NSF/IUCRC CAC PROJECT

INTEGRATED VISUALIZING, MONITORING, AND MANAGING HPC SYSTEMS

Jie Li

Doctoral Student, TTU

06/04/2021

Advisors:

Mr. Jon Hass, SW Architect, Dell Inc.

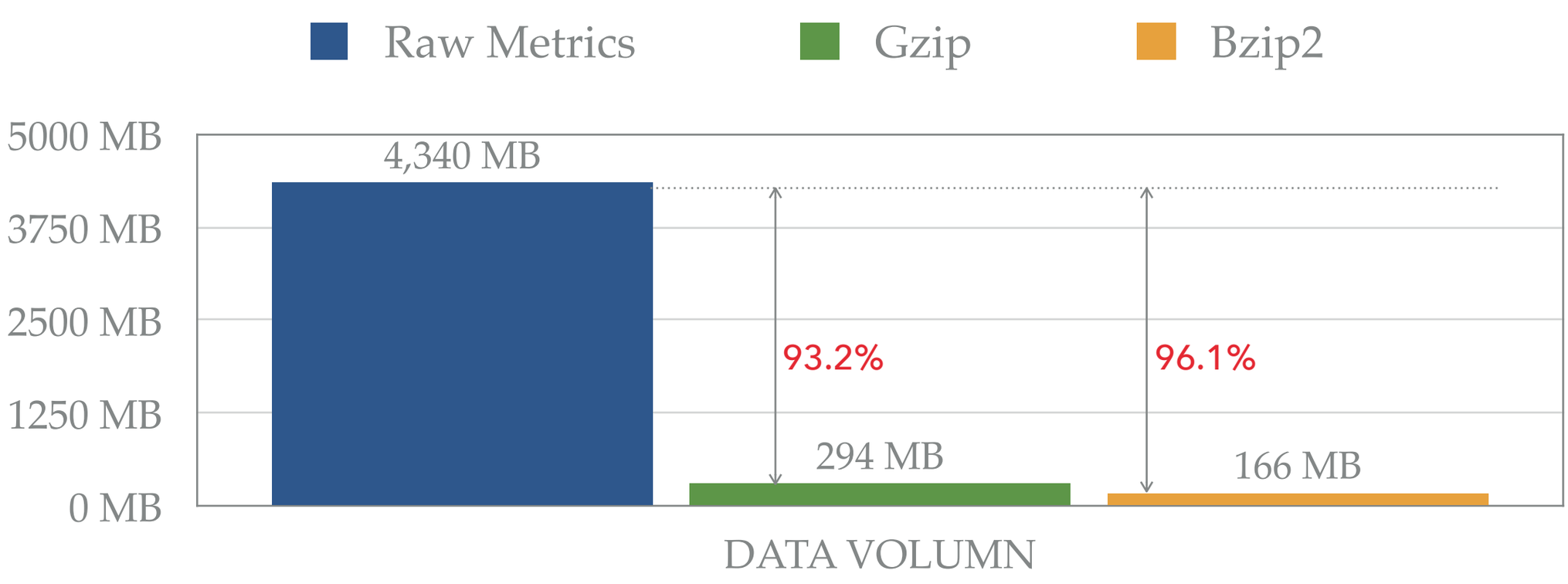
Dr. Alan Sill, Managing Director, HPCC, TTU

Dr. Yong Chen, Associate Professor, CS Dept, TTU

Dr. Tommy Dang, Assistant Professor, CS Dept, TTU

LAST MEETING

- Explored compressed monitoring metrics in zipped files
  - Compressing CSV file achieved a better performance:
    - Compression rate is about 96.1% using Bzip2



