Integrated Analysis of Room Performance

Michael Novak

May 28, 2012

0.1 Automated Analysis

The figures and tables in this document were created by analysis.hs; functions acting on the raw data produce LaTeX & gnuplot files, which are rendered as tables and figures; the colors used in the plots were the carefully chosen colors[1] to be distinguishable.

- Postmortem data gotten from /u/karavan/Public at rita.cat.pdx.edu

- 2011-05-05—The Low Density Workload—LD_A1_56p_2ppn_28n_IO-BASIC_even_RAWDATA

- 2011-04-25—The High Density Workload—HD_A1_224p_8ppn_28n_RAWDATA

0.2 Data

We've been given postmortem data, LD_A1_56p_2ppn_28n_IO-BASIC_even_RAWDATA, corresponding to the experiments described in [2]:

- Low Density Even Distribution – In this set, we used racks B1, B2, B3, and B4; and we scheduled two MPI tasks per node, using all of the nodes in all four racks.

- In the low density LD schedules, the 224 processes executed on 112 nodes, with two processes per node, and the processes were spread across racks B1, B2, B3, and B4.

- The nodes are connected via a DDR InfiniBand 1-layer fat tree; one rack (A2) contains the networking equipment and seven racks (A1, A3–4, B1–4) are for the compute nodes.

- *Power.* We collect power per rack, at 10 second intervals. This is converted by FRED to kWh.

- *Cooling.* Each spray-cooled rack has a Thermal Management Unit (TMU) which contains pumps, a liquid-to-liquid heat exchanger, and control electronics.

0.3 Power & Cooling

Figure 1 shows power consumption, in kWh (7.63¢ per kWh, in Oregon), for the racks used in the trials. It was hard to tell whether the gradient of one line was steeper than another; hence the accompanying tables were created to show overall kWh increase per trial, and to further scrutinize power consumption within the racks.

The tables, within Figure 1, can be used to compare how the two air-cooled racks (A_1, A_2) faired, energy performance-wise, against the water-cooled racks $(A_1, B_1, \ldots, B_4)$. Note: our data seems to be corrupted by faulty sensor readings; at some nodes, in certain racks, sensor readings gave negative temperature values; these readings were discarded; so the average temperature does not include every node in every rack; and, most temperature readings remained static with time, which made calculating the average trivial.

For the water-cooled racks, Figure 2 is a visualization of TMU utilization for each trial (t01–03)[2]: the degree of cooling which the TMU provided, captured by the water-temperature-delta metric; and the mass flow rate of water, captured by the manifold pressure metric. TMU utilization metrics were collected by FRED at 10 second intervals; so, for simplicity, each trial's time series starts at 0 and is incremented accordingly.

---

[1]http://www.sron.nl/~pault/colourschemes.pdf

[2]Within RAWDATA, the application and system level data for each trial are in the trial directories: t1, t2, t3.

0.3.1 Chillers

Within the cooling infrastructure, most of the energy is spent on chillers, which refrigerate the coolant—water, in our case—used to extract heat from the equipment in the data center. The following metrics, derived from [3], are displayed in Figures 3 & 4. Chiller utilization measurements were collected by FRED at 30 second intervals; so, for simplicity, each trial's time series starts at 0 and is incremented accordingly. Note: only three chillers units (C_1, C_2, C_4) were turned on during the trials, as evidenced by KW consumption—C_3 and C_5 were turned off.

- **IT cooling load.** This is the amount of heat that is generated (and thus needs to be dissipated) at a data center. It is approximately equivalent to the power consumed by the equipment since almost all of it is dissipated as heat. It is commonly specified in kilowatts (KW).

- **COP.** The coefficient of performance (COP) of a chiller unit indicates how efficiently the unit provides cooling, and, is defined as the ratio between the cooling provided and the power consumed, i.e.,

$$COP_i = \frac{L_i}{P_i} \tag{1}$$

  where L_i is the cooling load on the ith chiller unit and P_i is the power consumed by it.

- **Chiller utilization.** Chiller utilization depends on the degree of cooling provided; i.e., the difference between the inlet and outlet water temperatures.

0.4 Data Mining

If only we had access to more data...

- It would be nice if FRED measurements were stored in a NoSQL graph database. A NoSQL database would be much faster than PostgreSQL. And having the data structured as a graph seems like it would make the data easier to navigate.

- We could compare high-density workload with low-density workload; and , perhaps medium-density workload—if it exists.

- We could try to use the temporal data-mining techniques in [3, 4] on TMU and Chiller data. But we might need longer time slices; no motifs can be seen at such a fine granularity—within just 3 sets of 9 minute trials—clustering would just produce nonsense.

- [4, p. 34:7]

  Motifs are repetitive patterns of occurrence in time-series data. In understanding multivariate time-series data about chiller utilizations, we seek to identify motifs that underly how different chillers are involved in meeting the varying demand posed by data centers. We would like to identify regions of time-series progression that demonstrate better/improved sustainability measures than others.

- [3, pp. 1308–1309]

  A multivariate time series $T = \langle t_1, \ldots, t_m \rangle$ is an ordered set of real-valued vectors of a particular variable. Each real-valued vector t_i captures the utilizations across all the chiller units.

- You look at the time series to detect motifs; the number of motifs is used in k-means clustering. For example, if you spot 5 motifs [4, p. 34:12]:

First, the time information is stripped from the data and clustering is performed with number of clusters, $k = 5$. The cluster transitions are overlaid in the plot. These transitions are used to encode the multivariate numeric data into symbolic form and serve as the input to frequent episode mining. One episode is discovered which occurs five times and this episode is unpacked into the original data and overlaid on the dataset as shown. This serves as an example of a motif that we embedded and were able to uncover.

- [3, p. 1307]

  High amplitude and frequent variations in utilization due to varying load or some failure condition result in decreased lifespan, and, thus, need to be minimized.

- Algorithms can be developed for processing multivariate time-series data to characterize sustainability measures of the patterns mined.
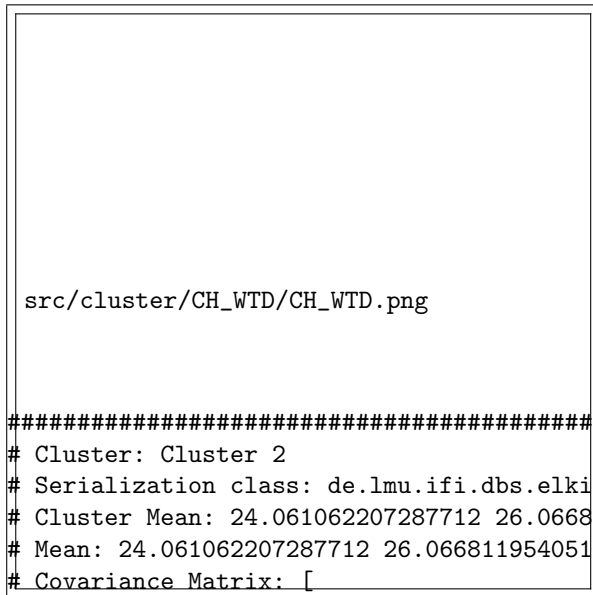
For clustering, I'm using ELKI (release 0.4 [1]). The algorithms being used are: k-means and expectation maximum (EM). ELKI has nice visualization features; for instance, pictured to the left are the Gaussian clusters discovered by EM clustering with $k = 3$, on the three trials for the Chiller water-temperature delta data.

I've labelled the input data in such a way that the time-series information is retained after clustering; so the cluster transition can be overlaid in the corresponding plot. As an example, here's what EM clustering output looks like (data from ManifoldPress—Figure 2):

src/cluster/CH_WTD/CH_WTD.png

```
###############################################################
# Cluster: Cluster 2
# Serialization class: de.lmu.ifi.dbs.elki.data.model.EMModel
# Cluster Mean: 24.061062207287712 26.066811954051424 24.640384129310334 24.961621059814675 22.481076640
# Mean: 24.061062207287712 26.066811954051424 24.640384129310334 24.961621059814675 22.48107664091767
# Covariance Matrix: [
#   [  0.00204516, -0.00007624,  0.00010518,  0.00067192, -0.00217584 ]
#   [ -0.00007624,  0.00030667, -0.00026587,  0.00044926,  0.00030202 ]
#   [  0.00010518, -0.00026587,  0.00093819, -0.00030461, -0.00072843 ]
#   [  0.00067192,  0.00044926, -0.00030461,  0.00338988, -0.00047919 ]
#   [ -0.00217584,  0.00030202, -0.00072843, -0.00047919,  0.00426668 ]
# ]
ID=324 24.13 26.04 24.65 24.94 22.34 t03_SEC_530
ID=323 24.07 26.08 24.67 25.0 22.39 t03_SEC_520
ID=320 24.07 26.09 24.66 25.01 22.46 t03_SEC_490
ID=321 24.06 26.08 24.63 25.01 22.41 t03_SEC_500
ID=307 24.03 26.1 24.61 25.02 22.57 t03_SEC_360
```

$\vdots$

And here's what k-means output looks like (data from COP—Figure 3):

```
###############################################################
# Cluster: Cluster 0
# Serialization class: de.lmu.ifi.dbs.elki.data.model.MeanModel
# Cluster Mean: 0.39636290564044135 0.39964999744800694 0.3177650733048521
```

3

```
ID=1 0.3962167689161554 0.40261407398762983 0.31213983864771416 t01
ID=2 0.3962167689161554 0.40261407398762983 0.31213983864771416 t01
ID=7 0.3962167689161554 0.40260509177027826 0.32062181199902845 t01
ID=8 0.3962167689161554 0.40260509177027826 0.31574415155719276 t01
ID=9 0.4026073619631902 0.40260509177027826 0.3109601492608716 t01
```
.
.
.

I've started putting together experiments in `src/clustering`; the Haskell script, which feeds input into ELKI, is cluster.hs. Meaningful clustering will have to wait until we get more data; as, for example, in [4, pp. 34:14–34:16 (see Tables I–V)].

Low-Density Workload

PowerUnit PowerKWH	A1	A3	A4	B1	B2	B3	B4
t01	1.28	1.03	1.07	1.11	1.17	1.31	1.27
t02	1.28	1.03	1.06	1.11	1.17	1.31	1.27
t03	1.30	1.05	1.09	1.13	1.19	1.33	1.35
Total PowerUnit	1	1	1	1	1	1	1

Total kWh increase.

CPU Temp	A1	A3	A4	B1	B2	B3	B4
t01	90.49	82.88	82.94	99.91	90.40	101.47	108.81
t02	90.49	82.88	82.94	99.91	90.40	101.47	108.81
t03	90.49	82.88	82.94	99.91	90.40	101.47	108.81
Total CPU	37	40	4	5	6	4	5

Average Per-Node CPU Tempurature Average for Each Rack

CPU Temp	A1	A3	A4	B1	B2	B3	B4
t01	92.61	141.70	85.69	127.22	92.47	112.55	114.63
t02	92.61	141.70	85.69	127.22	92.47	112.55	114.63
t03	92.61	141.70	85.69	127.22	92.47	112.55	114.63
Total CPU	37	40	4	5	6	4	5

Maximum Per-Node CPU Tempurature Maximum for Each Rack

NodeAir Delta	A1	A3	A4	B1	B2	B3	B4
t01	26.10	24.44	26.30	25.51	24.55	23.51	26.39
t02	26.10	24.44	26.30	25.51	24.55	23.51	26.39
t03	26.10	24.44	26.30	25.51	24.55	23.51	26.39
Total NodeAir	24	24	22	25	23	23	16

Average Per-Node Air Temperature Delta for Each Rack

Table 1: Rackwide Per-Node Metrics

PowerUnit PowerKWH	A1	A3	A4	B1	B2	B3	B4
t01	13.35	9.89	10.24	10.73	11.12	13.07	14.34
t02	6.87	4.25	4.41	4.56	4.86	5.69	6.14
t03	18.90	12.08	12.48	12.95	13.73	16.13	17.51
Total PowerUnit	1	1	1	1	1	1	1

Total kWh increase.

