(Byzantine) Vertical Paxos

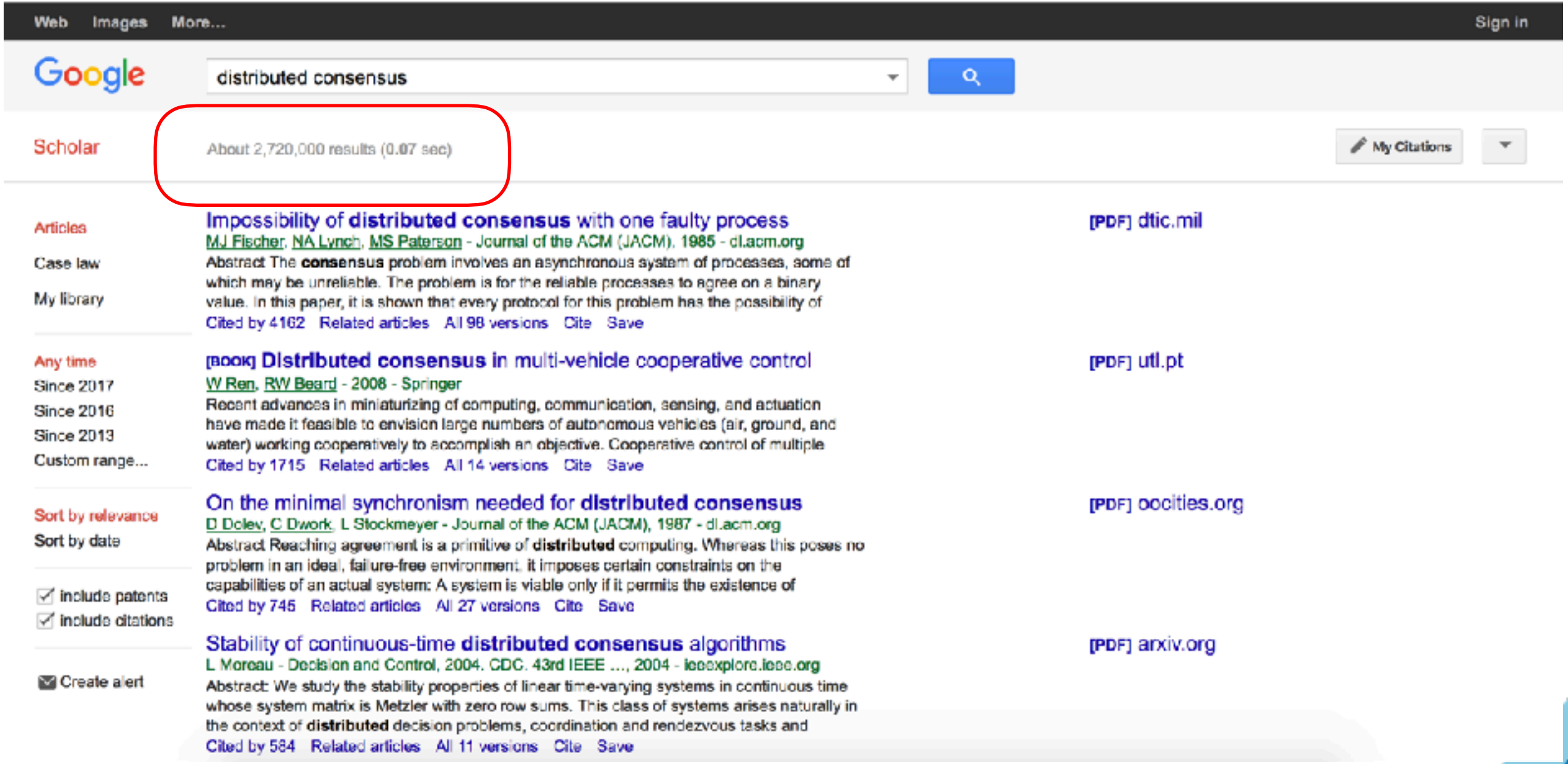
Dahlia Malkhi

EPFL Summer School on Blockchains, June 2017

vmware®

