

Does Software Engineering Research have Impact on Software Engineering Practice?

A Brief Introduction to the Impact Project

Alexander L. Wolf
Imperial College London

Assumptions

There is a software engineering practice

- engineers apply techniques, use tools, and follow processes

There is a body of work that can be identified as software engineering research

- both academic and industrial

There is skepticism about the impact of research on practice

- witness: the relatively low level of academic and industrial research funding vis-à-vis the increasingly critical role played by the practice

From where does technology come?

Facile answers are misleading

It comes from...

- Sun, Microsoft, IBM, Google, "the Web", ...

Yes, but!

- from where *did* they get it? and how?

It comes from...

- Dr. X, who published a seminal research paper and produced a popular prototype

Yes, but!

- we *didn't* read the paper
- it was only a prototype, not engineered/licensed for industrial use

It comes from...

- being "in the air", everybody knows it

Yes, but!

- how *did* it get there? who nurtured it?

Ideas have many parents...

Researchers

Scientific and technical communities

Technology transfer agents

Students with new degrees

New hires

Early adopters

Commercializers

... and their contributions differ

Initial conceptualization of *idea*

Evangelism

Prototype *demonstration*

Public promulgation

Nurturing by community activities

Education, training, and indoctrination

Product commercialization

Why should we care?

Facile answers are misleading

Some technology is not very good

- why are we stuck with it?
- why is it not better?

Some technology seems *useful*

- how can we get more of it?
- how can we speed its appearance?
- are there approaches that need to be strengthened/nurtured, despite their slow adoption?

How do we evaluate the contributions?

Facile answers are misleading

Ideally

- qualitatively
- quantitatively

But a challenging task

- different parties have different motivations
- apportioning contributions is difficult
- long timescales attenuate measurements and memories

CiP: Pressure to measure in the UK

UK Government is seeking accountability

- demonstrating economic impact of tax-payer investment in basic research and improving exploitation of research outputs

2006 DTI report on "Increasing the Economic Impact of the Research Councils"[†]

- provides several recommendations on how to consider economic impact in funding decisions
- example: an individual competent in the economic impact of research should be accommodated on each review panel

[†]<http://www.dti.gov.uk/files/file32802.pdf>

But measuring is easier said than done

2006 EPSRC report on "International Review of ICT Research in the UK"[†]

- panel noted the difficulty in conducting macro-economic analysis of ICT commercial impact

2007 Russell Group response to DTI report[‡]

- "There is no evidence to date of any rigorous way of measuring economic impact other than in the very broadest of terms and outputs."

[†]<http://epsrc.ac.uk/ResearchFunding/Programmes/ICT/ReviewsAndConsultations/InternationalReview/>

[‡]<http://www.russellgroup.ac.uk/news/2007/rcuk-consultation-on-the-efficiency-and-effectiveness-of-peer-review.html>

Goals of the Impact Project

Scholarly, objective, case-based evaluation

Desired deliverables

- peer-reviewed papers
- presentation materials and outreach activities
- expertise

Community building

Prospective for future research investment

Lessons learned for "successful" research

- but only with respect to transfer into practice (there are other measures of research success)

Administration

An initiative of ACM SIGSOFT

- volunteers mostly paying their own way
- modest funding from US NSF, UK IEE, and various agencies in Italy, UK, Germany, and Japan

International executive committee



Leon Osterweil



Carlo Ghezzi



Jeff Kramer



Alexander Wolf

Method

Form teams around practices

- important and widespread
 - » examples: configuration management, programming languages, middleware, assertions, walkthroughs, ...
- recruit volunteers, including researchers (academic and industrial) and practitioners

Start from practice and trace backward

Use accepted historical tools

- qualify conclusions by solidity of evidence; use references, oral histories, ...

Michael Mahoney
Professor of (Science) History



A quick tour of two completed reports

Software configuration management

- first report to emerge from the project
- ACM TOSEM, October 2005

Middleware technology

- latest report to emerge from the project
- ACM TOSEM, to appear in 2008

Software configuration management

Investigation into the research origins of successful SCM vendor products (ca. 2003)

- version control
- product models
- change control
- composition/selection
- build management
- workspace management

Lead authors



Jacky Estublier



David Leblang

Other team members

- G. Clemm, IBM (ClearCase)
- R. Conradi, U. Trondheim
- A. van der Hoek, U. California
- W. Tichy, U. Karlsruhe (RCS)
- D. Wiborg-Weber, Telelogic (Continuus)

Is there a practice?

