

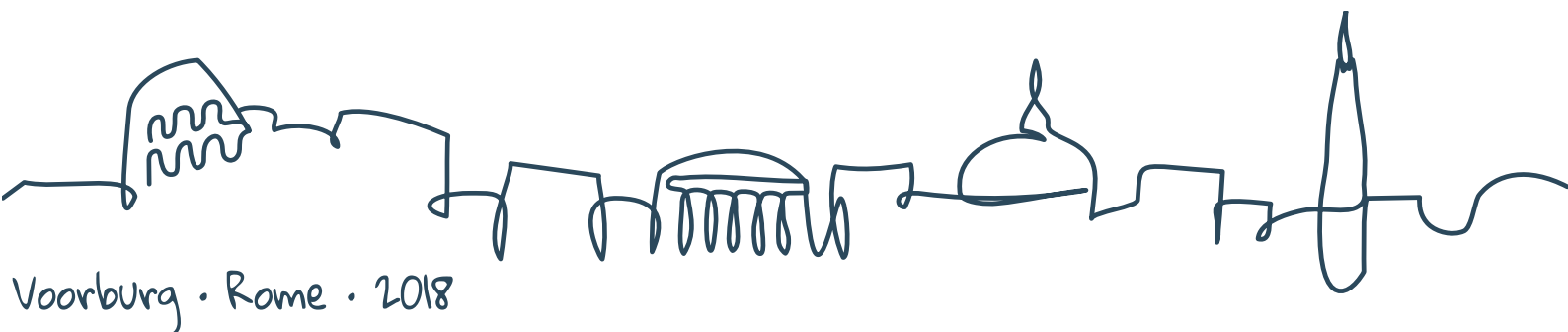


ROME
24 | 28
SEPTEMBER
2018

ISIC 6311 Data processing, hosting, and related services

Andrew Baer, Thomas Hill, and Thomas Ponikowski
Census Bureau – United States

Bonnie Murphy, Melanie Santiago, and Jason Carnival
Bureau of Labor Statistics – United States



ISIC 6311 Data processing, hosting, and related services

1. Descriptions and characteristics of the industry

1.1. Definition of the industry

ISIC class 6311 Data processing, hosting and related activities includes the following activities:

- Provision of infrastructure for hosting, data processing services and related activities
- Specialized hosting activities such as:
 - Web hosting
 - Streaming services
 - Application hosting
- Application service provisioning
- General time-share provision of mainframe facilities to clients
- Data processing activities:
 - Complete processing of data supplied by clients
 - Generation of specialized reports from data supplied by clients
- Provision of data entry services

This ISIC class concords with NAICS industry 518210, which similarly comprises establishments primarily engaged in providing infrastructure for hosting or data processing services.

Most of the services classified in this industry are currently marketed in the U.S. as cloud computing services. The U.S. National Institute of Standards and Technology defines cloud computing as, “a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction.”¹

There are three primary categories of cloud computing services: software as a service (SaaS), platform as a service (PaaS), and infrastructure as a service (IaaS). SaaS allows users to access software that is hosted remotely on the service providers’ infrastructure. PaaS allows developers to remotely access a computing environment where they can develop and maintain applications. Finally, IaaS allows users to remotely access storage and server capacity.

In the U.S. this industry includes companies that primarily provide remote access to software for which they do not hold copyright. If the company does hold copyright for most of the software that is provided remotely, the activity is classified in NAICS industry 511210, Software Publishers. A similar distinction holds for content streaming services. The service of providing infrastructure

¹ <https://csrc.nist.gov/publications/detail/sp/800-145/final>

for hosting streaming content is included in Data processing, hosting, and related services. However, content owners that stream original or purchased content on a subscription or advertising supported basis are classified elsewhere in the Information sector.

1.2. Market conditions and constraints

The following table provides size statistics for the U.S. Data processing, hosting and related activities industry based on the 2012 U.S. Economic Census.

TABLE 1
Size of NAICS industry 51821 in U.S. in 2012 and 2007

	Information Sector	Information Sector	51821	Percent of Information sector	51821	Percent of Information sector
Year	2012	2007	2012		2007	
Establishments	138,341	141,566	14,689	11%	17,129	12%
Annual payroll (\$1,000)	269,069,624	228,836,587	41,126,425	15%	26,428,444	12%
Revenue (\$1,000)	1,238,463,251	1,072,342,856	107,994,265	9%	66,651,853	6%

Establishments classified in NAICS 51821 contributed approximately 9% of U.S. Information sector revenue in 2012, up from 6% in 2007. The decline in the number of establishments is reflective of a continued trend of consolidation among the industry's largest businesses. This industry has continued to grow substantially since 2012. The 2016 U.S. Annual Services Report indicates that revenue grew 10.3% and 7.9%, in 2016 and 2015, respectively².

