The Big Idea

There is a revolution taking place

Technology is disappearing into the walls of the buildings to the point of becoming invisible.

Process and (web) service orchestration is king.
Into the walls!?
More Evidence
The Disappearing Infrastructure

Resting Data Bases
Getting your data base into RESTful state.

NOVEMBER 13, 2006

A question of style

In the last note, we looked at the base url format. This was:

www.resting-db.com/test/country?rep=xml

We can easily extend the syntax in this url to include a style sheet.

www.resting-db.com/test/country?rep=xml/style=linked_country.xslt

Using this link, you will get a formatted table of countries together with links to single country data. Choose one and you will see:

www.resting-db.com/test/country:code.EQ.'cc'?rep=xml,style=one_country.xslt'

where 'cc' is the country code.

In order to keep the url human readable, we have chosen to use the 'EQ.' or 'LIKE.' to represent '=' and 'like'. So to get a shorter list of countries whose codes start with 'K' we would construct the following url.

www.resting-db.com/test/country:code.LIKE.'K'*?rep=xml,style=linked_country.xslt

The syntax also contains a sort instruction after the '?'. So,

www.resting-db.com/test/country:code.LIKE.'K'*?rep=xml,style=linked_country.xslt:sort=alpha

OCTOBER 22, 2006

A RESTful database protocol

It is refreshing to come across a concept that has the promise of being substantially
The Disappearing Large Scale Infrastructure

Service Highlights

- **Elastic**
  Amazon EC2 enables you to increase or decrease capacity within minutes, not hours or days. You can commission one, hundreds or even thousands of server instances simultaneously. Of course, because this is all controlled with web service APIs, your application can automatically scale itself up and down depending on its needs.

- **Completely Controlled**
  You have complete control of your instances. You have root access to each one, and you can interact with them as you would any machine. Each instance predictably provides the equivalent of a system with a 1.7Ghz x86 processor, 1.75GB of RAM, 160GB of local disk, and 250Mb/s of network bandwidth.

- **Designed for use with Amazon S3**
  Amazon EC2 works in conjunction with Amazon Simple Storage Service (Amazon S3) to provide a combined solution for computing and storage across a wide range of applications.

- **Reliable**
  Amazon EC2 offers a highly reliable environment where replacement instances can be rapidly and reliably commissioned. The service runs within Amazon's proven network infrastructure and datacenters.

- **Secure**
  Amazon EC2 provides web service interfaces to control network security. You define groups of instances and their desired accessibility.

- **Inexpensive**
  Amazon EC2 passes on to you the financial benefits of Amazon's scale. You pay a very low rate for the compute capacity you actually consume. Compare this with the significant up-front expenditures traditionally required to purchase and maintain hardware, either in-house or hosted. This frees you from many of the complexities of capacity planning, transforms what are commonly large fixed costs into much smaller variable costs, and removes the need to over-buy "safety net" capacity to handle periodic traffic spikes.

Pricing

- Pay only for what you use.
- $0.10 per instance-hour consumed (or part of an hour consumed).
- $0.20 per GB of data transferred into/out of Amazon (i.e., Internet traffic).
- $0.15 per GB-Month of Amazon S3 storage used for your images (charged by Amazon S3).

Data transferred within the Amazon EC2 environment, or between Amazon EC2 and Amazon S3, is free of charge (i.e., $0.00 per GB).
The Disappearing File Servers
The Disappearing Office Suites

Office Suite

- **Zoho Writer**
  - Online word processor with collaboration features.
  - No download, no install, just sign up to create documents.
  - 100% free

- **Zoho Sheet**
  - Online alternative to traditional spreadsheet applications with powerful features like charting, collaboration & more.
  - 100% free

- **Zoho Show**
  - Online presentation tool to create, edit, publish, and show presentations.
  - 100% free

- **Zoho Wiki**
  - Wiki that is as easy to use as a word processor with group concept, versions, sub-pages and more...
  - 100% free

- **Zoho Notebook**
  - Create, Aggregate and Collaborate on multiple types of content online.

- **Zoho Virtual Office**
  - A web-based collaboration groupware that includes Email Client, Documents, Calendar and more.
  - On-Demand - Free for individuals

Utilities

- **Zoho Challenge**
  - The easiest way to evaluate candidates online.
  - Free up to 25 candidates. For more, price starts at $199

Productivity Tools

- **Zoho Projects**
  - Project management software to create, manage & collaborate online.