Practice measured by success of vendor products

- Ovum: \$1B (1998), \$2B (2000), \$3.3B (2002)
 - » 25% mainframe
 - » 15%-20% workstations
 - » 5%-10% PC
- Gartner: \$6B (2003)

This ignores the "success" of freeware and shareware

- example: CVS

Historical method

1. Examine characteristics/features of leading products in SCM market

2. Assume products are used in practice

3. Trace product characteristics/features back to research ideas and prototypes

4. Make arguments for/against influence of research on practice via products

When was it introduced?

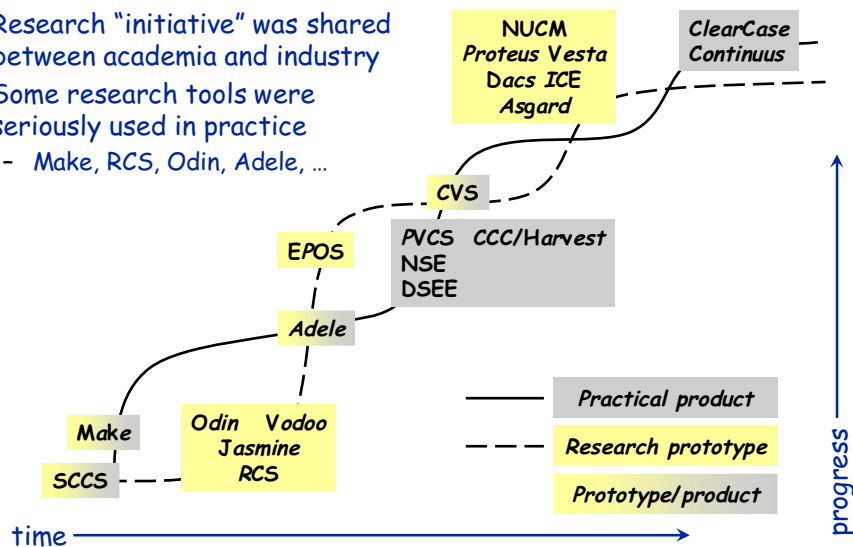
| | Academic Research | Industrial Research | Industrial Product |
|------|--|---|--------------------|
| 1972 | | SCCS (Bell Labs) | |
| 1976 | | Diff (Bell Labs) | |
| 1977 | | Make (Bell Labs) | |
| 1980 | Variants, RCS (Purdue) | Change-sets (Xerox Parc) | |
| 1982 | Merging, and/or graph (Purdue) | | |
| 1983 | | Change-sets (Aide-de-Camp) | |
| 1984 | Selection (Grenoble) | | |
| 1985 | | System model (DSEE) | |
| 1988 | Process support (Grenoble) | | |
| | NSE Workspaces (Carnegie Mellon, Sun) | | |
| 1990 | | nDFS file system (Bell Labs) | |
| 1994 | | Virtual file system and MultiSite (ClearCase) | |
| 1996 | | Activities (Asgard, Bellcore) | |
| 2000 | WebDAV (California, Microsoft, ClearCase, ...) | | |

An argument: research/product timing

Research "initiative" was shared between academia and industry

Some research tools were seriously used in practice

- Make, RCS, Odin, Adele, ...



An argument: professional interaction

SCM research community organized regular workshops beginning in 1988

Product architects were present at all

- Cagan, Clemm, Dart, Leblang, Wiborg-Weber, ...
- Some presented work, while others simply attended and participated in discussion

The meetings put ideas "in the air", and helped to keep them there

The role of creativity: vendor's view

Vendors tend to consider that research impact is restricted to...

algorithms (e.g., differencing)

pieces of reusable code (e.g., RCS)

and not...

concepts (e.g., hierarchical workspaces)

architectures (e.g., peer-to-peer repositories)

which are often seen as "**engineering common sense**"

The role of creativity: researcher's view

Researchers tend to consider that...

precedence

concepts

prototypes

are sufficient as impact and ignore...