The following table provides industry data on firm concentration from the 2012 Economic Census.

² <https://www.census.gov/programs-surveys/sas/newsroom/updates/latest-sas.html>

TABLE 2
Concentration of NAICS industry 51821 in U.S. in 2012

Firm concentration	Number of establishments	Receipts (\$1,000)	Revenue of largest firms as percent of total revenue (%)	Annual payroll (\$1,000)	First-quarter payroll (\$1,000)	Number of paid employees for pay period including March 12
All firms	14,689	107,994,265	100	41,126,425	10,865,590	497,300
4 largest firms	1,063	17,149,626	15.9	5,524,686	1,360,048	85,098
8 largest firms	2,036	24,458,671	22.6	7,556,578	1,915,042	113,044
20 largest firms	2,416	35,379,028	32.8	11,342,115	3,077,387	144,476
50 largest firms	2,926	48,956,046	45.3	16,523,947	4,501,733	197,180

The Data processing, hosting, and related services industry was not very concentrated in 2012. The four largest firms earned almost 16% of revenue while the 50 largest firms comprised only 45.3% of the industry revenue total. There are many large to medium sizes firms that provide these types of services in the United States.

The degree of concentration in this industry has increased in recent years. According to Synergy Research Group, the top 4 U.S. companies providing cloud infrastructure services (IaaS and PaaS) had a collective market share of more than 50% in the 2nd quarter of 2017³.

The level of regulation in this industry is moderate, but is increasing over time as a result of greater access and storage of personally identifiable information and intellectual property. There is pressure in the United States to create additional legislation to oversee online operations to protect intellectual property. Individual states have a variety of laws concerning privacy protections for personal information, particularly personal medical information that is stored electronically in the cloud.

³ <https://awsinsider.net/articles/2017/08/01/aws-market-share-3x-azure.aspx?m=1>

The following product lines generated revenue in NAICS 518210, Data processing, hosting, and related services for the 2012 Economic Census.

TABLE 3
NAICS industry 51821 product lines in U.S. in 2012

Product Line Description	Receipts (\$1,000)	% of Total
Industry total	107,994,265	100
Application service provisioning, with or without integration of related services	24,294,678	22.5
Business process management services, including financial, human resources, supply-chain, customer relations, and vertical markets management	18,385,487	17
Business process management services - financial	6,519,566	6
Other business process management services	7,061,972	6.5
Website hosting services, with or without integration of related services	12,081,543	11.2
Website hosting, with integration of related services	11,064,365	10.2
Data management services	15,240,362	14.1
Information technology (IT) technical support services	10,062,127	9.3
All other receipts	5,726,240	5.3

The following table provides the largest North American Product Classification System (NAPCS) product lines that appear on the 2017 Economic Census.

518210 - Data processing, hosting, and related services		
NAPCS	DESCRIPTION	DEFINITION
7014250000	Web site hosting services	<i>Providing the infrastructure to host a customer's web site and related files in a location that provides fast, reliable connection to the Internet.</i>
7014250003	Web site hosting services without integration of related applications	<i>Providing the infrastructure to host a customer's web site and related files in a location that provides fast, reliable connection to the Internet. The service is limited to storage on a single server, in either shared or dedicated capacity, without the service provider managing or integrating software applications. Software hosted on the server is the client's responsibility. Service level guarantees are standardized and limited in scope.</i>
7014250006	Web site hosting services with integration of related applications	<i>Providing a bundled service package that consists of the hosting and management of the web site and related applications. An important characteristic of this service is the promise of a secure and reliable site and Internet connections that can be quickly scaled to accommodate variations in traffic use. Package frequently includes consulting, customization and systems integration services. Applications are frequently E-commerce related and enable on-line storefronts, shopping carts and catalogs with advanced and complex features such as order processing, fulfillment, procurement, invoicing, transaction processing, customer relational management and back-end database and data warehouse integration and migration services.</i>
7012028000	Web site design and development services	<i>Designing the structure and content of a web page and/or of writing the computer code necessary to create and implement a web page.</i>
7014275000	Application service provisioning, including cloud services	<i>Providing leased software applications from a centralized, hosted, and managed computing environment</i>
7014275006	Application service provisioning with integration services	<i>Providing leased software applications from a centralized, hosted, and managed computing environment with integration to the systems and infrastructure of the client. Consulting, customization and systems integration services are frequently bundled with the hosting and management of the application.</i>

