# Ancelus Database Query Efficiency

The Ancelus Database Management System is a graph database, not a Relational Database. It does, however, have the ability to duplicate queries based on the Structured Query Language (SQL) commonly found in Relational Database Systems.

The Ancelus Database Management System addresses these issues using a collection of optimization methods to produce extreme transaction performance that does not significantly degrade when the system requires big data and complex data.

To demonstrate this feature, the following seven table join benchmark prototype is provided.
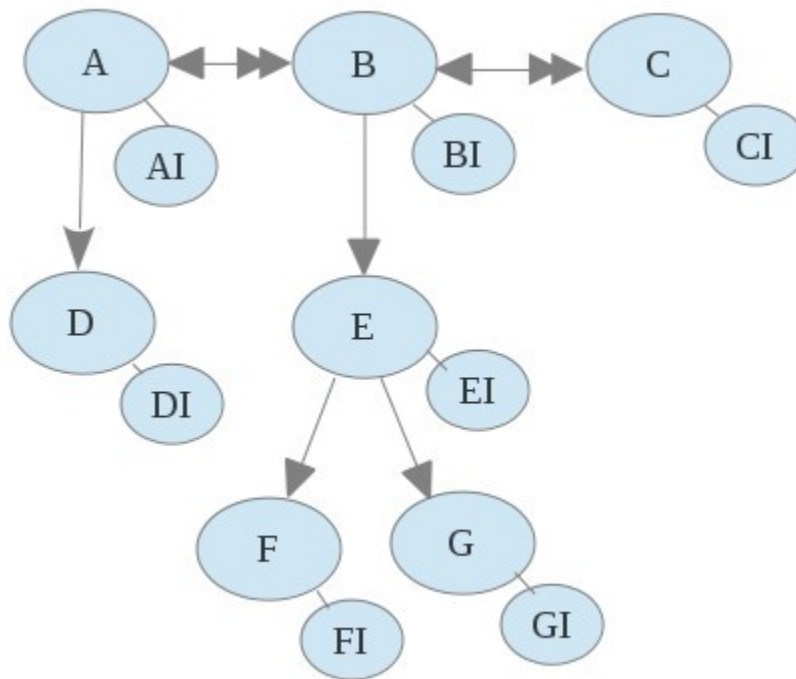


Table Size:
A    1 million rows
B    10 million rows
C, D, E, F, and G    1 billion rows

The Structured Query Language (SQL) statement that will be modeled follows:

```
SELECT A,AI,B,BI,C,CI,D,DI,E,EI,F,FI,G,GI
FROM A,B,C,D,E,F,G
WHERE B.A=A AND C.B=B AND A.D=D and B.E=E AND E.F=F and E.G=G
AND A BETWEEN ? AND ?
```

The Ancelus transaction benchmarks are compared to the Postgresql database management system. In each case, the results are observed to produce identical results.

**Create Database Tables (Schema)**
Ancelus and Postgres can create empty tables using negligible system resources (a few seconds). The difference is insignificant and not relevant to the overall performance comparisons.

**Load (Provision) Tables**
The Ancelus tables (and relations) and Postgres tables were loaded using comma separated variable (CSV) files. These files contained randomly generated character and floating point values.

**Load Times**

| Ancelus (minutes) | Postgres (minutes) | Ratio (P/A) |
|---|---|---|
| 207 | 5170 | 25 |

**Single Query Benchmark**
The seven table join query described above is now performed. The "A" table range (BETWEEN operation) was chosen to produce in excess of 100 million records.

**7 Table Join Benchmark (115.8 million records)**

| Ancelus (minutes) | Postgres (minutes) | Ratio (P/A) |
|---|---|---|
| 0.8 | 57 | 71 |

**Stress Tests - Multi-processor**
The stress test consists of 16 simultaneous 7 table join queries. Each query used a separate "A" table range select. The system power requirement was monitored. Approximately 340 Watts power utilization was observed in both cases.

**16X 7 Table Join Benchmark (1846.5 million records)**

| Ancelus | Postgres | Ratio (P/A) |
|---|---|---|
| 0.8 minutes | 436 minutes | 545 |
| .0045 KWH | 2.5 KWH | 545 |

**Overall Transaction Rate (Transactions per Minute)**

| Ancelus | Postgres | Ratio (A/P) |
|---|---|---|
| 2.3 Billion TPM | 4.2 Million TPM | 545 |

Ancelus scales up linearly on multiprocessor

**Stress Test Schema Change**

The next test measures the performance of the database if a simple schema change is attempted while the 7 table join query is in progress.  An item ("DI2") will be added to the table "D" 20 seconds after the start of the 16 simultaneous 7 table join queries.

**16X 7 Table Join Benchmark (1846.5 million records) with schema change**

| Ancelus | Postgres | Ratio (P/A) |
|---------|----------|-------------|
| 0.8 minutes | 608 minutes | 760 |

**Add column "DI2" to table "D"**

| Ancelus | Postgres | Ratio (P/A) |
|---------|----------|-------------|
| 0.077 minutes | 556 minutes | 7220 |

**Database Size**

| Ancelus | Postgres | Ratio (P/A) |
|---------|----------|-------------|
| 275GB | 787GB | 2.9 |

The Ancelus database is approximately one third the size of the Postgres database. In fact, the Ancelus database is smaller than the total size of the CSV files (321GB)  used to load both databases.  This is possible because the foreign keys in the Ancelus database are links to data rather than actual data.  The 40 character foreign keys  found in the CSV files require 4 bytes (32-bit integer) in Ancelus.

The Ancelus query performance improvements are enabled by these foreign key links. The Ancelus "join" operation requires 0 (zero) compares.  The links are generated at data load time, not at query time.  The "one to many" schema relationships are represented by internal link list structures in the Ancelus database and are also generated at data load time.

The Ancelus database reduction in power requirement may not be significant in a single server installation.  However, a large data center (or supercomputer) might include 100,000 server platforms.  At that scale, the full load power requirement would be reduced by ~300 million KWH per year.

**System Specifications**
Supermicro  AS-1123US-TR4 Server
CPU: 32 Core (2X AMD EPYC 7302)
Memory: 2048 gigabyte (32 X 64GB 2667MT/sec.)
Disk: 4TB U.2 NVME SSD  (2X 2TB Intel SSDPE2KX020T8)
        8TB SSD (Samsung 870 QVO MZ-77Q8TO)



Note:  This system is more than 4 years old.  On a currently available server platform, the Ancelus performance improvements would be significantly greater.

**Scale Down**
This level of query efficiency provides significant savings when the system requirements are scaled down.



Minisforum UM773
CPU: AMD Ryzen 7 7735HS 8 Core 16 Thread
Memory: 64 GB (2X 32GByte) DDR5 5600 MTPS
1 TB M.2 SSD - 2.5 Gigabit Ethernet port

This minicomputer costs ~$500 and consumes less than  500 KWH (~$100) per year.  It is capable of managing a large Manufacturing Execution System (MES) that interfaces to >500 machines and tracks >100,000 manufactured units and > 5 million transactions. (Ancelus RealTime Track and Trace).