# $\mu PandOS$

## Contents

1	Cos	s'è?	2
<b>2</b>	Fase	se 1 - Definizione operazioni su liste di pcb e messaggi	2
	2.1	Obiettivi	2
	2.2	Prototipi delle funzioni	2
		2.2.1 Allocazione e deallocazione dei PCB:	2
		2.2.2 PCB Queue:	2
		2.2.3 PCB Trees:	3
		2.2.4 Allocazione e deallocazione dei messaggi:	3
		2.2.5 Message	3
•	-		
3		se 2 - Definizione del Nucleo, Scheduler, SSI, Interrupt ed eccezioni	<b>4</b>
	3.1	Utility	4
	3.2	·	4
	3.2	Inizializzazione nucleo	
		0 0	4
		· · · · · · · · · · · · · · · · · · ·	4
		· · · · · · · · · · · · · · · · · · ·	4
	2.2	3.2.4 Processi SSI e Test	5
	3.3	Scheduler	5
	3.4	SSI	5
		3.4.1 SSIRequest	6
4	Ges	store delle eccezioni	7
	4.1	Interrupt Handler	7
		4.1.1 Gestione Interrupt Processor Local Timer	7
		4.1.2 Gestione Interrupt Interval Timer	7
		4.1.3 Gestione Interrupt Device	7
	4.2	Pass Up or Die	8
	4.3	Eccezioni causate da SYSCALL	9
5	Fase	se 3	10
6	$\mathbf{Cre}$	<del></del>	11
	6.1	Github	11
	6.2	Autori	11

#### 1 Cos'è?

μPandOS è un sistema operativo microkernel sviluppato per fini didattici; in particolare questa implementazione è fatta al fine di svolgere un progetto propedeutico all'esame per il corso 08574 - Sistemi Operativi (anno accademico 2023/24) per l'università di Bologna.

## 2 Fase 1 - Definizione operazioni su liste di pcb e messaggi

#### 2.1 Obiettivi

In questa fase andremo a scrivere le basi per quanto riguarda questo progetto, ovvero definiremo i metodi di due strutture fondamentali per quanto riguarda PandOS, ossia i messaggi e i PCB

#### 2.2 Prototipi delle funzioni

#### 2.2.1 Allocazione e deallocazione dei PCB:

- void initPcbs(): tramite la funzione freePcb, vengono aggiunti in coda gli elementi della pcbTable (da 1 a MAXPROC) nella lista dei processi liberi;
- void freePcb(pcb\_t \*p): mette l'elemento puntato da p nella lista dei processi liberi;
- pcb\_t \*allocPcb(): rimuove il primo elemento dei processi liberi, inizializza tutti i campi e ritorna un puntatore ad esso.

#### 2.2.2 PCB Queue:

- void mkEmptyProcQ(struct list\_head \*head): inizializza una variabile come puntatore alla testa della coda dei processi;
- int emptyProcQ(struct list\_head \*head): se la coda la cui testa è puntata da head è vuota ritorna TRUE, altrimenti FALSE;
- void insertProcQ(struct list\_head \*head, pcb\_t \*p): inserisce il PCB puntato da p in fondo alla coda dei processi puntata da \*head;
- pcb\_t \*headProcQ(struct list\_head \*head): ritorna NULL se la coda dei processi è vuota, altrimenti il PCB in testa;
- pcb\_t \*removeProcQ(struct list\_head \*head): rimuove la testa della coda dei processi puntata da \*head e ritorna un puntatore dell'elemento in questione; se la lista è vuota ritorna NULL;
- pcb\_t \*outProcQ(struct list\_head \*head, pcb\_t \*p): cerca mediante un for\_each il PCB p nella lista puntata da head e lo rimuove; se lo trova ritorna p stesso, altrimenti NULL.