1 day of iDRAC9 monitoring data on the RedRaider Cluster

TIMESCALE DB COMPRESSION

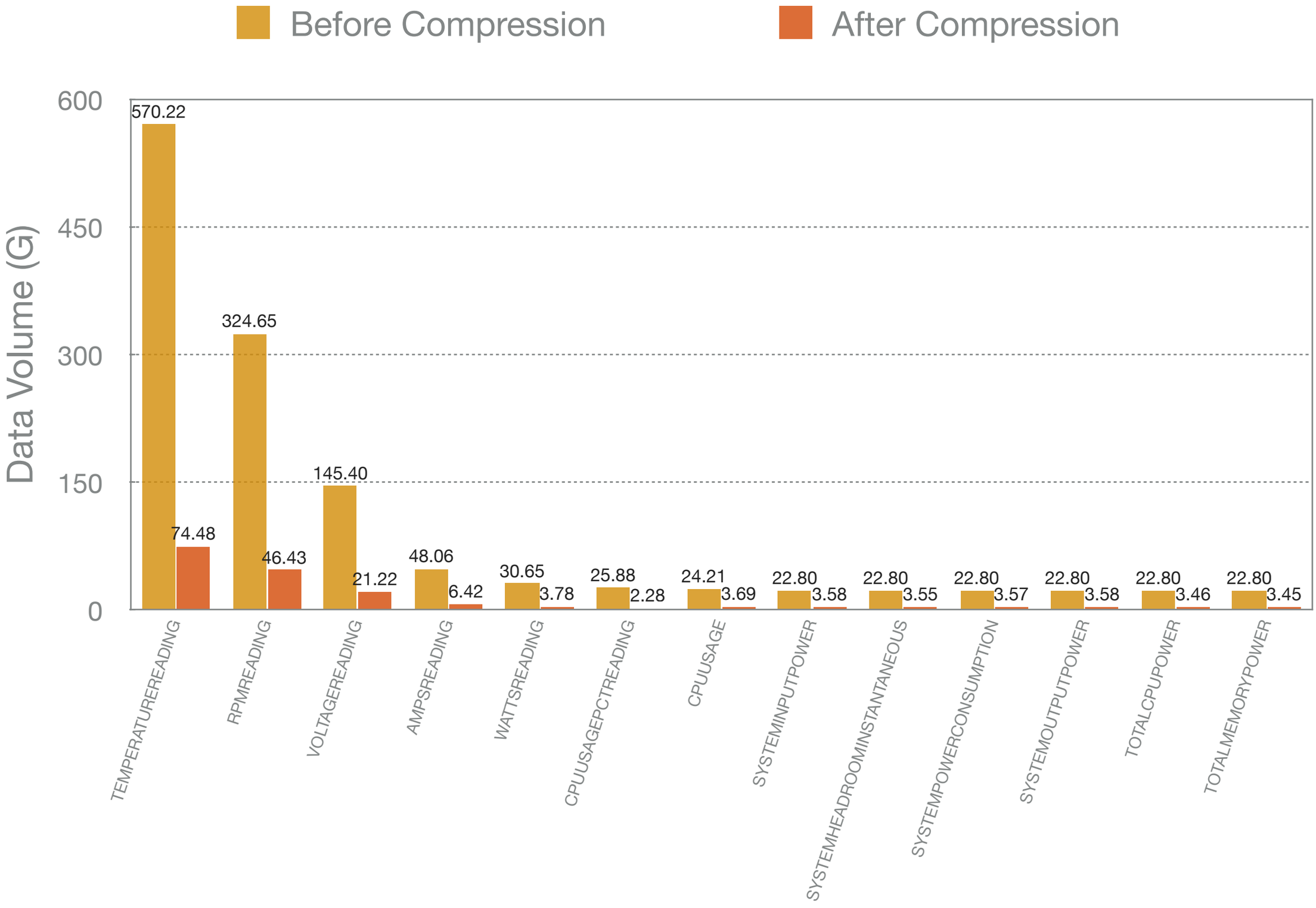
- Upgrade TimescaleDB from version 1.74 to version 2.3.
- Explore TimescaleDB built-in compression mechanism.
- Use the following commands to enable the compression on all idrac9 tables.

```
ALTER TABLE idrac9_measurements SET (
  timescaledb.compress,
  timescaledb.compress_segmentby = ‘nodeid'
);

SELECT add_compression_policy(‘idrac9_measurements', INTERVAL '7 days');
```

