Cloud Computing *for* Media People

Media Dojo
www.media-dojo.com

Volume 1.0 | Winter 2010

Preface

Welcome to the Media Dojo Guide to Cloud Computing for Media People.

The Guide translates cloud computing for business level media and marketing executives. It surveys some of the foundation concepts, technologies, players, and potential impact of cloud computing on the media and marketing industries.

Think of this document like a travel guide for media people to a country called the Cloud. The basic concepts and technologies comprise the "Getting There and Getting Around" section. The three main service categories can be thought of as "Hotel and Lodging". When I lay out how I see cloud computing affecting the media industry, that's the "Eating and Drinking" section.

Some people want a travel guide to precisely plan their journey step-by-step. The guide is their bible for each and every holiday moment. Others use a travel guide for a quick and organized introduction to an interesting spot where they want to go. They consult it but quickly start looking around on their own.

I hope you'll use this Guide in the latter fashion.

Circa 2010, we are all amateurs at implementing the cloud model for media and marketing. Anyone who claims to have nailed it cold is trying to sell you something. Consequently, much of the information and/or prediction in this first volume will be obsolete within six months to a year.

Nevertheless, this is a good time to plant an initial stake in the ground about Cloud Computing's potential impact on media and marketing. Players like Amazon, Google and others have productized their distributed computing knowledge and assets into realistic commercial offers. A flood of digital media businesses are launching based on the new IT economics of cloud computing. As this process iterates and scales, the media industry is in for massive change. Smart people have little problem imagining such things.

But successful people will be those who get the timing right. They already know the points of leverage in their organizations, their markets and their industries. What they need is to understand some of the core moving parts of a new model for telling a story.

So welcome indeed, target reader.

John du Pre Gauntt
Media Dojo
March 2010

Foreward: The Next Lurch

Chances are the next killer business in media and marketing will be hosted on a computing cloud.

This sudden, explosive value creation probably won't advance the state of cloud technology one jot. More likely, successful media innovators will identify some powerful yet unarticulated human demand for a new type of entertainment, learning, communication and community. They will use cloud computing to serve that need.

This whitepaper examines cloud computing concepts like virtualization, multi-tenancy, metadata, IaaS, PaaS, SaaS and other cloud terms for a business level media and marketing audience. As a media or marketing professional, you need this knowledge to become literate. But cloud computing technologies and services aren't your next media business. Your next business comes from discovering a core demand just before it becomes conscious. Only with that knowledge, can cloud computing become a true innovation platform for you rather than a digital fashion statement.

All this talk about unarticulated needs sounds theoretical, right?

Climb into your time machine and go back to 1995 to ask people about their online search habits. You might find 10 people at a UNIX conference who could give you a decent answer. Good luck finding anyone who cared about search in the studios or advertising agencies in Hollywood, London or Madison Avenue.

Fast forward five years to 2000. Internet hype is deafening. The money is gushing. New companies bloom like algae. Search has become important. But the search market is largely locked up by Alta Vista, Lycos, Excite and especially Yahoo! And besides, everyone knows that push technology is the next big thing.

Fire up your time machine again and travel to 2005. What a rocky ride you took through stock market implosions, terrorism, war, and the meteor strike of Google's IPO. The 1990s are a sepia toned memory. But the search market is well established. We have a currency. We know digital is the future. And we know that bidding on keywords will get us to that future.

It must be time for social networks, mobile Internet and the real-time web to scramble things yet again. Make no mistake. The future has NEVER been a smooth march to the upper right corner of the chart.

Futures *lurch.*

When the future lurches in a different direction, most market analysts and executives assume that successful businesses will be based on meeting today's needs faster, better, and cheaper. Sure, you can make money doing that. You might even get rich.

But you won't make a killing.

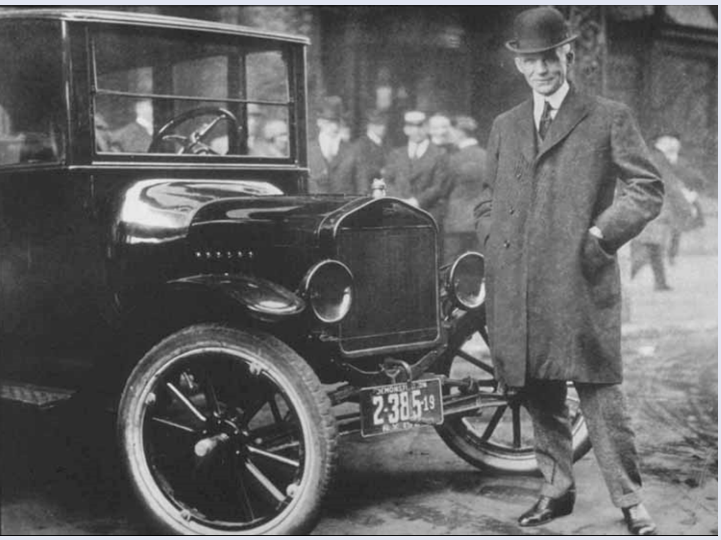


IMAGE SOURCE: www.futurelab.net

To do that, you've got to radically upset the prevailing balance of productivity and investment in a given industry. Before electricity came into the workplace, 19th century manufacturing productivity largely tracked investment in steam power and machinery. But a 20th century capitalist re-organizing work around electricity and electric machinery could realize huge efficiency gains without making a near equal corresponding investment.

And once customers understood that they really "needed" electric irons, refrigerators, transportation and power tools, it didn't matter that your water wheel, steam engine and belts were fully amortized. It didn't matter that there were still plenty of applications for traditional power sources and methods.

Try as you might, you could no longer make a killing by using steam or water power. Competition had lurched in a different direction.

I believe cloud computing will radically upset the balance of media productivity and investment over the next few years. The new organizational models and investment profiles enabled by OPEX-based, on-demand, as-needed access to computing resources will rip most media and marketing production out of the piece-work orientation that dominates today. Computing clouds will enable mass customization of media on the production side and direct-to-consumer on the distribution side. The barriers to entry will never be lower. The barriers to success will never be higher.

The most important things won't change, however. Before designing or embarking on a campaign, marketers will still need to answer who buys, why they buy and how they buy. No intelligent media creator will try to substitute technology for a compelling story, vivid characters, and unique takes on age-old human dramas. Cloud computing won't change those imperatives and thank goodness for that.

However, cloud computing will change the environment in which the media and marketing industries approach these challenges.

When you're in the midst of a lurch, the crucial test isn't engineering.

It's imagination.

Therefore, my prime directive in this work is to give business people some tools to better imagine how to make the next killing in media and marketing by using the cloud.

Good hunting.

TABLE OF CONTENTS

Story Becomes Software

You're a media person. You're good at telling stories. That's why you got into the media business in the first place. Storytelling is the only consistent advantage you'll ever have.

But how you tell stories, how you propagate stories, how you co-evolve stories with your audience, how you monetize stories, well — all of that is now up for grabs. It's up for grabs because of a flip in the media equation.

You previously put the medium — not the *media* — at the center of your business. The media was lashed to the unique assets and constraints of television, film, print, music, radio and yes, even the World Wide Web. Thirty minute situation comedies and late night variety shows work mainly when they're experienced through television. Above the fold placements of news and ads reflect the physicality and logic of newsprint. Good luck trying to draw people to the local cinema to see a 20min film, no matter how good it is. Web pages, banners, overlays, ad bugs, and other digital coinage have evolved in tandem with people clicking with a computer mouse.

But that's all changing fast. We're sorry to inform you Marshall McLuhan, but very soon the medium will no longer be the message. The message will be the **story** that finds and engages audiences who experience it across a palette of mediums and co-evolve it with the story creator. Mediums will be that which are organized around the assets and constraints of stories, not the other way around.

Consider a transmedia franchise like *Heroes*. A non-exhaustive list of the entry points to the story beyond television include web clip summaries, behind the scenes documentaries, mock news reports, dual screen episodes, polls and quizzes, console games, online games, interactive graphic novels, comics, mobile games, text updates, trivia contests, fake web sites, blogs, Facebook pages…the list goes on.

In its ideal form, this type of storytelling allows each medium to reach an audience on its own terms. Each entry point enables an audience to experience a particular doorway into the world of *Heroes*. Marketers use the various doors to engage with different demographic and behavioral slices of the *Heroes* fanbase. The popularity and/or emphasis of these mediums may ebb and flow. But the franchise hangs together because each *Heroes* media experience originates from and feeds back to the mothership of the story canon and mythology developed by Tim Kring and the *Heroes* writers.

Kring and his collaborators didn't create all those entry points simply to show off how many different media formats are out there. In a visceral sense, the decision to tell the *Heroes* story across all these platforms is acknowledgment that for an on-demand media world, the game is shifting away from pushing out content for people to consume in favor of creating story-based worlds for them to dwell and interact.

Heroes and other content franchises like *Lost* rightfully deserve credit for breaking transmedia storytelling into the mainstream. That said, these properties emerged from the top-down medium of television. They created a branded narrative for television first then extended the narrative across various media platforms.

However, the next evolution of storytelling will do more than extend a narrative across multiple platforms. Next generation storytelling will directly bundle functionality and data with narrative. The story will "behave" differently, not just "display" differently, depending on the given medium and context an audience chooses. Effectively, many of the mechanics we associate with gaming will be integrated into media content. Episodes and other story elements will act more like computer objects, enabling an audience to experience the media and do something at the same time.

Moreover, media won't be restricted only to the mediums we currently associate with entertainment and information. Media will be integrated directly with enterprise computer applications, 3rd party data sources and services, new devices and new forms of social interaction. People who master this different orientation to storytelling will resemble brand managers as much as they resemble traditional content creators.

IMAGE SOURCE:HTTP://WWW.LIFEINITALY.COM/DECOR/DOORS.ASP

It is extremely likely that their story-based worlds will live on a computing cloud rather than a cable head-end, a film reel or a print run. The cloud[1] itself will be the medium, not simply because it offers better economics for hosting and transmitting digital content. Cloud-native media will be able to transact directly with the audience in real-time for almost any interaction that makes sense from a social or business point-of-view.

That's why cloud computing matters to the media industry.

1 For our purposes, "the cloud" refers to the ecosystem of cloud computing rather than a single entity. It's similar to how we use "the Internet" in conversation. And just like using a catch-all term to describe a network of networks, there will be instances where using "the cloud" as a term is inappropriate.

We're All Nerds Now

Three forces are converging that are driving more of the media business into the cloud: gadgets, mobility and multitasking.

The first factor is the sheer number of networked gadgets and services that have media access as part of their value proposition. The CES show in Las Vegas now groans with networked consumer electronics devices that expect to be filled with media and services. Whether we're talking about Amazon's Kindle, Google's Nexus 1 or any of the numerous permutations, the current conventional wisdom is that hardware and OS determine who controls and draws excessive profit from the customer relationship.

iPad currently owns the news cycle because it positions itself as a game changing device for media consumption. We won't know definitively whether that proposition pans out until after iPad launches and spends a few quarters in the market. One thing is certain, however. Apple has morphed from being a computer manufacturer to being a media products and services company. Whether iPod, iTunes, iPhone, the App Store, and now iPad, the current world leader in electronic product innovation has staked its future on creating ecosystems based on the triangulation of product design, media and mobility.

The Outside World is the Desktop

One area where narrative form meets IT function and personal data involves mobile location based services. Companies such as Foursquare, Gowalla and MyTown, plus mobile augmented reality players like Layar, Mobilizy and acrossair are turning the outside physical world into an interactive social desktop for media.

In the case of Foursquare, end-users receive points for broadcasting the physical places they visit through "checking in" via the Foursquare mobile application. Each time a user checks into a location, they earn points and unlock digital candy in the form of virtual badges. Users can check in anywhere as well as adding new locations to the Foursquare system. As users check-in, their Foursquare social contacts can see where they are. Foursquare users can also send location status updates to other social networks such as Facebook or Twitter. If the user checks into a given spot more than any other Foursquare user, they can become "mayor" of that particular location.

Media companies have started to see value to including their content with the Foursquare service. One of Canada's free daily newspapers, *Metro News*, offers restaurant reviews, city tips, as well as stories that Foursquare users can discover as they encounter physical points-of-interest. Foursquare has created a custom badge for *Metro News*, which tries to get its readers to check-in each time they pick up a copy of the print publication. *Metro News* plans to introduce special deals for Foursquare "mayors" with each Friday edition. The combined *Metro News*/Foursquare offer gives businesses a unique ad buy which specifically targets repeat customers who are likely to act as influencers on other potential customers. Other Foursquare forays into combined print/location offers include a deal with *Explore Chicago* that mimics the *Metro News* arrangement as well as a tie-up with ZAGAT, which will introduce a "foodie" badge to reward readers who check in at ZAGAT reviewed establishments.

IMAGE SOURCE: *http://www.core77.com/gallery/images/ces2010-03.jpg*

Foursquare extended its direct tie-ups with media companies in February 2010 by announcing a deal with Bravo! Television. In this case, Bravo! is offering badges and prizes to viewers who check-in at some 500 Bravo-tagged locations around the United States. The Bravo! badges are tied to content properties such as *Top Chef, The Real Housewives, Top Chef Masters* and *Shear Genius* to name a few. The network will create Foursquare tie-ins to Bravo! advertisers which could be coupons or other incentives related to the various themed virtual badges and shows.

HBO is adapting the Bravo! model to help launch new shows like *How to Make It in America*, which debuted February 14, 2010. According to the promo, "Hustle your way into the NYC scene with HBO's *How to Make It in America*. Unlock one or all four of the badges: Culture, Living, Cocktails, and Nightlife. Need some tips? Check out these sites, Blackbook, Eater, Flavorpill and Urban Daddy. Then tune-in to HBO on Sundays at 10pm to follow Ben and Cam, two enterprising twentysomethings as they hustle their way through New York City, determined to achieve the American Dream". According to *Ad Age*, the goal of the tie-up is to use content from the show to help audience members live vicariously through the characters and their lifestyles.[2]

The current dance between media companies and GPS providers is a glimpse into a new vein of storytelling that dispenses with the idea of a passive audience. Innovation of plot, character and setting is now augmented with innovation focused on specific audience behaviors that can be tied to the dramatic arc of a story. But we've done this before. It's called live theater. However, the theater setting is expanding beyond wooden or digital structures to become the physical world itself. Shakespeare was ahead of himself in more ways than one when he wrote in *As You Like It*, "all the world is a stage and all the men and women merely players…"

2 http://adage.com/digitalnext/post?article_id=141977

The mobility factor arguably exerts a bigger impact on the media industry than Apple itself. Cisco released its latest Visual Networking Index[3] in February 2010 to the collective gasps of mobile network operators. According to the networking giant, the amount of global mobile data traffic in 2009 increased 160% to reach 90 petabytes per month, which is roughly the same as 23 million DVDs. Cisco went on to say that mobile data traffic world-wide is growing nearly 2.5 times faster than fixed broadband, with smartphones and various types of air cards driving more than 90% of mobile data traffic world-wide by 2014. Most important, Cisco expects that by 2014 (4 years from now), more than 400 million of the world's Internet users will access it predominately or solely through a mobile connection.

The upshot for media providers from the Cisco VNI report is that mobility is being baked into daily life for nearly all Internet users from the triad economies to emerging media markets. Moreover, with the veritable algae bloom of e-readers, netbooks, networked toys and other consumer electronics expected to hook into these mobile data networks, ambient access to interactive media will be expected by most market segments.