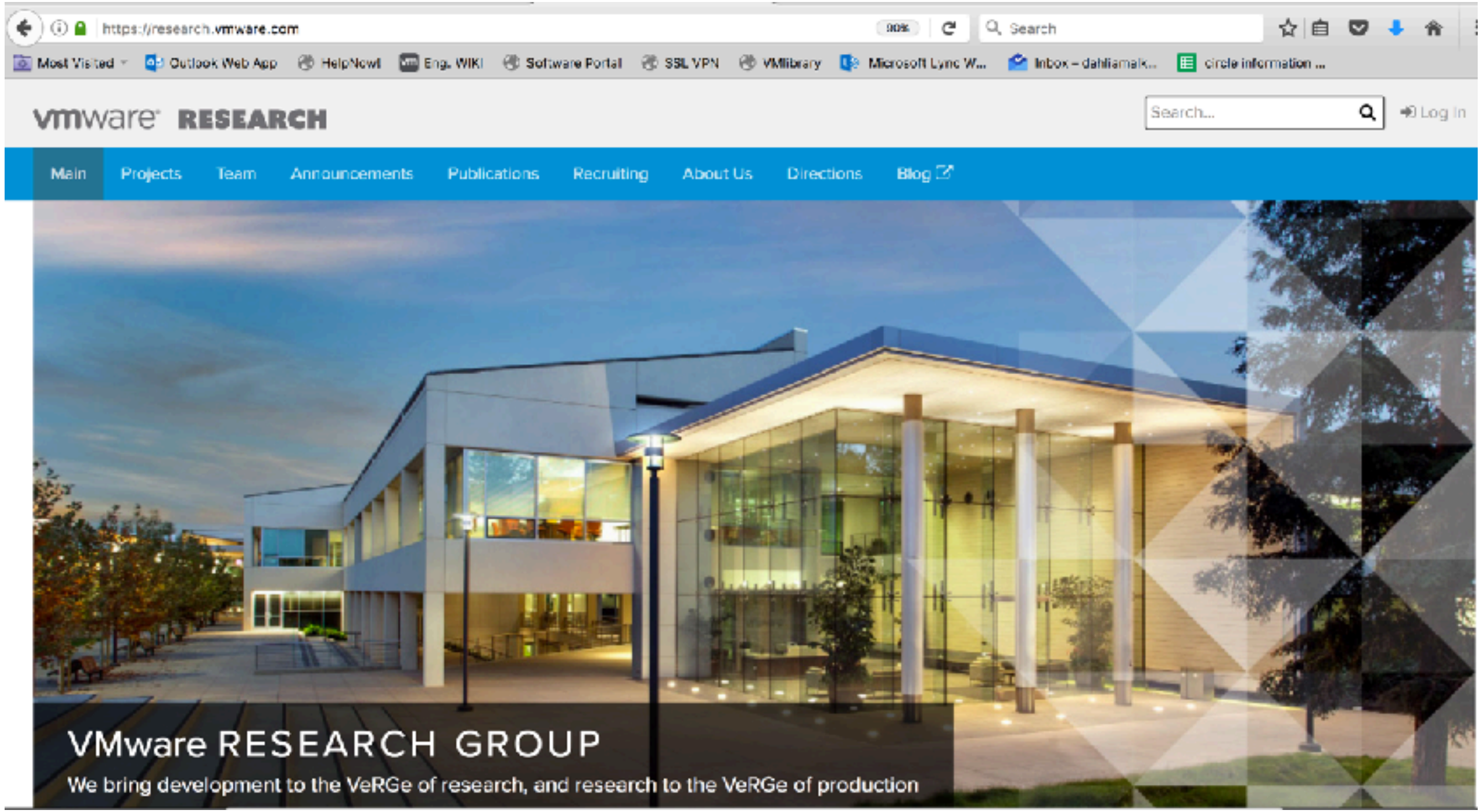
How a typical application becomes distributed

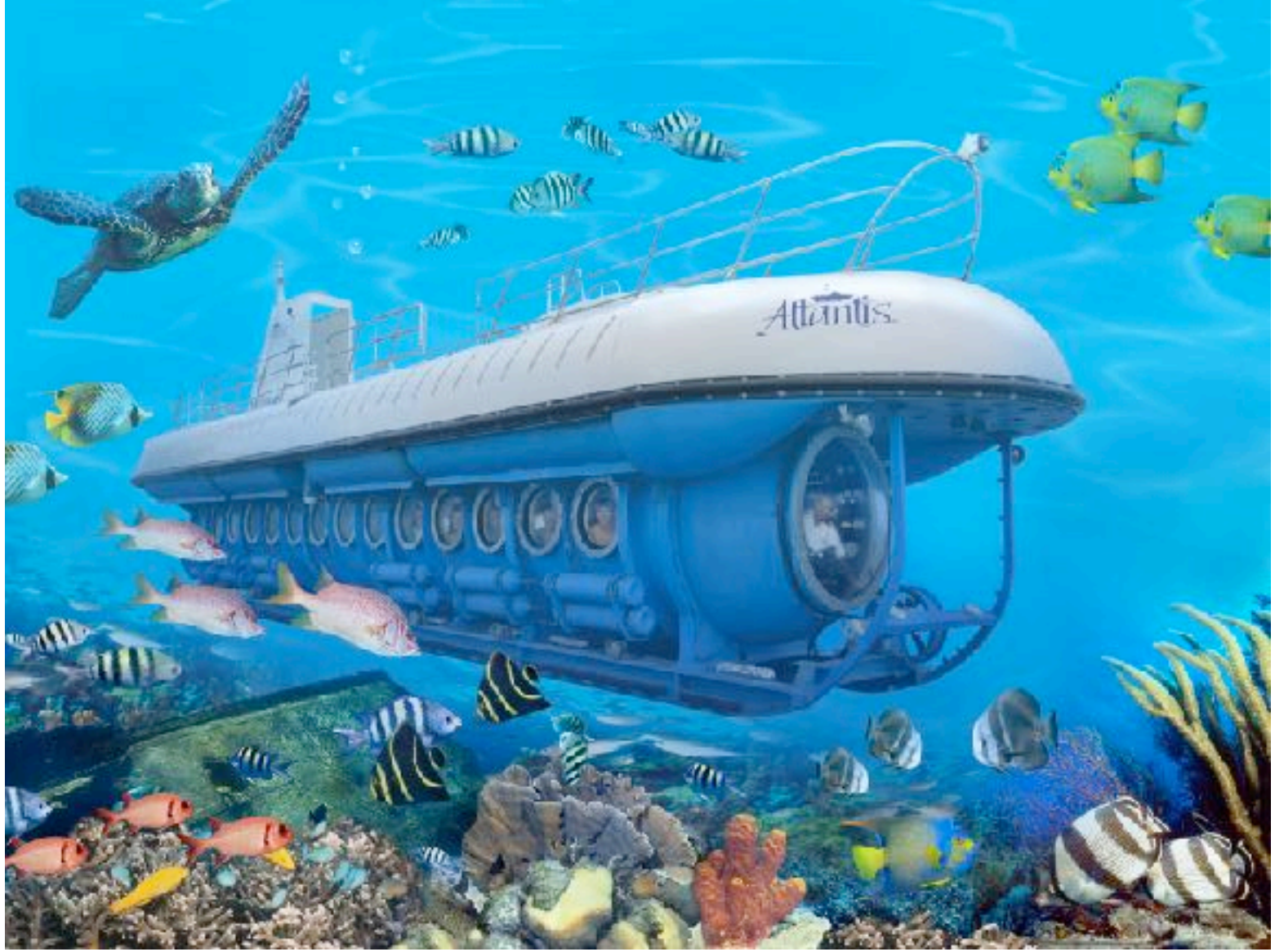
How a typical application becomes distributed


