Aufgabe 7: Uhrensynchronisation: NTP

Literatur:

[1] Mills, David L.:
Internet Time Synchronization: The Network Time Protokoll
IEEE Transactions on Communications, vol. 39, No. 10, October 1991

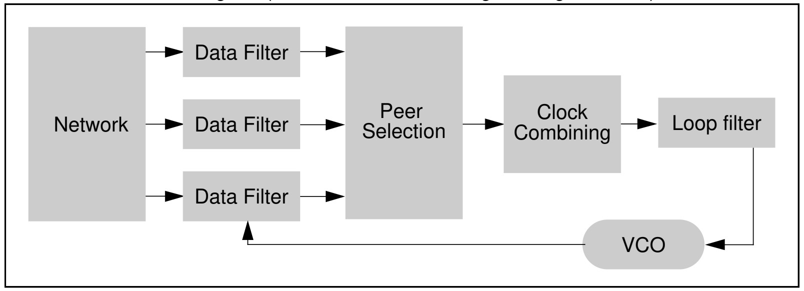
[2] Mills, David L.:
Network Time Protokoll (Version 3): Specification, Implementation and Analysis
http://www.ee.udel.edu/~mills/database/rfc/rfc1305/rfc1305{a,b,c}.pdf

[3] Coulouris, G.; Dollimore, J.; Kindberg, T.:
Distributed Systems: concepts and Design, 2nd Edition
Addison Wesley, Wokingham, GB: 1994

[4] NTP FAQ:
http://www.eecis.udel.edu/~ntp/ntpfaq/NTP-a-faq.htm

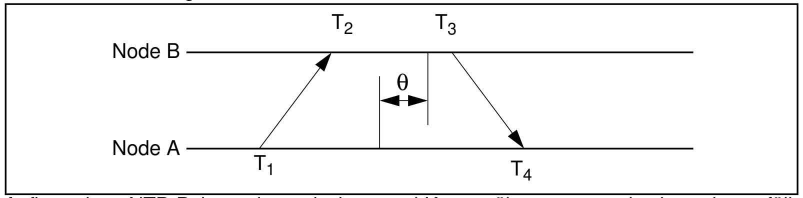
Lokale Installation der Uni-Erlangen:
siehe http://www4.informatik.uni-erlangen.de/Services/NTP

NTP berechnet die Zeitabweichung in 3 Stufen und korrigiert die Frequenz der lokalen Uhr mit Hilfe eines I-Reglers (vereinfachte Darstellung des Algorithmus!).



Es werden zu einer Menge von Knoten NTP-Pakete zur Synchronisationsermittlung geschickt. Die eintreffenden Daten werden mit statistischen Methoden gefiltert. Dann wird der vertrauenswürdigste Rechner ausgewählt und der Zeitunterschied zu diesem Rechner ermittelt. Dieser dient als Vorgabe für einen PLL-Regler, der die Frequenz der lokalen Uhr (VCO = Voltage Controlled Oscillator) regelt.
Ein wesentliches Merkmal des Protokolls ist die Abschätzung des Fehlers, der in den verschiedenen Phasen auftreten kann.

Modell zur Ermittlung des Zeitunterschieds θ zwischen zwei Rechneruhren:



Aufbau eines NTP-Pakets, das zwischen zwei Knoten übertragen und sukzessive gefüllt wird.

LI	VN	Mode	Stratum	Poll	Precison
Root Dispersion					
Root Delay					
Reference Timestamp					
T_1					
T_2					
T_3					
Authenticator					

(Die grau hinterlegten Felder sind für die Übungsaufgabe relevant.)
Stratum bezeichnet den Abstand zur Zeitquelle (z.B. DCF77-Empfänger). Rechner mit einer Möglichkeit, UTC direkt zu empfangen, haben Stratum 1. **Root Dispersion** ist der summierte Fehler, der bisher bei den Berechnung in den Zwischenknoten gemacht wurde. **Root Delay** ist die summierte Laufzeit des NTP-Pakets vom Root-Rechner über die Zwischenknoten. Die Zeiten T_k sind dem vorhergehenden Bild zu entnehmen.