The use and re-use of media assets is no longer a strategy, but a requirement for media companies to stay relevant.

The third factor that will accelerate media industry adoption of cloud computing is multitasking, which received a powerful validation in January 2010 with release of new research by the Kaiser Foundation[4]. The Kaiser media survey revealed that even though today's 8-18 year-olds devote an average of 7 hours 38 minutes to entertainment media across a typical day, their proclivity to multi-task (simultaneously using more than one medium) bumps that figure up to the equivalent of 10 hours 45 minutes worth of media crammed into those chronological 7.5 hrs of daily exposure.

Peter Hirshberg[5] from The Conversation Group noted two main outcomes from the Kaiser findings. The first is that there is no longer such a thing as being offline for an increasingly large segment of the population. Screens, media and interactivity follow kids and young adults through their days as surely as electric lighting. The second outcome is that only in a very few, specific instances can we assume that an audience member is engaged with only one medium. Multitasking and partial attention is the basic reality for trying to engage these audience segments.

Given the explosion of networked media devices plus a mobile population that regularly multitasks, the use and re-use of media assets is no longer a strategy but is a requirement in order for media companies to stay relevant. Media providers and their partners are racing to keep up with their audiences by innovating the story to fit with their lifestyle choices surrounding media. This calls for the rapid-fire ability to mix and remix media assets, audience insights, and applications across devices and user contexts. The craft-based tradition of media and marketing organizations creating their content *sui generis* is about to get a major overhaul[6]. Mash-ups of content assets will become the norm rather than the exception in a world of many devices pulling on ambient media. The cloud has started to become the platform for precisely that sort of innovation.

3 http://www.cisco.com/en/US/solutions/collateral/ns341/ns525/ns537/ns705/ns827/white_paper_c11-520862.html

4 http://www.kff.org/entmedia/mh012010pkg.cfm

5 A great tandem presentation with Tim Kring at DLD 10 in Munich [http://www.dld-conference.com/video.php].

6 Suffice to say that I don't see mashing up media assets as a good/bad, right/wrong issue, but simply a fact of life.

The Cloud as an Innovation Platform

We've been re-using media assets for a long time in the DVD market. Media providers realized that DVDs enabled them to re-purpose a lot of material from the cutting room floor in the form of alternate endings, deleted scenes, documentaries and other collateral. The cloud offers media providers the equivalent of all those DVD bonus materials within a network-hosted story. The difference, however, between bonus content on a DVD disk with bonus content in the cloud is that the latter lends itself to being integrated with services and special offers that can be tailored to the individual viewer or their particular audience segment.

This works not simply because the cloud enables quick integration of applications and services, but also because many of the targeted media distribution channels such as Facebook and MySpace are cloud-based as well. This makes it far more simple, fast and inexpensive to combine applications and distribution in one package. Moreover, some of the most valuable combinations are at the B2B level. A media provider or one of their marketing partners might want to attach Saleforce.com to a social networking widget that is carrying a content payload. Integrating Salesforce.com with another application to run on Facebook or MySpace likely would be time or cost-prohibitive if you had to implement the solution with on-premise software and infrastructure.

The cloud also enables innovation by making media itself to become more "clickable" in real-time.

Using the cloud for integrating functionality, media and distribution in what feels like a single environment enables media creators to start looking at almost any enhancement to a story world as an application. Take communications. Currently, many media properties only offer email, SMS or registrations as a return path for the audience to communicate directly with the originators. However, it might be that for whatever reason, a media property or one of its marketing partners wants to exchange phone calls with its audience.

Twilio is a Seattle-based company that uses the cloud to turn voice and SMS into modular applications that can be baked into most any web application by web developers rather than telecom specialists. Before Twilio, the developer would need to understand telecom programming languages, set up a stack of PBX software or bring in an expensive consultant if they wanted to interact with callers or texters. However, by using the Amazon cloud, Twilio puts the equivalent of a phone network in the sky that a developer can hook into their apps. The app is then paid for by how it's being used by the audience. This model allows a media developer to use their existing web skills, code base, servers and databases to "just add voice" or "just add text" to whatever can enhance the value of a story.

The cloud also enables innovation by making media itself to become far more "clickable" in real-time. Take online video, for example. There are a clutch of cloud-oriented players out there like Ooyala, Veeple, Overlay.TV, as well as familiar online video companies like Brightcove, that are finding points in online video streams that lend themselves to direct interaction by the audience. Perhaps a member of the audience is watching *Swordfish* online and is attracted to a hat worn by actress Halle Berry. They want to find out about the hat and where to

possibly buy it. The hat is tagged or painted in the video stream to draw attention by the user, who clicks on it to pull in information or travel out to another site where they can take action or purchase it.

Making online video interactive is a several step process. First, the media provider needs to populate the video stream with targets for people to click or skip if they want. Those targets need to be made discoverable by the audience (perhaps through a specific icon, a key word box or another interface design). The discoverable targets then need to be linked to whatever interactive experience the media provider wishes to offer. And it's not always an e-commerce experience. In Q1 09, Overlay.TV worked with the teen band The Jonas Brother to offer tools that let online fans sing along and record themselves with the band's new single, "Lovebug". According to the CEO Rob Lane of Overlay.TV, 54% of the people who visited Jonasbrothers.com during a 30 day period created their own karaoke video using the toolset. Fans had the option of sharing their videos across various social sites.

IMAGE SOURCE: www.overlay.tv

Whether the video programmer is linking out to commerce or enabling the audience to integrate their expression with the content, the important point is that it's entirely up to the user regarding when and how they want to engage with the media. Just watch the video and when you see something you like, you take action.

Flexibility at this stage is fundamental because interactive video doesn't yet work as an out-of-the-box solution. What will trigger user engagement in a fashion show is different than a sports event or a comedy or a drama or even an instructional video. So it's important for content owners to be sophisticated about envisioning how interactivity can enrich the user's experience, and by extension, lead to better chances at monetization. What happens, the different calls to action, and what is engaging depends very much on the content itself and how it is programmed.

This last point of a new approach to editorial programming is often overlooked in the rush to make everything clickable and see what sticks. Historically, content providers organized their stories around interruptive advertising. A 30min television time slot was diced into story and ad pods. It may be true that a cliffhanger or sudden plot turn advanced the story. But it also created the opportunity for the advertiser to address the audience. The more interesting aspect of interactive video going forward is that editorial programming might shift subtly to focus on placement, flow and discoverability of targets within a content stream that can entice the audience to interact.

Cloud computing offers storytellers both a canvas and functional palette of options to allow a story to engage its audience as deep or as shallow as both parties want. There are times when a member of the audience is dog tired from a long day and just wants a simple, lean-back experience of the story straight up with no interactive chaser thank you very much. Other contexts might find the same audience

member eager to engage at far deeper levels because they've developed a passion for dwelling with the story world. They may even want to add part of their own expression and share it. The point of cloud computing is that the media provider now can offer such options using a single chassis rather than needing to build multiple versions of the story.

However, such flexibility only works if the media provider has detailed knowledge of the audience. This brings out the other side of cloud-native media, in which the cloud becomes not simply an innovation platform creating content, it becomes a business platform for commercializing that content.

Competing on Data

Data dominance is the new high ground for media competition.

Media professionals since day one have had an inventory and supply chain management problem similar to airlines and seat pricing. The value of the asset like an ad impression or a seat changes rapidly. The asset value changes according to who wants it, changes according to how it's bundled with other offers. It's a whirl of activity right up to the point when the plane takes off or the ad impression is served. In both instances, the value of any unsold asset plunges to zero.

Data dominance is the new high ground for media competition.

Savvy operators in both media and air travel have long known what they wanted to do to crack this problem: take raw data in the form of server logs or ticket price searches; enhance it with other data including that from 3rd parties; refine the data into information through an analytics package; process that information into insight about opportunities to maximize revenue at the margin; and then take action based on that insight.

It's a relatively straight forward process that, technically speaking, requires a **shitload** of data, storage, processing, and reporting resources. Through the 90s and well into the first part of the 2000s, these problems were basically unsolvable with the computing resources and capabilities of the time. The volumes of data that needed to be processed, the analytics people wanted to run, the kind of insight that people wanted to extract from the data was difficult to obtain cost-effectively.

Now a slew of data optimization and analytics companies like MediaMath, Pubmatic, Peer 39, YieldEx and others are applying computer science and new capabilities offered by cloud computing to advance the state of the art for media publishers and marketers.

Consider an impression coming in that displays as male 25-34 on the US west coast looking at sports content. There's probably a marketer wanting to buy sports content opportunities, another wants to buy exposure to young males (no pun), still others want to buy US west coast. So there's an optimization challenge to figure out how would a media publisher maximize revenue given a set of X impressions

that expire at Y rate in the context of Z demand by different marketers. Bump that volume up to tens of millions of impressions served monthly and it becomes reasonably clear that without serious data and analytic horsepower, a lot of money and ad performance is being left on the table. Moreover, the proliferation of social platforms like Facebook, new devices like iPhones and Kindles, plus new calls-to-action have mushroomed the amount of data needing to be analyzed by many orders of magnitude.

Media providers will become data sources and sellers in their own right.

If anything, media providers will become data sources and sellers in their own right by combining their impression data with 3rd party impression data to distill even more accurate profiles of their audiences. This emphasis isn't data for its own sake. We can expect a lot of laundry to be aired over the next few years as the fabled "click" becomes even more of a football to be kicked around by publishers, media buyers and search providers. For most of the online media world's existence, the only consistent, measurable currency has been a click, or more precisely, the last click before an end-user gets to a transaction with a marketer. For a number of marketing objectives, especially for branding, clicks are actually a bad measure.

What media providers want, and what they might be able to get with better data management and analysis, is attribution tracking. Attribution is knowing, "sure the guy went on Google and searched for a product. But he also saw two banners, played 2 related games, saw a video beforehand and consulted his friends on Facebook. He wouldn't have known what to search for if he hadn't done that."

How does a media provider get credit for those higher level marketing activities?

He won't get credit unless he can prove his media had an impact.

And that requires data.

Silver Buckshot Trumps Silver Bullets

All this talk about transmedia storytelling, innovation platforms and competing on data probably sounds like yet another heaping platter of consultant jargon about cloud computing. But it all translates into something rather simple for media providers: if you've managed to catch lightning in a bottle in the form of a compelling story, you can't base your business anymore just on selling copies that story or restricting access to it.

The story needs to become a world in which the audience can dwell and participate; a world that provides multiple avenues for monetization; and a world in which competitive advantage involves how fast and how well you can customize media or marketing at the margin. Mass customization is the enduring reality for media going forward.

In the online video space alone, over the next few years we're looking at 10-30X increase in video content being served across all devices, even though there will be only a 2-3X increase in the number of online video viewers. So what users are doing with your video becomes a more important metric to track than just who is being exposed to your video.

This switch from exposure to activity gains flesh once dynamic pricing of interactive content is introduced, either as a direct-to-consumer or direct-to-sponsor proposition. Should a given content stream price at $0.02 or $0.05? Well, the answer depends on the target user, their specific context, and which ad inventory is available at that exact moment. Most everyone in the media business will have access to the same ad inventory. But the winners will be those who make cream out of the last marginal 15% of possible revenue.

IMAGE SOURCE: www.futurelab.net

Such a level of execution requires a media organization that can learn and take action fast enough to be in the right place with a good experience right when the best people want it most. Organizing media assets in the cloud can be an important first step. But it's not a panacea. The biggest change for media organizations to internalize is that in a world of unlimited media shelf space, you can no longer base a media business on carving out and isolating readers, listeners, viewers or game players.

Instead, tomorrow's leaders will use the cloud to serve media customers.

They are King.

Cloud Computing 101

There are conditions in America which necessitate the use of such instruments more than here. Here we have a superabundance of messengers, errand boys and things of that kind…The absence of servants has compelled America to adopt communications systems for domestic purposes

– SIR WILLIAM PREECE, CHIEF ENGINEER FOR THE BRITISH POST OFFICE 1878

Accept the fact right now that your technical staff will probably hate the following descriptions of cloud computing. They will tell you that a huge amount of important detail has been left out.

Guilty as charged.

The aim of the following section is to combine metaphors with some formal description to allow business layer people to grasp the basic principles and elements of cloud computing then move on.

That's not a license for the business side of media and marketing shops to forget about the technology. Many fired media and marketing executives happily ignored Internet until they were leapfrogged by protocols and technologies surrounding IP, TCP, DNS, SMTP, and HTML. Technology matters.

You don't need to be an engineer but you should at least become literate in the basic moving parts so that when an engineer purses their lips and says to your great new idea, "That's not trivial", you get a decent handle whether s/he should be believed.

So think of the following section as 101 level cloud computing as opposed to cloud computing for dummies.

Key Concepts:

- A working definition of cloud computing
- 3 rules of thumb and one Iron Law for using cloud computing effectively
- The basic cloud technology stack
- Types of delivery models for cloud computing
- Types of computing clouds

Why it's important:

- This is the conceptual bucket for the technology stack and the cloud vendor ecosystem

Cloud Computing Technical Overview

Let's start by imagining actual clouds.

Scientifically, you have condensed water vapor or other fine particles (dust) suspended in the air. Descriptively, a cloud is a soft, billowy mass that floats and changes shape according to any number of reasons. We also use "cloud" to shroud something (mist clouded the hills). Metaphorically, a cloud can be something that obscures (clouded judgment), or fills a space with darkness, gloom or suspicion (walking under a cloud).

Are you surprised, then, that the penultimate cloud computing definition remains elusive?

For now, split the world into cloud computing services and cloud computing technologies. Most of us are already quite experienced with cloud computing services whether we know it or not. If you've ever accessed a social network, YouTube, or a search engine with your PC or Internet capable mobile phone, you've engaged with a cloud computing service in one way or another. Here's a sample of some typical cloud computing behaviors:

- Uploading a photo or video to a social profile
- Backing up documents, mail, sound or image files continuously and securely via multiple devices
- Syncing data across multiple devices from a server or portal living in the cloud
- Accessing a multi-player game via a browser interface

Each of the previous examples involves consumption of data storage, processing, applications and services delivered over the Internet — the first case of cloud computing services.

The aim of this section is to look at the second case — cloud computing technologies. These are the assets that ultimately enable the service. As a media or marketing organization, you may or may not carry these technologies as assets on your books. It depends on your individual business. More important is that you understand and can roughly classify these technologies in order to decide what you keep in-house and what you procure as a service from a cloud computing vendor.

For our purposes, cloud computing is on-demand, as-needed access to computer processing, storage and networking resources/services on a pay-as-you-go basis.

For our purposes, cloud computing is on-demand, as-needed access to computer processing, storage and networking resources/services on a pay-as-you-go basis.

Already, there are massive semantic wars by academics, vendors, consultants and many others to craft the "standard" definition of cloud computing. Some focus their definition on the technology attributes of cloud computing. Others concentrate on the economics or the commercial aspects. Both approaches are valid.

Take a look at the following crude comparison of traditional IT infrastructure versus cloud computing:

TRADITIONAL IT INFRASTRUCTURE	CLOUD COMPUTING
File Server	Google Docs
MS Outlook, Apple Mail	Gmail, Yahoo!, MSN
SAP/Oracle/Siebel CRM	SalesForce.com
MS Office/Lotus Notes	Google Apps
Quicken/Oracle Financials	Intacct/Netsuite
Off-site data backup	Amazon S3
Servers, Racks, Firewall	Amazon EC2, GoGrid, Mosso, Joyent

Cloud Computing Architectures, Reese George, O'Reilly Media 2009

Circa 2010, you're not going to see which one of these approaches ends up owning the definition. Far more track record needs to be established by cloud vendors and cloud customers operating at scale. Until such operating data starts to provide unambiguous proof of success or failure, buy-side people are likely to do better with rules of thumb about cloud computing rather than some all-singing all-dancing definition.

