**IOCP 정리**

**2015182025**

**게임공학과**

**오현호**

**2021 01 09 ~ ing**

**P3 객체 생성 CreateIoCompletionPort**

**P4 소켓 등록 CreateIoCompletionPort + 임시 Worker Thread**

**P5 GetQueuedCompletionStatus,**

**P6 PostQueuedCompletionStatus**

**P7 IOCP 객체 구조 모양**

**P8 IOCP 서버 설계 + Accept 처리**

**P9 클라이언트 객체 + 오버랩 구조체**

**P10 오버랩IO 포인터의 확장 + 완료 처리**

**P11 버퍼 처리 – Recv**

**P12 버퍼 처리 – Send**

**P12 버퍼 처리 – Send**

**P12 버퍼 처리 – Send**

**P13 오버랩IO에서 IOCP로 추가할 것 + 프로토콜 정의 + 패킷 처리 루틴**

**P14 Recv 구현 + 패킷 재조립**

**P15 Recv 패킷 재조립 구현 + AcceptEx 함수**

**P16 WSASend + Send 완료 구현**

**HANDLE CreateIoCompletionPort (**

**HANDLE FileHandle,**

**HANDLE ExistingCompletionPort,**

**ULONG\_PTR CompletionKey,**

**DWORD NumberOfConcurrentThreads );**

**HANDLE hIOCP = CreateIoCompletionPort(INVALID\_HANDLE\_VALUE, NULL, NULL, 0};**

커널 객체 생성 API

FileHandle에는 무조건 INVALID\_HANDLE\_VALUE가 들어가야함

그래야 IOCP 객체를 생성함

NumberOfConcurrentThreads - IOCP 객체가 동시에 컨트롤 할 수 있는 쓰레드 수

0을 넣으면 컴퓨터에 돌아가는 코어의 갯수만큼 사용하겠다는 뜻

**HANDLE CreateIoCompletionPort(socket, hIOCP, key, 0};**

IOCP 객체를 만들었으니 이제 객체를 이용해 send recv를 해야함

데이터 통신을 하려면 소켓을 IOCP 객체에 등록을 해줘야 함

등록을 해줘야 소켓의 send recv가 IOCP 객체를 통해서 이루어짐

자세히 보면 위에 객체 생성API랑 이름이 같음. 이래서 IOCP가 어려움

socket을 hIOCP 객체에 연결을 하겠다는 의미

key는 socket을 어떤 숫자로 부를까 정하는 것

IOCP를 IO할 때 소켓을 주지 않고 키를 줌, 그래서 소캣을 등록시킨 다음엔 키를 통해서access함,

key 값은 개발자 마음이므로 주로 플레이어의 ID나 클라이언트 자료를 저장하는 자료구조의 주소값으로 설정함, 중요한건 값이 유니크해서 중복이 없어야함

==========================================================

IOCP에서 관리하는 쓰레드를 생성하는 API는 따로 없음

멀티스레드로 동작할 때 IOCP 객체에 스레드들을 어떻게 연동을 시키느냐?

스레드 생성은 프로그래머 알아서 생성하고 나중에 등록을 시킴

IOCP는 굳이 멀티스레드 아니어도 동작 가능

메인스레드를 IOCP 스레드로 취급을 해서 가능

그러나 고성능 프로그래밍 하려면 멀티스레드 필수임

IOCP로 오버랩IO할 때 완료 검사는 어떻게 하느냐?

**while(true) { //서버 메인 루프**

**GetQueuedCompletionStatus(**

**hIOCP,**

**&dwIOSize,**

**&key,**

**&lpOverlapped,**

**INFINITE;**

**);**

**lpOverlapped에 따른 처리;**

**}**

**GetQueuedCompletionStatus** 함수 호출을 하면 멈춰있다가 오버랩IO가 완료되면 함수 종료되면서 정보들을 파라미터에 넣어줌, 우리는 이걸 가지고 (**lpOverlapped에 따른 처리)** 여기서 오버랩IO CallBack에서 했던걸 해주면 된다

