F5 White Paper

SOA Infrastructure Reference Architecture: Defining the Key Elements of a Successful SOA Infrastructure Deployment

The purpose of this paper is to document the infrastructure components and their relationships to each other—as well as to core SOA (Service Oriented Architecture) application components—that are necessary to enable the successful implementation of a robust SOA.

By Lori MacVittie
Technical Marketing Manager, Application Services

Contents

About This Document

Introduction

While the basic concept of accelerating the return on IT investments through reusable, loosely coupled services is not new, the widespread adoption of a set of architectural principles and standards through which the concept may be implemented, i.e. SOA, *is* new.

SOA is more than simply technology; it is a strategy designed to support business agility and therefore has implications throughout the organization. Given the distributed nature of a SOA it is more important than ever to consider the capabilities of the underlying infrastructure across which services will be deployed and delivered, as this foundation can have a tremendous impact on the ability of an SOA to achieve its goals and affect positive change in business operations.

Many aspects of SOA must be addressed in order for the organizational implementation to succeed. Most of the focus within the industry and across organizations thus far has encompassed the discovery, definition, design, and deployment of the composite and atomic services that comprise the building blocks on which a SOA is built. Most of these services are built upon enterprise application platforms such as those offered by Oracle, BEA, IBM, and Microsoft. Consequently, there was very little consideration for the incorporation of infrastructure-based services or the deployment of infrastructure-focused products as a means to enable the anticipated ROI of a SOA implementation.

As organizations increasingly implement more complex SOA, it becomes important to consider the ramifications of a fully implemented SOA on the existing infrastructure. The choice of infrastructure can negatively impede the success of an SOA implementation, and should therefore be given proper consideration before deployment is complete.

The purpose of this paper is to document the infrastructure components and their relationships to each other—as well as to core SOA application components—that are necessary to enable the successful implementation of a robust SOA.

Intended Audience

This document is intended for the following audiences:

- Business and IT leaders who are concerned about the delivery of services and the impact of the implementation of an SOA on existing mission-critical applications.
- Enterprise architects who need to drive the vision and roadmap of the SOA program and the architecture of each implementation under it.
- Program managers who need to manage a portfolio of sub-projects within an overall SOA business strategy.
- Standards bodies, which need a better understanding of use-cases of how business and IT plan to leverage technology to meet their objectives.

Benefits of the SOI Reference Architecture

SOI delivers a number of business benefits to the enterprise both for IT and the business.

IT Benefit	Business Benefit
Greater IT automation	Higher IT productivity Lowered operational costs
Dynamic resource allocation	Greater resource efficiency Reduce capital costs Higher reliability
Greater business accountability	Regulatory compliance Greater efficiency of IT processes

From the perspective of IT management, SOI (Service Oriented Infrastructure) offers additional benefits:

- Reduced operations workload
- Optimized resource utilization, which lowers capital expenditures
- Greater flexibility through dynamic resource allocation
- Simple, cost-effective upgrades eliminate forklift upgrades
- Support for on-demand usage models
- Removes resource management and security responsibilities from application code
- Enables reuse of existing infrastructure services

This document helps readers to:

- **Compare alternatives:** Identify and define the key technology infrastructure components of SOA to establish a baseline reference for comparison of options.
- **Improve collaboration:** A common language clarifies the nature of SOA infrastructure components defined in this document.
- **Accelerate implementations:** This guide defines the infrastructure services available along with the requirements.
- **Avoid potential risks:** The guide identifies some problem areas not yet addressed by the vendor community.

SOI Reference Architecture

SOI Reference Architecture

The following diagram illustrates a complete SOA reference architecture.