Web Images More... Sign in
Google
distributed consensus
Scholar
About 2,720,000 results (0.07 sec)
My Citations
Articles
Case law
My library
Any time
Since 2017
Since 2016
Since 2013
Custom range...
Sort by relevance
Sort by date
include patents
include citations
Create alert
Impossibility of distributed consensus with one faulty process
MJ Fischer, NA Lynch, MS Paterson - Journal of the ACM (JACM), 1985 - dl.acm.org
Abstract The consensus problem involves an asynchronous system of processes, some of which may be unreliable. The problem is for the reliable processes to agree on a binary value. In this paper, it is shown that every protocol for this problem has the possibility of
Cited by 4162 Related articles All 98 versions Cite Save
[PDF] dtic.mil
[BOOK] Distributed consensus in multi-vehicle cooperative control
W Ren, RW Beard - 2008 - Springer
Recent advances in miniaturizing of computing, communication, sensing, and actuation have made it feasible to envision large numbers of autonomous vehicles (air, ground, and water) working cooperatively to accomplish an objective. Cooperative control of multiple
Cited by 1715 Related articles All 14 versions Cite Save
[PDF] utl.pt
On the minimal synchronism needed for distributed consensus
D Dolev, C Dwork, L Stockmeyer - Journal of the ACM (JACM), 1987 - dl.acm.org
Abstract Reaching agreement is a primitive of distributed computing. Whereas this poses no problem in an ideal, failure-free environment, it imposes certain constraints on the capabilities of an actual system: A system is viable only if it permits the existence of
Cited by 745 Related articles All 27 versions Cite Save
[PDF] ocities.org
Stability of continuous-time distributed consensus algorithms
L Moreau - Decision and Control, 2004. CDC. 43rd IEEE ..., 2004 - ieeexplore.ieee.org
Abstract: We study the stability properties of linear time-varying systems in continuous time whose system matrix is Metzler with zero row sums. This class of systems arises naturally in the context of distributed decision problems, coordination and rendezvous tasks and
Cited by 584 Related articles All 11 versions Cite Save
[PDF] arxiv.org