7014275009	Application service provisioning without integration services	<i>Providing leased software applications from a centralized, hosted, and managed computing environment where the leased application is not customized and not integrated with other applications of the client. The application is usually accessed over the word-wide web. A common example is office suite software applications.</i>
7014300000	Business process management services	<i>Providing a bundled service package that combines information-technology-intensive services with labor (manual or professional depending on the solution), machinery, and facilities to support, host and manage a business process for a client.</i>
7014300003	Financial business process management services	<i>Providing a bundled service package that combines information-technology-intensive services with labor (manual or professional depending on the solution), machinery, and facilities to support, host and manage a financial business process for a client, such as financial transaction processing, credit card processing, payment services, and lending services</i>
7014300006	Human resources business process management services	<i>Providing a bundled service package that combines information-technology-intensive services with labor (manual or professional depending on the solution), machinery, and facilities to support, host and manage a human resource business process for a client, such as benefits administration, payroll processing, and personnel administration.</i>
7014300009	Supply-chain business process management services	<i>Providing a bundled service package that combines information-technology-intensive services with labor (manual or professional depending on the solution), machinery, and facilities to support, host and manage a supply chain management business process for a client, such as inventory management, procurement services, logistics services, production scheduling and order processing.</i>
7014300012	Customer relations business process management services	<i>Providing a bundled service package that combines information-technology-intensive services with labor (manual or professional depending on the solution), machinery, and facilities to support, host and manage a customer relations management business process for a client, such as help desk, call center, and customer service.</i>
7014300015	Vertical markets business process management services	<i>Providing a bundled service package that combines information-technology-intensive services with labor (manual or professional depending on the solution), machinery, and facilities to support, host and manage a vertical market business process for a client.</i>

7014300018	Other business process management services	<i>Providing a bundled service package that combines information-technology-intensive services with labor (manual or professional depending on the solution), machinery, and facilities to support, host and manage other business processes for a client.</i>
7014325000	Collocation services	<i>Providing rack space within a secured facility for the placement of servers and enterprise platforms. Includes provision of space for the client's hardware and software, connection to the Internet or other communication networks, and routine monitoring of servers. Clients are responsible for the management of the operating system, hardware, and software.</i>
7014350000	Data storage services, including remote backup, storage of computer data, and cloud services	<i>Managing or administrating the storage and back-up management of data such as remote back-up services, storage, or hierarchical storage management (migration)</i>
7014375000	Data management services	<i>Providing ongoing management and administration of data as an organizational resource. May include performing data modeling, data mobilization, data mapping/rationalization, data mining and system architecture</i>

Software as a service (SaaS) is a service that allows users to run existing applications available over the internet. The service is an independent platform, does not require software on the user's computer, runs on a single instance of the software, and is accessible via web browsers or lightweight client applications. Some positive elements of this service is that its universally accessible for any platform, it is effective for collaborative working, the vendor provides the software tools, and the applications are either free or low cost and paid on a subscription basis. Some negative elements are that users are susceptible to browser portability issues and inconsistent internet performance.

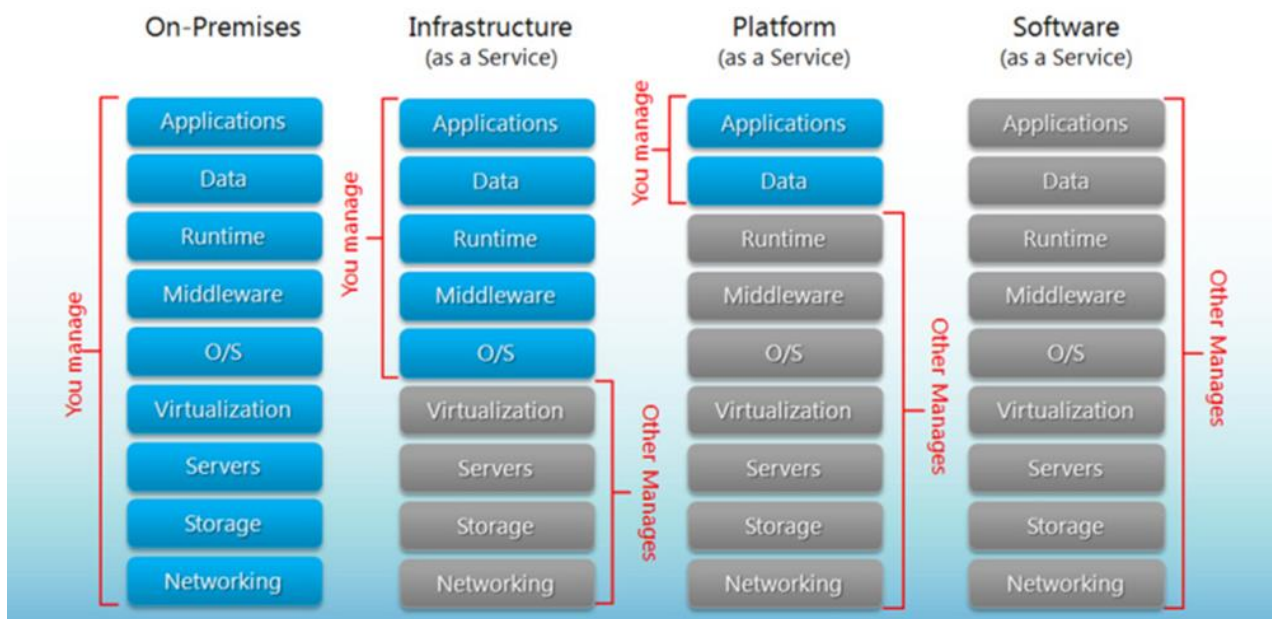
Platform as a service (PaaS) provides a programming language execution environment, an operating system, a web server and a database that software developers can use to run, build, and compile programs without working in the underlying infrastructure. Users can manage data and application resources while the vendor manages the remaining resources and provides the user the environment and tools for creating new online applications. Some positive elements of platform as a service is that it has cost effective rapid deployment. A negative element is that developers are limited to the provider's language and tools and migration problems could occur during deployment.

