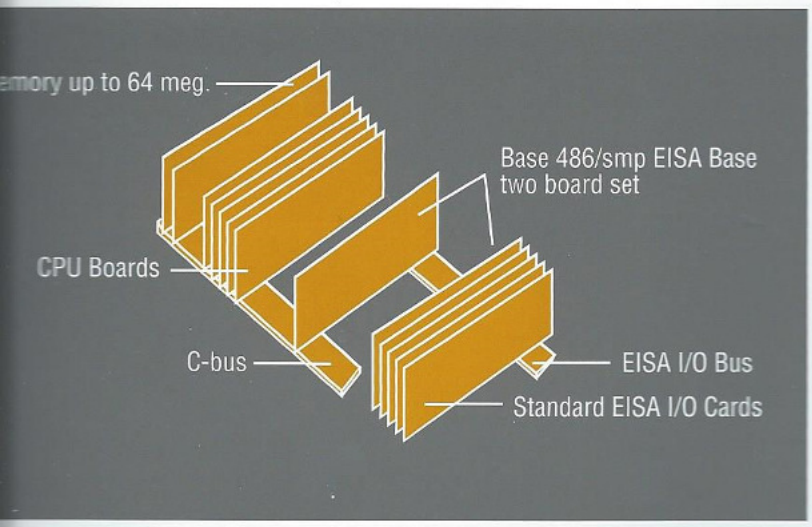
# Turning PCs into real computers.

## Corollary 486/smp symmetric multiprocessor core system for OEMs and systems integrators.




With the advent of the 386 and 486, PCs have reached a level of performance competitive with minicomputers, allowing them to act as database servers and multiuser systems.

However, these systems were limited to the speed of the latest Intel processor until Corollary brought the advantages of multiprocessor architecture to the PC.

The 486/smp supports up to ten 486 processors.

Yet it has a PC-bus that supports standard, off-the-shelf I/O cards.

To take advantage of multiprocessor architecture, Corollary developed a multiprocessor version of UNIX which is marketed by the Santa Cruz Operation. Called SCO MPX, the software not only runs on the 486/smp, but also the Compaq SystemPro. SCO MPX is 100% compatible with off-the-shelf applications and drivers.

Because of its technological innovations, Corollary is recognized as the source for multiprocessor hardware and software. Leading computer OEMs, including Digital Equipment Corporation, Advanced Logic Research, Mitac and Zenith Data Systems have adopted the Corollary/smp technology for their multiprocessor PC products.

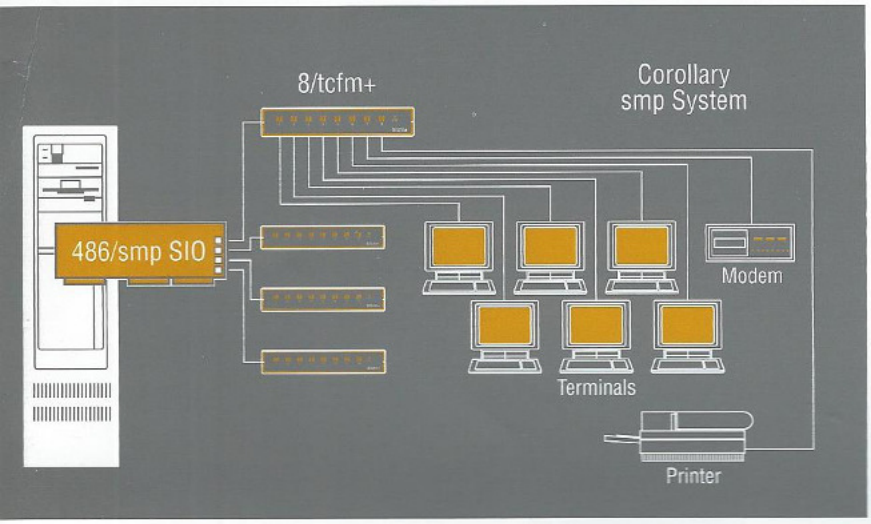
## 486/smp system basics

The 486/smp is a high performance multiprocessor system compatible with off-the-shelf PC UNIX software. It is designed primarily for multiuser applications, and provides a binary-compatible upgrade path from a single user PC to a 256+ port system. Although based on PC technology, it rivals supermini performance at significantly lower cost. Key features of the smp are:

- Up to ten tightly coupled CPUs, either 386 or 486 based, in any mixture
- Up to 64 megabytes of RAM
- Runs off-the-shelf XENIX/UNIX application software
- Plug-in standard PC I/O cards
- Dual-bus architecture
- Uses existing I/O drivers, unchanged

The 486/smp architecture is a tightly-coupled, shared memory architecture, where multiple processors have access to common memory over a high-speed multiprocessor bus.

