

Cloud Computing & Data Management (Υπολογιστικά Νέφη & Διαχείριση Δεδομένων)



Δημήτρης Ζεϊναλιπούρ
Λέκτορας
Τμήμα Πληροφορικής
Πανεπιστήμιο Κύπρου



Αξιοποίηση Νέων Τεχνολογιών και η Εφαρμογή τους στα
Κυπριακά Κέντρα Πληροφόρησης, 8-9 Νοεμβρίου 2012, Παν. Κύπρου

Υπολογιστικά Νέφη (Ορισμός)



- Οι **ορισμοί** ως προς το τι είναι τα υπολογιστικά νέφη **διαφέρουν!**
- **Κοινό** υπόβαθρο των διάφορων ορισμών:
 - Υπολογιστική **ισχύς**, **αποθηκευτικός χώρος** και **λογισμικό** είναι **αγαθά** διαθέσιμα από μια **μεγάλη υποδομή**
 - **Utility Computing** (Υποδομές Κοινής Ωφελείας / **Pay-as-you-go** (Πληρώνουμε για τους πόρους που καταναλώνουμε)
 - **Virtualization**: Εικονικά Συστήματα
 - **Everything as a Service (*aaS)**

Υπολογιστικά Νέφη (Μεγάλες Υποδομές)



Google's Datacenter in Oregon



Υπολογιστικά Νέφη (Μεγάλες Υποδομές)



Microsoft Azure in Chicago



112 containers x 2000 servers = 224,000 nodes

Υπολογιστικά Νέφη (Utility Computing)



Το Amazon EC2 Σύστημα Διαπροσωπείας

The screenshot shows the AWS Management Console interface for Amazon EC2. The browser address bar displays the URL <https://console.aws.amazon.com/ec2/home#c=EC2&s=Instances>. The page title is "AWS Management Console - Windows Internet Explorer". The user is logged in as "Welcome, kishorekumar_62@yahoo.com".

The main content area is titled "My Instances" and shows a table of instances. One instance is selected, and a context menu is open over it. The instance details are as follows:

Instance	AMI ID	Zone	Security Group	Type	Status	Platform	Key Pair
<input checked="" type="checkbox"/> i-2392ff4a	ami-0529ce6c	us-east-1a	basicwin32	m1.small	running	windows	basicwin

The context menu options are: Reboot, Terminate, Launch more like this, Connect Help, Get System Log, Get Default Administrator Password, and Bundle Instance.

Below the table, the details for the selected instance are shown:

Instance:	i-2392ff4a		
AMI ID:	ami-0529ce6c	Zone:	us-east-1a
Security Groups:	basicwin32	Type:	m1.small
Status:	running	Owner:	970757705789
Reservation:	r-8bf871e2	Ramdisk ID:	-

The footer of the console shows the copyright notice: "© 2008, Amazon Web Services LLC or its affiliates. All right reserved." and links for Feedback, Support, Privacy Policy, and Terms of Use. The page is identified as an Amazon.com company.

Υπολογιστικά Νέφη (Pay-as-you-go)



To Amazon RDS (Relational Database Service)

Pay by the hour your DB Instance runs.

US – N. Virginia	US – N. California	EU – Ireland	APAC – Singapore
DB Instance Class			Price Per Hour
Small DB Instance		963\$ / year	\$0.11
Large DB Instance		↓	\$0.44
Extra Large DB Instance			\$0.88
Double Extra Large DB Instance			\$1.55
Quadruple Extra Large DB Instance		27,156 \$ / year	\$3.10

DB Instance Classes

Amazon RDS currently supports five DB Instance Classes:

