

Microsoft



Microsoft® Research

Faculty Summit 2012

Riviera Maya, Mexico | May 23-25 | In partnership with CONACYT



Why Software Engineering is the Best Job in 2012

Judith Bishop
Director of Computer Science
Microsoft Research Connections

From a report in March 2012

Based on

- Work Environment
- Physical Demands
- Job Outlook
- Income Levels
- Stress



<http://www.careercast.com/jobs-rated/10-best-jobs-2012>

Some History

Eclipse software development environment with its extensible plug-in system.

- From IBM Canada, 2001
- Now in a Foundation
- Free and open source under its own license
- Strong community base



Java programming language and its run-time platform

- From Sun Microsystems in 1995
- Now with Oracle
- Free and Open Source under GNU public license
- Strong research, education and developer base
- Part of browser technology



Microsoft collaborates with universities

- Why?

- Expand a lab's research base from 100s to 10,000s
- Verify the work in practice
- Launch new applications
- Support tomorrow's leaders

- How?

- Awards, grants, prizes
- Interns (1,000 a year)
- Visitors
- And more





Amazing numbers

- More than **100** top PhD research students from leading European universities in Microsoft Research Fellowships
- More than **25 000** scientists, academics and students have attended Microsoft-sponsored conferences, summits and workshops since 2005
- Nearly **1 000** interns selected for internships at our labs in Redmond, China and India annually
- **465** students from more than 35 countries have interned at the Cambridge lab in the past 7 years
- **50** conferences sponsored in the past 5 years for \$350,000

Computer Science Activities

- Innovation Foundation (SEIF)
- Interaction with RiSE
- Promoting Visual Studio through ICSE events

Software Engineering



- Interaction with NLP, ISRC, Bing
- Developed Language Models Services
- Bing-MSR Speller Challenge
- Developing Knowledge Services

Semantic Computing



- Interaction with Network Research team in MSR-R
- Student projects on WP7 phones at 17 universities
- Promoting Azure and WP7

Mobile Computing



- Book on Patterns and Practices group
- Course on Concurrency and Parallelism Promoting C# and F#

Parallelism and Concurrency



- Developing a Try F# system in a browser
- Workshops on "F# in Education"
- Supporting Pex4Fun
- TouchDevelop outreach and book

Programming Languages



- CS EdWeek
- Work with the NSF, ACM, IEEE and IFIP
- MSR Software Summits
- Support for 12 conferences

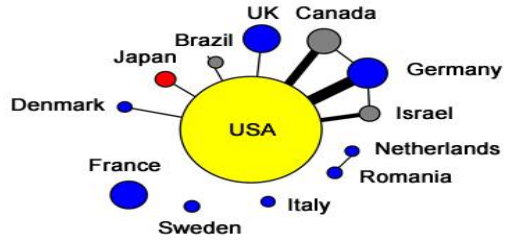
Community interaction



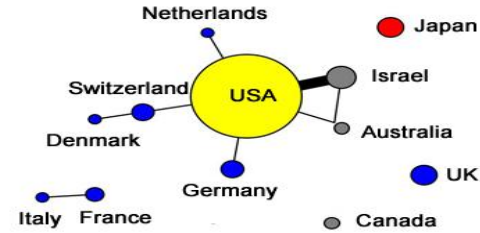
CS is research-oriented with strong internal and external impact

How collaboration has grown, 1996-2011

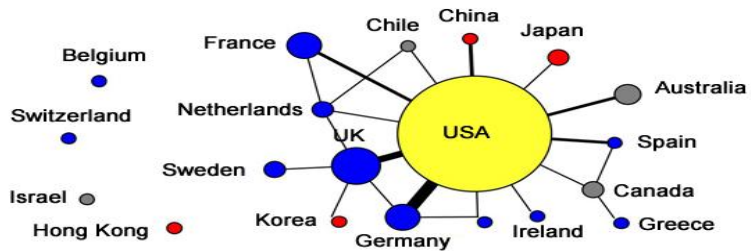
Mapping Collaboration Networks in 'Programming' Conferences