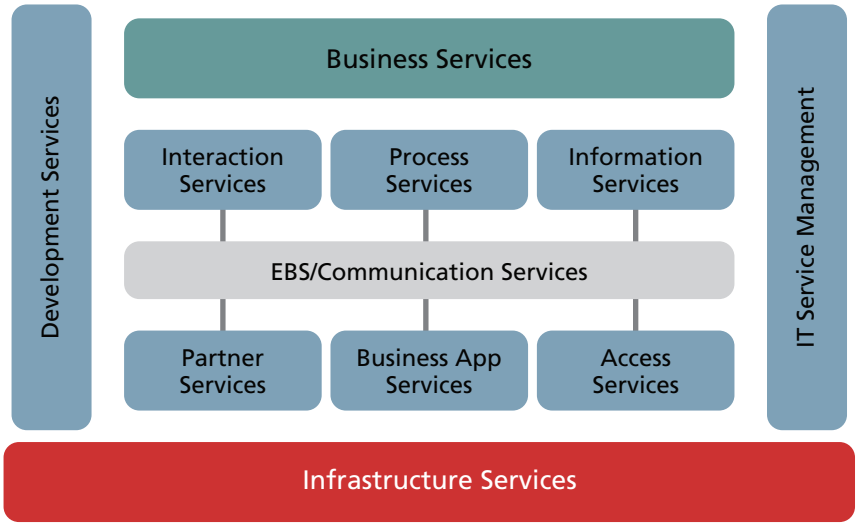

Figure 1: SOA Reference Architecture

Not all of the components in the reference architecture will be touched directly by the infrastructure components necessary to implement a fully SOA-compliant reference architecture. The following components can be directly impacted by the infrastructure reference architecture:

- Development Services
- Access Services
- IT Service Management
- Partner Services
- Infrastructure Services

Figure 2 illustrates the implementation of the SOI Reference Architecture using F5® solutions.

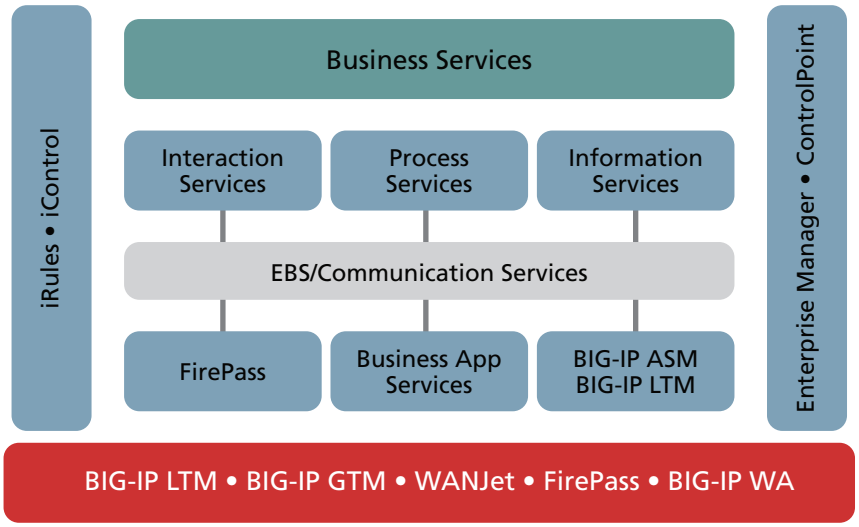

Figure 2: F5 SOI Reference Architecture

Development Services

Development services are those that enable rapid, agile development of services and orchestration of processes for deployment into the SOA ecosystem. Development services include infrastructure components such as registries and design-time governance as well as the frameworks within which services are developed.

Developers can use existing services for distributed, real-time, event-driven, multi-platform applications by taking advantage of standardized SOI. Reuse of existing infrastructure services enables rapid deployment models, resulting in greater agility of IT and through it, the business.

In order for network and security infrastructure to be incorporated into the SOA, it is necessary for those components to provide development services that can be integrated using SOA-focused methodologies. These include:

- Standards-based interfaces to supported capabilities. These support agility through dynamic configuration and modification of the metadata based policies that govern the delivery and security of services within the SOA ecosystem.
- A standards-based platform through which shared network and security services can be implemented and deployed, encouraging reuse and enforcement of organizational standards upon service implementations.

Standards-based Interfaces

F5 supports standards-based interfaces through its iControl™ API, a set of services designed to provide remote, ubiquitous access through which developers and administrators can configure, manage, and monitor all aspects of the F5 BIG-IP® platform.

Administrators can utilize iControl to dynamically configure and modify security and delivery policies in real-time from any Web Services capable client and language. Developers can use iControl to incorporate dynamic modification of the profiles and rules that manage the delivery of applications, from security policies to SLA (Service Level Agreement) enforcement, to application-specific policies that govern all aspects of application delivery.

F5 further supports these capabilities through integration with popular development platforms such as Eclipse and Visual Studio.

