HUNTING THE DRAGON-KING: MULTI-DIMENSIONAL TESTING

Robert V. Binder

Co-Director, UIC Internet Assurance Lab

Midwest Verification Day - UIC

September 21, 2013

Overview

- Is software eating the world?
- Bugs
- Black Swans
- Dragon Kings
- Multi-dimensional Testing
- Challenges

“Software is Eating the World” -*Marc Andreessen*

Why Software Is Eating The World

By MARC ANDREESSEN

This week, Hewlett-Packard (where I am on the board) announced that it is expl
PC business in favor of investing more heavily in software, where it sees better p
Google plans to buy up the cellphone handset maker Motorola Mobility. Both m
But both moves are also in line with a trend I've observed, one that makes me op
of the American and world economies, despite the recent turmoil in the stock ma

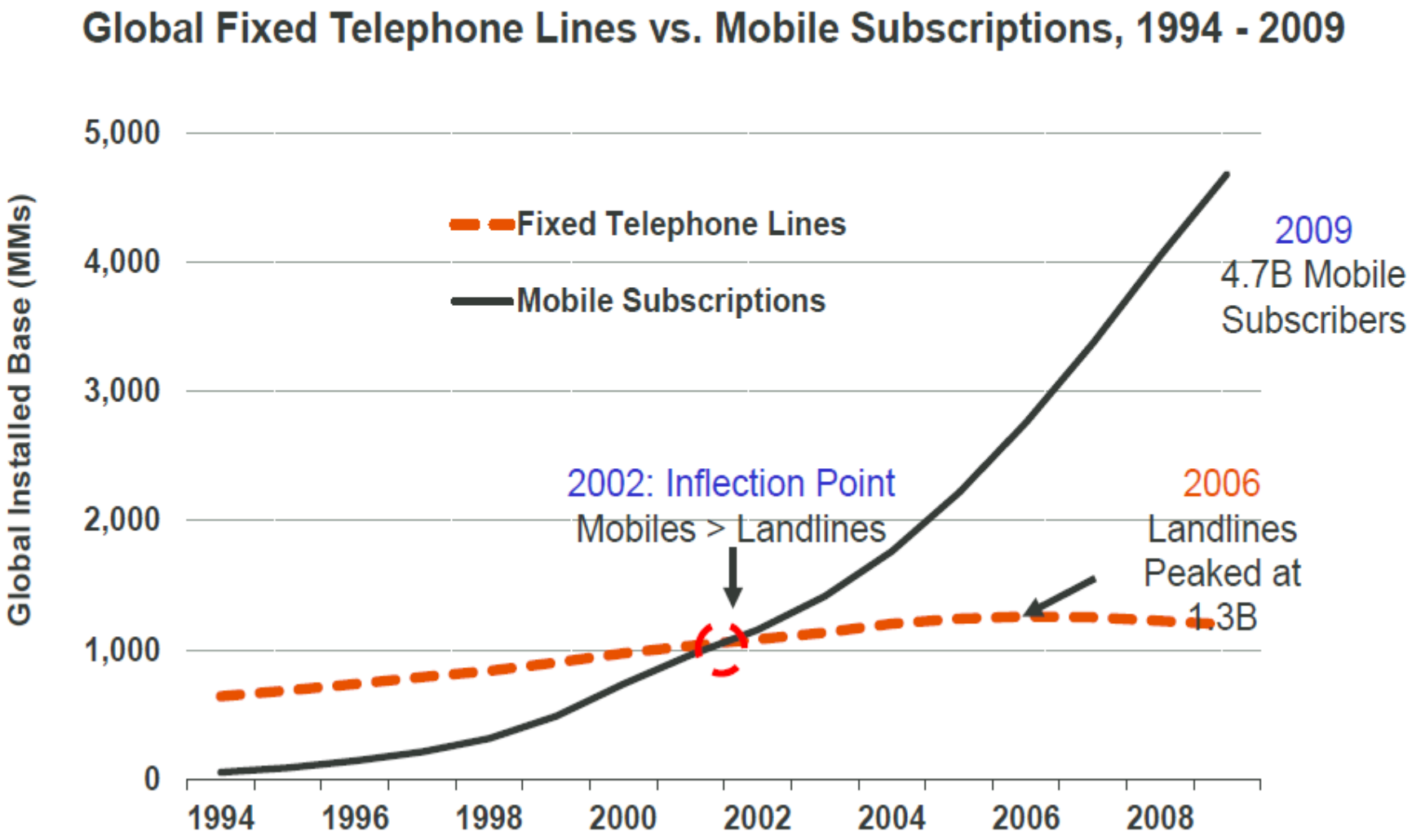
In an interview with WSJ's Kevin Delaney, Groupon and LinkedIn investor Marc Andreessen insists that the recent popularity of tech companies does not constitute a bubble. He also stressed that both Apple and Google are undervalued and that "the market doesn't like tech."

In short, software is eating the

More than 10 years after the pea
a dozen or so new Internet com
Twitter are sparking controvers
rapidly growing private market
occasional successful IPO. With
Webvan and Pets.com still fresh
are asking, "Isn't this just a dan

