The God Architecture

A Big Idea

Michael L. Brodie
mlbrodie@mit.edu

Abstract

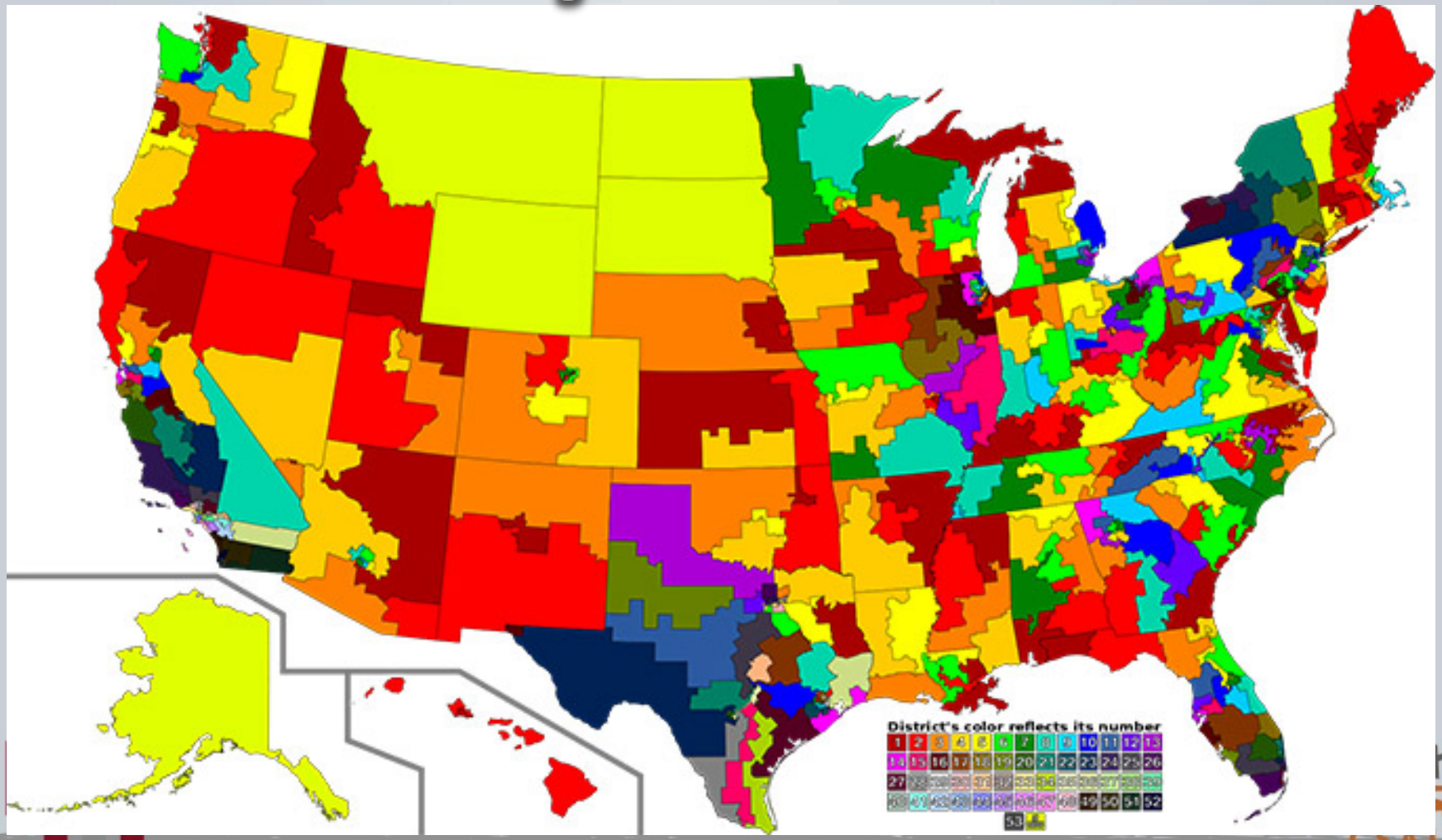
Computing architectures are *Gerrymandered* and *Balkanized* by computer science disciplines and by software products. Pernicious consequences include massively sub-optimal enterprise systems that are legacy before being operational and are almost impossible to interoperate with or migrate [1]; and inherently sub-optimal future systems. Colossal increases in computing and storage power have enabled disciplines to independently re-invent outside their core competencies. As advances slow, cost minimization may depend on collaboration over core competencies. A Big Idea for computing is an architecture that dynamically achieves some defined metric (e.g., cost minimization) by adapting to ever changing workloads and data sets to be developed collaboratively across computer science disciplines by setting aside some preconceptions and implied or explicit architectures. For example, database concepts and architectures have limited the potential benefits of data management. Two first steps could be to identify core competencies that each discipline could contribute, and to look for inspiration to autonomous, decentralized systems that work at scale, e.g., the nervous system of cephalopods [2] and humans [3]. If particle physics can find the God Particle for $10BN over a decade, computing should aim to provide greater value to the world by finding the God Architecture in half the time at half the cost**.

