AT&T's MPLS OAM Architecture, Experience, and Evolution

Jerry Ash, Li Chung, Kevin D'Souza, Wai Sum Lai, Harmen Van der Linde, and Yung Yu, AT&T Labs

ABSTRACT

This article provides an overview of AT&T's MPLS OAM architecture, and gives examples of operational experience. Hallmarks of the architecture are a single, converged, and integrated MPLS/optical network, and the evolution to fully automated, zero-touch network operation. The Concept of One converged IP/MPLS architecture will reduce operations, development, and capital costs. The Concept of Zero aims to bring full automation for every human-to-computer interaction currently required for setting up and maintaining network services, delivering services to customers in real time with zero defects and cycle time, and supporting both a network as well as an operational environment with six nines reliability. This approach effectively opens the network to the customer, enabling new levels of customer network management, service creation, and ordering, and empowering enterprise customers with the tools to create their own network services as they transform their own internal networks. In the article we describe AT&T's MPLS-enabled services, the corresponding MPLS operations architecture (including MPLS MIBs), our MPLS OAM operational experience, and MPLS OAM evolution needs for MPLS MIB enhancements and new network capabilities. By applying technologies such as artificial intelligence, self-healing/self-identifying network elements, expert systems, rules-based processes, and automatic speech recognition, the architecture will migrate from a predictive network that monitors, correlates, and recommends action; to an adaptive network that monitors, correlates, and takes action; to a cybernated network that has integrated components that dynamically manage by business rules and policies. We give several examples of how AT&T is already investing in and implementing this future vision, and conclude by challenging network researchers, developers, and key industry players to apply new technologies in fully realizing the operational vision.

INTRODUCTION

A guiding principle of AT&T's architecture evolution is a converged Concept of One IP/multiprotocol label switching (MPLS) network, which through network consolidation will reduce operations, development, and capital costs. As illustrated in Fig. 1, all layer 3 and 2 services will be supported on a common MPLS backbone network and multiservice edge (MSE) platform. The MSE provides access to all AT&T services onto the IP/MPLS network and the required protocol conversion/encapsulation for supporting the customer services over the MPLS core network. AT&T layer 3/2 services can be supported on any port and at any speed.

Customers connect to AT&T's network services at the MSE through the multiservice access (MSA) network, which provides grooming, aggregation, transport, and protection/restoration. It also supports services confined to a metropolitan area. The functional architecture of the MSA network provides packet-aware transport to the MSE and MPLS core network, which enables virtual circuit access to services delivered by the MSE. MSA packet transport uses virtual circuits with MPLS-based layer 2 encapsulation that are independent of the protocol and interface through which the customer or third-party access provider meets the AT&T network. The MSA architecture provides dual-homing capabilities for enhanced reliability. With reference to the transport of packet services, virtual circuit access enables:

- More flexibility in bandwidth, interfaces, and protocols
- Reduced dependence on digital crossconnects
- Capital cost savings by reducing the need for channelized time-division multiplexing (TDM) interfaces
- Faster, simpler provisioning of packet services
- Support for switched metro Ethernet services

Along with the IP/MPLS network convergence, a fully automated Concept of Zero network operation will evolve, to include automated processes, operations support systems (OSSs), network, and services. Process automation includes the consolidation of similar functions across organizations, and deployment of automated rules, workflow, auto-inventory, e-enabled services, and supply chain. Operations processes will evolve from semi-manual activities in shared work centers to fully automated operations activities in consolidated centers. OSS automation

0163-6804/04/$20.00 © 2004 IEEE

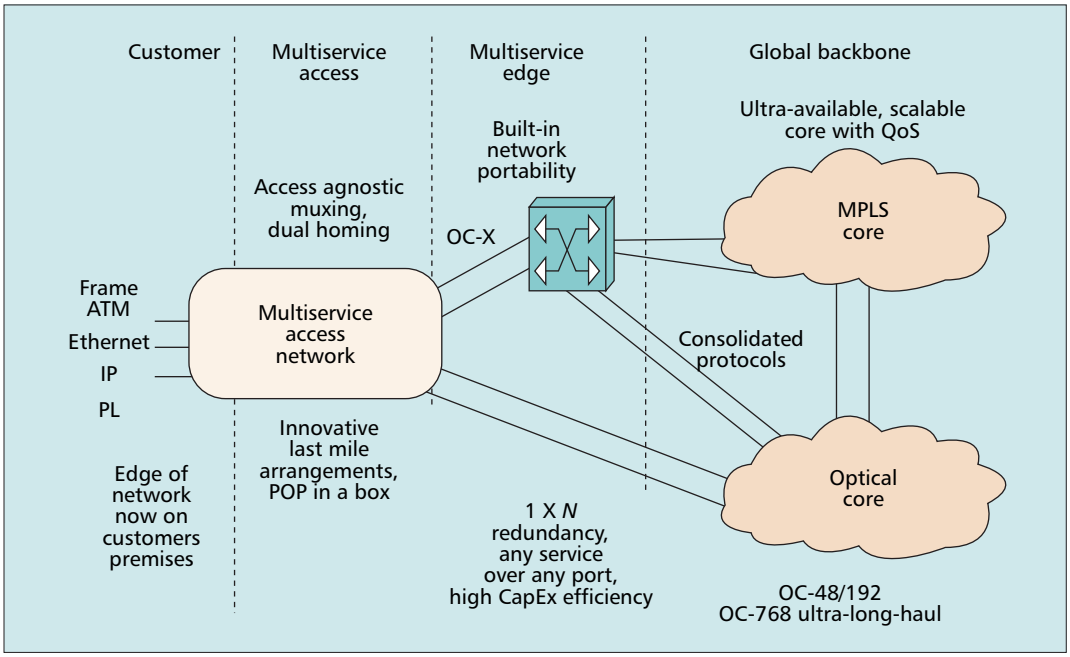

Figure 1. *MPLS Concept of One converged architecture.*

By applying these technologies, the architecture will migrate from a predictive network that monitors, correlates, and recommends action, to an adaptive network that monitors, correlates, and takes action, to a cybernated network that has integrated components that dynamically manage by business rules and policies.

envisions consolidation of systems, retiring legacy systems, scrubbing databases of record, and choking sources of database errors. This migration will greatly reduce the number of systems AT&T must manage and operate. Separate OSS network topology databases will migrate to a single network database of record with common maintenance and provisioning capabilities. Network automation envisions deploying a hands-free, intelligent, self-healing network and retiring legacy network elements. Service automation includes the automation of every human-to-computer interaction currently required for setting up and maintaining network services. This involves e-bonding between AT&T's network and the customer's own network management systems, which effectively opens the network to the customer, enabling new levels of customer network management, service creation, and ordering. This is largely accomplished through new OSS and business support system (BSS) developments, and is aimed at empowering enterprise customers with the tools to basically create their own network services as they transform their own internal networks. The goal is delivering services to customers in real time, with zero defects and cycle time, and to support both a network as well as an operational environment with six nines reliability.

The MPLS converged network is based on an open standards architecture. Standards employed include MPLS/Border Gateway Protocol (BGP) IP virtual private network (VPN) [1, 2], label distribution protocol [3], multiprotocol BGP [4], open shortest path first [5], ATM/frame relay/Ethernet MPLS encapsulation [6], and others. MPLS operations are also based on open standards, including the MPLS/BGP VPN management information base (MIB) [7], label distribution protocol MIB [8], multiprotocol BGP MIB [9], data plane connectivity verification and fault detection [10], and others.

