Supporting Characterisation Communities with Interactive HPC (Characterisation Virtual Laboratory)

www.cvl.org.au

Dr Lance Wilson (lance.wilson@monash.edu)
Dr Chris Hines (christopher.hines@uwa.edu.au)
Mr Jafar Lie (jafar.lie@monash.edu)
Dr Wojtek Goscinski (wojtek.goscinski@monash.edu)
Demonstration Outline

What is the CVL (Characterisation Virtual Laboratory)?

What is a virtual laboratory?

What is a workbench?

What do we provide?

What are tools we have developed? What are we reusing?

How do we manage research software? (How do we receive contributions from the community?)

How are we making CVL sustainable (Federation of sites nationally)
What is the CVL?

The **Researcher** view:
A single place where you can access your characterisation data and software tools, in a user friendly format.

**Project (C-DeVL) view:**
A program of work to connect Australian Characterisation instruments with data management environments, tools and analysis pipelines on the Australian research cloud.
CVL Desktop Unique Users per month (CVL@MASSIVE) 08/2016 - 04/2019

Total cumulative desktop sessions: 41,206
Unique users: 643
What do we provide?
(Interactive HPC)
One Caveat ....

$ ping m3.massive.org.au

PING m3.massive.org.au (118.138.254.157): 56 data bytes

64 bytes from 118.138.254.157: icmp_seq=0 ttl=45 time=178.006 ms
64 bytes from 118.138.254.157: icmp_seq=1 ttl=45 time=192.210 ms
64 bytes from 118.138.254.157: icmp_seq=2 ttl=45 time=177.006 ms
64 bytes from 118.138.254.157: icmp_seq=3 ttl=45 time=173.014 ms
Live Demo of Strudel Desktop V1
Strudel Desktop V1 - Login process
Strudel Desktop V1 - Desktop Usage
Strudel Web & Desktop
Strudel Web & Desktop

Web interface
desktop.cvl.org.au

SSH AuthZ

LDAP

Select a remote system

CVL@UWA (BETA)
CVL (AAF)
CVL (LDAP)
DesignHouse@MASSIVE
M3 and CVL@MASSIVE
Strudel Web & Desktop

- Web interface or desktop app
- Job Control
- Guacamole
- TurboVNC (Strudel Desktop)
Live Demo of Strudel Web V1
Strudel Web V1 - Login Process
What are tools we have developed?
Strudel V2 User flow

Steps
1. Website delivered
2. Login
3. Create job/s and tunnels
4. Proxy web sockets

Frontend/CDN

Web browser

Identity/Authentication

OIDC

SSHAuthZ

Batch Job Builder

Site #1

Docker Backend

TES

TWS

Cluster

Job Scheduler

Web App (Jupyter)
Live Demo of Strudel Web V2
Strudel Web V2 - Login Process
How do we manage research software? (How do we get contributions from the community?)
Federated Research Software Stacks

- Procedures for contributions have been developed ([link](https://github.com/Characterisation-Virtual-Laboratory/CharacterisationVL-Software))
- A public repository is now available for container build scripts

  https://github.com/Characterisation-Virtual-Laboratory/CharacterisationVL-Software

- 40 Packages have been committed to the repositories (and growing)
- Pre-built containers are available from a public repository (Singularity Hub)
- Ongoing work to accommodate disparate clusters

What communities (new and/or existing) are important to you?
How are we making CVL sustainable (Federation of nodes nationally)
CVL “Cluster in a Box” Deployment Technology
Timeline of development and reuse

C-DeVL Activities

MASSIVE M3

CVL on NeCTAR

MonARCH

CVL@MASSIVE

CVL@UWA

2014

2015

2016

2018
Strudel Deployments

- University of Queensland
- University of Western Australia
- Central Queensland University
- NCI
- Pawsey
- Julich
- TERN
- TPAC
- Swinburne University
- University of Melbourne
- University of Southern Queensland
Reusable Software and Infrastructure

- Strudel and Strudel Web
- MyTardis ( www.mytardis.org )
- AuthZ ( authentication certificates )
- Ansible scripts for all services/applications
- Data repatriation scripts from ANSTO Australian Synchrotron
- Software containers
Demo

Strudel Web V1
How do we manage research data?

MyTardis Ecosystem

https://store.synchrotron.org.au/dataset/54331
https://store.erc.monash.edu/dataset/13
Live Demo

MyTardis - Monash University
What is MyTardis?
A data management system for instruments - http://mytardis.org/

MyTardis is a data management system, developed by Monash, for instrument data.
The Monash instance of the MyTardis service is called Store.Monash
Store.Monash
Monash University instance of the MyTardis

94 Instruments

12 Million Data Files

0.75 PB Managed Storage

1974 Users

68,189 Datasets

1023 Groups
My Tardis Version 4.1: New Features

**Scalable Platform**
- Dynamic Priority Queue
- Sustainable / Scalable Data Storage
- Monolithic to Microservices

**Easier to Operate and Maintain**
- Automated Provisioning, Deployment, and Configuration Management
- Automatic Scaling
- Better Monitoring & Log Aggregation

**Better User Experience**
- Search Improvements
- User Interface Improvements
- Client-side interaction Improvements

**Better Monitoring & Log Aggregation**

**Search Improvements**

**Client-side interaction Improvements**
Acknowledgments

https://github.com/Characterisation-Virtual-Laboratory
https://docs.cvl.org.au
help@cvl.org.au

http://www.mytardis.org/
https://github.com/mytardis/mytardis
https://mytardis.readthedocs.io
store.star.help@monash.edu

CVL and MyTardis has been supported by:

Australian Research Data Commons
Summary

● **Reusable Components**
  ○ Strudel Desktop and Strudel Web ([desktop.cvl.org.au](http://desktop.cvl.org.au))
  ○ MyTardis ([www.mytardis.org](http://www.mytardis.org))
  ○ AuthZ ([https://github.com/monash-merc/ssh-authz](https://github.com/monash-merc/ssh-authz))
  ○ Software containers -
    [https://github.com/Characterisation-Virtual-Laboratory/CharacterisationVL-Software](https://github.com/Characterisation-Virtual-Laboratory/CharacterisationVL-Software)

● **Federation**
  ○ [www.cvl.org.au](http://www.cvl.org.au)
  ○ [https://github.com/Characterisation-Virtual-Laboratory](https://github.com/Characterisation-Virtual-Laboratory)