CORROLLARY




Each 486/smp SIO can connect up to 32 serial devices via Corollary's terminal concentrators.

To prevent memory accesses from becoming a bottleneck, there is a write-back cache on each CPU board and a very high-speed bus optimized for multiprocessor CPU-to-memory operations. The system uses a dual-bus architecture, featuring a standard ISA or EISA bus for peripherals and a high-performance 32-bit bus (C-bus™) for processor/memory traffic. This results in a system which preserves compatibility with AT and EISA peripherals yet achieves a performance level only possible with a bus designed specifically for multiprocessing.

## Modular system organization

The main hardware components of the system are the base CPU, C-bus, additional processor boards, ECC memory boards, and the EISA bus I/O subsystem.

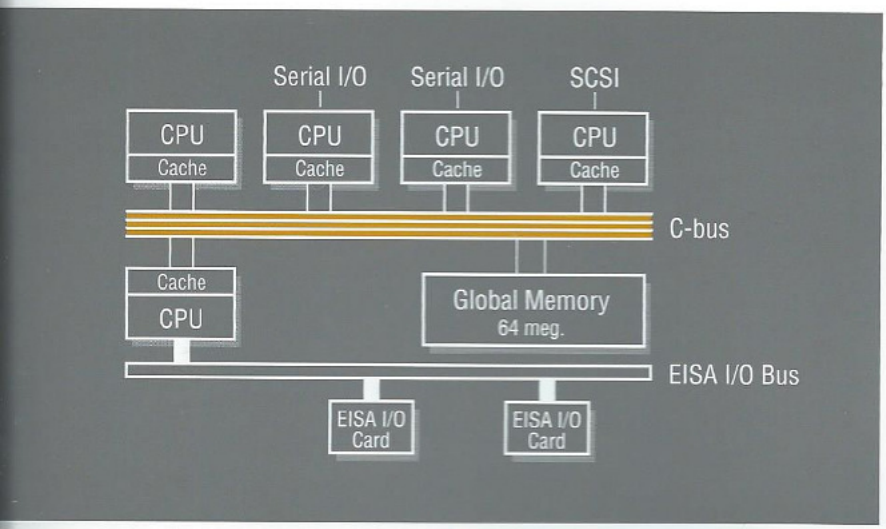
The base 486/smp CPU acts as a bridge between the C-bus and the PC bus. Its access to the C-bus is direct and identical to the manner in which all other smp CPUs access the C-bus. In addition to its C-bus connection, the base CPU can directly control the EISA bus. Since the base CPU is an EISA compatible computer, existing driver code can be executed on that CPU and operate as it would on a standard PC with no changes.

## Extensive peripheral support

Since the smp system can execute existing driver code unmodified, hundreds of low cost interface cards are immediately available. However, Corollary also provides direct C-bus connections for some classes of I/O, for example: Serial I/O and SCSI. Combining these I/O interface functions with a general purpose CPU board relieves the EISA bus, and spreads the low-level I/O system software burden over several processors.

## Core OEM technology

This is a core multiprocessor system, not a complete computer. The smp components include processors, memory, backplane, and support I/O such as serial I/O and SCSI I/O. OEMs add value and differentiation by packaging, peripheral support, and software.




Block diagram illustrates the Corollary dual bus architecture.

The core system approach allows unusual flexibility in configurations.

For OEMs, the core computer concept reduces development time and risk, but allows for significant differentiation and added value.

While the smp architecture is certainly not a "personal computer", the base processor is, in fact, a top speed PC in its own right.