- **Zoho CRM**
  - Elegant & Affordable On-Demand CRM Solution.
  - Free up to 3 users. For more, price starts at $12

- **Zoho Creator**
  - Create Online Database Applications in minutes.
  - Build Forms, Collect Data and Manage.

- **Zoho Planner**
  - Online organizer to maintain your todo's, reminders, notes, attachments etc.

- **Zoho Chat**
  - Unique and intuitive way to make group decisions faster.
The Disappearing Applications
The Disappearing Workforce
Achieving Massive Scale
The Search for Jim Gray

Jim Gray Summary Home Page

Microsoft eScience Group
Gray@microsoft.com
455 Market St, 16th fl., San Francisco, CA, 94105
tel: (415) 778-8222,

Jim Gray is a researcher and manager of Microsoft Research’s eScience Group. His primary research interests are in databases and transaction processing systems -- with particular focus on using computers to make scientists more productive. He and his group are working in the areas of astronomy, geography, hydrology, oceanography, biology, and health care. He continues a long-standing interest on building supercomputers with commodity components, thereby reducing the cost of storage, processing, and networking by factors of 10x to 100x over low-volume solutions. This includes work on building fast networks, on building huge web servers with CyberBricks, and building very inexpensive and very high-performance storage servers.

Jim also is working with the astronomy community to build the world-wide telescope and has been active in building online databases like http://terraService.Net and http://skyserver.sdss.org. When the entire world’s astronomy data is on the Internet and is accessible as a single distributed database, the Internet will be the world’s best telescope. This is part of the larger agenda of getting all information online and easily accessible (digital libraries, digital government, online science …). More generally, he is working with the science community (Oceanography, Hydrology, environmental monitoring, …) to build the world-wide digital library that integrates all the world’s scientific literature and data in one easily-accessible collection. He is active in the research community, is an ACM, NAE, NAS, and AAAS Fellow, and received the ACM Turing Award for his work on transaction processing. He also edits a series of books on data management.

What’s new?
Performance of a Sun X4500 under Windows, NTFS and SQL server 2005. “(pdf) Sun loaned this storage/compute brick to JHU for some of the eScience internet services we are building here. This preliminary performance report shows it to be a balanced system (4 cpus, 160GB ram, 45 disks, 24TB all in 4U using 400W.) Here is the spreadsheet with the numbers for the graphs. And here is a zip of the test tools and scripts.

A radical view of Flash Disks: document and talk.

“SkyServer Traffic Report – The First Five Years,” is a study of the traffic on Skyserver.sdss.org, an eScience website. Done jointly with Vinod Singh, Alex Szalay, Ani Thakar, Jordan Peddie, Bill Brooks, Svetha Lakshmi, and Stan Varny it analyses the traffic trying to see how people and programs use the site, the data, and the batch job system.

“Cross-Matching Multiple Spatial Observations and Dealing with Missing Data”, with Alex Szalay, Tanas Budavari, Robert Lupton, Maria Nieto-Santisteban, and Ani Thakar explains how to spatially correlate observations of the same area (of the sky or earth or …).

“Life Under Your Feet: An End-to-End Soil Ecology Sensor Network Database, Web Server and Analytic Service,” with Katrina Slavova, Andreas Terzis, Renan Musallam-E, Joshua Cogan, Sam Smale, Stuart Ozer, Randall Bean, and Alex Szalay of JHU we built a end-to-end soil monitoring system deployed as a Baltimore urban forest sensor.
Amazon Mechanical Turk will have a limited scheduled outage on the morning of Sunday, March 11th. The outage will begin at about 12:00 am Pacific Standard Time, and is scheduled to last approximately three hours. During the outage you will not be able to accept or submit any HITs. Please do not accept any HITs that you cannot complete before then. Thank You.

**Timer: 00:00:00 of 60 minutes**

**Enter Product & Nutrition Label Information**
- Requester: Product Data Entry
- Qualifications Required: Product Data Entry Qualification is 100, Location is US
- Reward: $0.09 per HIT
- HITs Available: 49
- Duration: 60 minutes

**Enter Food Product Information**
- Grab some food items (see notes) from your kitchen cupboards, and find the UPC (under the barcode) for each item - it is usually twelve digits long (including smaller check digits on either side).
- Open this url in a separate tab or browser window:
  

- Enter in the full UPC from an item, to check if we already have it in our system. If so, try another, otherwise you'll be presented with a form requesting further details.
- Once successful, the UPC entered is reserved for you to complete, for up to an hour. Enter in all the required details, and as much optional information as you can provide. The form will usually take between 2 - 6 minutes to complete.
- When completed, the form will give you a token code, to ensure you get paid for the work completed. Copy/paste it into the input box below, and submit the HIT.
- Your work will be approved within 24 hours.

