

# **Southern California Earthquake Data Center (SCEDC) 2019 Operations**

Caltech Technical Report for SCEC Award #19061

Investigators: Robert W. Clayton (Caltech), Ellen Yu (Caltech)

Submitted March 27, 2020

## Technical Report

1. Archived continuous seismic data for 583 stations and parametric and waveform data for 61,236 local events and 227 teleseismic earthquakes. (Figure 1)
2. The SCEDC continued to support its web services and ensure they are FDSN compliant. Our statistics show that in 2019, there was a large increase in the number of requests made with the STP client. FetchData and Obspy were popular FDSN web services methods. (see Figure 2).
3. The SCEDC hosted its continuous waveform archive (100TB) in the cloud as part of the Amazon Open Data Set program. Tutorials on how to get started are available at <https://scedc.caltech.edu/cloud/> Tim Clements and Marine Denolle at Harvard University ran a pilot project with the data set and whose work was also presented at the 2019 SCEC Annual meeting. <https://nextjournal.com/thclements/cactus-to-clouds/#accessing-the-scedc-dataset-on-aws>  
Our statistics show that 30,693,886 requests were made against the public data set in 2019.
4. The SCEDC loaded data from portable USGS stations that were deployed in response to the 2019 Ridgecrest aftershock sequence, into its archive. These data are available as both continuous and triggered waveforms, and event phases.  
<https://scedc.caltech.edu/research-tools/ridgecrest-info.html>
5. The SCEDC also implemented new web services – the data availability query endpoint. This service, modeled after the IRIS availability web service, shows the start and end times of continuous waveform segments present in the archive.
6. To further assist SCEC researchers in data discovery – the SCEDC re-vamped its clickable station map. <https://service.scedc.caltech.edu/SCSNStationMap/station.html>. Users can filter based on time and instrument type. They can also download the search results into a csv file.
7. To improve waveform completeness, the Data Center has implemented an automated gap fetching system for SCSN stations using Q330 loggers. The software will determine gaps in the waveform archive, and then make a request to the logger in the field for the duration the gap. If it is able to retrieve data, it will populate the gap. During this period the SCEDC recovered 436,672 hours from 244 stations through this method.
8. The SCEDC is assisting a SCEC funded project to associate hypocenters with faults in the SCEC Community Fault Model. Software from this project was configured to run automatically on suitable events detected by the SCSN, producing an email message that is sent to project participants for further feedback and improvement.
9. In keeping with best data security practices, Data Center staff regularly patched operating systems and databases throughout the reporting period. The system remained functional throughout the 2019 Ridgecrest sequence.
10. The SCEDC continued to make improvements to the Station Information System (SIS) with the Southern California Seismic Network (SCSN). We have completed developing the user

interface to track all data hops from a station to processing center. Efforts continue for loading telemetry data paths for the remaining stations and developing code to distribute this information. This information will help researchers in Earthquake Early Warning studies analyze how complex telemetry paths can affect data retrieval from stations as well as possible changes in network detection capability, which will be useful in CSEP testing.

11. The SCEDC continued to host the developmental database SCEC WGCEP group for UCERF3.

### **Presentations and Publications**

Yu, E., P. Acharya, A. Bhaskaran, S. L. Chen, J. Andrews, V. Thomas, Z. Ross, E. Hauksson, R. Clayton; Cloud Computing and Big Data – Using the Southern California Earthquake Data Center (SCEDC) and the Southern California Seismic Network (SCSN) Products and Services for Earthquake Research; (Abstract; Poster 301, SCEC Contribution 19061) presented at the *2019 SCEC Annual Meeting Palm Springs, 11 Sept. 2019*.

Yu, E., P. Acharya, A. Bhaskaran, R. Bhadha, S. L. Chen, J. Andrews, V. Thomas, Z. Ross, E. Hauksson, R. Clayton; Southern California Earthquake Data Now in Amazon Cloud; (Abstract; Poster S21H-0630) presented at the *2019 AGU Annual Meeting San Francisco, 10 Dec. 2019*.