CPU Temp	A1	A3	A4	B1	B2	B3	B4
t01	97.18	83.52	81.71	100.15	89.51	94.54	110.43
t02	104.83	83.99	82.30	100.44	89.75	94.22	111.05
t03	116.73	83.89	82.26	100.42	89.68	97.33	111.66
Total CPU	37	41	4	5	6	4	5

Average Per-Node CPU Tempurature Average for Each Rack

CPU Temp	A1	A3	A4	B1	B2	B3	B4
t01	107.68	141.97	85.05	127.94	91.73	106.14	120.02
t02	126.13	142.13	85.20	127.94	91.86	106.93	120.95
t03	126.13	142.13	85.20	127.76	91.86	108.96	120.95
Total CPU	37	41	4	5	6	4	5

Maximum Per-Node CPU Tempurature Maximum for Each Rack

NodeAir Delta	A1	A3	A4	B1	B2	B3	B4
t01	26.80	26.20	26.62	24.52	25.28	24.37	26.24
t02	27.43	26.06	26.44	24.76	25.28	24.43	26.92
t03	29.16	25.87	26.66	24.16	25.30	24.38	27.17
Total NodeAir	23	23	21	25	23	23	15

Average Per-Node Air Temperature Delta for Each Rack

Table 2: Rackwide Per-Node Metrics
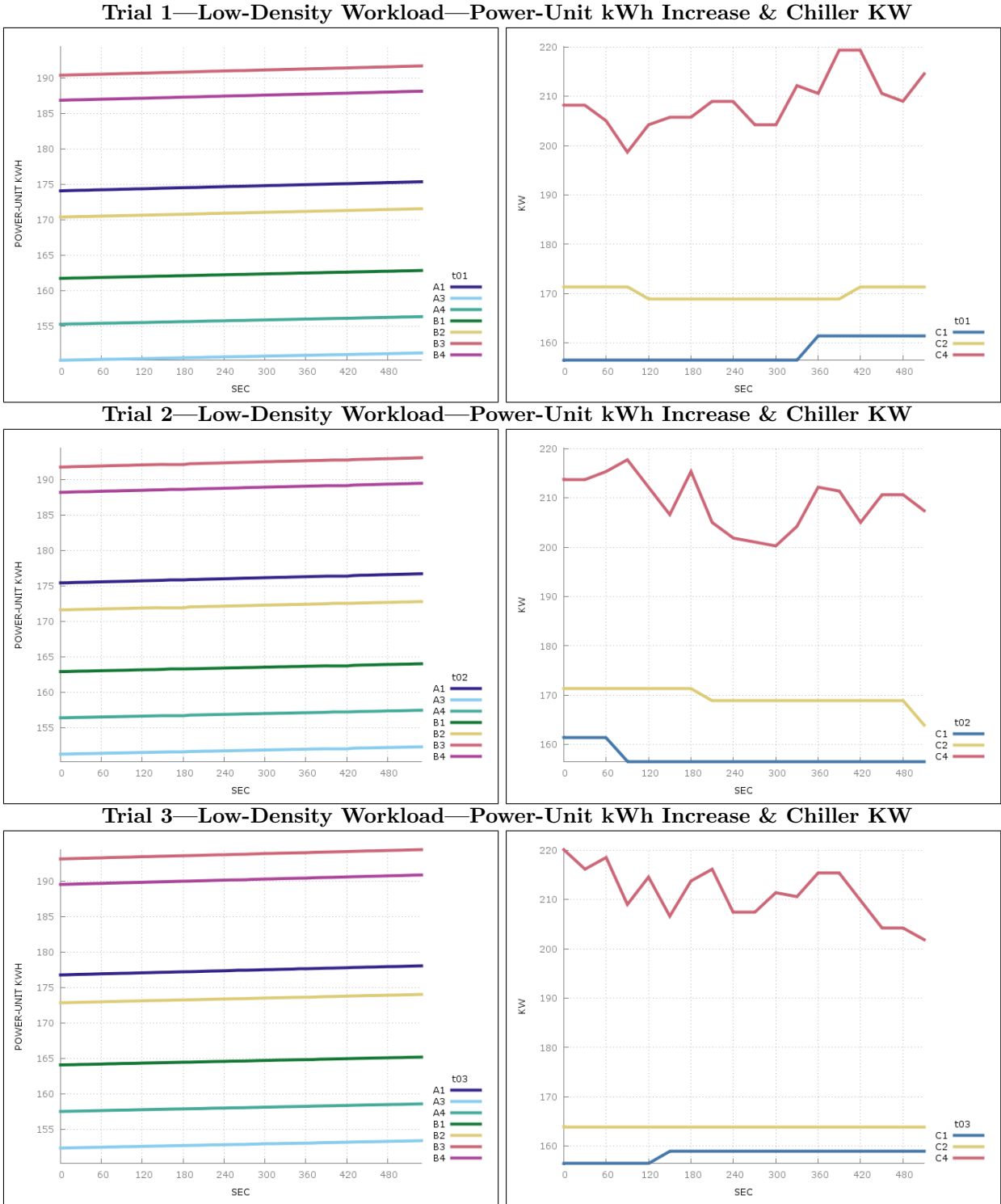
Figure 1: The IT cooling load for the Chillers in use, specified in kilowatts (left), and the Power Unit kWh increase (right).
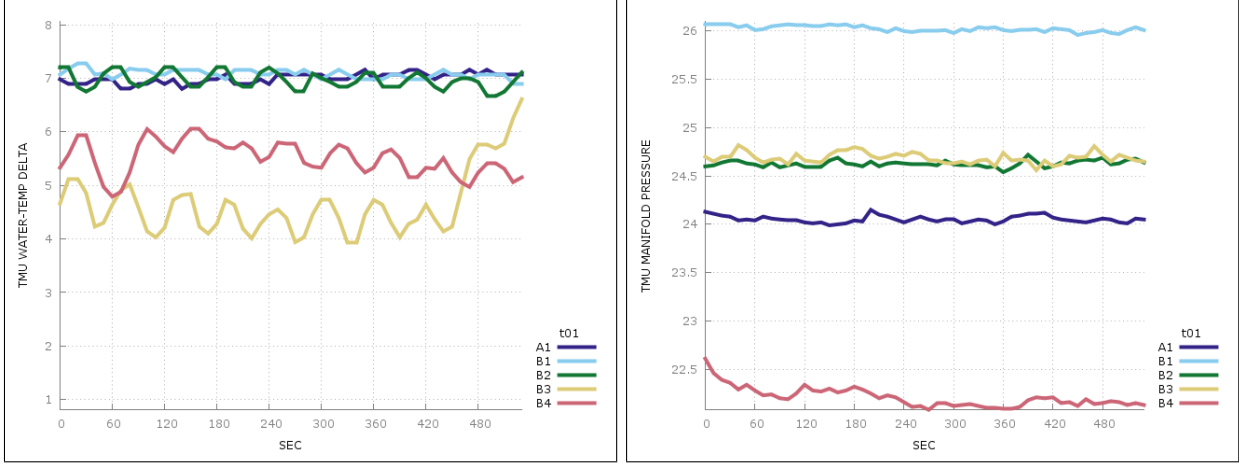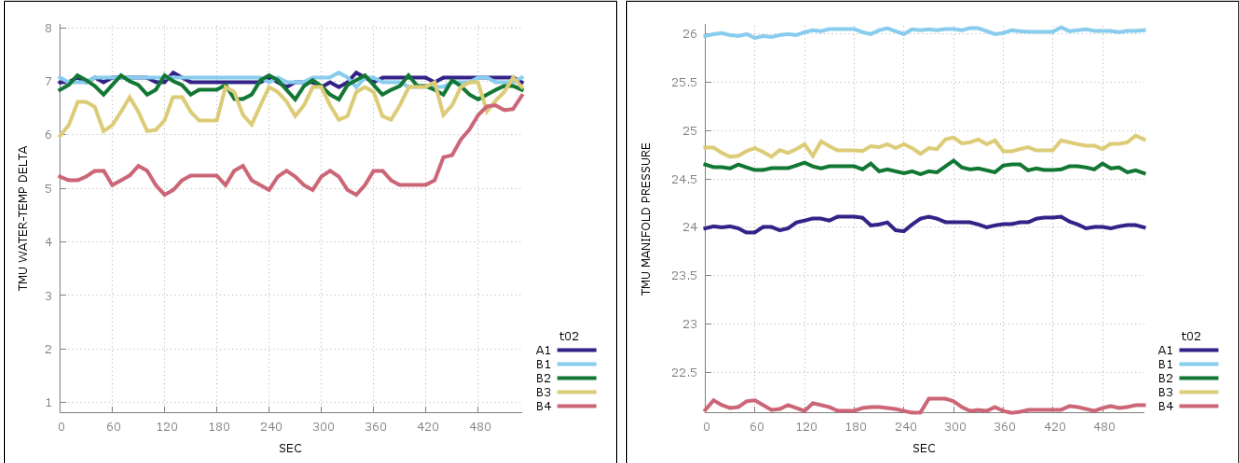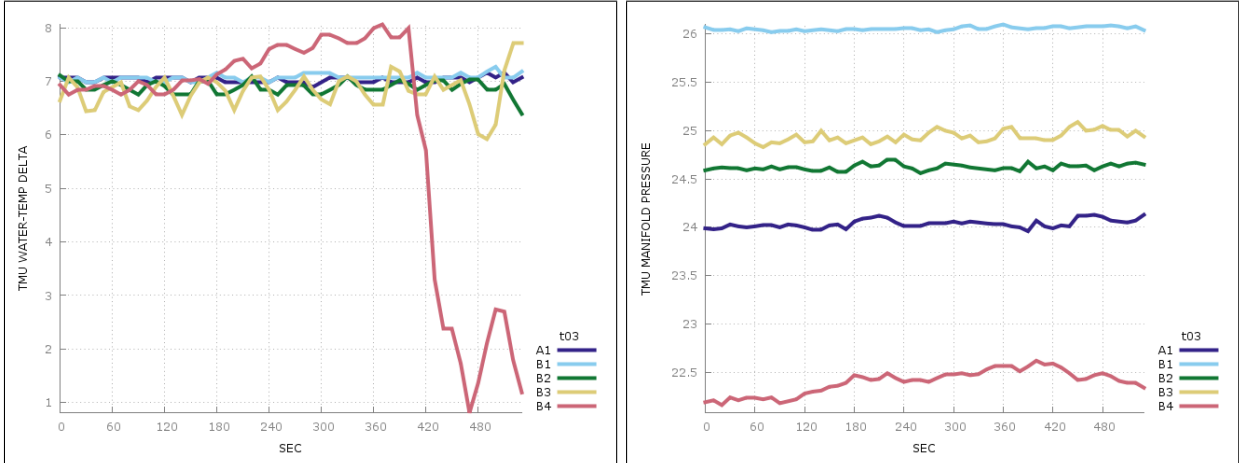
Figure 2: TMU utilization. The water temperature delta (WTD), which is obtained by subtracting inlet from outlet temperatures, and the pressure of the hot/cold liquid distribution manifold (right).

