ABSTRACT

Systems and methods are described for providing a gateway that enables cloud-based service exposure. The gateway can allow a particular operator to expose its services and to control, manage and monetize the communication traffic that accesses these services. In accordance with one use case, the gateway can be utilized to expose the services of a Web based application to other external service providers and applications and to manage, control and monetize the requests received from the external providers to the exposed service. In accordance with another use case, the gateway can be utilized to expose the excess capacity of a telecom network, such as a code division multiple access (CDMA) network or a global system for mobile communications (GSM) network and to manage the access to the exposed capacity of the network.
FIGURE 1
FIGURE 3
Gateway Communication Service 400

Application-Facing Interface 406

Authentication 408

XML Serialization 410

Begin Transaction 412

Validation, Policy Enforcement, Routing 414

Protocol Translation and State 416

FIGURE 4
Incoming Service Request Gateway 500

Services Gateway Interceptor Stack Database 508

Databases, LDAP,
Protocol Translation 504

Underlying Service (e.g. network node, application) Rule Repository 512

FIGURE 5
GATEWAY FOR ENABLING CLOUD-BASED SERVICE EXPOSURE

CLAIM OF PRIORITY

[0001] The present application claims the benefit of U.S. Provisional Patent Application No. 61/294,766, entitled “GATEKEEPER SERVICE EXPOSURE PLATFORM FOR MOBILE COMMUNICATIONS,” by Shanath Rajasekar et al., filed on Jan. 13, 2010, which is incorporated by reference herein in its entirety, including all Appendices filed therewith.

COPYRIGHT NOTICE

[0002] A portion of the disclosure of this patent document contains material which is subject to copyright protection. The copyright owner has no objection to the facsimile reproduction by anyone of the patent document or the patent disclosure, as it appears in the Patent and Trademark Office patent file or records, but otherwise reserves all copyright rights whatsoever.

FIELD OF THE INVENTION

[0003] The current invention relates to mobile communications and gateways for managing access to services in the cloud.

BACKGROUND

[0004] With the rising pervasiveness of wireless communications and mobile devices, more and more objects are becoming integrated with the World Wide Web. Everything from cellular phones to laptops, electronic books, tablets, personal digital assistants (PDAs) and even automobiles are striving to maintain a constant connection to the internet. As a result of all this interconnectivity, many websites and online content providers have been showing an increasing interest in providing services to mobile clients.

[0005] As the worlds of TCP/IP applications and telephony networks continue to converge, subscribers want services that provide them with functionality and flexibility that cross the traditional boundaries between the world of the Internet and the world of their phones. Operators of the telecom networks and of various online content applications want to be responsive to those desires and to provide services that will satisfy subscriber demands, promote subscriber loyalty, increase revenues and drive traffic to their networks.

[0006] However, for both mobile network operators (MNOs) that own the telecom network and for online content providers, integrating these services has historically been complex and uneasily. What is needed is an improved way to manage access to the services provided in the internet and telecom network clouds and to expose those services to external entities.

BRIEF SUMMARY OF THE INVENTION

[0007] In accordance with various embodiments of the invention, systems and methods are described for providing a gateway that enables cloud-based service exposure. The gateway can allow a particular operator to expose its services and to control, manage and monetize the communication traffic that accesses these services. In accordance with one use case, the gateway can be utilized to expose the services of a Web based application to other external service providers and applications and to manage, control and monetize the requests received from the external providers to the exposed service. In accordance with another use case, the gateway can be utilized to expose the excess capacity of a telecom network, such as a code division multiple access (CDMA) network or a global system for mobile communications (GSM) network and to manage the access to the exposed capacity of the network.

BRIEF DESCRIPTION OF THE DRAWINGS

[0008] FIG. 1 is an illustration of one use case of the gateway platform, in accordance with various embodiments of the invention.

[0009] FIG. 2 is an illustration of another use case of the gateway platform, in accordance with various embodiments of the invention.

[0010] FIG. 3 is an illustration of a possible deployment of the gateway platform, in accordance with various embodiments of the invention.

[0011] FIG. 4 is an illustration of a communication service in the gateway, in accordance with various embodiments of the invention.

