Who’s Down wit’ MPP?

hadoop and the Post-Relational Revolution

David Teplow

Who am I?

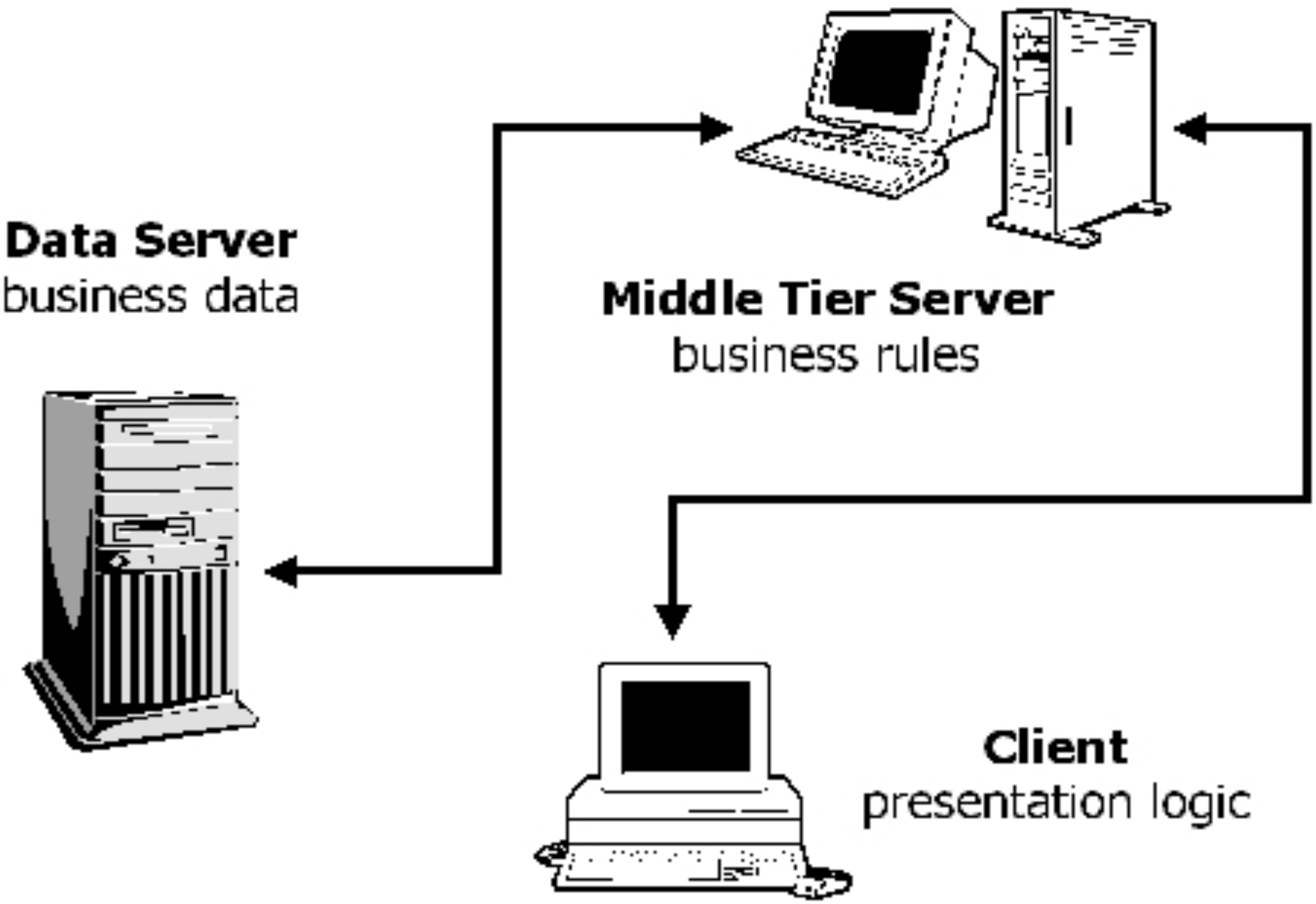
- Database architect / developer since the first release of the first RDBMS (Oracle v2; 1981)
- Helped launch the Northeast OUG in 1983
- Founded Database Technologies in 1986
- NOUG President from 1992 – 1999
- Founded Integra Technology Consulting in 2000
- Served on the IOUG board from 2003 – 2008
- Focused on Hadoop / Big Data since 2012

