AXD 301—A new generation ATM switching system

Staffan Blau and Jan Rooth

The AXD 301 is a new, general-purpose, high-performance ATM switch from Ericsson that can be used in several positions in a network. In its initial release, the AXD 301 is scaleable from 10 Gbit/s—in one subrack—up to 160 Gbit/s. The AXD 301 supports every service category defined for ATM, including ABR. An advanced buffering mechanism allows services to be mixed without compromising quality.

Designed for non-stop operation, the AXD 301 incorporates duplicate hardware and software modularity, which enables individual modules to be upgraded without disturbing traffic. The switching system, which supports both ATM Forum and ITU-T signaling, is easily managed using an embedded Web-based management system.

The authors describe the basic 10- and 20-Gbit/s systems as well as the 160-Gbit/s system. They also touch upon product structure, the control system and call handling.

The AXD 301 is a new asynchronous transfer mode (ATM) switching system from Ericsson. Combining features associated with data communication, such as compactness and high functionality, with features from telecommunications, such as robustness and scaleability, the AXD 301 is a very flexible system that can be used in several positions in a network.

Compactness and scaleability

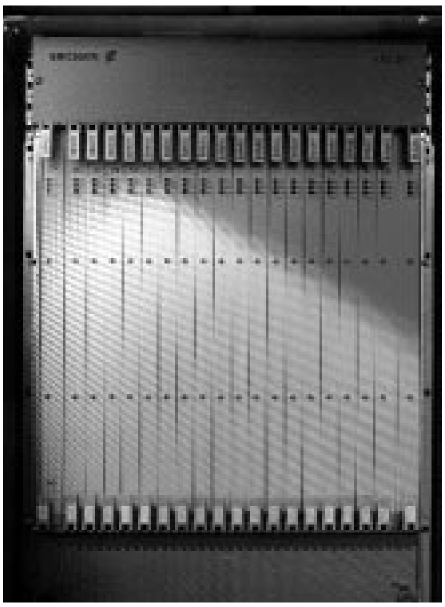
The basic AXD 301 module is contained in a single subrack and has a switching capacity of 10 Gbit/s. The subrack can be configured as a full-fledged, stand-alone switching system with complete duplication of all key components.

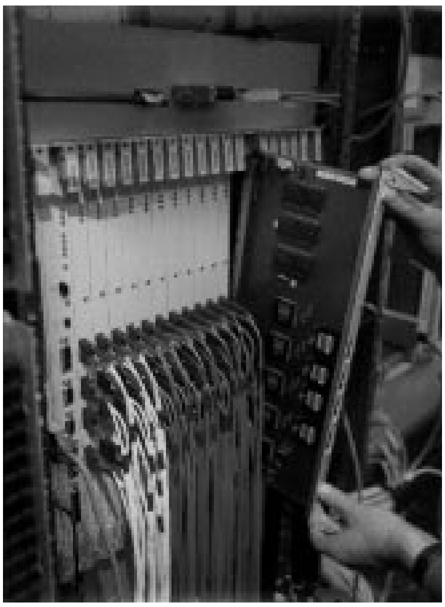
Because the single-subrack system is actually half a 20-Gbit/s system, two 10-Gbit/s subracks can be interconnected to form a non-blocking 20-Gbit/s system.

For switching capacities greater than 20 Gbit/s, a central, fully non-blocking 160-Gbit/s switch matrix is introduced. Regardless of capacity, all installed equipment can be used together with the central matrix. Capacity is extended by adding access subracks and switch boards in the central switch-matrix subrack. Thus, scaleability is achieved in a linear fashion.

Linear scaleability also applies to functions; for example, to call control and management. As more switching capacity is added, more processors can be introduced, thereby increasing processing capacity in what becomes a fully distributed control system.

Figure 1
A complete 10-Gbit/s system in one subrack.





Non-stop operation

The AXD 301 system is designed for non-stop operation. Duplicated hardware minimizes the impact of equipment failure on traffic. Furthermore, protection-switching and equipment-protection options can be applied to external interfaces. Hardware can be inserted or removed without interrupting service. The software, which is modular, has been structured to facilitate the introduction of new functionality without disturbing traffic.

The system is highly robust, with advanced functions for capturing software faults, isolating hardware faults, and protecting against overload.