Infrastructure as a service (IaaS) allows users to run any desired application on their supplier's cloud hardware of their own choice. This offers users all computing resources, architecture and infrastructure, but in a virtual environment so that multiple users can access them. The vendors are responsible for managing data storage, virtualization, servers, and networking.

The following diagram shows the components of computing that are managed either by the cloud vendor or the customer. The first column shows the traditional model where the customer

manages all aspects of the computing environment using their own infrastructure, while the latter columns describe cloud computing service models.

FIGURE 1
Cloud computing models vs. traditional on-premises enterprise IT management



Source: <https://www.hostingadvice.com/how-to/iaas-vs-paas-vs-saas/>

There are four different deployment models for cloud services:

- Private cloud – cloud infrastructure used solely by one organization
- Public cloud – cloud infrastructure shared by many users
- Community cloud – cloud infrastructure shared by a group of affiliated users; and
- Hybrid cloud – a service offering where certain data and applications are placed on a private cloud while others are placed on a public cloud

The Data processing and Hosting Services industry is attractive to new and large companies. New companies demand these services to avoid the high costs associated with investments into hardware and software. Large companies require these services because the cost of internal data management is greater than that of outsourcing. Large companies have found that the resources required to manage data increased rapidly as the amount of data that needs to be analyzed grows. Over the last decade, the number of companies capable of managing data of this magnitude within their own infrastructure have declined. This has led to consolidation amongst the largest players as they continue to acquire smaller companies that provide specialized services.

2. Turnover/output measurement

2.1. General framework

The U.S. Census Bureau provides output data for the Data Processing, Hosting, and Related Services industry as part of the Quarterly Services Survey, Services Annual Survey, and the Economic Census conducted every five years.

The U.S. Bureau of Labor Statistics (BLS) Produce Price Index program uses Economic Census data to update industry and product index weights.

BLS also uses industry turnover and expense data as a main data source for service industry labor productivity and cost measures.

The Bureau of Economic Analysis (BEA) uses turnover data for annual industry accounts, including both GDP by industry and annual input-output accounts. The Economic Census data is used as the measure of output in the benchmark input-output tables.

Data Processing, hosting, and related services output is primarily the provision of remote access to applications, computing platforms, and computing infrastructure hosted by the service provider.

2.2. Measurement issues

The U.S. Economic Census measures company output at the establishment level. This creates a significant challenge for this industry since these companies do not typically record hosting and related activity turnover to a particular location. For example, a cloud computing company may offer on-demand remote access to server processing. Where should the turnover generated from this activity be recorded? It could be allocated to the location of the server facility, but cloud infrastructure is typically designed to optimally allocate server capacity amongst many server locations. So the service may be provisioned from different server facilities at different points in time. In practice, US Census Bureau staff often allocate turnover to establishments based on the proportion of total company payroll at each location. Since the marginal labor production cost associated with these activities is frequently quite small, this allocation by payroll may not be logical.

Data processing and hosting services are frequently offered in customizable bundles, where customers can choose from a large selection of distinct software, storage, and data processing offerings. Based on the guidance provided at the 2016 meeting of the Voorburg Group, the U.S. will consider adding survey questions to better understand the products that are being bundled and how frequently they are transacted together⁴.

⁴ <https://www.dzs.hr/voorburg/docs/Paper%20Bundling.pdf>

2.3. Description of methods of measurement

The Service Annual Survey (SAS) uses a stratified simple random sample design. The sample design for the Quarterly Services Survey (QSS) is a stratified, systematic probability proportional to size design. Sampling units expected to have a large effect on the precision of the estimates are selected with certainty. All sampling units not selected with certainty are assigned a weight representative of non-selected units.