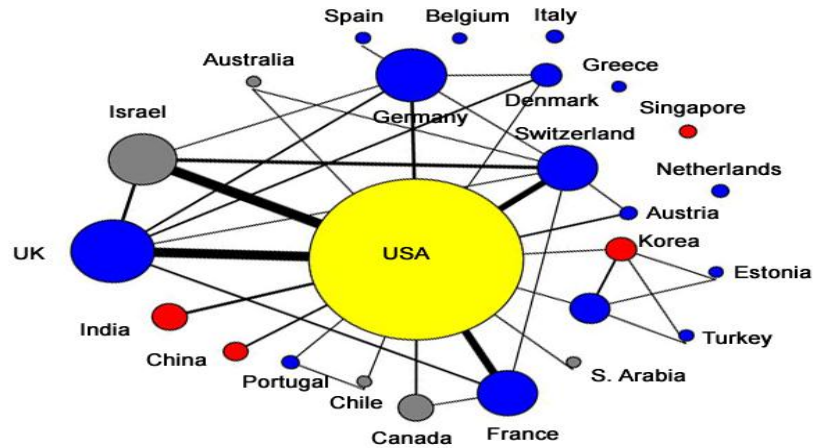
1996



2001



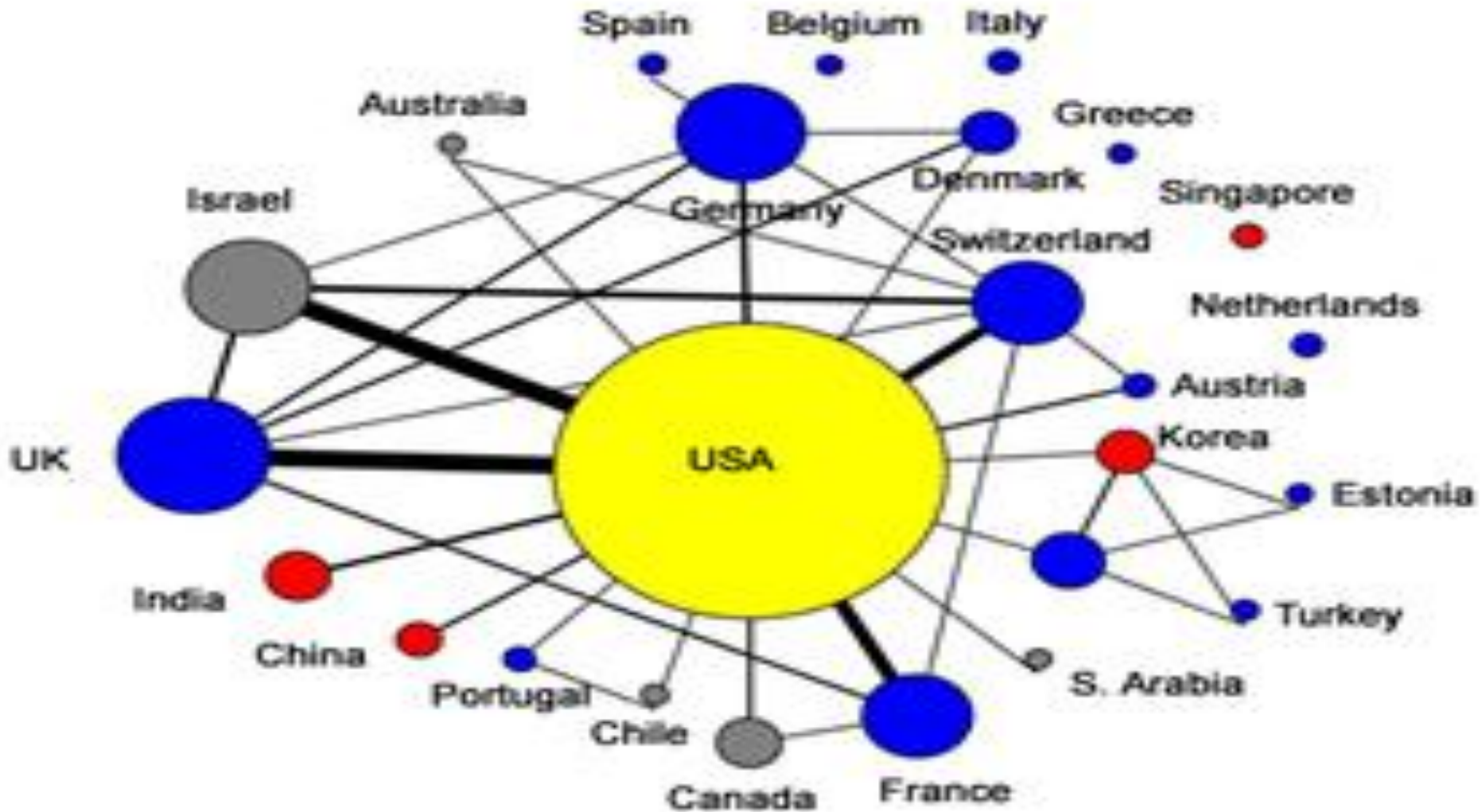
2006



2011

* Circle size represents number of publications authored by each country. Lines between countries indicate publication co-authorship. Line width indicate the degree of collaboration (wider = more collaborations). Conferences analyzed include PPOPP, POPL, PLDI, and OOPSLA.

PL Collaboration 2011



Top 12 CS Research Organizations



Advanced Search

- Author »
- Publication »
- Conference »
- Journal »
- Organization »
- Keyword »

Academic > Top organizations in Computer Science

1 - 100 of 9,599 results

Organization	Publications	H-Index
Stanford University	35354	327
University of California Berkeley	32888	315
Massachusetts Institute of Technology	36272	312
Microsoft	36967	303
Carnegie Mellon University	38826	272
IBM	47624	261
Princeton University	11242	211
Cornell University	13488	208
University of Illinois Urbana Champaign	26523	204
AT&T Labs Research	9471	203
University of Washington	15524	202
Google Inc.	9459	192

Top Software Engineering organizations last 5 years



[Advanced Search](#)

- [Authors »](#)
- [Publications »](#)
- [Conferences »](#)
- [Journals »](#)
- [Organizations »](#)
- [Keywords »](#)

[Academic](#) > Top organizations in Software Engineering

1 - 100 of 4,095 results

Organizations	Publications	H-Index
Microsoft	741	29
Carnegie Mellon University	568	23
IBM	796	20
The French National Institute for Research in Computer Science and Control	448	19
Massachusetts Institute of Technology	265	19
University College London	208	19
University of Washington	192	19
Georgia Institute of Technology	270	18
University of Maryland	256	18
University of California Irvine	223	18



Agenda

1. Software engineering in the large
2. Software in the browser
3. Software social experience
4. Social aspects in the large
5. Software development kits



2. SOFTWARE ENGINEERING IN THE LARGE



The Challenge

Microsoft ships software to **1 billion** users around the world

How do we find out when things go wrong?

We want to

fix bugs **regardless of source**

prioritize bugs that **affect most users**

generalize the solution for any programmer

get the **solutions out to users** efficiently

try to **prevent bugs** in the first place



Debugging in the Large with WER...



Minidump



!analyze



5

17

23,450,649



Do you want to send more information about the problem?

Additional details about what went wrong can help Microsoft create a solution.



Show Details

Send information

Cancel



!analyze

Engine for WER bucketing heuristics

Extension to the *Debugging Tools for Windows*

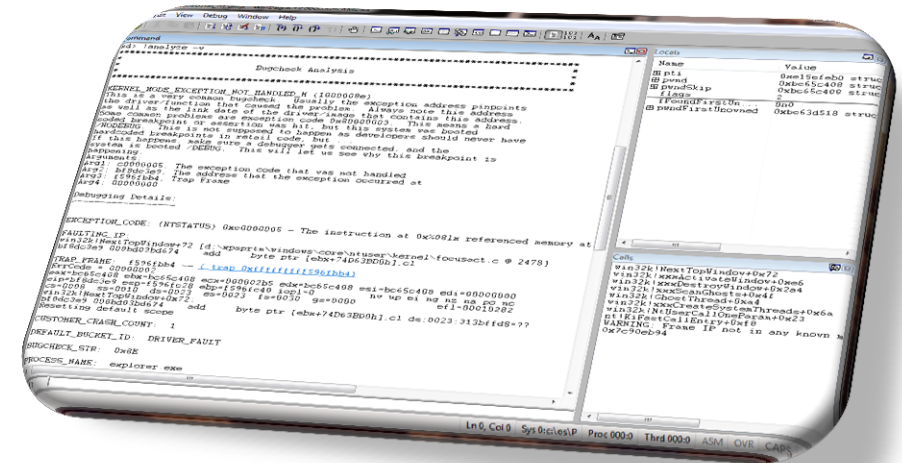
input is a minidump, output is bucket ID

runs on WER servers (and programmers desktops)