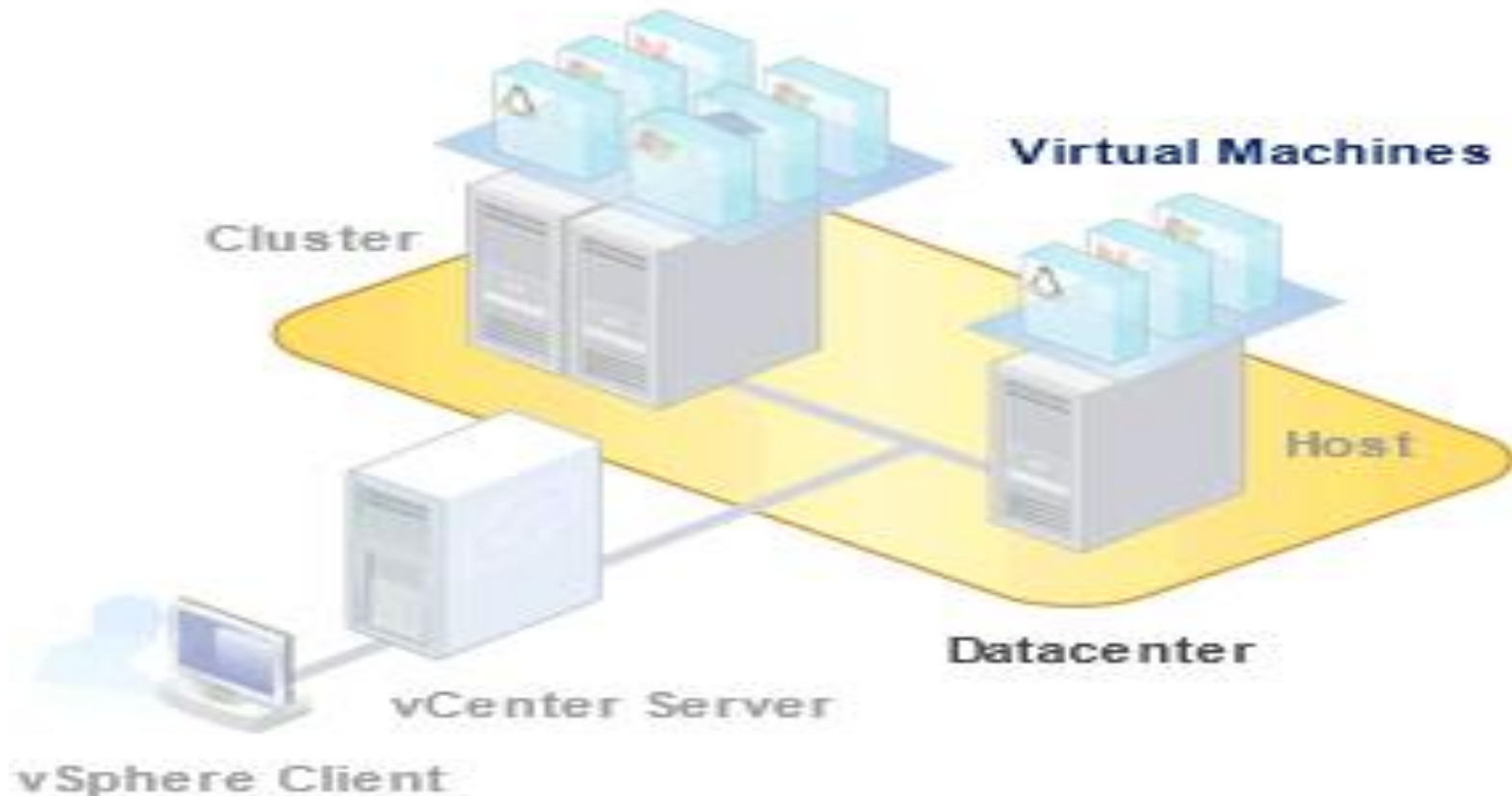
- Small DB Instance: 1.7 GB memory, 1 ECU (1 virtual core with 1 ECU), 64-bit platform, Moderate I/O Capacity
- Large DB Instance: 7.5 GB memory, 4 ECUs (2 virtual cores with 2 ECUs each), 64-bit platform, High I/O Capacity
- Extra Large DB Instance: 15 GB of memory, 8 ECUs (4 virtual cores with 2 ECUs each), 64-bit platform, High I/O Capacity
- Double Extra Large DB Instance: 34 GB of memory, 13 ECUs (4 virtual cores with 3.25 ECUs each), 64-bit platform, High I/O Capacity
- Quadruple Extra Large DB Instance: 68 GB of memory, 26 ECUs (8 virtual cores with 3.25 ECUs each), 64-bit platform, High I/O Capacity

For each DB Instance class, RDS provides you with the ability to select from 5GB to 1TB of associated storage capacity. One ECU provides the equivalent CPU capacity of a 1.0-1.2 GHz 2007 Opteron or 2007 Xeon processor.