Technologies critical to the architecture include:

- Artificial intelligence (AI) embedded in the network: will constantly analyze customers' traffic to anticipate needs and proactively suggest additional capacity or new service features designed to match customers' evolving needs.
- Self-healing proactive network: to provide automatic restoration, rerouting, or repair in milliseconds before a customer sees any impact
- Extensible markup language (XML), self-identifying intelligent network elements, expert systems, rules-based processes, and speech technology to automate the customer interface

By applying these technologies, the architecture will migrate from a predictive network that monitors, correlates, and recommends action to an adaptive network that monitors, correlates, and takes action, to a cybernated network that has integrated components that dynamically manage by business rules and policies. Of course, migration to these goals will be achieved over time, as the improved capabilities are developed and applied.

We describe MPLS-enabled services and the corresponding MPLS operations architecture, including MPLS MIBs. We give examples of MPLS operations, administration, and management (OAM) operational experience, including applications of the Concept of Zero principles and benefits, and illustrations of MPLS network monitoring, statistics gathering, and customer reports. We identify MPLS OAM evolution needs, including needed MPLS-MIB enhancements and new network capabilities. In concluding our article we challenge network researchers, developers, and key industry players to apply new technologies in fully realizing the operational vision.

When a fault condition in the MPLS VPN network occurs, traps/alarms are sent to the fault management system for monitoring. An interesting challenge is to reduce the number of alarms generated for a single event, since a network event (a trap) could indicate a logical failure

MPLS-Enabled Services

MPLS [1] is rapidly emerging as a key technology for next-generation networks, and provides a viable solution for many of the challenges posed by the growing Internet. With the converged IP/MPLS network, all services will be based on MPLS technology. MPLS fits well with current IP quality of service (QoS) frameworks, provides robust mechanisms for traffic engineering and restoration, and is a very suitable platform for providing value-added services such as VPNs. In a VPN, a set of sites communicate over a shared backbone network, but they "feel," in terms of access and security, like they are on their own private network. The VPN is defined by a set of policies, controlled by customers, that enables connectivity and QoS. Known well before MPLS, VPNs using layer 2 technologies such as asynchronous transfer mode (ATM), frame relay, X.25, and layer 3 technologies such as IPSec, GRE tunneling, and L2TP have been offered for a long time.

MPLS, however, brings a new perspective to the service and allows scalability. The ability to tunnel MPLS label switched paths (LSPs) inside other MPLS LSPs is one aspect of MPLS that improves scalability, and makes it possible for backbone routers to not be VPN-aware [2, 11]. With MPLS, the provider will provision and manage the VPN, thereby simplifying the customer's management task. In particular, the BGP/MPLS VPN solution [2], based on extensions of BGP [4], is scalable since the backbone routers are not VPN-aware, and the provider edge (PE) routers only hold the routing information of the VPNs directly connected to them. The BGP/MPLS VPN solution uses multiple VPN routing and forwarding tables (VRFs) to separate the routes and isolate the VPNs. A method for constrained distribution of routing information is used to distribute routing information for each VPN, through a BGP extended community attribute called a *route target*. A *route distinguisher* construct is combined with normal IPv4 addresses to enable private VPN addresses, which support very flexible addressing.

Three scenarios for MPLS-VPN implementation [2] are as follows:

- Enterprise VPN. Called an interregional VPN in AT&T's service definition, it is the basic type of MPLS VPN, with no exchange of MPLS labels between the VPN provider and the customer. The VPN customer is often an enterprise, but can also be a service provider.
- Carrier's carrier (CsC) VPN. An MPLS VPN carrier network provides MPLS VPN services to another carrier. MPLS labels are exchanged between the two carriers.
- Interprovider backbone: Two sites of a VPN are connected to different autonomous systems, which could be provided by different service providers. BGP may be needed to distribute the customer's VPN information.

AT&T's MPLS VPN uses all three scenarios based on service needs. Although services and network architecture are built with all scenarios, operationally the MPLS VPN services are converged as a single operations architecture, as described in the following section. We use MPLS VPN technology to illustrate MPLS OAM, keeping in mind that with the converged IP/MPLS network, all services will be based on MPLS technology and use the same MPLS OAM architecture.

MPLS Operations Architecture

The design of AT&T's MPLS operations architecture, illustrated in Fig. 2, begins with the simple guiding principles of converged network operations systems, and operations with process automation. Such simple principles lay a foundation for AT&T's overall operations architecture including network design, network management systems development, and work flow automation. These principles, together with the network automation inherent in the MPLS network elements and service automation provided by service creation technology, as discussed in the introduction, further simplify network operation for the provider.

For network management systems design, design principles and strategies used to implement converged network systems include the following:

- Modular platform design strategy (Fig. 2), whereby each module deploys a single platform to perform the intended functions. For example, there will be one fault management platform, one performance management platform, one configuration management platform, and so on.
- Shared databases of record with a common data model and a single database.
- Policy-based configuration to simplify the MPLS VPN configuration design

Examples given later further illustrate these design principles and strategies.

For day-to-day operation of the MPLS network, process design principles and directions include zero touch automation as an operations strategy. This means that ideally no human intervention is needed in the process automation to maintain the network. In reality, 100 percent zero touch may be impossible to achieve; however, this is a very powerful principle to guide operations design and implementation. For example, maintenance process automation is modeled in seven steps:

- Fault identification
- Network event analysis, including correlation
- Create tickets
- Pick tickets
- Isolate trouble, including diagnosis
- Restore and repair network, including remote restore and repair
- Test and turn up

When a fault condition in the MPLS VPN network occurs, traps/alarms are sent to the fault management system for monitoring. An interesting challenge is to reduce the number of alarms generated for a single event, since a network event (a trap) could indicate a logical failure (e.g., an LDP session or MPLS VRF down). A hard failure (e.g., a link down) often triggers multiple logical channel failures; thus, many