500 heuristics

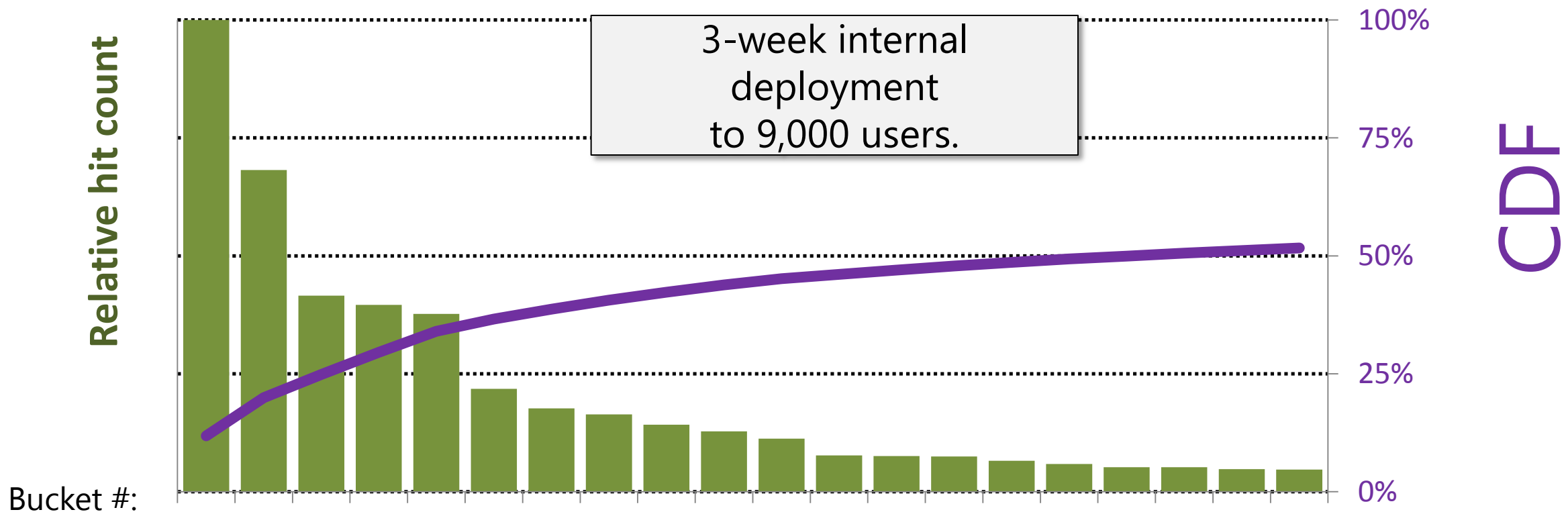
grows ~ 1 heuristic/week

Kirk Glerum, Kinshuman Kinshumann, Steve Greenberg, Gabriel Aul, Vince Orgovan, Greg Nichols, David Grant, Gretchen Loihle, and Galen Hunt, Debugging in the (Very) Large: Ten Years of Implementation and Experience, in *Proceedings of the 22nd ACM Symposium on Operating Systems Principles (SOSP '09)*, Association for Computing Machinery, Inc., Big Sky, MT, October 2009





Top 20 Buckets for *MS Word 2010*



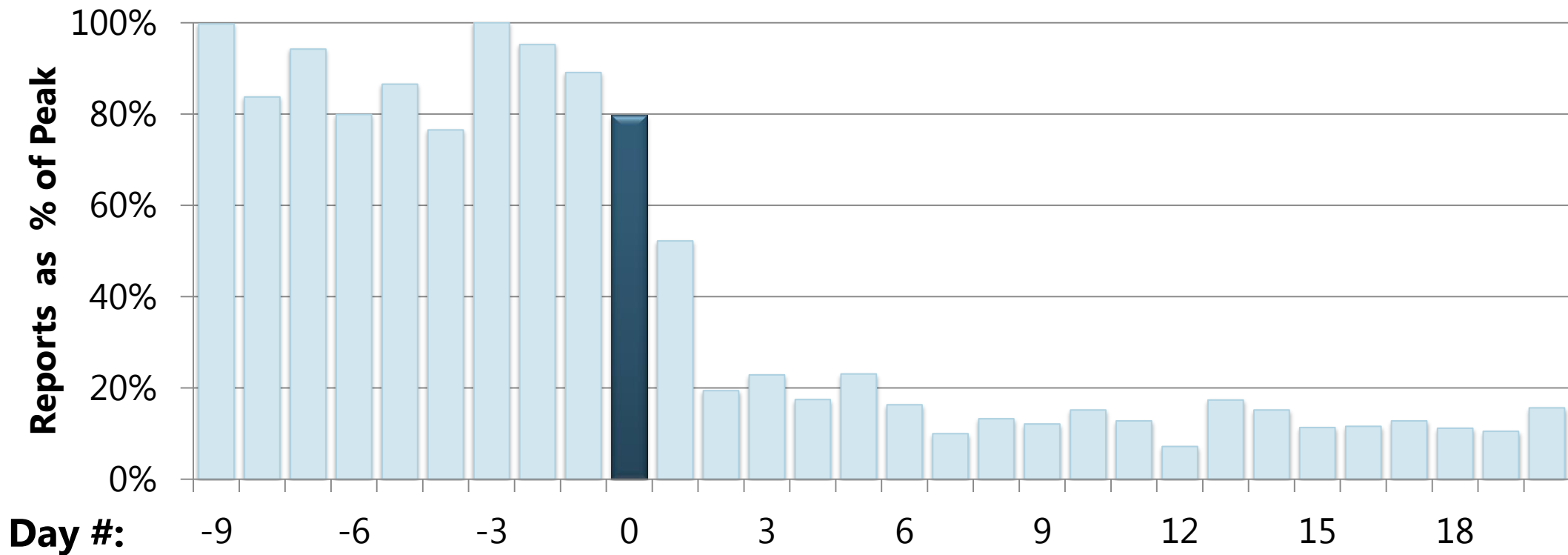
Just 20 buckets account for **50% of all errors**
Fixing a small # of bugs will help **many users**



Fixing bugs in software

- First use found ≥ 5 -year old heisenbugs in Windows
- Windows Vista team fixed 5,000 bugs in beta
- Anti-Virus vendor fixed top 20 buckets and dropped from 7.6% to 3.6% of all kernel crashes
- Office 2010 team fixed 22% of reports in 3 weeks
- And you can fix yours...