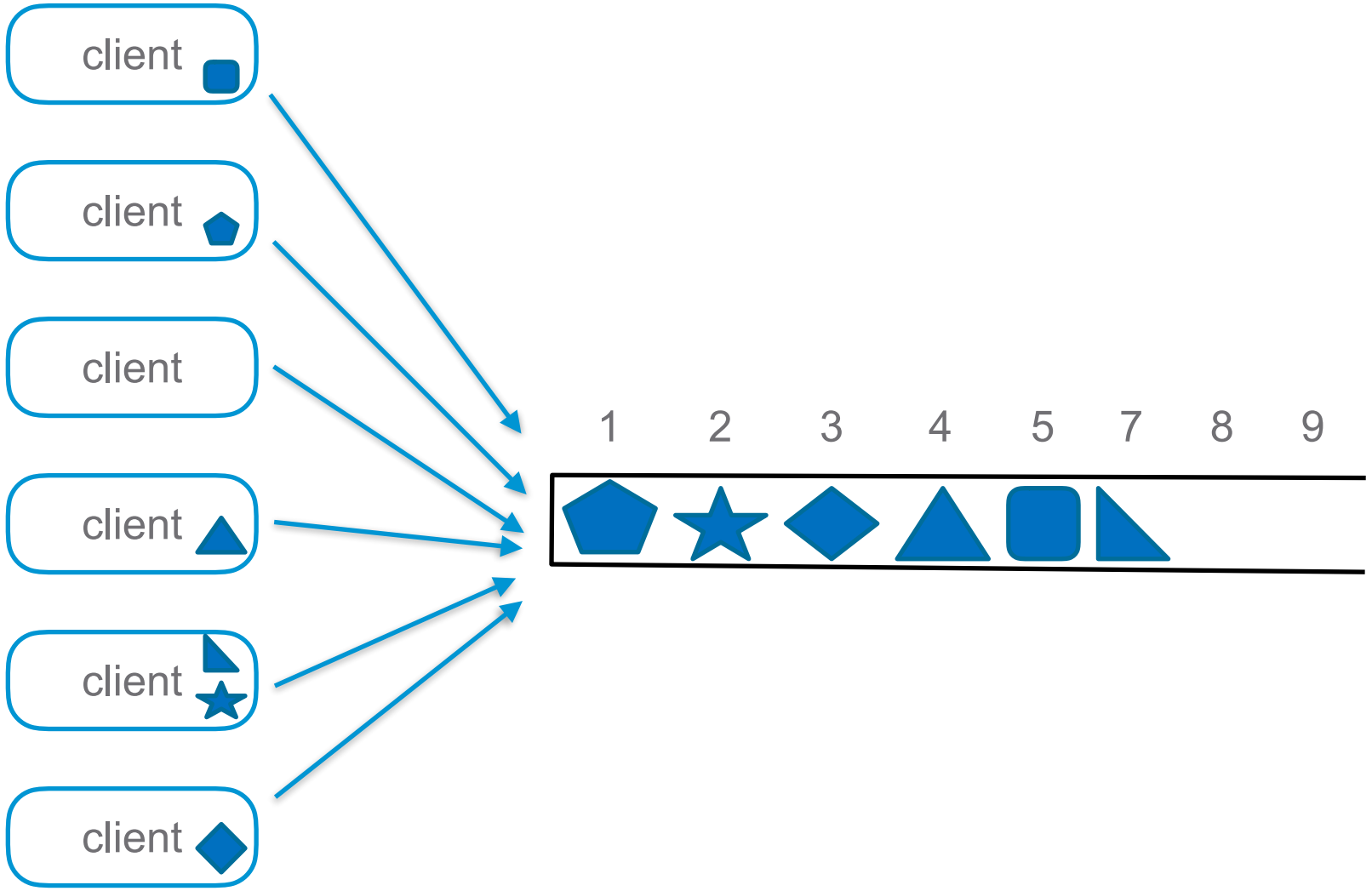
SMR Reconfiguration

- Overview SMR
- (Vanilla) Paxos reconfiguration
- Contiguous log reconfiguration: VR, ZK, Raft
- Explicit reconfiguration: Cheap Paxos, Vertical Paxos, Virtually Synchronous Paxos, 700 BFTs
- Byzantine VP

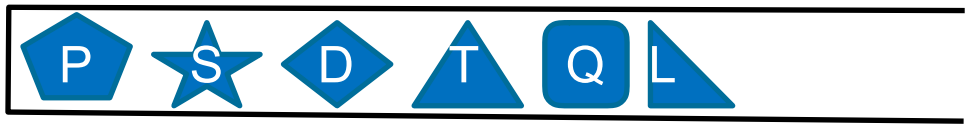

https://research.vmware.com
Search
Most Visited
Outlook Web App
HelpNow!
Eng. WIKI
Software Portal
SSL VPN
VMlibrary
Microsoft Lync W...
Inbox – dahliamalk...
circle information ...
vmware RESEARCH
Search...
Log In
Main
Projects
Team
Announcements
Publications
Recruiting
About Us
Directions
Blog
VMware RESEARCH GROUP
We bring development to the VeRGe of research, and research to the VeRGe of production


Atlantis

State Machine Replication (SMR)

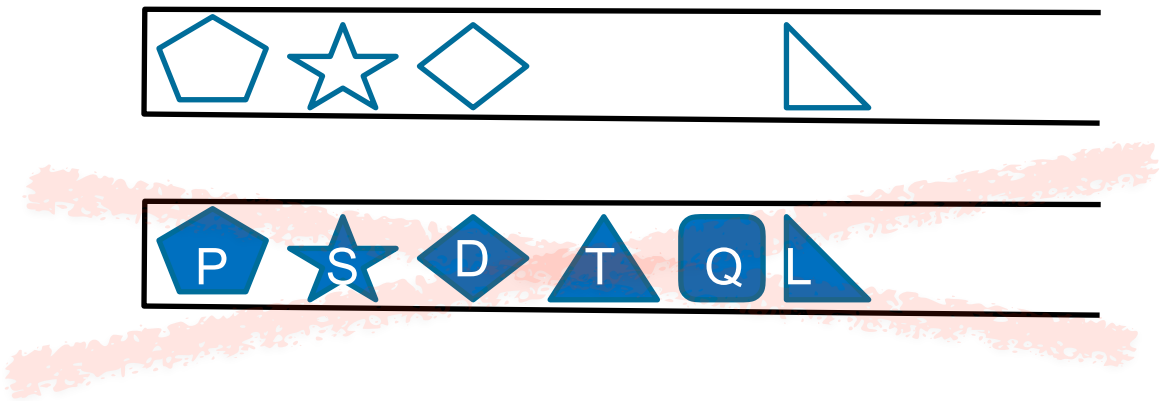


State Machine Replication (SMR)