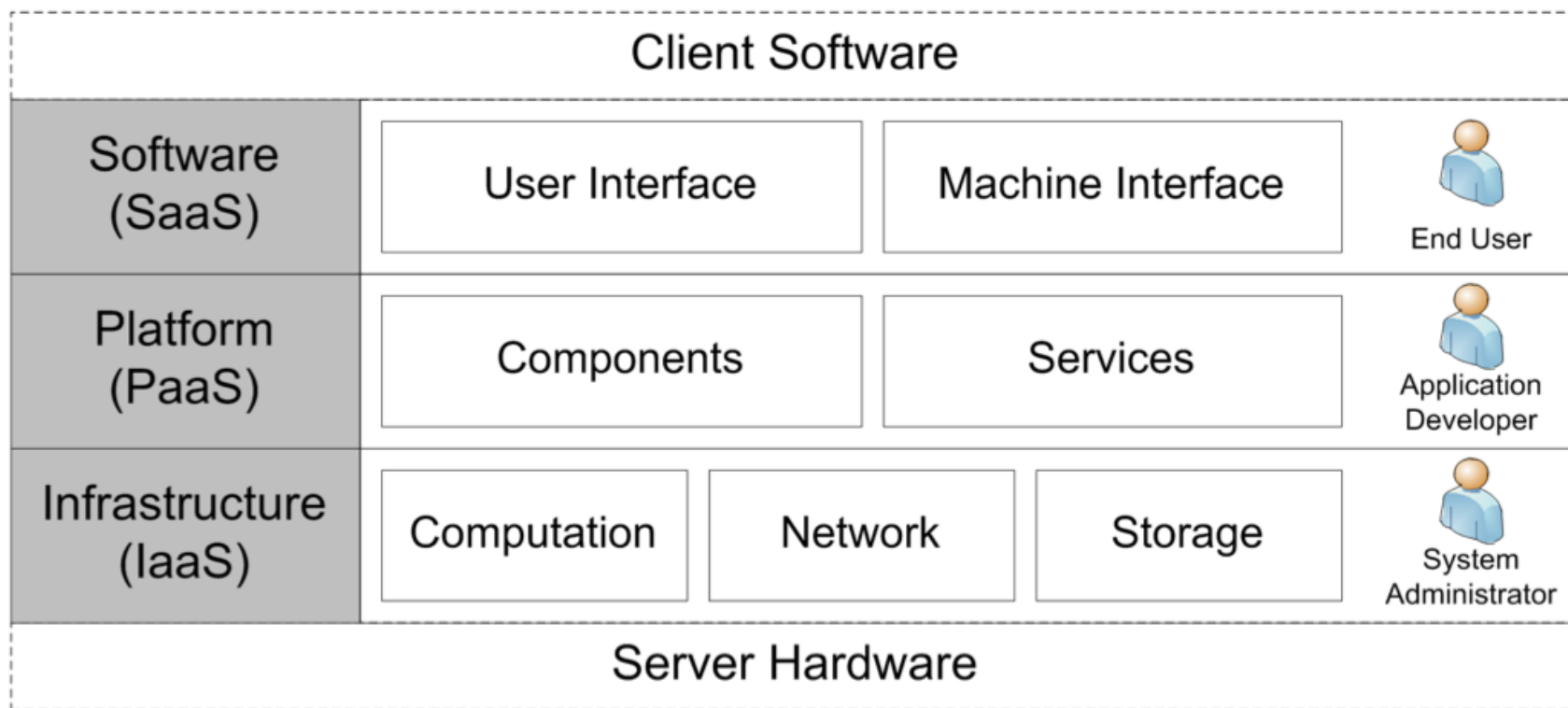
Υπολογιστικά Νέφη (Virtualization)



Εικονικά Συστήματα (CPU, RAM, Network, Storage, Fault-tolerance...)



Υπολογιστικά Νέφη (Everything as-a-Service)



Source: Wikipedia (<http://www.wikipedia.org>)

SaaS Examples: Google Apps, Quickbooks Online and Salesforce.com.

PaaS Examples: Amazon Elastic Beanstalk, Heroku, EngineYard, Google App Engine, and Microsoft Azure.

IaaS Examples: Amazon CloudFormation (and underlying services such as Amazon EC2), Rackspace Cloud, Google Compute Engine, and RightScale.

Δημόσια - Ιδιωτικά Νέφη (Public - Private Clouds)



- **Δημόσια Νέφη (Public Clouds)**
 - Amazon Web Service, Microsoft Azure, Google AppEngine, κτλ.
- **Ιδιωτικά "Νέφη" (Private Clouds)**
 - Κέντρα μέσα σε οργανισμούς (Datacenters)
 - Συνήθως **ΔΕΝ** περιλαμβάνονται στο ευρύτερο ορισμό των **Υπολογιστικών Νεφών**, αλλά αποτελούν λύση για πολλούς οργανισμούς
 - Καλύτερη Ιδιωτικότητα (EU Directive 95/46/EC on the protection of personal data / Safe Harbor)
 - Χαμηλότερη Καθυστέρηση Πρόσβασης
 - Απαιτούν γνώσεις διαχείρισης (backup, maintenance,...)

VCenter: Our IaaS Cloud



To Datacenter του
εργαστηρίου μας!



VCenter: Our IaaS Cloud



The screenshot displays the VMware vCenter Management console. The left-hand pane shows a hierarchical tree view of the vCenter environment, including hosts like 'esx5.in.cs.ucy.ac.cy' and 'esx1.in.cs.ucy.ac.cy', and various folders such as 'ADMIN', 'BSC', 'Images', 'RESEARCH', 'XP-Cluster', and 'vcenter'. The main pane is titled 'zenix' and shows the 'Summary' tab. The 'Status' section indicates the overall health is 'Normal'. The 'Guest OS Details' section shows the VM is 'Powered Off' and running 'CentOS 4/5/6 (64-bit)'. The 'VM Hardware' section lists resources: 2 CPU(s), 2048 MB of memory, and a 16.00 GB hard disk. The 'Related Items' section shows the VM is associated with 'esx5store' storage, 'DMSLGuestNet' network, and 'esx5.in.cs.ucy.ac.cy' host.

