I've Looked at Clouds from Both Sides Now

Tobin J. Lehman
IBM Almaden Research Center
toby@almaden.ibm.com

Almost

- Almost a producer
  - Efforts to interest users/investors
- Almost a consumer
  - Efforts to evaluate clouds for use

Cloud Computing Lineage

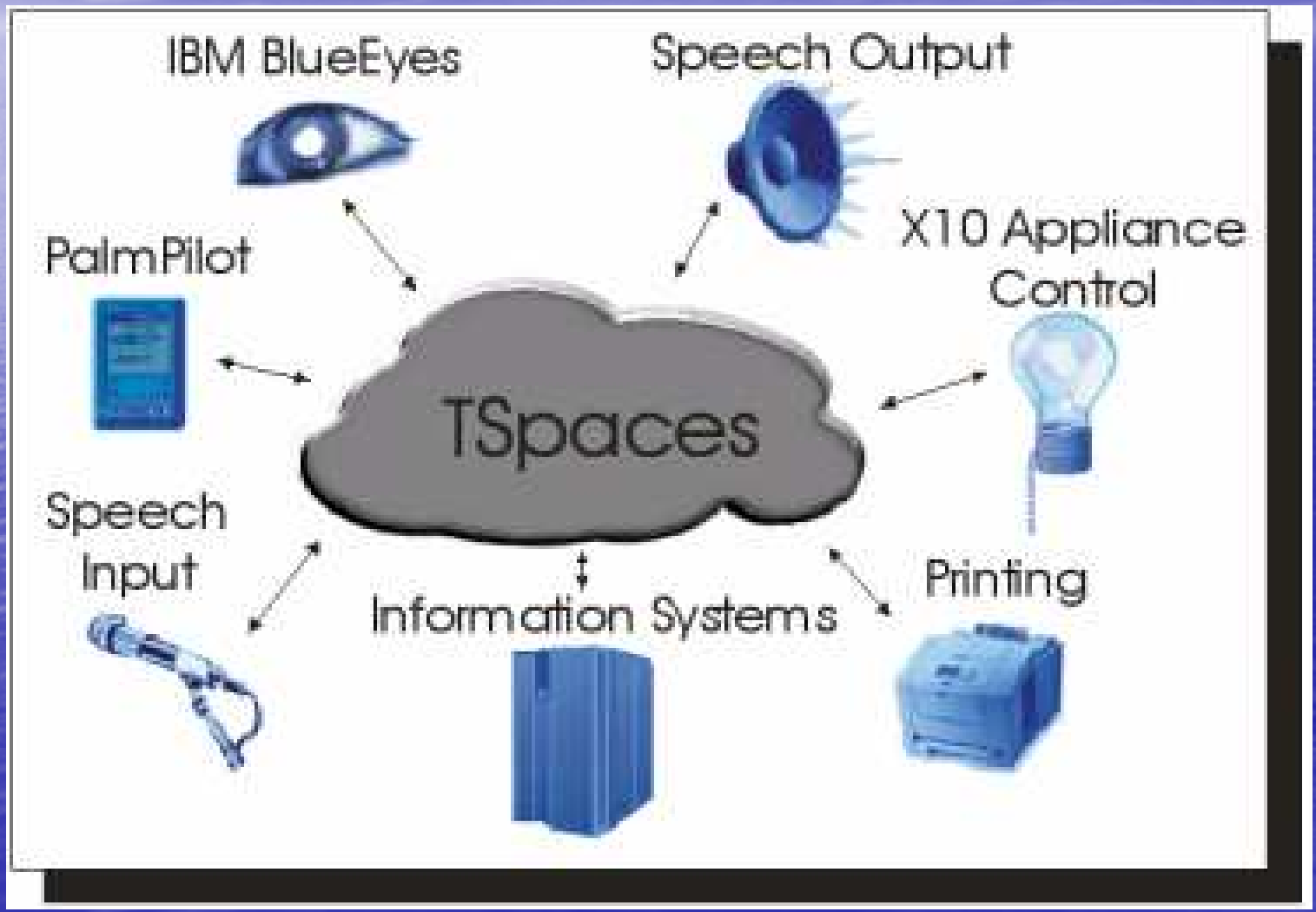
- Parallel and Distributed Computing
- Cluster Computing
- Application Hosting
- Grid (computation, data, function) computing
- Software as a Service

My History of Pre-Cloud Activities

- TSpaces ConnectionWare (1998 – 2001)
  - Ubiquitous Computing Infrastructure
  - Easy connection to processes living anywhere
  - Be Cool, Be Obscure
- OptimalGrid (2002-2004)
  - Coordinate distributed processes to organize large parallel computations
  - Create "The People's Grid" (inexpensive MIPS)
  - Apply autonomic grid techniques to scientific, IT management and game applications
  - Be Smart, Be Cheap, Be Fast, Be successful
- Solution Definition Manager (2005-2009)
  - High End Financial App -> Possible Cloud Client

What is TSpaces?

It is "intelligent connection-ware" that enables communication between disparate things.