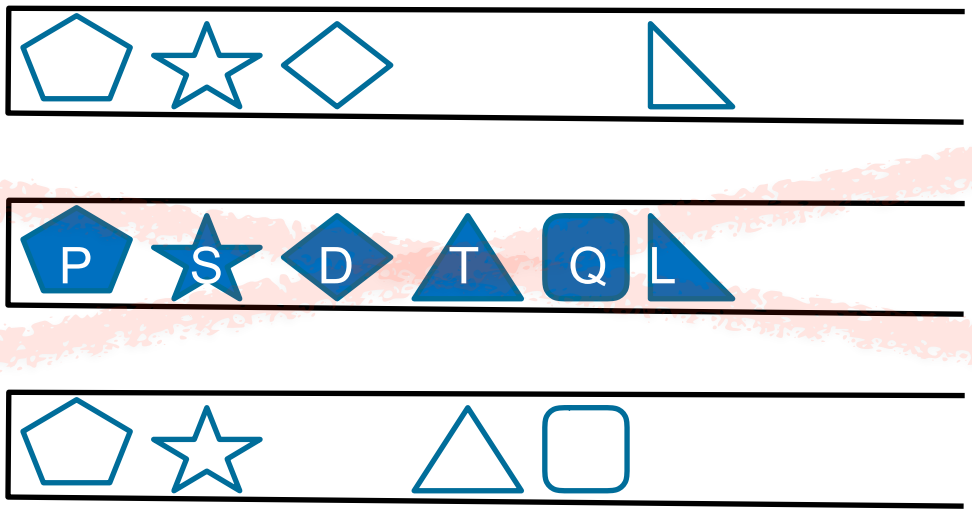
State Machine Replication (SMR)

F-resilience



State Machine Replication (SMR)

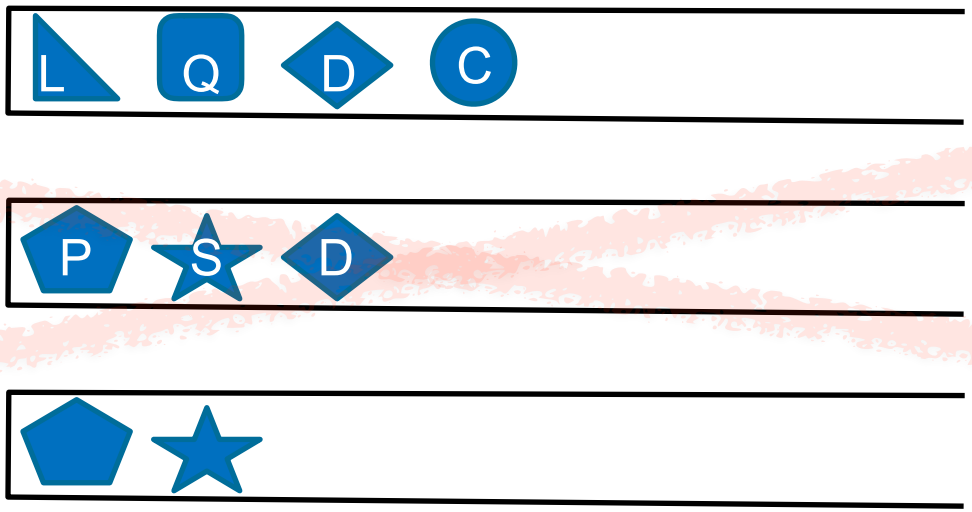
F-resilience



Glossary

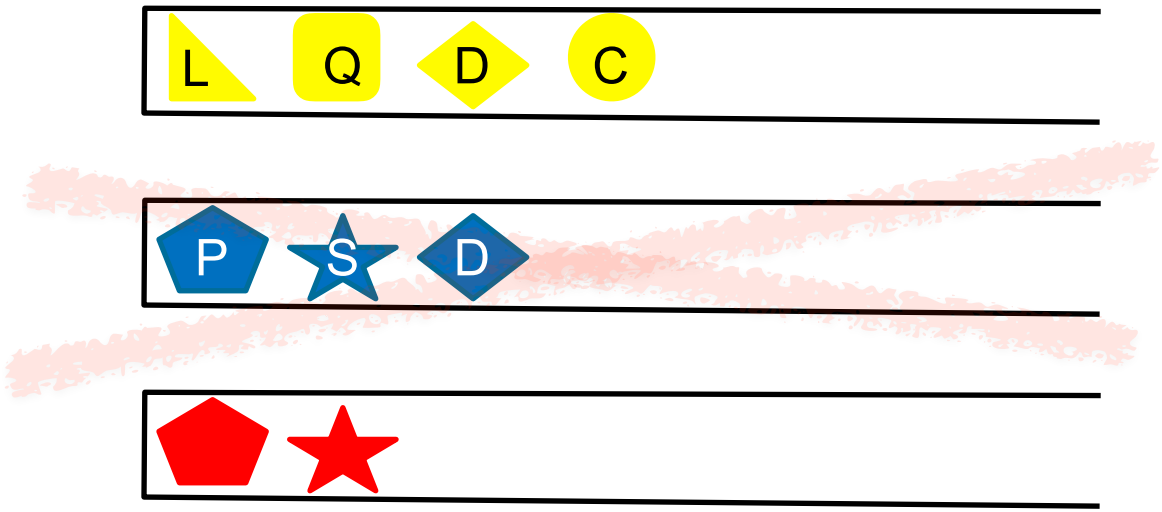
- replica(s)
  - leader/coordinator
- command sequence
- clients
- learn