New samples are drawn for the SAS and QSS every five years. During the period for which the samples are used, updates are made on a quarterly basis to reflect changes in the business universe. These updates are designed to account for new businesses (births) and businesses that discontinue operations (deaths). The samples are also updated to reflect mergers, acquisitions, divestitures, splits, and other changes to the business universe.

The Economic Census occurs once every five years, in years ending in 2 and 7; the SAS and QSS occur annually and quarterly, respectively. Copies of Data Processing, Hosting, and Related Services dealing forms from the Economic Census, SAS, and QSS are available in the appendix. Data for the Economic Census and SAS are collected exclusively through the Internet, with respondents logging onto a secure website to provide their data.

Estimates are computed for totals, ratios, percent contribution estimates, and trends. Estimates for both SAS and QSS are computed by specific tabulation levels. Estimates for SAS are computed by industry and employer status. Non-employer data is collected by a separate survey, but are published with SAS data releases. Estimates for QSS are computed for each industry, but not for employer status. The unadjusted estimates for each tabulation level are calculated as the sum of the weighted data of all units contained in the tabulation level. Adjusted estimates are those which have been adjusted to the most recent Economic Census. Because QSS is a sample of SAS, and SAS is a sample of the Economic Census, they all have the same sampling frame. The Census Bureau revises previously published quarterly estimates to reflect historical corrections to data for the current SAS and QSS.

The Census Bureau employs several imputation methods which can be divided between two categories: values derived by logical edits and values derived from statistical modeling. In a logical edit, the replacement value is derived from other reported values. For example, some respondents provide only basic data items such as turnover, but do not report product details. Missing product data may be imputed based on other reported items. Replacement values derived from statistical models apply historic trends or industry averages.

3. Measurement of SPPI

1.1. General framework

The Bureau of Economic Analysis (BEA) publishes GDP-by-industry and input-output data at the NAICS three-digit level, 514, encompassing Data processing, Internet Publishing, and Other Information Services. BEA publishes input-output data at the NAICS four-digit level – 5182, encompassing Data Processing, Hosting, and Related Services.

1.2. Measurement issues

The following table shows the U.S. PPI structure for NAICS 518210, Data processing, hosting, and related services.

TABLE 4
U.S. SPPI product structure

Index Codes	Index Title
518210	Data processing, hosting, and related services
518210P	Primary services
5182101	Business process management
5182104	Data management and storage, information transformation and other services
5182105	Web hosting, ASP and other IT infrastructure provisioning services
518210SM	Other receipts

The U.S. PPI for data processing, hosting, and related services closely follows the NAPCS structure, however the NAPCS structure includes additional detail.

The service line “Business process management” (BPM) includes activities performed for businesses to optimize and adapt their processes. BPM is a bundled service package that combines information technology intensive services with labor, machinery, and facilities to support, host and manage a business process for a client. In contrast with the other primary service lines for this PPI, BPM services are not limited only to the provision of technology. Examples of services that can be outsourced to firms in this industry are procurement, billing, accounting, and human resources. Many of these services may also be referred to as transaction processing.

The service line “Data management and storage, information transformation and other services” includes data management services, commonly referred to as content management. This service encompasses ongoing management and administration of data as an organizational resource. These services include:

- Data modeling: The analysis of data objects and their relationships to other data objects. Data modeling is often the first step in database design and object-oriented programming as the designers first create a conceptual model of how data items relate to each other.

Data modeling involves a progression from conceptual model to logical model to physical schema.

- **Data mining:** A class of database applications that look for hidden patterns in a group of data that can be used to predict future behavior. For example, data mining software can help retail companies find customers with common interests. Data mining is popular in the science and mathematical fields but also is utilized increasingly by marketers trying to distill useful consumer data from Web sites.
- **Data mobilization:** Brings data together from a variety of sources, formats, and platforms and integrates the data into a database environment of one's choice.
- **Data mapping/rationalization:** Rationalization is the process of identifying the commonalities among data in several different physical locations, and either creating an index to map the relationships, or to change the data or structures.
- **Data transfer:** Refers to the act of digital information moving from one computer or network to another computer, device or network within a given amount of time.

Information transformation services, also referred to as integration services or data conversion, are defined as providing technical expertise and equipment to transform information and documents from one format or media to another. This includes integrating two different applications so that data can be transferred from one application to another. Related services include:

- **Streaming:** A technique for transferring data such that it can be processed as a steady and continuous stream. Streaming technologies are becoming increasingly important with the growth of the Internet because most users do not have fast enough access to download large multimedia files quickly. With streaming, the client browser or plug-in can start displaying the data before the entire file has been transmitted.
- **Scanning:** Using an input device to convert an image or text into digital form for storage or display.