Even though we can't yet nail down the technical and business aspects of cloud computing in one definition, the core questions that cloud computing customers need to answer aren't terribly different from any other business decision:

- What am I buying?
- What am I being promised?
- Am I getting a good deal?

If you combine our working definition of cloud computing above with the following three rules of thumb and one Iron Law, the result isn't intellectually pretty. But it's actionable in the here and now for analyzing proposals coming across your desk or developing a cloud computing strategy for your media shop:

RULE 1: Unless you can whip out a credit card or a purchase order and be up and running in 10min or less, it's not cloud computing.

COROLLARY: Cloud computing requires little if any human intervention to set up.

RULE 2: Unless you can release computing resources as easy as you accessed them, it's not cloud computing.

COROLLARY: You are buying a service, not assets. Any cloud computing contract that suggests some kind of ownership of a computing asset should be suspect.

RULE 3: Unless you're paying only for what you're using, it's not cloud computing.

COROLLARY: A cloud computing vendor should be able to produce an accounting of your usage and accumulated cost that's as granular as a phone bill.

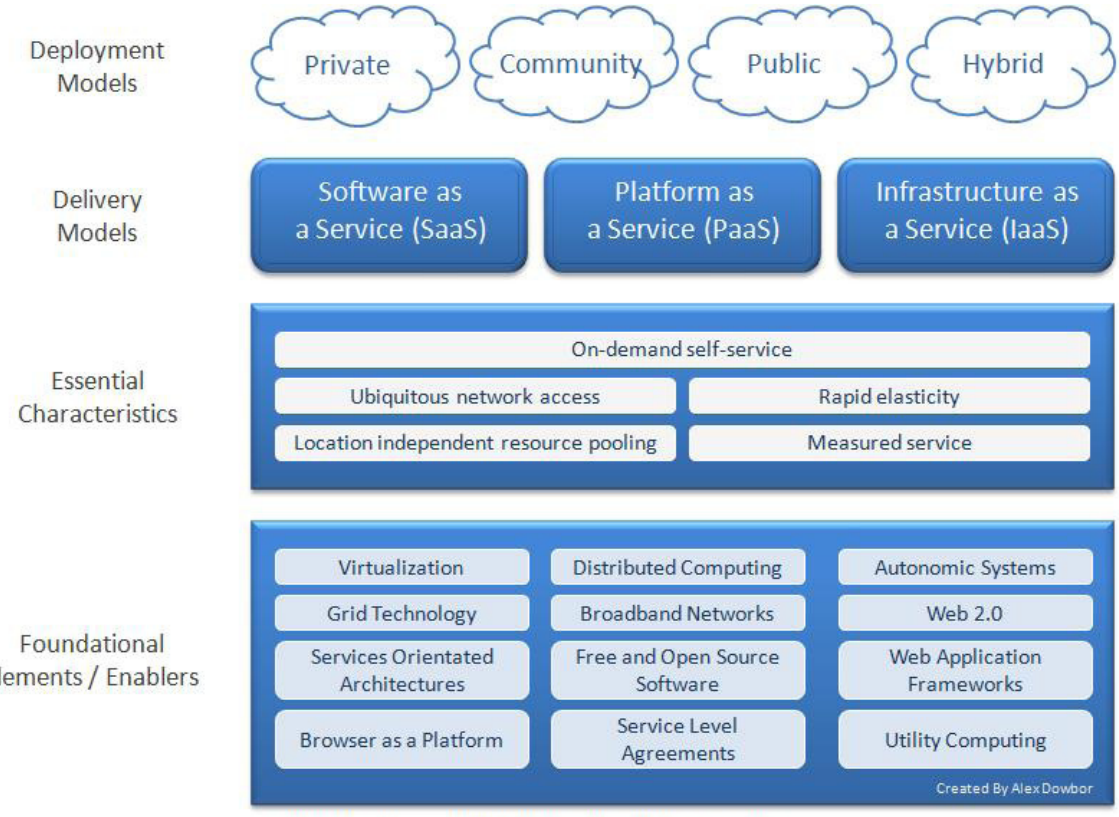
THE IRON LAW: At the outset of your voyage into cloud computing, more than half of anything proposed to you probably does NOT apply to the individual circumstances of your organization or is NOT an immediate priority.

The Iron Law will keep you out of the most trouble. Don't forget that we're still in the missionary stage of cloud computing. The cloud vendors and consultants are fired up with the spirit and eager to show you the mountaintop. This doesn't mean that cloud computing is false or that these people are dishonest. It simply means that believing passionately in a new model is important but not always the best way to deliver economic value to the customer.

With that in mind, it's time to check out what's under the hood.

Conceptual Model of Cloud Computing

In June 2009, the U.S. Government's National Institute of Standards and Technology (NIST) published a RFP that provided one of the better overviews of cloud computing. The following diagram was derived from the NIST spec by the Cloud Computing Use Cases Group [1]. There are a lot of boxes and layers. Don't worry so much about the lower layers. As media or marketing executives, you will play at the top two layers of delivery and deployment. When you contract with a cloud provider, those are the two sections that contain most of the action. At the same time, your technical staff needs to understand how the lower level parts enable the service.



Based on: *NIST Working Definition of Cloud Computing v14 and NIST Presentation on Effectively and Securely Using the Cloud Computing Paradigm v22*
http://www.csrc.nist.gov/groups/SNS/cloud-computing/index.html

1. http://groups.google.com/group/cloud-computing-use-cases?hl=en

Like a traditional media plot, we will start in the middle. According to NIST, there are five essential characteristics that differentiate a cloud-based computing environment from a networked computing environment:

- **Rapid Elasticity:** The customer has the ability to scale resources both up and down as needed.
- **Measured Service:** Performance is the responsibility of the cloud provider, who has control of billing, access control, optimizing resources, capacity planning and other tasks.
- **On-Demand Self-Service:** The customer can use the cloud service as needed without any human interaction with the cloud provider.
- **Ubiquitous Network Access:** The cloud provider's capabilities are accessible over the network by any client device that can access the web.
- **Location-Independent Resource Pooling:** Physical and virtual resources are assigned and reassigned according to customer demand. The location of physical resources underneath the cloud infrastructure is not known by the customer, and can change dynamically.

Think of the preceding attributes more as the goals or purpose of a cloud computing solution rather than what the feature set should look like. Vendors are going to be twiddling the knobs on any one of those goals to differentiate themselves from other vendors. Good for them. When you're looking at the needs of your business, you're likely to weight one or more of these five goals as more important than the others.

Moving up the stack, there are three primary delivery models for cloud computing: Infrastructure as a Service (IaaS), Platform as a Service (PaaS), and Software as a Service (SaaS). By definition, once you get above those levels, you're a 100% service customer anyway.

Before you learn about individual elements making up each category, it's important to understand the primary focus of each category:

IaaS: the physical hardware environment of computing

PaaS: the operating and developments environments for applications

SaaS: software enabled business processes

Of course, there are cloud offers that cut across layers like SalesForce.com, which can be both PaaS and SaaS. The space is evolving rapidly so don't think the above distinction is carved in stone. Still, you can have a more efficient preliminary conversation with your internal staff and vendor

candidates if you perform a loose mapping between what you want to do (e.g. saving hardware costs, faster innovation, automating business processes) with the main layers of cloud computing.

Let's take a quick look at the attributes of each layer. Starting from the base layer, NIST defines Infrastructure-as-a-Service (**IaaS**) as customer control of "fundamental computing resources" such as processing, storage, networking or middleware. The customer controls the operating system (OS), storage, deployed apps and even networking components like firewalls or load balancers, but not the physical cloud infrastructure below them. Think of IaaS as you having responsibility for almost everything but the hardware. You've got the most control. But you've also got the most management overhead. IaaS is the realm of a dedicated cloud provider like Amazon Web Services (AWS), Joyent, and Rackspace Cloud among others.

Moving up the stack, you have Platform-as-a-Service (**PaaS**). PaaS lets the customer access a hosting environment for developing and/or deploying their applications. The customer controls how an app is built and/or deployed but doesn't control the OS, hardware or network infrastructure. The platform is an application framework like Google's App Engine. Another example would be Force.com, through which developers and some customers can create their own Salesforce.com applications. Think of a PaaS solution as both a development and operating environment for your apps.

				CUSTOMER IMPLICATIONS
Software as a Service		Google Apps	salesforce	+ Application logic, platform and infrastructure abstracted + Significant reduction in effort to deploy, run and manage - Apps can be configured but may not meet highly customized requirements
Platform as a Service			force	+ Platform & infrastructure abstracted + Custom apps can be built an order of magnitude more quickly and cheaply - Custom apps still need to be supported and managed
Infrastructure as a Service	amazon web services			+ Physical infrastructure abstracted + Can be scaled up & down as needed - Needs to be provisioned/managed - Higher levels of stack still need to be managed, maintained and supported

When you migrate further up the stack to Software as a Service (**SaaS**), you start getting into the mainstream understanding of cloud computing. In a SaaS environment, the customer doesn't own the infrastructure or the application. They own their data and perhaps some of the permitted modifications or customizations made to the app. SalesForce.com is an example of SaaS. As a customer, you upload your data and manipulate it. But you do so under the rules set by the application, which is designed to scale across many different customers.

Finally, NIST documents four main deployment models for cloud computing:

- **PUBLIC CLOUD:** This is a cloud service available to clients from a 3rd party service provider via the Internet. All legal and contractually permitted uses of the cloud infrastructure are governed by standard terms and conditions. AWS, Joyent, Rackspace, Micrsosoft and Azure are examples of public cloud providers.
- **PRIVATE CLOUD:** a cloud service existing behind an organizational and technical firewall. Access is restricted by various policies and permitted uses are defined by the cloud owner. Typically, these are enterprise-level efforts to bring the economics of cloud computing into a single organization.
- **COMMUNITY CLOUD:** A cloud infrastructure controlled and used by a group of organizations that have shared interests or a common mission. The members of the community share access to the data and applications in the cloud. Examples here might be government agencies that pool their infrastructure resources to gain economies of scale.
- **HYBRID CLOUD:** This is where public and private clouds interoperate. Users typically outsource non-business critical information and processing to the public cloud, while keeping mission-critical services and data under their control. Another instance might be a private cloud owner adding extra capacity (e.g. "cloud bursting") to handle a seasonal spike in demand.

Don't think that the previous model is the final word just because a government agency like NIST has organized cloud computing in this manner. Check back to your mid-90s government reports on the nascent Internet and you will understand that the above framework will have a severely limited shelf life. Still, it's a decent starting point. In that sense, NIST has accomplished its mission.

To recapitulate, here's the basic breakdown of cloud computing:

- **WORKING DEFINITION:** cloud computing involves on-demand, as-needed computer processing, data storage and networking resources/services on a pay-as-you-go basis.
- **ESSENTIAL CHARACTERISTICS:** access/release computing resources as needed, pay only for what you use, no upfront capital cost to the customer
- **DELIVERY MODELS:** Infrastructure as a Service (IaaS), Platform as a Service (PaaS), Software as a Service (SaaS)
- **DEPLOYMENT MODELS:** public cloud, private cloud, community cloud, hybrid cloud
- **RULES OF THUMB:** shouldn't need to own or manage any hard computing asset, billing should be almost 100% usage based and granular

IRON LAW: throwing all of your business into the cloud is a BAD idea until you've done a massive amount of homework.

You should have the tools to plow through the following sections, which will drill deeper into the cloud computing stack. The first section will be a primer on virtualization. From there, we will explore each of the three main layers — IaaS, PaaS, and SaaS.

More Power to the Shields!

What you'll find here:

- A primer on virtualization
- Types of virtualization
- Why virtualization matters to your business
- The main trade-offs with virtual environments

Why it's important:

- You can't understand anything about cloud computing without knowing about virtualization

It's happened on any number of Star Trek episodes. The *Enterprise* is being pounded by the Klingons, the Romulons or some other galactic nasty. While the bridge alarm wails, Captain Kirk or Captain Picard barks out the classic trekkie line, "More Power to the Shields!"[2]

Do you ever wonder if there are parts of the Enterprise that go dark when that happens?

Imagine Dr. McCoy getting primed for some R&R inside a 3-D immersive simulation, when all of a sudden — poof! He's back to a green screen because power and computing cycles have been diverted to protect the ship. Goddamn Klingons....

When you hear people go into excruciating technical detail about virtualization and IT flexibility, just think of Mr. Scott directing more power to the shields. The fundamental value propositions for the customer are similar:

1. A problem or an opportunity needs computing resources ASAP
2. There's no time to build or configure new physical assets. We've got to pull in resources as if switching on more lights
3. But we live in a rational universe so there are performance/cost trade-offs.

STAR TREK IMAGE: http://www.blogcdn.com/www.joystiq.com/media/2008/08/spock-chess.jpg

2 Here's a hat tip to Alan Williamson for spinning up that effective metaphor. Make sure to send your technical management to one of his Cloud Computing Boot Camps. http://www.aw20.co.uk/

4. If the problem/opportunity is pressing enough, we'll opt for performance and worry about the money later. But if it's something that's not as important, we can shut it down to cap the cost.
5. Either way, we can build up and tear down capacity as we need it. The infrastructure responds to our command without a lot of human intervention
6. When the danger/opportunity passes, we can return to our typical operating state.

The practical value of virtualization is better utilization of hardware.

So what happens technically when you virtualize computing resources?

Basically, virtualization uses a software simulation to fool an operating system (OS) like Windows, Linux etc. into "seeing" the same hardware interfaces as it would with a real, physical machine. The software simulation is a functional snapshot of physical input/output ports, data storage, network or graphics access cards that make up a server, a storage drive, a network router or any other piece of computer hardware that an OS or application uses to carry out its job.

For example, instead of directly binding individual disk or tape drives to individual physical servers, storage virtualization creates a "virtual" drive. To human programmers, applications, servers, administrators and even end-users, the virtual drive is a single address. However, the actual stored file might exist as pieces on any number of physical machines.

As customers we don't need to care where the physical storage location of the file happens to be. Storage virtualization gives us a general P.O. Box where we can throw the data. To a piece of digital content or an application or a human programmer for that matter, the virtual storage address offers exactly the same interfaces as a dedicated physical computer.

The practical value of virtualization is better utilization of hardware. Instead of individual machines running at 10-15% utilization (not uncommon) because they're "reserved" for a specific application or business process, the cloud provider can put pieces of Tom's job, pieces of Dick's job and pieces of Harry's job on the same machine or cluster of machines to bump utilization up to 75%. The math gets pretty clear at that point.

Another important aspect of virtualization is that it helps you plan for failure. How many times has the "server" gone down in your shop with the result being that your staff enjoys a longer lunch hour? Virtualization enables a cloud provider to mirror the data or processing job across multiple machines so that if any one piece of hardware fails, the job is still preserved while the end-user doesn't really see a significant performance difference.

Storage is only one flavor of virtualization among many. Here are some of the main classes of virtualization. The list is not 100% complete because like any important concept, people are plastering "virtual" on everything in order to capture marketing territory. However, if you remember these basic classes, you're in the 85-90%+ range of what's being done.

