

---

# **ECE 697J – Advanced Topics in Computer Networks**

ACE Programming Model and SDK

11/13/03

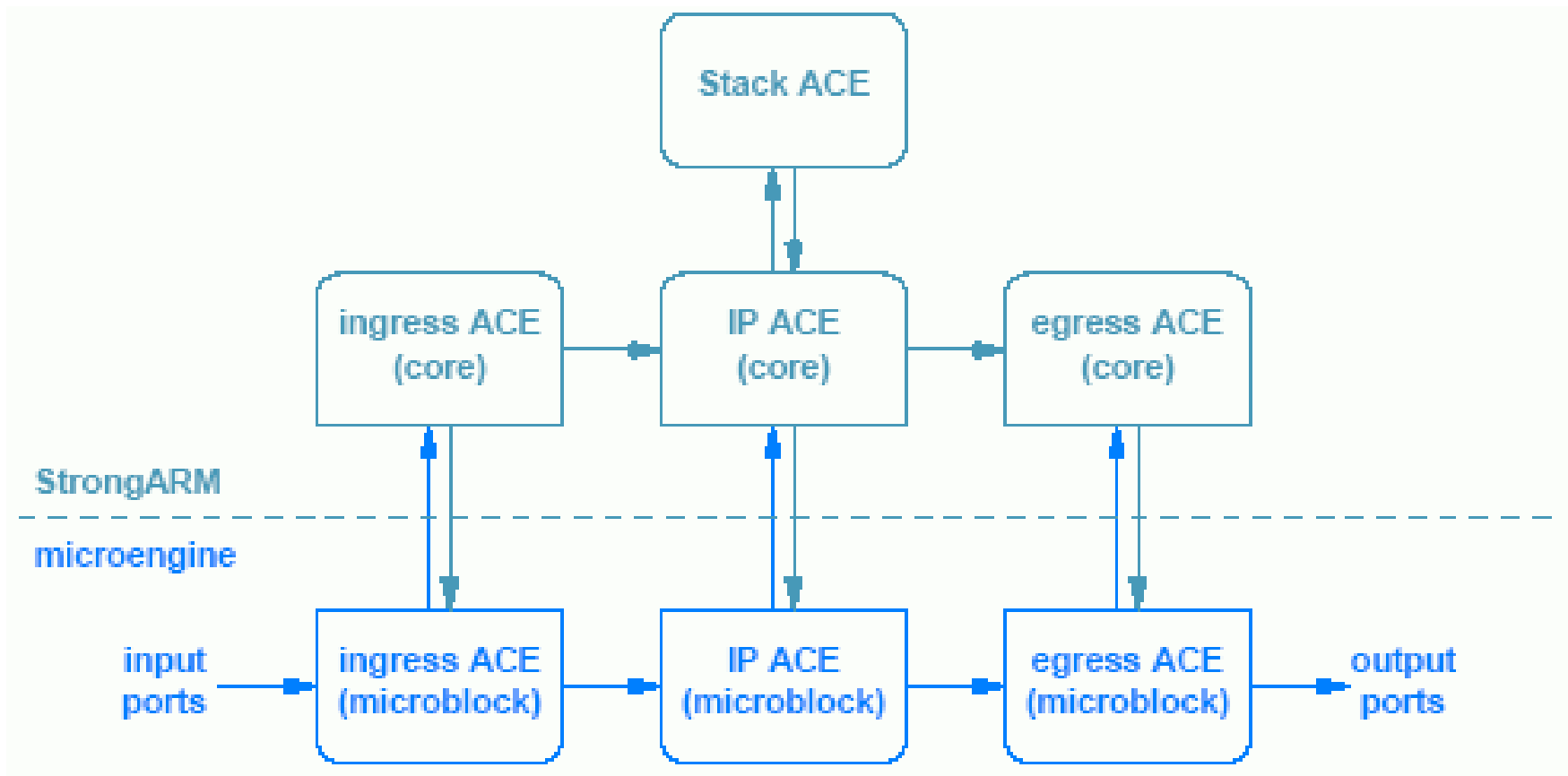
# Overview

---

- Programming Model
  - Active Computing Element (ACE) Abstraction
  - Allocation of ACEs to microengines
  - Packet Queues
- Software Development Kit
  - Simulator
  - Example: IP forwarding
- Lab 2: IP forwarding and classification on IXP1200

# Last Class

- Active Computing Element (ACE) abstraction:

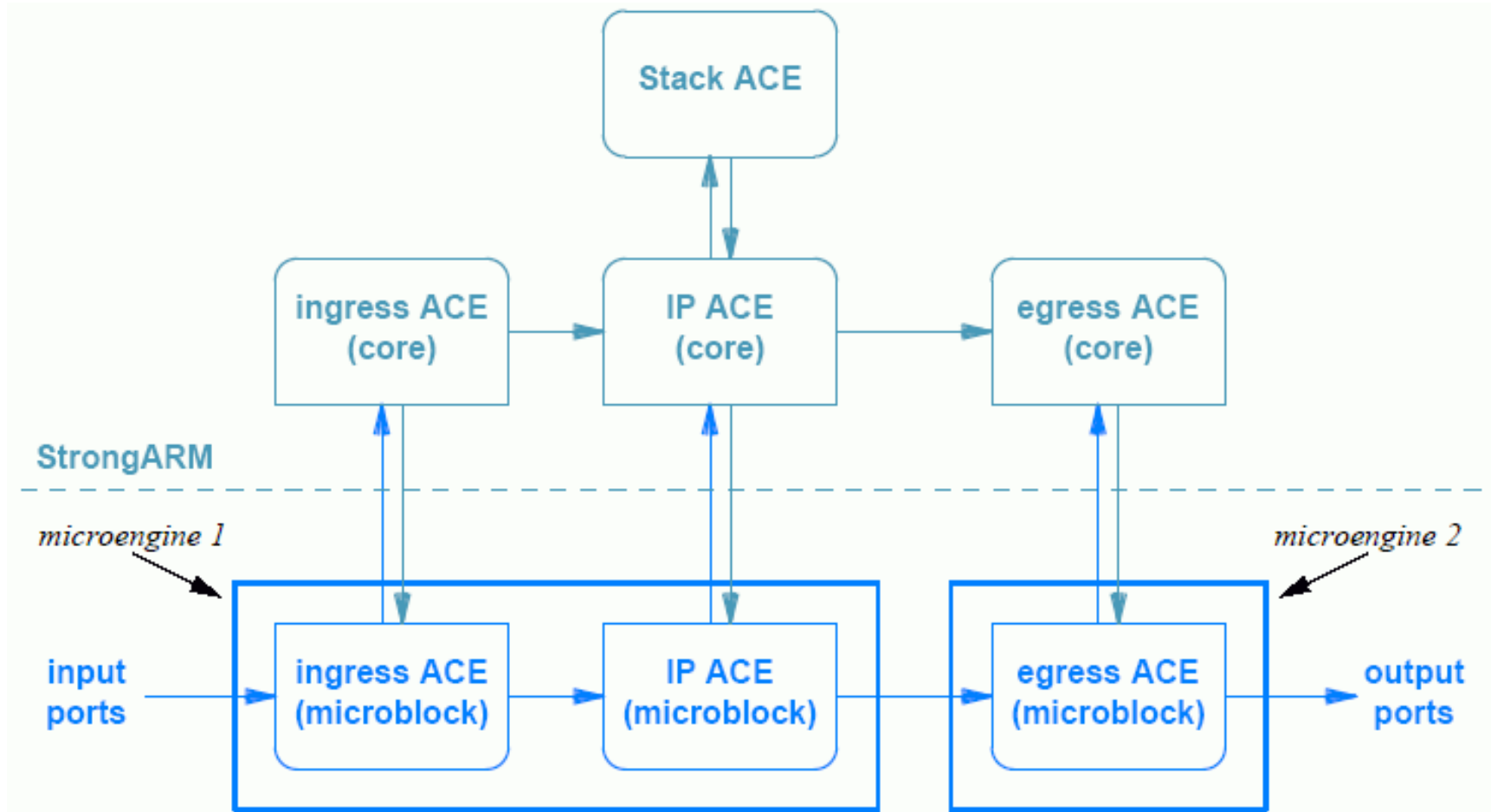


# Microengine Assignment

---

- Packet processing involves several microblocks
- How should microblocks be allocated to microengines?
  - One microblock per microengine
  - Multiple microblocks per microengine (in pipeline)
  - Multiple pipelines on multiple microengines
- What are pros and cons?
  - Passing packets between microengines incurs overhead
  - Pipelining causes inefficiencies if blocks are not equal in size
  - Multiple blocks per microengine causes contention and requires more instruction storage
- Intel terminology: “microblock group”
  - Set of microblock running on one microengine