Standards-based Platform

F5 offers a standards-based platform and provides visibility and control over application messages through a standards-based scripting language, iRules™. iRules, based on Tool Command Language (TCL), provides developers and administrators with the ability to adapt the rules that govern delivery of applications to best fit organizational and business needs.

iRules can further provide services-based functionality that can be reused across applications, giving developers and administrators the ability to tailor delivery policies according to the unique needs of applications and other infrastructure that comprise the corporate SOA.

F5 supports developers of iRules by providing a stand-alone Windows-based editor as well as a web GUI (Graphical User Interface) through which rules can be developed.

Access Services

Access services are those that provide control over access to services as well as threat prevention services. This includes functionality and features such as secure remote access, encryption of in-flight data, AAA (Authentication, Authorization, and Audit), service virtualization, data integrity validation, and threat prevention.

Developers and network and security administrators can take advantage of access services hosted on a standardized SOI platform. The advantage of utilizing services hosted on a standardized platform includes a reduction in performance penalties due to the lower cost of executing services in a shared context. Another benefit is the ability to reuse services implemented through as policies across applications, departments, and lines of business.

In order for network and security infrastructure to be incorporated into the SOA, it is necessary for those components to provide access services that can be integrated using SOA-focused methodologies. These include:

- Threat prevention services accessible to developers through developer services and available as reusable policy-based services.
- Access control services that take advantage of existing access control resources, such as identity stores.
- Service virtualization that allows the implementation of a proxy façade pattern to restrict the execution of services.

Threat Prevention

F5 provides threat prevention services through a combination of products and development services.

BIG-IP® Application Security Manager™ (ASM) offers pre-packaged, integrated XML and web application threat prevention focusing on preventing exploitation of language, application, and platform-specific vulnerabilities such as SQL injection and XSS (Cross Site Scripting). These services, in addition to custom threat prevention capabilities with F5's iRules scripting language, provide comprehensive perimeter threat prevention services.

Additionally, the BIG-IP platform offers network-layer attack mitigation against a varied category of attacks including denial of service and protocol manipulation.

Access Control

Access control services, whether intended to govern application or network access on the LAN or across the WAN, are supported by F5 in a number of ways.

F5 FirePass® SSL VPN is well suited to providing access control services for both mobile clients and partners alike. It offers comprehensive network admission policy enforcement as well as securing the channel over which mobile clients and partners may access services. FirePass is well suited to the volatile SOA environment. It supports a variety of connectivity methods and clients as well as the ability to craft access control policies, which are dynamic and based on existing conditions at the time the client requests access.

F5's Advanced Client Authentication™ Module, available for the BIG-IP platform, supports identity store based authentication to ensure that clients or partners accessing services are authorized to do so. The Advanced Client Authentication Module takes advantage of existing identity stores within the organization.

Finally, BIG-IP ASM offers fine-grained access control over services, enabling security policies to be enforced not only at the service level, but also at the operational level, an important facet of supporting an SOA implementation.

Service Virtualization

Service Virtualization enables the organization to obscure service endpoints and therefore subtle implementation details from partners and clients by presenting a virtualized endpoint. This virtualized endpoint becomes the location seen by the client or partner, and upon submitting a request, the virtual endpoint is translated to the real endpoint.

For example:
Client endpoint: http://soa.example.com/myendpoint
Real endpoint: http://soa.example.com/service/myservice.asmx

F5 supports service virtualization in its BIG-IP platform through its iRules scripting language.

IT Service Management

IT Service Management provides visibility into applications as well as the supporting infrastructure. Service Management services provide reporting and monitoring services.

F5 supports IT service management through both features core to its BIG-IP platform as well as stand-alone products designed specifically to provide core service management features as well as additional functionality.

F5 Enterprise Manager™ is a stand-alone device designed to aid in the configuration and management of large BIG-IP platform installations. Enterprise Manager provides configuration management services, and can act as an aggregation point for reporting and monitoring services across all BIG-IP platform instances. Additionally, F5's BIG-IP platform contains a number of features that support reporting and monitoring services, as well as the means by which both can be easily tailored to provide better visibility into both the organizational SOA as well as its supporting Service-Oriented Infrastructure.

Reporting Services

F5s BIG-IP platform supports a standard set of reporting services through the aggregation and reporting of performance and usage focused statistics. These statistics include a variety of network-oriented statistics as well as those specifically related to application delivery and its components such as:

- Throughput
- Connections
- Requests

These performance-related statistics are gathered on a device and virtual server basis.