Functionality

The AXD 301 supports all standardized ATM service categories, including constant bit rate (CBR), variable bit rate (VBR), unspecified bit rate (UBR) and available bit rate (ABR). The system supports ITU-T and ATM Forum signaling specifications. The signaling protocols allow operators to build a network that combines plug-and-play functionality with network control and advanced features. Signaling protocols include the user network interface (UNI 4.0), private network-network interface (PNNI 1.0) and the broadband ISUP (B-ISUP).

The internal structure of the AXD 301 facilitates extensions and upgrades, and complements the overall system with new services. In its initial release, the AXD 301 supports the following add-on modules:

- Multiprotocol label-switching module (IP switched over ATM)—this module will be available as soon as the standard is set in the Internet Engineering Task Force (IETF).
- Wide-area data-service module—this module provides support for frame relay and related services.
- Narrowband transport and interworking module—which enables AXE to function as a voice server in an ATM network.

Low-cost operation

The AXD 301, which has an embedded Web-based element-management system, is designed to accommodate local and remote management using a Web browser on an ordinary personal computer or workstation. The element manager can be used for local and remote management. Similarly, the system can be managed remotely via the simple network management protocol (SNMP). Support is provided for all relevant, standardized management information bases (MIB).

Network management can be provided using Ericsson's management system for multiservice ATM networks, which manages not only AXD 301 switches but also edge switches, concentrators and other ac-

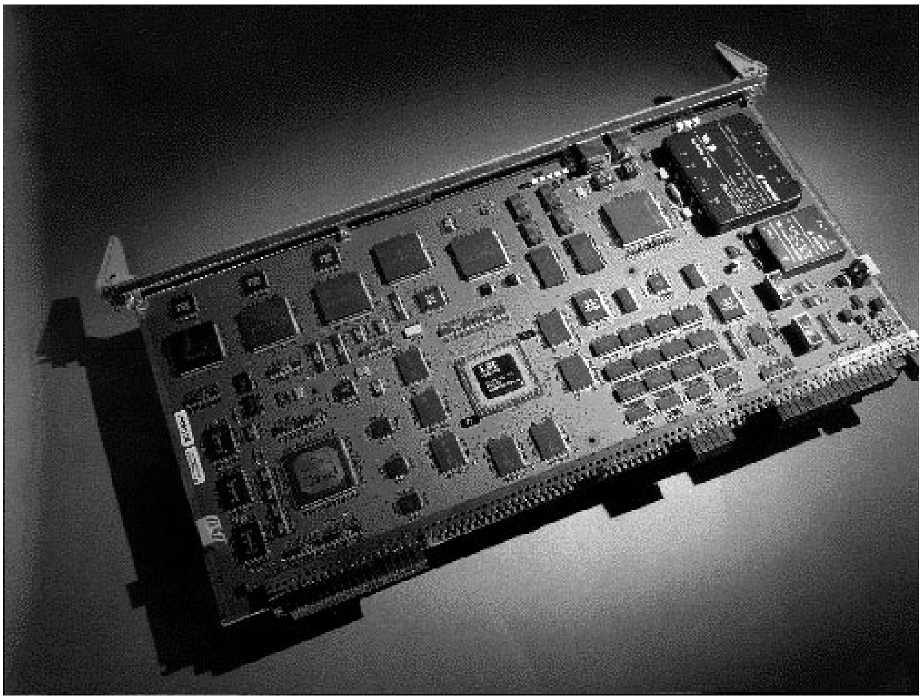


Figure 2
The ATM terminiation board handles all ATM functions.

Box A
Abbreviations

ABR	Available bit rate
ATB	ATM termination board
ATM	Asynchronous transfer mode
BICI	Broadband intercarrier interface
B-ISUP	Broadband ISDN signaling user part
CBR	Constant bit rate
CLP	Cell loss priority
E1	A physical 2-Mbit/s link
E3	A physical 34-Mbit/s link
EFCI	Explicit forward-connection indication
ET	Exchange terminal
ETSI	European Telecommunications Standards Institute
HCI	Half-call interface
HTTP	Hypertext transfer protocol
IETF	Internet Engineering Task Force
ILMI	Integrated local-management interface
IP	Internet protocol
ISDN	Integrated services digital network
ITU-T	International Telecommunication Union – Telecommunications Standardization Sector
LIFO	Last in, first out
MIB	Management information base
MPLS	Multiprotocol label switching
NEBS	Network equipment building system
NPC	Network parameter control
OC-3	Optical carrier, 155 Mbit/s link
OC-12	Optical carrier, 622 Mbit/s link
OTP	Open telecom platform
PNNI	Private network-network interface
PVC	Permanent virtual connection
QoS	Quality of service
RM	Resource management
SC	Switch core module
SNMP	Simple network management protocol
SPVC	Soft permanent virtual connection
STM	Synchronous transfer mode
STM-1	Synchronous transfer mode, 155 Mbit/s link
STM-4	Synchronous transfer mode, 622 Mbit/s link
STM-16	Synchronous transfer mode, 2400 Mbit/s link
T1	A physical 1.5 Mbit/s link
T3	A physical 45 Mbit/s link
UBR	Unspecified bit rate
UNI	User network interface
UPC	User parameter control
VBR	Variable bit rate
VC	Virtual channel