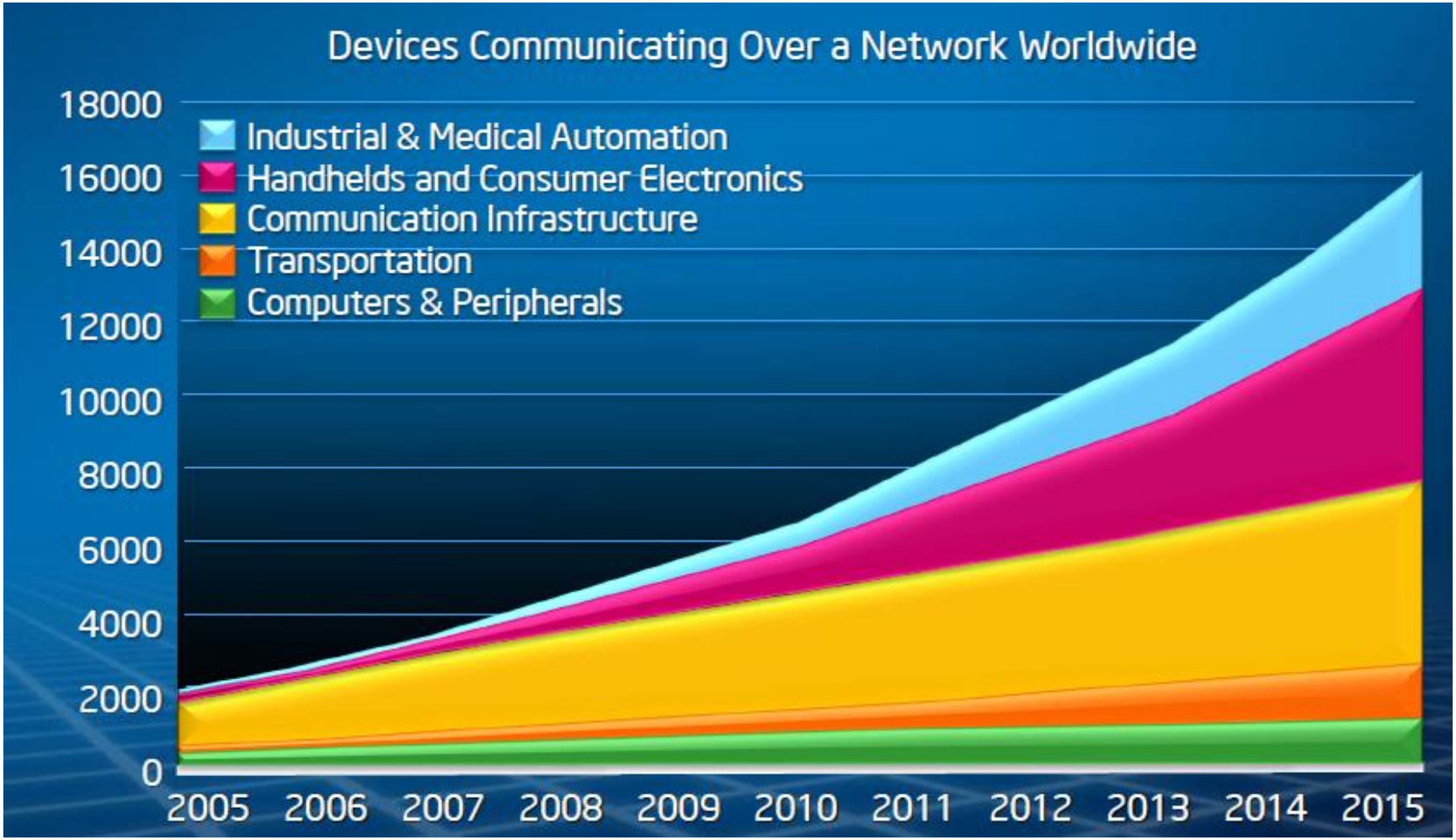
I, along with others, have been
case. (I am co-founder and gene
firm Andreessen-Horowitz, which has investe

First Generation of Re-Imagination - After 125 Years, Landlines Were Surpassed by Mobiles in 2002



Internet Trends. Mary Meeker, Kleiner-Perkins. D10 Conference May 2012

Embedded and Connected



2015 Connected Device Landscape: The Expanding Computing Continuum.
David Perlmutter, Intel. 2009

IT Disruption

By 2015, mobile application development (AD) projects targeting smartphones and tablets will outnumber native PC projects by a ratio of 4-to-1.

Through 2016, the financial impact of cybercrime will grow 10% per year, due to the continuing discovery of new vulnerabilities.

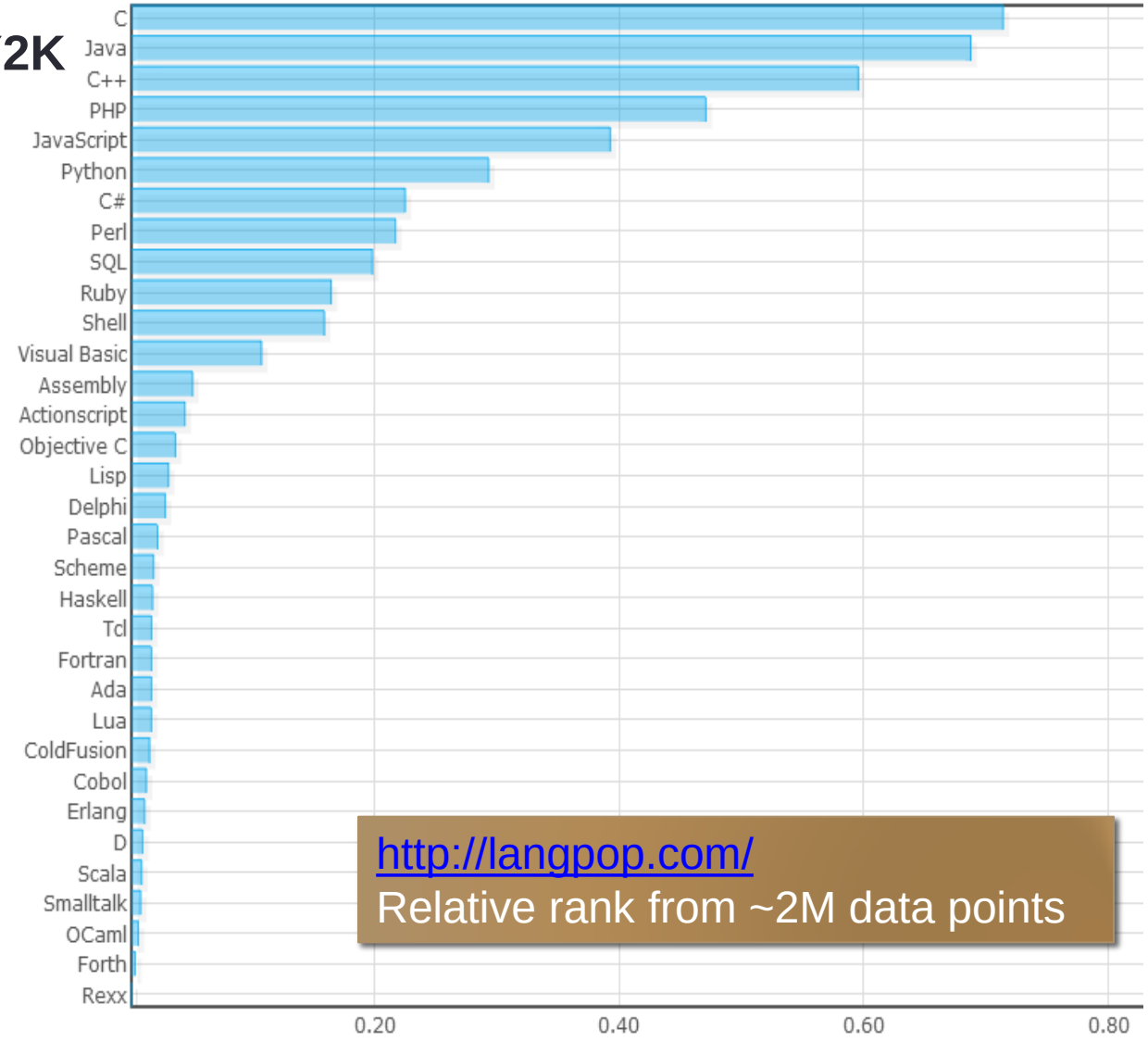
By 2016, 40% of enterprises will make proof of independent security testing a precondition for using any type of cloud service.

Gartner Top Predictions for 2012: Control Slips Away, Daryl Plummer. *March 2012*

Is *Software* really eating the world?