Verwendete Symbole:

- δ Delay, Gesamtlaufzeit eines Pakets
- θ Zeitdifferenz zu einem anderen Rechner
- ε Fehler
- ε_σ Filterfehler
- ε_ξ Auswahlfehler
- φ Uhrabdrift, Frequenzfehler
- ρ Lesefehler
- ν Gewichtsfaktor
- w Gewichtsfaktor
- n Anzahl der Pakete, aus denen das 'beste' herausgefiltert wird
- m Anzahl der Knoten (Peers, Rechner), die als vertrauenswürdig. dh. als nicht fehlerhaft angesehen werden.
- τ Alter der NTP-Pakete in Millisekunden (Differenz aus T_4 des aktuellen (neuesten) Pakets und T_4 beim Empfang des Pakets)

Kurzbeschreibung des Algorithmus:

1 Polling
alle Knoten (Peers), die in einer Pollingliste enthalten sind, werden nach ihrer Zeit gefragt.

2 Datenfilterung
Mit Hilfe einer Statistik wird für jeden Knoten getrennt aus den letzten ***n*** Telegrammen das vertrauenswürdigste herausgesucht. Das vertrauenswürdigste ist dabei das mit der kürzesten Übertragungszeit.
Für jeden Knoten werden die einzelnen Werte nach aufsteigender Übertragungsdistanz $\delta/2+\varphi\tau$ sortiert.
Der Filter-Fehler ε_σ wird nach folgender Formel berechnet:

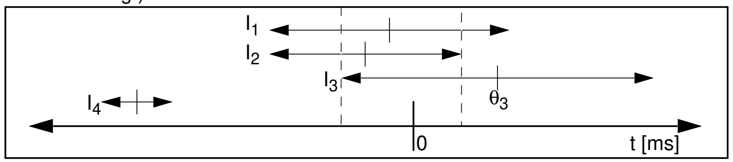
$$\varepsilon_\sigma = \sum_{j=0}^{n-1} |\theta_j - \theta_0| \nu^j$$

Wobei im Originalalgorithmus n=8 und als Gewichtsfaktor ν=0.5 verwendet wird. In der folgenden Übungsaufgabe begnügen wir uns mit n=4.
Für die weiteren Berechnungen wird nur das beste Wertepaar (θ_0, δ_0) zusammen mit dem abgeschätzten Fehler ε_σ verwendet, alle anderen Paare werden nicht berücksichtigt.

3 Auswahl der vertrauenswürdigen Knoten
Mithilfe der Tripel $(\theta_0, \delta_0, \varepsilon_\sigma)$ von allen befragten Knoten und dem summierten Gesamtfehler $\varepsilon = \rho + \varphi\tau + \varepsilon_\sigma$, der sich aus dem Lesefehler ρ, der möglichen Uhrabdrift $\varphi\tau$ seit dem Empfang des NTP-Pakets und dem Filterfehler ε_σ zusammensetzt, werden die Korrektheitsintervalle nach folgender Formel ermittelt:

$$I = \left[\theta - \frac{\delta}{2} - \varepsilon, \theta + \frac{\delta}{2} + \varepsilon\right]$$

Daraus wird der Durchschnitt aller Intervalle gebildet (siehe Skizze), wobei Intervalle die außerhalb der Mehrheit liegen, nicht berücksichtigt werden (-> I_4). Ein Knoten ist nur dann vertrauenswürdig, wenn sich seine Abweichung θ im Durchschnitt befindet (-> I_3 wird verworfen weil θ_3 nicht innerhalb der beiden gestrichelten Linien liegt).



Bleibt kein Rechner übrig, muß das gesamte Verfahren mit Polling usw. neu eingeleitet werden.