efficiency

usability

reliability

dismissing them as "*engineering common sense*"

Both are right and both are wrong

A good idea is had more than once

Vendors have disincentives for distributing credit for ideas

Researchers have incentives for claiming credit for ideas

Research and productization both require *engineering creativity*

Middleware technology

Investigation into the research origins of successful middleware technology (ca. 2007)

- web services
- distributed object systems
- application servers
- message queues
- transaction monitors
- remote procedure call systems

Lead authors



Wolfgang Emmerich

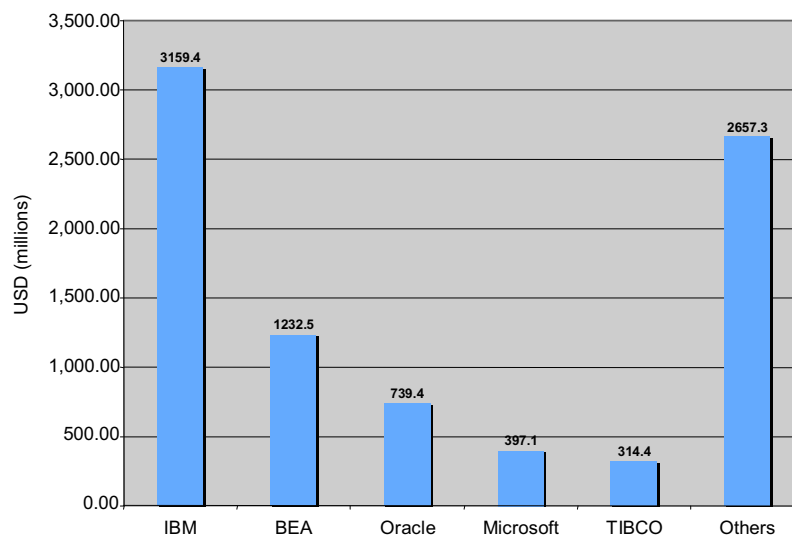


Mikio Aoyama



Joe Sventek

Is there a practice?

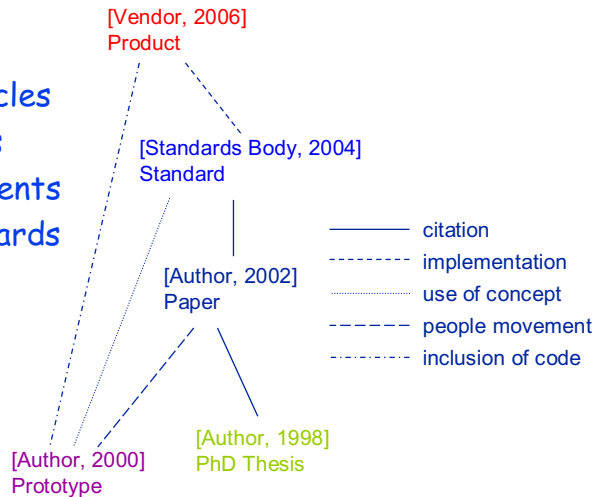


Middleware License Market in 2005 [Gartner 2006]