Status	
Overall	Normal

Guest OS Details	
Power State	Powered Off
Guest OS	CentOS 4/5/6 (64-bit)
IP Addresses	
DNS Name	
VMware Tools	Not running (Current)

VM Hardware	
CPU	2 CPU(s), 0 MHz used
Memory	2048 MB, 0 MB used
Hard disk 1	16.00 GB
Network adapter 1	DMSLGuestNet disconnected
CD/DVD drive 1	Disconnected
Floppy drive 1	Disconnected
Other	Additional Hardware
HW Version	8

Related Items	
Storage	esx5store
Networks	DMSLGuestNet
Host	esx5.in.cs.ucy.ac.cy
Resource Pool	esx5.in.cs.ucy.ac.cy



SmartLab: Our Smartphone IaaS Cloud!





<input checked="" type="checkbox"/>	ONLINE	HTC Hero (A6262)	3.2	320x480	Qualcomm Snapdragon MSM7600A	512	288								
<input checked="" type="checkbox"/>	ONLINE	HTC Desire S (S510e)	3.7	480x800	Qualcomm Snapdragon MSM8255	1024	768								
<input checked="" type="checkbox"/>	ONLINE	HTC Desire S (S510e)	3.7	480x800	Qualcomm Snapdragon MSM8255	1024	768								
<input checked="" type="checkbox"/>	ONLINE	HTC Desire (A8181)	3.7	480x800	Qualcomm Snapdragon QSD8250	1024	512								
<input type="checkbox"/>	ONLINE	HTC Desire	3.7	480x800	Qualcomm Snapdragon	1024	512	576	4	Gyroscope, G-Sensor, Magnetometer	Android 3.1	SH0APPL00958	CS7226		

<http://smartlab.cs.ucy.ac.cy/>

"Μεγάλα Δεδομένα" ("Big Data")

