



# **PMINJ Chapter 01 May Symposium 2017**



## ***Cloud Computing and Big Data. What's the Big Deal?***

**Arlene Minkiewicz, Chief Scientist  
PRICE Systems, LLC  
[arlene.minkiewicz@pricesystems.com](mailto:arlene.minkiewicz@pricesystems.com)**

# Agenda

- Introduction
- Cloud Computing
  - Defined
  - Benefits and Risks
- Big Data
  - Defined and Applied
  - Examples
- Wrap Up



# Introduction

- Cloud computing is a paradigm that opens the door for utility computing
- Instead of investing in hardware, software and infrastructure, organizations can access through the cloud on an as-needed basis
- Still lots of hype – some vendors have their head further in the clouds than their technology



# Introduction

- **Cloud computing is a technology that's here to stay**
  - A report by Cisco found that 83% of all data center traffic will be based in the cloud within the next 3 years
  - According to Synergy Research Group – the cloud computing market grew 21% in 2015 to the tune of \$110 billion
  - According to IDC (International Data Corporation) spending on public cloud services should double from \$70 B in 2015 to \$141 B in 2019
  - In 2016 Hybrid cloud adoption rose from 58% to 71% as more companies are embracing this type of solution
  - Gartner predicts that by 2020 a Corporate “No-Cloud” policy will be as rare as a “No-Internet” Policy today



# You probably already are a cloud consumer!



twitter



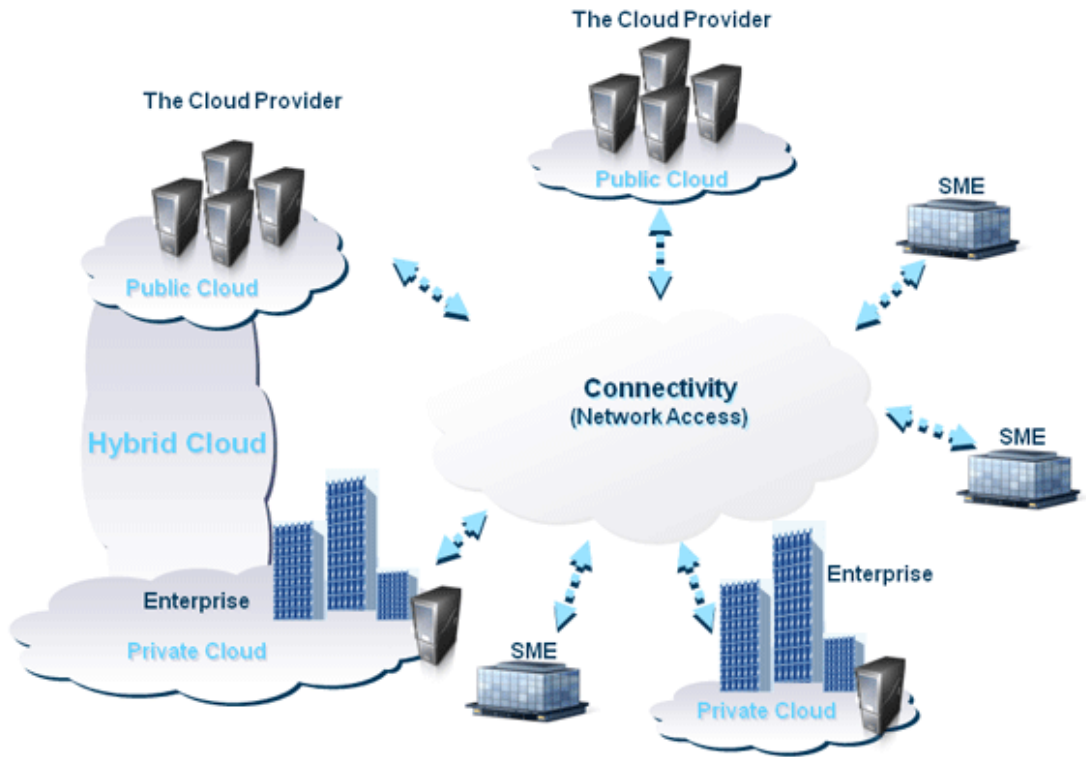
# Cloud Computing

- Consumers of cloud computing access hardware, software and networking capabilities from third party providers
- The cloud can be defined as resources and applications that are available on the Internet or other network via any device that connects to the Internet or other network
- According to National Institute of Standards and Technology (NIST), cloud computing delivers the following...
  - On demand self service
  - Ubiquitous network access
  - Location independent resource pooling
  - Rapid elasticity
  - Measured services