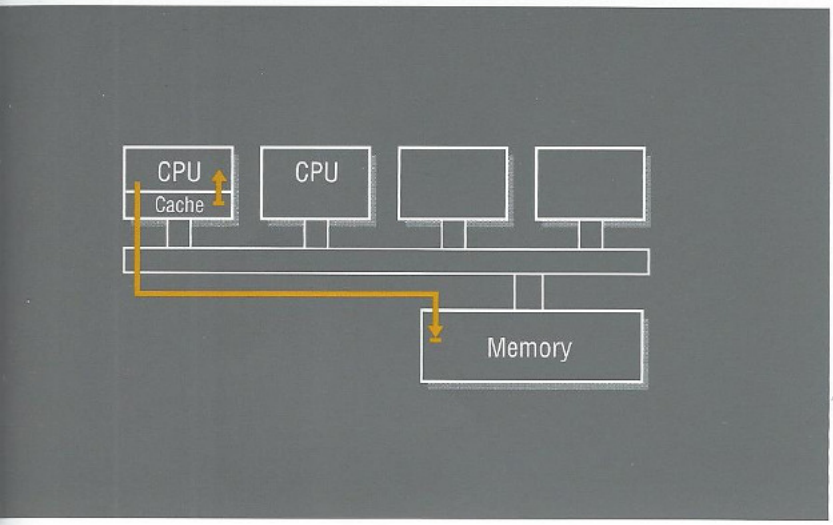
## CPU architecture

The high memory bandwidth required to satisfy multiple 486 processors is provided by a local cache for each processor. These caches provide each CPU with zero-wait-state access to data located in the cache for both read and write operations. The cache is a write-back type which transfers written cache data to the main memory only when needed. By using a write-back cache, the memory bandwidth required for each processor is greatly reduced, allowing up to ten processors to be connected to the shared memory with minimal bus contention.

The cache is organized into lines of four 32-bit words. This means that on a cache miss four 32-bit words are brought in from main memory at a time, producing a higher hit rate and reducing the amount of bus traffic by using the bus more efficiently. The "line" size and overall capacity of the cache are designed to yield a 96%+ hit rate. The 486/smp boards have a 256k byte cache while the 386/smp boards have a 64k byte cache.

## ECC memory

Global memory is physically separate from the processor boards, on dedicated memory cards. Each memory board contains Error Correction Code (ECC) logic. The ECC hardware uses redundant memory cells to detect and correct all single bit errors and detect and report all double bit errors. Up to 64 megabytes of memory can be installed in a system.




WRITE-THROUGH CACHE
With this type cache, all writes go over ths system bus.

## Write-back cache coherency

In a multiprocessor architecture, a critical issue is keeping all the caches in sync, called "the cache coherency problem". Cache coherency provides each processor with an identical view of the memory contents. A data value can be cached in more than one cache. If the duplicate data in one cache is modified, then the data in all other caches is automatically invalidated.

Cache coherency is straightforward with a write-through cache. If a value changes, the new value is immediately written to main memory; therefore main memory always has the correct value. During the write process the other processors can detect that a change has occurred by watching the common bus.

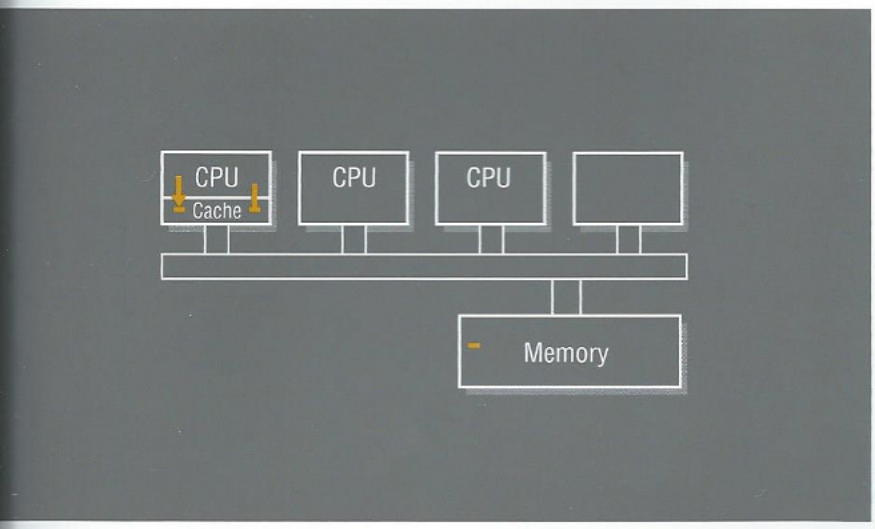
With a write-back cache, cache coherency becomes more complicated. Main memory no longer always contains the up-to-date values, and cached values change with no activity visible on the common bus.