#### 2.2.3 PCB Trees:

- int emptyChild(pcb\_t \*p): ritorna l'esito della chiamata alla funzione list\_empty, alla quale viene passato come parametro l'indirizzo del list\_head p\_child di p;
- void insertChild(pcb\_t \*prnt, pcb\_t \*p): si assegna prnt al puntatore p\_parent di p. Dopo si aggiunge p alla lista dei fratelli, tramite list\_add (se non ci sono altri figli) e list\_add\_tail (per rispettare la FIFOness), alle quali viene passato come parametro gli indirizzi del list\_head p\_sib di p e del list\_head p\_child di prnt (p diventa fratello dei figli di prnt e quindi figlio di prnt).
- pcb\_t \*removeChild(pcb\_t \*p): il controllo sulla presenza o meno di figli avviene con la funzione emptyChild. Se ci sono figli, si sceglie il primo figlio tramite la macro container\_of, chiamata sull'elemento successivo al list\_head p\_child. In seguito il figlio viene eliminato tramite la funzione list\_del e viene troncato il legame con il padre, assegnando il valore NULL al puntatore p\_parent del figlio.
- pcb\_t \*outChild(pcb\_t \*p): se p ha un padre, rimuovo p dalla lista dei suoi fratelli chiamando list\_del a cui passo come parametro l'indirizzo di p\_sib di p, in seguito rimuovo il legame con il padre assegnando NULL al puntatore p\_parent di p.

#### 2.2.4 Allocazione e deallocazione dei messaggi:

- void freeMsg(msg\_t \*m): Inserisce l'elemento puntato da m in testa alla lista dei messaggi.
- msg\_t \*allocMsg(): Ritorna NULL se la lista dei messaggi è vuota. Altrimenti rimuove un elemento dalla testa, imposta a 0 la variabile m\_payload di ogni messaggio presente nell'array msgTable e ritorna un puntatore all'elemento rimosso.
- void initMsgs(): Inserisce gli elementi presenti nell'array msgTable in coda alla lista dei messaggi.

#### 2.2.5 Message

- void mkEmptyMessageQ(struct list\_head \*head): Inizializza una una lista di messaggi vuota.
- int emptyMessageQ(struct list\_head \*head): Ritorna 1 se la lista puntata da head è vuota, altrimenti 0.
- void insertMessage(struct list\_head \*head, msg\_t \*m): Inserisce il messaggio puntato da m in coda alla lista puntata da head.
- void pushMessage(struct list\_head \*head, msg\_t \*m): Inserisce il messaggio puntato da m in testa alla lista puntata da head.
- msg\_t \*popMessage(struct list\_head \*head, pcb\_t \*p\_ptr): Rimuove il primo messaggio trovato nella lista puntata da head che è stato inviato dal thread p\_ptr.

  Se p\_ptr è NULL, ritorna il primo messaggio in coda.

  Se head è vuota o se non viene trovato alcun elemento mandato dal thread p\_ptr, ritorna null.
- msg\_t \*headMessage(struct list\_head \*head): Se la lista puntata da head è vuota ritorna NULL, altrimenti ritorna il messaggio in testa ad essa.

## 3 Fase 2 - Definizione del Nucleo, Scheduler, SSI, Interrupt ed eccezioni

Di seguito sono riportate le scelte progettuali per quanto riguarda i moduli sviluppati:

#### 3.1 Utility

#### **3.1.1** timer.c

In questo modulo abbiamo delle funzioni/procedure ausiliarie richiamate degli altri moduli per la gestione dei vari timer:

- unsigned int getTOD(): ritorna il valore del time of day clock, che viene nel nostro caso salvato nella variabile globale start: utilizzata per il calcolo del CPU time.
- void updateCPUtime(pcb\_t \*p): chiama la funzione qui sopra descritta per aggiornare il valore del campo p\_time del processo passato alla funzione.
- void setIntervalTimer(unsigned int t): funzione che imposta il valore dell'interval timer.
- void setPLT(unsigned int t): funzione che imposta il valore del processor local timer.
- unsigned int getPLT(): funzione che permette di ottenere il valore del processor local timer.