1. M.L. Brodie and M. Stonebraker. 1995. Legacy Information Systems Migration: Gateways, Interfaces, and the Incremental Approach, Morgan Kaufmann.
2. Octopuses, genes and intelligence: Tentacles that think: Studying cephalopods may illuminate the evolution of brains of all sorts, The Economist, August 15, 2015
3. What is consciousness? The hard problem The Economist, September 19, 2015

** In 1964 Peter Higgs and three others hypothesized the Higgs Boson as a fundamental particle that must exist to complete part of the standard model of physics. It was proven to exist in 2012 after 40 years of research and experiments including a $10BN 10-year development of the LHC at CERN, Genève. The Physicists called the Higgs Boson "The God Particle", but it was not "the supreme fundamental particle", but rather just another fundamental particle. And together with other bosons constitutes 0.3% of all matter leaving undiscovered Dark Matter and Dark Energy constituting 95.1% of all matter. The moniker may have contributed to enthusiasm to find it. Similarly, a flexible, adaptive computing architecture has been hypothesized for years. If achieved it would not only save trillions in IT costs, it would enable and accelerate discovery of all kinds. It is now time for the computer science community to do the research and empirical effort to realize the vision.

Most Pernicious US Political Problem

Gerrymandered



Pernicious Global Political Problem

Balkanized

Pernicious Consequences: Industry

- Verizon 10,000 systems are impossible to
  - Tune
  - Modify
  - Interoperate
  - Migrate

"We had systems from [the Occupational Safety and Health Administration] that were so old that they were buying parts on eBay, literally. We had a 30 year old system where the people that developed it were dead."