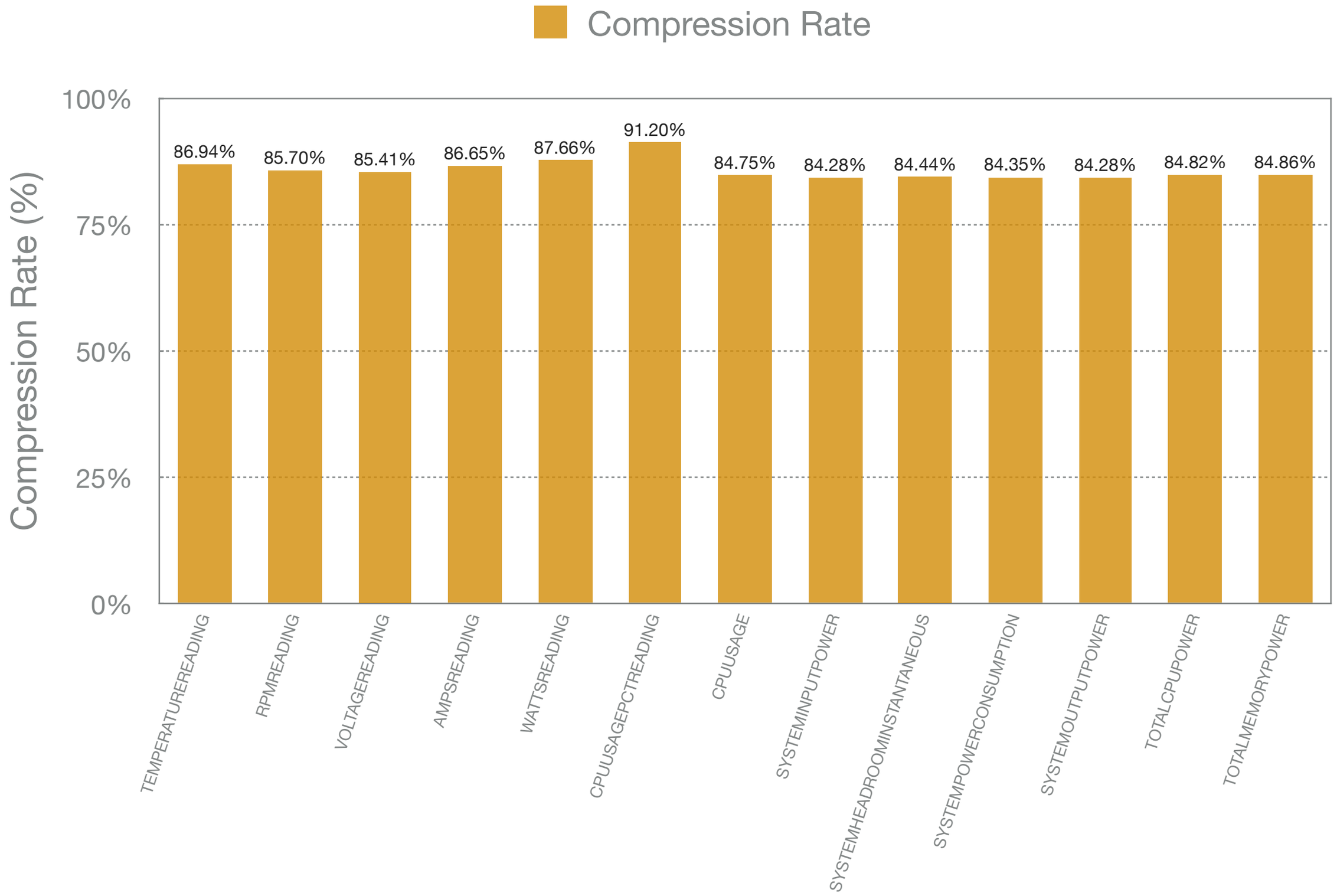
Configure compression and compress chunks older than 7 days.