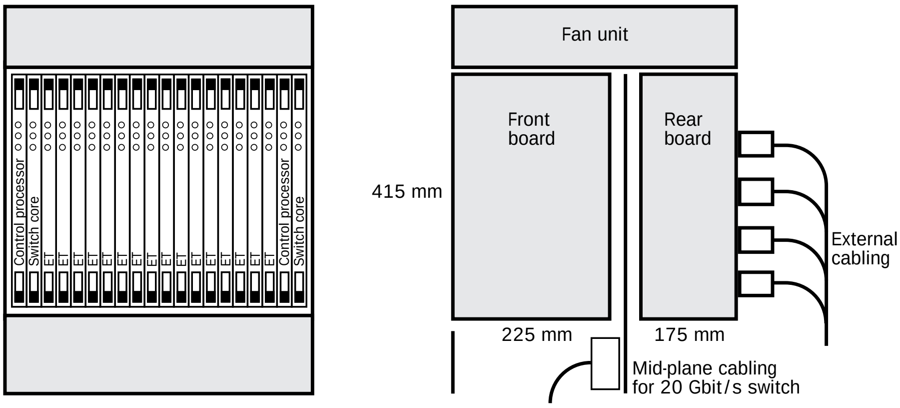
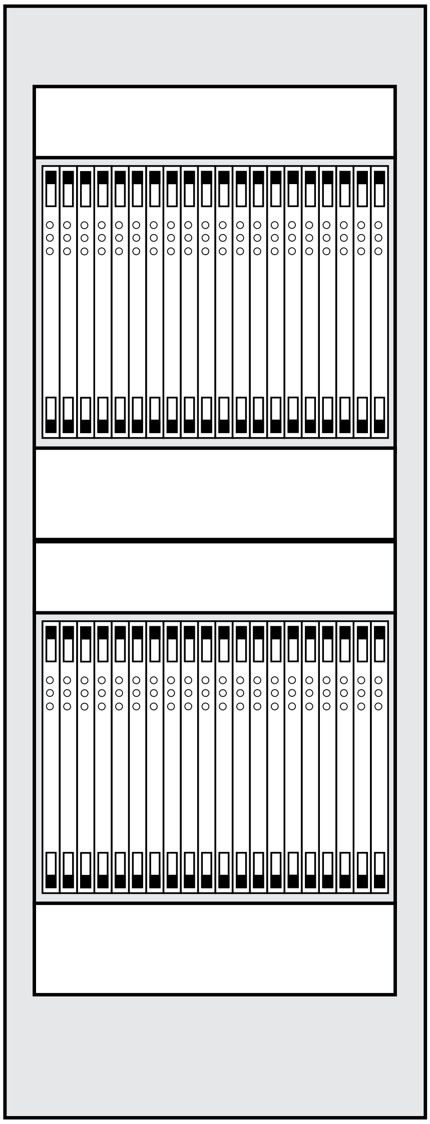


Figure 3
Physical outline of the 10-Gbit / s system.

Figure 4
Two 10-Gbit / s systems can be directly interconnected, forming a 20-Gbit / s system.



cess devices that make up the multiservice ATM network solution.

Basic 10- and 20-Gbit / s systems

The 10-Gbit/s ATM switching system is contained in a single ETSI subrack. Its physical dimensions, including the fan unit are:
Width 450 mm
Depth 400 mm
Height 700 mm
The subrack fits into any ETSI cabinet or network-equipment building-system-compliant (NEBS) equipment frame.