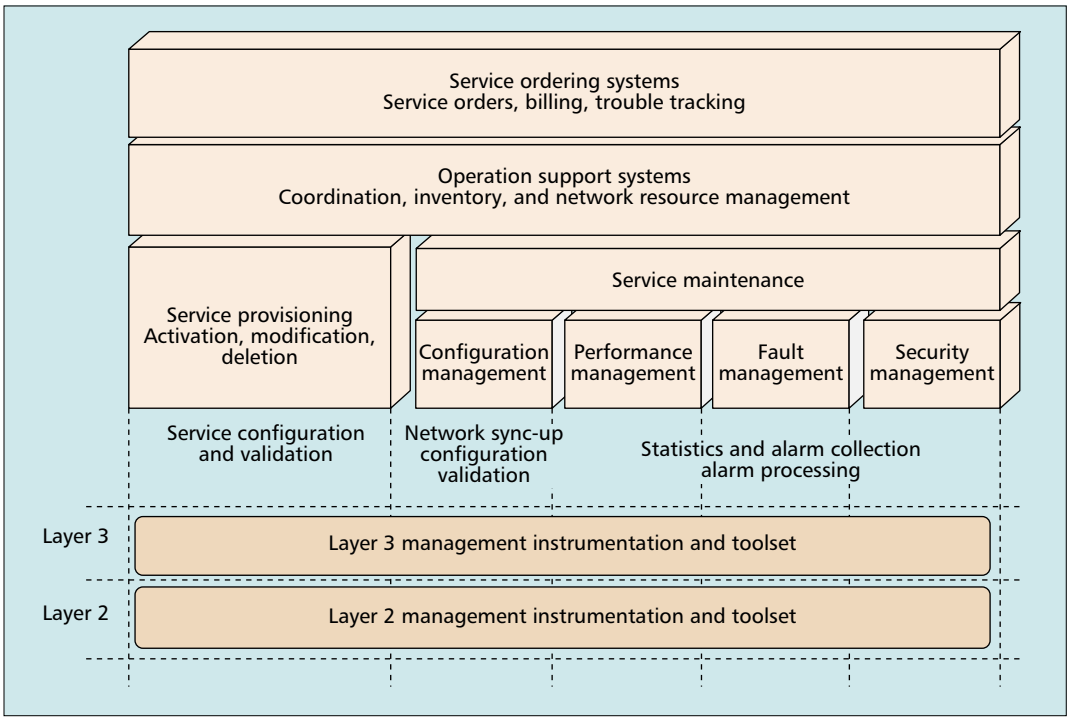

Figure 2. *MPLS operations architecture.*

A fundamental change to the existing operational connectivity model is required with the imposition of MPLS-based VPNs. There are new challenges, such as how to connect to managed assets that lie within customer VPNs, as they are invisible to the core network.

traps are generated and sent within a very short time interval. Network event correlation is needed to filter or suppress the related traps and identify the root cause of the event. A ticket can be automatically generated when the root cause of an event is found. Once a ticket is generated, auto-testing can be performed based on the ticket information to isolate the network problem. Once the root cause of a network problem is diagnosed, restoration or repair can be done automatically or in some cases manually (e.g., by changing a router card.) Once the service is restored, another test will be done before service is turned up for operation.

This strategy leads to the following network management approach:

- Use rule-based domain-specific and cross-domain correlation applied to layer 2 and 3 MPLS domains, where rules are developed based on operational experience.
- Event correlation is performed as close to the domain as possible.
- Self-healing is done via auto-corrective actions, such as using $1 \times N$ backup for automatic switchover during failure.
- Auto-testing, auto-diagnosis, and auto-repair functions can be invoked from multiple functional areas, such as using MPLS ping to auto-test VPN interfaces, detect problems, and initiate repair.

The existing operations architecture is being extended to cover the technologies enabled with MPLS, such as VPNs and traffic engineering. We now focus on these operations architecture extensions:

- Obtaining connectivity to the service provider managed assets within customer VPNs
- Modeling of virtual network topologies across the service provider network
- Fault management of additional network protocols associated with MPLS
- Performance management of MPLS-specific network extensions such as MPLS VPNs

Connectivity to Managed Assets

The cornerstone of a network management solution is providing the operations organization with connectivity to all managed assets within the network. In traditional service provider networks that provide customers with any-to-any connectivity services with few restrictions imposed, this simply consists of connecting the OSSs into the core network with the imposition of a somewhat simplistic security model, such as tweaking a few security parameters. However, a fundamental change to the existing operational connectivity model is required with the imposition of MPLS-based VPNs. There are new challenges such as how to connect to managed assets that lie within customer VPNs, as they are invisible to the core network. A separate and carefully orchestrated connectivity design is required to re-establish connectivity from the operations center into the managed assets that lie within customer VPNs, while maintaining the protective measures taken to secure the customer's VPN from unauthorized access originating outside the customer's VPN. Assets that might need to be managed and lie within the scope of the customer's VPN include the WAN links connecting the PE to the CE (PE-CE link), as well as customer premises equipment (CPE) such as the CE router.

In very large-scale networks such as those AT&T manages, connectivity solutions include the implementation of extranet VPNs to exchange appropriate routes between the element management systems (EMSs) and managed assets that lie within the customer's VPN.

With VPNs and traffic engineering, the traditional connectivity model needs to be augmented with an understanding of the virtual topology the customer network and corresponding service is constrained to operate within.

Additional objectives factored into developing such solutions include:

- Maintaining the security of networks that lie within the customer's VPN and protecting them from unauthorized access external to the VPN (e.g., via the operations center).
- Scalability of the connectivity solution to include growth in the number of managed customer VPNs as well as the number of managed assets within those customer VPNs.
- Seamless integration of new customer VPN sites into the management solution. This needs to account for the fact that multiple customers might use the same IPv4 address block to number managed assets, whereas a common EMS requires unique identification of the individual elements that constitute its management domain.
- Flexibility in accommodating multiple management domains based on the service type the customer has purchased. This would range from unmanaged service to different shades of managed service including PE-CE and CPE management, or a hybrid of some managed sites with the remainder of the customer's sites being unmanaged by the service provider.
- Automation of configuration of the PE routers to enable large-scale and accurate implementation of such a management connectivity design.

Modeling of VPN Topologies

Traditional network management solutions applied to layer 3 networks typically use a simplistic any-to-any connectivity model with few connectivity restrictions on the routing between any arbitrary points within the network. With VPNs and traffic engineering, the traditional connectivity model needs to be augmented with an understanding of the virtual topology in which the customer network and corresponding service is constrained to operate. For the network management solution to produce meaningful and actionable alerts, it is imperative that this virtual topology model be at the heart of all monitoring functionality. In this paradigm, alerts are only relevant if they can be related to impact or potential impact on one or more customer's connectivity within the context of the customer's logical network topology. Therefore, an accurate and scalable representation of the VPN topologies across the service provider network is an essential component of any MPLS-based network management solution. This model needs to be fed into every network management functional component to enhance the context of the monitoring function and align the resulting information with the service being offered across the network.

The topology model should provide information on VPNs defined across the physical network topology in terms that include the following relationships:

- The relationship between PE-CE network interfaces to customer VPNs
- Connectivity topology among customer sites that lie within the same VPN (e.g., full mesh, partial mesh)
- Connectivity topology spanning multiple VPNs (i.e., extranet)

Such information in turn can be used by the following management functional areas to determine the importance of alerting the operations center as to whether a problem is or has the potential to be service affecting. For example, a VPN-based topology model is useful in determining which VPNs and customers are affected corresponding to a PE router being unreachable from the network core.

Fault Management

Fault management is mainly focused on monitoring the up/down operational status of various elements within the network domain. For MPLS-based networks this would include new protocols associated with the MPLS implementation such as [3, 12], as well as the services delivered across the MPLS-based network, such as individual VPNs. Fault monitoring needs to cover several dimensions of the network, including:

- Element-level monitoring: The ability to monitor operational impairments associated with logical or physical elements within the network, such as the operational status of individual LDP sessions and the relevant route exchange protocols such as multiprotocol BGP. This would be in addition to traditional monitoring capabilities such as the operational status of every link and node.
- Path-level monitoring: Since customer traffic has a choice of traversing multiple paths across a network between a specific source and destination, it is imperative that the operational status of individual paths be known. This is complicated by the fact that in IP/MPLS networks, paths may be either explicitly defined or dynamically set up. Path-level monitoring poses scalability problems, since with MPLS-based VPNs or traffic engineering tunnels, path-specific probes would need to be executed within the context of the customer VPN. Such probes can be either internal to the protocol or externally generated.
- Service-level monitoring: This provides an indication of whether the service is operational or not. Such monitoring is typically done in the context of a service provided across the network, such as within the context of a customer VPN. A typical result of such monitoring would be the ability to create a reachability matrix indicating operational status between the service endpoints.

