

## Designing a Mobility Services Architecture

### Introduction

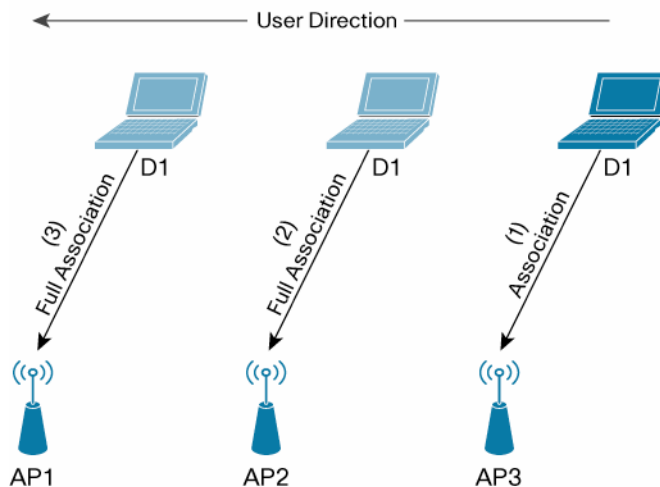
To deliver true business mobility, IT must take a practical approach focused on unifying networks, managing the wave of mobile devices, and enabling mobile application development. IT must evolve existing wireless networks to support a variety of new mobility applications. What's more, these applications must be able to extend across multiple networks and scale from small businesses to the very largest enterprises.

To achieve this transition, IT must transform the wireless LAN into a mobility network by abstracting the application layer from the network layer, effectively allowing for the delivery of mobile applications across networks, including Wi-Fi, Ethernet, cellular, WiMAX and RFID. The Cisco® 3300 Series Mobility Services Engine is at the heart of this mobility architecture evolution and provides an open API that allows a broader ecosystem of partners to access network intelligence to develop industry relevant mobility solutions.

### Evolution of Wireless Architecture

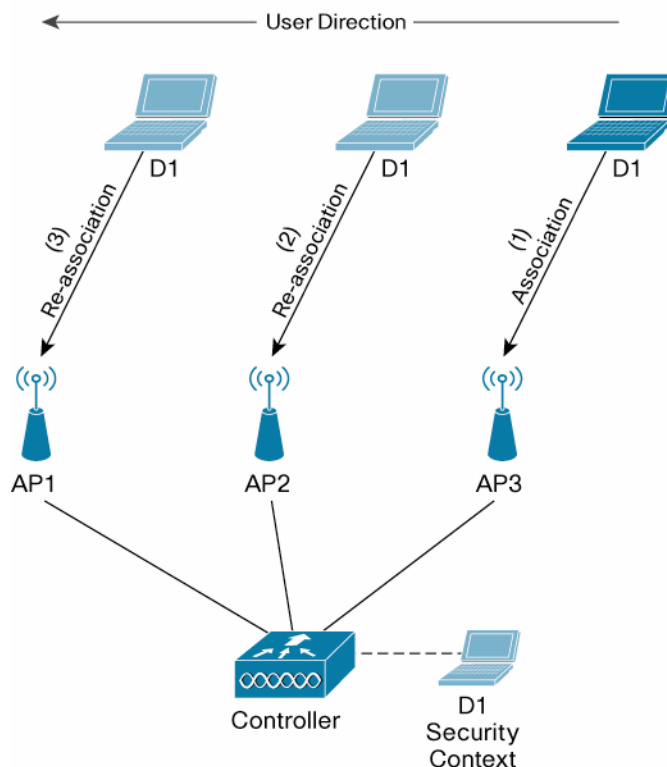
The initial deployments of enterprise wireless networks primarily using 802.11 technologies provided the user with basic connectivity to the corporate network. As shown in Figure 1, access points (APs) were standalone systems that acted as gateways between the wireless client and the corporate backbone network. There was no communication or coordination between access points and therefore they were managed as independent systems. As the network grew in size, this became a major problem in managing the network and limited the adoption of wireless technology. For the wireless client, each interaction with an access point would be treated as an independent system and therefore no context sharing could occur. Simple activities such as client roaming were often difficult to perform. As users roamed across access points, inefficient and costly communication would be repeated to reestablish the client's connection to the corporate backbone network.

**Figure 1.** Standalone Wireless System



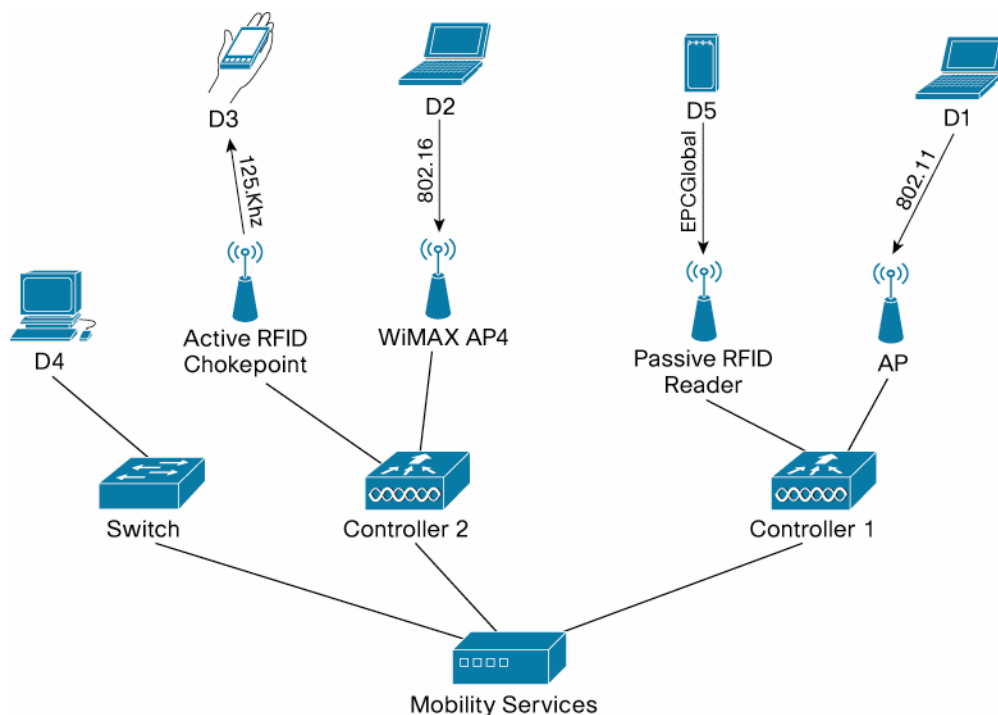
The next phase of enterprise wireless network development introduced unified 802.11 wireless networks incorporating a wireless controller, as shown in Figure 2. The unified wireless network solved several key problems that existed with standalone wireless systems. First of all, the wireless controller provided management and control plane aggregation for all of the access points connected to it. This significantly improved manageability and ease of deployment of large numbers of access points in a corporate network. For the first time, access points could be managed in a coordinated manner to help solve the problems introduced by the dynamic wireless environment encountered in offices, hospitals, shop floors, and so on. Secondly, the wireless controller provided the ability to store the client's security context. Now, as clients roamed from access point to access point, the controller maintained the security context and avoided time consuming, full reauthentication of the client. This improved performance and scalability for both the client and the wireless network as it reduced load on both systems.