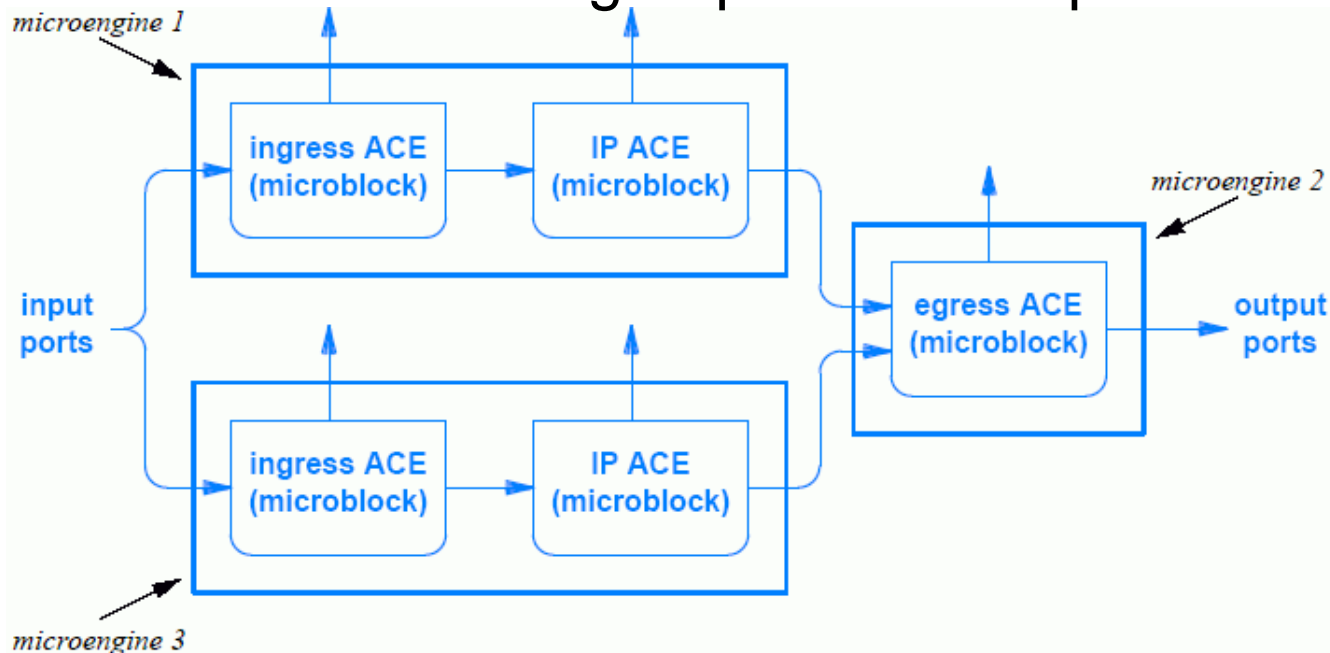
# Microblock Groups



- Microblock groups can be replicated to increase parallelism

# Microblock Group Replication

- Performance critical groups can be replicated:



- Additional complexity:
  - Single core component (not replicated) communicates with multiple groups
  - Multiple inputs, multiple output

# Control of Packet Flow

---

- Packets require different processing blocks
  - IP requires different microblocks than ARP
  - Special packets get handed off to core
- “Dispatch Look” control packet flow among microblocks
  - Each thread runs its own dispatch loop
  - Infinite loop that grabs packets and hands them to microblocks
  - Return value from microblock determines the next step
- Invocation of microblock is similar to function call

# Dispatch Loop

- Example:
  - Two microblocks (ingress + IP)

```
while (1) {  
    Get next packet from input device(s);  
    Invoke ingress microblock;  
    if ( return code == 0 ) {  
        Drop the packet;  
    } else if ( return code == 1 ) {  
        Send packet to ingress core component;  
    } else { /* IP packet */  
        Invoke IP microblock;  
        if ( return code == 0 ) {  
            Drop packet;  
        } else if ( return code == 1 ) {  
            Send packet to IP core component;  
        } else {  
            Send packet to egress microblock;  
        }  
    }  
}
```



# Dispatch Loop Conventions

---

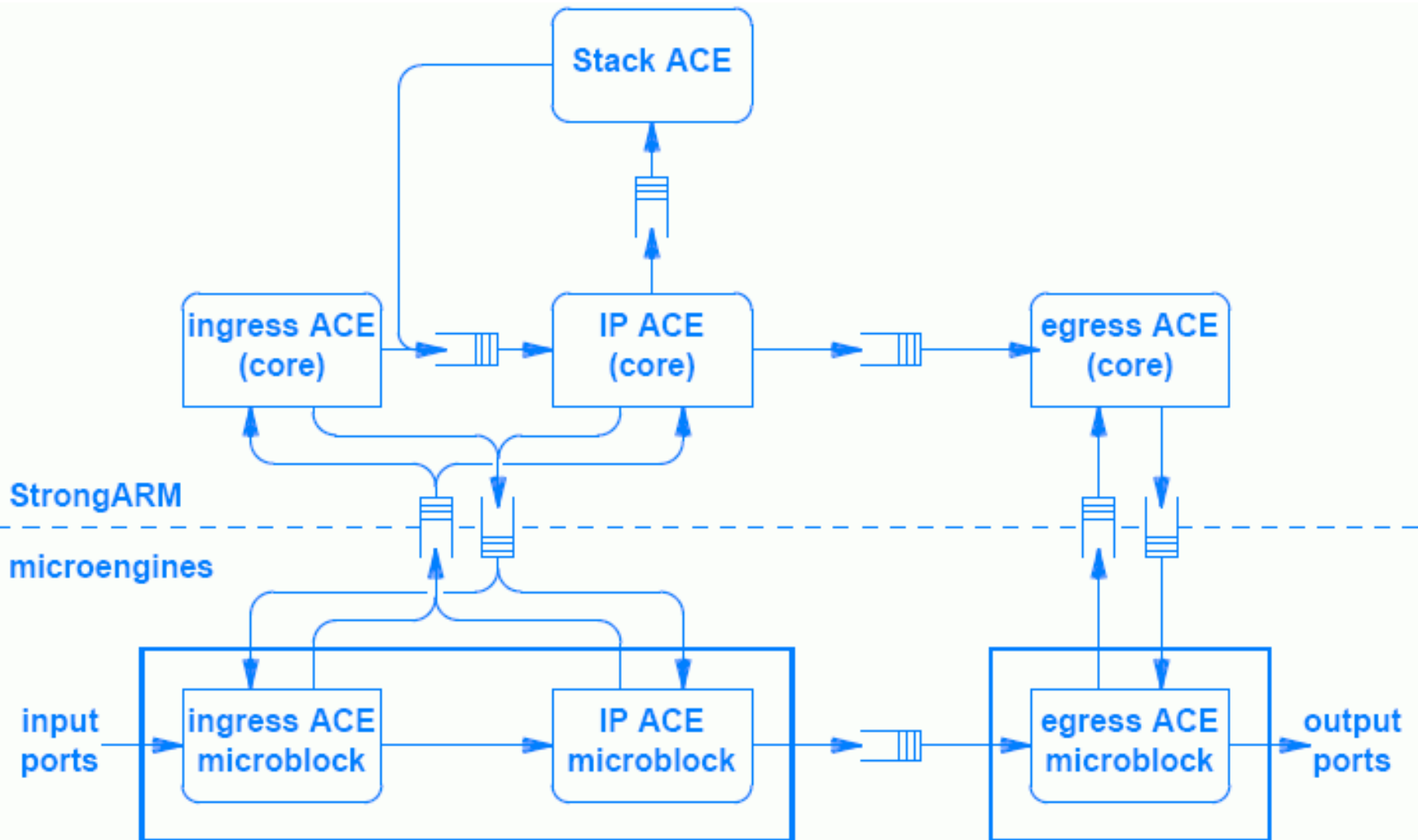
- Parameters passed to microblock:
  - Buffer handle for frame that contains a packet
  - Set of state registers that contain information about the frame
  - A variable called `dl_next_block` in which return value gets stored
- State registers:
  - Information about packet: length
  - Information generated by software: classification result
  - Registers can be changed by microblock
- Return values:
  - Meaning assigned by programmer
  - Conventions: zero = “drop packet”, other values for “pass on” and “send to core” etc.