[0012] FIG. 5 is an illustration of the communication service utilizing an interceptor stack to apply service level agreements and policy enforcement to incoming requests, in accordance with various embodiments of the invention.

DETAILED DESCRIPTION

[0013] The invention is illustrated by way of example and not by way of limitation in the figures of the accompanying drawings in which like references indicate similar elements. References to embodiments in this disclosure are not necessarily to the same embodiment, and such references mean at least one. While specific implementations are discussed, it is understood that this is done for illustrative purposes only. A person skilled in the relevant art will recognize that other components and configurations may be used without departing from the scope and spirit of the invention.

[0014] In the following description, numerous specific details are set forth to provide a thorough description of the invention. However, it will be apparent to those skilled in the art that the invention may be practiced without these specific details. In other instances, well-known features have not been described in detail so as not to obscure the invention.

[0015] In accordance with various embodiments of the invention, a gateway is described for enabling cloud based service exposure. The gateway can allow a particular service provider (hereinafter owner or operator) to expose its services and to control, manage and monetize the communication traffic that access these services. In accordance with an embodiment, the gateway can supply a policy enforcement engine, security enforcement, billing integration and traffic throttling/shaping functionality on the request traffic flowing through it. The policy enforcement engine uses service level agreements (SLAs) to regulate external access by clients and other service providers to the particular service being offered by the operator and exposed by the gateway. The security enforcement can provide authentication and authorization for the request traffic and the billing integration enables the operator to monetize and charge for the traffic flowing through the gateway. Finally, the traffic shaping and throttling features enable the gateway to manage and control the amount of request traffic being sent to the service being provided by the operator.
In accordance with various embodiments, two primary use cases of the gateway are described in the present disclosure—a use case for exposing the services of a Web based application to other service providers and a use case for exposing the excess capacity of a telecom network, such as a code division multiple access (CDMA) network or a global system for mobile communications (GSM) network.

Web Application Use Case

In accordance with an embodiment, the gateway can be deployed by an owner of a web application that wishes to expose its services to other websites/service providers and to be able to monetize and control access to its services by such external service providers. This can enable the owner of a service to utilize the policy enforcement, traffic shaping/throttling, billing integration and other features of the gateway and to apply these features on the request traffic being sent to the web application.

FIG. 1 is an illustration of one use case of the gateway platform, in accordance with various embodiments of the invention. Although this diagram depicts components as logically separate, such depiction is merely for illustrative purposes. It will be apparent to those skilled in the art that the components portrayed in this figure can be combined or divided into separate software, firmware and/or hardware. Furthermore, it will also be apparent to those skilled in the art that such components, regardless of how they are combined or divided, can execute on the same computing device or can be distributed among different computing devices connected by one or more networks or other suitable communication means.

As illustrated, a service provider (operator) 126 of a particular web site or web application can publish its service 102 on the Internet 100. This service can be available to a multitude of users/clients 120 that can access it by way of a browser on a computing device, as well as to a number of other service providers and owners 106, 108 that may wish to access it, such as by using Web Services.

In accordance with an embodiment, the owner 126 of the web service 102 can deploy the gateway 104 to manage access to its service by any number of other providers 106, 108 that wish to make use of the operator’s services. By way of an example, the operator of web service A 102 may be an owner of a small business that can make boat reservations on a particular lake. A different provider 106 may be a travel website that wishes to incorporate the ability to make boat reservations for users 122 that are booking their travel plans using the travel web service 106. Similarly, another provider 108 may sell waterskiing equipment and may also wish to provide the ability for its users 124 to make a boat reservation over its own website by using Web Services. In accordance with an embodiment, the gateway 104 is positioned to manage access to the operator’s web service 102 by the various external providers and to monetize and control that access.

In accordance with an embodiment, the gateway performs a variety of functions to the incoming requests, including but not limited to service and API exposure 118, security 116, billing integration 112, policy enforcement 110 and traffic shaping/throttling 114. The service exposure 118 portion can expose an interface to access the service 102 provided by the operator and to invoke its functionality. This can be performed by publishing a generic Web Services Definition Language (WSDL) file that describes how to use web services to access the service provided by the operator. In alternative embodiments, the gateway can expose the service as several different façades (views) that can be accessible by the external providers, each façade being accessible over a different protocol or API.