**Figure 2.** Unified Wireless System



The newest phase of enterprise network development expands on the unified wireless network architecture and applies unified management and control plane techniques to a new set of network access technologies at the edge of the network, including WiMAX, active RFID, passive RFID, and Ethernet (Figure 3). As the unified network architecture scales to incorporate multiple means of accessing the network, it becomes important to provide services that extend across all networks in an integrated manner. As networks and management and control planes unify, it becomes possible to deliver new mobility services that were previously difficult to create because of the disparate nature of each network.

**Figure 3.** Mobility Services Architecture



The following sections describe in more detail the Mobility Services Architecture and its most important components.

### Mobility Services Architecture

Mobility services are a set of value-added network services that consolidate intelligence from various points in the network to enable and optimize the delivery of business mobility applications. This intelligence has typically been highly distributed throughout the network, resulting in complex service provisioning and management. The combination of the services, control, and data planes adds complexity and limits the network's ability to adapt to new services while maintaining consistent performance. As businesses start to design their networks to natively support mobility, the combination of the services, control, and data planes becomes a limiting factor in the flexibility and scale the network can provide to support mobile applications.

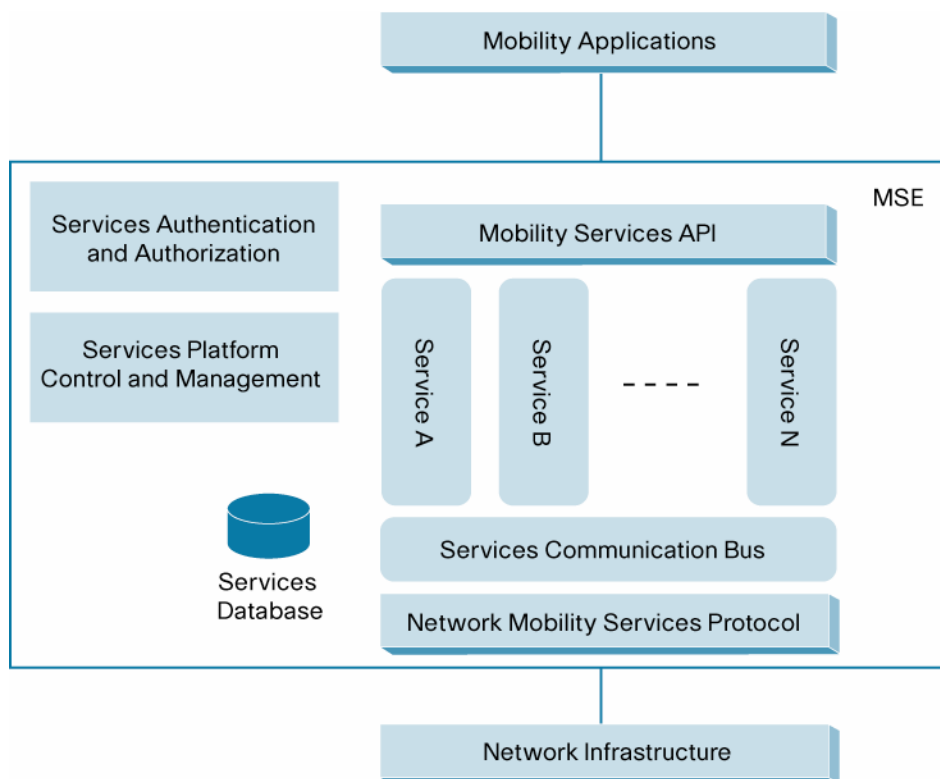
The answer lies in a centralized services architecture. While still critical to the ability of networks to provide the intelligence for optimal mobile application performance, mobility services should be abstracted from the control and data planes in order to be centralized into a services engine. This centralization of services offers several benefits, including scalability and improved provisioning and management. Additionally, a centralized services architecture removes the direct linkage between service and network, allowing services to extend across a variety of networks, including Wi-Fi, Ethernet, WiMAX, and cellular.

Increasingly, the mobility network must be able to support a multitude of applications. The true value of mobility services is delivered via their ability to enhance application performance by providing real-time information from the network and related applications. This cross-pollination of network and application intelligence has a synergistic effect, augmenting the richness and breadth of the types of mobility solutions that can be delivered. At the same time, a critical component of services delivery is helping to ensure that third-party applications have a standard interface by which they can access this network and application intelligence. The Cisco Mobility Services Engine supports an open API based on Simple Object Access Protocol/Extensible Markup

Language (SOAP/XML), which provides northbound access to these services to an ecosystem of mobility application partners. With service intelligence centralized from the control network into the Mobility Services Engine, IT can open access to the API without concern about disruption to the underlying production network.

The Cisco Mobility Services Engine can run one or more mobility services software instance. Figure 4 shows the high-level architecture of the Mobility Services Engine.

**Figure 4.** Cisco Mobility Services High-Level Architecture



A mobility service is a software instance running on the MSE and has the following characteristics:

- The service acts across multiple edge technologies, such as 802.11 wireless and 802.3 wired networks.
- The service provides a value-added function across multiple network elements.
- The service provides an interface to that value-added function to external applications and servers using a mobility service API.
- The service adds intelligence to the network through its function to enhance the usability of the network.
- The service provides visibility into the network that applications and servers would not otherwise easily obtain.
- The service can be combined with other mobility services to achieve higher-order functions.
- The service is manageable via the mobility service API.
- The service can be deployed across multiple MSE to scale the function it provides.

### The Limits of Conventional WLAN Architectures

Existing wireless LANs are optimized for the delivery of wireless data. The use of a WLAN controller helps to ensure scalable, high performance wireless connectivity. However, existing WLAN equipment such as WLAN controllers is not well equipped to deliver the services that the business requires to support the full breadth of emerging mobility applications. WLAN architectures excel at delivering reliable, pervasive wireless connectivity, which is especially important for mobile data applications such as e-mail, Web browsing, and file transfers. The architectures are largely designed to optimize throughput of wireless packets while ensuring a low-operational impact for deploying and managing elements such as mobile devices and wireless access points.