Labor Department CIO Dawn Leaf, September 15, 2015

Pernicious Consequences: Research

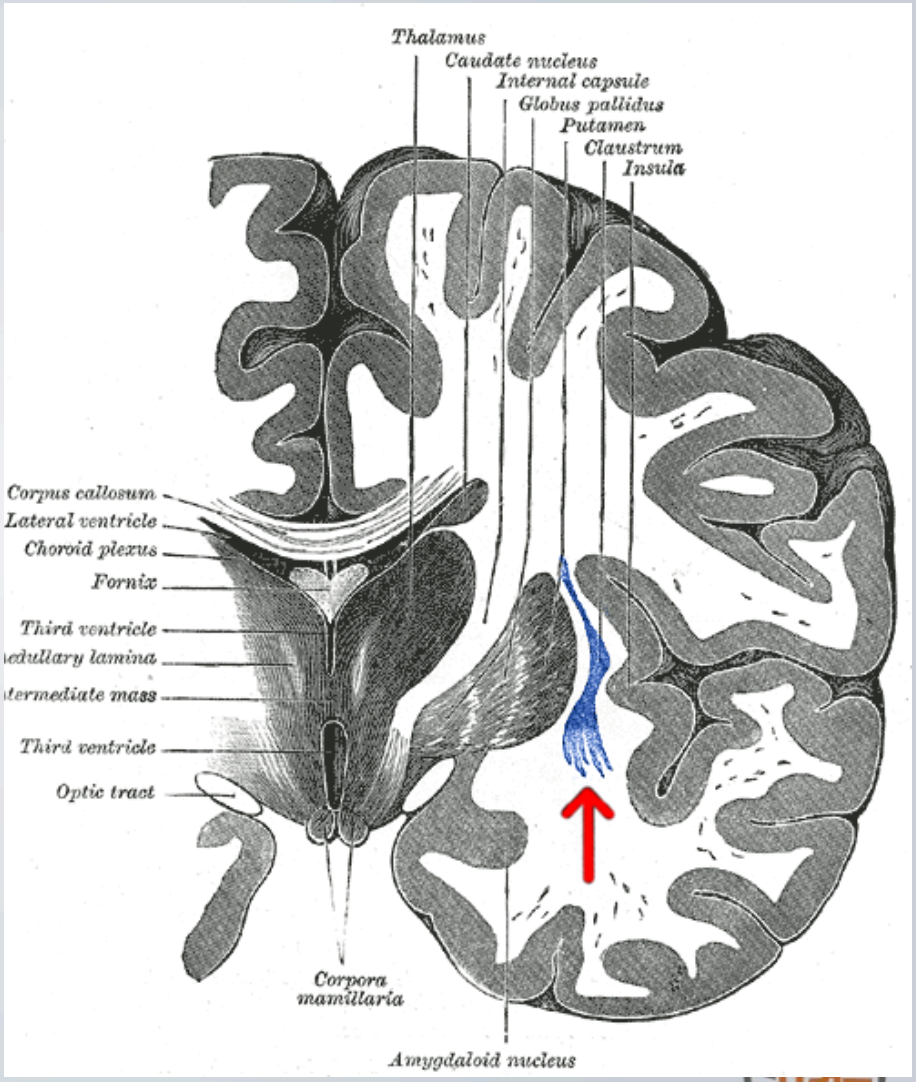
First Class (Data Model)	Array DMS*	Probabilistic Inference DMS**
Objects	Arrays	Imprecise data
Operators	Array operators (linear algebra)	Probabilistic Inference (statistics)
Performance		
Competitor	PCA, Hierarchical Data Format 5	MCMC / lifted inference
Improvement X	10 – 1,000	10-1,000
Applications	**Complex analysis**	**Complex analysis**
	Life sciences (genomics, drug discovery), astrophysics, …	Ranking, Top-K, K-means, recognition, mining, synthesis, …

* SciDB, tileDB, ** ProbDMS, Trio, Deepdive, PrDB, Orion

Big Idea
The God Architecture

An architecture that dynamically achieves a defined global metric by adapting to ever changing workloads & data sets

Inspiration: Nervous Systems



Metric

Cost minimization

Global

Approach

- Engineering POCs: Google, Facebook, Paypal, …
- Research: decade long activity
  - Collaborative (not Gerrymandered, Balkanized)
    - E.g., Computer Science Discipline core competencies
  - For given (workload, data set, metric)
    - Find compute / architecture patterns based on core competencies across
      - Computation
      - Communication
      - Storage

Core Competencies

Databases**

- ~50 ➔ 4
- Data language (algebra, SQL: R / W / search)
- Query optimization
- ACID
- Data recovery

** “Databases” considered harmful

Networking*

- Multi-party control systems over
  - Distributed State Abstraction
  - Specification Abstraction
  - Forwarding Abstraction

* Scott Shenker, UC Berkeley;
Mohammad Alizadeh, Stanford/CSAIL, MIT

Big Idea
The God Architecture

Particle Physics found the “God Particle” in 10 years for $10+ BN

Computer Science should find the “God Architecture” in ½ the time for ½ the cost