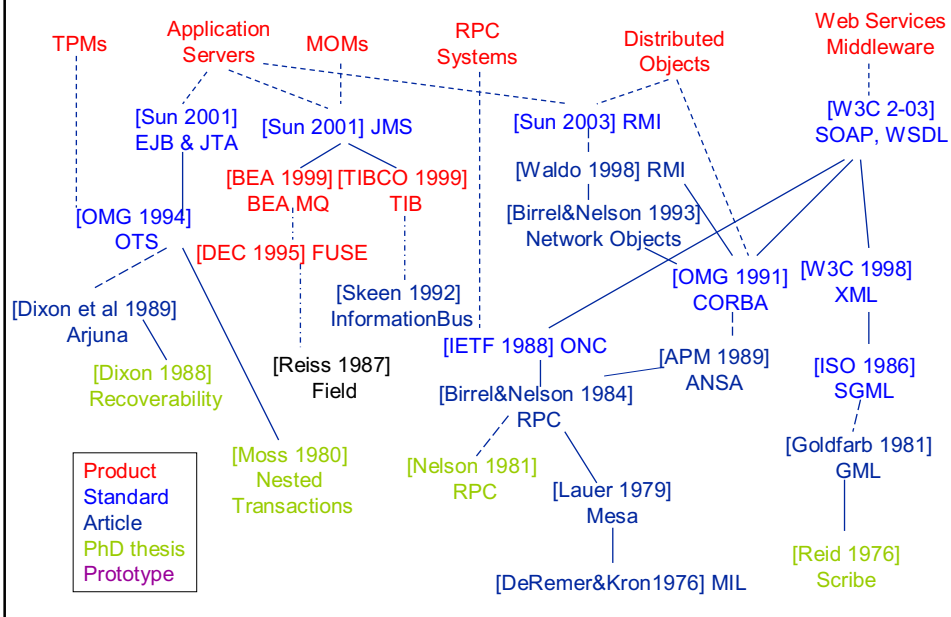
Historical method

1. Seek sources 2. Derive "impact trace graph"

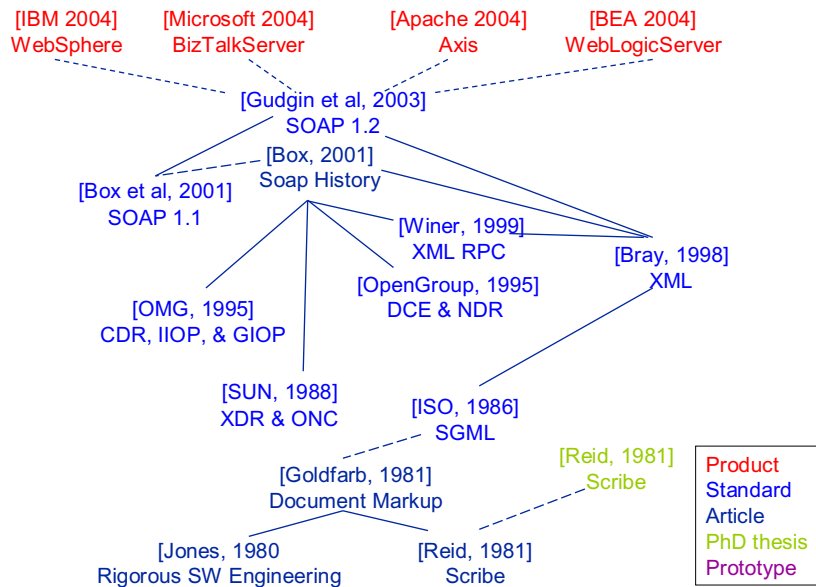
- market analysis reports
- professional articles
- technical reports
- standards documents
- minutes of standards meetings
- people movement
- PhD theses
- software
- interviews