4 Auswahl des Knotens mit dem geringsten Fehler
Die verbliebenen ***m*** Werte werden zuerst nach Stratum und dann nach Root Delay geordnet. Dann wird für alle Knoten der Fehler aufgrund der unterschiedlichen Herkunft berechnet.

$$(\forall j (0 \le j < m)) \quad \varepsilon_j = \sum_{k=0}^{m-1} |\theta_j - \theta_k| w^k$$

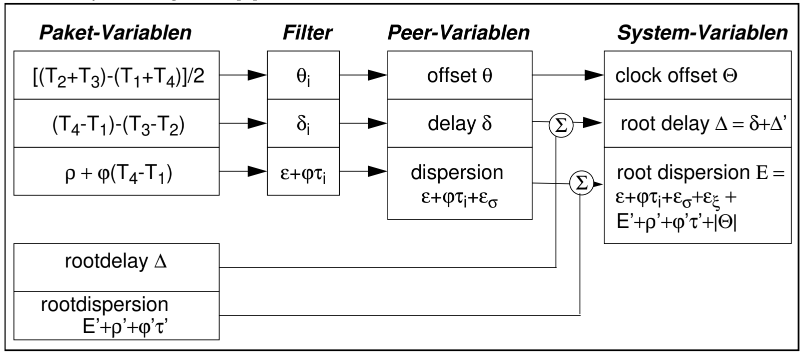
Für den Gewichtsfaktor w wird 0.75 im Originalalgorithmus verwendet. Der Wert mit dem größten ε_j wird verworfen. Das wird solange wiederholt, bis nur noch ein Wert übrigbleibt oder alle ε_j kleiner sind als das Minimum der ε_σ. Im Konfliktfall gewinnt das kleinere j. Der Fehler ε_j des übriggebliebenen Wertes wird als Auswahlfehler ε_ξ bezeichnet.

5 Berechnung der Abweichung (Offset Θ)
Dieser im vorherigen Schritt übriggebliebene Rechner wird als neuer Referenzknoten ausgewählt. Die neue Abweichung Θ wird gespeichert. Das Stratum wird auf den um 1 erhöhten Wert des Referenzrechners gesetzt. Das Root Delay ergibt sich aus dem Root Delay des Referenzrechners und der geschätzten Übertragungszeit. Die Gesamtfehlerabschätzung (Root Dispersion) ergibt sich mit $\varepsilon=\rho+\varphi(T_4-T_1)$ zu:

$$E = \varepsilon + \varphi\tau_i + \varepsilon_\sigma + \varepsilon_\xi + E' + \rho' + \varphi'\tau' + |\Theta|$$

6 Weitergabe der berechneten Werte an den Regler zur Frequenzanpassung

Fehlerfortpflanzung nach [2] S. 103:



Reale Werte an der Uni Erlangen:

tcsh: /local/NTP/bin/ntpq
ntpq> host faui01
current host set to faui01.informatik.uni-erlangen.de

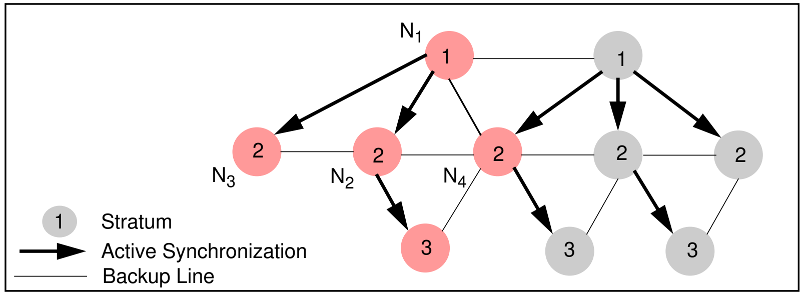
ntpq> rv
status=06f4 leap_none, sync_ntp, 15 events, event_peer/strat_chg
system="UNIX", leap=00, stratum=2, rootdelay=8.00, rootdispersion=3.60,
peer=1787, refid=faui45.informatik.uni-erlangen.de,
reftime=afc2c566.9f572000 Fri, Jun 11 1993 10:48:38.622, poll=7,
clock=afc2c5a0.d2b08000 Fri, Jun 11 1993 10:49:36.823, phase=0.102,
freq=95551.67, compliance=3

ntpq>
...