The Evolution of Computing

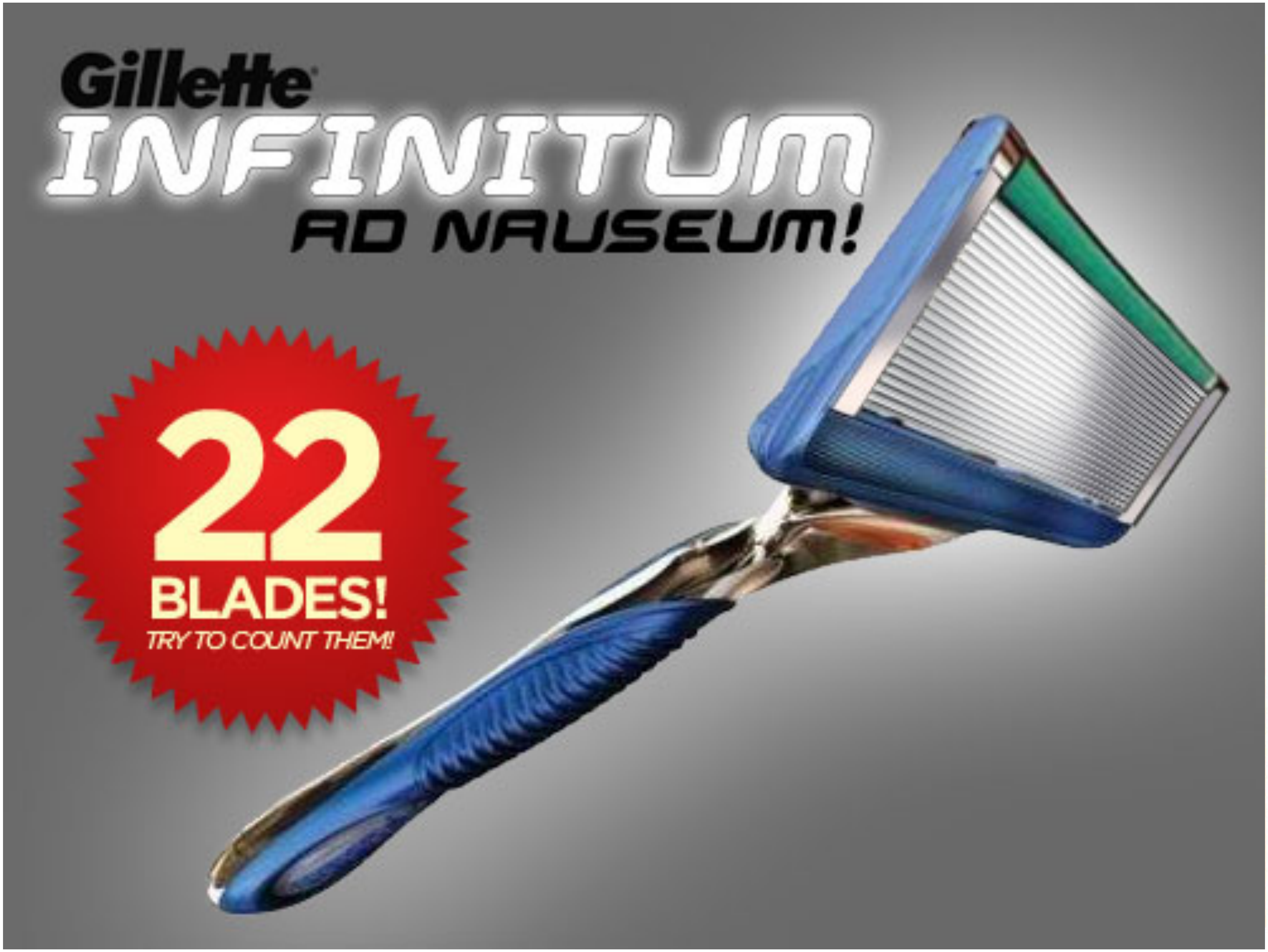
Mainframe

Client-Server

Three Tier



The Evolution of Shaving



The MPP Breakthrough

- In 2003 & 2004, Google published papers on:
  - Google File System: A way to distribute data across a self-healing cluster of servers.
  - MapReduce Processing: A way to process data that takes advantage of data locality and parallel processing.
- A team at Yahoo! borrows these techniques for its purposes, then creates ***Hadoop*** as an open source project in 2006