## Cache operation details

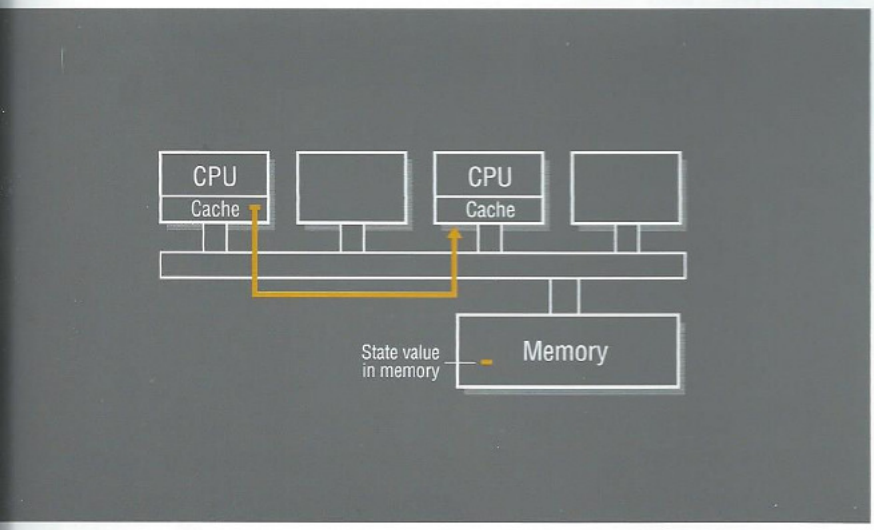
Since the type of cache used in this system is unusual outside of super minicomputers, we will go into some detail regarding its operation.

Cache coherency schemes enable multiple caches to communicate between themselves to ensure that they always have the correct values. In the 486/smp, each line in a cache has a tag, containing the high order address for the data location and a set of access bits. Four access states are supported: "modified", "exclusive", "shared", and "invalid" (MESI protocol).

When a processor requests data not in its cache, the data is read from memory. If no other processor has a copy of the data, the data is marked "exclusive". If that processor then writes over the "exclusive" data, it writes only to its cache without copying back to memory, but marks the data tag as "modified". The processor then can write the data any number of times. Hence most writes occur without accessing the system bus.




WRITE-BACK CACHE
As long as needed data is in cache, no reads or writes over system bus.




WRITE-BACK CACHE
Data needed by one cache can be supplied by another cache.

Once a processor modifies data, the original data in main memory becomes stale. When another processor requests that location, the hardware of the cache owning the modified line recognizes that the correct data is not in main memory but in its cache. The cache with the correct data prevents the memory from responding and puts the valid data on the bus itself instead.

If a processor tries to write shared data, it must first get "exclusive" use of that data. In that operation, any other caches holding the shared data are told to mark that cache entry "invalid". The requesting processor now has the data marked as "exclusive" and can write over it. Once it does actually change the data, it marks its copy as "modified".

## Why a new bus?

The design of a 386/486 shared memory multiprocessor architecture requires a high performance, 32-bit bus which supports cache coherency in order to support more than two or three processors. However, since this cache management technology is fairly new, it is not available with industry standard buses. For example it is not supported in the Microchannel or even more "industrial" buses like VME, EISA, Multibus II, or NuBus.

Write-back caches are critical to eliminate bus saturation in a shared memory multiprocessor system. Alternatives such as more expensive 64-bit buses do not provide enough usable bandwidth. The features to implement cache coherency must be supported in the bus. Given the need for a special bus, Corollary created a bus uniquely optimized for multiprocessors.