# Packet Queues

---

- Packet flow depends on packet data
- Processing time depends on packet data
- Packet movement can't be predicted
  - Microblocks need to continue processing without waiting
- Packets need to be buffered
  - “Communication Queues”
  - Unidirectional FIFO (yes, really FIFO)
  - Bidirectional communication requires two queues
- Also between microblocks and core
  - Single queue for all microblock group instances
  - Uses exception mechanism “IX\_EXCEPTION”
  - Exception handler in core determines further steps

# Packet Queue Example



# Crosscalls

---

- Mechanism for non-packet communication between ACEs
  - Similar to remote procedure calls and remote method invocations
- Caller and callee need to agree on parameters
  - Interface Definition Language (IDL) specifies details
  - IDL compiler creates “stubs” to handle marshaling
- Types of crosscalls
  - Deferred: caller does not block, asynchronous notification
  - Oneway: caller does not block, no return value
  - Twoway: caller blocks, callee returns value
- ACEs are prohibited from twoway calls
  - No blocking allowed
- Other control software (non-ACE) may use all types

# SDK

---

- Software Development Kit:

Software	Purpose
C compiler	Compile C programs for the StrongARM
NCL compiler	Compile NCL programs for the StrongARM
MicroC compiler	Compile C programs for the microengines
Assembler	Assemble programs for the microengines
Simulator	Simulate an IXP1200 to debug code
Downloader	Load software into the network processor
Monitor	Communicate with the network processor and interact with running software
Bootstrap	Start the network processor running
Reference Code	Example programs for the IXP1200 that show how to implement basic functions

# Software Setup

Figure 3: Linux software configuration for the Intel IXP1200 Advanced Development Platform using only a Windows NT development environment

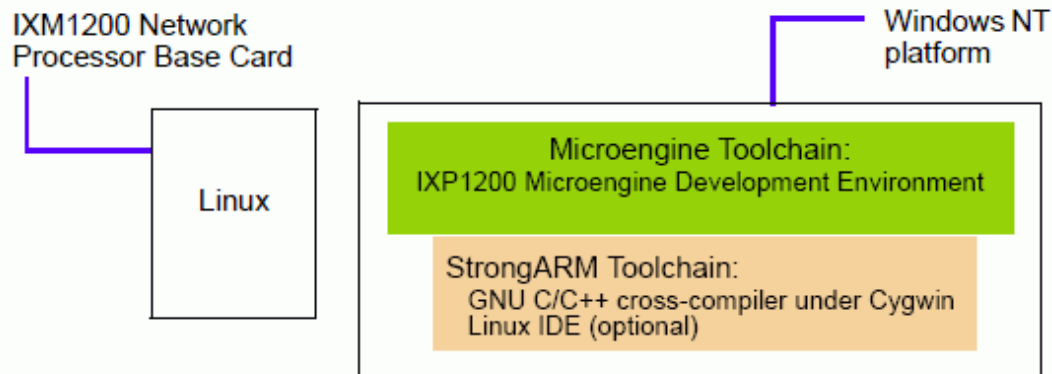
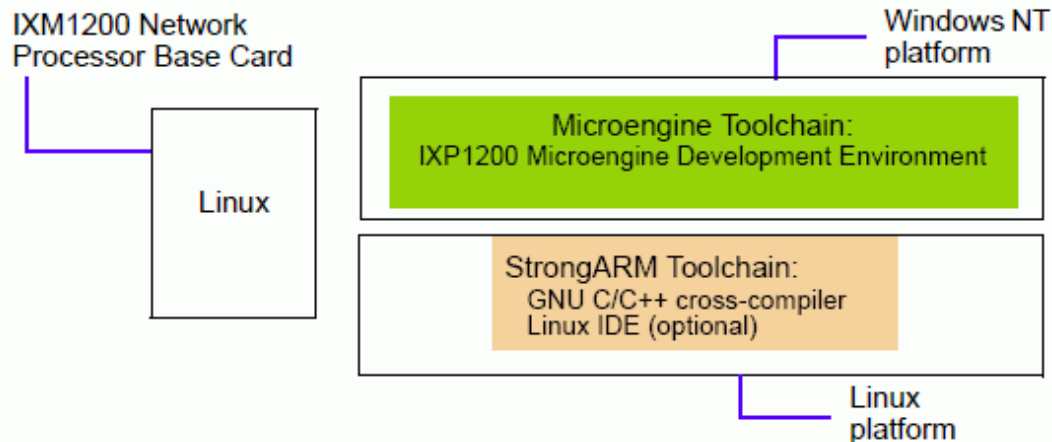


Figure 4: Linux software configuration for the Intel IXP1200 Advanced Development Platform using both Windows NT and Linux development environments



# Simulator

---

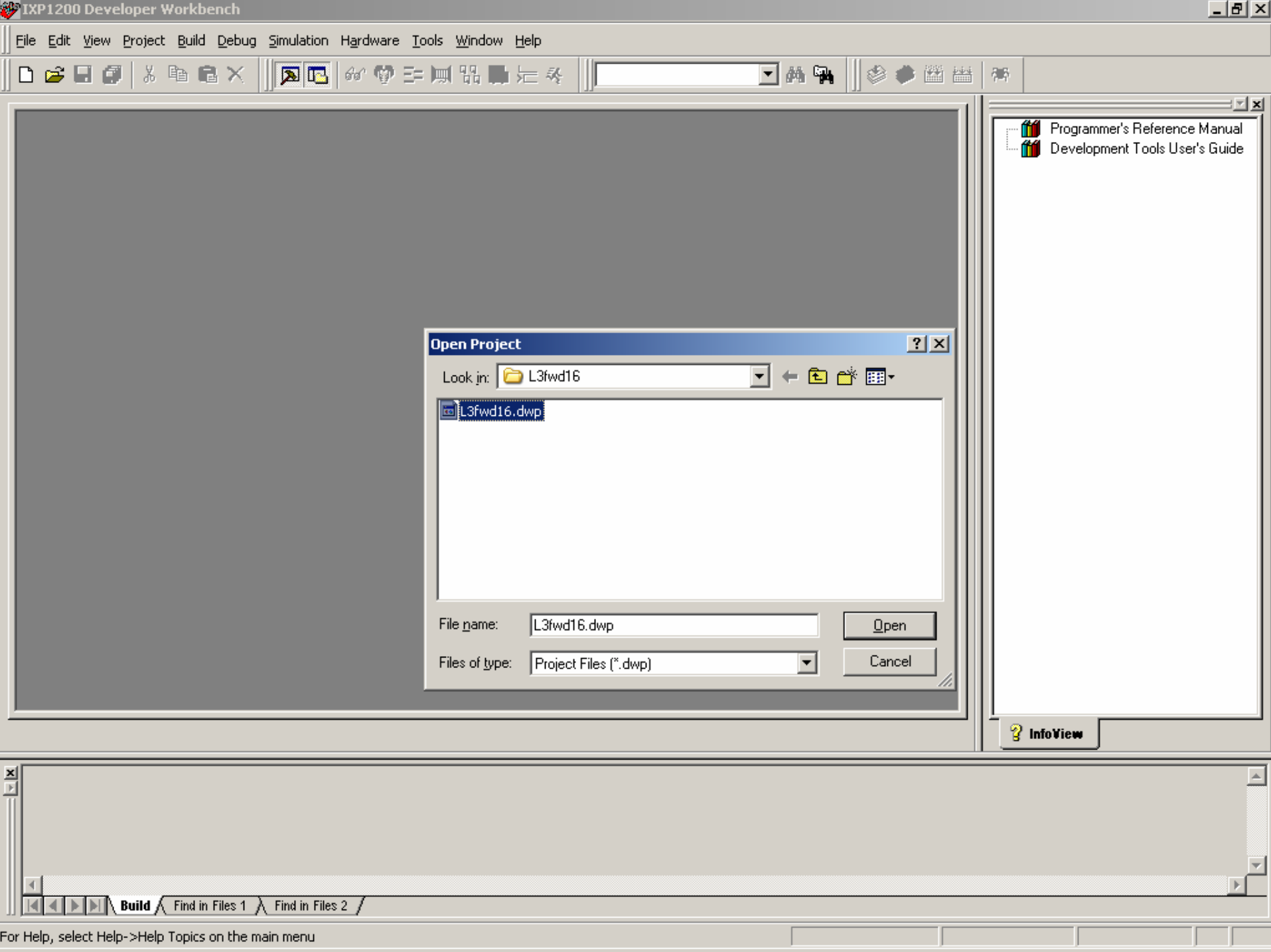
- Cycle-accurate simulation of IXP1200
- Allows for easy experimentation
  - Packet generator
  - Visualization for thread behavior, memory accesses
  - Runs under Windows
- We will use simulator for Lab 2
  - Part I: run existing IP forwarding example, collect statistics
  - Part II: make a minor modification for classification
- We have lab machines set up for you
  - You can also install simulator on your own machine (big!)