Trial 1—Low-Density Workload—Chiller Utilization
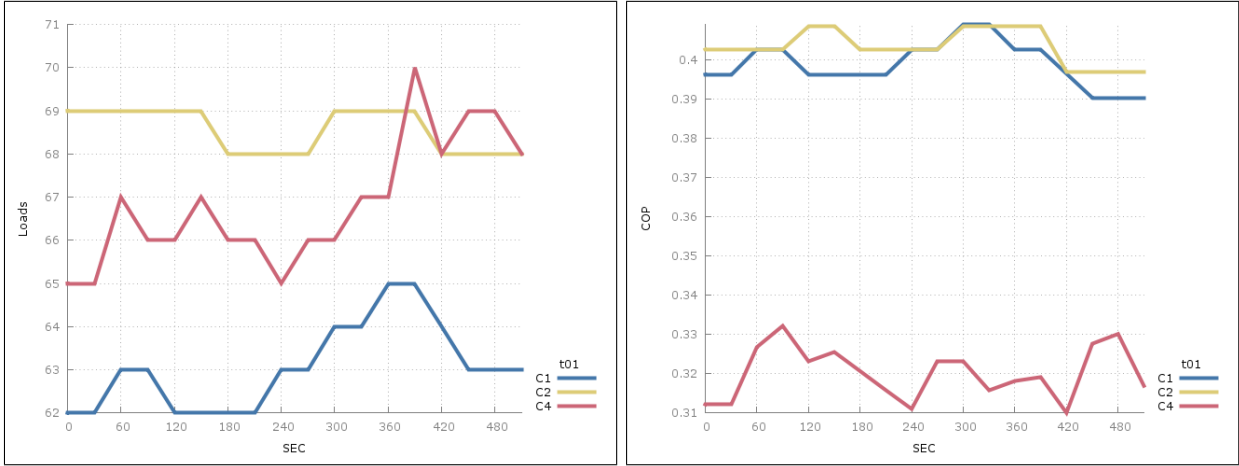
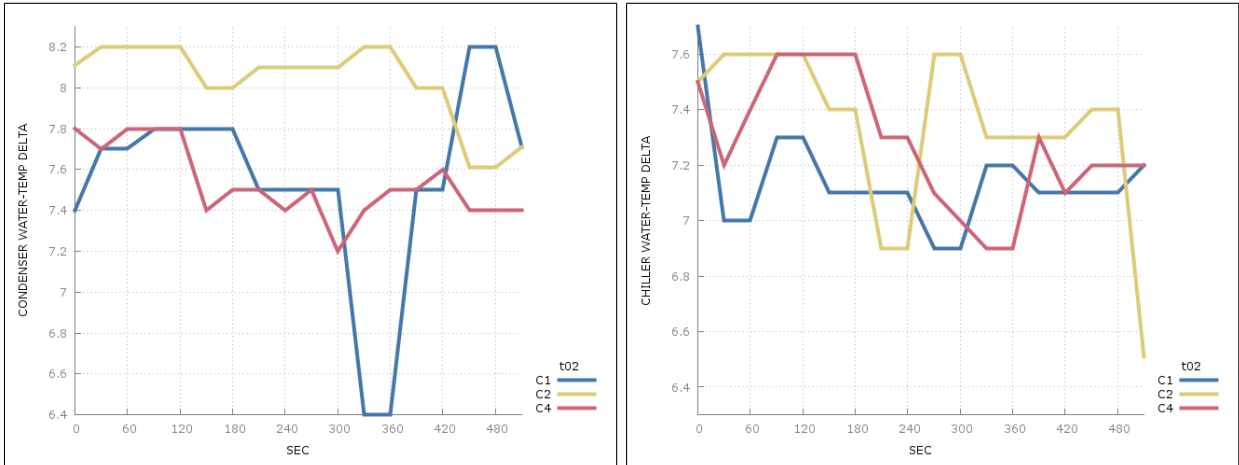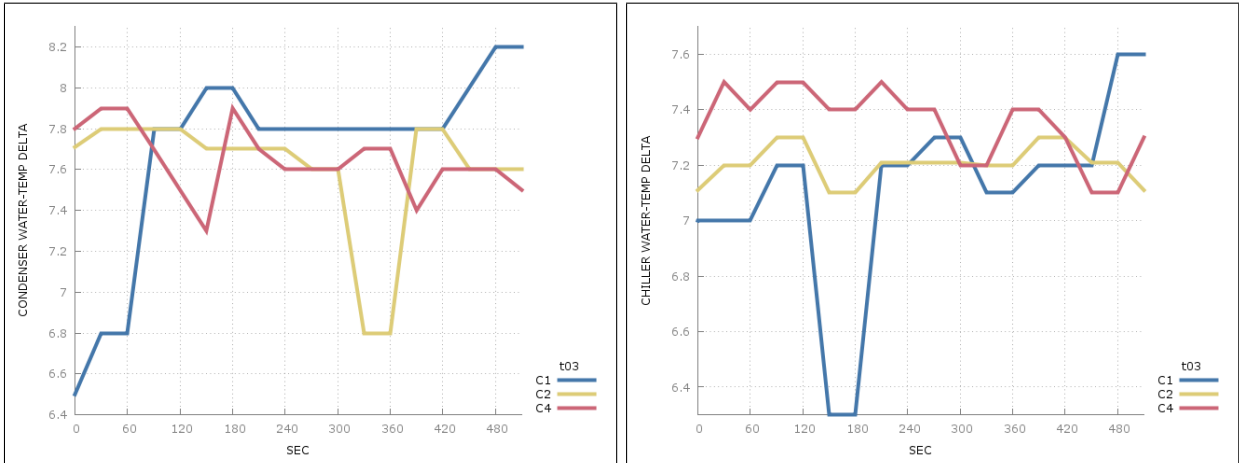**Trial 2—Low-Density Workload—Chiller Utilization**
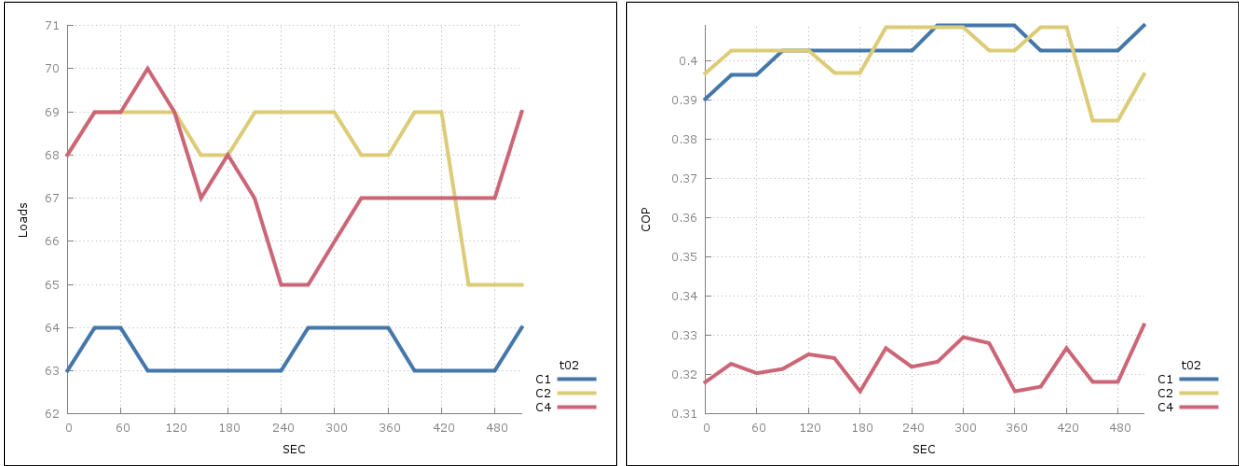
Trial 3—Low-Density Workload—Chiller Utilization
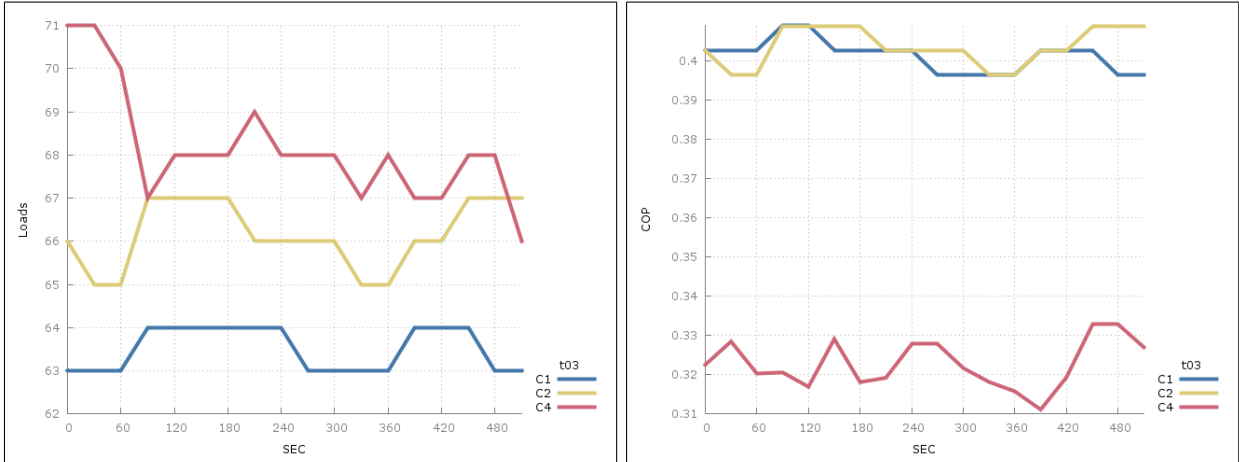
Figure 3: Chiller Loads (left) and Chiller coefficient of performance (COP, right).

Trial 1—Low-Density Workload—Chiller Utilization
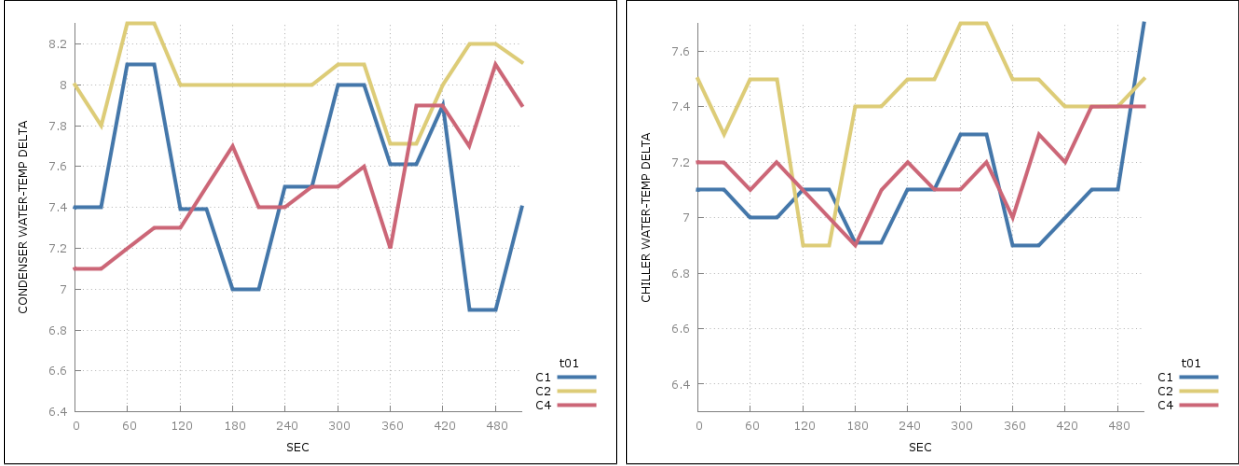
Trial 2—Low-Density Workload—Chiller Utilization

Trial 3—Low-Density Workload—Chiller Utilization

Figure 4: Chiller utilization. The condenser water-temperature delta ($LCWT - ECWT$, left), and the Chiller water-temperature delta ($ECHWT - LCHWT$, right).