As the central part of existing WLAN architectures, the WLAN controller delivers real-time management and network optimization for all network elements. Its primary function is to centralize the management and policy enforcement for network operations. The WLAN controller is not optimized as a service delivery platform. Existing controllers are designed to provide a high-performance data path and have only a modest CPU for control processing. At best, they have limited memory and storage space and generally rely on an operating system that is optimized for data path control.

Additionally, traditional WLAN architectures are incapable of delivering the support for mobility applications that businesses demand. First of all, alternative WLAN offerings have no ability to incorporate other network access technologies, but are limited to Wi-Fi only. These networks are closed platforms that have no standard interface to support third-party development of applications. A significant number of development cycles is therefore required for any new application or service to be integrated into the network. Second, the architectures of WLAN only networks have no ability to centrally manage multiple network environments, nor can they deliver scalable services and applications across a variety of networks.

Only by evolving to a mobility network where applications and services are separated from the control plane, can IT truly deliver the services required. A true mobility network centralizes services delivery, allowing for a modular approach to services enablement. In this way, services can scale across multiple places in the network and over a variety of networks.

As an extension to existing wireless networks, the Cisco Mobility Services Engine integrates directly with the WLAN controller and allows the controller to fully dedicate its resources to the consistent and reliable delivery of packets, while allowing service enablement and scale to be handled by the Mobility Services Engine. In this way, the business can adapt more rapidly to shifts in application requirements.

In summary, the mobility service architecture is preferable to existing WLAN architectures because it has the following attributes:

- Unification of services across multiple networks (for example, Ethernet, Wi-Fi, cellular)
- Support for an open API for the development of enterprise applications
- Ability to manage services across multiple networks
- MSE support for software modularity, leaving the controller to process traffic
- Services-oriented hardware platform
- Platform that scales for application evolution (investment protection)

## Mobility Services Software

The Cisco Mobility Services Engine is a platform that is designed to support a variety of services, loaded onto the platform as a suite of software. This section introduces some of the major types of mobility services software.

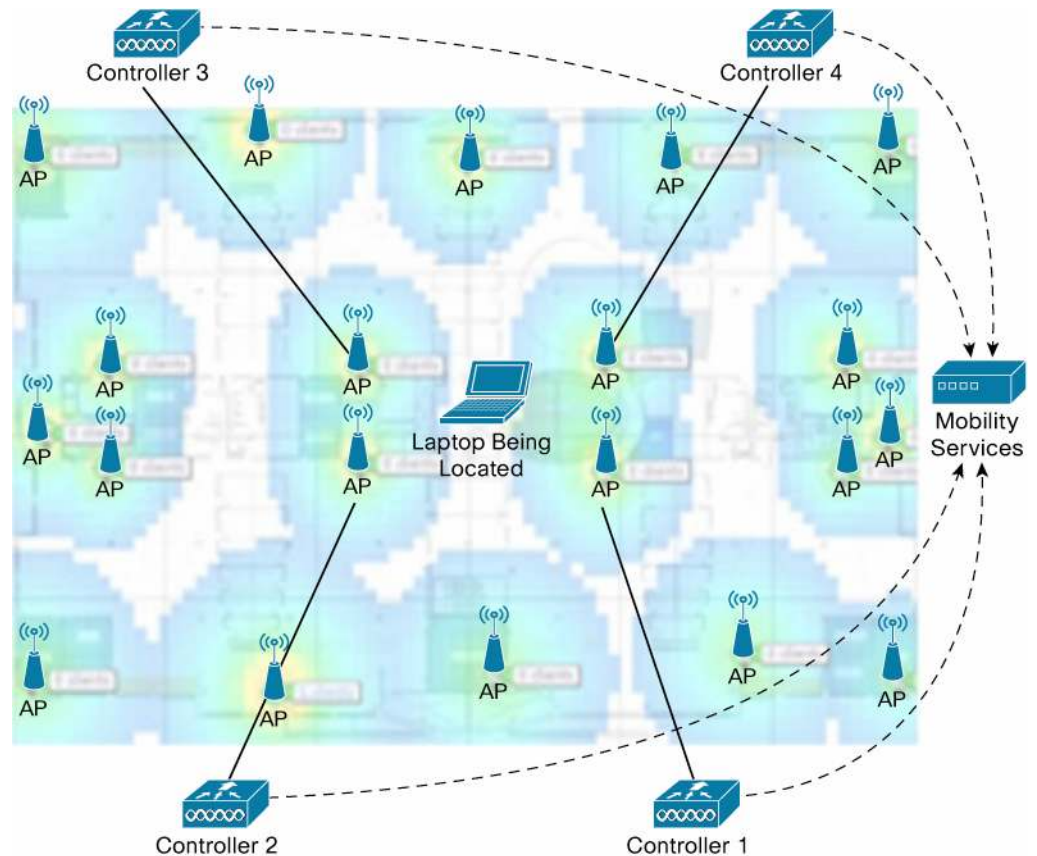
### Context-Aware Software

Context-aware software captures and integrates into business processes detailed contextual information about such things as location, temperature, availability, and applications used. Functionally, context-aware software is a prime example of a mobility service that covers multiple network access types.

Applications and servers that consume contextual location information care primarily about that information on the client side, regardless of how that information was determined or how the client connects to the underlying network. Through its mobility services API, context-aware applications and services provide a single, unified view of contextual information, including location information for network clients. Context-aware services enable both queries for context information as well as registration for asynchronous events based on movement, containment, absence, and other advanced context-aware event triggers. Applications and servers that consume context-aware information no longer have to be written to multiple disparate APIs across a wide variety of technologies. The context-aware service running on the Cisco Mobility Services Engine has simplified the task of building context-aware applications and services.

Having the service run on the Cisco Mobility Services Engine rather than on each individual switch or wireless controller also provides a more scalable network design. This is because of the way in which location is determined for a network comprised of wireless access points. In a typical enterprise wireless deployment, there are many access points deployed in a physical space such as a floor. For redundancy and scalability, those access points are connected to multiple wireless controllers rather than a single wireless controller, as shown in Figure 5.