Methods/Tools since Y2K

- 2nd gen IDE
- MBT, MDD
- Open Source
- UML 2.0
- Cont. Integration
- Web 2.0
- Agile
- ISO 12027
- LAMP Stack
- Many protocols
- HTML 5
- iOS/Android
- Big Data
- Software Def Netwks



Is *Software* really eating the world?

Source of In-House Developed Software Code for Current Project
(Mean of Respondents)

	Auto	Mil/Aero	Medical
Hand Coded for the Current Project (e.g., C, C++, Assembly, Java, etc.)	56.0%	49.1%	62.8%
Generated Using Modeling Tools (e.g., UML/SysML, HMI, etc.)	15.5%	16.7%	9.5%
Leveraged from Previous In-House Projects/Designs	28.5%	34.2%	27.7%

Engineering Change to Accelerate Product and Systems Development.
Christopher Rommel, VDC Research. August 2012

Hardware is eating the world

Everything Computes & Connects

- 10x to 100x
  - ten years
- 100x to 1000x
  - twenty years

- Cheaper, better, faster
  - *Disrupt or die*

Software Methods and Tools

- ?X Productivity
- ?X Quality
- Last ten years?
- Last twenty years?

- Cheaper, better, faster
  - *“Choose two”*

Old bugs, new bugs

- No fundamental or transformative improvements in software
- **Bugs still as likely, more code**

- End of the uni-stack
- Multi-stack, mobile, connected
- **New failure modes**

What do you think will happen?

Old Bugs

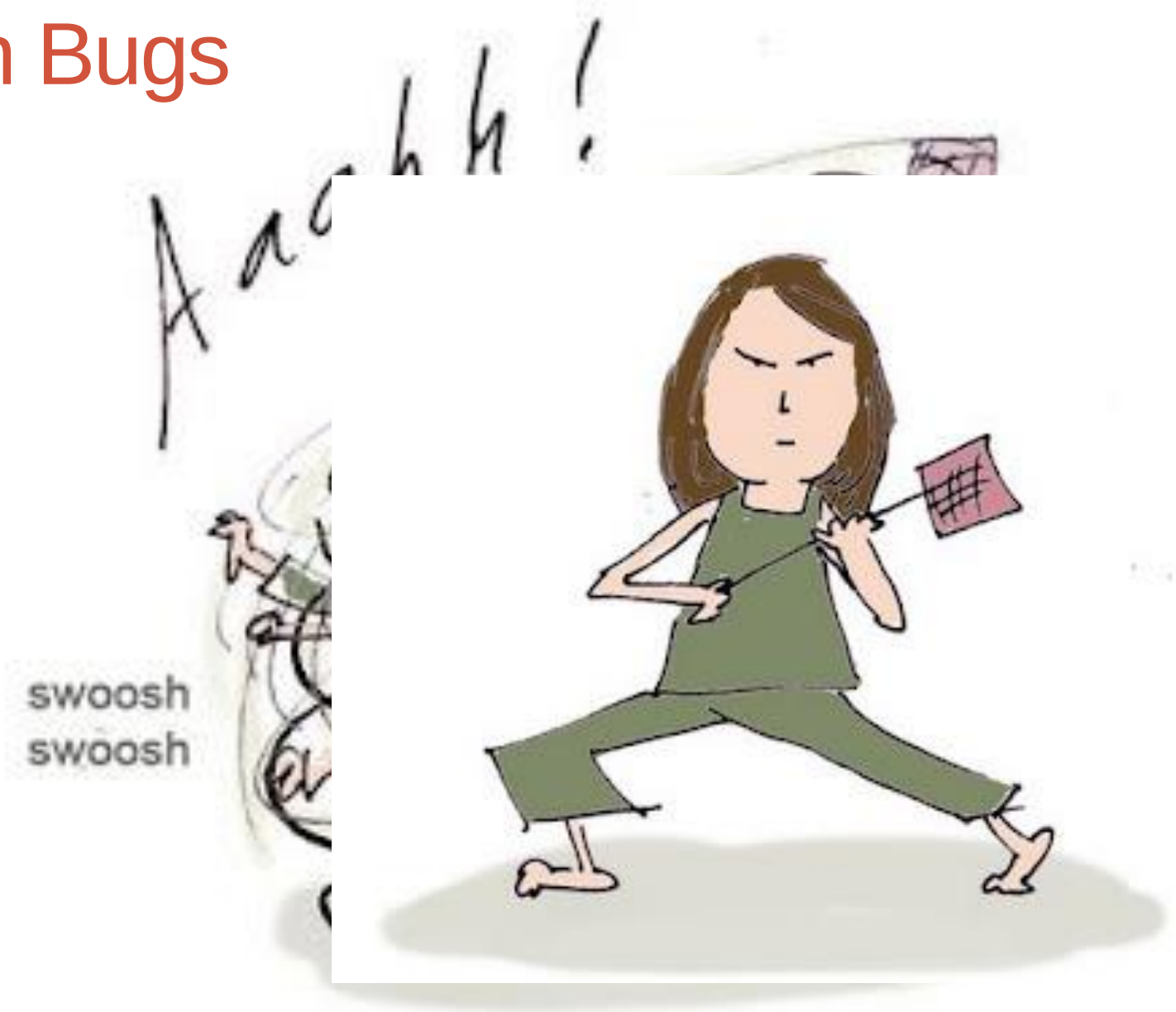
- Incorrect output
- Missing feature
- Abnormal end
- Missed deadline