2019 Stats

In 2019, the SCEDC archived:

- 61,248 local events
- 76,412,592 triggered waveforms
- 2,180,326 arrivals
- 9,365,478 amplitudes (strong motion)

Magnitude	Number of local events (le):
-1.0 - 0.0	317
0 - 1	24598
1 - 2	29236
2 - 3	5905
3 - 4	1119
4 - 5	108
5 - 6	3
6 - 7	1
7 - 8	1

2019 event type break down:

# events:	Event type
61,248	le (local event)
653	qb (quarry blast)
116	re (regional event)
218	ts (teleaseism)
4	ex (explosion)
2	th (thunder)
62,250	Total

Data transferred via STP client in 2019:

	waveforms	waveforms/day	GB	MB/day	KB/sec
2018 Q4	35,999,347	391,297	1,048	11,669	138
2019 Q1	86,622,845	962,476	1,968	22,396	265
2019 Q2	38,119,117	418,891	659	7,413	88
2019 Q3	164,669,896	1,789,890	4,917	54,725	649
2019 Q4	59,466,672	646,377	1,876	20,886	248

Data transferred via Web Services in 2019:

	Event (MB)	Station (MB)	Dataselect (GB)	WebSTP (GB)	RESP (MB)	SACPZ (MB)
2018 Q4	19,693	72,253	8,872	473	20.03	0.24
2019 Q1	41,547	99,684	6,857	298	2.22	0.35
2019 Q2	92,327	77,789	3,370	400	418	1.74
2019 Q3	153,729	176,305	1,470	805	1.30	1,263
2019 Q4	60,013	193,647	2,609	122	56	35