9
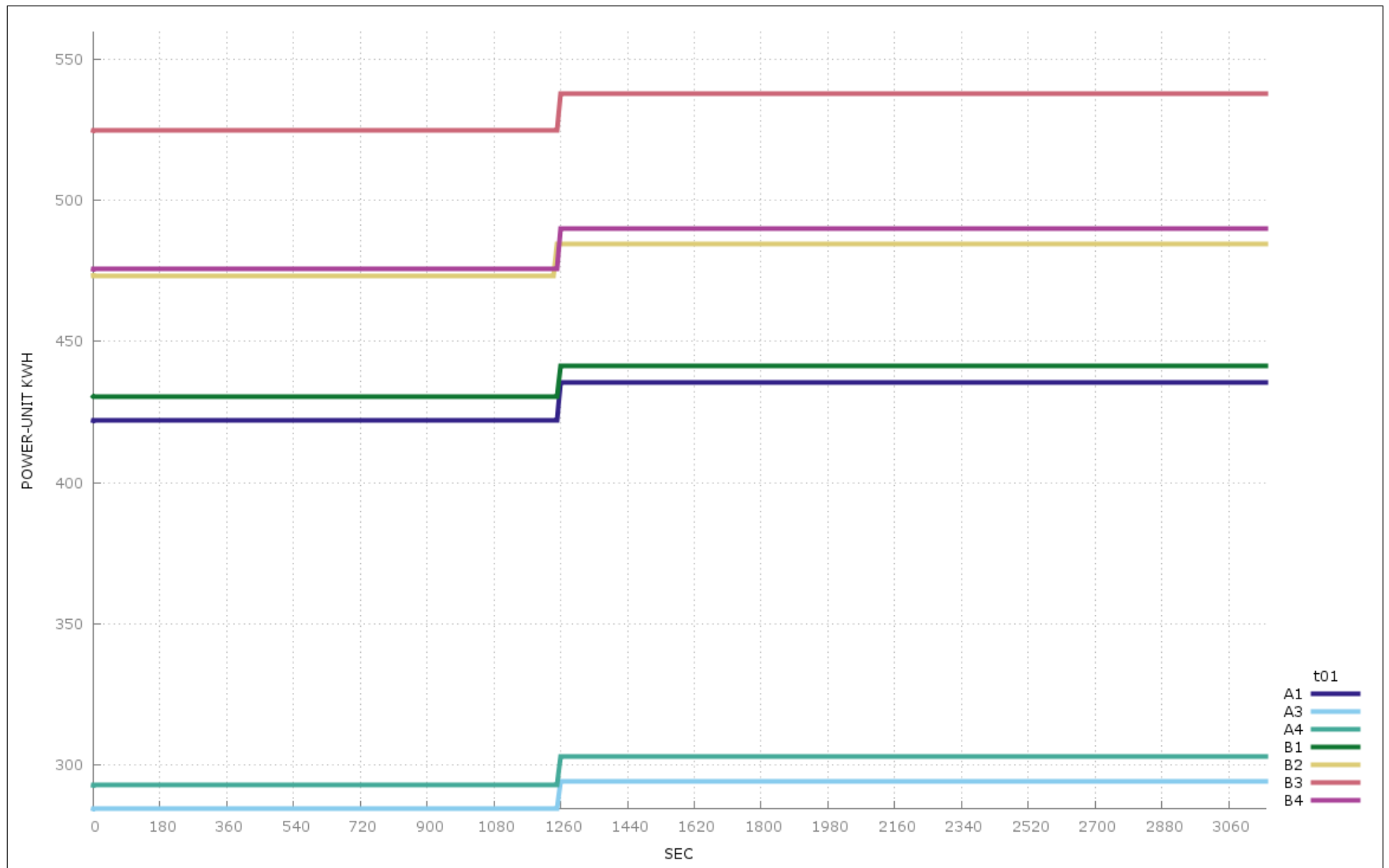
Figure 5: Trial 1—High-Density Workload—kWh Increase
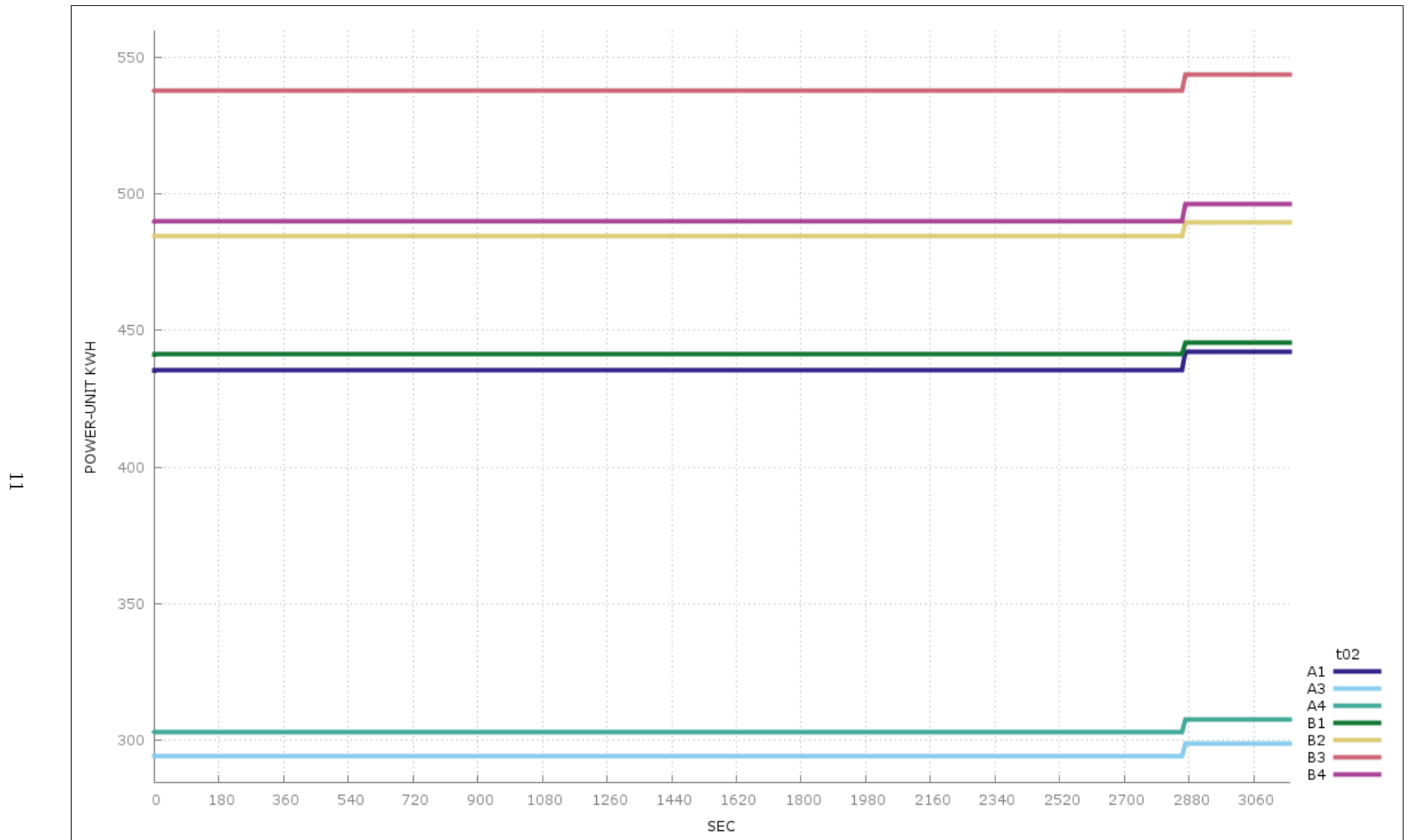
Figure 6: Trial 2—High-Density Workload—kWh Increase
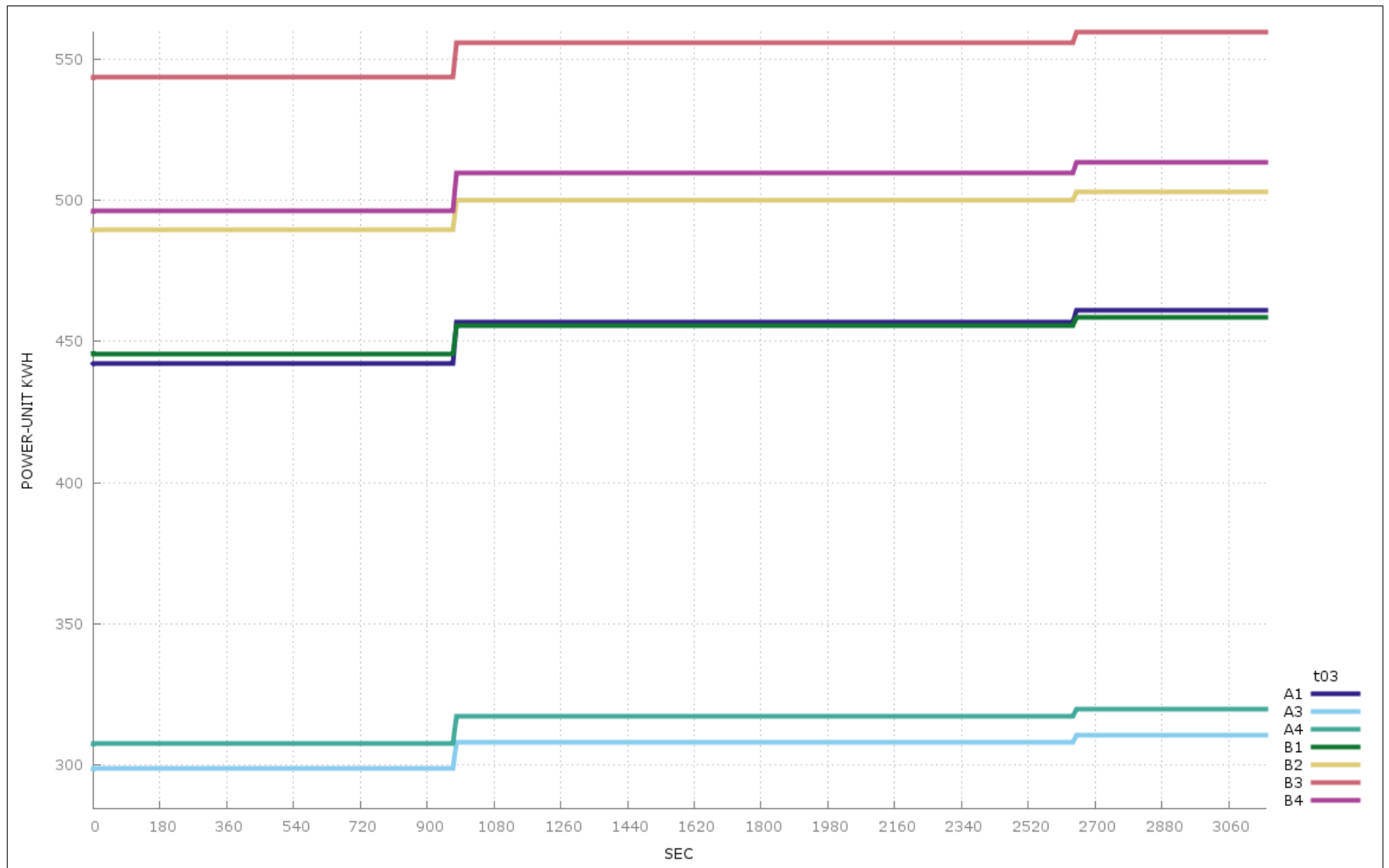
Figure 7: Trial 3—High-Density Workload—kWh Increase
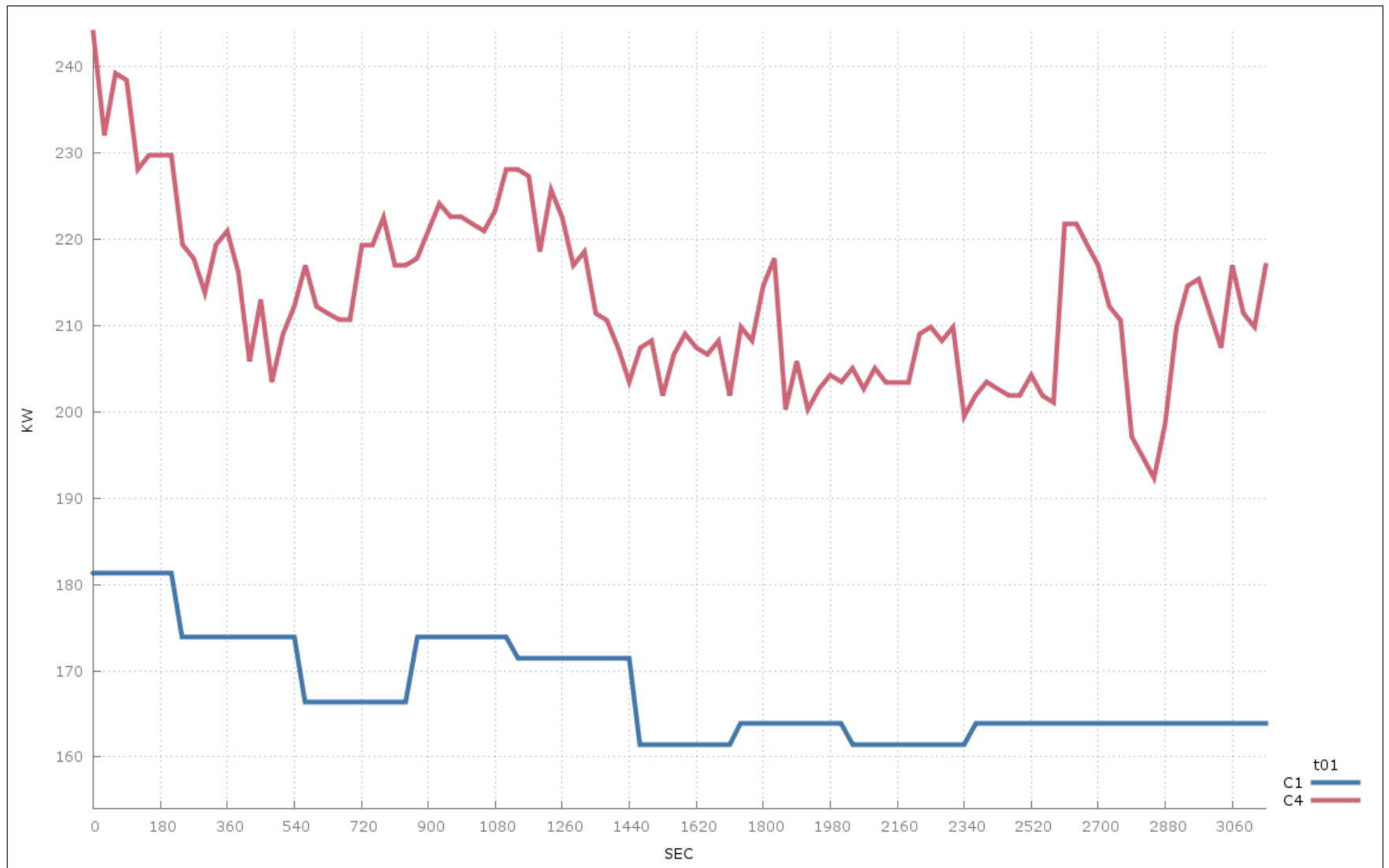
Figure 8: Trial 1—High-Density Workload—Chiller KW
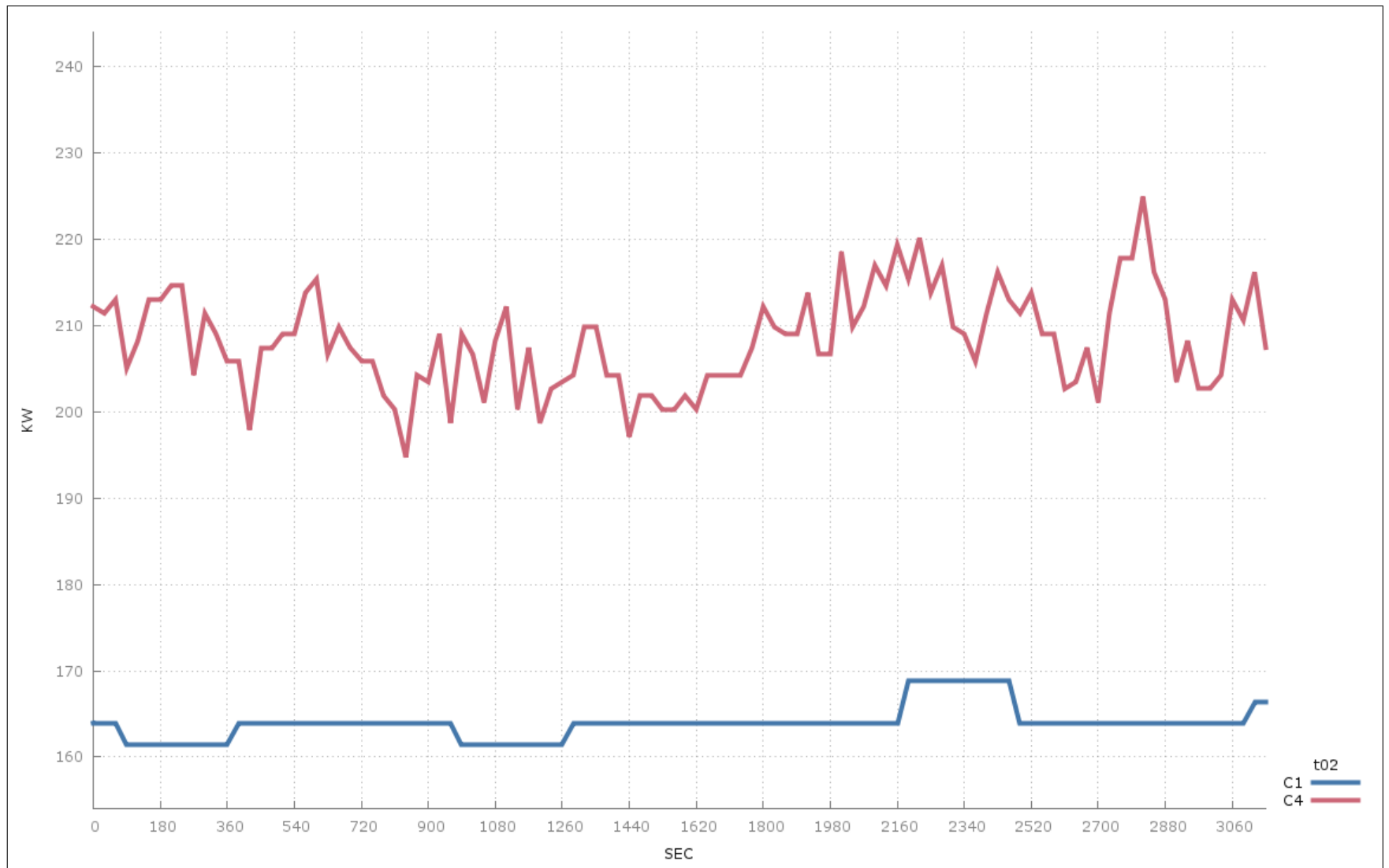
Figure 9: Trial 2—High-Density Workload—Chiller KW
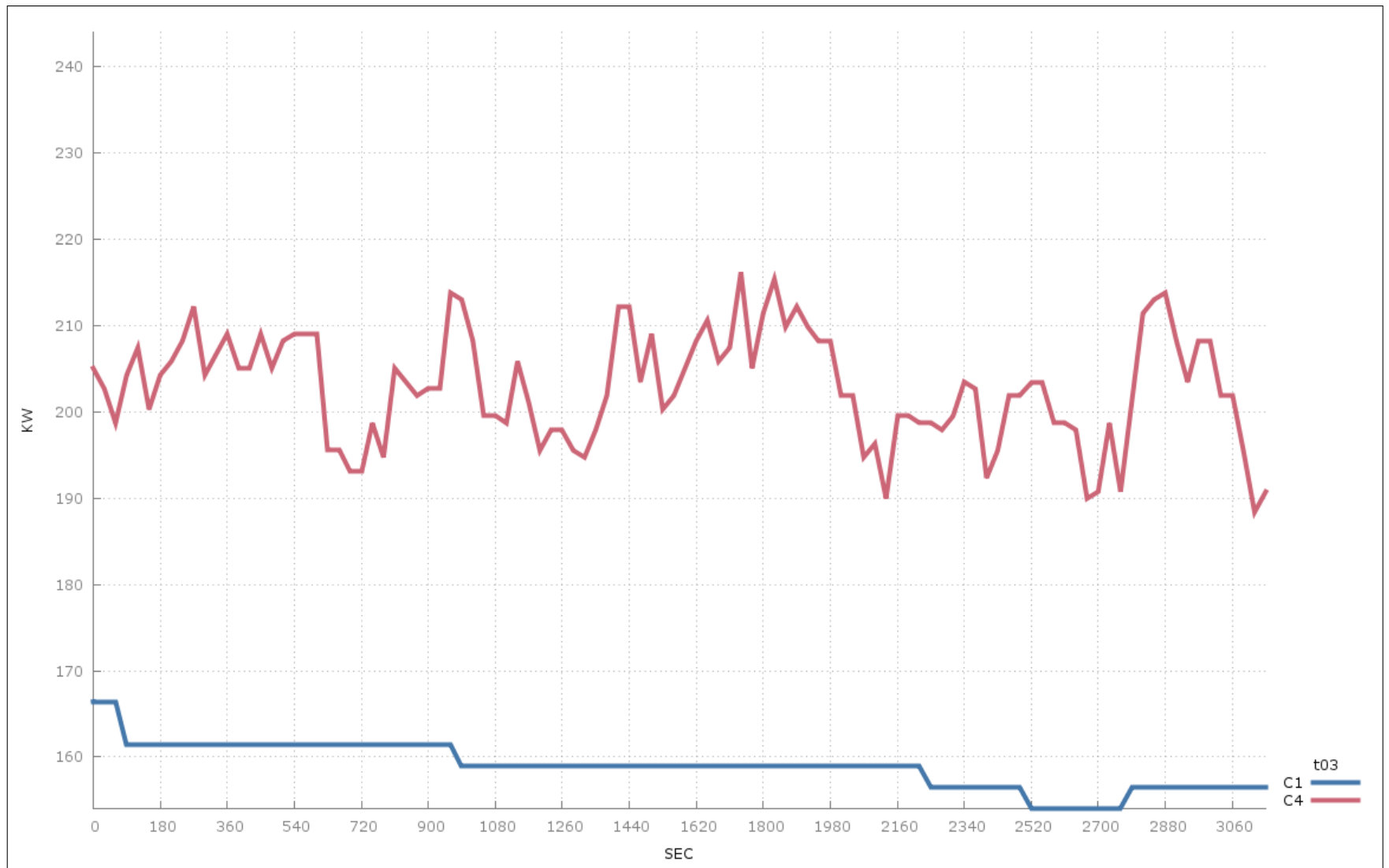
Figure 10: Trial 3—High-Density Workload—Chiller KW
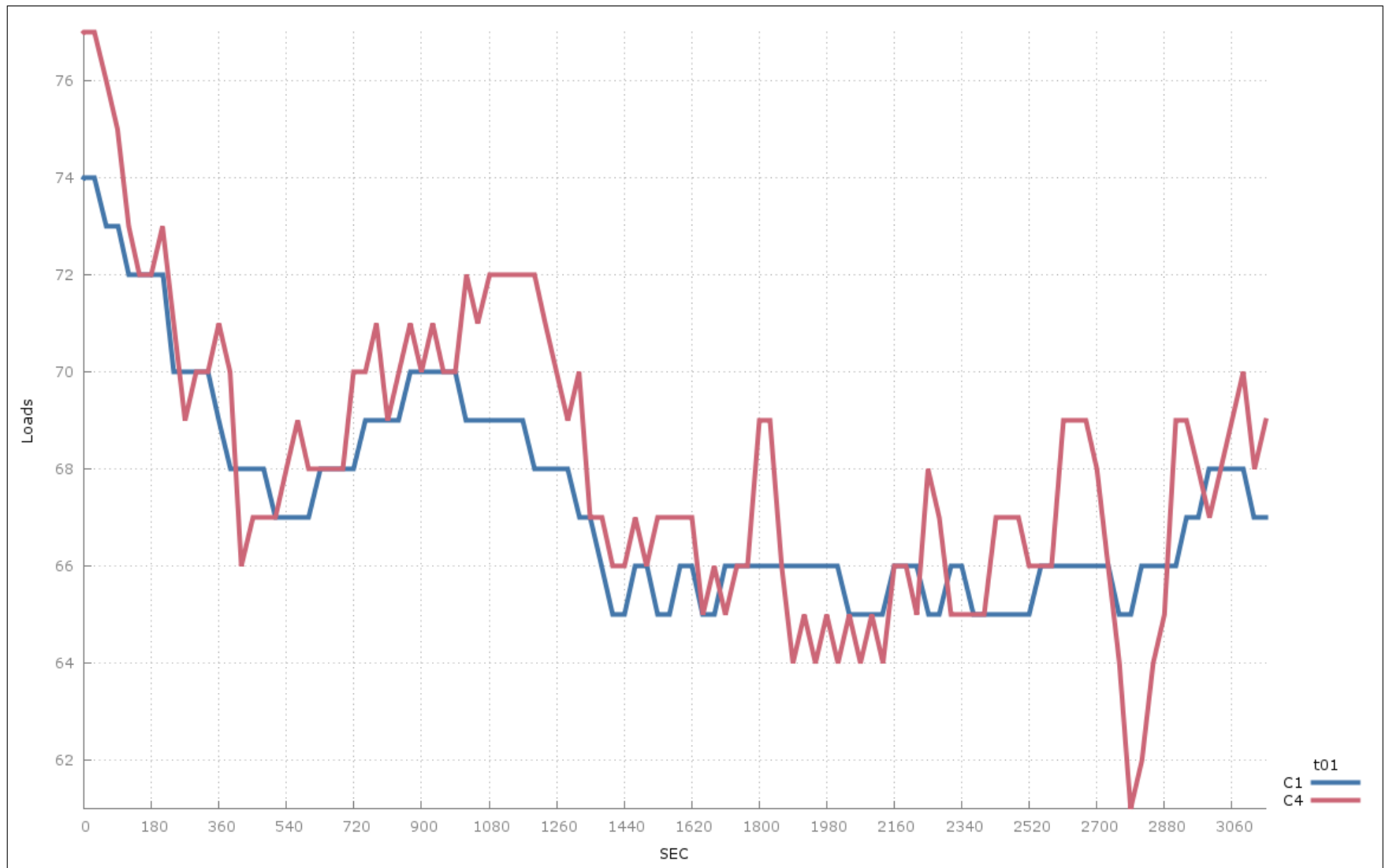
Figure 11: Trial 1—High-Density Workload—Chiller Loads
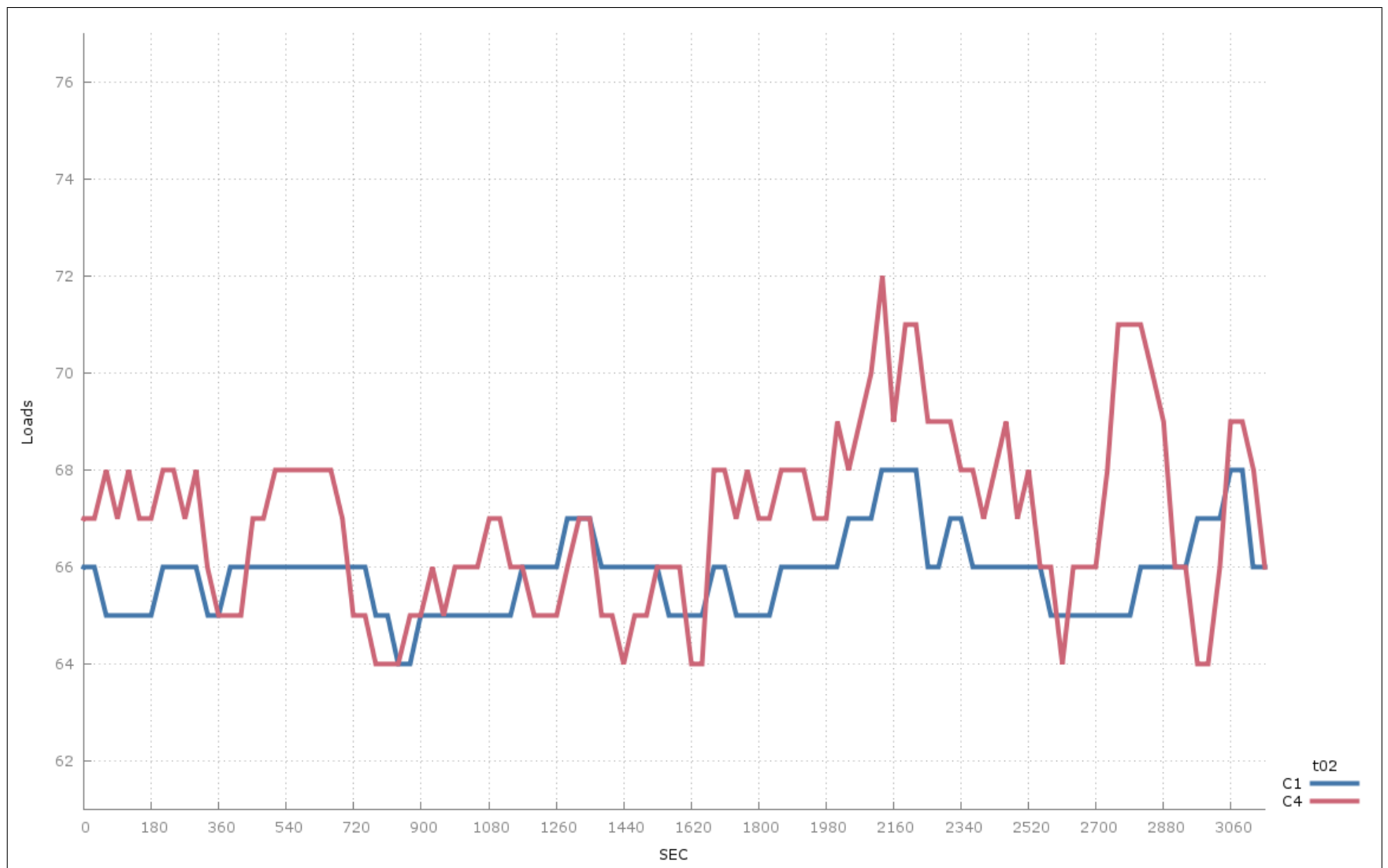
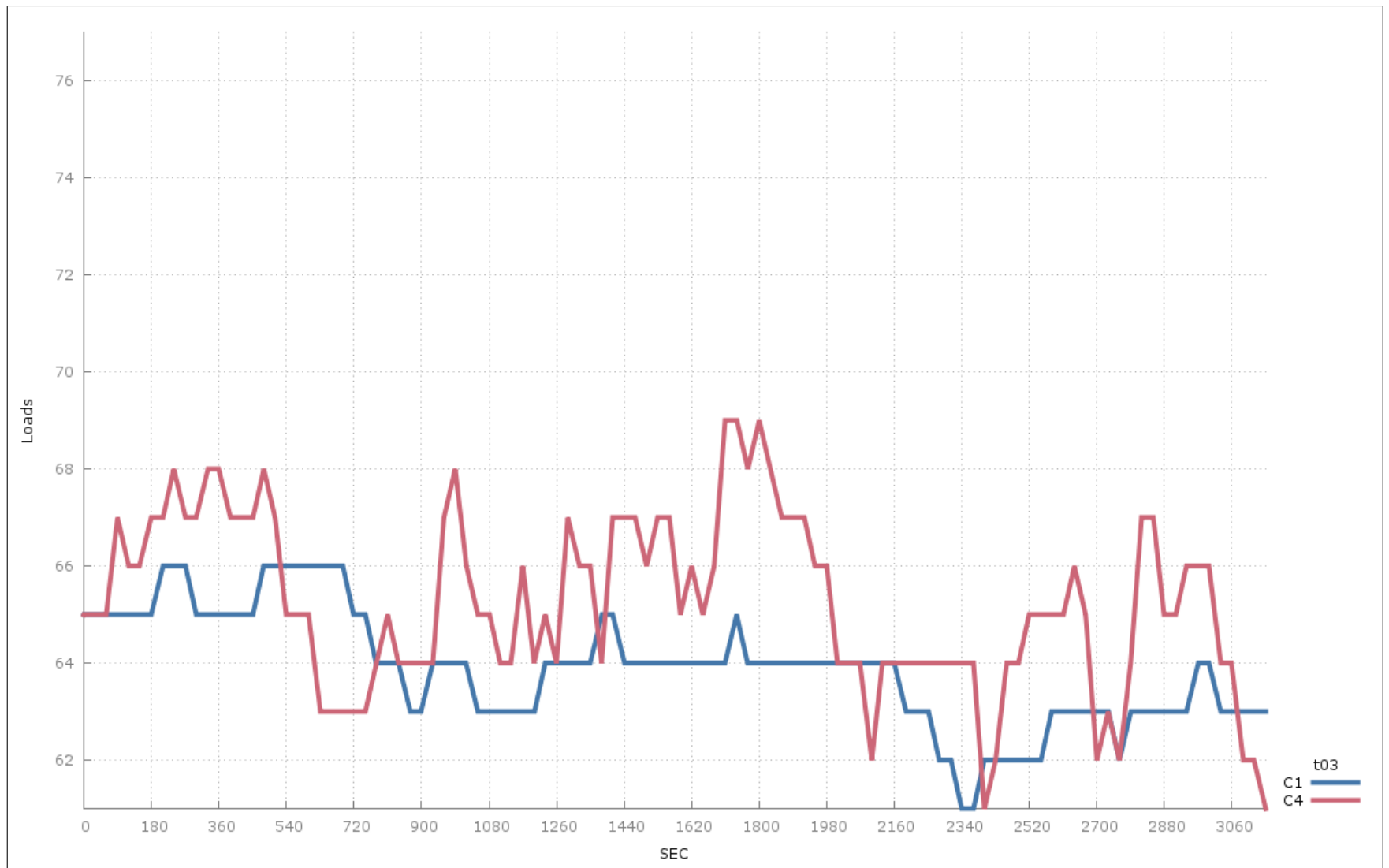Figure 12: Trial 2—High-Density Workload—Chiller Loads
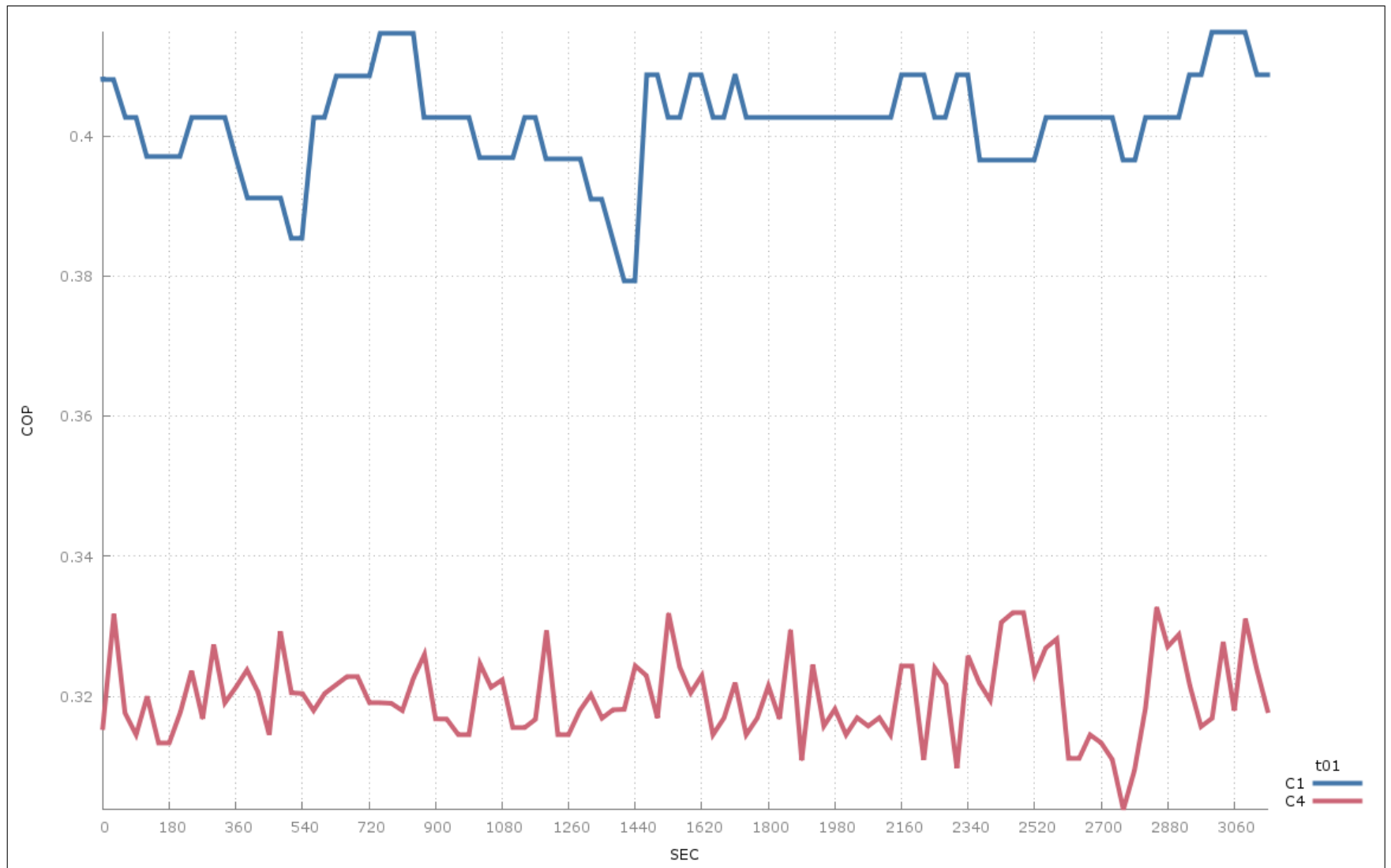
Figure 13: Trial 3—High-Density Workload—Chiller Loads