- Bohr Bugs
- Mandel Bugs
- Heisen Bugs

Bohr Bugs

Mandel Bugs

Heisen Bugs

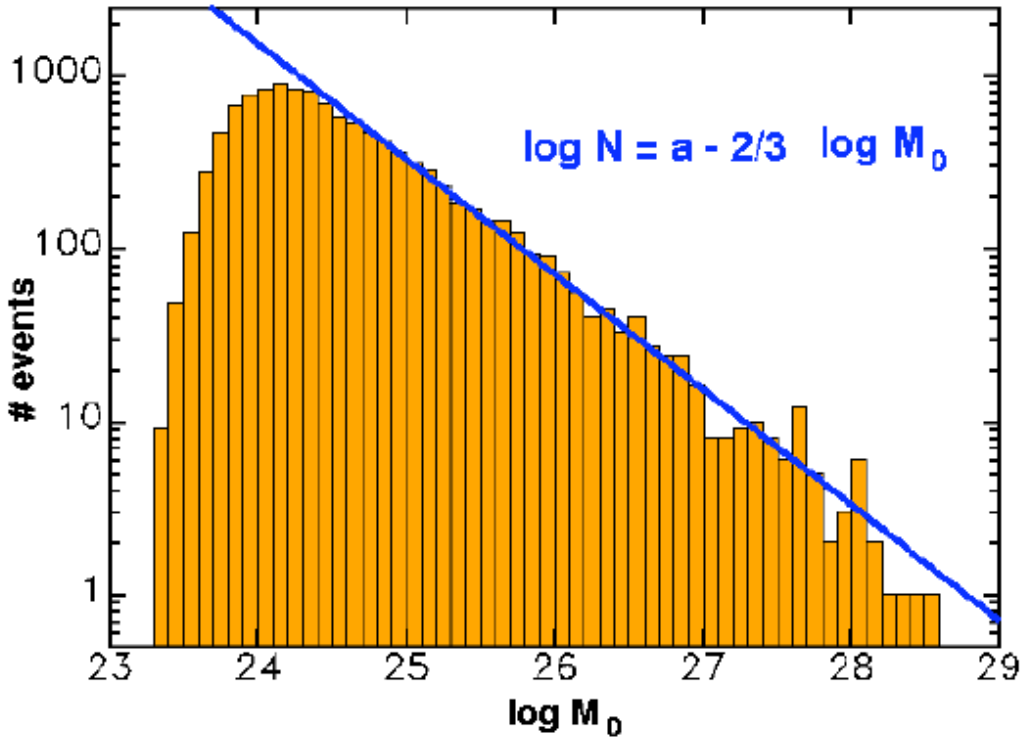


New Bugs

- Power Law Models
- Black Swans
- Dragon Kings

Power Law Models

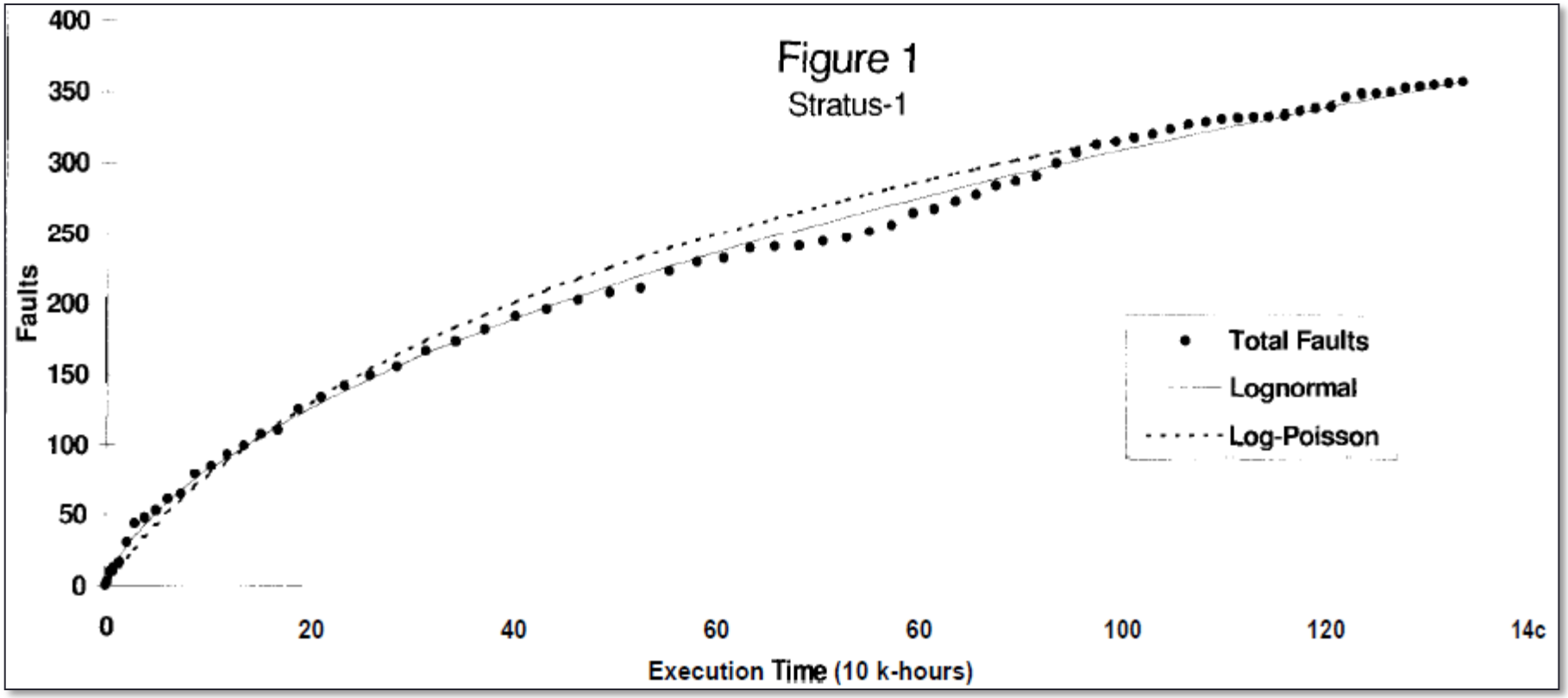
- Many natural phenomena
- “Fat Tails”



