

Intertask Kommunikation in Echtzeitbetriebssystemen

Stefan Lindörfer

stefan.lindoerfer@stud-mail.uni-wuerzburg.de

Universität Würzburg

Institut für Informatik

Lehrstuhl V - Technische Informatik

Seminar: Embedded Systems

Prof. Dr. Reiner Kolla

Wintersemester 2020/21

Abstract Die Informationsübermittlung zwischen Tasks in Echtzeitbetriebssystemen ist von fundamentaler Bedeutung und kann je nach Anforderungsszenario auf verschiedene Arten erfolgen. In der vorliegenden Arbeit wird zunächst ein Überblick über die verschiedenen Möglichkeiten gegeben sowie darauffolgend jeweils vertieft auf die Mechaniken eingegangen und durch Beispiele demonstriert.

1 Einführung

Echtzeitbetriebssysteme (auch **Real Time Operating Systems**, kurz RTOS genannt) sind aus der Informatik nicht mehr wegzudenken und begegnen uns oft unbewusst im Alltag in zahlreichen verschiedenen Anwendungsbereichen. So werden sie etwa in modernen Automobilen als Betriebssysteme eingesetzt, zur Steuerung und Regelung industrieller Anlagen verwendet oder im Verkehrswesen (z.B. Schienenverkehr- oder Ampelsteuerungen) genutzt [vgl. 1, S. 157]. Auch als Betriebssysteme von Satelliten in der Raumfahrt sowie für den Einsatz in Flugzeugen sind sie geeignet [2].

Die Besonderheit dieser Art von Betriebssystemen ist die Fähigkeit, Echtzeit-Anforderungen umsetzen zu können. Das bedeutet (anders als bei Nicht-RTOS) als Softwareentwickler die Zusage seitens des Betriebssystems zu besitzen, dass eine bestimmte Aufgabe innerhalb eines vordefinierten Zeitfensters entweder ausgeführt und abgeschlossen (oder unterbrochen) wird. Betriebssysteme anderer Kategorien führen ihre Aufgaben meist schnellstmöglich aus und passen kein Zeitfenster ab, garantieren also keine Echtzeit, sondern setzen auf größtmögliche Performanz.

Aufgaben in Echtzeitbetriebssystemen sind in sogenannte Tasks (oder auch Threads) unterteilt, die vom Betriebssystem gekapselt voneinander verwaltet werden. Abhängig vom jeweiligen Szenario können Tasks nicht ausschließlich getrennt voneinander operieren und ihren jeweiligen Aufgaben nachgehen, sondern ein Informationsaustausch zwischen ihnen ist erforderlich: Um etwa einen Ablauf abzubilden, bei dem zuerst Task A und anschließend Task B ausgeführt werden soll, muss eine Information von A nach B übermittelt werden können, die den entsprechenden Prozessfortschritt anzeigt sodass B beginnen kann. Oder auch wenn Daten eines Tasks von einem anderen zur (synchronisierten) Weiterverarbeitung benötigt werden, ist ein entsprechender Kommunikationskanal erforderlich um die Daten zu transferieren.

Dieser gesamte interne Informationsaustausch wird als Intertask-Kommunikation bezeichnet. Es existieren mehrere allgemeine Möglichkeiten und Techniken, Informationen und Daten zwischen Tasks auszutauschen, abhängig vom jeweiligen Szenario und den Anwendungsanforderungen. Diese Techniken werden im Folgenden aufgezeigt und nacheinander erläutert.

1.1 Überblick

Grundsätzlich unterteilt man Intertask-Kommunikation in drei Kategorien [vgl. 3, S. 79]:

1. **Synchronisation und Koordination von Tasks ohne Datentransfer**
2. **Datentransfer zwischen Tasks ohne Synchronisation**
3. **Datentransfer zwischen Tasks mit Synchronisation**

Hier
eventuell
die Re-
ferenzen
der je-
weiligen
Kapi-
tel ab-
bilden.

Hier fehlt
noch eine
Quelle

Punkt 1 enthält anders als (2) und (3) keinen Datentransfer und unterscheidet sich von diesen - wie bereits in 1 ausgeführt - dadurch, dass lediglich eine Information ausgetauscht bzw. vom Empfänger abgefragt wird, um Tasks zu synchronisieren oder einen Ablauf umzusetzen. Während bei (2) und (3) ein Datentransfer insofern stattfindet, als das Daten ausgetauscht und vom Empfänger für die Weiterverarbeitung genutzt werden, sie also nicht für eine Ablaufsteuerung verwendet werden.