# IP Forwarding Example

---

- Full-blown RFC1812-compliant IP forwarding
  - Lots of special cases
  - Look for main program structure
  - 4 uE for IP processing (0-3)
  - 3 uE for output queuing (4-5)
- Run program and collect workload statistics
  - Thread behavior
  - Memory accesses
  - Instruction coverage
  - Etc.







```
// main loop
.while (1)

    xbuf_alloc($pop_xfer, 1);

    .if (packet_buf_addr == UNALLOCATED)
        buf_pop($pop_xfer[0], FREELIST_HANDLE, sig_done); // if
    .endif

    port_rxdy_chk(@rdready_inflight, rec_req);
    critsect_enter[@req_inflight] ; block other contexts from sending
    port_rx_request(rec_req); // get mpacket

#ifdef RFC1812
    port_rx_receive(exception, recv_port, rec_state, ETHER_100M); // get mpacket status
#else
    port_rx_receive(exception, rec_state, ETHER_100M); // get mpacket status
#endif //RFC1812

mpacket_received#: // wait for

    .if (packet_buf_addr == UNALLOCATED)
        buf_wait();
        #if (FREELIST_ID == 0)
            #define BASE_ADDR SRAM_BUFF_DESCRIPTOR_BASE
        #else
            #define_eval HALF_BUFFER_COUNT (BUFFER_COUNT / 2)
            #define_eval BASE_ADDR (SRAM_BUFF_DESCRIPTOR_BASE + (HALF_BUFFER_COUNT * 4))
        #endif
        .while ($pop_xfer[0] == BASE_ADDR)
            buf_pop($pop_xfer[0], FREELIST_HANDLE, ctx_swap); // if no 1
        .endw
        buf_dram_addr_from_sram_addr(packet_buf_addr, $pop_xfer[0], FREELIST_HANDLE);
        move(descriptor_addr, $pop_xfer[0]);
    .endif
    xbuf_free($pop_xfer);
```

## Assembler Source Files

- buf.uc
- constants.uc
- critsect.uc
- cycle.uc
- dram.uc
- endian.uc
- ether.uc
- field.uc
- ip.uc
- ixplib.uc
- mailbox.uc
- mem\_map.h
- packetq.uc
- port.uc
- ports\_validcard.h
- project\_config.h
- rifo.uc
- rx\_ether100m.uc**
- scratch.uc
- sem.uc
- sig.uc
- sram.uc
- stdmac.uc
- tfifo.uc
- tx.uc
- tx\_ether100m.uc**
- tx\_ether100m\_fill.uc
- xbuf.uc

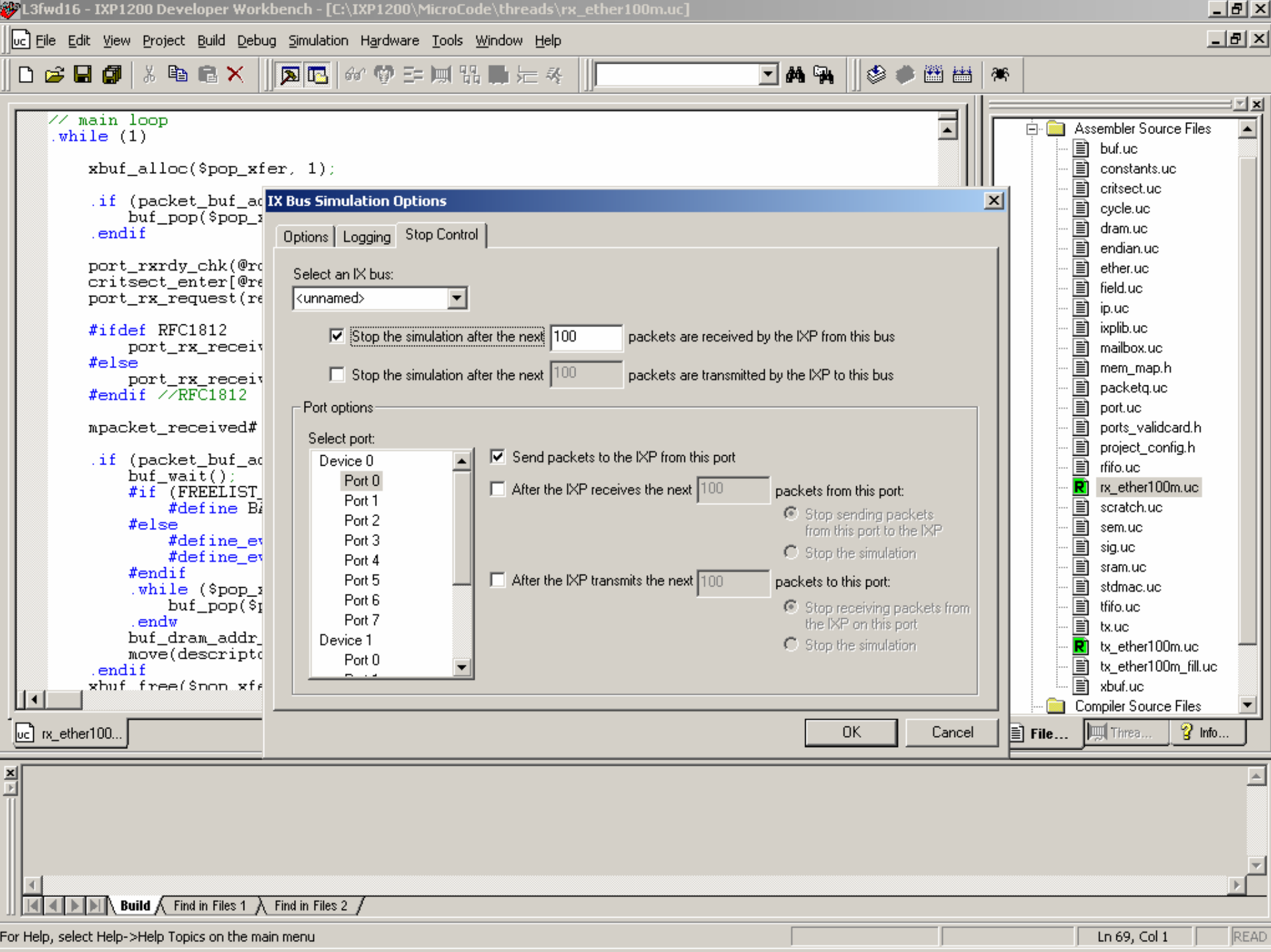
## Compiler Source Files

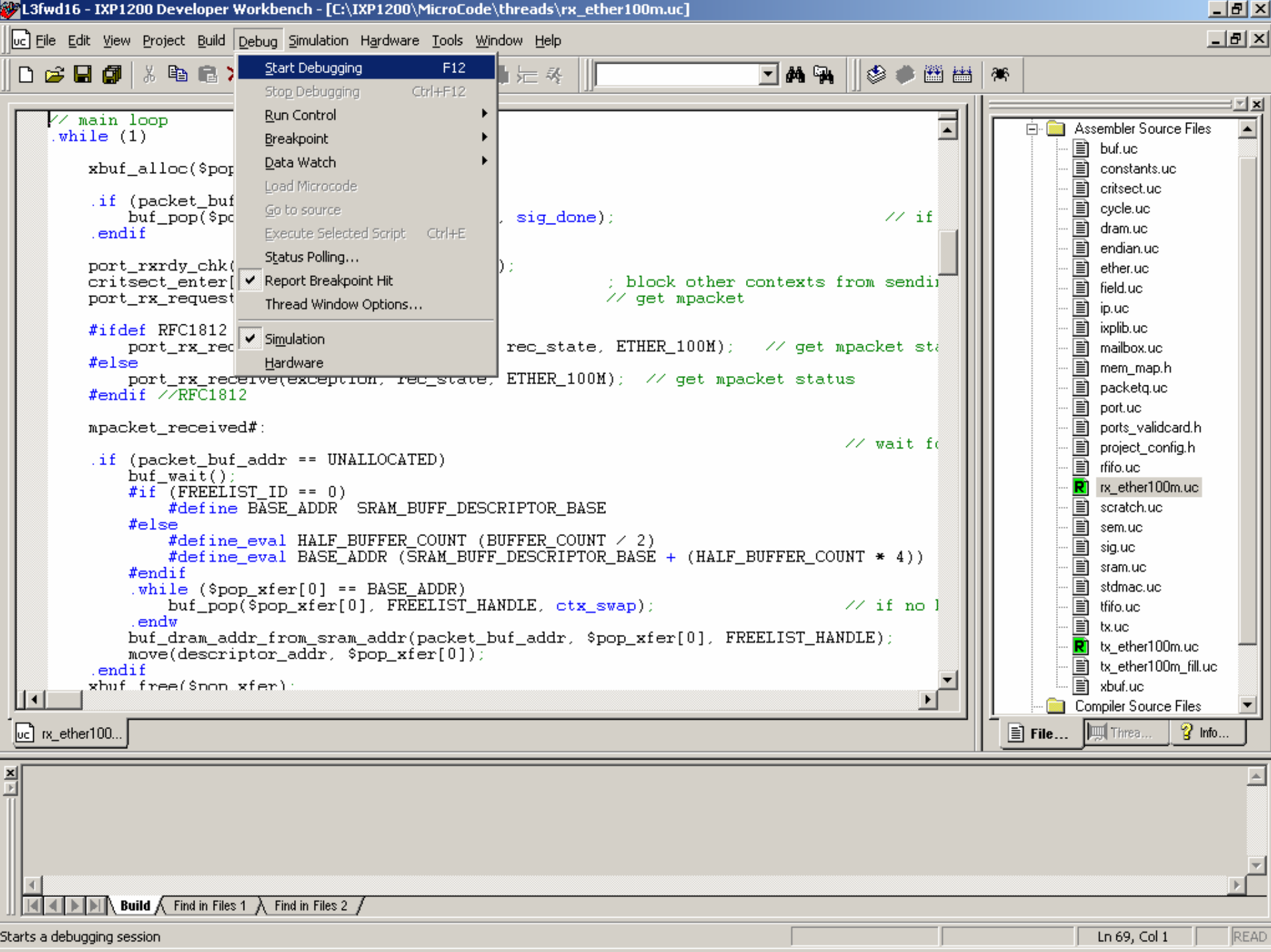
File...