The policy enforcement engine 110 of the gateway 104 uses service level agreements (SLAs) to regulate service provider and application access to the particular service functionality provided by the owner 126. In accordance with an embodiment, the policy engine supports a range of quality-of-service guarantees that can be modulated by time of day/day of week, rates and quotas. If desired, further rules covering access can also be added. The service provider and application accounts can be divided into groups to simplify SLA management and maintenance. The gateway can come with a set of default SLAs and also provide tools (e.g., wizard) to create custom SLAs tailored for the particular service of the owner. In accordance with an embodiment, if the owner 126 has a list of subscribers, the subscriber permissions and preferences can be reflected in separate subscriber SLA created by the operator 126. For example, one subscriber may indicate that they wish to allow service provider 106 to obtain information regarding the subscriber, while other subscriber 108 to not be able to access their information. The subscriber SLA can contain the preferences of the operator's subscriber and enforce the requests coming from the service providers accordingly.

The security framework 116 can enable the gateway 104 to function as a single point of contact for access to the functionality provided by the owner 126. It can provide common authentication, authorization and access control procedures for all applications that are external to the owner 126, as well as internal applications. In accordance with an embodiment, the applications can be authenticated using plaintext or digest passwords, X.509 certificates or SAML 1.0/1.1 tokens. Service requests can use XML encryption based on the W3C standard. To ensure message integrity, the requests can be digitally signed, using W3C XML digital signature standard.

The billing integration mechanism 112 can enable the owner 126 to incorporate charging and payment functionality into the gateway according to the request traffic being sent to its service 102. In accordance with an embodiment, the gateway can include a data record listener which listens for various data records generated within the gateway and issues Diameter offline charging requests to an external billing server upon detecting that a particular data record has been generated. In addition, the gateway can include a credit control interactor that intercepts requests in the communication traffic flowing through the gateway and initiates Diameter online charging requests to external billing servers upon intercepting said requests. The ability to integrate the gateway with the Diameter online and offline charging requests can enable the owner 126 to charge for its services more dynamically, according to the traffic being handled by the service 102.

The traffic shaping and throttling 114 can protect the service 102 and the various underlying cluster nodes that are providing the service from traffic overload during periods of high traffic volume. When a request is received by the gateway, a budget service can be invoked to determine a current available budget for the request. The budget can be based on a service level agreement (SLA) for the service provider or application that sent the request. Additionally, the SLA can include preference information from the subscribers of the owner 126. In accordance with an embodiment, the requests can be assigned a priority, such as high or low priority. For
example, the operator may assign requests coming from its subscribers a high priority and requests coming from the external service providers a lower priority. If the budget is under a particular threshold value, the low priority requests can be denied, while the high priority requests can be processed as long as there is some available budget left. If the budget for the request has reached the restricted level, the request can be denied and optionally en-queued to a traffic shaping queue to be processed at a later time.

[0027] In this manner, the gateway can be used by the operator 126 to expose its services and to manage, control and monetize access to these services from external entities.

[0028] Excess Network Capacity Use Case

[0029] In another embodiment, the gateway can be utilized to expose the excess capacity of a telecommunications (telecom) network, such as a CDMA or GSM network. The telecom network usually includes a set of resources (e.g. SMSC, MMSC, etc.) that handle communications between mobile subscriber devices (e.g. cellular phones, etc.). These resources of the telecom network are typically provisioned to handle very large amounts of traffic, however, for the majority of the time, they are usually operating at a much lower usage than their maximum capacity. By way of example, the telecom network may be provisioned to handle communications during the busiest periods, such as New Year’s evening, times of natural disasters or other emergency situations and the like. As such, at any given time that does not involve such periods of high traffic, the telecom network may be operating at 20-30% capacity. Furthermore, there may be certain periods of time when the network can be reasonably expected to operate at much lower usage rates, such as during the hours of midnight until 6 am, for example. In light of this, it can be desirable for the network operator to utilize the excess capacity of the network and to monetize that use for various purposes.