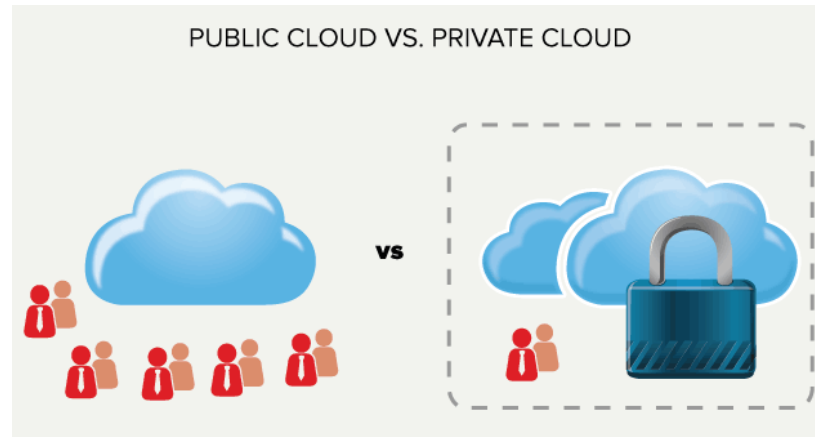
# Cloud Computing

- Three types of clouds
  - Public cloud
  - Private cloud
  - Hybrid cloud



# Public and Private Clouds

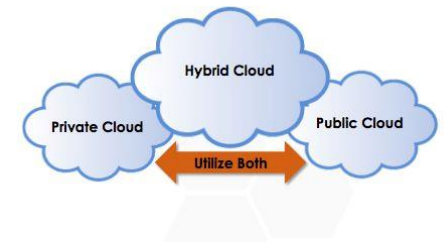
- **Public Cloud**
  - Any cloud service that is delivered via the Internet
  - Data on public cloud services is that which is most likely to be compromised in a cloud environment
- **Private Clouds** are used by those organizations that want to take advantage of cloud technology but don't want to risk vulnerabilities introduced by the Internet
  - Private clouds can be within an organizations fire wall (on premise)
  - Private clouds could be housed with a cloud vendor but with no multi tenancy and no access to the Internet





# Hybrid Clouds

- Combines onsite private clouds with resources from the Public Cloud
- Organizations gain the benefits of cloud computing while using public cloud services in situation where data or applications are sensitive
- Cloud and compliance have advanced significantly since its introduction – Lockheed Martin, Microsoft, and the Dept. of Agriculture are all running federally compliant clouds
- In 2016 hybrid cloud adoption grew from 58% to 71% over 2015
- Gartner predicts that by the end of 2017 nearly half of large enterprises will have hybrid cloud deployments



# Cloud Computing

## ■ Cloud computing offerings include:

### – Software as a Service (SaaS)

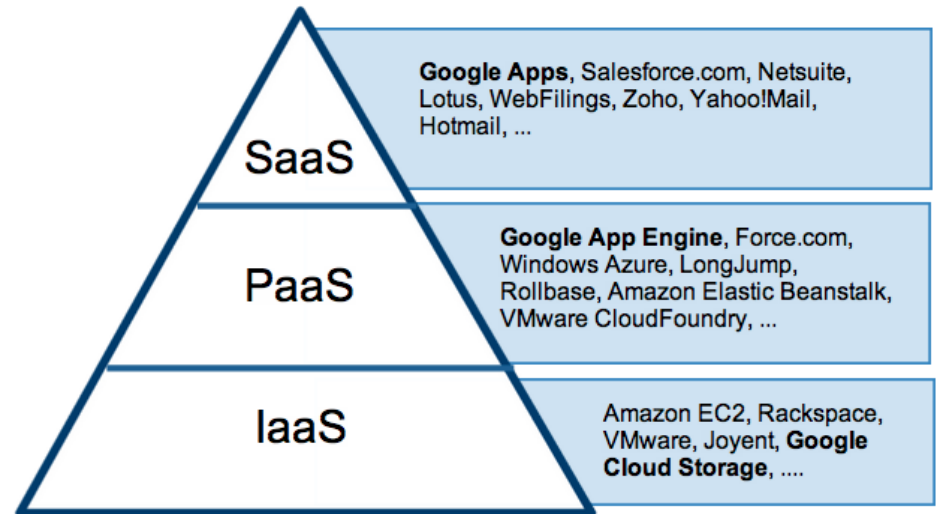
- *Project management*
- *Customer Relationship Management (CRM)*
- *Human Resources (HR)*