Hardware: Processor Bug



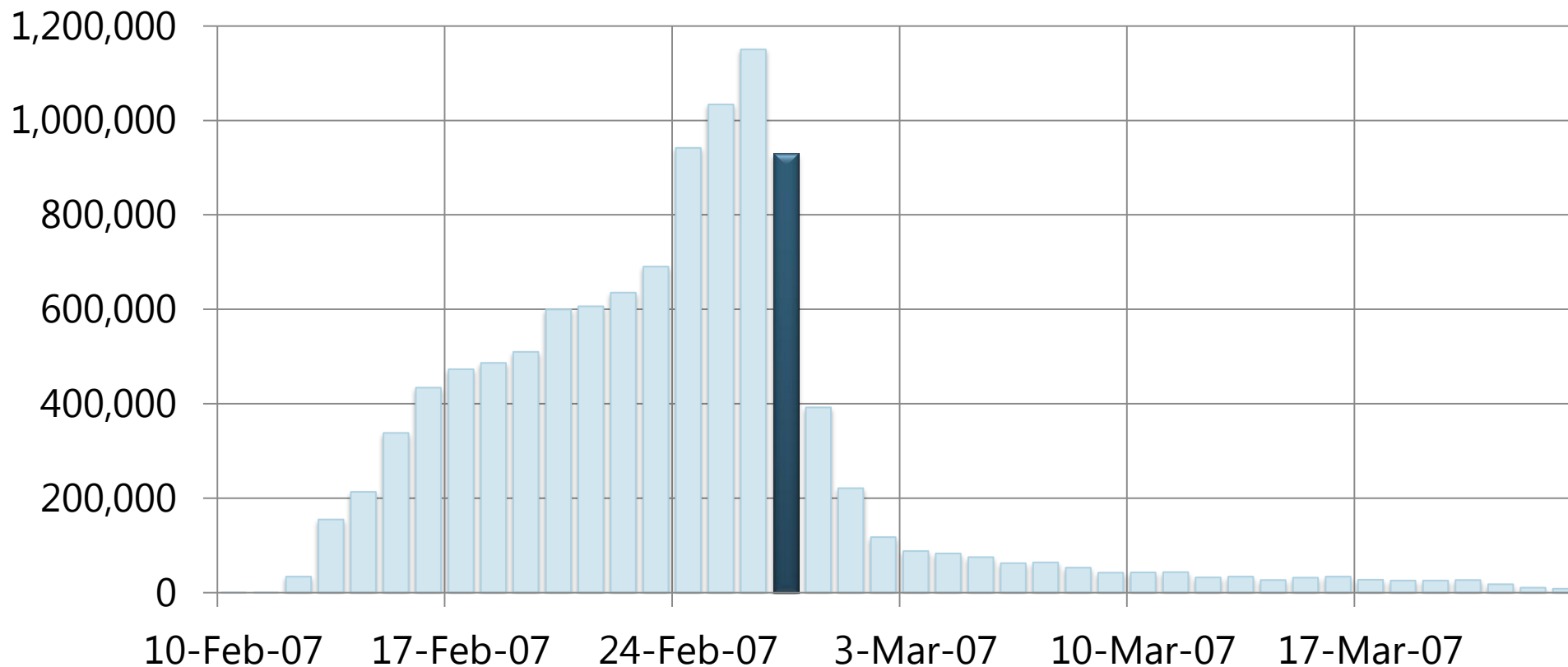
WER helped fix hardware error

Manufacturer **could** have caught this earlier with WER



Renos Malware

Reports per Day



Early detection w/o user action (renos, blaster, slammer, etc.)

WER **scales** to handle global events



3. SOFTWARE IN THE BROWSER



Types of software

1. Only a browser, e.g., Explorer, Firefox, Safari

2. A platform and language(s), e.g., a CLI implementation and C# or the JVM

3. An integrated development environment (IDE), e.g., Visual Studio or Eclipse



Browser based software

Sandbox Approach

- Download a Silverlight/ Moonlight control with a complete compiler
- All interaction directed to the control from the browser
- Computation on client
- Con: Effort to create the system
- Pro: No additional hardware needed

Server Approach

- Maintain a server (or cloud) presence
- All interaction directed to the server from the browser
- Computation on server
- Con: Scalability issue
- Pro: Can gather data on usability



Building Rich User Interfaces

```
Canvas.SetTop(label, 20.0)
Canvas.SetLeft(label, 20.0)

let updateLabel() =
    label.Text <- string System.DateTime.Now
    updateLabel()

let timer = new DispatcherTimer()
timer.Interval <- new System.TimeSpan
(0,0,0,1)
timer.Tick.Add(fun args -> updateLabel())
timer.Start()

App.Dispatch (fun() ->
    App.Console.ClearCanvas()
    App.Console.CanvasPosition <-
    CanvasPosition.Right
    clock2()
)
```

[load & run ▶](#)

A fancier clock

Since Try F# is built in Silverlight, the full power of the Silverlight platform is available. As an example, consider the clock sample from the Silverlight SDK after it has been ported to F#.