그래서 사실상 얘가 스레드를 IOCP 객체에 등록해주는 함수임

그래서 지금은 다른 스레드를 선언하지 않았지만 얘를 메인 스레드에 호출하면 메인 스레드를 IOCP에 넘겨주는거임, 하위는 자세한 파라미터

**BOOL GetQueuedCompletionStatus(**

**HANDLE CompletionPort,**

**LPDWORD lpNumberOfByte,**

**PULONG\_PTR lpCompletionKey,**

**LPOVERLAPPED \*lpOverlapped,**

**DWORD dwMillisocnds );**

**CompletionPort –** IOCP 객체 핸들 넣어주면 된다

**lpNumberOfByte –** 총 전송된 데이터 양

**lpCompletionKey –** 오버랩IO니까 여러 개의 IO가 중첩되서 진행되고 있고 그중 하나의 IO가 완료됐을 때 누가 완료됐는지 확인용, 아까 위에서 말했던 키 값을 알려줌

**\*lpOverlapped –** 오버랩IO할땐 오버랩IO 구조체 반드시 사용해야 하는데 IOCP도 오버랩IO이니까 사용해야함, 오버랩 구조체 리턴해줌

뒤에 진도 배우면 여길 통해서 여러가지 정보를 구할 수 있음

**dwMillisocnds –** 이 함수의 타임아웃, 얼마나 기다릴건지 설정

INFINITE 쓰면 IO가 하나라도 종료하지 않으면 무한루프 상태

적절한 숫자를 설정하면 그 시간에 리턴함, 타임아웃이므로 타임아웃 에러코드가 뜸

**BOOL PostQueuedCompletionStatus(**

**HANDLE CompletionPort,**

**DWORD NumberOfByte,**

**ULONG\_PTR dwCompletionKey,**

**LPOVERLAPPED lpOverlapped);**

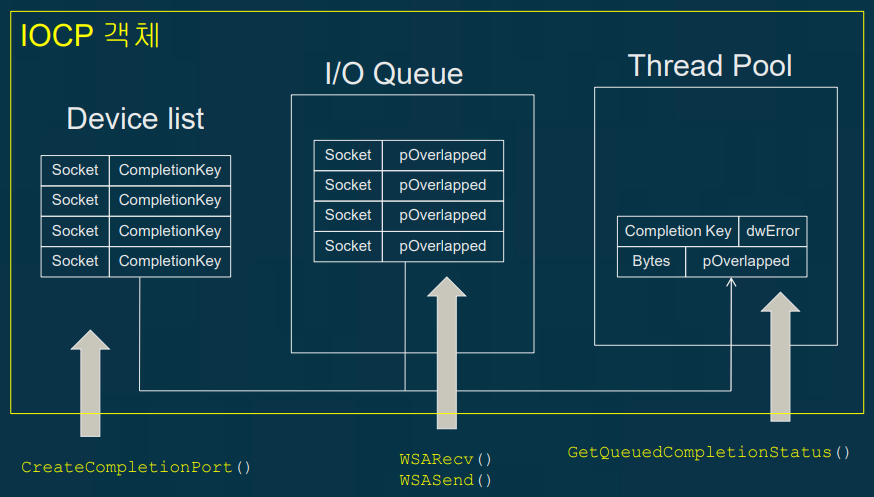
IOCP 객체에서 완료를 한 개 전달해주는 함수

이 함수의 뜻은 파라미터들을 **GetQueuedCompletionStatus**에 넘겨주라는 뜻

**PostQueuedCompletionStatus**를 호출하면 **GetQueuedCompletionStatus**가 전달받은 파라미터들로 그대로 리턴함, 검사 하나도 안함 진짜 넘겨준 그대로 리턴함

그래서 이건 프로그래머가 임의의 값을 넘겨주는 용도로 사용

스레드 관리에 유용하게 사용됨, 나중에 NPC 할 때 쓸거임



그래서 이 IOCP 함수들을 사용하면 이런 구조가 형성된다.