**Figure 5.** Typical Wireless-Controller-to-Access-Point Deployment for Location Services



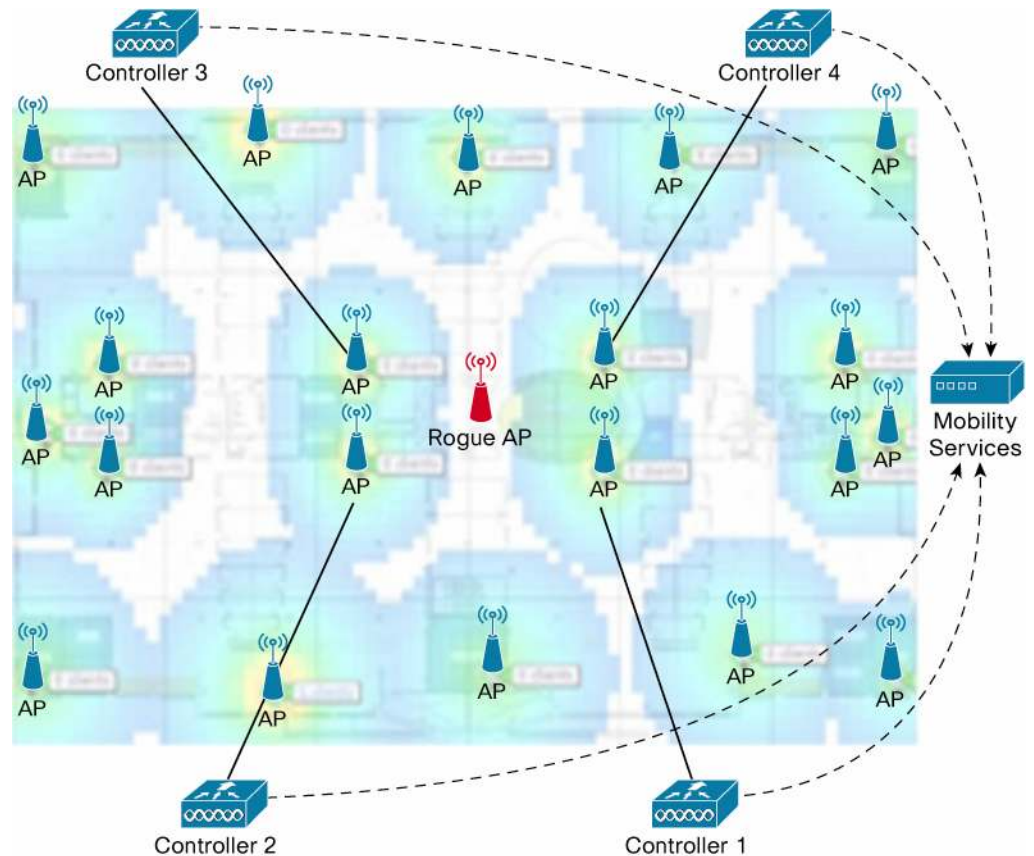
To calculate the location of the laptop shown in Figure 5, a MSE running the context-aware software has to collect information from all controllers (1 to 4) with the access points in the physical environment around the client, not just the set of access points connected to a single wireless controller. For this reason, the context-aware software must run on an appliance or server that aggregates all access point measurements from multiple wireless controllers.

In addition, location calculations have to be performed at a very high rate, and within a matter of a few seconds, for context consumers to be able to take advantage of context-aware information. Having a dedicated platform that is tuned for this task results in a more scalable service that can meet the needs of high-performance applications that use contextual information.

### Adaptive Wireless IPS Service

An adaptive wireless intrusion protection system (IPS) service is another example of a mobility service that benefits from the centralization of security intelligence into the MSE. The adaptive wireless IPS service collates and analyzes the RF information gathered from each wireless controller, allowing the IPS service to detect attacks by rogue devices or anomalies such as rogue access points that may lead to a security breach. Figure 6 shows how the access points monitor the RF environment in each part of a floor. Each access point connects to different wireless controllers and those wireless controllers connect to the wireless IPS service.

**Figure 6.** Typical Wireless Controller Detection of Rogue Access Points



Because each wireless controller may have only a partial view of the network, the MSE running the adaptive wireless IPS service is best placed to provide the analysis necessary to detect all security issues. Furthermore, the centralization of intrusion prevention capabilities provides for greater extensibility across multiple network access technologies, including Wi-Fi, Ethernet, and WiMAX.

### Future Services

The mobility services architecture has been designed to add new mobility services to the product line easily and quickly. Adopting an architecture that can be easily adapted to new requirements and new functions is critical for customers investing in the MSE technology. As customers and their devices become more mobile, the applications and the network will become more mobility-friendly, aided by MSE architecture. New mobility services, such as client provisioning and mobile intelligent roaming, will become available to further enhance the MSE product line.

### Mobility Service Interfaces

One of the important design aspects of the mobility services architecture is the introduction of secure and scalable interfaces between service components.

The mobility service architecture interfaces fall into three categories, with each category representing a particular area or layer of the network.

Providing a consistent design in each category is an important architectural goal. For each category of interface, the following guidelines were followed:

#### Common Authentication and Authorization

Principle:

- Each component and service using an interface uses the same authentication and authorization mechanisms.

Benefits:

- The authentication and authorization mechanisms are robust and extremely secure.
- Authorization policies can be applied uniformly across each interface.

### **Highest Efficiency Possible**

Principle:

- Each service interface must be tuned to provide the highest performance for the service.

Benefits:

- High-performance interfaces remove any potential latency from the service.
- High performance provides maximum service applicability.

### **Common Object Model**

Principle:

- Each component and service using an interface represents common, shared objects using the same schema as other components and services.

Benefits:

- A common object model provides maximum reuse of objects and interactions between services.
- A common object model enables services to be combined easily for  $1+1=3$ . The combination of multiple services delivers synergies that ensure a greater benefit than the sum of the inputs.

### **Common Semantics and Behavior**

Principle:

- Each component and service using an interface behaves the same way regardless of the service being provided.

Benefits:

- Adhering to common semantics and behavior reduces the development time for new components or services.
- Consistent behavior makes it easier to debug problems in a customer environment.

### **Common Syntax and Style**

Principle:

- Each component and service interface provides a consistent terminology and syntax for methods, objects, and function calls.

Benefits:

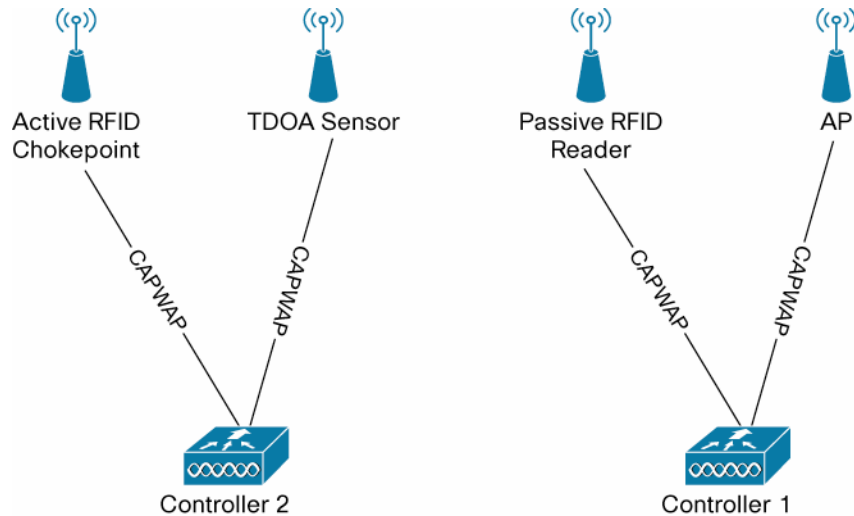
- Common syntax and style makes it easier to get familiar with new services and functions over time and thus reduces development time.