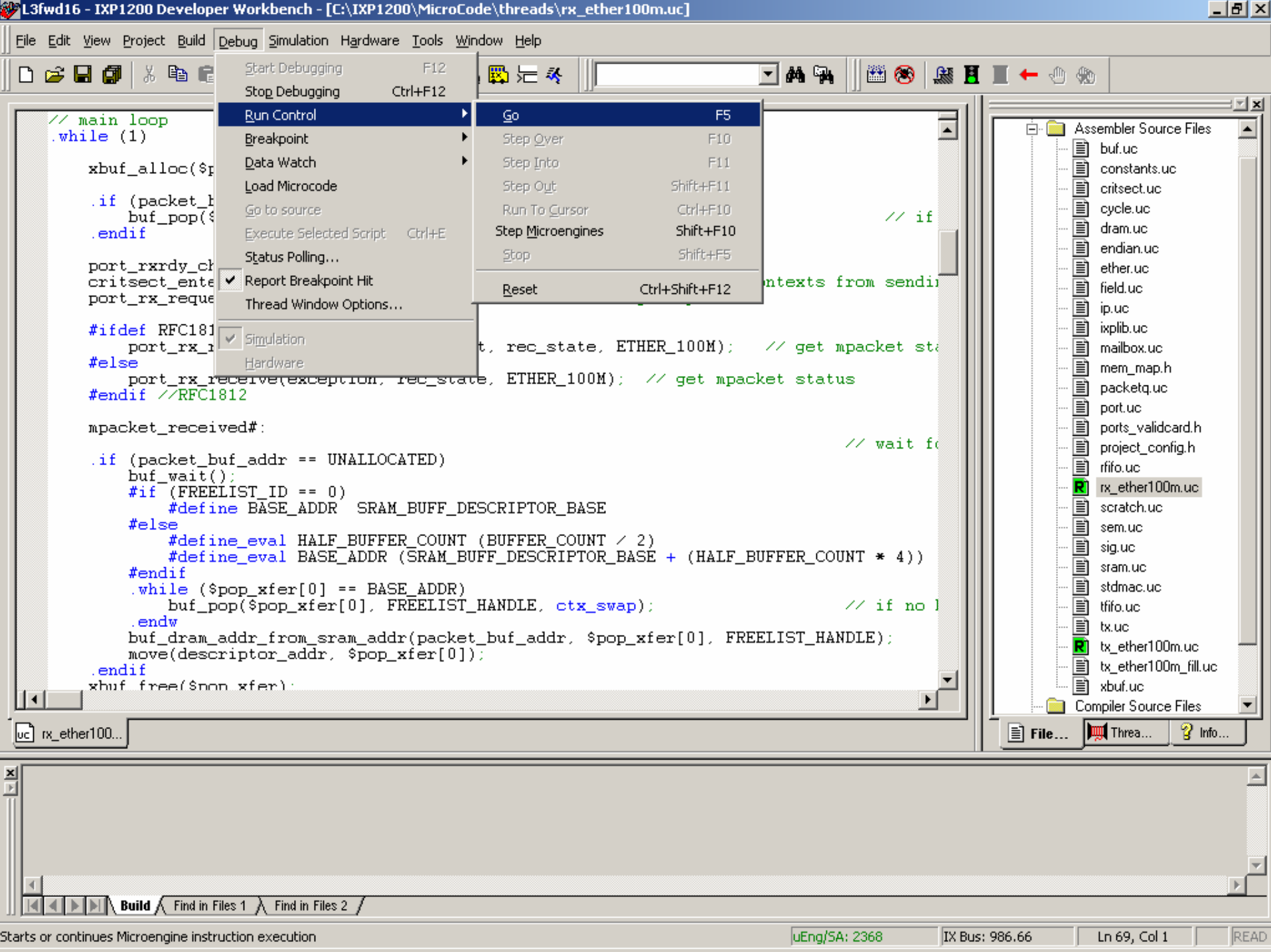
Threa...

Info...

uc rx\_ether100...









```
// main loop
.while (1)

    xbuf_alloc($pop_xfer, 1);

    .if (packet_buf_addr == UNALLOCATED)
        buf_pop($pop_xfer[0], FREELIST_HANDLE, sig_done); // if
    .endif

    port_rxdy_chk(@rdready_inflight, rec_req);
    critsect_enter[@req_inflight] ; block other contexts from sending
    port_rx_request(rec_req); // get mpacket

#ifdef RFC1812
    port_rx_receive(exception, recv_port, rec_state, ETHER_100M); // get mpacket status
#else
    port_rx_receive(exception, rec_state, ETHER_100M); // get mpacket status
#endif //RFC1812

mpacket_received#:

    .if (packet_buf_addr == UNALLOCATED)
        buf_wait();
        #if (FREELIST_ID == 0)
            #define BASE_ADDR SRAM_BUFFER
        #else
            #define_eval HALF_BUFFER_COUNT
            #define_eval BASE_ADDR (SRAM_BUFFER + HALF_BUFFER_COUNT * 4)
        #endif
        .while ($pop_xfer[0] == BASE_ADDR)
            buf_pop($pop_xfer[0], FREELIST_HANDLE, ctx_swap); // if no 1
        .endw
        buf_dram_addr_from_sram_addr(packet_buf_addr, $pop_xfer[0], FREELIST_HANDLE);
        move(descriptor_addr, $pop_xfer[0]);
    .endif
    xbuf_free($pop_xfer);
```

## IXP1200 Developer Workbench - Error



The simulation is being stopped because the IXP received 100 packets from IX bus.