**<IOCP 서버 설계> (싱글스레드 기준)**

1. 초기화

바인드 리슨 등등 + IOCP 핸들 생성

1. 리슨 소켓 IOCP에 등록

Accept 해야되는데 그러면 클라가 연결될때까지 멈춰있어야함

물론 겜서버는 accept 기다려야되는데 그 와중에 send recv 끊임없이 해야함

멀티스레드라면 Accept 전용 스레드 하나 열면 되는데 싱글스레드이니까

오버랩IO로 accept 받아야함, 그러기 위해선 리슨 소켓을 IOCP에 등록하고 AcceptEX (오버랩IO 버전의 accept) 호출해야함

1. 서버 메인 루프

새로운 소켓 접속하면 IOCP 연결 후 WSARecv 호출

WSARecv 완료되면 클라이언트에서 준 데이터로 패킷 처리

WSASend 완료되면 자료구조 정리(메모리 재사용)

**<Accept 처리>**

Accept는 블로킹 처리되면 안 되므로 AcceptEx 사용해서 비동기 처리

리슨 소켓을 IOCP 등록하고 AcceptEx 호출하면 IOCP 객체에서 완료 알려줌

완료되면 무한 루프에서 클라이언트 정보를 저장하는 클라이언트 객체를 만들고

접속된 소켓은 IOCP에 등록, 그래야 클라이언트가 IOCP로 동작을 함

그다음 WSARecv 호출하고 오버랩IO recv callback을 만들어서 동작시켜야함

그리고 다시 AcceptEx() 호출

**<클라이언트 객체>**

서버는 클라 정보 가지고 있는 객체가 필요함

ID, 네트워크 접속 정보 (소켓, 네트워크 버퍼), 연결 상태, 게임 정보 (hp, 위치)

**GetQueuedCompletionStatus** 리턴 값 받았을 때 클라이언트 객체를 찾을 수 있어야함

IOCP에서 오고 가는건 key, 총 데이터 크기, 오버랩IO 포인터 이므로 key를 클라이언트 객체의 포인터로 하거나 클라이언트 객체의 ID 혹은 Index로 하면 된다.

주로 클라이언트 ID로 많이 함

**<오버랩 구조체>**

모든 send recv는 오버랩 구조체가 필요함, send recv는 순서대로 아닌 동시다발적으로 일어남, 그래서 구조체가 스크립 당 하나만 있으면 되는게 아니라 여러 개가 필요함

왜냐면 send recv가 중첩되서 실행되고 있는데 하나의 구조체를 동시에 여러 호출에서 사용하는 것은 불가능하다

왜 불가능하냐, 데이터가 에러가 남, 호출할땐 안나는데 동작하면서 에러 및 데이터 깨짐

여러 개의 구조체를 어떻게 관리하고 할당할지 잘 생각해야함

소켓당 recv 호출은 무조건 한 개여야함, 여러 개일 경우 어떤 데이터가 어떤순서로 들어올지 모름, 두개 이상 호출해도 두번째부터 에러가 날거임

소켓이 100개면 100개의 호출이 중첩되야 하고 다 각각의 다른 소켓에 대한 recv호출

그래서 클라이언트 recv 호출용 오버랩 구조체 하나 만들어놓고 돌려쓰는 게 좋음

왜냐? New/delete 오버헤드 줄이기 위해

소켓당 send 호출은 여러 개가 될 수 있음 -> send 버퍼도 같은 개수만큼 필요함

그리고 개수를 알 수 없으므로 new/delete 사용해야함

Send 할때 new, send가 complete되었을 때 delete -> 근데 new/delete 자체가 별로 좋은 방법은 아니니까 Free list나 available list 사용해서 따로 관리하는 게 성능 더 좋음, 만약 따로 관리할 시 send 버퍼뿐만 아니라 오버랩 구조체도 같이 관리해줘야함

그래서 그냥 send 버퍼랑 오버랩 구조체를 같이 묶어서 관리할거임(제일 흔한 방법)