[0030] In accordance with an embodiment, the gateway can be used to expose the telecom network resources of the network operator to other entities, such other operators, service providers and web applications. For example, the gateway can expose the native protocols of the telecom network resources, such as native MM7 and other signaling protocols, as a set of generic web services APIs that are accessible to these external entities. In addition, a set of policies and configurations can be established on the gateway as SLAs to regulate the exposure of the excess capacity in order to ensure that the network is not overloaded.

[0031] FIG. 2 is an illustration of another use case of the gateway platform, in accordance with various embodiments of the invention. Although this diagram depicts components as logically separate, such depiction is merely for illustrative purposes. It will be apparent to those skilled in the art that the components portrayed in this figure can be combined or divided into separate software, firmware and/or hardware. Furthermore, it will also be apparent to those skilled in the art that such components, regardless of how they are combined or divided, can execute on the same computing device or can be distributed among different computing devices connected by one or more networks or other suitable communication means.

[0032] As illustrated, the telecom network operator 230 can deploy the gateway 208 to manage access to the internal resources of the telecom network 200, such as the short messaging service centers (SMSC 204), multimedia messaging service centers (MMSC 202) and other resources 206. The gateway can enable the operator to expose the excess capacity of the network resources to other mobile network operators (MNOs), mobile virtual network operators (MVNOs) 222, service providers and web applications 224 and other entities 226 that wish to access the mobile network.

[0033] In accordance with an embodiment, the gateway 208 can provide the service orchestration and exposure 210 of the internal network capabilities. The protocols required by underlying network capabilities are often complex and the learning curve associated with achieving competence in them can be steep. The gateway can expose access to standard network capabilities, such as SMS, MMS and Call Control as a set of interfaces tailored to the particular entity accessing the network. For example, depending on the needs of the external provider, the gateway can offer a simple object access protocol (SOAP) style façade (interface), a Parlay X 2.1 and 3.0 façade, a representational state transfer (RESTful) style façade and native protocol façade interfaces. The SOAP Web Service interface can be published as standard WSDL files and RESTful Web Services can be designed to use in pure HTTP environments. In accordance with an embodiment, the gateway can also enable the ability to orchestrate multiple services in a single call by invoking the façade interfaces provided thereon.

[0034] In accordance with an embodiment, the gateway can be structured as a container optimized for the task of executing a set of communication services. Communication services are specialized components that can allow the operator to provide application developers and other providers a simple way to access the network services, including messaging, call control, terminal location and presence. A communication service can be comprised of a client-facing interface (service façade) and a service enable that provides the protocol translation and routing the request to the network resource. Communication services will be described in further detail below, with reference to FIG. 4.

[0035] In accordance with an embodiment, the gateway enables network integration 218 by employing a plurality of plugins to communicate with the network resources. A plugin is a connection established between the gateway and a particular network resource. The plugin communicates in the native protocol of the network resource, however, the gateway can expose this native protocol as a generic interface or some other façade to the client accessing its capabilities. By utilizing a set of such plugins, the gateway can provide an internal system for routing of service requests directly to appropriate network nodes based on a variety of parameters, including sending application, destination address, or any arbitrary request parameter. There can be in-production deployment of multiple instances of network plugins on an as needed basis; as a result routing can be managed in a very fine-grained way.

[0036] In addition, as previously described in connection with FIG. 1, the gateway can include the policy enforcement 220, billing integration 214 and traffic throttling 216 features that can be applied to the communications flowing to and from the telecom network 200.

[0037] In accordance with an embodiment, all of these features of the gateway can be used to expose the excess capacity of the telecom network to the third party service providers 222, 224, 226. For example, the policy enforcement engine can be used in conjunction with the traffic throttling capabilities to set up a service level agreement that allows a particular third party service provider to utilize the network resource for a reduced charge while the network resource is running at less
than a pre-specified capacity. For example, the network operator 230 may set up one tier of pricing for utilizing its network when it is running at 20% capacity, another pricing tier while it is running at 50% capacity and may preclude access altogether when the network is running at 90% capacity. Various other use cases of this are also possible.