Beyond basic networking and application delivery statistics, F5 provides the means by which organizations can customize reporting services. These customized reporting services can include both performance and network related metrics, as well as business-oriented statistics; this data is gathered through F5 iRules by inspecting and aggregating data carried within business transactions.

Monitoring Services

The F5 BIG-IP platform also offers monitoring services that provide insight into both the platform's performance and health as well as the ability to monitor the health and performance of applications within the organization. Traditional network-oriented monitoring services are supported through the use of SNMP and F5 custom MIBs.

Health, performance, and other statistics can be automatically logged on the F5 platform or exported to a number of industry standard logging mechanisms such as syslog. These statistics and information can also be extracted from the BIG-IP platform through the use of F5's open-standards based iControl API, a set of Web Services that provide interoperable access to the BIG-IP platform via industry SOA standards such as WSDL (Web Services Description Language) and SOAP (Simple Object Access Protocol).

Partner Services

Partner services provide secure, remote access to business processes. Partner services control access to processes and applications based on a variety of parameters; most often the parameters include the role of the partner within the business process. For example, a supplier would likely have different access levels to services within a business process than would a distributor.

These services ensure that the appropriate access rights to business processes and services are granted to partners. They also provide a level of security through encryption services, often implemented using SSL encrypted tunnels.

These services encompass two varieties:

- Access control services should be dynamic, graded, and based on a variety of organizational specific attributes and values that are guided by the corporate security policy.
- Secure access should be a matter of course when exchanging sensitive corporate transactions across public networks and therefore partner services should always be provided via a secured channel for the protection of both the consumer and the producer.

Access Control

F5 FirePass SSL VPN is well suited to providing access control services for both mobile clients and partners alike. It offers comprehensive network admission policy enforcement as well as securing the channel over which mobile clients and partners may access services. FirePass is well suited to the volatile SOA environment, and supports a variety of connectivity methods and clients as well as the ability to craft access control policies which are dynamic and based on the conditions that exist at the time the client desires access.

FirePass can base access to network and application resources based on real-time endpoint attributes that include information such as location, applications running, level of security provided from anti-virus daemons, as well as custom scripts which can be implemented to include more detailed real-time system level information.

F5's ACA (Advanced Client Authentication) Module, available as a module for its BIG-IP platform, supports identity store based authentication to ensure that clients or partners accessing services are authorized to do so. ACA takes advantage of existing identity stores within the organization.

Finally, BIG-IP Application Security Manager offers fine-grained access control over services, allowing security policies to be enforced not only at the service, but at the operational level, an important facet of supporting an SOA implementation.

Secure Access

F5's FirePass SSL VPN secures access to resources through comprehensive access control services, and by securing the channel over which partners communicate with services using SSL. FirePass provides secure access to corporate applications and data using a standard web browser, thus preventing eavesdropping and theft of sensitive data.

Infrastructure Services

The core purpose of any SOI should be to support the availability, scalability, and optimization of SOA-based services and applications. These functions are provided through product features such as:

- Load balancing
- Compression
- Caching
- TCP Connection management
- Advanced health monitoring of services

Availability

Availability—often referred to as high availability, or HA—is the ability of a system to continue performing even when faced with a component failure. In the case of SOA, this generally refers to the ability to continue servicing requests even if one instance of a service becomes unavailable.

In the broader sense, availability has also come to imply adherence to Service Level Agreements (SLA). In this sense, a service not only needs to respond even in the event of a failure, but that it needs to respond within a specified amount of time.

F5 supports high availability environments with a number of features on the BIG-IP® Local Traffic Manager™ (LTM) such as load balancing, layer 7 switching (or content based routing), Quality of Service (QoS), and advanced health monitoring capabilities. Availability services can also be achieved at a global level, distributing load amongst multiple data centers or simply providing failover capabilities to ensure that if one site is inaccessible, messages are routed to a secondary site to ensure continuous access.

BIG-IP load balancing capabilities extend beyond traditional industry standard algorithms to better fit within an SOA through Content-Based Routing (CBR) services. CBR services enable businesses to adjust load and route messages based on factors such as SOAP (Simple Object Access Protocol) headers, the value of elements within messages, and a variety of transport protocol level parameters including port, source IP address, cookies, and destination. CBR services provide value at the edge of the network, in a federation pattern, or as a service attached to an ESB (Enterprise Service Bus), providing more thorough and flexible load balancing options than offered by standard ESB solutions.

