QCon

Please evaluate my talk via the mobile app!

From Instability to Resilience: The Story of a Web Site

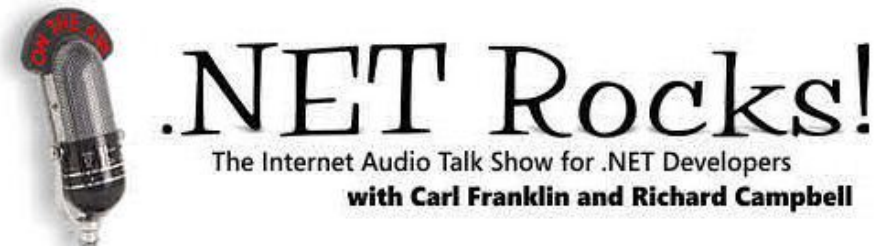
Richard Campbell

@richcampbell

Richard Campbell

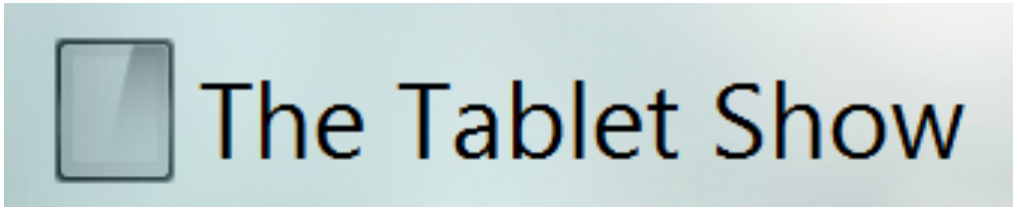
- Background
  - First laid hands on a microcomputer in 1977, it's been all downhill from there
  - Spent the last fifteen years helping companies scale software on a variety of platforms
- Currently
  - Architect and Consultant
  - Organizer of DevIntersection
  - Rabid Podcaster

Podcasts



For .NET Developers
First published 2002
Two shows a week
955 episodes in the archive

For IT Pros
First published 2007
Once a week
357 episodes in the archive



For Tablet Developers
First Published 2011
Once a week
126 episodes so far

What is Resilience?

Resilience Definition

re·sil·ience [ri-**zil**-y*uh* ns, -**zil**-ee-*uh* ns] Show IPA

noun

1. the power or ability to return to the original form, position, etc., after being bent, compressed, or stretched; elasticity.
2. ability to recover readily from illness, depression, adversity, or the like; buoyancy.

Also, **re·sil·ien·cy.**

Origin:

1620–30; < Latin *resili* (*ēns*), present participle of *resilīre* to spring back, rebound (see resilient) + -ence

The Web Site Story

- Vertical market e-commerce system
- More than 200,000 discrete items
- Considered mature (version 3!)
- Busier than ever, but making less money
- Increasing tech support calls and complaints

Diagnostic Instrumentation

- Performance Monitor
  - CPU utilization
  - Requests/Sec
  - Requests Queued
  - .NET Heap Allocated

Initial Diagnosis

- CPU utilization high, not pinned
- 20-30 Requests/sec, steady
- Request Queued jumps occasionally
- .NET Heap Climbs to 800MB, then dumps
  - Memory leak?
- Worker process recycling every 20 minutes

Taking Action

- Speed of response vs. comprehensive response
- Must have measurements before and after
- Addressing a memory leak with more memory!

First Actions & Results

- Switch to 64 Bit OS (but compile to 32 bit)
- Added 4GB of Memory (for a total of 8GB)
- Memory leak continued
  - Just took longer
  - Dump occurred a 2GB (maximum pool size)
  - Worker process recycling every two hours
- Complaint level drops dramatically

Character of the Web Site

- Accuracy
- Reliability
- Scalability
- Performance

- Where does Resilience fit?

Better Living through Hardware

- Adding more web servers
  - Spreading failure around
- Moving to the Cloud
  - Paying for failure by the hour

Increasing Understanding

- What is causing this memory leak?
  - Using .NET Memory Profiling to analyze
- Most memory consumed by cache

<foreshadow>

- Noticed large blocks of static memory (session objects)

</foreshadow>

When Cache Runs Amok

- Caching added to improve performance
- Cache objects being created around searches
- What are the chances of that cache object ever being used again?
- How do you expire a cache object?

Instrumenting Cache

- Wrapped cache objects in a dictionary abstraction
- Kept counters for re-use

- Two hours of run time (one worker process recycle) results in 95% of cache objects never reused

Fixing Cache

- Don’t allow free form search cache
- Don’t let the users populate cache at all
- Treat the real problem

Valuing Resiliency

- Nothing like failure to help determine value
- Owning your ROI
- Spending CapEx to reduce OpEx

Adding More Resiliency

- Bend further
  - Multiple servers provide redundancy
- Elasticity
  - Can we expand or shrink based on need?
- Better feedback
  - How do we know when we're bent, when to expand and when to shrink?

Clusters of Joy

- Successful failover is hardware, software & configuration together
- If it hasn't been tested, it doesn't work
- Routine failover is your friend!

The Plague of State

- Getting session out of the web server
- The battle of performance over scalability
- As little state as possible (and no less)

Instrumenting Production

- The only source of truth
- How can you build software that can test in production without impacting the user?
- Instrumentation driving features

Building Resiliency

- Resiliency is a journey, not a destination
- Know the character of your system
- Work from facts, not conjecture to move the system forward

QCon

Please evaluate my talk via the mobile app!