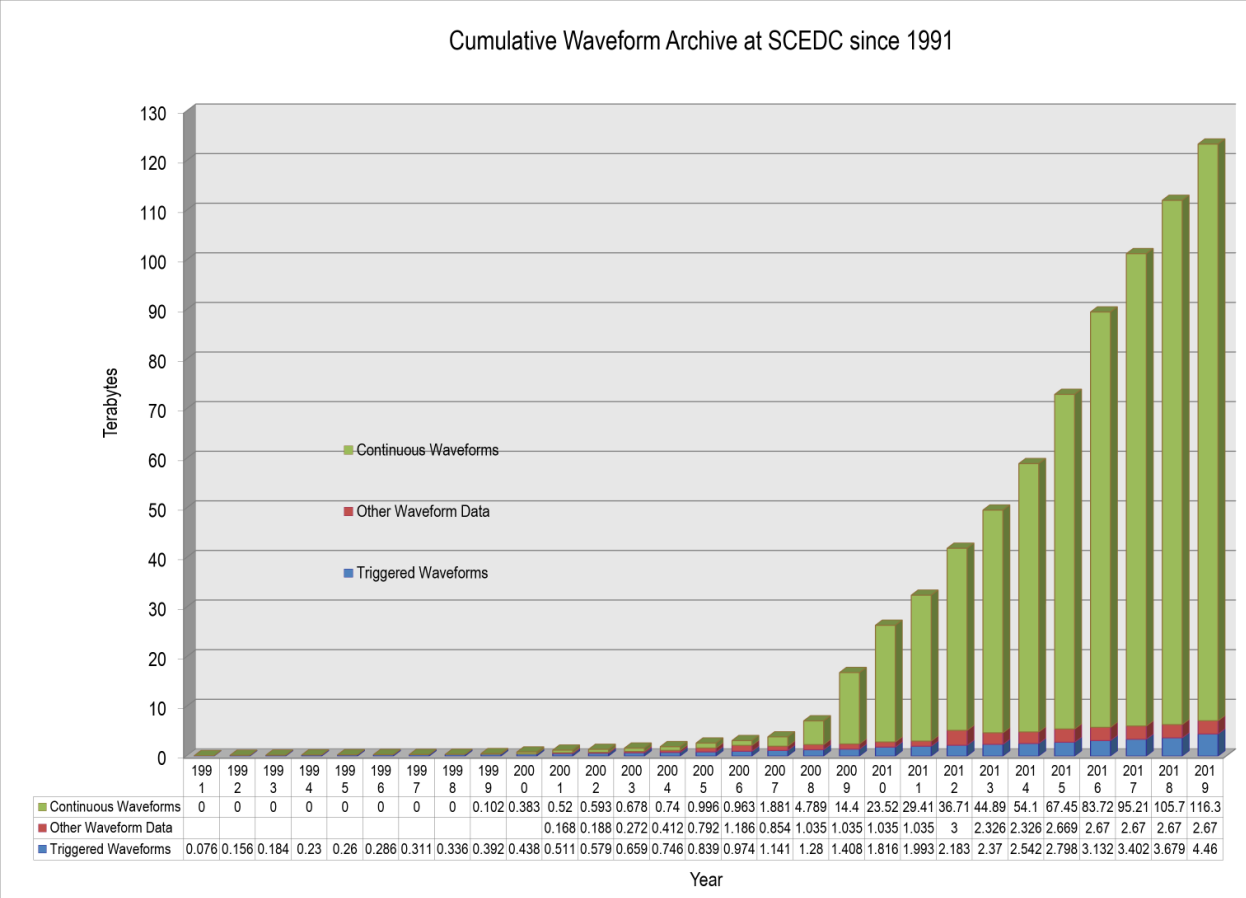
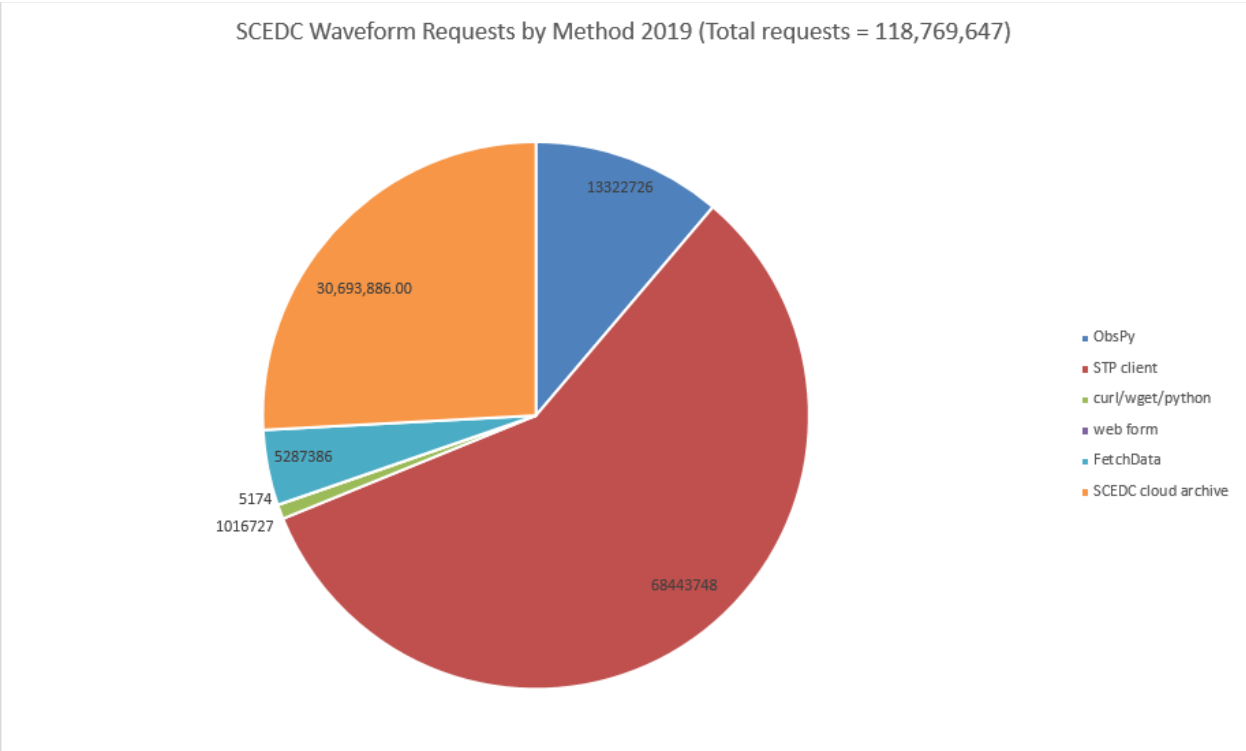


Figure 1. (Exemplary Figure) Data volumes stored at the SCEDC for seismological research.