[0038] FIG. 3 is an illustration of a possible deployment of the gateway platform, in accordance with various embodiments of the invention. Although this diagram depicts components as logically separate, such depiction is merely for illustrative purposes. It will be apparent to those skilled in the art that the components portrayed in this figure can be combined or divided into separate software, firmware and/or hardware. Furthermore, it will also be apparent to those skilled in the art that such components, regardless of how they are combined or divided, can execute on the same computing device or can be distributed among different computing devices connected by one or more networks or other suitable communication means.

[0039] As illustrated, for high availability and security, the gateway deployment 300 can be split into two tiers: an access tier 302 and a network tier 304. In accordance with an embodiment, the client requests from the external service providers and applications can be received at the access tier, which provides the interfaces (facades) for accessing the operator's (owner's) services. The network tier, on the other hand, can connect to the internal resources of the operator, such as to the internal network resources (in the use case of exposing excess network capacity) or the web application, and can handle the translations and communication with these internal resources.

[0040] In accordance with an embodiment, each tier is comprised of at least one cluster, with at least two server instances per cluster, and all server instances run in active mode, independently from each other. The servers in all clusters can be managed servers 306, 308, 310, 312. Together, the clusters make up a single domain, which can be controlled through an administration server 314. Communication between the access tier and the network tier can be performed using Java remote method invocation (RMI). The requests received to the gateway can be load balanced between the access tier and the network tier, and failover mechanisms can be provided between the two.

[0041] In accordance with an embodiment, there is an additional tier 316 containing the database. Within the cluster, the data can be made highly available using a cluster-aware storage service which ensures that all state data is made available across all network tier instances.

[0042] FIG. 4 is an illustration of a communication service in the gateway, in accordance with various embodiments of the invention. Although this diagram depicts components as logically separate, such depiction is merely for illustrative purposes. It will be apparent to those skilled in the art that the components portrayed in this figure can be combined or divided into separate software, firmware and/or hardware. Furthermore, it will also be apparent to those skilled in the art that such components, regardless of how they are combined or divided, can execute on the same computing device or can be distributed among different computing devices connected by one or more networks or other suitable communication means.

[0043] As previously mentioned, the communication service can be a specialized component that allows an operator to provide external service providers and application developers access to its services. In accordance with an embodiment, the communication service is a component that runs within the gateway and provides the functionality of the gateway, exposing capabilities of the operator's service. In accordance with an embodiment, all traffic in the gateway is processed through these services.

[0044] In accordance with an embodiment, the communication service 400 is comprised of a service façade 402 and a service enabler 404. The service façade includes an application-facing interface 406 which communicates with the application, a security layer which handles authentication 408 and the XML serialization layer 410 for communicating with the service enabler. The service enabler includes a processing layer 412, where requests are validated, evaluated according to service level agreements 414 and routed. The service enabler further includes a protocol translation layer (plugin) 416 which communicates with the underlying resource (e.g. network element).

[0045] In accordance with an embodiment, each request coming into the gateway can be processed using a communication service 400. The communication service can utilize a stack or sequence of interceptors for processing each request. The interceptor stack is described in further detail below, with reference to FIG. 5.

[0046] FIG. 5 is an illustration of the communication service utilizing an interceptor stack to apply service level agreements and policy enforcement to incoming requests, in accordance with various embodiments of the invention. Although this diagram depicts components as logically separate, such depiction is merely for illustrative purposes. It will be apparent to those skilled in the art that the components portrayed in this figure can be combined or divided into separate software, firmware and/or hardware. Furthermore, it will also be apparent to those skilled in the art that such components, regardless of how they are combined or divided, can execute on the same computing device or can be distributed among different computing devices connected by one or more networks or other suitable communication means.