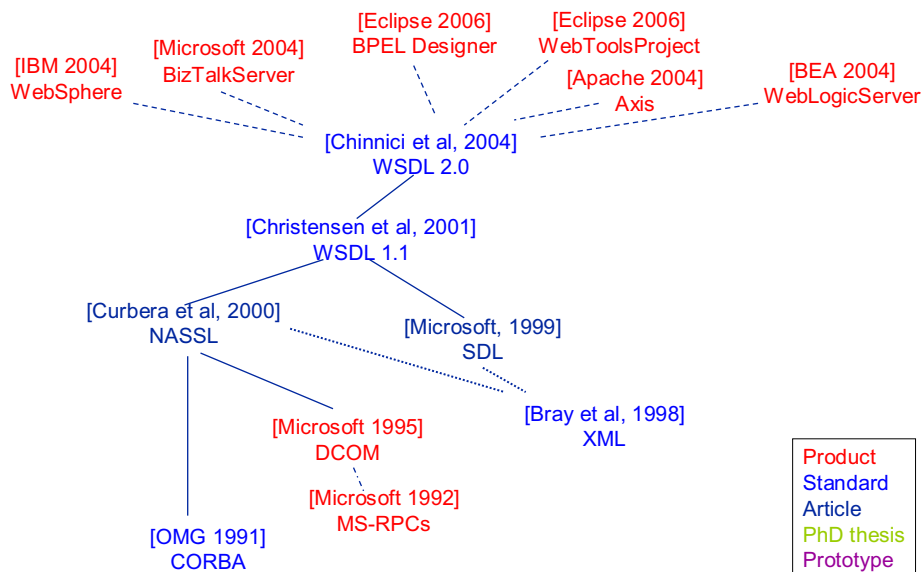
The big picture of impact



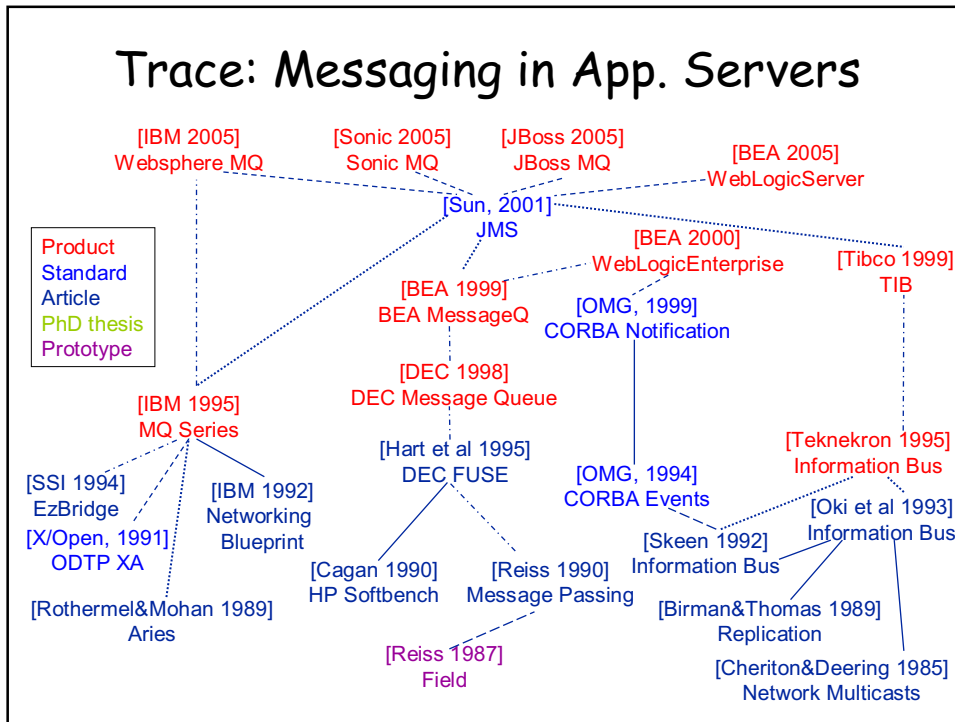
Trace: Simple Object Access Protocol



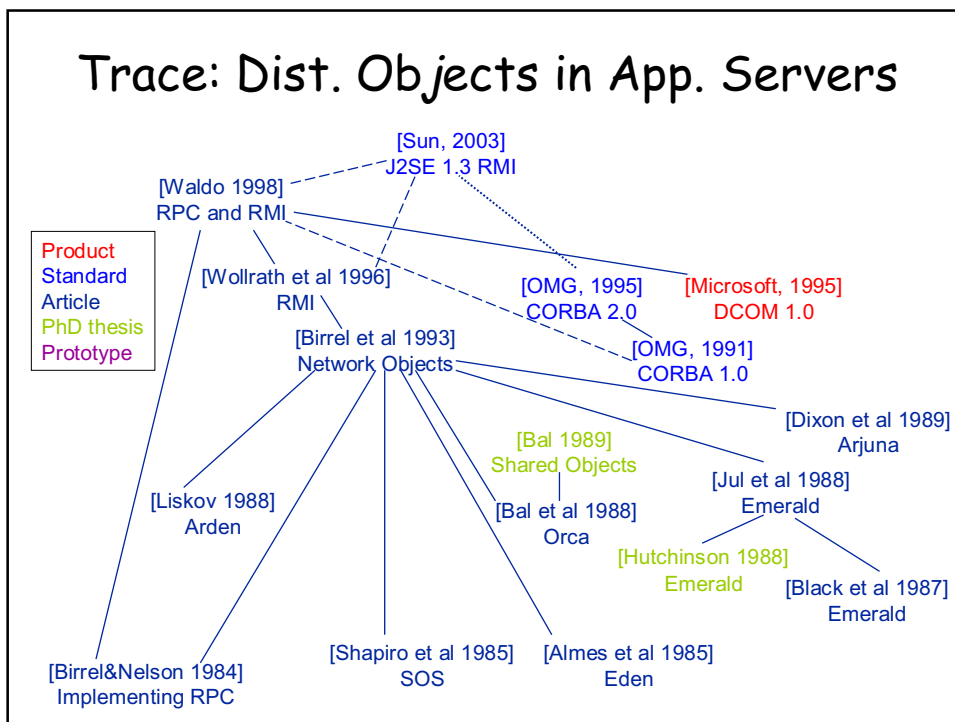
Trace: Web Services Description Lang.