[load ▶](#) [run ▶](#)

Don't forget to click the Run button after the file has finished loading.

```
// Silverlight Clock sample implemented and running in the browser
// http://code.msdn.microsoft.com/silverlightsdk
//
open System
open System.Windows
open System.Windows.Controls
open System.Windows.Markup
open System.Windows.Media
open System.Windows.Media.Animation
open System.Windows.Shapes
open Microsoft.TryFSharp

let buildClockUI (canvas:Canvas) =
    let xaml = "
```

[run](#)

```
> //
// Silverlight Clock sample implemented and running in the browser
// http://code.msdn.microsoft.com/silverlightsdk
//
open System
open System.Windows
open System.Windows.Controls
open System.Windows.Markup
open System.Windows.Media
open System.Windows.Media.Animation
[... and 115 more line(s)]

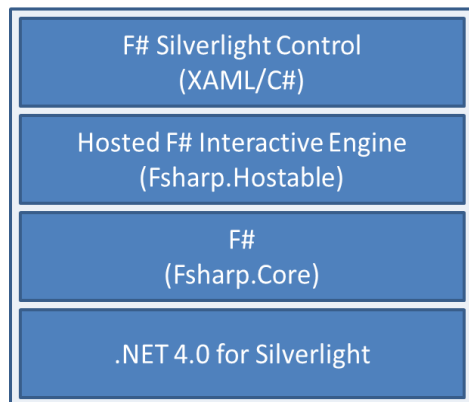
val buildClockUI : System.Windows.Controls.Canvas -> unit
>
```



Don Syme, Dean Guo, Christophe Poulain, Joe Pamer, Laurent le Brun, Nigel Horspool, Judith Bishop, 2010

Inside the site

Try F# Control Built in Silverlight



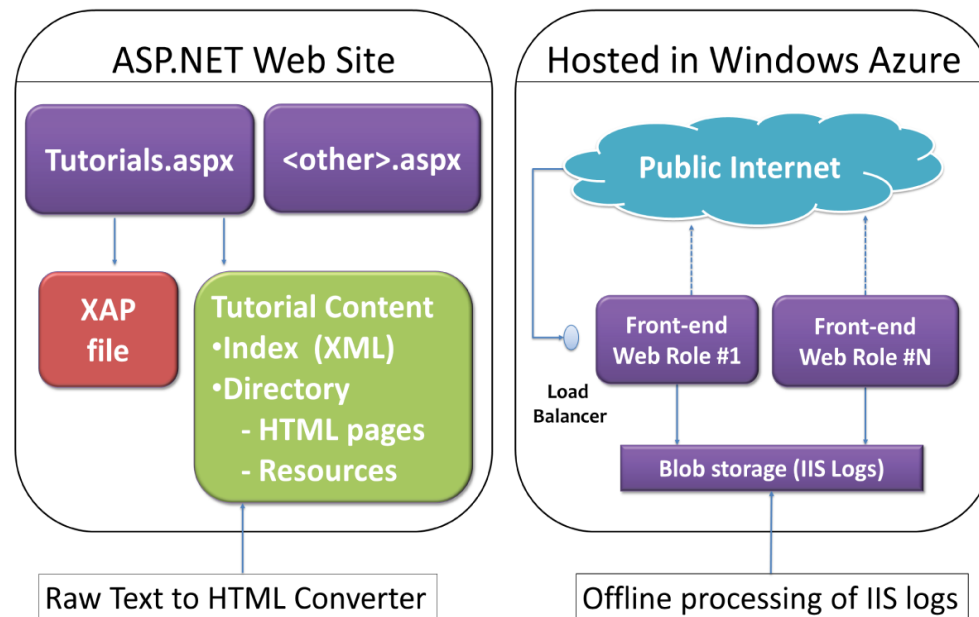
Compiled into XAP file

- Manifest
- DLL for our control
- FSharp.Hostable.dll
- FSharp.Core.dll
- .NET Silverlight DLLs

Control is scriptable:

```
string Script { get; set; }
string Output { get; set; }
void ClearOutput()
int LinesToEcho { get; set; }
Canvas Canvas { get; set; }
void ClearCanvas()
CanvasPosition CanvasPosition { get; set; }
void LoadFromString(string script)
void LoadFromUrl(string url)
void LoadFromFile()
bool IsDirty { get; set; }
void Save()
void Execute(string script)
void Cancel()
void Reset()
```

Try F# Web Site Deployed in Windows Azure



Demo of F#



- **Pex** - Visual Studio 2010 Power Tool developed by Microsoft Research to help unit testing of .NET applications.
 - Can be launched from the command line and run as Type 2 or Type 3 software.
- **Pex4Fun**
 - radically simplified version of the fully featured Pex
 - accessed via a browser
 - all the work happens on one of Microsoft Research servers
 - creates a game out of unit testing by providing existing or user entered code puzzles in C#, Visual Basic, or F#
 - Users determine from the unit tests what code needs to be added or changed.



Random Puzzle Learn New

205,285 clicked 'Ask Pex!'

C# Visual Basic F#

The code is a puzzle. Do you understand what the code does? Click Ask Pex! to find out.

```
using System;

public class Program {
    // What values of v can cause an exception? Ask Pex to find out!
    public static void Puzzle(int[] v) {
        if (v != null &&
            v.Length > 0 &&
            v[3] == 12345)
            throw new Exception("hidden bug!");
    }
}
```

Ask Pex!

Done. 5 interesting inputs found. [How does Pex work?](#)

Permalink

	v	Output/Exception	Error Message
✓	null		
✓	{}		
✗	{0}	IndexOutOfRangeException	Index was outside the bounds of the array.
✓	{0, 0, 0, 0, 0}		
✗	{0, 0, 0, 12345, 0}	Exception	hidden bug!