- **SERVER VIRTUALIZATION:** puts the entire server environment (OS, hardware, connectivity) onto a virtual machine so it can be run on any compatible hardware to increase utilization. Server virtualization can be run in parallel to where the same bare-metal box is running multiple OS environments.
- **STORAGE VIRTUALIZATION:** puts together multiple physical data stores (e. g. tapes, drives, disks) into one logical pool that can be spread over multiple physical machines
- **NETWORK VIRTUALIZATION:** creates logically separated communication channels within and across Local Area Networks (LANs) to allow multiple types of network channels on the same physical infrastructure
- **DESKTOP VIRTUALIZATION:** puts the entire PC environment on a virtual machine so it can be provisioned to users from a common centrally stored "image".
- **APPLICATION VIRTUALIZATION:** takes apps and their supporting OS and creates a logical package for each app that runs as an independent layer on an OS so apps don't effect each other or the underlying OS[3].

Virtualization is the secret sauce that enables media and marketing shops to start switching from a CAPEX-led IT cost base to an OPEX-led IT cost base.

Potentially, the most important flavor of virtualization for media and marketing is device virtualization. This is where virtualization technologies and techniques allow designers to emulate DVRs, set-top boxes, radios, TVs, computer monitors, any of which can be targeted and/or customized for various hardware. Think forward to CES 2011-2013. It's highly likely that you'll find consumer electronics plays in which the hardware consists of an awesome screen, killer speakers, universal remotes/game players, an Ethernet and Wi-Fi broadband connection—but little more.

The action takes place when the consumer connects the hardware to broadband and selects the cloud-hosted services, content, channels, applications and features that will be configured to the device. Do you want a game-centric set-up that also gets MTV, CNN and Sirius satellite radio? No problem. You want to turn that 1.5 meter wall mounted HDTV into surround sound Hulu? Check. Do you want a bigger screen in a year or two? Excellent, just save your profile, swap out the hardware and start again. Of course, business models for these scenarios are an entirely different subject. This is just an example of what can be done technically.

3 If you want to drill deeper into virtualization and cloud building blocks, check out PricewaterhouseCoopers' Global Technology Center Summer 2009 issue devoted to cloud computing. http://www.pwc.com/us/en/technology-forecast/summer2009/index.jhtml

Your IT environment has gone from living in a detached house with a white picket fence to living in a giant condominium.

Devices are where virtualization gets sexy for media and marketing. But there are more profound reasons for media and marketing executives to start thinking about all things virtual. Virtualization is the secret sauce that enables media and marketing shops to start switching from a CAPEX-led IT cost base to an OPEX-led IT cost base. On the revenue side, virtualization allows you to innovate and test new ideas faster, cheaper and better than if you needed to reserve a piece of physical infrastructure for each new project.

But, as mentioned earlier, we live in a rational universe. There are trade-offs.

Perhaps the most important trade-off involves the understanding that a virtual operating environment is, by definition, a shared operating environment. At the end of the day, that wonderful cloud into which you're throwing all your media files and business processes actually consists of massive physical data centers operated by cloud providers. These data centers enable you to encode and serve up your digital content assets for everything from iPhone to Xbox. But along with your job, the data centers are backing up other people's email. They're storing photo albums and student records for the local middle school. Progressive hospitals and clinics are using them to make mirror images of their patient files. The data center is crunching massive amounts of enterprise sales reports or other corporate data.

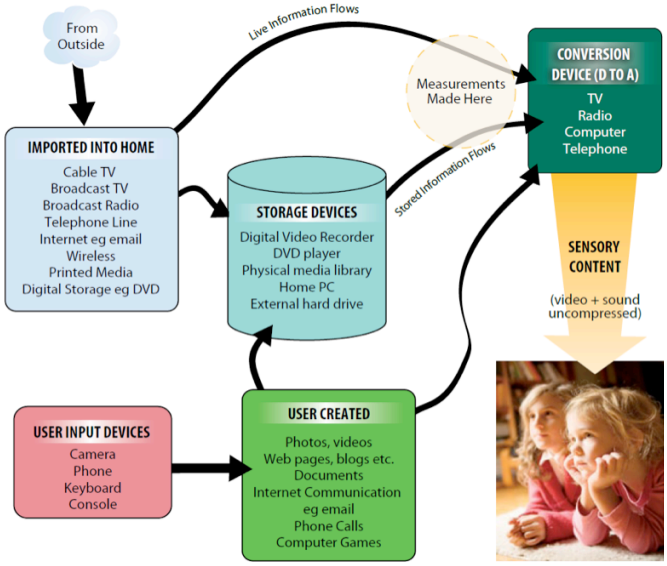


CHART COURTESY OF: Global Information Industry Center University of California San Diego

Your IT environment has gone from living in a detached house with a white picket fence to living in a giant condominium. What you're saving in scalability, maintenance and hassle, comes at the price of learning how to be a good neighbor.

For example, you might want to digitize an entire movie library into several different formats so it can play on multiple screens. It's 1030AM on a working Wednesday. You spin up a lot of virtual machines and quickly start cursing because the performance sucks. Sure, the data center can probably crunch whatever you throw at it. But the *data* has to get there through the public Internet and that depends on the speed of your network connection.

If we accept that data centers are the computing equivalent of electric power plants, why are we shocked that they exhibit many of the same performance issues? Heavy electricity users move their jobs around as best they can during the summer when demand for air conditioning pulls hard on the grid. Unless it is mission critical, why would you schedule a heavy upload/download job in the middle of the business morning when everyone is awake, video chatting and using social/streaming media?

Cloud providers would love to tell you that they're infinitely scalable, meaning that you can get all you want whenever you want at one low price. The ethical

ones don't say that because it's simply impossible. Amazon has limits to what it can digest at any one time when the real world hits its cloud. Google has limits.

Moreover, data centers will eventually suffer a black-out just like power plants. Anyone who denies it can or will happen is smoking something. When data centers go down, the typical response by bloggers and/or other media is to claim that cloud computing isn't ready for prime time. It makes for quick and loud copy but that doesn't really help your business.

So remember the following:

A. Virtualization enables a data center owner to create and operate a shared infrastructure. A shared infrastructure results in much better utilization of hardware, human management and other capital intensive computing resources. That's the better economics you want.

B. The ultimate performance of shared infrastructures comes from how well the data center owner forecasts demand peaks and valleys to where the end-customer has the perception that they're getting as much as they need when they need it. We're still on the learning curve here.

C. You, the customer of the cloud computing service, must take responsibility for modeling your business risk and exposure to temporary sluggishness of the cloud service all the way to failure of the external data center.

D. You, the customer of the cloud computing service, need to establish the internal business processes and policies that ensure your employees or end-users do not consciously or inadvertently put the business at risk.

Concentrating only on the financial cost savings of cloud computing without thinking through the new risks qualifies you as dumb money.

That's the true price you're paying in order to get the savings and flexibility you're after. Concentrating only on the financial cost savings of cloud computing without thinking through the new risks qualifies you as dumb money.

Google, Amazon, Rackspace, Microsoft or any other cloud infrastructure provider are NOT selling data centers. They're selling you a cloud computing service. It's little different from the service you get from your bank (a shared infrastructure), or your airline (another shared resource), or UPS/FedEx (ditto). You trust these organizations to manage assets or execute business processes on your behalf. If they fail to meet their obligations per the contract, you suffer damage.

So sue them like a normal company.

But don't blame finance, aeronautical engineering, or operations research if you misjudge interest rates, miss your flight or your package doesn't get there on time. Likewise, don't blame virtualization, virtual machines or cloud computing if you don't understand and mitigate your risks.

Pipes and Joints: Infrastructure-as-a-Service (IaaS)

What you'll find here:

- The three main elements of IaaS
- Computer
- Storage
- Network

Why it's important:

- This is the plumbing for cloud computing

In a nutshell, Infrastructure-as-a-Service (IaaS) provides a virtual machine that functions as a server. It's primary value is to mimic physical infrastructure. The three main subcategories are COMPUTE, which relates to the processing aspect of a server, STORAGE which is how files are stored persistently or content is prepped for delivery via the web and NETWORK which includes inbound/outbound bandwidth with internal cloud communication. Most IT shops pull these three elements together into one unit called infrastructure. But it's possible to contract with a cloud vendor in a more a la carte manner. As always, it depends on your business.

The first thing you'll notice when you try to access **compute** services is that the marketing terms and definitions are all over the map. For example, Amazon's basic compute unit in EC2 is called an "instance". GoGrid calls them "cloud servers". Joyent offers an "accelerator". Rackspace provides a "server instance". Regardless of the particular name given by the vendor, you're paying to access some kind of *computer* to process your job.

You might be encoding video files. You might be analyzing your media or ad server logs for trends. You might be offering a gaming service that scales a virtual world across users. Within reason, the actual job shouldn't matter at the compute level. Naturally, as you move up the stack, the particulars of what you want to do will matter more. However, just think that you're accessing the IT equivalent of horsepower whenever someone is talking to you about IaaS compute services.

As far as a shopping list goes, you will configure a virtual computer that specifies the CPU, OS, memory and disk space you need. Usually, you will choose from a set of pre-rolled options offered by the cloud provider. Call these basic configurations an "image". An image is a blueprint that allows you to spin up as many or as few virtual machines as you need to do the job. You should be able to add applications (e.g. your security or other corporate app) to your image that you want baked into each virtual machine. You should be able to take snapshots of your customized image to replicate new virtual machines to your spec fast and as-needed. The cloud provider should support various OS such as Windows, Linux flavors like Ubuntu/Fedora/CentOS, or Open Solaris to name a few. You should also get an Application Programming Interface (API) so you can pull in 3rd party cloud management tools like Rightscale or Elastic Server if you don't want to use the cloud provider's in-house management kit.

Some of the primary IaaS offers out there are Amazon Web Services (AWS), GoGrid, Joyent, or Rackspace cloud. Certainly, there are others like AppNexus, Flexiscale etc. As buy-side people, you can expect to have abundant choices for cloud infrastructure when you start out. As you work through what moving parts of your business are best suited for which type of cloud service, however, the list of possible vendors will shorten considerably. So don't sweat the fact that there's a lot of choice out there.

Billing models for compute services typically revolve around the configuration you choose (e.g. CPU, RAM, OS, disk space), the amount of time you access the virtual machine, and the data transfer bandwidth in-bound/out-bound. Typically, a cloud provider will give you a range of virtual machine configurations from which to choose. The differences among machines are processors, RAM, and disk memory for the most part. For example, AWS chops its EC2 offer into standard instances and high CPU instances. The default configuration for an EC2 instance is 1.7GB of RAM, 1 virtual core CPU (e.g. the equivalent of 1.0-1.2 GHz 2007 Opteron or 2007 Xeon processor), 160GB of storage, a 32-bit platform and moderate input/output (I/O) performance. All of this comes at $0.10 per hour in the US market for Linux, $0.125 for Windows. GoGrid's entry level virtual machine is 1 Intel Xeon equivalent, 500MB RAM, and 30GM of disk storage for $0.095 per hour.

Conceptually, cloud storage is little different from one of those physical storage facilities where you rent a numbered stall and receive a key.

Cloud-based Storage

The next important category of IaaS involves cloud-based **data storage**. Don't confuse this with the disk memory supplied with your COMPUTE service. In many, if not most, cases once you terminate a virtual machine, you kill whatever data was stored in that instance as well. Even more important, don't confuse storing data files with having a database. Storage means exactly that. You're renting buckets or whatever the vendor term is to stick your files. Conceptually, it's little different from one of those physical storage facilities where you rent a numbered stall and receive a key.

There are two broad reasons for cloud-based data storage. The first reason is the need for persistent backup for large files. Typically, this involves frequent WRITES to the data store and less READS of the data. You might do this for disaster recovery. You might want a single version of truth of your data while a slew of instances chew on copies of it. A second main driver for cloud-based data storage is to keep your content ready for web delivery to end-users or other applications. This involves a lot of READS and much fewer WRITES to the data source.

You shouldn't let a vendor lock you into a particular language to upload/ download data. It needs to be an open API to where you can use your own tools.

It's time for data storage shopping list. First and foremost, you want to abstract away all the hard logistics of storing data. As a customer, you shouldn't need to care about the physical addressing schemes for where the data is going. All you want is a virtual address to store the data. The file backup should be automatic after you've set up the back-up schedule. You shouldn't let a vendor lock you into a particular language to upload/ download data. It needs to be an open API to where you can use your own tools. Further to that, you want to be able to define the level of security model and/or levels of encryption. To improve READ/WRITE performance you probably want your data to have a pipe to a Content Delivery Network (CDN) so it can live as close as possible to where it will be used. Naturally, you want to pay only for what you use. Depending on your business, you might want a cloud vendor tool for moving stonking large files (>20GB), otherwise standard bandwidth charges and performance might start to bite.

Big cloud players in the data storage services business include AWS S3, Rackspace cloud, and Nirvanix. These firms have their own physical storage infrastructure. There are also a large number of storage service providers who piggyback on the big players' infrastructure and differentiate via their management tools and intuitive user interfaces (UIs). Companies like JungleDisk and Elephant Disk are in that category and sell mainly to consumer and SME markets.

Billing for storage services typically includes a raw per GB rate, plus transfer bandwidth, plus certain data command requests like COPY, GET, LIST, POST, PUT. There are other commands such as DELETE that are offered free. For example, Rackspace Cloud charges $0.15/GB per month for an unlimited number of files that are less than or equal to 5GB in size for any one file. To upload a file to the Rackspace cloud, you pay $0.08 per GB in bandwidth charges. Downloading data from the cloud is more expensive at $0.22 per GB. You pay $0.01 for the first 500 PUT, POST, LIST requests for small data files (<250kb) with command requests for larger data files (>250kb) being free. Alan Williamson has a nice on-line tool for playing with storage costs across the three vendors at [http://www.aw20.co.uk/tools/storagecosts.cfm]. Please note that the actual pricing on the tool is from November 2008. The more important thing is to see how the numbers start jumping around once you start inputting different amounts of data in GB and whether your usage trends to uploading or downloading.

Networking and the Cloud

The final piece of IaaS is the **network** aspect of cloud services. Networking in the cloud involves uploading data or messages into the cloud for processing or storage, downloading data/messages from the cloud to some kind of client or end-user, and transferring data/messages within a cloud or across clouds that are part of some parallel or distributed computing process. For our purposes, just split cloud networking into two main buckets: inbound/outbound bandwidth + message queues.

When you think about inbound/outbound bandwidth, think content delivery network (CDN). CDNs are sets of computers (sometimes called "edge nodes") placed at various geographic or logical parts of a network to speed up data access throughout the network. When a client uses a CDN, they are accessing a copy of the data nearest to them instead of trying to hit the same, centralized server where the original file lives. A CDN can move web objects, downloadables like media files, software and documents, applications themselves, media streams as well as other types of networking data such as routes or DNS queries.

The logic for using CDNs to deliver cloud computing is pretty straightforward. It may well be true that the supply side for cloud computing is centralizing into behemoth-sized data centers. However, the customers of these services are widely distributed. Moreover, the idea of customer or end-user needs to be expanded beyond John Q. Public trying to pull down a web page. Far more important are situations in which the end-user of a server call isn't a human being at all, but is another application or service delivered via the cloud. Many Software-as-a-Service (SaaS) like SalesForce.com wouldn't be viable without being hooked directly in a CDN.

