

SERL Technical overview

Darren Bell – Associate Director Technical Services

Energy Data for Research 13 July 2020

Copyright © 2020 UK Data Service.





SERL infrastructure in 6 icons





Participant Portal

- Participant Portal live in August on Amazon Web Services <u>https://serl.ac.uk/portal</u>
- All data encrypted and site has been pen tested
- Currently holding around 1700 consent records
- AWS "Serverless" architecture





AWS "Serverless" architecture





Hadoop at the heart of the data store

- HDP Hortonworks Data Platform 3.1
- Hadoop is a *suite* of different products (like Office is a suite of Excel, Access, Word, Powerpoint, Publisher etc.)





The bits of Hadoop we are using



• On top of HBase, we use JanusGraph for querying



Technical Infrastructure – data ingest (1)

- Messaging pipeline for Participant Portal triggers onboarding of smart meter data
- Basic goal is to have as minimal human intervention as possible





Technical Infrastructure – data ingest (2)

 Once DCC schedules are set up, smart meter data is retrieved daily over a secure, encrypted connection





Building capacity

• Load Balancing Spreading the data ingest load over more machines





Building resilience

- Dashboards
- Alerting systems







Inbound data challenges

- Not just readings!
- Duplicate postings
- Missing postings
- Vendor-specific issues

<pre><?xml version="1.0" encoding="UTF-8" standalone="yes"?></pre>
SOLResponse schemaVersion="3.0" xmlns="http://www.cgi.com/DccAdapter/SOL/ServiceResponse">
<pre></pre>
<responsecode>IO</responsecode>
<responsedatetime>2020-03-05T03:13:33.954Z</responsedatetime>
-
E <body></body>
<pre><responsemessage></responsemessage></pre>
<pre><servicereferencevariant>D.7.3</servicereferencevariant></pre>
<pre></pre>
<pre><deviceid>0F-C1-75-FF-FF-DF-43-B4</deviceid></pre>
<dspscheduleid>32345</dspscheduleid>
CDCCAlert>
<dccalertcode>N53</dccalertcode>
-
<pre>- </pre>
-
L



DCC data challenges

- As one of the first "Other User" clients, we are subject to significant teething challenges
- Battling "Alert storms" from the DCC at one point 2/3 of the postings were alerts – 14,000 in addition to the 7000 postings we were expecting. The DCC have now rectified this.
- Historic data has been problematic meter devices were to cope with one call for 13 months historic data: it was to break this down into 13 individual monthly calls.
- The reality has been different many devices will not return data for a request for in excess of 2 days of data.
- Creates verbose workflows which run-up against restrictive (in our opinion) thresholds set on the DCC Adapter



Technical Infrastructure – Researcher Portal

- Now that consent and ingest has commenced and is relatively stable, more focus on front-end
- Secure lab infrastructure on AWS workspaces

SMART ENERGY RESEARCH LAB	Dashboard Data resou	irces Projects	Secure labs	Data product builder	Submissions	Searches	A Darren Bell
Library Catalogue	Build data product "	SERL-EPC 1"	0				💾 Save progress
My workspace	Selec	1 2	yerage Identify li	A Filter	5 Aggregate	6 Build	
Learn							
Developers	Previous	to more data	Next				
Access	Select variables to filter on for Case count: 11, SMRP Observatory data linked to EPC Drag and drop variables here Variable count: "						
	Variable name	Source dataset	Variable type	Selected filter varia	ble	Filter to apply	
	Electricity Reading (epv)	SMRP Observatory Data	Continuous	Gas Reading (SMRP Observatory	<u>Data</u>		•
	Gas Reading (gpv)	SMRP Observatory Data	Continuous	HEAT LOSS (EPC	CORRIDOR	No corridor	× Unheated corridor × ▼
	HEAT LOSS CORRIDOR	EPC	Category				
_	Timestamp_(ets)	SMRP Observatory Data	Text	_			
	Variable 5	EPC	Category				
	Variable 6	SMRP Observatory Data	Text				
	Pi	revious	Next				
	«	1 2 3 4 5 6	789»				



Q2 Researcher Portal development underway

 Datasets Browse and Submit Project functionality completed so far <u>https://serl.ac.uk/researcherportal</u>

SMRP Researcher Portal	× +		- 0
$\cdot \rightarrow C$ (1) localhost:30	000/workspace/NewProject		☆ → 🔀 😐 📑 📀
SMART ENERGY RESEARCH LAB			<mark>ک</mark> Logout
brary	Create new project	SUBMIT FOR APPROVA	L SAVE CHANGES CANCEL
Data Documents	Key Information		^
ly Workspace	Application Type	Research Project	
Projects	Title	Title of research project	
	Project start	dd/mm/yyyy	
	Project end	dd/mm/yyyy	
	AR number	If you have an Accredited Research number, please enter it here	
		To become an Accredited Researcher, please complete the Accredited Researcher application form available on the <u>UK Statistics Authority website</u> .	
		NB The project lead and all members of the research team must have Accredited Researcher Status and an associated AR number before a project application is submitted.	
D Type here to search	th 🗄 🧲	👝 👩 🚓 刘 🚓	へ \$⊃ <i>(</i> ∉ \$)) ENG ^{15:25}



Key activities in Q2/Q3

- Onboarding full complement of weather data
- Addition of tariff data to researcher datasets
- Completed historic data collection (extremely onerous)
- First static datasets generated for UKDA Secure Lab
- Implementing infrastructure for cloud-based secure desktops (Amazon Workspaces)
- Optimizing workflows for CoT/WoC
- Optimizing workflows for Reporting



Key activities in Q3

- Wave 2 Onboarding and scaling up for 5k households
- Ingest of Covid-19 survey from Qualtrics
- Researcher Portal MVP
- Complete technical documentation
- Possibly more SENS data