Some key features are:

- Up to 64 megabytes of shared memory
- 64 megabyte per-second-data transfer rate
- Uses block transfers exclusively
- Multiplexed address and data lines
- PC AT form factor for low-cost packaging

The C-bus is a high-bandwidth system bus that is optimized for block transfers to/from the CPU board caches. The C-bus also

implements the protocols required to maintain cache coherency. The C-bus is loosely patterned after the NuBus with the addition of critical features for cache support.

The C-bus clocks at 16 MHz. and supports transfers of 64 megabytes per second. This fast clock rate and proprietary bus transaction protocol (optimized for cache line replacement), give the C-bus much higher performance than even industry standard buses like VME, NuBus and Multibus II.

## Bus Signal Summary

| | |
|---|---|
| CLK | Synchronizes bus arbitration and data transfers. Nominally 16 MHz. with a duty cycle of 25%. |
| RESET- | Used to return all modules to their initial power-up state. |
| AD<31...0> | These lines are multiplexed to carry address information at the beginning of a transaction, and 32 bits of data later in the transaction. |
| TM<2...0> | At the beginning of a transaction, these three lines indicate the type of transaction being initiated. At the end of a transaction they are used to indicate the status of the requested data. |
| STRT- | asserted at the beginning of a transaction to indicate a valid address is on the bus. |
| ACK- | asserted at the end of a transaction to indicate the last data transfer. |
| ID<3...0>- | These lines specify the geographical location of the module. |
| ARB<3...0>- | Contending modules compare these lines with the binary value of their own ID<3...0>- lines, and drive the ARB<3...0>- lines according to the rules of the distributed arbitration logic. |
| FAIR- | Line to prevent modules that have acquired the bus from re-arbitrating for the bus until this signal is unasserted. |

## Excellent diagnostics for field test and production

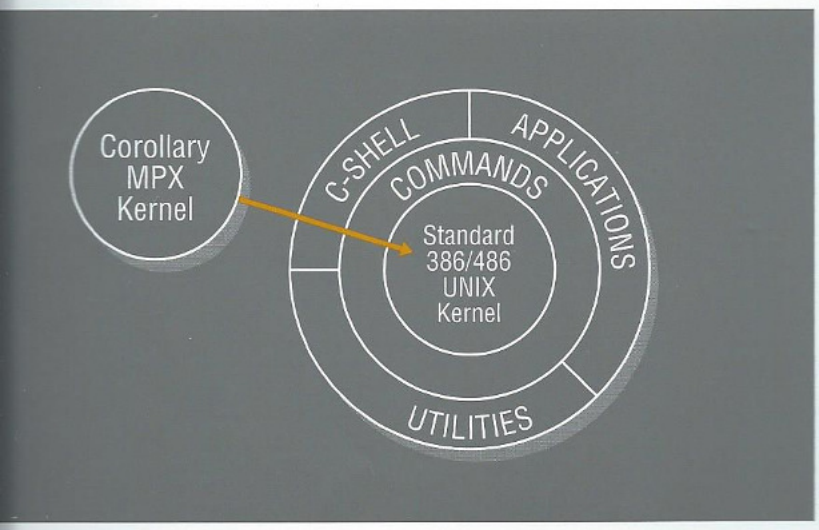
Corollary's ROM Resident Diagnostic (RRD) is an interactive multiprocessor diagnostic program. It is used during production to assure that the product is operational and can be used in the field to test the smp system components. RRD uses the COM1 serial port, therefore it can run locally from a terminal or remotely over a modem.

RRD has 65 multiprocessor diagnostic tests including memory tests, CPU tests, SCSI tests, Serial Communication Controller, DMA and floating point.

For engineering use, there is also a full set of interactive commands that give the user access to system resources such as memory, buses and CPUs. Some of the commands include peek and poke, memory display, memory search, memory test, CPU commands and ECC commands. As a sophisticated language a user can generate definable arguments, create iterative loops and gain assistance through a help facility.