[0047] In accordance with an embodiment, the gateway 500 can use service level agreements (SLAs) to govern the access to the operator's resources by external applications. These SLAs can group service providers into groups and define whether a member of a service provider group is able to (1) have access to a particular communication service (which can be regulated down to supported methods and parameters), (2) participate in any quality of service (QoS) agreements. Participating in the QoS agreement can specify things such as the guaranteed number of requests a service provider may send through a particular communication service in a given period of time. These guarantees can be modulated by time of day, day of week, rate (invocations per time period) and quota (aggregated number of invocations). SLA enforcement can be provided by using an interceptor stack. Interceptors can be components that are invoked by the communication service in sequence for each request. In accordance with an embodiment, the interceptors can be registered in the gateway by using a configuration file. Once registered, the interceptors can be invoked in a particular sequence for the requests coming into the gateway. Each interceptor can deny the request, allow the request, abstain from processing the request or pass the request to the next interceptor specified by the sequence. As such, the some interceptors can be skipped in the sequence. The interceptors can also modify the various data associated with the request. Interceptors can
collaborate in request processing. In addition, new custom interceptors can be created and registered at the gateway to process incoming requests.

In accordance with an embodiment, a communication service 502 receives an incoming request and invokes the interceptor stack 506 to process the request. At least one interceptor in the interceptor chain can evaluate the request parameters and then either deny or allow the request according to the SLA associated with the service provider that sent the request. The interceptor stack can implement data from a variety of data sources 508, 510, 512 when evaluating the request. For example, a rule repository 512 can be maintained to manage a set of rules for processing the request. Once the interceptor stack is finished processing the request, it can be forwarded to the protocol translation component 504 which can translate the request into the particular protocol used by the service of the operator.

It should be noted that the structure, deployment and other various components of the gateway illustrated herein can be modified and customized as needed, within the scope of the various embodiments described herein. The particular structure is described for purposes of illustration and understanding and is not intended to limit all of the embodiments of the use cases described herein. Further details and information on the gateway and other functionality described throughout this disclosure can be found in the U.S. Provisional Patent Application No. 61/294,766, which is incorporated herein by reference in its entirety, including all of the Appendices filed therewith.

Throughout the various contexts described in this disclosure, the embodiments of the invention further encompass computer apparatus, computing systems and machine-readable media configured to carry out the foregoing systems and methods. In addition to an embodiment consisting of specifically designed integrated circuits or other electronics, the present invention may be conveniently implemented using a conventional general purpose or a specialized digital computer or microprocessor programmed according to the teachings of the present disclosure, as will be apparent to those skilled in the computer art.

Appropriate software coding can readily be prepared by skilled programmers based on the teachings of the present disclosure, as will be apparent to those skilled in the software art. The invention may also be implemented by the preparation of application specific integrated circuits or by interconnecting an appropriate network of conventional component circuits, as will be readily apparent to those skilled in the art.

The various embodiments include a computer program product which is a storage medium (media) having instructions stored thereon/in which can be used to program a general purpose or specialized computing processor(s)/device(s) to perform any of the functions presented herein. The storage medium can include, but is not limited to, one or more of the following: any type of physical media including floppy disks, optical discs, DVDs, CD-ROMs, microdrives, magneto-optical disks, holographic storage, ROMs, RAMs, PRAMS, EPROMs, EEPROMs, DRAMs, VRAMs, flash memory devices, magnetic or optical cards, nanosystems (including molecular memory ICs); paper or paper-based media; and any type of media or device suitable for storing instructions and/or information. The computer program product can be transmitted in whole or in parts and over one or more public and/or private networks wherein the transmission includes instructions which can be used by one or more processors to perform any of the features presented herein. The transmission may include a plurality of separate transmissions. In accordance with certain embodiments, however, the computer storage medium containing the instructions is non-transitory (i.e. not in the process of being transmitted) but rather is persisted on a physical device.

The foregoing description of the preferred embodiments of the present invention has been provided for purposes of illustration and description. It is not intended to be exhaustive or to limit the invention to the precise forms disclosed. Many modifications and variations can be apparent to the practitioner skilled in the art. Embodiments were chosen and described in order to best explain the principles of the invention and its practical application, thereby enabling others skilled in the relevant art to understand the invention. It is intended that the scope of the invention be defined by the following claims and their equivalents.

What is claimed is:

1. A system for exposing and managing cloud-based services, said system comprising:
   a web application that provides at least one service accessible over the internet to a plurality of users;
   a gateway deployed to manage access to said web application by one or more external service providers, wherein the gateway exposes an interface to invoke the service of the application, wherein the gateway receives incoming requests from the external service providers to invoke the interface and applies a set functions to manage said incoming requests, wherein the set of functions include one or more of the following:
   a set of authentication, authorization and access control procedures for said incoming requests;
   a policy enforcement engine that uses service level agreements to regulate the external service providers accessing the service provided by the application; and
   traffic throttling and traffic shaping of the incoming requests from the external service providers.