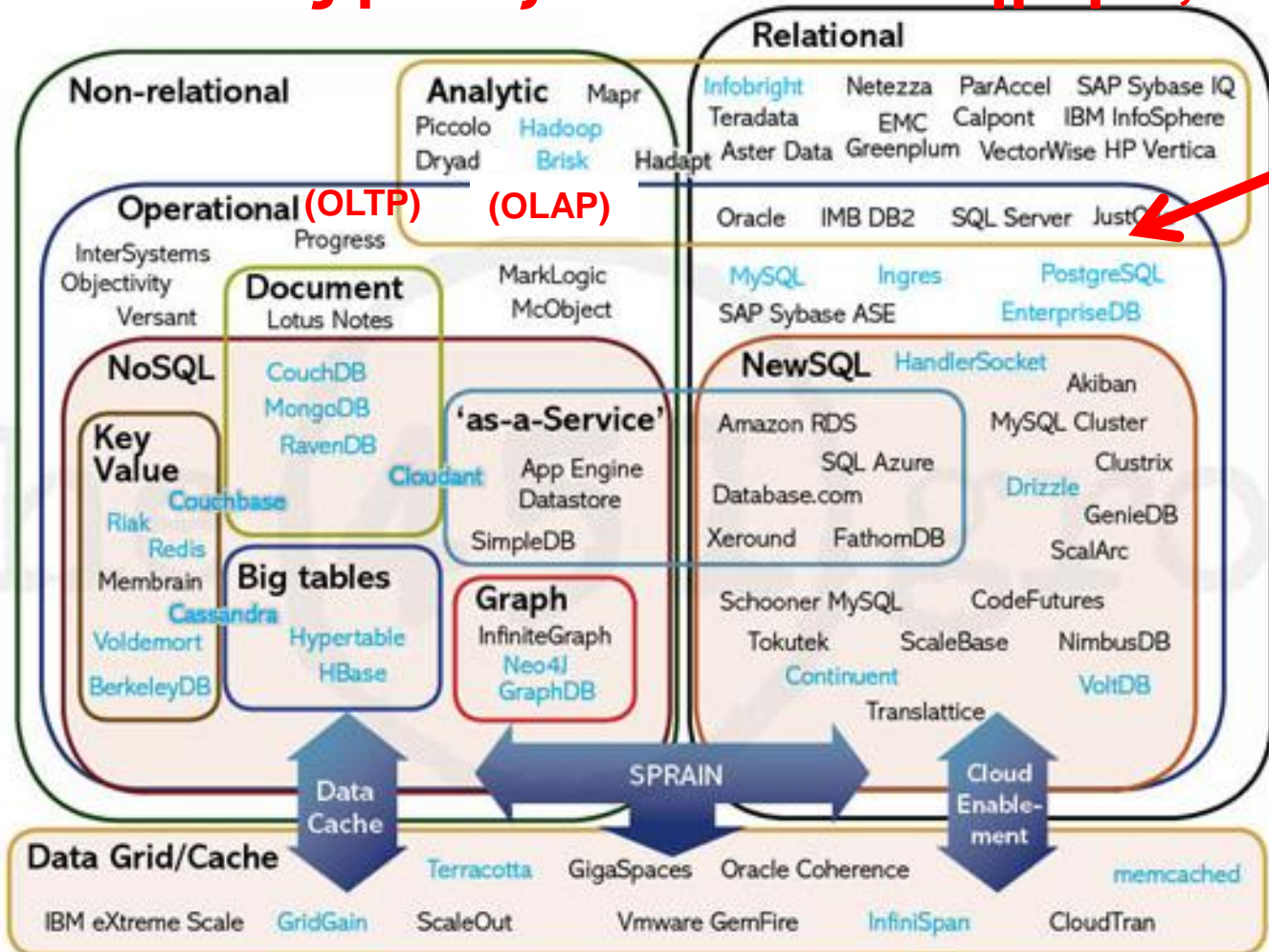


- Συλλογές δεδομένων που είναι **τόσο μεγάλες** και **περίπλοκες** που γίνεται **δύσχρηστη** η λειτουργία **συμβατικών συστημάτων** διαχείρισης δεδομένων (για μικρό σύνολο υπολογιστών)
 - **Παραδείγματα**
 - Το **Facebook** διαχειρίζεται πάνω από **40 δισεκατομμύρια** φωτογραφίες με **2000 υπολογιστές** (21PB HDFS Cluster). 
 - Το **Yahoo** δημιουργεί υποδομή **4000 κόμβων** για διαχείριση 60 εκατομμύριων αρχείων 
 - Έχουμε **αντίστοιχα παραδείγματα** δεδομένων στην τοπική αγορά ; **Όχι ακόμη** αλλά στο μέλλον θα **προκύπτει όλο** και περισσότερο αυτή η ανάγκη.
 - *Smart meters* σε Σπίτια, Δικτυακά Δεδομένα Τηλεπικοινωνιακών Οργανισμών, Βιβλιοθήκες Πολυμέσων του **Μέλλοντος**, κτλ.;
- Δημήτρης Ζεϊναλιπούρ – Τμήμα Πληροφορικής – ΠΚ (<http://dmsl.cs.ucy.ac.cy/>)

Συστήματα Διαχείρισης Δεδομένων



Πως μοιάζει το τοπίο σήμερα;



Συμβατικές RDBMSs

Venn Diagram by 451 group

NoSQL και NewSQL



- **NoSQL:** Μια μεγάλη κλάση DBMSs που δεν ακολουθούν το σχεσιακό μοντέλο (δεν χρησιμοποιούν πίνακες αλλά ούτε και SQL)
 - π.χ., JSON Stores, GraphDBs, Map-Reduce DBs
- **NewSQL:** Σχεσιακές DBMS κατασκευασμένες για το Cloud π.χ., το Amazon RDS, VoltDB, κτλ.
- **Κοινά Χαρακτηριστικά**
 - Κατασκευασμένα να δουλεύουν πάνω σε **αναξιόπιστες κατανεμημένες** υποδομές (2-3 failures / 1000 nodes / day)
 - Λιγότερες Εγγυήσεις ορθότητας (consistency)
 - Ψηλότερη επίδοση και επεκτασιμότητα
 - Π.χ., μια Hbase τρέχει πάνω σε 4000 nodes!

