
      ec2: ec2.amazonaws.com

Resources:
  eksVPC:
    Type: AWS::EC2::VPC
    Properties:
      CidrBlock: !Ref VpcBlock
      EnableDnsSupport: true
      EnableDnsHostnames: true
    Tags:
      - Key: Name

```

(continues on next page)

(continued from previous page)

```

    Value: !Sub '${AWS::StackName}-VPC'
  - Key: Project
    Value: aws-eks

eksInternetGateway:
  Type: AWS::EC2::InternetGateway
  Properties:
    Tags:
      - Key: Name
        Value: !Sub '${AWS::StackName}-InternetGateway'
      - Key: Project
        Value: aws-eks

eksVPCGatewayAttachment:
  Type: AWS::EC2::VPCGatewayAttachment
  Properties:
    InternetGatewayId: !Ref eksInternetGateway
    VpcId: !Ref eksVPC

eksPublicRouteTable:
  Type: AWS::EC2::RouteTable
  Properties:
    VpcId: !Ref eksVPC
    Tags:
      - Key: Name
        Value: !Sub '${AWS::StackName}-RouteTable'
      - Key: Project
        Value: aws-eks

eksPublicRoute:
  DependsOn: eksVPCGatewayAttachment
  Type: AWS::EC2::Route
  Properties:
    RouteTableId: !Ref eksPublicRouteTable
    DestinationCidrBlock: 0.0.0.0/0
    GatewayId: !Ref eksInternetGateway

eksPublicSubnet01:
  Type: AWS::EC2::Subnet
  Properties:
    AvailabilityZone: !Ref AvailabilityZonePublicSubnet01
    MapPublicIpOnLaunch: true
    CidrBlock:
      Ref: PublicSubnet01Block
    VpcId:
      Ref: eksVPC
    Tags:
      - Key: Name
        Value: !Sub "${AWS::StackName}-PublicSubnet01"
      - Key: Project
        Value: aws-eks

eksPublicSubnet02:
  Type: AWS::EC2::Subnet
  Properties:
    AvailabilityZone: !Ref AvailabilityZonePublicSubnet02
    MapPublicIpOnLaunch: true

```

(continues on next page)

(continued from previous page)

```

    CidrBlock:
      Ref: PublicSubnet02Block
    VpcId:
      Ref: eksVPC
    Tags:
      - Key: Name
        Value: !Sub "${AWS::StackName}-PublicSubnet02"
      - Key: Project
        Value: aws-eks

eksPublicSubnet01RouteTableAssociation:
  Type: AWS::EC2::SubnetRouteTableAssociation
  Properties:
    SubnetId: !Ref eksPublicSubnet01
    RouteTableId: !Ref eksPublicRouteTable

eksPublicSubnet02RouteTableAssociation:
  Type: AWS::EC2::SubnetRouteTableAssociation
  Properties:
    SubnetId: !Ref eksPublicSubnet02
    RouteTableId: !Ref eksPublicRouteTable

eksSecurityGroup:
  Type: AWS::EC2::SecurityGroup
  Properties:
    GroupDescription: Cluster communication with worker nodes
    VpcId: !Ref eksVPC
    Tags:
      - Key: Name
        Value: !Sub "${AWS::StackName}-SecurityGroup"
      - Key: Project
        Value: aws-eks

eksIAMRole:
  Type: AWS::IAM::Role
  Properties:
    AssumeRolePolicyDocument:
      Version: '2012-10-17'
      Statement:
        - Effect: Allow
          Principal:
            Service:
              - eks.amazonaws.com
          Action:
            - 'sts:AssumeRole'
    RoleName: !Ref EKSIAMRoleName
    ManagedPolicyArns:
      - arn:aws:iam::aws:policy/AmazonEKSClusterPolicy

eksCluster:
  Type: AWS::EKS::Cluster
  Properties:
    Name: !Ref EKSClusterName
    Version: !Ref EKSVersion
    RoleArn:
      "Fn::GetAtt": ["eksIAMRole", "Arn"]
    ResourcesVpcConfig:

```

(continues on next page)

(continued from previous page)

```

    SecurityGroupIds:
      - !Ref eksSecurityGroup
    SubnetIds:
      - !Ref eksPublicSubnet01
      - !Ref eksPublicSubnet02
    DependsOn: [eksIAMRole, eksPublicSubnet01, eksPublicSubnet02, eksSecurityGroup]

eksNodeInstanceRole:
  Type: AWS::IAM::Role
  Properties:
    AssumeRolePolicyDocument:
      Version: "2012-10-17"
      Statement:
        - Effect: Allow
          Principal:
            Service:
              - !FindInMap [ServicePrincipals, !Ref "AWS::Partition", ec2]
          Action:
            - "sts:AssumeRole"
    ManagedPolicyArns:
      - !Sub "arn:${AWS::Partition}:iam::aws:policy/AmazonEKSWorkerNodePolicy"
      - !Sub "arn:${AWS::Partition}:iam::aws:policy/AmazonEKS_CNI_Policy"
      - !Sub "arn:${AWS::Partition}:iam::aws:policy/
↪AmazonEC2ContainerRegistryReadOnly"
    Path: /

eksNodeGroup:
  Type: AWS::EKS::Nodegroup
  Properties:
    ClusterName: !Ref EKSClusterName
    NodeRole:
      "Fn::GetAtt": ["eksNodeInstanceRole", "Arn"]
    AmiType: AL2_x86_64
    InstanceTypes:
      - !Ref EKSWorkerNodeInstanceType
    NodegroupName: !Ref EKSNodeGroupName
    RemoteAccess:
      Ec2SshKey: !Ref EKSKeyPair
    ScalingConfig:
      MinSize: !Ref NodeAutoScalingGroupMinSize
      DesiredSize: !Ref NodeAutoScalingGroupDesiredCapacity
      MaxSize: !Ref NodeAutoScalingGroupMaxSize
    Labels:
      Project: aws-eks
    Subnets:
      - !Ref eksPublicSubnet01
      - !Ref eksPublicSubnet02
    DependsOn: [eksCluster, eksNodeInstanceRole]

Outputs:
  SubnetIds:
    Description: Subnets IDs in the eksVPC
    Value: !Join [ ",", [ !Ref eksPublicSubnet01, !Ref eksPublicSubnet02 ] ]
  SecurityGroups:
    Description: Security group for the cluster control plane communication with_
↪worker nodes
    Value: !Join [ ",", [ !Ref eksSecurityGroup ] ]
  VpcId:

```

(continues on next page)

(continued from previous page)

Description: The eksVPC Id
Value: `!Ref eksVPC`

Parameters

- `EKSClusterName (String)` - The desired name of your AWS EKS Cluster.
- `EKSVersion (String)` - The desired version of your AWS EKS Cluster. (1.16, 1.17, 1.18, 1.21 - Default: 1.21)
- `EKSNodeGroupName (String)` - The desired name of your AWS EKS Node Group. (Default: Node-Group01)
- `NodeAutoScalingGroupDesiredCapacity (Number)` - Number of desired worker nodes. (Default: 2)
- `NodeAutoScalingGroupMinSize (Number)` - Minimum size of Node Group ASG. (Default: 1)
- `NodeAutoScalingGroupMaxSize (Number)` - Maximum size of Node Group ASG. Set to at least 1 greater than `NodeAutoScalingGroupDesiredCapacity`. (Default: 5)
- `EKSWorkerNodeInstanceType (String)` - EC2 instance type for the node instances. (Default: t2.large)
- `EKSRoleName (String)` - Description: The name of the IAM role for the EKS service to assume. (Default: EKSClusterRole)
- `EKSKeyPair (AWS::EC2::KeyPair::KeyName)` - The name of Key Pair to establish connection with Worker Node. (Default: "devopskey")
- `VpcBlock (String)` - The CIDR range for the VPC. This should be a valid private (RFC 1918) CIDR range. (Default: 10.0.0.0/16)
- `PublicSubnet01Block (String)` - CidrBlock for public subnet 01 within the VPC. (Default: 10.0.0.0/24)
- `PublicSubnet02Block (String)` - CidrBlock for public subnet 02 within the VPC. (Default: 10.0.1.0/24)
- `AvailabilityZonePublicSubnet01 (CommaDelimitedList<AWS::EC2::AvailabilityZone::Name>)` - Availability Zone for the Public Subnet 01. (Default: us-east-1a)
- `AvailabilityZonePublicSubnet02 (CommaDelimitedList<AWS::EC2::AvailabilityZone::Name>)` - Availability Zone for the Public Subnet 02. (Default: us-east-1b)

Parameters File

Create a `parameters.json` file for use when deploying the CloudFormation template:

CloudFormation template

```
[
  {
    "ParameterKey": "NodeAutoScalingGroupDesiredCapacity",
    "ParameterValue": "2"
  },
  {
    "ParameterKey": "EKSSNodeGroupName",
    "ParameterValue": "NodeGroup01"
  },
  {
    "ParameterKey": "NodeAutoScalingGroupMinSize",
    "ParameterValue": "1"
  },
  {
    "ParameterKey": "NodeAutoScalingGroupMaxSize",
    "ParameterValue": "5"
  },
  {
    "ParameterKey": "EKSSWorkerNodeInstanceType",
    "ParameterValue": "t2.large"
  },
  {
    "ParameterKey": "EKSClusterName",
    "ParameterValue": "l7eksdevops"
  },
  {
    "ParameterKey": "EKSKeyPair",
    "ParameterValue": "devopskey"
  },
  {
    "ParameterKey": "EKSIAMRoleName",
    "ParameterValue": "EKSClusterRoleLatest"
  }
]
```

Provisioning

Create a profile for use with AWS CLI:

```
~$ aws configure --profile <profile>
AWS Access Key ID [None]: AKIAIOSFODNN7EXAMPLE
AWS Secret Access Key [None]: wJalrXUtnFEMI/K7MDENG/bPxRfiCYEXAMPLEKEY
Default region name [us-east-1]:
Default output format [None]:
```

Switch to the correct AWS profile for current shell session:

```
~$ export AWS_PROFILE=<profile>
```

Deploy CloudFormation stack:

```
$ aws cloudformation deploy \
  --stack-name l7esp-example \
```

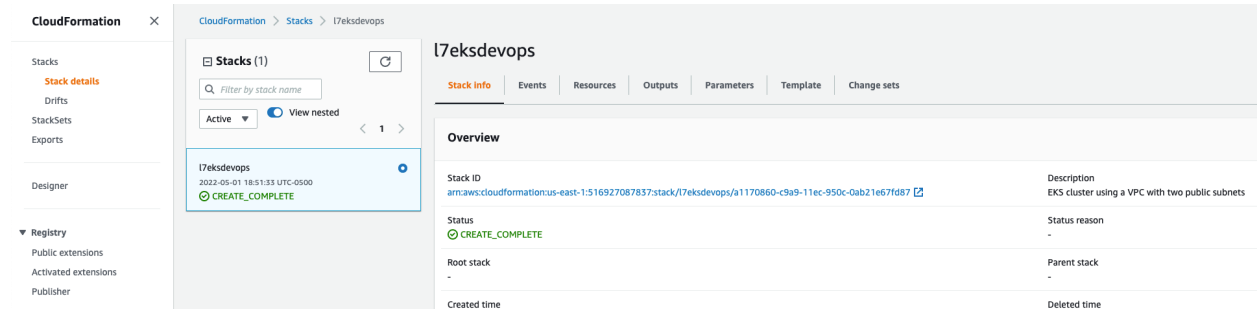
(continues on next page)

(continued from previous page)

```
--template-file eks.yml \
--parameter-overrides file://eks.parameters.json \
--capabilities CAPABILITY_NAMED_IAM
```

Validate Resource Creation

Log into the AWS Console and navigate to CloudFormation. Under “Stacks”, validate that the CloudFormation stack was created:



The screenshot shows the AWS CloudFormation console. On the left, the 'Stacks' section is active, displaying a list of stacks. The 'l7eksdevops' stack is highlighted, showing its status as 'CREATE_COMPLETE'. The main panel shows the 'Overview' tab for the 'l7eksdevops' stack. The overview table contains the following information:

Stack ID	Description	Status	Status reason	Root stack	Parent stack	Created time	Deleted time
arn:aws:cloudformation:us-east-1:1516927087837:stack/l7eksdevops/a1170860-c9a9-11ec-950c-0ab21e67fd87	EKS cluster using a VPC with two public subnets	CREATE_COMPLETE	-	-	-	-	-

Access the Kubernetes cluster

Use the AWS CLI to download the Kubernetes credentials into your local config:

```
~$ aws eks update-kubeconfig --name <eks_cluster_name>
Updated context arn:aws:eks:us-east-1:123456789012:cluster/<eks_cluster_name> in /
~$ cd ~/<user>/<user>/.kube/config
```

Check that Kubernetes nodes were successfully provisioned and are healthy:

```
~$ kubectl get nodes
NAME                                STATUS    ROLES    AGE    VERSION
ip-10-0-0-123.ec2.internal          Ready     agent    1h     v1.21.5-eks-9017834
ip-10-0-1-123.ec2.internal          Ready     agent    1h     v1.21.5-eks-9017834
```

You should see a list of nodes with a **Ready** status.

Installing L7|ESP Helm chart

To install L7|ESP on the new Kubernetes cluster, see the *Helm deployment guide*.



Azure Kubernetes Service

Prerequisites

The L7|ESP Helm chart will be installed on a managed cloud-native Kubernetes service, therefore it is imperative to have:

- Azure Subscription
- Azure Service Principal Account

CLI tools

The `kubectl` and `az` command line tools will be use respectively to programmatically access the Kubernetes cluster and the Azure subscription. To install the CLI tools, follow the link below and choose the correct operating sytem:

- [Kubernetes CLI](#)
- [Azure CLI](#)

ARM template

The latest ARM template is as follows:

Bicep

```

1  @description('The location of AKS resource.')
2  param azureLocation string = resourceGroup().location
3
4  @description('The name of the Managed Cluster resource.')
5  param clusterName string = 'l7esp-example'
6
7  @description('The number of nodes for the cluster. 1 Node is enough for Dev/Test and
8  ↳minimum 3 nodes, is recommended for Production')
9  @minValue(1)
10 @maxValue(100)
11 param clusterNodeCount int = 3
12
13 @description('The size of the Virtual Machine.')
14 param clusterNodeSize string = 'Standard_D4_v4'
15
16 @description('Disk size (in GiB) to provision for each of the agent pool nodes. This
17 ↳value ranges from 0 to 1023. Specifying 0 will apply the default disk size for that
18 ↳agentVMSize.')
19 @minValue(0)
20 @maxValue(1023)
21 param clusterNodeDiskGB int = 0
22
23 @description('Optional DNS prefix to use with hosted Kubernetes API server FQDN.')
24 param clusterDNSPrefix string = 'l7esp-example'
25
26 resource cluster 'Microsoft.ContainerService/managedClusters@2020-09-01' = {
27   location: azureLocation
28   name: clusterName
29   tags: {

```

(continues on next page)

(continued from previous page)

```

27     displayname: 'AKS Cluster'
28   }
29   identity: {
30     type: 'SystemAssigned'
31   }
32   properties: {
33     dnsPrefix: clusterDNSPrefix
34     agentPoolProfiles: [
35       {
36         name: 'agentpool'
37         osDiskSizeGB: clusterNodeDiskGB
38         count: clusterNodeCount
39         vmSize: clusterNodeSize
40         osType: 'Linux'
41         type: 'VirtualMachineScaleSets'
42         mode: 'System'
43       }
44     ]
45   }
46 }
47
48 output clusterFQDN string = cluster.properties.fqdn

```

JSON

```

1  {
2    "$schema": "https://schema.management.azure.com/schemas/2019-04-01/
↪deploymentTemplate.json#",
3    "contentVersion": "1.0.0.0",
4    "metadata": {
5      "_generator": {
6        "name": "bicep",
7        "version": "0.6.18.56646",
8        "templateHash": "16874195123538177185"
9      }
10   },
11   "parameters": {
12     "azureLocation": {
13       "type": "string",
14       "defaultValue": "[resourceGroup().location]",
15       "metadata": {
16         "description": "The location of AKS resource."
17       }
18     },
19     "clusterName": {
20       "type": "string",
21       "defaultValue": "l7esp-example",
22       "metadata": {
23         "description": "The name of the Managed Cluster resource."
24       }
25     },
26     "clusterNodeCount": {
27       "type": "int",
28       "defaultValue": 3,
29       "maxValue": 100,

```

(continues on next page)

(continued from previous page)

```

30     "minValue": 1,
31     "metadata": {
32         "description": "The number of nodes for the cluster. 1 Node is enough for Dev/
↪Test and minimum 3 nodes, is recommended for Production"
33     }
34 },
35     "clusterNodeSize": {
36         "type": "string",
37         "defaultValue": "Standard_D4_v4",
38         "metadata": {
39             "description": "The size of the Virtual Machine."
40         }
41     },
42     "clusterNodeDiskGB": {
43         "type": "int",
44         "defaultValue": 0,
45         "maxValue": 1023,
46         "minValue": 0,
47         "metadata": {
48             "description": "Disk size (in GiB) to provision for each of the agent pool
↪nodes. This value ranges from 0 to 1023. Specifying 0 will apply the default disk
↪size for that agentVMSize."
49         }
50     },
51     "clusterDNSPrefix": {
52         "type": "string",
53         "defaultValue": "l7esp-example",
54         "metadata": {
55             "description": "Optional DNS prefix to use with hosted Kubernetes API server.
↪FQDN."
56         }
57     },
58 },
59 "resources": [
60     {
61         "type": "Microsoft.ContainerService/managedClusters",
62         "apiVersion": "2020-09-01",
63         "name": "[parameters('clusterName')]",
64         "location": "[parameters('azureLocation')]",
65         "tags": {
66             "displayname": "AKS Cluster"
67         },
68         "identity": {
69             "type": "SystemAssigned"
70         },
71         "properties": {
72             "dnsPrefix": "[parameters('clusterDNSPrefix')]",
73             "agentPoolProfiles": [
74                 {
75                     "name": "agentpool",
76                     "osDiskSizeGB": "[parameters('clusterNodeDiskGB')]",
77                     "count": "[parameters('clusterNodeCount')]",
78                     "vmSize": "[parameters('clusterNodeSize')]",
79                     "osType": "Linux",
80                     "type": "VirtualMachineScaleSets",
81                     "mode": "System"
82                 }

```

(continues on next page)

(continued from previous page)

```

83     ]
84   }
85 }
86 ],
87 "outputs": {
88   "clusterFQDN": {
89     "type": "string",
90     "value": "[reference(resourceId('Microsoft.ContainerService/managedClusters',
↳parameters('clusterName'))).fqdn]"
91   }
92 }
93 }

```

ARM template parameters

When deploying the ARM template, you can override parameters like so:

```

1  {
2    "$schema": "https://schema.management.azure.com/schemas/2015-01-01/
↳deploymentParameters.json#",
3    "contentVersion": "1.0.0.0",
4    "parameters": {
5      "agentCount": {
6        "value": 4
7      },
8      "agentVMSize": {
9        "value": "standard_d11_v2"
10     }
11   }
12 }

```

- `azureLocation`: The location of AKS resource. Default value: same region/location as the resource group you are deploying ARM template into.
- `clusterDNSPrefix`: Optional DNS prefix to use with hosted Kubernetes API server FQDN. Default value: `aks-esp`
- `clusterName`: The name of the Managed Cluster resource. Default value: `aks101cluster-vmss`
- `clusterNodeCount`: The number of nodes for the cluster. One node is enough for Dev/Test and minimum 3 nodes is recommended for Production. Default: 3
- `clusterNodeDiskGB`: Disk size (in GiB) to provision for each of the agent pool nodes. This value ranges from 0 to 1023. Specifying 0 will apply the default disk size for that `agentVMSize`. Default value: 0
- `clusterNodeSize`: The size of the virtual machine. Default: `Standard_D4_v4`

Deploy ARM template

Login to Azure. For example, if you are using a service principle account that has Contributor role, you can login to Azure using the following command:

```
$ az login \
  --service-principal \
  --username <service-principal-id> \
  --password <service-principal-password> \
  --tenant <tenant-id>
```

Create a resource group to deploy into, if you haven't already:

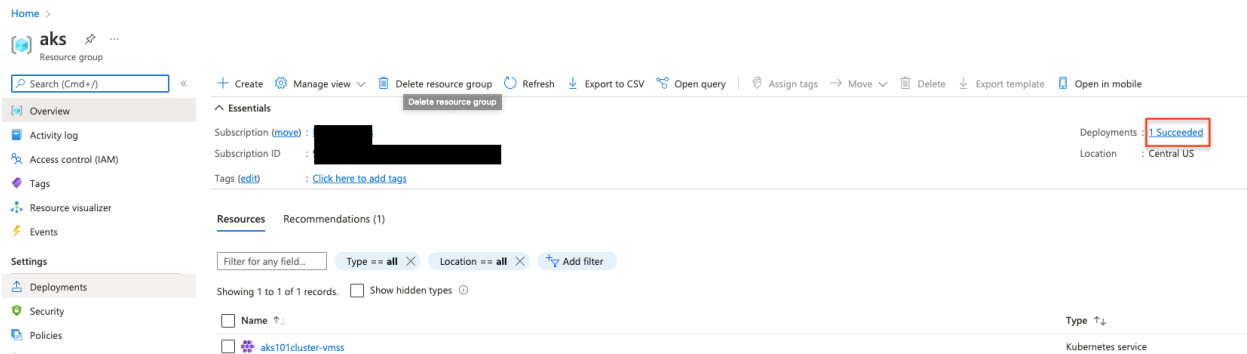
```
$ az group create \
  --name <resource-group-name> \
  --location <resource-group-location>
```

Create ARM deployment in the resource group:

```
$ az group deployment create \
  --resource-group l7esp-example-rg \
  --name l7esp-example \
  --template-file ./aks.template.json \
  --parameters ./aks.parameters.json \
  --rollback-on-error \
  --verbose
```

Validate Resource Creation

In the Azure portal, navigate to the resource group and verify that all resources are listed:



Access the Kubernetes cluster

Log into your Azure Subscription to access cluster:

```
$ az login
```

Use the Azure CLI to download the Kubernetes credentials into your local config:

```
$ az aks get-credentials \
  --name <cluster-name> \
```

(continues on next page)

(continued from previous page)

```
--resource-group <cluster-resource>  
Merged "<my-cluster>" as current context in /home/<user>/.kube/config
```

Check that Kubernetes nodes were successfully provisioned and are healthy:

```
$ kubectl get nodes  
NAME                                STATUS    ROLES    AGE    VERSION  
aks-agentpool-21130059-vmss000000  Ready    agent    1h     v1.22.6  
aks-agentpool-21130059-vmss000001  Ready    agent    1h     v1.22.6  
aks-agentpool-21130059-vmss000002  Ready    agent    1h     v1.22.6  
aks-agentpool-21130059-vmss000003  Ready    agent    1h     v1.22.6
```

You should see a list of nodes with a **Ready** status.

Installing L7|ESP Helm chart

To install L7|ESP on the new Kubernetes cluster, see the [Helm deployment guide](#).



Canonical MicroK8s

MicroK8s is the simplest production-grade conformant K8s. Lightweight and focused. Single command install on Linux, Windows and macOS.

Prerequisites

The L7|ESP Helm chart can be installed onto a multi-node MicroK8s cluster running on traditional VMs:

- A fleet of virtual machines (e.g. Ubuntu) to convert into MicroK8s nodes

Requirements (Amazon EC2)

To install the L7|ESP Helm chart on Amazon EC2 Instance, you must have:

- An AWS account to deploy to
- Access keys for the AWS account (see [Managing access keys for IAM users](#))

aws command line tool will be use to programmatically access the AWS account. To install the AWS CLI tool, follow the link below and choose the correct operating sytem:

- [AWS CLI](#)

Infrastructure as Code

CloudFormation Template

The latest CloudFormation template is as follows:

CloudFormation template

```
AWSTemplateFormatVersion: 2010-09-09
Parameters:
  InstanceTypeParameter:
    Type: String
    Default: t3a.medium
    Description: Enter instance size. Default is t3a.medium.
  SSHLocation:
    Description: The IP address range that can be used to SSH to the EC2 instances
    Type: String
    MinLength: '9'
    MaxLength: '18'
    Default: 0.0.0.0/0
    AllowedPattern: '(\d{1,3})\.\d{1,3})\.\d{1,3})\.\d{1,3})/(\d{1,2})'
    ConstraintDescription: Must be a valid IP CIDR range of the form x.x.x.x/x
  AMI:
    Type: String
    Default: ami-08c40ec9ead489470
    Description: The Linux AMI to use.
  Key:
    Default: l7esp-ec2-linux
    Description: Name of an existing EC2 KeyPair to enable SSH access to the
↪instance
    Type: String
  VPC:
    Type: String
    Default: vpc-0278eb63cfc078a74
    Description: The key used to access the instance.
Resources:
  InstanceSecurityGroup:
    Type: AWS::EC2::SecurityGroup
    Properties:
      GroupName: "Internet Group"
      GroupDescription: "SSH and web traffic in, all traffic out."
      VpcId: !Ref VPC
      SecurityGroupIngress:
        - IpProtocol: tcp
          FromPort: '22'
          ToPort: '22'
          CidrIp: !Ref SSHLocation
      SecurityGroupEgress:
        - IpProtocol: -1
          CidrIp: 0.0.0.0/0
  Linux:
    Type: 'AWS::EC2::Instance'
    Properties:
      ImageId: !Ref AMI
      InstanceType:
```

(continues on next page)

(continued from previous page)

```

    Ref: InstanceTypeParameter
KeyName: !Ref Key
SecurityGroupIds:
  - Ref: InstanceSecurityGroup
BlockDeviceMappings:
  - DeviceName: /dev/sda1
    Ebs:
      VolumeSize: 100
Tags:
  -
    Key: Appplication
    Value: Linux Server
  -
    Key: Domain
    Value: None
  -
    Key: Environment
    Value: Test
  -
    Key: LifeTime
    Value: Transient
  -
    Key: Name
    Value: l7esp-ec2-linux-MicroK8s
  -
    Key: OS
    Value: Linux
  -
    Key: Purpose
    Value: Support Test Instance
Outputs:
  PublicIp:
    Value:
      Fn::GetAtt:
        - Linux
        - PublicIp
    Description: Server's PublicIp Address

```

Parameters

- InstanceTypeParameter (String) - The desired name of your AWS EKS Cluster.
- SSHLocation (String) - The IP address range that can be used to SSH to the EC2 instances. (Default: 0.0.0.0/0)
- AMI (String) - The desired AMI ID (Here are few suggested AMI's [AMI Amazon Linux \(Recommended\)](#), [Debian Stable](#), [Ubuntu LTS](#) releases)
- Key (String) - This for Tag with Keyan Name as required for your AWS EC2 Instance
- VPC (String) - This is a VPC ID where you want to provision EC2 Instance

Parameters File

Create a `parameters.json` file for use when deploying the CloudFormation template:

CloudFormation template

```
[
  {
    "ParameterKey": "InstanceTypeParameter",
    "ParameterValue": "c6g.4xlarge"
  },
  {
    "ParameterKey": "SSHLocation",
    "ParameterValue": "0.0.0.0/0"
  },
  {
    "ParameterKey": "AMI",
    "ParameterValue": "ami-0f69dd1d0d03ad669"
  },
  {
    "ParameterKey": "Key",
    "ParameterValue": "l7esp-ec2-linux"
  },
  {
    "ParameterKey": "VPC",
    "ParameterValue": "vpc-0278eb63cfc078a74"
  }
]
```

Provisioning

Create a profile for use with AWS CLI:

```
~$ aws configure --profile <profile>
AWS Access Key ID [None]: AKIAIOSFODNN7EXAMPLE
AWS Secret Access Key [None]: wJalrXUtnFEMI/K7MDENG/bPxRfiCYEXAMPLEKEY
Default region name [us-east-1]:
Default output format [None]:
```

Switch to the correct AWS profile for current shell session:

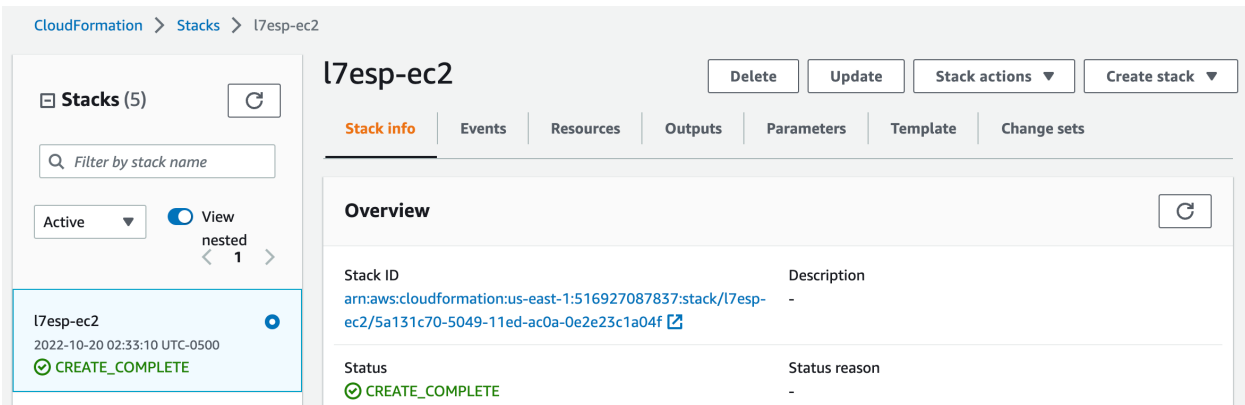
```
~$ export AWS_PROFILE=<profile>
```

Deploy CloudFormation stack:

```
$ aws cloudformation deploy \
  --stack-name l7esp-ec2 \
  --template-file amazon_linux_ec2.yml \
  --parameter-overrides file://ec2.parameters.json \
  --capabilities CAPABILITY_NAMED_IAM
```

Validate Resource Creation

Log into the AWS Console and navigate to CloudFormation. Under “Stacks”, validate that the CloudFormation stack was created:



Install MicroK8s

The installation steps for Ubuntu are, roughly, as follows:

```
$ sudo apt-get update
$ sudo apt-get upgrade -y
$ sudo snap install core
$ sudo snap install microk8s --classic --channel=1.25/stable
$ microk8s enable dns
$ microk8s enable hostpath-storage
$ sudo microk8s status --wait-ready
$ sudo usermod -a -G microk8s l7esp
$ sudo chown -f -R l7esp ~/.kube
$ echo -e "l7esp\nl7esp" | passwd root
$ echo "alias kubectl='microk8s kubectl'" >> ~/.bash_rc
```

Note: Please refer to the official documentation for latest advice on how to install MicroK8s. The steps above are merely for demonstration purposes. [Install a local Kubernetes with MicroK8s](#)

Access the Kubernetes cluster

SSH into one of the cluster’s master node and check that all nodes were provisioned and are in **Ready** state, with the following command:

```
$ microk8s kubectl get nodes
```

NAME	STATUS	ROLES	AGE	VERSION
k8s-master	Ready	<none>	1h	v1.20.13-35+d877e7a8ac536e
k8s-node-1	Ready	<none>	1h	v1.20.13-35+d877e7a8ac536e
k8s-node-2	Ready	<none>	1h	v1.20.13-35+d877e7a8ac536e

Installing L7|ESP Helm chart

To install L7|ESP on the new Kubernetes cluster, see the *Helm deployment guide*.

Note: MicroK8s comes with its own packaged version of the Kubernetes CLI so remember to use the `microk8s kubectl` to ensure you are using the correct version when controlling the cluster.



Helm

Helm is a package manager for Kubernetes, and “charts” are the packages they install.

1. Confirm that you are connected to the correct Kubernetes cluster:

```
$ kubectl config use-context <cluster-name>
Switched to context "l7esp-example".
$ kubectl config current-context
l7esp-example
```

Note: If you do not have a *Kubernetes* cluster yet, please do that before proceeding.

2. Install Helm as documented on their website: <https://helm.sh/docs/intro/install/>
3. Deploying the L7|ESP server to a Kubernetes cluster is performed using a Helm chart.

```
$ helm dependency update ./charts/l7esp-server
$ helm upgrade \
  --atomic \
  --create-namespace \
  --install \
  --namespace l7esp-example \
  --reset-values \
  --timeout 10m \
  --values ./values.yaml \
  l7esp-example \
  ./charts/l7esp-server
```

Note: You must provide an appropriate `values.yaml` file. All the configurable options are documented in the Helm chart `README.md`.

4. Validate L7|ESP server was installed by getting a list of all deployed Kubernetes resources:

```
$ kubectl get all --namespace <l7esp-namespace>
```

NAME	READY	STATUS	RESTARTS	AGE
pod/l7esp-example-56595b4b8c-gmt98	1/1	Running	0	1h

NAME	TYPE	CLUSTER-IP	EXTERNAL-IP	PORT(S)
↪ service/l7esp-example	ClusterIP	10.0.123.123	<none>	8002/TCP
↪ 1h				

NAME	READY	UP-TO-DATE	AVAILABLE	AGE
deployment.apps/l7esp-example	1/1	1	0	1h

NAME	DESIRED	CURRENT	READY	AGE
replicaset.apps/l7esp-example-56595b4b8c	1	1	0	1h

5. If you want to see the logs of the L7|ESP server startup, you can do so by running:

```
$ kubectl logs --follow --namespace <l7esp-namespace> pod/l7esp-example-
↪56595b4b8c-gmt98
```

6. Forward the L7|ESP server port to your local machine:

```
$ kubectl port-forward --namespace <l7esp-namespace> svc/l7esp-example-
↪8002:8002
```

7. Verify that you can access the L7|ESP server's web UI in your browser by following this link:

localhost:8002

Setting up ingress

Ingress is a Kubernetes extension that allows you to expose services to the internet. The L7|ESP Helm chart does not install an Ingress controller, but you can install one yourself. A popular Ingress controller is the NGINX Ingress controller, which we will use in the following example:

1. Deploy the Nginx ingress controller:

```
helm upgrade \
  --atomic \
  --install ingress-nginx ingress-nginx \
  --repo https://kubernetes.github.io/ingress-nginx \
  --namespace ingress-nginx \
  --create-namespace
```

Note: For more information see the Nginx Ingress controller's website: <https://kubernetes.github.io/ingress-nginx/deploy/>

2. Update L7|ESP Helm chart values to apply any annotations required by the Ingress controller.

```
$ helm upgrade \
  --atomic \
  --create-namespace \
  --install \
  --namespace l7esp-example \
```

(continues on next page)

(continued from previous page)

```
--reset-values \
--values ./values.yaml \
l7esp-example \
./charts/l7esp-server
```

Note: See the `values.yaml` file in the L7|ESP Helm chart for more information.

3. Verify that the ingress controller is working by visiting the following URL:

<http://l7esp.example.com>

Note: If you are using a different domain name, you will need to update the ingress controller to use that domain name.

Installing L7|ESP Content on kubernetes

The following steps are to install your L7|ESP environment's project configuration and content, provided to you by the L7 Informatics implementation team.

First, scale down your deployment to one replica if you have more than one esp running using the following command (skip this step and last step if there is only one replica)

```
$ kubectl scale --current-replicas=<current_number> --replicas=1 deployment/<esp_
↳deployment name>
```

then, get a shell into the L7|ESP pod using the following command:

```
$ kubectl exec -it <pod name> -c <container-id> --namespace l7esp-example -- bash
```

Next, extract the L7|ESP 3.x deployment bundle tarball and change into the directory it unpacked (where `docker-compose.yml` resides) before running the rest of the commands:

```
$ cd /opt/l7esp/data/project
$ wget https://$download_link
$ tar xf $project-$version.tar.gz --strip-components=1
```

Next, issue the the following commands to install the configuration and content:

```
$ make install
```

Finally, scale back to the original replica set

```
$ kubectl scale --current-replicas=1 --replicas=<desire number > deployment/<esp_
↳deployment name>
```



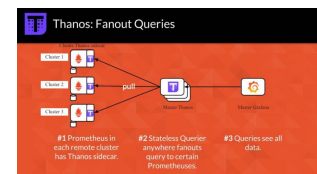
L7ESP K8s Monitoring with Grafana

Introduction

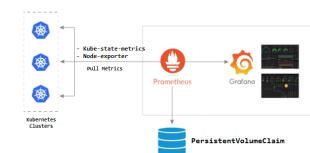
The Document below was created to give instruction to the DevOps support team to help them onboard the monitoring tool to enhance observability across internal assets and customer assets.

Grafana Architecture

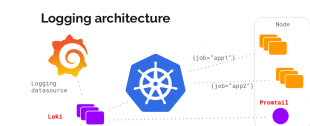
- Grafana architecture (multiple cluster monitoring) will consist of one master thanos deployment at the master cluster with the grafana front end application, this master node will play as the single pane for monitoring. The client cluster will need a deployment of the thanos stack to ship metrics and log to the master cluster. In case the assets to be monitored are not on k8s, we will leverage the numerous data sources grafana provides to fetch metrics.



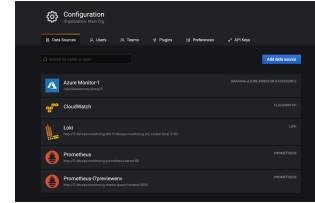
- For a single cluster monitoring you can just leverage a prometheus agent scraping metrics from nodeexporter/kube state metrics on your nodes. Metrics are then shipped to grafana.



- The logging of your application will be scraped and shipped to grafana by loki running jobs with promtail on the nodes to request logs from your k8s application.



- At last grafana can also leverage API data sources to fetch metrics from other assets, such as Azure or AWS cloud environment. This will be done through their data sources.



Requirements

To install the grafana stack, on your kubernetes cluster, you must have

- Kubernetes cluster to deploy your helm charts.

The `kubectl` and `helm` command line tools will be use respectively to programmatically access the Kubernetes cluster and the AWS account. To install the CLI tools, follow the link below and choose the correct operating sytem:

- [Kubernetes CLI](#)
- [Helm CLI](#)

Grafana

The Grafana application can be installed using the helm chart. Download the helm chart on your laptop from the following website: [Grafana helm chart](#) them install it using helm.

Following are values.yaml parameters to managed your dashboards in grafana.

- Automated uploads of new dashboard so user can upload new dashboard.json on the data folder of your confimap. That folder lable should match the label on the values file specified below.