A replicated system after a stormy night



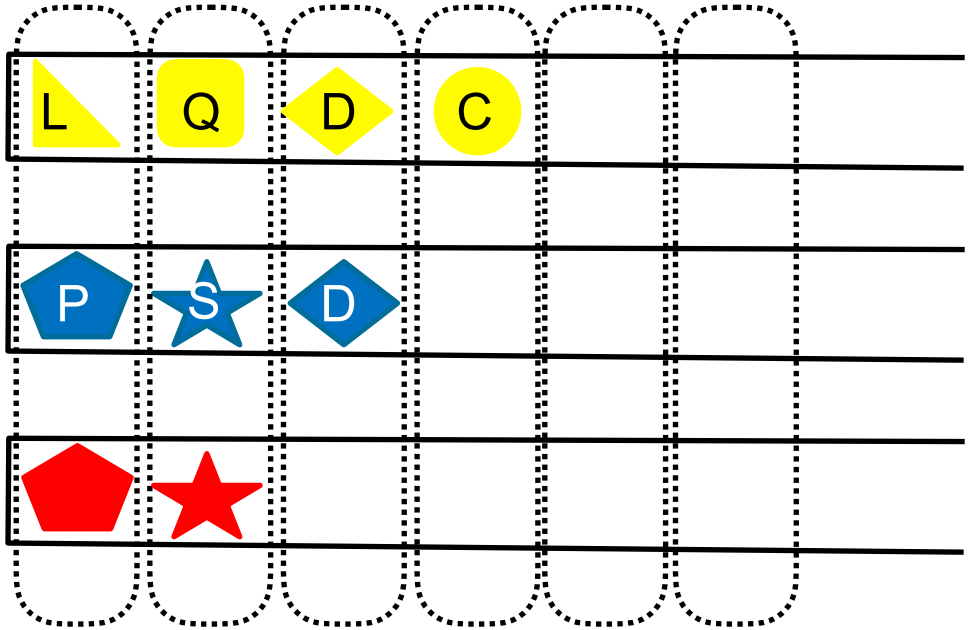
A replicated system after a stormy night