### **CAPWAP: Wireless Controller to Access Technology**

This interface category provides management, control, and data plane communication between the wireless controller and the set of access technologies, such as an access point, that interface

with the endpoint devices. As shown in Figure 7, the protocol between the wireless controller and many access networks is based on the protocol [RFC 3990] being developed by the Control and Provisioning of Wireless Access Points (CAPWAP) Working Group of the IETF.

**Figure 7.** CAPWAP Protocol Interfaces

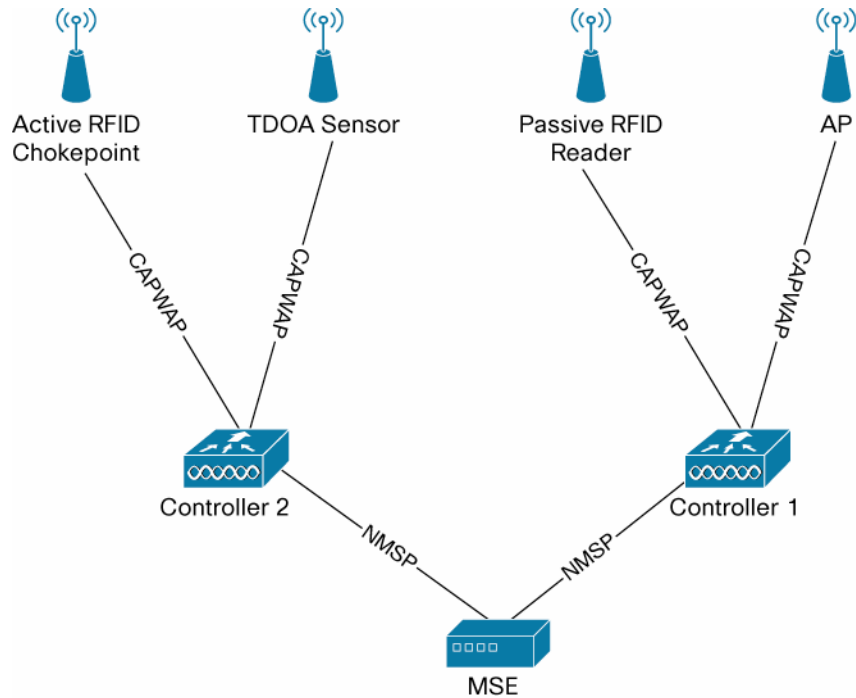


When there is a common protocol between the wireless controller and access technologies, the network administrator can administer the network devices in a consistent manner because the protocol provides consistent functional behavior for security, configuration, and monitoring features. As new access technologies adopt the CAPWAP standard, IT can unify networks in a consistent way without being limited by proprietary solutions.

#### **NMSP: Mobility Services Engine to Wireless Controller and Switch**

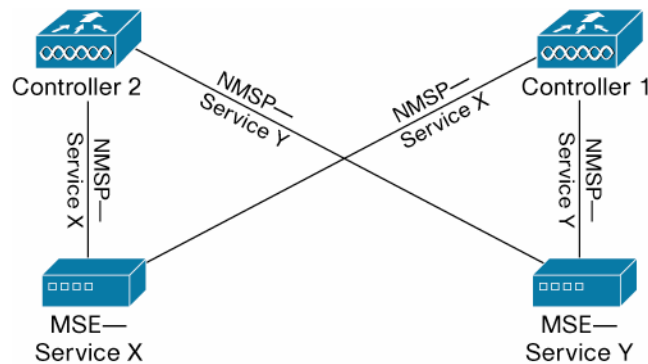
This interface category provides management and data plane communication between the Mobility Services Engine and the wireless controllers. As shown in Figure 8, the protocol between the MSE and wireless controllers is based on Network Mobility Services Protocol (NMSP).

**Figure 8.** NMSP Protocol Interfaces



NMSP provides a single, common protocol between the MSE and wireless controllers to communicate all service-level information. Each wireless controller advertises the services that it provides to any of the mobility services engines that may connect to it. When a mobility services engine connects to the wireless controller, it subscribes to the set of services that it wishes to consume information for. Information shared by the mobility controller is categorized by service and includes location measurements, statistical data, security context data, and so on. As the wireless controller knows which services are being consumed by a particular MSE, the wireless controller can intelligently tune its own performance for a particular service consumer and also what information is transmitted to a particular MSE, as shown in Figure 9.

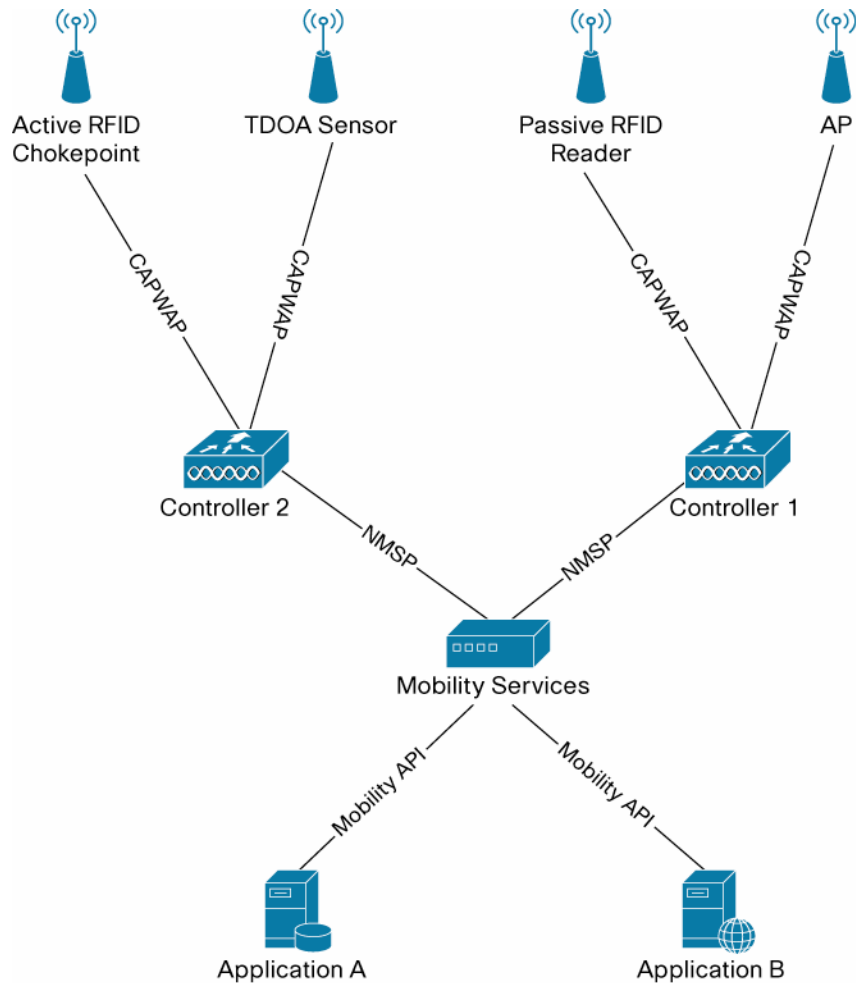
**Figure 9.** Multiple MSEs with Different Services and Controllers



