

P2P, DSM, and Other Products from the Complexity Factory

Willy Zwaenepoel
EPFL

Impact of Research



Impact of Research

- Not so great

- Many research ideas lost out
- Many non-research developments won out

Impact of Research

● Not so great

- Many research ideas lost out
- Many non-research developments won out

● Why?

- We make things too complex
- Not: things are too complex

Impact of Research

● Not so great

- Many research ideas lost out
- Many non-research developments won out

● Why?

- We make things too complex
- Not: things are too complex

● Why?

- Publishing/reviewing pushes us to complexity

Apologies, Caveats and Excuses

- Talk is rather polemic in nature
... things are said a little crassly

P2P (Peer-to-Peer)

- No (central) server
- Easier to operate, maintain, scale, make more reliable ...
- Started as an application (file sharing)
- Proposed as infrastructure for applications

Research on P2P

- Concentrated on DHT (a beautiful concept)
- Chord, Pastry, ...
- Applications: backup, streaming, ...

The Problem with P2P

- Little application other than illegal file sharing

Reality Check

- If we have learned anything in the last 25 years of distributed computing:
- Distributed is harder than centralized
- Must have compelling reason for it

Reasons for Distribution

- You cannot handle it in one place
 - Performance – controlled replication
 - Availability – controlled replication

Reasons for Distribution

- You cannot handle it in one place
 - Performance – controlled replication
 - Availability – controlled replication
- Geographical distribution
 - Google!

Reasons for Distribution

- You cannot handle it in one place
 - Performance – controlled replication
 - Availability – controlled replication
- Geographical distribution
 - Google!
- Illegality – P2P
 - From Napster to Gnutella, Kazaa, ...
 - “Raw” traffic numbers high, but static content
 - Could be handled by conventional replication (?)

Difficulties for P2P as Infrastructure

- Hard to find anything
- Hard to make anything secure
- Hard to write anything

Advantages for P2P Infrastructure Research

- Complex to find anything
- Complex to make anything secure
- Complex to write anything

Advantages for P2P Infrastructure Research

- Complex to find anything
- Complex to make anything secure
- Complex to write anything

- Complexity begets papers
- P2P = Paper-to-Paper

There are Applications

- Large file multicast
- Can be handled by very simple techniques
 - BitTorrent
- It should worry us that these come from non-research corners of the world!

A 3D grid of spheres on a blue background. The spheres are arranged in a regular, repeating pattern that recedes into the distance, creating a perspective effect. The spheres are light blue and connected by thin, light blue lines. The background is a solid, dark blue color.

Mirror, mirror on the wall, ...

DSM (Distributed Shared Memory)

- Parallel computing on clusters
- Shared memory programs easier to write
- Single shared address space
- Portions cached in physical memory
- Usually done by page faulting
- TreadMarks (ParallelTools)

Reality Check

- Cluster *hardware* only suitable for coarse-grained parallel computation
- A fortiori true for any cluster software
- A fortiori true for DSM

However, most DSM Research

- Is about fine-grain parallelization

Difficulties with Fine-Grained DSM

- Expensive synchronization
- Expensive fine-grained data sharing
 - Smaller than a page
 - True or false sharing

Advantages for Fine-Grain DSM Research

- Complex fine-grain synchronization
- Complex fine-grain data sharing
 - Compiler, language, runtime, ...

Advantages for Fine-Grain DSM Research

- Complex fine-grain synchronization
- Complex fine-grain data sharing
 - Compiler, language, runtime, ...
- Complexity begets papers ...

TreadMarks

● Academic experience

- (Almost) every paper or grant for research on fine-grain DSM was *accepted*
- (Almost) every paper or grant for research on coarse-grained DSM was *rejected*

TreadMarks

● Academic experience

- (Almost) every paper or grant for research on fine-grain DSM was accepted
- (Almost) every paper or grant for research on coarse-grained DSM was rejected

● Industrial experience

- Only coarse-grain applications
- Real applications: *a page is not large enough!*

Competition is Message Passing

- MPI (Message Passing Interface)
- Low abstraction
- No room for complexity fabrication
- As a result more successful

Competition is Message Passing

- MPI (Message Passing Interface)
- Low abstraction
- No room for complexity fabrication
- As a result more successful
- It should worry us that MPI did not come from distributed systems research

Painful Observations (1)

- Many research ideas lost out
- Non-research designs won out
- Has to do with this fabricated complexity

Painful Observations (2)

- Has to do with publishing/reviewing
 - Simple papers tend to get rejected
 - Complex papers tend to get in

Your Average Review Form

- Novelty
- Excitement
- Writing
- Confidence

Some Likely Review Comments

- « Incremental »
- « Engineering »
- « Nothing new »
- « Boring »

Some Questions to Add?

- Does the added functionality justify the increase in complexity?
- Does the performance improvement justify the increase in complexity?
- Could this system be maintained by an above-average programmer in industry?
- ...

It IS Possible: Virtual Machines

- Provide simple solutions to real problems
 - Server consolidation
 - Migration

Virtual Machines History

- IBM VM
- VMWare
- Xen
- KVM

Virtual Machines History - Simplification

- IBM VM – specialized hardware
- VMWare – x86 hardware
- Xen – paravirtualization
- KVM – part of Linux

Conclusion

- Brute force often (not always) works
- Our publishing and reviewing system pushes us in the opposite direction
- Hence, our work has little impact