The service line "Web Hosting, ASP and other IT infrastructure provisioning services" includes:

Web hosting: A web host is an interconnected group of web servers containing web site data that others can access through the Internet. This service is generally divided into three categories: shared hosting, dedicated/managed hosting, and colocation hosting.

- **Shared hosting** is the most basic model where numerous web sites are shared on one "server". The server is owned by the hosting company and used for multiple clients.
- **Dedicated hosting** is where a web site is allotted its own server. The server is still owned by the hosting company, but dedicated solely to that client. This is more flexible than shared hosting as webmasters generally tend to have full control over the back end of the server, including choice of operating system.

- Colocation hosting: The customer provides the hardware to the hosting company, but the server is housed on the hosting company's premises. The customer is essentially renting space in the hosting company's network operations center receiving bandwidth facilities to get the web server up and running.

Application service providers (ASPs) as well as independent software vendors provide leased software applications from a centralized, hosted, and managed computing environment owned and operated by the ASP. The application resides on the ASP's server(s) and is accessed by users via a web browser or special purpose software provided by the ASP.

Other hosting and infrastructure provisioning services include:

- Computing infrastructure services: Infrastructure services include the provision of physical or, more likely, virtual computers. Clients can install their own software and operating system and are billed based on computing time. (IaaS)
- Data storage: Data storage services are defined as managing or administering the storage and back-up management of data such as remote back-up services, storage, or hierarchical storage management (migration).
- Platform services: Platforms are provided to developers for building applications on top of a computing infrastructure. The platform typically includes an operating system, programming language execution environment, database, and web server. (PaaS)
- Other services: These services can include hosting a client's custom application (not the same as hosting a client's website) and computer time share.

A significant challenge in this industry is correctly classifying firms and collecting products and services that are in-scope. It is possible that multiple services could be provided in a contract and that these services could cross industries. For example, companies may also provide network design or IT consulting services in addition to hosting application or providing network infrastructure. In those instances, only services falling under data processing, hosting, and related services are considered in-scope and collected. Additionally, a sample unit may provide accounting, procurement, payroll, and/or financial transaction processing services. For the sample unit to be correctly classified in the PPI for data processing and hosting, it cannot be an accounting, procurement, or payroll department of a company that is classified in another industry (e.g. accounting department of a manufacturing firm); and it must primarily provide other data processing and hosting services in conjunction.

Furthermore, companies in this industry are likely to be involved in areas of computer services other than data processing and hosting, like professional services or information services. These companies are continuously adding new and complementary services and expanding into other aspects of information services such as electronic mail, electronic database services, systems integration, and software development.

Sampling design

A firm's probability of selection is based on its employment size. Due to the wide variety of services offered by data processing, hosting, and related services firms, multiple frame sources are used to create the frame universe. Sample units are initially stratified based on their primary service which is determined by analyst research. There are three strata designations:

- BP – business process management
- H – hosting
- DM – data management and all other related services

After a firm is selected and agrees to participate in the survey, a probability sampling technique called disaggregation is used to determine which specific services will be included in the PPI. Services offered by establishments in this industry are provided on a contractual basis. The length of the contract can vary from monthly agreements up to several year agreements. Prices are typically quoted on the basis of time (hourly, monthly) or bandwidth, storage capacity, and memory size (gigabytes, terabytes).

Once a survey unit in this industry is initiated, post stratification is performed whereby each establishment's strata designation is determined by the plurality of turnover information collected at initiation. Typically, an establishment will remain in the same strata in which it was sampled.

Data sources for weights

Sampled transactions are weighted by a measure of their size and importance. In the first stage of PPI computation, price indexes are constructed for narrowly-defined groupings of goods or services. The individual transactions included in these indexes are weighted by the producing establishment's turnover for the product line. In the second stage of PPI computation, indexes for individual goods and services are combined into aggregate indexes. Data for weighting together the product line indexes comes primarily from the Economic Censuses of the U.S. Census Bureau. These weights are updated every 5 years.

1.1. Description of pricing methods and criteria for choosing the method

Price determining characteristics

In general, all primary services provided by establishments in this industry can be classified as the outsourcing of a business function. The cost of outsourcing such processes to each customer varies depending upon one or more of the following price determining characteristics factors:

- Number of users
- Computer storage capacity (gigabytes)
- Computer connection speed
- Processor speed of the server
- Amount of memory used
- Level of service/uptime

- Number of transactions to be processed
- Sophistication of technology needed
- Disaster recovery plan/server back-up
- Amount of data being transferred
- Storage space (physical space to store clients hardware)
- Network/server redundancy
- Multi-tenant or single tenant architecture (shared or dedicated server)
- Price floors/ceilings written into contract
- Implementation fees (includes training)
- Penalty fees
- Customization/optional features

Pricing methods

Several pricing schemes may be utilized by companies in this industry. The following are the most common types and examples of prices for data processing and hosting.