On the billing side of CDNs, you typically pay a per GB rate that changes according to peak and off-peak times. Think of it similar to a US mobile phone voice plan. There is a price per minute that discounts depending on what kind of "bucket" you sign up for and for how long. In the case of a global CDN, the smallest data bucket of data you can contract for may be several terabytes. Typically, a CDN contracts more at the B2B level with another service provider that's baking in the CDN cost with their service to you. However, if you're big enough, you can cut your own deal.

It doesn't matter if you're getting a great deal on storage or processing if your customer feels the services is slooww.

It's a natural fit for cloud computing services to hook up with CDNs. One of the first deals out of the gate came in late 2008 when Rackspace announced a deal with Limelight Networks. Limelight operates a large optical network with thousands of interconnection and/or peering arrangements with other networks. The deal is that Rackspace storage and processing services are connected to Limelight's CDN to amp up their performance. Another cloud-oriented CDN effort is Amazon CloudFront.

CloudFront integrates with other AWS offers like S3 to offer low latency, high throughput data transfer. A request for a data object stored in S3 is routed to the nearest edge location. In CloudFront, S3 objects are organized into "distributions", which specify the location of the original version of the data. A distribution gets a unique CloudFront.net domain name (e.g. ReallyCoolStuff.cloudfront.net) that references the objects throughout the network of edge locations. When the object is requested using a distribution domain name, the object is automatically routed to the nearest edge location for faster delivery.

The 800lb gorilla in the CDN world, Akamai, is also going after cloud computing with a vengeance, albeit at the B2B level. Targeting SaaS providers, Akamai is positioning its CDN as a means of keeping performance high and latency low for SaaS-based applications. CDNs are massively capital and expertise intensive. Akamai claims nearly 50,000 edge nodes in its network. If your media demands low latency like video or gaming, the CDN choice looms large. It doesn't matter if you're getting a great deal on storage or processing if your customer feels the service is *slooww.*

The last category of cloud networking involves queues. Message queues are what enable communications among distributed computing systems. Just like the beloved British term for standing in a line, queues manage the flow of messages that pass control/content/instructions across applications that might be living on multiple machines or different parts of a cloud. Amazon Simple Queue Service (Amazon SQS) is an example. Developers use SQS to move data between distributed components of their applications that perform different tasks, without losing messages or requiring each component to be always available. The reason you want to be able to do this is so you can chop a job like video encoding to run across multiple commodity (read "stupid") machines, whether virtual or not. Aside from efficiency gains, parallel processing and good queuing enables an individual machine to fail without killing your job. Like CDNs, you buy a bucket of message requests with Amazon SQS ($0.01/10,000 requests).

You're renting a *computer*

To recapitulate, IaaS marries computer processing, data storage and networking to give you the equivalent of a server via a virtual machine. To get that complete virtual machine, you'll pay a rate for raw processing, You'll pay a second rate for inbound/outbound bandwidth. That charge might be baked into your processing rate or it may be separate. You'll pay a rate to store your data, which may or may not include inbound/outbound bandwidth. If you do a lot of parallel or distributed processing, you'll pay

another rate for a queuing service. You can buy these services a la carte or bundled into one line item.

But take a look at things you're not directly paying for any more. The capital cost of a server is straightforward. Even though it's the easiest to see, it's actually one of the smaller categories of cost savings. Potentially larger are IT human resource costs, electric power, real estate, insurance, software licensing, and maintenance to name a few. The potential game changers for your media shop, however, are TIME, HASSLE and UTILIZATION.

Think back to the last time you ordered a physical server. It doesn't matter if it's in your basement or in some hosting provider's rack. It's yours. How long did it take from the precise moment you signed off on the budget to the precise moment that server crunched your first job? How many meetings did it take to get that server included as a line item? How many times has your staff needed to take that server off-line for maintenance or re-set it after a crash? And, what is the current utilization of that server?

If you answered "a lot", "a lot", "a little", you're probably a good candidate to be a cloud computing customer. If, on the other hand, your company runs its servers hard 24/7 to crunch data, perhaps your choice whether to adopt IaaS isn't as clear cut.

IaaS marries computer processing, data storage and networking to give you the equivalent of a server.

Platform-as-a-Service (PaaS)

What you'll find here:

- The difference between PaaS and IaaS
- Top level PaaS elements
- PaaS examples

Why it's important:

- PaaS is your development environment for innovation

It's not a 100% hard and fast rule, but here's a decent rule of thumb for differentiating IaaS and PaaS:

IaaS is *machine-centric*.

PaaS is *application-centric*.

If you think about an automobile, then roughly speaking IaaS comprises the engine, the chassis, the transmission, wheels and the body. PaaS gets into the steering wheel, the gear shift, accelerator/brake pedals, and the instrumentation. IaaS is about creating and accessing IT resources. PaaS is about developing and deploying IT applications.

IMAGE SOURCE: *www.nytimes.com*

The difference between PaaS and IaaS is more about goals than methods. Both IaaS and PaaS include compute, data and networking elements. However, IaaS is an operations environment where the value of compute, data and networking concentrates on increasing the productivity of infrastructure. PaaS, on the other hand, is an environment where the value of compute, data and networking focuses on creating and increasing the productivity of applications.

For our purposes, PaaS solutions are application development and operations environments that are hosted in the cloud. Using PaaS, developers can build applications without installing any specialized tools locally to a machine. Additionally, developers can use PaaS to deploy their applications without specialized systems administration skills. At no time should a PaaS customer worry about service availability, load balancing, scaling, OS maintenance, security and a host of similar infrastructure concerns.

Using PaaS may result from much faster development of custom applications that a business needs to differentiate from a competitor. Another flavor of innovation is the integration of existing applications in a novel manner. PaaS can be especially important for stitching together various in-house applications, external data sources and SaaS applications into a unique cloud-based configuration.

To date, there are four primary deployment models for PaaS solutions:

- **Social Application PaaS:** this is the realm of Facebook and other web 2.0 companies that provide APIs so that 3rd parties can write new applications that add value to the user experience. The platform is not simply a development environment but is also where the developer finds their user audience.
- **Raw Computing PaaS:** this is where developers upload their traditional software development stack to an IaaS solution like Amazon Web Services. This is the realm for the hardcore software developer.
- **Web Application PaaS:** This is where a big player like Google provides APIs that lets developers pull service elements like maps, calendars, spreadsheets and other services to bake into the developing app. Innovation is usually the result of some unique mashup combination as opposed to genuinely new technology or programming technique.
- **Business Application PaaS:** This is where platforms like Force.com provide the application infrastructure that is specifically focused on transactional business apps such as database, e-commerce, integration, workflow and user interface levels.

Generally speaking, PaaS offers three main elements:

1. An Integrated Development Environment (IDE)
2. A database
3. A runtime environment

An IDE is a software workshop for geeks.

As a PaaS customer, you control those three elements. But you don't control the hardware or the network infrastructure on which your applications are running. Let's look at each main element that make up PaaS.

Integrated Development Environment (IDE)

An IDE is a software workshop for geeks. Like your grandfather's woodshop, an IDE contains a workspace and a toolset. Instead of planes, saws and

drills, IDEs have source code editors, compilers, build automation tools, and debuggers. The purpose of both a woodshop and an IDE are similar. Both aim to transform raw materials into a finished product.

Assume you know what you want an application to do (e.g. analyze server log activity to let you price your display ads better). Your application developer will need to write and edit source code, catch syntax errors (bad software grammar), automate repeated tasks, compile code (translate it into machine-readable format), build connections to databases, and debug logical problems (grammatically correct gibberish).

Do a LOT of homework to understand how your data is organized by the PaaS database.

Most IDEs are configured to support a specific programming environment like Windows or JAVA. The IDE provides a feature set that tries to match the programming models that make up the environment. You've got tools for creating, debugging, mounting, and integrating an application with a given computing environment. There are some multi-lingual IDEs such as Eclipse, which is an open-source IDE that can be used for JAVA but also for C, C++, Perl, Python, PHP etc. Other IDEs such as Microsoft Visual Studio focus on developing applications for all flavors of Windows.

Like wood or metal working shops, IDEs run the gamut for offering tools and sophistication. The main point for a customer of a PaaS offer to remember is that IDEs are the workspace for developing applications.

Database

If IDEs are the workspace for applications, then databases supply the raw materials. Don't confuse data storage offers like Amazon S3 or Nirvanix with a database. Data storage is the file cabinet. The cabinet says you can stuff X number of GB here for Y price. On the other hand, a database is the logical filing system (e.g. alphabetic, by date, by customer value) for organizing what you've stored. Hence, one of the most important links in any "platform" is between the IDE and the database. Do a LOT of homework to understand how your data is organized by the PaaS database. This is where a significant amount of vendor lock-in takes place in the cloud.

Database technologies and techniques can be mind-numbing once you get under the hood. For our purposes, however, it's more important to remember that all that black magic and rocket science focuses on structuring data so you or your application can request data or ask it questions. Whether you've rolled up the 3rd grade parent mailing list in Excel or you're generating several terabytes of data each day like a MySpace or a Facebook, the fundamental principles don't change.

The currency for structuring data consists of data objects, their attributes and their relationships. For example, imagine that you've got a HR recruiting app. Assume for now that you've got a nice web browser user interface for your people. Behind the scenes in the database, you've created data objects for the HR app like POSITION and CANDIDATE. But currently they're empty shells. So you create attributes that live inside them. For now, assume that we document those attributes using columns like a spreadsheet. CANDIDATE becomes an object that has columns for name, address, email, university degree etc. POSITION might have columns for department, title, location, compensation etc.

Under those columns, the database will create rows containing the specific sub-fields related to Jim Smith or the unfilled position of VP of Sales. After you've created data objects and declared their attributes through sub-fields, it's time to make relationships. Those relationships can exist within a data object (e.g. location is flexible for the Sales department but fixed for the Operations department) or across objects (e.g. Operations doesn't consider candidates with less than a masters degree).

Data objects, attributes and relationships are the set-up for a database. The action, however, happens at the level of a request or a query. You or an application can dream up extraordinarily complex questions. But typically, the two query types that hit a database boil down to "Give me everything related to X" or "Give me only that data that has Y attribute or is related to Z". The first instance ("give me everything…") can be called an information retrieval query, which is what you do when you fire up a search engine. The second flavor ("give me only…") is often called a database query. There are all sorts of query languages to help you interrogate databases.

Everyday computer end-users employ runtime environments and libraries whether they know it or not.

The purpose for all this, however, doesn't really shift. You need to answer a question so that you can make a decision.

Now substitute the anthropomorphic use of "You" in the preceding examples with an IDE or software application that needs to query a database to pull the information it needs to make its next processing decision. Very quickly, it should be apparent why database decisions can have a profound impact on your business in the cloud.

Runtime environment

When you pull the trigger on a gun, it fires. All the various sub-processes (e.g. hammer being released, striking the cap, spark igniting the powder, a focused explosion driving the bullet out of the barrel) come together in one BANG! When you spark at application with the RUN command (however delivered), you've pulled the trigger so to speak.

A PaaS solution is the fundamental platform through which a creative idea becomes an executable application.

Runtime environments do two things for an application. First, they hold the pre-canned libraries and subroutines built into a "runtime library". A runtime library enables almost any application to send instructions to the computer's processor, grab RAM memory and other system resources. Basically, they enable the application to do its stuff.

Everyday computer end-users employ runtime environments and libraries whether they know it or not. When you watch a FLASH movie on the web, an embedded FLASH player creates a separate runtime environment for the FLASH file to execute on the page. The most common runtime environment is the JAVA Runtime Environment (JRE) that enables the applets and full applications to run on any machine with an installed JRE.

At the same time, runtime environments are very important for application developers. Not all errors are caught by source code editors or the compilers that make up an IDE. In many cases, the errors won't happen until you try to run the application. A runtime environment enables the developer to track instructions being processed by the application and debug runtime errors. If an application melts, the runtime software often keeps generating telemetry that can provide information about why the program crashed. That diagnostic information isn't restricted to just internal errors within an application. It's more often the case that applications need to exchange data or commands with each other. A development team needs to see how these components interact with each other real-time and at scale to catch problems. Additionally, the diagnostic data generated by a runtime environment can be used to increase an application's performance.

You don't want to be doing that in the production environment of IaaS with the meter ticking.

To sum up, a PaaS solution is the fundamental platform through which a creative idea becomes an executable application. Your developers need to have a workspace and tools (IDE) to create applications. The applications need raw materials (database). And you need to find out how the applications behave when you release them into the wild (runtime environment).

PaaS Examples

force

FORCE.COM: The canonical example of PaaS is the Force.com platform from SalesForce.com. The Force.com platform consists of a stack of database, integration, logic and user interface (UI) capabilities. The Force.com customer brings their application requirements, design and data to the table. They use Force.com's IDE, database, integration, workflow and UI tools and services to bake an application, test it in a runtime environ-

ment, and then deploy it on the SalesForce.com infrastructure. Effectively, SalesForce.com has integrated IaaS resources and capabilities under the Force.com PaaS offer. Whether those IaaS resources are wholly owned by SalesForce.com or accessed from another IaaS provider isn't relevant to the Force.com user.

The infrastructure layer supporting Force.com is the same infrastructure that delivers SalesForce.com's main CRM application. It's a combination of data centers and security technologies in three geographically separate data centers. The SalesForce.com data centers use near real-time replication to mirror the data at each location.

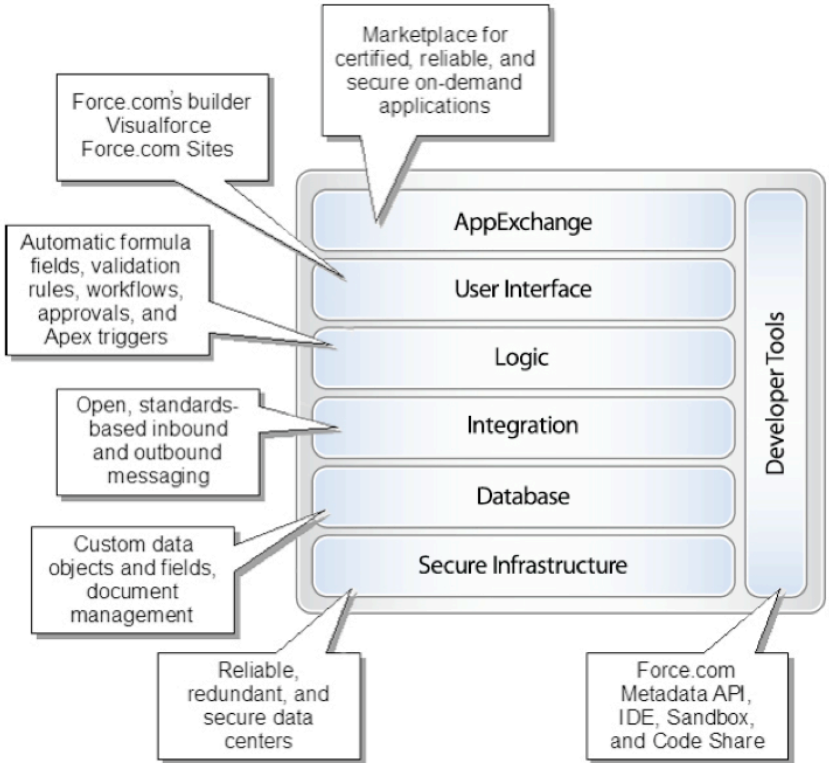


IMAGE SOURCE: salesforce.com

Secure data centers are almost table stakes to get into the PaaS game. They are required but add little differentiation from one provider to another, especially as more data centers come online. More important for PaaS is how a vendor organizes their database as a service offer. Especially for business applications, databases and tools provide much of the development firepower. The Force.com database uses objects to store instances (records) of data. The database has several similarities to and differences from traditional relational databases. Think of a data object and its fields in Force.com as similar to a relational table with its columns. However, unlike relational database tables that separate data objects with separate storage, Force.com maintains the structure of a data object as metadata (e.g. data about data). There are a few huge data storage tables in Force.com that hold the data until it's required at execution. At runtime, Force.com then renders the virtual object records by analyzing their metadata.