**Open API: Mobility Services Engine to Application**

This interface category provides management and data access to the services running on the MSE, as shown in Figure 10. The mobility services API is an interface that provides management and data access to applications that interact with mobility services.

**Figure 10.** Mobility Services API

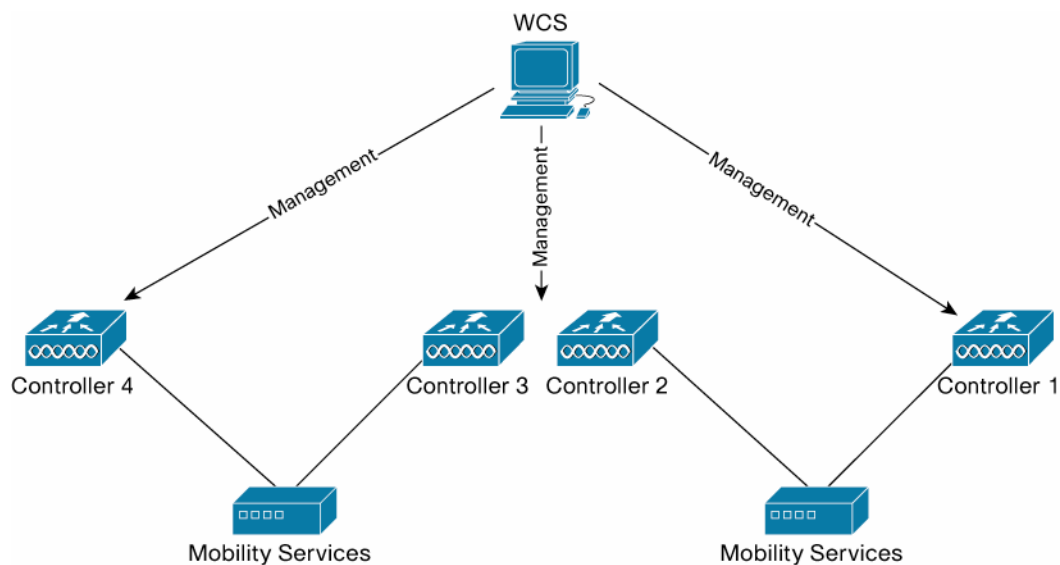


The mobility services API is the critical interface to external applications and servers that make use of the mobility services. It is both the administrative interface for management of the services and also real-time interface for consumers of the services. The main objective of the mobility services API is to represent services and services data to service consumers in a consistent manner. Instead of individual services that have separate mechanisms for authentication and authorization, separate style guides for how common functions are performed, separate data representations for common objects such as a controller or access point, the mobility services API provides a consistent view of this data. This provides application integrators with a simplified task when considering using multiple services on the MSE. It also allows the network administrator to easily manage services—for example, by applying policies for access control across a set of services in the same way—rather than having to adapt policy management for each individual service.

### Services Management

The mobility services architecture includes a management solution that covers the network components, Mobility Services Engine, and also the services that run in the network (Figure 11). By providing an integrated management solution for configuration, fault-analysis, and monitoring, enterprise administrators do not have to install, learn, or maintain separate products for separate services or the network components providing those services.

**Figure 11.** WCS Management System



With the integrated management system, the administrator can administer the Mobility Services Engine to configure the services running on a particular MSE, start or stop a particular service, and view the status of each service. For each service, the administrator can use configurable performance parameters from a single management system to manage its behavior across the various network components providing that service. With the advantage of an integrated service view in the management system, the administrator can easily choose which services run and where those services are in the network.

### Scaling

The scalability of the mobility services architecture is a fundamental requirement. Scalability covers two main aspects: high availability and clustering.

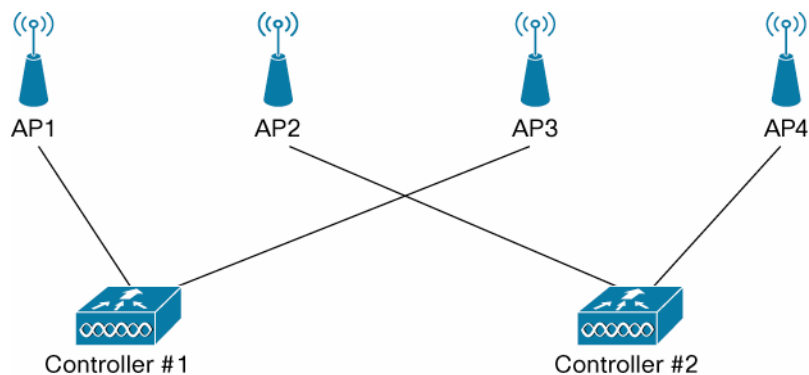
#### High Availability

High availability of the services and network components providing those services is a critical requirement. The following sections describe the various ways high availability is achieved by the mobility services architecture.