Fault Management Correlations — The above monitoring dimensions are not discrete, meaning that one dimension of fault information could impact another. For example, an LDP session being down could cause a path to become unavailable or congested, and/or could also impact a service to become unavailable. One root cause of the LDP session down could therefore cause multiple network events to be generated; one for the LDP down event, one for an unavailable path between a specific source and destination, and one for a service impact alert. To achieve end-to-end automated opera-

tions, correlation is needed to group all three related fault conditions into one network event. The correlated network event will be able to identify the root cause of the related events and its level of service impacts (e.g., percent or number of packet dropped), so a single trouble ticket can be generated with service impact information. The service impact information can be used to identify the severity of the network event as well as to assist in the priority of service restoration.

Correlation is a powerful tool to drastically reduce the number of trouble tickets generated by the above monitoring schemes. Reduction of the number of trouble tickets is necessary to improve troubleshooting time as well as time to repair (or service restoration). To assist in the engineering of correlation, a few basic rules are developed:

- Container-based rule — Group all the traps from ports on the same bay into one report.
- Link-topology-based rule — Group the traps from both the far and near ends that are physically connected into one report.
- Protocol-based rule — Group the alarms from higher protocol layers under a subtitle as secondary alarms and report the lower protocol layer alarms as primary ones. For example, group synchronous optical network (SONET) section failure traps (layer 1 trap) with link down traps (layer 2 trap) with LDP session traps (layer 3 trap).
- Cross-event rule — Group the alarms resulting from a sequence of events that relate to each other, such as when a slot fails, the function switches to the standby slot, which becomes active, and the failed one becomes standby. Other examples are as follows:
  - –Switch `sffca81ck` has been restarted.
  - –Switch `sffca81ck` CPU utilization exceeds 81 percent.
  - –Switch `sffca81ck` CPU utilization exceeds 77 percent.

Security Considerations — The MPLS-VPN architecture [2] provides two choices on how a PE router can handle the time-to-live (TTL) field of a packet when it first enters the MPLS network: to propagate the TTL value into the TTL field of the label header or not. Assuming that TTL propagation is enabled, a PE router will forward the packet with a label stack imposed to reach the destination address. This labeled packet will have the same TTL value as contained in the original IP header. For security purposes, TTL propagation can be disabled. As a consequence, traceroute will not show any MPLS network internal router hops, thus hiding network topology and router addresses from the outside world. This increases the difficulty of sending packets destined to the MPLS network routers, and thus protects against attacks on those routers, so this method is adopted. However, it also prohibits customers and network operations from using traceroute for analysis within the MPLS network, whereas traceroute is normally an operational capability within an Internet service provider's network. Therefore, further enhancements to enable both traceroute capabilities and security are needed.

PERFORMANCE MANAGEMENT

Due to the large size of the networks AT&T manages and the high standards to which such networks are bound as defined in service level agreement (SLA) contracts with AT&T's customers, it is imperative that network behavior be quantified in multiple dimensions. Such quantification covers several aspects related to monitoring the network's health and accurately representing the customer's experience across the network. Performance management in the AT&T implementation therefore includes service-level, element-level, and traffic flow monitoring.

Service-level monitoring measures the customer's experience across the network to identify potential violations of SLAs prior to the customer identifying such impairments. This function requires the deployment of monitoring systems and software across the network, and requires a good understanding of a representative sample of customer traffic traversing the network. Both passive monitoring and active probing technologies are used to quantify the customer experience in terms of specific criteria such as latency or delay, packet loss, and packet jitter, corresponding to the representative customer traffic. In large-scale MPLS-based VPN networks, such monitoring needs to be representative of the customer experience within the context of the customer VPN. Building a VPN-specific monitoring solution requires being able to address scalability in terms of the increasing size of the customer's VPN and an increase in the number of VPNs across the network.

Element-level monitoring tracks the utilization of individual active network assets with the intention of identifying utilization trends that might point to root causes associated with potential overutilization of that element. This provides a low-level perspective of how individual network assets behave in response to customer traffic traversing the network. Additional requirements include monitoring logical elements such as traffic engineering tunnels and new MPLS parameters such as the number of routes within a VPN.

Traffic flow monitoring quantifies traffic flows across key network boundaries, such as to/from the Internet, between a customer VPN and the core network, and within the network to understand if network resources are being utilized per network design. This requires monitoring instrumentation that can efficiently examine the embedded contents of MPLS labeled packets. Traditional instrumentation that operates on IPv4 forwarded packets will be augmented for MPLS awareness, especially at interfaces where MPLS switching is used to exchange packets. This also poses a challenge to the monitoring systems that will need to provide a means to attribute the traffic flows to elements such as a VPN, or a source/destination within a VPN. Judicious placement of such monitoring systems also becomes a challenge as there are limited points wherein such monitoring systems can be placed. For example, since elements within the core of the network that carry traffic aggregated across multiple VPNs are oblivious to VPNs car-

Building a VPN-specific monitoring solution requires being able to address scalability in terms of the increasing size of the customer's VPN and an increase in the number of VPNs across the network.

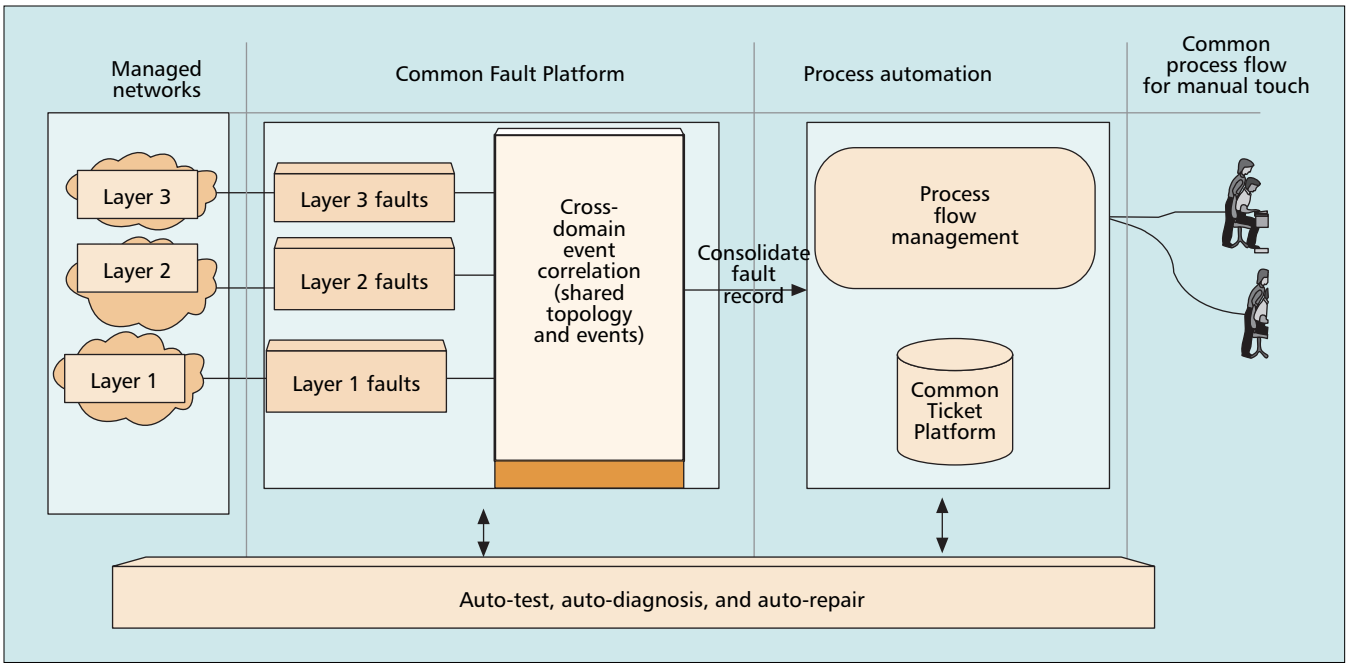

Figure 3. *Common fault platform. Converged platform to achieve zero touch.*

ried across them, such monitoring systems are constrained to be placed at the edge of the network, resulting in scalability issues related to their deployment.