Aufgabenstellung:

Der Knoten N_2 soll neu synchronisiert werden.
In seiner Liste möglicher Referenzrechner befinden sich N_1, N_3 und N_4.

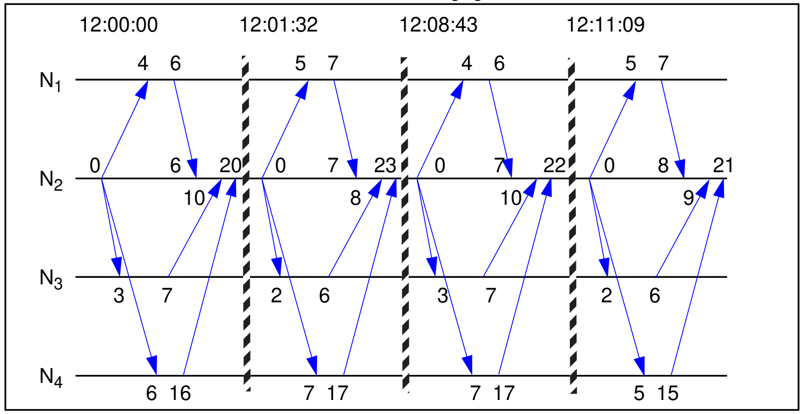
Übersichtsbild der Knoten:



Nachrichtenaustausch, die Länge des Polling-Intervall liegt im Bereich von [64s .. 1024s]

Im folgenden Bild sind die Laufzeiten der NTP-Pakete bei den letzten 3 Polling-Runden und der aktuellen angegeben. Alle Zeiten sind in Millisekunden angegeben und sind gedanklich an die im "hh:mm:ss"-Format angegebenen absoluten Zeiten anzuhängen.

Die Paketlaufzeiten usw. sind in Millisekunden angegeben.



Beispiel:
Knoten N_2 schickte um 12:00:00.000 seiner lokalen Zeit ein NTP-Paket an N_1, das dort um 12:00:00.004 dortiger lokaler Zeit eintraf, bis um 12:00:00.006 bearbeitet wurde und dann zurückgeschickt wurde. Es traf um 12:00:00.006 wieder beim Knoten N_2 ein.

Lokale Daten der Knoten zum Zeitpunkt T=12:11:09

Knoten		N_1	N_2	N_3	N_4
clock offset	Θ	0	2	20	15
root delay	Δ	0	2	46	8
root dispersion	E	2	13	24	5
Lesefehler	ρ	10^{-3}	10	20	4
Uhrabdrift	φ	10^{-4}	10^{-4}	10^{-4}	10^{-4}
stratum		1	2	2	2

a) Berechnen Sie den Referenzknoten und den dazugehörigen Offset für Rechner N_2 !

b) Wie sieht das Ergebnis aus, wenn in den Stratum-1-Rechner N_1 eine bessere (und damit auch teurere) Uhr mit Abdrift $\varphi = 10^{-9}$ eingebaut wird.

Lösungsvorschlag Aufgabe 7: Uhrensynchronisation NTP

a) Berechnen Sie den Referenzknoten und den dazugehörigen Offset für Rechner N_2

1 Polling

	(δ, θ, $\delta/2+\varphi\tau$)			
	12:00:00	12:01:32	12:08:43	12:11:09
N_1	(4,2, 68.9)	(5,2.5, 60.2)	(5,1.5, 17.1)	(6,2, 3)
N_3	(6,0, 69.9)	(4,0, 59.7)	(6,0, 17.6)	(5,-0.5, 2.5)
N_4	(10,1, 71.9)	(13,0.5, 64.2)	(12,1, 20.6)	(11,-0.5, 5.5)