**<오버랩IO 포인터의 확장>**

오버랩IO 구조체는 반드시 필요함, 그러나 구조체 자체에서 쓸만한 정보는 없음

커널이 사용하는거지 게임서버 내에선 안씀, 그래서 오버랩IO 구조체 자체에다가 우리가 필요한 정보를 넣을 공간은 없음, 실습이랑 과제에서 hEvent 안에다가 소켓을 넣는데 표준도 아니고 안전하다는 보장도 없음, 그래서 구조체를 사용을 해야되는데 마땅히 방법이 없음 -> 꼼수 써야함, 우리는 정보가 필요하기 때문에 구조체를 써먹어야함

IOCP 등록 완료 후 **GetQueuedCompletionStatus**에서 주는 정보만으로 완료처리 해야하는데 num of byte, key를 알려주므로 key를 통해 어떤 소켓인지는 알 수 있는데 가장 중요한 send 인지 recv 인지 accept인지 알 수가 없음 그래서 정보를 추가 해야함

오버랩IO 구조체 안에는 추가할 수 없음, 구조체에 정보를 덧붙여서 전달 해줘야함(확장)

**GetQueuedCompletionStatus**에서는 IO가 완료되었다, 어느 소켓이다, 몇 바이트가 전송되었다 -> 그러나 그 바이트가 어느 버퍼인지(버퍼 위치) 안 알려줌

recv같은 경우는 버퍼가 한 개니까 거기서 확인 가능한데 send인 경우 어떤 버퍼에서 완료되었는지 알 수 없음. 버퍼의 내용은 알 필요 없지만 버퍼의 위치는 알아야함

왜? Send에 사용된 버퍼는 재사용 해야되므로, new했다하면 위치를 알아야 delete 가능

메모리 pool로 관리한다치면 위치를 알아야 반환 가능함

**<완료 처리>**

**GetQueuedCompletionStatus**를 통해 완료 정보를 얻고 에러 났을 경우 에러처리

소켓이 close 되면 접속종료처리 해줘야함

확장 오버랩IO 구조체 통해서 알아낸 send recv accept를 처리해줘야함

**Accept인 경우**

새 클라이언트 객체를 등록하고 recv 호출함

**Recv인 경우**

패킷이 다 왔는지 검사 후 다 왔으면 패킷 처리

여러 개의 패킷이 한 번에 왔을 때 하나씩 처리

패킷이 중간에 잘릴 수 있으니 다음 recv에서 합쳐서 처리-> 패킷 재조립

Recv 처리 끝나면 recv 호출을 더 해서 오버랩IO 상태 유지

**Send인 경우**

Send는 누가 호출하냐? Recv가 호출을 함

클라이언트가 패킷을 보내면 recv에서 처리를 하고 다른 클라이언트한테 정보를 알려줘야 하거나 본인 소켓에 응답을 해야할 때 send하고

Send 완료 후 send할 때 사용한 오버랩 구조체랑 버퍼를 free or 재사용 해야함

**<버퍼관리>**

**RECV** 버퍼 관리 잘 신경써야함

하나의 소켓에 대한 recv 호출은 언제나 하나이기 때문에 하나의 버퍼만 계속 사용

근데 패킷들이 중간에 짤린 상태로 도착할 수 있어서 recv 버퍼를 함부로 비우면 X

그럼 읽은곳 뒤에 이어서 읽어와야 되는데 그러면 recv 버퍼 관리가 어려워짐

그래서 일단 읽은 거를 저장해두는 버퍼가 추가로 필요함

그게 귀찮으면 ring buffer 사용->복잡해지지만 데이터 복사가 줄어서 성능이 올라감

우리는 링버퍼 안쓸건데 본인이 IOCP 이해도가 높아지면 링버퍼 시도해볼 것

**SEND** 하나의 소켓에 여러 개의 send 동시에 할 수 있음

클라이언트는 여러 개임, 클라 하나가 상태가 바뀌었을 때 다른 모든 클라이언트에서 브로드캐스팅 된다. 그래서 클라들은 본인에 관련된 send와 다른 클라 관련된 send를 동시다발적으로 받음.