```
dashboards:
  enabled: true
  SCPProvider: true
  # label that the configmaps with dashboards are marked with
  label: grafana_dashboard
```

- enable default dashboard at login by updating the following values parameters.

```
dashboards:
  default_home_dashboard_path: /tmp/dashboards/k8s-dashboard.json
```

- Set initial grafana login username and password

```
grafana:
  enabled: true
  # Pass values to Grafana child chart
  adminUser: admin
  adminPassword: admin
```

Datasource

Datasources are special back-end query API that query data from different assets. Below are a list of currently used data sources by ESP grafana.

- blackboxexporter: A probe that leverages prometheus to check health and metrics from SSL certificates
- thanos: High availability master-client architecture to scrape metrics from other clusters
- prometheus: Time series data queryier for metrics
- aws cloudwatch: Data source to ingest AWS cloudwatch data and logs
- azure monitor: Data source to ingest Azure data and logs
- Loki: Data source to ingest knative application logs.

Install datasources dependency helm charts

Datasources helm charts can be installed from the following websites: * Prometheus * prometheus-blackbox-exporter * Thanos * Loki * promtail

Values parameters configuration for each Datasources

Enable and configure each datasource from the values.yml:

- configure prometheus

```
    url: http://{{ .Release.Name }}-prometheus-server:80
prometheus:
  server:
    extraArgs:
      log.level: debug
      storage.tsdb.min-block-duration: 2h # Don't change this, see docs/components/
↪sidecar.md
      storage.tsdb.max-block-duration: 2h # Don't change this, see docs/components/
↪sidecar.md
    retention: 4h
    service:
      annotations:
        prometheus.io/scrape: "true"
        prometheus.io/port: "9090"
    statefulSet:
      enabled: true
    podAnnotations:
      prometheus.io/scrape: "true"
      prometheus.io/port: "10902"
    affinity:
      podAntiAffinity:
        requiredDuringSchedulingIgnoredDuringExecution:
          - labelSelector:
              matchExpressions:
                - key: app
                  operator: In
                  values:
                    - prometheus
          - key: component
```

(continues on next page)

(continued from previous page)

```

    operator: In
    values:
      - server
    topologyKey: "kubernetes.io/hostname"

```

- configure Thanos

```

    url: http://{ .Release.Name }-thanos-query-frontend:9090
thanos:
  objstoreConfig: |-
    type: s3
    config:
      bucket:
      endpoint:
      access_key:
      secret_key:
      insecure: true
  query:
    stores:
      -
        # - SIDECAR-SERVICE-IP-ADDRESS-2:10901
  bucketweb:
    enabled: true
  compactor:
    enabled: true
  storegateway:
    enabled: true
  ruler:
    enabled: true
    alertmanagers:
      - http://l7-devops-monitoring-prometheus-alertmanager.grafana.svc.cluster.
↪local:9093
    config: |-
      groups:
        - name: "metamonitoring"
          rules:
            - alert: "PrometheusDown"
              expr: absent (up{prometheus="monitoring/prometheus-operator"})

```

- configure prometheus-blackbox

```

prometheus.yml:
  scrape_configs:
    - job_name: "healthchecks"
      scrape_interval: 60s
      scheme: https
      metrics_path: projects/446fab8e-5a9f-4f41-a90b-f086581f64a5/metrics/
↪lthdUEmLy2xTibPH5Jngrm_BPMGopHT7
      static_configs:
        - targets: ["healthchecks.io"]

    - job_name: 'kubernetes-pods'
      kubernetes_sd_configs:
        - role: pod

      relabel_configs:
        - source_labels: [__meta_kubernetes_pod_annotation_prometheus_io_scrape]

```

(continues on next page)

(continued from previous page)

```

    action: keep
    regex: true
  - source_labels: [__meta_kubernetes_pod_annotation_prometheus_io_path]
    action: replace
    target_label: __metrics_path__
    regex: (.+)
  - source_labels: [__address__, __meta_kubernetes_pod_annotation_prometheus_
↪io_port]
    action: replace
    regex: (.+):(?:\d+);(\d+)
    replacement: ${1}:${2}
    target_label: __address__
  - action: labelmap
    regex: __meta_kubernetes_pod_label_(.+)

- job_name: "snykmetrics"
  static_configs:
    - targets: ["localhost:9090"]
    - targets: ["localhost:9532"]

- job_name: 'blackbox'
  metrics_path: /probe
  params:
    module: [http_2xx]
  static_configs:
    - targets:
      - https://clinops-stage.preview.l7esp.com
      - https://stage.preview.l7esp.com
      - https://250.preview.l7esp.com
      - https://qa-clinops300.preview.l7esp.com
      - https://preview.l7esp.com
      - https://inspiring-jepsen-5082.edgestack.me
      - https://lab7io.atlassian.net
      - https://scipher-qa.l7esp.com
      - https://cdn.l7esp.com
      - https://lab7io.slack.com
      - https://ci.l7informatics.com
      - https://registry.l7informatics.com
      - https://l7devopstest.azurecr.io

  relabel_configs:
    - source_labels: [__address__]
      target_label: __param_target
    - source_labels: [__param_target]
      target_label: instance
    - target_label: __address__
      replacement: l7-devops-monitoring-prometheus-blackbox-exporter:9115 #↪
↪Blackbox exporter scraping address

```

- configure loki/promtails

```

    url: http://{ .Release.Name }-loki.{ .Release.Name }.svc.cluster.
↪local:3100
loki:
  enabled: true

promtail:
  enabled: true

```

(continues on next page)

(continued from previous page)

```

scrapeConfigs: |
  # See also https://github.com/grafana/loki/blob/master/production/ksonnet/
  ↪promtail/scrape_config.libsonnet for reference

  # Pods with a label 'app.kubernetes.io/name'
  - job_name: kubernetes-pods-app-kubernetes-io-name
    pipeline_stages:
      {{- toYaml .Values.config.snippets.pipelineStages | nindent 4 }}
    kubernetes_sd_configs:
      - role: pod
    relabel_configs:
      - action: replace
        source_labels:
          - __meta_kubernetes_pod_label_app_kubernetes_io_name
        target_label: app
      - action: drop
        regex: ''
        source_labels:
          - app
      - action: replace
        source_labels:
          - __meta_kubernetes_pod_label_app_kubernetes_io_component
        target_label: component
      {{- if .Values.config.snippets.addScrapeJobLabel }}
      - action: replace
        replacement: kubernetes-pods-app-kubernetes-io-name
        target_label: scrape_job
      {{- end }}
      {{- toYaml .Values.config.snippets.common | nindent 4 }}

```

- Azure monitor

```

- name: Azure Monitor-1
  type: grafana-azure-monitor-datasource
  access: proxy
  jsonData:
    azureAuthType: clientsecret
    cloudName: azuremonitor
    tenantId:
    clientId:
    subscriptionId: 17 informatics
  secureJsonData:
    clientSecret:
  version: 1

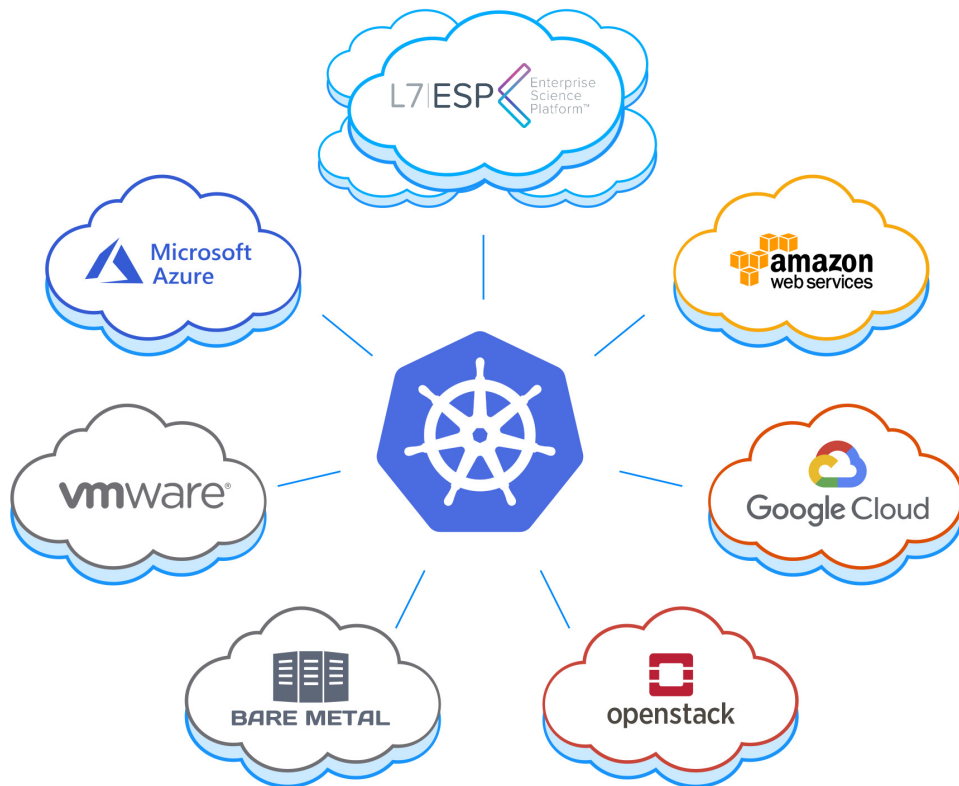
```

- AWS cloudwatch

```

- name: CloudWatch
  type: cloudwatch
  jsonData:
    authType: keys
    defaultRegion: us-east-1a
  secureJsonData:
    accessKey: ''
    secretKey: ''

```



Amazon Elastic Kubernetes Service

Amazon Elastic Kubernetes Service (EKS) is a managed service and certified Kubernetes conformant to run Kubernetes on AWS and on-premises.



Azure Kubernetes Service

Azure Kubernetes Service (AKS) offers serverless Kubernetes, an integrated continuous integration and continuous delivery (CI/CD) experience, and enterprise-grade security and governance.

The logo for MicroK8s, featuring the text "MicroK8s" in white on an orange square background.

Canonical MicroK8s

MicroK8s is the simplest production-grade conformant K8s. Lightweight and focused. Single command install on Linux, Windows and macOS.



Helm

Helm helps you manage Kubernetes applications — Helm Charts help you define, install, and upgrade even the most complex Kubernetes application.



L7ESP K8s Monitoring with Grafana

Grafana allows you to query, visualize, alert on, and understand your metrics no matter where they are stored. Create, explore, and share beautiful dashboards with your team and foster a data-driven culture.

1.15.2 Docker



Azure Container Instances

Prerequisites

Azure Container Instances is a quick way to deploy your container into Docker host within Azure, without having to provision the underlying servers first (VM + OS).

- Azure Subscription
- Azure Service Principal Account

CLI tools

The `az` command line tool will be used to programmatically access your Azure subscription. To install the CLI tool, follow the link below and choose the correct operating system:

- [Azure CLI](#)

Obtain the latest ARM template

The ARM template establishes some default values and general configuration for the ESP deployment. You will leave this template intact and use the parameters file (created in the next step) to fill in your specific details and credentials. The latest ARM template is as follows:

Bicep

```
@description('Name (default: l7esp-example)')
param containerName string = 'l7esp-example'

@description('Azure location to deploy containers (default: Azure Resource Group_
↳location)')
param azureLocation string = resourceGroup().location

@description('L7|ESP software license (optional; JSON string)')
param appLicense string = ''

@description('L7|ESP admin password (required; initial password)')
param appPassword string = 'changeme'

@description('L7|ESP web UI port (default: 8002)')
param appPortHTTP int = 8002

@description('Container Image (e.g. acme/l7esp:3.0.0)')
param containerImage string = ''

@description('Container CPU cores (default: 2 vCPU)')
param containerCPU int = 2

@description('Container Memory (default: 6GB RAM)')
param containerMemory int = 6

@description('Container Restart Policy (default: Never)')
@allowed([
  'Always'
  'Never'
  'OnFailure'
])
param containerRestartPolicy string = 'Never'

@description('Container Registry URL (default: docker.io)')
param registryURL string = ''

@description('Container Registry Username (e.g. Docker Hub user)')
param registryUsername string = ''
```

(continues on next page)

(continued from previous page)

```

@description('Container Registry Password (e.g. Docker Hub app password)')
param registryPassword string = ''

resource container_instance 'Microsoft.ContainerInstance/containerGroups@2021-09-01' _
↳= {
  name: containerName
  location: azureLocation
  properties: {
    containers: [
      {
        name: containerName
        properties: {
          image: containerImage
          ports: [
            {
              port: appPortHTTP
              protocol: 'TCP'
            }
          ]
          environmentVariables: [
            {
              secureValue: appPassword
              name: 'L7ESP_PASSWORD'
            }
            {
              secureValue: appLicense
              name: 'L7ESP_LICENSE'
            }
          ]
          resources: {
            requests: {
              cpu: containerCPU
              memoryInGB: containerMemory
            }
          }
        }
      }
    ]
  }
  osType: 'Linux'
  restartPolicy: containerRestartPolicy
  ipAddress: {
    type: 'Public'
    ports: [
      {
        port: appPortHTTP
        protocol: 'TCP'
      }
    ]
  }
  imageRegistryCredentials: [
    {
      password: registryPassword
      server: registryURL
      username: registryUsername
    }
  ]
}

```

(continues on next page)

(continued from previous page)

```

}

output containerIPv4Address string = container_instance.properties.ipAddress.ip

```

JSON

```

{
  "$schema": "https://schema.management.azure.com/schemas/2019-04-01/
↳deploymentTemplate.json#",
  "contentVersion": "1.0.0.0",
  "metadata": {
    "_generator": {
      "name": "bicep",
      "version": "0.6.18.56646",
      "templateHash": "17739852969827724894"
    }
  },
  "parameters": {
    "containerName": {
      "type": "string",
      "defaultValue": "l7esp-example",
      "metadata": {
        "description": "Name (default: l7esp-example)"
      }
    },
    "azureLocation": {
      "type": "string",
      "defaultValue": "[resourceGroup().location]",
      "metadata": {
        "description": "Azure location to deploy containers (default: Azure Resource_
↳Group location)"
      }
    },
    "appLicense": {
      "type": "string",
      "defaultValue": "",
      "metadata": {
        "description": "L7|ESP software license (optional; JSON string)"
      }
    },
    "appPassword": {
      "type": "string",
      "defaultValue": "changeme",
      "metadata": {
        "description": "L7|ESP admin password (required; initial password)"
      }
    },
    "appPortHTTP": {
      "type": "int",
      "defaultValue": 8002,
      "metadata": {
        "description": "L7|ESP web UI port (default: 8002)"
      }
    },
    "containerImage": {

```

(continues on next page)

(continued from previous page)

```

    "type": "string",
    "defaultValue": "",
    "metadata": {
      "description": "Container Image (e.g. acme/l7esp:3.0.0)"
    }
  },
  "containerCPU": {
    "type": "int",
    "defaultValue": 2,
    "metadata": {
      "description": "Container CPU cores (default: 2 vCPU)"
    }
  },
  "containerMemory": {
    "type": "int",
    "defaultValue": 6,
    "metadata": {
      "description": "Container Memory (default: 6GB RAM)"
    }
  },
  "containerRestartPolicy": {
    "type": "string",
    "defaultValue": "Never",
    "allowedValues": [
      "Always",
      "Never",
      "OnFailure"
    ],
    "metadata": {
      "description": "Container Restart Policy (default: Never)"
    }
  },
  "registryURL": {
    "type": "string",
    "defaultValue": "",
    "metadata": {
      "description": "Container Registry URL (default: docker.io)"
    }
  },
  "registryUsername": {
    "type": "string",
    "defaultValue": "",
    "metadata": {
      "description": "Container Registry Username (e.g. Docker Hub user)"
    }
  },
  "registryPassword": {
    "type": "string",
    "defaultValue": "",
    "metadata": {
      "description": "Container Registry Password (e.g. Docker Hub app password)"
    }
  },
  "resources": [
    {
      "type": "Microsoft.ContainerInstance/containerGroups",

```

(continues on next page)

(continued from previous page)

```

    "apiVersion": "2021-09-01",
    "name": "[parameters('containerName')]",
    "location": "[parameters('azureLocation')]",
    "properties": {
      "containers": [
        {
          "name": "[parameters('containerName')]",
          "properties": {
            "image": "[parameters('containerImage')]",
            "ports": [
              {
                "port": "[parameters('appPortHTTP')]",
                "protocol": "TCP"
              }
            ],
            "environmentVariables": [
              {
                "secureValue": "[parameters('appPassword')]",
                "name": "L7ESP_PASSWORD"
              },
              {
                "secureValue": "[parameters('appLicense')]",
                "name": "L7ESP_LICENSE"
              }
            ],
            "resources": {
              "requests": {
                "cpu": "[parameters('containerCPU')]",
                "memoryInGB": "[parameters('containerMemory')]"
              }
            }
          }
        }
      ],
      "osType": "Linux",
      "restartPolicy": "[parameters('containerRestartPolicy')]",
      "ipAddress": {
        "type": "Public",
        "ports": [
          {
            "port": "[parameters('appPortHTTP')]",
            "protocol": "TCP"
          }
        ]
      },
      "imageRegistryCredentials": [
        {
          "password": "[parameters('registryPassword')]",
          "server": "[parameters('registryURL')]",
          "username": "[parameters('registryUsername')]"
        }
      ]
    },
    "outputs": {
      "containerIPv4Address": {

```

(continues on next page)

(continued from previous page)

```

        "type": "string",
        "value": "[reference(resourceId('Microsoft.ContainerInstance/containerGroups', 'l7esp-example'), parameters('containerName')).ipAddress.ip]"
    }
}
}

```

You can use either the JSON or the Bicep version of the file, but Bicep is a bit easier to read and edit, and it's the Azure-native format. We'll assume you saved the file in your current directory as `arm_template.bicep`.

Create a parameters file with your specific customization

Save this file as `parameters.json` in your current directory.