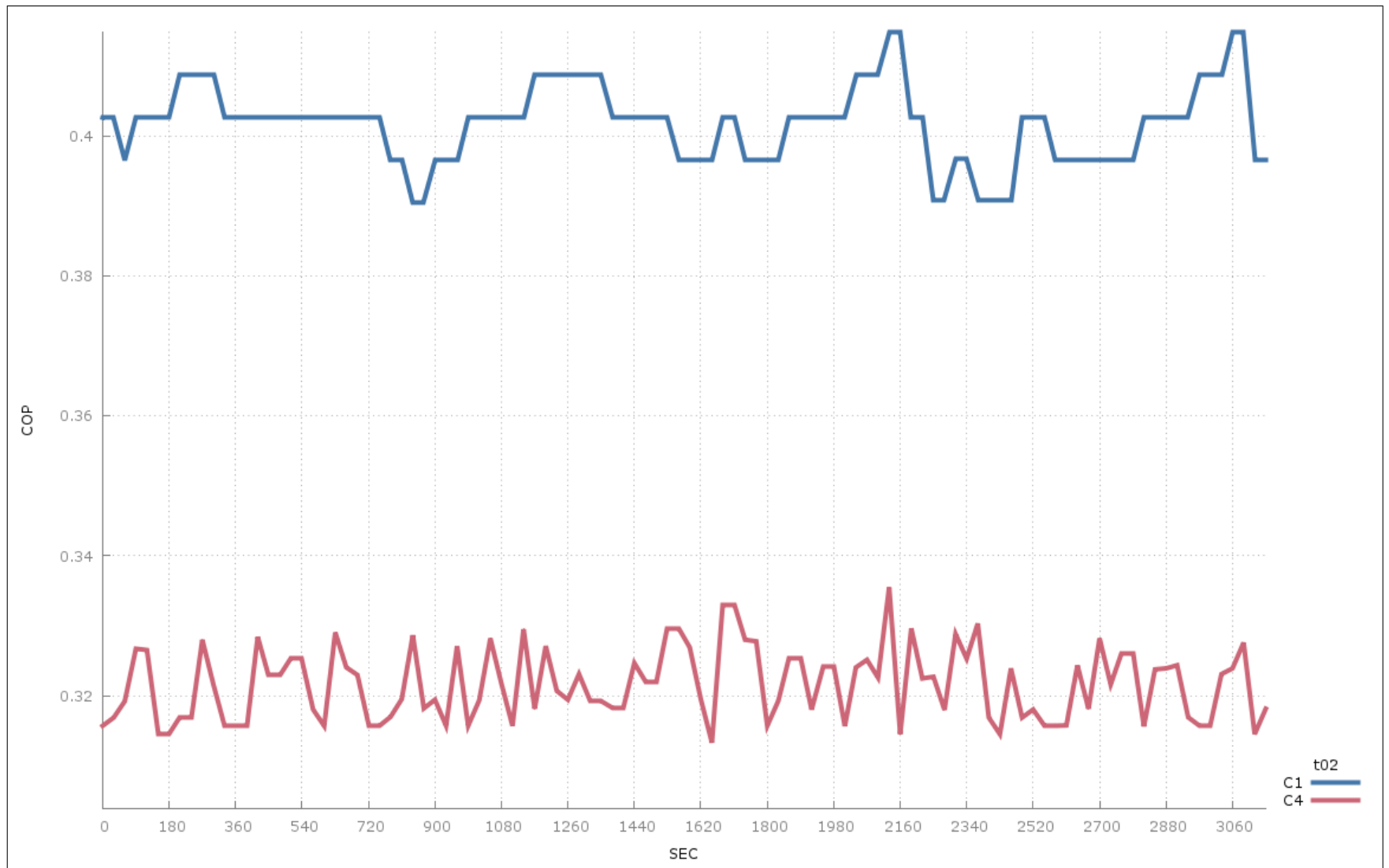Figure 14: Trial 1—High-Density Workload—Chiller COP