2. The system of claim 1, wherein the set of functions further include:
   protocol translation of said incoming requests between a protocol employed by the application and protocols used by said external service providers.

3. The system of claim 1, wherein the set of functions further include:
   a billing integration module that transmits Diameter offline or Diameter online charging requests to an external billing server based on the requests received at the gateway.

4. The system of claim 1, wherein the gateway further includes a communication service that invokes a sequence of interceptors to process each incoming request at the gateway.

5. The system of claim 1, wherein at least one service level agreement enforced by the policy engine specifies a guaranteed number of requests that an external service provider is allowed to send in a specified period of time.

6. The system of claim 5, wherein the guaranteed number of requests can be modulated by time of day, day of week, rate and quota.

7. The system of claim 1, wherein the gateway is deployed in two tiers: an access tier that provides a service interface to said external service providers and a network tier that interfaces with the service of said web application.

8. The system of claim 1, wherein the traffic throttling and shaping of said incoming requests further includes:
determining a budget for each incoming request and if the budget is under a particular threshold value, denying low priority requests.

9. The system of claim 1, wherein said external service providers are other web applications that access the service of said web application over web services application programming interfaces (APIs).

10. A system for exposing and managing cloud-based services, said system comprising:

a telecommunications (telecom) network including a set of resources that process communication traffic between mobile subscriber devices, wherein the telecom network is operating at less than a maximum capacity of the communication traffic that the telecom network is provisioned to handle; and

gateway positioned to manage access to the telecom network, wherein the gateway exposes, to one or more external service providers, an amount of excess capacity of the telecom network that is available between the used capacity and the maximum capacity, wherein the external service providers are enabled to utilize the excess capacity of the telecom network.

11. The system of claim 10, wherein the resources of the telecom network include a short messaging service center (SMSC) and a multimedia messaging service center (MMSC).

12. The system of claim 10, wherein the gateway exposes to said external service providers a generic interface to access said resources of the telecom network and wherein the gateway further performs translation between the generic interface and at least one native protocol implemented by said network resources.

13. The system of claim 10, wherein managing access by the gateway to the telecom network further includes applying, to requests transmitted between the telecom network and the external service providers, one or more of the following functions:

a set of authentication, authorization and access control procedures for said incoming requests;
a policy enforcement engine that uses service level agreements to regulate the external service providers accessing the service provided by the application; and

traffic throttling and traffic shaping of the incoming requests from the external service providers.

14. The system of claim 10, wherein the gateway further includes:

a billing integration module that transmits Diameter offline or Diameter online charging requests to an external billing server based on the requests received at the gateway.

15. The system of claim 10, wherein the gateway further includes a communication service that invokes a sequence of interceptors to process each incoming request at the gateway.

16. The system of claim 10, wherein at least one service level agreement enforced by the policy engine specifies a guaranteed number of requests that an external service provider is allowed to send in a specified period of time.

17. The system of claim 16, wherein the guaranteed number of requests can be modulated by time of day, day of week, rate and quota.

18. The system of claim 10, wherein the gateway is deployed in two tiers: an access tier that provides a service interface to said external service providers and a network tier that interfaces with the resources of said telecom network.

19. The system of claim 10, wherein the traffic throttling and shaping of said incoming requests further includes:

determining a budget for each incoming request and if the budget is under a particular threshold value, denying low priority requests.

20. A method for exposing and managing cloud-based services, said method comprising:

accessing a telecommunications (telecom) network including a set of resources that process communication traffic between mobile subscriber devices, wherein the telecom network is operating at less than a maximum capacity of the communication traffic that the telecom network is provisioned to handle;

managing access to the telecom network by deploying a gateway to said telecom network; and

exposing, by said gateway to one or more external service providers, an amount of excess capacity of the telecom network that is available between the used capacity and the maximum capacity, wherein the external service providers are enabled to utilize the excess capacity of the telecom network.