Other Force.com services include integration, which helps businesses tie in their legacy, largely non-cloud based applications such as ERP. Integration looms large as an enterprise concern as does the need to model a company's unique business processes and requirements. Force.com offers a mapping and workflow engine to enable Logic as a Service for those needs. The Force.com stack finishes with a layer to enable quick creation of customer user interfaces as well as the ability to buy rather than build Force.com applications through the aptly named, AppExchange.

Spanning all these layers are developer tools that provide full access to the data, logic, integration and UI capabilities of Force.com. The determining factor seems to be the internal sophistication of the customer's needs and development assets more than anything else.

From a pricing point-of-view, Force.com charges according to the number of deployed applications and their user base. Other factors such as storage and then number of data objects increase in proportion. There is a free version that enables a customer to develop and deploy one application to 100 users. The next level up is called Force.com Enterprise. It enables 10 applications per user, up to 200 data objects, access to CRM services from SalesForce.com, and more than 100 users. The cost is $50 per seat per month. Finally, there is Force.com Unlimited which costs $75 per seat per month for unlimited apps per user, 2,000 data objects, direct technical support from SalesForce.com and increased storage.

Force.com can be considered one of the most native PaaS offers in the sense that it was incubated in a cloud-native company in the first place. How long that distinction and operational experience advantage persists remains to be seen as other large IT infrastructure and services shops get into offering PaaS.

GOOGLE APP ENGINE: Another PaaS offering is Google App Engine. Unlike Amazon Web Services which offers the end user an a la carte menu of processing (EC2), data storage (S3), and content distribution (CloudFront), Google has opted for a menu fixée.

Google App Engine was first released in beta in April 2008. It's basically a platform for developing and hosting web applications in Google's managed data centers. In that sense, the much abused "end-to-end" marketing phrase can reasonably apply to the Google offer. The fees charged to a developer revolve around storage, bandwidth, and CPU cycles required by an application.

Currently, Google App Engine only supports development in the JAVA and Python programming environments. Google has said it intends to support more programming languages in the future. It's also true that the App Engine environment itself has been written to be language independent. But the reality in late 2009 is that unless you're accustomed to developing in JAVA or Python, you're largely out of luck.

In return for a trade-off on development environments, you're gaining the infrastructure and management skills to scale your web application to handle millions of users. App Engine offers the web app developer full

support for all common web technologies, persistent data storage with standard database tools for queries, sorting, and transaction. Other App Engine features include automatic scaling and load balancing and a robust set of APIs to hook into other Google services like Maps, Checkout, Gmail etc.

Of course, your technical staff can rifle through all the different trade-offs and advantages inherent in the App Engine environment at the level of the individual app. At the strategic level of your business, the question is a bit more abstract — but important. Fundamentally, the biggest advantage to using App Engine is that you can focus almost exclusively on the specific problem or opportunity your application or service is trying to address. All the "muck" (Jeff Bezos' term for the need to manage hardware and operations) is lifted off your books. Assuming you have good developers, you will be very fast out of the gate and will be able to scale your offer smooth and quick.

PaaS is integral to a company's capacity to innovate rapidly.

That's not trivial.

On the other hand, you're married to Google whether you admit it or not, like it or not. The more you develop, the more you're locked in. Don't worry so much about the fact you need to develop in JAVA or Python. More critical is that you need to format all your data to Google's way of doing things. You won't be using a quasi-standard relational database like SQL but will use Google's flavor known as Big Table.

Again, there's nothing inherently nefarious about this kind of arrangement. Google is offering hands down one of the biggest and most successful computing infrastructures in human history. They definitely know a thing or two about keep distributed computing services up and humming. They built the infrastructure. They write the rules on how it's used. That's capitalism.

OTHER PAAS SOLUTIONS: PaaS offers have begun to blossom with a vengeance. For example, Heroku, is one of the popular PaaS flavors for the Ruby on Rails community of web app designers. Microsoft has spent most of 2009 preparing for its November launch of the Azure Web Services platform. Azure offers a Windows-based cloud environment for running apps and storing data on servers in Microsoft data centers. Azure provides cloud-based data services through its SQL Azure component while it enables distributed infrastructure to cloud-based and local apps through its .NET system. Until Microsoft takes Azure out of beta and sets it loose, it will be difficult to evaluate whether Azure is a good bet for a media shop.

Regardless of the various feature/functions of each vendor's PaaS offer, the value proposition of PaaS can't be neatly boiled down to direct cost savings and simplicity. PaaS is more integral to a company's capacity to innovate rapidly. Those decisions, by definition, tend to cross organizational and technical boundaries. The key questions come back to what makes your firm special in the first place.

Software-as-a-Service (SaaS)

What you'll find here:

- SaaS definition
- How SaaS works (multi tenancy and metadata)
- SaaS examples
- Trade-offs

Why it's important:

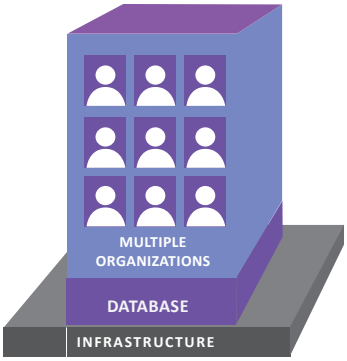
- You're trusting an end-to-end business process to SaaS

Software-as-a-Service (SaaS) is a deployment model in which a service provider licenses you an application to use on-demand and pay only for what you use. Here's a quick and dirty metaphor:

You've decided to join a club. You are issued your own table at the restaurant which you can decorate as you see fit. If you're so inclined, you can bring personal property to the table such as an ice bucket from home. The club issues you a membership card. The card documents your ID plus a description of services you can expect given your level of membership (e.g. hourly, daily, monthly, lifetime). The card also tells the club staff where to find your personal table, how it should be set, and where to find your tagged bottle of Scotch. All the information that makes this membership come "alive" happens when you show the card at the front door of the club. It shouldn't be your concern as a member what happens to your table or personal property when you're not at the club. You just expect everything to be available instantly when you want it. After all, membership has its privileges.

In a SaaS environment, all the infrastructure, licensing, application maintenance, upgrades, delivery and so forth are the responsibility of the service provider. As a SaaS customer, you pay a usage or subscription fee that goes up or down based on your customizations, scale and use of the core application. The service provider owns the app. You own the data generated by using the app.

SaaS is the cloud computing layer that has received the most attention of the general press. Whether industry news refers to Google's Gmail or Docs, SalesForce.com, Hotmail or other Web 2.0 implementations, most people get SaaS as a cloud computing concept even though that's not the same thing as being able to use SaaS successfully.



How SaaS works

Virtualizing physical infrastructure and application environments is the secret sauce for IaaS and PaaS.

Multi-tenancy is the secret sauce of SaaS.

Multi-tenancy means that each SaaS customer pulls from a single core version of a software application. If you recall from the IaaS section, virtualization uses a master blueprint that specifies hardware specs and other infrastructure resources called an "image", from which any number of identical running instances (e.g. individual virtual machines) can be rolled.

Take that IaaS image concept and substitute a core application. Instead of a master image, the SaaS provider uses a master version of an application. When a SaaS provider spins up an instance of the app from that master version, they're spinning up a customer specific instance, which imports the preferred look and feel, the customer data as well as any additional customer services (e.g. add-on security models and code).

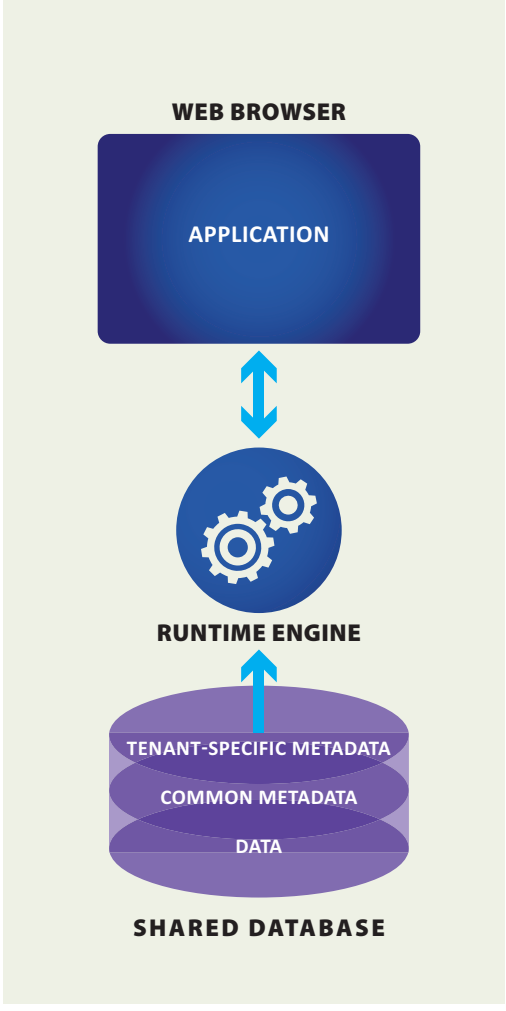
Don't think with SaaS that everything runs off one server. In reality, you're talking about huge data centers with thousands of servers humming along to keep the SaaS offer operating and scaling according to demand. The SaaS provider no doubt operates several geographically separate data centers that mirror both the core application and individual tenant data.

Multi-tenancy enables the SaaS provider to offer compelling economics in both speed and cost for broad applications like customer relationship management (CRM). Most SaaS users cite rapid deployment as a major benefit. SaaS also enables customers to deploy in tidy increments, commit to short-term rather than long-term agreements at the outset, and roll out new functions easily. SaaS customers can add new users or remove them without worrying whether that 501st new user requires an upgrade to a 1,000 seat version of a software application. Multi-tenancy enables SaaS providers to standardize on a web browser like user interface (UI). This hits user adoption and training issues head-on.

Multi-tenancy arose from the ashes of the Application Service Provider (ASP) market from the 90s. ASPs also delivered software applications remotely from a data center. The customer owned the software license with all the inherent maintenance, modification and upgrade headaches. The ASP managed operations, guaranteed uptime (not!) and took a monthly fee. A major problem with ASPs is because they didn't own software IP, it was often the case they needed to create a separate operating environment (including separate hardware in some cases) for each customer.

SaaS companies like SalesForce.com emerged at the end of the ASP era. They developed their own core software IP (CRM software in the case of SalesForce). The next step was to enable it to be delivered as a subscription-based service as opposed to being sold as a licensed product.

The core SalesForce.com application runs on the same operating system, on the same hardware, with the same data storage model. This enabled SalesForce to bring the per-unit cost of its core application far below traditional on-premise CRM offers. It also allowed SalesForce to offer a more robust service level agreement (SLA) than many others because the SaaS model allowed SalesForce to standardize on its infrastructure instead of needing to manage multiple operating environments as was the case with ASPs.



Metadata is data about data. Just think of it as the descriptive clothes you put on naked data to tell a system what the data is, what it can do, with whom (person or app) it can communicate.

At this point, sharp-eyed readers might start asking how a single master application allows each tenant to customize their environment and experience. How do tenants create custom extensions to standard data objects or even create new data objects? How does the SaaS provider secure tenant-specific data in a shared database so one tenant can't see another tenant's data? How does the SaaS provider patch or upgrade the core engine without melting tenant-specific customizations?

Most multi-tenant architectures meet these challenges by generating the individual tenant application only at runtime. Instead of storing a separate tenant application, database, and profile, a SaaS system typically uses metadata to separate the application services from the customer specific data from the customer look and feel. When the customer wants to access their environment, the metadata descriptions give the core SaaS engine the instructions to spin up a customized instance for that particular customer session.

Metadata is data about data. Just think of it as the descriptive clothes you put on naked data to tell a system what the data is, what it can do, with whom (person or app) it can communicate, who often needs it and so forth. For our purposes, metadata in a SaaS environment is like the barcoded baggage tags they put on your luggage at the airport. The SaaS operator takes your metadata claim tag to move your specific bundle of function and data from A to B on a shared infrastructure, not lose or damage it, get it there on time, handle special items etc.

Once you successfully log-in to a SaaS application, a virtual valet rolls your specific customer instance from the meta-data. For example, if a developer builds a custom app and defines a custom table for data, the core SaaS app doesn't create a unique table in a database or compile any specific runtime code. Instead, the platform stores meta-data that describes all these changes. When it's time to run, only then does the platform's runtime engine roll all the "virtual" components it needs to create a customer operating environment.

Accessing a service environment rather than buying a set of features and functions is an important departure by SaaS from "regular" software. Historically, software has been sold based on features, functions and versions; the more the better. That gave the vendor the ability to repeatedly sell you the same thing through different versions, upgrades, enhancements, maintenance, consulting etc. Moreover, all the internal decisions and external negotiations tended to revolve around how the software functioned rather than your business process needs. This isn't to cast software vendors in a disparaging light but simply to point out that the burden of adapting more often than not lay with the customer rather than the application.

It's not exactly a reversed situation with SaaS. There are core functions of a SaaS app just like any regular software app. But how you tie the cost-benefit to your business will be different. Signing up and paying for a SaaS app should probably be at the end of the process rather than its start. If you as a customer are to get full value out of SaaS — not to mention avoid unnecessary risks — here are a few areas suggested by GigaOm contributor Alistair Croll[4] you might consider beforehand:

ADAPTABILITY: How easily can you modify the application? This can be as simple as adding fields or building dashboards, or as advanced as accessing a programming platform.

RELIABILITY: How much can you depend on the system to function well? This boils down to four things: Performance, availability, scalability and security.

TASK PRODUCTIVITY: How effectively can your users accomplish their goals? How many cases-per-minute or entries-per-day can workers do, and how many errors do they make?

PRICE: How much will it cost — really? Because SaaS offerings are so varied in pricing, it's hard to compare them. A better model is to create several benchmark subscribers (a 10-, 100-, and 1,000-person organization) and compare upfront and ongoing costs for them.

BACK-END INTEGRATION: Can you plug it in to other things? Any enterprise SaaS offering will have to work with other systems, for everything from authentication to data sharing.

LONGEVITY: How long will the SaaS company be around, and what's your exit strategy? With ISVs, you could ask for software in escrow. When a SaaS provider closes down, your entire IT systems can vanish with the flick of an "off" switch. This isn't academic debate. Customers of Coghead's SaaS offer discovered in February 2009 that the services and license agreement had been terminated, effective immediately. Existing customer needed to sign new T&C to access their data and apps on an unsupported "as is" basis for 60 days with everything going dark at the end of April 2009. *Stuff* happens in the real world.

ECOSYSTEM: How many third-party developers and integrators surround a particular platform with plug-ins and add-ons, and how active are they? A vibrant ecosystem means a more extensible, flexible solution.