```

{
  "$schema": "http://schema.management.azure.com/schemas/2015-01-01/
  ↪deploymentParameters.json#",
  "contentVersion": "1.0.0.0",
  "parameters": {
    "appLicense": {
      "value": ""
    },
    "appPassword": {
      "value": ""
    },
    "containerImage": {
      "value": ""
    },
    "appPortHTTP": {
      "value": 8002
    },
    "containerCPU": {
      "value": 2
    },
    "registryUsername": {
      "value": ""
    },
    "registryPassword": {
      "value": ""
    },
    "registryURL": {
      "value": ""
    },
    "containerMemory": {
      "value": 6
    }
  }
}

```

Edit this file to set important values for your instance:

- `containerName`: This name will identify your container in Azure. For this example we will use 'acme'. (default: `l7esp-example`)
- `appPassword`: The admin password on the ESP instance that you are deploying (required; initial password)

- `appLicense` : The text of the entire JSON object representing the ESP license key. The object should be wrapped in single quotes. (required; JSON string)
- `containerImage`: The docker image that your deployed container will be based on. Example: `docker.io/acme/l7esp:3.0.0`
- `registryURL` : The *hostname* of the Docker registry where your image is hosted. The default is 'docker.io' for Docker Hub, but may be located somewhere in Azure like 'examplecompany.azurecr.io' (default: `docker.io`)
- `registryUsername` : The login name for the container registry (required)
- `registryPassword` : The password for the container registry user (required)

Optional parameters that you may want to change, depending on your specific deployment:

- `containerCPU` : Container CPU cores (default: 2 vCPU)
- `containerMemory` : Container Memory (default: 6GB RAM)
- `containerRestartPolicy` - Container Restart Policy (default: Never)
- `azureLocation` - Azure location to deploy containers (default: Azure Resource Group location)

Login to Azure

```
$ az login
```

This command will open a web browser, where you can authenticate to Azure using the credentials for a user with the ability to perform deployments. This brief command should be sufficient for most cases, but if you are using a Service Principal or Tenant, your command may look like this:

```
$ az login \  
  --service-principal \  
  --username <service-principal-id> \  
  --password <service-principal-password> \  
  --tenant <tenant-id>
```

Deploy a new Resource Group

You may already have a Resource Group created for this deployment, but if not, create one now: `az group create --name example-resource-group --location "Central US"`

Run ARM template pipeline to create AKS

Now run the deployment command:

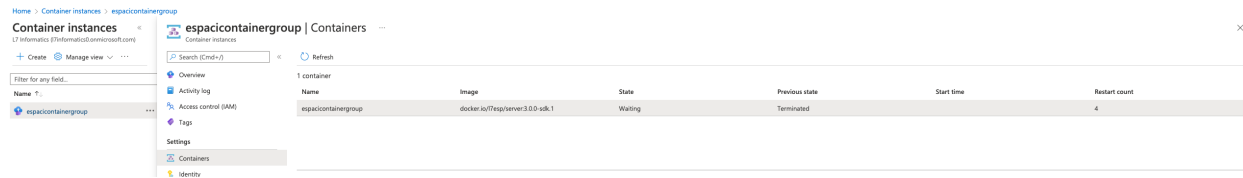
```
$ az deployment group create --resource-group example-resource-group --template-file .  
  ↪ /arm_template.bicep --parameters deployment.json --verbose
```

If you are updating an existing deployment, you may want to add the flag `--rollback-on-error` to this command. However, this flag will cause a new deployment to fail.

Validate Resource Creation

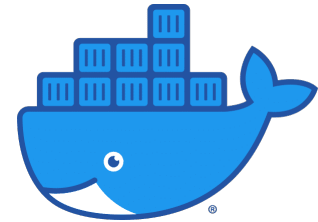
After about 3 minutes, the command will complete and you should see the new container running in the Azure portal.

Navigate to Container Instances and verify that the L7|ESP container instance was created and that it ends up in the **Running** state:



espaccontainergroup Containers						
1 container						
Name	Image	State	Previous state	Start time	Restart count	
espaccontainergroup	docker.io/l7esp/l7esp:3.0.0-adj.1	Waiting	Terminated		4	

ESP will not be available until the build process has completed (about 15 minutes). You can observe the resource usage on the Container Instances detail page. When the memory and CPU usage graphs stabilize and level off with a solid line, you will be able to access the application via the public IP address displayed in the Azure Portal – <http://<public IP address>:8002>



Docker Compose

Introduction

As of 3.0.0, L7|ESP is distributed as a container image. A container runtime, such as Docker, is the only requirement to run L7|ESP on any operating system. Cloud-based PaaS (platform as a service) solutions may also be used to run L7|ESP containers. In addition, container orchestration systems such as Kubernetes or Docker Swarm may be used to manage L7|ESP containers at scale. This particular guide assumes you are installing L7|ESP on a Linux server using the Docker container runtime. If this is not the case, please request documentation for your intended deployment approach.

Configure Linux

The simplest way to get started with L7|ESP is to use a Linux server with your operating system of choice. The first thing you will want to do is create a non-privileged Linux user that will be used to run the L7|ESP container and be the owner of any L7|ESP related files.

Optional L7|ESP User Account:

The following command should create a user called l7esp with home folder /home/l7esp:

```
$ adduser l7esp
```

The following command can be used to switch to this new user at any time:

```
$ sudo su - l7esp
```

Note: To get back to the root or previous, privileged user, simply type exit or press Ctrl+D.

Installing Docker

It is recommended that you follow the official installation documentation for your operating system: <https://docs.docker.com/engine/install/>

At the time of writing, Docker provides a convenience script at <https://get.docker.com/> that we will be using in this guide.

```
$ curl -fsSL https://get.docker.com -o get-docker.sh
$ sudo sh get-docker.sh
```

Follow any on-screen instructions to complete the Docker Engine installation. It is recommended to also review the post-installation instructions:

Optional, you can add the *l7esp* user to the docker group so it can run containers without requiring root or sudo privileges:

```
$ sudo usermod --append --groups docker l7esp
```

Note: For the Linux group change to take effect you will need to log out of your SSH session and back in again.

Installing Docker Compose

Docker Compose simplifies the running of containers by allowing you to specify all the container configuration in a YAML file, instead of running individual docker commands to create, configure and run containers.

It is recommended that you follow the official installation documentation for your operating system: <https://docs.docker.com/compose/install/>

At the time of writing, Docker Compose can be installed on Linux simply by downloading the executable:

```
$ sudo curl -L "https://github.com/docker/compose/releases/download/v2.3.3/docker-
→compose-$(uname -s)-$(uname -m)" -o /usr/local/bin/docker-compose
$ sudo chmod +x /usr/local/bin/docker-compose
```

Note: Replace the version v2.3.3 with the latest version available at the time of installation.

Installing L7|ESP Server

L7|ESP Deployment Bundle

Download and extract the archive provided to you by the L7 Informatics implementation team using roughly the following commands, but with the correct values where the dollar-signs are:

```
$ wget https://$download_link
$ tar xf $project-$version.tar.gz
```

L7|ESP Container Configuration

In the Deployment Bundle, there will be a Docker Compose configuration file that looks similar to the following:

```
$ cat docker-compose.yml
version: "3.9"
services:
  server:
    image: l7esp/server:3.0.0-rc.1
    environment:
      L7ESP_PASSWORD: "${L7ESP_PASSWORD?Set password in .env file}"
      L7ESP_LICENSE_FILE: "/opt/l7esp/data/project/conf/esp.license"
    ports:
      - published: "${L7ESP_PORT?Set port in .env file}"
        target: 8002
        protocol: tcp
        mode: host
    volumes:
      - type: volume
        source: data
        target: /opt/l7esp/data
      - type: bind
        source: .
        target: /opt/l7esp/data/project
volumes:
  data:
```

Things to note in the above configuration:

- The `L7ESP_PASSWORD` and `L7ESP_PORT` environment variables are required and will be read from the `.env` file automatically. Make sure the `.env` file contains the initial administrator password you wish to use and that the port is set to 8002.
- The exact image name will be provided to you by L7 Informatics who will be able to set you up with access to pull the required image. This will require you to run the docker login command on the server. (Alternatively, if you have already downloaded an archive containing the image to use, you may use the docker load command.)
- You should replace `L7ESP_PASSWORD` environment variable value with the initial administrator password you wish to use. This will be the `admin@localhost` user unless you override the username with the `L7ESP_USER` environment variable.
- The published port 8002 will be the publicly exposed port that the L7|ESP web interface will be accessible on. It is recommended to set up a proxy server that provides TLS offloading before forwarding traffic to L7|ESP.
- The data volume is where L7|ESP will persist all data created at runtime and should be a part of your backup plan. You can determine the exact location of this volume by using the docker inspect command on the running container.
- The project bind-mount source path should point to the directory that contains your L7|ESP environment's project configuration and content. This will be provided to you by L7 Informatics either as a tarball that can be extracted or as a Git repository that can be cloned, depending on the use-case.

If you wish to make changes to the `docker-compose.yml` file, please communicate these changes with the L7 Informatics implementation team so they can incorporate these into the next archive you receive, such that the changes are permanent and well tested. Alternatively, you may create your own copy of the `docker-compose.yml` file in a different directory if you wish to maintain this yourself, provided you adjust the project bind-mount source path to point to the archive directory.

Starting L7|ESP

From the same directory as the Docker Compose configuration file (*docker-compose.yml*), run the following command to start the L7|ESP container in the background:

```
$ docker-compose up --detach
```

Next, run the following command to watch the logs as the container starts to ensure there are no errors:

```
$ docker-compose logs --follow
```

Once the log output has stopped, you can issue the following command to check if the L7|ESP web interface is healthy:

```
$ curl localhost:8002/health  
PASS
```

Finally, visit L7|ESP in your web browser of choice (Chrome is recommended):

http://SERVER_IP:8002

You should now be able to log in with the following credentials:

- Username: *admin@localhost* (or L7ESP_USERNAME value)
- Password: *Password12!* (or L7ESP_PASSWORD value)

Installing L7|ESP Content

So far, you should have the core L7|ESP server software running successfully. However, you will likely notice that the software is not licensed or otherwise configured correctly at this point.

The next step is to install your L7|ESP environment's project configuration and content, provided to you by the L7 Informatics implementation team.

First, get a shell into the L7|ESP container using the following command:

```
$ docker-compose exec server bash
```

Next, issue the the following commands to install the configuration and content:

```
$ make install
```

You will now see the a series of installation logs as the L7|ESP SDK, written mostly in Ansible, performs a series of task, including but not limited to:

Synchronize source files with the corresponding target directories in the container
Install your L7|ESP software license
Install any curated L7|Hub content bundles
Import custom configuration into the L7|ESP Config app
Import custom content into the L7|ESP Builders section
Hook any server extensions your configuration requires and reload services

Appendix

Migrating Data

If you are upgrading from an L7|ESP 2.x instance that was not containerized or otherwise need to restore the data volume from a backup, you can start the L7|ESP container while overriding the entrypoint command. This will disable the initialization process that usually runs, allowing you to perform the restoration.

Note: Replace *server_base* with installation directory where the previous L7|ESP server resides, which can be found by running the following command from the 2.x tarball directory:

```
$ grep server_base roles/prod.yml
```

- First, stop the existing L7|ESP 2.x server that you are planning on migration to containerized L7|ESP 3.x platform by running the following l7 stop command:

```
$ /server_base/Lab7_ESP/current/bin/l7 stop $(/server_base/Lab7_ESP/current/bin/l7_
→status | grep -v database | grep RUNNING | awk '{print $1}')
```

- Second, once the L7|ESP 2.x server has stopped, create a backup of the database and temporary copy of the Data directory that will be used during the restore process:

```
$ /server_base/Lab7_ESP/current/sys/bin/pg_dump \
--host localhost \
--port 1487 \
--dbname lab7 \
--clean \
--if-exists \
--format=c > /tmp/backup/esp_db_$(date).pgdump

$ rsync \
--verbose \
--archive \
--exclude 'log/' \
--exclude 'run/' \
--exclude 'database/' \
--exclude 'conf/' \
/server_base/Lab7_ESP/Data/ \
/tmp/backup
```

- Next, extract the L7|ESP 3.x deployment bundle tarball and change into the directory it unpacked (where docker-compose.yml resides) before running the rest of the commands:

```
$ wget https://$download_link
$ tar xf $project-$version.tar.gz
$ cd $project/
```

Note: You will need to log into Docker Hub with the appropriate credentials before proceeding using the docker login command. Please contact L7 Informatics for more information.

The following command will launch a one-off L7|ESP 3.x container with the configuration specified in docker-compose.yml (mainly, there will be a Docker volume created at /opt/l7esp/data) but with the following settings overridden by command-line flags:

- The entrypoint process will be overridden to the Bash shell, such that a) L7|ESP won't automatically initialize itself and run the server processes, and b) instead you will be granted a shell session where you can perform the restoration process
- The container will be run as the root user (UID 1) such that you don't run into Linux permission issues given that the files mounted in from the host may be owned by user(s) not present inside the container and you will later have to adjust these permissions.
- The container will be automatically removed when you leave the Bash shell such that it doesn't remain in the stopped state on the host after the restoration process, as we only need this container during the restore process and any files you move into the data volume will remain since the volume will be persisted.
- The backup taken earlier to /tmp/backup on the host machine will be bind-mounted to /tmp/restore in the container file system such that the files will be available for you to restore from once you have a shell session inside the container.

```
$ docker-compose run \  
  --entrypoint bash \  
  --user root \  
  --rm \  
  --volume /tmp/backup:/tmp/restore server
```

Copy the backup you mounted into this container to the location you just freed up:

```
$ rsync -avh /tmp/restore/* /opt/l7esp/data
```

Modify the Linux file ownership of the restored files to match the container's default user:

```
$ chown -Rfv l7esp:l7esp /opt/l7esp/data
```

The data has now been restored and you may now exit the one-off container. The container will be automatically removed but the volume with the restored data will remain:

```
$ exit
```

Now you can exit the one-off container and it will be automatically removed. The data volume will be kept, and will now contain the data you restored from the previous L7|ESP instance.

```
$ docker-compose up --detach && docker-compose logs --follow
```

Once the logs have finished initialization login to the container console to complete the database restore.

```
$ docker-compose exec server bash
```

Use the following commands to stop the current ESP instance and perform the DB restore

```
$ l7 stop $(l7 status | grep -v database | grep RUNNING | awk '{print $1}')  
  
$ psql -h localhost -p 1487 -d postgres << EOF  
  DROP DATABASE lab7;  
  CREATE DATABASE lab7;  
  CREATE ROLE esp;  
  GRANT ALL PRIVILEGES ON DATABASE lab7 TO esp;  
EOF  
  
$ pg_restore \  
  --host localhost \  
  --port 1487 \  

```

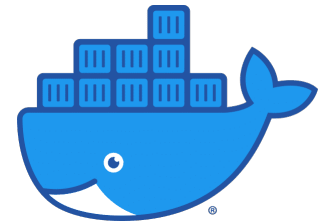
(continues on next page)

(continued from previous page)

```
--dbname lab7 \  
--clean \  
--if-exists \  
--no-owner \  
/opt/l7esp/data/backup/esp_db_${date}.pgdump  
  
$ l7 init --yes  
  
$ pg_esp_docker_migrate $server_base  
  
$ exit
```

Note: Replace *localhost* and *1487* with the appropriate hostname and PGSQL port when using a different ESP

See the Starting L7|ESP section above for commands to run to start up L7|ESP 3.0. The database you restored from 2.x will be automatically migrated to the latest 3.0 schema upon startup and if there are any errors you will see those in the container logs.



Docker Swarm

Prerequisites

To install L7|ESP onto a Docker Swarm cluster, you must have:

- At least 3 VMs with Docker Engine installed.

Note: As Docker Swarm is a distributed system, you must maintain a quorum of at least 3 nodes if you wish to be able to tolerate the failure of one of these nodes.

See [Add manager nodes for fault tolerance](#) in the official documentation for more information.

Creating the cluster

On the first node, run the following command to create the cluster:

```
$ docker swarm init
```

On the other nodes, run the following command to join the cluster:

```
$ docker swarm join-token manager
```

Note: It is recommended that the first two nodes are added as managers, per the fault tolerance recommendations above. Additional nodes can be added as workers also.

Compose file

Docker Swarm will take a Compose file as configuration so you should use the `docker-compose.prod.yml` file in root of your project, which will look something like this:

```
version: "3.2"
services:
  server:
    image: project_name/l7esp:3.0.0-sdk.1
    environment:
      L7ESP_PASSWORD: admin
    ports:
      - published: 8002
        target: 8002
        protocol: tcp
        mode: host
    volumes:
      - type: volume
        source: data
        target: /opt/l7esp/data
volumes:
  data: {}
```

Deploy Compose stack

From the project folder, run the following command to deploy the stack.

```
$ docker stack deploy --compose-file=./docker-compose.prod.yml l7esp-example
```

Validate Resource Creation

From a Swarm manager node, run the following command. Remember that the password is ‘vagrant’.

```
$ ssh manager.swarm.example.com
manager.swarm.example.com:~$ docker node ls
```

ID	ENGINE VERSION	HOSTNAME	STATUS	AVAILABILITY	MANAGER
ri27kwen3bu9s8384cl45rq2w *		docker-manager	Ready	Active	Leader
i87w5vzik6qag5ia3e79d46c9		docker-worker1	Ready	Active	

Validate L7|ESP Service

From a Swarm manager node, run the following command to validate that the L7|ESP container is running in the Swarm cluster:

```
$ ssh manager.swarm.example.com
manager.swarm.example.com:~$ docker node ps $(docker node ls -q)
```

ID	NAME	IMAGE	NODE	PORTS
↪ DESIRED STATE	CURRENT STATE	ERROR		
e2fgtesxlr9	esp_server.1	l7esp/server:3.0.0-sdk.2	docker-worker1	
↪ Running	Running 1 minute ago			



Amazon Elastic Container Service

Prerequisites

Amazon Elastic Container Instances is a quick way to deploy your container into Docker host within AWS, without having to provision the underlying servers first (VM + OS).

- AWS account
- AWS access profile
- AWS VPC and available subnet

CLI tools

The AWS command line tool will be used to programmatically access your AWS account. To install the CLI tool, follow the link below and choose the correct operating system:

- [AWS CLI](#)

Obtain the latest CloudFormation template

Save the following ESP CloudFormation template to your local environment as `ecs.template.yml`.

yaml

```

Parameters:
  appLicense:
    Type: String
    Description: L7 ESP application License
  appPassword:
    Type: String
    Description: L7 ESP application login password
  ImageRepo:
    Type: String
    Description: Azure Container Registry Image URL
  ContainerName:
    Type: String
    Description: ECS Container name
  containerCPU:
    Type: String
    Description: ECS CPU Resource Allocation
  appPortHTTP:
    Type: String
    Description: App Port
  containerMemory:
    Type: String
    Description: ECS Memory Resource Allocation
  MyTaskDefinition:
    Type: String
    Description: Task Definition Name
  ServiceName:
    Type: String
    Description: ECS Service name
  subnetID:
    Type: String
    Description: The SubnetID where ECS will be deployed
  RepoUsername:
    Type: String
    Description: Private Repository username
  RepoPassword:
    Type: String
    Description: Private Repository password
  NoEcho: True

Resources:
  ECSClusterRole:
    Type: AWS::IAM::Role
    Properties:
      RoleName: !Sub ${AWS::StackName}-Role
      AssumeRolePolicyDocument:
        Version: 2008-10-17
        Statement:
          Effect: "Allow"
          Principal:
            Service: "ecs-tasks.amazonaws.com"
          Action: "sts:AssumeRole"
      Description: ESP role for deployment
      Policies:
        - PolicyName: EmbeddedInlinePolicy
          PolicyDocument:

```

(continues on next page)

(continued from previous page)

```

    Version: "2012-10-17"
    Statement:
      - Effect: Allow
        Action: "ecs:*"
        Resource: "*"
      - Effect: Allow
        Action:
          - "kms:Decrypt"
          - "secretsmanager:GetSecretValue"
        Resource:
          - !Ref RepoCreds
    ManagedPolicyArns:
      - "arn:aws:iam::aws:policy/CloudWatchFullAccess"

ECSCluster:
  Type: AWS::ECS::Cluster
  Properties:
    ClusterName: !Sub '${AWS::StackName}-cluster'