with leader ranks



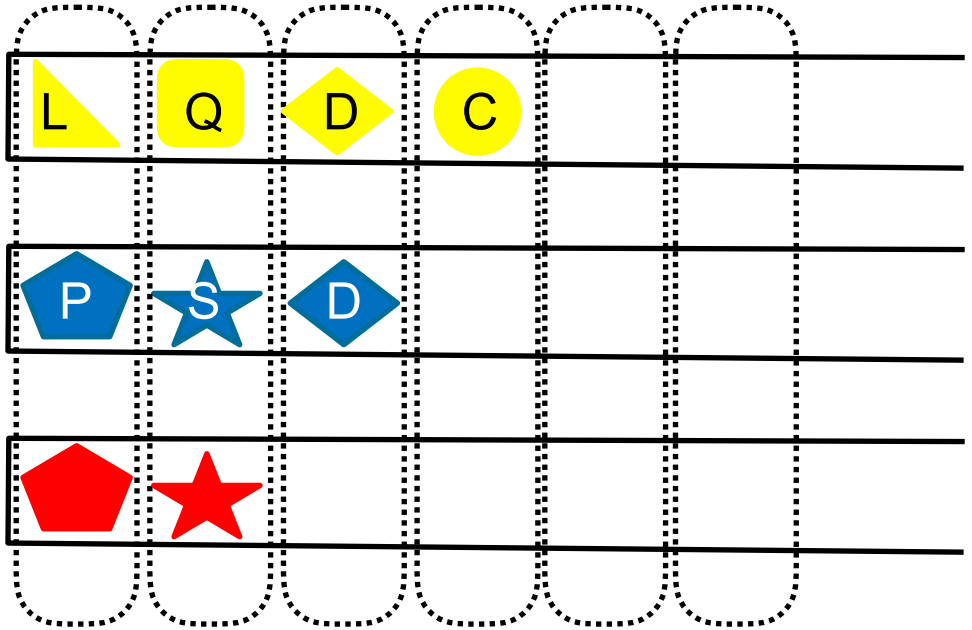
Paxos: Two-Phase SMR Protocol

- First phase: leader replacement, amortized
- Second phase: replication



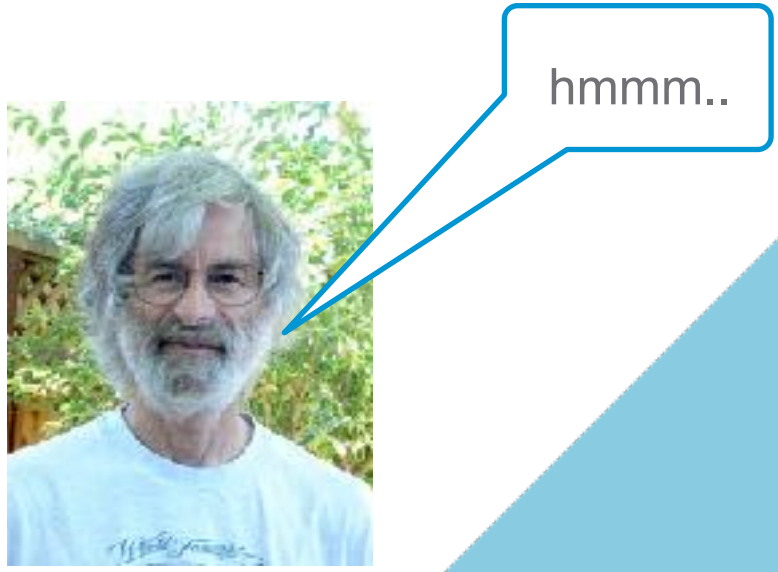
Contiguous log

- most recent log wins



Vertical Paxos

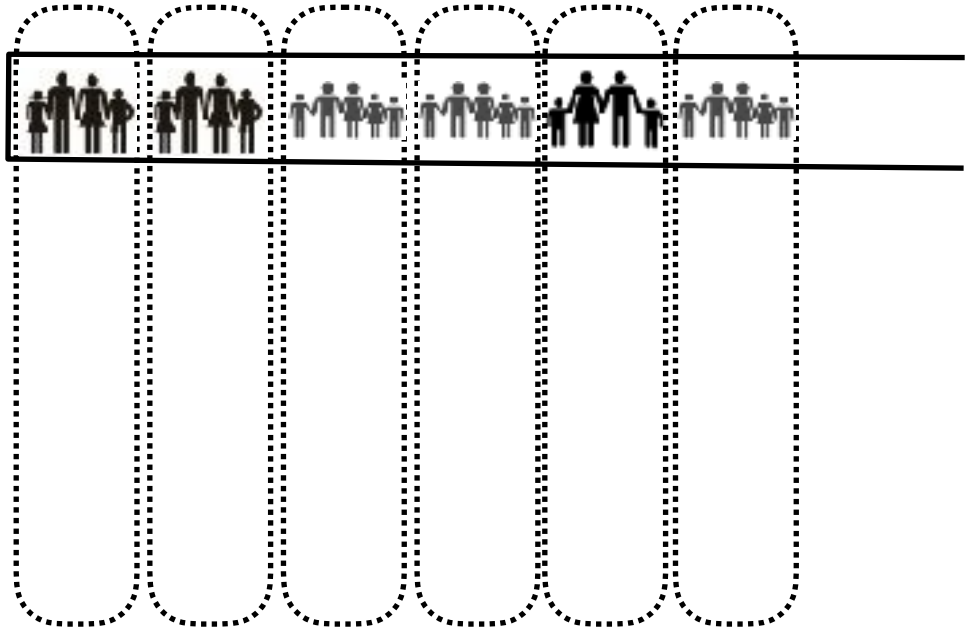
reconfiguration == leader replacement



vmware

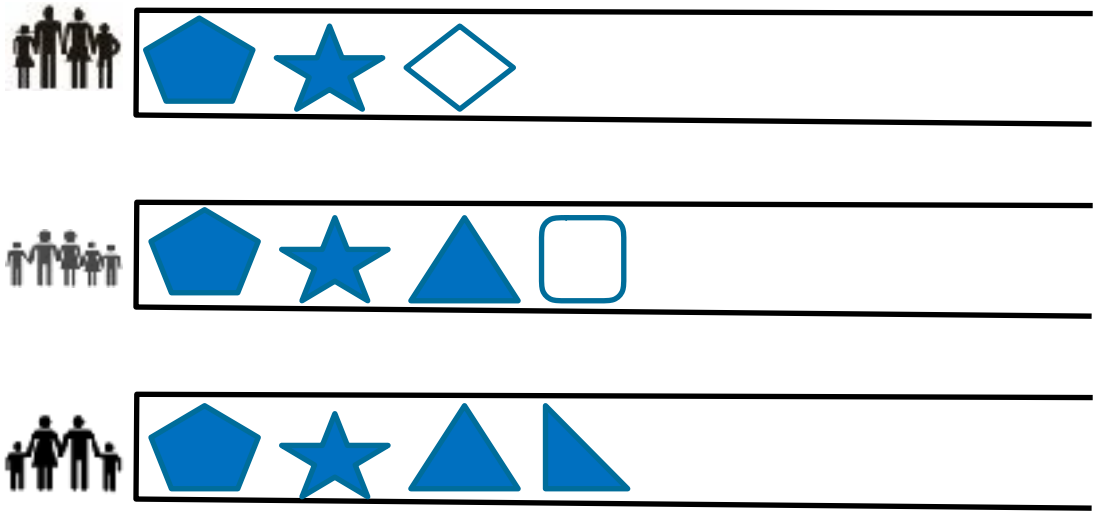
reconfiguration

- intra-sequence



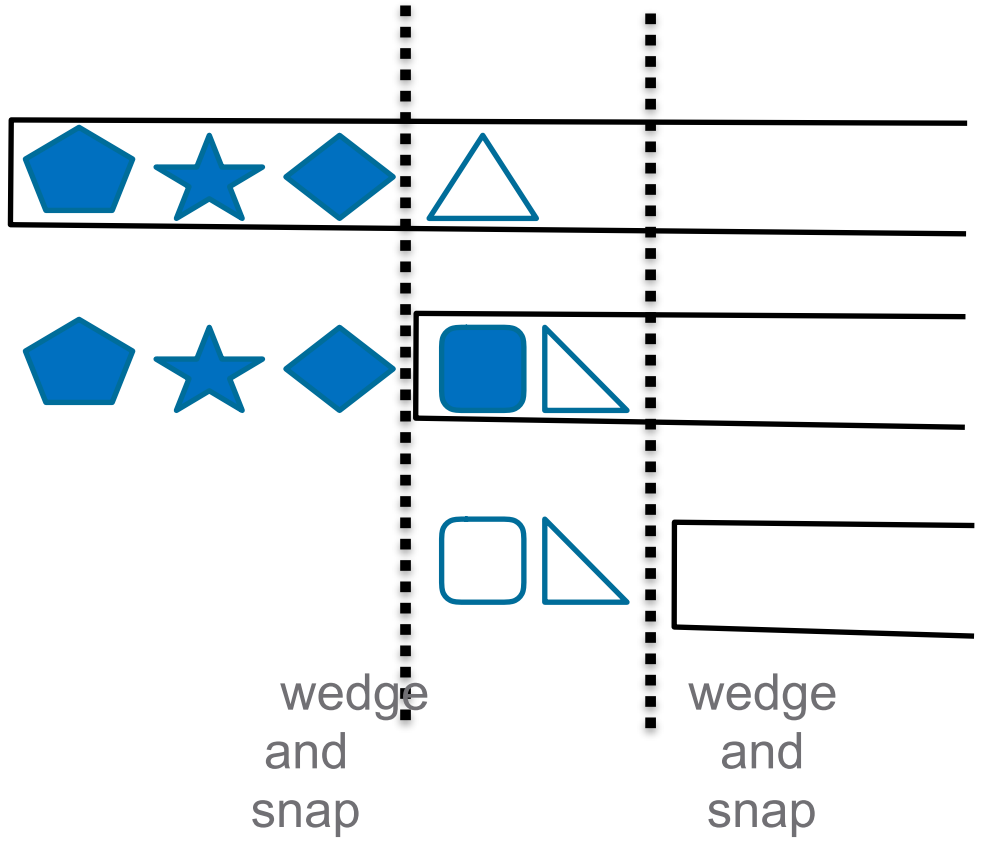
reconfiguration

- intra-sequence



Virtually Synchronous Paxos

- inter-sequence
- separate steady state from reconfiguration
- two consensus decisions: ***state*** and ***next***

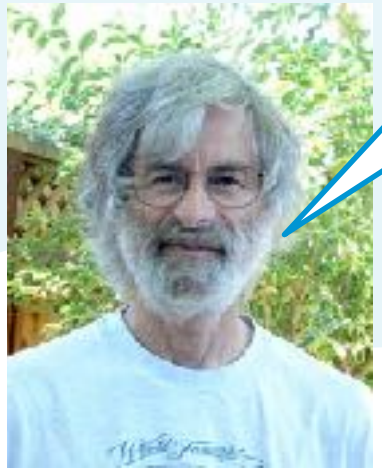


Vertical Paxos

- steady state is F+1
- aux reconfiguration master
- Primary backup: special case, implicit wedge-and-snap
- Cheap Paxos: special case (F+1)-of-(2F+1), F standby's

- why is it vertical?

Each of the phases of Paxos may use non-intersecting quorums. Only quorums from different phases are required to intersect. Majority quorums are not necessary as intersection is required only across phases.



nice!

Flexible Paxos: Use-Cases

BVP: Byzantine Vertical Paxos