##### Controller High Availability

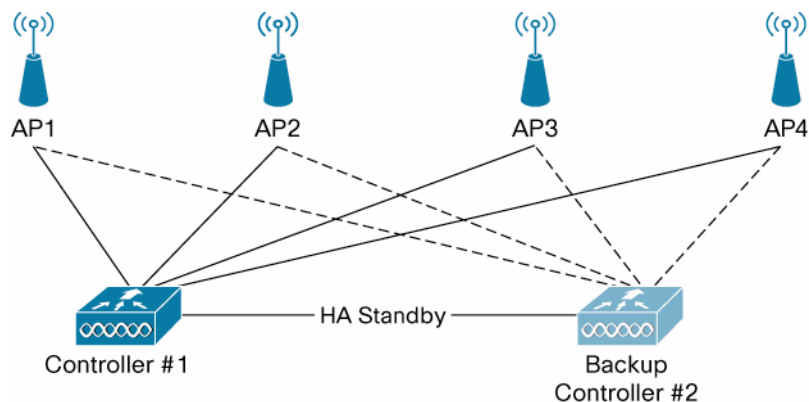
One mechanism commonly used for higher availability is to deploy access points in a manner where every other access point is terminated by a different controller. This is often referred to as a “salt-and-pepper” deployment (Figure 12).

**Figure 12.** Salt-and-Pepper Deployment



An alternative mechanism for achieving the high availability of controllers is to have a backup controller providing hot-standby capability in case the primary controller goes down for any reason (Figure 13). In this case, the access point switches from the primary controller to the secondary controller in the same manner as in the salt-and-pepper deployment. However, the secondary controller will already be in sync with the primary controller before it goes down, and so is able to take over as the primary controller much more easily and quickly.

**Figure 13.** Controller High-Availability Pair



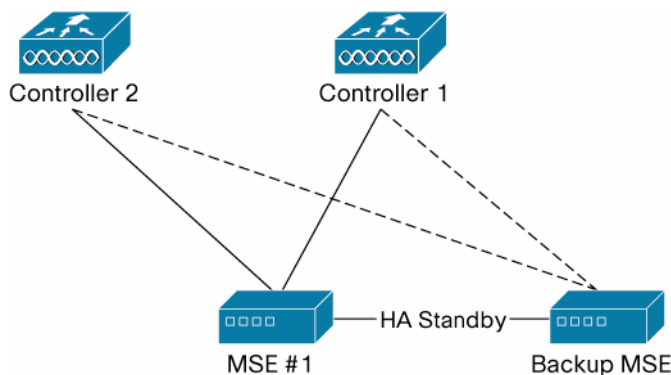
#### Mobility Services Engine High Availability

MSE high availability helps to ensure that services running on a particular MSE are active at all times with very little or no downtime due to failure. There are multiple levels of high availability for the MSE.

The first level is within a single hardware appliance, where the appliance has both redundant hard-drives and redundant power supplies.

The second level is across hardware appliances, where the entire physical system and the services running on that system are replicated using hot-standby technologies, as shown in Figure 14. In this example all service data is synchronized between the two systems in real-time to help ensure that if one of the physical appliances goes down, the mobility controllers and applications using the MSE can continue to provide and consume services.

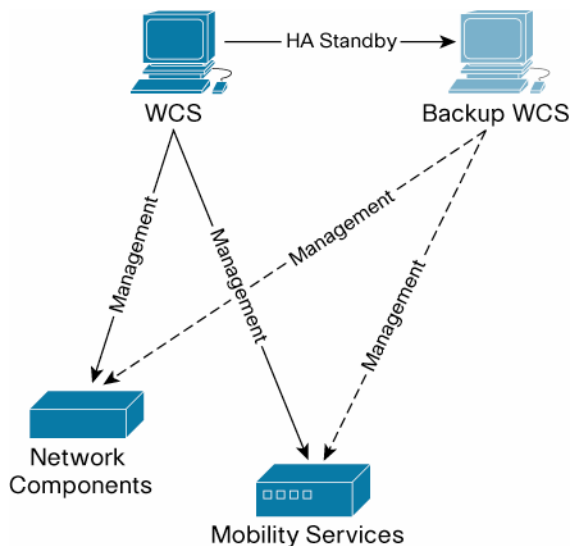
**Figure 14.** MSE Hot Standby



### Management System High Availability

The management system high availability helps to ensure that service management can be performed even if the system fails. As shown in Figure 15, the management system runs hot-standby protocols to ensure all service and network management data is available immediately on the backup management system as soon as failure of the primary management system is detected. Once the backup management system becomes active, the network components, including the mobility controller as well as the MSE, will automatically become aware of the new server providing management.

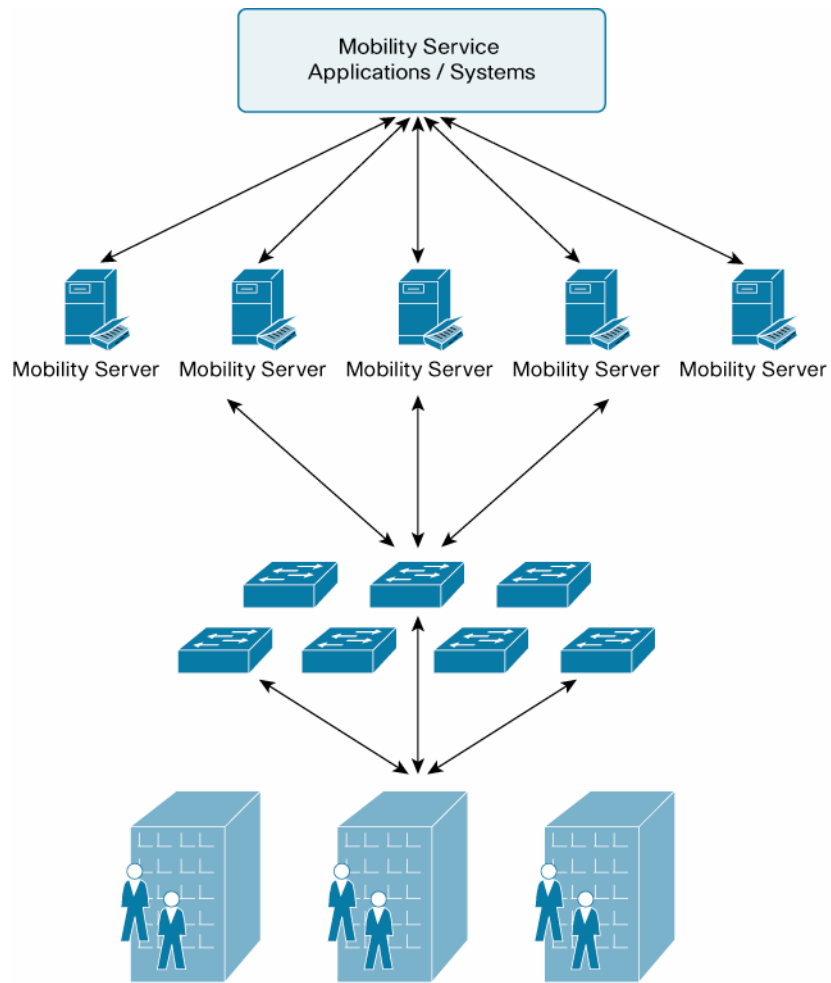
**Figure 15.** 15 Management System Hot Standby