Dragon-Kings, Black Swans and the Prediction of Crises. Didier Sornette

http://www.youtube.com/watch?v=FlTSbzOvKZI

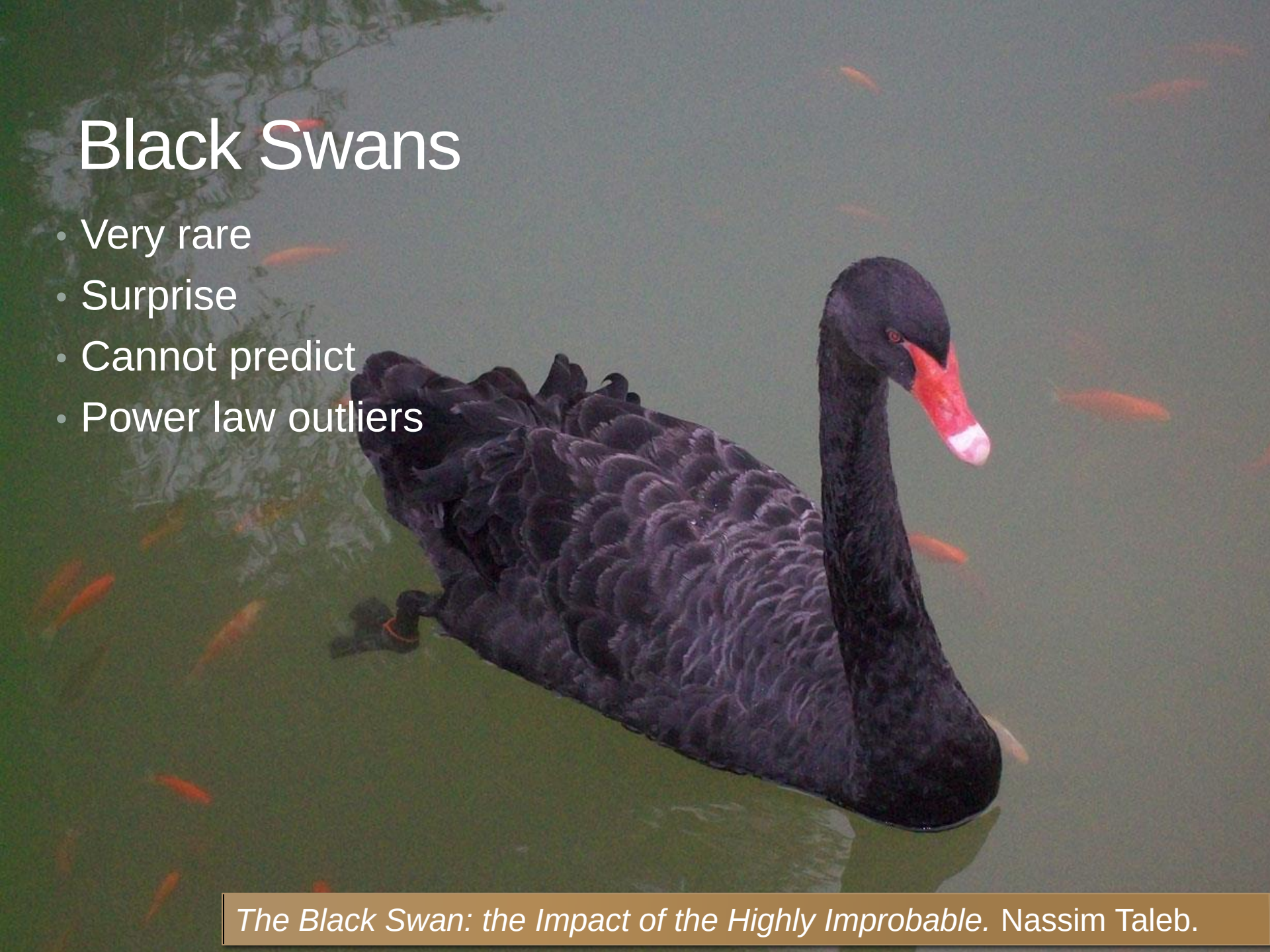
Reliability Growth



The Lognormal Distribution of Software Failure Rates: Application to Software Reliability Growth Modeling. Robert E. Mullen

Black Swans

- Very rare
- Surprise
- Cannot predict
- Power law outliers



The Black Swan: the Impact of the Highly Improbable. Nassim Taleb.

Software Black Swans

- Total eclipse
  - Once every 360 years
- Multi-cycle race conditions in embedded systems

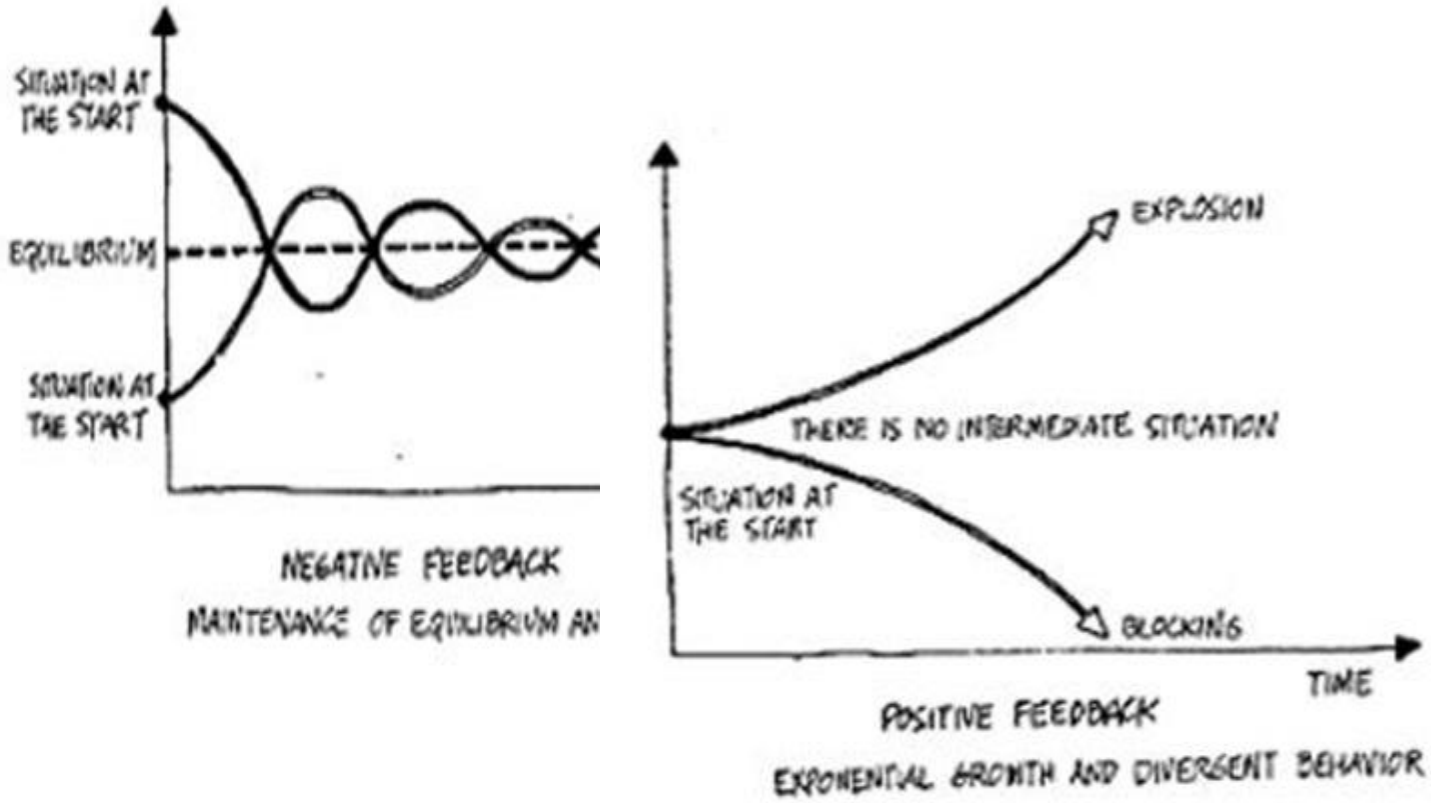
Dragon Kings

- Extreme Outlier
  - Beyond fat tail
- Not “random”
- Latent Positive Feedback
- Phase Shift
  - Earthquakes
  - Material Fractures
  - Epileptic Seizures
  - Market Crashes

Dragon Kings

Sornett calls these exceptional events dragon kings "to stress that we deal with a completely different kind of animal, beyond the normal, whose extraordinary characteristics [have] profound significance."