MPLS MIB Architecture

To support network operations various MPLS MIB modules will be used, including the LDP MIB [8], label switching router (LSR) MIB [13], and traffic engineering (TE) MIB [14] in the core network, and the VPN MIB [7], FTN MIB [15], BGP MIB [9], and IF MIB [16] at the network edge. The LDP MIB [8] is used to manage LDP sessions between different LDP peers. The LSR MIB [13] monitors the label switching behavior of LSRs and is used to manage LSPs, particularly for troubleshooting and isolation of impairments indicated through monitoring with the LDP MIB [8]. MPLS traffic engineering tunnels are managed by the TE MIB [14]. The VPN MIB [7] contains instrumentation to manage MPLS/BGP VPNs, and has tables to model VRF table entries and the interfaces associated with these VRFs. The ingress LER performs the mapping between incoming prefixes and outgoing LSPs, and this mapping is managed by the FTN MIB [15]. The BGP MIB [9] is used to manage MP-BGP sessions. The interface (IF) MIB [16] is used for the management and monitoring of physical and logical interfaces, such as operational status, utilization, and error counters.

Example Applications of Automated Operations and MPLS OAM Operational Experience

AT&T has beneficially applied automated operations to its operational implementation. A few examples are given in this section to illustrate automated processes, OSSs, network, and services, and we summarize the operational benefits derived. We also illustrate MPLS OAM operational experience through examples of MPLS network monitoring, statistics gathering, and customer reports.

Example Applications of Automated Operations and Benefits

Applying the Concept of Zero principles led to a whole new way of thinking about using the network functionality itself to automate functions previously done manually. This has resulted in automating many provisioning functions, with features built into the network software to automatically populate routing and transport capacity information in network elements. This in turn reduces cycle time and cost for new updates to the network, since most of the provisioning work is performed by the network element and there is no need for operations planning or OSS development work. As new equipment is added to the network, the network elements automatically detect the equipment and self-provision, just as when a new printer is added to a PC network, and the PC automatically detects and applies the appropriate software to activate the printer.

For example, when an end office (EO) is added to a voice over IP (VoIP) access router, the EO common-language-location-identification (CLLI) code is used to point to a list of reachable numbers in the network that are automatically loaded from the routing database. This is an application of the principle to automatically trigger the update of routing information when a network transport entity or switching entity is added to the network. In addition, mechanisms are applied to automatically derive routing and transport capacity parameters, as follows:

- Determination of bandwidth allocation on each link for each service category, based

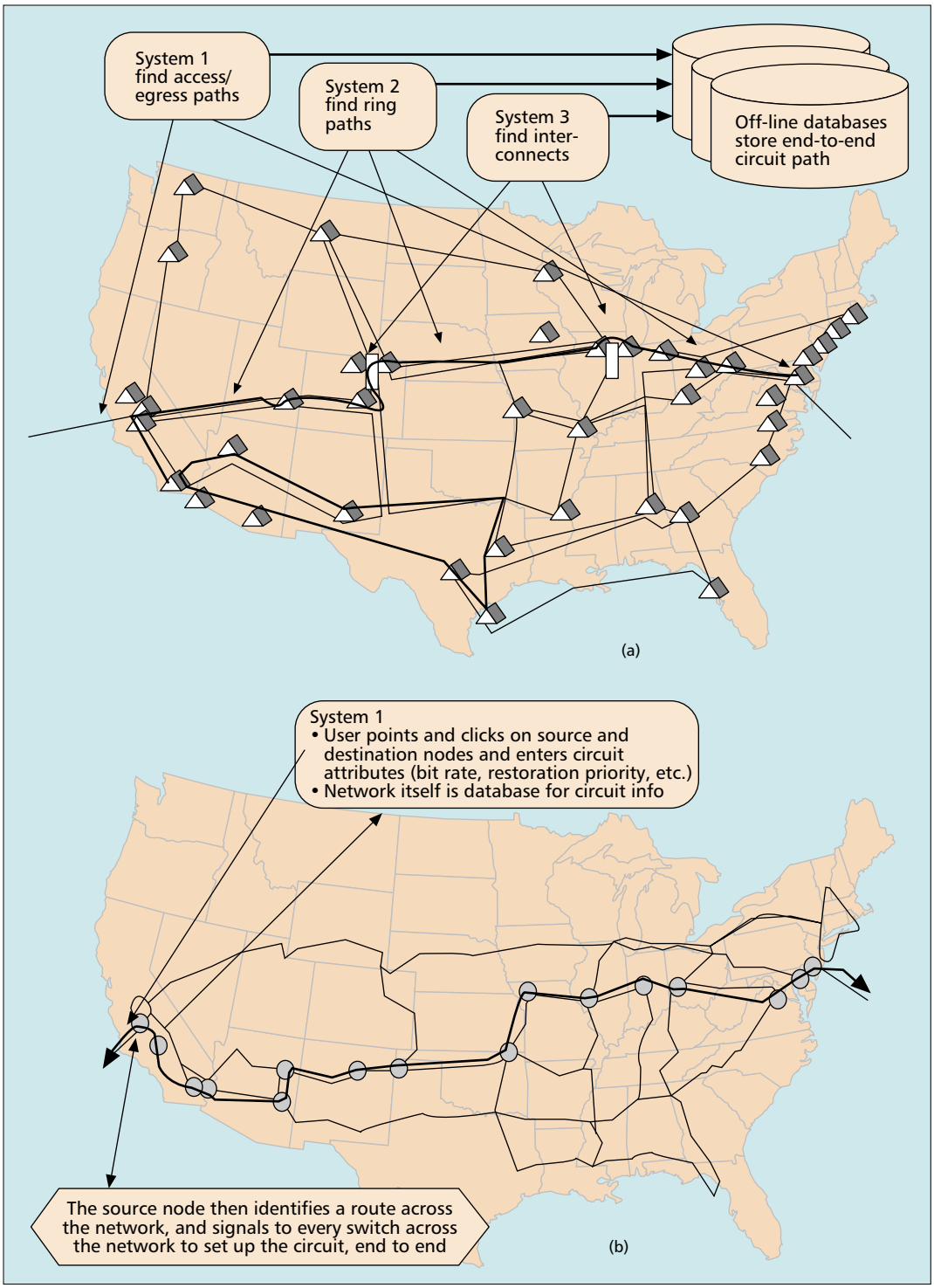

Figure 4. *a) Yesterday's operations of an optical network; b) operational migration to an automated intelligent optical network.*

on automatic detection of changes in transport capacity
- Capability selection to determine routing information for preference or avoidance of certain types of facilities (e.g., preference for fiber facilities and avoidance of satellite facilities where the additional delay is unacceptable)
- Provisioning of route lists and other routing information, such as automatic route selection mechanisms, based on learning principles
- Selection of overflow routes transiting other countries to increase call completion in times of congestion

Figure 3 illustrates the automated, converged, fault platform, called the Common Fault Platform, which implements zero-touch operations. The Common Fault Platform houses layer 1, 2, and 3 fault records, and business rules are used to perform cross-domain correlation to associate related faults into a consolidated fault record for further process automation. The process automa-

Traffic flow monitoring quantifies traffic flows across key network boundaries, such as to/from the Internet, between a customer VPN and the core network, and within the network to understand if network resources are being utilized per network design.

