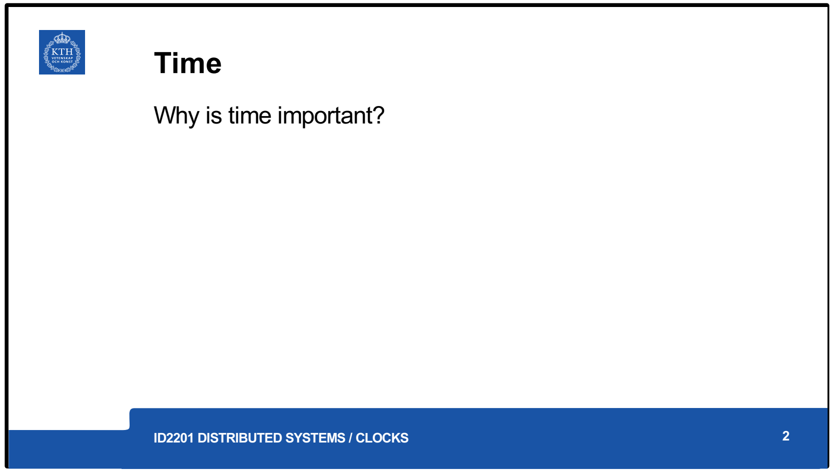
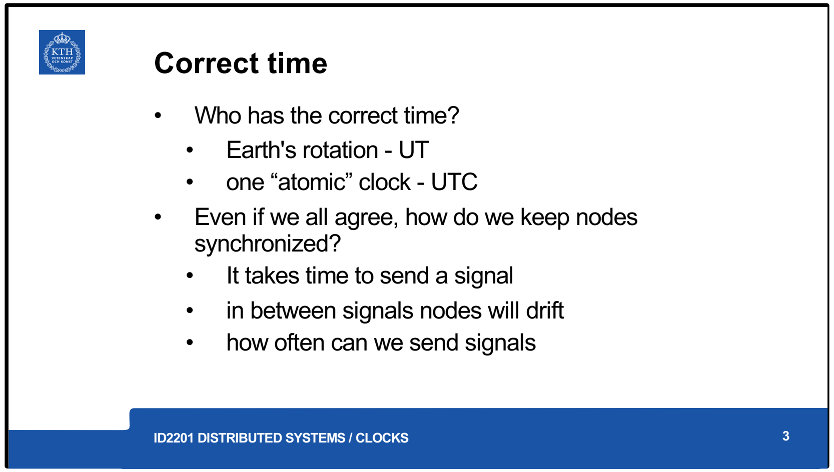
KTH ROYAL INSTITUTE
OF TECHNOLOGY

Clocks

Johan Montelius and Vladimir Vlassov



The importance of accurate timekeeping for distributed systems is rather apparent. **First**, time is a quantity we often want to measure accurately, e.g., for monitoring or ordering: To know at what time of day a particular event occurred at a particular computer. For example, an eCommerce transaction involves events at a merchant’s and bank’s computers. It is essential, for auditing purposes, that those events are timestamped accurately. **Second**, algorithms that depend upon clock synchronization have been developed for several problems in distributed computing, e.g., maintaining the consistency of distributed data (the use of timestamps to serialize transactions) to asses freshness (staleness) of replicas or cashed copies, order updates and eliminating the processing of duplicate updates, checking the authenticity of a request sent to a server, etc.



UT = Universal time. Defined by the Earth's rotation, formerly determined by astronomical observations, today, GPS satellites are used instead. This time scale is slightly irregular. The Earth rotates once in about 24 hours with respect to the sun and once every 23 hours, 56 minutes, and 4 seconds with respect to the stars. Earth's rotation is slowing slightly with time; thus, a day was shorter in the past. This is due to the tidal effects the Moon has on Earth's rotation. Atomic clocks show that a modern-day is longer by about 1.7 milliseconds than a century ago, slowly increasing the rate at which UTC is adjusted by leap seconds. **UTC (French: Temps universel coordonné), Coordinated Universal Time**, is the primary time standard by which the world regulates clocks and time. UTC is based on **TAI (International Atomic Time)**, but a so-called ‘leap second’ is inserted – or, more rarely, deleted – occasionally to keep it in step with astronomical time. UTC signals are synchronized and broadcast regularly from land-based radio stations and satellites covering many parts of the world. Satellite sources include GPS.