The 10-Gbit/s switch system (Figure 3) includes two control processors, two switch planes, 16 exchange terminals, and fans. The subrack is divided internally into a front part and a rear part, with a mid-plane separating the two. All external cabling is handled via connection boards in the rear part.

The exchange terminals (ET) consist of two boards: an ATM-termination board (ATB) at the front of the subrack, with a switch-port capacity of 622 Mbit/s; and a corresponding line-termination board at the rear. In its initial release, the AXD 301 supports the following line-termination boards:

- 16 x 2 Mbit/s (E1) with circuit emulation;
- 8 x 34 Mbit/s (E3);
- 8 x 45 Mbit/s (T3);
- 4 x 155 Mbit/s (STM-1/OC-3), electrical and optical;
- 1 x 622 Mbit/s (STM-4/OC-12), optical.

Two 10-Gbit/s subracks can be directly connected with cabling between the backplanes. This gives a fully non-blocking 20-Gbit/s switching system.

Switch fabric

In 10- and 20-Gbit/s configurations, the switch core is implemented on one board per switch plane. The corresponding rear boards are used for the clock system and network synchronization. The ATM switch fabric can be divided into five major functional blocks (Figure 5).

Transmission and line-termination functions

The line-interface boards terminate the physical layers and provide transmission-related functions. Support is provided for physical interfaces ranging from E1 up to STM-4.

ATM layer functions

ATM layer functions handle every aspect of the ATM layer, buffering, and congestion.

Buffering is handled in virtual channel (VC) queues. Each ATB card can buffer up to 64,000 cells. The early packet discard and tail-packet discard functions maximize the number of complete packets that are transmitted. Cell-loss priority (CLP) can be used for discarding cells.

ATM layer functions support every ATM service class defined by the ITU-T and ATM Forum, including available bit rate. ABR is supported with explicit-rate feedback, where the allowed cell rate is inserted into the returning resource-management (RM) cells. To ensure that the AXD 301 can be evolved to fulfill future requirements, its internal implementation supports 16 quality-of-service (QoS) classes with subclasses. Peak and sustainable cell rates are policed through monitoring, tagging or discarding cells.

Switch-core-interface functions

The switch-core-interface functions adapt cells to the internal cell format and insert routing tags. One switch-core board is active; and for redundancy, the other board operates in hot standby mode.

Switch-core functions

In terms of delay, delay variation, and cell loss, the structure of the switch core is optimized to support good ATM characteristics and smooth, cost-effective upgrades from a 10-Gbit/s switch to 20 Gbit/s, 160 Gbit/s, and beyond.

The switch core is a bufferless space switch for point-to-point connections, but uses a buffer for point-to-multipoint connections. Its functionality accommodates space switching, feedback to ingress buffers, point-to-multipoint cell-copying and expansion to larger switches. For point-to-multipoint connections, the switch core can copy cells to several destinations.

The 10-Gbit/s and 20-Gbit/s configurations employ a two-stage space-switching matrix. In the first stage, the switch-core interface distributes cells randomly into the switch core. The first and second stages switch the cells based on the routing tag included in the internal cell format. More stages are added to accommodate a larger switch-core capacity. Because the switch matrix is bufferless, delay is constant regardless of the number of stages involved.

Network synchronization and clock generation

A network synchronization function, which extracts timing from one of several synchronization references, synchronizes the timing of the node with the network or with an external clock source.

The 160-Gbit/s system

The 160-Gbit/s switch core is an upscaled version of the 10-Gbit/s or 20-Gbit/s switch—it uses the same switch chips and maintains all basic principles for transporting and buffering cells, thereby providing a non-blocking switch capacity of 160 Gbit/s. The primary difference between a 160-Gbit/s system and a 10- or 20-Gbit/s system, is that the switch core in the 160-Gbit/s system uses a four-stage switch matrix that resides in a cabinet of its own.

To upgrade 10- and 20-Gbit/s configurations to larger capacities, a 160-Gbit/s switch core is installed in a separate cabinet, and the switch-core board in existing switch subracks is connected to the central switch-core cabinet. All previously installed interface boards (exchange terminals) will work with the new switch core.

Further scaleability

The switch matrix is composed of a unique, bufferless, multistage network into which cells are sent at random—this maintains the balance of load within the network. Thus, given the same basic switching chips, a four-stage network can be scaled up to 320 Gbit/s; and a five-stage network may be scaled up to 2 Tbit/s.