#### 3.2 Inizializzazione nucleo

#### 3.2.1 Dichiarazione e inizializazione variabili globali

Nel modulo initial.c viene implementato il main(), la dichiarazione delle variabili globali:

- int process\_count ossia il contatore dei processi attivi;
- int soft\_blocked\_count ossia il contatore dei processi bloccati;
- int start ...
- int pid\_counter, usato per assegnare in maniera sequenziale i PID ai processi man mano che vengono creati;
- pcb\_t \*current\_process ossia il puntatore al PCB del processo corrente;
- pcb\_t \*ssi\_pcb, che è il puntatore al PCB del SSI;

#### 3.2.2 Dichiarazione e inizializazione strutture dati

Vengono inoltre implementate le strutture dati principali:

- attraverso le funzioni initPcbs() e initMsgs() vengono inizializzate le strutture della fase 1;
- Ready\_Queue, ossia la lista dei proessi pronti ad essere eseguiti;
- 8 liste per i processi bloccati in attesa dei device o per il terminale (una per input e una per output);
- void initPassupVector() è una procedura che viene richiamata per definire il pass up vector, ossia è la struttura dati a livello hardware che indica a quale funzione passare il controllo quando si verifica un interrupt.

#### 3.2.3 Interval timer

Viene caricato l'interval timer a 100 ms attraverso la chiamata alla procedura ausiliaria setIntervalTimer(PSECOND) definita in timers.c

#### 3.2.4 Processi SSI e Test

Infine, prima di richiamare lo Scheduler, attraverso la procedura void initFirstProcesses() vengono inseriti nella Ready Queue i processi del SSI e del test. Questi avranno lo status settato in modo da avere la maschera dell'interrupt abilitata, l'interval timer abilitato e che siano in modalità kernel. Avranno rispettivamente pid 1 e 2.

#### 3.3 Scheduler

Lo Scheduler è il componente che gestisce la coda dei processi pronti ad essere eseguiti (**Ready Queue**); la procedura principale che svolge tutto ciò è void scheduler(); questa parte con un controllo iniziale sulla Ready Queue vedendo se è vuota (con emptyProcQ(&Ready\_Queue)):

- se non è vuota prendo il processo che deve essere preso in carico dalla CPU (current\_process) con la funzione removeProcQ(&Ready\_Queue), setto il Timer attraverso la funzione setPLT() a 5 ms (con la costante TIMESLICE) per implementare il Round Robin, e infine viene caricato lo stato del processo corrente nel processore (con LDST());
- altrimenti (se vuota), si effettua la Deadlock detection; in particolare può decidere se effettuare un HALT() quando non ci sono più processi da eseguire; se ci sono altri PCB entrerà in WAIT(); se la ready queue è vuota e ci sono processi bloccati si entra in deadlock invocando PANIC() fermando così l'esecuzione;

#### 3.4 SSI

Essendo che µPandOS è un microkernel, le uniche syscall implementate sono la Send e la Receive; queste vengono usate dai processi per chiedere al processo SSI risorse; quanto detto è implementato nell'apposito modulo ssi.c, in particolare nella funzione SSILoop(), che implementa il polling del processo SSI: questa è eternamente in attesa di ricevere un messaggio da un qualsiasi processo che necessità una risorsa, prova a soddisfarlo attraverso l'apposita funzione unsigned int SSIRequest(pcb\_t\*sender, ssi\_payload\_t \*payload) e se riesce viene inviato un riscontro al processo che ha effettuato la richiesta tramite la syscall send. Di seguito si forniranno maggiori dettagli riguardo quest'ultima funzione;

#### 3.4.1 SSIRequest

All'interno di questa funzione vengono analizzati i parametri pcb\_t\* sender, ssi\_payload\_t \*payload che contengono rispettivamente il processo che ha richiesto il servizio e il messaggio mandato col servizio richiesto; in particolare nel messaggio è determinante il service\_code, che stabilisce l'oggetto della richiesta del sender. Attraverso l'utilizzo di uno switch su quest'ultimo parametro, la ssi fornisce il servizio richiesto. Si distinguono in particolare i casi:

- 1 (CREATEPROCESS): viene richiesta la creazione di un processo; questa richiesta viene soddisfatta solo se c'è spazio nella tabella dei processi liberi; in caso affermativo viene invocata la funzione ssi\_new\_process() con parametri il sender, che fungerà da parent, e un puntatore a una variabile di tipo ssi\_create\_process, contenete i dati necessari alla creazione del nuovo processo; Nella funzione ssi\_new\_process viene chiamata allocPcb() per l'allocazione in memoria del nuovo processo e successivamente questo viene inserito nelle ready queue e nella lista child del pcb sender.
- 2 (TERMPROCESS), viene chiamata la procedura ssi\_terminate\_process() passando come parametro il processo da terminare. In questa funzione avviene ricorsivamente la terminazione di tutti i processi figli del sender e dei realtivi fratelli.
- 3 (DOIO), viene chiamata la procedura ssi\_doio con parametri il puntatore al pcb del sender e un puntatore a una variabile di tipo ssi\_do\_io\_t contenente i dati necessari a tale richiesta. Viene successivamente chiamata la procedura addrToDevice() che, dato l'indirizzo del device passato con la richiesta di DOIO, ne determina i rispettivi numero di device e riga, per poi chiamare la funzione blockProcessOnDevice. Quest'ultima, passati come argomenti il pcb del sender, il numero del device e un intero che indica se si sta facendo una lettura o una scrittura, rimuove il sender dalla ready queue e lo inserisce nella lista di processi bloccati del relativo device. La funzione addrToDevice() invidua la corretta riga e il corretto numero di device confrontando il command\_address passato come argomento con tutti i campi di DEV<sub>R</sub>EG<sub>A</sub>DDR[][].Perunamaggioreef ficienza, ilcommand\_addressvieneprimaconfrontatoconicampidiDEV<sub>R</sub>E
- 4 (GETTIME), viene restituito come unsigned int un puntatore alla variabile p\_time del processo sender.
- 5 (CLOCKWAIT), viene chiamata la funzione ssi\_clockwait con argomento il puntatore alla struct pcb\_t del processo sender. Quest'ultima inserisce il sender nella lista dei processi bloccati da pseudo clock. La procedura pseudoClockInterruptHandler() definita in interrupts.c si occupa della gestione di questa lista.
- 6 (GETSUPPORTPTR), viene restituito come unsigned int un puntatore alla variabile p\_supportstruct del processo sender.
- 7 (GETPROCESSID), viene invocata la funzione int ssi\_getprocessid. Quest'ultima prende come parametri il puntatore al pcb del sender e un puntatore generico arg, e ritorna il pid del sender se l'argomento è NULL, altrimenti il pid del processo padre del chiamante.
- Se il service code non contiene nessuno dei seguenti codici viene terminato il sender con la funzione ssi\_terminate\_process().

In alcune delle funzioni menzionate vengono inoltre modificati i valori di variabili globali quali:

- soft\_blocked\_count : viene decrementata in ssi\_terminate\_process() se il processo viene rimosso da una delle liste di processi bloccati dei device, o incrementata in ssi\_clockwait() quando viene aggiunto.
- process\_count : viene incrementata alla creazione di un nuovo processo in ssi\_create\_process e decrementata quando un processo viene terminato in ssi\_terminate\_process().

#### 4 Gestore delle eccezioni

La funzione che si occupa della gestione delle eccezioni è la funzione void exceptionHandler() dichiarata nel file phase2/include/exceptions.h e definita nel file phase2/exceptions.c. Questa funzione salva lo stato al tempo dell'eccezione dalla BIOSDATAPAGE ed in seguito trova il codice dell'eccezione eseguendo operazioni di manipolazione dei bit sul registro cause, ottenuto con la funzione getCAUSE. In particolare si esegue l'operazione cause & GETEXECCODE che permette di mantenere solo i bit che definiscono il codice dell'eccezione, i quali vengono shiftati a destra di 2 posizioni (costante CAUSESHIFT).

#### 4.1 Interrupt Handler

Nel caso il codice dell'eccezione abbia valore 0 (costante IOINTERRUPTS) viene invocata la funzione per la gestione degli interrupt void interruptHandler(int cause, state\_t\* exception\_state). Qui viene utilizzata la macro CAUSE\_IP\_GET(cause, line), grazie alla quale, passando il cause register e il valore di una linea di interrupt, è possibile sapere se c'è un interrupt su quella linea. Il controllo viene fatto per tutte le linee, seguendo l'ordine di priorità che va dall'interrupt causato dal processor local timer, all'interrupt causato da un dispositivo terminale. In base alla linea su cui avviene l'interrupt viene invocato un'opportuna funzione per la gestione di quello specifico interrupt.