At a global level, BIG-IP Global Traffic Manager (GTM) provides the mechanism through which global availability is maintained, providing load distribution and failover services between multiple sites to ensure services are always available. Global services are not restricted to simple failover and availability scenarios. These services can also be used to provide CBR services on a global level, routing messages to specific data centers based on identifying elements in a manner similar to BIG-IP® Local Traffic Manager™ (LTM)'s content-based routing services, but with a much broader scope.

Advanced health monitoring capabilities include the ability to dynamically adjust routing decisions based on the message returned. This provides centralized exception handling that includes removal of potentially sensitive application information as well as rerouting failed transactions to another instance of the service, insuring the service in the event of a failure. This monitoring goes beyond the general "ping" based health checking offered by most software-based load balancing options and is fully customizable to the organization's unique needs.

Scalability

Scalability is the ability to service more consumers and, to do so transparently. F5 supports scalability of services through BIG-IP Local Traffic Manager (LTM) and BIG-IP Global Traffic Manager (GTM) with local and global load balancing and CBR services. These services allow for additional instances of services to be introduced into the SOA and begin servicing consumers transparently.

Through BIG-IP LTMs connection management services, F5 improves the scalability of existing resources. By reusing existing connections when possible, the additional burden of managing TCP sessions is removed from SOA services, which enables those services to handle more concurrent messages. This reduces the hardware and licensing costs required, as well as long term maintenance, power, and cooling costs.

Optimization

Optimization is a huge benefit to SOA implementations, improving the overall performance and utilization of resources across the entire infrastructure. Optimization includes the implementation of protocol-specific standard enhancements that increase the performance of core protocols such as TCP and HTTP, as well as additional features such as dynamic caching, compression, and connection management.

In addition to general optimization of protocols, F5 BIG-IP WebAccelerator™ optimizes the HTTP and TCP protocols specific to the conditions present during a connection between two endpoints. This enhances consumer and producer performance, and takes into account the differences inherent in the implementations of many common service platforms. This aids in enhancing service-to-service performance, as well as client-to-service performance.

Also part of the optimized SOA infrastructure is acceleration. Acceleration occurs either as the result of optimization—by offloading certain resource intensive operations such as SSL encryption and decryption from services to the infrastructure—or through the use of compression technology to reduce the total size of messages. Compression is particularly useful in accelerating the exchange of messages in inter-organizational or inter-office scenarios when large messages must traverse the bandwidth limited WAN or the Internet. XML-based messages are particularly well suited to compression because they are text-based. Compression services are provided by BIG-IP WebAccelerator and BIG-IP LTM.

Offloading SSL encryption and decryption from services also provides gains in performance and efficiency within the SOA. Generally speaking, when SSL services are offloaded to the infrastructure, the task of managing the certificates and keys are also offloaded, providing gains in ease of administration by centralizing management and enlisting the aid of hardware-assisted acceleration of cryptographic services. F5 supports SSL offload and centralization of certificate and key management with BIG-IP LTM.

Conclusion

The infrastructure services layer within the SOI Reference Architecture is often glossed over and rarely discussed other than as a layer within the overall reference architecture. But the infrastructure services layer can have a positive impact on the reliability, security, and performance of a SOA implementation. The infrastructure should be given consideration in all phases of the maturity process, but especially in the web application development stage. Consider carefully those services that can be enabled, integrated with, or moved to the infrastructure layer and what impact such a decision may have on the organizational SOA.

F5 Networks, Inc.
Corporate Headquarters
401 Elliott Avenue West
Seattle, WA 98119
+1-206-272-5555 Voice
(888) 88BIGIP Toll-free
+1-206-272-5556 Fax
www.f5.com
info@f5.com

F5 Networks
Asia-Pacific
+65-6533-6103 Voice
+65-6533-6106 Fax
info.asia@f5.com

F5 Networks Ltd.
Europe/Middle-East/Africa
+44 (0) 1932-582-000 Voice
+44 (0) 1932-582-001 Fax
emeainfo@f5.com

F5 Networks
Japan K.K.
+81-3-5114-3200 Voice
+81-3-5114-3201 Fax
info@f5networks.co.jp

WP-SOA-INFRASTRUCTURE-REF-ARCH 01/08