RepoCreds:
  Type: AWS::SecretsManager::Secret
  Properties:
    Description: Docker Hub credentials for pulling ESP image for ECS deployment
    Name: !Sub '${AWS::StackName}-repocreds'
    SecretString: !Sub '{"username" : "${RepoUsername}", "password" : "${
↪RepoPassword}" }'

TaskDefinition:
  Type: AWS::ECS::TaskDefinition
  DependsOn: ECSClusterRole
  Properties:
    Family: !Ref MyTaskDefinition
    Cpu: !Ref containerCPU
    Memory: !Ref containerMemory
    ExecutionRoleArn: !Ref ECSClusterRole
    NetworkMode: awsvpc
    RequiresCompatibilities:
      - FARGATE
    ContainerDefinitions:
      - Name: !Ref ContainerName
        Image: !Ref ImageRepo
        RepositoryCredentials:
          CredentialsParameter: !Ref RepoCreds
        MountPoints:
          - ContainerPath: /opt/l7esp/data/project
            SourceVolume: !Sub ${ContainerName}-volume
        PortMappings:
          - ContainerPort: !Ref appPortHTTP
            HostPort: !Ref appPortHTTP
        Environment:
          - Name: L7ESP_PASSWORD
            Value: !Ref appPassword
          - Name: L7ESP_LICENSE
            Value: !Ref appLicense
    LogConfiguration:
      LogDriver: awslogs
    Options:

```

(continues on next page)

(continued from previous page)

```

        awslogs-group: !Ref LogGroup
        awslogs-region: !Ref AWS::Region
        awslogs-stream-prefix: !Ref ContainerName
    Volumes:
        - Name: !Sub ${ContainerName}-volume
    Service:
        Type: AWS::ECS::Service
        DependsOn: ECSCluster
        Properties:
            Cluster: !Ref ECSCluster
            ServiceName: !Ref ServiceName
            TaskDefinition: !Ref TaskDefinition
            DesiredCount: 1
            LaunchType: FARGATE
            NetworkConfiguration:
                AwsvpcConfiguration:
                    AssignPublicIp: ENABLED
                Subnets:
                    - !Ref subnetID

    LogGroup:
        Type: AWS::Logs::LogGroup
        Properties:
            LogGroupName: !Sub /ecs/${AWS::StackName}

    Outputs:
        ClusterName:
            Value: !Ref ECSCluster
        ServiceName:
            Value: !Ref Service

```

Deploy Cloudformation template with desired parameters

Save the below example parameters file in your local environment as `ecs.parameters.json`:

```

[
  {
    "ParameterKey": "appLicense",
    "ParameterValue": ""
  },
  {
    "ParameterKey": "appPortHTTP",
    "ParameterValue": "8002"
  },
  {
    "ParameterKey": "appPassword",
    "ParameterValue": ""
  },
  {
    "ParameterKey": "containerCPU",
    "ParameterValue": "2048"
  },
  {
    "ParameterKey": "containerMemory",
    "ParameterValue": "8192"
  }
]

```

(continues on next page)

(continued from previous page)

```

    },
    {
      "ParameterKey": "ImageRepo",
      "ParameterValue": ""
    },
    {
      "ParameterKey": "ContainerName",
      "ParameterValue": "l7esp"
    },
    {
      "ParameterKey": "RepoUsername",
      "ParameterValue": ""
    },
    {
      "ParameterKey": "RepoPassword",
      "ParameterValue": ""
    },
    {
      "ParameterKey": "MyTaskDefinition",
      "ParameterValue": "ExampleESPTaskDefinition"
    },
    {
      "ParameterKey": "ServiceName",
      "ParameterValue": "L7 ESP"
    },
    {
      "ParameterKey": "subnetID",
      "ParameterValue": ""
    }
  ]

```

Edit this file to include your specific AWS environment parameters:

- `containerName` - This will be the name of the ESP container that runs in your CloudFormation stack (default: `l7esp-example`)
- `appLicense` : L7|ESP software license (optional; JSON string)
- `appPassword` - L7|ESP admin password (required; initial ESP password)
- `appPortHTTP` : L7|ESP web UI port (default: 8002)
- `ImageRepo`: Container Image (e.g. `docker.io/acme/l7esp:3.0.0`)
- `RepoUsername`: The username for your container repository
- `RepoPassword`: The password for your container repository
- `containerCPU` : Container CPU cores (default: 2 vCPU, expressed as '2048') [\[AWS vCPU values explained\]](#)
- `containerMemory` : Container Memory (default: 6GB RAM)
- `ClusterName` : ECS cluster name (e.g. L7ECS)
- `TaskDefinitionName` : ECS Task name (e.g. `ExampleESPTaskDefinition`)
- `ServiceName`: ECS Service name (e.g. `ExampleESPService`)
- `subnetID` : The ID of your AWS subnet (e.g. `subnet-<id>`)

Run Cloudformation template pipeline to create ECS

Configure AWS profile from your AWS CLI to access aws account.

```
$ aws configure
```

List your AWS Profile:

```
$ aws configure list-profiles
```

Select the correct profile by setting the profile environmental variable

```
$ export AWS_PROFILE=profilename
```

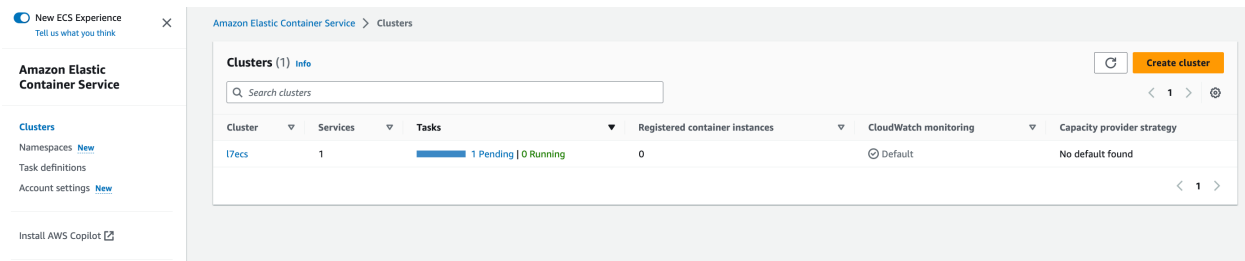
Deploy the AWS cloudformation

```
$ aws cloudformation deploy \
  --stack-name ESP-ECS-example \
  --template-file ecs.template.yml \
  --parameter-overrides file://ecs.parameters.json \
  --capabilities CAPABILITY_NAMED_IAM
```

You can add the `--debug` flag for additional debugging output.

Validate Resource Creation

In the AWS portal, navigate to ecs and verify that the L7|ESP container instance was created and that it ends up in the **Running** state:

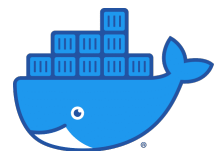


Azure Container Instances

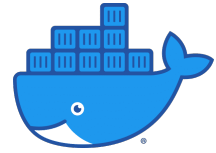
Azure Container Instances (ACI) allows you to run Docker containers in a managed, serverless cloud environment, without having to set up VMs, clusters, or orchestrators.



Docker Compose



Compose is a tool for defining and running multi-container Docker applications. With Compose, you use a YAML file to configure your application's services. Then, with a single command, you create and start all the services from your configuration.



Docker Swarm

Docker in swarm mode is a container orchestration tool that allows you to manage multiple containers deployed across multiple host machines.



Amazon Elastic Container Service

Elastic Container Instances (ECI) allows you to run Docker containers in a managed, serverless cloud environment, without having to set up VMs, clusters, or orchestrators.

1.16 Content Installation

The sections below provide details on mechanisms for creating bundles and deploying content created with the L7|ESP SDK. After reading this documentation, you'll understand how production bundles of L7|ESP are built and how to install a production bundle in either an existing Linux server or inside of a container.

1.16.1 Deployment Bundle

For deploying the application in production, L7 provides an application bundle that contains specific versions of L7|ESP and related tools, along with all content that customers have defined in their L7|ESP SDK. These bundles must be generated on request by L7, and once created, can be used to install L7|ESP and custom content on a production system.

Deployment bundles are structured in a way that mirrors the L7|ESP SDK structure, since a lot of the commands for installing the bundle in production are the same as those that install the bundle in the development VM. Generally, the only differences between the L7|ESP SDK structure and the deployment bundle structure are a few boilerplate files (the `README.md`, `Makefile`, and the `docs/` folder).

1.16.2 Installation

Requirements

The deployment bundle must be installed on one of the following Linux-based operating systems:

- Ubuntu 20.04+
- RHEL/CentOS 7+

Aside from the operating system, the only additional requirement is *make*, which comes by default on most Linux-based operating systems.

Install

After extracting this package, you can use the Makefile to install L7|ESP and all custom content created for this deployment:

```
~$ tar -zxvf <package-name>.tar.gz
~$ cd <package-name>
~$ make install
```

This command will prompt for several configuration settings (hit <enter> to use defaults) and then proceed to install L7|ESP and import all custom content.

You can check the status of the installation at any time by running:

```
~$ make status
L7|ESP web server      : Available
L7|ESP system services : Available
L7|ESP database       : Available
```

Update

To update an existing installation with a new L7|ESP bundle or related software/content, you can use a similar process as the install process. However, instead of `make install`, you'll use `make update` to perform the update:

```
~$ tar -zxvf <package-name>.tar.gz
~$ cd <package-name>
~$ make update
```

This command will run through a lot of the same tasks as `make install`, but will preserve the database state and also perform any necessary database migrations for the update process.

1.16.3 Administration

Reload

If at any point you need to reload the application after install (i.e. if the application gets into a strange state or if you've rebooted the server), you can run:

```
~$ make reload
```

This command will take down the instance (if it's running) and reload the L7|ESP service (including all related web, logging, and pipeline services).

Import

If at any point you've deleted your database and need to re-import content defined in the customer bundle, you can run:

```
~$ make import
```

All of the content that you've defined as part of your production configuration will be loaded.

1.16.4 Database

Reset

If at any point you need to reset the application after install (i.e. completely wipe the application database), you can run:

```
~$ make reset
```

This command will take down the instance (if it's running), remove the database, re-run database migrations and setup, and reload the L7|ESP service (including all related web, logging, and pipeline services).

Note: Please be careful when resetting a production database. All data will be lost. If you set `db_archive: true` in your production playbook, the database will always be backed up in a folder called `archive` in the application install directory.

Archive

To create a backup of the application database (including all content and data created by the application), you can run:

```
~$ make archive
```

This command will create a `.sql` file (in a folder called `archive` inside of the application install directory) containing a database dump from the application.

1.16.5 Other Commands

Below are the list of available commands in the deployment bundle's Makefile:

```
~$ make help
info          list info about package
init          install prerequisites
install       provision production instance
update        update existing esp installation with ansible
reset         reset database with ansible tasks
archive       backup/archive esp database
stop          stop running esp instance
reload        reload or start application
import        import bundled content with ansible
status        check status of running esp instance
```

1.17 Hardware Recommendations

The most effective hardware configurations reduce application risk, but requirements will vary based on ESP instance activity. Consult with an L7 representative before committing to a configuration.

Note: Requirements are current as of October 2023 and are subject to change.

1.17.1 Baseline Configuration

Minimum recommended specifications to get up and running with a Single-Server setup. Suitable for test and development deployments.

- Server: 4-core, 16GB RAM, 500GB SSD storage
- Operating System: 64-bit Linux distro (e.g., RHEL 8, Ubuntu 20.04 LTS)

1.17.2 Performance Configuration

Increase performance with a two-server setup. Recommended minimum configuration for a production deployment.

- Application Server: 8-core, 64GB RAM, 500GB SSD storage
- Database server: 8-core, 64GB RAM, 500GB SSD storage
- Operating System: 64-bit Linux distro (e.g., RHEL 8, Ubuntu 20.04 LTS)
- Database: PostgreSQL 13.X+

1.17.3 Optimal Performance Configuration

Minimize risk and increase performance with a customized multiple-server setup. (Does not include bioinformatics processing.)

Servers: The most reliable server configurations involve a combination of load balancers, shared storage, application servers, container orchestration software, and clustered database servers.

Minimum multi-server example:

- 2 x 8-core, 64GB RAM, 100GB disk application servers
- 2 x 8-core 64GB RAM, 500GB disk database servers
- 1TB SSD class shared storage server
- Operating System: 64-bit Linux distro (e.g., RHEL 8, Ubuntu 20.04 LTS)
- Database: PostgreSQL 13.X+

1.18 Database Configuration

1.18.1 Introduction

The ESP application doesn't need any permissions outside of its own PostgreSQL database and it is recommended to create a single database for the application, as well as a dedicated role with matching name, and grant all permissions for that role to that database.

Currently, there isn't a role separation between "read-only", "read/write", "read/write/DDL modification permissions" users at the application level. The same user is used with the *l7-init* (DDL migrations) as in normal application execution. These credentials are stored in the database config file */opt/l7esp/data/conf/database.json*.

Prerequisites 64-bit Linux VM with the following minimum specs:

- Debian or RHEL based OS
- 8vCPUs

- 64GB of RAM
- 500GB of available storage
- **Timezone configured for UTC/GMT**
- PostgreSQL version 13.X+

1.18.2 Cloud

If you are hosting L7|ESP in the cloud, see the following guides:

- [Amazon Aurora PostgreSQL](#)
- [Azure Database for PostgreSQL](#)

Amazon Aurora PostgreSQL

Prerequisites

The Amazon Aurora PostgreSQL is a manage database hosted on aws, which can then be used as the L7|ESP backend database

- AWS Subscription

CLI tools

To install AWS aurora database you must have:

- An AWS account to deploy to
- Access keys for the AWS account (see [Managing access keys for IAM users](#))

The `awscli` command line tools will be used to programmatically access the Kubernetes cluster and the AWS account. To install the CLI tools, follow the link below and choose the correct operating sytem:

- [AWS CLI](#)

Infrastructure as Code

CloudFormation Template

The latest CloudFormation template is as follows:

CloudFormation template

```

---
AWSTemplateFormatVersion: 2010-09-09
Description: A basic CloudFormation template for an RDS Aurora cluster.
Parameters:
  DatabaseInstanceType:
    Default: db.r5.large
    AllowedValues:

```

(continues on next page)

(continued from previous page)

```

    - db.r5.large
    - db.r5.xlarge
    - db.r5.2xlarge
    - db.r5.4xlarge
    - db.r5.8xlarge
    Description: The instance type to use for the database.
    Type: String
DBEngineVersion:
    Default: 13.7
    AllowedValues:
        - 11.16
        - 12.11
        - 13.7
        - 14.3
    Description: Select Database Engine Version
    Type: String
DatabasePassword:
    AllowedPattern: "[a-zA-Z0-9]+"
    ConstraintDescription: must contain only alphanumeric characters.
    Description: The database admin account password.
    MaxLength: '41'
    MinLength: '8'
    NoEcho: 'true'
    Type: String
    Default: "xekyxwxcu2mzy1KXA"
DatabaseUsername:
    AllowedPattern: "[a-zA-Z0-9]+"
    ConstraintDescription: must contain only alphanumeric characters.
    Description: The database admin account user name.
    MaxLength: '16'
    MinLength: '1'
    Type: String
    Default: "admin"
DatabaseBackupRetentionPeriod:
    Type: String
    Default: 7
    AllowedValues:
        - 1
        - 7
    Default: 7
    Description: The database backup retention period in days.
DatabaseSubnets:
    Description: The subnets to place database instances in.
    Type: List<AWS::EC2::Subnet::Id>
    Default: "subnet-09c7fe8aea62f0361, subnet-02206efa4f182c7fb"
DatabaseSecurityGroups:
    Type: List<AWS::EC2::SecurityGroup::Id>
    Description: Security groups to apply to the RDS cluster.
    Default: "sg-0baeabc6d976c41cf"
Metadata:
    AWS::CloudFormation::Interface:
    ParameterGroups:
        - Label:
            default: Database Configuration
        Parameters:
            - DatabaseInstanceType
            - DatabaseName

```

(continues on next page)

(continued from previous page)

```

        - DatabaseUsername
        - DatabasePassword
        - DatabaseSubnets
        - DatabaseSecurityGroups
        - DatabaseBackupRetentionPeriod
        - DBEngineVersion
    ParameterLabels:
        DatabaseInstanceType:
            default: Database Instance Type
        DatabasePassword:
            default: Database Password
        DatabaseUsername:
            default: Database Username
        DatabaseBackupRetentionPeriod:
            default: Database Backup Retention Period
        DatabaseSubnets:
            default: Database Subnets
        DatabaseSecurityGroups:
            default: Database Security Groups
Resources:
    StackAlarmTopic:
        Type: AWS::SNS::Topic
        Properties:
            DisplayName: Stack Alarm Topic
    DatabaseSubnetGroup:
        Type: AWS::RDS::DBSubnetGroup
        Properties:
            DBSubnetGroupDescription: CloudFormation managed DB subnet group.
            SubnetIds:
                Ref: DatabaseSubnets
    DatabaseCluster:
        Type: AWS::RDS::DBCluster
        Properties:
            Engine: aurora-postgresql
            EngineVersion:
                Ref: DBEngineVersion
            MasterUsername:
                Ref: DatabaseUsername
            MasterUserPassword:
                Ref: DatabasePassword
            BackupRetentionPeriod:
                Ref: DatabaseBackupRetentionPeriod
            PreferredBackupWindow: 02:00-03:00
            PreferredMaintenanceWindow: mon:03:00-mon:04:00
            DBSubnetGroupName:
                Ref: DatabaseSubnetGroup
            VpcSecurityGroupIds:
                Ref: DatabaseSecurityGroups
    DatabasePrimaryInstance:
        Type: AWS::RDS::DBInstance
        Properties:
            Engine: aurora-postgresql
            EngineVersion:
                Ref: DBEngineVersion
            DBClusterIdentifier:
                Ref: DatabaseCluster
            DBInstanceClass:

```

(continues on next page)

(continued from previous page)

```

        Ref: DatabaseInstanceType
    DBSubnetGroupName:
        Ref: DatabaseSubnetGroup
DatabaseReplicaInstance:
    Type: AWS::RDS::DBInstance
    Properties:
        Engine: aurora-postgresql
        EngineVersion:
            Ref: DBEngineVersion
        DBClusterIdentifier:
            Ref: DatabaseCluster
        DBInstanceClass:
            Ref: DatabaseInstanceType
        DBSubnetGroupName:
            Ref: DatabaseSubnetGroup
DatabasePrimaryCPUAlarm:
    Type: AWS::CloudWatch::Alarm
    Properties:
        AlarmDescription: Primary database CPU utilization is over 80%.
        Namespace: AWS/RDS
        MetricName: CPUUtilization
        Unit: Percent
        Statistic: Average
        Period: 300
        EvaluationPeriods: 2
        Threshold: 80
        ComparisonOperator: GreaterThanOrEqualToThreshold
        Dimensions:
            - Name: DBInstanceIdentifier
              Value:
                  Ref: DatabasePrimaryInstance
        AlarmActions:
            - Ref: StackAlarmTopic
        InsufficientDataActions:
            - Ref: StackAlarmTopic
DatabaseReplicaCPUAlarm:
    Type: AWS::CloudWatch::Alarm
    Properties:
        AlarmDescription: Replica database CPU utilization is over 80%.
        Namespace: AWS/RDS
        MetricName: CPUUtilization
        Unit: Percent
        Statistic: Average
        Period: 300
        EvaluationPeriods: 2
        Threshold: 80
        ComparisonOperator: GreaterThanOrEqualToThreshold
        Dimensions:
            - Name: DBInstanceIdentifier
              Value:
                  Ref: DatabaseReplicaInstance
        AlarmActions:
            - Ref: StackAlarmTopic
        InsufficientDataActions:
            - Ref: StackAlarmTopic
DatabasePrimaryMemoryAlarm:
    Type: AWS::CloudWatch::Alarm

```

(continues on next page)

(continued from previous page)