OK

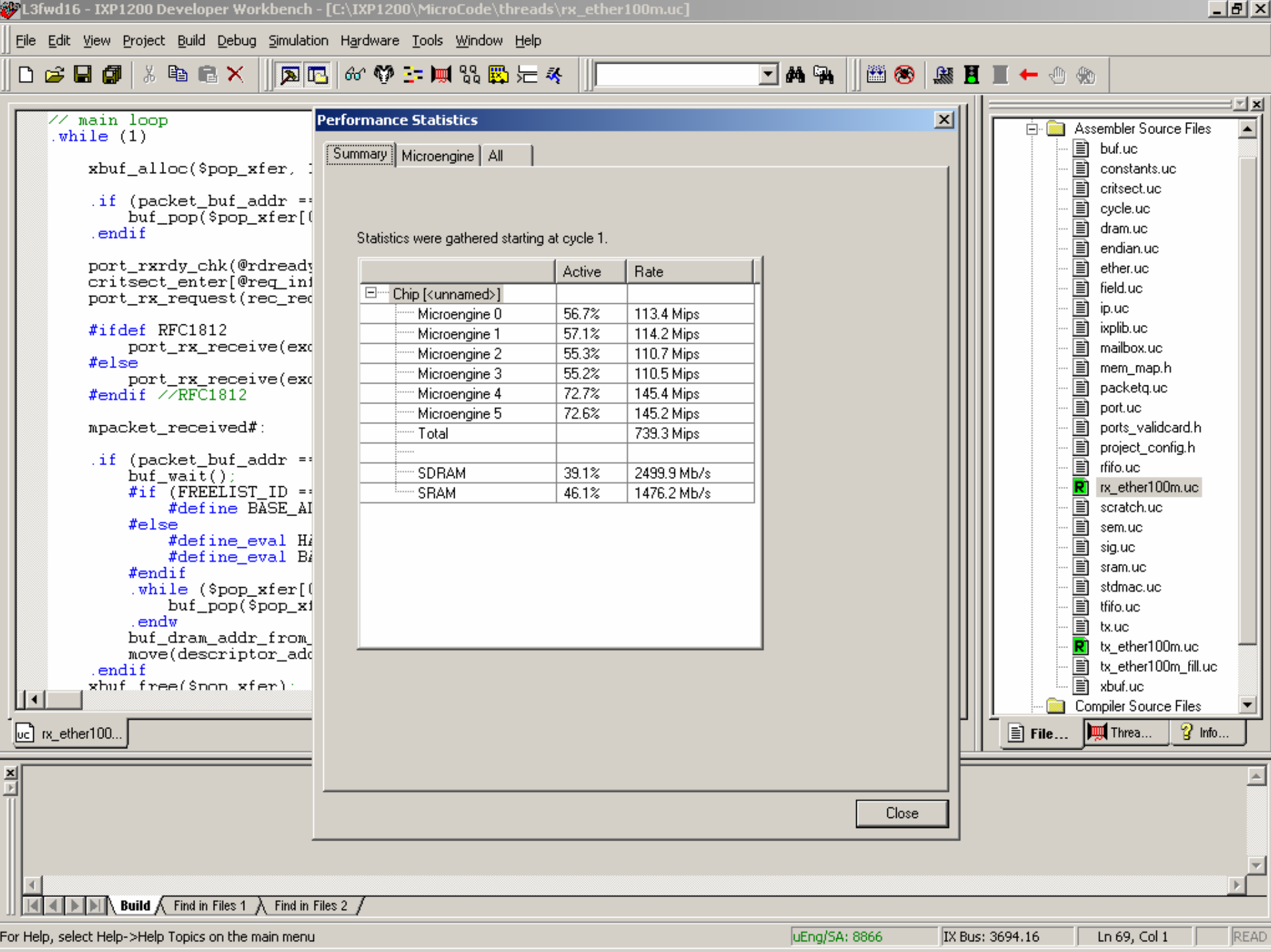
## Assembler Source Files

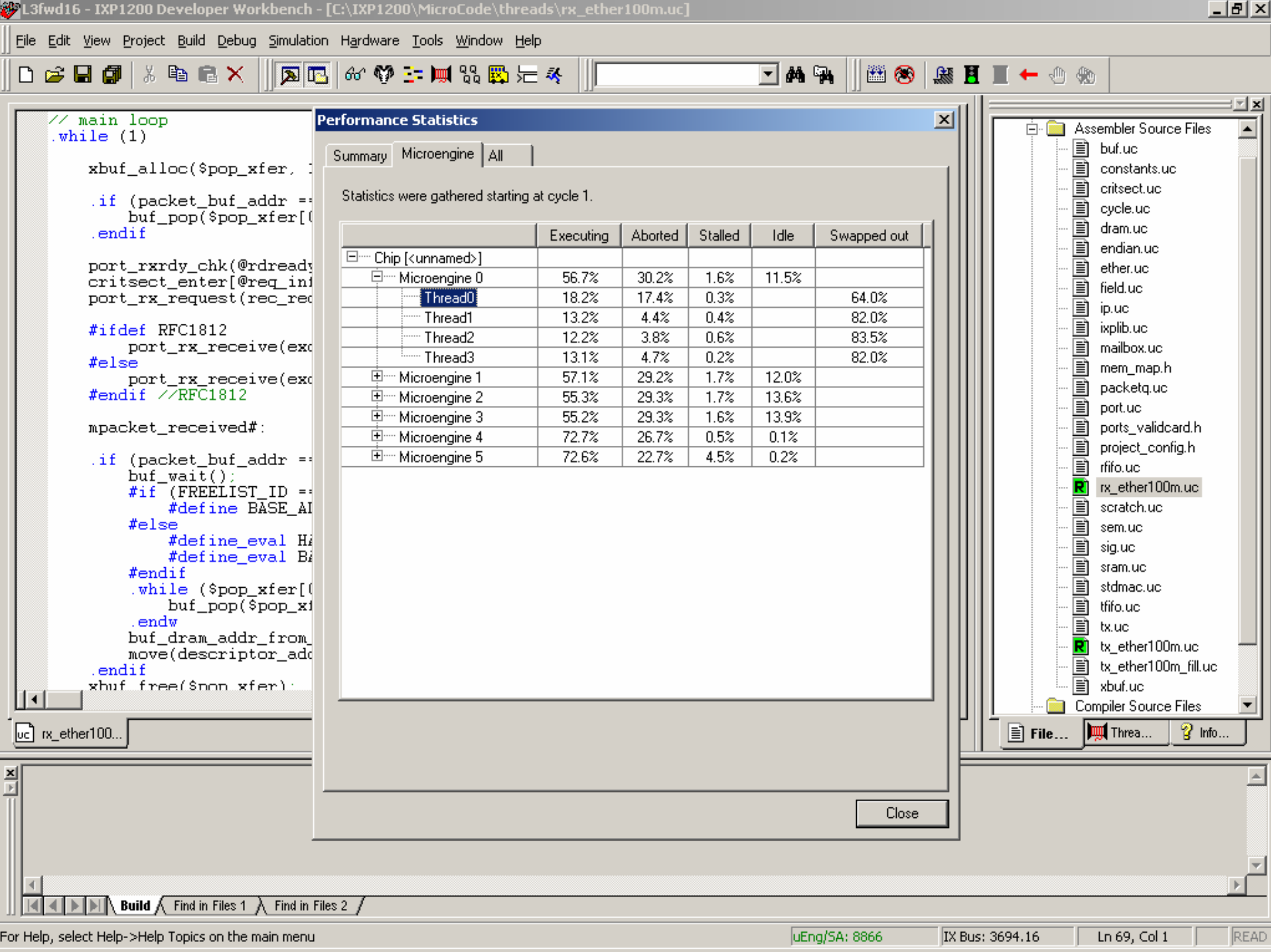
- buf.uc
- constants.uc
- critsect.uc
- cycle.uc
- dram.uc
- endian.uc
- ether.uc
- field.uc
- ip.uc
- ixplib.uc
- mailbox.uc
- mem\_map.h
- packetq.uc
- port.uc
- ports\_validcard.h
- project\_config.h
- rifo.uc
- rx\_ether100m.uc**
- scratch.uc
- sem.uc
- sig.uc
- sram.uc
- stdmac.uc
- tfifo.uc
- tx.uc
- tx\_ether100m.uc**
- tx\_ether100m\_fill.uc
- xbuf.uc

## Compiler Source Files

File... Threa... Info...







```
// main loop
while (1)

xbuf_alloc($pop_xfer, 0)

.if (packet_buf_addr == 0)
    buf_pop($pop_xfer, 0)
.endif

port_rxdy_chk(@rdready)
critsect_enter[@req_in]
port_rx_request(rec_req)

#ifdef RFC1812
    port_rx_receive(exc)
#else
    port_rx_receive(exc)
#endif //RFC1812

mpacket_received#:

.if (packet_buf_addr == 0)
    buf_wait()
    #if (FREELIST_ID == 0)
        #define BASE_ADDR 0
    #else
        #define_eval HZ
        #define_eval BZ
    #endif
    .while ($pop_xfer[0] != 0)
        buf_pop($pop_xfer, 0)
    .endw
    buf_dram_addr_from
    move(descriptor_addr, buf_dram_addr)
.endif
xbuf_free($pop_xfer)
```

### Performance Statistics

Summary

Microengine

All

Statistics were gathered starting at cycle 1.

	Executing	Aborted	Stalled	Idle	Swapped out
Chip [unnamed]					
Microengine 0	56.7%	30.2%	1.6%	11.5%	
Thread0	18.2%	17.4%	0.3%		64.0%
Thread1	13.2%	4.4%	0.4%		82.0%
Thread2	12.2%	3.8%	0.6%		83.5%
Thread3	13.1%	4.7%	0.2%		82.0%
Microengine 1	57.1%	29.2%	1.7%	12.0%	
Microengine 2	55.3%	29.3%	1.7%	13.6%	
Microengine 3	55.2%	29.3%	1.6%	13.9%	
Microengine 4	72.7%	26.7%	0.5%	0.1%	
Microengine 5	72.6%	22.7%	4.5%	0.2%	

### Assembler Source Files

- buf.uc
- constants.uc
- critsect.uc
- cycle.uc
- dram.uc
- endian.uc
- ether.uc
- field.uc
- ip.uc
- ixplib.uc
- mailbox.uc
- mem\_map.h
- packetq.uc
- port.uc
- ports\_validcard.h
- project\_config.h
- rfifo.uc
- rx\_ether100m.uc
- scratch.uc
- sem.uc
- sig.uc
- sram.uc
- stdmac.uc
- tfifo.uc
- tx.uc
- tx\_ether100m.uc
- tx\_ether100m\_fill.uc
- xbuf.uc

### Compiler Source Files

File...