가장 간단한 방법은 send 할때마다 오버랩 구조체와 버퍼를 new/delete로 할당

아니면 클라마다 send형 구조체를 만들어서 send 할때마다 queue에 쌓아놨다가 하나씩 send를 한다치면 프로그래밍도 복잡해지고 여러 개의 send를 동시에 보낼 수 없어서 성능저하 발생함-> 오버랩 구조체와 버퍼를 다이나믹하게 할당을 받아서 비동기 식으로 그때그때 send 해줘야함

근데 소켓이 하나 있는데 여러 개의 클라이언트가 동시에 send를 하면 데이터가 섞이지 않냐? -> 그렇지 않음, send를 실행한 순서대로 버퍼에 넣어서 전송하기 때문, 버퍼에서 어떤 순서로 보내질지는 모르지만 확실한건 데이터가 섞이진 않음

내부 버퍼가 차서 send가 중간에 잘렸는지는 number of byte로 확인 가능

아까 위에서 순서는 모르지만 데이터는 안섞인다 했지만 내부 버퍼가 차서 잘린 경우에는 섞여버림 (100바이트 중 70바이트 보내고 다른패킷 보내고 남은 30보낼 경우) 그래서 위처럼 모아서 차례차례 보내야함, 버퍼에 모아놨다가 차례차례로 보내고 패킷이 하나가 완료가 되면 다음꺼 보내고 하는식으로 -> 결국 성능 저하

두번째는 중간에 데이터가 짤렸다? 그러면 그냥 클라 접속 종료시키는 것도 방법

그런데 충격적이게도 두번째 방법을 써야함 왜냐?

커널의 내부 버퍼가 다 차서 내부 버퍼로 send 데이터를 다 복사 못했다는 뜻임

즉, 운영체제에 할당된 데이터를 다 썼다는 얘기고 그 말은 new 나 말록조차 관리를 못해줌, 결국 게임서버가 메모리를 너무 많이 사용한거임 -> 게임서버가 오동작을 해서 제대로 실행될 가능성이 없다는 뜻임 -> 결국 서버 셧다운이 답임

**<오버랩IO에서 IOCP로 추가할 것>**

다중 접속 관리할 방법 추가 및 패킷 포맷, 프로토콜 정의 해야함

기본 패킷 : 길이 + 타입 + 데이터 순서

그래서 자주 주고받아야 하는 패킷은 사이즈를 작게하고 길드 정보, 파티원 정보 같이 정보량 많은건 패킷 사이즈 크고 한 번씩 주고받게 설정

클라 -> 서버: 로그인 요청, 이동 패킷, (스킬 사용, 아이템 사용은 추가 구현)

서버 -> 클라: 로그인 수락(캐릭터 초기화 같이), 위치 지정, ID 접속 알림, 로그아웃 알림

(추가 요소: 상태 이상, 레벨 업, 마법 스킬 등), 실제 게임은 패킷이 100종류 정도 있음

**<프로토콜 정의>**

어차피 서버랑 클라 둘다 PC이므로 구조체가 편함, 모든 패킷은 chat size랑 char type이 맨앞에 있음, int id는 누가 접속을 했다, 누가 어디로 이동을했다 할 때 누가에 해당됨

근데 만약 숫자로 관리 안하고 캐릭터 이름으로 id를 저장한다 -> 문제가 많아짐

문자열로 되면 길이가 길어져 네트워크 대역폭(bandwidth)이 낭비된다

여기서 대역폭이란 일정 시간 내에 데이터 연결을 통과할 수 있는 정보량의 척도임

Char o\_type을 통해 이게 플레이어인지 npc인지 구분, char name을 통해 이름 표시

#pragma pack(push, 1) 패킷에 빈공간 다 지워버리는 역할, 이걸 안하면 패킷 내부에 변수들이 int형에 맞춰서 데이터 크기 설정해버림