Dynamic Symbolic Execution in Pex

➤ Generates test data systematically

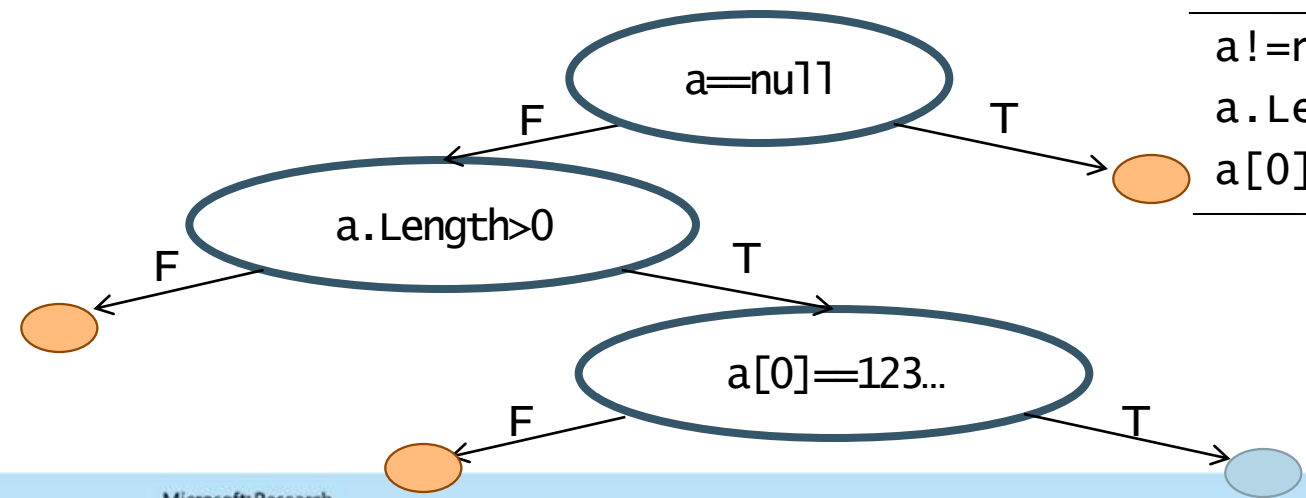
```

void coverMe(int[] a)
{
    if (a == null) return;
    if (a.Length > 0)
        if (a[0] == 1234567890)
            throw new Exception("bug");
}
    
```



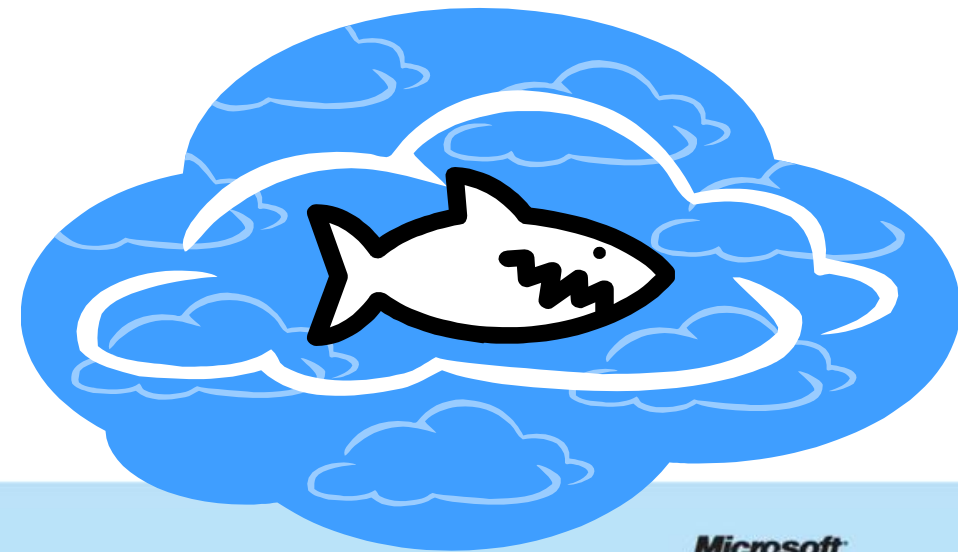
Constraints to solve	Input	Observed constraints
	null	a==null
a!=null	{}	a!=null && !(a.Length>0)
a!=null && a.Length>0	{0}	a==null && a.Length>0 && a[0]!=1234567890
a!=null && a.Length>0 && a[0]==123456890	{123..}	a==null && a.Length>0 && a[0]==1234567890

Done: There is no path left.



Dynamic Symbolic Execution Summary

- “Ask Pex” sends code to cloud
- Code is compiled and analyzed in cloud
- Dynamic Symbolic Execution automatically finds relevant interesting test inputs that achieve high code coverage
- Results are shown in browser




Auto-Completion

Pex for fun - from Microsoft Research - Windows Internet Explorer

http://pex4fun.com/

My Duels | Settings | Sign In

Curious? Learn More!

Pex  *for fun*

Random Puzzle Learn New

205,301 clicked 'Ask Pex!'

C# Visual Basic F#

The code is a puzzle. Do you understand what the code does? Click **Ask Pex!** to find out.

```
using System;

public class Program {
    public static void Puzzle() {
        Console.W
```

<< 12 / 19 >>

- WindowHeight
- WindowLeft
- WindowTop
- WindowWidth
- Write (18)
- WriteLine (19)**

void **WriteLine**(
string value)

Writes the specified string value, followed by the current line terminator, to the standard output stream.

value: The value to write.

throws System.IO.IOException

Research in Software Engineering (RiSE)

RiSE coordinates Microsoft's Research in Software Engineering in Redmond, USA. Our mission is to advance the state of the art in Software Engineering and Programming Languages, and to bring those advances to Microsoft's business.



The screenshot shows a browser window with the URL <http://rise4fun.com>. The main heading is "RiSE4fun" with the subtext "Gave 348 answers!". Below this are several tool categories: Agl, Bek, Boogie, Code Contracts, Concurrent Revisions, Dafny, Esm, Fine, Pex, Rex, Spec#, Vcc, and Z3. A "Click on a tool to load the next sample." instruction is present. A code editor displays a snippet of Dafny code for bit vector multiplication. Below the code is an "Ask Z3!" button and a question: "Is this formula satisfiable? Ask Z3!". Social media sharing options for "Tweet" and "75 people" are also visible.

<http://rise4fun.com>, our tools in your browser.



[see all...](#)

Using video clips on Channel9



4. SOCIAL EXPERIENCE



Try F# on Facebook

facebook 3 2 Search

tryF#

Computers · [Edit Info](#)

Wall TryFS

Share: [Status](#) [Photo](#) [Link](#) [Video](#) [Questi](#)

Write something...

TryFSharp created an event.

F# in Education Workshop, Nov 5, 2010
Wednesday, December 29, 2010 at 9:00am
Cambridge, Massachusetts

1,430 Impressions · 0.28% Feedback

[Like](#) · [Comment](#) · [Share](#) · December 29, 2010 at 2:08p

[Like](#) 3 people like this.