TAI (French: Temps Atomique International), International Atomic Time, is a weighted average of the time kept by over 400 atomic clocks in over 50 national laboratories worldwide.
Atomic clock: the second as the duration of 9192631770 (about 9B) cycles of the radiation corresponding to the transition between two energy levels of the cesium-133 atom at rest at a temperature of 0 K.

A correct clock

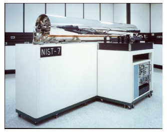
Drift is a change in how well one clock can measure a time interval.

Monotonic is the property that time always moves forward.

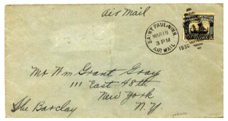
Correctness often means monotonic and low drift.

A correct clock might not be synchronized.

How to synchronize



033-415783



Synchronization

The accuracy or external synchronization:

- Each node in our network is synchronized with an external (global) source within a bound.

Precision or internal synchronization:

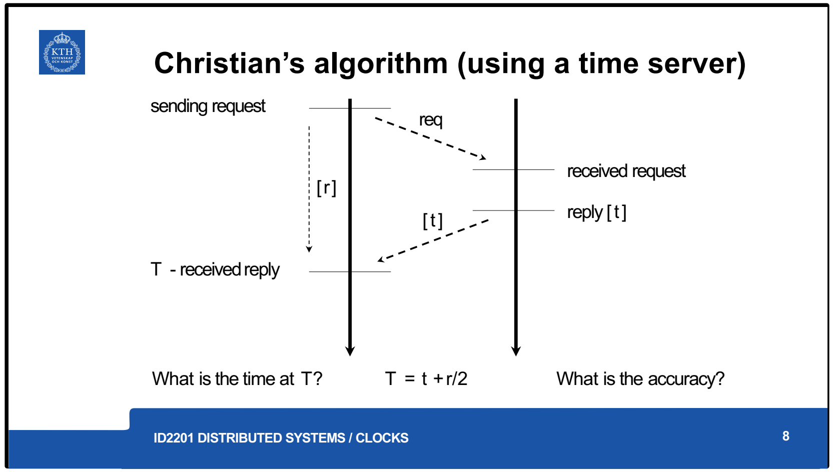
- Every pair of nodes in our network are synchronized within a bound.

Asynchronous networks

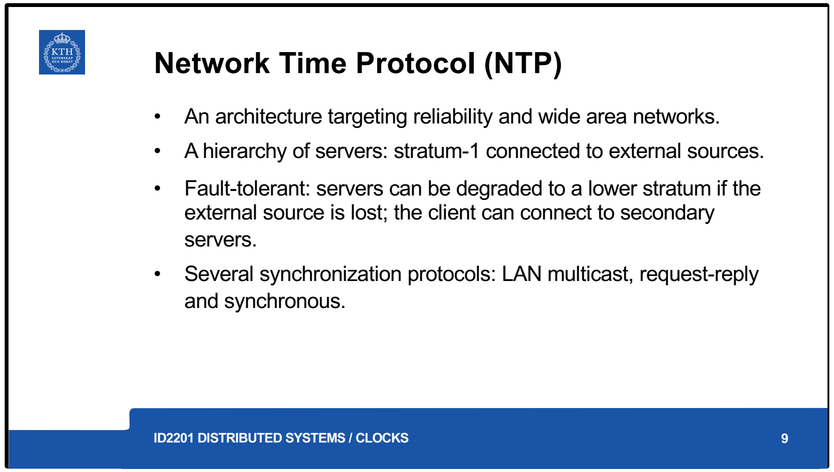
One server is connected to an external source and used to synchronize other nodes in the network.

The problem is, of course, that round-trip times are unknown and vary.