TIMESCALE DB COMPRESSION



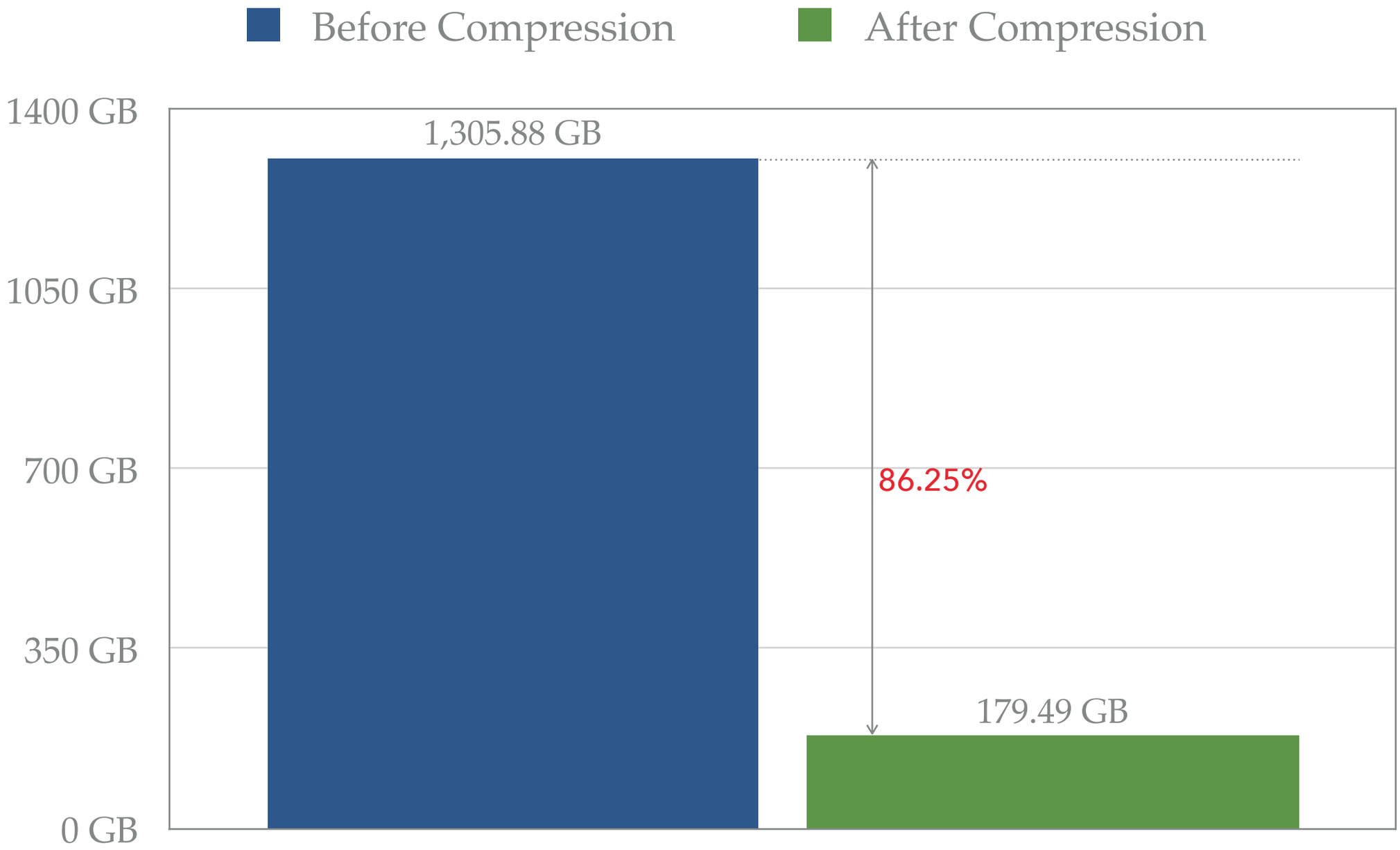
(The monitoring data is collected from March 24th to Jun 2nd)

TIMESCALE DB COMPRESSION



(The monitoring data is collected from March 24th to Jun 2nd)

TIMESCALE DB COMPRESSION



(The monitoring data is collected from March 24th to Jun 2nd)

QUESTIONS?/COMMENTS?

