Εθνικό και Καποδιστριακό Πανεπιστήμιο Αθηνών
Τμήμα Πληροφορικής & Τηλεπικοινωνιών

Πληροφοριακά Συστήματα

7ο Εξάμηνο 2021-22

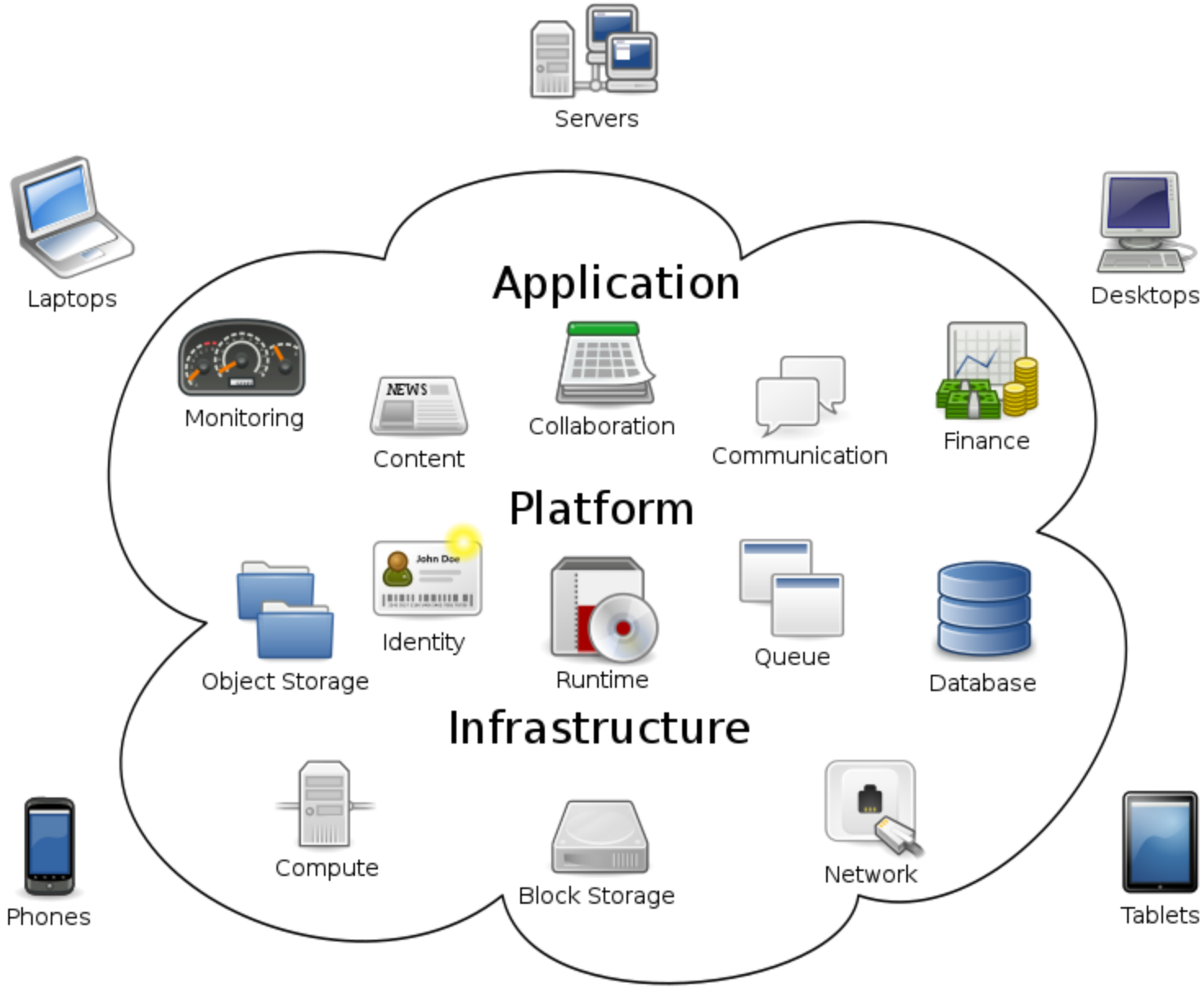
Υπολογιστικό νέφος

Δρ. Κώστας Σαΐδης (saiko@di.uoa.gr)

Υπολογιστικό νέφος (γενικά)

Περί τίνος πρόκειται

Η κατ' αίτηση (on-demand) διάθεση υπολογιστικών πόρων στο διαδίκτυο, χωρίς να απαιτείται από τον χρήστη η ενεργή του συμμετοχή στη διαχείρισή τους.