MPP Proliferates

2006 - **Hadoop** <http://hadoop.apache.org/>

2006 - **Google BigTable** <https://cloud.google.com/bigtable/>

2006 - **MarkLogic** <http://www.marklogic.com/>

2007 - **HBase** <http://hbase.apache.org/>

2008 - **Cassandra** <http://cassandra.apache.org/>

2008 - **CouchDB** <http://couchdb.apache.org/>

2009 - **MongoDB** <https://www.mongodb.com/>

2009 - **Redis** <http://redis.io/>

2009 - **Riak** <http://basho.com/products/>

2009 - **Voldemort** <http://www.project-voldemort.com/voldemort/>

2011 - **Couchbase** <http://www.couchbase.com/>

2012 - **AWS DynamoDB** <https://aws.amazon.com/dynamodb/>

2013 - **AWS Redshift** <https://aws.amazon.com/redshift/>

2015 - **Snowflake** <https://www.snowflake.net/>

INTEGRA
TECHNOLOGY CONSULTING

A Note About Spark

- Spark is an Open Source processing framework, (spark.apache.org) like MapReduce but better because it:
  - Uses Resilient Distributed Datasets (RDD) to harness a cluster’s memory (the way Hadoop harnesses a cluster’s disk and CPU)
  - Includes libraries for database access (Spark SQL)

    As well as for Streaming, Machine Learning and Graphics
  - ***Does not*** have its own storage layer, therefore ***requires*** HDFS, HBase, Cassandra, Amazon EC2, etc.

Classifying MPP Systems

- **By Structure**
  - ✓ Relational
  - ✓ Key-Value
  - ✓ Column-Oriented
  - ✓ Document-Oriented

✗

- **By Feature**
  - ✓Consistency
  - ✓Availability
  - ✓Partition Tolerance

✓

CHOOSE TWO

The CAP Theorem

The CAP Theorem states that you can only have two of the following features:

✓ **Consistency** - each client always has the same view of the data

✓ **Availability** - all clients can always read and write

✓ **Partition Tolerance** - the system works well across physical network partitions

CAP Properties

- **Consistency** is a continuum:

  Strong <------- consistency ------> **Eventual** / **Strict** (a.k.a. **ACID** versus **BASE**)

- **Availability** is also a continuum - measured in **Latency** (a.k.a. Response Time)
- **Partition Tolerance** is binary - either it does or it doesn’t

Consistency and Availability (C+A)

- Relational Databases
  - Oracle
  - SQL Server
  - MySQL
  - PostgreSQL
- Database Appliances
  - Teradata/Aster
  - Greenplum
  - Vertica

C+A Limitations

- Massive data volumes or massive concurrency requires MPP
- MPP requires Partition Tolerance
  - If each node in a cluster is 99.9% reliable

    = 1 failure every 3 years, on average
  - A 50-node cluster would be 95.1% reliable

    = 1 failure every 3 weeks, on average

Availability and Partition Tolerance (A+P)

- Cassandra *
- CouchDB
- Riak
- Voldemort
- AWS DynamoDB **

A+P Limitations

- Eventual Consistency might be adequate for shopping carts or certain purchases
- Fund transfers and bidding systems require Strong Consistency

Consistency and Partition Tolerance (C+P)

- Hadoop Distributed File System (HDFS)
- BigTable
- MarkLogic
- HBase
- MongoDB
- Redis
- Couchbase
- AWS Redshift
- Snowflake

C+P Limitations

- Availability takes precedence over Consistency when even small increases in latency cause business issues
  - Amazon claims to lose 1% in sales when latency increases 1/10 of a second
  - Google found that traffic drops 20% when latency increases 1/2 of a second

Git Down wit' MPP

- MPP is an elegant, affordable and scalable way to handle extremely high data volumes, throughput and/or availability requirements
- MPP requires Partition Tolerance
- Business requirements will dictate whether to choose Consistency or Availability when network failures (inevitably) occur

Thank You

Q & A

I welcome your further questions or comments:
DTeplow@IntegraTC.com

INTEGRA
TECHNOLOGY CONSULTING