Latent Positive Feedback = Dragon King

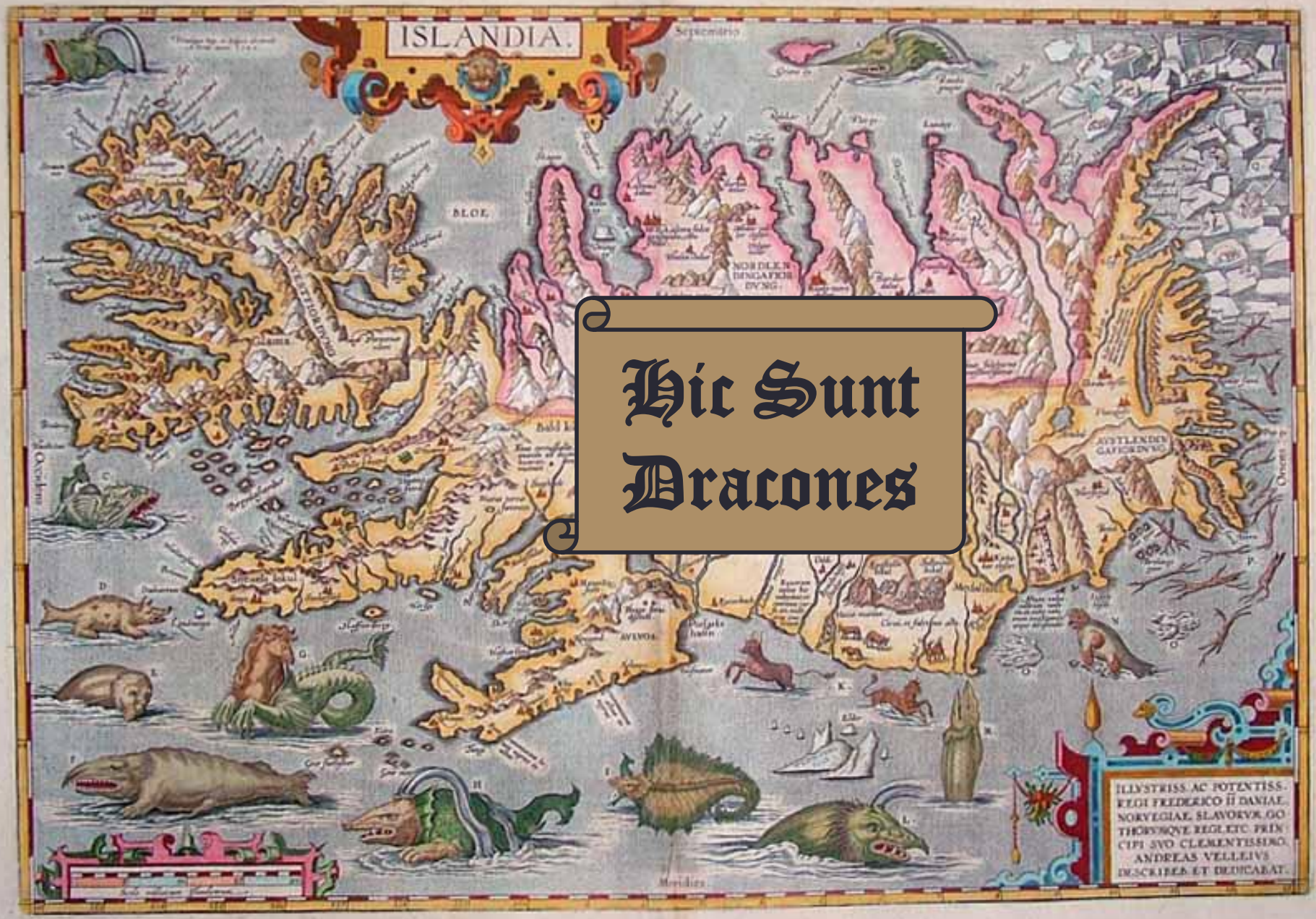


New Failure Modes: Amazon Cloud Outage

- Amazon web service Elastic Compute Cloud crashes (power failure)
- The Elastic Load Balancing (ELB) system frantically tries to assign workloads to available servers.
- As Amazon's cloud rebooted, "a large number of ELBs came up in a state which triggered a bug we hadn't seen before"
- ELB reaction: try to allocate more larger servers.
- Backlog in the "control plane" results.
- Demand from customers in unaffected availability zones continues
- System swamped and crashes again.

What kind of failure was this?

- Why didn’t test find it?
- What would we have to do to find these kind of bugs?


ISLANDIA.
Septentrio
Hic Sunt Dracones
ILLVSTRISS. AC POTENTISS.
REGI FREDERICO II DANIAE,
NORVEGIAE, SLAVORVM, GO
THORVMQVE REGI ETC. PRIN
CIPI SVO CLEMENTISSIMO,
ANDREAS VELLEIVS
DESCRIBEB. ET DEDICABAT.
Meridies

Here be Dragons

- ✓ **Latent positive feedback**
  - **External disruptors**
  - **Malicious attack target**
- ✓ **Partial degradation**
  - **Provoked panic**
- ✓ **Low testability**
  - **Multi-stack**
  - **Field-infeasible**
  - **Uncontrollable Inputs**

Looking under the Model-based Lampost?

Logic errors and incorrect assumptions are inversely proportional to a path's execution probability.

Boris Beizer

- 9 in 10 Functionality only
- Climbs a tree, not the twigs
- Typically no load test

Multi-Dimensional Testing

Behavior

- Realistic variation
- Markov event/data
- Include
  - Fat finger
  - Abuse cases
  - Sneak paths
  - Panics

Interfaces

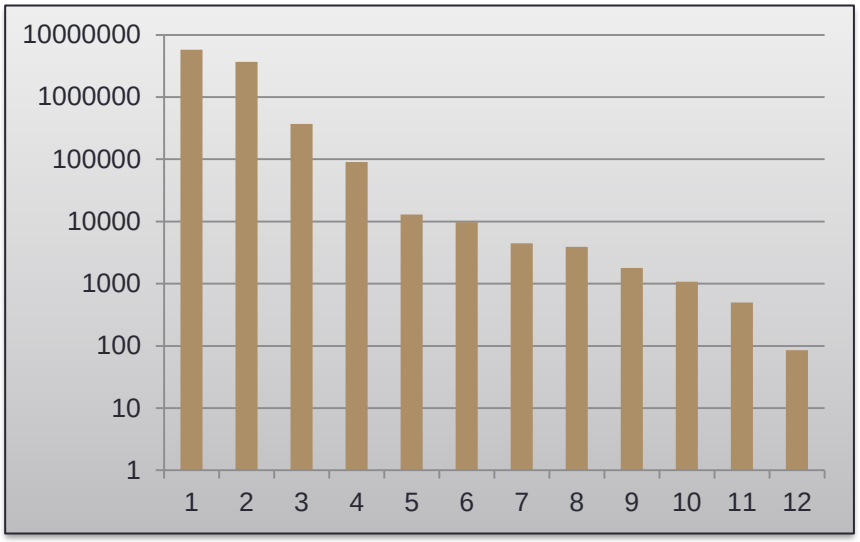
- Drive all interfaces
- Disrupt resources