Cloud computing

Χαρακτηριστικά (κατά NIST)

- On-demand self-service: Κατά το δοκούν ενεργοποίηση πόρων/υπηρεσιών από τους χρήστες
- Broad network access: Μεγάλες ταχύτητες, "κοντά" στην υποδομή του δημόσιου internet
- Resource pooling: Πολλές διαθέσιμες επιλογές πόρων
- Rapid elasticity: Άμεση -και αυτοματοποιημένη αν απαιτείται- διάθεση πρόσθετων πόρων
- Measured service: Μέτρηση της "κατανάλωσης" των πόρων

Κατηγορίες

- Private (enterprise) cloud
- Public cloud
- Hybrid cloud

Vendor lock-in

- Σημαντική έννοια
- Το "τεχνολογικό κλείδωμα" μιας λύσης σε έναν συγκεκριμένο προμηθευτή

Μοντέλα Υπηρεσιών στο Υπολογιστικό νέφος

Infrastructure as a Service

- Virtual machines
- Servers
- Storage
- Load balancers
- Networking

Platform as a Service

- Database
- Web server
- Application server
- Development tools

Software as a Service

- Information Systems (CRM, ERP, etc)
- Email
- Collaboration
- Gaming

Backend as a Service

- APIs for backend storage and computing services
  - User management
  - Push notifications
  - Integrations
  - etc
- Custom SDKs
- Started as a backend solution for Mobile app developers

Serverless Computing & Function as a Service

- Simply run a command to deploy your code
- Pay the invocations of your code
- No need to pre-purchase units of capacity
- Pay as you go

Οργανισμό

- Ευελιξία
- Μείωση κόστους
- Ανεξαρτησία από τη συσκευή
- Ανεξαρτησία από την τοποθεσία
- Ευκολότερη συντήρηση
- Καλύτερη αξιοποίηση πόρων
- Μέτρηση της απόδοσης του συστήματος
- Αύξηση παραγωγικότητας
- Κλιμάκωση
- Ασφάλεια

Αναλυτικότερα

Agility for organizations may be improved, as cloud

Cost reductions are claimed by cloud providers. A public-cloud delivery model converts capital expenditures (e.g., buying servers) to operational expenditure. This purportedly lowers barriers to entry, as infrastructure is typically provided by a third party and need not be purchased for one-time or infrequent intensive computing tasks. Pricing on a utility computing basis is "fine-grained", with usage-based billing options. As well, less in-house IT skills are required for implementation of projects that use cloud computing.]

Costs savings depend on the type of activities supported and the type of infrastructure available in-house.

Device and location independence enable users to access systems using a web browser regardless of their location or what device they use (e.g., PC, mobile phone). As infrastructure is off-site (typically provided by a third-party) and accessed via the Internet, users can connect to it from anywhere.

Maintenance of cloud computing applications is easier, because they do not need to be installed on each user's computer and can be accessed from different places (e.g., different work locations, while travelling, etc.).

Multitenancy enables sharing of resources and costs across a large pool of users thus allowing for:

- centralization of infrastructure in locations with lower costs (such as real estate, electricity, etc.)
- peak-load capacity increases (users need not engineer and pay for the resources and equipment to meet their highest possible load-levels)
  (utilisation and efficiency improvements for systems that are often only 10–20% utilised.

Performance is monitored by IT experts from the service provider, and consistent and loosely coupled architectures are constructed using web services as the system interface.

Productivity may be increased when multiple users can work on the same data simultaneously, rather than waiting for it to be saved and emailed. Time may be saved as information does not need to be re-entered when fields are matched, nor do users need to install application software upgrades to their computer.

Availability improves with the use of multiple redundant sites, which makes well-designed cloud computing suitable for business continuity and disaster recovery.

Scalability and elasticity via dynamic ("on-demand") provisioning of resources on a fine-grained, self-service basis in near real-time, without users having to engineer for peak loads. This gives the ability to scale up when the usage need increases or down if resources are not being used. Emerging approaches for managing elasticity include the use of machine learning techniques to propose efficient elasticity models.

Security can improve due to centralization of data, increased security-focused resources, etc., but concerns can persist about loss of control over certain sensitive data, and the lack of security for stored kernels. Security is often as good as or better than other traditional systems, in part because service providers are able to devote resources to solving security issues that many customers cannot afford to tackle or which they lack the technical skills to address. However, the complexity of security is greatly increased when data is distributed over a wider area or over a greater number of devices, as well as in multi-tenant systems shared by unrelated users. In addition, user access to security audit logs may be difficult or impossible. Private cloud installations are in part motivated by users' desire to retain control over the infrastructure and avoid losing control of information security.