Preferably, the interconnection between large nodes will make use of STM-4 or STM-16 links. In a later release, a high-speed access module will be introduced to house STM-16 and STM-4 interface boards. Eight of the access module slots can be used, accommodating any combination of these board types. Thus, the capacity of the access subrack is 20 Gbit/s.

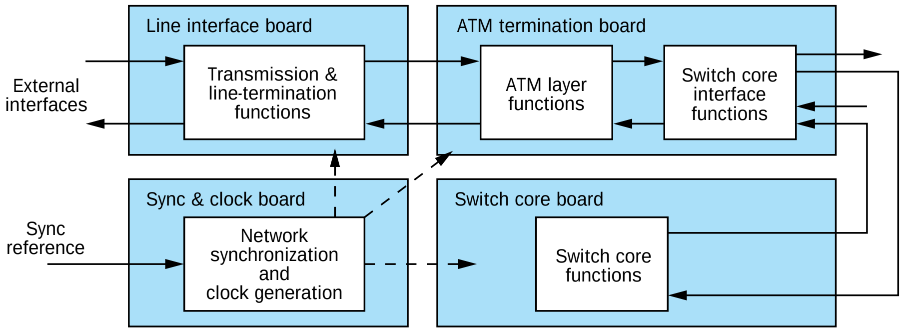


Figure 5
Functional model of the switching system.

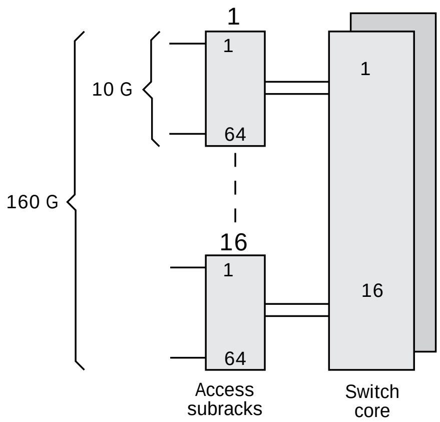


Figure 6
Architecture of the 160-Gbit/s system.

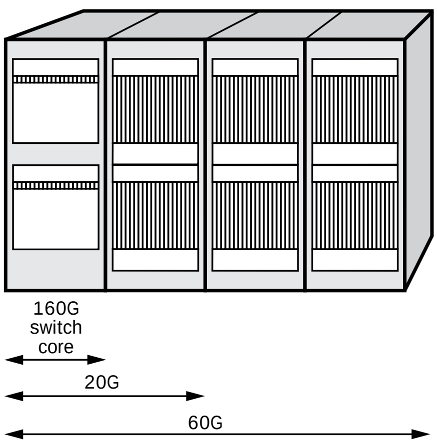


Figure 7
Configuration example—a 60-Gbit/s system that can be scaled up to 160 Gbit/s.

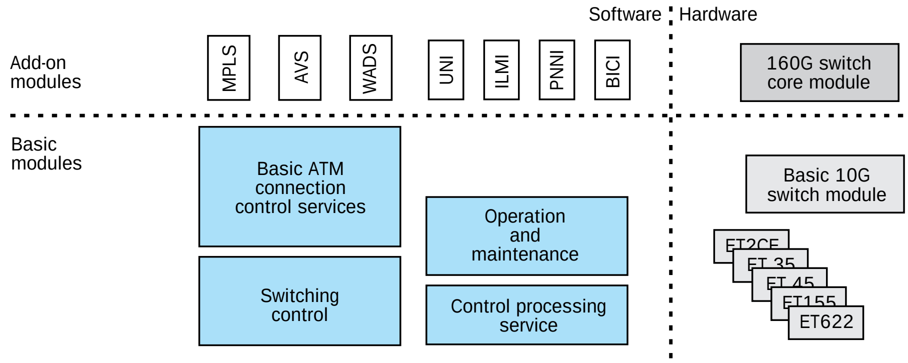
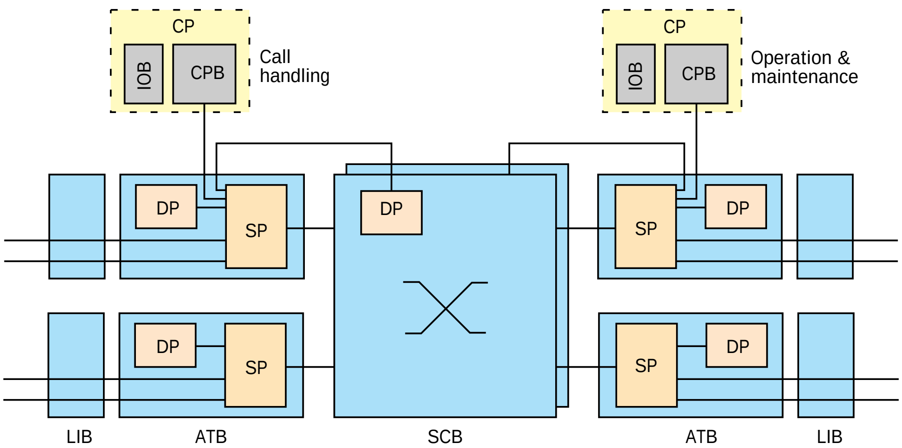