**<패킷 처리 루틴>**

**bool PacketProcess(const unsigned char\* pBuf, int client\_id);**

클라->서버, 서버->클라 패킷이 왔을 때 어떻게 처리할지 정해주는 함수

**<Recv 구현>**

**OverEx** 오버랩 구조체 안에 추가할 것

**unsigned char m\_IOCPbuf[MAX\_BUF\_SIZE];** -> IOCP recv/send 버퍼 추가

**enumOperation m\_Operation;** -> send/recv/accept 구별용

**class ClientInfo** 안에 추가할 것

**unsigned char m\_recv\_packet\_buf[MAX\_PACKET\_SIZE];**

recv 할때 데이터 크기가 커서 중간에 끊길 때 위에 오버랩 구조체 안에 있는 버퍼에 있는 데이터를 보관하기 위한 버퍼

**int m\_prev\_size;** ->그리고 그 버퍼의 크기 저장, 위에랑 얘는 사실상 패킷 재조립용

**<패킷 재조립>**

리시브가 완료되었다 -> 루프 돌면서 패킷 재조립을 함

Start:

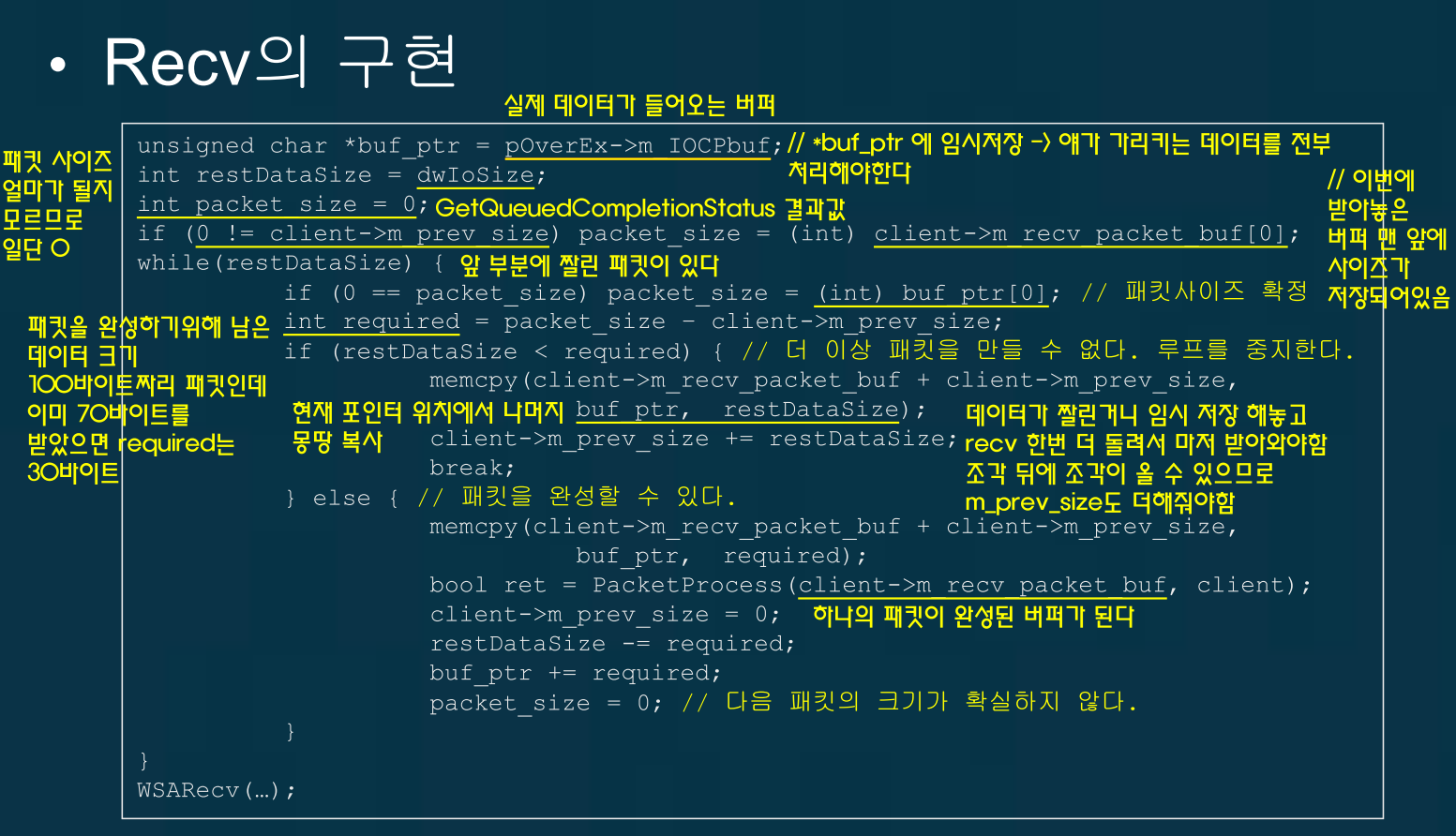
1. 모든 데이터를 처리했으면 End
2. 모든 데이터를 처리하지 않았다 -> 처리하지 않은 남은 데이터로 패킷 완성 가능?