SmartLib: Νοητές Βιβλιοθήκες



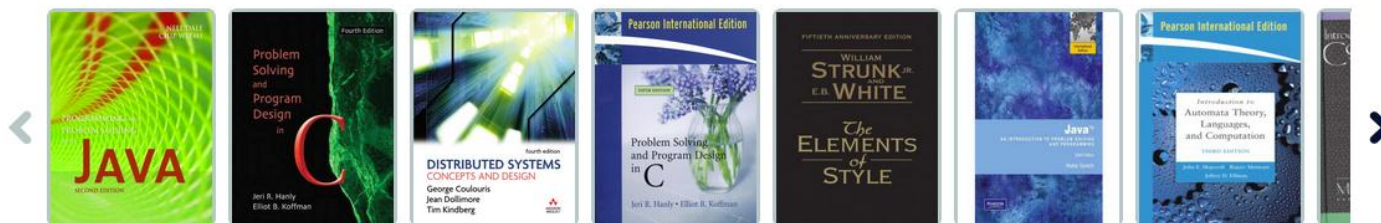
<http://smartlib.cs.ucy.ac.cy/>

SmartLib

HOME REGISTER DOWNLOADS ABOUT CONTACT

Username
Password
Login

Popular Books



All Books

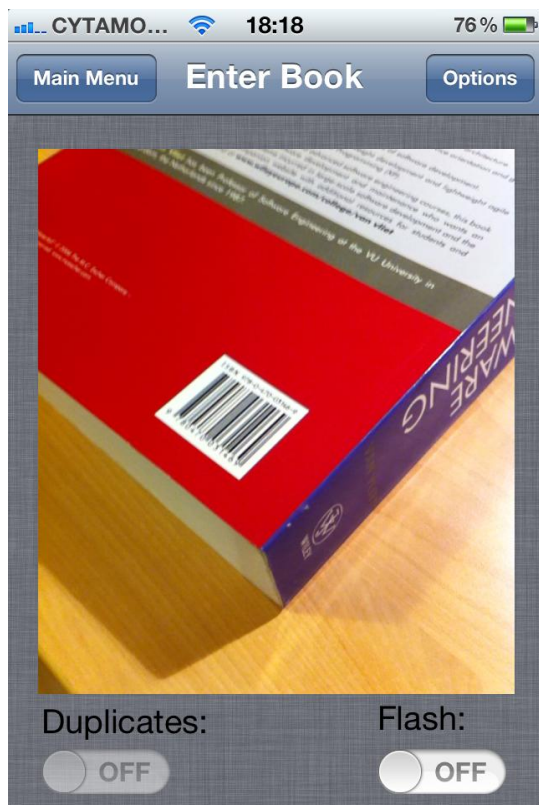
Search

Cover	Title	Authors	Published	Pages	Language	ISBN
	3D computer graphics	Alan H. Watt	2000	570	en	9780201398557
	3D games: Animation and advanced real-time rendering	Alan H. Watt, Fabio Policarpo	2003	547	en	9780201787061

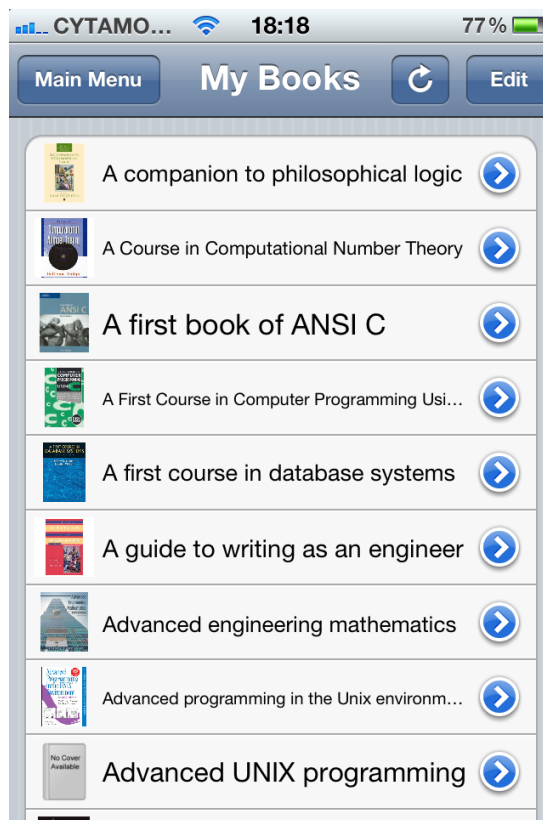
SmartLib: Νοητές Βιβλιοθήκες



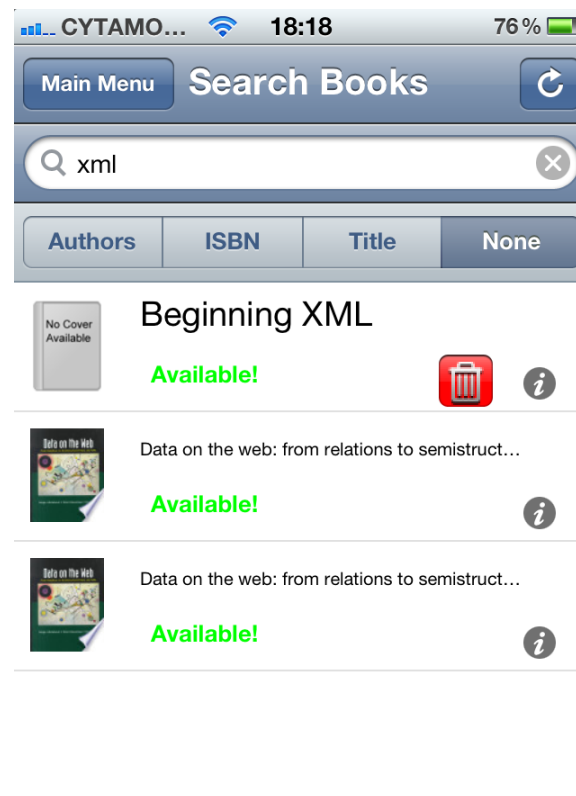
Scan



Manage



Share



SmartLib: Νοητές Βιβλιοθήκες



Web2.0: The Semi-structured Web!

<https://www.googleapis.com/books/v1/volumes?q=databases>

The screenshot shows a Google search interface with the search term 'databases'. The search results are displayed in a list format. The first result is 'Databases: Organizing Information' by Greg Roza, published in 2010. The second result is 'Advanced Database Systems: 10th British National Conference on ...' edited by Peter M.D. Gray and Rob J. Lucas, published in 1992. The third result is 'Databases Illuminated' by Catherine Ricardo, published in 2011. The fourth result is 'Fuzzy Databases: Modeling, Design And Implementation' by ... published in 2010. The search results are displayed in a list format with a sidebar on the left containing navigation options like 'Web', 'Images', 'Maps', 'Videos', 'News', 'Shopping', 'Books', and 'More'. The 'Books' section is currently selected.