Figure 8
The AXD 301 product structure.

Figure 9
The control processors use the ATM switching fabric to communicate with the external network and device processors.

ATB	**ATM termination board**
CP	**Control processor**
CPB	**Control processor board**
DP	**Device processor**
IOB	**I/O board**
LIB	**Line interface board**
SP	**Switch port**



Product structure

The AXD 301 base system product provides functionality for a cross-connect ATM switching system. Add-on modules provide further service capabilities. The add-on software modules available in the first release provide:

- basic signaling support, including UNI signaling;
- ILMI signaling for UNI;
- PNNI signaling;
- BICI (B-ISUP) signaling;
- multiprotocol label switching (MPLS);
- support for allowing AXE 10 to control voice connections over an ATM network;
- wide-area data services (WADS).

The control system

The internal computing resources of the 10-Gbit/s switching system consist of

- two general-purpose control processors—which handle network-signaling termination, call control, and operation and maintenance;
- simple device-control processors—one on each ATM termination board and switch-core board for low-level control of the switch hardware.

For inter-processor communication and network signaling, every processor is connected to the ATM switch core. The device processors on the ATM termination boards are connected through the local switch port; the control processors and device processors on the switch core boards are connected (via the subrack backplane) to the switch port of the nearest ATM termination board.

Dual, standby processor operation

During normal operation, one control processor handles calls, while the other processor handles operation and maintenance. In addition, each processor acts as standby for its counterpart. In the event that one of the processors should fail or be taken out of operation, the system automatically switches over to single-processor mode.

Data persistency and backup

An internal, distributed, real-time database-management system copies data to each control processor, ensuring that configuration data and all data relating to operator-ordered connection setup is protected from processor failure.

The database-management system also maintains continuous backup on each processor disk. If the system must be completely restarted, the backup is used for reloading the database.

For greater security, backup data may be transferred to an external storage medium that can be used to restore system data after a catastrophic failure. Orders to generate external backup can be given manually or they can be regularly scheduled. Moreover, it is customary to generate external backup each time the system software is upgraded.

Upscaling processing capacity

When the ATM cell-switching capacity is upscaled, it may be necessary to increase call-processing capacity. For this purpose, each extra 10-Gbit/s access subrack may optionally be equipped with one or two control processors. This way, the two-processor configuration can be extended with additional call-handling processors, where each processor handles the resources of, and signaling that originates from, a group of exchange terminals.

Upgrading in-service software

System upgrades are controlled from remote management systems using IP-based communication carried over the ATM network. During the upgrade, the new software is installed on the processor disks and then loaded into memory alongside the old software; the runtime system then performs an on-the-fly state transfer from the old software to the new software.

Aside from some fundamental modules, such as the operating system nucleus, all system software can be changed during full system service with little or no disturbance to ongoing services.

Automatic recovery from software errors

The high-level characteristics of the Erlang programming language (Box B) yield high-quality software programs. However, since it is virtually impossible to ascertain that a software application is absolutely bug-free, and since bugs or corrupt data almost always manifest themselves as a runtime execution error, the Erlang language and runtime system provide several mechanisms for detecting execution errors and for triggering error handlers. These mechanisms are used extensively in AXD 301 software. Ordinarily, when an execution error occurs, a single call is lost or a single ongoing management operation is aborted. If runtime recovery occurs too many times within a given time