**Figure 2.** Access to the SCEDC waveform requests by method. Except for the STP client, and SCEDC cloud archive, all other methods utilize SCEDC's FDSN dataselect webservice.

**Figure 3a.**

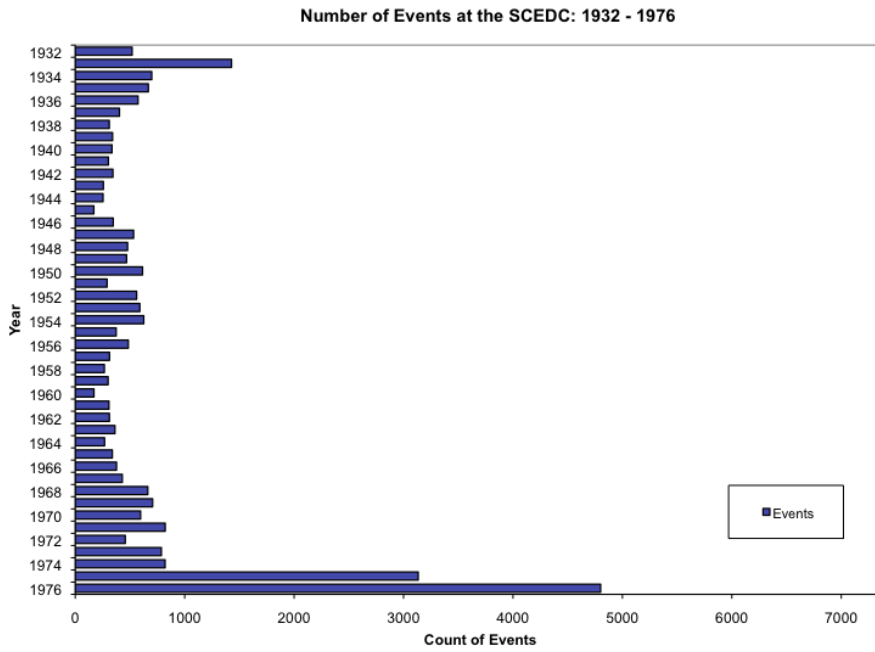


Figure 3b.

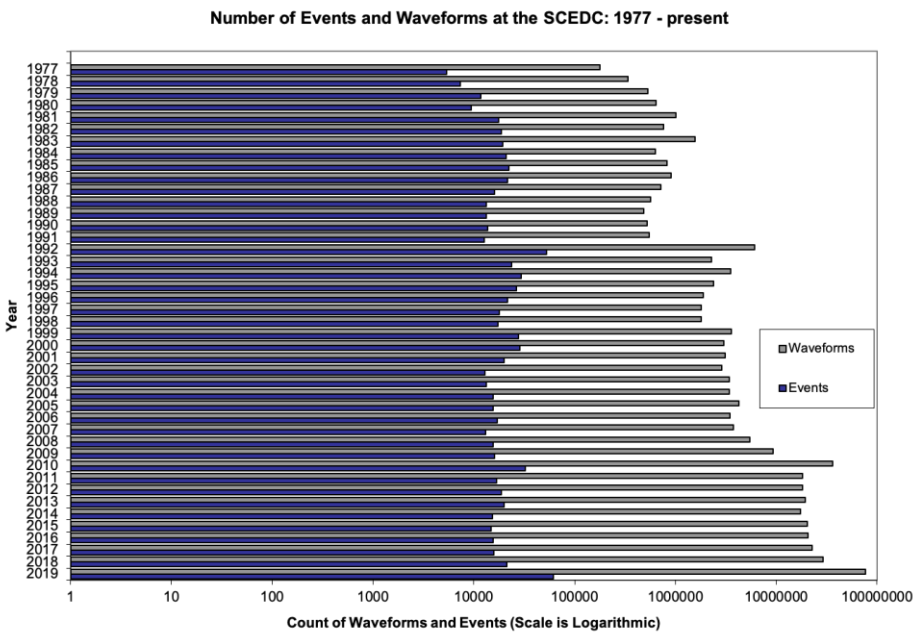


Figure 3a and b. Number of events and event associated waveforms in the archive over time