1.2 Begriffsgrundlage

Für eine genauere Betrachtung der einzelnen Kategorien aus 1.1, müssen zunächst die beiden Begriffe Koordination und Synchronisation definiert und pragmatisch voneinander abgegrenzt werden:

- **Koordination:** „Das Integrieren und Anpassen (einer Reihe von Teilen oder Prozessen), um eine reibungslose Beziehung zueinander herzustellen.“ [vgl. 3, S. 80]
- **Synchronisation:** „Etwas verursachen, bewegen oder ausführen, genau zur exakten Zeit.“ [vgl. 3, S. 80]

anderes
Wort!

Es fällt auf, dass die Definition der Koordination keinen Bezug zur Zeit beinhaltet. Der wesentliche Unterschied zwischen Koordinierung und Synchronisierung ist somit der Zeit-Faktor. Während mit einer Koordination ein theoretisch zeitunabhängiger, sequentieller Ablauf von Tasks angestrebt wird, meint Synchronisation dagegen das zeitliche Abgleichen von Vorgängen und legt damit verstärkt Fokus auf die Kerneigenschaft von Echtzeitbetriebssystemen. Dennoch haben beide Begriffe ihre klare Daseinsberechtigung und ihren verhältnismäßigen Anwendungsspielraum.

2 Task-Interaktion ohne Datentransfer

Müssen, wie schon in 1.1 erwähnt, keine Daten im eigentlichen Sinne zwischen Tasks transferiert werden, sondern nur ein Arbeitsablauf gesteuert werden, spricht man von Task-Interaktion ohne Datentransfer. Dies kann sowohl mit dem Ziel einer Task-Synchronisation durchgeführt werden als auch ohne den bereits festgestellten Zeit-Faktor mittels Koordination, siehe 1.2. Demzufolge wird bei den Möglichkeiten dieser Kategorie auch in diese beiden Fälle unterschieden. Tabelle 1 zeigt diese Unterscheidung auf und gibt gleichzeitig einen Überblick über die jeweils zu verwendenden Konstrukte dieser Kategorie. Für das Ziel einer Koordination, werden Condition Flags verwen-

Koordination	Synchronisation	
Condition Flags	Event Flags	Signale
<u>Operationen:</u>	<u>Operationen:</u>	<u>Operationen:</u>
Set	Set	Wait
Clear	Clear	Send
Check	Check	Check

Tabelle 1: Koordinierungs- und Synchronisationskonstrukte [vgl. 3, S. 82]

det. Ein Flag ist ein Statusindikator, der einen bestimmten Zustand anzeigt. Ist Synchronisierung erforderlich, dass also der gesteuerte Prozess zeitkritisch auszuführen ist, dann können - je nach Anwendungsfall - Event Flags oder Signale verwendet werden. Alle drei Konstrukte und deren Operationen werden im Folgenden genauer erläutert.

2.1 Task-Koordinierung mit Condition Flags

Die einfachste Möglichkeit der Koordination ist das Condition Flag. Tabelle 1 zeigt die auf Condition Flags anwendbaren Operationen **Set** (setzen), **Clear** (zurücksetzen) und **Check** (überprüfen).

Abbildung 1 zeigt exemplarisch, wie Condition Flags verwendet werden: Task A übernimmt in diesem Fall die Steuerung des Ablaufs, während Task B darauf wartet, ausgeführt zu werden

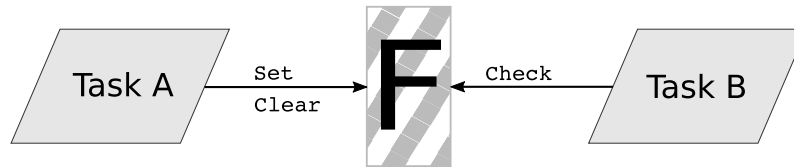


Abbildung 1: Einfache Benutzung von Condition Flags [vgl. 3, S. 83]

und fortlaufend überprüft (**Check**), ob das Flag gesetzt wurde. Im einfachsten Fall wird dafür eine globale Boolean-Variable eingesetzt, bei dieser *true* den Zustand **Set** und *false* den Zustand **Clear** repräsentieren kann. Ist es erforderlich, mehrere Zustände einzusetzen, können Aufzählungstypen (enums) verwendet werden. Diese bietet auch den Vorteil der besseren Lesbarkeit des Quellcodes, da sofort eindeutig verifizierbar ist, welchen Flag-Zustand **Set** bzw. **Clear** darstellt. Für die meisten Anforderungen ist dieses Vorgehen absolut ausreichend. [vgl. 3, S. 84]