```

Properties:
  AlarmDescription: Primary database freeable memory is under 700MB.
  Namespace: AWS/RDS
  MetricName: FreeableMemory
  Unit: Bytes
  Statistic: Average
  Period: 300
  EvaluationPeriods: 2
  Threshold: 700000000
  ComparisonOperator: LessThanOrEqualToThreshold
  Dimensions:
    - Name: DBInstanceIdentifier
      Value:
        Ref: DatabasePrimaryInstance
  AlarmActions:
    - Ref: StackAlarmTopic
  InsufficientDataActions:
    - Ref: StackAlarmTopic
DatabasePrimaryMemoryAlarm:
  Type: AWS::CloudWatch::Alarm
  Properties:
    AlarmDescription: Replica database freeable memory is under 700MB.
    Namespace: AWS/RDS
    MetricName: FreeableMemory
    Unit: Bytes
    Statistic: Average
    Period: 300
    EvaluationPeriods: 2
    Threshold: 700000000
    ComparisonOperator: LessThanOrEqualToThreshold
    Dimensions:
      - Name: DBInstanceIdentifier
        Value:
          Ref: DatabaseReplicaInstance
    AlarmActions:
      - Ref: StackAlarmTopic
    InsufficientDataActions:
      - Ref: StackAlarmTopic
DatabasePrimaryReplicationAlarm:
  Type: AWS::CloudWatch::Alarm
  Properties:
    AlarmDescription: Database replication latency is over 200ms.
    Namespace: AWS/RDS
    MetricName: AuroraReplicaLag
    Unit: Milliseconds
    Statistic: Average
    Period: 300
    EvaluationPeriods: 2
    Threshold: 200
    ComparisonOperator: GreaterThanOrEqualToThreshold
    Dimensions:
      - Name: DBInstanceIdentifier
        Value:
          Ref: DatabaseReplicaInstance
    AlarmActions:
      - Ref: StackAlarmTopic
DatabaseReplicaReplicationAlarm:

```

(continues on next page)

(continued from previous page)

```

Type: AWS::CloudWatch::Alarm
Properties:
  AlarmDescription: Database replication latency is over 200ms.
  Namespace: AWS/RDS
  MetricName: AuroraReplicaLag
  Unit: Milliseconds
  Statistic: Average
  Period: 300
  EvaluationPeriods: 2
  Threshold: 200
  ComparisonOperator: GreaterThanOrEqualToThreshold
  Dimensions:
    - Name: DBInstanceIdentifier
      Value:
        Ref: DatabaseReplicaInstance
  AlarmActions:
    - Ref: StackAlarmTopic

```

Parameters

- DatabaseInstanceType (String) - The instance type to use for the database.
- DatabasePassword (String) - The database admin account password.
- DatabaseUsername (String) - The database admin account user name.
- DatabaseBackupRetentionPeriod: (Number) - The database backup retention period in days.
- DatabaseSubnets (String) - The subnets to place database instances in.
- DatabaseSecurityGroups (String) - Security groups to apply to the RDS cluster.

Parameters File

Create a `parameters.json` file for use when deploying the CloudFormation template:

CloudFormation template

```

[
  {
    "ParameterKey": "DatabaseInstanceType",
    "ParameterValue": "db.r5.large"
  },
  {
    "ParameterKey": "DatabasePassword",
    "ParameterValue": ""
  },
  {
    "ParameterKey": "DatabaseUsername",
    "ParameterValue": ""
  },
  {
    "ParameterKey": "DatabaseBackupRetentionPeriod",

```

(continues on next page)

(continued from previous page)

```

    "ParameterValue": "7"
  },
  {
    "ParameterKey": "DatabaseSubnets",
    "ParameterValue": ""
  },
  {
    "ParameterKey": "DatabaseSecurityGroups",
    "ParameterValue": ""
  },
  {
    "ParameterKey": "DBEngineVersion",
    "ParameterValue": "13.7"
  }
]

```

Provisioning

Create a profile for use with AWS CLI:

```

~$ aws configure --profile <profile>
AWS Access Key ID [None]: AKIAIOSFODNN7EXAMPLE
AWS Secret Access Key [None]: wJalrXUtnFEMI/K7MDENG/bPxrFiCYEXAMPLEKEY
Default region name [us-east-1]:
Default output format [None]:

```

Switch to the correct AWS profile for current shell session:

```

~$ export AWS_PROFILE=<profile>

```

Deploy CloudFormation stack:

```

$ aws cloudformation deploy \
  --stack-name aurora-example \
  --template-file aws_aurora.yml \
  --parameter-overrides file://aurora.parameters.json \
  --capabilities CAPABILITY_NAMED_IAM

```

Validate Resource Creation

Log into the AWS Console and navigate to CloudFormation. Under “Stacks”, validate that the CloudFormation stack was created:

```

$ aws rds describe-db-clusters \
  --db-cluster-identifier <cluster name>

```

Connect ESP instance to AWS Aurora

Once you validate that the aurora database have been successfully created on the aws portal, gather the database connection strings. You can find this under the database options of RDS portal.

The screenshot shows the Amazon RDS console interface. On the left is a navigation sidebar with options like Dashboard, Databases, Query Editor, Performance insights, Snapshots, Automated backups, Reserved instances, Proxies, Subnet groups, Parameter groups, Option groups, Custom engine versions, Events, Event subscriptions, Recommendations (0), and Certificate update. The main content area shows the breadcrumb path 'RDS > Databases > aurora-example-databasecluster-1o70oziwymna' and the instance ID 'ad18cv41ro0m5fe'. Below this, a 'Related' section contains a search bar and a list of database identifiers. The identifier 'ad18cv41ro0m5fe' is selected. At the bottom, the 'Connectivity & security' tab is active, displaying the 'Endpoint & port' section with the endpoint 'ad18cv41ro0m5fe.c8abnghvesd3.us-east-1.rds.amazonaws.com' and port '3306'.

Next, we can add the database connection string on the esp helm chart. on the values.yml file under the database section add the connection string informations:

```
database: |
{
  "host": "<postgres_server_url_or_ip>",
```

(continues on next page)

(continued from previous page)

```

    "port": <postgres_server_port>,
    "user": "<esp_db_user>",
    "pass": "<esp_db_password>",
    "name": "<esp_db_name>",
    "schema": "<esp_db_schema>",
    "multitenant": false,
    "start_service": true,
    "l7-esp.concierge": {
        "pool_size": 3
    },
    "l7-esp.notifications": {
        "pool_size": 3
    },
    "l7-esp.hub": {
        "pool_size": 3
    },
    "l7-esp.http": {
        "pool_size": 3
    }
}

```

Now deploy the helm chart (see helm documentation). Verify that esp is up and the database schema was successfully created

Azure Database for PostgreSQL

Prerequisites

The Azure Database for PostgreSQL is a manage database hosted on Azure, which can then be used as the L7|ESP backend database

- Azure Subscription

CLI tools

The `az` command line tools will be used to programmatically access/manage the PostgreSQL instance and the Azure subscription. To install the CLI tool, follow the link below and choose the correct operating sytem:

- [Azure CLI](#)

ARM template

The latest ARM template is as follows:

Bicep

```

1 param administratorLogin string
2
3 @secure()
4 param administratorLoginPassword string
5 param location string = resourceGroup().location
6 param serverName string
7 param serverEdition string = 'GeneralPurpose'
8 param skuSizeGB int = 128
9 param dbInstanceType string = 'Standard_D4ds_v4'
10 param haMode string = 'ZoneRedundant'
11 param availabilityZone string = '1'
12 param version string = '12'
13 param virtualNetworkExternalId string = ''
14 param subnetName string = ''
15 param privateDnsZoneArmResourceId string = ''
16
17 resource serverName_resource 'Microsoft.DBforPostgreSQL/flexibleServers@2021-06-01' =
↪ {
18   name: serverName
19   location: location
20   sku: {
21     name: dbInstanceType
22     tier: serverEdition
23   }
24   properties: {
25     version: version
26     administratorLogin: administratorLogin
27     administratorLoginPassword: administratorLoginPassword
28     network: {
29       delegatedSubnetResourceId: (empty(virtualNetworkExternalId) ? json('null') :
↪ json('${virtualNetworkExternalId}/subnets/${subnetName}'))
30       privateDnsZoneArmResourceId: (empty(virtualNetworkExternalId) ? json('null') :
↪ privateDnsZoneArmResourceId)
31     }
32     highAvailability: {
33       mode: haMode
34     }
35     storage: {
36       storageSizeGB: skuSizeGB
37     }
38     backup: {
39       backupRetentionDays: 7
40       geoRedundantBackup: 'Disabled'
41     }
42     availabilityZone: availabilityZone
43   }
44 }

```

JSON

```

1 {
2   "$schema": "https://schema.management.azure.com/schemas/2019-04-01/
↪deploymentTemplate.json#",
3   "contentVersion": "1.0.0.0",
4   "parameters": {
5     "administratorLogin": {
6       "type": "string",
7     },
8     "administratorLoginPassword": {
9       "type": "secureString"
10    },
11    "location": {
12      "type": "string",
13      "defaultValue": "[resourceGroup().location]"
14    },
15    "serverName": {
16      "type": "string"
17    },
18    "serverEdition": {
19      "type": "string",
20      "defaultValue": "GeneralPurpose"
21    },
22    "skuSizeGB": {
23      "type": "int",
24      "defaultValue": 128
25    },
26    "dbInstanceType": {
27      "type": "string",
28      "defaultValue": "Standard_D4ds_v4"
29    },
30    "haMode": {
31      "type": "string",
32      "defaultValue": "ZoneRedundant"
33    },
34    "availabilityZone": {
35      "type": "string",
36      "defaultValue": "1"
37    },
38    "version": {
39      "type": "string",
40      "defaultValue": "12"
41    },
42    "virtualNetworkExternalId": {
43      "type": "string",
44      "defaultValue": ""
45    },
46    "subnetName": {
47      "type": "string",
48      "defaultValue": ""
49    },
50    "privateDnsZoneArmResourceId": {
51      "type": "string",
52      "defaultValue": ""
53    }
54  },

```

(continues on next page)

(continued from previous page)

```

55  "resources": [
56    {
57      "type": "Microsoft.DBforPostgreSQL/flexibleServers",
58      "apiVersion": "2021-06-01",
59      "name": "[parameters('serverName')]",
60      "location": "[parameters('location')]",
61      "sku": {
62        "name": "[parameters('dbInstanceType')]",
63        "tier": "[parameters('serverEdition')]"
64      },
65      "properties": {
66        "version": "[parameters('version')]",
67        "administratorLogin": "[parameters('administratorLogin')]",
68        "administratorLoginPassword": "[parameters('administratorLoginPassword')]",
69        "network": {
70          "delegatedSubnetResourceId": "[if(empty(parameters('virtualNetworkExternalId
71 ↪')), json('null'), json(format('{0}/subnets/{1}', parameters(
72 ↪ 'virtualNetworkExternalId'), parameters('subnetName')))]",
73          "privateDnsZoneArmResourceId": "[if(empty(parameters(
74 ↪ 'virtualNetworkExternalId')), json('null'), parameters('privateDnsZoneArmResourceId
75 ↪'))]"
76        },
77        "highAvailability": {
78          "mode": "[parameters('haMode')]"
79        },
80        "storage": {
81          "storageSizeGB": "[parameters('skuSizeGB')]"
82        },
83        "backup": {
84          "backupRetentionDays": 7,
85          "geoRedundantBackup": "Disabled"
86        },
87        "availabilityZone": "[parameters('availabilityZone')]"
88      }
89    }
90  ]
91 }

```

ARM template parameters

When deploying the ARM template, you can override parameters like so:

- administratorLogin: The default administrator username.
- administratorLoginPassword: The default administrator password
- location: The location of AKS resource. Default value: same region/location as the resource group you are deploying ARM template into.
- serverName: The name of the server.
- serverEdition: The edition of the server.
- skuSizeGB: the storage size, in Gigabyte, of the Azure Database for PostgreSQL server.
- dbInstanceType: The size of the virtual machine. Default: Standard_D4_v4
- haMode: high availability mode.

- `availabilityZone`: availability of the database.
- `version`: the version of PostgreSQL server to deploy.

Deploy ARM template

```
$serverName = Read-Host -Prompt "Enter a name for the new Azure Database for
↳PostgreSQL server"
$resourceGroupName = Read-Host -Prompt "Enter a name for the new resource group where
↳the server will exist"
$location = Read-Host -Prompt "Enter an Azure region (for example, centralus) for the
↳resource group"
$adminUser = Read-Host -Prompt "Enter the Azure Database for PostgreSQL server's
↳administrator account name"
$adminPassword = Read-Host -Prompt "Enter the administrator password" -AsSecureString

New-AzResourceGroup -Name $resourceGroupName -Location $location # Use this command
↳when you need to create a new resource group for your deployment
New-AzResourceGroupDeployment -ResourceGroupName $resourceGroupName `
    -TemplateFile "postgres-flexible-server-template.json" `
    -serverName $serverName `
    -administratorLogin $adminUser `
    -administratorLoginPassword $adminPassword

Read-Host -Prompt "Press [ENTER] to continue ..."
```

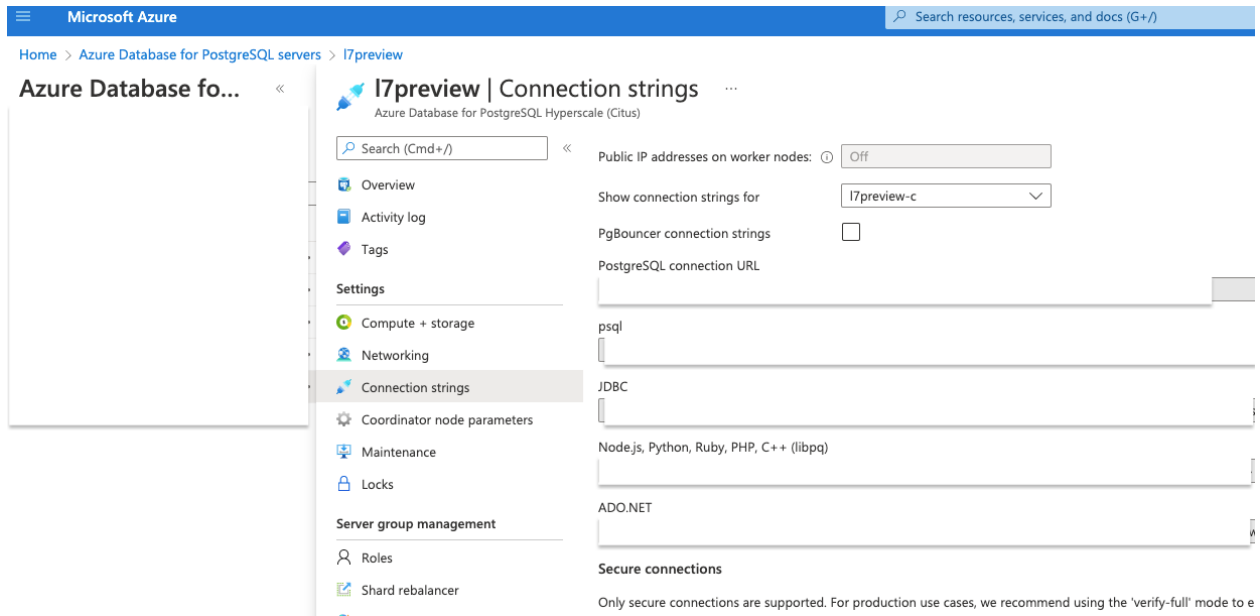
Review Deployed resources

Follow these steps to verify if your server was created in Azure.

```
echo "Enter your Azure Database for PostgreSQL Flexible Server name:" &&
read serverName &&
echo "Enter the resource group where the Azure Database for PostgreSQL Flexible
↳Server exists:" &&
read resourcegroupName &&
az resource show --resource-group $resourcegroupName --name $serverName --resource-
↳type "Microsoft.DBforPostgreSQL/flexibleServers"
```

Connect ESP instance to Azure PostgreSQL

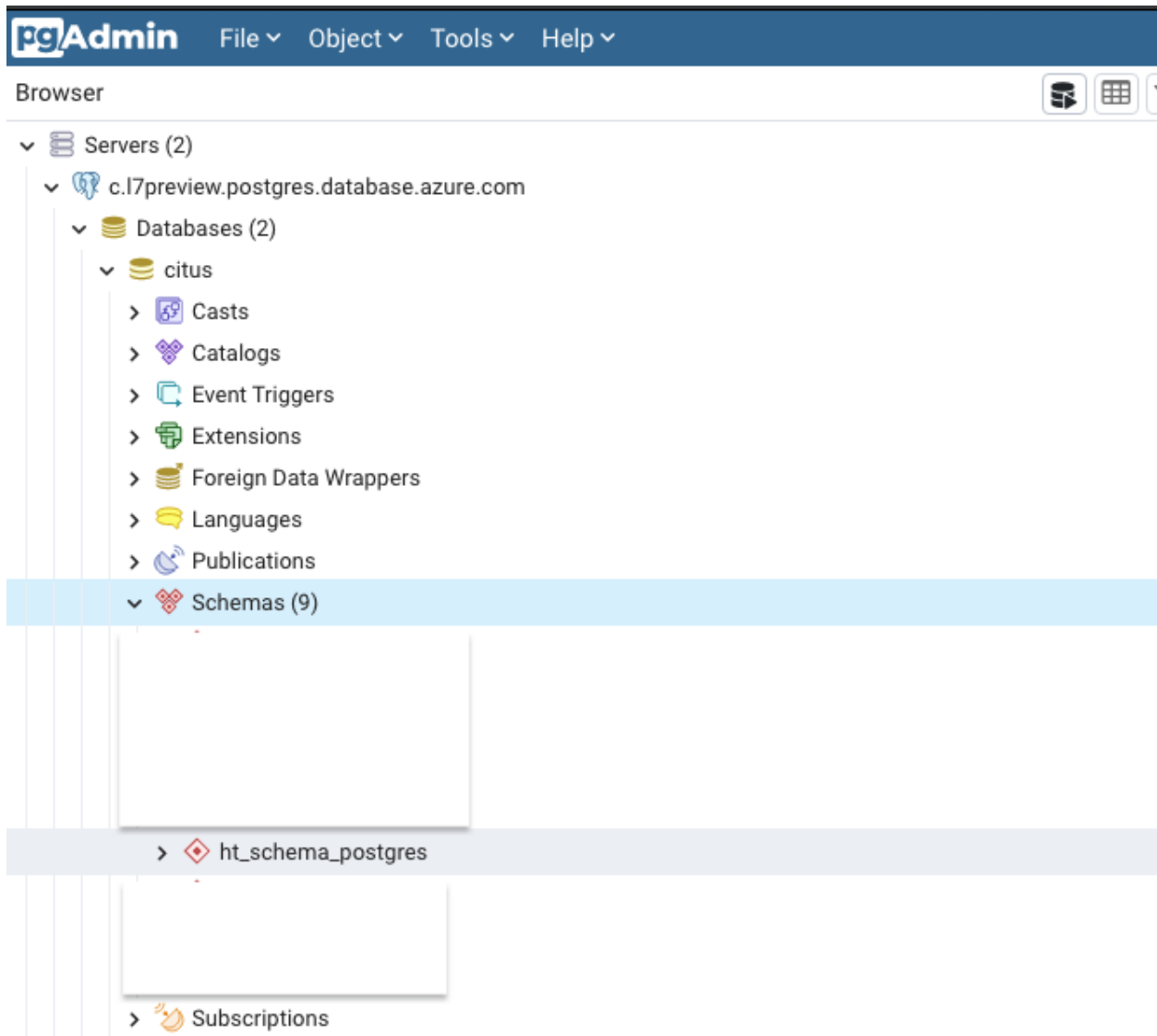
Once you validate that the postgres database have been successfully created on the Azure portal, gather the database connection strings. You can find this under the connection strings settings of the database on the portal.



Next, we can add the database connection string on the esp helm chart. on the values.yml file under the database section add the connection string informations:

```
database: |
{
  "host": "<postgres_server_url_or_ip>",
  "port": <postgres_server_port>,
  "user": "<esp_db_user>",
  "pass": "<esp_db_password>",
  "name": "<esp_db_name>",
  "schema": "<esp_db_schema>",
  "multitenant": false,
  "start_service": true,
  "l7-esp.concierge": {
    "pool_size": 3
  },
  "l7-esp.notifications": {
    "pool_size": 3
  },
  "l7-esp.hub": {
    "pool_size": 3
  },
  "l7-esp.http": {
    "pool_size": 3
  }
}
```

Now deploy the helm chart (see helm documentation). Verify that esp is up and the database schema was successfully created



1.18.3 Instructions

ESP DB creation

Most installations of PostgreSQL done via a package manager include a user called “postgres”. This user has full super admin privileges to the PostgreSQL instance installed on your system. Using the `su` command switch to the postgres user account then use the PostgreSQL interactive terminal (*psql*) to create the database, user role, and assign the permissions.

