

# Big Data 100: Is There A Typical Big Data Workload ?

Chaitan Baru, Milind Bhandarkar, Raghu Nambiar,  
Meikel Poess, Tilmann Rabl

# Vision

The **BigData100 List** is a new collaborative effort between academia and industry to develop definitions and specifications for big data benchmarking and support publication of benchmark results, blending approaches from high performance computing ([top500.org](http://top500.org)) and transaction processing and database systems (Transaction Processing Council, TPC).

# Challenge

- Workload Specification
  - SPEC vs TPC Model
- Benchmark Evolution
  - Perpetual Beta & Versioning
- Benchmark Process
  - Specification, Auditing

# Discussion Logistics

- 3 groups – 3 rooms
  - Chaitan – linear
  - Raghu – square
  - Tilmann – constant
- Generate a 5 min presentation
  - Proposal for the Big Data 100 benchmark
- You have to earn your dinner!

# Room Mapping

Room	Room	Room
Chaitan	Raghu	Tilmann
Vinayak Borkar	Heinrich Stamerjohanns	Boris Glavic
Florian Stegmaier	Gopar Tadepalli	Brobst Stephen
Sanjeev Kumar	Ahmad Ghazal	Yogesh Dandawate
Venkat Lakshminarasimha	Mainkar Sameer	Mod Marathe
Irfan Modak	Avinash Rao	Rekha Singal

Questions ?