These questions form a good step 2 in your media shop after the grand "what is our cloud strategy?" meme has been hammered out. However, the rubber meets the road once you start actually negotiating with a SaaS vendor. Assuming that the business case makes sense, the real value comes from how you translate things into a service level agreement (SLA).

4. http://gigaom.com/2009/02/25/7-questions-to-evaluate-saas/

You can be as cynical about SLAs as you want. Come up with a better way to make an agreement actionable and toothsome and you will be a hero. In the interim, there are several areas to think about when you're negotiating a SLA:

What's under the hood: You're basically giving up a lot of things that used to be under the control of your in-house IT department (thankfully!). But in return you need to get assurances with teeth about operational functions like:

- Availability (how fast does the app spark up)
- Reliability (a 99.9% uptime guarantee means you can be down 8hrs + each year and still be in agreement)
- Performance (how many simultaneous users can it support)
- Maintenance
- Backup
- Disaster and Recovery

Reputable SaaS vendors submit to external security audits.

This part of the SLA is probably going to be the most standard aspect. Basically, the vendor is giving a promise about the attributes and performance of the infrastructure that enables the application.

OPERATING RISK: This part of a typical SLA starts getting away from the general infrastructure to more specifics about how customers like you will use it. Expect to find a lot of language about risks surrounding your data and how it's secured. When a SaaS vendor talks security to you, it's about keeping the external bad guys out. Rightfully, you can expect very little protection from a SaaS vendor if you have a corrupt insider. Reputable SaaS vendors submit to external security audits. You can bet that as SaaS gains traction your business liability insurer will start giving you boxes to tick off on forms regarding how your SaaS vendor handles data.

BUSINESS RISK: This separates the tourists from the hard-core in the SaaS world. You can't nail these parts of a SLA with as many formal measures and benchmarks as you can with the previous areas. That said, as certain SaaS vendors like SalesForce or Intacct (accounting services) gain experience and track record, you'll start seeing language that addresses technical support quality, expected ROI, or even business value received. You can bet that this will be the area of most negotiation if you're big enough to do a custom deal with a SaaS provider.

CARROTS AND STICKS: Nothing works for either party unless the SaaS provider must legally pay for failure and the SaaS customer is willing to pay extra for killer performance that delivers clear value. If the vendor isn't on the hook for performance, they won't perform period. Likewise, if you as a customer expect 100% uptime and reliability, you need to pay for it.

There are numerous methods for operationalizing the above points in a SLA, which are beyond the scope of this paper. Just know that as a customer, you want to find SaaS vendors who are as transparent as possible. If they are afraid of public disclosure of their performance against a published SLA, there's no doubt a very good reason lurking somewhere.

As with everything to do with cloud computing, caveat emptor!

Media Specific SaaS Examples

Ooyala

800 W. El Camino Real Suite 350
Mountain View, CA 94040
http://www.ooyala.com

SUMMARY: Ooyala means cradle in Telugu, a South India language. Founded in early 2007, ooyala provides end-to-end services and technology for powering online video for a number of different screens. Strip out the jargon and ooyala's service model looks like UPS/FedEx for online video. Its main product ("Backlot") enables content owners to transcode, manage, deliver, analyze and monetize their video assets.

Another growth area for ooyala is interactive video in which elements of video streams such as a character's hat are made clickable in order to link out to e-commerce, search, maps and other services. Headquartered in Mountain View, CA with sales offices in New York City and London, ooyala was started by Google alumni from both the AdSense and web search area.

PROBLEM: Video is difficult to monetize except by the crudest methods — transactions, subscriptions or advertising for eyeballs. Ooyala's aim is to increase the scope for monetizing video by managing the entire ecosystem on behalf of the video content owner ranging from video ingestion to distribution to ad serving to analytics.

APPROACH: Since launch, ooyala has focused on bread-and-butter issues of managing video assets using the cloud and thereby drive a lot of the cost out of the system. With that trend well established, the next stage is to make video more interactive and engaging in order to open more chances for monetization. The founders realized they had a good bit of computer vision expertise in the company. The idea is to use computer vision techniques to mark or paint objects inside a video stream (e.g. character clothes, cars, other props) to make them clickable and then link that click to some kind of information or service.

CUSTOMER BASE: Ooyala works with brands and video content providers including Armani, AOL/Bebo, Joost, TV Guide, National Geographic Japan, NTT, Electronic Arts, Glam Media, Slide, Televisa, Wenner Media.

OWNERSHIP: privately owned with investment from Sierra Ventures. $10m raised.

UNIQUE POINT: True to its Google DNA, ooyala is betting that video is best monetized through micro-transactions that can be automated and quickly scaled. The cloud-based infrastructure reduces the overall cost of managing, distributing and ad serving for video. By making video clickable and able to link out to service, ooyala opens up new avenues for vacuuming digital pennies and nickels.

CLOUD ANGLE: Ooyala's management have been cloud fluent since day one, having come from one of the most cloud-intensive companies in the world. Amazon S3 = Google GSS, Amazon EC2= Google Borg. But ooyala isn't banging the cloud drum loudly. The cloud infrastructure is simply a better way to gain scale quick to offer more effective monetization of video assets and advertising. Sounds familiar, yes?

Playfish
125 Kensington High Street
London W8 5SF
http://www.playfish.com

SUMMARY: Playfish is one of the world's largest and fastest growing social games companies. The company's games target friends who want to play together over social and mobile platforms such as Facebook, MySpace, Google, Bebo, iPhone and Android. Each of the company's nine game titles has been a top 10 hit on Facebook. Pet Society is the most popular game, played by over 11 million people each month. Playfish has offices in London, San Francisco, Beijing and Tromso, Norway.

THE PROBLEM: The games industry is in the midst of a well documented problem with its business model. The cost of producing a top tier console or PC game is going up while the addressable market of people willing to shell out $50 a pop shrinks just as fast. Social games attempt to jump out of this problem by making gameplay part of a bigger social behavior that's monetized differently.

THE APPROACH: Playfish focuses the gaming value proposition away from selling copies of a game title and instead focuses on the social interaction among friends via a game. The company looks to Facebook and other social networks as the platform for its games rather than the gaming console or PC makers. The Playfish business model is based on in-game micro-transactions (e.g. buying accessories for characters) as well as in-game advertising.

CUSTOMER BASE: Since launch in October 2007, Playfish has distributed over 100 million games via the Internet and social networks. It's monthly active user base is around 40 million people. Daily access to Playfish games is about 9 million users. The geographic spread is global. The portfolio consists of 9 games, many of which are based around some kind of challenge for a group of friends or peers.

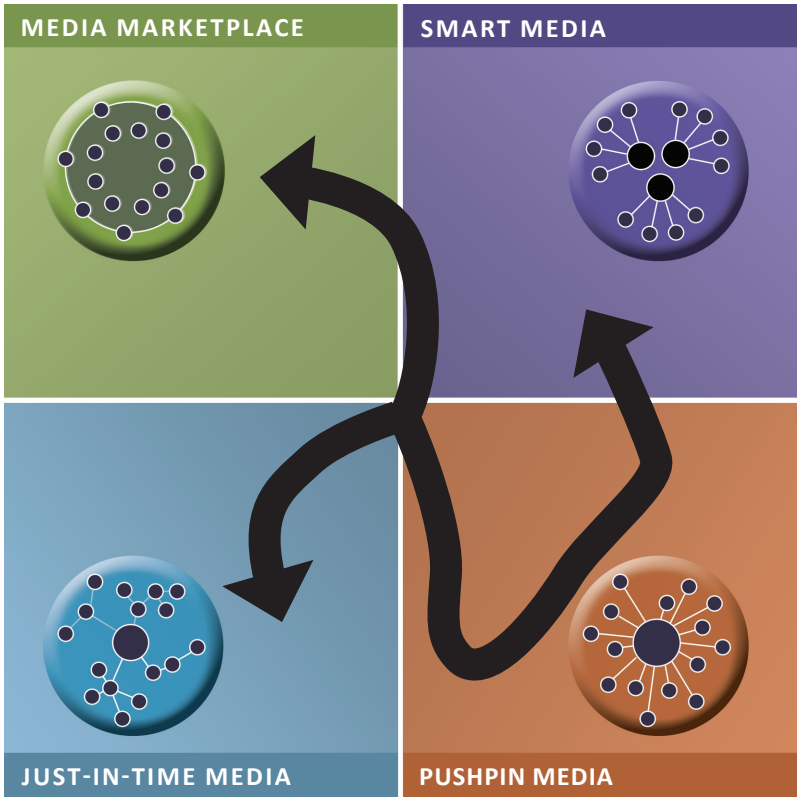
OWNERSHIP: Acquired by Electronic Arts (EA) in November 2009 for $300 million plus performance incentives.

UNIQUE POINT: Playfish is attempting a new business model for gaming in which the game is a service rather than a product. The company assumes its users will access Playfish games in a similar manner as they access any other cloud-based service such as Gmail. Additionally, Playfish games often keep the same session even if the user toggles between a PC and a mobile client to access the service. Indeed, mobile takes a prominent position. Playfish launched iPhone and iPod touch versions of its most popular game, "Who has the Biggest Brain" at the South by Southwest Festival in March 2009.

CLOUD ANGLE: Playfish is based 100% in the cloud. The company's staff operates on laptops. Playfish uses the full suite of Amazon services (S3, EC2) and is one of the largest users of AWS CloudFront, which it uses to push large Flash files around the fixed and mobile Internet for players to access.

Four Scenarios About Media 2020

It's like trying to predict back in 1910 the impact of the automobile on society – the highway system, gasoline refineries, motels instead of hotels, new dating patterns, increased social mobility, commuting to work, the importance of the rubber industry, smog, drive-thru restaurants, mechanized warfare, and on and on. The net will bring more than quantitative changes, it will bring "qualitative" changes. Things that were impossible will now become inevitable. – LARRY LANDWEHR, 1993



This section offers four scenarios about media in 2020. These scenarios are NOT predictions. They aren't forecasts.

Instead, scenarios are stories based on some plausible technical and organizational changes going on 2010. They provide a common context and lexicon for groups representing different disciplines and duties to imagine the future in a structured fashion.

When you read these scenarios, keep in mind three main points that will affect your organization:

A. How attainable is a scenario?

B. What is the positive or negative impact of a scenario if it happens?

C. How persistent or exclusive would a scenario be to the entity that controls it in 2020?

Don't think that these four scenarios cover all the ground for media and cloud computing in 2020. Don't think that they are in direct competition with each other.

Just use them as crude axes for chopping through the river of news and information crossing your desk every day. Regardless of how accurate they may or may not be, the actual future will undoubtedly hold many more surprises.

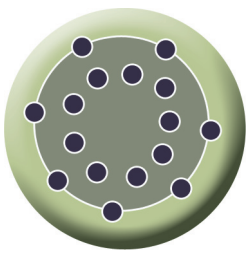
SCENARIO A:

Media Marketplaces 2020

Most professional media and marketing content has been subsumed into the larger world of electronic commerce.

MAIN THEME: Most professional media and marketing content has been subsumed into the larger world of electronic commerce. Products and services now come with directly embedded media or offer digital tokens for consumers to exchange for the media of their choice.

MARKETS: There are no longer separate media markets for readers, viewers, listeners or game players. There are only customers who access media as part of a larger purchase. They buy goods and services that ship with embedded media and marketing or tokens to exchange for media access. If you buy a microwave oven at a retailer, it comes with one year's access to the food channel along with a recipe database. The oven can send a shopping list to your printer or mobile phone. It even adjusts its settings for that perfect Beef Wellington. The traditional advertising market, in which commercial messages played in a pod as part of a scheduled program, is effectively dead. Titles and artists have trumped networks and channels. In 2020 there are any number of branded products and services that regularly serve an audience of 2-3 million with sponsored content and messages. Brands have gone direct and haven't looked back. The twin forces of user generated media plus direct-to-consumer marketing by brands have slashed the relative bargaining power of producers and distributors of content. Cable TV and IPTV operators have exited the content business to concentrate on connectivity and 3rd party billing services. Revenue splits with brands and retailers have cushioned the blow for pipe owners who had to cut well over half of their 2010 level workforces.



INDUSTRY STRUCTURE: New breeds of media aggregators and customizers work with content originators and brands in place of traditional broadcast networks. The intermediaries have focused their energies on two main objects: world-class supply chain management for media plus creating a flawless customer experience. Customer knowledge on the front-end combined with orchestration of media assets on the back-end have helped the players work with brands to go direct-to-consumer. Brand power has accrued to the "service integrator" rather than the systems integrator. The bandwidth providers, like search engine owners, knew a lot about their customers' behavior. But they couldn't convert that insight into new monetizable products. Like Xerox PARC, which gave the world so many pioneering PC innovations, Google and other search engines consistently dropped the ball at the goal line of consumer media and services commercialization.

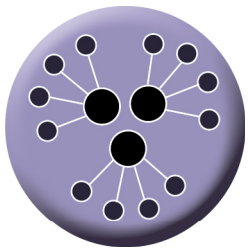
TECHNOLOGY: The cloud is the digital assembly line that enables mass customization of media and marketing to work in tandem with mass customization of products and services. Customers pick and choose among media titles and how they want to experience them based on their product or service purchases. The new breeds of content intermediaries operate the equivalent of Software as a Service (SaaS) for media. There is a shared, multi-tenant base version of content and code for a given media property. Wrapped around that core resource are various service layers that allow customization for different devices, viewing angles, tie-ins with other media or marketing experiences, and ability by consumers to roll their own versions of the content for sharing with friends. Metadata related to business rules and charging formulas are housed within the base code to ensure that revenue splits are paid up and down the value chain. The cloud also houses the integrated CRM and Business Intelligence (BI) applications that sift through mountains of purchase behavior to draw inferences that not only feed into product design but also *plot design*.

SCENARIO B:

Smart Media 2020*

REGULATION: Because the media intermediaries have wisely opted to focus primarily on entertainment content as opposed to news or educational content, regulators have used a light hand. At one point, there was a movement to create the equivalent of a government-run escrow service for consumer usage and preference data. However, the howls of protest from both privacy advocates and big business lobbies beat it back. For now, regulators are taking a "wait-and-see" approach. They haven't tipped their hand regarding what kind of security or other breaches will invite intervention and as of 2020, the media intermediaries haven't yet had the kind of blow up that would warrant it. Still, these are still early days.

MAIN THEME: Media providers and marketers have dispensed with the pretense that they can target mass media in a fragmented world. Networks and channels have largely disappeared. Media and marketing messages are tagged with information and function to become the equivalent of "smart bombs" that seek out the right audience regardless of channel or device.



Media and marketing messages are tagged with information and function to become the equivalent of "smart bombs" that seek out the right audience regardless of channel or device.