```
postgres=# CREATE DATABASE <esp_db_name>;
postgres=# CREATE USER <esp_db_user> WITH ENCRYPTED PASSWORD '<esp_db_password>';
postgres=# GRANT ALL PRIVILEGES ON DATABASE <esp_db_name> TO <esp_db_user>;
```

ESP DB configuration

The database configuration file at `/opt/l7esp/data/conf/database.json` should take the following format when using an external hosted PostgreSQL solution such as AWS RDS:

```
{
  "host": "<postgres_server_url_or_ip>",
  "port": "<postgres_server_port>",
  "user": "<esp_db_user>",
  "pass": "<esp_db_password>",
  "name": "<esp_db_name>",
  "schema": "<esp_db_schema>",
  "multitenant": false,
  "start_service": false,
  "l7-esp.concierge": {
    "pool_size": 3
  },
  "l7-esp.notifications": {
    "pool_size": 3
  },
  "l7-esp.hub": {
    "pool_size": 3
  },
  "l7-esp.http": {
    "pool_size": 3
  }
}
```

Note: When using an external PGSQL server, it will either need the data migrated from the existing PGSQL server or the new PGSQL server will need to be initialized, to prepare it for use with ESP.

Preparing PostgreSQL for ESP

Before starting ESP with an external database, you need to modify the PostgreSQL configuration file to work with L7|ESP. You can find the location of the configuration file by running the following query in the PostgreSQL terminal: `SHOW config_file;`

To optimize PostgreSQL for L7|ESP, refer to the [System Tuning](#) guide.

1.19 System Tuning

This section outlines some recommendations with respect to tuning L7|ESP application servers.

1.19.1 Operating System Tuning

As L7|ESP is cross-compiled and runs on many operating systems; we don't currently have any specific requirements for tuning at the operating system level (apart from hardware sizing recommendations which are in another document). Depending on the operating system, it's possible that some of the default values may not be suitable for a server environment. One such example of this pertains to the default per-user/application `ulimit` for open file descriptors, which can often be set to a very low value such as 1024 while a more suitable value for any server environment may be much higher (e.g. 24000) depending on the amount of traffic the server is handling. Our recommendation is to increase these limits if you find you are bumping against them, however the exact values chosen may vary widely depending on scale.

1.19.2 Application Tuning

L7 recommends adjusting the following settings in a production L7|ESP installation, contingent on the server running the DB having adequate resources to do so.

Postgresql Configuration

The postgresql configuration (in *Lab7_ESP/Data/conf/postgresql.conf*) should be modified as follows:

1. **max_connections:** should be a minimum of 6x the total number of L7|ESP processes. For instance, if `l7 status` shows:

```
$ l7 status
l7-esp.broker RUNNING pid 95176, uptime 4 days, 4:44:44
l7-esp.concierge RUNNING pid 95184, uptime 4 days, 4:44:44
l7-esp.database RUNNING pid 95113, uptime 4 days, 4:44:46
l7-esp.executor RUNNING pid 95185, uptime 4 days, 4:44:44
l7-esp.haproxy RUNNING pid 95178, uptime 4 days, 4:44:44
l7-esp.http:l7-esp.http.0 RUNNING pid 95182, uptime 4 days, 4:44:44
l7-esp.http:l7-esp.http.1 RUNNING pid 95181, uptime 4 days, 4:44:44
l7-esp.http:l7-esp.http.2 RUNNING pid 95180, uptime 4 days, 4:44:44
l7-esp.logger RUNNING pid 95179, uptime 4 days, 4:44:44
l7-esp.notification RUNNING pid 95175, uptime 4 days, 4:44:44
l7-esp.pipeline RUNNING pid 95183, uptime 4 days, 4:44:44
l7-esp.scheduler RUNNING pid 95177, uptime 4 days, 4:44:44
```

Then you should set a minimum # of connections of 72. For 3-worker configurations, L7 routinely sets this value to 100, and for systems with more web workers (the “http” processes), L7 routinely sets the value to 200 for systems with more web workers.

2. **shared_buffers:** should be set to ~25% of the available system RAM*, but not more than 8GB. (*assumes you have a dedicated DB server OR enough RAM to handle the load of the DB + the app servers. See below for app server requirements).
3. **temp_buffers:** Not less than 64MB. L7 routinely sets this to 128MB on dedicated DB servers in production.
4. **work_mem:** Not less than 256MB. L7 routinely sets this to 1GB on dedicated DB servers in production.
5. **maintenance_work_mem:** Not less than 128MB. L7 routinely sets this to 256MB on dedicated DB servers in production.

6. **effective_cache_size**: On a dedicated DB server - 75% of available RAM. On other servers, 2x the **shared_buffer** size.

7. **timezone**: This value should always be set to 'UTC '

L7 also recommends indexing the following columns of the “resource_val” table (note: these indexes will be applied by default starting in L7|ESP 2.4):

1. The `bound_resource_id` column
2. The `step_instance_sample_id` column.

Note: When running Postgres inside a container, it’s important to configure the shared memory size correctly. Docker provides containers with only 64MB of shared memory by default, which might be insufficient for Postgres to operate efficiently, especially during the execution of parallel queries. To increase the shared memory size available to Postgres, add the `shm_size` option to the Docker Compose file. Here’s an example:

```
services:
  server:
    image: l7esp/server
    shm_size: 1gb
```

The `shm_size` value in Docker corresponds to the shared memory segment in `/dev/shm` used by Postgres for parallel queries. The amount of shared memory a query uses is controlled by `work_mem`. Each query can use up to `work_mem` for each node in the `EXPLAIN` plan, for each process. This can significantly increase if you have many parallel worker processes and/or many tables/partitions being sorted or hashed in your query.

To calculate a reasonable `shm_size`, consider the total RAM available on the server, the amount allocated to Postgres, and the value of `work_mem`. Also, keep in mind these PostgreSQL settings that affect shared memory usage:

- `max_connections`: Each connection to the database server requires additional shared memory, typically a few kilobytes per connection.
- `max_locks_per_transaction`: More locks lead to more shared memory usage, with each additional lock using just a few bytes.
- `max_pred_locks_per_transaction`: This setting, used for predicate locking in Serializable transactions, also consumes additional shared memory, similar to `max_locks_per_transaction`.
- `max_prepared_transactions`: Each prepared transaction consumes additional shared memory, generally in the kilobytes range.
- `wal_buffers`: The memory used to store transaction logs, typically set to a few megabytes.
- `timezone`: The timezone set for PostgreSQL server instance.

On a memory-constrained system, you may want to consider disabling parallel queries. While parallel queries can make individual queries faster, they use more resources per query, which might not improve the overall throughput. To disable parallel queries, set `max_parallel_workers` to 0.

Remember to adjust the `shm_size` value based on actual usage and requirements, while also considering available resources on the server and potential performance issues. Allocating too much memory to shared memory may leave insufficient memory for other operations within the container.

L7|ESP Worker Configuration

L7|ESP's web workers are currently set, for most requests, to handle 1 request per worker*. Thus, the number of workers should be carefully considered when examining concurrent traffic load. (*Some routes such as adding/removing/failing samples to/from LIMS worksheets and file uploads allow concurrent requests to the same worker; the routes that support this will be expanding in future versions).

At rest, an L7|ESP worker consumes ~250-500mb of RAM, depending on a number of implementation-specific factors. Under large workflow loads (large batch sizes and/or large workflows), the worker memory can spike to ~2GB of RAM. If you observe memory spikes in excess of 2-3GB per worker, please notify L7.

This means a baseline system with L7|ESP running 3 web workers will require ~6-8 GB of RAM for routine operations, excluding the DB needs, to properly service ~5 users.

Note: Scaling workers based on available RAM can overload the database if it is already the bottleneck. Verify database CPU and I/O utilization before increasing worker count.

Another consideration is automated processes (pipelines). HAProxy routes requests based on the User-Agent header: requests with "python" in the User-Agent (i.e., pipeline/API client requests) are routed to a subset of workers called the "client backend", while UI requests can be routed to any worker.

The number of workers allocated to the client backend is controlled by the `num_python_backend_percent` setting in the `system` configuration block (default: "1/3"). The formula is: `ceil(num_workers * numerator / denominator)`.

For example:

- **3 workers** with "1/3": `ceil(3 * 1/3) = 1` client worker (worker 3)
- **4 workers** with "1/3": `ceil(4 * 1/3) = 2` client workers (workers 3-4)
- **6 workers** with "1/3": `ceil(6 * 1/3) = 2` client workers (workers 5-6)

Note that UI requests can still be routed to client backend workers, but pipeline requests are restricted to only the client backend workers. This ensures that heavy pipeline workloads don't consume all available workers.

For production configurations, L7|ESP recommends a minimum of 4 workers. For a configuration supporting 10 users with a standalone DB server and an application server with 16GB of RAM, L7 recommends a minimum of 6 web workers as follows:

1. **Set the `num_workers` key in the `executor` block of L7|ESP configuration to 2** (this can be done via the *Config* app in the user interface)
2. **Set the `num_workers` key in the `web_server` block of the L7|ESP configuration to 6** (this can be done via the *Config* app in the user interface)
3. **Set the `numprocs` value of the `[program:l7-esp.http]` stanza of the `supervisord.conf` file to 6**

For systems where the DB server is co-located with the application server, the system should have a minimum of 32GB of RAM for production use in environments where large worksheets are anticipated. (For this document, a large worksheet is any worksheet processing > 96 samples through a workflow with a combined total # of columns > 32 across all protocols. For instance, a server anticipating large loads should run with 32GB of RAM, an `effective_cache_size` (postgresql.conf) or 16GB, `shared_buffers` of 8GB, and number of web workers = 3-6 and number of executors = 1-2 depending on anticipated concurrent user use.

Isolating Specific Routes

For high-traffic integrations, you may want to dedicate workers to specific routes to prevent them from impacting other users. This requires modifying the haproxy configuration:

1. Add an ACL to the frontend to match the route:

```
" acl is_isolated_integration path_beg -i /api/isolated_integration",  
" use_backend isolated_integration_backend if is_isolated_integration",
```

2. Add a backend pointing to dedicated port(s):

```
"isolated_integration_backend": [  
  " option forwardfor",  
  " http-response set-header X-Server %s",  
  " server isolated_integration 127.0.0.1:8010"  
],
```

You can either hard-code to an existing worker port, or add a dedicated worker entry in `supervisord.conf` for full isolation:

```
[program:l7-esp.http.isolated_integration]  
command=python %(ENV_LAB7DIST)s/services/l7-esp.http.py -s 10 -n api  
numprocs=1  
; copy remaining keys from an existing service
```

1.20 HA Cluster Setup

L7|ESP application servers don't store files in the database, instead writing these to disk and storing a reference to the file path in the database. This is the case for any pipeline scripts that are generated and executed, the `stdout/stderr` logs we persist for the runs of such pipelines, and files that are generated by these pipelines and/or user-uploaded to the system via the Data application or related API endpoints.

L7|ESP writes all such state to a directory in the server root named "data". Depending on the installation, this may also contain logs generated by the running L7|ESP services and the PostgreSQL data directory if using the included local database. This works well for Dockerized use-cases where you may wish to make the filesystem read-only and capture all persistent state in a volume, though usually logging would be reconfigured to redirect to `stdout` in this case.

This can also be leveraged to create highly available architectures, where two or more L7|ESP application servers are placed behind a load balancer and are all accepting traffic, or you wish to have a hot standby that isn't accepting traffic in the event that the primary application server experiences a failure. In such mode, you will want to mount this directory to a network storage location (such as AWS EFS) and careful consideration should be given to the `supervisord` configuration, particularly to where log files and PID files are written and, if written to a shared storage, if the names of such files will collide and/or cause multiple processes to write to them.

Your options here vary depending on whether you wish these services to log to disk, to syslog, to stdout (e.g. Docker) and, if logging to disk, if you wish to log to the shared storage or not (e.g. for backup purposes). The simplest recommendation in a HA setup is to adjust the `LAB7LOGDIR` and `LAB7VARRUN` environment variables to some directory outside of the Data directory, though it's also possible to tailor logging in `supervisord.conf` to your exact requirements, such as including the `hostname` environment variable as part of the log filenames using the special `%(ENV_HOSTNAME)s` syntax.

In the event that the data volume should be mounted to a shared location but sharding occurs (e.g. the mount doesn't appear on one or more L7|ESP application servers for some reason and they begin writing these files to local disk), recovery should be as simple as merging the directories and files in these folders back together as all automatically generated file and directory names have either the database UUID they reference or a timestamp in their path or filename. Files that are

not dynamically generated, such as custom Python scripts that pipelines may execute should always be the same on each application server as these are usually placed there at install time.

1.21 Database User and Grants

The L7|ESP application doesn't need any permissions outside of its own PostgreSQL database so it is recommended to create a single database schema for the application, as well as a dedicated role with matching name, and grant all permissions for that role to that schema.

Currently, we don't separate "read-only", "read/write", "read/write/DDDL modification permissions" users at the application level. The same user is used in `l7 init` (DDL migrations) as in normal application execution. The credentials stored in the database config file `$LAB7DATA/conf/database.json` are used in both instances.

However, it is possible to switch these credentials with a more privileged account, run `l7 init` to perform any database migrations and then switch back to the less privileged account credentials for day-to-day application execution. Note that database migrations will only occur if you are upgrading the underlying L7|ESP version (e.g. L7|ESP 2.3.3 to L7|ESP 2.4) but won't take place during normal upgrades that only seed new content.

The database configuration file at `$LAB7DATA/conf/database.json` should take the following format when using an external hosted PostgreSQL solution such as AWS RDS:

```
{
  "host": "postgres",
  "port": 5432,
  "user": "l7esp",
  "pass": "password",
  "name": "l7esp",
  "start_service": false
}
```

Regarding encryption in transit, the PostgreSQL driver will automatically negotiate an appropriate connection based on the `SSL mode` defined on the database server.

1.22 Unix Application User

L7|ESP should be installed in Linux user-space and by default will only listen on user/registered ports, rather than system/well-known ports below 1024 that require root privileges.

The installation requires some basic tools, such as `make`, `wget`, `bzip2`, and `rsync`. This is enough for the installer to use Miniconda to bootstrap a Python environment where Ansible will be installed and automate the complete L7|ESP installation.

As `root`, you should install the basic requirements, and create a non-privileged Linux user/group for the application, and switch to that user:

```
root@l7espapp:~# yum install make wget bzip2 rsync
root@l7espapp:~# groupadd l7esp
root@l7espapp:~# useradd --groups l7esp l7esp
root@l7espapp:~# su - l7esp
```

As the L7|ESP user, you may fetch the deployment bundle, extract it and perform the installation:

```
17esp@17espapp:~$ wget "${ESP_DEPLOYMENT_BUNDLE_URL}"
17esp@17espapp:~$ tar xf "${ESP_DEPLOYMENT_BUNDLE_FILENAME}"
17esp@17espapp:~$ cd "${CUSTOMER_NAME}"
17esp@17espapp:~/${CUSTOMER_NAME}$ make install
```

To aide in system administration tasks, you may additionally wish to add the following lines to this Linux user's Bash profile so that common utilities will be made available via \$PATH the path environment variable, as well as some other useful environment variables:

```
17esp@17espapp:~$ cat ~/.profile
source /data/ESP/Lab7_ESP/current/bin/env.sh
PATH="/data/ESP/client/bin:$PATH"