Coding Duels

Pex for fun - from Microsoft Research - Windows Internet Explorer

http://pex4fun.com/default.aspx?language=CSharp&sample=ChallengeArithmetic1

Favorites Pex for fun - from Microsoft Research

Pex



Coding Duel
for fun

[My Duels](#) | [Settings](#) | [Sign In](#)

[Random Puzzle](#) [Learn](#) [New](#)

205,316 clicked 'Ask Pex!'

[C#](#) [Visual Basic](#) [F#](#)

This puzzle is an interactive Coding Duel. Can you write code that matches a secret implementation? Other people have already won this Duel 1184 times! [Help](#)

```
using System;

public class Program {
    // Can you fill the puzzle method to match the secret arithmetic operation?
    public static int Puzzle(int x) {
        return 0;
    }
}
```

Ask Pex! Done. 2 interesting inputs found. [How does Pex work?](#) [Permalink](#)

Pex found 1 difference between your puzzle method and the secret implementation. Improve your code, so that it matches the other implementation, and 'Ask Pex!' again. You are not signed in. Sign In to track your achievements. [Help](#)

	x	your result	secret implementation result	Output/Exception	Error Message
✓	0	0	0		
✗	1	0	-1	Mismatch	Your puzzle method produced the wrong result.

Coding Duels - Fun and Engaging


- Iterative gameplay
- Adaptive
- Personalized
- No cheating
- Clear winning criterion




Social Experience


Pex for fun - from Microsoft Research - Livefeed - Windows Internet Explorer


http://pex4fun.com/Livefeed.aspx


Pex  for fun


205,324 clicked 'Ask Pex!'


 User79033 asked Pex about a puzzle
11 seconds ago


 User79033 tried to win C# - «ChallengeArithmetic1»
14 seconds ago


 User79033 asked Pex about a puzzle
16 seconds ago

 User79032 asked Pex about a puzzle
20 seconds ago

 User79029 tried to win C# - «ChallengeDigits2»
32 seconds ago

 User79029 tried to win C# - «ChallengeDigits2»
59 seconds ago

 User79031 tried to win C# - «ChallengeArithmetic1»
2 minutes ago

 User79018 made 17th attempt to win C# - «ChallengeWordReverse»



Social Experience Summary

- Community
- High score lists, leaderboard
- Live feed



<http://pex4fun.com/Community.aspx>

<http://pex4fun.com/Livefeed.aspx>

Demo of Pex4Fun



5. SOCIAL ASPECTS IN THE LARGE



MS Products' (Bing) History is Huge

- Over one million change-sets.
- One change-set can contain 200K check-ins.
 - e.g., branch creations
- +700 branches, 6M different files, 150M different depot paths, 450M different revisions, 150M check-ins, 130M integrations
- Almost impossible to crawl incrementally
 - At least not in 3 months!



Speculative Analysis

- *Predict* what the user might do
 - *Do* that action in advance, store the results
 - If the user *chooses* to do it,
 - *Present* the pre-calculated information
-
- Aim: increase developer awareness with precise information

Development History (Recall)



child1

10th

12th

15th

16th

child2

11th

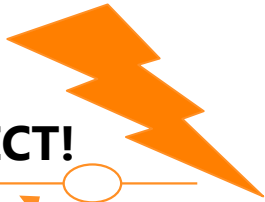
14th

18th

23rd

24th

CONFLICT!



Development History (with BEACON)



child1

10th

12th

15th

16th



CONFLICT!

child2

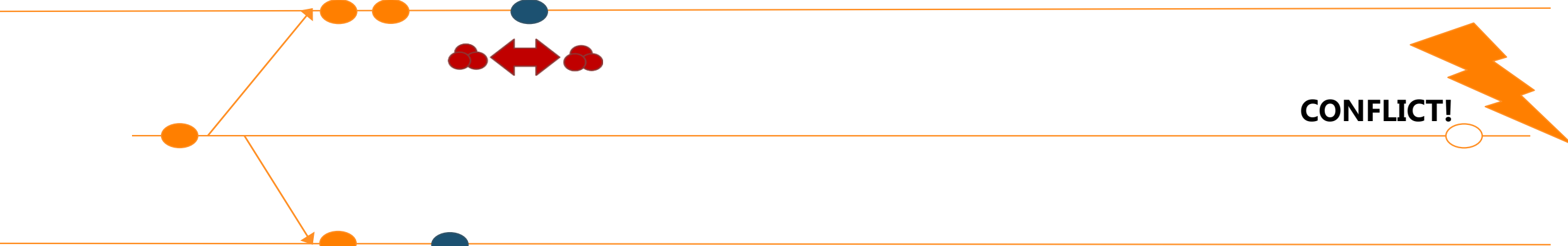
11th

14th

18th

23rd

24th



BEACON Example

Beacon Configuration - Microsoft Visual Studio - Experimental Instance

File Edit View Debug Team Data Architecture Test Tools Analyze Window Help

Start Page Beacon Configuration

Browse Enlistment **Configuration File:**

Load Enlistment **Development Branch:**

Beacon client name: beacon_bing

Main Branch: **UX_A**

Suggested Branches

<input checked="" type="checkbox"/> ux	<input type="checkbox"/> UX_01	<input type="checkbox"/> UX_02	<input type="checkbox"/> UX_03	<input type="checkbox"/> UX_04
<input type="checkbox"/> UX_07	<input type="checkbox"/> UX_08	<input type="checkbox"/> UX_09	<input type="checkbox"/> UX_11	<input type="checkbox"/> ux_12
<input type="checkbox"/> ux_13	<input type="checkbox"/> ux_14	<input type="checkbox"/> ux_15	<input type="checkbox"/> ux_16	<input type="checkbox"/> ux_17
<input type="checkbox"/> ux_18	<input type="checkbox"/> ux_19	<input type="checkbox"/> ux_20	<input type="checkbox"/> UX_Asia	<input checked="" type="checkbox"/> UX_B
<input type="checkbox"/> UX_C	<input type="checkbox"/> UX_Developer	<input type="checkbox"/> ux_domains	<input type="checkbox"/> UX_Html5	<input type="checkbox"/> ux_live