MARKETS: Performance finally won over exposure as the metric that counts most for media and marketing. No serious executive in the media or marketing industries cares anymore about CPMs. Technology has pretty much eliminated windowing as a media revenue strategy. Every release is, by definition, a global release for all platforms and devices. Instead of placing media or marketing messages on either pay-per-click networks (PPC) or specific publisher sites, content originators simply release their media or ads on the cloud. The media is tagged with metadata supplied by the originator or advertiser along with functionality. Once the media is let loose, fans or other audience types can simply cut and paste the content to their preferred media device, their social networking page, their own site, or any other means of experiencing and sharing the content with their peers. The media is paired with sponsored ads at the outset. However, if the pairing doesn't perform, new ads are wrapped with the media. The ads contain metadata supplied by the marketer that documents who the marketer wants to reach, what kind of environments they want to be in (or not be in such as racist or porn destinations). Most important, the smart media and its marketing payload specify how much they are willing to pay a "destination" (e.g. a social tastemaker, a local media source, a deli owner with a multimedia kiosk at the corner of 57th Street and Broadway). The media tracks its progress as it winds through all the various distribution choices and "phones home" information about who has pulled it down to its originator, how many interactions it has generated and how much juice is left in the account. The marketing payload propagates until it runs out of money, then….poof! However, if the media is getting traction, the marketer can re-fill the tank for another ride.

INDUSTRY STRUCTURE: The opportunities for fraud are legion with smart media. Until Verisign's successor Zergo cracked the security challenge to allow media to circulate freely with money attached, no one took the idea of smart media seriously. After the security challenge was resolved in 2016, a furious race developed to become a tastemaker/curator for various media verticals like travel & leisure, food, wine, autos and so forth. Aside from security and distribution, another major industry effort involved media configuration and matching with marketing. In this area, search engine players like Google re-oriented many of their algorithms away from web site links by webmasters and increasingly toward trying to find tastemaker-son social networks and other destinations.

TECHNOLOGY: To the end-user, smart media looks like a complete product. Actually, all that circulates is metadata. The core content and marketing payload isn't configured until the end-user selects and customizes it. Then, the "real" media experience is configured in the cloud and delivered as if part of a single session. In the cloud, content originators compose the equivalent of a chassis upon which a clutch of different plot and character variations can be wrapped. Instead of a restricted number of marketing messages bundled with the smart media, cloud infrastructures house templates with a wide variety of permutations on the color scheme, presentation, call-to-action and other real-time modifications to a given commercial offer. Once the basic media design is uploaded to the cloud, semi-intelligent agents offer the originator a palette of story options such as language localization, product tie-ins, even changes in characters that can be exposed to different audiences depending on the rules the originator prefers. Fully integrated with device libraries for optimal playback on the vast majority of media clients, the cloud-based media contains hooks into CRM and BI applications that are made available to content originators and marketers under a SaaS model.

REGULATION: At the outset, regulators were snarled with taxation and cultural content issues. Especially in Europe with its history of public funding for creative arts, the tax and culture issue loomed large. Regulators split the difference with industry by enacting a single excise tax, albeit low, to apply to all media and marketing messages. The excise tax is paid as part of the upload fee for placing the content in the cloud in the first place. In return, the upload date and data is time stamped by authorities to prove ownership of the content.

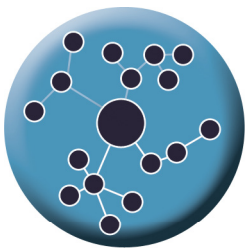
*This scenario was inspired by a 2004 post on John Battelle's blog entitled "Sell Side Advertising": http://battellemedia.com/archives/000844.php

SCENARIO C:

Just-in-Time Media 2020

MAIN THEME: The action in media has moved to how fast a provider can customize media and ads at the margin. Scale in media has become a massive optimization game where players compete to orchestrate massive ecosystems to deliver an individualized yet participatory experience.

MARKETS: Media end-users now expect to customize their entertainment experiences before they order them. Do you want to see a mixed martial arts bout from the point-of-view of the corner man? No problem, please mark that box when you order UFC 5000. For sports, concerts, literary gatherings, and other events, consumers can customize their media experience beyond the standard surround sound and individual camera angles, to include overlays and remixes of their own images and content with that of the performance. Now, you can watch and participate simultaneously. A re-mastering of Peter Jackson's 2003 Lord of the Rings Movie "The Return of the King" proved to be a major 2014 hit through its use of virtual reality technology to allow home theater fans to participate in the Battle of Pelenorr Fields. By connecting their next generation Wii device to the movie, consumers could immerse themselves in the action as an ax-wielding dwarf of a slender but deadly elf, watching and competing with other fans in real-time as the movie played. Studios and marketers returned to their vaults to turn more movies into MMORPG experiences similar to how they embraced the digital coloration craze of the early 2000s for their classic black and white properties. Marketers have proved adept at configuring personalized ads and micro-transactions for participatory digital theater. Weapons, clothing, and many other accoutrements for completing one's immersion in a media experience are drawn from a similar core implementation that is rapidly customized to the individual at that moment. Netflix has pioneered a recommendation engine that doesn't simply advise a consumer about a new film but suggests digital artifacts, services and social connections that can enhance the experience.



The action in media has moved to how fast a provider can customize media and ads at the margin.

IMAGES SOURCE: www.futurelab.net

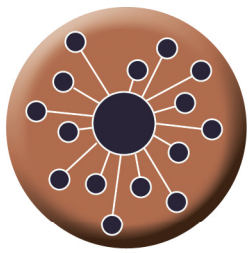
INDUSTRY STRUCTURE: The cloud has become the factory system for JIT media and marketing experiences. Media creators and marketers have borrowed organizing principles from manufacturing disciplines like LEAN and Total Quality Management (TQM) to wrest mass media and advertising production and distribution from its craft-based roots. Huge cloud-based intermediaries provide much of the production and distribution firepower previously associated with studios and media holding companies. These digital media equivalents of OEM contract manufacturers are responsible for executing the design vision of media and marketing providers, who provide a digital blue print of a media experience. The actual level of customization isn't as great as most consumers are led to believe. Visual angles, immersive reality, insertion into plots, different color schemes and language localization have become nearly standard offers for media creators. As the cloud infrastructure providers gain more experience with customizations, they are sharing and sometimes re-selling that information back to content creators to feed into their media design. Not surprisingly, the cloud ecosystem is as complicated as the media and marketing ecosystem. Numerous "middle layer" players have sprung up to coordinate the creative vision of the originator, the customization demand by the consumer, and the necessary technical and distribution requirements to pull things off.

SCENARIO D:
Pushpin Media 2020

TECHNOLOGY: JIT Media would not exist with extensive cloud computing. Variability in demand on the consumer side is more than matched by an entirely new swath of analytics demanded by media and marketing clients. It's no longer good enough to know who was watching but how they participated in the story. Along with the formidable cloud infrastructure for serving up custom experience is an equally impressive cloud infrastructure for customizing consumer electronic equipment. Since the Consumer Electronics Show (CES) of 2014, new electronics come to consumers more or less naked in terms of features and functions. Hardware consists of large screens, robust speakers, universal remotes and/or specialized game and participation devices, plus a broadband connection. The real action takes place when consumer configures the performance and features of their device in the cloud after purchase. Then the system comes alive.

REGULATION: Privacy has become a key driver and inhibitor to JIT Media. Without a rich customer profile, the ability to customize in real-time becomes exponentially more difficulty. Moreover, some of the richest profiling occurs when consumers aggregate their experiences with the same media. These group profiles have emerged as one of the most contentious ownership questions to cross the JIT media ecosystem precisely because they are probably the most valuable sort of information for both media creators and marketers. Regulators are still digesting the possible implications.

MAIN THEME: Augmented reality has finally won over virtual reality to create a new media market. Instead of trying to substitute a virtual world for the "real" world, media providers and marketers have discovered that a pinch of media in the right physical *place* has opened a huge entertainment and advertising opportunity.



MARKETS: Innovators hopped over the previously unbridgeable divide between the virtual and real worlds by declaring victory and moving on. Rather than using computers to perfect the illusion of a physical world experience, media providers are making the "real" world a little more fun and entertaining with computer generated media and stories. Instead of a consumer physically going to a theme park to experience an entertainment fantasy, the entertainment fantasy comes to the customer as superimposed media on a real physical location. Classical music fans can walk the streets of Liverpool as the Beatles lived it, with location specific music and information tied directly to their handheld devices. Concert halls have upended their traditional ban on recording by offering ticket holders the opportunity to wirelessly connect their mobile media devices to the venue's network to capture and create their own unique media experience to enjoy at home afterward. Every new film shot now embeds location data in its bonus materials so that fans can experiment with new plot angles or role play exercises when they visit the actual location. Looking through lens or layers on their mobile handsets and embedded multimedia implants, media fans have redefined "out-of-home entertainment". The most threatened ecosystem players are amusement parks, themed entertainment attractions and large scale arcades. Attendance at Disney Parks and Resorts has taken a nosedive as have other location-based entertainment providers. Micro-transactions based

Media providers and marketers have discovered that a pinch of media in the right place has opened a huge opportunity.

on electronic coupons have provided the primary business model for location-based media. These coupons not only unlock local commerce opportunities. They also feed back into lean back media experiences that will be made available outside of the physical location once the fan leaves.

INDUSTRY STRUCTURE: In a world of unlimited digital shelf space, venue owners have discovered that they have one of the only remaining scarce resources. Revenue share deals between media providers and private locations are the norm. In the case of public venues like parks and streets, the structure is more complicated. The Yellow Pages industry has re-invented itself. No longer selling space in big yellow books, the new local marketing company puts together a palette of media, marketing and offers based around physical locations. Transportation companies such as taxi fleets and bus companies have found the ultimate signage opportunity with marketing displays that change according to filters that are broadcast by consumers and their groups. Media providers, in turn, have discovered that tying media experiences to actual locations provides "roll-your-own" audience research. Rather than gathering focus groups in theatres, savvy media producers release beta copies at select physical locations world-wide to have a dynamic gauge of audience sentiment. Technical service providers controlling these wide area media experiences are reaping economies of scale plus audience insight that they share or sometimes sell back to media creators and marketers. Most of these players grew out of the mobile communications providers of the 2000s.

TECHNOLOGY: The most important application isn't media or ad serving. It's the algorithms and applications that match audience filters, preferences and privacy settings to what gets served to their AR devices. World cities such as Paris have streets which have inspired multiple films, songs and books. The task is to both scale and customize individual media experiences without having fans crowd out each other on the physical space. Cloud-based mobile service players provide most of the technical firepower to media creators and marketers.

REGULATION: Of all the media and marketing innovations, pushpin media has received the most attention from regulators. Privacy and safety concerns were raised to a fever pitch after six music tourists were mowed down by a disturbed individual who had hacked into their preferences and knew where they would be at a precise moment. Regulators don't have near the power to control private venues. However, it's a foregone conclusion that existing laws and procedures for granting live performance permits to public spaces are being extended to become simulated performance permits.

This page is intentionally blank.

Tossing the Rulebook

If you're losing at a game that you can't afford to lose, change the rules – DAN GEER

Today's incumbent media and marketing organizations are losing.

They're laying off talent, spinning out divisions, shutting down titles, and some have referred to their customers and partners as thieves. Scratch a traditional media or marketing executive and you'll probably find someone whose attention is consumed by threats, some of which are very real.

However, media and marketing's real enemy isn't file sharing or free content, government regulation or Google for that matter. The ultimate threat facing the current media and marketing industry is complexity. An inability to handle complexity is killing many of today's leading media and marketing organizations.

If you go to a media conference now, you'll see a conga line of experts talking about the fragmentation of media. Like any buzz word, "fragmentation" conjures an image of being a serious, scientific insight when it actually doesn't mean a damn thing. Media content has not fragmented. Media audiences have moved on. They've stopped adapting their attention around someone else's programming schedule and business model.

Moreover, consumers are way ahead of the media industry in handling the demand-side of complexity. Do you remember all those breathless predictions during the 90s about intelligent agents scouring the web for personalized content and services? Well, we now have a huge, robust set of intelligent agents that are getting very good at recommendations.

They're us.

We've become our own intelligent agents through social networking. There's now enough density in Facebook and other social worlds to make complex adaptive demand for media a reality.

Doesn't it follow, then, that the current media industry must develop complex adaptive supply?

If you forget every topic covered in this work about cloud computing, it's okay so long as you remember that the cloud is the best, single platform for exchanging and evolving media with multiple audiences who are using multiple devices in multiple contexts under multiple business models.

But even with that awesome capability, nothing happens unless you have a gripping story to tell.

Keep hold of that.

Acknowledgements

The following group of people have lent their time and insight speaking to me about the intersection of media, cloud computing and policy over the past year. Through interviews for the Media Dojo blogsite, participation in industry panels, my Knight Foundation fellowship, telephone conversations and email exchanges, they all have helped me try to navigate the rabbit hole. Naturally, opinions, all mistakes, faulty logic and salty language are my responsibility alone.

I give my sincere and heartfelt thanks.

Santiago Becerra, CEO of Roambi

Peggy de Bona, Acting Director of the Knight Center for Specialized Journalism, University of Maryland

Simone Brunozzi, Web Services Evangelist for Europe, Amazon Web Services

Srini Dharmaji, CEO of Goldspot Media

Alan Davidson, Director of Public Policy and Government Affairs, Google

Sebastien de Halleux, co-founder and COO, Playfish

Andrew Heyward, Senior Advisor, Marketspace LLC

Jayant Kadambi, co-founder and President, YuMe

A.J. Kinter, Director of East Coast Sales, Tumri

Sean Knapp, co-founder and CTO, Ooyala

Tariq Krim, founder and CEO of Jolicloud

Jeff Lawson, CEO of Twilio

John Lilly, CEO of Mozilla

Darcy Lorincz, CEO of Origin Digital

Dale McCrory, Product Director, Exact Target

Brock Meeks, Director of Communications, Center for Democracy and Technology

Ryan Nichols, Vice President of Product Development, Appirio

Beth Noveck, Deputy Chief Technology Officer, White House Office of Science and Technology Policy

Scott O'Neal, Deputy Assistant Director Cyber Division, FBI

Tom Shields, CEO of YieldEx

Jay Stanley, Public Education Director, ACLU Technology and Liberty Program

Doug Whatley, CEO of Breakaway Games

Why Call It a Dojo?

A dojo is where people go to learn, to take risks, to make mistakes, to keep showing up and advance a little more each time. I've been going to the dojo for 27 years now, through three Japanese martial arts. Each art has unique movement, technique, power, speed, accuracy and tactics.

However, the target for all these martial arts is a state of mind that some Japanese call Banpen Fugyo, ("10,000 Changes No Surprises"). Banpen Fugyo is the ability to adapt swiftly and decisively to chaos — both the good and the bad kind.

Until November 2008, I was a professional fortune teller for some of the world's top media and corporate brands. I concentrated on the intersection of technology, media and economics — making forecasts around areas where advanced technologies mixed with the media and marketing craft.

I quit my job for two reasons. First, I believe that media and marketing are poised to flip from being forecast-driven to being demand-driven industries. Second, I believe that cloud computing will be the institution through which this dramatic environmental shift happens. Chaos is virtually guaranteed.

You can't predict how chaos will create extraordinarily quick value or destruction. You can't predict its exact timing.

But you can prepare by opening your eyes and your mind. Through published research along with invitation-only virtual and physical meetings, Media Dojo aims to be a community-based training hall for a new breed of media and marketing innovators who make constant change a familiar friend.

Contact: John du Pre Gauntt,

Mobile: +1-206-618-8131

john.gauntt@media-dojo.com

www.media-dojo.com

Media Dojo

Published by Media Dojo | March 2010 | www.media-dojo.com

CONCEPT / RESEARCH / WRITING

John du Pre Gauntt

Media Dojo

Mobile: +1-206-618-8131

john.gauntt@media-dojo.com

DESIGN / PRODUCTION

Laura Urban Perry

Laura Urban Perry Design

+1-206-323-8966

laura@lauraurbanperry.com