Fees based on time, usage

Fees that are based on time or usage allow clients to pay for services by the hour, transaction, RAM units (gigabytes, terabyte, etc), or similar unit of measure. Clients can pay as they go with no long-term commitments, freeing them from the costs and complexities of planning, purchasing and maintaining hardware. This transforms what were commonly large, fixed costs into much smaller variable costs. Companies may refer to this type of pricing as “pay-as-you-go” or “on-demand” pricing. Alternatively, clients can prepay for services at a discounted rate.

The following table provides an example of fee-based pricing.

Standard instance*	Application usage
Small	\$0.08 per hour
Medium	\$0.16 per hour
Large	\$0.29 per hour
Extra Large	\$0.55 per hour

*Note that an instance is equal to one physical or virtual server or “license”.

Companies may present their fees as a tiered structure. With tiered pricing, services are offered at varied price levels based on the amount of that service which is purchased or used. Generally lower prices are offered for higher rates of use as an incentive for the customer to purchase more.

The following table shows an example of tiered pricing:

Data Transfer from Cloud to Internet	Price
First 1 TB/month	\$0.00 per TB
Next TB/month	\$0.21 per TB
Next 40 TB/month	\$0.16 per TB
Next 100 TB/month	\$0.07 per TB

Next 350 TB/month	\$0.03 per TB
-------------------	---------------

Calculating the fee-based price for a selected transaction typically includes the following steps:

1. Selecting a recent representative bill or record of service features for a client. If an actual bill is not available, the respondent is asked to develop one based on a typical service contract.
2. Record the service component(s) that are most representative of the package being priced.

Example

Service	Units	Unit of measure	Rate	Total charge
Small instance				
15 TB/hr memory usage				
Data transfer out				
Data transfer in				
IP addresses				
			Total	

3. The number of units (hours, gigabytes, etc.) and unit of measure for each service line are recorded.

Example

Service	Units	Unit of measure	Rate	Total charge
Small instance	900	hr		
15 TB/hr memory usage	56	hr		
Data transfer out	50	TB		
Data transfer in	150	TB		
IP addresses	19	address		
			Total	

4. The monthly average rate per hour, terabyte, etc. for each service component is collected. Monthly average rates are **preferred** so that discounting and “free” services can be captured. Monthly average rates are calculated by dividing total revenue for each service line by the total units or size for that service.

	Service	Total Size/Units	Total Revenue	Monthly Average Rate
	Standard instance usage			
(1)	Small	1075 hours	\$18.28	\$0.017/hr
	Medium	1034 hours	\$33.09	\$0.032/hr

	Large	1021 hours	\$43.90	\$0.043/hr
	Extra large	108 hours	\$6.05	\$0.056/hr
	Memory usage per hour			
	5 TB	378 hours	\$97.90	\$0.259/hr
	10 TB	221 hours	\$84.86	\$0.384/hr
(2)	15 TB	546 hours	\$258.26	\$0.473/hr
	20 TB	299 hours	\$197.94	\$0.662/hr
	Data transfer out			
(3)	< 100 TB	71 TB	\$15.7642	\$0.222/TB
	101-500 TB	395 TB	\$180.12	\$0.456/TB
	> 501 TB	620 TB	\$604.50	\$0.975/TB
	Data transfer in			
	< 100 TB	97 TB	\$0.00	\$0.000/TB
(4)	101-500 TB	333 TB	\$26.31	\$0.079/TB
	> 501 TB	561 TB	\$55.54	\$0.099/TB
	IP addresses			
	<10	2 addresses	\$2.69	\$1.345 /address
(5)	11-25	19 addresses	\$24.64	\$1.297/address
	26-50	37 addresses	\$44.81	\$1.211 /address
	50-75	54 addresses	\$54.11	\$1.002 /address
	>75	91 addresses	\$89.91	\$0.988 /address

The following table shows the selected transaction to be priced.

	Service	Units	Unit of measure	Rate	Total charge
(1)	Small instance	900	hr	\$0.017	\$15.30
(2)	15 TB/hr memory usage	56	hr	\$0.473	\$26.49
(3)	Data transfer out	50	TB	\$0.222	\$11.10
(4)	Data transfer in	150	TB	\$0.079	\$11.85
(5)	IP addresses	19	address	\$1.297	\$24.64
	Total				\$89.38

EXAMPLE B: In this example, the company can provide averages only at a broader level than was available in Example A. Here the firm provides average per hour rates for all instances provided, rather than an average only for small instances. The following table shows the revenue and units consumed for the services offered by Company B for which they can provide average rate data.