Other Branches

<input type="checkbox"/> //depot/dev/indexserve_relevance_tools_dev_bellevue_temp	<input type="checkbox"/> aar_index_ifm
<input type="checkbox"/> answers_01	<input type="checkbox"/> answers_alterations
<input type="checkbox"/> answers_commerce	<input type="checkbox"/> answers_commerce_01
<input type="checkbox"/> answers_domains_apps	<input type="checkbox"/> answers_domains_live
<input type="checkbox"/> answers_local	<input type="checkbox"/> answers_local_01

Update Configuration

Beacon Relation Details

Relationship details for branch: **ux**

Relationship = **CONFLICT**

Related files:

- 1- //depot/dev/UX/build.trigger
Edit by REDMOND\eliak on 8/26/2011 9:20:09 PM
Commit message: Trigger a new build after the ux_vnext FI which fixed a previous brea
- 2- //depot/dev/UX/private/frontend/deploy/FEX/fex.ini
Edit by REDMOND\v-brcard on 8/31/2011 1:14:54 PM
Commit message: Auto Integrator : Ux_vnext=> ux
- 3- //depot/dev/UX/private/frontend/deploy/FEX/ResultParts.ini
Edit by REDMOND\v-brcard on 8/31/2011 1:14:54 PM
Commit message: Auto Integrator : Ux_vnext=> ux

Beacon Status

Current branch: **UX_A**

ux

UX_B



6. SOFTWARE DEVELOPMENT KITS



Software Development Kits

- Used for access to
 - Proprietary hardware and their drivers
 - Large proprietary data

Research.microsoft.com/cs

Project Hawaii

- On WP7
- Services execute in the cloud (Azure)
- OCR, Speech to text etc
- **WP7 phones loaned to universities worldwide**
- C#

Kinect SDK

- Drivers and rich APIs for raw sensor streams and human motion tracking
- **Kinect unit is \$150**
- Available to academics
- C++, C#, VB

Web-NGram

- Content and model types
- N-gram availability to 5
- Training size: **All documents indexed by Bing in the en-us market**
- Updated Periodically

Arjmand Samuel, Stewart Tansley, Evelyne Viegas

Web N-Gram in Public Beta

Single Tag Cloud

Multi Tag Cloud



Ref: Dr. Li Ding, Rensselaer Polytechnic Institute

Concept Search

The screenshot displays a web interface for a concept search tool. At the top, a search bar contains the word "company". To its right are a "Concept" dropdown menu and a "Search" button. Below the search bar, a legend identifies three types of concepts: "Super Concept" (blue), "Sub Concept" (green), and "Similar Concept" (orange).

The central concept map shows "company" as a central node. It is surrounded by several "Sub Concepts" (green boxes) including Oracle, Google, HP, Apple, Dell, Nokia, Yahoo, IBM, Wal-Mart, Sony, Boeing, Intel, Cisco, General Elec..., Samsung, Microsoft, Hewlett-Packard, Motorola, and Nike. To the right, "organization" is a "Super Concept" (blue box) with sub-concepts like "information", "stakeholder", "body", "legal entity", "detail", "term", "institution", "factor", "party", "group", "Personal Inf...", "item", "field", "data", "noun", "enterprise", "corporate body", "personal inf...", and "entity".

At the bottom, "Similar Concepts" (orange boxes) include "world-famous...", "america's le...", "cell phone m...", "large US com...", "PC company", "communicatio...", "large softwa...", "import vehicle", "arketer", "technology i...", "massive company", "global auto ...", "internat...", "technology g.", "big-name vendor", and "online gi...".

On the left side, under "Matched Results:", a list of related terms is provided, such as "company", "large company", "leading company", "big company", "international company", "multinational company", "insurance company", "well-known company", "oil company", "local company", "global company", "private company", "technology company", "indian company", "U.S. company", "american company", "software company", "pharmaceutical company", "top company", "japanese company", "media company", "internet company", "cable company", and "successful company".

On the right side, under "Attributes:", a list of attributes is shown, including "foundation", "homepage", "industry", "company name", "location", "company logo", "product", "revenue", "name", "net income", "founder", "type", "logo", "parent", "asset", and "owner".

Table understanding

Query: *films budget*

Web Images Videos Shopping News Maps More | MSN Hotmail



films budget



Web

Web

Table

Videos

Ehow

More ▼

— Shrink

— Shrink table

Year	Movie	Worldwide gross	Budget	Distributor	Director
1977	Star Wars	\$ 782400000	\$ 11000000	20th Century Fox	George Lucas
1997	Titanic	\$ 1848813795	\$ 200000000	Paramount Pictures	James Cameron
1993	Jurassic Park	\$ 914691118	\$ 95000000	Universal Studios	Steven Spielberg
1995	Toy Story	\$ 365000000	\$ 90000000	Walt Disney Pictures	John Lasseter
1972	The Godfather	\$ 245066411	\$ 6000000	Paramount Pictures	Francis Ford Coppola
2009	Avatar	\$ 2606954237	\$ 237000000	20th Century Fox	James Cameron
1975	Jaws	\$ 470600000	\$ 7000000	Universal Studios	Steven Spielberg
1996	Independence Day	\$ 816969268	\$ 75000000	20th Century Fox	Roland Emmerich
1998	Armageddon	\$ 553709788	\$ 140000000	Touchstone Pictures	Michael Bay

SEIF Awards



Software Engineering Innovation Foundation (SEIF) with MSR's-RiSE group supports academic research in software engineering

- Technologies, tools and practices
- Teaching methods

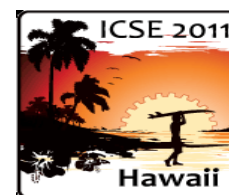
Awards of US\$ 25,000 for

- seed-funding larger initiatives,
- proofs of concept, or
- demonstrations of feasibility

In association with the RiSE and Mobile groups at Microsoft Research

research.microsoft.com/seif

- **LATAM Awards**
 - 2010 – 3 out of 12
 - 2011 – 1 out of 10
 - In open global competition
- **Connected to major conferences a workshop and an event**



SEIF DAY, July 18, 2012



Conclusion - best practices

- Software engineering is analytic and deals with huge data
- Technology transfer to academia by moving to
 - Browsers
 - SDKs
 - New devices
- Encouraging community through social media