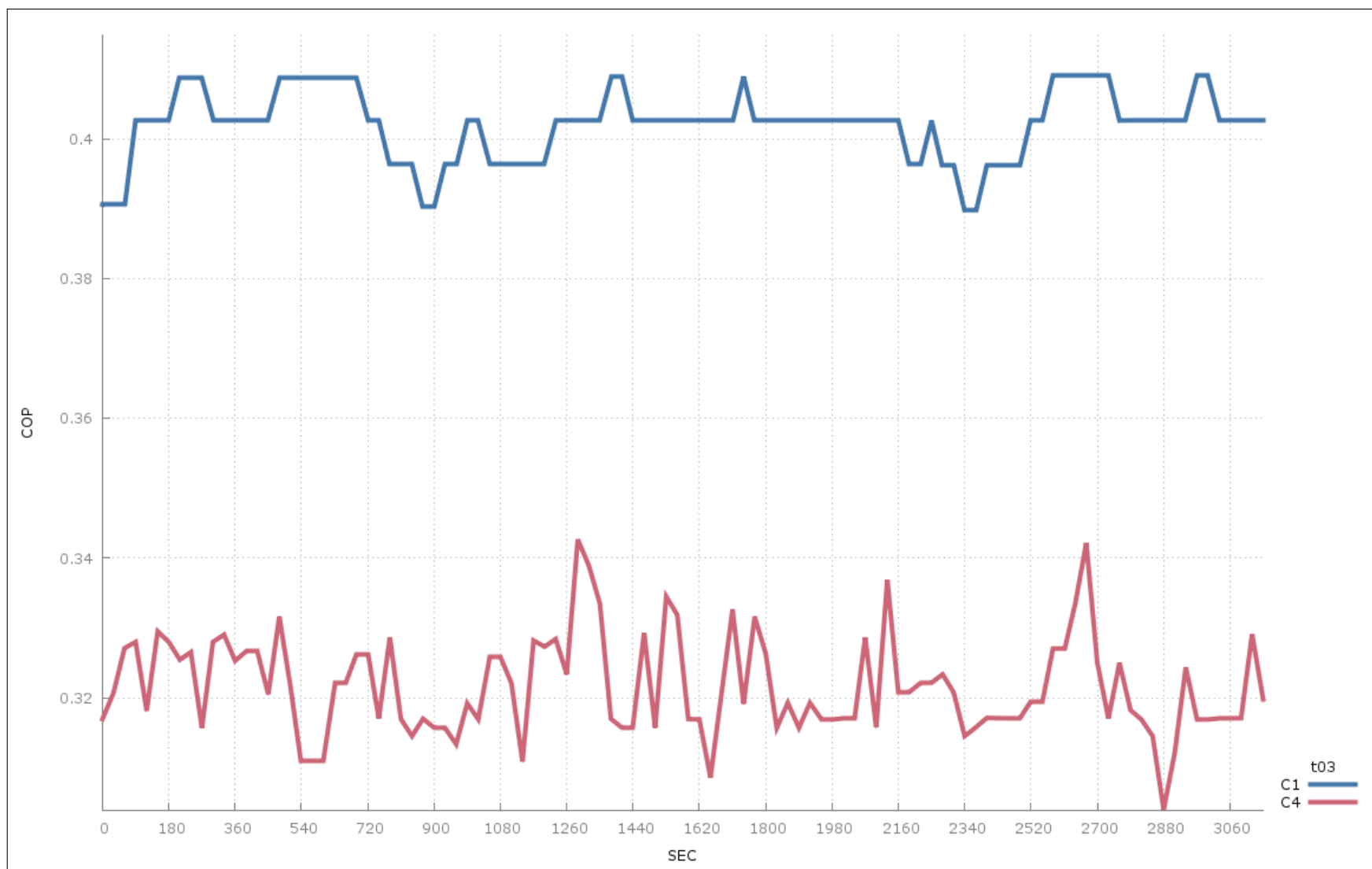Figure 15: Trial 2—High-Density Workload—Chiller COP