Load

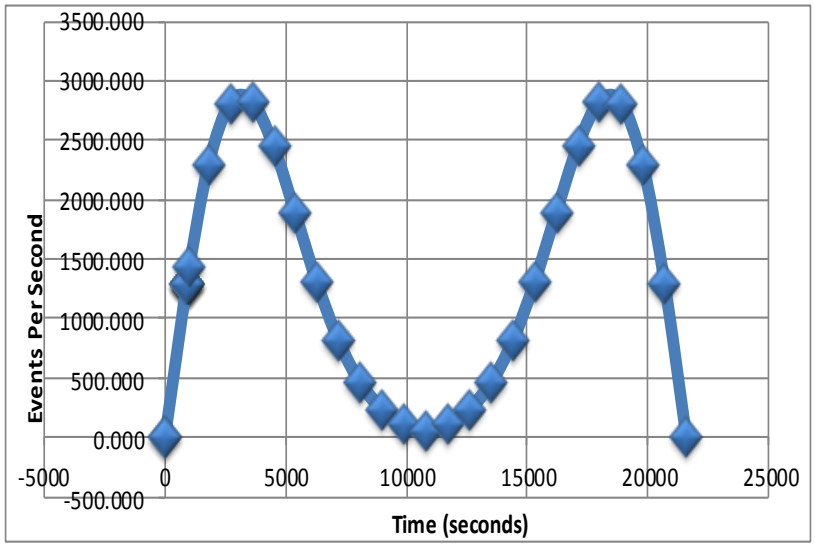
- Vary input rate wrt time

Not possible without full automation

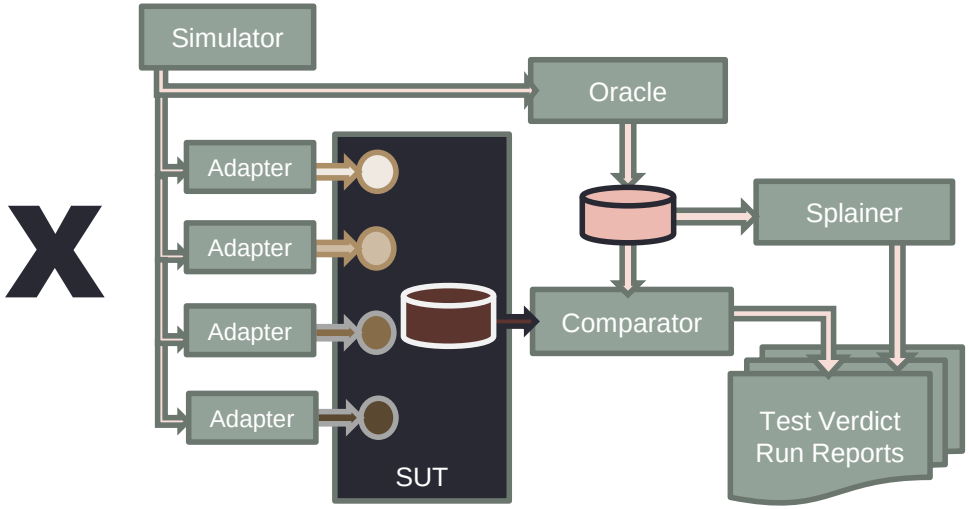
Interleave Behavior, Load, Interfaces



X



X



=

Hunting the Dragon King

	Common Bugs	Black Swans	Dragon Kings
Fragile	1	0	0
Massively Manual	2	0	0
Test Driven Dev	3	1	0
Exploratory	3	1	0
Crowdsourced	2	0	0
Functional MBT	5	3	1
Multi-dimensional MBT	5	4	3

- 0 You must be joking
- 1 Write if you get work
- 2 Even a broken clock is right…
- 3 One Sigma
- 4 Michael Jordan free throw (82%)
- 5 Sargent Preston

What have we learned?

- *Information Technology* is eating the world
- Common bug rates will **not** decline
- **More** new systems will have dragon kings
- Functional/model coverage MBT **will not** find dragon kings
- Multi-dimensional MBT **can** find dragon kings
- Opportunity for test engineering to **lead**

May you live in interesting times …

Challenges for the MBT Community

Tool Providers

- Start shipping multi-dimensional solutions

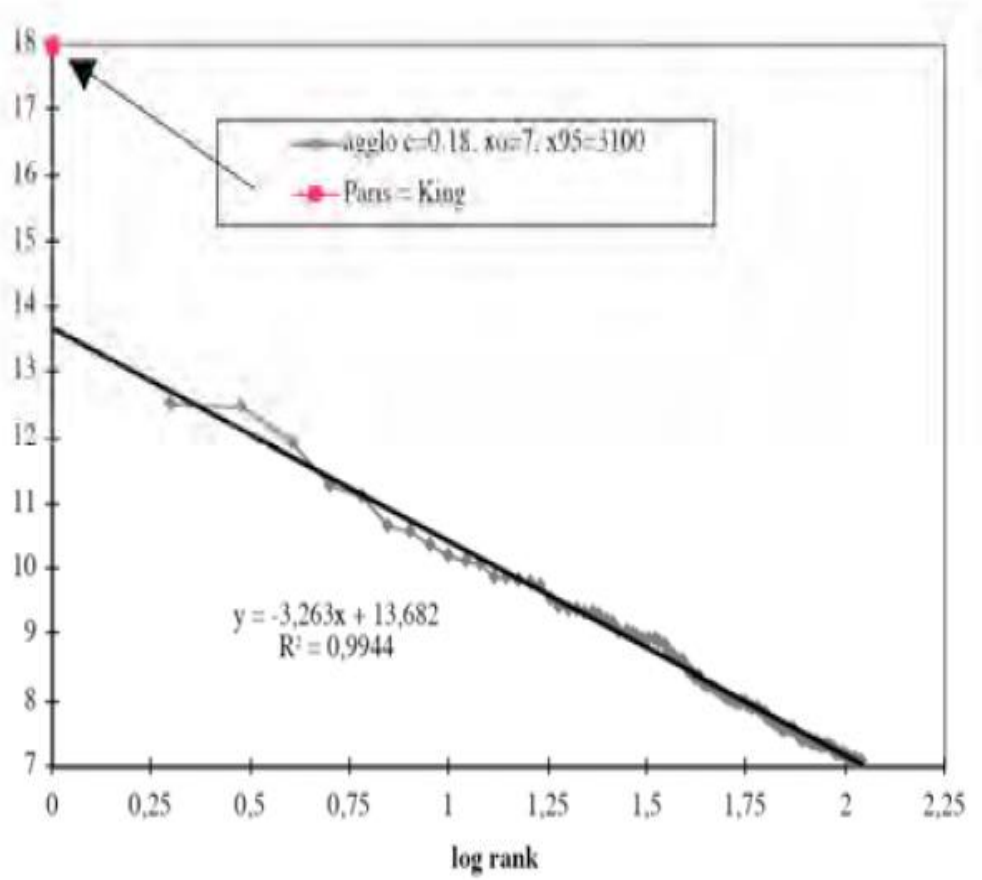
MBT Test Engineers

- Learn to lead
- Start hunting your dragon kings

Thank you !