- non-repudiation: interest in one correct replica
- wedge and snap: can “prove” decision

BVP in the synch model

- **3-message-delay**
- steady state 2F+1
  - client-leader
  - leader-all (signed)
  - all-client (signed)
  - proceed with 2F+1 signed echoes

} non-repudiate

- closing state for reconfig
  - synchronously probe 2F+1

} proof

- liveness
  - (2F+1)-of-(3F+1)

BVP in the synch model

- **4-message-delay**
- steady state F+1
  - client-leader
  - leader-all (signed)
  - all-all (signed)
  - all-client (signed)
  - proceed with F+1 signed echoes

} non-repudiate

- closing state for reconfig
  - XFT: synchronously probe F+1 by F+1

} proof

- liveness
  - (F+1)-of-(2F+1)

vmware

BVP in the asynch model

- w/TPM
  - similar to sync
  - steady state F+1
  - proof by HW attestation

A dynamic fault model

- interplay between adversary and system

- begin handover / end handover
- begin handover **to** speculative new: |new| - F correct
- end handover **from** current: all of current can be faulty

Take-aways

- going beyond Paxos
- reconfiguration-based approach

overflow

ask me questions here; otherwise I'll go on.

vmware®

Flexible Paxos: Use-Cases

Flexible Paxos: Use-Cases