Threa...

Info...

Close

Build

Find in Files 1

Find in Files 2

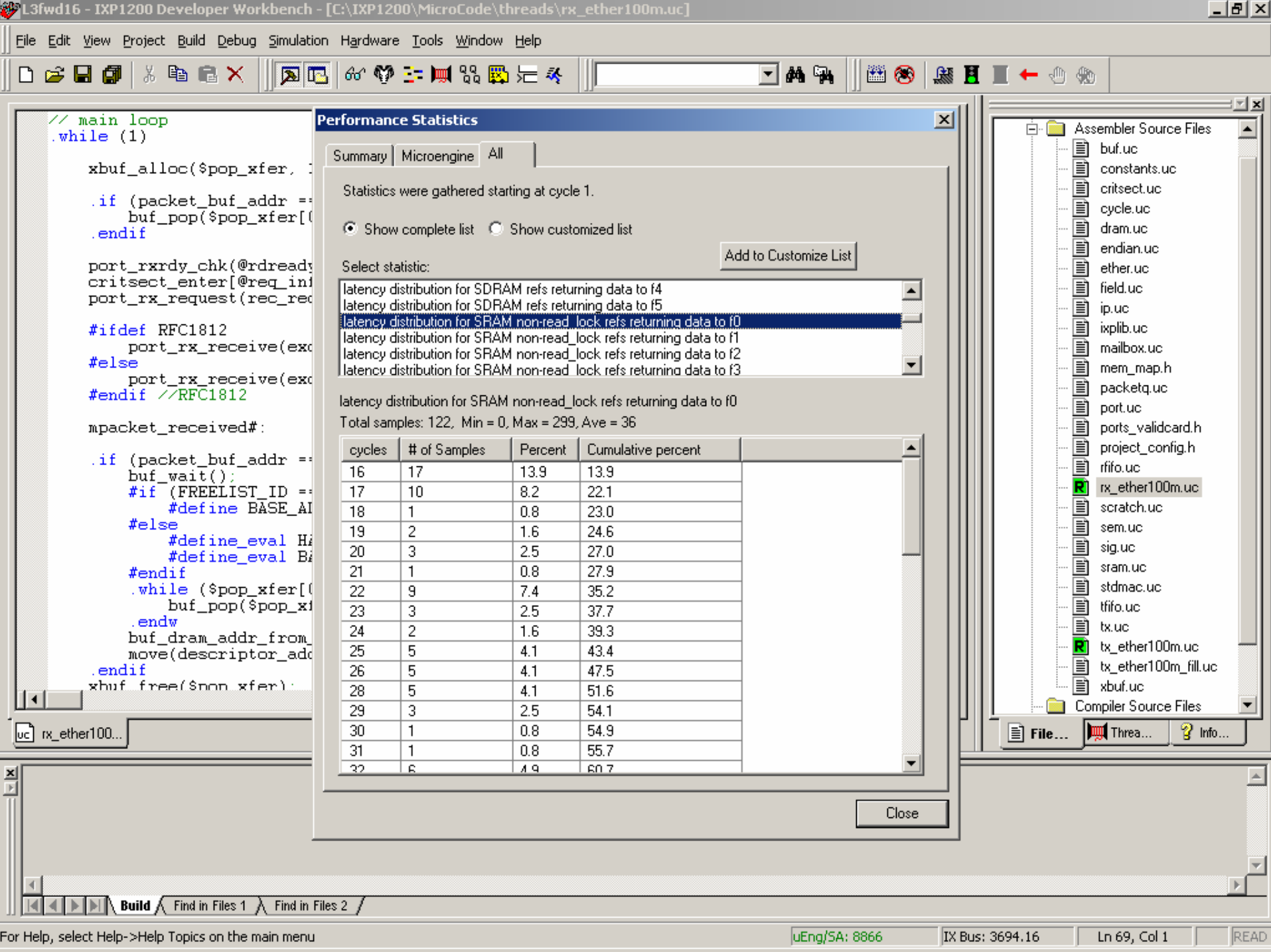
uEng/SA: 8866

IX Bus: 3694.16

Ln 69, Col 1

READ





## Execution Coverage

<unnamed>

Microengine 0

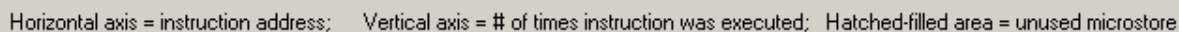
\_\_\_\_\_

☒ Context 2 (Thread2)