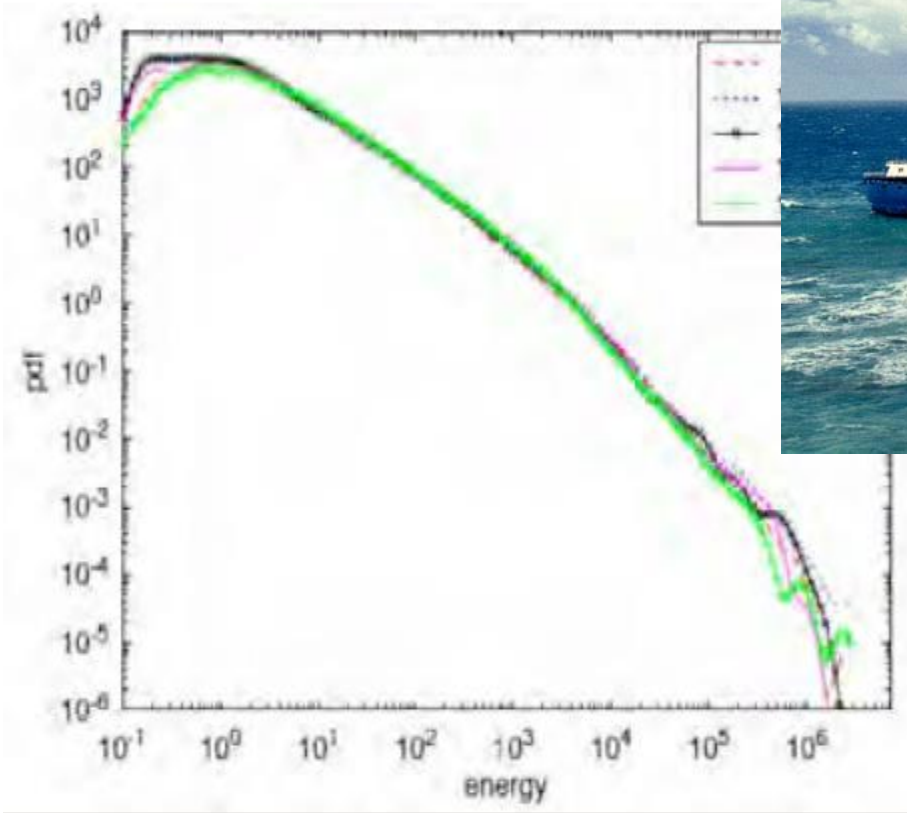
King Effect

- *More* extreme than power law outliers
- “King effect”



Dragon-Kings, Black Swans and the Prediction of Crises. Didier Sornette

Stress versus Fracture Energy



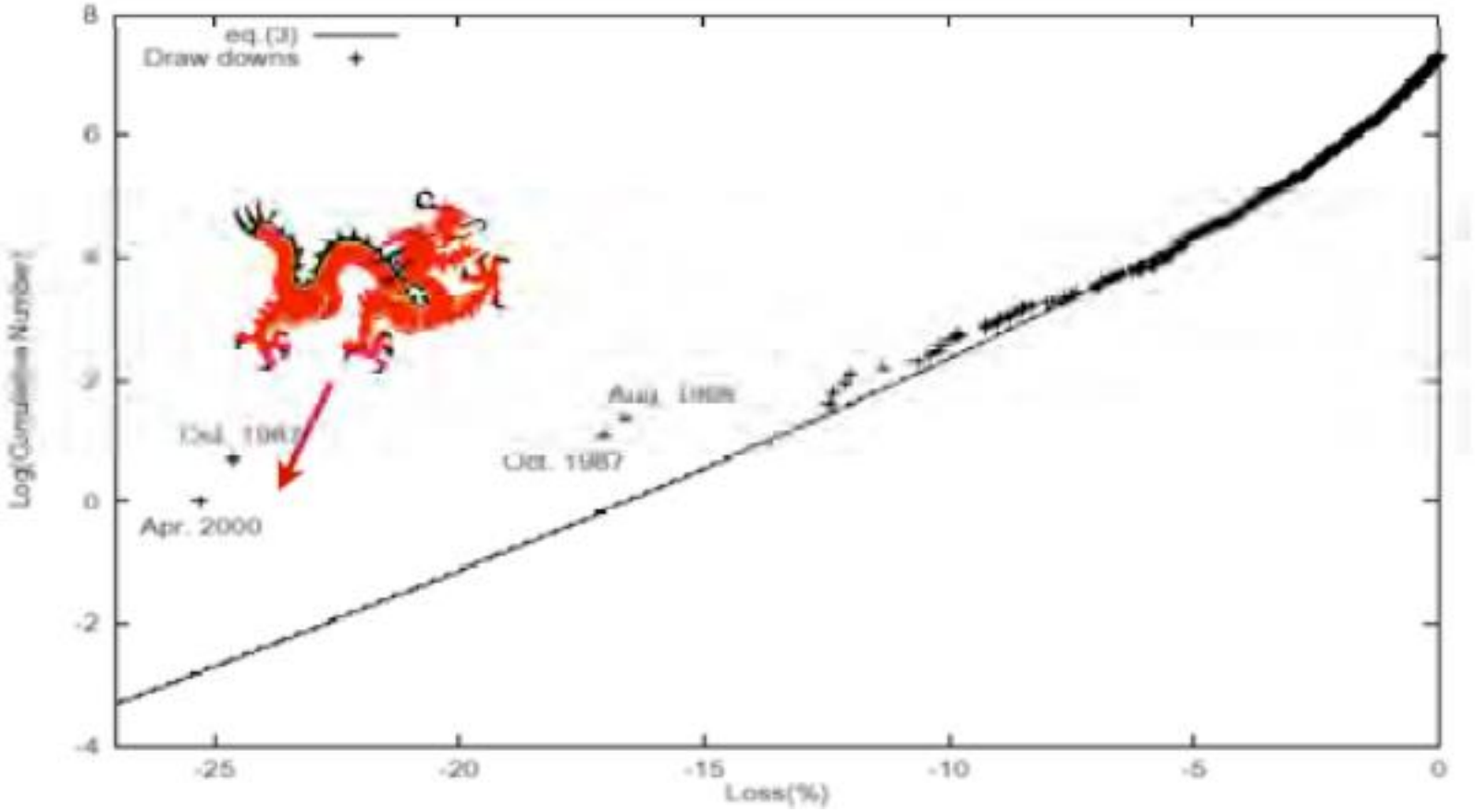
...

10^{10}

the final rupture associated with a run-away crack releases much more energy

Dragon-Kings, Black Swans and the Prediction of Crises. Didier Sornette

Market Crashes



Dragon-Kings, Black Swans and the Prediction of Crises. Didier Sornette

MBT and Agile

- Don’t waste your time with Fragile
- Agile Pain Points
  - Regression testing gap
  - Superficial testing
    - TD Development
    - Acceptance Tests
- The power of MBT is in the models
- **Test-Model** Driven Development
  - Rapid modeling
  - Some, maybe a lot of help for the RT trap