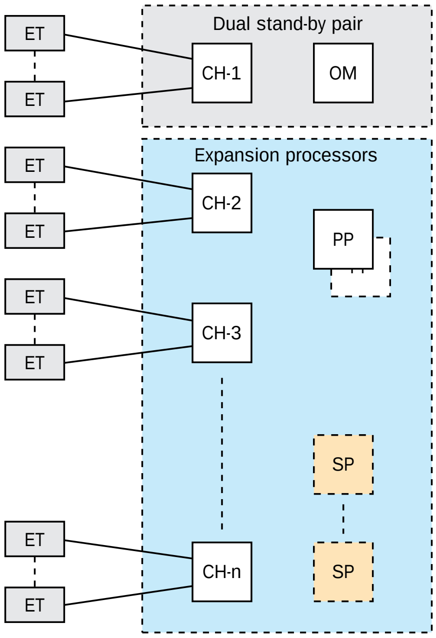


Figure 10
The basic two-processor configuration can be extended with more processors.

CH Call handling processor
ET Exchange terminal
OM Operation & maintenance processor
PP Pool processor (common stand-by)
SP Server processor

Box B
Software technology

The AXD 301 uses the open telecom platform (OTP) middleware developed by Ericsson.

The OTP[8]—which is specifically designed for use in embedded, real-time, distributed control applications—provides powerful support for distribution, concurrency, data persistency, recovery domains, Web and Java technology. It also permits the combination of an open software platform with telecommunications system characteristics traditionally achievable from proprietary software designs. The OTP runs on several standard computer platforms and operating systems.

The major part of AXD 301 software is written in Erlang, supplemented by C and Java. Erlang[9] is a high-level functional language that combines important attributes of declarative languages with constructs for supporting concurrency, distribution and error detection. Erlang gives very compact code with high quality and very high productivity. Several common causes of faults, such as corrupted pointers or memory leakage, cannot occur in Erlang since it does not allow these constructions.

frame, the recovery action is escalated, meaning that the application on the processor is restarted and reinitialized.

Call handling

The following connection types are supported in the AXD 301:

- Permanent connections set up by operator order.
  - Point-to-point virtual-path and virtual-channel connections.
  - Point-to-multipoint virtual-path and virtual-channel connections.
- On-demand, end-to-end connections set up from subscribers via signaling.
  - Point-to-point, virtual-channel connections.
  - Point-to-multipoint, virtual-channel connections.
- Soft permanent, edge-to-edge connections across a PNNI routing domain set up by operator demand.
  - Point-to-point, virtual-channel connections.
  - Point-to-multipoint, virtual-channel connections.

For on-demand control of connections, three different signaling protocols are presently supported:

- UNI and Q.2931—for users and private networks; this interface may also be used between public networks when BICI is not supported.
- PNNI—for use between nodes in the same network.
- BICI (B-ISUP)—for use between public networks, between nodes where PNNI is not supported, or between PNNI routing domains within a network.

Common resource platform

To enable friendly coexistence of many different connection-control applications, all ATM switching resources are shared through a common resource handler (see REH in Figure 11), which controls connection admission, reserves and allocates bandwidth, and sets up and releases all virtual-path or virtual-channel connections.

Also, to ensure that signaling-response delays for established calls are kept short, even when processor load is high, each new, on-demand (signaled) request for connection setup is passed to a load controller (see PLC in Figure 11). The load controller puts the requests in a last-in-first-out (LIFO) queue, from which it feeds the requests back to the call-control blocks at a rate that is inversely proportional to processor load. If, due to high processor load, requests are not fed to the call-control blocks within the maximum time allowed for delay, then they are rejected.