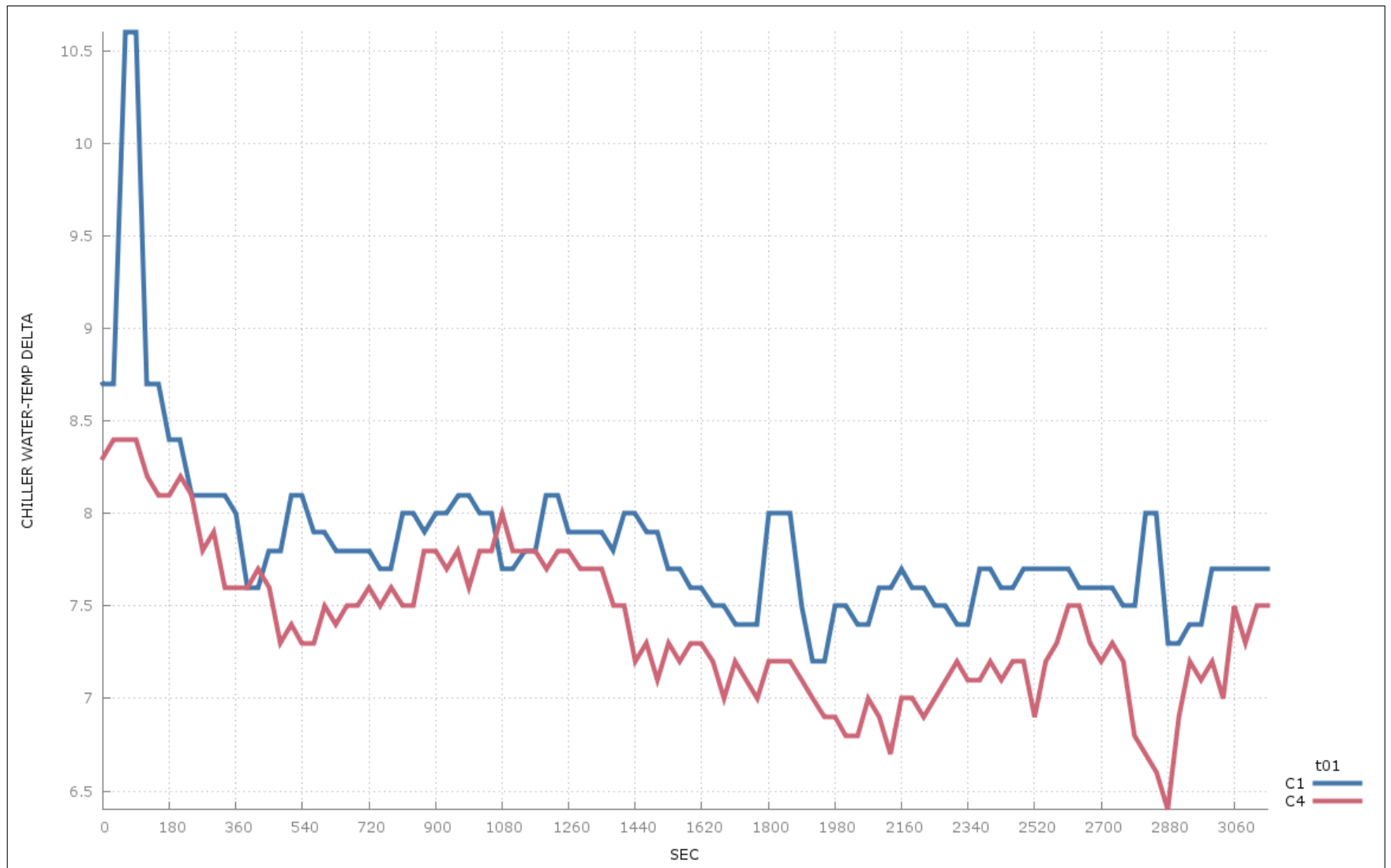Figure 16: Trial 3—High-Density Workload—Chiller COP