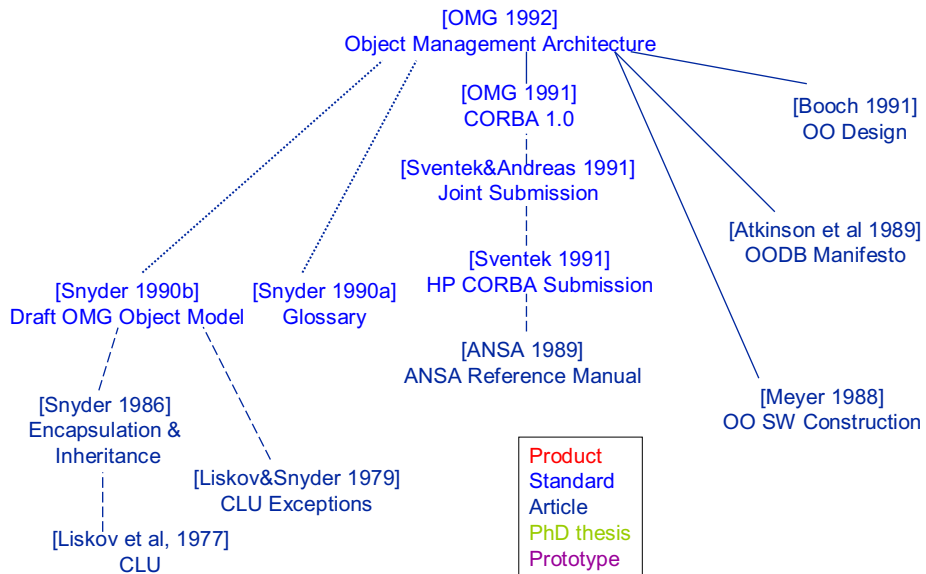
Trace: Messaging in App. Servers



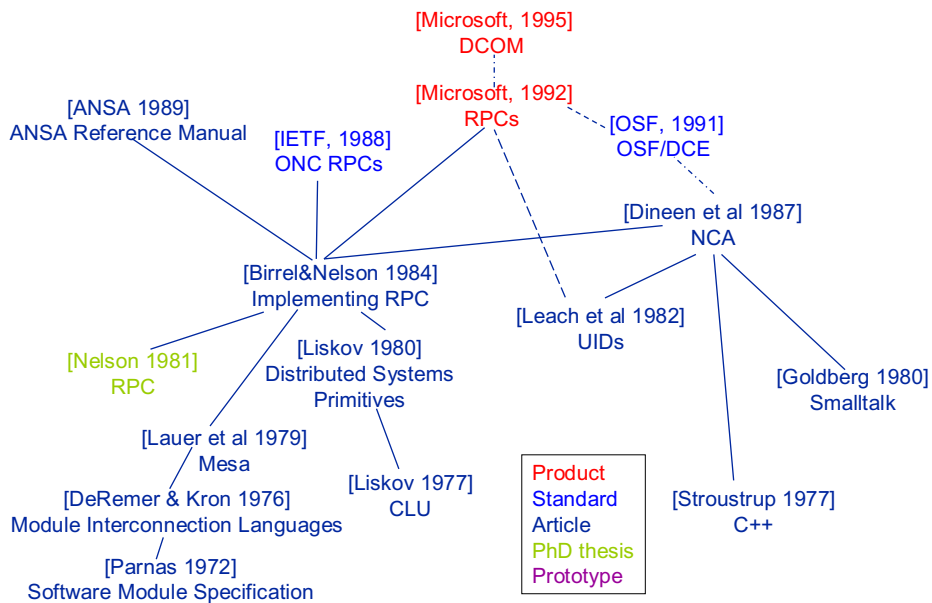
Trace: Dist. Objects in App. Servers



Trace: Distributed Objects in CORBA



Trace: Remote Procedure Calls



Some key findings/confirmations

Technology development is interdisciplinary

- often winds back and forth among disciplines

Technology maturation needs time

- 15-20 years between first publication of an idea and widespread availability in products