2 Datenfilterung
Sortieren nach Übertragungsdistanz, Berechnung der Fehler

	θ				ε_σ	$\rho+\varphi\tau_i$	ε
N_1	2	1.5	2.5	2	0.25 + 0.125 + 0	10^{-3} + 0	0.4
N_3	-0.5	0	0	0	0.25 + 0.125 + 0.0625	20 + 0	20.4
N_4	-0.5	1	0.5	1	0.75 + 0.25 + 0.1875	4 + 0	5.2

3 Auswahl der vertrauenswürdigen Knoten

	θ	$\delta/2$	ε	$I_{Untergrenze}$	$I_{Obergrenze}$
N_1	2	3	0.4	-1.4	5.4
N_3	-0.5	2.5	20.4	-23.4	22.4
N_4	-0.5	5.5	5.2	-11.2	10.2
Durchschnitt				-1.4	5.4

Wie leicht zu sehen ist, liegen alle θ_i im Durchschnittsintervall.

4 Auswahl des Knotens mit dem geringsten Fehler

	j	Stratum	Root Delay	θ	ε_j	ε_j
N_1	0	1	0	2	0 + 2.5*0.75 + 2.5*0.75^2	3.3
N_4	1	2	8	-0.5	2.5*1 + 0 + 0	2.5
N_3	2	2	46	-0.5	2.5*1 + 0 + 0	2.5

Das Minimum der ε_σ ist 0.375
Der Wert für N_1 wird verworfen da er die größte Abweichung relativ zu den anderen liefert.

	j	Stratum	Root Delay	θ	ε_j	ε_j
N_4	0	2	8	-0.5	0 + 0	0
N_3	1	2	46	-0.5	0 + 0	0

Nun wird N_3 verworfen und es wird N_4 wegen des kleineren Index j ausgewählt.

5 Berechnung der Abweichung

Stratum: 3
Clock Offset: -0.5
Root Delay: 8 + 11 = 19
Root Dispersion: $E = \rho + \varphi\tau + \varepsilon_\sigma + \varepsilon_\xi + E' + \rho' + \varphi'\tau' + |\Theta|$
$= 5.2 + 0 + 5 + 0.5 = 10.7$

6 Weitergabe der berechneten Werte an den Regler zur Frequenzanpassung

b) Wie sieht das Ergebnis aus, wenn in den Stratum-1-Rechner N_1 eine bessere (und damit auch teurere) Uhr mit Abdrift $\varphi = 10^{-9}$ eingebaut wird.

1 Polling

	(δ, θ, $\delta/2+\varphi\tau$)			
	12:00:00	12:01:32	12:08:43	12:11:09
N_1	(4,2, **2.0**)	(5,2.5, 2.5)	(5,1.5, 2.5)	(6,2, 3.0)
N_3	(6,0,69.9)	(4,0, 59.7)	(6,0, 17.6)	(5,-0.5, 2.5)
N_4	(10,1, 71.9)	(13,0.5, 64.2)	(12,1, 20.6)	(11,-0.5, 5.5)

2 Datenfilterung
Sortieren nach Übertragungsdistanz, Berechnung der Fehler

	θ				ε_σ	$\rho+\varphi\tau_i$	ε
N_1	**2**	1.5	2.5	**2**	0.25 + 0.125 + 0	10^{-3} + $.667*10^{-6}$	0.4
N_3	-0.5	0	0	0	0.25 + 0125 + 0.0625	20 + 0	20.4
N_4	-0.5	1	0.5	1	0.75 + 0.25 + 0.1875	4 + 0	5.2

3 Auswahl der vertrauenswürdigen Knoten

	θ	$\delta/2$	ε	$I_{Untergrenze}$	$I_{Obergrenze}$
N_1	2	**2**	0.4	-0.4	4.4
N_3	-0.5	2.5	20.4	-23.4	22.4
N_4	-0.5	5.5	5.2	-11.2	10.2
Durchschnitt				-0.4	4.4

Aufgrund der neuen kurzen Übertragungszeit bei N_1 fallen die beiden anderen Werten für θ_i nicht mehr in das Durchschnittsintervall.

Die Synchronisation erfolgt aufgrund des Rechners N_1.