### – Platform as a Service (PaaS)

- *Database*
- *Development and Testing*
- *Business Intelligence*

### – Infrastructure as a Service (IaaS)

- *Backup and Recovery*
- *Storage*
- *Computation*



# Cloud Computing Benefits

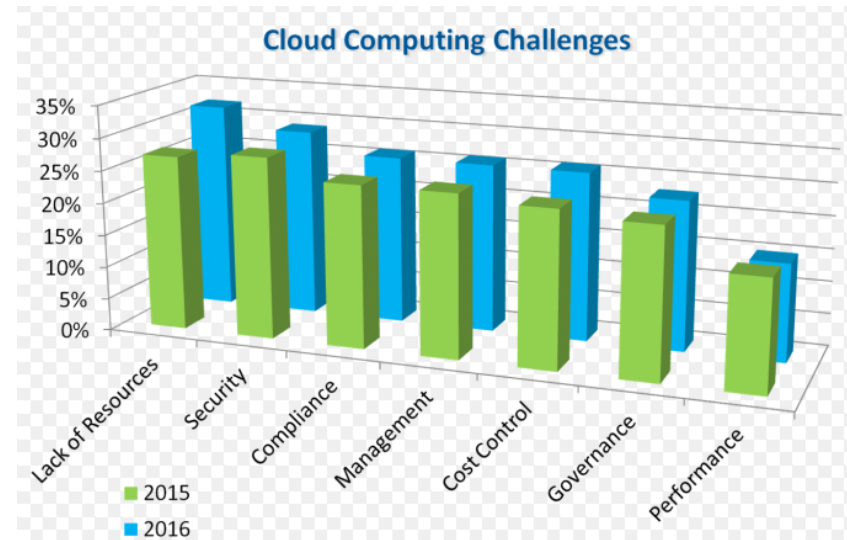
- Cost savings
- Agility
- Flexibility
- Scalability
- On-Demand Availability
- Data Centralization
- Portability
- Environmental
- Increased Innovation
- Disaster Relief



[Source: KPMG Cloud Survey Report 2014](#)

# Could Computing Risks and Challenges

- Lack of Available Resources
- Security
- Compliance
- Reliability/Performance
- Data Governance
- Legal Concerns
- Cost Management
- Vendor Lock In
- Loss of Control



[Source : Corpinfo Blog posting 8/2016](#)

# Big Data - Defined

- Defined by Forrester as “the techniques and technologies that make capturing value from data at extreme scales economical”
- Wikipedia defines it as “a collection of data sets so large and complex that it becomes difficult to process using on-hand database management tools or traditional data processing applications”
- Big data is defined by the following characteristics:
  - Rapid data accumulation
  - High data volumes
  - Diversity in types of data and data sources
- Or in plain language – number crunching of epic proportion accomplishing in minutes what may have taken weeks several years ago



# Big Data & the Cloud?

- The notion of Big Data can exist without cloud computing
- The question is whether the notion of Big Data would have been conceived without the cloud
- We share our thoughts, pictures, videos, etc. through cloud based apps such as Facebook, Twitter, Google+, etc.
- According to a Survey conducted by comScore and UPS for Forbes magazine – in 2016 more than 50% of purchases were made online than in stores