Technology transfer needs commitment

- people movement is most successful vehicle

PhD students are critical sources of ideas

- almost all impact traces lead back to PhD theses

Standards are critical enablers of ideas

- without widespread agreements on ideas there is no widespread adoption

Tech. development is interdisciplinary

Impact traces often cross CS disciplines

For middleware...

- software engineering
- networking
- programming languages
- distributed systems
- databases

Impact sometimes larger in area other than first publication

- e.g., message queues

Example: RPC IDLs

- infor. hiding [CACM 15(5), 1972]
- MIL [IEEE TSE SE-2(2), 1976]
- Mesa [ICSE-4, 1977]
- Cedar RPCs [ACM ToCS 2(1), 1984]
- Sun RPC [IETF RFC 1057, 1987]

Example: dist. transactions

- operating systems (Gray, 1976)
- nested transactions (Moss, 1981)
- concur. control (Bernstein, 1987)
- Arjuna (Dixon, 1989)
- OSF ODTP/XA (1991)
- CORBA CCS, OTS (1994)
- J2EE JTS, JTA (2001)

Technology maturation needs time

| RPCs | Dist. Transactions | Dist. Objects/RMI | |
|--|---------------------------------------|--|------|
| | | | 1970 |
| idea of module interconnection languages | | | |
| | | | 1980 |
| research on RPC systems | research on non-standard transactions | basic research and prototypes (Argus, Eden, Emerald) | |
| release of RPC into Apollo and Sun OSs | | | |
| | | | 1990 |
| standardization by IETF and OSF | standardization by OSF and OMG | consolidation as "network objects" | |
| | widespread use in application servers | standardization through JCP | |
| | | | 2000 |
| | | widespread use in Java and .NET | |

Technology transfer needs commitment

B. Nelson from CMU to Xerox PARC

- wrote definitive paper on RPCs with A. Birrel

B. Nelson and A. Birrel to DEC Research

- wrote Network Object paper providing the basis for Java RMI

A. Watson from APM to OMG

- lead CORBA standardization

A. Herbert from Cambridge to APM

- devised ANSA

G. Dixon from NCL to Transarc

- wrote OMG CORBA OTS and CCS service specs

J. Waldo from UMass, J. Sventek from APM to HP

- wrote CORBA 1.0 spec

J. Waldo from HP to Sun

- wrote RMI spec

PhD students are critical sources

Remote procedure calls

- architecture and failure semantics: Nelson (CMU 1981)
- orphan detection: Panzieri (Newcastle 1985)

Distributed transactions

- nested transactions: Moss (MIT 1981)
- object transactions: Dixon (Newcastle 1987)

Distributed object models

- general object models: Snyder (MIT 1978)
- RMI object model: Hutchinson (UW 1987) and Bal (Vrije 1989)

Web services

- Scribe: Reid (CMU 1981)

Impact reports roadmap

???

Software testing

Software architecture

Reviews and walkthroughs

Runtime assertion checking

Middleware technology

Modern programming languages

Software configuration management

Preliminary project "meta" findings

1. Lasting impact comes from repeated and sustained interactions
2. Interplay can be difficult to determine precisely and communicate clearly
3. More benefit from nurturing many and varied ways than single approach to impact
4. Community needs support to maintain the nurturing environment
5. Broad-based community funding may be more effective than targeted transitions

Further information and reading

Impact web site

- <http://www.acm.org/sigsoft/impact/>
- contains links to published reports on
 - » software configuration management (ACM TOSEM)
 - » modern programming languages (ACM TOSEM)
 - » middleware technology (ACM TOSEM)
 - » runtime assertion checking (ACM SIGSOFT SEN)