The Process Automation engine evaluates the fault record and performs initial process automation such as auto-diagnose, auto-test, and auto-repair. Based on the results of the initial process automation, rules are used to determine whether a ticket needs to be generated.

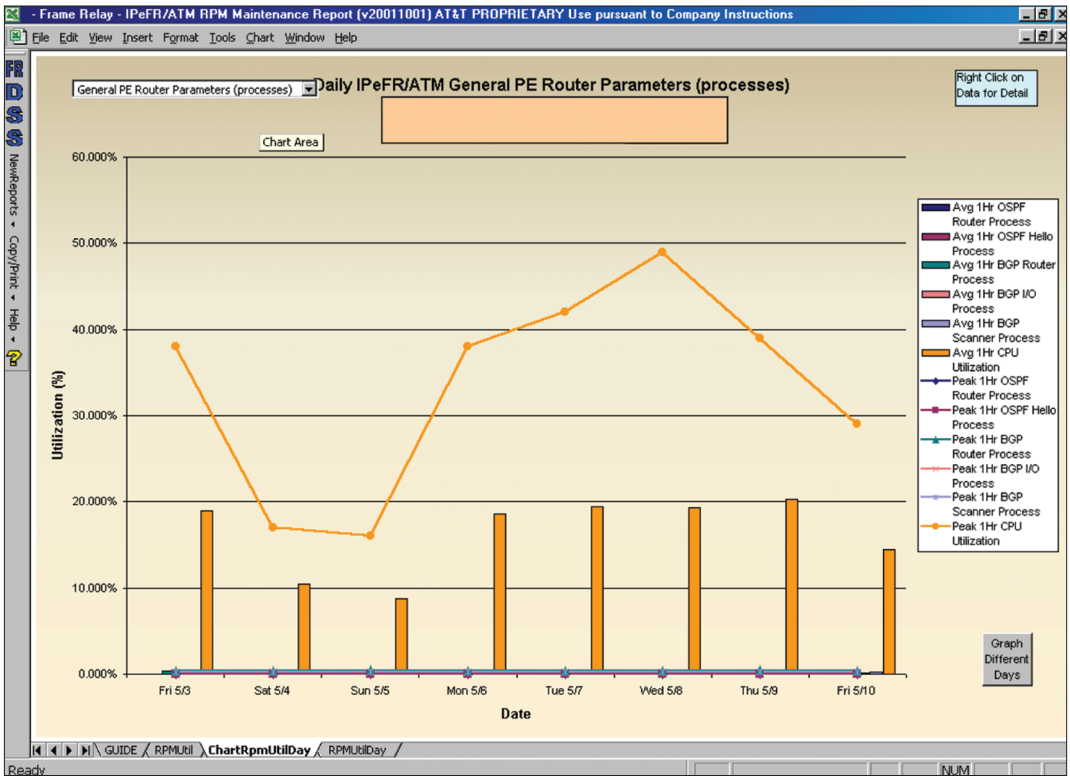

Figure 5a. *PE router resources — CPU/process utilization.*

tion engine is a centralized rule-based expert platform where field-based rules managers can build new rules as well as edit, activate, and deactivate existing rules for process flow automation. The rules are used to monitor equipment alarms and detect when field technician intervention is necessary. When process automation determines that a rule can be applied to an alarm, it processes the alarm and may request further information on the alarm by sending commands to the associated network element. The process automation engine evaluates the fault record and performs initial process automation such as auto-diagnose, auto-test, and auto-repair. Based on the results of the initial process automation, rules are used to determine whether a ticket needs to be generated.

Process automation can automatically create and refer out tickets through a work management system. If a ticket is to be created, the process flow assembles the required information and sends it to the Common Ticket Platform for ticket generation. In order to achieve automated operation, the Common Ticket Platform automatically moves the ticket through to resolution. Driven by ticket events, and as the ticket moves through its life cycle, additional process automation steps can be taken to reduce the manual work. If the ticket cannot be completed by the automated steps and needs manual intervention, a technician will manually work the ticket. Upon completion of the manual work, the ticket is put back into the automated steps. When the fault changes state, the Common Fault Platform sends an update message to the process flow automation engine to update the consolidated fault record and ticket status.

As an example, when the Common Fault Platform receives traps from network elements it does rule-based cross-domain event correlation and defines a network event. This network event is sent to the process automation engine, which then adds another layer of correlation to include process information, such as customer complaints and/or results from work center trouble identification, troubleshooting, and diagnostics. Process automation then does automatic testing, generates the trouble ticket, and automates the process steps so as to reduce or eliminate the manual work in resolving the problem and closing the trouble ticket.

Figures 4a and b illustrate the application of operations automation to the optical core network. Figure 4a illustrates yesterday's operations, which managed SONET rings connected by digital crossconnects or manual patch panels. Circuit provisioning involved finding available capacity across a series of rings and interconnects using several OSSs, where all data about paths is kept in offline databases. In contrast, Fig. 4b illustrates migration to automated operations of the intelligent optical core network. In this migration the network consists of intelligent switches, each with a complete map of all available routes to any destination, and one system initiates the end-to-end circuit provisioning, which is derived and routed automatically by the intelligent network elements.

The benefits of these applications are:

- Expense savings — saved work center personnel supporting routing/transport capacity provisioning, in that manual provisioning effort is avoided in the work centers
- OSS development cost reduction — OSS development and maintenance saved in automating the provisioning functions