# Big Data & the Cloud?

- Google processes over 40,000 searches per second, 3.5 B per day, 1.2 T per year worldwide
- All these applications exist in the cloud and their providers take Orwellian interest in every transaction and query that is made
- This is how Facebook knows who to recommend as your friend and how Amazon knows what books you might like to read or what movies you want to watch



# Big Data & Cloud



- Cloud computing is an enabling technology for Big Data
- Add to this the vast amount of data collected from other applications and devices that collect and transmit data
- This data is collected in many formats... text, video, still, audio, sensor reading, GPS coordinates, radio frequency identification readers(RIF), etc. – all thrown into the same pot
- Big Data is the tools and techniques that make it possible to process these large amounts of data in varying formats
- Full circle back to the cloud - where else can we find nearly unlimited access to storage and processing power?



# Big Data in Practice

- Potential uses include:
  - Threat detection
  - Battlefield analysis
  - Business intelligence
  - Machine reading
  - Medical research
  - Health care monitoring
  - And the list goes on.....
- Not to say we can't already do these things... Big Data opens the door to do them more efficiently and effectively



# Big Data in Practice

- The National Science Foundation along with 10 other research funding organizations are currently hosting round 4 of the Digging into Data challenge
- Some of the projects that Round 4 awardees are working on include...
  - Trees and Tweets: Mining Billions to Understand Human Migration and Regional Linguistic Variations
  - Analysis of Twitter data to identify how language patterns vary by regions and genders
    - *Use of uh vs Um*
    - *Alternation between Haha and LOL*
    - *Used visualization techniques to understand speech patterns across the US*



# Big Data in Practice

- Some of the projects that Round 4 awardees are working on include...
  - An Epidemiology of Information: Data Mining the 1918 Influenza Epidemic
    - *Analysis of digitized historic newspapers*
    - *Studied the spread of the flu through US and Canada*
    - *The project sought to gain understanding of how newspapers shaped public opinion about the 1918 pandemic*
- While very academic – these examples provide a window into the kinds of research Big Data makes possible
- More can be found at the Digging into Data Challenge website [www.diggingintodata.org](http://www.diggingintodata.org)



# Big Data in Practice



- A few more practical examples regarding the US DoDs applications of Big Data
  - Situational Awareness and Visualization – including Geospatial Intelligence (GEOINT) which exploits and analyzes geospatial information to improve situational awareness and decision making
    - *Imagery*
    - *Terrain data*
    - *Real time data*
  - Enterprise Cyber Security Data Management - distinguishing computer attacks and data threats from normal computer and network activity. Analyzing current activity based on information from:
    - *Computer error logs*
    - *Network attack information*
    - *Data from organizations that have been violated*
    - *Vulnerability reports*

# Big Data in Practice- Starbucks

- Starbucks will often expand in an area already saturated with Starbucks locations
- They use big data analytics to determine whether a particular location will be successful
- They do an analysis of success rate of a new location in terms of:
  - Location based data
  - Traffic data for the area
  - Demographic data
  - Customer based data



# Big Data in Practice

- A few more interesting applications
  - iPhones ResearchKit - turns your phone into a biomedical research device – tracking steps or asking questions about an illness for real time data collection on health
  - WeatherSignal – taps into sensors in the phone (barometer, hygrometer, ambient thermometer, light meter) to collect weather related data to feed into predictive models
  - Route (outdoor marketing company) uses big data to determine where best to advertise. Analysis of data from GPS, eye-tracking software and traffic patterns to determine where an advertisement will be most effective
  - Long Beach, CA uses smart water meters to detect illegal watering in real time and help homeowners cut consumption by as much as 80%



# Conclusion

- Cloud computing allows organizations to offload their IT resource requirements and responsibilities to outside providers
- In so doing, it creates an environment where huge amounts of data can be collected, stored and processed.
- Big Data is the notion of doing useful things with this data in an efficient manner. Cloud computing enables this notion
- Big Data has the potential to change the world through making it possible for researchers to efficiently and effectively comb through mountains of heterogeneous data
- Final Big Data Example - Discovering how children learn language

# Questions



[Arlene.minkiewicz@pricesystems.com](mailto:Arlene.minkiewicz@pricesystems.com)