예 : 패킷 버퍼 완성시키고 패킷 처리 함수 (process packet) 함수 호출, 이후 Start

아니요 : 이 경우 남은 데이터가 있는지 확인, 남은 데이터가 있다면 이건 짤린 패킷의 앞부분이라는 뜻, 그럼 패킷 버퍼로 모두 보내버리면 된다, 이후 Start

End:

Recv 호출

**<Recv 패킷 재조립 구현 + AcceptEx 함수>**

**SOCKET clientSocket = accept(listenSocket, (sockaddr\*)&clientAddr, &c\_AddrSize);**

**AcceptEx(listenSocket, clientSocket, buffer, NULL, sizeof(sockaddr\_in) + 16, sizeof(sockaddr\_in) + 16, NULL, over);**

기존의 accept는 리슨 소켓에서 소켓을 받아서 새로운 소켓을 하나 만들어서 리턴

매개변수1 : 리슨소켓

매개변수2 : 그러나 AcceptEx는 새로 만들어서 리턴 안하고 미리 소켓을 만들어서 거기다가 클라를 연결함

매개변수3 : accept를 받으면 연결된 클라 주소를 넣어줄 버퍼

접속을 할때 클라 주소도 필요하지만 클라가 가지고 있는 여러 카드중 어떤 네트워크 카드로 접속했는지 알아야해서 수신, 송신 주소 둘다 필요함

매개변수5,6 : 수신주소, 송신주소 (+16은 여유공간 마련)

매개변수7 : 버퍼에 주소가 몇바이트 들어갔냐, 근데 우리는 이거 안씀, 받는다 치더라도 여기서 처리 안할거임

**<WSASend>**

Send 하는 오버랩 구조체와 버퍼는 끝날 때까지 유지되어야 하고

오버랩 구조체와 버퍼 개수는 미리 할 수 없어서 다이내믹하게 관리해야 함

Send할 데이터가 100바이트인데 70바이트만 보내지고 30바이트가 남는 부분 send가 일어난 경우는, 즉 버퍼가 비워지지 않은 경우는 에러 처리하고 끊어버려야 함

이건 운영체제의 메모리가 꽉 찬 경우라는 거임, 이런 일이 일어나면 이미 답이 없는 경우이므로 이런 일이 안 일어나게끔 예방 해야함 -> send 데이터량 조절해야함

Send를 너무 많이 하거나, 클라로부터 받는 recv 속도가 너무 느리거나

-> 클라에서 보낸 데이터 빨리 처리, send 될 수 있으면 적게

-> 패킷을 자주 안보내고, 보내더라고 데이터량 적게 설계

근데 아무리 그렇게해도 최소한으로 필요한 양의 패킷이 있을거임

아무리 줄여도 recv가 쌓이고 send 공간이 없어서 메모리가 모자란다 -> 너무많은 동접

프로그램이 감당 가능한 동접을 넘어섰다는거임 -> 동접을 낮춰야함

**<Send 완료 구현>**

**if (dwIoSize < pOverEx->m\_IOCPbuf[0])**

**Disconnect(client);**

**delete pOverlappedEx;**

send는 무조건 패킷 통으로 보내야하는데 **dwIoSize < pOverEx->m\_IOCPbuf[0]**는 넘버 오브 바이트가 패킷 크기보다 작은 경우이므로 문제가 있다는거임 -> 연결 끊기