### Clustering

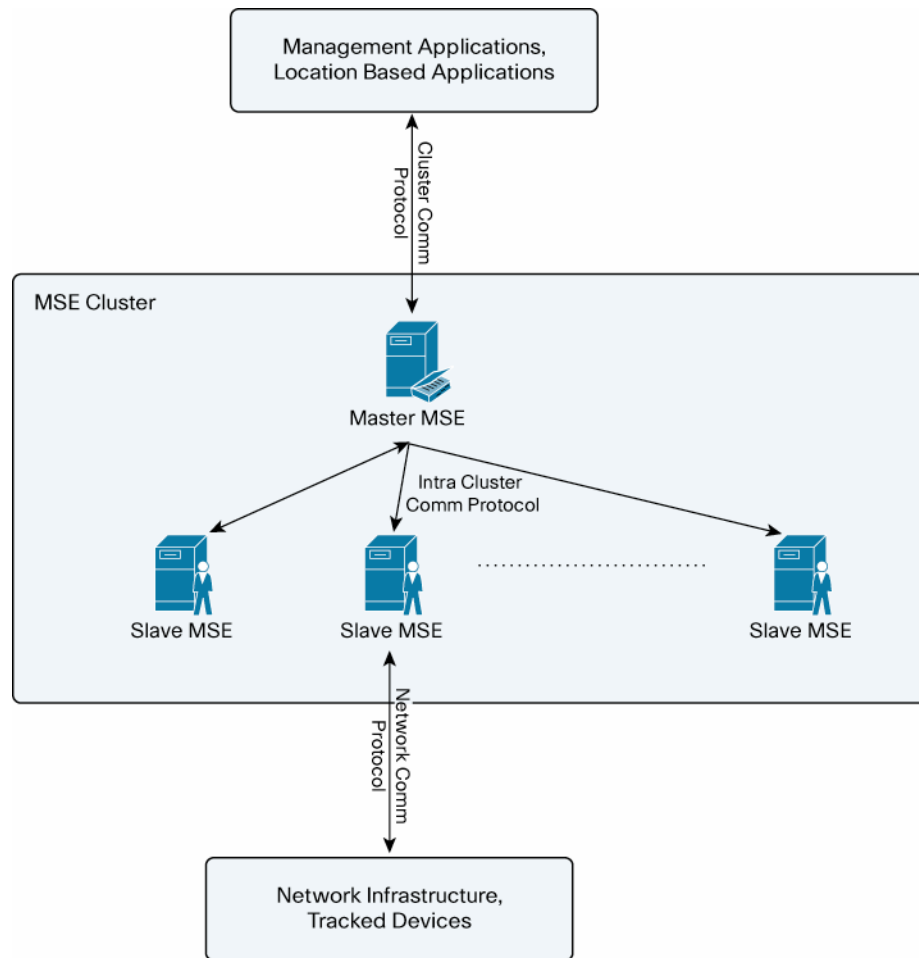
MSE clustering makes it possible to increase the number of services and the size of the networks that the services run in without applications or clients knowing that those services are provided by more than one physical MSE. As shown in Figure 16, as more mobility servers are deployed in the network, how the network and users of the network services map to those mobility servers is a critical architectural issue.

**Figure 16.** Network Scaling Issues



By introducing MSE clustering as shown in Figure 17, managing the services across multiple MSEs is greatly simplified. In this architecture, the master MSE coordinates the management of each slave MSE running one or more services.

**Figure 17.** 17 High Level Architecture of MSE Clustering



## Conclusion

Business has gone mobile. IT must respond to the emerging mobility challenges to enable this business transformation. By evolving the wireless LAN to a true mobility network, IT is better equipped to manage the wave of new mobile devices, unify multiple networks via a single control interface, and build an open platform for the development of mobility applications. The introduction of the Cisco Mobility Services Engine allows for the centralization of service enablement and delivery, leaving the wireless LAN controller to perform its primary task of efficiently managing the data plane. The Cisco Mobility Services Engine takes a modular approach to service delivery by supporting a suite of software capable of sourcing service intelligence, such as context and intrusion prevention information, from a variety of networks. The centralization of this service intelligence provides a logical opening point for an API that allows an ecosystem of third parties to develop industry-relevant solutions based on the inherent network intelligence. Only by evolving the WLAN into a true mobility network can the business put its employees, partners, customers, and assets in motion.



**Americas Headquarters**  
Cisco Systems, Inc.  
San Jose, CA

**Asia Pacific Headquarters**  
Cisco Systems (USA) Pte. Ltd.  
Singapore

**Europe Headquarters**  
Cisco Systems International BV  
Amsterdam, The Netherlands

Cisco has more than 200 offices worldwide. Addresses, phone numbers, and fax numbers are listed on the Cisco Website at [www.cisco.com/go/offices](http://www.cisco.com/go/offices).

CCDE, CCENT, Cisco Eos, Cisco Lumin, Cisco Nexus, Cisco StadiumVision, the Cisco logo, DCE, and Welcome to the Human Network are trademarks; Changing the Way We Work, Live, Play, and Learn is a service mark; and Access Registrar, Aironet, AsyncOS, Bringing the Meeting To You, Catalyst, CCDA, CCDP, CCIE, CCIP, CCNA, CCNP, CCSP, CVP, Cisco, the Cisco Certified Internetwork Expert logo, Cisco IOS, Cisco Press, Cisco Systems, Cisco Systems Capital, the Cisco Systems logo, Cisco Unity, Collaboration Without Limitation, EtherFast, EtherSwitch, Event Center, Fast Step, Follow Me Browsing, FormShare, GigaDrive, HomeLink, Internet Quotient, IOS, iPhone, iQ Expertise, the iQ logo, iQ Net Readiness Scorecard, iQuick Study, IronPort, the IronPort logo, LightStream, Linksys, MediaTone, MeetingPlace, MGX, Networkers, Networking Academy, Network Registrar, PCNow, PIX, PowerPanels, ProConnect, ScriptShare, SenderBase, SMARTnet, Spectrum Expert, StackWise, The Fastest Way to Increase Your Internet Quotient, TransPath, WebEx, and the WebEx logo are registered trademarks of Cisco Systems, Inc. and/or its affiliates in the United States and certain other countries.

All other trademarks mentioned in this document or Website are the property of their respective owners. The use of the word partner does not imply a partnership relationship between Cisco and any other company. (0805R)

Printed in USA

C11-475384-00 05/08