A minimum propagation time can be known.

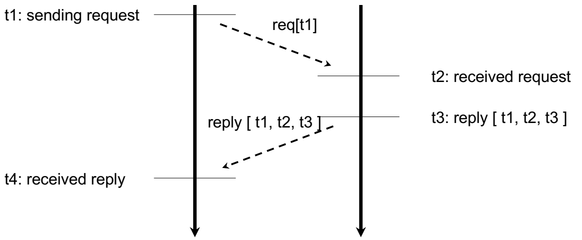


Cristian [1989] suggested using a time server connected to a device that receives signals from a source of UTC to synchronize computers externally. Upon request, the server process supplies the time according to its clock. The time at the server when the reply arrives is within [t + *min*, t+T_round-*min*] – the width of this range is T_round-2*min, so the accuracy is +-(T_round/2-min)
min is minimum transmission time. In the above slide, r is T_round.



Network Time Protocol (NTP) is a networking protocol for clock synchronization between computer systems over packet-switched, variable-latency data networks. Since 1985, NTP has been one of the oldest Internet protocols. The NTP service is provided by a network of servers across the Internet. Primary servers are connected directly to a time source, such as a radio clock receiving UTC; secondary servers are synchronized with primary servers. **The servers are connected in a logical hierarchy called a synchronization sub-net, whose levels are called strata. Primary servers occupy stratum 1: they are at the root.** Stratum 2 servers are secondary servers that are synchronized directly with the primary servers; stratum 3 servers are synchronized with stratum 2 servers, and so on. The lowest-level (leaf) servers execute in users' workstations.

NTP

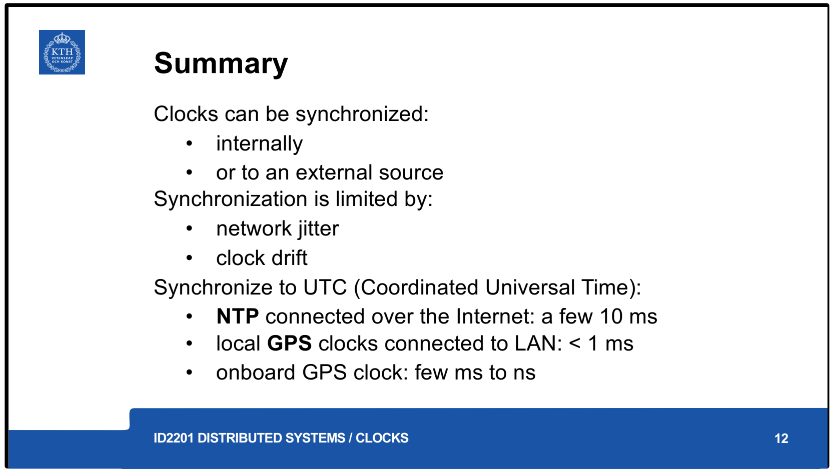


Similar to Christian's but with a better estimate of delay.
Stateless, no need to record *r*.

Berkeley algorithm

It is used to synchronize a network of nodes.

- send requests to all nodes
- collect it and calculate an *average* time T
- send out individual deltas to each node



NTP can usually maintain time to within tens of milliseconds over the public Internet and can achieve better than one-millisecond accuracy in local area networks under ideal conditions.
UTC signals are synchronized and broadcast regularly from land-based radio stations and satellites covering many parts of the world. **Satellite sources include GPS.**
GPS accuracy is ~ 1 microsecond. NTP accuracy is ~ 100 milliseconds. GPS is more accurate than NTP and has less latency compared to NTP. Clear sky, and after satellites are locked, the GPS master clock will send a signal every second to the slave clocks.
There are several approaches to correcting the times on computer clocks. For example, computers may use **radio receivers to get time readings from GPS with an accuracy of about 1 microsecond**. But GPS receivers do not operate inside buildings, nor can the cost be justified for every computer. Instead, a computer with an accurate time source such as GPS can send timing messages to other computers in its network. The

resulting agreement between the times on the local clocks is affected by variable message delays.