## Runs off-the-shelf applications

Unix systems are inherently multi-user and multi-tasking. The Corollary smp Kernel allows several processors to each independently execute UNIX processes. This approach to multiprocessor systems means that no changes to applications software (not even recompiling or relinking) are required to take advantage of multiple CPUs. Existing "shrink wrapped" 286 and 386 XENIX applications and 386 UNIX applications just work.




Only replace the kernel. All programs, utilities, and documentation are from the original packaged product UNIX.

## Multiprocessor kernel

Corollary's kernel, called SCO MPX, is a binary compatible extension of the SCO UNIX System V Release 3.2 kernel that supports multiple processors and was developed by Corollary's staff starting with SCO UNIX 386 source code. SCO's product was selected as the initial base operating system because of its market success and an ongoing strategic relationship between Corollary and The Santa Cruz Operation. SCO MPX will operate with both Corollary smp hardware and other multiprocessors like the Compaq SystemPro.

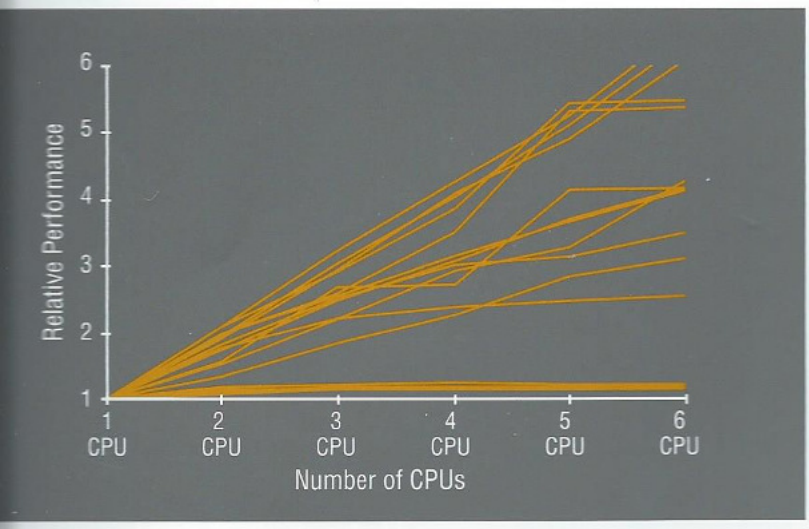
There are several approaches to developing a multiprocessor kernel. The "symmetric" approach used in the MPX kernel was chosen for its performance advantages over the less sophisticated "master/slave" architecture, and its compatibility advantages over "message-passing" schemes.

## Load balancing

All processors execute a single copy of the kernel and schedule themselves from a common run list. Load balancing across the CPUs is automatic. In effect it is self balancing; as processors finish tasks, they immediately search a common run list and schedule themselves to execute the process with the highest priority. All CPUs select processes from the run list according to the conventional UNIX priority scheme. Although the concept is simple, the implementation must deal with the appropriate locking of the run list and other data structures. In general, hundreds of separate locks are used to protect against race conditions and serialize access to common data structures.

The smp kernel can force a process or system call to execute on a particular processor. This allows the system to use third-party I/O drivers unmodified by assigning them to the base CPU.

All other programs and utilities are unchanged, and will automatically be scheduled across all the smp processors.




Each line represents one of the 18 Neal Nelson Business Benchmarks Tests (60 users). In general, as CPUs are added performance increases. The flat lines are disk access time tests, which do not improve as CPUs are added.

## About Corollary

Corollary is a pioneer in the field of multiprocessing and I/O subsystems. We developed the world's first 486-based multiprocessor and our multiprocessing technology is used by leading computer manufacturers including Digital Equipment Corporation, Zenith Data Systems, Advanced Logic Research, Mitac International and The Santa Cruz Operation.

Corollary's 8x2 and 8x4 serial I/O subsystems, which allow up to 128 users to be connected to a UNIX PC, are also used by a wide range of leading computer companies and computer resellers.

Because we're hardware and software innovators, you can count on Corollary for complete solutions.

If you would like more information about Corollary products, please call or write.

COROLLARY

**The Source for Multiprocessing**

Corollary, Inc.
800/338-4020
In California 714/250-4040
Fax 714/250-4043
P.O. Box 18977
Irvine, CA 92713-8977