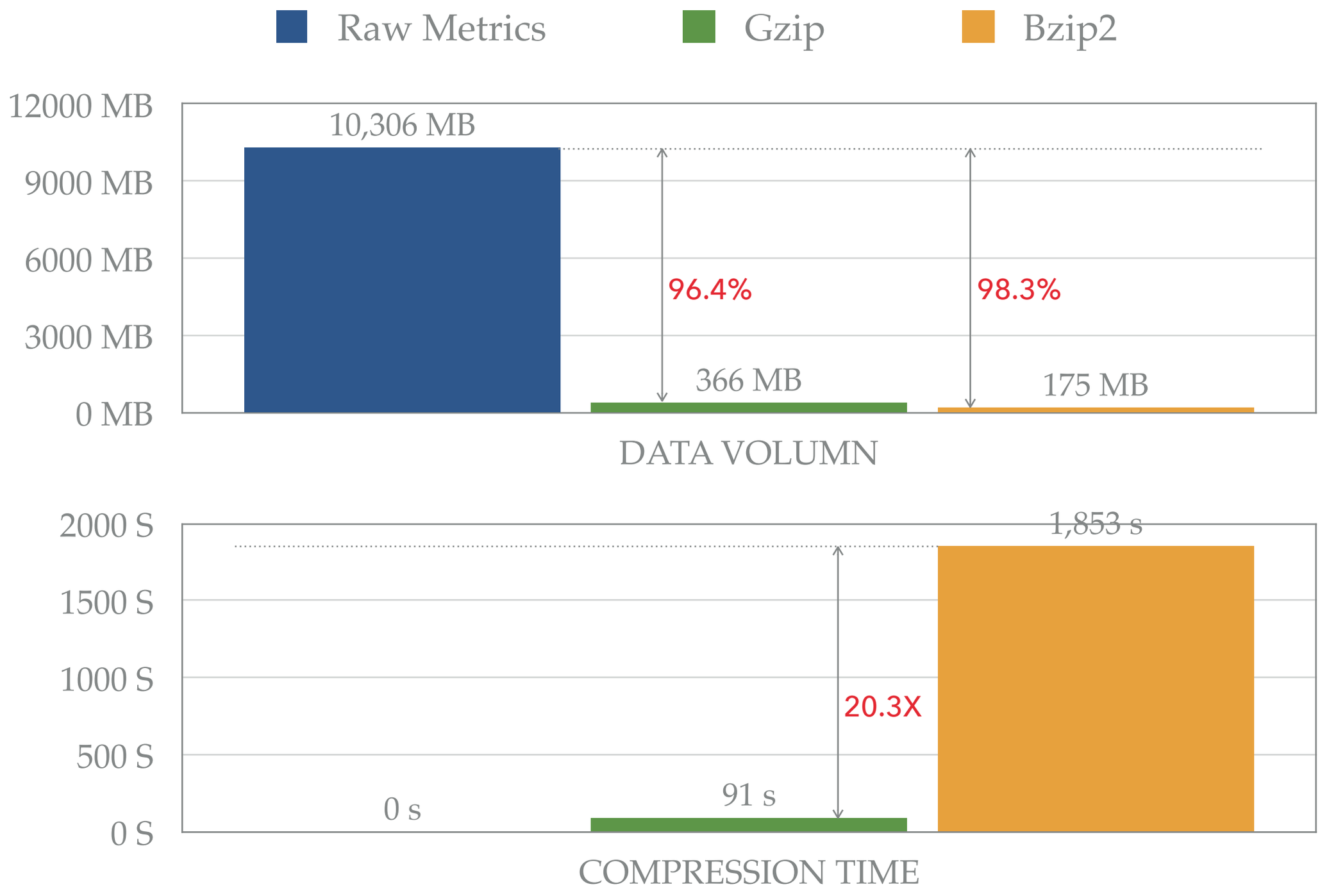
LAST MEETING

- Managing the growing amount of monitoring data.
  - Drop metrics that have values of 0, e.g., iousage, memoryusage, totalfanpower, etc.
  - NIC metrics (27 metrics)? e.g., txbytes, txbroadcast, etc.
  - Save old metrics (e.g., metrics collected one month ago) to zipped files.
- JSON: save json format data while collecting metrics.
- CSV: export csv files from TimescaleDB.

```
record = {
    'Timestamp': time,
    'Source': source,
    'FQDD': fqdd,
    'Value': value
}
```