Google databases

Search About 15,300,000 results (0.61 seconds)

Web **Databases: Organizing Information**
books.google.com/books?isbn=143589426X
Greg Roza - 2010 - Preview - More editions
Describes how databases work and how to use tables, files, and relational databases.

Images

Maps

Videos

News **Advanced Database Systems: 10th British National Conference on ...**
books.google.com/books?isbn=3540556931
Peter M.D. Gray, Rob J. Lucas - 1992 - Preview - More editions
The theme of this book is the potential of new advanced database systems. The volume presents the proceedings of the 10th British National Conference on Databases, held in Aberdeen, Scotland, in July 1992.

Shopping

Books **Databases Illuminated**
books.google.com/books?isbn=1449606008
Catherine Ricardo - 2011 - Preview - More editions
This Second Edition has been revised and updated to incorporate information about the new releases of Access 2010, Oracle 11g, and Intersystems Cache.

More

Browse books **Fuzzy Databases: Modeling, Design And Implementation**
books.google.com/books?isbn=1501402242

Any view
Preview and full view
Full view

"infoLink": "http://books.google.com/books?id=4Z6tFpuBmqaC&dq=databases&hl=&source=dbs_api",

Περισσότερες Πληροφορίες

<http://dmsl.cs.ucy.ac.cy/>



DATA MANAGEMENT SYSTEMS LABORATORY

Dept. of Computer Science, University of Cyprus



University
of Cyprus

Google™ Custom Search

Search x



SmartTrace
Crowdsourced Trajectory Similarity

DMSL News

Follow @dmslucy

DMSL to develop an innovative smartphone technology for a Finnish alliance backed by Nokia and Microsoft!

"Crowdsourcing with Smartphones" published in IEEE Internet Computing / SI: Crowdsourcing, vol. 16 no. 5 - <http://t.co/S0gB9XqV>

DMSL joins the EU FP7 Planet-Data Network of Excellence on Large Scale Data Management as an Associated Partner - <http://t.co/nNoM5Dap>

"Intelligent search in social communities of smartphone users" accepted in Springer's Distr. and Parallel DBs journal! <http://t.co/9wsMltxE>

"Crowdsourced Trace Similarity with Smartphones" to appear in upcoming IEEE TKDE journal (<http://t.co/Z0yQ208g>)

"Research in our laboratory focuses on Data Engineering Systems and Knowledge Discovery Solutions for emerging networks and systems (smartphone, sensor, crowd, peer-to-peer, etc.)"



Home



People



Publications



Courses



Projects



Code



Information



Contact

Copyright © 2010 Data Management Systems Laboratory (DMSL) - Department of Computer Science - University of Cyprus

Cloud Computing & Data Management (Υπολογιστικά Νέφη & Διαχείριση Δεδομένων)



Ευχαριστώ!

Δημήτρης Ζεϊναλιπούρ

Λέκτορας

Τμήμα Πληροφορικής
Πανεπιστήμιο Κύπρου



Αξιοποίηση Νέων Τεχνολογιών και η Εφαρμογή τους στα
Κυπριακά Κέντρα Πληροφόρησης, 8-9 Νοεμβρίου 2012, Παν. Κύπρου

JSON: Web 2.0 Data Interchange Format (Other Google APIs)



In fact, Web2.0 Services are omnipresent!

(Google, Twitter, Facebook, Youtube, Linkedin, ...)

<http://www.programmableweb.com/> - 7800 APIs!!! + 6800 Mashups!

The screenshot shows the Google APIs console interface. The URL is <https://code.google.com/apis>. The page title is "All services" and it says "Select services for the project." Below this is a table of services with columns for Service, Status, and Notes.

Service	Status	Notes
Ad Exchange Buyer API	<input type="checkbox"/> OFF	Courtesy limit: 1,000 requests/day
AdSense Host API	Request access...	Courtesy limit: 100,000 requests/day
AdSense Management API	<input type="checkbox"/> OFF	Courtesy limit: 10,000 requests/day
Analytics API	<input type="checkbox"/> OFF	Courtesy limit: 50,000 requests/day
Audit API	<input type="checkbox"/> OFF	Courtesy limit: 10,000 requests/day
BigQuery API	<input type="checkbox"/> OFF	Courtesy limit: 10,000 requests/day • Pricing
Blogger API v3	Request access...	Courtesy limit: 10,000 requests/day
Books API	<input checked="" type="checkbox"/> ON	Courtesy limit: 1,000 requests/day
Calendar API	<input type="checkbox"/> OFF	Courtesy limit: 10,000 requests/day
Custom Search API	<input type="checkbox"/> OFF	Courtesy limit: 100 requests/day • Pricing

JSON: Web 2.0 Data Interchange Format (Other Google APIs)