In kritischeren Situationen jedoch, z.B. bedingt durch hohe Zugriffsraten auf ein Flag (es können sich Lese- oder Schreibfehlern ergeben) oder wenn ein höheres Maß an Ausfallsicherheit gewünscht ist, empfiehlt sich eine Mehrfachabsicherung wie sie Abbildung 2 demonstriert:

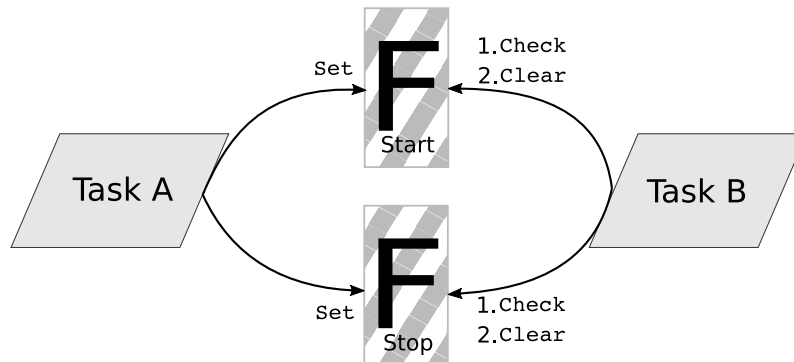


Abbildung 2: Verbesserte Benutzung von Flags zur Koordination [vgl. 3, S. 84]

Für jeden Zustand, der gesetzt werden kann (hier: Zwei), existiert ein eigenes Flag. Task A übernimmt in diesem Fall ausschließlich das Setzen der Zustände, führt dies jedoch nur durch, falls das entsprechende Flag vorher von Task B zurückgesetzt wurde [vgl. 3, S. 85]. Task B prüft jetzt zusätzlich den Zustand aller relevanten Flags bevor auf einen **Set**-Zustand reagiert wird [vgl. 3, S. 85]. Wenn Unstimmigkeiten auftreten, weil z.B. das Prüfergebnis keinen eindeutig definierten Gesamtzustand ausweist, kann ein Fehlverhalten rechtzeitig abgefangen werden, anders als dies bei Abbildung 1 möglich ist. Nach einem erfolgreichem **Check** wird das ausgeführte Flag wieder von B zurückgesetzt. Umgesetzt wird dies beispielsweise wieder mittels enums:

```
typedef enum {StartSet , StartClear} StartFlag;  
typedef enum {StopSet , StopClear} StopFlag;
```

Diese Herangehensweise führt, wie bereits ausgeführt, zu einer sichereren und zuverlässigeren Abwicklung.

Eine weitere Möglichkeit der Organisation mit Condition Flags sind Flag Gruppen. Diese bieten sich besonders dann an, wenn viele Zustände abgebildet werden sollen, die zudem logisch miteinander in Verbindung stehen können. Wenn beispielsweise etwa Zustand A nicht gleichzeitig mit Zustand B gesetzt sein darf. Dabei werden einzelne Flags auf Wortbreite zusammengefasst (auch eine geringere Breite ist möglich, z.B. durch eine Variable), in der jedes Bit ein Flag repräsentiert. Abbildung 2 zeigt dieses Konzept auf einer 2 Byte Variablen mit je 8 Bits.

Hier vielleicht ein neues Unterkapitel schaffen?

Wird der Wert der Variablen gelesen und als Integer (Ganzzahl) interpretiert (Abbildung 2 - Wert: 38806), kann er zudem bequem und einfach angepasst werden. Um etwa mehrere Flags, also

Bit	15	14	13	12	11	10	9	8	7	6	5	4	3	2	1	0
Wert	1	0	0	1	0	1	1	1	1	0	0	1	0	1	1	0

Tabelle 2: Condition Flag Gruppe [vgl. 3, S. 85]