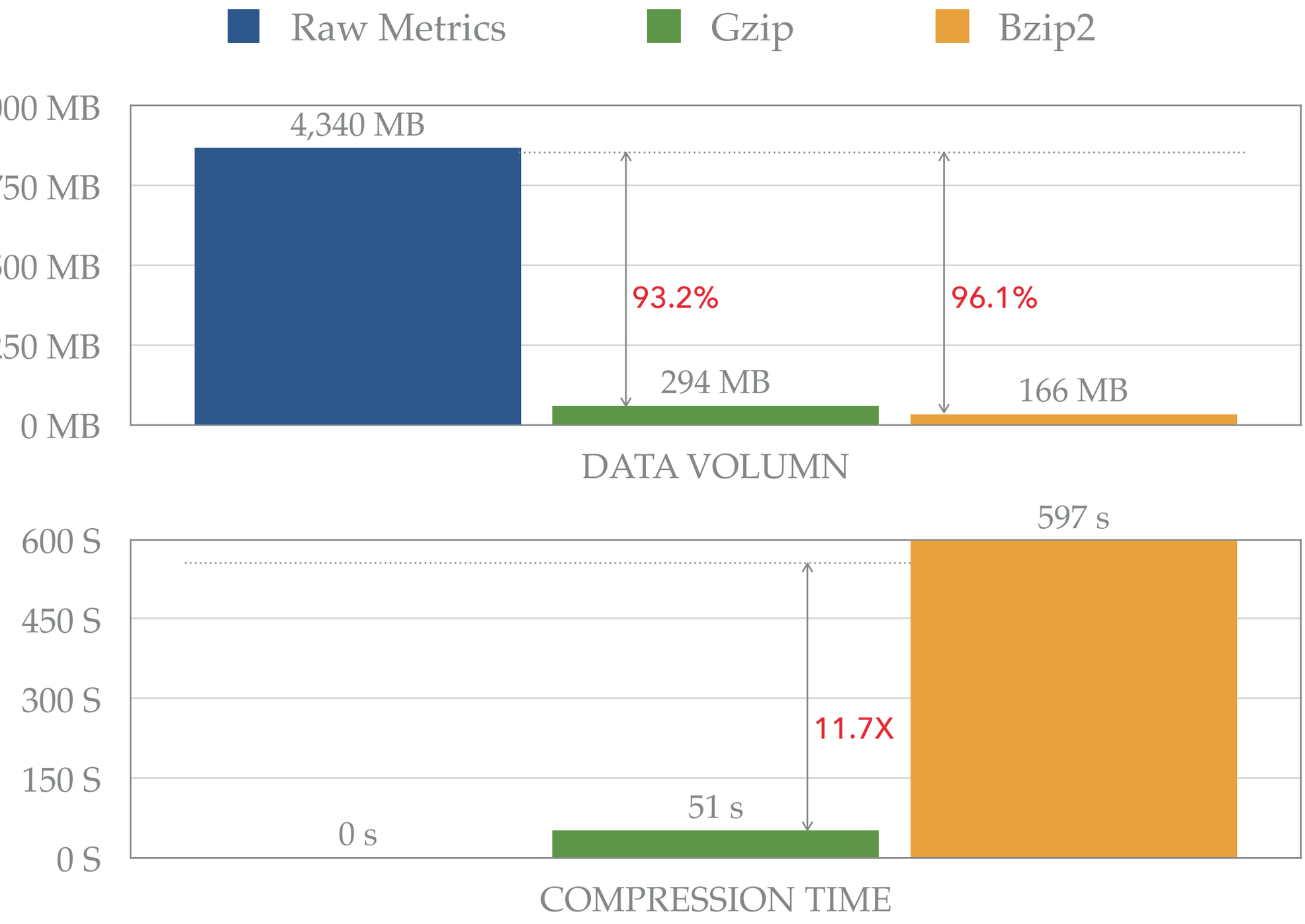
timestamp	nodeid	source	fqdd	value
2021-05-23 14:00:00-05	705	rawsensor	iDRAC.Embedded.1#SystemBoardPwrConsumption	308
2021-05-23 14:00:00-05	672	rawsensor	iDRAC.Embedded.1#SystemBoardPwrConsumption	672
2021-05-23 14:00:00-05	697	rawsensor	iDRAC.Embedded.1#SystemBoardPwrConsumption	308
2021-05-23 14:00:00-05	692	rawsensor	iDRAC.Embedded.1#SystemBoardPwrConsumption	308
2021-05-23 14:00:00-05	551	rawsensor	iDRAC.Embedded.1#SystemBoardPwrConsumption	658
2021-05-23 14:00:00-05	532	rawsensor	iDRAC.Embedded.1#SystemBoardPwrConsumption	308
2021-05-23 14:00:00-05	680	rawsensor	iDRAC.Embedded.1#SystemBoardPwrConsumption	308
2021-05-23 14:00:00-05	704	rawsensor	iDRAC.Embedded.1#SystemBoardPwrConsumption	644
2021-05-23 14:00:00-05	505	rawsensor	iDRAC.Embedded.1#SystemBoardPwrConsumption	308
2021-05-23 14:00:00-05	602	rawsensor	iDRAC.Embedded.1#SystemBoardPwrConsumption	308
2021-05-23 14:00:00-05	674	rawsensor	iDRAC.Embedded.1#SystemBoardPwrConsumption	322