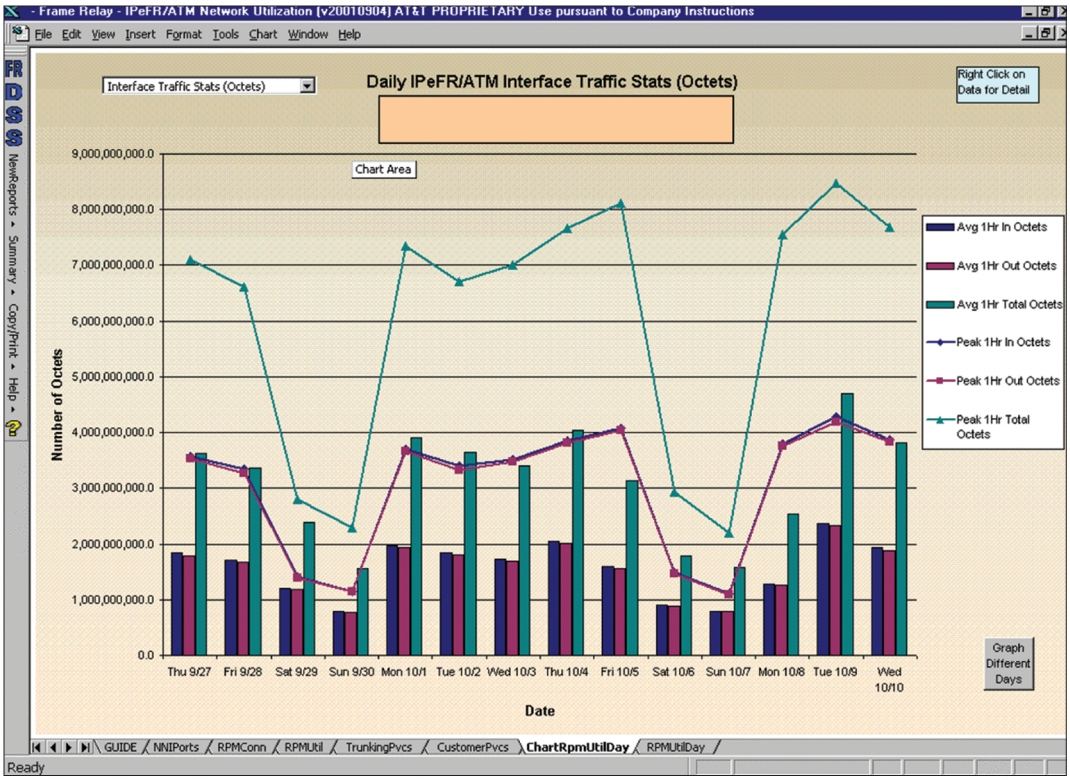

Figure 5b. *PE router traffic statistics.*

- Increased network throughput — increased network utilization, peak day and normal day, with implementation of automatic overflow routes
- Capital savings — transport capacity savings using automatic and efficient routing features

Illustrations of MPLS OAM Network Monitoring, Statistics Gathering, and Customer Reports

Examples of MPLS OAM network monitoring, statistics gathering, and customer reports are illustrated in Figs. 5a, b, and c. Figure 5a illustrates the CPU percent utilization, and Fig. 5b illustrates the daily traffic level in total octets, both during a given week for a particular PE router. Such data is readily available for any time period and any router. Figure 5c illustrates a typical customer report indicating traffic levels for each of their service classes. Such reports are readily available to customers, on demand, with a wide variety of options and flexibility.

MPLS OAM Evolution Needs

In this section we describe a few important examples of near-term needs for network management that can be achieved through extensions of MPLS MIBs. We then identify technology innovations and developments that can help achieve a fully automated operational network, and challenge network researchers and developers to apply such technologies toward meeting that goal.

Needed MIB Enhancements

AT&T's use of MPLS MIB modules is limited in some cases by shortcomings in the MIB modules. In [17] we present needs and requirements for the LDP-MIB, VPN-MIB, and BGP-MIB, based on AT&T laboratory testing of management capabilities for planned services such as MPLS VPNs. Here we highlight some of the issues.

LDP-MIB: Objects are needed to record MPLS performance usage statistics. Reference [18] proposes to preserve historical information related to LDP status and performance to ensure persistence of information when an LDP Entity goes up and down.

VPN-MIB: When the operator-defined VRF maximum route threshold is exceeded, the notification gives no information on the number of routes being dropped. Such a count is needed for capacity planning and threshold tuning purposes. Some examples of practical use are:

- Engineering the routing table size via the maximum route threshold.
- Understanding customer demand for prefix routes — increase in demand could be due to growth of the business, or rehoming. If so, re-engineering might be needed.
- Troubleshooting and understanding the customer service impact — a customer might inject too many routes due to provisioning errors.
- Router resource management to optimize router performance — the overflow counts can be summarized at the per-router level for forecasting and capacity planning.

BGP-MIB: The BGP neighbor maximum prefix limit should be used to limit the number of

In this migration, the network consists of intelligent switches, each with a complete map of all available routes to any destination, and one system initiates the end-to-end circuit provisioning, which is derived and routed automatically by the intelligent network elements.

Artificial intelligence, which is used to manage space exploration, will be a critical element of this network, wherein bandwidth is provided automatically and services are supported transparently.

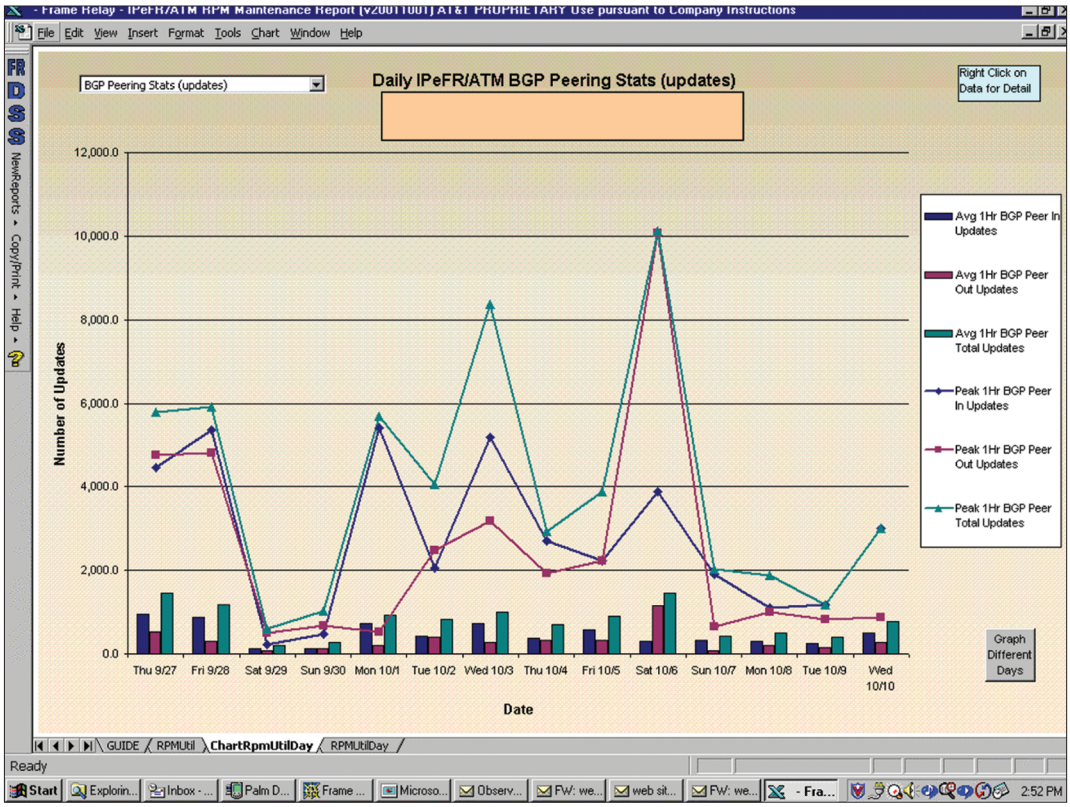

Figure 5c. *Customer report example.*

BGP routes injected in the VRF, and notification is needed to indicate when the maximum prefix threshold is exceeded. This notification can be used for capacity management, resource management, and network management.

Needed Technologies Innovation and Development to Achieve Fully Automated Network Operation

Needed technologies and standards developments to achieve a fully automated network are motivated in this section. In concept, this network would be as automatic as plugging in a fax machine is today. Once plugged in, the fax machine automatically configures itself, provisions itself, and adjusts itself to the network — it is ready to go without the user doing any of those network-related functions.