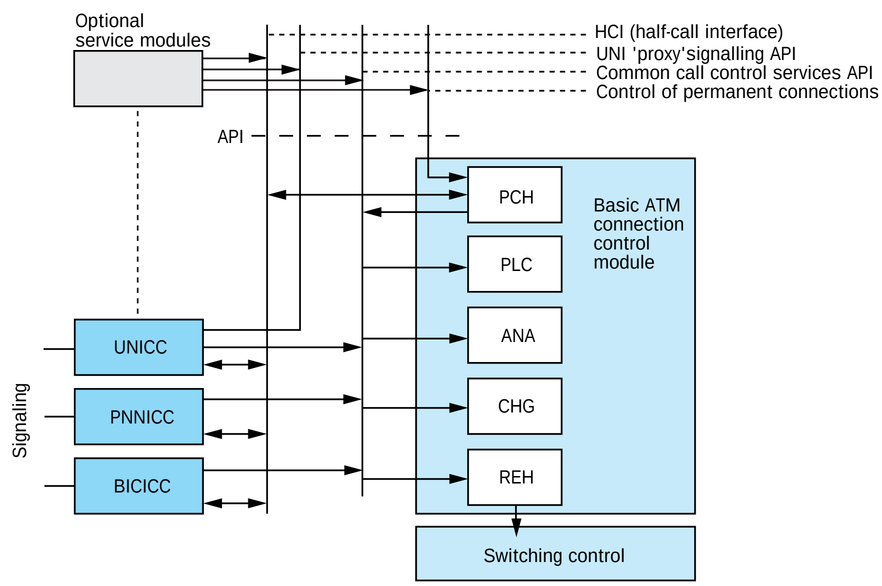


Figure 11
The call-control architecture allows many concurrent service applications in the switch.
ANA Address & routing analysis
CHG Charging data collection
PCH Management of permanent connections
PLC Processor load control
REH ATM resource handling
xxCC Half-call control for on-demand (switched) connections

Signaling protocol interworking

Signaling interworking enables different network interfaces to use different signaling protocols for controlling traffic over the interface. For example, two similar protocols, PNNI and B-ISUP, may need to exist between nodes in the network.

The AXD 301 supports protocol interworking and conversion between any combination of the signaling protocols. Signaling interworking is accomplished by controlling incoming and outgoing signaling parts separately, where a signaling protocol-independent half-call interface (HCI) is used for exchanging call-related information between the two parts (Figure 11). The information format, which minimizes format conversion in the half-call interface, requires little knowledge of the signaling type handled by the other half of the call. Thanks to the half-call interface, it will be easy to add signaling protocols in the future.

Flexible addressing and routing

Because PNNI and BICI use different schemes for routing, the AXD 301 also supports

- dynamic routing tables for PNNI-source routing;
- static routing tables for BICI hop-by-hop routing.

The routing analysis (see the ANA block in Figure 11) allows for mixed routing cases. Therefore, the same addresses can be reached independently of signaling protocols, which enables PNNI overflow traffic to be routed over a B-ISUP network, and B-ISUP overflow traffic to be routed over a PNNI network.

Charging support

Comprehensive charging data is provided on a per-connection basis for any type of connection; that is, for originating and terminating sides, and for duration- and usage-based charging. Charging data includes the number of the calling party, time of start and stop, quality of service, and number of cells sent.

To enable usage-based charging, counters tally the cells in ingress and egress directions per ATM connection. The value of each counter is collected by the CHG block (Figure 11) at the end of each call. For lengthy calls, the counter values are also collected at regular intervals.

When charging data—that is, records of call details—must be stored temporarily before it is transferred to an external system for post-processing, it is buffered in stable storage (on disk).

Ericsson can provide an external billing-gateway system that collects call-detail records from every AXD 301 switch in a network and formats the data as required by the network billing center.

Conclusion

The AXD 301 system comes with a multitude of ATM interfaces. Moreover, it supports several ATM-overlay services using optional functionality. Add to these properties an extensive base of functionality, scaleability, manageability, and upgradability and the AXD 301 becomes an outstanding choice for long-term investments in ATM network infrastructure—in the backbone core as well as at the edges of the ATM network.

References

1 Lundmark, T., Öster, G.: ATM for enhanced network capabilities. Ericsson Review 75(1998):1, pp 4-9.
2 Blau S., Rooth, J.: AXD 301—A new generation ATM switching system. Ericsson Review 75(1998):1, pp. 10-17.
3 Axell, J., Hellstrand, F.: ATM traffic management and resource optimization. Ericsson Review 75(1998):1, pp. 18-23.
4 Buhrgard, M., Westin, T., Wicklund, G.: A Scalable ATM switching solution. Ericsson Review 75(1998):1, pp. 24-31.
5 Hågård, G., Wolf, M.: Multiprotocol label switching in ATM networks. Ericsson Review 75(1998):1, pp. 32-39.
6 Höller, J.: Voice and telephony networking over ATM. Ericsson Review 75(1998):1, pp. 40-45.
7 Stockman, B., Wallers, A.: The BYB 501 equipment practice. Ericsson Review 74(1997):2, pp. 62-67.
8 Torstendahl, S.: Open telecom platform. Ericsson Review 94(1997):1, pp. 14-23.
9 Däcker, B.: Erlang—A new programming language. Ericsson Review 70(1993): 2, pp. 51-57.