2021-05-23 14:00:00-05	696	rawsensor	iDRAC.Embedded.1#SystemBoardPwrConsumption	308
2021-05-23 14:00:00-05	719	rawsensor	iDRAC.Embedded.1#SystemBoardPwrConsumption	644
2021-05-23 14:00:00-05	670	rawsensor	iDRAC.Embedded.1#SystemBoardPwrConsumption	574
2021-05-23 14:00:00-05	528	rawsensor	iDRAC.Embedded.1#SystemBoardPwrConsumption	616
2021-05-23 14:00:00-05	603	rawsensor	iDRAC.Embedded.1#SystemBoardPwrConsumption	308
2021-05-23 14:00:00-05	713	rawsensor	iDRAC.Embedded.1#SystemBoardPwrConsumption	308
2021-05-23 14:00:00-05	710	rawsensor	iDRAC.Embedded.1#SystemBoardPwrConsumption	336
2021-05-23 14:00:00-05	533	rawsensor	iDRAC.Embedded.1#SystemBoardPwrConsumption	616
2021-05-23 14:00:00-05	728	rawsensor	iDRAC.Embedded.1#SystemBoardPwrConsumption	308
2021-05-23 14:00:00-05	687	rawsensor	iDRAC.Embedded.1#SystemBoardPwrConsumption	308
2021-05-23 14:00:00-05	725	rawsensor	iDRAC.Embedded.1#SystemBoardPwrConsumption	308
2021-05-23 14:00:00-05	700	rawsensor	iDRAC.Embedded.1#SystemBoardPwrConsumption	308