#### 4.1.1 Gestione Interrupt Processor Local Timer

L'interrupt causato dal processor local timer si verifica quando il tempo nella CPU per il processo corrente si esaurisce. Per un'opportuna gestione di questo interrupt usiamo la funzione static void localTimerInterruptHandler(state\_t \*exception\_state). In questa routine viene riconosciuto l'interrupt con la chiamata setPLT(-1), in seguito si aggiorna il CPU time del processo corrente, si copia lo stato dell'eccezione nello stato del processo corrente, il quale infine viene inserito sulla ready queue. Dopo queste operazioni viene chiamato lo scheduler.

#### 4.1.2 Gestione Interrupt Interval Timer

In questo caso l'ACK dell'interrupt è eseguito con la chiamata setIntervalTimer(PSECOND). Dopodiché avvienelo sblocco di tutti i processi che erano in attesa dell'interrupt, rimuovendo ciascuno di essi dalla lista Locked\_pseudo\_clock, inserendoli sulla ready queue, dopo aver inviato loro un messaggio che consentirà ai processi interessati di sbloccarsi, quando rieseguiranno la SYS2 su cui si erano precedentemente bloccati. Ogni volta che viene rimossoun processo dalla lista dei processi in attesa dello pseudoclock tick, viene decrementata la variabile globale soft\_blocked\_count. Infine se il processo corrente è diverso da NULL, si esegue una LDST con lo stato dell'eccezione, altrimente viene chiamato lo scheduler.

#### 4.1.3 Gestione Interrupt Device

La gestione degli interrupt legati a tutti gli altri device viene affidata alla funzione static void deviceInterruptHandler(int line, int cause, state\_t \*exception\_state), la quale ricava la bitmap degli interrupt per i dispositivi della linea d'interesse. Questo viene realizzato accedendo all'area di memoria riservata ai device, all'inidirizzo BUS\_REG\_RAM\_BASE. In seguito si esegue l'and sui bit della bitmap con le costanti DEVXON con  $X \in \{0, \dots, 7\}$ , con questa operazione si ottiene il numero del device sulla line che ha causato l'interrupt, per l'ordine con cui queste operazioni sono effettuate,il numero del device calcolato sarà sempre quello a priorità maggiore. Calcolato il numero, data la linea si sblocca il processo dalla lista associata alla linea cercandolo tramite il device number, grazie al campo aggiuntivo dev\_no che abbiamo messo ai pcb. Questo campo viene settato dall'ssi quando viene bloccato il processo in attesa di interrupt durante il servizio DOIO. In caso di interrupt causato da un dispositivo terminale verifichiamo che il codice dell'operazione di trasmissione sia uguale a 5 (interrupt in attesa di essere riconosciuto), se così è allora significa che l'operazione è un'operazione di trasmissione di un carattere, altrimenti si tratta di un'operazione di ricezione. Per accedere al registro

```
del device usiamo la macro DEV_REG_ADDR nel modo seguente:
dtpreg_t *device_register = (dtpreg_t *)DEV_REG_ADDR(line, device_number);
In caso di dispositivi terminali, l'operazione è analoga:
termreg_t *device_register = (termreg_t *)DEV_REG_ADDR(line, device_number);
L'operazione di riconoscimento dell'interrupt avviene con l'istruzione
device_register->command = ACK;,
per i terminali a seconda di quale subdevice ha generato l'interrupt:
device_register->transm_command = ACK;
oppure
device_register->recv_command = ACK;
L'accesso allo status avviene con l'operazione seguente: device_register->status.
Per i terminali:
device_register->recv_status.
device_register->transm_status.
```

Infine se il processo sbloccato è diverso da NULL, si mette lo status nel suo registro v0, gli viene inviato un messaggio avente la ssi come mittente e come payload lo status del device, si inserisce il processo sulla ready queue e si diminuisce di un'unità soft\_blocked\_count. In seguito se il processo corrente è diverso da NULL si chiama lo scheduler, altrimenti si esegue una LDST dello stato ottenuto dalla BIOSDATAPAGE.