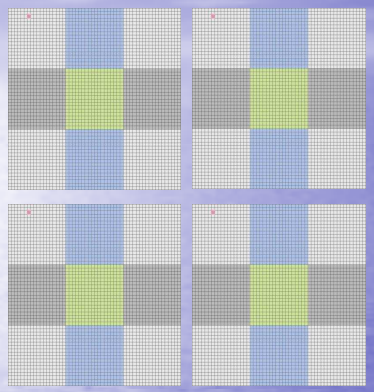
Thoughts on TSpaces

- We drew TSpaces as a "cloud" in order to stress the physical independence of the "ConnectionWare"
  - It was a bodiless msg provider (unlike MQ Series)
- Many were intrigued, but it lacked standards
  - Why not just be Jini/Javaspaces?
- Folks loved the simple ease of use, but it lacked enterprise-level (replicated) reliability
  - Had only "regular" log-based DB recovery
- Failed as a true internet service
  - But remains useful for intranet applications

TSpaces Summary

- Interesting, but not "Industrial Strength"

OptimalGrid:
Magically solving parallel problems on the Internet



Run this problem.

36 agents

4 whiteboards

1 master WB

1 APM

1 console

OptimalGrid: The Intent

- Easier to use than other grid systems
- Could solve parallel problems that otherwise would require expensive private clusters
- Free to the world of computational scientists.
- Bring easy to use parallel computing to the poor chemists, biologists and engineers.
- Oops. Did I say “free”?
- Also, dabbled in distributed game hosting.

OptimalGrid: Results and Realizations

- Lots of talk about the future of Grid
  - But, way more talk than money
- Some problems fun, some very hard
  - Parallel distributed autonomic grid: medium
  - Internet job submission of parallel jobs with flawless oversight security: hard
  - Two-way program verification: hard
  - Getting folks to vote with dollars: Very Hard
- Intranet use: Challenging, but possible
- Internet use: Impossible (for us)
  - Much harder than what we had planned for

OptimalGrid Summary

- Larger Problem Space than TSpaces
- Also, interesting, but not industrial strength

- Today, when I hear people say, “hey, just run that on the cloud,” I remember when I used to say, “Hey, run that on the grid.”

Solution Definition Manager (SDM)

- Model, Cost and Price the solution for outsourcing engagements
- Data point: a $500million deal has on the order of 50,000 cost items, each with 10 to 50 different cost attributes, 10 factors, 4 cost streams and 5 to 20 cost asset types.
- A ten year deal (120 months) could trigger:
  50k * 50 * 120 * 10 * 4 * 20 = 240 billion FPMs
  For each worksheet/attribute update
- Offloading calculations for report writing and "what if analysis" could be a large benefit to our user community (20,000 users)
- Are we a candidate for cloud computing?
  - Intranet server offload is not "cloud" (that's a server)
  - Would we let our calcs be hosted outside the firewall?

General Purpose compute and storage support would be nice... but

- Is our data protected?
  - Is it always safe, no matter what?
- Is privacy guaranteed?
  - Our data is extremely confidential
  - No one else should EVER be able to read/touch it
- Is our hosted program guaranteed to run in the cloud exactly as it runs on our server?
  - How do we verify the calculations?
  - Can we reliably prove “no tampering”?

Head in the Clouds

- What exactly are Clouds anyway?
- I'm not sure what it is, or where it is, but it seems to solve my computing problems.
- Are Clouds ready to hold my data?
  - Let's take a look.

Clouds: The Light and the Dark

- The Light:
  - You don’t own the hardware
  - Resources not in any specific location
  - Maintaining resource availability and integrity is someone else’s problem
- The Dark:
  - You don’t own the hardware
  - Resources not in any specific location
  - Maintaining resource availability and integrity is someone else’s problem

Rain Clouds

- Security Issues
  - Why should you trust them?
  - Why should they trust you?
- Audit Trails
  - Will the cloud provider track usage the same way you would? And, do you want that?
- Compliance issues
  - Does the cloud follow the same standards or compliance rules that you do?
- Legal Protection
  - Exactly what happens in the event of search and seizure at Cloud Headquarters?
  - Who is the warrant issued for?

Thunder Clouds

- Google problems with cloud computing (no quotes)

3,760,000 for problems with cloud computing.

- Google "problems with cloud computing"

715,000 for "problems with cloud computing".

  - The Real Problems with Cloud Computing
  - The Not So Obvious Problems with Cloud Computing
  - Dark Problems with Cloud Computing

Funnel Clouds: Your Data is whisked away ...

Leslie Lamport quote:

"You know you have [a distributed system] when the crash of a computer you never heard of stops you from doing any work"

- Gmail fail casts dark cloud on cloud computing
- Microsoft (Danger) and T-Mobile

Why is Cloud Computing Interesting?

- It Adds a Level of Indirection
  - Indirection enables technical flexibility
  - Indirection does not solve the safety, security and responsibility problem
  - Responsibility Indirection is Outsourcing.
- Cloud computing vs true IT Outsourcing
  - Differences in contract duration
  - Differences in costs
  - Differences in responsibilities

Conclusion

I've looked at clouds from both sides now,
From up and down, and still somehow,
It's cloud illusions I recall,
I really don't know clouds, at all.

-- My thanks and apologies to Joni Mitchell

References

- Paul Wallis, A Brief History of Cloud Computing: Is the Cloud There Yet?, Cloud Computing Journal http://cloudcomputing.sys-con.com/node/581838
- IBM's KittyHawk Project (get ref and details)
- Robert Lemos, 5 lessons from the dark side of cloud computing, http://www.infoworld.com/d/cloud-computing/5-lessons-dark-side-cloud-computing-669?source=IFWNLE_nlt_wrapup_2009-08-07
- http://www.katescomment.com/difference-it-outsourcing-cloud-computing/