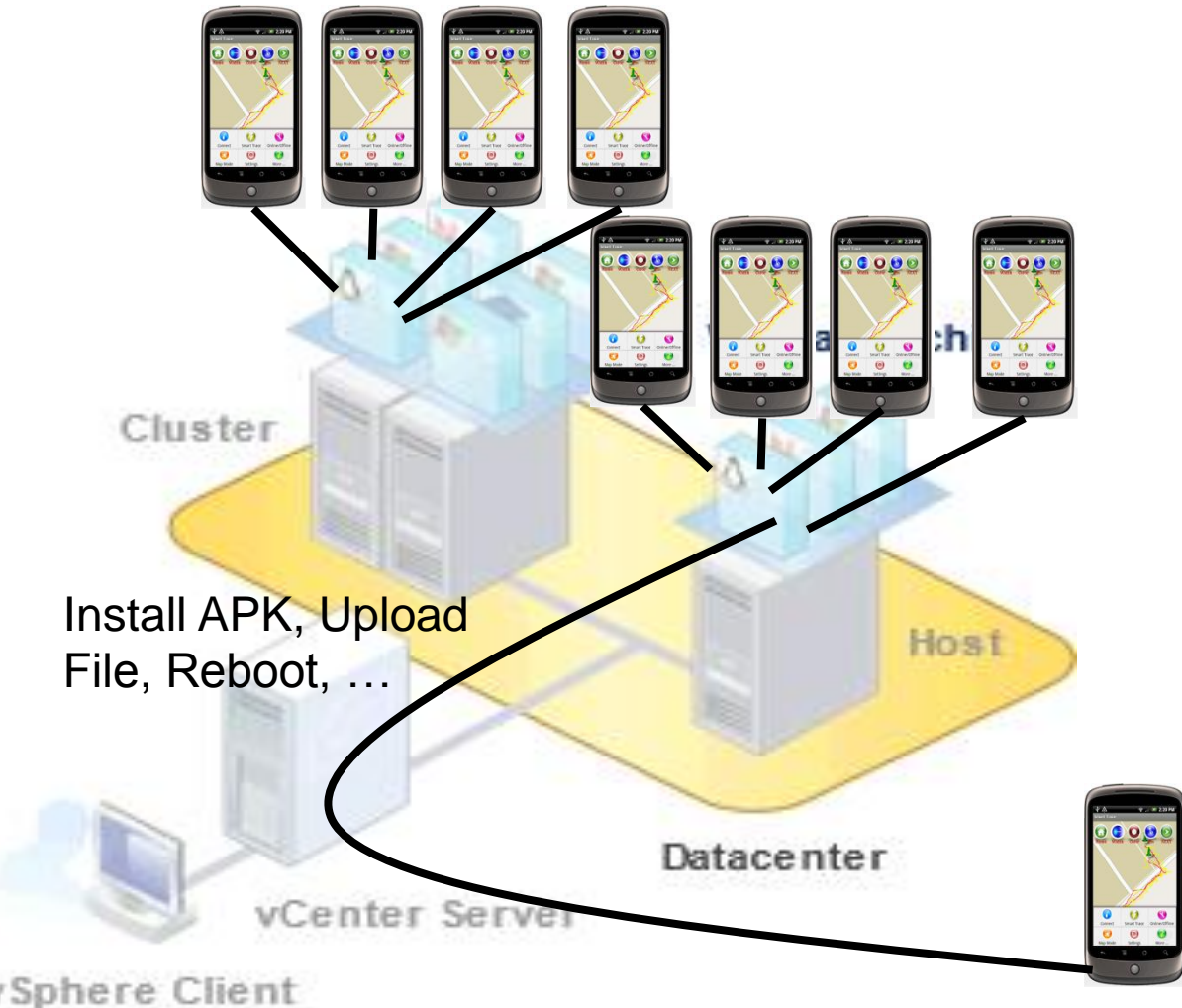


Yahoo Pipes Mashup

The screenshot shows the Yahoo Pipes Mashup editor. The browser address bar displays the URL: `pipes.yahoo.com/pipes/pipe.edit?_id=xOE_1Z8C3RGmkQrul7okhQ`. The page title is "Title Mangler". The interface includes a sidebar with categories like Sources, User Inputs, Operators, and Favorites. The main workspace contains several interconnected widgets: "Prefix (text)", "FeedURL (text)", "Postfix (text)", "Fetch Feed", "Loop", and "String Builder". The "Loop" widget is configured to "For each item in input feed" and contains a "String Builder" widget with three "text [wired]" fields and an "assign results to item.title" option. A "Pipe Output" button is visible at the bottom right of the workspace.

A **mashup**, in web development, is a web page, or web application, that **uses and combines data, presentation or functionality from two or more sources to create new services.** Wikipedia

SmartLab Programming Cloud



URL: <http://smartlab.cs.ucy.ac.cy/>

Δημήτρης Ζεϊναλιπούρ – Τμήμα Πληροφορικής – ΠΚ (<http://dmsl.cs.ucy.ac.cy/>)