#### 4.2 Pass Up or Die

Tramite Pass Up or Die il kernel gestisce tutte le eccezione che non sono syscall o interrupt, abbiamo quindi implementato un'apposita funzione static void passUpOrDie(int i, state\_t \*exception\_state) che controlla che la support struct del processo corrente sia diversa da NULL, se ciò è vero allora si salva lo stato dell'eccezione nello stato corretto della struttura di supporto:

saveState(&(current\_process->p\_supportStruct->sup\_exceptState[i]), exception\_state);
ed in seguito si esegue LDCTX, passando come parametri i valori del giusto constesto della struttura
di supporto:

```
LDCXT(current_process->p_supportStruct->sup_exceptContext[i].stackPtr,
current_process->p_supportStruct->sup_exceptContext[i].status,
current_process->p_supportStruct->sup_exceptContext[i].pc
):
```

L'indice i, parametro della funzione, può assumere due valori a seconda del tipo di eccezione:

- GENERALEXCEPT: per trap generiche, con codici 4...7 e 9...12;
- PGFAULTEXCEPT: per eccezioni TLB, con codici 1...3;

In caso di puntatore nullo, chiamiamo la funzione per la terminazione dei processi, che viene utilizzata dall'ssi per fornire il servizio TERMINATEPROCESS.

#### 4.3 Eccezioni causate da SYSCALL

L'eccezioni da SYSCALL sono quelle il cui codice ha valore SYSEXCEPTION. Quando queste si verificano, viene chiamata la funzione static void syscallexceptionHandler(state\_t \*exception\_state). In primis la funzione controlla il valore del bit KUc, se questo è 1 (user-mode) viene generata una trap generica, in cui il valore del codice dell'eccezione viene settato a EXC\_RI (istruzione riservata). Se il processo è in kernel-mode allora si continua e si controlla il registro a0 per capire quale servizio è stato richiesto:

- reg\_a0 = SENDMESSAGE: il processo corrente vuole eseguire una SYS1, dal registro al viene estratto l'indirizzo del PCB al quale si vuole inviare un messaggio, si controlla se il destinatario è nella ready queue o se è il processo corrente, in entrambi questi casi si inserisce il messaggio nell'inbox del destinatario. Se il destinatario è sulla lista dei processi liberi, allora esso è inesistente e si aggiorna solamente il registro v0 al valore DEST\_NOT\_EXISTS. Se il processo destinatario non si trova sulla ready e queue e nemmeno nella lista dei processi liberi, allora dopo l'invio del messaggio, lo si inserisce nella ready queue (se il messaggio inviato non è quello per cui era in attesa, il processo si bloccherà di nuovo). La funzione che invia i messaggi è int send(pcb\_t\*sender, pcb\_t\*dest, unsigned int payload). In questo caso come payload viene passato il valore del registro a2. Infine si incrementa il PC e si esegue LDST(exception\_state).
- reg\_a0 = RECEIVEMESSAGE: il processo corrente vuole eseguire una SYS2, dal registro a1 si accede all'indirizzo del PCB dal quale il processo richiedente vuole ricevere un messaggio, questo valore viene passato, insieme alla inbox del processo corrente, alla funzione popMessage, la quale restituirà il messaggio o NULL se non lo trova. Nel primo caso la SYS2 non è bloccante, quindi si assegna al registro v0 l'indirizzo di chi ha inviato il messaggio, e si mette nell'aria di memoria puntata dal registro a2 il payload del messaggio. Si incrementa il PC e si esegue LDST(exception\_state), poiché non è una syscall bloccante. Se invece la popMessage ha restituito NULL, allora la syscall sarà bloccante, lo stato dell'eccezione viene salvato nello stato del processo corrente, si aggiorna il CPU time del processo corrente ed infine si chiama lo scheduler.
- reg\_a0 >= 1: viene generata una trap generica.

5 Fase 3 - ...

## 6 Crediti

## 6.1 Github

Il sorgente del progetto è reperibile nella seguente repository su Github.

### 6.2 Autori

- Fiorellino Andrea, matricola: 0001089150, andrea.fiorellino@studio.unibo.it
- $\bullet\,$  Po Leonardo, matricola: 0001069156, leonardo.po@studio.unibo.it
- $\bullet$  Silvestri Luca, matricola: 0001080369, luca.silvestri<br/>9@studio.unibo.it