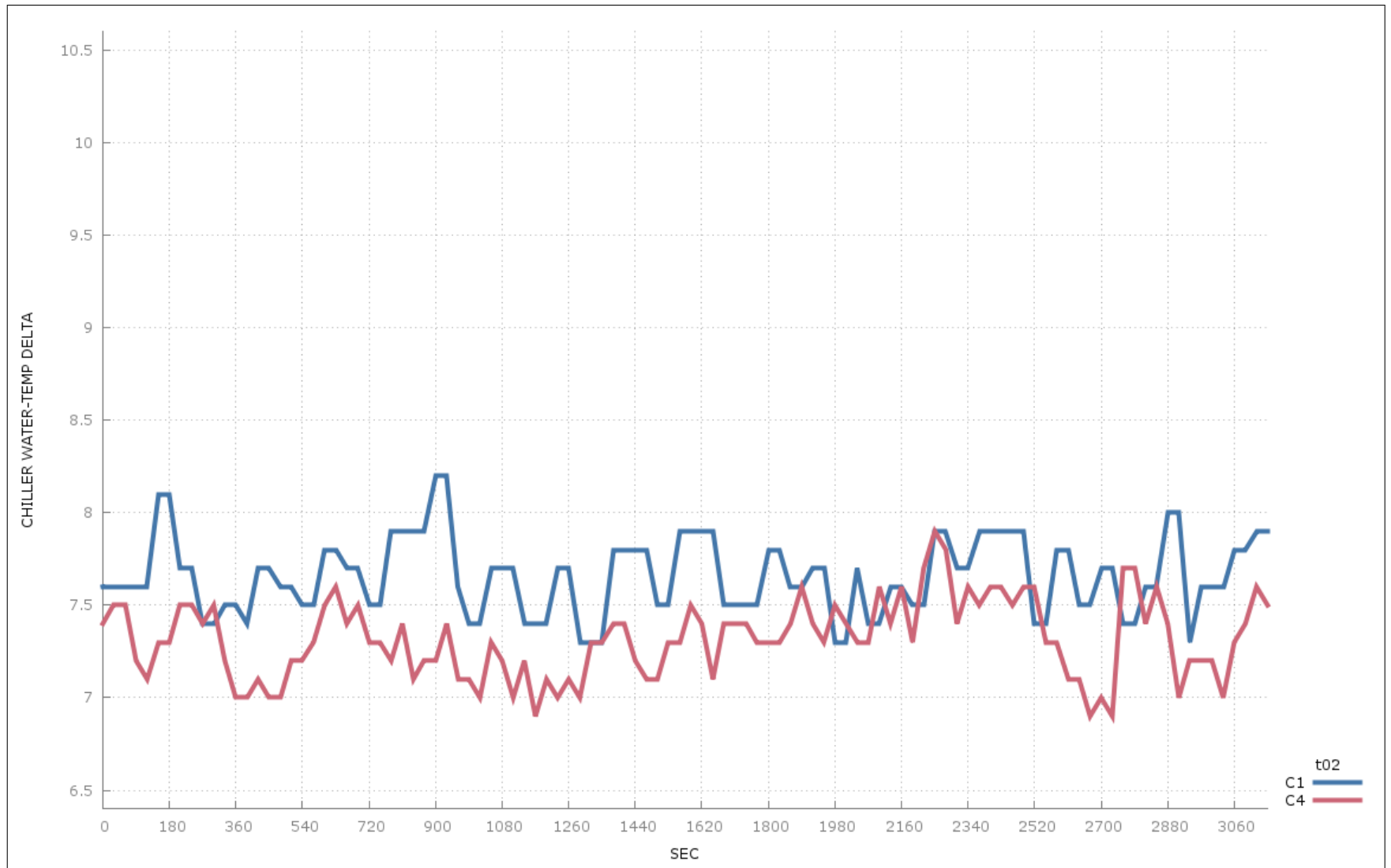Figure 17: Trial 1—High-Density Workload—Chiller Water Temperature Delta

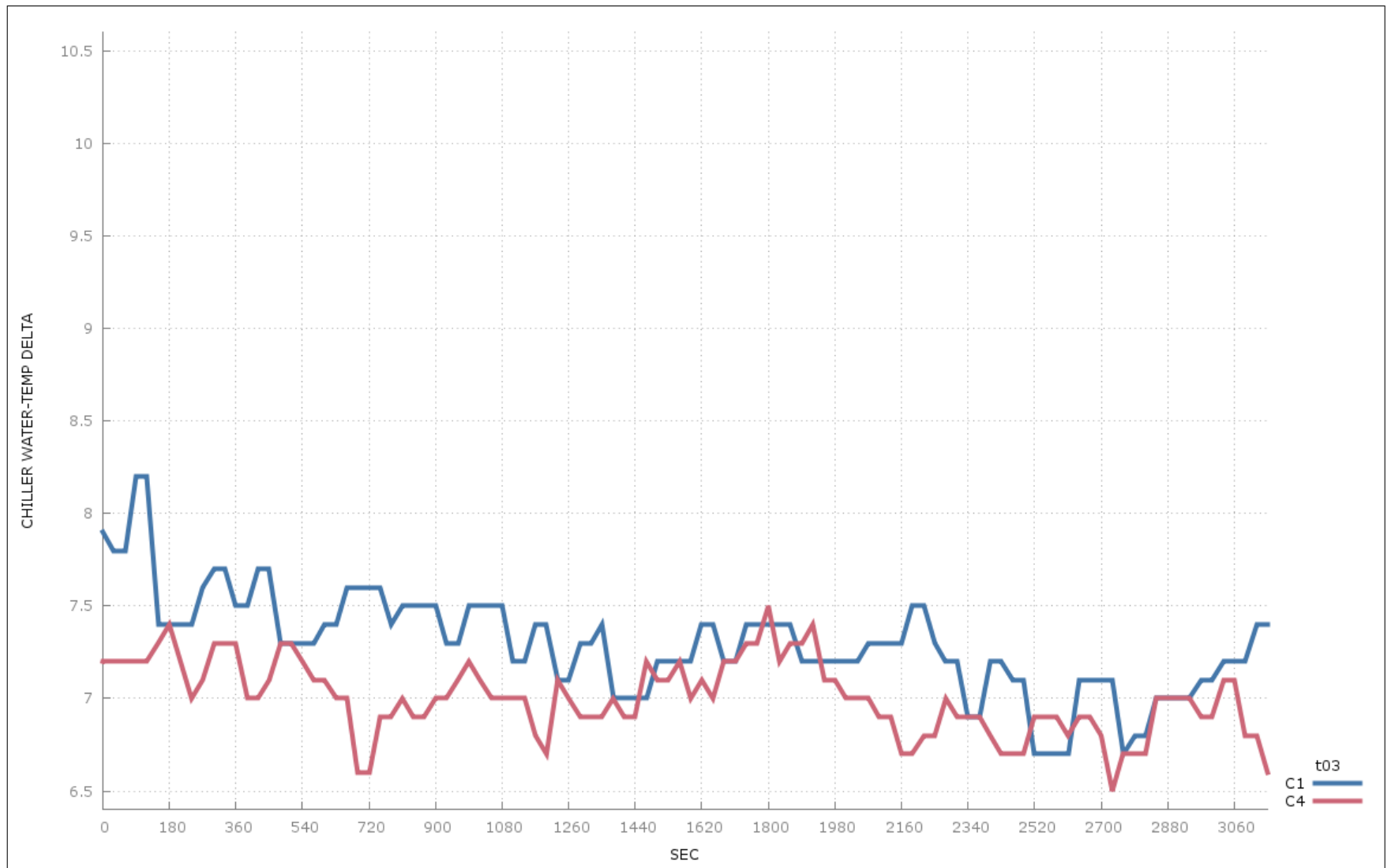Figure 18: Trial 2—High-Density Workload—Chiller Water Temperature Delta

Figure 19: Trial 3—High-Density Workload—Chiller Water Temperature Delta

Bibliography

[1] Elke Achtert, Ahmed Hettab, Hans-Peter Kriegel, Erich Schubert, and Arthur Zimek. Spatial outlier detection: data, algorithms, visualizations. In *Proceedings of the 12th international conference on Advances in spatial and temporal databases*, SSTD'11, pages 512–516, Berlin, Heidelberg, 2011. Springer-Verlag.

[2] Rashawn L. Knapp, Karen L. Karavanic, Sriram Krishnamoorthy, and Andres Marquez. Power- and cooling- aware parallel performance diagnosis. In *Parallel and Distributed Computing and Systems (PDCS 2011)*, Dallas, USA, December 2011.

[3] Debprakash Patnaik, Manish Marwah, Ratnesh Sharma, and Naren Ramakrishnan. Sustainable operation and management of data center chillers using temporal data mining. In *Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining*, KDD '09, pages 1305–1314, New York, NY, USA, 2009. ACM.

[4] Debprakash Patnaik, Manish Marwah, Ratnesh K. Sharma, and Naren Ramakrishnan. Temporal data mining approaches for sustainable chiller management in data centers. *ACM Trans. Intell. Syst. Technol.*, 2(4):34:1–34:29, July 2011.