**Take Note:**
- We can only accept entries for products which have a full Nutrition Facts label.
- No baby foods or minimal nutrient foods, such as salt, bottled water or diet soda.
- Once you have reserved a UPC, do not close the page, you will lose the unique url that gives you the reservation.
- Do not reserve more UPCs than HITs assigned to you. You can only submit work for one UPC on each HIT.

**Enter your token code**
The Big Idea

There is a revolution taking place

Technology is disappearing into the walls of the buildings to the point of becoming invisible.

Process and orchestration is king.
Key Issues

Agility
Growth
Reliability
Flexibility
Innovation
Scalability
In the absence of enterprise architecture we are fossilising our organisations …

- Long term viability of the organisation
- Amounts available to invest for innovation and growth
- Present infrastructure and applications

Culture

Organisational arrangements and the way in which decisions are made

IT Investment decisions
A Prescription for IT

- Start with the end in mind
- Start moving towards virtualised infrastructure
- Start understanding how business processes and services are related
- Start understanding your legacy portfolio in terms of distinct business services
Enterprise Architecture is **not** Enterprise-wide IT Architecture

How would “**old**” and “**new**” co-exist and evolve with “**COTS**” & “**Build**”?

**Service Consumer**

**Service Provider**

**Service Layer**

**COTS Packages**

**Executable Processes**

**Abstract Processes**

**Enterprise Process Repository**

**Model, Simulate, Redesign**

**Build**

**Components**

**Objects/Classes**

**Code**

**SOA infrastructure created using BPEL orchestration**

**Package(s)**

**Service Wrapper**

**Service Wrapper**
Any Grand Enterprise Architecture Plan is unlikely to succeed.
Enterprise Architecture is a coherent, group of principles and representations of the entire organisation that allows for rapid analysis, design and scenario planning.
A Minimalist Approach

• Layout the foundations
  – Meta-model is paramount
  – Single Shared Repository
  – Top-level as soon as possible
  – “Enterprise” is actually “end to end”

• On-going
  – Work where the work is happening
  – Use in design not in documentation
  – Make sure everything is linked into the right level
  – Build the repository one project at a time
  – Introduce principles when it is “right” to do so
The Meta-model is the Essence
Investec approach

- Buy IP
- Buy experience
- Build infrastructure
- Create awareness
- Train
- Support
- Manage the change
- In-house competence
- Create content
- Populate libraries
- Create publishing capability
- Common understanding
- Re-use
- Retain corporate memory
ARIS the second abstraction
Benefits (past 9-10 months)

- Three process projects to support new application roll-out
- HR meta-model development triggers a significant dialogue
- Web publishing of the repository content against the meta-model provides answers to complex questions
- In some areas, application architecture becomes process driven
- Tangible improvements in quality and speed of requirements management
- Every day people are communicating with the diagrams as a point of reference and backdrop to their planning
Requirements Management – Benefits of a Repository Approach

- Automated Gap Analysis
- Automated impact analysis
- Productivity of Specification and other design elements
- Completeness checking and trace-ability
- Collaboration across domains
- Power of re-use of repository
The IT Design Process – To get to Functional Specifications
The IT Design Process – To get to Functional Specifications
Repackaging to ensure best use of existing and new – The Art of Application Architecture
Around 250 deliverables, which translates into 1500 medium level tasks, like:

“Test interface between G and the Data Warehouse”,

Project Plan Generated and kept in sync with the Application Architecture
Concluding Comments

- As technology fades into the fabric of business and society
- Business model, process and its orchestration becomes central to firms
- This is best communicated via well thought out models and representations
- These models represent the essence of the business itself
- Providing a platform for business transformation
Questions
Principles act as filters to help our decision making

1. Organization Level Preferences
2. Enterprise Architecture Principles

1. Business Requirements
2. Business Architecture Principles

1. Information System Needs of the Organization
2. Data Requirements
3. Data Architecture Principles
4. Application Requirements
5. Application Architecture Principles

1. Technology, Infrastructure, Network etc. requirements to support above needs
2. Technology Architecture Principles

Decision