COMPRESS METRICS - JSON



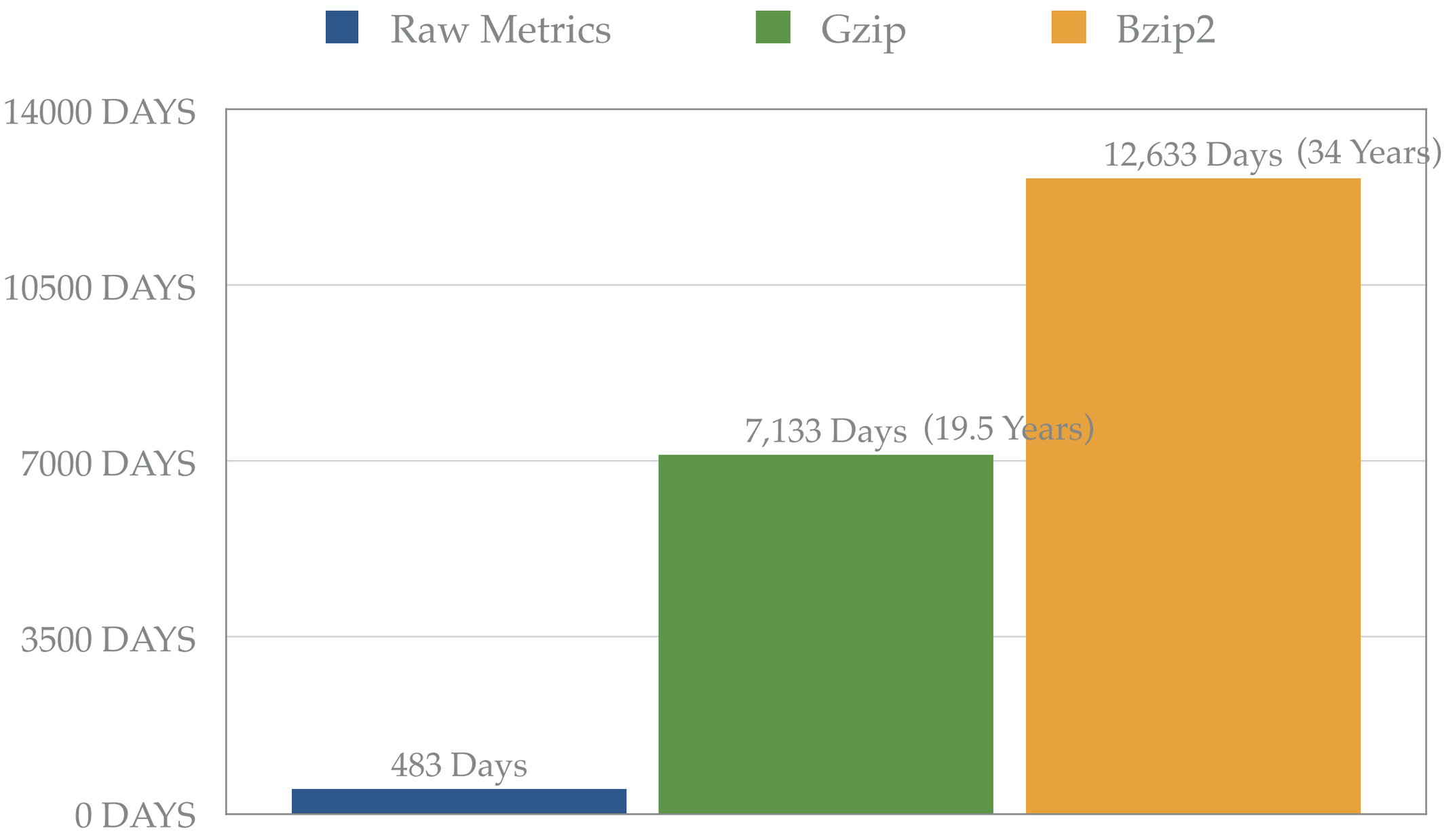
1 day of iDRAC9 monitoring data on the RedRaider Cluster

COMPRESS METRICS - CSV



1 day of iDRAC9 monitoring data on the RedRaider Cluster

COMPRESS METRICS - CSV



Suppose the available disk space is 2T

CONCLUSION

- Archiving monitoring metrics in files alleviates the data storage pressure.
- Storing data in CSV files consumes less storage.
- CSV files are much more simpler and more efficient to be loaded into TimescaleDB.

- TimescaleDB provides built-in compression mechanisms and deserves further exploring.