

For a calendar of technical society meetings in the Mid-Hudson Valley go to:
<http://pok.acm.org/calendar.html> and/or to MHVLUG's calendar at <http://hvstem.org/>
Poughkeepsie Chapter of the Association For Computing Machinery



```

aaa          ccccccc  mmmmmmm  mmmmmmm
 a  a        cc      cc      mm mm  mm mm
aa  aa      cc      c       mm mm mm mm
aaaaaaaaa   cc          mm      mmm  mm
aa  aa      cc      c       mm  m   mm  MEETING NOTICE
aa  aa      cc      cc      mm          mm
aa  aa      ccccccc  mm          mm  Free and open to the public

```



Program: Database Operations on Compressed Data

When: 7:30 pm, Monday, November 21st, 2016

Where: **Hancock Center, Room 2023**, Marist College

Directions: **Building 14** on the map at <http://www.marist.edu/about/map.html>

Parking: Please park at black dot #10 on <http://www.marist.edu/about/map.html> (the lot North of the Hancock Center #14) or in the lot East of Route 9, S/E of the former Main Entrance.

Speaker: Luther Woodrum

About the Program: A new data compression method, called mr6, stores data for large files with compression ratios of as much as 9 to 1, while gzip gets ratios of 5 or 6 to 1 for the same files. The new compression method reads and uncompresses data 50% faster than gzip, in 2/3 of the CPU time for gzip. The method stores only one unique copy of each field in a table column, which allows new ways to do queries, such as join and outer join. Pattern matching is much faster by matching on the unique values instead of every record. In addition, queries that use pattern matching can have millions of query patterns that are performed on every record, in a time that is sub-linear with the number of queries. New operations on lists of lists will be presented for queries that operate on compressed data without needing to uncompress it first. A new query language is in progress, and some of its features will be discussed.

About the Speaker: Luther began programming in 1957 as an actuarial programmer at an insurance company in Chicago. While there, he wrote sorts and, in 1958, implemented what is now called pivot tables. In 1960, he decided to make a change to more involvement with computer development, and left the insurance company. He joined IBM in 1961 to work on the 7070 sort and spent the following years working on sorting, indexing, and storage allocation using radix partition trees. Eventually this led to a component of MVS, called RPTS. While working on the sort, he developed a method of ensuring program correctness so that the 8,000 line sort had no errors from the first release forever. In 1968 he attended the graduate school, IBM SRI, and majored in probability, statistics, and queueing theory. This led to a paper on semi-markov processes, published in the IBM Systems Journal in 1970. In 1987 he retired from IBM as a senior programmer. He was granted many patents by that time, and more since then. In retirement he builds computers, does contract programming, and develops applications for sorting, pattern matching, data compression, and a query language for operating on compressed database files, for a licensed program product.

Cost: Our meeting is **Free** and open to the public

Dinner: 6:00 pm, Palace Diner, 845.473.1576
 Map and menu: www.thepalacediner.com
 All are welcome to join us for dinner.

We thank Marist College for hosting the chapter's meetings.
Refreshments are served after the meeting.
For further information, email collier@acm.org or call 845.522.1971.

P - L - E - A - S - E P - O - S - T