Einzelbits, gleichzeitig zu setzen oder rückzusetzen, kann Bitmanipulation verwendet werden. Auch die Anwendung logischer Operatoren (*AND*, *OR*, *XOR*, *NOT*, *XNOR*) ist auf die einzelnen Bits möglich. Mit solchen Flag Gruppen kann daher beispielsweise leicht die logische Aussage

$$((([15.] \wedge ([13.] \vee \neg [11.])) \oplus [5.]) \equiv 1) \quad (1)$$

implementiert und evaluiert werden, indem eine Bitmaske, die die relevanten Bits (hier: 15, 13, 11 und 5) markiert, mit dem *AND*-Operator auf den Variableninhalt angewendet wird:

$$\begin{array}{r} 1001011110010110 \quad (38806) \\ \& 1010100000100000 \quad (43040) \\ \hline 1000000000000000 \quad (32768) \end{array}$$

Das Ergebnis enthält für unmarkierte Bits stets eine 0 und für markierte eine 1, sofern das jeweilige Bit der Variablen gesetzt war, ansonsten ebenfalls eine 0 und kann damit komfortabel für weitere Vergleiche herangezogen werden. Beispielsweise erfüllt (die binäre Schreibweise von) 32768 bereits die oben angegebene Gleichung (1). Für ausführlichere Informationen zu diesem Thema sei auf die entsprechende Fachliteratur verwiesen.

Zusammengefasst sind Condition-Flags ein gutes Konstrukt um einen Ablauf von Tasks zu koordinieren und es bieten sich für verschiedenste Anforderungsprofile einfache und trotzdem leistungsstarke Lösungen an, um diesen Ablauf zu realisieren und die Kommunikation sicher und zuverlässig zu gestalten.

2.2 Task-Synchronisation über Event Flags

Event Flags übernehmen viele Eigenschaften von Condition Flags so wie sie in 2.1 vorgestellt wurden. Auf sie sind wie in Tabelle 1 dargestellt, auch die gleichen Operationen anwendbar. Event Flags können anders als Condition Flags trotzdem zur Task-Synchronisierung eingesetzt werden. Allerdings ist dies auf eine Richtung beschränkt, wie noch gezeigt wird. Die neu hinzukommende Komponente ist das Ereignis (Event).

Ereignisse können aperiodisch auftreten und - je nach Szenario - muss entsprechend schnell auf sie reagiert werden, das entsprechende koordinierte Verhalten also möglichst synchron zum Zeitpunkt des Events ausgelöst werden. Dazu wird eine Interrupt Service Routine (ISR) eingesetzt, die auf das Eintreten eines Events wartet und so lange keines eingetreten ist, untätig ist [vgl. 3, S. 87]. Ein Beispiel für ein solches Ereignis kann ein Hardware-Interrupt sein, etwa eine gedrückte Taste. Abbildung zeigt den schematischen Programmfluss im Einsatz mit Event Flags auf.

Wie eingangs bereits erwähnt, funktioniert diese Synchronisierung lediglich in eine Richtung: Die Interrupt Service Routine wird durch das vordefinierte Event (Interrupt) geweckt und setzt das entsprechende Flag, welches vom Empfänger-Task erkannt wird. Sollen zwei Tasks synchronisiert werden, sind die nun vorgestellten Signale dafür eine Möglichkeit.

Hier kommt Abbildung 5.9, S. 87 hin

2.3 Task-Synchronisation mittels Signale

3 Datentransfer ohne Synchronisation oder Koordination

Ist ein Datenaustausch zwischen Tasks erforderlich, aber nicht an Koordinierungs- und Synchronisationskriterien gebunden, unterliegt also weder einer zeitlichen Vorgabe noch einem Ablauf, werden verschiedene Datenstrukturen zum Austausch zwischen Tasks verwendet.

3.1 Überblick

3.2 Pools

3.3 Queues

4 Task-Synchronisation mit Datentransfer

Daten, die mit einer zeitlichen Priorität zwischen Tasks ausgetauscht werden, also genau zum richtigen Zeitpunkt beim Empfänger bereit stehen müssen, werden ebenfalls mit einer Datenstruktur transferiert. An diese besteht jedoch ein höherer Anspruch, da sie die Synchronisation sicherstellen muss.

4.1 Mailbox

5 Zusammenfassung

Literatur

- [1] Marco Winzker. *Elektronik für Entscheider: Grundwissen für Wirtschaft und Technik*. Vieweg, Wiesbaden, 2008.
- [2] Lehrstuhl VIII Universität Würzburg Institute für Informatik. *Rodos - Lehrstuhl für Informatik VIII*. 6. März 2019. URL: <https://www.informatik.uni-wuerzburg.de/aerospaceinfo/wissenschaftsforschung/rodos/>.
- [3] Jim Cooling. *Real-time Operating Systems: Book 1 - The Theory*. Independently Published, 2017.