### Reset Counts

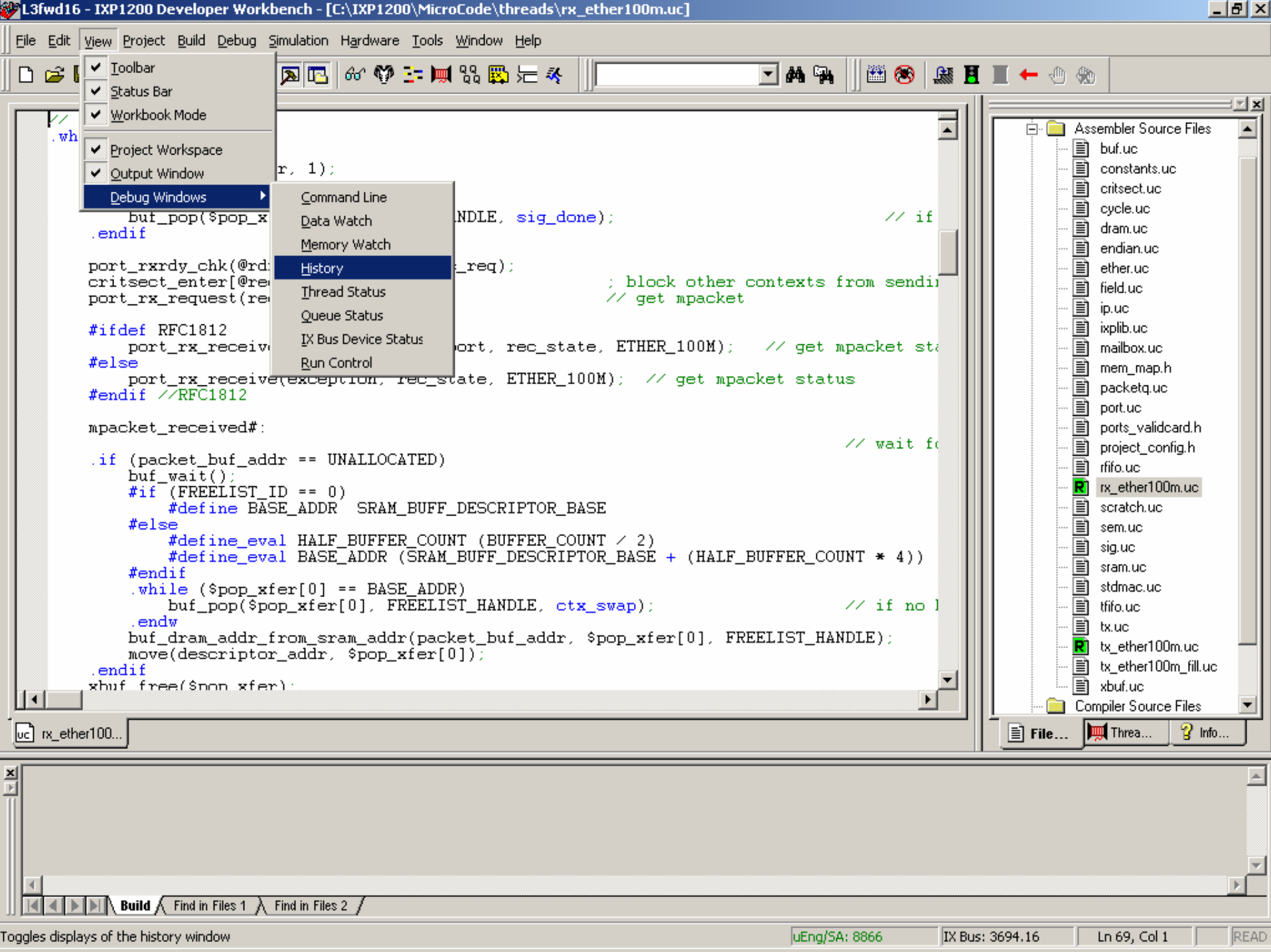
Close

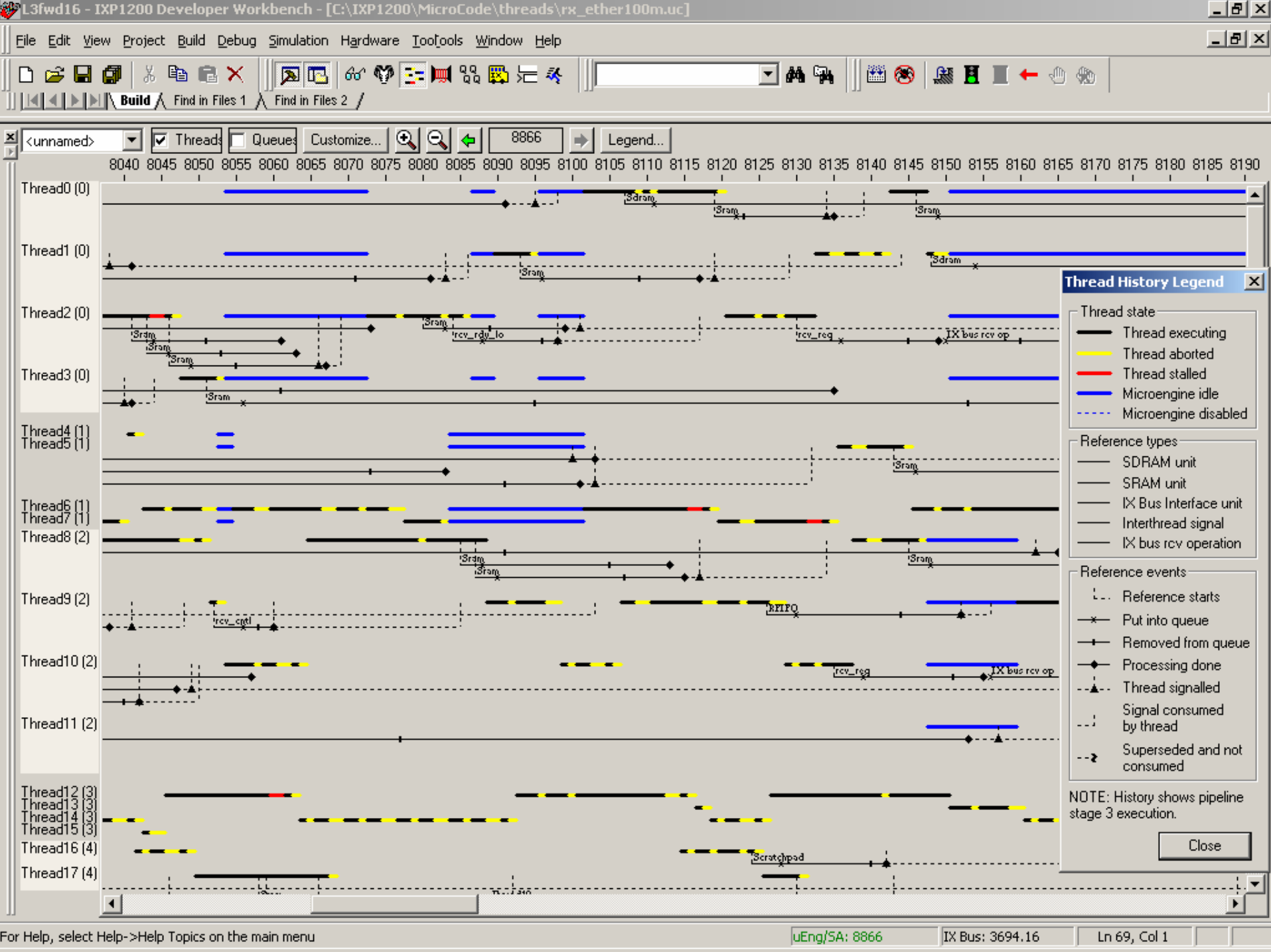
Values were collected starting at cycle 0.

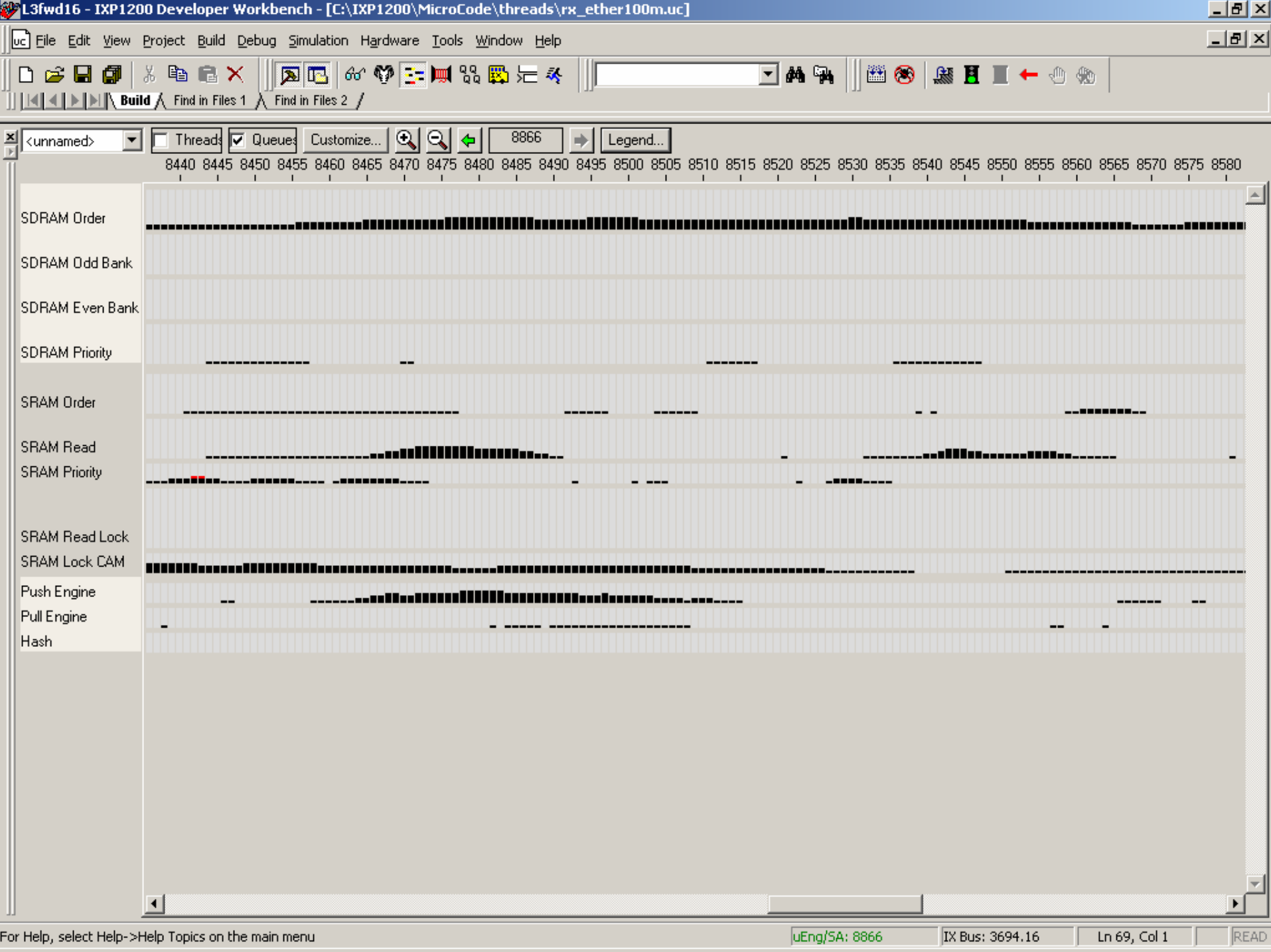


er Source Files


**?** Info







# Lab 2

---

- Part I: Collect statistics
  - Microengine utilization for all microengines
  - Detailed statistics of one thread from uE 0 and one from uE 5
  - Processing power of microengines (in MIPS).
  - Memory utilization and bandwidth.
  - Latency distribution for SDRAM refs for microengine 0 and SRAM non-read\_lock refs for microengine 0. Show a graph.
  - Show a screenshot for the thread history that shows overlapping SRAM and SDRAM requests by the same microengine.
  - Identify the overall delay for either request (in cycles). What factors contributed how much to the overall delay?
- DUE NEXT TUESDAY.

# Lab 1 Results

---

- Grading: 20 points total
  - Results: 10 points
  - Code: 3 points
  - TCP state machine + explanation: 2+1 points
  - IP and TCP headers: 1+1 points
  - Report (written content): 2 points
- Average: 16.6
- Max: 20
- Min: 14

# Next Class

---

- Microengine programming
  - Assembler
  - Instructions
  - Register access
  - Assembler directives
  - Etc.
- Read Chapter 24
- Turn in Part I of Lab 2