As articulated by AT&T CTO Hossein Eslambolchi [19], to achieve this goal we need a new network vision to leapfrog the incremental improvements of the past by creating a hands-free self-operating network. Artificial intelligence (AI), which is used to manage space exploration, will be a critical element of this network, wherein bandwidth is provided automatically and services are supported transparently. AI embedded in the network elements will constantly analyze customers' traffic to anticipate needs and implement capacity and/or service features as needed. A self-healing network that provides automatic restoration, rerouting, and repair before a customer sees any impact is already feasible, and a reality with MPLS fast reroute and shared mesh restoration technology. Other technologies that could be used to build this self-operating network are Web services with XML; intelligent network elements with self-identifying network processes; expert systems, rules-based processes, and trend analysis; and speech technology that automates the customer interface. By applying these technologies, the architecture will migrate from a predictive network that monitors, correlates, and recommends action to an adaptive network that monitors, correlates, and takes action to a cybernated network that has integrated components that dynamically manage by business rules and policies.

This vision needs key players in the telecom industry to create technologies that have higher performance, smaller size, lower power consumption, self-healing and redundancy, hitless reconfiguration and restoration, security protection, just-in-time engineering and installation, and the network element as the database of record. Achieving this future vision will require key players to participate and collaborate on an unprecedented and unparalleled scale. We give several examples in this article of how AT&T is already investing in and implementing this future vision, and would like to invite the collaboration of researchers, developers, and key industry players to apply these technologies toward meeting that goal.

Acknowledgments

We would like to acknowledge Hossein Eslambolchi for material from talks and interviews; Clayton Lockhart for his review, comments, and support; Diana Woo for material from architecture documents; Chris Chase for material from presentations; and the anonymous reviewers for helpful comments and suggestions.

REFERENCES

[1] E. Rosen *et al.*, "Multiprotocol Label Switching Architecture," RFC 3031.
[2] E. Rosen *et al.*, "BGP/MPLS IP VPNs," RFC 2547bis, work in progress.
[3] L. Andersson *et al.*, "LDP Specification," RFC 3036, Jan. 2001.
[4] Y. Rekhter and T. Li, "A Border Gateway Protocol 4 (BGP-4)," RFC 1771.
[5] J. Moy, OSPF Version 2, RFC 1247.
[6] L. Martini et al., "Encapsulation Methods for Transport of ATM Cells/Frame Over IP and MPLS Networks," work in progress.
[7] T. Nadeau *et al.*, "MPLS/BGP Virtual Private Network Management Information Base Using SMIv2," work in progress.
[8] I. J. Cucchiara *et al.*, "Definitions of Managed Objects for the Multiprotocol Label Switching, Label Distribution Protocol (LDP)," RFC 3815.
[9] J. Haas and S. Hares, "Definitions of Managed Objects for the Fourth Version of Border Gateway Protocol (BGP-4)," work in progress.
[10] K. Kompella *et al.*, "Detecting MPLS Data Plane Failures," work in progress.
[11] E. Rosen *et al.*, "MPLS Label Stack Encoding," RFC 3032.
[12] D. Awduche *et al.*, "RSVP-TE: Extension to RSVP for LSP Tunnels," RFC 3209.
[13] C. Srinivasan *et al.*, "Multiprotocol Label Switching (MPLS) Label Switching Router (LSR) Management Information Base," RFC 3813.
[14] C. Srinivasan *et al.*, Multiprotocol Label Switching (MPLS) Traffic Engineering Management Information Base," RFC 3812.
[15] T. Nadeau *et al.*, "Multiprotocol Label Switching (MPLS) Forwarding Equivalence Class To Next Hop Label Forwarding Entry (FEC-To-NHLFE) Management Information Base," RFC 3814.
[16] K. McCloghrie *et al.*, "The Interfaces Group MIB Using SMIv2," RFC 2863.
[17] W. Lai *et al.*, "Network Management Requirements for MPLS MIBs," work in progress.
[18] W. Lai *et al.*, "A Supplementary History Module for the MPLS LDP-MIB," work in progress.
[19] H. Eslambolchi, "Reasons for Optimism," *IEEE Commun. Mag.*, Mar. 2003.

BIOGRAPHIES

JERRY ASH [F] (gash@att.com) holds a B.S. degree from Rutgers University, and M.S. and Ph.D. degrees from the California Institute of Technology, all in electrical engineering. He is currently a technical leader of strategic standards at AT&T Labs. He joined Bell Laboratories in 1972, and was responsible for studies and analyses that led to the development of dynamic routing, which is now deployed throughout the global AT&T network. He is active in the Internet Engineering Task Force (IETF), particularly within the MPLS, GMPLS, routing, and traffic engineering working groups. He is author of over 180 articles and technical papers, a book entitled *Dynamic Routing in Telecommunications Networks* (McGraw Hill, 1998), and holds 14 patents. He is a Bell Labs Fellow, AT&T Fellow, recipient of the 1989 IEEE Alexander Graham Bell Medal, and was elected in 2001 to the New Jersey Inventors Hall of Fame.

LI-JIN W. CHUNG is technical manager of the AT&T Network Service Assurance Department at AT&T Labs, Middletown, New Jersey. She is responsible for systems engineering of performance management operations support systems including performance management support of MPLS VPN and VoIP services. She has a B.S. degree from Chung-Yung University, Taiwan, an M.S. degree from State University of New York, Stony Brook, and a Ph.D. degree from North Carolina State University, all in mathematics. She joined AT&T in 1976 and holds four U.S. patents.

KEVIN L. D'SOUZA is a technical consultant at AT&T Labs, where he is the technical lead on network management across AT&T's Global MPLS-Enabled Network. He has over 15 years of experience in network management of data networking services, having worked on several of AT&T's leading data services including IP and frame relay. He has been granted several patents in this area, and has an M.S. in computer science and a B.E. in electrical engineering.

WAI SUM LAI received his B.Sc. (First Class Honors) in electrical engineering from the University of Hong Kong), his M.S. in information and computer sciences from the University of Hawaii, and a Ph.D. in systems and computer engineering from Carleton University, Canada. He previously worked at Trans-World Electronics Ltd., Hong Kong, the ALOHA System at the University of Hawaii, and Bell-Northern Research, Ottawa, Canada. He is now at AT&T Labs, working on the traffic modeling and performance analysis of systems and services developed by AT&T. He has contributed to the standards development in ANSI T1S1.1 on packet mode services, CCITT SG 18 on ISDN, ITU-T SG 2 on packet cable telephony traffic engineering, and IETF on differentiated services-aware MPLS traffic engineering.

HARMEN VAN DER LINDE is a group manager in the Enterprise Services Division at AT&T. He has been involved in the design and development of AT&T's IP/MPLS network management architecture for many years. He is currently leading a team focused on network management architecture and operations planning for support of AT&T's layer 2 and 3 network infrastructure. He graduated from New York University, Stern School of Business, with an M.B.A. degree in economics and finance. Furthermore, he received an M.Sc. degree in electrical engineering from Twente University, The Netherlands.

YUNG-CHAO YU got an M.S. degree in computer science from the University of Southwestern Louisiana. From 1984 to 1988 he was a software developer for a network management system at Racal-Milgo, and then joined AT&T Bell Labs in 1988 as a developer for an OSI X500 directory platform. He has been on the architecture team for many of AT&T's IP network designs and services.

This vision needs key players in the telecom industry to create technologies that have higher performance, smaller size, lower power consumption, self-healing and redundancy, hitless reconfiguration and restoration, security protection, just-in-time engineering and installation, and the network element as the database of record.