17esp@17espapp:~$ which l7
/data/ESP/Lab7_ESP/current/bin/l7

17esp@17espapp:~$ which esp
/data/ESP/client/bin/esp

17esp@17espapp:~$ env | grep LAB7
LAB7LOGDIR=/data/ESP/Lab7_ESP/Data/log

17esp@17espapp:~$ tail -n 0 -f $LAB7LOGDIR/\*
==> /data/ESP/Lab7_ESP/Data/log/l7-esp.http.0.access.log <==
```

If for some reason, you need to migrate an existing L7|ESP instance to a different Linux account on the same machine (or any other machine for that matter), the steps would be as follows:

1. Stop the running L7|ESP instance with the `l7 stop` command.
2. If the L7|ESP server is installed in a user's home directory, move the entire server root directory to the target user's home directory.
3. Change file ownership on the entirely server root directory to the new Linux user/group (e.g. `chown -R 17esp:17esp /data/ESP`).
4. Start the L7|ESP instance back up again with the `l7 start` command.

1.23 Application Startup on Boot

The deployment bundle has the ability to install L7|ESP as a systemd service, however this will require the application user to be able to escalate to sudo privileges and that the Ansible variable `service: True` is set in the `roles/container.yml` file in your L7|ESP deployment bundle. You may also find the template used to create the service in the file `roles/esp/templates/service` and the related Ansible tasks can be found in the `roles/esp/tasks/run.yml` file. If you wish to make changes to any of this, please let us know so that we may include and test your desired defaults in future L7|ESP deployment bundle revision you receive from us.

1.24 L7|ESP Revision History

Simply put, all previous versions of workflows, protocols, and entities (samples) are stored in the database. In terms of old/new data values for a given field, L7|ESP captures the full history of changes in the `resource_action` table.

The `resource_action` database table provides an audit log for each resource in the system, which can typically be viewed under the History tab in the UI for a given resource.

When it comes to configuration items, such as workflow and protocol versions: a given version of a protocol (e.g.) is immutable. Each time you save a protocol or workflow, you are actually creating a new version of the protocol under the hood. The old version remains unmodified.

Currently, experiments use the most recent version of a workflow, in addition to the most recent version of the protocols nested in this workflow. Once submitted, the samples included in the experiment are “locked” into these versions - subsequent changes to the workflow or protocol definitions will not impact any “in-flight” samples.

1.25 Security Considerations

1.25.1 Introduction

This section covers topics related to protecting L7|ESP from accidental or malicious access and provides recommendations on overall security.

1.25.2 Ports and Protocols

The following ports should be opened between hardware and/or software firewalls between load balancer, application and database servers. (For example, security group rules and/or network ACLs in an AWS cloud-hosted environment, or `firewalld` software firewall on RedHat VMs, etc).

Table 1: L7|ESP Required Ports and Protocols

Protocol	Port Range	Purpose	Used By
TCP	443	Between Users and Load Balancer	HTTPS
TCP	8002	Between Application and Load Balancer	L7 ESP
TCP	5432	Between Application and Database Server	PostgreSQL

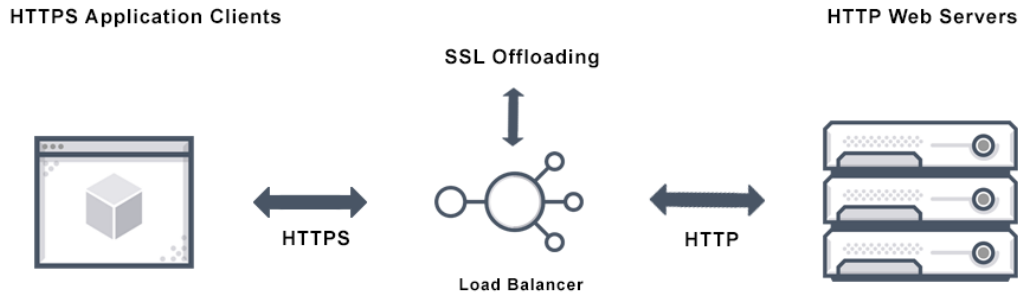
Key-based SSH access should also be configured for remote system administration on a port of your choosing (default is TCP 22). This will be required for initial software installation and occasional content deployments, but is not otherwise required for a functional L7|ESP system.

1.25.3 Accepting HTTPS traffic

To accept secure traffic, it is expected that the L7|ESP application will sit behind a TLS termination proxy. L7|ESP itself does not perform SSL offloading but provides some common configuration options you might expect from a typical web application in this regard.

There many are software options for performing SSL offloading, most advertise themselves as web servers, reverse proxies, load balancers, HTTPS front-ends, and similar terms. Popular options include HAProxy, Nginx, Apache but there are countless others such as, Caddy, Varnish, Pound, etc.

In a cloud setting, such as AWS, it is recommended to use one of the provided services if possible so you have one less component to maintain. For example, Amazon can perform SSL offloading on their load balancers (ELB/ALB) using certificates stored in ACM.



Referencing the illustration above, the load balancer would listen for HTTP traffic on ports 8002 and 443, and would have a DNS record associated with it so that users have a nice URL to use when visiting. In the following paragraphs, we will presume the DNS record is: `l7esp.example.com`.

Port 8002 is for HTTP traffic (ie. <http://l7esp.example.com> maps to `l7esp.example.com:8002`)

Port 443 is for HTTPS traffic (ie. <https://l7esp.example.com> maps to `l7esp.example.com:443`)

The HTTP endpoint should be configured to immediately redirect traffic to the HTTPS endpoint such that any insecure connections are converted to secure connections before the application may be accessed and users visiting `l7esp.example.com` without specifying the protocol will get the secure padlock icon.

Traffic to the HTTPS should be proxied to the L7|ESP application server behind the scenes. Assuming the DNS name of the L7|ESP application server is `L7ESPAPPSRV`, the proxy communication/HTTP traffic between `l7esp.example.com:443` and `L7ESPAPPSRV:8002` will be transparent and therefore unnoticed by the user.

While it is normally desirable to perform SSL offloading on another machine to alleviate the application server for the overhead of encrypting/decrypting HTTP traffic and security implications of store secure SSL certificates, it is certainly possible to install SSL offloading software directly to the application server.

Since L7|ESP is installed in Linux user-space and listens on user/registered ports, instead of system/well-known ports below 1024 that require root privileges, there is no chance of collision with HTTP software installed at the system-level (for example, using `sudo apt-get/yum install haproxy`).

This gives the administrator full ability to configure to surrounding system to their specific needs, if that includes installing SSL offloading or other software(s) onto the L7|ESP application server itself. Of course, this comes with extra maintenance overhead as that software will need configured separately from ESP.

If the administrator chooses this path, there will need to be some mechanism for providing the SSL offloading software with the required secure SSL certificates, but the L7|ESP listening port 8002 will no longer need exposed to the outside world as all traffic between that software and L7|ESP will be internal.

That said, it is still our suggestion that the activity of SSL offloading is performed outside the L7|ESP application where possible as it reduces the overall complexity of maintaining the L7|ESP application server.

1.25.4 L7|ESP Configuration

L7|ESP provides a number of configuration options that can be used to harden the application to your specific needs. The following sections will cover the most common security configuration options.

Web Server

web_server.blacklist A list of strings that, when sent in POST requests to L7|ESP, will be rejected by the server and result in 500 errors. This is to prevent XSS (cross-site scripting) attacks. The default list is: [">", "<", "&", "\"", "%3Cscript%3E"]. Normally, you won't need to change this, however it has been known to block valid requests so you may find it is necessary to add or remove items to tune it accordingly.

web_server.domain Used to configure the domain name where L7|ESP is hosted. You should set this to the DNS name of the server (e.g. `esp.example.com`) rather than `localhost`. If not set, the application will use `localhost` and `http://` as the base URL. This is used for generating links in emails and other places where the application needs to refer to itself. See the `web_server.high_security` setting if you are using `https://` (recommended).

web_server.https Used to configure the Secure attribute of cookies set by the L7|ESP server. When set to `true`, the Secure attribute will be set on cookies, which will prevent the browser from sending them over HTTP. This is useful if you are running L7|ESP behind a proxy that terminates SSL connections. If you are running L7|ESP without encryption, you should set this to `false` (default).

web_server.log_level While not specifically security-related, this setting can be used to tune the verbosity of the L7|ESP application logs. The default is `INFO`. Valid values are: `DEBUG`, `INFO`, `WARN`, `ERROR`, `FATAL`.

web_server.origin_whitelist Used to whitelist the only allowed origin of requests to the L7|ESP application. If this list is non-empty, the L7|ESP application will reject any requests that do not have an `Origin` header that matches one of the entries in the list.

web_server.use_high_security If set to `true`, this will prepend `https://` to the `web_server.domain` to form the base URL for the application. This is useful if you are running L7|ESP behind a proxy that terminates SSL connections. If you are running L7|ESP without encryption, you should set this to `false` (default).

web_server.ip_address_validation This setting is used to control whether or not L7|ESP will validate the IP address of the user when they log in. Not supported when users are logging in via SSO solution such as Keycloak.

web_server.ipware_header_precedence_order The IP address of a user is detected by checking the following HTTP request headers by default: ['HTTP_X_FORWARDED_FOR', 'X_FORWARDED_FOR', 'HTTP_CLIENT_IP', 'HTTP_X_REAL_IP', 'HTTP_X_FORWARDED', 'HTTP_X_CLUSTER_CLIENT_IP', 'HTTP_FORWARDED_FOR', 'HTTP_FORWARDED', 'HTTP_VIA', 'REMOTE_ADDR']. This setting allows you to define a different list of headers and the order to check them for the IP address of the user. L7|ESP will use the first matching header.

web_server.ipware_private_ip_prefix You may customize the prefixes to indicate an IP address is private. IP addresses matching a [predefined list](#) of prefixes are considered private & are **not** publicly routable.

System

system.concurrent_gui_user_sessions The maximum number of sessions a single user can have at a given time. This is to prevent a single user from logging into L7|ESP from multiple computers at the same time. The default is `-1` which is unlimited.

system.concurrent_client_user_sessions Same as above, however this is for non-GUI traffic - ie. Python scripts that use the L7|ESP API. The default is `-1` which is unlimited.

system.concurrent_session_handling How to handle sessions that exceed the above limits. Possible values `reject` to reject the new session with a “Too many sessions” message, or `kick` to log out the old session automatically. The default is `reject`.

Passwords

The following settings are available for configuring password requirements:

password.bcrypt_rounds The number of rounds to use when hashing passwords. The default is `13`. Increasing this number will make password cracking more difficult, but will also make login slower. The recommended value is `13`.

password.disallow_overlap Whether or not the password can contain you username or email address domain. The default is `false`.

password.expire_days The number of days after which a user’s password will expire. The default is `0` which is never.

password.min_length The minimum length of a password. Set to `0` (default) to disable.

password.require_digits Whether or not to require digits in passwords. The default is `false`.

password.require_mixed_case Whether or not to require mixed case in passwords. The default is `false`.

password.require_special Whether or not to require special characters in passwords. The default is `false`.

password.user_change_ok Whether or not the user can change their own password. Possible values: `never`, `notsso`, `always`.

Brute force protection

customize.max_login_failures The maximum number of failed login attempts before the user is locked out. The default is `3`.

customize.max_login_failures_timeframe The number of minutes to keep track of failed login attempts. The default is `5`.

login_failures_notify.enabled Whether to send notifications if brute force is detected. The default is `false`.

login_failures_notify.failure_threshold The number of failed login attempts before a notification is sent. The default is `6`.

login_failures_notify.notify_method The method to use for sending notifications. The default is `email`. Possible values are `email` for SMTP, `notification` for L7|ESP system notification, or `both`.

OpenID (SSO)

openid.verify_ssl_server This setting is used to control whether or not L7|ESP will verify the SSL certificate of the OpenID provider. This is useful if you are using a self-signed certificate for your OpenID provider. However, it is not recommended to disable this setting in production environments.

1.26 Backup and Disaster Recovery

1.26.1 Introduction

This document outlines strategies for backup, recovery and DR, as it pertains to the ESP server application. These suggestions can be implemented to achieve highly available (HA) infrastructure as well as help inform decisions when performing risk analysis related to your own internal business continuity (BC) guidelines and policies.

1.26.2 Risk Analysis

Determining the correct level of redundancy, HA, and backups should be defined by internal IT/DevOps/Business policies and processes to determine RTO (Recovery Time Objective) and RPO (Recovery Point Objective) values to meet any organizational defined SLO (Service Level Objective). One possible decision factor is the cost of implementing redundancy or HA versus lost revenue, reputation, and internal staff time associated with downtime.

1.26.3 Backup

Recommendations and examples

For atomic/PITR (*Point in Time Recovery*) backups, we recommend always backing up the database first, followed by the shared data volume (`/opt/l7esp/data`). This will ensure that all files referenced by the database backup are included in the whole backup.

At minimum L7 recommends whether performing a “hot” or “cold” backup:

- Keep copies of each deployment bundle that you deploy, since these contain information about the version of the software, as well as the configuration that was applied.
- Using `pg_dump --format="c"` for PostgreSQL database backups so you may be able to restore the backups with the `pg_restore` command. If you prefer to create backups in a different format with the `pg_dump` command, note that you will likely have to pipe the backup file into the `psql` PostgreSQL command-line utility to perform a restore.
- If the data volume exists outside of the default location (e.g. NFS/EFS mount) this should be backed up as well.
- Backup the deployed installation tarball to be able to reinstall if required.

One example of a backup strategy is to first backup the database to the shared data volume, then perform a backup of the data volume which will always result in atomic backups.

An option for AWS provisioned environments is the use of a managed service such as AWS Backup, which offers the following features:

- Centralized backup management
- Policy-based backup solution
- Tag-based backup policies

- Automated backup scheduling
- Automated retention management
- Backup activity monitoring
- Lifecycle management policies
- Incremental backups
- Backup data encryption
- Backup access policies
- Amazon EC2 instance backups
- Item-level recovery for Amazon EFS
- Cross-region backup
- Cross-account backup

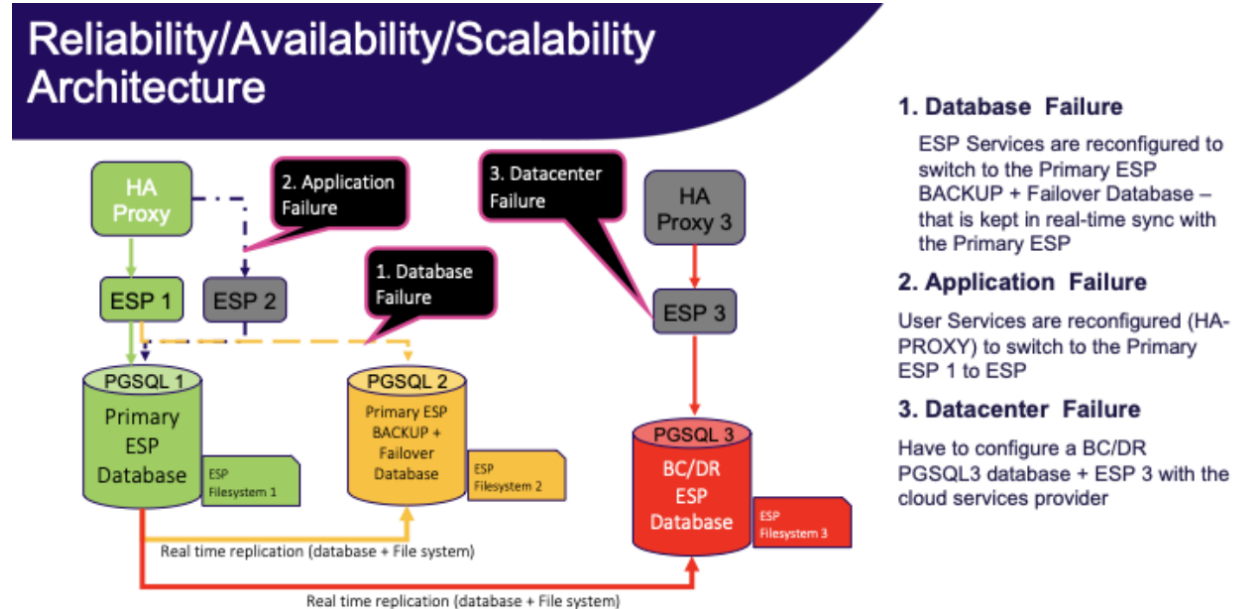
One thing to keep in mind when utilizing a managed service to perform a “hot backup” is that any RDS backup occurs before any EBS/EFS/NFS volumes to provide a valid restore point.

It is also highly recommended to regularly audit and test your backup/restore strategy to ensure it can be performed successfully as well as complies with any organizational policies and regulatory controls.

Common disaster-recovery instance requirements

- The disaster-recovery installation must have the same product version and patch level as the production installation.
- If any configuration or file changes (such as applying patches) are made to the production instance, the same changes must be repeated on the disaster-recovery instance.
- As the production system is used, all data changes must be replicated to the disaster-recovery instance. These changes can be database changes or file system changes, depending on the product in use.
- Replicating data changes imposes additional demands on the resources in the production system. To keep these demands to a minimum, the replication schedule should be carefully considered. If continuous replication is needed, the production system must be given additional resources (CPU and memory) to reduce the performance impact.

1.26.4 Disaster Recovery



Definitions

- **Hot standby** is a server that will automatically failover if the primary server fails.
- **Warm standby** is a server that will not automatically failover and that may not have all the latest transactions.
- **Cold standby** is a spare machine that needs to be turned on, backup restored (or even full staging of the machine).

Scenarios

Application failure

In the event of an ESP application server failure:

- **Hot standby:** Two ESP application servers behind a load balancer.

Note: User-uploaded files and pipeline scripts/log files that are referenced by the database are written as physical files to disk and are not stored inside the database as blobs for a number of reasons, such as performance. You should take care to mount these directories to a networked storage solution, such as an AWS EFS filesystem.

- **Warm standby:** Alternatively, you may cut traffic to another ESP application server of the same version, and sync the files in the ESP data volume using an out-of-band process, such as scheduled rsync. Using physical disks vs networked storage will increase file system performance and may reduce cost, but the file synchronization process will likely be eventually-consistent in nature.
- **Cold standby:** For this scenario, in the event of a failure, you can quickly bootstrap a new ESP application server either by installing a tarball on a fresh Linux server. Alternatively, if you create an image (e.g. AWS AMI) for an installed server, you can create a new VM from this image or even automate the process by using an AWS ASG with the correct policy and health checks.

Database failure

In the event of a PostgreSQL database server failure:

- **Hot standby:** The main use case for a hot standby is load balancing. You would use this to reduce the load on the database master server by delegating requests to one of more standby servers. To configure this, you must increase `wal_level` to `hot_standby` on the master database server and set `hot_standby` to on at the standby database servers. At a high level, most “clustered” PostgreSQL configurations can be considered hot standby, such as using more modern streaming replication modes.
- **Warm standby:** You can transfer a PITR backup to a standby database server and set it to always run an endless recovery process using WAL logs from the master database server. In this configuration, the standby database server is not accepting queries and sharing the load, but can be made available in the event of a failure. To configure this, you must specify `wal_level=replica; archive_mode=on` on the master database server and set `standby_mode` to on at the standby database server(s). When using a hosted service such as AWS RDS, you can simply enable the “Multi-AZ deployment” option when provisioning your database server.
- **Cold standby:** In this failure scenario, you would be configuring a new database server and restoring from backup. In AWS RDS, this would be equivalent to restoring a database instance from a snapshot.

Datacenter failure

In the event of a catastrophic datacenter (or regional, in cloud terms) failure:

- **Hot standby:** To achieve this scenario, you must have duplicate infrastructure running in another datacenter or region, with database and file synchronization between these sites. There is a cost vs risk tradeoff to be made as this can cost up to double the price, whereas the likelihood of this event may not mandate automated failover at this level.
- **Warm standby:** This is the same as a hot standby from a cost perspective as you would have duplicate infrastructure running, however you are failing over to the standby site by manually altering DNS records in the event of a failure.
- **Cold standby:** In this failure scenario, you would have a spare machine or the ability to provision one, in another network. The ESP application would need to be installed and backups restored to make it functional, and DNS updated to direct traffic here afterwards. In AWS, you could spin up a copy of the ESP infrastructure in another region using existing IaC in the unlikely event that this scenario ever occurs.

1.27 Questions/Feedback

For any questions about content development, testing, configuration, or anything else about the L7|ESP SDK, please contact L7 Informatics at support@l7informatics.com.