Service	Total Size/Units	Total Revenue	Average monthly rate
Standard instance usage	3238 hours	\$101.31	\$0.0313/hr
Memory usage per hour	1444 hours	\$638.96	\$0.4425/hr
Data transfer out	1086 TB	\$800.38	\$0.737/TB
Data transfer in	991 TB	\$84.85	\$0.0826/TB
IP addresses	203	\$18.33	\$1.0648/address

The following table shows the selected transaction to be priced. The average rates are entered from the corresponding rows in the table above:

Service	Units	Unit of measure	Rate	Total charge
Small instance	900	hr	\$0.0313	\$28.16
15 TB/hr memory usage	56	hr	\$0.4425	\$24.78
Data transfer out	50	TB	\$0.7370	\$36.85
Data transfer in	150	TB	\$0.0823	\$12.39
IP addresses	19	address	\$1.0648	\$20.23
			Total	\$122.41

- When monthly average rates are not available, the standard rate that would be charged in the current time period is collected.

EXAMPLE C: In this example, the company cannot provide average rate information. The following table shows all services offered by Company C and their current standard rate pricing structure. The highlighted rows are the applicable charges for the selected bill.

Service	Standard Rate
Standard instance usage	
Small	\$0.015/hr
Medium	\$0.030/hr
Large	\$0.045/hr
Extra large	\$0.060/hr
Memory usage per hour	
5 TB	\$0.225/hr
10 TB	\$0.380/hr
15 TB	\$0.470/hr
20 TB	\$0.665/hr
Data transfer out	
< 100 TB	\$0.200/TB
101-500 TB	\$0.400/TB
> 501 TB	\$0.900/TB
Data transfer in	
< 100 TB	\$0.001/TB
101-500 TB	\$0.075/TB
> 501 TB	\$0.110/TB
IP addresses	
<10	\$1.395 /address
11-25	\$1.250/address
26-50	\$1.200 /address
50-75	\$1.005 /address
>75	\$0.975 /address

The following table shows the selected transaction to be priced. The standard rates from the corresponding rows (as highlighted) are entered in the table below:

Service	Units	Unit of measure	Rate	Total charge
Small instance	900	hr	\$0.015	\$13.50
15 TB/hr memory usage	56	hr	\$0.470	\$26.32
Data transfer out	50	TB	\$0.200	\$10.00
Data transfer in	150	TB	\$0.075	\$11.25
IP addresses	19	address	\$1.250	\$23.75
Total				\$84.82

Flat fee

For some clients, it may be more cost effective to negotiate a contract to pay a single flat fee on a monthly or annual basis due to high capacity requirements and/or a large number of users. In these cases during repricing, companies will be asked to estimate the flat fee for the given service based on their current pricing structure.

Index estimation procedure, including estimation of missing prices

PPIs are calculated using the formula for a modified Laspeyres index. The Laspeyres index compares the base period turnover for a set of products or services to the current period turnover for the same set of products or services.

If no price report from a participating firm has been received in a particular month, the change in the price of the associated transaction is estimated by averaging the price changes for other transactions within the same detailed index line (i.e., for the same kind of services) for which price reports have been received.

Quality adjustment

The U.S. is currently studying the feasibility of using a hedonic quality adjustment for cloud computing, which would impact items in the SPPI for data processing, hosting, and related services. In developing the hedonic model, we have collected price data from three large U.S. cloud service providers and are evaluating six common variables: vCPUs, RAM, storage, SSD, cloud provider, and operating system (OS) (Linux vs Windows). Our objective is to create a statistical model that uses prices and qualities of hundreds of items from large producers to more accurately predict how even the slightest of quality changes affect the price of any given item in the index.

Frequency of collection

Data processing, hosting, and related services firms report prices for the selected transactions, usually on a monthly basis, using a form provided by the U.S. PPI. Firms are asked to report their prices as of the Tuesday of the week containing the 13th of the month. If the firm fails to report or reports incomplete information, it is called by an economist who requests the needed information.

Firms report prices through a web-based application. Firms continue to report until a new sample is selected for the industry— after 7 to 8 years, on average.

3.4. Evaluation of comparability of price data with output data

In general, output and price measures for data processing and hosting services are comparable. Due to the rapid growth of service offerings as well as changes in the service delivery process, the NAICS classification of some service lines need to be regularly updated. As such, the PPI works closely with the Census Bureau to ensure consistency among PPI prices and the Census output data.